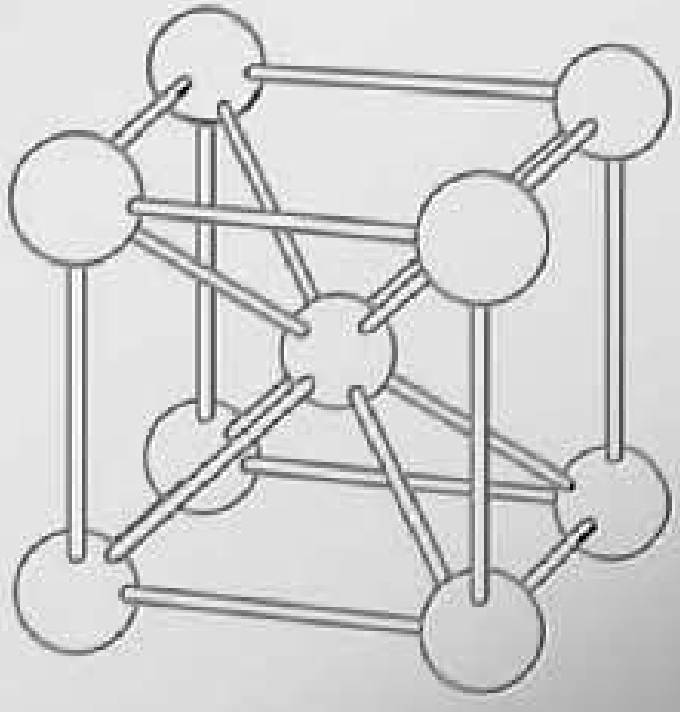
VOLUME ONE

ENCYCLOPEDIA OF CONDENSED MATTER PHYSICS

EDITED BY
FRANCO BASSANI
GERALD L. LIEDL
PETER WYDER



Editors

Franco Bassani, *Scuola Normale Superiore, Pisa, Italy*

Gerald L Liedl, *Purdue University, West Lafayette, IN, USA*

Peter Wyder, *Grenoble High Magnetic Field Laboratory, Grenoble, France*

Editorial Advisory Board

INTRODUCTION

Physics is the paradigm of all scientific knowledge. Over the centuries it has evolved to a complexity that has resulted in a separation into various subfields, always connected with one another and very difficult to single out. Freeman Dyson, in his beautiful book 'Infinite in All Directions', distinguishes two aspects of physics and two types of physicists: the unifiers and the diversifiers. The unifiers look for the most general laws of nature, like the universal attraction between masses and electric charges, the laws of motion, relativity principles, the simplest elementary particles, the unification of all forces, symmetry violation and so on. The diversifiers consider the immense variety of natural phenomena, infinite in their extension, try to explain them on the basis of known general principles, and generate new phenomena and devices that do not exist in nature. Even at the beginning of modern science Galileo Galilei, besides studying the laws of motion and laying down the principle of relativity, was interested in the phenomenon of fluorescence and disproved the theories put forward at his time. He was both a unifier and a diversifier. The full explanation of fluorescence had to await the advent of quantum mechanics, as did the explanation of other basic phenomena like electrical conductivity and spectroscopy.

The past century witnessed an explosive expansion in both aspects of physics. Relativity and quantum mechanics were discovered and the greatest of the unifiers, Albert Einstein, became convinced that all reality could be comprehended with a simple set of equations. On the other hand a wide range of complex phenomena was explained and numerous new phenomena were discovered. One of the great diversifiers, John Bardeen, explained superconductivity and invented the transistor.

In physics today we encounter complex phenomena in the behavior of both natural and artificial complex systems, in matter constituted by many particles such as interacting atoms, in crystals, in classical and quantum fluids as well as in semiconductors and nanostructured materials. Furthermore, the complexity of biological matter and biological phenomena are now major areas of study as well as climate prediction on a global scale. All of this has evolved into what we now call "condensed matter physics". This is a more comprehensive term than "solid state physics" from which, when the electronic properties of crystals began to be understood in the thirties, it originated in some way. Condensed matter physics also includes aspects of atomic physics, particularly when the atoms are manipulated, as in Bose–Einstein condensation. It is now the largest part of physics and it is where the greatest number of physicists work. Furthermore, it is enhanced through its connections with technology and industry. In condensed matter physics new phenomena, new devices, and new principles, such as the quantum Hall effect, are constantly emerging. For this reason we think that condensed matter is now the liveliest subfield of physics, and have decided to address it in the present Encyclopedia. Our focus is to provide some definitive articles for graduate students who need a guide through this impenetrable forest, researchers who want a broader view into subjects related to their own, engineers who are interested in emerging and new technologies together with biologists who require a deeper insight into this fascinating and complex field that augments theirs.

In this Encyclopedia we have selected key topics in the field of condensed matter physics, provided historical background to some of the major areas and directed the reader, through detailed references, to further reading resources. Authors were sought from those who have made major contributions and worked actively in the

area of the topic. We are aware that completeness in such an infinite domain is an unattainable dream and have decided to limit our effort to a six-volume work covering only the main aspects of the field, not all of them in comparable depth.

A significant part of the Encyclopedia is devoted to the basic methods of quantum mechanics, as applied to crystals and other condensed matter. Semiconductors in particular are extensively described because of their importance in the modern information highways. Nanostructured materials are included because the ability to produce substances which do not exist in nature offers intriguing opportunities, not least because their properties can be tailored to obtain specific devices like microcavities for light concentration, special lasers, or photonic band gap materials. For the same reasons optical properties are given special attention. We have not, however, neglected foundation aspects of the field (such as mechanical properties) that are basic for all material applications, microscopy which now allows one to see and to manipulate individual atoms, and materials processing which is necessary to produce new devices and components. Attention is also devoted to the ever-expanding role of organic materials, in particular polymers. Specific effort has been made to include biological materials, which after the discovery of DNA and its properties are now being understood in physical terms. Neuroscience is also included, in conjunction with biological phenomena and other areas of the field. Computational physics and mathematical methods are included owing to their expanding role in all of condensed matter physics and their potential in numerous areas of study including applications in the study of proteins and drug design. Many articles deal with the description of specific devices like electron and positron sources, radiation sources, optoelectronic devices, micro and nanoelectronics. Also, articles covering essential techniques such as optical and electron microscopy, a variety of spectroscopes, x-ray and electron scattering and nuclear and electron spin resonance have been included to provide a foundation for the characterization aspect of condensed matter physics.

We are aware of the wealth of topics that have been incompletely treated or left out, but we hope that by concentrating on the foundation and emerging aspects of the infinite extension of condensed matter physics these volumes will be generally useful.

We wish to acknowledge the fruitful collaboration of the members of the scientific editorial board and of the Elsevier editorial staff.

Special thanks are due to Giuseppe Grosso, Giuseppe La Rocca, Keith Bowman, Jurgen Honig, Roberto Colella, Michael McElfresh, Jaap Franse, and Louis Jansen for their generous help.

Franco Bassani, Peter Wyder, and Gerald L Liedl

Permission Acknowledgments

The following material is reproduced with kind permission of Nature Publishing Group

Figure 12 of Ionic and Mixed Conductivity in Condensed Phases

Figure 1 of Porous Silicon

Figure 5 of Quasicrystals, Electronic Structure of

Figure 2 of Superconductivity: Flux Quantization

http://www.nature.com/nature

The following material is reproduced with kind permission of the American Association for the Advancement of Science

Figure 6 of Biomolecules, Scanning Probe Microscopy of

Figure 4 of Excitons: Theory

Figures 5, 6 and 7 of Genetic Algorithms for Biological Systems

Figures 1 and 3 of Rhodopsin and the First Step in Vision

Figures 4 and 5 of Scattering, Inelastic: Electron

http://www.sciencemag.org

The following material is reproduced with kind permission of Taylor & Francis Ltd

Figure 5 of Ionic and Mixed Conductivity in Condensed Phases

Figure 2 of Semiconductor Nanostructures

http://www.tandf.co.uk/journals

The following material is reproduced with kind permission of Oxford University Press

Table 1 of Crystal Tensors: Applications

Figure 2 of Cyclotron Resonance: Metals

Figure 4 of Magnetoresistance Techniques Applied to Fermi Surfaces

http://www.oup.com

A

Acoustics: Physical Principles and Applications to Condensed Matter Physics

J D Maynard, The Pennsylvania State University, University Park, PA, USA

Introduction

For many condensed matter systems, including liquids as well as solids, acoustic measurements provide a crucial probe of important and fundamental physics of the system. In the case of solids, one of the first fundamental properties to be determined would be the atomic structure, defined by the minimum in the free energy with respect to the positions of the atoms. The next fundamental characteristic of interest might be the curvature of the free energy in the vicinity of the minimum, and this would be manifest in the elastic constants for the material. As derivatives of the free energy, elastic constants are closely connected to thermodynamic properties of the material; they can be related to specific heat, Debye temperature, the Gruneisen parameter, and they can be used to check theoretical models. Extensive quantitative connections may be made if the elastic constants are known as functions of temperature and pressure. Acoustic measurements not only probe lattice properties, they are also sensitive probes of the environment in which all interactions take place, and may be used to study electronic and magnetic properties (e.g., through magnetostriction effects). As will be discussed later, acoustic measurements involve tensor quantities, and thus can probe anisotropic properties of crystals. The damping of elastic waves provides information on anharmonicity and on coupling with electrons and other relaxation mechanisms. One of the most important features of acoustic measurements is that they provide a sensitive probe of phase transitions and critical phenomena; important examples, in addition to the obvious example of structural transitions in solids, include the superconducting transition and the superfluid transition. Indeed, one of the most impressive successes in critical phenomena has been the use of acoustics to study the lambda line of liquid helium. For the field of superconductivity, a paper on acoustic attenuation was included as one of the relatively small number of selected reprints on superconductivity published by the American Institute of Physics. Acoustic measurements are among the first performed whenever a material involving novel physics is discovered. Modern acoustic techniques, discussed below, can probe the properties of samples only a few hundred microns in size and nanoscale thin films, and may be utilized in practical applications such as micro-electro-mechanical systems (MEMS). Acoustic measurements provide significant information about condensed matter systems, and their accurate and precise measurement is certainly important.

Acoustics in Solids

For solids, acoustic phenomena reflect the elastic properties of the material. Interest in elasticity dates back to Galileo and other philosophers in the seventeenth century, who were interested in the static equilibrium of bending beams. With the basic physics introduced by Hooke in 1660, the development of the theory of elasticity followed the development of the necessary mathematics, with contributions from Euler, Lagrange, Poisson, Green, etc., and the resulting theory was summarized in the treatise by A E H Love in 1927.

Acoustic and elastic properties of solids are quantified in a set of elastic constants. These constants are like spring constants, relating forces, and displacements, and they may be measured with a static technique, in which a displacement is measured as a linear response to a small applied force. However, it was long ago learned that a better method is to measure an elastic vibration, as found, for example, in a propagating sound wave. Most existing complete sets of elastic constants for materials have been determined by measuring the time of flight of sound pulses. Recently, a relatively new method, resonant ultrasound spectroscopy (RUS), is being used. In the RUS method, rather than measuring sound velocities,

one measures the natural frequencies of elastic vibration for a number of normal modes of a sample, and processes these in a computer, along with the shape and mass of the sample. With a proper configuration, a single measurement yields enough frequencies to determine all of the elastic constants for a material (as many as 21 for a crystal with low symmetry). Samples may be prepared in a wide variety of shapes, including rectangular, spherical, etc., and it is not necessary to orient crystalline samples. A compelling reason for using RUS has to do with the nature of samples of new materials. Whenever a new material is developed, initial single crystal samples are often relatively small, perhaps on the order of a fraction of a millimeter in size. Also, with new developments in nanotechnology and the possibility of applications in the microelectronics industry, there is a great interest in systems which are very small in one or more dimensions, such as thin films and one-dimensional wires. For such small systems, pulse measurements are difficult, if not impossible, but RUS methods may be readily used.

Physical Principles for Acoustics in Solids

To begin a theory for acoustics in solids, one may imagine a spring, extended with some initial tension, and consider two points at positions x, and $x + \mathrm{d}x$. If one applies an additional local tension, or stress, σ, then the spring stretches and the two points are displaced by $\psi(x)$, and $\psi(x + \mathrm{d}x)$ respectively. The separation between the two points will have changed by $\mathrm{d}\psi$, and the fractional change in the separation, defined as the strain, is $\varepsilon \equiv \mathrm{d}\psi/\mathrm{d}x$. Hooke's law for the spring takes the form $\sigma = c\varepsilon$, where c is a one-dimensional elastic constant. For a three-dimensional elastic solid, one may use indices (i, j, etc.) which can take on the values 1, 2, and 3, referring to the x, y, and z coordinate directions, and generalize the strain to

$$\varepsilon_{ij} \equiv \frac{1}{2}\left(\frac{\partial \psi_i}{\partial x_j} + \frac{\partial \psi_j}{\partial x_i}\right) \qquad [1]$$

The symmetric form of ε_{ij} avoids pure rotations, which do not involve stress. The stress is generalized to σ_{ij}, a force per unit area acting on a surface element, where the first index refers to the coordinate direction of a component of the force, and the second index refers to the coordinate direction of the unit normal to the surface element. Hooke's law becomes

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl} \qquad [2]$$

where c_{ijkl} is the $3 \times 3 \times 3 \times 3$ (81 element) elastic tensor, and where a summation over repeated indices is implied. For a small volume element, the net force in the i-direction is $\partial\sigma_{ij}/\partial x_j$, and Newton's law may be written as

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 \psi_i}{\partial t^2} \qquad [3]$$

where ρ is the mass density.

The symmetric nature of the definitions, and the assumption that the elastic energy must be quadratic in the strains, reduces the number of independent elements of c_{ijkl} from 81 to 21. A basic symmetry has c_{ijkl} invariant if the indices are exchanged in the first pair or second pair of the four subscripts ($c_{jikl} = c_{ijkl}$, etc.); thus a reduced system of indices may be used: $11 \to 1$, $22 \to 2$, $33 \to 3$, $23 \to 4$, $13 \to 5$, $12 \to 6$, so that $c_{ijkl} \to c_{\mu\nu}$. The reduced system is used when tabulating values of elastic constants; however, the full four-index tensor must be used in calculations. Additional symmetries of a particular crystal group will reduce the number of independent elastic constants further below 21; for example, orthorhombic crystals have nine independent elastic constants, cubic crystals have three, and isotropic solids have only two.

Later, the case of an isotropic elastic solid will be useful for the purposes of illustration. In this case, one has $c_{11} = c_{22} = c_{33}$, $c_{44} = c_{55} = c_{66}$, $c_{12} = c_{13} = c_{23} = c_{11} - 2c_{44}$, and all other elements of the elastic tensor are zero. The two independent elastic constants may be taken as c_{11} and c_{44}, but other combinations, such as Young's modulus $Y = c_{44}(3c_{11} - 4c_{44})/(c_{11} - c_{44})$ and the bulk modulus $B = c_{11} - 4c_{44}/3$ are also used. The bulk modulus appears in an important thermodynamic identity involving γ, the ratio of the specific heat at constant pressure c_p to that at constant volume, c_v:

$$\gamma = \frac{c_p}{c_v} = 1 + \frac{T\beta^2 B}{\rho c_p} \qquad [4]$$

Here T is the temperature and β is the thermal expansion coefficient (TEC).

Anharmonic Effects

The basic formulation of acoustics in solids involves the expansion of energy minima about equilibrium to second order, or equivalently, assuming a harmonic potential, quadratic in strain. However, there are a number of effects which require going beyond second order. Some effects are related to exceeding "small displacements" from equilibrium, such as in quantum solids with large zero-point motion, and solids at high temperatures (near melting) where thermal motions are large. Other effects occur in equilibrium at normal or low temperatures; these include thermal

expansion, lattice thermal conductivity, and acoustic dissipation.

The relationship between anharmonic effects and thermal expansion is worth discussing. As temperature is increased, the amplitude of atomic oscillations increases, or equivalently the occupation of higher quantized energy levels increases. If the potential energy was exactly quadratic in displacements, then the center of oscillation, or the expectation value of displacement, would remain the same. With the same average positions for the atoms, the system would not expand with increasing temperature, and the TEC β would be zero. From the thermodynamic identity in eqn [4], one would also have $c_p = c_v$. On the other hand, when potentials are anharmonic, one may have (in the typical case) a stiffer repulsion at short interatomic distances, and weaker attractive forces at larger distances. The result is that at higher energy levels, the "center" position between classical turning points moves to larger distances, and the system thermally expands. This situation is illustrated in **Figure 1**.

Anharmonicity and thermal expansion can also be readily related to nonlinear acoustics in fluids. For fluid acoustics, nonlinear effects are proportional to a dimensionless second order parameter $(\rho/v)(\partial v/\partial \rho)$, where v is the sound speed, and the derivative is at constant entropy. For gases, this parameter is $(\gamma - 1)$, which by the thermodynamic identity in eqn [4], is proportional to the thermal expansion coefficient β. Thus the absence of nonlinear acoustic effects coincides with a vanishing thermal expansion.

That an anharmonic potential results in acoustic dissipation and lattice thermal conductivity may be understood by noting that with a harmonic potential, one gets a linear second-order wave equation, whose wave solutions can superimpose and pass through one another with no effect. An anharmonic potential allows sound waves (lattice vibrations) to interact and scatter from one another, permitting the transfer of energy from an ordered to a disordered form (acoustic dissipation), and allowing a change in distribution functions in passing from one location to another (lattice thermal conductivity).

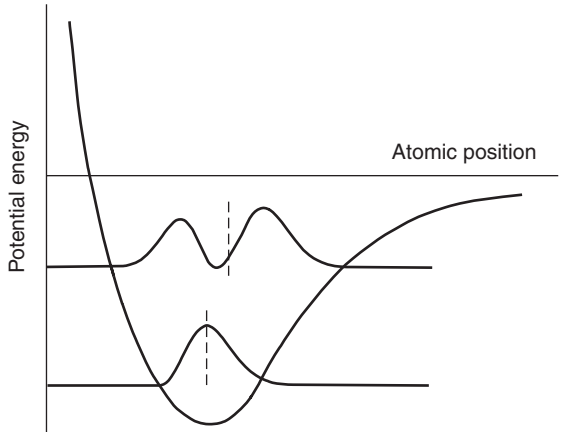


Figure 1 Illustration of the relationship between a nonquadratic potential energy curve (and nonlinear acoustics) and the phenomenon of thermal expansion.

Anharmonic effects may be probed with acoustic experiments by measuring the changes in the elastic constants as the sample is subjected to increasing uniaxial or uniform hydrostatic pressure. The coefficients which relate the changes to the pressure are referred to as "third-order elastic constants." How elastic constants themselves are determined with acoustic measurements is discussed next.

Determining Elastic Properties Experimentally

To determine the nature of sound propagation in solids, one must solve eqns [1] through [3] with some specified boundary conditions. Because of the tensor nature of the equations, the relation between particle displacement and the direction of wave propagation is quite complicated. To tackle the complexity and make a connection between ultrasound measurements and the elastic constants, two approaches may be taken. The first approach, used in conventional pulse ultrasound, is to note that if one had a sample with a large (infinite) plane surface which is perpendicular to one of the principle axes of the elastic tensor, and if a plane wave could be launched from that surface, then the tensor equations would uncouple, and a longitudinal wave or one of two transverse waves could propagate independently. In this case, for each wave, the relationship between the sound velocity and the independent elastic constants is fairly straightforward. While the determination of the principle axes and the relationships between the three sound speeds and the relevant elastic constants may be done analytically, the manipulations are complicated and must be done on a case-by-case basis; there is no elucidating general formula. The simplest case of an isotropic elastic solid will be presented here for purposes of illustration. In this case, Newton's law may be written in terms of the two independent elastic constants, c_{11} and c_{44}:

$$\rho \frac{\mathrm{d}^2 \boldsymbol{\psi}}{\mathrm{d}t^2} = c_{11} \nabla^2 \boldsymbol{\psi} - c_{44} \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{\psi}) \qquad [5]$$

$$= c_{44} \nabla^2 \boldsymbol{\psi} + (c_{11} - c_{44}) \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{\psi}) \qquad [6]$$

where the two equations are related by an identity for the $\boldsymbol{\nabla}$ operator. If one has $\boldsymbol{\nabla} \times \boldsymbol{\psi} = 0$, then the first

equation becomes a simple wave equation for a longitudinal wave with speed $v_l = \sqrt{c_{11}/\rho}$, and if one has $\nabla \cdot \boldsymbol{\psi} = 0$, then the second equation becomes a simple wave equation for transverse waves with speed $v_t = \sqrt{c_{44}/\rho}$.

Although the pulse ultrasound method has been used extensively in the past, it has a number of disadvantages, including problems with transducer ringing, beam diffraction, and side-wall scattering, and the inconvenience that the sample must be recut, repolished, and reattached to a transducer if one wants more than the three elastic constants accessible with one measurement. The second approach to determining elastic constants avoids all of the disadvantages.

The second approach is the one described earlier as RUS, which involves the use of a computer to numerically solve the elastic constants given a set of measured natural frequencies for a solid with a given shape and boundary conditions (usually stress-free conditions). The computer processing involves solving a "forward problem" (finding the natural frequencies in terms of the elastic constants) first and then inverting. Unlike the conventional pulse ultrasound approach, the forward problem does not provide a simple relationship between the modes of vibration and the elastic constants; the displacements in the various modes involve all of the elastic constants in a complicated manner, and a numerical computation is required to sort it all out.

The forward problem may be posed as the minimization of a Lagrangian L given by

$$L = \iiint (\rho\omega^2 \psi_i \psi_i - c_{ijkl}\varepsilon_{ij}\varepsilon_{kl})\, \mathrm{d}V \qquad [7]$$

The minimization is accomplished numerically with a Rayleigh–Ritz method, and the results yield discrete resonance frequencies, $f_n = 2\pi\omega_n$, given the elastic constants, c_{ijkl}. For the RUS technique, what is needed is the inverse. In most cases, there will be more measured frequencies than independent elastic constants; so what is required is to find a set of independent elastic constants which best fits the measured frequencies, usually in a least squares sense. Furthermore, when there are more measured frequencies than independent elastic constants, then other parameters may be varied in order to best fit the measured frequencies. Such parameters may include the shape and dimensions of the sample (although one known length is necessary), and the orientation of the crystallographic axes relative to the faces of the sample. In any case, it is not necessary that crystallographic axes be oriented with respect to faces of a sample, although computations are greatly simplified if they are oriented.

Experimental Methods for Acoustic Measurements in Solids

Acoustic measurements with the pulse method are fairly straightforward; emphasis is on careful bonding of transducers to samples and the use of suitable high-frequency pulse electronics. The RUS method is less well known, and can be briefly described as follows.

In a general RUS measurement, the natural frequencies of a sample with stress-free boundary conditions are determined by measuring the resonance frequencies of the sample when held (lightly, with no bonding agents, at two positions on the sample surface) between two transducers. One transducer acts as a drive to excite vibrations in the sample at a tunable frequency, and the second measures the amplitude (and possibly the phase) of the response of the sample; as the frequency of the drive is swept, a sequence of resonance peaks may be recorded. The positions of the peaks will determine the natural frequencies f_n (and hence the elastic constants), and the quality factors (Q's, given by f_n divided by the full width of a peak at its half-power points) will provide information about the dissipation of elastic energy.

RUS may also be used to measure the properties of thin films on a substrate, to determine the effects of induced strain from lattice mismatch, etc. In this case, the natural frequencies of the substrate alone are measured, then the same sample is again measured with the film in place. From the shifts in the natural frequencies, the properties of the film may be determined.

A simple apparatus for making RUS measurements is illustrated in **Figure 2**. In the illustration, a rectangular parallelepiped sample is supported by transducers at diametrically opposite corners. Corners are used for contact because they provide elastically weak coupling to the transducers, greatly reducing loading, and because the corners are always elastically active (i.e., they are never nodes), and thus can be used to couple to all of the normal modes of vibration.

Acoustics in Fluids

The thermo-hydrodynamic state of a fluid may be specified with five fields, which may be taken as the mass density $\rho(\boldsymbol{r}, t)$, the pressure $p(\boldsymbol{r}, t)$, and the mean flow velocity $\boldsymbol{u}(\boldsymbol{r}, t)$. The five equations needed to determine the five fields are conservation of mass, Newton's law for the motion of the center of mass of a fluid element (three components), and conservation of energy for motion about the center of mass. These

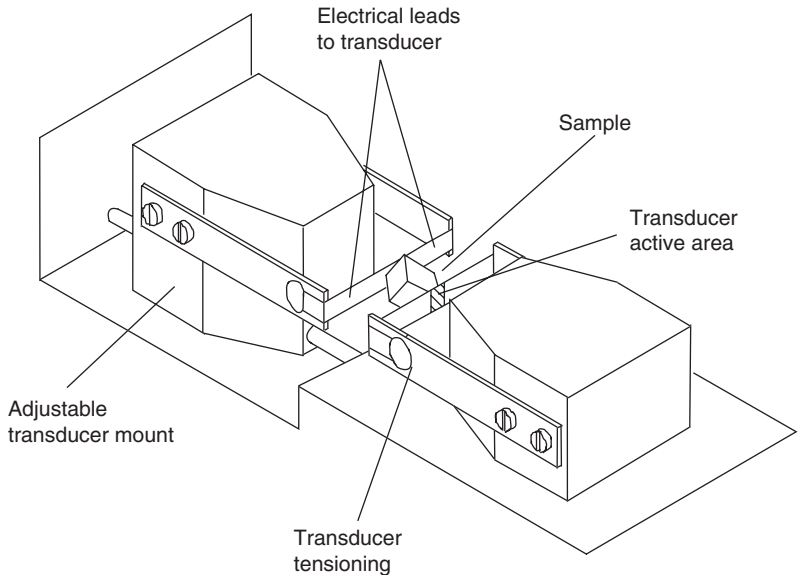


Figure 2 An apparatus for measuring elastic constants and acoustic attenuation with RUS.

equations are respectively:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0 \qquad [8]$$

$$\rho \left[\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \mathrm{u} \right] = -\boldsymbol{\nabla} p + \eta \nabla^2 \boldsymbol{u} + \left(\frac{1}{3}\eta + \xi \right) \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{u}) \qquad [9]$$

$$T \left[\frac{\partial S}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} S \right] = \frac{2\eta}{\rho} \left[e_{ij} - \frac{1}{3} (\boldsymbol{\nabla} \cdot \boldsymbol{u}) \right]^2 + \frac{\xi}{\rho} (\boldsymbol{\nabla} \cdot \boldsymbol{u})^2 + \frac{1}{\rho} \boldsymbol{\nabla} \cdot (\kappa \boldsymbol{\nabla} T) \qquad [10]$$

where S is entropy, $e_{ij} = (\partial u_i / \partial x_j + \partial u_j / \partial x_i)/2$, η and ξ are the shear and bulk viscosity, and κ is the thermal conductivity. The transport terms (involving η, ξ, and κ) give rise to dispersion and dissipation. If these terms are dropped, then the last equation simply gives $S =$ constant, and the first two (for small displacements from equilibrium) are easily combined to give a wave equation with a sound speed given by $(\partial p / \partial \rho)_S$.

Superfluids are modeled as having a superfluid component (in a macroscopic quantum ground state) and a normal-fluid component (a gas of excitations above the ground state). With two fluid components, nine fields are required to specify the state of the system; in addition to the ones for a classical fluid, one also has the mass density and the mean flow velocity for the superfluid component alone. Now the linearized equations admit two sound speeds for the unconstrained fluid, and two more sound speeds when the normal fluid component is held fixed through its viscosity, and the inviscid superfluid component is still free to flow. The four sound modes are (1) a pressure wave with the two components moving together (referred to as "first sound"), (2) a temperature wave with the two components moving in counterflow ("second sound"), (3) a pressure wave with the normal fluid clamped ("fourth sound"), and (4) a temperature wave with the normal fluid clamped ("fifth sound"). An illustration of these modes is presented in **Figure 3**. "Third sound" is a surface wave which propagates on a thin film of superfluid; it is not a fundamental sound mode because its restoring force is not intrinsic to the superfluid, but rather is determined by the substrate on which the superfluid film is formed. With both mechanical and thermal properties represented in the sound modes, the sound speeds, measured as functions of temperature and pressure, may be used to determine all of the thermodynamics of the superfluid.

Applications of Acoustics in Condensed Matter Physics

In areas of condensed matter physics which involve the development of exotic materials, it is often of great value to use acoustics simply to measure and tabulate sound speeds or elastic constants. This is the case for such materials as alloys, composites, porous materials, sintered materials, polymers (plastics, epoxies, elastomers, etc.), cements, piezoelectrics, viscoelastic and non-Newtonian fluids, and fluid mixtures. In some cases, it is of significant benefit to use acoustics to monitor systems which evolve in

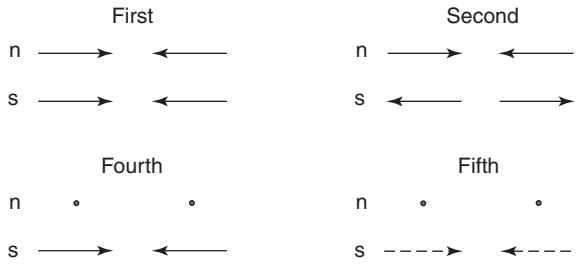


Figure 3 An illustration of the basic sound modes in superfluid helium. The "n" and "s" lines indicate the motion of the normal fluid and superfluid components, respectively. The dots indicate that the normal fluid has been clamped by its viscosity. The dashed lines indicate motion under "pressure release" conditions.

time, as during the curing of epoxy, the hardening of cement, the progress of chemical reactions, etc. Other special condensed matter systems which benefit from basic acoustic measurements are quantum solids, superfluids, quasicrystals, granular media, foams, rocks, etc. High amplitude acoustic fields are used as driving mechanisms in sonochemistry, cavitation studies (including cavitation in quantum fluids), acoustic levitation (facilitating noncontact sample manipulation), and sonoluminescence, in which a collapsing bubble creates such extreme conditions that light may be emitted. Acoustic emission, where signals are generated by the system itself, is important in studies of cracks, fracture, earthquakes, etc.

A particular application of acoustic measurements in solid state physics is the determination of the contribution of lattice acoustic modes to the specific heat, c_{lat}. This is given by

$$c_{\text{lat}} = 9\left(\frac{n}{\rho}\right)k_{\text{B}}\left(\frac{T}{\Theta_{\text{D}}}\right)^3\int_0^{\Theta_{\text{D}}/T}\frac{x^4 e^x dx}{(e^x-1)^2} \qquad [11]$$

where n is the number of lattice sites per unit volume, k_{B} is Boltzmann's constant, and Θ_{D} is the Debye temperature. It is the Debye temperature which may be expressed in terms of acoustic parameters, as follows:

$$\Theta_{\text{D}} = \frac{\hbar}{k_{\text{B}}}\left[\frac{1}{18\pi^2 n}\sum_s\int\frac{d\Omega}{4\pi}\frac{1}{v_s(\hat{k})^3}\right]^{-1/3} \qquad [12]$$

Here the sum in s is over the three different acoustic modes in a particular direction $\hat{k}$, which have sound speeds $v_s(\hat{k})$. For isotropic solids, the average of $1/v_s^3$ is $(1/v_{\text{l}}^3 + 2/v_{\text{t}}^3)$.

As mentioned in the introduction, an important application of acoustics is in the study of phase transitions. Acoustic studies are typically used in conjunction with a model of the transition, for example, a Landau expansion in the case of second-order transitions. In this case the free energy F, including elastic strain energy, may be expressed as a fourth-order equation in some order parameter Ψ:

$$F = \tfrac{1}{2}c_0\varepsilon^2 + \tfrac{1}{2}a_1(T-T_{\text{c}})\Psi^2 + \tfrac{1}{4}a_2\Psi^4 + \tfrac{1}{2}a_3\varepsilon\Psi^2 \qquad [13]$$

where c_0 and ε are some nominal elastic constant and strain (ignoring the tensor nature), T_{c} is the critical temperature, and a_1, a_2, and a_3 are constants. The effective elastic constant is $(\partial^2 F/\partial\varepsilon^2)$. Minimizing the free energy with respect to the order parameter gives the result that the effective elastic constant is c_0 when $T > T_{\text{c}}$, and is $(c_0 - a_3^2/2a_2)$ when $T < T_{\text{c}}$. Thus an experimentally measured jump in an elastic constant at a second-order phase transition gives access to the parameters in a Landau expansion.

An extensive application of RUS has been in geophysics, where the measurement of the thermodynamic properties and anharmonic effects of materials at high temperatures (exceeding twice the Debye temperature) are a high priority. Elastic data can check theoretical models and their extension to high temperature and pressure, where some asymptotic behavior may be convenient for other geophysical calculations and for extrapolations to even higher temperatures. Anharmonic effects are evident in the Gruneisen relation and in the departure of heat capacity from the Law of Dulong and Petit, $c_{\text{lat}} = 3(n/\rho)k_{\text{B}}$.

Acoustic Dissipation

Acoustic dissipation may be discussed with two basic pictures. In one picture, energy in an ordered form in the acoustic field is lost to some disordered form, so that the amplitude of a sound wave decreases with distance x as $e^{-\alpha x}$, where α is the attenuation coefficient. From the functional form $e^{-\alpha x}$, the attenuation coefficient α may be considered as an imaginary part of a complex wave vector, $\tilde{k} = k + i\alpha$. Acoustic dissipation in this picture is typically modeled with a coupling term in a Hamiltonian between phonons and some other system into which energy is lost (electrons, magnetic spins, etc.).

In the second picture, one notes that changes in stress (or strain) are not immediately followed by changes in strain (or stress), and there is a time lag, with a characteristic relaxation time, between energy being stored in kinetic and potential forms. If there were no time lag, then acoustic variations would oscillate back and forth along a single path (typically an isentrope) in the stress–strain plane. If there is a time lag, then the path opens up into a loop, and acoustic energy would be dissipated as "lost work" equal to the nonzero area of the loop. The time lag may be represented with a time dependence $f(t) = (1 - e^{-t/\tau})$, where τ is the characteristic relaxation time from a

sudden change. The time lag is incorporated in the equations of motion with additional time derivatives; however, the equations may be analyzed by Fourier transforming in time, and the result involves the Fourier transform of $f(t)$, proportional to $\hat{f}(\omega)$ given by

$$\hat{f}(\omega) = \frac{1}{1+(\omega\tau)^2} + \mathrm{i}\,\frac{\omega\tau}{1+(\omega\tau)^2} \qquad [14]$$

The real part gives rise to dispersion in sound propagation, and the imaginary part results in attenuation; the effective attenuation coefficient is $\alpha = \alpha_{\max} 2\omega\tau/(1+\omega^2\tau^2)$, where the attenuation peak $\alpha_{\max}$ occurs when $\omega\tau = 1$.

For fluids, the "time lag" derivatives are already in place in eqns [8]–[10]. If the terms containing the transport coefficients η, ξ, and κ are not dropped, and the equations are solved to first order in these coefficients, then solutions will involve a complex wave vector $\tilde{k} = k + \mathrm{i}\alpha$, with the attenuation coefficient α given by

$$\alpha = \frac{\omega^2}{2\rho v^3}\left[\left(\frac{4}{3}\eta + \xi\right) + \frac{\kappa}{\gamma c_v}(\gamma - 1)\right] \qquad [15]$$

For solids, forms of acoustic dissipation which are most readily analyzed with the first picture include electron–phonon scattering and phonon–phonon scattering. The picture for electron–phonon scattering is particularly important for studying Bardeen–Cooper–Schrieffer (BCS) superconductors; as more electrons pair and enter the superconducting state, the phonon interaction gets used up in the pairing interaction, and part of the electron–phonon scattering no longer contributes to acoustic attenuation. Thus, a measured drop in acoustic attenuation tracks the number of paired electrons.

Forms of acoustic dissipation which are most readily analyzed with the second picture include time lags for energy transfer to magnetic spin systems, electric dipole systems, defect and impurity motion (mechanical diffusion or viscosity), other forms of energy storage (thermal diffusion), etc. Models based on the second picture are particularly useful in analyzing acoustic attenuation measurements in glasses, "two-level" systems, impurity doped systems, etc.

The list of mechanisms related to acoustic attenuation just presented is by no means exhaustive. Indeed, the broad range of interacting mechanisms make acoustic attenuation one of the most widespread applications of acoustics in condensed matter physics.

See also: Crystal Symmetry; Crystal Tensors: Applications; Lattice Dynamics: Anharmonic Effects; Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Structural Instability and Soft Modes; Lattice Dynamics: Vibrational Modes; Mechanical Properties: Anelasticity; Mechanical Properties: Elastic Behavior; Specific Heat; Thin Films, Mechanical Behavior of.

PACS: 43.20.−f; 43.25.−x; 43.35.−c; 43.58.+z; 62.65.+k; 62.20.Dc; 62.80.+f; 74.25.Ld

Further Reading

Ahlers G (1980) Critical phenomena at low temperature. *Reviews of Modern Physics* 52: 489–503.

American Association of Physics Teachers. *Superconductivity, Selected Reprints*, published by the American Physical Society.

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Toronto: Thomas Learning.

Banys J, Macutkevic J, Samulionis A, and Vysochanskii Y (2004) Dielectric and ultrasonic investigation of phase transitions in CuInP2S6 crystals. *Phase Transitions* 77: 345–358.

Bhatia AB (1967) *Ultrasonic Absorption*. New York: Dover.

Goto T and Anderson OL (1988) Apparatus for measuring elastic constants of single crystals by a resonance technique up to 1825 K. *Review of Scientific Instruments* 59: 1405–1409.

Ichitsubo T, Ogi H, Nishimura S, Seto T, and Inui H (2004) Elastic stiffness and ultrasonic attenuation of superconductor MgB_2 at low temperatures. *Physical Review B* 66: 052514-1–052514-4.

Levy M (2000) *Handbook of Elastic Properties of Solids, Liquids, and Gases*. New York: Academic Press.

Love AEH (1927) *Treatise on the Mathematical Theory of Elasticity*. Cambridge: Cambridge University Press.

Maynard JD (1996) Resonant ultrasound spectroscopy. *Physics Today* 49: 26–31.

Migliori A and Sarrao JL (1997) *Resonant Ultrasound Spectroscopy*. New York: Wiley.

Pandry JD, Dey R, and Bhatt BD (2003) Estimation of molecular radius of liquids and liquid mixtures from sound velocity. *Journal of Molecular Liquids* 111: 67–71.

Putterman SJ (1974) *Superfluid Hydrodynamics*. Amsterdam: North-Holland.

Schrieffer JR (1964) *Theory of Superconductivity*. New York: W.A. Benjamin.

Truell R, Elbaum C, and Chick BB (1969) *Ultrasonic Methods in Solid State Physics*. New York: Academic Press.

Nomenclature

B	bulk modulus (Pa)
c	elastic constant (Pa)
c_{lat}	lattice specific heat ($\mathrm{J\,kg^{-1}\,K^{-1}}$)
c_p	specific heat at constant pressure ($\mathrm{J\,kg^{-1}\,K^{-1}}$)
c_v	specific heat at constant volume ($\mathrm{J\,kg^{-1}\,K^{-1}}$)
F	free energy per unit mass ($\mathrm{J\,kg^{-1}}$)
k	wave vector ($\mathrm{m^{-1}}$)
k_{B}	Boltzmann's constant ($\mathrm{J\,K^{-1}}$)
L	Lagrangian (J)
n	number of lattice sites per unit volume ($\mathrm{m^{-3}}$)

S	entropy per unit mass ($J\,kg^{-1}\,K^{-1}$)	ε	strain (dimensionless)
t	time (s)	η	shear viscosity ($N\,s\,m^{-2}$)
T	temperature (K)	Θ_D	Debye temperature (K)
$\boldsymbol{u}$	fluid element velocity ($m\,s^{-1}$)	ξ	bulk viscosity ($N\,s\,m^{-2}$)
v	sound speed ($m\,s^{-1}$)	κ	thermal conductivity ($W\,m^{-1}\,K^{-1}$)
V	volume (m^3)	ρ	mass density ($kg\,m^{-3}$)
x	spatial coordinate (m)	σ	stress (Pa)
α	attenuation coefficient (m^{-1})	τ	relaxation time (s)
β	thermal expansion coefficient (K^{-1})	ψ	displacement (m)
γ	specific heat ratio (dimensionless)	ω	angular frequency (s^{-1})

Allotropy and Polymorphism

D R Gaskell, Purdue University, West Lafayette, IN, USA

The thermodynamic origins of allotropy and polymorphism are presented. Allotropy and polymorphism are, respectively, the ability of elements and of chemical compounds to have stable existences in more than one crystal form. The allotropies of iron, carbon, and sulfur, and the polymorphisms of silica and zirconia are discussed.

Introduction

A chemical element exhibits allotropy when it can have a stable existence in more than one crystal form. Polymorphism is the same phenomenon exhibited by a chemical compound. When a range of possible states of existence is available to an element or compound, the stable state is that which has the lowest molar Gibbs free energy at the constant values of pressure and temperature of interest. The molar Gibbs free energy, G, in turn, is determined by the molar enthalpy, H, the molar entropy, S, and the temperature (in kelvins), T, as

$$G = H - TS$$

Low values of G are obtained with low values of H and high values of S. As only changes in enthalpy can be measured by the transfer of thermal energy between a thermodynamic system and a thermostating reservoir, the enthalpy H does not have a definite value and thus, also, G does not. In contrast, S, which is a measure of the thermal and configurational disorder in a thermodynamic system, does have a definite value. Consequently, changes in the molar Gibbs free energy which accompany phase transformations or chemical reactions can be obtained from the corresponding changes in H and S as

$$\Delta G = \Delta H - T\Delta S$$

The molar enthalpy of an element, relative to that of the state in which the atoms are at infinite distances from one another, is the thermal energy transferred to a thermostat when the atoms come together and occupy the sites in a regularly arrayed crystal lattice. The relative enthalpy is then a measure of the bonding energy in the crystal.

The Allotropy of Iron

Iron has two allotropes: a face-centered cubic (f.c.c.) crystal form and a body-centered cubic (b.c.c.) crystal form. In the former, the unit cell has atoms located at each of the eight corners, each one of which contributes one-eighth of an atom to the unit cell. Atoms located at the centers of each of the six faces of the cell, each contribute one-half of an atom to the unit cell, to give a total of four atoms per unit cell. Alternatively, the f.c.c. crystal structure can be considered to consist of planes of close-packed atoms stacked in the sequence ABCABC. The b.c.c. unit cell has atoms located at each of the eight corners of the unit cell and one atom located at the center of the cell, giving two atoms per unit cell.

The variations, with temperature at a constant pressure of 1 atm, of the molar Gibbs free energies of the f.c.c. and b.c.c. allotropes of iron are shown schematically in **Figure 1a**. In **Figure 1a**

$$\left(\frac{\partial G}{\partial T}\right)_P = -S$$

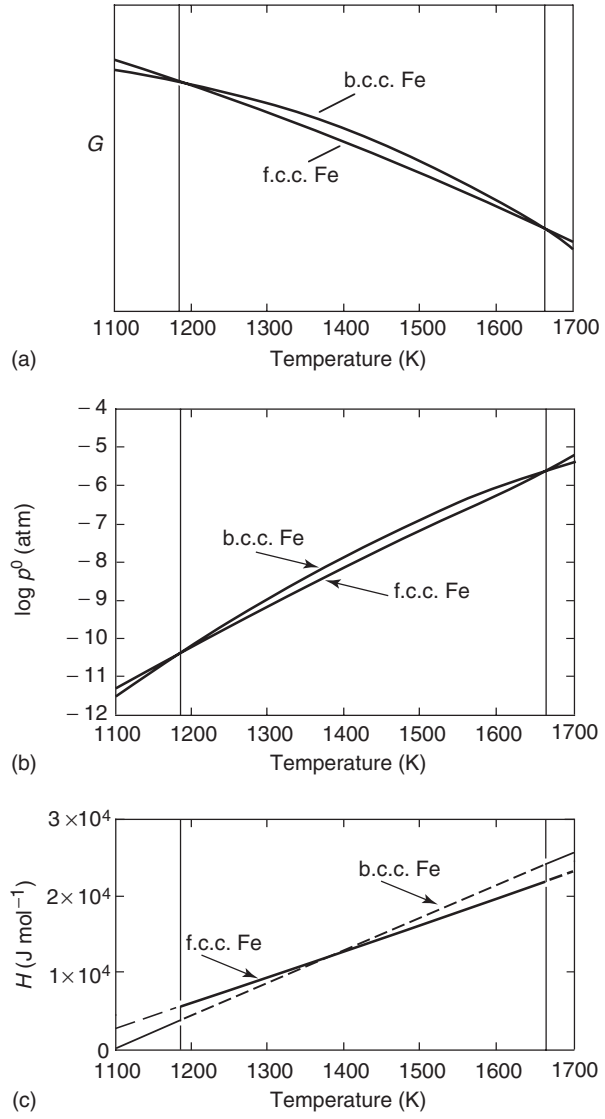


Figure 1 (a) The variations, with temperature, of the molar Gibbs free energies of b.c.c. Fe and f.c.c. Fe (schematic). (b) The variations, with temperature, of the saturated vapor pressures of b.c.c. Fe and f.c.c. Fe. (c) The variations, with temperature, of the relative molar enthalpies of b.c.c. Fe and f.c.c. Fe.

and

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_P = -\frac{c_p}{T}$$

where c_p is the constant pressure molar heat capacity of the phase. This quantity is formally defined as

$$c_p = \left(\frac{\partial H}{\partial T}\right)_P \qquad [1]$$

that is, it is the ratio of the heat absorbed by a mole of substance, at constant pressure, to the consequent increase in temperature.

The shapes of the curves in **Figure 1a** are determined by the constant pressure molar heat capacities and the molar entropy of the two crystal forms. In the temperature range 1100–1700 K, the value of c_p for b.c.c. Fe is 7.9 J K^{-1}, larger than that for f.c.c. Fe and the highly ordered f.c.c. crystal structure has a lower molar entropy than that of the b.c.c. structure. Thus, in **Figure 1**, since the rate of decrease of G with increasing temperature for the b.c.c. Fe is greater than that for the f.c.c. Fe and the second derivative of the line for b.c.c. Fe is greater than that of the f.c.c. line, the lines intersect twice at 1187 and 1664 K. Thus, the b.c.c. form is stable at temperatures lower than 1187 K and at temperatures from 1664 K to the melting temperature of 1809 K. The f.c.c. form is stable in the range of temperature 1187–1664 K. The low-temperature b.c.c. form is referred to as α-Fe, the f.c.c. form is referred to as γ-Fe and the high-temperature b.c.c. form is referred to as δ-Fe. The original assignment of β-Fe to the low-temperature b.c.c. form, between the Curie temperature of 1033 and 1187 K, is no longer used.

Figure 1b shows the variations, with temperature, of the saturated vapor pressures exerted by b.c.c. Fe and f.c.c. Fe in the range of temperature 1100–1700 K. The saturated vapor pressures are related to the difference between the Gibbs free energies of the two crystal forms by

$$\Delta G(\alpha \rightarrow \gamma) = RT \ln \frac{p_\gamma^0}{p_\alpha^0}$$

A negative value of $\Delta G(\alpha \rightarrow \gamma)$ makes the p_γ^0 value lower than the p_α^0 value.

As has been stated, the constant pressure molar heat capacity of b.c.c. Fe is 7.9 J K^{-1} larger than that of f.c.c. Fe. Consequently, from eqn [1], the rate of increase of the molar enthalpy of b.c.c. Fe is greater than that of f.c.c. Fe. The variations, with temperature, of the relative molar enthalpies of the two crystal forms are shown in **Figure 1c**, in which the reference state is chosen as b.c.c. Fe at 1100 K. Increasing the temperature from 1100 to 1187 K causes the molar enthalpy to increase along the b.c.c. Fe line. At 1187 K, the b.c.c. form transforms to the f.c.c. form with the required increase in enthalpy (the latent heat of the transformation). Further heating causes the enthalpy to increase along the f.c.c. Fe line (which intersects with the metastable b.c.c. Fe line at 1397 K) and, at 1664 K, the f.c.c. Fe transforms back to the b.c.c. Fe, again with the required increase in enthalpy.

The combination of the allotropy of iron and Henry Bessemer's patent no. 356, dated 12 February 1856 "On the Manufacture of Malleable Iron and

Steel without Fuel," gave rise to the Industrial Revolution. Carbon, which is the alloying element in plain carbon steel, occupies interstitial sites in the f.c.c. Fe and b.c.c. Fe lattices. The solution of carbon in f.c.c. Fe produces a phase called austenite, and its solution in b.c.c. Fe produces a low-temperature phase called α-ferrite and a high-temperature phase called δ-ferrite. As the interstitial sites in f.c.c. Fe are larger than those in b.c.c. Fe, the solubility of carbon in austenite is larger than in α-ferrite and δ-ferrite. Thus, carbon is an "austenite stabilizer." **Figure 2**, which is the phase diagram for the system Fe-Fe_3C, shows the extent of the austenite phase field which terminates at the eutectoid point (0.78 wt.% C, 723°C). When the temperature of austenite of this composition is decreased to a value lower than 723°C, the austenite undergoes a eutectoid decomposition to produce a structure consisting of alternating layers of ferrite and cementite (the metastable iron carbide, Fe_3C). This structure is called pearlite and the fineness of the structure is determined by the rate of cooling of the eutectoid austenite through the eutectoid temperature. The coarseness of the microstructure increases with decreasing cooling rate. With increasing rate of cooling a limit is reached, beyond which nucleation of the ferrite and the cementite from the austenite is inhibited. The austenite transforms to a metastable body-centered tetragonal phase by means of a diffusionless shear mechanism. This very hard and brittle structure is called martensite and, in it, the shear stresses which cause the brittleness are relieved by tempering at some temperature less than 723°C. The occurrence of allotropy in iron allows the mechanical properties of plain carbon steels to be manipulated by a choice of carbon content and heat treatment. Thus, steels can be produced, which, at one extreme, are soft enough to be plastically deformed to the shape of a paper clip, or, at the other extreme, hard and tough enough to be used as a bearing material.

Nickel, which has the f.c.c. crystal structure, forms substitutional solid solutions with iron and hence stabilizes the f.c.c. structure. Under equilibrium conditions, a γ-Fe,Ni containing 53 wt.% Ni undergoes a eutectoid decomposition to α-Fe and $FeNi_3$ at 345°C. However, as nucleation and growth of α-Fe from γ-Fe,Ni requires significant diffusion by migration of vacant lattice sites, the presence of a few percent Ni in solid solution in the γ-Fe,Ni produces a metastable γ-Fe,Ni phase at room temperature. In contrast, chromium, which has the b.c.c. structure, also forms substitutional solid solutions with iron and thus stabilizes the b.c.c. structure. This causes the formation of a "γ-loop" shown in **Figure 3**. At the minimum temperature of the γ-loop (7 wt.% Cr, 831°C), the γ-Fe and α-Fe,Cr phases have the same composition. In Fe–Cr solid solutions containing less than 12 wt.% Cr, the product of oxidation is the highly defective spinel $FeO \cdot Cr_2O_3$. With Cr contents greater than 12 wt.%, the oxide Cr_2O_3 is the product of oxidation. Thus, a "stainless" steel can be produced by having sufficient Ni in solid solution in Fe to give a single-phased f.c.c. structure at room temperature and by having sufficient Cr in solid solution to produce a protective

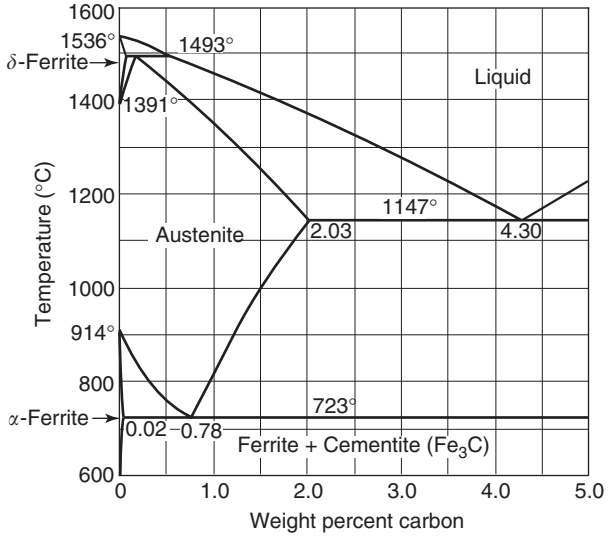


Figure 2 The phase diagram for the system Fe–Fe_3C.

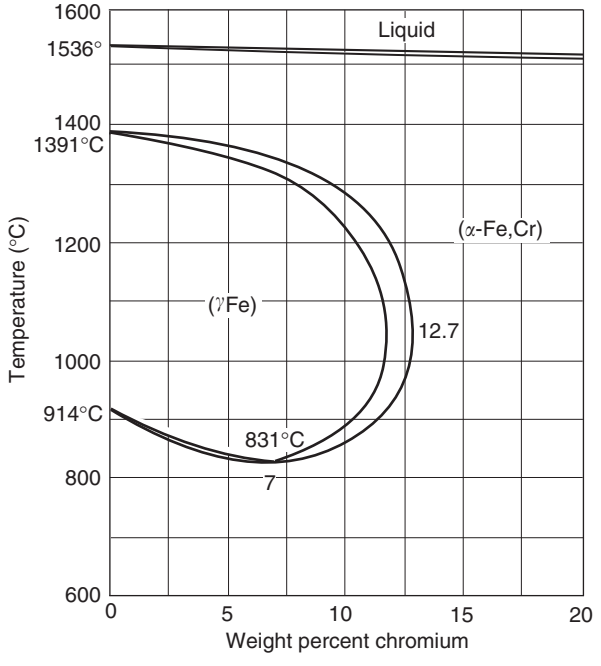


Figure 3 The iron-rich end of the Fe–Cr phase diagram.

coating of Cr_2O_3 on the surface of the alloy. Such an alloy is type 304 stainless steel (or 18-8 stainless steel) which is Fe containing 18 wt.% Cr and 8 wt.% Ni.

The Allotropy of Carbon

Carbon (element number 6), located at the top of group four in the periodic chart of the elements, has the electron configuration $2s^1 2s^2 2p^2$, which would indicate that it has a valence of 2. However, hybridization of the electron orbits of the carbon atom gives rise to the occurrence of the familiar allotropes, diamond, and graphite. In sp^3 hybridization, one of the 2*s* electrons is promoted to the 2*p* orbital and the four electrons in the second shell undergo hybridization to produce four energetically equivalent electrons. The 272 kJ mol^{-1} required to promote the electron from the 2*s* to the 2*p* shell is more than compensated for by the decrease in energy which occurs upon hybridization. The mutual repulsion between the hybrid orbitals causes them to point toward the corners of a tetrahedron, as shown in **Figure 4**. The unit cell for the diamond cubic crystal structure, which is shown in **Figure 5**, can be regarded as being based on the f.c.c. unit cell containing eight subcubes. Carbon atoms occupy the upper-back-left, upper-front-right subcubes, the lower-front-left and lower-back-right subcubes with the bond length being 15.4 nm.The rigidity of the bonds between the atoms makes diamond the hardest material in existence (number 10 on the Moh scale of hardness). It has a high melting temperature, 3800°C, and a large latent heat of melting, 105 kJ mol^{-1} at a pressure of 48 kbar. Also, it has a high index of refraction, 2.417, making it a desirable gem stone. In 1913, the determination of the structure of diamond was one of the early successes of X-ray analysis.

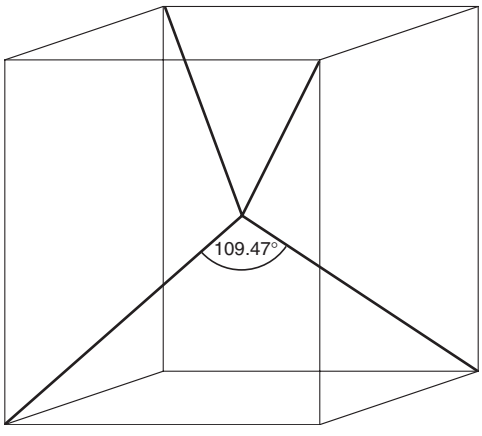


Figure 4 Lines directed to the corners of a tetrahedron.

The crystal structure of graphite is determined by sp^2 hybridization in which a 2*s* electron is promoted to the $2p_z$ orbital and the remaining 2*s* and the $2p_x$ and $2p_y$ orbitals form a trigonal hybrid containing three energetically equivalent orbitals which lie in a plane, forming angles of 120° with their neighboring orbitals. The crystal structure of graphite, which is shown in **Figure 6**, consists of sheets of σ-bonded atoms in the *xy*-plane, arranged in hexagons which are bonded to one another in the *z*-direction by π-bonds formed by overlap of the p_z orbitals. The lengths of the σ-bonds and the π-bonds are, respectively, 14.2 and 34 nm. As the π-bonds are much weaker than the σ-bonds, this property facilitates easy shearing of the sheets of hexagons and makes graphite a good lubricating material.

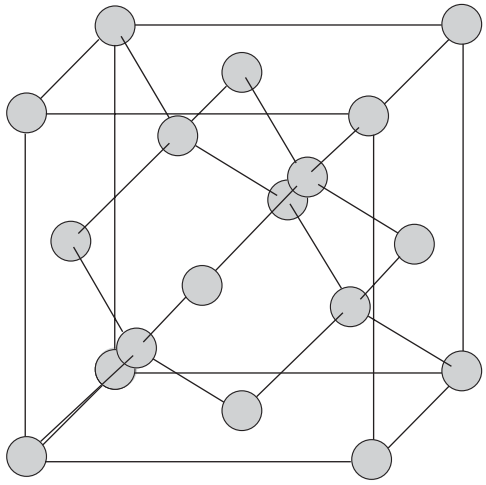


Figure 5 The unit cell for the diamond cubic crystal structure.

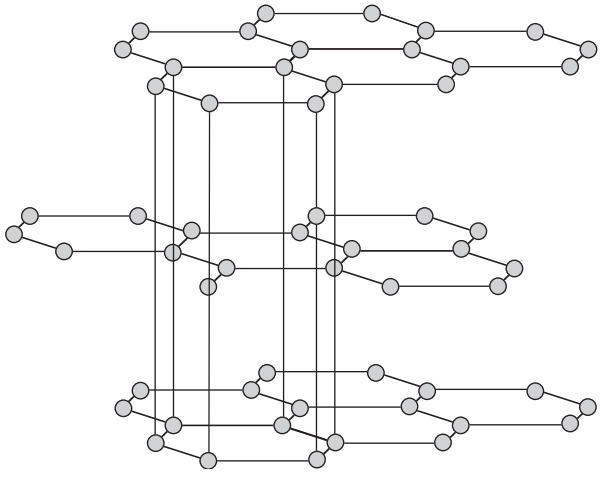


Figure 6 The crystal structure of graphite. The lines join atoms in successive sheets which are aligned vertically.

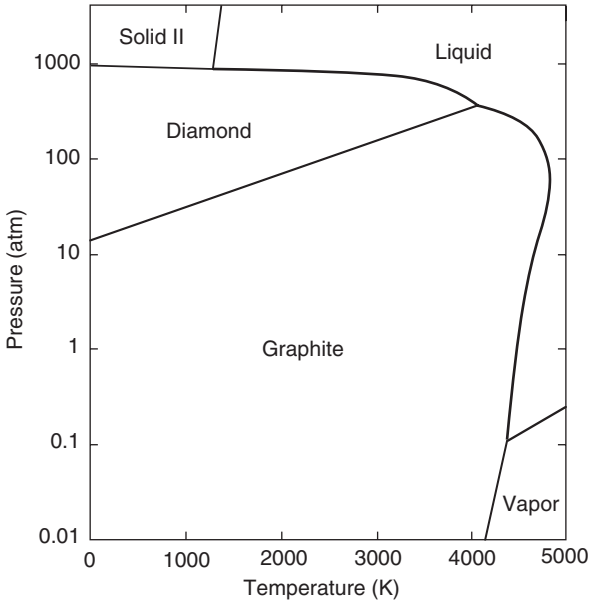


Figure 7 The phase diagram for carbon.

The phase diagram for carbon, presented as **Figure 7**, shows that a third allotrope, solid II, exists at pressures in excess of 1000 atm, and that diamond is metastable at low pressures. Diamonds are formed under conditions of high pressure experienced in clay pipes in the earth's crust. It exists in the metastable state at room temperature and pressure because of the extensive rearrangement of atoms required for the transformation from the diamond structure to the graphite structure. At lower pressures, the melting temperature of graphite increases with increasing pressure, and at higher pressure it decreases with increasing temperature. Thus, at lower pressures, the molar volume of graphite is lower than that of liquid carbon and, at higher pressures, the reverse is the case. At the "nose" of the melting curve, the molar volumes of graphite and liquid have the same value. The densities of diamond and graphite at 298 K and 1 atm pressure are, respectively, 3.515 and 2.2 g cm^{-3}.

The Allotropy of Sulfur

The phase diagram for sulfur, presented as **Figure 8**, shows that sulfur can exist in a rhombic crystal structure and in a monoclinic crystal structure. The unit cell of the rhombic form contains 128 atoms, existing as 16 puckered rings of eight atoms which form S_8 molecules. Sulfur melts to form a translucent liquid, which, when heated above 200°C, transforms to a red, highly viscous form produced by the breaking of the rings and the entanglement of the chains

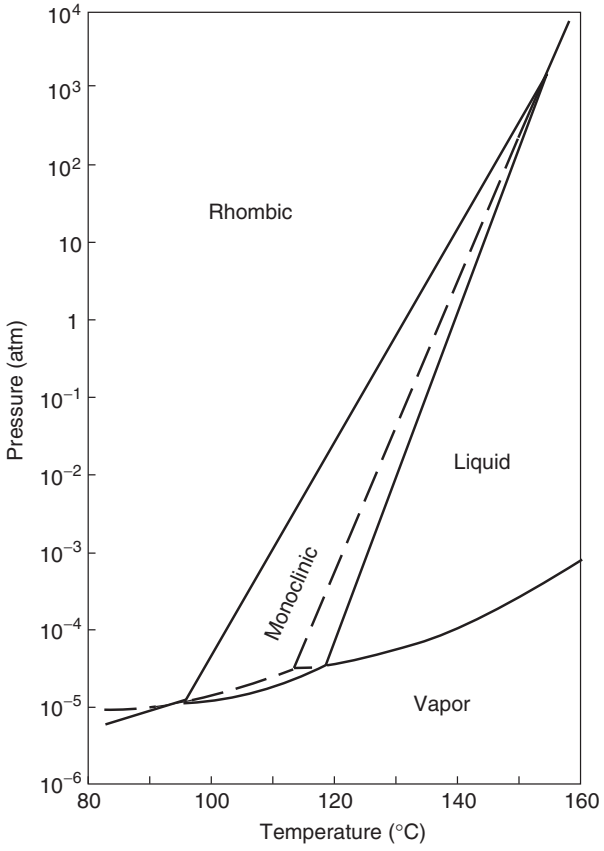


Figure 8 The phase diagram for sulfur.

produced. Amorphous or "plastic" sulfur can be obtained by fast cooling of the crystalline form and this amorphous phase also has a helical structure with eight atoms per spiral. Sulfur is soluble in carbon disulfide and is insoluble in water.

The complexity of rhombic sulfur allows easy supercooling of the monoclinic form and has been used as a means of providing experimental evidence to substantiate Nernst's heat theorem, also known as the third law of thermodynamics. This states that, at 0 K, the entropy of any homogeneous substance, which is in complete internal equilibrium, can be taken as being zero. Summation of the increase in molar entropy of rhombic sulfur when heated from 0 K to the temperature of transformation (38.86 J K^{-1}), the molar entropy of transformation of rhombic to monoclinic sulfur at this temperature (1.09 J K^{-1}), and the change in the molar entropy of monoclinic sulfur when cooled from the transformation temperature to 0 K (−37.8 J K^{-1}) gives the molar transformation of rhombic to monoclinic sulfur at 0 K as 0.15 J K^{-1}. This is less than the experimental error involved and, thus, is taken as an experimental verification of Nernst's heat theorem.

The Polymorphism of Silica

Silicon, occurring below carbon in group 4 in the periodic chart of the elements, undergoes sp^3 hybridization of the 3*s* and 3*p* electrons to produce energetically equivalent electron orbitals, which point to the corners of a tetrahedron. In the crystal structure of silica (SiO_2), the sp^3 orbitals of neighboring silicon atoms form bonds with oxygen atoms, such that each silicon atom is tetrahedrally coordinated by four oxygen atoms and each oxygen atom is bonded to two silicons. The basic building block in the silica structure is thus a tetrahedron, at the center of which is a silicon atom with oxygen atoms located at the four corners. The polymorphism of silica arises from the number of ways in which tetrahedra can be arranged to fill space and one such arrangement is shown in **Figure 9**. The phase diagram for silica, presented as **Figure 10**, shows that, at pressures less than 10^4 atm, the polymorphs are low quartz, high quartz, tridymite, and cristobalite. At pressures between 10^4 and 10^5 atm, silica exists as the polymorph coesite (first described by L Coes, Jr. in 1953) and at pressures higher than 10^5 atm, stishovite is the polymorph. Stishovite differs from the other polymorphs, in that the silicon is in octahedral coordination with oxygen. This polymorph was discovered in meteorites found in Russia by Stishov. It is believed that the high force exerted on the meteorite by contact with the surface of the earth caused the tetrahedral coordination in silica to transform to an octahedral coordination.

The differences in standard free energies of formation, $\Delta(\Delta G^\circ)$, among the various polymorphs of silica, using cristobalite as the reference state, are shown in **Figure 11**. The relatively small differences in the standard free energies among the polymorphs arise because the enthalpies of formation (the –Si–O–Si– bond energies) of the polymorphs and the configurational entropies of packing SiO_4 tetrahedra together are similar.

Figure 11 shows another representation of the polymorphism of silica. The horizontal arrows represent reconstructive transformations, which require bond breaking and complete rearrangement of the SiO_4 tetrahedra. Thus, although the differences in the standard free energies of formation of the polymorphs, shown in **Figure 11**, are small, the high energies of activation for these transformations are such that the high-temperature polymorphs can be undercooled easily. Used extensively in laboratory and medical applications, silica glass is obtained by supercooling liquid silica to form a metastable amorphous phase. The vertical arrows in **Figure 12** represent displacive transformations, which do not involve the breaking of bonds and, consequently, occur at relatively high rates. The sequence of transformations during gradual heating of β-quartz depends on the purity of the quartz. When high-purity quartz

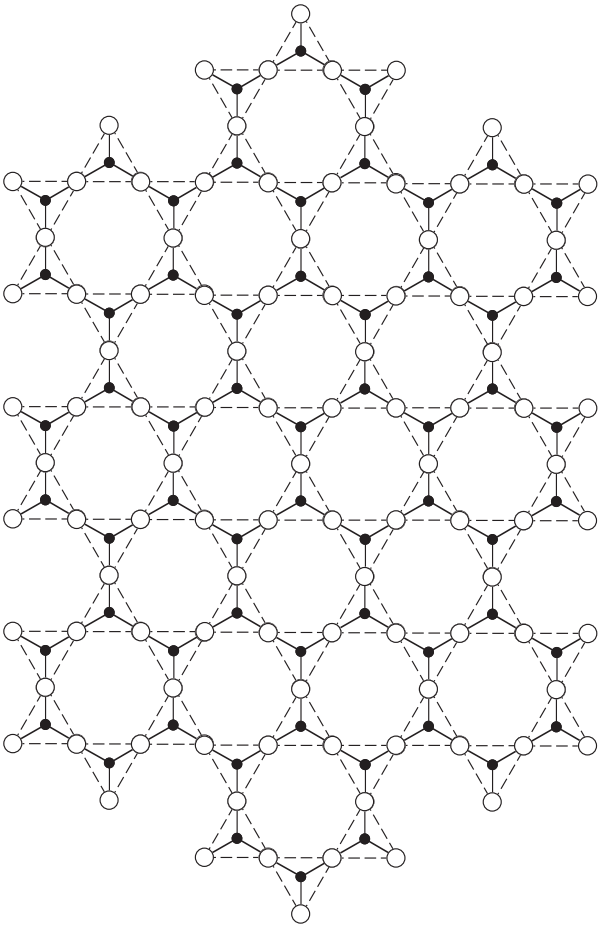


Figure 9 The arrangement of SiO_4 tetrahedra.

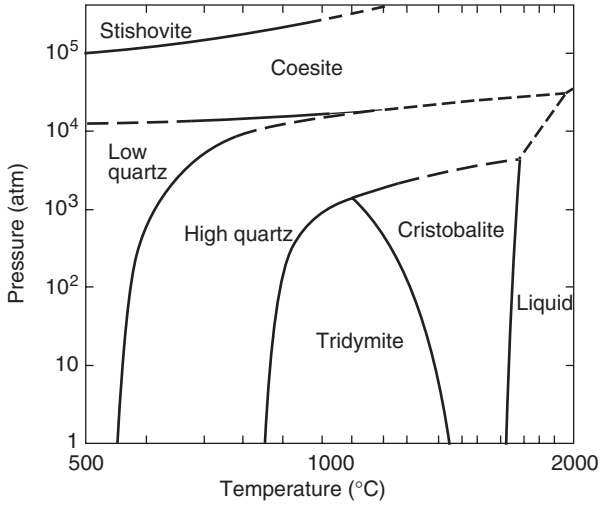


Figure 10 The phase diagram for silica.

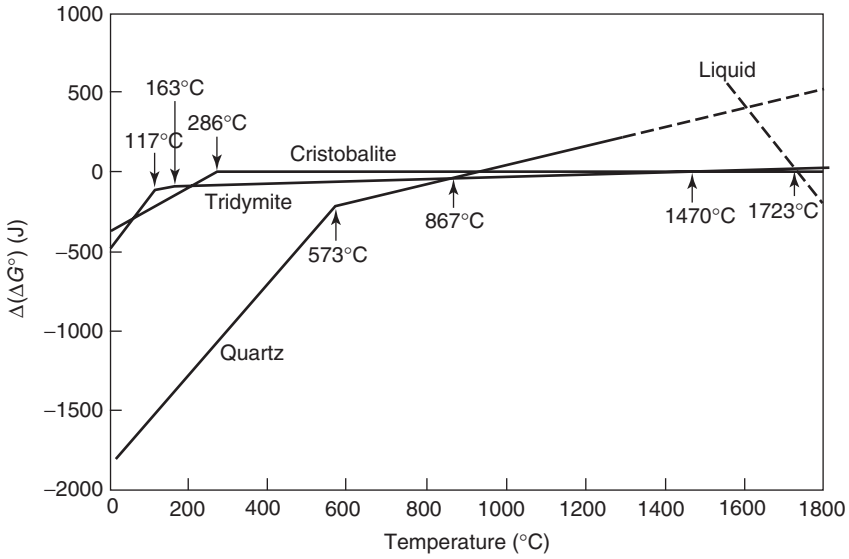


Figure 11 The differences in standard free energies of formation, $\Delta(\Delta G^\circ)$, among the various polymorphs of silica, using cristobalite as the reference state.

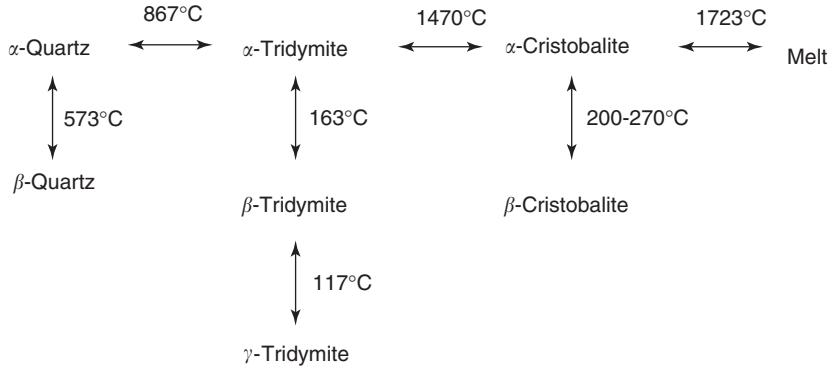


Figure 12 Phase transformations in silica.

($<10^{-2}$ wt.% impurities) is heated, the β-quartz transforms rapidly into α-quartz at 573°C. The α-quartz is stable to ~1025°C, at which temperature it transforms into α-cristobalite. However, if the quartz contains impurities in solid solution, α-quartz transforms to α-tridymite at ~870°C, which, on further heating, transforms into α-cristobalite at 1470°C.

The observation of the influence of impurities and the fact that most of the laboratory studies of phase equilibria in silica were conducted under hydrothermal conditions cast doubt on the validity of the phase relations shown in **Figure 10**. Evidence, which suggests that tridymite is not a stable phase in pure silica, but owes its existence to the presence of impurity ions in the structure, has been provided. This would make tridymite a polytype (same structure, different composition) rather than a polymorph. Two opinions have been expressed on this topic. One suggests the elimination of the tridymite phase and places the temperature of equilibrium between quartz and cristobalite at ~1025°C. The other suggests that a sharp polymorphic transformation exists between stable tridymite (designated "tridymite-S") and stable cristobalite at 1470°C and that quartz and stable tridymite coexist at ~870°C. The latter opinion is in accordance with the "classical" picture shown as **Figure 10**.

The precious mineral opal, which has the formula $SiO_2 \cdot nH_2O$, is a form of silica that is wholly amorphous.

The Polymorphism of Zirconia

Zirconia, ZrO_2, has three polymorphs at atmospheric pressure: a high-temperature cubic structure which exists from 2370°C to the melting temperature of 2680°C, a low-temperature monoclinic structure which exists at temperatures lower than 1174°C,

and an intermediate tetragonal structure which is stable between 1174°C and 2370°C. The high-temperature cubic form has the fluorite, CaF_2, structure shown in **Figure 13**. In this structure, the Zr^{4+} cations occupy the corner and face-centered positions in the f.c.c. lattice and the O^{2-} anions occur at the centers of the eight subcells described in **Figure 5**. The unit cell, thus, contains four Zr^{4+} and eight O^{2-} ions. The intermediate tetragonal form is a slightly deformed fluorite structure, and the low-temperature monoclinic form, which occurs naturally as the mineral baddeleyite, has a structure in which Zr ion is coordinated by seven oxygen ions. This structure is shown in **Figure 14**. The densities and specific volumes of the three polymorphs are listed in **Table 1**.

The 4.6% decrease in density accompanying the tetragonal to monoclinic transformation cannot be sustained and the solid exfoliates by means of a martensite shearing mechanism. The temperature at which this transformation occurs on cooling decreases with decreasing particle size. However, the high-temperature cubic form can be stabilized to room temperature by the substitution, for Zr^{4+}, of appropriately sized cations of valence less than 4. The cations Ca^{2+} and Y^{3+} are used to produce lime-stabilized zirconia (LSZ) and yttria-stabilized zirconia (YSZ). LSZ is considerably less expensive than YSZ. Several versions of the phase diagram for the system ZrO_2–CaO have been proposed and two of them are shown in **Figures 15** and **16**.

Figure 15 shows the existence of the three polymorphs of ZrO_2 and contains the line compound $CaO \cdot 4ZrO_2$, which undergoes incongruent decomposition at 1310°C. **Figure 15** also shows that the

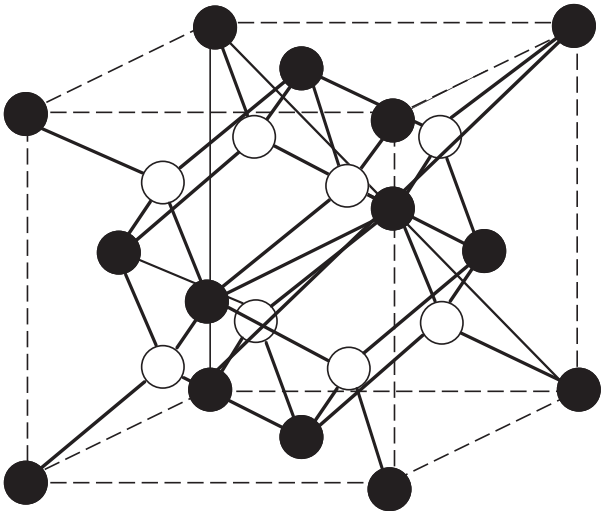


Figure 13 The unit cell for the zirconia cubic (fluorite) crystal structure. The solid and white circles represent Zr and O, respectively.

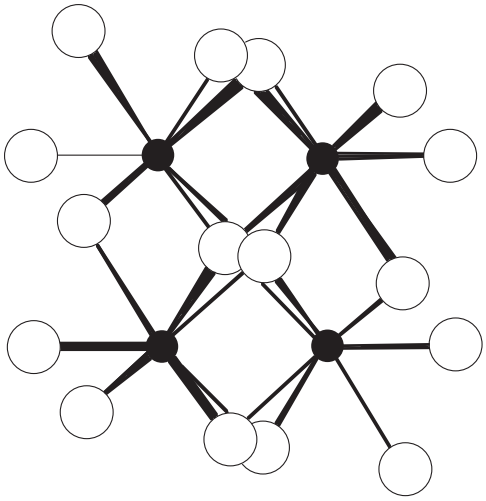


Figure 14 The crystal structure of monoclinic zirconia (baddeleyite). The solid and white circles represent Zr and O, respectively.

Table 1 The densities and specific volumes of the polymorphs of zirconia

Polymorph	ρ (kg m^{-3})	V (m^3 kg^{-1})
Cubic	6090	0.001 63
Tetragonal	6100	0.001 64
Monoclinic	5830	0.001 72

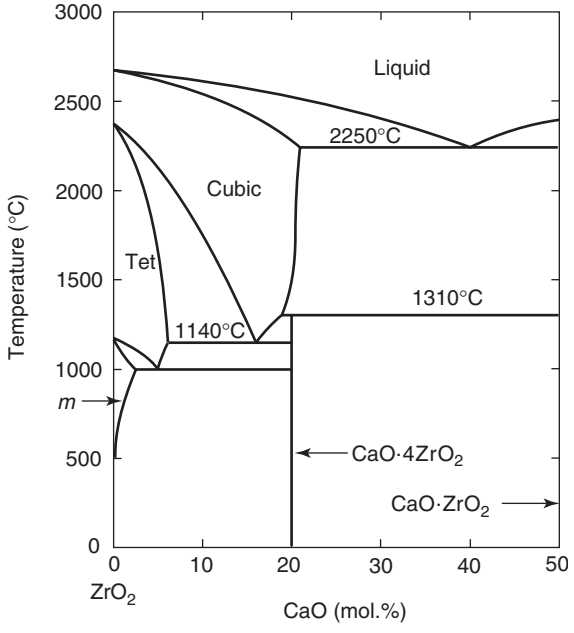


Figure 15 One version of the phase diagram for ZrO_2–CaO.

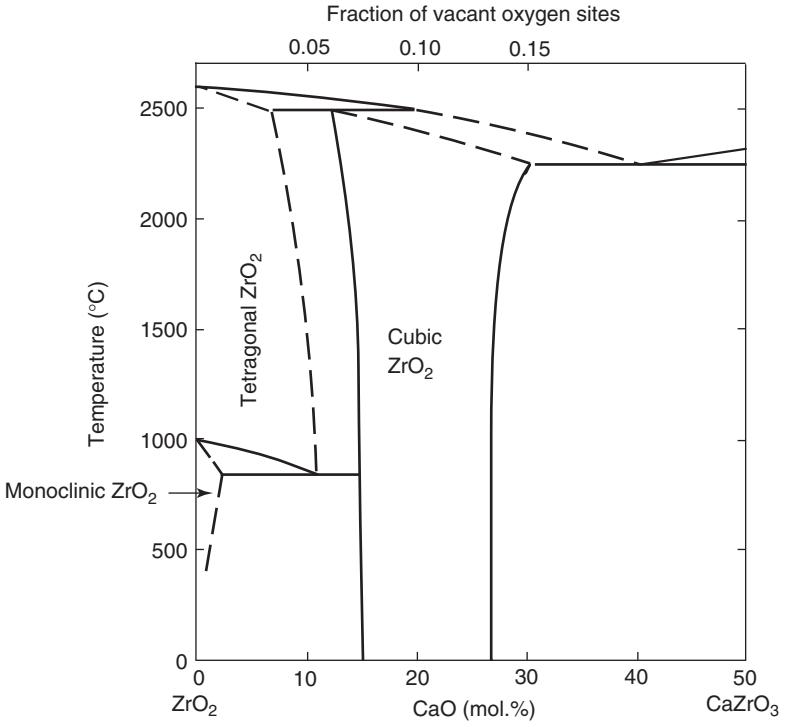


Figure 16 Another version of the phase diagram for the system ZrO_2–CaO.

substitution of Ca^{2+} ions for Zr^{4+} in cubic zirconia stabilizes the cubic structure only to 1140°C, at which temperature it undergoes a eutectoid decomposition to tetragonal zirconia and $CaO \cdot 4ZrO_2$. It is known that the substitution of Ca for Zr stabilizes the cubic structure to room temperature, but it is unlikely that supercooled cubic zirconia will maintain a metastable existence at temperatures near 700°C, at which it is used as an oxygen sensing device. On the other hand, **Figure 16** shows the existence of lime-stabilized-cubic zirconia at room temperature, but does not show the existence of cubic zirconia in the pure state. **Figure 17** shows the phase diagram for the system ZrO_2–$YO_{1.5}$, in which it is seen that the substitution of Zr^{3+} ions for Zr^{4+} ions stabilizes the cubic phase at room temperature.

Stabilized cubic zirconia is a practical ceramic material used as the solid-state electrolyte in a cell used for measuring the partial pressure (or thermodynamic activity) of oxygen in a gaseous or liquid medium. Electroneutrality requires that the substitution of a Ca^{2+} ion for a Zr^{4+} ion in cubic zirconia be accompanied by the formation of a vacant site of the oxygen anion sublattice. Hence, x moles of CaO + $(1-x)$ moles of ZrO_2 contain $(x + 2 - 2x) = (2 - x)$ moles of O^{2-} and x moles of vacant sites. Thus, the fraction of vacant sites is $x/(2 - x + x) = 0.5x$. **Figure 16** shows that the percent of vacant oxygen sites in stabilized cubic zirconia can be as high as 12–13%. This high percentage of vacant sites imparts high diffusivity of oxygen in cubic zirconia at elevated temperatures. The working of a lime-stabilized cubic zirconia EMF cell is illustrated in

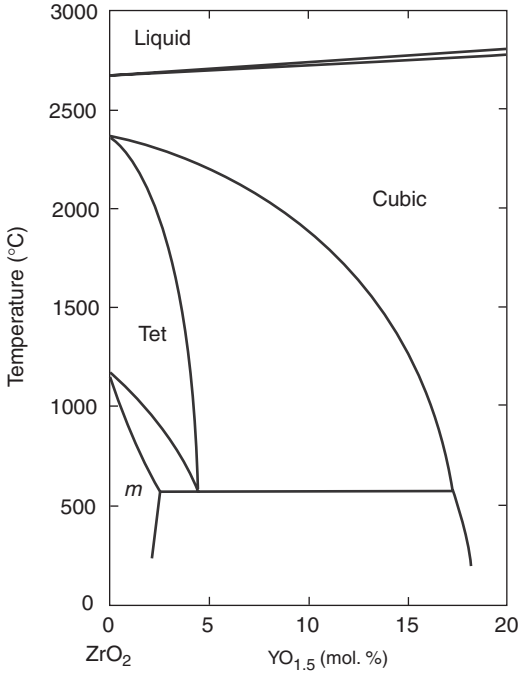


Figure 17 The phase diagram for the system ZrO_2–$YO_{1.5}$.

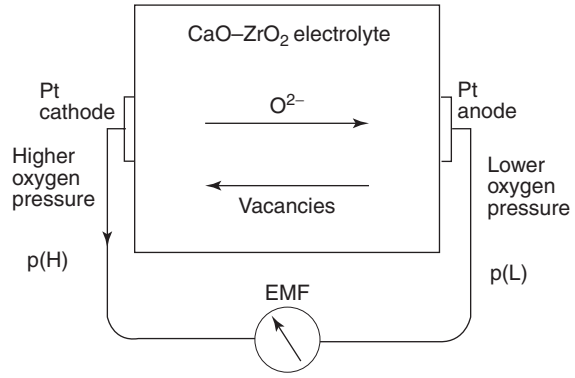


Figure 18 The working of an LSZ solid-state oxygen sensor.

Figure 18. The half-cell reaction at the cathode is

$$\tfrac{1}{2}O_{2(gas)} + 2e^- \rightarrow O^{2-}$$

and the half-cell reaction at the anode is

$$O^{2-} \rightarrow 2e^- + \tfrac{1}{2}O_{2(gas)}$$

This gives the cell reaction as $O_2(P = P_{high}, T) = O_2(P = P_{low}, T)$ for which

$$\Delta G = RT \ln \frac{P_{low}}{P_{high}}$$

and as

$$\Delta G = -zFE = RT \ln \frac{p_{O_{2(low)}}}{p_{O_{2(high)}}}$$

then

$$E = -\frac{RT}{4F} \ln \frac{P_{O_{2(low)}}}{p_{O_{2(high)}}}$$

where E is the measured EMF of the cell, F is Faraday's constant, and z is the number of electrons transferred by the electrochemical cell reaction. In the use of the oxygen sensor where the value of one of the oxygen pressures is known, measurement of the EMF of the cell and the temperature allows the unknown oxygen pressure to be determined. The oxygen content of liquid steel is measured routinely by immersion of a disposable LSZ cell in the steel bath in the converter.

Stabilized cubic zirconia is also used as an electrolyte in a hydrogen generator, in which water vapor is decomposed to produce hydrogen gas at the cathode and oxygen gas at the anode. Also, it is used in a fuel cell in which electric power is produced by the oxidation of CO or H_2 at the anode and the reduction of oxygen gas at the cathode.

See also: Alloys: Iron; Alloys: Overview; Ceramic Materials; Electronic Structure (Theory): Molecules; Irreversible Thermodynamics and Basic Transport Theory in Solids; Molecular Crystallography; Phase Transformation; Phases and Phase Equilibrium; Thermodynamic Properties, General.

PACS: 82.60.Fa; 81.30.Bx; 81.30.Dz; 61.50.Ks; 81.05.Je

Further Reading

Bragg L and Claringbull GF (1965) *The Crystalline State, Vol IV: Crystal Structures of Minerals*. Cornell, NY: Cornell University Press.

Levin EM, Robbins CR, and McMurdie HF (1964, 1969) Phase diagrams for ceramists. *Amer. Ceramic Society* (suppl.). Metals Park, OH: American Society for Materials.

Massalski TB (ed.) (1986) *Binary Alloy Phase Diagrams*, vols. 1 and 2. Metals Park, OH: ASM.

Weast RC and Astle MJ (eds.) (1996) *CRC Handbook of Chemistry and Physics*, 77th edn. Boca Raton, FL: CRC Press.

Nomenclature

c_p	constant pressure molar heat capacity ($J K^{-1}$)
E	electromotive force (V)
F	Faraday's constant = 96 487 C mol^{-1}
G	molar Gibbs free energy (J)
H	molar enthalpy (J)
P	pressure (atm or Pa)
p^0	saturated vapor pressure (atm or Pa)
R	universal gas constant (8.3144 $J K^{-1} mol^{-1}$)
S	molar entropy ($J K^{-1}$)
T	temperature (K)
V	specific volume ($m^3 kg^{-1}$)
z	number of electrons transferred in an electrochemical reaction
ρ	specific density ($kg m^{-3}$)

Alloys *See* Alloys: Aluminum; Alloys: Copper; Alloys: Iron; Alloys: Magnesium; Alloys: Overview; Alloys: Titanium.

Alloys: Aluminum

E A Starke Jr., University of Virginia, Charlottesville, VA, USA

Introduction

Aluminum is the third most abundant element in the earth's crust but because it is so reactive with other elements it is not found in the native state. Hans Christian Oersted (1777–1851), a Danish physicist and chemist, was successful in isolating aluminum in a pure form in 1835. Sir Humphrey Davy (1778–1829) had previously been unsuccessful at such attempts, but it was Davy who named the element "aluminum," the name used in the US. The rest of the world uses the term "aluminium." Until 1886 a chemical process was used to produce aluminum, which employed crystals of corundum that were chemically converted to aluminum chloride and then reduced with metallic sodium to form salt and metallic aluminum. From 1825 to 1886, aluminum was primarily used for jewelry and as expensive tableware due to the difficulty and cost of extracting it from its ore. During the construction of the Washington Monument, the world's tallest structure of that period, a material was needed to top off the structure and to serve as a lightning rod. A cast aluminum pyramid was produced by William Frishmuth in 1884 and mounted on the top. It was the largest aluminum casting ever produced and was the first architectural application of this metal.

In 1886, Charles Martin Hall (1863–1914) in the US and Paul-Louis Toussaint Héroult (1863–1914) in France simultaneously developed an economical electrochemical method of producing aluminum, which ultimately led to its widespread use throughout the world. Their invention replaced the chemical reduction process and lowered the metal's cost from \$15/pound in 1884 to \$0.50/pound in 1890. The versatility of aluminum has resulted in it replacing many older, more established materials, and it is now consumed, on a volumetric basis, more than all other nonferrous metals combined, including copper, lead, and zinc.

Aluminum is light, ductile, has good electrical and thermal conductivity, and can be made strong by alloying. It has a low density of $2.7\,g\,cm^{-3}$ compared to that of iron ($7.9\,g\,cm^{-3}$) due to its low atomic mass of 27. The ductility and formability of aluminum is due to the high symmetry and thermodyic stability of the face-centered cubic (f.c.c.) lattice and its high stacking-fault energy. In the pure form, aluminum has a low stiffness, E, of 70 GPa compared to 211 GPa for iron, and low tensile strength, 80 MPa, compared to 300 MPa for iron. However, its specific modulus, that is, modulus divided by density, is almost equal to that for iron, titanium, and magnesium. An advantageous chemical property of aluminum is its reactivity with oxygen, which leads to the formation of a dense layer of Al_2O_3 on the surface, which shields the base metal from further environmental interactions.

Pure aluminum is not used commercially because of its low strength, but this property can be improved by alloying elements. A major metallurgical development occurred in 1906 when Alfred Wilm (1869–1937) discovered the process of "age hardening" in aluminum alloys. Wilm was conducting research directed toward improving the strength of aluminum alloys. He knew that steel could be strengthened if the right compositions were cooled fast enough from high temperatures, so following this recipe he heated some alloys of aluminum containing 3.5–5.5 wt.% copper, plus less than 1% magnesium and manganese, to a high temperature and quenched them in water. To his frustration, many of the alloys he tested were softer after quenching than before the heat treatment. However, after a few days he found that their hardness and tensile properties had increased considerably. One of the alloys, designated Duralumin, is still in use today. Much later, in 1919, Merica, Waltenberg, and Scott explained the phenomenon, which is due to nano-sized clustering and precipitation of solute atoms from the supersaturated solid solution. However, these nano-microstructural features are too small to be resolved by optical microscopy and were only inferred by the X-ray diffraction studies of Guinier and Preston in 1938. Direct proof was not obtained until the development of transmission electron microscopy in 1959. Duralumin is probably the first example of a nanostructured material developed by humans.

Aluminum alloys are classified as heat-treatable or non-heat-treatable, depending on whether or not they undergo precipitation (age) hardening. There are also two different product forms of aluminum alloys; wrought alloys that have been worked or deformed after casting and casting alloys, which are used in the "as-cast" condition. Wrought aluminum alloys are normally designated by a four-digit numerical system developed by the Aluminum Association. The nomenclature has now been accepted by most countries and is called the International Alloy Designation System (IADS). The system used for wrought alloys is slightly different from that used for

cast alloys; however, the first digit designates the alloy group and is essentially the same for both wrought and cast alloys. The alloy group is associated with the major alloy addition and the second digit indicates modification of the original alloy or impurity limits. The last two digits identify the specific aluminum alloy. Experimental alloys also use this system, but are indicated as experimental by the prefix X. The 1XXX alloys contain a minimum of 99% aluminum. The major alloying addition in the 2XXX series is copper; it is manganese for 3XXX, silicon for 4XXX, magnesium for 5XXX, magnesium and silicon for 6XXX, zinc for 7XXX, and 8XXX is used for others (tin for casting alloys). A first digit of 9 is not used for wrought alloys but is used for other alloy additions for casting alloys.

The designation system is slightly different for casting alloys although, as mentioned, the first digit still refers to the major alloying element. The second and third digits serve to identify a particular composition. These three digits are followed by a decimal point, which is followed by a zero to indicate a casting. Often, a letter prefix is used to denote either an impurity level or the presence of a secondary alloying element. These letters are assigned in alphabetical sequence starting with A but omitting I, O, Q, and X. X is reserved for experimental alloys. For example, A201.0 has a higher purity than the original 201.0.

The heat-treatment or temper-nomenclature system developed by the Aluminum Association has also been adopted as part of the IADS by most countries. It is used for all forms of wrought and cast aluminum alloys with the exception of ingot. The system is based on the treatments used to develop the various tempers and takes the form of letters added as suffixes to the alloy number. One or more digits following the letter indicate subdivisions of the tempers, when they significantly influence the characteristics of the alloy. Alloys supplied in the as-fabricated or annealed condition are designated with the suffixes F and O, respectively. The letter W designates those supplied in the solution heat-treated condition. Alloys supplied in the strain-hardened condition are designated with the letter H and those in the heat-treated condition with the letter T. Digits following H represent the degree of strain hardening and those following T the type of aging treatment.

Processing of Aluminum Alloys

In fabricating aluminum alloy products, the aluminum alloy composition is made by adding the alloying elements to molten aluminum, usually in the form of a concentrated hardener or master alloy, with the requisite purity. Two types of alloying elements are normally added: those for strength and those to control the grain structures that precipitate as "dispersoids." During solidification of the ingot, some of the alloying elements and impurities may precipitate out of the aluminum, forming coarse "constituent" particles within the ingot. For wrought products, additional precipitation of the alloying elements may occur as the ingot is worked down to the final product form (such as sheet or plate).

The "constituent" particles potentially impair the properties of the final product in several ways. By tying up alloying elements that are added on purpose to develop the desired properties, these elements are not available to impart strength and other beneficial properties to the aluminum alloy. In addition, the constituent particles are more brittle than the surrounding aluminum and, by fracturing under stress, form and promote the growth of cracks and hence impair beneficial mechanical properties. One typical goal of aluminum alloy processing is to reduce the size and amount of the constituent particles in the aluminum alloy product.

Beginning with the as-cast ingot, conventional processing includes a homogenization treatment. This is the stage where dispersoids normally form. For alloys used in the as-cast condition, homogenization may be the final treatment for non-heat-treatable alloys, but for age-hardenable alloys further heat treatments will be required. For wrought alloys, hot working follows the homogenization for ingot breakdown and shape change to the appropriate product form. Since aluminum and its alloys have high stacking-fault energy, sufficient dynamic recovery normally occurs during hot deformation to give rise to a stable polygonized substructure, an example of which is shown in the transmission electron micrograph of **Figure 1**. The grains elongate in the direction of metal flow and do not recrystallize. An example of the grain structure of a hot-worked aluminum-alloy plate is shown in **Figure 2**. The dispersoids, and any constituents and primary phases that may be present, are strung out in the working direction; their spacing increases in the working direction and decreases normal to the working direction. In some cases large constituent phases, if present, are broken up during the working operation. The fine distribution of dispersoids associated with Cr, Mn, or Zr additions delays or prevents static recrystallization and aids in retaining the elongated or "pancake-shaped" grains during subsequent processing. After processing for shape change, non-heat-treatable alloys may then be cold worked and/or annealed to develop the desired strength, and heat-treatable alloys may be solutionized, quenched, in some cases stretched to remove residual stresses developed during quenching or for

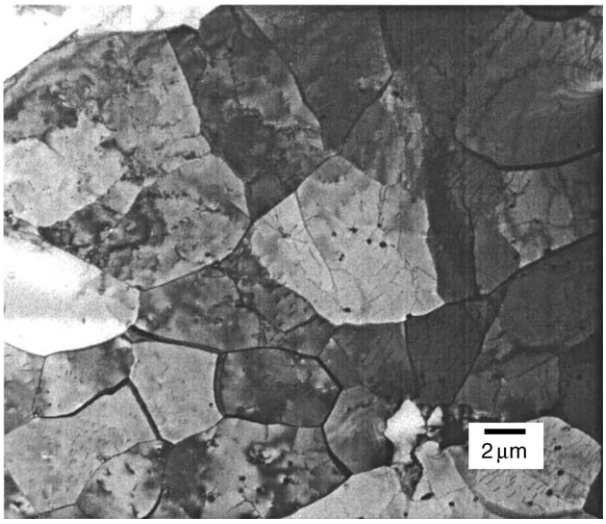


Figure 1 Subgrain structure in a hot-rolled Al–Cu–Mg–Li–Zr alloy.

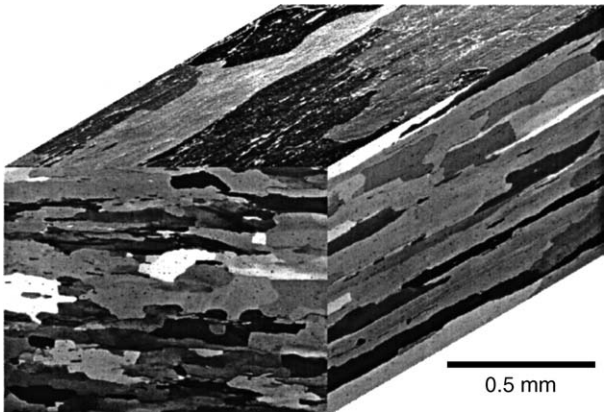


Figure 2 Optical micrograph illustrating the pancake grain structure of a hot-rolled Al–Cu–Mg–Li–Zr alloy.

creating nucleation sites for precipitation, and either naturally aged at room temperature or artificially aged at some moderate temperature. An Al–2.7Cu–1.8Li–0.6Zn–0.3Mg–0.3Mn–0.12Zr alloy naturally aged at room temperature after a rapid quench from the solution heat-treatment temperature is shown in **Figure 3** and the same alloy artificially aged at 150°C for 6 h after a rapid quench from the solution heat-treatment temperature is shown in **Figure 4**.

Alloys that are susceptible to corrosion in moist air may be clad with pure aluminum, which is not susceptible and can offer protection in an aggressive environment, or they may be anodized. Anodizing is a method for producing a very thick, protective oxide film on the surface of aluminum. The anodized film normally contains chemical compounds such as chromates, which are collected from the anodizing bath and render the film more corrosion resistant than films that form naturally in air.

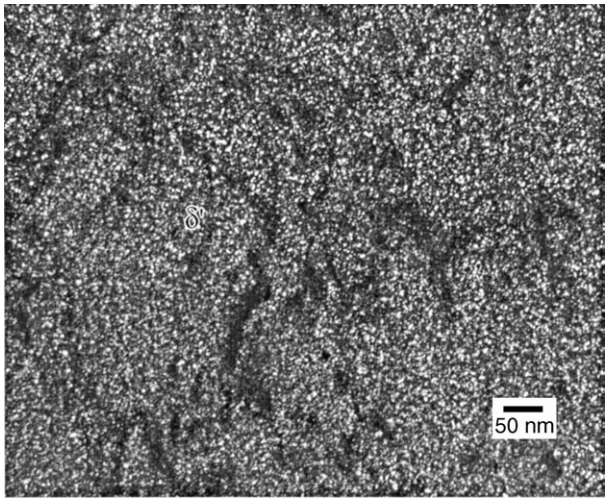


Figure 3 Dark-field TEM micrograph of the matrix precipitation from an Al–2.7Cu–1.8Li–0.6Zn–0.3Mg–0.3Mn–0.12Zr alloy naturally aged for seven days at room temperature after a rapid quench from the solution heat-treatment temperature. The matrix precipitation is dominated by fine δ' (Al_2Li) spheres.

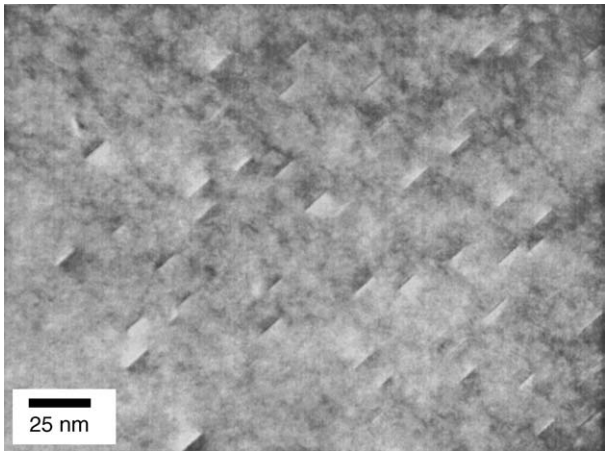


Figure 4 Bright-field TEM micrograph of the matrix precipitation from an Al–2.7Cu–1.8Li–0.6Zn–0.3Mg–0.3Mn–0.12Zr alloy (T6 temper) artificially aged at 150°C for 6 h after a rapid quench from the solution heat-treatment temperature. The matrix precipitation is dominated by fine θ''/θ' (Al_2Cu) plates and δ' spheres (not shown).

Non-Heat-Treatable Alloys

The 1XXX, 3XXX, 5XXX, and some of the 8XXX series alloys are non-heat-treatable. These alloys are primarily strengthened by elements in solid solution and by deformation structures. One of the first non-heat-treatable alloys, other than the 1100 alloy that only contained 0.95 wt.% Fe and Si as impurities, was an alloy containing 1.25 wt.% Mn and 0.12 wt.% Cu, now designated as 3003, that was introduced in 1906, the same year that Wilm discovered age hardening. Alloy 3003 can develop yield

strength up to 400 MPa by special thermomechanical treatments. Magnesium is added to non-heat-treatable alloys for its solid-solution strengthening effect. It also enhances work hardening and makes aluminum more anodic. In general, solid-solution alloys are more resistant to corrosion than two-phase alloys. Al–Mg alloys have a high resistance to corrosion, particularly in seawater and alkaline solutions. However, when magnesium exceeds the solid solubility in binary alloys, it precipitates at grain boundaries as Al_3Mg_2, which is anodic to the matrix and promotes intergranular attack.

Aluminum combines readily with transition metals, Ti, V, Cr, Mn, Fe, Co, Ni, and Zr to form intermetallic phases with little or no solubility in the aluminum matrix. The intermetallics phases increase the strength by enhancing work hardening during working operations, and by refining the grain structure. They increase the work hardening since they are usually incoherent with the matrix, are nondeformable, and must be looped or bypassed by moving dislocations. This increases the dislocation density and blocks the dynamic recovery processes. During hot working, the subgrain structure that is developed can also increase the strength. The smaller (less than 0.6 μm) intermetallic dispersoids aid in the stabilization of the substructure. The intermetallic particles generally do not add a component of the particle strengthening because of their low volume fraction, large size, and interparticle spacing.

It is customary to control both composition and cooling rates in order to prevent large primary phases from forming. For example, commercial Al–Mn alloys most often contain less than 1.25 wt.% Mn, although as much as 1.82 wt.% is soluble in pure aluminum. The 1.25 wt.% limit is imposed because Fe, present in most aluminum alloys as an impurity, decreases the solubility of Mn in aluminum. This increases the probability of forming large primary particles of Al_6Mn, which can have a disastrous effect on ductility.

Silicon is a principal addition to most aluminum casting alloys because it increases the fluidity of the melt. In the solid state, the hard silicon particles are the major contributor to the strength of non-heat-treatable casting alloys. An additional improvement in strength can be obtained by minor additions of Mg or by trace alloy additions, for example, Na that refines the cast structure. The latter also minimizes porosity and increases ductility.

Heat-Treatable Alloys

The 2XXX, 4XXX, 6XXX, 7XXX, and some 8XXX series alloys are considered heat-treatable. These alloys contain elements that decrease in solubility with decreasing temperature, and in concentrations that significantly exceed their equilibrium solid solubility at room and moderately higher temperatures. One of the transition elements Cr, Mn, or Zr, is added to age-hardenable wrought alloys to control the grain structure; however, the grain and deformation structure play only a secondary role in the strengthening of this class of materials.

A normal heat-treatment cycle after deformation processing includes a soak at a high temperature to dissolve the soluble constituent particles and any other precipitates that may be soluble (dispersoids containing the elements Cr, Mn, or Zr have already precipitated at this point). After the aluminum alloy has been held at the solutionizing temperature for a sufficient time, the alloy is rapidly cooled or quenched to prevent precipitation of coarse primary phases and to obtain a solid solution supersaturated with both solute elements and vacancies. The next step involves aging at room temperature (natural aging) or at an intermediate temperature (artificial aging).

Quenching is accompanied by a change in free energy, which increases progressively as the difference between the solutionizing temperature and the quenching temperature increases. The volume free energy change is the driving force for precipitation and is associated with the transfer of solute atoms to a more stable phase. However, when precipitation occurs there are other factors, which increase the free energy, that is, formation of the interface between the matrix and the precipitate requires an increase in the surface free energy, and, if there is a volume change or interfacial strains associated with the precipitate, there is an increase in elastic strain energy. The change in free energy when a precipitate forms, ΔG, is the sum of these free energy changes and can be expressed mathematically as

$$\Delta G = V\Delta G_V + A\gamma + V\Delta G_s \qquad [1]$$

where V is the volume of the new phase; ΔG_V is the free energy decrease due to creation of a volume, V, of the precipitates and is therefore negative; A is the area of the interface between the matrix and the precipitate; γ the energy of the new surface formed; and $V\Delta G_s$ is the increase in elastic strain energy per unit volume of precipitate.

The critical increase in free energy, ΔG^*, required for a nucleus to become an equilibrium or metastable precipitate is known as the activation energy barrier for nucleation and can be expressed as

$$\Delta G^* = (16\pi\gamma^3)/(3[\Delta G_V - \Delta G_s]^2) \qquad [2]$$

The nucleation rate, N, can be expressed as

$$N = C \exp(-Q/kT) \exp(-\Delta G^*/kT) \qquad [3]$$

where C is the number of nucleating sites per unit volume, Q is the activation energy for diffusion, k is the Boltzmann constant, and T the absolute temperature.

There are a number of requirements for an effective age-hardenable alloy. First, there has to be a sufficient volume fraction of the second phase for the desired strength. An examination of aluminum binary diagrams suggests that the most attractive alloying additions would be, in decreasing volume fraction of the second phase, Ag, Mg, Li, Zn, Ga, and Cu. The required volume fraction will depend on the effectiveness of the second phase on impeding dislocation motion, that is, whether it is a "soft" or "hard" phase, and its size and spacing. Second, the aging potential must be adequate and the nucleation of the precipitates should be as close as possible to being homogeneous. This is generally accomplished by the addition of a second alloying element. The addition of the second alloying element may: (1) reduce the solubility of the first element, (2) increase the diffusion rates by trapping vacancies, (3) increase the driving force for nucleation, or (4) reduce the activation energy against nucleation. If the strain term in eqn [1] is large, nucleation may be aided by defects, for example, dislocations. For this type of nucleation, the precipitate distribution depends on the distribution of the defects.

Microstructural Features

Chemical composition and processing control the microstructure and thus the physical, mechanical, and corrosion properties of aluminum alloy products. In general, microstructural features of importance for property control include the following:

1. *Coherency, volume fraction, and distribution of strengthening precipitates*. These form intentionally during aging treatments and unintentionally during quenching. The precipitates normally range in size from 1 to 10 nm.
2. *Size, distribution, and coherency of dispersoid particles*. These form during the ingot preheat or homogenization treatment by precipitation of the transition elements chromium, manganese, or zirconium. They are present by design to control the grain structure and degree of recrystallization. Dispersoids normally range in size from 10 to 200 nm.
3. *Degree of recrystallization*. This is determined by the thermomechanical history. During bulk metal deformation processing at elevated temperatures, the dislocations in aluminum alloy products arrange themselves into a structure of subgrains by the process known as dynamic recovery. As temperature decreases and strain rate increases, the stored energy increases. This increase in stored energy is manifested by an increase in number and consequent decrease in the size of subgrains and an increase in the dislocation density in cell walls. During subsequent heat treatments, the stored energy of deformation may decrease by either static recovery or recrystallization. The final product may be completely unrecrystallized, partially recrystallized, or completely recrystallized.
4. *Grain size and shape*. These are also controlled by thermomechanical history and may be influenced by ingot grain refining practice.
5. *Crystallographic texture*. The deformation process used to produce the product and the thermomechanical history determines this.
6. *Intermetallic constituent particle*. These form by a liquid–solid eutectic reaction during solidification, primarily from impurities, for example, iron and silicon. Because the low solubility of iron in pure aluminum is reduced by alloying elements, constituent particles containing iron are insoluble. The size and size distribution of insoluble constituent particles are controlled by the rate of ingot solidification, the chemical composition, and the extent and nature of bulk deformation. Constituents generally range in size from 1 to 30 μm.
7. Porosity may be a factor in thick plates and, particularly, in castings. Individual pores can range up to and beyond 200 μm.

Aluminum Alloy Design

After Wilm discovered age hardening of the Al–Cu–Mg–Mn alloy through serendipity, traditional aluminum alloy development primarily relied on a trial-and-error philosophy. Most aluminum alloys were developed by simply identifying a given alloying element that led to gains in a certain mechanical property, typically strength, and continuing to add that element until the property began to degrade. The amount of secondary element used would then be the percentage that corresponded with maximum performance. This process would then follow in ternary and higher-order systems in order to optimize a given mechanical property. Often this occurred with little knowledge of the corresponding microstructural components or appropriate thermodynamic understanding, for instance the phase field of operation. Unfortunately, this practice led to the slow insertion of novel alloys into various applications creating a disparity with the design community.

Table 1 Property–microstructure relationships in aluminum alloys

Property	*Desired microstructural feature(s)*	*Function of feature(s)*
Strength	Fine grain size with a uniform dispersion of small, hard particles	Inhibit dislocation motion
Ductility and toughness	Fine structure with clean grain boundaries and no large or shearable precipitates	Encourage plasticity and work hardening, inhibit void formation and growth
Creep resistance	Thermally stable particles within the matrix and on the grain boundaries	Inhibit grain boundary sliding and coarse microstructure
Fatigue crack initiation resistance	Fine grain size with no shearable particles and no surface defects	Prevent strain localization, stress concentrations, and surface slip steps
Fatigue crack propagation resistance	Large grain size with shearable particles and no anodic phases or hydrogen traps	Encourage crack closure, branching, deflection, and slip reversibility
Pitting	No anodic phases	Prevent preferential dissolution of second-phase particles
Stress corrosion cracking and hydrogen embrittlement	Hard particles and no anodic phases or interconnected hydrogen traps	Homogenize slip and prevent crack propagation due to anodic dissolution of phases

More recent aluminum-based alloy development has begun to utilize the knowledge of phase diagrams and the productive interplay between empirical research and theoretical science. Phase diagrams represent the state of an alloy as a function of temperature, pressure, and alloy concentration and can be used for alloy design and process development. There are currently software programs available that use thermodynamic functions for the calculation of phase diagrams. Such calculations reduce the effort required to determine equilibrium conditions in a multicomponent system. Materials design can now be viewed as the best application of available models for the prediction and synthesis of alloys with desired properties. Of course, all theoretical calculations must be verified experimentally, but the experiments should be based on sound theory and not on a trial-and-error approach.

The most general understanding of alloy design must derive itself from the metallurgical paradigm of processing–structure–property relationships. In order to optimize a given property, one must determine the necessary ideal microstructural components, which are dictated by the alloy composition and processing from solidification through final manufacture. Fundamentally, to properly determine the resulting microstructure for a given thermomechanical processing routine, the pertinent literature and equilibrium phase diagrams must be consulted. A summary of various empirical structure–property relationships found in aluminum alloys is summarized in **Table 1**. This table highlights the general structural features desired to optimize a given property.

Applications of Aluminum Alloys

Aluminum alloys are lightweight, resistant to atmospheric corrosion, and have good mechanical properties, which make them a popular material for use in a wide variety of applications. The 1XXX alloys are strain-hardenable, have high formability, corrosion resistance, and good electrical conductivity. They are often used for electrical applications and as foil and strip for packaging. The 2XXX alloys are heat-treatable, and possess excellent combinations of high strength and toughness, but are not resistant to atmospheric corrosion; so, they are usually painted or clad in such exposures. They are primarily used for aircraft and truck body applications. The 3XXX alloys have high formability, corrosion resistance, and can be joined by a variety of methods. They are widely used in cooking utensils, beverage cans, chemical equipment, roofing and siding, and in heat exchangers. The 4XXX alloys are heat-treatable and have good flow characteristics along with medium strength. They are primarily used in forging applications, for example, aircraft pistons and for welding wires. The 5XXX alloys are strain-hardenable and have excellent corrosion resistance, even in salt water, toughness, and weldability along with moderate strength. They find wide application in highway bridges, storage tanks, marine and automotive applications, and pressure vessels. The 6XXX alloys are heat-treatable, have high corrosion resistance, excellent extrudibility, and moderate strength. They are primarily used in building and construction, automotive and marine applications, and recently in aerospace applications. The 7XXX alloys are heat-treatable and can provide the highest strengths of all aluminum alloys. They are primarily used for automotive and aerospace applications. The 8XXX series is used for alloys that contain lesser-used alloying elements such as nickel and lithium. Some are used for electrical conductors and the lithium-containing alloys have found application in the aerospace industry due to lithium's effect in reducing density.

Summary

Aluminum alloys possess a number of very attractive characteristics that, together with their light weight, make them extremely attractive for many applications. Since aluminum comprises 8% of the Earth's crust by weight, the overall reserves are adequate to cope with anticipated demands for the foreseeable future. Cost differentials with respect to competing materials will probably dictate the extent to which aluminum will be used. The cost includes the price of the electrical energy needed for the extraction of the metal from its minerals, and this may be greatly reduced in the future since aluminum can be recycled for many applications. In addition, aluminum often pays off when lifecycle costs, instead of acquisition costs, are considered.

See also: Deformation Processing; Mechanical Properties: Strengthening Mechanisms in Metals; Phase Transformation; Phases and Phase Equilibrium; Recovery, Recrystallization, and Grain Growth; Solidification: Models and Simulations.

PACS: 60; 61.66; 62.20

Further Reading

Altenpohl DG (1998) *Aluminum: Technology, Applications, and Environment*. Washington, DC: The Aluminum Association.

Aluminum Association (1997) *Aluminum Standards and Data*. Washington, DC: The Aluminum Association.

Hatch JE (ed.) (1984) *Aluminum: Properties and Physical Metallurgy*. Metals Park, OH: American Society for Metals.

Hornbogen E (2001) Hundred years of precipitation hardening. *Journal of Light Metals* 1: 127–132.

Lorimer GW (1978) Precipitation in aluminum alloys. In: Russell KC and Aaronson HI (eds.) *Precipitation Processes in Solids*, pp. 87–119. New York: The Metallurgical Society (AIME).

Polmear IJ (1995) *Light Alloys, Metallurgy of the Light Metals*. London: Arnold.

Sanders RE Jr., Baumann SF, and Stumpf HC (1986) Non-heat-treatable aluminum alloys. In: Starke EA and Sanders TH (eds.) *Aluminum Alloys: Their Physical and Mechanical Properties*, vol. III, pp. 1441–1484. West Midlands, UK: EMAS.

Starke EA Jr. and Staley JT (1996) Application of modern aluminum alloys to aircraft. *Progress in Aerospace Science* 32: 131–172.

Vasudevan AK and Doherty RD (eds.) (1989) *Aluminum Alloys – Contemporary Research and Applications*. London: Academic Press.

Zhu AW, Gable BM, Shiflet GJ, and Starke EA Jr. (2002) The intelligent design of age hardenable wrought aluminum alloys. *Advanced Engineering Materials* 4(11): 839–846. Germany: Wiley-VCH Verlag GmbH.

Alloys: Copper

W F Hosford, University of Michigan, Ann Arbor, MI, USA

Copper Production

Copper is among the few metals known to ancient civilizations. Copper, gold, silver, platinum, and meteoric iron are the only metals found in their native state. However ~90% of copper comes from sulfide ores. Most copper ores contain 0.5% Cu or less, iron being the principal metallic impurity. The ores are ground to a fine powder and concentrated (normally by floatation) to 20–25% Cu. This concentrate is melted as a matte of mixed Cu and Fe sulfides containing up to 60% Cu. The matte is oxidized to remove the iron as an oxide and burn off the sulfur. The product is called blister copper which contains 98.5% Cu. Blister copper is fire refined to a tough-pitch (99.5% Cu). It is, often, then electrolytically refined to 99.95% + Cu. Au, Ag, and Pt metals are recovered from slime.

Chile is the single largest producer of copper. The other major producers are US, Indonesia, and Australia as shown in **Figure 1**. Other sources include Zambia, Peru, Canada, Russia, and Mexico. Recycle scrap accounts for ~50% of copper consumed in the US.

Copper has good strength and ductility as well as reasonably good corrosion resistance. The unique features of copper are its color and very high electrical and thermal conductivities.

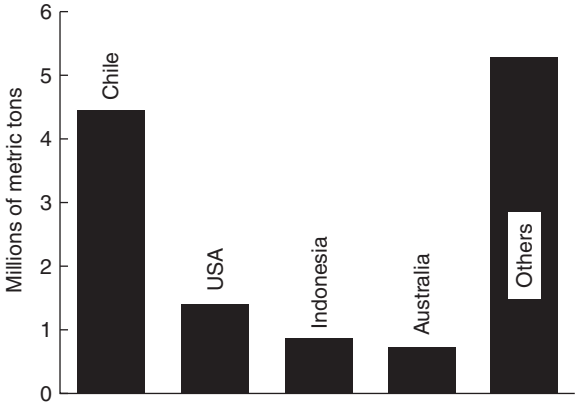


Figure 1 World production of copper in 2000. (Data from Chilean Copper Commission.)

Table 1 Physical properties of copper

Physical property	*Value*
Atomic number	29
Atomic mass	63.54 amu
Stable isotopes	^{63}Cu 69%, ^{65}Cu 31% All other isotopes have very short half-lives (<13 h)
Crystal structure	f.c.c.
Lattice parameter (25°C)	0.3615 nm
Density (20°C)	8930 kg m^{-3}
Electrical conductivity	103.06 IACS $= 59.77 \times 10^6$ ohm^{-1} m^{-1}
Melting point	1084.9 K
Heat of fusion	205 J kg^{-1}
Volume change on solidification	4.92%
Boiling point[a]	2967, 2995 K
Heat of vaporization	4730 J kg^{-1}
Linear thermal expansion coefficient (20°C)	16.7 μm mK^{-1}
Thermal conductivity (0°C)	403 W mK^{-1}
Specific heat (20°C)	386 J kg^{-1} K^{-1}
Young's modulus (GPa)	
polycrystal	128
⟨1 1 1⟩ direction	191.7
⟨1 0 0⟩ direction	99.9
Shear modulus	46.4 Gpa
Poisson's ration	0.34
Self-diffusion coefficient, D[b]	
830–1030°C	$47 \times 10^{-4} \exp(259\,000/RT)$ m^2 s^{-1}
750–950°C	$11 \times 10^{-4} \exp(240\,000/RT)$ m^2 s^{-1}
	where $R = 8.314$ J mol^{-1} K^{-1}
Energy of vacancies[c]	84 kJ mol^{-1}
Liquid/solid surface energy[d]	17.7 J m^{-2}
High-angle grain-boundary energy[e]	60 J m^{-2}
Solid/gas surface free energy[f]	170 J m^{-2}
Stacking fault energy[g,h]	70 J m^{-2}

All data are from (1979) *Metals Handbook,* 9th edn. vol. II. Metals Park, OH: ASM. except where noted.
[a](2004–2005) *Handbook of Chemistry and Physics*, 85th edn. Cleveland: CRP.
[b]Jost W (1952) *Diffusion in Solids, Liquids, Gases*. New York: Academic Press.
[c]Reed-Hill R (1973) *Physical Metallurgy Principles*. Princeton, NJ: Van Nostrand-Reinhold.
[d]Hollomon JH and Turnbull D (1953) *Progress in Metal Physics* 4.
[e]Gjostein NA and Rhines FA (1959) *Acta Mettallurgica* 7: 319.
[f]Udin H, Schaler AJ and Wulff J (1944) *Journal of Metals* 1: 1936.
[g]Fullman F (1951) *Journal of Applied Physics* 22: 448.
[h]Thoron PR and Hirsch PB (1958) *Philosophical Magazine* 3: 738.

The physical properties of copper are listed in **Table 1**.

Uses

Figure 2 shows the principal uses of copper. Its excellent electrical conductivity accounts for most of its use, principally as wire. Corrosion resistance, thermal conductivity, formability, and unique color account for almost the rest of the consumption of copper.

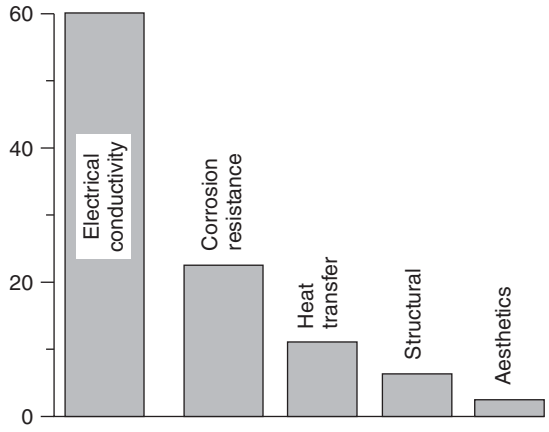


Figure 2 Uses of copper. (Data from ASM (2001) *Copper and Copper Alloys*, *ASM Speciality Handbook*. Materials Park, OH: ASM.)

Table 2 Electrical and thermal conductivities of several metals

Metal	*Electrical resistivity* (nΩ m)	*Electrical conductivity* (% IACS)	*Thermal conductivity* (W m^{-1} K^{-1})
Silver	14.7	108.4	428 (20°C)
Copper	16.73	103.06	398 (27°C)
Gold	23.5	73.4	317.9 (0°C)
Aluminum	26.55	64.94	247 (25°C)
Beryllium	40	43	190 (20°C)
Magnesium	44.5	38	418 (20°C)
Zinc	59.16	28.27	113 (25°C)
Nickel	68.44	25.2	82.9 (100°C)
Iron	98	17.59	80 (20°C)
Platinum	106	16.3	71.1 (0°C)

Data from *Metals Handbook*, 9th edn. vol II. Metals Park, OH: ASM.

Electrical Properties

The International Association of Classification Societies (IACS) standard of conductivity was set up with what was thought to be the conductivity of pure copper as 100%. Later it was found that copper had a higher conductivity than what was originally thought, but the standard was not changed. Pure copper has a conductivity of 104% according to the IACS standard.

Only silver has higher electrical and thermal conductivities than copper. **Table 2** lists the conductivities of various metals. The resistivity increases by 43% between 0°C and 100°C.

All impurities raise the electrical resistivity of copper. At low concentrations, the increase is

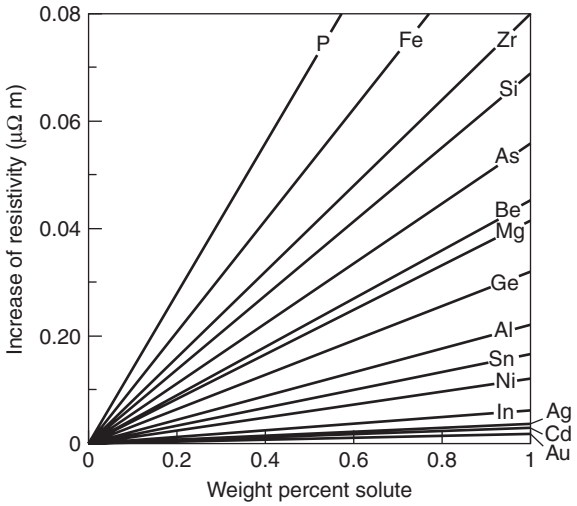


Figure 3 All impurities increase the resistivity of copper. The increase is proportional to the concentration. (Adapted from ASM (2001) *Copper and Copper Alloys. ASM Speciality Handbook*. Materials Park, OH: ASM.)

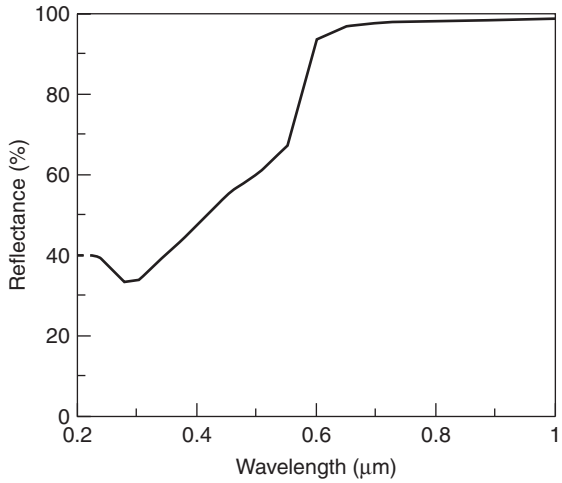


Figure 4 The color of copper is a result of the dependence of reflectivity on wavelength. (Data from ASM (2001) *Copper and Copper Alloys*, *ASM Speciality Handbook*. Materials Park, OH: ASM.)

proportional to the concentration of the impurity. In general, the effect of solutes in raising the resistivity is greater for large differences between the atomic diameter of the solute and that of copper as shown in **Figure 3**. Also solutes of high valency (P, Si, and As) have a greater effect on resistivity than those of low valency such as Ag, Au, and Cd.

Color

Probably the most distinctive property of copper is its color. The reddish hue is a result of how the reflectivity varies with wavelength. The dependence of reflectivity on the wavelength of a polished copper surface is shown in **Figure 4**.

Commercial Grades of Copper

There are several commercial grades of copper. The principal use of oxygen-free copper (C10100) with a minimum of 99.99% Cu is for wire. Fire-refined tough-pitch copper (C2500) is fire-refined copper containing between 0.02% and 0.05% oxygen in the form of Cu_2O and ~0.5% other elements. The copper–oxygen phase diagram is shown in **Figure 5**. There is a eutectic at 0.38% oxygen and 1066°C. **Figure 6** is a typical microstructure of cast tough-pitch copper showing primary copper dendrites surrounded by a copper–copper oxide eutectic.

Often tellurium (~0.5%) or sulfur (~0.5%) is added to copper to promote free machining. Copper is frequently deoxidized with phosphorus. Deoxidation leaves ~0.01% residual phosphorus in solid solution. This lowers the conductivity to below that of oxygen-free copper. Lead is often added to copper and copper alloys to form free machining characteristics. Lead is virtually insoluble in copper as shown by the copper–lead phase diagram (see **Figure 7**). Lead particles appear as a separate phase in the grain boundaries (see **Figure 8**).

Zirconium copper (C15000) contains 0.13–0.20% Zr. It can be heat treated to yield strengths of 400 MPa while retaining a conductivity of 84% IACS. Copper dispersion strengthened by 0.2–0.7% Al_2O_3 retains reasonable strength at temperatures up to 1000°C. It finds applications as electrodes for resistance welding.

Bismuth is a very detrimental impurity. It completely wets the grain boundaries and because it is brittle, its presence renders copper and copper-based alloys brittle. It must be kept $<0.003\%$.

Copper Alloys

Copper is too soft for structural applications. Its strength is markedly increased by alloying. For dilute substitutional solutions (see **Figure 9**), the yield strength increases in proportion to the solute concentration. The rate of this increase is proportional to the 4/3 power of a misfit parameter defined as $\varepsilon = (da/a)/dc$, where da/a is the fractional change in the lattice parameter with concentration c, expressed as an atomic fraction of solutes,

$$\frac{\Delta\tau}{\Delta c} = CG\varepsilon^{4/3} \qquad [1]$$

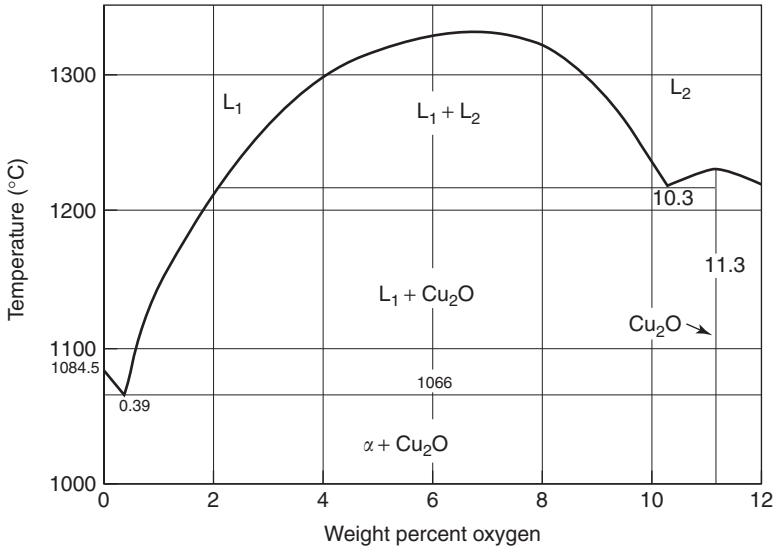


Figure 5 The copper–oxygen phase diagram. (Adapted from (1973) *Metals Handbook*, 8th edn., vol. VIII. Materials Park, OH: ASM.)

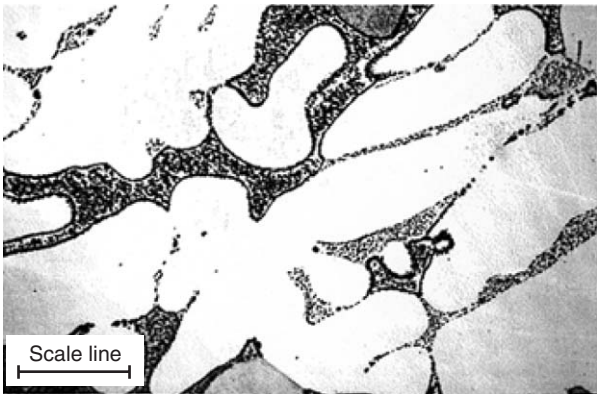


Figure 6 The microstructure of cast toughpitch copper. The scale line is ~125 μm. Note the primary copper dendrites and the Cu–Cu_2O eutectic. (Reproduced with permission of the Copper Development Association Inc.)

where G is the shear modulus of copper and C is a constant. The misfit, ε, is the fractional difference of atomic diameters of copper and solute. **Figure 10** shows the misfits for several solutes. On an atomic basis Sb, Sn, and In are the most potent hardeners.

Although different measurements of the stacking fault energy of pure copper have ranges from 40 to 169 mJ m^{-2}, the best estimate is probably 80 mJ m^{-2}. Zinc, tin, and aluminum form solid solutions that lower the stacking fault energy to <5 mJ m^{-2} at high concentrations as shown in **Figure 11**. The stacking fault energy depends on the ratio of valence electrons to atoms. A lower stacking fault energy leads to more strain hardening. The value of n in the power-law equation,

$$\sigma = K\varepsilon^n \qquad [2]$$

increases from ~0.4 for pure copper to ~0.6 or 0.65 at 35% Zn. The greater strain hardening raises the tensile strength as shown in **Figure 12**.

Finer grain size also increases strength. The grain-size dependence of the yield strength follows the Hall–Petch relation

$$\sigma_Y = \sigma_O + Kd^{-1/2} \qquad [3]$$

Deformation Mechanisms

The slip system in copper, as in all other f.c.c. metals, is $\{1\,1\,1\}\langle 1\,1\,0\rangle$. The critical shear stress for slip is ~3–4 MPa. The extent of work hardening increases as the temperature is lowered below room temperature as indicated in **Figure 13**.

Until the early 1950s, it was believed that mechanical twinning did not occur in copper. However, it is now known that copper can deform by mechanical twinning on $\{1\,1\,1\}$ planes, and in $\langle 11 \gg 2\rangle$ directions, at low temperatures. The critical shear stress for twinning at 4.2 K is ~150 MPa. In copper alloys, twinning depends on the stacking fault energy. The critical resolved-shear stress for twinning is lowered by solid solutions that reduce the stacking fault energy (see **Figure 14**). It is significant that copper is not embrittled at low temperatures and does not fracture by cleavage, because of which copper finds use in cryogenic equipment.

Commercial Alloys

Copper alloys are classified as being either wrought alloys or casting alloys, depending on their usage.

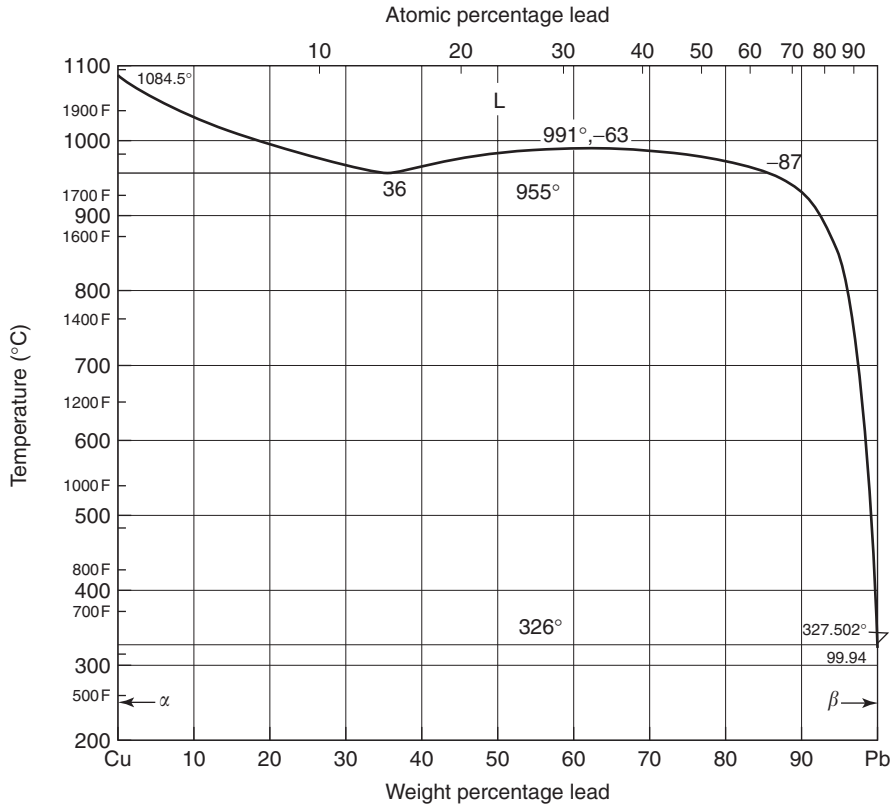


Figure 7 The copper–lead phase diagram. (Data from *Metals Handbook* 8th edn. vol. VIII. p. 296. Materials Park, OH: ASM.)

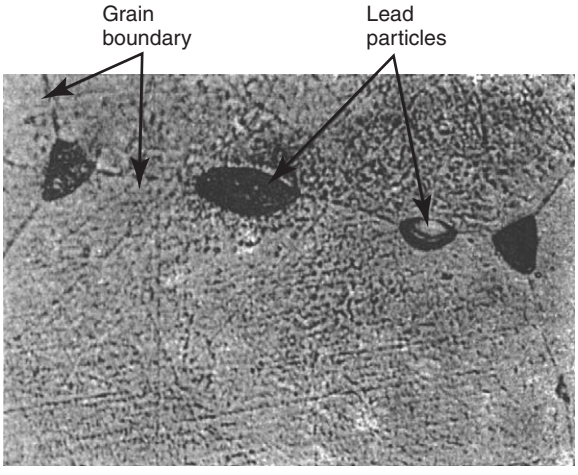


Figure 8 Lead appears as a separate phase in copper alloys. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*. Materials Park, OH: ASM.)

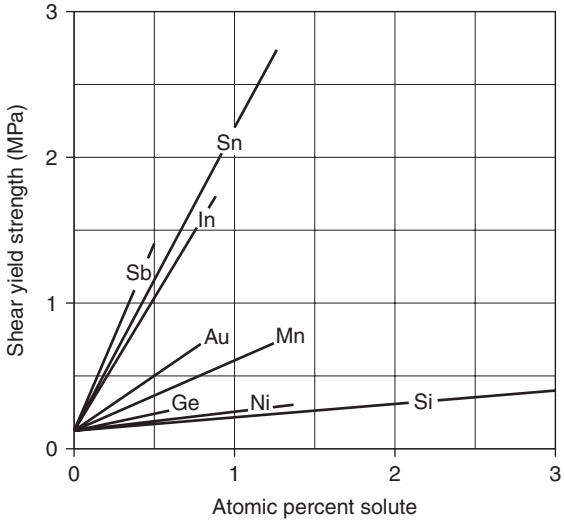


Figure 9 Effect of solute concentration on shear yield strength of dilute copperbase alloys. Note that the strength is proportional to concentration. (Data from Linde JO, Lindell BO, and Stade CH (1950) *Arkiv für Fysik* 2.)

The unified numbering system (UNS) designations for some of the wrought alloys are listed in **Table 3**.

Sheets produced by cold rolling and rods or wire produced by drawing are sold with a temper designation indicating the amount of cold reduction. The system is outlined in **Table 4**.

Brass

Brasses are alloys of copper and zinc. **Figure 15** is the copper–zinc phase diagram. The solubility of zinc in

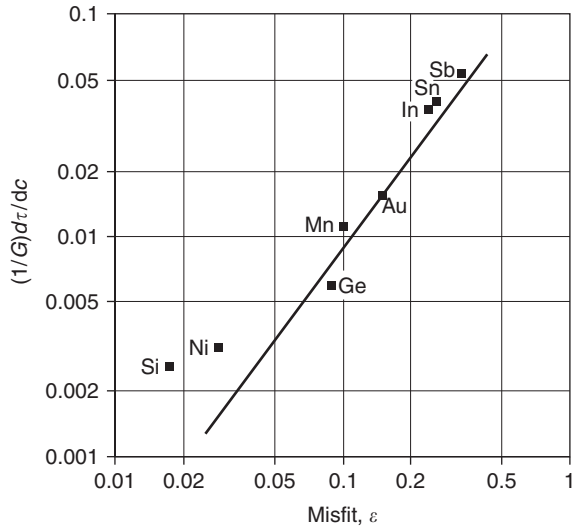


Figure 10 The effect of the atomic misfit on the strengthening effect of the solute. (Data from Linde JO and Edwardson S (1954) Arkiv für Fysik, 8.)

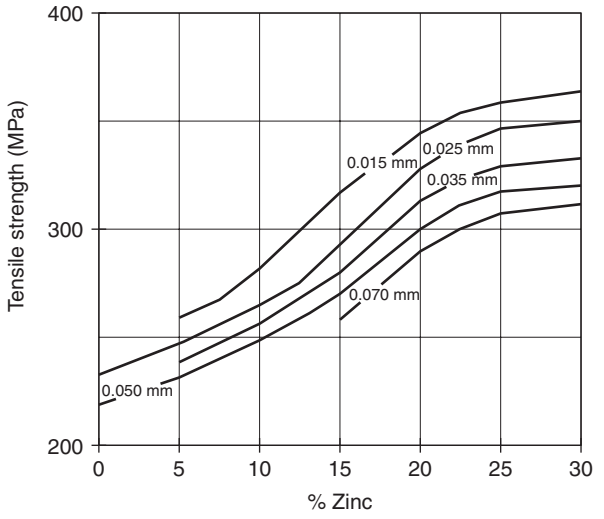


Figure 12 The effect of zinc content on increasing the tensile strength. Finer grain size has a similar effect. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Speciality Handbook*. Materials Park, OH: ASM.)

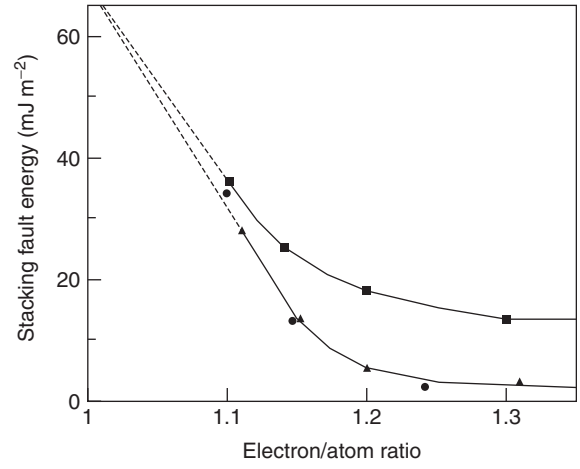


Figure 11 The stacking fault energy of alloys. (■ Cu–Zn, ▲Cu–Al, • Cu–Ge) (Adapted from Venables JA (1963) Deformation twinning in fcc metals. In: Reed-Hill RE (ed.) *Deformation Twinning. Proceedings of the Metallurgical Society Conference*, vol. 25, p. 77. Gordon and Breach.

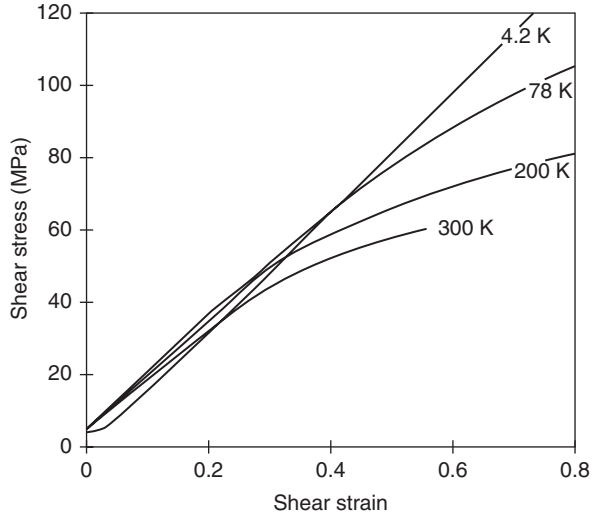


Figure 13 Tensile stress–strain curves for copper crystals at temperatures down to 4.2 K. (Adapted from Blewitt TH, Coltman RR, and Redman JK (1955) *Proceedings of the Conference on Defects in Crystalline Solids*, Bristol, p. 369. London: Physical Society.)

the f.c.c. lattice of copper is $>35\%$. At nearly equal atomic percentages of copper and zinc, there is a b.c.c. intermetallic phase β, that has a wide range of solubility. At temperatures $<\sim 460^{\circ}C$, β-brass undergoes an ordering reaction to form an ordered b.c.c. phase β' with a B2 (CsCl) structure. Each copper atom is surrounded by eight zinc and eight copper atoms. A eutectoid transformation, $\beta' \rightarrow \alpha + \gamma$ at 250°C has been reported. However, the transformation is rarely, if ever, encountered.

Zinc provides solid solution hardening to copper. Both the yield strength and the tensile strength increase with zinc content, as indicated in **Figures 16** and **17**.

As the lower stacking fault energy increases the exponent n, the uniform elongation also increases in the power-law approximation of the stress–strain curve. Another effect of the lower stacking fault energy is that annealing twins are much more frequent. **Figure 18** compares the annealed microstructure of copper with brass containing 30% Zn. The microstructures of brass containing 40% zinc or more consist of two phases, α and β. The two-phase microstructure of Muntz metal (40% Zn) is shown in **Figure 19**. At 800°C the microstructure is entirely β,

but as the alloy is cooled, α precipitates. The effects of other alloying elements to brass on the solid solubility of zinc can be approximated by "zinc equivalents." A concentration of 1% tin is equivalent to 2% zinc, so tin has a zinc equivalent of 2. **Table 5** lists the zinc equivalents of the common alloying elements. The relative values of the zinc equivalents of the elements can be understood in terms of the number of valence electrons per atom. Silicon and

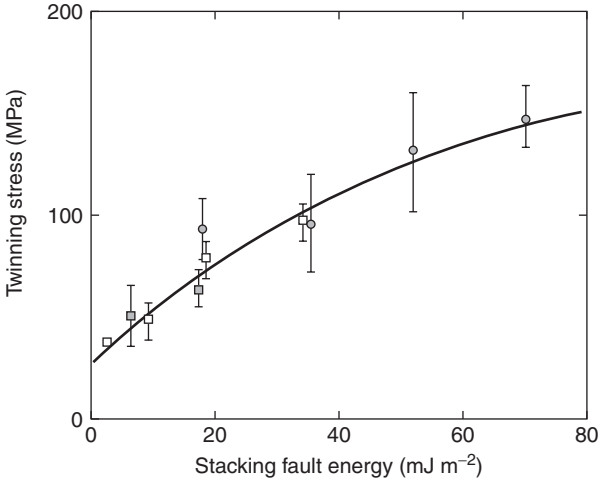


Figure 14 The critical shear stress for twinning decreases with decreasing stacking fault energy. (⌽, Cu–Zn; ⌻, Cu–Al; ⌻, Cu–Ge) (Adapted from Venables JA, In: Hirth and Rogers (eds.) *Deformation Twinning*, Reed-Hill: AIME.)

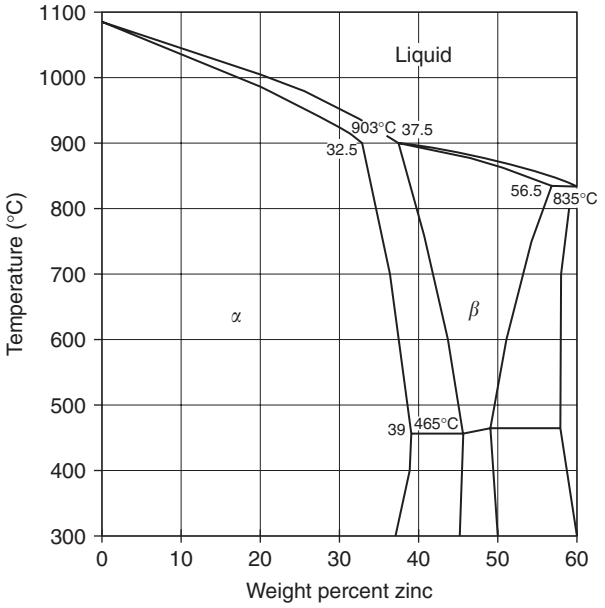


Figure 15 The copper–zinc phase diagram. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*. Materials Park, OH: ASM.)

Table 3 Commercial wrought-copper alloys

Designation	*Composition*	*Common name*	*Uses*
C21000	5% Zn	Gilding metal	Coinage, jewelry
C22000	10% Zn	Commercial bronze	Architectural, jewelry
C23000	15% Zn	Red brass	Architectural, plumbing
C24000	20% Zn	Low brass	Ornamental, musical
C26000	30% Zn	Cartridge brass	Lamps, cartridge cases
C27000	35% Zn	Yellow brass	Architectural, lamps
C28000	40% Zn	Muntz metal	Architectural, heat exchangers
C36000	35.5% Zn, 2.5% Pb	Free cutting brass	Screw machine parts
C44300	28% Zn, 1% Sn	Admiralty brass	Heat exchangers
C51000	5% Sn, 0.2% P	Phosphor bronze	Hardware
C71500	30% Ni	Cupro nickel	Heat exchangers
C75200	18% Ni, 17% Zn	Nickel silver	Hardware, jewelry
C172000	1.8–2.0% Be, 0.2–0.8% Ni + CO	Beryllium copper	Springs, instruments

Table 4 Temper designations

Temper designation		*B&S Gage numbers*	*Rolled sheet Thickness reduction* (%)	*Strain*	*Drawn wire Diameter reduction* (%)	*Strain*
H01	1/4 hard	1	10.9	0.116	10.9	0.232
H02	1/2 hard	2	20.7	0.232	20.7	0.463
H03	3/4 hard	3	29.4	0.347	29.4	0.694
H04	hard	4	37.1	0.463	37.1	0.926
H06	extra hard	6	50.1	0.696	50.1	1.39
H08	spring	8	60.5	0.925	60.5	1.86
H10	extra spring	10	68.6	1.16	68.6	2.32

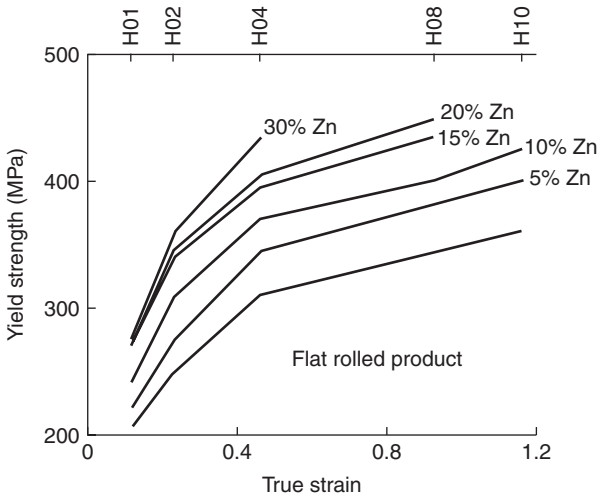


Figure 16 The yield strength of brasses increases with the amount of cold work and the zinc content. The true strain, $\varepsilon = \ln(t_o/t)$ where t is the thickness and t_o is the thickness before cold rolling. (Adapted from (1979) *Metals Handbook*, 9th edn., vol. II. Materials Park, OH: ASM.)

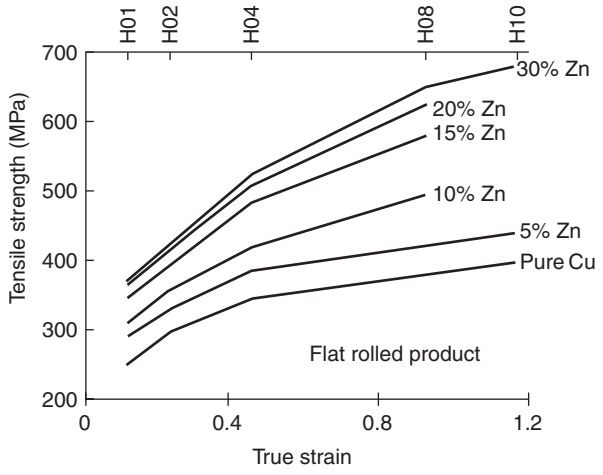


Figure 17 The tensile strength of brasses increases with the amount of cold work and the zinc content. The true strain, $\varepsilon = \ln(t_o/t)$ where t is the thickness and t_o is the thickness before cold rolling. (Data from (1979) *Metals Handbook*, 9th edn., vol. II. Materials Park, OH: ASM.)

aluminum have high valencies (4 for silicon and 3 for aluminum) and low atomic masses (28 for silicon and 27 for aluminum), so 1 wt. % causes a large increase of the electron-to-atom ratio. Magnesium, like zinc has a valency of 2, but the atomic mass of magnesium is less than half that of zinc (24 for magnesium vs. 65 for zinc), so 1 wt.%Mg has over twice the effect of 1 wt.%Zn. Tin has a valence of 4, but its atomic mass (119) is roughly twice that of zinc.

Other Alloys

True bronzes are copper–tin alloys. Tin is much less soluble in copper than zinc as shown in the copper–tin phase diagram (see **Figure 20**). The ε phase is very sluggish to form and is rarely encountered. Tin is a potent solid solution strengthener. However, alloys with $>8\%$ Sn are too brittle to be formed mechanically. Because bronzes have a much greater difference of the liquidus and solidus temperatures than brass, they are much easier to cast. Casting alloys often contain up to $\sim 20\%$ Sn. These are used extensively for pipe fittings and bells.

Alloys of copper with aluminum and silicon are called aluminum-bronze and silicon-bronze, even though they contain no tin. Manganese-bronze is a brass with only 1% Sn and 0.1% Mn. Aluminum-bronzes containing up to 7% Al behave much like brass. The phase diagram (see **Figure 21**) shows that a b.c.c. β-phase occurs at $\sim 12\%$ Al.

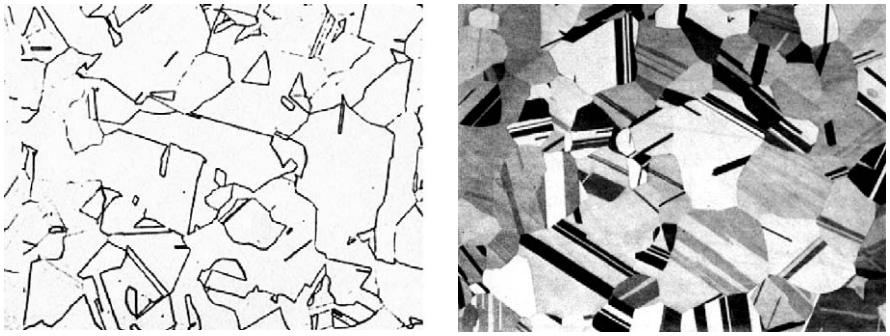
There is a eutectoid transformation of $\beta \rightarrow \alpha + \gamma_2$ at 565°C and 11.8% Al. **Figure 22** shows the pearlite-like microstructure of alternating platelets of α and γ_2 in an alloy of Cu–11.8% Al after slowly cooling from 800°C. With rapid cooling, the eutectoid transformation is suppressed and the alloy transforms by martensitic shear to a new hexagonal phase β'. **Figure 23** shows the martensitic β' needles in the same alloy after rapid cooling. The martensitic reaction is almost instantaneous. It starts when β is quenched below the M_s temperature and is virtually complete when the temperature is at the M_f. The M_s and M_f temperatures depend on the composition as shown in **Figure 24**. The oxidation resistance of aluminum bronzes is somewhat superior to brass.

Silicon bronzes are more easily cast than brass and have better resistance to oxidation and acid corrosion.

Copper and nickel are completely miscible in both the liquid and solid state as shown in the copper–nickel phase diagram (see **Figure 25**). The low-temperature miscibility gap is rarely encountered.

Nickel imparts improved corrosion and oxidation resistance to copper. Cupronickels containing 20–30% Ni are widely used for heat exchangers.

Alloying copper with 20% Ni causes a loss of the copper's familiar yellow–reddish color. Coinage accounts for $\sim 1\%$ of the consumption of copper. The five-cent pieces ("nickels") in the US and Canada are alloys of 75% copper and 25% nickel. The 10-cent ("dime") and 25-cent ("quarter") pieces are sandwiches of pure copper core clad with 75% copper–25% nickel alloy as shown in **Figure 26**. The Susan B Anthony US dollar is a sandwich of pure copper clad with 77Cu–12 Zn–7 Mn–4 Ni. European coins are



(a) (b)

Figure 18 Microstructures of (a) pure copper and (b) 70/30 brass (200×). Note that the frequency of annealing twins is higher in the brass. ((a) Adapted from (1972) *Metals Handbook*, 8th edn., vol. VII. Materials Park, OH: ASM. (b) Adapted from Hosford WF (1993) *The Mechanics of Crystals and Textured Polycrystals*. London: Oxford Science Pub. Courtesy A Graf).

Figure 19 Photomicrograph of a Muntz metal (40% Zn) after air cooling. The dark phase is β and the light phase is α, which has precipitated during cooling. (Adapted from Brick RM, Gordon RB, and Phillips A (1965) *Structure and Properties of Alloys*. New York: McGraw-Hill.)

Table 5 Zinc equivalents of several alloying elements in brass

Element	*Zn equivalent*	*Element*	*Zn equivalent*
Si	10	Al	6
Mg	–2	Sn	2
Fe	0.9	Mn	0.5
Co	–0.8	Ni	–1.3

either copper-based alloys or combinations of two copper-based alloys. Constantan (45% Ni) has a nearly zero-temperature coefficient of resistivity, hence is useful in instruments.

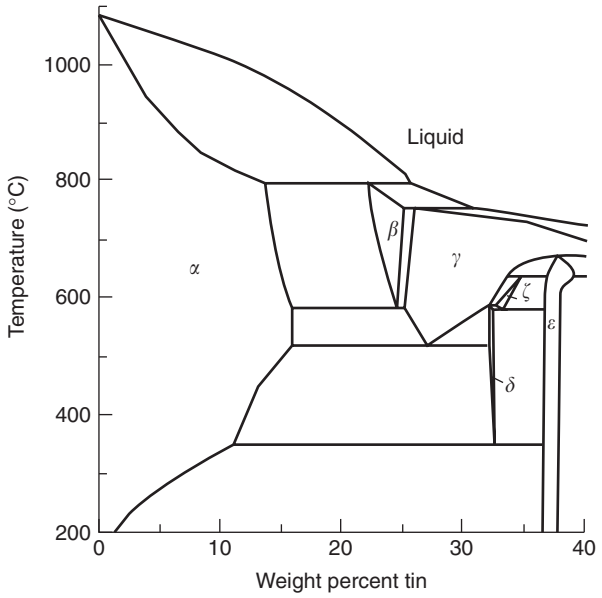


Figure 20 The copper–tin phase diagram. β is a b.c.c. phase. In alloys containing ~20% Sn, formation of the ε-phase is so sluggish that it is seldom encountered. (Adapted from (1973) *Metals Handbook*, 8th edn., vol. VIII. Materials Park, OH: ASM.)

Nickel Silvers

Nickel silvers are ternary alloys containing 12–18% Ni and 17–27% Zn. They have good strength and good corrosion resistance. They are used for springs and corrosion-resistant hardware. Costume jewelry is often made from nickel silver because of its pleasant silvery color.

Beryllium Copper

Copper alloys containing 1.8–2.0% Be and 0.2–0.8% Ni + Co are precipitation-hardenable. The alloy C172000 is solution treated at ~800°C and rapidly cooled. The precipitation is done between

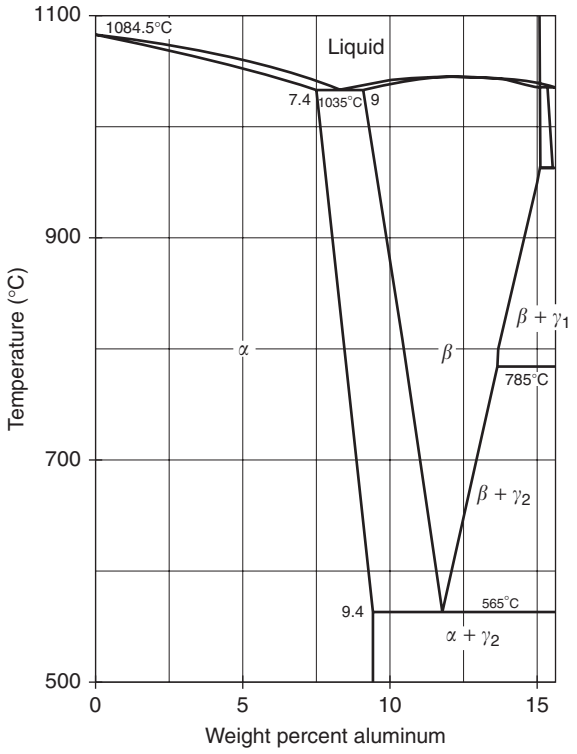


Figure 21 The copper–aluminum phase diagram. β is a b.c.c. phase. Note that eutectoid transformation at 565°C. (Adapted from (1973) *Metals Handbook*, 8th edn., vol. VIII. Materials Park, OH: ASM.)

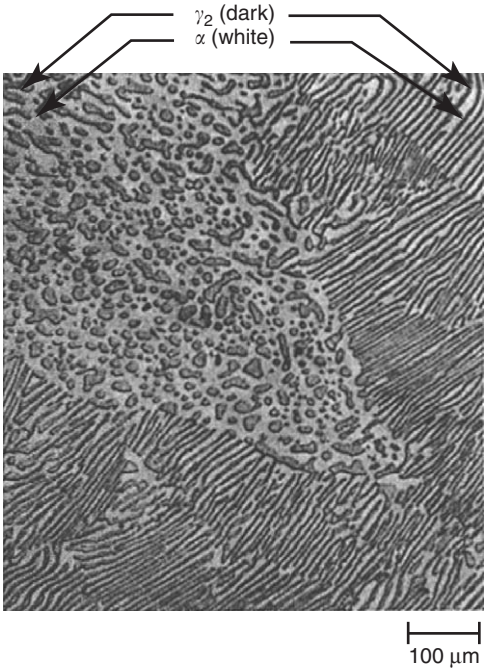


Figure 22 Pearlite-like structure in Cu–11.8% Al after slow cooling from 800°C. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*, p. 48. Materials Park, OH: ASM.)

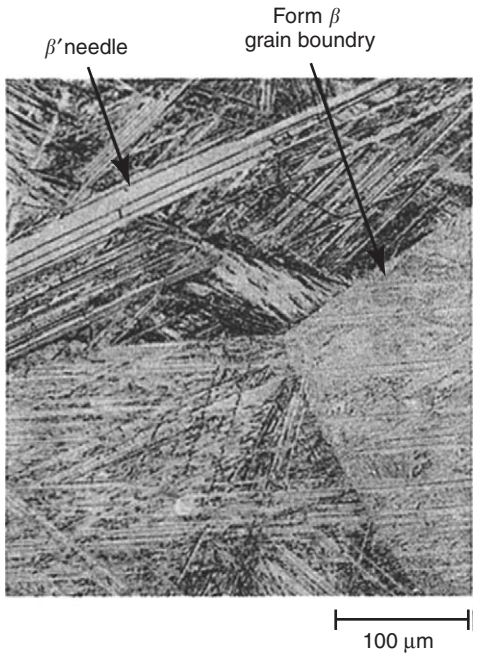


Figure 23 Martensitic structure of Cu–11.8% Al after rapid cooling from 800°C. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*, p. 48. Materials Park, OH: ASM.)

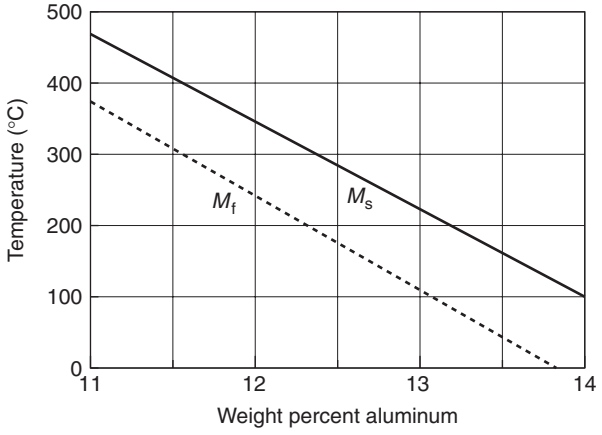


Figure 24 The M_s and M_f temperatures of aluminum bronzes. (Adapted from Flinn RA (1961) *The Non-Ferrous Copper, Brass, and Bronze Castings*, p. 20. Cleveland, OH: Founders Society.)

300°C and 400°C. **Figure 27** shows the copper-rich end of the Cu–Be phase diagram. Precipitation hardening produces very high strengths. Even higher strengths can be achieved by cold working after the solution treatment and before precipitation as shown in **Table 6**. The principal drawback is that the toxicity of beryllium is an extreme danger in preparing this alloy.

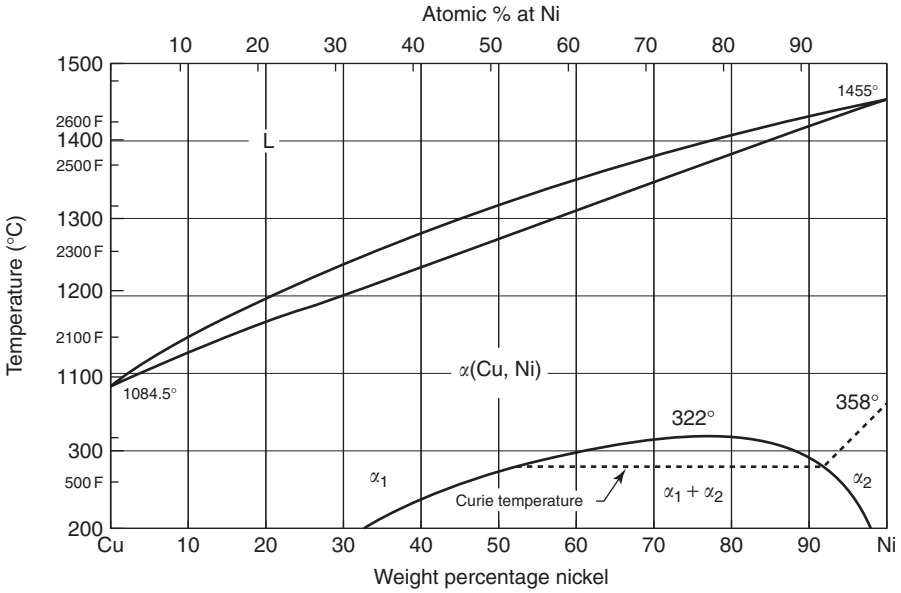


Figure 25 The copper–nickel phase diagram shows complete miscibility. (Adapted from (1973) *Metals Handbook*, 8th edn., vol. VIII. Metals Park, OH: ASM.)

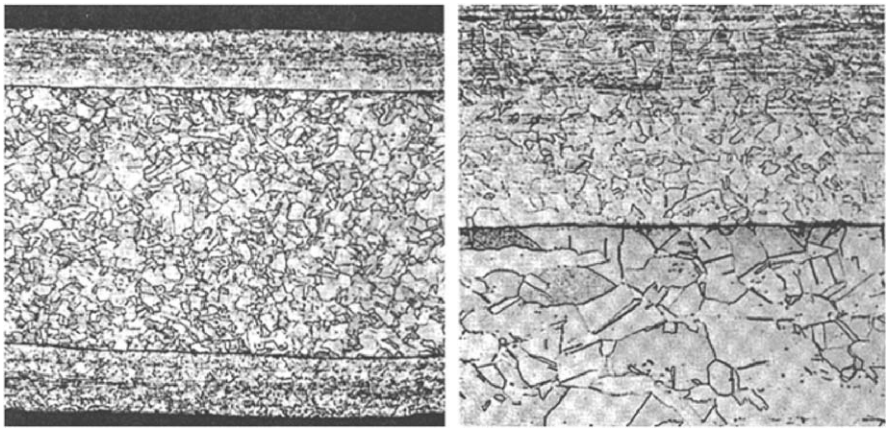


Figure 26 Cross section of a US dime. (Adapted from *Copper and Copper Alloys*, *ASM Speciality Handbook*. Materials Park, OH: ASM.)

Shape Memory Alloys

There are two groups of shape memory alloys: one group consists of Cu–Zn–Al alloys with 10–30% Zn and 5–10% Al. Cu–Al–Ni alloys with 11–14.5% Al and 3–5% Ni form the other group. The recoverable strain for both groups is 4%. The martensitic transformation temperature is $<120^{\circ}$C for the Cu–Zn–Al alloys and $<200^{\circ}$C for the Cu–Al–Ni alloys. The exact temperatures depend on composition. The temperature hysteresis is in the range of 15–25°C.

Deformation Textures

The wire texture of copper consists of 35% $\langle 1\,0\,0 \rangle$ and 65% $\langle 1\,1\,1 \rangle$. The ratio of $\langle 1\,0\,0 \rangle$ to $\langle 1\,1\,1 \rangle$ increases with alloying. The compression texture of copper, like all f.c.c. metals, has $\langle 1\,1\,0 \rangle$ directions parallel to the axis of compression.

Sheet textures are best described by orientation distribution functions (ODF). The rolling texture of copper may consist of two continuous fibers or tubes in the ODF. The major fiber runs between $\{1\,1\,0\}\langle 1\,1\,2 \rangle$ (B) at $\Phi = 45^{\circ}$, $\phi_2 = 90^{\circ}$, $\phi_1 = 35^{\circ}$ through $\{1\,2\,3\}\langle 6\,3\,4 \rangle$ (S) at $\Phi = 37^{\circ}$, $\phi_2 = 63^{\circ}$, $\phi_1 = 59^{\circ}$ to $\{1\,1\,2\}\langle 1\,1\,1 \rangle$ (C) at $\Phi = 35^{\circ}$, $\phi_2 = 45^{\circ}$, $\phi_1 = 90^{\circ}$. A second minor fiber runs between $\{1\,1\,0\}\langle 0\,0\,1 \rangle$ (G) at $\Phi = 45^{\circ}$, $\phi_2 = 90^{\circ}$, $\phi_1 = 0^{\circ}$ and $\{1\,1\,0\}\langle 1\,1\,2 \rangle$ (B) at $\Phi = 45^{\circ}$, $\phi_2 = 90^{\circ}$, $\phi_1 = 35^{\circ}$. The corresponding ODF is shown in **Figure 28** and represented schematically in **Figure 29**. This texture has been characterized as consisting of

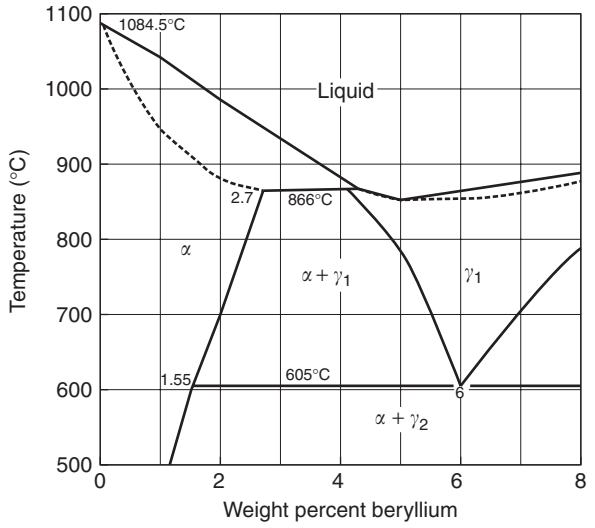


Figure 27 The copper–beryllium phase diagram. (Adapted from (1973) *Metals Handbook*, 8th edn., vol. VIII. Materials Park, OH: ASM.)

Table 6 Strength of beryllium copper sheet (C17200)[b]

Condition[a]	*Yield strength* (MPa)	*Tensile strength* (MPa)
Annealed	250	465
1/4 H	485	570
1/2 H	555	605
H	690	730
Annealed & HT	1060	1260
1/4 H & HT	1125	1290
1/2 H & HT	1180	1325
H & HT	1195	1360

HT indicates age-hardening treatment. (Solution treatments were done prior to work hardening.)

[a] 1/4 H, 1/2 H, and H indicate increasing strengths by work hardening.

[b] Data from ASM (2001) *Copper and Copper Alloys*, *ASM Speciality Handbook*, pp. 256. Materials Park, OH: ASM.

27% C, 38% S, 8% B, 18% between S and B, 3% G, and 6% other.

The rolling texture of low stacking fault energy alloys is different. The orientation $\{1\,1\,0\}\langle 1\,1\,2\rangle$ (B) at $\Phi = 45°$, $\phi_2 = 90°$, $\phi_1 = 35°$ is much stronger and the fiber G–B is more prominent. There are two additional fibers. Fiber γ corresponds to orientations having $\{1\,1\,1\}$ aligned parallel to the rolling plane. It extends from $\{1\,1\,1\}\langle 1\,1\,2\rangle$ at $\Phi = 55°$, $\phi_2 = 45°$, $\phi_1 = 0°$ to $\Phi = 55°$, $\phi_2 = 45°$, $\phi_1 = 90°$. Fiber τ consists of orientations with $\langle 1\,1\,0\rangle$ parallel to the transverse direction. For alloys with a stacking fault energy of $\sim 40\,\mathrm{mJ\,m^{-2}}$, there is an additional fiber (τ). **Figure 30** shows the ODF of cold-rolled brass and **Figure 31** is a schematic representation of the

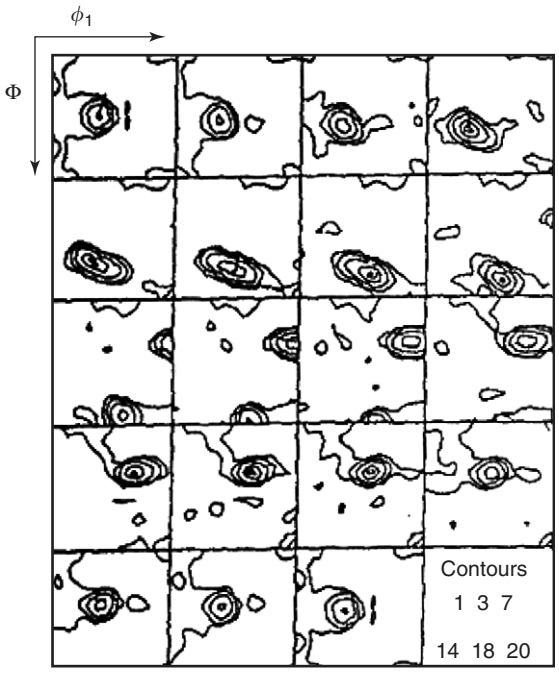


Figure 28 ODF for cold rolled copper. (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

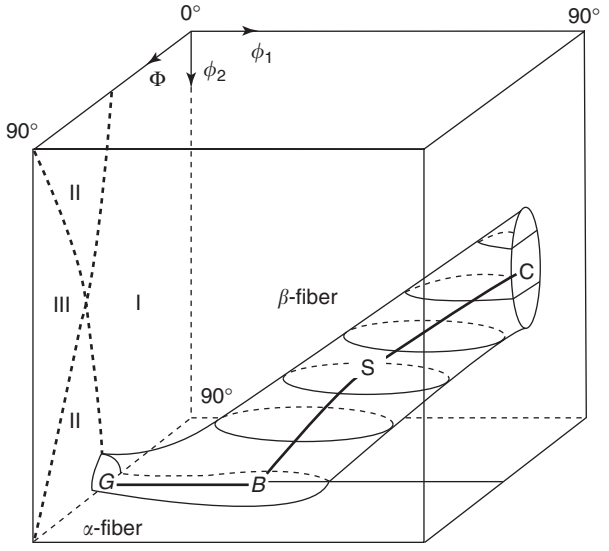


Figure 29 Schematic representation of the ODF shown in **Figure 28**. (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena.* London: Elsevier.)

fibers. As the amount of deformation increases the α-fiber remains strong, the β-fiber increases, and the C and S components decrease.

Textures are often represented by pole figures. **Figures 32a** and **32b** are the pole figures of copper rolled at room temperature.

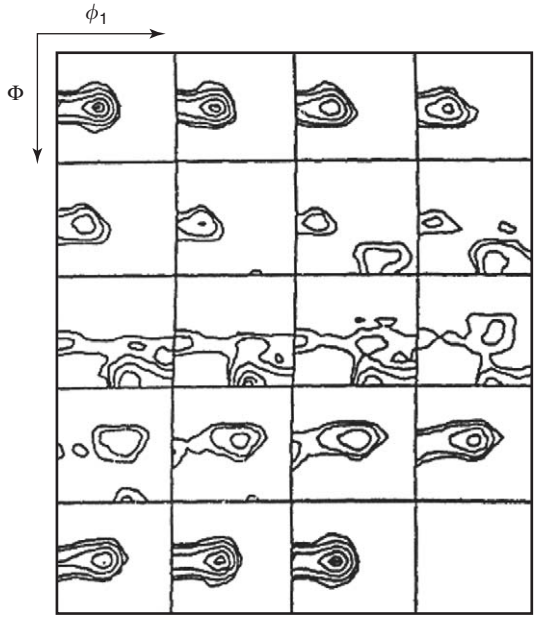


Figure 30 ODF for cold-rolled brass. (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

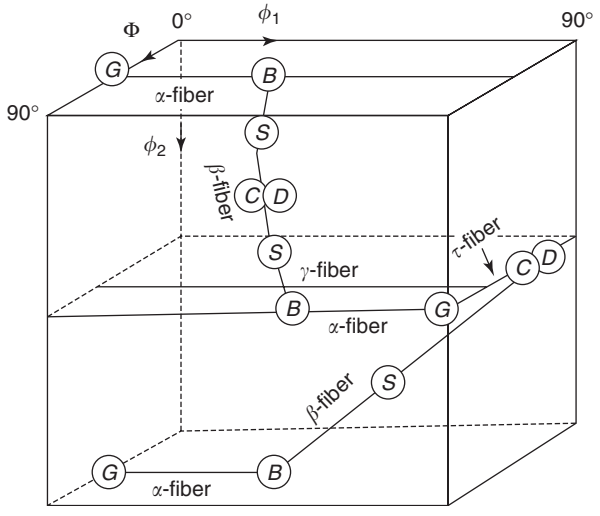


Figure 31 Schematic representation of the ODF in **Figure 30**. (Adapted form Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

This texture has been described as consisting of ideal components of ~39% {1 2 3}⟨41≫2⟩, ~34% {1 4 6}⟨21≫1⟩, ~14% {1 1 0}⟨1≫12⟩ and ~13% {1 1 2}⟨11≫1⟩. Pole figures for 70:30 brass are shown in **Figures 33a** and **33b.**

The relative amounts of these components change as the temperature of rolling decreases or as the copper is alloyed with zinc or some other element that lowers its stacking fault energy. The transition from copper-type recrystallization texture to the brass-type texture depends on the stacking fault energy of the alloy.

The rolling texture also depends on the rolling temperature. The pole figure of copper rolled at −196°C (see **Figure 34**) is very similar to brass rolled at room temperature. **Figure 35** shows that the texture transition with temperature is gradual.

Recovery

The effects of annealing cold-worked metal are usually divided into three stages: recovery, recrystallization, and grain growth. During recovery, the excess vacancies anneal out and dislocations are rearranged into lower-energy configurations. There are no major changes in hardness or strength. **Figure 36** shows the changes of stored energy and conductivity during recovery and recrystallization of copper.

Recrystallization

During recrystallization after cold work, new strain-free grains are formed. The extent of recrystallization depends on both temperature and time. It is customary to define the recrystallization temperature as the temperature at which 50% of the grains will be recrystallized in a half-hour. For most metals of commercial purity, the recrystallization temperature is between 1/3 and 1/2 of the melting point on an absolute scale. However, the recrystallization temperature depends on purity, the extent of cold work before annealing, and the prior grain size. **Figure 37** shows the kinetics of recrystallization of copper that is 99.999% pure after cold reduction of 98%.

The fact that the recrystallization temperature of 112°C in this case is only 28% of the melting temperature is explained by the extreme purity and very heavy reduction. The time, t, to achieve 50% recrystallization is given by the Arrhenius equation,

$$t = t_0 \exp\left(\frac{Q}{RT}\right) \qquad [4]$$

where Q is the activation energy, R is the gas constant, and T is the temperature. **Figure 38** is an Arrhenius plot of the data in **Figure 37**. The activation energy is 93 kJ mol^{-1}.

According to the Johnson–Mehl–Avrami equation, the fraction, f, of the microstructure recrystallized at any temperature should be given by

$$f = 1 - \exp(-Bt^n) \qquad [5]$$

where the exponent n is the sum of the contributions from nucleation n_n and growth n_g. If the nucleation rate is constant, $n_n = 1$ and if the growth rate is

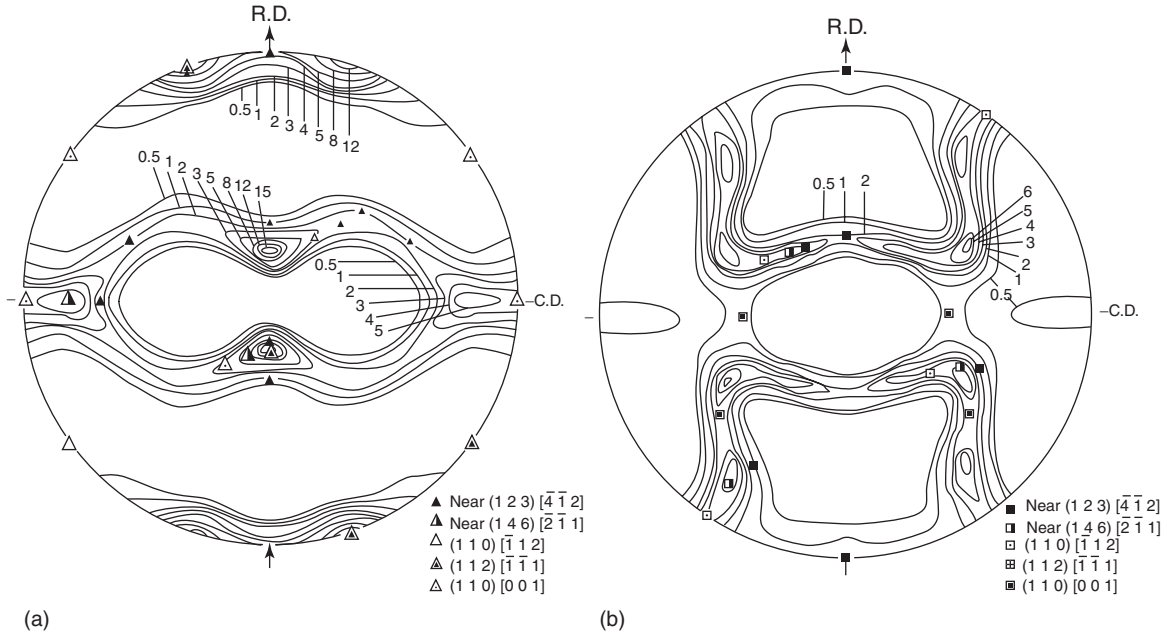


Figure 32 The (1 1 1) and (2 0 0) pole figures of electrolytic copper rolled 96.6% (Adapted from Hu H and Goodman SR (1963) Texture transition in copper. *Trans. Met. Soc. AIME* 227: 627–639.)

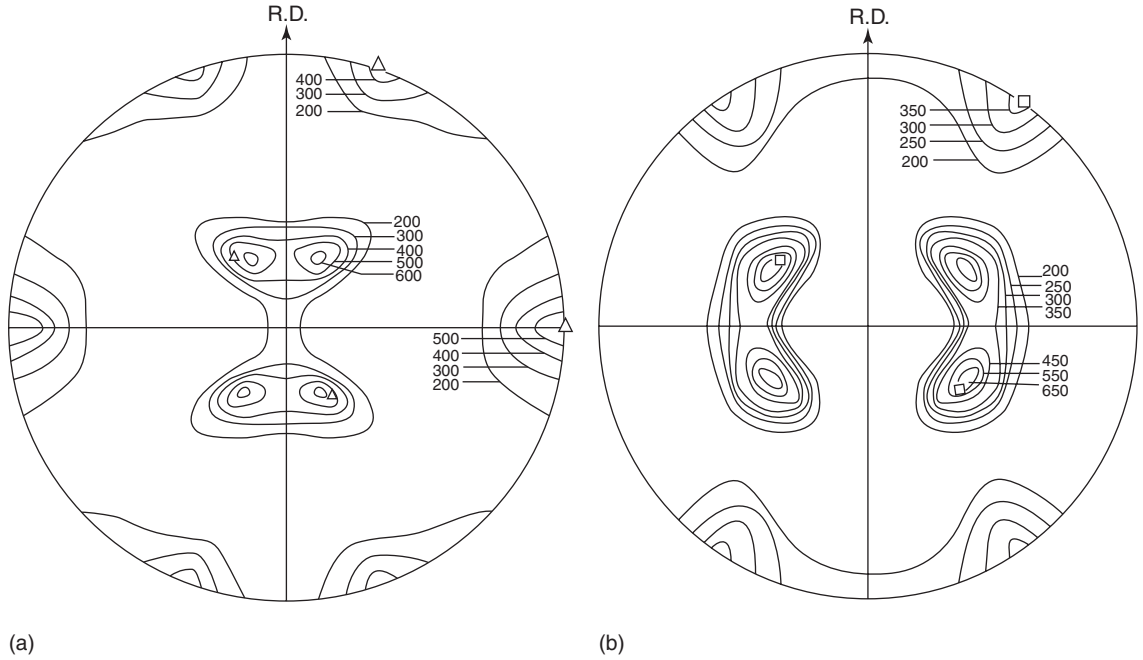


Figure 33 The (1 1 1) and (2 0 0) pole figures 70:30 brass cold rolled 95%. (From Hu H, Sperry PR, and Beck PA (1952) Rolling textures in face centered cubic metals. *Journal of Metals* (4)1: 76–81.)

constant $n_g = 3$. **Figure 39** is a log–log plot of $\ln(1 - f)$ versus t of the data for **Figure 34** at 135°C. The slope is close to 4, indicating constant nucleation and growth rates.

The effect of the amount of cold work on the recrystallization of brass is illustrated in **Figure 40**.

"Commercially pure" copper recrystallizes at 180°C while zone-refined copper recrystallizes at 80°C. Alloying elements raise the recrystallization temperature as shown in **Figure 41**. The recrystallization temperature also depends on prior grain size as shown in **Figure 42**. With greater amounts of prior

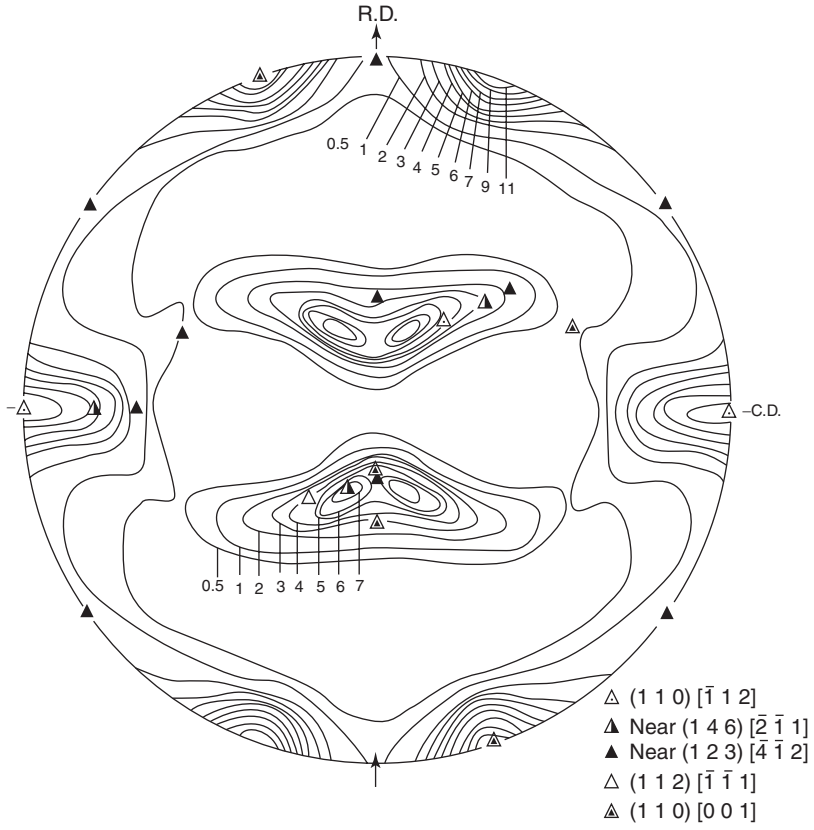


Figure 34 (1 1 1) Pole figure of electrolytic copper, rolled 96.6% at − 196°C. (Adapted from Hu H and Goodman SR (1963) Texture transition in copper. *Trans. Met. Soc. AIME* 227: 627–639.)

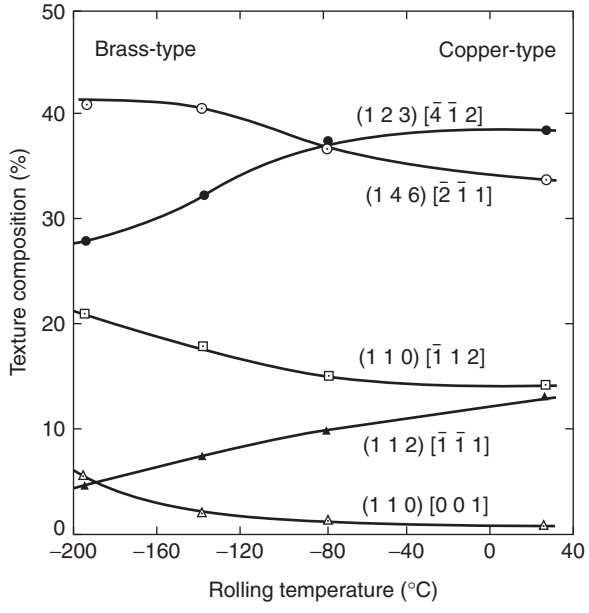


Figure 35 Change texture of cold-rolled copper with temperature. (Adapted from Hu H and Goodman SR (1963) *Trans. AIME* 227.)

cold work, the recrystallization temperature is lowered. (see **Figure 43**).

The grain size resulting from recrystallization depends on the relative rates of nucleation and growth of recrystallized grains. Increasing amounts of cold work before recrystallization decreases the resulting grain size because it increases the nucleation rate more than the growth rate. Solutes in solid solution decrease the recrystallized grain size because they decrease the growth rate by slowing grain-boundary motion. Second-phase particles have the same effect but they may increase the nucleation rate. Finally, the grain size before recrystallization affects the recrystallized grain size. With a finer prior grain size, there are more sites for nucleation, so the resulting grain size is also smaller.

Grain Growth

With continued heating, the larger recrystallized grains grow at the expense of other smaller recrystallized grains, with the result that the grain size

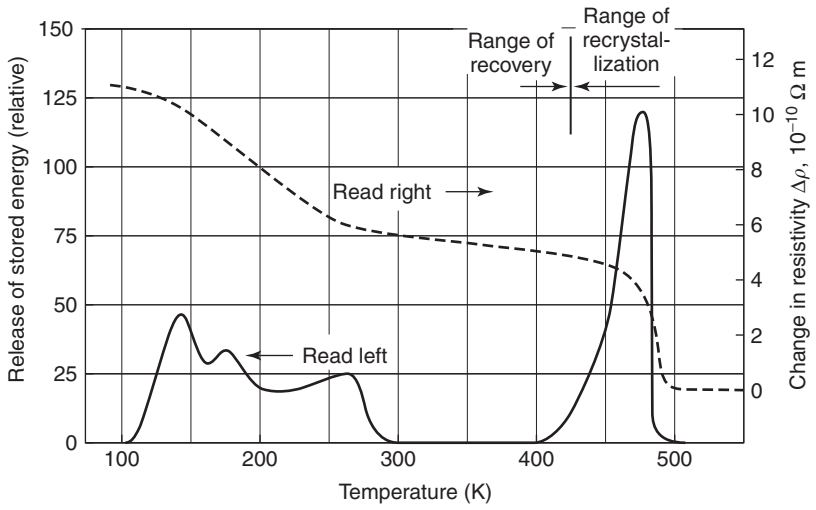


Figure 36 Energy release and changes of conductivity on annealing of cold worked copper. The changes $<300^{\circ}$C are caused by annealing out of vacancies. (Adapted from Guy AC and Hren JJ (1974) *Elements of Physical Metallurgy*. Reading, MA: Addison-Wesley.)

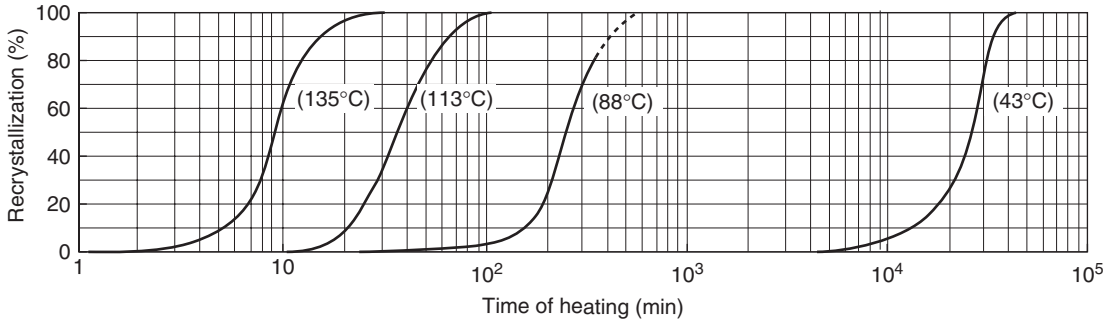


Figure 37 Isothermal recrystallization of 99.999% pure copper cold-worked 98%. (Adapted from Guy AG and Hren JJ (1974) *Elements of Physical Metallurgy*. Reading, MA: Addison-Wesley.)

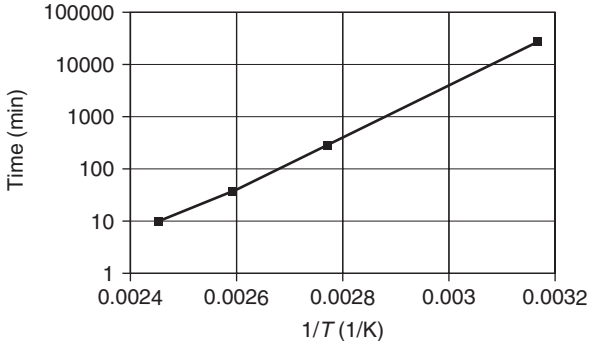


Figure 38 An Arrhenius plot of the time–temperature relation for 50% recrystallization of the copper in **Figure 37**.

increases. The increase of grain size is often characterized by

$$d^{1/n} - d_0^{1/n} = Ct \qquad [6]$$

Simple theory gives the exponent $1/n$ as 1/2, but it is usually much lower. Solutes and second-phase particles have a large effect.

Annealing Textures

ODFs for copper and copper with increasing amounts of zinc are given in **Figure 44**.

The transition from copper-type recrystallization texture to the brass-type texture depends on the stacking fault energy of the alloy. This correlation is shown in **Figure 45**. A marked change occurs at 8% Zn, 3%Al, 2% Ge, or a little over 1% P. The strength of the cube texture is reduced by light rolling reductions, large grain sizes, and low recrystallization temperatures.

Figure 46 is the {1 1 1} pole figure for copper cold rolled 97% at room temperature and recrystallized at 200°C. It consists almost entirely of cube texture and

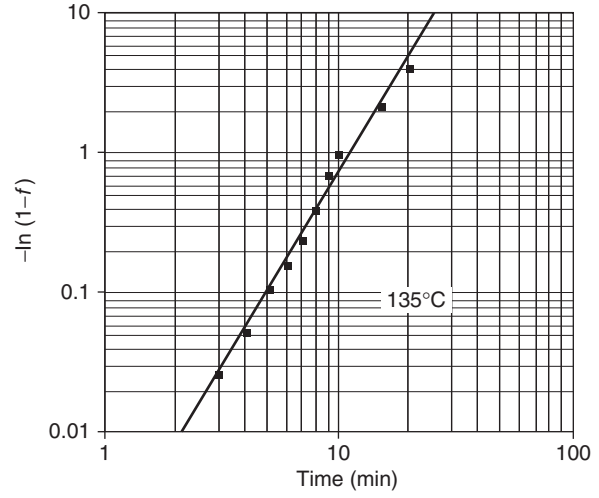


Figure 39 A Johnson–Mehl–Avrami plot of the recrystallization of the copper in **Figure 34** at 135°C. The slope is nearly 4, which indicates that the nucleation rate is constant and there is a constant rate of growth in three dimensions.

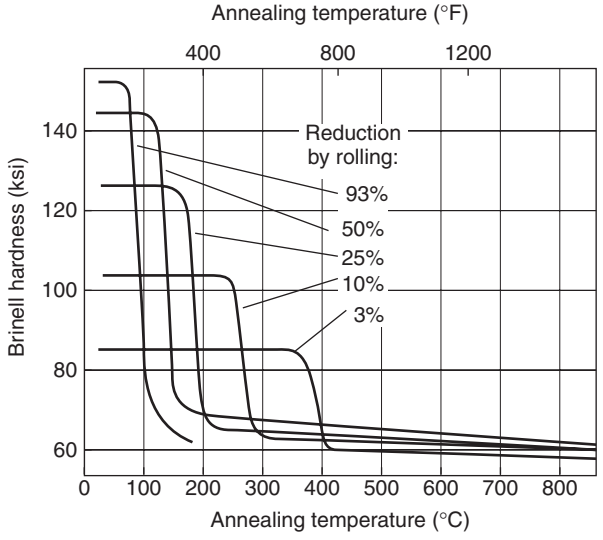


Figure 40 The recrystallization temperature of copper depends on the amount of cold work. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Speciality Handbook*. Materials Park, OH: ASM.)

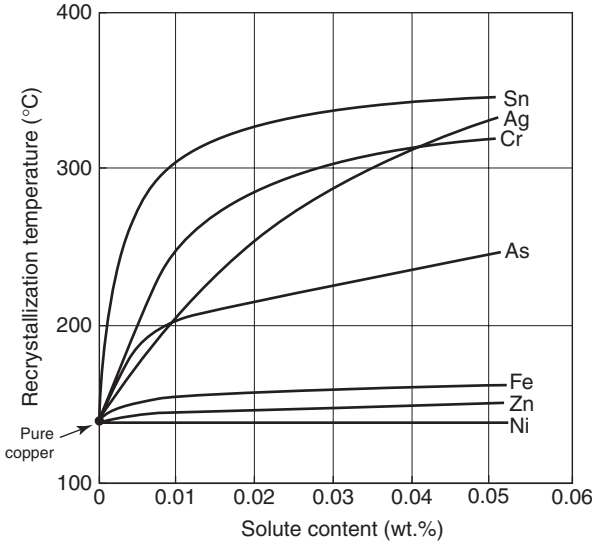


Figure 41 The effect of various solutes on the recrystallization temperature of copper. (Adapted from ASM (2001) *Copper and Copper Alloys*, ASM Speciality Handbook. Materials Park, OH: ASM.)

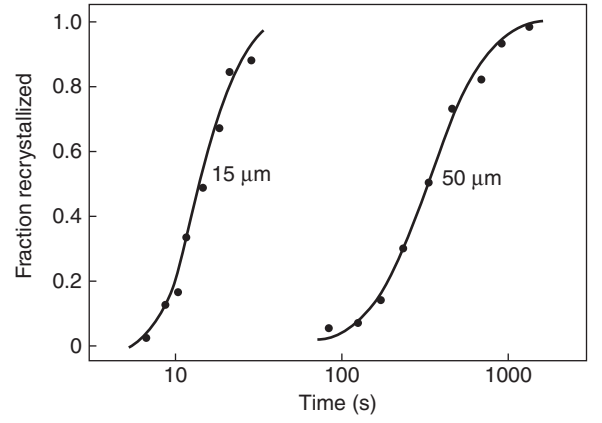


Figure 42 Dependance of recrystallization kinetics on prior grain size. Copper cold-rolled 93% and recrystallized at 225°C. (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

its twins. Other textural components of copper, rolled at lower temperatures and recrystallized at 400°C, are indicated in **Table 7**. The recrystallization texture of copper rolled at −196°C is very similar to the recrystallization texture of cold-rolled brass (see **Figure 47**).

Table 7 Texture components of copper, rolled, and recrystallized at 400°C

Component	*Percentage of texture*		
	−80°C	−140°C	−196°C
(2 1 4)[≫5 ≫2 3]	0	21.4	57.4
(3 1 0)[0 0 1]	0	8.1	21.3
(1 0 0)[0 0 1]	94.5	60.4	12.8
(1 2 2)[≫2 ≫1 2]	5.7	10.1	8.5

Adapted from Hu H and Goodman SR (1952) *Trans. Met. Soc. AIME* 277: 637.

Anisotropy

The anisotropy of plastic behavior of sheet material is conventionally documented by the Lankford parameter or *R*-value which is defined by the ratio of

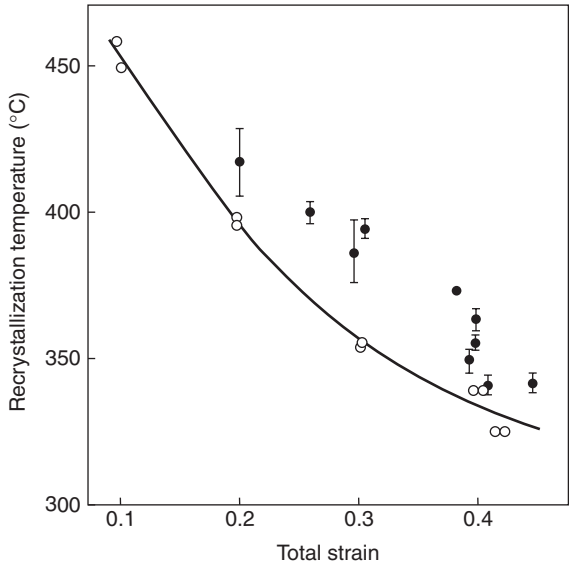


Figure 43 The dependence of the recrystallization temperature on prior strain (○, simple tension, ● compression + tension). (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

contractile strains during a tension test,

$$R = \frac{\varepsilon_w}{\varepsilon_t} \qquad [7]$$

where w and t denote the width and thickness directions in a tension test. Because R changes with the direction of the tensile axis, it is common to measure the values at 0°, 45°, and 90° to the rolling direction and to define an average $\gg R$ as

$$\gg R = \frac{(R_0 + 2R_{45} + R_{90})}{4} \qquad [8]$$

Like all f.c.c. metals, copper and copper alloys have R-values considerably <1.0. For annealed copper, typically, $R_0 = 0.9$, $R_{45} = 0.4$, $R_{90} = 0.9$, and $\gg R = 0.6$ or 0.7 and for annealed 70:30 brass, typically, $R_0 = 0.9$, $R_{45} = 0.9$, $R_{90} = 0.8$, and $\gg R = 0.9$. The cube texture component lowers R_{45} considerably, while components with {1 1 1} parallel to the sheet raise the R-values.

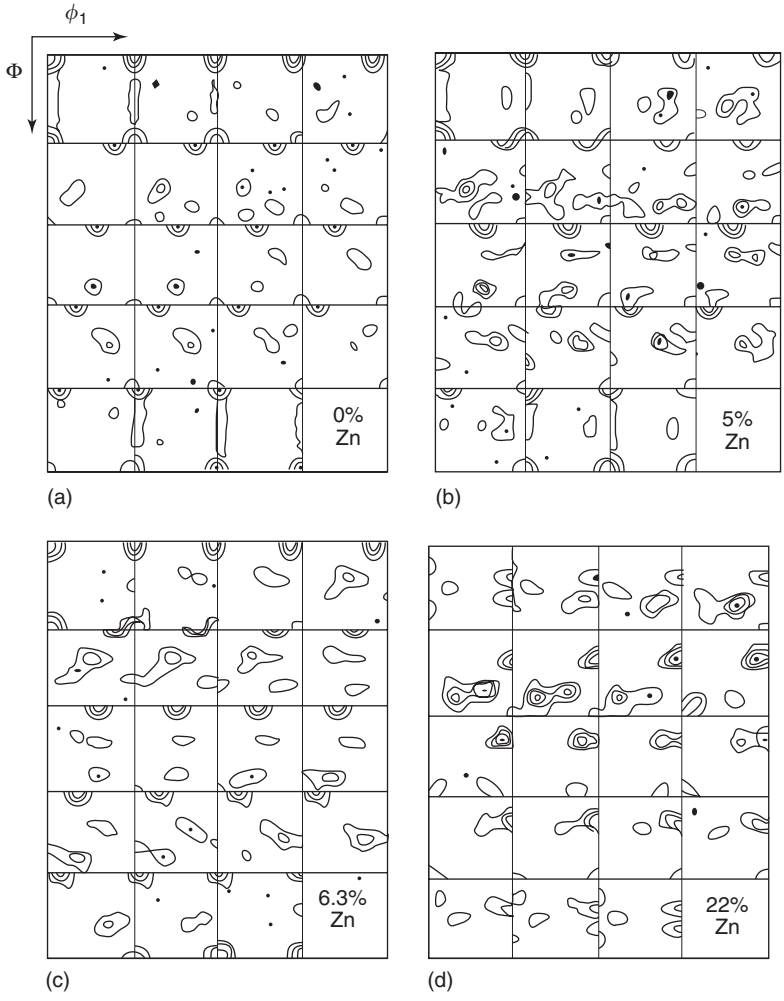


Figure 44 Recrystallization textures of rolled and recrystallized copper and copper alloys; (a) Cu, (b) 5% Zn, (c) 6.3% Zn, and (d) 22% Zn. (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

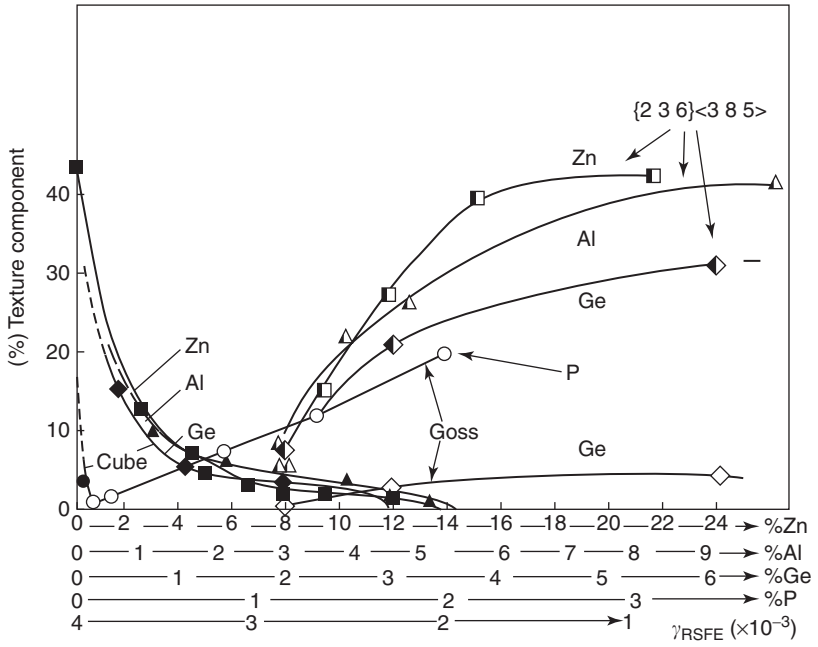


Figure 45 The transition from copper to brass texture depends on the stacking fault energy of the alloy. (Cube: Zn (■), Al (▲), Ge (◆), P (●); {2 3 6} ⟨3 8 5⟩: (Zn) (◧), Al (◭), Ge (◭); Goss: Ge (◇), P (○).) (Adapted from Humphreys FJ and Hatherly M (1996) *Recrystallization and Related Phenomena*. London: Elsevier.)

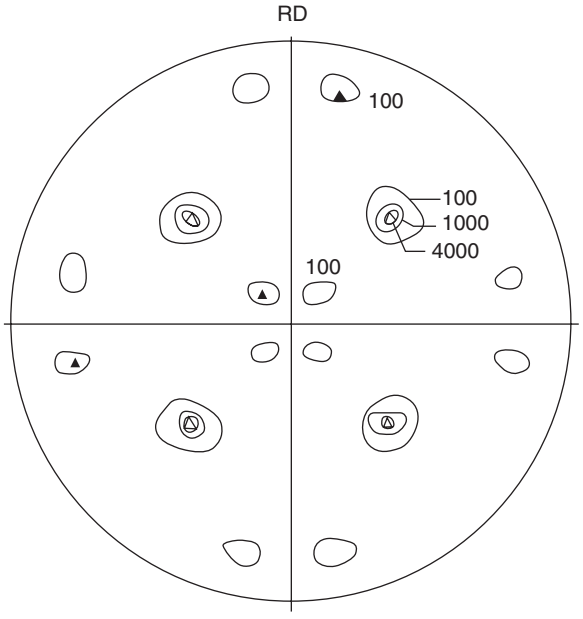


Figure 46 The {1 1 1} pole figure of copper cold rolled 97% at room temperature and annealed at 200°C. △ show the cube texture and the ▲ show its twin components (1 2 2)[2 1 2]. (Adapted from Hu H, Sperry PR, and Beck P (1952) *Trans AIME* 194.)

Forming

Sheet formability benefits from high strain-hardening exponents and high *R*-values. Forming limit diagrams (see **Figure 48**) show the combinations of strains that lead to the formation of localized necks. Minima occur near plane strain, $\varepsilon_2 = 0$. The level of ε_1 at this point depends strongly on the strain-hardening exponent, n in eqn [2]. Cold-worked sheets tend to have lower forming limits. The forming limits tend to rise with sheet thickness. A high strain-rate exponent m, in

$$\sigma = C\varepsilon^m \qquad [9]$$

is also useful in forming. However, it is found to be very low (0.002–0.005) for copper and brass at room temperature.

The compositions and conditions of these sheets are listed in **Table 8**.

The limiting drawing ratio (LDR) for flat bottom cups depends mainly on the *R*-value of the sheet material. Because the *R*-values of copper and brass sheets tend to be <1, the values of LDR tend to fall between 2.1 and 2.2. Formability in sheet forming that involve biaxial stretching depends mainly on the strain-hardening exponent n. Because of this, copper and especially brass have very good stretchability.

Extrusion and Forging

Hot forging temperatures for most copper alloys lie in the range of 600–750°C.

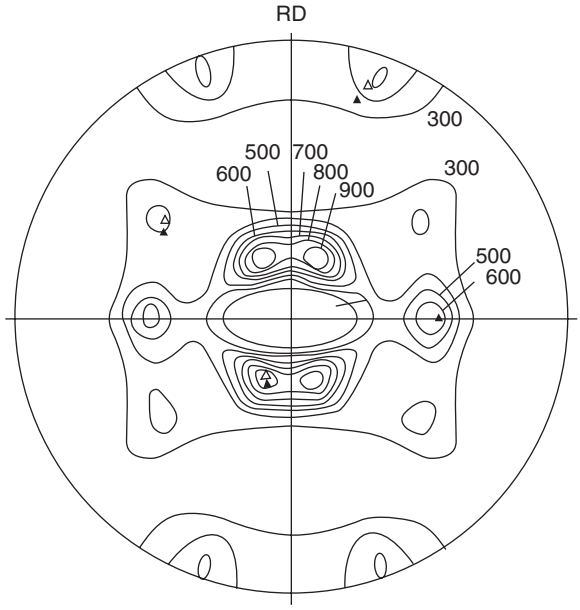


Figure 47 The {1 1 1} pole figure of 70:30 brass cold rolled 97% at room temperature and annealed at 340°C. △ show the ideal (2 2 5)[3 7 4] orientation and the ▲ show the (1 1 3)[2 1 1] orientation. (Adapted from Hu H, Sperry PR, and Beck P (1952) *Trans. Met. Soc. AIME* 194.)

Powder Processing

Parts of copper and copper alloys are sometimes fabricated by conventional powder metallurgy that involves pressing powder to the desired shape and sintering to achieve nearly full density. Porous bearings and filters are made by stopping the sintering before complete densification.

Casting

In general, alloys with a greater separation of liquidus and solidus are regarded as easier to cast. In a fixed thermal gradient, the length of dendrite L, arms is proportional to this temperature difference, ΔT. The liquid-to-solid shrinkage can occur interdendritically. This simplifies risering. Lower liquidus temperatures are also beneficial.

The solubility of hydrogen, like other gases, is much greater in liquid copper than solid copper. **Figure 49** shows the solubility of hydrogen in copper and copper–aluminum alloys. Alloying with aluminum and some other elements decreases the solubility.

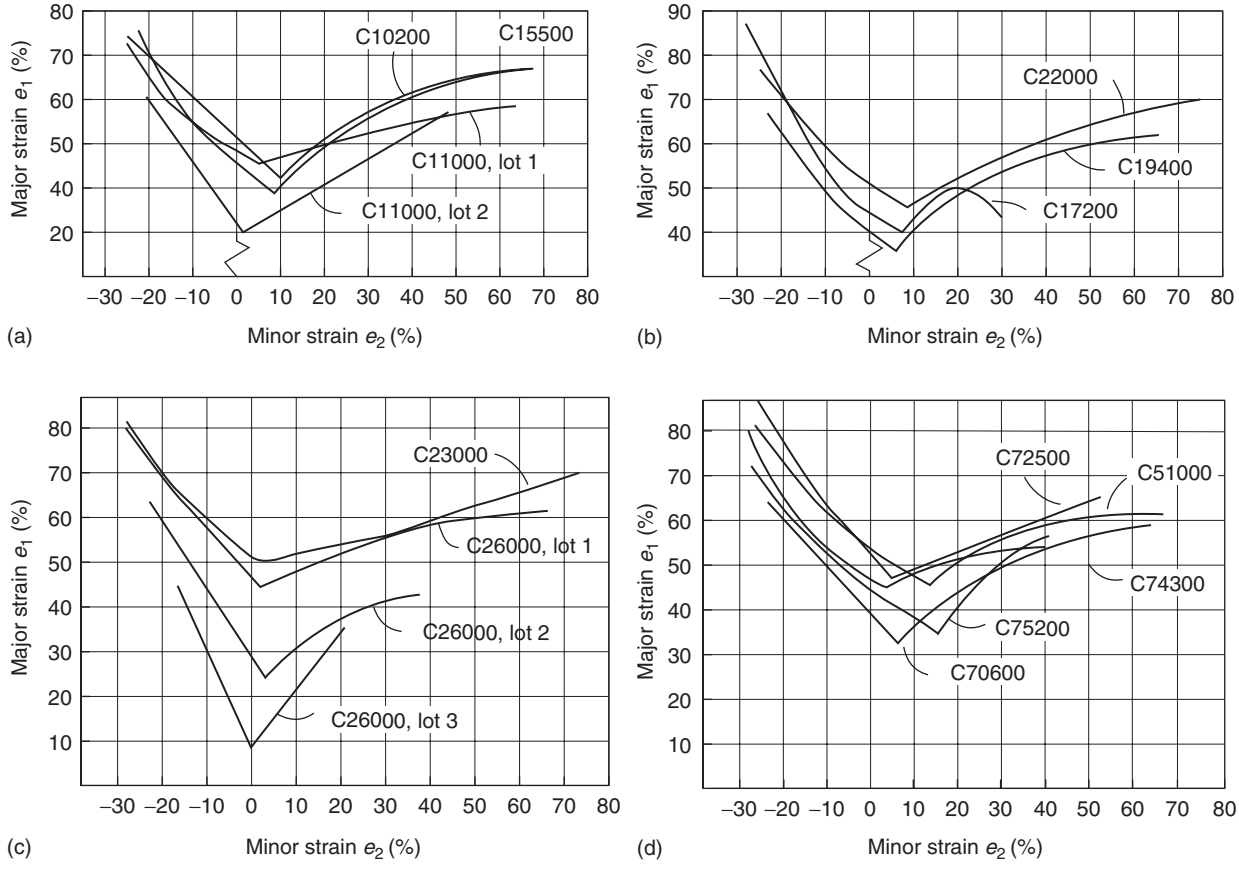


Figure 48 Forming limit diagrams for copper and copper alloy sheets.

Table 8 Alloys shown in **Figure 40** and their conditions

Designation	*Common name*	*Condition*	*Thickness* (mm)	*Grain size* (mm)	UTS[a] (MPa)
C10200	Oxygen free copper	annealed	0.66	0.014	234
C11000 lot 1	ETP copper	annealed	0.74	0.016	224
C11000 lot 2	ETP copper	H4	0.69		268
C15500	Silver copper	annealed	0.71	0.009	288
C220000	Commercial bronze	annealed	0.69	0.0067	234
C19400	HSM copper	annealed	0.69		319
C17200	Beryllium copper	annealed	0.25	0.019	491
C2300	Red brass	annealed	0.69	0.024	293
C2600 lot 1	Cartridge brass	annealed	0.64	0.025	345
C2600 lot 2	Cartridge brass	H4	0.69		407
C2600 lot 3	Cartridge brass	H8	0.51		531
C72500	Cu–Ni–Sn	annealed	0.69	0.023	356
C51000	Phosphor bronze	annealed	0.69	0.014	374
C74300	Nickel silver	annealed	0.69	0.035	387
C72500	Nickel silver	annealed	0.69	0.020	405
C70600	Cu–10% Ni	annealed	0.81	0.016	361

[a]UTS – ultimate tensile strength.
Adapted from Copper and Copper Alloys, ASM Specialty Handbook. Materials Park, OH: ASM.

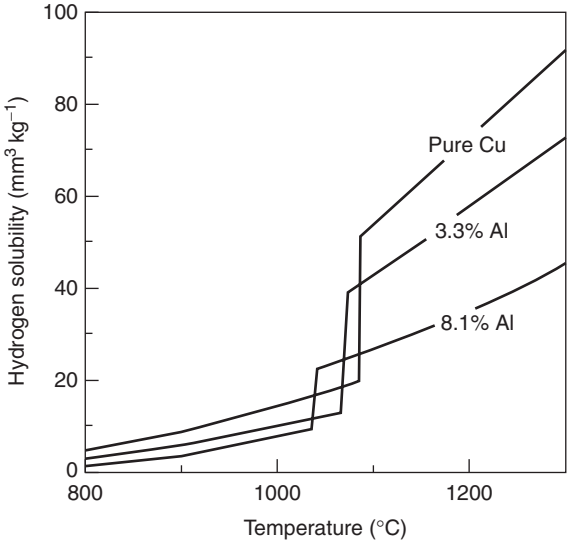


Figure 49 Solubility of hydrogen in copper and Cu–Al alloys. (Adapted from Flinn R A (1963) *Fundamentals of Metal Casting*. Reading, MA: Addison-Wesley.)

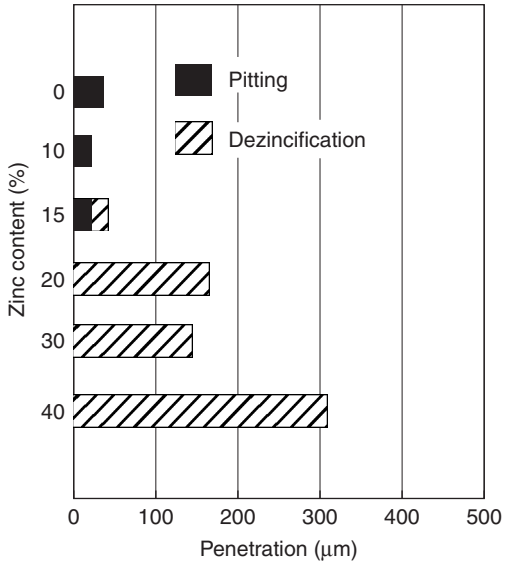


Figure 50 The effect of zinc content on corrosion of brass when exposed to 0.01 M NH_4Cl at 45°C for 60 days. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*. Materials Park, OH: ASM.)

As the metal solidifies, dissolved gas is released. For hydrogen,

$$H_2O + 2Cu \rightarrow 2\underline{H} + Cu_2O \qquad [10]$$

This is likely to cause gas porosity. The source of dissolved hydrogen may be from wet scrap or furnace gases.

Corrosion

Many of the applications of copper and copper-based alloys depend on their generally good corrosion resistance. A few specific forms of corrosion are given below. One is pitting corrosion in which the attack is very localized. Pitting attacks cause much more damage than the same amount of metal loss uniformly spread over the corroded surface. A second type of corrosion is dezincification of brasses containing >15% Zn. Both copper and zinc atoms go into solution, but copper then precipitates as a porous plug. **Figure 50** shows the effect of composition on

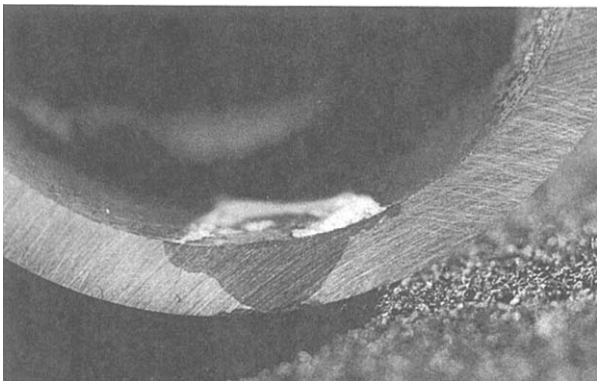


Figure 51 A plug of dezincified metal on a brass pipe. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*. Materials Park, OH: ASM.)

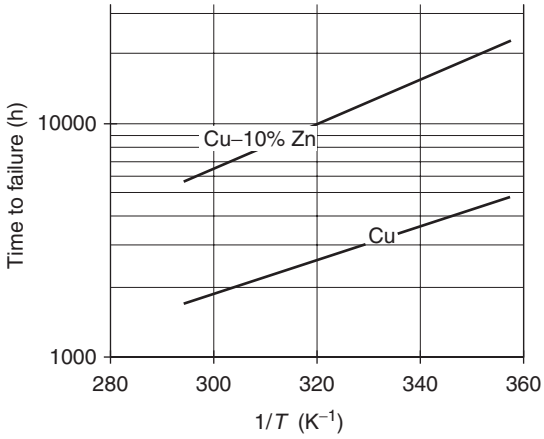


Figure 53 Arrhenius plot of time to failure of copper and brass under constant load in NH_4OH in the temperature range of 40–70°C. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*, Materials Park, OH: ASM.)

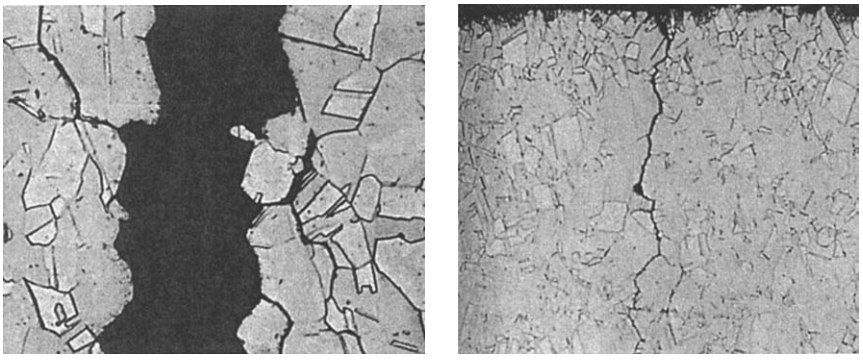


Figure 52 Examples of stress-corrosion cracking in brass. Both condenser tubes of drawn C12200 copper. (Adapted from ASM (2001) *Copper and Copper Alloys*, *ASM Specialty Handbook*. Materials Park, OH: ASM.)

the tendency to pitting and dezincification. An example of dezincification is shown in **Figure 51**.

Brass containing $\geqslant 15\%$ Zn is susceptible to stress corrosion or season cracking. It occurs only when the brass is under tensile stresses and specific environments. For brass, ammonia is the most common agent. **Figure 52** shows examples of stress-corrosion cracking of brass. The propensity to stress corrosion increases with zinc content and temperature as shown in **Figure 53**. Removal of residual stresses by stress-relief anneals is sufficient for many applications. Hydrogen is not a problem for most copper alloys. However, if tough-pitch copper containing Cu_2O is exposed to hydrogen at high temperatures, the reaction is

$$2\underline{H} + Cu_2O \rightarrow H_2O + 2Cu \qquad [11]$$

Such exposure could, for example, occur during torch welding. The formation of H_2O in the form of steam causes embrittlement.

See also: Deformation Processing; Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Mechanical Properties: Tensile Properties; Orientation Texture; Powder Processing: Models and Simulations; Solidification: Models and Simulations.

PACS: 81.05.Bx; 81.20.–n; 81.30.Dz; 81.40.–z; 81.70.–q

Further Reading

Haasen P (1978) *Physical Metallurgy*. Cambridge: Cambridge University Press.

Hu H and Goodman SR (1952) *Trans. AIME* 194.

Joseph G (1999) *Copper; it Trade, Manufacture, Use and Environmental Status*. Materials Park, OH: ASM.

Mendenhall JH (ed.) (1980) *Understanding Copper Alloys*. Olin Brass, New York: Wiley.

Reed-Hill RE and Abbaschian R (1994) *Physical Metallurgy Principles*, 3rd edn. Boston: PWS-Kent.

Smithells Metals Reference Book, London; Butterworths, 1983.

West EG (1882) *Copper and Its Alloys*. New York: Wiley.

Alloys: Iron

M Ferry, University of New South Wales, Sydney, NSW, Australia

Introduction

Iron alloys are arguably the most important class of engineering materials as a huge number of types can be produced by controlling the type and quantity of alloying elements, casting method, thermal and mechanical processing, etc. Factors contributing to the attractiveness of iron as the principal component of many engineering alloys include: the ability to be produced relatively inexpensively; it has a high melting temperature (1538°C) which allows thermally activated processes over a range of temperatures and it is allotropic thereby allowing its alloys to undergo several useful phase transformations to generate a wide variety of microstructures, and mechanical and physical properties.

Figure 1 presents a subgroup of a much wider range of commercially available iron alloys. The alloys given in this figure contain carbon, are generally classed as steels or cast irons, and are by far the most widely used iron alloys but it is pertinent to note that many others are also important. Depending on alloying additions and the complexity of the various processing stages, microstructures (and properties) of steels and cast irons vary dramatically, and include one or more of the following phases: ferrite, pearlite, bainite (range of structures), austenite, and martensite as well as second-phase dispersions such as cementite, graphite, intermetallic compounds, and precipitates. A detailed treatment of the structure and properties of Fe–C alloys is beyond the scope of this overview but several excellent books are given in the "Further reading" section.

There are a range of essentially carbon-free iron alloys that also generate microstructures with attractive properties. Examples include Fe–Be and FePt that exploit the martensitic reaction to generate shape memory behavior and those alloys that exhibit important physical properties such as Fe–Ni invar alloys with very low thermal expansion coefficients and multicomponent alloys such as (Fe, $Co)_{44}Zr_{7}$-Cu_1B_{15} (HITPERM) and $Nd_2Fe_{14}B$ with exceptional magnetic and magnetomechanical properties. The ability to generate iron alloys with such a large range

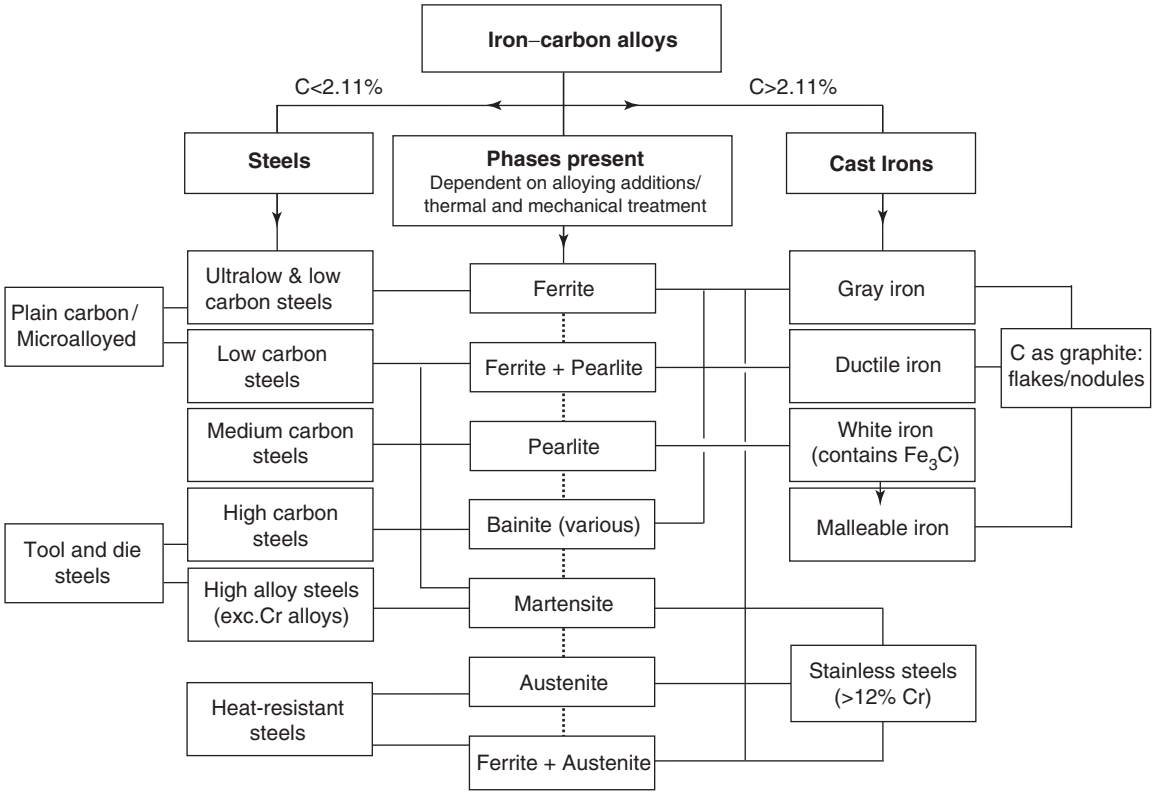


Figure 1 Classification scheme for iron–carbon alloys (range is not exhaustive).

of microstructures and properties has resulted in the use of these materials in an enormous number of applications.

The Production of Iron Alloys

Iron alloy components have been produced from the liquid state for many centuries with mass production of steel commencing in the late nineteenth century. Molten steel was traditionally cast into standing moulds but continuous casting was developed in the 1950s which enabled the production of huge tonnages of steel. Most of the world's steel production is now obtained through this casting route. Following casting, as-cast slabs are thermal and mechanical processed to the desired final shape. In the last few decades, however, near-net-shape casting processes have been developed that are capable of producing near-final-shape products directly from liquid with secondary processing such as hot rolling reduced to an absolute minimum. On a reduced scale, there are various types of novel iron alloys produced in the amorphous, nanocrystalline, and intermetallic form. These materials are either manufactured by conventional casting or by rapid solidification processing, spray deposition, mechanical attrition, and sintering or physical and chemical vapor deposition, etc. Iron alloys produced by these more exotic techniques are generally restricted to specific applications.

Structure and Properties of Iron

The production of pure iron is not possible, but numerous studies have been carried out on the properties of high-purity iron. Iron is situated toward the end of the first transition series in the periodic table with an electron configuration $1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s^2$. Iron lies at the left of the larger atomic species Co, Ni, and Cu which are known to crystallize with a close-packed structure. To the left of iron are the less-filled states of Ti, V, Cr, and Mn which subsequently crystallize with a body-centered cubic (b.c.c.) crystal structure. Iron lies between these groups of elements and, depending on temperature, may exist either as the b.c.c. structure ($1394^\circ C > T_\alpha < 912^\circ C$) or the face-centered cubic (f.c.c.) structure ($912^\circ C < T_\gamma < 1394^\circ C$). The high-temperature b.c.c. structure is not expected under normal circumstances but is a result of magnetic energy effects. The b.c.c. structure exists at standard temperature and pressure, but extreme pressures will generate a hexagonal close-packed (h.c.p.) structure (ε-iron).

An important difference between α-iron and γ-iron is the structure of their dislocations: the Burgers vector of the minimum energy configuration in α-iron is $a/2\langle 1\,1\,1\rangle$ and while dislocations can theoretically dissociate into segments of $a\langle 0\,0\,1\rangle$, the stacking fault energy (SFE) is so high that substantial dissociation has not been observed. In contrast, γ-iron has a relatively low SFE (<80 mJ m^{-2}) and dislocations readily dissociate by the reaction $a/2\langle 0\,1\,1\rangle \rightarrow a/6\ \langle 1\,1\,2\rangle + a/6\langle 2\,1\,1\rangle$, which generates a ribbon of stacking fault. The difference in both crystal structure and SFE of α-iron and γ-iron strongly influences the deformation and annealing behavior of these phases.

Table 1 Approximate diffusivity of various interstitial and substitutional elements in iron

Phase	*Element*	*Activation energy, Q (kJ mol^{-1})*	*Frequency factor, D_o (mm^2 s^{-1})*	*Temperature range (°C)*
α-Iron	Fe	240	200	700–750
	C	80	0.62	700–750
	N	76	0.3	700–750
	Co	226	20	700–790
	Ni	258	970	700–900
γ-Iron	Fe	284	49	900–1050
	C	135	15	900–1050
	Co	364	3000	1050–1250
	Ni	280	77	930–1050
δ-Iron	Fe	239	190	1400–450

Due to the various allotropic transformations in iron ($\delta \leftrightarrow \gamma \leftrightarrow \alpha$), many physical properties, such as specific volume and electrical conductivity, change discontinuously through these critical temperatures. A notable consequence of the f.c.c. $\leftrightarrow$ b.c.c. transformation is the change in magnitude of the self-diffusion coefficient, D, which follows an Arrhenius-type relationship:

$$D = D_o \exp(Q/RT)$$

where D_o is a frequency factor that is virtually independent of absolute temperature (T) and Q is the activation energy for self diffusion. **Table 1** shows the values of D_o and Q for iron and other alloying elements. It is easily shown that the rate of self-diffusion of iron decreases by two orders of magnitude when f.c.c. iron transforms to the more open b.c.c. structure.

Alloying Elements in Iron

A number of elements are added to iron to produce a variety of alloys with a wide range of microstructures and properties (**Figure 1**). Alloying elements have a strong influence on the stable temperature range of b.c.c. and f.c.c. phases. The Wever classification

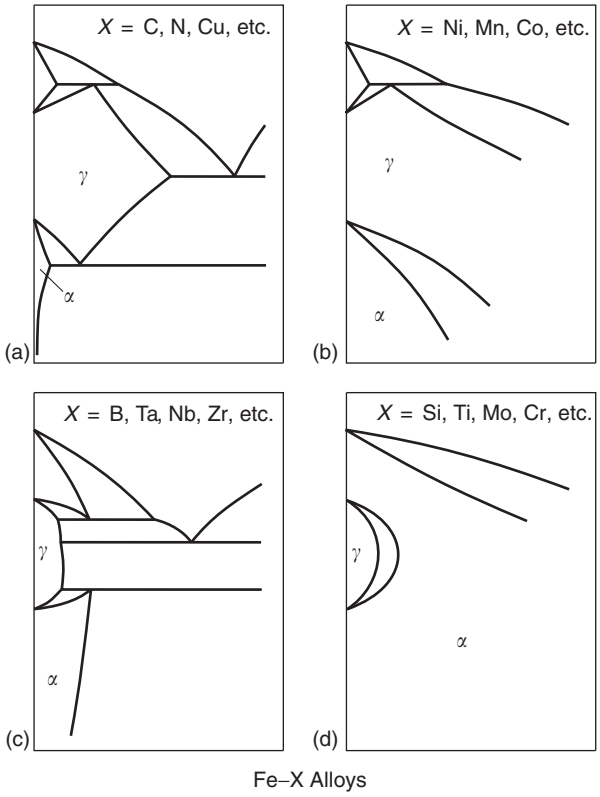


Figure 2 Classification of iron alloy binary phase diagrams: (a) and (b) austenite stabilizers; (c) and (d) ferrite stabilizers.

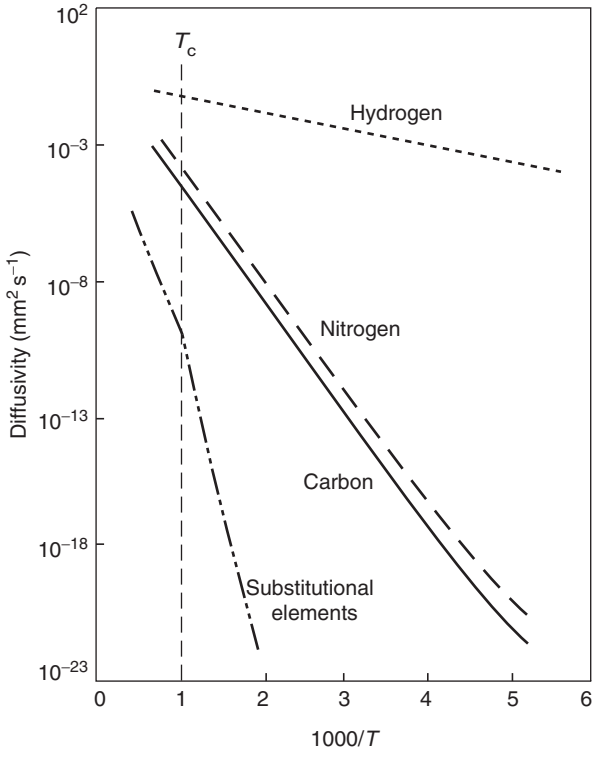


Figure 3 Variation of average diffusivity with temperature for various substitutional and interstitial elements in α-iron.

(**Figure 2**) has shown that most alloying elements fall into four categories with each generating a particular phase diagram. This figure shows that a particular alloying element in iron may either expand the γ-field (austenite stabilizer) or contract this field (ferrite stabilizer). Elements may also form iron-rich compounds that reduce the extent of the γ-field. The overall behavior is a result of the different enthalpies of solution in both the crystal structures of iron which leads to either an increase or decrease of the temperature range in which austenite is stable. Alloying elements can also be grouped into three major classes: those that form interstitial and substitutional solid solutions, and those that are almost completely immiscible with the crystal lattices of iron.

Interstitial Alloys

The elements H, C, N, and O are considerably smaller than Fe and, therefore, tend to occupy interstitial sites in both the lattices of iron. **Figure 2a** shows how the γ-field is stabilized for alloys containing interstitial elements since the f.c.c. lattice provides sites of lower strain energy than does the b.c.c. lattice. These elements have a high rate of diffusion in both the crystal structures of iron, particularly in the less densely packed b.c.c. structure, and are considerably more mobile than substitutional elements (**Figure 3** and **Table 1**). Nevertheless, interstitial elements have low solubility, particularly in α-iron (ferrite), due to the less favorable geometry of the b.c.c. lattice. The solubility of interstitial elements decreases markedly with temperature and obeys a similar relation as diffusivity:

$$C_{\mathrm{Fe-X}}(T) = A \exp(B/T)$$

where A and B are constants and X is either an equilibrium or metastable phase such as graphite, Fe_2B, Fe_3C, and Fe_3N_4 in equilibrium with either ferrite or austenite. **Table 2** gives values of the maximum solubility of various interstitial elements in ferrite and austenite which shows that hydrogen has the highest solubility in ferrite which is demonstrated by a small value of the constant, B (not shown). **Table 2** also shows that the solubility of C and N in both forms of ferrite are considerably lower than in austenite.

In the Fe–C system, graphite does not usually form in either ferrite or austenite due to its high activation energy of nucleation. The cementite phase (Fe_3C) or other less stable carbides tend to form instead. The

Table 2 Solubility of various interstitial elements in iron

Phase	*Element*	*Equilibrium phase*	*Temperature of maximum solubility* (°C)	*Maximum solubility* (at.%)
α-Iron	H	H_2	905	$1-2\times10^{-3}$
	B	Fe_2B	915	<0.02
	C	Fe_3C	727	0.095
	N	Fe_4N	590	0.4
γ-Iron	C	Fe_3C	1148	8.8
	N	Fe_4N	650	10.3
δ-Iron	C	Fe_3C	1495	0.09

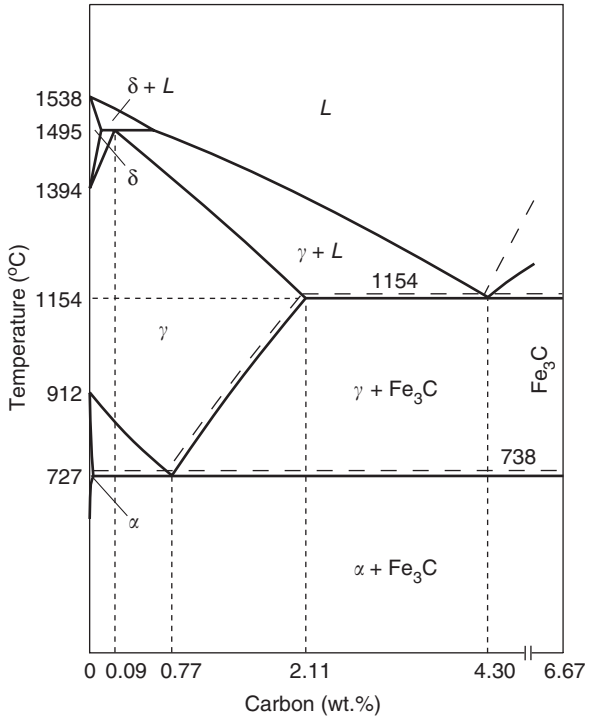


Figure 4 Schematic diagram of the metastable (constitutional) Fe–Fe_3C and the stable (equilibrium) Fe–graphite diagram (dashed) showing important temperatures and compositions.

difference between the equilibrium Fe–graphite and metastable (constitutional) Fe–Fe_3C phase diagram is shown in **Figure 4**, which illustrates the subtle differences in the phase boundaries. Fe–Fe_3C is probably the most important binary system with the following array of characteristics: (1) various types of invariant phase transformations are possible (peritectic, eutectic, and eutectoid) and (2) the solubility of C in α-iron at 727°C is only 0.022 wt.% whereas γ-iron can accommodate 0.77 wt.% C at this temperature. Carbon levels below ~2.11 wt.% are classified as steels while higher carbon alloys are classified as cast iron. Steels are classified further into hypoeutectoid (C < 0.77 wt.%) or hypereutectoid (C > 0.77 wt.%) while cast irons are classified as either hypoeutectic (C < 4.30 wt.%) or hypereutectic.

Substitutional Alloys

There are many substitutional elements that are partially or completely soluble in both the crystal structures of iron. Elements with the largest solubility fall close to Fe in the periodic table: Cr, Ni, Co, Mn, and V. A number of elements, such as Ti, Ca, and K, have almost no solubility in iron. For alloying elements with an f.c.c. or h.c.p. structure, the γ-field is stabilized (**Figures 2a** and **2b**) whereas the γ-field is reduced by elements that form Hume–Rothery phases with Cu, Ag, and Cu (**Figure 2c**) or when iron is alloyed with the b.c.c. transition metals V, Ti, Mo, W, Cr, etc. (**Figure 2d**). For those elements that restrict the γ-field completely, austenite cannot exist which eliminates the possibility of generating several important decomposition products (pearlite, bainite, or martensite). These alloys are non heat-treatable in the sense that phase transformations cannot be exploited to control the microstructure and properties. To achieve the desired mechanical properties such as yield strength, these alloys must either be work hardened (which can also recrystallize to produce a fine grain size) or elements are used to promote solid solution strengthening. The heat-treatable alloys are far more amenable to strengthening over a wide range by controlling the transformation products during austenite decomposition.

Most alloys of iron usually consist of a combination of interstitial and substitutional elements which results in various complex interactions involving atom species. For example, the rate of diffusional transformation is usually controlled by the more sluggish substitutional elements (as shown by the lower diffusivity of these elements, **Figure 3**) which favors the formation of nonequilibrium microstructures following austenite decomposition. In addition, elements can combine to produce enhanced strengthening by the generation of atom complexes (Mn–C, etc.) in the lattice or from the formation of alloy carbides ($[Fe_aX_bY_{c\ldots}]C_j$).

Phase Transformations in Iron Alloys

The allotropic transformations that occur in pure iron can result in a number of important phase transformations in its alloys. As indicated in **Figure 2**, a given concentration of a ferrite stabilizer can eliminate the austenite phase field, whereas an austenite stabilizer will expand this field to allow the alloy to undergo several useful transformations. For

example, **Figure 4** shows that a number of transformations are possible in the Fe–C system. The following brief discussion is devoted mainly to this system as carbon is the most important alloying addition to iron.

During solidification, the peritectic reaction can occur over a certain compositional range. For the peritectic composition (0.09 wt.% C), the following invariant reaction can occur at 1495°C: $L \rightarrow L + \delta \rightarrow \gamma$. In this reaction, liquid + δ-iron transforms isothermally to austenite and has several important consequences during continuous casting of steels.

At higher carbon levels, that is, 0.77 wt.% C, the eutectoid reaction occurs at 727°C: $\gamma \rightarrow \alpha + Fe_3C$. For slow cooling rates, austenite (0.77 wt.% C) decomposes via a diffusional transformation to produce pearlite which is a lamellar structure containing successive layers of ferrite (0.022 wt.% C) and cementite (6.67 wt.% C). With increasing cooling rate, the formation of diffusional products such as pearlite and ferrite is suppressed thereby generating nonequilibrium microstructures ranging from a series of bainitic structures to martensite.

At very high cooling rates, diffusional transformation products do not form since atomic diffusion is necessary for nucleation and growth of these phases. This results in the formation of martensite; a term not restricted to iron-base alloys but used to define the product of a nondiffusional phase transformation both in metals and ceramics. The martensitic transformation has the following important characteristics: each atom retains its original neighbors and there is no interchange among the atoms; the transformation does not involve individual atomic jumps characteristic of diffusion-controlled and interface-controlled transformations; the reaction is diffusionless which means the martensite product has the same composition as the parent phase. In Fe–C alloys, the martensite transformation generates a hard, brittle phase whereas, for essentially carbon-free alloys, the martensite is ductile.

For very low carbon levels (< 0.022 wt.%), Fe–C may undergo a series of precipitation reactions after cooling through the solvus where supersaturated carbon in ferrite decomposes at low temperatures to produce a series of precipitates. This can result in considerable hardening that may or may not be useful depending on the desired application of the alloy.

For carbon levels greater than ~2.11 wt.%, Fe–C undergoes the eutectic reaction: $L \rightarrow \gamma + Fe_3C$. At the eutectic composition (4.3 wt.% C), liquid iron transformations take place at 1147°C to produce austenite and cementite with subsequent cooling to room temperature promoting additional phase changes. It is pertinent to note that alloying additions such as Si promote the formation of graphite rather than cementite (**Figure 1**). Similar to steels, cast irons can generate a wide range of microstructures and properties depending on alloying additions, cooling behavior, etc.

It is clear that the transformation products that form during austenite decomposition in iron-base alloys are expected to be diverse and are influenced by a large number of processing and material-related variables. A particular transformation is achieved by the control of the type and quantity of alloying elements, the cooling rate through the transformation range, and thermomechanical treatments prior to or during transformation, etc. In addition to the classic phase transformations associated with Fe–C, several additional transformations are known to occur in iron alloys. The ability of the alloys of iron to undergo such a diverse range of phase transformations is a major factor contributing to their immense popularity as an engineering material.

Thermomechanical Behavior

An understanding of deformation of ferrite and austenite is extremely important as these phases undergo a significant amount of plastic deformation during thermomechanical processing, the most common method for producing useful final components. During deformation, many microstructural changes occur: the original grains change shape and an internal substructure forms, texture changes take place, precipitation may occur, dynamic recovery (DRV) or dynamic recrystallization (DRX) processes are possible and the constituent particles may fracture and redistribute. The most notable material factors that affect both the deformation microstructure of iron alloys include SFE, crystal structure, initial grain size and shape, and the size, shape, and volume fraction of a second phase.

Cold Deformation

Cold deformation is restricted to low homologous temperatures ($T < 0.5T_m$), where T_m is the absolute melting temperature. During cold deformation of either ferrite or austenite, there is a marked increase in dislocation density resulting in considerable work hardening and grains subdivide in a complex manner to produce a range of features such as a cellular substructure, microbands, deformation twins, deformation bands, and larger-scale heterogeneities such as shear bands. In high SFE b.c.c. α-iron, slip is the principal deformation process and occurs in the close-packed $\langle 1\,1\,1 \rangle$ direction but the slip plane may be any of the planes {1 1 0}, {1 1 2}, or {1 2 3}. In

contrast, γ-iron has a relatively low SFE ($<80\,\mathrm{mJ\,m^{-2}}$) which results in deformation by slip on $\{1\,1\,1\}\langle 0\,1\,1\rangle$ systems as well as by twinning. The different modes of deformation result in differences in the work-hardening rate of these phases as well as in the development of deformation textures.

Hot Deformation

Hot deformation is carried out at $T>0.5T_{\mathrm{m}}$ and is distinguished from cold working by the absence or near absence of strain hardening and lower matrix dislocation content. Deformation becomes more homogeneous with increasing temperature with the frequency of microstructural inhomogeneities such as microbands, transition bands, and shear bands reduced considerably. The driving pressure for softening by recovery and recrystallization (either dynamic or static) is the stored energy of deformation. The process of recovery reduces the internal energy of the metal through mechanisms such as annihilation dislocations and their re-arrangement into low-angle grain boundaries.

The relatively low SFE of austenite does not allow for substantial DRV whereas DRV occurs extensively in ferrite since dislocation climb occurs readily and, for a given temperature, the rate of diffusion of iron atoms in ferrite is $\sim 100x$ greater than in austenite. In ferrite, a significant proportion of the stored energy is released by DRV which reduces the driving force for both dynamic and static recrystallization (SRX). Since the reorganization of dislocations is more difficult in austenite, DRX can occur readily. These differences in the dynamic restoration processes generate two distinctive stress–strain (flow) curves, **Figure 5**. Deformation of austenite results in an increase in flow stress via work hardening but at a critical strain (ε_{c}), there is sufficient stored energy to initiate DRX which tends to eventually decrease the flow stress since the rate of work hardening is offset by the softening caused by recrystallization. In contrast, extensive DRV in ferrite leads to a balance between dislocation generation and annihilation which rapidly results in a steady-state flow stress. There is also insufficient stored energy to dynamically recrystallize this phase. During hot working, successive cycles of DRX in austenite have the capability of refining the grain size whereas grain refinement in ferrite can only be achieved by SRX.

Static Annealing

The principal material factors influencing static annealing in iron alloys include composition, initial grain size, initial texture, and second-phase precipitates. Processing variables are also important and include the mode of deformation, strain, strain rate, and temperature of deformation. In general, both SRX and DRX are initiated by nucleation of strain-free grains at microstructural inhomogeneities such as transition bands, shear bands, and grain boundaries. Cold-deformation of fine-grained ferrite results in nucleation predominantly at prior grain boundaries but a coarse grain size results in substantial shear band nucleation, particularly in those iron alloys containing high levels of interstitial elements.

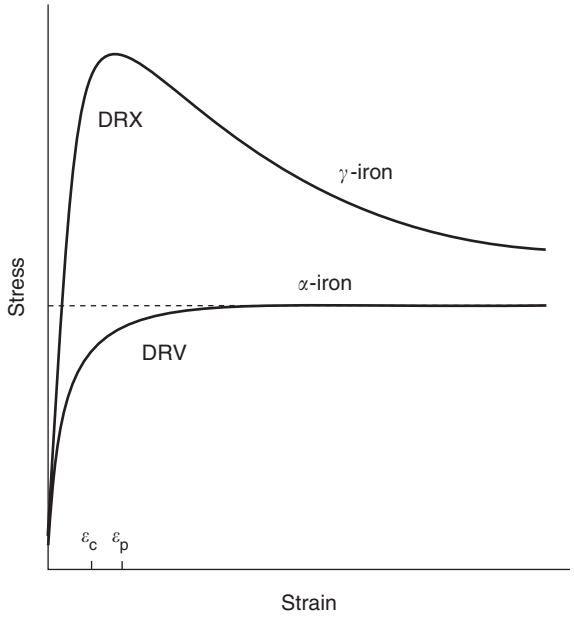


Figure 5 Schematic diagram of typical flow curves associated with the hot deformation of ferrite (α-iron) and austenite (γ-iron).

The most desirable nucleation sites in cold-rolled α-iron are grain boundaries as these nuclei often have orientations where $\langle 1\,1\,1\rangle$ is almost parallel to the normal direction (ND) of the sheet. This is significant since a strong $\langle 1\,1\,1\rangle \| \mathrm{ND}$ texture in cold-rolled and annealed steels results in superior sheet formability. Austenitic alloys recrystallize in a similar manner to ferritic alloys, that is, nucleation occurs mainly at grain boundaries but other microstructural heterogeneities characteristic of deformed austenite also play an important role. Due to the physical processes associated with slip and twinning during cold rolling, austenitic alloys do not produce a high fraction of favorable $\langle 1\,1\,1\rangle \| \mathrm{ND}$ nuclei resulting in reduced formability.

Recent Developments in Iron Alloys

Magnetic Alloys

Iron is ferromagnetic below the Curie temperature and these alloys exhibit a remarkably wide range of

Table 3 Some notable ferromagnetic iron alloys showing typical properties

Material	*Initial relative permeability* (μ_e)	*Hysteresis loss per cycle* ($J\,m^{-3}$)	*Saturation induction* ($Wb\,m^{-2}$)	$(BH)_{max}$ ($kJ\,m^{-3}$)
Soft magnetic alloys				
Commercial iron ingot	1.5×10^2	270	2.14	
Fe-3Si (oriented silicon iron)	1.5×10^3	40–140	2.0	
Fe-45Ni (PERMALLOY 45)	2.7×10^3	120	1.60	
$Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ (FINEMET)	5×10^3–2×10^5		1.0–1.2	
$Fe_{88}Zr_7B_4Cu_1$ (NANOPERM)	1×10^4–1×10^5		1.5–1.8	
$Fe_{44}Co_{44}Zr_7B_4Cu_1$ (HITPERM)	2×10^3–3×10^4		1.6–2.1	
Hard magnetic alloys				
Fe-1Mn-0.9C (martensitic)				3.8
Fe-3.5Cr-0.9C-0.3Mn (martensitic)				5.0
Fe-24.5Co-13.5Ni-8Al-2Nb (Alnico-XII)				76.8
$Fe_{60}Pt_{40}$				120
$Nd_2Fe_{14}B$ (oriented)				320–400
$Nd_2Fe_{14}B$ (sintered nanophase)				450

magnetic properties. Magnetic materials are broadly classified into two main groups with either soft or hard magnetic characteristics. Soft magnetic alloys have high permeability, low coercivity, and low hysteresis loss whereas hard magnetic (or permanent) magnetic alloys have a sufficiently large resistance to demagnetizing fields due to their high coercivity, remanence, and energy product.

Table 3 gives some properties of a select range of ferromagnetic alloys which includes some of the more recently developed alloys that possess either an amorphous, or a nanocrystalline phase embedded in an amorphous matrix. These materials are usually produced as ribbons or fibers by rapid solidification techniques and are often followed by a series of heat treatments. Soft magnetic alloys based on these structures are generally complex and contain elements such as Cr, Mo, and Al and varying amounts of metalloids (B, Si, C, and P) that help to form the glassy phase. These alloys have low anisotropy combined with high resistivity (which reduces eddy current losses) and have high permeability and low power losses. The most widely investigated alloys in this new class of soft magnetic materials are $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ (FINEMET), $Fe_{88}Zr_7B_4Cu_1$ (NANOPERM), and $Fe_{44}Co_{44}Zr_7B_4Cu_1$ (HITPERM) which are produced by rapid solidification and crystallization annealing to produce a nanocrystalline structure. Compared with more conventional iron alloys, these materials have very high permeabilities and large inductions (**Figure 6**).

Permanent magnetic iron-base materials include a variety of alloys and intermetallic compounds. The most widely studied materials in recent years are neodymium–iron–boron ($Nd_2Fe_{14}B$) and iron–platinum alloys. For permanent magnets, a parameter

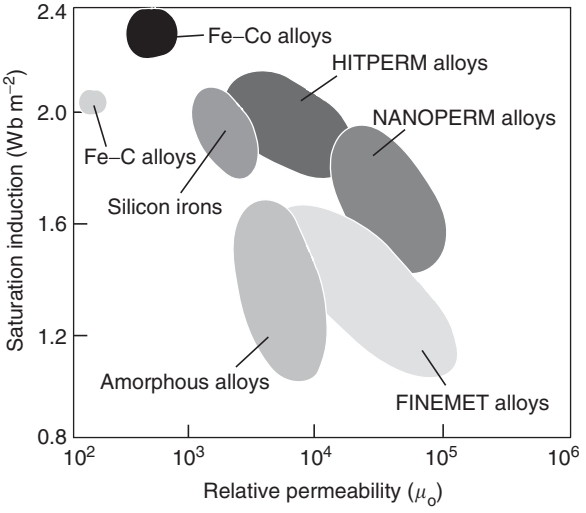


Figure 6 Estimated relationship between permeability and saturation induction for some soft magnetic iron alloys.

that specifies its performance characteristic, that is, its resistance to demagnetization, is the maximum energy product $(BH)_{max}$ where B and H are the magnetic induction and magnetic field strength, respectively. To date, the highest values of $(BH)_{max}$ are obtained in oriented Nd–Fe–B alloys (320–400 $kJ\,m^{-3}$) and sintered fine-grained versions of these alloys ($\sim$450 $kJ\,m^{-3}$).

Shape Memory Alloys

Iron alloys can exploit the martensitic reaction to enable a component to be deformed at one temperature with subsequent recovery of the original shape upon heating. This so-called shape memory effect (SME) is associated with the ability of some martensites to

undergo thermal reversion on reheating which, on a macroscopic scale, corresponds to up to 15% strain recovery. There are various types of SME including pseudoelasticity, one- and two-way shape memory behavior and magnetic field-induced shape memory. Iron-base shape memory alloys are produced either by conventional casting and thermomechanical processing, or mechanical attrition and sintering, or rapid solidification processing. While the various types of shape memory phenomena occur in alloys of the type Fe–Mn–Si–Cr, FeBe, FePt, and Fe–Pd–Pt, they are largely under development and more work is required to explore the potential of shape memory iron alloys.

Intermetallic Compounds

Iron alloys in the form of B2-type ordered intermetallic compounds based on the stoichiometric ratios, FeCo, FeAl, and Fe_3Al, have received considerable attention in high-temperature structural applications due to their low density, high electrical resistivity, high strength at temperature, and excellent resistance to oxidation and corrosion in various aggressive chemical environments. These alloys are either thermomechanically processed in a manner similar to metals or produced by more novel processing routes. While binary iron-base intermetallics are inherently brittle and have low creep resistance, considerable work is being carried out to improve these critical properties both by alloy design and by carefully controlling the various processing stages.

Direct Strip Cast Alloys

There have been substantial developments in the mass production of iron alloys by direct strip casting (DSC) which produces as-cast strip products of a thickness less than 2 mm. Iron alloys currently produced by DSC include carbon and alloy steels, stainless steel, iron–silicon alloys, and cast iron. This casting process often produces a microstructure that is markedly different from the same material produced by conventional casting and thermomechanical processing (TMP). This is a result of the high solidification rates that tend to produce far-from-equilibrium microstructures. By a careful control of casting parameters and alloying additions, as-cast low carbon steel strip can be generated with a final microstructure consisting of one or more of the following phases: martensite, bainite, acicular ferrite, and polygonal ferrite.

See also: Ferromagnetism; Magnetic Materials and Applications; Powder Processing: Models and Simulations; Solidification: Models and Simulations; Alloys: Overview; Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Recovery, Recrystallization, and Grain Growth; Mechanical Properties: Strengthening Mechanisms in Metals; Diffusionless Transformations; Phase Transformation; Phases and Phase Equilibrium.

PACS: 61.43.Dq; 61.44.+w; 61.50.Ks; 61.66.Dk; 61.72.Cc; 62.20.−x

Further Reading

Dunne DP (2000) Functional memory metals. *Materials Forum* 24: 95–108.

Gladman T (1997) *The Physical Metallurgy of Microalloyed Steels*. London: Institute of Materials.

Honeycombe RWK (1985) *The Plastic Deformation of Metals*. London: Edward Arnold.

Honeycombe RWK and Bhadeshia HKDH (1995) *Steels – Microstructure and Properties*. London: Edward Arnold.

Hume-Rothery W (1967) *The Structure of Alloys of Iron*. UK: Pergamon Press.

Humphreys FJ and Hatherly M (2004) *Recrystallization and Related Annealing Phenomena*, 2nd edn. Oxford: Elsevier Science.

Jiles DC (2003) Recent advances and future directions in magnetic materials. *Acta Materialia* 51: 5907–5939.

Leslie WC (1981) *The Physical Metallurgy of Steels*, 1st edn. London: McGraw Hill.

McHenry ME and Laughlin DE (2000) Nano-scale materials development for future magnetic applications. *Acta Materialia* 48: 223–238.

Manohar PA, Ferry M, and Hunter A (2000) Direct strip casting of steel: historical perspective and future direction. *Mater Forum* 24: 15–29.

Manohar PA, Ferry M, and Chandra T (2001) Recrystallization of ferrite and austenite. In: Buschow KHJ *et al*., (eds.) *Encyclopedia of Materials Science and Technology*, vol. 4, pp. 3019–3024. Oxford: Elsevier.

Munroe PR (2000) Intermetallic compound development for the 21st century. *Mater Forum* 24: 5–18.

Ray RK, Jonas JJ, and Hook RE (1994) Cold rolling and annealing textures in low and extra low carbon steels. *International Materials Review* 39: 129–172.

Sinha AK (1989) *Ferrous Physical Metallurgy*. Boston: Butterworths.

Zackay VF and Aaronson HI (eds.) (1962) *Decomposition of Austenite by Diffusional Processes*. USA: TMS-AIME.

Alloys: Magnesium

H Jones, University of Sheffield, Sheffield, UK

Fundamental Characteristics

Magnesium has a density of $1.74\,\mathrm{mg\,m^{-3}}$ at 20°C, melts at 650°C, has a Young's modulus of 45 GPa, and is hexagonal close-packed in crystal structure. Slip at ambient temperatures occurs mainly on the basal plane $(0\,0\,0\,1)\ \langle 1\,1\,\bar{2}\,0\rangle$ with pyramidal and prismatic planes becoming more active with increasing temperature. Deformation of room temperature therefore involves twinning also (initially on $(1\,0\,\bar{1}\,2)\langle 1\,0\,\bar{1}\,2\rangle$), especially when a compressive stress acts parallel to the basal plane.

Pure magnesium, in the absence of moisture, is protected by an adherent thin film of MgO. This MgO reacts with any moisture present to form $Mg(OH)_2$ which breaks down in the pressure of acids which then attack the magnesium. Saltwater is a particularly harsh environment for magnesium, especially when Fe, Ni, or Cu are present as precipitated impurities (**Figure 1**).

There are more than 20 elements with maximum equilibrium solid solubility exceeding 1 at.% in magnesium. Of these, eight are used in commercial magnesium alloys (Al, Zn, Mn, Zr, Li, Ag, Y, Nd) together with Si, Cu, Ca, and Ce. Al and Zn have large effects in reducing the lattice parameters of magnesium, and also raise the critical resolved stress for slip; so potent are solid solution hardeners in magnesium. Al is a less potent precipitation hardener than Zn or Nd, because it precipitates incoherently (e.g., as $Mg_{17}Al_{12}$ at grain boundaries) rather than via Guinier–Preston zones within the grains. The small grain size has a specially marked hardening effect on magnesium (**Figure 2**), approximately double that in aluminum, for example; the small grain size also improves ductility.

Magnesium and its alloys are commonly melted in mild steel crucibles because of their low solubility for iron. The melt surface, however, must be protected by a flux or inert atmosphere. Grain refinement of the cast microstructure can be achieved by treating

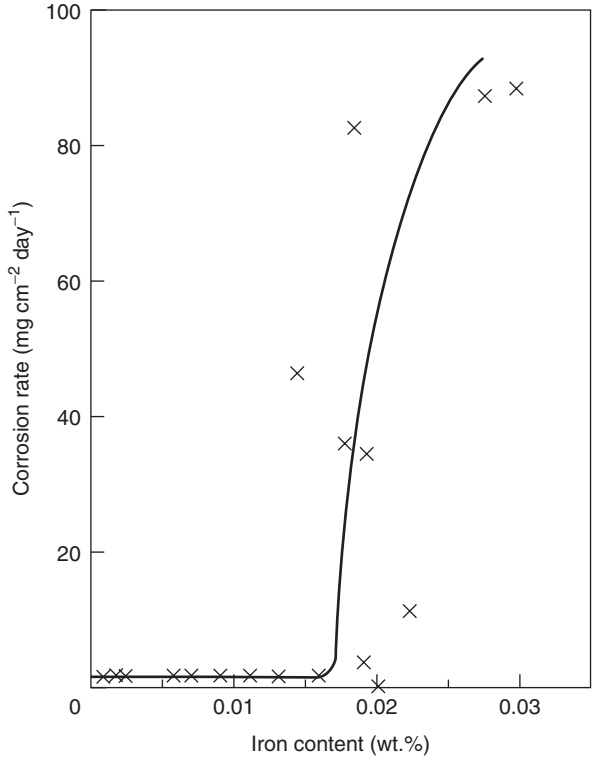


Figure 1 Effect of iron content on corrosion of pure magnesium produced by alternate immersion in 3% NaCl solution. (Reproduced with permission from Hanawalt JD, Nelson CE, and Peloubet JA (1942) *Trans. AIME* 147: 273–299.)

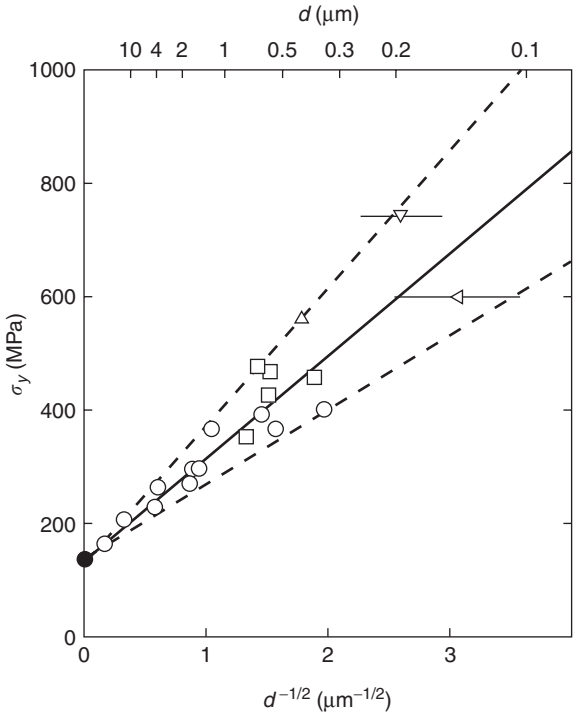


Figure 2 Effect of grain size d on yield strength σ_y of high-strength magnesium alloys compared with AZ91 alloy ingot (•). Key: (○) AZ91, (□) Mg–Al–Zn–Re base, (△) Mg–Y–MM, (◁) Mg–Ca–Al, (▽) Mg–Cu–Y alloy rapidly solidified particulate extrusions. (Reproduced with permission from Jones H (1995) *Key Engineering Materials* 97/98: 1–12; © Trans Tech Publishing Ltd.)

with CCl_6 (when Al is present) or with Zr (when it is not) prior to casting. The nucleant is expected to be Al_4C_3, and Zr itself, in the two cases.

Most magnesium alloy components are produced by high-pressure die casting, using hot chamber machines except for the largest castings (up to 10 kg). They can also be sand-cast if certain precautions are taken, for example, to inhibit reaction with the sand and moisture. Ingots are produced by Direct Chill casting, either for remelting or as feedstock for wrought products.

ASTM alloy designations identify the main alloying elements by letters followed by numbers indicating nominal wt.% of each rounded up to a single digit, e.g., AZ91 = Mg–9 wt.%Al–1 wt.%Zn. Letter codes for the main alloying elements are A = Al, C = Cu, E = rare earths, K = Zr, L = Li, M = Mn, Q = Ag, S = Si, W = Y (yttrium), and Z = Zn.

Casting Alloys

AZ91 (9.5Al, 0.5Zn, 0.3Mn, wt.%) is the most common of the Mg–Al-based die casting alloys, achieving ultimate tensile and yield strengths of 215 and 120 MPa, respectively, in the fully heat treated (T6) condition. High-purity versions (maximum 0.004Fe, 0.001Ni, and 0.015Cu and minimum 0.17Mn, wt.%) are available when corrosion resistance is paramount. Reduced Al-contents, such as in AM60 (6Al, 0.3Mn, wt.%) give increased ductility and fracture toughness, when required. Die cast Mg–Al–Zn alloys are susceptible to hot cracking at intermediate Zn contents as shown in **Figure 3**. Improved creep resistance results from added Si (AS21 and AS41: 2 or 4Si, 0.4 or 0.3Mn and 1Si, wt.%) or rare earths (AE42: 4.0Al, 2.5RE, 0.3Mn, wt.%) in which the formation of Mg_2Si and Al_xRE compounds is more effective in pinning grain boundaries than is the relatively soluble $Mg_{17}Al_{12}$. ZE41 (4.2Zn, 1.3RE, 0.7Zr, wt.%) combines precipitation hardening from Zn and Ce with grain refinement by Zr to allow service up to 150°C. EZ33 (2.7Zn, 3.2RE, 0.7Zr, wt.%) and QE22 (2.5Ag, 2.5Nd, 0.7Zr, wt.%) offer creep resistance or high proof strength up to 250°C, again with grain refinement by Zr. WE54 and WE43 (5 or 4.1Y, 3.25 heavy rare earths, 0.5Zr, wt.%) combine high strength at room and elevated temperatures with better corrosion resistance than other high-temperature magnesium-based alloys. The development of a die-castable magnesium-based alloy that has creep resistance comparable with A380 aluminum casting alloy (Al–8.5 Si–3.5Cu–0.5Mn, wt.%) for automotive engine applications is the focus of major R + D activity in Europe, Japan, and the USA currently (**Figure 4**).

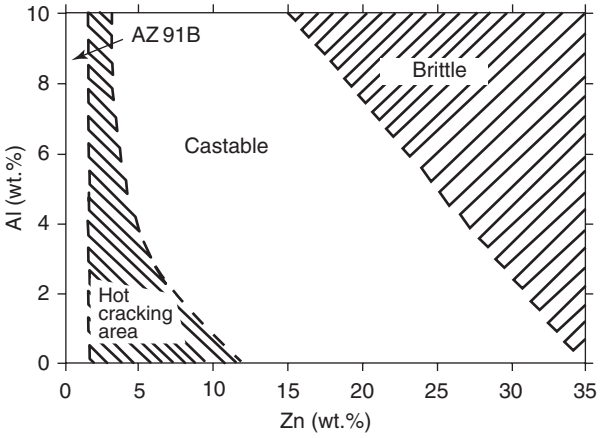


Figure 3 Effect of Al and Zn contents (wt.%) on die castability of Mg–Al–Zn alloys. (Reproduced from Foerster GS (1976) In: *Proceedings of the 33rd Annual Meeting of the International Magnesium Association*, Montreal, pp. 35–39.)

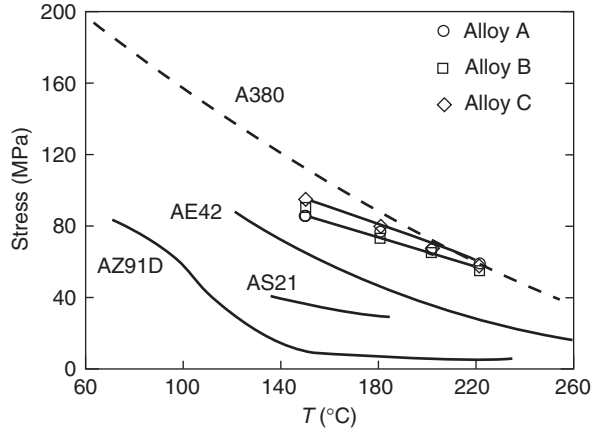


Figure 4 Effect of test temperature *T* on stress for 0.1% creep strain in 100 h for cast alloys based on the Mg–Al system compared with aluminum-based casting alloy A380. Alloys A to C are developmental Mg–Zn–RE alloys. (Reproduced from Villa Fanjul F, Srimanosaowapak S, Mc Nee KR, Greenwood GW, and Jones H (2003) The effect of Nd substitution for CeMM on creep performance of M–2.5mm–0.35Zn–0.3Mn (MEZ) alloy. *Z. Metallkunde* 44: 25.)

Wrought and Powder Metallurgy Alloys

Working of magnesium is carried out at 300–500°C by extrusion, rolling, or forging. Basal planes and $\langle 1\,0\,\bar{1}\,0 \rangle$ directions tend to orient in the direction of working which can result in proof strengths in compression that are 0.5–0.7 of those in tension. This difference is attributable to the occurrence of twinning more readily under compression, and is reduced at small grain sizes.

AZ31 (3Al, 1Zn, 0.3Mn, wt.%) is the most widely used wrought sheet magnesium-based alloy, comparable in proof and tensile strength with cast AZ91, but with high ductility and being weldable. ZK31

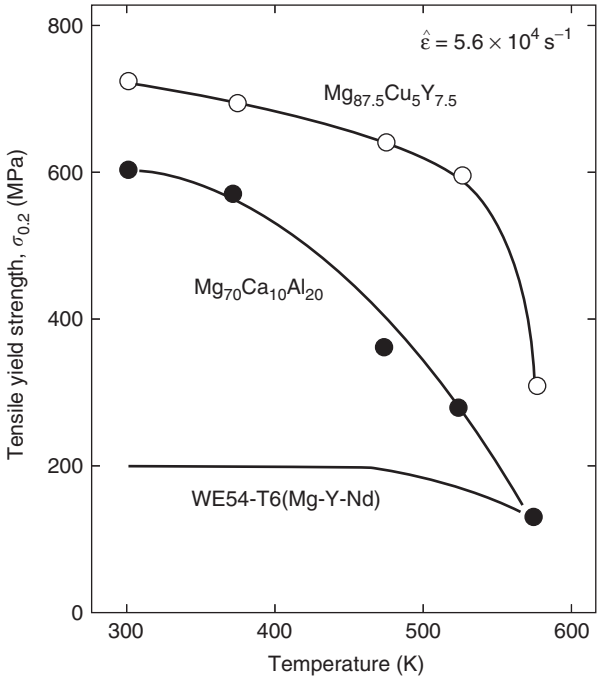


Figure 5 Temperature dependence of tensile yield strength of Mg–5Cu–7.5Y (at.%) and Mg–20Al–10Ca (at.%) rapidly solidified alloy powder extrusions compared with WE54-T6 conventionally processed alloy. (Reproduced from Inoue A and Masumoto T (1993) *Material Science and Engineering A* 173: 1–8, with permission from Elsevier.)

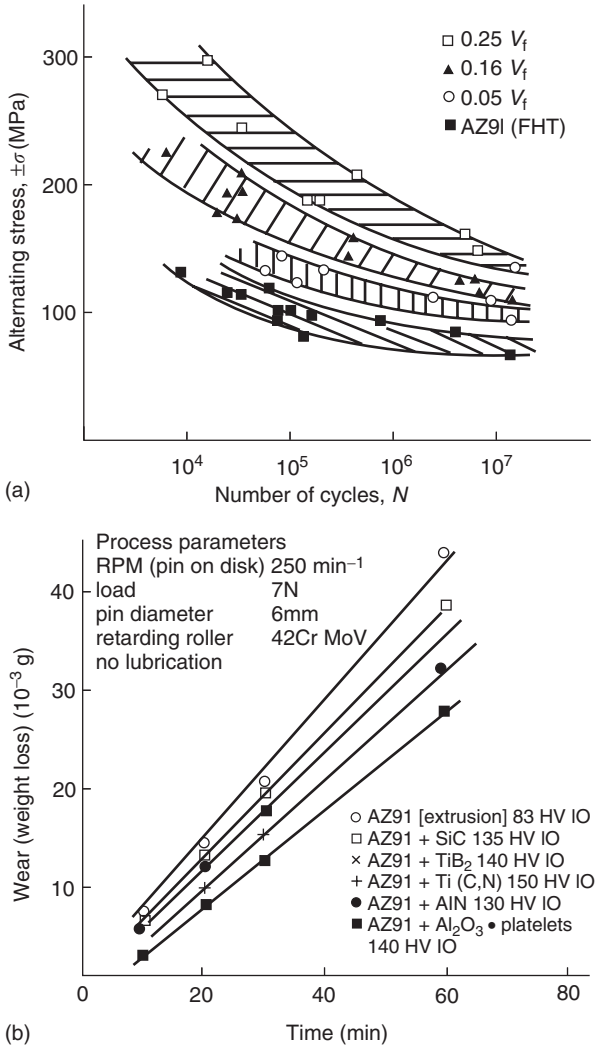


Figure 6 Effects of increasing volume fraction V_f of Saffil (α-Al_2O_3) short fibers and 15 vol.% of various ceramic particulates on (1) fatigue resistance and (2) wear resistance of AZ91 alloy. ((a) Reproduced from Llorca A *et al.* (1991) *Material Science and Engineering A* 135: 247–252, with permission from Elsevier. (b) Reproduced from Schröder J *et al.* (1989) *Proc. ECCM3*, pp. 221–226, with permission from Elsevier.)

(3.0Zn, 0.6Zr, wt.%) is stronger, but at the expense of weldability. LA141 (14Li, 1.2Al, 0.15 minimum Mn, wt.%) has lower strength than either of these but has very low density (1.35 mg m^{-3}) because of the high Li content which confers a duplex hexagonal close-packed/body-centered cubic microstructure. Limitations are poor corrosion resistance and susceptibility to overaging in service which results in excessive creep at relatively low loads. Sheets of all of these alloys can be hot-formed readily at 230–350°C, but cold formability is limited.

AZ61 (6.5Al, 1Zn, 0.3Mn, wt.%) and ZK61 (6Zn, 0.8Zr, wt.%) are examples of extrusion alloys, the latter giving a tensile strength exceeding 300 MPa in the T5 condition, at the expense of weldability. Tensile strengths can be raised to >500 MPa in EA55RS (5Al, 5Zn, 5Nd, wt.%) by rapid solidification processing, which involves planar flow casting to sheet <0.1 mm thick, followed by pulverization and then consolidation by vacuum hot pressing to provide a bulk feedstock for extrusion, forging, or rolling. The very high strengths obtained are mainly attributable to ultrafine grain size (down to 0.3 μm) stabilized by finely dispersed rare-earth aluminides. Strengths as high as 740 MPa have been developed in partially or completely devitrified initially glassy Mg–Cu–Y or Mg–Al–Ca rapidly solidified alloys (**Figure 5**). Some of these compositions exhibit the phenomenon of "bulk metallic glass formation," where section thicknesses of up to several millimeters can be produced in the glassy state directly by chill casting.

Finally, an increased Young's modulus, fatigue, and wear resistance have been achieved in metal matrix composites based on alloys such as AZ91 with ceramic reinforcements (**Figure 6**).

Applications

Figure 7 summarizes the consumption of extracted magnesium in the Western world in 1999. This shows that some 43% was used as an alloying addition to

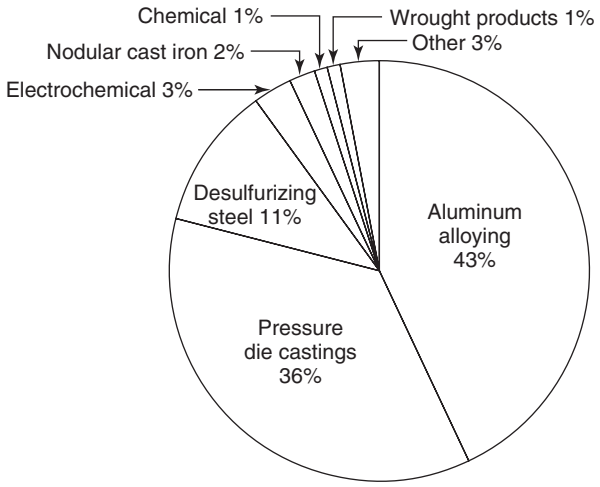


Figure 7 Breakdown of usage of magnesium in the Western world in 1999.

aluminum followed by 36% for magnesium-based alloy die castings, and 13% to desulfurize liquid steel and to nodularize graphitic cast irons, with smaller percentages for electrochemical, chemical, wrought alloy and other purposes. A comparison with the corresponding data for 1991 shows an increase in total shipments from 250 to 380 ktons (kt) (involving a 7% annual growth rate since 1993) due mainly to an increase in demand for magnesium alloy die castings from 38.6 kt (15%) to 133.4 kt (36%), mostly for applications in the automotive industry (driven by weight reduction requirements), but also for components of computers, laptops, cell phones, etc. Typical automotive applications include cylinder head covers, transfer case and transmission housings, seat frames, brake and clutch pedal brackets, steering wheel cores, lock housings, inlet manifolds, cross car beams, and instrument panel substrates. Predictions are that the annual demand for magnesium alloy die castings should exceed 200 kt by the end of 2004. A further growth in general usage of magnesium alloys is likely to depend on (1) expanding the range of alloys with good creep resistance to higher temperatures and even better corrosion resistance, (2) further development of new casting techniques such as thixoforming and squeeze casting, (3) improving the properties of wrought magnesium alloys and their formability, including superplastic forming and diffusion bonding, and (4) continued development of advanced magnesium-based materials, and their metal matrix composites, by such techniques as mechanical alloying, rapid solidification, and other powder technologies.

See also: Electron Gas (Theory); Mechanical Properties: Strengthening Mechanisms in Metals; Phase Transformation; Phases and Phase Equilibrium; Recovery, Recrystallization, and Grain Growth; Solidification: Models and Simulations.

PACS: 81.05.Bx; 71.20.Dg

Further Reading

Avedesian MM and Baker H (eds.) (1999) *Magnesium and Magnesium Alloys. Speciality Handbook*. Materials Park, OH: ASM International.

Neite G, Kubota K, Higashi K, and Hehmann F (1996) Magnesium-Based alloys. In: Matucha KH (ed.) *Materials Science and Technology*, vol. 8, Weinheim: VCH.

Polmear IJ (1995) *Light Alloys*, 3rd edn., ch. 5. London: Arnold (see also Material Science and Technology (1994), 10: 1–16).

Alloys: Overview

M Baricco, Università di Torino, Torino, Italy

Introduction

Metallic alloys constitute a fundamental class of engineering materials. Their properties are related to electronic and crystallographic structures, as well as to microstructures. In this article, the fundamental aspects of metallic alloys are described. For details on single classes of metallic alloys, the reader is addressed to specific sections.

Crystal Structures in Metals and Alloys

Chemical Bonds and Crystal Structures

Material properties are related to crystal structure, which depends not only on the chemistry of constituent atoms, but more strongly on the nature of chemical bonds. In fact, the electronic configurations of different atoms determine different kinds of chemical bonds. Though the concept of chemical bond has been developed for molecules, it can also be applied to solid state chemistry. The types of chemical bonds involved in solids may be distinguished as follows: ionic, covalent, metallic, van der Waals, and hydrogen bonds. Metals and alloys are mainly based on

the metallic bond, where valence electrons are delocalized over the whole crystal, even if some contribution from covalent and ionic bonds may be observed in special cases. Alkali metals are purely metallic bonded, whereas metals close to metalloids, such as zinc, have increasing covalent bond contributions.

Geometrical factors also play a crucial role in defining the crystal structure of metals and alloys. Because of the isotropic nature of the metallic bond, the structure of metals may be described in terms of hard-sphere packing. In the case of pure metals, noninterpenetrating equal spheres are interconnected so that the volume is filled as much as possible. Considering a simple layered structure of spheres, hexagonal close-packed (h.c.p.) and face-centred cubic (f.c.c.) crystal structures may be easily built up. On the other hand, considering a slightly lower packing factor, a body-centred cubic (b.c.c.) structure is formed. The crystal structure of metals is often changed on heating (allotropy), as observed in iron, which shows a transition from b.c.c. to f.c.c. at 910°C and a further transition from f.c.c. to b.c.c. at 1400°C before melting at 1539°C. The crystal structures of metals follow the hard-sphere packing as long as pure geometrical factors are concerned. As soon as covalent bonding contributions and electronic effects become important, less dense and more complicated structures are observed, such as in Sn. Similarly, different packing factors are observed in metals with a similar crystal structure. For the same reason, metals showing an h.c.p. structure may have a different ratio between lattice constants (c/a), which turns out to be equal to 1.63 and 1.86 for Mg and Zn, respectively.

In metallic alloys, the constituent elements may be simply mixed in a solid solution or may form an intermediate phase. In the case of alloys, the difference in electronegativity between constituent atoms becomes the main factor in defining the nature of the chemical bond. In the case of a high difference in electronegativity, an ionic contribution to the chemical bond may be observed. As a consequence, rather complicated crystal structures may be formed. Crystallographic parameters relative to various alloys are available in structural databases.

Solid Solutions

Solid solutions may be distinguished as substitutional, when the volume of constituent elements is similar, and interstitial, when a volume difference (size factor) higher than $\sim 15\%$ is observed. A typical example of substitutional solid solution of Zn in Cu is observed in brass, whereas an interstitial solid solution of C in f.c.c. Fe is observed in austenite. Complete solid solubility may be obtained only in alloys when components have the same crystal structure and the size factor is lower than $\sim 8\%$. Au–Ni and Ag–Au–Pt are examples of systems showing complete solid solubility.

Sometimes the distribution of atoms in a solid solution deviates from randomness. When similar atoms group themselves preferentially, a clustering effect is observed. On the other hand, when an atom is preferentially surrounded by different atoms, the solid solution is said to show a short-range ordering.

By means of rapid solidification techniques, the disordered structure of the liquid phase may be retained in the solid, giving a metallic glass. The formation of amorphous alloys is due to the strong undercooling of the liquid, which gives an increase in the viscosity up to the glass transition.

Intermediate Phases

When the limit of solid solubility is exceeded on alloying, a second phase is formed. It may be the primary solid solution of the alloying elements, such as in simple eutectic systems, but more often it appears as an intermediate phase. When only metallic components are present in the system, the intermediate phase is called an intermetallic compound.

The simplest intermediate phase is due to long-range ordering of components in the solid solution (ordered phases). In this case, crystallographic positions are preferentially occupied by a specific element so that an ordering parameter may be defined. Different sublattices constituted by single components may be defined in the structure, as evidenced by diffraction techniques. The ordering parameter changes continuously in a limited temperature range, according to the second-order thermodynamic transition. Ordering of phases is effective in improving mechanical properties and it is fundamental in superalloys.

When the atoms constituting the intermediate phase show a particular value of the valency electron concentration (e/a), electron phases are formed. For instance, 50% of Zn atoms in a Cu matrix give an electron concentration of 1.50, where a b.c.c. β-phase appears. The complex cubic γ-phase and h.c.p. ε-phase are formed for electron concentrations equal to 1.62 and 1.75, respectively. The formation of electron phases is based on the empirical Hume–Rothery's observations, later explained in terms of density of states for valency electrons. Electron phases show a metallic behavior and exist over a range of compositions, such as in brass and bronze.

When the intermetallic phase contains a large number of vacant lattice sites, a defect phase may be

formed in a limited composition range. For instance, in the ordered NiAl phase, the presence of vacancies in the Ni sublattice leads to an increase in the Al-concentration up to ~54%. Vacancy phases play an important role in superalloys.

Very common intermetallic compounds are known as Laves phases. They have a general formula AB_2 and may have a cubic structure ($MgCu_2$) or hexagonal structure ($MgZn_2$ and $MgNi_2$). For the Laves phases, the radii of components have the ratio $r_A/r_B = 1.225$, so that a very high packing density is achieved. Laves phases have recently received attention because of high hydrogen absorption performances, useful in modern batteries.

The interaction of transition metals (M) with small nonmetallic elements (X) generally leads to the formation of interstitial compounds, such as hydrides, borides, nitrides, and carbides. In these compounds, the metal atoms form the matrix and the alloying elements are in interstitial positions. If the ratio between atomic radii r_X/r_M is lower than 0.59, simple compound structures are obtained, where the transition metal forms an f.c.c. or h.c.p. lattice. For cementite (Fe_3C) $r_X/r_M = 0.61$, so that a complex orthorhombic structure is formed. Carbides and nitrides are very useful for strengthening in alloyed steels.

When the components of the intermediate phase have a strong difference in electronegativity, ionic contributions to the chemical bond lead to the formation of valency compounds, which are essentially nonmetallic. They have a fixed stoichiometry and are often formed between transition metals and metalloids. Silicides are an example of such a type of phase.

In recent years, rather complex intermediate phases have been continuously discovered. Frank–Kasper phases show a huge number of atoms in the unary cell. Quasicrystalline phases show a fivefold symmetry, which was considered forbidden according to the classical crystallographic rules.

Phase Stability and Transformations in Metallic Alloys

Thermodynamics and Kinetics of Phase Transformations

Phase stabilities and transformations in metallic alloys are ruled by thermodynamic and kinetic factors. As schematically shown in **Figure 1**, in order to transform a phase (α) into a phase (β), a gain in free energy is necessary. This free energy difference ($\Delta G = G_\beta - G_\alpha$) represents the driving force for the phase transformation. The phase transformation

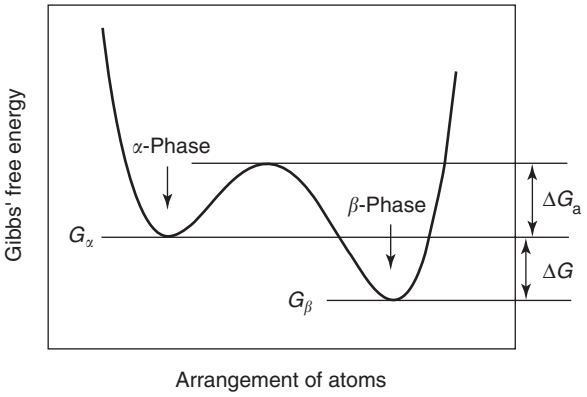


Figure 1 Free energy as a function of the arrangement of atoms in a phase transformation.

becomes possible if the activation barrier (ΔG_a) is overcome and the new phase is formed at a reasonable rate.

The kinetics of a phase transformation is strongly related to the mechanism. In fact, when the crystallography and/or the chemical composition of the two phases are significantly different, a reconstructive transformation occurs and a motion of single atoms is necessary. The whole kinetics of phase transformation may be controlled either by the diffusivity or by interface migration. A typical example is given by perlite formation in eutectoidic steels. On the contrary, a cooperative movement of neighboring atoms occurs in displacive transformations, which are diffusionless and generally very fast, because no activation barrier is present. The formation of martensite in steel by quenching represents an example of such a kind of phase transformation.

The kinetics of phase transformation in metallic alloys is often described in terms of time–temperature–transformation curves (TTT), which are a graphical representation of the time necessary to start and to complete isothermally the phase transformation in a specific temperature range. In the case of cooling treatments, continuous-cooling-transformation curves (CCT) are more appropriate. In order to draw thermal treatments in steels, databases of TTT and CCT diagrams are available. The kinetics of phase transformations may be described in terms of the Avrami equation $y = 1 - \exp(-kt)^m$, where y is the transformed fraction, t is the time, k is a temperature-dependent rate constant, and m is an empirical parameter ranging from 0.5 up to 4, which depends on the type of transformation.

Free Energy and Phase Diagrams

The free-energy of a phase depends on several factors, such as composition, temperature, pressure,

strain, surface, and interfacial energy. In order to define the equilibrium condition for a generic multicomponent system, the free energy of all phases must be known as a function of composition, temperature, and pressure. The thermodynamic equilibrium is reached when the chemical potential of all elements is the same in all phases (common tangent rule). The number of equilibrium phases is defined by the well-known Gibbs' phase rule and the amount of each phase can be calculated according to the lever rule. The description of the equilibrium phases is generally reported as composition–temperature plots, known as phase diagrams.

In the case of a simple binary metallic system, the equilibrium thermodynamics may be estimated by means of suitable models. For instance, the regular solution model considers only solution phases, where the free energy is described by means of an ideal entropy term and a temperature-independent term for the enthalpy of mixing. This term considers the chemical interactions between the constituent elements and it turns out positive in the case of repulsion and negative in the case of attraction. When the enthalpy of mixing becomes zero, the solution is ideal. On the basis of the regular solution model, several binary phase diagrams may be calculated, giving solubility or immiscibility either in the liquid or in the solid phases. Examples of phase diagrams calculated with the regular solution model are reported in **Figure 2**.

As an example, a full miscibility for the liquid and solid phases is shown in Cu–Ni, simple eutectics are observed in Cu–Ag and in Al–Si, and a single peritectic is observed in Ag–Pt. Intermediate phases are shown in the phase diagrams as single lines (line compounds) or as compounds with limited solubility, such as in bronze and brass. The equilibrium conditions for a ternary system may be represented by the Gibbs' triangle, where isothermal sections of the phase diagrams are reported as a function of the composition of components. The method for plotting compositions in a ternary phase diagram is shown in **Figure 3**. For multicomponent systems, only sections of the phase diagram are usually reported (pseudo-binary).

Phase diagrams are generally determined experimentally. Physical properties of alloys with different compositions are followed as a function of temperature, giving an experimental evidence of transition temperatures and equilibrium points. Calorimetric techniques are used for the determination of thermophysical quantities. Modern computing techniques allow the description of the thermodynamics of multicomponent systems of industrial interest. For instance, by means of the CALPHAD (CALculation of

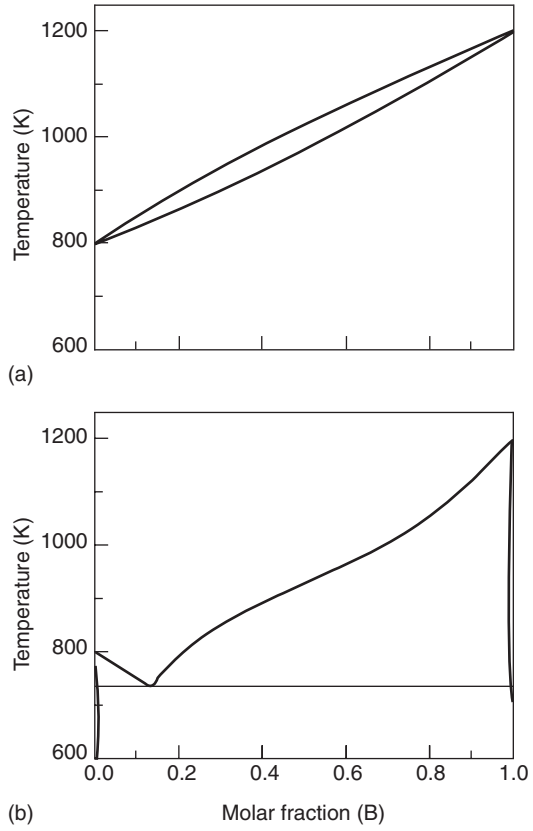


Figure 2 Examples of binary phase diagrams calculated according to the regular solution model. (a) Ideal solution; (b) Interaction parameter equal to $10^4\,J\,mol^{-1}$ and $3 \times 10^4\,J\,mol^{-1}$ for the liquid and solid phases, respectively.

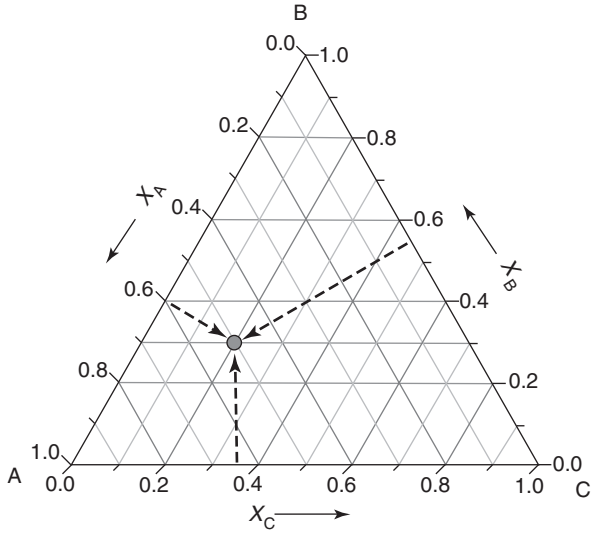


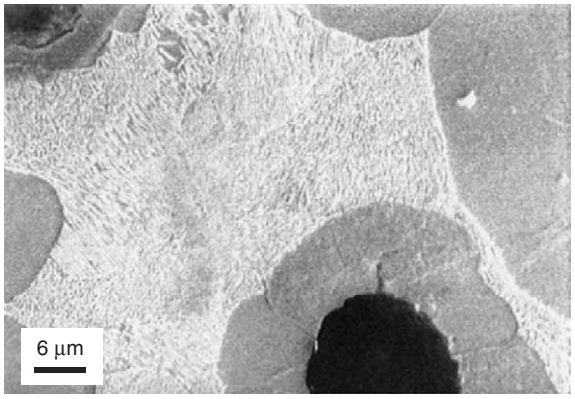
Figure 3 Representation of compositions for ternary alloys. The vertices represent 100% of A, B, or C. The binary systems are represented in the boundaries of the triangle. Ternary alloys are reported in the middle of the triangle. The distances of the points from the boundaries represent the relative proportions of the components indicated in the opposite vertex of the triangle. The point indicated in the figure indicates the composition with 20% of C, 50% of A, and 30% of B ($A_5B_3C_2$).

PHAse Diagrams) method, the free energy of all phases is described as a function of temperature and composition through parameters obtained by an assessment of experimental data. Databases of parameters are available for metallic alloys of industrial interest, allowing the calculation of equilibrium conditions, phase diagrams, and thermodynamic quantities.

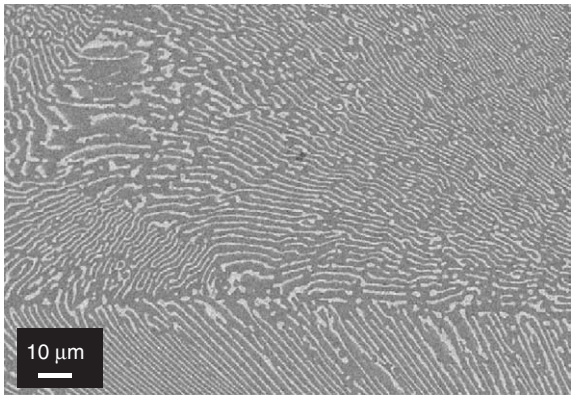
Microstructure of Metallic Alloys

Properties of metallic alloys are often related to the microstructure, which depends on phase transformations and mechanical treatments involved in the processing. Examples of microstructures are shown in **Figure 4**.

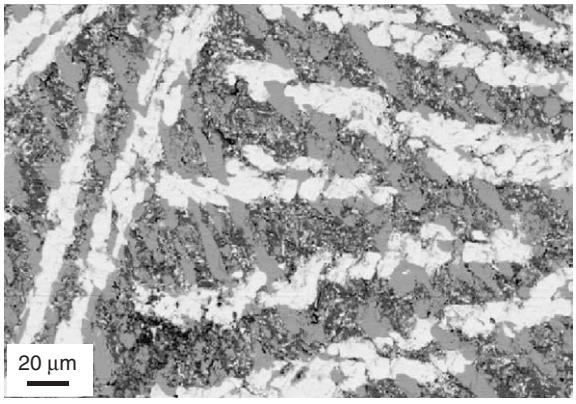
Reconstructive phase transformations usually proceed by nucleation and growth mechanisms. In the case of a single phase, the grain size determines the grain boundary area, which affects the strength through the Hall–Petch relationship. Grain size is ruled by solidification rate, plastic deformation and recrystallization treatments. Eutectic and eutectoid transformations give a lamellar microstructure. The significant composition difference between the parent and the product phases needs atomic diffusion, so the interlamellar spacing is related to the temperature of the phase transformation. As an example, the temperature of thermal treatments in steels selects the final microstructure, ranging from coarse perlite to fine bainite. Similar fine microstructures may be also obtained by tempering of martensite, previously obtained by quenching. For light alloys, precipitation hardening is related to the microstructure. Metastable phases with interfaces coherent to the matrix may be produced by suitable aging treatments of supersaturated solid solutions obtained by quenching. Similar microstructural effects are observed in spinodal decompositions, where composition fluctuations occur up to the gradient of concentration (uphill diffusion). In recent years, very fine microstructures may be obtained by suitable processing, giving nanostructured alloys. As an example, ball milling of metallic powders, electrodeposition or severe plastic deformation of ingots, may give a grain size down to few tenths of nanometers, leading to improved properties with respect to coarse-grain materials.



(a)



(b)



(c)

Figure 4 Examples of microstructures. (a) Microstructure in spheroidal cast iron: black zone is the nodular graphite, surrounding gray areas is ferrite, and the light area is perlite. (b) Microstructure in eutectic $Pb_{26}Sn_{74}$: white areas are Pb-rich solid solution and dark areas are the Sn-rich solid solution. (c) Microstructure in as-cast $Al_{87}Ni_7Ce_6$: white area is $Al_{11}Ce_3$; gray area is Al_3Ni; dark area is f.c.c. Al. (Courtesy of Rizzi P, University of Torino, Italy.)

Properties of Metallic Alloys

Mechanical Properties

Synthesis and processing of metallic alloys is mainly aimed at mechanical applications. In fact, metallic materials show a very broad range of mechanical properties, which may be often modified by suitable thermomechanical treatments. Basic mechanical properties of metallic alloys are generally determined by standard tensile tests, leading to the classical stress–strain curves shown in **Figure 5**.

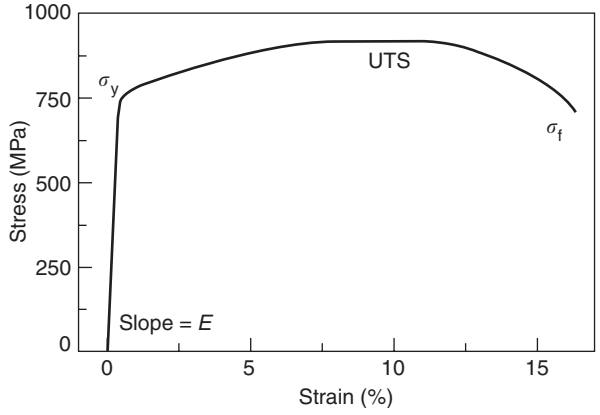


Figure 5 Example of a stress–strain curve for steel (0.4% C, 4% Cr). Basic mechanical parameters are indicated: E, Young modulus; σ_y, yield stress, UTS, ultimate tensile stress, σ_f, fracture stress.

Elastic properties are obtained from the elastic range of tensile tests and they are mainly related to the composition of metallic alloys. The Young modulus (E) ranges from $\sim$40 GPa from Mg-based alloys up to more than 400 GPa for W and Os. The Young modulus for Al-based alloys is $\sim$70 GPa, for Cu-based and Ti-based alloys is $\sim$100 GPa, and for ferrous alloys is $\sim$200 GPa. Metallic alloys have a shear modulus (G) $\sim$3/8 of E and a Poisson ratio (ν) of 0.33. After thermomechanical treatments and changes of the microstructure, the elastic properties remain rather constant, so that they may be considered an intrinsic property of the alloys. The Young modulus of metallic alloys generally decreases as a function of temperature ($dE/dT < 0$). For special applications, such as precision springs, specific compositions have been developed showing $dE/dT \approx 0$ (Elinvar).

The strength and elongation properties may be obtained from the plastic range of the stress–strain curve. The yield stress (σ_y) represents the transition from the elastic to the plastic regime. It depends on the dislocation mobility, so it can be significantly modified by thermomechanical treatments. The strengthening mechanisms in metallic alloys are related to the hindering of dislocation movements. This goal may be obtained by alloying, because the presence of alloying atoms may generate a stress field. In work-hardening processes, the presence of a high-dislocation density leads to an entanglement between crossing dislocations. For precipitation-hardening, widely used in light alloys, the presence of precipitates with interfaces coherent with the matrix produces efficient stress fields against dislocation movements. With the reduction of the grain size, a dislocation pile-up occurs at the grain boundaries and a strengthening in the alloy is observed. For superalloys, a precipitation of the ordered phase may be induced in the disordered matrix, so that dislocations may be blocked at the antiphase boundaries.

Often the yielding point cannot be identified easily, so the stress corresponding to a plastic deformation of 0.2% ($\sigma_{0.2\%y}$) is considered. In the case of annealed low-carbon steels, a nonuniform yield phenomenon may be observed, with the resulting production of Luder's lines at the surface. In this case, an upper-(σ_{uy}) and a lower-(σ_{ly}) yield stress are identified. As long as a uniform plastic deformation acts in the metallic alloys, a hardening effect is observed, according to the general equation $\sigma = K\varepsilon^n$, where σ is the true stress, ε is the true strain, K is the strength coefficient, and n is the strain-hardening exponent. The last parameter measures the rate at which an alloy becomes strengthened as a result of plastic deformation. When the plastic deformation becomes nonuniform, a neck begins to form in the test specimen and an apparent maximum is observed in the nominal stress–strain curve. Necking phenomena are very dangerous for metallic alloys aimed at high-plastic deformation, such as in deep-drawing. Certain metallic materials, under special deformation conditions, behave as superplastic alloys, so they resist necking and show uniform elongations up to 1000 times the normal amount.

Fracture in metallic alloys may be ductile or fragile. In ductile fracture, a significant plastic deformation acts before rupture and a characteristic cup–cone fracture is observed at the surface of the test specimen. Brittle fracture is more dangerous, because it happens suddenly without prior evident plastic deformation. Fracture may be intergranular when it follows the grain boundaries or intragranular when it crosses the interior of the grains. Fracture by cleavage is often observed instead of by shear. A ductile-to-brittle transition may be observed in a limited temperature range by standard impact tests (Charpy type). For instance, for ferritic steels it occurs below room temperature and may be crucial for low-temperature applications, so that austenitic steel is preferred. The tendency for crack propagation is defined by the fracture toughness (K_c). It is obtained from a specific test, where increasing stresses are applied to a pre-cracked specimen. Depending on the mode of loading, different fracture toughness parameters are defined (I, II, and III).

Fracture is generally pictured as consisting of two stages: crack formation and growth. Crack formation may be due to processing or use of the alloy. It generally happens at the surface and it is often related to repeated stresses in the elastic range, a phenomenon known as fatigue. Due to this reason, for

metallic alloys it is necessary to define an endurance limit, which is about one half of the tensile strength: below this value, the alloy can withstand an unlimited number of stress cycles without fracture. Unfortunately, some alloys (e.g., Al-based alloys) do not show this limit, so that the fatigue strength must be defined for a given number of cycles.

Hardness of metallic alloys is determined by specific tests, which are based on the resistance to penetration. A commercial hardness tester forces a small sphere, pyramid, or cone into the body of the alloy by means of an applied load. Hardness numbers may be roughly related to tensile strength. When the applied load is small, the hardness may be determined for single phases evidenced in the microstructure of the alloy. For instance, a resolution down to 200 μm may be obtained by the microhardness Vickers test.

For applications at high temperature, the creep phenomena may occur in metallic alloys. They are due to the occurrence of a plastic strain when the alloy is stressed in the elastic regime. Creep is related to dislocations climbing induced by the high temperature. For moderate stresses, grain boundary sliding may even occur, because of atomic diffusion. Creep may be avoided in superalloys, where precipitates of ordered phases hinder the dislocation movements.

Electrical and Magnetic Properties

Electrical properties of metallic alloys are closely linked to the electronic structure, and the conduction of electrons depends on partially filled bands. According to Mathiessen's rule, the electrical resistivity of alloys increases linearly with temperature near room temperature and above, although the behavior at very low temperature is more complex. Electrical resistivity is strongly enhanced by the presence of foreign atoms in a solid solution in the metal matrix. As an example, the effect of the addition of 0.1 wt.% of different atoms in Cu is shown in **Figure 6**.

It is clear that the effect is strongly dependent on the electronic configuration of the foreign atom. Oxygen has a very big effect and, for the production of copper wires, oxygen-free material is necessary. Cold rolling increases resistivity only slightly and it is often used as a means for strengthening alloys for electrical conductors. Metallic alloys behave as superconductors only at very low temperatures and they have been fully replaced by high-temperature oxide superconductors for practical applications. Thermoelectric effects of metallic alloys are widely used for devices such as thermocouples, thermoresistors, and Peltier junctions.

The magnetic properties of metallic alloys are of fundamental interest and have several industrial

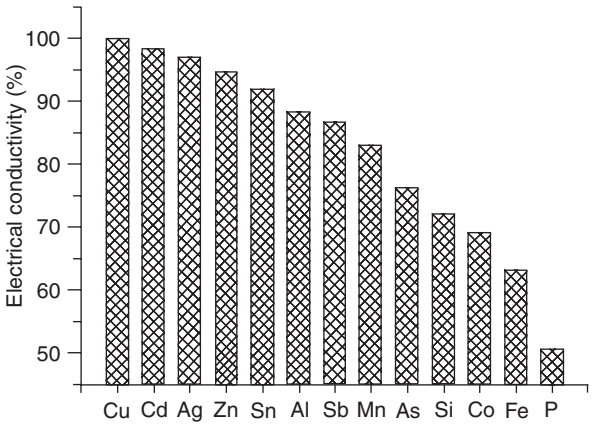


Figure 6 Effect of the addition of 0.1 wt.% of various elements in solid solution on the electrical conductivity of copper.

applications. The primary applications arise in ferromagnetic alloys, based on Fe, Ni, Co, and some rare-earth metals. On the basis of the values of magnetization induced in the material when subjected to a magnetizing field, represented graphically by the hysteresis loop, ferromagnetic metallic alloys are distinguished in soft and hard magnets.

Soft magnetic alloys are characterized by high magnetic permeability, low coercive field, and low core losses. For these materials, a high saturation magnetization is generally desirable. These materials are suitable for electrical motors, transformers, and relays, because of the fast response of magnetization to the applied magnetic field. In Fe, there is a strong relationship between permeability and crystallographic directions, so that the magnetic flux runs better along the (1 0 0) direction (i.e., the edge of the cubic unary cell), giving low-energy losses. For this reason, commercial steel containing 3% Si is processed so that a preferred orientation of grains is obtained (Goss texture). Amorphous alloys show still lower power losses and they find applications in small-sized transformers. Magnetic shields may be obtained using materials with very high magnetic permeability, such as Fe–Ni alloys.

On the contrary, hard magnets have a high coercive field and a high saturation magnetization. A parameter representative of the magnetic energy that can be stored per unit of mass is given by the maximum external energy, $(BH)_{max}$, which is calculated on the demagnetizing portion of the hysteresis loop. These materials are mainly used as permanent magnets, often used to convert electrical energy to mechanical motion (electrical engines) or to convert mechanical motion to electrical energy (microphones). Defects able to hinder the movements of dislocations are usually also able to hinder the movement of magnetic domains, increasing the

magnetic hardness of the material. So, high carbon and alloyed steels are often used as permanent magnets. In recent years, new metallic alloys, characterized by a strong magnetic anisotropy in the unary cell, have been developed. These materials contain rare earths as alloying element, such as in $SmCo_5$ and in $Fe_{14}Nd_2B$.

Chemical Properties

Metallic alloys are not chemically stable. In fact, metallurgical processes devoted to the production of metals from their oxides are often counter-balanced by corrosion and oxidation, which progressively transform the metallic alloys in more stable compounds. Corrosion is a wet phenomenon, which is possible only in the presence of humidity. The corrosion resistance of metallic alloys is generally related to the properties of parent elements, according to the electromotive series. A control of the microstructure of the alloys may change the mechanism of corrosion, which should not be localized but uniformly distributed on the whole surface. In several cases, the addition of elements prone to passivity (Cr, Al) is used to increase the corrosion resistance of alloys, such as in the case of stainless steel. Corrosion might be also avoided by deposition of thin layers of a metal on the surface of the alloy, such as in zinc coating and tinning of steel. Oxidation is a high-temperature process, due to the reaction of the metallic alloy with the atmosphere. At the very beginning of the reaction, a thin layer of the oxide is formed at the surface. If the volume of the oxide is at least as great as the volume of the metal from which it formed, the oxide is protective and oxidation proceeds slowly. On the contrary, if the volume of oxide is less than this amount, the layer is not continuous and it is less effective in preventing the access of oxygen to the alloy surface, so that the reaction rate becomes high. Often the simultaneous formation of two or more oxide layers is observed, such as in steel, where a variable amount of FeO, Fe_3O_4, and Fe_2O_3 layers are formed at different temperatures.

Materials Selection

For specific applications, it is a combination of material properties (material index) that characterizes the performance. Properties of alloys are available in databases, so that merit indices, combined with Ashby's charts, allow optimization of the material selection process.

See also: Alloys: Aluminum; Alloys: Copper; Alloys: Iron; Alloys: Magnesium; Alloys: Titanium; Diffusionless Transformations; Intermetallic Compounds, Electronic States of; Ionic Bonding and Crystals; Mechanical Properties: Anelasticity; Mechanical Properties: Creep; Mechanical Properties: Elastic Behavior; Mechanical Properties: Fatigue; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Mechanical Properties: Tensile Properties; Metallic Bonding and Crystals; Phase Transformation; Phases and Phase Equilibrium; Recovery, Recrystallization, and Grain Growth; Thermodynamic Properties, General; Thin Films, Mechanical Behavior of; van der Waals Bonding and Inert Gases.

PACS: 61.66.Dk; 61.82.Bg; 71.22.+i; 71.55.Ak; 73.61.At; 74.70.Ad; 75.20.En; 75.47.Np; 78.30.Er; 78.66.Bz; 81.05.Bx; 81.30.Bx; 81.40.Cd; 81.40.Ef; 81.40.Rs

Further Reading

Ashby MF (1999) *Materials Selection in Mechanical Design*. Oxford: Butterworth-Heinemann.

Christian JW (1975) *The Theory of Transformations in Metals and Alloys*. Oxford: Pergamon Press.

Guy A and Hren JJ (1974) *Elements of Physical Metallurgy*. Reading: Addison-Wesley.

Haasen P (1996) *Physical Metallurgy*. Cambridge: Cambridge University Press.

Saunders N and Miodownik AP (1998) *CALPHAD: A Comprehensive Guide*. New York: Elsevier Science.

(1998) *Metals Handbook* Materials Park: ASM International.

Nomenclature

$(BH)_{max}$	maximum external energy ($J\,m^{-3}$)
E	Young modulus (GPa)
G	free energy ($J\,mol^{-1}$), shear modulus (GPa)
k	rate constant (s^{-1})
K	strength coefficient (MPa)
m	Avrami exponent (dimensionless)
n	strain-hardening exponent (dimensionless)
r	atomic radius (m)
t	time (s)
T	temperature (°C)
UTS	ultimate tensile strength (MPa)
ν	Poisson ratio (dimensionless)
y	transformed fraction (dimensionless)
ε	strain (dimensionless)
σ	stress (MPa)

Alloys: Titanium

T R Bieler, R M Trevino, and L Zeng,
Michigan State University, East Lansing, MI, USA

Introduction

Titanium and titanium alloys offer a unique combination of physical and mechanical properties that makes titanium an excellent choice for applications that require high strength at temperatures below 550°C, good stiffness (104–133 GPa), toughness, and corrosion resistance. Another desirable characteristic of titanium is its relatively low density of 4.54 g cm^{-3}, which is between iron (7.87 g cm^{-3}) and aluminum (2.70 g cm^{-3}). The high specific strength along with good corrosion resistance makes titanium an enabling material in the aerospace industry (e.g., airframe components, rotors, and compression blades in jet engines), and it is important in the growing biomedical engineering field due to its biocompatibility with living tissues. Titanium alloys have historically been too expensive for the auto industry, but titanium alloys have been recently used for exhaust systems, valves, and springs (its high-yield strength and moderate modulus value gives titanium alloys tremendous elastic resilience). Titanium aluminide intermetallic alloys are likely to become important in the future, due to their lower density and superior high-temperature capabilities.

Titanium ranks ninth as the most plentiful elements, and is the fourth most abundant structural metal in the Earth's crust exceeded only by aluminum, iron, and magnesium. Despite this abundance, the difficulty in processing titanium makes it expensive (~$50/kg for Ti, compared to $4/kg for Al and $1/kg for Fe); yet it is a strategic metal due to which its corrosion resistance and specific strength outweigh its high cost. The cost is likely to drop in the next decade, as new processing methods that reduce the energy requirements and allow continuous processing are scaled up.

The crystallography is considered first, to establish how phase transformations between low- and high-temperature crystal structures occur. From this, the rationale and the complexity surrounding alloying and processing strategies can be appreciated. Finally, a few problematic issues associated with titanium alloys are noted.

Crystallographic Structure of Phases in Titanium Alloys

Titanium exists in two allotropic crystal forms, α, which is a hexagonal structure with a *c*/*a* ratio of ~1.59 (slightly squashed compared to the close packed ratio of 1.63), and β, which has the body-centered cubic (b.c.c.) crystal structure. In pure titanium, the α-phase is stable up to the β transus temperature (β_T), 883°C. The α-phase is transformed upon heating above 883°C to the b.c.c. β-phase. The idealized orientation relationship between α (thin lines) and β (dashed lines) is illustrated in **Figure 1**. However, the actual lattice spacings for α and β do not match perfectly. **Figure 2** shows how the atomic positions of these phases are related to each other based on a common origin using lattice constants from a Ti–6Al–4 V alloy. With heating, the $\alpha \rightarrow \beta$ transformation requires lattice strains of ~10% expansion along $[2\,\bar{1}\,\bar{1}\,0]_\alpha$ to become a $[1\,0\,0]_\beta$ direction, ~10% contraction along the $[0\,1\,\bar{1}\,0]_\alpha$ to become $[0\,1\,\bar{1}]_\beta$, and ~1% contraction along $[0\,0\,0\,1]_\alpha$ that becomes $[0\,1\,\bar{1}]_\beta$. Due to this imperfect geometrical relationship, one of the two $\langle 1\,1\,1 \rangle_\beta$ directions in the {1 1 0} plane will be rotated ~5° to align itself with one of the $\langle \bar{2}\,1\,1\,0 \rangle_\alpha$ directions, leading to six possible variants of β from one α orientation, as shown in **Figure 3a**. This means that two β variants can be misoriented from another by ~10°. Thus, β variants arising from one parent α crystal can be misoriented from each other by ~10°, 50°, 60°, or 70° rotation about a $\langle 0\,1\,1 \rangle_\beta$ axis (a prior $\langle 0\,0\,0\,1 \rangle_\alpha$ axis).

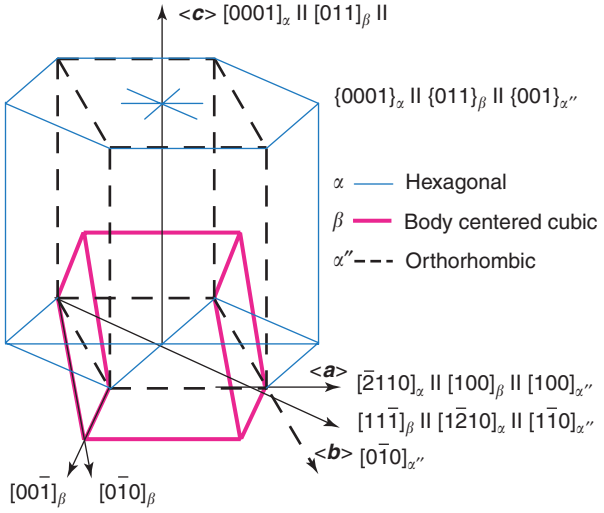


Figure 1 Idealized orientation relationship between α, β, and α″ crystal structures.

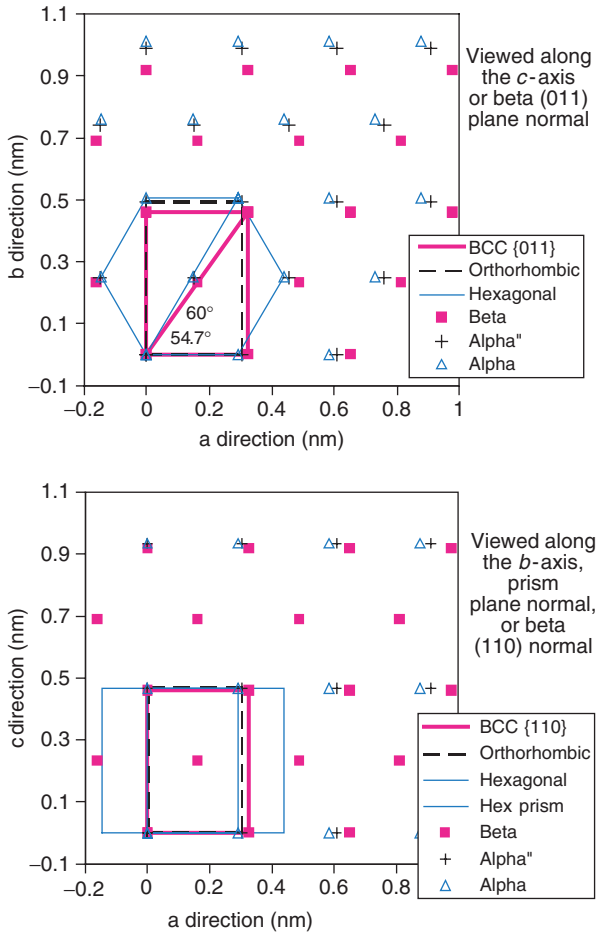


Figure 2 Atomic lattice positions of the α, β, and α'' crystal structures.

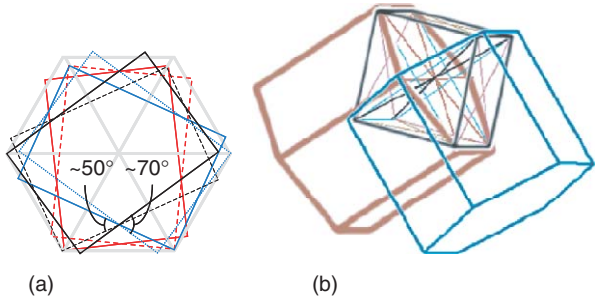


Figure 3 (a) Possible β {0 1 1} plane orientations based upon parent α during $\alpha \rightarrow \beta$ transformation, and (b) two possible α orientations based upon parent β during $\beta \rightarrow \alpha$ transformation.

With the reverse $\beta \rightarrow \alpha$ transformation upon cooling, a $\langle -2\,1\,1\,0 \rangle_\alpha$ vector can align with one of the two $\langle 1\,1\,1 \rangle_\beta$ directions on each of the six {1 1 0} planes in the β-phase. Since the angles between $\{1\,1\,0\}_\beta$ planes are either 60° or 90°, the c-axes of daughter α can point in as many as six directions, and each direction has two variants that are rotated $\sim 10°$ about the c-axis. An example of a 90° misorientation from the same parent β crystal is shown in **Figure 3b**.

Although there are 12 variants of the α orientation possible for each parent β orientation, most prior β grains will only choose a few of the 12 α variants during the α–β transformation process. Recent research has shown that the preferred transformation variants are those that have higher populations of dislocations present on $\{1\,1\,0\}_\beta$ planes. Therefore, it is possible to control the transformation process by controlling the dislocation density in the β-phase. The dislocation density depends on dislocation generation (plastic deformation) and annihilation (recovery) processes that are sensitive to prior deformation history, heating rates, and time at temperature. The understanding of what controls variant selection is just emerging, and it is likely to be exploited profitably in the coming decade to allow the manufacturing processes to be designed to achieve optimized properties for a given application.

In addition to α-and β-phases, there are a number of subtle variations, including α' and α'' martensites that may form during quenching, and ordered phases such as α_2- and ω-phases. With quenching, there is no opportunity for diffusional segregation of alloying elements; so the resulting hexagonal supersaturated α' has lattice constants slightly different from equilibrium α. The α''-phase has an orthorhombic crystal structure, which is outlined in bold lines in **Figure 1**. This crystal structure is between the α and the β crystal lattice (**Figure 2**). The α_2-phase (Ti_3Al) is an ordered version of the hexagonal lattice with Al atoms spaced regularly on the hexagonal lattice, which increases the resistance to dislocation motion. There has been much interest in ordered orthorhombic intermetallic Ti–Al–Nb alloys in the past decade. The ω-phase is a different hexagonal ordered-crystal structure that could form in prior β-regions after long-term aging or in α-regions after stress-induced transformation from $\alpha \rightarrow \omega$. Ordered phases resist dislocation motion, which can increase the strength, but the continuous ω-phase is known to reduce the ductility and facilitate crack nucleation.

Figure 2 shows that the atomic arrangement of the α-, α''-, and β-phases are very similar; but small differences in atomic position lead to different crystal structures that have significantly different properties. Thus, phase boundaries in Ti alloys are typically coherent or semi-coherent. These phase boundaries are quite easily observed in polished and etched microstructures, often giving the impression of a fine grain size, but it is important to realize that the domains of similar orientations are often much

larger than they appear, and these "invisible" mesotextural features can account for early dwell fatigue crack nucleation.

Unlike steels, the martensitic transformations in titanium result in small volume changes and modest shears, but the paradigm of using martensitic transformations in steels to manipulate microstructures and properties is also used in titanium alloys, though the details are completely different (e.g., α'' is softer than α). Ordered intermetallic alloys based on Ti_3Al and TiAl can also be manipulated with similar strategies as those used in titanium alloys, but again, the details differ. The technology of titanium alloys is highly dependent on the ability to control the phase transformation from the β- to the α-phase using strategic alloying and heat treatment history, which gives titanium alloys flexible and designable properties similar to that possible with steels.

Effects of Crystal Orientations on Mechanical Properties

The hexagonal crystal structure has considerable elastic and plastic anisotropy, which are strongly affected by processing history. Young's modulus is highest (143 GPa) along the *c*-axis, and lowest (104 GPa) in any direction in the basal plane. Titanium deforms with eight slip and twinning systems are illustrated in **Figure 4**. Each of these systems require different critical resolved shear stresses to operate, making the plastic deformation process much more complicated than cubic metals. Prism slip in $\langle a \rangle$ directions is the easiest, leading to plastic deformation that leaves the crystal dimension along the *c*-axis and the *c*-axis orientation unchanged. The next easiest slip system is basal slip in $\langle a \rangle$ directions, which also does not change the dimension of the crystal along the *c*-axis, but it causes crystal rotation about a prism plane normal axis. It is possible to have $\langle a \rangle$ slip on pyramidal planes as well.

To change the crystal dimension along the *c*-axis, twinning or $\langle c+a \rangle$ slip on pyramidal planes is required, but the critical resolved stress for $\langle c+a \rangle$ slip is high, so twinning systems are usually more easily activated. Consequently, Ti crystals are intrinsically hard when stressed along the *c*-axis. Furthermore, the relative strengths of these deformation systems change with temperature, for example, twinning does not occur at elevated temperatures. Consequently, the distribution of crystal orientations can have a large impact on both elastic and plastic properties. Even pure titanium behaves like a composite material; depending on the crystal orientation with respect to the resolved shear stresses, any crystal can be either hard or soft. In **Figure 5**, shades of gray represent the Taylor factor (proportional to the plastic flow stress of the crystal) in a rolled plate deformed in uniaxial tension.

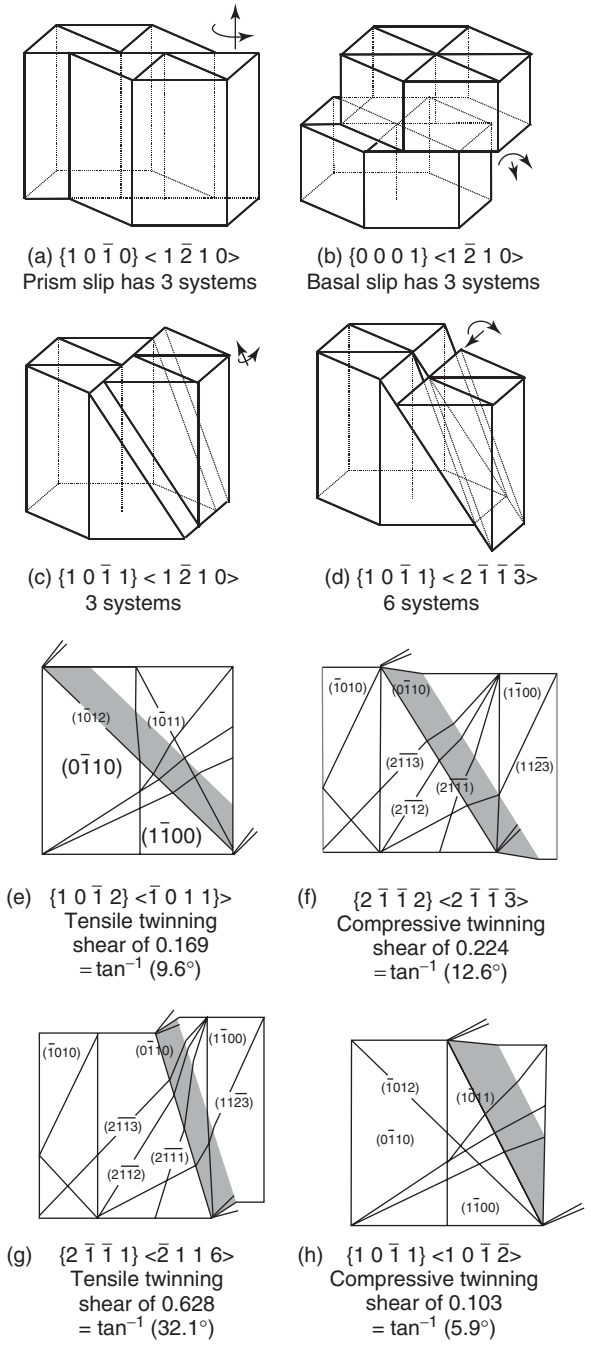


Figure 4 (a) Prism slip, (b) basal slip in the $\langle a \rangle$ direction, (c) $\langle a \rangle$ slip, (d) $\langle c+a \rangle$ slip on pyramidal planes, and (e–h) twinning systems; all types have 6 systems. Only $\langle c+a \rangle$ slip or twinning causes a change in crystal dimension in the *c* direction. Arrows illustrate axis of crystal rotation due to slip; only prism slip does not change the orientation of the *c*-axis.

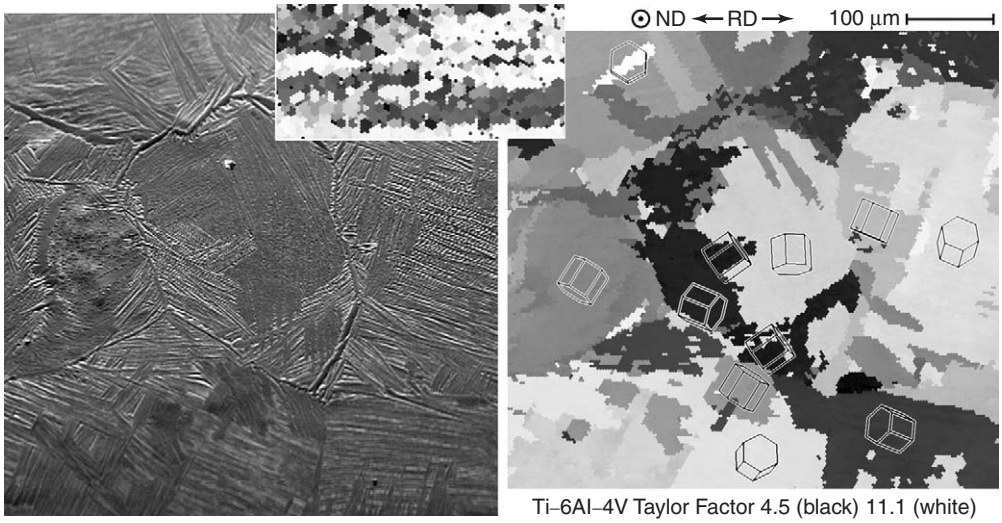


Figure 5 Rolled plate (rolling direction and plate normal are indicated) SEM and Orientation Imaging Microscopy Taylor factor maps before (inset) and after annealing above β transus, followed by controlled slow cooling. Taylor factor was computed based upon uniaxial deformation in the rolling direction with CRSS ratios of 0.7:1:3 for prism:basal:pyramidal $\langle c+a \rangle$ slip. (Images obtained at the Air Force Research Laboratory, WPAFB, OH.)

The Classification of Titanium Alloys

Titanium alloys are classified according to the phases present within their microstructure. Alloys consisting of mainly α-phase are called α-alloys, but if small amounts of β-phase are present, the alloy is classified as a near-α alloy. Alloys consisting of a mixture of both α- and β-phases are termed α–β alloys. Finally, titanium alloys that have the majority of β-phase at room temperature are called β-alloys. **Figure 6** shows the effects that several important alloying elements have on the phase diagram, indicating that the equilibrium between α, $\alpha+\beta$, and β phase fields (α and β transus lines) are highly sensitive to alloy composition. Aluminum is one of the most important alloying elements because it is a potent solid–solution strengthener and it reduces the density, so it is found in virtually all titanium alloys. Ti_3Al (α_2) forms with more than 6 wt.% Al, which can strengthen, but also embrittle the alloy. Molybdenum and vanadium are the two most popular β-stabilizing additives, which also provide strengthening of the β-phase. Tin and zirconium are the only two alloying elements used which have neither an α-stabilizing nor β-stabilizing effect on the crystal structure (**Figure 6a**), but both provide solid–solution strengthening to the α-phase; also, Sn is known to improve weldability. All other elements either stabilize the α- or β-phase, as summarized in **Table 1**.

The aluminum equivalent and molybdenum equivalent equations (in wt.%) are used to describe the degree of stability of the α- and β-phases, and to define what class an alloy is in. Oxygen and nitrogen are potent α-stabilizers, and iron is the most potent β-stabilizer:

$$Al_{equiv} = \%Al + \tfrac{1}{3}(\%Sn) + \tfrac{1}{6}(\%Zr) + 10(\%O)$$

$$Mo_{equiv} = \%Mo + 0.67(\%V) + 2.9(\%Fe) + 1.6(\%Cr) - \%Al$$

$$O_{equiv} = \%O + 2(\%N) + 0.67(\%C)$$

Commercially Pure Titanium; 99–99.5% Ti

Commercially pure (CP) titanium is considered an α-alloy, because the α-phase is the only phase present. Oxygen is the main alloying element that determines the grade and strength of the alloy (**Figures 6b** and **7**). CP titanium has lower strength than other alloys, but it is the alloy of choice for applications requiring corrosion resistance, and it has better elevated temperature creep resistance, and is less expensive than the other titanium alloys. **Figure 7** shows how the strength and hardness increase with interstitial element concentration. Since oxygen is the main alloying element in CP titanium, the O_{equiv} describes the strengthening effect of the interstitial elements O, N, and C. Each 0.1% O_{equiv} increases the strength of unalloyed titanium by $\sim$120.5 MPa.

Near α and α Titanium Alloys; $Al_{equiv} < 8$ and $Mo_{equiv} < 1$

Alpha titanium alloys are primarily used in the chemical and processes engineering industry. These applications require excellent corrosion resistance and ductility. Aluminum is the most important alloying element in order to stabilize α-phase, add

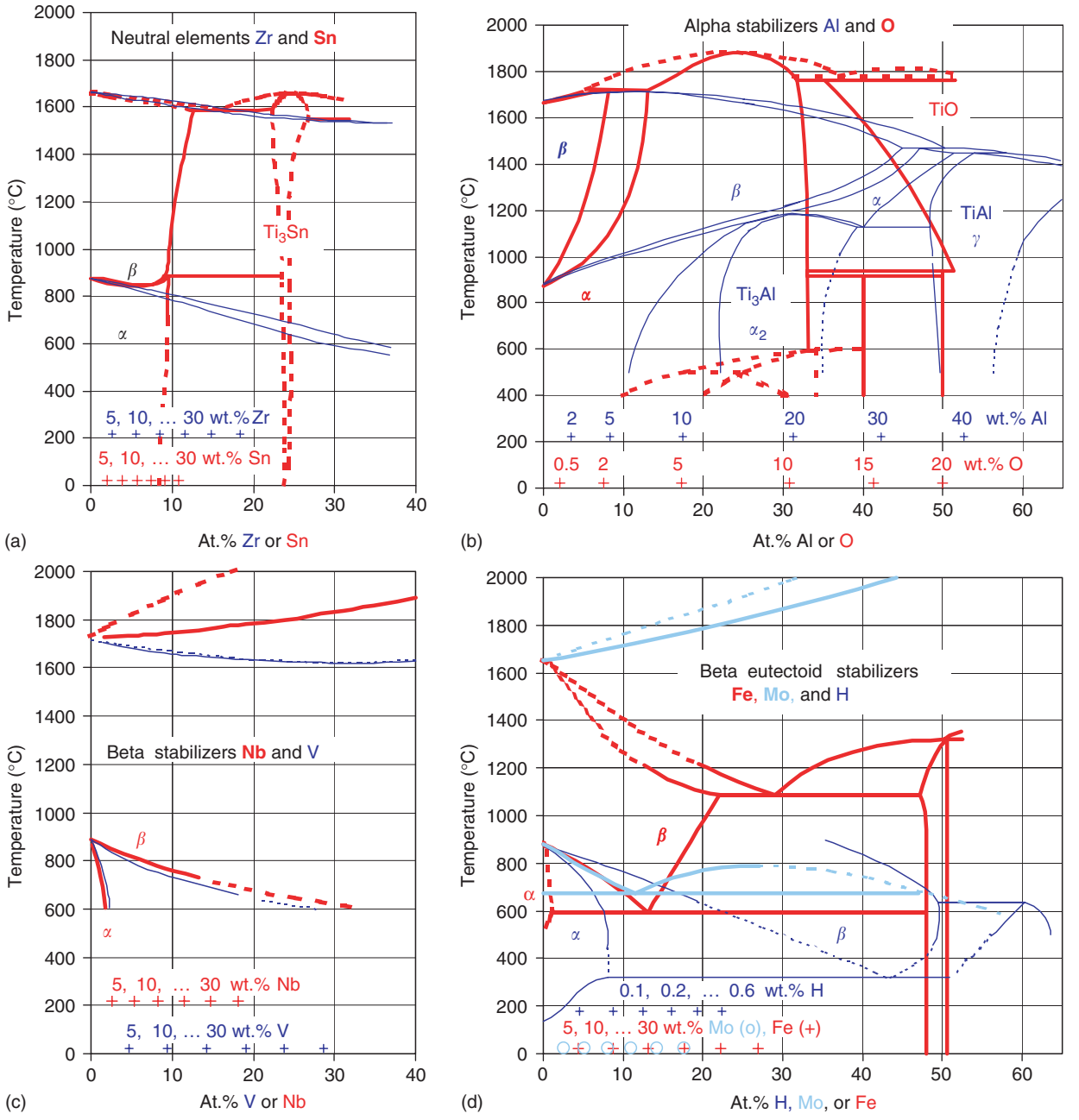


Figure 6 Exemplary phase diagrams for elements that have (a) neutral, (b) α stablizing, (c) β stabilizing, and (d) β eutectoid stablizing effects. (Adapted from Smithells Metals Reference Book, 7th edn., 1992, London: Butterworth. McCullough C, Valencia JJ, Levi CG, and Mehrabian R (1989) Phase equilibria and solidification in Ti–Al alloys. *Acta Metallurgica* 37(5): 1321–1336, for update of Ti–Al system.)

strength, and lower the density (**Figure 6b**, **Table 2**). Ti–5Al–2.5Sn is the most commonly used α alloy, and like most α alloys it cannot be age-hardened, but it does has excellent weldability.

Near-α Titanium Alloys; $6 < Al_{equiv} < 10$ and $Mo_{equiv} < 2$

Near-α alloys are ideal for high temperature applications ~500–550°C, where its excellent creep resistance can be combined with higher strength due to a small amount of dispersed β-phase. Small amounts of molybdenum and vanadium are added to retain some β-phase at room temperature. Ti–8Al–1Mo–1V is the most commonly used near-α alloy, however the high aluminum content can cause stress corrosion cracking (SCC) problems; therefore, most alloys used today are limited to 6 wt.% Al to avoid SCC problems. This alloy has good weldability, but a limited degree of hardenability due to a small amount of β-phase.

Table 1 Effects of alloying elements

Alloying element	*Range* (wt.%)	*Effect on structure and properties*
Aluminum, Al	2–7	α-Stabilizer, solid solution strengthener, reduces density, improves oxidation resistance
Carbon, C	0.05–0.1	α-Stabilizer, interstitial element
Chromium, Cr	2–12	β-Eutectoid stabilizer
Cobalt, Co		β-Eutectoid stabilizer
Copper, Cu	2–6	β-Eutectoid stabilizer, improves weldability, α and β strengthener
Hydrogen, H	0.008–0.02	β-Eutectoid stabilizer, interstitial element
Iron, Fe		β-Eutectoid stabilizer
Manganese, Mn		β-Eutectoid stabilizer
Molybdenum, Mo	2–20	β-Isomorphous stabilizer, moderate solid solution strengthener of β phase
Nickel, Ni		β-Eutectoid stabilizer
Niobium, Nb		β-Isomorphous stabilizer, known to improve oxidation behavior of Ti alloys, moderate solid solution strengthener of β phase
Nitrogen, N	0.015–0.07	α-Stabilizer, interstitial element
Oxygen, O	0.1–0.4	α-Stabilizer, interstitial element, strengthens α phase
Silicon, Si	0.05–1	β-Eutectoid stabilizer, Si atoms tend to segregate at dislocations and thus effectively prevent dislocation climb, improving creep resistance, also strengthens α phase
Tantilum, Ta		β-Isomorphous stabilizer
Tin, Sn	2–6	Neutral stabilizer, improves weldability, solid solution strengthener of α phase
Vanadium, V	2–20	β-Isomorphous stabilizer, moderate solid solution strengthener of β phase
Zirconium, Zr	2–8	Neutral stabilizer, solid solution strengthener of α phase, Zr also tends to homogenize fine silicide precipitates

Various sources, including (1990) Properties and selection: nonferrous alloy and special-purpose materials. *ASM Metal Handbook*, vol. 2, p. 605. Metals Park, OH: ASM International.

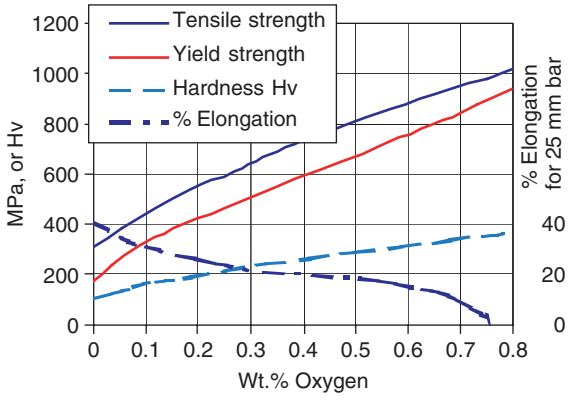


Figure 7 Effect of oxygen on strength and elongation in CP titanium. (Adapted from Jaffee RI (1958) The physical metallurgy of titanium alloys. *Progress in Metal Physics* 7:109.)

α–β Titanium Alloys; $5 < Al_{equiv} < 10$ and $2 < Mo_{equiv} < 8$

The α–β alloy Ti–6Al–4V is the most popular of all titanium alloys, representing more than 50% of the titanium market (**Figures 6b** and **6c**). The α–β alloys can be solution heat-treated, quenched, and aged to medium-/high-strength levels and have good formability, but the creep resistance and weldability are lower than the α and near-α alloys, primarily due to the presence of the β-phase, which has a much higher diffusivity and more slip systems.

β Titanium Alloys; $Al_{equiv} < 6$ and $Mo_{equiv} = 15$–30 (Metastable β), > 30 (Stable β)

The β titanium alloys are heat treatable, to achieve the highest strength levels of the five types of titanium alloys (**Figures 6c** and **6d**). The high strength arises from precipitation of very fine α-phase during an aging heat treatment. The b.c.c. crystal structure also gives good formability properties prior to heat treatment. Beta titanium alloys contain large amounts of alloying elements such as Mo, Cr, V, and Fe which are needed to stabilize the β-phase. The high degree of alloying elements makes β alloys the densest of the five alloy types (**Table 2**).

Tables 2–4 provide further details, using normal font to identify CP, italics to identify near-α and α alloys, bold italics to identify α–β alloys, and bold to identify β alloys. **Table 2** is sorted by density to illustrate the influence of alloying. **Table 5** describes weldability issues, and **Tables 3** and **4** describe heat treating methodologies, which are discussed next.

Table 2 Effect of alloying on properties (sorted by density)

Titanium alloy – CP, *near* α *and* α, **α − β, near β, metastable β, or β**	*Condition*	*E* (GPa)	*Poisson's ratio*	*Density* (g cm^{-3})	*Room temperature properties*			*K*(IC) (MPa m$^{1/2}$)	*V-notch* (J)
					Y.S. (MPa)	*T.S.* (MPa)	*% Elong.*		
Ti–8Al–1Mo–1V	*Duplex annealed*	*120*	*0.32*	*4.37*	*951*	*1000*	*15*	*82*	*32*
Ti–6Al–4V	***Annealed***	***110–140***	***0.342***	***4.43***	***924***	***993***	***14***	***33–110***	***19***
	Solution + age			***4.43***	***1103***	***1172***	***10***	***33–110***	
Ti–5Al–2.5Sn	*Annealed*	*115*	*0.31*	*4.48*	*807*	*862*	*16*	*96*	*26*
Ti–6Al–2Nb–1Ta–1Mo	*As rolled 2.5 cm plate*	*113.8*	*0.31*	*4.48*	*758*	*855*	*13*		*31*
Ti–3Al–2.5V	***Annealed***			***4.48***	***586***	***689***	***20***		
Ti–7Al–4Mo	***Solution + age***			***4.48***	***1034***	***1103***	***16***		***18***
99.5%Ti Grade 1	Annealed	102.7	0.34	4.51	241	331	30		
99.2%Ti Grade 2	Annealed	102.7	0.34	4.51	354	434	28		43
Ti–0.3Mo–0.8Ni (Ti Code 12)				*4.51*					
Ti–6Al–2Sn–4Zr–2Mo	*Duplex annealed*			*4.54*	*896*	*979*	*15*		
	Annealed	***110–117***		***4.54***	***1000***	***1069***	***14***	***30–70***	***18***
Ti–6Al–6V–2Sn (Cu + Fe)	***Solution + age***			***4.54***	***1172***	***1276***	***10***	***30–70***	
Ti–6Al–2Sn–2Zr–2Mo–2Cr–0.25Si	***Solution + age***	***110–120***	***0.327***	***4.57***	***1138***	***1276***	***11***	***65–110***	
Ti–6Al–2Sn–4Zr–6Mo	***Solution + age***	***114***		***4.65***	***1172***	***1269***	***10***	***30–60***	
Ti–5Al–2Sn–4Mo–2Zr–4Cr (Ti–17)		***112***		***4.65***	***1050***	***1100–1250***	***8–15***	***30–80***	
Ti–10V–2Fe–3Al		**110**		**4.65**	**1000–1200**	**1000–1400**	**6–16**	**30–100**	
Ti–15V–3Al–3Cr–3Sn		**80–100**		**4.78**	**800–1000**	**800–1100**	**10–20**	**40–100**	
Ti–13V–11Cr–3Al	**Solution + age**		**0.304**	**4.82**	**1172**	**1220**	**8**	**77**	**11**
Ti–3Al–8V–6Cr–4Zr–4Mo (β C)	**Solution + age**			**4.82**	**1379**	**1448**	**7**	**50–90**	**10**
	Annealed	**86–115**		**4.82**	**834**	**883**	**15**	**50–90**	
Ti–8Mo–8V–2Fe–3Al	**Solution + age**			**4.85**	**1241**	**1310**	**8**		
15Mo–3Al–0.3Fe–2.8Nb–0.2Si (β-21S)			**0.33**	**4.94**					
Ti–15Mo–5Zr–3Al				**5.01**					
Ti–11.5Mo–6Zr–4.5Sn (β III)	**Solution + age**	**83–103**		**5.06**	**1317**	**1386**	**11**	**50–100**	

Y.S. = Yield strength; T.S. = Tensile strength.

Sources: (1990) Properties and selection: Nonferrous alloys and special-purpose materials. *ASM Metals Handbook*, vol. 2, p. 621. Metals Park, OH: ASM International; (1993) *ASM Metals Reference Book*, 3rd edn. Metals Park, OH: ASM International.

Table 3 Heat treating of titanium alloys

Titanium alloy – CP, *near* α *and* α α + β, near β, metastable β, β	*T(β)* °C ± 15	*Stress–relief*		*Annealing*			*Solution treating*			*Aging*	
		°C	*Time* (h)	°C	*Time* (h)	*Cooling*	°C	*Time* (h)	*Cooling*	°C	*Time* (h)
99.5%Ti Grade 1	910	480–595	1/4–4	650–760	1/10–2	Air	N/A	N/A	N/A	N/A	N/A
99.2%Ti Grade 2	913	480–595	1/4–4	650–760	1/10–2	Air	N/A	N/A	N/A	N/A	N/A
α or near-α alloys											
Ti–5Al–2.5Sn	*1050*	*540–650*	*1/4–4*	*720–845*	*1/6–4*	*Air*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
Ti–8Al–1Mo–1V	*1040*	*595–705*	*1/4–4*	*790*	*1–8*	*Air or furnace*	*980–1010*	*1*	*Oil or water*	*565–595*	…
Ti–6Al–2Sn–4Zr–2Mo	*995*	*595–705*	*1/4–4*	*900*	*1/2–1*	*Air*	*955–980*	*1*	*Air*	*595*	*8*
Ti–6Al–2Nb–1Ta–1Mo	*1015*	*595–650*	*1/4–2*	*790–900*	*1–4*	*Air*					
Ti–0.3Mo–0.8Ni (Ti code 12)	*880*	*480–595*	*1/4–4*								
α–β alloys											
Ti–6Al–4V	***1000 (±20)***	***480–650***	***1–4***	***705–790***	***1–4***	***Air or furnace***	***955–970***	***1***	***Water***	***480–595***	***4–8***
Ti–6Al–6V–2Sn (Cu + Fe)	***945***	***480–650***	***1–4***	***705–815***	***3/4–4***	***Air or furnace***	***885–910***	***1***	***Water***	***480–595***	***4–8***
Ti–3Al–2.5V	***935***	***540–650***	***1/2–2***	***650–760***	***1/2–2***	***Air***					
Ti–6Al–2Sn–4Zr–6Mo	***940***	***595–705***	***1/4–4***	***N/A***	***N/A***	***N/A***	***845–890***	***1***	***Air***	***580–605***	***4–8***
Ti–5Al–2Sn–4Mo–2Zr–4Cr (Ti–17)	***900***	***480–650***	***1–4***	***N/A***	***N/A***	***N/A***	***845–870***	***1***	***Air***	***580–605***	***4–8***
Ti–7Al–4Mo	***1000***	***480–705***	***1–8***	***705–790***	***1–8***	***Air***					
Ti–6Al–2Sn–2Zr–2Mo–2Cr–0.25Si	***970***	***480–650***	***1–4***	***705–815***	***1–2***	***Air***	***870–925***	***1***	***Water***	***900–1100***	***4–8***
β or near-β alloys											
Ti–13V–11Cr–3Al	**720**	**705–730**	**1/12–1/4**	**705–790**	**1/6–1**	**Air or water**	**775–800**	**1/4–1**	**Air or water**	**425–480**	**4–100**
Ti–11.5Mo–6Zr–4.5Sn (β III)	**760**	**720–730**	**1/12–1/4**	**690–760**	**1/6–2**	**Air or water**	**690–790**	**1/8–1**	**Air or water**	**480–595**	**8–32**
Ti–10V–2Fe–3Al	**805**	**675–705**	**1/2–2**	**N/A**	**N/A**	**N/A**	**760–780**	**1**	**Water**	**495–525**	**8**
Ti–15V–3Al–3Cr–3Sn	**760**	**790–815**	**1/12–1/4**	**790–815**	**1/12–1/4**	**Air or water**	**790–815**	**1/4**	**Air**	**510–595**	**8–24**
Ti–8Mo–8V–2Fe–3Al	**775**										
Ti–15Mo–5Zr–3Al	**785**										
15Mo–3Al–0.3Fe–2.8Nb–0.2Si (β-21S)	**800**										
Ti–3Al–8V–6Cr–4Zr–4Mo (β C)	**795**	**705–760**	**1/6–1/2**	**790–815**	**1/4–1**	**Air or water**	**815–925**	**1**	**Water**	**455–540**	**8–24**

Adapted from (1991) Heat treating. *ASM Metals Handbook*, vol. 4, pp. 914–917. Metals Park, OH: ASM International.

Table 4 Heat treating of α–β alloys

Heat treatment designation	*Heat treatment cycle*	*Microstructure*
Duplex anneal	Solution treat at 50–75°C below $T(\beta)$, air cool and age for 2–8 h at 540–675°C	Primary α, plus Widmanstätten α–β regions
Solution treat and age	Solution treat at ~40°C below $T(\beta)$, water quench(a) and age for 2–8 h at 535–675°C	Primary α, plus tempered α' or a β–α mixture
Beta anneal	Solution treat at ~15°C above $T(\beta)$, air cool and stabilize at 650–760°C	Widmanstätten α–β colony microstructure
Beta quench	Solution treat at ~15°C above $T(\beta)$, water quench and temper at 650–760°C for 2 h	Tempered α'
Recrystallization anneal	925°C for 4 h, cool at 50°C h^{-1} to 760°C, air cool	Equiaxed α with β at grain-boundary triple points
Mill anneal	α–β Hot work plus anneal at 705°C for 30 min to several hours and air cool	Incompletely recrystallized α with a small volume fraction of small β particles

In more heavily β-stabilized alloys such as Ti–6Al–2Sn–4Zr–6Mo or Ti–6Al–6V–2Sn, solution treatment is followed by air cooling. Subsequent aging causes precipitation of α-phase to form an α–β mixture.
Adapted from (1991) Heat treating. *ASM Metals Handbook*, vol. 4, p. 914. Metals Park, OH: ASM International.

Table 5 Welding of titanium alloys

Alloy type	*Condition*	*Description*
Unalloyed titanium	Annealed	Welding of cold-worked alloys anneals the HAZ and negates any strength produced by cold working
Alpha titanium alloys	Annealed	Ti–5Al–2.5Sn, Ti–6Al–2Sn–4Zr–2Mo, Ti–5Al–5Sn–2Zr–2Mo, Ti–6Al–2Nb–1Ta–1Mo, and Ti–8Al–1Mo–1V are always welded in annealed condition
Alpha–beta titanium alloys	Annealed or solution-treated and partially aged	Low weld ductility of most α–β alloys is caused by phase transformation in the weld zone or HAZ
Metastable beta titanium alloys	Annealed or solution heat treated	In as-welded condition, welds are low in strength but ductile; to obtain full strength the alloys are welded in annealed condition, the weld is cold worked by shot peening or planishing, and the weldament is then solution treated and aged
Weld type	*Thickness range*	*Description*
Gas–tungsten arc weld (GTAW)	Base metal up to 2.5 mm, for thicker base metal filler metal is required	
Gas–metal arc weld (GMAW)	More than 3 mm to greater than 13 mm	Applied using pulse current or spray mode and is less costly than GTAW
Plasma arc welding (PAW)	Plate up to 13 mm, filler metal may be used	Faster than GTAW and can be used on thicker sections
Electron-beam welding (EBW)	Plates 6 mm to more than 76 mm	Used in aircraft and aerospace industry for producing high quality welds; performed in high vacuum atmosphere therefore low contamination of weldament
Laser-beam welding (LBW)	Base metal usually cannot exceed 13 mm	
Friction welding (FRW)		Useful for joining tube, pipe, or rods
Resistance welding (RW)		Used to join titanium sheet by spot welds or continuous seam welds, also used to weld titanium sheet to dissimilar metals
Fluxes	N/A	Fluxes cannot be used because they combine with titanium to form brittleness and may reduce corrosion resistance

Excerpted from (1990) Welding brazing and soldering. *ASM Metals Handbook*, vol. 6, pp. 783–786. Metals Park, OH: ASM International.

Heat Treating of Titanium Alloys

Titanium and titanium alloys are heat treated for a number of reasons: to reduce residual stresses from the many different fabrication processes (stress relieving), to obtain optimum combinations of ductility, machinability, and dimensional stability (annealing), to increase strength (solution treating and aging in α–β and β alloys), and finally to optimize specific properties such as fracture toughness, fatigue strength, and high-temperature creep resistance. In general, a heat treatment is used to rearrange phase–volume fractions, phase boundaries in strategic ways. A quench from high temperature is commonly used to provide a thermodynamic driving force for nanoscale rearrangement at lower aging temperatures. More recently, benefits of rapid heating into the β-phase field to obtain desirable microstructures have been described. Standard practice is identified in **Tables 3** and **4**, and the preferred phase for alloying elements is in **Table 1**. However, alloy element local concentration is highly dependent on the history of the phase volume fraction and locations of phase boundaries in space and hence, on all prior heat treating and working history. For this reason, the same heat treatment on materials with different prior history can cause considerable variability in resulting properties.

Unalloyed α titanium alloys can be stress relieved and annealed, but cannot be heat treated to increase their strength. When the heat treatment of titanium alloys involves heating to temperatures near the β transus, it is important that T_β is known for each specific alloy. **Table 3** illustrates how the β transus temperature (β_T) depends on alloy composition.

Stress Relieving

Stress relieving (SR) is used to remove residual stresses from prior fabrication steps, or even prior heat treatments. Unbalanced residual stresses can result in part distortion, and cause problems in those alloys susceptible to hydrogen embrittlement. SR can result in aging of all alloy types, α can be aged by α_2 precipitates, and if metastable β is present, α precipitation that provides strengthening can occur. For α and α–β alloys, the SR temperature will be in the range of 480–815°C, and if these alloys were β-annealed, more rearrangement of phase boundaries is needed and the SR temperature should be increased by $\sim$55°C. When stress-relieving β alloys, care must be taken to avoid interfering with the final age-hardening treatment. If a β alloy has not been heat treated, the SR temperature should be substantially below the β_T to prevent preferred α precipitation in grain boundaries (which embrittles the material). If the β alloy is in its final age-hardened condition, the SR temperature should be at or below the aging temperature to prevent strength reduction.

Annealing

Annealing is similar to SR but usually done at higher temperatures. The annealing of titanium and titanium alloys serves primarily to increase fracture toughness, room temperature ductility, dimensional stability, and high-temperature creep resistance. This is often the final heat treatment for α and α–β alloys, as it yields a good balance of properties. Care must be taken when using β alloys in the annealed condition in service temperatures up to 400°C, which could result in embrittlement, and service temperatures in the range of 400–600°C could age-harden the β alloy, resulting in a strength increase at the expense of ductility. Alpha and α–β alloys are typically annealed in the temperature range of 700–900°C, while β alloys are normally annealed in the temperature range of 690–815°C.

There are four commonly used annealing treatments, mill annealing (MA), duplex annealing (DA), recrystallization annealing (RA), and β annealing (BA). MA is a general-purpose heat treatment given to all mill products, but it is not a full anneal and the material may retain effects of prior working processes. DA is used for α and α–β alloys; it involves an initial anneal high in the α/β-phase field to generate significant fractions of both α and β with small grain/phase sizes near 10 μm, followed by an MA to provide thermal stability in the prior β regions. DA improves creep resistance and fracture toughness (resistance to crack nucleation) of the material. An RA is done at a temperature high enough in α/β-phase field to ensure recrystallization. It is then slowly cooled to form a high-volume fraction of equiaxed α with islands of retained β at triple points, and some interfacial β at α/α boundaries. RA provides high-damage tolerance properties (fracture toughness, crack growth resistance) that are crucial for fracture critical applications. Finally, BA is done at temperatures above the β_T, leading to a large grain size followed by a subsequent MA (**Figure 5** shows how a BA converted a small grain size DA microstructure into large prior β grains that were transformed into several α variants during slow cooling). For α and α–β alloys, BA maximizes the damage tolerance properties, since resistance to crack propagation is best when there are fewer grain boundaries.

Solution Treating

Solution treating (ST) is used to transform a desired amount of α- to β-phase in near α and α–β alloys, by

heating the material to a temperature near β_T and then cooling strategically to produce a higher ratio of β-phase in a desired morphology. Beta alloys are normally solution treated above the β_T (700–815°C), while α and α–β alloys are normally solution treated slightly below the β_T (850–1000°C) (**Table 3**). For most β alloys, the object of ST is to retain 100% β-phase upon quenching, so that upon subsequent aging, decomposition of the metastable β-phase occurs to generate second-phase precipitates that provide high strength.

Quenching

The rate of cooling from the ST temperature strongly affects the strength of the material. Slow cooling rates allow diffusion to occur, resulting in decomposition of the β-phase and may prevent effective strengthening during the aging process. Also the quench delay, which is the time from which the material is removed from the furnace and subsequently placed into the quenching media, can affect properties. Longer quench delays can lower strength; this reduced strength is ascribed to the formation of coarse acicular α prior to immersion in quenchant. Alpha–beta alloys are typically quenched in water, a 5% brine solution, or a caustic soda solution because the rate is sufficient to retain the β-phase obtained by ST. For highly β-stabilized alloys and small sections, an air or fan cooling may be adequate.

Aging

Aging is the final heat treatment step needed in order to achieve high-strength levels. Aging causes the decomposition of the supersaturated β-phase retained upon quenching, through α-precipitation within the β-matrix. Martensites also decompose in a manner that sharpens interfaces or boundaries that provide strengthening. The aging temperature controls the size and volume fraction of the α-precipitates, and determines the final strength of the material. Aging at or near the annealing temperature results in overaging, which is desirable if good toughness and dimensional stability are needed with only a modest increase in strength. For α and α–β alloys, aging is usually done in the temperature range of 480–605°C, and β alloys are typically aged in the temperature range of 425–595°C.

Suggested heat treatments for various titanium alloys are listed in **Tables 3** and **4**. It is also important to note that not all heat treatments are applicable to all titanium alloys, because some alloys are designed for specific purposes, for example, Ti–5Al–2Sn–2Zr–4Mo–4Cr (Ti-17) and Ti–6Al–2Sn–4Zr–6Mo are designed for strength in heavy sections, Ti–5Al–2.5Sn and Ti–2.5Cu, for weldability, Ti–6Al–2Sn–4Zr–2Mo and Ti–6Al–5Zr–0.5Mo–0.2Si, for creep resistance.

Welding of Titanium Alloys

Table 5 provides relevant details for welding titanium alloys. Unalloyed titanium and α alloys are weldable. These alloys are typically welded in the annealed condition, and have good weldability because they are insensitive to heat treatments. Alpha–beta alloys and weakly β-stabilized alloys are also weldable, but should be welded in the annealed or solution-treated conditions prior to aging heat treatments. Most β alloys can be welded in annealed or solution-treated condition, but subsequent aging can cause the weld to become brittle. Most strongly β-stabilized alloys are embrittled by welding, since it is not possible to quench a weld (without quenching, second phases precipitate preferentially on grain boundaries).

Challenging and Problematic Issues with Titanium Alloys

Although titanium alloys have many desirable characteristics, they have problematic areas as well. These include their affinity to oxygen, hydrogen pickup, susceptibility to certain types of chemical attack, and damage generation arising from the elastic and plastic anisotropy of the hexagonal crystal structure.

Heat treating can cause contamination due to oxidation that causes an oxygen-rich brittle layer on the metal surface called "α case." Unless surface hardening is needed for wear resistance, this α case must be mechanically or chemically removed before the part can be put into service. Another concern is embrittlement due to very low levels of hydrogen, nitrogen, and carbon absorption at virtually every fabrication stage from casting to heat treating. Due to its great affinity to oxygen, alumina (Al_2O_3) cannot be used to machine titanium alloys, so WC tools (typically C-2 grades) are used for turning and face milling, while high-carbide, high-speed steels are appropriate for drilling, tapping, and end milling.

Though CP titanium is highly corrosion resistant, SCC is more likely with increasing alloy content, and in environments with red fuming nitric acid, N_2O_4, HF, methonal, HCl, and seawater. The susceptibility depends greatly on alloy elements, crystal orientation and misorientations at grain boundaries, and the details of microstructural evolution.

Due to its hexagonal crystal structure, slip in the $\langle a \rangle$ direction is much more facile than other

directions, so heterogeneous deformation and shear banding is common. Combinations of working and heat treatment can generate highly preferred crystal orientations, so between the coherent phase boundaries with transformation shears, and the anisotropy of the thermal expansion coefficient (which is 20% higher in the $\langle c \rangle$ than $\langle a \rangle$ directions) substantial residual stresses are common even after a stress relief anneal. Though titanium alloys can be processed for excellent high-temperature creep resistance, room-temperature creep occurs in some microstructures and alloys, and hence dimensional stability can be a problem. To fully optimize material property design for particular applications, control of texture and microstructure with innovative processing strategies requires predictive understanding of the rules for variant selection during the $\alpha \rightarrow \beta \rightarrow \alpha$ transformations, which are not yet well established.

See also: Dislocations; Mechanical Properties: Plastic Behavior; Periodicity and Lattices; Recovery, Recrystallization, and Grain Growth.

PACS: 61.66. – f; 62.20. – x; 64.70. – p; 81.05.Bx; 81.40. – z; 81.40.Cd; 81.40.Ef

Further Reading

(1991) Heat treating. *ASM Handbook*, vol. 4. Metals Park, OH: ASM Publications.

Bondarchuk VI, Ivasishin OM, Moiseyeva IV, Okrainets PM, and Pishchak VK (2001) Effect of rapid heat treatment on the high-temperature deformation of refractory titanium alloys. *Metal Physics and Advanced Technologies (UK)* 19(5): 743–753.

Boyer RR and Lutjering G (1996) Heat treatment of titanium alloys: overview. In: Weiss I, Srinivasan R, Bania PJ, Eylon D, and Semiatin SL (eds.) *Titanium Alloy Processing*, pp. 349–367. Warrendale, PA: The Minerals, Metals and Materials Society.

Boyer R, Collings EW, and Welch G (eds.) (1994) *Materials Properties Handbook: Titanium Alloys*. Materials Park, OH: ASM Publications.

Donachie MJ Jr. (2000) *Titanium: A Technical Guide*, 2nd edn. Metals Park, OH: ASM Publications.

Exner HE, Muller C, and Schmidt H (2004) Modification of titanium alloys for medical applications. *Zeitschrift fur Metallkunde* 95(7): 650–662.

Froes (Sam) FH, Imam MA, and Fray D (eds.) (2004) *Cost-Affordable Titanium*. Warrendale, PA: The Minerals, Metals and Materials Society.

http://doc.tms.org

http://www.asm-intl.org

http://www.matls.com

Leyens C and Peters M (2003) *Titanium and Titanium Alloys; Fundamentals and Applications*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Semiatin SL, Seetharaman V, and Weiss I (1996) Hot working of titanium alloys – an overview. In: Weiss I, Srinivasan R, Bania PJ, Eylon D, and Semiatin SL (eds.) *Titanium Alloy Processing*, pp. 3–74. Warrendale, PA: The Minerals, Metals and Materials Society.

Smith WF (1993) *Structure and Properties of Engineering Alloys*, 2nd edn., pp. 433–484. New York: McGraw-Hill.

Veeck S, Lee D, and Tom T (2002) Titanium investment castings. *Advanced Materials & Processes* 160(1): 59–62.

Amorphous *See* Alloys: Overview; Composites: Overview; Disorder and Localization Theory; Disordered Solids and Glasses, Electronic Structure of; Porous Silicon; Waves in Random Media.

Anelasticity *See* Mechanical Properties: Anelasticity; Mechanical Properties: Elastic Behavior.

Aperiodic Crystals *See* Lattice Dynamics: Aperiodic Crystals.

B

Ballistic Transport

M Dragoman, National Institute for Research and Development in Microtechnology (IMT), Bucharest, Romania
D Dragoman, University of Bucharest, Bucharest, Romania

Introduction

The transport of charged carriers depends greatly on the scale of the sample. At the macroscopic scale, in bulk semiconductors with dimensions larger than 1 mm, conduction electrons move randomly due to the high scattering rates with impurities and phonons. An applied electric field induces a diffusive electron transport described by the stochastic Boltzmann equation. At the microscopic scale of atoms and molecules, which is of the order of $\text{Å} = 10^{-10}$ m, the particles evolve according to the laws of atomic and molecular physics. At the intermediate, mesoscopic scale, distinct features in the particle transport appear, which include the emergence of the ballistic transport regime characterized by the coherent, wavelike behavior of charged particles over length scales much larger than the dimension of molecules and much smaller than macroscopic dimensions. Ballistic transport is defined as the collisionless transport regime, valid when the dimensions of the sample are smaller than the mean free path L_{fp} and the phase relaxation length L_{ph}. The mean free path is a measure of the distance between successive electron collisions with impurities or phonons that destroy the initial electron momentum. L_{fp} is about tens of nanometers in polycrystalline metallic films and about tens of micrometers in high-mobility semiconductors at low temperatures ($T < 4$ K); the latter materials are thus more likely to exhibit ballistic transport. The phase relaxation length is the distance after which the coherence (the phase memory) of electrons is lost. Phase memory is lost in time-reversal breaking processes such as electron–electron collisions, dynamic scattering processes, or scattering by impurities with internal degrees of freedom (spin) that change in the process; any static spin-independent scattering does not affect phase relaxation. L_{ph} is of the same order of magnitude as L_{fp} in high-mobility semiconductors.

Experimental conditions for ballistic transport observation have become achievable only in the last decades due to impressive developments in semiconductor technology, especially in advanced epitaxial growth and lithography techniques. It has thus become possible to fabricate band-engineered semiconductor heterostructures called quantum wells, quantum wires, and quantum dots that confine, respectively, the electron motion in one, two, and three dimensions. The electron density in such structures can be controlled through electric potentials applied on surface gates. Coherent propagation of electron waves, electron tunneling, and Coulomb blockade are examples of specific phenomena associated with ballistic transport. These phenomena have found hosts of applications in the design of new devices such as resonant tunneling devices and the single-electron transistor, the ballistic transport regime being indispensable for high-frequency electronic and optoelectronic devices. Not only electrons but also holes and even quasiparticles, such as phonons, experience ballistic transport regime under the same conditions as those defined above, but with the corresponding mean free path and phase relaxation length parameters.

Energy Subbands and Density of States in Ballistic Transport

Ballistic transport occurs when, at most, only a few elastic scattering processes take place, such that electron motion is characterized in this regime by a constant energy E, and can be described by a time-independent Schrödinger equation for the envelope electron wave function Ψ, which varies slowly over the dimensions of the unit cell:

$$-\frac{\hbar^2}{2}\{m^{\alpha}\nabla[m^{\beta}\nabla(m^{\alpha}\Psi)]\} + V\Psi = E\Psi$$

where α and β are material-related parameters for which $2\alpha + \beta = -1$, m is the electron effective mass that incorporates the effect of the periodic lattice

potential on electron motion, and V is the potential energy. This scalar equation is valid when coupling between different electron bands is negligible. Ballistic transport has been mainly evidenced in AlGaAs semiconductor compounds for which $\alpha = 0$ and $\beta = -1$. The potential energy V includes the conduction band discontinuities in heterojunctions (which is the dominant component in the low doping and the low free carriers case), the electrostatic potential due to ionized donors and acceptors (which is determined from a self-consistent solution of the coupled Schrödinger and Poisson equations), and the self-consistent Hartree and exchange potentials due to free carriers.

The solution of the Schrödinger equation requires proper boundary conditions. The constraints on electron motion present in any ballistic structure impose the appearance of discrete energy levels along the direction of the constraint and a resulting discontinuity in the density of states. For example, in a quantum well the conduction electrons are free to move along the x and y directions, and are confined by potential barriers along z in a region of width L_z. For abrupt infinite-height barriers, the electron wave function $\Psi(x,y,z) = \sqrt{2/L_z}\sin(k_z z)\sqrt{1/L_x L_y}\exp(\mathrm{i}k_x x)\exp(\mathrm{i}k_y y)$, obtained from the Schrödinger equation for $V = 0$, with L_x and L_y the sample dimensions along x and y, must satisfy the boundary conditions $\Psi(x,y,0) = \Psi(x,y,L_z) = 0$. These induce a discrete spectrum $k_z = p\pi/L_z$ for the momentum component along z, with p an integer, and an energy dispersion relation of the form

$$E(k_x,k_y,k_z) = E_c + \frac{\hbar^2}{2m}\left(\frac{p\pi}{L_z}\right)^2 + \frac{\hbar^2}{2m}(k_x^2+k_y^2) = E_{s,p} + \frac{\hbar^2}{2m}(k_x^2+k_y^2)$$

where E_c denotes the bottom of the conduction band and $E_{s,p}$ is the cut-off energy of the subband or transverse mode labeled by p. The energy spacing between discrete p levels increases as L_z decreases, that is, as electrons become more confined. Although periodic boundary conditions imposed on x and y also require the quantization of k_x, k_y according to $k_x = q(2\pi/L_x)$, $k_y = r(2\pi/L_y)$, with q and r integer values, the energy spectrum remains quasicontinuous in the $k_x - k_y$ plane, since $L_x, L_y \gg L_z$. The quantization of the wave vector components along x and y mainly influences the density of states by assigning to an individual state an area $(2\pi/L_x)\times(2\pi/L_y)$ in the $k_x - k_y$ plane. Then, the total number of spin-degenerate states with a wave number less than k (or with an energy less than E, for $E > E_{s,p}$), which occupy in the $k_x - k_y$ plane an area $\pi k^2 = \pi(k_x^2 + k_y^2)$, is $N_{tot}(k) = k^2(L_x L_y)/2\pi$ (or $N_{tot}(E) = m(L_x L_y)(E - E_{s,p})/\pi\hbar^2$). The density of states per unit area $S = L_x L_y$, and per unit energy in the pth subband is thus

$$\rho_{2D,p}(E) = \frac{1}{S}\frac{\mathrm{d}N_{tot}(E)}{\mathrm{d}E} = \frac{m}{\pi\hbar^2}\,\vartheta(E - E_{s,p})$$

with ϑ the unit step function, the total density of states $\rho_{2D}(E) = \sum_p \rho_{2D,p}$ represented in **Figure 1a** being discontinuous, unlike in bulk semiconductors where no size constriction is imposed upon the electron motion.

The equilibrium electron density per unit area is given by

$$n = \int_0^\infty \sum_p \rho_{2D,p}(E) f(E)\mathrm{d}E = k_B T\frac{m}{\pi\hbar^2}\sum_p \ln[1 + \mathrm{e}^{(E_F - E_{s,p})/k_B T)}]$$

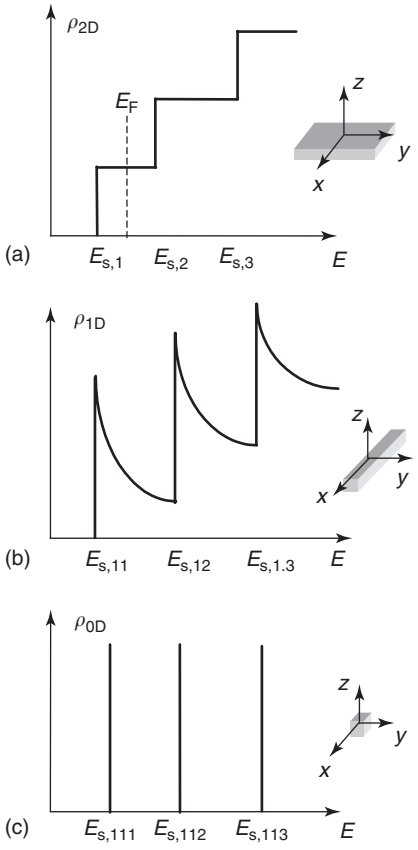


Figure 1 Density of states in (a) quantum wells, (b) quantum wires, and (c) quantum dots. The Fermi energy position in (a) corresponds to 2DEG systems.

the energy levels being occupied by electrons according to the Fermi–Dirac distribution function

$$f(E) = \frac{1}{1 + \exp[(E - E_F)/k_B T]}$$

At low temperatures or in the degenerate limit, when $k_B T \ll E_F$, the Fermi–Dirac distribution function becomes proportional to $\vartheta(E_F - E)$, and all electron subbands are filled up to the Fermi energy and are empty above it. In this case, all electrons that participate in transport have energies close to E_F, unlike in the diffusive transport regime in which electrons have a widespread energy distribution. The number of subbands $M(E)$ occupied at low temperatures by electrons with energy E is obtained by counting the number of transverse modes with cutoff energies smaller than E.

In a similar manner, if the electron motion is constrained by infinite-height potentials along two directions, y and z, the boundary conditions imposed on the electron wave function $\Psi(x, y, z) = 2/\sqrt{L_y L_z}\,\sin(k_y L_y)\sin(k_z L_z)\sqrt{1/L_x}\,\exp(\mathrm{i}k_x x)$ require that $k_y = p\pi/L_y$, $k_z = q\pi/L_z$, with p, q integers. In such a one-dimensional ballistic quantum wire, the electrons are free to move only along x, the energy dispersion relation being

$$E(k_x, k_y, k_z) = E_c + \frac{\hbar^2}{2m}\left(\frac{p\pi}{L_y}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{q\pi}{L_z}\right)^2 + \frac{\hbar^2 k_x^2}{2m} = E_{s,pq} + \frac{\hbar^2 k_x^2}{2m}$$

The total number of states with a wave vector component along x less than k_x, which occupy an "area" in the k_x plane equal to $2k_x$, is now $N_{tot}(k_x) = 2 \times 2k_x/(2\pi/L_x)$, where the first factor 2 accounts for spin degeneracy. The density of states in the k_x plane per unit length, $\rho_{1D}(k_x) = L_x^{-1}\mathrm{d}N_{tot}/\mathrm{d}k_x$, is constant and equal to $2/\pi$, half of this value representing the contribution of states for which k_x is positive, and half representing the contribution of states with negative k_x values. $\rho_{1D}(k_x)$ corresponds to a total energy-dependent density of states $\rho_{1D}(E_c) = (1/\pi)(2m/\hbar^2)^{1/2}\sum_{p,q}(E - E_{s,pq})^{-1/2}$, which is represented in **Figure 1b**.

In quantum dots the electron motion is confined along all directions, and the energy spectrum

$$E(k_x, k_y, k_z) = E_c + \frac{\hbar^2}{2m}\left(\frac{p\pi}{L_x}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{q\pi}{L_y}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{r\pi}{L_z}\right)^2 = E_{s,pqr}$$

is discrete, as in atoms or molecules. Quantum dots, with density of states $\rho_{0D} \propto \delta(E - E_{s,pqr})$ (see **Figure 1c**), can thus be considered as artificial atoms.

Two-Dimensional Electron Gas

If, as in **Figure 1a**, the position of the Fermi level in a quantum well is between the first and the second energy subband, the electron density per unit area, $n = (m/\pi\hbar^2)(E_F - E_{s,1})$, is related to the Fermi wave number k_F through $k_F = \sqrt{2\pi n}$. The Fermi wave number is determined by the kinetic energy of electrons: $E_{kin} = E_F - E_{s,1} = \hbar^2 k_F^2/2m$. The electrons form, in this case, a two-dimensional electron gas (2DEG), which has a metallic behavior since E_F is inside the conduction band.

Two-dimensional electron gas systems form, for example, at the interface between intrinsic GaAs and n-doped AlGaAs layers in modulation-doped GaAs/AlGaAs heterojunctions, the spatial separation of dopant atoms in AlGaAs from the free electrons that form an inversion layer on the GaAs side of the interface assuring extremely low scattering rates. The electron density in this 2DEG can be changed by modifying the electron kinetic energy via negative electrostatic voltages applied on Schottky surface gates situated in close proximity to the 2DEG; the bias depletes the 2DEG underneath the gate, and laterally from the geometric edge of the gate until no free electrons are left, a case in which a barrier for electrons is created. The coherence of the electron wave function in ballistic structures together with the possibility of manipulating the electron wave number allows the implementation of electron counterparts of both geometrical optical systems and interference devices. **Figures 2a–2d** show an electron prism, a refractive lens, an electron beam splitter, and an interferometer, respectively. As in the classical electromagnetic theory, the refraction at an interface between two 2DEG systems with electron densities n_1 and n_2 occurs according to Snell's law for ballistic electrons: $\sin\theta_1/\sin\theta_2 = (n_2/n_1)^{1/2}$, with θ_i the angle in region i with the normal to the interface. The electron wave function is partly reflected and partly transmitted at the interface. Narrow constrictions in the 2DEG created by the black depletion gates in **Figure 2** act as electron sources with a wide angular spectrum, the gray electrodes in **Figure 2** representing partially depleting gates that refract electron waves.

The analogy between electron propagation in ballistic conductors and electromagnetic wave propagation is based on the formal analogy between the time-independent Schrödinger equation for the electron wave function and the Helmholtz equation

$$\nabla^2 F + k^2 F = 0$$

satisfied by the electromagnetic field component F (electric field, magnetic field, or vector potential).

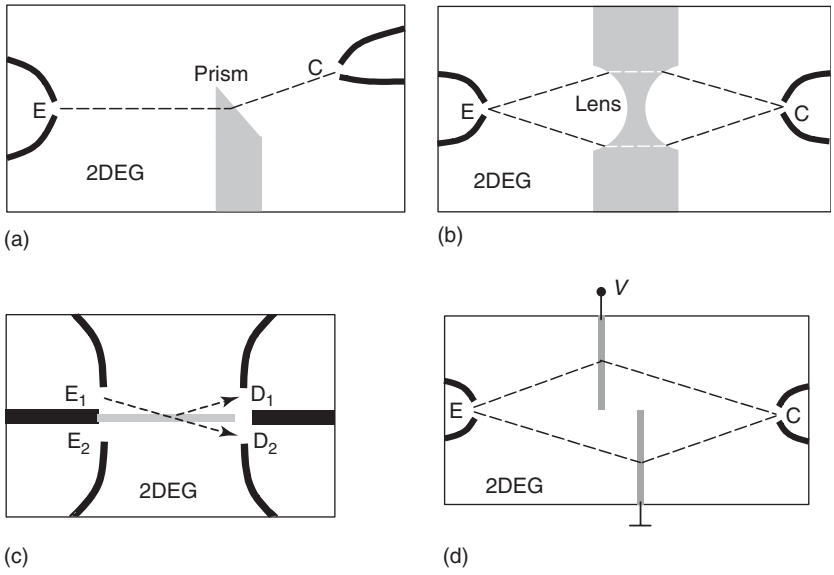


Figure 2 Schematic representation of (a) an electron prism, (b) a refractive electron lens, (c) a beam splitter, and (d) an interferometer. The electrons are emitted from the narrow orifice E and are collected by the collector C. Dashed lines represent electron trajectories.

The vectorial character of the latter implies that the eventual quantitative analogies between electron and light parameters (energy, effective mass, potential energy and frequency, electric permittivity, magnetic permeability, respectively) depend on which electromagnetic field component is compared to the scalar electron wave function. Different sets of analogous parameters between the electron wave function and the electromagnetic field are derived if, besides the evident similarity between the equation satisfied by $m^{\alpha}\Psi$ and $\boldsymbol{F}$ if k is replaced by $[2m(E-V)]^{1/2}/\hbar$, additional different boundary conditions for different field components $\boldsymbol{F}$ are considered.

Resistance of a Ballistic Conductor

A net electron transfer across a ballistic conductor sandwiched between two contacts that act as electron reservoirs is possible, when an external bias V is applied between the contacts. The bias drives the electron system in the ballistic conductor away from equilibrium such that no common Fermi energy exists; instead, one can define a spatially varying local quasi-Fermi level that takes the values E_{Fl} and E_{Fr} in the left and right contacts, respectively. The net current can then be easily calculated if the contacts are reflectionless, that is, if the electrons can enter them from the conductor without suffering reflections. If $E_{\mathrm{Fl}} > E_{\mathrm{Fr}}$, such that for a small applied bias $E_{\mathrm{Fl}} - E_{\mathrm{Fr}} = eV$, only electrons that flow from filled states on the left to empty states on the right contribute to the net current. At zero temperature, there is current flow only in the electron energy range $E_{\mathrm{Fr}} < E < E_{\mathrm{Fl}}$, the contribution to the current of electrons in each occupied subband being additive if the ballistic conductor has a constant cross section; in this case, there is no scattering of electrons from one subband (transverse mode) to another. The net current flow in one subband due to the extra density of electrons in the left contact, $\delta n = (\mathrm{d}n/\mathrm{d}E)eV$, is $I = ev\delta n$, where $v = \hbar^{-1}(\mathrm{d}E/\mathrm{d}k)$ is the electron velocity along the current flow direction. Then $I = (e^2/\hbar)V(\mathrm{d}n/\mathrm{d}E)(\mathrm{d}E/\mathrm{d}k)$, or $I = (2e^2/h)V$ for a one-dimensional conductor for which $\mathrm{d}n = N_{\mathrm{tot}}(k)(\mathrm{d}k/\mathrm{d}E)\mathrm{d}E$, with $N_{\mathrm{tot}}(k) = 1/\pi$ for electrons flowing in the direction of the applied bias. The total current can be expressed as $I = (2e^2/h)MV$, if the number of transverse modes $M(E)$ is constant over the energy range $E_{\mathrm{Fr}} < E < E_{\mathrm{Fl}}$, the ballistic conductor having a conductance $G_{\mathrm{c}} = I/V = 2e^2M/h$ or, equivalently, a resistance

$$R_{\mathrm{c}} = 1/G_{\mathrm{c}} = \frac{h}{2e^2M} \cong \frac{12.9\,\mathrm{k\Omega}}{M}$$

In the collisionless propagating regime, the resistance can only be caused by the mismatch between the finite number of transverse modes that are allowed to propagate in the ballistic conductor and the infinite number of transverse modes carried by contacts. The resistance is called contact resistance because this mismatch appears at the conductor/contact interface. R_{c} decreases with the number of transverse modes (energy subbands) in the ballistic conductor.

Unlike the conductance $G = \sigma S/L$ in bulk materials, where σ is the conductivity of the material, L the

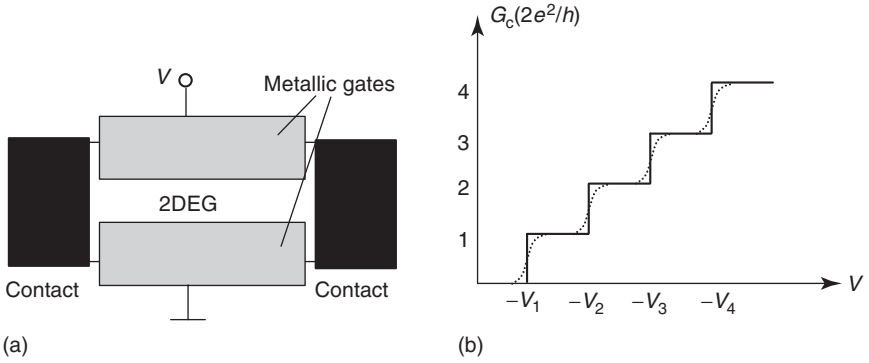


Figure 3 (a) Constriction with variable width in a 2DEG and (b) the conductance dependence on the gate voltage at zero temperature (solid line) and finite temperatures (dotted line).

length of the sample, and S its cross section, the conductance of ballistic structures does not depend on the length of the conductor. (Note that an unjustified extension of the domain of validity of the conductance formula in bulk materials would imply that in ballistic conductors, G becomes extremely large as L decreases.) In ballistic conductors, G_c depends however on S (the width of the one-dimensional conductor), since the number of transverse modes occupied by electrons propagating with the Fermi wave number can be estimated from $M = \mathrm{Int}[k_F S/\pi]$, where Int denotes the integer value. This dependence can be evidenced by measuring G_c of a ballistic conductor delimited from a 2DEG by a pair of metallic gates (**Figure 3a**), the conductor width depending on the negative voltage $-V_M$ applied on the gates. A step-like increase of the conductance with $2e^2/h$, as shown in **Figure 3b**, is observed at low temperatures and/or for confined electrons with large spacings between the energy levels each time M increases by one. This discontinuity is otherwise "smoothed" by a thermal motion.

Landauer Formula

The formula derived above for the conductance of ballistic structures assumed implicitly that electrons injected by the left contact are transmitted with unit probability to the right contact. This is not always the case. In particular, a partial (nonunity) transmission of electrons from one contact to the other occurs if the ballistic conductor is composed of several parts that differ in their width or potential energy. The conductance, in this case, can be easily calculated if the conductor with a transmission probability T is connected to two large reflectionless contacts by ballistic leads each of which have M transverse modes. In the model presented in **Figure 4**, T is the average probability that an electron injected in lead 1 is transmitted to lead 2, the conductance measured between the contacts at zero-temperature being given by the Landauer formula

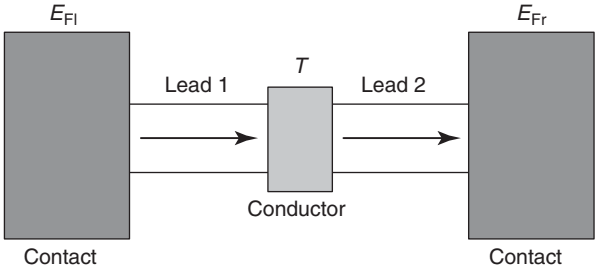


Figure 4 Model of a ballistic conductor with transmission probability T between the contacts.

$$G = \frac{2e^2}{h} MT$$

This expression can be viewed as a mesoscopic version of the Einstein relation $\sigma = e^2 \rho D$, if the conductivity σ is replaced by G, the density of states ρ is replaced by M, and the diffusion constant D is replaced by T. The total resistance $R = h/(2e^2MT)$ measured between the contacts is the sum between the contact resistance $R_c = h/(2e^2M)$ and the resistance of the "scatterer" element with transmission T, $R_s = h(1-T)/(2e^2MT)$, that would be measured between the two leads.

This identification allows the calculation of the resistance of multiple scatterers with transmission probabilities T_i connected in series as $R_s = \sum_i R_{si}$, with $R_{si} = h(1-T_i)/(2e^2MT_i)$, or the calculation of the total transmission probability as $(1-T)/T = \sum_i (1-T_i)/T_i$. This addition law, which can alternatively be obtained by adding the contribution of successive partial transmitted waves, is an expression of the coherent nature of the electron wave function in ballistic conductors.

At finite temperatures, when the Fermi–Dirac distribution is no longer step-like, electron transport through multiple energy channels takes place in the energy range $E_{Fr} - \Delta E < E < E_{Fl} + \Delta E$, where ΔE is of the order of a few $k_B T$. Since both the number of transverse modes M and the transmission probability T generally depend on E, the net current flow between the left and right contacts in the absence of inelastic scattering is

$$I = \frac{2e}{h} \int M(E) T(E) [f_l(E) - f_r(E)] \mathrm{d}E$$

where $f_l(E)$ and $f_r(E)$ are the Fermi–Dirac quasi-distribution functions in the left and right contacts.

Büttiker Formula

The expression for the net current flow or for the conductance of a ballistic structure can be generalized to account for the existence of several contacts or terminals. This is the case, for example, of four-terminal measurements (see **Figure 5**), where two additional terminals are used to measure the voltage drop along a conductor besides the two terminals through which the current flows. Modeling the additional floating or zero-external current terminals as scatterers characterized by transmission probabilities, the current through the pth terminal at zero-temperature is determined by the Büttiker formula

$$I_p = \frac{2e^2}{h} \sum_q [\bar{T}_{qp} V_p - \bar{T}_{pq} V_q] = \sum_q [G_{qp} V_p - G_{pq} V_q]$$

where $G_{pq} = (2e^2/h)\bar{T}_{pq}$ is the conductance associated to the electron transfer from terminal q with a quasi-Fermi energy level $E_{Fq} = eV_q$ to terminal p with a quasi-Fermi energy level $E_{Fp} = eV_p$. $\bar{T}_{pq}$ is, in general, the product between the number of modes M and the transmission probability per mode T_{pq} at the Fermi energy. The requirement that the net current flow is zero at equilibrium implies that $\sum_q G_{qp} = \sum_q G_{pq}$.

At finite temperatures, the net current flow through the pth terminal is given by

$$I_p = \frac{2e}{h} \int \sum_q [\bar{T}_{qp}(E) f_p(E) - \bar{T}_{pq}(E) f_q(E)] \mathrm{d}E$$

Transmission Probability Calculation

Ballistic electron transport is determined by the transmission probability between the leads that connect the conductor to several contacts. There are many methods to calculate the transmission probability: the scattering matrix formalism, the Green's function approach, the Kubo formalism, or the transfer Hamiltonian formalism. The most widespread for its ease of application is the matrix formalism which, in its simplest case, relies on approximating the spatially varying potential profile, electron effective mass, and/or dimensions of the conductor as a succession of regions in which this variation is step-like.

The application of the matrix method to ballistic conductors with a constant cross section is simplified by the absence of electron scattering from one transverse mode to another. Then, for a one-dimensional conductor composed of several regions with different but constant potential and electron effective mass, the solution of the Schrödinger equation in the ith region, $\Psi_i(x) = A_i \exp(\mathrm{i}k_i x) + B_i \exp(-\mathrm{i}k_i x)$, can be expressed as a superposition of forward- and backward-propagating waves in the x direction with wave number $k_i = \hbar^{-1}\sqrt{2m_i(E - V_i)}$ (see **Figure 6**). Inside each layer the wave function and its x-derivative must be continuous, whereas boundary conditions impose the continuity of the wave function and $\nabla\Psi \cdot \hat{x}/m^{\alpha+1}$ at the interface between two

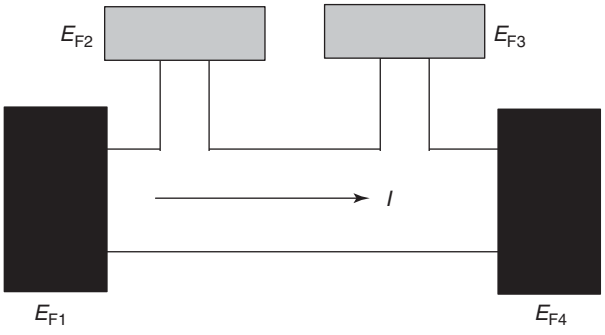


Figure 5 Four-terminal configuration: the current flows between the first and the fourth terminal, the second and third being used to measure the voltage drop along the conductor.

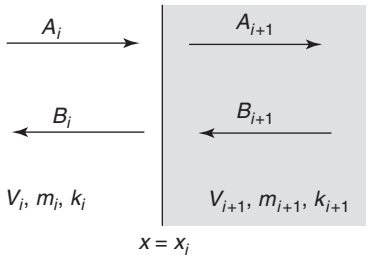


Figure 6 Electron wave function forward- and backward-propagating components at an interface between layers i and $i + 1$ with different characteristics.

layers, where $\hat{x}$ is the unit vector (the versor) along the x direction.

The matrix that expresses the electron wave function form at different x values depends on which parameters are related. For example, when $\alpha = 0$, a transfer matrix at the interface between the layers i and $i+1$ situated at $x = x_i$ can be defined as

$$\begin{pmatrix} A_i \exp(\mathrm{i}k_i x_i) \\ B_i \exp(-\mathrm{i}k_i x_i) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} (1+v_{i+1}/v_i) & (1-v_{i+1}/v_i) \\ (1-v_{i+1}/v_i) & (1+v_{i+1}/v_i) \end{pmatrix} \times \begin{pmatrix} A_{i+1} \exp(\mathrm{i}k_{i+1} x_i) \\ B_{i+1} \exp(-\mathrm{i}k_{i+1} x_i) \end{pmatrix}$$

where $v_i = \hbar k_i / m_i$ is the electron velocity in the ith layer, whereas at free propagation across the ith layer, between planes $x = x_{i-1}$ and $x = x_i$ the transfer matrix is diagonal, with elements $\exp[\mathrm{i}k_i(x_i - x_{i-1})]$ and $\exp[-\mathrm{i}k_i(x_i - x_{i-1})]$. The total transfer matrix with elements M_{pq}, $p, q = 1, 2$ is obtained by multiplying the corresponding matrices at each interface and across each layer, and the transmission probability across the whole structure, from the first to the last, Nth layer is $T = v_N |A_N|^2 / (v_1 |A_1|^2) = v_N / (v_1 |M_{11}|^2)$.

The matrix that relates the outgoing field amplitudes $B_i \exp(-\mathrm{i}k_i x_i)$, $A_{i+1} \exp(\mathrm{i}k_{i+1} x_i)$ to the incoming field amplitudes $A_i \exp(\mathrm{i}k_i x_i)$, $B_{i+1} \exp(-\mathrm{i}k_{i+1} x_i)$ on either side of an interface is called the scattering matrix. For free propagation across the ith layer the scattering matrix is antidiagonal, with elements $\exp[-\mathrm{i}k_i(x_i - x_{i-1})]$ and $\exp[\mathrm{i}k_i(x_i - x_{i-1})]$. In terms of the scattering matrix elements S_{pq}, $p, q = 1, 2$ of a succession of N regions, the total transmission probability can be expressed as $T = v_N |S_{21}|^2 / v_1$. In general, the elements of the scattering matrix are $S_{11} = r$, $S_{12} = t'$, $S_{21} = t$, $S_{22} = r'$, with r, t, r', t' the reflection and transmission amplitudes of the structure from left to right and right to left, respectively.

For a ballistic conductor with a variable cross section, there is scattering between transverse modes, which is not caused by collisions but by the transverse mode matching imposed by the electron wave function continuity. The difference from the former case is that not only do the electron wave numbers k_i in the adjacent regions differ, but so do the number of transverse modes M_i. More precisely, the electron wave function propagating along x has in the ith layer the expression $\Psi_i(x, y, z) = \sum_{k=1}^{M_i} [A_{ik} \exp(\mathrm{i}k_i x) + B_{ik} \exp(-\mathrm{i}k_i x)] \phi_{ik}(y, z)$, where $\phi_{ik}(y, z)$ is the transverse part of the kth mode in the ith layer. The transmission probability has a more complicated expression, but is also obtained by imposing the boundary conditions at each interface between different layers.

Electron Tunneling

The matrix formulation is valid for either real or imaginary values of the wave numbers k_i. An imaginary wave number in at least one region i implies that the electron energy is smaller than the potential energy value in this region, so that electrons are forbidden to enter the ith region from a classical point of view. Such a region acts as a barrier for electron propagation. The propagation of ballistic electrons in regions with imaginary wave numbers can be paralleled with an evanescent propagation of electromagnetic waves. As for evanescent electromagnetic waves, the electron wave function decays exponentially inside the region with imaginary k_i and therefore, the transmission probability through this region vanishes unless its dimension along the propagation direction is smaller than the distance over which the electron wave function decays to zero. Thus, the transmission probability can still be significant for thin barriers; electrons can pass through classically forbidden regions, a phenomenon known as electron tunneling. Tunneling can also occur through a succession of barriers for electrons separated by quantum wells, that is, by regions with real wave numbers. In ballistic structures, quantum tunneling is a coherent phenomenon and thus generates, in this case, constructive or destructive interference between partially reflected and transmitted waves. Then, resonances, which correspond to high-transmission probability values, can appear similarly to the occurrence of high intensity values in light diffraction due to interference between coherent light beams.

Coulomb Blockade

In quantum dots, more than in any other confined structure, a Coulomb interaction between tightly confined electrons leads to a high sensitivity of electronic states on the number of particles in the dot and to the appearance of specific phenomena such as the Coulomb blockade, which reflect the discrete nature of the electron charge. A Coulomb blockade has been observed in semiconductor quantum dots and small metallic clusters coupled to metallic leads through tunneling barriers. It consists in the opening of a gap in the energy spectrum at the Fermi energy when an electron is transferred at low temperatures from the reservoir into the dot or cluster (generically, island). This gap, caused by charge rearrangement and observed when the associated potential change is greater than the thermal energy $k_B T$, leads to the inhibition of further electron tunneling until this charging energy is compensated by an applied bias.

A quantitative description of the Coulomb blockade is simpler in metallic clusters, where size quantization effects are absent because the ballistic transport conditions are not met; in semiconductor quantum dots, the Coulomb blockade coexists with the quantization of energy levels. In either case, the Coulomb blockade can be understood in terms of a macroscopic capacitance associated with the system, if the electrons that tunnel through either junction relax in the dot. Modeling the thin tunnel junctions as a parallel combination of a tunneling resistance R_i and a capacitance C_i, $i = 1, 2$, as in **Figure 7**, a voltage source between the leads performs the work $W_s = eV(C_1 n_2 + C_2 n_1)/C$ to transfer n_2 charges through junction 2, and n_1 charges through junction 1. Here, V is the applied voltage and $C = C_1 + C_2$ is the capacitance of the island. The electrostatic energy stored in the capacitors is $E_s = (n_1 e)^2/2C_1 + (n_2 e)^2/2C_2 = [C_1 C_2 V^2 + (ne)^2]/2C$, with $n = n_1 - n_2$ the excess electrons accumulated on the island. Then, the change in the total energy $E(n_1, n_2) = E_s - W_s$ when a single particle tunnels through junction 1 is $\Delta E_1^{\pm} = E(n_1, n_2) - E(n_1 \pm 1, n_2) = (e/C)[-e/2 \mp (en + VC_2)]$, whereas the change in energy when a particle tunnels through the second junction is $\Delta E_2^{\pm} = E(n_1, n_2) - E(n_1, n_2 \pm 1) = (e/C)[-e/2 \pm (en - VC_1)]$. The system evolves at zero temperature only when $\Delta E_{1,2}^{\pm} > 0$, a condition which can only be satisfied for $n = 0$ if the voltage exceeds a threshold. Tunneling cannot occur below this threshold, and thus no current flows through the island because the Coulomb gap e^2/C that opens symmetric around the Fermi level at $V = 0$ and $C_1 = C_2 = C/2$ prevents tunneling, no states being available for electrons to tunnel into from the leads and no empty states existing for electrons in the island to tunnel out. This Coulomb blockade is a consequence of the additional Coulomb energy $e^2/2C$ that must be expended by an electron to tunnel in or out of the island, the Coulomb gap resembling the energy gap in semiconductors.

If a bias is applied now such that $V > e/C$, an electron can tunnel in the island from one lead, the Fermi energy in the island raises by e^2/C, and further tunneling is prohibited by the appearance of a new energy gap unless the voltage increases to $V > 3e/2C$ or the extra electron in the island tunnels out through the other junction, a case in which the island reverses to the $n = 0$ state. These correlated tunneling processes into and out of the island produce a net current flow. The I–V characteristic has a staircase-like behavior if the capacitances or the tunneling resistances of the two junctions are very dissimilar. The charging effects can be observed if $e^2/C \gg k_B T$ and $R_{1,2} \gg h/e^2$.

See also: Effective Masses; Meso- and Nanostructures; Nanostructures, Electronic Structure of; Tunneling Devices.

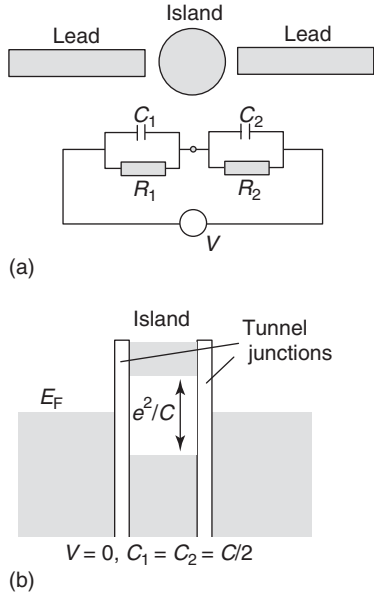


Figure 7 (a) Island connected to external leads through tunneling junctions and its equivalent model, (b) Opening of a gap in the energy spectrum at the Fermi energy when $V = 0$, $C_1 = C_2 = C/2$ due to the additional Coulomb energy that must be spent by an electron to tunnel in or out of the island.

PACS: **73.23. – b; 73.23.Ad; 73.40.Gk; 85.35.Be**

Further Reading

Datta S (1997) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.

Dragoman D and Dragoman M (1999) Optical analogue structures to mesoscopic devices. *Progress in Quantum Electronics* 23: 131–188.

Ferry DK and Goodnick SM (1997) *Transport in Nanostructures*. Cambridge: Cambridge University Press.

Nomenclature

C	capacitance
k_x, k_y, k_z	wave vector components along x, y, z
k_F	Fermi wave number
e	electron charge
E	electron energy
E_c	bottom of the conduction band
E_{kin}	kinetic energy of electrons
E_F	Fermi energy
E_{Fl}, E_{Fr}	quasi-Fermi levels in the left and right contacts
E_s	electrostatic energy

$E_{s,p}$	cut-off energy of the subband or transverse mode labeled by p
$\boldsymbol{F}$	electromagnetic field component
$f(E)$	Fermi–Dirac distribution function
$f_l(E)$, $f_r(E)$	Fermi–Dirac distribution functions in the left and right contact
G_c	conductance of a ballistic conductor
G_{pq}	ballistic conductance associated to the electron transfer from terminal q with a quasi-Fermi energy level E_{Fq} to terminal p with a quasi-Fermi energy level E_{Fp}
I	current
K_B	Boltzmann constant
L_{fp}	mean free path
L_{ph}	phase relaxation length
L_x, L_y, L_z	dimensions of the free space propagation region along x, y, z
m	electron effective mass
$M(E)$	number of subbands occupied by electrons with energy E
M_{pq}	transfer matrix elements
n	electron density per unit area
$N_{tot}(k)$, $N_{tot}(E)$	total number of spin-degenerate states with a wave number less than k, or with an energy less than E
r, r'	reflection amplitudes from left to right and right to left
R	resistance
R_c	contact resistance
S_{pq}	scattering matrix elements
t, t'	transmission amplitudes from left to right and right to left
T	temperature; also electron transmission probability
v	electron velocity
V	potential energy; also applied voltage
W_s	work
$\rho_{2D,p}$	density of states per unit area $S = L_xL_y$
$\rho_{1D}(k_x)$	density of states in the k_x plane per unit length
ρ_{0D}	density of states in quantum dots
ϑ	unit step function
θ	angle of propagation
$\Psi(x,y,z)$	envelope electron wave function

Bioelectronics

E Katz, The Hebrew University of Jerusalem, Jerusalem, Israel

Introduction

Bioelectronics is a rapidly progressing interdisciplinary research field that aims to integrate biomolecules and electronic elements, for example, electrodes, field effect transistors, or piezoelectric crystals, into functional systems. Surface engineering of biomaterials, such as enzymes, antigen–antibodies, or deoxyribonucleic acids (DNA), on the electronic supports controls the electrical properties of the biomaterial–transducer interface and enables the electronic transduction of biocatalyzed transformations and biorecognition events on the transducers. Bioelectronic sensing devices, biofuel elements, and biotemplated circuitries have been developed recently.

Bioelectronic Systems Based on Biocatalytic Redox Enzymes

Natural redox enzymes catalyzing oxidative or reductive transformations of organic molecules operate upon an electron transfer between the enzyme active center and the respective substrate molecules. For example, the enzyme glucose oxidase (GOx), which catalyzes oxidation of glucose, includes the active center (also called cofactor) flavin adenine dinucleotide (FAD) that accepts electrons from the substrate glucose and donates them to the natural acceptor O_2. If the enzyme is immobilized on a conductive support and an efficient electron transfer between them is provided, the electron transfer from the enzyme active center could be directed to the conductive support instead of the natural acceptor, thus producing an electronic signal. However, most of the redox enzymes deny direct electrical contact with electrodes. The electron-transfer theory developed by R Marcus implies that the rate constant of the electron transfer between an electron acceptor and a donor pair is given as follows, eqn [1]:

$$k_{et} \propto \frac{\exp[-\beta(d-d_0)] \cdot \exp(-\Delta G^0 + \lambda)^2}{4RT\lambda} \qquad [1]$$

In eqn [1], d and d_0 are the actual distance and the van der Waals distance, respectively, that separate the donor–acceptor pair; ΔG^0 and λ correspond to the free-energy change and the reorganization energy accompanying the electron-transfer process, respectively; and β is the electronic coupling constant.

Thus, theory predicts that the electron transfer rate constant between the donor and acceptor sites exponentially decreases upon increase of the distance separating the couple. As the enzyme active center and the conductive support may be considered as a donor–acceptor pair, the embodiment of the active centers in the protein matrices spatially insulates the active centers and blocks the electrical contact between the redox enzymes and conductive supports. The understanding of the fundamental limitations for the electrical contacting of the redox enzymes with electrodes enables the development of synthetic methods to electrically wire redox enzymes and electronic transducers. Realizing that the lack of electrical contact between redox enzymes and the electrodes originates from the distance separating the donor–acceptor couple, one may suggest that the shortening of the electron-transfer distances by intermediate electron relay units may facilitate the electron transport between the active centers and the electrode (**Figure 1**).

The electrical contacting of redox enzymes that defy direct electrical communication with electrodes can be established by using charge carriers as intermediates between the enzyme active center and the electrode. These artificial electron donor or acceptor molecules (in the case of reductive or oxidative enzymes, respectively), usually referred to as electron-transfer mediators, can be accepted by many redox enzymes in place of their natural oxidants or reductants. The redox enzymes were incorporated into polymer matrices consisting of electron-transfer mediator units (e.g., Os-complexes (**1**)) (**Figure 2a**), or the redox mediator molecules (e.g., ferrocene derivatives (**2**)) were covalently tethered to the enzyme molecules (**Figure 2b**) to provide electron transport between the active centers of the enzymes and electrodes. Although redox enzymes anchored to electrodes by these means

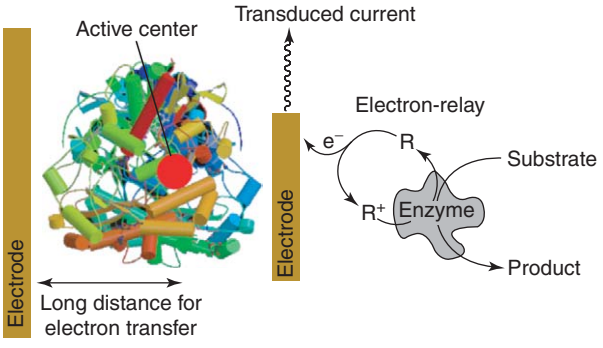


Figure 1 Spatially isolated redox active centers of enzymes could be electrically contacted on electrode surfaces with the use of electron relay molecules.

reveal electrical contact with the conductive support, the experimental electron transfer turnover rates are far lower than those observed for the enzymes with the natural donor or acceptor substrates. This was attributed to the random, nonoptimized, positioning of the relay units relative to the redox centers of the respective enzymes, that themselves lack a structural orientation in the integrated enzyme–electrode system. Thus, for optimal electrical communication between the redox centers of the enzymes and the electrodes, it is essential to develop means to align all enzyme molecules on the electrode in a configuration so that the redox centers are as close as possible to the electrode surface, and to place an appropriate electron relay unit between the redox center and the electrode support. The architecture of electrically contacted enzymes in an aligned configuration was accomplished by a surface-reconstitution method. The surface-confined cofactor unit (FAD, (**3**)) was incorporated into the GOx enzyme molecule instead of the natural cofactor, thus allowing the enzyme alignment on the electrode surface in a form of a stable complex. An electron-transfer mediator (e.g., pyrroloquinolino quinone (PQQ), (**4**)) located between the FAD-cofactor and the electrode surface provides efficient electron transfer from the enzyme FAD active center to the electrode (**Figure 3a**). Bioelectrocatalyzed oxidation of glucose to gluconic acid by the surface-reconstituted GOx enzyme results in the generation of an electronic signal in the electrode. Similarly, nicotine adenine dinucleotide (NAD^+)-dependent enzymes, for example, lactate dehydrogenase (LDH), were associated with a surface-confined NAD^+-cofactor (**5**) linked to the electrode surface via a catalyst molecule (PQQ) providing its fast electrochemical oxidation (**Figure 3b**). Since the NAD^+-cofactor is weakly bound to the respective enzymes, the enzyme molecules after formation of the surface-confined complex were additionally cross-linked to stabilize the biocatalytic interface. Bioelectrocatalyzed oxidation of lactate to pyruvate by the surface-integrated LDH enzyme results in the generation of an electronic signal in the electrode. Therefore, an efficient electron transport from the enzyme active centers (e.g., FAD or NAD^+) via the intermediate relay units (e.g., PQQ) to the conductive support was achieved, thus providing electronic signals generated by the biocatalytic systems. It should be noted that the relay units are chemically reduced and oxidized upon the electron-mediating process that could introduce some limitations to the rate of the process.

Application of conductive nanosized objects, such as metal nanoparticles or carbon nanotubes, for

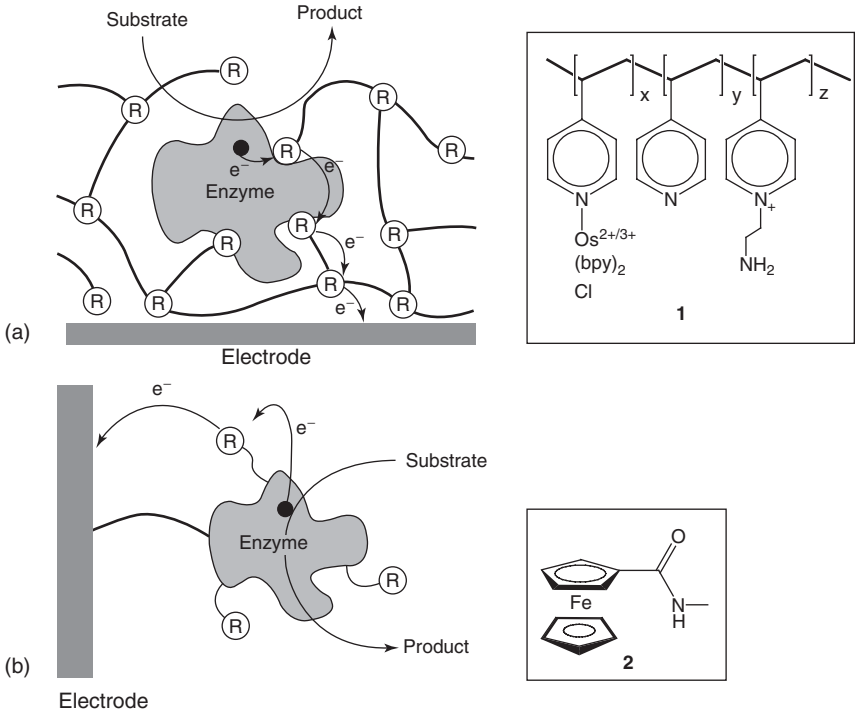


Figure 2 Redox enzymes electrically contacted on the electrode by means of: (a) the redox polymer (**1**) surrounding the enzyme molecules and (b) the redox mediator (e.g., ferrocene, **2**) tethered covalently to the enzyme molecules.

wiring of redox enzymes allowed the effective electron transfer between the enzyme active centers and the conductive supports using metal-like conductivity instead of the chemical redox-mediating process. Gold (Au) nanoparticles or single-walled carbon nanotubes (SWCNTs) were functionalized with the FAD-cofactor units and assembled on electrodes. Reconstitution of GOx enzyme molecules on the functionalized interfaces resulted in the biocatalytic interfaces, where the enzyme active centers are electrically wired to the conductive support via the Au nanoparticles (**Figure 4a**) or SWCNTs (**Figure 4b**). The nanosize of the wires allowed a short distance between them and the FAD cofactor, thus providing effective electron transfer, whereas the conductivity of the Au nanoparticles and SWCNTs provided the efficient electron transport to the macroconductive support.

The electric current provided by the enzyme–electrode systems depends on the rate of the biocatalytic reactions and could be proportional to the concentration of the substrate, if the entire process is limited by the substrate diffusion to the biocatalytic interface. Thus, the generated electronic signal could be used to design amperometric biosensors. For example, the GOx enzyme electrically contacted via an Au nanoparticle generates the current proportional to the glucose concentration in a certain concentration range.

The electrically wired redox enzymes could generate electrical power sufficient for construction of biofuel cells. The biofuel cells are primary sources of electrical energy extracted from organic materials using electrochemical processes catalyzed by biological substances (e.g., redox enzymes). They include two enzyme–biocatalytic electrodes: the first electrode (anode) provides oxidation of the organic substrate by an enzyme (e.g., oxidation of glucose by GOx), thus injecting electrons into the external electrical circuit, whereas the second electrode (cathode) is used to withdraw electrons from the circuit and transfer them to a chemical acceptor (e.g., O_2). The tailoring of efficient electron transfer at the enzyme-modified electrodes could enable specific biocatalytic transformations that compete kinetically with any chemical reaction of the electrodes or of the biocatalysts with interfering substrates (e.g., substrate transport from the counter compartment, oxygen). This would enable the design of noncompartmentalized biofuel cells, where the biocatalytic anode and cathode are immersed in the same phase with no separating membrane. In a working example, an anode consisting of GOx reconstituted onto a PQQ–FAD monolayer for the biocatalyzed oxidation of glucose was coupled to a cathode composed of an aligned cytochrome *c*/cytochrome oxidase (Cyt *c*/COx) couple that catalyzes

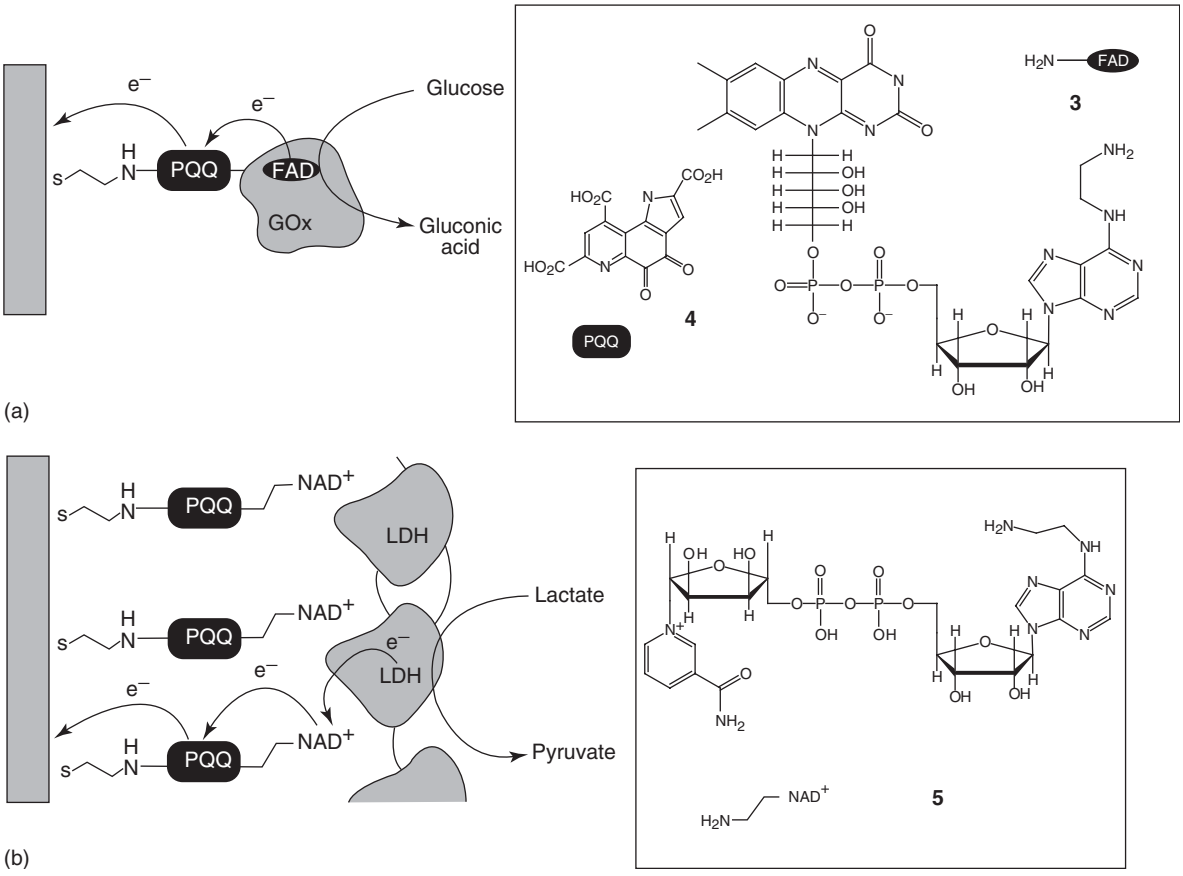


Figure 3 Redox enzymes electrically contacted on the electrode by means of: (a) reconstitution of the GOx on the electrode surface functionalized with the FAD (**3**)-cofactor and PQQ (**4**)-mediator; and (b) integration of LDH with the NAD^+ (**5**)-cofactor and the PQQ-mediator.

the reduction of O_2 to water (**Figure 5**). Since the reconstituted GOx provides extremely efficient biocatalyzed oxidation of glucose that is unaffected by oxygen, the anode can operate in the presence of oxygen. Thus, the biofuel cell uses O_2 as an oxidizer and glucose as a fuel without the need for compartmentalization. The glucose–O_2 biofuel cell element, or similar cells, might act as implanted self-fueled electrical energy generation devices in living organisms providing power for other bioelectronic elements.

Redox enzymes could be integrated with a gate of a field-effect transistor. An ion-selective field-effect transistor (ISFET) operates with a gate open to a chemical system, which provides electrical signals governing the current between source and drain electrodes. This signal is a potential change generated by chemical means on the gate. The redox enzymes associated with the ISFETs gate could reduce or oxidize co-immobilized electron relay units, thus providing the potential change on the gate further amplified by the respective change of the source–drain current (**Figure 6**). The amplified electronic signal is dependent on the rate of the biocatalytic reaction provided by the enzyme and proportional to the enzyme–substrate concentration, thus allowing potentiometric biosensors. Nano-sized bio-ISFETs were constructed with a single SWCNT operating as a gate modified with a few biomolecules (e.g., redox enzymes) providing a signal further electronically amplified by the source–drain current.

Bioelectronic Systems Based on Protein or DNA Biorecognition Systems

Biomolecules could provide very specific recognition processes resulting in the formation of complexes between them. For instance, biological immunosystems protecting organisms from foreign species are based on affinity complex formation between antibodies (recognition proteins produced by the organism) and antigens (foreign target molecules).

Biological genetic information is saved in deoxyribonucleic acid (DNA) molecules producing complexes with complementary DNA molecules. Association of a biorecognition element such as antigen (or antibody) or a DNA molecule with an electronic transducer allows electronic transduction of the affinity complex formation with the respective complementary biomolecule. The bioaffinity complex formation in a thin film near the conductive or semiconductive surface of the electronic transducer yields a chemically modified film on the surface that alters its interfacial properties. Blocking the electrode surface with bulky biomolecules upon immunorecognition between antigens and antibodies (**Figure 7a**), or formation of the charged interface upon hybridization of negatively charged DNA molecules (**Figure 7b**), perturbs the double-charged layer existing at the electrode–electrolyte interface resulting in the increase of its thickness, δ_{dl}, and in the insulation of the electrode surface with respect to redox labels added to the solution. This results in the capacitance change and electron-transfer resistance change at the interface, respectively, that could be read out by impedance spectroscopy. Two main approaches were used to increase these interfacial changes in order to amplify the electronic signal generated upon the biorecognition events: (1) secondary binding of bulky units (proteins, functionalized liposomes, or nanoparticles) complementary to the primarily biorecognition complex formed on the interface (**Figure 8a**), and (2) labeling the biorecognition elements with biocatalytic entities which provide a cascade of reactions on the surface, for example, resulting in the precipitation of an insoluble material on the sensing surface (**Figure 8b**). Both the amplifying processes are triggered by the primary

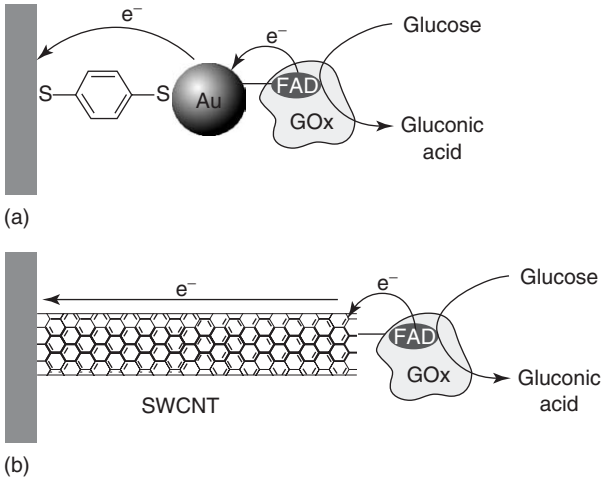


Figure 4 Redox enzymes electrically contacted on the electrode by means of: (a) reconstitution of the GOx on the electrode surface functionalized with the FAD-cofactor and an Au-nanoparticles plugging into the enzyme; and (b) reconstitution of the GOx on the electrode surface functionalized with the FAD-cofactors covalently linked to conductive SWCNTs assembled on an Au electrode surface.

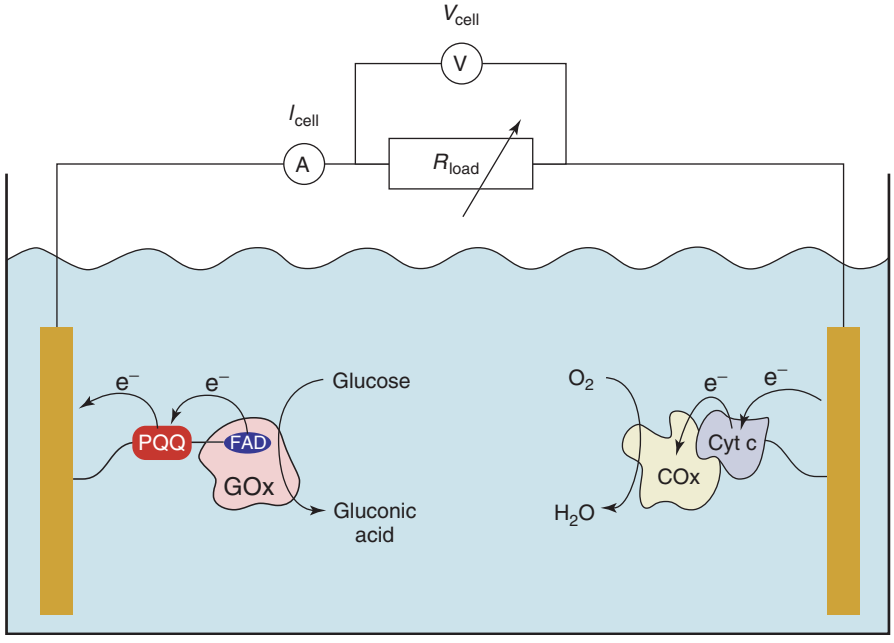


Figure 5 A noncompartmentalized biofuel cell, which generates electrical energy by the oxidation of glucose and reduction of O_2 biocatalyzed by the GOx reconstituted on the anode and Cyt *c*/COx-assembly on the cathode, respectively.

recognition event and cannot proceed in the absence of the antigen–antibody or DNA–complementary DNA complex at the surface. Combination of high specificity and sensitivity provided by the primary biorecognition process and the secondary biocatalytic amplification, respectively, with the electronic transduction of the biochemical signal allows bioelectronic immunosensing and DNA sensing.

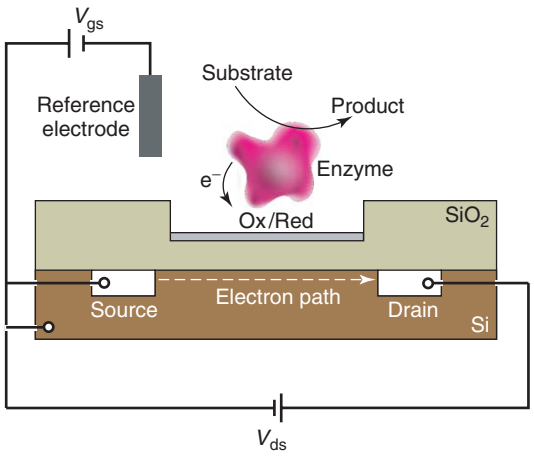


Figure 6 Redox-enzyme integrated with a gate of a field-effect transistor (V_{gs} – gate-source voltage; V_{ds} – drain-source voltage; Ox/Red – a redox-relay associated with the redox-enzyme).

Single-base DNA mismatches analysis, which is important for the analysis of genetic disorders, could exemplify the bioelectronic DNA sensing. The process is exemplified by the identification of the DNA mutant (**6**), in which the T-base in the normal gene (**6a**) is substituted with the G-base (**Figure 9**). A thiolated DNA primer (**7**) complementary to the analyzed gene up to one base before the mutation site is assembled as the sensing interface on an Au electrode. The normal gene (**6a**), as well as the mutant (**6**), associate with the sensing interface to generate the double-stranded DNA configuration. Reaction of the resulting interface with biotinylated-deoxycytosine triphosphate (dCTP) base in the presence of a polymerase enzyme results in the coupling of the biotinylated-tagged-base only to the layer which includes the mutant site. Subsequent association of the avidin-alkaline phosphatase (AlkPh)

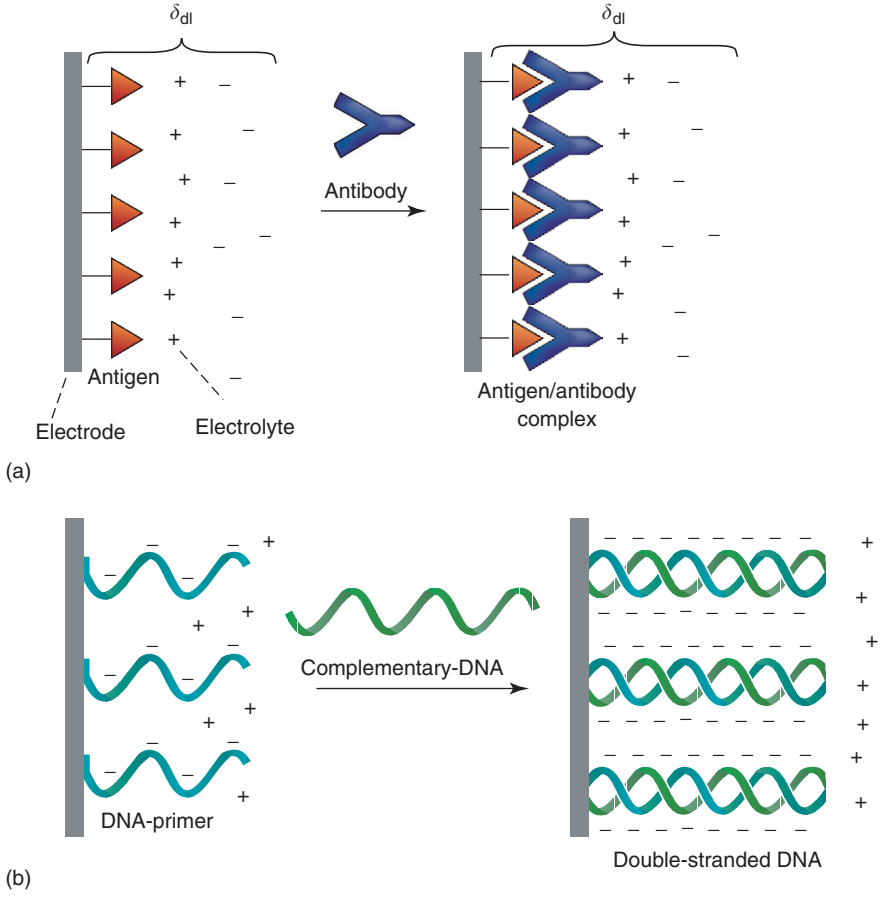


Figure 7 Alteration of interfacial properties of the transducer surface upon: (a) binding of bulky biological molecules (e.g., antibody) to the biorecognition elements (e.g., antigen) on the surface; (b) formation of the charged complexes (e.g., double-stranded DNA) on the transducer surface.

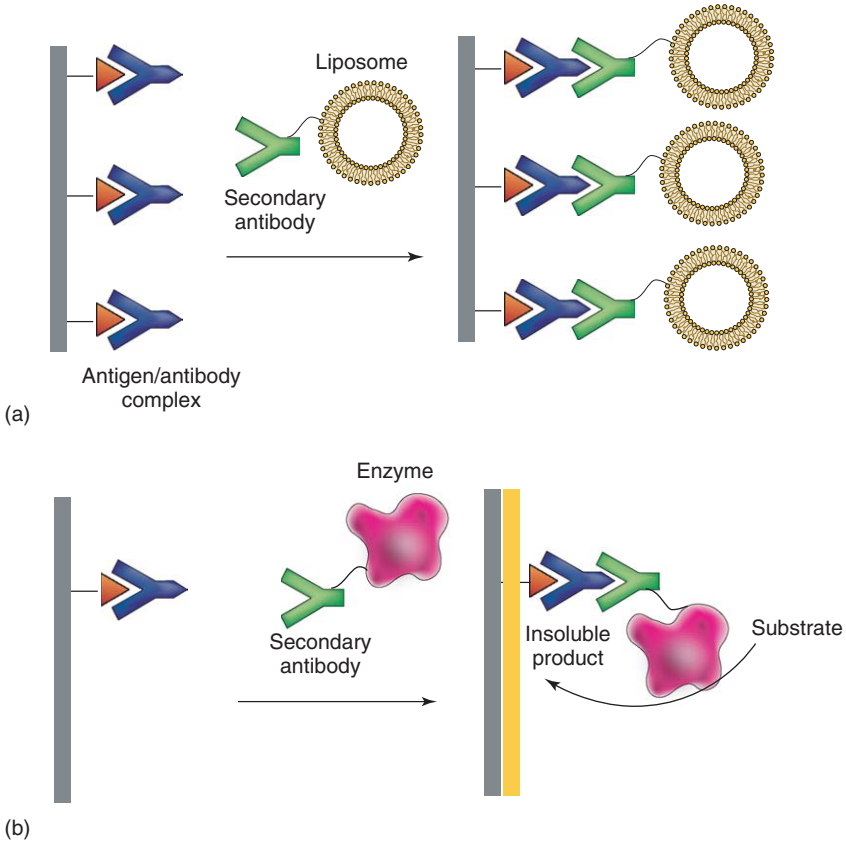
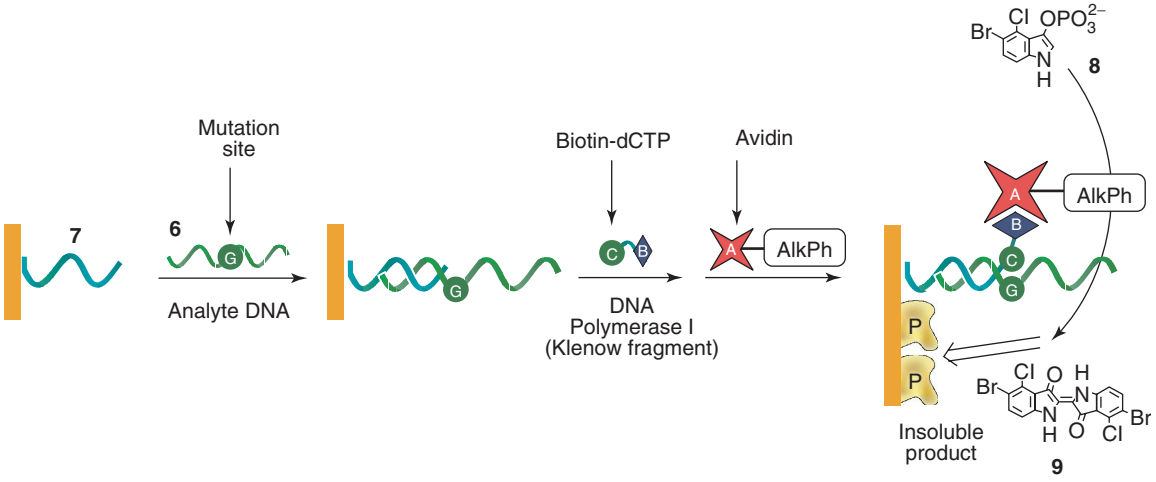


Figure 8 Amplification of the electronic read-out signal from the primary biorecognition events achieved by the enhancement of the interfacial changes upon: (a) secondary biorecognition processes; and (b) biocatalytic reactions yielding insoluble materials on the interface.



Mutation site

(**6**) 5′ - CTT TTC TTT TCT TTT **G**GA TCC GCA AGG CCA GTA ATC AAA CG - 3′

(**6a**) 5′ - CTT TTC TTT TCT TTT **A**GA TCC GCA AGG CCA GTA ATC AAA CG - 3′

(**7**) 5′ - HS-$(CH_2)_6$-CGT TTG ATT ACT GGC CTT GCG GAT C-3′

Figure 9 The detection of a single-base mismatch in DNA amplified by a biocatalytic reaction on the transducer interface and read-out by impedance spectroscopy.

biocatalytic conjugate, which stimulates the oxidative hydrolysis of 5-bromo-4-chloro-3-indoyl phosphate (**8**) to the insoluble indigo product (**9**), provides a route for the amplified transduction of the existence of the mutation site in (**6a**). Faradaic impedance spectroscopy was used as an electronic means to transduce the biochemical signal.

Another approach to the bioelectronic sensing of the recognition events uses incorporation of redox active species into the bioaffinity complex generated on a conductive transducer surface. Redox molecules, which are able to exchange electrons with the electrode support, are used as labels upon hybridization of complementary DNA molecules (**Figure 10a**), or upon biocatalytic replication of a long DNA sequence (**Figure 10b**). For example, the amplified amperometric transduction of DNA sensing with a bioelectronic system was demonstrated by the generation of a redox active DNA strand complementary to the analyzed DNA, which activates a secondary bioelectrocatalytic process (**Figure 10b**). The complex generated between the 7249-base M13 phage viral

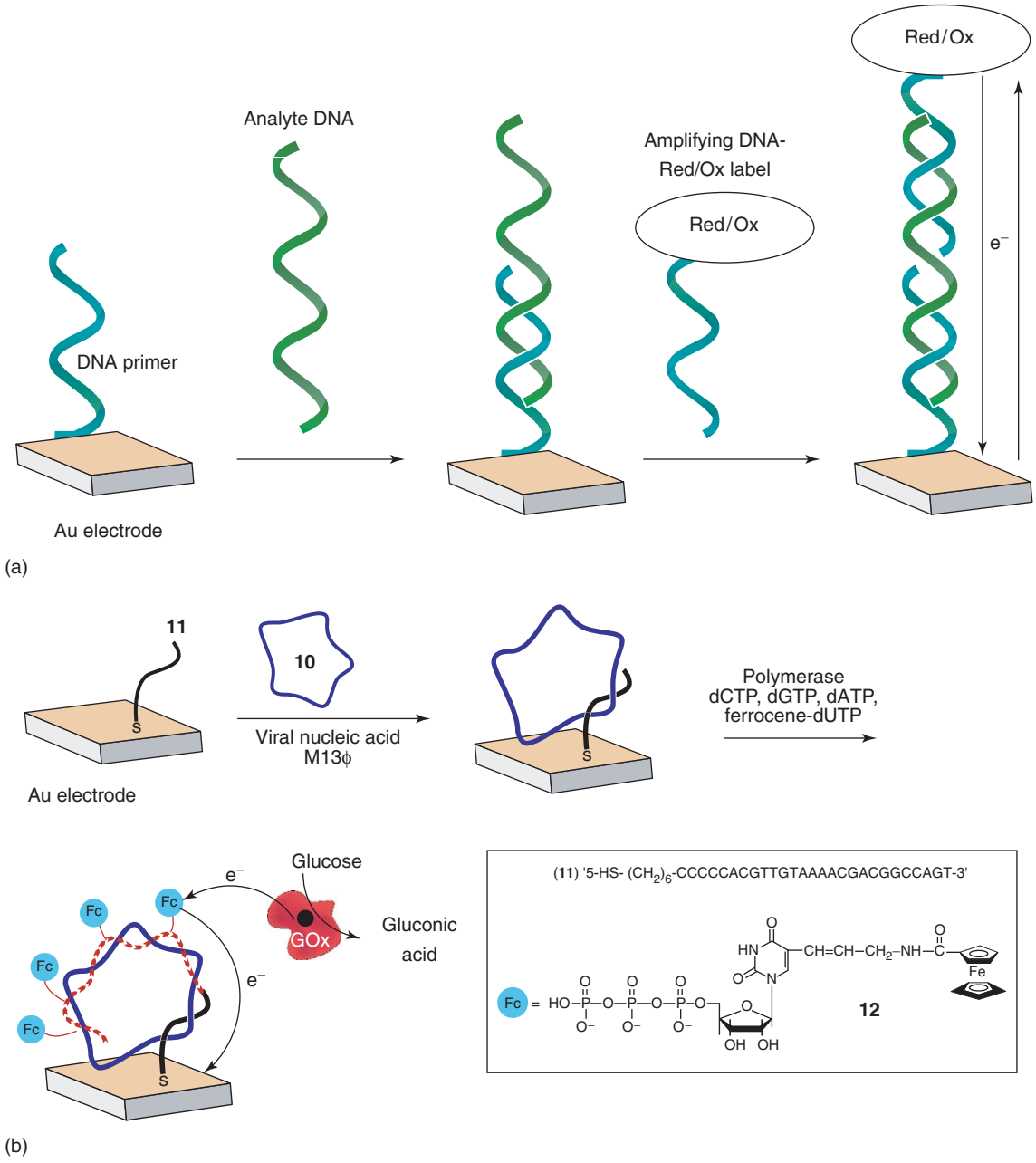


Figure 10 Labeling DNA molecules with redox units for amperometric read-out of the DNA reactions on the transducer interface. (a) Hybridization of complementary DNA molecules and (b) biocatalytic replication of DNA.

DNA (**10**) and a 27-base primer (**11**) associated with an Au electrode was replicated in the presence of a polymerase enzyme and a mixture of nucleotides that includes ferrocene-modified deoxyuridine triphosphate (dUTP) (**12**). The resulting ferrocene-labeled DNA replica was then coupled to GOx and, upon electrochemical oxidation of the ferrocene units, the GOx-biocatalyzed oxidation of glucose was activated. Because the enzyme exhibits a high turnover rate, the electrical output from this process represents an amplified electrical signal of the primary recognition event between the M13 phage viral DNA and the sensing nucleic acid associated with the electrode.

Metal nanoparticle labels linked to the biorecognition components (e.g., DNA molecules) were electrochemically oxidized, thus generating the electronic signal reporting on the formation of a bioaffinity complex on the sensing interface. Catalytic features of metal nanoparticles that enable the electroless deposition of metals on the nanoparticle clusters allow the enlargement of the particles, thus providing further amplification of the read-out signal.

Switchable and Tunable Bioelectronic Systems

Bioelectrocatalytic and biorecognition assemblies on electronic transducer interfaces that can be internally "tuned" or turned on and off during their operation are highly desirable. The use of such interfaces could allow the construction of variable output biofuel cells or variable range biosensors. Various physical input signals such as light, magnetic field, or electrical potential applied on the bioelectronic interfaces have been used to switch and tune bioelectronic transduction of the generated chemical signals.

For instance, photoswitchable biomaterials open a route to optobioelectronic systems. Photoisomerizable molecules undergo reversible structural changes when exposed to specific energies of light. Their two states often differ considerably, for instance, in their charge, polarization, or shape. If an enzyme molecule is functionalized with photoisomerizable groups, then the local environment of that enzyme will depend on the state of those groups. As the conformation of the enzyme is very sensitive to its environment, the shape and therefore the activity of the enzyme may be controlled by photonic input. To optimize the photoswitchable bioelectrocatalytic features of redox enzymes, the site-specific functionalization or mutation of the active-site microenvironment is essential. This has been accomplished, for example, by the reconstitution of GOx with a photoisomerizable FAD-cofactor. Nitrospiropyran carboxylic acid was covalently coupled to N^6-(2-aminoethyl)-FAD to yield a photoisomerizable nitrospiropyran-FAD cofactor (**13**). The native FAD cofactor was extracted from GOx and the modified FAD cofactor (**13**) was reconstituted into the apoglucose oxidase (apo-GOx) (**Figure 11a**). This reconstituted enzyme includes a photoisomerizable unit directly attached to the redox center of the enzyme, and hence the enzyme is predisposed for optimized photoswitchable bioelectrocatalytic properties. The photoisomerizable enzyme was assembled on an Au-electrode as outlined in **Figure 11b**. The bioelectrocatalytic interface was reversibly switched "off" and "on" by light signals resulting in photoisomerization of the molecular unit bound to the FAD cofactor between spiropyran (**13**) and merocyanine (**14**) isomeric states, respectively. The electron transport between the photoisomerizable enzyme cofactor and the electrode surface was achieved with the use of a diffusional electron-transfer mediator (e.g., ferrocene monocarboxylic acid (**15**)).

The distance between surface-confined redox proteins and the bioelectronic transducer surface could be controlled by an applied electrical potential. Biasing the electrode surface at potentials more positive than a zero-charge potential results in the positive charging of the surface, thus repelling positively charged protein molecules (e.g., cytochrome *c*). The longer distance, which is allowed by a long flexible molecular spacer, results in a slow kinetics of the electron transfer between the electrode and the protein redox active center. Opposite biasing yielding a negative electrode surface results in a short distance upon electrostatic attraction of the positively charged protein, thus providing a short electron-transfer distance and the fast electron exchange (**Figure 12**).

The magnetic control of bioelectrocatalysis is a novel concept of bioelectronics. **Figure 13** outlines the method to control by an external magnetic field the bioelectrocatalytic oxidative functions of a redox enzyme, for example, GOx. Magnetic particles (e.g., Fe_3O_4) functionalized with a relay unit, R (e.g., a ferrocene unit), are used to electrically contact the redox enzyme (e.g., GOx) and the electrode. Positioning of a magnet below the electrode attracts the functionalized-magnetic particles to the electrode. This enables the oxidation of the relay unit by the conductive support, and subsequently activates the bioelectrocatalyzed oxidation of the substrate (e.g., glucose) by the oxidized relay associated with the magnetic particles. Transfer of the external magnet to an upper position lifts the functionalized-magnetic particles upward, and removes them from the electrode support. This prevents the oxidation of

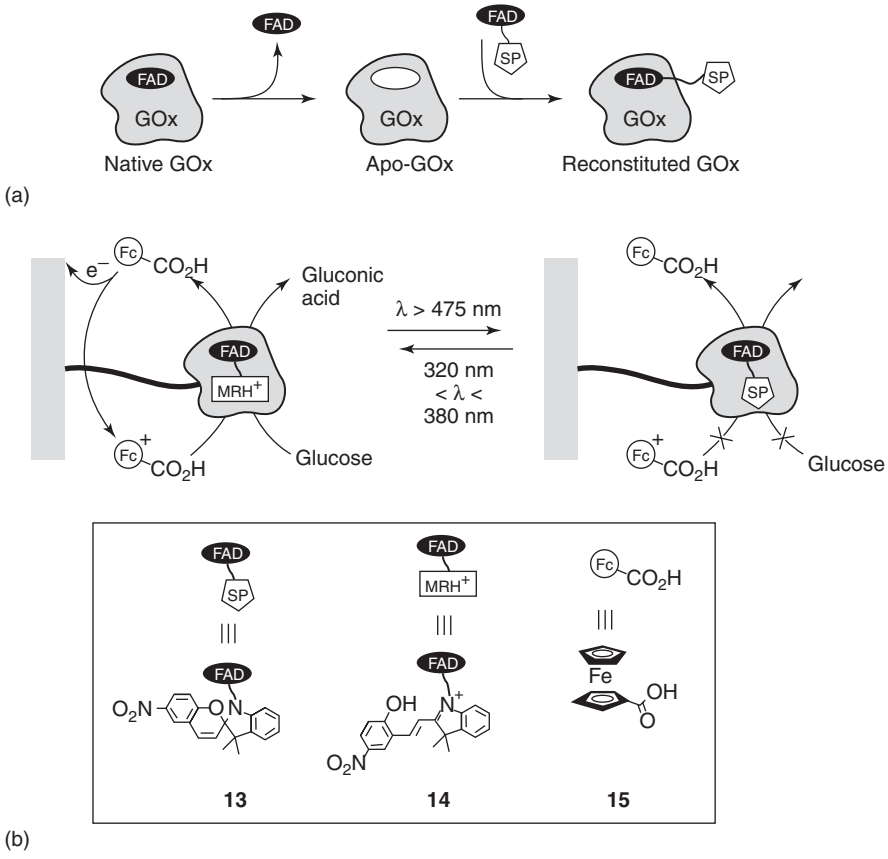


Figure 11 (a) Implanting of a photoisomerizable FAD derivative into apo-GOx to yield a photoisomerizable enzyme. (b) Switching "on" and "off" the photoisomerizable enzyme by light signals.

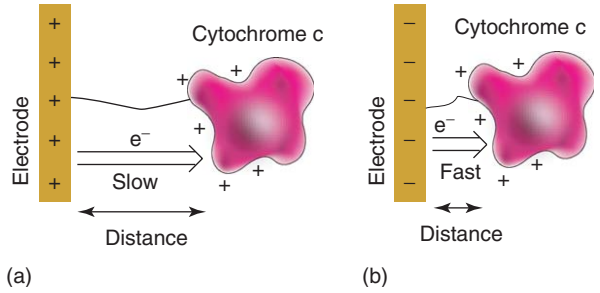


Figure 12 Potential-induced changes of the protein-surface distance resulting in the alteration of the electron transfer kinetics.

the relay units and the bioelectrocatalyzed oxidation of the substrate is switched off. By alternate positioning of the external magnet below and above the electrode surface, the bioelectrocatalyzed oxidation of the substrate is reversibly switched between "on" and "off" states, respectively. The magnetic control of a bioelectrocatalytic reduction process can be similarly controlled by the use of relay-functionalized magnetic particles that are reduced at the electrode support.

An important advance in the magnetic control of electrocatalytic and bioelectrocatalytic transformations was accomplished by magnetic attraction of the magnetic particles to the electrode support followed by the rotation of the magnetic particles on the electrode by means of an external rotating magnet. The rotation of the redox-functionalized magnetic particles turns them into circularly rotating microelectrodes. As a result, the redox-activated bioelectrocatalytic process mediated by the functional particles is controlled by convection rather than diffusion of the substrate to the microelectrodes. Accordingly, enhanced amperometric responses of the particle-mediated bioelectrocatalytic processes are anticipated, and the resulting currents should be controlled by the rotation speed of the particles. **Figure 14** depicts the amplified amperometric analysis of glucose by the rotation of ferrocene (**16**)-functionalized magnetic particles on the electrode support. The rotating magnetic particles are functionalized with an electron-transfer mediator, and the interaction with GOx and glucose is convection controlled. The bioelectrocatalysis in the presence of rotating magnetic particles can be used to enhance biosensing events and to yield the amplified electronic signal.

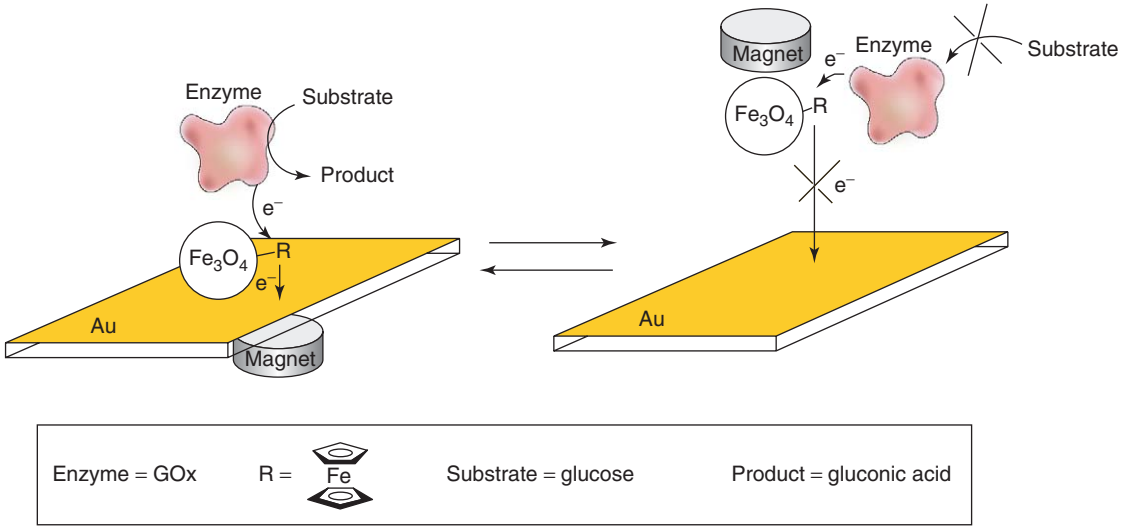


Figure 13 Magneto-switched bioelectrocatalyzed oxidation of glucose in the presence of electron relay-functionalized magnetic particles.

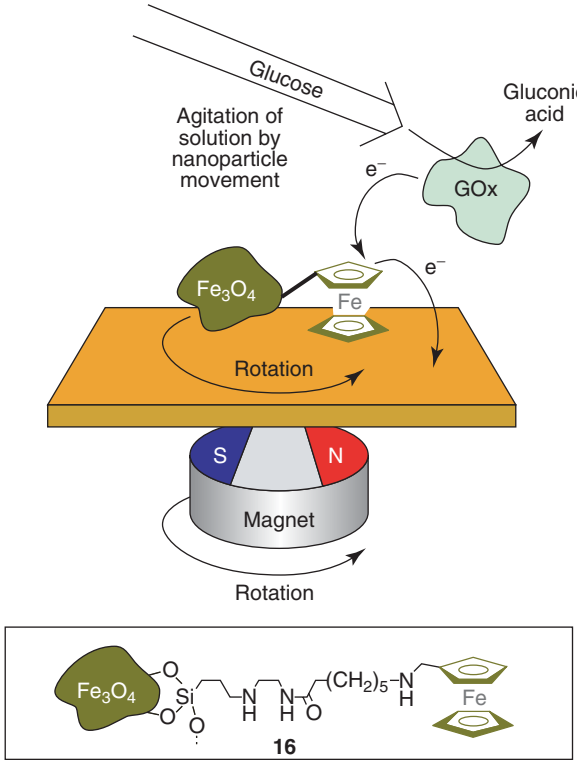
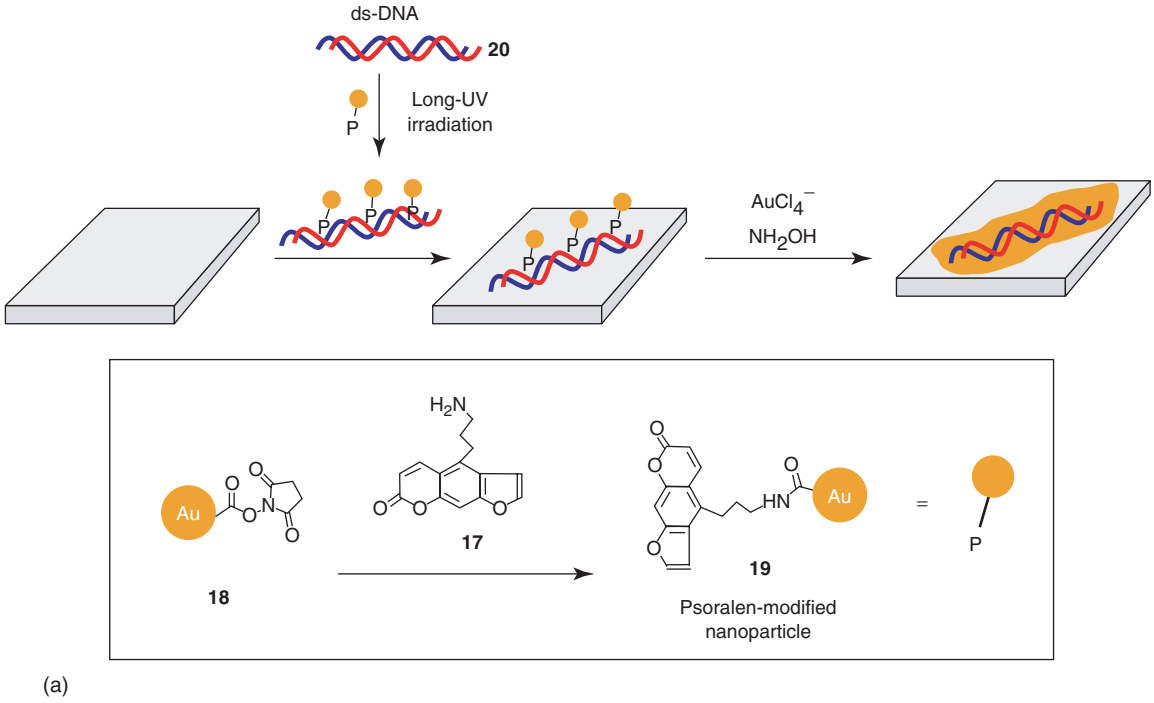


Figure 14 Bioelectrocatalytic oxidation of glucose in the presence of GOx and ferrocene-functionalized magnetic particles enhanced upon circular rotation of the particles by means of an external rotating magnet.

Biomaterial-Based Nanocircuitry for Bioelectronic Systems

The miniaturization of objects by lithographic methods reaches its theoretical limits. It is generally accepted that different miniaturization methods need to be developed in order to overcome this barrier. While the lithographic methods use "top" to "down" miniaturization of patterns, the alternative approach of the "bottom-up" construction of objects has been suggested as a means to overcome the lithographic limitations. That is, the construction of objects on molecular or supramolecular templates could generate nanometer-size features. Nanowires are considered as building blocks for self-assembling logic and memory circuits in future bionanoelectronic devices.

Different biomaterials, for example, proteins and nucleic acids, may act as important building blocks for functional bionanocircuitry, and eventually may provide bionano-elements for the construction of nanodevices. Biological templates have been used to direct the nucleation, deposition, and assembly of inorganic micro- and nanostructures. Among the different biomaterials, DNA is of specific interest as a template for the construction of nanocircuitries. Several arguments support the use of DNA as a future building block of nanostructures: (1) Nucleic acids of pre-designed lengths, base-orderings, and shapes can be synthesized, and complex structures generated by self-assembly methods. (2) Nature provides us with an arsenal of biocatalysts that can manipulate DNA. These enzymes may be considered as tools for shaping the desired DNA and eventually for the generation of nanocircuitry. For example, ligase ligates nucleic acid, endonucleases affect the specific scission of nucleic acids, telomerase elongates single-stranded nucleic acids by telomer units, and polymerase replicates DNA. These biocatalysts represent "cut" and "paste" tools for the formation



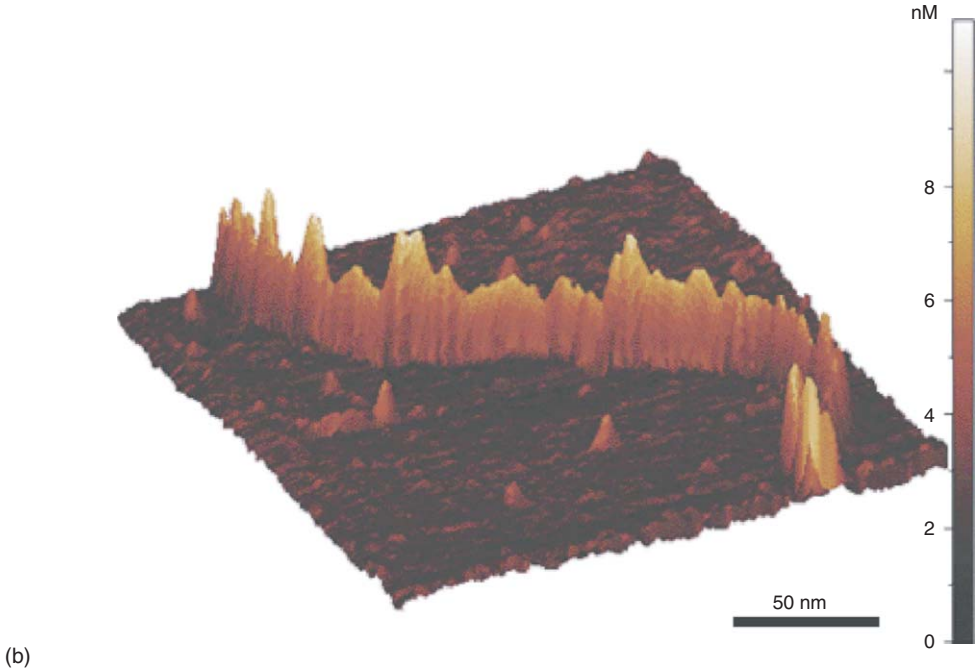


Figure 15 (a) Assembly of an Au-nanoparticle wire in the polyA/polyT template. (b) AFM image of an Au-nanoparticle wire in the polyA/polyT template. (Adapted from Patolsky F, Weizmann Y, Lioubashevski O, and Willner I (2002) Au-nanoparticle nanowires based on DNA and polylysine templates. *Angewandte Chemie International Edition* 41: 2323–2327, with permission.)

of DNA templates and by the application of the replication biocatalyst, the design of future "factories" of nanowires may be envisaged. (3) The intercalation of molecular components into DNA and the binding of cationic species, for example, metal ions to the phosphate units of nucleic acids allows the assembly of chemically active functional complexes. (4) Different proteins bind specifically to certain nucleic acid sequences. This allows the addressable assembly of complex DNA–protein structures. Such protein–DNA complexes may either act as addressable domains other than the DNA for the selective deposition of metals, or alternatively, may act as temporary shielding domains that protect the DNA

from metal deposition. Such insulated domains may then be used for the deposition of other metals or semiconductors, thus enabling the fabrication of patterned complex structures.

With the vision that DNA may act as a template for the generation of nanocircuitry, attempts were made to explore the possibility of organizing DNA-crosslinked semiconductor nanoparticles and DNA-based metal nanoparticle nanowires on surfaces. Since nanoparticles are loaded on the DNA template with gaps between them, the conductivity of metallic nanoparticle aggregates on a DNA template can be enhanced upon the chemical deposition of another metal (e.g., Ag deposition on Au aggregates) filling the gaps and forming a continuous conductive nanowire. The binding of the primary metallic clusters to the template DNA for the subsequent catalytic deposition of wires on the DNA frame may be accomplished by several means: (1) The reduction of metal ions linked to the phosphate groups to metallic seeds linked to the DNA. (2) The use of metal or semiconductor nanoparticles functionalized with intercalator units. Intercalation of the molecular components into double-stranded DNA leads to the association of the nanoparticle to the DNA template. (3) The synthesis of DNA with functional tethers that enable the covalent attachment of the metal or semiconductor nanoparticles to the DNA. (4) The synthesis of single-stranded DNA that includes constant repeat units (e.g., telomers) and the hybridization of metal or semiconductor nanoparticles functionalized with short nucleic acids, that are complementary to the single-stranded DNA repeat units.

Figure 15a exemplifies the method for assembling the Au-nanoparticles on the DNA template using Au-nanoparticles functionalized with an intercalator. The amino-functionalized psoralen (**17**) is reacted with the Au_{55}-nanocluster (diameter 1.3 nm) that includes a single *N*-hydroxysuccinimide active ester functionality (**18**), to yield the psoralen-functionalized Au_{55}-nanoparticle (**19**). As psoralen acts as a specific intercalator for A–T base pairs, the functionalized Au_{55}-nanoparticles were reacted with the pA/pT-ds-DNA (**20**). Subsequently, the assembly was irradiated with UV light to induce the $2\pi + 2\pi$ cycloaddition reaction between the psoralen units and the thymine base sites of DNA. This latter process fixes covalently the Au_{55}-nanoparticles to the DNA matrix. **Figure 15b** depicts the AFM image of the resulting nanoparticle wire. A $\sim$600–700 nm long nanoparticle wire is formed. Its width corresponds to $\sim$3.5–8 nm and it is controlled by the width of the DNA template. The height of the wire is $\sim$3–4 nm, consistent with the fact that the Au-nanoparticles intercalate into the DNA on opposite sides of the double-stranded DNA template. The continuous appearance of the Au-nanoparticle wire is due to the dimensions of the scanning AFM-tip, and in reality, most of the particles are not in intimate contact one with another. The possibility to arrange the Au-nanoparticles on the DNA template allows further catalytic enlargement of the particles by an electroless deposition process (e.g., reaction of $AuCl_4^-$ with NH_2OH) to yield continuous conductive nanowires.

Practical applications of the nanowires require their electrical contacting with macro- or microelectrodes. Toward this goal, a single nanowire produced on a DNA template bridging two microsize electrodes was constructed. Two microelectrodes facing each other (12–16 μm separation) were functionalized with 12-base oligonucleotides that were then bridged with a 16 μm long λ-DNA (**Figure 16**). The resulting phosphate units of the DNA bridge were loaded with Ag^+ ions by ion exchange, which were then reduced to Ag metal by hydroquinone. The small Ag aggregates produced along the DNA backbone were then used as catalysts for further reductive deposition of silver, eventually leading to the formation of an Ag nanowire. This micrometer-sized element had a typical width of 100 nm and a granular morphology (as determined by AFM).

The use of biomaterials as templates for the generation of nanostructures and nanocircuitry in the presence of nanoparticles is in an early phase of development. The viability of the concept has been proven, and nanowires of controlled shapes and electronic functions have been generated by the incorporation of nanoparticles on template biomaterials. The biggest challenges are ahead however, and exciting systems are envisaged for the future. Besides the interesting nano-architectures that may be generated, fundamental problems, such as probing charge transport phenomena in such nanoparticle-biomaterial

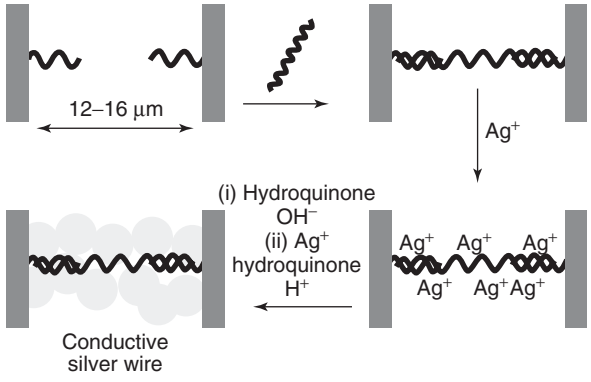


Figure 16 The construction of a wire bridging two microelectrodes by the deposition of nanoparticles on a bridging DNA strand, followed by silver deposition.

systems, may be highlighted. Nonetheless, the organization of nanoparticle architectures on biomaterial templates involves relatively simple motifs that do not make full use of the biomaterial's catalytic functions. The replication of nanoparticle-labeled DNA, and specifically the tailoring of polymerase chain reaction (PCR) for the formation of nanoparticle wires, represents a novel approach to design biological machines for the synthesis of nanoparticle wires and circuitry.

Conclusions

Three facets of bioelectronics – biomaterial-based electronic sensors, biofuel cells, and biomaterial-based electronic circuitry – continue to progress rapidly in research, and some of its applications have been developed commercially. Other topics, such as hybrid systems of neural network and electronic elements, biomaterial-based computers, and biomaterial-based micromachinery devices represent other opportunities in bioelectronics that show promise for future applications. Despite this outlook, however, challenges in bioelectronics remain, including the development of implantable fuel cells, biomaterial-based micromachines (e.g., prosthetic units), and the fabrication of functional electronic circuitry. Recent advances in nanotechnology, the availability of quantum-size nanoparticles and nanotubes of unique electronic and catalytic properties, and nanoscopic tools for manipulating surfaces could open nanobioelectronics as a new research field. The electronic detection of biorecognition events of single molecules, the optical or electronic read-out of biological processes by functional nanoparticles, the high-throughput analysis of numerous genes or protein functions on dense sensing arrays, the construction of biomaterial-based electronic devices of nanodimensions, and the tailoring of ultrasmall self-powered devices for the controlled release of therapeutic drugs may be possible as well.

See also: Biomedical Materials; Biomembranes; Biomolecules, Scanning Probe Microscopy of; DNA and RNA, Biophysical Aspects; Electric and Magnetic Fields in Cells and Tissues; Electromagnetic Biological Effects; Electrophysiology; Elementary Excitations in Biological Systems; Fluorescent Biomolecules; Fluorescent Proteins; Folding and Sequence Analysis; Genetic Algorithms for Biological Systems; Ionizing Radiation Effects in Biological Materials; Ion Transport in Biological Membranes; Membrane and Protein Aggregates; Metalloproteins, Electron Transfer in; Metalloproteins, Structural Determination of; Neuroscience; Photosynthesis, Physics of; Protein Folding and Aggregation; Protein Folding, Engineering of; Protein Folding, Evolution and; Red Blood Cells, Physical Properties of; Rhodopsin and the First Step in Vision; Single-Molecule Methods in Biophysics; Solid-State NMR Structural Studies of Proteins.

PACS: 81.16. Dn; 81.16. Fg; 82.39.Jn; 82.39.Pj; 82.39.Rt; 82.45.Fk; 82.45.Jn; 82.45.Mp; 82.45.Tv; 82.47.Rs; 82.65.+r; 82.80.Fk; 87.14.Ee; 87.14.Gg; 87.16.Xa; 87.68.+z; 87.80.Tq; 87.90.+y

Further Reading

Bone SS and Zaba B (1992) *Bioelectronics*. New York: Wiley.

Braun E, Eichen Y, Sivan U, and Ben-Yoseph G (1998) *Nature* 391: 775–778.

Carrara S, Facci P, Sivozhelezov V, Adami M, Erokhin V, *et al.* (1996) *Molecular Bioelectronics*. New York: World Scientific.

Chakraborty T, Peeters F, and Sivan U (eds.) (2002) *Nano-Physics and Bio-Electronics: A New Odyssey*. London: Elsevier.

Christof M, Niemeyer CM, and Mirkin CA (eds.) (2004) *Nanobiotechnology: Concepts, Applications and Perspectives*. Weinheim: Wiley-VCH.

Gizeli E and Lowe CR (eds.) (2002) *Biomolecular Sensors*. Cambridge: Taylor and Francis.

Hoffmann K-H (ed.) (2002) *Coupling of Biological and Electronic Systems*. Berlin: Springer.

Katz E, Shipway AN, and Willner I (2002) Mediated electron-transfer between redox-enzymes and electrode supports. In: Wilson GS, Bard AJ, and Stratmann M (eds.) *Encyclopedia of Electrochemistry*, vol. 9: *Bioelectrochemistry*, chapter 17, pp. 559–626. Weinheim: Wiley-VCH.

Katz E, Shipway AN, and Willner I (2003) Biofuel cells: functional design and operation. In: Vielstich W, Gasteiger H, and Lamm A (eds.) *Handbook of Fuel Cells – Fundamentals, Technology, Applications*, vol. 1, part 4, ch. 21, pp. 355–381. Weinheim: Wiley-VCH.

Nicolini CA (ed.) (1998) *Biophysics of Electron Transfer and Molecular Bioelectronics*. New York: Plenum Press.

Patolsky F, Lichtenstein A, Kotler M, and Willner I (2001) *Angewandte Chemie International Edition* 40: 2261–2265.

Patolsky F, Lichtenstein A, and Willner I (2001) *Nature Biotechnology* 19: 253–257.

Willner I and Katz E (eds.) (2005) *Bioelectronics: From Theory to Applications*. Weinheim: Wiley-VCH.

Willner I, Blonder R, Katz E, Stocker A, and Bückmann AF (1996) *Journal of the American Chemical Society* 118: 5310–5311.

Willner I, Katz E, and Willner B (2000) Layered functionalized electrodes for electrochemical biosensor applications. In: Yang VC and Ngo TT (eds.) *Biosensors and Their Applications*, chapter 4, pp. 47–98. New York: Kluwer.

Willner I, Katz E, and Willner B (2002) Amplified and specific electronic transduction of DNA sensing processes in monolayer and thin-film assemblies. In: Brajter-Toth A and Chambers JQ (eds.) *Electroanalytical Methods of Biological Materials*, pp. 43–107. New York: Dekker.

Willner I, Willner B, and Katz E (2003) Bioelectronics: development of biosensors, biofuel-cells and circuitry. In: Barsanti L, Evangelista V, Gualtieri P, Passarelli V, and Vestri S (eds.) *Molecular Electronics: Bio-sensors and Bio-computers*, NATO Science Series, II. Mathematics, Physics and Chemistry – vol. 96, pp. 311–339. Dordrecht: Kluwer.

Biological Structures

R M Ottenbrite and R Javan, Virginia Commonwealth University, Richmond, VA, USA

Polymers

Polymers are materials composed of molecules linked to one another in a linear or branched manner as multiple repeat units. Applications of biomedical polymer structures include: sutures, maxillofacial implants, dental constructs, joint replacements, cardiovascular devices, artificially engineered tissues, and drug-delivery and gene-transvection systems.

In order to understand biomedical polymer structures, one should have some knowledge of the physical properties that are essential for their function in a biological system. To accomplish specific biorelated tasks, the materials used must meet many physical and biological requirements. These include biocompatible and nontoxic, degradable or nondegradable, water soluble or insoluble, hard or soft, rigid or flexible, hydrophobic or hydrophilic, and bioactive or bioinactive materials. To help with the reading of this article some concepts such as biocompatibility, characterization, biodegradation, and applications such as biological scaffolds, hydrogels, polymeric drugs, and drug-delivery systems are addressed.

Biocompatibility

Biocompatibility refers to the ability of a material to perform with an appropriate host response in a specific application and incorporates blood and tissue compatibility. When a material is characterized for a bioapplication, mechanical properties, such as stress–strain relationships, surface properties, such as wettability, in addition to chemical and biological properties, are thoroughly evaluated.

Biodegradation

Biodegradation is the chemical breakdown of materials by the action of living organisms leading to changes in physical properties. There are other definitions of polymer degradation; however, they all involve a chemical breakdown in an aqueous environment caused by hydrolysis (reactions with water) and/or enzymes. Biodegradable polymers were first used surgically as sutures, staples, implants, and temporary devices. With current developments being made in drug-delivery systems, tissue engineering, and scaffolds for regenerative organs, there is a growing need for new biodegradable materials. The majority of the products presently in use are composed of poly(lactic acid), poly(glycolic acid), poly (ε-caprolactone), their copolymers and composites. In addition to these polymers, several other important degradable polymer systems are available and many more are currently being developed (**Figure 1**).

Biological Scaffold

Biological scaffold is an artificial support system to replace, temporarily or permanently, a biological structure. Sutures, bone-fixation devices (bone nails, screws, or plates), and vascular grafts are early examples of such support devices. The degradable scaffold implant provides temporary mechanical support until the natural tissue has developed and regained its strength. As the tissue assumes integrity, the scaffold material should concomitantly slowly degrade.

Adjusting the rate of degradation to the regeneration process of the surrounding tissue presents one of the major challenges in the design of a temporary scaffold. For temporary scaffolding to function effectively, a gradual stress transfer should occur. As the natural tissue develops, the degradable implant should gradually weaken so that regenerated tissue can resume its function. For example, when a bone is shattered, usually a metal pin is inserted or a metal

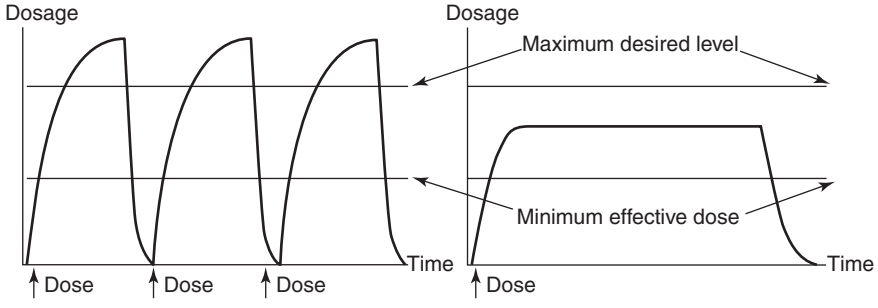


Figure 1 Ordinary drug delivery (left) and delivery using degradable polymer systems (right).

plate fixed to the bone in order to stabilize the injury. These supply support and alignment during the healing process. However, the pin or the plate must be surgically removed after ossification is complete. While these supports are in place, the bone to be mended has been stress free. After the pin or plate is removed, the regenerated bone is too weak to support normal activity until the bone regenerates sufficiently to assume the normal stress loads. The major advantage of biodegradable supports is that as the osteocytes proliferate, the polymer degrades and is absorbed while stress is gradually transferred from the scaffold to the bone. Consequently, the bone adjusts to the increasing stress as it mends. This significantly reduces rehabilitation time and eliminates a second surgical process to remove the support.

Tissue Engineering

Tissue engineering is a new frontier in bioscaffolding and is defined as "the application of engineering principles to create devices for the restoration, modification, and assembly of functional tissues from native or synthetic sources." The goal is to regenerate tissue, bone, or organs that have been disabled or weakened by disease or injury. The tissue scaffold has to be nontoxic and biodegradable as well as possess the right three-dimensional structure to support cell differentiation. The mechanical properties of the scaffold and the surrounding environment have to be in harmony during the process. To enhance cell proliferation, slow release of supportive bioactive recipients, such as specific enzymes, can be included in the scaffold which are then released as the scaffold degrades.

Polymeric Drugs and Drug-Delivery Systems

These systems constitute an important technology that has been developing over the last 30 years. Currently, polymeric drug delivery is a 35 billion dollar market and will continue to grow 5% per year for the next ten years. Many drugs are rapidly cleared from the circulatory system through the kidneys or taken up by the liver or spleen and, therefore, exhibit very little efficacy. Other difficulties that are associated with drug assimilation are due to low solubility, random distribution, and toxicity. Polymers are very versatile and, therefore, offer novel approaches so that systems can be tailored for an intended application and long-term delivery times (days, weeks, months, and even years). Polymeric drug-delivery systems are prepared in a variety of sizes and shapes, such as films, fibers, disks, slabs, pellets, rods, and monolithic forms. The development of new polymer methodology is critically important as new drug therapies, such as peptide, protein, and DNA type drugs, are being developed (**Figure 1**).

Polymeric drugs are drugs that are physiologically active polymers. In nature, these are proteins (enzymes) and polysaccharides (heparin). However, other than synthetic mimetics, there are few man-made polymers that are biologically active. Among these are the polyanionic and polycationic polyelectrolytes. Polyanions stimulate the immune system and have shown anticancer and antiviral activity (HIV and hoof and mouth disease). Polycations have antimicrobial activity and are currently being investigated as gene-delivery adjuvants.

Active agents can be attached to a polymer backbone and are known as drug-conjugates. Polymer–drug conjugates enhance circulation time (due to increased size) as well as increase solubility which enables transport in the circulatory system. However, conjugation usually inactivates the drug. Therefore, polymeric prodrugs were developed. Prodrugs are conjugate systems that have a drug attached to a polymer by a degradable linkage so that the drug can be released by hydrolysis or enzyme activity. Furthermore, polymer conjugate systems can be modified with solubilizing and targeting groups. This system has been successfully used for delivering anticancer drugs, such as daunarubacin.

Implantable drug-delivery systems involve incorporating drugs into a biodegradable polymer matrix. Using this technology, desirable dosages can be released over a prolonged period of time at an intermittent or constant rate. This achieves effective therapies and prevents under- or overdosing the patient. Biodegradable polymers that are most commonly used are poly(glycolic acids), poly(lactic acids), polyurethanes, and poly(ortho esters) with commercial names such as Decapeptyl®, Lupron Depot®, and Sandostatin LAR®. Degradation may occur through bulk hydrolysis in which the polymer degrades in a fairly uniform manner throughout the matrix. For some degradable polymers, mostly polyanhydrides and poly(ortho esters), degradation and erosion occurs only at the surface of the polymer which provides a release rate that is proportional to the surface area of the drug-delivery system.

With advancements in pharmaceutical sciences, it has been recognized that constant release is not necessarily the optimum method for delivering drugs. As a result, externally modulated or self-regulating drug-delivery systems, using biomedical polymer structures, are being engineered to respond to environmental stimuli, pH, temperature, or pressure. The glucose-sensitive insulin-release system is an example of this application; in this device, high glucose concentrations induce matrix disintegration

with polymer dissociation from the surface, leading to the release of insulin.

Hydrogels

Hydrogels are water-insoluble, three-dimensional networks formed by cross-linking a hydrophilic polymer. These networks absorb large quantities of water to form a gel. Natural hydrogels, such as acacia, agar, carrageen, and alginates, have been known for centuries and are used in food and personal products.

Research to develop synthetic hydrogels started in the 1970s and dramatically increased in the 1990s due to the recognition of the many potential applications that these constructs could have in medicine, agriculture, and industry. This work has led to the synthesis of many new hydrogels and a better understanding of the physicochemical properties involved. When a hydrogel is in the swollen state, the three-dimensional compartment holding the solvent is called a cell or pore. As the gel dries, it becomes concomitantly smaller in size and finally collapses to the solid state called a sol (thus the term sol–gel). The overall chemical structure remains the same during these processes. The three-dimensional cell structure is achieved through cross-linking linear polymer chains either chemically (by covalent bonding) or physically (by hydrogen bonding, ionic, or hydrophobic interactions). Since noncovalent bonds are reversible, physical gels have sol–gel reversibility. The extent of swellability depends on the hydrophilic nature of the polymer and the cross-linking density. Smaller cells (higher cross-link density) provide more stable hydrogels and appear to be more solid-like. Methods for preparing hydrogels include irradiative cross-linking using electron beams, gamma rays, X-rays, or UV light, and chemical reactions using di- or multifunctional cross-linking agents (**Figure 2**).

Most hydrogels have good mechanical stability, good refractive index, and oxygen permeability. Consequently, one of the early biomedical applications of hydrogels was for soft contact lenses. The most widely used hydrogel for bioapplications is cross-linked poly(hydroxyethyl methacrylate) (PHEMA). This structure provides a water content similar to that of living tissue and was one of the first materials used for soft contact lenses. This hydrogel is inert to biodegradation, permeable to metabolites, not absorbed by the body, endures heat sterilization, and can be fabricated into different shapes and forms. However, problems with protein deposits are still an issue. Hydrogels are now used for artificial tendons and muscles, wound-healing bioadhesives, artificial kidney membranes, artificial skin, and maxillofacial reconstruction materials.

Hydrogels have also been made with poly(*N*-vinyl-2-pyrrolidone), poly(methacrylic acid), and poly (methyl methacrylate-co-maleic anhydride), which are used in similar biomedical applications. For example, polyacrylamides are synthetic long-chain polymers designed to attract positively or negatively charged materials and are widely used in gel electrophoresis. In gel extraction processes, an acid is added to shrink the hydrogel, and conversely, a base is added to swell it (**Figure 3**).

Biodegradable hydrogels have emerged as an important class of biomaterials for controlled drug delivery. The release rate of an excipient is determined by drug and polymer properties, such as the swellability, degradability, porosity, permeability, and hydrophilicity of the hydrogels.

Smart (or intelligent) polymer-structured hydrogels are being developed for several biomedical applications. A unique aspect about these hydrogels is that their swelling ratio (swollen volume divided by dried volume) can be engineered to change abruptly with a slight external aberration. Hydrogel materials can undergo continuous or discontinuous changes in swelling in response to environmental stimuli, such as changes in pH, temperature, ionic strength, pressure, and light. These stimuli–response properties are advantageous in many potential bioapplications. For example, timed control release of insulin has been achieved with "smart" pH-sensitive hydrogels. This phenomenon is known as on-demand release of insulin in response to increased levels of glucose. Thermosensitive hydrogels, such as poly(*N*-isopropylacrylamide) and poly(*N*-vinyl caprolactam), have also been engineered for controlled drug delivery, where hydrophobic interactions respond to small changes in temperature

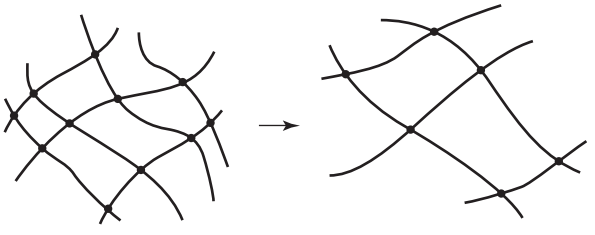


Figure 2 Network of cross-linking polymer chains in hydrogels (left) that swell after water absorption (right).

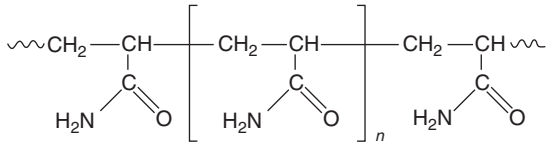


Figure 3 Polyacrylamide.

that can vary drug release rates. "Smart" hydrogels are being used extensively for robotics where responses to numerous and different stimuli are important features. Reversible contraction and expansion are essential in the development of advanced robotics with electrically driven muscle-like actuators. The transformation of electrochemical stimuli into mechanical work is being used for such artificial muscle responses.

Hybrid hydrogel systems are composed of at least two distinct classes of macromolecules. These are of particular interest because it is possible to combine and superimpose the component macromolecule properties into the hydrogels. Theoretically, it is possible to manipulate the DNA sequence and generate unlimited varieties of protein products with optimum structures and properties. For example, hybrid hydrogels have been prepared, using synthetic polymers and biological macromolecules, as engineered proteins and DNA-based biomaterials.

Microporous and macroporous hydrogels are being made with pore sizes in the 10–100 nm and 100 nm–10 μm range, respectively. These hydrogels are designed to swell within minutes regardless of the size of the matrix. In these cases, the pore size is an important property of a hydrogel. Now superporous hydrogels, a new type of hydrogel, with supersize pores are being developed. These pore sizes are larger than 100 μm and can reach millimeter range. Their fast swelling kinetics, in addition to super-absorbent properties, allow these superporous hydrogels to be used in special biomedical applications, such as long-term oral drug delivery, due to their ability to achieve high gastric retention times. Superporous hydrogels and their composites are also being engineered for artificial pancreas, cornea, skin, and joint cartilage. They are being used as cell growth substrates in tissue engineering scaffolding, burn dressings, soft tissue substitutes, and as surgical pads to control bleeding.

Hyaluronic Acid

Hyaluronic acid (HA, Hyaluronan®) is a naturally occurring, biocompatible, and biodegradable linear polysaccharide. It is composed of unbranched repeating disaccharide units of glucuronic acid and *N*-acetyl glucosamine linked by *β*-(1-3) and *β*-(1-4) glycosidic bonds. HA is present in all soft tissues of higher organisms but is in very high concentration in the synovial fluid and vitreous humor of the eye. It can be extracted from various sources of animal tissues, such as the umbilical cord, skin, and rooster combs.

Hyaluronic acid is used in many medical applications, such as promoting cell mobility and differentiation in wound healing, viscosupplementation and viscosurgery, ophthalmic surgery, and in cosmetic applications. Despite its highly attractive rheological properties, unmodified HA has a short residence time and poor mechanical properties. However, HA can be chemically modified to prolong degradation time and improve mechanical stability *in vivo*. Recently, it has been formed as a thin film that can be used for tissue separation. HA-based films or sponges have been made for implants loaded with therapeutic agents for delayed release and prolonged activity.

Chitosan

Chitosan, poly(*β*-(1-4)-D-glucosamine), is a water-soluble product obtained by *N*-deacetylation of chitin (**Figure 4**). Chitin is a naturally occurring polysaccharide and second to polysaccharides as the most abundant natural polymer. It is mainly isolated from crustaceans, such as crab and shrimp, but is also found in fungi and the hard shells of insects. The physical properties of chitosan are related to the number and the distribution of *N*-acetyl groups. The primary amino group at the 2-position forms quaternary groups in water that become soluble at $pH < 8.5$.

Chitosan has superior biocompatible and biodegradable properties and is used in biomedical applications such as wound dressings. It was first used for burn bandages because it maintains good hydration, allows toxins and metabolites to be transported away from the injury, and promotes healing. In wound healing, chitosan facilitates the formation of granulation tissue with angiogenesis. It induces fibroblasts to release interleukin-8, which is involved in the migration and proliferation of fibroblasts and vascular endothelial cells. However, chitosan's surface-induced clot-forming potential has limited its application in blood contact situations. Although chitosan itself has a hemostatic function, its derivatives, such as, *N*-hexanoyl and *N*-octanoyl chitosan, have antithrombogenic activity. The antimicrobial and wound-healing properties along with an excellent film capability make chitosan suitable for ocular lens bandages.

Although water soluble, chitosan can be formed in different shapes such as nanoparticles, microspheres,

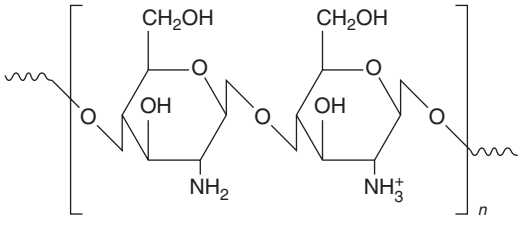


Figure 4 Chitosan [poly(*β*-(1-4)-D-glucosamine)].

membranes, sponges, rods, beads, and solutions. Consequently, the porous spongy chitosan matrices are excellent biomaterials for tissue engineering, especially artificial skin. When replacing skin, it is important that the substitute material completely adheres, wets, and conforms to the wound surface so as to prevent small air pockets where bacteria can proliferate. In addition, antibiotics and growth factors can be incorporated into the basic matrix.

The chitosan microspheres were found to be a very favorable form for parenteral controlled drug-delivery systems requiring low-acting drug delivery to the systemic circulation, as well as for active or passive targeting to the treatment sites. Several types of active reagents, such as DNA, Cyclosporine A, and insulin, have been encapsulated in chitosan nanoparticles. Chitosan membranes have good mechanical properties and are being evaluated for transdermal drug delivery.

With new biotechnology techniques, peptide analogs that are resistant to enzyme degradation have been prepared for oral drug delivery. However, their hydrophilicity and molecular size present problems with respect to absorption through tight junctions of the intestinal epithelium. Due to chitosan's mucoadhesive characteristics, it has the ability to partially open tight junctions and enhance the penetration of macromolecules across membranes, such as the intestinal barrier. Consequently, chitosan and its derivatives have been found to be very effective for oral peptide delivery systems.

Currently, chitosans are being investigated as nonviral gene-delivery vectors. However, when loaded with DNA at physiological and alkaline pH, they are insoluble. This originally presented a problem, but trimethyl quaternization has increased the soluble range of this system.

Polyhydroxyalkanoates

Polyhydroxyalkanoates (PHA)s are natural polyesters that are produced and stored by a wide variety of microorganisms. *Pseudomonades* are the most commonly known microorganism that produces and uses PHA as an intracellular energy reserve. When *Pseudomonades* are stressed under conditions of excess carbon substrates and limited essential nutrients, PHA is formed and stored as nodules that can be up to 90% of the organism's dry weight. Polyhydroxybutyrates (PHB) and polyhydroxyvalerates (PHV) comprise the majority of the bacterial polyesters produced in nature (**Figure 5**). These natural polyesters provide an important source of highly pure 3-D-hydroxybutanoic acid and 3-D-hydroxypentanoic acid, respectively. PHB has been developed as biocompatible

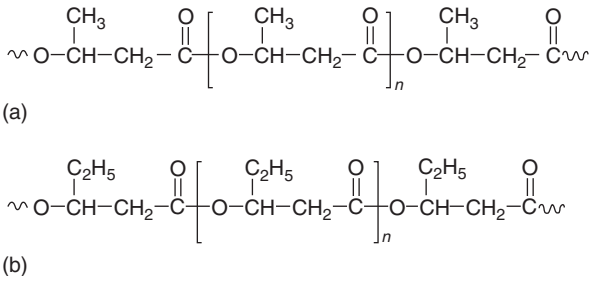


Figure 5 (a) Poly(hydroxybutyrate) and (b) poly(hydroxyvalerate).

and biodegradable materials based on properties such as low toxicity and degradation to 3-D-hydroxybutyric acid, which is a normal constituent in human blood. In addition, industries now have the technology to melt-process PHB materials into a variety of shapes and forms. These thermoplastic biopolymers can be produced as rigid brittle plastics, flexible materials, or strong elastomeric fibers.

Pure PHB or PHV are highly crystalline and brittle; however, their copolymers and composites provide materials that have more desirable properties such as enhanced rates of degradation and more flexibility. Consequently, PHAs have been developed for medical applications, such as sutures, artificial skin, drug delivery, and paramedical disposables. Currently, polyhydroxybutyrates are being investigated for bone-implants and bone-related treatments. It has been found that transected nerves heal and restore their function after being wrapped in thin PHB films.

These unique natural polyesters have not yet reached their full biomedical potential due to some biocompatibility issues. For example, on contact with blood, PHA induces platelet adhesion and subsequent thrombi formation. However, new composites are being developed, by combining PHB with poly(ethylene glycol) to form "natural–synthetic hybrid" block copolymers, that exhibit enhanced blood-compatibility as well as inherent bone-compatibility.

Aliphatic Polyesters

Aliphatic polyesters are derived from naturally occurring compounds, such as lactide (LA), glycolide (GL), and ε-caprolactone (CL). These represent an important class of biocompatible and biodegradable materials. Initially, the main application of these materials was for reabsorbable sutures. L-lactic and L-glycolic acids, commonly found in animal systems and in low concentrations, are nontoxic and readily absorbed by the host. However, pure poly(L-lactic acid) and pure poly(L-glycolic acid) are highly crystalline; consequently, they do not have good material properties, such as flexibility or degradability

(**Figure 6**). On the other hand, PLLA and PLGA copolymers and composites (mixtures) are amorphous and provide desirable properties such as tenacity, flexibility, and facile degradability. In addition, PLA/PGA copolymers can be "tailor-made" to meet the requirements of specific applications. Currently, these copolymers are being widely utilized as temporary scaffolds for the regeneration of tissues and as carriers to deliver bioactive molecules.

Poly(ε-caprolactone) (PCL) has a low melting point, high solubility, and an exceptional ability to form blends. Biodegradable PCL staples for wound closure are in clinical use in Europe. Studies have shown that ε-caprolactone and its copolymers are tissue-compatible and nontoxic. Semicrystalline PCL has good drug permeability and a very slow degradation rate. The requirements for optimal controlled drug release or mechanical properties for engineered tissue can be achieved through copolymering PCL with other monomers. When PCL is combined with amorphous PLA/PGA copolymers that have poor drug permeability, a much higher degradation rate is obtained.

These aliphatic polyester polymers usually degrade by hydrolysis, and to a lesser extent, by enzymatic processes in animal systems. Consequently, these materials have rapidly gained recognition in the area of tissue engineering. Here, cells and extracellular matrix components can be implanted into the porous scaffolds to make pseudo-anatomical shapes. Although approved by FDA, specific applications, biocompatibility, sterilization, and storage issues need further improvement.

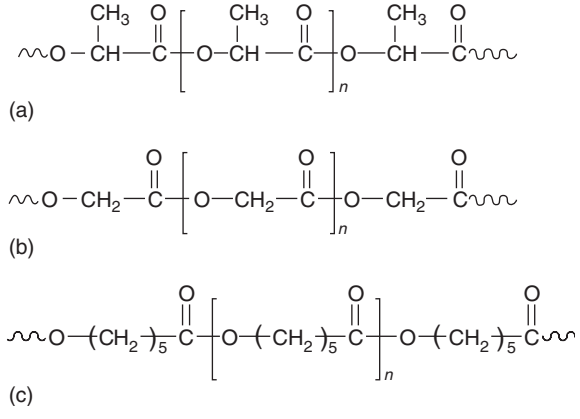


Figure 6 (a) Poly(lactic acid), (b) poly(glycolic acid), and (c) poly(caprolactone).

Current developments involve applications for bone, cartilage, and vessel engineering and for mechanically dynamic sites. Recently, unique biodegradable copolymer nanofibrous scaffolds were produced by electrospinning. The diameter of the fibers was approximately 500 nm with an aligned topography that mimics the orientation of cells and fibrils found in a native artery. This process provides an almost ideal tissue-engineering scaffold, especially for blood vessel engineering due to the nanometer-scale dimension that physically resembles the natural material but has the advantage of biodegradability.

Poly(ortho-esters)

Poly(ortho-esters) are unique polymeric structures that undergo degradation more rapidly than the aliphatic polyesters. In an aqueous environment, they undergo surface erosion and are able to release drugs at a constant rate. However, the degradation is not uniform and the acid that is released catalyzes a localized breakdown and causes the formation of local erosion forming large void areas. By varying the acidity and the amounts of neutralizing excipients incorporated into the polymer, the erosion rate can be controlled. Delivery systems are being produced for short-term applications such as oral 12–24 h applications, intermediate ophthalmic products lasting 1–7 days, and subdermal implants that last as long as one year or more.

Polyanhydride

Polyanhydride polymers are synthesized by dehydrating dicarboxylic acid functional monomers to form corresponding polyanhydride linkages (**Figure 7**). The anhydride groups in the polymer chain are very moisture sensitive and on contact with water rapidly hydrolyze to form carboxylic acid groups. Because of the hydrolytic instability of this linkage, these materials were developed as sutures, drug carriers, bioabsorbable prostheses, and vascular grafts. In general, they show some degree of crystallinity and most are soluble in common organic solvents. The decomposition rate decreases as the aromatic content and hydrophobicity of the polymer backbone is increased.

Well-defined and predictable degradation rates depend on the polymer composition and the size and shape of the implant. A unique feature of polyanhydrides is their ability to undergo uniform surface

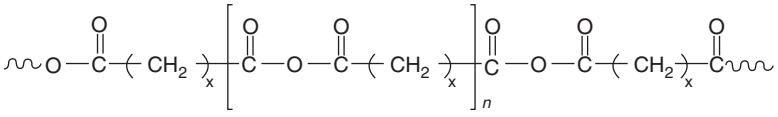


Figure 7 Polyanhydride.

erosion without additional additives to inhibit pitting and cavity formation. This increases performance predictability and decreases the presence of extraneous substances; both properties are highly desirable in controlled-release drug delivery, which is the most current application of this polymer system. Current investigations are under way to use polyanhydrides for controlled delivery of osteogenic proteins and insulin. Polyanhydrides and their degradation products have been shown to be neither mutagenic nor cytotoxic and were FDA approved in 1996 for local delivery of chemotherapeutics to brain tumors.

Polyphosphazenes

Polyphosphazenes are unique polymers with an inorganic backbone structure consisting of nitrogen–phosphorous bonds (**Figure 8**). By changing the side groups on the phosphorous, researchers can easily modify the bulk and surface properties of polyphosphazenes. The side groups can be amino, alkyl, aryl, alkoxy, aryloxy, inorganic, or organometallic units.

If both side groups on the repeat unit are identical, the polyphosphazene is homogeneous; if they are different, then the polyphosphazene is heterogeneous. Due to the facile modification of the polyphosphazene side chains, interesting new materials have been generated in order to develop inert biocompatible materials for use in cardiovascular and other biomedical devices as well as biological membranes and coatings. Currently, these polymers are being investigated as gene and drug-delivery systems. Several bioactive agents, such as, steroids, local anesthetics, catecholamines, and heparin, have been attached to the polyphosphazene side groups are being evaluated as a delivery system.

Nondegradable Polymers

Nondegradable polymers are materials that are not degraded in a biological environment. These materials are used in devices that are implanted for permanent use, such as joint prostheses, maxillofacial implants, and heart valves. Many of these materials are made by polymerization of vinyl groups ($-CH=CH_2$). These polymer structures include methyl methacrylate, polypropylene, and polyethylene and cyanoacrylates. Some are polyesters such as Dacron®. Other polymer materials include polyurethanes, silicones, and poly(ethylene oxide).

Polycyanoacrylates

Polycyanoacrylates are better known as "super glues" as they adhere to many different materials, such as metal, glass, wood, and tissue (**Figure 9**). Strong bonds are formed at room temperature in an aqueous environment without the need of adding a catalyst. The monomers polymerize very rapidly in the presence of weak bases such as water (moisture) or amino compounds. The higher alkyl cyanoacrylates, such as *n*-butyl cyanoacrylate, are more hydrophobic and, therefore, spread more rapidly on surfaces, polymerize more rapidly and degrade slower than methyl cyanoacrylate. In the body, these materials achieve hemostasis rapidly and bond to wet tissues strongly. Consequently, they are useful as bioadhesives.

However, due to adverse tissue response and production of tumors in laboratory animals, cyanoacrylates have not been approved for routine clinical applications. The current applications include surface wound dressing in dental surgery and ophthalmic surgical adhesives where monomers are applied directly to the tissue and almost instantaneously polymerize with strong adherence to the tissue. They are also being investigated as adhesives to stop bleeding of gastric varices – abnormally dilated and lengthened vessels. As the cyanoacrylate is injected into the varix, it forms a polymer plug that occludes the aperture in the vessel, which is eventually expelled into the lumen of the stomach. Cyanoacrylates have also been investigated as potential drug-delivery matrices. However, they have a tendency to induce significant inflammatory response at the implant site.

Polyurethanes

Polyurethanes are unique materials that offer the elasticity of rubber combined with the toughness and durability of metal (**Figure 10**). They are more commonly known as wood finishes that are more durable than varnish. For medical use, they can be manufactured as very strong and resilient products with good flexible and electrometric properties. Applications including artificial heart diaphragms, ventricular

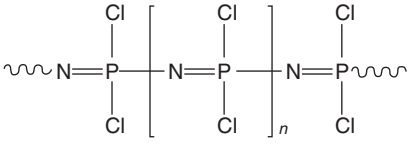


Figure 8 Polyphosphazene.

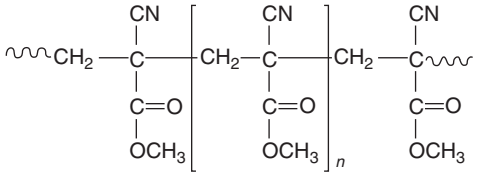


Figure 9 Polycyanoacrylate.

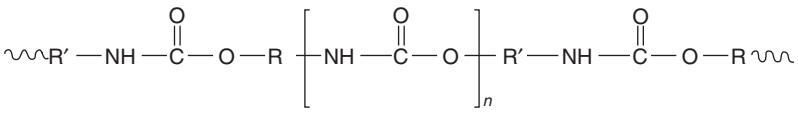


Figure 10 Polyurethane.

assist bladders, vascular grafts, mammary prostheses, and pacemaker leads.

Copolymers of urethane with monomers, such as esters, ethers, and ureas provide a wide range of mechanical properties that are suitable for tissue engineering and regenerative medicine. Polyurethanes have exceptional elasticity and flexibility that are extremely important qualities for soft tissue engineering and scaffolds. Porous, biodegradable, electrometric polyurethane scaffolds combined with the patient's own bone marrow have successfully facilitated bone regeneration. Studies show that polyurethane scaffolds induce crystalline calcium phosphate deposition similar to that formed in bone structure. When used in plastic and orthopedic reconstructive surgery, tissue ingrowth occurs with low acute inflammation. However, after several months these implants begin to show degradation. Poly(ether urethanes), however, are resistant to hydrolysis and are much less biodegradable.

Poly(methyl methacrylate)

Poly(methyl methacrylate) (PMMA) is an amorphous, transparent, and hydrophobic thermoplastic polymer that is very hard and stiff but brittle and notch-sensitive (**Figure 11**). PMMA, commonly known as "safety glass," is an "acrylate" with brand names such as Plexiglas, Diakon, and Lucite. Based on its strength, hardness, and adhesive qualities it is used in dentures, dental crowns, and caps.

It has excellent light transmittance and good abrasion and UV resistance, but poor low temperature, fatigue, and solvent resistance. It is, however, very biocompatible and was the original material used for hard contact lenses and intraocular lenses. Soft contact lenses evolved by substituting a $–CH_2–CH_2OH$ group for the alkyl methyl ester group which was engineered to form a hydrophilic hydrogel lens and other biorelated products.

Methyl methacrylate (PMMA) is used extensively as a medical adhesive. It is used as bone cement to secure prostheses, such as hip replacements and dental crowns. In orthopedic surgery, PMMA cement is injected into the collapsed vertebra to reconstruct back injuries. Although this procedure does not re-expend the collapsed vertebra, it seems to alleviate pain by reinforcing and stabilizing the fracture. PMMA is used extensively in maxillofacial augmentation to

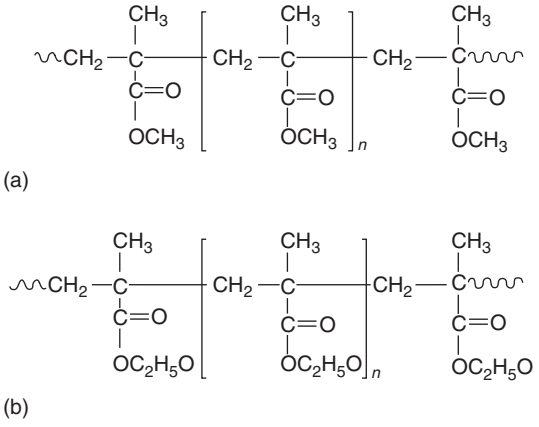


Figure 11 (a) Poly(methyl methacrylate) and (b) poly(hydroxyethyl methacrylate).

improve the skin contours and reduce depressions in the skin due to scars, injury, or lines. When applied in cosmetic surgery, PMMA is used as a suspension of polymer beads in a vehicle, such as bovine collagen (Artecoll®), hyaluronic acid (MetaCrill®), or other biocolloidal suspension.

Silicone

Silicone materials consist of repeating units of inorganic $–SiO_2–$ as the backbone structure, with methyl or other functional groups as –Si– side groups (**Figure 12**). These materials are biocompatible, very stable, nontoxic, and insoluble in body fluids. For medical applications, dimethylsiloxane is polymerized to form a silicone gel. Cross-linking the dimethylsiloxane polymer chains leads to the formation of silicone rubber-like elastomers.

High-viscosity silicone rubbers are used as tissue expanders. The more common applications of this material have been for maxillofacial, breast, chest, and calf augmentation. They are currently being evaluated for joint replacement and tendon reconstruction. Although silicone is considered biologically inert, it can elicit a mild foreign body reaction followed by tissue encapsulation. The use of silicone gels became controversial due to the concerns regarding monomer migration and toxicity as well as (unproven) human adjuvant diseases. These uncertainties lead to a limitation of the use of silicone gel implants by the FDA in 1992. Currently, silicone gel implants are being reevaluated and are

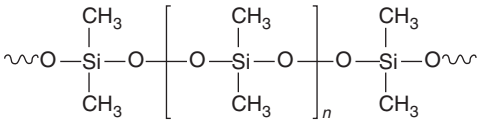


Figure 12 Silicone.

available only under specific guidelines for breast reconstruction.

Poly(ethylene oxide)

Poly(ethylene oxide) (PEO) is also known as poly(ethylene glycol) or (PEG) and is one of the few polymers that is approved by the FDA for clinical use (**Figure 13**). PEG compounds are nontoxic and nonimmunogenic as well as soluble both in water and polar solvents. In the pharmaceutical industry, the use of many drugs has been prevented or limited due to delivery and solubility problems. Getting a drug to target is a difficult task and failure to do so can produce serious toxic effects. Currently, the attachment of PEG to various drugs (also known as "PEGylation") is widely employed to enhance drug solubility and efficacy. PEGylation technology consists of linking PEG to a bioactive component. The resulting bioconjugate, when administered intravenously, is stable in the blood with enhanced drug delivery. The method of attachment may be with a permanent or a biodegradable linkage depending on the specific application. The solubility of the bioconjugate can be tailored since it is dependent upon the lipophilicity of the drug and the length of the PEG. Usually, the molecular weight of the PEG molecules used is between 200 and 20 000, and their shape can be either linear or branched. The *in vivo* half-life of PEG-bioconjugates in plasma is prolonged due to the larger molecular size of the PEG-conjugate, which diminishes the rate of renal excretion and distribution in the tissues.

In addition, PEG-bioconjugates can decrease antigenicity by masking the immune recognition of antigenic sites. This is known as a "stealth" component of the PEG-bioconjugates since they are not recognized by the immune system as a foreign material. Consequently, PEGulated drugs have become very important in the design of new pharmaceuticals due to the fact that they enhance bioavailability by increasing drug solubility, as well as increasing residence time in the body by increasing the size of the conjugate drug and so decreasing kidney elimination.

Pluronics®

Pluronics® represent an important class of biomedical polymers (**Figure 14**). They are unique materials

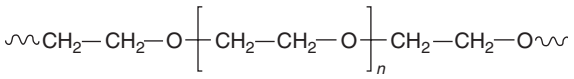


Figure 13 Poly(ethylene glycol).

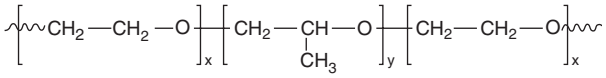


Figure 14 Pluronics $(PEO)_x$–$(PPO)_y$–$(PEO)_x$.

composed of triblock PEO–PPO–PEO copolymers of poly(ethylene oxide) (PEO) and poly(propylene oxide) (PPO). The pluronic PEO block is hydrophilic and water soluble while the PPO block is hydrophobic and water insoluble. In an aqueous environment, these block copolymers self-assemble into micelles with a hydrophobic PPO center core and a hydrophilic PEO outer shell that interfaces with water. Since these micelles are amphiphilic, they are able to accommodate lipophilic molecules in the central hydrophobic core area. Consequently, pluronic micelles are effectively used as drug carriers because their assemblies can act as passive drug containers. These assemblies deliver drugs into subcellular compartments by slowly releasing hydrophilic–hydrophobic encapsulated excipients into physiological fluids.

Pluronic micelles have many advantageous properties that include a relatively low *in vivo* toxicity and an appropriate size that restricts renal excretion. They provide the opportunity to deliver drugs from micelles in a spacial and temporal controlled manner with enhanced intracellular uptake via fluid-phase endocytosis rather than a passive diffusion. When drugs are encapsulated as pluronic micelles, their intracellular uptake by normal cells or drug-sensitive tumor cells is usually substantially reduced. This shielding effect is an advantage as it diminishes drug interaction with healthy tissues. Site-specific release of pluronic-encapsulated drugs can be activated under the influence of ultrasound with an enhanced penetration and retention effect that provides selective accumulation of the encapsulated drugs in solid tumor cells. The latter factor is important for overcoming drug resistance, which is a significant problem in cancer chemotherapy. In addition to sensitizing tumor anticancer agents and maximizing therapeutic outcomes, pluronic block copolymers also effect gene expression at the transcript level. The aggregation state of these micelle systems can be controlled by choosing the appropriate pluronic size and PPO/PEO block-length ratio. For example, pluronic P-105, which is often used in biosystems, has a molecular weight of 6500 and PEO/PPO repeat unit blocks of 37 and 56, respectively.

Summary

In summary, the application of biomedical polymer structures has become one of the most interesting and rapidly growing areas in polymer science with an annual economic growth of 5%. Based on current research results, conventional commodity polymers, such as poly(methylmethacrylate), polystyrene, polyethylene, and polypropylene, are being replaced with novel, high-performance materials. The design and synthesis of these exceptional polymer systems are providing new and more effective ways to enhance biocompatibility along with unique and efficacious bioapplications.

See also: Biomedical Materials; Polymer Structures.

PACS: 87.80.Rb; 81.05.Je; 87.64.Je

Further Reading

Chiellini E, Sunamoto J, Migliaresi C, Ottenbrite R, and Cohn D (eds.) (2001) *Biomedical Polymers and Polymer Therapeutics*. New York: Plenum Publishers.

Denkbas E, Ozturk E, Ozdemir N, Kecei K, and Ergun M (2003) *Journal of Bioactive and Compatible Polymers* 18(3): 177–190.

Ehrenfreund-Klienman T, Domb AJ, and Golenser J (2003) Polysaccharide scaffolds prepared by crosslinking of polysaccharides with chitosan or proteins for cell growth. *Journal of Bioactive and Compatible Polymers* 18(5): 323–337.

Heller J, Roskos KV, and Duncan R (1993) Use of poly(ortho esters) in the controlled release of therapeutic agents. *Makromoleculare Chemie, Macromolecular Symposia* 70/71: 163–171.

Hong Y, Gao C, and Shen J (2003) Influence of quaternized polyurethane membrane surfaces on human endothelial cell attachment and growth. 18(3): 191–206.

Hoste K, Schacht E, and Rihova B (2002) Macromolecular prodrugs of Mitomycin C. 17(2): 123–139.

Leong K, Domb A, Ron A, and Langer R (1990) Polyanhydrides. In: Kroschwitz JI, Menges G, Mark HF, Bikales N, and Overberger CG (eds.) *Concise Encyclopedia of Polymer Science and Engineering*. New York: Wiley-Interscience.

Lichun L, Charles AG, and Antonios GM (1999) *In vitro* degradation of thin poly(DL-lactic-*co*-glycolic acid) films. *Journal of Biomedical Materials Research* 46: 236–244.

Ottenbrite RM, Huang SJ, and Park K (eds.) (1996) *Hydrogels and Biodegradable Polymers for Bioapplications. ACS Symposium Series* 627, Washington DC, pp. 2–10, 68–92, 244–245.

Ottenbrite RM (1998) *Frontiers in Biomedical Polymers*. Lancaster PA: Technomic Publishing.

Ottenbrite RM and Kim SW (2001) *Polymeric Drugs & Drug Delivery Systems*. pp. 39–41, 131–155. Lancaster PA: Technomic Publishing.

Ottenbrite RM and Kim SW (2001) *Polymeric Drugs & Drug Delivery Systems*. Lancaster PA: Technomic Publishing.

Rapoport N (2003) In: Dinh S and Liu P (eds.) *Advances in Controlled Drug Delivery*. ACS Symposium Book Series, pp. 85–101. Washington DC.

Schacht E, Vandorpe, Lemmouchi Y, Dejardin S, and Seymour L (1998) Degradable polyphosphazenes for biomedical applications. *Journal of Bioactive and Compatible Polymers*. In: Ottenbrite RM (ed.) *Frontiers in Biomedical Polymer Applications*, pp. 27–42. Lancaster, PA: Technomics Publishing.

Vert M, Feijen J, Albertsson A, Scott G, and Chiellini E (eds.) (1992) *Biodegradable Polymers and Plastics*. Stockholm, Sweden: The Royal Society of Chemistry.

Xuejun X, Sannino A, Ambrosio L, Netti PA, and Nicolais L (2004) *Journal of Bioactive and Compatible Polymers* 19(1): 5–13.

Zanzig J, Marimuthu B, Werka, and Scholz C (2003) *J. Bioact. & Compat. Polym* 18(5): 339–354.

Biomedical Materials

J R Jones and L L Hench, Imperial College, London, UK

The Need for Biomedical Materials

Improving healthcare and technology are increasing life expectancy but as one ages the body parts cannot maintain their function. Tissues such as bone and cartilage are needed to support the aging body even though the cells that produce them become less active with age. The heart, kidneys, and liver have to operate for much longer than ever before. This entry discusses how biomedical materials and biophysics techniques are being used in the development of regenerative medicine procedures. The aim of regenerative medicine is to regenerate diseased and damaged tissue to its original state and function.

Bone Defects

Bone is a natural composite of collagen (polymer) and bone mineral (ceramic). Collagen is a triple helix of protein chains, a complex structure that has high tensile and flexural strength and provides a framework for the bone. Bone mineral is a crystalline calcium phosphate ceramic (hydroxyapaptite (HA), $Ca_{10}(PO_4)_6(OH)_2$) that provides stiffness and the high compressive strength of the bone. The two most important types of bone are cortical and cancellous bone. Cortical bone is a dense structure with high mechanical strength and is also known as compact bone. Cancellous or trabecular bone is an internal porous supporting structure present in the ends of long bones such as the femur or within the confines of the cortical bone in short bones. The trabecular bone is a network of struts (trabeculae) enclosing large voids (macropores) with 55–70% interconnected porosity.

The mechanism for natural bone generation/regeneration in the body is the secretion of extracellular matrix by osteogenic cells (osteoblasts), which have developed (differentiated) from stem cells. The extracellular matrix is collagen type I, which mineralizes to form bone mineral, creating a composite of orientated collagen fibrils and apatite. The bone is remodeled in response to its local loading environment by the body. Osteoclasts are cells that resorb old bone and bone that is not required (i.e., not under any load), while osteoblasts lay down new bone.

Osteoporosis is a disease where bone resorption occurs faster than new bone is produced, causing the trabeculae to become thinner which leads to a reduction in total bone density and strength. The disease eventually leads to fracture of bones especially in the hip, wrist, knee, and spine. At present, when osteoporotic fracture occurs in knees and hips, joint replacement is often required. Millions of orthopedic prostheses made of bioinert materials have been implanted, an example of which is the Charnley total hip replacement, which is heralded as one of the most successful surgical inventions.

Long-term monitoring of 20 000 Charnley joints has revealed that it has a survivability of 76% after 25 years implantation, that is, 24% of hip operations required revision surgery. Improved metal alloys, special polymers, and medical-grade ceramics are the basis for this success, which has enhanced the quality of life for millions of patients. Reasons for failure tend to be aseptic loosening of the femoral stem, where bone resorption occurred due to a mismatch in the Young's modulus of the bone and the metal stem. Many modifications and variations of the Charnley joint have been developed over the years, including coating the metal stem with a synthetic HA layer that can bond to the bone mineral in the bone; however, survivability studies are so far only medium term and have shown similar results to that of the Charnley prosthesis.

A New Direction

All present day orthopedic implants lack three of the most critical characteristics of living tissues: (1) the ability to self-repair; (2) the ability to maintain a blood supply; and (3) the ability to modify their structure and properties in response to environmental factors such as mechanical load. All implants have a limited lifespan and as life expectancy is continually increasing, it is proposed that a shift in emphasis from "replacement" of tissues to "regeneration" of tissues is required to satisfy this growing need for very long-term orthopedic repair.

One way to restore diseased or damaged tissue to its original state and function would be the successful engineering of the replacement tissue in the laboratory. In a typical tissue engineering application, cells would be harvested from the patient (i.e., osteogenic cells in the case of bone) and seeded on a synthetic scaffold that acts as a guide and stimulus for tissue growth in three dimensions creating a tissue engineering construct or living biocomposite. The biocomposite would then be implanted back into the patient. Over time, the synthetic scaffold should be resorbed into the body as nontoxic degradation products at the same rate that the cells produce their own extracellular matrix.

In this article, the development of scaffold technology is discussed, focusing primarily on porous glass scaffolds that bond to bone and the growth of bone and cartilage in the laboratory on such scaffolds. The biophysics techniques that are now being used to monitor cell behavior as the cells grow are also discussed.

Biomedical Materials

Any material that is implanted into the body should be biocompatible, that is, not cytotoxic (not toxic to cells). There are three classes of noncytotoxic materials; bioinert, bioactive, and bioresorbable. No material is completely inert on implantation, but the only response to the implantation of bioinert materials is encapsulation of the implant by the fibrous tissue. Examples of bioinert materials are medical-grade alumina, stainless steels, and high-density polyethylene that are used in the total hip replacements.

Resorbable materials are those that dissolve on contact with body fluids and the dissolution products can be secreted via the kidneys. The most common biomedical resorbable materials are polymers that degrade by chain scission such as polyglycolic (PGA) and polylactic acids (PLLA) and their co-polymers, which are commonly used as sutures.

Bioactive Materials

Bioactive materials stimulate a biological response from the body such as bonding to tissue. Bone is a natural composite of bone mineral (HA) and collagen and other proteins. Therefore, synthetic HA (with $Ca/P = 1.667$) and other calcium phosphate ceramics, including coral, have gained much attention as a bone mineral substitute.

There are two classes of bioactive materials; class B bioactive materials bond to hard tissue (bone) and stimulate bone growth along the surface of the bioactive material (osteoconduction). Examples of class B bioactive materials are synthetic HA and tricalcium phosphate ceramics. Synthetic HA has been used in

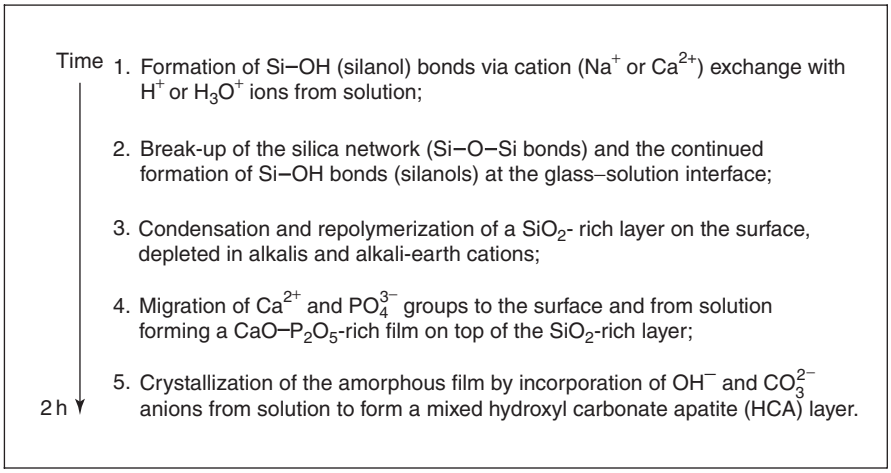


Figure 1 Mechanism of formation of the HCA layer on the surface of bioactive glasses in solution.

several clinical applications such as a bone defect filler and is used to coat the stainless steel or titanium alloy shaft of the total hip prosthesis so that it bonds to the thigh bone (femur). Tricalcium phosphate (β-TCP, with Ca/P = 1.5) is an osteoconductive material that is also resorbable in the body. β-TCP is usually used in conjunction with synthetic HA to improve the resorbability of HA in applications such as the filling of bone defects left by cysts, sinus floor augmentation, and bone cements.

Class A bioactive materials not only bond to bone and are osteoconductive but they are also osteoproductive, that is, they stimulate the growth of new bone on the material away from the bone/implant interface and can bond to soft tissue such as gingival (gum) and cartilage. Examples of class A bioactive materials are bioactive glasses. A certain composition of melt-derived bioactive glass (46.1% SiO_2, 24.4% Na_2O, 26.9% CaO, and 2.6% P_2O_5, in mol), called Bioglass® is used in the clinic as a treatment for periodontal disease (Perioglas®) and as a bone-filling material (Novabone®). Bioglass® implants have also been used to replace damaged middle ear bones, restoring hearing to thousands of patients.

The mechanism of bone bonding to bioactive materials is thought to be due to the formation of an HA layer on the surface of the materials after immersion in body fluid. This layer is similar to the apatite layer in bone and therefore a strong bond can form. The layer forms quickest on class A bioactive materials. The mechanism of formation of the HA layer on bioactive glasses is shown in **Figure 1**. Stage 1 involves release of cations. Stage 2 is the break up of the silica-based glass network. These two stages are dissolution processes; therefore, these glasses are not only bioactive but are also resorbable in the body. β-TCP ceramics are also resorbable.

Figure 2 shows suggested mechanisms for osteoproduction, where new bone grows on bioactive glasses away from the glass–host bone interface. The sequence follows on from the events in **Figure 1** but concentrates on the effects of the glass on the cells within the bone and bone marrow, however it does not fully explain osteoproduction. Recent findings by Professor Dame Julia Polak's team at the Tissue Engineering and Regenerative Medicine Centre at Imperial College, London, have shown that the dissolution products of bioactive glasses up-regulate seven families of genes that regulate osteogenesis and the production of growth factors.

Composites

To overcome the low toughness of HA, a polyethylene–HA composite was developed by Professor William Bonfield and colleagues at the University of London. It is clinically available under the name HAPEX™ and has been used as middle ear implants and orbital floor replacements. The polymer provides toughness strength and the HA increases the compressive strength and provides bioactivity so that implants are not encapsulated by scar tissues. The composite is class B bioactive and is easily shaped with a scalpel by the surgeon and so is now routinely used to replace the bones of the middle ear.

Scaffolds for Bone Repair

The Criteria for an Ideal Scaffold

To be able to regenerate the trabecular bone, a construct is required that will mimic the structure and mechanical properties of the trabecular bone and stimulate new bone growth in the shape dictated by the scaffold. Therefore, the scaffold should have a

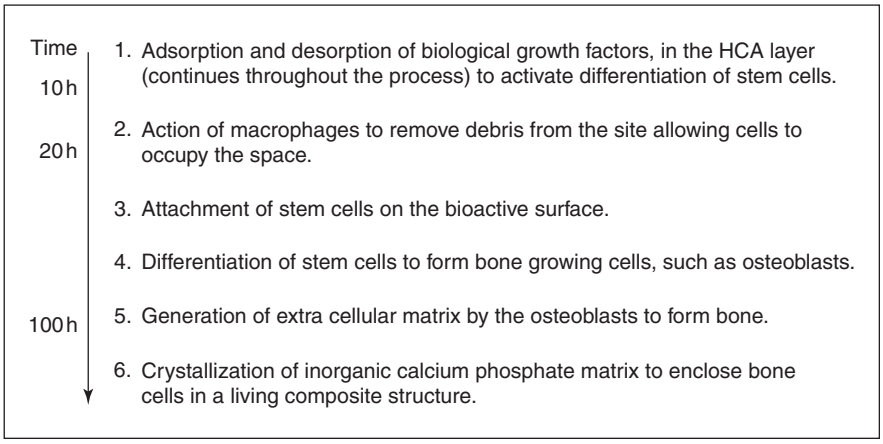


Figure 2 Suggested mechanism of cellular interactions between cells and bioactive glasses after implantation, after the formation of the HCA layer.

structure that acts as a template for tissue growth in three dimensions. The template must be an interconnected macroporous network containing pores with diameters in excess of 100 μm to promote cell migration, tissue ingrowth and vascularization, and nutrient delivery to the center of the regenerating tissue on implantation. The scaffold material should be able to bond to the host bone without the formation of scar tissue, that is, it should be made from an osteoconductive material. If the bone is to be restored to its original state, the scaffold should be resorbable so that eventually there is no trace of the scaffold's presence. The degradation rate of the scaffold must be controllable so that it can be tailored to match the rate of bone growth. The dissolution products would ideally influence the genes in the bone generating stem cells to stimulate efficient cell differentiation and proliferation. Importantly, the mechanical properties of the scaffold should match that of the host bone and not decrease too rapidly during degradation so that bone regeneration occurs at load-bearing sites.

If the scaffold is to be mass produced so that surgeons can use it in the clinic, it must be made from a processing technique that can produce irregular shapes to match that of the defect in the bone of the patient. It must have the potential to be producible to the required ISO (International Standards Organization) or FDA (Food and Drug Administration) standards and be easily sterilized.

The need for implant materials to pass the FDA standards may be necessary for the safety of patients, but it places a large financial obstacle in the path of new scaffold developments. Many researchers use materials that have already been passed by the FDA so that they are sure their scaffold will be safe for use in the clinic. They are unwilling to risk using materials that have not undergone FDA testing because of the time and financial commitment required for fulfilling the extensive tests.

Polymer Scaffolds

Resorbable polymeric scaffolds have been developed with porous structures similar to the trabecular bone. There are three reasons why much effort has been put into using polymers as scaffolds. First, polymers are easy to process in the shape of a 3D scaffold with a pore morphology suitable for a tissue engineering application with techniques such as phase separation, gas foaming, freeze drying, combined solvent leaching and extrusion, and computer-controlled rapid prototyping techniques. Polymer fibers can be woven together using textile processing techniques.

Second, polymers can have high tensile properties and high toughness. The mechanical properties of polymers can be controlled very easily by changing the molecular weight (chain length) of the polymer. Third, bioresorbable polymers have been successfully used as dissolving sutures for many years. Therefore, these degradable polymers, such as the esters PGA and PLLA, have passed FDA regulations and can be implanted into the body.

There are however some problems with using biodegradable polymers as scaffolds for bone regeneration. First, while a degradable scaffold is desired, sutures dissolve within 2 weeks of implantation, which is too rapid for bone regeneration applications.

Second, although resorbable polymers can be made with high tensile strength and toughness and their mechanical properties can be matched with collagen, their Young's modulus is much lower than that of bone, therefore these polymers cannot be used in load-bearing sites where they undergo compressive forces.

A third problem is that the mechanical strength of polymers decreases rapidly as they degrade. High tensile strength polymers have long chains (high molecular weight), which entangle with each other. As a tensile force is applied to the entangled chains, the chains begin to unravel until they become straight. The maximum tensile strength of polymers is reached when the chains are fully extended. Therefore, for a polymer to be tough, it must have a molecular weight over the entanglement value. The mechanism for degradation of biodegradable polymers is chain scission. The polymers undergo hydrolysis and the chains are cut in two. The average molecular weight of the polymer is therefore approximately halved with every chain scission event. The mechanical properties of the polymer are proportionate to the square root of the molecular weight and therefore decrease very rapidly as the polymer degrades. Biodegradable polymers have also been found to produce an inflammatory response due to the acidic by-products of the degradation.

Bioactive Ceramic Scaffolds

Ceramics are crystalline materials and tend to have high compressive strength and Young's modulus but low toughness, that is, they are brittle materials. Alumina and synthetic HA are ceramics that are most commonly used in biomedical applications. Alumina is a bioinert ceramic that is very hard and resistant to wear. It is therefore commonly used as a replacement to the ball of the femur in total hip replacement. Bioactive ceramics that have been used as bone regeneration materials are synthetic HA and β-TCA. Synthetic HA has been used most regularly because of its similarity to bone mineral; it has a similar structure and Young's modulus. β-TCP is similar to bone mineral in that they are both calcium phosphate ceramics, however β-TCP is resorbable.

In a porous form, HA and β-TCP ceramics can be colonized by bone tissue. A problem with introducing pores into a ceramic is that the compressive strength of the material decreases dramatically. The strength of the scaffold depends on the thickness and strength of the struts or pore walls. Generally, for the brittle compression of a foam

$$\sigma_{cr} \propto \rho_r^{3/2} \quad [1]$$

where σ_{cr} is the critical strength of the pore walls and ρ_r is the relative density, expressed as

$$\rho_r = \rho_b/\rho_s \quad [2]$$

where ρ_b is the bulk density of the scaffold and ρ_s is the skeletal (true) density of the material.

The simplest way to generate porous scaffolds from ceramics such as HA, is to sinter particles, preferably spheres of equal size. The scaffolds can then be pressed using cold isostatic pressing. As the sintering temperature increases, the pore diameter decreases, and mechanical properties increase as the packing of the spheres increases. Mechanical properties can be increased further by hot isostatic pressing (HIPing). High mechanical strengths can be achieved but the pore diameter is not high enough. Porosity can be increased by adding fillers such as sucrose, gelatine, and polymethyl methacrylate (PMMA) microbeads to the powder slurry, which burnout on sintering. However, this technique decreases the compressive strength to below that of the trabecular bone.

The majority of methods that are used to produce polymer foams cannot be applied to ceramic systems. However, a popular method for producing highly porous ceramics is to produce a polyurethane foam template that can be immersed in ceramic slurries under vacuum to allow the slurry to penetrate into the pores of the foam. The organic components are burnt out and the ceramic foams sintered (1350°C) producing a scaffold with interconnected pore diameters of up to 300 μm.

The ceramic slurries can also be foamed to obtain pores in the range of 20 μm up to 1–2 mm. The incorporation of bubbles is achieved by injection of gases through the fluid medium, mechanical agitation, blowing agents, evaporation of compounds, or by evolution of gas by *in situ* chemical reaction. A surfactant is generally used to stabilize bubbles formed in the liquid phase by reducing the surface tension of the gas–liquid interface. Surfactants are macromolecules composed of two parts, one hydrophobic and one hydrophilic. Owing to this configuration, surfactants tend to adsorb onto gas–liquid interfaces with the hydrophobic part being expelled from the solvent and a hydrophilic part remaining in contact with the liquid. This behavior lowers the surface tension of the gas–liquid interfaces, making the foam films thermodynamically stable, which would otherwise collapse in the absence of the surfactant.

The gel-casting method has been the most successful method used to produce macroporous bioactive HA ceramics with interconnected pores of greater than 100 μm in diameter. Suspensions of HA particles and organic monomers are foamed by agitation with the addition of a surfactant under a nitrogen atmosphere. *In situ* polymerization (cross-linking) of the monomers, which is initiated by a catalyst, creates a 3D polymeric network (gel). The porous gels are sintered to provide mechanical strength and to burnout the organic solvents. HA foams have been made with

compressive strengths in excess of 10 MPa, which is similar to that of trabecular bone. When the foams were implanted into the tibia of rabbits, bone partially filled the pores after 8 weeks and there was no inflammatory response. The compressive strength of the scaffolds that were colonized by bone trebled.

Hydrated silicon ($SiOH_4$) has been found to be a major contributor to the mineralization of bone and gene activation, which has lead to the substitution of silica (SiO_2) for calcia (CaO) into synthetic HA. *In vivo* results showed that bone ingrowth in silica-substituted HA granules was significantly greater than that into phase pure HA granules.

The disadvantages of an HA scaffold over a bioactive glass scaffold with similar morphology are that HA resorbs only very slowly, the dissolution products do not stimulate the genes in the osteogenic cells and HA is only osteoconductive as it generates bone at a slower rate than bioactive glasses. HA is still a bone replacement material rather than a regenerative material.

Bioactive Glass Scaffolds

Theoretically, gel-casting could be applied to melt-derived bioactive glass powders. However, bioactive glasses undergo surface reactions on contact with solutions to produce an HCA surface layer and it is desirable to have control over the reaction before a scaffold is ready for clinical use.

Sacrificial porogens and foaming agents have also been used to create melt-derived bioactive glass (Bioglass®) scaffolds. Large pores with diameters in the region of 200–300 μm were created but the total porosity was just 21% and there were large distances between pores, so this process failed to mimic the interconnectivity of the trabecular bone.

The most successful method for the production of bioactive glass scaffolds of similar structure to trabecular bone mineral is the foaming of sol–gel derived bioactive glasses. The sol–gel process involves polymer reactions of the glass precursors in a solution (sol). The sol is a solution of silica species that forms a gel by cross-linking together into a silica network. Sol–gel derived bioactive glasses tend to be more bioactive and resorb quicker than melt-derived glasses of similar compositions. The compositions of gel-glasses can also contain fewer components while maintaining bioactivity. This is due to them having a textural nano-sized porosity that is inherent to the sol–gel process, which provides a specific surface area of 150–600 $m^2 g^{-1}$, which is two orders of magnitude higher than melt-derived glasses. Along this surface there are many silanol groups that act as nucleation sites for HCA layer formation (**Figure 1**).

During the foaming process, air is entrapped in the sol under vigorous agitation as viscosity increases and the silica (–Si–O–Si–) network forms. A surfactant is added to stabilize the bubbles at short times. As the porous foam becomes a gel, the bubbles are permanently stabilized. The gel is then subjected to controlled thermal processes of aging (60°C) to strengthen it, drying (130°C), and thermal stabilization/sintering to remove organic species from the surface of the material (500–800°C). Bioactive glass foam scaffolds can contain macropores up to 600 μm in diameter, connected by pore windows with modal diameters in excess of 100 μm and compressive strengths up to 2.5 MPa. **Figure 3** shows a scanning electron microscopy (SEM) micrograph of a pore network of a typical bioactive glass foam. *In vitro* cell studies using primary human osteoblasts have shown the foams stimulate formation and mineralization of bone nodules within 2 weeks of culture (**Figure 4**). **Figure 4** shows osteoblasts that have attached to the concave surface of the pore. The cells have proliferated, coating the pore in an organic layer, and have released extracellular matrix, which has mineralized to form bone mineral. *In vivo* studies have shown that foam scaffolds implanted on rabbit crania stimulated new bone growth at a similar rate to that of the melt-derived bioactive glass powder available commercially.

Spectroscopy and Biophotonics

Vibrational spectroscopy is one of the most common techniques to study the surface reactions of bioactive materials. It is nondestructive, rapid, and can be used to analyze small areas of a sample surface. The two most commonly used vibrational spectroscopy

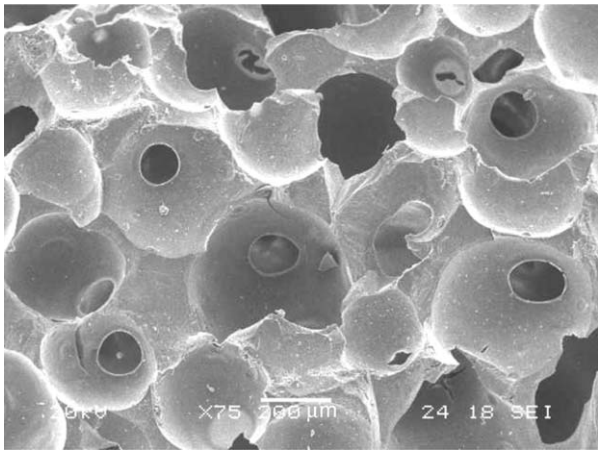


Figure 3 Scanning electron micrograph of a bioactive glass foam scaffold.

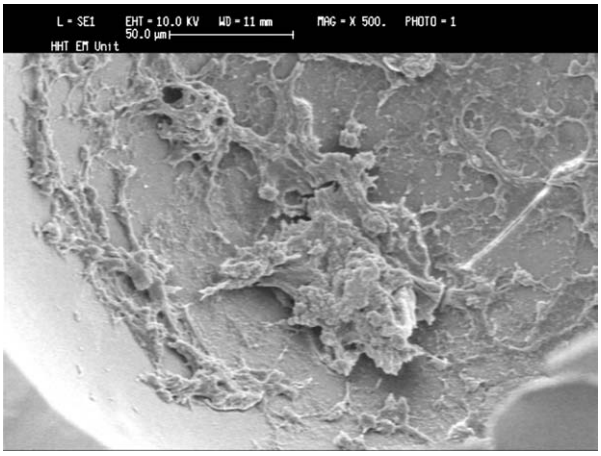


Figure 4 Scanning electron micrograph bone cells and mineralized bone extracellular matrix in a pore of a bioactive glass foam scaffold. (Courtesy of J Gough.)

techniques are Fourier transform infrared spectroscopy (FTIR) and Raman spectroscopy.

Fourier Transform Infrared Spectroscopy

FTIR spectra have been used to determine the rate of formation of the HCA layer on bioactive glasses and the technique is used as a quality assurance (QA) test for bioactivity. FTIR spectra of the glass surface area usually taken after a glass sample has been reacted in simulated body fluid (SBF) and dried. The crystalline HCA layer is characterized by the P–O bending vibration bands at 560 and 604 cm^{-1} (infrared wave number). Stretching and bending vibration bands for the amorphous silica network are also present in the spectra. The relative intensity of the bands can be used to monitor growth of the layer as a function of time. **Figure 5** shows FTIR spectra of three forms of the sol–gel derived bioactive glasses of the 58S composition (60 mol.% SiO_2, 36 mol.% CaO, 4 mol.% P_2O_5) that were soaked in SBF for 2 h. The spectra show that the HCA layer thickness was thicker for the powders and foams with respect to the monolith. Although FTIR can be used for the characterization of dried materials after soaking or after implantation and removal, *in situ* measurements are hindered by the strong adsorption of water in the infrared region.

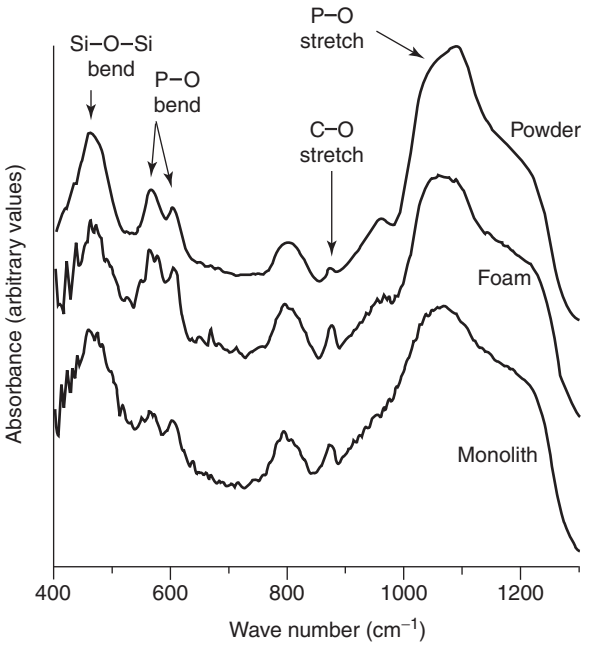


Figure 5 FTIR spectra for 58S glass powder, foam, and monolith after 2 h immersion in SBF at 37°C and 175 rpm agitation.

Bio-Raman Spectroscopy and Biophotonics

Raman spectra are fingerprints of the chemical composition of the material. Raman spectra are obtained by collecting electromagnetic radiation that has been scattered by molecules (Raman effect). The spectral shifts depend on the vibrational frequencies of the molecules. The low Raman signal and low Raman scattering of water, and the combination of Raman spectroscopy with optical microscopy (water immersion objectives) has allowed noninvasive *in situ* monitoring of cells in cell culture. This technique is called micro-Raman spectroscopy. The optical objective provides high spatial resolution and is used to locate the cell or part of the cell, such as the cytoplasm or the nucleus, from which a spectrum will be collected. A laser is then fired at the identified area and the scattered photons are collected by a detector.

Low scattering efficiency of the cells makes the measurement of Raman shifts difficult. Raman signals can be enhanced using UV lasers; however, high-powered lasers can denature the cells and change cell phenotype. Decreasing the laser power decreases the Raman signal, but reduces the risk of changes to the cell. A laser wavelength of 785 nm has been found to produce optimal *in situ* spectra of living cells. A single spectrum obtained from the nucleus of a cell can contain information about the protein and DNA content of the cell. **Figure 6** shows a Raman spectrum of a cultured lung cell containing vibrational bands corresponding to DNA (A, G, T, C adenine, guanine, thymine, cytosine), BK: backbone, RP: ribose-phosphate and the proteins phenylalanine (Phe) and tyrosine (Tyr). Raman spectroscopy using 785 nm lasers can be used to monitor the biochemical changes taking place in a cell. During cell death, for example, the degradation of proteins, DNA breakdown, and the formation of lipid vesicles can be detected in real time.

Cells cultured on bioactive materials can also be monitored by *in situ* Raman spectroscopy. A

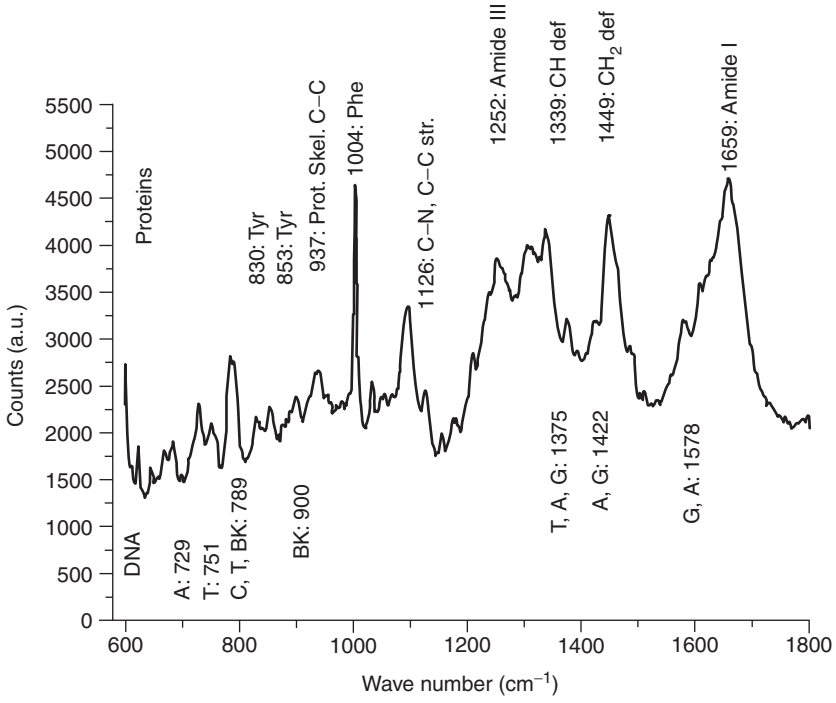


Figure 6 Raman spectrum of a cultured living lung cell. (Courtesy of I Notingher.)

background spectrum for the unreacted material can be subtracted from the spectrum of a cell on the material. The resultant spectrum will then also contain vibrational bands corresponding to the formation of any calcium phosphate layer on the surface of the bioactive material as a function of time.

Raman spectroscopy may hold the key for the *in situ* monitoring of bone cells growing on scaffolds. When acquisition techniques and interpretation of the spectra are perfected, the micro-Raman spectrometers can be linked to bioreactors and the effect of flow rates, growth factors, and dissolution products on cell behavior can be monitored in real time.

Summary

Numerous materials have been developed to replace living tissues. They are bioinert (stainless steel, alumina), bioresorbable (polyglycolic acid), or bioactive (synthetic HA and bioactive glasses). New biocomposites have been made that combine bioinert and bioactive properties. The latest generation of materials are used in combination with cells. Many techniques have been employed to produce porous scaffolds. The scaffold that most closely matches the criteria for an ideal scaffold and that most closely mimics the structure of the trabecular bone is the bioactive glass foam. However, these scaffolds do not yet have the mechanical properties to be directly implanted into load-bearing sites. Ideally, a bioactive glass/polymer foam composite would be developed with improved toughness. The properties of a scaffold must be optimized with respect to the cell response to the scaffold. Biophotonics techniques such as Raman spectroscopy have the potential to be able to monitor cell activity and intracellular processes of living cells in real time noninvasively.

Acknowledgments

Lloyds Tercentenary Foundation, EPSRC and MRC (UK), Dr. Julie Gough, Dr. Ioan Notingher.

See also: Bioelectronics; Biomembranes; DNA and RNA, Biophysical Aspects; Electromagnetic Biological Effects; Genetic Algorithms for Biological Systems; Ion Transport in Biological Membranes; Neuroscience.

PACS: 87.80.Rb; 81.05.Je; 81.05.Rm; 81.05.Kf; 81.05Qk; 81.20.Fw; 82.39.Wj; 87.64.Je

Further Reading

Brinker CJ and Scherer GW (1990) *Sol–Gel Science: The Physics and Chemistry of the Sol–Gel Process*. Boston: Academic Press.

Chalmers J and Griffiths PR (2001) *Handbook of Vibrational Spectroscopy Volume 5: Applications in Life, Pharmaceutical and Natural Sciences*. New York: Wiley.

Davies JE (2000) *Bone Engineering*. Toronto: EM^2.

Gibson LJ and Ashby MF (1988) *Cellular Solids Structure and Properties*. Oxford: Pergamon.
Hastings GW and Williams DF (1980) *Mechanical Properties of Biomaterials*. New York: Wiley.
Hench LL, Jones JR, Lenza RFS, and Vasconcelos WL (2003) Tissue engineering. In: Banner N, Polak JM, and Yacoub M (eds.) *Lung Transplantation*, pp. 367–373. Cambridge: Cambridge University Press.
Hench LL and Polak JM (2002) Third generation biomaterials. *Science* 295(5557): 1014–1020.
Hench LL and Wilson J (1993) *An Introduction to Bioceramics*. Singapore: World Scientific.
Jones JR, Sepulveda P, and Hench LL (2002) Bioactive materials for tissue engineering scaffolds. In: Polak JM, Hench LL, and Kemp P (eds.) *Future Strategies for Tissue and Organ Replacement*, pp. 3–19. London: Imperial College Press.
Notingher I, Verrier S, Haque S, Polak JM, and Hench LL (2003) Spectroscopic study of human lung epithelial cells (A549) in culture: living cells versus dead cells. *Biopolymers (Biospectroscopy)* 72: 230–240.
Park J and Lakes RS (1992) *Biomaterials: An Introduction*, 2nd edn. New York: Plenum.
Thomson RC, Yaszemski MJ, and Mikos AG (1997) Polymer scaffold processing. In: Lanza RP, Langer R, and Chick WL (eds.) *Principles of Tissue Engineering*, pp. 263–271. Austin, TX: Landes.
Wise DL (2000) *The Biomaterials and Bioengineering Handbook*. New York: Dekker.

Biomembranes

F Bordi and C Cametti, Università degli Studi di Roma "La Sapienza," Rome, Italy

Introduction

A biomembrane can be considered as a barrier of highly selective permeability that allows and regulates the traffic of a myriad of different molecular species between the interior of the cell and its surrounding environment (as, e.g., in a plasmatic membrane), or between different compartments within the same cell (e.g., in the membrane of the different organelles). Its basic structure is formed by a bilayer of amphiphilic molecules, the most common of which are three different classes of lipids (glycerolipids, sterols, and sphingolipids). On the extracellular side, this bilayer couples to a glycocalix, a carbohydrate network composed of oligosaccharides believed to be responsible for cell–cell recognition and adhesion. On the intracellular side, the bilayer couples to the cytoskeleton, that contributes to the mechanical properties of the overall composite structure.

The main functions exerted by a biomembrane can be summarized as follows: (1) the membrane is a selective filter which controls the transport and the permeation of ions, molecular aggregates, and even large particles between the extracellular medium and the cytosol; (2) the membrane is the site for energy producing processes and for hormone signal transduction; (3) the membrane acts as receptor for extracellular signals and mediates the communications between intra- and extra-cellular media; and (4) the membrane can perform mechanical tasks as in cellular motion or in eso- or endo-cytosis processes.

These demanding tasks are fulfilled thanks to a very complex, yet in principle, surprisingly simple, basic structure based on two-layered sheets of lipid molecules held together by a delicate balance between hydrophobic and hydrophilic interactions (**Figure 1**). The first important fact is that this is a self-assembling structure. The balance between a favorable interaction of the hydrophilic polar heads of the lipids with the aqueous solvent and the unfavorable hydrophobic interactions of their aliphatic chains, represents the driving force to the spontaneous formation in an organized closed structure. Due to hydrophobic/hydrophilic interactions, the lipid molecules arrange themselves into two-faced leaflets (the double layer) with the hydrophobic

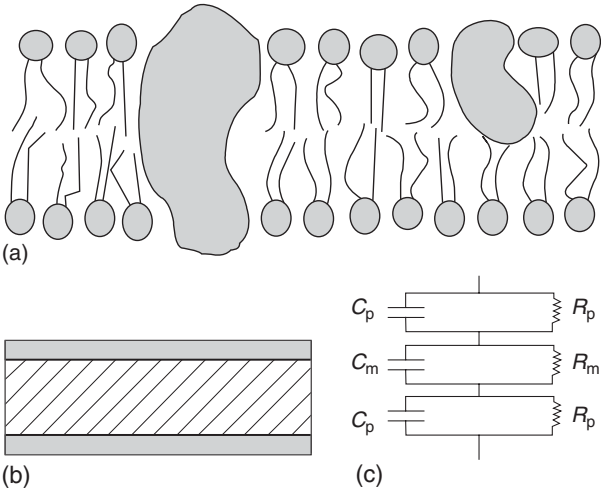


Figure 1 A sketch of a biological cell membrane as seen by a chemist ((a) a lipid bilayer with membrane proteins) and by a physicist ((b) a hydrophobic layer covered by a hydrophylic interface). (c) The passive equivalent circuit of a plasmatic membrane. C_p, R_p and C_m, R_m represent the capacitance and resistance per unit surface of the hydrophilic and hydrophobic regions, respectively.

moiety in the inside and "polar heads" on the outside, representing the interface with the aqueous environment. The presence of lipids of different shapes (different volume ratios of the hydrophobic/hydrophilic moieties) in the two leaflets stabilizes the radius of curvature of the double layer.

The second important aspect is that, being mainly stabilized by "structure-solvent" effects, without strong "links" between different molecules, the bilayer is a highly dynamic structure, where lipids (and proteins) can flex, rotate, and diffuse laterally, as in a "two-dimensional" fluid. The lipid distribution results from a continuous inward and outward movement between the two monolayers, the lipid asymmetry being maintained by specific mechanisms that counterbalance the concentration-driven transbilayer permeation (**Figure 2**). The lipid bilayer has often been considered, in the past, as a passive structure and only recently, more attention is being directed towards its implications in a variety of membrane biological functions.

Due to its peculiar structure, the overall shape of the membrane during the biological functionality can be easily changed, for example, during pseudopodia movement or eso-/endo-cytosis processes. In endocytosis, in particular, after a region of the membrane has surrounded the particle to be introduced within the cell, it detaches from the rest, forming a vacuum that can move within the cytoplasm. The whole process is reversed in eso-cytosis, when a substance has to be expelled from the cell. These processes well illustrate the "modularity" of the membrane structure, that can easily lose or acquire parts to accomplish particular tasks.

Among the enormous variety of possible lipids, only a few classes are used to build up a biomembrane. Lipids that exert a predominantly structural role can be grouped into three categories: sterols (e.g., cholesterol), phospholipids (e.g., phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), sphingomyelin (SPHM)), and sphingolipids (cerebrosides, glycolipids). Lipids with a predominantly functional role are: phosphatidylinositol, phosphatidic acid, and many types of gangliosides.

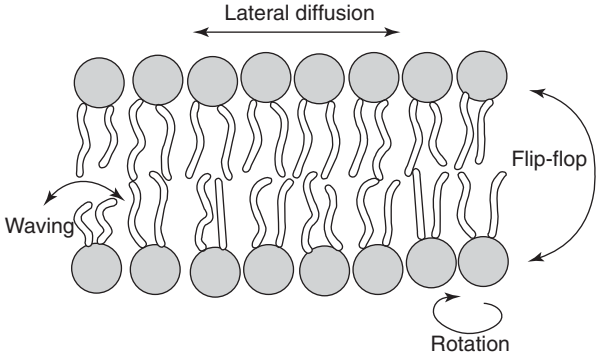


Figure 2 Movements of phospholipids in bilayers. The lipid distribution results from lateral diffusion, flip-flop, waving and rotation movements between the two leaflets of the membrane.

The lipid composition of different cell types shows great differences and within the same cell, the plasma membrane, the nuclear envelope, mitochondria, and other membranous organelles and structures have different lipid compositions. Depending on the cell type, proteins constitute 20–80% of the membrane mass. Nevertheless, by comparing the average molecular weight of lipid molecules (in the range 700–1000 Da) and the typical molecular weight of a protein (in the range of the order of tens of thousands Da), lipids are the more abundant components in terms of molar concentration.

Membrane proteins are usually divided into two classes, "intrinsic" and "extrinsic" ones. This differentiation is mainly based on the greater or lesser difficulty in separating the fraction of the protein of the two classes from the other membrane components by organic solvent extraction. However, intrinsic proteins are generally considered more deeply embedded into the nonpolar environment (the interior of the lipid double layer), while extrinsic proteins are assumed more "at the surface" or even simply adsorbed at the double-layer interface. However, most intrinsic proteins possess specific sites that are exposed to the aqueous environment on one, or on both the sides of the membrane.

The thickness of a lipid bilayer can be estimated in the range of 4–6 nm. However, in estimating the thickness of a real biological membrane, the fact that proteins and many lipids (glycolipids) bear groups that extend into the aqueous medium, has to be taken into account. This region (glycocalix), at the outer surface of the plasmatic membrane, is typically 10–100 Å thick.

Membranes of cells and organelles are sometimes corrugated and folded or show protrusions of different shape and size. For example, the needle-like protrusions of the plasmatic membrane of some cells (microvilli) typically measure 50–200 nm in diameter and may extend for several hundreds of nanometers and even more.

Functional Cell Membrane Domains and Influence of the Cytoskeleton on the Organization of the Cell Membrane

One intriguing issue in membrane biophysics is the presence within the plasma membrane of lipid

microdomains, termed rafts, which are believed to be essential for its complex activity. Although the present view of structural organization of a biological membrane is still deeply indebted to the fluid-mosaic model proposed by Singer and Nicholson in 1972, consisting of a fluid-lipid bilayer where lipids and proteins are more or less randomly distributed, recently a considerable amount of experimental work suggests that biomembranes are not laterally homogeneous, but floating domains with distinct lipid and protein composition. However, it is not yet completely understood what the mechanism at the molecular level is, that determines the composition of these domains and what their precise functional roles are. Moreover, in addition to the random motion predicted by the fluid mosaic model, a variety of lateral transport modes exist for several membrane proteins.

Microdomains have been detected in certain lipid mixtures, for example, PC and cholesterol, that exhibit coexisting liquid phases under definite conditions of temperature, lateral pressure, and composition. In this case, the size of membrane domains ranges from several hundred lipid diameters (100–200 nm) to one micrometer.

The biological membrane system where the existence of lateral domains has been now evidenced with certainty, is the plasma membrane of mammalian cells. In living cells, raft domains appear to be very small and probably heterogeneous. This may explain why they have escaped direct visualization by microscopical techniques. Indirect evidence for small rafts was obtained using single-particle tracking of the thermal position fluctuation, showing that raft-associated membrane proteins are stably associated to a small cholesterol-dependent lipid assembly of ~ 50 nm in diameter.

In this case, domains are lipid and protein supramolecular complexes which include, for example, multimers of receptor molecules or receptor and effector molecules that are thought to be the first trigger for the cellular reactions that follow ligand binding. Moreover, there are specialized membrane domains, such as synapses, caveolae, and cell–cell and cell–substrate adhesion structures, where specific proteins and lipids are assembled together to carry out specific functions.

The domain distribution does not cover the whole membrane area, and domains (rafts) are in dynamic equilibrium with more homogeneous and fluid membrane regions. These domains are thought to form because they represent the energetically most favorable packing of the membrane components. However, these rafts should be considered as being "soft" structures that, despite a high degree of local order, do not have a long-range order.

Where the fluid-mosaic model fails is in the assumption of a uniform protein distribution within the double layer. Limiting ourselves to the transport within the membrane, a great deal of evidence has now been accumulated, showing that a variety of 'intra-membrane' movement modes exist for membrane proteins. In addition to the simple Brownian diffusion assumed by the fluid-mosaic model, membrane proteins also appear in a stationary phase and experience different diffusion modes: directed diffusion, confined diffusion (the particle can diffuse freely but within a restricted area), diffusion in a harmonic potential well (this is, e.g., the diffusion mode experienced by a protein bound elastically to the cytoskeleton).

In particular, single-particle tracking techniques, together with a careful theoretical and statistical analysis of the particle tracks, have shown that the plasma membrane is compartmentalized with regard to lateral diffusion of transmembrane proteins although, in many cases, inter-compartmental hopping diffusion is observed.

The relationship of lipid rafts to the zones detected by single-particle tracking, where proteins do not experience free diffusion (transient confinement zones (TCZs)) is still controversial. The membrane lipid composition and, in particular, the abundance of cholesterol appears to have an effect on TCZ abundance. However, experimental evidence, the observed 'hopping' between different zones, in particular, indicates that trapping cannot simply be ascribed to particles encountering a much more viscous region, and suggests that some types of barriers are involved.

To account for the presence of these barriers, two models have been already proposed. These models, of course, do not exclude each other and some combinations of them occur. In the first model, which has been given the suggestive title of "membrane-skeleton fence model," the membrane skeleton provides the barrier to the free diffusion of membrane proteins. The idea is that the space between the membrane and the cytoskeleton is too small to allow the cytoplasmic portion of any membrane protein to pass. The membrane compartment would hence ensue from a steric hindrance, and would be different for proteins that extend differently into the cytoplasm. The hopping diffusion between adjacent compartments should result from the combination of two favorable conditions. As a result of the dynamic properties of the cytoskeleton, the distance between the membrane and the skeleton may fluctuate (the membrane skeleton probably dissociates and reassociates from the membrane in a dynamic equilibrium). In this way, the membrane protein molecules that have sufficient kinetic energy and that approach the compartment boundaries will be able to cross the

barrier. This model has received strong support from experiments on red blood cells. It has been observed that an impaired cytoskeleton (as it is found in erythrocytes of patients affected by different forms of hereditary hemolytic anemia) causes an enhancement of the lateral diffusion of "band 3" protein, a typical transmembrane protein of erythrocyte membrane, while its rotational diffusion remains practically unaffected.

In the second model, TCZs would be delimited by the presence at their boundaries of a high concentration of "obstacles." The protein confinement within a small zone should result from a repeated bouncing of the protein between the obstacles, in a sort of "pinball effect." The obstacles could be membrane proteins directly linked to the cytoskeletal network. This mechanism may justify the transient confinement of proteins that are known not to have parts in the protrusion into the cytoplasm and that, for this reason, cannot directly interact with the cytoskeleton.

Also, the directed diffusion of membrane protein has been ascribed to two different mechanisms. Analogous to contractile movements that allow cell locomotion, the first mechanism involves the concerted action of a number of molecular motors to move entire regions of the cortical cytoskeleton and, with it, patches of cross-linked membrane proteins.

Moreover, a protein can in principle undergo directed diffusion exploiting the skeleton fence mechanism again, with elements of the cytoskeleton tracing a sort of corridor underneath the membrane, within which the protein can freely diffuse.

Membrane Electrostatics

Biological membranes carry a great number of polar or ionized groups. The net charge is usually negative, because most of the membrane proteins and all charged natural lipids have their isoelectric point below neutral pH (in human blood, e.g., physiological pH must be maintained within narrow limits ~ 7.4). The charge density is not usually very high, typically less than $0.05\,\mathrm{C\,m^{-2}}$ (corresponding to about 1 elementary charge on a surface of $250\,\text{Å}^2$) and much lesser than the charge density in artificial lipid double layer, where the surface charge density can reach values as high as $0.25\,\mathrm{C\,m^{-2}}$ (1 elementary charge per $50\,\text{Å}^2$).

From an electrical point of view, a biomembrane can be considered as an extremely thin ($\sim 50\,$Å) hydrophobic isolating interface between two conducting phases. The passive electrical behavior of a membrane can be described as an R–C element network, whose specific capacity C_s is given by $C_s = \varepsilon_0\varepsilon_s/d$ and the specific conductance G_s by $G_s = \sigma_s/d$, where ε_s and σ_s are the membrane permittivity and the membrane conductivity, d the membrane thickness, and ε_0 the dielectric constant of free space. For a normal cell membrane, different experimental procedures yield $C_s \simeq 1\ \mu\mathrm{F\,cm^{-2}}$ and $G_s \simeq 1\ \Omega^{-1}\,\mathrm{cm^{-2}}$ that, assuming a membrane thickness $d \approx 80\,$Å allows one to evaluate the membrane permittivity and the electrical conductivity to be of the order of $\varepsilon_s \simeq 5$ and $\sigma_s \simeq 10^{-4}\ \Omega^{-1}\,\mathrm{cm^{-1}}$.

Fixed surface charges give rise to an electrostatic surface potential that, in turn, results in the formation of a diffuse electrical double layer. An accumulation of charged species near the membrane surface influences the transmembrane transport mechanisms. The conformation and function of many molecules is generally affected by electrostatic fields. As a consequence, the membrane charged surface acts as a catalytic site and, besides the transmembrane transport, the changes in the surface charge can also influence recognition and other biochemical reactions that take place at its surface.

Using as an obligatory starting point the Poisson–Boltzmann equation with the exponential space charge density expanded to the first order in the electrostatic potential, the electrostatic potential Φ, and consequently the space charge density ρ, in the simplified case of a 1 : 1 valent electrolyte solution, can be calculated as a function of the distance from the membrane surface (assumed as an infinite plane) from the expression (Gouy–Chapman theory)

$$\Phi(x) = \frac{2k_BT}{e}\ln\frac{1+\gamma e^{-kx}}{1-\gamma e^{-kx}} \qquad [1]$$

Here, γ is a constant that depends on the electrostatic potential calculated at the surface Φ_0

$$\gamma = \tanh\left(\frac{e\Phi_0}{4k_BT}\right) \qquad [2]$$

and k is the Debye constant

$$k = \left(\frac{2\rho_\infty e^2}{k_BT\varepsilon_0\varepsilon_r}\right)^{1/2} \qquad [3]$$

that depends on the ionic force of the solution. k_BT is the thermal energy.

In eqn [3], ρ_∞ is the electrolyte concentration far away from the surface. The Debye length can be considered approximately as the thickness of the diffuse double layer at the membrane surface. Equation [1] combined with the boundary condition $d\Phi/dx|_{x=0} = -\sigma/\varepsilon_0\varepsilon_r$ gives the following relation, connecting the potential Φ_0 and the surface charge

density σ

$$\sigma = \sqrt{8\rho_{\infty}\varepsilon_0\varepsilon_r k_B T}\,\sinh\left(\frac{e\Phi_0}{2k_B T}\right) \qquad [4]$$

It must be considered that, in the case of biological membranes, the charge cannot be considered, not even approximately, homogeneously distributed on a flat surface. Charges are, in part, borne by the groups attached to the lipid head groups and to the surface proteins that extend into the solution, forming the glycocalix. The thickness of this zone is often comparable with the electrical double-layer thickness obtained from eqn [3]. Consequently, the charge should be considered distributed tri-dimensionally within a layer at the membrane surface, and not with a bi-dimensional distribution. Besides all the approximations of the model, this is another reason for taking the values calculated in the framework of the Gouy–Chapman theory with some warning. Models based on Bogolyubov–Born–Yvon–Green or Ornstein–Zernike equations, together with a suitable closure relation, give a better description of double-layer behavior and characteristics, but at the cost of a more complicated formalism. In the understanding of the electrokinetic phenomena that take place at the membrane surface, the Gouy–Chapman theory often gives helpful hints as a working model, provided that the higher-order effects, mainly arising from ion–ion interactions and from the nonuniform, tri-dimensional charge distribution, are negligible.

Across many biological membranes exists an electrostatic transmembrane potential also, $\Delta\Phi$, that is, a difference in the electrostatic potential between the two surfaces of the membrane.

Different ion-surface affinities lead to the formation of a nonuniform layer of ions more or less intimately associated with the membrane surface. This affinity can be different in the two surfaces of the membrane and contribute to the transmembrane potential.

Another contribution stems from the "dipole potential" that arises from the dipole moment of the membrane constituents, in particular from the headgroups of (zwitterionic/charged) lipids, the carbonyl groups of their hydrocarbon chains, and from the interfacially bound water layers.

However, when the electrolyte composition at the two membrane sides is different, the main contribution to the transmembrane potential comes from the Nernst diffusion potential. To the first approximation, Nernst potential is proportional to the logarithm of the ratio of the concentrations on the two sides of the membrane, according to

$$\Delta\Phi_N = \frac{k_B T}{ze}\ln\left(\frac{c_1}{c_2}\right) \qquad [5]$$

Such a transmembrane potential is found in almost all living cells.

The efficiency of the membrane as a barrier to the passage of charged molecules is mainly due to the high energy cost of transferring a charge from the aqueous environment into the apolar core of the membrane. Membrane pores or the formation of ion-carrier complexes can reduce this energy requirement dramatically. The membrane surface electrostatics plays a role in transmembrane transport by affecting the interfacial ion concentration and by increasing the probability of a pore or an ion-carrier complex formation.

Experimental Techniques

A large variety of experimental methods for the study of the biomembrane structure and dynamics have been developed, the most important of which are nuclear magnetic resonance (NMR), electron spin resonance (EPR), patch clamp, and Langmuir–Blodgett technique, single-particle tracking, radiowave dielectric spectroscopy, and many more. A brief survey on the applicability to the membrane of each of them follows.

Single-Particle Tracking

In "single-particle tracking," the membrane molecular species whose diffusion has to be investigated is marked by specifically bound colloidal gold particles or by fluorescent molecules. The concentration of these probes is kept very low, in such a way that only few of them appear within the field of an optical microscope, equipped with a digital camera. From a statistical study of deviations from a purely diffusive Brownian motion, information on membrane structure and particle–particle interactions can be obtained. In a different approach, known as "fluorescence correlation spectroscopy," the analysis of the motion of individual molecules is replaced by the analysis of the time-averaged fluctuations of the fluorescence signal, deriving from a very small volume, at the membrane surface. The requirement of selecting the signal from a volume small enough to make fluctuations observable is fulfilled by using a confocal optical microscope. In this way, an "open" volume optically selected $\sim 10^{-15}$ l, in which molecules can freely go in and out, can be investigated.

Patch Clamp

Electrical currents flowing through the cell membrane can be measured by means of a particular technique known as "patch clamp" developed in the early 1970s by B Sakmann and E Neher, who for their research in this field obtained the Nobel Prize in

physiology in 1991. A microelectrode of suitable size can be obtained by rapidly pulling a glass capillary with an inner diameter of one millimeter or less in size, intensely heated in the middle (a specially designed "puller" has been devised to this scope). The capillary ultimately breaks into two hollow spikes at the tip, with a diameter of a fraction of a micrometer. The rim of the hole on the tip is then polished and smoothed with a microforge. These "micropipettes" are filled with a suitable electrolyte. The micropipette is then implanted through the membrane into the cell. Two metal electrodes, one within the micropipette and the other in the bathing solution around the cell, connected to a picoamperometer, allow the measurement of the electrical current flowing through the membrane. In appropriate configurations, the current through the membrane of the whole cell or through "patches" cut from the cell membrane can be measured. In the last case, it is possible to measure the current flowing through an individual protein channel. With this technique, the protein channels responsible for the selective transport of different ions and charged molecules have been identified and their properties studied thoroughly, discovering, for example, that many of them are controlled by the voltage difference across the membrane (voltage-gated channels). Two different electrical arrangements are usually employed. In the "current-clamp" technique, the current is kept constant and the voltage recorded. In "voltage-clamp," conversely, the voltage is kept constant and the current is the measured signal. Voltage-clamp results are more suitable for very low-current intensities and in studies on voltage-gated channels, where the voltage level is used to control the opening and closing of the channel.

Langmuir–Blodgett Film

A suitable model of a biological membrane is represented by a lipid monolayer, where the nature and the packing of the lipid molecules, the composition of the subphase, and the temperature can be chosen without any limitation. The main characteristics of these systems consist in their homogeneity, stability, and planar geometry, with lipid molecules undergoing a well-defined orientation, providing a bi-dimensional system.

Langmuir monolayers are produced by dissolving at the air–water interface, a known amount of the amphiphile under investigation in a Langmuir trough. After volatile solvent evaporation, the monolayer extension is compressed by two movable barriers and monitored with the surface-pressure area isotherms, that is, a plot of the change of the surface pressure as a function of the area available to each molecule at the aqueous interface. The lowering of the surface tension at the interface caused by the presence of the amphiphile is given by $\Pi = \gamma^0 - \gamma$, where γ^0 and γ are the surface tension in the presence and in the absence of the amphiphile. Π is a measure of the surface pressure needed to prevent the film from spreading. At very low pressure, Π obeys a two-dimensional equivalent ideal gas law

$$\Pi A = k_B T \qquad [6]$$

where A is the area occupied by each molecule and $k_B T$ the thermal energy.

These monolayers undergo a surface potential due to the group dipole moments of the film-forming molecules. By considering the monolayer as a parallel plate condenser, comprising a sheet of uniformly distributed dipoles, the surface potential ΔV is given by

$$\Delta V = \frac{4\pi N_0 \mu_n}{\varepsilon_r A} \qquad [7]$$

where N_0/A represents the array of N_0 dipoles by area A and μ_n the normal component of the dipole moment per molecule. The potential of a monolayer-free subphase is taken as reference. The measurement of the surface potential can be carried out by the ionizing electrode method or the vibrating plate method. In the first case, the ionization of the air above the film, induced by an α-emitter, allows the potential difference between two electrodes, above and below the aqueous subphase, to be measured. In the latter, the vibrating electrode of one of the plates of a condenser generates an alternating signal that can be measured using a high-gain amplifier.

NMR and EPR Spectroscopies

A good picture of the characteristics of lipids and protein motion in membranes can be derived from both EPR and NMR techniques. In EPR, the study of the conformational flexibility of the membrane involves the use of spin labels containing an unpaired electron available to observe transition between different energy levels induced by the presence of the external magnetic field. Spin labels of different structures have been used to probe the behavior of lipid hydrocarbon chains in the membrane.

NMR spectroscopy can provide information about the structure and dynamics of membrane proteins at a molecular level and, in particular, the molecular arrangement in the bilayer. For example, because of the anisotropy nature of the bilayer, the order parameter, related to the quadrupole splitting arising in a ^{2}H-NMR spectrum, measured at different positions

in the alkyl chain, offers a detailed picture of the molecular bilayer arrangement. During the past decade, considerable advances in both NMR and computational methodologies have taken place, opening up the possibility of investigating structural and dynamic properties of systems that may not be suitable for crystallography. NMR can reveal dynamic information about regions of macromolecular structure, and conformation or structural changes that occur as a result of ligand binding. For example, by the ^{31}P-NMR method, the fraction of the phospholipid molecules at the outer leaflet of the bilayer can be determined. Moreover, dynamical parameters such as rotational and translational diffusion coefficients, the rate of conformational changes, the activity of membrane embedded proteins up to interactions between membranes and other macromolecules can be appropriately investigated.

Radiowave Dielectric Spectroscopy Technique

Dielectric spectroscopy in the radiowave frequency range furnishes a valuable method to investigate the electrical properties of a biological cell membrane, both concerning the structural properties (charge and dipole moment distribution) described by the permittivity, and the dynamical properties (ionic transport across the membrane) described by the electrical conductivity. The method takes advantage from the fact that, due to the different electrical parameters associated with the intracellular medium (the cytosol), the extracellular medium and the membrane, a well-defined dielectric dispersion appears, generally centered $\sim$1–10 MHz. The analysis of this dispersion, both on the basis of specific dielectric models of a biological cell suspension and on the basis of a phenomenological approach, results in the evaluation of some characteristic parameters, particularly the permittivity and the electrical conductivity of the cell membrane, which describe the structural and functional behavior of the membrane. This method, that, from an experimental point of view requires the measurement of the electrical impedance of the sample under investigation, has been applied to a variety of biological cell suspensions and the parameters associated with the cell membrane have been determined.

See also: Ion Transport in Biological Membranes; Membrane and Protein Aggregates.

PACS: **87.16.Uv; 87.16.Dg; 87.17. – d**

Further Reading

Cevc G and Marsh D (1987) *Phospholipid Bilayers: Physical Principles and Models*. New York: Wiley.

Gaber BP and Easwaran KRK (1992) *Biomembrane structure and function, The State of the Art*. New York: Adenine Press.

Lasic DD (1993) *Liposomes: From Physics to Applications*. Amsterdam: Elsevier.

Lipowsky R and Sackmann E (eds.) (1995) *Structure and Dynamics of Membranes*. Amsterdam: Elsevier.

Biomolecules, Scanning Probe Microscopy of

C Ascoli, CNR-Istituto Processi Chimico Fisici, Pisa, Italy

Introduction

The relevance of microscopy for science, and in particular for biology, has always been clear. In the famous talk of Feymann, given at Caltech for the annual meeting of the American Physical Society in 1959, the requirement of a better microscope for biologists was strongly underlined. The possibility of high-resolution probe microscopy was also clear for the physicists. The use of local probes to illuminate and/or to detect light in a near-field configuration was first proposed in 1928 by Synge, who later suggested the use of piezoelectric actuators. The actual possibility of measuring subatomic movements or distances by piezoelectric actuators was experimentally proved in 1971 by measuring the evanescent light on the surface of a prism placed in the condition of total reflection.

Scanning tunneling microscopy (STM), was first realized in 1982 and, subsequently, many different types of scanning probe microscopy (SPM) techniques followed, which are now used extensively. Clearly, the widespread development of SPMs was made possible by the easy use of personal computers.

The general structure of each SPM is similar: a thin local probe, placed close to the sample (at a distance of the order of the resolution), scans the sample and a quantity that depends on an interaction between the probe and the sample is measured. It is then reported in a plot $f(x,y)$ in a color scale or with a three-dimensional (3D) perspective. Virtually each kind of interaction between the protruding tip of the probe and the sample can originate a

particular SPM, but the key for high resolution is always the steep dependence of the measured quantity on the tip–sample distance. A steep dependence, with space constant of the order of 1 Å or less, as for STM, implies that at short distances the predominant interaction is between the most protruding atoms on the two sides, and then atomic resolution is possible. The most relevant SPM techniques, in particular for the biological systems, are discussed briefly below.

STM is based on the quantum tunnel effect. In quantum mechanics, electrons are described by a wave function. The Schrödinger equation implies that the wave function is not zero beyond a potential barrier even if the electron energy is lower than the barrier itself. Thus, in quantum physics, particles have a nonvanishing probability to be found beyond a potential barrier, in a region forbidden by classical physics: this is known as the tunneling effect. In STM, a sharp conducting tip (the probe) and a conducting sample are separated by an insulator (e.g., a vacuum gap). By the tunneling effect, electrons have some probability of passing from the tip to the conducting sample. When a difference of potential (typically between 1 mV and 1 V) is applied between the tip and the sample, a tunnel current (typically between 0.1 and 10 nA) runs through the insulator for working distances of the order of angstroms. **Figure 1** shows schematically how to get images by measuring the tunneling current: piezoactuators move the sample in the *xy* plane and the tunneling current, measured and stored in a computer, changes according to the underlying surface.

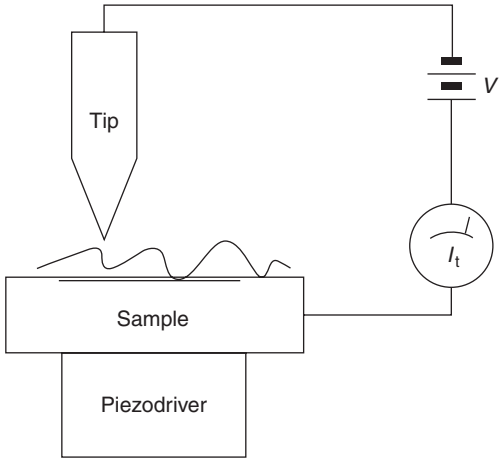


Figure 1 Scheme of STM. The tunneling current I_t flows from the tip to the sample because of the applied bias potential V. The piezodriver moves the sample forward and backward along the x direction; at the end of each line, it makes a step in y. The measured current I_t changes according to the tip–sample distance and gives tunneling current maps of the surface.

Atomic force microscopy (AFM) was developed some years later by Binnig, Quate, and Gerber. AFM is based on mechanical interactions: a small flexible cantilever, with a tip on the sample side, is carried in contact with the sample and makes a regular scanning of the surface. While doing so, the cantilever deflection changes according to the surface profile. The measured deflections (or other equivalent signals) give information on the sample topography and on other local features. An AFM essentially works as old phonographs did, but the difference in scale implies a fundamental difference. For instance, the meaning of "contact," which is clear on a macroscopic scale, is rather undefined on the atomic or nanometer scale. This holds for each SPM, and in order to get a better understanding of what is possible to measure, it is required to consider, in some detail, the interaction between the probe and the sample.

AFM can easily work in water solution, namely, in the normal conditions for biological molecules. Under appropriate conditions, STM can also work in liquids.

Another SPM technique, which necessarily works in ionic solution, is scanning ion conductance microscope (SICM). It was initially proposed by Hansma and co-workers, and it is based on the idea of using, as a scanning probe, a pipette-shaped microelectrode of the type employed by electrophysiologists for intracellular recording (the radius of the tip is of the order of hundreds of nanometers). The current flowing in such a probe is limited when the electrode is in the proximity of a nonconducting (or only partially conducting) object, typically the membrane of a cell. After some years of latency, this kind of microscopy is now applied successfully to living cells, with very interesting results. Two points have been of crucial importance for the development of SICM. One is the use of a small modulation of the tip–sample distance; the other is the coupling with classical electrophysiological experiments. The use of a position modulation allows for a better stabilization of the tip–sample distance (particularly critical at cellular scales). Combining this technique with other microscopies or with electrophysiological recordings, it is possible to observe the time evolution of submicrometric cellular structures and the localization of specific ion channels on the cell surface.

The scanning near-field optical microscopy (SNOM; also called NSOM, PSTM, and by some other names, according to different experimental configurations), the first one conceptually proposed, uses various kinds of proximal probes. By using proximal probes to collect light, or as sources of local lighting, it is possible to overcome the half-wavelength limit of the classical optical microscopy (Abbe or Rayleigh criterion). In fact, the Rayleigh criterion

is not a stringent theoretical limit. It takes into account only the propagating components of the electromagnetic field, which are commonly called "far-field components" and are those involved in traditional optical devices. However, light diffracted, refracted, or reflected by small objects contains more information than its propagating part; for instance, very close to a narrow slit, the light intensity obviously mimics the slit profile. The working distance and the probe used are very important for the resolution. Frequently, tapered optical fibers with a metal coating opened on the tip and suitably modified AFM cantilevers are employed. Moreover, apertureless probes (gold nanoparticles, quantum dots down to single molecules) are also in use now. Spatial resolutions up to a few tenths of nanometers have been reached. SNOM is also used experimentally to obtain nanolithography; its specificity in biological systems is connected with its capability to detect fluorescence of single molecules with spatial resolution.

The AFM Microscopy

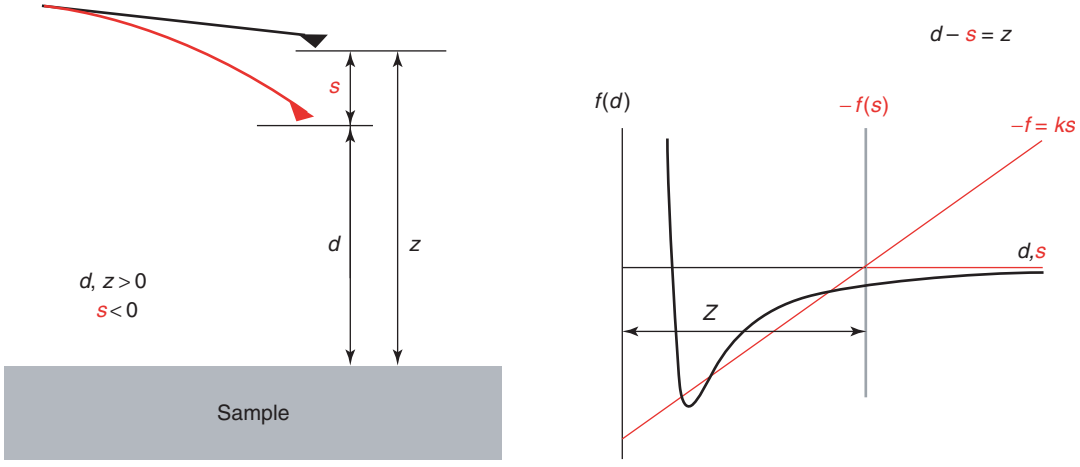
AFM can be considered as an extension of the sense of touch, in all its aspects. Through fingers one is able to perceive the relieves on a surface, to evaluate if it is smooth or rough, compliant or stiff, and how sticky it is. Accordingly, by an AFM it is possible to measure topographic relieves, the local elasticity and plasticity of a body, the friction between the sample and the tip and their adhesion, etc; but all this is now on a nanometer scale and sometimes on an atomic scale.

With regard to the atomic resolution available with AFM, it should be noted that most of the older data show atomic lattices with no well-defined atomic-scale local defects, while STM data clearly show the presence of defects at the atomic level. A partial answer to this issue is that a typical radius of curvature of the tip ranges 5–50 nm, while typical lattice constants are of the order of a few angstroms, so that in the tip there are a number of atoms that interact simultaneously with the sample surface. This number increases when a strong force is applied, because of the deformation of the tip. The force acting on the tip is the composition of the forces due to each single interacting atom. The periodicity of the lattice can be detected, because during the scanning the force due to each interacting atom on the tip has the lattice periodicity. So the tip vertical displacements reproduce the lattice period and not the profile of single atoms. However, it has been proved that, working at distances where repulsive and attractive forces balance, the interaction of the most protruding atom is enhanced. In this case, true atomic resolution has been achieved by AFM, revealing atomic-scale defects.

Forces between Tip and Substrate

Interactions between single atoms obey the Lennard-Jones (LJ) potential. The LJ potential is characterized by two terms: the attractive one, due to the van der Waals forces, acts on a relatively long range, while the repulsive term becomes most important at short distances. For AFM tip–substrate interaction, it is to be taken into account that the number of significantly interacting atoms varies with the distance. Far from the sample, many atoms on the tip and on the surface are in a significantly equivalent situation and contribute equally to the attractive force, while, as the tip approaches the sample, only the most protruding atoms on the two surfaces are relevant in the interaction. This implies that at a short distance, the main contribution comes from the repulsive force between the two nearest atoms on the two sides, while at greater distances the resulting attractive force is due to the integration on many atoms. Therefore, an LJ-like potential, with increased attractive forces, is expected to hold and actually is able to account for the basic features of experimental force–distance curves. **Figures 2** and **3** outline a graphical method to obtain the force–distance curves from an LJ-like potential. Let z be the distance (controlled by the piezodriver) between the resting position of the tip and the sample, d the actual tip–sample distance, and s the deflection of the cantilever. Which deflection and distance are realized for a given z can be determined by assuming that the elastic force of the cantilever (shown in red in **Figure 2**) balances the LJ-like force. Depending on the value of the cantilever elastic constant k, the intersection between $f = -ks$ and the LJ force can have more than one solution, as shown in **Figure 2**, where two stable solutions and an intermediate unstable solution occur. This determines two typical discontinuities in the experimental $F(z)$ force–distance curves, as shown in **Figure 3**. The "jump to contact" occurs during the approach when the attractive force suddenly pushes the cantilever down. The "jump off contact" occurs while retracting the tip from the sample. After the "jump to contact," the tip and the sample move together (for an ideally rigid sample) and this region of FD curves is called contact line. The tip–sample adhesion force is measured by the "jump off contact" height and, as discussed above, can be simply due to the unstable equilibrium between the LJ force and the elasticity of the cantilever. However, adhesion is due also to different chemical or physical reasons. For instance, the adhesion in air is very high for the meniscus formed because of the water adsorbed on the surfaces.

Actually, real force–distance curves give much more information than that predicted from the schematic



z Distance between sample and cantilever resting position

d Tip–sample distance

s Cantilever deflection

Figure 2 (a) Schematic representation of cantilever and sample relative position. The sign of *s* is assumed negative when the cantilever is deflected down (attractive force), so $z = d - s$. (b) Lennard-Jones force (black curve) and cantilever elastic force (red curve with the vertical axis inverted) are reported as functions of *d*. The elastic force, given by $f_{elastic} = -ks = kz - kd$, is zero for *d* equal to *z*, that is controlled by the piezodriver. The cantilever deflection is obtained by the balancing of Lennard-Jones and elastic forces, and is given by the ordinate of the intersection of the two graphs, where $f_{LJ} = -f_{elastic}$. Three possible solutions occur in the graph above. Multiple solutions disappear for the cantilever with *k* greater of the maximum slope of LJ force.

discussion above. By suitably choosing and preparing the cantilever, tip, and sample, the aspecific adhesion does not mask other interactions, and it becomes possible to measure forces between molecular partners (specific adhesion) or unfolding of proteins. In biological systems, the measurement of the force required for unfolding and the force responsible for chemical bonds involved in biological functions is very stimulating, as will be discussed below in some detail.

Nevertheless, AFM was originally considered mainly as a microscopy, namely as a tool to get images. The different ways to get and to handle AFM images are discussed next.

AFM Scheme and Operating Modes

The most-used scheme for detecting the deflection of an AFM cantilever is schematized in **Figure 4**: a Gaussian laser beam is focused on the back of the cantilever and the reflected light is collected by a four-quadrant photodetector, so the displacement of the laser spot gives a signal proportional to the cantilever deflection or torsion, and hence to the interacting force. In particular, with a properly mounted detector, the deflection signal is given by

$$I_D = (I_1 + I_2) - (I_3 + I_4)$$

with $I_{1,2}$ the upper-quadrants signal and $I_{3,4}$ the lower-quadrants signal, while

$$I_T = (I_1 + I_4) - (I_2 + I_3)$$

gives a measure of the torsion of the cantilever, proportional to the friction the tip encounters while scanning the sample.

The signal-to-noise ratio depends basically on the photodetector shot-noise, whose spectral density is given by

$$i_N = \sqrt{2eI\Delta f}$$

where *e* is the charge of the electron, *I* is the generated photocurrent proportional to the laser intensity, Δf is the detection band. The important elements in determining the sensitivity of this detection method are the dimensions of the light spot reflected by the cantilever, which are due to the diffraction induced by the cantilever on the beam.

In analogy with STM and other SPMs, AFM could operate either at constant height or at constant force (**Figure 5**). In the constant-height mode, the piezodriver moves the sample in the *xy* plane maintaining it at a constant *z* and the cantilever deflections are acquired. It is possible to employ this mode only on very flat surfaces, since steep steps can

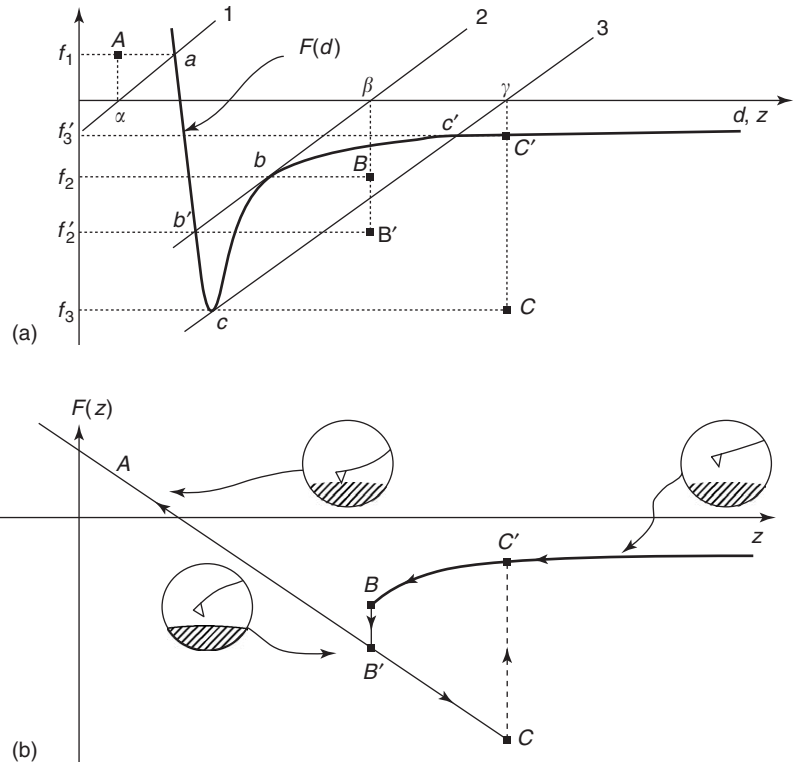


Figure 3 Graphical method to obtain experimental force–distance curves from the Lennard-Jones force F_{LJ}. In (a), lines 1, 2, and 3 represent the elastic force for three different values (α, β, γ) of the height z. The intersection of these lines with F_{LJ} determines the value of the cantilever deflection s. A single solution is possible for line 1 (intersection a), two solutions for lines 2 and 3 (intersection points b and b' and c and c'). Lines between 2 and 3 (not drawn) have three intersections with F_{LJ}. While the tip approaches the sample, the solution moves from c' to b (the corresponding forces can be read on the vertical axis following the dashed lines and range from $f_{3'}$ to f_2); further, approaching the tip and sample, the solution will jump from b to b'. During the withdrawal, a bigger jump will occur from c and c'. (b) Force vs. distance curve during the cantilever approach to and retract from the sample, obtained from the graphs in (a). In the insets, a representation of the cantilever deflection is shown for different positions. (Adapted from Cappella B, Baschieri P, Frediani C, Miccoli P, and Ascoli C (1997) Force-distance curves by AFM. *IEEE Engineering in Medicine and Biology* 58–65.)

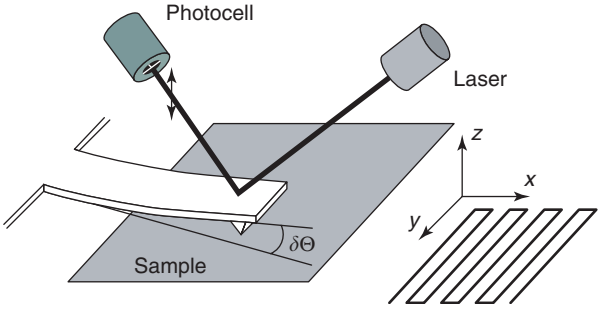


Figure 4 Left – Poggendorf detection method. A laser beam focused on the free end of the cantilever deflection is reflected on a four-quadrant photocell. When the cantilever deflects an angle $\delta\Theta$, the balance between the upper and lower quadrants of the position-sensitive photodetector changes proportionally to $\delta\Theta$. The cantilever deflection can be stabilized by a feedback which moves the sample holder on the z axis. Right – an illustration of the raster scanning of the surface.

cause breaking of the tip or damage of the sample. The use of the constant-force mode is more common. In this mode, the cantilever deflection is maintained constant via a feedback controlling the height z of the piezodriver. This mode decreases the imaging speed because of the limited response time of the feedback loop.

In the constant-force mode, it is usual to acquire the height signal (i.e., the z signal) as well as the so-called error signal. In fact, although the tip deflection is stabilized by the feedback, small deflections of the cantilever occur (they are the error signal). The error signal behaves as the derivative of the height signal along the scanning line and it is useful to enhance the borders of small objects and to highlight periodic patterns, when they are not immediately visible on the topographic images. The error-signal images are like the sample surface lighted by a lateral source.

The height signal is a direct measurement of the sample topography, which, however, depends on the applied force via the local elastoplastic properties of the surface. Generally, AFM images are represented as color-scale 2D images. When represented in

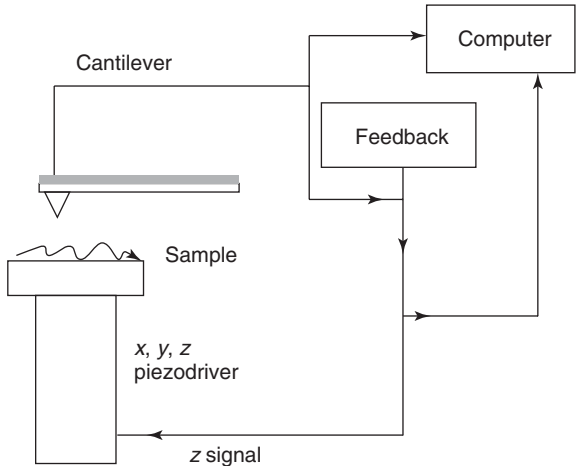


Figure 5 Scheme of an AFM setup. The deflection is used as input for a feedback loop that stabilizes it at a preset value. The output of the feedback controls the voltage that moves the piezodriver in the *z* direction. Both cantilever deflection and feedback output are sampled and stored by the computer.

perspective (3D images), it is easier to appreciate details that are lost in the flat rendering.

During scanning, the cantilever is deflected, but lateral forces, which depend mainly on the tip–sample friction, also twist it. The measurement of the cantilever torsion makes it possible to obtain friction maps, which give an information qualitatively different from the topographic one. Friction maps have been one of the first steps in the direction of obtaining chemical or the so-called spectroscopic information in AFM microscopy. A deep discussion on the tribology at a nanoscopic level is due to Bhushan and others (see the "Further reading" section).

When working in contact, as described above, the friction force can damage soft samples. A possibility of reducing friction forces is achieved in the "tapping" mode. In this mode, the cantilever is maintained in oscillation orthogonal to the sample at a frequency close to its resonant frequency. When approaching the sample, the oscillation, measured via a photodetector, is progressively limited by the sample contact. A feedback loop keeps the oscillation amplitude constant by adjusting the sample height *z*.

Since the tip touches the sample for a very limited time, there is basically no friction and the risk of damaging the tip or sample is greatly reduced compared to the contact mode. Furthermore, beyond the height information obtained from the feedback, in AC mode it is possible to use the phase signal. In fact, the oscillation phase varies during the scan compared to that of the drive signal, and this phase difference is determined by the local elastoplastic properties of the investigated sample. Thus, the phase image contains information on the sample composition.

Sample Preparation

A key step to get good images by AFM is sample immobilization. Graphite and mica and other inorganic materials can be simply glued to the support, but cells or cell fragments or large molecules, such as DNA, require specific preparation of the support to be firmly attached. Glass and mica are the most-used supports. Mica must be freshly cleaved and glass must be cleaned from contaminants, mainly hydrocarbons (e.g., with a detergent or via exposure to UV light). There are many methods for getting the biological samples firmly attached to the support and are based on the knowledge of their interactions with the surface and with the solution where they are suspended. Various ways to firmly anchor the sample to the support are:

- the glass surface may be chemically modified, as in the case of silanized glass, to cause covalent bonding between support and sample;
- coating the support with poly-L-lysine (the support surface becomes positively charged and cell plasma membranes, usually negatively charged, adsorb on it); and
- adjusting the pH and valence of the electrolyte solution can ensure the attachment.

The last of the above methods is possible due to the fact that biological macromolecules are usually charged, and their charge and isoelectric point depends on pH. The same is true for the support surface. Charged surfaces induce counter-ions around them and thus the formation of electrical double layers. A chemical–physical theory of colloidal stability (DLVO theory, developed by Derjaguin–Landau and Verwey–Overbeek) may be applied in order to determine the buffer conditions necessary to obtain a net attractive force as the result of the van der Waals attraction and double-layer repulsion. In conclusion, sample adhesion may be ensured by adjusting the pH and the electrolyte concentration, and by selecting the appropriate electrolyte valence.

A good buffer for sample adsorption is not necessarily good for sample imaging (the force at which the tip and sample are in contact depends on the buffer) and it is very common to perform sample adsorption in a buffer, then rinse it gently and carry out imaging in a different buffer. Once firmly attached to the support, the sample usually stays there even after the buffer has changed.

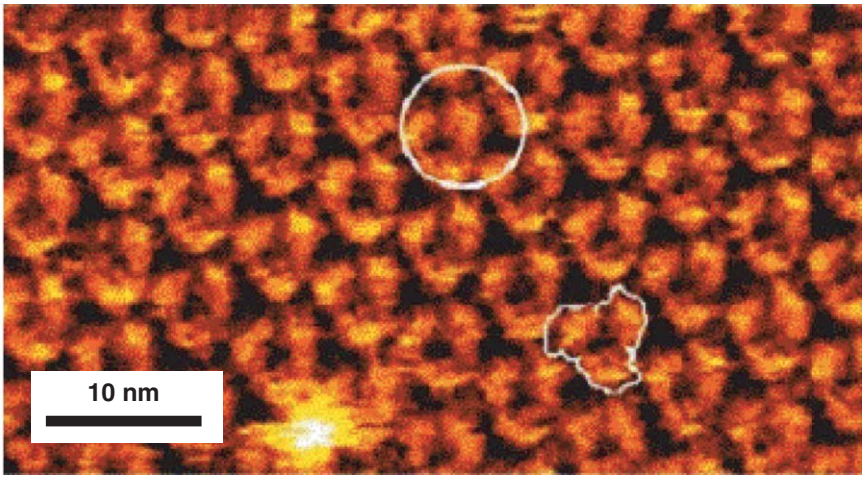
When using FD curves to study specific interactions between molecular partners (antigen–antibody, receptor and its ligand), the preparation of a sample includes suitable functionalization of the sample and tip. Once the stylus has been functionalized with one

of the partners, for instance, via silanization of the surface and subsequent adhesion of proteins, it can be approached to the sample (already functionalized with the other partner) to induce a chemical bond between the two. In this way, protein–protein interaction forces have been studied. Usually, suitable spacers are also interposed between the support and the molecules to be studied, so that the specific detachment on the FD curve occurs far enough from the nonspecific adhesion. Often chemical modifications of the terminal groups in the investigated proteins are required to obtain a suitable preparation.

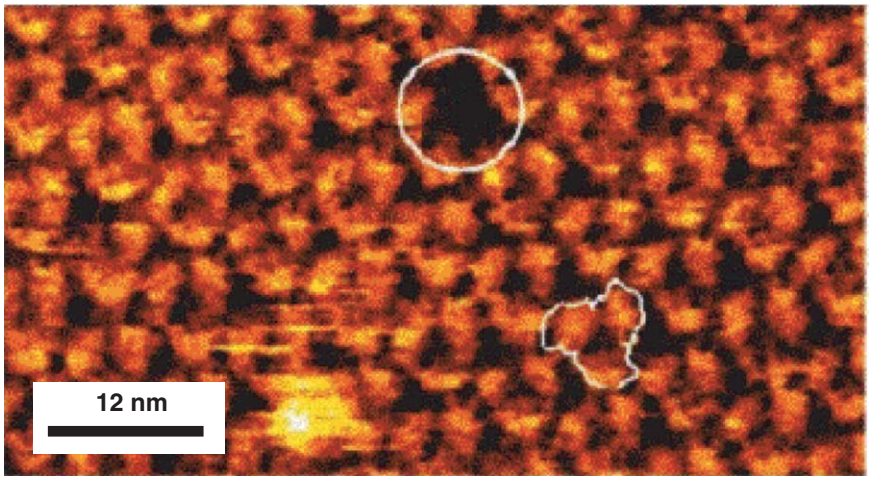
Some Results

It is difficult to give, in a short space, a review of the vast results obtained by SPM in biological systems. Below are reported just a few examples of what is possible to obtain, which are also illustrative of the methods used to extract information from the raw data.

Figure 6 shows a high-resolution image of bacteriorhodopsin (BR) in purple membranes extracted from *Halobacterium salinarum*. BR molecules, in their native membrane, are assembled into trimers, arranged to form a trigonal bidimensional crystal lattice. When the cantilever tip is gently pressed on a molecule, in some cases, the molecule adsorbs on the tip, so once the latter is retracted it extracts a single BR molecule from the membrane. The comparison of the two panels of **Figure 6** shows the empty place of the missing molecule. This figure is illustrative not only of the molecular resolution achieved by AFM on



(a)



(b)

Figure 6 (a) High-resolution AFM images of the cytoplasmic side of purple membranes extracted from *Halobacterium salinarum*. The 2D crystal lattice of bacteriorhodopsin molecules and the organization in trimers are clearly visible. A trimer is contoured for clarity. The length of the white bar is 10 nm. The same area has been imaged twice, before and after stretching the molecules that were adsorbed on the tip. (b) The missing molecule was extracted by the tip during the withdrawal. (Adapted from Oesterhelt F, Oesterhelt D, Pfeiffer M, Engel A, Gaub HE, *et al.* (2000) Unfolding pathways of individual bacteriorhodopsins. *Science* 288: 143–146.)

soft samples but also of its possible use as nanomanipulator. This is one of the directions in which AFM development is focused.

Figure 7 reports a beautiful example of imaging conformational changes at molecular resolution. On the right-hand side of the figure, the scanning in the *y*-direction was stopped and the same molecules were sampled in subsequent scans, showing the dynamical association and dissociation of GroEL–GroES complexes.

More information on the formation and the breaking of molecular bonds or the folding and unfolding

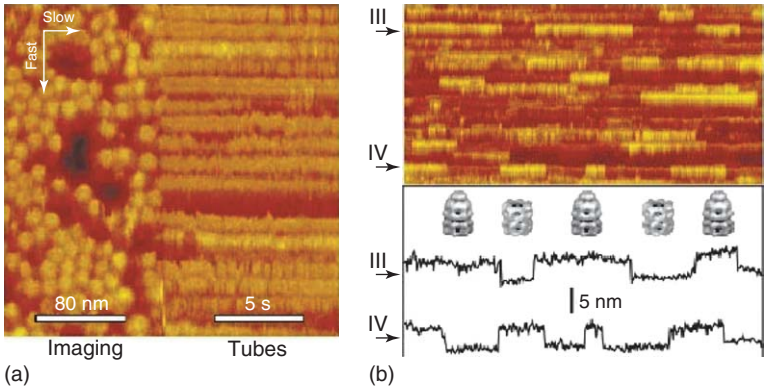


Figure 7 (a) Topography of GroEL proteins adsorbed on mica. The *x* axis, in this case, is in the vertical direction (indicated by the arrow of the fast scan direction). At half the way, the scanning along the *y* direction (the horizontal one in this case, indicated by the arrow of the slow scan direction) was disabled and the scan was repeated on the same line of proteins, generating the protein "tubes" of the right-hand part of the image. The tubes are the evolution over time of the height of the proteins along that line, monitored with a temporal resolution of 100 ms. (b) GroES (144 nM) and ATP (2.5 nM) are added to the buffer solution and the tubes are imaged. Under these conditions, the association and dissociation of GroEL–GroES complexes is possible, in fact, large repeated variations in height along the length of many tubes are observed. The height profile of the tubes indicated with III and IV are reported below. They show a stepping of the height between two levels that differ for 3.6 ± 1 nm. This height difference is consistent with that of GroEL and GroEl–GroES complex seen by X-ray crystallography, confirming that the observed phenomenon is the association and dissociation process. (Adapted from Viani MB, Pietrasanta LI, Thompson JB, Chand A, Gebeshuber IC, *et al.* (2000) Probing protein–protein interactions in real time. *Nature Structural Biology* 7(8): 644–647.)

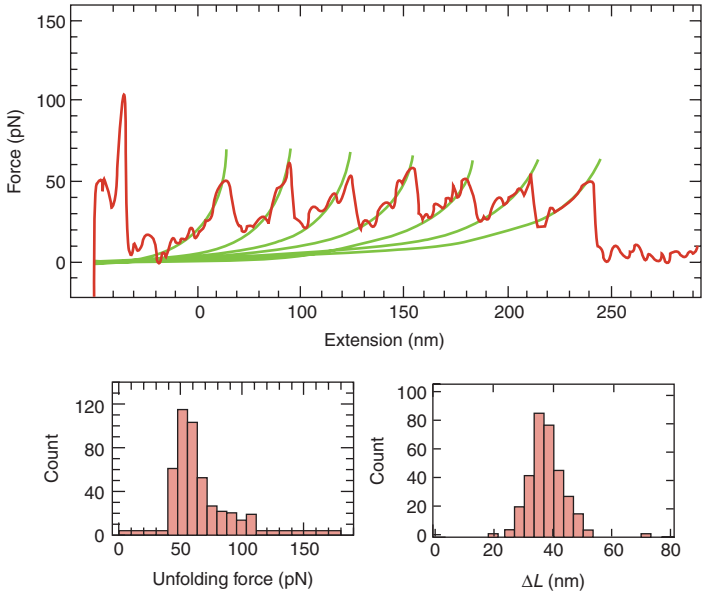


Figure 8 Unfolding of T4 lysozyme in a polymer. Successive unfolding of different units are recorded. The fitting curves are obtained by the WLC model. The two histograms show the distribution of the forces at which unfolding occur and the variation in length obtained by the fitting. (Adapted from Yang G, Cecconi C, Baase WA, Vetter IR, Breyer WA, *et al.* (2000) Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme. *Proceedings of the National Academy of Sciences of the USA* 97: 139–144.)

of polymers can be obtained by AFM by using the force–distance curves. An alternative technique, very suitable for studying force–distance curves, is that of optical tweezers. **Figure 8** is an example of studying by AFM, the unfolding process in a globular polymer. The figure shows the results as force–extension curves; in this kind of representation, the measured deflection is reported as a function of the actual tip–sample distance given by $d = z - s$ (see **Figure 2**). The dependence of the measured force on the elongation maps the potential of the folding–unfolding process and can be fitted by appropriate models, which come from theoretical studies of polymer dynamics. The reference book for polymer modeling appeared in 1969 and is due to Flory, who won the Nobel prize for chemistry in 1975. Models like freely jointed chain (FJC) or worm-like chain (WLC), and others can be used to fit force–extension curves, and this allows one to determine parameters as the contour length of the polymer or its permanence length.

The other kind of information that can be obtained from results such as those of **Figure 8** is the amplitude of the jumps due to the sudden occurrence of unfolding, that is, the estimated value of the force at which unfolding occurs. These values are somewhat variable, as shown from their histogram reported in the insets of **Figure 8**. This is not a mere effect of the noise that affects the measurements, but is partially due to the stochastic nature of the unfolding process. This is better visible in **Figure 9**, which was obtained by stretching the biotin–streptavidin bond; in this case, the measurements were made at several different velocities of the z-scanning, and the histograms of the measured strengths of the bond depend on this velocity. The peak of the histogram shifts on the right, and the histogram width increases with the velocity on the z-axis.

To understand why the measured strength depends on the rate of force variation, it is important to consider that the lifetime of a bond due to weak noncovalent interactions depends, in the most trivial case, on a potential barrier that separates the two conformations; the height of the barrier is the difference ΔG of the free energy between the bound state

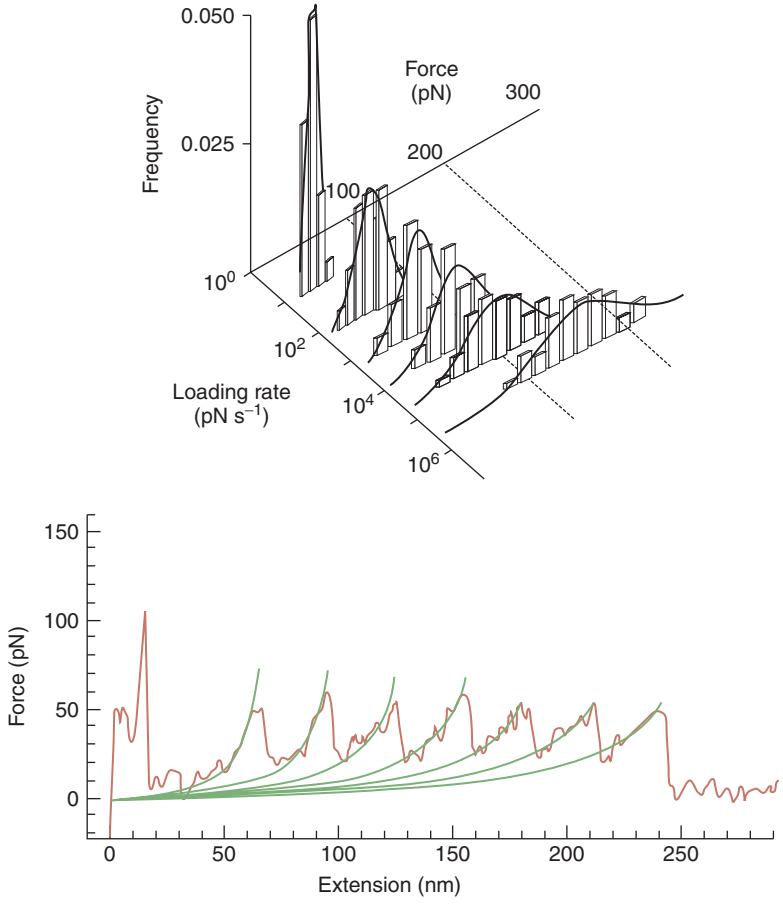


Figure 9 Histogram of biotin–strepavidin bond strength. The histogram peak shifts to right and the histogram width increases with increase in loading rate. (Adapted from Merkel R, Nassoy P, Leung A, Ritchie K, and Evans E (1999) Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature* 397: 50–53.)

and the transition state. Spontaneous breaking of the bond at temperature T occurs with a probability that depends exponentially on the ratio between ΔG and the thermal energy $k_B T$. Under stretching, the potential profile is modified, the energy landscape is tilted by force so that the height of the barrier decreases and the transition becomes more likely. Complex macromolecular bonds involve many interactions that create a series of barriers. However, for each barrier the probability of breaking in a short interval dt is an exponential function of the applied force, which varies as the stretching proceeds. At a given value, f, of the stretching force, the probability of breaking in a short interval dt is constant; say, $p(f)dt$. The process is Poissonian and the probability $P(t)$ to wait for a time t before the breaking depends exponentially on t:

$$P(t) = \frac{1}{p(f)}\, e^{-p(f)t}$$

Most of the breakings occur in time intervals, shorter than $3/p(f)$. If the stretching force varies quickly in this timescale, breaking will occur at higher values of the force. The process being an inhomogeneous Poisson process, $P(t)$ is given by

$$P(t) = \frac{\exp\left[-\int_0^t p(f(t))\, dt\right]}{\exp\left[-\int_0^\infty p(f(t))\, dt\right]}$$

Experiments are possible where the vertical sweep stops at a preset value of the deflection, keeping the force constant until the conformation change. If the process under study is reversible, these experiments allow one to determine the histogram of the waiting times and then $p(f)$. It is clear from the discussion above that the interpretation of results is much more direct in experiments made at constant stretching force, while the results obtained by retracting the tip without a control on the applied force do not allow a straightforward evaluation of $p(f)$. Some experiments at constant force have been reported recently, and it is to be expected that stretching at constant force will be a rule in the future.

See also: Protein Folding and Aggregation; Quantum Mechanics: Foundations; Scanning Near-Field Optical Microscopy; Scanning Probe Microscopy; Single-Molecule Methods in Biophysics.

PACS: 07.79. – v; 68.37. – d; 82.37.Gk

Further Reading

Bhushan B, Israelachvili JN, and Landman U (1995) Nanotribology: friction, wear and lubrification at the atomic scale. *Nature* 374: 607–616.

Broglia RA, Shakhnovich EI, and Tiana G (eds.) (2001) *Protein Folding, Evolution and Design*, Volume 145 International School of Physics Enrico Fermi. Amsterdam: IOS Press.

Bushan (ed.) (1998) *CRC Handbook of Micro/Nanotribology*. Boca Raton, FL: CRC Press.

Derjaguin BV and Landau LD (1941) Theory of the stability of strongly charged lyophobic sols and the adhesion of strongly charged particles in solutions of electrolytes. *Acta Physicochimica. URSS* 14: 633–662.

Flory PJ (1969) *Statistical Mechanics of Chain Molecules*. New York: Wiley. (also reprinted by Hanser Publishers, Müchen, 1989).

Israelachvili J (1992) *Intermolecular & Surface Forces*, 2nd edn. London: Academic Press.

Lennard-Jones JE and Dent BM (1926) The forces between atoms and ions. *Proceedings of the Royal Society (London)* 112A: 230–234.

Ohnesorge F and Binnig G (1993) True atomic-resolution by atomic force microscopy through repulsive and attractive forces. *Science* 260: 1451–1456.

van der Ziel (1955). *Noise*. London: Chapman and Hall.

Verwey EJW and Overbeek JTG (1948) *Theory of Stability of Lyophobic Colloids*. Amsterdam: Elsevier.

Bose–Einstein Condensation

M Inguscio, Università di Firenze and Istituto Nazionale, Florence, Italy
S Stringari, Università di Trento and Istituto Nazionale, Trento, Italy

Quantum Mechanics Reigns at Low Temperature

At room temperature, the dynamic behavior of a gas is governed by the laws of classical mechanics. In fact, the thermal de Broglie wavelength

$$\lambda_T = \sqrt{\frac{2\pi\hbar^2}{m k_B T}} \qquad [1]$$

which gives the smearing of the position of the atoms, due to the Heisenberg uncertainty principle ($\hbar$ is the Planck constant divided by 2π, k_B is the Boltzmann constant, m is the atomic mass, and T is the temperature of the gas), is much smaller than the average spacing between atoms, which then behave

as classical objects. As the gas is cooled, however, the smearing increases and the wave functions of adjacent atoms overlap, causing the atoms to lose their identity. If the overlap is large, the effects of quantum mechanics cannot be ignored.

When quantum effects become important, it is crucial to distinguish between Bose and Fermi statistics. Bosons are particles with even total spin and their many-body wave function is symmetric with respect to the exchange of two particles. As a consequence, bosons like to occupy the same state. Conversely, fermions (particles with odd spin) are described by an anti-symmetric wave function and cannot occupy the same state due to the Pauli exclusion principle. At very low temperatures, bosons and fermions behave quite differently (see **Figure 1**), giving rise to distinct dynamic and thermodynamic behaviors.

Bosons are known to undergo a phase transition below a critical temperature T_c. This transition is characterized by the macroscopic occupation of a single particle state and is called Bose–Einstein condensation (BEC). A unique peculiarity of the transition is that it can occur even in the absence of interactions, being driven by genuine quantum statistical effects.

A Einstein predicted the occurrence of this transition in 1925, on the basis of a paper of the Indian physicist S N Bose (1924), devoted to the statistical description of the quanta of light. For a long time, Einstein's predictions had no practical impact and only after the discovery of superfluidity in liquid helium (1938), the phenomenon of BEC became the object of theoretical investigation again with the pioneering works by London, Bogoliubov, Landau, Lifshitz, Penrose, Onsager, and Feynman.

At low enough temperatures, all interacting systems, with the exception of helium, undergo a phase transition to the solid phase. This behavior is illustrated in **Figure 2**, where a typical pressure–temperature phase diagram is shown. In the figure, the pressure–temperature line characterizing the BEC phase transition of an ideal gas is also drawn. Above this line, a dilute gas would be Bose–Einstein condensed. However, this configuration is unstable since thermodynamic equilibrium would correspond to the crystal phase for such values of pressure and temperature. This shows that BEC can be actually achieved only in conditions of metastability, and that the density of the gas should be small enough to suppress the collisional processes (three-body collisions) responsible for bringing the system into the thermodynamically stable solid phase. In practice, typical densities reached in the BEC phase are 10^{13}–$10^{15}\,\mathrm{cm}^{-3}$, so that the predicted values of T_c are extremely low (a few microkelvin or even less). This explains why BEC was experimentally realized in atomic gases only 70 years after Einstein's historical paper.

The realization of BEC in atomic gases was achieved for the first time in 1995 and was recognized with the Nobel prize in 2001. This achievement is the result of extraordinary efforts made in atomic physics during the 1990s, through the development of advanced techniques of cooling and trapping atoms, which were recognized with the award of the Nobel prize in 1997. At present BEC has been achieved in

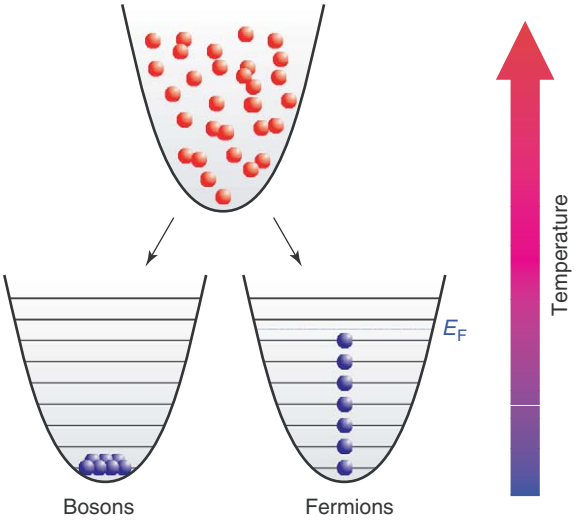


Figure 1 Schematics of the level occupancy for a harmonically trapped gas at ultralow temperature. The ground state of the many-body system is completely determined by the quantum statistics. The bosons (even spin particles) all occupy the same single-particle ground state, forming a Bose–Einstein condensate. The fermions (odd spin particles), obeying the Pauli exclusion principle, pile up until they reach the Fermi energy E_F.

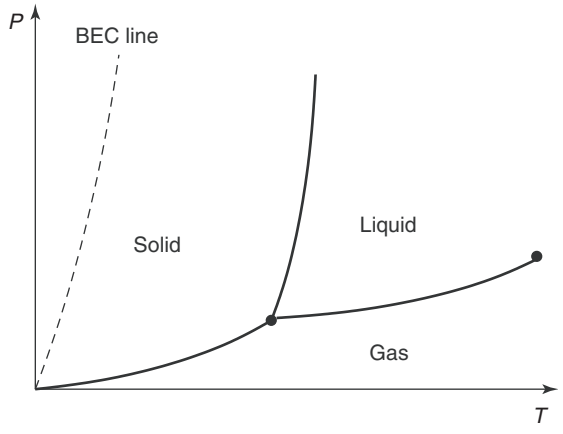


Figure 2 A typical pressure–temperature phase diagram. The dashed line corresponds to the BEC phase transition for an ideal gas.

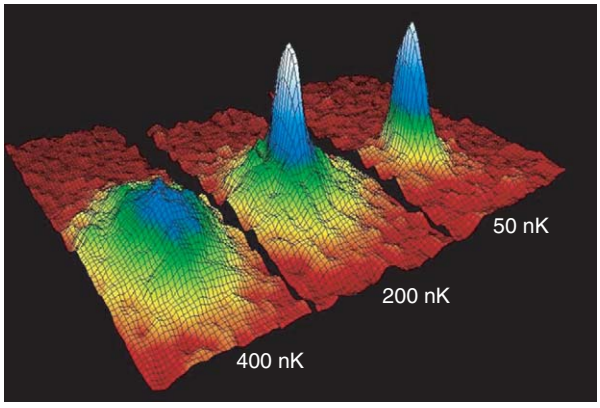


Figure 3 Density distribution of an ultracold sample of ^{87}Rb across the BEC transition recorded during one of the first experiments at JILA (1995) in which BEC has been reached.

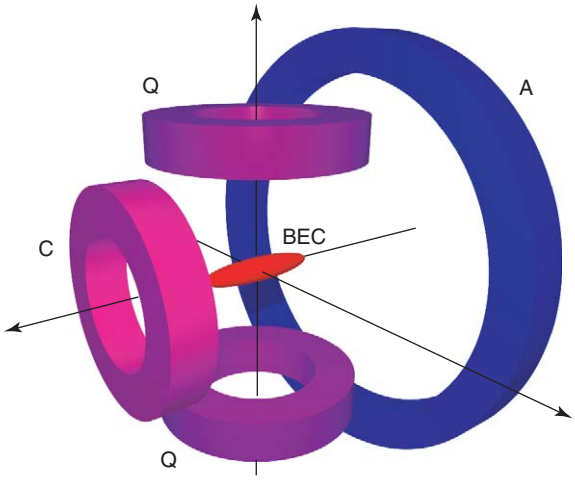


Figure 4 A four-coil magnetostatic trap for neutral atoms, as the one used in the experiments at LENS (Florence). The magnetic field has a minimum in the center of the trap and is, in good approximation, harmonic around this point (Ioffe–Pritchard configuration). The atomic cloud trapped in the minimum of the potential is cigar-shaped because of the cylindrical symmetry of the trap.

^{87}Rb and ^{85}Rb, ^{23}Na, ^{7}Li, H, metastable ^{4}He, ^{41}K, ^{133}Cs, ^{174}Yb, and ^{52}Cr (**Figure 3**).

How to Reach BEC in Dilute Atomic Gases

The first experimental studies on BEC were focused on spin-polarized hydrogen that was considered the most natural candidate because of its light mass. For this purpose, cryogenic and evaporative cooling techniques were developed since the early 1970s. However, BEC was first realized in alkali atoms despite their relatively high masses, because of their suitability to be cooled with laser techniques. In a typical setup, the strategy to achieve BEC of alkali gases consists of the following steps: (1) laser cooling of the atoms, (2) magnetic trapping, and (3) evaporative cooling.

The first cooling stage is performed in a magneto-optical trap (MOT) that consists of the combination of counterpropagating laser beams and a magnetic field gradient. The physical mechanism at the basis of the cooling process is the transfer of momentum between photons and atoms. The laser beams are red-detuned with respect to an atomic transition, in such a way that the resulting effect is a viscous force slowing the motion of atoms. In addition, the presence of the magnetic field gradient allows one to spatially confine the atoms in a region of a few cm^3. In the case of rubidium, one can typically trap $\sim 10^9$ atoms with a temperature of $\sim 10\,\mu$K (inferred from the velocity distribution of the atoms).

The density and temperature achievable in a MOT are far from the values required for BEC (the fundamental limit for laser cooling is the recoil temperature, i.e., the kinetic energy gained by an atom after absorption of one photon, $\sim 400\,$nK for rubidium). The next stage consists in transferring the atoms from the MOT to a magnetostatic trap, where higher densities and lower temperatures can be achieved. An example of a typical coil configuration is shown in **Figure 4**. The current flowing in the coils produces a magnetic field that has a minimum in the center of the trap, and is, in good approximation, harmonic around this point. Due to their magnetic moment, atoms in a low-field seeking Zeeman state are trapped in the center of the trap.

The final cooling stage is the so-called evaporative cooling that is performed by selectively removing the hottest atoms from the magnetic trap. RF transitions are used to excite the atoms located at the border of the trap (the ones with the highest energy) to untrapped states. This process is successful only if the energy of the RF cut is lowered slowly enough in such a way that the remaining atoms have time to collide and thermalize to lower temperatures. At the end of this process, if the density of the sample is high enough, the phase transition to a Bose–Einstein condensate can take place (typical values for rubidium are $n \sim 10^{14}\,\text{cm}^{-3}$ and $T \sim 100\,$nK). BEC, in a trap, occurs both in momentum and coordinate space, with the atomic density distribution reflecting the shape of the trapping potential (cigar-shaped for the trap in **Figure 4**).

All these stages are performed in a ultra-high-vacuum environment (background pressure $\sim 10^{-12}\,$torr), in order to ensure a perfect thermal isolation and a long lifetime of the trapped atomic samples against collisions with the background gas.

Imaging the New Macroscopic Quantum State

Images of the ultracold atomic sample are achieved by shining a nearly resonant laser beam. The absorption of light by the atoms creates a shadow that is recorded by a CCD camera. The typical size of a condensate is of the order of a few micrometers and a very high resolving power is hence necessary for *in situ* imaging. More frequently, images are taken by switching the trap off and allowing the gas to expand to larger sizes. Typical images obtained after expansion are shown in **Figure 5**. From these images one extracts relevant information on the physical properties of the system. At temperatures higher than T_c, the sample expands ballistically and quickly reaches an isotropic distribution, well described by a Maxwellian. From this distribution one extracts the temperature of the gas. At the onset of the BEC transition, the shape of the cloud is characterized by a pronounced increase of the density in the center. At lower temperatures, the density takes a characteristic bimodal shape (**Figure 5**): the noncondensed atoms determine the wing of the distribution, still given by a Maxwellian form, while the condensed atoms give rise to a narrow central peak which becomes more and more pronounced as the temperature is lowered. The shape of the central peak is no longer a Maxwellian, but is given by an inverted parabola as predicted by the Gross–Pitaevskii theory, and exhibits a typical nonisotropic expansion. The diagnostic of the images also provides the number N of atoms in the sample.

In the presence of harmonic confinement, the theory predicts both the value of the critical temperature T_c and the fraction of condensed atoms for $T \leqslant T_c$. For an ideal gas trapped by a harmonic potential, one finds

$$k_B T_c = 0.94 \hbar \omega_0 N^{1/3} \qquad [2]$$

and

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^3 \qquad [3]$$

where ω_0 is the geometric average of the trapping frequencies and N_0 is the number of atoms in the condensate. A typical theory–experiment comparison is shown in **Figure 6**, where the value of T_c is a few hundreds of nanokelvin. Despite the relatively small number of atoms used in this experiment ($\sim 10^4$), the evidence for the phase transition is very clear. The good agreement with the predictions of the ideal gas model indicates that the effects of the interactions on the condensate fraction as well as on the value of T_c are small. This is the consequence of the extreme diluteness of the trapped gas.

In the following sections, some of the major achievements obtained in these years on these ultracold gases will be discussed. These include the study of their coherence and superfluid properties, as well as the novel features exhibited in the presence of a

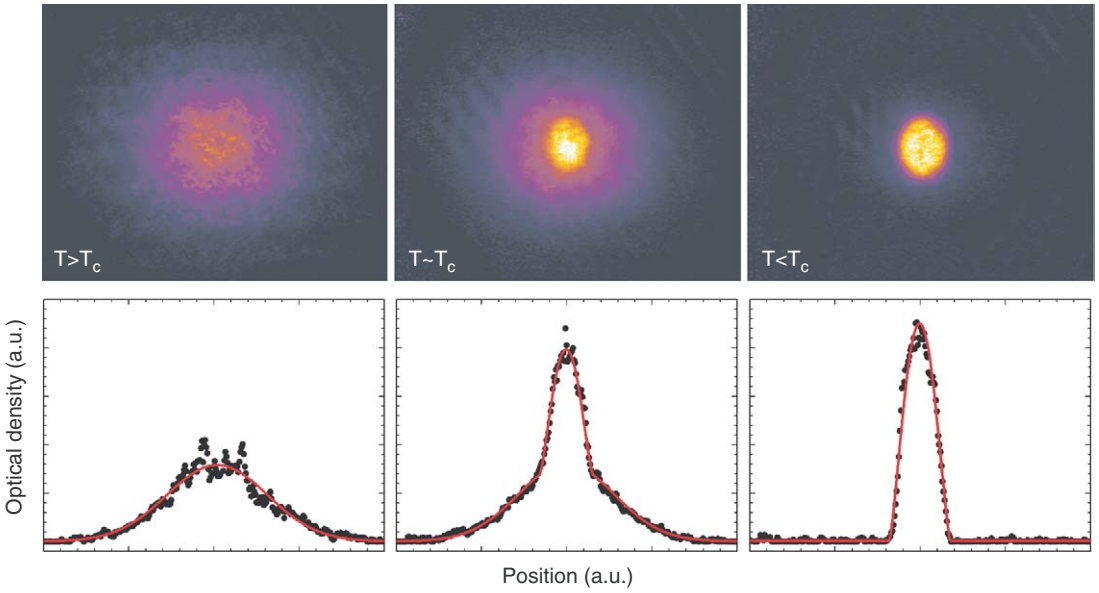


Figure 5 (Top) Absorption images of an expanded atomic cloud varying the temperature T across the BEC transition at T_c. From left to right, the images show a thermal cloud, a partially condensed cloud, and a pure BEC. (Bottom) Horizontal cross section of the density distribution. The lines are fits of the experimental points with a Gaussian, a bimodal distribution and an inverted parabola, respectively.

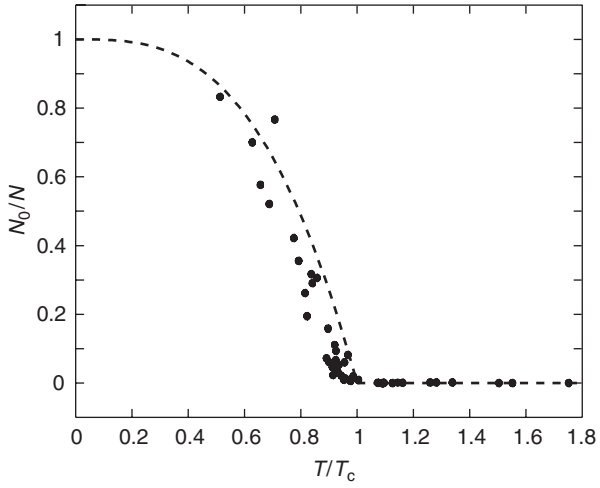


Figure 6 Condensate fraction N_0/N as a function of T/T_c. Circles are experimental results from JILA (1996), while the dashed line is the theoretical prediction of eqns [2] and [3].

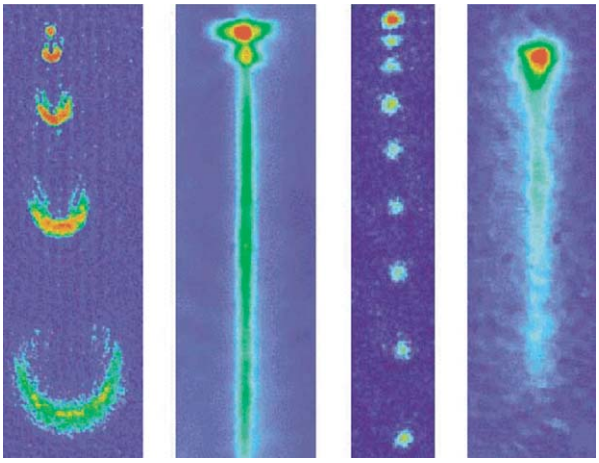


Figure 7 A collection of atom lasers, obtained outcoupling coherent atoms from a Bose–Einstein condensate with different methods. The first two lasers have been achieved using pulsed or cw RFs in order to induce transitions toward untrapped states. The third one has been obtained taking advantage of the tunneling induced by gravity in a vertical lattice. The fourth laser has been obtained using optical Raman transitions.

periodic potential. A short discussion on recent results with ultracold Fermi gases is also presented.

Coherence and Superfluidity

Bose–Einstein condensates are characterized by a complex order parameter (the so-called wave function of the condensate)

$$\Phi = \sqrt{n_0}\mathrm{e}^{\mathrm{i}S} \qquad [4]$$

The modulus is fixed by the condensate density n_0 which, in dilute gases and at temperatures much smaller than T_c, in practice, coincides with the density of the gas. The phase S of the order parameter is at the origin of important properties of the system, both concerning coherence and superfluid effects.

Coherence phenomena make the physics of these systems similar to the one of a laser, if one replaces photons with atoms. Bright sources of atoms, the so-called atom lasers, can be obtained outcoupling the atoms from the condensate (see **Figure 7**) that is the analog of the laser cavity in which coherent photons are stored. This has been accomplished with different methods (RF transitions, Raman pulses, gravity-induced tunneling) that have allowed one to create cw and pulsed sources of coherent atoms.

Another example of this analogy is revealed by the interference phenomena observed with Bose–Einstein condensates. **Figure 8** shows the interference fringes produced by two separate condensates overlapping after expansion. This experiment, which first demonstrated the coherence of the condensed state, is the analog of the most famous double-slit experiment with light.

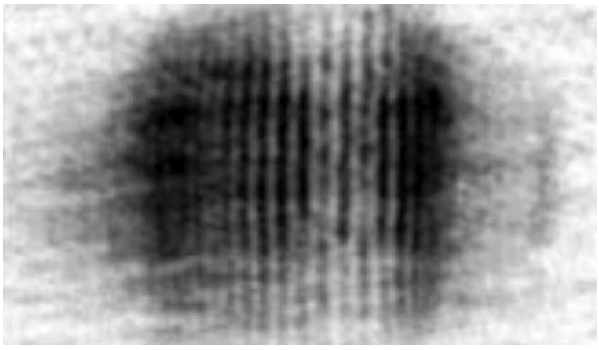


Figure 8 Density profile of two condensates overlapping after expansion. In the central region one clearly sees interference fringes, similar to the ones observed in optics in the famous Young experiment. This image has been recorded in the first experiment in which interference of matter waves has been observed (MIT 1996).

Differently from photons, atoms are interacting particles. In the mean field approach, interactions give rise to important nonlinear phenomena. A beautiful analogy with nonlinear optics is provided by the experiments on four wave mixing (see **Figure 9**), in which three matter waves, interacting through nonlinearities, can produce a fourth matter wave. This can only be achieved if phase-matching conditions (energy and momentum conservation) are satisfied. While in the case of light, nonlinearities are a property of the medium in which the mixing process takes place, in the case of atoms nonlinearities are an intrinsic property of the mixing waves and are due to

the elastic s-wave collisions between the atoms forming the matter wave. Interactions in Bose–Einstein condensed gases are in most cases repulsive. Profiting from the occurrence of resonant states between two interacting atoms, it is now possible to tune the value and even the sign of the scattering length by simply applying an external magnetic field. When the scattering length becomes negative, the nonlinearities in the Gross–Pitaevskii equation are responsible for solitonic solutions (bright solitons) which propagate without dispersion in 1D configurations. The existence of bright solitons has been confirmed in a recent series of experiments.

Another spectacular consequence of the phase of the order parameter is related to superfluidity. In fact, the gradient of the phase S is proportional to the superfluid velocity:

$$\boldsymbol{v}_s = \frac{\hbar}{m}\nabla S \qquad [5]$$

Before 1995, superfluid phenomena had been observed only in dense liquids, such as ^{4}He. The possibility of extending such investigations to dilute gases has opened new challenging perspectives for a better understanding of the microscopic origin of superfluidity and its relationship to the phenomenon of BEC. Among the various manifestations of superfluidity, it is worth mentioning the rich variety of collective oscillations exhibited by these confined systems, which have been the object of intense theoretical and experimental work in the last few years. **Figure 10** shows the shape oscillation of a Bose–Einstein condensate. From these images one can extract the frequency of the collective oscillation with high precision. The dynamic behavior of the condensate is well described by the hydrodynamic equations of superfluids, which predict values for collective frequencies in excellent agreement with the experiment.

An even more spectacular prediction of superfluidity concerns the rotational properties of Bose–Einstein condensates. Superfluids cannot rotate like classical objects because of the irrotationality constraint imposed by eqn [5]. A striking consequence is that angular momentum can be carried only by quantized vortices. If the confining trap rotates slowly, vortices cannot be formed because they are energetically unfavorable and the sample does not carry any angular momentum. However, when the angular velocity increases, quantized vortices are formed (see **Figure 11**). Eventually, at high angular velocities, one can generate a vortex lattice of regular geometrical shape.

Another important manifestation of superfluidity is given by the Josephson oscillations characterized by the coherent tunneling of atoms through the barriers generated by an external potential (see the next section).

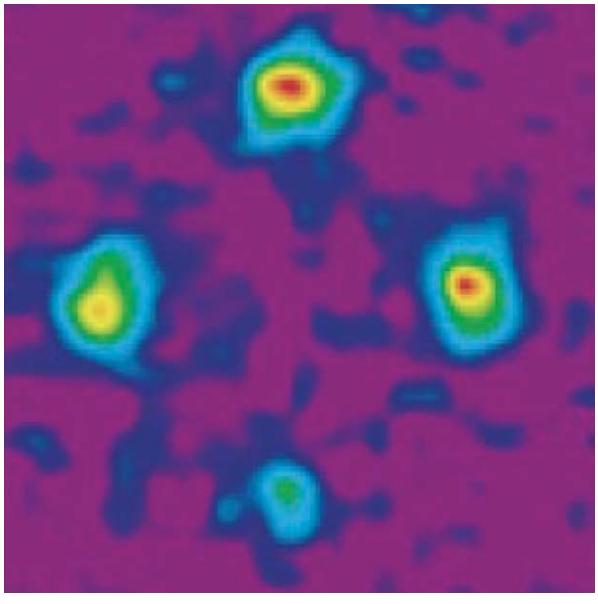


Figure 9 Four wave mixing with matter waves (NIST 1998). In this absorption image, taken after a time of flight, each peak represents a matter wave with different momentum. The three most intense peaks (up, left, right) are created by splitting a Bose–Einstein condensate with laser light. The fourth smaller peak (down) is the result of the mixing process caused by the nonlinear interaction between the three initial waves.

BEC in Periodic Potentials

Ultracold atoms can also be trapped with nonresonant laser light taking advantage of the interaction with the induced dipole moment. Differently

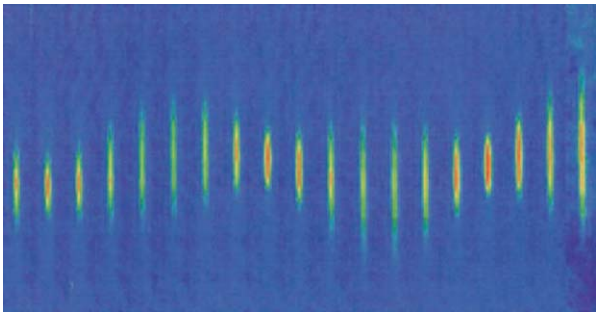


Figure 10 Collective excitations of a Bose–Einstein condensate (MIT, 1998). The excitations were produced by modulating the magnetic field used to trap the condensate and then letting the cloud evolve freely. The field of view in the vertical direction is about 600 μm and the time step is 5 ms per frame. The measured frequency of these collective excitations is a clear signature of the BEC superfluidity.

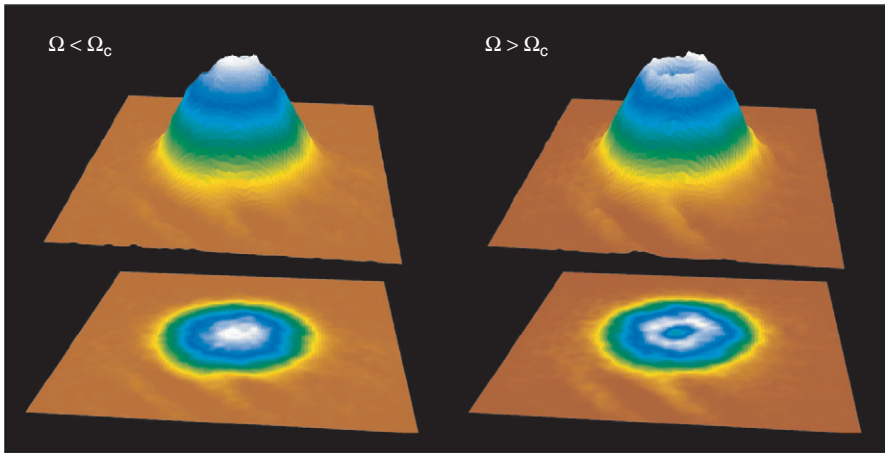


Figure 11 Formation of vortices in a Bose–Einstein condensate (ENS, Paris, 1999). The images refer to a BEC confined in a trap rotating at angular velocity Ω. Above a critical velocity Ω_c, a vortex forms in the condensate, as can be seen by the reduction of density in correspondence of the center.

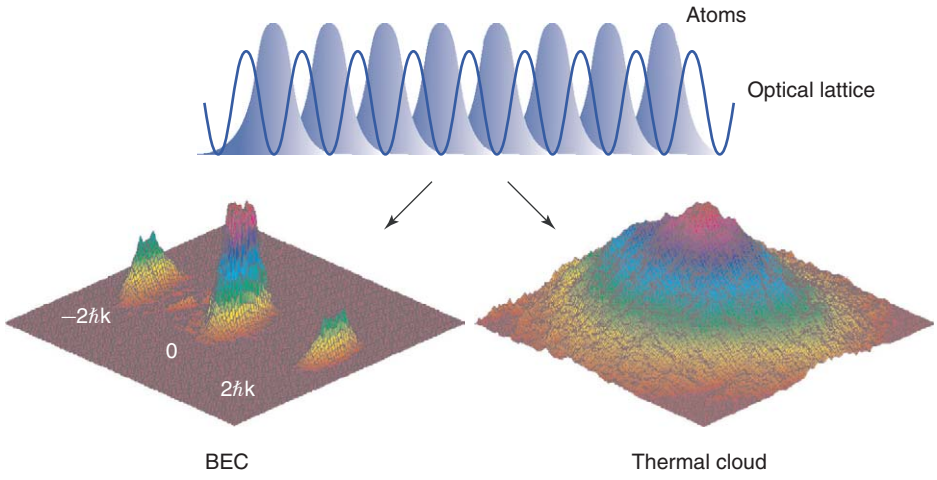


Figure 12 Expansion of a Bose–Einstein condensate from an optical lattice (LENS, Florence, 2001). The atoms are trapped in a periodic potential generated with laser light and then released from this trap. The figure on the left, showing the density distribution after expansion, exhibits resolved peaks originated from the interference of the condensates initially located at the lattice sites. The observed structure, signature of the long-range coherence of the system, cannot be observed with a thermal incoherent sample (figure on the right). This effect is analogous to the diffraction of light by a matter grating.

from magnetic trapping, this allows one to trap atoms of any spin state and to produce new geometrical configurations. In particular, working with counterpropagating laser beams, it is possible to generate periodic potentials, the so-called optical lattices. These optical configurations give rise to the possibility of producing new crystals where the ordered atomic structure is ensured by the laser light.

Arrays of condensates can be produced with an optical lattice high enough to trap the atoms in the lattice sites (see **Figure 12**). The density distribution of the atoms released from this optical trap shows well-resolved peaks originated by the quantum interference of the condensates initially located at the lattice sites and overlapping during the expansion. This feature, which is a signature of the long-range coherence of the system, cannot be observed in a thermal sample, which instead presents a broader unstructured distribution. This is analogous to the diffraction of light from a matter grating, where the role of matter and light is interchanged.

A condensate in a periodic potential is also an ideal tool to study the key phenomena of solid-state physics related to the Bloch description. In particular, it is possible to point out directly the typical band structure and related transport properties in the ideal situation of a perfect crystal, avoiding the presence of phonons or defects, typical of solid-state configurations (**Figure 13**).

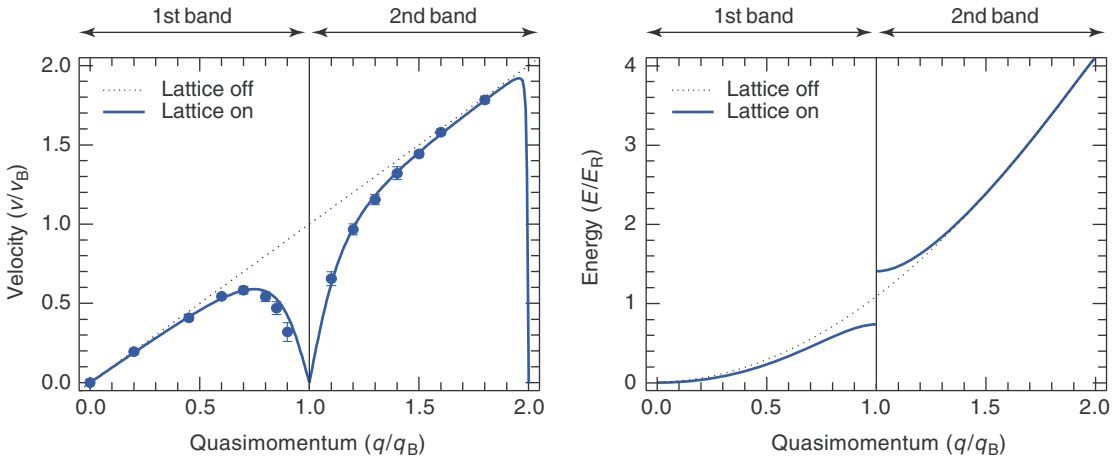


Figure 13 A BEC in an optical lattice is the ideal system for studying problem of solid-state physics. By directly observing the propagation of the matter wave packet in a periodic potential, one can measure the energy band spectrum.

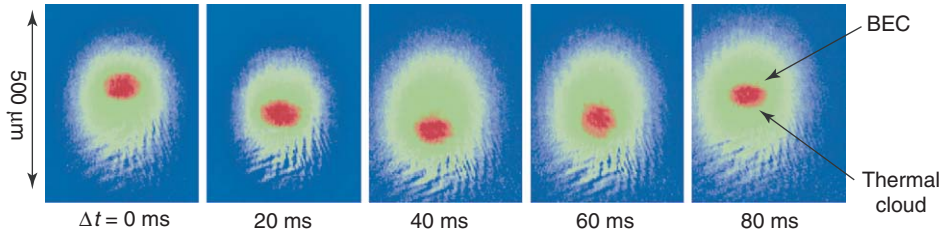


Figure 14 A partially condensed cloud is oscillating in the harmonic trap in the presence of an optical lattice (along the vertical direction, in figure). While the thermal cloud center of mass is stuck in the initial position, the condensed part moves through the periodic potential because of the coherent tunneling between the lattice sites, analogous to the Josephson effect (LENS, Florence, 2001).

An interesting possibility provided by these periodic potentials is the study of Josephson-like oscillations in the superfluid phase. These oscillations are characterized by a coherent tunneling of the gas through the optical barriers. Experimentally, one can easily excite center-of-mass oscillations of the gas confined in the combined potential created with a harmonic trap and an optical lattice. **Figure 14** shows that only the condensed component is able to tunnel coherently, while the thermal component is localized. Theoretically, one can predict the frequency of the oscillation, which turns out to be in good agreement with the experiment.

A spectacular achievement with these new configurations is the possibility of studying the transition from a superfluid to a Mott insulator phase (**Figure 15**). By increasing the intensity of the laser (and hence the height of the barriers generated by the 3D optical lattice), one reduces the tunneling rate between neighboring sites, thereby favoring the localization effects. At some point, the coherence of the sample is lost and the system exhibits an insulating phase. **Figure 15** shows the experimental evidence for this transition, which is proven by the disappearance of the interference peaks in the density profile imaged after expansion.

Fermions

Differently from the case of bosons, an ideal Fermi gas does not exhibit a phase transition at low temperature. The system, however, is still characterized by important quantum phenomena originating from the Pauli exclusion principle. This is well illustrated in **Figure 1**, where one sees that at very low temperatures the gas exhibits single occupancy of the particle states up to a maximum energy, called Fermi energy.

The experimental procedure to reach ultracold regimes in Fermi gases cannot be the same as in the case of bosons. Indeed, in a Fermi gas of a single spin species, collisions are suppressed at low temperature by the antisymmetrization requirement. This causes the absence of thermalization processes which are crucial for the mechanism of evaporative cooling to be efficient. This difficulty has been overcome by working with two spin species, which can interact also at low temperature. Alternatively, one can use

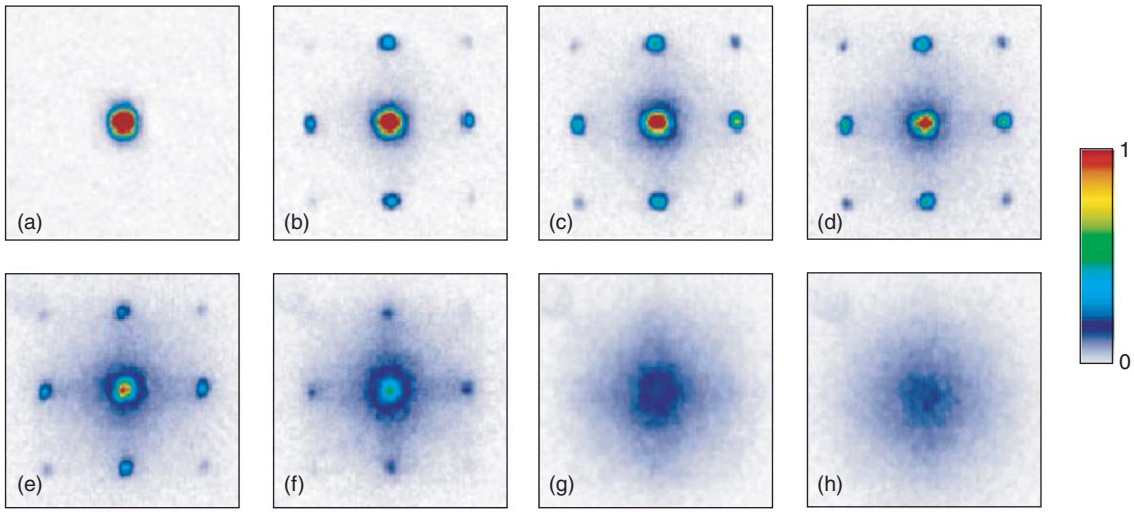


Figure 15 BECs in 3D optical lattices have allowed the observation of the quantum phase transition from a superfluid to a Mott insulator. In the latter phase (achieved for the first time in Munich in 2001) a number Fock states are created at the lattice sites when the tunneling between the optical barriers is lower than a critical value. Consequently, phase fluctuations increase and the long-range coherence typical of the superfluid phase is lost. This is evidenced by the disappearance of the interference pattern in expansion when the lattice height is increased (from a to h).

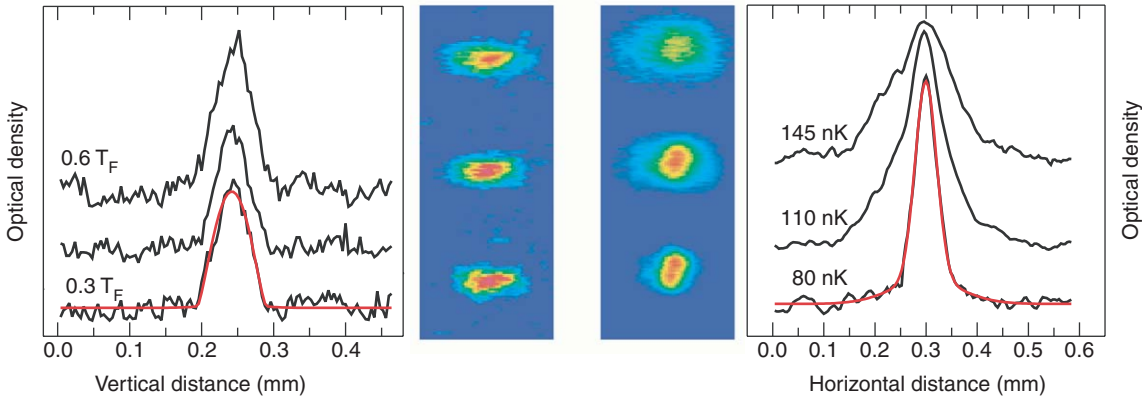


Figure 16 Sympathetic cooling of fermionic ^{40}K with bosonic ^{87}Rb (LENS, Florence). The bosons are directly cooled with standard evaporative cooling techniques and the fermions thermalize sympathetically at lower temperatures via collisions with the bosons. The figure shows the simultaneous onset of Fermi degeneracy for ^{40}K (left), and of Bose–Einstein condensation for ^{87}Rb (right). The absorption images are taken for decreasing temperatures (from top to bottom) and the graphs show the cross sections of the density distribution after expansion.

mixtures of fermions and bosons. Bosons are cooled down with traditional techniques and fermions then thermalize sympathetically. **Figure 16** shows this second type of approach, achieved with potassium and rubidium.

The absence of interactions of polarized fermions in a gas, on the one hand, complicates the cooling procedure and, on the other hand, can provide important advantages. In fact, noninteracting degenerate Fermi gases exhibit peculiar features which have been the object of recent investigations. An example is given by the occurrence of undamped Bloch oscillations in the presence of periodic potentials, which are very promising for future precision measurements (see **Figure 17**). Also, the achievement of efficient registers for quantum computation might be strongly favored by the use of ultracold noninteracting Fermi gases.

On the other side, the possibility of trapping fermions in different spin states is opening new perspectives related to the physics of interacting systems. In particular, recent experimental efforts have been aimed at realizing superfluid configurations. For dilute gases interacting with attractive scattering

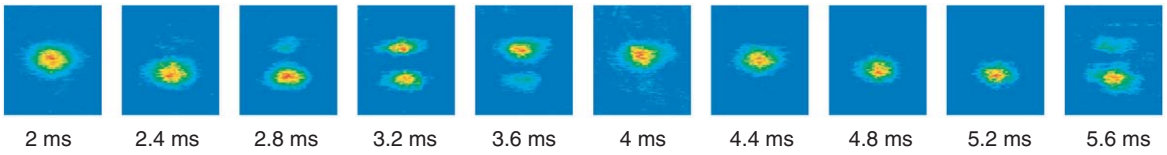


Figure 17 Bloch oscillations of ultracold fermions trapped in a vertical optical lattice, as observed in a recent experiment at LENS, Florence (2004). The images show the momentum distribution of the fermionic cloud moving under the action of gravity. Measuring the frequency of the oscillation allows an accurate determination of the local gravity acceleration g, and can be, in principle, used for high precision measurements of forces on a short distance scale.

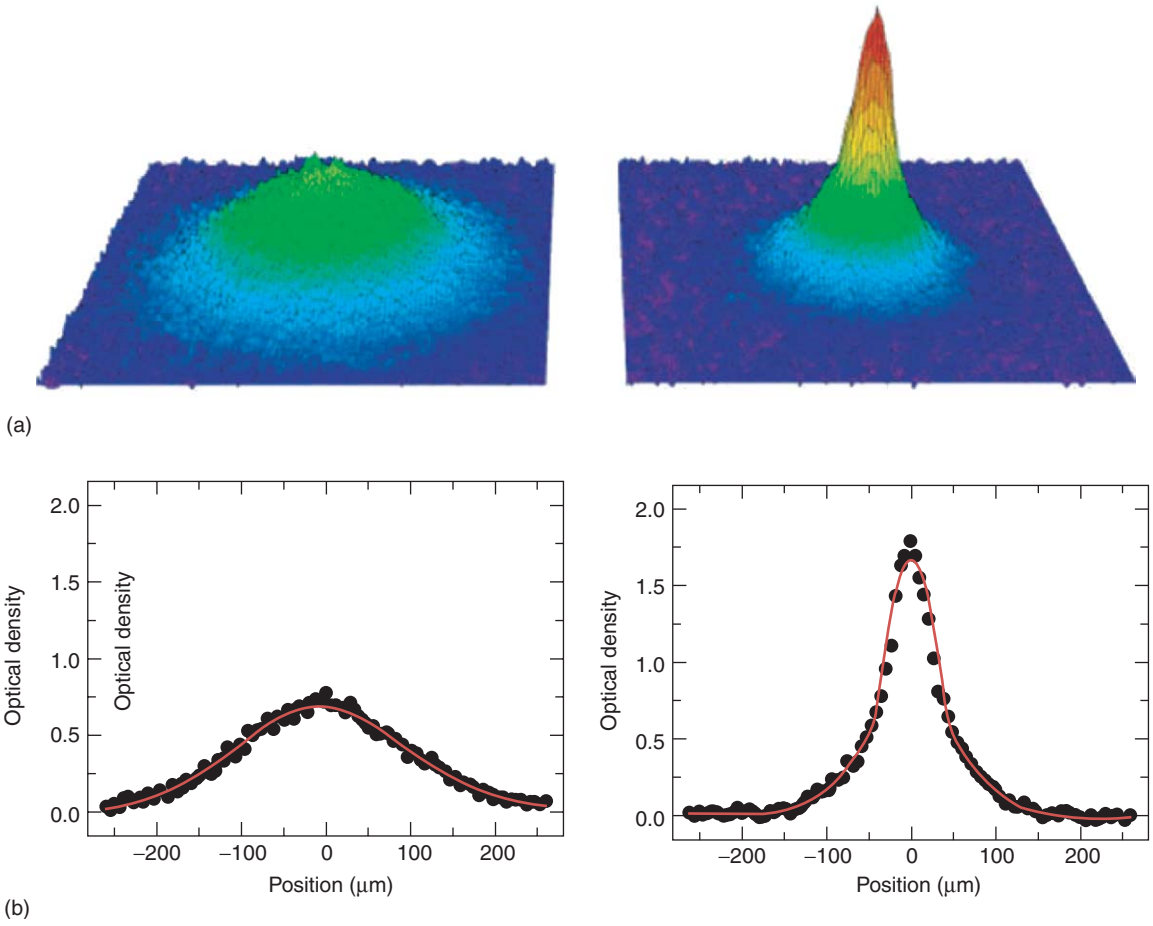


Figure 18 BEC of $^{40}K_2$ molecules (JILA, 2003). The (bosonic) dimers are produced starting with quantum degenerate ^{40}K fermions in two spin states and sweeping the value of an applied magnetic field across a resonance. (a) Absorption images of the molecular density distribution after expansion. When the initial temperature of the fermionic cloud is lowered across a critical value (from left to right), a narrow peak appears in the momentum distribution. (b) Cross section and fit with a bimodal distribution. The change in the density profile is a signature that a BEC of molecules has formed.

length, the superfluid phase is described by the traditional BCS theory, first developed to describe the phenomenon of superconductivity. If the scattering length of fermions in two spin states becomes large or positive, new scenarios take place. Such configurations can now be achieved experimentally by tuning the scattering length in the vicinity of a resonance. In particular, this technique has proved quite successful in the efficient production of ultracold molecules in a gas of atoms interacting with positive scattering length. Since these molecules have a bosonic nature, being composed of an even number of fermions, they can undergo a phase transition to BEC, similar to the one exhibited by a gas of bosonic atoms.

Molecular BEC has been already realized starting from Li and K fermionic atoms (see **Figure 18**). At present, significant efforts are being made to study the BEC–BCS crossover, which occurs by varying continuously the scattering length through the application of an external magnetic field. In particular at resonance, where the scattering length becomes much larger than the average distance between atoms, the system exhibits new challenging features which cannot be simply described using the theory of weakly interacting gases and requires the use of more advanced many-body approaches.

See also: Phase Transformation; Phases and Phase Equilibrium; Quantum Mechanics: Atoms; Quantum Mechanics: Critical Phenomena; Statistical Mechanics: Quantum.

PACS: 03.75.Fi; 03.75.Kk; 05.30.Jp

Further Reading

Cornell EA and Wieman CE (2002) Nobel Lecture: Bose–Einstein condensation in a dilute gas, the first 70 years and some recent experiments. *Reviews of Modern Physics* 74: 875.

Dalfovo F, Giorgini S, Pitaevskii LP, and Stringari S (1999) Theory of Bose–Einstein condensation in trapped gases. *Reviews of Modern Physics* 71: 463.

Inguscio M, Stringari S, and Wieman CE (eds.) (1999). Bose–Einstein condensation in atomic gases. In: *Proceedings of the International School of Physics "Enrico Fermi," Varenna (Italy) – Course CXL*. IOS Press.

Ketterle W (2002) Nobel Lecture: when atoms behave as waves: Bose–Einstein condensation and the atom laser. *Reviews of Modern Physics* 74: 1131.

Martelluci S, Chester AN, Aspect A, and Inguscio M (eds.) (2000) Bose–Einstein condensates and atom lasers. In: *Proceedings of the International School of Quantum Electronics, Erice (Italy) – 27th Course*. Kluwer Academic/Plenum.

Pethick CJ and Smith H (2002) *Bose–Einstein Condensation in Dilute Gases*. Cambridge: Cambridge University Press.

Pitaevskii LP and Stringari S (2003) *Bose–Einstein Condensation*. Oxford: Oxford University Press.

Bravais Lattice *See* Periodicity and Lattices.

C

Carbon Materials, Electronic States of

P Lambin, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium
J Fink, Leibnitz Institüt für Festkörper- und Werkstof-Forschung Dresden, Dresden, Germany

Introduction

Carbon is a remarkable element that shows a variety of stable structures, ranging from three-dimensional (3D) diamond to 2D graphite and 1D carbyne. The different forms of carbon are the consequence of two important characteristics of its electronic $1s^2 2s^2 2p^2$ configuration. First, carbon may form different hybridizations, called sp^n, between its $2s$ orbital and n of its $2p$ orbitals ($1 \leqslant n \leqslant 3$). Next, the C–C distance is small enough to allow for a large π bonding between those $2p$ orbitals of neighboring atoms that are not hybridized with the $2s$ orbital (when $n < 3$). These two properties together give many structural flexibilities, thanks to which the chemistry of carbon is so rich. Already with pure carbon, different structures can be realized, which have very different electronic properties. These structures derive from the three allotropic crystalline forms of carbon listed in **Table 1**.

Hybrid Orbitals

When two carbon atoms are brought close together to form a C_2 molecule, the electrons move in an effective potential that has cylindrical symmetry around the line joining the two nuclei. A molecular wave function may be totally invariant under a rotation by an angle φ about this axis, thus forming a σ state, or it may transform like $\cos(m\varphi)$, thus forming a π state ($m = 1$), a δ state ($m = 2$), etc. The ground state of the C_2 molecule is a σ state. Its expression can be written as a linear combination of atomic orbitals (LCAO) restricted to the $2s$ and $2p_z$ valence orbitals, which have full rotational symmetry about the molecular axis chosen as the z direction. In addition, the wave function of the ground state is even with respect to the inversion about the molecule center. This means $\psi_{\sigma_g} = (\Phi_{A+} + \Phi_{B-})/\sqrt{2}$, where $\Phi_{A+} = \alpha\phi_{A_s} + \beta\phi_{A_z}$ and $\Phi_{B-} = \alpha\phi_{B_s} - \beta\phi_{B_z}$. In that expression, ϕ_s and ϕ_z are the $2s$ and $2p_z$ orbitals located on one C atom, either at location A or B. Φ_{A+} and Φ_{B-} are linear combinations of these orbitals, called sp^1 hybrids (see **Figure 1a**). They are symmetric to each other with respect to the inversion center. The numerical coefficients α and β are determined by solving a secular LCAO equation.

In a C_2 molecule, two electrons occupy the σ_g ground state (see **Figure 1b**) and form a strong σ bond. Two electrons occupy the next σ state, which is a nonbonding state with wave function $\psi_{\sigma_u} = (\Phi_{A-} - \Phi_{B+})/\sqrt{2}$. It is a combination of sp^1 hybrids on A and B sites that point opposite to the other atom. These hybrid orbitals are ready to overlap with an H atom each, such as in the acetylene C_2H_2 molecule, or with other C atoms such as in carbyne, which is a 1D carbon structure (pure carbyne should be fully sp^1-hybridized). The last four electrons of the C_2 molecule occupy the doubly degenerate π_u bonding state (see **Figure 1b**) obtained by combining

Table 1 Structural and electronic properties of the carbon allotropes at room temperature

	Diamond	C_{60} *fullerite*	*Graphite*
σ-Bond hybridization	sp^3	$sp^{2+\varepsilon}$	sp^2
Structure	f.c.c.	f.c.c.	hex
Lattice parameters (nm)	$a = 0.357$	$a = 1.420$	$a = 0.246$, $c = 0.671$
Density (g/cm^3)	3.52	1.67	2.27
C–C distance (nm)	0.154	0.140 – 0.145	0.142
Bandgap (eV)	5.48	2.2	− 0.03
Type	Large-gap semiconductor	Semiconductor	Semimetal
Binding energy (eV/atom)	7.35	6.96	7.37

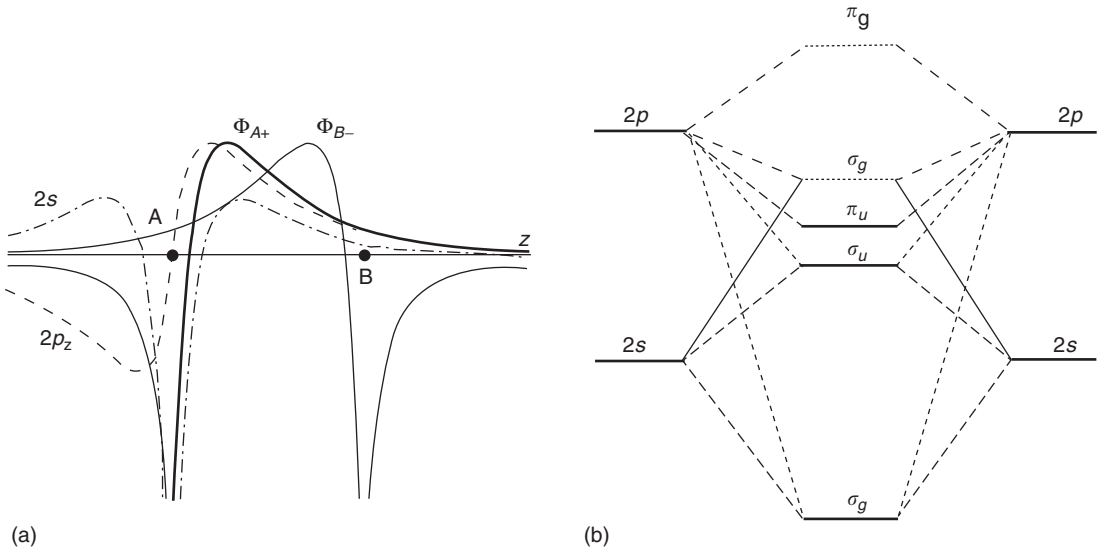


Figure 1 (a) Variation of the 2*s* (dot-dashed line) and $2p_z$ (dashed line) atomic orbitals on site A along the axis *z* of the C_2 molecule. The solid line curves represent the Φ_{A+} (thick line) and Φ_{B-} (thin line) sp^1 hybrid orbitals, (b) Formation of the electronic levels of C_2 from the 2*s* and 2*p* levels of the atoms. The highest σ_g and the π_g levels are unoccupied.

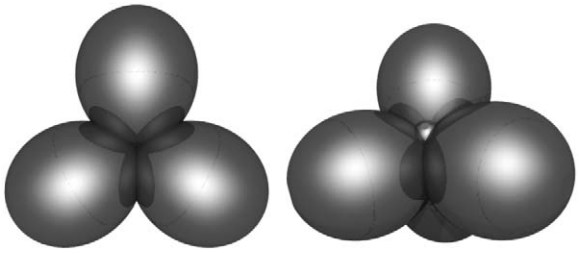


Figure 2 Carbon sp^2 (left) and sp^3 (right) hybrid orbitals. Each orbital is represented by its angular probability distribution for an electron at 0.15 nm from the C nucleus.

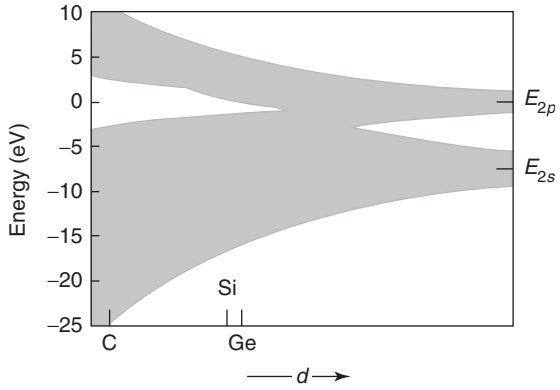


Figure 3 Schematic illustration of the band formation in diamond structure as a function of the interatomic distance (tight-binding description). The positions of Ge, Si, and C are indicated.

the $2p_x$ orbitals and the $2p_y$ orbitals. These electrons form two π bonds.

Other hybridizations of the 2*s* and 2*p* valence orbitals may be formed. The most symmetric combinations correspond to the sp^2 and sp^3 hybrids illustrated in **Figure 2**. In the first case, each carbon atom forms three sp^2 hybrids, by mixing the 2*s* orbital with $2p_x$ and $2p_y$ orbitals. In graphene and graphite, the sp^2 hybrid orbitals have maxima of probability density along three directions at 120° in the *xy* plane. They form σ bonds with neighboring atoms. sp^3 hybrids are combinations of all valence orbitals. In diamond, each atom forms four such hybrids along four directions at 109° in space with tetrahedral symmetry.

Diamond

When carbon atoms are brought together to form a solid, their 2*s* and 2*p* levels broaden and form energy bands. The formation of these bands is illustrated in **Figure 3** for the case of diamond. At large interatomic distance, the bands are narrow and conserve most of the 2*s* and 2*p* characters of the individual atoms. When the distance decreases, the width of the bands increases and their *s* and *p* characters start to mix. At some distance, there is a crossing between the bottom of the band that originates from the 2*p* level, and the top of the band that originates from the 2*s* level. From now on, one may consider that the electronic states result from the σ overlapping of sp^3 hybrid orbitals. By reducing the interatomic distance further, a gap opens between σ bands that have acquired bonding and antibonding characters, respectively. There are four hybrid orbitals per atom,

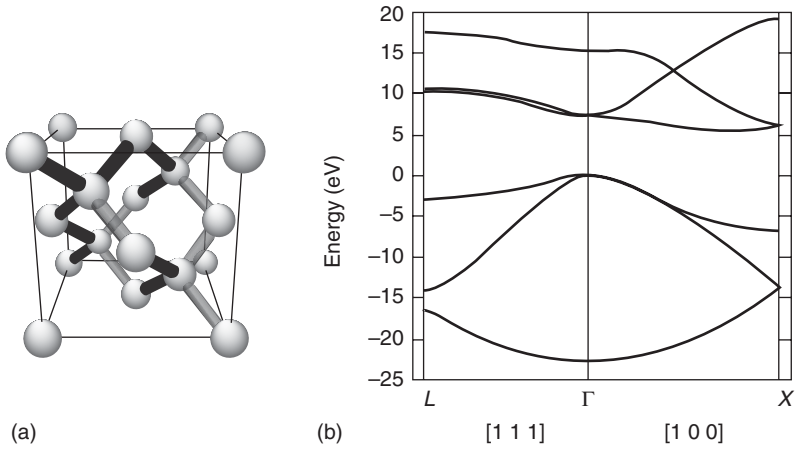


Figure 4 (a) Cubic diamond crystal structure, (b) Band structure of diamond along high-symmetry lines calculated with the all-electron GW approximation. (Courtesy of M Alouani.)

giving rise to two states in each of the bonding and antibonding bands. Each band may therefore accommodate four electrons, due to spin degeneracy. Since there are four electrons per atom, the bonding band is full, the antibonding band is empty, and the system is a semiconductor. The so-called hybridization gap that separates the bonding and antibonding σ bands increases with decreasing distance. Interestingly enough, this picture remains qualitatively correct for the tetrahedral *sp* semiconductors that adopt the same diamond structure. From the decreasing interatomic distance in going from Ge to Si to C, one may understand the increase of their band gap. The picture also remains qualitatively correct for tetrahedral amorphous carbon.

For a crystalline material, the most complete description of the electronic states is provided by its band structure. A theoretical band structure of diamond is shown in **Figure 4b** for Bloch vectors parallel to a principal diagonal of the cubic cell (ΓL line) and parallel to an edge of the cubic cell (ΓX line). The valence band extends from -22.6 to 0 eV. The width of the hybridization gap at the Γ point is 7.4 eV. There are four branches in the valence band, the two atoms per unit cell of diamond (**Figure 4a**) bringing two bonding σ states each. The symmetry of the wave function along the ΓL and ΓX lines is high enough to allow for a twofold degeneracy of the top valence branch and the bottom conduction branch. The computed minimum band gap is 5.6 eV. It is an indirect gap, corresponding to a transition between the Γ point and the minimum of the bottom conduction branch along ΓX. The gap is so large that diamond does not absorb any visible radiation. Pure diamond is transparent, as a consequence. Diamond can be doped with boron and other elements, and can reach relatively high *p*-type or *n*-type conductivity. The σ bonds in diamond are very strong. As a result, diamond is the hardest material known. Its Young modulus for uniaxial strain applied along a fourfold symmetry axis is 1.05 GPa.

Graphite

Graphite can exist in different phases according to the stacking of its layers, each of which is a single carbon sheet with a honeycomb structure called graphene. The most common crystalline form of graphite, discovered by J D Bernal in 1924, has the ABABAB… stacking illustrated in **Figure 5a**. The σ states of graphite are formed by overlapping sp^2 hybrid orbitals oriented along the C–C bonds in each layer. Qualitatively, the picture of **Figure 3** still applies to these σ states. In Bernal graphite, there are four atoms per unit cell (two per layer), each of them builds three σ states, of which half have a bonding character. In the band structure of **Figure 5c**, the bonding σ states contribute $1.5 \times 4 = 6$ branches between -20.1 and -3.4 eV. Since there is very little σ interaction between two successive atomic planes of graphite, the σ branches are nearly twofold degenerate. The antibonding σ^* states lie above $+3.7$ eV, the hybridization gap is 7.1 eV. Unlike the case of diamond, this gap is bridged by the π states that originate from the $2p_z$ orbitals, perpendicular to the atomic layers. These orbitals form two bonding bands and two antibonding bands. They appear as two pairs of branches split by the interlayer coupling, which is much more efficient with the out-of-plane $2p_z$ orbitals than with the in-plane sp^2 hybrids. The bonding π and antibonding π^* branches have an appreciable dispersion, ~ 8 eV, which means a high degree of delocalization of the electrons originating from the $2p_z$ orbitals. The π and π^* branches cross

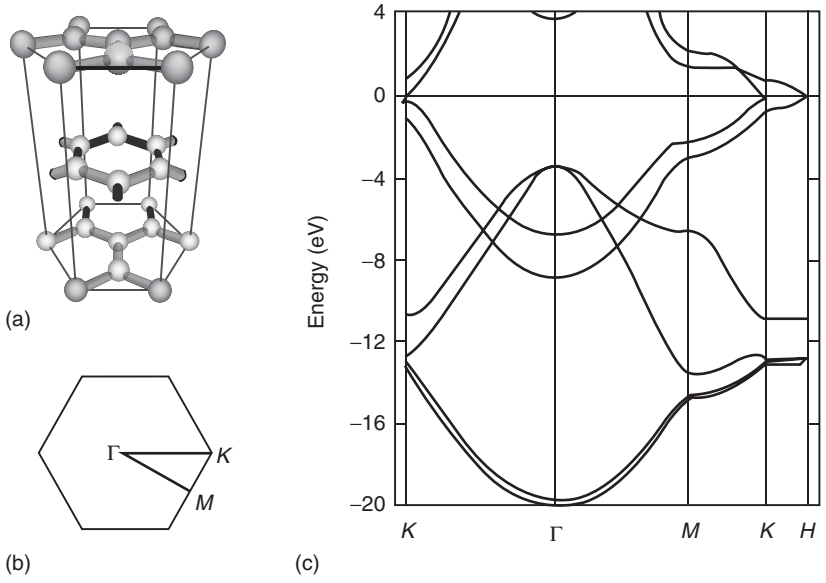


Figure 5 (a) Graphite crystal structure, (b) high-symmetry points in the basal plane of the first Brillouin zone of graphite, and (c) band structure of graphite calculated with the local-density approximation and the pseudopotential approach. (Courtesy of J C Charlier.)

each other for Bloch vectors very close to the *KH* line, which corresponds to an edge of the hexagonal prism defining the first Brillouin zone of graphite. There is a small overlap of $\sim 30\,\text{meV}$ between the π and π^* bands, and the Fermi level is located in this region. This means that graphite is a semimetal. It is a poor electrical conductor (room temperature resistivity of crystalline samples $\rho_{||} \approx 10^{-6}\,\Omega\text{m}$) because the number of charge carriers (electrons and holes) is small.

The high cohesive energy of graphite (see **Table 1**) comes almost exclusively from the strong σ interactions between the sp^2 hybrids and the less strong π interactions between the $2p_z$ orbitals. The atomic layers bind together by weak van der Waals forces. If the distance between the atomic layers of graphite is allowed to increase, each pair of branches, whose degeneracy was lifted off by the interlayer coupling in the graphite band structure, will degenerate into a single branch. The resulting band structure is that of graphene. In graphene, there is one π branch and one π^* branch. These branches cross at the corner K of the hexagonal first-Brillouin zone defined in **Figure 5b**. The Fermi level coincides with the crossing energy, and graphene is a zero-gap semiconductor.

Nanotubes

Carbon nanotubes are cylindrical forms of graphite. A single-wall nanotube (SWNT) is composed of one graphene sheet rolled up on a cylinder. The circumference of the nanotube is a translation vector of graphene, wrapped around in such a way that its two ends coincide. This vector has the form $\boldsymbol{C} = n\boldsymbol{a}_1 + m\boldsymbol{a}_2$, where n and m are two integer numbers (the wrapping indices of the nanotube), $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ being two Bravais vectors of graphene. As in graphite, the electronic structure of a nanotube close to the Fermi energy is dominated by π and π^* states. An SWNT can be metallic or semiconducting, depending on its wrapping indices. When the radius is not too small, the rule is that the (n, m) nanotube is a metal if $n - m$ is a multiple of three, otherwise it is a semiconductor, the gap of which is inversely proportional to its diameter d. This rule derives from the condition that the Bloch vector $\boldsymbol{k}_K$ of the K point (see **Figure 5**), where the π and π^* branches of graphene meet, satisfies cyclic boundary conditions around the circumference: $\boldsymbol{C} \cdot \boldsymbol{k}_K = l2\pi$, with l an integer. In fact, only the armchair (n, n) nanotubes are metallic, the nanotubes with $n - m$ a nonzero multiple of three have a small gap that scales like $1/d^2$ (see **Table 2**). The density of states of two nanotubes, a metal and a semiconductor, are shown in **Figure 6**. In the band structure of a metallic nanotube, two branches derived from the π and π^* branches of graphene meet at the Fermi level (see **Figure 6**). These branches, with nearly linear dispersion, give rise to a plateau of density of states around the Fermi level. Each band crossing the Fermi level contributes one quantum $2e^2/h$ to the conductance of the nanotube when it is connected to two macroscopic electrodes. This means that the minimum resistance of a metallic SWNT is $h/4e^2 = 6.45\,\text{k}\Omega$. Below room temperature, the intrinsic resistance of an isolated nanotube increases upon cooling. This unconventional behavior for a metallic system may be the signature of electron

Table 2 Electronic properties of the single-wall (n, m) nanotube

$n-m \neq \mathcal{M}(3)$	$n-m=\mathcal{M}(3)\neq 0$	$n-m=0$
Semiconductor	Small-gap semiconductor	Metal
$E_g = 2\gamma_0 d_{CC}/d$	$E_g = c_g\lvert\cos(3\theta)\rvert(d_{CC}/d)^2$	$n(E_F) = 2\sqrt{3}d_{CC}/(\pi^2\gamma_0 d)$

d is the diameter, $\theta = \tan^{-1}[\sqrt{3}m/(2n+m)]$ is the chiral angle, $d_{CC} = 0.142\,\text{nm}$, E_g is the bandgap of the semiconducting tubes, $c_g \approx 3.1\,\text{eV}$, $n(E_F)$ is the density of states per atom at the Fermi level for the metallic tubes, and $\gamma_0 = 2.9\,\text{eV}$ is the π-electron interaction.

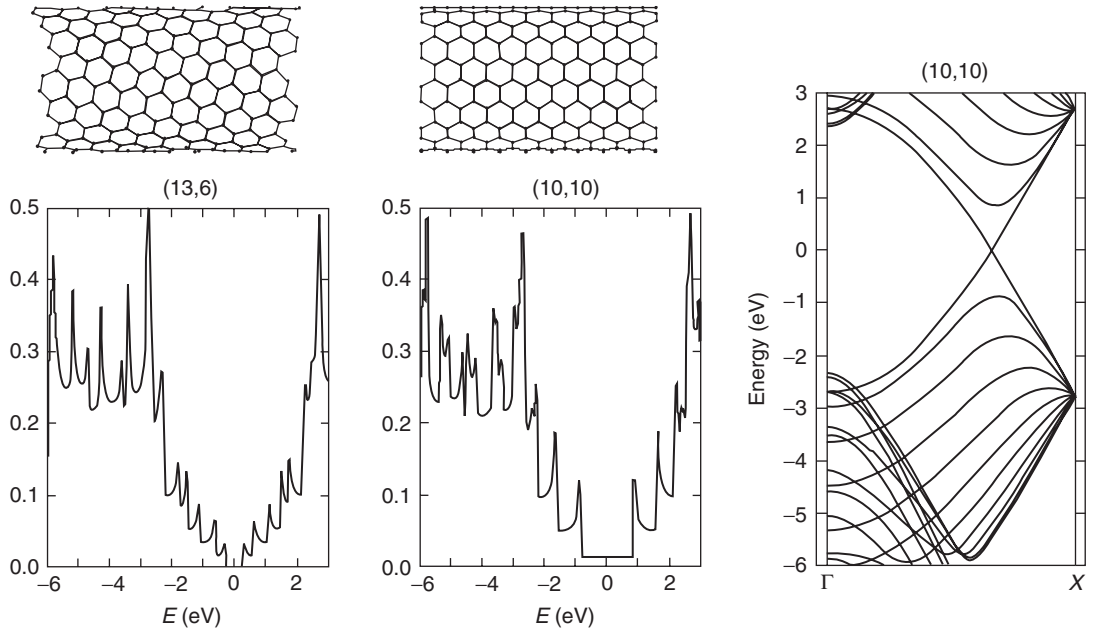


Figure 6 Tight-binding $\sigma + \pi$ electron density of states of the semiconducting (13,6) chiral nanotube and the metallic armchair (10,10) nanotube. The atomic structures of these two nanotubes are illustrated on the top. The band structure of (10,10) is shown in the right-hand side panel, van Hove singularities in the density of states occur at energies where an energy branch has a maximum or a minimum. The Fermi level is at zero energy.

correlation effects (Luttinger liquid), or it may indicate an increase of elastic backscattering cross section of the electrons by defects due to quantum interferences.

SWNTs most often pack together on a triangular lattice, in the form of bundles of 10–200 tubes. The van der Waals attraction between the tubes has small effects on the electronic density of states of the individual components. Multi-wall carbon nanotubes are composed of several single-wall tubes arranged coaxially, with interlayer distance $\sim 0.34\,\text{nm}$. The electronic structure of large multi-wall nanotubes is similar to that of graphite; roughly speaking, they behave like a semimetal. In principle, about one third of the coaxial layers are metallic, the remaining being semiconducting. Experiments reveal that, under high potential bias, each tubular layer of a multi-wall nanotube can transport a current $\sim 20\,\mu\text{A}$ before breaking, which corresponds to a very large current density.

A remarkable characteristic of the SWNTs, directly linked to their 1D periodicity, is the presence of spikes in the density of states (see **Figure 6**), called van Hove singularities. These spikes form real fingerprints of the nanotube structure. They can be probed by optical absorption, resonant Raman spectroscopy, electron-energy-loss spectroscopy (EELS), and scanning tunneling spectroscopy. An example of optical absorption spectrum of bundles of SWNTs with 1.3 nm average diameter is shown in **Figure 7a**. The peaks ~ 0.7 and 1.3 eV are due to transitions between the first pair (E_{11}^S) and the second pair (E_{22}^S) of van Hove singularities that border the gap of the semiconducting nanotubes. The absorption peak $\sim 1.9\,\text{eV}$ is due to interband transitions (E_{11}^M) across the plateau of density of states of the metallic nanotubes (see **Figure 6**). As in all sp^2 carbon systems, collective electron excitations from occupied π to unoccupied π^* bands of a nanotube form the so-called π plasmon, which can be probed by EELS. Momentum-dependent EELS experiments performed in transmission on SWNTs show the π-plasmon peak located at 5.2 eV when the electrons are collected in a

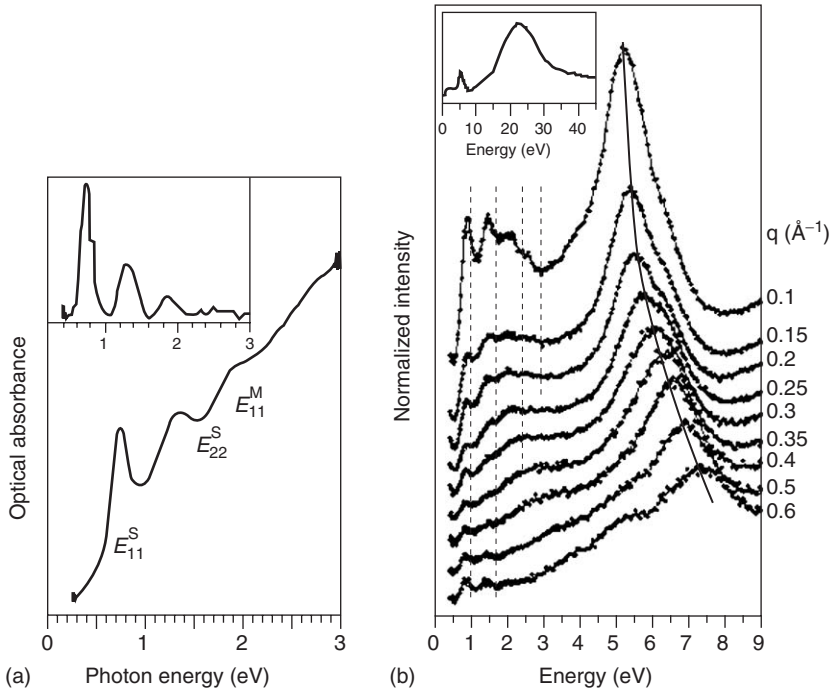


Figure 7 (a) Optical absorption spectrum of bundles of SWNTs, with interband transition peaks superimposed on the wing of the broad π-plasmon absorption band. The same spectrum after removal of the plasmon background is shown in the inset. (b) Electron-energy loss spectrum of purified, nonaligned bundles of single-wall carbon nanotubes as a function of momentum transfer q. The inset shows the loss spectrum in a wider energy range for $q = 0.15\,\text{Å}^{-1}$.

direction close to the incident direction (see **Figure 7b**). In this geometry, the wave vector q transferred from the electrons to the plasmon excitation is small. By increasing the detection angle of the spectrometer with respect to the incident direction, q increases, which produces an upward shift of the π-plasmon. The observed plasmon dispersion indicates dispersive bands of delocalized π electrons, as in graphite. At low energy, there are a few, well-resolved loss peaks, whose positions do not vary with q, and which are due to the interband transitions E_{11}^{S}, E_{22}^{S}, and E_{11}^{M}. In the inset of **Figure 7b**, the loss spectrum for $q = 0.15\,\text{Å}^{-1}$ is shown, the π plasmon is near 5 eV and the $\sigma + \pi$ plasmon, due to collective excitations of all the valence electrons, is near 23 eV.

Fullerites

Fullerites are solids realized by close packing of fullerene molecules, among which are C_{60}, C_{70}, C_{76}, C_{78}, and C_{84}. The cohesion of these crystals is assured by van de Waals interactions between the C clusters. In the C_{60} molecule, which is the most abundant fullerene, the atoms are located on a sphere of 0.7 nm diameter. They are positioned at the vertices of 12 pentagons, the edges of which have a length of 0.145 nm, slightly longer than the 30 bonds (0.140 nm) shared by the hexagons that connect the pentagons together. Due to curvature, the hybridization of the σ bonds in fullerene is not exactly sp^2. Nevertheless, one still may refer to $2p$ orbitals locally normal to the molecular shape surface, that will form π states, and $sp^{2+\varepsilon}$ hybrids, which point along the C–C bonds and will build σ states. The electronic states of the molecule have dominant π or σ characters, with some mixing of both characters due to the strong curvature of the molecule. All together, there are 60 π states, of which 30 are occupied. Similarly, there are 90 σ occupied states. Due to the icosahedral symmetry of the molecule, most of the electronic levels are degenerate. The highest occupied molecular orbital (HOMO) of C_{60} has h_u symmetry (fivefold degeneracy), and the lowest unoccupied molecular orbital (LUMO) has t_{1u} symmetry (threefold degeneracy). The HOMO–LUMO separation is 2 eV. The states near the HOMO–LUMO gap have a dominant π character. Because both the HOMO and LUMO wave functions are antisymmetric with respect to the inversion center of the molecule, the HOMO–LUMO transition is dipole forbidden. The lowest measured optical transition of the C_{60} molecule is ~ 3 eV.

Due to its lower symmetry, the C_{70} molecule has eight different bonds, whose length varies between 0.137 and 0.147 nm. The one-electron states of C_{70} are singly or doubly degenerate, the HOMO–LUMO separation is 1.8 eV. Like all molecules, the fullerenes

are highly-correlated electronic systems. An effective Coulomb energy U can be estimated by $U = I - A - \Delta$, where I is the ionization energy, A is the electron affinity, and Δ is the HOMO–LUMO separation. From the data of **Table 3**, one obtains $U \approx 2.8$ eV, which is a bare value for an isolated fullerene. It can be reduced by 1 eV in the solid state due to screening effects. In the solid, the electronic levels of the individual fullerene molecules are only weakly broadened by intermolecular coupling; they form bands with little dispersion, typically 0.5 eV. The C_{60} crystal (**Figure 8b**) is a semiconductor with a 2.2 eV band gap. The distribution of the occupied states in C_{60} fullerite is clearly revealed by photoemission spectroscopy (**Figure 8b**), whereas the unoccupied states can be probed by inverse photoemission and also by EELS (**Figure 8c**).

The fullerenes can be doped by intercalation (see the next section, "Intercalation compounds"), by incorporation of atoms in the cage (endohedral doping, e.g., Gd@C_{82}, Sc_2@C_{84}), and by atomic substitution (heterofullerenes, such as $C_{59}N$). X-ray photoemission provides valuable information on the electronic structure of these compounds. Inversely, various fullerenes and endohedral metallo-fullerenes can be encapsulated inside single-walled carbon nanotubes.

Table 3 Electronic properties of fullerene molecules

	C_{60}	C_{70}
Symmetry	I_h	D_{5h}
HOMO degeneracy	5	1 + 2*
LUMO degeneracy	3	1 + 2*
HOMO–LUMO gap (eV)	2.0	1.8
Ionization energy (eV)	7.5	7.3
Electron affinity (eV)	2.7	2.8

*Nearly degenerate states

In these remarkable structures, called carbon peapods, the encapsulated molecules form 1D chains. The fullerenes in a peapod can be more densely packed than in bulk fullerite. Their electronic levels form bands, which hybridize partly with the electronic states of the host nanotube.

Intercalation Compounds

Graphite intercalation compounds (GIC) are composed of planes of intercalant atoms or small molecules separated by a few graphitic sheets, in a sequence that repeats periodically along the *c*-axis as in a superlattice. The number of graphitic layers stacked between two successive intercalant planes is called the stage of the GIC. The intercalants can be electron donors (such as alkali metals) or electron acceptors (such as iodine). Depending on the intercalants used and their concentration, a huge class of GICs can be realized with tailored electronic properties, ranging from metals ($\rho_{||} \approx 10^{-8}\ \Omega$m) to insulators.

Similar to graphite, bundles of single-wall carbon nanotubes can be doped by intercalating either electron acceptors (*p*-type doping) or donors (*n*-type doping) between, and sometimes, inside the tubes. In all cases, there is a charge transfer, a noticeable modification of the optical response of the nanotubes, and a strong reduction of the electrical resistance of the bundles. The Fermi level of the nanotube can be shifted by more than 1 eV by electron or hole doping. In the case of alkali intercalation, for instance, the Fermi level moves up across the conduction bands of the nanotubes. As a consequence, the lowest optical transitions E_{11}^S, E_{22}^S, and E_{11}^M, are progressively suppressed as their final states become occupied. The corresponding resonant behavior of the Raman cross section disappears at the same occasion.

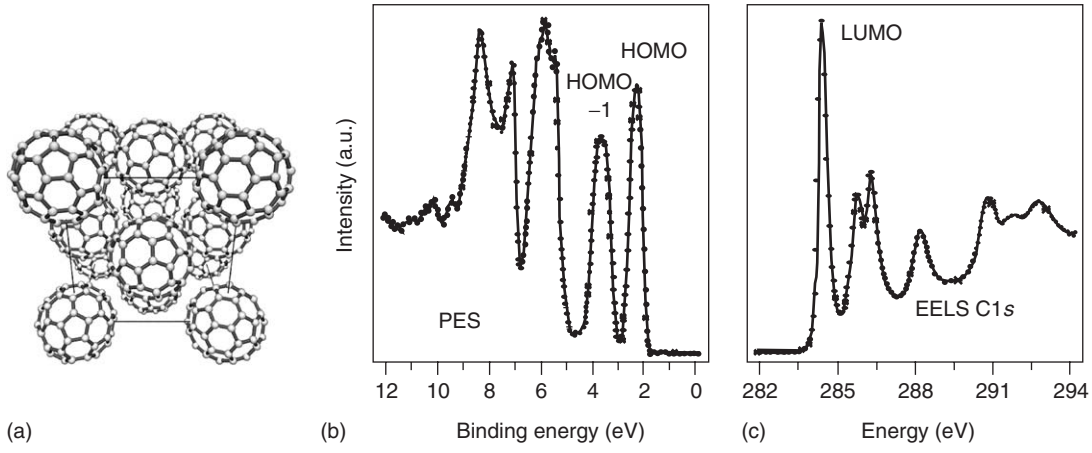


Figure 8 (a) Conventional cubic cell of the C_{60} fullerite, (b) Room temperature photoemission spectrum, and (c) C1*s* excitation EELS spectrum of C_{60} thin film, revealing the distribution of occupied and unoccupied electronic states, respectively.

The fullerites can be doped by intercalation of atoms and small molecules in the octahedral and tetragonal cavities of their crystal lattice. Since the electron affinity of the fullerene molecules is high, they easily accept electrons from the intercalants. In the case of C_{60}, the phase diagram of the A_xC_{60} compounds, with A an alkali metal, is remarkably rich. The x electrons given by the ionized alkali atoms fill in the t_{1u} LUMO-derived band, which may accommodate six electrons. This simple rigid band-structure picture predicts that all the A_xC_{60} compounds should be metallic except for $x = 0$ and 6. In fact, there are strong electron correlation effects in fullerites, as mentioned above, which invalidate this model, and only A_3C_{60} is metallic. Remarkably enough, most of the compounds having that composition are superconductors. The critical temperature T_c increases with increasing the ionic radius of the intercalant: 19 K in K_3C_{60}, 30 K in Rb_3C_{60}, and 32 K in $RbCs_2C_{60}$. The raising of T_c is due to the narrowing of the LUMO-derived band and the resulting increase of density of states at the Fermi energy.

Disordered Carbons

Carbon may form different partially disordered solids depending on the preparation technique used and experimental conditions. All these disordered forms of carbon have considerable technological interest. They belong to two structural families: turbostratic carbons, which derive from graphite (stack of graphene layers, now with orientational disorder in their sequence), and amorphous carbons, which mix variable proportions of sp^2 and sp^3 units in a nonperiodic arrangement. Carbon black is a turbostratic carbon obtained by combustion of carbon-rich precursors under oxygen depleted conditions; its texture can be viewed as a folded version of the graphite network. Glassy carbon, also a turbostratic carbon, with a density ranging from 1 to 1.5 g cm^{-3}, is formed by vacuum pyrolysis of polymerized organic resins; it is built up from nanosized graphene stacks that enclose 1–5 nm sized pores. Graphite-like amorphous carbon is produced by sputtering or evaporation; its density varies between 1.8 and 2.27 g cm^{-3} (density of graphite). Diamond-like amorphous carbon, obtained in general by C atom or ion deposition, has a density intermediate between graphite and diamond (3.52 g cm^{-3}).

The amorphous carbons form a wide range of structures intermediate between the two extreme allotropic forms, diamond and graphite. The density of a carbon sample can be determined from the position of its $\sigma + \pi$ plasmon measured in EELS, by the formula $\omega_p = \sqrt{ne^2/\varepsilon_0 m^*}$, where n is the valence-electron density, directly related to the atomic density, ε_0 is the dielectric permittivity, and m^* is the electron effective mass ($\approx 0.87\, m_e$). The bonding type of a nonhydrogenated, disordered carbon can be characterized by measuring the ratio between sp^3 and sp^2 hybridization states, though there may be a small proportion of sp^1 states as well. EELS and C1s photoelectron spectroscopy give access to the sp^3/sp^2 fraction. Experiments indicate that the sp^3/sp^2 ratio increases linearly with the atomic density of the sample.

When the sp^2 and sp^3 sites are intermixed at the atomic scale, the electronic structure of amorphous carbon near the Fermi level is dominated by π and π^* states, arising from the sp^2 sites, and to a much lesser extent, by disorder-induced σ-band tails due to both sp^2 and sp^3 sites. When the sp^3/sp^2 ratio is large, as in most diamond-like carbons, one speaks of tetrahedral amorphous carbon. This material is hard, electrically insulating, and transparent. The minority sp^2 sites tend to form small clusters; they control the band gap, which can be reduced by a factor of 2 or more with respect to crystalline diamond. In sp^2-rich systems, the sp^3 sites can be randomly and uniformly distributed among the atoms, leading to a non-graphitic disordered structure. The electrical conductivity of such a sample is small and exhibits a semiconducting behavior. By heat treatment, and concomitant migration of the sp^3 defects, graphite-like amorphous carbon is obtained. It can be viewed as a random assembly of nanometer-sized turbostratic entities, partly embedded in a matrix of four-fold atoms. The electronic structure and the

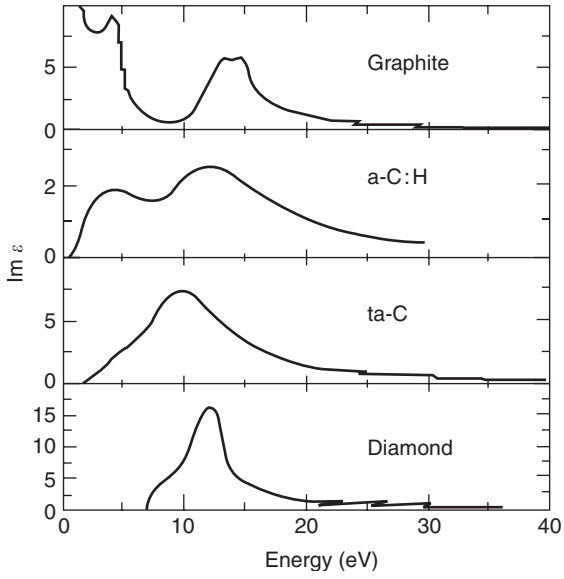


Figure 9 Plot of the imaginary part of graphite, graphitic amorphous carbon, tetrahedral amorphous carbon, and diamond vs. energy.

conductivity, then, depend critically on the size distribution of the turbostratic entities and on their degree of disorder.

To illustrate how the electronic states of amorphous carbons may interpolate between those of graphite and diamond, **Figure 9** shows the imaginary part of the dielectric function of graphite, together with a graphitic-like amorphous carbon which contains some hydrogen (a-C:H), a tetragonal amorphous carbon (ta-C), and diamond. The peaks in these spectra reflect excitations of the π and σ valence states. In graphite, one sees the presence of $\pi \rightarrow \pi^*$ transitions ($\sim 4\,\text{eV}$) and $\sigma \rightarrow \sigma^*$ transitions ($\sim 14\,\text{eV}$). In a-C:H, the $\pi \rightarrow \pi^*$ transitions are slightly reduced (consistent with a small fraction of sp^3 sites), in ta-C, they are strongly reduced (consistent with a large sp^3/sp^2 ratio), and in diamond, they are absent.

See also: Disordered Solids and Glasses, Electronic Structure of; Elemental Semiconductors, Electronic States of; Electronic Structure (Theory): Molecules; Valence Photoemission.

PACS: 71.20. – b; 81.05.Uw; 81.07.De; 81.05.Tp

Further Reading

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Dresselhaus MS (1998) The wonderful world of carbon. In: Yoshimura S and Chang RPH (eds.) *Supercarbon, Synthesis, Properties and Applications*, pp. 10–29. Berlin: Springer.

Forró L and Mihály L (2001) Electronic properties of doped fullerenes. *Reports of Progress in Physics* 64: 649–699.

Gunnarsson D (1997) Superconductivity in fullerides. *Review of Modern Physics* 69: 575–606.

Kuzmany H, Fink J, Mehring M, and Roth S (eds.) (1993) *Electronic Properties of Fullerenes*. Berlin: Springer.

Saito R, Dresselhaus G, and Dresselhaus MS (1996) *Physical Properties of Carbon Nanotubes*. London: Imperial College Press.

Silva SRP (ed.) (2003) *Properties of Amorphous Carbon*. Stevenage: IEE INSPEC series.

Zabel H and Solin SA (eds.) (1992) *Graphite Intercalation Compounds II: Transport and Electronic properties*. Berlin: Springer.

Catalysts: Combinatorial Catalysis

A Hagemeyer and A Volpe Jr., Symyx Technologies, Inc., Santa Clara, CA, USA

Introduction to Combinatorial Heterogeneous Catalysis

The use of combinatorial and high-throughput methodologies to accelerate the research and development process in heterogeneous catalysis is increasing at a rapid pace in both industrial and academic laboratories. For industry, the key drivers include the need to reduce the time-to-market for new and optimized catalysts and processes, increased probability of success due to the ability to perform far greater numbers of experiments than in the past, better intellectual property protection made possible by the thoroughness with which a given technical area can be explored, shorter/more projects possible per unit time, the benefits of "early failure," and the increased organizational efficiency resulting from improved data storage, access, analysis, and sharing. Similar advantages have already been addressed by the pharmaceutical industry, where long development times and high research costs have forced the development of high-throughput approaches to accelerate the drug discovery process.

Traditional methods for the discovery of new heterogeneous catalysts are not very efficient because discovery protocols are primarily trial-and-error processes. The ability to predict the required catalyst composition, structure, and formulation for a given chemical transformation is low and, for complex multicomponent catalysts, almost nonexistent. In addition, there are many variables that affect a catalyst, including not only the elemental composition but also metal precursor types used, wet synthesis variables, method of post-treatment such as calcination conditions, presence and type/shape of catalyst support, and operating conditions such as temperature, pressure space velocity, and reactant gas composition. These variables are too numerous to adequately explore using conventional methods, especially in discovery programs, but also in catalyst optimization work where there may be fewer variables but where catalyst complexity (catalysts are often multicomponent systems) is high.

The high-throughput experimental process in heterogeneous catalysis involves the software assisted design of diverse, high-density assemblies or arrays of potential catalytic materials (known as "libraries"), and high-throughput synthesis, characterization, and screening techniques that are characterized by using robotics and advanced software. The integrated synthesis and screening of a plurality of catalysts in library format has been recognized as an essential factor. Equipment miniaturization and integrated data management systems are also key

aspects of successful workflows. The development and implementation of these methods requires the involvement of unconventional engineering and software resources not commonly available at chemical, refining, and petrochemical companies where heterogeneous catalysis is practiced. High-throughput research is optimally performed by interdisciplinary teams typically made up of chemists, engineers, and programmers.

Work beginning in the mid-1990s, led by Symyx, has significantly advanced the field of high-throughput experimentation and initiated the many efforts that are currently underway. The number of experiments that can now be performed using state-of-the-art high-throughput workflows can be an order of magnitude or more higher than was possible only a few years ago using conventional research. For example, a high-throughput program can yield 50 000 experiments per year compared to 500–1000 experiments using traditional methods. This article describes the current state-of-the-art synthesis and screening techniques for high-throughput experimentation in heterogeneous catalysis with a focus on technology developed over the last 2–3 years. For focus, this discussion is being limited to gas–solid systems. Earlier work has been extensively reviewed.

The Hierarchical Workflow in Heterogeneous Catalysis

The research and scale-up phases leading to commercialization are shown in **Figure 1**. The high-throughput workflow can be divided into primary, secondary, and, in some cases, tertiary screening. Primary screening approaches are typically very high throughput qualitative or semiquantitative screens performed on small samples often using unconventional reactor designs, and are most often focused on discovery. The objective during this phase is to broadly screen a large and diverse set of material types that may perform the desired catalytic reaction. Discoveries, termed "hits," are then taken to the secondary screen and, importantly, compositional space that is not useful is discarded. Although it is critical that primary screening results correlate with the real catalytic process, this can usually be accomplished by screening for qualitative trends and using relative performance rankings generated by using a simplified analog of the real process parameter. The primary screen must be designed to minimize both false positives and especially false negatives. Primary screening removes the key bottleneck in the R&D process.

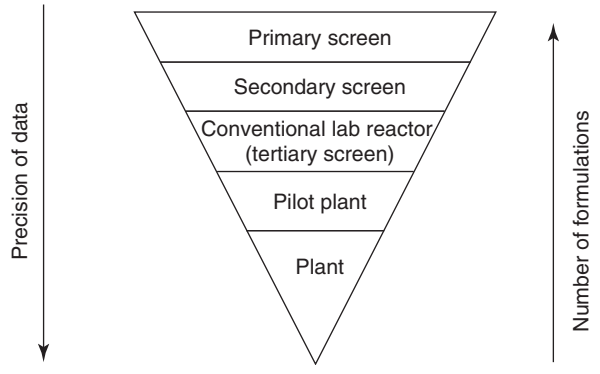


Figure 1 Stages in catalyst discovery, scale-up, and commercialization.

Secondary screening is used for confirmation and optimization of primary hits. In contrast to the primary screen, the catalyst form, reactor, and process analytics are designed to closely represent those of the real bench-scale material and reaction. The data quality and precision should be equivalent to that of a standard laboratory reactor since the goal is to observe small improvements in performance as a function of catalyst modification. Optimized hits, termed "leads," are taken to the tertiary screening phase to generate commercial development candidates. If necessary, tertiary screening can be performed using conventional fixed-bed microreactors with full reactant and product detection and full mass balance, and in some cases is also parallelized. Secondary synthesis and screening technology has, in many cases, evolved to the point where the quality of the data obtained is equivalent to that obtained using conventional laboratory technologies. Accordingly, in some instances, pilot plant studies can proceed directly from secondary screening, without tertiary screening.

A generalized set of high-throughput workflows used in heterogeneous catalysis is depicted in **Figure 2**. The cycle of library design, synthesis, screening, and data analysis is illustrated. Software tools and databasing of synthesis and performance data are key factors.

High-Throughput Synthesis Technologies

For primary screening, catalyst libraries can be efficiently prepared and tested on two-dimensional substrates. Symyx has described the automated preparation of 16×16 catalyst arrays of approximately 1 mg each on 4 in quartz wafers (**Figure 3**). To prepare these libraries, precursor solutions are first premixed in microtiter plates and then volumetrically transferred to quartz or glass catalyst substrates using automated liquid handling robots, either in a rapid serial or parallel manner. For impregnation synthesis, catalyst supports are predispensed onto the

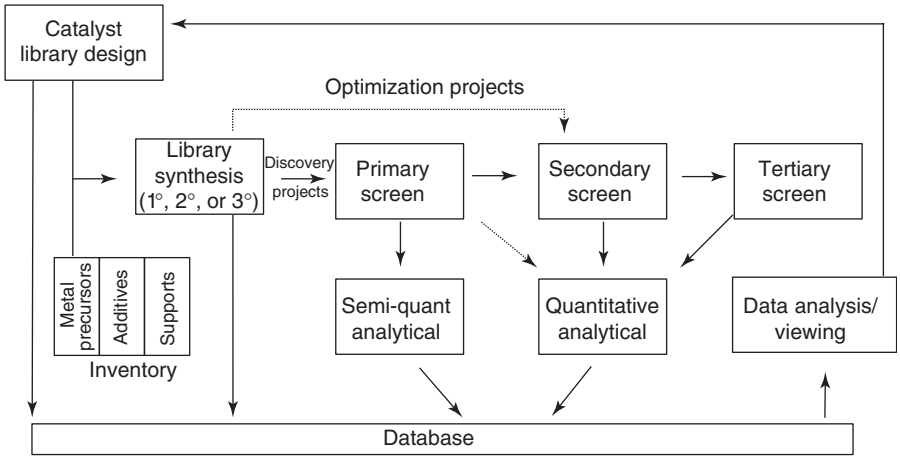


Figure 2 High-throughput catalyst discovery workflow.

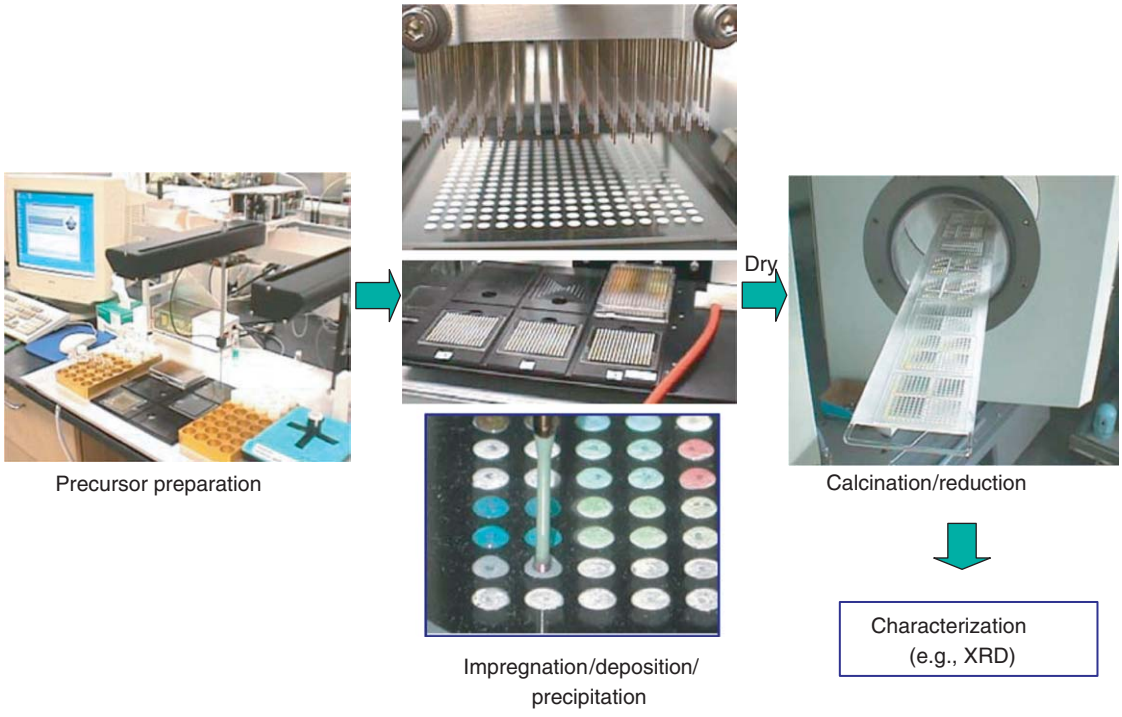


Figure 3 Example of a primary synthesis workflow. Catalysts can be prepared by impregnation, deposition/evaporation, or precipitation.

wafers as slurries resulting in an array of thick film support spots. After drying, the premixed catalyst precursor solutions are added using advanced, often fully parallel, liquid-dispensing techniques allowing precise addition of appropriate liquid volumes (this is very important for incipient wetness impregnation, for example). After drying, the wafers are calcined/reduced to form the final catalyst oxides, metals, etc., and then tested for catalytic activity in primary screening reactors. **Figure 3** shows hardware components for impregnation synthesis, including the liquid handling robot, parallel liquid transfer (impregnation) tool, and post-treatment furnace. This method produces catalysts using preparative methods that closely resemble those used on much larger scales, and can also be used for precipitation and solvent evaporation synthesis.

Apart from parallel synthesis, methods of pooled synthesis have also been reported for primary screening. In one recent proof-of-concept example that is

an alternative to spatially addressable libraries, Mallouk *et al.* have described the synthesis of split-pool material libraries consisting of noble metals on gamma alumina beads. With this technique, the support beads are placed in small vials and impregnated with several combinations of many metal precursor solutions. After drying and thermal treatment, the beads are mixed together and then split again into separate vials. This process is repeated several times to generate a library. The theoretical number of different bead compositions is given by $N\ (\mathrm{nm}) = (n+m-2)!/(n-m)!(m-1)!$, where n is the number of components and m is the number of split-pool operations. The use of fluorescent dyes and micro-X-ray fluorescence as tags for post-synthesis bead identification is proposed. Although this method makes it possible to prepare large libraries in a simple manner, a rapid method of analyzing the individual bead compositions still needs to be developed. Conceivably, some type of single-bead reactor could be used for the screening step. The use of continuous films (as opposed to discrete spots or grids) is another approach for fast parallel synthesis of compositionally diverse flat libraries, prepared, for example, by physical vapor deposition techniques.

Larger amounts of bulk catalysts (up to approximately 1 g) are usually prepared for secondary screening because the testing reactors are generally fixed-bed units. Synthesis methods utilize partially or fully automated procedures that are similar to larger-scale conventional preparations. For example, an impregnation synthesis workflow for secondary screening in use at Symyx consists of several stations for precursor preparation, support dispensing, impregnation, catalyst washing, and thermal post-treatment. The metal precursor solution mixtures from the experimental design are prepared in an array of vials from a set of stock solutions. The supports are pre-sized, either by pressing, crushing, and support powders, or by crushing and sieving larger-scale commercial supports such as extrudates or spheres. These supports are dispensed into an array of vials using a powder-handling robot prior to impregnation. Obtaining good liquid–solid contacting during the impregnation is important for obtaining a uniform catalyst. An effective way to achieve this is by mechanically fluidizing the supports while adding the precursor solutions. After drying and sometimes calcination, the catalysts may be washed with water or another solvent to remove undesirable components and thermally treated under the appropriate gas flow conditions for oxidation, reduction, sulfidation, etc. Inspection of catalyst particles by optical microscopy as well as SEM/EDS is often used to investigate catalyst distribution within the support particles. Precipitation and solvent evaporation workflows to synthesize bulk catalysts are also available.

In the zeolite area, high-throughput hydrothermal secondary synthesis workflows have also been reported. For example, the preparation of new, open-framework zinc phosphate compounds has been recently reported by Xu *et al.* using a multi-well autoclave and automated XRD analysis using a GADDS microdiffractometer for characterization. The authors demonstrate the utility of a crystallization reactor containing 64 Teflon chambers at the 800 µl scale. Reagent dispensing and mixing as well as product separation and characterization are automated. Several other multi-autoclaves for the synthesis of molecular sieves can be found in the literature.

Some of the challenges in these secondary synthesis workflows, especially when making unsupported catalysts, are rapidly and reproducibly sizing the catalyst particles properly for the secondary reactor-bed dimensions, and performing solids processing steps such as grinding, which are often important synthetic variables. Processing steps such as parallel pressing, grinding/crushing, and sieving of catalysts are beginning to be addressed.

High-Throughput Screening Technologies

The focus of high-throughput heterogeneous catalysis over the past few years has been on the development of new reactors. Designs range from scaled-down traditional laboratory reactors sharing some common components such as feed systems and detectors to fully parallel units allowing the evaluation of hundreds of catalysts simultaneously. Several analytical techniques for the high-throughput screening of catalysts have recently been proposed or demonstrated. These include gas and liquid chromatography, thin layer chromatography, resonance-enhanced multiphoton ionization, time-resolved and differential IR thermography, scanning mass spectrometry, photoacoustic analysis, gas sensors, and colorimetry/fluorescence. While GC, IR thermography, and fluorescence or colored-dye assays have been some of the most commonly used methods, especially in early work, new techniques have emerged to identify active catalysts in an array. The need for fast screening of libraries together with the complexity involved in parallel or rapid serial detection of the reactants, products, and side products make it difficult for any single screening method to offer a general solution to the problem. It is important, therefore, to develop screening technologies and methods suitable to the detection requirements of the reaction(s) of interest. In many cases, it is essential to have a combination of screening tools available for screening

under different conditions and to adequately characterize the product streams.

Symyx researchers have recently described a primary screening reactor system based on a microfabricated fluidic flow distribution and equalization device, a 256-element catalyst array, and optical detection methodology allowing fully parallel reaction and detection (**Figure 4**). Microreactors are formed by connecting ("sandwiching") together the catalyst wafer and a gas distribution wafer. Each of the 256 individual reactant streams contacts a 2 mm diameter $\times$ 0.2 mm deep well containing $\sim$1 mg of catalyst, and then flow simultaneously out of the reactors and through a temperature gradient to a cooled absorbent plate/wafer where the products of interest are trapped by absorption or chemical reaction. After sufficient amounts of products have accumulated, the absorbent array is removed from the reactor and sprayed with a dye solution in a separate spray station. A dye is chosen that selectively interacts with the reaction product(s) of interest to cause a change in the absorption spectrum, or the enhancement or bleaching of fluorescence. The developed absorbent array is then imaged by a CCD camera, and the intensities are integrated and the data are stored. Standard catalysts and control spots are present on every wafer to allow normalization and to check the performance of the known standard. This reactor system has recently been used to screen catalysts for the partial oxidation of ethane to acetic acid and the ammoxidation of propane to acrylonitrile, and the results are reported in the literature. A continuous flow parallel 49-channel reactor using a manifold of capillaries as connectors to an analytical device containing filter paper impregnated with an organic dye for the colorimetric assay has been reported by Schuth and demonstrated for NOx abatement.

A fast serial scanning mass spectrometer either independently or in combination with a photothermal deflection detector has been reported for the primary screening of heterogeneous catalyst libraries. Local heating and local sampling of catalyst spots prevents cross talk and interference between neighboring spots. In the case where the products can be identified without overlapping signals, mass spectrometry alone is applicable. In cases where the MS signals overlap, the mass spectrometer is used in conjunction with the additional detection system. For example, the sensitivity of the PTD detector to ethylene is below 0.1 ppm. Related work in this area has been described by the groups of Maier and Senkan, whose scanning MS devices are characterized by common heating of a wafer-formatted array of wells or metal-impregnated shaped beads in parallel channels, respectively.

Willson, and subsequently Maier, independently also describe a parallel method for screening catalyst libraries by differential IR thermography. IR thermography measures the heat evolved due to catalytic activity in exothermic reactions. In Maier's protocol, the image of the IR radiation emitted by the catalyst was used to obtain information on the surface

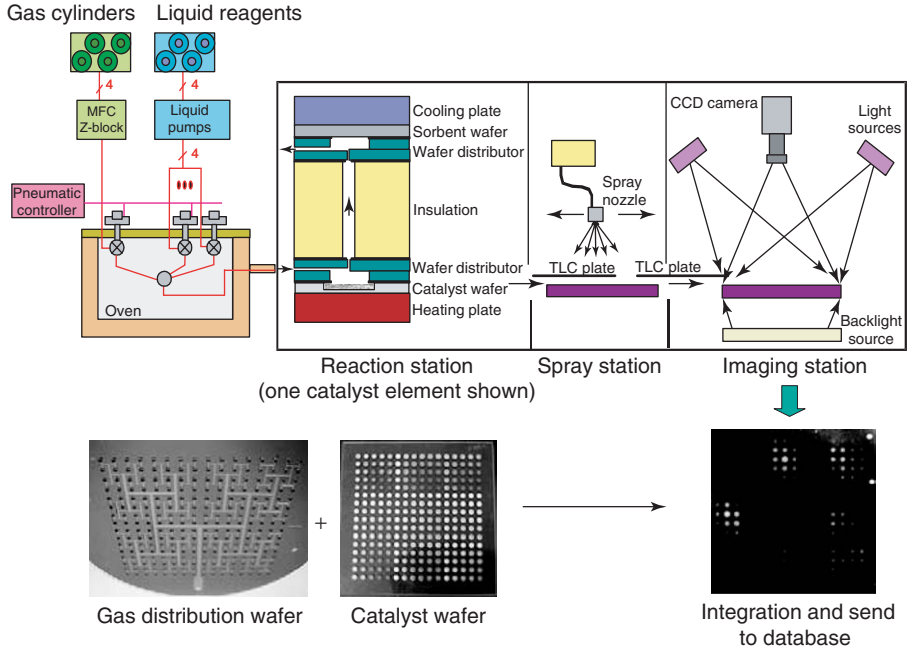


Figure 4 Primary screening reactor for evaluating 256 catalysts in parallel.

temperature and hence the catalyst activity. The IR images of a catalyst library in a reaction block were obtained using an IR camera with a PtSi array detector. The method was demonstrated using a number of examples. In one example, mixed oxides were prepared as amorphous microporous materials (AMMs) by sol–gel. The compositions of the AMM-mixed oxides consisted of a small percentage of a transition metal oxide in silica or titania. The catalysts were prepared *in situ* in the wells of a reaction block. The library was used to study the hydrogenation of cyclohexene as well as the oxidation of isooctane and toluene in air. The mixed oxides with catalytic activity could be seen in the thermographic image of the catalyst library during reaction.

Another interesting method for screening catalysts reported by Smotkin *et al*. utilizes a composite membrane system made up of a semipermeable polymer onto which is overlaid a porous carbon paper on which a catalyst library has been deposited. Detection is by mass spectrometry and the catalysts are locally heated by a CO_2 laser. The membrane separates the catalyst reaction chamber and the vacuum of a QMS. Products flow through the membrane to the mass spectrometer. Note that a loss of detection sensitivity may be a limitation if library density were to be increased over the 12 catalysts currently used.

Secondary screening is normally carried out in parallel fixed-bed reactors sized for catalyst loadings between 15 and 500 mg. **Figure 5** depicts a 48-channel reactor system with sets of eight reactor wells (channels) contained in six individual modules. The effluents of six reactors are analyzed in parallel. Reactant feeds are generated by vaporizing liquid flows from HPLC pumps, high-pressure syringe pumps, and mixing this vapor with gas components metered through mass flow controllers. The reaction feed is divided equally between the 48 channels using flow restrictors such as silica capillaries. The capillaries feed into the inlets of the reactor modules. Each of the eight reaction product streams leaving a reactor module are fed into a stream selection valve where one of the eight effluents is selected and sent to a GC for on-line analysis. The remaining channels are combined and sent through a back-pressure controller to waste.

A key aspect of running a high-throughput reactor successfully is to ensure high system reliability and data integrity. For the above system, for example, after the catalysts are loaded into the reactor and the modules are sealed, a leak check is performed. If a leak is detected, additional valving is used to isolate the location of the leak for repair. After a leak check has been successfully completed, a flow rate check is performed through each of the 48 channels to ensure that the flows are within specifications. Once these checks are completed successfully, the reaction gasses are mixed and stabilized. Reaction temperatures are then set and must reach the set point or the run is stopped. Each analytical selection valve has a bypass input to allow the feed to be directly analyzed on each GC. This allows analytical measurements to be normalized to the reactant feed compositions and allows tracking of any feed compositional changes. During screening, parameters such as temperatures, pressures, and flow rates are monitored and recorded. At the end of a run, the flow rates are measured again to ensure that all flows are still within range. All data are stored in a database for future reference.

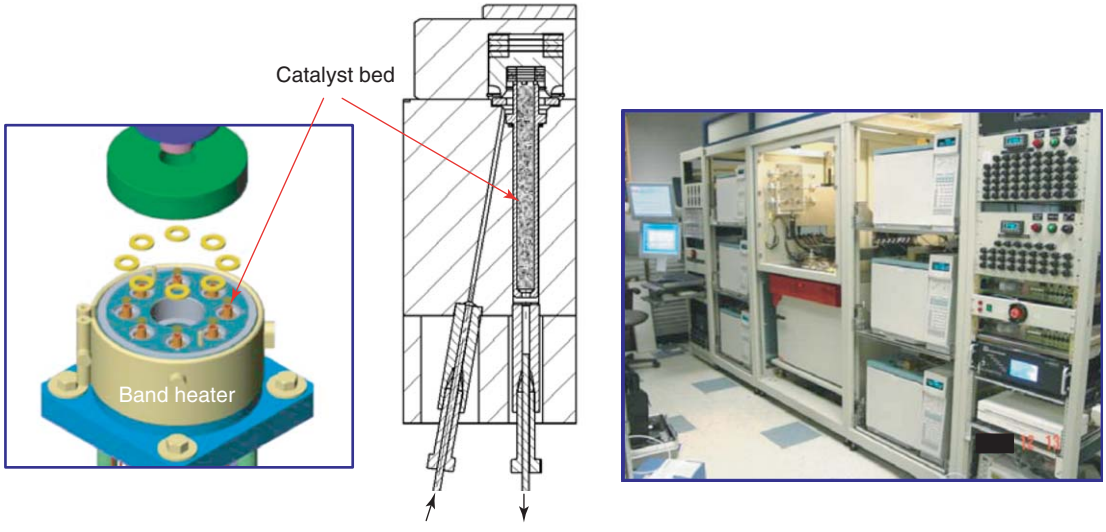


Figure 5 48-Channel fixed-bed reactor for secondary screening in use at Symyx Technologies.

Researchers at IMM, Mainz, have developed a modular parallel 48-channel primary-/secondary-screening reactor. It consists of a gas distribution module, a catalyst titer plate/reaction module, an insulation module to separate hot and cool sections of the reactor, and a multiport valve module to deliver products to a GC. The catalysts are prepared by sputtering or by impregnating catalyst precursor solutions onto a predispensed support on the titer plate. The catalytic combustion of methane was used as a demonstration reaction.

An interesting high-temperature, high-pressure 16-channel secondary screening reactor that utilizes a fast-response stainless-steel photo-etched fluidic switching valve plate that has no moving parts and can operate at temperatures above those of conventional valves has been described by researchers from University of Sheffield. The valve is operated serially, directing one stream to an FTIR detector and 15 streams to a vent line. It is designed to minimize dead volume between the reactor and detector, which consists of a low-volume commercial IR light pipe. The reactor has been used for the evaluation of catalysts for the hydrogenation of CO to ethanol, with quantification of products in ~ 2 min. A 49-channel parallel flow reactor for use at pressures of up to 5 MPa has been developed by Schuth's group and applied to the high-pressure methanol synthesis over Cu/ZnO catalysts.

High-Throughput Characterization Techniques

It is often of interest to characterize properties of catalysts synthesized by various high-throughput methods, especially in the secondary screening phase. Currently, a mix of conventional techniques and high-throughput methods are in use. The current trend is to adapt conventional single-channel characterization tools with movable x–y stages or auto samplers to increase throughput by automated serial analysis of catalyst libraries. Pore volume, acid site density, and metal surface area can be obtained by gas adsorption; however, there have been few, if any, reports of high-throughout methods for these analyses, although auto samplers are often used. Particle-size distribution via light scattering, morphology via SEM or TEM, phase and structure by XRD, and elemental composition via EDS and XRF have been implemented for use on arrays of catalysts.

Applications

This section lists examples of the application of high-throughput methods to heterogeneously catalyzed chemistries. Combinatorial catalysis is ideally suited for the discovery of novel noble metal and mixed metal oxide catalyst formulations for total combustion/VOC removal, emissions control from stationary and mobile sources (NOx abatement, CO oxidation, automotive three-way catalysis) due to the uncomplicated feeds and product mixes that often allows truly parallel detection and very high sample throughputs in primary screening. Furthermore, combinatorial methods are advantageously applicable to multicomponent catalysts (mixed metal oxides, alloys) for selective oxidation, hydrogenation, and dehydrogenation. The demand for higher selectivities and efficiencies in refining has triggered the search for new zeolite structures by parallel hydrothermal synthesis. Additionally, improved fast analytical techniques are making it increasingly possible to perform high-throughput experimentation on complex real plant feeds, especially in the secondary screening phase.

Published examples of extended and comprehensive gas-phase combinatorial screening programs include ethane oxidative dehydrogenation leading to the discovery of novel Ni–Nb–Ta catalysts (Symyx), combustion catalysis due to the demand for better low-temperature activity (Maier) in exhaust gas and air cleaning, automotive emission, stoves, and explosion prevention sensors, the water gas shift reaction aiming at high activity nonpyrophoric catalysts for fuel processors in future fuel cell driven vehicles (Symyx, Mirodatos), preferential CO oxidation in excess hydrogen (Mirodatos, Wolf), propylene oxidation over supported metals (Senkan), isobutane oxidation over mixed metal oxides (Maier), low-temperature light paraffin isomerization catalysts (Corma), methanol synthesis catalysts (Schuth), ethylbenzene oxidative dehydrogenation (Degussa), HCN synthesis over supported noble metals (Baerns, Degussa), new anode and cathode formulations for PEM fuel cells (Symyx, Mallouk, Johnson Matthey), and diesel tailpipe emissions control (HTE).

Future Trends in Combinatorial Catalysis

Combinatorial methods are allowing the exploration of very large and diverse compositional, structural, and process spaces, much of which would otherwise go unexplored, but from which new and unexpected discoveries often arise. The result is an increasing probability of discovering new catalytic materials, facilitated catalyst and process optimization, and the availability of large amounts of information to aid the chemist in the development of new heterogeneous catalysts.

Looking forward, it is clear that the field of combinatorial heterogeneous catalysis will continue to advance and be applied at an ever-increasing rate.

Areas in which advances are needed and are being pursued include: acceleration of scale-up activities using commercial-sized catalysts (e.g., large commercial-size pellets with control over active catalyst distribution in the support), performing process optimization in high-throughput reactors, development of new equipment that can handle harsher reaction feeds and conditions, including high temperature, pressure, corrosive and heavy feeds, improved and faster analytical methods for analyzing reaction products from screening (including analyzing complicated real plant feeds) and for faster/better characterizing catalyst materials, and further miniaturization and automation.

See also: Catalysts: Materials; Scattering, Elastic (General).

PACS: −82.65.+r, 81.16.Hc, 82.45.Jn

Further Reading

Bergh S, Cong P, Ehnebuske B, Guan S, Hagemeyer A, *et al.* (2003) Combinatorial heterogeneous catalysis: oxidative dehydrogenation of ethane to ethylene, selective oxidation of ethane to acetic acid, and selective ammoxidation of propane to acrylonitrile. *Topics in Catalysis* 23: 65–79.

Busch OM, Hoffmann C, Johann TRF, Schmidt HW, Strehlau W, *et al.* (2002) Application of a new color detection based method for the real-time parallel screening of $DeNO_x$ catalysts. *Journal of the American Chemistry Society* 124(45): 13527–13532.

Cawse JN (2001) Experimental strategies for combinatorial and high-throughput materials development. *Accounts of Chemical Research* 34: 213–221.

Guliants VV (ed.) (2001) Current developments in combinatorial heterogeneous catalysis. *Catalysis Today* (Special Journal Issues) 67.

Guram A, Hagemeyer A, Lugmair CG, Turner HW, Volpe AF Jr, *et al.* (2004) Application of high throughput screening to heterogeneous liquid and gas phase oxidation catalysis. *Advances in Synthesis Catalysis* 346(2+3): 215–230.

Hagemeyer A, Jandeleit B, Liu Y, Poojary DM, Turner HW, *et al.* (2001) Applications of combinatorial methods in catalysis. *Applied Catalysis A* 221: 23–43.

Hagemeyer A, Strasser P, and Volpe A (eds.) (2004) *High Throughput Screening in Chemical Catalysis: Technologies, Strategies and Applications*. Weinheim: Wiley-VCH Verlag GmbH.

Jahnisch K, Hessel V, Lowe H, and Baerns M (2004) Chemistry in microstructured reactors. *Angewandte Chemie International Ed.* 43: 406–446.

Liu Y, Cong P, Doolen RD, Guan S, Markov V, *et al.* (2003) Discovery from combinatorial heterogeneous catalysis. A new class of catalysts for ethane oxidative dehydrogenation at low temperatures. *Applied Catalysis A* 254: 59–66.

Maier WF (ed.) (2003) Combinatorial catalysis. *Applied Catalysis* 254.

Müller A, Drese K, Gmaser H, Hampe M, Hessel V, *et al.* (2003) Fast preparation and testing methods using a microstructured modular reactor for parallel gas phase catalyst screening. *Analytical Chemistry* 74: 1933–1938.

Murphy V, Volpe AF Jr., and Weinberg WH (2003) High throughput approaches to catalyst discovery. *Current Opinion in Chemical Biology* 7: 427–433.

Perego C (ed.) *Catal. Today* 81 (2003) (Proceedings of the Euro-CombiCat, Ischia, Italy, June 2–5, 2002).

Schüth F, Busch O, Hoffmann C, Johann T, Kiener C, *et al.* (2002) High-throughput experimentation in oxidation catalysis. *Topics in Catalysis* 21: 55–66.

Senkan S (2001) Combinatorial heterogeneous catalysis – a new path in an old field. *Angewandte Chemie International Ed.* 40: 312–329.

Song Y, Yu J, Li G, Li Y, Wang Y, *et al.* (2002) Combinatorial approach for hydrothermal syntheses of open-framework zinc phosphates. *Chemical Communications*. 1721–1721.

Sun Y, Chan BC, Ramnarayanan R, Leventry WM, Mallouk TE, *et al.* (2002) Split-pool method for synthesis of solid-state material combinatorial libraries. *Journal of Combinatorial Chemistry* 4: 569–575.

Xiang XD, Sun XD, Briceno G, Lou Y, Wang KA, *et al.* (1995) A combinatorial approach to material discovery. *Science* 268: 1738–1740.

Catalysts: Materials

A Hagemeyer and A Volpe Jr., Symyx Technologies, Inc., Santa Clara, CA, USA

Definition of Catalysis

A "catalyst" is defined as a compound that increases the rate of a chemical reaction, but which is itself not consumed or changed by the reaction. A catalyst does not, however, affect the thermodynamics of a reaction or the equilibrium composition of the reaction components.

The Importance of Catalysis

Catalysis is very important and widely applied in the chemical and pharmaceutical industries, in refining, and in environmental protection. A large number of catalysts and catalyst types are utilized commercially, including heterogeneous catalysts (porous solids), homogeneous catalysts (dissolved in liquid reaction mixture), and biological catalysts (as enzymes). Catalytic processes include single-phase homogeneous liquid, two-phase heterogeneous liquid–solid and gas–solid (the most prevalent), and three-phase heterogeneous liquid–solid–gas, implemented using a

variety of reactor types. Catalyst activity, selectivity, chemical and mechanical stability, lifetime, and form determine the economics and capital cost of each process. Control factors to tune catalyst properties are bulk chemical composition, surface composition, stability of active phases, active site distribution, texture (i.e., surface area and porosity), and mass and heat transport. The method of preparation is also crucial to final catalyst performance.

Catalytic processes are superior to noncatalytic synthetic routes mainly because of energy, raw material savings, and selectivity enhancements. Advantages of heterogeneous catalysis compared to homogeneous catalysis include easy separation of the catalyst and reaction products, minimized reactor corrosion, and larger ranges of possible reaction conditions, which is important when the equilibrium limits the reaction to extreme conditions. Typical heterogeneous catalysts comprise metals or metal oxides and inorganic microporous materials such as zeolites. This article focuses on heterogeneous catalysis in the vapor phase.

Applications of Catalysis in Chemical Processes

Thermodynamics frequently limits the concentration of the desired product(s) of a reaction, and conditions (temperature, pressure, space velocity, and reactant composition) must be optimized to maximize the equilibrium concentration of those products. The next step is to develop a catalyst to increase the reaction rate since it is frequently too slow to be practical. It might seem that the optimal catalyst should produce the highest possible reaction rate; however, this is not always the case since such a catalyst might exhibit mass and energy transfer limitations and could be unstable under feasible reaction conditions.

The catalyst and the reactor form an integrated unit. Common reactor types (and associated catalyst formulations) are fixed beds (shaped millimeter-sized bodies), fluidized beds (50–150 μm diameter fluidizable particles), stirred tank reactors (slurried powders), and trickle beds (for three-phase reactions). Reactor operation can be continuous, semicontinuous, or batch, with adiabatic or isothermal temperature control. Consideration of mass and energy balance from the macroscopic catalyst bed scale down to the active center is needed to address the complexity of heterogeneous catalysis (**Figure 1**).

Status of Catalysis Research and Catalyst Discovery

Current approaches used for the understanding and rational design of catalysts include various types of spectroscopy on catalysts and catalyst models, acquisition of kinetic data for catalytic reactions, quantum-chemical and thermodynamic calculations for the elementary steps and molecules involved in a reaction, and microkinetic modeling.

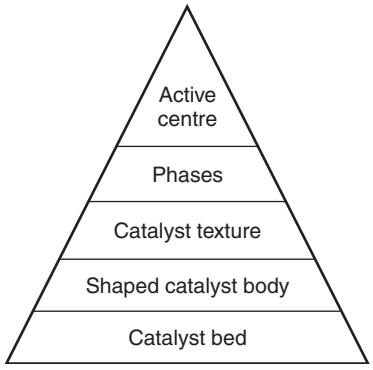


Figure 1 Heat and mass balances can be formulated on five different macroscopic and microscopic scales.

Although catalysis has been used in industry for more than 150 years, the experimental techniques for investigating catalysis at the atomic level did not become routine until about 25 years ago, and the computational techniques have not yet become routine. For this reason, the field has been largely an empirical one. Recently, high throughput methodologies for catalyst discovery and optimization (particularly for complex multicomponent systems where theory is lacking) have emerged and are rapidly expanding.

Kinetic Considerations

The slow step in a reaction mechanism is known as the "rate-limiting step" since it determines the rate of the overall reaction. The rate-limiting step can usually be described as an energy barrier the system must cross. The rate constant has the temperature dependence:

$$k = A \exp(-E_A/RT)$$

where A is the pre-exponential factor and E_A is the activation energy. The activation energy can be determined graphically from experimental determination of the rate constant as a function of reaction temperature through an Arrhenius plot. A catalyst provides an energetically easier reaction pathway to transform reactants into products that is characterized by a lower activation energy (**Figure 2**).

The kinetics of a catalytic reaction is usually measured in a reactor under relevant process conditions and often power law equations such as $r = k\, C_A^n\, C_B^m \ldots$ that are used to model the macrokinetics (n, m are reaction orders for A and B).

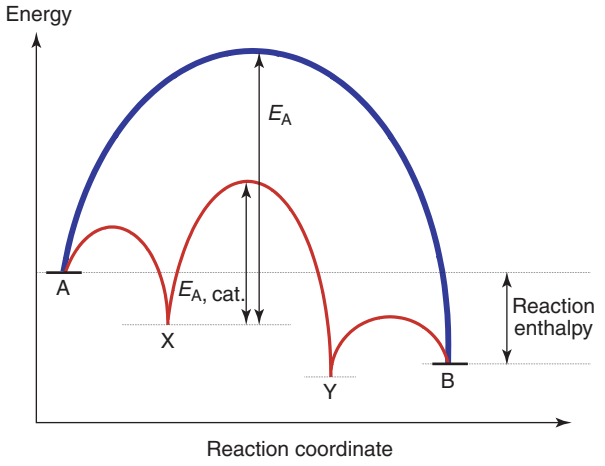


Figure 2 Simplified free energy diagram (red curve – with catalyst; blue curve – without catalyst) for the reaction A→B showing lowering of the activation energy through the use of a catalyst (not the intermediate adsorbed species X and Y).

Aspects of Adsorption

Mass transport and adsorption are important aspects of catalysis. Mass transport exists between the fluid phase (gas, liquid) and catalyst particles, and within a catalyst particle. It is important to determine the influence of mass transport on the catalytic reaction rate and product selectivity. The first step in a catalytic event is adsorption of reactant(s) onto a catalyst surface, followed by chemical reaction(s) at the surface and eventually desorption of products into the fluid phase. There are two main classes of adsorption: physisorption and chemisorption. In physisorption the attractive force is a weak van der Waals interaction with an adsorption energy of $\sim$5–10 kJ mol^{-1} and the chemical bonds in the adsorbing molecules remain intact. Multiple layers of adsorbed molecules are possible. In chemisorption the adsorption energy is comparable to that of a chemical bond, and the molecule may adsorb intact or it may dissociate. Chemisorption energies are $\sim$30–70 kJ mol^{-1} for molecules and 100–400 kJ mol^{-1} for atoms.

The number of adsorption sites on a catalyst is constant and the competition for those sites has important consequences for the macroscopic kinetics and thus catalysis. This is the reason for treating the surface sites as if they were a reactant in the reaction equations. Measurements of adsorption isotherms are used to characterize catalyst porosities and surfaces, to model the reaction kinetics, and help in understanding reaction mechanisms. An isotherm is the coverage considered as a function of temperature and pressure, $\theta(T, p)$. The Brunauer–Emmett–Teller (BET adsorption isotherm for multi-layer physisorption is used in the determination of catalyst surface areas and porosities. Let the surface area be given by

$$\mathrm{SA} = \frac{V_\mathrm{m}}{22414} A_\mathrm{m} N$$

with SA = surface area, V_m = monolayer volume (cc STP), 22414 = molar volume of ideal gas (cc STP), A_m = adsorbate molecular cross-sectional area, N = Avogadro's number. The BET model is applied to isotherm data from $P/P_0 = 0.05 - 0.35$ as fallows:

$$\frac{P/P_0}{V_\mathrm{ads}(1 - P/P_0)} = \frac{1}{V_\mathrm{m} C} + \frac{(C-1)}{V_\mathrm{m} C} \frac{P}{P_0}$$

and from the linear relationship, $V_\mathrm{m} = 1/(\mathrm{slope} + \mathrm{intercept})$ and $C = \mathrm{slope/intercept} + 1$ (the C value is related to the heat of adsorption). The Langmuir adsorption isotherm, on the other hand describes adsorption for single-layer chemisorption, i.e., when the adsorbate forms a monolayer with sites having the same adsorption energy. The Langmuir calculation is applicable only to microporous samples:

$$\frac{P/P_0}{V_\mathrm{ads}} = \frac{1}{V_\mathrm{m} B} + \frac{P/P_0}{V_\mathrm{m}}$$

and from the linear relationship (typically for $P/P_0 < 0.05$), $V_\mathrm{m} = 1/\mathrm{slope}$.

Catalysis

In a catalytic reaction, active sites react and are regenerated and reused in a cyclic manner. The catalyst activity depends on "the number of active sites" and "the turnover frequency." The reaction mechanism consists of many different steps each of which may be of a different type. For example, the Langmuir–Hinshelwood mechanism consists of the following sequence of steps: (1) adsorption from the gasphase, (2) dissociation of molecules at the surface, (3) reactions between adsorbed molecules, and (4) desorption to the gas phase. The Eley–Rideal mechanism covers another important class of reactions and consists of: (1) adsorption from the gas phase, (2) dissociation of molecules at the surface, (3) reactions between gas and adsorbed molecules, and (4) desorption to the gas phase. The Mars–van Krevelen mechanism is often observed for selective oxidations over reducible metal oxide catalysts when lattice oxygen from the oxide is inserted into the hydrocarbon reactant and the lattice defect is subsequently reoxidized by reaction with gaseous oxygen. The oxidation of the hydrocarbon is thus mediated by the catalyst which is the direct oxidizing agent proviing nucleophilic lattice oxygen that is often

more selective than molecular or chemisorbed oxygen species.

The Principle of Sabatier

When different metals are used to catalyze the same reaction, it has been found that the rate of reaction correlates well with the location of the metal on the periodic table. A plot of the rate versus adsorption results in a volcano curve (see **Figure 3b**). This relationship is known as the "Principle of Sabatier." Catalytic activity is highest for the catalyst on which the reactants are most weakly chemisorbed but on which there is still good surface coverage. For this case an increase in adsorption strength causes the adsorbed molecules to be less reactive and a decrease leads to lower activity because the surface is not being fully utilized. In general, the activation energy for the catalyzed dissociation of diatomic molecules decreases when moving left from the noble metals in the periodic table. For CO, for example, the antibonding $2\pi^*$ states are partially filled when chemisorbed and they fill further during dissociation. The antibonding molecular states must be close to the Fermi level of the metal, and when the metal has *d*-states close to the Fermi level, as in the transition metals, there is a strong covalent interaction between the antibonding states and the metal *d*-states. For the early transition metals with low *d* filling, the antibonding metal–C orbital is almost empty due to strong donation resulting in a stronger metal–CO bond, negative adsorption enthalpy, and dissociation (e.g., Fischer Tropsch metals Fe, Co, and Ru), whereas for the late transition metals and noble metals with high *d* filling, the antibonding metal–C orbital is almost full due to strong back-donation resulting in weaker metal–CO bond, positive adsorption enthalpy, and no dissociation (e.g., WGS and methanol synthesis metals: Cu, Pt; see **Figure 3**).

Microkinetic Modeling

The detailed microscopic description of a chemical reaction in terms of the motion of the individual atoms taking part in the event is known as the "reaction dynamics." The link between the microscopic description of the reaction dynamics and the macroscopic kinetics that can be measured in a catalytic reactor is the microkinetic model. Such a model for a given reaction begins with binding energies and

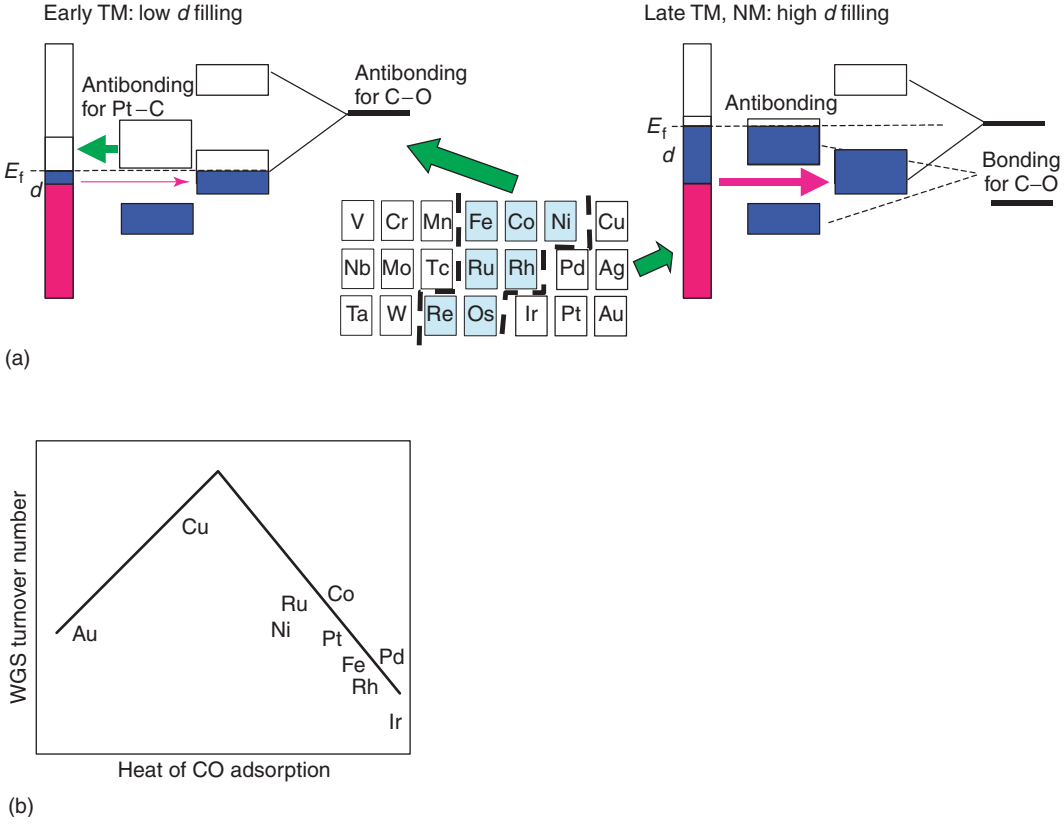


Figure 3 (a) Reactivity trends on alloys for CO dissociation. (b) Volcano curve.

reaction rate constants deduced from surface science experiments on well-defined single-crystal surfaces, and relate these to the macroscopic kinetics of the reaction. The aim is to identify the basic parameters of the reactants and the surface that determine the reaction dynamics (activation barriers, etc.) and relate them to factors determining the catalytic activity. The models are useful for understanding variations in the catalytic activity from one system to another. The stability of the intermediates and the activation barriers are among the input parameters for the microkinetic model, and it is straightforward to calculate the effects of changes in stability for some or all the intermediates.

Catalyst Properties, Shapes, and Preparation

Catalysts for continuous-flow fixed-bed processes have to be shaped to allow the fluid reactant stream to freely pass through the packed bed with an acceptable pressure drop and come into contact with a surface area as large as possible. Catalyst bodies of various sizes, shapes, and geometries have been developed for different applications, e.g., pellets, cylinders, rings, spheres, and monoliths (**Figure 4**). The shape determines the void fraction, packing density, and mechanical strength. The fluid mechanics, flow characteristics and pressure drop, and diffusion restrictions are influenced by the shape and size of the catalyst bodies. Small bodies are often easier to manufacture but suffer from higher pressure drop.

Catalyst pellets used in commercial fixed-bed reactors range in diameter from 1.5 to10 mm. Typical ring sizes are 6–20 mm. Extrudates are 1–5 mm in diameter and 10–30 mm long. Pellets can have several channels and are usually 20–40 mm in diameter and 10–20 mm long (compared to a ring, the use of several smaller channels gives improved mechanical strength with good mass- and heat-transport). Monoliths are used for high gas flow rates or dust (fly ash, soot) containing gas feeds. For the control of emissions in power plant exhaust gas steams, for instance, typical honeycombs are sized 150 mm in edge and 1 m in length, with ~1000 channels per monolith. A stack consists of two to three layers of several hundred honeycombs each. Solid catalysts can be entirely composed of the catalytically active material (bulk catalysts) or supported on porous carriers to improve the mechanical stability or finely distribute the active material thereby creating a large number of active sites.

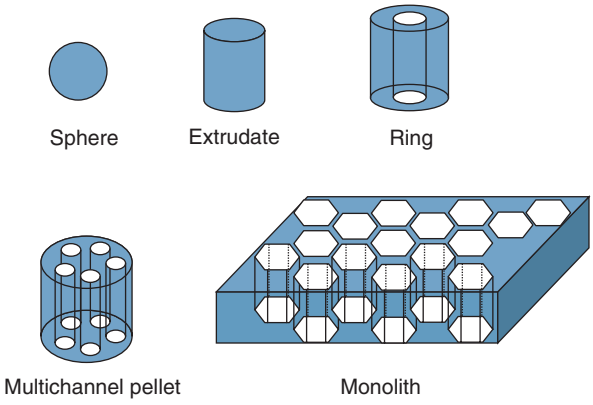


Figure 4 Shaped catalyst bodies.

High activity and selectivity, long-term stability (lifetime), resistance to poisons, tolerance to deviations from operating conditions such as temperature excursions (hot spots), good mechanical strength, and lack of toxic or hazardous ingredients – make high demands on technical catalysts. A variety of synthetic tools are available for the preparation of catalysts, and the performance often more strongly depends on the details of the synthetic route rather than the final composition.

Precipitation (Bulk Catalysts)

Catalysts are precipitated by mixing one solution with another solution or a suspension to form an insoluble compound. Typical precipitation agents are acids or bases to induce a pH change to a metal salt solution. The precipitate is filtered, washed, dried, calcined, and crushed to a fine powder. A binder is added and the powder is tableted. The catalyst may also be shaped before calcination by extruding the filter cake. Typical binders are graphite or stearic acid. Common catalysts prepared by precipitation include alumina-supported Cu–Zn for methanol synthesis and the water gas shift reaction.

Fusion (Bulk, Premolten Alloys, and Metal Oxides)

Catalysts based on metals and metal oxides may be prepared by fusion, crushing, and screening. The finished catalyst consists of irregular pieces of a narrow range of sizes. Fusion is only possible for catalysts that are conductors at high temperatures. During the preparative process, the catalyst powder is placed in an electro-furnace and heated by passing a large current through graphite electrodes. At the start of the cycle, solid catalyst is added to the furnace and the electrodes are lowered. The voltage is fixed and the current is regulated through the height of the electrodes. The molten mass is stirred by the electromagnetic fields generated by the current flowing between the electrodes. At the end of the process, the

electrodes are raised and the molten catalyst is poured out.

The catalyst has no porosity before reduction. The pore system is created when the metal oxide is reduced during the activation of the catalyst (the catalyst may also be supplied in prereduced form). The prereduced catalyst is first completely reduced then passivated by gentle oxidation to allow it to be transported and loaded safely. Common catalysts prepared by fusion include alloy catalysts (e.g., Pt–Rh gauze for ammonia oxidation) and oxides (e.g., Fe oxide for ammonia synthesis or Fischer–Tropsch synthesis), and fused V_2O_5–$K_2S_2O_7$ for sulfuric acid synthesis.

Impregnation (Supported Catalysts)

Impregnated catalysts are prepared by impregnating (e.g., spraying) a solution of a metal salt or mixture onto pellets of a porous support. The metal loading in the finished catalyst is typically 1–5% but can be as low as 0.1%. When liquid is slowly added to a porous solid powder, the liquid is first absorbed in the pores and the powder flows as if it is dry. When the pores are filled, the outer surface of the granules suddenly becomes wet, the granules tend to stick together, and the powder forms lumps instead of flowing freely. Incipient wetness is the point when the pores are filled, but the outer surface of the granules is dry, and can easily be determined by shaking or stirring the powder. Following impregnation to incipient wetness, the pellets are dried and calcined to transform the metal into an insoluble form. The metal salt can be deposited homogeneously through the pellet or most of the metal may be deposited near the outside of the pellet. The distribution of the metal is controlled through the interaction of the metal solution(s) with the support, e.g., through varying the metal counterion, through control of pH, or through addition of chelating agents to the impregnation liquid. The impregnation method (wet versus dry, fast, and slow) and drying conditions can also affect distribution. Typical distributions include eggshell, egg white, and egg yolk. Shell catalysts are often used for selective oxidations to overcome pore diffusion limitations and suppress consecutive over-oxidation reactions, whereas egg-white catalysts are used in refining to protect the inner catalyst layer from feed poisons.

Compared to other catalyst preparation methods, impregnation offers several advantages. For example, the pellets may be shaped before the metal is added; the filtering and washing of the catalysts are not necessary; low metal loadings are easily achieved; and control over the distribution of the metal in pellets is possible. There are also some disadvantages. For example, high metal loadings are not possible without multiple impregnation steps, and an impregnation solution may not be available.

Common catalysts prepared by impregnation include carbon-supported Pt and Pd for hydrogenations, ceria-zirconia supported Pt–Rh for the three-way catalytic converter, and titania-supported V_2O_5 for NO_x abatement.

Evaporative Methods

Wet mixing of (soluble or insoluble) metal precursors in water or organic solvents and subsequent removal of the solvent by evaporation, freeze drying, or spray drying may be used for the production of mixed multimetal oxide catalysts. Often, preformed catalyst powders are slurried and spray dried once again to produce a fluffy and compactable powder for subsequent shaping by extrusion or tablet pressing.

Catalyst Supports

Many heterogeneous catalysts used in industries consist of very small metal particles on a support. Since the catalysis occurs at the surface of the metal particles, the catalysts are prepared so as to expose a large metal area, typically 10–100 $m^2 g^{-1}$of catalyst. The function of the support is primarily to increase and stabilize the area of the metal particles. The size of the metal particles may be reported as the total area of the metal surface, as the diameter of the metal particles (commonly 3–15 nm), and as the so-called metal dispersion (defined as the fraction of all metal atoms that are present at the metal particle surface, commonly from 20 up to $\sim$80%).

A support contains pores having irregular shapes. These pores fall into two size ranges. Macropores have diameters of 100 nm or more and are formed from cracks/spaces between crystallites. Micropores have diameters of 1 nm or less and are due to the "roughness" of the solid surface itself, or the structure of the crystal lattice.

Common and commercially available porous supports are refractory ceramic materials (synthetic such as silica, alumina, titania, zirconia, ceria, magnesia, and niobia, or naturally occurring such as clays, pumice, kieselguhr, and bentonite) and carbon carriers. Their properties are specified by composition, phase, crystallinity, surface area, porosity, packing density, crush strength, and shape. Supports are also used for metal oxide catalysts.

Hydrothermal Synthesis (Zeolites)

Typically, Si and Al precursors together with an organic structure-directing agent are heated up to 200°C under autogeneous pressure in autoclaves for up to several days to synthesize crystalline microporous

zeolites or other molecular sieves such as aluminophosphates. These materials have well-defined pore structures (pore diameters ranging from ~3 to 8 Å) as compared to the broad pore size distributions found in traditional carrier materials. As catalysts, they may contain metal in the pore. Applications of zeolites as supports, solid acids, or shape selective catalysis are common. Zeolites are widely used in refining and petrochemicals, e.g., ultrastabilized Y zeolite for fluid catalytic cracking and ZSM-5 for xylene isomerization.

Raney Metals

Highly porous metal sponges can be made in aqueous solutions by dissolution of aluminum alloys using a strong base. The metal sponges are typically used for hydrogenation in laboratory synthesis. Raney nickel is made by dissolving an (Al, Ni)-alloy containing 50% Ni in 20% NaOH in water. The surface area is 80–100 $m^2 g^{-1}$. The most commonly used Raney metals are Ni, Co, Cu, and Fe.

Colloids

Metal colloids can be prepared by chemical reduction of metal salts in solution using CO, hydroxylamine, oxalic acid, formaldehyde, citric acid, or sodium citrate, or by electro-reduction in the presence of stabilizing salts, ligands, or polymers. Reaction can be carried out directly in colloidal solution or carriers can be impregnated with these preformed nanoparticles.

Catalyst Characterization

Catalysts are characterized for chemical composition, atomic order (crystallinity), phase identification and distribution, porosity, and to identify deactivation mechanisms. Commonly applied techniques include inductively coupled plasma, X-ray fluorescence, physi-/chemisorption measurements (BET, Hg porosimetry, various titration techniques), electron microscopy (SEM, TEM), Auger spectroscopy, X-ray photoelectron spectroscopy, atomic force microscopy, secondary ion mass spectrometry, X-ray absorption, and X-ray diffraction. These methods gather information from different microscopic scales and are complementary.

Catalyst Deactivation

A number of different processes that contribute to the loss of catalyst activity with time on stream are described below. This problem is most often dealt with commercially by slowly increasing the catalyst temperature over time to maintain constant activity/productivity. When it is no longer possible to achieve the necessary performance, the catalyst is replaced.

Sintering

For supported metal catalysts, the active component is present on the catalyst in the form of small particles which are inherently unstable. A catalyst slowly loses activity due to growth of these particles (favored by the resulting decrease in surface energy) and the resulting loss of surface area. There are two mechanisms for sintering: atoms may detach from one particle and move to another (Ostwald ripening), or crystallites may move along the support surface and coalesce. Sintering depends on the nature of the metal, the support, the strength of the metal–support interaction, the gases present, temperature, pressure, and time. Catalyst sintering is irreversible.

Structure Decomposition or Structure Collapse

Crystalline catalysts such as zeolites and multi-metal oxides can deactivate due to destruction of the crystalline phase or lattice structure. Another problem can be loss of elements from the catalyst (or volatile compounds such as Mo, Re, etc.). This can be solved in some cases by the slow addition of the volatile element to the catalyst bed. In the case of attrition losses (e.g., for fluid catalytic cracking (FCC) zeolite catalysts), it is necessary to add in new catalyst slowly over time to the riser.

Fouling and Poisoning

Catalytic activity may also be lost due to the formation of carbon ("coking") or due to the deposition of impurities or of dust or other particulates in the reactant stream onto the catalyst surface. The formation of carbon in many cases is reversible. For example, the catalyst can be taken off line and the carbon removed by oxidation ("catalyst regeneration").

Examples of fouling by particulates occur in removal of NO_x from industrial flue gases or in catalytic converters used in automobiles. In these cases, the catalyst is manufactured in a shape that allows the particulates to pass through the reactor with minimum accumulation (e.g., a catalyst was coated onto a ceramic monolith). A serious problem with catalyst dust may occur if the catalyst bed is improperly loaded, if there is a problem with excessive vibration in the reactor, or if physically weak catalysts are used. In these cases, dust may be formed by attrition, leading to dust accumulation in the reactor or elsewhere in the process impeding or blocking the gas flow. The highly reactive catalyst dust presents a serious safety hazard since it can cause an explosion when pipes or heat exchangers are opened for maintenance.

Examples of poisons that can be present in feeds include metallic impurities such as V, Na, and Ni in FCC processes, metallo-organics in desulfurization

processes, and thiophenic sulfur compounds in various refining processes such as reforming. These poisons react with the catalyst and reduce or destroy its activity. It is often necessary to remove the poisons with a guard bed or dedicated upstream catalytic process (as is done with hydrodesulfurization, for example). Sometimes, the impurity is completely absorbed in the first few percent of the catalyst bed due to its strong interaction with the catalyst.

Future Trends in Heterogeneous Catalysis

Commercial catalysts are used in refining (e.g., alkylation, dimerization, disproportionation, isomerization, catalytic cracking, FCC, hydroprocessing (e.g., hydrotreating, hydrodesulfurization, hydrocracking), hydrogen production, reforming), chemical catalysis (e.g., aromatics production, monomer synthesis, polymerization, organic synthesis, cyclic and acyclic intermediates, medical, flavors and perfumes, rubber processing, plasticizers, surface active agents, agricultural, using oxidation, hydrogenation, dehydrogenation, phase transfer catalysis), biocatalysis, and environmental applications (automobile and industrial pollution control). There are numerous references available covering these applications in detail.

Competitive pressures of a mature industry, the trend toward simpler and less capital intensive processes, access to cheaper feedstocks, and the desire for more environmentally benign processes drives the development of new, high-performance catalysts in the future. New approaches to the discovery of potentially new catalytic materials, including innovative preparation methods and combinatorial/high-throughput catalysis, are having an impact and pushing catalyst development forward. Increased molecular understanding of the mechanism of catalytic transformations based on *in situ* catalyst characterization is also having an impact.

See also: Alloys: Overview; Catalysts: Combinatorial Catalysis; Mass Transport; Thermodynamic Properties, General.

PACS: 81.16. – Hc; 82.30.Vy

Further Reading

Augustine RL (1996) *Heterogeneous Catalysis for the Synthetic Chemist*. New York: Dekker.

Bowker M (1998) The basis and applications of heterogeneous catalysis. *Oxford Chemistry Primers*, number 53. New York: Oxford University Press.

Cornils B, Herrmann WA, Schlogl R, and Wong C-H (eds.) (2003) *Catalysis from A to Z*. Weinheim: Wiley-VCH.

Derouane EG (2001) Catalysis in the 21st century. *CATTECH* 5(4): 214–225.

Ertl G, Knoezinger H, and Weitkamp J (1999) *Preparation of Solid Catalysts*. Weinheim: Wiley-VCH.

Farrauto RJ and Bartholomew CH (1998) *Fundamentals of Industrial Catalytic Processes*. London: Blackie Academic and Professional.

Gates BC (1992) *Catalytic Chemistry*. New York: Wiley.

Hagemeyer A, Strasser P, and Volpe Jr. AF (eds.) (2004) *High Throughput Screening in Heterogeneous Catalysis*. Weinheim: Wiley-VCH.

Le Page JF (1978) *Applied Heterogeneous Catalysis*. Paris: Editions Technip. (Distributed in US by Gulf Publishing Company, Houston (1987).)

Morbidelli M, Gavriilidis A, and Varma A (2001) *Catalyst Design: Optimal Distribution of Catalyst in Pellets, Reactors, and Membranes*. Cambridge: Cambridge University Press.

Rase HF (2000) *Handbook of Commercial Catalysts: Heterogeneous Catalysts*. Boca Raton, FL: CRC Press.

Satterfield CN (1991) *Heterogeneous Catalysis in Industrial Practice*. New York: McGraw-Hill.

Stoltze P (—) *Introduction to heterogeneous catalysis*. Denmark: Aalborg University (http://www.aaue.dk/~stoltze/catal/).

Szostak R (1997) *Molecular Sieves*. London: Blackie Academic and Professional.

Thomas JM and Thomas WJ (1997) *Principles and Practice of Heterogeneous Catalysis*. Weinheim: VCH.

Twigg MV (ed.) (1989) *Catalyst Handbook*. London: Wolfe Publishing Co.

Wijngaarden RJ, Kronberg A, and Westerterp KR (1998) *Industrial Catalysis*. Weinheim: Wiley-VCH.

Ceramic Materials

K P Constant, Iowa State University, Ames, IA, USA

Introduction

Although some of the earliest artifacts from ancient civilizations bear witness to man's long history of utilization of ceramic materials, presently ceramic science is advancing at an unprecedented pace. Ceramics remained an art (often with closely guarded secrets) for a major part of recorded history, and has become the science it is only in the past century. With the pressures of increasingly demanding high-tech applications, whether smaller, larger, hotter, tougher, cleaner, brighter, or smarter, ceramics are often the material of choice.

State-of-the-art characterization techniques allow one to achieve a new level of understanding of atomic-scale processes and capability for manipulation of atomic structure. In effect, ceramics can often be built atom-by-atom, layer-by-layer.

That said, what are ceramics? The definition is becoming increasingly more difficult to articulate. For a significant fraction of society, the definition relates to the art of making and decorating pottery, or other items of "baked" clay. However, from the viewpoint of the scientist or engineer, ceramics are often defined by the chemistry, behavior, or processing methods used to create them. For example, a definition might be: an artifact made of hard, brittle, electrically insulating material produced from nonmetallic minerals by firing at high temperatures. This definition specifies the chemistry, properties, and processing. While this definition is accurate for a large fraction of ceramics, it is not all-inclusive, and examples of ceramics that do not conform to one or more parts of this definition are becoming increasingly common. There are examples of ceramics (some discussed here), which are not brittle and electrically insulating, are not made from nonmetallic minerals, or not fired at high temperatures. In fact, ceramic superconductors are among the most conductive materials known. An acceptable, but not complete definition might be a refractory, inorganic, and nonmetallic material. Even this attempt at description is flawed as the boundaries between biology and ceramics blur.

The field of ceramics (and all materials) is often represented as the four corners of a tetrahedron-labeled structure, properties, processing, and performance. In this article, the state of the art of ceramic science and engineering is presented through discussion of properties, novel processing, and general trends, with examples of the applications that have driven these advancements. The structure and performance are addressed within the context of both the properties and processing of these advanced ceramics.

Properties

Ceramics have long been associated with characteristic properties such as high hardness, high stiffness, resistance to high temperatures, and chemical inertness, but today's ceramics stretch conventional applications of ceramics beyond, what in the past were considered, their fundamental limitations. A huge body of work has been generated since the 1990s, some of which are highlighted here with examples in which mechanical, electrical, and optical properties are critical. Although not represented in the examples below, ceramics are also often used in thermal and magnetic applications in which there are many examples of exciting and novel advances.

Mechanical/Structural

Ceramics are often used where high hardness is required but are usually limited by their lack of toughness. A number of innovations have improved upon this mechanical deficiency including those demonstrated in the development of transparent ceramics for ballistic protection, defense against artillery shells, bullets, and shrapnel in warfare. Toughness can also be improved by the use of hard coatings on tough substrates. Although carbide coatings have long been used for grinding and cutting tools, even harder diamond coatings are becoming increasingly popular. Because the processing techniques required are the most novel aspect of that field, thin film diamond coatings are discussed in the section, "Processing."

Transparent ceramics for ballistic protection One of the most aggressive mechanical applications of ceramics is that of armor. With increasing attention paid to terrorist threats, there is escalating interest in the development of armored vehicles. The chief requirement of armor is that of the ability to sustain repeated impact, but in this application, armor has two other significant constraints. In many instances, it must be transparent and should also be relatively lightweight. As shown in **Table 1**, the ceramic must be optically transparent, and if sensors or guidance systems are used, may also need to be transparent in the infrared or ultraviolet spectra. Although a variety of polymers have been used in many applications, they are not suitable for applications that require optical transparency from the ultraviolet through the mid-infrared (IR) wavelengths. Such applications not only include military applications

Table 1 Wavelength transparency requirements for ceramic armor

Wavelength	*Electromagnetic spectrum*	*Applications*
10–400 nm	Ultraviolet	Sensor windows and radomes
400–700 nm	Visible light	Face shields, windows, windshields for vehicles, Blast shields
700–400 000 nm	Infrared	Sensor windows and radomes

such as forward-looking IR windows, missile domes, underwater sensors, transparent armor, and personnel protection, but also have commercial promise for supermarket scanner windows, scratchproof lenses, lighting applications, and semiconductor wafer carriers. A tough, lightweight, transparent ceramic material has been developed through the US Air Force Research Laboratory working with Raytheon Electronic Systems. This polycrystalline aluminum oxynitride, known commercially as ALON, is optically transparent, has low density, high strength, and high durability. Processing techniques for fabricating sintered corundum are being developed for applications where impact resistance is not critical. A technique that has been shown to be successful involves liquid shaping of sintered and hot isostatically pressed Al_2O_3 ceramics.

Electrical

Ceramics have long been used in the electronics industry as passive devices such as capacitors and insulators, and continue to be critical in those functions; however, in a variety of other electrical applications, ceramics have become increasingly important including sensors (gas and chemical), conductive oxides (mixed conductors), and high-power/high-temperature electronic materials.

Fuel cells The exponential growth of portable electronic devices, cellular phones, PDAs, pagers, computers, etc., provides a strong motivation for the development of fuel cells and high-energy-density batteries. Fuel cells can provide clean energy for a variety of portable energy needs. Ceramic materials play a large role in the development of commercially viable products. Solid oxide, proton exchange membranes, and direct methanol fuel cells, lithium-ion batteries, oxide-ion electrolytes, proton conductors, and mixed ionic/electronic conductors are all ceramic-based technologies. Solid oxide fuel cells (SOFC) are electrochemical energy conversion devices based on ion-conducting oxide electrolytes and are used in a broad spectrum of power-generation applications from small lightweight compact devices to large SOFC/turbine hybrid systems. The cells are solid-state devices operating at high temperatures. Because of this, there is great design flexibility (no liquid management is required as in other "batteries"). There are a variety of designs and geometries currently under investigation. The most common electrolyte materials include yttria (Y_2O_3) stabilized zirconia (ZrO_2), often called 8YSZ, as it contains 8% yttria. This ceramic is ion conducting in both oxidizing and reducing environments. The most common anode material is a cermet (ceramic–metal composite) composed of Ni and 8YSZ. Nickel provides the electronic conductivity and catalytic activity while the 8YSZ supports the Ni particles and improves the anode thermal expansion match. The most common cathode material is a strontium-doped lanthanum manganite (Sr-doped $LaMnO_3$) – called LSM. The interconnect material is usually doped $LaCrO_3$ or a metal that is stable in both oxidizing and reducing environments; however, metal interconnects require lower operating temperatures. The cost of these fuel cells is still a limiting factor in terms of broad applications and ways to reduce the costs are being aggressively sought, including reducing operating temperatures to allow the use of metallic interconnects and increasing the power density. Much effort has focused on the techniques to make dense ceramic coatings using the techniques of the semiconductor industry including chemical vapor deposition (CVD), electrochemical vapor deposition (EVD), physical vapor deposition (PVD), and screen printing as well as more traditional techniques such as sol–gel, tape casting, slurry casting, and centrifugal casting. The former techniques often have film quality and performance advantages, but also have low deposition rates (they are slow) and are costly. The latter techniques are efficient and economical but often result in lower-quality films.

Optical

Although glass has long been used as an optical material, more recently in fibers for communication and laser surgery, ceramics are gaining importance in a variety of optical applications, including wireless technologies, microwave dielectrics, luminescent materials, and nonlinear optical materials. One such emerging field is that of photonic bandgap materials.

Photonic bandgap materials Photonic materials are those in which light is manipulated in novel ways. It is a critically important technology that has wide-ranging potential in a variety of applications, including lasers, catalysis, and optical computing. The field of "photonics" refers to light emitting materials, optical waveguides, nonlinear optical materials, and photonic bandgap materials among others. Most of these applications make use of glasses and dielectric materials. In the specific case of photonic bandgap materials, a high dielectric constant is required. The unique properties of photonic crystals, which allow or inhibit propagation of certain frequencies in all directions, have spurred a flurry of activity in the development of these periodic dielectric structures.

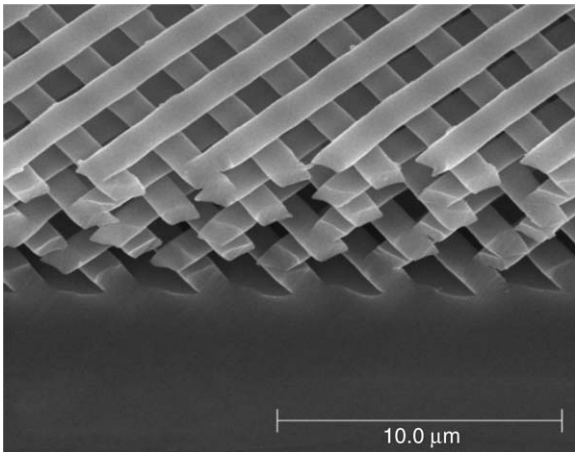


Figure 1 An eight-layer polymer template used to mold a ceramic titania bandgap crystal. (K Constant, unpublished work.)

One method of creating these structures is through a process called microtransfer molding, where a sacrificial polymer template is created which represents the inverse of the desired structure. The creation of this template is a multiple-step process, but requires relatively simple technology, and is quick and economical. A master mold containing parallel bars is created using standard photolithographic techniques. A flexible reproduction (or mold) is created from polydimethylsiloxane (PDMS) by pouring the prepolymer over the master (**Figure 1**). The channels of this mold are then filled with a polymer that is cured and then transferred to a substrate. This process is repeated, rotating the bars perpendicular to the previous layer and aligning them over the channel of the layer, two layers down. A titania ceramic precursor is then infiltrated into the polymer structure and allowed to gel. The gel is converted to a ceramic (TiO_2) and the polymer is removed through heat treatment. The result is a ceramic structure with a configuration that resembles the diamond cubic structure and has a photonic bandgap. If the structure is built with a periodicity $\sim$400–700 nm, a photonic bandgap at optical frequencies will be present. Currently, 1 μm periodicity has been achieved and work toward smaller structures is ongoing.

Processing

The processing of advanced ceramics can be divided into two general classes with respect to the starting material – those that involve powders (powder processing) obtained by a variety of methods, and those that do not. In this section, recent advances in each method are presented through a specific example of each.

3D Rapid Prototyping

A recent advance in the forming stage of ceramic processing is that of 3D rapid prototyping. There are a variety of techniques that fall into this general category. The concept is to translate a 3D computer-generated design (CAD model) to a physical structure without the traditional step of making a mold in which to shape the product or the time and cost associated with machining a ceramic to desired dimensions. These methods can often make very complex geometries in a variety of shapes, sizes, surface finishes, and materials compositions. Most of these techniques rely on cutting 2D "slices" from the 3D model and building them one on top of the other. These techniques usually begin with a bed of fine ceramic powder. Some techniques deposit a binder into specific areas of the powder to "select" that which will remain after the entire shape is formed.

Other techniques use a laser or directed light to cure a polymer component added to the ceramic powder for that purpose. Layer upon layer of powder is added to the bed, each layer representing a subsequent 2D slice of the desired product. Regardless of the technique, the result is that some ceramic powder is held in place as part of the shape to be made, while the remainder can be shaken out of the part. The powder is then further consolidated through heat treatment and sintering to form a monolithic ceramic part. These techniques have a variety of applications including those that are obvious such as prototyping parts for testing before investing in tooling for mass production, as well as the less obvious, such as one-of-a-kind designs, for example, prosthetic devices for biomedical uses.

One blend of rapid-prototyping and biomimetics recently reported is that of computer tomography (CT) assisted artificial bone design. (Biomimetics is discussed in further detail in a later section.) In this work, a 3D image of natural bone is acquired using CT, and used to create a duplicate by rapid prototyping. Currently, the resulting bone-structured ceramic is being used to study osteogenesis (the building of bone), but it could be envisioned that it could be used to fabricate a replacement for missing or diseased bone. Another approach to rapid prototyping and its role in the fabrication of MEMS is discussed later.

Thin Film Techniques

A variety of thin film techniques have emerged from the semiconductor industry, many involving ceramic materials such as oxides, nitrides, and carbides. Most of them rely on liquid or gas phase precursors and sophisticated control of chemistry, temperature, and

pressures. These techniques often result in the growth of thin films of chemistries and morphologies not achievable in bulk form through traditional powder processing in similar temperature or pressure regimes (or sometimes in any temperature or pressure regime). One example of such techniques is presented here: CVD (chemical vapor deposition) of diamond thin films.

Thin film diamond coatings Diamonds are truly an extreme material, having several properties that can only be described in superlatives: the hardest, the stiffest, the highest thermal conductor, highest sound propagation velocity, etc. Additionally, diamonds are very resistant to chemical corrosion, are biologically compatible, are optically transparent over a huge range of frequencies, and are excellent electrical insulators. However, their scarcity and availability only in single-crystal form have prevented widespread xploitation of most of these properties. With the exception of industrial diamonds created for cutting and grinding, the use of synthetic diamonds has been very limited. Rather than attempting to recreate nature's approach to diamond making (extreme temperatures and pressures), an alternative method has been discovered which has two significant advantages over nature's approach – easier processing resulting in a more convenient form. This method, which induces diamond growth by adding carbon atoms one-at-a-time to a template, can be achieved without high temperature and pressure, and results in a thin film coating rather than isolated single crystals. Although this idea was conceived several decades ago (in the 1950s), it was not until the mid-1980s that good-quality films were grown. The concept, while relatively simple, presents numerous engineering challenges. At the heart of the technique is CVD. In this technique, a gas-phase chemical reaction occurs above a solid surface (often called a substrate as it forms the mechanical support for the resulting film). For the case of diamond films, a method of activating the gas-phase carbon-containing precursors is required. These include a hot filament or electrical discharge, or a combustion flame.

The most common method of diamond thin film growth uses methane, CH_4, as a precursor gas. The gas concentration (usually diluted with H_2) and temperature of the substrate must be controlled to suppress the deposition of graphite or amorphous carbon. The method used is normally tailored to the desired application. For example, hot-filament techniques, which use a refractory metal (such as tungsten or tantalum), can be used to make good-quality diamond films for mechanical application, but they are not suitable for electronic applications because they contain metallic impurities from the filament that degrade the electrical properties.

Although requiring more expensive equipment than hot-filament growth, microwave plasma CVD reactors are currently the most widely used technique for diamond growth as they produce high-quality, chemically pure films. The substrate requirements for this process are significant. Among them, the substrate must be able to withstand the thermal and chemical conditions of the reactor, which precludes a large number of materials including all polymers, most metals, and many electronic materials. Also, because diamond has such a low thermal expansion coefficient, α, it is susceptible to delamination or flaking off a substrate which contracts more on cooling from the reaction temperature (usually about 700°C, so a substrate with a low α is required). In addition, the substrate must be able to form a carbide to some extent, as it has been suggested that the diamond films actually grow off an initial carbide interfacial layer. Currently, silicon wafers are the most commonly used substrate material.

The chemical and physical processes which occur during thin-film growth are complex, but a very simplified model can be described by the schematic in **Figure 2**. Not all reactions result in diamond; the process is slow and requires very careful control of both the chemical and physical conditions of the system. Much effort has been expended to understand the gas-phase chemistry of this system, and progress continues to date, resulting in improved, more efficient methods and higher-quality thin films.

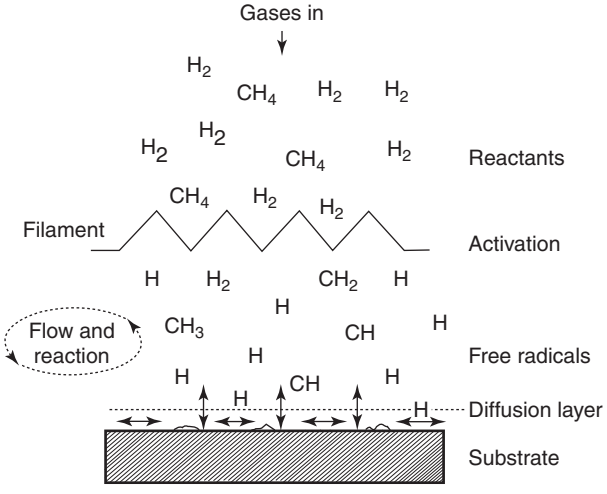


Figure 2 Schematic of the physical and chemical processes occurring during diamond CVD. (Reproduced with permission from May PW (2000) *Philosophical Transactions of the Royal Society of London. Series A* 358: 473–495; © Royal Society.)

The widespread application of diamond thin films is still limited by economic restrictions, but they are beginning to be used for cutting tools, thermal management where rapid heat dissipation is required, IR windows, and electronic devices. Recently, work has been reported which uses diamond thin films as an ideal substrate for integrating microelectronics with biological sensing functions. More on this topic is discussed in the section, "Trends in ceramic materials."

Ceramics and the Environment

Increasingly strict legislative controls and escalating public concern for the local and global environment have affected most manufacturers of goods, including those of ceramics as is testified to by the recent creation of ISO 14000 – the international standards organization that is focused on certifying manufacturers and organizations in environmental management systems.

The environmental impact of any class of materials can be considered to have many aspects including: the source of raw materials, the energy and by-products associated with all stages of manufacture, and the disposal of the product when it is no longer useful. Likewise, ceramics can also be considered in terms of their use in remediation of the environmental impact of energy production. There has been increased awareness of each of these aspects in the manufacture of ceramics as well as in other materials. One such environmental application of ceramics has already been discussed in SOFC for clean energy storage solutions. With respect to energy production, research continues in the area of coal gas cleanup – that is in the development of ceramic sorbents for use in integrated coal gasification combined cycle (IGCC) systems that burn coal for the production of electricity. These calcia-based sorbents have been designed not only to remove sulfur from the gas used to drive the turbines but also to be regenerable – so that the sorbent itself does not create an environmental burden as a waste product. In this work, natural lime is encased in a ceramic shell that provides structural integrity as well as a porous coating for gas flow (**Figure 3**).

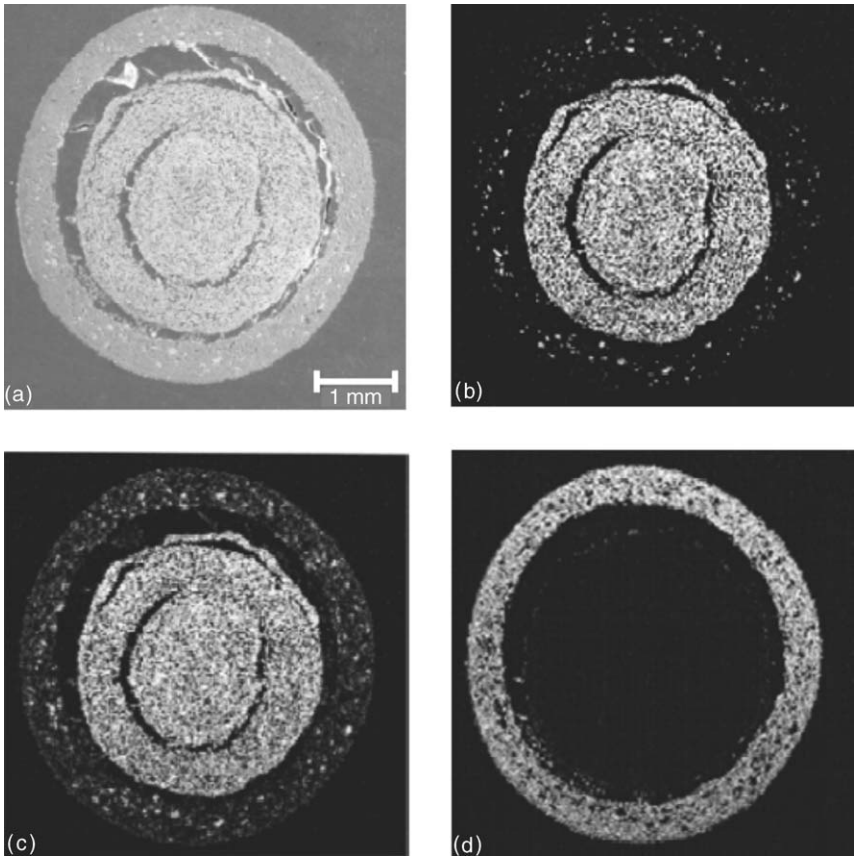


Figure 3 Micrograph of a cross section of a sulfided pellet used as a coal-gas sorbent. (a) SEM view, (b) sulfur map, (c) calcium map, and (d) aluminum map. (K Constant, unpublished work.)

Eco-ceramics Environmentally friendly ceramics usually refer to those that undergo processing using environmentally benign solvents or resulting in harmless by-products. NASA recently reported the development of a method to make a SiC ceramic using wood as the preform to supply both the structure and the carbon for the process. In this technique, wood is pyrolyzed to form a carbon structure that is then infiltrated with silicon or a silicon alloy. The result is a SiC-based ceramic with a microstructure imposed by the original wood and a shape that duplicates that of the pyrolyzed form. The structure (and therefore, properties) is tailorable by the choice of wood and can be further adjusted by laminating wood layers in various arrangements. This technique has been used to produce porous and nonporous structures of various compositions based on SiC. This technique could be considered akin to biomimicry, which is discussed in the next section.

Trends in Ceramic Materials

The most notable general trends in ceramics are those of the entire materials industry, and perhaps science and engineering as a whole. The first includes an ever-aggressive shrinking scale of our technological devices – requiring a more complete understanding of how material properties vary for very thin films and small grains (sometimes composed of fewer than 100 atoms). The second includes a renewed interest and focus on biological systems, which can serve as inspiration for the structure, as manufacturing aids, or even as end-users of ceramic materials. Third, a renewed commitment to national security and aggressive pursuit of technologies to support these efforts have resulted in increased attention on sensors that can detect any number of chemical and biological hazards quickly, effectively, and efficiently. Previous examples of properties and processing have already exposed these trends. The ceramic armor already discussed, photonic bandgap materials, and sensors for detection of biological agents, and micromechanical devices all have potential use in national security functions.

Nanotechnology/MEMS

As the scale of electronic devices shrinks, so has the scale of both the starting materials and the components of such devices. The trend toward nanoscale control has impacted ceramics as it has most fields of engineering and science. In a like manner, mechanical devices have continued to shrink to smaller and smaller dimensions, resulting in what is termed micro-electromechanical systems (MEMS). Ceramics are particularly well suited for MEMS applications in biological and chemical sensing systems (they tend to be chemically inert, even at high-temperatures and severe environments). The advantages they bring to full-scale mechanical devices (e.g., engines) are also present at the microscale. Low density and high-temperature capability make ceramics an attractive choice for high-speed devices, such as nanoturbines.

Ceramic MEMS have been made by microcasting a liquid ceramic precursor (e.g., polysilazanes) in a mold produced from photoresist by conventional lithographic methods (**Figure 4**). The resulting form is then pyrolyzed to convert to a Si–C–N amorphous ceramic. A second method of producing MEMS involves photopolymerization, a technique becoming increasingly widespread in the rapid prototyping industry. The technique involves adding a photoinitiator to a liquid ceramic precursor that will promote polymerization upon exposure to light of a specific wavelength. Shapes are constructed by exposing a layer of the precursor through a mask that has the desired shape. 3D structures are made by successive exposure to another layer, which has either the same, or a different geometry depending upon the desired result. Exposure can also be accomplished through the "writing" of a directed laser beam. The exposed material polymerizes, and solidifies, allowing the remaining unexposed liquid to be washed away. In this way, a 3D structure can be created. The resolution achievable by such efforts is related to several factors both of the chemistry of the precursor and photoinitiator and the physical characteristics of the mask or direct-write system. A similar approach to direct writing of structures is found in computer-controlled ink-jet printing of ceramic parts that is limited only by the length scale of the droplet. This technique relies on a colloidal, gel-based ink that forms a self-supporting structure. This technique, called robocasting has produced features as small as 100 μm, and has been made out of silica, alumina, lead-zirconate-titanate, and hydroxyapatite.

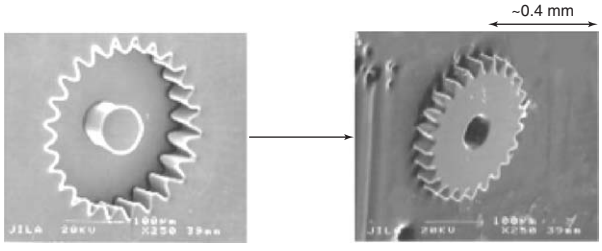


Figure 4 Left: a photoresist mold about 200 μm across. Right: The fired ceramic gear. (Reprinted with permission from The American Ceramic Society; © 2001.)

Bioceramics

Perhaps one of the most transformational trends in ceramics is that of bioceramics. Biology can take a variety of roles in the development of advanced ceramics from serving as the inspiration (or more directly, the template) for novel structures to providing the motivation for the development of materials to be used in biomedical applications. Biomimicry is really nothing new – ancient people observed animals to learn the best methods to stalk prey, to identify safe foods to eat, and to predict weather changes, but until the tools to examine nature at the microscopic (and now, atomic) level were developed, few efforts were made to copy nature at these length scales. Today, however, many scientists and engineers are again carefully studying natural materials striving to glean their secrets. Examples include the study of fracture toughness and interfacial design of the fiber-matrix ceramic composite that constitutes sea-urchin teeth and the self-assembled opal structure used as a template for photonic bandgap materials. Ceramics for biomedical applications have also become increasingly important in improving the quality of life for many of the world's aging. Ceramics can be bioinert (such as alumina or zirconia) or bioactive (including hydroxyapatite, and bioactive glass and glass ceramics). Ceramics have been used or are being developed for use in replacement joints for hips and knees, replacements for diseased or damaged bone, and reconstruction of teeth and jawbones. Ceramic and carbon coatings are also used in a variety of biomedical devices including prosthetic heart valves. The bonding between hard and soft tissue is being studied to improve such applications. Ceramics are also being used for controlled drug delivery. For example, therapeutic treatment of cancer has been achieved using glass beads doped with radioactive isotopes. In addition, in the drug-delivery arena, rapid prototyping, or 3D printing techniques have been used to fabricate specialized drug-delivery devices. These ceramics are functionally graded structures that can yield complex release profiles in which there is control time, dosage, and even the chemistry of the drug.

Biosensors

Diamond, because of its electrical and chemical properties, is a good candidate for integrated sensing materials, and recently, selective biological modification and adsorption has been achieved. In this work, DNA oligonucleotides were covalently bonded to nanocrystalline diamond thin films resulting in an extremely stable, highly selective platform for subsequent surface hybridization processes. This integration of DNA and other biological materials with microelectronics has broad-ranging implications for the development of completely integrated bioelectronic sensing systems.

Concluding Remarks

While these examples testify to the extent and breadth of the research and development in advanced ceramics, there remains much to be understood and accomplished. The ever-expanding variety of analytical tools will provide the technology to further probe the structure and properties of ceramic materials, often at the nanoscale. The blurring of lines between ceramics, biology, and other fields of engineering will most certainly persist requiring that tomorrow's ceramists be cognizant of a broad range of fields and have command of a variety of analytical tools and techniques. While evolving, ceramics will continue to have a large role in the shaping of our society.

See also: Biomedical Materials; Carbon Materials, Electronic States of; Composites: Overview; Glasses; Optical Properties of Materials; Photonic Bandgap Materials, Electronic States of; Powder Processing: Models and Simulations; Ruthenates; Vapor Transport Processing: Mechanisms; Vapor Transport Processing: Models and Simulations.

PACS: 81.05. – t; 81.15. – z; 81.16. – c

Further Reading

Ball P (1997) *Made to Measure: New Materials for the 21st Century.* Princeton, NJ: Princeton University Press.

Barsoum MB (1997) *Fundamentals of Ceramics*. New York: McGraw-Hill.

Chiang YM, Birnie DP, and Kingery WD (1997) *Physical Ceramics*. New York: Wiley.

Hench LL (1998) Bioceramics. *Journal of American Ceramics Society* 81(7): 1705–1728.

http://www.afrlhorizons.com/Briefs/Sept03/ML0303.html (transparent armor).

Leung WY, Kang H, Constant K, Cann D, Kim C-H, Biswas R, Sigalas MM, and Ho K-M (2003) Fabrication of photonic bandgap crystals using microtransfer molded templates. *JAP* 93: 5866.

Liew L, Zhang W, An L, Shah S, Luo R, Liu Y, Cross T, Dunn M, Bright V, Daily J, Raj R, and Anseth K (2001) Ceramic MEMS. *American Ceramic Society Bulletin* 80(5): 25.

May PW (2000) Diamond thin films: a 21st-century material. *Philosophical Transactions of the Royal Society Of London Series A* 358: 473–495.

Minh N (2003) Recent advances in solid oxide fuel cell technology – SOFCs for power generation applications. *American Ceramic Society Bulletin* 82(7): 12, July 2003.

Rahaman MN (2003) *Ceramic Processing and Sintering*, 2nd edn. Dekker.
Reed JS (1995) *Principles of Ceramics Processing*, 2nd edn: Wiley.
Rice and Roy W (2003) *Ceramic Fabrication Technology*. New York: Dekker.
Singh M, Martinez-Fernandez J, and deArellano-Lopez AR (2003) Environmentally conscious ceramics (eco-ceramics) from natural wood precursors. *Current Opinion in Solid State and Materials Science* 7(3): 247–254.
Smay JE, Gratson GM, Shepherd RF, Cesarano J, and Lewis JA (2002) Directed colloidal assembly of 3D periodic structures. *Advanced Materials* 14(18): 1279–1283.
Tian J, Wang X, Zhang C, and Guo Z (2002) Biomimetic artificial bone design and manufacture. *American Ceramics Society Bulletin* 81(12): 39–41.
Wang R (1998) Fracture toughness and interfacial design of a biological fiber-matrix ceramic composite in sea urchin teeth. *Journal of the American Ceramic Society* 81(4): 1037–1077.
Yand W, Auciello O, Butler JE, Wei-Cai, Carlisle JA, *et al.* (2002) DNA-modified nanocrystalline diamond thin-films as stable, biologically active substrates. *Nature Materials* 1: 253–257.

Ceramics *See* Ceramic Materials.

Ceramics, History of

I Yasui, University of Tokyo, Tokyo, Japan

Introduction

A ceramic is defined as a nonmetallic inorganic solid, but in a narrow sense it is a massive solid obtained by the heat treatment process of inorganic powders.

Ceramics have a long history. One of the first materials used by humans was a stone implement. Obsidian in composition, such a material can be classified as being glassy. Similarly, porcelain and earthenware also have a long history. The same can be said of glass too; human beings produced glassy solids in 3000 BC. Ceramic powder was also used as an adhesive at the time of the construction of the pyramids.

Ceramics are classified, from the viewpoint of structure, as: (1) single crystal, (2) sintering object, (3) glass, and (4) powder.

1. The technology to produce single crystals artificially is difficult; moreover, it is not suitable for mass production. Commercial ceramic products in the form of single crystals are, therefore, rather new historically. At present, some of the single crystals such as a quartz oscillation unit and an optical crystal are of much importance, because it is impossible to obtain the desired properties by using materials other than single crystals.
2. Sintered bodies are obtained by a forming process where starting powders are pressurized and the massive solid is obtained by subsequent heat treatment. Although earthenware was the origin of this technique, ceramics with various functionalities were developed.
3. Glass is a material obtained through a melting state by heating starting raw powders at a high enough temperature. Since the composition of common glasses was similar to that of earth, it was made artificially as early as in 3000 BC. However, these early glasses were not transparent. Although glass as a transparent material was developed later, it was only in the twentieth century that the technology of manufacturing sheets of glass with larger area became available.
4. Powders have been used from the BC period by grinding natural ores. The history of pigments also goes back a long way. Powder today means such materials as can be used in making high-performance sintered ceramic bodies. Original materials for ceramics were natural minerals, especially silicates. Clay, which is a natural silica mineral and very suitable for the production of porcelains, was the primary raw material for a long time. It is, therefore, understandable that the industry which used clay emerged first. The ceramic industry showed a marked progress after materials other than silicates came into use. Alumina is the first example and titanium oxide as a dielectric followed in industrial use. These simple oxides offered an easy control of the properties of their powders, suggesting that substances with easy handling became raw materials for practical use.

Silicate industries have also undergone constant development. The wall thickness of ordinary china used for tableware is ~ 1 mm; $\sim 50\,\mu$m thick silicate ceramics are manufactured in large quantities as catalyst carriers for cars (in the absence of this technology, automobiles would perhaps be contributing to air pollution at levels even higher than they do now).

As a next step, complex oxides constituted the mainstream of the practical use of ceramics. The first generation of complex oxide ceramics are barium titanates and ferrites. These substances came to be used, due to a great deal of development in the management of purity, fine-tuning of composition, surface modification, etc. For example, the control of grain boundaries is one of the most important parameters that determines the total characteristics of a ceramic material. New types of ceramic heaters, which are capable of self-control of temperature with a composition of barium titanates, were developed and used. These are characterized by controlled grain boundaries and the so-called positive temperature coefficient (PTC).

Most industrial applications of ceramics exploit the characteristics of the substrates on the one hand and the electronic or ionic properties of their constituents on the other.

Nowadays, the most useful example of functional ceramics are lead zirconium titanates (PZTs). For instance, cellular phone designs are materialized by using these piezoelectric ceramics. Barium titanates and ferrites, sometimes called electronic ceramics, are extensively used in the electronics industry. This is because the electronic properties play a crucial role in determining the function of these ceramics.

Simultaneously, materials using the ionic properties of ceramics have also emerged and are used practically. The purification of the exhaust gas of a car is attained due to the use of zirconium oxide ceramics as an oxygen sensor, thus utilizing the ionic conductivity of oxide ions. It is called a solid electrolyte because the substance has conductivity due to an ionic motion within a solid. One of the notable examples is beta alumina, which is now used as an electrolyte for an Na–S battery, a large-sized rechargeable battery for load leveling of electric power.

A compositional survey of the possible properties of oxides containing two or three metals has already been conducted. Although additive agents are used to promote sintering or optimization of properties, compounds which principally contain more than four kinds of metals are seldom examined. In this unexplored area, a new category of ceramic materials has emerged, that is, superconducting ceramics.

It is to be noted that the most important characteristic of ceramics is their tolerance to high temperature. For many years, oxides have been used as refractory materials. It is quite clear that "refraction" supports the basic technology in the metal-refining industry. By extending the function of an oxide as a high-temperature material to the extreme, its application in the area of energy production etc., can be expected. However, it was not oxides, but nitrides and carbides, which attracted attention as such materials in the 1980s. As a result, a ceramic turbocharger was developed for automobiles. Although non-water-cooled engines made from ceramics were proposed to be used, this has not yet been realized.

The glass industry at present is very large in scale. This is one of the several reasons for natural silica to be the major raw material in the industry. An example of the application of glass, which uses transparency to its extreme, is the optical fiber. An optical fiber is made from very high purity silica although it includes a very small quantity of additives to control the contour of refractivity.

The composition of the often used glasses resembles that of sheet glass or bottles. The insulated film used in semiconductor devices is also glassy. Although only glassy material with a simple composition can be used in device processes, one can expect new compositions to find future applications.

Cement is a material which has been used in large quantities for a long time. Of course, it is also a product of the silicate industries. The use of a material in such large quantities is bound to pose a problem of depletion of resources, sooner or later.

History of Development of Ceramics

In the preceding section, an outline of ceramics was given from the historical point of view. The present section selects some ceramics as examples to consider the development process in more detail.

Large-Sized Insulator and Ceramic Carrier

The largest ceramic (1.2 m $\varnothing \times 12$ m length) ever made was the insulator for 1000 kV power transmission. In consideration of it being corrosion-proof and earthquake-resistant, it had a special form of a bamboo hat and was made of large-sized porcelain with a thick wall.

The catalyst system was established as a primary means of automobile exhaust gas purification. Although started in the form of pellets, the monolithic honeycomb-type catalyst carrier was developed in 1976. This type of catalyst opened up the possibility of stricter emission control of automobiles.

Cordierite, with a composition of $2MgO \cdot Al_2O_3 \cdot 5SiO_2$, was used because of excellent heat resistance due to a low expansion coefficient. In the initial stage, the number of cells was $\sim$200/square inches (1 inch = 2.54 cm) but later reached 900 square inches, and the wall thickness decreased from 12 mil (1 mil = 0.0254 mm) to 2 mil.

Alumina (Aluminum Oxide) and Substrate Material

Some aluminum oxides have been used for many years as refractory materials. Examples range from highly pure alumina crucibles to low-purity alumina pipes or plates for low-heat-resistance applications. Their application as an insulator for spark plugs was developed in 1931, and alumina started being used in automobile and electronics industries.

The decisive extensive use of ceramic materials in electronic industries occurs in the form of substrate materials used as a part of large-scale integrated circuits. Alumina substrates for electronic circuits cover a wide range of products, such as those made by the thick-film process, the thin-film process, and multilayer-type circuit boards. Alumina powders with purity higher than 99.5%, with 0.2% of magnesium oxide as additives, are generally used to give a flat and smooth surface, which is excellent for use as substrates.

The multilayer aluminum circuit board was first developed by IBM for the CPU used in the 308 X series mainframe computers. The CPU was $90 \times 90\,mm^2$ in size and the number of layers was 33. It was evaluated that this circuit board had superior properties both in the delay time of signal propagation and heat dissipation. Silver, gold, copper, etc., were used as conductive materials for wiring.

Because of the success of the alumina circuit board for the CPU, other substrate materials with higher heat dissipation capability, such as in the AlN and SiC boards, were investigated. Since AlN and SiC have simple crystal structures, both materials have excellent thermal conductivity, provided the grain boundaries are controlled. Since AlN is an insulator, it is only necessary to control the grain boundaries to obtain a high heat conduction, but since SiC has electronic conductivity, it is inevitable to control the grain boundaries so that it is electrically insulating. It was found that the addition of BeO as a sintering agent changed ceramics to insulating substrates.

Although ceramic substrates for large-scale integrated circuits were excellent in performance, the production cost was a prohibitive factor for use of this product. The semiconductor packaging process was changed, and it was found that even plastic could be used as a substrate. If ceramic substrates continue to be used in the electronics industries, they will be limited to high-performance type of products, such as single crystals or diamond.

Development of Ceramic Capacitors

Natural materials such as mica have historically been used in capacitors for a long time. Electrolyte capacitors use an aluminum oxide thin film as a dielectric. Several other oxides have been used as capacitor materials for many years.

It is well-known that barium titanate has a high dielectric constant (several 1000 or more). This ceramic was independently developed and studied in USA, Japan, and the Soviet Union. The abnormalities of the dielectric constant of barium titanate were discovered by Wainer and Salomon in 1942, which was an important breakthrough in the application of ceramics after World War II. Subsequently, Vul and co-workers, von Hippel, and Ogawa and Waku studied the ferroelectric properties of barium titanate.

However, from the point of its practical use, it is important to note that it took a remarkably long time from the discovery of barium titanate as a substance to its development as a ceramic capacitor. It was actually in the 1990s that the ceramic capacitor with a multilayer-type structure using barium titanate came to be loaded into almost all electric devices. This is because it took time to develop the process of manufacturing multilayer ceramic capacitors on a large scale. A process using the doctor blade method is first used to form thick films, and it is necessary to carry out the lamination of films. The technical developments, other than that of a ceramic layer, were also important. The development of conductive layers using nickel for an internal electrode was also a key factor in the development of the technology.

History of Piezoelectric Ceramics

The piezoelectric effect was discovered by Jack and Pierre Curie in 1880. Materials that display the piezoelectric effect need to be in the form of single crystals. Therefore, in order to use this effect for practical purposes, single crystals of appropriate materials were produced industrially. In Japan, Toyo Communication Equipment industrialized the production of quartz by the hydrothermal method in 1959.

Other materials were also studied for use as piezoelectric ceramics. Since it was reported by Bell Laboratories in 1963 that the thin film of CdS worked as a piezoelectric transducer, development of piezoelectric thin films and thick films started. The development of the thick-film process and materials that were more suitable for mass production attracted much attention. PZT, that is, Pb (Zr, Ti) O_3, discovered in the late 1950s, became the primary candidate for practical applications. A phase transition occurs when the Zr/Ti ratio is 53/47. In the vicinity of the phase boundary with this composition, the dielectric constant and the piezoelectric constant become the maximum.

In order to make a piezoelectric body of this composition, materials with random orientation cannot be used. Therefore, polarization processing was indispensable during sintering and the related process itself had to be developed. A number of sheets must be laminated as a layer, especially in the case of an actuator; thus, the manufacturing technology acquires key importance.

In Europe, a restriction in the use of lead in electronic products will be in effect from the year 2006. PZT, of course, contains lead. It does not necessarily mean that this material will be prohibited from use in Europe soon, but it is also possible, that if a new material for replacement is developed, PZT will then be forbidden. Therefore, development of piezoelectric materials which do not contain lead is now in severe competition.

Ion Conductivity Ceramics

It has been known for many years that glass and other ceramics show an alkali ion conduction. For example, a group of substances called beta alumina, containing Na as a component (in spite of being known as alumina), have been known as Na ionic conductors for many years. Their crystal structures were analyzed in the 1950s.

The Na–S battery, which uses beta alumina as an electrolyte, was proposed as a large-sized rechargeable battery for load leveling after the oil crisis of the 1970s. The Na–S battery attained the level of practical use at last by a joint development of NGK Insulators and the Tokyo Electric Power Co., (the delivery track record in 2002 was set at 5000 kW).

Although historically, ZrO_2 has been used due to its very high chemical durability at very high temperature, it was already known at the end of the nineteenth century that ZrO_2, with the addition of CaO or Y_2O_3, has oxygen ion conductivity. This material attracted attention after its use as an oxygen sensor for automobile exhaust gas purification was proposed. As an oxygen sensor, ZrO_2 plays the role of a solid electrolyte similar to that in fuel cells. A solid electrolyte-type fuel cell, on the other hand, has not yet reached a stage where it can be used in practical applications.

Development of Ceramic Sensors

It must be noted that a ceramic material used as a chemical sensor brings about an improvement in the performance of various electrical appliances. The range of home electrical appliances which were originally apparatuses that could only detect temperature, now also include appliances (such as air conditioners, microwave ovens, or refrigerators) which can detect humidity, carbon dioxide, infrared radiation, etc. This development has been required from the viewpoint of increasing energy efficiency in relation to global warming and other environmental issues. In addition, detection of flammable gas, such as hydrogen, methane, and a liquefied petroleum gas, is indispensable because of concerns of safety. A gas sensor that made use of SnO_2 was developed. The principle of a gas sensor itself is rather conventional. The ZrO_2 sensor already mentioned is the so-called concentration cell type, and the principle is also classic.

The SnO_2-type gas sensor utilizes the characteristics of compounds as semiconductors, along with the capability of gas adsorption. It has, therefore, been an important target in its development as to how the most suitable textures and microstructures could be formed in ceramics to enhance the properties of adsorption of thick-film materials.

There are several ways for a modified usage of ion-conducting ceramics to detect a certain gas. One is to use ion-conducting ceramics together with the auxiliary electrode, which consists of the conduction ion and detection gas components. For example, in the case of the carbon dioxide sensor using beta alumina, $NaCO_3$ (containing Na^+, the conduction ion of beta alumina) and the carbon dioxide of the detection gas will be used as the auxiliary electrode. Many approaches for development range between such theoretical understanding and the practical design of a sensor (see section “Further reading”).

Ferrite Materials and Magnetic Recording

A ferrite, which is a magnetic oxide body, was invented by Takei and Kato in the 1930s. Ferrites are widely used as a core material for a transformer or magnetic heads. The oxide permanent magnet was also invented during almost the same period, and after that, it has been used as a low-cost permanent magnet material. Iron oxide ceramic powder is used as a magnetic recording material in videotapes such as VHS. A record system may change to a perpendicular type and oxide powder may be used again, although currently metallic materials are used for the purpose of high-density recording.

Ceramic Turbocharger and Ceramic Turbine

Silicon nitride and silicon carbide, which are the main components of nonoxide ceramics, are artificial compounds. Since these compounds have a strong covalent bond in nature, till 1960 it was accepted that they could not be sintered.

The first oil shock pushed the development of efficient energy devices, and the use of ceramic materials, which can be employed at much higher

temperatures, was pursued. It was proposed in the 1980s that nonoxide ceramics could be used as parts of a gas turbine. The ceramic gas turbine project (CGT) was actually undertaken from 1988 to 1999 in Japan. A CGT with 300 kW power for co-generation and a 100 kW CGT for automobile use were developed. However, it has not reached the stage of commercial production, mainly because of production costs.

Ceramic turbocharger turbine blades made of silicon nitride are an example of a practical application of a nonoxide. Ironically, the material originally developed for raising energy efficiency was used in the automobile industry, which aimed at producing fuel-inefficient cars. However, the ripple effect was large. Nonoxide ceramics now have several applications, which are growing steadily.

Superconducting Ceramic Materials

In 1986, an oxide superconductor of the lanthanum–barium–copper system was first announced by Bednorz and Muller of the IBM Zurich Research Institute and its critical temperature T_c was 35 K, which exceeded the historical T_c 23 K of the NbGe. The Y–Ba–Cu system oxide was discovered in the following year and the critical temperature exceeded 77 K, which is the temperature of liquid nitrogen. Then, various new compound oxides, such as the Hg system, the TI system, and the Bi system, were announced.

A F Hebard and others have reported that the powder of C_{60} with doped potassium (K_3 C_{60}) shows a superconducting nature at a critical temperature of $T_c = 18$ K, and the same thin film of K_3 C_{60} shows the superconducting nature at a critical temperature of $T_c = 16$ K. MgB_2 has recently been reported to become superconducting at 39 K (for further details, see section "Further reading"). Thus, superconductivity is serving as a frontier of the science of new compounds in the ceramics system.

Optical Fibers

In 1964, Nishizawa succeeded in developing graded-index-type optical fibers for communications.

In 1966, Kao found that if heavy metals contained in glass were removed, the absorption loss of light in glass decreased dramatically. This epoch-making discovery was the key to utilizing an optical fiber and one that enabled the material development of low-loss fibers. Kao also had the foresight of proposing the structure of today's optical fibers consisting of a clad with a low refractive index and a core with a high a refractive index glass.

Following the basic work of Nishizawa and Kao, Corning Inc. succeeded in making low-loss (20 $dB\,km^{-1}$) fibers. Subsequently, the basic technology of the present-day information technology grew through a stiff competition in the development of optical fibers.

Conclusion

Ceramics are practically useful materials. From their discovery, through application development and testing (the so-called seed-driven route), ceramics have come a long way. However, new materials with superior properties continue to be desired in a variety of application areas. Given the limitations of the global environment, exploring new ceramic materials is a real challenge for further sustainability.

See also: Ceramic Materials.

PACS: 77.84.Dy; 81.05.Je

Further Reading

Bednorz JG and Muller KA (1986) *Zeitschrift fur Physik B* 64: 189.
Burger WG and Weigei CW (1982) *IBM Journal of Research and Development* 297.
Hebard AF, *et al.* (1991) *Nature* 350: 600.
Nagamatsu J, Nakagawa N, Muranaka T, Zenitani Y, and Akimitsu J (2001) Superconductivity at 39 K in MgB_2. *Nature* 410: 63.
Ogawa T (1947) Physical-properties theory research. No. 6, 1.
Tsuji S, Mizuishi K, Nakayama Y, Shimaoka M, and Hirao M (1983) *Japanese Journal of Applied Physics* 22: 239–242.
von Hippel, Breckenridge RG, Chesley FG, and Tisza L (1946) *Ind. Eng. Chem.* 38: 1097.
Vul M and Goldman LM (1945) *C. R. USSR* 46: 139.
Wainer and Salomon AN (1942) *Electronic Reports* 8 – Titanium Alloy Mfg. Co.
Wu MK, Ashuburn JR, Torng CJ, Hor PH, Meng RL, *et al.* (1987) *Physical Review Letters* 58: 908.
Yamazoe N and Miura N (1994) *Sensors and Actuators B* 20: 95–102.

Composites: Carbon Matrix

B Rand, University of Leeds, Leeds, UK

Introduction

Carbon matrix composites (or carbon–carbon composites) are advanced composites in which carbon fibers are used to reinforce a carbon matrix. Both the fibers and matrices may encompass a vast range of crystallinity and the composites themselves may be constructed from a variety of different fiber architectures and volume fractions such that this class of composites may display a huge diversity of properties. These are the most refractory of all composites (in nonoxidizing conditions), exhibiting exceptional dimensional stability in extremes of temperature, high toughness and mechanical reliability, specific thermal conductivity many times that of copper, and remarkable wear resistance at high temperatures and speeds. These properties lead to such applications as:

- furnace parts, rocket nozzles, fusion reactor linings,
- construction materials for space applications,
- components in electronic packaging,
- thermal management,
- aircraft and formula one brake linings, and
- first wall linings in fusion reactors.

The origin of these properties lies in the intrinsic properties of the graphite crystal, since the various carbon forms that are developed in the composites display structures and properties that resemble those of graphite to different degrees.

Graphite Properties

The structure of graphite is well known, comprising graphene layers stacked in *abab* (hexagonal) or, rarely, *abcabc* (rhombohedral) sequence. The carbon atoms in the graphite crystal are sp^2 hybridized, forming three strong covalent bonds within the graphene layer whilst the layers are held together by van der Waals type interaction. Thus, the crystal displays the strongest and the weakest bonding possible in its two principal crystallographic directions, resulting in the most extreme anisotropy in physical and chemical properties. The principal physical properties are listed in **Table 1**. The strong "in-plane" bonding results in a Young modulus greater than 1000 GPa and a thermal expansion coefficient that is negative at room temperature and becomes slightly positive above $\sim 150^{\circ}$C. Although there is an electronic contribution at low temperatures, thermal transport is largely by a phonon mechanism above room temperature and reaches values comparable with those of diamond in the "*a*" direction. Delocalization of the π electrons makes graphite a semimetal with a small-band overlap. The "*a*"-axis electrical resistivity is 0.4–0.8 $\mu\Omega$ m at room temperature. In the "*c*"-direction most of the properties are significantly lower, as shown in the table, but the "*c*"-axis thermal expansion coefficient is large due to the weak bonding in this direction. The development of engineering carbon and graphite materials requires careful control of this anisotropy (**Figure 1**).

Carbon Fibers and Matrices – Structures and Properties

All synthetic carbon forms are produced by the thermal degradation in an inert atmosphere of organic precursors, gas, solid, or liquid (carbonization). The structure of the resultant carbon is determined by the nature of the precursor and the thermal history. The essential structural units are defective graphene layers, limited usually to nanometer dimensions, and stacked imperfectly with no regular stacking sequence, **Figure 2**. The precursors fall into two fairly well-defined groups depending upon whether they develop into synthetic graphite with the hexagonal *abab* stacking of graphite when heat treated to temperatures $\sim 3000^{\circ}$C. Those which do are graphitizing carbons whilst those which show relatively little

Table 1 Major physical properties of crystalline graphite

Property		*"a"-axis*	*"c"-axis*
Density (kgm^{-3})	2.265×10^3		
Young modulus (GPa)		1060 ± 20	36.5 ± 1
Shear modulus (GPa)		4.5 ± 0.5	
Thermal conductivity (Wm^{-1}K^{-1})		>2000	<10
Electrical resistivity ($\mu\Omega$ m)		0.4–0.8	>100
Coefficient of thermal expansion (K^{-1})		$\sim -1 \times 10^{-6}$ at room temperature	25.6×10^{-6} at room temperature

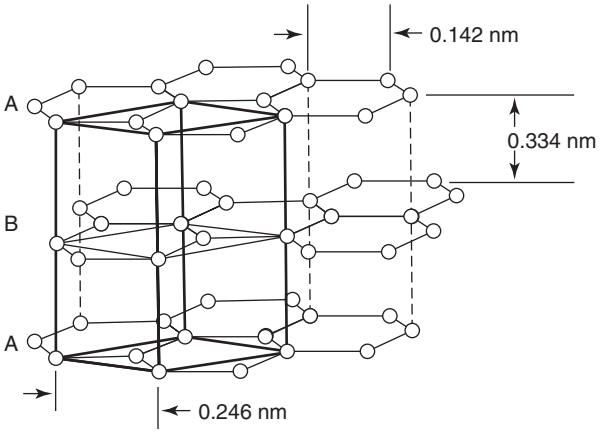


Figure 1 The structure of an ideal crystalline graphite.

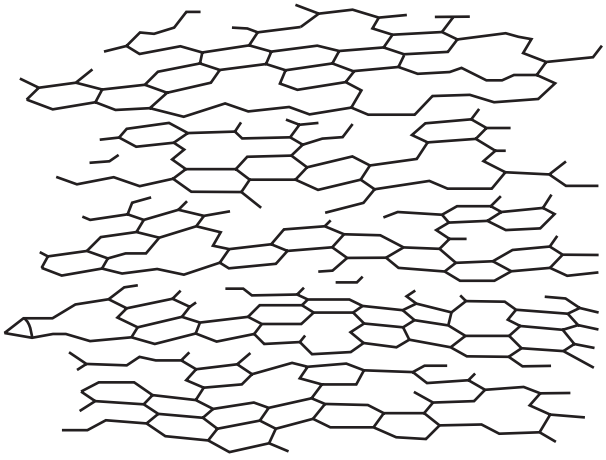


Figure 2 Schematic representation of the structure of nongraphitic carbons, showing defective and bent layers, defects and approximate layer stacking resulting in limited two-dimensional ordering.

structural development following such drastic thermal stimulus are the nongraphitizing carbons. Carbon fibers from rayon and from polyacrylonitrile (PAN) fall into the latter group (even though in the commercial world they are frequently and incorrectly referred to as graphite fibers), whilst those from mesophase pitch and hydrocarbon gases are of the graphitizing variety. In low-temperature carbons, the defective graphene layers may contain vacancies, and have five- and seven-membered rings which bend the layers in different directions. Hence it is not possible for the layers to stack in any regular array and the interlayer spacing, d_{002} is significantly greater than that in graphite. Only very broad 001 and *hk* X-ray diffraction lines are observed. A line-broadening analysis, using the Scherrer equation, enables X-ray coherence lengths in the "*a*" and "*c*" directions (L_a and L_c) to be identified from the 002 and 10 diffraction peaks. These are initially of the order of 1–2 nm, but during graphitization they increase progressively with heat treatment temperature (HTT) up to about 50–100 nm, when the additional *hkl* lines appear, indicating that a 3D structure has been formed. Simultaneously, there is steady decrease in d_{002} to the value for graphite.

X-ray diffraction gives no information about the spatial distribution of these coherently diffracting regions. High-resolution transmission electron microscopy (HRTEM), **Figure 3**, shows that they may be randomly arranged as in the nongraphitizing carbons or show extensive regions of preferred orientation in the graphitizing materials.

Matrix Carbons

The most common matrix material is chemically deposited from hydrocarbon vapors, CVD. When the deposition takes place in the interstices of the fiber preform onto the surfaces of the fibers, it is known as chemical vapor infiltration (CVI). The structure of the deposit is very dependent upon the deposition conditions, temperature, partial pressure, and residence time. Various different microstructures are possible which are classified as isotropic, laminar, rough laminar, and surface nucleated. Infiltration conditions must be carefully controlled not only to deposit the desired microstructure for the particular application, but also to ensure a uniform deposition throughout the fiber architecture and avoid preferential deposition on the surface of the fiber preform. This requires a careful balance between the rate of deposition and of vapor transport into the porous fiber preform. The infiltration process is slow and the manufacture of large items may take several weeks. It may also be necessary to machine the outer surface intermittently to remove the surface deposit which seals off the interior. The product is graphitizable. For aircraft brake linings, the preferred microstructure is the rough laminar variety which seems to give the optimum combination of thermomechanical properties and wear resistance. Since the deposition takes place onto the fiber surfaces, the graphene layers show a preferred orientation with the fibers, **Figures 4** and **5**.

Alternative matrix sources are pitches and thermosetting polymeric resins. Pitches are complex mixtures of polynuclear aromatic molecules, many having aliphatic side chains. They are thermoplastic and can have low viscosities at temperatures ~100–150°C above the glass transition temperature, T_g. Heat treatment increases the average molecular weight by evaporating volatile species, resulting in a gradual shift in the T_g to higher temperatures and

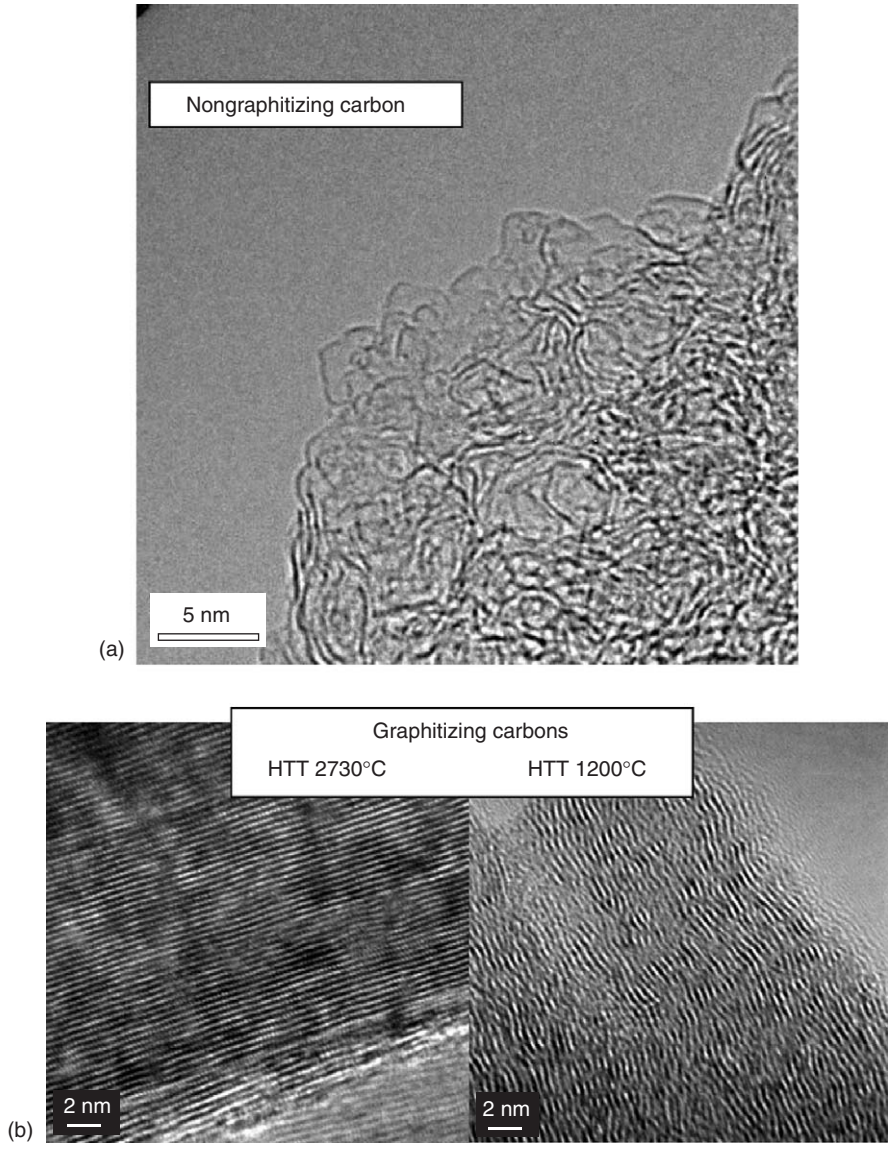


Figure 3 High-resolution transmission electron micrographs showing 002 lattice fringes for (a) a nongraphitizing carbon and (b) graphitizing carbons at two different heat treatment temperatures.

ultimately the solidification of the material to a carbon product. Intermediate in this process is a key step in which the disk-like molecules stack up to form a nematic discotic liquid crystalline phase, the carbonaceous mesophase. This is the basis of the graphitizability of the subsequent carbon, the basic long-range preferred orientation of lamellar structure being determined at the solidification stage of this transient phase.

Phenolic resins are the most common thermoplastic precursors, although polymers with higher carbon yields, such as poly-paraphenylenes have also been used. The resins are thermally cross-linked prior to carbonization. This prevents any liquid crystal formation and the materials do not graphitize. Normally, only minor microstructural changes take place after heat treatment to 3000°C. However, both the pitch and resin precursors may bond to the carbon fibers during the fabrication heat treatment and form an aligned interphase at the fiber–matrix interface. Thus, all three matrix precursors tend to give matrix oriented with the fibers, although in the case of the pitch and resin matrices, this may only extend a few micrometers out from the surface. **Figure 6** shows optical micrographs of typical examples.

Carbon Fibers

The main form of carbon fiber used is from PAN. A wide range of properties are available, depending

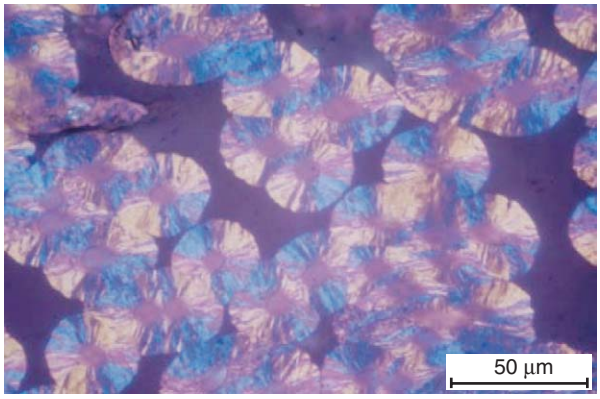


Figure 4 Polarized light optical micrograph of CVD carbon deposited onto the surfaces of carbon fibers within a carbon–carbon composite (crossed polars). The yellow and blue interference colors denote regions where the layers are oriented at 90° to each other. In this instance the layers are oriented parallel to the fiber surfaces.

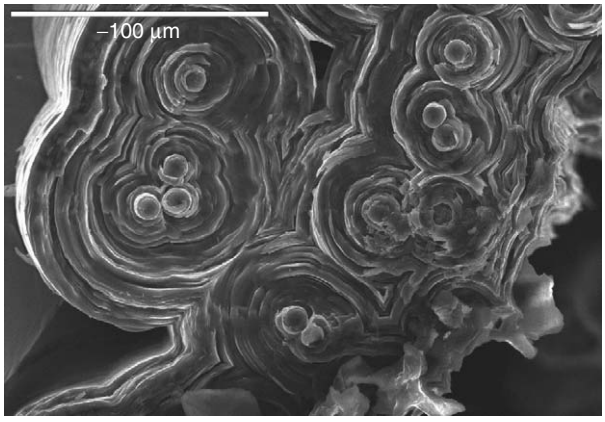


Figure 5 Scanning electron microscope image of CVD carbon layers deposited around fibers within a carbon–carbon composite.

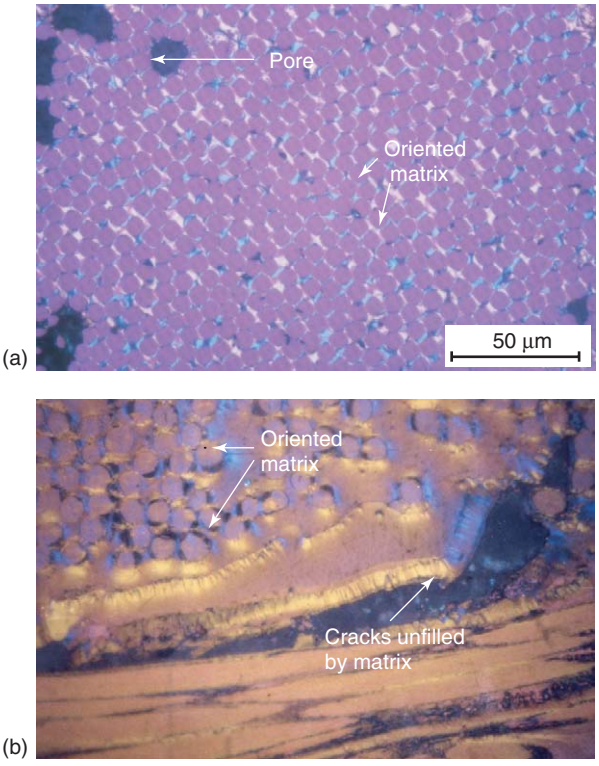


Figure 6 Preferred orientation of matrix around fibers in (a) a pitch-based matrix and (b) a resin-based matrix. Polarized light optical micrograph. Arrows point to regions of preferred orientation and also to pores and cracks.

mostly upon the degree of preferred orientation of the graphene layers along the fiber axis and the perfection of those layers, controlled by the final HTT. Young moduli up to ~500 GPa are possible with strengths in the range 2–7 GPa. **Table 2** lists typical data. Fibers are also spun from mesophase pitch and these are highly graphitizable displaying much higher extents of preferred orientation of the planes along the fiber axis and consequently higher moduli, which may approach that of the perfect graphene layer itself. These fibers also show very high thermal and electrical conductivities. Although in all cases the graphene layers are aligned predominantly along the fiber axis, transversely they may show various textures, whilst still maintaining the high stiffness. This effect is most strongly displayed in the range of mesophase pitch-based fibers, **Figure 7**, where the texture is largely controlled by the conditions in the spin capillary. Indeed even ribbon-shaped fibers have been produced with significant orientation of the graphene planes perpendicular to the plane of the ribbon as well as parallel to the ribbon long axis. The strengths of all carbon fibers are dominated by process-induced defects, flaws, the dimensions of which control the strength in accordance with the Griffith fracture criteria. The fiber microstructure has a secondary effect.

An early form of carbon fibers was produced from rayon in the US. This form is not so widely used worldwide, but still finds some defense applications in the US. Its demise is largely due to the fact that the PAN process is cheaper, whilst that for rayon required a stretching operation at graphitization temperatures to develop the high elastic moduli, by aligning the layers with the fiber axis. These fibers tend to have a "dog-bone" transverse shape and can therefore be recognized readily easily in micrographs.

Carbon fibers may also be produced from hydrocarbon vapors by a catalytic route; their dimensions and structure may vary considerably with the process conditions from filaments of micron diameter down to carbon nanotubes. Some of these materials are highly graphitizable and display excellent properties, but as yet are not used in the fabrication of C–C composites to any significant degree.

Table 2 Physical properties of typical carbon

Fiber type	*Fiber diameter* (μm)	*Density* ($g\,cm^{-3}$)	*Young modulus* (GPa)	*Tensile strength* (GPa)	*Elongation at failure* (%)	*Thermal conductivity* ($Wm^{-1}\,K^{-1}$)	*Electrical resistivity* (μΩ m)
PAN-based							
High strength	5.5–8.0	1.75–1.80	228–300	4.3–7.06	1.8–2.4	8–9	15–18
Intermediate modulus	5.0–5.2	1.73–1.80	230–310	3.1–7.06	1.3–2.4	~38	11–13
High modulus	5.4–8.4	1.78–1.96	350–540	1.8–3.6	0.4–0.7	64–120	6.5–10
MP-based							
Low HTT	11	1.9	140	1.4	1.0	100	8
↓							
High HTT	10	2.15	820	2.2	0.2	1100	~1.3
Rayon-based							
Heat treated		1.43–1.70	20–55	0.35–0.7			
Stretch graphitized		~2.0	~620	3.1			

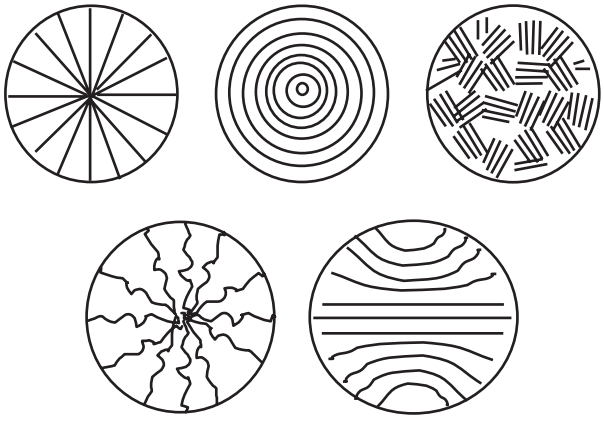
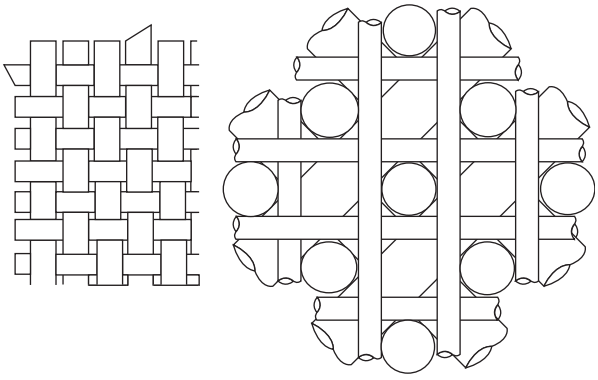


Figure 7 Schematic representation of the various transverse textures exhibited by mesophase pitch-based carbon fibers. Transverse texture is controlled during the spinning stage by the effects of shear and extension on the discotic liquid crystal domains.



Plain weave 5D woven structure

Figure 8 Schematic representation of one typical weave structure and a 5D multidirectional weave.

Fiber Architectures

Composites may be unidirectional, 1D, angle ply laminates, woven cloths, or more complicated 3-, 4-, or multi-directional woven structures, designed to deliver more or less isotropic properties. Pierced, "needled" laminated plies or cloths may be used to give improved properties in the "*z*"-direction, whilst the "in-plane" isotropy will depend upon the cloth weave or lamination sequence of angled 1D plies. General purpose applications rely on laminates of various types and the expensive multi-D structures are mainly used in sophisticated and demanding aerospace and defense applications. High-pressure infiltration and carbonization of pitch and resin precursors may be required to densify these more complicated structures. **Figure 8** shows examples of some of these structures, schematically. Angle ply laminates are commonly used in the fabrication of aircraft and formula one car brake disks, comprising not only unidirectional layers, but also random staple-fiber regions of significantly lower fiber fraction. These staple layers assist in the transport of the pyrolytic vapor into the preform.

Fabrication

Figure 9 shows typical fabrication routes for all precursor types. The fiber preform is densified by infiltration of liquid pitch or resin or by CVI. In some cases, the preform may be rigidized with resin first. In the CVI process, it is difficult to completely fill all the internal porosity since the entrances to large pores

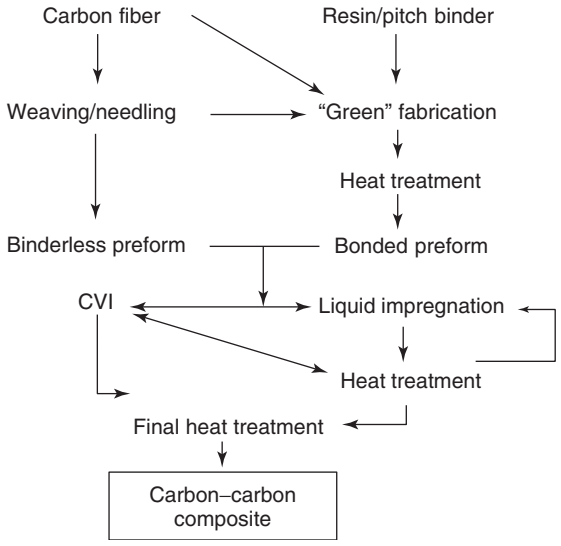


Figure 9 Generalized fabrication routes for carbon–carbon composites.

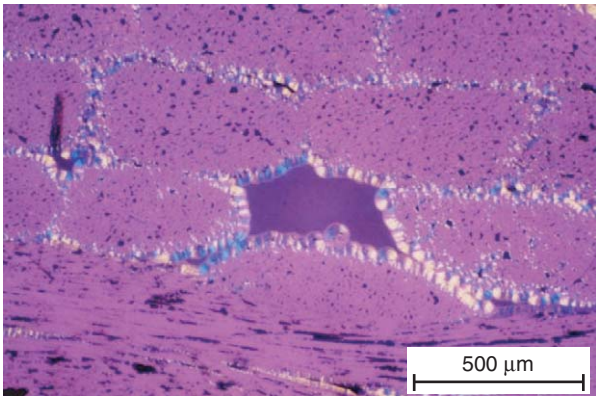


Figure 10 Polarized light optical micrograph of CVI densified C–C composite showing the tendency for the deposition of carbon to take place on the surface of bundles and not to completely penetrate the interstices within the bundle.

may be blocked by the pyrolytic deposit, **Figure 10**, and the process is slow. In the liquid infiltration route, the resins and pitches have only 50–80% carbon yields due to the loss of volatiles. Also, the density of the carbon–graphite product is significantly less than that of the precursor organic, resulting in volumetric yields that are even lower. Consequently, repeated infiltrations and re-carbonization steps are required to reduce porosity. Since graphitization of the matrix increases its density, it may even be necessary for the most advanced composites to be re-impregnated after this step and then re-graphitized, resulting in very expensive products.

Various combinations of precursors and impregnation stages may be adopted and this can result in materials with varieties of carbon structures within the matrix and in the fibers. The most extreme fabrication routes are adopted for aerospace and defense applications. For more commercial applications, such as aircraft brake disks, there are vigorous programs currently underway to develop more rapid fabrication routes.

Composite Microstructure

C–C composites are complex materials. They display structural features at many different dimensional scales all of which contribute to the observed physical properties. **Figure 11** shows some examples of various internal microstructures. At the nanoscale, there are the different degrees of structural organization in the various carbon forms (fiber, pitch, resin, or CVD-derived matrix) that may be present as well as the possibility of nanoscale porosity (especially within the carbon matrix derived from thermosetting resins). At the microscale, there is the distribution of anisotropic regions within the matrix, the fibers, and at the fiber–matrix interfaces, as well as porosity and microcracks. At the millimetric scale there are the fiber bundles (which vary in dimensions according to the fiber type and manufacturer), the laminae, large pores which have not been filled by matrix and cracks formed by shrinkage of bundles during carbonization.

Lamellar cracks develop on cooling from the fabrication temperature to relieve thermal stresses arising from the anisotropic thermal contraction of the components of the microstructure. A thermal expansion mismatch, leading to such microcracking, may exist

- at fiber–matrix interfaces within fiber bundles,
- at interfaces between carbon forms from different precursors,
- between regions of different preferred orientation within the matrix,
- between laminae in 2D materials, and
- between differently oriented bundles in multi-D composites.

Thus, cracks and voids are an integral part of the composite structure and must be taken into account in the interpretation of all physical and chemical properties. The matrix carbon increases its true density on heat treatment up to graphitization temperatures leading to contraction which also is largely in the "*c*"-direction; thus this mechanism is also a source of microcracking. However, whilst those fissures arising from the anisotropic thermal contraction may close on reheating, those from the

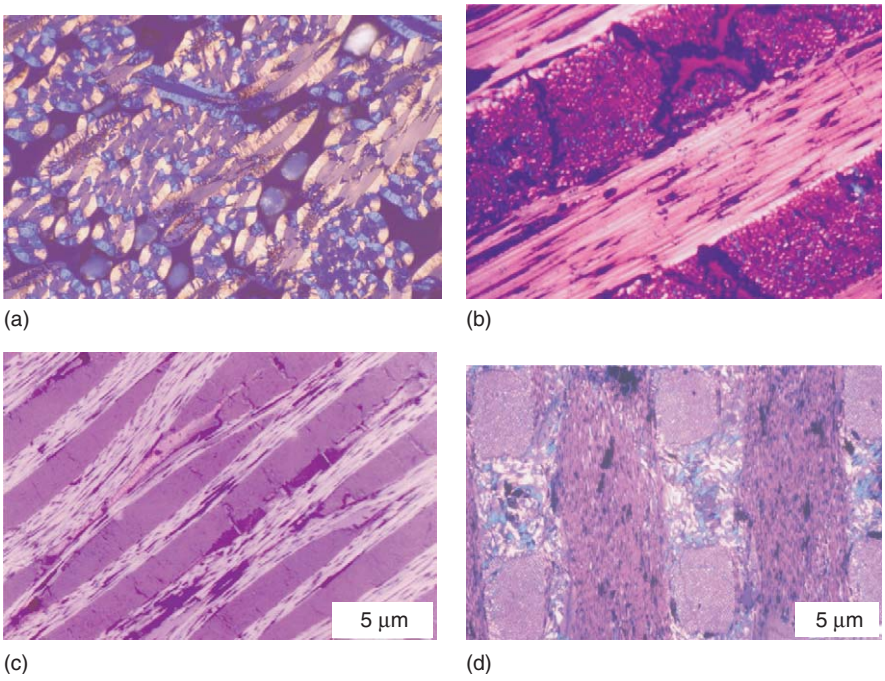


Figure 11 Optical micrographs (polarized light) showing microstructures of various types of C–C composite. (a) CVD carbon around fibers in a random fiber architecture. (b) Section of a 0/90 laminate showing shrinkage cracks in the transverse ply, partially filled with resin-based carbon. Note how the carbon in the crack has shrunk as well, displaying the shape of the surrounding pore. (c) Section through a woven cloth composite showing shrinkage cracks in the transverse bundles. (d) Section through a 3D orthogonal composite. The matrix pocket that exists between the *X*, *Y*, and *Z* bundles is in the case partially filled with pitch-based carbon matrix.

densification process will not. **Figure 11c** shows a typical example of transverse cracks in a woven cloth due to the differential shrinkage during carbonization of fiber bundles oriented perpendicular to each other and **Figure 12** is a schematic representation of the microcracks in an orthogonal 3D composite.

Physical Properties

The most dominant feature controlling the physical properties is the fiber architecture and type. The mechanical property in any principal direction is determined by the volume fraction of fibers oriented in that direction. This decreases as the directionality increases in multi-D lay-ups. **Tables 3** and **4** give approximate values for different types of material.

Mechanical Properties

Young Modulus

The Young modulus varies widely according to the composition and structure of the composite, see **Table 3**. Unidirectional materials may show values as high as 800 GPa after graphitization when the highly graphitizing MP fibers are used. Since both pitch and CVD matrices are highly graphitizable and also orient with the fiber, they may both contribute significantly to the observed Young modulus value, provided that the degree of filling of the fiber preform is high, **Figure 13**. One problem in modeling the mechanical properties of C–C composites is that it is not usually possible to obtain experimental data on the basic properties of the matrix material, because its structure is controlled by the fabrication conditions and the interaction with fibers. Simple mixture rules do not always work, because the level of microcracking and the residual porosity in incompletely densified composites may limit the load transfer between the fiber and the matrix as illustrated in **Figure 13**. For laminates and multi-D composites, the values in specific directions are more or less as expected from the fiber orientations and volume fractions. **Table 3** gives some typical values.

Strength and Fracture Behavior

C–C composites are examples of brittle matrix composites. Bulk carbon and graphite materials are porous and of generally low strength, elastic modulus, and fracture toughness. Therefore, in contrast to ceramic matrix composites, for which the matrix is intrinsically strong and where the fibers are used mainly to modify the mode of fracture and induce damage tolerance and mechanical reliability, in C–C the fibers do act to stiffen and reinforce as well as to

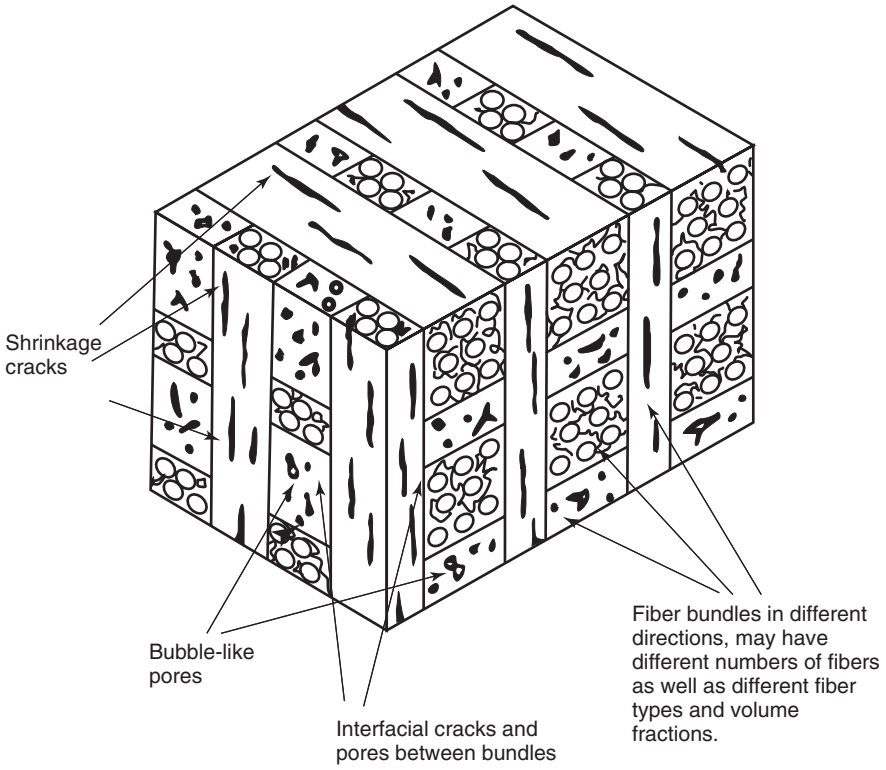


Figure 12 Schematic representation of the structure of a 3D orthogonal C–C composite depicting microcracks in fiber bundles and at bundle–bundle interfaces, pores and variable bundle dimensions, and character.

toughen the matrix. In common with other brittle matrix composites, the strength of the fiber–matrix bond is critical. The lower strain to failure of the matrix leads to the generation of cracks within this phase which, if the fiber–matrix bonding is strong, may propagate through the fiber, leading to brittle fracture at the matrix failure strain. However, if debonding can take place at the fiber–matrix interface, the cracks may be deflected along the fiber axis allowing the fibers to bridge the crack leading to classical multiple matrix cracking. In 1D composites, when this cracking is saturated with approximately equally spaced cracks, the elastic modulus should approximate to $v_f E_f$ and failure should occur at the failure strain of the fibers, leading to full utilization of the fiber strength. **Figure 14** shows stress–strain plots for 1D composites with weak and strong fiber–matrix bonding, illustrating this behavior. The crack bridging and fiber pull-out lead to high work of fracture and excellent damage tolerance. However, weak bonding at the fiber–matrix interface also reduces the intralaminar shear strength in 1D materials. The fiber–matrix bond strength with pitch and resin-derived matrices arise from the interaction of the precursor with the complexes on the fiber surfaces, which depend on the previous treatments given to the fibers by their manufacturers to control fiber–resin interactions. In the absence of such complexes, the bonding is very weak or even absent. The fiber–matrix bonding is also influenced by the heat-treatment. Graphitization increases the solid density of the matrix and fiber, if it has not experienced that temperature regime in its production. This introduces internal stresses and shrinkages which tend to debond the fiber from its matrix. This occurs with all matrix types. Graphitization tends to "polygonize" the oriented matrix sheath around the fiber and also shrink the fiber radius, so debonding the interface, **Figure 15**. Thus, a material that displays brittle fracture after carbonization in the range 1000–1600°C may show multiple matrix cracking and damage tolerance after graphitization. **Figure 14** shows these effects for a 1D material in tensile loading. In flexural testing, any weak interfaces may promote shear failure at stresses significantly lower than the tensile fracture stress unless very large span to depth ratios are adopted.

The behavior of 2D laminates and of multi-D structures may also be influenced by the interfacial region, but other factors come into play. Typical strength values for composites of different architecture are shown in **Table 3**. The 3D and multi-D materials generally show low strength values in specific directions because of the lower fiber volume fraction.

Table 3 Typical mechanical of various types of carbon–carbon composite

Composite type	*Fiber volume fraction in measurement direction*	*Tensile modulus* (GPa)	*Tensile strength* (MPa)	*Tensile strain at failure* (%)	*Flexural modulus* (GPa)	*Flexural strength* (MPa)	*Compressive strength* (MPa)	*Shear strength* (MPa)	*Fracture energy* ($kJ\,m^{-1}$)	*Poisson's ratio*
Unidirectional										
-fiber axis	0.5–0.7	up to 800	100–1300[a]	0.2–0.7[a]	up to 800	100–1300[a]	up to 400	20–30	20–100	0.25
-transverse		<10	<15		<10	<20	20–50		<1	
2D										
-typical in-plane (*XY*)	0.25	10–100	100				150–200	30		
-typical *z* direction		~60								
-8 harness satin (*XY*)					10–70	100–300				
3D										
-orthogonal (*XY*)	0.16	56	100	0.2			70			
-orthogonal (45° to *XY*)	0	5	20							
-3D/226 (*Z*)	0.3	140	300				160			
-3D/226 (*XY*)	0.1	60	105				120			
7D (Z)	0.07	25	50							
Felt	0.1–0.25				10–20	25–300				

Actual values depend on exact fiber type and volume fraction in measurement direction, and volume fraction and nature of the carbon matrix.
[a] Depends on interfacial bond strength.

Table 4 Typical thermal properties of various types of carbon–carbon composite

Composite type	*Coefficient of thermal expansion at* 25°C (10^{-6} K^{-1})	*Coefficient of thermal expansion at* ~1000°C (10^{-6} K^{-1})	*Thermal conductivity at* 25°C (Wm^{-1} K^{-1})
Unidirectional			
-fiber axis	− 1–0	1–2	50–800[a]
-transverse	0–10	0–10	10–200[a]
2D			
-in plane (*XY*)	~0	1.5–2.6	50–250
-*z* direction			30–100[a]
3D			
Orthogonal	0	2–4	10–250

Actual values depend on exact fiber type and volume fraction in measurement direction, and volume fraction and nature of the carbon matrix.
[a] High values obtained with graphitized pitch based fibers and matrix.

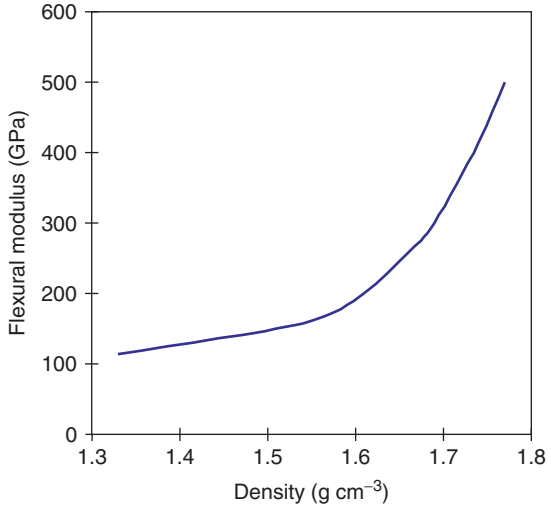


Figure 13 Example of the effect on the elastic modulus of successively densifying a 1D composite by a series of reimpregnation and recarbonization steps using pitch. The final composite would be graphitized. As the pore space between fibers is progressively filled, the load transfer to the fibers is dramatically improved. These data would be typical of a 50–60% volume of PAN-based fibers which after graphitization would display a Young modulus of ~450 GPa. Note that the composite modulus exceeds this value showing that the pitch-based matrix oriented around the fibers develops a Young modulus in excess of that of the fibers.

In directions with minimal fiber fraction, for example, at 45° to an orthogonal direction in orthogonal 3D composites, the fracture stress is similar to the values for normal un-reinforced porous graphites. Multi-D composites tend to show a very gradual type of failure mode (sometimes called "pseudoplastic"), indicating high damage tolerance, **Figures 16** and **17**. This is probably due to the many cracks and interfaces present that may deflect the main crack and provide a variety of crack-bridging elements (fibers and fiber bundles) that make a major contribution to the fracture energy. In woven cloth laminate structures, the presence of needled tows may enhance the strength and elastic modulus in the "*z*"-direction but this may be at the expense of some disruption of the "in-plane" fiber arrangement resulting in a small decrease in the "in-plane" properties. Similarly, the weave pattern can influence the "in-plane" properties due to the crimping effect of the criss-crossing fiber tows. The various bundles in the multi-D composites may be of different sizes or contain different fibers, enabling the composite to be designed to withstand the anticipated stresses in various directions to be experienced in the specific application.

Effect of Temperature on Mechanical Properties

The Young moduli and the fracture stress of most carbon–graphite materials are remarkable in that they increase with measurement temperature. The effect is attributed to the re-expansion into the microcracks created by thermal expansion mismatch, which lowers the porosity and increases the Young modulus and strength. These effects are present in C fiber–C matrix composites to a significant degree. Thus, C–C composites in nonoxidizing conditions are the strongest materials at very high temperatures when the strengths of other materials are drastically reduced. **Figure 18** shows a comparison, whilst **Figure 19** shows schematically how the fracture behavior of 3D orthogonal material may be drastically altered at extreme temperatures.

Creep and Fatigue

It is because of the strong covalent bonding that the mechanical properties do not degrade significantly

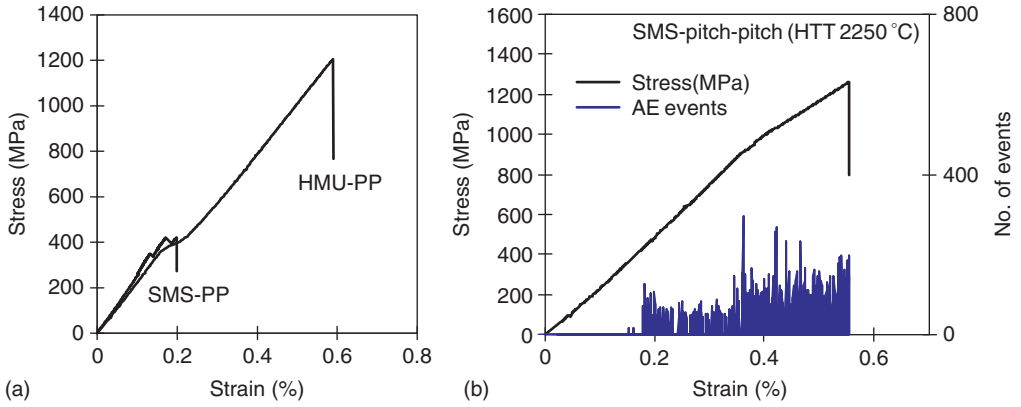


Figure 14 Tensile behavior of 1D carbon composites. (a) A surface-treated intermediate modulus PAN-based fiber (SMS-PP) composite is compared with one of the similar fiber volume fractions using high modulus carbon fibers. The former has an active surface whilst the latter has an inert, largely basal, surface. In both cases, the matrix is derived from pitch and the carbonization temperature is in the region of 1000°C. The fracture is catastrophic at the failure strain of the resin for the surface-treated fiber composite, due to a strong interfacial bond, whilst in the latter case multiple matrix fracture occurs around the same value of strain and the ultimate fracture stress is determined by the fibers, giving strength greater than 1 GPa. (b) The behavior of the surface-treated fiber composite is displayed after graphitization. The matrix has now pulled away from the fiber weakening the interface and dramatically increasing the strength and the strain at ultimate failure. During this heat treatment the Young modulus of the fiber increases as well. (b) Also shows the acoustic emission accompanying the test, indicating some subcritical events prior to fracture.

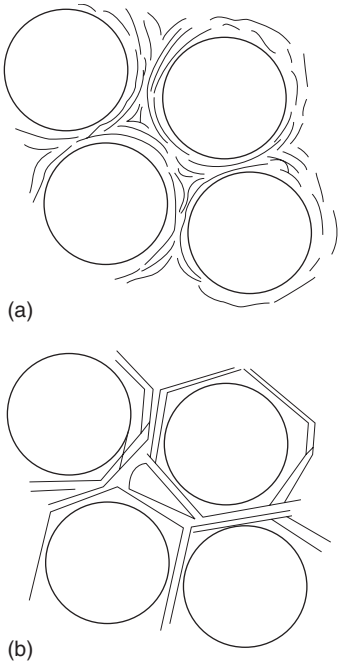


Figure 15 Schematic representation of the preferred orientation of graphene layers with fiber surface in C–C composites. (a) Carbonized matrix and (b) after graphitization.

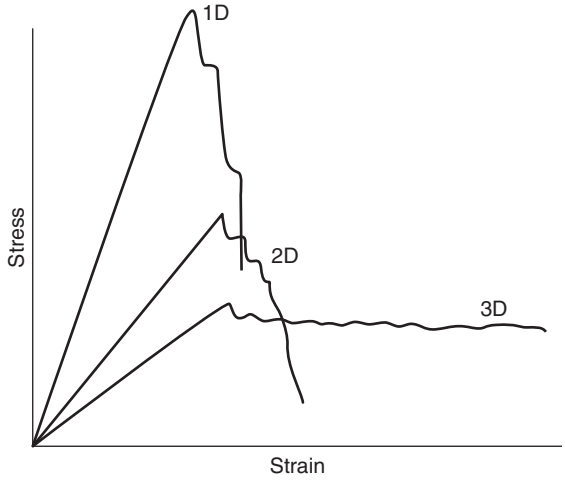


Figure 16 Schematic representation of the different general stress–strain behavior of C–C composites of different fiber architecture.

until very high temperatures are achieved. Thus, the creep resistance of C–C composites is very high, there being virtually no creep below ~2000°C, but above this region the creep rate becomes finite and shows a steady increase as the temperature rises to 3000°C.

The fatigue resistance of C–C composites is also good. Studies in flexure and in torsion have shown that the opening and closing of slit-like fissures (pores) in the structure is a significant feature. At stresses greater than about 1/3 the fracture stress, microcracks are formed and internal damage to the matrix can lead to the loss of small particles, a so-called "dusting out" process.

Friction and Wear

The tribological properties are critical in the major commercial application of C–C composites, in aircraft brake linings. The friction coefficient increases with temperature to a value as high as 0.6 at the

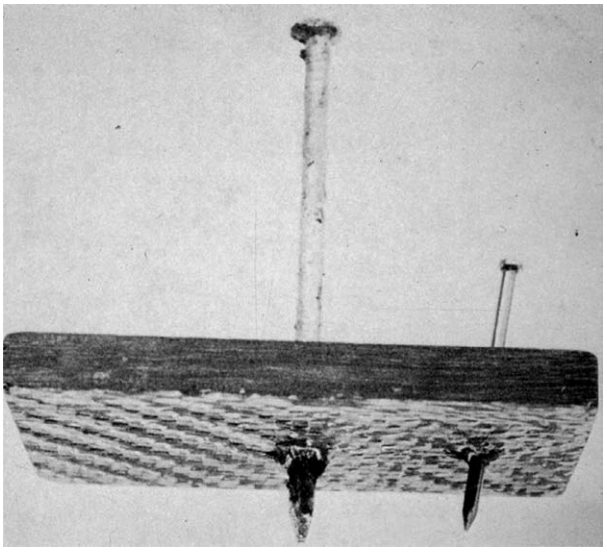


Figure 17 Nails driven through a C–C composite, illustrating the general toughness (damage tolerance) of this material.

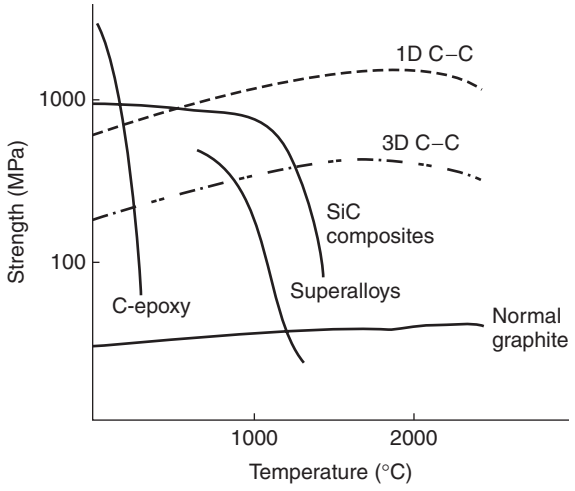


Figure 18 Comparison of the temperature dependence of strength of C–C composites with other advanced materials.

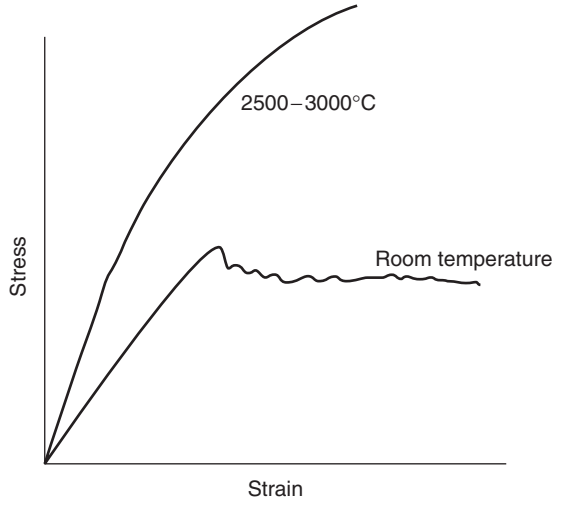


Figure 19 Schematic representation of the effect of temperature on the stress–strain behavior of a 3D orthogonal C–C composite.

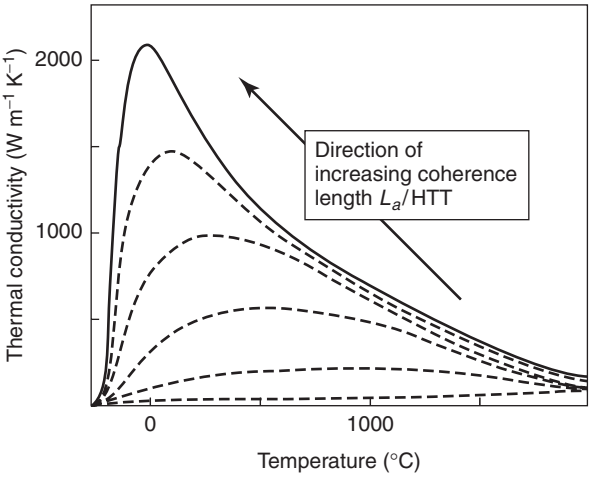


Figure 20 General trend of thermal conductivity with temperature for carbon–graphite materials of different order (extent of graphitization).

same time as the wear resistance increases. The key factor is the structure of the film of wear debris at the interface which may take on different structures in different wear regimes, that is, low temperature, during taxi-ing or high temperatures during landing and aborted take off.

Thermal and Electrical Properties

Thermal Expansion

Carbon–carbon composites are extremely dimensionally stable over very wide temperature ranges. The thermal expansion in fiber directions is controlled by the values of the fibers themselves, which for high-modulus fibers and fibers that have been heat-treated to graphitizing temperatures during the composite processing is initially negative rising to a small positive value above 150°C. The result is that over a temperature range of 2800°C, the linear expansion is only $\sim$0.7%. This value prevails more or less within the plane of laminates and in fiber directions in multi-D composites. In transverse directions, the situation is more complicated, the expansion coefficient being determined by the extent to which the large "*c*-axis" expansion of the graphite crystal is accommodated in microcracks, and the orientation of the lamellar regions and the associated cracks to the direction of measurement. Some typical values are listed in **Table 4**.

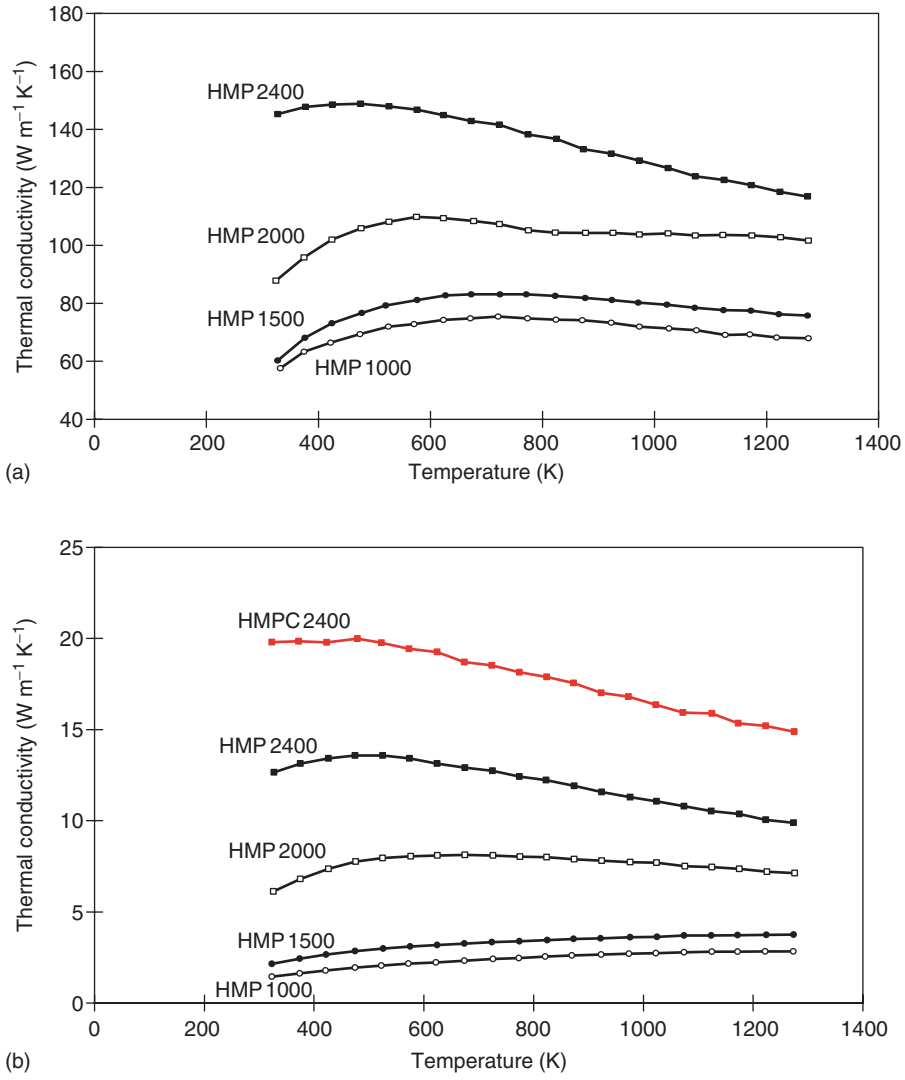


Figure 21 Thermal conductivity data for 1D model laminates measured (a) in the longitudinal and (b) in the transverse directions. The composites all comprise high modulus PAN-based carbon fibers with a pitch-based matrix, heat treated to temperatures in the range 1000–2400°C.

Thermal and Electrical Conductivity

Thermal and electrical conductivity both increase with the degree of order of the graphene layers and also with the extent of preferred orientation of the layers with respect to the thermal or potential gradient. Both parameters increase with the *a*-axis coherence length, L_a. The thermal transport in graphite is dominated by lattice vibrations (phonons). Only near 0 K is the electronic contribution to heat conduction significant. The thermal conductivity is highly temperature and structure dependent. Phonon–phonon interaction (Umklapp scattering) and scattering at boundaries between ordered regions are the principal processes controlling the conductivity at elevated temperatures and the relative effects of these two mechanisms depends on the coherence length, L_a. When L_a is large (graphitized materials), the phonon scattering mechanism is dominant, but when it is small, as in nongraphitizable carbons and materials heat treated only to modest temperatures, the Umklapp mechanism is significant only when the phonon mean free path becomes smaller than the coherence length, that is, at high temperatures. The result is that thermal conductivity–temperature curves for carbons of varying structural perfection show maxima at different temperatures, the value of the maximum increasing with the degree of graphitization, **Figure 20**. The maximum for graphite and highly graphitized fibers and matrices is below room temperature, but for poorly graphitized materials it may be as high as 500–1000°C. At very high

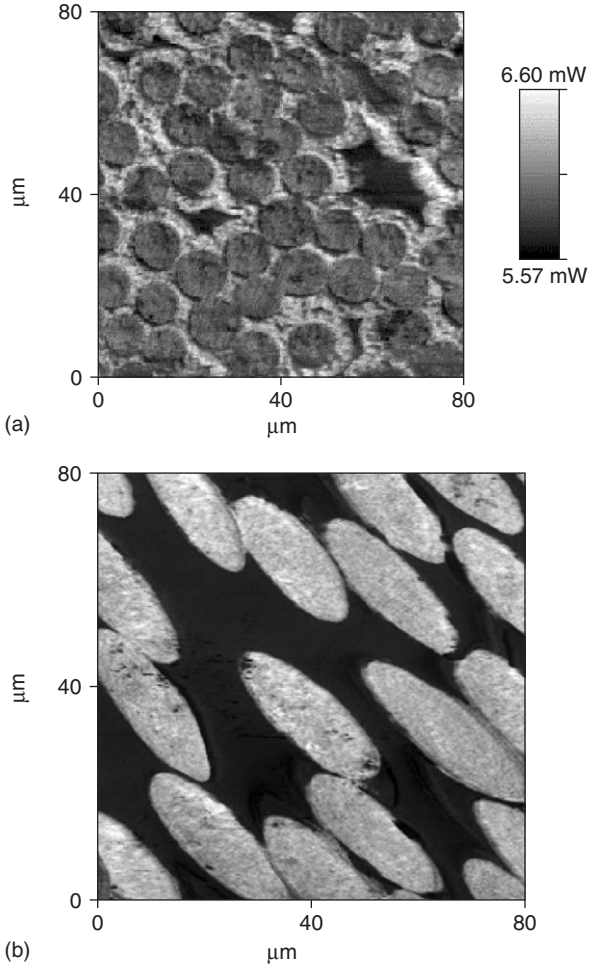


Figure 22 Thermal micrograph of transverse sections of (a) a PAN fiber-based composite with a CVD matrix and (b) a high thermal conductivity mesophase pitch-based fiber composite with a resin-derived matrix. The brightness reflects the relative thermal conductivity, which in (a) is greatest in the graphitizable matrix, but in (b) is greatest in the graphitized fiber.

temperatures, the curves tend to come together resulting in very little difference in thermal transport of different forms of carbon in this region. **Table 4** lists thermal conductivities of various types of composites and **Figure 21** shows data for a graphitized series of 1D laminates, showing the increase in thermal conductivity as the matrix is progressively improved in order. PAN-based carbon fibers are nongraphitizing and display limited thermal conductivities even for the high-modulus forms that have been heat treated to near 3000°C. Thus, pitch- and CVD-derived matrices, after graphitization, may have significantly higher conductivities than PAN-based fibers. Graphitized MP fibers, on the other hand, have high conductivities. **Figure 22** illustrates these trends. When mesophase pitch-based fibers are combined with these two types of matrix and graphitized, then extremely high thermal conductivities may be obtained.

The electrical conductivity of composites will generally vary in the same way as thermal conductivities. An empirical relationship between the two properties has been developed for carbon fibers. It is possible that a similar relationship may prevail for 1D materials, in the fiber direction. This arises because both properties vary with structure in the same way.

Thermal Shock

The combination of low expansion coefficient with high thermal conductivity, high strength at elevated temperatures and very high work of fracture result in exceptional thermal shock resistance of C–C composites, a critical factor in refractory applications, such as furnace parts, first wall materials in fusion reactors, rocket nozzles, and shielding (e.g., shuttle nose cone and leading edges of wing tips).

Oxidation

Because of the oxidation of graphite, that is prevalent above $\sim 500^{\circ}C$, the exceptional properties can only be achieved at high temperatures in inert atmospheres or for limited periods of time in oxidizing ones. There have been many attempts to protect carbon against oxidation. The main approaches are:

- incorporation of various glassy phases in the pore structure to reduce oxygen transport to the internal structure,
- inclusion of carbide-forming elements during fabrication. The carbides oxidize to protective oxides,
- coating the outer surface with refractory carbides. Thermal expansion mismatch may lead to cracking of the layer on cooling, but the internal glassy phases are designed to help seal these cracks on reheating.

Concluding Remarks

The properties of C–C composites can vary over a vast range due to the variety of carbon structures that may be engineered within the fibers and the carbon matrices, and the variety of fiber architectures that exist. The carbon structures may change dramatically during thermal processing. This variation enables materials to be designed with properties that may approach those intrinsic to the graphite crystal in 1D materials which are highly anisotropic, or may have almost isotropic properties in multi-D materials with properties in specific directions that reflect the

amount of fiber so oriented. The materials may be the most refractory and dimensionally stable of all in inert atmospheres or for limited periods in oxidizing ones. The technology is fairly mature, the most active field currently being to reduce processing costs and time, and to develop C–C–SiC composites with improved wear for terrestrial braking systems.

See also: Composites: Overview; Composites: Polymer–Polymer; Mechanical Properties: Elastic Behavior; Mechanical Properties: Tensile Properties; Scattering, Elastic (General); Shubnikov–de Haas and de Haas–van Alphen Techniques; Transmission Electron Microscopy; van der Waals Bonding and Inert Gases; Vapor Transport Processing: Mechanisms.

PACS: 61; 61.43Er; 61.66Bi; 61.72Nn; 62; 62.20. – x; 65; 65.40. – b; 65.40.De

Further Reading

Buckley JD and Edie DD (1992) Carbon–Carbon Materials and Composites. NASA Reference publication 1254.

Donnet JB and Bansal RC (1990) *Carbon Fibers*. New York: Dekker.

Dresselhaus MS, Dresselhaus G, Sugihara K, Spain IL, and Goldberg HA (1988) *Graphite Fibers and Filaments*. Berlin: Springer.

Figueiredo JL, Bernardo CA, Baker RTK, and Huttinger KJ' (eds.) (1989) *Carbon Fibres, Filaments and Composites'*. Dordrecht: Kluwer, NATO ASI Series E: Vol. 177.

Fitzer E and Manocha LM (1998) *Carbon Reinforcements and Carbon/Carbon Composites*. Berlin: Springer.

McAllister E and Lachman WL (1983) Multidirectional carbon–carbon composites, In: Kelly A and Mileiko ST (eds.) *Fabrication of Composites*, vol. 4 of Kelly, A and Rabotnov, YN (series eds.) Handbook of Composites, North Holland publishers, Amsterdam.

Ngai T (1990) Carbon–carbon composites. In: Lee SM (ed.) *International Encyclopedia of Composites*., vol. 1. New York: VCH publishers.

Savage G (1992) *Carbon–Carbon Composites*. London: Chapman & Hall.

Thomas CR (ed.) (1993) *Essentials of Carbon–Carbon Composites*. Cambridge: Royal Society of Chemistry.

Trefilov VI (ed.) (1995) *Ceramic– and Carbon-Matrix Composites*. London: Chapman & Hall.

Walker PL, Jr and Thrower P (eds.) *Chemistry and Physics of Carbon*. New York: Dekker. (An important series of volumes containing comprehensive research reviews of the whole field of carbon and graphite science.)

Watt W and Perov BV (eds.) (1985) *Strong Fibres*, vol. 1 of Kelly A and Rabotnov YN (series eds.) Handbook of Composites, North Holland publishers, Amsterdam.

Nomenclature

D	Fiber directionality
E	Young modulus
E_f	Young modulus of fiber
K	thermal conductivity
v_f	volume fraction of fiber

Composites: Overview

C H Zweben, Devon, PA, USA

Introduction

A composite material is two or more materials bonded together. This definition distinguishes composite materials (composites) from materials such as alloys. The most important composites consist of matrices reinforced with various types of fibers, whiskers, or particles. Using composites technology, it is possible to design new, multifunctional materials with unique properties and combinations of properties that cannot be obtained with any monolithic material.

Many widely used materials are composites, but not generally recognized as such. For example, tungsten carbide cutting tools are actually composites consisting of tungsten carbide particles bound together with a cobalt matrix, rather than monolithic tungsten carbide. The reason for using the composite is that it has a much higher fracture-toughness than monolithic tungsten carbide, a brittle ceramic.

Composites are material systems. Their properties depend on those of the constituents (reinforcement and matrix), reinforcement coating, if any, and the process by which they are made. All of these have pronounced effects, and altering any of these can greatly affect their properties.

The development of composites and the related design and manufacturing technologies is one of the most important advances in the history of materials. Composites are widely used, and are enabling materials in many applications. For example, Voyager, which was in effect an all-composite aircraft, was the first to fly around the world without refueling. However, it is emphasized that modern composites technology is only about five decades old. In the history of materials, this is hardly the blink of an eye.

Composite materials are considered primarily as structural materials. However, there are numerous examples of applications for which nonmechanical properties, such as thermal, electrical, magnetic, and

piezoelectric, are of primary interest. Important applications based on nonmechanical properties include glass–fiber-reinforced polymer printed circuit boards, magnetic particle-reinforced polymer audio and video recording tape, and ceramic particle-reinforced polymer sensors.

Another unique characteristic of composites is that many are made with processes that have advantages over those used for monolithic materials. To cite one example, manufacturing methods for fiber-reinforced polymers allow fabrication of aircraft structures having shapes that are aerodynamically superior to those made from metals. Another illustration is the magnetic particle-reinforced polymer permanent magnets that can easily be made into complex shapes by processes such as injection molding.

Composites are usually categorized by the type of material used for the matrix. The primary classes are polymer matrix composites (PMCs), metal matrix composites (MMCs), intermetallic matrix composites (IMCs), ceramic matrix composites (CMCs), and carbon matrix composites (CAMCs). The last category, CAMCs, includes carbon–carbon composites (CCCs), which consist of carbon matrices reinforced with carbon fibers. Currently, PMCs are the most widely used.

It is worth noting that biological structural materials occurring in nature are typically composites. Common examples are wood, bamboo, bone, teeth, and shell. Further, use of artificial composite materials is not new. Bricks made from straw-reinforced mud were employed in biblical times. This material has also been widely used in the American Southwest for centuries, where it is known as adobe. In current terminology, it would be described as an organic fiber-reinforced ceramic matrix composite.

As discussed earlier, there are many types of composites. This article concentrates on those that are used primarily because of their mechanical and thermal properties.

Table 1 Types of composite materials

Matrix	*Reinforcement*			
	Polymer	*Metal*	*Ceramic*	*Carbon*
Polymer	X	X	X	X
Metal	X	X	X	X
Ceramic	X	X	X	X
Carbon	X	X	X	X

Overview of Composite Materials and Applications

For the purposes of this article, solid materials are divided into four categories: polymers, metals, ceramics, and carbon. Carbon as a separate class is considered because of its many unique forms and characteristics. There are reinforcements and matrix materials in all four categories of solids. This results in the potential for a limitless number of new material systems having unique properties that cannot be obtained with any single monolithic material. Composites consisting of all possible combinations of matrix materials and reinforcements have been produced (see **Table 1**).

For decades, CCCs were the only significant type of CAMCs. However, there are now other types of composites utilizing a carbon matrix. Notable among these is silicon carbide fiber-reinforced carbon, which is being used in military aircraft gas turbine engine components.

Table 2 shows the radical differences in the characteristics of the four classes of matrix materials used in composites. This is reflected in the properties of the resulting composites.

Composites are now important commercial and aerospace materials. Currently, PMCs are the most widely used. MMCs are employed in a significant and increasing number of commercial and aerospace applications, such as automobile engines, electronic packaging, cutting tools, circuit breakers, contact pads, high-speed and precision machinery, and aircraft structures. CCCs are used in high temperature, lightly loaded applications, such as aircraft brakes, rocket nozzles, glass processing equipment, and heat

Table 2 Properties of selected matrix materials

Material	*Class*	*Density* ($g\,cm^{-3}$)	*Modulus* (GPa)	*Tensile strength* (MPa)	*Tensile failure strain* (%)	*Thermal conductivity* ($W\,m^{-1}\,K^{-1}$)	*CTE* ($10^{-6}\,K^{-1}$)
Epoxy	Polymer	1.8	3.5	70	3	0.1	60
Aluminum (6061)	Metal	2.7	69	300	10	180	23
Titanium (6A1-4 V)	Metal	4.4	105	1100	10	16	9.5
Silicon carbide	Ceramic	2.9	520		< 0.1	81	4.9
Alumina	Ceramic	3.9	380		< 0.1	20	6.7
Glass (borosilicate)	Ceramic	2.2	63		< 0.1	2	5
Amorphous carbon	Carbon	1.8	20		< 0.1	5–9	2

treatment furnaces. The space shuttle "Orbiter" has CCC leading edges. Although CMCs are not as widely used at present, there are notable production and developmental applications that are indicative of their potential.

Composites technology makes possible the use of an entire class of solid materials, for example, ceramics, in applications for which monolithic versions are unsuited because of their great strength scatter and poor resistance to mechanical and thermal shock. Further, some manufacturing processes for CAMCs are well adapted to the fabrication of large, complex structures. This allows consolidation of parts, which can reduce the manufacturing costs.

In addition to their excellent structural properties, composites have unique physical properties that are of great interest in applications such as thermal management and packaging of microelectronic, optoelectronic, and microelectromechanical (MEMS) devices. For example, carbon fibers with thermal conductivities much greater than that of copper are now commercially available. These reinforcements are being used in polymer, metal, and carbon matrices to create composites with high thermal conductivities that are being used in applications for which thermal management is important. Discontinuous versions of these fibers are also being incorporated in thermoplastic injection molding compounds, improving their thermal conductivity by as much as two orders of magnitude or more. This greatly expands the range of products for which injection-molded polymers can be used. In addition, there are an increasing number of composites reinforced with diamond particles, which impart very high thermal conductivities and low coefficients of thermal expansion (CTEs). Thermal management materials are covered in a separate section later.

Composites are now used in a wide and increasing range of important commercial and aerospace applications, including

- internal combustion engines;
- machine components;
- thermal management;
- electronic, optoelectronic, and MEMS packaging;
- automobile, train, ship, spacecraft, launch vehicle, and aircraft structures;
- aircraft and commercial gas turbine engines;
- mechanical components, such as brakes, drive shafts, and flywheels;
- tanks and pressure vessels;
- dimensionally stable components;
- process industries equipment requiring resistance to high-temperature corrosion, oxidation, and wear;
- offshore and onshore oil exploration and production;
- sports and leisure equipment;
- biomedical devices;
- civil engineering structures;
- instrument structures; and
- antennas.

Over the years, increasing production volumes have helped to reduce material prices, increasing their attractiveness in cost-sensitive applications.

As mentioned earlier, composites technology is in its infancy. It is anticipated that new and greatly improved materials are likely to emerge. It is also expected that new concepts will emerge, such as "smart materials" that incorporate greater functionality, including the integration of electronics, sensors, and actuators.

Reinforcements

The main types of reinforcements used in composite materials include continuous (cont.) fibers, discontinuous (disc.) fibers, whiskers (elongated single crystals), particles (including flakes), and numerous forms of fibrous architectures produced by textile technology, such as fabrics and braids. In addition, there is an increasing interest in composites reinforced with various types of nanoparticles, such as carbon nanotubes, graphite nanoplatelets, and silica (clay) nanoparticles.

As mentioned earlier, use of natural fiber-reinforcement is far from a new idea. However, in recent years there has been considerable work on composites reinforced with a variety of naturally occurring fibrous materials, including wood fiber, kenaf, hemp, flax, jute, sisal, banana leaf, china reed, and rice hulls. Matrices include thermoplastic and thermoset polymers and Portland cement. A number of material systems have been used in production applications, such as construction and automobile parts.

Increasingly, designers are using hybrid composites that combine different types of reinforcements and reinforcement forms to achieve greater efficiency and reduce cost. For example, fabrics and unidirectional tapes are often used together in structural components. In addition, carbon fibers are combined with glass or aramid to improve impact resistance. Laminates combining composites and metals, such as "Glare," which consists of layers of aluminum and glass fiber-reinforced epoxy, are being used in aircraft structures to improve fatigue resistance. There are also examples of composites reinforced with combinations of fibers and particles.

The great importance of composites and the revolutionary improvements in properties they offer

Table 3 Properties of key reinforcing fibers

Fiber	*Density* ($g\,cm^{-3}$)	*Axial modulus* (GPa)	*Axial tensile strength* (MPa)	*Axial CTE* ($10^{-6}\,K^{-1}$)	*Axial thermal conductivity* ($W\,m^{-1}\,K^{-1}$)
E-glass	2.6	70	2000	5	0.9
HS glass	2.5	83	4200	4.1	0.9
Aramid	1.4	124	3200	−5.2	0.04
Boron	2.6	400	3600	4.5	
SM carbon (PAN)	1.7	235	3200	−0.5	9
UHM carbon (PAN)	1.9	590	3800	−1	18
UHS (IM) carbon (PAN)	1.8	290	7000	−1.5	160
UHM carbon (pitch)	2.2	895	2200	−1.6	640
UHK carbon (pitch)	2.2	830	2200	−1.6	1100
SiC monofilament	3.0	400	3600	4.9	
SiC multifilament	3.0	400	3100		
Si–C–O	2.6	190	2900	3.9	1.4
Si–Ti–C–O	2.4	190	3300	3.1	
Aluminum oxide	3.9	370	1900	7.9	
High density polyethylene	0.97	172	3000		
High-modulus PBO	1.58	270	5800	−6.0	

are derived to a great extent from the development of synthetic fibers with unprecedented properties. The key synthetic fibers are made from glass, carbon (sometimes called graphite), ceramics, and high-modulus organics, such as aramids. Most fibers are produced in the form of multifilament bundles called strands or ends in untwisted forms, and yarns when twisted. Some fibers are produced as monofilaments, which generally have much larger diameters than strand filaments. **Table 3** presents typical properties of key fibers, which are discussed in the following subsections.

Fiber strength requires some discussion. Most of the key fibrous reinforcements are made of brittle ceramics or carbon. It is well known that the strengths of monolithic ceramics decrease with increasing material volume because of the increasing probability of finding strength-limiting flaws. This is called size effect. As a result of size effect, fiber strength typically decreases monotonically with increasing guage length and diameter. Flaw sensitivity also results in considerable strength scatter at a fixed test length. Consequently, there is no single value that characterizes fiber strength. This is also true of key organic reinforcements, such as aramid fibers. Consequently, the values presented in **Table 3** should be considered as approximate values, and are useful primarily for comparative purposes. It is to be noted that because unsupported fibers buckle under very low stresses, it is very difficult to measure their inherent compression strength, and these properties are almost never reported. Instead, composite compression strength is measured directly.

Glass Fibers

Glass fibers are used primarily to reinforce polymers. The leading types of glass fibers are E-glass, high-strength (HS)-glass, and corrosion resistant (CR)-glass. E-glass fibers, the first major synthetic composite reinforcement, was originally developed for electrical insulation applications (that is the origin of the "E"). E-glass fibers are, by many orders of magnitude, the most widely used of all fibrous reinforcements. The primary reasons for this are their low cost and early development compared to other fibers. Glass fibers are produced as multifilament bundles. Filament diameters range from 3 to 20 μm. **Table 3** presents the representative properties of E- and HS-glass fibers.

E-glass fibers have relatively low elastic moduli when compared to other reinforcements. In addition, E-glass fibers are susceptible to creep and creep (stress) rupture. HS-glass is stiffer and stronger than E-glass, and has better resistance to fatigue and creep.

The thermal and electrical conductivities of glass fibers are low, and glass fiber-reinforced PMCs are often used as thermal and electrical insulators. The CTE of glass fibers is also low when compared to most metals.

Carbon (Graphite) Fibers

Carbon fibers, often called graphite fibers in the US, are used to reinforce polymer-, metal-, ceramic-, and carbon matrices. There are dozens of commercial carbon fibers, with a wide range of strengths and moduli. As a class of reinforcement, carbon fibers are

characterized by high stiffness and strength and low density and CTE. Fibers with nominal tensile moduli as high as 895 GPa and with tensile strengths of 7000 MPa are commercially available. Carbon fibers have excellent resistance to creep, stress rupture, fatigue, and corrosive environments, although they oxidize at high temperatures. Some types of carbon fibers also have extremely high thermal conductivities – many times that of copper. This characteristic is of considerable interest in electronic packaging and other applications, where temperature control is important. Carbon fibers are workhorse reinforcements in high-performance aerospace and commercial PMCs and some CMCs. Of course, as the name suggests, carbon fibers are also reinforcements in carbon–carbon composites.

Most carbon fibers are highly anisotropic. Axial modulus, tension and compression strength, and thermal conductivity are typically much greater than the corresponding properties in the radial direction. Carbon fibers generally have small, negative axial CTEs (which means that they get shorter when heated) and positive radial CTEs. Diameters of common reinforcing fibers, which are produced in the form of multifilament bundles, range from 4 to 10 μm. Carbon fiber stress–strain curves tend to be nonlinear. The modulus increases with increasing tensile strain, and decreases with increasing compressive strain.

The three key precursor materials for carbon fibers are polyacrylonitrile (PAN), petroleum pitch, and coal tar pitch. Rayon-based fibers, once the primary CCC reinforcement, are far less common in new applications. Another type of carbon fiber, referred to as "vapor grown" by its manufacturer is made by a chemical vapor deposition (CVD) process. Some of the latter have exhibited axial thermal conductivities as high as 2000 $W\,m^{-1}\,K^{-1}$, five times that of copper.

Carbon fibers made from PAN are the most widely used. There are dozens available in the market. Fiber axial moduli range from ~235–590 GPa. They generally provide composites with excellent tensile and compressive strengths, although strength properties tend to drop off as the modulus increases. Fibers having nominal tensile strengths, as high as 7 Gpa, are available. **Table 3** presents properties of three types of PAN-based carbon fibers and two types of pitch-based carbon fibers. The PAN-based fibers are standard modulus (SM), ultrahigh strength (UHS), and ultrahigh modulus (UHM). SM PAN fibers are the most widely used type of carbon fiber reinforcement. They are one of the first types commercialized, and tend to be the least expensive. UHS PAN carbon fibers are the strongest type of another widely used class of carbon fiber called intermediate modulus (IM), because of the fact that their axial modulus falls between those of SM and UHM carbon fibers. IM fibers are also widely used in aircraft and other aerospace structural applications. PAN fibers have relatively low thermal axial and transverse thermal conductivities.

A key advantage of pitch-based fibers is that they can be produced with much higher axial moduli and thermal conductivities than those made from PAN precursors. For example, UHM pitch fibers with nominal moduli as high as 895 GPa are available. In addition, some pitch fibers, designated as ultrahigh thermal conductivity (UHK), have extremely high axial thermal conductivities. For example, there are commercial UHK fibers with a nominal axial thermal conductivity of 1100 $W\,m^{-1}\,K^{-1}$, almost three times that of copper. However, composites made from pitch-based carbon fibers are generally somewhat weaker in tension and shear, and much weaker in compression than those using PAN-based reinforcements.

As mentioned earlier, carbon fibers display nonlinear stress–strain behavior, making the method of calculating the modulus critical. Various tangent and secant definitions are used throughout the industry, resulting in confusion in reported properties. The moduli presented in **Table 3** are based on tangents to the stress–strain curves at the origin.

Boron Fibers

Boron fibers are primarily used to reinforce PMCs and, to a lesser extent, MMCs. Boron fibers are produced as monofilaments (single filaments) by CVD of boron on a tungsten wire or carbon filament, the former being the most widely used. They have relatively large diameters (100–140 μm) compared to most other reinforcements. **Table 3** presents representative properties of boron fibers having a tungsten core and diameter of 140 μm. The ratio of overall fiber diameter to that of the tungsten core influences effective fiber properties. For example, fiber specific gravity is 2.57 for 100 μm fibers and 2.49 for 140 μm fibers. Because boron fibers are more expensive than many types of carbon fibers, their use is much more restricted.

Fibers Based on Silicon Carbide

Silicon carbide-based fibers are primarily used to reinforce metals and ceramics. There are a number of commercial fibers based on silicon carbide, one of which, a monofilament, is produced by CVD of high-purity silicon carbide on a carbon monofilament core. Some versions use a carbon-rich surface layer that serves as a reaction barrier in MMCs and CMCs.

There are a number of muitifilament silicon carbide-based fibers made by pyrolysis of polymers. Some of these contain varying amounts of silicon, carbon, oxygen, titanium, nitrogen, zirconium, and hydrogen. **Table 3** presents properties of selected silicon carbide-based fibers.

Fibers Based on Alumina

Alumina-based fibers are primarily used to reinforce metals and ceramics. As for silicon carbide-based fibers, they have a number of different chemical formulations. The primary constituents, in addition to alumina, are boria, silica, and zirconia. **Table 3** presents properties of high-purity alumina fibers.

Aramid Fibers

Aramid, or aromatic polyamide fibers are high-modulus organic reinforcements, primarily used in PMCs and for ballistic protection. There are a number of commercial aramid fibers produced by several manufacturers. As for other reinforcements, they are proprietary materials with differing properties. **Table 3** presents properties of one widely used aramid fiber, "Kevlar 49." As for carbon fibers, aramid fibers have nonlinear stress–strain curves.

High-Density Polyethylene Fibers

High-density polyethylene fibers are primarily used to reinforce polymers and for ballistic protection. **Table 3** presents properties of one such type. The properties of high-density polyethylene tend to decrease significantly with increasing temperature, and they are susceptible to creep deformation (time-dependent deformation under constant stress), even at low temperatures.

PBO Fibers

Many polymeric fibers have been developed over the years. One that appears to be making successful inroads is poly-*p*-phenylenebenzobisoxazole (PBO). There are two versions, regular and high modulus (HM). **Table 3** presents properties of the latter, which has the highest modulus of any commercial organic fiber. Applications of this type include aerospace components, sports equipment, and audio speaker cones.

Characteristics and Properties of Composite Materials

Composites are strongly heterogeneous materials. That is, the properties of a composite material vary considerably from point to point in the material. For example, the properties of a point located in the matrix are typically much different from one in the reinforcement phase. Most monolithic polymers, ceramics, metallic alloys, and intermetallic compounds are usually considered as homogeneous materials, to a first approximation.

Many artificial composites, especially those reinforced with fibers, are anisotropic, which means that their properties vary with direction (the properties of isotropic materials are the same in every direction). They share this characteristic with a widely used natural fibrous composite, wood. As for wood, when structures made from artificial fibrous composites are required to carry loads in more than one direction, they are typically used in a laminated form known as plywood. It is to be noted that the strength properties of some metals also vary with direction. This is typically related to manufacturing processes, such as rolling. In addition, all single crystals are anisotropic.

With the exception of MMCs, composites do not display plastic behavior as monolithic metals do, which makes composites more sensitive to stress concentrations. However, the absence of plastic deformation does not mean that composites should be considered brittle like monolithic ceramics. The heterogeneous nature of composites results in complex failure mechanisms that impart an effective toughness.

Fiber-reinforced materials have been found to produce durable and reliable structural components in countless applications. For example, PMCs have been used in production boats, electrical equipment, and solid rocket motors since the 1950s, and extensively in aircraft since the early 1970s. Technology has progressed to the point where the entire empennages (tail sections) of commercial aircraft are made of carbon–epoxy. Passenger planes under development are scheduled to have virtually all-composite structures, including wings and fuselages.

There are a large and increasing number of materials that fall in each of the four types of composites, which makes generalization difficult. However, as a class of materials, composites tend to have the following characteristics: tailorable mechanical and physical properties, high strength, high modulus, low density, low CTE, excellent resistance to fatigue, creep, creep rupture, corrosion, and wear. Composites are available with tailorable thermal and electrical conductivities that range from very low to very high.

As for monolithic materials, each of the four classes of composites has its own particular attributes. For example, CMCs tend to have particularly good resistance to corrosion, oxidation, and wear, along with high-temperature capability.

The outstanding mechanical properties of composite materials has been a key reason for their

extensive use in structures. However, composites also have important physical properties, especially low, tailorable CTE and high thermal conductivity, resulting in the use of an increasing number of applications. Key examples are electronic packaging and thermal management, as discussed later.

Many composites, such as PMCs reinforced with carbon and aramid fibers, and silicon carbide particle-reinforced aluminum, have low CTEs, which are advantageous in applications requiring dimensional stability. Examples include spacecraft structures, instrument structures, optical benches, and optoelectronic packaging. By appropriate selection of reinforcements and matrix materials, it is possible to produce composites with near-zero CTEs.

CTE tailorability provides a way to minimize thermal stresses and distortions that often arise when dissimilar materials are joined. For example, the CTE of silicon carbide particle-reinforced aluminum depends on particle content. By varying the amount of reinforcement, it is possible to match the CTEs of a variety of key engineering materials, such as steel, titanium, and alumina (aluminum oxide). This characteristic of composites has particular value in electronic packaging, because thermal stresses can cause failure of ceramic substrates, semiconductors, and solder joints.

There are a large and increasing number of thermally conductive PMCs, MMCs, and CAMCs. One of the most important types of reinforcements for these materials is pitch fibers. As discussed earlier, PAN-based fibers have relatively low thermal conductivities. However, pitch-based fibers with thermal conductivities more than twice that of copper are commercially available. These UHK reinforcements also have very high stiffnesses and low densities. As pointed out earlier, fibers made by CVD exhibit thermal conductivities as high as $2000\,\mathrm{W\,m^{-1}\,K^{-1}}$, about five times that of copper. Fibers made from another form of carbon, diamond, also have the potential for thermal conductivities in this range. PMCs and CCCs reinforced with UHK carbon fibers are being used in a wide range of applications, including spacecraft radiators, battery sleeves, electronic packaging, and motor enclosures. MMCs reinforced with diamond particles have reported thermal conductivities as high as $1200\,\mathrm{W\,m^{-1}\,K^{-1}}$, three times that of copper. The latter materials are of great interest in electronic and optoelectronic packaging.

Overview of Mechanical and Physical Properties

As discussed earlier, initially, the excellent mechanical properties of composites were the main reason for their use. Today, however, there are an increasing number of applications for which the unique and tailorable physical properties of composites are key considerations. For example, the extremely high thermal conductivity and tailorable CTEs of some composite material systems are leading to their increased use in electronic packaging. Similarly, the extremely high stiffness, near-zero CTE, and low density of carbon fiber-reinforced polymers have made these composites the materials of choice in a variety of applications, including spacecraft structures, antennas, and optomechanical system components such as telescope metering structures and optical benches.

As stated before, composites are complex, heterogeneous, and often anisotropic material systems. Their properties are affected by many variables, including, *in situ* constituent properties such as reinforcement form, volume fraction, and geometry; properties of the interphase, the region where the reinforcement and matrix are joined (also called the interface); and void content. The process by which the composite is made affects many of these variables. Composites containing the same matrix material and reinforcements, when combined by different processes, may have very different properties.

It is important to keep several other important things in mind when considering composite properties. First, most composites are proprietary material systems made by proprietary processes. There are few industry or government standards for composites and reinforcements, as there are for many structural metals. However, this limitation also applies to many monolithic ceramics and polymers, which are widely used engineering materials. Despite their inherently proprietary nature, there are some widely used composite materials made by a number of manufacturers which have similar properties. Notable examples are standard-modulus- and intermediate-modulus carbon fiber-reinforced epoxy.

Another critical issue is that properties are sensitive to the test methods by which they are measured, and there are many different test methods used throughout the industry. Further, test results are very sensitive to the skill of the technician performing them. As a consequence of these factors, it is very common to find significant differences in the reported properties of what is nominally the same composite material. Because of these considerations, the properties of composite materials in this article should be considered as approximate values. As for all materials, composite properties depend on temperature.

There is often a great deal of confusion among those unfamiliar with composites, about the effect of reinforcement form. The properties of composites

are very sensitive to reinforcement form, reinforcement volume fraction, and internal reinforcement geometry.

It is important to keep in mind that a key problem with disc. fiber–reinforcement is the difficulty to control fiber orientation. For example, material flow during processing can significantly align fibers in some regions. This affects all mechanical and physical properties, including modulus, strength, CTE, thermal conductivity, etc. This has been a frequent source of failures.

Traditional fabric reinforcements have fibers oriented at 0° and 90°. For the sake of completeness, it is noted that triaxial fabrics, which have fibers at 0°, +60°, and −60°, are now commercially available. Composites using a single layer of this type of reinforcement are approximately quasi-isotropic, which means that they have the same inplane elastic (but not strength) properties in every direction. Their thermal conductivity and CTEs are also approximately isotropic in the plane of the fabric. "Perfect" fabrics would have exactly isotropic thermal properties.

Overview of Manufacturing Processes

Manufacturing processes for some composites have significant advantages over those used for monolithic metals and ceramics. For example, fiber-reinforced polymers and ceramics can be fabricated in large, complex shapes that would be difficult or impossible to make with other materials. The ability to fabricate complex shapes allows consolidation of parts, which reduces machining and assembly costs. For example, one-piece PMC grill opening panels are widely used in automobiles, replacing as many as 12 metal parts that have to be joined by welding or bolting.

Some processes allow fabrication of parts in their final shape (net shape) or close to their final shape (near-net shape), which can also reduce manufacturing costs. The relative ease with which smooth shapes can be made is a significant factor in the use of composites in boats, aircraft, and other applications for which fluid dynamic and aerodynamic considerations are important. The manufacturing methods for the key composites are discussed in the respective sections.

Polymer Matrix Composites

PMCs are the most efficient structural materials that have ever been developed for moderate-temperature applications. As a result, they are now used as baseline materials in numerous applications, including aircraft, spacecraft, boats, solid-fuel launch vehicles, industrial equipment, and sports equipment. At present, thermosetting polymers are the key matrix materials for structural applications, but use of thermoplastics is gradually increasing.

The high thermal conductivities of some PMCs have led to their increasing use in applications such as spacecraft structures and electronic packaging components, such as printed circuit boards, heat sinks, heat spreaders, and heat sinks used to cool microprocessors. The addition of disc. thermally conductive carbon fibers and ceramic particles to thermoplastics significantly increases thermal conductivity, opening the door for the use of injection molded parts in an increasing number of applications, such as heat sinks and motor covers.

Polymers have low-stiffness and are relatively weak, viscoelastic materials with low thermal conductivities and high CTE. In order to obtain materials with mechanical properties that are acceptable for structural applications, it is necessary to reinforce them with cont. or disc. fibers. The addition of ceramic or metallic particles to polymers results in materials that have an increased modulus. However, as a rule, strength typically does not increase significantly, and may actually decrease.

There are many particle-reinforced polymers used in electronic packaging, primarily because of their physical properties. For these applications, ceramic particles, such as alumina, aluminum nitride, boron nitride, and even diamond are added to obtain electrically insulating materials with higher thermal conductivities and lower CTEs than those of the base polymer. Metallic particles such as silver, copper, and aluminum are added to create materials that are both electrically and thermally conductive. These materials have replaced solders in some applications. Magnetic composites are made by incorporating ferrous or magnetic ceramic particles in various polymers.

As stated earlier, for a wide range of applications, composites reinforced with cont. fibers are the most efficient structural materials at low to moderate temperatures. Consequently, the focus is on them. **Table 4** presents room temperature mechanical properties of unidirectional polymer matrix composites reinforced with key fibers: E-glass, aramid, boron, SM PAN carbon, IM PAN carbon, UHM PAN carbon, UHM pitch carbon, and UHK pitch carbon. The fiber volume fraction is 60%, which is a typical value.

The properties presented in **Table 4** are representative of what can be obtained with a well-made PMC employing an epoxy matrix. Epoxies are widely used, provide good mechanical properties, and can be considered as a reference matrix material. Properties of composites using other resins may differ from these.

The properties of PMCs, especially strength, depend strongly on temperature. The temperature dependence of polymer properties differs considerably. This is also true for different epoxy formulations, which have various glass transition temperatures.

The properties shown in **Table 4** are axial, transverse and shear moduli, Poisson's ratio, tensile and compressive strengths in the axial and transverse directions, and inplane shear strength. The Poisson's ratio presented is called the major Poisson's ratio. It is defined as the ratio of the magnitude of transverse strain to the axial strain when the composite is loaded in the axial direction. It is to be noted that transverse moduli and strengths are much lower than corresponding axial values. Unidirectional composites share this characteristic with wood, which is stronger and stiffer along the grain than perpendicular to it.

Elastic moduli are based on tangents to the stress–strain curves at the origin. Using this definition, tensile and compressive moduli are usually very similar. However, this is not the case for moduli which are computed using various secants. These typically produce compression moduli that are significantly lower than tensile moduli because the stress–strain curves are nonlinear.

Due to the low transverse strengths of unidirectional laminates, they are rarely used in structural applications. The design engineer selects laminates with layers in several directions to meet requirements for strength, stiffness, buckling, etc. There are an infinite number of laminate geometries that can be selected. For comparative purposes, it is useful to consider quasi-isotropic laminates, which have the same elastic properties in all directions in the plane of the fibers. It is to be noted that through-thickness properties of quasi-isotropic laminates are somewhat similar to the transverse properties of unidirectional composites.

Laminates have quasi-isotropic elastic properties when they have the same percentage of layers every $180^\circ/n$, where $n \geqslant 3$. The most common quasi-isotropic laminates have layers that repeat every 60°, 45°, or 30°. It is to be noted, however, that strength properties in the plane are not isotropic for these laminates, although they tend to become more uniform as the angle of repetition becomes smaller. Laminates have quasi-isotropic CTEs and CTEs, when they have the same percentage of layers in every $180^\circ/m$, where $m \geqslant 2$. For example, laminates with equal numbers of layers at 0° and 90° have quasi-isotropic thermal properties.

Table 5 presents the mechanical properties of quasi-isotropic laminates having equal numbers of layers at 0°, $+45^\circ$, -45° and 90°. The elastic moduli of all quasi-isotropic laminates are the same for a given material. It is worth noting that the moduli and strengths are much lower than the axial properties of unidirectional laminates made of the same material. In many applications, laminate geometry is such that the maximum axial modulus and tensile and compressive strengths fall somewhere between axial unidirectional and quasi-isotropic values.

Table 6 presents physical properties of selected unidirectional composite materials having a typical fiber volume fraction of 60%. The densities of all of the materials are considerably lower than that of aluminum, while some are lower than that of magnesium. This reflects the low densities of both fibers and matrix materials. The low densities of most polymers give PMCs a significant advantage over most MMCs and CMCs at low-to-moderate temperatures, all other things being equal.

As **Table 6** shows, all of the composites have relatively low axial CTEs. This results from the combination of low fiber axial CTE, high fiber stiffness, and low matrix stiffness. The CTE of most polymers is very high. It is noteworthy that the axial CTEs of PMCs reinforced with aramid fibers and some carbon fibers are negative. This means that, contrary to the general behavior of most monolithic materials, they contract in the axial direction, when heated.

The transverse CTEs of the composites are all positive, and their magnitudes are much larger than the magnitudes of the corresponding axial CTEs. This results from the high CTE of the matrix and a Poisson effect caused by a constraint of the matrix in the axial direction and lack of constraint in the transverse direction. The transverse CTE of aramid composites is particularly high, in part, because the fibers have a relatively high, positive radial CTE.

It is to be noted that PMCs also undergo dimensional changes due to moisture absorption and desorption. These changes are usually described by coefficients of moisture expansion, which are analogous to CTEs. This subject is beyond the scope of the current article.

The axial thermal conductivities of composites reinforced with glass, aramid, boron, and a number of the carbon fibers are relatively low. In fact, E-glass and aramid PMCs are often used as thermal insulators. As **Table 6** shows, most PMCs have low thermal conductivities in the transverse direction as a result of the low thermal conductivities of the matrices and the fibers in the radial direction. Through-thickness conductivities of laminates tend to be similar to the transverse thermal conductivities of unidirectional composites.

Table 7 presents the inplane thermal conductivities and CTEs of quasi-isotropic laminates made from

Table 4 Representative mechanical properties at room temperature of selected unidirectional polymer matrix composites (nominal fiber volume fraction = 60%)

Fiber	*Axial modulus* (GPa)	*Transverse modulus* (GPa)	*Inplane shear modulus* (GPa)	*Poisson's ratio*	*Axial tensile strength* (MPa)	*Tranverse tensile strength* (MPa)	*Axial compressive strength* (MPa)	*Tranverse compressive strength* (MPa)	*Inplane shear strength* (MPa)
E-glass	45	12	5.5	0.28	1020	40	620	140	70
Aramid	76	5.5	2.1	0.34	1240	30	280	140	60
Boron	210	19	4.8	0.25	1240	70	3310	280	90
SM carbon (PAN)	145	10	4.1	0.25	1520	41	1380	170	80
UHM carbon (PAN)	310	9	4.1	0.20	1380	41	760	170	80
UHM carbon (pitch)	480	9	4.1	0.25	900	20	280	100	41

Table 5 Mechanical properties of selected quasi-isotropic polymer matrix composites (fiber volume fraction = 60%)

Fiber	*Axial modulus* (GPa)	*Transverse modulus* (GPa)	*Inplane shear modulus* (GPa)	*Poisson's ratio*	*Axial tensile strength* (MPa)	*Tranverse tensile strength* (MPa)	*Axial compressive strength* (MPa)	*Tranverse compressive strength* (MPa)	*Inplane shear strength* (MPa)
E-glass	23	23	9.0	0.28	550	550	330	330	250
Aramid	29	29	11	0.32	460	460	190	190	65
Boron	80	80	30	0.33	480	480	1100	1100	360
SM carbon (PAN)	54	54	21	0.31	580	580	580	580	410
IM carbon (PAN)	63	63	21	0.31	1350	1350	580	580	410
UHM carbon (PAN)	110	110	41	0.32	490	490	270	270	205
UHM carbon (pitch)	165	165	63	0.32	310	310	96	96	73
UHK carbon (pitch)	165	165	63	0.32	310	310	96	96	73

Table 6 Physical properties of selected unidirectional polymer matrix composites (fiber volume fraction = 60%)

Fiber	*Density* ($g\,cm^{-3}$)	*Axial CTE* ($10^{-6}\,K^{-1}$)	*Transverse CTE* ($10^{-6}\,K^{-1}$)	*Axial thermal conductivity* ($W\,m^{-1}\,K^{-1}$)	*Transverse thermal conductivity* ($W\,m^{-1}\,K^{-1}$)
E-glass	2.1	6.3	22	1.2	0.6
Aramid	1.38	−4.0	58	1.7	0.1
Boron	2.0	4.5	23	2.2	0.7
SM carbon (PAN)	1.58	0.9	27	5	0.5
IM carbon (PAN)	1.61	0.5	27	10	0.5
UHM carbon (PAN)	1.66	−0.9	40	45	0.5
UHM carbon (pitch)	1.80	−1.1	27	380	10
UHK carbon (pitch)	1.80	−1.1	27	660	10

Table 7 Physical properties of selected unidirectional quasi-isotropic polymer matrix composites (fiber volume fraction = 60%)

Fiber	*Density* ($g\,cm^{-3}$)	*Axia CTE* ($10^{-6}\,K^{-1}$)	*Transverse CTE* ($10^{-6}\,K^{-1}$)	*Axia thermal conductivity* ($W\,m^{-1}\,K^{-1}$)	*Transverse thermal conductivity* ($W\,m^{-1}\,K^{-1}$)
E-Glass	2.1	10	10	0.9	0.9
Aramid	1.38	1.4	1.4	0.9	0.9
Boron	2.0	6.5	6.5	1.4	1.4
SM carbon (PAN)	1.58	3.1	3.1	2.8	2.8
IM carbon (PAN)	1.61	2.3	2.3	6	6
UHM carbon (PAN)	1.66	0.4	0.4	23	23
UHM carbon (pitch)	1.80	−0.4	−0.4	195	195
UHK carbon (pitch)	1.80	−0.4	−0.4	335	335

the same materials as in **Table 6**. Here again, a fiber volume fraction of 60% is assumed.

It is to be noted that the CTEs of the quasi-isotropic composites are higher than the axial values of the corresponding unidirectional composites. However, the CTEs of quasi-isotropic composites reinforced with aramid and carbon fibers are still very small. By appropriate selection of fiber, matrix, and fiber volume fraction, it is possible to obtain quasi-isotropic materials with CTEs very close to zero. The through-thickness CTEs of these laminates are positive and relatively large. However, this is not a significant issue for most applications. One exception is optical mirrors, for which through-thickness CTE is an important issue.

The inplane thermal conductivity of quasi-isotropic laminates reinforced with UHM pitch carbon fibers is similar to that of aluminum alloys, while UHK pitch carbon fibers provide laminates with a conductivity over 50% higher. Both materials have densities ~35% lower than that of aluminum.

As mentioned earlier, through-thickness thermal conductivities of laminates tend to be similar to the transverse thermal conductivities of unidirectional composites, which are relatively low. If laminate thickness is small, this may not be a severe limitation. However, low through-thickness thermal conductivity can be a significant issue for thick laminates and for very high thermal loads. Through-thickness conductivity can be increased by addition of thermally conductive reinforcements, such as disc. carbon fibers, and ceramic particles, such as boron nitride or carbon nanotubes. Fiber-reinforced thermoset PMCs are made by a wide variety of processes, many of which are highly automated, including hand lay-up, filament winding, tape placement, fiber placement, and various types of resin transfer molding.

A significant recent advance in PMC technology is the development of injection moldable carbon fiber-reinforced thermoplastics with much higher thermal conductivities than those available in the past. Unreinforced polymers have thermal conductivities $\sim 0.2\,W\,m^{-1}\,K^{-1}$. A number of commercially available PMCs consisting of thermoplastic matrices reinforced with disc. carbon fibers have reported thermal conductivities ranging from $2\,W\,m^{-1}\,K^{-1}$ to as high as $100\,W\,m^{-1}\,K^{-1}$. Matrices include PPS, nylon 6, polycarbonate, and liquid crystal polymers. These composites are also electrically conductive. Electrically insulating PMCs reinforced with thermally conductive disc. ceramic particles, such as boron nitride, have reported thermal conductivities of up to $15\,W\,m^{-1}\,K^{-1}$.

Metal Matrix Composites

MMCs consist of metals reinforced with a variety of ceramic fibers, carbon fibers, whiskers, and particles. There are wide ranges of materials that fall in this category. An important example cited earlier is a material consisting of tungsten carbide particles embedded in a cobalt matrix, which is used extensively in cutting tools and dies. This composite (often referred to as cermet, cemented carbide or simply, but incorrectly as "tungsten carbide") has much better fracture toughness than monolithic tungsten carbide, which is a brittle ceramic material.

Another interesting MMC, tungsten carbide particle-reinforced silver, is a key circuit breaker contact pad material. Here, the composite provides good electrical conductivity and much greater hardness and wear resistance than monolithic silver, which is too soft to be used in this application. Ferrous alloys reinforced with titanium carbide particles have been used for many years in numerous aerospace and commercial production applications, including dies, engine valves, and aircraft fuel pumps. Compared to the monolithic base metals, they offer better wear resistance, higher stiffness, and lower density.

MMCs are also used in automobile engine blocks. In one design, the cylinder walls of an aluminum engine block are reinforced with a combination of disc. aluminum oxide (alumina) and carbon fibers, enabling elimination of cast iron cylinder liners. MMCs have also been used in high-speed electronics manufacturing equipment and in the equipment used for production of microprocessor chips, such as photolithography tables. Other applications include diesel engine pistons, aircraft structures, exit-guide vanes of aircraft engine fan, actuators, and automobile and train brake rotors.

The most important uses for MMCs is in electronic packaging and thermal management. For example, silicon carbide particle-reinforced aluminum, often called Al–SiC in the electronics industry, is being used in high volume production parts, such as microprocessor lids and power modules for hybrid electric vehicles. Other MMCs used in packaging are carbon fiber-reinforced aluminum, beryllium oxide (beryllia) particle-reinforced beryllium and silicon–aluminum. Here, the advantages are high stiffness, high thermal conductivity and low density, and CTE. Two traditional packaging materials, copper–tungsten and copper–molybdenum, are also MMCs. The CTEs of all of these composites can be tailored by varying the ratio of the two constituents. Another major advantage of the newer composites is that they have relatively low densities.

Monolithic metallic alloys are among the most widely used structural materials. By reinforcing them with cont. fibers, disc. fibers, whiskers, and particles, new materials are created with enhanced or modified properties, such as higher strength and stiffness, better wear resistance, and lower CTE. In some cases, the improvements are dramatic.

The greatest increases in strength and modulus are achieved with cont. fibers, at least in the direction parallel to the fibers, called the axial or longitudinal direction. As for PMCs, transverse properties are dominated by the properties of the matrix and interface. However, because the metal matrices are in themselves structural materials, transverse strength properties are frequently great enough to permit use of unidirectional MMCs in some structural applications. This is usually not possible for PMCs. The boron fiber-reinforced aluminum struts used on space shuttle "Orbiter" are a good example. Other key MMCs reinforced with cont. fibers include silicon carbide-reinforced titanium and carbon fiber-reinforced aluminum. The key particle-reinforced MMCs, include titanium carbide-reinforced steel, aluminum reinforced with silicon carbide and alumina particles, titanium carbide particle-reinforced titanium, and titanium boride-reinforced titanium.

Aluminum reinforced with silicon carbide particles is arguably the most important of the newer types of structural MMCs. The low cost of the aluminum matrix and silicon carbide particles makes these composites particularly interesting. There is a wide range of materials falling into this category. They are made by a variety of processes. Properties depend on the type of particle, particle volume fraction, matrix alloy, and the process used to make them. **Table 8** presents representative composite properties for three particle volume fractions, 25%, 55%, and 70%. Properties of common steel, aluminum, and titanium alloys are shown for comparison.

It is seen that as particle volume fraction increases, modulus and yield strength increase, and fracture toughness, tensile ultimate strain, and CTE of particle-reinforced composites decrease. Particulate reinforcement also improves the elevated temperature strength properties and, perhaps surprisingly, the fatigue resistance. The ability to tailor CTE by varying particle volume fraction is a key attribute of these materials.

There are a variety of processes to make silicon carbide particle-reinforced aluminum, including powder metallurgy, stir casting, and pressure and pressureless infiltration. The last two, as well as remelt casting can make net shape or near-net shape parts. Fiber-reinforced MMCs are made by a variety of processes, including pressure infiltration and diffusion bonding.

Table 8 Properties of silicon carbide particle-reinforced aluminum, aluminum, titanium and steel

Property	*Aluminum 6061-T6*	*Titanium 6Al-4V*	*Steel 4340*	*Composite particle volume fraction* (%)		
				25	*55*	*70*
Modulus (GPa)	69	113	200	114	186	265
Tensile yield strength (MPa)	275	1000	1480	400	495	
Tensile ultimate strength (MPa)	310	1100	1790	485	530	225
Elongation (%)	15	5	10	3.8	0.6	0.1
Specific modulus (GPa)	5	26	26	40	63	88
CTE ($10^{-6}\,K^{-1}$)	23	9.5	12	16.4	10.4	6.2
Density (g cm^{-3})	2.77	4.43	7.76	2.88	2.96	3.00

Carbon Matrix Composites

CAMCs consist of a carbon matrix reinforced with any combination of fibers, whiskers, or particles. For many years, the only significant CAMCs were CCCs, in which the reinforcements are disc. or cont. carbon fibers. In the last few years, a new proprietary carbon matrix material system has been developed which has a silicon carbide fiber-reinforcement. This material is now being used for engine flaps on a military aircraft engine. One of the key reported advantages of this new material is that it has a higher CTE than CCCs, reducing the tendency of protective ceramic coatings to crack. The focus of this section is on CCCs.

CCCs are used in a variety of applications, including electronic packaging, spacecraft radiator panels, rocket nozzles, reentry vehicle nose tips, the space shuttle "Orbiter" leading edges and nose cap, aircraft brakes, heat treating furnaces, and glass making equipment.

As for PMCs, there are many different CCC materials having widely different mechanical and physical properties. The primary advantages of CCCs are: (1) high strength compared to competing materials at very high temperatures, (2) high stiffness, (3) ablation resistance, (4) high thermal conductivity (some systems), (5) low CTEs, (6) low density, and (7) absence of outgassing. In addition, CCCs are less brittle than monolithic carbon.

The primary disadvantages are: (1) Susceptibility to oxidation at temperatures above $\sim$37–500°C (700–930°F), (2) low interlaminar (through-thickness) tensile and shear strengths for materials with 2D reinforcement, (3) microcracking at low stresses in some directions for 3D composites, and (4) high cost of many systems. Because of the low interlaminar strength properties of CCCs, many applications, particularly those with thick walls, often use 3D reinforcement.

As mentioned earlier, one of the most significant limitations of CCCs is oxidation. Addition of oxidation inhibitors to the matrix and protective coatings raises the threshold substantially. In inert atmospheres, CCCs retain their properties to temperatures as high as 2400°C (4300°F).

The combination of high thermal conductivity and low density makes CCCs attractive candidates for thermal management and electronic packaging. In addition, CCCs have very low CTEs, leading to their use as thermal doublers with carbon fiber-reinforced PMC structures. The unique combination of properties possessed by CCCs, combined with a lack of outgassing, also makes them attractive for optical systems.

Two leading types of processes are used to make CAMCs. The first type is the chemical vapor infiltration (CVI) which is a process in which gaseous chemicals are reacted or decomposed, depositing a solid material on a fibrous preform. In the case of CAMCs, hydrocarbon gases such as methane and propane are broken down, and the material deposited is the carbon matrix. The second type of process involves infiltration of a preform with polymers or pitches, which is then converted to carbon by pyrolysis, (heating in an inert atmosphere). After pyrolysis, the composite is sometimes heated to high temperatures to graphitize the matrix. To minimize porosity, the process is repeated until a satisfactory density is achieved. This is called densification. Common matrix precursors are phenolic and furan resins, and pitches derived from coal tar and petroleum.

Ceramic Matrix Composites

As a class of materials, monolithic ceramics are characterized by high stiffness and hardness, resistance to wear, corrosion and oxidation, and high-temperature operational capability. However, they also have serious deficiencies, which have severely limited their use in applications that are subjected to significant tensile stresses.

A fundamental problem is that ceramics have very low fracture-toughnesses, which makes them very sensitive to the presence of small flaws. This results in

great strength scatter and poor resistance to thermal and mechanical shock. Civil engineers recognized this deficiency long ago, and do not use ceramic materials such as stone and concrete to carry tensile loads. In the latter, this function has been relegated to reinforcing bars or prestressing cables made of steel or, more recently, PMCs. An important exception has been in lightly loaded structures where dispersed reinforcing fibers of asbestos, steel, glass, and carbon allow modest tensile stresses to be supported.

CMCs can be thought of as an improved form of CAMCs in which the carbon matrix is replaced with ceramics that are stronger and much more resistant to oxidation. CMCs employ a variety of reinforcements including cont. fibers, disc. fibers, whiskers, and particles. Cont. fibers provide the best properties. There are many different types of CMCs, which are at various stages of development. As discussed earlier, straw-reinforced mud is an ancient CMC, as is concrete, which consists of a cement matrix reinforced with stone and sand.

The key advantage of CMCs is that, when properly designed and manufactured, they have many of the advantages of monolithic ceramics, such as much lower density than high-temperature metals, but with greater durability. That is, CMCs have higher effective fracture-toughnesses, so that they are less susceptible to failure when subjected to mechanical and thermal shock. As a consequence, it is possible to consider CMCs for applications where they are subjected to moderate tensile loads. However, CMCs are the most complex of all types of composites, and CMC technology is less developed than that of PMCs, MMCs and CAMCs.

CMCs are being used in a number of commercial production applications. One of the most successful is silicon carbide whisker-reinforced alumina cutting tool inserts, which have greater fracture-toughness, and are, therefore, more durable than monolithic alumina. Another application is silicon carbide whisker-reinforced aluminum nitride crucibles, which are used for casting molten aluminum. In this application, the key advantage of the CMC over monolithic ceramics is thermal shock resistance. Silicon carbide particle-reinforced alumina is being used in slurry pumps because of its good durability and wear resistance. In this application, the process makes it possible to fabricate reliable, complex parts that would be hard to make out of monolithic ceramics. Other high-temperature CMC applications include coal-fired power plant candle filters used for particulate removal, natural gas burner elements, and U-tubes. In addition, there are a wide variety of candidate applications including stationary gas turbine combustor liners and shrouds, abradable rim seals, reverberatory screens, particle separators, tube shields, recuperators, turbine tip shoes, pipe hangers, heat treating furnace fans, hot gas filters, and natural gas burner elements.

Aerospace applications of ceramic matrix composites to date have been limited. Perhaps, the most significant are fighter aircraft engine flaps. There are two types, both using silicon carbide matrices. One is reinforced with carbon fibers, and the other a multifilament silicon carbide fiber. Another application is a missile diverter thruster made of carbon fiber-reinforced silicon carbide. Again, the process used to make this part is CVI. The space shuttle "Orbiter" thermal protection system (TPS) makes extensive use of tiles composed of a 3D network of disc. oxide fibers with silicate surface layers. While there is no continuous matrix for most of the tile, the surface region is a form of CMC. In a sense, this can be considered as a type of functionally graded material.

The addition of cont. fibers to a ceramic matrix can significantly change failure modes. Monolithic ceramics have linear stress–strain curves, and fail catastrophically at low strain levels. However, well-designed and fabricated CMCs display nonlinear stress–strain behavior with much more area under the curve, indicating that more energy is absorbed during failure, and that the material has a less catastrophic failure mode.

Reinforcements that have been used for CMCs include cont. fibers, disc. fibers, whiskers, and particles. Key cont. fibers used in CMCs include carbon, silicon carbide-based, alumina-based, alumina-boria-silica, quartz, and alkali-resistant glass. Steel wires are also used. Disc. CMC fibers are primarily silica-based. Silicon carbide is the key whisker reinforcement. Particulate reinforcements include silicon carbide, zirconium carbide, hafnium carbide, hafnium diboride, and zirconium diboride.

A large number of ceramics have been considered for matrix materials, including alumina, glass, glass-ceramic, mullite (aluminum silicate), cordierite (magnesium aluminosilicate), yttrium alumina garnet (YAG), barium aluminosilicate (BAS), barium magnesium aluminosilicate (BMAS), calcium aluminosilicate (CAS), barium and strontium aluminosilicate (BSAS, or celsian), "Blackglas" (silicon oxycarbide or Si–O–C), silicon nitride, silicon carbide, silicon nitride-bonded silicon carbide, silicon carbide and silicon, hafnium carbide, tantalum carbide, zirconium carbide, hafnium diboride, zirconium diboride, and molybdenum disilicide.

The most mature CMCs consist of silicon carbide matrices reinforced with silicon carbide-based fibers (SiC–SiC) and silicon carbide reinforced with carbon fibers (C–SiC).

As for other classes of composite materials, there are many processes that can be used to make CMCs. Key considerations in process selection are porosity and reactions between reinforcements, reinforcement coatings, and matrices. The most important processes for making CMCs at this time are CVI, melt infiltration, preceramic polymer infiltration and pyrolysis (PIP), slurry infiltration, sol-gel, hot pressing, and hot isostatic pressing. In addition, there are a number of reaction based processes, which include reaction bonding and direct metal oxidation ("Dimox").

Thermal Management Materials

One of the most significant new areas for composites is in thermal management. There are many engineering uses of these materials, but perhaps, the most important is in the packaging of microelectronics, optoelectronics, and MEMS. Packaging provides support and protection to semiconductors and ceramics, which typically have low CTEs, and removes heat by conduction. The material requirements for these applications are high thermal conductivity, low CTE to minimize thermal stresses, and low density to minimize weight. Composites offer advantages in all the three areas.

Table 9 presents properties of traditional packaging materials, including thermal conductivity, CTE, density, and specific thermal conductivity, which is defined as the ratio of density to specific gravity. Specific thermal conductivity is a useful figure of merit where both thermal conductivity and weight are important. It is seen that, aside from CVD diamond (diamond produced by chemical vapor deposition), the maximum thermal conductivity is for copper, which has a high CTE. All materials with low CTEs have thermal conductivities that are no better than that of aluminum. In addition, all of them have relatively high densities. These are all significant deficiencies. Increasing heat fluxes has resulted in a need for improved materials.

Table 10 shows properties of some of the increasing number of new monolithic and composite packaging materials that have been developed in response to the universally recognized, thermal management problems. A number of these materials are being used in production applications. Reinforcements include cont. and disc. thermally conductive carbon fibers, and a variety of particles, including diamond, silicon carbide, and beryllia. Monolithic materials, all of which are carbonaceous, include carbon foam, natural graphite, and highly oriented pyrolytic graphite (HOPG). It is seen that these materials offer great advantages over traditional packaging materials.

The Future of Composite Materials

Composites are now baseline materials in countless applications. However, the technology, which is only several decades old, is still in its infancy. In the future, significant improvements in properties of existing fibers, matrices and processes, and development of new composites are likely to be seen. So far, composites have been used primarily for their excellent mechanical properties, environmental resistance, and durability. This is likely to continue. In addition, an increasing use in applications for which nonmechanical properties, such as thermal conductivity and CTE are important is foreseen. Among these are packaging of microelectronics, optoelectronics, and MEMS. These are extremely large and growing markets.

Development is under way on multifunctional and smart composites, which incorporate electronics,

Table 9 Properties of traditional packaging materials

Reinforcement	*Matrix*	*Thermal cond.* ($W\,m^{-1}\,K^{-1}$)	*CTE* ($ppm\,K^{-1}$)	*Specific gravity*	*Specific thermal cond.* ($W\,m^{-1}\,K^{-1}$)
	Aluminum	218	23	2.7	81
	Copper	400	17	8.9	45
	CVD Diamond	1100–1800	1–2	3.52	310–510
	Invar	11	1.3	8.1	1.4
	Kovar	17	5.9	8.3	2.0
	C-I-C	164	8.4	8.4	20
	C-Mo-C	182	6.0	9.9	18
	Titanium	7.2	9.5	4.4	1.6
Copper	Tungsten	157–190	5.7–8.3	15–17	9–13
Copper	Molybdenum	184–197	7.0–7.1	9.9–10.0	18–20
	Solder – Sn63/Pb37	50	25	8.4	6.0
	Epoxy	1.7	54	1.2	1.4
E-glass fibers	Epoxy	0.16–0.26	11–20	2.1	0.1

Table 10 Properties of new monolithic and composite packaging materials

Reinforcement	*Matrix*	*Thermal cond.* ($W\,m\,K^{-1}$)	*CTE* ($ppm\,K^{-1}$)	*Specific gravity*	*Specific thermal cond.* ($W\,m\,K^{-1}$)
	Carbon foam	135–245	−1	0.6–0.9	220–270
	HOPG	1300–1700	−1.0	2.3	740–850
	Natural graphite	150–500			
Invar	Silver	153	6.5	8.8	17
Cont. carbon fibers	Aluminum	218–290	−1–+16	2.3–2.6	84–126
Disc. carbon fibers	Aluminum	185	6.0	2.5	74
Disc. carbon fibers	Polymer	20–290	4–7	1.6–1.8	12–160
Silicon	Aluminum	126–160	6.5–17	2.5–2.6	49–63
SiC particles	Aluminum	170–220	6.2–16.2	3.0	57–73
Beryllia particles	Beryllium	240	6.1	2.6	92
Natural graphite	Epoxy	370	−2.4	1.94	190
Cont. carbon fibers	Polymer	330	−1	1.8	183
Disc. carbon fibers	Copper	300	6.5–9.5	6.8	44
SiC particles	Copper	320	7–10.9	6.6	48
Cont. carbon fibers	SiC	370	2.5	2.2	170
Cont. carbon fibers	Copper	400–420	0.5–16	5.3–8.2	49–79
Cont. carbon fibers	Carbon	400	−1.0	1.9	210
Cont. carbon fibers	SiC	370	2.5	2.2	170
Graphite flake	Aluminum	400–600	4.5–5.0	2.3	174–260
Diamond particles	Aluminum	550–600	7.0–7.5	3.1	177–194
Diamond & SiC particles	Aluminum	575	5.5		
Diamond particles	Copper	600–1200	5.8	5.9	330–670
Diamond particles	Cobalt	>600	3.0	4.12	>145
Diamond particles	Magnesium	550	8		
Diamond particles	Silicon	525	4.5		
Diamond particles	SiC	600	1.8	3.3	182

sensors, actuators, and microprocessors. Potential applications include everything from spacecraft to cell phones.

Nanotechnology is one of the most exciting and challenging areas for composites. Polymers reinforced with low cost silicate (clay) nanoparticles are already being used in commercial production applications. Carbon nanotubes and graphite nanoplatelets have impressive mechanical, thermal, and electrical properties that make them attractive candidate reinforcements. In addition to these trends, it is likely that important new materials and applications will emerge.

See also: Alloys: Aluminum; Alloys: Copper; Alloys: Overview; Alloys: Titanium; Ceramic Materials; Ceramics, History of; Composites: Carbon Matrix Composites; Composites: Polymer–Polymer; Conductivity, Thermal; Glasses; Highly Porous Metals and Ceramics; Mechanical Properties: Elastic Behavior; Mechanical Properties: Tensile Properties; Meso- and Nanostructures; Polymer Structures; Powder Processing: Models and Simulations; Semiconductor Nanostructures; Thermal Conductivity at Low Temperatures; Vapor Transport Processing: Mechanisms.

PACS: **81.05.Ni; 81.05.Pj; 81.05.Qk**

Further Reading

Bader MG, Kedward KK, and Sawada Y (2000) Design and applications. In: Kelly A and Zweben C (eds.) *Comprehensive Composite Materials*, vol. 6. Oxford: Pergamon, Elsevier Science.

Carlsson L, Crane RL, and Uchino K (2000) Test methods, nondestructive evaluation, and smart materials. In: Kelly A and Zweben C (eds.) *Comprehensive composite Materials*, vol. 5. Oxford: Pergamon, Elsevier Science.

Chou T-W (2000) Fiber reinforcements and general theory of composites. In: Kelly A and Zweben C (eds.) *Comprehensive Composite Materials*, vol. 1. Oxford: Pergamon Press, Elsevier.

Clyne TW (2002) Metal matrix composites. In: Kelly A and Zweben C (eds.) *Comprehensive Composite Materials*, vol. 3. Oxford: Pergamon, Elsevier Science.

Hayashi T, Ushio H, and Ebisawa M (1989) The properties of hybrid fiber reinforced metal and its application for engine block. *SAE Technical Paper No. 890557*, Warrendale, Pennsylvania, USA: Society of Automotive Engineers.

Kelly A (1994) Introduction. In: Kelly A (ed.) *Concise Encyclopedia of Composite Materials*, revised edn. Oxford: Pergamon.

Ko FK (1993) Advanced textile structural composites. In: Moran-Lopez IJ and Sanchez JM (eds.) *Advanced Topics in Materials Science and Engineering*. New York: Plenum.

Rawal SP, Misra, MS, and Wendt RG (1990) *Composite Materials for Space Applications*. NASA CR-187472. Hampton, Virgina: National Aeronautics and Space Administration.

Savage G (1993) *Carbon–Carbon Composites*. London: Chapman and Hall.

Talreja R (2000) Polymer matrix composite. In: Kelly A and Zweben C (eds.) *Comprehensive Composite Materials*, vol. 2. Oxford: Pergamon, Elsevier Science.
Warren R (2000) Carbon/carbon, cement, and ceramic matrix composites. In: Kelly A and Zweben C (eds.) *Comprehensive Composite Materials*, vol. 4. Oxford: Pergamon, Elsevier Science.
Zweben C (1994) Metal matrix composites: aerospace applications. In: Flemings MC (Eds.) *Encyclopedia of Advanced Materials*. Pergamon.
Zweben C (1998) Composite materials and mechanical design. In: Kutz M (ed.) *Mechanical Engineers' Handbook*, 2nd edn. New York: Wiley.
Zweben C (2001) Heat sink materials for electronic packaging. In: Buschow KHJ (ed.) *The Encyclopedia of Materials: Science and Technology.* Cambridge: Elsevier Science.
Zweben C (2002) Metal matrix composites, ceramic matrix composites, carbon matrix composites and thermally conductive polymer matrix composites. In: Harper J (ed.) *Handbook of Plastics, Elastometers and Composites*, 4th edn., pp. 321–344. New York: McGraw-Hill.
Zweben C (2002) Thermal management and electronic packaging applications. In: *ASM Handbook*, vol. 21, Composites. Materials Park, Ohio: ASM International.

Composites: Polymer–Polymer

R Talreja, Texas A&M University, College Station, TX, USA

Introduction

Alongside metals and ceramics, polymers form a main class of materials that offer a range of interesting properties for many engineering applications. Polymers in a monolithic or single-phase form may not possess the desired properties for a given application. They can be modified to have two phases or combined with other polymers to construct composites. Usually, the term polymer–polymer composite (PPC) refers to the case of rigid, rod-like polymer molecules dispersed at a molecular level in a flexible coil-like polymer. The rigid part is thus viewed as reinforcement and the flexible part as matrix. This article also includes the class of composites where polymer fibers are combined with a polymer matrix. Such composites are designated as fiber-reinforced polymer–polymer composites (FRPPC). To begin with, the polymer microstructure is briefly reviewed.

The Polymer Microstructure

A polymer is composed of molecules consisting of long sequences of atoms linked by covalent bonds. Most polymers have a linear skeletal structure, which may be represented by a chain with two ends. This gives the name "linear polymers." However, many polymers have a branched structure in which the main chain has side chains, or branches, or a three-dimensional network structure, in which the individual chains are connected to other chains at junction points. The network polymers are said to be cross-linked, and such nonlinear polymers can be formed by polymerization. The skeletal structure determines the properties of a polymer.

The polymers are commonly classified as thermosets, thermoplastics, and elastomers. Thermosets are normally rigid network polymers in which the chain motion is restricted by a high degree of cross-linking. Heating normally activates the cross-linking agent, which is the reason for calling them thermosets. Once formed, the thermosets tend to degrade rather than melt on application of heat. Thermoplastics are linear or branched polymers, which can be melted by application of heat. On cooling they solidify, allowing molding and remolding to different shapes. The cool-down from the molten state does not always cause crystallization in thermoplastics. In cases where crystallization occurs, it is often partial, giving the name "semicrystalline thermoplastics." The amorphous structure transition from a glassy (hard) state to a rubbery (soft) state occurs at the glass transition temperature (T_g). Elastomers are cross-linked polymers in a rubbery state. They are characterized by large, rapidly recoverable deformation. Elastomers, like thermosets, are intractable once formed and degrade rather than melt on heating.

Other variations of polymers exist. "Liquid crystalline polymers" have molecular alignment in one direction, but not a three-dimensional crystalline order. In recent years, such polymers have been developed with distinct regions of straight and rigid molecular chains. "Blends" are physical mixtures of two or more polymers, while "grafts" are formed by chemically attaching long side chains of a second polymer to the molecular chains of the base polymer. Copolymers have chemical combinations of molecules in the main chain. The set of molecules that repeat in the chain can be arranged in blocks, giving the designation "block copolymers," or be randomly distributed in a "random copolymer."

PPC Systems – Processing, Structure, and Properties

The rod-like molecules in PPC systems have high melting point, which makes it difficult to use melt processing for producing a single-phase material with a homogeneous dispersion of rigid molecules. Ternary solutions (a solvent, a rod-like solute, and a random-coil chain) have been shown theoretically to be good means of achieving this. This process requires a high degree of molecular orientation of the rod-like molecules in the random-coil matrix. Since the concentration of rod-like molecules needed is at or near the critical value to form anisotropic domains, the resulting viscosity of the solution is high, which makes orientation of the molecules by external shear suitable. Alternatives to physical blending are possible, for example, synthesis of block copolymers composed of flexible-coil segments and rigid-rod block segments.

The intended microstructure of aligned rigid rod-like molecules within the amorphous flexible coil-like matrix has been achieved in some PPC systems. For instance, a blend of a rod-like poly(*p*-phenylene benzobisthiazole) (PBT) and coil-like poly(2-5(6)-benzimidizole) (ABPBI) vacuum-cast from solution with up to 50 wt.% of PBT typically shows 3 μm long ellipsoids composed mainly of 10 nm PBT crystallites moderately well-aligned with the long axis. Similar morphology results in a triblock copolymer of ABPBI/PBT/ABPBI, but with smaller ellipsoids.

The rod-like molecules of polymers can potentially provide higher stiffness and strength than fibers that tend to have imperfect alignment and flaws. Furthermore, in an FRPPC imperfect adhesion between fibers and matrix can be a source of reduced properties. **Table 1** illustrates the tensile properties of PBT and ABPBI fibers and composites of these polymers. As shown in the data, better modulus and strength values are achieved in copolymers than in blends due to the smaller size of the PBT ellipsoids, leading to higher reinforcement efficiency. The fracture studies show that the PBT fiber fibrillates extensively upon fracture, while the PBT/ABPBI blend fibrillates moderately and the copolymer shows no evidence of fibrillation.

FRPPCs

Using polymer fibers as reinforcement in polymer matrix develops a wide range of polymer composites. Several high-strength polymer fibers have been developed for this purpose in recent years. The most widely used FRPPCs are aramid fibers in epoxy, polyethylene fibers in epoxy, and high-density polyethylene fibers in polyethylene.

Processing of polymers to produce high-modulus and high-strength fibers has two generic routes: (1) spinning of stiff, nematogenic polymers to form oriented structures in solid state, and (2) morphological manipulation of flexible polymers into highly oriented extended-chain fibers through complex processing. Aramid fibers are examples of the first processing route. They are spun from liquid-crystal dopes through a dry-jet wet-spinning process. Polyethylene fibers are produced by a gel-spinning process in which a low-concentration solution of ultrahigh molecular weight polyethylene is extruded to form a gel precursor fiber, which is then hot-drawn to draw ratios of 20 or more to produce a very high-oriented fiber with an extended-chain fibrillar structure. Other linear polymers such as polypropylene, polyoxymethylene, and aliphatic polyamides, nylon 6, and nylon 66, have also been used to produce high-modulus, high-strength fibers. These polymers have melt temperatures higher than polyethylene. The molten polymer is spun into filaments, which are then drawn to fibers of 1.5–6 times their undrawn length. Poly(ethylene terephthalate) (PET) fibers are also produced by this process. To improve properties, modification to PET is done by synthesizing copolyesters, which contain typically 90 mol.% PET and small amounts of ester-forming ingredients such as glycols and dicarboxylic acids.

The high degree of molecular orientation in the axial direction of polymer fibers produces high tensile modulus and strength in that direction. Typical values for some fibers are listed in **Table 2**. As shown in the table, the transverse strength values are low. The tensile transverse modulus is also low, typically less than 10% of the fiber-axis tensile modulus. The axial compressive strength is also much lower than

Table 1 Properties of PBT and ABPBI fibers and composites

Material	*Tensile modulus* (GPa)	*Tensile strength* (MPa)	*Strain to fracture* (%)
PBT fiber	320	3100	1.1
ABPBI fiber	36	1100	5.2
30%PBT/70% ABPBI blend film	1.1	35	5.6
30%PBT/70% ABPBI blend fiber	120	1300	1.4
30%PBT/70% ABPBI copolymer film	2.4	220	43
30%PBT/70% ABPBI copolymer fiber	100	1700	2.4

Table 2 Typical modulus and strength values of high-modulus, high-strength organic fibers

Fiber	*Tensile modulus* (GPa)	*Tensile strength* (GPa)	*Transverse strength* (GPa)
Gel PE	140	2.7	0.1
Kevlar 49	120	3.4	0.6
Kevlar 29	60	3.4	0.6
Thermotropic copolyester	60	3.0	0.1
Ordered polymer	265	3.0	1.5
Cellulose	25	1.5	N/A

the axial tensile strength. This limits the use of FRPPC in applications where compression and/or bending are involved. The molecular orientation in polymer fibers is also the cause of poor abrasion resistance.

Aramid fibers have a characteristic fibrillar structure. The Kevlar fibers exhibit a skin-core configuration with skin displaying a higher degree of fibrillar orientation than the core. This structure yields high tensile properties due to the axial fibrillar orientation but is not capable of sustaining large compressive stress due to buckling of the fibrils.

Polymer Fiber–Polymer Matrix Interfaces

One way to achieve good bonding at a polymer–polymer interface is by interdiffusion of molecular chains to develop entanglements of the two polymers. Due to the high molecular orientation along the axis of the polymer fiber, however, this is usually difficult to achieve. Heating the fiber to increase the polymer chain mobility often results in relaxation of the oriented molecules leading to degradation of fiber-axis properties. An alternative way is to use covalent, hydrogen and van der Waals interactions, or mechanical interlocking at the fiber–matrix interface. Effective bonding requires good wetting of the fiber surface by the matrix polymer during the composite processing. The surface energy of the matrix polymer should be lower than that of the fiber polymer to achieve good wetting. As examples, the surface energies of epoxy, polyethylene, polypropylene, and polyamide are 43, 33, 35, and 39 $mJ\,m^{-2}$, respectively. Another way of improving the interfacial bonding is by surface treatment of fibers. Typical surface treatments are plasma oxidation, fluorination, and acid or alkali wash. The plasma oxidation treatment of polyethylene fibers has shown an improvement of the interlaminar shear strength of epoxy reinforced with these fibers from 10 MPa in untreated state to 27 MPa. The accompanying reduction of tensile strength is found to be small. Still another way to improve fiber–matrix bonding in FRPPCs is by creating an interfacial region (interphase) of good properties. It is, for instance, possible in semicrystalline thermoplastic matrix composites to nucleate crystallization from the fiber surface. The transcrystalline layer formed tends to have laminae oriented normal to the fiber axis, resulting in good transverse strength.

Processing of FRPPCs

A common way to produce FRPPCs is by impregnation of fibers with the polymer matrix. The objectives of impregnation are to wet all fibers by the matrix and distribute the fibers uniformly in the matrix. If fiber bundles are used, the matrix should penetrate into the bundles and wet all fibers. Incomplete or imperfect impregnation causes dry fibers, resin-rich areas, and uneven distribution of fibers, leading to lower composite properties. Three common methods of impregnation exist, viz., by using melt, powder, or solution. In melt impregnation, the fibers (or fiber bundles) are fixed in a mold and the molten resin is transferred to the mold, often under pressure. Alternatively, the matrix polymer is used as a film, which is interspersed with layers or bundles of fibers. In this case, the mold is heated to melt the matrix polymer. The melt impregnation is not very effective in penetrating fiber bundles and wetting fibers if the melt viscosity is high. In powder impregnation, the matrix polymer is used as a fine powder. The particle size of the powder should be in the range of fiber diameter, that is 10–25 μm. This method works better at lower fiber volume fraction since at high fiber volume fraction it is difficult to uniformly distribute the powder between fibers. The third method is solution impregnation where the matrix polymer is dissolved in a liquid. For example, for polyethylene (PE) matrix a xylene solution with 5 wt.% PE is used. The impregnated fiber bundles are consolidated into a unidirectional prepreg, which is then laminated to produce a composite laminate. The deficiencies of this method are segregation and crystallization of matrix on fiber surfaces. These deficiencies can be partly overcome by using suspension impregnation where particles of matrix polymer are suspended in a solution. The solution is stirred during impregnation to prevent particle segregation.

Table 3 Typical properties of unidirectional glass fiber-reinforced epoxy (GFRP), carbon fiber-reinforced epoxy (CFRP), and two aramid fiber-reinforced epoxies (AFRP). The designation 29 and 49 are for Kevlar 29 and Kevlar 49 fibers, respectively. Fiber volume fraction is 60% in all cases.

Material	*Density* ($\times 10^3$ Kg m^{-3})	*Axial tensile modulus* (GPa)	*Trans. tensile modulus* (GPa)	*Shear modulus* (GPa)	*Axial tensile strength* (MPa)	*Trans. tensile strength* (MPa)	*Axial compr. strength* (MPa)	*Inter-laminar shear strength* (MPa)
GFRP	2.0	40	10	4.5	780	28	480	40
CFRP	1.5	140	6.9	5.0	1620	34	1200	80
AFRP-29	1.4	50	5.0	3.0	1350	30	238	44
AFRP-49	1.4	76	5.6	2.8	1380	30	276	60

Properties of FRPPCs

The mechanical properties of polymers reinforced with carbon or glass fibers are generally superior to FRPPC systems. A comparison of typical values obtained in a 60% fiber volume fraction unidirectional lamina for a high-strength carbon fiber-reinforced epoxy, an E-glass fiber-reinforced epoxy, and two aramid FRPPCs (Kevlar 29 and Kevlar 49) is shown in **Table 3**. As illustrated by these values, the inferior compressive strength of aramid fiber-reinforced composites limits their use in some structural applications. On the other hand, these composites possess high ductility and toughness, which makes them attractive for structures subjected to impact. Increasingly, hybrid composites, combining aramid and carbon (and/or glass) fibers, are being used to achieve good stiffness, strength, and toughness properties.

For PE fiber-reinforced composites the main drawback is the temperature-dependent degradation of properties. Typically, the PE fibers display unstable properties beyond 130°C, and have significant creep at lower temperatures. When combined with a polymer matrix, for example epoxy, the composite shows strain rate dependent behavior. Consequently, in cyclic stressing a hysteresis heat accumulation occurs, which reduces the stiffness and strength properties.

In PE fiber-reinforced composites the stiffness and strength properties depend on the temperature at which the composite is processed by impregnation. For instance, if processing temperature is 120–130°C, the tensile modulus has a value in a relatively high range of 30–35 GPa, which comes down to 20 GPa at 140°C processing temperature. The tensile strength shows less sensitivity to the processing temperature. The compressive behavior is affected by the processing temperature similarly as in tension.

See also: Biological Structures; Composites: Carbon Matrix; Composites: Overview; Liquid Crystals; Polymers, History of; Polymer Structures.

PACS: 81.05.Lg; 81.05.Qk; 82.35.Ab; 82.35.Lr

Further Reading

Hu X-C and Yang HH (2000) Polyamide and polyester fibers. In: Chow TW (ed.) *Comprehensive Composite Materials*, vol. 1, ch. 11. Amsterdam: Elsevier.

Peijs T, Jacobs MJN, and Lemstra PJ (2000) High performance polyethylene fibers. In: Chow TW (ed.) *Comprehensive Composite Materials*, vol. 1, ch. 9. Amsterdam: Elsevier.

Prevoresek DC and Mills PJ (1989) Polymer–polymer composites. In: Kelly A (ed.) *Concise Encyclopedia of Composite Materials*, pp. 235–238, Pergamon.

Schulte K and von Lacroix F (2000) High-density polyethylene fiber/polyethylene matrix composites. In: Talreja R and Månson J-E (eds.) *Comprehensive Composite Materials*, vol. 1, ch. 7. Amsterdam: Elsevier.

Computer Simulation Techniques in Condensed Matter Physics

K Binder, Johannes Gutenberg Universität Mainz, Mainz, Germany

Introduction

Computer simulation techniques in condensed matter physics aim to compute the structure and the physical properties (including also dynamic response functions and transport) from atomistic input. The theoretical basis of these approaches are quantum mechanics and statistical thermodynamics. There exist many variants of these methods, and it depends on the nature of the problem to find out which variant is most appropriate. Such simulations can complement both analytical theory and experiment, and due to the availability of cheap computing power

simulations have become an extremely widespread tool of research.

The conceptually simplest approach is the classical molecular dynamics (MD) method: one simply solves Newton's equations of motion numerically for the many-particle system (assuming that the N atoms or molecules in a volume V interact with effective potentials that are either phenomenologically assumed or fitted to electronic structure calculations). The basis of this method thus is classical mechanics, and one creates a deterministic trajectory in the phase space of the system. The idea is to take time averages of the observables of interest along this trajectory. The ergodicity hypothesis of statistical mechanics asserts that these time averages are equivalent to ensemble averages of the appropriate microcanonical (NVE) ensemble. Of course, Newton's equations of motion conserve the total energy E, and hence the conjugate intensive thermodynamic variables (temperature T, pressure p) can only be indirectly inferred and they exhibit statistical fluctuations. Since N is finite and sometimes even fairly small, such fluctuations cannot be neglected and need to be considered with care.

Sometimes one desires to directly realize other ensembles of statistical mechanics with a simulation, for example, the NVT or NpT ensemble. This is possible by introducing a coupling to suitable "thermostats" or "barostats."

An alternative way to carry out an MD simulation with $T = \text{const}$ introduces a weak friction force, together with random forces whose strengths are controlled by the fluctuation dissipation theorem. This method is closely related to the "Brownian Dynamics" (BD) method where one simulates Langevin equations (the inertial terms in the equations of motion being omitted). Clearly, this is a coarse-grained description of a physical system as only a subset of the degrees of freedom is explicitly considered, the remaining ones form the "heat bath." While this is not always a good description of the physical dynamics, such methods can be advantageous for the computation of static properties from averages along the stochastic trajectory in phase space.

A similar description applies to the importance sampling Monte Carlo (MC) method, where one creates a random walk-like trajectory in configuration space, controlled by transition probabilities that ensure the approach to thermal equilibrium through the detailed balance condition. Many of the practical limitations of simulations, such as "statistical errors," as well as systematic errors due to the finite size of the simulated system or the finite "length" of the simulated trajectory (or observation time, respectively) are common to all these simulation methods.

Of course, it is quantum mechanics and not classical mechanics that describes the basic physics of condensed matter. But attempting a numerical solution of the Schrödinger equation for a system of many nuclei and electrons is premature and still not feasible even with the fastest computers. Thus, approximations are needed: one very popular approach is the "*ab initio* MD" or "Car–Parrinello method" (CP), where some electronic degrees of freedom are included in MD via the density-functional theory (DFT). The huge advantage of this technique is that one no longer relies on effective interatomic potentials, which often are only phenomenologically chosen *ad hoc* assumptions, lacking any firm foundation in quantum chemistry. However, the disadvantage of this technique is that it is several orders of magnitude slower than classical MD, and hence only very short timescales and very small systems are accessible. Furthermore, the method cannot handle van der Waals-like forces well (e.g., in noble gases, one is still better off with the simple Lennard–Jones potential, amended by three-body forces).

In the standard CP method, the ionic motion is still treated classically. Alternatively, one can still use effective potentials between ions and/or neutral atoms as in classical MD or MC, but rely on quantum statistical mechanics for the ionic degrees of freedom. This is achieved by the path integral Monte Carlo (PIMC) or path integral molecular dynamics (PIMD). Such techniques are indeed crucial for a study of solids at low temperatures, to ensure that their thermal properties are compatible with the third law of thermodynamics. For most fluids, however, classical MD is sufficient (of course, quantum liquids such as helium are an exception).

Other variants of quantum Monte Carlo (QMC) have been developed in order to deal with electronic degrees of freedom on rigid lattices, desiring to describe magnetic systems (e.g., via the Heisenberg model where spin operators at neighboring lattice sites are coupled by exchange interactions), models for high T_c superconductors (e.g., using the Hubbard Hamiltonian, describing a hopping of electrons from site to site, and an on-site Coulomb repulsion), etc. These techniques ("diffusion Monte Carlo," "Green's function Monte Carlo," "variational Monte Carlo," "worldline Monte Carlo," etc.) are fairly specialized and complicated, and hence are not treated further here.

What information does one then wish to extract from the simulations? When systems in thermal equilibrium are considered, the first task is to calculate static properties. For example, in a fluid a basic property is the static structure factor $S(k), S(k) = \langle |\delta\rho(\boldsymbol{k})|^2 \rangle_T, \delta\rho(\boldsymbol{k})$ being a spatial Fourier transform of density fluctuations, $\boldsymbol{k}$ being the wave vector of a

scattering experiment. In addition, one may want to calculate time-dependent correlation functions that describe the decay of small thermal fluctuations in equilibrium systems with time. A quantity of this type is the intermediate scattering function $S(k,t) = \langle \delta\rho(-\boldsymbol{k},0)\delta\rho(\boldsymbol{k},t)\rangle$ or its Fourier transform with respect to time, $S(k,\omega)$. This "dynamic structure factor" is accessible to inelastic scattering of neutrons, X-rays, or light (with energy transfer $\hbar\omega$).

It is also possible to consider systems out of thermal equilibrium; for example, in solids under strong mechanical deformation, a crack may form and propagate. An important application of MD for fluids also is the study of systems exhibiting a steady state flow under shear deformation. The purpose of such nonequilibrium molecular dynamics (NEMD) work can be the estimation of transport coefficients (e.g., the shear viscosity) if the deformation is weak enough so that the system is still in the regime of linear response. However, the study of nonlinear phenomena also ("shear thinning," a decrease of the effective viscosity with increasing shear rate, or "shear melting" of colloidal crystals, etc.) may be of interest. In addition, one can study nonsteady state transient behavior, as it occurs, for example, after a sudden change from one thermodynamic state to another, and one wishes to study the approach of the system to its new thermal equilibrium. Classical examples of this problem are nucleation of fluid droplets in a supersaturated gas, or the kinetics of phase separation in a binary mixture ("spinodal decomposition") after a temperature quench. However, such processes are often too slow, and cannot be studied by NEMD, and one has to resort to a simulation of a coarse-grained model by nonequilibrium Monte Carlo (NEMC). This approach is suitable to deal also with irreversible growth processes far from equilibrium, such as diffusion-limited aggregation (DLA). This is an example of a mechanism of structure formation on mesoscopic scales of length and time that was discovered by simulations, and still lacks a good understanding in terms of the analytical theory.

In the following, the salient features of a few important simulation methods (classical MD, MC, PIMC) are briefly described. Also, limitations of simulations (e.g., due to the small size of the simulated system) are mentioned, and the extent to which these limitations can be overcome are discussed.

MD Algorithms

Consider a system of N particles (atoms) with Cartesian coordinates, $\boldsymbol{X} = \{\boldsymbol{r}_i\}, i = 1, \ldots, N$, in a d-dimensional space. The dynamics then is described by Newton's equations of motion, $m_i\ddot{\boldsymbol{r}}_i = \boldsymbol{f}_i = -\partial U_{\text{pot}}/\partial \boldsymbol{r}_i$, m_i being the mass of the ith particle, $\boldsymbol{f}_i$ the force acting on it. This force is assumed to be entirely due to interactions with other particles: $U_{\text{pot}}(\boldsymbol{X}) = U_{\text{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \sum u(|\boldsymbol{r}_i - \boldsymbol{r}_j|)$, where the sum extends once over each pair i, j of particles, and it has been assumed for simplicity that U_{pot} is pairwise additive. Then $\boldsymbol{f}_i = -\sum_j \partial u(|\boldsymbol{r}_i - \boldsymbol{r}_j|)/\partial \boldsymbol{r}_i$. Note that $E = E_{\text{kin}} + U_{\text{pot}} = \sum_i m_i \dot{\boldsymbol{r}}_i^2/2 + U_{\text{pot}}$ is a constant of motion ($\dot{E} = \mathrm{d}E/\mathrm{d}t = 0$).

MD now means that Newton's equations of motion are integrated numerically, by a computationally efficient scheme, such as Verlet's algorithm

$$\boldsymbol{r}_i(t+\delta t) = 2\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta t) + \frac{1}{m_i}(\delta t)^2 \boldsymbol{f}_i(t) + 0\{(\delta t)^4\} \quad [1]$$

where δt is the MD time step, and the velocity $\boldsymbol{v}_i(t)$ is updated similarly,

$$\boldsymbol{v}_i(t) = [\boldsymbol{r}_i(t+\delta t) - \boldsymbol{r}_i(t-\delta t)]/[2(\delta t)] + 0\{(\delta t)^3\} \quad [2]$$

This algorithm is manifestly time-reversible and in spite of inevitable errors due to the discreteness of the time step δt, it keeps an important symmetry property of Newton's equations. To understand how large δt can be chosen, consider argon as an example, where atoms interact with a Lennard–Jones potential, $u_{\text{LJ}} = 4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, with $\sigma \approx 3.4$ Å, $\varepsilon/k_{\text{B}} \approx 120$ K ($k_{\text{B}} =$ Boltzmann's constant), and $m \approx 6.6 \times 10^{-23}$ g. Rescaling coordinates ($\boldsymbol{r}^* \equiv \boldsymbol{r}/\sigma$) yields ($\boldsymbol{r}_{ij}^* \equiv \boldsymbol{r}_i^* - \boldsymbol{r}_j^*$)

$$\boldsymbol{r}^*(t+\delta t) = 2\boldsymbol{r}^*(t) - \boldsymbol{r}^*(t-\delta t) - (\delta t/\tau_0)^2 \frac{\boldsymbol{r}_{ij}^*}{|\boldsymbol{r}_{ij}^*|} \times \sum_{j(\neq i)} [(r_{ij}^*)^{-13} - (r_{ij}^*)^{-7}/2] \quad [3]$$

where the natural time unit τ_0 of MD was defined as $\tau_0 = (m\sigma^2/48\varepsilon)^{1/2}$ which is roughly $\tau_0 \approx 3.1 \times 10^{-13}$ s for argon. In order to keep numerical errors small, one needs to choose $\delta t = 0.03\,\tau_0$ (or smaller), that is, $\delta t \approx 10^{-14}$ s. So even with a million time steps, one only reaches a real time of ~ 10 ns.

As mentioned above, from a statistical mechanics point of view this algorithm realizes the microcanonical NVE ensemble. Temperature T is then inferred from the kinetic energy,

$$T = \langle \hat{T} \rangle_{\text{NVE}}, \quad \hat{T} = 2E_{\text{kin}}/(3k_{\text{B}}N) = \sum_{i=1}^{N} m_i \boldsymbol{v}_i^2/(3k_{\text{B}}N) \quad [4]$$

Note that N is finite (typically $10^2 \leqslant N \leqslant 10^6$) and hence temperature fluctuations are non-negligible.

These fluctuations contain information on the specific heat C_V, since $[\langle \hat{T}^2 \rangle - \langle \hat{T} \rangle^2]/\langle \hat{T} \rangle^2 = 2(1-3k_B/2C_V)/(3N)$. Since the "length" of an MD run (i.e., its number of time steps) is finite, these fluctuations of T cause a statistical error in the estimation of temperature.

The pressure p can be estimated using the virial theorem,

$$p = \langle \hat{P} \rangle_{\mathrm{NVE}}, \quad \hat{P} = \left(2E_{\mathrm{kin}} + \sum_i \boldsymbol{r}_i \cdot \boldsymbol{f}_i \right) / 3V \quad [5]$$

and again statistical fluctuations need to be considered.

It is often desirable to realize the canonical (NVT) rather than the microcanonical (NVE) ensemble. In an MD framework, this can be done by extending the Lagrangian of the system by a variable, representing the thermostat which has a fictitious "mass" Q. This yields the Nose–Hoover algorithm, where Newton's equations of motion are extended by a "friction"-like term,

$$\ddot{\boldsymbol{r}}_i = \boldsymbol{f}_i/m_i - \zeta(t)\dot{\boldsymbol{r}}_i, \quad \dot{\zeta} = \left[\sum_i m_i \boldsymbol{v}_i^2 - 3Nk_B T \right] / Q \quad [6]$$

Thus the "friction coefficient" $\zeta(t)$ fluctuates around zero, responding to the imbalance between the instantaneous kinetic energy and the intended canonical average. The total energy no longer is conserved, and energy fluctuations are linked to the specific heat as ($\hat{H}$ is the Hamiltonian)

$$NC_V/k_B = [\langle \hat{H}^2 \rangle_{\mathrm{NVT}} - \langle \hat{H} \rangle_{\mathrm{NVT}}^2]/(k_B T)^2 \quad [7]$$

It is important to note, however, that dynamical correlation functions between observables such as $\langle A(0)A(t) \rangle$, where A is the observable of interest, are not precisely identical to the microcanonical ones. This is even more true for MD runs that realize the isothermal–isobaric (NpT) ensemble, where the pressure is given and the volume V rather fluctuates, by coupling to the so-called "Andersen barostat." Therefore, one often uses MD in NVT or NpT ensembles for equilibration (and the computations of static averages) only, generating a set of well-equilibrated system configurations, which are then used as initial states for runs where thermostats and/or barostats are switched off, realizing the runs in the NVE ensemble, where $\langle A(0)A(t) \rangle$ is computed as a time average, $\int_0^{t_{\mathrm{obs}}-t} \mathrm{d}t_0 [A(t_0)A(t_0+t)]/(t_{\mathrm{obs}}-t), t_{\mathrm{obs}}$ being the total "observation time" over which the run is extended.

In order to judge whether t_{obs} is large enough, it is important to know what the slow observables in the system are. A typical reason for slow relaxation is the presence of large "objects" of correlated sets of particles in the systems, which show slow motions. An example is "critical slowing down" near second-order phase transitions, where large "clusters" of the variable representing the local order parameter occur. For example, near the liquid–gas critical point, large density fluctuations occur which are correlated over a correlation length ξ which diverges at criticality in the thermodynamic limit ($N \to \infty$), and then also the corresponding relaxation time τ diverges. Another example are polymer chains in solution or melt – a flexible polymer consisting of N monomers relaxes in the melt on a timescale $\tau \approx N^2$ in the Rouse limit (unentangled chains) or even $\tau \propto N^3$ in the reptation limit of entangled chains (i.e., a snakelike motion of the chains along their own contour). Thus, MD simulations of critical phenomena as well as of polymeric systems or other soft matter systems with objects containing many subunits (membranes, microemulsions, etc.) are intrinsically difficult.

However, it should also be noted that another reason for slow relaxation are conservation laws. In the NVE and NVT ensembles, the average density $\langle \rho \rangle = N/V$ is strictly conserved. Assuming for local fluctuations in density $\delta\rho(\boldsymbol{r},t) - \langle \rho \rangle$ that Fick's law holds, a gradient $\nabla(\delta\rho(\boldsymbol{r},t))$ causes a density current $\boldsymbol{j}(\boldsymbol{r},t)$, D being a diffusion coefficient, $\boldsymbol{j}(\boldsymbol{r},t) = -D\nabla(\delta\rho(\boldsymbol{r},t))$. The conservation law implies the continuity equation $\partial\rho(\boldsymbol{r},t)/\partial t + \nabla \cdot \boldsymbol{j}(\boldsymbol{r},t) = 0$, and hence a diffusion equation for $\delta\rho(\boldsymbol{r},t)$ results, $\partial(\delta\rho(\boldsymbol{r},t))/\partial t = D\nabla^2(\delta\boldsymbol{\rho}(\boldsymbol{r},t))$. The Fourier components $\delta\rho_{\boldsymbol{k}}(t)$ of $\delta\rho(\boldsymbol{r},t)$ then satisfy the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\rho_{\boldsymbol{k}}(t) = -Dk^2\delta\rho_{\boldsymbol{k}},$$
$$\delta\rho_{\boldsymbol{k}}(t) = \delta\rho_{\boldsymbol{k}}(0)\ \exp(-Dk^2 t) \quad [8]$$

Therefore, the dynamic correlation function of density fluctuations at long wavelengths decays very slowly,

$$S_\rho(\boldsymbol{k},t) \equiv \langle \delta\rho(-\boldsymbol{k},0)\delta\rho(\boldsymbol{k},0) \rangle$$
$$= S(k)\ \exp(-t/\tau_{\boldsymbol{k}}), \quad \tau_k = (Dk^2)^{-1} \quad [9]$$

Equation [9] demonstrates the "hydrodynamic slowing down" $\tau_k \to \infty$ for $k \to 0$. Of course, due to finite size effects, this divergence of the relaxation time is cut off; in a (hyper)cubic system of linear dimension L with periodic boundary conditions, the smallest wave number is $k_{\min} = (2\pi)/L$, and hence the largest relaxation time is $\tau_{\max} = L^2/(4\pi^2 D)$. Clearly, this phenomenon makes the equilibration of very large systems difficult.

A brief mention is made of the way transport coefficients are estimated from simulations. The simplest case are self-diffusion coefficients, where one can apply the Einstein relation, following the mean square

displacements of "tagged" particles, $\langle [\boldsymbol{r}(t)-\boldsymbol{r}(0)]^2 \rangle = 2\mathrm{d}Dt$, $t \to \infty$. Transport coefficients relating to collective properties can be extracted from Green–Kubo relations, for example, the shear viscosity is related to time correlations of the off-diagonal components of the pressure tensor σ_{xy},

$$\eta = \int_0^\infty \mathrm{d}t \langle \sigma_{xy}(0)\sigma_{xy}(t) \rangle / (Vk_\mathrm{B}T) \qquad [10]$$

where

$$\sigma_{xy} = \sum_{i=1}^{N} \left\{ m_i v_i^x v_i^y + \frac{1}{2} \sum_{j(\neq i)} x_{ij} f_i^y(r_{ij}) \right\} \qquad [11]$$

Similar relations can be written down for thermal conductivity, electrical conductivity, interdiffusion coefficient in mixtures, etc. Alternatively, transport coefficients can be estimated from NEMD, creating via suitable boundary conditions, a steady state of the appropriate quantity (mass density of a particle species, heat, momentum density) through the system.

Importance Sampling Monte Carlo Methods

Most MC calculations aim at the estimation of thermodynamic averages by importance sampling methods. Considering a variable $A(\boldsymbol{X})$ where $(\boldsymbol{X})$ denotes a point in phase space, such averages are defined in classical statistical mechanics by

$$\langle A \rangle = \int \mathrm{d}\boldsymbol{X} P_\mathrm{eq}(\boldsymbol{X}) A(\boldsymbol{X}),$$

$$P_\mathrm{eq}(\boldsymbol{X}) = (1/Z) \exp[-\hat{H}(\boldsymbol{X})/k_\mathrm{B}T] \qquad [12]$$

where $\hat{H}(\boldsymbol{X})$ is the Hamiltonian of the system, and Z is the partition function

$$Z = \int \mathrm{d}\boldsymbol{X} \exp[-\hat{H}(\boldsymbol{X})/k_\mathrm{B}T] \qquad [13]$$

Note that when one deals with the simulation of a fluid, momenta of the particles need not be included in $\boldsymbol{X}$ since they cancel out from all averages. In addition, $\boldsymbol{X}$ may denote any subset of variables only, for the problem of interest. For example, dealing with the phase transition of an anisotropic magnet, one may approximate the crystal lattice as rigid, and then each lattice site i carries as single degree of freedom an Ising spin, $S_i = \pm 1$. The Ising model Hamiltonian then is

$$\hat{H}_\mathrm{Ising} = -J \sum_{\langle i \neq j \rangle} S_i S_j - h \sum_i S_i \qquad [14]$$

assuming a nearest neighbor pairwise exchange energy J and a coupling to a magnetic field h (choosing units such that the magnetic moment per spin is unity). Here $\boldsymbol{X}$ stands for $S_1, S_2, \ldots, S_N$, that is, the phase space is discrete, and $\int \mathrm{d}\boldsymbol{X}$ stands symbolically for summation over all 2^N microstates of the system.

Now the basic idea of MC sampling is to consider the task of computing $\langle A \rangle$ as a problem of numerical integration over the phase space, where one would approximate $\langle A \rangle$ by an average $\bar{A}$ over M points only,

$$\bar{A} = \frac{1}{M} \sum_{v=1}^{M} P_\mathrm{eq}(\boldsymbol{X}_v) A(\boldsymbol{X}_v) \qquad [15]$$

The difficulty, of course, is that the integration space is very high dimensional, and it is a problem to choose the sample of M points $\boldsymbol{X}_v$ appropriately. It is easy to see that a regular grid of points $\boldsymbol{X}_v$ does not work, but a uniform random sampling of phase space ("simple sampling") does not work either in most cases. A notable exception, of course, is the sampling of nonthermal random distributions, for example, the random occupation of lattice sites in a mixed crystal A_xB_{1-x} by A and B atoms can be simulated straightforwardly. Choosing random numbers η uniformly distributed between zero and unity, an A-atom is assigned to the considered lattice site if $\eta < x$, while otherwise the site is taken to be occupied by a B-atom. This straightforward use of MC sampling works only if the occupation probabilities of neighboring lattice sites are independent of each other, however. Other physically relevant problems where simple sampling works are the generation of random walks and (short) self-avoiding walks, the bond percolation problem (nearest neighbor links connecting lattice sites are chosen at random conducting or isolating, to understand the electrical conductivity of disordered solids), etc. However, simple sampling is not useful for the estimation of thermal averages when N is large. The reason is that it is only a very small region of phase space where the important contributions come from, and via simple sampling only very few points $\boldsymbol{X}_v$ would be drawn from this important region. This is seen when one considers the reduced distribution function $P_N(m)$ of the density m of some extensive variable, for example, the magnetization of the Ising magnet, $m = (1/N)\sum S_i$, where the sum runs over all lattice sites, or the internal energy, etc., $P_N(m) = \int \mathrm{d}\boldsymbol{X} \delta(m - (1/N)\sum S_i) P_\mathrm{eq}(\boldsymbol{X})$. Such distributions exhibit sharp peaks at the (*a priori* unknown) average value $\langle m \rangle$, of width proportional to $1/\sqrt{N}$. It is clear, however, that from simple sampling for large N, almost no states would be generated in the region where $P_N(m)$ has its peak; most of the computational effort would be wasted for exploring a completely uninteresting part of the phase space.

Thus a method is needed that leads one automatically to the important region of the phase space. Such a method is provided by the Metropolis importance sampling algorithm, which yields states $\boldsymbol{X}_\nu$, with a probability proportional to the Boltzmann factor. This is achieved by generating a Markov chain of states recursively one from the other, $\boldsymbol{X}_\nu \rightarrow \boldsymbol{X}_{\nu+1} \rightarrow \boldsymbol{X}_{\nu+2} \rightarrow \ldots$, using a properly chosen transition probability. For a fluid, the move from $\boldsymbol{X}$ to $\boldsymbol{X}'$ may mean a random displacement of a single randomly chosen particle; for an Ising model, it may mean a single spin is chosen at random and considered for a spin flip, or for exchange with one of its neighbors; but also flips of large clusters of spins in one move together may be carried out. Also moves that have no counterpart in reality are conceivable (e.g., in a dense polymer melt, one may select a pair of long chains that are close by each other, cut them in the middle, and reconnect the parts differently than previously). Even though such moves need not represent any physically plausible dynamics, they may provide an efficient sampling of the phase space. The great flexibility in the choice of such moves, adapted to the problem one wishes to study, is one of the particular strengths of the MC methods.

The states $\boldsymbol{X}_\nu$ are distributed according to the Boltzmann factor $P_{\text{eq}}(\boldsymbol{X})$ if the transition probability $W(\boldsymbol{X} - \boldsymbol{X}')$ satisfies the condition of detailed balance

$$P_{\text{eq}}(\boldsymbol{X})\, W(\boldsymbol{X} \rightarrow \boldsymbol{X}') = P_{\text{eq}}(\boldsymbol{X}')\, W(\boldsymbol{X}' \rightarrow \boldsymbol{X}) \qquad [16]$$

For instance, a choice that satisfies eqn [16] is simply

$$W(\boldsymbol{X} \rightarrow \boldsymbol{X}') = \begin{cases} \tau_0^{-1}, & \text{if } \delta\hat{H} \equiv \hat{H}(\boldsymbol{X}') - \hat{H}(\boldsymbol{X}) < 0 \\ \tau_0^{-1} \exp(-\delta\hat{H}/k_{\text{B}}T), & \text{if } \delta\hat{H} \geqslant 0 \end{cases}$$

Here, a constant τ_0 has been arbitrarily introduced to set a timescale, so that W gets the meaning of a transition probability per unit time (but one may choose $\tau_0 = 1$, of course, and normally one considers one Monte Carlo step (MCS) per degree of freedom as the natural unit of the MC "time"). This notion is useful, since MC sampling can be interpreted as "time averaging" along stochastic trajectories in phase space, described by a master equation for the probability $P(\boldsymbol{X}, t)$ that a state $\boldsymbol{X}$ occurs at time t,

$$\frac{\mathrm{d}}{\mathrm{d}t} P(\boldsymbol{X}, t) = -\sum_{\boldsymbol{X}'} W(\boldsymbol{X} \rightarrow \boldsymbol{X}') P(\boldsymbol{X}, t) + \sum_{\boldsymbol{X}'} W(\boldsymbol{X}' \rightarrow \boldsymbol{X}) P(\boldsymbol{X}', t) \qquad [17]$$

This rate equation describes the balance between the loss of probability by all processes $\boldsymbol{X} \rightarrow \boldsymbol{X}'$ that lead away from the considered state and the gain of probability due to the inverse process. Obviously, $\mathrm{d}P/\mathrm{d}t = 0$ if $P = P_{\text{eq}}(\boldsymbol{X})$, due to the detailed balance principle, eqn [16].

Of course, the dynamical properties of a system described by such a stochastic trajectory through phase space generated via MC will differ in general from dynamic properties deduced from the deterministic MD trajectories. In fact, the MC dynamics is of physical significance only when one deals with systems for which the considered degrees of freedom are a (slow!) subset weakly coupled to the remaining (fast!) degrees of freedom, that behave like a heat bath. This is the case for the Ising magnet and for diffusion in solid alloys $A_x B_{1-x}$, for instance. In both cases the phonons of the crystal act like a heat bath, causing random spin flips or random jumps of A or B atoms to vacant sites, respectively. Often the MC dynamics is considered as fairly realistic for random growth or aggregation phenomena on mesoscopic scales, such as diffusion-limited aggregation (DLA).

The dynamic interpretation of MC sampling in terms of a stochastic trajectory in phase space is important for the understanding of "statistical error," since the subsequently generated states $\{\mathbf{X}_\nu\}$ are highly correlated with each other. In fact, one can show that the expected mean square statistical error becomes

$$\langle (\delta A)^2 \rangle = \frac{1}{n} [\langle A^2 \rangle - \langle A \rangle^2](1 + 2\tau_{AA}/\delta t) \qquad [18]$$

where

$$\tau_{AA} = \int_0^\infty \Phi_{AA}(t)\mathrm{d}t, \quad \Phi_{AA}(t) \equiv [\langle A(0)A(t) \rangle - \langle A \rangle^2]/[\langle A^2 \rangle - \langle A \rangle^2] \qquad [19]$$

if n observations $\{A_\mu\}$ of the observable A separated by time increment δt are taken. If these n observations were statistically independent, the enhancement by the "dynamic factor" $(1 + 2\tau_{AA}/\delta t)$ would be absent. As noted in the discussion of MD methods, relaxation times τ_{AA} get large near critical points, or when one considers long wavelength Fourier components of a conserved quantity ("critical slowing down," "hydrodynamic slowing down," also hamper MC!). Also the questions of whether or not an algorithm is ergodic, and how fast equilibrium is reached or whether hysteresis may occur, etc., are important for both MD and MC.

A further limitation of simulations that may need consideration are finite size effects. Considering Fourier components such as $\delta\rho_{\boldsymbol{k}}$ (eqns [8] and [9]), one must note that in a finite system with periodic boundary conditions $\boldsymbol{k}$-space is quantized. For example, if one has a cubic simulation cell $L \times L \times L$, the allowed

wave vectors are $\boldsymbol{k}_n = (2\pi/L)(n_x, n_y, n_z)$ where n_α are integers. With respect to static phenomena, the gap in the spectrum of long wavelength modes affects physical properties for which these modes are important (e.g., long wavelength phonons cause the T^3 law for specific heats, if the crystal is treated by quantum statistical mechanics; long wavelength magnons in a Heisenberg ferromagnet cause a divergence of the longitudinal and transverse susceptibilities; capillary waves cause a divergence of interfacial widths, etc.). Thus for many physical properties, it is advisable to repeat the simulation for a range of values of L, and carry out a suitable extrapolation of the results to the thermodynamic limit, $L \to \infty$.

The problem of finite size effects is particularly cumbersome when one deals with phase transitions. However, the study of phase transitions is a problem for which simulations are often necessary, and, in fact, useful! One approach to overcome the limitation of finite size for such problems relies on the concept of finite size scaling. Consider, as a simple example, the transition of the Ising model {eqn [14]} from the paramagnetic phase at $T > T_c$ to a ferromagnetic state at $T < T_c$. Near the critical temperature T_c, spontaneous magnetization m_{sp} and susceptibility χ are singular in the thermodynamic limit, $m_{sp} \propto (1 - T/T_c)^{\beta}$, $\chi \propto |1 - T/T_c|^{-\gamma}$, where β, γ are "critical exponents." Varying H for $T < T_c$, however, m at $H = 0$ has a jump singularity from $-m_{sp}$ to $+m_{sp}$, χ then has a delta-function singularity: this is a special case of a first-order transition.

In a finite system, however, these singularities are always rounded. For example, χ for $H = 0$ as a function of T does not diverge, there is a peak of finite height only, which also is shifted away from T_c. Thus, it is a nontrivial problem to estimate both T_c and other critical properties (such as exponents) with sufficient accuracy.

For critical phenomena, finite size scaling rests on a comparison of lengths: the linear dimension L scales with the correlation length $\xi \propto |1 - T/T_c|^{-\nu}$ of the system. As a consequence, near T_c moments of the order parameter distribution depend on L and ξ basically via a scaling function of the ratio L/ξ, apart from a power-law prefactor $\langle |m|^k \rangle = L^{-k\beta/\nu} \tilde{M}_k (L/\xi)$, $L \to \infty, k$ integer, $\tilde{M}_k$ being a "scaling function"! Particularly useful are, thus, suitable ratios of moments where the power law prefactors cancel out, such as the reduced fourth-order cumulant

$$U_L(T) = 1 - \langle m^4 \rangle / [3 \langle m^2 \rangle^2] = \tilde{U}(L/\xi), \quad L \to \infty \qquad [20]$$

Since $\xi \to \infty$ as $T \to T_c$, all curves $U_L(T)$ must intersect at $T = T_c$ in a universal value $\tilde{U}(0)$; note that $U_L(T) \to 2/3$ for $T < T_c$ as $L \to \infty$, while $U_L(T) \to 0$ for $T > T_c$. If such a unique intersection point is found, it thus serves to find the location of T_c. If the intersections of different pairs $U_L(T)$, $U_{L'}(T)$ scatter, on the other hand, one can conclude that corrections to scaling are still pronounced, and the study of larger systems is required. At T_c, one also finds that $\langle |m| \rangle \propto L^{-\beta/\nu}, \chi \propto L^{\gamma/\nu}$, and thus one can estimate critical exponents.

Also, first-order transitions are rounded; while in the Ising model for $T < T_c$, two phases with $\pm m_{sp}$ occur precisely at $H = 0$, the distribution function of the magnetization simply is $P_\infty(m) = [\delta(m - m_{sp}) + \delta(m + m_{sp})]/2$, and for finite L, one observes two broadened peaks not only at $H = 0$ but over a range ΔH of fields, with $\Delta H \propto L^{-d}$, inversely proportional to the volume of the system. But the location of the first-order transition is found accurately from the "equal weight rule" for the two peaks.

The sampling of $P_L(m)$ is very useful since one can show that for $T < T_c$, one has $P_L(0)/P_L(m_{sp}) \propto \exp(-2L^{d-1}\sigma/k_B T)$, σ being the interfacial tension between the coexisting phases. Use of this formula, in fact, is one of the most powerful recipes to estimate interfacial free energies from simulations. Of course, $P_L(0)/P_L(m_{sp})$ for large L is extremely small, and thus this ratio cannot be sampled directly, but needs more advanced schemes that are not described here due to lack of space, such as "umbrella sampling" or "multicanonical MC."

Of course, the above comments discuss the study of phase transitions by simulations only in a nutshell, and the reader is directed to the "Further reading" section for further details.

Path-Integral Monte Carlo (PIMC)

Thermal averages for a quantum system are calculated, and thus eqns [12] and [13] are rewritten appropriately

$$\langle \hat{A} \rangle = Z^{-1} \operatorname{Tr} \exp(-\hat{H}/k_B T)\hat{A}, \qquad Z = \operatorname{Tr} \exp(-\hat{H}/k_B T) \qquad [21]$$

using a notation that emphasizes the operator character of the Hamiltonian $\hat{H}$ and observable $\hat{A}$. For simplicity, consider first the case of one particle in a potential $V(x)$ in one dimension, where $\hat{H} = -(\hbar^2/2m)\mathrm{d}^2/\mathrm{d}x^2 + V(x)$, m being the mass of the particle. In position representation $Z = \int \mathrm{d}x \langle x| \exp(-\hat{H}/k_B T)|x\rangle, |x\rangle$ being an eigenvector of the position operator. The problem now is that $\hat{H} = \hat{E}_{kin} + \hat{V}(x)$, where $\hat{E}_{kin}$ and $\hat{V}(x)$ do not commute. Applying the Suzuki–Trotter formula, P being a positive

integer

$$\exp[-(\hat{E}_{\text{kin}} + \hat{V})/k_{\text{B}}T] = \lim_{P\to\infty} [\exp(-\hat{E}_{\text{kin}}/k_{\text{B}}TP)\, \exp(-\hat{V}/k_{\text{B}}TP)]^P \quad [22]$$

one can reduce Z to ($x_{P+1} = x_1$)

$$Z = \lim_{P\to\infty} \left(\frac{k_{\text{B}}TmP}{2\pi\hbar^2}\right)^{P/2} \int \mathrm{d}x_1 \ldots \int \mathrm{d}x_P \times \exp\left(\left[-\frac{\kappa}{2}\sum_{S=1}^{P}(x_s - x_{s+1})^2 - P^{-1}\sum_{S=1}^{P} V(x_s)\right]/k_{\text{B}}TP\right) \quad [23]$$

where $\kappa \equiv (k_{\text{B}}T/\hbar)^2 mP$. Apart from the prefactor, eqn [23] is the configurational partition function of a ring polymer within the framework of classical statistical mechanics: P beads are coupled by harmonic springs with spring constant κ, and are exposed to a potential $V(x)/P$.

This approach is straightforwardly generalized to a system of N interacting quantum particles, if the quantum mechanical exchange is neglected: one ends up with a system of N classical cyclic "polymer chains." As a result of this isomorphism, methods for simulating classical systems (MC, MD) can be carried over to such quantum-mechanical problems, too. At high temperatures, κ gets very large, the cyclic chain contracts essentially to a point, while at low T the chains are spread out, representing zero-point motions. The lower the T is, the larger the P has to be: if σ is a characteristic distance over which the potential $V(x)$ changes, one must have $\hbar^2/m\sigma^2 \ll k_{\text{B}}TP$ so that two neighbors along the "polymer chain" are at a distance much smaller than σ. The typical size of the ring polymer, of course, is of the order of the thermal de Broglie wavelength λ_{T}: at very low temperatures (or for very light particles such as He atoms) λ_{T} becomes comparable to the interparticle spacing, and the precise formulation of the problem requires the correct incorporation of the statistics of the particles (Bose–Einstein versus Fermi–Dirac) into the formalism. For fermions, this leads to the famous difficulty known as "minus sign problem" – the distribution to be sampled loses its interpretation as a probability distribution, because it is no longer non-negative everywhere. Thus the study of systems such as liquid ^{3}He (or fermions on a lattice, a problem to which a similar formulation based on the Suzuki–Trotter decomposition is applicable) is still an active area of research. However, one may stress that the simplified formulation of PIMC (or PIMD, respectively) that ignores the statistics of the particles is useful (and indispensable!) to study the thermal properties of crystals (such as solid Ne or Ar, SiO_2, polyethylene etc.) at low temperatures (and obtain temperature dependences of lattice parameters, elastic constants etc., that are compatible with the third law of thermodynamics).

Concluding Remarks

In this article, the basic foundations of three important simulation methods (classical MD and MC, PIMC) have been briefly described, and some technical limitations as well as recipes of how they can be overcome have also been mentioned. However, there are many "tricks of the trade" that could not be described, due to lack of space: for example, one can switch on or off arbitrary perturbations or parts of the interparticle forces, and combine this with suitable thermodynamic integration methods to gain information on thermodynamic excess potentials, etc. Very important variations are also possible by a suitable choice of various boundary conditions (e.g., a combination of "tilted periodic" with "antiperiodic" boundary conditions allows the study of tilted interfaces in Ising systems). An important aspect also is that an analysis of conceptually important quantities (e.g., Voronoi polyhedra distributions in amorphous solids, ring statistics in molten silica, orientational order parameters in two-dimensional melting, etc.) is readily possible with simulations, while one can infer these quantities from experiment at best indirectly. Many such extensions of the techniques described here, as well as further simulation methods, and last but not the least, numerous application examples can be found in the "Further reading" section.

See also: Molecular Dynamics Calculations; Molecular Dynamics, Car–Parrinello.

PACS: 71.15. – D; 61.20.J; 02.50.N

Further Reading

Allen MP and Tildesley DJ (1987) *Computer Simulation of Liquids*. Oxford: Clarendon Press.

Binder K (ed.) (1979) *Monte Carlo Methods in Statistical Physics*. Berlin: Springer.

Binder K (ed.) (1992) *The Monte Carlo Method in Condensed Matter Physics*. Berlin: Springer.

Binder K and Ciccotti G (eds.) (1996) *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*. Bologna: Societa Italiana di Fisica.

Binder K and Heermann DW (2002) *Monte Carlo Simulation in Statistical Physics. An Introduction*, 4th edn. Berlin: Springer.

Car R and Parrinello M (1985) Unified approach for molecular dynamics and density functional theory. *Physical Review Letters* 55: 2471–2474.

Doll JD and Gubernatis JE (eds.) (1990) *Quantum Simulations*. Singapore: World Scientific.
Frenkel D and Smit B (2002) *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd edn. San Diego: Academic Press.
Grotendorst J, Marx D, and Muramatsu A (eds.) (2002) *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*. Jülich: NIC.
Landau DP and Binder K (2000) *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge: Cambridge University Press.
Rapaport DC (1995) *The Art of Molecular Dynamics*. Cambridge: Cambridge University Press.
Suzuki M (ed.) (1992) *Quantum Monte Carlo Methods in Condensed Matter Physics*. Singapore: World Scientific.

Conductivity, Electrical

J Bass, Michigan State University, East Lansing, MI, USA

Introduction

One important way to characterize a material is in terms of its electrical conductivity (or the inverse, its electrical resistivity). A metal conducts electricity well, and its resistivity usually increases with increasing temperature. Most elements (82 of the 104 naturally occurring ones) are metals. A semiconductor conducts electricity less well, and its resistivity usually decreases with increasing temperature. An insulator hardly conducts electricity at all; its resistivity is very large. This article deals with the electrical resistivity of pure metals and alloys. The focus is mainly on nonmagnetic metals – magnetic metals are covered elsewhere. References to both the theory of the electrical resistivity of metals, and to data on the resistivities of pure metals and alloys, are given in the Further Reading section.

Formally, the electrical conductivity tensor $\boldsymbol{\sigma}$ (and its inverse – the resistivity tensor $\boldsymbol{\rho}$) are defined by Ohms law:

$$\boldsymbol{J} = \boldsymbol{\sigma} \cdot \boldsymbol{E} \qquad [1]$$

and

$$\boldsymbol{E} = \boldsymbol{\rho} \cdot \boldsymbol{J} \qquad [2]$$

where $\boldsymbol{J}$ is the current density and $\boldsymbol{E}$ is the electric field (both vectors). In the general case, $\boldsymbol{\rho}$ is the inverse of the tensor $\boldsymbol{\sigma}$. However, if no magnetic field is present, and if the crystal structure of the metal is cubic – simple cubic (s.c.), face-centered cubic (f.c.c.), or body-centered cubic (b.c.c.) – then $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ reduce to scalars, and $\rho = 1/\sigma$. Since many of the most common metals have cubic structures, the focus is on the case where ρ and σ are scalars, in which case it suffices to consider ρ, because σ is simply its inverse. Where appropriate, noncubic metals are discussed.

Near room temperature ($\sim$295 K), convenient units for ρ are micro-ohm cm ($\mu\Omega$ cm), since the resistivities of all pure elemental metals at 295 K range from $\sim$1.6 $\mu\Omega$ cm for Ag to somewhat over 100 $\mu\Omega$ cm for Mn, Gd, and Tb.

Measurements

The resistivity ρ at temperature T, $\rho(T)$, of a given metal or alloy is derived from a "four-probe" measurement of the electrical resistance $R = V/I$ of a long thin wire of known cross-sectional area $A = \pi r^2$, or a long, narrow film of area $A = Wt$, as shown in **Figure 1**. Here r is the radius of the wire, W and t are the width and thickness of the film, and the voltage leads are made long and thin to minimize perturbing the current flow. These geometries minimize "end effects," thereby ensuring that the current density $J = I/A$ is uniform across A, and contact effects by eliminating current flow into the voltage contacts. In such a case, measuring the voltage V across a well-defined length L of the sample gives $\rho = RA/L = VA/IL$.

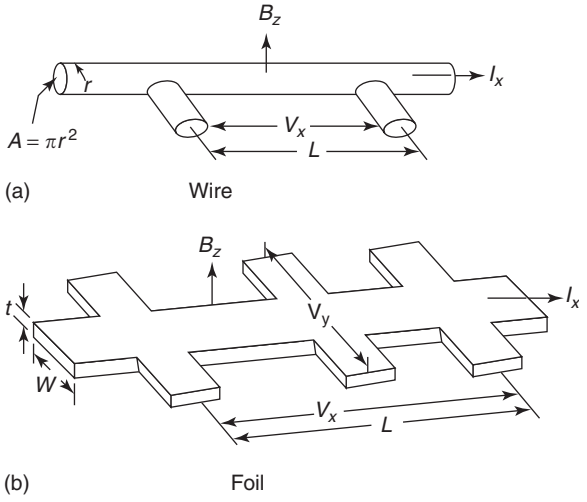


Figure 1 (a) Wire and (b) film geometries for measuring resistivities so that complications due to nonuniform current flow and contact resistances are minimal.

Strictly, as noted above, a unique value of $\rho(T)$ can be determined for either a single crystal or a polycrystalline sample of a cubic metal, but only for a single crystal of a noncubic metal that is oriented along a symmetry axis. Even in these cases, it will be seen below that $\rho(T)$ can depend upon the size of the sample if r or t are very small, and upon the content of impurities or other defects such as grain boundaries or dislocations, especially at low temperatures.

Theoretical Background

The first derivation of Ohm's law and σ was made by Drude in 1900, who used the classical kinetic theory to treat an assumed gas of free electrons in a metal. His result can be obtained by assuming that an electron of charge e and mass m is accelerated by an electric field E and is subject to a frictional retarding force proportional to its velocity v and inversely proportional to a relaxation time τ (a measure of the average time between scattering events). Then, at steady state, the electron is not accelerated, and

$$m\,\mathrm{d}v/\mathrm{d}t = 0 = eE - mv/\tau \;\rightarrow\; v = eE\tau/m \qquad [3]$$

Taking

$$J = nev \qquad [4]$$

(where e is the electron's charge and n is the number of electrons per unit volume), Ohm's law is obtained:

$$J = nev = (ne^2\tau/m)E \qquad [5]$$

with

$$\sigma = ne^2\tau/m \qquad [6]$$

The Bohr–Sommerfeld analysis of a quantum free-electron gas gives a similar expression, but with τ evaluated at the Fermi energy. If a mean free path λ between scattering events is defined as $\lambda = v_F\tau$, eqn [6] can be rewritten as

$$\sigma = ne^2\lambda/mv_F \qquad [7]$$

Inverting eqn [7] then gives

$$\rho\lambda = mv_F/ne^2 \qquad [8]$$

For a given metal or alloy at temperature T, eqn [7] shows that $\sigma \propto \lambda$, and eqn [8] shows that the product $\rho\lambda$ is a constant for a given metal, depending upon the particular metal only through n and v_F. Equation [8] can be used to estimate λ for a given value of ρ.

To generalize eqns [7] and [8] to real (as opposed to free-electron) metals, with real Fermi surfaces, one must turn to a more general analysis, such as the Boltzmann transport equation. Here, one defines an electron distribution function $f(\boldsymbol{k},\boldsymbol{r},t)$ such that $(4\pi^3)^{-1}\,f(\boldsymbol{k},\boldsymbol{r},t)\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\boldsymbol{r}$ gives the number of electrons of wave vector $\boldsymbol{k}$ (or crystal momentum $\hbar\boldsymbol{k}$, where $\hbar$ is Planck's constant divided by 2π) at point $\boldsymbol{r}$ within an element $\mathrm{d}\boldsymbol{r}$ in real space and $\mathrm{d}\boldsymbol{k}$ in wave vector space at time t. $f(\boldsymbol{k},\boldsymbol{r},t) = 1$ if all states with $\boldsymbol{k}$ and $\boldsymbol{r}$ are filled and 0 if all such states are empty. If the electron gas is at thermal equilibrium at a temperature T,

$$f_0 = 1/(\mathrm{e}^{(\varepsilon-\varepsilon_F)/k_BT} + 1) \qquad [9]$$

is the usual Fermi–Dirac distribution function. The Boltzmann transport equation for $f(\boldsymbol{k},\boldsymbol{r},t)$ is then

$$\partial f/\partial t + (\partial f/\partial \boldsymbol{r})(\partial \boldsymbol{r}/\partial t) + (\partial f/\partial \boldsymbol{k})(\partial \boldsymbol{k}/\partial t) = (\partial f/\partial t)_{\mathrm{scatt}} \qquad [10a]$$

where $(\partial f/\partial t)_{\mathrm{scatt}}$ is the rate of change of f due to scattering from defects, quantized lattice vibrations (phonons), or the sample surface. Since $\partial \boldsymbol{r}/\partial t = v$, and $\hbar(\partial \boldsymbol{k}/\partial t) = \boldsymbol{F}$, any external force $\boldsymbol{F}$, eqn [10a] may be rewritten as

$$\partial f/\partial t + v(\boldsymbol{k})(\partial f/\partial \boldsymbol{r}) + (1/\hbar)\boldsymbol{F}(\partial f/\partial \boldsymbol{k}) = (\partial f/\partial t)_{\mathrm{scatt}} \qquad [10b]$$

If both an electric field $\boldsymbol{E}$ and a magnetic field $\boldsymbol{B}$ are present, $\boldsymbol{F} = e\boldsymbol{E} + v\times\boldsymbol{B}$. In the absence of a magnetic field, eqn [10b] can be rewritten as

$$\partial f/\partial t + v(\boldsymbol{k})(\partial f/\partial \boldsymbol{r}) + (e/\hbar)\boldsymbol{E}(\partial f/\partial \boldsymbol{k}) = (\partial f/\partial t)_{\mathrm{scatt}} \qquad [10c]$$

If the focus is only upon the electrical conductivity σ of the bulk material, the temperature can be taken as constant across the sample, any changes in f with $\boldsymbol{r}$ can be neglected, and only the steady-state solution can be considered. Equation [10c] then reduces to

$$(e/\hbar)\boldsymbol{E}(\partial f/\partial \boldsymbol{k}) = (\partial f/\partial t)_{\mathrm{scatt}} \qquad [10d]$$

Because changes in f with $\boldsymbol{r}$ are neglected, this equation will not allow one to treat scattering at the sample surface (size effects). Such effects are discussed briefly below.

Finally, for the physics of present interest, the deviations of $f(\boldsymbol{k})$ from $f_0(\boldsymbol{k})$ will be small. These deviations can be considered by writing $f(\boldsymbol{k}) = f_0(\boldsymbol{k}) + g(\boldsymbol{k})$. In order to obtain Ohm's law, eqn [10d] must then be linearized (i.e., drop terms of higher order in $\boldsymbol{E}$), giving

$$(e/\hbar)\boldsymbol{E}(\partial f_0/\partial \boldsymbol{k}) = (\partial f/\partial t)_{\mathrm{scatt}} \qquad [11]$$

Equation [11] is the linearized Boltzmann transport equation in the absence of temperature gradients, magnetic fields, and size effects. In introductory and intermediate texts, it is usually simplified one more time by making the "relaxation time approximation,"

that is by taking

$$(\partial f/\partial t)_{\text{scatt}} = -g_k/\tau_k \qquad [12]$$

where τ_k is called the relaxation time. Since f_0 (see eqn [9]) is only a function of the energy ε_k, and $\boldsymbol{v} = (1/\hbar)\partial\varepsilon/\partial\boldsymbol{k}$, eqn [11] can be rewritten in the final form

$$e\boldsymbol{E}v_k(\partial f_0/\partial\varepsilon_k) = -g_k/\tau_k \qquad [13]$$

Generalizing eqn [4] gives $J = \int \mathrm{d}\boldsymbol{k}^3 e g_k v_k$ and some manipulation of eqn [13] leads to the relation

$$\sigma = (1/12\pi)(e^2/\hbar) \int \lambda_k \mathrm{d}S_k \qquad [14]$$

where λ_k is the mean free path of electrons of wave vector $\boldsymbol{k}$, and $\mathrm{d}\boldsymbol{S}_k$ is a differential area on the Fermi surface around wave vector $\boldsymbol{k}$. Equation [14] is the generalization of eqn [7].

For the present purposes, only one more result (the Bloch–Grüneisen equation for the temperature-dependent electrical resistivity of a free-electron gas (i.e., a spherical Fermi surface), in which the electrons are scattered only by phonons and with no participation of reciprocal lattice vectors (i.e., no Umklapp scattering)) is needed. The result, taken as their approximation for the resistivity of an ideally pure metal, ρ_{P}, is

$$\rho_{\text{P}} = (K/\vartheta)(T/\vartheta)(T/\vartheta)^4 \int_0^{\vartheta/T} 4z^4 \, \mathrm{d}z/(\mathrm{e}^z - 1) \qquad [15]$$

Since the interest here is only in the Bloch–Gruneisen prediction for the temperature dependence of ρ_{P}, several constants and parameters have been collected together into K, and only scaling parameter, ϑ, is left with the units of temperature.

At high temperatures, where $T \gg \vartheta$, the term $\mathrm{e}^z - 1$ in the denominator of the integral reduces to just z, leaving four powers of z inside the integral. The lower limit of the integral then gives zero, and the upper limit is proportional to $(\vartheta/T)^4$, which cancels the $(T/\vartheta)^4$ outside the integral, leaving just (T/ϑ). In this limit, ρ_{P} is proportional to T.

At low temperatures, where $T \ll \vartheta$, the upper limit can be approximated as infinity, turning the integral into just a number. The temperature dependence in this limit is then just that outside the integral, or T^5.

The Bloch–Gruneisen calculation thus predicts that ρ_{P} should be proportional to T at high T, and to T^5 at low T. A linear variation with T is often seen at high T (but not necessarily extrapolating to $\rho_{\text{P}} = 0$ at $T = 0$ as would be predicted by Bloch–Gruneisen). A strict T^5 variation is rarely, if ever, seen at low T. Deviations from the simple prediction occur for a variety of reasons, including: real Fermi surface effects such as Umklapp scattering; electron–electron scattering (which can give a T^2 variation); deviations from Matthiessen's rule; and the onset of superconductivity.

Experimental Temperature and Impurity-Concentration Dependence

From Bloch's theorem, the resistivity of a perfectly periodic metal should be zero. Resistivity results from scattering of conduction electrons by deviations from perfect periodicity, either due to the presence of defects at atomic sites (e.g., impurities, vacancies, dislocations, grain boundaries) or due to vibrations of the atoms of the lattice away from their equilibrium sites (quantized lattice vibrations are called phonons).

In a high-purity metal, scattering by phonons dominates at all but the lowest temperatures. As noted above, for a free-electron metal, the Bloch–Gruneisen model of $\rho(T)$ predicts a linear dependence upon T at high temperatures, and a more rapid decrease at lower temperatures, eventually reaching T^5 at cryogenic temperatures. **Figures 2–4** illustrate some of the behaviors of $\rho(T)$ at higher temperatures, using data for the simple f.c.c. metal Cu, which looks qualitatively like the expectation from Bloch–Gruneisen,

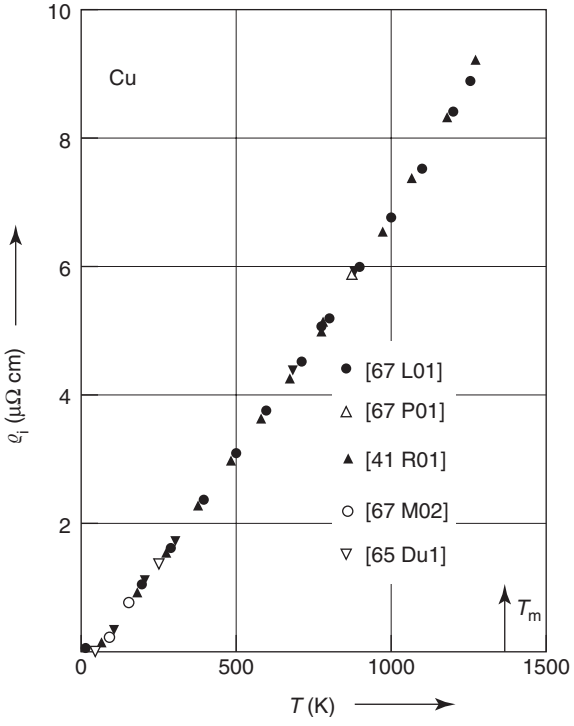


Figure 2 Resistivity as a function of temperature for Cu, a metal where the form of the resistivity is similar to that predicted by the Bloch–Gruneisen model.

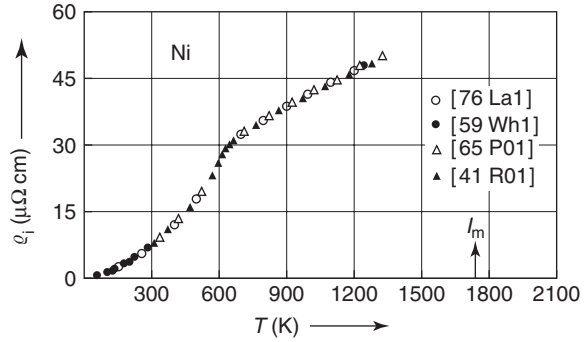


Figure 3 Resistivity as a function of temperature for Ni, a metal that becomes ferromagnetic below its Curie temperature of ~650 K.

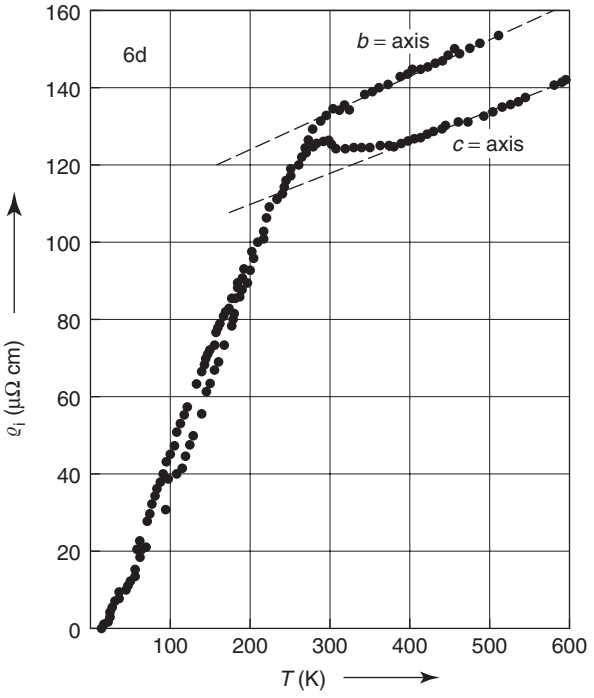


Figure 4 Resistivity as a function of temperature for a single-crystal sample of Gd, which has a noncubic crystal structure.

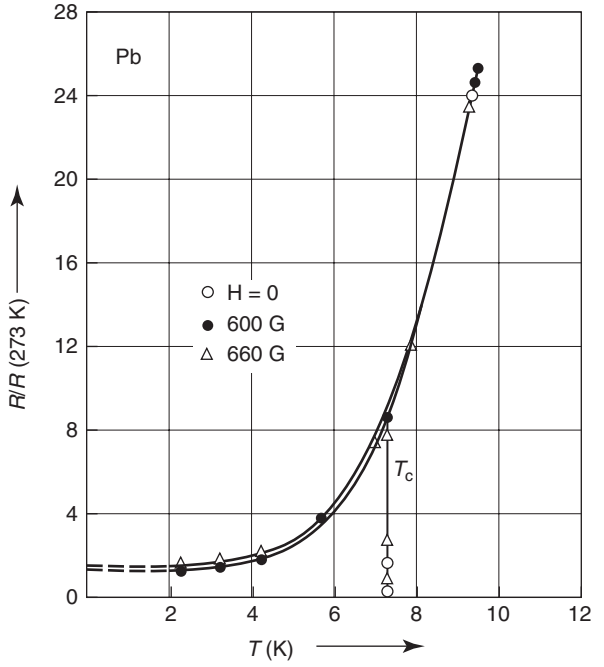


Figure 5 Low-temperature resistivity of Pb, which becomes superconducting (the resistivity drops to zero) in zero magnetic field at ~7.2 K. In a magnetic field of 600 Oe, Pb remains normal to the lowest temperature measured in this figure, and the resistivity at the lowest temperatures becomes constant, at a value determined by residual impurities.

the f.c.c. metal Ni, which undergoes a ferromagnetic phase transition at 630 K, and the hexagonal close-packed (h.c.p.) metal Gd, which undergoes magnetic phase transitions at 240 and 293 K. In the simplest case, Cu, at temperatures above ~100 K, $\rho(T)$ is approximately linear in T. At cryogenic temperatures, $\rho(T)$ decreases more rapidly than linearly with T. If the metal was perfectly pure, $\rho(T)$ would approach 0 as $T \rightarrow 0$ K, either smoothly if the metal remains "normal" (i.e., nonsuperconducting), or abruptly at a characteristic temperature, T_c, if the metal becomes superconducting. **Figure 5** shows an example of a superconductor, Pb, both in the zero magnetic field (open circles) where it becomes superconducting (resistance drops to zero) at $T_c = 7.2$ K and in magnetic fields (triangles and filled circles) large enough to keep it normal down to at least 2 K. In this article, the concentration is on nonsuperconducting behavior. Below ~8 K, the normal resistivity of Pb in a magnetic field decreases more rapidly than linearly with temperature, and then saturates to a constant value at the lowest temperatures. **Figure 6** shows, in Pd, a more complex low-temperature resistivity than just the Bloch–Gruneisen form; here $\rho(T)$ contains two terms, a T^2 term arising from electron–electron scattering, plus an approximately T^5 Bloch–Gruneisen-like term. As T decreases, **Figures 5** and **6** show that the temperature-dependent part of $\rho(T)$ eventually becomes smaller than the resistivity due to residual defects, and the measured resistivity (designated as residual resistivity, ρ_0) becomes closely independent of temperature (constant) at a value determined by the nature and concentration of these defects.

If a concentration c of a given defect is added to a "pure" metal, the resulting total resistivity $\rho_T(c, T)$ can often be approximated by simply summing the temperature-dependent resistivity, $\rho_P(T)$, of the ideally pure host metal and a temperature-independent term, $\rho_0(c)$, due to the presence of the defect. This

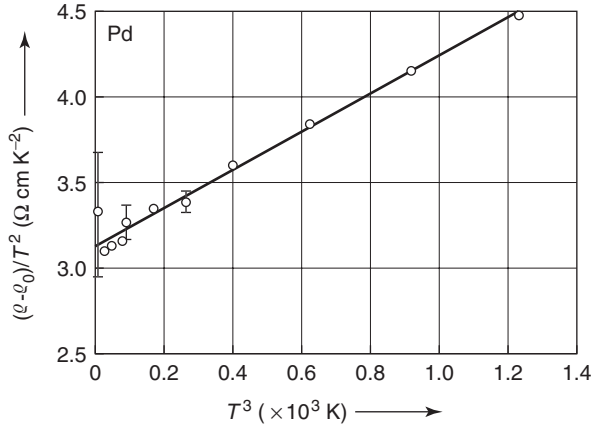


Figure 6 $(\rho - \rho_0)/T^2$ vs. T^3 for the transition metal Pd at low temperatures, where ρ_0 is the (constant) residual resistivity. In addition to an approximately T^5 Bloch–Gruneisen-like term, the data contain also a T^2 term due to electron–electron scattering.

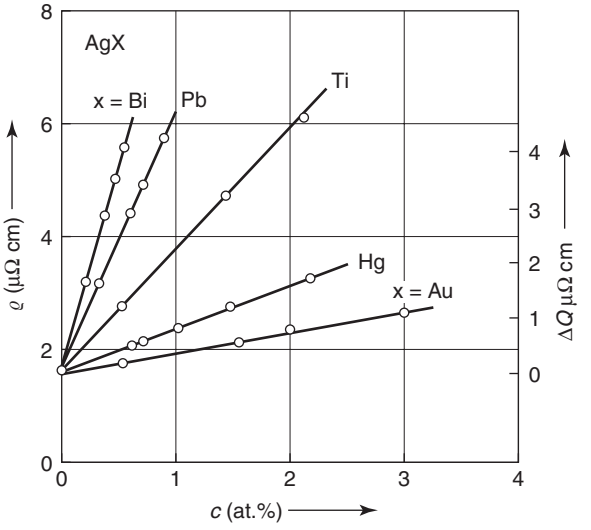


Figure 7 Resistivity vs. concentration (c in at.%) for Bi, Pb, Tl, Hg, and Au impurities in Ag, showing the expected linear variation with c, for small c.

result is called Matthiessen's rule:

$$\rho_T(c,T) \approx \rho_0(c) + \rho_P(T) \qquad [16]$$

The content of eqn [16] is twofold: (1) the most important feature of the defects is simply their presence, rather than that they vibrate differently from the host atoms; and (2) scattering of electrons from the defects is independent of scattering from the phonons. In practice, eqn [3] is used to define $\rho_P(T)$, by subtracting the low temperature constant limit of the data, ρ_0, from the measured $\rho(T)$ for the highest-purity samples available.

For small c, $\rho_0(c)$ is proportional to c, $\rho_0(c) = Kc$. The scale for K is $\mu\Omega\,\text{cm at.}\%^{-1}$, with values ranging from $0.1\,\mu\Omega\,\text{cm (at.}\%)^{-1}$ to $10\,\mu\Omega\,\text{cm (at.}\%)^{-1}$ for different impurities in different host metals. **Figure 7** shows examples of linear variations of $\rho_0(c)$ with c for dilute Ag-based alloys. Note that K increases with increasing difference between the number of "conduction electrons" for the impurity (Au = 1, Hg = 2, Tl = 3, Pb = 4, and Bi = 5) and for the host metal, Ag = 1. If the values of K for impurity A in host B and for impurity B in host A are closely the same, then the relation between $\rho_0(c)$ and c can be approximated over the whole range of alloy concentrations as

$$\rho_0(c) = Kc(1-c) \qquad [17]$$

In practice, the K values for different pairs of metals are usually different enough that eqn [17] is only a very rough approximation to reality. **Figure 8** shows an example of experimental data for PdAu alloys.

Careful measurements of $\rho_T(c,T)$ show that Matthiessen's rule is never exact. The resulting deviations from Matthiessen's rule (DMR) are defined by the

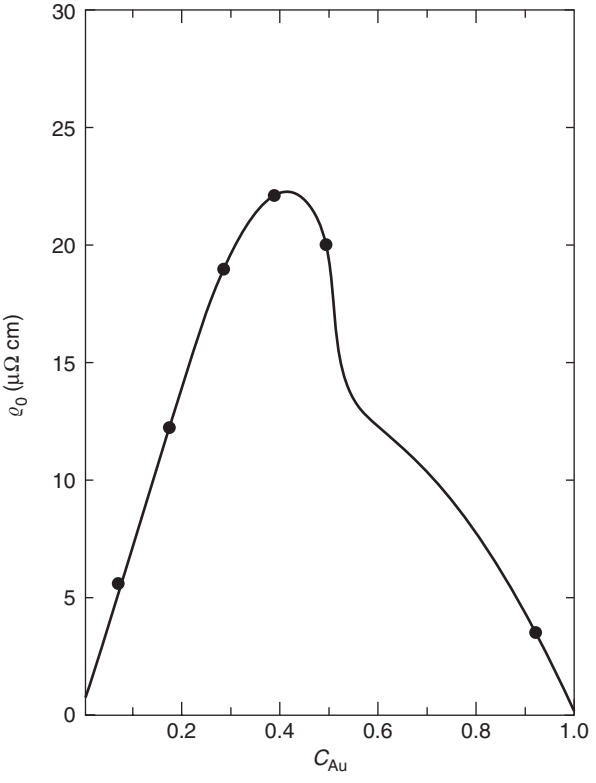


Figure 8 Residual resisitivity vs. Au concentration C_{Au} for PdAu alloys, showing that the residual resistivity need not be a symmetric function of C_{Au}.

relation

$$\Delta(c,T) = \rho_T(c,T) - \rho_0(c) - \rho_P(T) \qquad [18]$$

Such deviations have been studied for a variety of impurities and host metals. $\Delta(c,T)$ usually consists of

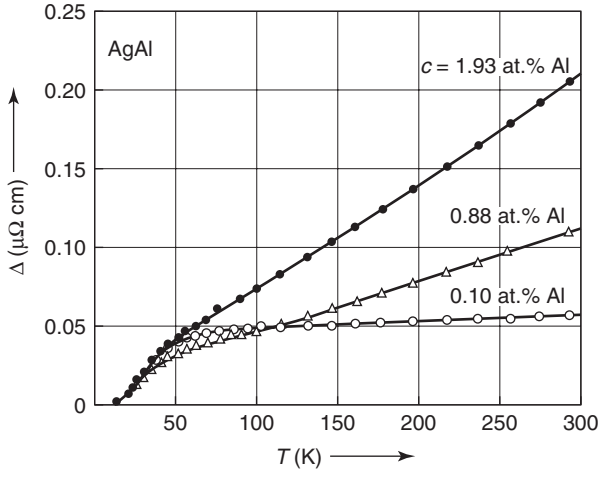


Figure 9 Deviation from Matthiessen's rule, Δ, vs. temperature for three alloys of Al in Ag.

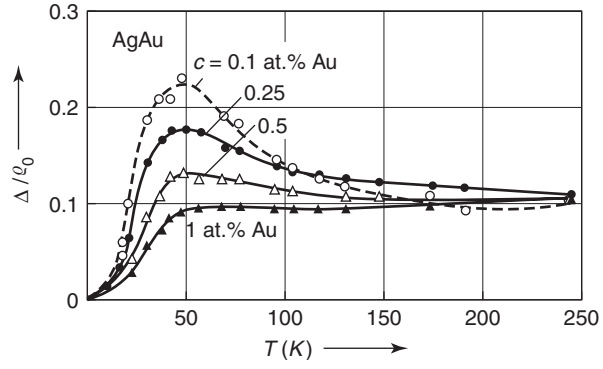


Figure 10 Normalized deviation from Matthiessen's rule, Δ/ρ_0, for a series of alloys of Au in Ag.

a low-temperature peak followed by a linear component at high T that can have positive, negative, or zero slope. **Figures 9** and **10** illustrate the forms seen.

Pressure

Measurements of ρ are almost invariably made at constant pressure P, usually at atmospheric pressure or in vacuum. Calculations of ρ, in contrast, are normally made by assuming constant volume V. Thus, to compare measurements and calculations, knowledge of how ρ changes with P is required. Studies of $\rho(P)$ are useful for providing such corrections (which are usually small) and also because measuring R is often the easiest way to study phenomena induced by an application of pressure, such as pressure-induced phase transitions. The relative ease of measuring R also makes it one of the properties of choice for studying phase transitions induced by temperature.

Size Effects (Mean Free Path)

If a sample wire or film is made so thin that there is a significant probability of electrons being scattered by the sample surface, and if such scattering is at least partly diffuse (i.e., with the final momentum of the scattered electron randomized in direction), then the resistance of the sample will increase due to such scattering. In the simplest case of a thin cylindrical wire of a free-electron metal of diameter d and completely diffuse scattering (i.e., the component of the electron's momentum along the wire or film is randomized by each collision), the total resistance $\rho_T(d)$ can be approximated by the relation

$$\rho_T(d) = \rho_B + (\rho\lambda)/d \qquad [19]$$

Here, ρ_B is the resistivity of the bulk material, and $\rho\lambda$ is a constant representative of the particular metal involved, as shown in eqn [8]. If one replaces ρ by the experimental resistivity of the metal at a particular temperature, then λ would be the experimental "mean free path" for free electrons in that metal at that temperature – that is, λ is the typical distance moved between scattering events that significantly change the direction of the electron's momentum. A qualitative insight into eqn [19] can be obtained by analogy with Matthiessen's rule, wherein the second term of eqn [19] represents the increase in resistivity when the sample diameter is equal to its mean free path. For partially specular scattering (mirror-like, in which the momentum component along the wire or film remains unchanged), or for the different geometry of a film, the equations are more complex, and the interested reader is referred to the Further Reading section. **Figure 11** shows examples of the applicability of eqn [19] for samples with different amounts of surface specularity.

Some Related Phenomena Outside the Purview of This Article

Magnetoresistance

The standard quasiclassical treatment of electrical conduction in metals outlined above can also be used to describe the phenomenon of magnetoresistance in nonmagnetic metals and alloys, as well as the changes in electrical resistivity that occur upon application of an external magnetic field.

Kondo Effect

In contrast, this standard quasiclassical analysis always yields, as a low-temperature limit, a constant resistivity, corresponding to the scattering of the conduction electrons by residual impurities. It could,

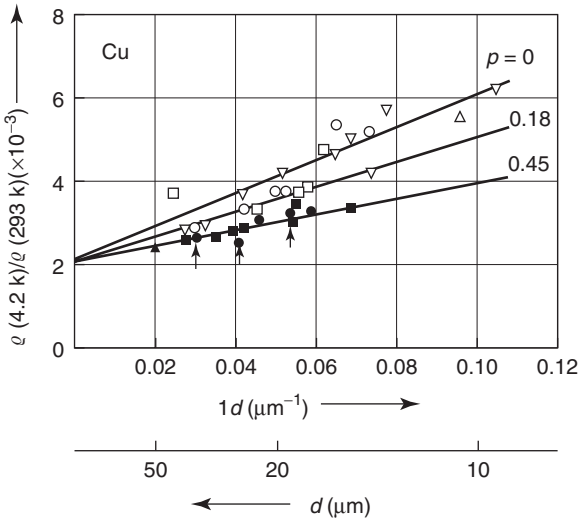


Figure 11 $\rho(4.2\,\mathrm{K})/\rho(293\,\mathrm{K})$ vs. inverse wire diameter, $1/d$, for a series of Cu wires with surfaces giving different amounts of diffuse scattering ($p = 0$ indicates estimated completely diffuse scattering).

thus, not explain observed deviations from this prediction for some magnetic impurities in nonmagnetic hosts, where the resistivity increases with decreasing temperature. This phenomenon was dubbed the "Kondo effect" when it was explained many years later by Kondo as arising from an anomalous scattering by a dilute concentration of localized magnetic moments in a metallic host. This phenomenon is described elsewhere in this encyclopedia.

Mesoscopic Effects

The standard quasiclassical analysis of metallic transport is also unable to explain mesoscopic effects in metallic samples with at least one dimension less than the phase coherence (or dephasing) length, which can be microns or more at cryogenic temperatures. These effects include weak localization in highly disordered metals, and oscillations with the magnetic field in multiply connected samples of size comparable to the dephasing length. To properly treat both the temperature dependence of transport in highly disordered metals, and the oscillatory response to magnetic fields of metallic rings with dimensions smaller than the dephasing length, one must take account of the interference between waves of electrons that follow different trajectories between two points in the sample.

See also: Ionic and Mixed Conductivity in Condensed Phases.

PACS: 72.15. – v; 72.15. – Eb; 72.10. – d

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Saunders College.

Bass J (1982, 1985) *Electrical Resistivity of Pure Metals and Alloys*. Landolt-Bornstein, New Series, Group III, V15a, 1; V15b, 1. Berlin: Springer.

Dugdale JS and Myers A (1985) *Electrical Resistivity of Pure Metals and Alloys*: *Pressure Effects*. Landolt-Bornstein, New Series, Group III, V15b, 13. Berlin: Springer.

Meaden GT (1965) *Electrical Resistance of Metals*. New York: Plenum Press.

Rossiter PL (1987) *The Electrical Resistivity of Metals and Alloys*. Cambridge: Cambridge University Press.

Rossiter P and Bass J (1992) Electronic transport properties of normal metals. In: Cahn R, Haasen P, and Kramer EJ (eds.) *Materials Science and Technology*, vol. 3A, pp. 257–320. Weinheim: VCH.

Schroder K (1983) *Handbook of Electrical Resistivities of Binary Metallic Alloys*. Boca Raton, FL: CRC Press.

Ziman JM (1964) *Principles of the Theory of Solids*. Cambridge: Cambridge University Press.

Conductivity, Thermal

K Behnia, Ecole Supérieure de Physique et de Chimie Industrielles, Paris, France

Definition

Heat travels inside solids due to a property called thermal conductivity. The magnitude of thermal conductivity is specific to each element, alloy, or compound and is defined by the equation relating the heat flow to the thermal gradient produced by this flow:

$$J_q = -\kappa \nabla T \qquad [1]$$

Here, J_q and ∇T are vectors representing respectively heat current density and temperature gradient. Thermal conductivity κ is, in general, a second-rank tensor. However, when heat flows along a high-symmetry axis, the temperature gradient created is parallel to the axis and the thermal conductivity along that particular axis may be treated as a scalar. In SI units, J_q is expressed in $\mathrm{W\,m^{-2}}$ and ∇T in $\mathrm{K\,m^{-1}}$. The SI unit of thermal conductivity is therefore $\mathrm{W(K\,m)^{-1}}$.

Equation [1] is reminiscent of Ohm's law defining the electrical conductivity as the quantity relating the electric field to current density $\boldsymbol{J} = -\sigma \boldsymbol{\nabla} V$. Indeed, there are a number of analogies between the two transport properties. There is a big difference however: to conduct electric charge, a solid should host itinerant electrons. Otherwise, it would qualify as an (electric) insulator. In the world of heat conduction, such a sharp distinction between conductors and insulators does not exist. The fundamental reason behind this resides in the existence of phonons, which can carry heat even in the absence of any electronic heat conductivity.

Thermal Conductivity and Thermal Diffusivity

Equation [1] is relevant in the case of thermal equilibrium. In general, the exchange of heat implies a variation of temperature with time. Conservation of energy implies

$$c\frac{\partial T}{\partial t} = -\nabla \boldsymbol{J}_q \qquad [2]$$

where c is the heat capacity per unit volume of the compound. Since the specific heat of various compounds is usually expressed as heat capacity per unit mass (or molar mass), it should be multiplied by the density (or the molar density) of the compound to yield the c used here. A finite $\boldsymbol{J}_q$ means that the balance of heat entering to and exiting out of a volume element is not zero and the energy surplus (or deficit) changes the local temperature.

The combination of eqns [1] and [2] yields the so-called equation of heat, also called the Fourier equation:

$$\frac{\partial T}{\partial t} = D\nabla^2 T \qquad [3]$$

where $D = \kappa/c$ is called the thermal diffusivity. It has the dimensions of an area divided by time ($\mathrm{m^2\,s^{-1}}$ in SI units). Thermal diffusivity is thus thermal conductivity divided by specific heat.

Kinetic Theory of Heat Carriers

Electrons and phonons are the principal carriers of heat in a solid. Phonons are the elementary excitations of a vibrating crystal. In terms of solid-state physics, it is a particle with a definite energy and wave vector. Electrons here mean itinerant electrons capable of moving from one atom to the other. In order to grasp the basic differences between these two types of carriers, it is instructive to focus on a formula derived from the kinetic theory of gases. In this picture, the thermal conductivity is expressed as

$$\kappa = 1/3cvl \qquad [4]$$

where c is the specific heat per volume, v is the average velocity of the particles composing the gas, and l is the mean free path.

Phonons are bosonic quasiparticles with a specific heat displaying a T^3 temperature dependence at low temperature and then saturating to a constant value at temperatures of the order of the Debye temperature, Θ_D. The relevant velocity is the speed of sound (of the order of several $\mathrm{km\,s^{-1}}$ in a solid) and does not change much with temperature. Finally, the mean free path of phonons is limited by whatever impedes their free propagation in a crystal. At sufficiently low temperatures, nothing can scatter extended phonons of long wavelength. In this so-called "ballistic" regime, the mean free path of the phonon becomes comparable to the sample's dimensions.

Electrons are fermions. Their specific heat is a linear function of temperature in ordinary temperatures, which remain much smaller than the Fermi temperature of most metals. The relevant velocity for them is the Fermi velocity (of the order of several hundreds of $\mathrm{km\,s^{-1}}$ in common metals). Finally, even in the cleanest samples and at low temperatures, their mean free path seldom exceeds a micrometer.

Thermal Transport in Insulating Crystals

The case of insulators is simpler as there are no electrons to conduct or to scatter, and their thermal conductivity reflects the capacity of the lattice alone to conduct heat.

The Debye temperature sets the most important energy scale in this context, Θ_D. The thermal conductivity of an insulator peaks at a temperature, which is a fraction (often one-tenth) of Θ_D. This maximum marks the border between the two regimes: while at higher temperatures it is a phonon–phonon scattering which determines the phonon mean free path, l_{ph}, at lower temperatures, collision with defects becomes the predominant limit on l_{ph}. These two regimes are considered here separately:

1. *Low temperature* (*below* $T_{peak} \sim 0.1\Theta_D$). At this temperature range, the T^3 specific heat dominates the temperature dependence of thermal conductivity since l_{ph} displays a weak temperature dependence. With increasing temperature, l_{ph} decreases from the maximum attained in the $T = 0$ limit. Its temperature dependence is governed by the presence of defects in the crystal which scatter phonons in various ways. The cross section of a

phonon with a specific scattering center (such as a point defect, a line defect, a planar defect, or a dislocation) is a function of the phonon wavelength λ_{ph}. As the temperature rises, the dominant wavelength in the phonon population increases ($\lambda_{ph} \propto k_B T$). Therefore, the relative importance of different scattering mechanisms is shifted with the change in temperature.

2. *High temperature* (*above* $T_{peak} \sim 0.1\Theta_D$). The maximum is related to the inflection in the temperature dependence of C_{ph} and l_{ph}. It marks the beginning of the predominance of the phonon–phonon scattering, which leads to a sharper decrease in l_{ph}. This is a consequence of the rise in the phonon population with increasing temperature. On the other hand, the specific heat increases less rapidly and eventually saturates to a constant value. Therefore, the temperature dependence of κ is governed by $l_{ph}(T)$. The expected temperature dependence of the latter is T^{-1}. The minimum thermal conductivity at high temperature is imposed by the lattice spacing. Indeed, l_{ph} cannot become significantly lower than the latter.

The Case of Glasses

In amorphous solids, the mean free path of the phonon, even at low temperatures, does not exceed a few lattice constants. Therefore, the magnitude of thermal conductivity is drastically reduced. A generic feature of glasses is the absence of a maximum in the temperature dependence of thermal conductivity. Since l_{ph} presents a weak temperature dependence, the overall behavior is governed by the variation of the specific heat. The difference of conductivities between a glass and a crystal is largely reduced at high temperatures, because of the short l_{ph} for traveling phonons in both cases. In the intermediate temperature range, the peak is often replaced with a plateau (**Figure 1**).

Thermal Conductivity of Metals and the Wiedemann–Franz Law

In metals, heat is carried both by phonons and electrons. Lattice conductivity in metals suffers from the presence of electrons as additional scattering centers. In the peak region of thermal conductivity (around $0.1\Theta_D$), clean insulators can surpass clean metals as the best thermal conductors. In other words, the additional electronic thermal conductivity cannot compensate the loss in lattice thermal conductivity. The latter is proportional to the strength of electron–phonon coupling and introduces a T-square term to the temperature dependence of the phonon thermal conductivity.

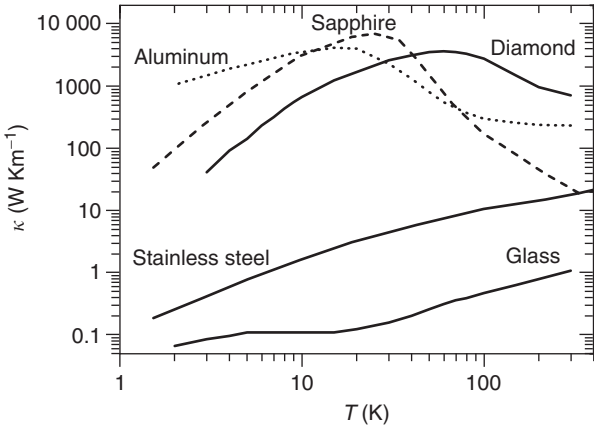


Figure 1 Thermal conductivity of insulating and metallic solids.

In most cases, the predominant contribution to heat transport in a metal is of electronic origin. The application of eqn [2] to electrons yields:

$$\kappa_{el} = 1/3 C_{el} v_F l_{el} = 1/3 C_{el} v_F^2 \tau \qquad [5]$$

where τ is the scattering time. This equation has a parenthood with the classical expression for electric conductivity:

$$\sigma = \frac{ne^2\tau}{m} \qquad [6]$$

with e and m representing the charge and the mass of electron, and n the carrier density. In both expressions, the conductivity is proportional to the scattering time. Thus, the ratio of the two conductivities is expected to be independent of scattering. Divided by temperature, this ratio yields the Lorentz number: $L = \kappa_{el}/\sigma T$. The Wiedemann–Franz law which establishes an intimate relationship between thermal conductivity and electric conductivity of a given metal, states that the Lorentz number is equal to a universal constant:

$$\frac{\kappa_{el}}{\sigma T} = \frac{\pi^2}{3}\left(\frac{k_B}{e}\right)^2 = L_0 \qquad [7]$$

where k_B is the Boltzmann constant and L_0 is the Sommerfeld number. Wiedemann and Franz formulated their empirical law in 1853 after measuring the conductivities of various metals at room temperature. The quantum theory of metals gave a solid foundation to this law which is expected to hold in the $T = 0$ limit irrespective of almost any particularity of any known metal. The fundamental reason behind its universality is easy to grasp. Whatever impedes the transport of charge by electrons would also affect the transport of entropy by them. The

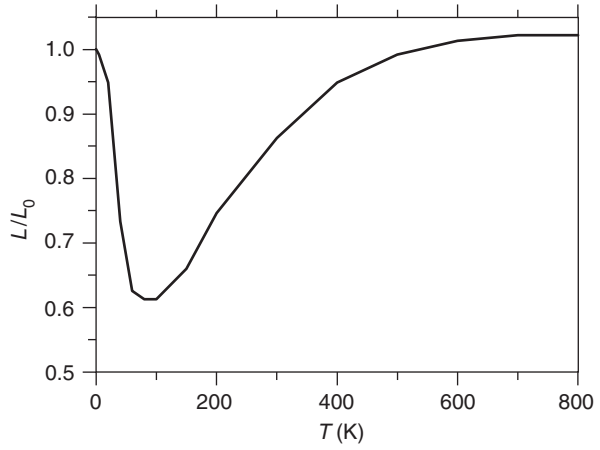


Figure 2 The temperature dependence of the Lorentz number in a common metal (aluminum).

ratio of the two conductivities, therefore, does not depend on the specific electronic properties of the metal.

The argument remains rigorously valid in the absence of inelastic collisions, that is, scattering events which do not imply a transfer of energy during the scattering. Such events occur at finite temperatures when electrons are scattered by phonons and disturb the transport of heat more efficiently. Indeed, such a collision may lead to a partial loss of the electron's energy (i.e., heat) without affecting its momentum producing a thermal resistance with no counterpart in the charge channel. At high temperatures, however, the loss of energy during inelastic collisions becomes negligible compared to the thermal energy, and the validity of the Wiedemann–Franz law is more or less restored (see **Figure 2**).

Superconductivity and Heat Transport

Many metals become superconductors below a critical temperature, T_c. The transition has drastic consequences on heat transport both by electrons and by phonons. At the onset of superconductivity, a gap opens up in the energy spectrum of the electronic spectrum and individual electrons end up in a superconducting condensate. This condensate lacks any degree of freedom in order to qualify as a carrier of entropy. Therefore, in the superconducting state, only the "normal fluid," that is, electronic excitations which happen to exist due to their thermal energy at finite temperature, can carry heat. The number of such excitations decreases exponentially with temperature and at sufficiently low temperature, they cease to play any significant role in thermal transport.

Phonon heat transport is also affected by the superconducting transition. At the onset of superconductivity in many metals, electrons are a principal source of scattering for phonons. When this is the case, the condensation of electrons leads to a sharp enhancement in the phonon mean free path and the phonon thermal conductivity. At very low temperatures, an ordinary superconductor conducts heat in a manner analogous to an insulator.

The other consequences of the superconducting transition on the thermal conductivity of electrons and phonons generate a wide variety of behaviors in real superconductors. If the electrons are the dominant heat carriers at the onset of superconductivity, then the superconducting transition is accompanied by a rapid decrease in the thermal conductivity of the system. On the other hand, if for any reason (a low carrier density, a low mean-free path, a high critical temperature, etc.) the lattice contribution happens to be significant in the normal state, the thermal conductivity can increase below T_c reflecting the enhancement of l_{ph}. The superconductor should be clean enough for this to happen, otherwise the impurity scattering of phonons is such that the disappearance of electrons would not present any significant change in phonon scattering. Thus, according to the numerous relevant parameters of each case, the superconducting transition can lead to a reduction, to an enhancement, or to no change in the thermal conductivity of the system.

The picture of heat transport in superconductors in this article has been somewhat modified by the results of the last two decades of research on unconventional superconductors. In high-T_c cuprates, the entry of the system into the superconducting state is accompanied with an upturn in thermal conductivity. It is now generally believed that part of this enhancement is due to an increase in the electronic component of thermal conductivity. While there is still no microscopic understanding of what happens, this extraordinary situation is interpreted as a signature of strong electron–electron scattering present in the normal state of the cuprates. In other words, the suppression of electrons leads to a lower number of heat carriers with a longer mean-free path in such a way that the overall result is a significant enhancement of thermal conductivity. A second feature of interest is the complex structure of the superconducting gap in unconventional superconductors. Since their gap is anisotropic with a magnitude which becomes zero along definite orientations, low-energy electronic excitations can survive down to $T = 0$. The electronic thermal conductivity is not expected to be an exponentially decreasing function of temperature. This has been confirmed by experiments probing heat transport by these "nodal" quasiparticles.

Magnetic Excitations as Heat Carriers

In a magnetically ordered solid, a new type of heat carriers appear. The elementary excitations of a magnetic lattice are called spin waves or magnons and can transport heat in a manner analogous to phonons. Most of the considerations regarding phonons are relevant for magnon heat transport with Curie or Néel temperature replacing the Debye temperature as the relevant energy scale. A quantitative isolation of the magnon contribution to heat transport is, however, a difficult experimental task as the deconvolution of various terms in their presence is far from obvious.

Entropy carriers associated with spin degrees of freedom are also expected in metals close to a magnetic stability. Such excitations, often called spin fluctuations or paramagnons, are reported to produce a substantial enhancement of thermal conductivity in a number of exotic metals.

Thermal Magnetoresistivity and the Righi–Leduc Effect

A magnetic field has no direct effect on heat transport by phonons. Electronic thermal conductivity, on the other hand, is affected by a magnetic field in a manner analogous to the electric conductivity. Usually, it is diminished as the magnetic field disfavors transport by charged particles.

In type II superconductors, a magnetic field penetrates the solid in the shape of mesoscopic filaments called vortices. These superconducting vortices constitute a new type of scattering centers for heat carriers. The contribution of their own movement to heat transport is often negligible.

In the presence of a magnetic field, the electronic thermal conductivity tensor presents a nondiagonal term, κ_{xy}. The trajectory of charged carriers is skewed under the influence of the Lorentz force (or magnetically aligned scattering centers). This creates a transverse heat current, which is balanced by a transverse thermal gradient. This is the Righi–Leduc effect, sometimes dubbed the thermal Hall effect in analogy to the famous electric one.

Measuring Thermal Conductivity

Experimental methods to determine thermal conductivity fall into two broad groups. Steady state or "static" methods study a sample time-independent temperature profile and measure κ using eqn [1]. In contrast, non-steady-state or "dynamic" methods work with a time-dependent temperature. They often yield a quantity akin to thermal diffusivity.

Measuring the thermal conductivity of a solid implies a reliable determination of the temperature gradient produced by a heat current, and the quality of an experimental setup depends on its ability to yield accurate values of these two quantities. This proves to be more difficult than the analogous case of electrical conductivity since charge can only travel through conductors but heat can be transmitted by conductors, insulators, and even vacuum.

See also: Magnetoresistance Techniques Applied to Fermi Surfaces.

PACS: 72.15.Eb; 66.70.+f; 72.15.Cz

Further Reading

Berman R (1976) *Thermal Conduction in Solids*. Oxford: Oxford University Press.

Tye RP (ed.) (1969) *Thermal Conductivity*, vols. I and II. London and New York: Academic Press.

Confocal Optical Microscopy

M J Solomon and M Kogan, University of Michigan, Ann Arbor, MI, USA

Introduction

Confocal optical microscopy is a direct visualization method that is increasingly applied in condensed matter physics to study the structure and dynamics of complex fluids. The focus of this article is the theory and practice of confocal microscopy as it relates to the areas of colloid physics, experimental statistical mechanics, and microrheology. Available instrumentation and methods are classified and their limitations discussed. Auxiliary methods, such as fluorescent particle synthesis and quantitative image processing, are also summarized. Finally, current applications of confocal optical microscopy in condensed matter physics are discussed.

Principle of Confocal Optical Microscopy

Standard optical microscopy techniques, including microscopy by bright field and epifluorescence illumination, fail for the optically dense specimens that are typical of applications in materials science and condensed matter physics. The origin of the optical density is often a combination of scattering due to strong refractive index contrast, emission due to high fluorophore density, or attenuation due to the long optical path of an object plane located deep in the specimen. This problem was first encountered and addressed in the life sciences. There, methods were sought to image deep into cellular tissue without the need for laborious sectioning. The technology developed, called confocal optical microscopy, combines optical components typical of microscopy and scattering devices. In particular, the detection path typical of optical microscopy is modified by the addition of a confocal pinhole at the image plane. This pinhole arrangement is typical of light scattering detection optics in that the aperture rejects light not originating from the vicinity of the object plane. The rejection renders an image of the object plane detectable that would otherwise be overwhelmed by unfocused out-of-plane emission. When the confocal detection scheme is combined with the capability to actuate the object plane through the specimen, a powerful technique for the imaging of three-dimensional (3D) volumes is realized. While fluorescence imaging is the most common application of confocal microscopy, visualization of structure by scattered or reflected light is also feasible. **Figure 1** shows the difference in image forming capability for a gelled colloidal suspension (volume fraction of 10%) imaged by confocal and epifluorescence microscopy.

The direct visualization of image volumes by confocal microscopy has been fruitfully applied in many areas of condensed matter physics. In experimental statistical mechanics, application of quantitative image processing to 3D intensity maps generated by confocal microscopy has been used to quantify the thermodynamic and structural properties of glasses and gels comprised of colloidal particles. The field-induced and template-induced crystallization of novel colloidal phases has been observed and quantified. The 3D structure of mesoporous phases has been discovered. The diffusivity of fluorescent probes in polymer networks has been quantified. Particle tracking routines applied to time series of confocal image volumes have observed cooperative motion in colloidal suspensions. These observations can be used to understand the fundamental origins of the glass transition.

These aforementioned applications require understanding of the spatial and temporal resolution of confocal microscopy, application of image processing tools to quantify acquired image volumes, and synthesis of model materials suitable for confocal imaging. These issues are discussed in subsequent sections. First, the device configurations commonly used in confocal microscopy are classified.

Confocal Laser Scanning Microscopy

The conceptual idea of rejecting out-of-focus emission from the detector by adding an aperture to the detection system has been implemented in two principal ways. **Figure 2a** shows the confocal laser scanning microscope (CLSM). The confocal aperture is

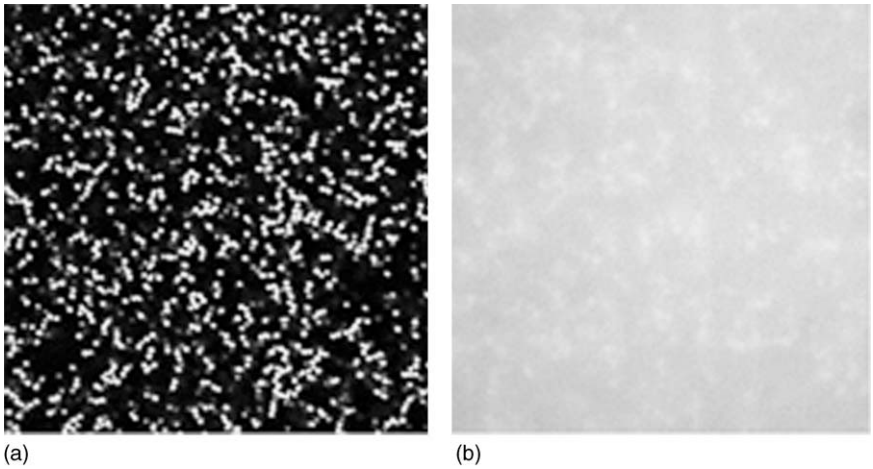


Figure 1 A depletion gel of fluorescent poly(methyl methacrylate) colloids $\sim 1.9\,\mu m$ in diameter dispersed in a refractive index and density-matched solvent imaged with a 1.4 NA objective by (a) confocal fluorescence and (b) epifluorescence microscopy. The comparison demonstrates the image-forming capability of confocal microscopy for the optically dense specimens that are characteristic of condensed matter physics and materials science.

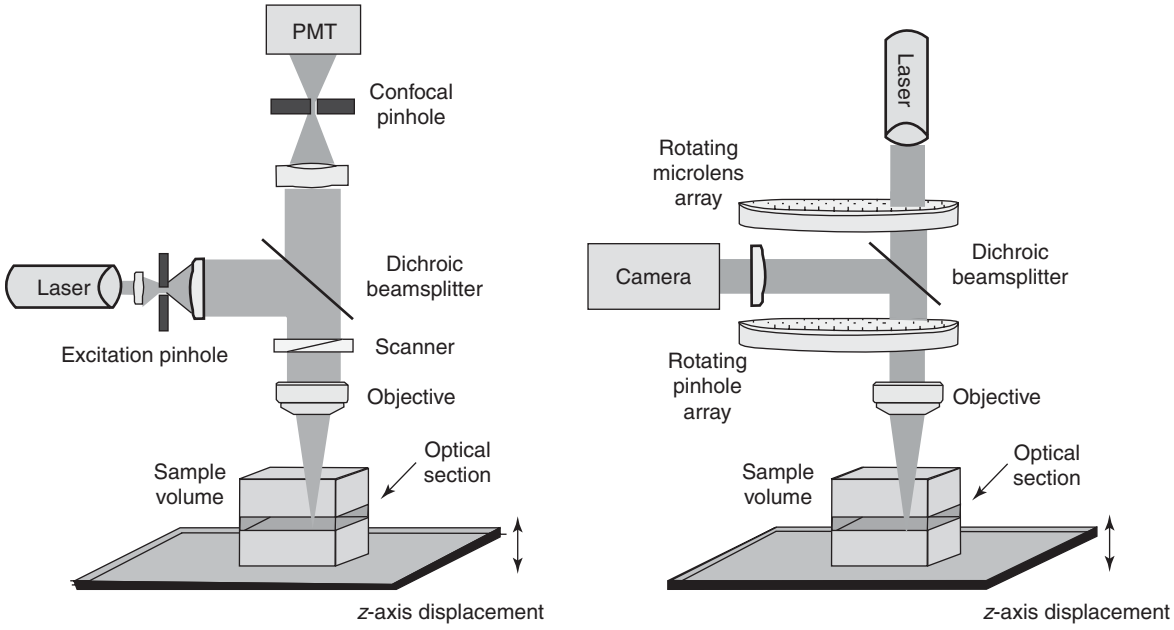


Figure 2 Schematics of the: (a) confocal laser scanning microscope and (b) Nipkow disk confocal microscope.

adjusted to accomodate the emission wavelength and to optimize the trade-off between axial resolution and emission collection efficiency inherent to confocal microscopy. In this configuration, the image is collected pointwise with a single detector, usually a photodiode or photomultiplier tube. In order to ensure sufficient emission from the excitation point in the objective plane, laser illumination is typical. Scanning mirrors manipulate the path of an expanded, spatially filtered source so that the illuminated point is rastered across the object plane. The speed of the galvanometrically actuated scanning mirrors determines the acquisition rate of a 2D image. At current capabilities a line scan can be executed at rates up to $\sim 1\,\text{kHz}$. Thus, a 512×512 pixel scan can be accomplished in less than 1 s. 3D image volumes are acquired by piezoelectric actuation of the specimen relative to the objective. This actuation can be conducted with nanoscale resolution and micrometer-scale repeatability. A $512 \times 512 \times 200$ 3D scan, a typical resolution, requires a few minutes with a best available CLSM.

Just as in the epifluorescence technique, in CLSM the excitation and emission illumination traverse the microscope objective in reverse directions. Typically, a dichroic beam splitter ensures that only emission from the specimen passes to the detector optics. Alternatively, an acoustooptic beam splitter provides flexibility in selecting transmission bands. CLSM is commonly used in fluorescence mode, and in some applications, it is of interest to monitor the emission of multiple fluorophores. Such detection is accomplished by excitation with multiple laser sources. The detection system must then separate the emission spectrum into bands corresponding to the wavelength ranges relevant to the multiple fluorophores. This decomposition can be accomplished with a cascade of dichroic filters. Alternatively, after the confocal pinhole, the emission can be spectrally resolved with a prism and adjustably filtered with mechanical slits before detection with a photomultiplier tube. Detection of up to four spectral regions is common. While fluorescence visualization is a frequent application of CLSM, some situations require analysis of reflected or scattered light. This method, called reflection CLSM, is of particular interest for measurements of topography in materials science and solid-state electronics.

The confocal imaging system is used in conjunction with the components of a standard optical microscope. In particular, the microscope objectives and specimen preparation for CLSM and optical microscopy are identical.

Variations on the design of **Figure 2a** include substitution of the confocal pinhole for a confocal slit. This configuration allows faster scanning with reduced resolution.

Nipkow Disk Confocal Microscopy

The pointwise scanning of CLSM imposes a significant limitation on the image acquisition speed of this method. The Nipkow disk confocal microscope, shown in **Figure 2b**, is an alternate method that can achieve video rate imaging. Current configurations

use two synchronized disks rotating at a rate that scans the image at 360 frames s^{-1}. One disk is patterned with an array of many thousands of pinholes. The pinhole diameter is optimized for high numerical aperture (NA) objective microscopy (optimization for NA = 1.4 is typical). The excitation illumination is chopped by the pinhole array. A second rotating disk, located prior to the pinhole disk, consists of a microlens array. Each microlens focuses excitation light onto a corresponding pinhole, thereby significantly increasing the excitation intensity supplied to the specimen. That is, the pinhole spacing must be significant relative to its diameter to reduce collection of defocused emission from one pinhole by a neighboring one. Thus, the pinhole disk has inherently low transmission efficiency without incorporation of the microlens array.

Emission from the specimen propagates through the rotating pinhole array in the reverse sense. In this case, each pinhole serves as the confocal aperture. The emission paths are diverted to the detection system by a dichroic beamsplitter. A photodiode array images the emission paths. Because the primary excitation is subdivided into a number of sources by the pinhole disk, multiple points of the objective plane are simultaneously interrogated. The read-out rate of the digital camera detector is currently the limiting step in the time required to acquire a 2D image. 3D image volumes are acquired by relative actuation of specimen and objective, just as for CLSM.

At full frame resolution, image acquisition rates for Nipkow disk confocal microscopy are about an order of magnitude faster than CLSM. However, because the primary incident laser source is subdivided into a number of beams, the excitation supplied to a particular point in the object plane is typically less than CLSM. In addition, the confocal aperture is not adjustable as is usually the case in CLSM.

Limits of Optical Resolution

Confocal microscopy presents unique opportunities for 3D imaging of microstructure; however, the method is an optical one and thus resolution limits are dictated by the wavelength of light. Here, two limits of optical resolution are considered. The first, resolution in the object plane (lateral optical resolution), is governed by the Rayleigh criterion. The second, resolution perpendicular to the object plane (axial optical resolution or optical section thickness), is dictated by the point spread function. The effect of confocal aperture diameter on resolution is also assessed.

The Rayleigh criterion labels two adjacent points in the object plane distinguishable if their airy diffraction disks (due to the finite aperture of objective) are resolved in the image plane. For unpolarized light, this lateral optical resolution is $0.61\lambda/\mathrm{NA}$ where λ is the wavelength of light and NA is the numerical aperture of the objective. For NA = 1.4, lateral resolution is $\sim 0.2\,\mu m$ and this is the best resolution currently possible for optical microscopy. The Rayleigh criterion is a useful estimate of lateral resolution in confocal microscopy.

In confocal microscopy, the criterion for axial object resolution commonly adopted is that two points oriented perpendicular to the object plane can be resolved if the first minima in their point spread functions (evaluated along the optical axis) are distinct. This criterion has the same physical basis as the one for the Rayleigh resolution and leads to a characteristic axial optical resolution of $2\lambda/(\mathrm{NA})^2$. For NA = 1.4, axial resolution is $\sim 0.5\,\mu m$, slightly greater than twice the lateral object resolution.

These resolution criteria are lower bounds. Actual devices achieve lower resolution due to polarization effects and nonidealities. For example, the confocal aperture diameter has a profound effect on optical section thickness. In scanning devices, this diameter is adjusted to optimize the trade-off between image intensity and axial resolution. Units scaled on characteristics of the diffraction point spread function are a convenient basis to specify the aperture diameter and its performance. The dimensionless lateral scale in the image plane is then $(2\pi/\lambda)(\mathrm{NA}/M)r_a$, where M is the magnification in the image plane and r_a is a particular image plane dimension such as the aperture radius.

Calculations and experiments for confocal visualization of well-defined geometries show that optical section thickness within a few percent of minimum is attained with a confocal aperture dimension that corresponds to one airy diffraction disk (cf. dimensional scaling above). Aperture reduction below this value yields only small improvement in resolution at the expense of reduced image intensity. As the confocal aperture is opened, progressively greater amounts of out-of-focus emission are admitted to the detector; however, resolution deteriorates. One way in which this effect has been quantified is by imaging a plane reflective coating. In this experiment, the half-width at half maximum of the axial intensity, a measure of the optical thickness, doubles as the aperture is increased from one airy disk to two. This experiment demonstrates that axial resolution is a severe function of the confocal aperture. **Figure 3** displays images that show the qualitative effect of confocal aperture on image intensity and resolution as the aperture is opened for a colloidal crystal of interest to soft condensed matter physics and materials science.

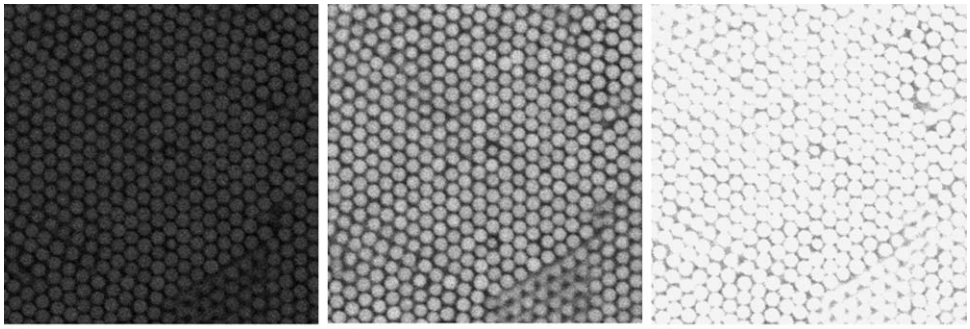


Figure 3 The effect of pinhole aperture on the image-forming capability of confocal laser scanning microscopy. In optical units nondimensionalized on the size of the airy diffraction disk, the images are for confocal apertures of 0.5, 1.0, and 2.0 (from left to right). The specimen is a colloidal crystal of colloidal poly(methyl methacrylate), diameter $\sim 1.9\,\mu m$, in a refractive index and density-matched solvent for which electrostatic interactions between the colloids have been screened by the addition of an organic salt.

The resolution of scanning confocal microscopes with slit apertures, developed for fast imaging, has also been studied by imaging a plane reflective mirror. In this case, the minimum axial thickness in the specimen achieved by a slit is $\sim 25\%$ greater than for a pinhole. In addition, in optical units, the half-width of the slit must be about a factor of 3 smaller than the pinhole radius to realize such a minimum axial resolution. This slit width reduces the intensity of the image received at the detector.

Model Materials for Confocal Imaging

Most applications of confocal microscopy in condensed matter physics have used fluorescent imaging. The need to incorporate fluorophores into structures to be imaged imposes a constraint on materials that can be studied. Thus, studies of experimental statistical mechanics and self-assembly using colloidal models require methods for syntheses of monodisperse colloidal particles. Determination of bulk behavior requires imaging deep within the sample, where effects of the sample boundary are minimized. To accomplish such imaging, the particle and suspending medium must be at least approximately refractive index matched; otherwise scattering reduces the intensity and image-forming quality of light returned to the detector.

A common approach to this challenge is to synthesize monodisperse colloids to which an oligomeric stabilizing layer is grafted to the particle surface. This steric layer ensures colloid stability in mixtures of organic solvents that match the refractive index of the colloid. A number of strategies exist to incorporate fluorescent dyes into silica or poly(methyl methacrylate) colloids during particle syntheses. For polymer colloids, the postsynthesis method of colloid swelling and dye absorption is also possible; however, these processes change the potential interactions between the colloids in ways that are poorly understood.

In addition to refractive index matching, solvent mixtures are often selected to match the density of particle and solvent so as to minimize colloid sedimentation (or flotation). Such particle motion, even when small, can destroy fragile gel or crystalline structures that arise due to the physical interactions between particles. Sedimentation also interferes with the study of Brownian and cooperative particle motion by particle tracking.

Long duration kinetic and dynamic studies require fluorescent dyes that are resistant to photobleaching. In the life sciences, the limitation of cytotoxity on dye concentration means that photobleaching is an important factor in dye selection. While the situation is less severe in condensed matter physics (because cytotoxic effects are absent), recent work has favored more photostable dyes over traditional ones such as fluorescein and rhodamine.

Quantitative Image Processing for Confocal Microscopy

The 3D map of fluorescent intensity provided by confocal microscopy is a powerful qualitative indicator of complex fluid behavior. An even more fruitful approach is to use image processing to extract quantitative measures of structure and dynamics from the fluorescence maps. In 2D microscopy, a number of approaches for the interrogation of particle locations using edge detection and thresholding have been developed. Alternatively, particles can be identified from local intensity maxima and their locations then refined to subpixel resolution according to the distribution of brightness in the vicinity of the

maxima. This latter approach, although computationally intensive, is readily extendable to the 3D particle location problem and has thus found wide application in confocal microscopy of colloidal systems. While estimates of the errors in particle location for these routines vary widely, the most significant determinant of the precision and accuracy of particle location is the delocalizing effect of Brownian motion on the timescale of image volume acquisition. For time series acquired at a rate such that the typical particle displacement is much less than the mean separation between particles, algorithms have been developed that reliably link particle locations in each frame into trajectories. This trajectory analysis yields dynamical information about Brownian and cooperative motion of the particles.

Once particle centroids and trajectories have been located, the information contained in the confocal image volumes has been transformed into output in the same format that would result from a computer simulation of a corresponding model system. (In a typical image volume of $40 \times 40 \times 20\,\mu m^3$ resolved by a high-resolution objective, the number of micrometer-sized particles is of order 10^4. This number is not too different from the state-of-the-art computer simulation.) Thus, all the tools of statistical mechanics developed for the analysis of molecular, Brownian, and Stokesian dynamics simulations can be applied to the experimental results. This formal correspondence between experiment and simulation offers an unusually rich avenue for direct comparison between the two. **Figure 4** displays the results of quantitative image processing of a depletion colloidal gel comprised of monodisperse fluorescent poly(methyl methacrylate) colloids of diameter $\sim 1.9\,\mu m$. The projections of the confocal image volume in two orthogonal planes qualitatively demonstrate the degree to which image processing methods locate the imaged particles.

Applications in Condensed Matter Physics

The qualitative fluorescence imaging and 3D sectioning capability of confocal microscopy have found wide application in the qualitative characterization of microstructure in colloidal suspensions, polymer gels, immiscible blends, and composites. Recently, the additional step of generating quantitative measures of complex fluid structure and dynamics from confocal microscopy has been taken. In this brief summary, the focus is on these new developments. These advances have been driven by three recent steps forward: (1) methods to synthesize monodisperse fluorescent colloidal particles have been discovered; (2) confocal microscopy, borrowed from the life sciences, has been applied to image materials comprised of colloidal particles; (3) techniques of quantitative image processing have been developed to extract the locations and trajectories of the colloidal particles.

Confocal microscopy has been used to interrogate the phase behavior of concentrated colloidal suspensions interacting through screened electrostatic forces. The self-assembly of colloidal particles into ordered crystals is a step in one method to produce photonic materials. The effect of an applied electric field on colloidal crystal structure has also been discovered. The nucleation and growth of crystal nuclei in colloidal suspensions has been directly visualized. The particle correlation function and coordination

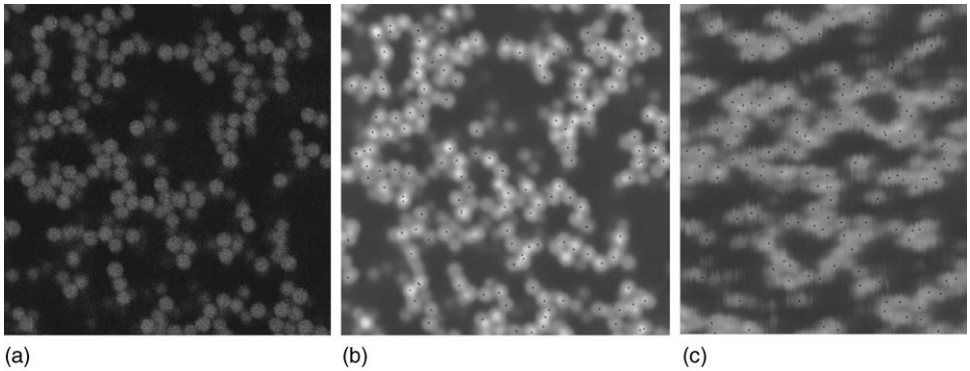


Figure 4 Illustration of methods for quantitative image processing. Image (a) is a representative 2D slice from a 3D image volume of a depletion gel of poly(methyl methacrylate) colloids (volume fraction of 10%, diameter 1.9 μm) imaged by CLSM. Image (b) shows a projection of ~20 2D lateral slices with an overlay of particle locations found within that volume by quantitative image processing. Image (c) shows a similar projection, but now in a plane that includes the axial direction of the microscope. Note that the optical resolution in the axial direction is reduced relative to the lateral direction because of the different characteristics of the 3D point spread function along the optical axis and in the objective plane.

number of hard-sphere glasses have also been quantified by direct visualization with confocal microscopy. Short-range attractive forces, included by conformational changes of grafted steric layers or by the depletion interactions induced by nonadsorbing polymer, lead to the assembly of amorphous gel structures. The clusters and voids that comprise these structures have been directly visualized with CLSM. Cooperative particle dynamics in the vicinity of the colloidal hard-sphere glass transition have been visualized by time-resolved confocal microscopy. The thermal fluctuations of gel backbone segments have likewise been quantified as have the flow-induced structures of colloidal gels as they are subjected to shear and compressional flow. These direct visualization studies of structure and dynamics in colloidal systems by confocal optical microscopy represent a new and fruitful approach to understand the behaviors of complex fluids that are of broad interest in soft condensed matter physics and materials science.

See also: Crystal Optics; Geometrical Optics; Imaging and Optical Processing; Interaction of Light and Matter; Optical Microscopy.

PACS: 87.64.Tt; 87.64.Rr; 42.30.Va; 82.70.Dd

Further Reading

Crocker JC and Grier DG (1996) Methods of digital video microscopy for colloidal studies. *Journal of Colloid and Interface Science* 179(1): 298–310.

Dinsmore AD, Weeks ER, Prasad V, Levitt AC, and Weitz DA (2001) Three-dimensional confocal microscopy of colloids. *Applied Optics* 40(24): 4152–4159.

Habdas P and Weeks ER (2002) Video microscopy of colloidal suspensions and colloidal crystals. *Current Opinion in Colloid and Interface Science* 7: 196–203.

Pawley JB (ed.) (1995) *Handbook of Biological Confocal Microscopy*, 2nd edn. New York: Plenum.

Van Blaaderen A (1998) From the de Broglie to visible wavelengths: Manipulating electrons and photons with colloids. *MRS Bulletin* 39–43.

Copolymers

C Hagiopol, Georgia-Pacific Resins, Inc., Decatur, GA, USA

Introduction

Polymers and copolymers are macromolecular compounds. Polymers (homopolymers) are macromolecules with a uniform structure. Copolymers show heterogeneous compositions. Copolymer synthesis is a key to understanding the structure and properties of copolymers. Many organic compounds could be labeled as "monomers." However, despite the diversity of potential monomers, the number of industrially developed homopolymers is rather small. There are some applications, which demand a combination of physical and mechanical properties not available in any single polymer. It might seem possible to develop new materials showing a combination of properties by blending different homopolymers. Unfortunately, the mixing together of two dissimilar homopolymers usually does not result in a single compatible material, but rather in a biphasic mixture. Properties of physical blends are strongly dependent upon the degree of compatibility of the components. A very small number of high-molecular-weight amorphous polymer pairs are thermodynamically compatible, that is, truly soluble in each other. Incompatibility is observed in diluted solutions as well. This is a direct consequence of the well-known relationship for free energy (ΔG), given in the equation: $\Delta G = \Delta H - T\Delta S$. Due to their high molecular weight, polymers have very small entropies of mixing (ΔS). Therefore, even a slightly positive enthalpy (ΔH), generated by endothermic mixing, is sufficient to produce a positive free energy, thus resulting in incompatibility.

The incompatibility of the polymer blend provides a driving force for each to aggregate in separate phases. These two-phase morphological systems are coarse dispersions in which the particles are usually (a) large, (b) inhomogeneous, and (c) characterized by poor interphase adhesion. The poor interphase adhesion results in very poor mechanical properties. The properties of such mixtures are frequently unattractive because of defects arising from weakness, opacity, etc. It is possible to overcome these shortcomings by polymerizing a mixture of the two different monomers and obtain a mixture of the functional groups along the polymer chain.

The chain polymerization process in which two (or more) monomers are simultaneously polymerized is termed a copolymerization; its product is a copolymer.

During the copolymerization process, the macromolecular chain dimension depends on three statistic processes: initiation, propagation, and termination,

which are basically the same as those involved in polymerization processes. Different molecular chains will have different numbers of monomer units. A copolymer is a mixture of macromolecules having different lengths. Thus, a copolymer is always associated with an "average" molecular weight and with a molecular weight distribution. In the copolymer case, a new source of diversity is added to the molecular weight distribution: the chemical composition distribution. In terms of chemical composition distribution, two different aspects are recognized: on the one hand, there is a chemical distribution within a macromolecule (random or statistical, alternating, block copolymers), and on the other hand a difference in chemical composition between two molecules (e.g., molecules obtained at different conversions). This complexity can be further enhanced by changing the number of comonomers involved in copolymerization: two comonomers (binary copolymers), three comonomers (ternary copolymers), and so on. Very complex copolymer compositions are difficult to study but involve an immense potential for a large number of applications.

Almost all monomers with a carbon–carbon double bond undergo free-radical copolymerization. Ionic polymerization can be developed only on a limited number of monomers: monomers with electron-releasing substituents undergo cationic polymerization. Anionic polymerization occurs when electron-withdrawing groups are present.

Binary Copolymerization

Copolymerization reactions have been under investigation ever since the systematic study of polymer science began in the 1930s. Copolymers could not necessarily be expected to have the same composition as the mole ratio of the monomers in the feed mixture, but this was more likely to drift as the reaction progressed (conversion effect on copolymer composition). In binary copolymerizations, four chain-growth reactions [1]–[4] take place instead of a single chain growth in homopolymerization reactions. The rate of addition of a monomer to a radical (k_{ii} and k_{ij}) is assumed to be independent of the size and nature of the radical chain and is influenced only by the nature of the radical end group:

$$-M_1\bullet + M_1 \overset{k_{11}}{\rightarrow} -M_1\bullet \qquad [1]$$

$$-M_1\bullet + M_2 \overset{k_{12}}{\rightarrow} -M_2\bullet \qquad [2]$$

$$-M_2\bullet + M_2 \overset{k_{22}}{\rightarrow} -M_2\bullet \qquad [3]$$

$$-M_2\bullet + M_1 \overset{k_{21}}{\rightarrow} -M_1\bullet \qquad [4]$$

The assumption of a steady-state condition – $k_{ij}[M_i\bullet][M_j] = k_{ji}[M_j\bullet][M_i]$ – for the elementary reaction leads to an equation for instantaneous (m_i) copolymer composition

$$\frac{dM_1}{dM_2} = \frac{m_1}{m_2} = \frac{M_1}{M_2}\frac{r_1M_1 + M_2}{M_1 + r_2M_2} \qquad [5]$$

The parameters r_1 and r_2 are the monomer reactivity ratios and provide a measure of the preference a free radical has for reacting with its own monomer rather than with the comonomer in the mixture: $r_1 = k_{11}/k_{12}$ and $r_2 = k_{22}/k_{21}$. The reactivity ratios express the monomer relative reactivity with a copolymerization process. They are always positive. Any negative value is a miscalculation. Equation [5] shows that the copolymer composition depends on the comonomers (feed) composition and on the reactivity ratios. If $r_i < 1.0$, then the cross-propagation reactions – [2] and [4] – are faster than the homopropagation reactions – [1] and, respectively, [3]. Large values of r_i will then be indicative of a tendency to form long homopolymer sequences, whereas small values of r_i imply rapid cross-propagation reactions and represent a tendency towards alternation.

Styrene is the most used comonomer in copolymerization studies. An empirical scheme (Q–e scheme, where Q is the resonance stabilization parameter and e is the electronegativity parameter) was developed in order to characterize the monomer capabilities to be involved in a copolymerization process. In that empirical scheme, styrene was chosen as the standard ("reference monomer"). **Table 1** shows the copolymer composition (obtained at a very low conversion, $<1.0\%$) calculated by using eqn [5], for styrene copolymerization with different comonomers. For a given styrene concentration in the feed, the resultant copolymer will have different compositions: 24.6% styrene in the copolymer with *p*-methylstyrene, 35% styrene in the copolymer with vinyl methyl ketone, 50% in the copolymer with maleic anhydride, and 95% in the copolymer with vinyl acetate. When $r_1 \sim r_2 \sim 1$ (styrene/*p*-methyl styrene copolymerization), the copolymer composition is almost identical to the feed composition for the entire range of the feed composition ("ideal" copolymerization). That is a rare case. When $r_1 \ll 1.0$ and $r_2 \ll 1.0$ (styrene/maleic anhydride copolymerization), the comonomer ratio is ~ 1.0 (alternating copolymer) regardless of the feed composition. Due to a much more reactive styrene monomer ($r_1 \gg r_2$), the copolymer with vinyl acetate is very rich in styrene even at low styrene concentrations in feed.

Within a copolymer macromolecule, the probability that a monomer M_1 will add onto a radical

chain terminating in its own species, thereby forming an $-M_1-M_1\bullet$ sequence is

$$p_{11} = \frac{r_1}{r_1 + (M_2/M_1)}$$

In a similar way,

$$p_{22} = \frac{r_2 M_2}{r_2 M_2 + M_1}, \quad p_{12} = \frac{M_2}{r_1 M_1 + M_2}$$
$$p_{21} = \frac{M_1}{r_2 M_2 + M_1}$$

(also $P_{12} = 1 - P_{11}$ and $P_{21} = 1 - P_{22}$). The copolymer composition shows a statistical distribution of monomer units. Therefore, the sequence length reciprocals of conditional probabilities could be expressed as a number-average: $\bar{N}_1 = r_1(M_{10}/M_{20}) + 1$ and $\bar{N}_2 = r_2(M_{20}/M_{10}) + 1$. For alternating copolymers, $r_1 \ll 1.0$ and $r_2 \ll 1.0$ and $\bar{N}_1 = 1$ and $\bar{N}_2 = 1$.

This type of copolymer characterization is valid at very low conversion values only. Due to the differences in comonomer reactivity ($r_1 \neq r_2$), the unreacted comonomer composition (instantaneous "feed" composition, M_i and M_j) and the instantaneous copolymer composition ($m_i = (M_{i0} - M_i)/(1 - M_i - M_j)$) may drift with conversion (C) during batch copolymerization (eqn [6]).

$$C = 1 - \left(\frac{M_1}{M_{10}}\right)^{r_2/(1-r_2)} \left(\frac{M_2}{M_{20}}\right)^{r_1/(1-r_1)} \times \left(\frac{M_{10} - \frac{1-r_2}{2-r_1-r_2}}{M_1 - \frac{1-r_2}{2-r_1-r_2}}\right)^{(1-r_1r_2)/(1-r_1)(1-r_2)} \quad [6]$$

For styrene-vinyl-acetate copolymerization, **Figure 1** shows the unreacted styrene concentration (calculated with eqn [6]), which decreases (from the initial value of 0.25 mole fraction) as the conversion increases. Consequently, a greater amount of the other comonomer (vinyl acetate) is incorporated in the instantaneous copolymer. At ~30% conversion, there is no styrene left and a vinyl acetate homopolymerization starts.

The final material is a blend of a styrene–vinyl acetate copolymer having a styrene content ranging from 0.95 to 0 mole fraction. About 70% of that material is polyvinylacetate. That can result in less valuable heterogeneous products. By controlled feeding of the monomers into the polymerization reactor, it is possible to regulate the sequence distribution in random copolymers and obtain homogeneous copolymers up to high conversions.

Ternary Copolymers

The relative ratio between two comonomers is different in ternary copolymerization from that in the corresponding binary copolymerization. For styrene-*p*-methyl styrene copolymerization (almost an "ideal" copolymerization case – see **Table 1**), the presence of a third comonomer (acrylonitrile, see **Figure 2**) will change the ratio between styrene and *p*-methyl styrene in the copolymer. This change is more important at higher acrylonitrile concentrations and higher conversions.

Block Copolymers

Block copolymers are macromolecules of two or more chemically different chains joined together by covalent bonds. A variety of molecular architectures are

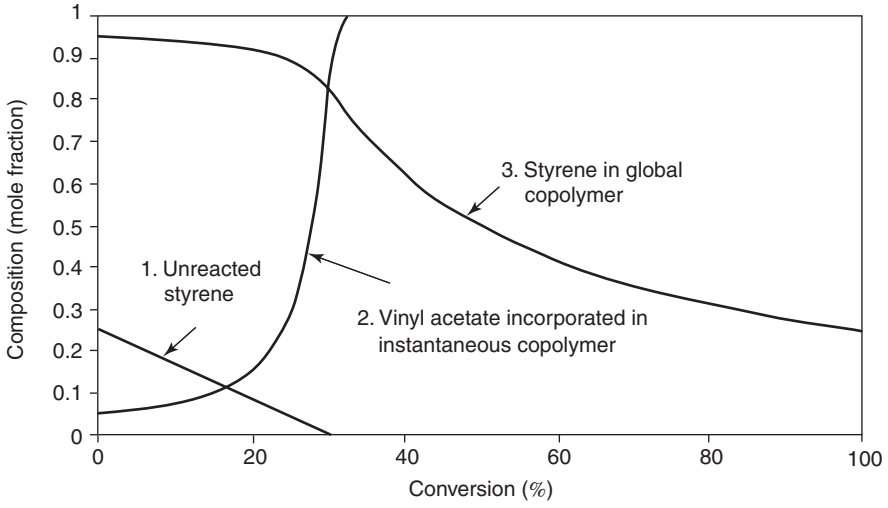


Figure 1 Styrene copolymerization with vinyl acetate: copolymer composition (instantaneous and overall composition) and the unreacted comonomer composition as a function of conversion.

Table 1 Styrene copolymerization

Styrene in feed (mole fraction)	*Styrene (M_1) in copolymer (mole fraction) for different copolymerizations*					
	M_2: p-methyl styrene		*M_2: maleic anhydride*		*M_2: vinyl acetate*	
	$r_1 = 0.89$	$r_2 = 0.99$	$r_2 = 0.05$	$r_1 = 0.005$	$r_2 = 56.0$	$r_2 = 0.01$
0.05	0.050		0.478		0.768	
0.25	0.246		0.500		0.950	
0.50	0.487		0.511		0.983	
0.75	0.735		0.536		0.994	

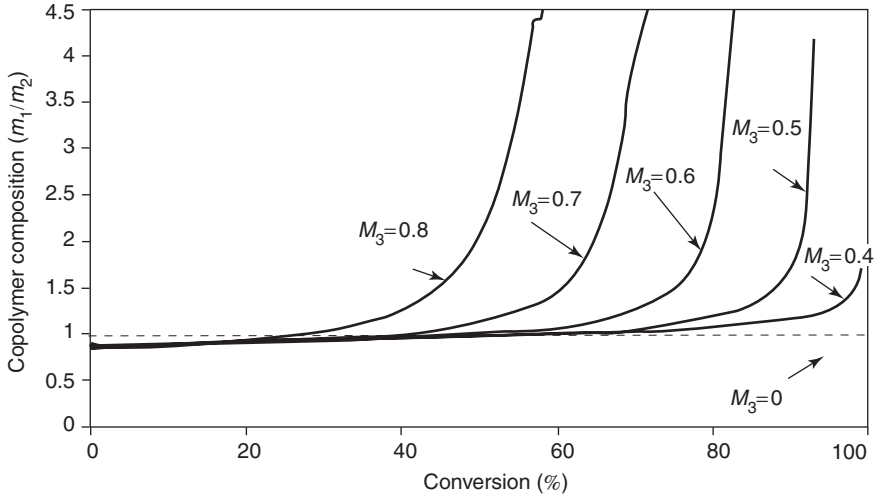


Figure 2 Ternary copolymerization of styrene (M_1) with *p*-methyl styrene (M_2) and acrylonitrile (M_3): the change in copolymer composition (styrene/*p*-methyl styrene ratio) with conversion, for different initial concentrations of acrylonitrile.

possible, for example, AB diblock, ABA triblock, and starblock copolymers. A block may have just one or a maximum of two covalent bonds with other blocks. The bonds are located at the ends of the blocks. There is an extreme case of statistical copolymers obtained through free-radical copolymerization ($r_1 \gg 1.0$ and $r_2 \ll 1.0$) when long sequences of $-M_1-$ units (blocks) are obtained. After the M_1 block, only one unit of M_2 can be inserted. For styrene (M_1) copolymerization with vinyl acetate ($r_1 = 56$ and $r_2 = 0.01$), and a molar ratio of monomers in feed of 1:1, the styrene block length will be of 57 units ($\bar{N}_{St} = 57$ and $\bar{N}_{VAc} = 1$): –[–A–]$_{57}$–[–B–]–[–A–]$_{57}$–[–B–]–. This type of copolymer is not yet a block copolymer.

Block copolymers may be synthesized by a special technique, which represents somehow a limitation in the final block copolymer structures. Special monomer structures and special catalysts are required in order to get a desired block copolymer structure. In ionic polymerization, monomers show a different reactivity as compared to the free-radical mechanism. In anionic copolymerization with butadiene, styrene has a reactivity ($0.02 < r_{St} < 0.04$) of ~ 10 times lower than in free-radical copolymerization ($0.4 < r_{St} < 0.6$). With the same type of process, butadiene is 10 times more reactive in anionic copolymerization ($11 < r_{Bu} < 14$) than in free-radical copolymerization with styrene ($1.3 < r_{Bu} < 1.5$). When styrene and butadiene (1:1 molar ratio) are charged together with the lithium alkyl catalyst solution into a batch reactor, the butadiene molecules polymerize first and with almost total exclusion of any styrene present: a long block of butadiene units is developed ($\bar{N}_{Bu} = 15$) before a styrene unit is incorporated. Only when all the butadiene monomers are consumed does styrene enter the polymer chain.

Living anionic copolymerization provides the maximum degree of control for the synthesis of copolymers with predictable well-defined structures. In order to retain the growing free-radical activity and for a narrow-molecular-weight distribution of polymers, living polymerization requires the absence of chain transfer and termination reaction. Block copolymers can be prepared by sequential monomer addition. Living cationic polymerizations are much more difficult to achieve. It is possible

to obtain block copolymers by a free-radical mechanism by reversible addition–fragmentation transfer (RAFT).

Much of the interest in block copolymers has arisen because of their remarkable microphase morphology. Block copolymers show properties quite different from either a random copolymer or a macroscopic blend of the corresponding two homopolymers. The component blocks are usually thermodynamically incompatible. That is why one of the characteristic features of block copolymers is the tendency for microphase separation to occur in the solid. This separation imparts many of the interesting and desirable properties displayed by block copolymers. The morphology of such multiphase systems influences their behavior and can be controlled by altering the ratio of block lengths, the molecular weight distribution, and the nature of the comonomers. The phase separation occurs in a block copolymer when the block molecular weights are higher than those required to produce phase separation in a blend of the corresponding homopolymers. The increasing tendency for segregation is stronger as the temperature decreases. In the majority of cases, the two types of polymer chains are not thermodynamically compatible but they are constrained to coexist at a molecular level by the bonds between them. In an AB poly(styrene-*b*-butadiene), up to a polystyrene content of $\sim 65\%$ the polystyrene is in the dispersed phase. At higher polystyrene contents, phase inversion takes place. In the presence of a solvent, as individual blocks are mutually incompatible, there is a ternary system in which the component blocks may interact with the solvent to different degrees. When a solvent is "good" for one block but "poor" for the other, anomalies in intrinsic viscosity–molecular-weight relations were observed. These anomalies are due to the tendency of the blocks with poor solvation to aggregate and form a micellar structure. The phase separation in the solution looks like the solid-state microstructure.

Graft Copolymers

Graft copolymers are macromolecules of two or more different chemical chains in which a chain (named backbone) has multiple branches formed from macromolecular chains with a chemical composition different from that of the backbone. The covalent bonds are randomly distributed along the backbone chain and are located at the end of the grafted branch. Unlike block copolymers, the graft copolymer can be synthesized more easily. Any macromolecular compound and almost any comonomer can be involved in a grafting process. Acrylonitrile–styrene–butadiene (ABS) grafted copolymers are synthesized by styrene–acrylonitrile copolymerization in the presence of the polybutadiene homopolymer. The grafted branches are formed through a chain transfer to polybutadiene, or by copolymerization of the vinyl double bond from polybutadiene. Macromonomers copolymerization is an elegant way to synthesize a graft copolymer. Polyethyleneoxide acrylate can be copolymerized with another acrylic ester, and a polyacrylate grafted with polyethylene-oxide will result.

The chemical structure of cellulose (poor solubility in common solvents, lack of thermoplasticity, etc.) can be modified by preparing a graft copolymer of cellulose, a process in which attempts have been made to combine synthetic polymers with natural polymers to produce materials with the best properties of both. Cellulose molecules are modified through the creation of branches of synthetic polymer that confer certain desirable properties to cellulose without destroying its intrinsic properties.

Copolymers in Blends

Many copolymers obtained at high conversion are blends. Their behavior in a complex mixture is of great interest. The compatible blends are characterized by a single-phase morphology. In the case of a homogeneous blend of two homopolymers, the glass transition temperature (T_g) will follow the simple additive rule (Fox equation): $1/T_g = (W_1/T_{g1}) + (W_2/T_{g2})$. An inhomogeneous blend will display at least two major glass transitions. A free-radical copolymer with a narrow chemical composition distribution has only one glass transition temperature. It has been shown that the T_g of a copolymer does not always obey the Fox equation. It is better described by the equation below, which takes into account the two consecutive monomer units (DYAD) distribution (AA, BB, AB, BA linkages on the chain backbone):

$$\frac{M_{cop}}{T_g} = \frac{[AA]M_a}{T_{gaa}} + \frac{[BB]M_b}{T_{gbb}} + \frac{[AB]M_a + [BA]M_b}{T_{gab}}$$

Most of the copolymers are actually blends of various copolymers with different chemical compositions. Styrene–methyl methacrylate copolymers with only a difference of 5% in styrene content become incompatible. They are blends and the "blends" issue is open to many choices. At similar compositions, both block copolymers and graft copolymers should be subjected to similar thermodynamic driving forces for phase separation, phase-size constraint, and compatibility with homopolymer mixture.

Block and graft copolymers are amphiphilic molecules because they may have affinity for two different types of environments. In a dispersed system, the graft and/or block copolymers self-organize at the interface and modify interfacial properties to a great extent, thus enhancing phase compatibility. In order to do so, their blocks or branches should be different enough and with a real affinity for those phases. Block and graft copolymers act as oil in oil emulsifiers. They may be absorbed at various surfaces: liquid/gas, liquid/liquid, solid/liquid, etc. The surface-active properties of block and graft copolymers have made them useful as dispersants, emulsifiers, foam stabilizers, and wetting agents. A large class of compatibilizers has been developed. Very attractive blends have been envisaged: inorganic fillers dispersed in rubber of plastic materials, wood flower dispersed in polyolefines, and glass fibers as reinforcement for tires. ABS copolymers consist of a dispersion of polybutadiene rubber particles in a matrix of styrene–acrylonitrile copolymer. Some of the styrene–acrylonitrile–copolymer (SAN) copolymer molecules are grafted to the rubber during polymerization. The grafted SAN acts as a dispersing agent.

See also: Polymer Structures; Polymers, History of.

PACS: 61.41.+e; 82.35 Gh; 82.35 Jk

Further Reading

Alexandridis P and Lindman B (2000) *Amphiphilic Block Copolymers*. Amsterdam: Elsevier.

Cowie JMG (1985) *Alternating Copolymers*. New York, NY: Plenum.

Folkes MJ (1985) *Processing, Structure and Properties of Block Copolymers*. London: Elsevier.

Goodman I (1982) *Developments in Block Copolymers-1*. London: Applied Science.

Goodman I (1985) *Developments in Block Copolymers-2*. New York, NY: Elsevier Applied Science.

Guillot J, Guyot A, and Pichot C (1995) Radical copolymers in dispersed media. In: *Macromolecular Symposia*, 92, p.354. Zug, Switzerland: Huething & Wept Verlag.

Hagiopol C (1999) *Copolymerization: Toward a Systematic Approach*. New York, NY: Kluwer Academic/Plenum.

Hebeish A and Guthrie JT (1981) *The Chemistry and Technology of Cellulosic Copolymers*. Berlin: Springer.

Krentsel BA, Kissin YV, Kleiner VI, and Stotskaya LLZ (1997) *Polymer and Copolymers of Higher α-Olefins*. C Hanser Verlag.

Molau G (1971) *Colloidal and Morphological Behavior of Block and Graft Copolymers*. New York, NY: Plenum.

Odian G (2004) *Principles of Polymerization*. 4th edn. New York, NY: Wiley-Interscience.

Platzer NAJ (1975) *Copolymers, Blends, and Composites*, Advance in chemistry series – 142. Washington, DC: American Chemical Society.

Core Photoemission

P Padova and P Perfetti, CNR-ISM, Rome, Italy

Introduction

The photoemission technique is based on the extraction of an electron from a system (atoms, molecules, and solids) by a photon of energy $h\nu$. In core photoemission spectroscopy (CPS), the photon energy is sufficient to excite the electron from the core level into the vacuum. The first applications of CPS, also known as electron spectroscopy for chemical analysis (ESCA), used $Mg_{k\alpha}$(1254 eV) and $Al_{k\alpha}$(1486.6 eV) photons emitted by conventional X-ray sources. ESCA made it possible to study core levels and opened up an extraordinary as well as prolific field of research in physical chemistry. This kind of spectroscopy benefited from the advent of synchrotron radiation (SR), which is a brilliant, continuous radiation ranging from the infrared to X-rays. The electrons circulating at relativistic energy in a storage ring emit SR. The photon energy tunability of SR and the high-energy resolving power of the monochromators ($E/\Delta E \approx 10\,000$, from vacuum ultraviolet to X-rays) of third-generation SR machines have increased the accuracy of spectroscopic data. Thus, it is now possible to explore with more details the physical phenomena related to clean reconstructed surfaces and to the formation of adsorbate/substrate systems and band discontinuity at interfaces. Since the escape depth of electrons from a solid can be varied by an appropriate choice of the photon energy, it is also easier to separate bulk from surface effects and thereby improve the agreement between experiments and theory.

Here, the basic principles of core-level photoelectron spectroscopy are presented first and the origin of core-level lineshape broadening is then reviewed with the help of some experimental examples. Later, the concept of core-level binding energy is analyzed, taking into account all the possible mechanisms responsible for core-level chemical shifts, and some *ad hoc* experimental results are

illustrated. Finally, photoelectron diffraction for the determination of the surface/interface geometry and Schottky barrier height are examined, with the use of a few examples of CPS applications.

Basic Principles

In photoemission, a system containing N electrons in the initial state is left, after excitation, in a final state with $N-1$ electrons and one hole. The energy of the outgoing electron is influenced by the interaction between the photoelectron itself and the ionized system that is left behind. If the interaction is negligible (the so-called sudden approximation) and the photoelectron kinetic energy (E_k) is high, one can derive E_k (see **Figure 1a**), using Einstein's argument (proposed in 1905) to explain the photoelectric effect

$$E_k = h\nu - |E_B| - \Phi$$

where $h\nu$ is the energy of the incoming photon, Φ the work function of the emitting material, and E_B the one-electron binding energy (BE) of the core-level.

In the one-electron approximation, it is assumed that the $(N-1)$ electron orbitals will be the same in the final state as they were in the initial state. This is the so-called "frozen-orbital approximation." **Figure 1a** is an energy diagram of the photoexcitation of electrons in a solid; the vacuum energy E_{Vac} represents the zero of energy. In a solid, the core-level binding energy usually refers to the Fermi level, E_F. By measuring the kinetic energies of the photoelectrons emitted in the vacuum, it is possible to probe the occupied electron states of the system. Since every element has its own characteristic core-level spectrum, irrespective of the material it belongs to, CPS analysis is an extraordinary tool for identifying different atomic species in a chemical compound. An example of CPS analysis is the photoelectron spectrum of the indium arsenide clean surface shown in **Figure 1b**. In addition to the valence band emission at high kinetic energy, and secondary contribution peaked at low kinetic energy, two doublets are clearly visible. They have the binding energies characteristic of indium (In) 4*d* shells and arsenic (As) 3*d* shells, with angular momentum 5/2 and 3/2, respectively. Each level is separated by the spin–orbit interaction, which gives rise to the formation of a doublet. The spectrum in **Figure 1b** distinctly shows the In and As atomic species.

When there is strong interaction between the electrons and the potential of the photohole, the valence electrons get rearranged, with subsequent excitation of the rest of the system. This process costs energy and, as a result, the excited photoelectrons will

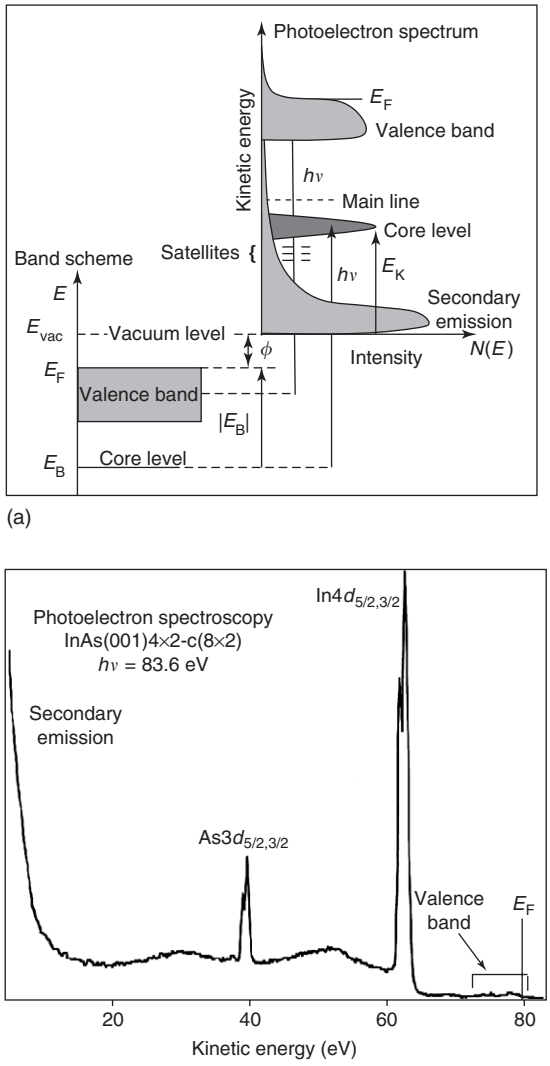


Figure 1 (a) Energy diagram of photoexcitation of electrons in a solid. Secondary emission, core-level, and valence band form the photoemission spectrum for no electron–electron interaction. Dotted lines: core-level main and satellite lines for electron–electron interaction (b) Photoelectron spectrum of the InAs clean surface. Valence band emission at high kinetic energy (in the vicinity of the Fermi level), secondary contribution, and two doublets of In4$d_{5/2,3/2}$ and As3$d_{5/2,3/2}$ at low kinetic energy.

appear at a lower binding energy (main line) than in the case of no interaction; the main line will be accompanied by extra lines, called "satellite lines," at higher binding energy (**Figure 1a**).

In a metal, the relaxation process can be due to electron–hole pair excitations at the Fermi level giving rise to an asymmetric core lineshape, or to collective electron oscillations related to the positive lattice ions (plasmons). Note that this phenomenon is intrinsic to the photoemission process and should be distinguished from extrinsic interaction of the photoexcited electrons with the other electrons of the system, along their path to the vacuum. Extrinsic

interactions can be simply described as scattering events that give rise to a smooth background of secondary electrons superimposed on the intrinsic primary photoelectron contribution and on a huge structure at very low kinetic energy (see **Figure 1b**).

Here, the expression of the photocurrent detected in a CPS experiment, with some approximations, is given

$$I \propto \sum_{f,i,k} |\langle \phi_{f,Ek}|\boldsymbol{r}|\phi_{i,k}\rangle|^2 \sum_s |c_s|^2 \cdot \delta(E_{f,k} + E_s(N-1) - E_0(N) - h\nu)$$

It is assumed that the final state, with $N-1$ electrons, has s excited states with energy $E_s(N-1)$; E_0 is the ground-state energy of the N electrons and $E_{f,k}$ is the kinetic energy of the photoelectron promoted from the initial state $\phi_{i,k}$ to the final state $\phi_{f,Ek}$. $|c_s|^2$ is the probability that the removal of one electron from the ground-state orbital ϕ_k is accompanied by the excitation of s states in the $N-1$ electron system. The energy conservation derived from the delta function in the relation tells one that, in addition to the main peak, a certain number of satellite structures due to the excitation of the system are left after the creation of a hole.

In a CPS experiment, the photoelectron intensity measured by the detector is the result of three separate processes: (1) the optical excitation of an electron, viewed as a purely bulk transition between the initial and final Bloch states of the crystal, (2) the propagation of the excited electron toward the surface, and (3) the escape of the photoelectron from the solid into the vacuum. This representation is known as the "three-step model" (**Figure 2**) and is most commonly used for interpreting CPS spectra because it is simple and the photoemission results are directly comprehensible.

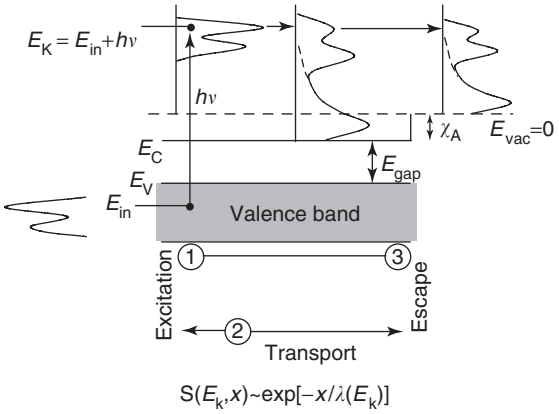


Figure 2 Three-step model. Step one: primary electron excitation. Step two: electron transport to the surface. Scattering processes reduce the primary-emission intensity and produce the secondary electron contribution, which is superimposed on the primary. Step three: transmission into the vacuum; here, the electron distribution is modified by the potential step at the surface. All energies are referred to E_{vac} which is set to zero.

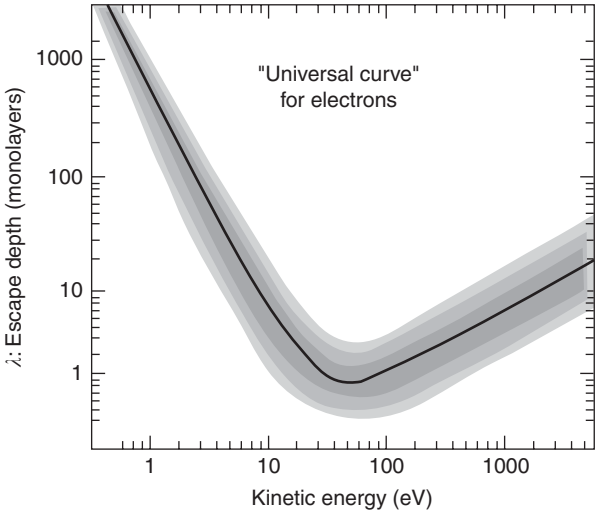


Figure 3 "Universal curve" for the kinetic energy dependence of the electron escape depth in a solid (λ).

Escape Depth and Surface Sensitivity

It has been seen that the photoelectrons lose energy while they travel to the surface and that many of them are unable to escape into the vacuum. The probability $S(E_k,x)$ that an electron with kinetic energy E_k, photoexcited at a distance x from the surface, will escape is given, in a first approximation, by

$$S(E_k, x) \propto \exp[-x/\lambda(E_k)]$$

where $\lambda(E_k)$ is the escape depth. **Figure 3** shows the so-called universal curve for the excited electrons escape-depth dependence on the kinetic energy. This curve, which is an average of measurements made on several different materials, establishes the dependence of λ on the kinetic energy of the emitted electrons. One can see from the figure that λ varies from tens of angstroms at low kinetic energy to a minimum of ~ 3 Å for $E_k = 50$ eV and then again up to ≈ 40–50 Å for $E = 1000$ eV. An interesting aspect of this behavior is that the photon energy can be chosen such that the surface or the bulk contribution shows up clearly in the CPS spectrum. It is to be noted that these contributions have different core-level binding energies because the surface and bulk atoms are not equivalent owing to the different chemical environments. However, this concept is explained in detail in the section on Si. For most of the elements, surface sensitivity is obtained for electron kinetic energies ranging between 20 and 50 eV. For Si, this minimum

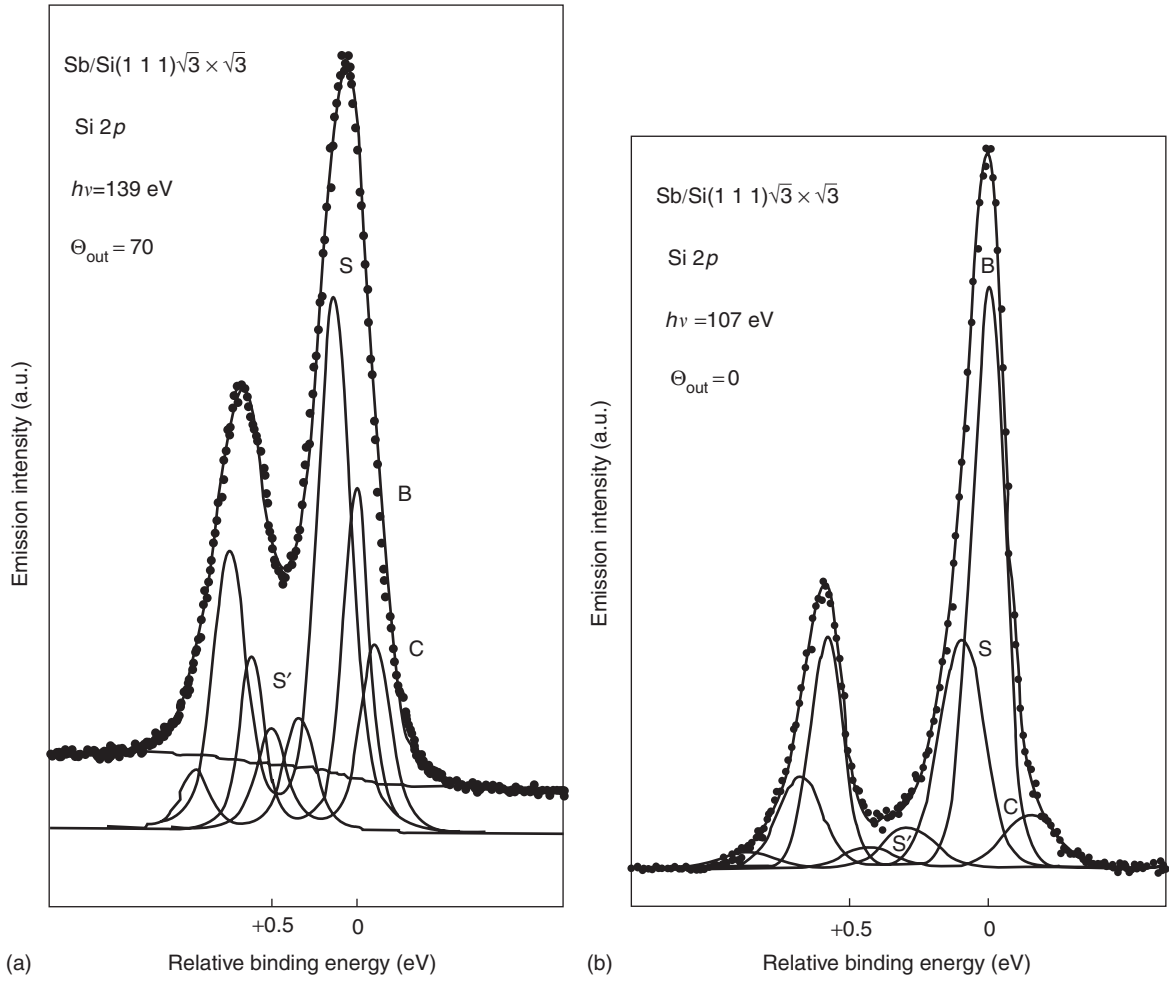


Figure 4 (a) Surface-sensitive ($\Theta = 70^\circ$, photon energy $h\nu = 139$ eV) and (b) bulk-sensitive ($\Theta = 0^\circ$, photon energy $h\nu = 107$ eV) Si 2*p* core levels on Sb/Si(1 1 1)$\sqrt{3} \times \sqrt{3}$. Also shown is bulk- and surface-component convolution of the spectra. (Reproduced from Cricenti, *et al.* (2000) Sb-terminated Si(1 1 0), Si(1 0 0), and Si(1 1 1) surfaces studied with high resolution core-level spectroscopy. *Applied Surface Science* 162–163: 380–383, with permission from Elsevier.)

corresponds to two layers (the Si lattice parameter is $a = 5.43$ Å). This small escape depth makes it possible to detect below one monolayer (ML) of chemically altered atoms at the surface or at an interface. (One ML is defined as the number of deposited atoms that equals the number of surface atoms of the substrate.) If it is impossible to exploit a tunable photon energy source, the surface effect can be enhanced by changing either the light incidence angle or the electron emission angle, or by taking advantage of the selection rules that support surface excitation processes. **Figure 4** shows the surface- and bulk-sensitive Si 2*p* core levels taken on the Sb/Si(1 1 1)$\sqrt{3} \times \sqrt{3}$ interface using SR. The spectra were collected at different surface sensitivities with emission angles $\Theta = 70^\circ$ ($\Theta = 0^\circ$) and photon energies of $h\nu$ 139 eV (107 eV). Convolution of the spectra will give rise to different components (for the convolution procedure, see the following section). To the left of the figure, the surface S contribution, due to interaction between Sb and Si atoms, dominates the spectrum because the escape depth parameter is minimum. On the other hand, the spectrum taken at lower photon energy and $\Theta = 0^\circ$ shows a dominant bulk contribution because of the increased escape depth λ.

Core-Level Lineshapes

In a core-level spectrum, three main contributions affect the lineshape:

1. *Core-hole lifetime*. The hole does not live infinitely. It can recombine with one of the outer electrons and the recombination energy can be released, with the emission of a photon (fluorescence), or transferred to an electron residing in another shell. The electron is then excited into the vacuum, and this is known as the Auger process

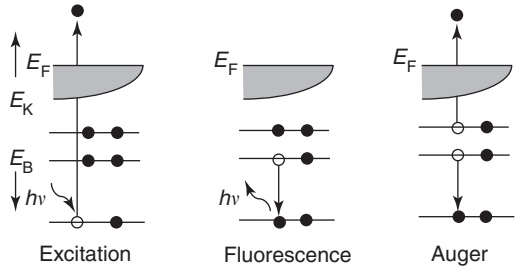


Figure 5 Schematic diagram of two possible electron–hole recombination processes after photoexcitation. In the fluorescence process, recombination arises with the emission of a photon. In the Auger process the recombination energy is transferred to another electron, with the formation of two holes.

(see **Figure 5**). The Heisenberg uncertainty principle states that the shorter the recombination time, the larger the uncertainty in one's knowledge of the energy of the emitted core electron, and hence one gets a wider lineshape. The intensity of the fluorescence and Auger processes depends on the atomic number Z of the atomic species considered, and the fluorescence channel is dominant for a high Z value. Both processes are responsible for the so-called intrinsic core-level lifetime.

2. *Phonon broadening*. This is a final-state effect and is due to the excitation of very low energy phonons (quantum lattice vibrations).
3. *Instrumental resolution*. This mainly depends on the monochromator and electron-energy analyzer and is usually calculated by a convolution of the two contributions.

The three sources of core-level broadening are generally well described by a Lorentzian and two Gaussian lineshapes. The Lorentzian function is characterized by a very sharp, narrow peak with most of its intensity located in the tails and is given by

$$F_L(E) \propto \Gamma_L^2/[(E - E_0)^2 + \Gamma_L^2]$$

where E_0 is the peak energy position and Γ_L is the full width half maximum (FWHM). The broadened lineshape, termed Gaussian, tends to have much smaller wings and greater intensity under the center peak than the Lorentzian lineshape and is given by

$$F_G \propto \exp(E - E_0)^2/\Gamma_G^2$$

where E_0 is the peak energy position and Γ_G is the FWHM. A best fit to the experimental data is usually performed with a theoretical Voigt function, which is a convolution between a Lorentzian and a Gaussian function. In a CPS spectrum, this analysis applies both to the surface and to the bulk contributions. The phonon broadening is easily derived from the Gaussian width once the experimental resolution is known.

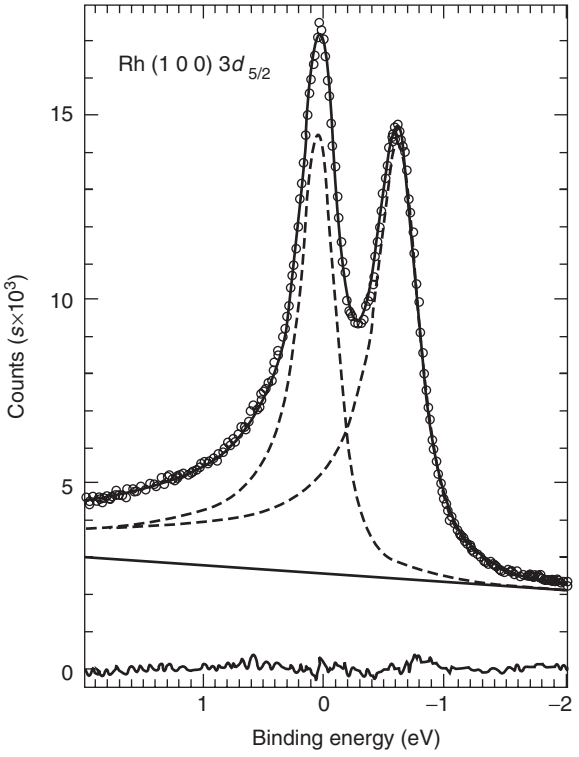


Figure 6 Core-level spectrum of Rh3$d_{5/2}$, showing the bulk peak and the surface chemical-shifted component (dashed lines). The data were taken on the clean (1 0 0) surface of Rh at a photon energy of 390 eV. The Doniach–Sunjic convolution (full line) closely reproduces the experimental result (open dots), particularly the asymmetric part due to the electron–hole pair excitation caused by core-hole screening. (Reproduced from Zacchigna, *et al.* (1996) Photoemission from atomic and molecular adsorbates on Rh(1 0 0). *Surface Science* 347: 53–62, with permission from Elsevier.)

Other possible sources of broadening are imperfections in the crystal, such as defects, surface reconstructions, and inhomogeneous strains, which are generally considered to affect the Gaussian width. Since these contributions are difficult to quantify, it is practically impossible to determine the intrinsic core-level lifetime. One of the goals of experimentalists is to reduce sample imperfections so as to obtain from the fitting procedure, the best upper limit for the core-level intrinsic lifetime. In metals, the intrinsic lifetime of the core hole is affected by the formation of electron–hole pairs, whose excitation energy is reflected in the low kinetic-energy part of the detected core-level spectrum. This phenomenon gives rise to an asymmetric lineshape. In this case a Doniach–Sunjic function is used for the fitting procedure, as reported in **Figure 6** for the asymmetric lineshape of the $3d_{5/2}$ peak of rhodium (Rh).

The Case of Si

Figure 7 shows an example of analysis of high-resolution CPS of Si 2p collected on the Ge/Sb/Si(0 0 1)

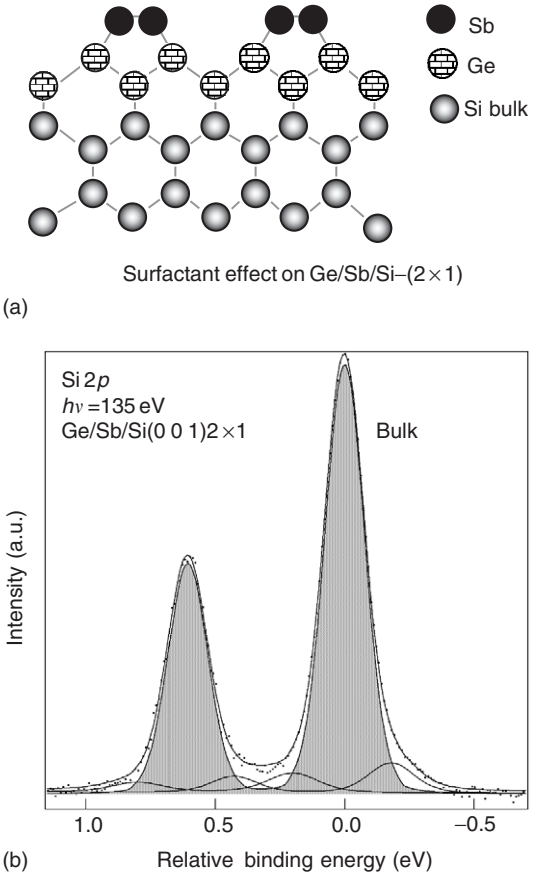


Figure 7 Side view of ball-and-stick model of (a) Sb surfactant effect on Ge/Sb/Si(0 0 1)2 × 1, and (b) Si 2*p* spectrum taken on hydrogenated Ge/Sb/Si(0 0 1) system. The spectrum shows the convolution of the Si 2*p* core level in surface (blank doublets) and bulk contributions (shaded doublet). The "bulk" component dominates the spectrum and demonstrates a nearly unreconstructed (1 × 1) configuration for the last layers of the Si in the Ge/Sb/Si structure. (Reproduced with permission from De Padova P, Larciprete R, Quaresima C, Ottaviani C, Ressel B, and Perfetti P (1998) Identification of the Si 2p surface core level shifts on the Sb/Si(0 0 1) – (2 × 1) interface. *Physical Review Letters* 81: 2320–2323; © American Physical Society.)

system. This interface was artificially created to organize the last silicon layer in a geometrical configuration very much like the bulk structure. The system was prepared by depositing 2.5 MLs of Ge on the interface formed of 1 ML of Sb on Si(0 0 1). This is an epitaxial growth, where the process of surface-energy minimization leads to an Sb–Ge site exchange process that segregates the antimony at the top of the surface and traps the Ge atoms between the Sb adlayer and the Si substrate. Antimony acts as a surfactant, and the insertion of the epitaxial layers of Ge leaves the interfacial Si atoms arranged in an almost perfect bulk-like (1 × 1) configuration. The final result of this process is shown schematically in **Figure 7a**. The system is characterized by high homogeneity and atomic order. The bulk-like atomic arrangement of the top silicon layers can be verified by convoluting the Si 2*p* core level with the use of the spin–orbit split Voigt functions. One can clearly see from **Figure 7** that the lineshape is completely dominated by the "bulk" component, and a comparison with **Figure 10** emphasizes the effect of the different atomic arrangement of the last silicon layers; here, indeed, the Si 2*p* lineshape contains many core-level components due to surface reconstruction.

The peak fit of the Si 2*p* core level on the Ge/Sb/Si system gives a Lorentzian FWHM $\Gamma_L = 20$ meV. This is one of the lower Lorentzian width values found for silicon. Higher gamma values have been found in other high-resolution results for a clean Si (1 1 1) surface with a complicated 7 × 7 reconstruction. The gamma value is an indication of the core-hole lifetime, and the above reported value of 20 meV can be considered as an upper limit of the intrinsic lifetime for silicon.

Binding Energy, Surface Core-Level Shift, and Chemical Shift

The binding energy E_B of a core level can be defined as the difference between the total energy of the unperturbed state with N electrons in its ground state $E_0(N)$ and the energy of the state formed by the hole and the $N-1$ remaining electrons after photoexcitation, $E_f(N-1)$

$$E_B = E_f(N-1) - E_0(N)$$

The absolute value of the measured core-level binding energy is influenced by several factors that can affect either $E_0(N)$ or $E_f(N-1)$. To explain the physical origin of the variation in the binding energy, one can define ΔE_B as the difference between the measured binding energy and the binding energy of an electron in its fundamental state in an isolated atom. From the preceding equation, one obtains

$$\Delta E_B = \Delta E_0 + \Delta E_f$$

Initial-State Effects

Initial-state effects in the binding energy of core electrons (ΔE_0) originate mainly from a change in the electrostatic interaction between the core electrons and the outer electrons. The change is determined by the variation in the charge distribution on the emitting atom and by the different physical situations. Consequently, the following can contribute to the variation in ΔE_0:

1. The charge transfer, which is due to the chemical bonds formed by the same atomic species in different chemical compounds.

2. The kind of atomic-species configuration, which depends on whether an atom is isolated or arranged in a molecule or a solid. For example, in a solid the atoms can exist either in a surface or in a bulk configuration; accordingly, on the emitting atom, the valence-electron distribution changes and, as a consequence, also the measured electron binding energy.
3. The "Madelung contribution," which is due to charge transfer between the emitter and the neighbor atoms, could be important in ionic compounds.

For surface atoms, a charge transfer causes binding energy shifts, called surface core-level shifts (SCLS). Contributions (1) and (2) are only schematic, since each change in configuration is always accompanied by bond modification; this is what happens to an atom that exists in different chemical compounds.

Core photoemission spectroscopy studies do not usually deal with evaluation of the absolute binding energy of a particular core-level but with the binding energy changes between two different chemical forms of the same atom. This energy difference is known as a "chemical shift."

Chemical shifts are quite important in identifying the chemical compounds present in a system and can also be extremely helpful in determining the structural properties by, for example, the surface diffraction technique. In a chemical bond between two different atoms, there is a charge transfer from the more electropositive atom to the more electronegative one. Recall that electronegativity is a measure of the attraction of an atom for electrons. Due to a Coulomb interaction between core and valence electrons, the charge transfer results in a more or less effective screening of the core hole, which affects the photoelectron kinetic energy. Consequently, if the CPS spectra of different compounds are compared, one sees that photoelectrons emitted from atoms in a higher oxidation state display higher kinetic energy (lower binding energy) and vice versa.

The Case of the Trans-Dinitrobis (Ethylene-Diamine) Cobalt(III) Nitrate Compound

The N1*s* core-level photoemission spectrum in the $[Co(NH_2CH_2CH_2NH_2)_2(NO_2)_2]NO_3$ compound, where the chemical shift concept is applied to identify nonequivalent nitrogen atoms in different compounds, is shown in **Figure 8**. The chemical shift is roughly proportional to the charge transfer and can be used to follow the chemical reactions of each *N* atom in different molecules. The binding energy of the core-level N1s increases with increasing "oxidation" number of the nitrogen atoms. This is the case of NH_2, NO_2, and NO_3 in $[Co(NH_2CH_2CH_2NH_2)_2(NO_2)_2]NO_3$.

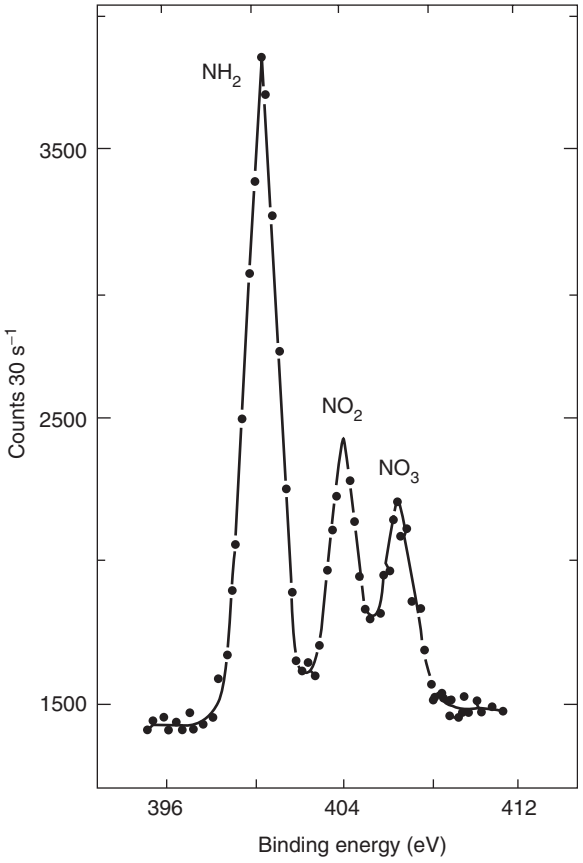


Figure 8 Chemical shift of the N1*s* core level in the $[Co(NH_2CH_2CH_2NH_2)_2]NO_3$ compound. (Reproduced with permission from Hendrickson, *et al.* (1969) *Inorganic Chemistry* 8: 2646; © American Chemical Society.)

The Case of a Clean Si(0 0 1)2 × 1 Surface

The $Si(0\,0\,1)2 \times 1$ surface can be considered as a pedagogical example to show the SCLSs. An ideal surface, considered as a truncation of the bulk lattice, for instance a covalent semiconductor of the IV group, is formed of a large number of broken bonds pointing out of the surface. These broken bonds, or dangling bonds, are energetically unstable, so the surface atoms reorganize to minimize the total energy at the surface, thus reducing the number of dangling bonds. The electronic properties of a clean reconstructed surface are different from those of an ideal surface due to the new type of bonds created. Several reconstructions can be induced on a surface sample as a function of parameters such as temperature, surface and bulk strains, crystallographic orientation, atomic species, and adsorbates. A schematic representation of the $Si(0\,0\,1)$ (1×1) and (2×1) reconstructions is shown in **Figure 9**. The

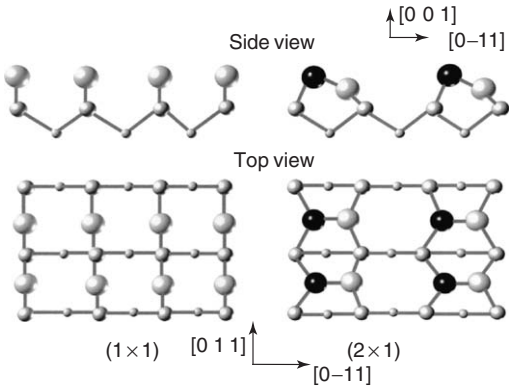


Figure 9 Ball-and-stick model of an ideally terminated (1 × 1) Si(0 0 1) and asymmetric dimer reconstructed structure Si(0 0 1)2 × 1.

(2 × 1) reconstruction is formed of buckled dimers. At liquid nitrogen temperature, the surface reconstruction changes again, forming a larger $c(4 \times 2)$ unit cell consisting of buckled dimers with up-and-down opposite orientation in the [1 1 0] and [1 – 10] directions. Various surface components associated with the reconstruction of clean Si(0 0 1) have been reported for the Si 2*p* core level, and attributed to nonequivalent dimer atoms and to the subsurface layer in contact with the dimers. **Figure 10** shows surface-sensitive, high-resolution photoemission spectroscopy of the Si 2*p* core level of clean Si(0 0 1)$c(4 \times 2)$. Many peaks are well resolved in the CPS spectrum. By convoluting the Si 2*p* spectrum with spin–orbit split Voigt functions, four components related to the $c(4 \times 2)$ reconstruction can be identified, in addition to the bulk component B. The four surface contributions *U*, *D*, *S*, and *C* are interpreted as "up" and "down" dimer atoms, second layer atoms, and subsurface layers or defects. The energy shifts with respect to the B component are $U = -498$ meV, $C = -200$ meV, $D = +60$ meV, and $S = +200$ meV. The up-and-down dimer atom contributions, at higher and lower binding energy relative to the bulk silicon B, are easily identified since, in the reconstruction process, there is a charge transfer from down to up atoms. A similar argument can be applied to the chemically shifted peak *S*. More uncertain is the interpretation of the *C* component.

Final-State Effects

Final-state effects influence the variation ΔE_B through ΔE_f. In a CPS process, the outer electrons probe a higher effective charge of the nucleus because of the creation of the hole. Thus, they relax and decrease both the total final energy and the photoelectron binding energy. ΔE_f is zero when the relaxation energy of the remaining electrons in the photoemission process

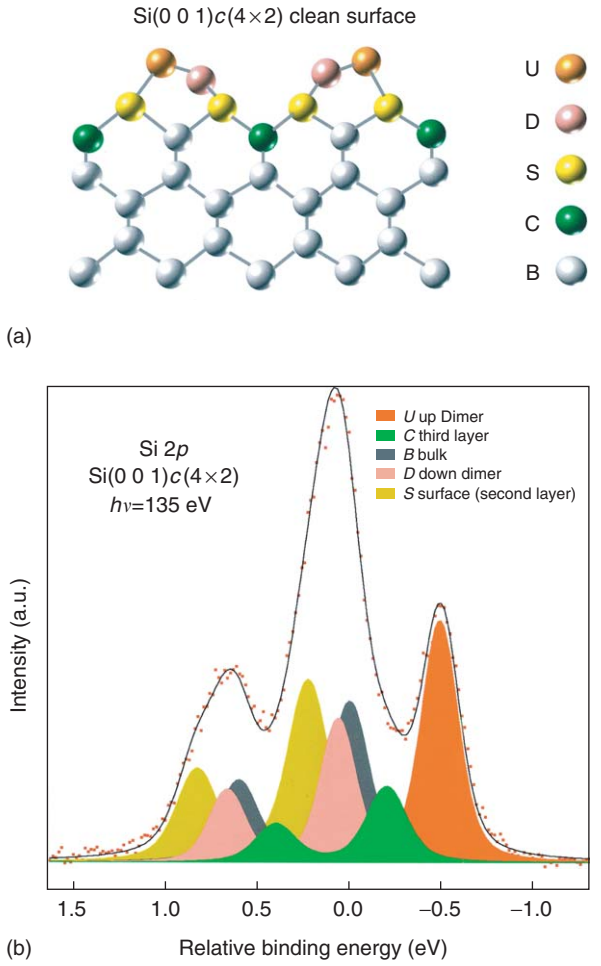


Figure 10 Surface-sensitive Si 2*p* core-level photoemission spectrum for the clean Si(0 0 1)-$c(4 \times 2)$ surface: (a) ball-and-stick model for the asymmetric dimer $c(4 \times 2)$ reconstructed structure; (b) convolution of the Si 2*p* spectrum in four surface contributions *U*, *D*, *S*, and *C*, in addition to the bulk component B. Each patterned atom corresponds to a nonequivalent Si atom. (Reproduced with permission from De Padova P, Larciprete R, Quaresima C, Ottaviani C, Ressel B, and Perfetti P (1998) Identification of the Si 2p surface core level shifts on the Sb/Si(0 0 1) – (2 × 1) interface. *Physical Review Letters* 81: 2320–2323; © American Physical Society.)

is negligible; this is called the Koopmans' approximation, and the photoelectron binding energy is that of the orbital from which it is removed. Finally, one may have two contributions to the relaxation process, one intra-atomic and the other extra-atomic.

Intra-atomic relaxation is practically independent of the neighbor atoms of the emitter (unless the emitter changes its valence electrons because of a charge transfer) and is typical of isolated atoms. It is already seen that the relaxation process often leaves the system in an excited state, which gives rise to extra "satellite lines," also called "shake-up lines," in the photoemission spectra.

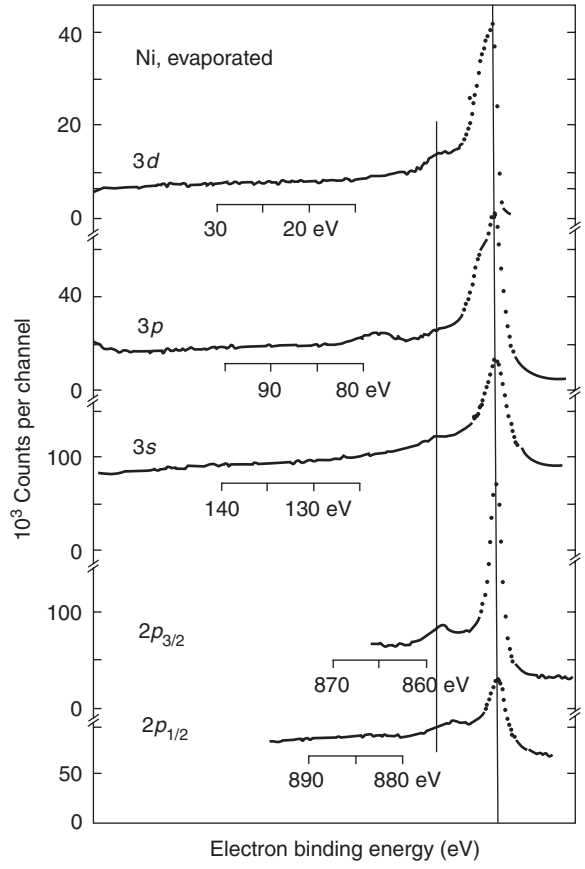


Figure 11 X-ray photoelectron spectroscopy spectra of the 3*d*, 3*p*, 3*s*, $2p_{3/2}$, and $2p_{1/2}$ core levels of Ni metal. The main lines are aligned. (Reproduced from Hüffner, *et al.* (1975) Multielectron effects in the XPS spectra on Nickel. *Physics Letters* 51A: 299–300, with permission from Elsevier.)

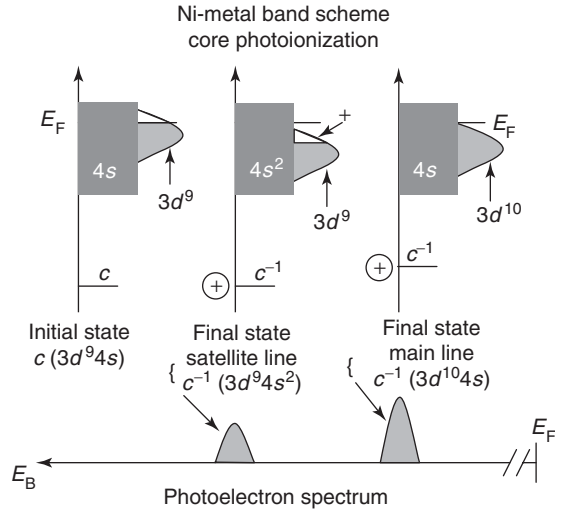


Figure 12 Schematic diagram of the Ni density of states, indicating the origin of the main and the satellite line for core ionization. C denotes a core-level; c^{-1} a core hole. The initial state is $c(3d^9 4s)$, and the two final states are $c^{-1}(3d^9 4s^2)$ (satellite) and $c^{-1}(3d^{10} 4s)$ (main line).

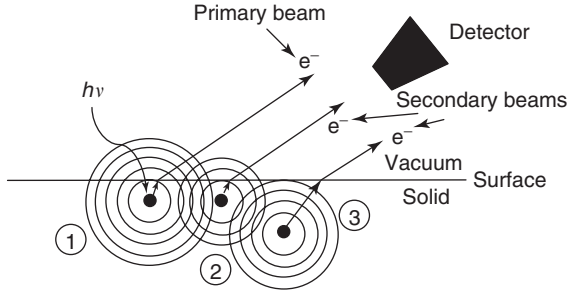


Figure 13 Photoelectron diffraction: A photon impinging on an atom (1) excites an electron represented by a primary spherical wave (primary beam), which is the part of the wave that goes directly into the vacuum. The part scattered by the other ions (2) and (3) inside the solid gives rise to secondary beams, which can interfere with each other on the detector.

Extra-atomic relaxation takes into account the effect of the relaxation energy of the valence electrons on the binding energy of the core orbital of the excited atom. This effect could be strongly influenced by the type of bonds that the emitting atom forms with the neighbor atoms. Note that the binding energy shifts depend on the more or less effective screening of the core hole, which is induced by the relaxing electrons. The shifts will be higher for *s* or *p* electrons than the more localized *d* electrons.

The Case of Ni

Figure 11 shows the CPS spectra of the 3*d*, 3*p*, 3*s*, $2p_{3/2}$, and $2p_{1/2}$ levels of Ni metal. In each case the spectrum consists of a main line accompanied by a satellite line, at about 6 eV higher binding energy. The main lines are aligned to demonstrate the constant distance of the satellite position. The origin of the 6 eV satellite can be explained in terms of final-state effects. In **Figure 12**, the basic physical model is summarized to illustrate the problem. The representation of the band structure of the Ni metal in the initial and final states shown in the figure indicates the origin of the main line and the satellite for the core ionization (c^{-1}). In the initial state $c(3d^9 4s)$, the $3d^9$ state is degenerate with the 4*s* band (left in the figure). In the final state, the electronic Coulomb attraction, due to the core hole created, pulls the unoccupied states below E_F, and the electron screening of the hole produces two effects. Consequent to the first effect, the $3d^9$ level gets filled and the electronic configuration $c^{-1}(3d^{10} 4s)$ (right in the figure) is produced. This is responsible for the main line visible in **Figure 11**. In the second effect, the *d* electrons at the Fermi level are scattered from the Ni ion on which the core hole resides, and a two-hole state

is created (center in the figure). This state is lower in energy than the one-hole state by about 6 eV (satellite line in **Figure 11**).

Photoelectron Diffraction

It has been already mentioned that CPS data can be used to give structural information about surfaces or interfaces by analyzing the diffraction pattern of the electrons emitted by a particular core level. **Figure 13** shows a schematic representation of photoelectron diffraction. The photoelectron, excited by a photon of energy, $h\nu$ can be directly revealed by a detector as a primary beam, or scattered by the neighbor atoms. The elastically scattered waves form secondary beams, which can reach the detector and interfere with each other and with the primary beam, giving rise to an interference pattern. By measuring the photoelectron intensity as a function of the electron wavelength and/or crystal orientation, one can derive structural information. The diffraction pattern must be compared with calculations performed with the use of a hypothetical crystallographic model for the structure considered.

An example of a photoelectron diffraction experiment on an interface is reported in **Figure 14**. The system is formed of one ML of copper embedded between the Ni substrate and a nickel overlayer of increasing thickness. The photoelectron intensity of the Cu $2p_{3/2}$ core level, excited by photons of 1254 eV, is detected as a function of the polar angle ϑ and the Ni overlayer coverage. Note that at high photoelectron energy the forward scattering is dominant, and one expects high photoelectron intensities in the direction connecting the emitting atom and the scatterer. As one can presume from **Figure 14**, the maximum intensity of the Cu $2p_{3/2}$ core level will be reached at $\vartheta = 45°$, for two MLs of Ni; in this case, the two Ni atoms are aligned in the forward direction. For a Ni coverage of four MLs, the maximum is attenuated because of defocusing effects.

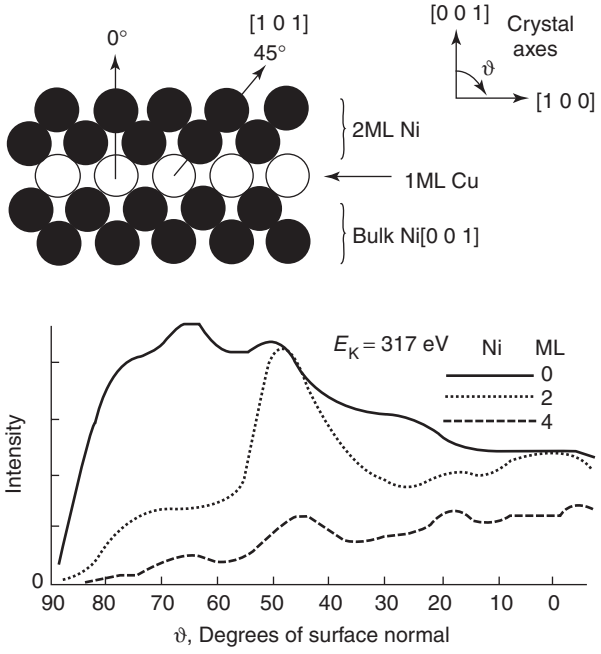
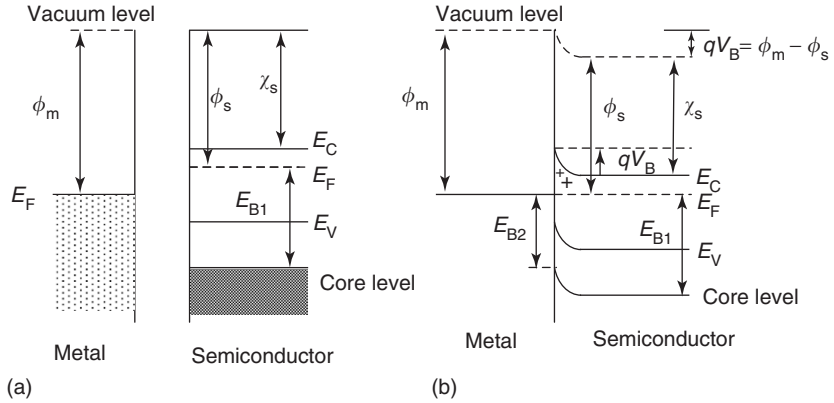


Figure 14 Photoelectron diffraction patterns obtained by scanning the intensity of Cu $2p_{3/2}$ in the (0 1 0) plane as a function of the polar angle ϑ. Inset: schematic diagram of 1 ML of Cu embedded between the Ni bulk and 2 MLs of Ni overlayers. (Reproduced with permission from Egelhoff F, Jr. (1987) Role of multiple scattering in X-ray photoelectron spectroscopy and Auger-electron diffraction in crystals. *Physical Review Letters* 59: 559–562; © American Physical Society.)



$$qV_B = \phi_m - \phi_s = E_{B1} - E_{B2}$$

Figure 15 Band diagram of a metal–semiconductor contact, for the case of an *n*-type semiconductor and for $\phi_m > \phi_s$. (a) Metal and semiconductor are separated; (b) Metal and semiconductor are in contact, and the Schottky barrier is formed by aligning the Fermi levels after thermodynamic equilibrium is reached.

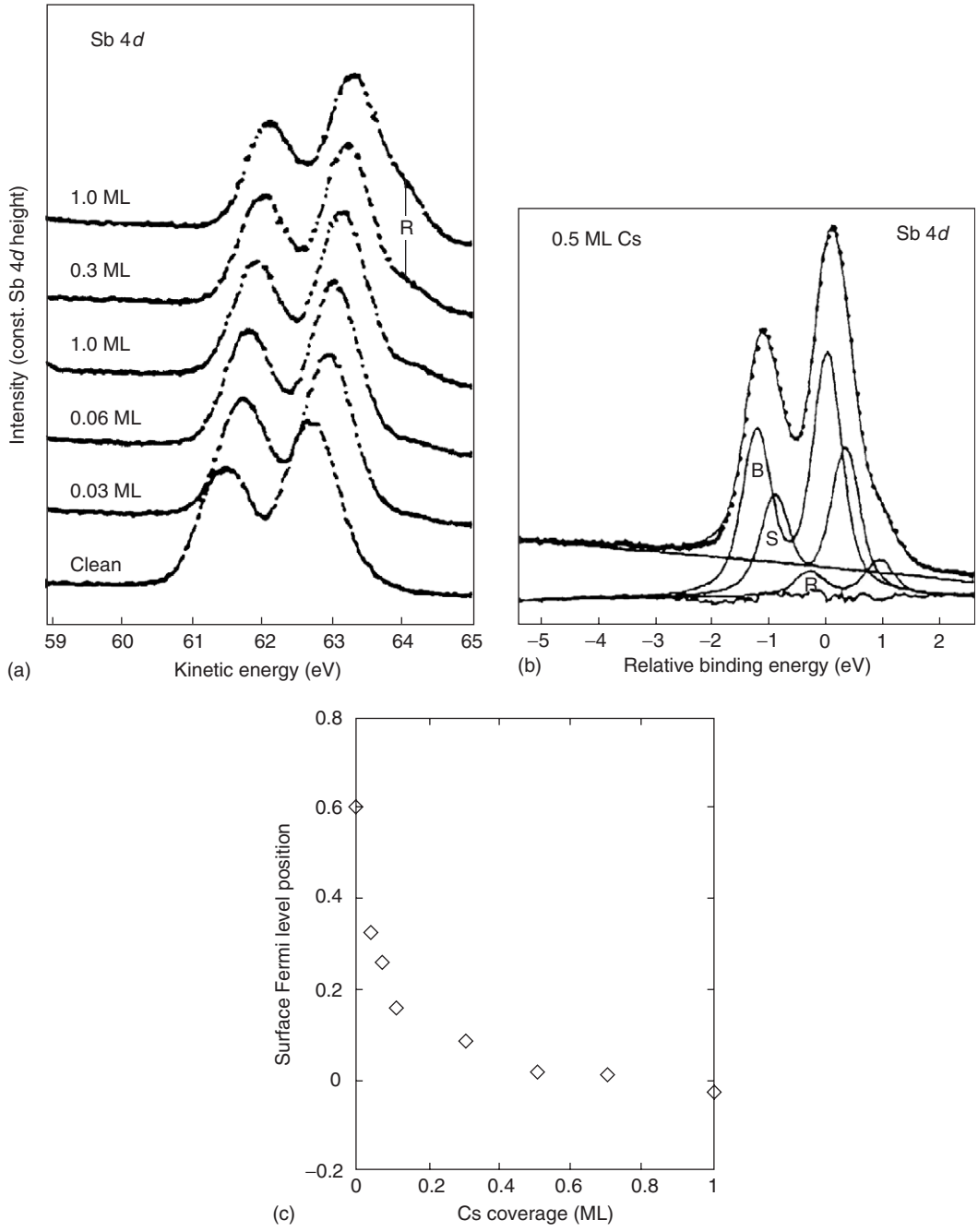


Figure 16 (a) Sb 4*d* core level collected at photon energy of $h\nu = 90$ eV on the Cs/GaSb(1 1 0) interface; (b) peak decomposition of the Sb 4*d* core-level spectrum for 0.5 ML Cs/GaSb(1 1 0); (c) band bending for the interface formation of Cs/GaSb(1 1 0) at room temperature. (Reproduced with permission from Schirm KM, Soukiassian P, Mangat PS, Soonckindt L (1994) Schottky-barrier and interface formation of Cs/GaSb (1 1 0) and Rb/GaSb (1 1 0) at room temperature. *Physical Review B* 49: 5490–5497; © American Physical Society.)

How to Measure the Schottky Barrier by Core-level Photoemission: The Case of Cs/GaSb(1 1 0)

Core-level analysis can provide a lot of information about the first stages of the formation of a metal–semiconductor interface and about the physics of the process. Here a useful application of the CPS technique through the measurement of the Schottky barrier height at a metal–semiconductor junction is illustrated. **Figure 15a** shows the band diagram of a metal and an *n*-type semiconductor before they come into contact. The bands are aligned to the vacuum level and, in the semiconductor, a flat band situation

is assumed, neglecting any possible band-bending due to the presence of surface states. The valence band and the core levels have well-defined binding energies relative to the vacuum level. ϕ_m and ϕ_s are the work functions, with $\phi_m > \phi_s$. When the two systems come into contact, as in **Figure 15b**, a thermodynamic equilibrium is reached by aligning the Fermi levels, with a subsequent charge transfer from the semiconductor to the metal. Band-bending forms at the interface, together with a charge depletion layer that extends some hundred Å inside the semiconductor, depending on the doping concentration. The Schottky barrier is proportional to the built-in potential V_B at the junction; it is given by $qV_B = (\phi_m - \phi_s)$ and is formed entirely after the deposition of a very small amount (a few MLs) of metal. A CPS spectrum, taken from the side of the metal, explores only the first layers of the semiconductor at the interface, picking up the binding energy of the core-level when it has reached the maximum band-bending. The measurement of the binding energy difference of the core levels $E_{B2} - E_{B1}$ before and after the interface growth provides the Schottky barrier height, once the position of the Fermi level on the clean semiconductor surface is known. The position can be found from the Fermi-level measurement on the gold sample in contact with the system under analysis. The Schottky barrier height can also be evaluated by measuring the shift of the top of the valence band as a function of metal coverage. This procedure is affected by an uncertainty in the measurement of the valence-band edge, due to the possible presence of adsorbate states. In **Figure 15a**, a flat-band semiconductor has been hypothesized, which is only an exception. A clean semiconductor surface usually has its own band-bending because of the presence of surface states in the bandgap, which pin the Fermi level. E_F can be obtained from knowledge of the doping concentration, and the photoemission valence band edge is measured with respect to E_F.

The photovoltage effect, which is an accumulation of positive charge in the depletion layer, can lead to additional band-bending. It is induced by incident light creating electron–hole mobile pairs that move in opposite directions in the electric field at the junction. As a result, a charge accumulation takes place at the interface, thereby reducing the band-bending. The photovoltage effect is strongly dependent on temperature, photon flux, and doping concentration and must be taken into account when performing a Schottky-barrier measurement by photoemission spectroscopy.

An example of an application of the Schottky-barrier height measurement is given in **Figure 16** for a Cs/GaSb(1 1 0) metal–semiconductor interface. **Figure 16a** shows the evolution of the Sb 4*d* core-level spectra for different Cs metal deposition thicknesses. The movement of the core-level doublet as a function of the cesium coverage gives an indication of band-bending formation. For a correct evaluation of the band-bending, the bulk contribution is obtained from a best fit to the experimental data (see **Figure 16a**). The position of the surface Fermi level, derived from the energy shift of the bulk peak as a function of Cs coverage, is plotted in **Figure 16c**. Here, one can note that the built-in potential at the junction is completed for a very small quantity of metal ($\sim$1 ML) and that the Schottky-barrier height is $\sim$0.5 eV. This value is not affected by the photovoltage effect, since GaSb is a small-gap semiconductor and heavily *n*-doped.

See also: Interaction of Light and Matter; Luminescence; Optical Absorption and Reflectance; Optical Properties of Materials; Semiconductor Optics; Time-Resolved Optical Spectroscopies; X-Ray Absorption Spectroscopy.

PACS: 79.60. – i; 82.80. – d; 68.35.Bs

Further Reading

Barr TL (1994) *Modern ESCA, The Principles and Practice of X-ray Photoelectron Spectroscopy*. Boca Raton: CRC Press.

Briggs D and Seah MP (eds.) (1990) P*ractical Surface Analysis*, *Auger and X-ray Photoelectron Spectroscopy*, 2nd edn. vol. I. Chichester: Wiley.

Ernst-Eckhard K (1983) In: Eastman DE and Farge Y (eds.) *Handbook on Synchrotron Radiation*, vol. 1a. Amsterdam: North-Holland.

Feuerbacher B, Fitton B, and Willis RF (eds.) (1978) *Photoemission and the Electronic Properties of Surfaces*. Chichester: Wiley.

Lamberti C (2004) Surface science report Elsevier. *The Use of Synchrotron Radiation Techniques in the Characterization of Strained Semiconductor Heterostructures and Thin Films*. Amsterdam: Elsevier.

Sharma BL (1984) *Metal-semiconductor Schottky Barrier Junctions and Their Applications*. New York: Plenum.

Stefan H (2003) *Photoelectron Spectroscopy, Principles and Applications*, 3rd edn. Berlin: Springer.

Nomenclature

e	1 elementary charge (C)
E_B	binding energy (eV)
E_F	Fermi level (eV)
E_k	kinetic energy (eV)
h	Planck constant (eV)
qV_B	Schottky barrier (eV)
Γ_G	gamma$_G$ (eV)
Γ_L	gamma$_L$ (eV)
λ	escape depth (A)
Φ	work function (eV)

Creep *See* Mechanical Properties: Creep.

Crystal Field Splitting

I B Bersuker, University of Texas at Austin, Austin, TX, USA

Introduction

The basis of the crystal field theory (CFT) was created by Bethe in 1929 in his classical work "Term splitting in crystals." This publication contains, in essence, all the main elements of the modern theory. In the 1930s, some important results illustrating the efficiency of the CFT were obtained, including explanation of magnetic behavior of ions in weak and strong crystal fields, in particular, the reduction of the orbital magnetic moment, temperature dependence of the magnetic susceptibility, and the Jahn–Teller effect. More intensive development of CFT began in the 1950s when it was shown that this theory successfully explains the origin of absorption spectra in the visible and related regions, as well as a series of other optical, electric, magnetic, thermodynamic, and electron spin resonance properties of impurity centers in crystals and molecular coordination systems.

The main assumption of CFT is that while the impurity center (IC) or the central atom in clusters or molecular coordination compounds (hereafter denoted as IC) is considered with its detailed electronic structure, the atoms (ions) of the environment are implied to be "structureless" sources of electrostatic fields (sometimes allowing for their polarization in the field of the IC and other atoms). This assumption allows one to consider phenomena that take place mainly within the electronic shell of the IC, but do not involve the electronic structure of the environment explicitly. In spite of this significant limitation, the CFT, within the limits of its applicability, is a rather efficient means to investigate the many aspects of the electronic structure and properties of impurity centers in crystals and other local properties of polyatomic systems.

In accordance with the basic statements of the CFT, the electronic structure of the IC system is determined by the Schrödinger equation with the Hamiltonian H:

$$H = H_0 + V + W \qquad [1]$$

where H_0 includes all the interactions in the IC: the kinetic energy of its n electrons, the interaction between them and with the nucleus; V is the interaction between the IC electrons with the atoms of the environment taken as point charges q_i or dipoles μ_i; and W is the electrostatic interaction of the positive charge Ze of the IC nucleus with the crystal field charges or dipoles. Taking the origin of the polar coordinate system at the nucleus, the N ligand coordinates are denoted as $\boldsymbol{R}(R_i, \vartheta_i, \phi_i), i = 1, 2, \ldots, N$. Then

$$V = -\sum_{i=1}^{N} \sum_{j=1}^{n} \frac{eq_i}{|\vec{r}_j - \vec{R}_i|} \qquad [2]$$

and

$$W = \sum_{i=1}^{N} \frac{Zeq_i}{R_i} \qquad [3]$$

If $q_i < 0$, the value of V in [2] is positive, and hence the electron–ligand interaction destabilizes the IC; for $q_i > 0$, V is negative. The term W in [3] is usually not explicitly considered in the CFT, since it is independent of the electron coordinates and does not influence the electronic properties considered in CFT. Usually, the valence electrons of the IC are most important in CFT, so the number of terms in the first sum of eqn [2] equals the number of the valence electrons, while Ze is the effective charge of the remaining core.

Crystal Field Splitting (CFS) of the Energy Levels of One *d* Electron

Qualitative Aspects and Visual Interpretation

In accordance with the assumptions of the CFT, the crystal field influences the IC properties by changing its electronic structure. The main crystal field effect on the states of the IC is the splitting of its energy levels. The origin of this splitting is known in quantum mechanics as the Stark effect.

Consider the simplest (from the point of view of the CFT) case when the central ion of the IC in an octahedral environment contains only one d electron over the closed shell, for instance, in the $TiCl_6^{3-}$ cluster. The ground state of the Ti^{3+} ion is ^2D with an orbital momentum $L = 2$ (orbital degeneracy $2L + 1 = 5$) and a spin $S = 1/2$ (doublet state). The five orbital states are just the five possible angular states of the only d electron with identical radial parts. The real functions representating the three orbitals d_{xy}, d_{xz}, d_{yz} (known as t_{2g} orbitals) are oriented in space in such a

way that their distribution maxima (lobes) lie in the region between the coordinate axes. The remaining two orbitals d_{z^2} and $d_{x^2-y^2}$ (known as e_g orbitals) have their lobes oriented exactly along the axes.

Compare the electron distributions in the two types of d states, e_g and t_{2g}, say $d_{x^2-y^2}(e_g)$ and $d_{xy}(t_{2g})$, illustrated in **Figure 1**. Taking into account that the Cl^- ions have negative charges, it can be easily concluded that in octahedral complexes the $d_{x^2-y^2}$ electron is subject to a stronger electrostatic repulsion from the chlorines, than the d_{xy} one. Hence, the electron energies of these two states (which are equal in the free ion), under the electrostatic influence of the ligands, become different: the $d_{x^2-y^2}$ energy level is higher than the d_{xy} one. All the three t_{2g} states (d_{xy}, d_{xz}, d_{yz}) are fully equivalent with respect to the six ligands, and therefore they have the same (lower) energy forming a threefold degenerate term. It can be shown that the two e_g states also have equal energies, forming a twofold degenerate term.

Thus, the five d states that have the same energy in the free ion are divided into two groups of states with different energies under the octahedral field of the ligands. In other words, the fivefold degenerate D term of the free ion is split in the field of the ligands of an octahedral complex into two terms: threefold degenerate $^2T_{2g}$ and twofold degenerate ^2E_g:

$$^2\mathrm{D} \rightarrow {}^2T_{2g} + {}^2E_g \qquad [4]$$

It follows from this consideration that the CFS occurs as a result of the smaller repulsion (from the environment) of the t_{2g} states, than e_g states, but all the five states are equally destabilized in the field of the chlorine ions by an amount of E_0. This is illustrated in **Figure 2**.

For a tetrahedral complex, the qualitative picture of the term splitting is inverse to that of the octahedral case. Indeed, in the tetrahedral environment of four ligands, the t_{2g} orbitals are oriented with their lobes much closer to the ligands than the e_g orbitals; hence, the former are subject to stronger repulsion than the latter. Therefore, the energy levels of the t_{2g} orbitals are higher than those of e_g. Symmetry considerations and the calculations, given below, show that even in the tetrahedral system, the three t_{2g} states, as well as the two e_g states, remain degenerate forming the T_2 and E terms, respectively. Hence in a tetrahedral complex the splitting of the D term is (**Figure 3b**)

$$D \rightarrow T_2 + E \qquad [5]$$

which is very similar to that in an octahedral field, eqn [4]. But unlike the octahedral case, the T_2 has greater energy when compared to E, and the splitting magnitude, as well as the destabilization energy, is smaller.

Both the octahedral and tetrahedral symmetries appertain to the cubic groups of symmetry. In **Figure 3**, the splitting of the atomic D term in the field of eight ligands at the corners of a cube is also shown. If the symmetry of the ligand field is lowered, the terms T_2 and E may be subject to further splitting. **Figure 4** illustrates the expected CFS pattern for different symmetry environments of a d electron IC.

If the number of d electrons above the closed shell equals 9, the visual interpretation of the splitting becomes possible again; the d^9 configuration can be

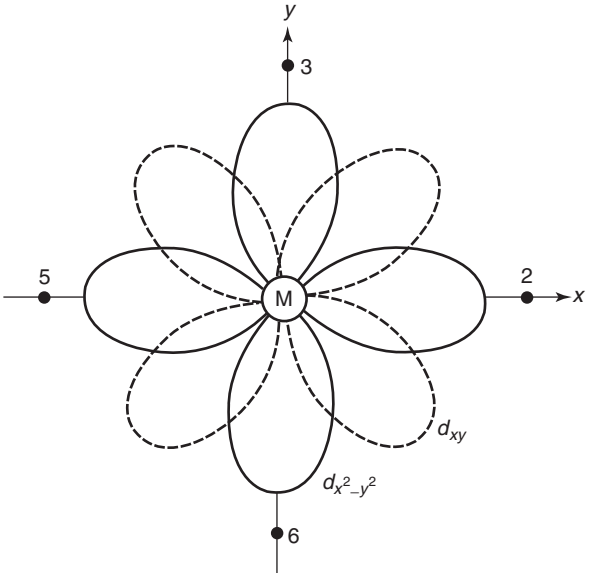


Figure 1 Comparison of electron distributions in the two d states, $d_{x^2-y^2}$ (solid line) and d_{xy} (dashed line) with respect to the four ligands in the xy-plane.

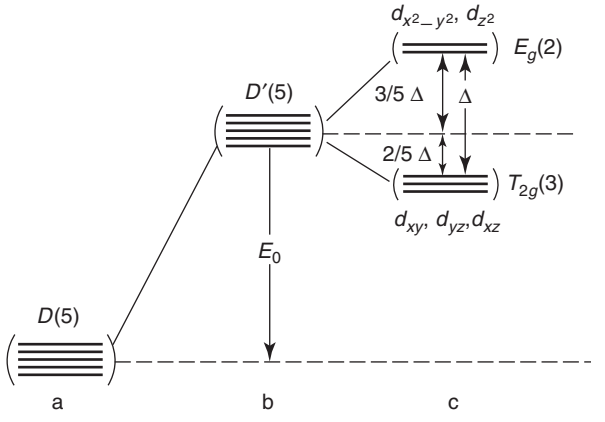


Figure 2 Destabilization E_0 and splitting $\triangle$ of the atomic D term in octahedral crystal fields (the degeneracies are indicated in parenthesis): (a) free atom, (b) spherical averaged crystal field, (c) octahedral field. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

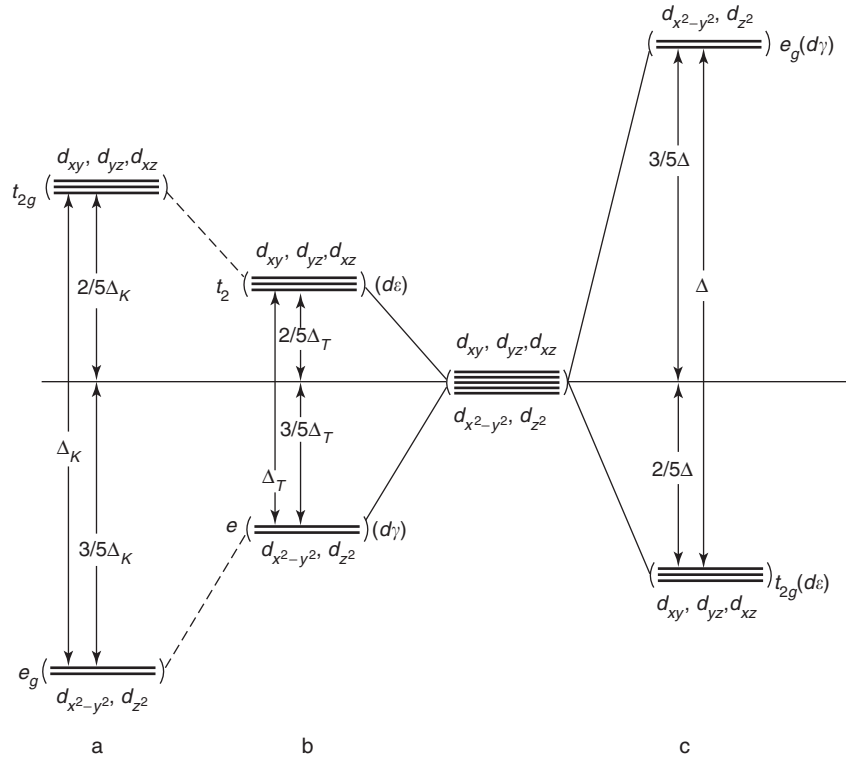


Figure 3 Splitting of the *D* term in (a) cubic, (b) tetrahedral, and (c) octahedral crystal fields. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

formally presented as one positive electronic charge above the closed-shell d^{10} configuration, that is, a d hole in the d^{10} shell. The ground state of the d hole is also ^2D, as for the d^1 configuration, and its splitting in crystal fields of different symmetries has the same components as d^1, but with their mutual arrangement inverse to that of d^1 (repulsion changed to attraction). This rule of inverse term splitting for complimentary electron configurations is valid for any pair of electronic configuration, d^n and d^{10-n} ($n = 1, 2, 3, 4$), in which the number of, respectively, d electrons and d holes over the closed shell d^{10} is the same.

Qualitatively, the expected CFS can be obtained directly from a group-theoretical analysis without solving eqn [6] below. **Table 1** illustrates the expected CFS of different atomic states in crystal fields of different symmetry.

Calculation of CFS Magnitudes

Assume first that the term V in eqn [1] is much smaller than the interatomic interactions H_0, and hence V can be considered as a perturbation to the solutions of H_0. This assumption is valid if the resulting term splitting obtained in this way is smaller than the energy gap between the terms of the free atom (or ion), solutions of H_0. Consider the case of one d electron above the closed shell, that is, the electron configuration of the IC $[A]d^1$, where $[A]$ is a closed shell. The solution for the free ion with the Hamiltonian H_0 yields the fivefold orbitally degenerate ^2D term. To reveal the modifications of this term under the perturbation V from [2], one has to solve the perturbation theory problem, which for fivefold degeneracy reduces to the secular equation of the fifth order with respect to the energy level corrections ε (m is the magnetic quantum number of the atomic d functions ψ_{lm} with $l = 2$):

$$\| V_{mm'} - \varepsilon\delta_{mm'} \| = 0, \quad m,\ m' = 2, 1, 0, -1, -2 \qquad [6]$$

where, in accordance with [2],

$$V_{mm'} = \sum_i eq_i \int \psi_m^* \psi_{m'} / |\boldsymbol{r} - \boldsymbol{R}_i| \, \mathrm{d}\tau \qquad [7]$$

The general expression for $V_{mm'}$ can be evaluated directly:

$$\begin{aligned} V_{mm'} = \sum_i eq_i \, [& A_{mm'} F_4(R_i) Y_4^{m-m'}(\vartheta_i, \phi_i) \\ & + B_{mm'} F_2(R_i) Y_2^{m-m'}(\vartheta_i, \phi_i) \\ & + D_{mm'} F_0(R_i) Y_0^{m-m'}(\vartheta_i, \phi_i) \end{aligned} \qquad [8]$$

where $A_{mm'}$, $B_{mm'}$, and $D_{mm'}$ are numerical constants, determined by corresponding Clebsh–Gordan

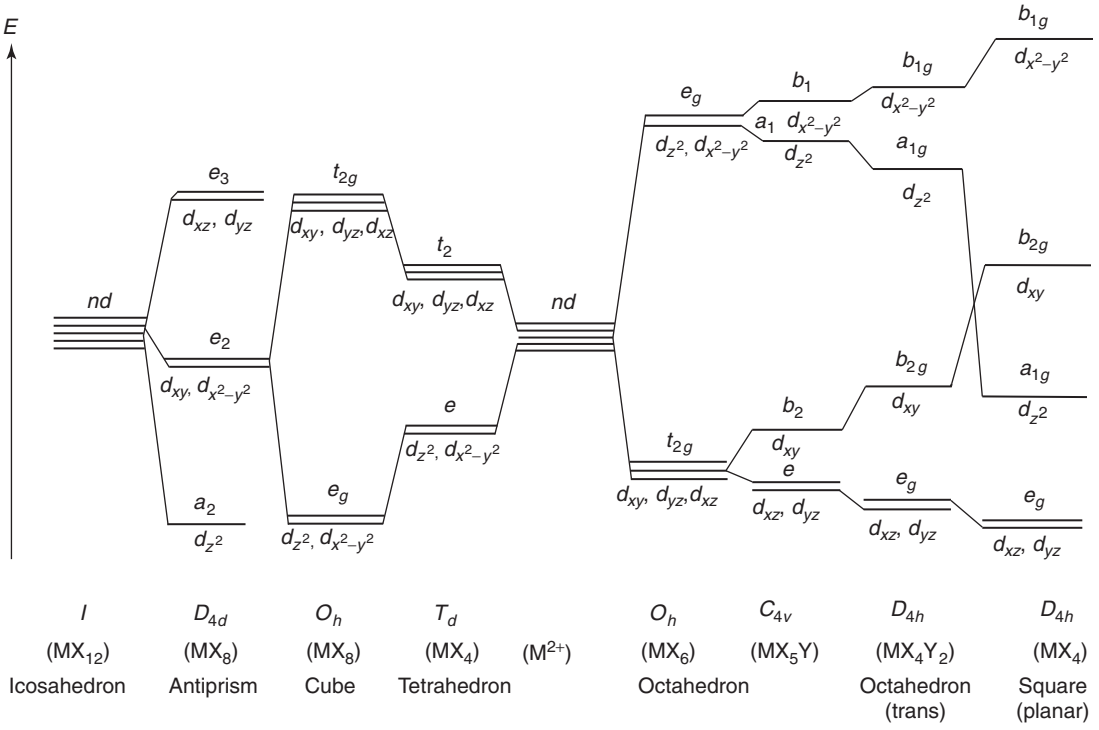


Figure 4 Splitting of *d* orbital energy levels in crystal fields of different symmetries. In MX_5Y and MX_4Y_2 clusters, the splitting of the T_{2g} and E_g terms can be inverted depending on the ratio of field strengths *X*/*Y*. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory.* New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

Table 1 Types of symmetry (irreducible representations) to which the atomic states with given quantum numbers *L*(*l*) or *J*(*j*) belong in point groups of different symmetry[a]

L or *J*	O_h	T_d	D_3	D_{4h}	C_{4v}	C_{2v}
0	A_{1g}	A_1	A_1	A_{1g}	A_1	A_1
1	T_{1u}	T_2	$A_2 + E$	$A_{2u} + E_u$	$A_1 + E$	$A_1 + B_1 + B_2$
2	E_g	E	E	$A_{1g} + B_{1g}$	$A_1 + B_1$	$2A_1$
	T_{2g}	T_2	$A_1 + E$	$B_{2g} + E_g$	$B_2 + E$	$A_2 + B_1 + B_2$
3	A_{2u}	A_1	A_2	B_{1u}	B_2	A_2
	T_{1u}	T_2	$A_2 + E$	$A_{2u} + E_u$	$A_1 + E$	$A_1 + B_1 + B_2$
	T_{2u}	T_1	$A_1 + E$	$B_{2u} + E_u$	$B_1 + E$	$A_1 + B_1 + B_2$
4	A_{1g}	A_1	A_1	A_{1g}	A_1	A_1
	E_g	E	E	$A_{1g} + B_{1g}$	$A_1 + B_1$	$2A_1$
	T_{1g}	T_1	$A_2 + E$	$A_{2g} + E_g$	$A_2 + E$	$A_2 + B_1 + B_2$
	T_{2g}	T_2	$A_1 + E$	$B_{2g} + E_g$	$B_2 + E$	$A_2 + B_1 + B_2$
5	E_u	E	E	$A_{1u} + B_{1u}$	$A_2 + B_2$	$2A_2$
	T_{1u}	T_2	$A_2 + E$	$A_{2u} + E_u$	$A_1 + E$	$A_1 + B_1 + B_2$
	T_{2u}	T_1	$A_1 + E$	$B_{2u} + E_u$	$B_1 + E$	$A_1 + B_1 + B_2$

[a] Belonging to several types of symmetry means corresponding splitting.

coefficients, while the functions $F_k(R)$ are

$$F_k(R) = R^{-(k+1)} \int_0^R r^k R_{nl}^2(r) r^2 \, dr + R^k \int_R^\infty r^{-(k+1)} R_{nl}^2(r) r^2 \, dr \qquad [9]$$

with $R_{n2}(r)$ as the radial function of the *d* electrons with $l = 2$.

As an example, one can calculate the splitting of a *d* electron term in an octahedral crystal field of six identical ions at the corners of a regular octahedron. The ligand charges and coordinates are

$$\begin{aligned} &q_i = q, \qquad R_i = R, \quad i = 1, 2, \ldots, 6 \\ &\vartheta_1 = 0, \qquad \vartheta_2 = \vartheta_3 = \vartheta_5 = \vartheta_6 = \pi/2, \qquad \vartheta_4 = \pi \\ &\phi_2 = 0, \qquad \phi_3 = \pi/2, \qquad \phi_5 = \pi, \qquad \phi_6 = 3\pi/2 \end{aligned} \qquad [10]$$

Substituting these values into eqn [8] and solving eqn [6], the following result is obtained:

$$\begin{aligned}\varepsilon_1 = \varepsilon_2 = \varepsilon(E_g) = E_0 + (3/5)\Delta \\ \varepsilon_3 = \varepsilon_4 = \varepsilon_5 = \varepsilon(T_{2g}) = E_0 - (2/5)\Delta\end{aligned} \quad [11]$$

with

$$\Delta = \varepsilon(E_g) - \varepsilon(T_{2g}) = (5/3)eqF_4(R) \quad [12]$$

and

$$E_0 = 6eqF_0(R) \quad [13]$$

Thus, in accordance with the above described qualitative results, the d electron energy levels (^2D term) are split by the octahedral ligand field into a twofold (E_g) and threefold (T_{2g}) degenerate terms. Since $F_k > 0$, the twofold degenerate E_g term has higher energy than the T_{2g} term. The splitting magnitude Δ is the main CFS parameter. Note that the CFS preserves the center of gravity: the sum of the energy level displacements (from the E_0 value) multiplied by their degeneracies equals zero.

As compared to the octahedral crystal field, the corresponding splitting in the tetrahedral ones is inverted and smaller (**Figure 3**). Calculations yield

$$\begin{aligned}\varepsilon(T_2) = E_0^T + (2/5)\Delta_T \\ \varepsilon(E) = E_0^T - (3/5)\Delta_T\end{aligned} \quad [14]$$

where the splitting parameter Δ_T and destabilization energy E_0^T are

$$\Delta_T = (20/27)eqF_4(R) \quad [15]$$

$$E_0^T = 4eqF_0(R) \quad [16]$$

It is observed that with the same R and q, $\Delta_T = -(4/9)\Delta$, that is, the splitting magnitude in tetrahedral symmetry "ceteris paribus" is 4/9 times the octahedral splitting. In the cubic coordinate system (eight ligands at the corners of a cube), the splitting and destabilization energy are qualitatively similar to the tetrahedral case, but two times larger in magnitude:

$$\mathbf{\Delta}_C = (40/27)eqF_4(R) = 2\mathbf{\Delta}_T \quad [17]$$

$$E_0^C = 8eqF_0(R) = 2E_0^T \quad [18]$$

If the ligand is a point dipole, the term splitting is qualitatively the same, as in the case of point charges. Assuming that the dimensions of the dipole with the dipole moment μ_i are much smaller than the distance to the IC, one can obtain the following expression for the splitting magnitude in the octahedral case:

$$\Delta = -(5/3)\mu_i F_4'(R) \quad [19]$$

where the prime at F means its derivative; note that $F_4' < 0$.

For lower symmetry fields, the calculations can be carried out in a similar way.

CFS for Configurations with Several *d* Electrons

The Case of Weak Crystal Field

If the electronic configuration of the IC contains more than one d electron above the closed shell, the picture of possible energy terms and their splitting in the ligand fields is significantly complicated by the interaction between the d electrons. If the ligand field is not very strong, the atomic terms can still be classified by the quantum number of the total atomic momentum L, and the influence of the ligands can be taken as a perturbation of the atomic terms.

For the electronic configuration of the IC $[A](nd)^2$, the possible atomic terms are ^3F, ^3P, ^1G, $1D$, and ^1S, the ^3F term belonging to the ground state. The perturbation operator, following eqn [2], has two terms and the calculation of its matrix elements is similar to the d^1 case. For instance, for the ground-state term ^3F of the d^2 configuration in the octahedral field of O_h symmetry, the sevenfold orbital degeneracy ($L = 3,\ 2L + 1 = 7$) splits into three terms; $^3F \rightarrow {}^3A_g + {}^3T_{2g} + {}^3T_{1g}$, with energies:

$$\begin{aligned}\varepsilon(^3A_{2g}) = 2E_0 + (6/5)\Delta \\ \varepsilon(^3T_{2g}) = 2E_0 + (1/5)\Delta \\ \varepsilon(^3T_{1g}) = 2E_0 - (3/5)\Delta\end{aligned} \quad [20]$$

Since Δ and E_0 are positive, $^3T_{1g}$ is the ground term, and $^3T_{2g}$ and $^3A_{2g}$ follow consecutively (**Figure 5**). Their wave functions can be obtained as solutions of the secular equation [6] in the form of linear combinations of the atomic functions.

The splitting of other terms of the $[A](nd)^2$ configuration in obtained similarly (**Figure 5**):

^1D term:

$$\begin{aligned}\varepsilon(^1E_g) = 2E_0 + (12/35)\Delta \\ \varepsilon(^1T_{2g}) = 2E_0 - (8/35)\Delta\end{aligned} \quad [21]$$

^3P term:

$$\varepsilon(^3T_{2g}) = 2E_0 \quad [22]$$

^1G term:

$$\begin{aligned}\varepsilon(^1A_{1g}) = 2E_0 + (2/5)\Delta \\ \varepsilon(^1E_g) = 2E_0 + (2/35)\Delta \\ \varepsilon(^1T_{2g}) = 2E_0 - (13/35)\Delta \\ \varepsilon(^1T_{1g}) = 2E_0 + (1/5)\Delta\end{aligned} \quad [23]$$

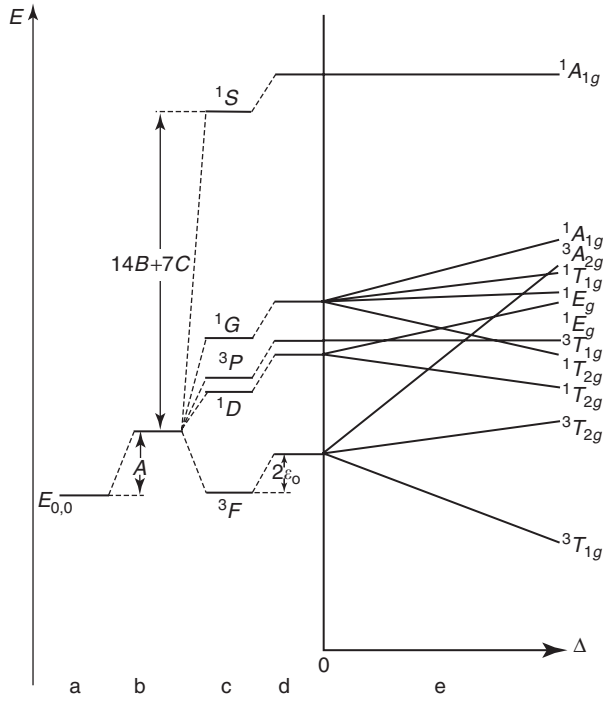


Figure 5 Splitting of the terms of the electronic d^2 configuration in octahedral crystal field (weak field limit): (a) *d* electron energy level, (b) electron interaction – spherical averaged part, (c) electron interaction – formation of atomic terms, (d) crystal field destabilization, (e) ligand field splitting as a function of △. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

^1S term:

$$\varepsilon(^1A_{1g}) = 12eqF_0 = 2E_0 \qquad [24]$$

The criterion of validity of the weak field approximations is that the term splitting is much smaller than the energy gap between the terms. As seen from **Figure 5**, for large Δ values, the components of the split terms even cross each other making the approximation of the weak field invalid.

As with one *d* electron, the tetrahedral splitting is similar to the octahedral one, but with the inverse order of the energy levels: $\varepsilon(^3A_2) < \varepsilon(^3T_2) < \varepsilon(^3T_1)$. Again, $E_0^T = (2/3)E_0$ and $\Delta_T = -(4/9)\Delta$, provided the ligand charges q and their distances R to the IC are the same, as in the octahedron. For systems with lower symmetries the calculations are more difficult, but they can be reduced by using the method of equivalent operators or irreducible tensor operators.

The qualitative picture of term splitting for electronic configurations $[A](nd)^n$ with $n > 2$ in fields of lower symmetry can be found directly by using the complimentary rule: the configurations d^n and d^{10-n} have mutually inverted schemes of term splitting. In the case of weak fields, this rule is also valid for the pairs of configurations d^n and d^{5-n}.

Strong Fields. Low-Spin and High-Spin Centers

In the other limit case, opposite to the weak field one, the effect of the ligand field on the states of the IC is strong; it surpasses the electrostatic interaction between the electrons. In this case, the orbital coupling between the electrons is broken and, in fact, the states with a definitive total momentum quantum numbers, *L* (*S*, *P*, *D*, states etc.) cease to exist. In other words, each *d* electron chooses its orientation in space under the influence of the ligand field rather than the other *d* electrons. A formally similar situation occurs when the orbital coupling between the electrons is broken by the spin–orbital interaction (cf. the *jj*-coupling scheme). This is the so-called "strong ligand field limit."

It follows that when the ligand field is strong, atomic term splitting cannot occur since the terms themselves are destroyed. To determine the states in this case, one should first find out the orientations of each of the *d* states in the ligand field, neglecting the electron interaction, and then evaluate the possible terms of the system taking into account the interaction of the electrons in these crystal-field-oriented electronic states. For one *d* electron in the octahedral field of the ligands, there are two nonequivalent orbital states: the more stable t_{2g} state (d_{xy}, d_{xz}, d_{yz}) in which the electrostatic repulsion from the six ligands is smaller, and the less stable (higher in energy by △) state e_g ($d_{x^2-y^2}$, d_{z^2}) in which the repulsion from the ligands is larger. Hence in the strong ligand field, neglecting the electron interaction, the *d* electrons first occupy the t_{2g} orbitals (maximum six electrons) and then the e_g orbitals (four electrons); the electron configuration is $(t_{2g})^n$ for $n < 6$, and $(t_{2g})^6(e_g)^{n-6}$ for $n > 6$. The energy terms can be obtained from these configurations by including the electron interaction.

Consider the case of the atomic electron configuration $[A](nd)^2$. The two *d* electrons in the ground state of an octahedral complex occupy two t_{2g} orbitals (the state with two electrons in one orbital is higher in energy) forming the $(t_{2g})^2$ configuration. In the excited states, one of the two *d* electrons can occupy the e_g orbital and form the $(t_{2g})^1(e_g)^1$ configuration, which has higher (by △) energy than $(t_{2g})^2$; the two electrons can occupy the e_g orbitals, resulting in the excited $(e_g)^2$ configuration, which has high energy compared to $(t_{2g})^1(e_g)^1$ (by △) and also compared to $(t_{2g})^2$ (by 2△).

Thus, the d^2 configuration in the strong octahedral field forms three configurations $(t_{2g})^2$, $(t_{2g})^1(e_g)^1$, and $(e_g)^2$ – situated consecutively with an energy

spacing $\triangle$ (**Figure 6c**). In each of these configurations, the electron interaction yields several terms, similar to the term formation in the free atom. The resulting terms for the $(t_{2g})^2$ configuration expressed by atomic Racah parameters A, B and C (obtainable from spectroscopic data) are as follows:

$$\begin{aligned}\varepsilon(^3T_{1g}) &= A - 5B \\ \varepsilon(^1T_{2g}) &= A + B + 2C \\ \varepsilon(^1E_g) &= A + B + 2C \\ \varepsilon(^1A_{1g}) &= A + 10B + 5C\end{aligned} \qquad [25]$$

The splitting of all the above configurations that emerge from d^2 is shown in **Figure 6e**. In particular, the ground state of the $(t_{2g})^2$ configuration $^3T_{1g}$ is the same as in the weak field limit. However, the sequence and spacing of the excited states is essentially different. The differences occur for $n = 4, 5, 6, 7$ in octahedral symmetry, and for $n = 3, 4, 5, 6$ in tetrahedral systems. It is important that in these cases, the spin multiplicity of the ground state is always lower in the strong field limit than in the weak field. Therefore, the centers with strong ligand fields are called low-spin centers, as distinct from complexes with weak ligand fields known as high-spin centers.

The criterion of validity for the strong field limit requires that the CFS is smaller than the energy gap $\triangle$ between the three configurations: $(t_{2g})^2, (t_{2g})^1\ (e_g)^1$, and $(e_g)^2$. For the splitting of the $(t_{2g})^2$ configuration the largest distance between its components in [25] is $15B + 5C$; hence, the condition of validity of the strong field approach is $15B + 15C \ll \triangle$. Otherwise the terms of the same symmetry from different configurations become strongly mixed (e.g., $^1T_{2g}$ from $(t_{2g})^2$ with $^1T_{2g}$ from $(t_{2g})^1(e_g)^1$); this is known as "configuration interaction."

For more than two d electrons, the criterion of validity of the strong field approximation can be established similarly. Of special interest are the cases of d^4, d^5, d^6, and d^7 in octahedral complexes and d^3, d^4, d^5, and d^6 in tetrahedral systems, for which the two-limit cases differ by the spin of the ground state. The pairing energy Π is defined as the difference between the energies of interelectron interactions in the low-spin and high-spin complexes, respectively, divided by the number of pairings destroyed by the low-spin → high-spin transition. It is obvious that the low-spin state is preferable if

$$\Pi < \triangle \qquad [26]$$

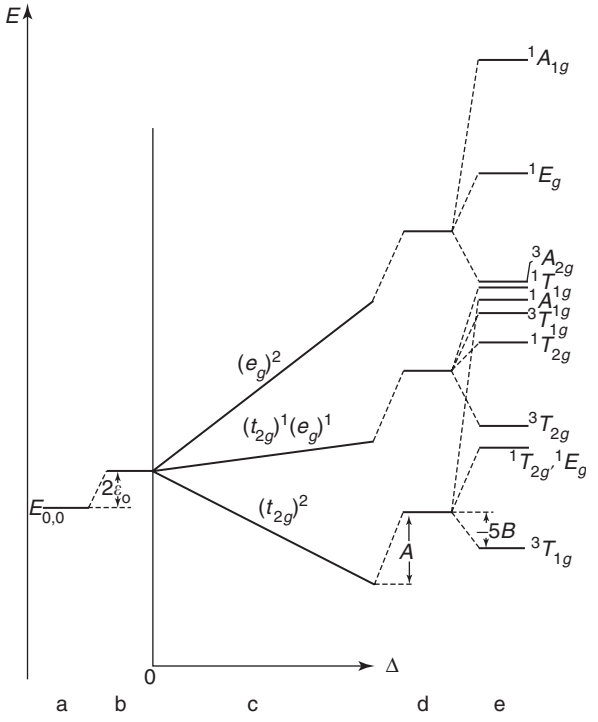


Figure 6 Splitting of the terms of the d^2 configuration in strong fields of octahedral symmetry: (a) d electron energy level, (b) crystal field destabilization, (c) crystal field splitting as a function of $\triangle$, (d) electron interaction destabilization, (e) electron interaction splitting. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

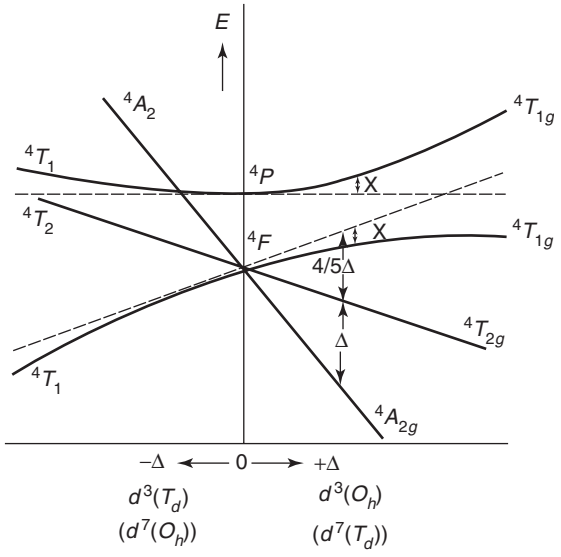


Figure 7 Splitting of the IC ground-state term ^4F of the $d^3(d^7)$ electron configuration in octahedral O_h and tetrahedral T_d crystal fields as a function of the CFT parameter $\triangle$ with the $T_{1g}(F) - T_{1g}(P)$ interaction included; x indicates the deviations of energy terms of the same symmetry due to the nonintersection rule. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

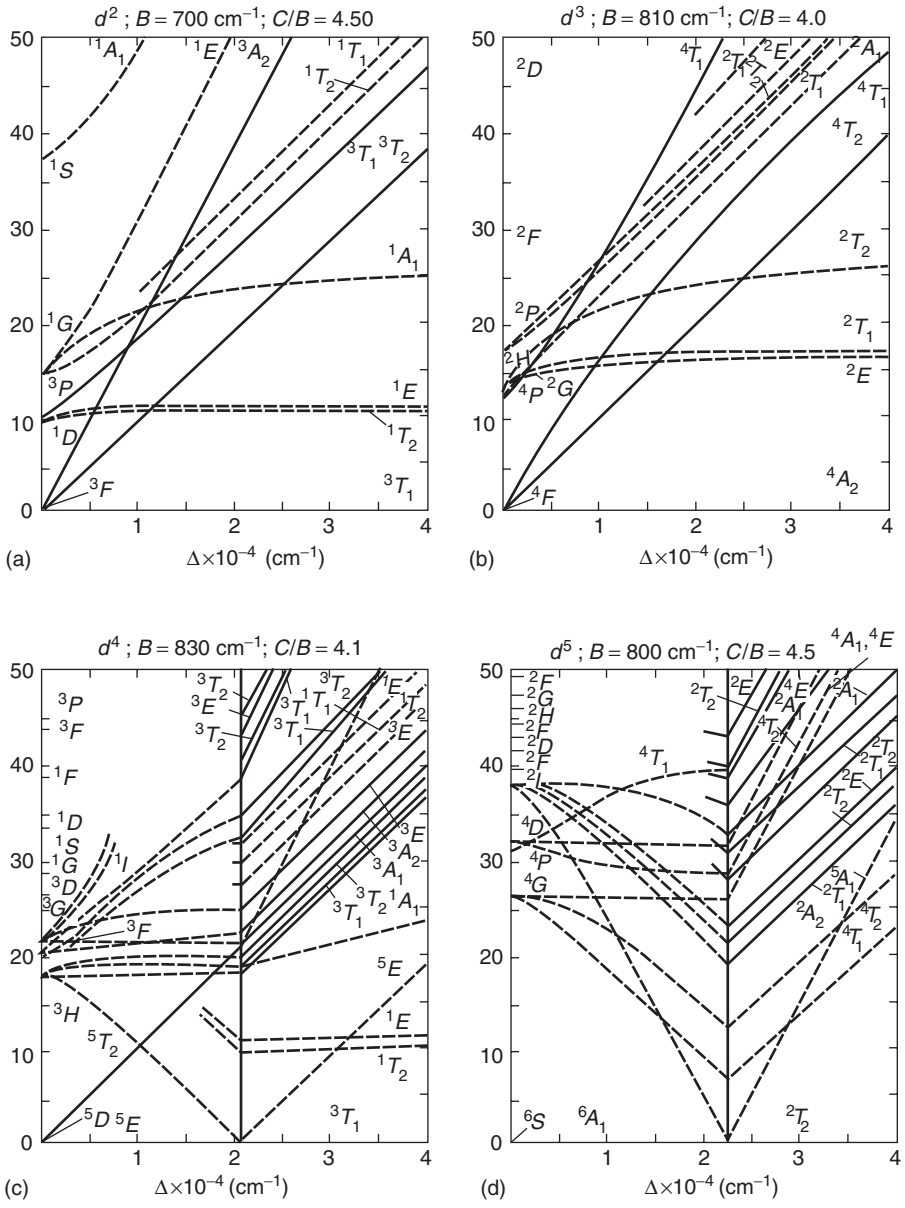


Figure 8 (a–g) The Tanabe–Sugano diagrams: energy terms of IC with d^n configurations, $n = 2, 3, \ldots, 8$, as a function of octahedral crystal fields $\triangle$ in 10^3 cm^{-1}. For convenience, the energy terms with multiplicities different from the ground state ones are given by dashed lines. The indices *g* and *u* are omitted. Some levels of minor significance are not shown. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

On the contrary, if

$$\Pi > \triangle \qquad [27]$$

the high-spin state is the ground state. The pairing energy is the lowest for the d^6 configuration, and hence the low-spin state is preferable in octahedral fields with this configuration as compared with others ceteris paribus. Since the $\triangle$ value in tetrahedral fields is significantly smaller than that for octahedral systems, the low-spin configuration for the former is much less probable than for the latter.

Tanabe–Sugano Diagrams

If the ligand field has intermediate strength for which neither the criterion of weak field nor of strong field is realized, the problem should be solved with the ligand field and electron interactions considered simultaneously to account for the configuration interaction, which produces repulsion of terms with the same symmetry ("nonintersection rule"). **Figure 7** illustrates this effect for the $^4T_{1g}(F)-{}^4T_{1g}(P)$ repulsion in octahedral ($^4T_1(F)-{}^4T_1(P)$ in tetrahedral)

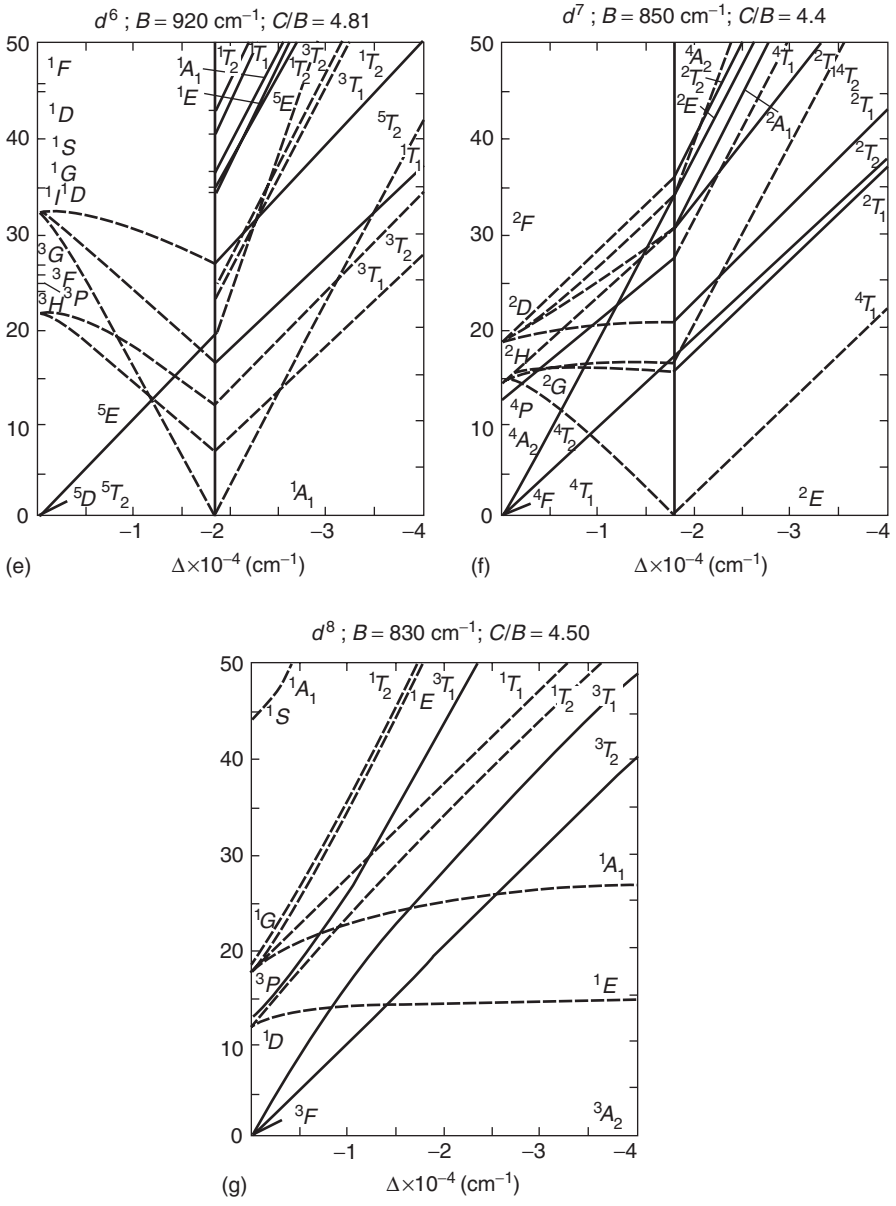


Figure 8 (Continued)

$d^3(d^7)$ IC. Energy level diagrams as functions of the CFT parameter Δ for all the d^n configurations ($n = 2$, 3, 4, 5, 6, 7, 8) were calculated numerically by Tanabe and Sugano (**Figures 8a–8c**). In these diagrams, the energy read off is taken at the ground state. Therefore, at a certain value of Δ (more precisely, Δ/B), there is a term crossing, the ground state changes and all the energy levels on the diagram are subject to a break. Usually at this break the ground-state multiplicity also changes, and there is a transition from the weak ligand field to the strong field. The Tanabe–Sugano diagrams give the most complete information about the electronic structure of the system in the CFT model.

f-Electron Term Splitting

One of the important features of *f* electrons is that they are usually screened from the ligand field by the outer *s*, *p*, *d* electrons and hence they are less affected by the ligands than the *d* electrons. On the other hand, the *f* electrons are subject to stronger spin–orbital coupling. For *f* electrons in weak fields, the atomic terms should be characterized (in addition to L and S) by the quantum number $J = L + S - 1,\ L + S - 2, \ldots, |L - S|$ that takes into account the spin–orbital interaction. Since the total spin S can be half-integer, J may also be half-integer. For instance, the states of one *f* electron are characterized by $L = 3, S = 1/2, J = 7/2$ and $5/2$,

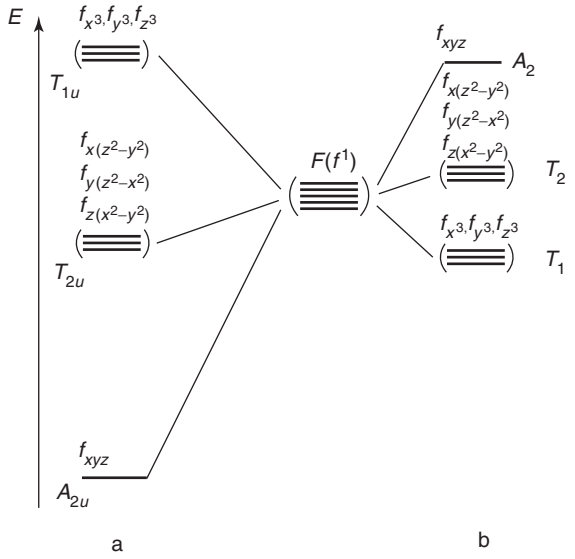


Figure 9 Splitting of the atomic energy levels of one *f* electron in (a) octahedral and (b) tetrahedral crystal fields (the spin–orbital interaction is neglected). (Reprinted with permission from *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

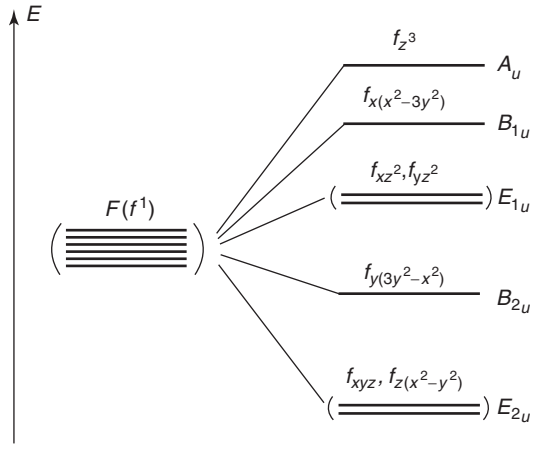


Figure 10 Splitting of one *f* electron energy levels in the field of a hexagonal biprism. (Reprinted with permission from Bersuker IB (1996) *Electronic Structure and Properties of Transition Metal Compounds. Introduction to the Theory*. New York: Wiley; © Wiley. This material is used by permission of John Wiley & Sons, Inc.)

and with the spin–orbital interaction included there are two terms: $^2F_{7/2,}$ $^2F_{5/2}$. The weak field approximation here means that the splitting of each of these terms by the ligand field can be considered separately.

A visual interpretation of the CFS can be obtained in a simplified model by neglecting the spin–orbital coupling. In particular, considering the angular distributions of atomic *f* electron functions (from the cubic set) and the corresponding electrostatic repulsion of the electron in these states from six point charges (or dipoles) of an octahedral cluster, one can conclude that in the three states f_{x^3}, f_{y^3}, and f_{z^3}, the repulsion is the greatest (and equal for all of them). In the other three states $f_{x(y^2-z^2)}$, $f_{y(z^2-x^2)}$, $f_{z(x^2-y^2)}$, the repulsion is equally strong but smaller than in the previous three states, and in the f_{xyz} state it is the smallest. Thus, the sevenfold orbitally degenerate *F* term of the free atom (ion) with one *f* electron is split by the octahedral crystal field into three terms from which one is nondegenerate and two are threefold degenerate, $F \rightarrow A_{2u} + T_{2u} + T_{1u}$ (**Figure 9**). In tetrahedral fields, similarly to *d* electron states, the CFS picture is inverted. By way of example, one can also show the scheme of splitting of the energy levels of an *f* electron in the field of a hexagonal biprism with the sixfold axis along the *z*-axis (**Figure 10**), which is realized, for instance, in uranyl centers (here the low-symmetry set of one-electron atomic angular *f* functions is used).

With the spin–orbital interaction included, the *f* electron CFS can be determined qualitatively by symmetry considerations. If *J* is an integer, the splitting coincides completely with that expected for the corresponding *L* value given in **Table 1**. For semi-integer *J* values, the double groups of symmetry should be employed, the terms being classified by their irreducible representations E_1', E_2' (twofold degenerate), G' (fourfold) or, respectively, $\Gamma_6, \Gamma_7, \Gamma_8$ in Bethe's notations (**Table 2**).

Table 2 Splitting of atomic terms with semi-integer *J* values in crystal fields of different symmetry

J	*Cubic symmetry* O'	*Tetragonal symmetry* D_4'	*Hexagonal symmetry* D_6'
1/2	E_1'	E_1'	E_2'
3/2	G'	$E_1' + E_2'$	$E_2' + E_3'$
5/2	$E_2' + G'$	$E_1' + 2E_2'$	$E_1' + E_2' + E_3'$
7/2	$E_1' + E_2' + G'$	$2E_1' + 2E_2'$	$E_1' + 2E_2' + E_3'$
9/2	$E_1' + 2G'$	$3E_1' + 2E_2'$	$E_1' + 2E_2' + 2E_3'$
11/2	$E_1' + E_2' + 2G'$	$3E_1' + 3E_2'$	$2E_1' + 2E_2' + 2E_3'$

See also: Insulators, Impurity and Defect States in; Metals and Alloys, Impurity and Defect States in; Semiconductors, Impurity and Defect States in.

PACS: 71.70.Ch; 71.55. – i

Further Reading

Bethe H (1928) Term splitting in crystals. *Annals of Physics* 3: 133–208.

Gerloch M and Slade RS (1973) *Ligand Field Parameters*. Cambridge: Cambridge University Press.

Griffith JS (1962) *The Theory of Transition Metal Ions*. Cambridge: Cambridge University Press.

Nielson CW and Koster GF (1963) *Spectroscopic Coefficients for the p^n, d^n, and f^n Configurations*. Cambridge: MIT Press.
Schlafer HL and Gliemann G (1969) *Basic Principles of Ligand Field Theory*. New York: Wiley.
Sugano S and Tanabe Y (1970) *Multiplets of Transition Metal Ions in Crystals*. New York: Academic Press.
Wyborne BG (1965) *Spectroscopic Properties of Rare Earth*. New York: Interscience.

Crystal Growth, Bulk: Methods

G Müller, University of Erlangen-Nuernberg, Erlangen, Germany
J Friedrich, Fraunhofer Institute IISB, Erlangen, Germany

Introduction

Many modern technological systems would not exist without the availability of synthetic single crystals. Therefore, the technology to prepare and to produce bulk single crystals – which is called "crystal growth" – is one of the current key technologies.

This article gives a brief overview about the techniques which are used to grow bulk single crystals of inorganic materials. The fundamental physical and physicochemical mechanisms which govern the crystal growth processes are treated elsewhere. The growth of epitaxial films as well as of organic materials are also not considered.

The structure of the article is sketched in **Figure 1**.

Melt Growth

The most frequently used and most important method of producing bulk single crystals is by solidifying

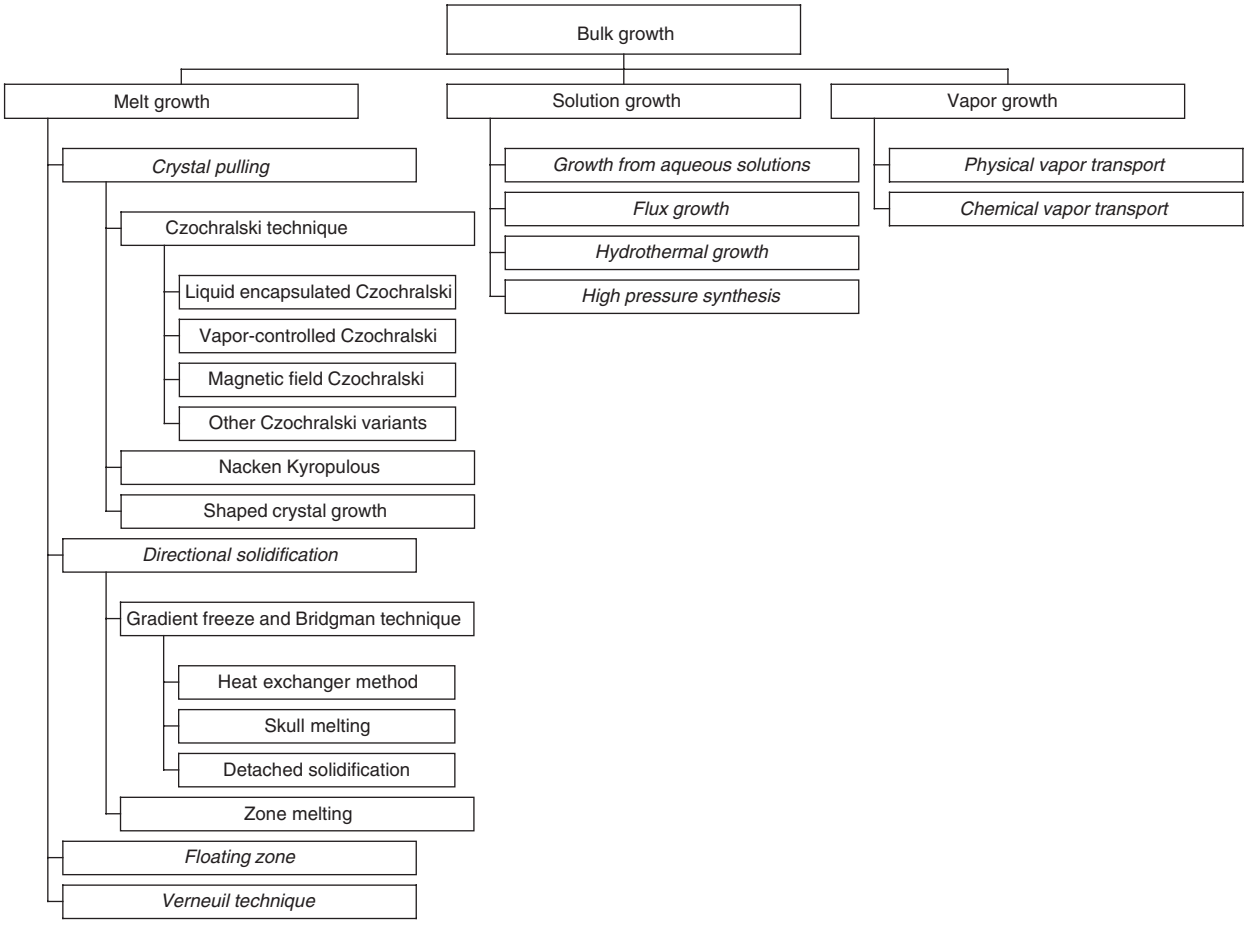


Figure 1 Overview of methods for bulk crystal growth.

the material from its melt (melt growth). Only if the crystal to be grown has no congruent melting point, or if the melting point or vapor pressure are too high, the crystals have to be grown either from solutions or from the vapor phase.

Crystal Pulling

Czochralski technique The Czochralski (Cz) method is the most important method for the production of bulk single crystals of a wide range of electronic and optical materials (**Figure 2**). At the beginning of the process, the feed material is put into a cylindrically shaped crucible and melted by resistance or radio-frequency heaters. After the feed material is completely molten, a seed crystal with a diameter of typically a few millimeters is dipped from the top into the free melt surface and a melt meniscus is formed. Then, after re-melting of a small portion of the dipped seed, the seed is slowly withdrawn from the melt (often under rotation) and the melt crystallizes at the interface of the seed by forming a new crystal portion. During the further growth process, the shape of the crystal, especially the diameter, is controlled by carefully adjusting the heating power, the pulling rate, and the rotation rate of the crystal. Therefore, an automatic diameter control is generally applied. This diameter control is based, either on the control of the meniscus shape (e.g., for silicon) or on the weighing of the crystal (e.g., for GaAs, InP) or of the melt (for oxides).

In order to control the convective heat and species transport in the melt including the shape of the solid–liquid interface, which plays one of the most dominant roles in terms of the crystal quality, a proper combination of crystal and crucible rotation is used during the whole process.

The most important technical application of the Cz method is the growth of dislocation – free silicon crystals with diameters of 300 mm and a weight up to 300 kg (**Figure 3**). Germanium crystals are also mainly produced by the Cz method as well as several technically important oxide and fluoride crystals like

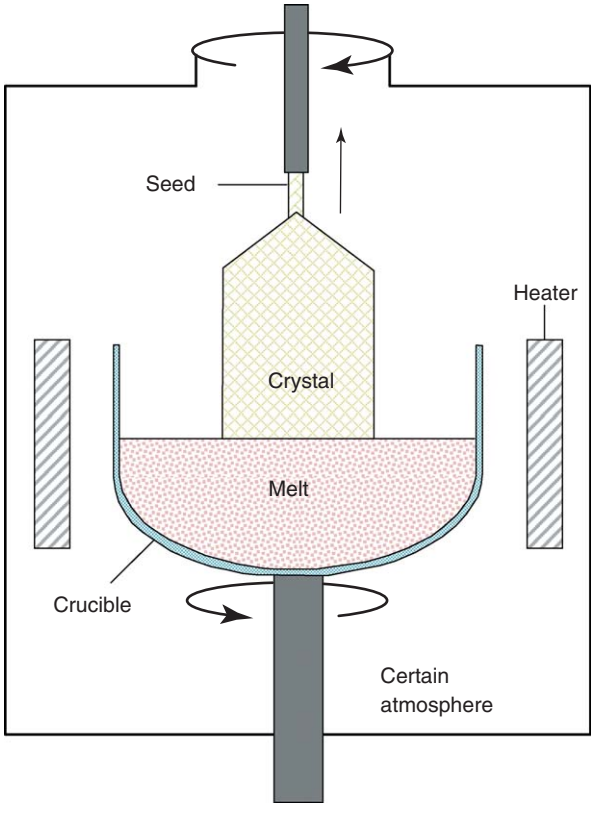


Figure 2 Schematic of the principle of the Cz method.

Figure 3 Silicon crystal with a diameter of 300 mm and a weight exceeding 250 kg, grown by the Cz method. (Reproduced with permission from Siltronic.)

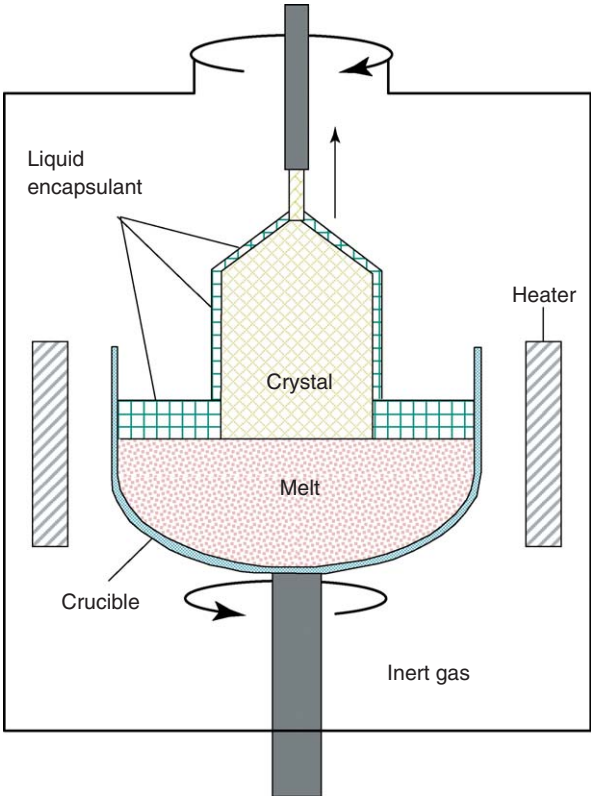


Figure 4 Schematic of the principle of the LEC method.

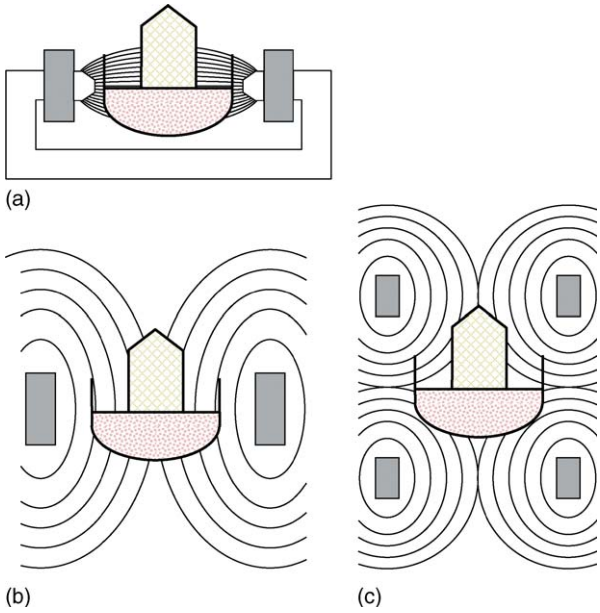


Figure 5 Examples of typical configurations for a magnetic field in the VCz method. (a) Transversal; (b) axial; and (c) cusp. Gray areas are sections of magnetic coils, the curved lines denote the magnetic field lines.

garnets, niobates, tantalates, silicates, vanadates, aluminates, and germanates.

If the material to be grown has a high partial pressure of one or more components, a modified Cz-setup is used (**Figure 4**), which is called the liquid-encapsulated Czochralski (LEC) method. In the LEC-setup a liquid, generally boric oxide, is placed at the top of the melt surface to encapsulate the melt. This liquid layer prevents the loss of a volatile component by evaporation from the melt if the pressure in the growth chamber exceeds the partial pressures of the components.

Today, the LEC-technique is the most widely used growth method for the compound semiconductors GaAs, GaP, InP as well as for PbSe, PbTe. For example, GaAs crystals are industrially produced with diameters of 150 mm and LEC growth of GaAs crystals with 200 mm in diameter is already demonstrated.

However, the low thermal conductivity of the liquid encapsulant causes large temperature gradients and large temperature nonlinearities in the growing crystal. This is unfavorable with respect to the crystal quality as a relatively large number of structural defects (many dislocations) are generated in the crystal compared to other growth techniques.

The high-temperature gradients of the Cz/LEC-method can be reduced by installing an additional inner growth chamber, with hot walls. In this inner growth chamber, a vapor pressure is generated by an extra source to prevent a decomposition of the crystal (the so-called vapor-controlled Czochralski technique (VCz)).

Although the VCz-method is a rather complex technique, it is successfully used to grow GaAs, InP, and BGO crystals with reduced defect densities not only in R&D, but also in industrial production.

The heat and species transport in the melt has a very strong influence on the crystal properties as they are responsible for the uniformity of dopants on the micro- and macro-scale as well as for the shape of the solid–liquid interface and, therefore, for the thermal stress generated in the crystal. During growth, the use of stationary or time-dependent electromagnetic fields enables the control of the flow in electrically conducting melts (e.g., semiconductors). A Lorentz force is generated in the melt, which depends on the magnetic field configuration and leads either to a damping of the flow or to a stimulation of a certain flow pattern.

Today, magnetic fields (**Figure 5**) are used in the industrial production of silicon crystals with 300 mm diameter as well as for growing InP crystals by the LEC and VCz methods.

In order to improve the axial uniformity of the dopant distribution in the grown crystal, the crystal can be grown in such a way that the melt volume is kept constant by supplying the solidified portion from a source. Such a source can be realized either by

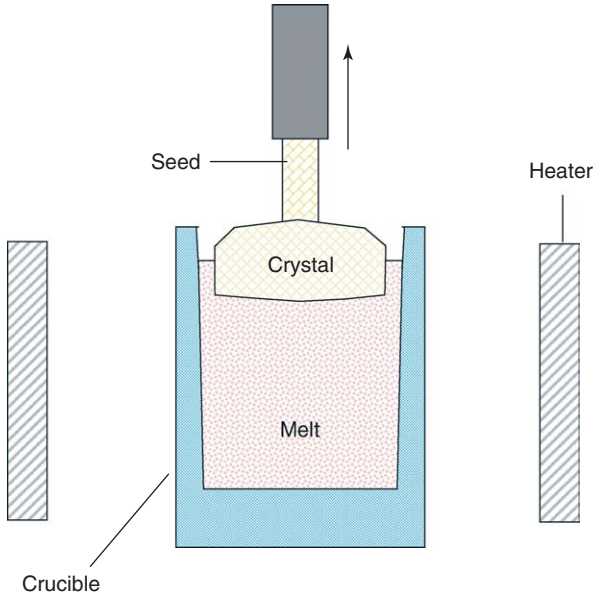


Figure 6 Schematic of the principle of the Nacken–Kyropulous method.

continuously supplying the melt with a feed (continuous Czochralski (CCz) method) or by placing an inner crucible with holes in a larger outer crucible (double crucible Czochralski technique). Although the feasibility of these methods was demonstrated, its use for production is limited owing to technical and material problems preventing a high yield.

Nacken Kyropulous The Kyropulous (or Nacken Kyropulous) method is rather similar to the Cz method. The main difference is that the crystal is not growing at the top of the melt as in the Cz method but is partly immersed in the melt (**Figure 6**). In order to extract efficiently the heat from the growing crystal, usually a seed with a large diameter is used. Growth is achieved by either lowering continuously the heating power or by pulling out the growing crystal carefully from the melt.

This method is widely used to produce alkali halide crystals for optical components (NaCl, KBr) and alkali iodides for scintillators (NaI, CsI). Crystals with dimensions over 500 mm have been achieved.

Shaped crystal growth For a variety of technological applications, crystals of specified size and shape such as plates, rods, tubes, fibers are required. Often it is more cost efficient to grow shaped crystals instead of preparing the needed shape from cylindrical boules as they are grown by, for example, the Cz method.

The diameter of a crystal in the Cz pulling techniques is controlled by the shape and position of the melt meniscus at the triple point crystal-melt-gas (or encapsulant). This control can be improved by using

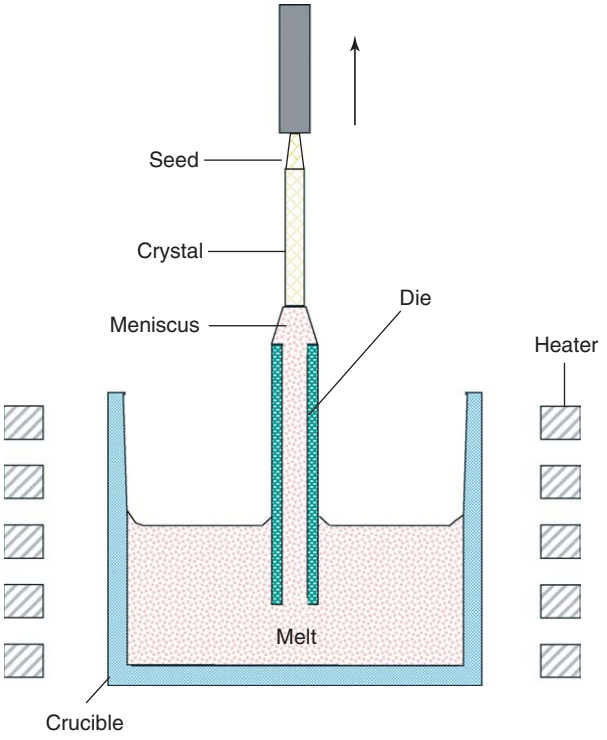


Figure 7 Schematic of the principle of the EFG method.

a die which is in contact to the melt. The melt raises in a narrow channel in the die due to capillary effects. When the seed crystal is dipped onto the melt portion at the top of the die, the meniscus is formed. Liquid solidifies at the seed while the seed is pulled away (**Figure 7**). This causes new liquid to raise up in the die. The heat of crystallization is removed from the solid–liquid interface very efficiently by conduction and by radiation.

The method for growing shaped crystals as described above is commonly known as the Stepanov method (nonwetting die-melt system) or edge-defined film-fed growth (EFG) (wetting die-melt system). In the case of fiber crystals, the method is also called micro pulling down method (μPD) because a crystal fiber with a diameter of only some hundred microns can be grown from a die by pulling the crystal downwards (**Figure 8**). Variants of the EFG method are the noncapillary shaping (NCS) method where the diameter of the die channel is greater than the value of the capillary constant (**Figure 8**). The NCS method offers advantages in order to avoid the generation of bulk micro defects in the growing crystal caused by gaseous or solid inclusions.

Applying a continuous displacement to the growing crystal or the die allows one to grow crystals with more complex shapes (**Figure 8**), for example, domes or hollow cones. This method is named growth from an element of shape (GES).

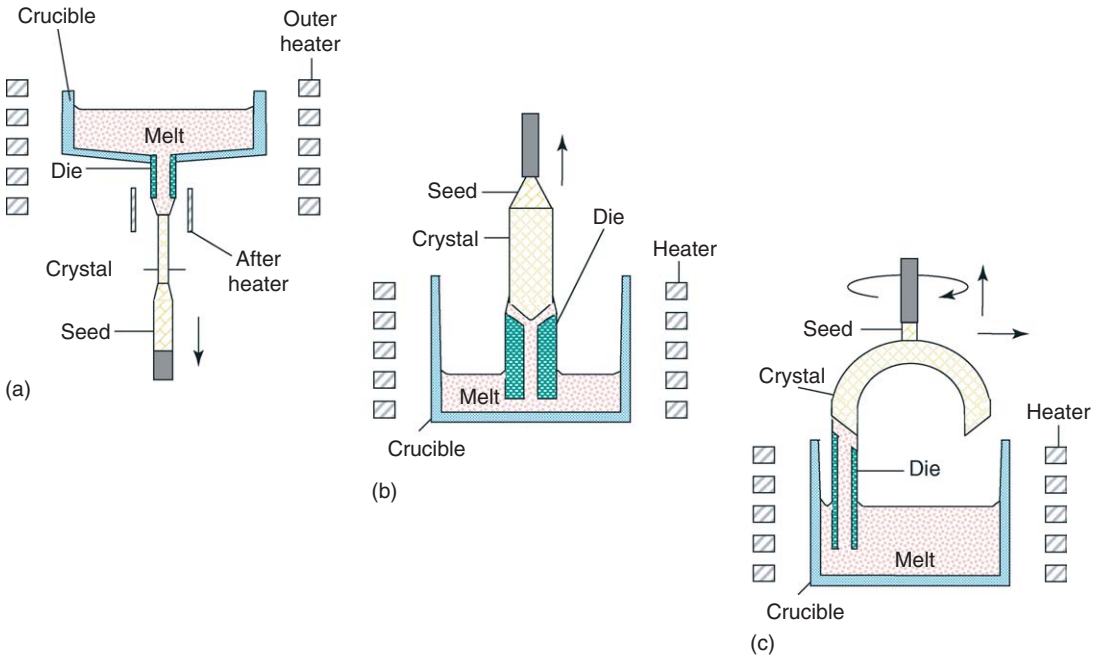


Figure 8 Schematic of the principle of the μPD (a), NCS (b), and GES (c) methods.

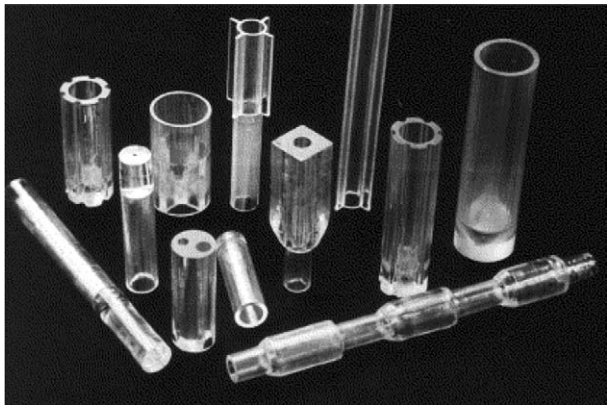


Figure 9 Shaped sapphire crystals grown by the EFG, NCS, and GES techniques. (Reproduced from Kurlov VN (2001) Properties, growth, and applications. In: *Encyclopedia of Materials: Science and Technology* pp. 8259–8265, with permission from Elsevier.)

Nowadays, shaped crystal growth is mainly used for production of silicon and sapphire. Sapphire crystals in the form of ribbons up to 150 mm in width, tubes up to 85 mm in diameter, fibers, near-net-shaped domes (up to 80 mm in diameter), rods of various cross sections, rods with capillary channels, etc., were grown by the shaped growth techniques described above (**Figure 9**).

For photovoltaic applications, the use of near-net shape silicon substrates reduces material losses and process costs by minimizing the efforts for slicing, grinding, and polishing, which is required in bulk-grown silicon. Therefore, the share of thin silicon ribbons or thin wall (octagonal) tubes grown by the EFG method (**Figure 10**) is continuously increasing in the substrate market for solar cells.

Directional Solidification

As against crystal pulling, the growth of single crystals in directional solidification is achieved by melting a charge in a crucible and by controlled freezing of the melt from one end (seed) to the other (tail). Directional solidification is the basic principle of controlled solidification of a melt. It has been used for many materials for a long time and has its greatest importance in the field of metallurgical casting.

Directional solidification has several advantages compared to the Cz methods. It operates usually under stable hydrodynamic conditions, it is well suited for computer modeling and for automatic process control, the crystal grows under low thermal stress conditions, cylindrical crystals are grown without any diameter control, and the equipment is less expensive.

However, a crucial limitation of directional solidification is that a strong interaction between the growing crystal and the crucible material might occur, resulting in crystal imperfections which limits the yield and hence the industrial applicability. Furthermore, the failures generated in the crystal during growth, for example, by twinning or polycrystal formation cannot be observed *in situ* during growth. A correction by re-melting as it is done in the Cz methods is, therefore, not possible.

Gradient freeze and Bridgman–Stockbarger techniques The crystal growth configuration consists typically of a tube furnace which provides a temperature profile with a negative gradient parallel to the growth direction (**Figure 11**). The method can be carried out by moving the growth interface in a horizontal or vertical direction. The single-crystal seed is positioned at one end of the boat or the lower end of the vertical crucible. Then, the crystal is directionally solidified by a controlled shifting of the temperature profile relative to the solid–liquid interface. Three possibilities exist: moving the crucible relative to the fixed furnace (Bridgman–Stockbarger method); moving the furnace relative to the fixed crucible; and without any mechanical movement, only shifting the temperature profile by a controlled change of the heating powers of the furnace (Tamann–Stoeber or gradient freeze method).

Figure 10 Silicon octagons grown by the EFG method. (Reproduced from Kalejs JP (2002) Silicon ribbons and foils – state of the art. *Solar Energy Materials and Solar Cells* 72(1–4): 139–153, with permission from Elsevier.)

Based on one of these principles, one calls the technique the horizontal Bridgman (HB) or the vertical Bridgman (VB) or the vertical gradient freeze (VGF) method. From an industrial point of view, the vertical configurations (VGF, VB) are preferred because they result in a higher yield of round wafers compared to the HB method.

The increasing interest in the use of the directional solidification results from the fact that it uses the simplest principle of melt growth and that the structural perfection of the single crystals with respect to thermal stress and dislocation formation is better than those of crystals produced by Cz methods.

Nowadays, the VGF or VB method is mainly used for the production of compound semiconductors such as high-quality GaAs and InP single crystals to be used as substrates for laser diodes or CdTe crystals for infrared and X-ray detectors. A variety of oxide and fluoride crystals are grown by VGF/VB methods for high-value applications. For example, CaF_2 crystals (**Figure 12**) are utilized as lenses for the deep ultraviolet wavelength in the semiconductor lithography technology. In the field of metallurgy,

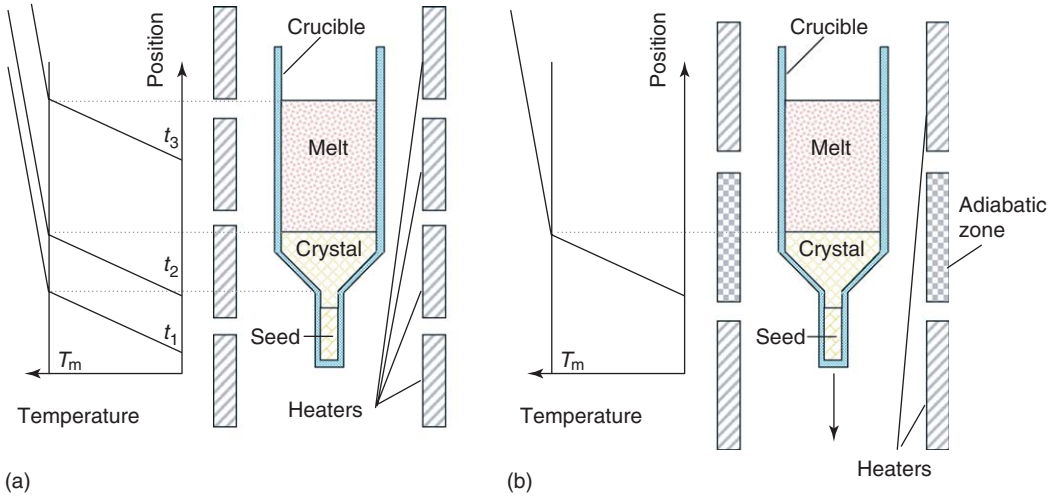


Figure 11 Schematic of the principle of directional solidification methods; (a) VGF and (b) VB method.

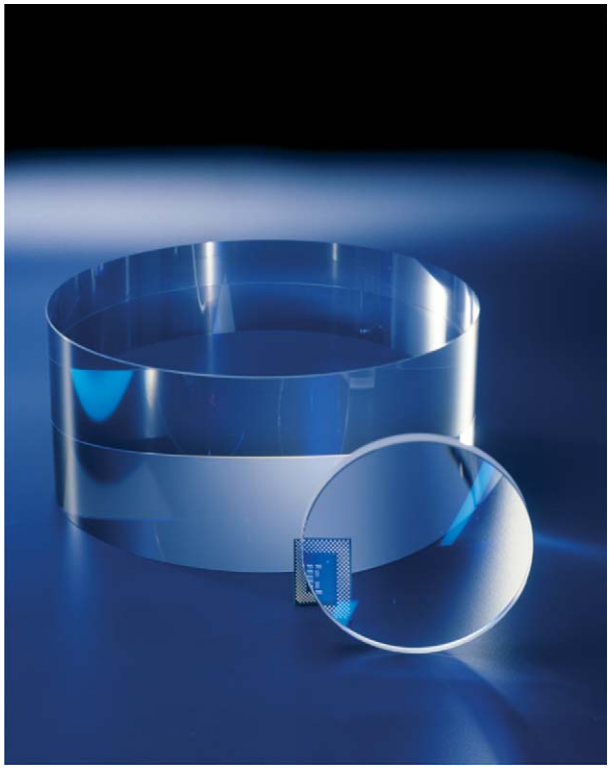


Figure 12 CaF_2 crystals as lense material for the microlithography are grown by modified Bridgman–Stockbarger methods. (Source: SCHOTT Lithotec.)

directional solidification is used to manufacture single crystalline turbine blades made from nickel-based superalloys (**Figure 13**).

The heat exchanger method (HEM) is a modified directional solidification technique. The original characteristic of HEM is that a gas-cooled heat exchanger is placed below the seed-crystal (**Figure 14**). After melting and homogenization, growth is initiated and carried out by increasing the cooling gas flow and decreasing the furnace temperature.

HEM was developed mainly for the growth of large sapphire crystals with diameters up to 500 mm. Several tons of solar-grade polycrystalline silicon as well as CdTe and BGO crystals are also grown by HEM. For these materials, the HEM technology is often fluently merged into the VGF/VB technology.

For refractory materials with high chemical reactivity and melting points above 2000°C, often no suitable crucible material exists. This problem is overcome by the "skull melting" technique, where the material to be grown serves as the crucible. RF heating is used to melt the charge contained in a water-cooled crucible (e.g., copper). As a result of the water cooling, a skull of sintered material is formed which acts as a noncontaminating crucible for the melt. Crystals are typically grown by directional solidification of the molten charge without a seed crystal.

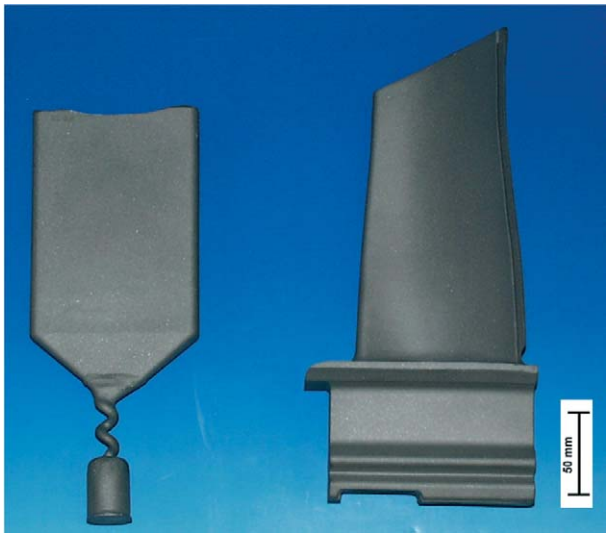


Figure 13 Single crystalline turbine blades made by directional solidification of nickel-based superalloys. (Reproduced with permission from Institute of Science and Technology of Metals, University Erlangen-Nuernberg.)

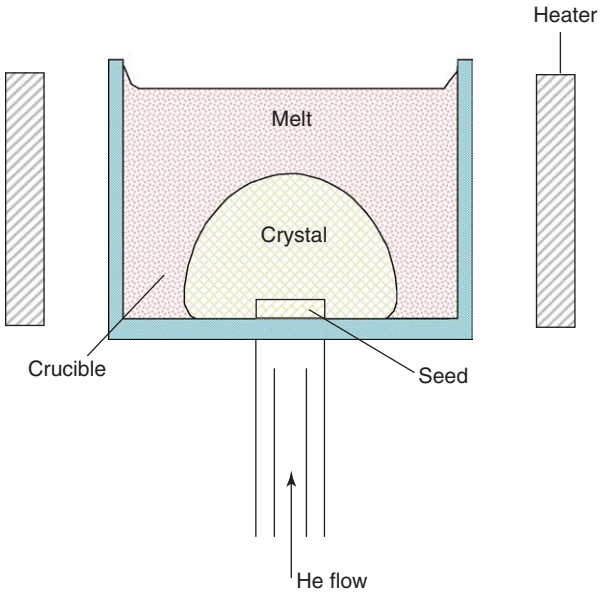


Figure 14 Schematic of the principle of the HEM method.

The method is used, for example, to grow 50–100 tons of cubic ZrO_2 per month for gemstones. Sometimes this technique is also named "cold crucible technology," especially when the melt is levitated by an RF field in such a way that no contact exists between the molten charge and the water-cooled crucible.

One interesting variant of the directional solidification method is the so-called detached solidification (**Figure 15**). In detached solidification, which is also called the de-wetting growth process, a small gap between the growing crystal and the crucible wall

exists. The formation and stability of such a gap depends on the wetting behavior between the melt and the crucible material, the contact angle, the growth angle, and external forces (especially a certain gas pressure difference between the gap and the top of the melt).

The experimental results achieved so far for CdTe, Ge, GeSi, and GaSb on earth and also under microgravity demonstrate that the structural perfection of the crystals is generally improved when the dewetting effect occurs. However, there still remains a lot of research to be done until this technology can be transferred into industrial production.

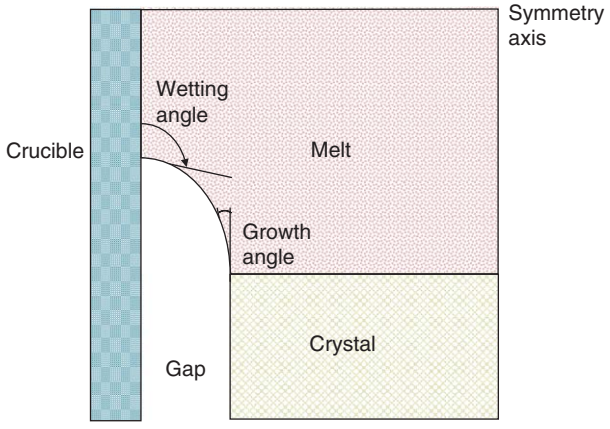


Figure 15 Schematic of the principle of detached solidification.

Zone melting in crucibles In zone melting only a part of the solid is melted in a crucible or a boat by using a resistance or induction zone heater (**Figure 16**). Growth is achieved by moving the heat source relatively to the axis of the container. Nowadays, the importance of zone melting for bulk crystal growth is mainly limited to its use for purification of the feed material. Impurities with a segregation coefficient $k_0 < 1$ can be accumulated very efficiently by multiple passes of melt zones toward the end of the ingot.

Floating Zone (Crucible-Free Zone Melting)

The floating zone (FZ) technique is a crucible-free crystal growth method. In FZ growth, the molten zone is kept between two vertical solid rods by its own surface tension (**Figure 17**). A single crystal is grown by dipping a seed crystal into one end of the zone and translating the molten zone toward the feed stock. The main advantage of the FZ technique is the absence of the container, which precludes a contamination by the crucible material and the generation of crystal defects caused by the interaction between the growing crystal and the container. Therefore, the technique is especially used for highly reactive materials, intermetallic compounds, and refractory materials. The heating of the molten zone can be achieved by several methods including RF induction, optical, electron beam, laser and resistance heating.

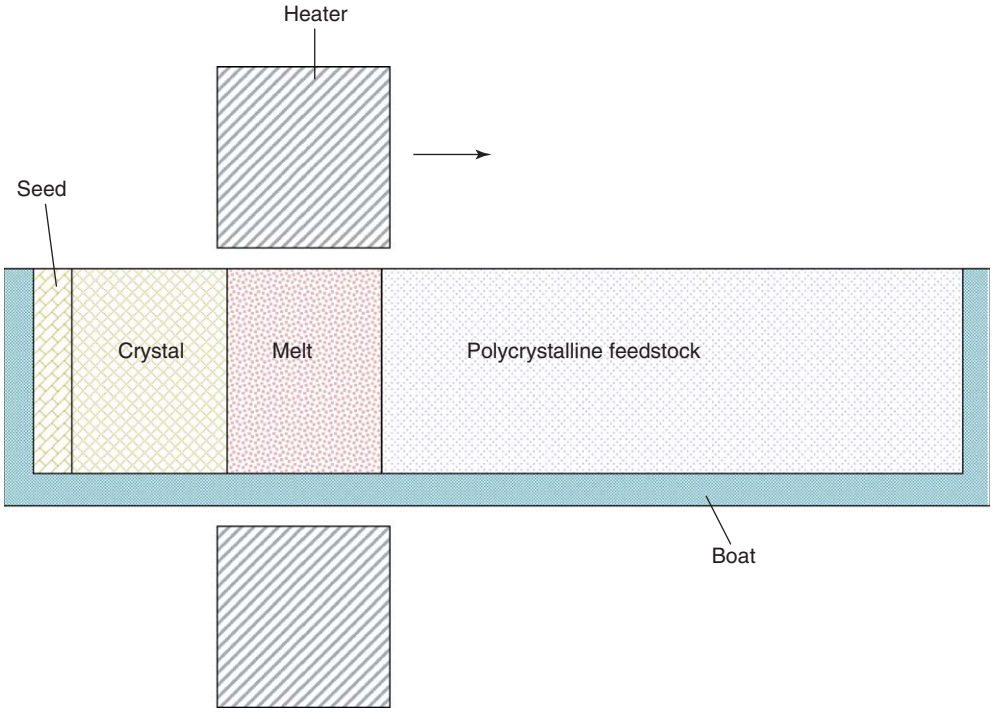


Figure 16 Schematic of the principle of zone melting.

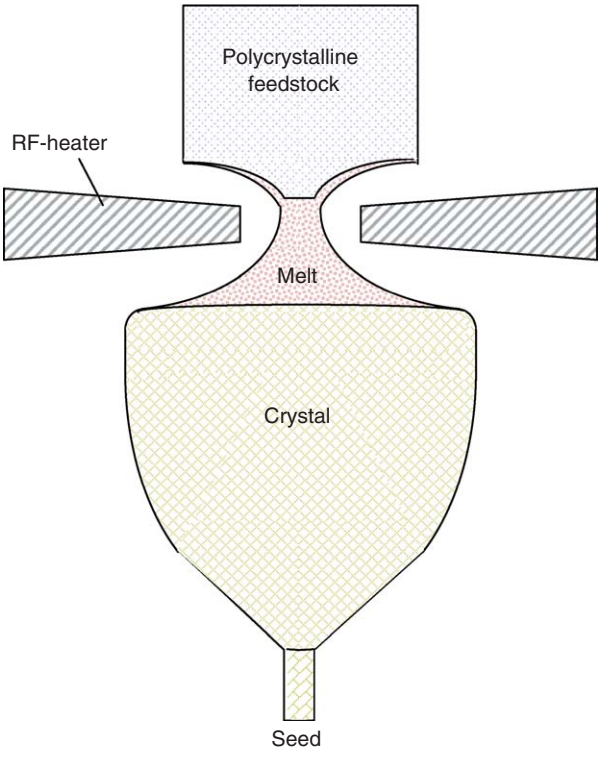


Figure 17 Schematic of the principle of the FZ method for growing Si crystals.

Figure 18 GaAs crystals grown by the FZ technique; left: sample grown in space under microgravity during the German Spacelab mission D2, right: sample grown on earth with maximum size. (Source: Crystal Growth Laboratory, Fraunhofer IISB, Erlangen.)

The industrial application of the FZ technique is rather limited because the maximum sample diameter is relatively small due to the difficulties in maintaining a stable molten zone. Usually the maximum height of the molten zone is a few millimeters, which is determined by the ratio between the surface tension and the density of the melt.

These difficulties were overcome in the industrial FZ growth of silicon by using the needle-eye technique. In the needle-eye technique, the shape of the molten zone is stabilized by the electromagnetic pressure field generated with a specially shaped RF-induction coil. The silicon crystals produced by this FZ technique have diameters of up to 200 mm and extremely low oxygen contamination ($< 10^{16}\,\mathrm{cm}^{-3}$). Such pure silicon substrates are needed for power electronic devices.

Crystals of other materials with higher density and lower surface tension, for example, GaAs can only be grown with larger diameters under microgravity conditions (**Figure 18**).

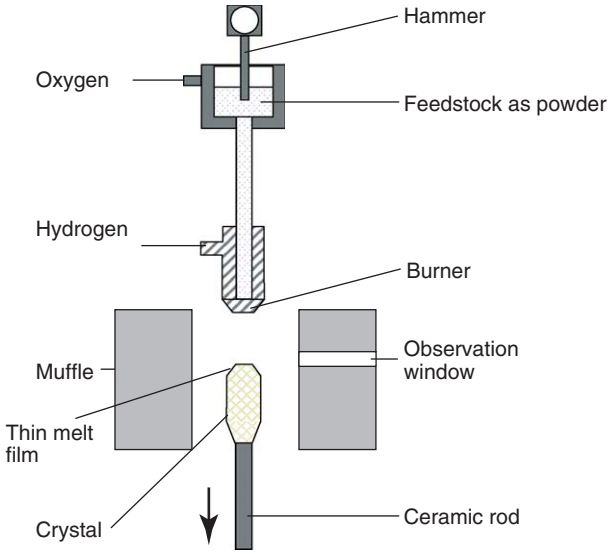


Figure 19 Schematic of the principle of the Verneuil method.

Verneuil Technique

The Verneuil technique (**Figure 19**), also called the flame fusion technique, is one of the oldest methods for growing refractory oxide crystals. The feed is a fine powder which is fused by an oxygen–hydrogen flame burner while falling down toward the growing crystal. A thin melt film is formed on top of the growing crystal, which is slowly pulled down. It is important to balance the rate of powder feed and the rate of pulling to maintain a constant growth rate

and crystal diameter. Otherwise, the thermal field in the vicinity of the solid–liquid interface is changed which results in poor crystal quality.

The advantage of this process is the low equipment costs. The main disadvantage is the extremely large temperature gradient causing a high density of stress-induced crystal defects. Therefore, often an after heater is used in order to prevent the crystal from cracking.

Nowadays, the Verneuil technique is exploited for the economic mass production of sapphire and ruby used in jewelry and precision bearings ($\sim$700 tons per year). Most high-melting oxide compounds can be grown using this technique.

Solution Growth

Opposite to melt growth where the crystal solidifies from its own melt, in solution growth the elements or compounds of the material to be crystallized are dissolved in a suitable solvent.

Principle

For crystal growth, the solution has to be supersaturated. This can be achieved by different means:

1. In the temperature changing technique, the solution is supersaturated by slow cooling as the solubility is usually decreasing with decreasing temperature (supersaturation is achieved by slow heating, if the solubility decreases with increased temperature)
2. In the evaporation technique, the supersaturation is obtained by controlled evaporating of the solvent at a constant temperature.
3. In the temperature gradient technique, two regions of different temperatures $T1$ and $T2$ are established. At the higher temperature $T1$, the material is dissolved, while at lower temperature $T2$ crystal growth takes place.

The growth rate is mostly limited by the transport rate of the solute to the growth interface. Therefore, an active mixing by an "accelerated crucible rotation technique" is used.

An important issue in solution growth is the selection of a suitable solvent. At the growth temperature, the solvent should generally have a sufficient solubility of the material to be grown as well as a low vapor pressure and low viscosity.

Low-Temperature Solution Growth

Low-temperature solution growth is one of the simplest methods to grow single crystals especially if aqueous solutions are used (**Figure 20**). Industrially important crystals grown by low-temperature solution growth are potassium dihydrogen phosphate (KDP, **Figure 21**) ammonium dihydrogen phosphate (ADP), both for electrooptic applications, and, for example, triglycine sulfate (TGS) for IR imaging applications.

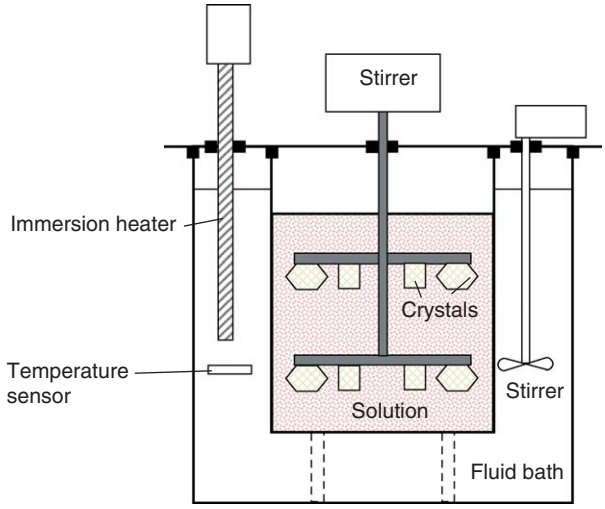


Figure 20 Apparatus for crystal growth by the temperature changing method using a fluid bath. (Reproduced from Sangwal K (2001) Growth from solutions. In: *Encyclopedia of Materials: Science and Technology*, pp. 3670–3680, with permission from Elsevier.)

Figure 21 KDP single crystal grown from aqueous solution. (Source: CEA.)

High-Temperature Solution Growth

Practically any material can be grown by high-temperature solution growth (often also called flux growth). Mainly low-melting temperature solvents

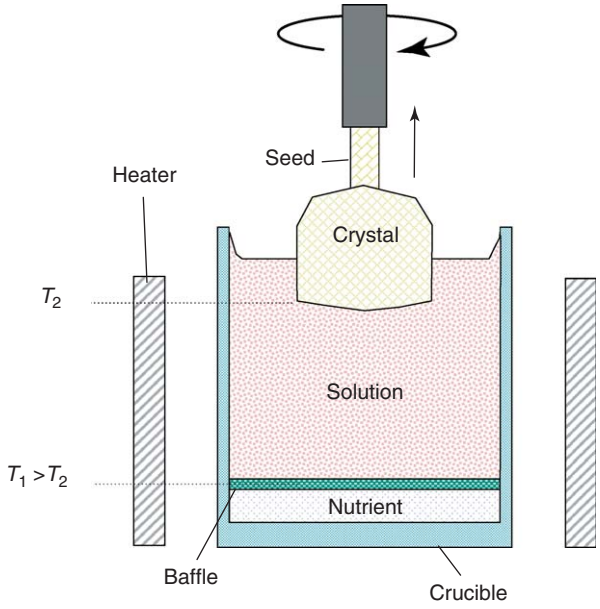


Figure 22 Typical equipment for the temperature gradient technique with top seeding (TSSG).

(fluxes), molten salts, and oxides are used. This method can be applied to congruent and non-congruent melting materials. Solution growth is also the preferable growth technique, if the material undergoes a destructive solid–solid phase transition below its melting point. It is also applied as liquid phase epitaxy to produce epitaxial layers on substrates. However, the growth rates are small compared to melt growth techniques and sometimes the solvent contaminates the crystal.

For a variety of technically important materials such as $Sr_{1-x}Ba_xNb_2O_6$ (SBN), $Pb_{1-x}Ba_xNb_2O_6$ (PBN), $LiNbO_3$, $BaTiO_3$, BBO, the so-called top-seeded solution growth method (TSSG) is used (**Figure 22**). TSSG corresponds to the CZ method; the only difference is that a solution is used instead of the melt.

In the field of semiconductor crystals, the use of solution growth is limited. In research laboratories the solution growth of ZnSe, ZnS, ZnTe, as well as of SiC and GaN is investigated.

Hydrothermal Growth

Hydrothermal growth refers usually to heterogeneous reactions in aqueous solvents or mineralizers above 100°C and under high pressure (several kbars). In hydrothermal growth, the high-pressure conditions are needed to enhance the solubility of the solute in the solvent. A precisely controlled temperature gradient is used to achieve the crystallization at the seed location (principle (3.) in the section "Principle").

Hydrothermal growth is extensively exploited to grow synthetic quartz (SiO_2) crystals. It is carried out in special vessels also called autoclave or bomb under pressures of typically 170 MPa and around 400°C. This technique has also been used to grow berlinite ($AlPO_4$), calcite ($CaCO_3$), and zincite (ZnO) crystals.

High-Pressure Synthesis

High-pressure synthesis refers to the growth of diamond crystals. In this method, anvils generate a pressure of around 60 kbar in the heated growth cell. The growth cell contains graphite mixed with a metal or alloy that serves as a solvent/catalyst. At around 1500°C diamond is the stable modification whereas graphite becomes meta-stable and is partially dissolved in the solvent. The solubility difference acts as a driving force for the conversion of carbon from graphite to diamond crystals in the solvent.

Vapor Growth

Vapor growth is like solution growth usually applied when melt growth is impractical. Due to the lower temperatures as compared to melt growth, many thermally activated processes like impurity incorporation, compositional uniformity, structural imperfection, are usually decelerated. Therefore, the crystal quality is enhanced. However, in vapor growth the growth rates are usually also considerably decreased due to the low density of crystal material in the gas phase, the low transport rate of the vapor to the growth region, and the decreased interface kinetics as the temperature is decreased.

Sublimation Technique

In physical vapor transport (PVT) often called as sublimation growth, a source material held at a temperature $T1$ sublimates, and its vapor is transported by diffusion and convection to the seed crystal held at a lower temperature $T2$ where it can crystallize. The supersaturation at the crystallization front depends on the difference of the partial pressures of the sublimated material in the source and growth region as well as on the total vapor pressure in the system. The supersaturation has to be precisely controlled in order to avoid parasitic nucleation.

Nowadays, the sublimation technique is industrially used to produce silicon carbide crystals with diameters of up to 100 mm (**Figure 23**). Typically, the growth takes place in an inductively heated graphite crucible at elevated temperatures ($T = 2100$–2400°C). A transport of Si, Si_2C, and SiC_2 gas species is established from the SiC source to the growing SiC

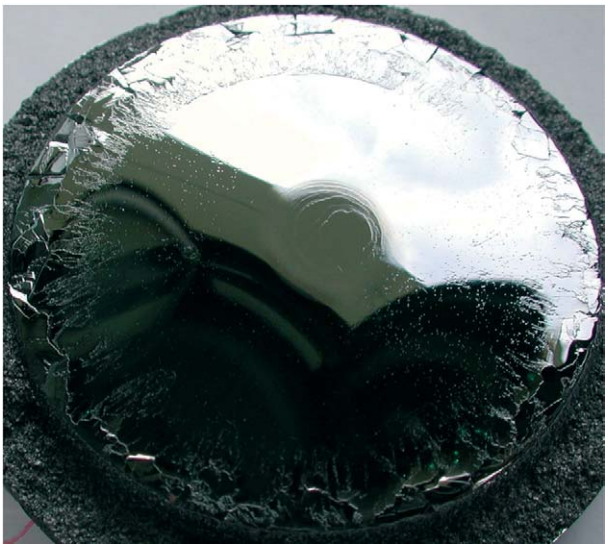


Figure 23 SiC crystal with 50 mm diameter grown by the PVT method. (Reproduced with permission from Institute for Material Science, Department of Electrical Engineering Materials, University Erlangen-Nuernberg.)

single crystal in an Argon ambient atmosphere ($p = 10$–100 mbar).

Also aluminum nitride is grown by PVT methods at temperatures above 2000°C. The PVT method is also used for II–VI compounds.

Chemical Vapor Transport and Chemical Vapor Deposition

In the chemical vapor transport (CVT) method, a reactive gas is used which reacts with the source material or gaseous chemical compounds containing the crystal components. This source material is transported as gaseous species to the growth zone of the reactor. Here, the reactions take place, forming the crystal components which are deposited at the growth interface.

CVT has less importance in bulk crystal growth, but is the dominant method for producing thin films. In bulk crystal growth, CVT is applied only on a laboratory level to grow a variety of materials especially chalcogenides. Nowadays, hydride vapor phase epitaxy (HVPE) as a variant of the CVT principle is used to grow thick free-standing gallium nitride substrates (**Figure 24**).

Conclusions

Bulk growth of inorganic crystals has become an important industrial key technology. Although methods to grow any inorganic crystal material are known in principle, continuous R&D efforts are necessary for a further development of the processes

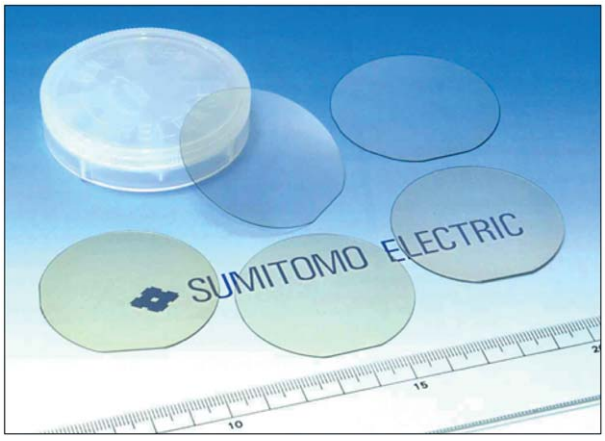


Figure 24 GaN substrates with 2'' diameter grown by the HVPE method. (Reproduced with permission from Sumitomo Electric Industries, Ltd.)

to fulfill the demands on crystal properties coming from the various fields of application. Among the most important R&D tasks are:

- up-scaling of the growth systems in order to grow crystals of larger dimension for higher productivity and yield,
- avoiding of crystal defects which are deleterious for the performance of devices made from the crystals,
- better control of crucial processing parameters in order to tailor the physicochemical properties of the crystals according to the needs of their applications, and
- an increased uniformity of the relevant crystal properties in the micro- and macroscale.

See also: Crystal Growth, Bulk: Theory and Models; Film Growth and Epitaxy: Methods; Solidification: Models and Simulations; Vapor Transport Processing: Mechanisms; Vapor Transport Processing: Models and Simulations.

PACS: 02.60.Cb; 02.70.−c; 07.35.+k, 44.; 47.27.−l; 47.65.+a; 61.; 61.72.−y; 64.70.−p; 64.75.+g; 68.08.−p; 81.05.−t; 81.10.Aj; 81.10.Bk; 81.10.Dn; 81.10.Fq; 81.10.Jt; 81.10.Mx; 81.30.Fb

Further Reading

Chernov AA (1999) *Modern Crystallography III*, 2nd edn. Berlin: Springer.

Hurle DTJ (1993–94) *Handbook of Crystal Growth*, vols. 1–3. Amsterdam: Elsevier.

Müller G (1988) Convection and crystal growth from melt. In: Freyhardt HC (ed.) *Crystals*, vol. 12. Berlin: Springer.

Müller G (1998) Melt growth of semiconductors. In: Fornari R Paorici C (eds.) *Theoretical and Technological Aspects of Crystal Growth*, vol. 276–277, pp. 87–108. Materials Science Forum.

Müller G and Rudolph P (2001) Crystal from melt growth. *Encyclopedia of Materials: Science and Technology*, pp. 1866–1873. Amsterdam: Elsevier.
Pfann WG (1958) *Zone Melting*. New York: Wiley.
Rosenberger F (1979) *Fundamentals of Crystal Growth I*. Berlin: Springer.
Rudolph P (1998) Elements of thermodynamics for the understanding and design of crystal growth processes. In: Fornari R Paorici C (eds.) *Theoretical and Technological Aspects of Crystal Growth*, vol. 276–277, pp. 1–26. Materials Science Forum.
Wilke KTh and Bohm J (1988) *Kristallzüchtung*. Germany: Deutsch & Thun, Frankfurt am Main.

Crystal Growth, Bulk: Theory and Models

J J Derby, University of Minnesota, Minneapolis, MN, USA

Introduction

The growth of bulk crystalline materials remains one of the most challenging and astonishing technical endeavors of materials processing. For example, electronic-grade silicon grown by the Czochralski method is one of the purest and most perfect materials ever produced by mankind. Current production technology routinely achieves impurity levels of less than parts per billion in single-crystal ingots of up to 300 mm in diameter and over 200 kg in mass, and these crystals are completely free of dislocations. Indeed, it is even possible today to control the distribution of micro-defects, such as voids, in bulk silicon crystals. Advances in the production technology of bulk crystal growth to current levels would have been impossible without the application of theory and modeling to understand and improve these processes.

While silicon remains the foremost in economic importance, many other materials are produced in the form of bulk single crystals. Most of these are semiconductors or oxides and used as substrates for the fabrication of electronic, optical, or optoelectronic devices. The most common means for the production of these bulk crystals are melt-growth and solution-growth techniques. For more detailed information, one should consult the excellent overview on bulk crystal growth by Müller and Friedrich in this encyclopedia. Vapor, plasma, or molecular beam techniques, because of their slower growth rates, are rarely employed for the production of bulk crystals, though these methods are extremely important for the growth of thin films needed for device fabrication. Topics of film growth and epitaxy are covered elsewhere in this encyclopedia.

This article presents an overview of the theory and modeling of bulk crystal growth. It must be emphasized from the start that crystal growth is a fascinating, yet complex process that depends on a host of coupled physical phenomena ranging from thermodynamics to kinetics. While many behaviors can be described by relatively simple theories, there is still much that remains to be understood. Modeling, which can be broadly described as applied theory, is usefully employed for clarifying many aspects of crystal growth. However, the extent to which modeling can realistically represent bulk crystal growth is also limited, mainly due to the challenges in representing the many phenomena occurring at disparate length and timescales, and the limits of current computers and numerical algorithms. Nevertheless, advances in theory and modeling will continue to aid important advances in crystal growth technology.

Crystal Growth Theory

Crystal growth theory consists of mechanistic models of how and why a crystal grows from a fluid phase (either liquid or vapor). There are many ways that crystals of different materials grow under different conditions, and no single simple theory describes all of these behaviors. The ensuing discussion highlights some of the more useful theories for the description of the growth of bulk single crystals, starting first with thermodynamic concepts, followed by issues involving kinetics. Also, since bulk crystals are almost always grown from an existing seed crystal, crystal nucleation phenomena will not be addressed here.

Thermodynamics

The formation of a crystalline solid from a fluid is a first-order phase transformation, accompanied by discontinuous changes in enthalpy, entropy, and specific volume. However, under equilibrium conditions, the Gibbs energy is continuous from the solid to the fluid phases. Crystal growth occurs under nonequilibrium conditions, when the thermodynamically stable crystal phase coexists with a thermodynamically metastable phase, typically a supersaturated or supercooled fluid. The driving force for transition to the stable, crystalline phase is the difference in Gibbs energy between the two phases. All crystal growth

processes are designed to set up and maintain a precise driving force, ΔG, for crystallization, where G denotes the Gibbs energy associated with each phase. This driving force takes the form of an undercooling in temperature, ΔT, in melt growth and a supersaturation σ, or compositional difference, in solution growth. At the larger length scales involved with bulk crystal growth, the crystallization driving force is strongly impacted by the kinetics associated with energy and materials transport by diffusion and convection via fluid flow. The effects of continuum transport will be discussed in more detail in ensuing sections.

A solid crystal is separated from the surrounding fluid phase by a phase boundary or interface. The ordered crystalline structure of the solid, in contrast to the nearly disordered structure of the fluid, leads to interfacial energetic effects which are highly anisotropic and a function of interface orientation. Typically, these energetic differences are small compared to the crystallization driving force between the bulk phases; however, surface effects are important in many situations. Indeed, it is these energetics that prescribe the equilibrium shape that a crystal exhibits (via a shape minimization of the Gibbs energy of the crystal).

Under many situations, the surface shape which minimizes the Gibbs energy corresponds to an atomically smooth plane associated with the underlying crystalline structure and referred to as a singular face or facet. Under conditions not far from equilibrium, growth is characterized by how the addition of atoms to the surface affects the Gibbs energy of the surface. Often, growth occurs layer by layer, leading to faceted crystal shapes. The kinetics of mechanisms involved in layer growth are discussed in the ensuing section.

Some single crystals, especially metals grown form the melt, do not exhibit faceted interfaces under equilibrium conditions. This behavior is explained by a phenomenon known as an interface-roughening transition. The basic idea involves the change in Gibbs energy of an initially flat crystal surface with the addition of atoms during growth. There is an enthalpic driving force to add atoms to the surface, as they form new bonds to the surface and to each other. However, if there is a large change in entropy between the crystal and fluid phases, the Gibbs energy is lowest when there are just a few extra atoms on the surface or a few atoms missing in the plane below. Under these conditions, the surface is smooth, and the crystal exhibits a faceted shape; see **Figure 1a**. On the other hand, if the change in entropy between phases is small enough, as is common for most metals, then the Gibbs energy of the surface is minimized when half of the available surface sites are filled, resulting in a rough interface and a crystal shape which may be arbitrary; see **Figure 1b**. For melt growth, the Gibbs energy depends only on the temperature, and the nature of the interface is set only by the properties of the material. For solution growth, the Gibbs energy depends on temperature and composition, so it is possible to change the interface from smooth to rough, for example, by increasing the temperature of the system while maintaining a crystallization driving force in supersaturation.

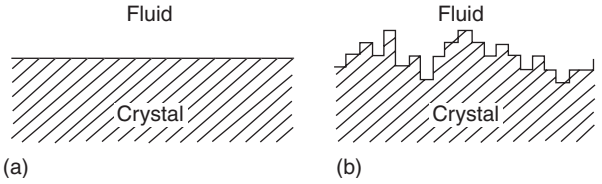


Figure 1 Schematic diagram of the limiting behaviors of a fluid–crystal interface: (a) smooth surface and (b) rough surface.

A final topic involving the thermodynamics of crystal growth is that of lattice defects arising during growth. Again, the basic idea goes back to thermodynamics and the minimization of the Gibbs energy for a crystalline material. If one considers a perfect crystal, the process of ordering of the atoms into a perfect lattice is characterized by a minimization of the enthalpy of the system. However, the total entropy of the system will be increased by disorder, that is, by defects in the perfect lattice. Since the Gibbs energy, G, contains both enthalpic and entropic components, that is, $G = H - TS$, where H is enthalpy, T is absolute temperature, and S is entropy, a minimization of G may be accomplished by some amount of disorder, becoming more important with increasing T. In a pure material, this disordering is manifested by the defective incorporation of atoms during growth, leading to point defects (either vacancies or interstitials) in the resulting crystal. Since the equilibrium of point defects is strongly dependent upon temperature, the details of the cool-down process of the crystal is very important in determining their ultimate fate during the growth of high-melting-point crystals, such as silicon. This relationship is much more complicated for multicomponent crystals, such as compound semiconductors. The interested reader is referred to the "Further Reading" section for more extensive discussions.

Kinetics

In addition to thermodynamic considerations, kinetic factors are always important in the growth of bulk crystals. It is useful to consider two limiting situations

according to the structure of the crystal surface, as depicted in **Figure 1**. If the crystal surface is atomically smooth, the crystal grows layer by layer and is limited by interfacial kinetics. On the other hand, if the surface is atomically rough, the incorporation of atoms into the crystal is very fast and the growth behavior is largely determined by the continuum transport of mass and energy. While real crystallization behavior is much more complicated, these two scenarios can be usefully employed to model the growth of many bulk crystals.

A simple theoretical model that gives insight to growth mechanisms is the Kossel–Stranski model, depicted in **Figure 2**, which assumes that a crystal is a simple cubic lattice comprised of atoms, or growth units. Each lattice site is either empty or occupied, and empty sites are forbidden below occupied sites. Growth then occurs by the addition of atoms to preexisting surfaces of the crystal. The change in the Gibbs energy of the surface by the addition of a growth unit is determined by its position relative to the other atoms, with the lowest surface energy corresponding to an atomically flat, singular surface, or facet. This is also often referred to as a Terrace–Ledge–Kink model for crystal growth.

The rate of growth of the crystal is determined by how fast new layers can be added. For a singular surface, the rate-determining step is the two-dimensional nucleation of a new layer, which is typically slow compared to transport. For a surface orientation far away from a facet direction, there are many steps along the surface so that the incorporation of atoms into the available steps is fast, and the growth of the crystal normal to the surface is limited by the rate of transport from the fluid phase to the surface. A general outcome of this mechanism is that the growth of any crystal is anisotropic and that the slowest growth will occur in directions normal to singular faces. If this is true and there are no transport limitations, then a growing crystal will eventually be bounded by the slowest growing faces, leading to a crystal with a faceted shape, which is often referred to as its growth habit.

However, there is a problem with the above theory,. If two-dimensional nucleation must occur to start a new layer on a singular surface, far higher driving forces than observed would be needed to obtain the growth rates obtained in experiments. Burton and Cabrera postulated that a near-singular, or vicinal, surface, such as depicted in **Figure 3a**, would contain enough steps to rectify this large discrepancy between theory and observation. At the same time, Frank noted that dislocations are present in nearly all crystals and that a screw dislocation intersecting a face would

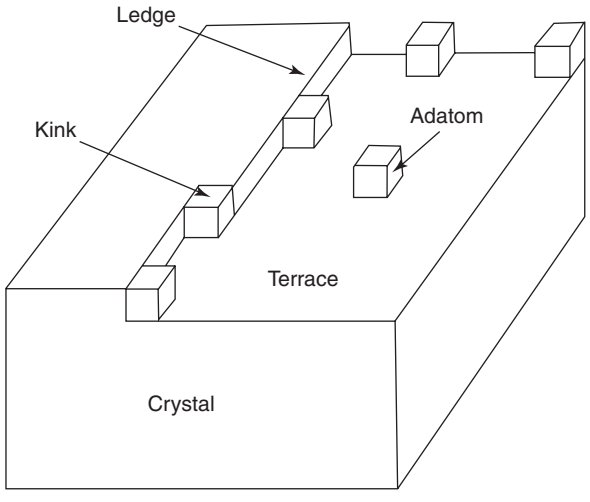


Figure 2 The Kossel–Stranski model depicts a crystal as a simple cubic lattice. Under growth conditions, adatoms adsorb to a terrace, then diffuse and incorporate into lower-energy ledge or kink sites.

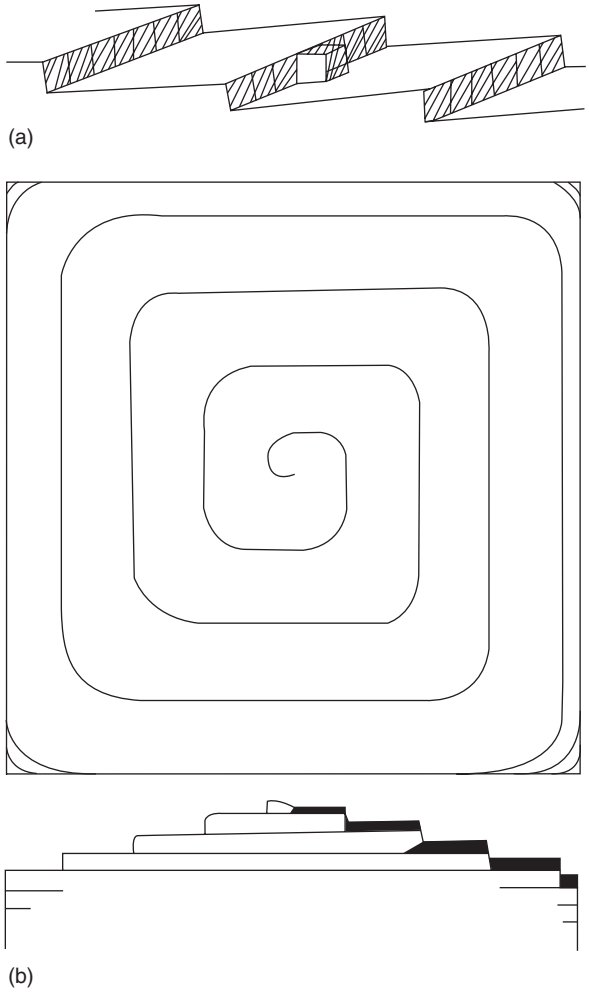


Figure 3 Schematic depictions of (a) steps along a vicinal surface and (b) steps around a screw dislocation.

act as a continuous source of steps, as indicated in **Figure 3b**. Burton–Cabrera–Frank, or BCF, theory describes the growth of singular crystal faces as a mechanism of step generation and movement.

The notion of rough versus smooth interfaces must also be modified to account for kinetic factors. Notably, if the driving force for crystallization is sufficiently high, a two-dimensional nucleation of new layers on a singular surface can occur fast and often enough to make the surface atomically rough. This phenomenon is known as kinetic roughening and is observed in vapor and molecular beam growth. It is much less likely to be observed during melt growth, because of the very high undercooling that would be required, or in bulk solution growth, since there are usually step sources available to allow the BCF mechanism of growth.

These theories give a qualitatively correct picture of the mechanisms behind crystal growth, although real behavior is much more complicated. The rather simple theories discussed above do not consider such effects as clustering of atoms on the surface, roughening to several layers within the crystal, the ordering of a liquid preceding the crystal interface, fluctuations, surface reconstruction, impurities or other chemical interactions involving noncrystallizing species, along with many others. There are still challenges for the theory of crystal growth.

Crystal Growth Modeling

Crystal growth modeling involves the construction of a mathematical description of a crystal growth system, the solution of the governing equations of the model, and the interpretation of the modeling results. The challenge for modeling bulk crystal growth is combining an appropriate description of crystallization behavior with a proper description of continuum transport in the bulk. The crystallization interface is part of a self-consistent solution to a mathematical moving-boundary problem, which also poses special challenges. Below, brief discussions of governing equations and solution techniques are presented for modeling continuum transport and the crystal interface in bulk crystal growth.

Continuum Transport

The transport of heat, mass, and momentum is important in bulk crystal growth processes. Governing equations are written to describe the conservation of these quantities within the solid crystal and the accompanying fluid phase. A brief accounting of these phenomena and their governing equations is given below.

Flow in the fluid phase is especially important for the transport of heat and mass via convection in bulk crystal growth systems. For liquids, such as the molten phase in melt growth or the solution phase in solution growth, flows are described by conservation equations written for momentum and continuity of an incompressible, Newtonian fluid with the application of the Boussinesq approximation to describe fluid density changes:

$$\rho\left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\nabla\boldsymbol{v}\right) = -\nabla p + \mu\nabla^2\boldsymbol{v} + \rho_0\boldsymbol{g}[1 - \beta_T(T - T_0) - \beta_s(c - c_0)] + \boldsymbol{F}(\boldsymbol{v}, \boldsymbol{x}, t) \quad [1]$$

where ρ is the density of the fluid, $\boldsymbol{v}$ is the velocity field, t is time, ∇ is the gradient operator, p is the pressure field, μ is the fluid viscosity, $\boldsymbol{g}$ is the gravitational vector, β_T and β_s are the thermal and solutal expansivities, respectively, T is temperature, c is concentration, the subscript zero denotes the reference state about which the linear dependence of the density is approximated, and $\boldsymbol{F}(\boldsymbol{v}, \boldsymbol{x}, t)$ is an additional body force, for example, that which results from the application of a magnetic field to a conducting fluid. The requirement for continuity of the fluid phase takes on the following form

$$\nabla\cdot\boldsymbol{v} = 0 \quad [2]$$

which states that the divergence of the velocity field must be zero everywhere. Collectively, these two expressions constitute the celebrated Navier–Stokes equations.

Flows driven by buoyancy, referred to as natural convection, are important in all bulk crystal growth systems due to thermal or solutal gradients. In addition, temperature or compositional gradients along a liquid–gas interface can drive very strong flows arising from the variation of surface tension along the surface. Often referred to as Marangoni flows, these are important in many meniscus-defined melt-growth systems, such as the Czochralski and floating zone methods. Forced convection flows are driven by applied rotation of crystal or crucible or, in solution-growth systems, by pumping or stirring mechanisms. Also of great importance in many bulk growth systems are time-dependent or turbulent flows, which arise naturally if the driving forces are strong enough. These flows dramatically affect the nature of continuum transport to the crystal interface. The application of strong stationary or rotating magnetic fields has been applied to electrically-conductive melts in attempts to control such flows.

The transport of heat and mass can be determined from the solution of the appropriate governing

equations. The energy balance equation is given by the following expression:

$$\rho C_p \frac{\partial T}{\partial t} + \rho C_p \boldsymbol{v} \cdot \nabla T = \kappa \nabla^2 T \qquad [3]$$

where C_p is the heat capacity and κ is the thermal conductivity of the fluid. For media which possess some transparency to infrared radiation, the transport of energy via internal radiative transport may also be important in high-temperature crystal growth systems, and an extra term must be added to the above equation. The exchange of thermal radiation among surfaces is also very important in many high-temperature growth systems. These radiative processes are described by terms which are strongly nonlinear in temperature, thus posing significant modeling challenges. Heat transfer is also very strongly influenced by thermal boundary conditions, that is, the conditions imposed on the crystal and fluid by the system design. The modeling of heat transfer in high-temperature melt growth furnaces is itself a significant technical challenge due to complicated geometries and radiation heat transfer.

The conservation equation for a dilute species in a fluid is given by

$$\frac{\partial c}{\partial t} + \boldsymbol{v} \cdot \nabla c = \mathscr{D} \nabla^2 c \qquad [4]$$

where $\mathscr{D}$ is the diffusion coefficient of the species. Mass transfer in bulk crystal growth systems is largely determined by the interactions of diffusion and convection near the liquid–crystal interface. At a growing crystal surface, the equilibrium partitioning of a solute between solid and fluid phases coupled with diffusion and convection results in segregation, that is, the inhomogeneous distribution of a solute in a grown crystal. These phenomena are discussed in more detail in ensuing sections.

The Crystal Interface

The manner in which the crystal interface is represented is a central feature of bulk crystal growth models. A self-consistent growth model requires that the interface geometry be computed as part of the solution to the transport problem, that is, as a free or moving boundary. For the case of melt growth, a mathematical expression of the normal growth velocity of the interface, V_g is given by

$$V_g = \beta \Delta T \qquad [5]$$

where β denotes a kinetic coefficient and ΔT is the driving force for crystallization,

$$\Delta T = T_i - T_m\left(1 - \frac{\gamma}{L}\mathscr{H}\right) \qquad [6]$$

where T_i represents the interface temperature, T_m is the melting temperature of a planar interface, γ is a capillary coefficient, L is the latent heat of solidification, and $\mathscr{H}$ depicts the local mean curvature of the interface. For an atomically rough interface, the kinetic coefficient becomes large enough so that the undercooling ΔT goes to zero and $T_i = T_m(1 - \gamma\mathscr{H}/L)$. Here, the rate of interface movement is controlled by the flow of latent heat away from the interface. In many bulk melt-growth systems, the interface is flat enough so that capillarity is unimportant, and the interface is located along the melting point isotherm of the system, $T_i = T_m$.

From an algorithmic point of view, two methods are primarily employed for computing the location of the crystal interface. Front-tracking methods define a discrete moving surface to separate the interface between crystal and melt, and diffuse-interface methods treat the interface as a region of finite thickness across which physical properties vary rapidly but continuously from one bulk value to the other. Both methods have been used with great success to model different crystal growth problems. Front-tracking methods, such as the isotherm method, have an advantage in accuracy and numerical efficiency, but their implementation becomes problematic when the shape of the interface becomes complicated, for example, as in representing the shape of a dendrite. Diffuse-interface methods, such as the phase field method, come to the fore in such situations; they are able to compute interface shapes of great complexity, albeit at a higher computational cost for a given level of accuracy.

Modeling the growth velocity of a crystalline surface in bulk solution growth is much more problematic than in melt-growth systems, since interfacial kinetics are much more important. The simplest representation of interface velocity and growth is

$$V_g = \beta_k \sigma \qquad [7]$$

where β_k denotes a kinetic coefficient and σ is the supersaturation. The supersaturation is defined as $\sigma \equiv \Delta\mu_g/k_B T = \ln(C/C_e)$, where $\Delta\mu_g$ is the change in the chemical potential between the crystal and liquid, k_B is the Boltzmann constant, and C and C_e are the actual and equilibrium molar concentrations. The kinetic coefficient in this expression varies strongly as a function of the detailed nature of the surface, posing great challenges for realistic modeling. Indeed, predicting the shape of crystals growing from the solution phase is still a formidable undertaking; see the "Further reading" section for additional details.

Modern atomistic simulation techniques, such as molecular dynamics and kinetic Monte Carlo

methods, are being applied to model crystal interfaces. These techniques are showing great promise for obtaining a deeper understanding of crystal growth; however, they are typically restricted to describing systems characterized by very small length scales over very short time scales (see the "Further reading" section).

Crystal Growth Analysis

As previously emphasized, bulk crystal growth is characterized by strongly interacting and often very nonlinear phenomena. Theory and modeling have allowed many of these behaviors to be analyzed and understood; a few examples are discussed below.

Segregation

For a dilute species (often called a dopant) added to a pure material, the phase behavior of the nearly pure material can be described by a phase diagram of temperature plotted as a function of concentration (at constant pressure), as shown in **Figure 4a**. The important behaviors, from the point of view of crystal growth, are that the solid in equilibrium with the liquid at a fixed temperature differ in composition and that the equilibrium melting (or freezing) temperature is a function of composition. The first of these phenomena gives rise to segregation, referring to the partitioning of a dopant between the solid and melt which results in a crystal of inhomogeneous composition, even when grown from an initially uniform-composition melt. The second phenomenon, when coupled with segregation during growth, is responsible for the morphological instability known as constitutional supercooling and will be discussed in the next section.

Classic explanations of dopant segregation involve directional solidification, where a crystal is growing into a melt in one direction and both phases initially have the same solute composition, c_0. For the case of no melt mixing and a partition coefficient, k, less than unity, the dopant is rejected from the growing crystal and diffuses away from it. This flux causes the melt-dopant concentration at the interface to increase in time, until it is exactly c_0/k, as depicted schematically in **Figure 4b**. Under these conditions, axial segregation is said to be diffusion-limited and, barring initial and final transients, gives rise to a constant composition c_0 along the length of the crystal.

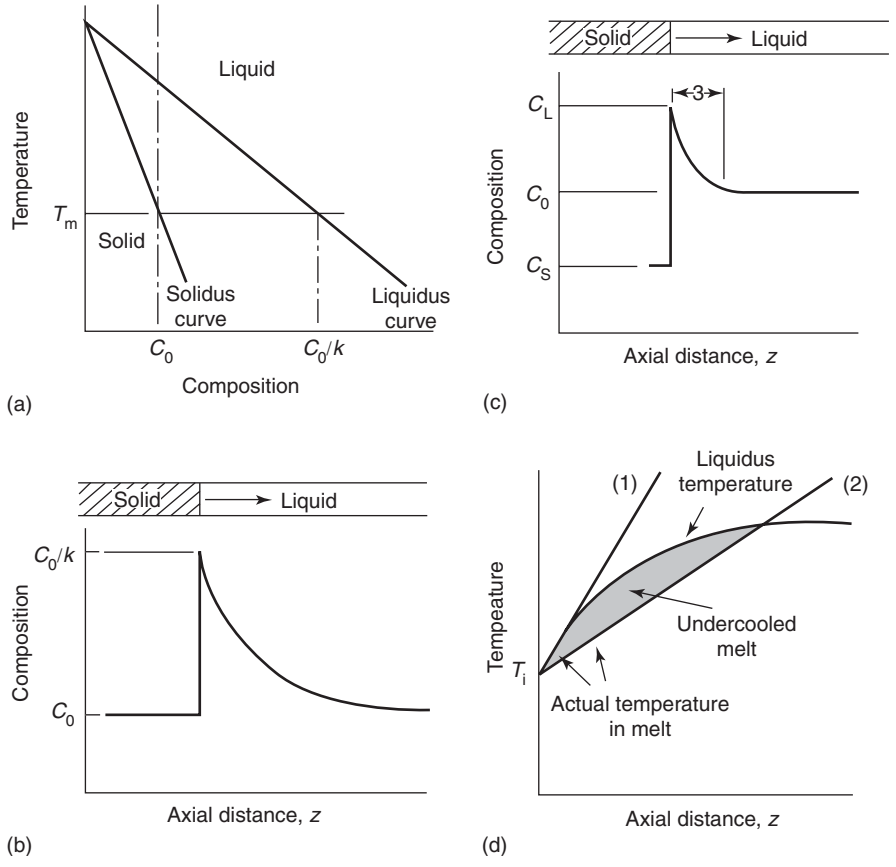


Figure 4 The solidification of a dilute binary alloy: (a) simplified phase diagram, (b) diffusion-limited axial segregation, (c) axial segregation as depicted by the BPS model, and (d) conditions leading to constitutional supercooling.

Convection in the melt disrupts this appealing diffusion-controlled scenario by altering the form of the concentration profile in the melt. The BPS model of Burton, Prim, and Schlicter postulates a stagnant film in front of the interface of thickness δ. Transport is assumed to occur only by diffusion within the film and by perfect convective mixing beyond the film, leading to the dopant concentration profile depicted in **Figure 4c**. Under this scenario, the axial distribution of a dopant in the crystal is given by

$$\frac{c_s}{c_0} = k_{eff}(1-f)^{(k_{eff}-1)} \quad [8]$$

where c_s is the concentration of dopant in the crystal, c_0 is the initial concentration of solute in the bulk, f is the fraction of melt solidified, and k_{eff} is an "effective" distribution coefficient given by

$$k_{eff} = \frac{k}{k + (1-k)\exp(-V_g\delta/\mathscr{D})} \quad [9]$$

where k is the equilibrium distribution coefficient, V_g is the growth velocity of the crystal, and $\mathscr{D}$ is the diffusion coefficient of the dopant in the melt. If $\delta \to \infty$, the effective segregation coefficient approaches unity, and the constant axial concentration profile for diffusion-controlled growth is obtained. As $\delta \to 0$, complete mixing of the melt is implied, $k_{eff} \to k$, and the Scheil equation for axial segregation is recovered. The BPS model is often very effective for fitting experimental segregation data by suitable choice of the parameter δ.

The simplicity and elegance of the BPS expression is perhaps a bit misleading, since it has virtually no predictive capabilities. The underlying idea for this model is too simple; there is no precise physical meaning of the parameter δ, since flows in real systems are never completely stagnant nor intense enough to produce perfect mixing. In addition, segregation across the face of the crystal can be more important than segregation along its length, and this effect is ignored in the one-dimensional solidification model discussed above. Numerical models are required for a more complete analysis of segregation in crystal growth (see the "Further reading" section).

Morphological Instability

The one-dimensional directional solidification model presented above can be used to understand the morphological stability of a crystal interface during growth. The situation in the melt in front of a growing interface is depicted in **Figure 4d**. An axial temperature gradient is applied to the system to maintain growth, with the interface at T_i and the temperature rising as it extends into the melt. Typically, thermal diffusion processes are much faster than mass diffusion, so the temperature profile is linear over the length scale characteristic of the dopant diffusion layer in front of the solidification interface. Also shown in this graph is a curve that denotes the equilibrium freezing temperature of the melt, or the liquidus temperature, in front of the interface. This changes with position due to the compositional profile in the melt set up by segregation and mass diffusion (see **Figure 4c**) and the dependence of the equilibrium freezing temperature on composition (see **Figure 4a**). Under the conditions depicted by temperature profile (1), the growth of the crystallization front into the melt is stable, since the actual temperature is always above the freezing temperature. However, if the thermal gradient is small enough, a situation depicted by profile (2) can arise. Here, much of the melt in front of the interface is below its freezing point, and a perturbation to the interface will grow rapidly into the supercooled region, resulting in an unstable growth interface. This situation is known as constitutional supercooling, and the criterion for it to arise is,

$$\frac{dT_L}{dz} < \frac{mc_0(1-k)V_g}{k\mathscr{D}} \quad [10]$$

where dT_L/dz denotes the axial temperature gradient of the melt at the interface and m is the slope of the liquidus curve on the binary phase diagram. The resultant behavior, which is often referred to as the Mullins–Sekerka instability, is a wavy interface, which, under larger driving forces, becomes successively more unstable in the form of cells and, eventually, dendrites.

Morphological instability can also arise during the bulk growth of crystals from a liquid solution; experiments have demonstrated that steps growing along a vicinal surface can be strongly affected by solution flowing above the surface. Under certain conditions, the density of growth steps redistributes along the surface, giving rise to regions where the steps are very closely spaced, which are known as step bunches. This step-bunching morphological instability is understood to be driven by the interaction of solution flow and mass transfer with the underlying distribution of steps. Flow in the same direction of the step can drive variations in the solute supersaturaton field to be out of phase with the distribution of steps, thereby destabilizing a train of initially evenly spaced steps. Flows opposite to the direction of step motion will generally stabilize perturbations to step spacing. It is postulated that step bunching may be a prelude to the formation of solvent inclusions during growth, but the detailed mechanisms involved are not yet well understood. Chernov provides an excellent overview of these phenomena.

Crystal Defects

Modeling has been gainfully employed to understand the origins of several types of defects in bulk crystals grown from the melt and to optimize growth conditions to reduce their numbers. During the growth of semiconductor crystals, a large number of dislocations can be produced via dislocation formation and multiplication processes driven by thermal stresses generated during growth and cooling. While the necking process developed by Dash for Czochralski growth has allowed large elemental semiconductor crystals (e.g., silicon and germanium) to be grown dislocation-free, large compound semiconductor crystals, such as gallium arsenide and cadmium telluride, are plagued by high levels of dislocations. Modeling heat transfer in growth systems has been successfully employed to minimize thermal stresses and dramatically reduce defect densities of bulk compound semiconductor crystals. Völkl provides an extensive review of this area.

Intrinsic point defects also arise naturally during the growth of all bulk crystals, and their fate under growth conditions determines the properties of the resultant material. As discussed previously, point defects arise from their entropic contribution to lowering the Gibbs energy of the crystal, which varies exponentially with temperature. Point defects also diffuse through the lattice via processes which are strongly thermally activated. There has been much recent progress in understanding how point defects arise, move, and interact in silicon to produce microdefects, such as voids arising from the condensation of excess vacancies and networks of dislocation loops formed by excess interstitials. Accurate modeling is required to predict these effects and optimize growth conditions. Such efforts have been referred to as defect engineering (see the "Further reading" section).

Process Engineering

Advances in computers and numerical algorithms are enabling increasingly powerful models for crystal growth systems to be used for scientific inquiry and engineering optimization. Modeling is a particularly powerful tool for obtaining a better understanding of melt crystal growth, since experimental observations of growth process are made difficult by high temperatures, long process times, and the extreme difficulty of measuring the state of the growing crystal *in situ*. From a technical perspective, the optimization of industrial-scale crystal growth processes via modeling is growing more important with each passing day.

An example of what modern process modeling can do is depicted in **Figure 5**, where the effects of a ampoule tilt and rotation are assessed for a Bridgman growth system. A schematic of the process is shown in **Figure 5a**, and a cut-away view of the three-dimensional finite element mesh employed by the model is shown in **Figure 5b**. This model solves for continuum transport coupled with the position of the melt–crystal interface using a front-tracking technique. Heat transfer boundary conditions are supplied by a furnace thermal analysis code. **Figure 5c** shows isoconcentration contours for a dopant in the melt, and pathlines for the melt flow field are depicted in **Figure 5d**. This information is needed to understand how operating conditions affect crystal growth and how the process can be improved.

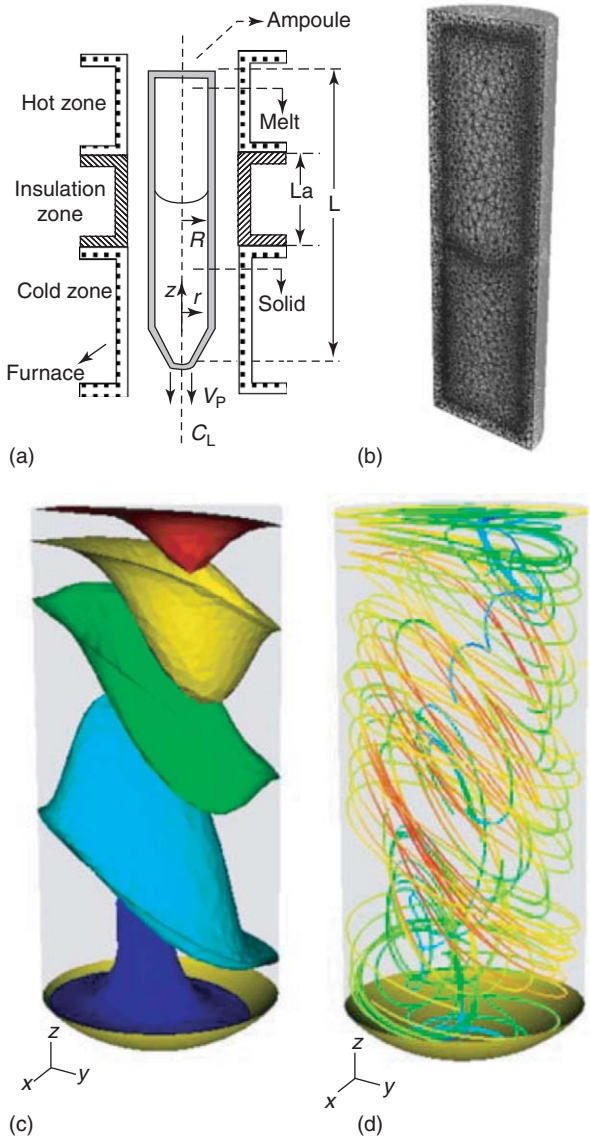


Figure 5 Process model for vertical Bridgman crystal growth system: (a) schematic diagram of process, (b) finite element mesh used in model, (c) isoconcentration surfaces caused by segregation, and (d) melt pathlines showing fluid flow in a tilted, rotated system.

See also: Computer Simulation Techniques in Condensed Matter Physics; Crystal Growth, Bulk: Methods; Film Growth and Epitaxy: Methods; Irreversible Thermodynamics and Basic Transport Theory in Solids; Mass Transport; Phase Transformations, Mathematical Aspects of; Phases and Phase Equilibrium; Specific Heat; Thermodynamic Properties, General.

PACS: 02.60.Cb; 02.70.−c; 44.; 47.27.−i; 61.; 61.72.−y; 64.70.−p; 64.75.+g; 68.08.−p; 81.10.Aj; 81.10.Dn; 81.10.Fq; 81.30.Fb

Further Reading

Brown RA (1988) Theory of transport processes in single crystal growth from the melt. *AIChE Journal* 34: 881–911.

Burton WK, Cabrera N, and Frank FC (1951) The growth of crystals and the equilibrium structure of their surfaces. *Philosophical Transaction of the Royal Society of London* 243: 299–358.

Chernov AA (1961) The spiral growth of crystals. *Soviet Physics Uspekhi* 4: 116–148.

Chernov AA and Komatsu H (1995) Topics in crystal growth kinetics. In: Van der Eerden JP and Bruinsma OSL (eds.) *Science and Technology of Crystal Growth*, pp. 67–80. Dordecht: Kluwer Academic.

Derby JJ (1995) Macroscopic transport processes during the growth of single crystals from the melt. In: Van der Eerden JP and Bruinsma OSL (eds.) *Science and Technology of Crystal Growth*, pp. 97–110. Dordecht: Kluwer Academic Publishers.

Derby JJ, Daoutidis P, Kwon YI, Pandy A, Sonda P, *et al.* (2002) High-performance-computing, multi-scale models for crystal growth systems. In: Breuer M, Durst F, and Zenger C (eds.) *High Performance Scientific and Engineering Computing*, Lecture Notes in Computational Science and Engineering, pp. 185–200. Berlin: Springer.

Derby JJ, Edwards K, Kwon YI, Rojo JC, Vartak B, *et al.* (1998) Large-scale numerical modeling of continuum phenomena in melt and solution crystal growth processes. In: Fornari R and Paorici C (eds.) *Theoretical and Technological Aspects of Crystal Growth*, Material Science Forum, vol. 276–277, pp. 119–134. Zurich: Trans Tech Publishing Ltd.

Hurle DTJ and Rudolph P (2004) A brief history of defect formation, segregation, faceting, and twinning in melt-grown semiconductors. *Journal of Crystal Growth* 264: 550–564.

Mullins WW and Sekerka RF (1964) Stability of a planar interface during solidification of a dilute binary alloy. *Journal of Applied Physics* 35: 444–451.

Jackson KA (1999) Computer modeling of atomic scale crystal growth processes. *Journal of Crystal Growth* 198–199: 1–9.

Rudolph P (2003) Thermodynamic fundamentals of phase transition applied to crystal growth processes. In: Scheel HJ and Fukuda T (eds.) *Crystal Growth Technology*, pp. 15–42. West Sussex: Wiley.

Sinno T, Dornberger E, Von Ammon W, Brown RA, and Dupret F (2000) Defect engineering of Czochralski single-crystal silicon. *Materials Science and Engineering: R: Reports* 28(5–6): 149–198.

Yeckel A and Derby JJ (2003) Computational simulations of the growth of crystals from liquids. In: Scheel HJ and Fukuda T (eds.) *Crystal Growth Technology*, pp. 115–137. West Sussex: Wiley.

Vökl J (1994) Stress in the cooling crystal. In: Hurle DTJ (ed.) *Handbook of Crystal Growth*, vol. 2B, pp. 821–871. Amsterdam: North-Holland.

Winn D and Doherty MF (2000) Modeling crystal shapes of organic materials grown from solution. *AIChE Journal* 46: 1348–1367.

Nomenclature

c	concentration, mass ($\mathrm{kg\,m^{-3}}$)
C	concentration, molar ($\mathrm{mol\,m^{-3}}$)
C_p	heat capacity at constant pressure ($\mathrm{J\,K^{-1}\,kg^{-1}}$)
$\mathscr{D}$	diffusivity ($\mathrm{m\,s^{-2}}$)
f	fraction of melt solidified (–)
$\boldsymbol{F}$	body force per unit volume ($\mathrm{N\,m^{-3}}$)
$\boldsymbol{g}$	gravitational constant ($\mathrm{m\,s^{-2}}$)
G	Gibbs energy ($\mathrm{J\,mol^{-1}}$)
H	enthalpy ($\mathrm{J\,mol^{-1}}$)
$\mathscr{H}$	mean curvature ($\mathrm{m^{-2}}$)
k	distribution coefficient (–)
k_B	Boltzmann constant ($\mathrm{J\,K^{-1}}$)
L	latent heat of solidification ($\mathrm{J\,m^{-3}}$)
m	Liquidus curve slope ($\mathrm{K\,m^3\,kg^{-1}}$)
p	pressure (Pa)
S	entropy ($\mathrm{J\,mol^{-1}\,K^{-1}}$)
t	time (s)
T	temperature (K)
$\boldsymbol{v}$	velocity ($\mathrm{m\,s^{-1}}$)
V_g	growth velocity ($\mathrm{m\,s^{-1}}$)
$\boldsymbol{x}$	position coordinate vector (m)
z	axial distance (m)
β	kinetic coefficient for melt growth ($\mathrm{m\,s^{-1}\,K^{-1}}$)
β_k	kinetic coefficient for solution growth ($\mathrm{m\,s^{-1}}$)
β_s	solutal expansivity ($\mathrm{m^3\,kg^{-1}}$)
β_T	thermal expansivity ($\mathrm{K^{-1}}$)
γ	capillary coefficient ($\mathrm{J\,m^{-1}}$)
δ	stagnant film thickness (m)
κ	thermal conductivity ($\mathrm{W\,m^{-1}\,K^{-1}}$)
μ	dynamic viscosity (Pa s)
μ_g	chemical potential (J)
ρ	density ($\mathrm{kg\,m^{-3}}$)
σ	supersaturation
∇	gradient operator

Crystal Optics

D A Lyashenko and Yu P Svirko, University of Joensuu, Joensuu, Finland

Introduction

Crystal optics describes the optical phenomena in anisotropic media when light waves may behave differently depending on direction propagation and polarization. This branch of optics is named after natural crystals that are used in optical devices for many years. The crystal optics techniques and approaches can also be employed to describe light propagation in noncrystalline bulk media (e.g., in an isotropic liquid) when static external fields induce optical anisotropy or in artificial structures, which can be created by recent advances in material science.

Conventionally, crystal optics is restricted to the investigation of the optical propagation in transparent dielectric crystals leaving behind the dramatic enrichment of the propagation phenomena in the vicinity of the absorption resonances in crystalline media where the frequency and spatial dispersion may play decisive roles. Also, crystal optics studies the linear optical phenomena, which are governed by the second-rank tensor of the dielectric susceptibility. However, it is necessary to notice here that the nonlinear optical phenomena, which arise due to dependence of the light propagation effects on the light wave amplitude and are governed by the higher-rank tensors of nonlinear susceptibility, are much more sensitive to the crystal symmetry. Nevertheless, the core part of crystal optics "crystallized" in the pre-laser era and these intensity-dependent effects are studied in the framework of nonlinear rather than crystal optics. In this article, major effects in light propagation in anisotropic nonmagnetic media are discussed.

Maxwell Equations

Spatial and temporal evolutions of the electric field $\boldsymbol{E}$ and magnetic induction $\boldsymbol{B}$ of a light wave in nonmagnetic medium are described by the Maxwell equations:

$$\begin{aligned} \operatorname{curl} \boldsymbol{E} &= \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \\ \operatorname{div} \boldsymbol{B} &= 0 \\ \operatorname{curl} \boldsymbol{B} &= \frac{1}{c}\frac{\partial \boldsymbol{D}}{\partial t} \\ \operatorname{div} \boldsymbol{D} &= 0 \end{aligned} \qquad [1]$$

The mutual orientation of the vectors $\boldsymbol{E}$, $\boldsymbol{B}$ and the electric displacement $\boldsymbol{D}$ in an electromagnetic wave propagating in a nonabsorbing anisotropic medium is determined by the Maxwell equations [1]. In the plane monochromatic wave with frequency ω and wave vector $\boldsymbol{k}$,

$$\begin{aligned} \boldsymbol{E}(t,\boldsymbol{r}) &= \boldsymbol{E}\,\mathrm{e}^{-\mathrm{i}\omega t+\mathrm{i}\boldsymbol{k}\boldsymbol{r}} + \boldsymbol{E}^{*}\mathrm{e}^{\mathrm{i}\omega t-\mathrm{i}\boldsymbol{k}\boldsymbol{r}} \\ \boldsymbol{B}(t,\boldsymbol{r}) &= \boldsymbol{B}\,\mathrm{e}^{-\mathrm{i}\omega t+\mathrm{i}\boldsymbol{k}\boldsymbol{r}} + \boldsymbol{B}^{*}\mathrm{e}^{\mathrm{i}\omega t-\mathrm{i}\boldsymbol{k}\boldsymbol{r}} \\ \boldsymbol{D}(t,\boldsymbol{r}) &= \boldsymbol{D}\,\mathrm{e}^{-\mathrm{i}\omega t+\mathrm{i}\boldsymbol{k}\boldsymbol{r}} + \boldsymbol{D}^{*}\mathrm{e}^{\mathrm{i}\omega t-\mathrm{i}\boldsymbol{k}\boldsymbol{r}} \end{aligned} \qquad [2]$$

the vectors $\boldsymbol{D}$, $\boldsymbol{B}$, and $\boldsymbol{k}$ are perpendicular to one another:

$$\begin{aligned} \boldsymbol{k}\times\boldsymbol{B} &= -\frac{\omega}{c}\boldsymbol{D} \\ \boldsymbol{k}\times\boldsymbol{E} &= \frac{\omega}{c}\boldsymbol{B} \end{aligned} \qquad [3]$$

That is, in any anisotropic medium, $\boldsymbol{D}\perp\boldsymbol{B}$ and $\boldsymbol{E}\perp\boldsymbol{B}$. However, vectors $\boldsymbol{E}$ and $\boldsymbol{D}$ are not necessarily collinear. The plane containing vectors $\boldsymbol{D}$ and $\boldsymbol{B}$ (and perpendicular to the wave vector $\boldsymbol{k}$) is called the wave front plane, while those containing pair $\boldsymbol{k}$, $\boldsymbol{H}$ and $\boldsymbol{k}$, $\boldsymbol{D}$ are called the polarization plane and vibration plane, respectively. The plane containing orthogonal vectors $\boldsymbol{E}$ and $\boldsymbol{B}$ is perpendicular to the energy flux vector $\boldsymbol{S} = (c/4\pi)[\boldsymbol{E}\times\boldsymbol{B}]$, which is not necessarily parallel to the wave vector $\boldsymbol{k}$ and determines the direction of the ray propagation in the medium (see **Figure 1**). Since the number of variables in Maxwell's equations is greater than the number of equations, an

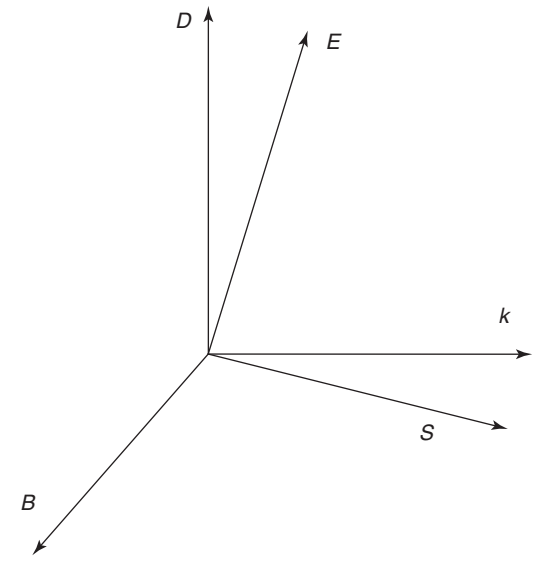


Figure 1 Mutual orientation of vectors $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$, $\boldsymbol{k}$, and $\boldsymbol{S}$ in the anisotropic medium.

additional relationship (constitutive equation) is needed between electric displacement $\boldsymbol{D}$ and the wave fields. However, the electric and magnetic fields of the light wave in a nonmagnetic medium are related by the first Maxwell equation, that is, the constitutive equation can be reduced down to $\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{E})$. This functional relationship between electric displacement and electric field strength should account for the material properties of the medium.

Constitutive Equation in Anisotropic Medium

Constitutive equations of the medium must obey the causality principle that ensures a retrospective relationship between $\boldsymbol{D}$ and $\boldsymbol{E}$: displacement at the moment is determined by the preceding electric field, but not by the future one. Similarly, the displacement is not necessarily local, that is, it could depend on the electric field not only at the point of observation, but also in some vicinity around it. With account for the causality and nonlocality, the constitutive equation for linear optics can be presented in the following form:

$$D_i(t,\boldsymbol{r}) = \sum_{j=x,y,z} \int_{-\infty}^{t} \mathrm{d}t_1 \int_V \mathrm{d}\boldsymbol{r}_1 \times \psi_{ij}(t-t_1, \boldsymbol{r}-\boldsymbol{r}_1) E_j(t_1, \boldsymbol{r}_1) \qquad [4]$$

Here, the subscripts label the Cartesian axes x, y, z of the laboratory frame, $\psi_{ij}(t-t_1, \boldsymbol{r}-\boldsymbol{r}_1)$ is the response function that describes the contribution into displacement at point $\boldsymbol{r}$ and time t resulting from the finite electric field at point $\boldsymbol{r}_1$ at time t_1. Since $\boldsymbol{E}$ and $\boldsymbol{D}$ are vector quantities, the response function is tensorial by nature and relates projections of the displacement and electric filed in the laboratory Cartesian frame. The properties of the response function are determined by the mechanism of the optical response in the medium and the medium symmetry. If the optical response is instantaneous, $\psi_{ij}(t-t_1, \boldsymbol{r}-\boldsymbol{r}_1) \propto \delta(t-t_1)$, that is, displacement at the moment t is given by the electric field at the same moment of time. If the response in a particular location only depends on excitation at this location, $\psi_{ij}(t-t_1, \boldsymbol{r}-\boldsymbol{r}_1) \propto \delta(\boldsymbol{r}-\boldsymbol{r}_1)$. In molecular media, the response function would normally have a maximum at $\boldsymbol{r} = \boldsymbol{r}_1$ and diminish for $|\boldsymbol{r}-\boldsymbol{r}_1| > a$, where a is the typical size of a molecule. In crystals, the effects of the spatial dispersion are often associated with excitation of quasiparticles such as excitons, which may have effective radii of up to hundreds of angstroms.

The constitutive equation can be reduced to a more familiar form if one considers propagation of a monochromatic wave with frequency ω and wave vector $\boldsymbol{k}$ (eqn [2]):

$$D_i = \sum_{j=x,y,z} \varepsilon_{ij}(\omega, \boldsymbol{k}) E_j \qquad [5]$$

where

$$\varepsilon_{ij}(\omega, \boldsymbol{k}) = \int_0^{\infty} \mathrm{d}\tau \int_V \mathrm{d}\boldsymbol{\rho} \psi_{ij}(\tau, \boldsymbol{\rho}) \mathrm{e}^{\mathrm{i}\omega\tau - \mathrm{i}\boldsymbol{k}\boldsymbol{\rho}} \qquad [6]$$

is the dielectric tensor that governs the propagation of the electromagnetic waves in the medium.

In the following, the spatial dispersion effects (i.e., the dependence of the dielectric tensor on the wave vector), which in most cases result in minor correction in the wave propagation in anisotropic media, are ignored. However, if the spatial dispersion is the only source of the optical anisotropy (e.g., circular birefringence in chiral liquids), it should be taken into account to describe the light propagation effects.

Dielectric tensor is symmetric with respect to permutation of its indices, $\varepsilon_{ij} = \varepsilon_{ji}$. This fundamental symmetry relation arises from the Onsager principle of symmetry of kinetic coefficients and holds in nonmagnetic media at thermal equilibrium. Another important property of the dielectric tensor is its Hermicity for nonabsorbing media, that is, $\varepsilon_{ij} = (\varepsilon_{ji})^*$ in the frequency region where the absorption losses are negligible. Correspondingly, the energy dissipation rate in the medium at frequency ω is given by the non-Hermitian part of the dielectric tensor:

$$Q = \frac{\mathrm{i}\omega}{8\pi} \sum_{i,k} (\varepsilon_{ij}^* - \varepsilon_{ji}) E_i E_j^* \qquad [7]$$

In the transparent medium the dielectric tensor is real, $\varepsilon_{ij} = (\varepsilon_{ji})^* = (\varepsilon_{ij})^*$, and, therefore, there exists a Cartesian frame in which it can be presented in the diagonal form:

$$\varepsilon_{ij} = \begin{pmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{pmatrix} \qquad [8]$$

$\varepsilon_1, \varepsilon_2$, and ε_3 are called the principal dielectric constants (or principal permittibilities) of the crystal that are always positive. The coordinate axes of this system are called principal dielectric axes of the crystal. In this coordinate frame, the constitutive equation [5] reduces down to $D_1 = \varepsilon_1 E_1$, $D_2 = \varepsilon_2 E_2$, $D_3 = \varepsilon_3 E_3$, that is, displacement and electric field will not parallel one another unless $\boldsymbol{E}$ coincides with one of the principal axes or the principal dielectric constants are equal.

If the optical response is not instantaneous, the dielectric tensor depends on the frequency, and,

therefore, both the orientation of the principal axes and the magnitude of the principal dielectric constants may vary with frequency. On the other hand, crystal symmetry imposes restrictions on the structure of the dielectric tensor and, therefore, on the number of nonequal principal dielectric constants. The competition between symmetry restrictions and frequency (also referred to as color) dispersion allows one to introduce three distinct groups in the optical classification of the crystals.

In cubic crystals, three crystallographically equivalent, mutually orthogonal directions may be chosen. In crystals of a cubic system, the principal dielectric constants are equal, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon$. Therefore, the constitutive equation in cubic crystals is the same as that in an isotropic medium, $\boldsymbol{D} = \varepsilon \boldsymbol{E}$, that is, cubic crystals are optically isotropic. In these crystals, the frequency dispersion may change only the magnitude of the permittivity.

In crystals of trigonal, tetragonal, and hexagonal systems, the plane containing two equivalent perpendicular directions is perpendicular to the three-, four-, or sixfold rotation axis, respectively. Conventionally, this direction is chosen along the third principal axis. In these crystals, which are called uniaxial, relationship $\varepsilon_1 = \varepsilon_2 \neq \varepsilon_3$ holds at all frequencies.

In crystals of orthorhombic, monoclinic, and triclinic systems, no two equivalent directions can be chosen. These crystals in which $\varepsilon_1 \neq \varepsilon_2 \neq \varepsilon_3$, are called biaxial. In crystals belonging to the orthorhombic system, the directions of the principal axes are fixed and coincide with twofold rotation axes and/or symmetry planes. This ensures their independence from the wave frequency. This situation changes for crystals of the monoclinic system, in which only the twofold rotation axis exists, and correspondingly, orientation of just one principal axis does not show color dispersion. In crystals of the triclinic system, both the directions of the principal axes and magnitudes of the principal dielectric indices may change with frequency.

Fresnel Equation, Wave Vector Surface, and Ray Surface

Maxwell equations [3] allow one to arrive at the following equation for displacement:

$$\frac{\omega^2}{c^2}\boldsymbol{D} = [\boldsymbol{k} \times [\boldsymbol{E} \times \boldsymbol{k}]] = \boldsymbol{k}^2\boldsymbol{E} - (\boldsymbol{k}\boldsymbol{E})\boldsymbol{k} \qquad [9]$$

This gives three linear homogeneous equations for the Cartesian components of electric field in the principal crystal frame:

$$\begin{aligned}
&\left(k_2^2 + k_3^2 - \frac{\omega^2}{c^2}\varepsilon_1\right)E_1 - k_1k_2E_2 - k_1k_3E_3 = 0\\
&-k_1k_2E_1 + \left(k_1^2 + k_3^2 - \frac{\omega^2}{c^2}\varepsilon_2\right)E_2 - k_2k_3E_3 = 0\\
&-k_1k_3E_1 - k_2k_3E_2\\
&\quad + \left(k_1^2 + k_2^2 - \frac{\omega^2}{c^2}\varepsilon_3\right)E_3 = 0
\end{aligned} \qquad [10]$$

The compatibility condition for these equations is that the relevant determinant should vanish. This condition gives the so-called Fresnel's equation that defines the frequency dependence of the wave vector (or dispersion relation). By introducing vector $\boldsymbol{n} = (c/\omega)\boldsymbol{k}$ along the direction of propagation and effective refraction index $n = |\boldsymbol{n}|$, the Fresnel equation is conventionally presented in the following form:

$$\begin{aligned}
&n^2(\varepsilon_1 n_1^2 + \varepsilon_2 n_2^2 + \varepsilon_3 n_3^2)\\
&\quad - [n_1^2\varepsilon_1(\varepsilon_2 + \varepsilon_3) + n_2^2\varepsilon_2(\varepsilon_1 + \varepsilon_3)\\
&\quad + n_3^2\varepsilon_3(\varepsilon_1 + \varepsilon_2)] + \varepsilon_1\varepsilon_2\varepsilon_3 = 0
\end{aligned} \qquad [11]$$

When the direction of propagation is given, eqn [11] gives two different magnitudes of the effective refractive index. Fresnel's equation in coordinates $\{n_1, n_2, n_3\}$ defines the so-called wave vector surface.

Normal to the wave vector surface gives the direction of the energy propagation (Poynting vector) in the medium. Usually, it is described by the so-called ray vector $\boldsymbol{s}$, which is introduced by condition $(\boldsymbol{n}\boldsymbol{s}) = 1$. Since the Poynting vector is perpendicular to $\boldsymbol{E}$ and $\boldsymbol{B}$, one can readily arrive at

$$\begin{aligned}
\boldsymbol{s} \times \boldsymbol{B} &= -\boldsymbol{E}\\
\boldsymbol{s} \times \boldsymbol{D} &= \boldsymbol{B}
\end{aligned} \qquad [12]$$

By replacing $\boldsymbol{E}$ with $\boldsymbol{D}$ and $\boldsymbol{k}$ with $-(\omega/c)\boldsymbol{s}$ and vice versa in [3], one may arrive at [12]. This corresponds to the rule of duality in crystal optics. According to this rule, an equation valid for vectors $\boldsymbol{E}$, $\boldsymbol{n}$, and tensor ε_{ij} is also valid if these variables are replaced with $\boldsymbol{D}$, $\boldsymbol{s}$, and tensor $(\varepsilon_{ij})^{-1}$, respectively. For example, from Fresnel's equation [11], one can obtain the equation which defines the ray vector surface:

$$\begin{aligned}
&s^2(\varepsilon_3\varepsilon_2 s_1^2 + \varepsilon_1\varepsilon_3 s_2^2 + \varepsilon_1\varepsilon_2 s_3^2)\\
&\quad - [s_1^2(\varepsilon_2 + \varepsilon_3) + s_2^2(\varepsilon_1 + \varepsilon_3)\\
&\quad + s_3^2(\varepsilon_1 + \varepsilon_2)] + 1 = 0
\end{aligned} \qquad [13]$$

One can observe that when the direction of $\boldsymbol{s}$ is given, two rays with two different wave vectors can propagate in the crystal.

The above analysis can be visualized by means of simple geometrical constructions. One may draw the

ellipsoid that corresponds to the tensor $(\varepsilon_{ij})^{-1}$ with semiaxes equal to the square roots of the principal dielectric constants. This equation is called the ellipsoid of wave normals (also known as the index ellipsoid), and in the Cartesian frame with principal dielectric axes, it is described by the following equation:

$$\frac{x^2}{\varepsilon_1}+\frac{y^2}{\varepsilon_2}+\frac{z^2}{\varepsilon_3}=1 \qquad [14]$$

If the ellipsoid is cut by a plane which is orthogonal to vector $\boldsymbol{n}$, the intersection of the plane with the ellipsoid will be an ellipse (see **Figure 2**). The length of its axes determines the two values of the effective refractive index, while the axes determine the directions of the relevant vector $\boldsymbol{D}$. It is necessary to notice that unit vectors in these two directions and unit vector along $\boldsymbol{n}$ form Cartesian basis.

In order to find the orientation of the electric field in the medium, the ellipsoid that corresponds to the tensor ε_{ij} with semiaxes equal to the square roots of the inverse principal dielectric constants (the Fresnel or ray ellipsoid) is considered:

$$\varepsilon_1 x^2+\varepsilon_2 y^2+\varepsilon_3 z^2=1 \qquad [15]$$

If the ellipsoid is cut by a plane with normal $\boldsymbol{s}$, the intersection of the plane with the ellipsoid will be an ellipse, whose axes determine the two values of the effective refractive index and orientation of the relevant vector $\boldsymbol{E}$.

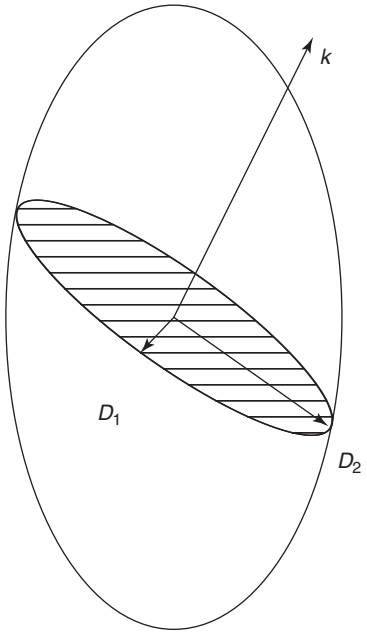


Figure 2 Ellipsoid of wave normals. The normal to the wave vector section of the ellipsoid is ellipse (shaded). The length of its semiaxes determines the values of the refractive index, while the orthogonal axes determine the directions of oscillations, i.e., transverse components $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ of the electric the displacement vector.

Optical Properties of Uniaxial Crystals

In uniaxial crystals, two principal dielectric constants are equal to one another ($\varepsilon_1=\varepsilon_2=\varepsilon_\perp$, $\varepsilon_3=\varepsilon_{||}$). This reduces the Fresnel equation [11] down to the product of two factors:

$$(n^2-\varepsilon_\perp)\times$$
$$[\varepsilon_{||}n_3^2+\varepsilon_\perp(n_1^2+n_2^2)-\varepsilon_{||}\varepsilon_\perp]=0 \qquad [16]$$

This implies that the wave vector surface for uniaxial crystals consists of two separate surfaces, a sphere and spheriod, which touch one another at opposite poles on the n_3-axis. Depending on the sign of $\varepsilon_\perp-\varepsilon_{||}$, the crystal is said to be a negative or a positive uniaxial crystal (the sphere lies outside or inside the spheroid, respectively; see **Figure 3**).

Correspondingly, there may exist two types of waves (ordinary and extraordinary waves) that propagate along a given direction in a uniaxial crystal. The relevant refractive indices are the following:

$$n_0=\sqrt{\varepsilon_\perp}$$
$$n_e=\sqrt{\frac{\varepsilon_{||}\varepsilon_\perp}{\varepsilon_\perp\sin^2\theta+\varepsilon_{||}\cos^2\theta}} \qquad [17]$$

where θ is the angle between the optical axis (the principal axis "3") and the wave normal. Two

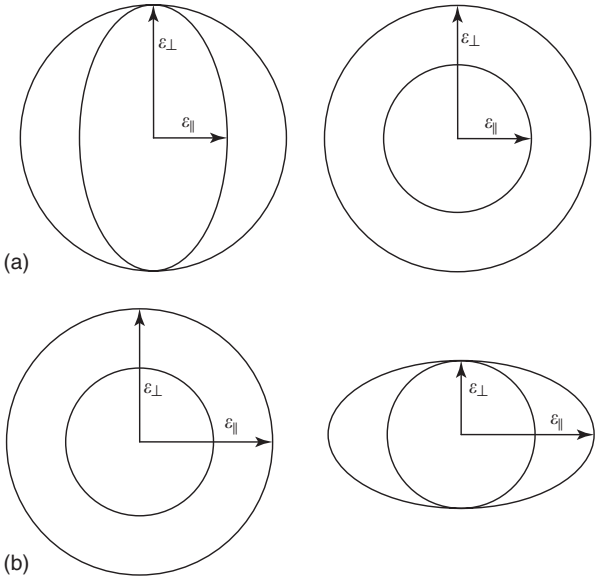


Figure 3 Orthogonal sections of normal surfaces for uniaxial crystals: (a) negative crystal – the sphere with radius equal to $\varepsilon_\perp$ lies outside the spheroid with main semiaxes $\varepsilon_\perp$ and $\varepsilon_{||}$. (b) positive crystal – the sphere with radius equal to $\varepsilon_\perp$ lies inside the spheroid with main semiaxes $\varepsilon_\perp$ and $\varepsilon_{||}$.

refractive indices are equal only at $\theta = 0$ when the wave normal coincides with the optical axis.

The ray surface can be found by replacing $\boldsymbol{n}$ with $\boldsymbol{s}$ and $\varepsilon_{\perp,||}$ with $(\varepsilon_{\perp,||})^{-1}$ in [16]. For the ordinary wave, the directions of the wave vector and ray vector are the same, that is, with respect to ordinary waves the crystal behaves like an isotropic medium. For the extraordinary wave, the directions of the ray vector and wave vector do not coincide; however, they lie in the plane containing the wave vector and the optical axis. This plane is called the principal plane. The angle θ' between the ray vector of the extraordinary wave and the crystal axes is given by the following equation: $\tan\theta' = (\varepsilon_{\perp}/\varepsilon_{||})\tan\theta$.

Since the four vectors $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{s}$, and $\boldsymbol{n}$ always lie in the same plane, the extraordinary wave is polarized so that vector $\boldsymbol{E}$ lies in the principal plane. Therefore, in the ordinary wave, vector $\boldsymbol{E}$ lies in the plane, which is perpendicular to the principal one.

Optical Properties of Biaxial Crystals

In biaxial crystals, the three principal values of the tensor are all different, that is, the wave normally becomes ellipsoid rather than spheroid as in uniaxial crystals. The direction of the principal axes may, in biaxial crystals, depend on frequency. The Fresnel surface can be obtained by solving eqn [11]. In the case of $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$, the shape of one octant of the wave vector surface is shown in **Figure 4**.

The singular point of self-intersection (there are four such points, one in each quadrant) gives the direction of the optical axes or binormals of the crystal. The dashed line in **Figure 4** passes through two opposite singular points and is at the following angle to the "3"-axis:

$$\tan\beta = \sqrt{\frac{\varepsilon_3(\varepsilon_2 - \varepsilon_1)}{\varepsilon_1(\varepsilon_3 - \varepsilon_2)}} \qquad [18]$$

The directions of the optical axes are evidently the only ones for which the wave vector has only one magnitude.

The properties of the ray surface are entirely similar to those of the wave surface. The directions of the optical ray axes or biradials can be derived from [18] by replacing $\varepsilon_{1,2,3}$ with $(\varepsilon_{1,2,3})^{-1}$:

$$\tan\gamma = \sqrt{\frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_3 - \varepsilon_2}} \qquad [19]$$

The directions of the wave vector and the ray vector are the same only for light propagating along one of the principal axes.

Near a singular point, the inner and outer parts of the wave vector surface are cones with a common vertex and, therefore, the direction of the normal to the surface becomes indeterminate. This implies that the wave vector along the binormal corresponds to an infinity of ray vectors, whose directions occupy a certain conical surface, called the cone of internal conical refraction. Similar results hold for the wave vectors corresponding to a given ray vector. The ray vectors along the biradial correspond to an infinite number of wave vectors, whose directions occupy the cone of external conical refraction.

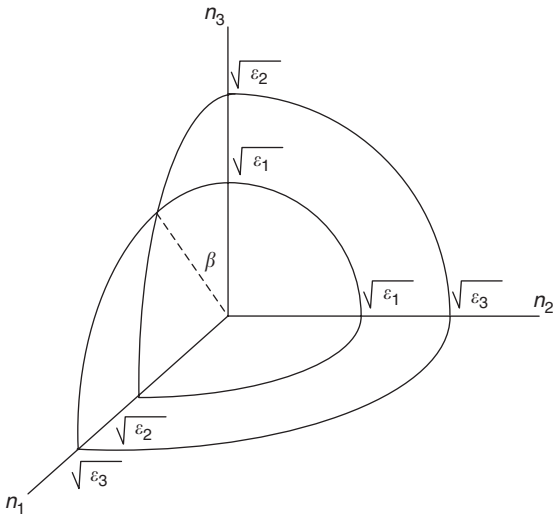


Figure 4 One octant of the Fresnel surface of a biaxial crystal with $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$. Dashed line shows the optical axis (binormal) of the crystal. One can observe that a section of the Fresnel surface by the plane normal to one of the coordinate axes consists of an ellipse and a circle. In the plane $\{n_1, n_2\}$, ellipse with semiaxes ε_1 and ε_2 lies inside the circle with radius ε_3, while in the plane $\{n_2, n_3\}$, the circle with radius ε_1 lies inside the ellipse with semiaxes ε_2 and ε_3. In the plane $\{n_1, n_3\}$, the ellipse with semiaxes ε_1 and ε_3 crosses the circle with radius ε_2.

In observations of the internal conical refraction, one can use a crystal plate cut perpendicular to the binormal. When monochromatic light wave incidents normally on one of the crystal faces, the wave vector of the transmitted wave is parallel to the binormal and so the rays are on the cone of internal refraction. Therefore, a collimated light beam that enters the crystal spreads out into a hollow cone emerging from the crystal as a hollow light cylinder.

In order to observe the external conical refraction, the crystal plate must be cut perpendicular to the biradial, and small apertures should be placed in exactly opposite positions on both crystal facets. When the first aperture is illuminated by a focused beam, only rays that propagate along the biradial can reach the second aperture. Therefore, the relevant wave vectors occupy the cone of the external conical refraction. It is necessary to notice, however, that the cone of light emerging from the crystal does not exactly coincide

with the cone of external refraction because of the refraction on crystal-vacuum inteface.

Polarization Devices Based on the Crystal Optics

Since in the crystal two linear polarized waves with different wave vectors are allowed along a given direction, one can observe double refraction or birefringence at the vacuum–crystal interface. That is, an arbitrary polarized light beam entering an anisotropic crystal is divided into two beams with orthogonal linear polarizations. The direction of propagation of the beams is determined by the orientation of the incident wave vector with respect to the principal crystal axes and does not necessarily lie in the incident plane. These properties of the vacuum–crystal interface have enabled one to develop various devices to control and to investigate the polarization properties of light. The polarization devices belonging to two most common types are discussed below. These are polarizers that produce linear polarized light by splitting an elliptically polarized beam into two linearly polarized ones, and retarders that control the ellipticity of the light beam by changing the phase difference between its orthogonally polarized components.

The Nicol prism was invented in the nineteenth century and consists of a natural rhombic crystal of calcite, which is cut into two parts along a diagonal plane, and joined together again with Canada balsam (see **Figure 5**). For a given angle of incidence, at the vacuum–crystal interface a light beam splits into ordinary and extraordinary rays with effective refractive indices $n_0 = 1.6584$ and $n_e = 1.4864$, respectively. Since the refractive index of the Canada balsam $n_{balsam} = 1.5260$ satisfies the condition $n_e < n_{balsam} < n_0$, the ordinary ray undergoes total reflection at the Canada balsam joint while the extraordinary ray passes through. Therefore, the Nicol prism allows one to achieve linear polarization with practically no lateral displacement of the incident light beam. It is necessary to notice that a number of polarizing prisms (e.g., Glan–Foucault polarizer that employs total internal reflection similar to the Nicol prism and Wollaston prism that produce two orthogonally polarized output beams) have been devised.

The difference in the phase velocity of the ordinary and extraordinary waves in the birefringent crystal enables one to achieve a desired phase difference between them by changing the crystal thickness and/or the relevant refractive indices. Since the directions of vibrations in these waves are mutually orthogonal, the total phase difference accumulated between them determines the polarization of the transmitted light. Such devices are called compensators because they

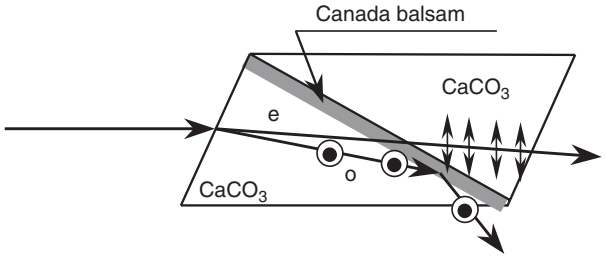


Figure 5 The Nicol prism produces the light beam polarized in the plane of the picture (e-beam). The orthogonally polarized o-beam undergoes total internal reflection on the calcite–Canada balsam interface.

enable one to compensate the phase difference between orthogonal components of the elliptically polarized light wave. It is noticed here that these devices may also be referred to as retarders, because they introduce a finite retardation of one component with respect to another.

A compensator is usually made from a parallel plate of birefringent crystal that cuts so that the optic axis is parallel to the front and back plates of the plate. At the vacuum–crystal interface ($z = 0$), the input wave at normal incidence divides into ordinary and extraordinary ones that propagate with phase velocities $v_0 = c/n_0$ and $v_e = c/n_e$, respectively, and linearly orthogonally polarized. If the components of the light field in eqn [2] at $z = 0$ are equal to E_x and E_y, respectively, those of the light waves emerging from the plate are given by $E_x \exp\{2i\pi n_0 d/\lambda\}$ and $E_y \exp\{2i\pi n_e d/\lambda\}$, where λ is the wavelength in vacuum. Therefore, the phase difference, which the plate introduces between the ordinary and extraordinary waves, is given by the following equation:

$$\Delta\phi = \frac{2\pi d}{\lambda}(n_0 - n_e) \qquad [20]$$

The quarter-wave plate has a thickness of $d = \lambda/4|n_0 - n_e|$ and introduces a phase difference $\Delta\phi = \pm\pi/2$ to the transmitted wave. This wave plate transforms linearly polarized wave into circularly polarized one and vice versa. In order to obtain circularly polarized wave from linearly polarized, the optical axis of the plate is oriented at 45° to the incident polarization azimuth so that $E_x(z = 0) = E_y(z = 0)$. In such a case, the amplitudes of the transmitted ordinary and extraordinary polarized waves remain equal in magnitude. However, they acquire the $\pm\pi/2$ phase shift, that is, $E_x(z = 0) = \pm\, iE_y(z = 0)$. This implies that the wave becomes left- or right-circular polarized depending on the relationship between n_0 and n_e.

The wave plate with thickness of $d = \lambda/2|n_0 - n_e|$ is called the half-wave-plate, which allows one to rotate the polarization azimuth by 90°. If the optical

axis of the plate is oriented at 45° to the incident polarization azimuth (that is, $E_x(z=0) = E_y(z=0)$), in the transmitted wave, the extraordinary polarized components acquire the π phase shift, that is, $E_x(z=d) = E_x(z=0)$ while $E_y(z=d) = -E_y(z=0)$. One can readily find that this corresponds to the rotation of the polarization azimuth by 90° or to the transformation of the left circular polarized beam to the right circular polarized one.

The frequency dispersion results in color sensitivity of the wave plates. In order to achieve compensation of the phase shift for any wavelength, devices based on the combination of edges are used. The most common examples are the Babinet and Soleil compensators.

See also: Electrodynamics: Continuous Media; Nonlinear Optics; Optical Instruments; Optical Properties of Materials.

PACS: 45.25.Lc; 42.25.Bs; 78.20. – e; 61.50.Ah

Further Reading

Born M and Wolf E (1965) *Principles of Optics*. Oxford: Pergamon.

Ditchburn RW (1976) *Light*. London: Academic Press.

Landau LD and Lifshitz EM (1966) *Electrodynamics of Continuous Media*. Oxford: Pergamon.

Nye JF (1985) *The Physical Properties of Crystals*. Oxford: Clarendon Press.

Svirko YuP and Zheludev NI (1998) *Polarization of Light in Nonlinear Optics*. Chichester: Wiley.

Crystal Structure

J S Rutherford, National University of Science and Technology, Bulawayo, Zimbabwe

Introduction

The bulk of our experimental knowledge of chemical structure, the arrangement of atoms in space, comes from the study of crystalline solids. This article intends to explain the nature of this chemical information, and how it may be extracted from the experimental results, given the form of the theoretical model used to describe the crystal structure.

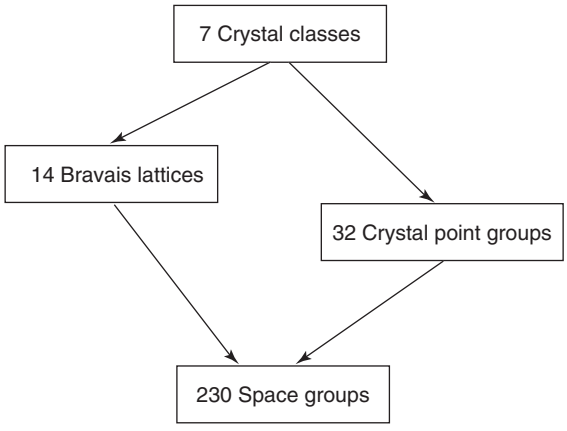
This article first discusses some relevant elements of crystal symmetry and explains how the crystal structure is coded in a typical theoretical model. The information available from experimental techniques are used to establish crystal structures, especially X-ray diffraction and also diffraction of neutrons and electrons. A very large number of crystal structures have now been determined by these methods, and the resulting information is held in various crystallographic databases; these will therefore be described, focusing on the information they contain on individual determinations, which will typically include lattice parameters and space group, as well as the details of the unit cell contents as described by the refinement model used.

It is important to be able to evaluate the reliability of geometric descriptors derived from such a model, so refinement models and the published measures of the quality of the experimental structures determined are discussed. This is followed by a review of derived parameters, that is, those geometric descriptors of chemical interest.

Crystal Symmetry

Ignoring defects, the crystal is a regular arrangement of identical building blocks, the unit cells, in three dimensions. The set of points equivalent to each other by the translation from one unit cell to another form a quasi-infinite three-dimensional geometric lattice, and, in order to describe a crystal structure, it is sufficient to define both this lattice and the structure or contents of an individual unit cell.

Crystal structures are characterized by their space group symmetry, that is, each structure falls into one of the 230 space group types detailed in International Tables for Crystallography. Each of these space groups falls into one of each of the broader categories of crystal system, Bravais lattice and crystal point group as shown below, that is, it will belong to a Bravais lattice and a crystal point group within its crystal class.



In general, a space group symmetry operation consists of a matrix $\boldsymbol{R}$ and a vector $\boldsymbol{t}$, which generate

Table 1 Some properties of the crystal classes

Crystal class	*Restrictions on lattice parameters*	*Possible lattices*	*Point groups and their orders (in parentheses)*			
			Nonpolar nonchiral	*Polar nonchiral*	*Nonpolar chiral*	*Polar chiral*
Cubic	$\mathbf{a}(=\mathbf{b}=\mathbf{c})$ $(\alpha=\beta=\gamma=90^\circ)$	P,I,F	*m* $\bar{3}$*m* (*48*) $\bar{4}$3m (*24*), *m*$\bar{3}$(*24*)		432(*24*), 23(*12*)	
Hexagonal	$\mathbf{a}(=\mathbf{b})\neq\mathbf{c}$ $(\alpha=\beta=90^\circ,$ $\gamma=120^\circ)$	P	6/*mmm*(*24*), $\bar{6}$m2(*12*), 6/m(*12*), $\bar{6}$(*12*)	6 *mm*(*12*)	622(*12*)	6(*6*)
Trigonal	$\mathbf{a}(=\mathbf{b})\neq\mathbf{c}$ $(\alpha=\beta=90^\circ,$ $\gamma=120^\circ)$	P,R	$\bar{3}$*m*(*12*), $\bar{3}$(*6*)	3 *m*(*6*)	32(*6*)	3(*3*)
Tetragonal	$\mathbf{a}(=\mathbf{b})\neq\mathbf{c}$ $(\alpha=\beta=\gamma=90^\circ)$	P,I	4/*mmm*(*16*), $\bar{4}$2*m*(*16*), 4/*m*(*8*), $\bar{4}$(*8*)	4 *mm*(*8*)	422(*8*)	4(*4*)
Orthorhombic	$\mathbf{a}\neq\mathbf{b}\neq\mathbf{c}$ $(\alpha=\beta=\gamma=90^\circ)$	P,I,C,F	*mmm*(*8*)	*mm*2(*4*)	222(*4*)	
Monoclinic	$\mathbf{a}\neq\mathbf{b}\neq\mathbf{c}$ $(\beta\neq(\alpha=\gamma=90^\circ)$	P,C	2/*m*(*4*)	*m*(*2*)		2(*2*)
Triclinic	$\mathbf{a}\neq\mathbf{b}\neq\mathbf{c}$ $\alpha\neq\beta\neq\gamma$	P	$\bar{1}$(*2*)			1(*1*)

Table 2 Lattice types and their equivalent points under translation

Lattice type	*No. of points*	*Equivalent points under translation are (x,y,z)+*
P	1	(0, 0, 0)
C	2	(0, 0, 0) (1/2, 1/2, 0)
I	2	(0, 0, 0) (1/2, 1/2, 1/2)
R	3	(0, 0, 0) (2/3, 1/3, 1/3) (1/3, 2/3, 2/3)
F	4	(0, 0, 0) (1/2, 1/2, 0) (1/2, 0, 1/2) (0, 1/2, 1/2)

from a position vector $\boldsymbol{x}$ an equivalent point $\boldsymbol{x}'$ as

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$$

where $\boldsymbol{R}$ is an operation of the corresponding crystal point group. When $\boldsymbol{R}$ is the identity operation (100/010/001) the overall operation is a lattice translation, and $\boldsymbol{t}$ will be integral unless the lattice is centered in some way. Other point symmetry operations $\boldsymbol{R}$ may be combined with fractional translations $\boldsymbol{t}$ to form alternate space group operations (screw axes and glide planes). When the possible combinations of Bravais lattices and crystal point groups are considered, and include the cases where screw axes and glide planes replace the pure rotation axes or mirror planes of the point group, the 230 space groups result. **Table 1** gives the crystal classes, along with, for each class, the lattice parameters required to define the unit cell, the possible centerings and the point groups, classified according to the presence of polarity and chirality. The equivalent points in the various centered lattices are shown in **Table 2**.

Unit Cell Contents

The unit cell of the crystal will normally contain a number of formula units of the constituent substance. Although the chemical formula may be easy to define in the case of elements and synthetic compounds, it may be problematic in the case of biological materials because of their complexity, and also for many minerals. Assuming that the formula mass, M, can be reasonably estimated, the number, z, of formula units per unit cell can be calculated from its dimensions, provided the density of the crystal, ρ, is known, since

$$z = \rho N_{\mathrm{A}}\, U/M$$

where N_{A} is Avogadro's number and U the cell volume.

For macromolecular crystals such as proteins, this process is complicated by the large and uncertain amount of solvent generally present in the crystal and so contributing to M, and by the difficulties in obtaining accurate density measurements.

Once the number of each type of atom is known, the next stage is to describe how the mean atomic positions are distributed within the unit cell. In a simple case, this merely means describing the contents of that normally smaller volume which may be used to generate the complete unit cell through the space group operations. This volume, in mathematical terms the fundamental domain of the space group, is referred to as the asymmetric unit. The asymmetric unit is not normally uniquely defined; however, when rotation or inversion axes are present, they must lie at the borders of the asymmetric unit. (Centers of symmetry $\bar{1}$ and mirror planes $m = \bar{2}$, are special cases of inversion axes.)

The volume of the asymmetric unit may be calculated using data from **Tables 1** and **2**. The number of equivalent points generated by each crystal point group is given in parentheses against it in **Table 1**. The

number of points equivalent by translation is given in **Table 2** for each lattice type. Multiplying these two numbers together will give the factor by which the asymmetric unit is smaller than the unit cell. For example, the trigonal space group $R3c$ has a rhombohedral lattice (triply primitive = 3) and belongs to point group $3m$ (order 6). The unit cell is therefore 18 times larger than the asymmetric unit. This number is termed the multiplicity of the general position of the space group; the explanation of this follows.

We can define another number z', the number of formula units in the asymmetric unit. For many structures $z' = 1$, and each of the atoms of the formula unit occupies a so-called general position within the asymmetric unit. In such a case, each atom's position is described by three arbitrary fractional coordinates (x_i, y_i, z_i). However, various exceptions exist. The most common is that $z' < 1$; this normally means that the positions of at least some of the atoms must coincide with one or more point symmetry elements (in addition to the identity), so that when the corresponding transformations are applied, their coordinates are unchanged. These atoms will lie on the boundaries of the asymmetric unit, and their coordinates will be restricted in some way. Such atoms are said to lie on special positions of the space group.

The various possible special or "Wyckoff" positions are listed, together with their multiplicities and local site symmetries, in International Tables for Crystallography. Where one or two coordinates are constricted, that special position type may be occupied by more than one type of atom (obviously with different free coordinates), whereas only one atom may occupy a special position comprising a set of points.

An example which illustrates many of these points, and where all the possible Wyckoff positions of the space group are used, is that of the isostructural crystals $M_3As_2O_8$ (M = Mg, Co). These have $z = 6$ in space group $I\,\bar{4}2d$. The Wyckoff positions for this space group are shown in **Table 3**; the general position with symbol e has multiplicity 16. This makes $z' = 3/8$ on the basis of the formula, and it immediately follows that elements M and As at least must lie on special positions or be disordered. The number of As atoms in the unit cell (12) cannot be satisfied by the multiplicity of one special position alone, and that of M (18) cannot even be satisfied by any combination of the special positions, which only have multiplicities of 8 or 4, unless some disorder is also present. The way in which these restrictions are resolved in the structure is also shown in **Table 3**, where all five Wyckoff position types are used. The As atoms are tetrahedrally coordinated (site symmetries 2 and $\bar{4}$ are subgroups of the tetrahedral group $\bar{4}3m$), and most of the M atoms are octahedrally coordinated (2 is a subgroup of $m\,\bar{3}m$), with one-ninth of the M atoms disordered among eight coordinate sites of symmetry $\bar{4}$.

On the other hand, for molecular crystals, it is possible for the molecule (formula unit) to itself have some internal symmetry, and that $z' < 1$ indicates that some point symmetry elements of the molecule coincide with the site symmetry of a special position. Depending on the situation, this might not require any individual atoms to lie in special positions. Also, for molecular crystals, cases where $z' > 1$ are not uncommon, and rare cases exist where $z' = 1$ does not correspond to one molecule in a general position, but more than one in distinct special positions.

Crystallographic Databases

The information on an individual structure will be held as a record on a crystallographic database. The main databases are CRYSTMET for metals and minerals, the Inorganic Crystal Structure Database (ICSD) for inorganic structures, the Cambridge Structural Database (CSD) for organic and organometallic structures, and the Protein Data Bank (PDB) for macromolecular structures. These are augmented by the Powder Diffraction File, the Surface Structure Database, and the Nucleic Acid Database for special purposes.

Table 3 Distribution of the atoms over the Wyckoff positions of I $\bar{4}2d$ in the tetragonal $M_3As_2O_8$ phases

Site multiplicity	*Wyckoff symbol*	*Oriented site symmetry*	*Representative coordinates*	*Atoms*	*Occupancy*	*Contribution to cell contents*		
						M	*As*	*O*
16	e	1	(*x*,*y*,*z*)	O(1)	1			16
				O(2)	1			16
				O(3)	1			16
8	d	•2•	(*x*, 1/4, 1/8)	M(1)	1	8		
				As(1)	1		8	
8	c	2••	(0, 0, *z*)	M(2)	1	8		
4	b	$\bar{4}$	(0, 0, 1/2)	M(3)	0.5	2		
4	a	$\bar{4}$	(0, 0, 0)	As(2)	1		4	
					Totals	18	12	48

The information held in a database record (individual crystal structure) will include composition, lattice parameters, and space group, as well as basic details of the experiment and refinement. Typically, the databases as a whole are relational, or are otherwise structured to facilitate a wide range of possible searches. They are regularly updated with new crystal structures, but only after each submission has undergone an extensive validation procedure to ensure consistency and completeness, and to identify duplicate or related entries. Specialized software for search, retrieval, and visualization is made available to subscribers, along with regular updates to the databases themselves.

Applications of the databases fall into two main categories. First, there are searches for a single record or a limited group of records, as an aid in identification, structure description, property prediction, or possibly even structure determination through structure type. However, increasingly the databases are resources for "data mining", the discovery of new information through statistical analysis of the entire database.

The Refinement Model

How to construct the standard refinement model of an inorganic or "small molecule" crystal structure is discussed here. The least-squares refinement, the predominant method, is assumed here. The mean positions of the nuclei are defined by their fractional coordinates – some of which may be constrained for atoms in special positions – and, for X-ray and electron diffraction, their surrounding electron density is assumed to be spherical, and is described by its Fourier transform, the scattering factor. The distribution of the instantaneous nuclear position about its mean is described by the Debye–Waller or temperature factor, either isotropic or anisotropic. The latter form requires six components for an atom in a general position or on a center of symmetry, but will be constrained for other types of special positions.

There will be an overall scale factor (sometimes more, depending on the details of the data collection), in addition to the parameters mentioned above. Atomic site occupancies may be varied to take account of disorder or substitution. Modern refinements may also vary the relative contributions from the various twin orientations which may be present, or the Flack parameter that describes the absolute structure (chirality) of a chiral crystal structure.

However, in evaluating the results of a crystal structure determination, it is important to distinguish those quantities which are determined in an unbiased way by the refinement, and those for which presumed chemical information was applied. In many cases, any attempt to refine all possible parameters is doomed to failure; but this does invalidate the results of a more cautious refinement. In practice, the parameters to be refined must be chosen in such a way as to maintain a well-behaved refinement process. The inclusion of parameters with little impact on the agreement between calculation and observation will produce near-singular least-squares matrices and hence excessive parameter shifts. For this reason chemical information is frequently included in the form of geometric constraints or, increasingly often these days, restraints. A typical constraint is to refine a molecule or group as a rigid body. In the case of restraints, the assumed dimensions of the molecule or group are treated as additional observations with appropriate estimated errors.

In X-ray diffraction, the weakly-scattering hydrogen atoms are normally constrained or restrained in some way, and their precise positions are usually unreliable.

Because of the limited quality of diffraction data from macromolecular crystals, it is not possible to refine such structures using the same sort of least-squares model as for small molecules, and so constraints or restraints are invariably necessary.

An alternative approach to crystal structure determination that is growing in popularity and reliability is that of Rietveld analysis of powder diffraction data.

Measures of Quality

The quality of a structure refinement is generally reported as the conventional R-factor

$$R = \frac{\sum_i |\Delta F_i|}{\sum_i |F_i^o|}$$

This quantity is related to the actual measure normally minimized, the weighted R-factor based on $|F|^2$

$$_2R_w = \frac{\sum_i w_i |\Delta F_i|^2}{\sum_i w_i |\Delta F_i^o|^2}$$

typically being about half. Another measure which may be reported is the goodness-of-fit S,

$$S = \frac{\sum_i w_i |\Delta F_i|^2}{n - m}$$

which should be about unity if the standard errors of the observations are correctly estimated, and the model used is adequate to fit the observations (i.e., the errors in the model are negligible).

The conventional R-factor is a good rough guide to the quality of a structure determination, but has shortcomings in terms of comparing refinements on different crystal structures, because it does not take account of the distribution of structure factors in each particular data set. Data sets of high variance (hypercentric

distributions, superlattices, and other pseudosymmetric structures), tend to refine to higher values of the *R*-factor, both because the higher proportion of very strong and very weak diffraction peaks that cannot be measured with the same precision as the others, and because such data sets are more sensitive to limitations in the model.

Typically, the databases allow some screening of the quality of the individual structure determinations in the search process. The CSD, for example, allows a cut-off maximum to be applied to the conventional *R*-factor, as well as the elimination of structures where disorder is present. It also allows the comparison of structures based on the quality of the derived geometric descriptors (see next section), in its case the average standard error in C–C bond lengths.

Derived Geometric Parameters

Much of the importance of crystal structure determination lies in the information available on bond lengths (i.e., internuclear distances), bond angles, and other geometric descriptors. These may be calculated by standard methods; either the fractional coordinates may be transformed to orthogonal distance coordinates,

$$X = ax$$

$$Y = ax\cos\gamma + by\sin\gamma$$

$$Z = ax\cos\beta + by\sin\beta\cos\alpha^* + cz\sin\beta\sin\alpha^*$$

where α^* is an angle of the reciprocal lattice, or the calculations may be made using the metric tensor for the crystal,

$$G = \begin{pmatrix} a\cdot a & a\cdot b & a\cdot c \\ b\cdot a & b\cdot b & b\cdot c \\ c\cdot a & c\cdot b & c\cdot c \end{pmatrix}$$

The first approach would be more appropriate for simple hand calculations, whereas the second is that used in molecular geometry programs.

It is also worthwhile to examine the validity of the anisotropic temperature factors determined by least-squares. These parameters may be visualized as ellipsoids using the ORTEP program (see **Figure 1**). Very large or distorted thermal ellipsoids on individual atoms may indicate large amplitude internal motion or even short-range static disorder. The TLS (T – translation, L – libration, S – screw correlation) matrix formalism may be used to test whether individual atomic displacements within a molecule or group are consistent with rigid-body motion.

In addition to bond lengths and angles, metal and inorganic structures are frequently described in terms of coordination polyhedra. For organic and organometallic structures, additional descriptors are required to describe the molecular or group conformation; these include torsion angles and the deviation from planarity of a group of atoms.

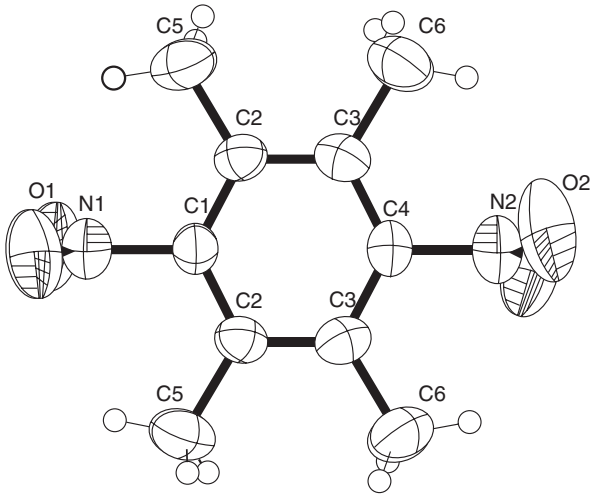


Figure 1 ORTEP drawing of an organic molecule in its crystal structure, showing hydrogen atoms as small spheres and non-hydrogen atoms as 50% probability ellipsoids. The elongation of the oxygen (labeled "O") ellipsoids indicates large amplitude librations about the C–N bonds.

The description of protein crystal structures is somewhat different. Precise bond lengths and angles are not experimentally accessible, but the various levels of organization are, these being:

1. primary structure, the sequence of amino acids present in each chain,
2. secondary structure, the local hydrogen-bonding patterns and chain torsion angles, which result in turn in the typical regular structures (helices and sheets),
3. tertiary structure, the three-dimensional relationships between the secondary structure elements, and
4. quaternary structure, the combination of protein chains to form the overall protein molecule.

See also: Biological Structures; Crystal Structure Determination; Crystal Symmetry; Magnetic Point Groups and Space Groups; Molecular Crystallography; Periodicity and Lattices; Point Groups; Polymer Structures; Protein Folding and Aggregation; Quasicrystals; Structure Types and Classifications.

PACS: 61.50.Ah; 61.66. – f

Further Reading

Allen FH (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallographica* B58: 380–388 (International Union of Crystallography, Chester).

Belsky A, Hellenbrandt M, Karen VL, and Luksch P (2002) New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. *Acta Crystallographica B* 58: 364–369 (International Union of Crystallography, Chester).

Berman HM, Battituz T, Bhat TN, Bluhm WF, Bourne PE, *et al.* (2002) The protein data bank. *Acta Crystallographica* D58: 899–907 (International Union of Crystallography, Chester).

Dunitz JD (1995) *X-ray Analysis and the Structure of Organic Molecules*. Basel: Verlag Helvetica Chimica Acta.

Giacovazzo C (1992) Crystallographic computing. In: Giacovazzo C (ed.) *Fundamentals of Crystallography*, pp. 61–140. Oxford: Oxford University Press.

Giacovazzo C (1992) Symmetry in crystals. In: Giacovazzo C (ed.) *Fundamentals of Crystallography*, pp. 1–60. Oxford: Oxford University Press.

Müller U (1993) *Inorganic Structural Chemistry*. Chichester: Wiley.

Sands DE (1982) Molecular geometry. In: Sayre D (ed.) *Computational Crystallography*, pp. 421–429. Oxford: Clarendon Press.

Stout GH and Jensen LH (1989) *X-ray Structure Determination: A Practical Guide*, 2nd edn. New York: Wiley.

Wells AF (1982) *Structural Inorganic Chemistry*, 5th edn. Oxford: Clarendon Press.

White PS, Rodgers JR, and Le Page Y (2002) CRYSTMET: a database of the structures and powder patterns of metals and intermetallics. *Acta Crystallographica B* 58: 343–348 (International Union of Crystallography, Chester).

Zanotti G (1992) Protein crystallography. In: Giacovazzo C (ed.) *Fundamentals of Crystallography*, pp. 535–597. Oxford: Oxford University Press.

Crystal Structure Determination

R Spagna, Istituto di Cristallografia CNR, Rome, Italy

Introduction

X-rays were discovered in 1895 by W C Roentgen. To ascertain the wave motion of X-rays, M von Laue, stimulated by Ewald's thesis, suggested an experiment where X-rays are made to strike a crystal of copper sulfate. Since a crystal is an arrangement of atoms that repeats itself in three dimensions, it should constitute a three-dimensional diffraction grating. The diffraction pattern obtained clarified definitively that X-rays are an electromagnetic radiation having proper wavelength (in the angstrom range, $\sim 10^{-10}$ m, like the interatomic distances). Then, W H Bragg and W L Bragg determined the first crystal structure, NaCl, which gave rise to crystallographic studies. The interactions between light and the object to be imaged are similar in both visible and X-ray cases. The similarities end when the scattered light is to be combined to form the image; opticians working with visible light have lenses to reconstruct the image. Unfortunately, there is no known way to focus X-rays with a lens. It is only possible to collect the intensities of the diffracted beams and the process of combining them must be done mathematically via Fourier synthesis. Since electrons are the objects that diffract the X-rays, a map of electronic density can be obtained and because the electrons are concentrated around the nuclei of atoms, the maxima indicate the atomic positions. In the Fourier synthesis, the coefficients are amplitudes of the scattered wave which are complex numbers: magnitudes and phases. From the experimental data, only the magnitudes and not the phases can be obtained. Therefore, it is not possible to calculate the electronic density map directly from experimental data: phases must be derived by other means. This is the so-called phase problem in crystallography.

In the following, the specimens that receive attention are single crystals, even if recent studies allow one to determine the crystal structure up to ~ 70 atoms from polycrystalline materials, that is, ensembles of a large number of small crystals.

Crystal Diffraction

A crystal is a periodic arrangement of a motif and it is convenient to think of a set of imaginary points which have a fixed relation in space to the atoms of the crystal. A set of points so formed constitutes a point lattice, which is defined as an array of points, each of which has identical surroundings. The unit cell is the smallest repeating unit that can generate the crystal with only translation operations. The space of the crystal is called "real space" (**Figure 1**).

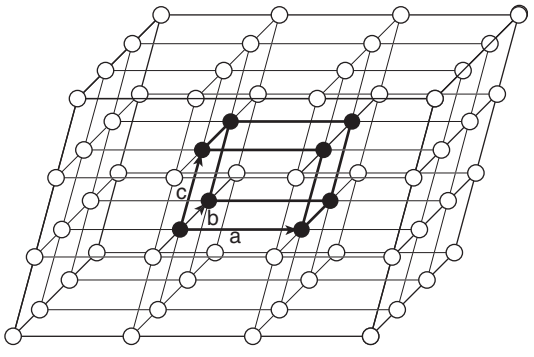


Figure 1 Perspective view of a crystal lattice.

Laue applied general wave-optical principles with a three-dimensional lattice concept to deduce three equations which must be simultaneously satisfied to explain the fact that X-rays are scattered selectively in certain well-defined directions.

The requirement for constructive interference is that the path length difference $(p_1 - p_2)$ between the incoming and the outgoing beams should be an integer number of wavelengths (**Figure 2**):

$$a(\cos\alpha_2 - \cos\alpha_1) = h\lambda$$
$$b(\cos\beta_2 - \cos\beta_1) = k\lambda$$
$$c(\cos\gamma_2 - \cos\gamma_1) = l\lambda$$

where a is the spacing along the crystallographic x axis, α_2 and α_1 are the angles of the outgoing and incoming beams with the x axis respectively, h is an integer; the other two equations have the same meaning.

Bragg saw that the conditions for constructive interference of X-rays were equivalent to that of a family of lattice planes, d_h, described by hkl Miller indices, ($\boldsymbol{h}$), reflecting the X-rays and considering spacing between the planes instead of spacing between atoms (the path length difference $AB + BC$ should be an integer number n of wavelengths):

$$n\lambda = 2d_h \sin\theta$$

where θ is the angle of reflection and n defines the order of reflection from that plane, (Bragg's law) (**Figure 3**).

Ewald gave a beautiful geometric interpretation of the diffraction condition by introducing the concept of "reciprocal lattice." Families of planes, **h**, can be represented by their normals specified as vectors having the lengths, d_h^*, inversely proportional to the d_h. In doing so, a three-dimensional array of points is produced having an inverse relation between the spacing of the points and the planes giving rise to them (**Figure 4**).

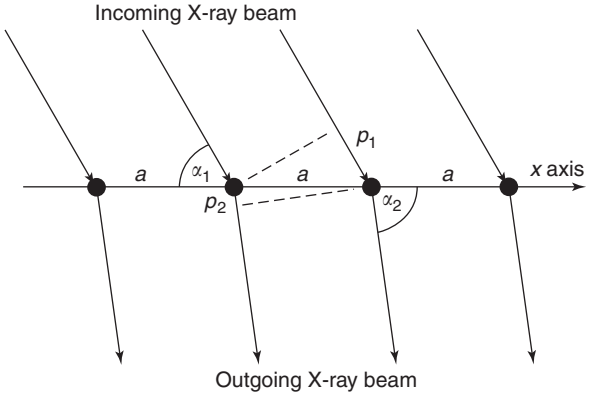


Figure 2 Diffraction from a row of lattice points, called ***x*** axis, with spacing *a*.

Then, he drew a sphere with radius $1/\lambda$ and imagined the crystal to be at C, while the origin of the corresponding reciprocal lattice was at O (**Figure 5**).

A set of (hkl) planes perpendicular to the paper and inclined to the incident X-ray beam XC at the appropriate Bragg angle θ is represented. CP is the direction of the diffracted beam. The perpendicular to OP at P intersects the incident beam at X, and the

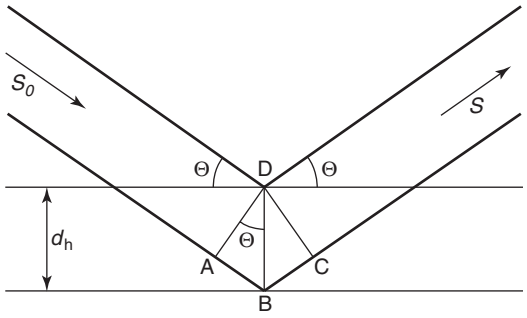


Figure 3 Reflection of X-rays from two lattice planes belonging to the family **h**. $d_{\mathbf{h}}$ is the interplanar spacing.

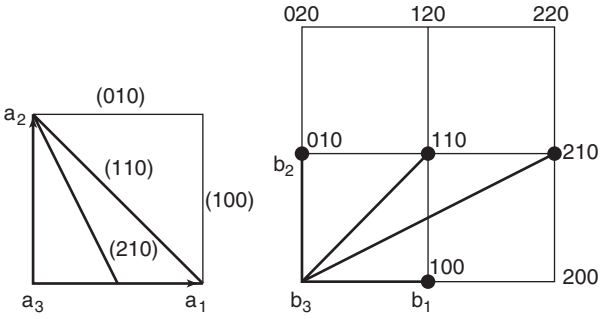


Figure 4 The crystal lattice and the corresponding reciprocal lattice. The axes $\mathbf{a}_3$ and $\mathbf{b}_3$ are normal to the drawing.

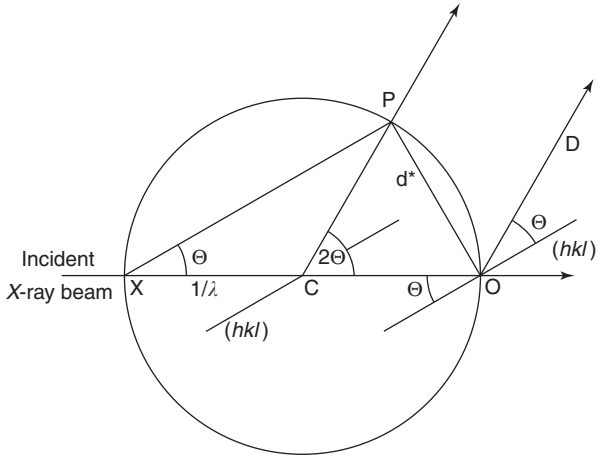


Figure 5 The Bragg equation interpreted in terms of the reciprocal lattice by Ewald.

angle PXO is equal to θ, then

$$\mathrm{PO} = \frac{2}{\lambda}\sin\theta$$

PO is normal to the set of $\boldsymbol{h}$ planes, and following the definition of the reciprocal space:

$$\mathrm{PO} = d^* = \frac{1}{d}$$

then

$$\frac{1}{d} = \frac{2}{\lambda}\sin\theta$$

according to Bragg's law.

To record a diffraction pattern, the crystal at C is rotated and the reciprocal lattice with its origin at O moves jointly. The sphere is stationary and each time a reciprocal lattice point passes through its surface, the Bragg condition for the corresponding set of planes is satisfied. A reflected beam is then produced and its direction is given by the line CP joining the crystal to where the reciprocal lattice point intersects the Ewald sphere.

Experimental Methods

In the laboratory sources, X-rays are produced when a beam of electrons, accelerated by high voltage, strikes a metal target, fixed or evenly rotating. The result is a filtered flux of photons, called a monochromatic beam, having a spectral bandpass, $\delta(\lambda)/\lambda$, very small and picked on the value of the characteristic line of the metal used as a target.

X-rays, as well as other types of electromagnetic radiation, are also generated by synchrotrons. In these installations, either electrons or positrons are accelerated at relativistic velocities and forced by special insertion devices to oscillate around the mean orbit. Some important properties of the radiation produced are the very high flux of photons, a broad continuous spectral range of energy, the narrow angular collimation, and a high degree of polarization. Using particular arrangements, a beam of broad wavelength bandpass, the polychromatic beam, or a beam having a narrow wavelength distribution at the wanted energy, that is, a monochromatic beam, can be selected. An experimental technique which uses a polychromatic beam is the Laue method. It was the first to be used and fell into disuse until a renewed interest by the advent of synchrotrons.

In a diffraction experiment, a crystal having a dimension generally in the range 0.5–0.1 mm is attached to a glass fiber which is fixed onto a pin on a goniometer head (**Figure 6**). The head is mounted onto a goniometer which allows one to set the crystal in different orientations in the X-ray beam. The crystal position is ideally in the center of the Ewald sphere.

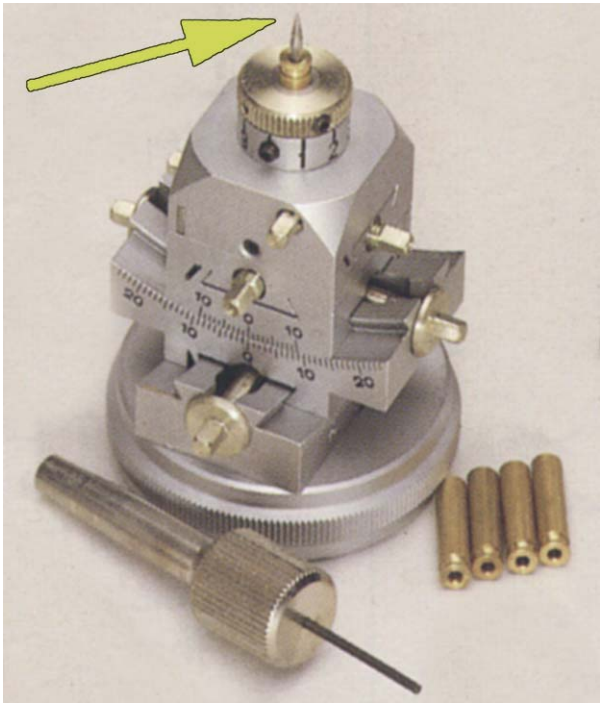


Figure 6 A goniometer head. The arrow shows the location of the crystal. (Reproduced with permission from Huber Diffraktionstechnic, Rimsting, Germany.)

The major differences between the experimental methods reside in the mechanics operating on the crystal and in the detector type. The film-based camera techniques are rotation/oscillation, Weissenberg and precession methods; the single counter or CCD device is used by the single-crystal diffractometer (**Figure 7**).

Nowadays, the oscillation method using electronic area detectors is commonly used in biocrystallography (**Figure 8**), while the diffractometer is widely used for small–medium size molecules.

Irrespective of the method used, two types of experimental data may be derived from measurements of the diffraction pattern: (a) the angles or directions of scattering, which can be used to measure the size and shape of the unit cell, and the corresponding Miller indices of the reflecting planes and (b) the intensities of the diffracted beams, $I_{\boldsymbol{h}}$.

Data Reduction

The amplitude of the scattered wave by a crystal is called the structure factor, $F_{\boldsymbol{h}}$, which is a function of the scattering atoms:

$$F_{\boldsymbol{h}} = \sum_{j=1}^{N} f_j \exp(2\pi \mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}_j)$$

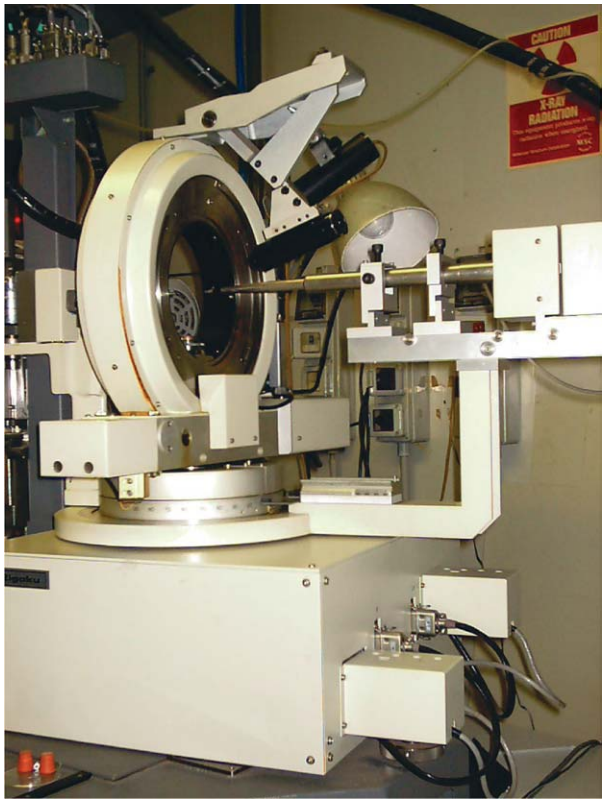


Figure 7 A single crystal 4-circle diffractometer using a rotating anodo X-ray source (Istituto di Cristallografia – Sezione di Monterotondo, Rome, Italy). (Reproduced with permission from Huber Diffraktionstechnic, Rimsting, Germany.)

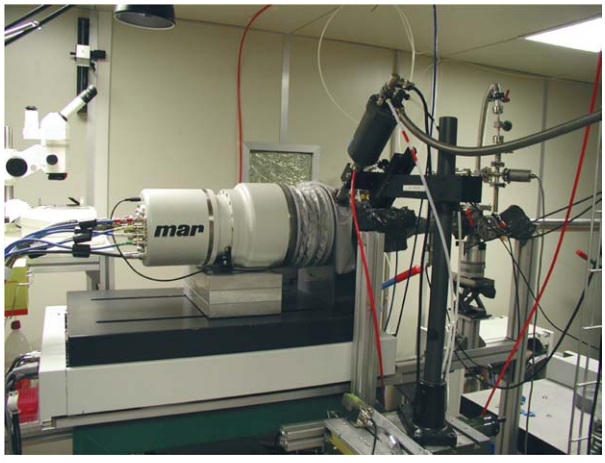


Figure 8 CCD setup on ELETTRA diffraction beam line. (Courtesy of Kristina Djinovic, Structural Biology Laboratory, Sincrotrone Trieste, Italy.)

where f_j is the scattering factor for the jth atom having $\boldsymbol{x}_j$ coordinates in the unit cell containing N atoms. The scattering factor gives the amplitude scattered by an atomic species in terms of the amplitude scattered by a free electron. This means that f depends on the atomic number of the species. The values of atomic scattering factors for neutral atoms are listed in the International Tables for Crystallography for values of $(\sin\theta)/\lambda$. The measured intensity is a function of $|F_h|^2$ and the kinematic theory gives a good approximation of the relationships between I_h and $|F_h|^2$:

$$I_h = KI_0LPTE|F_h|^2$$

where I_0 is the intensity of the incident beam, K takes into account universal and experimental dependent constants, P is the polarization factor, L is the Lorentz factor, T is the transmission factor, E is the extinction coefficient.

Lorentz Correction

Diffraction arises whenever reciprocal lattice nodes cross the Ewald sphere. Depending on the method used to record the reflection intensity and on the position of the reciprocal lattice node, the times required for different nodes to cross the Ewald sphere are different. The Lorentz correction takes into account this difference.

Polarization Correction

The intensity of the diffracted beam changes depending on the state of polarization of the incident beam and on the scattering angle.

Transmission Correction

This correction is related to the absorption of the incident and diffracted beams by the crystal, which depends on the material and the length of the path traveled by the radiation in it. There are analytical methods to calculate this correction which require a precise knowledge of the crystal shape, often very difficult to obtain. An experimental correction is the most widely used and it is designed to correct intensities measured with the diffractometer. It is based on the relative transmission factor plotted as a function of the scanning angle.

Extinction Coefficient

This coefficient depends on the mosaic structure of the crystal, that is, a real crystal can be schematized like a mosaic of very small crystalline blocks tilted very slightly toward each other at small angles, and has two components:

1. the primary one takes into account the loss of intensity due to multiple reflections from different lattice planes. Each scattering causes a phase lag of $\lambda/4$, thus a diffracted radiation is joined by a double scattered radiation with a phase lag of π, consequently destructive interference will result.

2. the secondary one takes into account the fact that the lattice planes first encountered by the primary beam will reflect a significant fraction of the primary intensity so that deeper planes receive less primary radiation. This causes a weakening of the diffracted beam observable for high-intensity reflections at low $\sin\theta/\lambda$ values in sufficiently perfect crystals.

By applying these coefficients to the measured intensities, $|F_h|^2$ is obtained and then:

$$|F_h| = (|F_h|^2)^{1/2}$$

Statistical Analysis

Wilson proposed a simple method to put the observed amplitudes $|F_h|_{obs}$ on an absolute scale, based on a statistical analysis. In fact, he showed that the average of the squared structure amplitudes should be

$$\langle |F_h|^2 \rangle = \sum_{j=1}^{N} f_j^2 = \sigma^2$$

assuming that thermal motion is isotropic and equal for all the atoms,

$$\langle |F_h|^2 \rangle_{obs} = K \langle |F_h|^2 \rangle \exp(-Bs^2)$$

where K is the scale factor and s is equal to $\sin\theta/\lambda$. A plot of the logarithm of the function versus s^2 at some average value of s^2 would derive the values of K and B.

Normalized Structure Factors

Once the intensities have been corrected for thermal motion and placed on an absolute scale, it is simple to obtain the normalized structure magnitudes $|E_h|$ from

$$|E_h|^2 = |F_h|^2/\varepsilon\sigma^2$$

where ε are numbers that vary with the space group and type of reflections. These quantities are independent of the scattering angle and correspond to idealized point atom structures:

$$E_h = \sum_{j=1}^{N} \exp(2\pi i \boldsymbol{h} \cdot \boldsymbol{x})$$

Owing to this peculiarity, the E_h plays a key role in direct method procedures (see below).

Probability Distribution

Hughes empirically, and Wilson by taking the analogy with the random-walk problem found that the X-ray intensities of a centrosymmetric crystal obey a Gaussian distribution. Expressed in terms of $|E|$, the centric distribution is

$$P_{\bar{1}}(|E|) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{|E|^2}{2}\right)$$

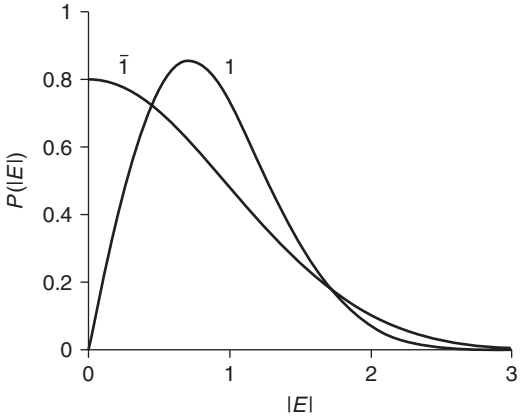


Figure 9 Probability distributions for centro ($\bar{1}$) and for non-centrosymmetric (1) structures.

In further studies, the acentric distribution was obtained (**Figure 9**):

$$P_1(|E|) = 2|E| \exp(-|E|^2)$$

These distributions are completely independent of the structure complexity. Theoretical values of the moments of higher order and of the percentages of $|E| < t$ (where $t = 1,2,3\ldots$) can be calculated from the distributions. These values can be compared with the corresponding experimental values to argue about the presence or absence of a center of symmetry.

The Patterson Function

Patterson introduced a function which is the convolution of electron density $\rho(\boldsymbol{x})$ with itself:

$$P(\boldsymbol{u}) = \int_V \boldsymbol{\rho}(\boldsymbol{x})\boldsymbol{\rho}(\boldsymbol{u}+\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

where V is the volume of the unit cell. The Patterson function can be directly calculated from the set of squared but not phased reflection amplitudes $|F_h|^2$

$$P(\boldsymbol{u}) = \frac{1}{V}\sum_{h} |F_h|^2 \cos 2\pi \boldsymbol{h} \cdot \boldsymbol{u}$$

The map computed gives a vectorial representation of the scattering objects. This depends upon the fact that while F_h is related to the distribution of the atoms in the crystal, $|F_h|^2$ depends on the distribution

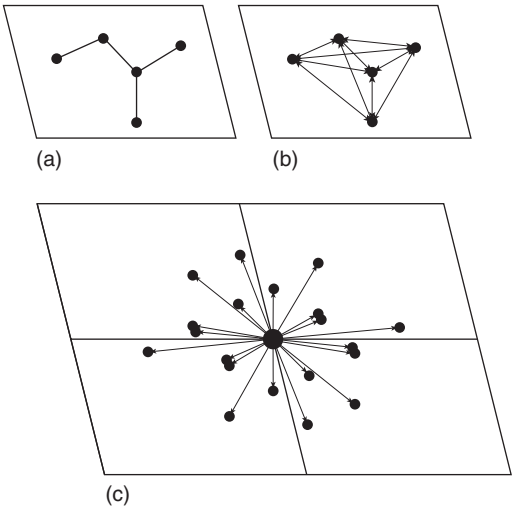


Figure 10 (a) Scheme of a five point atoms molecule, (b) drawing all interatomic vectors, (c) the Patterson function shows all vectors in (b) translated to a common origin.

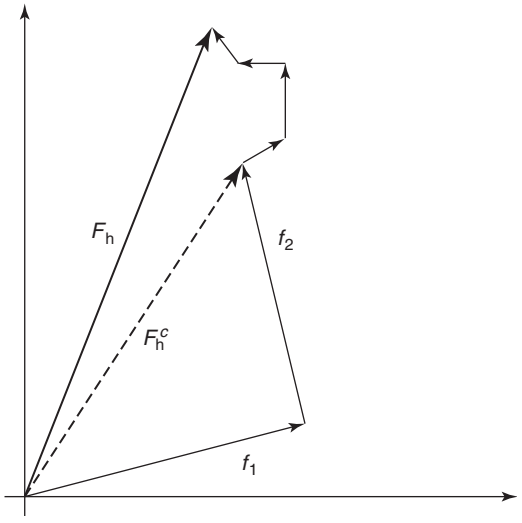


Figure 11 Argand diagram in which two heavy atoms (with atomic scattering factors f_1 and f_2) and four light atoms contribute to the structure factor F_h; the resultant of the contributions of the two heavy atoms (dashed vector) is quite close to F_h.

of interatomic distances. In fact,

$$|F_h|^2 = F_h \cdot F_h^* = \sum_{j=1}^{N} f_j^2 + \sum_{i \neq j=1}^{N} f_i f_j \cos[2\pi \boldsymbol{h} \cdot (\boldsymbol{x}_i - \boldsymbol{x}_j)]$$

depends on the distances $(\boldsymbol{x}_i - \boldsymbol{x}_j)$ between the i and j atoms.

In **Figure 10**, a planar structure of N atoms (a) and the corresponding interatomic vectors (b) are reported. Then this set of vectors has been translated to the origin (c), corresponding to the distribution of $N(N-1)$ peaks in the Patterson function. This higher density of peaks with respect to the electron density makes the Patterson map difficult to interpret even with a moderate number of atoms.

The Heavy Atom Method

Because the density in the Patterson map goes as squares of the numbers of electrons of the scattering atoms, Patterson maps of crystals that contain heavy atoms are dominated by the vectors between them.

Harker noticed that the symmetry of a crystal might provide particular vectors that would lead to a direct measure of the coordinates of individual atoms. For example, if the crystal's symmetry included an inversion center, there would be equivalent atoms at positions (x,y,z) and $(-x,-y,-z)$. The vector between these atoms would be found at $(2x,2y,2z)$ in the Patterson map, indicating the original atomic positions. This procedure greatly simplifies the interpretation of the Patterson map and allows one to obtain a good initial model; in fact, the structure factors F_h^c calculated with the heaviest atoms, f_1 and f_2, represent the predominant contribution to the structure factors F_h (**Figure 11**).

The phase φ_h^c is a good approximation of the true phase of F_h, and an electron density map using the observed structure factors and the calculated phases φ_h^c as coefficients can be computed. The remaining atoms will be found following the so-called method of Fourier synthesis recycling, that is described later.

Direct Methods

Direct methods are those methods that exploit relationships among the intensities to determine the crystal structure directly. They are today the most widely used tool for solving small- and medium-size crystal structures through powerful and sophisticated computer programs, such as *SIR97*, *SHELX*, *SAYTAN*.

In general, the phase, φ, and the amplitude, $|F|$, of a wave are independent quantities and only the amplitudes are obtainable from experiment. However, in X-ray diffraction it is possible to relate these quantities considering two properties of the electron density: (1) it is positive everywhere and (2) it is composed of discrete atoms.

The electron density, $\rho(\boldsymbol{x})$, expressed in terms of Fourier synthesis, is

$$\rho(\boldsymbol{x}) = V^{-1} \sum_{\boldsymbol{h}=-\infty}^{\infty} F_h \exp(-2\pi i \boldsymbol{h} \cdot \boldsymbol{x}) = V^{-1} \sum_{\boldsymbol{h}=-\infty}^{\infty} |F_h| \exp(i_n) \exp(-2\pi i \boldsymbol{h} \cdot \boldsymbol{x})$$

To compute, one needs to know the structure factors $F_{\boldsymbol{h}}$:

$$F_{\boldsymbol{h}} = |F_{\boldsymbol{h}}| \exp(\mathrm{i}\boldsymbol{\varphi}_n) = \sum_{j=1}^{N} f_j \exp(2\pi \mathrm{i}\boldsymbol{h} \cdot \boldsymbol{x}_j)$$

Experimentally, a large number, M, of amplitudes can be measured and then, a system of simultaneous equations formed by the definition of crystal structure factors can be written. The unknown quantities are the phases $\varphi_{\boldsymbol{h}}$ and the atomic positions $\boldsymbol{x}_j$, the known quantities are the $|F_{\boldsymbol{h}}|$ obtained from the measured intensities. Since the equation involves complex quantities, two equations, one for real and one for the imaginary part, can be considered. Therefore, there are $2M$ equations to determine $(M + 3N)$ unknown quantities and since the number of equations exceeds, by far, the number of unknowns, the problem to determine the phases is in principle over-determined.

Sayre derived a basic relationship between structure factors. Considering the structure composed of fully resolved and identical atoms, the two functions $\rho(\boldsymbol{x})$ and $\rho^2(\boldsymbol{x})$ must be very similar and show maxima at the same position (see the one-dimensional structure case reported in **Figure 12**).

The $\rho^2(\boldsymbol{x})$ can be also expressed in terms of Fourier synthesis

$$\rho^2(\boldsymbol{x}) = V^{-1} \sum_{\boldsymbol{h}=-\infty}^{\infty} G_{\boldsymbol{h}} \exp(-2\pi \mathrm{i}\boldsymbol{h} \cdot \boldsymbol{x})$$

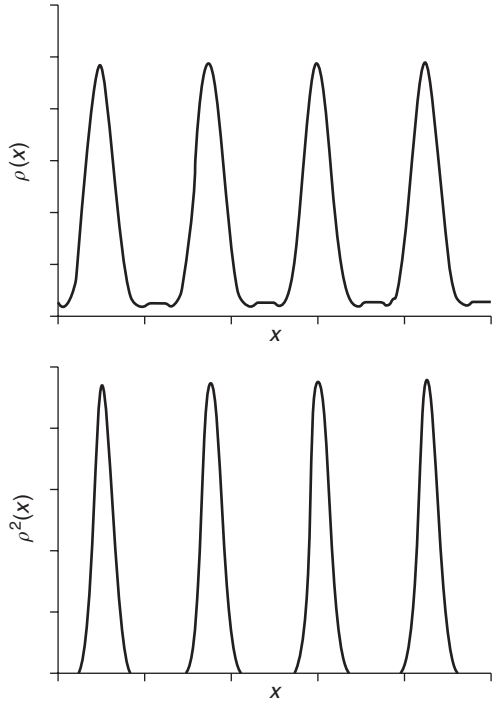


Figure 12 Calculation of $\rho(\boldsymbol{x})$ and $\rho^2(\boldsymbol{x})$ from 1D structure having four equal atoms at equal distance of 2.5 Å along the $\boldsymbol{x}$ axis. The graphics are not in the same scale and show the same position of the peaks.

The mathematical operation which relates the electron density and the structure factors is the Fourier transform, Γ:

$$\Gamma(\rho) = \frac{1}{V} F_{\boldsymbol{h}} = \frac{1}{V} f \sum_{j=1}^{N} \exp(2\pi \mathrm{i}\boldsymbol{h} \cdot \boldsymbol{x})$$

$$\Gamma(\rho^2) = \frac{1}{V} G_{\boldsymbol{h}} = \frac{1}{V} g \sum_{j=1}^{N} \exp(2\pi \mathrm{i}\boldsymbol{h} \cdot \boldsymbol{x})$$

where g is the scattering factor of the squared atom, and because of the convolution theorem, the $\Gamma(\rho^2)$ corresponds to the convolution product of $(1/V)F_{\boldsymbol{h}} * (1/V)F_{\boldsymbol{h}}$ obtaining

$$G_{\boldsymbol{h}} = \frac{1}{V} \sum_{\boldsymbol{k}} F_{\boldsymbol{k}} F_{\boldsymbol{h}-\boldsymbol{k}}$$

From the ratio of the previous equations

$$F_{\boldsymbol{h}} = \frac{f_{\boldsymbol{h}}}{g_{\boldsymbol{h}}} G_{\boldsymbol{h}} = \boldsymbol{\theta}_{\boldsymbol{h}} G_{\boldsymbol{h}}$$

then

$$F_{\boldsymbol{h}} = \frac{\boldsymbol{\theta}_{\boldsymbol{h}}}{V} \sum_{\boldsymbol{k}} F_{\boldsymbol{k}} F_{\boldsymbol{h}-\boldsymbol{k}}$$

which is the Sayre's equation, able to calculate the value of the structure factor $F_{\boldsymbol{h}}$ from the sum of the products $F_{\boldsymbol{k}} F_{\boldsymbol{h}-\boldsymbol{k}}$. Multiplying both sides by $F_{-\boldsymbol{h}}$,

$$|F_{\boldsymbol{h}}|^2 = \frac{\boldsymbol{\theta}_{\boldsymbol{h}}}{V} \sum_{\boldsymbol{k}} |F_{\boldsymbol{h}} F_{\boldsymbol{k}} F_{\boldsymbol{h}-\boldsymbol{k}}| \exp[\mathrm{i}(\boldsymbol{\varphi}_{-\boldsymbol{h}} + \boldsymbol{\varphi}_{\boldsymbol{k}} + \boldsymbol{\varphi}_{\boldsymbol{h}-\boldsymbol{k}})]$$

For large values of $|F_{\boldsymbol{h}}|$, the left-hand side will be large, real, and positive. It is therefore likely that the largest terms in the sum on the right-hand side will also be real and positive. It follows that, if $F_{\boldsymbol{k}}$ and $F_{\boldsymbol{h}-\boldsymbol{k}}$ also have large moduli, the phases are linked by the relationship:

$$\Phi_{\boldsymbol{h}\boldsymbol{k}} = \boldsymbol{\varphi}_{-\boldsymbol{h}} + \boldsymbol{\varphi}_{\boldsymbol{k}} + \boldsymbol{\varphi}_{\boldsymbol{h}-\boldsymbol{k}} \cong 0 (\mathrm{mod}\, 2\pi)$$

This probabilistic condition for three reflections with large structure factors has proved to be the most important original approach for the practical use of direct methods.

Structure Invariants

A structure invariant (s.i.) is a quantity which remains unchanged when the origin is arbitrarily shifted. The intensities are of course the most simple example of structure invariants while the phases are, in general, dependent on the choice of the origin. Under certain conditions, phase relationships between structure

factors are independent of this choice, since their values depend only on the structure and can be estimated. It is easy to show that the most general structure invariant relationship is represented by:

$$F_{h_1}F_{h_2}\cdots F_{h_m} = |F_{h_1}F_{h_2}\cdots F_{h_m}|\exp[\mathrm{i}(\varphi_{h_1}+\varphi_{h_2}+\cdots+\varphi_{h_m})]$$

where

$$h_1+h_2+\cdots+h_m=0$$

Since the moduli of the structure factors are invariants themselves, the angular part:

$$\Phi_m=\varphi_{h_1}+\varphi_{h_2}+\cdots+\varphi_{h_m}$$

is also an s.i.

Probability Methods

Probability methods were introduced by Hauptman and Karle, and led to the joint probability distributions of a set of normalized structure factors on the basis that the atomic coordinates were the primitive random variables uniformly and independently distributed, and the reciprocal vectors were assumed to be fixed. It may also be assumed that the reciprocal vectors are the primitive random variables while the crystal structure is fixed. The probabilistic approach yielded formulas to calculate phases or combination of phases which are s.i.'s by the $|E_h|$ alone. Considering the most important class of s.i's, the three-phase s.i. (triplets), the distribution associated derived by Cochran for a non-centrosymmetric structure is given by:

$$P(\Phi_{hk})=\frac{1}{L}\exp(G_{hk}\cos\Phi_{hk})$$

where L is a normalization factor and for equal atoms

$$G_{hk}=\frac{2}{\sqrt{N}}|E_hE_kE_{-h-k}|$$

where N is the number of atoms in the unit cell.

In **Figure 13**, the probability distributions for different values of the parameter G_{hk} are shown to have a maximum at Φ_{hk} equal to zero, and the variance decreases as G_{hk} increases.

In centrosymmetric crystals, the relation becomes:

$$S(-\boldsymbol{h})S(\boldsymbol{k})S(\boldsymbol{h}-\boldsymbol{k})\cong+$$

where $S(\boldsymbol{h})$ stands for the sign of the reflection $\boldsymbol{h}$ and the symbol $\cong$ stands for "probably equal." In this case, the basic conditional formula for sign determination is given by Cochran and Woolfson:

$$P_+=\frac{1}{2}+\frac{1}{2}\tanh\left(\frac{1}{\sqrt{N}}|E_{-h}E_kE_{h-k}|\right)$$

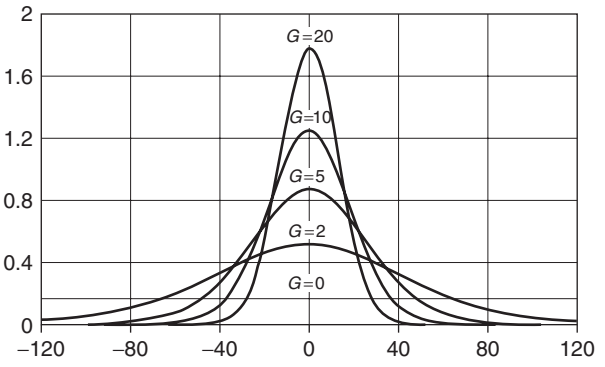


Figure 13 Trends of the probability distributions for different values of the parameter G_{hk}.

The larger the absolute value of the argument of tanh, the more reliable is the sign indication.

For fixed $\boldsymbol{h}$, let the vector $\boldsymbol{k}$ range over the set of known values E_kE_{-h-k}, then the total probability distribution of φ_h is given by the product of the single distributions, increasing the reliability of the estimate, in mathematical terms.

Calculation of the conditional joint probability distributions for φ_h, under the condition that several phases φ_k and φ_{h-k} are known, led to the tangent formula:

$$\tan\varphi_h=\frac{\sum_k|E_kE_{h-k}|\sin(\varphi_k+\varphi_{h-k})}{\sum_k|E_kE_{h-k}|\cos(\varphi_k+\varphi_{h-k})}$$

in which the summations are taken over the same sample of reciprocal vectors $\boldsymbol{k}$.

Formulations of the "nested neighborhood principle" and of the "representation theory" gave the basis to obtain joint probability distributions which improve the estimates for s.i.'s Furthermore, the latter formulated precise general rules for identifying the phasing magnitudes using the space-group symmetry. More complex formulas derived were also able to give indication of Φ_{hk} away from 0°.

Modified tangent techniques are used in almost all computer programs for the phase determination process.

Phase Determination

Once the structure factors are normalized, they are sorted in decreasing order of $|E|$. The s.i.'s are searched in the set of reflections with $|E|$ larger than a given threshold ($\sim$1.3). A typical procedure for phase determination uses a multisolution approach and for each trial, the following steps are performed:

1. assign random values to the phases of the selected reflections.
2. cyclically apply the tangent formula. Since this formula contains information on the positivity of

the electronic density, the phases may converge at the correct values, so defining the structure.
3. for each trial, compute a figure of merit, to estimate its goodness.

The electron density map corresponding to the trial with the highest figure of merit is calculated first.

For each electron density map, a procedure for its interpretation starts involving four steps:

1. peak search;
2. separation of peaks into clusters;
3. application of stereochemical criteria to produce molecular fragments to be processed in subsequent calculations; and
4. labeling of peaks in terms of atomic species. A cyclic procedure consisting of the structure factor and the diagonal least-squares calculations together with a Fourier synthesis provides a way of refining the structure model.

In Progress

Procedures based on direct methods more or less similar to those which have been described are able to solve the crystal structure for small- and medium-size molecules (up to 200 atoms in the asymmetric unit) in a routine way, but fail for bigger molecular structures. An analysis of the problem has been provided by Giacovazzo *et al*. Intuitively, it is seen that the reliability of the probability distributions for s.i.'s is an inverse function of N and as this number increases, the reliability of the relationships decreases whatever the value of the triple product of $|E|$s. The result is that the tangent formula is inadequate for macromolecules.

In spite of this consideration, recent programs – *Shake-and-Bake*, *SHELX-D*, *ACORN*, *SIR2002* – have enlarged the size of crystal structures solvable by direct methods. The key to this success is based on the cyclical use, in both real and reciprocal space, of the refinement of atomic parameters or phases, including the electron density modification procedure. Small proteins (up to 2000 atoms in the asymmetric unit and with data at atomic level resolution) have been solved *ab initio* by these new algorithms, rather expensive in terms of computer time.

Completing the Structure

The model obtained both by Patterson and by direct methods is often incomplete because not all the atoms have been localized. Assuming that the phase φ_h^c of $|F_h^c|$ is a good approximation of the true phase of F_h, an electron density map, using as coefficients the observed amplitudes with the corresponding calculated phases, is computed. This map will reveal

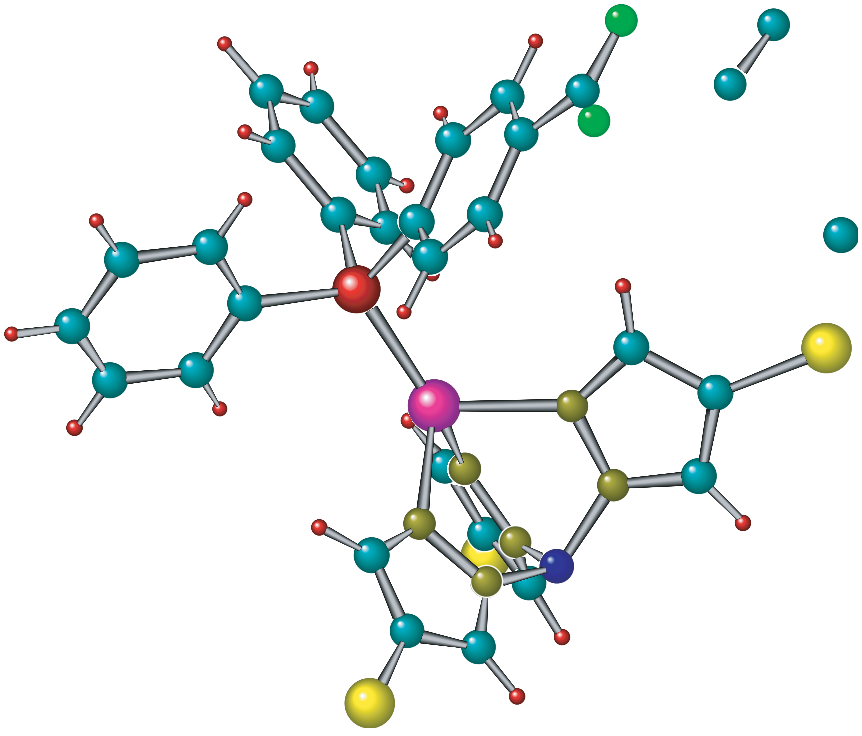


Figure 14 Example of a crystal structure.

new atoms of the structure and then new structure factors can be obtained. The new calculated phases will be used to compute new electron density maps. Each cycle will reveal new atoms until the structure is completed.

Difference Fourier Method

The electron density map

$$\rho_c(\boldsymbol{x}) = \frac{1}{V}\sum_{\boldsymbol{h}} F_{\boldsymbol{h}}^{c} \exp(-2\pi i \boldsymbol{h}\cdot\boldsymbol{x})$$

shows peaks at the positions of the given atoms, while

$$\rho_0(\boldsymbol{x}) = \frac{1}{V}\sum_{\boldsymbol{h}} F_{\boldsymbol{h}}^{0} \exp(-2\pi i \boldsymbol{h}\cdot\boldsymbol{x})$$

where $F_{\boldsymbol{h}}^{0} = |F_{\boldsymbol{h}}^{0}| \exp i\varphi_{\text{true}}$, represents the true structure. To compute the difference electron density map, assume $\varphi_{\text{true}} \approx \varphi_{\boldsymbol{h}}^{c}$:

$$\begin{aligned}\Delta\rho(\boldsymbol{x}) &= \rho_0(\boldsymbol{x}) - \rho_c(\boldsymbol{x}) \\ &= \frac{1}{V}\sum_{\boldsymbol{h}} (|F_{\boldsymbol{h}}^{0}| - |F_{\boldsymbol{h}}^{c}|) \exp(-2\pi i \boldsymbol{h}\cdot\boldsymbol{x} + i_{\boldsymbol{h}}^{c})\end{aligned}$$

If in the model an atom is missing, then $\rho_c(\boldsymbol{x})$ will be zero at the corresponding position, while $\rho_0(\boldsymbol{x})$ will show a maximum. The difference synthesis will also show a peak at the same position but it will be almost zero at the positions of the correct atoms in the model where $\rho_0(\boldsymbol{x}) \approx \rho_c(\boldsymbol{x})$.

Refining the Structure

Least Squares Method

Structure refinement involves "improving" the parameters of the model that give the best fit between the calculated diffraction intensities, F^c, and those observed in the experiment, F^0, to extract precise information about interatomic distances, bond and torsion angles, and other geometric features. One of the most widely applied is the method of least squares in which a weight $w_{\boldsymbol{h}}$ has to be associated to each $F_{\boldsymbol{h}}^{0}$. According to the theory, the quantity:

$$S = \sum_{\boldsymbol{h}} w_{\boldsymbol{h}} (|F_{\boldsymbol{h}}^{0}| - |F_{\boldsymbol{h}}^{c}|)^2$$

has to be minimized. Because the intensities do not depend linearly on the atomic parameters, the function $|F^c|$ has to be expanded in a Taylor series and the normal equations can be derived. The solution of the matrix provides shifts to be applied to the parameters. The procedure will be carried out until convergence is reached. To evaluate the average discrepancies between the calculated model values and the observed values, several statistical descriptors are computed. The most common parameter is the R factor:

$$R = \frac{\sum_{\boldsymbol{h}} (|F_{\boldsymbol{h}}^{0}| - |F_{\boldsymbol{h}}^{c}|)}{\sum_{\boldsymbol{h}} |F_{\boldsymbol{h}}^{0}|}$$

A critical examination of the statistical descriptors informs the crystallographer on the quality of the refined model and can reveal the presence of systematic errors or model inadequacy. Generally, for small- or medium-size structures, assuming good quality crystals, one should expect R values in the range 0.03 – 0.05 (**Figure 14**).

See also: Crystal Structure; Crystal Structure Determination, History of.

PACS: 7.85.Jy; 61.10. – i; 61.10.Nz; 61.50. – f

Further Reading

Bragg WL (1913) The diffraction of short electromagnetic waves by crystal. *Proceedings of the Cambridge Philosophical Society* 17: 43–57.

Bragg WH and Bragg WL (1913) The reflection of X-rays by crystals. *Proceedings of the Royal Society of London (A)* 88: 428–438.

Buerger MJ (1959) *Vector Space and Its Application in Crystal-Structure Investigation*. New York: Wiley.

Cruickshank DWJ (1965) *Computing Methods in Crystallography*, pp. 112–116. Oxford: Pergamon.

Ewald PP (1921) Das 'reziproke Gitter' in der Structur theorie. *Zeitschrift fur Kristallographie (A)* 56: 129–156.

Friedrich W, Knipping P, and von Laue M (1912) Interferenz-Erscheinungen bei Roentgenstrahlen. In: *Sitzungsberichte der Mathematisch-Physikjalischen Klasse der Koeniglich Bayerischen Akademie der Wissenshaften zu Muenchen*, 303–322 (reprinted in *Naturwiss*. (1952) 361–367).

Giacovazzo C (1998) *Direct Phasing in Crystallography – Fundamentals and Applications*. Oxford: Oxford University Press.

Giacovazzo C (2002) *Fundamentals of Crystallography*, 2nd edn. Oxford: Oxford University Press.

Hammond C (2001) *The Basics of Crystallography and Diffraction*, 2nd edn. Oxford: Oxford University Press.

Hauptman HA and Karle J (1953) Solution of the phase problem. I. The centrosymmetric crystal. *ACA Monograph*, no.3. Ohio: Polycrystal Book Service.

IUCr, *International Tables for Crystallography* (1996) vol. B, Corr. reprint; (1999) vol. C, 2nd edn. Dordrecht: Kluwer Academic.

von Laue M (1912) Eine quantitative Prüfung der Theorie fur die Interferenz–Erscheinungen bei Roentgenstrahlen. In: *Sitzungsberichte der Mathematisch-Physikalischen Klasse der Koeniglich Bayerischen Akademie der Wissenshaften zu Muenchen*, 363–373 (reprinted in *Naturwiss*. (1952) 368–372).

Patterson AL (1934) A Fourier series method for the determination of the components of interatomic distances in crystals. *Physical Review* 46: 372–376.

Prince E (1994) *Mathematical Techniques in Crystallography and Material Science*, 2nd edn. Berlin: Springer.

Rollett JS (1965) *Computing Methods in Crystallography*. Oxford: Pergamon.

Sayre D (1952) The squaring method: a new method for phase determination. *Acta Crystallographica* 5: 60–65.
Schwarzenbach D, Abrahams SC, Flack HD, Gonshorek W, Hahn T, *et al.* (1989) Statistical descriptors in crystallography. Report of International Union of Crystallography Subcommittee on Statistical Descriptors. *Acta Crystallographica A* 45: 63–75.
Sheldrick GM (1998) SHELX applications to macromolecules. In: Forties S (ed.) *Direct Methods for Solving Macromolecules Structures*, pp. 401–411. Dordrecht: Kluwer Academic.
Wilson AJC (1942) Determination of absolute from relative X-ray intensities data. *Nature* 150: 151–152.

Crystal Structure Determination, History of

D Viterbo, Università del Piemonte Orientale "A. Avogadro," Alessandria, Italy

Introduction

The basic principles for obtaining accurate structural information from diffraction data have been outlined elsewhere in this encyclopedia. It has been shown that X-rays allow one to "see" the electron density distribution in a crystalline medium when the diffracted amplitudes are properly combined. These amplitudes, also called "Structure factors," are complex quantities and must be defined both in magnitude and phase, but unfortunately, because of the high energy of X-rays, in almost all diffraction experiments the phase information is lost, giving rise to the "phase problem in crystallography." The whole history of crystal structure determination is a fascinating account of the different mathematical and physical attempts at overcoming the phase problem. The success of these efforts is testified by the significant number of Nobel prizes awarded to the scientists who have contributed to this subject.

The main steps of the story will be described and it will be seen how, from the initial determination of the crystal structure of NaCl by W H and W L Bragg, structures of very complex biological assemblies such as viruses and ribosomes are solved nowadays. The focus is mainly on the methods concerning single crystals or polycrystalline materials, but mention will also be made of less ordered materials, such as fibers or membranes.

Basic Concepts

From the measured intensities of the beams diffracted by a crystal, what can only be derived are the magnitudes of the complex "structure factors," which are related to the atomic positions $\boldsymbol{x}_j = (x_j, y_j, z_j)$ of the N atoms in the crystal unit cell (i.e., to the structure) by the relation

$$F_{\boldsymbol{h}} = \sum_{j=1}^{N} f_j \exp(2\pi i \boldsymbol{h} \cdot \boldsymbol{x}_j) = |F_{\boldsymbol{h}}| \exp(i\varphi_{\boldsymbol{h}}) \qquad [1]$$

where f_j is the amplitude scattered by each atom, and the vector $\boldsymbol{h} = (hkl)$ indicates the diffraction direction in terms of the Miller indices of the crystal planes originating the corresponding Bragg reflection. The intensity $I_{\boldsymbol{h}}$ is related to $|F_{\boldsymbol{h}}|^2$ and the phase $\varphi_{\boldsymbol{h}}$ is lost in the squaring process. For centrosymmetric structures, the structure factors become real quantities and the phase reduces to a sign.

The electron density function within the crystal unit cell is related to the structure factors by the Fourier synthesis:

$$\begin{aligned}\rho(\boldsymbol{x}) &= \frac{1}{V}\sum_{\boldsymbol{h}} F_{\boldsymbol{h}} \exp(-2\pi i \boldsymbol{h} \cdot \boldsymbol{x}) \\ &= \frac{1}{V}\sum_{\boldsymbol{h}} |F_{\boldsymbol{h}}| \exp(i\varphi_{\boldsymbol{h}}) \exp(-2\pi i \boldsymbol{h} \cdot \boldsymbol{x}) \qquad [2]\end{aligned}$$

where V is the volume of the unit cell.

The Early Times and the Trial and Error Methods

After Laue's discovery (in 1912) of diffraction of X-rays by crystals (for which he was awarded the Nobel Prize in physics in 1914), in 1913 W L Bragg reformulated the theory of diffraction in terms of reflection by crystal planes, and, together with his father W H Bragg, he initiated the era of crystal structure determination.

W H and W L Bragg (**Figure 1**) were awarded the Nobel prize in physics in 1915, but because of the war, W L Bragg could only deliver the Nobel lecture in 1922. The reading of this lecture is fascinating as one is taken through the simultaneous interpretation of both X-ray spectra and crystal structures (with almost no information on atomic structure and dimensions), using the ***X-ray spectrometer***, the ancestor of modern diffractometers. "In this instrument X-rays coming from a tube are limited to a narrow beam by

Figure 1 Swedish postage stamp for the Nobel prize to W H and W L Bragg.

slits, and fall on a crystal at the center of the spectrometer table by which they are reflected; the reflected beam is received and measured in an ionization chamber."

The steps required to analyze a crystal structure by X-rays are clearly outlined:

- "Find the dimensions of the unit cell of the space lattice."
- "Calculate haw many atoms or molecules are contained within it, using as data the density of the crystal and the mass of each molecule."
- "Determination of the manner in which the atoms are grouped together to form each unit of the structure. It is here that the principal difficulty is experienced."

In order to overcome this difficulty, which is nothing but the phase problem, the essential ingredients of the so-called "trial and error" methods are outlined. "The structure of the group of atoms which composes the unit of the crystal grating influences the strength of the various reflections …. The rays are diffracted by the electrons grouped around the center of each atom. In some directions the atoms conspire to give a strong scattered beam; in others, their effects almost annul each other by interference. The exact arrangement of the atoms is to be deduced by comparing the strength of the reflections from different faces and in different orders."

The space group symmetry can be derived from the diffraction pattern (as is described elsewhere in this encyclopedia). For the simplest structures belonging to high symmetry space groups, such as NaCl and KCl, the atomic positions are uniquely defined by symmetry. For less simple structures, plausible atomic arrangements can be deduced on the basis of symmetry and of physical and chemical properties, and then, each hypothesis tested against the observed intensities: for each structural hypothesis, $|F_h|_{\text{calc}}$ can be calculated for all measured reflections using eqn [1] and compared with $|F_h|_{\text{obs}}$ derived from the corresponding observed intensities. An overall agreement factor

$$R = \frac{\sum_h ||F_h|_{\text{obs}} - |F_h|_{\text{calc}}|}{\sum_h |F_h|_{\text{obs}}} \qquad [3]$$

is then evaluated. A low value of R (say $R < 0.4$) will indicate a plausible structural model. This model will then be further validated by its ability of being completed and refined by the methods described elsewhere in this encyclopedia.

A detailed account of the "trial and error" methods is given in the book by H Lipson and W Cochran. These methods require great effort, ingenuity, and skill and were only used with simple structures. They are seldom used today to carry out standard structure analyses, but may still be used to obtain structural information when the diffraction pattern is less well-defined. Here, just two examples will be given. The first, in the realm of inorganic chemistry, is the determination of the structure of potassium alum $KAl(SO_4)_2 \cdot 12H_2O$ by Cork (1927) and C A Beevers and H Lipson (1934, 1935). As first determined by R W G Wyckoff (1923), the crystals belong to the centrosymmetric cubic space group $Pa\bar{3}(T_h^6)$, for which the equivalent positions are

	$x,y,z;\quad z,x,y;\quad y,z,x$	
	$\frac{1}{2}+x,\frac{1}{2}-y,\bar{z};\quad \frac{1}{2}+z,\frac{1}{2}-x,\bar{y};\quad \frac{1}{2}+y,\frac{1}{2}-z,\bar{x}$	
	$\bar{x},\frac{1}{2}+y,\frac{1}{2}-z;\quad \bar{z},\frac{1}{2}+x,\frac{1}{2}-y;\quad \bar{y},\frac{1}{2}+z,\frac{1}{2}-x$	
	$\frac{1}{2}-x,\bar{y},\frac{1}{2}+z;\quad \frac{1}{2}-z,\bar{x},\frac{1}{2}+y;\quad \frac{1}{2}-y,\bar{z},\frac{1}{2}+x$	
24 *d* 1	$\bar{x},\bar{y},\bar{z};\quad \bar{z},\bar{x},\bar{y};\quad \bar{y},\bar{z},\bar{x}$	
	$\frac{1}{2}-x,\frac{1}{2}+y,z;\quad \frac{1}{2}-z,\frac{1}{2}+x,y;\quad \frac{1}{2}-y,\frac{1}{2}+z,x$	
	$x,\frac{1}{2}-y,\frac{1}{2}+z;\quad z,\frac{1}{2}-x,\frac{1}{2}+y;\quad y,\frac{1}{2}-z,\frac{1}{2}+x$	
	$\frac{1}{2}+x,y,\frac{1}{2}-z;\quad \frac{1}{2}+z,x,\frac{1}{2}-y;\quad \frac{1}{2}+y,z,\frac{1}{2}-x$	
8 *c* .3.	$x,x,x;\ \frac{1}{2}+x,\frac{1}{2}-x,\bar{x};\ \bar{x},\frac{1}{2}+x,\frac{1}{2}-x;\ \frac{1}{2}-x,\bar{x},\frac{1}{2}+x$	
	$\bar{x},\bar{x},\bar{x};\ \frac{1}{2}-x,\frac{1}{2}+x,x;\ x,\frac{1}{2}-x,\frac{1}{2}+x;\ \frac{1}{2}+x,x,\frac{1}{2}-x$	
4 *b* .$\bar{3}$.	$\frac{1}{2},\frac{1}{2},\frac{1}{2};\quad \frac{1}{2},0,0;\quad 0,\frac{1}{2},0;\quad 0,0,\frac{1}{2}$	
4 *a* .$\bar{3}$.	$0,0,0;\quad 0,\frac{1}{2},\frac{1}{2};\quad \frac{1}{2},0,\frac{1}{2};\quad \frac{1}{2},\frac{1}{2},0$	

The unit cell contains 4 K, 4 Al, 8 S and 32 O atoms and 48 water molecules. The K and Al atoms should then be located in the fixed positions on the inversion centers *a* and *b*, with multiplicity 4. Locating the K atoms in *a* and the Al atoms in *b* or vice versa corresponds to the choice between two equivalent origins of the unit cell and is arbitrary. The eight S atoms must occupy the eightfold position *c*, but they are not definitely located since the *x* parameter (along the threefold axis) is not defined. The 32 O atoms cannot all be equivalent, since there is no position with a 32-fold multiplicity; they must occupy one general *d* position and one special *c* position. This choice is compatible with the known tetrahedral stereochemistry of the SO_4 group. Similarly, the most likely position of the 48 oxygen atoms (hydrogen atoms are not considered) of the water molecules (W)

is on two sets of general positions *d*. The most probable arrangement of atoms is thus:

- K and Al atoms on fixed special positions *a* and *b*;
- S atoms on special positions *c* with one parameter;
- O atoms on a set of special positions *c* with one parameter and a set of general positions *d* with three parameters; and
- W atoms on two sets of general positions *d* with six parameters.

This is a striking instance of the usefulness of space-group symmetry, which allows the simplification of the problem of determining the position of 96 atoms to that of determining just 11 parameters. Stereochemical considerations can further reduce the number of parameters, and other considerations, such as atomic radii and known bond distances and angles can further help in defining the most likely atomic arrangements to be tested against the observed diffraction intensities.

The second example, in the realm of organic chemistry, describes the structure determination of some aromatic compounds. As pointed out by W H Bragg (1929), "the major interest in the attempt to solve an organic structure lies in the determination of the position of the atoms in the molecule. This is a very difficult problem: but there are clear signs that the difficulties are yielding." The structure of hexamethylbenzene determined by K Lonsdale (1928, 1929) is the first case of a structure solved "with satisfactory completeness and certainty." On the basis of the available knowledge of the structure of graphite (H Ott, 1928) and of a careful analysis of the measured intensities, a good model of the structure was obtained, which gave a satisfactory agreement between observed and calculated structure factors. Some preliminary results on the structure of the aromatic rings had been obtained by W H Bragg (1921) from the study of the crystals of naphthalene and anthracene, but their correct structures were only determined later by K Banerjee (1930). The solution of the structure of hexamethylbenzene established that: (1) the molecules exist as separate entities in the crystalline state, (2) the benzene ring is hexagonal in shape and is planar, (3) the approximate ring C–C distance is 1.42 ± 0.03Å, and the C–CH_3 distance is 1.54 ± 0.12Å, (4) the ring possesses a center of symmetry and so the Kekulé static model with its three fixed double bonds must be excluded. Only the determination of the crystal structure could answer these fundamental questions, which are now basic concepts in organic chemistry.

Patterson and Fourier Methods

Without the phase information, it is impossible to calculate the electron density function $\rho(\boldsymbol{x})$ using eqn [2]. In order to overcome this difficulty, A L Patterson (1934, 1935) proposed to use a Fourier series similar to eqn [2] in which the coefficients were the squares of the structure factor moduli, $|F_h|^2$, directly derivable from the measured intensities. The Patterson function is the self-convolution of $\rho(\boldsymbol{x})$, and its maxima give information about the length and orientation of all the N^2 interatomic vectors linking the N atoms in the unit cell. In his 1935 paper, Patterson proved that some structures (KH_2PO_4, C_6Cl_6, and $CuSO_4 \cdot 5H_2O$; **Figure 2**) could be solved by his method.

The following step was the discovery by D Harker (1936) that the space group symmetry was reflected in the concentration of Patterson peaks on certain planes or lines. M J Buerger's (1946) "implication diagrams" and the "superposition methods" first proposed by D M Wrinch (1939) are general methods for recovering the atomic positions from a vector set and are described in the book by Buerger (1959). Although several fairly complex structures, composed of almost equal light atoms, have been solved from their Patterson functions, when the number of atoms increases, the number of interatomic vectors becomes very large and the Patterson map becomes almost featureless. However, if the structure contains some heavy atoms, prominent peaks corresponding to vectors between these atoms dominate the map. This is the basis of the very successful "heavy atom methods," which have since been used for the solution of thousands of crystal structures. One of the first cases is the solution of the structure of $NiSO_4 \cdot 7H_2O$ by C A Beevers and C M Schwartz (1935); the diffraction data from the "isomorphous" $MgSO_4 \cdot 7H_2O$ crystals were also used to help the solution. As will be seen, these methods turned out to be the key to the solution of protein structures. The great power of the heavy atom methods was shown by the solution of the large molecule of vitamin B12, for which D Crowfoot Hodgkin (**Figure 3**) won the 1964 Nobel prize for chemistry. Another important effect related to the presence of heavy atoms is that of "anomalous dispersion," which can be used not only to obtain phase information, but also to reveal the absolute configuration of a chiral compound (J M Bijvoet 1949).

Fourier methods are generally used for completing and refining the models obtained in the initial stages of structure solution, but in some special cases they can also be used in crystal structure determination. One of the first examples was the determination of the rather complex structure of platinum phthalocyanine by J M Robertson and I Woodward (1940), in which, because of symmetry, the heavy Pt atom must be located at the origin on the inversion center. The

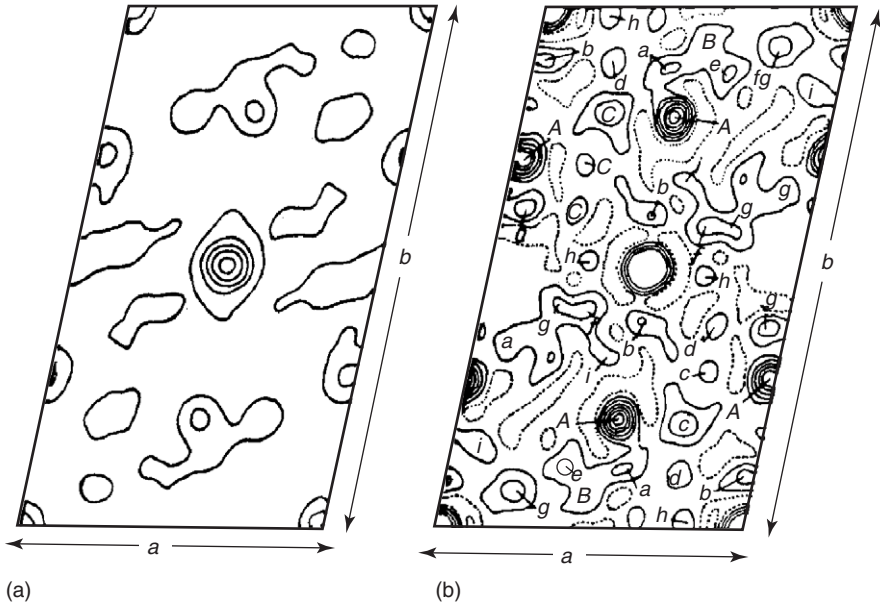


Figure 2 (001) projection of the Patterson function of $CuSO_4 \cdot 5H_2O$: (a) computed with $|F_h|^2$ coefficients; (b) computed with modified coefficients to increase resolution of peaks.

Figure 3 Dorothy Hodgkin and J D Bernal (in the center) with I Fankuchen and D Fankuchen.

(010) electron density projection (**Figure 4**) computed with all positive signs for the (*h*0*l*) reflections, as implied by the Pt position, revealed the light carbon and nitrogen atoms of the phthalocyanine molecule.

Direct Methods

Trial and error and Patterson methods require a great deal of human intervention and a search for more direct methods to derive the crystal structure from the diffraction data started quite early on, when H Ott (1927), K Banerjee (1933), and M Avrami (1938) proposed some algebraic methods for the solution of the structure factor equations as given in eqn [1]. Unfortunately these methods were of very limited practical use because they could only be applied to very simple cases.

The initial step toward the present direct methods (DM) is usually regarded as the work of D Harker and J S Kasper (1948), which appeared on the first issue of *Acta Crystallographica*, the journal of the newly born International Union of Crystallography. Their inequality relations between structure factors played an important part in the determination of the structure of decaborane, $B_{10}H_{14}$, by J S Kasper, C M Lucht, and D Harker (1950). J Karle and H A Hauptman (1950), by imposing the physical condition that the electron density is positive or zero everywhere, were able to generalize the inequalities into a determinantal form. The next step was the derivation by D Sayre (1952) of an equality relation between structure factors for structures composed of equal and fully resolved atoms. In 1953, Hauptman and Karle established the basic concepts and the probabilistic foundations of DM and derived formulas for the estimate of some linear combinations of phases called "structure invariants"; their results were collected in a monograph with the far-seeing and ambitious title *Solution of the Phase Problem. I. The Centrosymmetric Crystal* (**Figure 5**).

The theoretical development of DM continued for more than ten years, with important contributions also by other authors (among which W Cochran, W H Zachariasen, E F Bertaut, M M Woolfson), but the practical applications were very few. Indeed, until

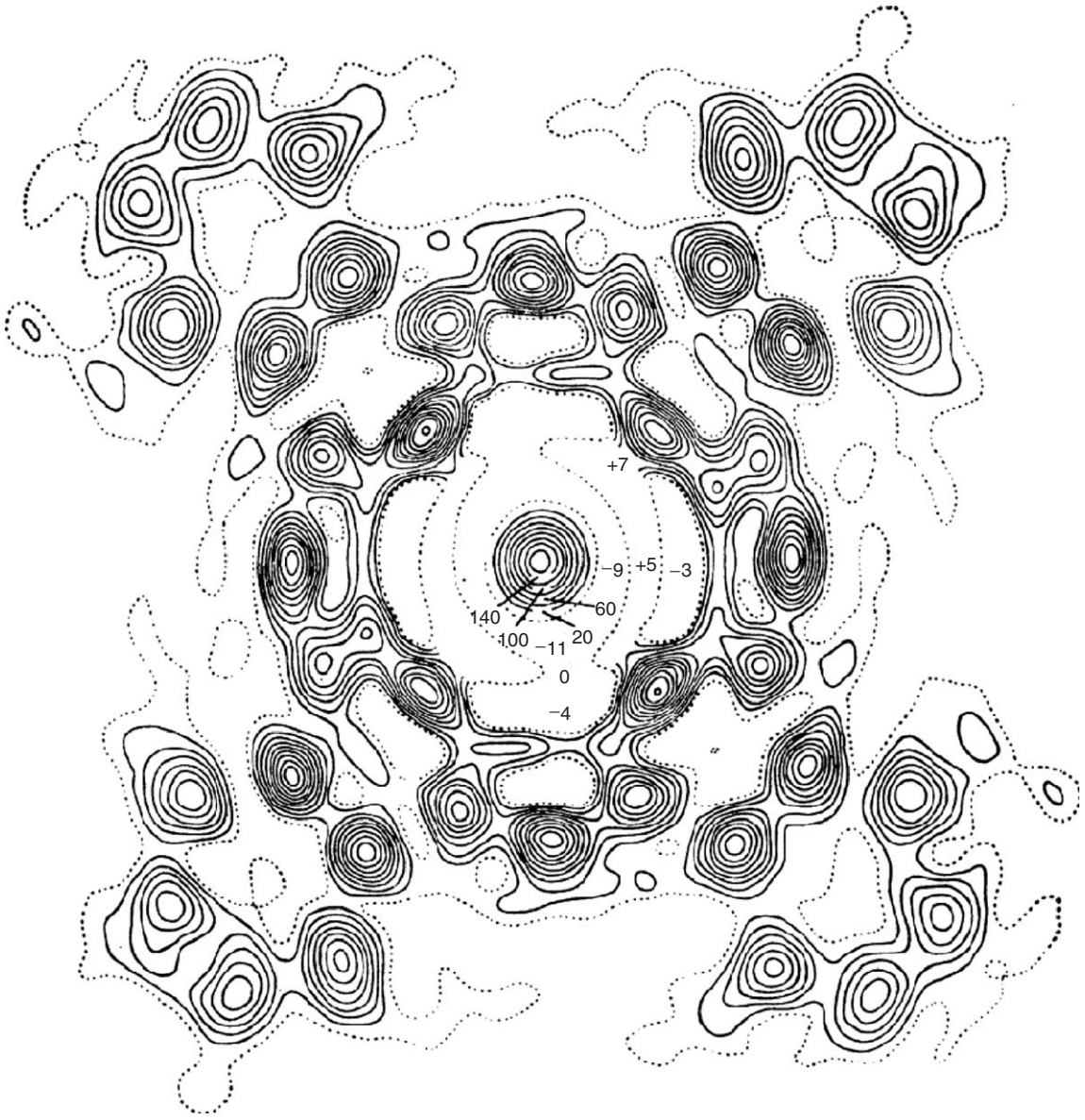


Figure 4 (010) projection of the electron density of Pt-phthalocyanine.

1967, less then 10% of the structures were solved by DM, and only the availability of faster computers and suitable programs triggered the fast growth of their use. By the mid-1970s, this exceeded 50%. The theoretical work became more strictly linked to the development of practical procedures. In 1966, J Karle and I Karle formulated the basic rules for the practical application of the "symbolic addition" methods, first proposed by W H Zachariasen (1952), and in 1968 G Germain and M M Woolfson proposed the "multisolution" methods, which were the basis for the very successful MULTAN program (G Germain, P Main and M M Woolfson, 1970). The continuous progress and success of DM in solving increasingly complex crystal structures was the motivation for conferring in 1985 the Nobel prize in chemistry on the mathematician H Hauptman and the physicist J Karle. The theoretical and practical developments of DM are still continuing and today they are also capable of solving small proteins from diffraction data at atomic resolution.

Polycrystalline Materials

So far only the use of X-ray diffraction by single crystals has been considered, but it is well known that many materials cannot be easily obtained in this form, but only as polycrystalline samples composed of very small crystallites (a powder). Very early on, W Friedrich (1913), P Debye and P Scherrer (1916), and

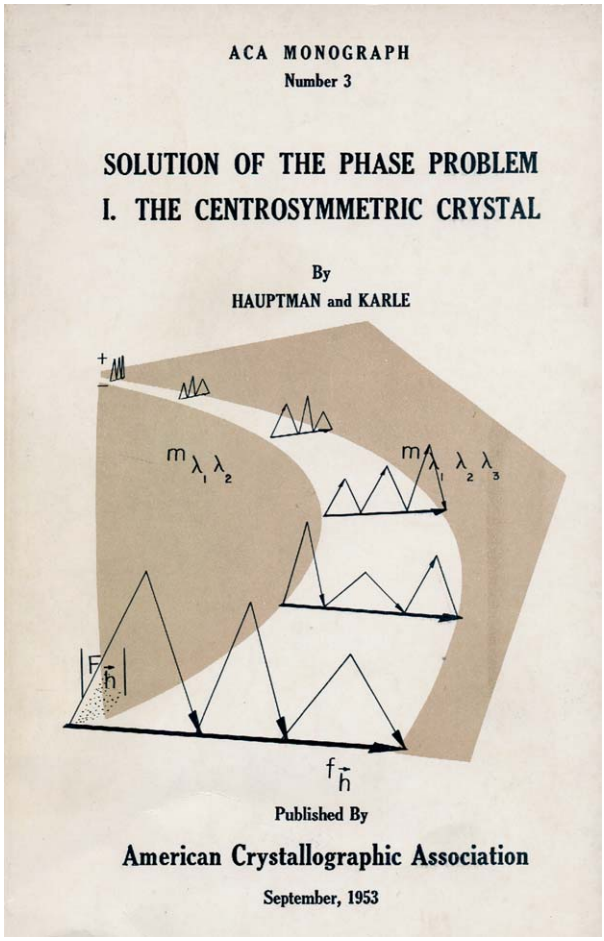


Figure 5 Cover of the 1953 monograph by H Hauptman and J Karle.

A W Hull (1917) obtained diffraction diagrams from powder samples and gave their theoretical interpretation. The diffraction pattern of every polycrystalline sample is like its fingerprint, and powder patterns soon became a widely used analytical method for the identification of crystalline species. Unfortunately in powder patterns, different reflections tend to overlap and this problem becomes more severe for more complex structures. The derivation of the indices and intensity of the individual reflections, of the cell parameters, and of the space group symmetry becomes much more difficult. In more recent years, several algorithms for the analysis of diffraction patterns were proposed and implemented in computer programs, but the most relevant step was the use of the full profile fitting method proposed by H M Rietveld (1969). These theoretical developments, together with the advent of more accurate instrumentation and the use of synchrotron radiation, have opened the way to the *ab initio* solution of fairly complex crystal structures (both by Patterson and by DMs) of compounds which cannot be obtained as suitable single crystals.

Impact of the Development of Experimental and Computational Techniques

The progress of crystal structure determination was also marked by the developments in X-ray sources and detectors. In the early days sealed tubes were used, and the most common detector was a photographic film on which intensities were usually estimated by a painstaking visual comparison with a suitable scale. Around the 1960s, diffractometers with counter detectors appeared and for special applications, they were equipped with the more intense X-ray source of a rotating anode generator. The most recent advances are the advents: (1) of the very intense, collimated, and tuneable synchrotron radiation, (2) of powerful focusing devices, and (3) of electronic area detectors.

In the early days the computational effort needed to carry out a crystal structure determination appeared quite formidable and almost impossible to be carried out by hand. The summation of a 3D Fourier series such as eqn [2], involving hundreds or thousands of terms, could not be conceived. Only 2- and 1D projections involving fewer terms could be computed by some ingenious devices such as Lipson and Beevers (1936) strips or some optical methods. The advent of punched-card machines in the 1940s and their rapid development into electronic computers also marked a breakthrough in crystal structure determination. In the early 1960s, huge machines using punched cards or paper tapes and magnetic tapes were employed; their computing power was less than that of a modern palm-top computer and the calculation of a 3D electron density map could take a few hours (a few seconds on modern PC). As already mentioned, the advent of faster computers allowed and triggered the formulation of more demanding algorithms, and of the corresponding software for the solution and refinement of crystal structures of increasing complexity. Additionally, information technology allowed the setting up of structural databases collecting almost half a million structures; their search is an invaluable source of physical, chemical, and biological information.

Solution of Biological Macromolecules

Parallel to the development of the structural crystallography of small and medium size molecules, it was soon recognized that X-ray diffraction methods could provide detailed structural information also for large biological macromolecules. Indeed, globular proteins can often be crystallized and these crystals, although usually with a poorer degree of order, have the same regularities in external form and the same

optical characteristics as the crystals of simple substances. In 1934 J D Bernal and D Crowfoot, and in 1935 D Crowfoot obtained the first X-ray single crystal photographs of pepsin and insulin. The patterns contained thousands of reflections, as expected for crystals of large molecules. The route was opened, but these studies also emphasized the difficulty of the task. In 1936 M Perutz joined Bernal's lab and was soon inspired by his "visionary faith in the power of X-ray diffraction to solve the structures of molecules as large and complex as enzymes or viruses at a time when the structure of ordinary sugar was still unsolved." This was the beginning of his lifelong dedication to the study of crystalline hemoglobin. After a long battle in trying some very ingenious but unsuccessful methods to derive phase information, in 1953, triggered by the recent success in binding mercury to hemoglobin, he was led to propose the use of the isomorphous replacement method for the solution of protein structures. The full success of this idea was only achieved six years later in 1960, when it led to the determination of the hemoglobin structure at 5.5 Å resolution. Meanwhile, J Kendrew had also applied isomorphous replacement to the smaller myoglobin molecule and succeeded in building an atomic model of the structure at 2 Å resolution. In 1962, M Perutz and J Kendrew were awarded the Nobel prize in chemistry (**Figure 6**). The following years were rich with theoretical, computational, and instrumental developments, which allowed the solution of more protein structures. In 1972, the Protein Data Bank was established and the number of deposited coordinates of macromolecular structures has increased almost exponentially since then. The improvement of the isomorphous replacement technique was paralleled by new phasing methods, such as anomalous dispersion and molecular replacement, new refinement algorithms, their automation, and the use of computer graphic tools for the interpretation of the electron density maps. There were progresses in instrumentation for data collection and synchrotron radiation became available. Increasingly larger structures of viruses and other biological assemblies could be solved: in 1982, A Klug was awarded the Nobel prize in chemistry for "his structural elucidation of biologically important nucleic acid-protein complexes," in 1988 J Deisenhofer, R Huber, and H Michel were awarded the same prize "for the determination of the three-dimensional structure of a photosynthetic reaction centre," and in 2003 R MacKinnonn, together with P Agre, was awarded the prize "for structural and mechanistic studies of ion channels."

Less Ordered Structures

So far the use of X-ray diffraction from crystals has been considered, that is, from regular 3D arrangements of atoms or molecules. The contrasting situation is that of totally amorphous materials with no long-range order, but rather a continuity of intermediate partially ordered states is present in nature. The "paracrystal" model for the general interpretation of the diffraction patterns of these systems was proposed by R Hosemann (1950).

Already in 1913 W Friedrich had found that paraffin gives a strong X-ray powder reflection, and between 1923 and 1929, several diffraction studies were conducted on long-chain fatty acids.

Several natural substances, such as cellulose, silk, hair, nerve, sinew, and muscle occur as fibers; X-ray fiber diagrams were first reported by R O Herzog and W Jancke (1921). Later the work of W T Astbury and others (1932, 1934) resulted in a classification of

Figure 6 The Nobel prize winners in 1962: from left M Wilkins, M Perutz, F H C Crick, J Steinbeck, J D Watson, and J Kendrew.

Figure 7 Rosalind Franklin and her diffraction pattern of the B form of DNA.

the fibrous proteins. In particular keratinous proteins, such as those which constitute the skin and horny structures of mammals, the feathers of birds, the scales of reptiles, and the flagella of some bacteria, were investigated.

The most famous work based on the interpretation of a fiber diffraction pattern was that on the structure of DNA by F H C Crick, J D Watson, M Wilkins, and R Franklin. The first paper was published by Watson and Crick (1953), who, together with M Wilkins, were awarded the Nobel prize in medicine in 1962 (**Figure 6**). The essential contribution of Rosalind Franklin (**Figure 7**) could not be recognized because of her death in 1958 at the age of 37, and the story of DNA turned out to be a tale of competition and intrigue, told from different points of view by J Watson (2001), A Sayre (2000), B Maddox (2002), and M Wilkins (2003) in fascinating books.

Among other paracrystalline systems, liquid crystals were first analyzed in 1922 by G Friedel, who opened the route to the structural elucidation of the different types of mesomorphic phases and of natural and artificial membranes.

Conclusions

Structure determination by X-ray diffraction is one of the most powerful tools for seeing the arrangement of atoms and molecules in a material and has greatly contributed to the advances of different branches of science.

In physics, it contributed to the understanding of the electronic structure of solids and was the background for the discovery of new materials such as the superconducting ceramic materials by J G Bednorz and K A Müller (Nobel prize in physics in 1987).

In chemistry, it allowed a more profound understanding of the nature of the different types of chemical bonds, as was clearly shown by L C Pauling, who was awarded the Nobel prize in chemistry "for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances." He was the first to show that polypeptide chains in proteins could form α-helices and β-pleated sheets (1951), basic information that was essential in the analysis of the structures of all biological macromolecules. More recently, the very

fruitful concept of supramolecular chemistry was the motivation of the 1987 Nobel prize in chemistry to J M Lehn together with D J Cram and C J Pedersen.

In life sciences, knowledge of the ternary and quaternary structure of proteins and biological assemblies opened the route to the elucidation of the molecular bases of biological processes and is an essential part of molecular biology. The 3D structure of drug molecules and of some of their receptors is the basis for modern drug design strategies.

Mineral crystals were the first to be discovered and studied by man, but the advent of X-ray crystallography opened a completely new scenario in the field of earth sciences. For instance, the knowledge of the structures of the different types of silicates and of rock-forming minerals allows a better understanding of petrogenetic processes.

In materials sciences, structural information is the basis for the tailoring of new materials with specific properties. An example is the stereospecific polymerization and the synthesis of sterically regular polymers for which K Ziegler and G Natta were awarded the Nobel prize in chemistry in 1963.

See also: Crystal Structure; Crystal Structure Determination; Space Groups.

PACS: 01.65.+g; 61.50.−f; 61.66.−f

Further Reading

Bijvoet JM, Burgers WG, and Hägg G (ed.) (1972) *Early Papers on Diffraction of X-rays by Crystals*. International Union of Crystallography. Utrecht: A. Oosthoek's Uitgeversmaatschappij N.V.

Bragg WL (1922) Nobel lecture on *The Diffraction of X-Rays by Crystals*, From *Nobel Lectures*, Physics 1901–1921. Amsterdam: Elsevier Publishing Company (http://www.nobel.se/physics/laureates/1915/wl-bragg-lecture.html).

Buerger MJ (1959) *Vector Space*. New York: Wiley.

Ewald PP (ed.) (1962) *Fifty Years of X-Ray Diffraction*. Internation Union of Crystallography. Utrecht: A. Oosthoek's Uitgeversmaatschappij N.V. (available in PDF format from http://journals.iucr.org/iucr-top/publ/50YearsOfXrayDiffraction/).

Hahn T (ed.) (2002) *International Tables for Crystallography* (2002), 5th edn., International Union of Crystallography. Dordrecht: Kluwer Academic. This is the last edition of the tables, but there has been a long and interesting story of these tables, starting from the first German edition of 1935 (*Internationale Tabellen zur Bestimmung von Kristallstrukturen*, Borntaeger, Berlin) followed by the more complete English edition in 1952 (*International Tables for X-ray Crystallography*, Kynoch Press, Birmingham) and by all subsequent editions. Before 1935, the space group symmetry tables derived by E S Fedorov (in Russian, *Zap. Min. Obshch.*, 1885 and 1891) and A. Schoenflies (*Kristallsysteme und Kristallstruktur*, Leipzig, 1891) were used.

Lipson H and Cochran W (1953). In: Bragg WL (ed.) *The Determination of Crystal Structures*. The Crystalline State vol. III. London: G. Bell & Sons LTD.

Maddox B (2002) *Rosalind Franklin: The Dark Lady of DNA*. New York: Harper Collins Publishers.

Monaco H, Artioli G, Viterbo D, Ferraris G, Gilli G, *et al*. (2002). In: Giacovazzo C (ed.), *Fundamentals of Crystallography*, 2nd edn. Oxford: Oxford University Press.

Sayre A (2000) *Rosalind Franklin and DNA*. New York: WW Norton & Company.

Watson JD (2001) *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. New York: Touchstone Books.

Wilkins M (2003) *The Third Man of the Double Helix: The Autobiography of Maurice Wilkins*. Oxford: Oxford University Press.

Crystal Symmetry

Th Hahn, Institut für Kristallographie, RWTH Aachen, Germany
H Klapper, Mineralogisch-Petrologisches Institut, Universität Bonn, Bonn, Germany

Introduction

Among the various states of condensed matter, the crystalline state is governed by symmetry more than any other state (e.g., liquids, glasses). This is because all atoms of one type in a solid compound display the same structural behavior and adopt, as far as possible, the same atomic environment around them. This applies to all kinds of atoms in the compound and results from the tendency of any system to minimize its free energy. As a first consequence, this leads to three-dimensional periodicity (translational symmetry) of the atoms (ions, molecules) in a crystal structure. As a second consequence, further symmetries, such as rotations, reflections, and inversions (through a point), may occur.

As a rule, symmetry in science is described mathematically in terms of the "group theory." For crystals, the main symmetry groups are point groups, translation groups, and space groups (these are discussed elsewhere in this encyclopedia). Group elements are the symmetry operations discussed in the section "Crystallographic symmetry operations and symmetry elements." Fundamentals of crystallography and of crystallographic symmetry groups can be found in text books of crystallography and in *ITA* (2002) (cf. section "Further reading").

Two levels of crystalline symmetry must be distinguished:

1. "Microscopic" symmetry of the crystal structure: this implies a three-dimensional periodicity which is described by translation groups and their lattices. Furthermore, all permissible combinations of translations with reflections, rotations and inversions, including screw rotations and glide reflections, can occur as elements of space groups.
2. "Macroscopic" symmetry of the single (bulk) crystal: it is derived from the microscopic symmetry (space group) of the crystal structure by the suppression of all translations, both lattice translations and translation components of screw rotations and glide reflections. This leads to the point group of the crystal, which represents its morphological symmetry as manifested by the general face form, which may develop under ideal conditions during crystal growth. Furthermore, the symmetry of physical properties of crystals, in particular of tensor properties, is also described by point groups. Note that in this description, the macroscopic crystal is considered as an anisotropic homogeneous continuum.

Lattices and Their Symmetries

The symmetry operations of a crystal are "isometries" or "rigid motions", that is, mappings which preserve distances and, hence, angles and volumes also. The simplest motion is a translation which is a parallel displacement of an object by a translation vector $\boldsymbol{t}$. The three-dimensional periodicity (translation symmetry) of the (infinitely extended) crystal structure is described by the infinite set $\{\boldsymbol{t}\}$ of all translation vectors $\boldsymbol{t}$ mapping the structure onto itself, that is, by the lattice. It is represented by the equation

$$\{\boldsymbol{t}\} = \{m \cdot \boldsymbol{a} + n \cdot \boldsymbol{b} + p \cdot \boldsymbol{c}\} \quad (-\infty < m, n, p < +\infty) \qquad [1]$$

with $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ three linearly independent basis vectors spanning a three-dimensional parallelepiped called unit cell. Each translation t maps the crystal structure onto itself and is, hence, a symmetry operation of the crystal.

Equation [1] implies that for a given basis vector set $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, one unique lattice $\{\boldsymbol{t}\}$ is defined. On the other hand, one and the same lattice can be generated by infinitely many different triplets $\boldsymbol{a}'$, $\boldsymbol{b}'$, $\boldsymbol{c}'$ of linearly independent lattice vectors, each triplet spanning a "primitive" (smallest) unit cell (this triplet is called "primitive basis" here). All these triplets are equally correct lattice bases. As will be shown below, one (or at most two) particularly symmetry-adapted basis vector sets are conventionally accepted by all scientists.

There exist four different types of lattice symmetries in two dimensions and seven in three dimensions. These so-called "lattice point symmetries" are defined as the symmetries of the infinite surrounding of each lattice point by the other lattice points. These lattice symmetries are also called "holohedries" (holohedral point groups), a term which stems from the old morphological classification, meaning "full face form." These holohedries determine the seven three-dimensional crystal systems and crystallographic coordinate systems: triclinic (anorthic), monoclinic, orthorhombic, tetragonal, trigonal/rhombohedral, hexagonal, cubic (see **Table 1**, see footnote there about trigonal/rhombohedral). Geometrically, these holohedral lattice symmetries are brought out directly in the so-called "Wigner–Seitz cells" or "Dirichlet domains" (see *ITA* (2002)).

Crystallographic Coordinate Systems

"Primitive" Crystallographic Coordinate Systems

Whereas in physics and chemistry normally Cartesian coordinate systems are used, in crystallography quite different systems are applied. Here it is customary that every crystal receives its own specific crystallographic coordinate system, based on the symmetry of the crystal and characterized as follows:

1. Basis vectors are three linearly independent, shortest ("primitive") translation vectors of the crystal lattice (see eqn [1]). This implies six lattice parameters: three axial lengths a, b, c, and three interaxial angles α, β, γ. ("Primitive" means that the unit cell spanned by the three basis vectors contains only one lattice point, in contrast to "centered cells," see the next section.)
2. The origin of the coordinate system is free in principle, but for convenience in computation, it is usually chosen in a center of symmetry (inversion center), if present, otherwise in a point of high site symmetry of the space group (e.g., in an intersection of two or more symmetry axes).

The advantage of such a crystal-specific coordinate system is that all lattice points receive integer coordinates, and this applies to all choices of a primitive coordinate system for a given crystal. Note that the lattice parameters of a crystal vary with temperature and pressure.

Table 1 Crystal systems, lattice point symmetries, conventional crystallographic coordinate systems and Bravais lattice types in two and three dimensions. The lattice point symmetries (column 2) are given in Hermann–Mauguin (left) and in Schoenflies symbols (right); see Section "Symbols for point groups and space groups"

Crystal system	*Lattice point symmetry (holohedry)*		*Conventional coordinate system (restrictions on cell parameters $a, b, c, \alpha, \beta, \gamma$)*	*Bravais lattices*[a]
Two dimensions				
Oblique (monoclinic)	2		None	*mp*
Rectangular (orthorhombic)	2*mm*		$\gamma = 90^\circ$	*op*, *oc*
Square (tetragonal)	4*mm*		$a = b, \gamma = 90^\circ$	*tp*
Hexagonal	6*mm*		$a = b,\ \gamma = 120^\circ$	*hp*
Three dimensions				
Triclinic (anorthic)	$\overline{1}$	C_i	None	a*P*
Monoclinic	2/*m*	C_{2h}	$\alpha = \gamma = 90^\circ$ (unique axis *b*)	*mP*, *mC* *mA*, *mI*
			$\alpha = \beta = 90^\circ$ (unique axis *c*)	*mP*, *mA* *mB*, *mI*
Orthorhombic	2/*m* 2/*m* 2/*m*	D_{2h}	$\alpha = \beta = \gamma = 90^\circ$	*oP*, *oI*, *oF* *oA*, *oB*, *oC*
Tetragonal	4/*m* 2/*m* 2/*m*	D_{4h}	$a = b,\ \alpha = \beta = \gamma = 90^\circ$	*tP*, *tI*
Trigonal[b]				
Rhombohedral lattice	$\overline{3}$ 2/*m*	D_{3d}	$a = b = c,\ \alpha = \beta = \gamma \neq 90^\circ$	*hR*
Hexagonal lattice	6/*m* 2/*m* 2/*m*	D_{6h}	$a = b,\ \alpha = \beta = 90^\circ, \gamma = 120^\circ$	*hP*
Hexagonal	6/*m* 2/*m* 2/*m*	D_{6h}	$a = b,\ \alpha = \beta = 90^\circ, \gamma = 120^\circ$	*hP*
Cubic	4/*m* $\overline{3}$ 2/*m*	O_h	$a = b = c,\ \alpha = \beta = \gamma = 90^\circ$	*cP*, *cI*, *cF*

[a] Bravais lattices (column 4): The lower-case first letter of the symbol identifies the crystal system. Note that the trigonal and the hexagonal system both receive the letter *h*. Second letter of the symbol: *p*, *c* in two dimensions: primitive and centered cell. Three dimensions: *P* = primitive cell; *A*, *B*, *C* = one-face centered cell on (b, c), (*a*,*c*) or (*a*,*b*) face; *I* = body centered cell; *F* = all-face *centered cell*; *R* = rhombohedral centering of the hexagonal cell.

[b] In the trigonal crystal system two Bravais lattices, *hR* and *hP*, with different holohedries, $\overline{3}$ 2/*m* and *6/m 2/m 2/m*, occur. Both are normally described by "hexagonal axes."

Conventional Crystallographic Coordinate Systems

In order to define a unique crystallographic coordinate system for each crystal, special conventions have to be devised. It turns out, however, that the use of a special primitive coordinate system, as described in the preceding section, would indeed always characterize the crystal lattice uniquely, but in many cases would not bring out the symmetry of the crystal. This, however, is a universally accepted objective for the definition of the conventional crystallographic coordinate system in the form of a symmetrical parallelepiped as unit cell.

For this reason, the 14 Bravais types of lattices are considered, with the provision that all Bravais lattices with the same holohedry should have the same conventional coordinate system. A consequence of this procedure is that only for primitive Bravais lattices, all lattice points receive integer coordinates, whereas for centered lattices of the same lattice, symmetry points with fractional coordinates 1/2, 1/3, and 2/3 (and no others) appear. In this way, seven conventional coordinate systems result, which apply to the crystals of all 14 Bravais lattice types. As a typical example, the cubic crystal system is considered: for the primitive cubic lattice *cP*, the conventional coordinate system employs the three shortest lattice vectors with right interaxial angles $\alpha = \beta = \gamma = 90^\circ$. In this case, the conventional coordinate system is at the same time a primitive crystallographic basis. For the body-centered and all-face-centered cubic Bravais lattices *cI* and *cF*, however, only a fraction of the lattice points (1/2 for *cI* and 1/4 for *cF*) have integer coordinates represented by 0,0,0, whereas the remaining centering lattice points have fractional coordinates: 1/2, 1/2, 1/2 for *cI*, and 1/2, 1/2, 0 & 0, 1/2, 1/2 & 1/2, 0, 1/2 for *cF*. If only primitive cells were used, the coordinate system for *cP* would remain the same (cube = 90° rhombohedron), whereas for *cI* and *cF* rhombohedral cells with $\alpha = 109.47^\circ$ and $\alpha = 60^\circ$, respectively would result which would not display the cubic symmetry.

The conventional crystallographic coordinate systems are collected in **Table 1**, which lists the four plane and the seven spatial lattice symmetries with their point-group symbols and the restrictions on the general lattice parameters $a, b, c, \alpha, \beta, \gamma$ imposed by the lattice symmetry. The symbols of the Bravais lattices which belong to a given coordinate system are listed too. These symmetry criteria uniquely determine five of the seven three-dimensional coordinate systems because these five exhibit three or more symmetry directions ("blickrichtungen"): orthorhombic, tetragonal, rhombohedral, hexagonal, cubic. For the two exceptions, triclinic and monoclinic, no or only one symmetry direction exists. In these cases, metrical criteria, such as reduced bases (i.e., choice of the three (triclinic) or two (monoclinic) shortest lattice translations as basis vectors), have to be applied. It should be noted that for historical reasons, in the monoclinic system two different "settings" are in use, which are distinguished by the labeling of the unique symmetry axis: "unique axis *b*" and "unique axis *c*" settings. The *b*-axis setting is the more frequently used one (see Part 2 of *ITA* (2002)).

There is one case in which not one but two different conventional coordinate systems are in use, for trigonal crystals with a rhombohedral *hR* lattice. These crystals are alternatively described by "rhombohedral axes" or, more frequently, by "hexagonal axes," whereby the first case corresponds to a primitive rhombohedral unit cell, the second to the triple R-centered hexagonal unit cell. In contrast, crystals with a hexagonal *hP* lattice are always described by "hexagonal axes."

Detailed descriptions of the conventional crystallographic coordinate systems and reduced bases can be found in sections 1.2, 2.1, 8.3.1, 9.1 and 9.2 of *ITA* (2002).

Crystallographic Symmetry Operations and Symmetry Elements

Symmetry Operations

The following types of symmetry operations exist in crystals and are elements of crystallographic symmetry groups:

1. Translations (already treated in the section "Lattices and their symmetries").
2. Rotations around an axis, reflections across a plane and inversions in a point, as well as their combinations, that is, rotoinversions and rotoreflections. These operations constitute the so-called point-group operations, which leave at least one point of space invariant.
3. Combinations between point-group operations and translations, leading to screw rotations and glide reflections. Both types occur only in space groups, not in point groups. Screw rotations are rotations by an angle $\varphi = 360^\circ/n$, followed by a translation (screw component) parallel to the rotation axis. Glide reflections are reflections across a plane, followed by a translation (glide component) parallel to the plane. The screw and glide components are rational fractions of a lattice translation.

An important fact of crystallographic symmetry is the restriction of the rotation angles φ by the translation lattice to $\varphi = 180^\circ, 120^\circ, 90^\circ, 60^\circ$, defining twofold, threefold, fourfold, and sixfold rotations and screw rotations. This is the consequence of the three-dimensional periodicity of crystals. Apart from a simple mathematical proof (see section "Matrix representation of symmetry operations"), this result is immediately obvious from the fact that a plane can be covered without gaps and overlaps only by parallelograms and their special forms (rectangles, rhombuses, squares), and by triangles and hexagons, but not by pentagons (corresponding to a rotation angle of 72°), heptagons, octagons, etc. Hence, rotation axes other than two-, three-, four-, and sixfold axes are not possible in crystals.

Symmetry Elements

In addition to symmetry operations (which represent "action" on an object), symmetry can also be described in (static) geometrical terms, known as symmetry elements. They form the geometric locus (geometric element), oriented in space, at which a symmetry operation or a cyclic group of operations is performed (plane for a reflection, line for a rotation, point for an inversion), together with a description of these operations. In other words, the symmetry element is the "invariant subspace" of a given operation, that is, the set of fixed points which are invariant under the operation.

Symmetry elements are reflection (mirror) planes, glide planes, rotation axes, screw axes and centers of symmetry. For rotoinversions and rotoreflections, the symmetry element consists of two items: axis and inversion point on the axis, or axis and mirror plane normal to the axis. For screw rotations and glide reflections, the screw and glide components have to be removed first in order to obtain the "fixed points" for screw axes and glide reflections.

Note that the group elements of a symmetry group are the symmetry operations, not the symmetry elements. This discrepancy of nomenclature has historical reasons and is the source of frequent misunderstandings.

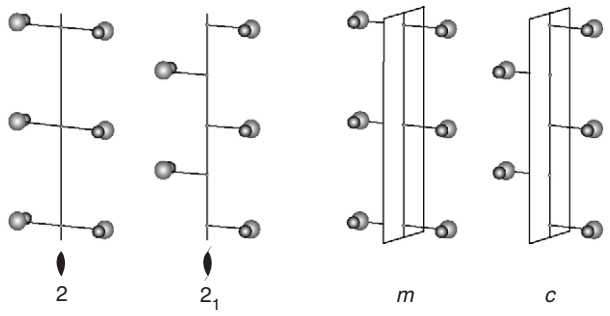


Figure 1 Symmetry elements in crystals: twofold rotation and screw axes 2 and 2_1 (left pair), and reflection and glide planes *m* and *c* (right pair). The "objects" are polar two-atomic molecules, in order to bring out the orientational differences between rotations and reflections. Note that each diagram extends over two translation periods along *c*. The graphical symbols underneath the diagrams are explained in **Table 2**.

The crystallographic symmetry elements (reflection and glide plane, rotation and screw axes, and rotoinversion axes) are illustrated in **Figures 1–3**. Printed and graphical symbols of all symmetry elements are collected in **Table 2**. Further details are given in Parts 1 and 8.1 of *ITA* (2002) and the references quoted therein. Complete "frameworks of symmetry elements" for the 17 plane groups, the 230 space groups, and the 32 crystallographic point groups are presented in Parts 6, 7 and 10.1 (and explained in section 2.2.6) of *ITA* (2002).

Matrix Representation of Symmetry Operations

General Relations

In the mathematical treatment of symmetry operations, normally the coordinate system, that is, the set of basis vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and the origin, is considered to be at rest, whereas the object is mapped onto itself (special affine mapping). (Equivalently, it is equally correct to transform the coordinate system, whereby the object is considered to be at rest.) A general symmetry operation transforms a point X with coordinates x, y, z into a symmetrically equivalent point $\tilde{X}$, with coordinates $\tilde{x}$, $\tilde{y}$, $\tilde{z}$ according to

$$\begin{aligned}\tilde{x} &= W_{11}x + W_{12}y + W_{13}z + w_1\\ \tilde{y} &= W_{21}x + W_{22}y + W_{23}z + w_2\\ \tilde{z} &= W_{31}x + W_{32}y + W_{33}z + w_3\end{aligned} \qquad [2]$$

or, in matrix notation:

$$\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ \tilde{z}\end{pmatrix} = \begin{pmatrix}W_{11} & W_{12} & W_{13}\\ W_{21} & W_{22} & W_{23}\\ W_{31} & W_{32} & W_{33}\end{pmatrix}\cdot\begin{pmatrix}x\\ y\\ z\end{pmatrix} + \begin{pmatrix}w_1\\ w_2\\ w_3\end{pmatrix}$$

$$\tilde{\boldsymbol{r}} = \boldsymbol{W}\cdot\boldsymbol{r} + \boldsymbol{w} = (\boldsymbol{W},\boldsymbol{w})\cdot\boldsymbol{r} \qquad [3]$$

The (3×3) matrix $\boldsymbol{W}$ is the rotation part, and the (3×1) column matrix $\boldsymbol{w}$, the translation part of the symmetry operation. The matrices $\boldsymbol{W}$ represent all point-group operations, rotations, reflections, inversions, and rotoinversions. The two parts $\boldsymbol{W}$ and $\boldsymbol{w}$ can be assembled into an (augmented) (4×4) matrix $\mathbb{W}$ according to

$$\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ \tilde{z}\\ 1\end{pmatrix} = \left(\begin{array}{ccc:c} W_{11} & W_{12} & W_{13} & w_1\\ W_{21} & W_{22} & W_{23} & w_2\\ W_{31} & W_{32} & W_{33} & w_3\\ \hdashline 0 & 0 & 0 & 1\end{array}\right)\cdot\begin{pmatrix}x\\ y\\ z\\ 1\end{pmatrix}$$

$$= \left(\begin{array}{ccc:c} & \boldsymbol{W} & & \boldsymbol{w}\\ \hdashline 0 & 0 & 0 & 1\end{array}\right)\cdot\begin{pmatrix}x\\ y\\ z\\ 1\end{pmatrix} = \mathbb{W}\cdot\mathbb{r} \qquad [4]$$

Since every symmetry transformation is a "rigid-body motion," the determinant of the matrices $\boldsymbol{W}$ and $\mathbb{W}$ is

$$\det \boldsymbol{W} = \det \mathbb{W} = \pm 1 \qquad [5]$$

The sequence of two symmetry operations (successive application) is given by the product of their (4×4) matrices $\mathbb{W}_1$ and $\mathbb{W}_2$:

$$\mathbb{W}_3 = \mathbb{W}_2\cdot\mathbb{W}_1 \qquad [6]$$

whereby, $\mathbb{W}_3$ is again a symmetry operation.

The inverse operation $\mathbb{W}^{-1}$ is defined by $\mathbb{W}\cdot\mathbb{W}^{-1} = \mathbb{I}$, whereby $\mathbb{I}$ is the (4×4) unit matrix representing the identity operation:

$$\mathbb{W}^{-1} = \left(\begin{array}{ccc:c} & \boldsymbol{W}^{-1} & & -\boldsymbol{W}^{-1}\cdot\boldsymbol{w}\\ \hdashline 0 & 0 & 0 & 1\end{array}\right) \qquad [7]$$

Finally, the augmented matrix for a translation $\boldsymbol{t}$ is

$$\mathbb{T} = \left(\begin{array}{ccc:c} & \boldsymbol{I} & & \boldsymbol{t}\\ \hdashline 0 & 0 & 0 & 1\end{array}\right), \qquad \boldsymbol{t} = \begin{pmatrix}t_1\\ t_2\\ t_3\end{pmatrix} \qquad [8]$$

and $\boldsymbol{I}$ the (3×3) unit matrix.

Note on the matrices $\boldsymbol{W}$ and $\theta\cdot k$ of eqns [3] and [4]: The rotation parts $\boldsymbol{W}$ of these matrices leave an atom located in the origin of the coordinate system invariant, that is, the symmetry element of $\boldsymbol{W}$ intersects the

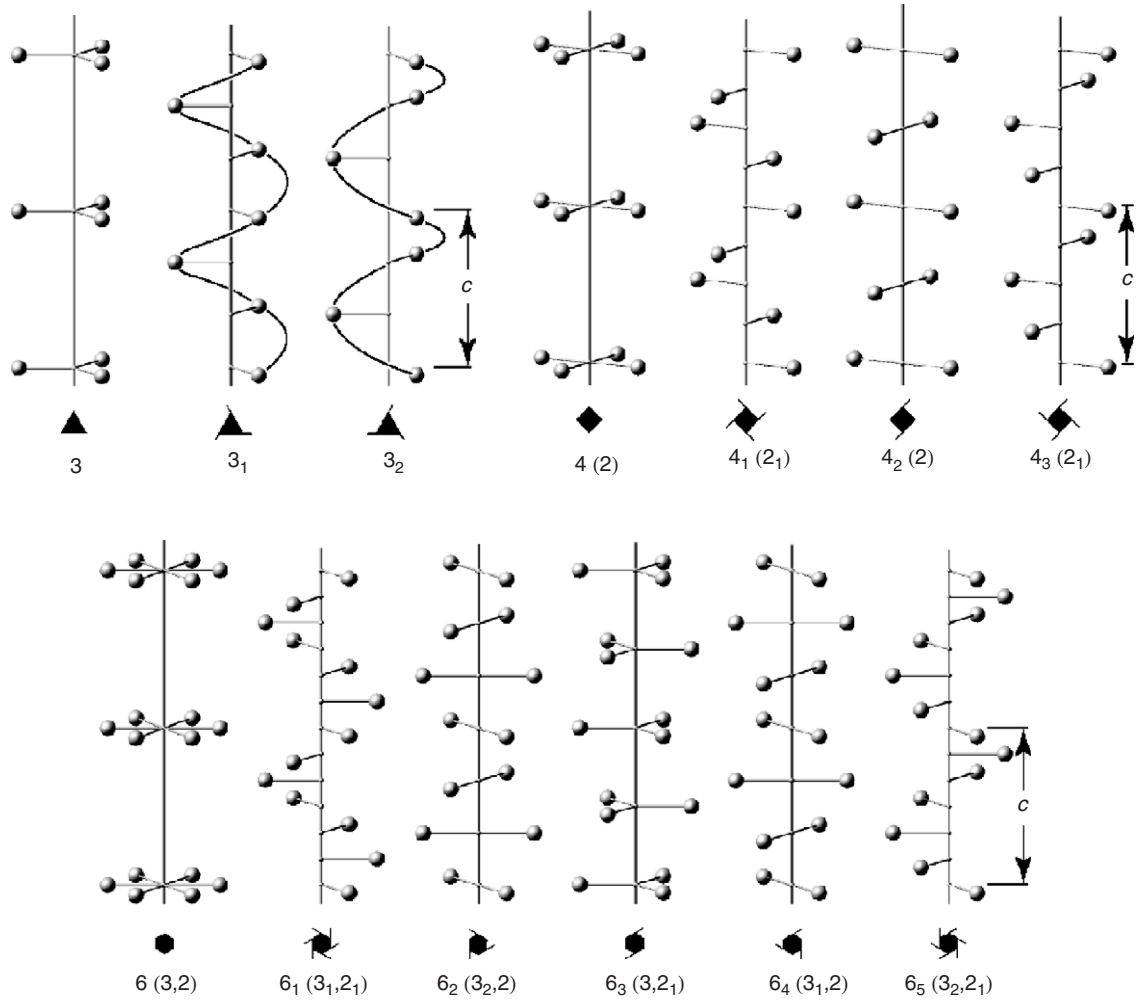


Figure 2 Symmetry elements in crystals: threefold rotation and screw axes (top row), fourfold rotation and screw axes (top row), and sixfold rotation and screw axes (bottom row). Each diagram extends over two translation periods along *c*. For the screw axes 3_1 and 3_2, the course of the screws (helices) is indicated in order to show the enantiomorphism (mirror images) of the two screw axes. The subelements (subgroups) of each symmetry element (cyclic group) are given in parentheses. The graphical symbols are explained in **Table 2**.

origin. The translation parts $\boldsymbol{w}$ consist of two components, first the (intrinsic) "screw or glide component" of a screw rotation or glide reflection (if present), second the "location parameter," which depends on the location of the symmetry element with respect to the origin. Thus, for a pure rotation or reflection through the origin, $\boldsymbol{w} = (w_1, w_2, w_3) = (0, 0, 0)$.

Geometric Interpretation of (Crystallographic) Symmetry Matrices

A (3×3) matrix $\boldsymbol{W}$ of a rigid motion has two essential invariants, that is, two quantities which do not change under a transformation of the coordinate system and, hence, can serve to identify the "nature" of the motion. These quantities are the determinant, $\det \boldsymbol{W}$, and the trace, $\mathrm{tr}\, \boldsymbol{W} = W_{11} + W_{22} + W_{33}$. Their properties are as follows:

1. A rotation by an angle φ and its representation in a Cartesian coordinate system x_1, x_2, x_3 are assumed, whereby the rotation axis intersects the origin and is parallel to x_3. The rotation matrix is

$$\boldsymbol{W} = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The trace is $\mathrm{tr}\, \boldsymbol{W} = 1 + 2\cos\varphi$, valid for any coordinate system. Since this rotation maps a lattice point with integer coordinates onto another lattice point, all elements of the matrix, if referred to a crystallographic coordinate system, must be integers,

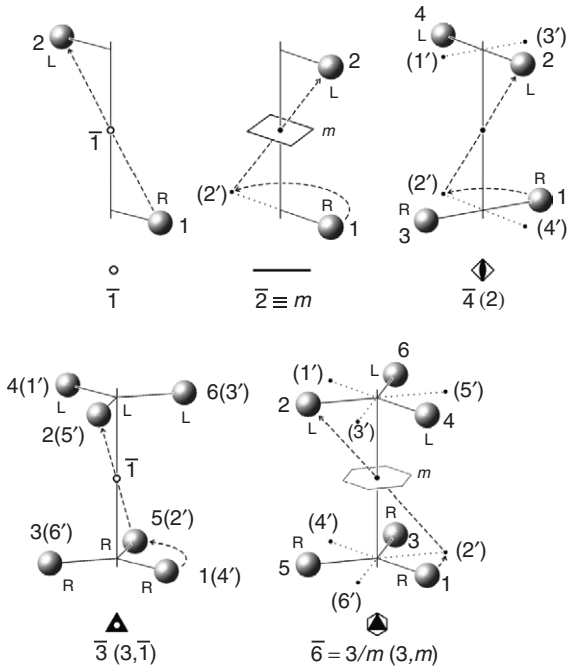


Figure 3 Symmetry elements in crystals: the five crystallographic rotoinversion axes $\bar{1}$, $\bar{2}$, $\bar{4}$, $\bar{3}$, $\bar{6}$ (orders 2, 2, 4, 6, 6, resp.), and their action on a spherical object. Starting from object 1, objects 2, 3, … are successively generated. The intermediate positions of the object (between the rotation and the inversion part of the operation) are labeled by numbers in parentheses. For all groups the first operation $1 \rightarrow (2') \rightarrow 2$ is indicated by dashed arrows. The labels R and L denote the handedness (right or left), if the object is chiral (starting object 1: R). All odd powers of rotoinversions are hand-changing (R↔L), even powers hand-preserving (R→R, L→L). Note that for axes $\bar{n}$ with n odd (i.e., for $\bar{1}$ and $\bar{3}$), the inversion point is a center of symmetry $\bar{1}$ (i.e., a subelement). For axes $\bar{n}$ with $n = 4k + 2$ (i.e., $\bar{2}$ and $\bar{6}$), a mirror plane m normal to the axis and intersecting the inversion point is a subelement, whereas for $n = 4k$ (i.e., $\bar{4}$) neither $\bar{1}$ nor m occur as subelements. The subelements of the rotoinversion axes are noted in parentheses.

hence tr $W = 1 + 2\cos\varphi = m$, and $\cos\varphi = 1/2\,(m-1) = 1/2p$ (m, p integers). Since $\cos\varphi$ is restricted to the range -1 to $+1$, only the following values of $\cos\varphi$ are compatible with lattice symmetry: $\cos\varphi = +1, +1/2, 0, -1/2, -1$, corresponding to rotation angles of $0^\circ \equiv 360^\circ$, 60°, 90°, 120°, and 180°. This is the famous "crystallographic restriction," which filters out the five permissible φ-angles (and their multiples) from the infinity of "n-fold rotations" with $\varphi = 360^\circ/n$.

2. Since symmetry operations are rigid motions which preserve distances, angles, and volumes, the determinant of their matrices must be $+1$ or -1:

   det $W = +1$: rotation; handedness (left or right) of object retained;

   det $W = -1$: reflection, inversion, rotoinversion : handedness of object reversed;

3. By combining the possible values of det W and tr W, the "nature" of a symmetry operation by means of its (3×3) symmetry matrix W is identified, as shown in **Table 3**.
4. By more detailed analysis of the (3×3) "point-group" matrices W and the augmented (4×4) "space-group" matrices $\mathbb{W}$, further characteristics of symmetry operations can be derived, that is, the orientation of the axis or of the plane in the given coordinate system, as well as the location of the symmetry element (plane, axis, point) in the unit cell relative to the chosen origin (see *ITA* (2002), Part 11).
5. In order to illustrate the previous somewhat abstract considerations, some examples of (4×4) matrices for (glide) reflections and (screw) rotations, together with their geometric elements, are given in **Table 4**. A list of all (3×3) matrices occurring in crystallographic point and space groups is available in Part 11 of *ITA* (2002).
6. It should be noted that in quasicrystals (i.e., forms of solid matter not based on exact three-dimensional periodicity but rather on "quasiperiodicity"), symmetries may occur which go beyond the "crystallographic restriction" discussed above under point 1. For instance, quasicrystals with icosahedral symmetry (containing fivefold rotations), as well as with tenfold and twelvefold rotations have been found.

Symbols for Point Groups and Space Groups

So far, "individual" symmetry operations and symmetry elements have been treated with their printed and graphical symbols, see **Table 2**. In order to combine these into symbols for crystallographic point and space groups, rules for the combination have to be specified. There are two types of group symbols in use: the oldest are the Schoenflies symbols contained in Schoenflies' classic work of 1891. The second are the Hermann-Mauguin symbols, derived around 1928 in several papers by Hermann and Mauguin. The latter are now predominantly used in crystallography and crystal-structure analysis, and are customarily called "international symbols." Schoenflies symbols are often preferred in spectroscopy and in studies of molecules.

Schoenflies Symbols

These symbols are based on group theory and use capital letters to indicate the major types of symmetry groups.

C: Cyclic groups C_1, C_2, C_3, C_4, C_6 of order $n = 1, 2, 3, 4$ and 6;

Table 2 Printed and graphical symbols for symmetry elements in two and three dimensions. For glide planes only a few selected examples are given

Printed symbol	*Symmetry element and its orientation, glide and screw vectors* $\boldsymbol{t}$	*Graphical symbol normal to plane of projection*
m	Reflection (mirror) plane (three dimensions) Reflection (mirror) line (two-dimensions)	
a, *b*, or *c*	"Axial" glide plane	
a	$\perp[0\,1\,0]$ or $\perp[0\,0\,1]$ $\quad \boldsymbol{t} = 1/2\boldsymbol{a}$	
b	$\perp[1\,0\,0]$ or $\perp[0\,0\,1]$ $\quad \boldsymbol{t} = 1/2\boldsymbol{b}$	
c	$\perp[1\,0\,0]$ or $\perp[0\,1\,0]$ $\quad \boldsymbol{t} = 1/2\boldsymbol{c}$	
g	Glide line (two dimensions) $\perp[0\,1]$; $\perp[1\,0]$; $\quad \boldsymbol{t} = 1/2\boldsymbol{a}$; $1/2\boldsymbol{b}$ (resp.)	
n	"Diagonal" glide plane $\perp[0\,0\,1]$; $\perp[1\,0\,0]$; $\perp[0\,1\,0]$; $\boldsymbol{t} = 1/2(\boldsymbol{a}+\boldsymbol{b})$; $1/2(\boldsymbol{b}+\boldsymbol{c})$; $1/2(\boldsymbol{a}+\boldsymbol{c})$ (resp.)	
d	"Diamond" glide plane (pair of planes)	
	Orthorhombic: $\perp[0\,0\,1]$; $\perp[1\,0\,0]$; $\perp[0\,1\,0]$; $\boldsymbol{t} = 1/4(\boldsymbol{a}\pm\boldsymbol{b})$; $1/4(\boldsymbol{b}\pm\boldsymbol{c})$; $1/4(\boldsymbol{c}\pm\boldsymbol{a})$ (resp.) Cubic (examples): $\perp[1\,1\,0]$; $\perp[0\,1\,1]$; $\perp[1\,0\,1]$; $\boldsymbol{t} = 1/4(-\boldsymbol{a}+\boldsymbol{b}\pm\boldsymbol{c})$; $1/4(\pm\boldsymbol{a}-\boldsymbol{b}+\boldsymbol{c})$; $1/4(\boldsymbol{a}\pm\boldsymbol{b}-\boldsymbol{c})$ (resp.)	
1	None (identity)	
2, 3, 4, 6	*n*-fold rotation axis, *n* (three dimensions) *n*-fold rotation point, *n* (two dimensions)	
$\overline{1}$	Center of symmetry, inversion center	
$\overline{2} = m, \overline{3}, \overline{4}, \overline{6}$	Rotoinversion axis, $\bar{n}$, with inversion point on the axis[a]	
2_1		
3_1, 3_2	*n*-fold screw axis, n_p	
4_1, 4_2, 4_3	The screw component is the fraction p/n $\boldsymbol{t}$ of the shortest lattice translation $\boldsymbol{t}$ parallel to the axis	
6_1, 6_2, 6_3, 6_4, 6_5		

[a] The inversion point is a center of symmetry (inversion center) if $\boldsymbol{n}$ is odd.

Notes: 1. The graphical symbols (column 3) refer to planes and axes oriented normal to the plane of projection.

2. Diamond glide planes *d* occur only in orthorhombic *F* space groups, in tetragonal *I* space groups, and in cubic *I* and *F* space groups. They always occur in pairs with alternating vectors, for instance, 1/4 ($\boldsymbol{a}+\boldsymbol{b}$) and 1/4 ($\boldsymbol{a}-\boldsymbol{b}$). The second power of a glide reflection *d* (i.e., the result of two successive applications) is a lattice centering vector. The name "diamond glide" is due to its pronounced occurrence in the diamond structure.

3. Rotoreflections are not listed here, because they correspond to rotoinversions according to the following scheme: two-, three-, four-, and sixfold rotoreflections are identical with the rotoinversions $\overline{1}$ (inversion), $\overline{6}$, $\overline{4}$ and $\overline{3}$, respectively. Note that the even two- and sixfold rotoreflections (corresponding to the odd rotoinversions $\overline{1}$ and $\overline{3}$) contain a symmetry (inversion) center.

4. The following pairs of screw axes are "enantiomorphic," i.e. each pair consists of a right- and a left-handed screw axis: (3_1, 3_2), (4_1, 4_3), (6_1, 6_5) and (6_2, 6_4). The screw axes 2_1, 4_2, 6_3 are "neutral." Note that all rotation and screw axes may accommodate enantiomorphic structures, whenever the building units (atomic groups, molecules) of the structure are "chiral," i.e., either all right- or all left-handed.

5. A right-handed screw rotation is customarily defined as a counter-clockwise rotation with the screw vector pointing toward the observer.

Table 3 Crystallographic symmetry operations as derived from their traces tr $\boldsymbol{W}$ and determinants det $\boldsymbol{W} = +1$ and -1

	det $\boldsymbol{W} = +1$					det $\boldsymbol{W} = -1$				
tr $\boldsymbol{W}$	$+1$	$+1/2$	0	$-1/2$	-1	$+1$	$+1/2$	0	$-1/2$	-1
H.-Mauguin symbol of operation	1	6	4	3	2	$\overline{1}$	$\overline{6} \equiv 3/m$	$\overline{4}$	$\overline{3}$	$\overline{2} \equiv m$
Schoenflies symbol of operation	C_1	C_6	C_4	C_3	C_2	C_i	C_{3h}	S_4	C_{3i}	C_s
Order n of operation[a]	1	6	4	3	2	2	6	4	6	2

[a] The "order" n of the operation $\boldsymbol{W}$ is the order of the cyclic group generated by $\boldsymbol{W}$, i.e., the power n of $\boldsymbol{W}$ for which the identity 1 is reached: $\boldsymbol{W}^n = 1$.

Table 4 Examples of augmented (4×4) matrices of symmetry operations (referred to the appropriate conventional coordinate system)[a]

Inversion with center in x_0, y_0, z_0

$$\left(\begin{array}{cccc} -1 & 0 & 0 & 2x_0 \\ 0 & -1 & 0 & 2y_0 \\ 0 & 0 & -1 & 2z_0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Twofold rotation along z in x_0, y_0, z

$$\left(\begin{array}{cccc} -1 & 0 & 0 & 2x_0 \\ 0 & -1 & 0 & 2y_0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Twofold screw rotation along z in x_0, y_0, z

$$\left(\begin{array}{cccc} -1 & 0 & 0 & 2x_0 \\ 0 & -1 & 0 & 2y_0 \\ 0 & 0 & 1 & 1/2 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Glide reflection n with plane $\perp x$ in x_0, y, z

$$\left(\begin{array}{cccc} -1 & 0 & 0 & 2x_0 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Fourfold (screw) rotation 4_p with $p = 0, 1, 2, 3$ along z in 0, 0, z (§)

$$\left(\begin{array}{cccc} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & p/4 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Fourfold rotoinversion $\bar{4}$ along z in 0, 0, z

$$\left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Threefold (screw) rotation 3_p with $p = 0, 1, 2$ along z in 0, 0, z (§)

$$\left(\begin{array}{cccc} 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & p/3 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Sixfold (screw) rotation 6_p with $p = 0, 1, 2, 3, 4, 5$ along z in 0, 0, z (§)

$$\left(\begin{array}{cccc} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & p/6 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Threefold rotation 3 around body diagonal x, x, x (cubic system) (#)

$$\left(\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

[a] The matrices in this table represent the "first powers" of the operations. Higher powers are obtained by repeated multiplication of the matrix. For point-group operations of order 2 ("binary" operations: inversions, twofold rotations, reflections), the second power is the identity, i.e., binary operations are their own inverses.
(§) Note that $p = 0$ stands for a normal (nonscrew) rotation.
(#) Cyclic permutation of the coordinates. This matrix also represents the one threefold rotation around x, x, x in the rhombohedral coordinate system. There are four equivalent cubic body diagonals x, x, x; $-x, x, x$; $x, -x, x$ and $x, x, -x$ with different cyclic permutations of $+1$ and -1.

C_s = reflection group (s = *Spiegel*symmetry) of order 2;
C_i = inversion group of order 2;

D: Dihedral groups D_2, D_3, D_4, D_6, that is, a cyclic group C with additional twofold rotation axes normal to the main symmetry axis. The order m of these groups is twice that of the corresponding C_n group: $m = 2n$.

T: Tetrahedral group (order $n = 12$);

O: Octahedral group (order $n = 24$).

Subscripts h, v, and d indicate the presence of mirror planes perpendicular (h = horizontal) and parallel (v = vertical, d = diagonal) to the main symmetry axis, which is considered to be vertical, for example, $C_{2h}, D_{4h}, T_h, O_h, C_{2v}, C_{6v}, D_{3d}$. The Schoenflies symbols are independent of any coordinate system.

These symbols apply to three-dimensional crystallographic point groups. Space groups are indicated by superscripts added to the corresponding point-group symbols, for example, $C_{2h}^5, D_2^4, D_{2h}^{16}, T_h^7, O_h^9$. This implies that the Schoenflies symbols carry no information about the nature of the space group. No Schoenflies symbols exist for two-dimensional groups.

Hermann–Mauguin (International) Symbols

These symbols have, as their basis, the seven different point symmetries of the Bravais lattices. Every lattice, if viewed along a general direction, exhibits only inversion symmetry $\overline{1}$. There are, however, between zero (triclinic) and 13 (cubic) specific directions, along which symmetries above $\overline{1}$ are seen. These so-called "lattice symmetry directions" (blickrichtungen) and the symmetries seen along them are characteristic for the seven types of lattice symmetries (holohedries), see **Table 1**. They can be grouped

Table 5 Representative lattice symmetry directions for two and three dimensions. The number in parentheses refers to the number of equivalent symmetry directions of the given set, e.g., cubic [1 0 0] (3) refers to the three equivalent directions [1 0 0], [0 1 0], [0 0 1]

Lattice	*Symmetry direction (position in Hermann–Mauguin symbol)*		
	Primary	*Secondary*	*Tertiary*
Two dimensions			
Oblique (monoclinic)	Rotation	------	------
Rectangular (orthorhombic)	point	[1 0] (1)	[0 1] (1)
Square (tetragonal)	In	[1 0] (2)	$[1\bar{1}]$ (2)
Hexagonal	plane (1)	[1 0] (3)	$[1\bar{1}]$ (3)
Three dimensions			
Triclinic (anorthic)	None		
Monoclinic	[0 1 0] (1) (unique axis *b*) [0 0 1] (1) (unique axis *c*)	------	------
Orthorhombic	[1 0 0] (1)	[0 1 0] (1)	[0 0 1] (1)
Tetragonal	[0 0 1] (1)	[1 0 0] (2)	$[1\,\bar{1}\,0]$ (2)
Hexagonal	[0 0 1] (1)	[1 0 0] (3)	$[1\,\bar{1}\,0]$ (3)
Rhombohedral (hexagonal axes)	[0 0 1] (1)	[1 0 0] (3)	------
Rhombohedral (rhombohedral axes)	[1 1 1] (1)	$[1\,\bar{1}\,0]$ (3)	------
Cubic	[1 0 0] (3)	[1 1 1] (4)	$[1\,\bar{1}\,0]$ (6)

into none (triclinic), one (monoclinic), two (rhombohedral), and three (all other lattices) classes of equivalent symmetry directions (e.g., the fourfold cubic rotation axes [1 0 0], [0 1 0], [0 0 1] form one class). These classes are arranged in a well-defined sequence (called "positions"), which is again characteristic for each crystal system. They form, thus, the basis of the Hermann–Mauguin (international) point-group and space-group symbols.

The representative lattice symmetry directions for these classes are presented in **Table 5**, both for two- and three-dimensional lattices. The three columns define the "position" of an entry in a Hermann–Mauguin group symbol (primary, secondary, and tertiary position). The entries in an actual symbol are those symmetry axes and symmetry planes, which are "observed" for the corresponding blickrichtung. A slash indicates the occurrence of a mirror (glide) plane perpendicular to an axis (e.g., $4/m$, $2_1/c$). In this way, the international point-group symbols are obtained. This process is illustrated, first for "point groups," by the following examples:

1, $\bar{1}$ triclinic; no symmetry direction, since $\bar{1}$ is a point (Schoenflies symbols C_1, C_i);

2, m, $2/m$ monoclinic; only one symmetry direction along which is "seen" a twofold axis, the normal to a mirror plane or the combination of both (Schoenflies symbols C_2, C_s, C_{2h});

$mm2$ orthorhombic; three perpendicular symmetry directions, two of which carry normals to mirror planes, the third a twofold axis (Schoenflies symbol C_{2v});

$4/m$ tetragonal, short for $4/m11$, since only the primary blickrichtung carries symmetry (Schoenflies symbol C_{4h});

$6/m2/m2/m$ (short: $6/mmm$) hexagonal holohedry; three classes of symmetry directions, the primary single direction is populated by a sixfold rotation axis and a perpendicular mirror plane: $6/m$. The three equivalent secondary and three equivalent tertiary directions are populated by twofold axes with perpendicular mirror planes, leading to a total of seven "occupied" blickrichtungen (Schoenflies symbol D_{6h});

$4/m\bar{3}2/m$ (short: $m\bar{3}m$) cubic holohedry; point group of highest crystallographic symmetry; along the three cube edges of the cubic unit cell, symmetry axes/planes $4/m$ occur, the four body-diagonal directions of the cube are populated by $\bar{3}$ inversion axes, and the six face diagonals carry $2/m$ symmetries. In total, there are 13 symmetry directions with 48 symmetry operations ($=$group order) (Schoenflies symbol O_h).

The space-group symbols are modifications of the point-group symbols in two ways: first, the letter with the centering type of the appropriate Bravais lattice, *P*, *A*, *B*, *C*, *I*, *F*, *R* is added in front of the symbol. Second, as appropriate, the symbols for rotation axes are substituted by those of screw axes,

and the symbols of reflection planes by those of glide planes. Lattice symmetry directions, which do not carry any symmetry, are represented by the symbol "1", as in $P31m$ and $P3m1$, or are left empty as in $P6_3/m$, short for $P6_3/m11$.

This procedure leads to the following full (and short) symbols, given as examples:

Monoclinic: $P2_1/c$, $C2/c$ (two of the frequently encountered space groups);

Orthorhombic: $P2_12_12_1$, $P2_1/n\ 2_1/m\ 2_1/a$ ($Pnma$), $C2/m\ 2/c\ 2_1/m$ ($Cmcm$), $I2_1/b\ 2_1/c\ 2_1/a$ ($Ibca$);

Trigonal: $P31m$, $P3m1$, $R3m$;

Hexagonal: $P6_3/m$, $P6_522$, $P\overline{6}2m$, $P6_3/m\,2/c\ 2/m$ ($P6_3/mcm$);

Cubic: $I2_13$, $I2_1/a\overline{3}$ ($Ia\overline{3}$), $P4_332$, $F\overline{4}3c$, $F4/m\overline{3}\ 2/m$ ($Fm\ \overline{3}m$), $I4_1/a\overline{3}\ 2/d$ ($Ia\overline{3}d$).

It is not possible to treat here the details of the derivation and the nomenclature of space groups. This topic requires a certain amount of practical acquaintance. The *International Tables,* Vol. A [*ITA* (2002)], contain complete listings and explanations of the 17 plane (two-dimensional) groups, the 230 space groups and the 32 crystallographic point groups. The so-called "subperiodic groups" (number of periodic dimensions lower than the total number of dimensions of the group), i.e., frieze, layer and rod groups, are presented in Vol. E (2002) of the *International Tables*. For textbooks, please see "Further reading" below.

Acknowledgments

The authors thank R A Becker, H Schmidt, and J Simons (Aachen) for the preparation of the diagrams.

See also: Crystal Tensors: Applications; Crystal Tensors: Introduction; Periodicity and Lattices; Point Groups; Space Groups.

PACS: 61.50.Ah; 33.15.Bh; 02.10.Yn; 02.20.Hj

Further Reading

Bloss ED (1971) *Crystallography and Crystal Chemistry*. New York: Holt, Rinehart and Winston.

Buerger MJ (1965) *Elementary Crystallography*. New York: Wiley.

Buerger MJ (1971) *Introduction to Crystal Geometry*. New York: McGraw-Hill.

Burns G and Glaser AM (1990) *Space Groups for Solid State Scientists*. New York: Academic Press.

Giacovazzo C (ed.) (1992) *Fundamentals of Crystallography*. Oxford: Oxford University Press.

Hahn Th and Wondratschek H (1994) *Symmetry of Crystals. Introduction to International Tables for Crystallography*, Vol. A. Sofia: Heron Press.

International Tables for Crystallography (2002) Vol. A, *Space-group Symmetry*, 5. Edition, Hahn Th (ed.). Dordrecht: Kluwer Academic Publishers. [1. Edition 1983, 5. Edition 2002, abbreviated as *ITA* (2002)].

International Tables for Crystallography (2002) Vol. E, *Subperiodic Groups*, 1. Edition, Kopsky V and Litvin DB (eds.). Dordrecht: Kluwer Academic Publishers.

Ledermann W (1976) *Introduction to Group Theory*. London: Longman.

McKie D and McKie Ch (1986) *Essentials of Crystallography*. Oxford: Blackwell.

Nye JF (1957) *Physical Properties of Crystals. Their Representation by Tensors and Matrices*. Oxford: Clarendon Press. [1. Edition 1957, Revised Edition 1985].

O'Keefe M and Hyde BG (1996) *Crystal Structures I. Patterns and Symmetry*. Washington, DC: Mineralogical Society of America.

Schwarzenbach D (1996) *Crystallography*. Chichester: Wiley. [Original French edition 1996, German edition 2001].

Shubnikov AV and Koptsik VA (1974) *Symmetry in Science and Art*. New York: Plenum Press.

Vainstein BK (1994) *Modern Crystallography, Vol. 1, Fundamentals of Crystals. Symmetry and Methods of Structural Crystallography*, 2. Edition. Berlin: Springer.

Wondratschek H and Neubüser J (1967) *Acta Crystallographica* 23: 349–352.

Nomenclature

$\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$	Basis vectors of a lattice
a, b, c, α, β, γ	Lattice parameters
t; $\boldsymbol{t}$	Translation; translation vector
P, A, B, C, I, F, R	Centering types of lattices
X; $\tilde{X}$	Point; transformed point
x, y, z; $\tilde{x}$, $\tilde{y}$, $\tilde{z}$	Coordinates of points X and $\tilde{X}$
$\boldsymbol{r}$; $\tilde{\boldsymbol{r}}$	Location vector; transformed location vector
W	(3×3) matrix, rotation part of a symmetry operation
$\boldsymbol{w}$	(3×1) column, translation part of a symmetry operation
W_{ij}; w_k	Elements of W and $\boldsymbol{w}$
$\mathbb{W}$, $\mathbb{T}$, $\mathbb{r}$	(4×4) augmented matrix; (4×4) augmented matrix of pure translation; (4×1) augmented vector
(hkl)	Miller indices of a crystal face (lattice plane)
$[uvw]$	Direction indices of a crystal edge (lattice direction)

Schoenflies Symbols

$C_n = C_1$, C_2, C_3, C_4, C_6	Cyclic group of order $n = 1, 2, 3, 4, 6$
C_s; C_i	Reflection group; inversion group (order 2)
S_4	Rotoinversion (rotoreflection) group of order 4
$D_n = D_2, D_3$, D_4, D_6	Dihedral group of order $m = 2n = 4, 6, 8, 12$

T	Tetrahedral group (order 12)
O	Octahedral group (order 24)

Hermann–Mauguin (International) Symbols

$n = 1, 2, 3, 4, 6$	n-fold rotation axes (order n)
$\overline{n} = \overline{1}, \overline{2} \equiv m,$ $\overline{4}, \overline{3}, \overline{6} \equiv 3/m$	n-fold rotoinversion axes (orders 2, 2, 4, 6, 6, resp.)
$n_p = 2_1, 3_1, 3_2, 4_1, 4_2, 4_3, 6_1, 6_2, 6_3, 6_4, 6_5$	n-fold screw axes (order ∞)
m ($\equiv \overline{2}$)	reflection (mirror) plane (order 2)
a, b, c, n, d	glide reflection planes (order ∞)
g	glide line (two dimensions, order ∞)

Crystal Tensors: Applications

D R Lovett, University of Essex, Colchester, UK

Introduction

The application of crystal tensors to crystals with different structures and hence different symmetries is discussed in this article. For most applications, crystals need careful selection to process the appropriate symmetry for the particular application. Many of the terms and equations used have been introduced in this encyclopedia elsewhere to which the reader is also referred.

Electrical and Thermal Properties

Much of the early applications of tensors to crystal properties concerned the electrical and thermal resistivities and conductivities of materials and more complicated effects, such as the Hall effect. The consequence of applying voltages or thermal gradients across anisotropic crystals in arbitrary directions is demonstrated in this encyclopedia elsewhere. The strength of the theory is in suggesting the principal directions in which measurements should be made. Resistivity and conductivity are symmetrical second-rank tensors relating two vector quantities, and for the majority of crystal systems the maximum number of independent coefficients is two, although this rises to three in the orthorhombic system, four in the monoclinic system, and six in the triclinic system. Early measured examples of room-temperature coefficients were for metals, such as cadmium (hexagonal $6/mmm$) with ρ_1 and ρ_2 having the value 6.80 and ρ_3 having the value 8.30 in units of $10^{-8}\Omega$ m. Another example is tin (tetragonal structure, $4/mmm$) with ρ_1, ρ_2 equal to 9.9 and ρ_3 equal to 14.3. These values show that any measurements off the principal directions would require interpretation by a tensor theory. Similarly, notable examples for thermal conductivites are graphite (hexagonal, $6/mmm$) with k_1 of 6.5 and k_3 of 11.3 W m^{-1} °C^{-1}, and quartz (trigonal, 32) with k_1, k_2 of 9.24 and k_3 of 6.65.

In the Hall effect, a magnetic field is applied perpendicular to the electric current flow and a voltage is established perpendicular to both. It can be expanded from the relationship between electric field and current, $\boldsymbol{E}_{\mathrm{I}} = \rho_{ij}\boldsymbol{J}_i$, where $\boldsymbol{J}$ is the current density, $\boldsymbol{E}$ is the electric field, and ρ_{ij} is the resistivity:

$$\rho_{ij} = \rho_{ij}^{(0)} + \rho_{ijk}B_k + \rho_{ijkl}B_kB_l$$

where $\rho_{ij}^{(0)}$ represents resistivity as just discussed, ρ_{ijk} the Hall tensor, and ρ_{ijkl} the longitudinal magnetoresistance. ρ_{ijk} is a third-rank axial tensor (i.e., its sign can be $\pm$), but it can be rearranged as a second-rank antisymmetric tensor. Details are omitted here, but in the most complex case of a triclinic crystal, there are up to nine independent values. Most crystals have a maximum of three independent coefficients.

Elasticity

This is the classic case of a fourth-rank tensor. The reduction of the number of coefficients is discussed in this encyclopedia elsewhere and, in addition, there is an extensive literature regarding relationships between stiffness (equivalent to Young's modulus) and compliance coefficients. Examples of well-established values of the compliance matrix (s_{ij}) are given in **Table 1** to demonstrate variability.

Optoelectronic Effects and Devices

Currently, the application of crystal tensors is playing an essential role in the understanding of optoelectronic effects and the setting up of electrooptic devices. The application of an electric field alters the dielectric constant of many materials and even if a crystal is anaxial (i.e., it behaves isotropically) without the field, it may become uniaxial and thus double refracting with the field. In addition, crystals that are already uniaxial or biaxial usually show a change in

Table 1 Examples of values of compliance at room temperature

Crystal	*Class*	s_{11}	s_{12}	s_{44}	s_{33}	s_{13}	s_{14}	s_{66}
Sodium chloride	*m*3*m*	22.1	−4.5	78.3				
Aluminium	*m*3*m*	15.9	−5.8	35.2				
Copper	*m*3*m*	14.9	−6.3	13.3				
Nickel	*m*3*m*	7.99	−3.12	8.44				
Tungsten	*m*3*m*	2.57	−0.73	6.60				
Tin	4/*mmm*	18.5	−9.9	57.0	11.8	−2.5		135
ADP	$\bar{4}\,2m$	18	7	113	43	−11		162
Quartz	32	12.7	−1.7	20.1	9.7	−1.5	−4.3	

Units are given in $\mathrm{T\,Pa^{-1}}$.
Source: Nye JF (1985) *Physical Properties of Crystals*. Oxford: Oxford University Press.

their refractive index components in the presence of an appropriate electric field.

The electric displacement $\boldsymbol{D}$ can be written as

$$\boldsymbol{D} = \varepsilon\varepsilon_0\boldsymbol{E} = \varepsilon_0(1+\chi)\boldsymbol{E} = \varepsilon_0\boldsymbol{E} + \varepsilon_0\chi\boldsymbol{E}$$

where χ, the electric susceptibility, represents the way the material responds to the electric field, ε is the relative electric permittivity, and ε_0 the permittivity of free space. Alternatively, $\boldsymbol{P}$, the polarization term, can be written as

$$\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}$$

In general, the polarization term can be expanded as a power series in $\boldsymbol{E}$:

$$\boldsymbol{P} = a\boldsymbol{E} + b\boldsymbol{E}\times\boldsymbol{E} + c\boldsymbol{E}^3 + \text{higher terms}$$

The first term on the right-hand side of the equation is the linear term, the next term is the nonlinear quadratic term, and so on. The coefficients a, b, etc., are such that the series converges. It is only in some materials that the coefficients are such that the nonlinear terms are useful. It is common practice to use the reciprocal dielectric constant, the optical impermeability η, and write out the optical impermeability in the form of the change that occurs with application of the field. Thus,

$$\Delta\eta_{ij} = \eta_{ij}(\boldsymbol{E}) - \eta_{ij}(\boldsymbol{E}=0) = r_{ijk}E_k + s_{ijkl}E_kE_j$$

The r_{ijk} components are the linear electrooptic coefficients and are often called the Pockels coefficients. They give rise to the linear Pockels effect. The s_{ijkl} components produce the quadratic effect that is often referred to as the Kerr electrooptic effect. The latter effect can be expected to be small compared with the linear effect, but the linear effect cannot be present in centrosymmetric crystals. Judicious selection of crystals makes both these effects and further effects practicable. In general, both the linear and quadratic coefficients vary with the wavelength of light. This means that the coefficients need to be known for different wavelengths but this can have its advantages; for instance, it can lead to useful phase matching.

The Pockels Effect

The optical indicatrix can be written in the presence of an electric field $\boldsymbol{E}$ as $\eta_{ij}(\boldsymbol{E})x_ix_j = 1$ (where $i,j = 1,2,3$ produce the general form for the indicatrix) and in the absence of a field as $\eta_{ij}x_ix_j = 1$. The η tensor is symmetric so that i and j can be permuted. The coefficients can also be reduced into a contracted form but the contracted forms cannot be transformed according to the rules for tensors. Equalities for the contracted subscripts for the linear electrooptic coefficients become as a consequence $r_{11k} = r_{1k,}$ $r_{22k} = r_{2k}$, $r_{33k} = r_{3k}$, $r_{23k} = r_{32k} = r_{4k}$, $r_{13k} = r_{31k} = r_{5k}$, and $r_{12k} = r_{21k} = r_{6k}$ for $k = 1,2,3$. If these subscripts are now used for linear change of the optical impermeability only, the optical indicatrix becomes

$$\left(\frac{1}{n_1^2} + r_{1k}E_k\right)x_1^2 + \left(\frac{1}{n_2^2} + r_{2k}E_k\right) \times x_2^2 + \left(\frac{1}{n_3^2} + r_{3k}E_k\right)x_3^2 + 2r_{4k}E_kx_2x_3 + 2r_{5k}E_kx_3x_1 + 2r_{6k}E_kx_1x_2 = 1$$

where $k = 1,2,3$. The above equation is a general form of the ellipsoid with the principal axes not assumed to coincide with the principal Cartesian axes x_1, x_2, and x_3. If it is possible for them to coincide, then the equation simplifies and leaves out the crossed x terms. Crystals that are widely used for Pockels devices include lithium niobate, ammonium dihydrogen phosphate (and potassium dihdrogen phosphate, for which the same arguments apply), and gallium arsenide. These three cases are examined to see how the application of crystal tensors leads to slight differences in their use.

Lithium Niobate

The lithium niobate crystal has a trigonal cell of crystal symmetry $3m$. Although the cell can be represented in rhombohedral form, it is easier to use a hexagonal cell (containing six molecules). There is then a threefold rotational axis, which is parallel to the [0 0 0 1] direction (using hexagonal indices) and mirror planes of type $(2\,\bar{1}\,\bar{1}\,0)$. The usual configuration is to apply the electric field along the c- (i.e., x_3-) axis. The x_1-axis is defined such that one of the mirror planes is perpendicular to it in order to obtain the electrooptic components. (The components would come out in a slightly different form if x_2 rather than x_1 was defined to have the mirror plane perpendicular to it.) Carrying out the symmetry operations for $3m$, the contracted form r_{mn} can be further reduced to the form

$$\begin{pmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{pmatrix}$$

With the electric field along the x_3-axis, $k = 3$, and cross-product terms in x_1, x_2, and x_3 disappearing, the general equation becomes

$$\left(\frac{1}{n_o^2} + r_{1k}E_k\right)x_1^2 + \left(\frac{1}{n_o^2} + r_{2k}E_k\right) \times x_2^2 + \left(\frac{1}{n_e^2} + r_{3k}E_k\right)x_3^2 = 1$$

In addition to the changes mentioned, n_1 and n_2 have been made equal to the ordinary refractive index n_o and n_3 to the extraordinary refractive index n_e. It can now be seen that when an electric field $\boldsymbol{E}$ is applied, the semi-axes of the ellipsoid become (see **Figure 1**)

$$n_{x_1} = n_{x_2} = n_o - n_o^3 r_{13} E/2$$

and

$$n_{x_3} = n_e - n_e^3 r_{33} E/2$$

Consequently, for a light beam passing along the x_1-axis, the birefringence becomes

$$n_{x_3} - n_{x_2} = (n_e - n_o) - (n_e^3 r_{33} - n_o^3 r_{13})E/2$$

Thus, the birefringence can be written as

$$\Delta n = \Delta n_{E=0} - \alpha E$$

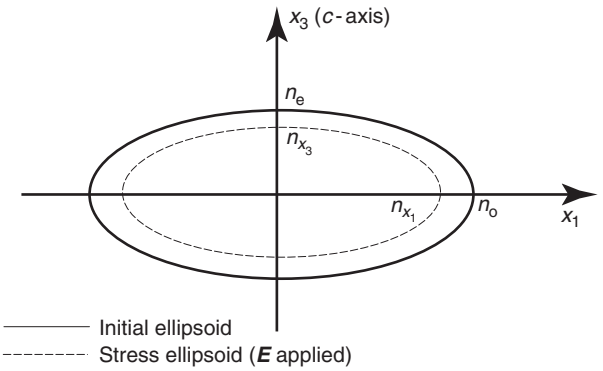


Figure 1 Refractive index ellipsoids for lithium niobate with and without applied electric field.

where

$$\alpha = \tfrac{1}{2}(n_e^3 r_{33} - n_o^3 r_{13})$$

For lithium niobate, in this configuration the magnitude of α is $1.1 \times 10^{-10}\ \mathrm{m\,V^{-1}}$. By altering the magnitude of the electric field, any desired state of polarization can be obtained. Alternatively, by correctly choosing an initial thickness of the crystal, using crossed linear polarizers in front and behind the crystal, and modulating the electric field, the amplitude of the transmitted light can be modulated.

Ammonium Dihydrogen Phosphate

With ammonium dihydrogen phosphate (ADP), a similar setup can be used but ADP is uniaxial. ADP belongs to the crystal class $\bar{4}2m$ and the higher symmetry leads to fewer coefficients:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & r_{41} & 0 \\ 0 & 0 & r_{63} \end{pmatrix}$$

The presence of the r_{63} terms leads to a modified indicatrix ellipsoid. The two alternative cases are discussed. First, an electric field parallel to the x_3 (c-axis) is considered again. The equation for the ellipsoid can be written as

$$\frac{x_1^2 + x_2^2}{n_o^2} + \frac{x_3^2}{n_e^2} + 2r_{63}Ex_1x_2 = 1$$

The presence of the electric field means that the ellipsoid is modified such that it is no longer symmetrical about the x_3-axis and the material acts as biaxial. What was a circular cross section now becomes elliptical with new axes Ox_1' and

Ox_2' rotated by 45° from the original axes. The semi-axes become

$$n_{x_1'} = n_o - \tfrac{1}{2}n_o^3 r_{63}E \quad \text{and}$$
$$n_{x_2'} = n_o + \tfrac{1}{2}n_o^3 r_{63}E$$

As a consequence, a slice of crystal cut perpendicular to the Ox_3-axis will act like a birefringent plate with a fast axis Ox_1' and a slow axis Ox_2', and a birefringence of

$$\Delta n = n_{x_1'} - n_{x_2'} = n_o^3 r_{63}E$$

This configuration is useful for switching. With faces cut perpendicular to Ox_3 (perpendicular to the c-axis) and with no field applied, the crystal appears isotropic to a beam of light passing through in the x-direction. Once an electric field is applied, birefringence is established. If a voltage V is applied to a crystal of thickness L, the electric field is V/L and a phase difference for the two perpendicular components is

$$\delta = \frac{2\pi}{\lambda}L\Delta n = \frac{2\pi}{\lambda}n_o^3 r_{63}V$$

With crossed polarizers before and after the crystal, a sinusoidal form of output is obtained. For no applied voltage, the transmission is zero; for applied voltage, it is finite. Alternatively, the applied voltage can be modulated and a phase modulation produced.

Gallium Arsenide

Gallium arsenide (also indium arsenide, cadmium telluride, and similar materials) has crystal class $\bar{4}3m$. As this is a high-symmetry class, there is only one independent term

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & r_{41} & 0 \\ 0 & 0 & r_{41} \end{pmatrix}$$

and so $r = r_{41}$. With n the refractive index before application of the electric field, the equation applicable in the presence of an electric field becomes

$$\frac{1}{n^2}(x_1^2 + x_2^2 + x_3^2) + 2r \times (x_2x_3E_1 + x_1x_3E_2 + x_1x_2E_3) = 1$$

It is normal to use a configuration in which the electric field is in the $\langle 1\,0\,0\rangle$, $\langle 1\,1\,0\rangle$, or $\langle 1\,1\,1\rangle$ direction. Again, this has application to optical modulation.

The Quadratic Electrooptic Effect (The Kerr Effect)

The effect can occur in crystals of any symmetry. If the Kerr tensor components are represented by s_{ijkl} (notation can vary), then pairs of i, j and k, l permute, but in the reduced notation it is not necessary that $s_{mn} = s_{nm}$. There is a maximum of 36 components, but once again crystal symmetry criteria will reduce the number.

Barium titanate, $BaTiO_3$, is taken as an example. Below 120°C, it has a tetragonal structure of class $4mm$. A suitable direction in which to apply the electric field is [1 1 0]. In this case, the electric field $\boldsymbol{E}$ can be resolved as $E/\sqrt{2}$ along x_1- and x_2-directions and it is zero in the x_3-directions. By applying symmetry arguments, it can be shown that the pattern of components is

$$\begin{pmatrix} s_{11} & s_{12} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{12} & s_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{44} \end{pmatrix}$$

The indicatrix takes the form

$$\left(\frac{1}{n^2} + \frac{s_{11}E^2}{2} + \frac{s_{22}E^2}{2}\right)(x_1^2 + x_2^2) + \left(\frac{1}{n^2} + s_{12}E^2\right)x_3^2 + 2x_1x_2s_{44}E^2 = 1$$

All other terms are zero. However, to get the ellipsoid into a form related to the principal axes, it is necessary to carry out a 45° rotation so that the principal axis is in the same direction as the electric field, this direction becoming a direction of symmetry. If this is done, the revised indicatrix takes the form

$$\left(\frac{1}{n^2} + \frac{s_{11}E^2}{2} + \frac{s_{22}E^2}{2} + s_{44}E^2\right)x_1'^2 + \left(\frac{1}{n^2} + \frac{s_{11}E^2}{2} + \frac{s_{22}E^2}{2} - s_{44}E^2\right) \times x_2'^2 + \left(\frac{1}{n^2} + s_{12}E^2\right)x_3'^2 = 1$$

Comparing this equation with the standard equation for the indicatrix gives

$$(n_{x_1})^{-2} = n^{-2} + (s_{11} + s_{12})E^2/2 + s_{44}E^2$$
$$(n_{x_2})^{-2} = n^{-2} + (s_{11} + s_{12})E^2/2 - s_{44}E^2$$
$$(n_{x_3})^{-2} = n^{-2} + s_{12}E^2$$

Usually, the change in n will be small so that expansion of n to third order gives

$$\begin{aligned} n_{x_1} &= n - n^3(s_{11} + s_{12})E^2/4 \\ &\quad - n^3 s_{44}E^2/2 \\ n_{x_2} &= n - n^3(s_{11} + s_{12})E^2/4 \\ &\quad + n^3 s_{44}E^2/2 \\ n_{x_{31}} &= n - n^3 s_{12}E^2/2 \end{aligned}$$

and the quadratic dependence on the applied electric field.

It is already indicated that ADP and KDP are of crystal class $\bar{4}2m$ and so their Kerr effect matrix is

$$\begin{pmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{13} & 0 & 0 & 0 \\ s_{31} & s_{131} & s_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66} \end{pmatrix}$$

It is common to apply the electric field along the Ox_3 optic axis. The crystal remains uniaxial with the indicatrix now of the form

$$\left(\frac{1}{n_o^2} + s_{13}E_k\right)x_1^2 + \left(\frac{1}{n_o^2} + s_{13}E_k\right)x_2^2 + \left(\frac{1}{n_e^2} + s_{33}E_k\right)x_3^2 = 1$$

If the birefringence before the application of the electric field is $\Delta n_{E=0}$ and after is Δn, then the change in birefringence is

$$\Delta n_{E=0} - \Delta n = \tfrac{1}{2}(n_e^3 s_{33} - n_o^3 s_{13})E^2$$

The first-order change in birefringence due to the Pockels effect was $(1/2)n_o^3 r_{63}E$, so that if it is assumed that $n_e \approx n_o$, the ratio of the second-order effect to the first-order one is $((s_{33} - s_{13})/r_{63})E$. For ADP and KDP, the ratio for $(s_{33} - s_{13})/r_{63}$ is $\sim 10^{-6}$ m V^{-1}. However, because the constants for optical effects are very dependent on the frequencies used (the materials show optical dispersion), it is not realistic here to give tables of magnitudes for specific materials for any of the optical coefficients.

Second Harmonic Generation and Phase Matching

Frequency doubling, also called second harmonic generation, has had important application for increasing the frequency range available from lasers. What is required is nonlinear coupling between two light beams, normally of the same frequency. However, it is necessary that the second harmonic waves generated are all in phase. Birefringent crystals are normally used and matching of indices of refraction is required, necessitating the application of crystal tensors.

In an application of nonlinear optics such as this, it is normal to use polarization $\boldsymbol{P}$ in the formulation. The change in optical impermeability is written as

$$\begin{aligned} \Delta\eta_{ij} &= \eta_{ij}(\boldsymbol{P}) - \eta_{ij}(\boldsymbol{P} = 0) \\ &= f_{ijk}P_k + g_{ijkl}P_kP_j \end{aligned}$$

where $\boldsymbol{P}$ has replaced $\boldsymbol{E}$ and the Pockels and Kerr coefficients are now defined as f_{ij} and g_{ijkl}. The polarization can be expanded in the form

$$\begin{aligned} P_i &= \varepsilon_0\chi_{ij}E_j + 2d_{ijk}E_iE_k \\ &\quad + 4\chi_{ijkl}E_jE_kE_l \end{aligned}$$

where χ_{ij}, d_{ijk}, and χ_{ijkl} are the linear, second-order nonlinear, and third-order nonlinear susceptibilities, respectively. Third-order effects can lead to Raman and Brillouin scattering but it is the second-order effect that can be used for frequency doubling. If the sinusoidal optical fields associated with the two light rays are of the form $E_j^{\omega_1}e^{i\omega_1 t}$ and $E_k^{\omega_2}e^{i\omega_2 t}$, then these can give rise to field variations at frequencies $m\omega_1 + n\omega_2$, where m and n are integers. Taking m and n each as 1 to give a variation $\omega_3 = \omega_1 + \omega_2$,

$$P_i^{\omega_3} = 2d_{ijk}E_j^{\omega_1}E_k^{\omega_2}$$

and if $\omega_1 = \omega_2$, then $P_i^{2\omega} = 2d_{ijk}E_j^{\omega}E_k^{\omega}$. There will be no physical difference between the jk and kl components so that notation suffixes can be once again replaced with matrix notation suffixes. Similar symmetry restrictions apply as they do for the electrooptic matrix components r_{mn} and magnitudes are related but different. Hence, there is a close connection between the electrooptic effects and second harmonic generation.

Now it is necessary to match the indices such that $k^{2\omega} = 2k^{\omega}$ or $\Delta k = k^{2\omega} - 2k^{\omega} = 0$. But $\Delta k = (2\omega/c)(n^{2\omega} - n^{\omega})$, where $n^{2\omega}$ and n^{ω} are the refractive indices at the two frequencies and c the velocity of light in vacuo. Equal values of $n^{2\omega}$ and n^{ω} can be achieved in a birefringent crystal by a matching between refractive indices for the ordinary and extraordinary rays by choosing a correct direction Θ for the light for Δk to be zero. This is shown in **Figure 2** for a negative uniaxial crystal, where differences between the surfaces have been greatly magnified. The condition for Θ is given by

$$\frac{\cos^2\Theta}{(n_o^{2\omega})^2} + \frac{\sin^2\Theta}{(n_e^{2\omega})^2} = \frac{1}{(n_o^{\omega})^2}$$

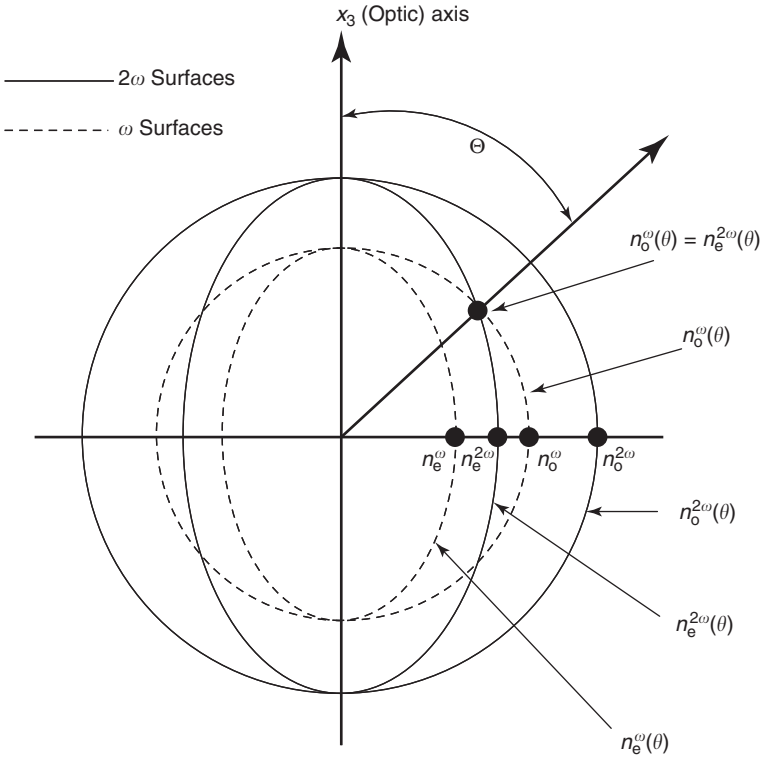


Figure 2 Phase matching between ordinary and extraordinary rays in a negative uniaxial crystal.

Crystalline materials that are often used for such matching are KDP and ADP, which have significant differences between their ordinary and extraordinary refractive indices.

Some Further Applications

There are many other applications of tensors to the interpretation of effects in crystals that have not been mentioned. For instance, in the pyroelectric effect a change in temperature leads to a change of electric dipole moment Δp_i that is related to the temperature change ΔT by the pyroelectric coefficients p_i, which are vector components. A small change in temperature ΔT leads to an electric dipole moment according to $\Delta P_i = p_i \Delta T$. A complication is that the size of the effect also depends on whether the sample is clamped or not clamped during heating. Only crystals with no center of symmetry can exhibit the effect. All materials, which are pyroelectric, must also be piezoelectric. For the latter case, the relationship becomes $P_i = d_{ijk} s_{jk}$, where d_{ijk} are the piezoelectric moduli, which would seem to have 27 components in general but whose number of components can be reduced to 18 for even the most general (triclinic) case. There is also the inverse case of a strain produced by an electric field. As in the case of pyroelectricity, there are differences for the piezoelectric effect and the converse effect according to the applied conditions, for example, whether the sample is under constant stress or constant strain.

When stress is applied to a crystal, there may not only be an electric moment produced but the refractive index may be altered as well. The stress is directional and so an isotropic crystal may become either uniaxial or biaxial. This gives rise to the photoelastic effect (and hence the photoelastic tensor). This effect underlies the Perspex modeling of objects and machinery to study stresses. The application illustrates how even a noncrystalline material that starts isotropic can become anisotropic under stress.

Rather than applying macroscopic stresses to a crystal, acoustic waves can be passed through, leading to another interesting set of applications, namely acoustooptics.

This article has shown the wide range of applications of crystal tensors and it has demonstrated how, especially with low-symmetry crystals, the effects can become complicated and difficult to interpret. Perhaps the main power of the theory of crystal tensors, and hence its application, is to predict the best directions in which to cut and use crystals so as to produce maximized effects and keep the complications to a minimum.

See also: Crystal Symmetry; Crystal Tensors: Introduction.

PACS: 81.40.Jj

Further Reading

Huard S (1977) *Polarization of Light*. Chichester: Wiley and Paris: Masson.
Kelly A, Groves GW, and Kidd P (2000) *Crystallography and Crystal Defects*, 2nd edn. New York: Wiley (1st edn., (1970) London: Longman).
Lipson SG, Lipson H, and Tannhausen DS (1995) *Optical Physics*, 3rd edn. Cambridge: Cambridge University Press.
Lovett DR (1999) *Tensor Properties of Crystals*, 2nd edn. Bristol: Institute of Physics.
McKie D and McKie C (1974) *Crystalline Solids*. New York: Wiley.
Popov SV, Svirko YP, and Zheludev NI (1995) *Susceptibility Tensors for Nonlinear Optics*. Bristol: Institute of Physics.
Press ED (ed.) (1985) *Handbook of Optical Constants of Solids*. Orlando: Academic Press.
Sands DE (1982) *Vectors and Tensors in Crystallography*. Reading, MA: Addison-Wesley.
Yariv A and Yeh P (1984) *Optical Waves in Crystals*. New York: Wiley.

Nomenclature

B	magnetic field
c	velocity of light in vacuo
$\boldsymbol{D}$	dielectric displacement (electric flux density)
d_{ijk}	piezoelectric modulus, second-order nonlinear susceptibility
$\boldsymbol{E}$	electric field
k	wave vector
$\boldsymbol{L}$	thickness
n, n_i	refractive index
n_o	ordinary refractive index
n_e	extraordinary refractive index
P, P_i	polarization
p_i	pyroelectric coefficient
r_{ijk}, r_{ij}	Pockels coefficients, reduced matrix coefficients
s_{ij}	elastic compliance (matrix form)
s_{ijkl}, s_{ij}	Kerr electrooptic coefficients, reduced matrix coefficients
ε_0, ε	permittivity of free space, relative permittivity
η_{ij}	optical impermeability
λ	wavelength
ρ, ρ_{ij}	electrical resistivity
ρ_{ijk}, ρ_{ijkl}	Hall effect, longitudinal magnetoresistance
χ, χ_{ijkl}	electric susceptibility, third-order electric susceptibility
ω	angular frequency
Θ	phase matching angle

Crystal Tensors: Introduction

D R Lovett, University of Essex, Colchester, UK

Introduction

A crystal is inherently symmetrical, hence it presents the same aspect from a number of different directions. The properties of a single crystal generally depend on the direction in which the properties are measured. However, as a result of the crystal symmetry, there are different directions in which the physical properties of the crystal are the same. What is needed is a method of exploiting the inherent symmetry in order to simplify the formulation of the physical properties. The use of tensors is one such tool. This article explains what tensors of different ranks are, and shows how such tensors can be used to describe the directional variation of the physical properties within crystals.

Second-Rank Tensor Illustrated by Electrical Resistivity

The application of the technique and a demonstration of the symbols used can be shown mostly easily by looking at electrical resistivity. Assume an electric current density J passing through a crystalline sample when an electric field E is established within the sample. These quantities are related by

$$E = \rho J$$

where ρ is the electrical resistivity of the sample. If the sample is isotropic, that is, its properties are the same in all directions, then the electric current and the electric field will be in the same direction and ρ is a scalar. This is not necessarily so in a crystal. Resolving the electric field into components in different directions, generally using Cartesian coordinates, leads to components in high-resistivity and low-resistivity directions, and when combined these produce an overall current which is rotated compared with the applied electric field. An everyday analogy is that when a supermarket trolley is pushed it may not move off in exactly the direction of the force! Mathematically, one can now express the relationship between current density and electric field in the form:

$$E_1 = \rho_{11}J_1 + \rho_{12}J_2 + \rho_{13}J_3$$
$$E_2 = \rho_{21}J_1 + \rho_{22}J_2 + \rho_{23}J_3$$
$$E_3 = \rho_{31}J_1 + \rho_{32}J_2 + \rho_{33}J_3$$

Linearity between E and J components has been assumed here, equivalent to assuming Ohm's law. Alternatively, the three combined equations can be written in a matrix form or even more simply by:

$$E_p = \sum_{q=1}^{3} \rho_{pq} J_q = \rho_{pq} J_q$$

Here, the convention is that when the dummy suffix (the repeated q) appears on the right-hand side, then the summation is assumed with the dummy suffix taking values of 1, 2, and 3. Resistivity ρ now has nine components and the resistivity is in the form of a so-called second tensor, which is often represented more generally as T_{pq}. Not all these components are necessarily independent, and for instance, by considering its symmetry, it can be shown (and this is perhaps fairly obvious) that a crystal with a simple hexagonal structure has only two independent components, $\rho_{11} = \rho_{22}$ in the hexagonal (basal) plane, and ρ_{33} perpendicular to this plane.

For the above, the Cartesian axes have been selected carefully with axis x_3 perpendicular to the basal plane. However, the axes may be selected arbitrarily, or alternatively a crystal may have been cut for physical examination in a different way. A significant aspect of using tensors to represent physical quantities is that the tensors can be transformed for appropriate rotation of the Cartesian axes. Consider as the common origin a point O (**Figure 1**) and original axes Ox_1, Ox_2, and Ox_3. One has new axes Ox'_1, Ox'_2, and Ox'_3, where the superscript $'$ is used to distinguish a new axis from an old axis. Then l_{ip} is used to represent the cosine of the angle between the old axis Ox_p and the new axis Ox'_i. A table showing the so-called direction cosines is now constructed. For instance, l_{21} represents the direction cosine of Ox'_2 with respect to Ox_1 and so on:

	x_1	x_2	x_3
x'_1	l_{11}	l_{12}	l_{13}
x'_2	l_{21}	l_{22}	l_{23}
x'_3	l_{31}	l_{32}	l_{33}

Not all the ls are independent, as the mutual orientation of the two sets of axes can be specified by three parameters, so the dependence of the ls can be expressed as:

$$l_{ik} l_{pk} = \delta_{ip}$$

where $\delta_{ip} = 1$ if $i = p$, and $\delta_{ip} = 0$ if $i \neq p$. The squares of the direction cosines must add to one across the rows and down the columns.

Having set up these direction cosines, it can be shown that the resistivity component δ'_{ij} relative to one set of axes is related to the resistivity component δ_{pq} relative to the other set by:

$$\delta'_{ij} = l_{ip} l_{jq} \delta_{pq}$$

and more generally, any second-rank tensor will usually transform (but see later, the section concerning axial tensors) as:

$$T'_{ij} = l_{ip} l_{jq} T_{pq}$$

The tensor has nine components. Similarly, a vector has three components (that is, it is a first-rank tensor). A third-rank tensor has 27 components and transforms according to:

$$T'_{ijk} = l_{ip} l_{jq} l_{kr} T_{pqr}$$

and so on.

For the higher-rank tensors, a calculation of physical effects would become very complex if all the components were independent. However, Neumann's principle, which states that the symmetry elements of any physical property of a crystal must include the symmetry elements of the crystal, can be utilized. The physical properties may and often do include more symmetry than the point group. Physical properties that can be characterized by second-rank tensors must be centrosymmetric, whereas the crystal itself may not be. This may be seen by expressing

$$p_i = T_{ij} q_j$$

Reversing the directions of p and q changes the sign of each of their components but leaves T_{ij} unchanged.

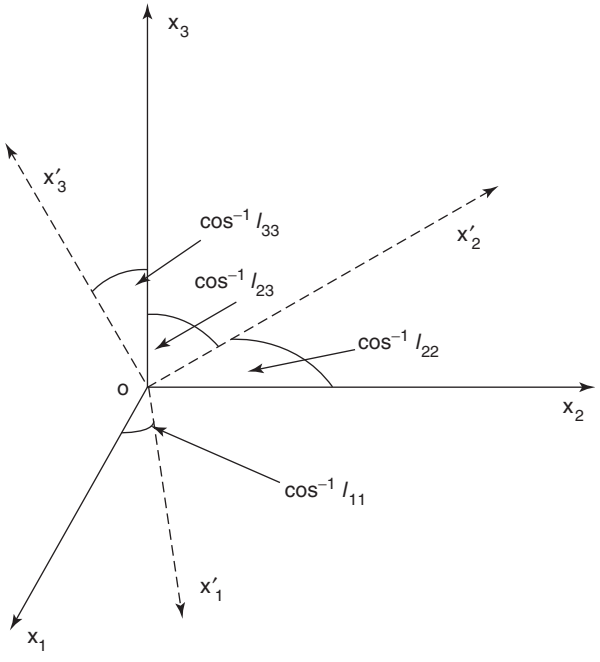


Figure 1 Rotation of axes.

Representation Quadric for Symmetrical Second-Rank Tensor

Second-rank tensors are said to be symmetrical if $T_{pq} = T_{qp}$. Most common second-rank tensors describing physical properties are symmetrical although the thermoelectric tensor is an exception. Alternatively, if $T_{pq} = -T_{qp}$, the tensor is said to be antisymmetrical. A quadric can be represented by the equation

$$S_1x_1^2 + S_2x_2^2 + S_3x_3^2 = 1$$

provided it is described relative to the principal axes. Similarly, if a symmetrical second-rank tensor of general form,

$$T_{pq} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}$$

is described with respect to the principal axes, it becomes:

$$T_{pq} = \begin{bmatrix} T_{11} & 0 & 0 \\ 0 & T_{22} & 0 \\ 0 & 0 & T_{33} \end{bmatrix} = \begin{bmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ 0 & 0 & T_3 \end{bmatrix}$$

where T_1, T_2, and T_3 are the principal coefficients of the tensor T_{pq}. Further consideration of the condensed (single subscript) notation for the two subscripts will come a little later. If T_1, T_2, and T_3 are all positive, the tensor can be represented by an ellipsoid whose semiaxes have lengths of $1/\sqrt{T_1}, 1/\sqrt{T_2}$, and $1/\sqrt{T_3}$. If two of the principal components are positive and one is negative, the surface is a hyperboloid of one sheet; if one is positive and two are negative, it is a hyperboloid of two sheets, and if all the three are negative, the surface is an imaginary ellipsoid. However, for conductivity and resistivity the coefficients are positive, thus giving a real ellipsoid. Only three components are ever necessary to define a second-rank tensor referred to as the principal axes. However, if these axes need to be defined in relation to conventional crystallographic axes as is required for monoclinic and triclinic crystals, then up to a further three coefficients are necessary.

Second-Rank Tensor as Illustrated by Thermal Conductivity and Resistivity

In order to see easily and visually what happens if a crystal is not aligned with a principal axis in the same direction as the driving force within the crystal, it is advantageous to look at the heat flow. In an isotropic medium, heat conductance can be represented by the equation:

$$h = -k \text{ grad } T$$

where k is the thermal conductivity and is a constant having the same value in any direction in the medium. Also,

$$\text{grad } T = \partial T/\partial x_1 + \partial T/\partial x_2 + \partial T/\partial x_3$$

and this is the temperature gradient. The heat flow h per unit area is in the direction of the greatest fall of temperature, that is, in the direction of –grad T. In suffix notation, the equation for heat conduction may be written as

$$h_i = -k\partial T/\partial x_i$$

But just as in the case of electrical resistivity, thermal conductivity may vary with crystallographic direction, and hence the heat flow may not be in the same direction as the maximum temperature gradient. The thermal conductivity equation is then written as

$$h_i = -k_{ij}\partial T/\partial x_j$$

where k_{ij} is the second-rank thermal conductivity tensor. Alternatively, one can use the thermal resistivity tensor r_{ij} to obtain

$$\partial T/\partial x_i = -r_{ij}h_j$$

k_{ij} and r_{ij} are symmetrical tensors having three components when referred to their principal axes. If this is done, then $k_1 = 1/r_1$, etc. Otherwise, in general, reciprocals cannot be taken.

To show the more general use of tensors, consider the heat flow in two crystalline samples of different geometries as shown in **Figure 2a**. The first example shows a very thin sample with a large heat source

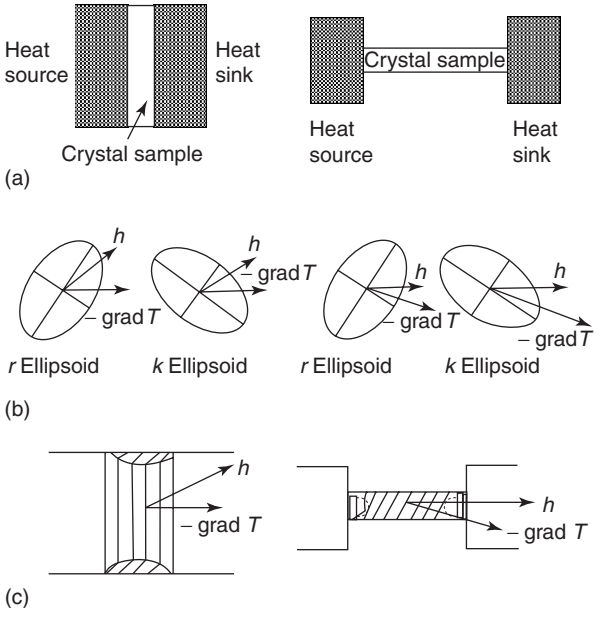


Figure 2 (a) A thin wide and a long narrow sample, (b) cross sections of the conductivity and resistivity ellipsoids, and (c) distortion of the isothermals.

and sink on opposite faces in a setup that is similar to the classical disk experiment for measuring the thermal conductivity of an isotropic sample of poor conductivity. The isothermal surfaces must be parallel to the end faces and it is easier, when one has an arbitrarily oriented crystal, to represent the setup by the conductivity equation. Alternatively, if a long narrow sample is used as in **Figure 2b**, then the heat flow is parallel to the sides and the resistivity equation is preferable. **Figures 2c** and **2d** show the k and r ellipsoid in each case (cross sections of the three-dimensional ellipsoids), whereas **Figures 2e** and **2f** show that some distortion of the isothermals occurs near the edges. It can be seen from these diagrams that measurements and their interpretation are easier if the crystal is aligned with major axes parallel to the heat source and sink.

Third-Rank Tensors

These arise when a tensor relates a vector to a second-rank tensor. If no simplification were possible, then there would be 3^3, that is, 27, components. In practice, the largest possible number is 18 in triclinic crystals without inversion symmetry, but the number reduces rapidly with crystals showing higher symmetry. Properties that involve third-rank tensors include the piezoelectric effect, the linear electrooptic (Pockels) effect, and the Hall effect. However, to demonstrate the reduction of the number of tensor components further, consider the fourth-rank tensors.

Fourth-Rank Tensor as Illustrated by Elasticity

Young's modulus is given by stress divided by strain and, as both stress and strain can be represented by second-rank tensors, the modulus can be represented by a fourth-rank tensor. Stress is the force acting in any direction divided by the area. It is easiest to consider forces acting on a cube of unit dimensions (**Figure 3**). Here, it can be seen that σ_{22} is the normal force acting on one of the faces, and σ_{12} and σ_{32} are the shear components. For there to be no turning of the overall cube, $\sigma_{ij} = \sigma_{ji}$.

Strain e is the extension divided by the original length, and can either be tensile of the form e_{11} or a shear of the form e_{12}, etc. The problem in defining the shear strain is that any pure rotation must be eliminated. Any tensor can be divided into the sum of a symmetrical and an antisymmetrical tensor. Without going into detail here, to remove the rotation, the symmetrical parts of the shear strain components e_{ij} are taken, and the new tensor is called ε. Also, $\varepsilon_{ij} = \varepsilon_{ji}$ for no overall rotation of the body. The diagonal components of the tensor are the tensile strains (positive) or compressive (negative) strains, and the off-diagonal components are the shear strains.

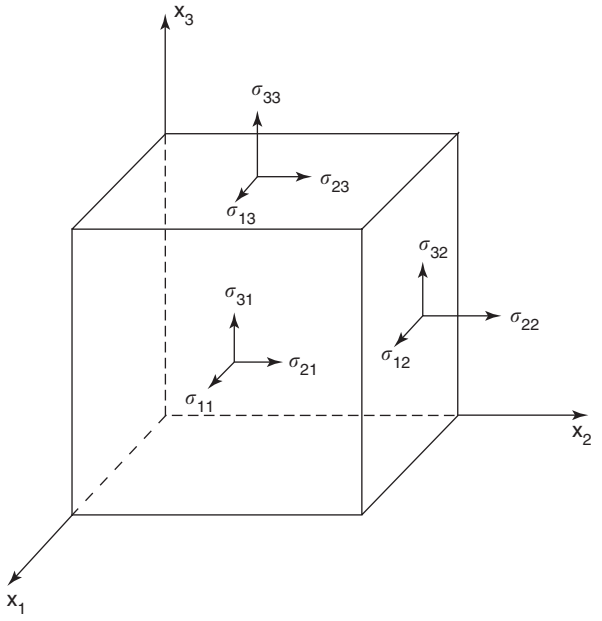


Figure 3 Normal and shear components of stress on a cubic element.

Strain is the response of the crystal to an influence. If this influence arises from a change within the crystal itself, then the resulting tensor will reflect the symmetry properties of the crystal. If however, as is more usual, the strain arises from external forces, then these will not be related to the crystal properties. The tensor will not conform to Neumann's principle and is an example of a field tensor. Tensors that are governed by Neumann's principle are called "matter" tensors.

Having looked at stress and strain, it is possible to set up the relationship for elasticity. As before, starting with an isotropic material such as a polycrystalline sample, Hooke's law may be expressed in the form:

$$\sigma = c\varepsilon$$

where c is Young's modulus or the elastic stiffness. Alternatively,

$$\varepsilon = s\sigma$$

where s is the elastic compliance. The above two equations become in tensor notation:

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl} \qquad \text{and} \qquad \varepsilon_{ij} = s_{ijkl}s_{kl}$$

where c_{ijkl} is the fourth-rank tensor for stiffness, and s_{ijkl} is the fourth-rank tensor for compliance. As one already has $\sigma_{ij} = \sigma_{ji}$ and $\varepsilon_{ij} = \varepsilon_{ji}$, the result obtained is:

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{jilk} \qquad \text{and}$$
$$s_{ijkl} = s_{jikl} = s_{ijlk} = s_{jilk}$$

Thus the number of components has been reduced to 36. These can be represented in a 6×6 matrix, a type of representation which is used fairly generally. What must be remembered is that as soon as components are represented in a matrix form, the rules of transformation cannot be followed for change of axes.

Making use of the symmetry in the first and last suffixes and following standard convention:

Tensor notation suffixes	11	22	33	23	32	31	13	12	21
Matrix notation suffixes	1	2	3	4		5		6	

As a consequence, the relationships are expressed as

$$\sigma_m = c_{mn}\varepsilon_n \quad \text{and} \quad \varepsilon_m = s_{mn}s_n$$

where m and n take on the values 1 to 6. Thus, the principal components of stress take on the notation $\sigma_1, \sigma_2,$ and σ_3. This is a common form of notation for the principal components of a second-rank tensor. This is so even if the tensors are not involved in the setting up of a higher-rank tensor in which a matrix notation is fully used. However, there is a complication in that one does not end up with the correct matrix forms of the equations for the relationships between stress and strain without some modification of the transformations. There are more pairs of type 23 than of those type 11, for instance. The conventional way around for stress and strain is to transform as follows:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \rightarrow \begin{pmatrix} \sigma_1 & \sigma_6 & \sigma_5 \\ \sigma_6 & \sigma_2 & \sigma_4 \\ \sigma_5 & \sigma_4 & \sigma_3 \end{pmatrix}$$

$$\begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} \rightarrow \begin{pmatrix} \varepsilon_1 & \varepsilon_6/2 & \varepsilon_5/2 \\ \varepsilon_6/2 & \varepsilon_2 & \varepsilon_4/2 \\ \varepsilon_5/2 & \varepsilon_4/2 & \varepsilon_3 \end{pmatrix}$$

In addition, factors of 2 and 4 are to be introduced as follows:

$$\begin{aligned} s_{ijkl} &= s_{mn} \quad \text{for } m = 1,2,3 \text{ and } n = 1,2,3 \\ 2s_{ijkl} &= s_{mn} \quad \text{for either } m \text{ or } n = 1,2,3 \\ 4s_{ijkl} &= s_{mn} \quad \text{for both } m \text{ and } n = 1,2,3 \end{aligned}$$

But

$$c_{ijkl} = c_{mn} \text{ for all } m \text{ and } n$$

The matrix for stiffness becomes:

$$\begin{pmatrix} c_{11} & c_{12} & .. & .. & .. & c_{16} \\ c_{21} & c_{22} & .. & .. & .. & c_{26} \\ c_{31} & .. & .. & .. & .. & .. \\ c_{41} & .. & .. & .. & .. & .. \\ c_{51} & .. & .. & .. & .. & .. \\ c_{61} & .. & .. & .. & .. & c_{66} \end{pmatrix}$$

and similarly for compliance.

For other conversions of fourth-rank tensors to the matrix notation, it is necessary to check just how the transformation has been carried out.

Having reduced the number of components so far, it can be further shown (by energy considerations when stresses are applied) that c_{mn} will be the same as c_{nm}, and s_{mn} will be the same as s_{nm}. The number of independent components is then reduced to 21. This is still a large number but just as with second-rank tensors, one can look at the effect of different crystal symmetries. For higher-symmetry crystals, this can be achieved more easily by inspection to see how pairs of suffixes transform with a change of axes as symmetry operations specific to the crystal are carried out. **Table 1** shows the maximum number of independent components that are left for fourth-rank tensors and for other tensors for a selection (but only a selection) of crystal systems. The table gives an idea of the variation of the number of coefficients. Although, in principle, one can go to higher-rank tensors above the fourth, this is less usual.

Axial Tensors

There are some cases where axes and hence vectors transform not according to $r'_i = l_{ip}r_p$ (i.e., as $r'_1 = l_{11}r_1 + l_{12}r_2 + l_{13}r_3,$ etc.) but according to $r'_i = \pm\, l_{ip}r_p$ (i.e., as $r'_1 = \pm l_{11}r_1 \pm l_{12}r_2 \pm l_{13}r_3,$ etc.), where $+$ means a transformation that leaves the hand of the axes unchanged and $-$ changes the hand (whether left-handed or right-handed). Such vectors are called axial vectors whereas true vectors are called polar vectors. A polar vector can be shown to be an arrow pointing in a specific direction. An axial vector needs a right-handed or left-handed screw to represent it. Polar vectors (angular velocity would be such an example) are referred to as right-handed axes.

Axial tensors can similarly exist and they are sometimes called pseudotensors. The most common example is optical activity which is a rotation of the plane of polarization as light passes through a crystal and where the property transforms according to $T'_{ij} = \pm l_{ip}l_{jq}T_{pq}$. In this case, the plus and minus

Table 1 Maximum number of independent tensor components for a selection of examples

	Examples					
International notation/ crystal system	*Resistivity (second-rank symmetric (polar))*	*Hall effect (second-rank antisymmetric (polar))*	*Linear electrooptic effect (third-rank (polar))*	*Elastic compliance (fourth-rank double symmetric (polar))*	*Piezo-optic effect (fourth-rank optical (polar))*	*Optical activity (axial second-rank (symmetric))*
*m*3*m* cubic	1	1	0	3	3	0
432 cubic	1	1	0	3	3	1
23 cubic	1	1	1	3	4	1
6/*mmm* hexagonal	2	2	0	5	6	0
6 hexagonal	2	3	4	5	8	2
3*m* trigonal (rhombohedral)	2	2	4	6	8	0
4/*mmm* tetragonal	2	2	0	6	7	0
4 tetragonal	2	3	4	7	10	2
1 triclinic	6	9	18	21	36	6

signs give rise to positive and negative rotations of the plane of polarization.

To see what happens in particular crystals for particular tensor applications, it is necessary to look at the transformations for different crystal symmetries. Both rotation planes and mirror planes need to be considered. Consider first the transformation of an axial vector. It is necessary to combine the effects of both the transformation of the coordinate axes and the sense of rotation. Doing this, it can be shown that for a diad (twofold rotation) axis and for threefold, fourfold, and sixfold axes, axial vectors behave in a similar way to polar tensors. So $r'_i = l_{ip} r_p$ for axes of rotation. On the other hand, when the effect of mirror planes is looked at and also when the effect of a centre of symmetry is considered, there is a reversal of sign compared with the situation for a polar tensor. Hence, $r'_i = \pm l_{ip} r_p$ for planes and centers of symmetry.

Similarly for an axial tensor, the specific symmetry needs to be considered. For a second-rank tensor applied to a crystal having a centre of symmetry, 1 goes to -1, 2 to -2, 3 to -3. In addition, there is a change of rotation so that with two negative signs for the change of axes and the additional change for rotation, $T'_{ij} = T_{pq}$ is the result.

Tensors for Crystal Optics

Crystal optics incorporates a more complex situation than that for physical effects such as electrical and thermal conductivity. Passage of light involves propagation of transverse electromagnetic waves. Unless plane-polarized light is used, there will be wave vibrations (electric and magnetic vectors) in all directions orthogonal to the direction of propagation. If the direction of propagation is taken to be x, then for a single wave, if the electric vector is varying in the y direction, the magnetic vector varies in the z direction. It is the electric vector that is taken as the defining vector. For nonpolarized light, the electric vectors for all the waves can be resolved into orthogonal y and z directions. It is the interaction of these electric vector components, with possibly different strengths, with the atoms of the crystal that can give rise to the light being slowed down by different amounts for the two components. These two components, the component resolved in the y direction and the component resolved in the z direction, may each have its own velocity. The crystal in question exhibits a so-called birefringence or double refraction. So in understanding what will happen and making use of the properties of crystal tensors, one must in general also know about the nature of the light being used, that is, whether it is polarized (and in which direction it is polarized) or unpolarized. Even if a crystal does not exhibit birefringence in its natural state, it may do so under stress or under the application of independent electric or magnetic forces. This can lead to a number of interesting and useful effects (refer to "See also" section).

Here, the main implications for uniaxial and biaxial crystals are considered. In an isotropic medium, the dielectric properties at optical frequencies are represented by the equation

$$D = \varepsilon_0 K E = \varepsilon E$$

E is the electric field, ε_0 is the permittivity of free space, K is the dielectric constant for the isotropic medium, ε is the permittivity of the medium, and D is the dielectric displacement or electric flux density. The corresponding equation for an anisotropic

medium involves K and ε becoming symmetrical second-rank tensors:

$$D_i = \varepsilon_0 K_{ij} E_j = \varepsilon_{ij} E_j$$

This equation is similar to that for heat flow using the conductivity tensor. Also, just as the reciprocal tensor can be used in heat flow, that is the thermal resistivity tensor, the reciprocal dielectric tensor $(K^{-1})_{ij}$ can be similarly used. It is normally represented by the symbol η_{ij} and is called the optical impermeability. The ellipsoid is given the special name of the optical indicatrix. If the principal values of the refractive index are defined as $n_1, n_2,$ and n_3, and $x_1, x_2,$ and x_3 are used for the principal axes of this representational ellipsoid, the equation for the ellipsoid is

$$\frac{x_1^2}{n_1^2} + \frac{x_2^2}{n_2^2} + \frac{x_3^2}{n_3^2} = 1$$

The principal values of the refractive index are related to the principal components of the dielectric tensor and the optical impermeability tensor by

$$n_1 = \sqrt{K_1} = 1/\sqrt{\eta_1}$$
$$n_2 = \sqrt{K_2} = 1/\sqrt{\eta_2}$$
$$n_3 = \sqrt{K_3} = 1/\sqrt{\eta_3}$$

An alternative way of writing the equation for the ellipsoid is

$$\eta_1 x_1^2 + \eta_2 x_2^2 + \eta_3 x_3^2 = 1$$

For cubic crystals the principal refractive indices are equal, the indicatrix is consequently a sphere, and no double refraction occurs. For the cases of tetragonal, hexagonal, and trigonal crystals, the indicatrix becomes an ellipsoid of revolution about the principal symmetry axis. This principal axis is called the optic axis. There are two principal values of the refractive index and the crystals are called uniaxial. The refractive index, as measured in the direction of the principal axis, would correspond to n_1 in the above equation. In this case it is generally referred to as n_E, the extraordinary refractive index. The other principal value corresponding to $n_1 = n_2$ is referred to as n_O, that is the ordinary refractive index. The electric vectors of the light rays passing in the direction of the optic axes lie within this cross-sectional plane with $n_1 = n_2$, and all the rays travel in the direction of the optic axis with equal speed (there is no birefringence).

For orthorhombic, monoclinic, and triclinic crystal systems, the ellipsoid for the indicatrix is triaxial. There are now two circular cross sections (**Figure 4**), and the two directions perpendicular to these circular cross sections are referred to as the optic axes. Optically, the crystals are referred to as biaxial. There is no birefringence exhibited in these two directions as the electric vectors for the light lie within one of the circular planes of the indicatrix.

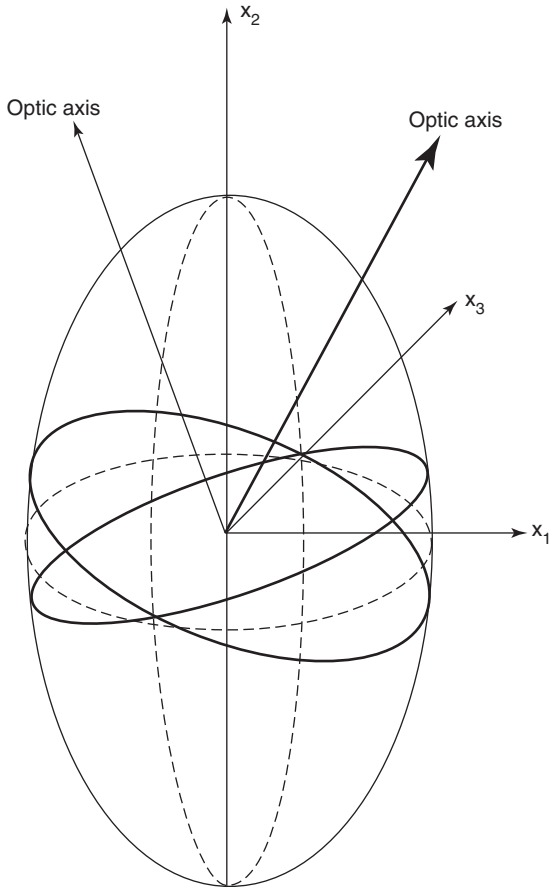


Figure 4 Ellipsoid for the indicatrix in a biaxial crystal.

Wave Surfaces

The wave surface, that is the optical front, for light passing through a crystal is different from the indicatrix but related to it. Consider a point surface of light and see where the light gets to over a particular time. If the light travels outward with the same speed in all directions, the surface is a sphere. But as has been seen not only will there be a variation of speed with direction for many crystals, but the speed will also vary for variation of orientation of the electric vectors perpendicular to the direction of passage of the light. As indicated above, in the direction of the optical axis in a uniaxial crystal, all the rays proceed with the same speed and this is determined by the refractive index n_O. Perpendicular to the optic axis, rays move with the full range of possible velocities determined by the extreme (principal) values of the refractive index, n_O and n_E. Consequently, the wave surfaces for a uniaxial crystal can take one of the two alternatives shown in **Figure 5**. Which alternative applies depends

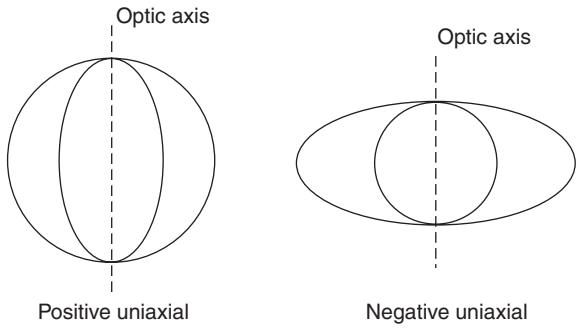


Figure 5 Wave surfaces for uniaxial crystals.

on whether $n_O > n_E$ (positive uniaxial) or $n_O < n_E$ (negative uniaxial). There is an extra complication in that surfaces are also often used to show the variation of the wave vector k with direction. As $k = 2\pi/\lambda$, the wave vector surface gets larger as the wavelength gets smaller, whereas the distance traveled and hence the size of the wave surface gets smaller.

The situation is more complicated in biaxial crystals and this is discussed for specific applications elsewhere in the encyclopedia (refer "See also" section). It is the application of crystal tensors to optical materials as used in electronic devices which has proved to be of particular importance in recent years.

See also: Conductivity, Electrical; Crystal Tensors: Applications; Group Theory; Piezoelectricity; Point Groups; Polarizabilities.

PACS: 81.40.Jj

Further Reading

Kelly A, Groves GW, and Kidd P (2000) *Crystallography and Crystal Defects*, 2nd edn. New York: Wiley. (1st edn. (1970) London, Longman).

Lipson SG, Lipson H, and Tannhausen DS (1995) *Optical Physics*, 3rd edn. Cambridge: Cambridge University Press.

Lovett DR (1999) *Tensor Properties of Crystals*, 2nd edn. Bristol: Institute of Physics.

Nye JF (1985) *Physical Properties of Crystals*. Oxford: Oxford University Press.

Sands DE (1982) *Vectors and Tensors in Crystallography*. Reading, MA: Addision-Wesley.

Crystalline Organic Metals

T Ishiguro, Kyoto University, Kyoto, Japan

Introduction

Organic materials showing high electrical conductivity with a weak temperature dependence are called organic metals. The number of the charge carriers in metals is kept almost constant against temperature variation, in contrast to semiconductors, in which the number is determined by thermal activation processes. In the organic conductors, the charge carriers consist of the delocalized π-electrons at the highest occupied molecular orbital (HOMO) or the lowest unoccupied molecular orbital (LUMO), consisting of a $2p_z$ orbital of carbon (C) atom, a $3p_z$ orbital of sulfur (S) atom, or a $4p_z$ orbital of selenium (Se) atom. Structurally, organic metals are classified into two categories. Polymeric materials represented by polyacetylene and polypyrrole form the first group. A typical structural motif of these materials is an assembly of fibrils formed by bundled polymer chains, which normally cannot be prepared in single crystalline forms. In most cases, the polymers themselves are insulators and the metallic properties are induced by chemical doping. Crystalline organic materials, forming the second group, are characterized by the regular arrays of molecules (TTF, TMTTF, etc.) as shown in **Figure 1**.

The vast majority of crystalline organic metals consists of two kinds of molecules combined with the charge transfer interaction. By transferring the charge from one type of molecules (donors) to the other type (acceptors), partial filling of the bands is achieved, allowing movement of conduction of the π-electrons. In special cases, however, single molecular metals can be formed, such as in single crystals of $Ni(tmdt)_2$. In these materials, the minima of a LUMO are lower than the maxima of a HOMO so that the π-electrons are transferred between the orbitals within the molecule, and the conductance becomes metallic without the charge transfer. Also, the assembly of C atoms connected with sp^2-bonding and containing delocalized π-electrons, such as a fullerene C_{60} molecule, functions as a metal, when it is combined with an alkali metal (A) to form A_3C_{60}, through a charge transfer from A to C. Starting with the molecular and electronic structure, the features of the conducting and superconducting states are presented in this article.

Molecular Structure and Electronic Structure

Molecular Structure

The molecules shown in **Figure 1** are nearly planar and can be assembled to form face-to-face stacks so

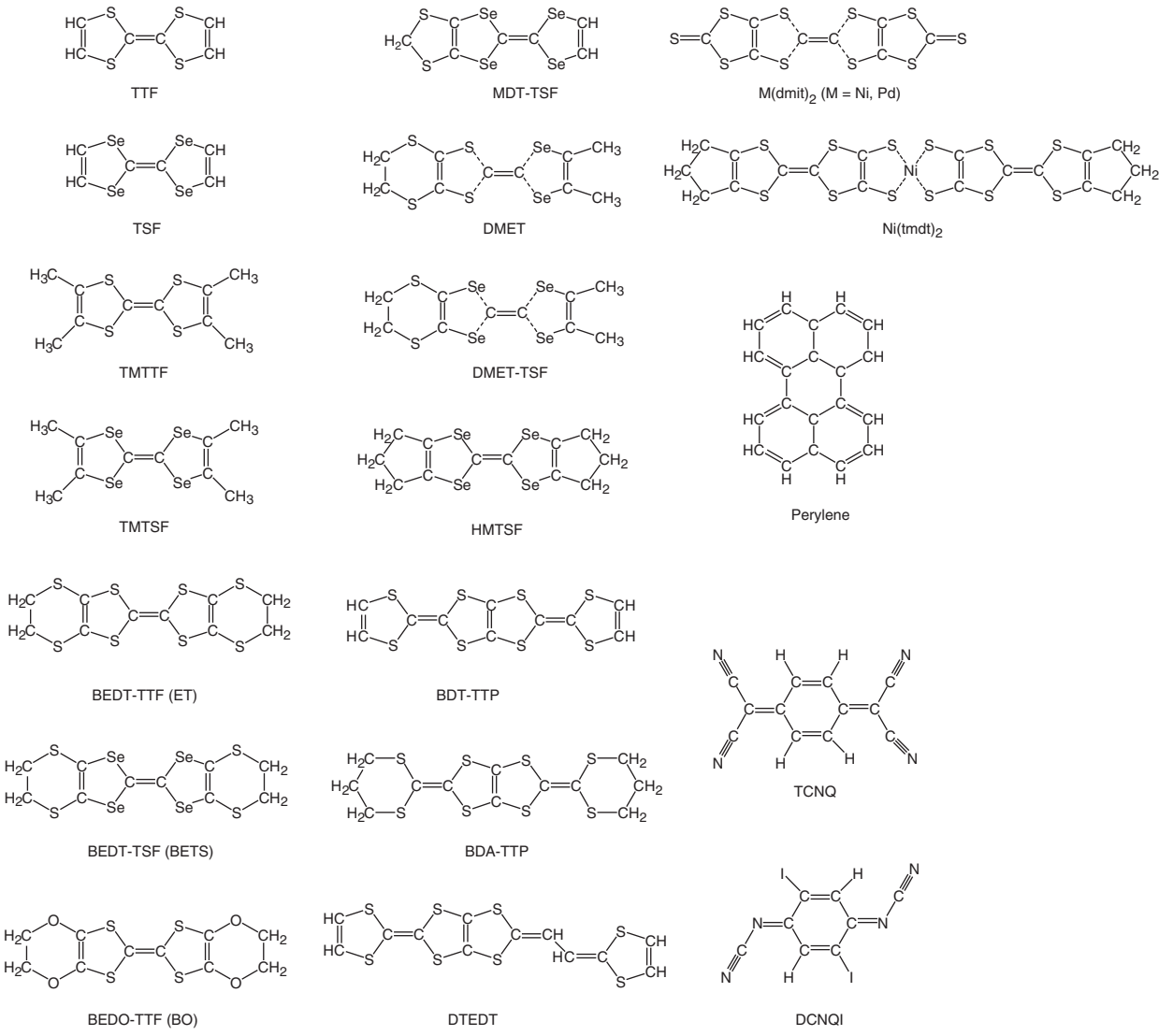


Figure 1 Chemical formulas of basic molecules constructing crystalline organic conductors.

that the π-electron orbitals overlap on adjacent molecules. When the donor and acceptor molecules are stacked separately to form segregated columns, each column works as a one-dimensional electrical conductance channel. When the overlap of π-electron orbitals in the side-by-side direction is not negligible, the conductivity can be quasi-two-dimensional (Q2D) or quasi-three-dimensional (Q3D). The pronounced anisotropy in the electron transfer brings about conductors of a restricted dimensional nature.

In the charge transfer salt TTF·TCNQ, which triggered an excitement about the organic metal in the 1970s, the electronic charge is transferred from TTF to TCNQ. Since the charge transfer is incomplete (0.59 e per molecule, where e is the charge of the electron), the columns of TTF and TCNQ work as hole and electron conductors, respectively. Since the electron transport between the columns is much weaker than within the columns, the electronic structure is characterized by two channels of Q1D nature (**Table 1**). In the case of $(TMTSF)_2X$ and $(TMTTF)_2X$

Table 1 Material types, their conducting channels and the conductance dimensionality (Q1D, Q2D, and Q3D)

Materials	*Conducting channels*	*Dimensionality*
$(TMTSF)_2X$	TMTSF columns	Q1D
TTF·TCNQ	TTF columns and TCNQ columns	Q1D
β-$(BEDT\text{-}TTF)_2X$	BEDT-TTF columns with intercolumn coupling	Q2D
κ-$(BEDT\text{-}TTF)_2X$	BEDT-TTF layers	Q2D
$Ni(tmdt)_2$	$Ni(tmdt)_2$ columns with intercolumn coupling	Q3D
A_3C_{60}	f.c.c.-packed C_{60}	3D

($X = PF_6$, AsF_6, ClO_4, etc.), the columns of TMTSF and TMTTF work as Q1D conductors, while the counter anions X^- having a closed electronic shell do not form a conduction path. For molecules such as BEDT-TTF and BEDT-TSF (also abbreviated as ET and BETS, respectively), forming charge transfer salts (BEDT-TTF)$_2X$ ($X = I_3$, $Cu(NCS)_2$, $Cu[N(CN)_2]Cl$), and (BEDT-TSF)$_2X$ ($X = GaCl_4$, $FeCl_4$), the side-by-side type contacts of the molecules bring about a π-electron transfer that is as substantial as the face-to-face type contacts. Consequently, the molecules form conducting layers separated by the layers of counter anions of X^-. Within the layer plane, BEDT-TTF and BEDT-TSF molecules can be arranged in several different patterns and the resulting crystalline structures are differentiated by Greek letters such as α, β, κ, λ, θ, etc. The arrangement in the κ-type, however, has no columnar motif but a checkerboard pattern with pairs of BEDT-TTF molecules. In these cases, the electronic structure has a Q2D character.

For A_3C_{60} compounds, the spherical C_{60} molecules form a face-centered-cubic (f.c.c.) lattice, while the alkali metal atoms occupy the interstitial openings of the cubic lattice regularly. The resultant structure remains 3D providing the isotropic conductor. The single component conductor $Ni(tmdt)_2$ also shows the quasi-three-dimensionality due to the intercolumn interaction.

Electronic Structure

The electronic structure of low-dimensional organic metals, such as (TMTSF)$_2X$, (BEDT-TTF)$_2X$, has been calculated by the tight-binding band model, taking into account the electron transfer between the molecular orbitals, derived by the extended Hückel method, but neglecting the transfer in the least conductance direction. This relatively simple method has been applied successfully to derive Fermi surfaces in these conductors, due to the fact that the HOMO bands, serving for the charge transport, are energetically separated from the other bands. In these salts, according to the chemical formula with the monovalent anion X^-, the charge per donor molecule is $+0.5\,e$, so that the bands formed by the orbitals of the donor molecules should be quarter-filled. In reality, however, the resultant electronic band is half-filled because of the dimerization of the donor molecules. The cross sections of the Fermi surfaces for typical organic metals are illustrated in **Figure 2**. To obtain more detailed and a 3D electronic structure including the least conduction direction, methods such as the local density-functional calculation and the first-principles calculation are applied. It is also important to take into account the electron correlation energy, which can be comparable to the transfer energy and change the metal to be an insulator for the half-filled case.

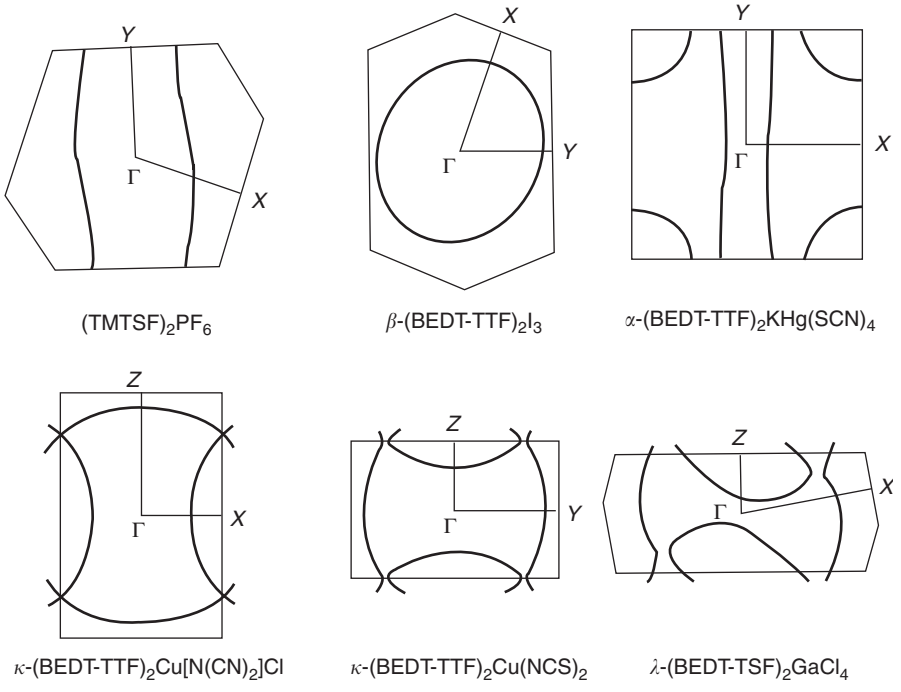


Figure 2 Fermi surface cross sections of typical crystalline organic metals.

Electron Dynamics in Crystalline Organic Metals under Magnetic Field

For the 2D organic metals, the magnetoquantum oscillations (the Shubnikov–de Haas and the de Haas–van Alphen effect) have been observed very clearly in the resistivity, the magnetization, the heat capacity and so on, due to the high density of states, characteristic to the band structure with restricted dimensionality, and the high-crystalline perfection of the compounds. From the experimental results, the cross-sectional areas of the Fermi surface and the effective mass of the charged carriers have been derived. The reported effective mass values are listed in **Table 2**. The larger mass values from the de Haas effect were obtained in the higher magnetic field region, associated with the magnetic breakdown phenomena between the adjacent Fermi surfaces. The cyclotron resonance due to the transitions of the electrons between the Landau levels gave lower mass values compared to those obtained from the de Haas effect. The difference has been ascribed to the effect of the electron correlation on the mass derived from the quantum oscillation. The effective masses derived from the plasma frequency found in the optical spectra are also listed in **Table 2**. The difference in the mass values is related to the electronic polarization within the conducting plane.

In the Q2D metal with a weak interlayer interaction such as (BEDT-TTF)$_2X$ salts, the Fermi surface, possessing a cylindrical shape in a purely 2D case, exhibits periodical swelling (warping) along the cylinder axis. The magnetic phenomena such as magnetoresistance, magnetization, and magneto-acoustic absorption exhibit an oscillatory change with the angle of the magnetic field from the cylinder axis without showing a shift in the oscillation phase by the field strength, in contrast to the de Haas effect. This is called the angle-dependent magnetoresistance oscillation (AMRO) and is well explained by the semiclassical theory, demonstrating that all of the electron orbits near the Fermi surface enclose the same cross-sectional area at certain magnetic field directions. From the analysis of the oscillation phase, the diameters of the Fermi surface cross section can be derived.

For Q1D metals such as (TMTSF)$_2X$ salts, the electron orbits on the Fermi surface are not closed in the momentum space (**Figure 2**), except in a very limited region. When the magnetic field is applied almost perpendicular to the most conducting axis, the electrons near the Fermi surface traverse the Fermi surface, being driven by the Lorentz force. In the real space, this makes a wavy motion due to a back-and-forth oscillation corresponding to the Bragg reflection at the Brillouin zone boundaries. The angular dependence of the magnetoresistance shows the extrema at certain crystalline directions, when the trajectories of the electrons, traversing the momentum space, satisfy certain commensurability conditions with respect to the geometry of the Brillouin zone. The observed angle dependence enables one to derive the transfer integrals and the geometrical parameters on the Fermi surface of the Q1D metal. The wavy amplitude decreases inversely proportional to the magnetic field, resulting in the confinement of the electron motion within the layer at a very high field.

Table 2 Effective masses in units of the free electron mass m_0

Materials	*Effective mass (m_0)*
β_L-(BEDT-TTF)$_2$I$_3$	2.0[c]
β_H-(BEDT-TTF)$_2$I$_3$	4.65[a]
β-(BEDT-TTF)$_2$IBr$_2$	4.2[a]
κ-(BEDT-TTF)$_2$I$_3$	1.9[a] 3.9[a] 2.2[c] 3.0[c]
κ-(BEDT-TTF)$_2$Cu(NCS)$_2$	3.5[a] 6.5[a] 1.18[b] 4.1[c] 5.5[c]
κ-(BEDT-TTF)$_2$Cu[N(CN)$_2$]Br under 9 kbar	0.95[a] 6.4[a]
κ-(BEDT-TTF)$_2$Cu[N(CN)$_2$]Cl under 6 kbar	1.7[a] 3.5[a]
θ-(BEDT-TTF)$_2$I$_3$	1.8[a] 3.5[a] 1.5[c]
α(BEDT-TTF)$_2$NH$_4$Hg(SCN)$_4$	2.1/2.5[a] 1.17[b]
α-(BEDT-TTF)$_2$KHg(SCN)$_4$	1.4[a] 2.4/2.5[a] 0.40[b] 0.94[b]
α-(BEDT-TTF)$_2$RbHg(SCN)$_4$	1.5[a]
λ-(BEDT-TSF)$_2$GaCl$_4$	3.6[a] 6.3[a]
κ-(BEDT-TSF)$_2$GaCl$_4$	1.2[a] 2.4[a]
κ-(BEDT-TSF)$_2$FeBr$_4$	4.7[a] 8.0[a] (2.0[a] unassigned)
κ-(BEDT-TSF)$_2$FeCl$_4$	2.8[a] 5.3[a]
K$_3$C$_{60}$	2.4[c]

[a] Data were derived from the de Haas effect.
[b] Data were derived from the cyclotron resonance.
[c] Data were derived from the optical reflection spectra.

Ground States in Normal Conductors

Since the low-dimensional metallic state is intrinsically unstable, organic metals show a variety of specific types of order at low temperatures as their ground states such as the Peierls insulator, the spin-density wave (SDW) state, the spin-Peierls state, the Mott insulating state, and the Wigner crystalline state.

Peierls Insulator

The 1D metal responds in a divergent way to the external perturbation with $2k_F$ periodicity (k_F: Fermi wave number). This instability of the electron system is removed by a $2k_F$ periodic modulation of the electron density as predicted by R E Peierls. **Figure 3** shows the interaction of the $2k_F$ phonons with the electrons near the Fermi level. In this case, through the electron–phonon interaction, the lattice is modulated

with the same wave number and the mixed wave called a charge-density wave (CDW) is formed. The resultant electronic state has a gap at the Fermi level for single electron excitations and hence becomes insulating. However, the collective CDW excitation can slide through the crystal if there are no obstacles to pin it. The corresponding Fröhlich mode was first introduced to explain the superconductivity mediated by an electron–phonon interaction. In reality, however, the free-sliding mode of the CDW is restricted by defects and by interactions between CDWs on different chains. As a result, the CDWs are pinned to form the static lattice deformation and thereby the insulating state is called the Peierls insulator.

Experimentally, the transition to the Peierls insulating phase is detected as a change in the resistivity from a metallic decrease to a nonmetallic increase on cooling and a reduction of the paramagnetic spins by free carriers (**Table 3**). The accompanying lattice modulation with a period of π/k_F is found as the superlattice, which can be detected by X-ray and neutron scatterings. The typical transition behavior was seen in TTF·TCNQ. It is noteworthy that the bond alternation in polyacetylene polymer has been ascribed to the Peierls interaction.

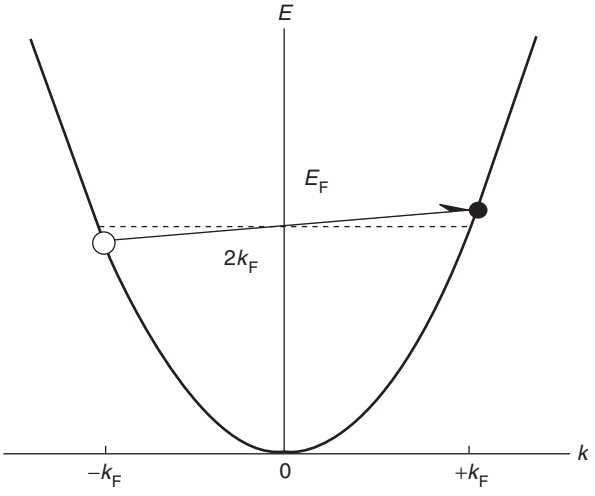


Figure 3 Interaction of $2k_F$ perturbation exciting the electron near the Fermi level E_F in a 1D metal.

For 1D metals, the Fermi surface is represented by the parallel planar sheets separated by $2k_F$: the Fermi surfaces can be superposed to the adjacent ones by a $2k_F$ shift in the momentum space. In general, when the Fermi surface can be superposed partially or entirely to the other Fermi surface by a parallel shift, it is said that the Fermi surfaces are nested. Whenever the Fermi surface has a nesting property, the electronic state becomes unstable against the density wave formation. When the electron transfer in the direction perpendicular to the 1D axis is not negligible, the planar sheets of the 1D Fermi surface can be warped to lose the nesting property, resulting in the suppression of the Peierls insulating state.

SDW State

For a metal with Fermi surfaces satisfying the nesting condition, provided that the on-site Coulomb repulsion is predominant, the SDW phase is formed by suppressing the Peierls instability. In this case, the charges distribute uniformly on the sites, while the electron spins order antiferromagnetically with a $2k_F$ periodicity. No lattice modulation accompanies the SDW, in contrast to the CDW.

The SDW transition is found in $(TMTSF)_2PF_6$ at 12 K at ambient pressure as a transition to the insulating state. The spin susceptibility decreases on the transition for a magnetic field applied parallel to the easy axis of spins, whereas it is almost unchanged for a field perpendicular to the easy axis (**Table 3**). When the field along the easy axis is increased, the metamagnetic transition with spin flip is observed. The antiferromagnetic (AF) ordering is detected by the AF resonance and also as a broadening in the NMR line shape. The AF phase in $(TMTSF)_2PF_6$ is suppressed by an application of pressure due to an increase in the transfer energy in the transverse direction of the columns. It enhances the Fermi surface warping (**Figure 2**), resulting in the depression of the nesting condition. On the suppression of the AF phase, the superconductivity phase appears in this salt. On the other hand, when the magnetic field is applied perpendicular to the most conducting plane, the transfer in the direction perpendicular to the

Table 3 Comparison of nonmetallic ground states

	Peiers state	*SDW state*	*Spin–Peierls state*	*Mott state*	*Wigner crystal*
Dimensionality	one	one	One	two/one	two/one
Electron-lattice interaction	yes		Yes		
Electron–electron		yes	Yes	yes	yes
Interaction		(on-site)	(spin)	(on-site)	(long range)
Magnetism	NM	AF	NM	AF	PM
Lattice modulation	yes	no	Yes	no	no

NM: nonmagnetic, AF: antiferromagnetic, PM: paramagnetic.

most conducting plane is suppressed because of the confinement effect for the open orbit electrons toward the one-dimensionality, resulting in the enhancement of the SDW state and the field-induced SDW transition.

Spin-Peierls State

For the 1D metal with a half-filled band, when the adjacent spins form singlet pairs by an exchange interaction, the nonmagnetic ground state with a gap for single electron excitation, called the spin-Peierls state, is formed. In this case, the $2k_F$ lattice modulation appears to be associated with the dimerization of the spin structure. $(TMTTF)_2PF_6$ is known to show the spin-Peierls state. It has been demonstrated that, with an increase of the intercolumn interaction by pressure, $(TMTTF)_2PF_6$ undergoes a series of transitions from the spin-Peierls state, the SDW state, the superconducting state, and the normal metallic state.

Mott Insulating state

For several 2D half-filled electronic systems supposed to be metals, the electron localization takes place due to a strong on-site Coulomb interaction. When the energy of the on-site Coulomb interaction becomes greater than the transfer energy between the molecules, a transition to the so-called Mott insulating state takes place. The spins are ordered antiferromagnetically in the nonmetallic state. A typical example is found in κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Cl$. In this case, with an increase in the bandwidth by pressure, the metallic state is restored.

Wigner Crystal State

When the long-range Coulomb interaction works, the electrons are localized by mutual repulsion to form a Wigner lattice, even in the quarter-filled state of the 1D metal. As an example of the Wigner crystal, the formation of a $4k_F$ modulation in a 1D metal $(DI\text{-}DCNQI)_2Ag$ has been pointed out. In this case the electrons occupy half of the sites, resulting in a charge disproportionation to the sites. A similar ordering is found in the 2D electron system on the surface of superfluid He and at the semiconductor interface.

Organic Superconductors

In 1963, a model of an organic polymer superconductor possessing a high-transition temperature was proposed by W A Little. It had generated a strong interest in the development of organic superconductors. However, the superconducting organic polymer has not been synthesized to date. In 1979, superconductivity in organic materials was discovered in the Q1D charge transfer salt $(TMTSF)_2PF_6$ under pressure by the group of D Jérome and K Bechgaard. Subsequently, in 1984, the β-$(BEDT\text{-}TTF)_2I_3$ which possessed a layered structure, was found by the group of I F Schegolev and E B Yagubskii to be superconducting. This structure belongs to a second large family of organic superconductors including those with T_c exceeding 10 K. In the following, $(BEDT\text{-}TSF)_2X$ salts involving stages of interplay with magnetism, as well as the other types of exotic organic superconductors and the fullerene superconductors are presented. All known organic superconductors belong to the type-II superconductors.

TMTSF Superconductors

The charge transfer salt $(TMTSF)_2PF_6$ which showed an SDW state at ambient pressure displayed superconductivity at 1.2 K under pressure by suppressing the SDW. The pressure–temperature phase diagram has demonstrated that the SDW phase exists in proximity to the superconductivity phase, suggesting that the mechanism of the superconductivity is exotic and the pairing through a spin-fluctuation coupling is probable. The salt $(TMTSF)_2ClO_4$, showing superconductivity at ambient pressure, was experimentally useful to characterize the anisotropic nature. It showed a notable anisotropy in the critical magnetic field. The evaluated superconductivity coherence lengths are listed in **Table 4**. Although the anisotropy is obvious, even the shortest coherence length is longer than the intermolecular distance in each direction in this salt.

It is noteworthy that the size of the counter anion plays an essential role in determining the T_c of organic

Table 4 Superconductivity parameters of typical organic superconductors

Materials	T_c (K)	ξ_1 (nm)	ξ_2 (nm)	ξ_3 (nm)
$(TMTSF)_2ClO_4$	1.25	71	34	2.0
$(DMET\text{-}TSF)_2AuI_2$	1.0	100	40	2.0
β_L-$(BEDT\text{-}TTF)_2I_3$	1.4	63	61	2.9
β_H-$(BEDT\text{-}TTF)_2I_3$ under 1.6 kbar	7.2	13	[a]	1.0
β-$(BEDT\text{-}TTF)_2IBr_2$	2.3	46	44	1.9
β-$(BEDT\text{-}TTF)_2AuI_2$	4.2	25	[a]	1.9
κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$	8.7	2.9	[a]	0.31
κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Br$	10.9	2.3	[a]	0.58
κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Cl$ under 1.0 kbar	10.5	1.3	[a]	0.95
β-$(BDA\text{-}TTP)_2SbF_6$	6.3	10.5	[a]	2.6
α-$(EDT\text{-}TTF)[Ni(dmit)_2]$	1.3	31	[a]	2.4
λ-$(BEDT\text{-}TSF)_2GaCl_4$	6.4	12.5	[a]	1.6
K_3C_{60}	19.3	3.4	[b]	[b]
Rb_3C_{60}	30	3.0	[b]	[b]

T_c: transition temperature.
ξ_1, ξ_2, and ξ_3: coherence length ($\xi_1 \geqq \xi_2 \geqq \xi_3$).
[a] and [b] are derived under the assumption of $\xi_1 = \xi_2$ and $\xi_1 = \xi_2 = \xi_3$, respectively.

superconductors by working as a spacer between TMTSF molecules. The role is referred to as the chemical pressure effect, in contrast to the ordinary (physical) pressure effect. The superconductivity in organic materials shows a number of unconventional features. It is very sensitive to crystalline disorder. The alloying of TMTSF and the anion represented in the form of $(TMTSF)_{2-x}(TMTTF)_x(ClO_4)_{1-y}(ReO_4)_y$ can destroy the superconductivity with the composition of a few percent of x or y. A similar effect is caused by the radiation damage at the level of 100 ppm. The isotope effect studied by the replacement of H-atoms in TMTSF with deuterons gave a reduction in T_c too large for the pairing by the phonon mechanism. The proton NMR relaxation experiment showed the absence of the Hebel–Schlichter peak expected for isotropic pairing. All these unusual results motivated the detailed study of the mechanism of superconductivity, with special interest to a possibility of triplet pairing. In $(TMTSF)_2PF_6$, the upper critical field obtained in the applied magnetic field precisely oriented in certain crystalline directions, far exceeds the Pauli paramagnetic limit, suggesting that the superconducting pairs are not in the spin-singlet state. The NMR Knight-shift experiment on ^{77}Se asserted the lack of spin susceptibility variation on the superconducting transition, in accordance with the triplet pairing.

BEDT-TTF Superconductors

Salts of the general formula $(BEDT\text{-}TTF)_2X$ present a variety of superconductors due to the freedom in the stacking arrangement of BEDT-TTF molecules. The salt β-$(BEDT\text{-}TTF)_2I_3$ undergoes a superconductivity transition at 1.5 K at ambient pressure, while another phase, formed through the process which lets a hydrostatic pressure environment pass through, exhibits T_c of 8 K. Due to the values of their T_c, the two phases are called β_H- and β_L-type, respectively. The difference between the phases is ascribed to the presence of the incommensurate modulation in the position of the ethylene C_2H_4 fragments at the end of the BEDT-TTF molecule, accompanied by a weak potential disorder, for the β_L-type. The superconductivity is suppressed for an alloy, such as $(BEDT\text{-}TTF)_2(I_3)_{1-x}(IBr_2)_x$, with x exceeding 20%.

The κ-type salts represented by κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$, κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Br$, and κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Cl$ show typical 2D superconducting properties, where the superconducting coherence length in the direction perpendicular to the layer (**Table 4**) is short compared to the interlayer spacing (1.6, 1.5, and 1.5 nm, respectively). In these cases, the superconductivity transition is influenced by a thermal fluctuation, especially strongly under the magnetic field applied perpendicular to the superconducting plane, so that the upper critical field expected in the mean field theory cannot be determined. These 2D superconductors bear close similarity to the layered CuO_2 superconductors with respect to low carrier density, strong electron correlation, short coherence length, and proximity to the AF phase. In contrast to the CuO_2 superconductors, however, they undergo superconducting transition without doping and are characterized by weaker electron correlations. Experimental studies on the magnetic penetration depth, NMR, μSR, specific heat, thermal conductivity, electron tunneling spectroscopy, and magnetization suggest superconductivity in presence of the gap nodes. However, recently, some experiments have suggested superconductivity in the absence of the gap nodes.

BEDT-TSF Superconductors

BEDT-TSF molecules can form superconducting layers with anion layers containing the localized magnetic ion Fe^{3+} in a high 5/2-spins state as well as with nonmagnetic ions such as $GaCl_4^-$, $GaBr_4^-$, and $InCl_4^-$. The salt λ-$(BEDT\text{-}TSF)_2GaCl_4$ undergoes a superconducting transition at 6 K, while λ-$(BEDT\text{-}TSF)_2FeCl_4$ shows a transition to the AF insulating state at 8 K, after showing a similar temperature dependence of the resistivity down to the transition temperature. With the external field applied parallel to the conducting plane, λ-$(BEDT\text{-}TSF)_2FeCl_4$ is changed to be a ferromagnetic metal, and above 17 T it reveals superconductivity in the field interval up to 44 T. The reentrant superconductivity is explained in terms of the Jaccarino–Peter mechanism, in that the exchange field supplied from the magnetic ions to π-electrons is compensated by the external magnetic field to show superconductivity. In the high magnetic field applied precisely parallel to the superconducting plane, λ-$(BEDT\text{-}TSF)_2GaCl_4$ was found to show the spatially modulated order parameter phase predicted by Fulde–Ferrell and Larkin-Ovchinnikov. A similar situation was reported for κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$ and κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Cl$.

Alloys such as λ-$(BEDT\text{-}TSF)_2Ga_{1-x}Fe_xBr_yCl_{4-y}$ show superconductivity in a wide range of variations in x and y. This contrasts with the cases for TMTSF and BEDT-TTF superconductors which show sensitivity to disorder. The alloying varies the concentration of the magnetic ions with x and the anion volume with y. The latter, known as the chemical pressure effect, works consistently with the pressure effect.

For κ-$(BEDT\text{-}TSF)_2FeBr_4$, the AF state appearing at 2.6 K was found to remain in the superconducting state appearing at 1.1 K, resulting in the coexistence of the two phases.

Other Types of Organic Superconductors

The effort at synthesis brought about a notable variety of molecules forming superconducting salts. Unsymmetrical donors such as DMET, which is the hybrid molecule of TMTSF and BEDT-TTF, provide superconducting salts such as $(DMET)_2AuI_2$. Here, the unsymmetrical molecules in the columns are stacked alternately in direction so that the resultant assembly has a twofold symmetry. Apart from salts with a TTF-type skeleton, superconductors were found with a TTP-type skeleton such as BDA-TTP and DTEDT, among salts such as $(BDA\text{-}TTP)_2SbF_6$ and $(DTEDT)_3AuI_2$. Superconductivity was also found in materials based on the acceptor molecule $M(dmit)_2$ ($M = Ni$, Pd), and in salts such as $(TTF)[Ni(dmit)_2]_2$ and $N(CH_3)_4[Pd(dmit)_2]_2$.

A_3C_{60} Superconductors

By doping alkali metals such as K and Rb to C_{60} crystals, the superconductors K_3C_{60} and Rb_3C_{60} with T_c of 19.3 and 30 K, respectively, were synthesized. The maximum T_c reached for $RbCs_2C_{60}$ was 33 K. In these cases, the superconductivity is isotropic. The transition temperature increases with an increase in the intermolecular distance of the C_{60} molecules, which was demonstrated by the chemical pressure effect as well as the physical pressure effects by the replacement of alkali metals. A combination with ions of different valences such as alkali-earth atoms, has brought about superconductors such as Ca_5C_{60}.

See also: Molecular Crystallography; Organic Semiconductors; Polymers and Organic Materials, Electronic States of; Spin Density Waves and Magnons; Superconductivity: General Aspects.

PACS: 71.18.+y; 71.20.−b; 71.30.+h; 71.45.Lr; 72.80.Le; 75.30.Fv; 75.47.−m

Further Reading

Bourbonnais C and Jérome D (1999) *The normal phase of quasi-one-dimensional organic superconductors*. In: Bernier P, Lefrant S, and Beiden G (Eds.) *Advances in Synthetic Metals, Twenty Years of Progress in Science and Technology,* pp. 206-261. New York: Elsevier.

Devreese JT, Evrard RP, and van Doren VE (eds.) (1979) *Highly Conducting One-Dimensional Solids*. New York: Plenum.

Dresselhaus MS, Dresselhaus G, and Eklund PC (1996) *Science of Fullerenes and Carbon Nanotubes*. New York: Academic Press.

Ishiguro T, Yamaji K, and Saito G (1998) *Organic Superconductors*, 2nd edn. Berlin: Springer.

Jérome D and Caron LG (eds.) (1987) *Low-Dimensional Conductors and Superconductors*. NATO Advanced Study Institute Ser. B vol 155. New York: Plenum.

Kagoshima S, Nagasawa H, and Sambongi T (1988) *One-Dimensional Conductors*. Berlin: Springer.

Kagoshima S, Kato R, Fukuyama H, Seo H, and Kino H (1999) *Interplay of structural and electronic properties*. In: Bernier P, Lefrant S, and Beiden G (Eds.) *Advances in Synthetic Metals, Twenty Years of Progress in Science and Technology,* pp. 262-316. New York: Elsevier.

Nalwa HS (ed.) (1997) *Handbook of Organic Conductive Molecules and Polymers*. Chichester: Wiley.

Narlikar A (ed.) (2000) Organic superconductors and more on striped phase in cuprates. In: *Studies of High Temperature Superconductors*, vol. 34. New York: Nova Science Publishers.

Saito G, Kagoshima S, Kanoda K, Aoki H, Mori T, *et al.* (2003) *Proceedings of The Fourth International Symposium on Crystalline Organic Metals, Superconductors and Ferromagnets*, Synthetic Metals vol 133–134. Amsterdam: Elsevier.

Singleton J (2000) Studies of quasi-two-dimensional organic conductors based on BEDT-TTF using high magnetic fields. *Reports on Progress in Physics* 63: 1111–1208.

Singleton J (2001) *Band Theory and Electronic Properties of Solids*. London: Oxford University Press.

Williams JM, Ferraro JR, Thorn RJ, Carlson KD, Geiser U, *et al.* (1992) *Organic Superconductors (Including Fullerenes)*. New Jersey: Prentice Hall.

Wosnitza J (1996) *Fermi-Surfaces of Low-Dimensional Metals and Superconductors*. Berlin: Springer.

Nomenclature

A	alkali metal
e	the charge of the electron
k_F	Fermi wave number
T_c	transition temperature
X	counter molecule
ξ	coherence length

Cyclotron Resonance: Metals

J Singleton, Los Alamos National Laboratory, Los Alamos, NM, USA

Introduction

Landau quantization may occur when electrons in a solid are subjected to a magnetic field. The excitation of electrons from one Landau level to another by photons results in a resonant contribution to the magnetic-field-dependent, high-frequency conductivity. It is this phenomenon that is known as "cyclotron resonance." This article is chiefly concerned with the electronic "scattering rate" τ_{CR}^{-1} that can be deduced from a cyclotron resonance experiment; this is associated with the finite energy width of the Landau

levels caused by their limited lifetime. By convention, the scattering rate is defined as the reciprocal of a notional scattering time τ_{CR}, which is sometimes also referred to as the "Landau-level lifetime." It should be noted that a simple-minded first analysis of a cyclotron resonance experiment often yields an "apparent" scattering rate, which may bear little relationship to the actual scattering processes present; it can contain contributions due to experimental factors that mask the true result. Apparent cyclotron scattering rates are distinguished throughout by using the subscript "M," which stands for "measured," hence the notation τ_{CRM}^{-1}.

Another concern is the relationship of τ_{CR}^{-1} to the "apparent" scattering rates deduced from other techniques such as de Haas–van Alphen (or Shubnikov–de Haas) oscillations (τ_{dHvA}^{-1}) and measurements of the electronic conductivity $\sigma(\tau_{\sigma}^{-1})$. It should be noted that τ_{CR}^{-1} is preferable to τ_{dHvA}^{-1} as a guide to the true energy width of the Landau levels due to scattering. It is also found that there can be considerable differences between τ_{CR} and τ_{σ}.

In order to keep the discussion as general as possible, electrons in a solid are generally referred to as "quasiparticles". This enables the descriptions that follow to encompass not only both carrier types (electrons, holes) but also systems in which effects such as electron–electron and electron–phonon interactions are significant and describable by the Landau Fermi-liquid theory.

This article deals with cyclotron resonance in metals; the most fundamental definition of a metal is "a solid with a Fermi surface." From the cyclotron resonance viewpoint, this broad class of materials must be separated into three groups:

1. metals with a high electron density, including elements such as Cu, Au;
2. intermediate electron-density systems, often of reduced dimensionality, such as crystalline organic metals and certain oxides (e.g., Sr_2RuO_4); and
3. low electron-density systems, such as degenerate semiconductors, and semiconductor heterojunctions.

The distinction between these three groups results from the ease with which light of the appropriate frequency to excite cyclotron resonance (usually in the GHz to THz range) can penetrate the sample. Hence, the rest of the article is arranged as follows. The second section presents a brief introduction to Landau quantization. The third section deals with experimental techniques for elemental metals and the extraction of true scattering rates from the apparent rate produced by fitting experimental data. Intermediate electron-density systems and degenerate semiconductors are dealt with in the fourth section; the experimental techniques and the extraction of scattering rates from data are treated in separate subsections. However, the causes of the difference between τ_{CR} and τ_{dHvA} are similar in both types of "metal," and so are treated in a common subsection. Systems containing two or more distinct types of quasiparticles are described in the fifth section; often the interactions lead to a single cyclotron resonance rather than two distinct resonances. Finally, the sixth section covers the distinctions between the effective masses measured in different experiments, and the seventh section mentions two effects closely related to cyclotron resonance.

Landau Quantization: A Brief Description

In order to gain a qualitative understanding of the principles of the various measurement techniques for cyclotron resonance, it is useful to consider the semiclassical equation of motion of a quasiparticle of charge q, energy E, and wave vector quantum number $\boldsymbol{k}$, subjected to the Lorentz force:

$$\hbar\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = q\boldsymbol{v}\times\boldsymbol{B} \qquad [1]$$

Here, $\boldsymbol{B}$ is the magnetic induction and $\boldsymbol{v}$ is the quasiparticle's velocity, defined as

$$\boldsymbol{v} = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}E(\boldsymbol{k}) \qquad [2]$$

where $\nabla_{\boldsymbol{k}}$ is the gradient operator in k-space. Equations [1] and [2] imply that (1) the component of $\boldsymbol{k}$ parallel to $\boldsymbol{B}$ is constant; and (2) $\mathrm{d}\boldsymbol{k}/\mathrm{d}t$ is perpendicular to $\nabla_{\boldsymbol{k}}E(\boldsymbol{k})$. This means that the quasiparticle path is one of constant energy.

The Fermi surface is of course a surface of constant energy. Under the action of a magnetic field, quasiparticles on the Fermi surface remain on that surface while moving on paths which lie in planes perpendicular to the magnetic field (see **Figure 1**). The band

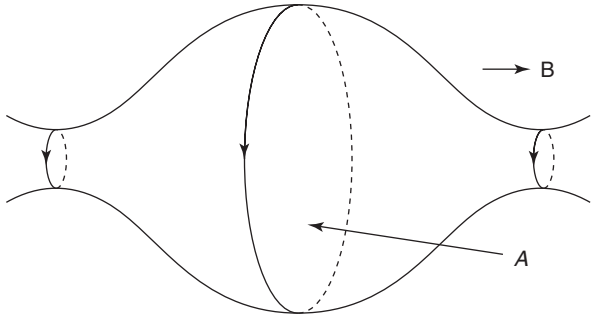


Figure 1 Orbits on a Fermi surface section are in planes perpendicular to $\boldsymbol{B}$. Here, A represents the k-space cross-sectional area of the orbit. The two orbits shown are "extremal" orbits.

structure of solids is such that in many cases, a constant energy path of this kind forms a closed orbit (or "cyclotron orbit") of cross-sectional area (in k-space) A (**Figure 1**). Standard texts (such as those by Ashcroft and Mermin or Singleton) show that the angular frequency ω_c associated with this periodic orbit is

$$\omega_c = \frac{qB}{m^*_{CR}} \quad [3]$$

where

$$m^*_{CR} = \frac{\hbar^2}{2\pi}\frac{\partial A(E, k_\parallel)}{\partial E} \quad [4]$$

The quantity defined in eqn [4] is known as the cyclotron mass or cyclotron effective mass.

From a quantum-mechanical viewpoint, the cyclotron orbital motion corresponds to the quantization of the quasiparticle's motion in the plane perpendicular to $\boldsymbol{B}$, resulting in an energy spectrum

$$E(\boldsymbol{B}, k_\parallel) = (l + \tfrac{1}{2})\hbar\omega_c + E(k_\parallel) \quad [5]$$

Here, l is an integer $(l = 0, 1, 2, \ldots)$ known as the Landau quantum number, and $E(k_\parallel)$ is the energy associated with the unaffected wave vector $k_\parallel$ in the direction parallel to $\boldsymbol{B}$. The quantization of the quasiparticle's energy in the plane perpendicular to $\boldsymbol{B}$ is known as "Landau quantization," and the energy levels defined by l are called "Landau levels."

Cyclotron resonance occurs when a photon promotes a quasiparticle from a full state in one Landau level to an empty state in another Landau level. Owing to the fact that photons of the appropriate energy carry very little momentum, such transitions conserve $k_\parallel$. Hence, cyclotron resonance occurs when the photon energy ($h\nu$ or $\hbar\omega$) corresponds to multiples of $\hbar\omega_c$.

Cyclotron Resonance in Elemental Metals

Experimental Considerations

The field components of electromagnetic radiation decay with distance z into a conducting material as $\exp(-z/\delta)$, where $\delta = \left(\frac{1}{2}\sigma\omega\mu_r\mu_0\right)^{-1/2}$ is the skin depth or anomalous skin depth; here σ is the conductivity of the material, $\mu_r\mu_0$ is its permeability, and ω is the angular frequency of the radiation. Metals have rather

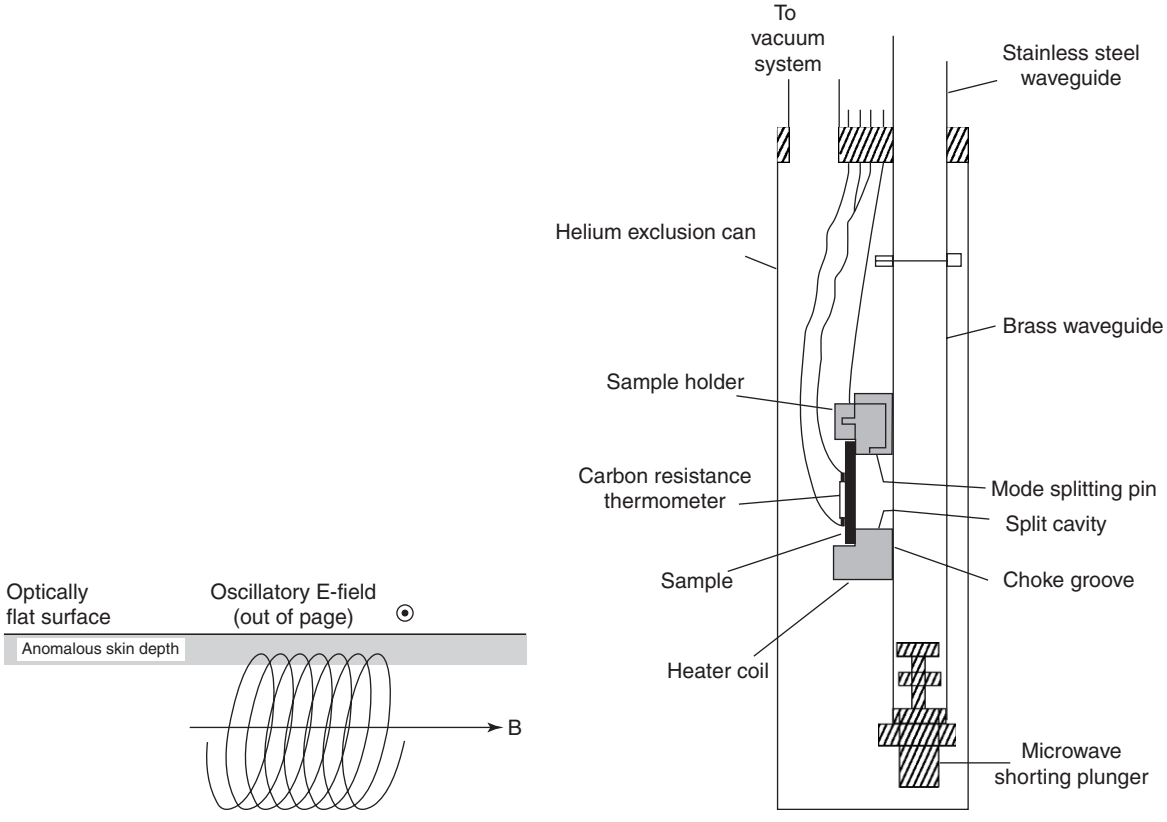


Figure 2 Left: geometry of a cyclotron resonance experiment in a metal. The shaded region indicates the skin depth for the radiation; the helical orbits of the electrons are shown schematically (from Singleton J, *Band theory and electronic properties of solids* (Oxford University Press, 2001), chapter 8). Right: a practical implementation of the Azbel'–Kaner technique. The sample acts as one wall of a microwave resonant cavity; the whole apparatus is immersed in a liquid helium dewar which sits inside an iron-yoked electromagnet. (Modified from Häussler P and Welles SJ (1966) *Physical Review B* 152: 675.)

high conductivities; moreover, typical effective masses in metals combined with the magnetic fields readily available in laboratories have meant that frequencies $\omega/2\pi \sim 1 - 100$ GHz have tended to be applied in cyclotron resonance experiments. This combination of factors results in a small skin depth; hence, radiation cannot penetrate far into the crystal. These considerations dictate the geometry of the cyclotron resonance measurement (see **Figure 2**); such a configuration is often referred to as the Azbel'–Kaner geometry.

The magnetic field is applied parallel to the surface of the crystal, which is placed in the region of an oscillating electric field in a resonant cavity; the E-field $\boldsymbol{E}$ of the radiation is arranged to be perpendicular to the quasistatic magnetic field $\boldsymbol{B}$ and parallel to the surface. Examination of eqns [1] and [2] shows that the projection of an electron's real-space orbit in a plane perpendicular to $\boldsymbol{B}$ is the same shape as its k-space orbit rotated by $\pi/2$ radians. The paths of electrons on closed Fermi surface orbits are therefore helices with axes parallel to $\boldsymbol{B}$ in real space.

If the angular frequency of the radiation $\omega = j\omega_c$, where j is an integer, then the electrons on a helical path corresponding to the cyclotron frequency ω_c will receive a "kick" from the radiation's electric field every time they come within a skin depth of the surface; this results in absorption of energy. Usually, ω is kept constant and the field is swept, so that absorptions, seen as resonances in the surface impedance, are uniformly spaced in $1/B$ (see **Figure 3**).

Extraction of Scattering Rates

Lindholm and Le Page and others have demonstrated that a magnetic field which is very accurately parallel to a very flat sample surface is required for a successful measurement; if this is not the case, the helical electron paths drift in and out of the skin depth, causing a smearing of the features in the surface impedance. If the surface is flat, and the magnetic field is accurately parallel to it, then the surface impedance Z becomes

$$Z \propto \frac{e^{i\pi/3}}{s} \qquad [6]$$

with

$$s^3 \approx \sum_j p_j R(\lambda_j) + R_{NR} \qquad [7]$$

to reasonable accuracy. Here, the index j labels the various Fermi surface orbits involved; in practice, it is found that only "extremal orbits" make a significant resonant contribution. Extermal orbits are defined by $dm^*_{CR}/dk_{\parallel} = 0$; frequently, these are just the maximum- and minimum-area closed orbits for a particular direction of the magnetic field (see **Figure 1**.) In such regions, a large number of electrons have almost identical masses (and thus almost identical ω_c), providing a large contribution to Z. The term R_{NR} encompasses nonresonant processes, whilst the resonant contributions are described by the function λ_j:

$$\lambda_j = 2\pi\frac{\omega}{\omega_{cj}}\left(\frac{1}{\omega\tau_{j\mathrm{CRM}}} + i\right) \qquad [8]$$

Note that a cyclotron frequency ω_{cj} and an apparent scattering time $\tau_{j\mathrm{CRM}}$ have been identified for the jth orbit. As long as the electronic dispersion relationship is nearly parabolic close to the Fermi surface, $R(\lambda_j)$ takes the simple form:

$$R(\lambda_j) = \frac{1}{1 - e^{-\lambda_j}} \qquad [9]$$

In practice, the sample orientation is chosen so that only one or two extremal orbits are possible, simplifying the features observed in Z; an example is shown in **Figure 3**.

The above expressions for Z have been used to fit experimental data under a variety of valid approximations; typical examples are given by Lindholm and Le Page, and Häussler and Wells and references therein. It is found that the apparent scattering time $\tau_{j\mathrm{CRM}}$ encompasses several contributions:

$$\frac{1}{\tau_{j\mathrm{CRM}}} = C_{msj}\omega + C_{srj}\omega^{1/3} + \frac{1}{\tau_{impj}} + \frac{1}{\tau_{phononj}} + \frac{1}{\tau_{e-ej}} \qquad [10]$$

The first two terms on the right-hand side of eqn [10] are due to the so-called "mass spread" ($C_{msj}\omega$) and the effects of surface roughness ($C_{srj}\omega^{1/3}$). The former term has the following origin; as has been mentioned above, the resonance spectrum is dominated by extremal orbits about the Fermi surface, defined by $dm^*_{CR}/dk_{\parallel} = 0$. In such regions, a large number of electrons have almost identical masses (and thus almost identical ω_c), providing a large contribution to Z. The fact that higher derivatives of m^*_{CR} do not vanish in real Fermi surfaces causes "mass-spread"; a distribution of cyclotron frequencies leads to a smearing of the resonant features which at first sight looks as though it is due to scattering. This effect is represented by the term $C_{msj}\omega$ in eqn [10].

Because both the mass-spread and surface roughness terms consist of a constant times ω raised to a small power, they may be distinguished from the other (frequency-independent) contributions by measuring the surface impedance at more than one frequency.

The last three (frequency-independent) terms in eqn [10] are caused by the actual scattering mechanisms; their sum is τ^{-1}_{CRj}. A weighted average of all of

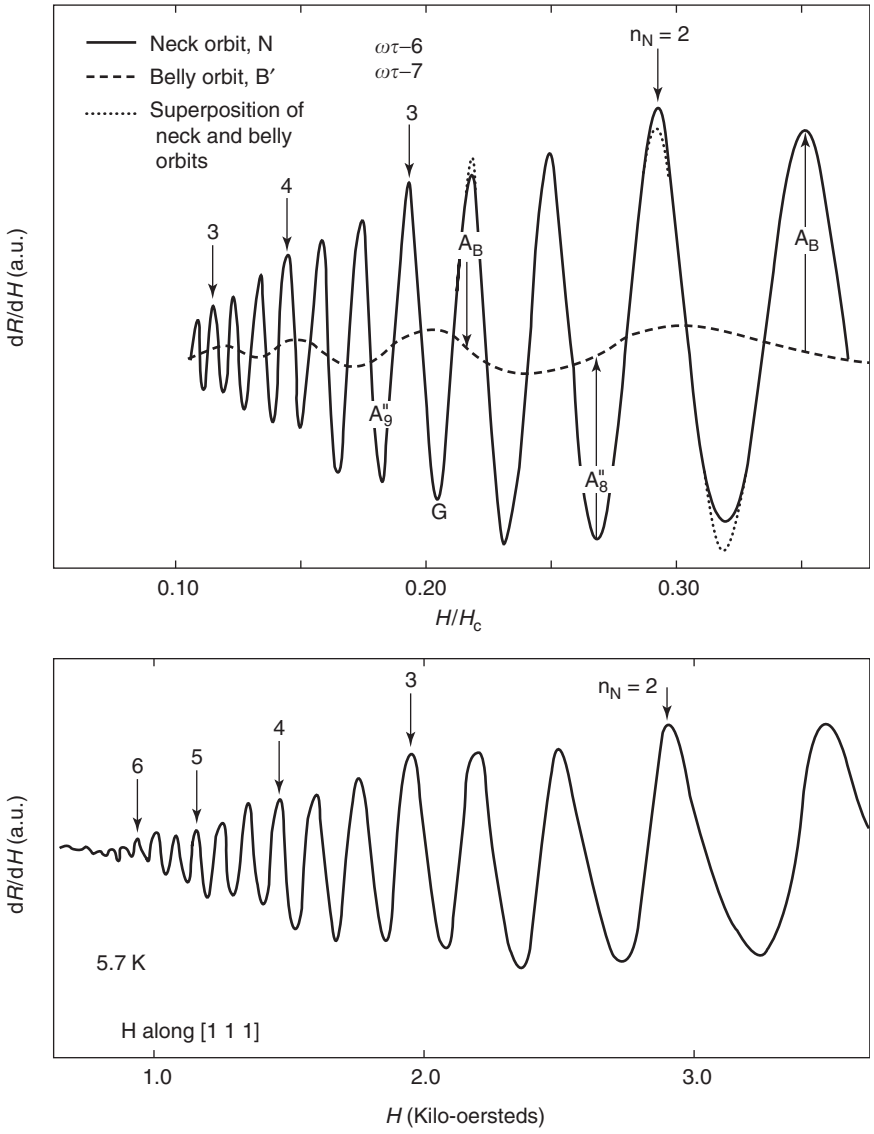


Figure 3 Combined cyclotron resonance signals from extremal belly orbits B' and neck orbits N in copper; the field is in the [1 1 1] direction. The upper curve is a theoretical simulation that shows the two contributions separately (dashed and dotted curves) and combined (solid curve). The lower part of the figure shows experimental data. The fortuitous factor of 3 between the cyclotron periods of the two orbits facilitates separation of the two contributions. (After Hässler P and Welles SJ (1966) *Physical Review B* 152: 675.)

the electron-impurity scattering processes leads to the term $\tau_{\text{imp}j}$, whilst $\tau_{\text{phonon}j}$ and $\tau_{\text{e}-\text{e}j}$ are due to electron–phonon scattering and electron–electron scattering events, respectively. The three terms may be distinguished by observing their temperature dependences:

$$\frac{1}{\tau_{\text{imp}j}} = \text{constant}, \qquad \frac{1}{\tau_{\text{phonon}j}} \propto T^3 \quad \text{and} \quad \frac{1}{\tau_{\text{e}-\text{e}_j}} \propto T^2 \qquad [11]$$

The origins of the temperature power laws for these mechanisms are dealt with in standard texts such as Ashcroft, Mermin, and Singleton.

Layered Metals and Degenerate Semiconductor Systems

Layered Metals: Experimental Considerations

Fermi surfaces and resonant cavity techniques There has been great recent interest in layered materials, as there are many correlated-electron systems which have very anisotropic electronic band structure. Examples include crystalline organic metals and superconductors, the "high-T_c" cuprates, and layered ruthenates and manganites. In the context of the current discussion, such systems have two distinctive features.

1. Typically, their quasiparticle densities are rather lower than those in elemental metals; hence, their

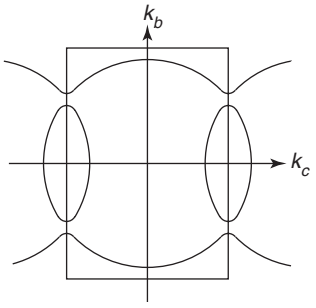


Figure 4 Brillouin zone and Fermi surface of κ-(BEDT-TTF)$_2$Cu(NCS)$_2$, showing the open, quasi-one-dimensional sections, and the closed, quasi-two-dimensional pocket. (Modified from Singleton J (2000) *Reports of Progress in Physics* 63: 1111.)

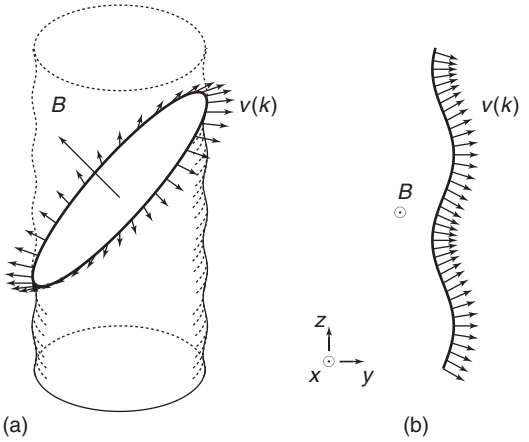


Figure 5 (a) Schematic side view of the warping (corrugation) of a closed section of a Fermi surface caused by finite interplane transfer integrals. The arrows indicate the velocities of a quasiparticle following a closed orbit about the Fermi surface in a plane perpendicular to the magnetic field ***B***. (b) An open section of the Fermi surface. In an in-plane magnetic field, quasiparticles are driven across the Fermi surface, so that their velocities (shown by arrows) rock from side to side. (After Singleton J (2000) *Reports of Progress in Physics* 63: 1111.)

skin depths tend to be larger, and the effects of screening are reduced.

2. These systems may often be described by a tight-binding Hamiltonian in which the ratio of the interlayer transfer integral $t_\perp$ to the average intralayer transfer integral $t_\parallel$ is $\ll 1$. This results in very anisotropic electronic properties.

Figure 4 shows the cross section of the Fermi surface of a typical example, the organic superconductor κ-(BEDT-TTF)$_2$Cu(NCS)$_2$, parallel to the highly conducting planes. The Fermi surface consists of both open and closed sections; it is customary to label such sections "quasi-one-dimensional" and "quasi-two-dimensional," respectively. The names arise because the Fermi surface is a surface of constant energy, and so eqn [2] shows that the velocities of quasiparticles at the Fermi surface will be directed perpendicular to it. Therefore, referring to **Figure 4**, quasiparticles on the closed Fermi surface pocket can possess velocities which point in any direction in the (k_b, k_c) plane; they have freedom of movement in two dimensions and are said to be quasi-two-dimensional. By contrast, quasiparticles on the open sections have velocities predominantly directed parallel to k_b and are quasi-one-dimensional.

Thus far, only a cross section is considered through the Fermi surface parallel to the highly conducting planes (**Figure 4**). In the interlayer direction (also known as the interplane direction), the dispersion is very small, due to the small interlayer transfer integrals. For the purpose of this discussion, the quasiparticle dispersion is written as

$$E(\boldsymbol{k}) = E(k_x, k_y) - 2t_\perp \cos(k_z d) \qquad [12]$$

where k_x (and k_y) and k_z are the components of the wave vector parallel (k_x, k_y) and perpendicular (k_z) to the conducting planes, and $t_\perp$ is again the interlayer transfer integral. The presence of dispersion in the interlayer direction results in a slight warping of the Fermi surface, shown schematically for a closed pocket in **Figure 5a**.

As before, when a magnetic field is introduced, eqns [1] and [2] indicate that the k-space path of the quasiparticle is defined by the intersections of surfaces of constant energy with planes perpendicular to ***B***. Therefore, for almost all directions of the magnetic field, quasiparticles on the closed sections of the Fermi surface will be able to complete the closed k-space orbits in the plane perpendicular to ***B*** (see **Figure 5a**); in contrast, the quasiparticles on quasi-one-dimensional sections of the Fermi surface will not complete the closed orbits (see **Figure 5b**).

The experimental detection of cyclotron resonance depends on the effect that the quasi-particle paths have on the high-frequency conductivity. Equation [2] shows that the quasiparticle velocity must always be perpendicular to the Fermi surface. Hence, as the quasiparticles traverse the orbits shown in **Figure 5a** and **5b**, the corrugations of the Fermi surface cause their velocity vectors to rock back and forth. It is the evolution of the quasiparticle velocities with time that affects the conductivity of the system, and thus enables cyclotron resonance to be detected. Cyclotron resonance is of course associated with the closed orbits of **Figure 5a**; the open orbits shown in **Figure 5b** lead to a related resonant contribution to the high-frequency magnetoconductivity, known as a Fermi surface

traversal resonance (FTR) (see the "Further reading" section for more information on FTRs).

As mentioned above, the low electron densities of these materials compared to those in typical elemental metals result in a larger skin depth; hence, with careful planning, an experiment can be arranged in which a substantial fraction of a typical crystalline sample (volume ~ 0.01–$0.1\,\mathrm{mm}^3$) can be penetrated by radiation with frequencies ~ 10–$100\,\mathrm{GHz}$. Therefore, the experimental response is typically much more determined by the frequency-dependent bulk conductivity, rather than the surface impedance.

The evolution of the quasiparticle's velocity can be related to its contribution to the frequency-dependent conductivity of the metal using the frequency-dependent Boltzmann transport equation (also known as the Chambers equation):

$$\sigma_{ij}(\omega) = \frac{e^2}{4\pi^3} \int \mathrm{d}^3k \left[-\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})}\right] v_i(\boldsymbol{k}, 0) \int_{-\infty}^{0} v_j(\boldsymbol{k}, t) \times \cos(\omega t) \mathrm{e}^{t/\tau}\, \mathrm{d}t \qquad [13]$$

Here, τ is the relaxation time (not necessarily the same as τ_{CR}), $v_i(\boldsymbol{k}, t)$ is the ith component of the velocity of a quasiparticle with wave vector $\boldsymbol{k}$ at time t, and $f_0(\boldsymbol{k}) = (\mathrm{e}^{(E(\boldsymbol{k})-E_{\mathrm{F}})/k_{\mathrm{B}}T} + 1)^{-1}$ is the Fermi function. This is an integral (over all states at the Fermi surface) of the velocity–velocity correlation function for each Fermi surface orbit.

The mechanism for detecting cyclotron resonance can be understood by examining **Figure 5a**. For a general direction of the magnetic field (as shown in **Figure 5**), the corrugations of the Fermi cylinder will cause a quasiparticle's interplane (z) velocity to acquire an oscillatory component as it is driven round an orbit about the Fermi cylinder. These oscillatory interplane velocities contribute at a particular frequency (and, perhaps, its harmonics) to eqn [13]; hence they are detectable as resonances in the high-frequency interplane conductivity $\sigma_{zz}(\omega)$.

Note that if the corrugations are not exactly perpendicular to the cylinder axis, then even a magnetic field parallel to the cylinder axis (i.e., in the interplane z-direction) will induce an oscillatory interplane component of the quasiparticle velocity and hence a resonant response in σ_{zz}. This resonant response in σ_{zz} occurs at integer harmonics of the cyclotron frequency, $j\omega_{\mathrm{c}}$; these can be caused both by the orbit traversing j corrugations or by the presence of nonsinusoidal corrugations.

In addition, Nam *et al.* proposed a mechanism for cyclotron resonance harmonics in the intraplane high-frequency conductivity components σ_{xx} and σ_{yy}; this results from Fermi surface cross sections that are nonelliptical. Unless the Fermi surface lacks inversion symmetry, this mechanism generates only odd harmonics. Both types of cyclotron resonance mechanisms with their attendant harmonics were reported by Rzepniewski *et al.* in the layered perovskite Sr_2RuO_4.

Cyclotron resonance in such materials is observed using millimeter-wave cavity perturbation techniques such as those described by Schrama *et al.* and Mola *et al.* (see **Figure 6**). Such measurements probe the bulk conductivity properties of anisotropic conductors. Changes in the dissipation of the cavity (Q-factor) are measured as a function of an external quasistatic magnetic field $\boldsymbol{B}$ with the millimeter-wave frequency ν held constant. It is assumed that changes in the dissipation of the cavity are due to changes in the dissipation of the sample inside.

The sample is placed in the cavity such that the millimeter-wave magnetic field $\boldsymbol{H}_{\mathrm{osc}}$ is polarized parallel to the sample's highly conducting planes (perpendicular to the c-axis) (see **Figure 6**). The response of an anisotropic conductor in this electromagnetic environment is understood by examining the polarizations of the currents induced in the sample by $\boldsymbol{H}_{\mathrm{osc}}$. For $\boldsymbol{H}_{\mathrm{osc}}$ parallel to the (a, b) plane (the highly conducting plane), both in-plane and interlayer currents are induced (**Figure 6**). These currents flow within a distance $\sim\delta$ (here δ is the skin depth) from the sample's edges and faces (**Figure 6b**). The in-plane ($\delta_{\parallel}$) and interlayer ($\delta_{\perp}$) skin depths give information about σ_{zz} and a combination of σ_{xx} and σ_{yy}, respectively.

Extraction of the apparent scattering rates in layered metals As described by Schrama *et al.* and Mola *et al.*, changes in the Q-factor of the resonant cavity can be related via the skin depths to changes in the components of the sample's conductivity tensor. Generally, the most reliable method of extracting the apparent scattering time is to fit the high-frequency conductivity directly (see **Figure 7**); for example, McKenzie suggests an interlayer conductivity

$$\frac{\sigma_{zz}(\omega)}{\sigma_{zz}(\omega = 0)} = \sum_{-\infty}^{\infty} \frac{[J_j(k_{\mathrm{F}} c \tan\theta)]^2}{1 + (\omega - j\omega_{\mathrm{c}} \cos\theta)^2 \tau_{\mathrm{CRM}}^2} \qquad [14]$$

for layered materials with weak interlayer coupling that has been successfully applied to organic conductors. Here, j is an integer, J_j is the jth Bessel function, c is the interlayer lattice spacing, k_{F} is the Fermi wave vector, and θ is the angle between the magnetic field and the normal to the conducting planes. Note that the apparent scattering time τ_{CRM} is always a fit parameter in such methods.

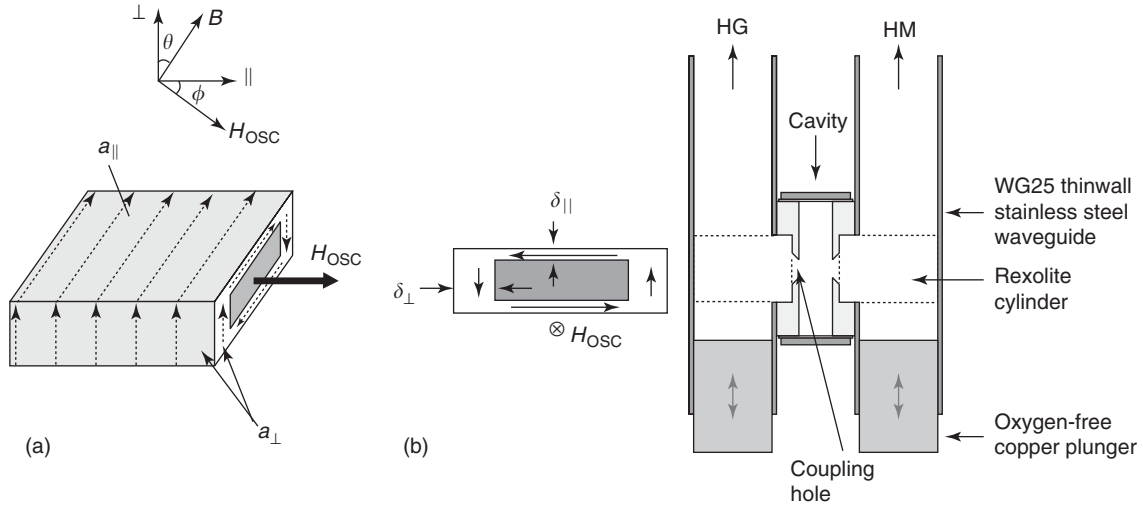


Figure 6 Experimental configuration for cyclotron resonance in a layered metal. Left: (a) the sample is placed in the cavity with the oscillating millimeter-wave field $\boldsymbol{H}_{osc}$ in the highly conducting planes ($\parallel$ direction). (b) Both in-plane and interplane oscillating currents J_{osc} are induced. In the skin-depth regime, in-plane currents flow within a surface layer of thickness $\delta_{\parallel}$ parallel to large sample faces, and interplane currents flow within a thicker layer $\delta_{\perp}$ parallel to sample edges. Right: schematic of the vertical cross section of a rectangular cavity that can be used for cyclotron resonance measurements of a layered metal. The cavity is placed between two waveguides that lead to the source (HG) and detector (HM). The cavity is held in this position by two Rexolite cylinders of 3 mm diameter which are fitted in the recessed area around the coupling holes and which extend through the holes (3 mm diameter) in the waveguides. The sample is placed on a thin mylar film in the cavity center, in the *H*-field antinode of the TE_{102} mode of the cavity. The advantage of such a cavity system is that it allows the cavity and the sample to be rotated in the vertical quasistatic magnetic field (provided by either a superconducting or a Bitter magnet); because the cavity rotates, the sample remains in the same high-frequency electrodynamic environment. Hence, the magnetic-field-orientation dependence of the cyclotron resonance can be studied (the mechanism for rotating the cavity has been omitted for clarity). (Modified from Schrama JM, *et al.* (2001) *Journal of Physics*: *Condensed Matter* 13: 2235.)

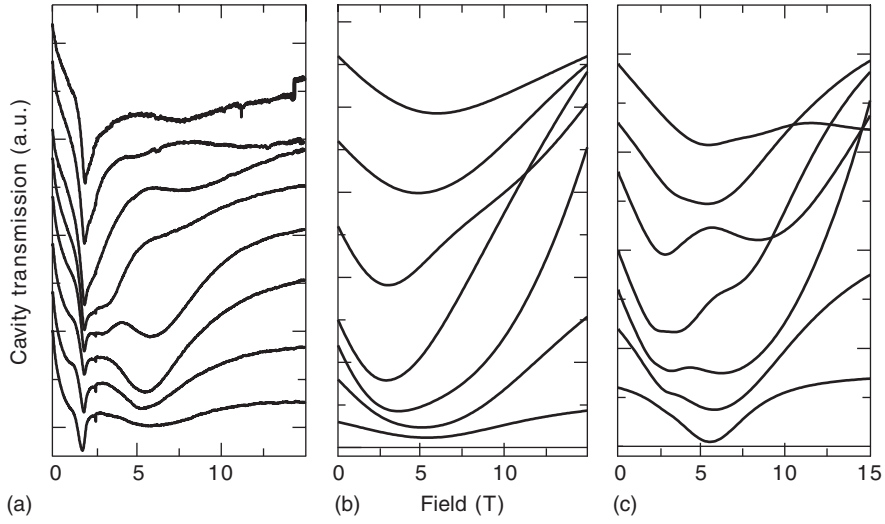


Figure 7 (a) Cyclotron resonance data for the quasi-two-dimensional organic conductor β''-(BEDT-TTF)$_2$$SF_5CH_2CF_2SO_3$. The data are in the form of cavity transmission versus magnetic field; the frequency of the measurement is 71 GHz, and the temperature is 1.5 K. Field sweeps for several orientations θ of the cavity and sample in the quasistatic field are shown (see **Figure 6**); $\theta = 0$ corresponds to the field being perpendicular to the highly conducting planes of the sample. The sharp slope downwards and small minimum at low fields are associated with the superconducting-to-normal transition of the sample and a background feature in the cavity response. The small sharp feature is the conduction-electron spin resonance of the sample. The cyclotron resonances are the broader features at higher fields. (b) A model of the data using eqn [14] with k_F calculated for this particular sample orientation and $\tau_{CRM} = 1.22 \times 10^{-12}$; the latter value was derived from a simple treatment (see eqn [16]) of Shubnikov–de Haas oscillations from the same sample. Note how the predicted resonances are too broad. (c) A repeat of (b) for $\tau = 7 \times 10^{-12}$ s; note how the predicted resonance widths are similar to the experimental ones. All figures are for θ between 0° and 70° in 10° steps. (Modified from Singleton J and Edwards RS (2003) In: Herlach F and Miura N (eds.) *High Magnetic Fields*, *Science and Technology*, vol. 2, 85 pp. Singapore: World Scientific.)

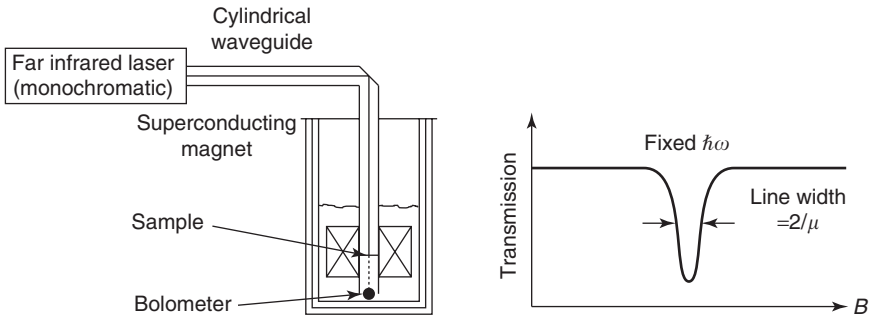


Figure 8 Schematic of a cyclotron resonance experiment in a semiconductor.

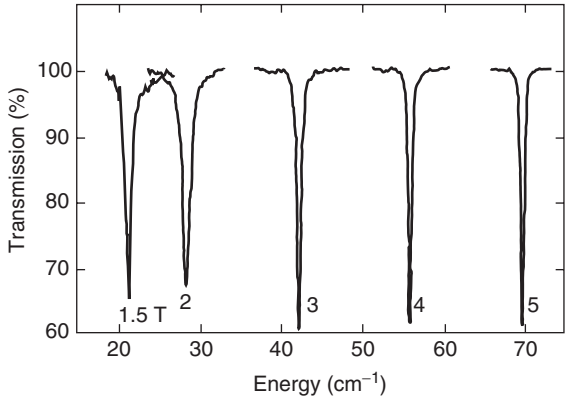


Figure 9 Cyclotron resonances in a GaAs-(Ga,Al)As heterojunction at $T = 4.2\,\mathrm{K}$; the magnetic field has been applied perpendicular to the two-dimensional electron layer, which has an areal carrier density of $\sim 9 \times 10^{10}\,\mathrm{cm}^{-2}$. The data have been recorded by fixing the magnetic field at 1.5, 2, 3, 4, and 5 T; at each field, a Fourier-transform spectrometer has been used to record the transmission of the sample as a function of energy. (Data from Wiggins G, University of Oxford.)

Cyclotron Resonance in Semiconductors

Experimental details The number densities of carriers in semiconductor samples are much lower than those in metals, so that the radiation can completely penetrate even large samples. A simple transmission arrangement is usually adopted (see **Figure 8**). The cyclotron resonance is usually recorded by measuring the sample transmission; experiments are carried out either by fixing the magnetic field and varying the energy of the radiation (see **Figure 9**), or by using a fixed-frequency source (such as a far-infrared laser, which can give a number of monochromatic laser lines in the wavelength range 40–1000 μm) and sweeping the magnetic field (see **Figure 8**).

The magnetic field is usually provided by a superconducting magnet ($B = 0$ to $\sim 20\,\mathrm{T}$). As the whole of the cyclotron orbit experiences the electric field of the radiation (cf. the situation in metals; see above), the quantum-mechanical selection rule for the Landau-level quantum number ($\Delta l = \pm 1$) holds. Thus, the resonance condition is

$$h\nu \equiv \hbar\omega = \hbar\omega_c = \frac{eB}{m^*_{CR}} \qquad [15]$$

where $\nu(\omega)$ is the frequency (angular frequency) of the radiation.

As stated earlier, the degenerate semiconductor systems (e.g., heterojunctions) are dealt with in the first section. In such samples, free carriers are present even at low temperatures; typical cyclotron resonance data are shown in **Figure 9**. However, it is worth mentioning that in lightly doped samples, the carriers must be excited into the bands by either raising the temperature to cause the impurities to ionize (but not so far as to broaden the Landau levels) or by additionally illuminating the sample with above bandgap radiation.

Extraction of apparent scattering times in semiconductor systems In some situations where both the scattering time (or mobility) and the density of the carriers are low (e.g., bulk degenerate semiconductors), the measured line width of the cyclotron resonance can give information about the true scattering rate τ^{-1}_{CR} directly. Via the uncertainty principle, the scattering induces a frequency uncertainty $\Delta\omega_c \sim \tau^{-1}_{CR}$. If the experiment is a fixed-frequency, swept-field one, this translates to an uncertainty (that is, resonance width) in magnetic field of $\Delta B = \Delta\omega_c m^*_{CR}/q \sim m^*_{CR}/q\tau_{CR}$. Provided that $m^*_{CR} \approx$ the effective mass for linear motion (and this is often true in direct-gap semiconductors such as GaAs), it is apparent that ΔB is of the order of the reciprocal of the carrier mobility. More precise derivations show that the full-width at half-maximum is in fact twice the reciprocal of the mobility.

However, in systems such as GaAs-(Ga, Al)As heterojunctions, where both the scattering time (or mobility) and the equivalent three-dimensional carrier density in the two-dimensional layer are relatively

large, there is a very marked contribution to the cyclotron resonance line width due to the impedance mismatch between the vacuum and the carriers. The effect is illustrated in **Figure 10**; often this "mismatch" contribution to the line width eclipses that due to scattering. The problem can be ameliorated by evaporating a surface metallic layer of carefully chosen thickness on the sample (see, for example, the paper by Watts *et al.* in the literature list), but the most reliable method to extract the apparent scattering rate is to perform a full modeling of the transmission or reflectivity of the sample, as carried out in **Figure 10**.

Contributions to the Apparent Scattering Rate in Layered Metals and Semiconductor Heterostructures

Comparisons with other methods A measure of the scattering rate in metallic systems is often derived from the rate at which magnetic quantum oscillations (such as de Haas–van Alphen oscillations) grow in amplitude with the increasing magnetic field; the dominant term in the Lifshtiz–Kosevich formula (see, e.g., the book by Shoenberg in the Further reading section) describing this phenomenon is $\exp[-14.7m\ \mathrm{CR}^*\ T_\mathrm{D}/B]$ (SI units). The constant describing the phase smearing of the oscillations due to Landau-level broadening is the so-called Dingle temperature, T_D.

If one were to assume that T_D is solely due to scattering (i.e., the energy width of the Landau levels detected by the de Haas–van Alphen effect is associated only with their finite lifetime due to scattering), then T_D would be related to the scattering rate τ_dHvA^{-1} by the expression

$$T_\mathrm{D} = \frac{\hbar}{2\pi k_\mathrm{B}\tau_\mathrm{dHvA}} \qquad [16]$$

In many experimental works, it is assumed that the τ_dHvA deduced from T_D in this way is a true measure of the scattering rate; it will be noted shortly that it is not!

Another measure of the scattering rate can be found by measuring the zero-magnetic-field electrical conductivity; in layered systems, one must be careful to ensure that this is the intralayer component of the conductivity. If there is a dominant single carrier type of known density n, then the Drude expression $\sigma = nq^2\tau_\sigma/m^*$ can be used to estimate τ_σ. It is universally found that $\tau_\sigma \geqslant \tau_\mathrm{dHvA}$; in some cases such as semiconductor heterojunctions, τ_σ can be one to two orders of magnitude greater than τ_dHvA.

Finally, if these scattering times are compared with the τ_CRM deduced from cyclotron resonance experiments, the following is found:

$$\tau_\sigma \geqslant \tau_\mathrm{CRM} \geqslant \tau_\mathrm{dHvA} \qquad [17]$$

For some layered metals (see, e.g., the work of Hill in "Further reading" section), the τ_CRM measured in

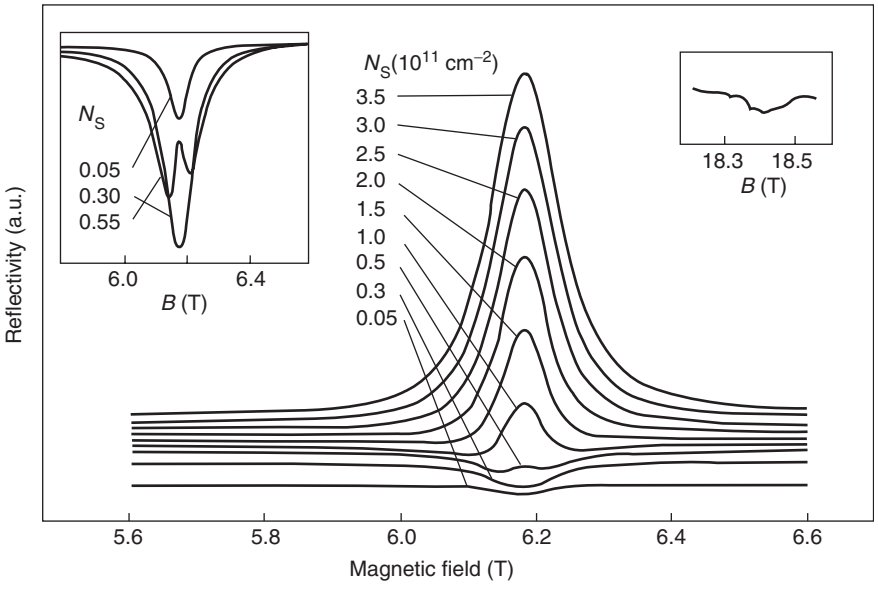


Figure 10 Model calculations for the reflectivity of a GaAs-(Ga,Al)As heterojunction at a photon energy $\hbar\omega = 10.43$ meV for several values of the areal electron density N_s; each calculation is displaced vertically for clarity. The cyclotron resonance shows up as a dip in the reflectivity signal at low N_s but becomes a peak at higher carrier densities. Note that although the scattering rate is kept constant ($\tau = 10.6$ ps), the full-width at half-maximum of the resonance grows with increasing N_s. Similar effects are seen in the transmission of the sample. The left-hand inset shows an expansion of the low-density part of the main figure; the right-hand inset contains experimental data for a photon energy of 31.05 meV and $N_s = 0.3 \times 10^{11}\ \mathrm{cm}^{-2}$. (Modified from Langerak CJGM, *et al.* (1988) *Physical Review B* 38: 13133.)

cyclotron resonance has been found to be four to ten times larger than τ_{dHvA}. An example of this is shown in **Figure 7**, where the insertion of the scattering rate inferred from Shubnikov–de Haas oscillations (**Figure 7b**) into a model for the cyclotron resonance produces line widths that are much too broad. A more realistic line width is obtained with a longer scattering time (**Figure 7c**).

The importance of spatial inhomogeneities The difference between any measurement of the Landau-level broadening and the scattering rate deduced from a conductivity measurement is easy to grasp; a quasiparticle will be removed from a Landau-level eigenstate by any scattering event, small- or large-angle. Conversely, small-angle scattering events hardly affect the conductivity, because they cannot randomize a quasiparticle's excess forward momentum (see **Figure 11** and standard texts such as Ashcroft and Mermin or Singleton). In the most thoroughly studied systems of reduced dimensionality, semiconductor heterostructures, and Si MOSFETs, it is found that $\tau_\sigma \approx \tau_{dHvA}$ in systems in which large-angle scattering (e.g., due to impurities in the conducting channel) dominates. On the other hand, in systems in which small-angle scattering (e.g., due to remote impurities) is pre-eminent, $\tau_\sigma \gg \tau_{dHvA}$. Having seen this, the remaining task is to explain how two measurements of the Landau-level width (cyclotron resonance and de Haas–van Alphen effect) can yield such different answers.

The difference between τ_{CRM} and τ_{dHvA} turns out to be due to spatial inhomogeneities or potential fluctuations. Screening is less effective in systems containing low densities of quasiparticles (such as layered metals or semiconductor heterojunctions), compared to that in elemental metals; hence, variations in the potential experienced by the quasiparticles can lead to a spatial variation of the Landau-level energies (see **Figure 11**). Even in the (hypothetical) complete absence of scattering, Harrison, for example, has shown that this spatial variation would give the Landau level a finite energy width (see **Figure 11**) and therefore lead to an apparent Dingle temperature

$$T_D = \frac{\bar{x}[1-\bar{x}]F'(\bar{x})^2 a}{\pi k_B m^*}\sqrt{\frac{\hbar e^3}{2F}} \qquad [18]$$

Here, F is the magnetic-quantum-oscillation frequency, and $F' = dF/dx$; x represents the (local) fractional variation of the quasiparticle density due to the potential fluctuations and $\bar{x}$ is its mean.

The Dingle temperature measured in experiments therefore normally represents a combination of the effects described by eqns [16] and [18]. Hence, the simple-minded use of eqn [16] to yield τ_{dHvA} from T_D tends to result in a parameter that is an overestimate of the true scattering rate (see **Figure 11**). By contrast, cyclotron resonance (shown by vertical arrows in **Figure 11**) is a "vertical" transition (due to the very low momentum of the photon); it measures just the true width of the Landau levels due to scattering (represented by shading).

Summary

This has been a complicated section, with many parameters and effects to grasp. It is therefore worth giving a summary of the factors affecting

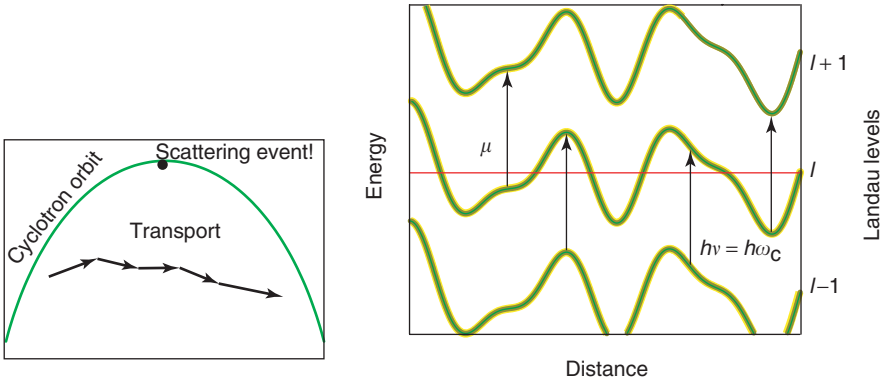


Figure 11 Left: an illustration of the difference in scattering rates in Landau quantization and electrical conductivity. Any scattering event (small- or large-angle) during the cyclotron orbit terminates that particular Landau state. By contrast, small-angle scattering is not effective at randomizing the quasiparticle's excess forward momentum in a transport (conductivity) measurement. Right: an illustration of the effect of spatial inhomogeneities on Landau-level widths and energies. The variations in the potential experienced by the electrons make the Landau levels (shaded curves) move up and down in energy as one moves through the sample. As the field is swept, the levels move up through the chemical potential μ and depopulate, resulting in the de Haas–van Alphen and Shubnikov–de Haas effects. The Dingle temperature essentially parametrizes the movement of the total energy width of a Landau level through μ; hence, it measures a width that includes the energy variation due to inhomogeneities. By contrast, cyclotron resonance (shown by vertical arrows) is a "vertical" transition; it measures just the true width of the Landau levels due to scattering (represented by shading).

measurements of the scattering rate in layered metals and degenerate semiconductor systems.

Provided a proper treatment of the sample's high-frequency magnetoconductivity is used, the scattering rate τ_{CRM}^{-1} deduced from a cyclotron resonance experiment is a good measure of the energy width of the Landau levels due to their finite lifetime. In other words, $\tau_{\mathrm{CRM}} \approx \tau_{\mathrm{CR}}$. (Once this has been realized, cyclotron resonance can be used to give quantitative details of the quasiparticle scattering mechanisms; see Further reading section for more details.)

By contrast, the apparent scattering rate deduced from a simple-minded treatment of de Haas–van Alphen and Shubnikov–de Haas oscillations can contain very significant contributions from spatial inhomogeneities.

Finally, the scattering rate τ_{σ}^{-1} measured in a zero-field conductivity experiment can be very significantly different from τ_{CR}^{-1}, because the two measurements are sensitive to different types of scattering processes.

Systems With Two Interacting Carrier Populations

Cyclotron resonance data can sometimes be difficult to interpret in systems containing two distinct quasiparticle populations that interact. An example is given by *p*-doped GaAs-(Ga,Al)As heterojunctions, containing two hole spin subbands with differing effective masses. One would expect two distinct cyclotron resonances; however, depending on the temperature, photon energy, and the field, a single resonance is frequently observed. This results from the interactions between the two types of quasiparticles; the overall energy of the system is reduced if these interactions constrain both populations to have the same cyclotron frequency.

Cole *et al*. have derived expressions for the high-frequency magnetoconductivity $\sigma(\omega)$ that can be very generally used to treat such situations; they allow the scattering rates and effective masses for each type of quasiparticle to be extracted, and also enable their interactions to be quantified.

Which Effective Mass Does a Cyclotron Resonance Experiment Measure?

Cyclotron resonance is potentially a very interesting measure of the interactions in narrow-bandwidth metallic systems. In a translationally invariant system, the mass m_{CR}^{*} measured in a cyclotron resonance experiment should not contain any contribution from the interactions between the carriers themselves; this is known as Kohn's theorem. In the language of the Fermi-liquid theory, in such a situation, cyclotron resonance measures the dynamical mass $m_{\lambda\mathrm{CR}}$, whereas a thermodynamic measurement such as the de Haas–van Alphen oscillations reveals the effective mass m^{*}, which also contains contributions from the Coulomb interactions between the quasiparticles.

A simple band-structure calculation essentially uses ions and molecules which are rigidly fixed in a perfectly periodic arrangement to obtain a periodic potential and hence the bands. However, the ions and/or molecules in a substance will in general be charged, or at the very least possess a dipole moment; as an electron passes through the solid, it will tend to distort the lattice around it owing to the Coulomb interactions between the ions and molecules and its own charge. This leads to the electron being accompanied by a strain field as it moves through the substance; alternatively, one can consider the electron being surrounded by virtual phonons. This affects the way in which the electron can move through the substance and usually acts to increase the apparent mass.

In the following, the orbitally averaged (cyclotron) mass is dealt with for consistency; other "effective masses" (e.g., for linear motion) are renormalized in an analogous way. If m_{bCR} is the cyclotron mass calculated in the band-structure calculations, then the electron–lattice interactions discussed above will result in a cyclotron dynamical mass, $m_{\lambda\mathrm{CR}}$, where

$$m_{\lambda\mathrm{CR}} \approx (1+\lambda) m_{\mathrm{bCR}} \qquad [19]$$

where λ is an electron–phonon coupling constant.

The interactions between the electrons themselves must also be taken into account. As electrons are highly charged, they repel each other; therefore, as an electron is moved across a sample under consideration by an external force, in effect there will be a backflow of electrons caused by this repulsion. Thus, it is "harder" to move electrons about than expected, that is, their apparent effective mass is heavier than $m_{\lambda\mathrm{CR}}$. The effective masses m^{*} measured in experiments such as the de Haas–van Alphen effect include this contribution from the interactions between the electrons, so that

$$m^{*} \approx \left(1 + \frac{F_{\mathrm{s}}^{1}}{3}\right) m_{\lambda} \qquad [20]$$

where F_{s}^{1} is a constant known as a Fermi-liquid parameter. (For a more detailed discussion of such effects, the reader is referred to the "Further reading" section.)

In practice, however, real systems do not possess translational invariance. Although the mass m_{CR}^{*} measured in a cyclotron resonance experiment does generally differ from that obtained from the de Haas–van Alphen effect (m^{*}), the simple expectations

of eqn [20] often do not hold; in one or two instances, m^*_{CR} was even found to be greater than m^*. Nevertheless, a comparison between a cyclotron resonance experiment and the de Haas–van Alphen effect is often a good guide to the strength of the interactions in a system.

Effects Related to Cyclotron Resonance

There are one or two effects related to cyclotron resonance that go beyond the scope of this article but which should nevertheless be mentioned for completeness. One is "surface cyclotron resonance," an effect due to excitation of quasiparticles between levels caused by crossed electric and magnetic fields at the surface of a sample. Interested readers are referred to the articles by Koch and Jensen and by Merkt in the literature list. The other one is the FTR due to the magnetic-field-induced motion of quasiparticles across open sections of a Fermi surface; a discussion and relevant literature are given in the article by Schrama *et al*.

See also: Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Shubnikov–de Haas and de Haas–van Alphen Techniques.

PACS: 76.40.+b; 72.10.Di; 72.10.Fk; 71.10.Ay; 71.70.Di; 73.21.−b; 78.20.Ls

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.
Cole BF, *et al.* (1997) *Journal of Physics: Condensed Matter* 9: 3163.
Gross EKU, Runge E, and Heinonen O (1991) *Many Particle Theory*. Bristol: Adam Hilger.
Harrison N and Singleton J (2001) *Journal of Physics: Condensed Matter* 13: L463.
Häussler P and Welles SJ (1966) Determination of relaxation times in cyclotron resonance in copper. *Physical Review* 152: 675.
Herlach F and Miura N (2003) *High Magnetic Fields*, *Science and Technology*, vol. 2, Theory and Experiments. Singapore: World Scientific.
Hill S (1997) *Physical Review B* 55: 4931; *Physical Review B* 62: 8699.
Koch JF and Jensen JD (1969) *Physical Review* 184: 643.
Langerak CJGM, *et al.* (1988) *Physical Review B* 38: 13133.
Lindholm DA and Le Page J (1969) *Physical Review* 187: 861.
McKenzie RH and Moses P (1999) *Physical Review B* 60: R11 241.
Moses P and McKenzie RH (1999) *Physical Review B* 60: 7998.
Merkt U (1985) *Physical Review B* 32: 6699.
Mola M, *et al.* (2000) *Review of Scientific Instruments* 71: 186.
Moss TS and Balkanski M (1994) *Optical Properties of Semiconductors*, chaps. 6,7, and 11. Amsterdam: Elsevier.
Nam MS, *et al.* (2001) *Journal of Physics: Condensed Matter* 13: 2271.
Quader KF, *et al.* (1987) *Physical Review B* 36: 156.
Rzepniewski E, *et al.* (2002) *Journal of Physics: Condensed Matter* 14: 3759.
Schrama JM, *et al.* (2001) *Journal of Physics: Condensed Matter* 13: 2235.
Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.
Singleton J (2001) *Band Theory and Electronic Properties of Solids*, chaps. 8 and 9. Oxford: Oxford University Press.
Vaughan TA, *et al.* (1996) *Physical Review B* 53: 16481.
Watts M *et al.* (1992) In: Landwehr G (ed.) *High Magnetic Fields in Semiconductor Physics*, Springer Series in Solid State Sciences, vol. 101. 581p. Berlin: Springer.

Nomenclature

A	cross-sectional area of the Fermi surface
$\boldsymbol{B}$	magnetic induction
E	energy of a quasiparticle
$\boldsymbol{E}$	electric field
j	a general integer used as a label
$\boldsymbol{k}$	quasiparticle wave vector quantum number
$k_\parallel$	the component of $\boldsymbol{k}$ parallel to $\boldsymbol{B}$
l	Landau-level quantum number
m^*_{CR}	cyclotron effective mass
m^*	the quasiparticle mass measured in a de Haas–van Alphen experiment
q	quasiparticle charge: $q = \pm e$
t	time
$t_\perp$, $t_\parallel$	transfer integrals
$\boldsymbol{v}$	quasiparticle velocity
Z	surface impedance
ν	frequency of photons used to excite cyclotron resonance
σ, σ_{ij}	conductivity, general element of conductivity tensor
τ_{CR}	scattering time associated with Landau-level broadening
τ_{CRM}	apparent scattering time from a cyclotron resonance measurement
τ_{dHvA}	apparent scattering time deduced from a de Haas-van Alphen measurement
τ_σ	scattering rate inferred from a zero-field conductivity measurement
ω	angular frequency of photons used to excite cyclotron resonance
ω_c	cyclotron (angular) frequency

Cyclotron Resonance: Semiconductors

D R Leadley, University of Warwick, Coventry, UK

Introduction

Semiconductors, as a class of materials, have attracted a huge amount of attention not only due to their intrinsic fascination but also due to their use in electronic and optoelectronic applications. At the most basic level, it is important to measure values for the electron and/or hole effective mass m^*, which is related to the band structure and determines how fast these charge carriers can be accelerated. Such measurements have been performed using the phenomenon of cyclotron resonance (CR) since it was first proposed, by Shockley in 1953, on a whole variety of elemental and compound semiconductors including bulk 3D and layered 2D structures. However, cyclotron resonance is capable of providing much more information than just the value of m^* concerning the band structure of the host material as well as the interactions of electrons with each other and the surrounding lattice.

This article reviews the basic mechanism of CR, considers how the resonance may be observed experimentally, and then discusses some of the information that can be obtained from the resonance position and line width. In particular, it addresses the issue of how the measured m^*_{CR} should be related to the band edge value m^*_0 and how energy level crossings, polaron coupling, spin splitting, and collective interactions affect the measured CR position.

So What Is Cyclotron Resonance?

In a magnetic field B, the previously continuous spectrum of electronic energies becomes quantized into a discrete set of Landau levels (LLs) (as described elsewhere in this volume) with energies

$$E_l(B, k_\parallel) = \left(l + \tfrac{1}{2}\right)\hbar\omega_c \pm \tfrac{1}{2}g\mu_B B + E(k_\parallel) \qquad [1]$$

where l is an integer, $\omega_c = eB/m^*$ is the cyclotron frequency, by analogy to cyclotron motion in free space when $m^* = m_0$, the second term accounts for spin splitting, and $E(k_\parallel)$ is the energy of motion (or confinement) parallel to the magnetic field. The number of occupied LLs is known as the filling factor ν. (Care is required with this number as different workers may, or may not, include spin and valley degeneracy: a full LL could be described as having $\nu = 1$ in a bulk semiconductor, $\nu = 2$ in a GaAs heterojunction, or $\nu = 4$ in an Si MOSFET.) Since the free carrier density in most semiconductors is much smaller than the atomic density, the Fermi surface only extends to a small fraction of the Brillouin zone and, to a first approximation, can be regarded as ellipsoidal; so the cyclotron mass is equivalent to the effective mass for transport in a plane perpendicular to B.

Transitions can be made between these LLs by absorbing a photon of the appropriate energy $\hbar\omega = \hbar\omega_c$ – this is the cyclotron resonance. The examples in this article are taken from degenerate semiconductors, where the optical transitions are within either the conduction or the valance band. Traditionally, nondegenerate semiconductors have also been studied by CR using cross-bandgap radiation. These interband transitions can reveal a wealth of information about the electronic structure of semiconductors, but are less straightforward to analyze as both bands are involved and are not considered further here.

Most CR experiments are conducted in the Faraday geometry, where the direction of propagation of the radiation is along B. Hence, the oscillating electric field is polarized in the plane perpendicular to the magnetic field, and this is the direction of the electron motion that is sensed. Another consequence of the low free carrier density is that the skin depth will be large and the same electric field is seen throughout the sample, irrespective of the magnetic field, so the radiation couples to the center of mass motion of the electrons and is insensitive to carrier–carrier interactions – Kohn's theorem. The CR transition has the selection rule $\Delta l = \pm 1$, which accounts for the unit of angular momentum absorbed from the photon, and is spin preserving. The choice of sign means that, by circularly polarizing the radiation only, certain transitions will occur: conduction band electrons will only be excited to higher LLs by σ^+ polarization, so this is termed electron-active, and the hole CR will generally only be excited with σ^- polarization (although σ^+ will excite certain transitions in the valence band between heavy and light hole levels).

Scattering processes within the semiconductor means that the LLs are not infinitely narrow, so the CR is broadened and the line width can be used as a quantitative measure of the sample quality. If the average carrier scattering time is τ, then CR can be observed provided $\omega_c\tau > 1$, which, in a semiclassical view, means the carrier must substantially complete a cyclotron orbit before scattering. This condition can be rewritten as $\mu B > 1$ (as mobility $\mu = e\tau/m^*$), setting a minimum magnetic field of $1/\mu$ for the observation, that is, 10 T for a $1000\,\mathrm{cm^2\,V^{-1}\,s^{-1}}$

sample. In practice, this means CR is normally recorded at low temperature (4.2 K), where the mobility is highest, and low-mobility semiconductors can only be accessed with the highest available magnetic fields.

Further broadening of the CR line can arise from dielectric mismatch, especially within layered structures and at frequencies close to poles in the dielectric response. In some cases, this can completely mask the effects of lifetime broadening and even change the line from the clean Lorentzian of **Figure 1** to a differential shape, from which it is much more difficult to extract the true resonance position. This effect can often be corrected for by modeling the dielectric response of the particular sample being investigated.

Magnetoplasmon Mode

The above description of CR as a resonant absorption is adequate in most cases, but is an oversimplification. The CR absorption spectrum can be calculated within the linear response theory as being proportional to the real part of the dynamical conductivity. FIR illumination presents a long wavelength perturbation to the electron plasma, which has a collective response. In a magnetic field, the plasma excitation modes occur close to the integer multiples of $\hbar\omega_c$. The lowest of these is referred to as the magnetoplasmon mode, and at small wave vector k, is equivalent to forming a magnetic exciton with an electron in the upper LL and a hole remaining in the lower LL. The excitation has an energy $E(k) = \hbar\omega_c + \delta E(k)$, where the second term is the magnetoexciton binding energy due to electron correlations. Pure CR corresponds to the $k = 0$ limit of the magnetoplasmon, in which limit $\delta E(k) \to 0$ due to Kohn's theorem. However, lattice defects, impurities, band nonparabolicity, and the structural asymmetry often found in 2D systems can break translational invariance and allow $k \neq 0$ modes to mix with the pure CR mode at $k = 0$. Coupling to other plasmon modes of the electron gas is then possible, especially when a metallic grating is applied to the surface of the semiconductor that imposes spatial periodicity on the FIR excitation and selects particular wave vectors. Although there have been suggestions of coupling to such magnetoplasmons in GaAs and GaN 2DEGs, the details of these measurements are currently not explained.

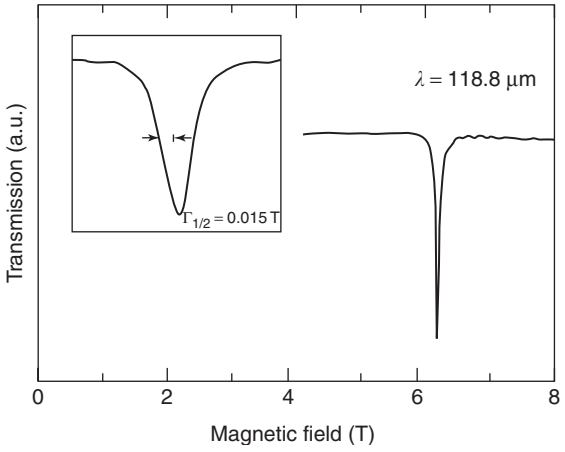


Figure 1 Typical FIR transmission through a high-mobility GaAs 2DEG showing a very narrow CR. (Reprinted with permission from Hopkins *et al.* (1987) *Physical Reviews B* 36: 4789; © American Physical Society.)

Experimental Techniques

Experimentally, the CR condition for a particular specimen can be satisfied and the resonance directly detected in one of two ways: either by applying a fixed frequency of radiation and varying the magnetic field, or by fixing the field and varying the frequency.

Swept Field, Fixed Frequency

Conceptually, the former is more straightforward. Magnetic fields up to ~ 30 T can now be generated in resistive magnets, that are designed to be swept, and wire-wound pulsed magnetic coils extend this field range to some 60 T. For a typical semiconductor, the effective mass is in the range of $0.01–1.0\,m_0$ which, at the now routinely accessible field of ~ 10 T, means sources of far infrared (FIR) radiation are required in the 10–1000 μm range (1–100 meV). Suitable monochromatic FIR can be obtained from a molecular gas laser, filled with methanol or deuterated methanol at low pressure, that is pumped by a CO_2 laser operating at 9–11 μm. Although thousands of lines can be obtained from such optically pumped FIR lasers, most experiments are limited to ~ 20 strong lines, with 118.8 μm being the most popular. The transmitted FIR is usually detected by a cooled C bolometer or Ge photodetector and noise reduced by phase locking with chopping of the CO_2 laser beam. To avoid interference fringes and distorted line shapes, samples are wedged by a few degrees. This experimental arrangement works well, but there are a number of practical difficulties: for example, the limited number of strong FIR lines, stabilization of the FIR intensity over the time of the magnet sweep, transmission of FIR to the sample through transparent windows and evacuated metal light pipes, which need to be highly polished on the inside to avoid attenuation, and the limited sensitivity of the detection arrangement. The last of these can sometimes be improved by taking the ratio of signals from

bolometers mounted just above and below the sample. It is also possible to detect the CR through changes in the resistance of the sample being investigated in a photoconductivity measurement. FIR experiments provide much higher resolution than the earlier microwave experiments in fields below 1 T, where the geometry of the microwave resonant cavity has to be considered and tuned carefully, although the quality of that data is a tribute to the skill of the researchers. Measurements at high fields can also be made in the "quantum limit," where only one LL is initially populated as opposed to the classical limit at low fields, where many LL transitions are involved and broaden the resonance. Between these limits, studies have been made of the effect of filling factor on the line width, which can oscillate with the density of available initial and final states.

Megagauss fields At higher pulsed magnetic fields now available in the megagauss region (>100 T), the more powerful CO_2 laser can be used directly, and other possible sources include water vapor, CO, and HeNe lasers. At the shorter wavelengths, fast HgCdTe photovoltaic detectors can be used to detect the rapid change in IR intensity. Magnetic fields up to 180 T can be produced using the single turn coil technique pioneered in the University of Tokyo. A capacitor bank is discharged through the coil to give currents of order 2 MA for $\sim 7\,\mu$s. The maximum field is set by the size of the capacitor bank and the coil diameter, which is ~ 10 mm. Although the coil explodes violently outwards, the cryostat and sample are left intact and the experiment can be repeated. Data from the detector and a pickup coil are transmitted via optical fiber to a shielded room where they are recorded on a fast multichannel digitiser. The pick-up voltage allows the field to be calibrated each time to an accuracy better than 2%. A set of typical CR recordings are shown for several II–VI semiconductors in **Figure 2**.

Even higher fields, beyond 600 T, can be obtained using an electromagnetic flux compression technique. A primary pulse current through an outer primary coil seeds a magnetic field and induces an opposing secondary current within an inner copper ring. The repulsive forces generated rapidly squeeze the ring causing it to implode and compress the trapped magnetic flux to a very high field density. Unfortunately, the entire cryostat and sample are destroyed in this one shot experiment! However, the extreme fields mean that CR can be studied in generally very low-mobility materials, such as the dilute magnetic semiconductor $In_{1-x}Mn_xAs$, where high doping with magnetic impurities reduces the mobility. Data are shown in **Figure 3** for a temperature-dependent study of CR in

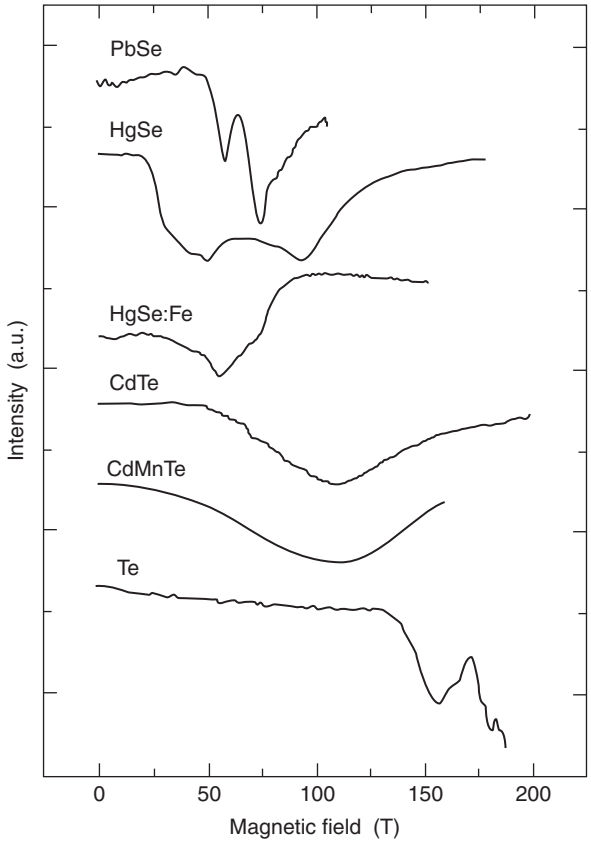


Figure 2 FIR transmission through a variety of II–VI semiconductors using a CO_2 laser with single turn coils. (Reproduced from Puhlmann *et al.* (1998) *Physica B* 246–247: 323–327, with permission from Elsevier.)

CdS. Notice that at room temperature the CR peak is over 100 T wide. (Further development of this technique continues in megagauss facilities, particularly in Tokyo University and the Humboldt University, Berlin, with fields of 300 T reported in 5 mm single turn coils and flux compression at 1000 T.)

By contrast, CR has also been observed at very low fields (20 mT), in very high quality GaAs/GaAlAs heterojunctions where the mobility exceeds $10^6\,cm^2\,V^{-1}\,s^{-1}$, using microwave radiation from 17 to 90 GHz Gunn diodes. At these low fields, the cyclotron radius is $\sim 0.2\,\mu$m which enables potential fluctuations in the sample to be observed on this macroscopic scale.

Fourier Transform Spectroscopy

The advent of reliable high-field superconducting magnets and Fourier transform infrared (FTIR) spectrometers in the late 1980s provided a real alternative to the swept field CR measurements at discrete energies, although that is still the only way for fast measurements in pulsed magnets. In this technique, the FIR source is a heated metal wire or globar producing a broad spectrum of radiation that passes

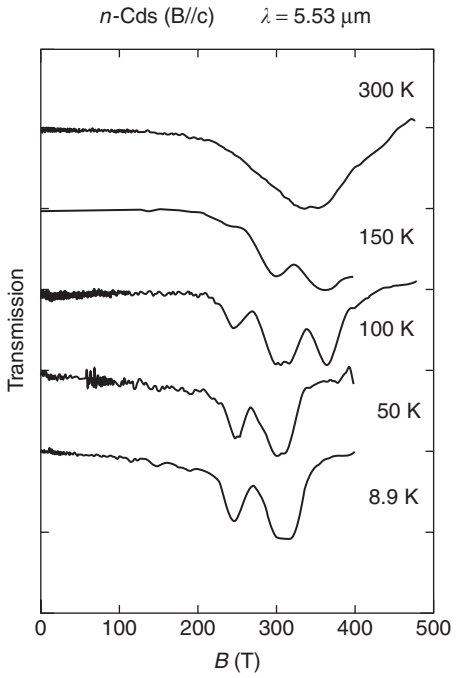


Figure 3 Temperature dependence of IR transmission through *n*-CdS at 5.53 μm using ultrahigh magnetic fields. The electron CR line at 300 K is replaced by other lines related to impurity and phonon-coupled CR as the temperature is reduced. (Reproduced from Imanaka *et al.* (1998) *Physica B* 246–247, 328–332, with permission from Elsevier.)

through a Michelson interferometer before passing through the sample and being detected by a cooled Si bolometer. By rapidly scanning the moving mirror in the interferometer, and recording the FIR intensity as a function of its position, an interferogram can be produced that is inverted by a hard-wired fast Fourier transform to yield a transmission spectrum with a resolution of $\sim 0.1\,\mathrm{cm}^{-1}$ (0.01 meV). Each scan only takes a fraction of a second, but can be repeated many times to increase the signal-to-noise ratio. Intensity variation in the source spectrum and detector response can be removed by taking the ratio of spectra with and without the sample present. In practice, a set of transmission spectra are usually recorded as the magnetic field is stepped through the range of interest (**Figure 4**) and differences are detected from the ratio of adjacent spectra. As a result of FTIR spectroscopy, the CR can be followed continuously as a function of magnetic field or energy, which allows measurements of band nonparabolicity and subband structure to be made.

Free-Electron Lasers

The limitation of a fixed number of FIR laser lines can also be overcome by using a free-electron laser

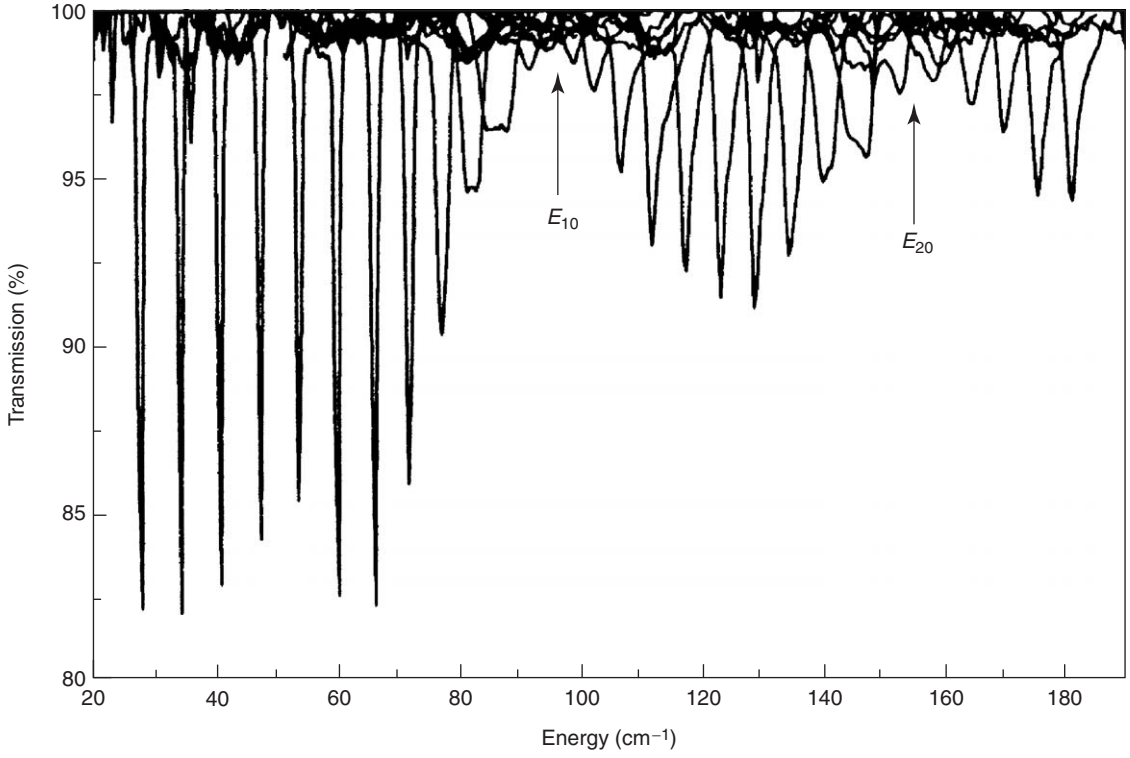


Figure 4 Series of FTIR spectra from 2 to 14 T at 0.5 T steps for a GaAs heterojunction at 3.5 K. The arrows indicate resonant intersubband coupling at 12.4 and 20 meV. (Reprinted with permission from Michels *et al.* (1995) *Physical Review B* 52: 2688; © American Physical Society.)

(FEL) that can provide radiation from UV to microwaves. The FEL system consists of an electron accelerator, an undulator in which the electrons emit synchrotron radiation, and an optical resonator. For a given undulator, the output is continuously tunable over a wide range, for example, 30–250 μm, and without changing the accelerator energy, the FIR can be rapidly scanned over a factor of 2 in wavelength. Since the electrons become bunched within the undulator, they emit radiation in a train of picosecond-duration micropulses within a macropulse that lasts for 0.01–5 ms with a ~10 Hz repetition rate, depending on the installation. At the FELIX facility in the Netherlands, a fast-pulsed magnet is synchronized to sweep during the arrival of the FIR macropulse enabling CR measurements at fields up to 60 T.

Time-resolved CR has been studied since the 1970s on the microsecond scale looking at, for instance, exciton lifetimes, but at the Stanford FEL installation, synchronization to the micropulses enables this to be performed on a picosecond timescale. This allows the carrier relaxation dynamics to be studied in a magnetic field, particularly when implemented in a pump-and-probe experiment with a mode-locked Ti:sapphire laser (**Figure 5**).

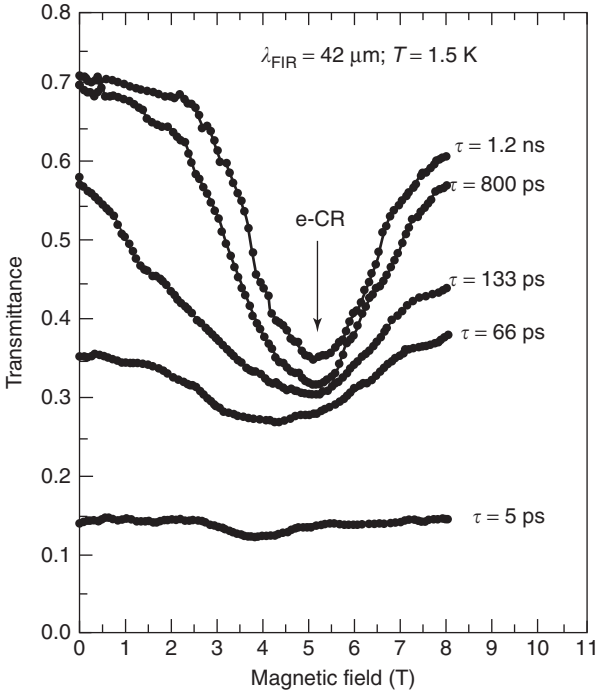


Figure 5 Picosecond time-resolved CR in InSb using micropulses from an FEL. Initially, the very high density of photocreated electrons produces a very wide CR due to carrier scattering, only after ~1 ns equilibrium is restored. (Reprinted with permission from Kono *et al.* (1999) *Applied Physics Letters* 75: 1119–1121.)

Optically Detected Cyclotron Resonance

Rather than monitoring the FIR absorption directly, optically detected cyclotron resonance (ODCR) detects the variation in photoluminescence (PL) excited by a visible laser. ODCR provides several benefits: undoped samples can be studied since the visible laser pump provides carriers, the conduction and valence band can be studied simultaneously, and the photomultiplier tubes (PMTs) or CCDs used to detect PL offer much greater sensitivity than the usual FIR detectors. Typically, the luminescence is dispersed into a PMT, with a monochromator having a sufficiently wide exit slit to record the integrated intensity from a single PL line without the need to adjust for the diamagnetic shift when sweeping the magnetic field, and detected at the chopping frequency (~2 kHz) of the visible laser. This output is fed into a second lock-in amplifier referenced to the more slowly chopped FIR to produce the ODCR signal. ODCR can then either be recorded as a function of the magnetic field or a function of the PL energy at fixed field.

ODCR is clearly only suited to direct bandgap semiconductors that have a strong PL signal. Studies in bulk GaAs show the free-electron CR transitions and also reveal transitions from the ground to excited states of donor-bound impurities (**Figure 6**). Chemical shifts due to different donor atoms also appear with improved resolution over photoconductivity measurements. The mechanism behind ODCR is thought to be that the FIR illumination liberates electrons from donor sites to the conduction band, increasing the impact ionization of bound excitons and so reducing the PL signal, which is predominantly due to excitonic recombination. The free electron-related luminescence makes a much smaller contribution to the PL and has been shown to increase in GaAs, as expected in this impact ionization mechanism, but decrease in other measurements where lattice heating is more important. Charged excitons, or trions labeled X^-, can also be considered as particles with energy quantized into a set of LLs, and CR between these can be used to extract the trion effective mass and lifetime. In quantum wells, the ODCR signal is strongly dependent on the filling factor and, when the cyclotron energy exceeds the inter-subband separation, can show a complete transfer of luminescence intensity from the ground to second subband at resonance.

A development of ODCR is far infra-red-modulated photoluminescence (FIRM-PL) whereby the whole PL spectrum is recorded at a fixed magnetic field, by dispersing the luminescence signal onto a CCD with typically a 1 m spectrometer. Two series of such spectra are taken as the magnetic field is

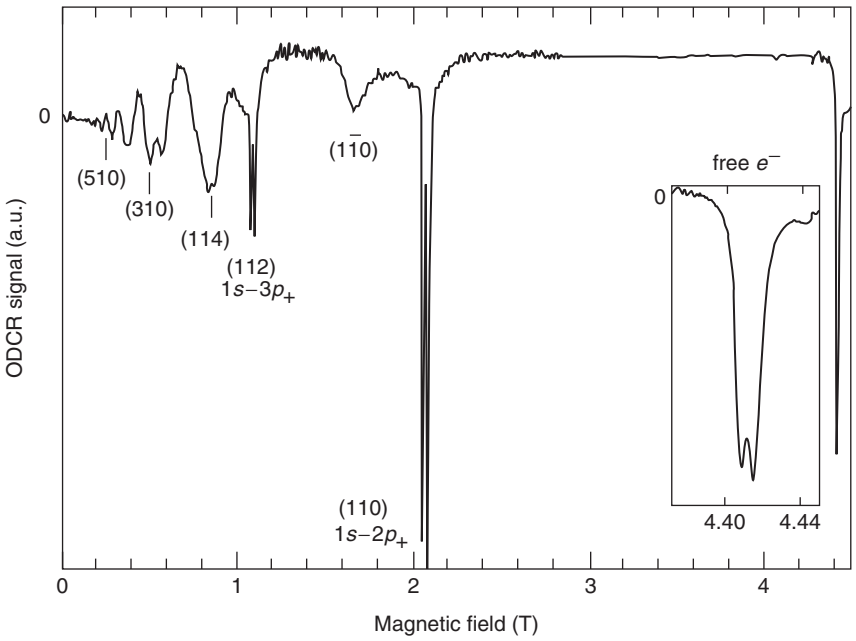


Figure 6 ODCR signal from bulk GaAs for an FIR wavelength of 163 μm, showing free-electron CR at 4.4 T, including a spin-split peak, and excitonic transitions at lower fields. (Reprinted with permission from Michels JG, Warburton RJ, Nicholas RJ, and Stanley CR (1994) An optically detected cyclotron resonance of bulk GaAs. *Semiconductor Science Technology* 9: 198–206; © IOP Publishing.)

stepped, with and without FIR radiation, and the CR contribution for each of the PL lines can be seen in their difference (**Figure 7**). By studying these differences, as functions of magnetic field, PL, and FIR power, a lot can be learned about the energy levels and carrier dynamics within the semiconductor. However, these data have to be interpreted carefully to account for the changes in carrier distribution and increase in nonradiative recombination due to FIR heating.

Effective Mass Measurements

Measurable features of CR include the peak position, amplitude, and line width. The integrated intensity of a CR line can be used to obtain the density of carriers responsible for the absorption, and the line width gives a measure of their mobility. Here, much emphasis will be on the position of the resonance, although discontinuities in position are usually accompanied by increases in line width. In many CR experiments, the position of the resonance is often recorded in terms of an effective mass, which removes a linear background compared to the energy of the transition. For an idealized, isotropic semiconductor with a single parabolic energy band (i.e., $E \propto k^2$), CR would yield the same value of effective mass m^*_{CR} irrespective of the photon energy absorbed or the direction of measurement. However, real materials are not that simple.

Crystalline Anisotropy

Crystalline anisotropy can be defined through a tensor as follows:

$$\left(\frac{1}{m^*_{\mathrm{CR}}}\right)_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j} \qquad [2]$$

which will depend on details of the electronic band structure and the underlying crystal structure. The value measured for m^*_{CR} in a CR experiment is an average over the extremal orbit taken by an electron around the Fermi surface perpendicular to the applied magnetic field. No crystalline structure can be spherically symmetric, so there is an inevitable amount of anisotropy. By rotating a sample about a magnetic field, applied along one of the three principal crystalline axes, this mass anisotropy can be mapped out for each electron and hole band (**Figure 8**). This early CR data on Si and Ge were instrumental in working out the band structure of these, and other materials, and continues to be used as high-quality crystals of new semiconductors, for example, SiC, ZnO, GaN, type IIb diamond, become available. As another example from the narrow-gap layered semiconductors, CR has shown that although their crystal structure is the same as that of Bi_2Se_3, it has a single conduction band minimum at the Γ point whilst Bi_2Te_3 has six ellipsoids, like Si. Further, by substituting some of the Bi by Sb, a material is formed with two distinct conduction bands.

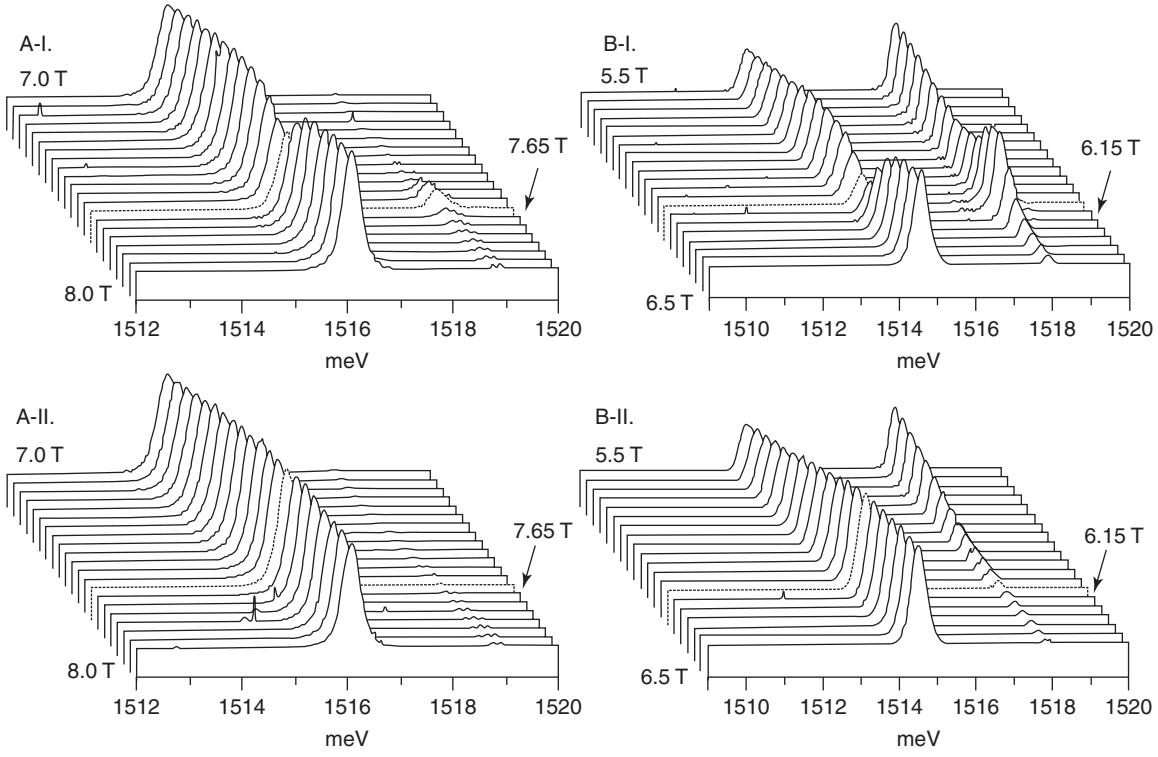


Figure 7 FIRM-PL: PL spectra of a GaAs2DEG with (I) and without (II) FIR radiation at (A) 96.5 μm and (B) 118.8 μm taken over a 1 T range about the CR position, indicated by arrows. Note the transfer of PL intensity from the E_0 to E_1 lines at resonance. (Reprinted with permission from Chang CC and Nicholas RJ (1999) A far infrared modulated photoluminescence FIRM-PL study of cyclotron resonance in a 2D electron gas in GaAs/AlGaAs hetero-junctions. *Semiconductor Science and Technology* 14: 768–774; © IOP Publishing.)

The anisotropy becomes very large in low-dimensional structures formed from semiconductor heterostructures, where there is considerable confinement energy in one (or more) direction(s), and the incorporation of strain into pseudomorphic layers, for example, $Si/Si_{1-x}Ge_x$ or $GaAs/In_xGa_{1-x}As$, adds further anisotropy especially to the valence band where heavy and light hole bands show considerable mixing. CR has been used as a valuable tool for verifying or adjusting band structure calculations, especially for *p*-type quantum wells with many crossing and anticrossing energy levels.

Nonparabolicity

Semiconductors also, inevitably by definition, have a small energy difference between the maximum of the valence band and the conduction band minimum. The perturbation theory shows that the effect of multiple bands is to change the curvature of each as a function of energy, known as nonparabolicity. This clearly affects m^*_{CR} which can be calculated, using a five-band $k \cdot p$ model, as

$$m^*_{CR} = m^*_0 \left\{ 1 - \frac{2K_2}{E_g}([l+1]\hbar\omega_c + \langle T_z \rangle) \right\} \quad [3]$$

where $\langle T_z \rangle$ is the kinetic energy parallel to the magnetic field, E_g is the bandgap, and K_2 is a factor characterizing the nonparabolicity. K_2 is normally negative, so the mass increases with energy. This increase can be observed (**Figure 9**) by changing the FIR frequency, as when the CR transition occurs at higher field, $\hbar\omega_c$ is greater and the final state is higher in the band. Similarly, at fixed frequency, transitions between different LLs may be resolved as distinct lines with, for example, $l = 1$ to 2 occurring 10% higher in field than $l = 0$ to 1. Low-temperature CR measurements show that m^*_{CR} also increases with carrier concentration, due to band filling increasing the energy of the initial state – the Moss–Burnstein shift. For these reasons, it is necessary to extrapolate a series of measurements back to zero frequency and zero carrier concentration if the band edge mass m^*_0 is required.

Measurements at higher temperatures show an apparent increase in mass due to thermal motion contributing to $\langle T_z \rangle$, but this can be corrected by using eqn [3] to show a temperature-independent m^*_0. If lines from different LLs are unresolved, the mass will also appear to increase with temperature as higher LLs contribute more through the Boltzmann factor. Another important contribution to $\langle T_z \rangle$ is

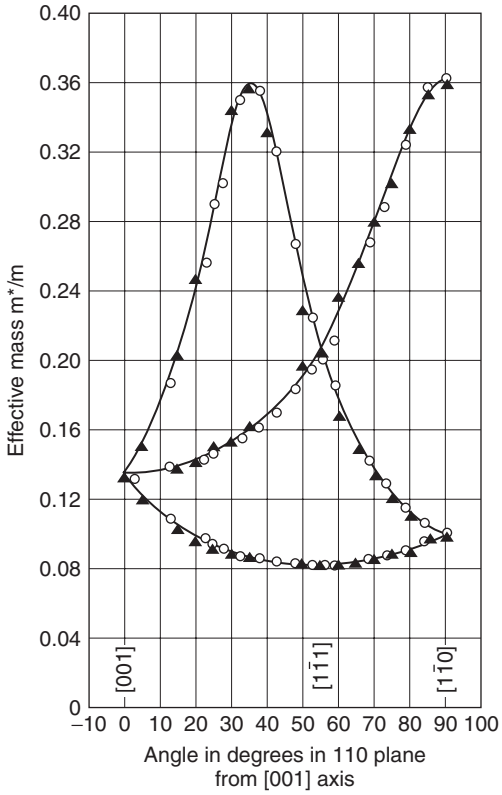


Figure 8 Effective mass of electrons in Ge measured by CR for different crystalline directions. Solid lines are calculations. (Reprinted with permission from Dresselhaus, Kip, and Kittel (1955) *Physical Review* 98: 368–384; © American Physical Society.)

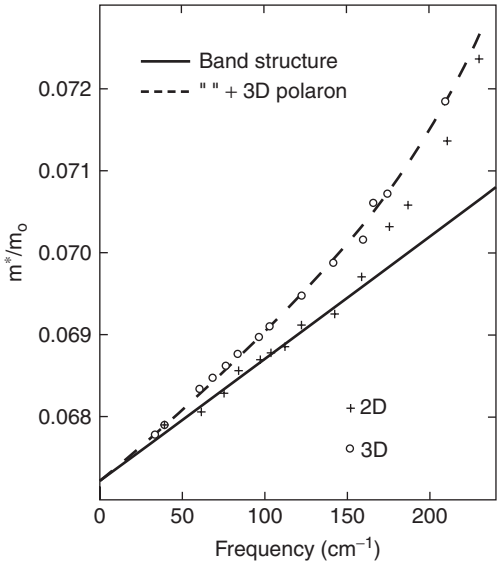


Figure 9 Energy dependence of m^*_{CR} in 3D and 2D GaAs, showing the band structure and polaron contributions to nonparabolicity. (The 3D results are shifted up by $0.0012\,m_0$ for comparison.)

the confinement energy in heterostructures, which generally results in 2D systems having larger m_0^* than the corresponding bulk semiconductor (**Figure 10**).

Polaron Coupling

In polar semiconductors, there is an additional contribution to the nonparabolicity from polaron coupling, of similar magnitude to that from the band structure. When an electron moves, it attracts (repels) the positive (negative) ions and the associated strain field dresses the electron into a composite particle known as a polaron, having a greater effective mass expressed as $m^*_{pol} = m^*(1 + \alpha/6)$, where α is the Fröhlich coupling constant that varies from 0.07 in GaAs to 1.2 in ZnO. This leads to a nonparabolicity contribution of

$$K_2^{pol} = -\frac{3\alpha E_g}{40\hbar\omega_{LO}} \qquad [4]$$

where ω_{LO} is the longitudinal optic phonon energy, which explains the apparent discrepancy between a calculated value of -1.4 for K_2 in GaAs with the -1.8 measured by CR. While this mass enhancement is clearly observed in bulk semiconductors, it can be screened out in a 2DEG at low temperatures, reappearing at higher temperatures where the screening is less efficient as shown in **Figure 10**.

Another way to view the effects of electron–phonon coupling is shown in **Figure 11**. Anticrossing between the second LL and first LL plus an LO phonon opens an energy gap and bends energy levels, even well away from the resonant polaron condition, where $\hbar\omega_c = \hbar\omega_{LO}$. The two branches of this upper level can be traced out using CR at energies above and below the reststrahlen band; however, close to this region, careful interpretation of the results is required to account for the FIR transmission and internal reflections especially in layered heterostructures. Some CR measurements within the opaque reststrahl of GaAs have been made by viewing CR in reflection, whereby the light passed through a thin 2D layer on the way to and back from the highly reflective bulk substrate, and by mounting very thin layers on a transparent substrate. For InSe, $\alpha = 0.3$ which leads to a very large resonant polaron splitting and allows CR lines from both branches to be seen outside the reststrahl over a wide field range.

Besides changing the CR energy, the additional electron–phonon scattering at resonance leads to an increase in the line width. Inter-Landau level transitions can also occur via phonon absorption/emission when the optic phonon energy is an integer multiple of the cyclotron energy – known as magnetophonon resonance (MPR). It is seen at temperatures $\sim$80–200 K, where there is a favorable balance between

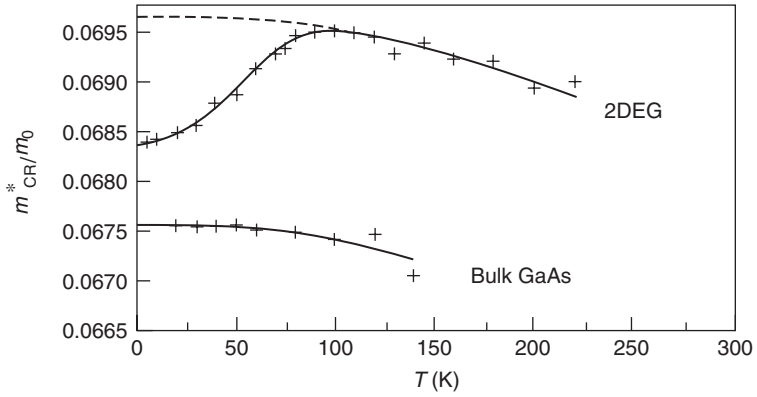


Figure 10 Comparison of the temperature dependence of m^*_{CR} in 3D and 2D GaAs. Temperature-dependent screening of the polaron contribution in 2D is seen by the difference between the dashed extrapolation from high temperature and the data. Note the increase due to confinement in 2D. (Adapted with permission from Hopkins *et al.* (1987) *Physical Review B* 36: 4789; © American Physical Society.)

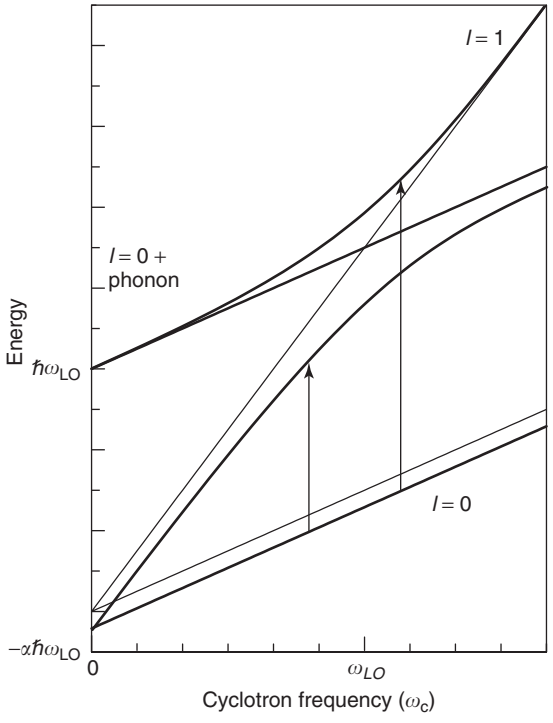


Figure 11 Schematic representation of the polaron coupling between the first two LLs. Dashed lines show the unperturbed levels, arrows are the CR transitions.

phonon population and LL broadening. While MPR is normally observed as peaks in the magnetoresistance, it can also be seen through oscillations in the CR line width.

Shifts and Coupling of the Resonance

It is already seen that the CR position can be altered by the electron–phonon interaction. This section discusses some of the other mechanisms responsible for moving and broadening the absorption lines.

Subband Structure

2D systems may have several electric subbands, due to confinement as distinct from the different symmetries of heavy and light holes. Transitions between LLs of different subbands are usually forbidden, but the presence of a magnetic field component in the 2D plane mixes these states forming resonant subband-Landau-level coupling (RSLC) between, for instance, the $l = 1$ LL of the E_0 subband and the $l = 0$ level of the E_1 subband. Thus, subband energies can be detected in tilted field CR as a discontinuity in m^*_{CR} as a function of energy (or field, as these are usually FTIR experiments), accompanied by an increase in the line width at the RSLC condition (**Figure 4**). This method has been successfully used to elicit the energy level spectra for a wide variety of structures, including heterojunctions, quantum wells, and superlattices. In high-quality samples, the CR splits to reveal transitions to the upper and lower branches, and shows a transfer of oscillator strength from one to the other on passing through the resonance. By contrast to direct inter-subband transitions, which require the light to be polarized perpendicular to the interface and exhibit a significant depolarization shift, in CR, the light is polarized parallel to the interface (even in tilted fields due to the high dielectric constant of most semiconductors) and depolarization shifts are very small.

Shifted CR

The resonance position can, however, be shifted if carriers are bound to a harmonic potential Δ when the CR position is described by $(\hbar\omega)^2 = (\hbar\omega_c)^2 + \Delta^2$. This is the case in *n*-ZnO and *n*-CdS, where the electrons are in very shallow traps (not bound impurity

states), resulting in trapped (not free) electron CR, and can also be seen in quantum wells, where fluctuations in well width lead to carriers being confined to localized islands. For shifted CR to be observed, the localization length needs to be greater than the cyclotron radius, but not so large that the perturbation is the same for the initial and final state.

Landau Level Hybridization

An interesting example of level crossing occurs in InAs/GaSb heterojunctions, where the conduction band minimum in InAs can be below the valence band maximum in GaSb. This leads to a semimetallic system with spatially separated electron and hole gases, unless the structure contains narrow quantum wells (usually of the InAs), where the additional confinement uncrosses the energy gap to produce a semiconducting arrangement. CR can be observed between both electron and hole LLs, and by following their position as a function of magnetic field, the relevant band edge energies can be found. In the semimetallic system, strong oscillations of the CR line width and amplitude have been observed that disappear when the electron and hole layers are separated by a thin barrier ($\sim$2 nm AlSb). These oscillations correspond to line width maxima whenever electron LLs and hole LLs cross, at which point energy gaps arise from hybridization of the electron and hole LLs. The hybridization only happens when the electron and hole wave functions can overlap by tunneling through the interface, which is why a thin barrier is sufficient to remove the effect. Similar hybridization effects have been seen in coupled GaAs quantum wells, forming symmetric and antisymmetric wave functions.

Spin Splitting

As evident from eqn [1], there are separate LL fans for each spin state and, in a nonparabolic band, the spin preserving CR transitions will be at different energies for spin-up and spin-down carriers, which enables the g-factor to be measured. FIR only excites bare electrons in a homogeneous layer (Kohn's theorem), so the g-factor measured in CR is the single-particle value and not the enhanced value detected in transport measurements. While this splitting is small in wide-gap materials, such as GaAs, where $g = -0.44$, and only observed in the quantum limit (see below), spin split CR is readily seen for narrow-gap semiconductors, such as InSb (**Figure 12**), where the proximity of conduction and valence bands leads to strong nonparabolicity as well as a large g-factor. For InAs with $g = -15$, the spin energy gap is so large that at low density and temperature, only the lower spin state is populated and only one CR line

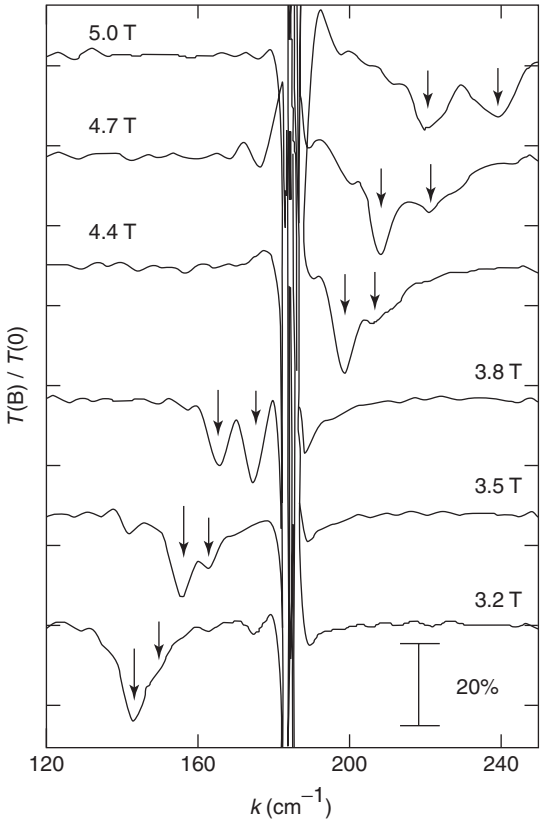


Figure 12 FRIR spectra showing clearly resolved spin splitting in InSb quantum wells on either side of the reststralen band due to absorption of InSb like optic phonons around 180 cm^{-1}. (Reproduced from Khodaparast *et al.* (2004) Spin effects in InSb quantum wells. *Physica E* 20: 386–391, with permission from Elsevier.)

observed, but once the second spin level becomes populated at higher temperature (or density), the second line appears. The onset of the second line can be used as a measure of carrier density.

Dilute magnetic semiconductors, such as $In_{1-x}Mn_xAs$, have been observed to have ferromagnetic phases below $\sim$50 K that are of interest for spintronic applications. The ferromagnetic exchange between Mn ions is believed to be hole mediated and strongly affect both the conduction and valence band parameters due to the narrow bandgap. Megagauss hole CR, in **Figure 13**, shows lines that shift to lower fields and intensify as the ferromagnetic phase is entered, consistent with the additional Mn exchange field increasing the spin splitting. In this regime, $k \cdot p$ calculations show that the spin splitting varies with temperature, magnetic field, and Mn concentration leading to g-factor that varies from -20 to $+100$. Again, CR experiments are vital to validate and refine these calculations.

Although spin–split electron CR is commonly observed in bulk GaAs, it has only been seen in 2DEGs at very low electron densities with $\upsilon < 1/10$.

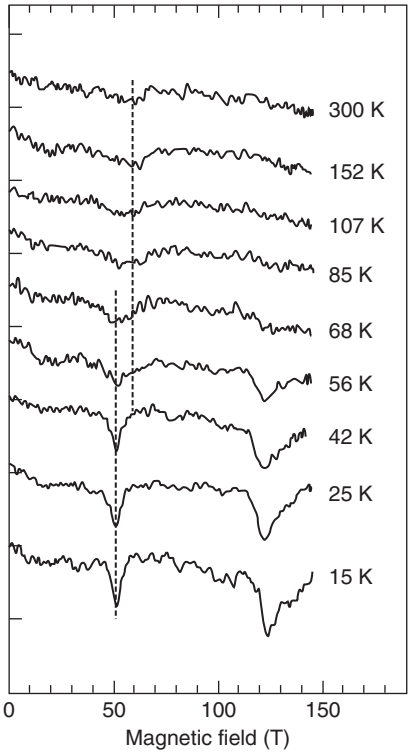


Figure 13 CR spectra at 10.6 μm from a InMnAs/GaSb heterojunction in magnetic fields up to 150 T. On cooling through the ferromagnetic transition at 55 K an additional line appears. (Reproduced from Khodaparast *et al.* (2004) High-field cyclotron resonance studies of InMnAs-based ferromagnetic semiconductor heterostructures. *Physica E* 21: 978–982, with permission from Elsevier.)

At this very low filling factor and 100 mK, only the lowest transition is seen at the position expected in bulk GaAs. The upper spin level becomes thermally populated and two lines appear by 1 K with their relative intensities following the level of populations governed by Boltzmann statistics and separated by the spin splitting expected in the single-particle picture. For $1/10 < \nu < 1/6$, this picture starts to break down with a smaller splitting than expected and for $\nu > 1/6$, through to $\nu = 2$, only one line is seen that shifts from the position of the lower spin transition by up to half the spin splitting as the temperature increases. These results have been explained by a model, whereby the cyclotron motions of the two spin states are coupled by the Coulomb interaction, which depends on the total electron density. The CR spectrum consists of hybridized spin-up/spin-down states with the position of the single peak determined by the relative spin population. Throughout the temperature and field region where fractional quantum Hall effect (FQHE) features are seen in transport experiments, there are no deviations of CR mass or line width showing that CR is not affected by the electron correlations responsible for FQHE.

A similar two-component situation arises in Si where there is a mass splitting between the two- and fourfold degenerate valleys. Again, only a single CR line is observed until the system is considerably perturbed by uniaxial stress. Attempts to explain this rely on electron–electron interactions perturbing the single-particle states.

Summary

The basis of CR in semiconductors has been discussed with an emphasis on the various experimental techniques that can be employed and the information that can be extracted. CR has been applied to a wide variety of different bulk materials and heterostructures, including elemental semiconductors, wide and narrow gap III–Vs, II–VIs, and magnetics. Although only a limited number of examples have been given here, the same principles apply to all materials. Each time better material becomes available, CR reveals more about the band structure, and as experimental facilities improve, the technique can be extended to more imperfect crystals. However, there are still unexplained couplings and anomalies reported that test our understanding of the way electrons interact with radiation in semiconductors.

This article has concentrated on FIR absorption as a measurement technique. Transitions from higher to lower LLs can also produce emission at the CR energy, and indeed, cyclotron resonance laser action was demonstrated in Ge in the 1980s. While this is an important aspect of CR in semiconductors, it is beyond the scope of the present article.

See also: Cyclotron Resonance: Metals; Defect Centers in Insulating Solids, Optical Properties of; Dielectric Function; Effective Masses; Electron–Phonon Interactions and the Response of Polarons; Electronic Structure Calculations: Scattering Methods; Elemental Semiconductors, Electronic States of; Excitons: Theory; Laser Radiation Sources; Magnetoresistance Techniques Applied to Fermi Surfaces; Optical Absorption and Reflectance; Polarons; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Lasers; Semiconductors, Optical and Infrared Properties of; Time-Resolved Optical Spectroscopies; Transport in Two-Dimensional Semiconductors.

PACS: 07.55. – w; 71.35. – y; 71.38. – k; 71.45. – d; 71.70. – d; 72.10. – d; 72.20. – l; 73.20. + q; 78.20. – e; 78.30. – j

Further Reading

Brummell MA, Leadley DR, Hopkins MA, and Nicholas RJ (1988) Cyclotron and magnetophonon resonance in semiconductor heterostructures. *Physica Scripta* T23: 73–81.

Cooper NR and Chalker JT (1994) Theory of spin-split cyclotron resonance in the extreme quantum limit. *Physical Review Letters* 72: 2057.

Devreese JT (1989) Polaron physics in 2D and 3D. *Physica Scripta* T25: 309–315.

Hopkins MA, Nicholas RJ, Brummell MA, Foxon CT, and Harris JJ (1987) Cyclotron resonance study of non-parabolicity and screening in GaAs-GaAlAs heterojunctions. *Physical Review B* 36: 4789.

Landwehr G and Rashba EI (eds.) (1991) Landau level spectroscopy. *Modern Problems in Condensed Matter Sciences*, vol. 27, Amsterdam: North-Holland.

Lax B (1958) Experimental investigations of the electronic band structure of solids. *Review of Modern Physics* 30: 122–154.

Mayer H and Rössler U (1991) Spin splitting and anisotropy of cyclotron resonance in the conduction band of GaAs. *Physical Review B* 44: 9048–9051.

Michels JG, Daly MS, Gee P, Hill S, Nicholas RJ, *et al.* (1996) Cyclotron resonance and spin states in GaAs/GaAlAs heterojunctions: experiment and theory. *Physical Review B* 54: 13807–13815.

Miura N, Osada T, and Takeyama S (2003) Research in super-high pulsed magnetic fields. *Journal of Low Temperature Physics* 133: 139–158.

Nicholas RJ, Hopkins MA, Brummell MA, and Leadley DR (1988) Polaron effects and optic phonon scattering in heterostructures. In: Leavens AR and Taylor R (eds.) *Interfaces, Quantum Wells and Superlattices, NATO ASI Series*, pp. 243–269. New York: Plenum.

Peeters FM (2004) Theory of electron–phonon interactions in semiconductors. In: Herlach F and Miura N (eds.) *High Magnetic Fields: Science and Technology*, vol. 2, pp. 23–45. New Jersey: World Scientific.

Zawadzki W and Pfeffer P (2004) Spin splitting of subband energies due to inversion asymmetry in semiconductor heterostructures. *Semiconductor Science and Technology* 19: R1–R17.

Nomenclature

B	magnetic field (T)
E_g	semiconductor energy gap (eV)
E_0, E_1	energy of confined subband (eV)
$E(k)$	electronic energy as a function of wave vector (eV)
g	Landé g-factor (dimensionless)
k	FIR or electronic wave vector (m^{-1})
$k_{\|\|}$	wave vector in a direction parallel to the magnetic field (m^{-1})
K_2	nonparabolicity factor (dimensionless)
K_2^{pol}	polaron nonparabolicity factor (dimensionless)
l	integer Landau level index (dimensionless)
m_0	free-electron mass (kg)
m^*	effective mass (electron or hole, unspecified) (kg)
m^*_{CR}	measured cyclotron resonance effective mass (kg)
m^*_0	band edge effective mass (kg)
$\langle T_z \rangle$	average electron kinetic energy parallel to magnetic field (eV)
α	Fröhlich coupling constant (dimensionless)
μ	carrier mobility ($cm^2\,V^{-1}\,s^{-1}$)
ν	Landau level filling factor (dimensionless)
τ	carrier scattering time (ps)
ω	frequency of radiation ($rad\,s^{-1}$)
ω_c	cyclotron frequency ($rad\,s^{-1}$)
ω_{LO}	longitudinal optic phonon frequency ($rad\,s^{-1}$)

D

Deep-Level Spectroscopy and Auger Spectra

R Colle, Università di Bologna, Bologna, Italy
S Simonucci, Università di Camerino, Camerino, Italy

Introduction

"Spectroscopy" can be defined as the study of processes that, in a very general and schematic way, can be described as follows

$$P(|i;\varepsilon,\boldsymbol{p},\alpha>)+T(|I;E,\boldsymbol{\pi},\beta>)\rightarrow \tilde{T}(|F;\tilde{E},\tilde{\boldsymbol{\pi}},\beta>)+\tilde{P}(|f;\boldsymbol{\varepsilon},\tilde{\boldsymbol{p}},\tilde{\boldsymbol{\alpha}}>)+\cdots$$

where P is a projectile (e.g., a photon, an electron, a proton, or a heavier particle) in a given initial state $|i\rangle$ (characterized by its energy ε, momentum $\boldsymbol{p}$, and/or quantum numbers α), and T is a target (e.g., an atom, or a molecule, or a condensed system) in a given initial state $|I\rangle$ with energy E, momentum $\boldsymbol{\pi}$, and quantum numbers β. Products of this generalized collision that conserves energy and momenta can be the same colliding species in different states, or different species resulting from the fragmentation and/or annihilation of the projectile and/or the target.

In greater detail, deep-level spectroscopy is a specific branch of spectroscopy that studies excitations or ionizations of the inner-shell (deep level) electrons of atoms, molecules, and condensed systems, produced by impact with X-ray photons, high-energy electrons, or (less frequently) heavy particles. Different emission products can be collected and lead to different experimental methods: electron spectrometry, fluorescence spectrometry, ion spectrometry, and their combinations in coincidence measurements, which completely define the full emission pattern following the collision.

A specific class of experiments is that in which the particles selected by the detector are electrons of energies characteristic of the target and, in good approximation, independent of the way in which the system has been ionized or excited. These electrons are called Auger electrons in the name of the French physicist who, in 1924, using a Wilson chamber to study X-ray photoionization of atoms, observed the appearance of fast electrons with energy dependent on the photon energy, and slow electrons with energy independent of that, but characteristic of the emitting atom. Auger also observed that each pair of fast (photo-)electron and slow (Auger-)electron originated from the same point (atom), but the directions of emission of the two electrons were random and mutually independent; furthermore, photoelectrons not accompanied by an Auger electron were also present.

The typical experimental apparatus used for deep-level spectroscopies consists of a source of projectiles prepared with specific characteristics and ejected through a collimated beam into the reaction volume where collisions between particles and targets take place. Specific reaction products are then accepted by a detector in the direction of its finite entrance slit and counted in terms of their kinetic energy (or momentum, or other properties) thus providing intensity distributions, for example, $I_\lambda(E_{\text{kin}};\hat{k})$, that represents the energy spectrum of the particles λ ejected along the direction $\hat{k}$. Global spectra, independent of the direction of emission, are then obtained by integrating these intensity distributions, that is, by collecting particles from all the directions of emission. In coincidence measurements and in more sophisticated experiments, several detectors are used simultaneously to obtain spectra that plot conditional intensities, for example, the probability of finding two or more particles with selected energies along specific directions.

The essential parameters that characterize these experiments are

- properties and preparation of the incident projectiles, for example, in the case of photons: energy (selected by a monochromator), intensity, and polarization;
- targets of the experiment, that is, observables of interest and dynamics of these observables induced by the collision; for example, in the case of atomic systems: energy levels, transition rates, angular distributions of ejected fragments, cross sections of fragmentation processes, and so on; and

- position and properties of the detector(s), that is, acceptance solid angle, transmission, resolution, and efficiency.

The experiment can be performed at a different level of accuracy, meaning with a different resolution: in an "ideal experiment," the target and projectile are prepared in a well-defined initial state and the collision products are collected in coincidence and completely characterized in their final states. Optimal resolution has been significantly approached in recent experiments of electron spectrometry on light atoms and on simple molecules, excited or ionized by monochromatized energy-tunable synchrotron radiation.

The theoretical analysis of these experiments is based on the well established and formally developed branch of quantum mechanics known as "quantum collision or scattering theory." Its effective implementation for a quantitative study of real experiments, however, is quite difficult and becomes progressively less accurate and more phenomenological as the number of particles in the system increases.

This article considers photoprocesses in atomic systems produced by the absorption of a monochromatized (λ) energy-tunable ($\hbar\omega$) radiation in the X-ray region, that is, 0.1 Å $< \lambda <$ 100 Å; 128 keV $> \hbar\omega >$ 128 eV. These processes provide a unique opportunity to explore the dynamics of the many-body interactions in quantum systems, since the radiation–matter interaction, which changes the internal structure of the target, has the special advantage of being known exactly.

The quantum theory of photoionization is summarized briefly.

Theory of Photoionization

Using the first-order time-dependent perturbation theory and dipole approximation, one can prove that the correct quantum-mechanical transition rate (probability per unit time) of a photoionization process, in which an atomic system in its ground state $|\Psi_o\rangle$ is promoted to a final state $|\Psi^-_{\alpha k}\rangle$ by an incident radiation field of frequency ω and polarization direction $\hat{\varepsilon}$ inside a volume V, can be expressed by

$$\frac{dW_{o\to\alpha}}{d\boldsymbol{k}} = \frac{2\pi}{\hbar}|\langle\Psi_o|\hat{H}_{int}|\Psi^-_{\alpha k}\rangle|^2 \times \delta(\hbar\omega + E_o - E)$$

$$\hat{H}_{int} = \left(\frac{2\pi\hbar\omega}{V}\right)^{1/2} \times e^{-i\omega t}\hat{\varepsilon} \times \sum_j q_j \boldsymbol{r}_j$$

where $\{q_j\}$ are the charges of the particles in the target, $E = E_\alpha + \varepsilon_k$ is the energy of the scattering state, and E_o that of the ground state. In this expression, $\Psi^-_{\alpha k}$ is a scattering wave function with normalization $\langle\Psi^-_{\alpha k}|\Psi^-_{\beta p}\rangle = \delta_{\alpha\beta}\delta(\boldsymbol{k}-\boldsymbol{p})$ and incoming-wave boundary condition (–); the latter assures the appropriate asymptotic behavior of the wave function for a process in which a particle is ejected from the target along a specific direction $\hat{k}$. The wave function $\Psi^-_{\alpha k}$ represents the final state of the process with one electron ejected with momentum $\boldsymbol{k}$ and energy $\varepsilon_k = \hbar^2k^2/2m$, and the ionized target left in a state $|\alpha, E_\alpha\rangle$. Upon dividing the transition rate by the incident photon current density c/V, one gets the (triple) differential cross section of the $|o\rangle \to |\alpha, \boldsymbol{k}\rangle$ process

$$\frac{d\sigma_{o\to\alpha}}{d\boldsymbol{k}} = \left(\frac{4\pi^2\omega}{c}\right)|\langle\Psi_0|\hat{O}_{\hat{\varepsilon}}|\Psi^-_{\alpha k}\rangle|^2 \times \delta(\hbar\omega + E_o - E)$$

where $\hat{O}_{\hat{\varepsilon}} = \hat{\varepsilon}\cdot\sum_j q_j \boldsymbol{r}_j$ is the dipole operator. Note that, if the state $|\alpha, E_\alpha\rangle$ of the ionized target is an excited state with finite lifetime $\tau = \hbar/\Gamma_\alpha$, the δ-function in the above expression is substituted by a Lorentzian function L_{Γ_α} such that

$$\lim_{\Gamma_\alpha\to 0}\frac{\Gamma_\alpha/2\pi}{(\hbar\omega + E_o - E)^2 + (\Gamma_\alpha/2)^2} = \delta(\hbar\omega + E_o - E)$$

The above cross section gives the angular distribution of a photoelectron ejected from the target along a direction $\hat{k}$ defined in the target frame. The corresponding direction of emission in the laboratory frame can be easily obtained if one knows the target orientation. If, however, the irradiated system consists of an ensemble of randomly oriented targets, the theory shows, in agreement with the experiment, that the measured angular distribution of the photoelectrons ejected along a given direction $\hat{r}$, defined with respect to the quantization axis $\hat{z}$ (which, for linearly polarized light, is the same as the electric field direction,) is given by the simple expression

$$\frac{d\sigma_{o\to\alpha}}{d\hat{r}} = \frac{\sigma_{o\to\alpha}}{4\pi}[1 + \beta\, P_2(\cos\theta)],$$
$$\cos\theta = \hat{r}\cdot\hat{z}$$

Here $\sigma_{o\to\alpha}$ is the total cross section of the process, $P_2(\cos\theta) = (1.5\cos^2\theta - 0.5)$ is the second Legendre polynomial, and β is called the angular distribution or anisotropy parameter.

If the X-ray photon absorption gives rise to a metastable state $|\Phi_r\rangle$ embedded in the continua of M decay channels $\{|\psi^-_{\alpha k}\rangle\}$, the appropriate scattering wave function for the final state of the process is given by Fano's theory of the interaction between

discrete and continuous states

$$|\Psi^-_{\alpha\boldsymbol{k}}\rangle = |\psi^-_{\alpha\boldsymbol{k}}\rangle + |\Phi^-_r\rangle \frac{\langle\Phi_r|\hat{H}-E|\psi^-_{\alpha\boldsymbol{k}}\rangle}{E-E_r-\mathrm{i}\Gamma_r/2}$$

$$|\Phi^-_r\rangle = \Phi_r\rangle + \sum_\beta \int \frac{|\psi^-_{\beta\boldsymbol{p}}\rangle \mathrm{d}\boldsymbol{p} \langle\psi^-_{\beta\boldsymbol{p}}|\hat{H}-E|\Phi_r\rangle}{E-(E_\beta+\varepsilon_p)-\mathrm{i}0^+}$$

where $\{E_\beta\}$ are the energies of the ionized target and E_r that of the metastable state, whose natural width is Γ_r. The cross section of the process is, therefore, given by

$$\frac{\mathrm{d}\sigma_{\mathrm{o}\to\alpha}}{\mathrm{d}k} = \left(\frac{4\pi^2\omega}{c}\right) \left| \langle\Psi_\mathrm{o}|\hat{O}_{\hat{\varepsilon}}|\psi^-_{\alpha\boldsymbol{k}}\rangle + \langle\Psi_\mathrm{o}\left|\hat{O}_{\hat{\varepsilon}}\right|\Phi^-_r\rangle \frac{\langle\Phi_r|\hat{H}-E|\Psi^-_{\alpha\boldsymbol{k}}\rangle}{E-E_r-\mathrm{i}\Gamma_r/2} \right|^2 \times \delta(\hbar\omega+E_\mathrm{o}-E)$$

and consists of three contributions: that due to the direct photoionization (square modulus of the first amplitude), that due to the two-step process (square modulus of the second term) with a characteristic Lorentzian profile, and an interference term which can distort the Lorentzian profile asymmetrically.

For two-step photoionization processes (emission of a photoelectron with momentum $\boldsymbol{p}$ and energy ε_p, followed by Auger decay with an electron ejected with momentum $\boldsymbol{k}$ and energy ε_k), one can approximate the scattering wave function with the product

$$|\Theta^-_{\alpha\boldsymbol{k}\boldsymbol{p}}(E)\rangle \simeq |\Psi^-_{\alpha\boldsymbol{k}}(E_p)\rangle |\eta_{\boldsymbol{p}}(\varepsilon_p)\rangle$$

where $\Psi^-_{\alpha\boldsymbol{k}}$ is the wave function for the decay process, $\eta_{\boldsymbol{p}}(\varepsilon_p)$ is the spin orbital of the ejected photoelectron and $E = E_\alpha + \varepsilon_k + \varepsilon_p = E_p + \varepsilon_p$ is the energy of the stationary state. The (sixfold) differential cross section, that gives the conditional probability of collecting the photoelectron with momentum $\boldsymbol{p}$ and the Auger electron with momentum $\boldsymbol{k}$ simultaneously, is given by

$$\frac{\mathrm{d}^2\sigma_{\mathrm{o}\to\alpha}}{\mathrm{d}\boldsymbol{k}\,\mathrm{d}\boldsymbol{p}} = \left(\frac{4\pi^2\omega}{c}\right) |\langle\Psi_\mathrm{o}|\hat{O}_{\hat{\varepsilon}}|\Theta^-_{\alpha\boldsymbol{k}\boldsymbol{p}}\rangle|^2 \delta(\hbar\omega-\Delta) \propto \frac{|\langle\Psi_\mathrm{o}|\hat{O}_{\hat{\varepsilon}}|\Phi^-_r;\eta_{\boldsymbol{p}}\rangle\langle\Phi_r|\hat{H}-E_p|\Psi^-_{\alpha\boldsymbol{k}}\rangle|^2}{(E_p-E_r)^2+\Gamma_r^2/4} \times \delta(\hbar\omega-\Delta)$$

where $\Delta = E - E_\mathrm{o}$ and the contribution of the direct double photoionization has been neglected. In Auger spectroscopy, only electrons produced by the decay process are recorded; therefore, the cross section relevant to an experiment, in which they are collected from all the directions and selected in terms of their kinetic energy ε, is given by

$$\frac{\mathrm{d}\sigma_{\mathrm{o}\to\alpha}}{\mathrm{d}\varepsilon} = \int \delta(\varepsilon-\varepsilon_k)\mathrm{d}\boldsymbol{k} \int \frac{\mathrm{d}^2\sigma_{\mathrm{o}\to\alpha}}{\mathrm{d}\boldsymbol{k}\,\mathrm{d}\boldsymbol{p}}\,\mathrm{d}\boldsymbol{p} \propto \frac{\mathscr{D}_r(\omega,\varepsilon)\mathscr{M}_{r\alpha}(\varepsilon)}{(\varepsilon_\mathrm{A}-\varepsilon)^2+\Gamma_r^2/4}$$

where $\varepsilon_\mathrm{A} = E_r - E_\alpha$ is the nominal Auger energy of the ejected electrons and

$$\mathscr{D}_r(\omega,\varepsilon) = \int |\langle\Psi_0|\hat{O}_{\hat{\varepsilon}}|\Phi^-_r;\eta_{\boldsymbol{p}}\rangle|^2 \delta(\hbar\omega-\Delta)\,\mathrm{d}\boldsymbol{p}$$

$$\mathscr{M}_{r\alpha}(\varepsilon) = \int |\langle\Phi_r|\hat{H}-E_p|\Psi^-_{\alpha\boldsymbol{k}}\rangle|^2 \delta(\varepsilon-\varepsilon_k)\,\mathrm{d}\boldsymbol{k}$$

The structure of this cross section shows that, in the two-step approximation, the Auger lines have a Lorentzian shape of width equal to the natural width of the core-ionized state. Note that, in this approximation, the angular distribution of Auger electrons ejected from an ensemble of randomly oriented targets has exactly the same structure as that given above for photoelectron emission. If, however, the decay process is not clearly separate from the primary photoionization, the scattering wave function $\Theta^-_{\alpha\boldsymbol{k}\boldsymbol{p}}$ cannot be approximated by a simple product, and antisymmetrization and correlation contributions have to be included.

Photoprocesses in Atomic Systems

Since, in the independent particle model of an atomic system, each electron is described by a single-particle function (called spin–orbital and corresponding to a specific energy level) and the dipole operator is a single-particle operator, one expects that the main photoprocesses are one-electron processes, that is, one-electron excitations with discrete spectrum and one-electron ionizations with continuous spectrum. However, since the electronic motion is correlated by the Coulomb interaction potential:

$$\hat{V}_\mathrm{c} = \frac{1}{4\pi\varepsilon_0} \sum_{i<j} \frac{e_0^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}$$

many-electron processes induced by the photointeraction are also possible. Among them, the most important ones are two-electron processes, that are classified as satellite or direct processes when two electrons are simultaneously excited or ionized or one is excited and the other ionized, and two-step processes when the primary excitation or ionization of an inner-shell electron is followed by radiative decay (fluorescence) or nonradiative (autoionization

or Auger) decay. It is clear that, in the case of two-step processes, the final state can also be reached through a direct path, which can compete and interfere with the two-step path modifying the spectral profiles appreciably. In this case, the distinction between direct and two-step processes breaks down and a unified description of the global process is necessary.

These processes are separately considered.

Direct Photoionizations

A direct photoionization is a one-step process in which a monochromatized X-ray photon of energy $\hbar\omega$ is absorbed by the target and leads to the emission of a photoelectron of energy $\varepsilon = \hbar\omega - E_I$, if $\hbar\omega$ is greater than the ionization energy (E_I) of the system and the recoil energy of the target (due to conservation of the total momentum) is negligible. This radiation–matter interaction mode is called photoelectric effect and the resulting structures in the energy spectrum of the ejected photoelectrons are called main photolines, if produced by a simple ionization, and satellites or discrete photolines, if produced by processes in which one electron is ionized and another one simultaneously excited. The continuous energy distribution due to direct double ionizations, in which two ejected electrons share the available energy, are called continuous satellites.

The photoelectron spectra show the dependence of the electron yield, that is, of the number of electrons recorded within a certain time interval at a given intensity of the incident light, on the measured kinetic energy ε or on the modulus of the binding energy $|E_B| = E_I = \hbar\omega - \varepsilon$ of the ejected electrons, or on the photon energy (see **Figure 1**). In these spectra, one observes discrete structures (peaks) at specific energy positions, with different height and width, two features that are related to the intensity (probability) of the specific process. The theory shows that the appropriate measure of the intensity is the area under the photoline, a number which strongly depends on the lineshape. This shape is the result of three contributions: (1) the energy distribution of the incident radiation, usually described by a Gaussian function G_B, centered at the nominal photon energy and with full width at half maximum (FWHM) equal to the "bandpass" (selected energy interval) of the monochromator; (2) the energy distribution of the core-hole state, that is described by a Lorentzian function L_Γ, whose FWHM Γ is called the natural or inherent width of the state; and (3) the spectrometer function, usually approximated by a Gaussian function G_{sp}, centered at a nominal kinetic energy ε_0 and with FWHM equal to the "pass-energy" of the spectrometer. The convolution of these three functions yields the experimental distribution function $F_{exp}(\varepsilon) = G_{sp} \otimes L_\Gamma \otimes G_B$ of photoelectrons detected at a given energy ε. The measured intensity along a given direction $I_{exp}(\varepsilon, \hat{k})$ is the area under the photoline recorded within a certain time interval, in which N_{ph} photons per second hit a sample of targets of density ρ. This intensity, corrected for the efficiency and luminosity of the spectrometer (S), is given by the

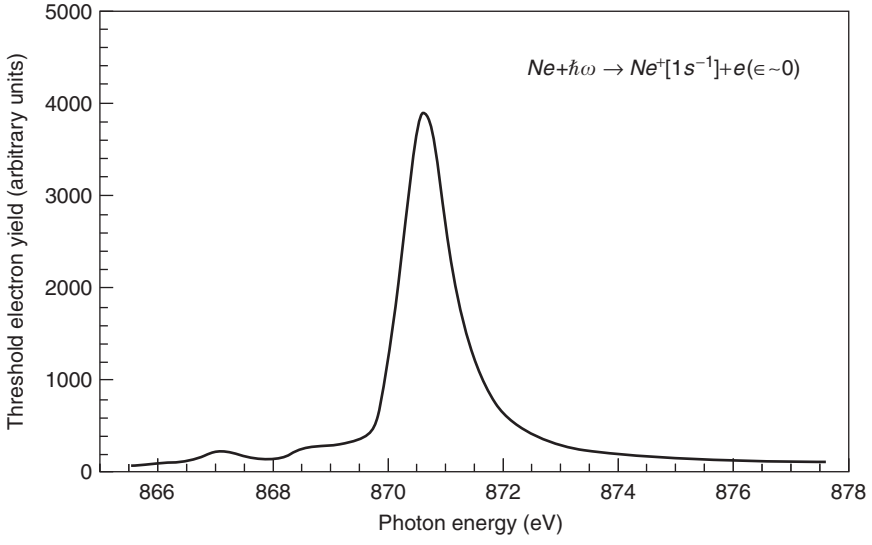


Figure 1 Threshold photoelectron spectrum in the region of Ne[$1s^{-1}$] ionization. The satellites on the low-energy side of the spectrum are due to the decay of metastable Ne[$1s\,2s^2 2p^6 np, n = 3, 4, \ldots$] states; the broadening on the high-energy side is due to post-collision interaction (PCI) between the slow photoelectron and the faster Auger electron emitted in the decay of the inner hole (see text). (Adapted from Avaldi L, Dawber G, *et al.* (1995) Measurement of Ne 1s and 1s nl satellite photoelectron spectra near threshold. *Physical Review A* 51: 5025.)

expression

$$I_{\exp}(\varepsilon, \boldsymbol{k}) = N_{\text{ph}}\ \rho S F_{\exp}(\varepsilon)\frac{\mathrm{d}\sigma}{\mathrm{d}\boldsymbol{k}}$$

where $\mathrm{d}\sigma/\mathrm{d}\boldsymbol{k}$ is the differential cross section of the process, which is independent of the specific features of the experiment and can be predicted *ab initio* by the theory. Note that X-ray photons interact with matter not only through photoelectric effect, but also through Compton and coherent scattering. In the Compton (incoherent) scattering, one photon collides with the target losing part of its energy and deflecting from the original direction. Results of this collision are electrons ejected with lower energies and a broad spectral distribution of scattered photons (Compton profile), related to the momentum distribution of the electrons in the target. Compton scattering is important especially at high X-ray energies and with elements of low atomic number. In the coherent (or elastic) scattering, instead, one photon is scattered without excitation or ionization of the target. This process occurs mostly at low energies and with elements of high atomic number.

Decay Processes

The absorption of a monochromatized X-ray photon can lead to the creation of a core-excited or a core-ionized state, which is metastable and spontaneously decays with a characteristic lifetime into a lower-energy state. In particular, when the frequency of the incident X-ray photon is tuned below or closely above a core-ionization threshold, core-excitations take place and the process is called resonant X-ray scattering or X-ray Raman scattering. The decay process can follow either a radiative path (i.e., emission of X-ray photons of energies $\{\hbar\omega\}$ and intensities $\{I_{\text{R}}(\hbar\omega)\}$ that constitute the fluorescence spectrum) or a nonradiative path (i.e., autoionization (after core-excitation) and Auger decay (after core-ionization) with emission of Auger electrons of characteristic energies $\{\varepsilon_j^{\text{A}}\}$ and intensities $\{I_{\text{A}}(\varepsilon_j^{\text{A}})\}$ that constitute the Auger spectrum). The global decay probability can be partitioned into a fluorescence yield: $\omega_{\text{R}} = I_{\text{R}}/(I_{\text{R}} + I_{\text{A}})$ and an Auger yield: $\omega_{\text{A}} = I_{\text{A}}/(I_{\text{R}} + I_{\text{A}})$, where $I_{\text{R}} = \sum_{\omega} I_{\text{R}}(\hbar\omega)$ and $I_{\text{A}} = \sum_j I_{\text{A}}(\varepsilon_j^{\text{A}})$ are the intensities of the radiative (R) and nonradiative (A) decay channels. One can predict from the theory and observe experimentally that, for atomic numbers $Z<30$, the Auger yield is bigger than the fluorescence yield, while for $Z>30$ the radiative decay becomes dominant.

The decay processes correspond to electronic transitions between levels that, in atoms, are classified according to their quantum numbers $(n, l,\ j = l \pm 1)$. These levels are grouped into shells and subshells that, in spectroscopic notation, are

- K-shell $(n = 1,\ l = 0,\quad j = 1/2|N_{\text{max}}^{\text{el}} = 2)$, the most tightly bound shell which contains a maximum number of electrons $N_{\text{max}}^{\text{el}} = 2n^2 = 2$;
- L-shell $(n = 2|N_{\text{max}}^{\text{el}} = 8)$, the next most tightly bound shell which consists of three subshells: $L_{\text{I}}(n = 2,\ l = 0,\ j = 1/2|N_{\text{max}}^{\text{el}} = 2j + 1 = 2)$, $L_{\text{II}}(n = 2,\ l = 1,\ j = 1/2|N_{\text{max}}^{\text{el}} = 2)$, $L_{\text{III}}(n = 2,\ l = 1,\ j = 3/2|N_{\text{max}}^{\text{el}} = 4)$;
- M-shell $(n = 3|N_{\text{max}}^{\text{el}} = 18)$ with five subshells: $M_{\text{I}}(n = 3,\quad l = 0, j = 1/2|N_{\text{max}}^{\text{el}} = 2)$, $M_{\text{II}}(n = 3,\ l = 1,\quad j = 1/2|N_{\text{max}}^{\text{el}} = 2)$, $M_{\text{III}}(n = 3, l = 1, j = 3/2|N_{\text{max}}^{\text{el}} = 4)$, $M_{\text{IV}}(n = 3,\ l = 2,\ j = 3/2|N_{\text{max}}^{\text{el}} = 4)$, $M_{\text{V}}(n = 3, l = 2, j = 5/2|N_{\text{max}}^{\text{el}} = 6)$;

and so on, until the last natural element, uranium $(Z = 92)$, with the last two electrons in the subshell O_{VIII}.

Fluorescence spectra The radiative decay of metastable states gives rise to emission of X-ray series, that are classified according to the shell that is filled and to the relative intensity of the emitted radiation. Thus, for example, the K-series corresponds to the filling of the K-shell and one distinguishes between the more intense K_α-rays produced by $L \rightarrow K$ transitions: $K_{\alpha_1}(L_{\text{III}} \rightarrow K)$ and $K_{\alpha_2}(L_{\text{II}} \rightarrow K)$, and the less intense K_β-rays produced by $M,N \rightarrow K$ transitions: $K_{\beta_1}(M_{\text{III}} \rightarrow K)$, $K_{\beta_2}(N_{\text{II,III}} \rightarrow K)$, $K_{\beta_3}(M_{\text{II}} \rightarrow K)$, $K_{\beta_4}(N_{\text{IV,V}} \rightarrow K)$, $K_{\beta_5}(M_{\text{IV,V}} \rightarrow K)$. The L-series corresponds to the filling of vacancies in the L-subshells and one distinguishes between the more intense $L_{\alpha 1}$- and $L_{\alpha 2}$-rays produced by $M_{\text{V,IV}} \rightarrow L_{\text{III}}$ transitions, and the less intense L_β-rays produced by $M,N,O \rightarrow L$ transitions: $L_{\beta_1}(M_{\text{IV}} \rightarrow L_{\text{II}})$, $L_{\beta_2}(N_{\text{V}} \rightarrow L_{\text{III}})$, $L_{\beta_3}(M_{\text{III}} \rightarrow L_{\text{I}})$, $L_{\beta_4}(M_{\text{II}} \rightarrow L_{\text{I}})$, $L_{\beta_5}(O_{\text{IV,III}} \rightarrow L_{\text{III}})$.

In addition to the main transitions, there are also satellite lines that are due to the radiative decay of two- and three-hole states produced in the primary photoionization, for example, $KL \rightarrow LL$ and $KLL \rightarrow LLL$ transitions. An example of fluorescence spectrum is shown in **Figure 2**, where the intensity of the X-rays produced by main and satellite transitions following 1*s* ionization of Ne is plotted as a function of the energy of the emitted radiation.

Auger spectra The nonradiative decay of core-ionized states gives rise to Auger spectra in which the electrons are detected with energies characteristic of the target and mostly independent of the photon energy. The transitions associated to the spectral lines are called normal (or diagram) Auger if the initial state of the decay is a core-hole state, produced by

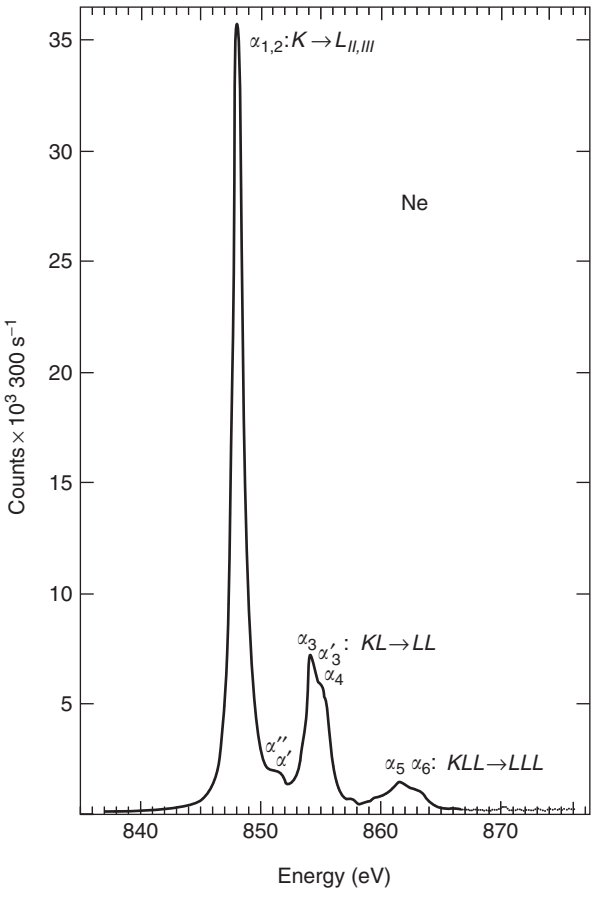


Figure 2 Fluorescence spectrum following 1*s* ionization of Ne: the main line is produced by the superposition of the $K \rightarrow L_{II,III}$ transitions; the less intense lines on the high-energy side of the spectrum are due to satellite transitions. (Adapted from *Physica Scripta* 7: 173 (1973).)

the absorption of a photon whose energy is higher than the core-ionization threshold. The emitted electrons have nominal energies $\{\varepsilon_j^{\mathrm{A}} = E_{\mathrm{i}}(1\mathrm{h}) - E_{f_j}(2\mathrm{h})\}$ resulting from the difference between the energy of the initial (i) one-hole (1 h) state and the energy of each (j) two-hole (2 h) final (f) state of the target. The Auger transitions are usually classified as main (or diagram) transitions, if the initial hole is in the *K*-shell, for example, $K \rightarrow L_1L_2$ in Ne, and Coster–Kronig and super Coster–Kronig transitions, if the final state has, respectively, one or two holes in the same shell in which the initial state has its hole. Examples of Coster–Kronig and super Coster–Kronig transitions are, respectively, the $L_{\mathrm{I}} \rightarrow L_{\mathrm{II,III}}\, M_{\mathrm{I}}$ transitions in Mg and the $N_{\mathrm{II,III}} \rightarrow N_{\mathrm{IV,V}}\, N_{\mathrm{IV,V}}$ transitions in Xe.

The shape of a typical Auger line, (i.e., that produced by a decay process completely separated from the primary photoionization (two-step process)) is obtained from the convolution of two Lorentzian functions $L_{\Gamma \mathrm{i}}$ and $L_{\Gamma \mathrm{f}}$ (that give the energy distributions of the initial (i) corehole and the final (f) two-hole state of the process) with the spectrometer function G_{sp} centered at a nominal kinetic energy. This convolution yields the experimental distribution function $F_{\mathrm{A}}^{\mathrm{fi}}(\varepsilon) = G_{\mathrm{sp}} \otimes L_{\Gamma_{\mathrm{f}}} \otimes L_{\Gamma_{\mathrm{i}}} = G_{\mathrm{sp}} \otimes L_{\Gamma_{\mathrm{fi}}}$ of the Auger electrons recorded at a given energy ε. In this expression $L_{\Gamma_{\mathrm{fi}}} = L_{\Gamma_{\mathrm{f}}} \otimes L_{\Gamma_{\mathrm{i}}}$ is a Lorentzian with FWHM $\Gamma_{\mathrm{fi}} = \Gamma_{\mathrm{f}} + \Gamma_{\mathrm{i}}$. The intensity of the Auger electrons produced by the transition is given by the expression

$$I_{\mathrm{A}}^{\mathrm{fi}}(\varepsilon) = N_{\mathrm{ph}}\ \rho S F_{\mathrm{A}}^{\mathrm{fi}}(\varepsilon) \sigma_{\mathrm{ph}} \omega_{\mathrm{A}}^{\mathrm{fi}}$$

where σ_{ph} is the photoionization cross section (first step of the process) and $\omega_{\mathrm{A}}^{\mathrm{fi}}$ the partial Auger yield (second step of the process). Note that, when the photon energy is close to the core-ionization threshold, the Auger emission can be influenced by the presence of a slow photoelectron. The resulting interaction between the two escaping electrons in the field of the doubly ionized target is called post-collision interaction (PCI). The main consequence of PCI is an energy distribution of the emitted electrons that is no longer described by a Lorentzian function, but becomes asymmetric, broadened, and displays maximum shift in energy.

Together with the lines associated with the normal Auger transitions, one also observes satellite lines, which are called photosatellites if produced by the decay of one-hole states in which another electron has been simultaneously ejected (shake-off), for example, $KL \rightarrow LLL$, or excited (shake-up) to an empty level (L^*). In the latter case, the electron can either be directly involved in the following decay process, for example, $KL^* \rightarrow LL$ (participator transition), or act as spectator, for example, $KL^* \rightarrow LLL^*$ (spectator transition). An example of Auger spectrum with photosatellites is given in **Figure 3**.

In addition to these satellites produced by two-electron processes in the primary photoionization, there are also satellites due to double Auger transitions that are three-electron processes in which the third electron is either excited to an empty level, for example, $K \rightarrow LLL^*$ (discrete double), or ejected into the continuum, for example, $K \rightarrow LLL$ (continuous double). A schematic representation of normal and satellite transitions from core-hole and shake-modified core-hole states is given in **Figure 4**.

A separate group of peaks is that in which the initial state of the decay process is a core-excited state of the neutral system, that is, a discrete metastable state produced by the absorption of a monochromatized photon of energy $\hbar\omega_{\mathrm{res}}$ below the core-ionization threshold, for example, $1s \rightarrow np$ in Ne. The autoionization of this state produces electrons with nominal energies $\{\varepsilon_j^{\mathrm{A}} = \hbar\omega_{\mathrm{res}} - E_{\mathrm{f_j}}(1\mathrm{h})\}$

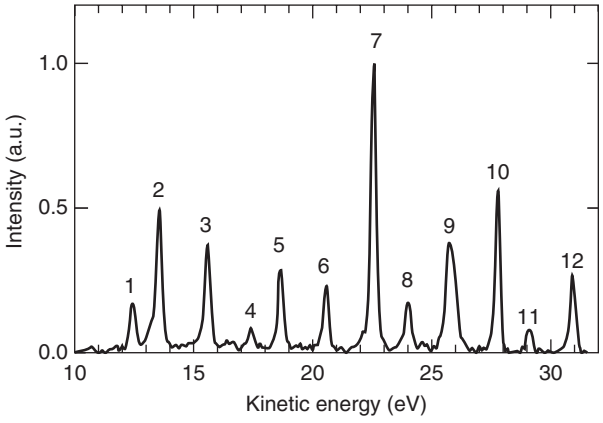


Figure 3 Neon $LL^* \rightarrow LL$ Auger spectrum. Lines 1,3,4,6 correspond to transitions $Ne^+[2s^12p^5(^3P)(3p,4p)] \rightarrow Ne^{2+}[2s^22p^4(^1D_2, ^3P)]$, lines 2,5,8,11 to transitions $Ne^+[2s^02p^6(^1S)(3p,4p)] \rightarrow Ne^{2+}[2s^12p^5(^1P,^3P)]$, lines 7,9,10,12, to transitions $Ne^+[2s^12p^5(^1P)(3p,4p)) \rightarrow Ne^{2+}[2s^22p^4(^1D_2, ^3P)]$. (Adapted from *Journal of Physics B* 33: 4343 (2000).)

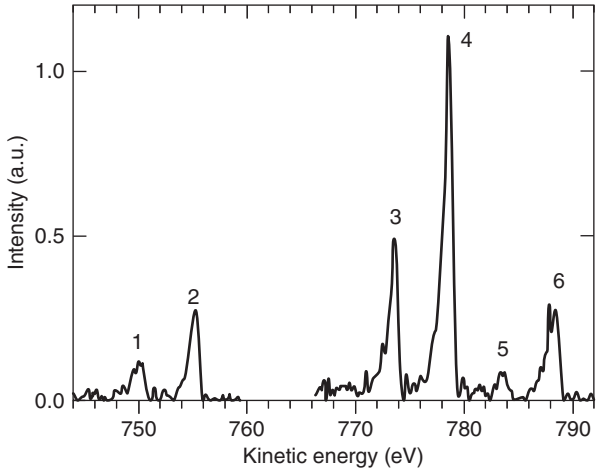


Figure 5 Resonant Auger transitions from the Ne $[1s^{-1}3p]$ state to the final $Ne^+[2s^{-2}np]$ (lines 1,2) and $Ne^+[2s^{-1}2p^{-1}np]$ (lines 3–6) ion states. (Adapted from *Journal of Physics B* 33: 4343 (2000).)

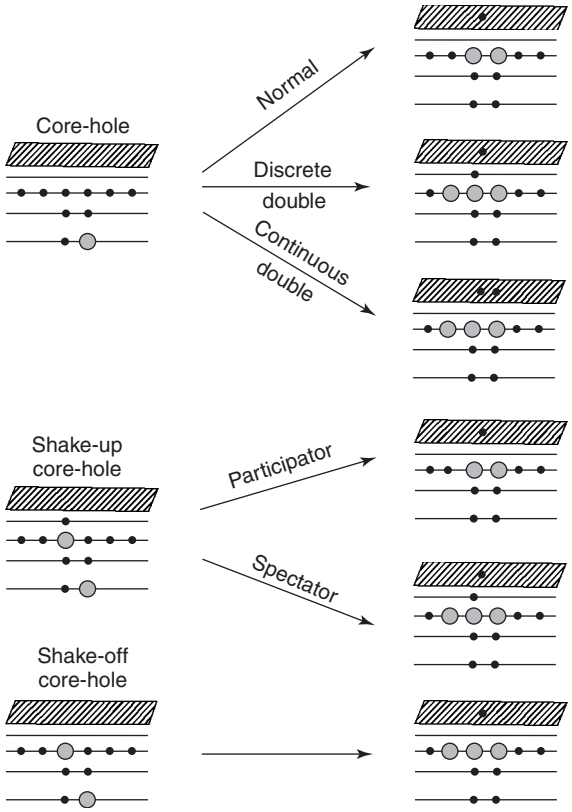


Figure 4 Selection of possible decay processes from core-hole and shake-modified core-hole states. The energy-level diagram (not to scale) is for electrons in 1*s*, 2*s*, 2*p* shells, in an excited orbital and in the continuum (shaded area).

either through a participator or a spectator transition. These excitations with their autoionization decay are called resonances and ought to be described as one-step processes (called Auger resonant or Auger resonant Raman), since the shape of the emission spectra are strongly dependent on the energy of the exciting radiation. The final state of these resonances can be reached via two indistinguishable pathways (a direct photoionization and a two-step process); it follows that the corresponding amplitudes can interfere and give rise to peculiar asymmetric lineshapes, called Fano profiles: see **Figure 5**.

Due to their similarity with a normal Auger process, the spectator transitions are also called resonance Auger or autoionization spectator transitions. Furthermore, due to correlation effects (orbital relaxation) during the decay process, it is also possible that the spectator electron can be promoted into a different orbital (shake-up) or even ejected into the continuum (shake-off). These processes are called shake-modified spectator transitions or resonance shake transitions.

Finally, there are also combinations of radiative and nonradiative transitions, called radiative Auger decay, in which the emission of a photon is accompanied by the simultaneous excitation or ejection of one electron.

Photon Impact Versus Electron Impact

Using high-energy electrons as projectiles instead of monochromatized photons, the energy transfer to the target covers a broad energy range: from zero up to the energy of the electron beam. It follows that the excitation/ionization of the target is not selective, but many different states of the neutral or ionized target can be populated; furthermore, the autoionization of these states produces additional structures in the electron spectrum. As a matter of fact, one collects a

large number of scattered and primary electrons with a wide range of kinetic energies. Hence, in noncoincidence experiments, the Auger lines are superimposed on a large and smooth background of electrons. Furthermore, at low-impact energies the production mechanism for satellites may be different for electron and photon impact. Finally, in the Auger decay of a metastable state produced by electron impact, there are, indeed, three electrons outgoing in the field of the ionized target; it follows that the presence of one electron with low kinetic energy is highly probable. This is the reason why, for electron impact, PCI does not vanish even for high energies of the projectile.

Molecular Spectra

The specificity of the molecular spectra resides on the presence of nuclear degrees of freedom in the target, that can be excited by the projectile and participate in the decay process. Final possible results of a collision between a molecule (M) and a photon are

$$
\begin{aligned}
& M + \hbar\omega \\
& \quad \rightarrow M^{n+} + n\, e^- \qquad \text{(ionization)} \\
& \quad \rightarrow M^* + \hbar\omega' \qquad \text{(resonance fluorescence)} \\
& \quad \rightarrow A + B \qquad \text{(neutral dissociation)} \\
& \quad \rightarrow A^+ + B + e^- \qquad \text{(dissociative ionization)} \\
& \quad \rightarrow A^+ + B^- \qquad \text{(ion pair formation)}
\end{aligned}
$$

In many cases, these channels are in mutual competition and characterized by the presence of intermediate metastable states, in which the target is electronically and/or vibrationally and/or rotationally excited. The result is a large number of spectral lines that, in principle, can be classified by the electronic, vibrational, and rotational quantum numbers of the final states of the ionized molecule. In many cases, however, the finite resolution of the electron spectrometer(s) does not allow one to separate these lines which, therefore, collapse into broad spectral profiles. Coincidence techniques (such as ion/photo- or Auger-electron, ion/ion, photo-electron/Auger-electron and their combinations) are the most powerful methods to investigate the dynamics of these processes that, in coincidence experiments, are treated as state-selected reactions.

The theoretical approach to study these processes is the same outlined above, but requires the use of wave functions that depend explicitly on both the electronic $\{\boldsymbol{r}\}$ and the nuclear $\{\boldsymbol{R}\}$ coordinates and, in Born–Oppenheimer approximation, are products of the type:

$$\Psi_{\alpha k}(\{\boldsymbol{r}\}\{\boldsymbol{R}\}) = \psi^{\mathrm{el}}_{\alpha_e \boldsymbol{k}}(\{\boldsymbol{r}\}, \{\boldsymbol{R}\}) \times \psi^{\mathrm{nu}}_{\alpha_n}(\{\boldsymbol{R}\})$$

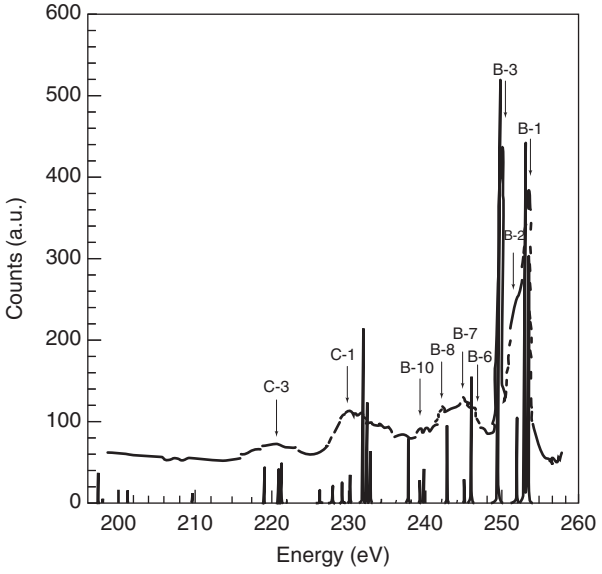


Figure 6 Calculated electronic transitions (see text) in the C1 *s*-Auger spectrum of the carbon monoxide, compared with the experimental spectrum (upper continuous line). (Adapted from Colle R and Simonucci S (1993) Multichannel resonance processes: theory and application to the Auger spectra of the CO molecule. *Physical Review A* 48: 392.)

The Auger spectroscopy has been applied to gas-phase molecules mainly to obtain information on their dicationic states. Auger spectra, however, also contain considerable chemical information on the local structure around the core-ionized atom, and it is often possible to recognize local similarities in Auger spectra of different molecules with the same functional group.

In **Figure 6**, the $C\,1s$ Auger spectrum of the carbon monoxide is shown in the region dominated by the normal $K \rightarrow YY$ transitions, where Y can be either a weakly ($Y = W$) or a strongly ($Y = S$) bound orbital; in particular, the C-lines correspond to $K \rightarrow WS$ transitions and the B-lines to $K \rightarrow WW$ transitions.

Deep-Level Spectroscopies of Atoms in Solids

Photoelectron, Auger-electron, and X-ray spectroscopies are increasingly used as tools for investigating solid-state systems in order to get information on electronic, structural, and dynamical properties of the system, and to know the type, arrangement, and properties of the atomic species on the surface and below the surface of the solid.

A powerful method for chemical analysis is electron spectroscopy for chemical analysis (ESCA), in which an X-ray beam of discrete energy is directed onto the specimen and the emitted electrons are collected and energy analyzed. The resulting spectra are, in general, simple and characterized by relatively

narrow and well-defined peaks that allow the identification of the emitting atoms.

Another method to have qualitative and quantitative information on the composition of the solid sample is the electron probe microanalysis, in which the X-rays emitted after ionization by electron impact are analyzed.

A more recent technique developed to obtain information on the structure of molecules (mainly low-Z molecules) bonded to or absorbed on a surface is near edge X-ray absorption fine structure (NEXAFS), also known as X-ray absorption near edge structure (XANES). This technique selects an atomic species through its K-edge and probes its bonds to intra-and extramolecular (i.e., surface atoms) neighbors by analyzing the prominent, structure-sensitive resonances in the first ~ 30 eV above the K-edge. An example is given in **Figure 7**.

The Auger spectroscopy of atoms in solids detects only the Auger electrons coming from the few outermost layers of the solid: the mean escape depth of these electrons may extend, indeed, from two to four atomic layers (4–8 Å), since Auger electrons produced below these layers lose enough energy to be scattered out of the peak into the inelastic background. The resulting Auger peaks contain information not only about the solid-state environment of the emitting atom, but also on the atom itself, and this is the reason of the importance of the Auger spectroscopy as an analytic tool. Auger spectroscopy also provides a useful technique for investigating two-particle density of states, hole–hole Coulomb interaction, and correlation effects. Unfortunately, the complexity of these spectra quite often obscures the results of this technique. The complexity is due to the overlap of adjacent lines, to the fact that, in general, the initial state of the decay process is not a simple one-hole state, to the difficulty in distinguishing between Auger and shake-modified transitions, and to the presence of plasma loss peaks (associated with both the elastic peaks and the more pronounced Auger peaks) and ionization loss peaks, due to incident electrons that lose part of their energy in the ionization of other atoms.

The Auger spectra of solids are thus characterized by a large background of electrons scattered inelastically by the atoms of the solid. These electrons have energies distributed over a wide range and constitute the so-called background of secondary electrons, whose intensity can be even two or three orders of magnitude bigger than the Auger signal. This is the reason why, instead of reporting the number of electrons recorded as a function of their kinetic energy, the Auger spectra of solid samples are often constructed by plotting the energy derivative of this number. An example is given in **Figure 8**.

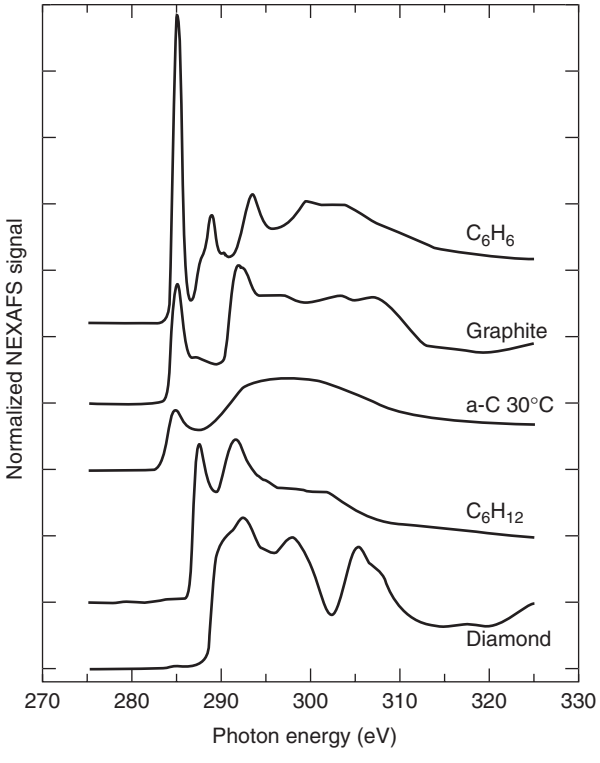


Figure 7 C1*s* NEXAFS spectra of various molecules and solids. (Adapted from Comelli G, Stöhr J, *et al.* (1988) Structural studies of argon-sputtered amorphous carbon films by means of extended X-ray-absorption fine structure. *Physical Review B* 38: 7511.)

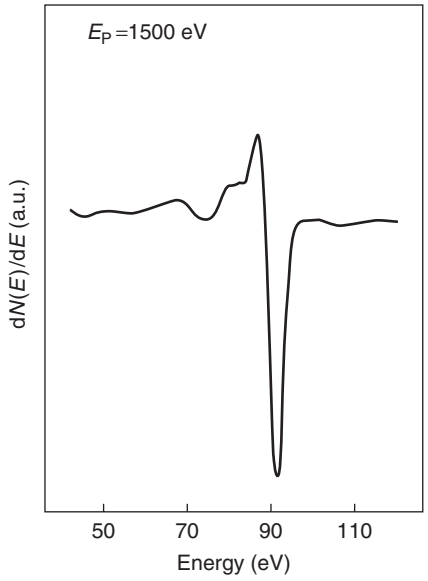


Figure 8 $L_{II,III}$ Auger-emission band of Si(1 1 1). (Adapted from Melles J, Davis L, and Levenson L (1974) Auger satellites of the $L_{2,3}$ Auger emission bands of Al, Si, and P. *Physical Review B* 9: 4618.)

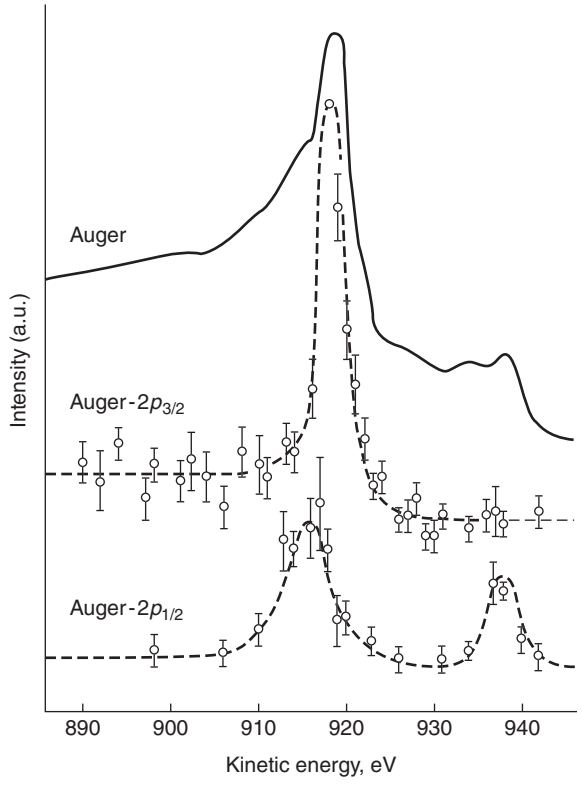


Figure 9 Auger and Auger–photoelectron coincident spectra of Cu. The second and third curves are the Auger spectrum in coincidence respectively with the $2p_{3/2}$ and $2p_{1/2}$ photoelectrons. Experimental points with error-bars are reported. (Adapted from *Physical Review Letters* 41: 1826 (1978).)

The complexity of the Auger spectra of atoms in solids is also the reason why much of the theoretical formulations is still phenomenological or semiempirical, and is quite often based on simplifying assumptions or model Hamiltonians, which try to simulate the real systems still allowing for analytical or numerical solutions.

Photoelectron/Auger-electron coincident spectroscopies have been recently explored to gain a better understanding of the correlations in solid-state systems. Detecting Auger- and photoelectrons simultaneously that are correlated in time and originate from the same core-ionization event, removes, indeed, most of the uncertainties and limitations of the standard Auger spectroscopy. Three different classes of experiments have been developed: (1) Auger spectra measured at fixed photoelectron energy; (2) photoelectron spectra measured at fixed Auger-electron energy; and (3) angular distribution patterns of Auger–photoelectron pairs. A highlighting example of the capability of these techniques to separate overlapping components of an Auger spectrum is given in **Figure 9**, where Auger and Auger–photoelectron coincidence spectra of copper are compared.

Finally, a recent technique, called either photoelectron or Auger-electron holography (depending on the process under consideration) is mentioned, which allows one to get three-dimensional images of the close vicinity of a given near-surface atom emitter. Using this technique, one records, on a photographic plate, the interference pattern (hologram) of a known reference wave (the photoelectron or Auger-electron wave emitted by the excited atom) with an unknown object wave (the superposition of the waves emitted by the surrounding atoms and due to multiple scattering processes); the obtained hologram is then illuminated with an appropriate decoding wave. The images of the atoms of the surface are obtained by means of computer reconstruction methods similar to those used in optical holography.

See also: Electronic Structure (Theory): Atoms; Electronic Structure (Theory): Molecules; Semiconductor Optics; Semiconductors, Electronic Transitions in; Time-Resolved Optical Spectroscopies; X-Ray Absorption Spectroscopy.

PACS: 82.80.Pv; 32.80.Hd

Further Reading

Aberg T and Howat G (1982) Theory of the Auger effect. In: Mehlorn W (ed.) *Corpuscles and Radiation in Matter. Encyclopedia of Physics XXXI*, p. 469. Berlin: Springer.

Agarval BK (1979) *X-Ray Spectroscopy*. Berlin: Springer.

Berkowitz J (1979) *Photoabsorption, Photoionization and Photoelectron Spectroscopy*. New York: Academic Press.

Chattarji D (1976) *The Theory of Auger Transitions*. New York: Academic Press.

Craig DP and Thirunamachandran T (1984) *Molecular Quantum Electrodynamics*. New York: Academic Press.

Newton RG (1982) *Scattering Theory of Wave and Particles*. Berlin: Springer.

Schmidt V (1997) *Electron Spectrometry of Atoms using Synchrotron Radiation*, 1st edn. Cambridge: Cambridge University Press.

Starace AF (1982) Theory of Atomic Photoionization. In: Mehlorn W (ed.) *Corpuscles and Radiation in Matter, Encyclopedia of Physics XXXI*, p. 1. Berlin: Springer.

Stöhr J (1992) *NEXAFS Spectroscopy*. Berlin: Springer.

Thompson M, *et al.* (1985) *Auger Electron Spectroscopy*. Berlin: Springer.

Defect Centers in Insulating Solids, Optical Properties of

G Baldacchini, ENEA, Frascati Research Center, Frascati, Italy

Introduction

As is well known from solid-state physics, crystalline materials (often also amorphous materials like glasses and ceramics) are of great interest in basic and applied research mainly because they contain defects. Indeed, properties of crystals or alloys such as mechanical strength, electrical and thermal conductivity, magnetic status, diffusion, and optical absorption and emission, are strongly dependent on crystalline irregularities and foreign impurities. For instance, impurities in metals increase their strength, decrease the electrical and thermal conductivity, and influence greatly the diffusion processes. Impurities in semiconductor and insulating crystals give them unusual electrical properties, which are used in electronic devices. Finally the optical properties are so deeply changed that new absorption and emission effects take place.

In so far as the optical properties are concerned, most insulating crystals (henceforth "crystals") are completely transparent, which is a consequence of their large bandgap. Indeed, the valence and conduction bands are separated by more than $\sim 3\,\text{eV}$ (400 nm), which is the upper limit for the detection of the human eye, and so the visible radiation cannot excite electrons from the valence band to the conduction band. In such crystals, the valence band is completely full of electrons, while the conduction band is empty, which explains their electrical and thermal insulating properties. **Figure 1** shows the absorption spectrum of a typical crystal, in this case an alkali halide (AH), which absorbs in the ultraviolet (UV) and in the infrared (IR) region of the spectrum, but not in the visible (VIS) part, that is, from ~ 1.7 to 3.1 eV (720 to 400 nm). The UV absorption is associated with the electronic excitations, whereas the IR band is due to lattice optical phonons, the so-called restrahl absorption.

When defects are introduced in the crystal, different effects can occur, depending on the nature of the defects. They can be extended, like dislocations, and so their electrons can be described by quasiperiodic or periodic wave functions as the electrons in the original crystal and, as a consequence, the optical properties change all over the spectrum. In practice, it is like dealing with a slightly new and different crystalline solid, and so this case is of no interest here. The situation is different when defects are localized, because the electrons bound to the defects cannot be described by the same wave functions as the electrons in the crystal, and so their eigenvalues too cannot be similar to those of the pure crystal. In this case new absorption bands appear in the wide region of the energy gap of **Figure 1**, where electron–crystal states do not exist. **Figure 2** shows a few typical absorption bands found in an AH crystal containing different kinds of point defects. The absorption bands are limited on the high-energy side by the ultraviolet absorption edge, E_g, of the crystal and on the low-energy side by the infrared (IR) lattice absorption. When these bands fall in the visible region of the spectrum, the crystal displays a

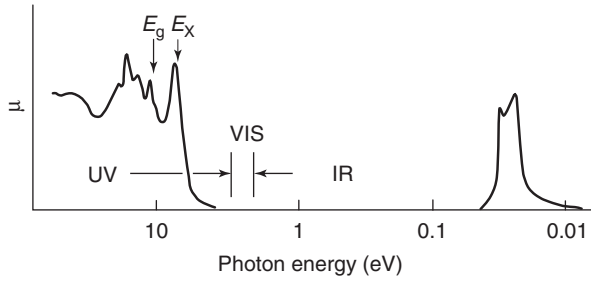


Figure 1 Absorption coefficient of a typical AH crystal as a function of energy in a log scale in order to show the wide range of the spectrum and the energy gap well inside the ultraviolet region. The absorption at high energy rises with electronic excitation (the labeled structures belong to excitons), while at low energy the peaks are associated with optical phonons. (Reproduced from Knox RS and Teegarden KJ (1968) Electronic excitations of perfect alkali halide crystals. In: Fowler WB (ed.) *Physics of Color Centers*. New York: Academic Press.)

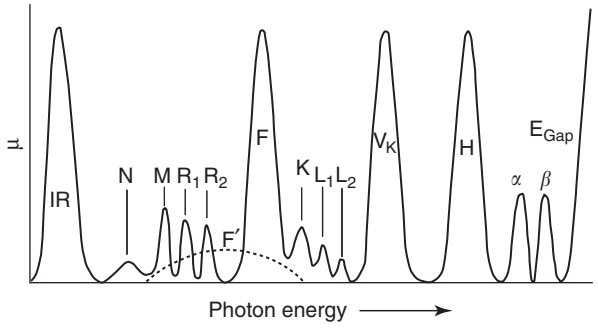


Figure 2 Schematic absorption coefficient of typical intrinsic point defects in AH crystals as a function of energy, which has been reversed with respect to **Figure 1**. The various absorption bands, which have been associated with well-defined color centers as labeled, rise in the wide crystal gap between the IR and ultraviolet E_g absorption of the crystal. (Reproduced from Townsend PD and Kelly JC (1973) *Colour Centres and Imperfections in Insulators and Semiconductors*. London: Sussex University Press.)

coloration that varies from blue to red, and for this reason point defects are usually referred to as color centers (CCs) in the literature.

There are several other kinds of crystals in nature, and it would be impossible even to simply consider them all in this article, which focuses only on the AHs, as shown in **Figures 1** and **2**. They are ionic crystals with a lattice cubic structure, and there are 20 of them, due to the combination of five alkaline metal atoms (Li, Na, K, Rb, Cs) with four halide atoms (F, Cl, Br, I), among them the well-known common salt, NaCl. However, it must be borne in mind that most of the optical properties of AHs are similar to those of other crystals, and so several approaches utilized in the AHs can also be extended to others. In other words, the simple AH crystals represent almost a model case for the crystals in general.

In so far as the point defects are concerned, they can be divided into intrinsic and extrinsic ones. Vacancies (empty lattice positions) and interstitials (lattice ions moved out of lattice positions) belong to the first class, while impurities belong to the second class. Impurities can be positive or negative atomic or molecular ions, which occupy the lattice positions or cause lattice distortions referred to as off-center situations; when they are monovalent, isoelectronic ions, they simply occupy the lattice positions, but when they are divalent or aliovalent ions, they produce vacancies nearby to keep the whole crystal electrically neutral. Finally, the intrinsic defects can bind themselves to the extrinsic ones, giving rise to defects associated with the impurities.

It is clear from the above discussion that the subject of defects in crystals is vast and, in particular, the extrinsic defects, due to their sheer large number represent an independent field of research in solid-state physics. In spite of their importance, these defects are not considered here and this article will be limited to addressing the intrinsic defects, which belong to the same constituents of the crystal (the CCs), and the intrinsic defects perturbed by foreign impurities (the perturbed CCs).

Color Centers

In 1863, a crystal of NaCl was colored at high temperature in Na vapors. However, this result and others obtained subsequently also in connection with the discovery of natural radioactivity and X-rays, which both produced coloration in several crystals in 1896, were considered mere scientific curiosities up to $\sim$1920 when a systematic research was started at Göttingen, Germany, by RW Pohl and co-workers.

Prior to 1940, CCs in AHs were characterized well enough as far as light absorption was concerned. At the same time, research in this field started to move outside Germany, where it was confined for almost a generation; important contributions were soon added elsewhere. For instance, there were high expectations for fluorescence in the near IR after optical pumping in the CC absorption bands in the visible region. After World War II, research was resumed and a lively debate started again about the existence of emission after optical excitation of the F center, which, as shown below, is the prototype of all CCs. Eventually, luminescence was found in 1954, and the optical cycle of the F center and other kinds of CCs was completely clarified subsequently.

As shown in **Figure 2**, there are several kinds of CCs that have been discovered and called differently. **Table 1** lists the known CCs with their old and actual denomination, without claiming to be a complete list. By looking at the basic structure of the CCs, there are only two species of them, that is trapped-hole and trapped-electron centers. Both of them appear, for instance, by irradiating AH crystals with X-rays, the former always having their bands, as V_k in **Figure 2**, at higher energies with respect to the latter, as F in the same figure. Although they were discovered at the same time long ago, the trapped-hole centers are not yet well known. Indeed, they produce so many absorption bands that it has proved to be difficult to associate each of them with well-defined centers and, as a result, only a few among them have been identified structurally. Moreover, trapped-hole centers never became as popular as trapped-electron centers. The latter have given rise to a much richer scientific production, as compared to the trapped-hole centers, which will not be discussed here.

CCs can be produced by resorting to chemical and physical methods. Among the first ones, the additive method is the most known and utilized to obtain homogeneously colored samples, while the electrolytic method is much simpler although the coloration is rather inhomogeneous. Usually, all crystals can be colored by irradiating them with ionizing radiation, UV light, X-rays, γ-rays, electron beams and elementary particles, ions, etc. **Figure 2** shows the absorption bands corresponding to several CCs produced with the previous methods in a typical AH crystal. If these bands are excited optically, it is possible to observe, especially from trapped-electron centers, emission bands that are strongly Stokes-shifted and often very intense. **Figure 3** gives the absorption and emission bands of the F center in KCl for various temperatures where, besides the Stokes shift, the broadening and peak positions show a sizeable temperature dependence. Similar properties belong also to the other CCs in KCl crystal, which are reported schematically in **Figure 4**, where on the left side there

Table 1 List of known color centers, with their previous and actual denomination

Actual notation	*Old notation*	*Description*
Trapped hole centers		
V_K	V_K	Self-trapped hole on a pair of halogen ions
H	H	Hole trapped along four halogen ions
V_4	V_4	A halogen di-interstitial
H_A, H_B	V_1	H center adjacent to alkali impurity ions
I^-	I	Interstitial halogen ion
I	[a]	Interstitial atom
[a]	F (halogen)$^-$	Mixed hole centers
[a]	V_2, V_3, …	Unidentified hole centers
Trapped electron centers		
F^+	α	A halogen ion vacancy
[a]	β	Not a CC but an excitation related to an F center
F	F	A halogen ion vacancy with one electron
F_2	M(C)	Two adjacent F centers
F_3	R_1(E), R_2(D)	Three adjacent F centers
F_4	N_1(G), N_2(G)	Four adjacent F centers
[a]	O	Unidentified electron center
F^-	F′	A halogen ion vacancy with two electrons
F_2^-	M′	Two adjacent halogen vacancies with three electrons
F_2^+	M^+	Two adjacent halogen vacancies with one electron
F_3^-	R′	Three adjacent halogen vacancies with four electrons
F_3^+	[a]	Three adjacent halogen vacancies with two electrons
F_4^-	N′	Four adjacent halogen vacancies with five electrons
F_A	A	An F center adjacent to one alkali impurity ion
F_B	B	An F center adjacent to two alkali impurity ions
F_C	[a]	An F center adjacent to three alkali impurity ions
F_H	[a]	An F center adjacent to a halide impurity ion
$(F_2)_p$	[a]	An F_2 center perturbed by defect centers
$(F_2^+)_{A/H}$	[a]	An F_2^+ center adjacent to an impurity positive/negative ion
$(F_2^+)_p$	[a]	An F_2^+ center perturbed by defect centers
F_4-like	[a]	An F_4 center perturbed by defect centers
$Tl^\circ(1)$	[a]	A halogen ion vacancy with one neutral Tl atom
F_Z	Z_1	An F_2^+ center with divalent positive ions
U	U	An F center adjacent to one hydrogen atom
U_1	U_1	An interstitial hydrogen negative ion
U_2	U_2	An interstitial hydrogen atom

[a] No correspondent notations exist.

is an essential design of the structure of a few centers as described in **Table 1**, while on the right side their absorption and emission bands are approximately sketched. It is immediately observed that CCs in KCl absorb light mostly in the VIS region and emit through the entire near-IR region, a property which is common to all CCs in the other AH crystals. Indeed, the elementary processes responsible for such behavior are practically the same in all of them, and the following sections will discuss the F center, as also the F_A and F_H centers, the three of them representing a good sample of all CCs.

F Centers

The F center, an electron trapped in an anion vacancy, is the simplest point defect, and for this reason it is also a prototype case study for all the other more complex CCs, as is the hydrogen atom with respect to atoms and molecules. In the normal state, the electron of the F center (see **Figure 4**) moves inside the cage of the six positive ions surrounding the anion vacancy, a situation equivalent to an electron in a box, a well-known problem in quantum mechanics. The separation of the first two quantum states, corresponding to the 1*s*, and 2*p* levels, is given by

$$E_{2p} - E_{1s} = 3\pi^2\hbar/8ma^2 = A/a^2 \qquad [1]$$

where $\hbar$ is the reduced Planck constant, m is the mass of the electron, and a is the nearest-neighbor lattice separation. This expression, known as the Mollwo–Ivey law, fits the experimental data of the absorption peak energies of the known F centers surprising well, as reported in **Figure 5**.

If eqn [1] is applied to the known emissions of the F centers, it fails completely. Indeed, the electron in the 2*p* state is not confined anymore in the previous

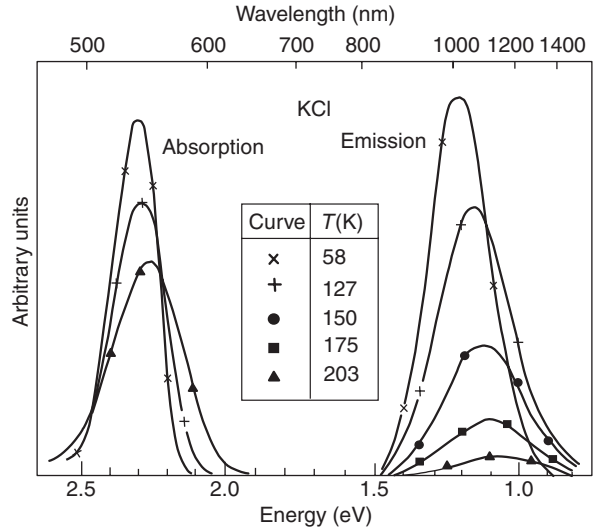


Figure 3 Absorption and emission bands of F centers in KCl as a function of temperature. (Reproduced from Luty F (1961) *Senderdruck aus Halbleiterprobleme Band VI*. Braunschweig: Verlag Friedr. Vieweg & Sohn.)

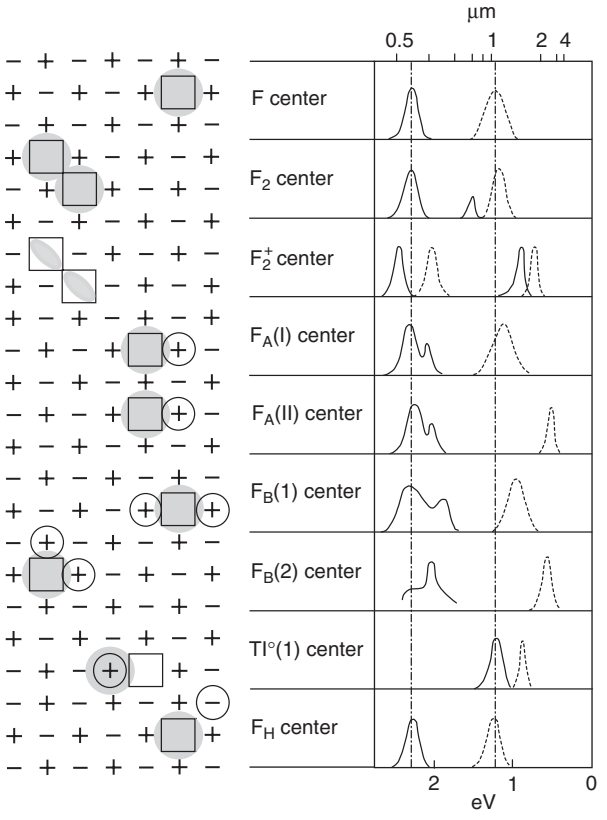


Figure 4 Structural models, absorption (solid line) and emission (dashed line) behaviors of the F center and of relatively simple aggregated centers in KCl at low temperatures. On the left, the square represents the halogen vacancy, the circle the foreign ion, and the shaded area the approximate electron density. (Reproduced from Baldacchini G (1997) Optical excitation and relaxation of solids with defects. In: Di Bartolo B (ed.) *Spectroscopy and Dynamics of Collective Excitations in Solids*. New York: Plenum.)

cage, but rather it moves outside it as well and feels an average potential that can be approximated by a Coulomb field with an effective mass m^* and a high frequency (ω_∞) dielectric constant. The separation of the first two quantum states can be derived by the Rydberg law, and is given by

$$E_{2p} - E_{1s} = 3e^4 m^* / 128\pi^2 \hbar^2 \omega_\infty^2 = E/\omega_\infty^2 \qquad [2]$$

The experimental data of the emission peak energies reported in **Figure 6** show that the functional dependence of eqn [2] is verified well enough. It is worth noting that the number of experimental points reported in **Figure 6** is smaller than that reported in **Figure 5**, an important occurrence, which will be discussed later.

It is known that the previous unusual properties are a consequence of the electron–lattice interaction, which, for instance, is missing in atoms where the absorption and emission spectra coincide. This peculiar situation of the CCs is better described by the configuration coordinate diagram shown in **Figure 7**. The two parabolas describe the energy of the F center in the ground, 1*s*, and excited, 2*p*, states as a function of the normal-mode coordinate, Q, which represents the displacement of the surrounding ions from the equilibrium positions. The amount of horizontal displacement between the two parabolas is a measure of the electron–lattice coupling, which is quite strong in the case of F centers in AHs. Indeed, they have very big Stokes shifts, as shown in **Figures 3**, **5**, and **6**, and usually the more the 2*p* parabola moves to the right, the less energetic is the emission, that is, the Stokes shifts are proportional to the electron–lattice couplings; this statement is valid for all point defects in crystals.

At low temperatures, the electron of the F center occupies the bottom of the 1*s* parabola, which is the ground state (GS). Optical excitation produces a vertical transition (Frank–Condon principle) to the 2*p* state in a highly excited vibrational level (i.e., absorption, A). The vibronic system, electron + phonon (phonon being an elementary excitation of the lattice vibration), relaxes very quickly, $\sim 10^{-12}$ s, to the bottom of the 2*p* parabola, which is called the relaxed excited state (RES). The electron can remain in this state for a relatively long time, $\sim 10^{-6}$ s, before returning to the unrelaxed GS with a vertical transition accompanied by light emission. In the end, a further lattice relaxation completes the optical cycle to the GS.

However, the emission, which is also called the ordinary luminescence (OL), is not the only radiative process that takes place after the excitation. Indeed, a

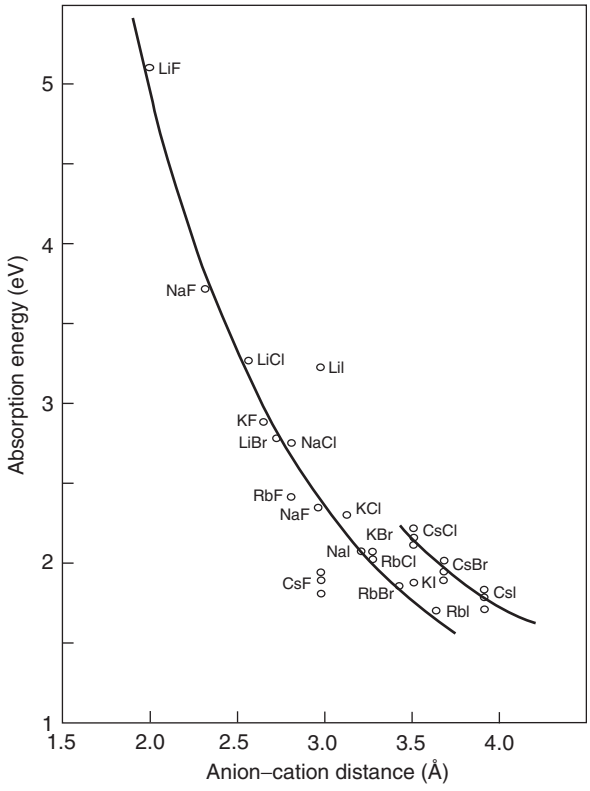


Figure 5 Peak position at low temperature of the F-center absorptions as a function of the nearest-neighbor distance. Continuous lines are best fits of eqn [1] to the experimental data. (Reproduced from Baldacchini G (1992) Relaxed excited states of color centers. In: Di Bartolo B (ed.) *Optical Properties of Excited States in Solids*. New York: Plenum.)

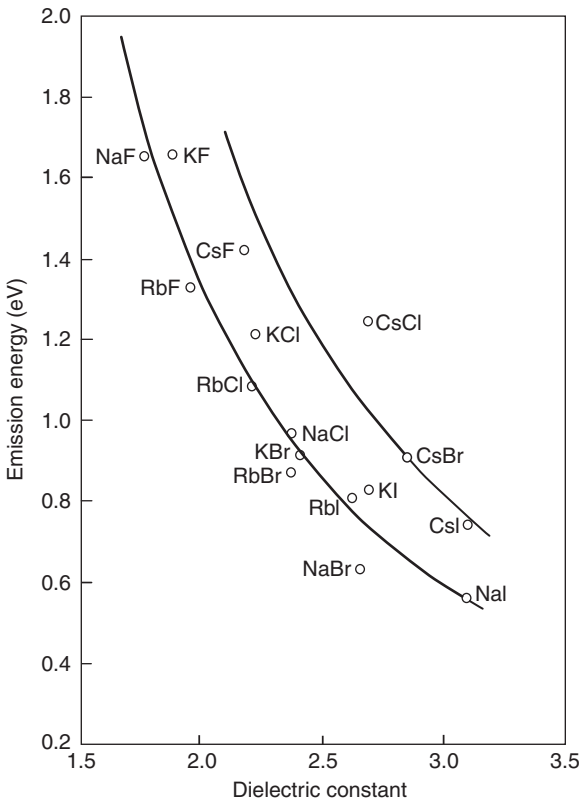


Figure 6 Peak position at low temperature of the F-center emissions vs. the dielectric constant at high frequency. Continuous lines are best fits of eqn [2] to the experimental data. (Reproduced from Baldacchini G (1992) Relaxed excited states of color centers. In: Di Bartolo B (ed.) *Optical Properties of Excited States in Solids*. New York: Plenum.)

little beyond the obvious Rayleigh line, a structured emission reveals the presence of Raman scattering (RS), which gives detailed information on the lattice modes coupled to the electronic transition. Still beyond, and up to the OL, there is a very weak emission tail that originates from decaying processes during the fast relaxation, and is commonly known as hot luminescence (HL).

So, with the previous model, which is based on the results from molecular chemistry and physics, the main optical properties of the F center and, for similitude, of the other point defects are well understood by now.

F_A Centers

As F centers are the simplest CCs, F_A centers are the simplest among the perturbed CCs, as shown in **Figure 4**. They were discovered during World War II in Gottingen and were originally called A centers (see **Table 1**).

Because of their axial symmetry due to the neighbor impurity ion, the absorption band splits into two separate bands, which is a characteristic common to all F_A centers, and related to reorientation processes which are optically induced and thermally activated. The case of their emission is different, which can be either similar to the F center (with a broad band with the usual Stokes shift), or a narrow band with a much bigger Stokes shift. Because of this peculiar difference, F_A centers have been traditionally divided in type I and type II, respectively, as shown in **Figure 4**. The reason for such a difference has been identified in the so-called saddle-point lattice configuration, formed by two vacancies separated by the impurity ion in an interstitial position. Thus, the ionic potential assumes the aspect of a double well, which can accommodate (much better than a single well) the two lobes of the $2p$ wave function. As a consequence, the RES become energetically deeper with a larger Stokes shift with respect to the F center.

Several attempts have been made to forecast the typology of F_A centers in AHs and, by taking into account the dimension of the ions, the effective charge parameter, and the residual covalent bonding, it is possible to give fairly good indications. So, the

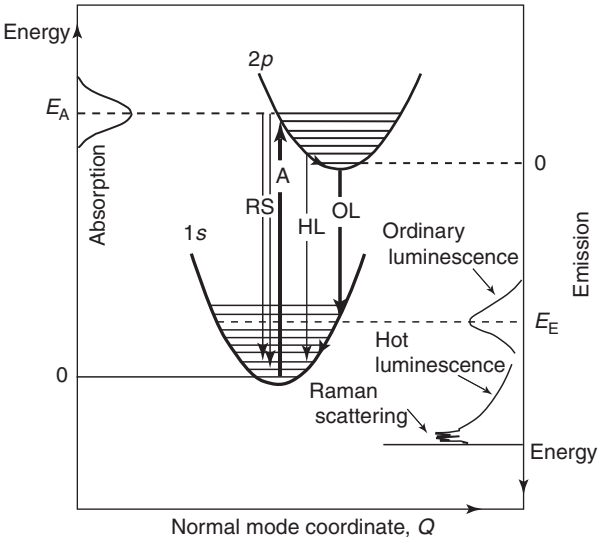


Figure 7 Configuration coordinate diagram illustrating the absorption and emission phenomena of the F center. The RS and HL are usually a few orders of magnitude less intense than OL (see text for details). (Reproduced from Baldacchini G (1992) Relaxed excited states of color centers. In: Di Bartolo B (ed.) *Optical Properties of Excited States in Solids*. New York: Plenum.)

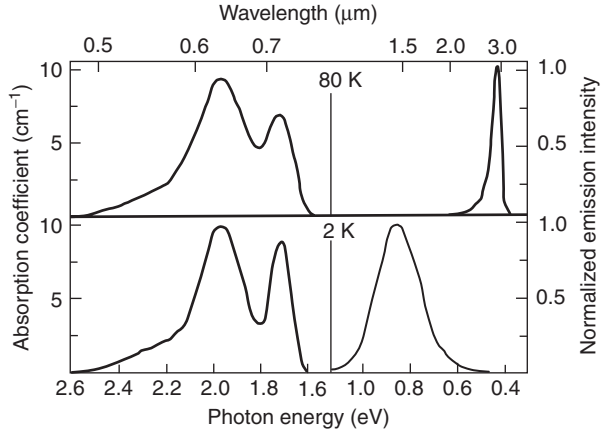


Figure 8 Absorption (left) and emission (right) at two temperatures of an $RbCl:Li^+$ crystal containing F_A centers. The luminescence is excited in both cases by an He–Ne laser. (Reproduced from Baldacchini G (1997) Optical excitation and relaxation of solids with defects. In: Di Bartolo B (ed.) *Spectroscopy and Dynamics of Collective Excitations in Solids*. New York: Plenum.)

experimental observation in 1988 that both type I and II centers can coexist in the same crystal host at different temperatures was an absolute surprise. **Figure 8** shows the absorption and emission spectra of the same crystal $RbCl : Li^+$ containing F_A centers at two temperatures. While the typical emission of type II centers is observed at 80 K, at 2 K the emission band is large and peaked at $\sim$0.85 eV, much like the F emission band in the same crystal peaked at 1.09 eV. It has also been shown that the two emissions are strongly temperature correlated, in the sense that while the intensity of type I band increases with temperature, the intensity of type II band decreases, and their temperature behavior can be described by a thermally activated process. This means that both types of F_A centers coexist in the same host crystal, separated by a potential barrier that can be overcome by a thermal excitation. This picture is very likely common to all F_A centers, because the thermally assisted type I to type II transition has also been observed in $KF : Na^+$, type II F_A center well known until recently for its characteristic emission at 0.6 eV at 100 K, which moves to 1.3 eV at 2 K.

F_H Centers

As simple as the F_A centers, the F_H centers differ from them by the charge of the impurity neighbor ion, which is negative in this case. These centers have been studied for long but, contrary to what happens to the F_A centers, their optical properties do not show sizeable changes with respect to the F centers, unless molecular ions such as OH^- and CN^- are used as impurities.

First, when F centers and molecular ions are aggregated, in general the absorption bands become broader but not much different from the original F band with both types of ions, while the emission is completely quenched with OH^- and almost unchanged by CN^- ions. Moreover, it has been shown that the electronic energy of RES is efficiently transferred to the stretching vibrational mode of the ions, where it is converted in lattice vibrations (phonons which ultimately are equivalent to heat) in the case of OH^-, and in vibrational IR emission at $\sim$0.25 eV in the case of CN^-. These novel properties are clearly very interesting and also very complex, because they reflect the elementary interactions of small atomic entities with all their peculiarities. As a consequence, they have produced many results, which are not considered in detail here. The present focus is on the recently discovered aspect concerning the optical and thermal bistability of the $F_H(OH^-)$ centers.

When OH^- ions are aggregated to F centers in a KBr crystal, the absorption spectrum changes as reported in **Figure 9**. The dotted absorption curve represents the band of the F centers in KBr in the same experimental conditions, for comparison. When the sample is cooled to 1.9 K, band I is observed first, but upon optical irradiation at 580 nm, as marked by the left arrow, band II is obtained. Conversely, band I appears again upon optical irradiation at 650 nm, as

marked by the right arrow. This bidirectional conversion can be repeated any number of times with no loss of absorption intensities. Moreover, if the temperature is increased up to 100 K, the original absorption band undergoes various conversions at well-defined temperatures, which indicates the existence of thermally assisted conversions. This complex situation has been ascribed to the different orientations assumed by the OH^- molecular dipole perpendicular or parallel to the F center–ion direction.

Besides the knowledge of the exact microscopic model, the F_H center displays an interesting and uncommon optical bistable process in KBr and recently also in KI, which means, again as for F_A centers, that it could be a general phenomenon in the AHs.

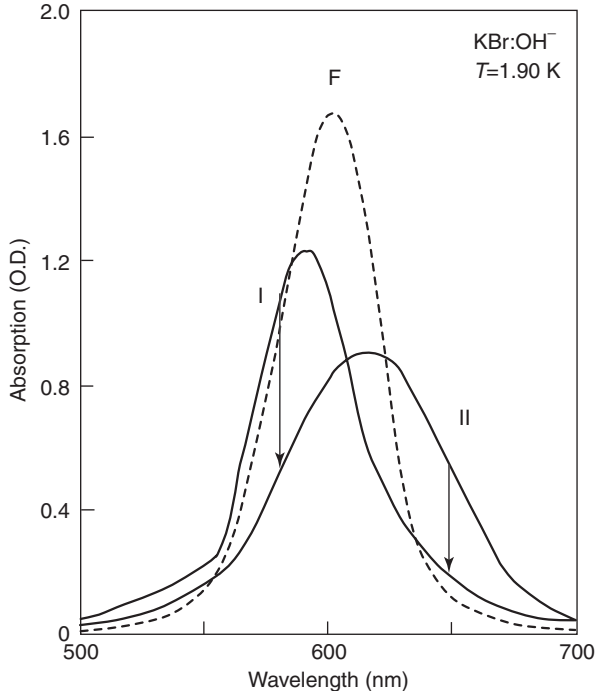


Figure 9 Optical absorption of the F center (dashed line) and of the two configurations I and II of the $F_H(OH^-)$ centers in KBr + 1.5% KOH at 1.9 K. (Reproduced from Baldacchini G (1994) Optical nonlinearities of color centers. In: Di Bartolo B (ed.) *Nonlinear Spectroscopy of Solids: Advances and Applications*. New York: Plenum.)

Applications

The basic research in CCs has also been accompanied by progress in their applications. Following the first experiments with natural radioactivity, there was an attempt, during World War II, to use the AHs to measure the artificial radioactivity from nuclear reactions, which failed because of the fading coloration in the majority of the crystals. At the same time, the process of coloration was also studied for the realization of cathode-ray screens for radar, but never really employed, in favor of more suitable solid-state materials. Later, after the discovery of the luminescence, stimulated emission was observed in F_A centers in 1965 and subsequently the first CC laser realized in 1974. Contemporarily, the possibility of storing images in three-dimensional holographies was demonstrated again with F_A centers. Moreover, the generation of hole burnings in inhomogeneously broadened absorption bands, the so-called zero-phonon lines, of F_3^+ centers opened the way to the practical fabrication of massive optical information storage.

In the subsequent years, CCs have been used mostly in the dosimetry techniques, where they had intrinsic

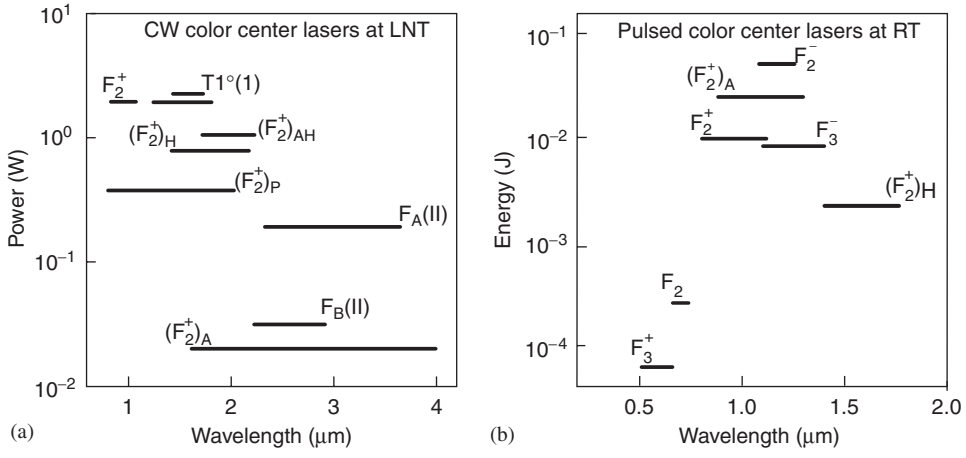


Figure 10 (a,b) Average power (energy) and tunability range in cw (pulsed) regime of the most common CC lasers in AHs working at liquid nitrogen temperature (room temperature) vs. wavelength. (Reproduced from Baldacchini G (1999) New luminescent phenomena in color centers. In: Struck CW, Mishra KC, and Di Bartolo B (eds.) *Electrochemical Society Proceedings*, vol. 98–24. Pennington, NJ: The Electrochemical Society.)

advantages over other materials, and in the laser field, where they reached a complete maturity by covering the spectral range between 0.8 and 4 μm, as shown in **Figure 10** which refers to continuous wave (cw) and pulsed-laser modes, respectively. These lasers, which are tunable and possess a very high spectral purity, have contributed greatly to the advancement of other scientific and technological sectors, such as high-resolution molecular spectroscopy, and communication via optical fibers where they have initiated the propagation of the soliton waves.

Nowadays, while the dosimetry use is still continuing, many CC lasers are no longer utilized, mainly because of new and simpler solid-state laser sources in the same spectral range. Indeed, as seen in **Figure 10**, most CC lasers operate at liquid nitrogen temperature, a practice which is rather cumbersome for users, especially when it is possible to avoid it. However, in special experiments, they still represent the only laser sources with the required characteristics, and this applies especially in the case of pulsed CC lasers at room temperature, which generally use the LiF crystal.

Conclusions

CCs represent one of the oldest fields in solid-state physics, having been studied for almost 150 years. They are well understood, but at the same time there are still many sectors which have to be clarified, although nowadays only a few of them are still tackled for use in applications.

First, as has already been mentioned earlier the known emissions of the F centers are much less than the number of AHs. Indeed, although all F centers display a normal light absorption, many of them do not produce any emission at all under optical excitation. This lack of emission refers to the phenomenon of quenching, which has plagued the CCs since the beginning, explaining also the delay in the discovery of emission. Many quenching factors are known to work effectively for CCs, but there are still several aspects waiting to be clarified, which are important also for other point defects in different materials, such as solid-state lasers.

Figures 5 and **6** also show that the experimental points cannot be fitted by a unique curve, a fact that underlines the hidden complexity of the F centers, which are in any case the simplest known defects in insulating crystals. Indeed, by taking into consideration all the properties such as the extension of the wave function, higher-lying levels, Stark effect, uniaxial stress effects, magnetic effects, RES lifetime, etc., it is possible to compare them quantitatively with a unique comprehensive theoretical model in the case of F centers in KCl. But when the same vibronic model is applied to the F centers in other AHs, there is no satisfactory agreement.

About two decades ago, an expert in this field said: "Evidently a vibronic model for the RES that provides a completely consistent quantitative explanation for all data has not been achieved." The situation does not look too different even now. This deadlock is certainly a consequence of the CCs being much more complex than they appeared to be at first sight, but it is also due to the scarce scientific interest in CCs. Indeed, many researchers switched their investigations to other more remunerative fields, as a consequence of the strong applications bias of most of the funding agencies.

However, there have been important contributions in the area of point defects lately, and among them those regarding LiF have been particularly notable. This AH crystal is distinguished by its peculiar properties. It is practically immune to moisture, possesses one of the lowest indexes of refraction and the largest bandgap, is colored only by means of ionizing radiation, and its CCs are stable at room temperature. Last but not the least, F centers do not produce any significant luminescence, while a large number of more complex CCs are optically active in the VIS and near IR region. They are also efficient light emitters, for example, almost all the laser emission in **Figure 10b** corresponds to LiF. In recent years, it has been applied successfully in the form of thin films, as passive and active waveguides, with interesting implications in fiber-optic communications. Moreover, it has proved to be a very good detector for soft X-rays; in several ways it's more efficient and sharper with respect to the detectors used at the moment for the same purposes. In conclusion, LiF salt is becoming a new interesting photonic material with promising developments in basic research and applications as well, mainly because of the long history of studies on CCs in AH crystals.

See also: Insulators, Optical Properties of; Ionic Bonding and Crystals; Optical Absorption and Reflectance; Optical Bistability; Point Defects; Polymers and Organic Compounds, Optical Properties of.

PACS: 42.55.Rz; 61.72.Ji; 61.82.Ms; 78.40.Ha; 78.55.Fv

Further Reading

Agullo-Lopez F, Catlow CRA, and Townsend PD (1988) *Point Defects in Materials*. London: Academic Press.

Basiev TT and Mirov SB (1994) *Room Temperature Tunable Color Center Lasers*. Switzerland: Harwood Academic.

Farge Y and Fontana MP (1979) *Electronic and Vibrational Properties of Point Defects in Ionic Crystals*. Amsterdam: North-Holland.
Henderson B (1972) *Defects in Crystalline Solids*. London: Edward Arnold.
Mott NF and Gurney RW (1964) *Electronic Processes in Ionic Crystals*, ch. IV. New York: Dover.
Schulman JH and Compton WD (1962) *Color Centers in Solids*. Oxford: Pergamon.
Seitz F (1954) Color centers in Alkali Halide Crystals II. *Review of Modern Physics* 26: 7–94.
Stoneham AM (1975) *Theory of Defects in Solids*. Oxford: Clarendon.
Teichmann J and Szymborski K (1992) Point defects and ionic crystals: color centers as the key to imperfections. In: Hoddeson L, Braun E, Teichmann J, and Weart S (eds.) *The History of Solid State Physics*. Oxford: Oxford University Press.

Deformation Processing

D Raabe, Max-Planck-Institut für Eisenforschung, Düsseldorf, Germany

Introduction

Modern manufacturing is to a large extent based on deformation processing. This applies in particular to elastic–plastic matter such as metallic alloys, composites, and polymers. Understanding of these materials and more advanced deformation processing techniques are paramount to modern manufacturing and product development. The current state in this field is characterized by the close integration of engineering concepts pertaining to manufacturing technology and materials physics on the one hand, and by the close integration of experiment, industrial processing, and computer simulation on the other.

Deformation processing encompasses such different techniques as forging, rolling, extrusion, drawing, or bending. Modern understanding of these techniques consists not only in attaining mere shape changes of plastic matter but also in optimizing their microstructures and properties.

This article gives a brief introduction to some important forming processes and simulation methods in this field.

Basic Types of Deformation Processes

Deformation processes use irreversible deformation to permanently change the shape of sufficiently ductile workpieces and to obtain desired microstructures and material properties at the same time. This means that deformation processes work in the elastic–plastic regime of the stress–strain curve of the material. The layout and understanding of such processes require knowledge about the mechanical and microstructural properties of the deformed materials, the external loads and contact boundary conditions such as friction, as well as the tool design. For a concise overview of the basic deformation processes, it is pertinent to focus on the four main bulk deformation processes, namely, rolling, forging, extrusion, and bulk drawing and on the four main sheet deformation processes, namely, bending, drawing, stretching, and shearing.

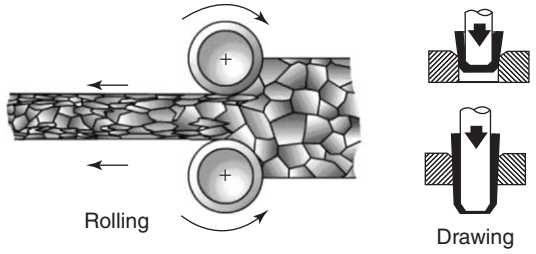


Figure 1 Schematics of two important deformation processes.

Rolling is characterized by two opposing rolls which reduce the thickness of a plate of material essentially under compression (see **Figure 1**). Forging is a process where two opposing dies form a part under large compressive stresses. This process is usually performed at elevated temperatures for reducing the required forces. Extrusion and drawing are compressive bulk forming processes where the material is forced (pushed or drawn) through a die opening, thereby reducing its cross section. Bending is a method of stressing a workpiece beyond its elastic limit, without substantial cross-sectional thinning, so that when stress is released the workpiece does not return to its original shape. Drawing, stretching, and shearing are related processes of working metal blanks in dies on a press into complex shapes.

Basics about the Simulation of Deformation Processes

Simulations of deformation processes are usually based on solving large sets of differential equations associated with a well-posed forming problem by use of nonlinear finite element methods in conjunction

with detailed physically based constitutive descriptions of the mechanical behavior of the processed materials. The underlying mechanical continuum formulations approximate force and momentum equilibrium as well as displacement compatibility on the basis of energy minimization considerations. The development of physically based laws for constitutive materials as input to such simulations has increasingly gained momentum in this field, because deformation processing does not only aim at shape changes of matter but also at optimizing microstructures and properties of the formed product.

Primary objectives of deformation simulations are the prediction of material shape, the minimization of failure, the optimization of material flow, and the calculation of final microstructures and properties after forming. Further applications are in the fields of optimizing tool designs, predicting tool forces, and simulating the final surface appearance of the part. The latter aspect involves both macroscopic (e.g., wrinkling) and microstructural (e.g., ridging, orange peel) mechanisms of surface plasticity. Related applications are encountered in the field of automotive crash simulations.

Rendering continuum-type forming simulations scientifically sound, predictive at the microstructure scale, in good accord with experiment, and at the same time economically rewarding requires that the materials involved are properly specified in terms of their respective constitutive behavior. For this purpose, finite element simulations can use four sets of material input data, covering hardening, forming limits, surface roughness development, and anisotropy (**Figure 2**).

Engineering of Tensorial Materials

The yield surface represents the generalization of the yield point from uniaxial tensile testing to general stress states. Expanding the yield point into a closed yield surface is required only if the material under inspection shows elastic–plastic anisotropy, that is, if it deforms differently in different directions. Such a behavior is the rule and not the exception in real materials. Strong elastic–plastic anisotropy is encountered, for instance, in metals, composites, and polymers.

The origin of elastic–plastic anisotropy in these materials is the crystalline arrangement of the atoms and the topological arrangement of the phases. Metallic matter, which is the typical material encountered in deformation processing, occurs in polycrystalline form, where each grain has a different crystallographic orientation, shape, and volume fraction. The distribution of the orientations is referred to as crystallographic texture. The anisotropy of the elastic tensor and the discreteness of crystallographic slip along certain lattice directions on preferred crystal planes also entails an anisotropic integral response of such specimens during loading. While the

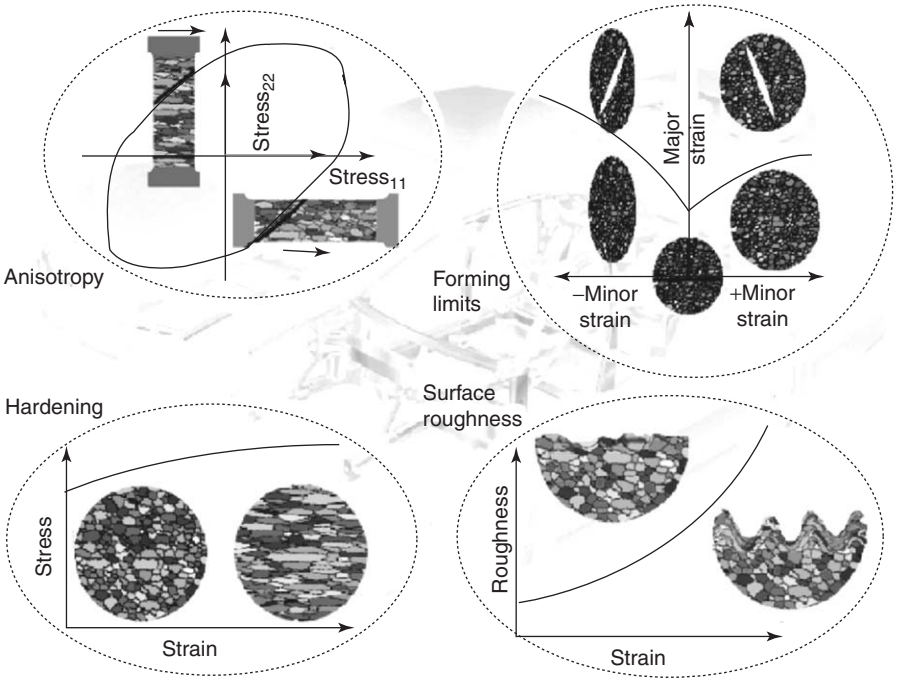


Figure 2 Modern finite element approaches that aim at simulating realistic forming operations typically require material input data about strain hardening, failure (forming limits), crystallographic and morphological anisotropy, and surface roughness.

deformation of single crystals is now well understood, plasticity of polycrystals is less clear. This is essentially due to the intricate elastic–plastic interactions occurring during co-deformation among the individual anisotropic crystals. This interaction leads to strong heterogeneity in terms of strain, stress, and crystal orientation. Another difficulty in the anisotropy of polycrystals lies in the fact that crystals rotate during forming, owing to the skew-symmetric portion of the displacement gradients created by crystal slip. This means that texture and anisotropy change during forming. In this context, it must be underlined that crystallographic orientation changes are principally nonreversible owing to the orientation sensitivity of strain path changes and the orientation dependence of strain hardening.

Elastic Anisotropy

Elastic anisotropy of crystals departs from the directionality of the electronic bond and the resulting crystalline structure. For small deviations of atoms from their equilibrium positions, the reversible elastic response to loads can be approximated by a linear relationship, which is referred to as Hooke's law. In this framework, the linear elastic constants can be derived as the components of the second derivative of the electronic potential. The elastic constants can be written in the form of a fourth-rank elastic stiffness tensor C_{ijkl} or in the form of a fourth-rank elastic compliance tensor S_{ijkl} according to

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl}, \quad \varepsilon_{ij} = S_{ijkl}\sigma_{kl}$$

Symmetry relations and thermodynamic considerations reduce the 81 elastic constants to a set of 3 (C_{1111}, C_{1122}, C_{2323}) in the case of cubic crystal symmetry (which corresponds to C_{11}, C_{12}, C_{44} in reduced matrix notation) and to a set of five independent numbers (C_{1111}, C_{1122}, C_{1133}, C_{3333}, C_{2323}) in the case of hexagonal crystal symmetry (which corresponds to C_{11}, C_{12}, C_{13}, C_{33}, C_{44} in reduced matrix notation).

Plastic Anisotropy

The plastic anisotropy of crystalline matter also departs from the directionality of the electronic bond and the resulting crystal lattice structure. Both aspects determine which slip planes and translation vectors (Burgers vectors) are important for the motion of lattice dislocations or the activation of plastically relevant athermal transformations. The main consequence of this anisotropy in the present context is that metals are deformed in a discrete rather than in a continuum fashion, rendering plasticity an intrinsically anisotropic property of metals. In the simple case of plastic isotropy, each arbitrary stress state imposed on a sample will entail plastic yielding at the same value. The sum of all these possible stress states at which plastic deformation is sustained is referred to as the yield surface. In cases where the material is anisotropic, the yield surface deviates from a spherical shape.

The shape of the yield surface can be understood in terms of the activation of single-slip systems. Assuming that the Burgers vectors b_j and the slip plane normals n_i of the s different slip systems available in a particular crystal lattice are known, their orientation factors m_{ij} are (**Figure 3**)

$$m_{ij}^{s} = n_i^s b_j^s$$

with the symmetric parts

$$m_{ij}^{\mathrm{sym},s} = \tfrac{1}{2}(n_i^s b_j^s + n_j^s b_i^s) \ \text{(crystal coordinates)}$$

when given in crystal coordinates. All vectors used in the equations are normalized. Transforming the latter equation into the sample coordinate system leads to

$$m_{kl}^{\mathrm{sym},s} = \tfrac{1}{2}(a_{ki}n_i^s\, a_{lj}b_j^s + a_{lj}n_j^s\, a_{ki}b_i^s)$$
(sample coordinates)

where a_{ki} and a_{lj} are the transformation matrices between the crystal coordinate system and the sample coordinate system. Using the s different orientation factors, $m_{kl}^{\mathrm{sym},s}$ of the s different available slip systems for the transformation of an external load into the slip geometry provides a kinematic formulation for the yield surface of a single crystal. Positions below the yield surface result in no plastic deformation on the assigned slip systems, and deformation begins

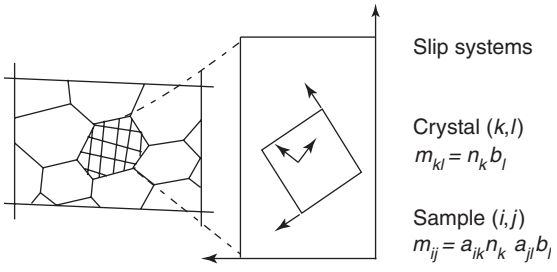


Figure 3 The plastic anisotropy of crystalline matter departs from the directionality of the electronic bond and the resulting crystal lattice structure. Both aspects determine the slip planes and translation vectors on which crystal lattice dislocations move during plastic deformation. The diagram shows the different coordinate system and the resulting relevant geometrical transformation operations.

at the lines corresponding to specific slip systems.

$$m_{kl}^{\mathrm{sym},s}\sigma_{kl} = \tau_{\mathrm{crit},+}^{s,\mathrm{active}}$$

$$m_{kl}^{\mathrm{sym},s}\sigma_{kl} = \tau_{\mathrm{crit},-}^{s,\mathrm{active}}$$

for the active slip systems, and

$$m_{kl}^{\mathrm{sym},s}\sigma_{kl} < \tau_{\mathrm{crit},+}^{s,\mathrm{nonactive}}$$

$$m_{kl}^{\mathrm{sym},s}\sigma_{kl} < \tau_{\mathrm{crit},-}^{s,\mathrm{nonactive}}$$

for the nonactive slip systems. One must note that the Einstein summation rule applies in all equations if not stated otherwise.

Most points on the single-crystal yield surface describe single-slip conditions. In the graphical representation of the yield surface, single-slip generally takes place when the stress tensor (in vector notation, using the tensor–vector transformation rule $\sigma_\lambda^{\mathrm{T}} = \{1/\sqrt{6}(2\sigma_{33} - \sigma_{11} - \sigma_{22}),\ 1/\sqrt{2}(\sigma_{22} - \sigma_{11}),\ \sqrt{2}\sigma_{23},\ \sqrt{2}\sigma_{13}, \sqrt{2}\sigma_{12},\}$) points at one of the planes rather than at one of the corners (see **Figures 4–6**).

Polyslip conditions, usually required for polycrystal deformation owing to the satisfaction of strain rate compatibility among the grains, are characterized by conus coordinates of the stress state. The conus positions for the stress can be calculated using a conventional polycrystal homogenization approach, for instance, the Taylor–Bishop–Hill theory of strain-rate homogeneity (indicated by σ^{TBH} in **Figure 6**). Such polycrystal models represent smoothing approaches that handle the compatibility of stress or strain required for a polycrystal to deform coherently. The corresponding multislip positions of the stress tensor, satisfying an externally imposed strain rate, are then denoted as Taylor positions. It must be noted, in this context, that the Taylor factor generally takes the form of a stress shape tensor for the crystal yield surface rather than that of a factor owing to its dependence on the strain rate tensor. Its magnitude for a given strain rate determines the kinematic size of the yield surface in the corresponding stress direction characterizing the correct polyslip conus and thus the kinematic portion of the corresponding stress state. The symbols $D^{s=1}$ and $D^{s=2}$ in the figure indicate the single-slip strain states from slip systems 1 and 2. Using these two slip systems allows one to realize any strain rate state in the respective conus by a linear combination of $D^{s=1}$ and $D^{s=2}$. For cubic crystals the yield surface reveals four classes of Taylor states for polyslip and one for single slip. These yield states are referred to as

penta-slip state (5 active slip systems):

$^5M_{pq}^{i}$ (f.c.c., b.c.c. (reduced): $i = 56$)

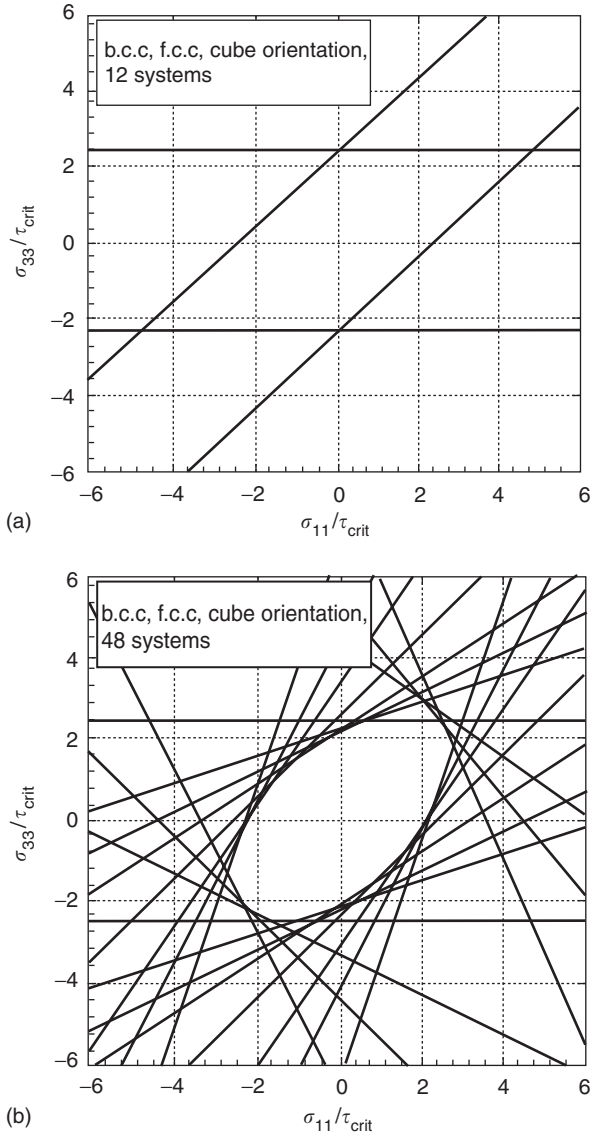


Figure 4 A simple Schmid-type formulation considering the different orientation factors of all available slip systems which essentially transforms an external load into shear stresses acting on the slip systems provides a kinematic formulation for the yield surface of a single crystal. The yield surface shown in (a) was derived using the 12 {1 1 0}⟨1 1 1⟩ slip systems. The yield surface shown in (b) was derived using the 12 {1 1 0}⟨1 1 1⟩, 12 {1 1 2}⟨1 1 1⟩, and 24 {1 2 3}⟨1 1 1⟩ slip systems (b.c.c.). The figure indicates that b.c.c. alloys, therefore, behave plastically principally different from f.c.c. alloys.

tetra-slip (4 active slip systems):

$^4M_{pq}^{j}$ (f.c.c., b.c.c. (reduced): $j = 108$)

tri-slip (3 active slip systems):

$^3M_{pq}^{k}$ (f.c.c., b.c.c. (reduced): $k = 135$)

bi-slip (2 active slip systems):

$^2M_{pq}^{l}$ (f.c.c., b.c.c. (reduced): $l = 66$)

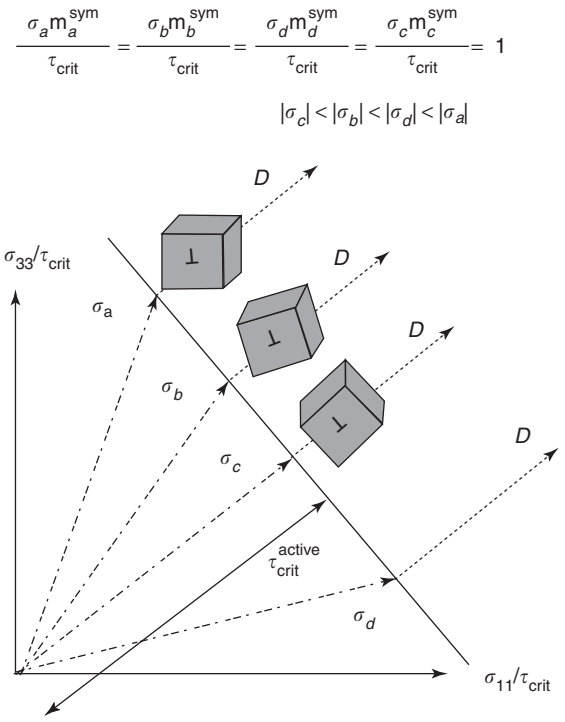


Figure 5 Most points on the single-crystal yield surface describe single-slip conditions. In the graphical representation of the yield surface, single slip generally takes place when the stress state (in vector notation) points at a hyperplane rather than a hyperconus. The strain rate tensor is indicated by ***D*** and *m* is the Schmid factor, that is, the dyadic product of the slip elements. The small cubes placed in the figure indicate the changing relative orientation between the external reference system and the crystal coordinate system.

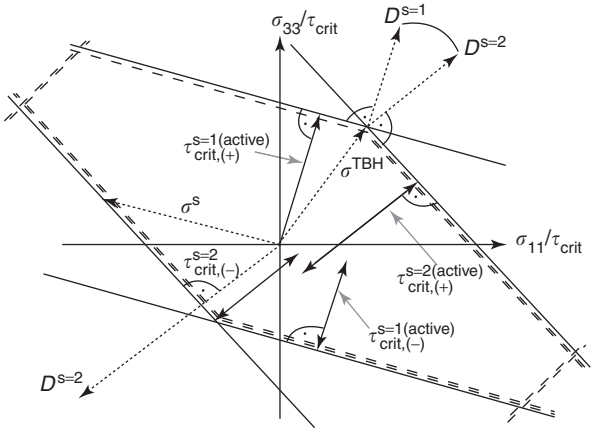


Figure 6 Polycrystal deformation requires polyslip conditions in order to satisfy strain rate compatibility among the grains. Polyslip states are crystallographically characterized by hyperconus coordinates of the stress state. The conus positions for the stress can be calculated using a conventional homogenization approach, for instance, the Taylor–Bishop–Hill theory (indicated by σ^{TBH}). The symbols $D^{s=1}$ and $D^{s=2}$ indicate the single-slip strain states from slip systems 1 and 2. Using these two slip systems allows one to realize any strain rate state in the respective conus by a linear combination of $D^{s=1}$ and $D^{s=2}$.

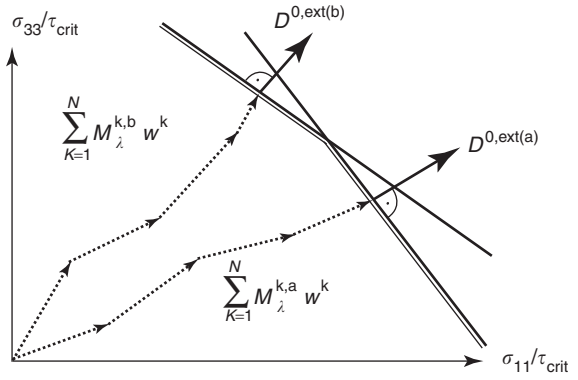


Figure 7 The Taylor stress state for a polycrystalline aggregate can, for a given external strain rate state, be integrated as a volume-weighted sum of all Taylor factors derived separately for each grain for the respective boundary condition. In the figure, *M* is the Taylor tensor, *D* the strain rate and *w* the volume fraction. The counter *k* sums over all crystals in the aggregate.

$$\text{single-slip (1 active slip system):}\quad {}^{2}M_{pq}^{n}\ (\text{f.c.c., b.c.c. (reduced):}\ n = 24)$$

where f.c.c. denotes face-centered cubic and b.c.c. denotes body-centered cubic crystal structure. The term "reduced" indicates that only the first 12 $\{1\,1\,1\}\langle 1\,1\,0\rangle$ b.c.c. slip systems have been considered here. The number at the end of each row gives the number of different conus cases (and single-slip cases) for the respective Taylor state. The total Taylor stress state for a polycrystalline aggregate can, for a given external strain rate state, then be integrated as a volume-weighted sum of all Taylor tensors derived separately for each grain for this boundary condition (**Figure 7**).

Empirical Approximations of the Yield Surface

The first empirical mathematical description of an anisotropic plastic yield surface was suggested in 1928 by von Mises in the form of a quadratic function. This approach, which was originally designed to empirically approximate the plastic anisotropy of single crystals, was in 1948 rendered by Hill into a generalized form using the Huber–Mises–Hencky approach. In Hill's form the yield surface amounts to

$$\begin{aligned} f(\sigma_{ij}) = (&F(\sigma_{22}-\sigma_{33})^2 \\ &+ G(\sigma_{33}-\sigma_{11})^2 + H(\sigma_{11}-\sigma_{22})^2 \\ &+ 2L\sigma_{23}^2 + 2M\sigma_{13}^2 + 2N\sigma_{12}^2)^{1/2} \end{aligned}$$

where *F*, *G*, *H*, *L*, *M*, and *N* are the anisotropy coefficients. The above equation can be

rewritten as

$$f(S_{ij}) = ((G+H)S_{11}^2 + (F+H)S_{22}^2 + (F+G)S_{33}^2 - 2HS_{11}S_{22} - 2GS_{11}S_{33} - 2FS_{22}S_{33} + 2LS_{23}^2 + 2MS_{13}^2 + 2NS_{12}^2)^{1/2}$$

where S_{ij} are the deviatoric stress components. The shape coefficients of Hill's quadratic yield function can be fitted from experimentally obtained mechanical data such as the Lankford values taken in different directions of a specimen. Scaling can be provided by the yield stress obtained from uniaxial tensile testing. While the Lankford coefficients and the yield stress can be determined from tensile testing, the direct measurement of mechanical response under complex loads is an intricate task. Although Hill-based anisotropy simulations (referring to the 1948 model by Hill) provide good approximations at least of the initial plastic anisotropy in some cases, it is less applicable when anisotropy is characterized by a sixfold slip symmetry due to its simple quadratic form.

Due to this principal shortcoming, a number of optimized empirical anisotropic yield surface concepts with higher-order polynomial forms have been proposed, such as those introduced later by Hill and by Barlat, that are better suited for f.c.c. alloys and many b.c.c. steels. In recent years, various authors have presented improved empirical yield surface approaches, where the yield function can be fitted using both mechanically obtained and even texture-based data.

The chief advantage of using an empirical anisotropic yield surface function as a constitutive law in metal-forming finite element simulations is time efficiency, and the simple mechanical methods with which it can be derived. The main disadvantage of empirical yield surface functions is that the anisotropy of polycrystalline matter generally changes during forming owing to the change of texture. This evolution of anisotropy is not mapped by a corresponding change of the shape of the yield surface. In other words, the same yield surface shape is used throughout a given finite element simulation without making a physically meaningful update of its steadily changing shape. Although empirical constitutive laws can be used to gradually change the yield surface shape during forming, their capability is typically constrained by a lack of physical information about the actual development of the crystallographic texture during forming.

Crystallographic Approximations of Elastic Anisotropy

An important problem in the field of deformation simulation is the approximation of the elastic response of a sample under load. The macroscopic elastic properties of a textured polycrystal can be calculated by formulating appropriate volume-weighted means of the individual elastic single-crystal tensor, rotated parallel to the respective local coordinate system of each individual crystal. This average value of the integral elastic tensor must therefore take into account all individual orientations of the grains which are described by the orientation distribution function.

An early homogenization approach for the elastic response under an external load was suggested by Voigt, who assumed that, in the case of a macroscopically prescribed strain rate state, each material portion is in the same strain rate state as the entire sample, irrespective of its spatial position in the specimen. The strain rate would then be homogeneous throughout the sample. However, in a polycrystalline sample, the elastic response typically varies from grain to grain, due to the spatially changing crystal orientation. Since in the Voigt model the prescribed strain rate is the same everywhere in the sample, the stress must vary. The Voigt limit for the elastic response of a polycrystalline sample can thus be calculated by weighting the tensor of the elastic stiffness as a function of orientation with the orientation distribution function. A different approach to treating the homogenization problem in an elastically loaded polycrystalline sample was suggested by Reuss. He suggested that in the case of a macroscopically prescribed stress state, each material portion is in the same stress state irrespective of its spatial position in the specimen. The stress would then be homogeneous throughout the specimen. The elastic response may then vary from grain to grain, in accord with the local orientation of the crystal. Since in the Reuss model the prescribed external stress is constant throughout the specimen, the strain must vary according to the local grain orientation. Consequently, the elastic Reuss limit can be calculated for a polycrystal by weighting the tensor of the elastic compliance as a function of orientation with the orientation distribution function. Since neither the Voigt nor the Reuss method provides reliable approximations to the elastic modulus of a polycrystal, Hill defined an average modulus that consists of the equally weighted results of both the above models.

Crystallographic Approximations of the Yield Surface

Materials subjected to forming operations develop or inherit morphological textures (e.g., elongated grains, chemical segregation, or second phases with elongated topology entailing directional effects) as

well as crystallographic textures (orientation distribution of the crystallites constituting polycrystalline matter). Both types of textures can directly serve as input data for the calculation of the yield surface shape using the Taylor–Bishop–Hill (crystallography) or self-consistent type (crystallography, morphology) approaches.

The experimental input data for deriving the yield surface from the crystallographic texture of polycrystals can be determined using X-ray, neutron, or electron diffraction experiments. Since texture-based yield surface approximations use the complete crystallographic anisotropy information of a specimen, they are often superior to empirical approaches which rely on a small set of mechanical parameters. However, modern empirical approaches for the approximation of the yield surface typically use the Taylor–Bishop–Hill type crystal simulations on the basis of experimental texture data to provide tangents for a better fit of the yield surface functions.

Besides its clear physical basis, another advantage of crystallographic yield surface approximations lies in its capability to incorporate both kinematical and kinetic plasticity effects. In this context, it must be considered that the crystallographic texture gives only the respective anisotropic shape function for a particular polycrystalline sample, but the texture dependence of the internal stress and the individual hardness of the different grains are typically ignored by the constitutive laws employed in homogenization approaches. However, it is principally feasible to generalize the crystallographic yield surface concept by enriching it with the individual strength of each grain. This leads to a formulation of the following kind:

$$f(S_{ij}) = \frac{1}{V}\int_V M_{ij}(\boldsymbol{g}, D_{ij})\tau_{\text{crit}}(D_{ij}, \boldsymbol{g})\ \mathrm{d}V \approx \sum_{k=1}^{N} M_{ij}^{k}\tau_{\text{crit}}^{k} w^{k}$$

where $f(S_{ij})$ is the yield surface, V the sample volume, M_{ij} the Taylor shape function obtained by homogenization theory as a function of strain rate D_{ij} and rotation matrix $\boldsymbol{g}$, τ_{crit} the flow stress of each individual grain, and w the volume fraction of each grain (**Figure 8**).

Although texture-based yield surface approximations have a crisp physical basis in that they incorporate crystal anisotropy in a genuine fashion, they have the shortcoming of ignoring texture changes during forming. This means that, as far as texture update during forming is concerned, there is basically little difference between the predictive capabilities of empirical and texture-based yield surface approximations, particularly if one considers that recent

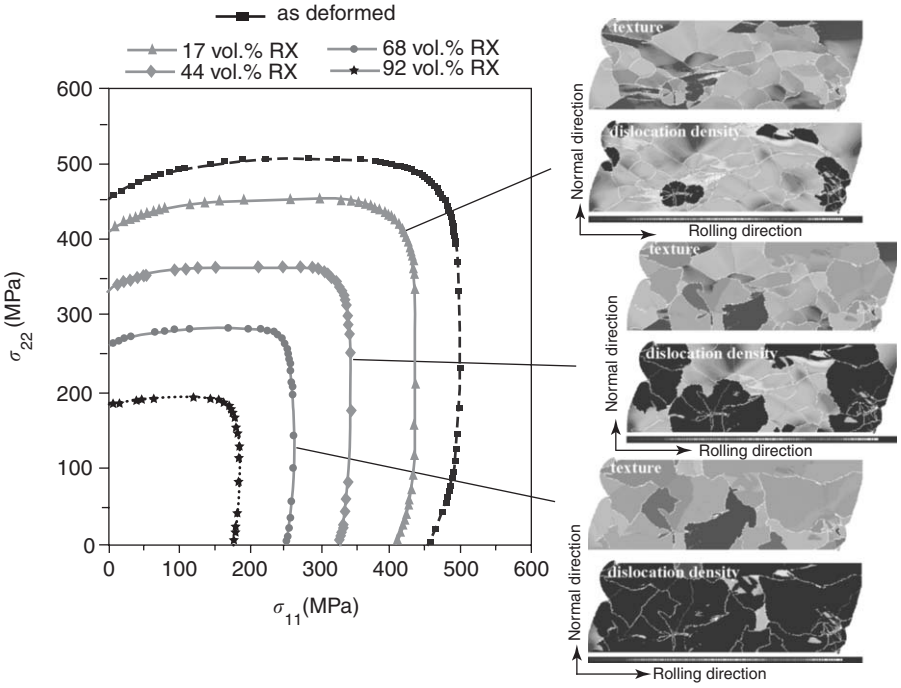


Figure 8 Coupled crystallographic–kinetic yield functions obtained by including both the texture and the texture-dependent flow stress of each individual grain weighted by its respective volume. The left-hand side of the diagram shows a portion of the planar yield surface as it anisotropically shrinks during partial recrystallization. The right-hand side shows three subsequent time steps of a simulation where the upper diagram gives the texture and the lower the dislocation density.

approaches use both mechanical and texture-related information to fit the yield function. These methods could be referred to as hybrid anisotropy yield criteria or semi-empirical yield criteria.

Constitutive Model for Crystal Plasticity Finite Element Simulations

Crystal plasticity phenomena can be directly linked with nonlinear variational formulations. The current approaches in this domain provide a direct means for updating the local crystallographic and hardening state of the material via integration of the evolution equations for the crystal lattice orientation and the critical resolved shear stress. The deformation behavior of the crystal volume elements are, at each integration point, governed by a crystal plasticity model which accounts for discrete plastic deformation by crystallographic slip and for the rotation of the crystal lattice during deformation. The crystal plasticity finite element models typically use space and time as independent variables and the crystal orientation and the accumulated slip as dependent variables (**Figure 9**).

In most large-strain constitutive crystal plasticity models, one assumes the stress response at each macroscopic continuum material point to be potentially given by one crystalline volume point. The constitutive equation for the stress in each grain is then expressed in terms of

$$\boldsymbol{T}^* = \boldsymbol{C}\boldsymbol{E}^*$$

where $\boldsymbol{T}^*$ is the Cauchy stress produced by the crystalline portion, $\boldsymbol{C}$ is the fourth-order elastic tensor, and $\boldsymbol{E}^*$ an elastic strain measure obtained by polar decomposition,

$$\boldsymbol{E}^* = \tfrac{1}{2}(\boldsymbol{F}^{*\mathrm{T}}\boldsymbol{F}^* - 1)$$

which leads to a stress measure, the elastic work conjugate to the strain measure $\boldsymbol{E}^*$,

$$\boldsymbol{T}^* = \boldsymbol{F}^{*-1}(\det(\boldsymbol{F}^*)\boldsymbol{T})(\boldsymbol{F}^*)^{-\mathrm{T}}$$

where $\boldsymbol{T}$ is the symmetrical Cauchy stress tensor in the grain, and $\boldsymbol{F}^*$ is a local elastic deformation gradient defined in terms of the local total deformation gradient $\boldsymbol{F}$ and the local plastic deformation gradient $\boldsymbol{F}^{\mathrm{p}}$. The relation between the elastic and the plastic portion of $\boldsymbol{F}$ amounts to

$$\boldsymbol{F}^* = \boldsymbol{F}(\boldsymbol{F}^{\mathrm{p}})^{-1}, \quad \det(\boldsymbol{F}^*) > 0, \ \det(\boldsymbol{F}^{\mathrm{p}}) = 1$$

The plastic deformation gradient is given by the flow rule

$$\dot{\boldsymbol{F}}^{\mathrm{p}} = \boldsymbol{L}^{\mathrm{p}}\boldsymbol{F}^{\mathrm{p}}$$

with $\boldsymbol{L}^{\mathrm{p}}$ being the plastic flow rate. In crystals, $\boldsymbol{L}^{\mathrm{p}}$ is composed of the superposition of the resolved crystallographic plastic shear rates, $\dot{\gamma}^{\alpha}$, such that

$$\boldsymbol{L}^{\mathrm{p}} = \sum_{\alpha} \dot{\gamma}^{\alpha}\boldsymbol{S}_0^{\alpha}, \quad \boldsymbol{S}_0^{\alpha} = \boldsymbol{m}_0^{\alpha} \otimes \boldsymbol{n}_0^{\alpha}$$

where $\dot{\gamma}^{\alpha}$ is the shearing rate on the slip system α, and $\boldsymbol{m}_0^{\alpha}$ and $\boldsymbol{n}_0^{\alpha}$ denote the slip direction and the slip plane normal of the slip system, respectively, in the initial unloaded configuration. The shearing rate on each slip system depends on the resolved shear τ^{α} and the slip resistance s^{α} of that slip system, and it is taken as

$$\dot{\gamma}^{\alpha} = \dot{\gamma}_0 \left| \frac{\tau^{\alpha}}{s^{\alpha}} \right|^{1/m} \operatorname{sign}(\tau^{\alpha})$$

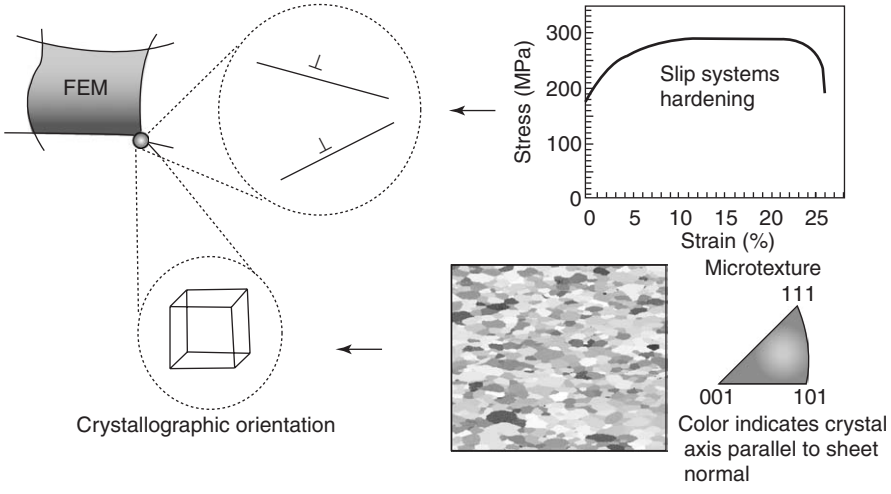


Figure 9 Schematical drawing showing the mapping of a set of orientations (microtexture) on the integration points of a crystal plasticity finite element mesh.

where $\dot{\gamma}_0$ denotes a reference value of the slip rate and m represents the strain rate sensitivity parameter. The evolution of the slip system resistance can be taken as

$$\dot{s} = \sum_{\beta} h^{\alpha\beta} |\dot{\gamma}^{\beta}|, \quad h^{\alpha\beta} = q^{\alpha\beta} h^{(\beta)},$$

$$h^{(\beta)} = h_0 \left(1 - \frac{s^{\beta}}{s_s}\right)^a$$

where $q^{\alpha\beta}$ are the components of a matrix that describe the latent hardening behavior of the crystal, and h_0, a, and s_s are slip system hardening parameters.

See also: Computer Simulation Techniques in Condensed Matter Physics; Crystal Symmetry; Crystal Tensors: Applications; Crystal Tensors: Introduction; Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Mechanical Properties: Tensile Properties; Space Groups.

PACS: 62.20.Fe; 81.20.Hy; 81.05.Bx; 62.20.Qp; 62.20.Dc

Further Reading

Banabic D, Bunge HJ, Pöhlandt K, and Tekkaya AE (2000) *Formability of Metallic Materials*. Berlin: Springer.

Hosford WF (1993) *The Mechanics of Crystals and Textured Polycrystals*. Oxford: Oxford University Press.

Hosford WF and Caddell RM (1995) *Metal Forming–Mechanics and Metallurgy*. 2nd edn. Prentice-Hall.

Kocks UF, Tóme CN and Wenk HR (1998) *Texture and Anisotropy*. Cambridge: Cambridge University Press.

Raabe D (1998) *Computational Materials Science*. Wiley-VCH.

Raabe D, Klose P, Engl B, Imlau KP, Friedel F, *et al*. (2002) *Advanced Engineering Materials* 4: 169.

Raabe D and Roters F (2004) *International Journal of Plasticity* 20: 339.

Density-Functional Theory

S Kurth, M A L Marques, and E K U Gross, Freie Universität Berlin, Berlin, Germany

Introduction

Density-functional theory (DFT) is a successful theory to calculate the electronic structure of atoms, molecules, and solids. Its goal is the quantitative understanding of material properties from the fundamental laws of quantum mechanics.

Traditional electronic structure methods attempt to find approximate solutions to the Schrödinger equation of N interacting electrons moving in an external, electrostatic potential (typically the Coulomb potential generated by the atomic nuclei). However, there are serious limitations of this approach: (1) the problem is highly nontrivial, even for very small numbers N and the resulting wave functions are complicated objects and (2) the computational effort grows very rapidly with increasing N, so the description of larger systems becomes prohibitive.

A different approach is taken in density-functional theory where, instead of the many-body wave function, the one-body density is used as the fundamental variable. Since the density $n(\boldsymbol{r})$ is a function of only three spatial coordinates (rather than the $3N$ coordinates of the wave function), density-functional theory is computationally feasible even for large systems.

The foundations of density-functional theory are the Hohenberg–Kohn and Kohn–Sham theorems which will be reviewed in the following section. In the section "Approximations for the exchange–correlation energy," various levels of approximation to the central quantity of DFT are discussed. The section "Results for some selected systems" will present some typical results from DFT calculations for various physical properties that are normally calculated with DFT methods. The original Hohenberg–Kohn and Kohn–Sham theorems can easily be extended from its original formulation to cover a wide variety of physical situations. A number of such extensions are presented in the section "Extensions of DFT," with particular emphasis on time-dependent DFT.

Hohenberg–Kohn and Kohn–Sham Theorems

In ground-state DFT one is interested in systems of N interacting electrons described by the Hamiltonian

$$\begin{aligned}\hat{H} &= \hat{T} + \hat{V} + \hat{V}_{\mathrm{ee}} \\ &= -\sum_{i=1}^{N} \frac{\nabla_i^2}{2} + \sum_{i=1}^{N} v(\boldsymbol{r}_i) + \frac{1}{2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i=j}}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\end{aligned} \qquad [1]$$

with the kinetic, potential, and interaction energy operators $\hat{T}$, $\hat{V}$, and $\hat{V}_{\mathrm{ee}}$, respectively.

The central statement of formal density-functional theory is the celebrated Hohenberg–Kohn theorem which, for nondegenerate ground states, can be summarized in the following three statements:

1. The ground state electron density $n(\boldsymbol{r})$ of a system of interacting electrons uniquely determines the external potential $v(\boldsymbol{r})$ in which the electrons move and thus the Hamiltonian and all physical properties of the system.
2. The ground-state energy E_0 and the ground-state density $n_0(\boldsymbol{r})$ of a system characterized by the potential $v_0(\boldsymbol{r})$ can be obtained from a variational principle which involves only the density, that is, the ground-state energy can be written as a functional of the density, $E_{v_0}[n]$, which gives the ground-state energy E_0 if and only if the true ground-state density $n_0(\boldsymbol{r})$ is inserted. For all other densities $n(\boldsymbol{r})$, the inequality
$$E_0 = E_{v_0}[n_0] < E_{v_0}[n] \qquad [2]$$
holds.
3. There exists a functional $F[n]$ such that the energy functional can be written as
$$E_{v_0}[n] = F[n] + \int \mathrm{d}^3 r v_0(\boldsymbol{r}) n(\boldsymbol{r}) \qquad [3]$$

The functional $F[n]$ is universal in the sense that, for a given particle–particle interaction (the Coulomb interaction in our case), it is independent of the potential $v_0(\boldsymbol{r})$ of the particular system under consideration, that is, it has the same functional form for all systems.

The proof of the Hohenberg–Kohn theorem is based on the Rayleigh–Ritz variational principle and will not be repeated here. The interested reader is referred to excellent reviews given in the "Further reading" section.

From the Hohenberg–Kohn variational principle, that is, the second statement given above, the ground-state density $n(\boldsymbol{r})$ corresponding to the external potential $v(\boldsymbol{r})$ can be obtained as a solution of the Euler equation
$$\frac{\delta E_v[n]}{\delta n(\boldsymbol{r})} = \frac{\delta F[n]}{\delta n(\boldsymbol{r})} + v(\boldsymbol{r}) = 0 \qquad [4]$$

The formal definition of the Hohenberg–Kohn functional $F[n]$ is well known:
$$\begin{aligned} F[n] &= T[n] + V_{\mathrm{ee}}[n] \\ &= \langle \Psi[n] | \hat{T} | \Psi[n] \rangle + \langle \Psi[n] | \hat{V}_{\mathrm{ee}} | \Psi[n] \rangle \end{aligned} \qquad [5]$$
where $\Psi[n]$ is that N-electron wave function which yields the density n and minimizes the expectation value of $\hat{T} + \hat{V}_{\mathrm{ee}}$. However, the explicit density dependence of $F[n]$ remains unknown. Approximations have been suggested, the oldest one being the well-known Thomas–Fermi approximation (which precedes the Hohenberg–Kohn theorem historically). Unfortunately, the accuracy of known approximations to $F[n]$ is rather limited in practical calculations. Therefore, eqn [4] is rarely used in electronic structure calculations today.

Instead, the Hohenberg–Kohn theorem provides the basic theoretical foundation for the construction of an effective single-particle scheme which allows for the calculation of the ground-state density and energy of systems of interacting electrons. The resulting equations, the so-called Kohn–Sham equations, are at the heart of modern density-functional theory. They have the form of the single-particle Schrödinger equation
$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{s}}(\boldsymbol{r}) \right] \varphi_i(\boldsymbol{r}) = \varepsilon_i \varphi_i(\boldsymbol{r}) \qquad [6]$$

The density can then be computed from the N single-particle orbitals occupied in the ground-state Slater determinant
$$n(\boldsymbol{r}) = \sum_i^{\mathrm{occ}} |\varphi_i(\boldsymbol{r})|^2 \qquad [7]$$

The central idea of the Kohn–Sham scheme is to construct the single-particle potential $v_{\mathrm{s}}(\boldsymbol{r})$ in such a way that the density of the auxiliary noninteracting system equals the density of the interacting system of interest. To this end one partitions the Hohenberg–Kohn functional in the following way:
$$F[n] = T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{xc}}[n] \qquad [8]$$
where $T_{\mathrm{s}}[n]$ is the noninteracting kinetic energy,
$$U[n] = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \qquad [9]$$
is the classical electrostatic energy of the charge distribution $n(\boldsymbol{r})$, and $E_{\mathrm{xc}}[n]$ is the so-called exchange–correlation energy which is formally defined by
$$E_{\mathrm{xc}}[n] = T[n] + V_{\mathrm{ee}}[n] - U[n] - T_{\mathrm{s}}[n] \qquad [10]$$

From the above definitions one can derive the form of the effective potential entering eqn [6]
$$v_{\mathrm{s}}[n](\boldsymbol{r}) = v(\boldsymbol{r}) + \int \mathrm{d}^3 r' \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} + v_{\mathrm{xc}}[n](\boldsymbol{r}) \qquad [11]$$

where the exchange–correlation potential v_{xc} is defined by

$$v_{xc}[n](\boldsymbol{r}) = \frac{\delta E_{xc}[n]}{\delta n(\boldsymbol{r})} \qquad [12]$$

Since $v_s[n](\boldsymbol{r})$ depends on the density, eqns [6], [7], and [11] have to be solved self-consistently. This is known as the Kohn–Sham scheme of density-functional theory.

Approximations for the Exchange–Correlation Energy

Clearly, the formal definition [10] of the exchange–correlation energy is not helpful for practical calculations and one needs to use an approximation for this quantity. Some of these approximations are discussed in the following.

While DFT itself does not give any hint on how to construct approximate exchange–correlation functionals, it holds both the promise and the challenge that the true E_{xc} is a universal functional of the density, that is, it has the same functional form for all systems. On the one hand this is a promise because an approximate functional, once constructed, may be applied to any system of interest. On the other hand, this is a challenge because a good approximation should perform equally well for very different physical situations.

Both the promise and the challenge are reflected by the fact that the simplest of all functionals, the so-called local density approximation (LDA), or its spin-dependent version, the local spin density approximation (LSD), has remained the approximation of choice for quite many years after the formulation of the Kohn–Sham theorem. In LDA, the exchange–correlation energy is given by

$$E_{xc}^{LDA}[n] = \int d^3r n(\boldsymbol{r}) e_{xc}^{unif}(n(\boldsymbol{r})) \qquad [13]$$

where $e_{xc}^{unif}(n)$ is the exchange–correlation energy per particle of an electron gas with spatially uniform density n. It can be obtained from quantum Monte Carlo calculations and simple parametrizations are available. By its very construction, the LDA is expected to be a good approximation for spatially slowly varying densities. Although this condition is hardly ever met for real electronic systems, LDA has proved to be remarkably accurate for a wide variety of systems.

In the quest for improved functionals, an important breakthrough was achieved with the emergence of the so-called generalized gradient approximations (GGAs). Within GGA, the exchange–correlation energy for spin-unpolarized systems is written as

$$E_{xc}^{GGA}[n] = \int d^3r f(n(\boldsymbol{r}), \nabla n(\boldsymbol{r})) \qquad [14]$$

While the input e_{xc}^{unif} in LDA is unique, the function f in GGA is not, and many different forms have been suggested. When constructing a GGA one usually tries to incorporate a number of known properties of the exact functional into the restricted functional form of the approximation. The impact of GGAs has been quite dramatic, especially in quantum chemistry where DFT is now competitive in accuracy with more traditional methods while being computationally less expensive.

In recent years, still following the lines of GGA development, a new class of "meta-GGA" functionals has been suggested. In addition to the density and its first gradient, meta-GGA functionals depend on the kinetic energy density of the Kohn–Sham orbitals,

$$\tau(\boldsymbol{r}) = \frac{1}{2}\sum_{i}^{occ} |\nabla \varphi_i(\boldsymbol{r})|^2 \qquad [15]$$

The meta-GGA functional then takes the form

$$E_{xc}^{MGGA}[n] = \int d^3r\, g(n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \tau(\boldsymbol{r})) \qquad [16]$$

The additional flexibility in the functional form gained by the introduction of the new variable can be used to incorporate more of the exact properties into the approximation. In this way, it has been possible to improve upon the accuracy of GGA for some physical properties without worsening the results for others.

Unlike LDA or GGA, which are explicit functionals of the density, meta-GGAs also depend explicitly on the Kohn–Sham orbitals. It is important to note, however, that one is still in the domain of DFT since, through the Kohn–Sham equation [6], the orbitals are functionals of the Kohn–Sham potential and therefore, by virtue of the Hohenberg–Kohn theorem, also functionals of the density.

Such orbital functionals, or implicit density functionals, constitute a wide field of active research. Probably the most familiar orbital functional is the exact exchange energy functional (EXX):

$$E_x^{EXX}[n] = -\frac{1}{2}\sum_{\sigma}\int d^3r \int d^3r' \times \sum_{i,j}^{occ} \frac{\varphi_{i\sigma}(\boldsymbol{r})\varphi_{i\sigma}^*(\boldsymbol{r}')\varphi_{j\sigma}(\boldsymbol{r}')\varphi_{j\sigma}^*(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|} \qquad [17]$$

At this point, some remarks about similarities and differences to Hartree–Fock theory are in order. If

one uses the exact exchange functional [17] and neglects correlation, the resulting total energy functional is exactly the Hartree–Fock functional. However, in exact-exchange DFT this functional is evaluated with Kohn–Sham rather than Hartree–Fock orbitals. Furthermore, the Hartree–Fock orbitals are obtained from a minimization without constraints (except for orthonormality) which leads to the nonlocal Hartree–Fock potential. The exact-exchange Kohn–Sham orbitals, on the other hand, are obtained by minimizing the same functional under the additional constraint that the single-particle orbitals come from a local potential. Therefore, the total Hartree–Fock energy is always lower than the total energy in exact-exchange DFT but the difference has turned out to be very small. However, single-particle properties such as orbitals and orbital eigenvalues can be quite different in both approaches. For example, Kohn–Sham exact-exchange energy spectra are usually much closer to experiment than Hartree–Fock spectra, indicating that exact-exchange DFT might be a better starting point to include correlation than Hartree–Fock spectra.

In order to go beyond exact-exchange DFT, one might be tempted to use the exact-exchange functional in combination with a traditional approximation such as LDA or GGA for correlation. Since both LDA and GGA benefit from a cancelation of errors between their exchange and correlation pieces, an approach using LDA or GGA for only one of these energy components is bound to fail. The obvious alternative is to use approximate, orbital-dependent correlation energy functionals. One systematic way to construct such functionals is known as Görling–Levy perturbation theory. Structurally, this is similar to what is known as Møller–Plesset perturbation theory in quantum chemistry. However, it is only practical for low orders. A more promising route uses the fluctuation–dissipation theorem which establishes a connection to linear response theory in time-dependent DFT.

A final class of approximations to the exchange–correlation energy are the so-called hybrid functionals which mix a fraction of exact exchange with GGA exchange,

$$E_{\mathrm{x}}^{\mathrm{HYB}}[n] = aE_{\mathrm{x}}^{\mathrm{EXX}}[n] + (1-a)E_{\mathrm{x}}^{\mathrm{GGA}}[n] \qquad [18]$$

where a is the (empirical) mixing parameter. This exchange functional is then combined with some GGA for correlation. Hybrid functionals are tremendously popular and successful in quantum chemistry but much less so in solid-state physics.

This last fact highlights a problem pertinent to the construction of improved functionals: available approximations are already very accurate and hard to improve upon. In addition, and this makes it a very difficult problem, one would like to have improved performance not only for just one particular property or one particular class of systems but for as many properties and systems as possible. After all, the true exchange–correlation energy is a universal functional of the density.

Results for Some Selected Systems

Over the years, a vast number of DFT calculations have been reported and only the scantiest of selections can be given here. Results are given for different functionals in order to illustrate the performance of the different levels of approximation. It should be kept in mind, however, that while there is only one LDA functional, there are several GGA, meta-GGA, or hybrid functionals which have been suggested and whose results for a given system will vary.

In solid-state physics, typical quantities calculated with ground-state DFT are structural properties. In **Table 1**, equilibrium lattice constants and bulk

Table 1 Equilibrium lattice constants a (in Å) and bulk moduli B_0 (in GPa) for some solids calculated with LSD, GGA, and meta-GGA approximations in comparison to experimental results. m.a.e. is the mean absolute error[a]

Solid	a^{LSD}	a^{GGA}	a^{MGGA}	a^{exp}	B_0^{LSD}	B_0^{GGA}	B_0^{MGGA}	B_0^{exp}
Na	4.05	4.20	4.31	4.23	9.2	7.6	7.0	6.9
NaCl	5.47	5.70	5.60	5.64	32.2	23.4	28.1	24.5
Al	3.98	4.05	4.02	4.05	84.0	77.3	90.5	77.3
Si	5.40	5.47	5.46	5.43	97.0	89.0	93.6	98.8
Ge	5.63	5.78	5.73	5.66	71.2	59.9	64.6	76.8
GaAs	5.61	5.76	5.72	5.65	74.3	60.7	65.1	74.8
Cu	3.52	3.63	3.60	3.60	191	139	154	138
W	3.14	3.18	3.17	3.16	335	298	311	310
m.a.e.	0.078	0.051	0.043	—	12.8	7.0	7.6	—

[a] Lattice constants from Perdew JP, Kurth S, Zupan A, and Blaha P (1999) *Physical Review Letters* 82: 2544, ibid. 82, 5179 (1999) (E); bulk moduli from Kurth S, Perdew J P, and Blaha P (1999) *International Journal of Quantum Chemistry* 75: 889.

moduli for a few solids are presented. LDA lattice constants are usually accurate to within a few percent. On average, GGA and meta-GGA give some improvement. For bulk moduli, the relative errors are much larger and improvement of GGA and meta-GGA as compared to LDA is statistical and not uniform. There are however some properties that are systematically improved by GGA. For example, GGA cohesive energies for transition metals are on average better by almost half an order of magnitude than the LDA values.

Possibly the best known success of GGA is the correct prediction of the ferromagnetic b.c.c ground state of iron for which LDA incorrectly gives a nonmagnetic f.c.c. ground state. The calculated magnetic moments shown in **Table 2** illustrate the improved performance of GGA also for this quantity.

One of the standard uses of DFT in solid-state physics is the calculation of band structures. Usually one interprets and compares the Kohn–Sham band structure directly with experimental energy bands. Strictly speaking, this interpretation has no sound theoretical justification since the Kohn–Sham eigenvalues are only auxiliary quantities. Experience has shown, however, that LDA or GGA band structures are often rather close to experimental ones, especially for simple metals. On the other hand, a well-known problem in semiconductors and insulators is that both LDA and GGA seriously underestimate the bandgap. This deficiency is corrected when using the exact-exchange functional (see **Table 3**).

Table 2 Magnetic moments (in μ_B) for some ferromagnetic solids[a]

Solid	M^{LSD}	M^{GGA}	M^{exp}
Fe	2.01	2.32	2.22
Co	1.49	1.66	1.72
Ni	0.60	0.64	0.61

[a] From Cho JH and Scheffler M (1996) *Physical Review B* 53: 10685.

Table 3 Fundamental bandgaps (in eV) of some semiconductors and insulators

Solid	LSD	EXX	exp
Ge[a]	–0.05	0.99	0.74
Si[a]	0.66	1.51	1.2
C[a]	4.17	4.92	5.50
GaAs[b]	0.49	1.49	1.52
AlN[a]	3.2	5.03	5.11
BeTe[b]	1.60	2.47	2.7
MgSe[b]	2.47	3.72	4.23

[a] Aulbur WG, Städele M, and Görling A (2000) *Physical Review B* 62: 7121.
[b] Fleszar A (2001) *Physical Review B* 64: 245204.

Since DFT also has gained a tremendous popularity in the world of quantum chemistry, finally some results for the atomization energies of small molecules are presented in **Table 4**. Similar to the cohesive energy of solids, one sees a clear improvement as one goes from LDA to GGA. Further improvement is achieved on the meta-GGA level but, for this particular quantity, the accuracy of the hybrid functional is truly remarkable. Unfortunately, however, the accuracy of hybrid functionals for molecular properties does not translate into an equally good performance for solids.

Extensions of DFT

The original theory of Hohenberg, Kohn, and Sham can only be applied to quantum systems in their (nondegenerate) ground-state. Over the years a number of extensions have been put forth (with varying degree of success). In fact, density-functional theories have been formulated for relativistic systems, superconductors, ensembles, etc. Here, only a small selection of extensions of DFT is reviewed. Special attention is paid to time-dependent DFT, as this theory is becoming particularly popular in the calculation of excited-state properties.

Spin Density-Functional Theory

Most practical applications of DFT make use of an extension of the original theory which uses the partial densities of electrons with different spin σ as independent variables,

$$n_\sigma(\boldsymbol{r}) = \sum_i^{\text{occ}} |\varphi_{i\sigma}(\boldsymbol{r})|^2 \qquad [19]$$

Table 4 Atomization energies (in units of $1\,\text{kcal}\,\text{mol}^{-1} = 0.0434\,\text{eV}$) of 20 small molecules. Zero-point vibration has been removed from experimental energies[a]

Molecule	ΔE^{LSD}	ΔE^{GGA}	ΔE^{MGGA}	ΔE^{HYB}	ΔE^{exp}
H_2	113.2	104.6	114.5	107.5	109.5
LiH	61.0	53.5	58.4	54.7	57.8
CH_4	462.3	419.8	421.1	420.3	419.3
NH_3	337.3	301.7	298.8	297.5	297.4
H_2O	266.5	234.2	230.1	229.9	232.2
CO	299.1	268.8	256.0	256.5	259.3
N_2	267.4	243.2	229.2	226.5	228.5
NO	198.7	171.9	158.5	154.3	152.9
O_2	175.0	143.7	131.4	125.6	120.5
F_2	78.2	53.4	43.2	37.1	38.5
m.a.e	34.3	9.7	3.6	2.1	—

[a] LSD, GGA, and meta-GGA results from Kurth S, Perdew JP, and Blaha P (1999) *International Journal of Quantum Chemistry* 75: 889; results for the hybrid functional from Becke AD (1998) *Journal of Chemical Physics* 98: 5648.

rather than using the total density of eqn [7]. Again, the single-particle orbitals are solutions of a noninteracting Schrödinger equation [6], but the effective potential eqn [11] now becomes spin-dependent by using the exchange–correlation potential

$$v_{\mathrm{xc}\sigma}[n_\uparrow, n_\downarrow](\boldsymbol{r}) = \frac{\delta E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]}{\delta n_\sigma(\boldsymbol{r})} \qquad [20]$$

instead of the one given in eqn [12]. Spin-density-functional theory (SDFT) relates to the original formulation of DFT as spin-unrestricted Hartree–Fock relates to spin-restricted Hartree–Fock theory. It allows for a straightforward description of systems with a spin-polarized ground state such as ferromagnets.

Density-Functional Theory for Superconductors

In 1988, triggered by the remarkable discovery of the high-T_c materials, Oliveira, Gross, and Kohn proposed a density-functional theory for the superconducting state. Their theory uses two independent variables – the "normal" density, $n(\boldsymbol{r})$, and a nonlocal "anomalous" density $\chi(\boldsymbol{r}, \boldsymbol{r}')$. This latter quantity reduces, under the appropriate limits, to the order-parameter in Ginzburg–Landau theory. With these two densities it is straightforward to derive a Hohenberg–Kohn-like theorem and a Kohn–Sham scheme. The resulting equations are very similar to the traditional Bogoliubov–de Gennes equations for inhomogeneous superconductors, but include exchange and correlation effects through a "normal,"

$$v_{\mathrm{xc}}[n, \chi](\boldsymbol{r}) = \frac{\delta E_{\mathrm{xc}}[n, \chi]}{\delta n(\boldsymbol{r})} \qquad [21]$$

and an "anomalous" exchange–correlation potential,

$$\Delta_{\mathrm{xc}}[n, \chi](\boldsymbol{r}, \boldsymbol{r}') = -\frac{\delta E_{\mathrm{xc}}[n, \chi]}{\delta \chi^*(\boldsymbol{r}, \boldsymbol{r}')} \qquad [22]$$

Unfortunately, only few approximations to the exchange–correlation functionals have been proposed so far. From these, both an LDA-like functional aimed at describing electron–electron correlations and the development of a scheme that incorporates strong electron–phonon interactions and electron screening are mentioned. In this framework it is now possible to calculate transition temperatures and energy gaps of superconductors from first principles.

Time-Dependent Density-Functional Theory

Several problems in physics require the solution of the time-dependent Schrödinger equation

$$\mathrm{i}\frac{\partial}{\partial t}\Psi(\underline{\boldsymbol{r}}, t) = \hat{H}(t)\Psi(\underline{\boldsymbol{r}}, t) \qquad [23]$$

where Ψ is the many-body wave function of N electrons with coordinates $\underline{\boldsymbol{r}} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, and $\hat{H}(t)$ is the operator (1) generalized to the case of time-dependent external potentials. This equation describes the time evolution of a system subject to the initial condition $\Psi(\underline{\boldsymbol{r}}, t = t_0) = \Psi_0(\underline{\boldsymbol{r}})$. The initial density is n_0.

The central statement of time-dependent density-functional theory (TDDFT), the Runge–Gross theorem, states that, if Ψ_0 is the ground state of $\hat{H}(t = t_0)$, there exists a one-to-one correspondence between the time-dependent density $n(\boldsymbol{r}, t)$ and the external potential $v(\boldsymbol{r}, t)$. All observables of the system can then be written as functionals of the density only. The Runge–Gross theorem provides the theoretical foundation of TDDFT. The practical framework of the theory is given by the Kohn–Sham scheme. A system of noninteracting electrons is defined, obeying the one-particle Schrödinger equation

$$\mathrm{i}\frac{\partial}{\partial t}\varphi_i(\boldsymbol{r}, t) = \left[-\frac{\nabla^2}{2} + v_{\mathrm{s}}(\boldsymbol{r}, t)\right]\varphi_i(\boldsymbol{r}, t) \qquad [24]$$

Similar to static Kohn–Sham theory, the Kohn–Sham potential is usually written as

$$v_{\mathrm{s}}[n](\boldsymbol{r}, t) = v(\boldsymbol{r}, t) + \int \mathrm{d}^3 r' \frac{n(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|} + v_{\mathrm{xc}}[n](\boldsymbol{r}, t) \qquad [25]$$

where $v(\boldsymbol{r}, t)$ includes both the static Coulomb potential of the nuclei and the explicitly time-dependent contribution of an external electromagnetic field. The exchange–correlation potential, v_{xc}, is chosen such that the density of the Kohn–Sham electrons

$$n(\boldsymbol{r}, t) = \sum_i^{\mathrm{occ}} |\varphi_i(\boldsymbol{r}, t)|^2 \qquad [26]$$

is equal to the density of the interacting system. Clearly, v_{xc} is a very complex quantity – it encompasses all nontrivial many-body effects – that has to be approximated. The most common approximation to $v_{\mathrm{xc}}(\boldsymbol{r}, t)$ is the adiabatic local density approximation (ALDA). In the ALDA it is assumed that the exchange–correlation potential at time t is equal to the ground-state LDA potential evaluated with the density $n(\boldsymbol{r}, t)$, that is,

$$v_{\mathrm{xc}}^{\mathrm{ALDA}}[n](\boldsymbol{r}, t) = v_{\mathrm{xc}}^{\mathrm{unif}}(n(\boldsymbol{r}, t)) \qquad [27]$$

Following the same reasoning it is simple to derive an adiabatic GGA or meta-GGA.

Two regimes can usually be distinguished in TDDFT (or more generally in time-dependent quantum mechanics), depending on the strength of the

external time-dependent potential. If this potential is "weak," the solution of the time-dependent Kohn–Sham equations can be circumvented by the use of linear (or low-order) response theory. In linear response, one looks at the linear change of the density produced by a small (infinitesimal) change in the external potential – information which is contained in the so-called linear density response function χ. Within TDDFT, χ can be calculated from the Dyson-like equation

$$\chi(\boldsymbol{r}t,\boldsymbol{r}'t') = \chi_s(\boldsymbol{r}t,\boldsymbol{r}'t') + \int d^3x \int d^3x' \int d\tau \int d\tau'$$
$$\chi_s(\boldsymbol{r}t,\boldsymbol{x}\tau)\left[\frac{\delta(\tau-\tau')}{|\boldsymbol{x}-\boldsymbol{x}'|} + f_{xc}(\boldsymbol{x}\tau,\boldsymbol{x}'\tau')\right]\chi(\boldsymbol{x}'\tau',\boldsymbol{r}'t') \qquad [28]$$

where χ_s is the (noninteracting) Kohn–Sham response function. The exchange–correlation kernel f_{xc} is defined as

$$f_{xc}(\boldsymbol{r}t,\boldsymbol{r}'t') = \left.\frac{\delta v_{xc}(\boldsymbol{r},t)}{\delta n(\boldsymbol{r}',t')}\right|_{n_0} \qquad [29]$$

where the functional derivative is evaluated at the initial ground-state density n_0. If one puts $f_{xc} = 0$ in [28], the well-known random phase approximation (RPA) to the excitation energies is obtained. Knowledge of χ allows for the calculation of linear absorption spectra. Here one has to distinguish between finite and extended systems. The optical absorption spectrum of a finite system is essentially given by the poles of χ. For an extended system, the spectrum is obtained from the imaginary part of the dielectric function, which can be written (in frequency space) as

$$\varepsilon^{-1}(\boldsymbol{r},\boldsymbol{r}',\omega) = \delta(\boldsymbol{r}-\boldsymbol{r}') + \int d^3x \frac{\chi(\boldsymbol{x},\boldsymbol{r}',\omega)}{|\boldsymbol{r}-\boldsymbol{x}|} \qquad [30]$$

Most calculations performed within this scheme approximate the exchange–correlation kernel by the ALDA. For finite systems, excitation energies are of excellent quality for a variety of atoms, organic and inorganic molecules, clusters, etc. Unfortunately, the situation is different for extended systems, especially for large-gap semiconductors. In **Figure 1** the ALDA optical absorption spectrum of silicon together with an RPA calculation, a Bethe–Salpeter calculation (a method based on many-body perturbation theory), and experimental results are shown. It is clear that ALDA fails to give a significant correction over RPA. In momentum space, the Coulomb interaction is simply $1/q^2$. It is then clear from [28] that if f_{xc} is to contribute for small q – which is the relevant regime for optical absorption – it will have to behave asymptotically as $1/q^2$. The ALDA kernel approaches a constant for $q \to 0$ and is therefore incapable to correct the RPA results. Recently several new kernels have been proposed that overcome this problem. One of them, the RORO kernel, has the simple form

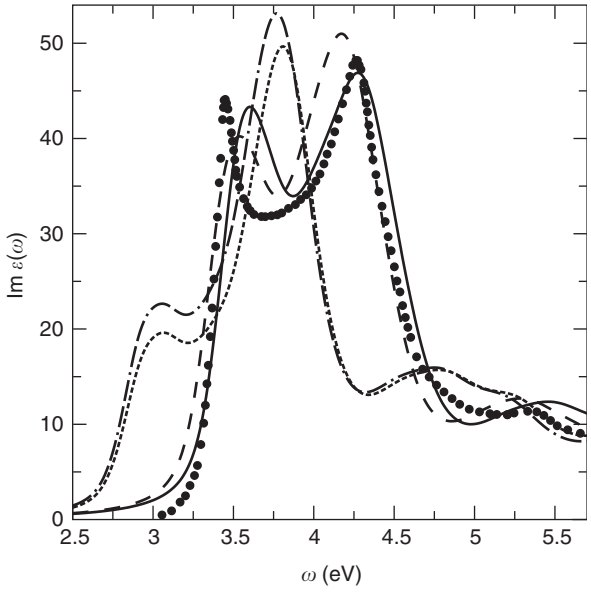


Figure 1 Optical absorption spectrum of silicon. The curves are: thick dots – experiment dotted curve – RPA; dot-dashed curve – TDDFT using the ALDA; solid curve – TDDFT using the RORO kernel; dashed curve – Bethe-Salpeter equation. (Reproduced with permission from Onida G, Reinig L, and Rubio A (2002) *Review of Modern Physics* 74: 601; Lautenschlager P, Garriga M, Viña L, and Cardona M (1987) *Physical Review B* 36: 4821; and Reining L, Olevano V, Rubio A, and Onida G (2002) *Physical Review Letters* 88: 066404.)

$$f_{xc}^{RORO}(\boldsymbol{r},\boldsymbol{r}') = -\frac{\alpha}{4\pi}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} \qquad [31]$$

where α is a phenomenologically determined parameter. Taking $\alpha = 0.12$ yields the curve shown in **Figure 1**, which is already very close to the experimental results.

Conclusions

There is no doubt that DFT has become a mature theory for electronic structure calculations. The success of DFT is based on the availability of increasingly accurate approximations to the exchange–correlation energy. Research into the development of still more accurate approximations continues. In static DFT, orbital functionals seem to be promising candidates to achieve this goal. In time-dependent DFT as well as in other extensions of DFT, development of approximations is still at a much earlier stage.

Applications of DFT abound. In solid-state physics, DFT is the method of choice for the calculation of material properties from first principles. While

traditional, static DFT calculations have mainly focused on structural ground-state properties such as lattice parameters, TDDFT also allows for the calculation of excited state properties such as absorption spectra. In quantum chemistry, the decade-long dream of electronic structure calculations with "chemical accuracy" appears within reach with DFT. Due to the favorable scaling of the computational effort with increasing system size and also to the ever-increasing computational power of modern computers, DFT calculations are now feasible for systems containing several thousand atoms. This opens the way for applications of DFT to biologically relevant systems.

See also: Ballistic Transport; Carbon Materials, Electronic States of; Crystal Field Splitting; Deterministic Aperiodic Solids, Electronic Properties of; Disorder and Localization Theory; Disordered Solids and Glasses, Electronic Structure of; Effective Masses; Electron Gas (Theory); Electronic Structure (Theory): Atoms; Electronic Structure (Theory): Molecules; Electronic Structure Calculations: Plane-Wave Methods; Electronic Structure Calculations: Scattering Methods; Electrons in Periodic Potentials with Magnetic Fields; Elemental Semiconductors, Electronic States of; Fermi Surface Measurements; Geometrical Phase and Polarization in Solids; Hartree and Hartree–Fock Methods in Electronic Structure; Insulators, Electronic States of; Insulators, Impurity and Defect States in; Intermetallic Compounds, Electronic States of; Jahn–Teller Effect; Liquids, Electronic Structure of; Liquids, Theory of: Fermi Liquids; Magnetoresistance Techniques Applied to Fermi Surfaces; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Metals and Alloys, Impurity and Defect States in; Molecular Dynamics Calculations; Molecular Dynamics, Car–Parrinello; Nanostructures, Electronic Structure of; Photonic Bandgap Materials, Electronic States of; Polymers and Organic Materials, Electronic States of; Pseudopotential Method; Quasicrystals, Electronic Structure of; Recursive Methods for Electronic States; Semiconductor and Metallic Clusters, Electronic Properties of; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Heterojunctions, Electronic Properties of; Semiconductors, Impurity and Defect States in; Semi-Empirical and Empirical Calculations; Surfaces and Interfaces, Electronic Structure of; Tight-Binding Method in Electronic Structure.

PACS: 31.15.Ew; 31.10.+z; 33.15.−e; 33.20.Kf; 33.20.Lg; 71.15.−m; 74.20.−z

Further Reading

Dreizler RM and Gross EKU (1990) *Density Functional Theory*. Berlin: Springer.

Fiolhais C, Nogueira F, and Marques MAL (eds.) (2003) *A Primer in Density Functional Theory*, Lecture Notes in Physics. Berlin: Springer.

Kohn W (1999) Nobel lecture: electronic structure of matter – wave functions and density functionals. *Reviews of Modern Physics* 71: 1253.

Parr RG and Yang W (1989) *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press.

Sen KD (ed.) (2002) A Celebration of the contributions of R. G. Parr. *Reviews in Modern Quantum Chemistry*. Singapore: World Scientific.

Deterministic Aperiodic Solids, Electronic Properties of

J Bellissard, Georgia Institute of Technology, Atlanta, GA, USA

Aperiodic Media

Since the early 1970s, several classes of aperiodic solids have been the focus of attention. Examples of such materials are the following:

1. Electrons in a uniform magnetic field are sensitive to the vector potential which cannot be constant in space at nonvanishing magnetic field, thus breaking the translation symmetry.
2. Randomly distributed impurities or defects break the translation invariance in crystal-like normal metals or doped semiconductors. However, electrons feel them only at low temperature. Similarly, only phonons of short wavelength are sensitive to disorder.
3. Amorphous materials, such as glass, silicon (in its amorphous phase), even liquids at short timescales, have been considered as such in the literature.
4. Artificial structures, such as superlattices or arrays of quantum dots, are technologically available and can be designed to break the translation symmetry.
5. Quasicrystals, such as $Al_{62.5}Cu_{25}Fe_{12.5}$ or $Al_{70.5}Pd_{22}Mn_{7.5}$ have a point-like diffraction spectrum with symmetries (such as the fivefold or tenfold symmetries) incompatible with spatial periodicity.

In view of their growing importance, there is a need for mathematical and numerical methods to study them, since the usual Bloch theory cannot be used,

due to the lack of translation symmetry. To avoid complexity, as a first approximation, the degrees of freedom (electrons, holes, phonons, lacunæ) are considered as independent and moving in an infinite volume limit ideal solid.

Transport Mechanisms

Various mechanisms of transport can be observed experimentally in aperiodic materials:

1. For metals, the conductivity increases as the temperature decreases. Moreover, the Fermi liquid theory (Landau) predicts
2. Thermally activated processes, (if a gap arises at the Fermi level), give

$$\sigma(T) \overset{T\downarrow 0}{\sim} T^{-2}$$

3. For weakly disordered systems (strongly doped semiconductors, normal metals), there is a residual conductivity at low temperature, due to quantum interferences

$$\sigma(T) \overset{T\downarrow 0}{\sim} \mathrm{e}^{-\Delta/T}$$

4. In strongly disordered systems (lightly doped semiconductors) with small density of states (DOS) at the Fermi level, Mott variable range hopping predicts in dimension d

$$\sigma(T) \overset{T\downarrow 0}{\sim} \sigma_0 > 0$$

5. High-quality quasicrystals are metallic alloys but their conductivity decreases with the temperature (see **Figure 1**) and exhibits a scaling behavior

$$\sigma(T) \overset{T\downarrow 0}{\sim} \mathrm{e}^{-(T_0/T)^{1/(d+1)}}$$

$$\sigma(T) \overset{T\downarrow 0}{\sim} \sigma_0 + aT^{\gamma}, \quad 1 < \gamma < 1.5$$

on a large range of temperatures (from 50 mK to 1000 K in some cases). In most cases $\sigma_0 > 0$ (residual conductivity), but there are examples for which $\sigma_0 \approx 0$ (**Figure 1**).

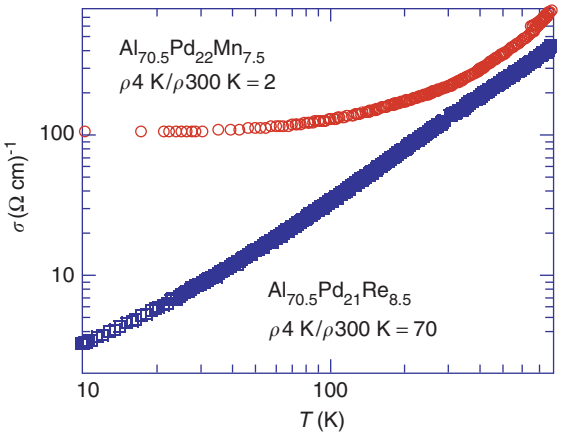


Figure 1 Conductivity of two examples of quasicrystals.

The Hull

The "Hull" of an aperiodic solid and its canonical transversal, or atomic surface, is the main construction (as the Brillouin zone for a crystal). It represents in a compactified way all possible samples of the material that can be ideally produced.

As a good approximation, atomic nuclei are supposed to be frozen at their ideal equilibrium position at zero temperature. The set $\mathscr{L}$ of atomic positions is thus a discrete subset of the space $\mathbb{R}^d$ (d is the dimension of the material). $\mathscr{L}$ is called uniformly discrete whenever the minimum distance between any two distinct points of $\mathscr{L}$ is positive. $\mathscr{L}$ is called relatively dense whenever there is a radius R large enough so that any ball of radius R meets $\mathscr{L}$ on at least one point. A set is Delone if it is both uniformly discrete and relatively dense. A Delone set is repetitive if, given any finite patch p (namely, a finite set of the form $p = \mathscr{L} \cap B$, where B is some open ball), and any small length $\varepsilon > 0$, there is a radius $R_p > 0$ such that any ball of radius R_p contains a translated copy of p modulo, an error smaller than ε. A Delone set has finite type if the set of vectors joining two of its points is itself discrete (this latter set is denoted by $\mathscr{L} - \mathscr{L}$). Lastly, a Meyer set is a Delone set such that the family $\mathscr{L} - \mathscr{L}$ of vectors joining two of its points is itself a Delone set.

Examples:

1. Throwing points randomly in space with a fixed density (Poisson's law) gives a discrete set which, with probability 1, is neither uniformly discrete nor relatively dense.
2. Randomly throwing impurities on sites of a silicon perfect diamond lattice gives a uniformly discrete set which, with probability 1, is not relatively dense. Other materials like zeolithes may have their atoms distributed on a fractal uniformly discrete but nonrelatively dense set.
3. In the ideal limit of zero temperature, most solids have their atoms distributed on a Delone set. But amorphous ones, such as glass or silicon, are neither of finite type nor are they repetitive because their two-point correlation function falls off quickly at a large distance. Perfect crystals and quasicrystals are described by a finite type repetitive Delone set.

4. Typical examples of Meyer sets are provided by crystals or quasicrystals (see **Figure 2**). These are the sets with a point-like diffraction pattern and no diffuse part.

If the solid is homogeneous, $\mathscr{L}$ and $\mathscr{L}+a$ have similar global properties. However, an observer sees only a finite sample in practice, namely, he sees only what happens in a bounded window $\Lambda\subset\mathbb{R}^d$. A sequence of translated $\mathscr{L}_n=\mathscr{L}+a_n$ converges to the discrete set $\mathscr{L}_\infty$ if and only if, for any bounded window Λ, the family $\mathscr{L}_n\cap\Lambda$ of points of $\mathscr{L}_n$ contained in Λ converges to $\mathscr{L}_\infty\cap\Lambda$. The Hull Ω of $\mathscr{L}$ is defined as the family of limit points of translated $\mathscr{L}$. By construction, any point $\omega\in\Omega$ corresponds to a set $\mathscr{L}_\omega$ obtained as the limit of a subsequence of translated $\mathscr{L}$. It can be proved that, as soon as $\mathscr{L}$ is uniformly discrete, Ω is a compact and metrizable set. Metrizability means that the topology defined above can be described through a metric. An example of such metric (it is not the only one) is the following: the distance $d(\omega,\omega')$ between the points ω,ω' of the Hull is the smallest number $\varepsilon>0$, such that $\mathscr{L}_\omega$ and $\mathscr{L}_{\omega'}$ coincide in the ball centered at the origin of radius $1/\varepsilon$ modulo, an error of ε. Compactness means that given any $\varepsilon>0$, there is a finite set $\{a_1,\ldots,a_d\}$ of vectors $\mathbb{R}^d$, such that any ω is within the distance ε of at least one of the $\mathscr{L}+a_i$'s. Hence in a ball of radius $1/\varepsilon$, $\mathscr{L}_\omega$ looks like $\mathscr{L}+a_i$ modulo, an error smaller than ε. In other words, compactness is a way of expressing that $\mathscr{L}$ describes a homogeneous solid. In particular, the minimum distance between distinct points in $\mathscr{L}_\omega$ is at least equal to the one in $\mathscr{L}$. Moreover the properties of being Delone, finite type, repetitive, or Meyer are inherited by all the $\mathscr{L}_\omega$s.

In addition, if $a\in\mathbb{R}^d$ is a vector in space, then $\mathscr{L}_\omega+a$ also belongs to the Hull, thus defining a point denoted by $T^a\omega$. It can be shown that the map $(a,\omega)\mapsto T^a\omega$ is continuous and satisfies $T^{a+a'}\omega=T^a(T^{a'}\omega)$ while $T^0\omega=\omega$, hence giving a topological action of the translation group on the Hull. Therefore, $(\Omega,\mathbb{R}^d,T)$ is a topological dynamical system.

To build the Hull it is convenient to use the canonical transversal Ξ, defined as the set of points ω in the Hull such that the set $\mathscr{L}_\omega$ contains the origin $0\in\mathbb{R}^d$ (see **Figure 3**). Starting from a point in Ξ and translating it, it takes a minimum distance between two points in $\mathscr{L}$ at least to have $0\in\mathscr{L}_\omega+a$ again. So, $\mathscr{L}_\omega$ can be seen as the trace of Ξ in the orbit of ω. Since such an orbit can be seen as the space $\mathbb{R}^d$ in which the solid is living and since the points of $\mathscr{L}_\omega$ are the positions of atoms, Ξ represents the position of the atom in the Hull. This is why Ξ is also called (especially in quasicrystallography) the atomic surface.

For periodic materials, the Hull is the d-torus obtained by identifying points in spaces that differ by a vector of the period group (the Bravais lattice). The transversal is the (finite) set of atoms of the unit cell. For quasicrystals, the Hull is a torus of higher dimension in which the physical space $\mathbb{R}^d$ has

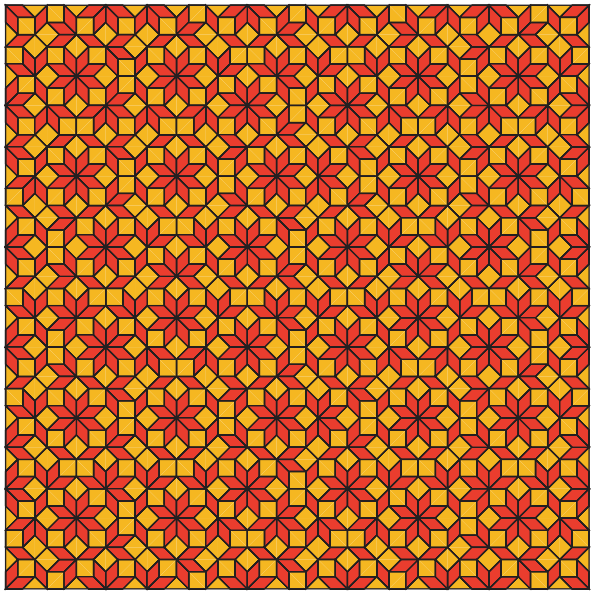


Figure 2 A Meyer set: the Ammann–Beenker tiling.

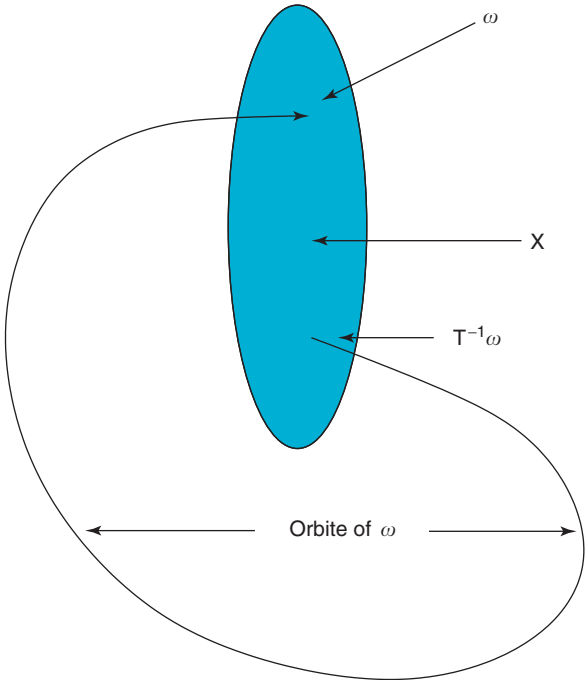


Figure 3 Schematic view of the atomic surface.

irrational slopes. Moreover, the Hull has the topology of a Cantor set, in the direction perpendicular to the physical space (see **Figure 4**).

Noncommutative Brillouin Zone

The description of dynamical degrees of freedom (electrons, holes, phonons) cannot be done through the Bloch theory since the translation invariance is broken. However, any observable describing them can be given by a covariant operator, namely, a family $A = (A_\omega)_{\omega\in\Omega}$ of self-adjoint bounded operators acting on the Hilbert space $\mathscr{H}$ of quantum states, such that: (1) $T(a)A_\omega T(a)^{-1} = A_{T^a\omega}$ if $T(a)$ is the translation operator by $a\in\mathbb{R}^d$ acting on $\mathscr{H}$, and (2) for any quantum state $|\psi\rangle$ the map $\omega\mapsto A_\omega|\psi\rangle$ is continuous (strong continuity). If the material is periodic, all such observables are matrix-valued functions over the Brillouin zone. Hence, the algebra of observables is the noncommutative analog of the space of continuous functions on the Brillouin zone, leading to the notion of Noncommutative Brillouin Zone (NCBZ). Such functions can be integrated: if $\mathbb{P}$ is an $\mathbb{R}^d$ invariant probability measure on the Hull, then $\mathscr{T}_\mathbb{P}(A)$ is defined as the average of the diagonal matrix elements of A_ω and is equal to the trace per unit volume of A. The map $A\mapsto\mathscr{T}_\mathbb{P}(A)$ is linear, positive (namely, $\mathscr{T}_\mathbb{P}(A^*A)\geqslant 0$), and tracial namely $\mathscr{T}_\mathbb{P}(AB) = \mathscr{T}_\mathbb{P}(BA)$. It is the noncommutative analog of the integral over the NCBZ. In addition, if $\boldsymbol{R} = (R_1, \ldots, R_d)$ is the position operator, then $\boldsymbol{\nabla}A = \imath[\boldsymbol{R}, A]$ gives a family of d commuting derivations (namely, $\boldsymbol{\nabla} = (\partial_1, \ldots, \partial_d)$ is linear and obeys the Leibniz rule $\boldsymbol{\nabla}(AB) = \boldsymbol{\nabla}(A)B + A\boldsymbol{\nabla}(B)$) generalizing the derivatives with reference to the quasimomentum in the NCBZ. Depending upon the degrees of freedom (electrons, holes with spin or spinless, or phonons), this algebra of observables is represented in various Hilbert spaces. Most formulas found in textbooks, valid for periodic crystals, extend in a similar way to aperiodic materials provided the integral over quasimomenta is replaced by $\mathscr{T}_\mathbb{P}$, and the derivatives $\boldsymbol{\nabla}_k$ by $\boldsymbol{\nabla}$.

In the independent particle approximation, the electron or holes are described by a Hamiltonian $H = (H_\omega)_{\omega\in\Omega}$ coming from a Schrödinger operator. The simplest example, for spinless electrons, acts on $\mathscr{H} = L^2(\mathbb{R}^d)$ and is given by

$$H_\omega = -\frac{\hbar^2}{2m}\Delta + \sum_{a\in\mathfrak{A}}\sum_{r\in\mathscr{L}_\omega^a} v_a(\boldsymbol{R}-\boldsymbol{r})$$

where $\mathfrak{A}$ is a finite set labeling various species of atoms, $\mathscr{L}_\omega^a$ represents the set of positions for the atoms of species $a\in\mathfrak{A}$, and v_a is the atomic pseudopotential created by atoms of species a acting on the charge carrier. In the tight-binding representation, the Hamiltonian becomes a matrix indexed by the atomic sites and the corresponding atomic species, acting on $\mathscr{H}_\omega = \ell^2(\mathscr{L}_\omega)$, of the form

$$\langle x, a|H_\omega|y, b\rangle = h_{ab}(T^{-x}\omega, y-x),$$
$$a, b\in\mathfrak{A},\ x\in\mathscr{L}_\omega^a,\ y\in\mathscr{L}_\omega^b$$

where h is some continuous matrix-valued function on the set Γ of pairs (ω, x) such that ω belongs to the transversal, and $x\in\mathscr{L}_\omega$.

In the harmonic approximation, phonons are described by the creation and annihilation operators of bosons attached to each atomic site with a polarization

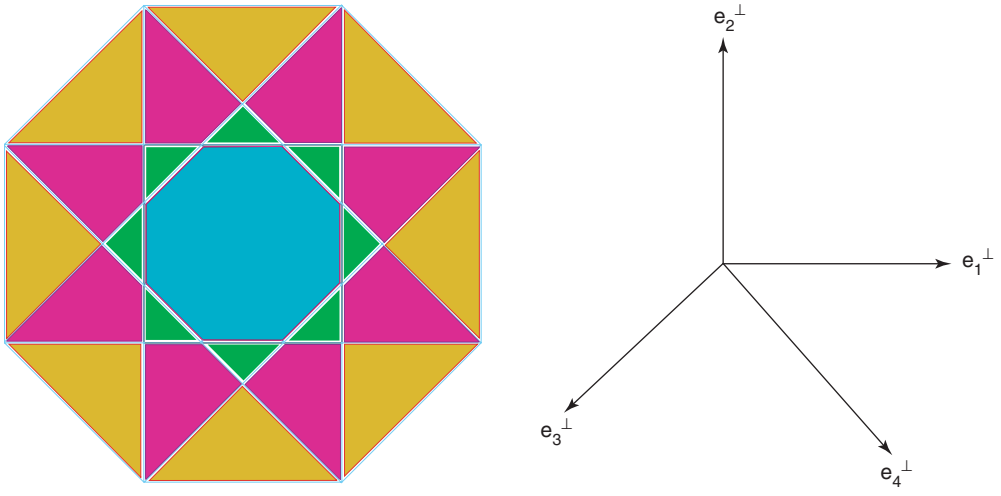


Figure 4 The atomic surface (canonical transversal) of the Ammann–Beenker tiling.

index, $a_{\omega,x;a}$, $a^\dagger_{\omega,x;a}$ such that $\left[a_{\omega,x;a}, a^\dagger_{\omega,y;b}\right] = \delta_{ab}\delta_{xy}\mathbf{1}$ whenever $(\omega, x) \in \Gamma$ and a,b are the polarization index (usually varying from 1 to d, if d is the dimension). Then the phonon Hamiltonian, restricted to a finite volume Λ, acts on the boson Fock space of $\ell^2(\mathscr{L}_\omega)$ and is given by

$$H^{(\mathrm{ph})}_{\omega;\Lambda} = \sum_{x,y \in \mathscr{L}_\omega \cap \Lambda} \hbar\kappa_{ab}(T^{-x}\omega, y-x) a^\dagger_{\omega,x;a} a_{\omega,y;b} + W_{\omega;\Lambda}$$

where W represents the anharmonic contributions. Here κ can be seen as the matrix elements of a covariant operator $K = (K_\omega)_{\omega\in\Omega}$. The eigenvalues of K are the square of the vibrational mode frequencies allowed to propagate through the solid.

The charge carrier DOS $\mathscr{N}(\Delta)$, in an energy interval Δ, is defined as the number of eigenstates of the charge carrier Hamiltonian describing the charge carrier motions per unit volume with energy within Δ. Similarly, the vibrational density of states (VDOS) in some frequency interval Δ is the number of phonon modes with frequency within Δ. The Shubin formula shows that they are given by

$$\mathscr{N}_{\mathbb{P}}(\Delta) = \mathscr{T}_{\mathbb{P}}(\chi_\Delta(H)) \text{ (charge carriers)},$$
$$\mathscr{N}_{\mathbb{P}}(\Delta) = \mathscr{T}_{\mathbb{P}}(\chi_\Delta(K^{1/2})) \text{ (phonons)},$$

where $\chi_\Delta(E) = 1$ if $E \in \Delta$ and 0 otherwise. A useful example of a formula expressed with this formalism is the Kubo formula for the electric conductivity in the relaxation time approximation (RTA) given by

$$\text{Kubo's formula}$$
$$\sigma_{ij} = \frac{e^2}{\hbar}\mathscr{T}_{\mathbb{P}}\left(\partial_f\left(\frac{1}{1+\mathrm{e}^{\beta(H-\mu)}}\right) \times \frac{1}{\hbar/\tau_{\mathrm{rel}} + \mathfrak{Q}_{\mathrm{H}} - \imath\hbar\tilde{\omega}}(\partial_i H)\right) \qquad [1]$$

where H is the one-particle Hamiltonian, $\mathfrak{Q}_{\mathrm{H}} = \imath[H, \cdot]$, $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature (k_{B} is Boltzmann's constant), $\tilde{\omega}$ is the current frequency, τ_{rel} is the inelastic collision time, μ is the chemical potential fixed by the charge carrier density n_{el} as

$$n_{\mathrm{el}} = \mathscr{T}_{\mathbb{P}}\left(\frac{1}{1+\mathrm{e}^{\beta(H-\mu)}}\right) = \int_{-\infty}^{+\infty} \mathrm{d}\mathscr{N}_{\mathbb{P}}(E)\frac{1}{1+\mathrm{e}^{\beta(E-\mu)}}$$

Exponents

The main difference between periodic and aperiodic solids can be seen in the electron energy spectrum or in the phonon vibrational spectrum (at least for short wavelength). The band spectrum can be a Cantor set, at least in 1D or near the groundstate if the dimension is higher. The electrons can be localized (Anderson localization) or can have an exotic long-distance behavior. The phonon spectrum is likely to have a large low energy optical band exhibiting quantum chaos. Various scaling exponents classify such a behavior. The spectral exponents are associated with the DOS, VDOS or with the local density of states (LDOS). The diffusion exponents are associated with the motion of a typical wave packet in space.

The DOS (or VDOS) regularity is characterized by its behavior as the size $|\Delta|$ of the energy or frequency interval goes to zero, namely $\mathscr{N}(\Delta) \sim |\Delta|^d$. Such a definition is ambiguous and, even in computer simulations, difficulties in interpreting results can be overwhelming. The symbol $\sim$ should then be defined as follows: if $f(\varepsilon)$ is a positive decreasing function of $0 < \varepsilon < 1$ (or $1 < \varepsilon < \infty$) then

$$\alpha^\bigstar = \lim_{\varepsilon\downarrow 0} \bigstar \frac{\ln f(\varepsilon)}{\ln \varepsilon}, \quad \text{denoted by } f \overset{\varepsilon\downarrow 0}{\sim} \varepsilon^{\alpha^\bigstar}$$

where $\bigstar$ stands for sup (or $+$) or inf (or $-$), and similarly if $\varepsilon \uparrow \infty$. Therefore, the DOS-exponent (and similarly for the VDOS) is defined by

$$\mathscr{N}_{\mathbb{P}}([E-\varepsilon, E+\varepsilon]) \overset{\varepsilon\downarrow 0}{\sim} \varepsilon^{d^\bigstar_{\mathrm{DOS}}(E)}$$

The LDOS, also called the spectral measure, is defined as follows: if $H = (H_\omega)_{\omega\in\Omega}$ is the one-particle Hamiltonian for the charge carriers, the LDOS $\mu_{\omega,\phi}$ relative to the quantum state $|\phi\rangle$ and its spectral exponent $d^\pm_{\omega,\phi}(E)$ are defined by

$$\mu_{\omega,\phi}([E-\delta, E+\delta]) := \int_{E-\delta}^{E+\delta} \mathrm{d}\mu_{\omega,\phi} = \frac{1}{\pi}\Im m \int_{E-\delta}^{E+\delta} \mathrm{d}E' \times \langle\phi|(E' - \imath 0^+ - H_\omega)^{-1}|\phi\rangle \overset{\delta\downarrow 0}{\sim} \delta^{d^\bigstar_{\omega,\phi}(E)}$$

Whenever Δ is an interval of energy and if $\bigstar = \sup$ (resp. inf), $d^\bigstar_{\omega,\phi}(\Delta)$ will be infimum (resp. supremum) over $\Delta' \subset \Delta$ of the supremum (resp. infimum) over $E \notin \Delta'$ of $d^\bigstar_{\omega,\phi}(E)$'s where $\int_{\Delta'} \mathrm{d}\mu_{\omega,\phi} = 0$. Then $d^{\bigstar,\pm}_{\omega}(\Delta)$ will be the supremum (resp. infimum) over the state $|\phi\rangle$ of the $d^\bigstar_{\omega,\phi}(\Delta)$'s. It turns out that $d^{\bigstar,\pm}_{\omega}(\Delta)$ is independent of ω with probability 1 and the common value is called $d^{\bigstar,\pm}_{\mathrm{LDOS}}(\Delta)$. In all cases

$$d^{-,\pm}_{\mathrm{LDOS}}(\Delta) \leqslant d^{-}_{\mathrm{DOS}}(\Delta)$$

The LDOS is absolutely continuous (ac) on an interval Δ, whenever for almost all $E \in \Delta$ there is $0 < C(E) < \infty$ such that $\mu_{\omega,\phi}([E-\delta, E+\delta]) \leqslant C(E)\delta$.

Then, the Radon–Nikodym theorem implies the existence of an integrable function F such that $\mu_{\omega,\phi}(\mathrm{d}E) = F(E)\mathrm{d}E$. (Here, $\mathrm{d}E$ is called the Lebesgue measure.) In such a case, $d^{\star}_{\omega,\phi}(E) = 1$ on Δ so that $d^{\star}_{\omega,\phi}(\Delta) = 1$. The LDOS is pure point (pp) on Δ if it can be written as a sum of Dirac measures $\mu_{\omega,\phi}(\mathrm{d}E) = \sum_i \delta(E - E_i)\mathrm{d}E$. In such a case, the E_i are eigenvalues and $d^{\star}_{\omega,\phi}(E_i) = 0$ on Δ, implying $d^{\star}_{\omega,\phi}(\Delta) = 0$. The LDOS is singular continuous (sc) if $\mu_{\omega,\phi}([E-\delta, E+\delta])\downarrow 0$ as $\delta\downarrow 0$ for all $E\in\Delta$, and if there is almost no $E\in\Delta$ for which $\mu_{\omega,\phi}([E-\delta, E+\delta]) \leqslant C(E)\delta$ with $0 < C(E) < \infty$. In particular, if $0 < d^{\star}_{\omega,\phi}(\Delta) < 1$, then the LDOS is sc on Δ.

A charge carrier moving in the solid is a quantum wave interfering with itself after colliding with atomic nuclei. The resulting effect of too many reflections is a slowing down of its motion so that the distance it goes after time t behaves like

$$r(t) \overset{t\uparrow+\infty}{\sim} t^{\overline{\beta}}$$

where $\overline{\beta}$ is called a diffusion exponent. As in the case of spectral exponents, there are many ways of defining $\overline{\beta}$. If $\boldsymbol{R}$ is the position operator for the charge carriers, then $\boldsymbol{R}(t) = \mathrm{e}^{\imath tH/\hbar}\boldsymbol{R}\mathrm{e}^{-\imath tH/\hbar}$ denotes the position after time t, and $|\boldsymbol{R}(t) - \boldsymbol{R}|$ is the distance that the particle goes during this time. To measure it, it is convenient to select particles having energy in some interval Δ and prepared in the quantum state located at the origin of space, denoted by $|0\rangle$. Since the system is not translation invariant, it is better to average over the choice of the origin (using $\mathbb{P}$). To avoid meaningless time oscillations, it is also convenient to average over time too, leading to the following formula

$$\begin{aligned}(L_{\Delta,q}(t))^q &= \int_{-t}^{+t} \frac{\mathrm{d}s}{2t} \int_X \mathrm{d}\mathbb{P}_{tr}(\omega) \\ &\quad \times \langle 0|\Pi_{\omega,\Delta}|\boldsymbol{R}_\omega(s) - \boldsymbol{R}|^q \Pi_{\omega,\Delta}|0\rangle \sim^{t\uparrow\infty} t^{q\overline{\beta}^{\star}_q(\Delta)}\end{aligned}$$

where $\Pi_{\omega,\Delta}$ projects on eigenvalues of H_ω located in Δ. The extra parameter q gives information about the distribution of distances in the initial state.

The diffusion exponents permit one to classify the motion of the charge carriers (or of the phonon modes) according to their values. In a periodic crystal, it is known that $\overline{\beta}^{\star}_q(\Delta) = 1$ for all q's because the motion is ballistic. On the other hand, in an Anderson insulator, where all states are localized, $\overline{\beta}^{\star}_q(\Delta) = 0$. In general, $0 \leqslant \overline{\beta}^{\star}_q(\Delta) \leqslant 1$ and $\overline{\beta}^{\star}_q(\Delta)$ increases with q. The value $q = 2$ is special and is amenable to experimental measurements. There are situations for which $\overline{\beta}_2 = 1/2$, in particular, in weakly disordered systems or if quantum chaos occurs: this is quantum diffusion. It is important to notice that such a diffusion is not due to some loss of information, but the net result of quantum interferences in a complex environment. The motion is called sub-ballistic if $1/2 < \overline{\beta}_2 < 1$ and subdiffusive if $0 < \overline{\beta}_2 < 1/2$.

These exponents are related by the Guarneri inequality

Guarneri's bound

$$\beta^{\star}_q(\Delta) d \geqslant d^{\star}_{\mathrm{LDOS}}(\Delta) \qquad [2]$$

Thus, in dimension $d = 1$, the ac spectrum implies (quasi) ballistic motion. In higher dimension, for any $\varepsilon > 0$, it is possible to find models of Hamiltonians with an ac-spectrum and diffusion exponent smaller than $1/d + \varepsilon$. In particular, the ac-spectrum is compatible with the quantum diffusion for $d \geqslant 2$.

Conductivity

One of the most striking consequences of the previous scaling analysis is the behavior of the conductivity at low temperature. The classical Drude formula asserts that the conductivity of a metal is given by

$$\sigma = \frac{e^2 n_{\mathrm{el}} \tau_{\mathrm{rel}}}{m^*} \overset{\tau_{\mathrm{rel}}\uparrow\infty}{\sim} \tau_{\mathrm{rel}}$$

where m^* is the effective mass of the charge carriers, and τ_{rel} the typical relaxation time for collisions. In view of the RTA Kubo formula [1], the scaling analysis leads to the anomalous Drude formula

$$\sigma \overset{\tau_{\mathrm{rel}}\uparrow\infty}{\sim} \tau_{\mathrm{rel}}^{2\beta_{\mathrm{F}}-1}$$

where $\beta_{\mathrm{F}} = \overline{\beta}^{\star}_2(E_{\mathrm{F}})$ is the diffusion exponent for $q = 2$ at the Fermi level E_{F}. The inelastic relaxation time increases as the temperature goes to zero so that (1) if the motion is ballistic or sub-ballistic, then the solid behaves like a conductor, (2) if the motion is diffusive, then there is a residual conductivity at zero temperature, (3) if the motion is subdiffusive, the solid behaves like an insulator. In the former case (metallic), the addition laws for conductivity show that if several relaxation mechanisms coexist, the one with the shortest relaxation time will dominate (Mathiessen rules). On the contrary, in the latter case (insulating) the one with the longer relaxation time will dominate. In most situations, the dominant dissipative mechanisms come from the phonon–electron collisions in two possible forms: (1) direct collisions leading to the Bloch law $\tau_{\mathrm{rel}} \sim T^{-5}$ as $T\downarrow 0$ (2) electron–electron interactions through the lattice (BCS term) leading to the Landau

Fermi liquid theory with $\tau_{\mathrm{rel}} \sim T^{-2}$. While in the metallic case the Fermi liquid theory applies, the Bloch law dominates in the insulating case leading to $\sigma \sim T^{5(1-2\beta_{\mathrm{F}})}$. If, as computer simulations suggest for $3D$-quasicrystals, $0.35 < \beta\ \mathrm{F} < 0.40$, it provides a mechanism for the results shown in **Figure 1** (AlPdRe-compound).

Another striking numerical observation is that at a small energy scale and for a finite volume, the level spacing distribution of eigenvalues for quasicrystalline lattices follows the Wigner–Dyson law: some quantum chaos occurs, namely the effective Hamiltonian behaves like a random matrix. Such a feature is likely to be universal in aperiodic media. If so, analysis provided for weakly disordered systems should apply, subsequently leading to a quantum diffusive regime at small energy or long timescales and, consequently, producing a residual conductivity at low temperature as shown for the AlPdMn-compound in **Figure 1**. But it might be an artifact of the finite volume. However, following an argument by Thouless, used in disordered systems, it follows that: (1) the time it takes for the particle with diffusion exponent $\overline{\beta}$ to reach the boundary of a sample of size L is $t_{\mathrm{Th}} \sim L^{1/\overline{\beta}}$ (called the Thouless time), (2) the time it takes for the particle to realize that the spectrum is actually discrete due to the finite volume is given by the Heisenberg principle $t_{\mathrm{H}} \Delta E \approx \hbar$ where $\Delta E \sim L^{-d}$ is the average level spacing, so that $t_{\mathrm{H}} \sim L^{d}$ (called the Heisenberg time). If $\overline{\beta} > 1/d$, then the particle reaches the boundary too fast and therefore it will dominantly see the effect of the finite volume, namely of the Wigner–Dyson level spacing distribution. If, on the other hand $\overline{\beta} < 1/d$, the particle never has the time to reach the boundary so that it behaves as if only the subdiffusion was allowed. Hence

$$\beta_{\mathrm{F}} < \frac{1}{d} \Rightarrow \sigma \overset{T \downarrow 0}{\sim} T^{\gamma}, \quad \beta_{\mathrm{F}} > \frac{1}{d} \Rightarrow \sigma \overset{T \downarrow 0}{\sim} \sigma_0 > 0$$

The Guarneri inequality [2] predicts an sc-spectrum near the Fermi level in the former case, while the latter is compatible with the ac-spectrum if $d > 2$. It is reasonable to expect that a slight change in the hopping parameters of the effective tight-binding Hamiltonian between atomic sites due to chemical constraints (e.g., by changing Mn into Re) might produce such a drastic change in behavior in **Figure 1**.

See also: Disordered Solids and Glasses, Electronic Structure of; Metals and Metallic Alloys, Optical Properties of; Nanostructures, Electronic Structure of.

PACS: 61.43. – j; 61.44. – n; 62.30. + d; 63.10. + a; 63.50. + x; 71; 71.20. – b; 71.23. – k; 72.10.Bg; 72.20. – i

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.

Connes A (1994) *Noncommutative Geometry*. San Diego: Academic Press.

Cornfeld I, Fomin S, and Sinai YaG (1982) *Ergodic Theory*. New York: Springer.

Gottschalk W and Hedlund GA (1955) *Topological Dynamics*. Providence: AMS.

Hiai F and Petz D (2000) *The Semicircle Law, Free Random Variables and Entropy*. Providence: American Mathematical Society.

Jones H (1960) *The Theory of Brillouin Zones and Electronic States in Crystals*. Amsterdam: North-Holland.

Katok A and Hasselblatt B (1995) *Introduction to the Modern Theory of Dynamical Systems*. Cambridge: Cambridge University Press.

Mahan GD (1990) *Many Particles Physics*, 3rd edn. New York: Plenum.

Mehta M (1990) *Random Matrices*, 3rd edn. New York: Academic Press.

Reed M and Simon B (1972) *Methods of Modern Mathematical Physics. I. Functional Analysis*. New York: Academic Press.

Reed M and Simon B (1975) *Methods of Modern Mathematical Physics II. Fourier Analysis, Self-Adjointness*. New York: Academic press.

Roche S, Mayou D, and Trambly de Laissardière G (1997) Electronic transport properties of quasicrystals. *Journal of Mathematical Physics* 38: 1794–1822.

Shklovskii BI and Efros AL (1984) *Electronic Properties of Doped Semiconductors*. Berlin: Springer.

Simon B (1978) *Methods of Modern Mathematical Physics. IV. Analysis of Operators*. New York: Academic Press.

Varilly JC, Figueroa H, and Gracia-Bondia JM (2000) *Elements of Noncommutative Geometry*. Boston: Birkhauser.

Voiculescu D, Dykema K, and Nica A (1992) *Free Random Variables*. Providence: American Mathematical Society.

Ziman JM (1979) *Models of Disorder. The Theoretical Physics of Homogeneously Disordered Systems*. Cambridge: Cambridge University Press.

Devices for Telecommunications

M Sugawara and N Hara, University of Tokyo, Meguro-ku, Japan

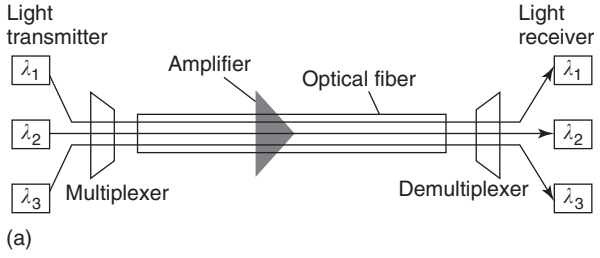
Introduction

The recent global spread of telecommunication networks is due to the advent of optical-fiber and wireless communication technologies. These technologies continue to develop to meet the explosive demand for data traffic on the Internet. In a ubiquitous society, where anyone can access the Internet anywhere at any time, diverse modes of services via various types of devices for connecting to the network are right across the corner.

The optical-fiber communication system basically consists of optical-fiber cables, light transmitters, amplifiers, and light receivers as illustrated in **Figure 1a**. Light transmitters convert electrical signals to optical signals with different wavelengths for different services. Optical signals propagate through the optical fiber to receivers at destinations, being amplified in the case of long fiber span over several tens of kilometers. The receivers reconvert the optical signals to electrical signals for data transmission to be accomplished. The multiplexing technology that uses many wavelengths in the same fiber increases the total transmission bandwidth, and is called wavelength division multiplexing (WDM). In 1996, the total bandwidth of $1.1\,\mathrm{Tb\,s^{-1}}$ was accomplished for a 150-km transmission by using 55 wavelengths with a spacing of 0.6 nm at $20\,\mathrm{Gb\,s^{-1}}$ per wavelength.

Wireless communication systems provide a variety of services such as cellular phones, wireless local area networks (WLAN), electric toll collection system (ETC), satellite communication, fixed wireless access (FWA), and automotive radar as shown in **Figure 1b**. Their progress is due to the development of high-performance transmitters as well as low-noise receivers in the frequency range of 1–100 GHz. A variety of high-speed transistors with output power of 0.01–500 W based on compound and silicon semiconductors play a significant role in this development.

This article explains structures and operation principles of major optical and electrical devices for telecommunications.



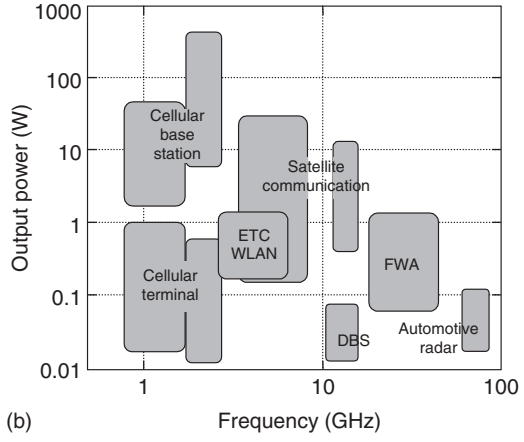


Figure 1 (a) WDM optical-fiber communication system. (b) Requirements for transistors in wireless communication system.

Optical Fiber

Optical fibers are transporting media of high-speed optical signals. Fibers have diverse properties superior to copper cables such as immunity to electromagnetic fields, anti-corrosion, low transmission loss of less than $0.2\,\mathrm{dB\,km^{-1}}$, and extremely broad bandwidth of 40 THz per single fiber. **Figure 2** shows (1) the structure of an optical fiber and (2) the internal light propagation. The fiber consists of the core, the clad surrounding the core, and resinous coating. The fiber cable used in long-haul communication consists of a bundle of optical fibers. A fiber is made by vertically drawing a cylindrical preform made of ultrapure silica in which dopants such as

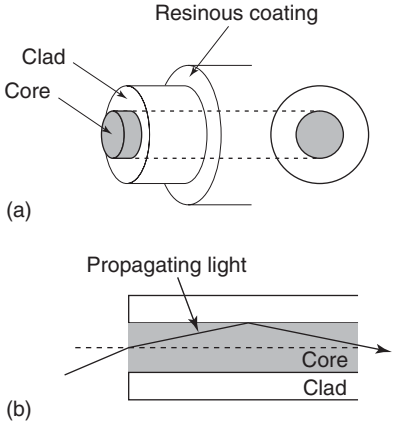


Figure 2 (a) Structure of optical fiber. (b) Internal light propagation.

GeO_2 and B_2O_3 are added in a controlled manner to adjust the refractive index profile of the fiber. The refractive index is increased by GeO_2 and decreased by B_2O_3. The diameter of the clad is 125 μm. The refractive index of the core is slightly larger than that of the clad, so that light entering the core propagates through the optical fiber by repeating the total reflection at the core–clad boundary.

There are two types of fibers with different core diameters for different applications. Fibers with a core diameter of $\sim$50 μm are known as multimode fibers, whereas fibers with a core diameter of $\sim$10 μm are known as single-mode fibers. Modes mean different electromagnetic-field profiles in the cross section of fibers, which can be calculated based on the Maxwell equations under appropriate boundary conditions of the electromagnetic fields. Multimode fibers have an advantage of easy installation owing to their large core diameter, and are used for LANs with fiber spans of less than $\sim$2 km. Their fiber span is limited by the distortion of optical signals due to intermode dispersion, that is, the difference in the group velocity of light pulses between modes. Single-mode fibers are suitable for long-span and high-speed transmission.

Loss, wavelength dispersion, and optical nonlinearity are primary factors that limit the optical-signal transmission performance of fibers. Single-mode fibers have two low-attenuation windows; one $\sim$1.3 μm with a loss of 0.3–0.4 dB km^{-1}, and the other $\sim$1.55 μm with a loss of 0.15–0.2 dB km^{-1}. The wavelength dependence of the dielectric constant of fiber materials causes wavelength dispersion of the light propagation constant, which is called "material dispersion." Another wavelength dispersion is "structural dispersion," which depends on the cross-sectional refractive-index profile of the fiber. The wavelength dispersion results in the broadening of short optical pulses and their interference with former and subsequent pulses to limit the fiber span. Conventional fibers have zero-dispersion point $\sim$1.3 μm. The optical nonlinear effects such as four-wave mixing, cross-phase modulation, and self-phase modulation cause the degradation of the signal-to-noise ratio, wavelength chirping, pulse broadening, and cross talk between different wavelength channels.

Semiconductor Laser

Semiconductor lasers work as standard light transmitters in the optical-fiber communication systems, owing to their small size of 0.2–1 mm length as well as their diverse excellent performances such as the capability of direct modulation up to 10–40 GHz, low-power consumption, single wavelength light, and high output power up to 1 W.

Figure 3a shows the structure of a semiconductor laser. The active region works as a light source via the electron–hole recombination due to spontaneous emission, and as the optical gain medium to amplify light due to stimulated emission. The active region is sandwiched by wide-gap *p*-type and *n*-type cladding layers. This semiconductor layer structure is called a double heterostructure, which effectively confines electrons and holes in the active region due to the well-like energy potential profile as well as light due to the refractive index profile with the active-region index larger than the clad index. The light waveguide is formed in the direction parallel to the layer. The facets on both sides work as optical mirrors to form a cavity for the light propagating in the waveguide. The light partly transmitting through the mirrors is the output. Quantum-well structures consisting of alternating ultrathin semiconductor lasers are widely used as the active region. Semiconductor nanocrystals called quantum dots are being developed for next-generation lasers.

When semiconductor lasers are operated, the forward bias is applied to the *pn* junction to inject electrons and holes into the conduction and valence bands of the active region, respectively, as seen in the band structure of **Figure 3b**. Then, the population inversion is realized where the number of electrons in the conduction band is larger than that of the valence band. Part of the spontaneous emission light emitted by the recombination between electrons and holes goes into the cavity mode, and goes back and forth in the cavity being amplified. When the optical gain balances the total loss, lasing oscillation starts.

Figure 3c shows the output light power versus injected current. Lasing starts at a threshold current to balance the gain and loss of light. The output power rises above the threshold current, where the conversion efficiency from the electrical to optical power reaches up to several tens percent. By directly modulating the injection current as shown in **Figure 3c**, the laser can output digital bit patterns with a variety of information. For long fiber span transmission, distributed feedback (DFB) lasers are used to enable stabilized single-light wavelength, where periodical refractive index distribution is made inside the cavity by an internal grating. Light-signal modulation is also possible by modulating the continuous output of a laser by an external light modulator.

Light Amplifiers

For very long fiber spans of tens of kilometers away, the optical signal may be so attenuated and distorted

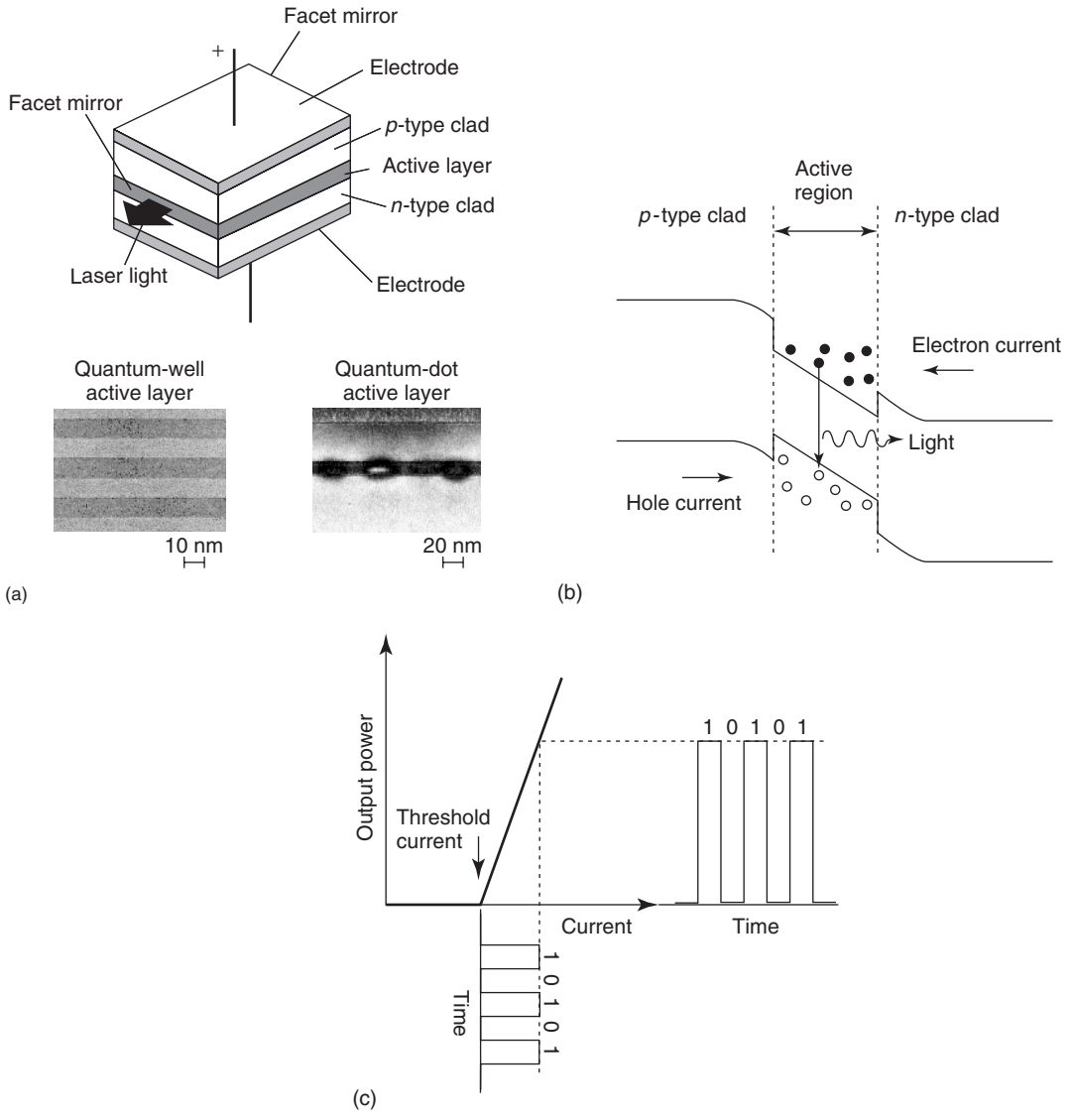


Figure 3 (a) Structure of semiconductor laser. (b) Band diagram under forward bias. (c) Output light power vs. injected current.

that it should be regenerated periodically to reach destinations. The regeneration includes reamplification to recover the signal power, reshaping to reduce noise and distortion, and retiming to reduce the timing jitter, and is called 3R regeneration from their initials. In optical communication networks, there are two types of regenerators: OEO regenerators and optical amplifiers. OEO regenerators receive a modulated optical signal by the photodetector, transform it to an electric signal of the same bit-rate with its timing recovered, amplify it, and then convert the electric signal back to the optical signal of the same modulation bit-rate by the laser. Optical amplifiers recover the optical signal power by the same mechanism as used in the laser; as the light goes through the gain media, its intensity is amplified due to the stimulated emission via the electron–hole recombination. Since only the signal power is recovered, this is a 1R regeneration. Optical amplifiers amplify simultaneously multiple optical signals, each of which is on a different wavelength and within an optical gain bandwidth of the amplifier. This is a significant advantage over the OEO regenerator, which led to the development of the WDM system of **Figure 1a**.

There are two types of optical amplifiers: fiber-type amplifiers (**Figure 4a**) and semiconductor optical amplifiers (SOAs) (**Figure 4b**). The most widely used fiber-type amplifier is the erbium-doped fiber amplifier (EDFA), which mainly consists of the erbium-doped fiber (EDF), pump laser, and coupler as shown in **Figure 4a**. The pump-laser light is coupled to the EDF to excite erbium atoms to generate gain. The optical signals pass through the EDF being amplified. The advantages of the EDFA are high gain, negligible

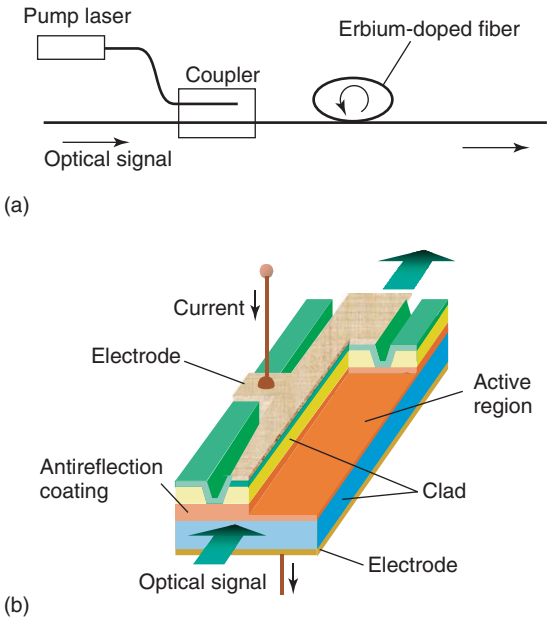


Figure 4 (a) Fiber-type amplifier. (b) Semiconductor optical amplifier.

pattern effect, low noise figure, low cross talk between different channels to enable multichannel amplification. The SOA has a structure similar to the semiconductor laser as shown in **Figure 4b**. By applying forward bias to the *pn* junction, carriers are injected to the active region, causing population inversion to generate gain. The antireflection coating is done on both sides of the facet. The optical signal enters the SOA from one side, propagates the waveguide being amplified, and comes from the other side. The advantages of SOAs are their compactness, low cost, capability of integration with other semiconductor devices, wide gain bandwidth of 50–100 nm, and their easy tunability. Their disadvantages are low saturation power and cross talk between different wavelength channels due to optical nonlinearity. SOAs with quantum dots in their active region are being studied intensively in order to increase the saturation power and realize broader gain bandwidth, so as to enable multi-wavelength amplification with negligible cross talk.

Photodetectors

Photodetectors with sufficiently high-speed response over GHz are positive intrinsic negative (PIN) photodiodes and avalanche photodiodes (APDs).

The PIN photodiode in **Figure 5** consists of an intrinsic region that is sandwiched by *p*-type and *n*-type wide-gap semiconductor layers, and is operated under the reversed bias. When the light signal comes into the device, it is absorbed in the intrinsic

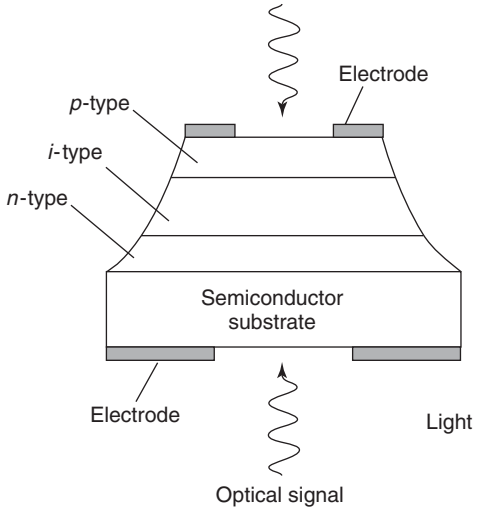


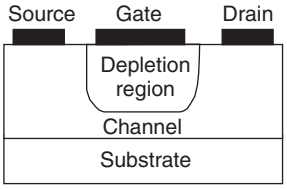
Figure 5 Structure of PIN photodiode.

region to generate electrons in the conduction band and holes in the valence band. The electrons flow toward the *p*-type layer, and the holes toward the *n*-type layer to produce a current pulse with the same duration and shape as the optical signal. The RC time constant determines its frequency response.

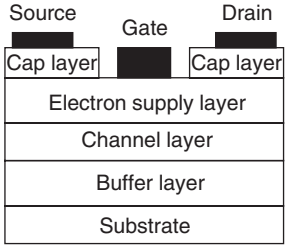
The APD is a detector designed to have a strong electric field of $5–6 \times 10^5\,\mathrm{V\,cm^{-1}}$ in the junction region under the reversed bias to enhance the photocurrent response by the impact ionization. When the light signal creates an electron–hole pair by the optical absorption, the electrons and the holes flow in the opposite direction as in the PIN photodiode. Because of the strong electric fields in the junction, the electron gains enough energy to cause secondary electron–hole pairs, which is called impact ionization, resulting in a signal multiplication process.

High-Speed Transistors

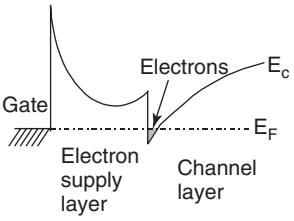
The transmitters and receivers in the optical-fiber communication systems include several kinds of transistors. High-speed/high-frequency operation of transistors is indispensable in increasing communication capacity. Transistors which operate at high frequency are also required in wireless communication systems, since the frequency range used in the systems becomes higher and higher. Compound semiconductors, such as GaAs, InP, and related alloy semiconductors, are suitable for high-speed applications due to the high electron velocity. In addition, utilization of heterostructures makes it possible to design the band lineup, and improves high-speed performances of transistors. Presently, Si-based transistors with heterostructures are also candidates for high-speed devices. They have the advantage of being



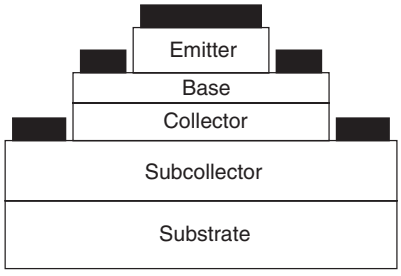
(a) Cross-sectional view of MESFET



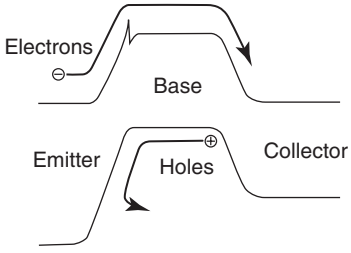
(b) Cross-sectional view of HEMT



(c) Conduction band diagram of HEMT



(d) Cross-sectional view of HBT



(e) Band diagram of HBT

Figure 6 Schematic diagrams of high-speed transistors.

easily integrated with matured Si process technologies.

There are two types of transistors with different principles of operation: field-effect transistors (FETs) and bipolar junction transistors (BJTs). In FETs, the applied gate bias voltage controls the current flow between the drain and the source. **Figure 6a** shows the schematic cross-sectional view of a metal–semiconductor FET (MESFET), one of the representative high-speed transistors. Unlike Si metal-oxide–semiconductor FETs (MOSFETs), a Schottky barrier gate is employed, because compound semiconductors do not have a stable oxide. The gate bias voltage controls the current flow by modulating the depletion region. MESFETs benefit from the high electron velocity of compound semiconductors. It is possible to further increase the electron velocity by using heterostructures. The schematic cross-sectional view and the conduction band diagram of a high-electron-mobility transistor (HEMT) are shown in **Figures 6b** and **6c**. Electrons from donors in a wider bandgap layer, an electron supply layer, transfer to a narrower bandgap layer, a channel layer. Electrons, therefore, are separate from the ionized donors, and ionized impurity scattering is significantly reduced, which results in high electron mobility. The gate voltage controls the current flow by changing the electron density.

BJTs consist of two *pn*-junctions combined with a thin semiconductor layer, the base layer. Electrons are injected from the emitter into the base; then they diffuse across the base layer, and are swept by the collector. The amount of the injected electrons, which can be changed by the base voltage, controls the current flow. **Figures 6d** and **6e** show the schematic cross-sectional view and the conduction band diagram of a heterojunction bipolar transistor (HBT), which is also a popular high-speed transistor. Employment of the wider gap emitter is a key to improving high-speed characteristics. This enables one to increase the doping density of the base to reduce the base resistance without seriously increasing the amount of holes injected from the base to the emitter.

See also: Laser Radiation Sources; Optical Fibers; Semiconductor Devices; Transistors.

PACS: 84.40.Ua

Further Reading

Agrawal GP (1995) *Nonlinear Fiber Optics*, 2nd edn. San Diego: Academic Press.

Akiyama T, Kawaguchi K, Sugawara M, Sudo H, Ekawa M, *et al.* (2003) A semiconductor optical amplifier with an extremely-high penalty-free output power of 20 dBm achieved with quantum dots. Post deadline paper in the 29th European Conference on Optical Communication.

Chuang SL (1995) *Physics of Optoelectronic Devices*. New York: Wiley.

Kartalopoulos SV (2000) *Introduction to DWDM Technology*. Piscataway: IEEE Press.

Mimura T, Hiyamizu S, Fujii T, and Nanbu K (1980) A new field-effect transistor with selectively doped GaAs/n-$Al_xGa_{1-x}As$ heterojunctions. *Japanese Journal of Applied Physics* 19: L225–227.

Onaka H (1996) 1.1Tb/s WDM transmission over a 150 km 1.3 mm zero-dispersion single-mode fiber. OFC '96 Technical Digest PD-19. San Jose.

Sugawara M (1999) *Self-Assembled InGaAs/GaAs Quantum Dots*. San Diego: Academic Press.

Diamagnetism

P F de Châtel, New Mexico State University, Las Cruces, NM, USA

Diamagnetic materials do not order magnetically and have a negative susceptibility, which means the magnetization is pointing opposite to the magnetic field. Possibly, there is also a paramagnetic susceptibility. The material is classified as diamagnetic if the sum of the two contributions is negative. Nevertheless, the diamagnetic susceptibility can be determined in both kinds of materials.

Diamagnetism is a very small effect, but all the more puzzling. It entails a response of all known materials to an external magnetic field, which manifests itself in a magnetization opposing that field. It appears thus that the magnetic field generates a magnetic dipole, whose orientation is most unfavorable, being at a maximum of the potential energy as a function of the angle between the magnetic field and the magnetization,

$$E = -\boldsymbol{B}\cdot\boldsymbol{M} = BM > 0$$

There is no simple, intuitive explanation for this paradox. This is not surprising though, considering that our intuition is mostly based on classical physics. The celebrated Bohr–van Leeuwen theorem states that the classical free energy of a system of charged particles moving in a magnetic field is independent of the field. Since the magnetization is proportional to the derivative of the free energy with respect to the field,

$$\boldsymbol{M} = -\frac{1}{V}\frac{\partial F}{\partial \boldsymbol{B}} \qquad [1]$$

this theorem implies that classical physics, in this case classical mechanics combined with classical statistics, cannot account for any field-induced magnetization, be it paramagnetic (aligned along the field) or diamagnetic (opposing). In the case of diamagnetic metals, whose charge carriers behave like free electrons, this conclusion contradicts what one would expect from electrodynamics. Charged classical free particles exposed to a magnetic field move in circles. Lenz's law of induction requires that the current loops represented by the orbiting electrons produce a magnetization opposed to the field which induces the currents; this should result in a diamagnetic response. It is no wonder that Niels Bohr was fascinated by this paradox. His main interest was the stability of atomic orbitals, which also contradicts classical expectation. He has formulated the rule of angular-momentum quantization, which, imposed on the classical treatment, accounted for the observed stable orbitals. As it will be shown below, the same rule, applied in the presence of a magnetic field, gives the correct answer for the diamagnetism of free atoms. The diamagnetism of free electrons cannot be understood in this semiclassical way, here a fully quantum mechanical treatment is needed.

The formal proof of the Bohr–van Leeuwen theorem is rather simple, but conceals the clue to its counterintuitive implication. Indeed, if one writes the free energy, $F = -k_B T \ln Z$, of a system of N identical particles (taking electrons of mass m_e and charge $-e$), the partition function has the form

$$Z = \prod_{i=1}^{N} \int \cdots \int \mathrm{d}\boldsymbol{r}_i\, \mathrm{d}\boldsymbol{p}_i \times \exp(-H(\{\boldsymbol{r}_i\},\{\boldsymbol{p}_i\})/k_B T) \qquad [2]$$

where H is the Hamiltonian, a function of all position vectors $\boldsymbol{r}_i$ and canonical momenta $\boldsymbol{p}_i$. The latter are written as

$$\boldsymbol{p}_i = m_e \boldsymbol{v}_i - e\boldsymbol{A}(\boldsymbol{r}_i)$$

where $\boldsymbol{A}$ is the vector potential describing the magnetic field, which enters the Hamiltonian via the kinetic energy,

$$E_{\mathrm{kin}} = \frac{1}{2} m_e v_i^2 = \frac{1}{2m_e}[\boldsymbol{p}_i + e\boldsymbol{A}(\boldsymbol{r}_i)]^2$$

To prove the independence of Z from $\boldsymbol{A}$, one changes the integration variables $\boldsymbol{p}_i$ to $\boldsymbol{p}'_i = \boldsymbol{p}_i + \boldsymbol{A}(\boldsymbol{r}_i)$ and realizes that this will not change the result of integration, because the shift of the limits of integration, $\pm\infty$, has no effect. The vector potential thus being eliminated from eqn [2], the partition function must be independent of the magnetic field.

To understand the implications of this freedom to shift parameters, it is instructive to carry out the differentiation in eqn [1] for the simple case of a system of N_a identical atoms. If the electron–electron interaction is treated in the mean-field approximation, the N electrons localized on the same atom are described by the Hamiltonian

$$H = \sum_{i=1}^{N} H^{(1)}(\boldsymbol{r}_i, \boldsymbol{p}_i)$$

with

$$H^{(1)}(\boldsymbol{r}, \boldsymbol{p}) = V(r) + [\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})]^2/2m_e \qquad [3]$$

where $V(r)$ is the potential consisting of the potential of the nucleus and of the "other" $N-1$ electrons, the latter contribution being spherically averaged (central-field approximation). The partition function (eqn [2]) becomes then the product of N identical integrals,

$$\begin{aligned} Z &= \left\{ \int\int \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{p} \exp(-H^{(1)}(\boldsymbol{r}, \boldsymbol{p})/k_B T) \right\}^N \\ &= \{Z^{(1)}\}^N \end{aligned}$$

and, consequently, the free energy is a sum of N identical terms, $F = -k_B T N \ln Z^{(1)}$. Choosing the gauge

$$\boldsymbol{A} = -\frac{1}{2}(\boldsymbol{r} \times \boldsymbol{B}) \qquad [4]$$

one obtains

$$\begin{aligned} \boldsymbol{M} &= -\frac{N_a}{V}\frac{\partial F}{\partial \boldsymbol{B}} \\ &= -\frac{N_a}{V}\frac{eN}{2m_e}\int\int \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{p}\, \boldsymbol{r} \times \left[\boldsymbol{p} - \frac{e}{2}(\boldsymbol{r} \times \boldsymbol{B})\right] \\ &\quad \times \frac{\exp(-H^{(1)}(\boldsymbol{r}, \boldsymbol{p})/k_B T)}{Z^{(1)}} \end{aligned} \qquad [5]$$

where N_a is the number of atoms. The result is seen to be proportional to the thermal average of the kinetic angular momentum, $\boldsymbol{r} \times m_e \boldsymbol{v}$. This is not unexpected, because currents, the sources of magnetization, are related to the velocity, not the canonical momentum $\boldsymbol{p}$, of the charge carriers. The same holds for the kinetic energy in the Hamiltonian [3], which appears in the statistical weight factor $\exp(-H^{(1)}/k_B T)$.

Consider now two points in the phase space of a single electron, which are given by the same position vector $\boldsymbol{r}$ and opposite velocities,

$$\begin{aligned} &\boldsymbol{r}_1 = \boldsymbol{r}_2 = \boldsymbol{r} \\ &\boldsymbol{p}_1 - \frac{e}{2}(\boldsymbol{r}_1 \times \boldsymbol{B}(\boldsymbol{r}_1)) = -\left(\boldsymbol{p}_2 - \frac{e}{2}(\boldsymbol{r}_2 \times \boldsymbol{B}(\boldsymbol{r}_2))\right) \end{aligned}$$

The statistical weight factors associated with these two points are identical, whereas the kinetic angular momenta have opposite signs and the same magnitude. Therefore, the sum of their contributions to the integral will vanish. The cancelation would have been more obvious, if $\boldsymbol{p}$ was shifted to $\boldsymbol{p}' = \boldsymbol{p} + (e/2)(\boldsymbol{r} \times \boldsymbol{B}(\boldsymbol{r}))$. In that case, the integrand would have become an odd function of $\boldsymbol{p}'$ resulting obviously in a vanishing $\boldsymbol{M}$. Again, the freedom to shift the canonical momentum at a given value of $\boldsymbol{r}$, which amounts to shifting the canonical angular momentum $\boldsymbol{L} = (\boldsymbol{r} \times \boldsymbol{p})$, turns out to be important in the derivation. As classical mechanics puts no constraints on $\boldsymbol{L}$ and, according to the ergodic theorem underlying classical statistics, a system will visit all possible states of different values of $\boldsymbol{L}$, it is understandable that $\boldsymbol{M}$ has to vanish on average.

These considerations suggest that the classical realm need not be left altogether. It may suffice to impose some constrains on the angular momentum, in the spirit of the Bohr model of atoms, to get sensible results for the magnetization. As a uniform magnetic field along the z-axis, $\boldsymbol{B} = (0, 0, B)$, leaves all rotations around the z-axis as symmetry operations, it is postulated that L_z be limited to the values $m\hbar$, $m = \pm 1, \pm 2, \pm 3, \ldots$. To impose this condition, each integral over the entire six-dimensional $(\boldsymbol{r}, \boldsymbol{p})$ phase space must be replaced by a sum of integrals, each over a subspace P_m containing only $(\boldsymbol{r}, \boldsymbol{p})$ points, for which $(\boldsymbol{r} \times \boldsymbol{p})_z = m\hbar$. In particular,

$$\begin{aligned} Z^{(1)} &= \sum_m \int\int_{P_m} \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{p} \exp(-H^{(1)}(\boldsymbol{r}, \boldsymbol{p})/k_B T) \\ &= \sum_m Z_m^{(1)} \end{aligned}$$

By the definition of the subspace P_m

$$\begin{aligned} &\int\int_{P_m} \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{p} (\boldsymbol{r} \times \boldsymbol{p})_z \exp(-H^{(1)}(\boldsymbol{r}, \boldsymbol{p})/k_B T) \\ &\quad = m\hbar Z_m^{(1)} \end{aligned}$$

so that eqn [5] can be written as

$$M_z = -\frac{N_a}{V}\sum_m \left\{ \frac{eN}{2m_e}\frac{Z_m^{(1)}}{Z^{(1)}}\hbar m + \frac{e^2 N}{4m_e} \langle \rho^2 \rangle_m^{\mathrm{th}} B \right\}$$

where $NZ_m^{(1)}/Z^{(1)} = p_m$ is the thermal occupation probability of the $L_z = m\hbar$ orbit. Introducing the

Bohr magneton, $\mu_B = e\hbar/2m_e$, the first term takes the form $p_m m\mu_B$, which can be recognized as the paramagnetic contribution to the magnetization. In the second term, which makes evidently a negative, diamagnetic contribution to the magnetization, $\langle \rho^2 \rangle_m^{th} = \int\int_{P_m} d\boldsymbol{r}\, d\boldsymbol{p}\, \rho^2 \exp(-H^{(1)}(\boldsymbol{r},\boldsymbol{p})/k_B T)/Z^{(1)}$ is the thermal average of ρ^2 for the same orbit m. Here, it is assumed that the $\boldsymbol{r}$ vector is in the $z = 0$ plane and $|\boldsymbol{r}| = \rho$, constant. It then follows that

$$M_{dia} = -\frac{N_a}{V}\frac{e^2}{4m_e}\sum_m \langle \rho^2 \rangle_m^{th} B$$

In practice, no thermal excitation out of the ground state will occur and $N\langle \rho^2 \rangle_m^{th}$ can be replaced by the average of the squared radii of occupied orbits. Clearly, at this stage the simple Bohr model fails. It would be a tortuous extension to introduce the principal quantum number n on the basis of the Bohr–Sommerfeld model and the angular momentum quantum number ℓ on an *ad hoc* basis. Even so, only for $T = 0$ would the introduction of the Pauli principle suffice to give the correct result. For $T > 0$, Fermi–Dirac statistics should be applied, which is a further step away from classical physics. In a proper quantum mechanical treatment, $(2/3)\langle r^2 \rangle$ enters, instead of $\langle \rho^2 \rangle^{th}$ where $\langle \ldots \rangle$ stands for the expectation value and the factor 2/3 accounts for the difference between the expectation values of $r^2 = x^2 + y^2 + z^2$ and $\rho^2 = x^2 + y^2$. The diamagnetic susceptibility is thus

$$\chi_{dia}^{(Langevin)} = \frac{M_{dia}}{H} = -\frac{N_a}{V}\frac{\mu_0 e^2}{6m_e}\sum_j \langle r_j^2 \rangle \qquad [6]$$

where the summation is over the occupied orbits j on each atom. Langevin's result for atomic diamagnetism was brought to this form by Pauli.

It is remarkable that Planck's constant does not appear in the expression for the diamagnetic susceptibility, whereas the importance of angular-momentum quantization in its derivation is obvious. Also, eqn [6] is an appealing result, as it enables a seemingly classical interpretation in terms of currents i_j, induced on each orbit j. Such currents will generate magnetic moments $m_j = i_j\alpha_j \propto \langle r_j^2 \rangle$, where α_j is the area enclosed by the orbit j. The sign of the moments (opposite to $\boldsymbol{B}$) can be understood in terms of Lenz's law and their magnitudes can be derived from Faraday's law of induction. This derivation, however, invokes well-defined orbits, which are modeled by tiny current loops. Bohr has grafted such orbits onto the classical treatment of particles moving in a potential proportional to r^{-1} and postulated the quantization of angular momentum. Remarkably, this supplement to the classical description was sufficient to understand quantized energy levels and line spectra, as well as diamagnetism.

The validity of eqn [6] turns out to extend beyond the case of atomic orbits. Most notably, it is applicable to ring-shaped molecules, whose structure offers extended orbits. The diamagnetic susceptibility of substances containing such molecules is known to be anisotropic, being larger in magnitude if the magnetic field is applied perpendicular to the plane of the ring-shaped molecules,

$$|\chi_{dia}^{\perp}| > |\chi_{dia}^{\parallel}|$$

The interpretation of this phenomenon in terms of "ring currents" goes beyond the qualitative argument that the induced magnetic moments are maximized in the perpendicular configuration. For benzene, C_6H_6, the anisotropic susceptibility

$$\chi^{an} = \chi_{dia}^{\perp} - \chi_{dia}^{\parallel}$$

is found to be close to what eqn [6] gives. The summation in eqn [6] is to be taken over the six carbon p electrons, which, not being localized in sp^2 hybrid states, are free to move around the molecule:

$$\chi_{benzene}^{an} = -6\frac{A}{V_m}\frac{\mu_0 e^2}{4m_e}\rho^2 = -\frac{A}{V_m}\frac{i_r}{H}\alpha_{hex} \qquad [7]$$

Here, A/V_m is the density of molecules, A being Avogadro's number and $V_m = 89 \times 10^{-6}\,m^3$ the molar volume of benzene, $\rho = 0.139$ nm is the distance from the axis of the molecule to the carbon nuclei and i_r is the total ring current carried by the six delocalized electrons. Pauling has shown in 1936 that the ratio of the anisotropic susceptibility of other, more complicated aromatic hydrocarbons to that of benzene is correctly given by a simple model of conducting networks, assuming that the "resistance" of a carbon–carbon bond is the same in all such molecules. The ultimate success of this model is the correct estimation of the anisotropic susceptibility of graphite. The crystal structure of graphite consists of honeycomb-like layers built of benzene-like units. Assuming that in each hexagonal unit a ring current is induced of the same magnitude as in benzene, one finds that the current of neighboring unit cells cancel and the current density inside the material vanishes. However, the outermost hexagons have two sides with uncompensated currents i_r flowing in a zigzag pattern at 30° from the orientation of the surface. The resulting surface current density is $j_S = i_r \cos 30°/d$, where $d = 0.335$ nm is the interlayer distance in graphite. The surface current density j_S generates a magnetization $M_{graphite} = -j_S = -i_r\sqrt{3}/2d$

inside the sample. This is to be compared with $M_{\text{benzene}} = -(A/V_{\text{m}})\, i_{\text{r}} \alpha_{\text{hex}}$ (cf. eqn [7]), leading to

$$\frac{\chi^{\text{an}}_{\text{graphite}}}{\chi^{\text{an}}_{\text{benzene}}} = \frac{\sqrt{3} V_{\text{m}}}{2A\alpha_{\text{hex}} d} = 7.6 \qquad [8]$$

which is close enough to the experimental value of 6.9, considering the simplicity of the model.

Semiclassical and empirical schemes like the ones used above for free atoms and aromatic compounds, respectively, cannot be applied to the diamagnetic response of conduction electrons in metals. Even the simplest model, the free-electron gas, has been a major challenge, until Landau gave his derivation of what has come to be called Landau diamagnetism. In fact, in most textbooks the Bohr–van Leeuwen theorem is only mentioned in this context, even though its validity is not violated by the presence of atomic or molecular potentials.

The Schrödinger equation was four years old, when Landau (who was twenty-one at the time) solved it for free electrons in the uniform magnetic field (eqn [6]). Instead of the symmetric gauge (eqn [4]), he has chosen

$$\boldsymbol{A} = (0, xB, 0) \qquad [9]$$

which enabled the separation of variables. The resulting differential equation in x was found to be identical with the Schrödinger equation of a harmonic oscillator with a force constant e^2B^2/m_{e}, while the ones in y and z were satisfied by plane-wave solutions. The latter is reasonable for the z dependence, considering that the Lorentz force is perpendicular to $\boldsymbol{B}$, so that the motion along the z axis is free. The density of the eigenstates described by $\mathrm{e}^{\mathrm{i}k_z z}$ is that of a one-dimensional electron gas,

$$\rho(\varepsilon) = \frac{2eB}{\hbar^2}\sqrt{2m_{\text{e}}\varepsilon} \qquad [10]$$

The motion in the (x,y) plane gives a discrete spectrum, with highly degenerate states at the energies

$$\varepsilon_n = \left(n + \frac{1}{2}\right)\frac{\hbar e}{m_{\text{e}}}B \qquad [11]$$

the Landau levels, which correspond to the stationary states of the harmonic oscillator. The resulting total density of states has a spiked structure. A singular contribution of the form [10] begins at each level [11]. These $\rho(\varepsilon - \varepsilon_n)$ contributions stand for the motion in the z direction while the motion in the (x, y) plane corresponds to the nth level of the harmonic oscillator. As to the x- and y-dependence of the wave functions, they are not suggestive of the circular motion resulting from the classical equations of motion for a charged particle subjected to a uniform magnetic field. This may seem to be due to the choice of the asymmetric gauge [9]. However, it is hardly surprising that the classical motion is not implied by the calculation of an effect, which does not occur at all in the classical theory.

The maxima in the density of states at energies separated by $\hbar eB/m_{\text{e}}$ result in a total energy, which is a periodic function of $1/B$. Through eqn [1], a similar periodicity arises in the magnetization, which is the signature of the de Haas–van Alphen effect. In the low-field limit, the same B-dependent total-energy expression yields the diamagnetic susceptibility

$$\chi^{(\text{Landau})}_{\text{dia}} = -\frac{n\mu_{\text{B}}^2}{\varepsilon_{\text{F}}}$$

where n is the electron density and ε_{F} is the Fermi energy. This is seen to be $-1/3$ times the Pauli paramagnetic susceptibility, a relationship, which cannot be expected to be valid for metals in general, but seems to hold for the so-called "simple" metals, for which the nearly free-electron approximation works.

See also: Magnetism, History of.

PACS: 75.20.Ck; 75.20.En; 75.30.Cr

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Reinhart and Winston.

Chikazumi S (1965) *Physics of Magnetism*. New York: Wiley.

O'Handley RC (2000) *Modern Magnetic Materials: Principles and Applications*. New York: Wiley.

Peierls RE (1955) *Quantum Theory of Solids*. Oxford: Oxford University Press.

Van Vleck JH (1932) *The Theory of Electric and Magnetic Susceptibilities*. Oxford: Clarendon Press.

Nomenclature

A	vector potential (A)
B	magnetic field (T)
L	canonical angular momentum ($\mathrm{kg\,m^2\,s^{-1}}$)
M	magnetization ($\mathrm{A\,m^{-1}}$)
$m_e \boldsymbol{r} \times \boldsymbol{v}$	kinetic angular momentum (of electron) ($\mathrm{kg\,m^2\,s^{-1}}$)
$m_e \mathrm{v}$	kinetic momentum (of electron) ($\mathrm{kg\,m\,s^{-1}}$)
p	canonical momentum ($\mathrm{kg\,m\,s^{-1}}$)

Diamond Anvil Cells

W A Bassett, Cornell University, Ithaca, NY, USA

Introduction

The extraordinary properties of diamond and the very high pressures that can be achieved simply by reducing sample size make the diamond anvil cell (DAC) capable of pressures up to and beyond those existing at the center of the earth. Resistance heating and laser heating can produce temperatures up to and beyond those at the center of the earth. The very simple principle of the DAC allows great versatility of design for accommodating many analytical techniques. Background, development, and applications of the DAC are described in this article.

History and Background

Alvin Van Valkenburg at the National Bureau of Standards (now the National Institute of Science and Technology) reasoned that he could make infrared absorption measurements on a sample under pressure by squeezing the sample between the flat faces of two diamond anvils. He did not know if the faces were perfectly parallel and so he placed his device on the stage of a microscope and looked through the diamonds at the sample. When he saw the sample, he immediately realized that he had something extraordinary. Although the diamonds were in fact not very parallel, he could see that parts of the sample looked different because they were at high pressure (HP); and the diamond anvil cell was born. Visual observation continues to the present as the most important technique used with DACs. Van Valkenburg found that he could solidify water, alcohol, and other liquids by placing them in gaskets formed by drilling holes in metal foils and placing them between the diamond faces. He also found he could see phase transitions in single crystals of calcite and other solids when surrounded by a fluid in a gasket. The National Bureau of Standards group soon found they could examine samples under pressure by X-ray diffraction as well.

Since that time there have been more than a dozen additions to the list of analytical techniques successfully used to study samples at high pressures and temperatures (HPT) in the DAC. These include a variety of applications employing visible light, infrared, X-rays, electrical resistance, and magnetic susceptibility.

Principles of the Diamond Anvil Cell

Diamond has many properties that make it ideal for use as anvils in high-pressure (HP) devices. It is the hardest material known, has elastic moduli with very large values, is transparent to large portions of the electromagnetic spectrum, is a good electric insulator, is an excellent thermal conductor, and is readily available as nearly perfect single crystals. Gem diamonds are available in high-quality, ready-cut shapes close to the most desirable shapes for anvils. Grinding and polishing the culet point on a brilliant-cut faceted diamond is all that is needed to turn it into an anvil face. Although metastable, diamond has a remarkable resistance to change with increased pressure and temperature up to hundreds of GPa and $\sim 1500\,\mathrm{K}$.

Great versatility is possible because the principle of the diamond anvil cell is simple: two diamond anvils driven together with a sample trapped between them (**Figure 1**). Because pressure is defined as force divided by area, it can be increased by increasing force or decreasing area. The DAC clearly benefits from the latter. As analytical techniques have become miniaturized, diamond anvil faces have been reduced to dimensions only microns across while forces continue to be generated by screws, levers, and small hydraulic or pneumatic devices. Because such small anvil faces are needed for generating very high pressures, precise parts with precise motions are essential for maintaining alignment. DACs need good support

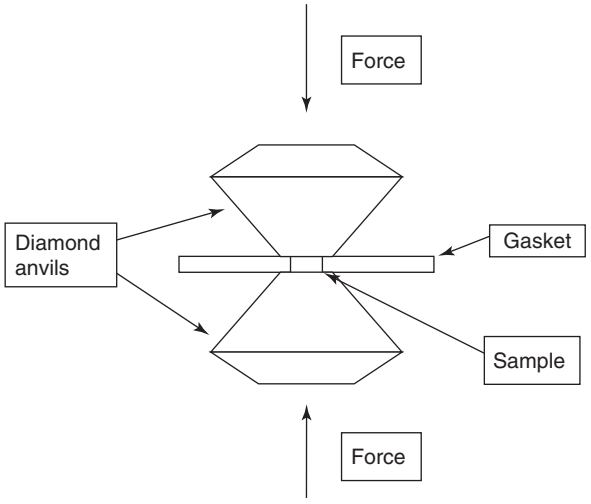


Figure 1 The diamond anvil cell (DAC) consists of two gem diamonds with their culet points ground and polished to form anvil faces. The diamonds are driven together against a sample contained in a gasket consisting of a hole drilled in a foil. This very versatile principle allows a wide variety of configurations depending on the type of analytical technique being applied.

for the anvils, usually a hard material such as tungsten carbide, while providing adequate access for analytical techniques, usually an opening in the support and/or in the plane of the sample. The principle is so simple that DACs can be adapted to the geometric needs of many analytical techniques. As a result, there are numerous DAC designs.

Gaskets play an important role in the many applications of DACs. The properties of gasket materials are crucial and are selected for specific applications. The rheologic properties are usually the most important. However, transparency to X-rays, electrical properties, and chemical properties are important as well. Gasket materials include rhenium, tungsten, iridium, spring steel, inconel, stainless steel, brass, boron–epoxy mixture, aluminum, and beryllium.

Heating a sample in a DAC can be achieved by

1. external resistance heating including use of heaters that surround the whole cell, heaters that surround just the anvils, heaters that surround the seats that the anvils sit on (**Figure 2**), and graphite heaters that are imbedded within the anvils;
2. internal resistance heaters are heaters that are insulated from the anvils and therefore do not heat the anvils to temperatures as high as sample temperature (**Figure 3**). They consist of an electrically conducting sample wire, a metal strip with a hole to receive the sample, or graphite sandwich; and
3. laser heating also heats the sample without heating the anvils (**Figure 4**).

External resistance heating (**Figure 2**) requires careful choice of materials that will resist softening or deformation with temperature change. Because some of the materials, including diamond, readily oxidize at high temperature, an inert or slightly reducing gas is needed to prevent oxidation. The main advantage of external heating is the very constant and uniform temperatures that can be achieved. Pressures can also be held constant when the cell is designed to eliminate or minimize the effects of dimensional changes with temperature. Pneumatic devices and springs that are kept cool are especially important for achieving this. The main disadvantage is the upper limit of 1500 K at which diamond begins to graphitize.

Internal resistance heating (**Figure 3**) has the advantage that it heats only the sample and can, therefore, generate much higher temperatures than can be tolerated by the diamond anvils. The upper limit for internal resistance heating is imposed by the melting point of the resistor. However, that can be very high if Re or W is used.

Laser heating (**Figure 4**) has employed yttrium aluminum garnet (YAG), yttrium lithium fluoride, and CO_2 lasers with their beams focused through the anvils onto the sample. The sample must be thermally insulated from the diamond anvils by a transparent

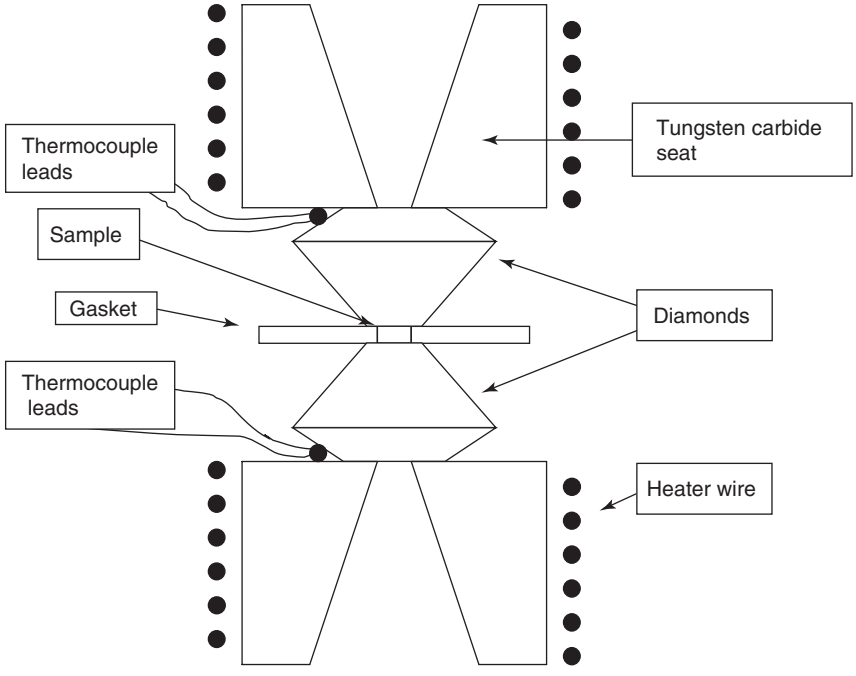


Figure 2 Samples can be heated to 1500 K while under pressure by means of resistance heaters wound around the seats supporting the diamond anvils. Thermocouples placed in contact with the diamond anvils measure temperatures to within a few kelvin. Corrections can be made by observing melting of standard materials in the sample chamber. Access can be through the diamonds or the gasket.

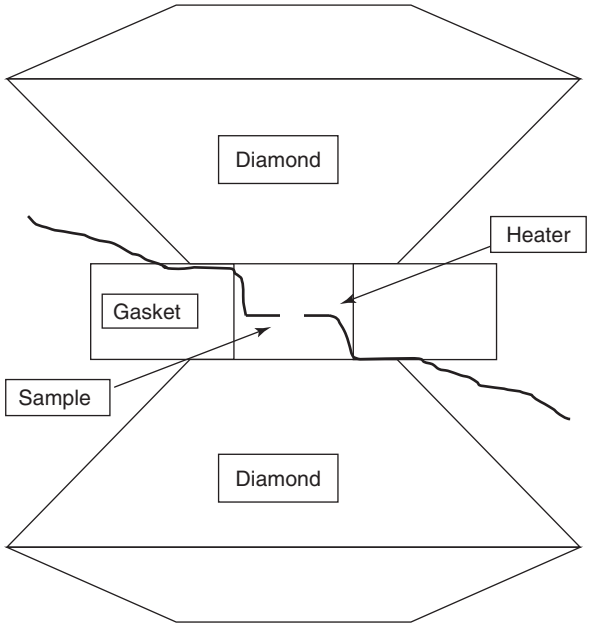


Figure 3 An internal resistance heater, usually constructed from rhenium, can be used to heat a sample to much higher temperatures (3000 K) than the diamonds can tolerate because it is insulated from them. Temperatures are measured by spectroradiometry.

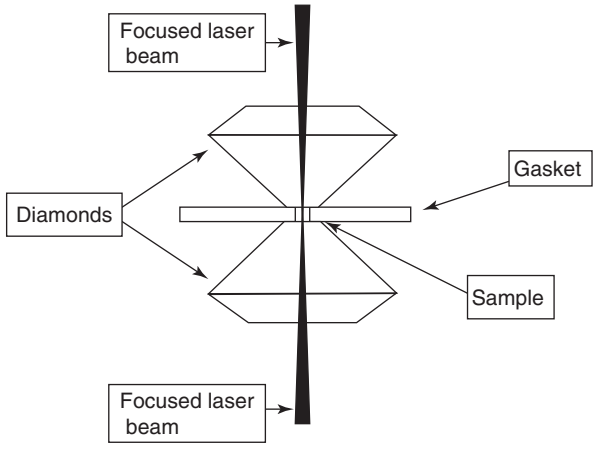


Figure 4 Laser heating can raise the temperatures of samples to several thousands of kelvin. When a laser beam is split and directed to the sample through both anvils, more uniform and constant temperatures can be achieved than heating from one side alone. Temperature gradients within the sample are minimized by using the larger TEM $_{01}$ laser mode and focusing the probing X-ray beam to a very small spot within the heated area.

material, which protects the diamonds and serves as a pressure medium. Although continuous power delivered to both sides of the sample has yielded the best quantitative results, single-side and pulsed laser heating have been used extensively. Temperatures of thousands of kelvin can be achieved, especially if very short pulses are used. Sample design for laser heating is much simpler than for internal resistance heating, but requires a rather large and complex optical system outside of the DAC.

Low temperatures down to that of liquid helium can be achieved by passing cold gases through the DAC or injecting liquid nitrogen or liquid helium directly into the cell and letting it boil there. Another approach has been to completely immerse the DAC into liquid nitrogen or liquid helium.

Pressure Measurement

Accurate methods for determining pressure can be divided into two groups: (1) primary, those that are independent and (2) secondary, those that must be calibrated using a primary pressure scale.

Primary

Force divided by area can provide reliable, accurate pressure measurements but only to pressures far less than what DACs are capable of because of the limitations of the apparatus for making such determinations, usually piston-cylinder devices. Phase transitions such as those in elemental bismuth were used as calibration points for transferring pressure measurements to other types of apparatus in which those same phase transitions could be observed.

One of the most significant advances was Decker's development of an equation of state (EoS) for NaCl from first principles. This could be applied directly when X-ray diffraction was used and NaCl was mixed with the sample in a DAC. Its use for calibrating secondary pressure scales, such as the ruby fluorescent scale, was also very valuable.

Dynamic shock-wave measurements are also an important source of primary pressure determination but rely on corrections when data are converted from high-temperature Hugoniot measurements to isothermal EoS. Nonetheless, dynamic measurements on some metals have provided an important means for calibrating the secondary pressure scales, especially the ruby fluorescence scale, to higher pressures.

Recent primary pressure scales have been based on combining data from direct compression measurement using X-ray diffraction (XRD) and sound velocity measurements, which yield data related to the pressure derivative of compression. From the two sets of data, it is possible to derive completely independent values of pressure as a function of lattice parameters. Brillouin scattering in the DAC and sound interferometry have been used to measure sound velocities.

Secondary

A secondary pressure scale is one that offers an observable change as a function of pressure. There

are many materials which display observable phase transitions at calibrated pressures – Bi, AgI, KCl, and NaCl, to mention a few. One or two of these can be included along with a sample as a pressure indicator. There are many materials whose EoS have been calibrated so that a measurement of their lattice parameters by XRD can be used for determining pressure. Sometimes the sample itself can be used in cases where its EoS is well defined and some other property is being investigated.

The most widely used secondary pressure scale is the ruby fluorescence scale. The R_1 and R_2 emissions from ruby when excited by some higher energy radiation shift monotonously with pressure. The fluorescence is typically excited by an argon ion laser and the emission is analyzed by a grating spectrometer. The positions of the R_1 and R_2 lines in the spectrum provide a convenient and inexpensive method for determining pressure. In addition, the shapes and separation of the R_1 and R_2 lines can be used as indicators of deviatoric stress and pressure gradient. Other fluorescent materials such as Cr:MgO and Sm:YAG have been used as well, but ruby continues to be the most popular.

Raman spectra of diamond show shifts with pressure and can be used as a secondary pressure scale. At lower pressures, diamond particles synthesized from ^{13}C are used to avoid interference from the diamond anvils. At higher pressures, natural diamond can be used.

Temperature Measurement

Temperature measurements in externally heated DACs can be accomplished by placing thermocouples in contact with the diamond anvils. Because diamonds are such good thermal conductors, discrepancies between sample temperature and thermocouple temperature are rarely more than a few degrees celsius even at the highest temperatures. Corrections to thermocouple readings can be determined by observing melting of standard materials such as $NaNO_3$ and NaCl placed in the sample chamber. The same thermocouples used for high-temperature measurements can be used for the low-temperature measurements.

Internal resistance heating and laser heating are utilized in order to achieve temperatures in excess of those that cause damage to the diamond anvils ($\sim$1500 K). The most commonly used method of temperature determination is spectroradiometry, which consists of fitting the spectrum of incandescent light from the hot sample to a blackbody or graybody curve. Steep temperature gradients and unknown emissivities can be sources of error.

Recently, the intensity ratio of stokes and anti-stokes peaks in X-ray Raman spectra has been used to determine temperatures. They were found to be in good agreement with those determined by spectroradiometry.

Pressure Measurement at High Temperatures

Primary

Accurate pressure measurement in samples at high temperature has been an area of much concern. Neither sample pressure (isobaric) nor sample volume (isochoric) can be assumed to remain constant as temperature is changed in most sample configurations. In these cases, an internal calibrant is needed. Decker's EoS for NaCl is a full EoS including the effect of temperature on molar volume and is, therefore, suitable for pressure measurement at high temperature. The full EoS of gold has also been determined and is desirable for samples in which only a small amount of pressure calibrant is wanted or where recrystallization needs to be avoided.

There are two configurations, however, that can greatly reduce or eliminate the effect of temperature on pressure. A very small opaque sample surrounded by a much larger volume of transparent pressure medium can be locally heated by laser and remain nearly isobaric because of the relative sizes. Isochoric conditions can be achieved at relatively low pressures by cycling a sample in a gasket, such as rhenium, to high temperature several times. If forces are not great, the gasket finishes relaxing on about the third cycle and the volume can be shown to remain constant with temperature within experimental error. If the full EoS of the contents is known, as with H_2O, the pressure can be calculated from the temperature. This approach has been especially valuable in fluid studies where the pressure is allowed to develop as a result of heating.

Secondary

Attempts to use ruby and Sm:YAG to measure pressure at high temperature have been less than satisfactory. The fluorescence emission from ruby fades with temperature and the emissions from Sm:YAG are complex to follow.

There are several rapidly-running reversible phase transitions that have been calibrated, for example, $BaTiO_3$, $PbTiO_3$, $Pb_3(PO_4)_2$, and α–β transition in quartz. These have specific trajectories in P–T space and, therefore, are not useful for determining pressure at just any location in the P–T space. They are, however, very useful for calibrating fluid EoS described above. Re and other metals have been calibrated for pressure measurement at very high temperatures using XRD.

Analytical Methods

Visual observation is the easiest and most useful method for studying a sample in a DAC. It requires only using a microscope to look at it through the diamond anvils. Visual observation is valuable for not only alignment and centering but can be used to observe phase transitions, coexistence of phases, and optical properties. Phase diagrams can be developed by observing phases as P and T are changed.

XRD of samples under pressure in DACs has been one of the most important applications of DACs. Angle dispersion X-ray diffraction (ADXD) using monochromatic radiation and energy-dispersion X-ray diffraction (EDXD) using white radiation are both used, usually by passing an X-ray beam through one anvil to the sample and collecting scattered X-rays through the other anvil (**Figure 5**). X-rays that pass through a transparent gasket can also be collected (**Figure 6**). Both polycrystalline and single-crystal samples can be studied. The advent of synchrotron radiation (SR) with its high intensities, very smooth white spectrum, and excellent collimation has reduced exposure times by orders of magnitude while improving resolution and offering a greater variety of configurations. Improved solid-state detectors have also played a major role.

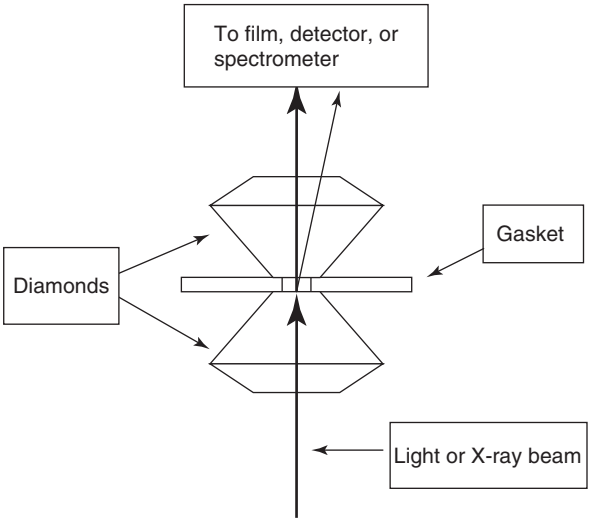


Figure 5 The most commonly used configuration is to pass electromagnetic radiation though one diamond anvil to the sample and to observe or analyze the radiation emerging through the other diamond anvil.

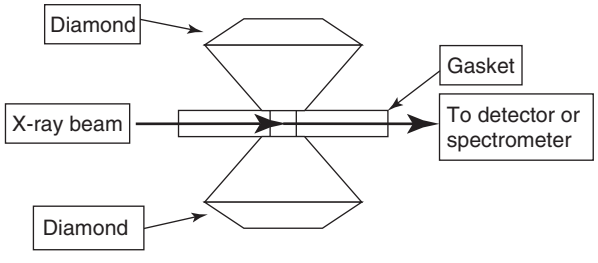


Figure 6 Gasket materials transparent to X-rays such as beryllium and boron–epoxy mixture allow incident X-rays to enter through one side of the gasket and scattered X-rays to be collected through the other side.

Emitted visible light from a sample in a DAC can be spectrally analyzed. Fluorescence emission may be excited by X-rays, ultraviolet light, or visible light of shorter wavelength. The trajectory of an X-ray beam as it passes through a fluorescent diamond can be located visually by emission of visible light. The ruby method for determining pressure depends on the excitation of fluorescent emission in small ruby chips included in a sample. Incandescent light from a hot sample contained in a DAC can be spectrally analyzed and fit to a blackbody curve to determine the temperature. Sample fluorescence has been an important source of fundamental information, especially for studying the effect of pressure and temperature on organic compounds. Lasers are favored for stimulating fluorescence because of their high intensity and because they make it possible to pinpoint parts of the sample.

Absorption spectra of light passing through a sample (**Figure 5**) as well as spectra of reflected light (**Figure 7**) provide useful information about bonding and band gap changes as a function of pressure and temperature. Imaging can allow selected portions of a sample to be analyzed.

Infrared (IR) absorption spectroscopy requires carefully selected diamonds (usually type II) to be transparent over the desired range of IR-active frequencies. IR absorption at high pressure has been especially important for understanding the effects of pressure and temperature on organic materials.

Raman scattering has been a very important analytical technique for HPT samples. Light scattered from the sample can be analyzed to determine Raman-active

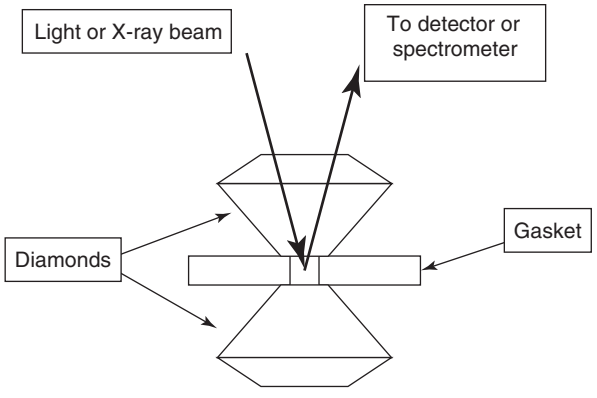


Figure 7 Observation and analysis of radiation scattered back through the same diamond it entered provide an important source of information for many types of studies.

vibrational frequencies. Vibrational frequencies are an important source of information for identifying phases, examining bonding, and studying the relationships between vibration frequencies and phase transitions. Hyper-Raman spectroscopy employs a three-photon process and can detect not only far IR vibrational frequencies but can also detect frequencies ("silent" modes) that are not normally IR or Raman active.

Brillouin scattering has been an important source of information on the effects of P and T on the elastic properties of single-crystal samples. A light beam from a laser, usually an Ar^+ laser, is directed through one of the diamond anvils onto an oriented single-crystal sample in a DAC. Light scattered by interaction with phonons in the sample contains frequency shifts (usually smaller than in Raman) that can be used to calculate the velocities of the phonons. If the anvil faces exactly bisect a 90° angle between the incident beam and scattered light, no corrections need to be made for the refractive index of the sample, medium, and anvils (**Figure 8**). These phonon velocities, in turn, can be used to determine the effects of P and T on elastic moduli.

Impulsive stimulated scattering (ISS) is another technique that yields information about sound velocities and elastic moduli. Sound waves in the ISS method are generated within the sample by means of laser radiation. Laser light from a different source is Bragg reflected from these sound waves. Its intensity fluctuation can then be used to determine sound velocities.

X-ray spectroscopy (XRS) which includes emission, absorption, and scattering is now the area of most active development of new techniques and applications in HPT research. X-ray emission spectroscopy (XES) consists of analyzing spectra of radiation excited by an incident X-ray beam as it passes out through the anvils (**Figures 5** and **7**) or the gasket (**Figure 6**). Laser-machined diamond anvils and very transparent gaskets are two modifications that now make it possible to analyze all elements heavier than Ca. While elemental analysis is the most important application, newly developed sub-eV energy resolution makes it possible to show a line shape which can be used to provide valuable information on filled electronic states as well as valence and magnetic states such as high-spin to low-spin transition in Fe.

X-ray absorption spectroscopy (XAS) consists of analyzing the spectra of photons absorbed when their energies exceed the excitation energy of deep-core electrons. Laser-drilled holes in the diamond anvils (**Figure 9**) can reduce the amount of diamond in the beam to 0.3 mm. In these DACs, X-ray absorption fine structure (XAFS) spectra down to 4 keV permit analyses of the K edges of elements as light as Sc as well as L edges of rare-earth elements. Near-edge X-ray absorption fine structure (XANES) provides information on the symmetry-projected conduction band density of states (DOS), while the extended X-ray absorption fine structure (EXAFS) provides local structure information. These element-specific probes are able to reveal electronic and structural changes that take place at HPT conditions. Intensity loss due to Bragg reflections as an X-ray beam passes through the diamond anvils causes unwanted sharp peaks. However, these

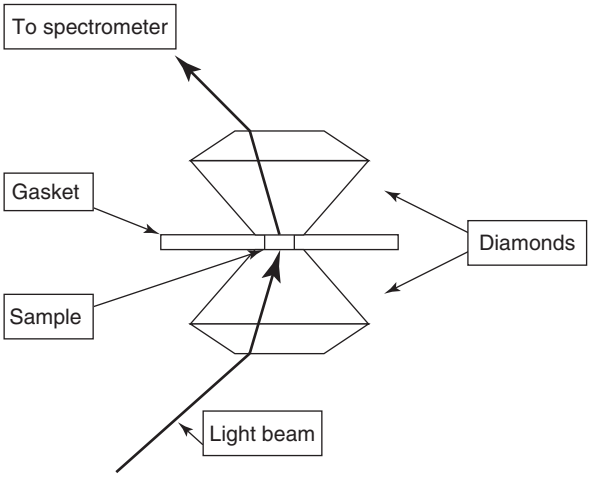


Figure 8 Brillouin scattering can be used to determine phonon velocities in high-pressure samples without making corrections for refractive indices if the anvil faces bisect the 90° angle between the incident and scattered light rays.

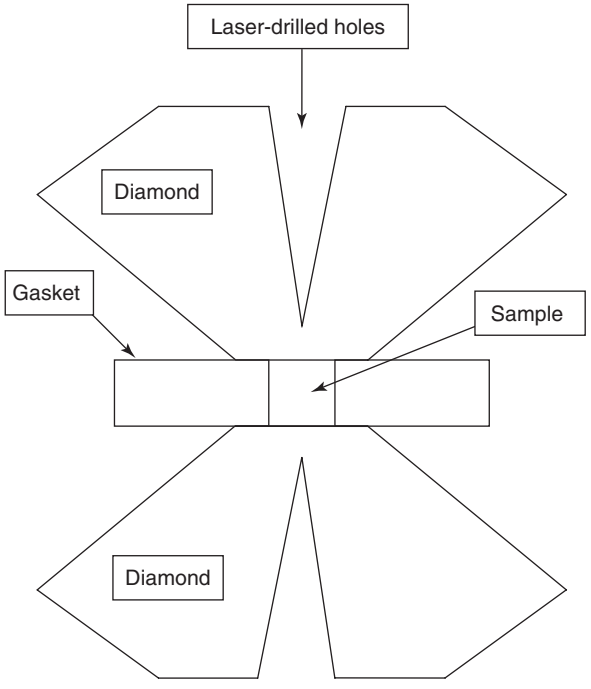


Figure 9 Holes laser-drilled to within 0.15 mm of the anvil face greatly reduce the absorption by diamond allowing X-ray absorption studies to be made at energies down to <4 keV.

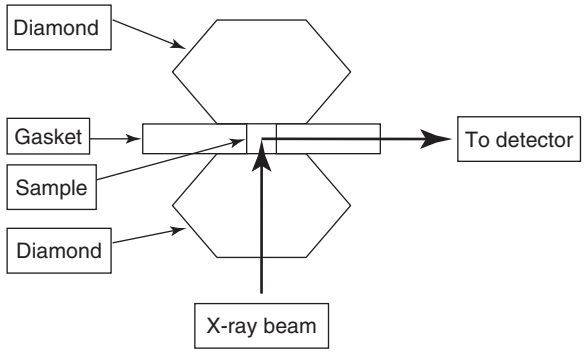


Figure 10 X-ray emission excited by an X-ray beam passing through one of the diamond anvils can pass out through a transparent gasket and be analyzed. Background due to Rayleigh and Compton scattering is minimized when the detector is located at 90° from the X-ray beam in the plane of polarization of an X-ray beam from a synchrotron source.

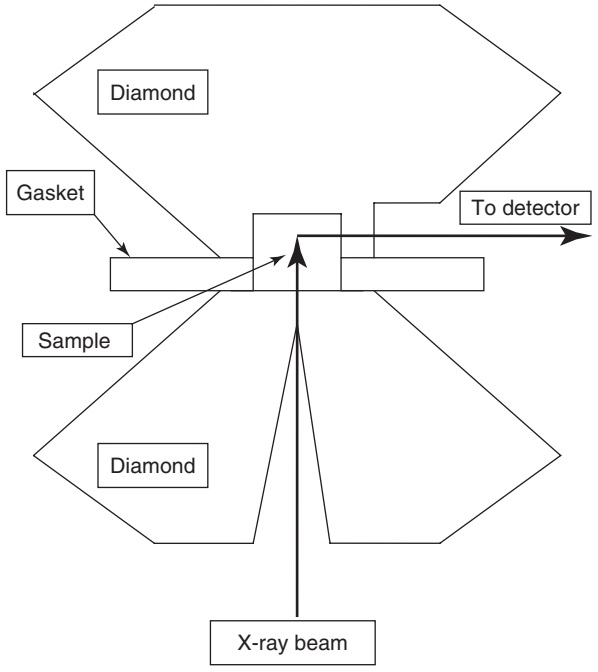


Figure 11 X-ray emission excited by an X-ray beam passing through one of the diamond anvils can pass out through the side of the other diamond anvil to be analyzed. Using a diamond with a laser-drilled hole for one anvil and a diamond with a thin-walled recess for the other greatly reduces absorption to permit high-pressure–temperature studies down to energies of 4 keV and less.

can be shifted by changing the orientation of the DAC slightly. An alternative and very successful way of obtaining absorption spectra is by measuring the intensity of fluorescence emission at the energy of one of the emission peaks while the energy of the incident beam is scanned over the absorption edge being studied. X-rays emerging through an X-ray transparent gasket (**Figure 10**) or side of one of the anvils (**Figure 11**) are collected and analyzed. The signal-to-noise ratio is much improved with this method.

Nuclear resonance X-ray spectroscopy (NRXS), also known as Mössbauer spectroscopy, deals with exceedingly narrow absorption peaks resulting from nuclear transitions. NRXS uses an in-line high-resolution monochromator (HRM) on an SR beam capable of narrowing the photon energy and fine-tuning the monochromatic X-ray beam with meV resolution. Avalanche photodiodes (APD) are used to collect signals from nuclear resonance absorption, and reject all other signals.

Inelastic X-ray spectroscopy (IXS) represents a major breakthrough for high-pressure research. As in Raman spectroscopy, the energies of the events being probed appear as differences in the energies of incident and scattered X-rays. For this reason, IXS is sometimes referred to as X-ray Raman spectroscopy. This principle makes it possible for experimenters to select X-ray energies that are high enough to easily penetrate diamond anvils and/or low-Z gaskets while probing low-energy phenomena. Recent advances in focusing of X-rays onto samples as small as 0.01 mm and dramatic improvements in monochromators and analyzers to achieve sub-eV resolution have opened the way for these techniques to be applied to investigations of a wide variety of low-energy events at HPT conditions. IXS comprises several techniques which yield different types of information. They are X-ray inelastic near-edge spectroscopy (XINES), electronic inelastic X-ray scattering (EIXS), resonant inelastic X-ray scattering (RIXS), nuclear resonant inelastic X-ray scattering (NRIXS), and phonon inelastic X-ray scattering (PIXS).

X-ray inelastic near-edge spectroscopy (XINES) is capable of making measurements that yield the same information as XANES on near core-electron absorption edge features. The principle is similar to that of electron energy loss spectroscopy (EELS). These techniques provide information about the nature of chemical bonding. Spectral features are particularly pronounced and the information particularly important for the very light elements. It is only since the development of XINES that such measurements can be made at HPT conditions because the low-energy X-ray and electron beams needed for XANES and EELS are completely blocked by any of the pressure vessels. Successful observations have been made on second-row elements from Li (56 eV) to O (543 eV) at high pressures. This opens a wide new field of near K-edge spectroscopy of the very light elements.

Electronic inelastic X-ray scattering (EIXS), which is very similar to XINES in principle and instrumentation, can be used to investigate the drastic effect that pressure can have on the energy and dispersion of electronic bands. The IXS principle makes it

possible to use a 10 keV X-ray beam for excitation and a monochromator capable of 0.3–1 eV resolution to probe high-energy electronic phenomena, including electronic band structure, Fermi surface, excitons, plasmons, and their dispersions at HP.

In resonant inelastic X-ray spectroscopy (RIXS) experiments, both incident and scattered X-ray energies need to be scanned. Of particular interest is the investigation of phase transitions driven by electron correlation effects that occur in transition elements and rare-earth elements and their compounds at high pressures.

Nuclear resonant inelastic X-ray scattering (NRIXS) utilizes inelastic scattering to collect nuclear resonance or Mössbauer spectra. An APD directly downstream of the sample collects nuclear resonant forward scattering (NRFS) spectra which provide a precise determination of the hyperfine interaction parameter and Lamb–Mössbauer factor. APDs surrounding the sample collect NRIXS spectra. Isotopes that exhibit a nuclear resonant Mössbauer effect, for example ^{57}Fe, offer an extraordinary opportunity to make phonon DOS measurements on samples containing them, measurements that provide valuable information on dynamic, thermodynamic, and elastic properties.

PIXS requires detecting a very weak sample signal in the presence of a very strong background signal, especially when a HP vessel contains the sample. New novel HRMs which are being developed will improve the feasibility of making nonresonant PIXS measurements at high pressure. Such measurements would be a valuable source of information on phonon dynamics in highly compressed materials, information important for understanding vibrational thermodynamic properties, elasticity, and phase transition mechanisms.

Electrical resistance has been measured on samples as a function of pressure and at high temperature in DACs. For these experiments, fine electric leads are connected to the sample and four-lead systems have been employed for the most quantitative measurements. A particularly novel approach, dubbed designer diamonds, has employed diamonds in which the leads are imbedded in the anvils by vapor depositing single-crystal diamond over the leads.

Magnetic measurements have been made in DACs constructed of nonmagnetic materials such as beryllium–copper. Coils for generating and picking up signals are placed as close to the sample as possible.

Sound velocity measurements have been made by three important methods, Brillouin scattering, impulse stimulated scattering, and interferometry of GHz acoustic signals. The first two of these use interactions between photons and phonons while the third generates GHz sound by a transducer and uses interferometry of the acoustic signals to measure travel time in a crystal only $\sim$0.1 mm across. Sound velocities are then calculated.

Gamma rays emitted by a radioactive sample in a DAC have been counted and the effect of pressure on the decay rate measured. Pressure has been shown to increase the decay rate of ^{7}Be in an electron-capture process.

Conclusions

Diamond anvil cells combined with many analytical techniques are a valuable source of information on fundamental properties of solids including crystal structures, equations of state, and nature of bonding. Encapsulation of fluids in gaskets has made it possible to learn about structures and properties of fluids, hydrothermal studies, properties of melts, interactions between fluids and solids, and other phase relationships. DACs have contributed to the search for superhard materials. Our knowledge of the planets has benefited from information about the physical properties of materials that form them. In the biological sciences, efforts are being directed to biological questions. These include organic chemistry, macromolecular studies, and microorganisms under pressure. Some of the methods described above are only now beginning to be applied to important questions. Future research with DACs is certain to yield important new results in all of the sciences.

See also: Core Photoemission; Crystal Structure; Crystal Structure Determination; Laser Radiation Sources; Luminescence; Optical Microscopy; Phase Transformation; Phases and Phase Equilibrium; Scattering, Inelastic: X-Ray (Methods and Applications); Scattering, Nuclear Resonant; Scattering, Resonant; Synchrotron Radiation; X-Ray Absorption Spectroscopy.

PACS: 07.35.+k

Further Reading

Bridgman PW (1963) General outlook on the field of high pressure research. In: Paul W and Warschauer DM (eds.) *Solids under Pressure*, vols. 1–13. New York: McGraw-Hill.

Paul W and Warschauer DM (eds.) (1963) *Solids under Pressure*, vols. 1–13. New York: McGraw-Hill.

Piermarini GJ and Block S (1975) Ultrahigh pressure diamond-anvil cells and several semiconductor phase transitions pressure in relation to the fixed point pressure scale. *Review of Scientific Instruments* 46: 973–979.

Van Valkenburg A (1962) Visual observation of high pressure transitions. *Review of Scientific Instruments* 33: 1462.

Weinstein BA and Piermarini GJ (1975) Raman scattering and phonons dispersion in Si and GaP at very high pressure. *Physical Review B* 12: 1172–1186.

Weir CE, Lippincott ER, Van Valkenburg A, and Buting EN (1959) Infrared studies in the 1–15 micron region to 30,000 atmospheres. *Journal of Research of the National Bureau of Standards A* 63: 55–62.

Dielectric Function

J Petzelt and I Rychetský, Institute of Physics of Academy of Sciences of the Czech Republic, Prague, Czech Republic

Definition and General Properties

Electromagnetic field in a condensed matter medium is described by Maxwell equations containing charge densities $\rho(\boldsymbol{r},t)$ and currents of electrons and ions in the medium. Dependence of charge density on the macroscopic electric field (averaged over local field variations due to the discreet atomic structure) constitutes a material relation characteristic for the material. The charge density can be expressed by means of the dipole moment vector per unit volume (polarization) $\boldsymbol{P}$ or by the electrical displacement vector $\boldsymbol{D}$:

$$\rho = -\mathrm{div}\,\boldsymbol{P}, \quad \boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P} = \varepsilon_0\chi\boldsymbol{E}$$

where ε_0 is the permittivity of vacuum and χ is the linear dielectric susceptibility tensor. Applying generally spatial and time-dependent electric field $\boldsymbol{E}(\boldsymbol{r},t)$, the polarization vector in the linear approximation is

$$P_i(\boldsymbol{r},t) = \varepsilon_0 \int_{-\infty}^{t} \mathrm{d}t' \int \mathrm{d}\boldsymbol{r}'\chi_{ij}(\boldsymbol{r}-\boldsymbol{r}', t-t')E_j(\boldsymbol{r}', t')$$

and may depend on the spatial surrounding (nonlocal response) and history of $\boldsymbol{E}$ (for causality reasons only on previous times). Assuming space and time homogeneity, the dependences are only on the differences $\boldsymbol{R}=\boldsymbol{r}-\boldsymbol{r}'$ and $\tau = t - t'$. Defining the time and spatial Fourier transform of the field $E_i(\omega,\boldsymbol{k}) = \int_{-\infty}^{\infty} \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\, E_i(\boldsymbol{r},t)\, \mathrm{e}^{\mathrm{i}(\omega t - \boldsymbol{k}\boldsymbol{r})}$, the Fourier components (i.e., plane waves) of polarization and electric displacement are $P_i(\omega,\boldsymbol{k}) = \varepsilon_0\chi_{ij}(\omega,\boldsymbol{k})E_j(\omega,\boldsymbol{k})$ and $D_i(\omega,\boldsymbol{k}) = \varepsilon_0\varepsilon_{ij}(\omega,\boldsymbol{k})E_j(\omega,\boldsymbol{k})$, where the linear dielectric function (relative dielectric permittivity or dielectric constant) $\varepsilon_{ij}(\omega,\boldsymbol{k}) = \delta_{ij} + \chi_{ij}(\omega,\boldsymbol{k})$ and susceptibility $\chi_{ij}(\omega,\boldsymbol{k})$ are complex functions of the real $\boldsymbol{k}$-vector and frequency:

$$\varepsilon_{ij}(\omega,\boldsymbol{k}) = \delta_{ij} + \int_0^{\infty} \mathrm{d}\tau \int \mathrm{d}\boldsymbol{R}\ \chi_{ij}(\boldsymbol{R},\tau)\mathrm{e}^{\mathrm{i}(\omega\tau - \boldsymbol{k}\boldsymbol{R})}$$

The dielectric function describes the linear relation between $\boldsymbol{D}$ and the macroscopic electric field $\boldsymbol{E}$.

The frequency and $\boldsymbol{k}$-vector can be mathematically considered as complex quantities. Since the fields are real, the dielectric function fulfills the relation $\varepsilon_{ij}(\boldsymbol{k},\omega) = \varepsilon_{ij}(-\boldsymbol{k}^*, -\omega^*)^*$ or, for real ω and $\boldsymbol{k}$, $\varepsilon_{ij}(\boldsymbol{k},\omega) = \varepsilon_{ij}(-\boldsymbol{k}, -\omega)^*$. In nongyrotropic (i.e., without optical activity) media where $\varepsilon_{ij}(\omega,\boldsymbol{k}) = \varepsilon_{ij}(\omega,-\boldsymbol{k})$, the dielectric function tensor is symmetric, $\varepsilon_{ij}(\omega,\boldsymbol{k}) = \varepsilon_{ji}(\omega,\boldsymbol{k})$. If the wave vector obeys the constraint $ak = \pi a/\lambda \ll 1$, where a is of the order of the characteristic system inhomogeneities (in crystals the interatomic distances), the dielectric function can be expanded in ak:

$$\varepsilon_{ij}(\omega,k) = \varepsilon_{ij}(\omega) + O((ak)^2) \approx \varepsilon_{ij}(\omega)$$

Neglecting the second term on the right-hand side, it is assumed that the response is local so that the spatial dispersion effects are neglected. Then only the frequency dependence has to be considered:

$$\varepsilon_{ij}(\omega) = \delta_{ij} + \int_0^{\infty} \mathrm{d}\tau\ \chi_{ij}(\tau)\mathrm{e}^{\mathrm{i}\omega\tau}$$

In most cases of practical interest, for macroscopically homogeneous media, this approximation is well justified and further discussion is limited to this case.

The complex dielectric function has the real and imaginary part, $\varepsilon_{ij}(\omega) = \varepsilon'_{ij}(\omega) + \mathrm{i}\varepsilon''_{ij}(\omega)$, which are even and odd functions of real ω, respectively:

$$\varepsilon'_{ij}(\omega) = \varepsilon'_{ij}(-\omega), \quad \varepsilon''_{ij}(\omega) = -\varepsilon''_{ij}(-\omega)$$

In the case of capacitors, the real part describes its capacity and the imaginary part determines the rate of losses of the radiation energy, which in the isotropic case is

$$Q = \tfrac{1}{2}\varepsilon_0|E(\omega)|^2\omega\varepsilon''(\omega)$$

In the case of electromagnetic waves, the real part describes the change of the wavelength compared to that in vacuum ($\lambda \propto 1/\sqrt{\varepsilon'}$) and the imaginary part its absorption (attenuation).

Instead of dielectric function, the AC electric conductivity tensor $\varepsilon_{ij}(\omega) = \mathrm{i}\sigma_{ij}(\omega)/\varepsilon_0\omega$ can be equivalently used. Along the principal axes of the dielectric tensor in nonmagnetic media where the magnetic permeability $\mu = 1$, the complex index of refraction $\hat{n}(\omega) = \sqrt{\varepsilon(\omega)} = n(\omega) + \mathrm{i}\kappa(\omega)$ is also frequently used to describe the dielectric response at optical frequencies. Here n is the real refractive index and κ index of absorption (characterizing the wave attenuation factor within one wavelength).

Kramers–Kronig Relations

The causality principle provides analyticity of $\varepsilon_{ij}(\omega)$ in the upper complex half-plane (in the lower half-plane $\varepsilon_{ij}(\omega)$ has singularities, mostly simple poles)

and results in Kramers–Kronig dispersion relations:

$$\varepsilon'_{ij}(\omega) - \delta_{ij} = \frac{2}{\pi} \boldsymbol{P} \int_0^{+\infty} \frac{x\varepsilon''_{ij}(x)}{x^2 - \omega^2}\, \mathrm{d}x$$

$$\varepsilon''_{ij}(\omega) - \frac{\sigma_{ij}(0)}{\varepsilon_0 \omega} = -\frac{2\omega}{\pi} \boldsymbol{P} \int_0^{+\infty} \frac{\varepsilon'_{ij}(x) - \delta_{ij}}{x^2 - \omega^2}\, \mathrm{d}x$$

Checking the validity of Kramers–Kronig relations is an important issue for experimentalists measuring independently the real and imaginary part of the dielectric response, even in a somewhat limited frequency range.

Sum Rules

The finite charges in the medium of interest result in the so-called sum rule for oscillator strengths (*f*-sum rule):

$$\int_0^{\infty} \omega\varepsilon''_{ij}(\omega)\, \mathrm{d}\omega = \delta_{ij} \frac{\pi}{2} \omega_{\mathrm{p}}^2, \quad \omega_{\mathrm{p}}^2 = \frac{Ne^2}{m\varepsilon_0}$$

$$\int_0^{\infty} \varepsilon'_{ij}(\omega)\, \mathrm{d}\omega = \delta_{ij} - \frac{2\pi^2}{\varepsilon_0} \sigma_{ij}(0)$$

where N is the density of charges, e and m their mass (assuming for simplicity only one type of charges of uniform density). These sum rules are of practical importance when discussing temperature or pressure dependences of the dielectric function, which undergoes pronounced changes (e.g., metal-insulator and superconductive phase transitions), since the total charge in the system is conserved.

Fluctuation–Dissipation Theorem

Dielectric function is closely related to thermal fluctuations of the polarization characterized by the autocorrelation function $\langle \delta P_i(t) \delta P_j(0) \rangle$, the Fourier component of which

$$S_{ij}(\omega) = \int_{-\infty}^{+\infty} \langle \delta P_i(t) \delta P_j(0) \rangle\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t$$

is called spectral density function. This function is proportional to dynamic structure factor measured in inelastic neutron scattering experiments and to inelastic light scattering (Raman and Brillouin) cross section. The fluctuation–dissipation theorem relates this function to dielectric losses:

$$\varepsilon_0 \varepsilon''_{ij}(\omega) = \frac{1}{2\hbar}\left(1 - \mathrm{e}^{-\hbar\omega/kT}\right) S_{ij}(\omega) \approx \frac{\omega}{2kT} S_{ij}(\omega)$$

where the last approximation is valid in the classical high-temperature limit $kT \gg \hbar\omega$.

Polaritons

Dielectric function $\varepsilon(\omega, \boldsymbol{k})$ describes the dielectric response to the plane-wave electric field $E(\omega, \boldsymbol{k})\mathrm{e}^{-\mathrm{i}(\omega t - \boldsymbol{kr})}$, which can be, in principle, realized for arbitrary values of ω, $\boldsymbol{k}$ by using appropriate external electric charges and currents. When one limits to propagating electromagnetic waves in macroscopically homogeneous media, the values ω, $\boldsymbol{k}$ become mutually dependent. For example, for transverse waves $\boldsymbol{E} \perp \boldsymbol{k}$ which propagate along the optical axes, $k(\omega)^2 = (\omega^2/c^2)\varepsilon(\omega)$. Such waves, whose propagation is characterized by a complex wave vector, characterize mixed electromagnetic and lattice vibrational (phonon) or electron excitations called polaritons. Knowing the complex dielectric function, polariton dispersion branches are fully determined.

Physical Mechanisms and Models of Dielectric Dispersion

Dielectric function provides direct and important information about the polarization mechanisms in the medium. In all systems, the dominant dispersion and absorption mechanism in the high-frequency optic range (visible and/or ultraviolet) is caused by electronic excitations (transitions from populated to empty electron levels). Since these transitions involve the whole Brillouin zone, there are no simple phenomenological models used for the description of this dispersion. Instead, microscopic calculations based on the known or calculated electron band structure are normally used and compared with the experiment.

In the infrared range, the dominating dispersion and absorption mechanism in dielectrics is lattice vibrations (one-phonon absorption). Only polar phonons connected with the dipole moment fluctuations in the $\boldsymbol{E}$ direction, that is, transverse phonons whose eigenvectors have polar vector symmetry are infrared active. Their dielectric response can be well described by that of damped harmonic oscillators. Assuming s infrared active (i.e., polar) modes, the dielectric function in approximation which neglects the mutual coupling among the oscillators is a simple sum of individual oscillator responses:

$$\varepsilon(\omega) = \varepsilon(\infty) + \sum_{j=1}^{s} \frac{\omega_{pj}^2}{\omega_{TOj}^2 - \omega^2 - \mathrm{i}\gamma_{TOj}\omega}$$

where $\varepsilon(\infty)$, ω_{pj}^2, ω_{TOj}, γ_{TOj} are the high-frequency (optic) permittivity due to electronic transitions, oscillator strength, square root of the force constant, and damping of the jth transverse optical phonon mode, respectively. A more general function frequently used particularly for fitting the experimental

infrared reflectivity is the so-called factorized form of the dielectric function:

$$\varepsilon(\omega) = \varepsilon(\infty) \prod_{j=1}^{s} \frac{\omega_{\mathrm{LO}j}^2 - \omega^2 - \mathrm{i}\gamma_{\mathrm{LO}j}\omega}{\omega_{\mathrm{TO}j}^2 - \omega^2 - \mathrm{i}\gamma_{\mathrm{TO}j}\omega}$$

where the frequencies and dampings of longitudinal optical phonons are introduced explicitly. All the parameters are real positive numbers. For finite damping, the transverse and longitudinal eigenfrequencies correspond to complex poles and zeros of the dielectric function, respectively. Only in the limit of small dampings, these eigenfrequencies approach the corresponding ω_{TO} and ω_{LO} frequencies. In the static limit $\omega = 0$, the factorized dielectric function yields the generalized Lyddane–Sachs–Teller relation. It shows that high values of the static permittivity can be reached particularly by low ω_{TO} frequencies. This is the case of displacive ferroelectrics, where the Curie–Weiss divergence $\varepsilon = C/(T - T_{\mathrm{C}})$ is achieved by softening of the soft phonon mode $\omega_{\mathrm{TO}}^2(\mathrm{soft}) \propto T - T_{\mathrm{C}}$ (Cochran law).

In highly conducting materials (metals), the lattice response is screened by free charges (electrons) and their absorption dominates in the whole infrared range. It can be usually well described by the simple Drude response:

$$\varepsilon(\omega) = \varepsilon(\infty) - \frac{\omega_{\mathrm{p}}^2}{\omega(\omega + \mathrm{i}\gamma)}$$

It is the response of a damped harmonic oscillator for zero force constant. The zero of this dielectric function is called screened plasma frequency (plasmon eigenfrequency) and characterizes longitudinal oscillations of the free carrier plasma density. Notice that unlike the case of oscillators, contribution of the Drude term to the static permittivity is negative.

Below the lattice absorption range, no polarization (absorption) mechanism exists in an ideal nonpiezoelectric dielectric crystal. In piezoelectric crystals the coupling between polar optic and acoustic waves activates the acoustic waves in a finite sample, which introduces additional dielectric resonances, typically in the MHz frequency range, whose frequencies are, however, dependent on sample size and shape.

However, structural disorder and strong lattice anharmonicities or all kinds of movable charged defects usually induce further dispersion in the microwave and lower frequency range. Such dispersion is as a rule of relaxation type, the simplest model being the Debye relaxation:

$$\varepsilon(\omega) - \varepsilon(\infty) = \frac{\omega_{\mathrm{p}}^2}{\omega_0^2 - \mathrm{i}\gamma\omega} = \frac{\Delta\varepsilon}{1 - \mathrm{i}\tau\omega}$$

where $\tau = \gamma/\omega_0^2$ and $\Delta\varepsilon = \omega_{\mathrm{p}}^2/\omega_0^2$ are relaxation time and dielectric strength, respectively. It is seen that in the low-frequency range $\omega \leqslant 1/\tau$, this function approaches that of the oscillator with frequency ω_0 and damping γ. If the oscillator is overdamped (pure imaginary eigenfrequency for $\gamma \geqslant 2\omega_0$), the dielectric losses are peaked approximately at ω_0^2/γ like the Debye relaxation. This frequency characterizes the dynamics of the defects or hopping frequency of the localized charges. In more complex cases, distribution of Debye relaxations have to be used to fit the broadband dielectric spectra. In these cases, frequently appearing, for example, in polymers, much effort was exerted to study the universal dispersion laws which rule the tails of the broad loss peaks. However, till now no generally accepted picture has been achieved.

The simplest model which yields the Debye dispersion in solids is hopping of charges over a potential barrier E_{a} much higher than the thermal energy kT. If the charges do not interact, the temperature dependence is described by the Arrhenius law $\tau = \tau_\infty \exp(E_{\mathrm{a}}/kT)$. If the charges interact, the effective barrier for hopping increases on cooling and the temperature law is usually modified either to Vogel–Fulcher law $\tau = \tau_\infty \exp(E_{\mathrm{a}}/k(T - T_{\mathrm{g}}))$, or to critical slowing-down $\tau \propto 1/(T - T_{\mathrm{C}})$. Here, T_{g} is the glass transition temperature and T_{C} the critical temperature (connected, e.g., with an order–disorder ferroelectric transition) where the relaxation time tends to infinity. In the latter case, the systems order below T_{C} and the relaxation vanishes. In the former case, the long-range order is prohibited by some frustration due to structural disorder so that the disorder remains frozen down to low temperatures (glass phase). Near and below T_{g}, the Vogel–Fulcher law gradually changes into the Arrhenius law and the relaxation broadens and weakens on cooling.

Dielectric Constant of Various Dielectrics

The static permittivity (dielectric constant) of homogeneous solids may acquire values reaching from ~ 2 up to the order of 10^4. Materials from both ends of this broad spectrum are required for contemporary microelectronic applications: low-permittivity materials such as organosilicate films for advanced on-chip interconnects, and high-permittivity materials such as capacitors, tunable ferroelectric films or low-loss ceramics for microwave passive elements in communication systems. As discussed above, the permittivity value is determined by several always-positive contributions in the dielectric spectra. In low-permittivity materials, both electronic and vibrational contributions are very weak, which is typical for weak bonding among nonpolar molecules in organic substances. In high-permittivity materials, mostly the vibrational contribution of strong and

low-frequency polar modes dominates, since the electronic contribution ε_∞ is always relatively small due to high frequencies of the electronic transitions. Also the order–disorder ferroelectrics may reach high permittivity values close to T_C, but these are usually of no practical use because of strong temperature dependences of ε' and high losses. The most perspective high-permittivity materials are the so-called relaxor ferroelectrics (special type of structurally disordered ferroelectrics), where only a short-range polar ordering is reached and the dominant dielectric dispersion is due to the high-frequency dynamics of the polar nanoclusters.

Effective Dielectric Function

In the case of macroscopically or mesoscopically inhomogeneous media (e.g., composites, ceramics, and films with granular structure), the standard experimental techniques are unable to determine the bulk dielectric functions of individual constituents. Instead, the so-called effective dielectric function can be determined. If the probing electric field is homogeneous on the scale of inhomogeneities and the boundaries between different homogeneous regions are sharp, effective medium approximation can be used for calculation of the effective dielectric function. For example, for a two-component composite of spherical particles with dielectric function $\varepsilon_1(\omega)$, $\varepsilon_2(\omega)$ and volume fraction f, $1-f$, respectively, the effective dielectric function $\varepsilon_{\text{eff}}(\omega)$ is given by

$$f\frac{\varepsilon_1-\varepsilon_{\text{eff}}}{\varepsilon_1+2\varepsilon_{\text{eff}}}+(1-f)\frac{\varepsilon_2-\varepsilon_{\text{eff}}}{\varepsilon_2+2\varepsilon_{\text{eff}}}=0$$

This expression holds as long as the probing AC field is homogeneous within individual particles, that is, up to the infrared range if the particle size is $\sim 1\,\mu$m. If the difference between permittivities of both components is pronounced, particularly strong changes in $\varepsilon_{\text{eff}}(\omega)$ (including effective polar mode frequencies) appear near the percolation thresholds of both components, which in the latter model appear for $f = 1/3$ and $2/3$. In ceramics and films, the percolation threshold of the bulk grain properties cannot be reached (each grain is separated from its neighbors by the grain boundary with different dielectric properties) and therefore, the effective dielectric spectra are often strongly influenced by grain boundaries. For example, combining semiconducting grains with insulating boundaries can lead to effective giant permittivities (values up to $\sim 10^6$ were actually observed) with strong dielectric relaxations at low or medium frequencies, which are absent in the response of both bulk components.

See also: Composites: Overview; Electrodynamics: Continuous Media; Electron–Phonon Interactions and the Response of Polarons; Ferroelectricity; Insulators, Optical Properties of; Lattice Dynamics: Vibrational Modes; Local Field Effects; Optical Absorption and Reflectance; Optical Sum Rules and Kramers–Kronig Relations; Polaritons; Semiconductors, Optical and infrared Properties of.

PACS: 63.20.Dj; 71.36.+c; 77.22.−d; 77.22.Ch; 77.22.Gm; 77.80.Bh; 78.20.Ci; 78.30.Gq

Further Reading

Agranovich VM and Ginzburg VL (1984) *Crystal Optics with Spatial Dispersion and Excitons*, 2nd edn. Berlin: Springer.
Barker AS Jr. (1975) *Physical Review B* 12: 4071.
Bergman DJ and Stroud D (1992) In: Ehrenreich and Turnbull (eds.) *Solid State Physics*. Boston: Academic Press.
Gervais F (1983) In: Button (ed.) *Infrared and Millimeter Waves*, vol. 8, 279p. New York: Academic Press.
Jonscher AK (1996) *Universal Relaxation Law*. London: Chelsea Dielectrics Press.
Klingshirn CF (1995) *Semiconductor Optics*. Berlin: Springer.
Landau LD and Lifshitz EM (1960) *Electrodynamics of Continuous Media*. Oxford: Pergamon.
Waser R (ed.) (2003) *Nanoelectronics and Information Technology*. Weinheim: Wiley.

Diffraction *See* Scattering, Elastic (General).

Diffusionless Transformations

H Ledbetter, A Saxena, and A Migliori, Los Alamos National Laboratory, Los Alamos, NM, USA

Introduction

Several topics on diffusionless phase transformations are briefly described. All solid-state phase transformations fall into two groups: diffusional and nondiffusional. The first involves thermally activated noncorrelated atomic jumps from one lattice position to another. The second (focused here) possesses several descriptors: diffusionless, displacive, nonreconstructive, coordinated-shear, military, and martensitic. Because martensite represents the prototypical diffusionless phase transformation, martensitic

and diffusionless are used nearly synonymously. Diffusionless transformations occur in nearly every material type: metals, alloys, ceramics, minerals, rare-gas solids, polymers, and (perhaps) biological systems. Although diffusionless, these transformations can occur at high temperatures.

The following topics are considered: energy relationships; vibration modes, phonons, and entropy; Landau theory, including microstructure and soft modes; elastic-inclusion theory. All these topics relate to the thermodynamic free energy. Free-energy differences among phases provide the driving force for a transformation from one phase to another.

Energy Relationships

For a first-order phase transformation, **Figure 1** summarizes the essential thermodynamics. The usual three energy functions relate by

$$F = E - TS \qquad [1]$$

where F denotes the Helmholtz free energy, E internal energy, T temperature, S entropy. A phase transformation occurs when $\Delta F = F^{(2)} - F^{(1)} = 0$, or $\Delta E = T\Delta S$. Often, ΔE is approximately temperature independent. (In a Debye model, ΔE is exactly temperature independent if $\Theta_D^{(2)} = \Theta_D^{(1)}$, where Θ_D denotes the Debye temperature.) Thus, a phase transformation occurs when the $T\Delta S$ term equals the ΔE term. Various *ab initio* methods exist for calculating ΔE. Here, *ab initio* means that no solid-state properties enter the calculation, only atomic properties do. Such calculations yield the basic solid-state cohesive properties: atomic volume, binding energy, compressibility.

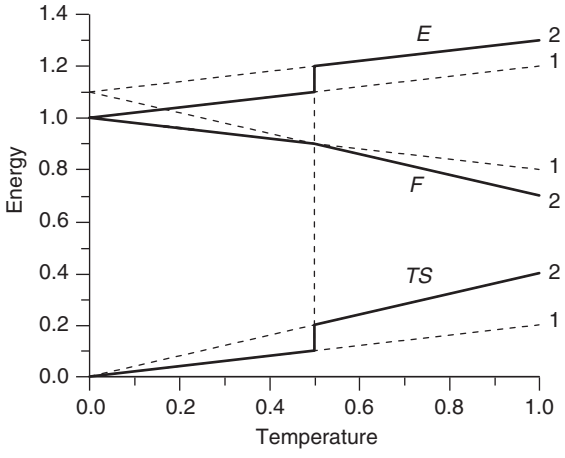


Figure 1 For a first-order phase transformation, schematic temperature dependence of three thermodynamic functions: internal energy E, Helmholtz free energy F, and entropy S. Numerals 1 and 2 denote low-temperature and high-temperature phases.

The Planck–Einstein oscillator model is the simplest model that captures most of the essential features of the thermodynamic functions. (At low temperatures, this model predicts a wrong specific-heat/temperature relationship.) This harmonic-oscillator model yields the following relationships:

$$(E - E_0)/3RT = x/(e^x - 1) \qquad [2]$$

$$S/3R = x/(e^x - 1) - \ln(1 - e^{-x}) \qquad [3]$$

$$(F - E_0)3RT = -\ln(1 - e^{-x}) \qquad [4]$$

$$C_V/3R = x^2 e^x/(e^x - 1)^2 \qquad [5]$$

where R denotes the universal gas constant, emphasizing the relationship between crystalline-solid and perfect-gas-model thermodynamics. The variable x denotes Θ_E/T, Θ_E being the Einstein temperature, where $\Theta_E = (3/4)\Theta_D$. Subscript zero denotes zero-temperature values. **Figure 2** shows the temperature dependence of these four functions. Both the $S(T)$ curve and the $F(T)$ curve show a strong correspondence to the bulk modulus (reciprocal compressibility) temperature dependence. (Cautionary note: Based on a harmonic model, these E–S–F relationships apply

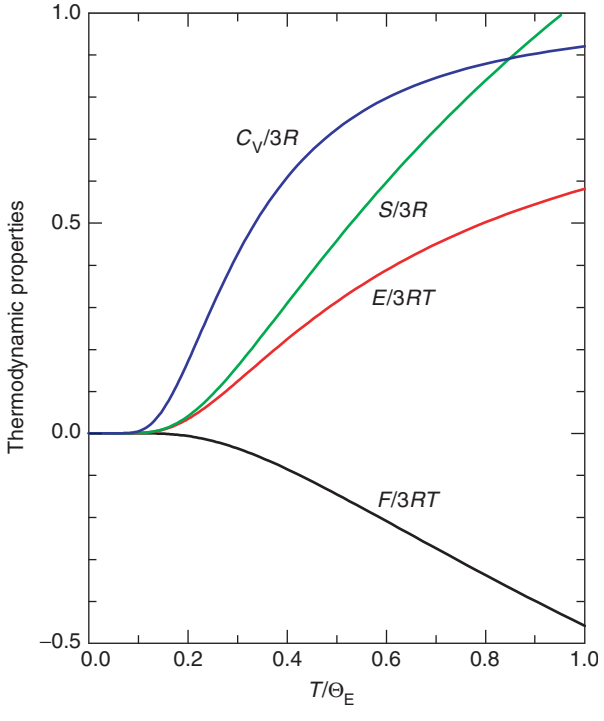


Figure 2 Based on a Planck–Einstein oscillator model, temperature dependence of four thermodynamic functions: specific heat C_V, entropy S, internal energy $E - E_0$, and Helmholtz free energy $F - E_0$. Note that while C_V and S represent absolute values, E and F represent values relative to zero temperature.

at temperatures below the melting temperature, near which strong anharmonic lattice vibrations may occur.) Among these three functions, only entropy vanishes at zero temperature because of the third law of thermodynamics. Some physical properties also vanish at zero temperature, for example, constant-volume specific heat

$$C_V(\approx C_P) = \int_0^T (S/T)\mathrm{d}T = -T(\partial^2 F/\partial T^2)_V \quad [6]$$

thermal expansivity

$$\beta = (1/V)(\mathrm{d}V/\mathrm{d}T)_P = -(1/B_T)(\partial^2 F/\partial T \partial V)_{V,T} \quad [7]$$

bulk-modulus (reciprocal compressibility) temperature derivative

$$(\partial B/\partial T) = \partial/\partial T(\partial^2 F/\partial V^2)_P \quad [8]$$

where B_T denotes the isothermal bulk modulus

$$B_T = (1/V)(\partial^2 F/\partial V^2)_T \quad [9]$$

These properties – B, β, C_V – provide useful probes of the phase transformation. (Caution: changes in these properties at first-order phase transformations differ dramatically from their changes at second-order phase transformations.) Interestingly, all three of these properties interrelate simply with the Gruneisen parameter: $\gamma = B_T\beta V/C_V = B_S\beta V/C_P$.

At a first-order phase transformation, discontinuities occur in the free-energy first derivatives, thus changes in entropy and volume:

$$S = -(\partial F/\partial T)_V \quad [10]$$

$$V = (\partial G/\partial P)_T \quad [11]$$

where $G = F + PV$ denotes the Gibbs free energy.

At low temperatures, tens of degrees, instead of the simple Planck–Einstein model, an improved model is needed, such as Debye's, which yields particularly simple low-temperature expressions for the thermodynamic functions:

$$S = AT^3 \quad [12]$$

$$E = E_0 + (3/4)ST^3 \quad [13]$$

$$F = E_0 - (1/4)ST^3 \quad [14]$$

where $A = 4\pi^4 k_B/5\Theta_D^3$. In simple harmonic models the characteristic temperature Θ_D determines all the thermodynamic functions together with their pressure, temperature, and volume derivatives.

The phase-transformation latent heat is

$$\Delta H = \Delta E + P\Delta V = T\Delta S \quad [15]$$

where $H = E + PV$ denotes enthalpy. Usually, the $P\Delta V$ term is small compared with ΔE. Latent heat is positive, that is, the system absorbs thermal energy during warming through a first-order phase transformation.

Vibration Modes, Phonons, Entropy

The fundamental electronic structure of a homogeneous solid determines all its properties. However, it is not currently possible to perform an *ab initio* electronic-structure calculation at finite temperature with an accuracy adequate to explain structural phase transitions. To simplify the general problem, it is useful to separate the Helmholtz free energy into four parts:

$$F = F_L + F_H + F_E + F_A \quad [16]$$

where F_L denotes the static lattice energy for electrons and ions in position on the lattice and in the ground state at zero temperature, F_H the harmonic phonon free energy, F_E the free energy of thermally excited electronic states, and F_A a lumped term including such anharmonic effects as phonon–phonon and electron–phonon interactions. At higher temperatures, F_H dominates. In metals, at lower temperatures, F_E dominates.

Consider a string of masses (atoms), a distance a apart and connected by springs (bonds) of stiffness c as shown in **Figure 3**. Looking at one of the masses, it is found that stretching the springs to the left or right of that mass produces an acceleration (Newton's force law $f = ma$) such that

$$m\frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} = -c(u_i - u_{i-1}) + c(u_{i+1} - u_i) \quad [17]$$

Equation [17] can be solved to find that there are many different vibrational frequencies with different wavelengths. For each angular frequency ω, the

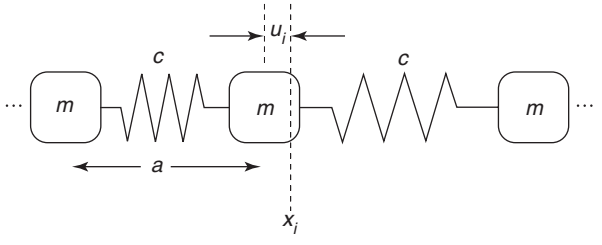


Figure 3 Simple one-dimensional mass/spring lattice.

displacements $u_i(t)$ are

$$u_i(t) = u_{i0}e^{i(kx_i - \omega t)} \quad \text{and} \quad k = \frac{2\pi}{\lambda}$$
$$\text{then } \omega = \sqrt{\frac{2c}{m}} \sin ka = s \sin ka \qquad [18]$$

Here λ denotes wavelength and s sound velocity. Considering that the full $\omega(k)$ dispersion curves describe an ensemble of simple harmonic oscillators, entropy can be computed for each element of the ensemble. For any individual vibration mode, the partition function can be used to compute its entropy. Remembering that the energies of each harmonic oscillator or normal mode i are equally spaced such that $E_i = (n + 1/2)\hbar\omega_i$, the average number of quanta per normal mode (or allowed vibration) is found to be

$$\bar{n}_i = \frac{1}{Z}\sum_{n=0}^{\infty} n \exp[-(n + 1/2)\hbar\omega_t/k_B T] = \frac{1}{e^{\hbar\omega_i/k_B T} - 1}$$
$$\text{If } k_B T \gg \hbar\omega_i \quad \text{then} \quad \bar{n}_i \approx \frac{k_B T}{\hbar\omega_i} \qquad [19]$$

Here the partition function Z is used to compute the expectation value of the quantum number. Similarly, expressions can be obtained for energy $\bar{E}_i$, free energy F_i, and entropy S_i, where

$$\bar{E}_i = \hbar\omega_i(\bar{n}_i + 1/2) \qquad [20]$$

$$F_i = -k_B T \ln Z \quad \text{and} \quad Z = \frac{1}{e^{\hbar\omega_i/2k_B T} - e^{-\hbar\omega_i/2k_B T}} \qquad [21]$$

$$TS_i = \bar{E}_i - F_i. \text{ If } k_B T \gg \hbar\omega_{\max}, \quad \text{then} \quad TS_i \approx k_B T(1 + \ln \bar{n}_i) \qquad [22]$$

where $\omega_{\max}$ denotes the highest frequency in the highest phonon branch. Computing the entropy of the entire system requires a three-dimensional average.

Differences between the phonon specific heat calculated from the full elastic tensor and that obtained by measurements is a key indicator of the importance of electronic and other nonphonon effects. At temperatures $k_B T \gg \Theta_{\max}$, where all the phonon modes populate fully, a different characteristic temperature Θ_0 (average of $\ln(\omega_i)$) is needed where

$$TS \approx 3Nk_B T[1 + \ln(T/\Theta_0)] \qquad [23]$$

Note that Θ_0 does not equal to Θ_D. Equation [23] describes the total vibrational entropy of a solid of N atoms at high temperatures, and can be calculated from the sound speed at all wavelengths and in all directions if optical modes can be neglected.

An Illustration: Plutonium

The simple approach outlined above, especially in the high-temperature limit, can establish quickly whether a theoretical model is sufficiently robust. This is illustrated by considering Pu, a mysterious metal, with at least five zero-pressure stable phases shown in **Figure 4**. The δ' b.c.t. phase is omitted as being thermodynamically unjustified. Electronic-structure calculations predict with reasonable certainty the low-temperature phases, but find that the two higher-temperature phases, f.c.c δ and b.c.c ε, are local energy saddle points and hence unstable.

One should not expect zero-temperature electronic-structure calculations (band-structure) to predict anything about such an elastically soft metal as Pu at elevated temperatures because the entropic contributions to the free energy will be unusually large. Remembering that such calculations assume zero temperature, results are only for the energy, not the free energy. In many other systems where Θ_D or other phonon averages are near or above room temperature (that is, relatively high), the free energy F is expected to include a much smaller fractional contribution from the entropy (see eqns [3] and [14].) In Pu, Θ_D is low because Pu has a very low $C' = (C_{11} - C_{12})/2$ shear modulus and very low bulk modulus $B = (C_{11} + 2C_{12})/3$, so that entropy is expected to play a much stronger role than in other f.c.c. metals. Failure to include the 740 meV/atom entropic contribution to the free energy at 750 K, the lower bound of the f.c.c. phase is doomed to failure in attempting to understand Pu's higher-temperature phase stability. For Pu, very likely, electronic-structure details do not stabilize the high-temperature crystal structures; vibrational entropy does.

Landau Theory

The focus now is on a particular class of group–subgroup materials known as ferroelastic martensites. Ferroelastics are a subclass of materials known as ferroics, which also include ferromagnetic and ferroelectric materials. The term martensite refers to a diffusionless first-order phase transition that can be described in terms of one (or several successive) shear deformation(s) from a parent to a product phase. The transition results in a characteristic lamellar microstructure resulting from transformation twinning or an equivalent lattice-invariant deformation. The morphology and kinetics of the transition are dominated by the strain energy. Ferroelasticity is defined by the existence of two or more stable orientation states of a crystal that correspond to different arrangements of the atoms, but are structurally

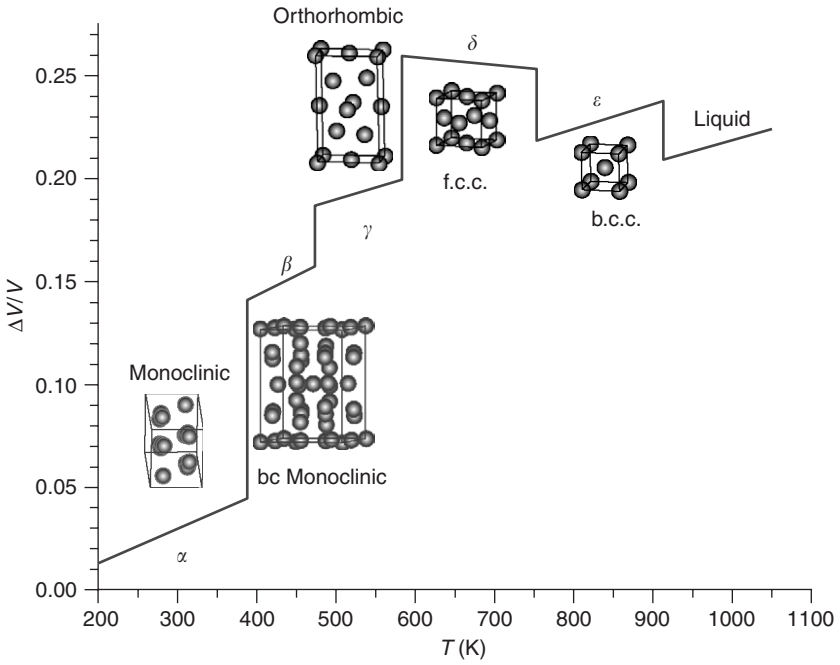


Figure 4 Measured volume changes and crystal structures of Pu at zero pressure.

identical, or enantiomorphous. Also, in the absence of mechanical stress, these orientation states are degenerate in energy.

Salient features of ferroelastic crystals include mechanical hysteresis and mechanically (reversibly) switchable domain patterns. Usually, ferroelasticity occurs as a result of a phase transition from a nonferroelastic high-symmetry "prototype" phase and is associated with the softening of a prototype-phase elastic constant with decreasing temperature or increasing pressure. Because the ferroelastic transition is usually weakly first order, or second order, it can be described to a good approximation by the Landau theory with spontaneous strain as the order parameter. Depending on whether the spontaneous strain, which describes the deviation of a given ferroelastic orientation state from the prototype phase, is the primary or secondary order parameter, the low-symmetry phase is called a proper or an improper ferroelastic, respectively. Although martensites are proper ferroelastics, examples of improper ferroelastics include ferroelectrics and magnetoelastics.

A few materials (either metals or alloy systems) are both martensitic and ferroelastic and show a shape-memory effect. They are characterized by high-mobility twin boundaries and often show precursor structures (such as tweed and modulated phases) above the transition. Furthermore, these materials show a small Bain strain, elastic shear-modulus softening, and a weak-to-moderate first-order transition. Some examples include $In_{(1-x)}Tl_x$, FePd, CuZn, CuAlZn, CuAlNi, AgCd, AuCd, $CuAuZn_2$, NiTi, and NiAl.

In many of these transitions, intracell distortion modes (or shuffles) can couple to the strain either as a primary or secondary order parameter. NiTi and titanium represent two such cases with technological importance. Other examples include actinide alloys: the shape-memory alloy UNb_6 and Ga-stabilized δ-Pu.

Landau-Theory Thermodynamics

To understand the thermodynamics of the phase transformation and the phase diagram, consider the Helmholtz free energy. The Landau free energy (LFE) is a symmetry-allowed polynomial expansion in the order parameter that characterizes the transformation, for example, strain tensor components and/or (intra-unit-cell) shuffle modes. A minimization of the LFE with the order-parameter components leads to conditions that give the phase diagram. Derivatives of the LFE with respect to temperature, pressure, and other relevant thermodynamic variables provide information about the specific heat, entropy, susceptibility, and so on. To study domain walls between different orientational variants (such as twin boundaries) or different shuffle states (such as antiphase boundaries), symmetry-allowed strain-gradient terms or shuffle-gradient terms must be added to the Landau free energy. These gradient terms are called Ginzburg terms and the augmented free energy is

referred to as the Ginzburg–Landau free energy (GLFE). Variation of the GLFE with respect to the order-parameter components leads to Euler–Lagrange equations whose solutions leads to the microstructure.

Symmetry-adapted strains are defined as a function of the Lagrangean strain-tensor components (for cubic symmetry) ε_{ij}:

$$\begin{aligned} e_1 &= \frac{1}{\sqrt{3}}(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}), \quad e_2 = \frac{1}{\sqrt{2}}(\varepsilon_{xx}-\varepsilon_{yy}), \\ e_3 &= \frac{1}{\sqrt{6}}(\varepsilon_{xx}+\varepsilon_{yy}-2\varepsilon_{zz}), \quad e_4 = \varepsilon_{xy}, \\ e_5 &= \varepsilon_{yz}, \quad e_6 = \varepsilon_{xz} \end{aligned} \qquad [24]$$

As an example, the Landau part of the elastic free energy for a cubic-tetragonal transition (for example, in FePd) in terms of the symmetry-adapted strain components is

$$\begin{aligned} F(e_2,e_3) =& \frac{A}{2}(e_2^2+e_3^2)+\frac{B}{3}(e_2^3-3e_2e_3^2) \\ &+\frac{C}{4}(e_2^2+e_3^2)^2+\frac{A_1}{2}e_1^2 \\ &+\frac{A_4}{2}(e_4^2+e_5^2+e_6^2) \end{aligned} \qquad [25]$$

Here, A, B, C denote second-order, third-order, fourth-order elastic constants, respectively; A_1 and A_4 denote bulk and shear moduli; e_2, e_3 denote order-parameter deviatoric-strain components. The non-order-parameter dilatation e_1 and shear strains e_4, e_5, e_6 are included to harmonic order. For domain walls (twinning) and microstructure, this free energy must be augmented by symmetry-allowed gradients of e_2, e_3. Coefficients in the GLFE are determined from a combination of experimental structural (lattice-parameter variation as a function of temperature or pressure), vibrational (e.g., phonon-dispersion curves along different high-symmetry directions), and various thermodynamic quantities (entropy, specific heat, elastic constants, and so on). Lacking sufficient experimental quantities, then electronic-structure calculations and molecular-dynamics simulations (using appropriate atomistic potentials) can provide the relevant information to determine some or all of the GLFE coefficients.

For simple phase transitions (for example, two-dimensional square-to-rectangle, or those involving a one-component order parameter), the GLFE can be written by inspection (from the parent-phase symmetry). However, in general, the GLFE must be determined by group-theoretic means, now readily available for all 230 crystallographic space groups.

Microstructure

For modeling the phase-transformation microstructure, there exist several related approaches: (1) GLFE-based as described above, (2) phase-field model in which strain variables couple in a symmetry-allowed manner to the morphological variables, and (3) sharp-interface models used by applied mathematicians.

Natural-order parameters in the GLFE are strain-tensor components. However, until recent years, researchers simulated the microstructure in a displacement variable by rewriting the free energy in displacement variables. This procedure leads to the microstructure without providing direct physical insight into its evolution. A natural way to gain insight is to work in strain variables only. However, if lattice integrity is maintained during the phase transformation, that is, no dislocation (or topological defect) generation is allowed, then one must obey the St. Venant elastic-compatibility constraints because various strain-tensor components are derived from the displacement field and they are not all independent. This can be achieved by minimizing the free energy with compatibility constraints treated with Lagrangean multipliers. This procedure leads to an anisotropic long-range interaction between the order-parameter strain components. This interaction (or compatibility potential) provides direct insight into the domain-wall orientations and various aspects of the general microstructure.

Mathematically, the elastic compatibility condition on the geometrically linear strain tensor $\overleftrightarrow{e}$ is given by

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \overleftrightarrow{e}) = 0 \qquad [26]$$

This represents one two-dimensional equation connecting the three components of the symmetric strain tensor. In three dimensions, two sets of three equations are needed, each connecting the six components of the symmetric strain tensor. For periodic boundary conditions in Fourier space, it becomes an algebraic equation, easily incorporated as a constraint.

The concept of elastic compatibility in a monocrystal can be generalized readily to polycrystals by defining the strain-tensor components in a global reference frame. By adding a stress term (bilinear in strain) to the free energy, one can compute the stress–strain constitutive response in the presence of a microstructure for both monocrystals and polycrystals and compare the recoverable strain upon cycling. Grain rotation and grain boundaries play an important role when polycrystals are subject to external stress. Similarly, calculation of the constitutive response can be generalized to improper ferroelastic materials such as those driven by shuffle modes, ferroelectrics, and magnetoelastics.

Finally, elasticity at the nanoscale can differ from macroscopic continuum elasticity. In this case, one must go beyond the usual elastic-tensor components and include intra-unit-cell deformation modes.

Dynamics and Simulations

The overdamped (or relaxational) dynamics can be used in simulations to obtain equilibrium microstructure: $\dot{e} = -(1/A')\delta(F + F^c)/\delta e$. Here, A' denotes a friction coefficient and F^c a long-range contribution arising from elastic compatibility. However, if the evolution of an initial nonequilibrium structure to the equilibrium state is important, one can use inertial strain dynamics with appropriate dissipation terms included in the free energy. The strain dynamics for the order-parameter strain-tensor components ε_l is given by

$$\rho_0 \ddot{\varepsilon}_l = \frac{c_l^2}{4} \vec{\nabla}^2 \left[\frac{\delta(F + F^c)}{\delta \varepsilon_l} + \frac{\delta(R + R^c)}{\delta \dot{\varepsilon}_l} \right] \qquad [27]$$

where ρ_0 denotes mass density, c_l a symmetry-specific constant, and $R = (A_l'/2)\dot{\varepsilon}_l^2$ denotes Rayleigh dissipation and R^c the contribution to dissipation arising from elastic compatibility. Overdots indicate time derivatives.

Some martensites show strain modulation (or tweed precursors) above the phase transition. These are attributed to disorder such as compositional fluctuations. They can be modeled and simulated by including symmetry-allowed coupling of strain to compositional fluctuations in the free energy. Some martensites exhibit supermodulated phases (for example 5R, 7R, 9R), which can be modeled within the Landau theory in terms of a particular phonon softening (and its harmonics) and coupling to the transformation shear.

Soft Modes

In the Landau phase-transition theory, the free energy F is expanded in some order-parameter Q. In particular, for a second-order (or continuous) transition with transition temperature T_c,

$$F(Q) = \frac{a}{2}(T - T_c)Q^2 + \frac{b}{4}Q^4 \qquad [28]$$

Q can be viewed as the atomic displacement in a mode pattern that takes the crystal from a higher-symmetry phase to a lower-symmetry phase. For small displacements, the Q^2 coefficient is effectively the spring constant for such a mode with frequency ω_0:

$$m\omega_0^2 = a(T - T_c) \qquad [29]$$

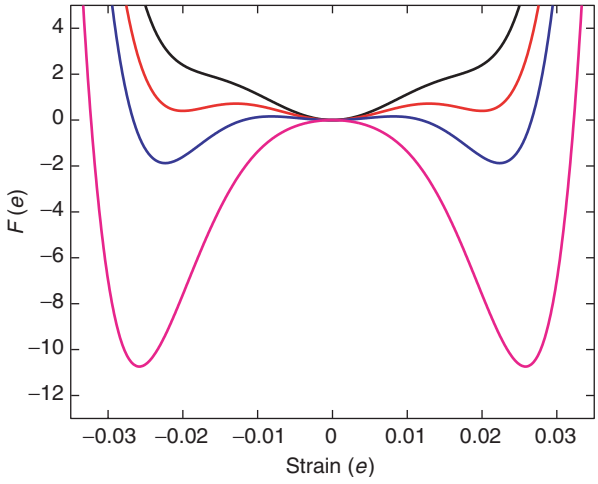


Figure 5 For a first-order phase transition, Helmholtz free energy F as a function of strain e (eqn [30] with representative parameters). The transition occurs at temperature T_c, at which the three energy minima become degenerate. Curves from top to bottom represent decreasing temperatures.

As $T \to T_c$, the phase transition is equivalent to the mode frequency $\omega_0 \to 0$. Thus arises the soft-mode concept, that is, the lattice becomes dynamically unstable; equivalently, the restoring force for the displacement in this mode vanishes. Such instabilities occur in many materials. $SrTiO_3$ provides a paradigm example. Structural phase transitions, martensites in particular, are first-order phase transitions; thus, they do not show complete mode softening at the transition temperature. The free energy in terms of a strain-tensor component is (**Figure 5**)

$$F = \frac{a}{2}(T - T_c)e^2 + \frac{b}{2}e^4 + \frac{c}{6}e^6 \qquad [30]$$

For $b < 0$, there is a discontinuous transition when $F(0) = F(e_0)$, where e_0 denotes the strain at the transition, corresponding to the two symmetric degenerate minima of $F(e)$. Because the curvature of F at $e = 0$ and at $e = e_0$ is upward and finite, the mode frequencies in both phases are nonzero; thus, the lattice is dynamically stable in both phases. However, the lattice is thermodynamically unstable. This instability is assisted by transfer of thermal energy from all other degrees of freedom in the material to move over the barrier between $e = 0$ and $e = e_0$ (**Figure 5**). Both the relevant elastic constant and the anomalous phonon soften only partially. Complete mode softening is not required.

Elastic-Inclusion Theory

The question of elastic-strain energy of an inclusion in a homogeneous matrix is old and general. It

connects not only with diffusionless phase transformations, but with many other familiar solid-state phenomena such as precipitation and twinning. A general inclusion shape poses a practically nonsolvable problem. But mathematical solutions exist for ellipsoid-shape inclusions. Ellipsoids include a surprisingly wide variety of observed cases. They range from thin disks ($c/a \ll 1$), through spheres ($c/a = 1$), to needles ($c/a \gg 1$). In 1957, Eshelby solved the general ellipsoid-inclusion/matrix elastic-energy problem.

Here, the computation of the elastic energy of a thin-disk inclusion in a homogeneous matrix is considered. This geometry corresponds to most martensites, bainites, mechanical twins, and planar precipitates. This geometry can possess zero-strain energy, unlike spheres and needles, which always show finite strain energy. Zero-strain energy arises for certain special disk orientations that lie in a plane called the "habit" plane, the austenite–martensite interface.

The backbone of the computations is the solution for the stress in, and elastic energy of, an arbitrary-orientation inclusion containing a transformation strain (eigenstrain) embedded in an anisotropic matrix. Considering an isotropic inclusion in an isotropic matrix, Eshelby laid the firm foundation for this computation. The stress in an inclusion having the same elastic constants C_{ijkl} as the matrix and with a uniform transformation strain ε_{ij}^T can be expressed as

$$\sigma_{ij} = C_{ijkl}(\varepsilon_{kl} - \varepsilon_{kl}^T) \qquad [31]$$

The total strain is

$$\varepsilon_{kl} = S_{klmn}\varepsilon_{mn}^T \qquad [32]$$

Here, S_{klmn} denotes the well-known (and tabulated) Eshelby tensor, a function only of C_{ijkl} and the inclusion shape. For ellipsoidal inclusions with semiaxes a_i, S_{klmn} can be expressed in terms of a surface integral over the unit sphere:

$$S_{klmn} = \frac{1}{8\pi} C_{pqmn} \int_{-1}^{1} \int_{0}^{2\pi} [G_{kplq}(\bar{\xi}) + G_{lpkq}(\bar{\xi})] \mathrm{d}\Theta \mathrm{d}\zeta_3 \qquad [33]$$

Here, $G_{ijkl}(\bar{\xi}) = \bar{\xi}_k \bar{\xi}_l N_{ij}(\bar{\xi})/D(\bar{\xi})$, $\bar{\xi}_\alpha = \xi_\alpha/a_\alpha$ (no sum on α), $\bar{\varsigma}_1 = (1 - \varsigma_3^{-2})^{1/2} \cos\Theta$, and $\bar{\varsigma}_2 = (1 - \varsigma_3^{-2})^{1/2} \sin\Theta$. Also, $D(\bar{\xi}) = \in_{mnl} K_{m1} K_{n2} K_{l3}$, $N_{ij}(\bar{\xi}) = (1/2) \in_{ijk} K_{lm} K_{ln}$, and $K_{ik}(\bar{\xi}) = C_{ijkl}\xi_j\xi_l$. Equation [33] is valid in a coordinate system aligned with the principal axes of the ellipsoidal inclusion. If the inclusion is spheroidal ($a_1 = a_2$), then its shape can be described solely in terms of the aspect ratio $\alpha = a_3/a_1$. The Eshelby tensor S_{klmn} simplifies considerably in the thin-disk limit ($\alpha \to 0$).

When the inclusion possesses anisotropic elastic constants C_{ijkl}^* that differ from those of the matrix, stress can be calculated by using Eshelby's equivalent-inclusion concept:

$$\sigma_{ij} = C_{ijkl}^*(\varepsilon_{kl} - \varepsilon_{kl}^T) = C_{ijkl}(\varepsilon_{kl} - \varepsilon_{kl}^T - \varepsilon_{kl}^*) \qquad [34]$$

The two terms on the right-hand side constitute an equivalency relationship that enables the inclusion with elastic constants C_{ijkl}^* to be simulated by an inclusion with elastic constants C_{ijkl}, but also containing a fictitious transformation strain ε_{ij}^*. In this case, one obtains ε_{ij} by the relationship $\varepsilon_{kl} = S_{klmn}(\varepsilon_{mn}^T + \varepsilon_{mn}^*)$.

A fundamental result obtained by Eshelby is that for the elastic-strain energy of an inclusion/matrix system:

$$E = -\frac{1}{2}\int_{\mathrm{inc}} \sigma_{ij}\varepsilon_{ij}^T \mathrm{d}V = -\frac{1}{2} V_{\mathrm{inc}} \sigma_{ij}\varepsilon_{ij}^T \qquad [35]$$

Integration is over the inclusion volume V_{inc}. The stress used in eqn [35] is computed either from eqn [31] or [34], depending on whether the inclusion elastic constants differ from the matrix elastic constants.

In general, computing the elastic-strain energy proceeds by first computing ε_{ij}^T based on the inclusion and matrix crystal structures. One then uses eqn [31] or [34] to compute the stress in the transformed inclusion, the stress caused by ε_{ij}^T. One then computes the energy from eqn [35]. One determines the minimum-energy orientation by specifying a habit-plane orientation (hkl), transforming the elastic-stiffness tensor to a coordinate system aligned with (hkl), and doing the energy computation. The preferred habit plane corresponds to that with minimum energy. With the assumption of a thin-disk inclusion, the calculation simplifies. In some cases, one can obtain explicit closed-form solutions.

As an example, the much-studied thin-disk martensite plates formed in an Fe-31Ni alloy is considered, an f.c.c.–b.c.c. transformation where the lattice-invariant deformation is twinning. Wayman's well-known book considered this alloy using both the Wechsler–Lieberman–Read and Bowles–Mackenzie approaches. **Figure 6** shows the main result of the inclusion-theory analysis described here.

After the habit plane is calculated, other phase-transformation features follow more easily. These features include magnitude of the lattice-invariant deformation (such as twinning), magnitude and direction of the macroscopic shape change, and orientation relationship.

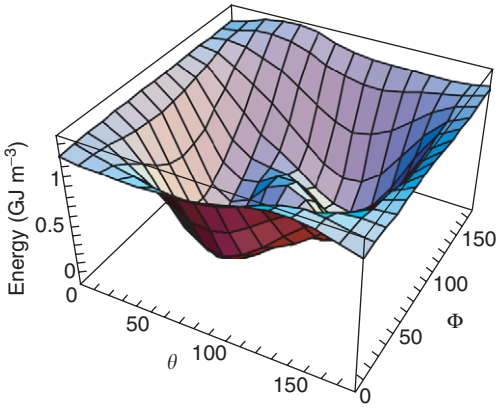


Figure 6 For an Fe-31Ni alloy, calculated elastic-strain energy vs. Θ, Φ rotation angles. The two minima correspond to two habit-plane variants. For this thin-disk-morphology case, the minima represent zero elastic-strain energy. Altogether, 24 variants occur: $(\pm h, \pm k, \pm l)$.

For the thin-disk zero-elastic-strain case, the inclusion theory gives results identical to those predicted by the invariant-plane-strain (i.p.s.) theories. But the inclusion theory allows one to move far beyond the i.p.s. theories. As examples, one can consider the following: (1) different inclusion-matrix elastic constants, (2) anisotropic elastic constants, (3) various shapes (disks, spheres, needles), (4) alternative pure distortions (such as the Bain strain), (5) various twinning systems, (6) anisotropic interface distortions, and (7) applied-stress effects.

Acknowledgments

This study was supported by the US Department of Energy.

See also: Alloys: Iron; Phase Transformation; Phases and Phase Equilibrium; Specific Heat; Superconductivity: Applications; Thermodynamic Properties, General.

PACS: 62.20.Dc; 63.20.–e; 63.70.+h; 65.40.Ba; 65.40.Gr; 81.30.Hf; 82.60.Fa

Further Reading

Bruce A and Cowley R (1981) *Structural Phase Transitions*. London: Taylor and Francis.
Christian J (1975) *Theory of Transformations in Metals and Alloys*. Oxford: Pergamon.
Fast J (1970) *Entropy; the Significance of the Concept and its Application in Science and Technology*. London: Macmillan.
Kaufman L and Hillert M (1992) *Thermodynamics of Phase Transformations, in Martensite*. Materials Park, OH: ASM International. (Many references therein.)
Landau L and Lifshitz E (1994) *Statistical Physics*. Oxford: Pergamon.
Mura T (1987) *Micromechanics of Defects in Solids*. Dordrecht: Kluwer.
Salje E (1990) *Phase Transformations in Ferroelastic and Co-elastic Solids*. Cambridge: Cambridge: University Press.
Toledano J and Toledano P (1987) *Landau Theory of Phase Transitions*. Singapore: World Scientific.
Toledano P and Dimitriev V (1996) *Reconstructive Phase Transitions*. Singapore: World Scientific.
Wallace D (1972) *Thermodynamics of Crystals*. New York: Wiley.

Dislocations

V Vitek, University of Pennsylvania, Philadelphia, PA, USA

Introduction

Dislocations are the line defects found in all crystalline materials. They are important in many branches of condensed matter physics, but most significantly they are defects, the motion of which produces plastic flow. The notion of dislocations has two starting points. First, the dislocation was introduced as an elastic singularity by considering the deformation of a body occupying a multiply connected region of space. Second, dislocations were introduced into crystal physics by Taylor, Orowan, and Polanyi in the context of analysis of the large discrepancy between the theoretical and experimental strength of crystals. These two approaches became almost immediately intertwined since the crystal dislocations are sources of long-ranged elastic stresses and strains that can be examined in the continuum framework. In fact, the bulk of the dislocation theory employs the continuum elasticity when analyzing a broad variety of dislocation phenomena encountered in plastically deforming crystals. The most comprehensive monographs devoted to the dislocation theory are books of Friedel, Nabarro, and Hirth and Lothe. An excellent simple introduction to this topic is the book by Hull and Bacon.

In this article, the crystallographic concept of dislocations is explained, the relation between dislocation movement and plastic strain is shown, the concept of the force acting on a dislocation due to externally applied stress is introduced, and the most important aspects of dislocation energetics and properties related to the long-ranged elastic fields are summarized. Then the dislocation nucleation and multiplication which is the most important ingredient

in understanding of the plastic deformation, and brittle-to-ductile transition is discussed. Finally, the properties that relate to the atomic structure of materials are paid attention to; in particular, the significance of dislocation dissociation and the dislocation core structure is elucidated.

Principal Characteristics of Dislocations

A dislocation line in a continuum is obtained by cutting the material along any surface that terminates on this line, giving the two cut surfaces a relative displacement and joining them again together. Suppose the cut is made along a planar surface terminating in a straight line, and that this line and the normal to the surface define an orthonormal coordinate system. Fixed vector translations of the cut surfaces parallel, respectively, to the three axes of this system define the three basic types of dislocation. When the displacement is parallel to the line, the result is a screw dislocation. The other two translations, both perpendicular to the line, produce (equivalent) edge dislocations; if the displacement is perpendicular to the cut, a material needs to be removed or added to rejoin the two surfaces. (Three more types of "dislocation," in which the two parts of the continuum on each side of the cut are given a relative rotation before rejoining, can be introduced. The line singularities of this type are called disclinations. These line defects play important role, for example, in liquid crystals but they are not the subjects of this article.) A straight dislocation of general mixed character is obtained when the fixed translation of the cut surfaces has components both parallel and perpendicular to the line. A curved dislocation in a plane or in three dimensions may be introduced in the same way with the surface of the cut (generally nonplanar) terminating on the curve. Since the dislocation line has been introduced as the edge of a surface, it can either terminate on external surfaces or internally by closing on itself and forming a loop. Alternatively, it may meet other dislocation lines at points where at least three dislocations meet and total sum of their Burgers vectors is zero; such points are called dislocation nodes. However, the dislocation line can never end abruptly inside the body.

The principal characteristics of a dislocation are the relative displacement of the two cut surfaces and the line terminating the cut. The former defines the Burgers vector of the dislocation and the latter the dislocation direction, the surface over which the cut has been made has no significance. In the continuum, the dislocation is defined mathematically by the relation $\oint \mathrm{d}\boldsymbol{u} = \boldsymbol{b}$, where $\boldsymbol{u}$ is the displacement vector, $\boldsymbol{b}$ the Burgers vector, and the integral is along any circuit encompassing the dislocation line; for circuits not encircling the dislocation $\oint \mathrm{d}\boldsymbol{u} = 0$. This relation is then the main boundary condition when determining within the elasticity theory, the displacement, strain, and stress fields associated with a dislocation.

Dislocations defined in the continuum have exact analogs in crystals. If the same virtual process of introducing a dislocation is employed in crystals then if the cut surfaces are to rejoin without introducing a fault along the cut, the translation vector, that is, the Burgers vector, must be a repeat vector of the Bravais lattice. A formal definition of the Burgers vector, analogous to the above path integral, invokes the concept of corresponding paths in the real crystal and a defect-free reference crystal. If a closed circuit, consisting of the same number of equivalent atoms to atom steps, is made either in the real or the reference crystal, the corresponding path in the other crystal will not, in general, be closed, and the closure failure is defined as the net Burgers vector of the dislocation lines that thread through the circuit in the real crystal.

The density of dislocations present in a body of material, ρ, is defined as the total length of dislocations per unit volume (dimension m^{-2}). The plastic strain, $\varepsilon^{\mathrm{pl}}$, which is produced if these dislocations move, on average, a distance $\bar{d}$, is $\varepsilon^{\mathrm{pl}} = \rho b \bar{d}$, where b is the magnitude of the Burgers vector of the dislocations.

Elastic Fields and Energy Associated with Dislocations

In a continuum the dislocation defined above is a line singularity invoking a long-ranged elastic field. The details of the evaluation of the strain and stress fields in the framework of linear isotropic and anisotropic elasticity can be found, for example, in the books by Nabanna and co-workers. In the case of straight dislocations their stress and strain fields decrease as d^{-1}, where d is the distance from the dislocation line. These fields have a complex angular dependence for a fixed distance from the dislocation line. Formally, both the stress and strain diverge at the dislocation line but physically this means that there is a region, centered at the dislocation line, in which the linear elasticity does not apply. This region, the dimensions of which are of the order of the lattice spacing in crystalline materials, is called the dislocation core and it can only be fully accounted for within atomic level models of dislocations. This is discussed in more detail later in this article. The stress and strain fields of a dislocation loop decrease

far away from its center as d^{-2}, similarly as the field of a dislocation dipole formed by two dislocations of opposite Burgers vectors.

The elastic energy associated with the long-range stress and strain fields of a straight dislocation is

$$E_{\mathrm{el}}^{\mathrm{sl}} = \frac{Kb^2}{4\pi}\ln\left(\frac{R}{r_{\mathrm{o}}}\right)$$

where R is the radius of a cylinder centered on the dislocation line and r_{o} is the core radius. K is a constant depending on the elastic moduli and orientation of the dislocation line. (In the isotropic case, $K = [\mu/(1-\nu)](1-\nu\cos^2\theta)$ where μ is the shear modulus, ν is the Poisson ratio, and θ is the angle between the dislocation line and the Burgers vector.) This energy diverges as $R \to \infty$ and thus depends on the size of the crystal, but it is generally assumed that when a material contains many dislocations of opposite signs, their stress fields cancel at about half of the average dislocation separation and R is identified with this distance. The elastic energy of dislocation loops is always finite and depends logarithmically on their size. For example, in an elastically isotropic medium, the energy of a glissile planar circular loop of radius r (the Burgers vector lies in the plane of the loop) is

$$E_{\mathrm{el}}^{\mathrm{loop}} = 2\pi r\frac{\mu b^2}{4\pi(1-\upsilon)}\left[\ln\left(\frac{8r}{r_{\mathrm{o}}}\right) - 1\right]$$

It is noteworthy that this energy can be interpreted as the energy per unit length of the dislocation times the length of the dislocation loop $2\pi r$. This suggests that the increase of the energy of a dislocation ΔE, associated with an increase of its length Δl, can be written as $\Delta E = \tau\Delta l$, where τ is the energy of the dislocation per unit length. In general, τ will be different for different dislocation shapes, but when the curvature of the dislocation line is small, and this concept is usually restricted to this situation, τ is well approximated as the energy per unit length of a straight dislocation. It is generally dependent on the orientation of the dislocation line, but in many analyses it has been approximated by a constant of the order of $\mu b^2/2$. Since $\tau = \Delta E/\Delta l$, it also has the meaning of a line tension, in analogy with the behavior of a taut string, and this approach to evaluation of the energy change with the change in dislocation length is called the line tension approximation. (Since the energy of the dislocation depends on the angle θ between the dislocation line and its Burgers vector, a more rigorous expression for the line tension is $\tau = E + \partial^2 E/\partial\theta^2$, analogously as in the case of anisotropic surface tension. This is not essential when using isotropic elasticity when τ is always positive and $\tau \approx E$ is a good approximation. However, in highly anisotropic crystals the more rigorously determined line tension may even become negative for certain orientations. In this case, the dislocations decrease their energy by increasing their length and changing orientation. This has been observed, for example, in β-brass.)

Forces on Dislocations and Interaction between Dislocations

According to the Colonnetti principle there is no elastic interaction between internal and external stresses, that is, the response of a medium to external loading is the same whether it is self-stressed or not. However, if the source of internal stresses is displaced, the external stress may be doing work. This is what happens in the case of a dislocation. Its motion induces the relative displacement of two parts of the body by the Burgers vector within the region swept by the dislocation. One may consider an element of the dislocation line $\Delta\boldsymbol{l}$ that is displaced by a vector $\Delta\boldsymbol{r}$ in a body subject to an external stress σ_{ij}^{e}. The area within which the material has been displaced by the Burgers vector is $\Delta S = |\Delta\boldsymbol{r} \times \Delta\boldsymbol{l}|$ and the unit vector normal to this surface has components $n_j = \varepsilon_{jlk}\Delta r_l \Delta l_k/\Delta S$ (Einstein summation convention is used). When displacing the material the external stress produced work $\Delta W = -\sigma_{ij}^{\mathrm{e}} n_j b_i \Delta S = -\sigma_{ij}^{\mathrm{e}} b_i \varepsilon_{jlk}\Delta r_l \Delta l_k$. This can be expressed alternatively as $\Delta W = F_l \Delta r_l |\Delta\boldsymbol{l}|$ with

$$F_l = \varepsilon_{ljk}\sigma_{ij}^{\mathrm{e}} b_i \varsigma_k$$

where ς_k is the k component of the unit vector parallel to the dislocation line. This expression can be interpreted such that F_l is the l component of the force, called the Peach–Koehler force, by which the stress σ_{ij}^{e} acts on the dislocation per unit length. The physical meaning of this force is that $F_l \Delta r_l$ is the work done when the dislocation of unit length is displaced by $\Delta\boldsymbol{r}$, produced by the stress σ_{ij}^{e} that is doing work when the material is being displaced by the Burgers vector in the area swept by the dislocation. For example, in the case of an edge dislocation parallel to the axis x_3 and the Burgers vector b parallel to the axis x_1, $F_1 = \sigma_{12}^{\mathrm{e}} b$ and $F_2 = -\sigma_{11}^{\mathrm{e}} b$. The component F_1 and thus the shear stress σ_{12}^{e} drive the glide in the slip plane x_1x_3, and the component F_2 and thus the tensile stress σ_{11}^{e} drive the climb perpendicular to the slip plane.

When using the Peach–Koehler formula, it is immaterial whether σ_{ij}^{e} is an externally imposed stress field or the stress arising from another internal source of stresses. Thus, if σ_{ij}^{e} is identified with the

stress associated with another dislocation, this formula determines the force between the two dislocations. The interaction energy between two dislocations marked, respectively (1) and (2), separated by a distance r, can then be evaluated as the work done when the dislocation (2) is brought in from a distance R, against the force $F^{(1-2)}$ exerted on dislocation (2) by the dislocation (1). This force obviously decreases as $1/r$. The interaction energy per unit length of the dislocations is then $E^{\text{int}}_{(1-2)} = -\int_R^r F^{(1-2)}\,\mathrm{d}r$. This interaction energy, proportional to $\ln(R/r)$, could also be evaluated as the part of the elastic energy corresponding to the cross terms between the stress and strain fields of dislocations (1) and (2), respectively. Similarly, elastic interaction between dislocations and other defects inducing internal stresses, such as inclusions, cracks, stress inducing point defects, is determined.

Dislocation Multiplication

Consider first the homogeneous nucleation of a circular glissile dislocation loop. The energy of such loop as a function of its radius is given above. The work done by the shear stress σ parallel to the Burgers vector acting in the plane of the loop during the growth of the loop is $\sigma b\pi r^2$. If the loop is to form and grow spontaneously, this work has to exceed its elastic energy already at the stage when the loop radius is only slightly larger than the radius of the core. Consider, for example, that this has to occur when $r = 2r_o$. The condition for spontaneous formation of the loops is then $\sigma = \mu b/(4\pi(1-\upsilon)r_o)\,(\ln 16 - 1)$. Since $r_o \approx b$, the stress needed for the dislocation nucleation is (taking $\nu = 1/3$) $\sigma \approx 0.2\mu$, which is of the same order of magnitude as the ideal shear strength of materials. This is a very high stress, and in most metallic, covalent, and even in ionic materials the plastic flow and intensive dislocation multiplication occur at stresses three to five orders of magnitude lower. Thus, homogeneous nucleation of individual dislocation loops cannot be the usual mechanism of dislocation nucleation. However, a collective homogeneous nucleation at finite temperatures by the mechanism akin to the Kosterlitz–Thouless transition is a possibility albeit not yet unequivocally proven.

An alternative heterogeneous mechanism, known as the Frank–Read source, requires stresses much lower than the ideal shear strength of the material. This process of generation of dislocations is shown schematically in **Figure 1**. A dislocation lying in a glide plane is pinned at two points (marked by bold dots), for example owing to intersection with dislocations in different glide planes. If these points are fixed, the dislocation can bow out under the action of a shear stress and adopt the successive configurations shown in (**Figure 1b–1f**). Eventually the two segments of opposite sign meet (**Figure 1e**), mutually annihilate, and a closed dislocation loop is produced (**Figure 1f**). A line segment between the pinning points remains available to repeat the whole process. The bowing out of the dislocation is opposed by the line tension of the dislocation that pulls the dislocation back by the force (per unit length) equal to $\tau\kappa$, where κ is the curvature of the line. In the isotropic case, the curvature is the largest when the dislocation forms the half-circle of radius d, where $2d$ is the separation of the pinning points (**Figure 1c**). Thus,

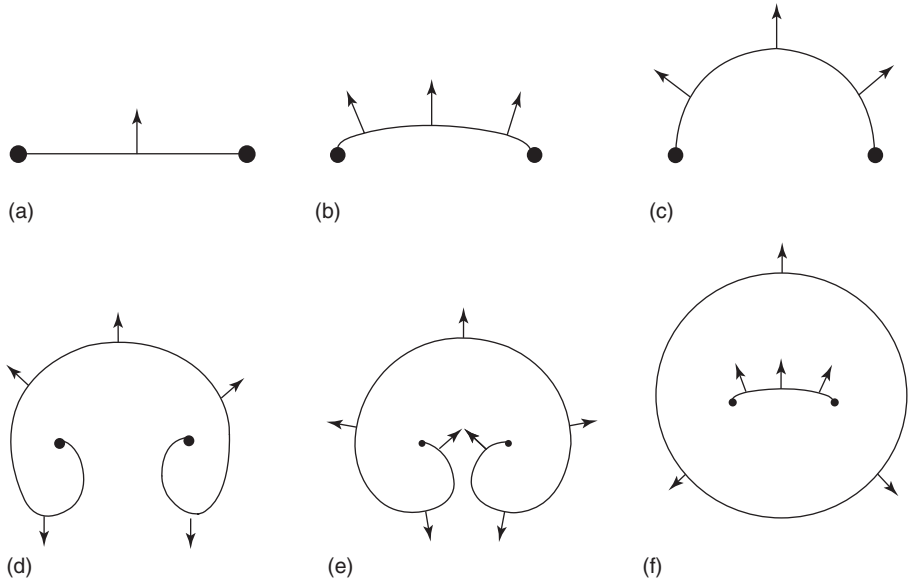


Figure 1 Successive stages in the operation of a double-needed Frank–Read dislocation source.

the condition for the onset of the operation of the Frank–Read source is $\sigma b = \tau/d$. Taking $\tau = \mu b^2/2$, the stress needed for the operation of the source is $\sigma = \mu b/2d$. Since d may be three or four orders of magnitude larger than b, this leads to a stress that is several orders of magnitude lower than the ideal shear strength of the material, in agreement with experimental observations. Hence, Frank–Read-type dislocation sources, the geometrical configurations of which may be more complex than that in **Figure 1**, represent the principal mechanism of dislocation multiplication during deformation. Nevertheless, homogeneous dislocation nucleation may still occur in the case where no sources are available. However, the stresses then reach the level of the ideal shear strength. This occurs, for example, in nanoindentation.

Dislocation Dissociation and Stacking Faults

A vital characteristic of dislocations in crystalline materials is their possible dissociation into partial dislocations with Burgers vectors smaller than the lattice vector. Such dislocation splitting can occur if the displacements corresponding to the Burgers vectors of the partials lead to the formation of metastable planar faults which then connect the partials. The reason for the splitting is, of course, the decrease of the dislocation energy when it is divided into dislocations with smaller Burgers vectors. For example, if the dislocation with the Burgers vector $\boldsymbol{b}$ splits into two partials with Burgers vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, such that $\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2$ and the angle between the vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ is acute, then $b^2 > b_1^2 + b_2^2$. Since the elastic energy of a dislocation is proportional to the square of its Burgers vector, the self-energy of the total dislocation is larger than that of two partials. This is the basis of the so-called b^2 criterion for estimation whether a dissociation is energetically favorable. However, to determine the energy balance precisely, one has to consider also the interaction energy between partials and the energy of the planar fault formed between them. A simple example is splitting of $1/2\langle 1\,1\,0\rangle$ dislocations into two Shockley partials with the Burgers vectors $1/6\langle 1\,1\,2\rangle$ on $\{1\,1\,1\}$ planes. The planar fault formed by the $1/6\langle 1\,1\,2\rangle$ displacement is an intrinsic stacking fault. The stacking of $\{1\,1\,1\}$ planes in an ideal face-centered cubic (f.c.c.) lattice can be described as a repeat sequence …ABCABCABC… and the $1/6\langle 1\,1\,2\rangle$ displacement between layers B and C produces the stacking sequence …ABCAB|ABCABC…; this corresponds to the stacking fault marked by |.

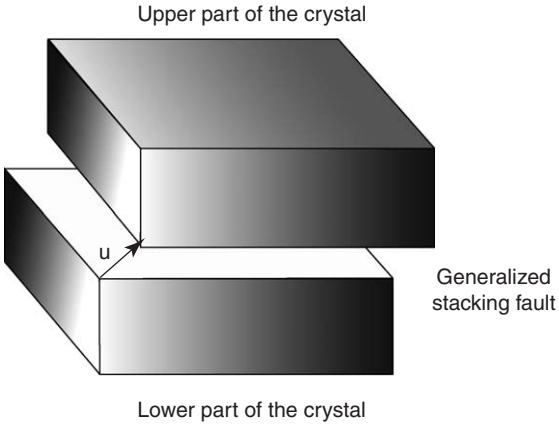


Figure 2 Formation of the generalized stacking fault.

In general, stacking-fault-like defects can be very conveniently analyzed using the notion of γ-surfaces (Stacking-fault-like defects include not only stacking faults but also other planar defects such as antiphase domain boundaries and complex stacking faults encountered in ordered alloys and compounds). To introduce the idea of γ-surfaces, a generalized stacking fault is defined first: imagine that the crystal is cut along a given crystallographic plane and the upper part displaced with respect to the lower part by a vector $\boldsymbol{u}$, parallel to the plane of the cut, as shown in **Figure 2**. The fault created in this way is called the "generalized stacking fault" and it is not, in general, metastable. The energy of such fault, $\gamma(\boldsymbol{u})$, can be evaluated using atomistic methods and/or density-functional theory based approaches; relaxation perpendicular to the fault has to be carried in such calculations. Repeating this procedure for various vectors $\boldsymbol{u}$ within the repeat cell of the given crystal plane, an energy–displacement surface can be constructed, commonly called the γ-surface. The local minima on this surface determine the displacement vectors of all possible metastable stacking-fault-like defects, and the values of γ at these minima are the energies of these faults.

Symmetry arguments can be utilized to assess the general shape of these surfaces. If a mirror plane of the perfect lattice perpendicular to the plane of the generalized stacking fault passes through the point corresponding to a displacement $\boldsymbol{u}$, the first derivative of the γ-surface along the normal to this mirror plane vanishes owing to the mirror symmetry. This implies that the γ-surface will possess extrema (minima, maxima, or inflexions) for those displacements, for which there are at least two nonparallel mirror planes of the perfect lattice perpendicular to the fault. Whether any of these extrema corresponds to minima, and thus to metastable faults, can often be ascertained by considering the change in the

nearest neighbor configuration produced by the corresponding displacement. Hence, the symmetry-dictated metastable stacking-fault-like defects can be ascertained on a crystal plane by analyzing its symmetry. Such faults are then common to all materials with a given crystal structure. The intrinsic stacking faults in f.c.c. crystals are symmetry-dictated since three mirror planes of the {1 0 1} type pass through the points that correspond to the displacements $1/6\langle 1\,1\,2\rangle$. However, other minima than those associated with symmetry-dictated extrema may exist in any particular material. These cannot be anticipated on crystallographic grounds, and their existence depends on the details of atomic interactions and can only be revealed by calculations of the γ-surface.

The primary significance of the dislocation dissociation is that it determines uniquely the slip planes, identified with the planes of splitting and corresponding stacking-fault-like defects, and, consequently, the operative slip systems. The fact that {1 1 1} planes are the slip planes in f.c.c. materials is a typical example. As discussed above, metastable stacking faults always exist on these planes and since their energy is in many cases relatively low, dislocations split in {1 1 1} planes and are confined to them.

Dislocation Cores

It has already been mentioned above that every crystal dislocation possesses a core region in which the linear elasticity does not apply and the structure and properties of the core can only be fully understood when the atomic structure is adequately accounted for. When a dislocation glides, its core undergoes changes that are the source of an intrinsic lattice friction. This friction is periodic with the period of the crystallographic direction in which the dislocation moves. The applied stress needed to overcome this friction at 0 K temperature is called the Peierls stress and the corresponding periodic energy barrier is called the Peierls barrier.

In general, the dislocation core structure can be described in terms of the relative displacements of atoms in the core region. These displacements are usually not distributed isotropically but are confined to certain crystallographic planes. Two distinct types of dislocation cores have been found, depending on the mode of the distribution of significant atomic displacements. When the core displacements are confined to a single crystallographic plane, the core is planar (a more complex planar core, called zonal, may be spread into several parallel crystallographic planes of the same type). Dislocations with such cores usually glide easily in the planes of the core spreading and their Peierls stress is commonly low. In contrast, if the core displacements spread into several nonparallel planes of the zone of the dislocation line, the core is nonplanar, extended spatially. The glide planes of dislocations with such cores are often not defined uniquely, their Peierls stress is high, and their glide well below the melting temperature is enabled by thermal activation over the Peierls barriers.

A planar dislocation core can be regarded as a continuous distribution of dislocations in the plane of the core spreading. The reason for the core extension is the decrease of the dislocation energy when it is redistributed into dislocations with smaller Burgers vectors. Such approach to the analysis of the dislocation core is the basis of the Peierls–Nabarro model. If the coordinate system in the plane of the core spreading is chosen such that the axes x_1 and x_2 are parallel and perpendicular to the dislocation line, respectively, the corresponding density of continuous distribution of dislocations has two components: $\rho_\alpha = \partial u_\alpha/\partial x_2 \ (\alpha = 1, 2)$, where u_α is the α-component of the displacement vector $\boldsymbol{u}$ in the x_1, x_2 plane. Owing to the dislocation distribution, the displacement $\boldsymbol{u}$ increases gradually in the direction x_2 from zero to the Burgers vector so that $\int_{-\infty}^{+\infty} \rho_\alpha \, dx_2 = b_\alpha$, where b_α is the corresponding component of the Burgers vector. In the continuum approximation, the elastic energy of such dislocation distribution can be expressed as the interaction energy of the dislocation distribution with itself:

$$E_{\mathrm{el}} = \sum_{\alpha,\beta=1}^{2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K_{\alpha\beta}\rho_\alpha\rho_\beta \ln(|x_2 - x_2'|)\, dx_2\, dx_2'$$

where $K_{\alpha\beta}$ are constants depending on the elastic moduli and orientation of the dislocation line. On the atomic scale, the displacement across the plane of the core spreading causes a disregistry that leads to an energy increase. The displacement $\boldsymbol{u}$ produces locally a generalized stacking fault, and in the local approximation the energy associated with the disregistry can be approximated as

$$E_\gamma = \int_{-\infty}^{+\infty} \gamma(\boldsymbol{u})\, dx_2$$

where $\gamma(\boldsymbol{u})$ is the energy of the γ-surface for the displacement $\boldsymbol{u}$. The continuous distribution of dislocations describing the core structure is then found by the functional minimization of the total energy $E_{\mathrm{tot}} = E_{\mathrm{el}} + E_\gamma$ with respect to the displacement $\boldsymbol{u}$. When the displacement vector is always parallel to the Burgers vector, the Euler equation corresponding to the condition $\delta E_{\mathrm{tot}} = 0$ leads to the well-known

Peierls equation

$$K\int_{-\infty}^{+\infty}\frac{1}{x_2-x_2'}\frac{\mathrm{d}u}{\mathrm{d}x_2'}\,\mathrm{d}x_2'=-\frac{\partial\gamma}{\partial u}$$

where u is the displacement in the direction of the Burgers vector and K is a constant depending on the elastic moduli and orientation of the dislocation line; in the isotropic elasticity K is given earlier. In the original Peierls–Nabarro model, $\partial\gamma/\partial u$ was taken as a sinusoidal function and the Peierls equation then has an analytical solution

$$u=\frac{b}{2\pi}\tan^{-1}\left(\frac{x_2}{\zeta}\right)$$

where 2ζ measures the width of the core (for $-\zeta<x_2<\zeta$, the disregistry is greater than half its maximum value) and it is proportional to the separation of the crystal planes into which the core spreads. The Peierls stress evaluated in the framework of this model is

$$\tau_{\mathrm{P}}=\frac{2\mu}{\alpha}\exp(-4\pi\zeta/b)$$

where μ is the shear modulus and α a constant of the order of 1. This stress is very small even when the core is very narrow; for example, if $\zeta=b$, $\tau_{\mathrm{P}}\approx10^{-5}\mu$. This is a general characteristic of the planar cores and this is the reason why in f.c.c. metals, in which dislocations possess planar cores, the Peierls stress is practically negligible.

The nonplanar cores can be divided into two classes: crosslip and climb cores. In the former case, the core displacements lie in the planes of core spreading, whereas in the latter case, they possess components perpendicular to the planes of spreading. Climb cores are less common and are usually formed at high temperatures by a climb process. The best-known example of the crosslip core is the core of the 1/2 $\langle1\,1\,1\rangle$ screw dislocations in body-centered cubic (b.c.c.) metals that is spread into three $\{1\,1\,0\}$ planes intersecting along the $\langle1\,1\,1\rangle$ direction. **Figure 3** shows two alternate structures of such core. The core shown in **Figure 3a** is spread asymmetrically into the $(\bar{1}\,0\,1)$, $(0\,\bar{1}\,1)$, and $(\bar{1}\,1\,0)$ planes that belong to the [1 1 1] zone and is not invariant with respect to the $[1\,0\,\bar{1}]$ diad and another energetically equivalent configuration, related by this symmetry, operation exists. The core in **Figure 3b** is invariant with respect to the $[1\,0\,\bar{1}]$ diad. These core structures were found by atomistic calculations employing (1) central-force many-body potentials and (2) a tight binding and/or density-functional theory based approach, respectively. Atomistic calculations employing appropriate descriptions of atomic interactions are the principal tool for investigations of such cores since direct observations are often outside the limits of experimental techniques. The example presented here shows that the structure of such cores is not determined solely by the crystal structure but may vary from material to material even when they crystallize in the same structure. The Peierls stress of dislocations with these cores is commonly several orders of magnitude higher than that of dislocations with planar cores. Furthermore, the movement of the dislocations is frequently affected not just by the shear

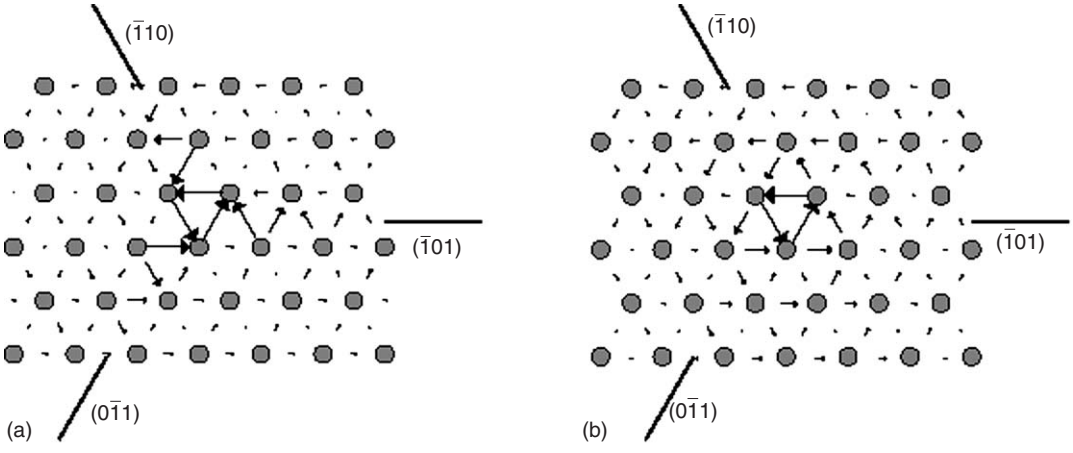


Figure 3 Two alternate core structures of the 1/2 [1 1 1] screw dislocation depicted using differential displacement maps. The atomic arrangement is shown in the projection perpendicular to the direction of the dislocation line ([1 1 1]) and circles represent atoms within one period. The [1 1 1] (screw) component of the relative displacement of the neighboring atoms produced by the dislocation is represented by an arrow between them. The length of the arrows is proportional to the magnitude of these components. The arrows, which indicate out-of-plane displacements, are always drawn along the line connecting neighboring atoms and their length is normalized such that it is equal to the separation of these atoms in the projection when the magnitude of their relative displacement is equal to |1/6 [1 1 1]|.

stress parallel to the Burgers vector in the slip plane but by other stress components. Thus, the deformation behavior of materials with nonplanar dislocation cores may be very complex, often displaying unusual orientation dependencies and breakdown of the Schmid law. The nonplanar dislocation cores are the more common the more complex is the crystal structure, and thus these cores are more prevalent than planar cores. In this respect, f.c.c. materials (and also some hexagonal close-packed (h.c.p.) materials with basal slip) in which the dislocations possess planar cores and, consequently, the Peierls stress is very low, are a special case rather than a prototype for more complex structures.

Additional complex features of the dislocation core structures arise in covalent crystals where the breaking and/or readjustment of the bonds in the core region may be responsible for a high lattice friction stress, and in ionic solids where the cores can be charged which then strongly affects the dislocation mobility. Such dislocation cores affect not only the plastic behavior but also electronic and/or optical properties of covalently bonded semiconductors and ionically bonded ceramic materials.

See also: Mechanical Properties: Anelasticity; Mechanical Properties: Creep; Mechanical Properties: Elastic Behavior; Mechanical Properties: Fatigue; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Mechanical Properties: Tensile Properties; Periodicity and Lattices; Recovery, Recrystallization, and Grain Growth; Thin Films, Mechanical Behavior of.

PACS: 61.72.Fp; 61.72.Hh; 61.72.Lk; 72.10.Fk

Further Reading

Colonnetti G (1915) *Accadutti. Licwinaz R. C.* 24: 404.
Eshelby JD (1956) In: Seitz F and Turnbull D (eds.) *Solid State Physics*, vol. 3, p. 79. New York: Academic Press.
Friedel J (1964) *Dislocations*. Oxford: Pergamon.
Hirth JP and Lothe J (1982) *Theory of Dislocations*, 2nd edn. New York: Wiley.
Hull D and Bacon DJ (2001) *Introduction to Dislocations*. Oxford: Butterworth-Heinemann.
Khantha M, Pope DP, and Vitek V (1994) *Physical Review Letters* 73: 684.
Kléman M and Lavrentovich OD (2003) *Soft Matter Physics: An Introduction*. New York: Springer.
Kosterlitz JM and Thouless DJ (1973) *Journal of Physics C: Solid State Physics*. 6: 1181.
Lardner RW (1974) *Mathematical Theory of Dislocations and Fracture*. Toronto: University of Toronto Press.
Nabarro FRN (1967) *Theory of Crystal Dislocations*. Oxford: Clarendon Press.
Orowaon E (1934) *Zeitschrift für Physik* 89: 634.
Paidar V and Vitek V (2002) In: Westbrook JH and Fleisher RL (eds.) *Intermetallic Compounds: Principles and Practice*, vol. 3, p. 437. New York: Wiley.
Peach MO and Koehler JS (1950) *Physical Review* 80: 436.
Polanyi M (1934) *Zeitschrift für Physik* 89: 660.
Taylor GI (1934) *Proceedings of Royal Society London A* 145: 362.
Vitek V (2004) *Philosophical Magazine* 84: 415.
Volterra V (1907) *Ann. Sci. Ec. Norm. Sup. Paris* 24: 401.

Disorder and Localization Theory

E N Economou, University of Crete, Greece
C M Soukoulis, Iowa State University, Ames, IA, USA

Introduction

A plane electronic wave, $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$, impinging upon a local potential, $v(|\boldsymbol{r}-\boldsymbol{r}_n|)$, centered at $\boldsymbol{r}_n$, is scattered, thus giving rise to a spherical wave of the form $f(\theta,\phi)\exp(i\boldsymbol{k}\cdot\boldsymbol{r})/r$ as $r\to\infty$. The scattering amplitude, f, is related to the differential $d\sigma/d\Omega(=|f(\theta,\phi)|^2)$, and the total scattering cross section, σ. The total scattering cross section, $\sigma=\int|f|^2\,d\Omega$, gives essentially the part of the area of the incoming wave front intercepted by the scattering potential. For transport properties, the relevant cross section is $\sigma_t=\int(1-\cos\theta)|f|^2\,d\Omega$.

An electron propagating within a solid or liquid is subject to many local scattering potentials stemming from the atoms comprising the solid or liquid, and giving rise to a total scattering potential $V(\boldsymbol{r})=\sum_n v(|\boldsymbol{r}-\boldsymbol{r}_n|)$. The response of electrons to $V(\boldsymbol{r})$ is of central importance in understanding and controlling the mechanical, electrical, magnetic, and optical properties of materials.

Identical and periodically placed scatterers allow a systematic constructive interference of the scattered waves, which to a great extent, compensates the effects of the scattering. As a result, a periodic potential, $V(\boldsymbol{r})$, produces plane-wave-like solutions, called Bloch waves, of the form $u_{kj}(\boldsymbol{r})\exp[i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_{kj}t)]$, associated with bands of allowed values of ω_{kj}; u_{kj} is a periodic function of $\boldsymbol{r}$ with the same period as $V(\boldsymbol{r})$. Gaps in the spectrum also appear because of systematic destructive interference (another effect of

periodicity) and/or well-separated atomic or molecular levels giving rise to nonoverlapping bands.

"Classical" Treatment of Disorder

In actual solids, there are always static, chemical and, structural deviations from periodicity, located randomly within the solid; in other words, there is always disorder. The thermal vibrations of atoms (or ions), although dynamical in nature, can be treated for fast processes and high temperatures as additional sources of static disorder. However, at low temperatures, lattice vibrations and other dynamical processes lead to inelastic scatterings, which are clearly distinct from the elastic scatterings considered up to now; these inelastic scatterings are responsible for an energy uncertainty, ΔE, which in turn defines an inelastic dephasing time $\tau_\phi \sim \hbar/\Delta E$; for $t \gtrsim \tau_\phi$, the phase is randomized.

Disorder is crucial in determining important transport properties, such as electrical and thermal conductivities, magnetoresistances, photoconductivities, and metal–insulator transitions. Under normal conditions, the main role of disorder is to change the propagation of electrons (or more generally, of whatever carriers) from ballistic to diffusive; the latter involves a new characteristic quantity, the so-called transport mean free path, l, which is connected to the diffusion length $L_D = (D\tau)^{1/2}$ by the relation $L_D = l/\sqrt{d}$. D is the diffusion coefficient, $\tau = l/v$, v is the carrier velocity, and d is the space dimensionality. The transport mean free path, for low concentration, $n_s = N_s/V$, of scatterers is related to the transport cross section σ_{ti} of each scatterer by the relation

$$\frac{1}{l} = \frac{1}{V}\sum_{i=1}^{N_s} \sigma_{ti} \qquad [1]$$

which means that over a length l, the N_s scatterers within the volume $V = Sl$ will intercept the whole cross section, S, of the incoming wave (i.e., $S = \sum_i \sigma_{ti}$).

The summation of σ_{ti}s in eqn [1] implies that interference effects have been omitted; thus the term "classical" in the title of this subsection. Actually, a scattered wave from one scatterer can be scattered again by other scatterers; this is especially true when the σ_{ti}s are large and their concentration, n_s, is high. All these single and multiple scattered waves can, in principle, interfere with one another, possibly modifying the basic result in eqn [1]. The argument in support of eqn [1] is that constructive and destructive interference more or less cancel each other as a result of the interfering waves having random phase differences (associated with the random positions and phase shifts of the scatterers).

Weak Disorder and the Role of Interferences

Consider an electron initially ($t = 0$) at the point $\boldsymbol{r}$. The probability density amplitude, $A(t)$, to again find the electron at $\boldsymbol{r}$, after time t, is given by $A(t) = \sum_v A_v(t)$, where the sum is over all directed path integrals, $A_v = \exp[(i/\hbar)\int_0^t L(\boldsymbol{r}_v(t'))\,dt']$, starting and ending at $\boldsymbol{r}$. The probability density, $dP(t)/d^3r$, to find the electron at $\boldsymbol{r}$ after time t is $dP/d^3r = |A(t)|^2 = \sum_v |A_v(t)|^2 + \sum_{v\neq\mu} A_v(t)A_\mu^*(t)$. Omitting the last double sum on the basis of random phases is equivalent to the approximation leading to eqn [1]. However, not all terms with $\mu \neq v$ have random phases; indeed, for each path $\mu = \bar{v}$, where $\bar{v}$ is the same as path v but run in the opposite direction, $A_\mu = A_{\bar{v}} = A_v$ (the last equality stems from time-reversal symmetry), and consequently $\sum_{v\neq\mu} A_v A_\mu^* = \sum_v A_v A_{\bar{v}}^* + \sum_{\mu\neq v,\bar{v}} A_v A_\mu^*$. Thus, making the reasonable assumption that $\sum_{\mu\neq v,\bar{v}} A_v A_\mu^*$ is really zero because of random phases, one may come to the conclusion that dP/d^3r is twice as big as the classical diffusion would predict, since $\sum_v A_v A_{\bar{v}}^* = \sum_v |A_v|^2$. Actually, $dP(t)/d^3r$ is less than twice as large because the equality $A_{\bar{v}} = A_v$ is not valid for very long paths, whose length far exceeds the inelastic dephasing length $L_\phi \equiv (D\tau_\phi)^{1/2}$; for such long paths, repeated inelastic scattering destroys the phase equality of $A_{\bar{v}}$ and A_v. In any case, the conclusion is that interference effects make quantum diffusion slower than classical diffusion, and lead to a decrease of the diffusion coefficient and the electrical conductivity. This decrease is proportional to the integral $\int dt \sum_v A_v A_{\bar{v}}^*$.

The latter can be estimated by taking into account that the paths which contribute are inside a tube of cross section λ^2 (or λ^{d-1} for a d-dimensional system, where λ is the wavelength) around the classical trajectory. Furthermore, the probability of returning is directly related to the probability of self-intersection of this orbital tube. Within time dt, the wave would move by $v\,dt$ and would sweep a volume $dV = \lambda^{d-1} v\,dt$; hence, the probability of self-intersection during the time interval dt, after time t has elapsed, is equal to the ratio of the volume $dV = \lambda^{d-1} v\,dt$ over the total volume swept up to t; the latter is of the order of $L_D^d = (Dt)^{d/2}$. Hence, the probability of self-intersection at any time between $t = \tau$ and $t = \tau_\phi$ is proportional to $\int_\tau^{\tau_\phi} dt \lambda^{d-1} v/(Dt)^{d/2}$. The elastic collision time, $\tau = l/v$, has been set as the lower limit because, for $t \ll \tau$, the motion is ballistic and the probability of self-intersection is zero; the inelastic dephasing time τ_ϕ is the upper limit since for $t \gtrsim \tau_\phi$,

the phase equality of A_v and $A_{\bar{v}}$ is lost. Notice that at high temperatures, where $\tau_\phi \approx \tau$, the quantum corrections are negligible. On the other hand, for low temperatures, where $\tau_\phi \gg \tau$, there is an appreciable increase in the probability, p, of returning to the initial point and consequently an appreciable correction, $\delta\sigma = \sigma - \sigma_0$, to the classical conductivity, σ_0, as a result of interference effects:

$$\delta\sigma \sim -p \sim 1 - \left(\frac{\tau_\phi}{\tau}\right)^{1/2} \approx -\left(\frac{\tau_\phi}{\tau}\right)^{1/2} = -\frac{L_\phi}{L_D} \text{ (1D systems)} \quad [2a]$$

$$\delta\sigma \sim -p \sim -\frac{\lambda}{L_D}\ln\frac{\tau_\phi}{\tau} = -\frac{2\lambda}{L_D}\ln\frac{L_\phi}{L_D} \text{ (2D systems)} \quad [2b]$$

$$\delta\sigma \sim -p \sim -\frac{\lambda^2}{L_D^2}\left[1 - \left(\frac{\tau}{\tau_\phi}\right)^{1/2}\right] = -\frac{\lambda^2}{L_D^2}\left(1 - \frac{L_D}{L_\phi}\right) \approx 0 \text{ (3D systems)} \quad [2c]$$

The above formulas are valid when the disorder is weak, so that $|\delta\sigma|$ is considerably smaller than σ_0, since otherwise, one could end up with negative conductivity $\sigma = \sigma_0 + \delta\sigma$. The condition $|\delta\sigma| \ll \sigma_0$ implies that $\lambda \ll L_D$, which in turn means that under the present conditions, $\delta\sigma$ is negligible for 3D systems. On the other hand, for 1D and 2D systems, $\delta\sigma$ can become appreciable when $L_\phi \gg L_D$; furthermore, since τ_ϕ is proportional to T^{-1} for 2D and $T^{-2/3}$ for 1D systems, and for very low temperatures, T, it follows that σ for $d = 1$ and $d = 2$ decreases with decreasing T. Such a T-dependence suggests an insulating behavior as $T \to 0$K, and consequently, the nonexistence of one- and two-dimensional metallic behavior. In other words, the possibility is raised that in 1D and 2D systems, as $L_\phi \to \infty$, the conductivity may become zero, no matter how weak the disorder is. The possibility of $\sigma \to 0$ as $L_\phi \to \infty$ may be realized even in 3D systems when the disorder is strong enough, so that $\lambda \gtrsim L_D$. This question is taken up again in the next subsection, in which formulas [2a]–[2c] are generalized to the strong disorder case.

This subsection is concluded by pointing out that the presence of a magnetic field, $\boldsymbol{B}$, breaks the time-reversal symmetry, and as a result, breaks the equality $A_v = A_{\bar{v}}$, and consequently, reduces the magnitude of $|\delta\sigma|$, and thus, tends to restore the classical behavior. Actually, the magnetic field adds a term $q\boldsymbol{A}\cdot\boldsymbol{v}$ to the Lagrangian, where q is the charge of carriers ($-e$ for electrons), and $\boldsymbol{A}$ is the vector potential such that $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. So the closed path integrals, A_v, $A_{\bar{v}}$, in the presence of $\boldsymbol{B}$, acquire an extra phase, that is $A_v = A_{v0}\exp(i\phi_v)$, $A_{\bar{v}} = A_{v0}\exp(-i\phi_v)$, where $\phi_v = q\Phi_v/\hbar$, and Φ_v is the magnetic flux through the closed path. As a result $A_v A_{\bar{v}}^* + A_{\bar{v}} A_{\bar{v}}^* = 2|A_{v0}|^2\cos(2\phi_v)$, instead of $2|A_{v0}|^2$ as in the case $\boldsymbol{B} = 0$. This implies a periodic variation of the conductance as a function of the applied magnetic field if Φ_v is a constant independent of v. Indeed, this phenomenon has been observed in cylindrical tubes of cross section S with very thin walls, and with the magnetic field parallel to the axis of the tube. The configuration is such that $\Phi_v = nBS$, where $n = 1, 2, \ldots$. Because L_ϕ^2 is not much larger than S, the dominant contribution comes from $n = 1$ and the conductance, G, varies as $\delta G \sim \cos(2eBS/\hbar)$; this is the so-called $h/2e$ Aharonov–Bohm effect. In ring configurations, such as in **Figure 1c**, where the conduction paths are between diametrically opposite points, the variation of the conductance, δG, is proportional to $\cos(eBS/\hbar)$; this is the so-called h/e Aharonov–Bohm effect. Notice that the Aharonov–Bohm effects can provide an operational definition of the length L_ϕ, since for lengths $L \gtrsim L_\phi$, the characteristic conductance oscillations gradually disappear. In thin wires, such as in **Figures 1d** and **1e**, the oscillatory magnetic field dependence of the conductance is more complicated as shown in **Figure 2**. It must be pointed out that these oscillations are reproducible (for the same system) and that their size is universal and $\sim e^2/\pi\hbar$. Furthermore, the conductance in the system of **Figure 1d** is different from that in the system of **Figure 1e**, in spite of the classical current carrying paths being identical in the two configurations; this shows the wave nature of electronic propagation, the non-negligible role of interference effects associated with the existence of the ring in **Figure 1e**, and the nonlocal nature of the current–electric field relation.

Disorder, Density of States, and Wave Functions

Three-Dimensional systems

In **Figure 3a**, the density of states (DOS) of a band associated with the s-level of identical atoms placed periodically in a simple cubic lattice is shown; the bandwidth is $12|V_2|$, where V_2 is the matrix element transferring electrons between two neighboring atoms. Disorder is introduced (**Figures 3b–3d**) by assuming that the s-levels of atoms are independent random variables possessing a Gaussian distribution of standard deviation w. The effects of the disorder are the following: (1) The analytical singularities

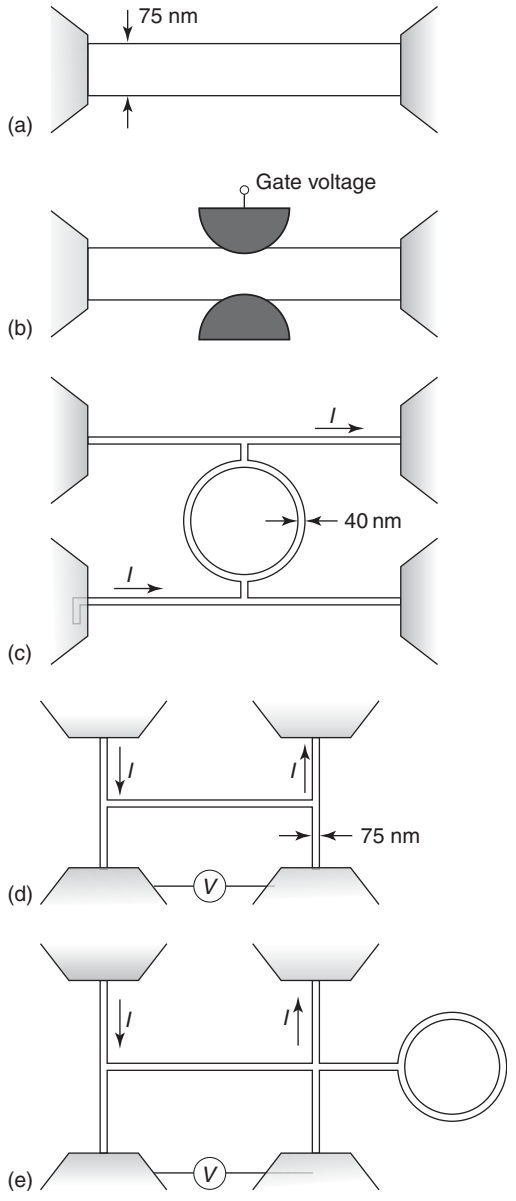


Figure 1 Various mesoscopic wires. (a) Simple quantum wire. (b) The gate voltage, V_G, creates a restriction which gradually, by changing V_G, leads to the 1D limit. (c) Four probe ring configuration which splits the current to the two semicircular paths. (d, e) Four probe configurations differing by the presence of the ring in case (e). The cross sections of the wires are $\sim 10^3\,\text{nm}^2$, and their length $\sim \mu\text{m}$.

shown in **Figure 3a** disappear. (2) The bandwidth increases as the band edges $\pm E_B^{(0)}$ move to new positions $\pm E_B$, where $E_B - E_B^{(0)} = c_1 w^2/|V_2|$ ($c_1 \approx 0.25$ in the present case). (3) Besides this widening of the band, tails in the DOS, $\rho(E)$, appear which are of exponential nature, $\rho(E) \sim \exp(-|E|/E_0)$ for $E_B \lesssim |E|$, where $E_0 = c_2 w^2/|V_2|$ ($c_2 \approx 0.12$ in the present case); these tails give rise to tails in the optical absorption of semiconductors for $\hbar\omega$ less than the gap (Urbach tails). (4) Two characteristic energies, $\pm E_c$, termed mobility edges, appear at which the nature of the eigenstates changes from propagating (or extended) to nonpropagating (or localized, shaded regions in **Figures 3b–3d**). The localized eigenstates are trapped either around a cluster of atoms or around a single atom whose *s*-energy level is much different than the average. For low disorder ($w \lesssim 2.5|V_2|$), the mobility edges follow the band edges closely, $|E_B| - |E_c| \sim w^4/|V_2|^3$. As the disorder increases, the mobility edges start moving toward the center of the band, and eventually they merge together making all eigenstates of the band localized; this is known as the Anderson transition (in the present case, this transition occurs when $w \simeq 6.2|V_2|$).

For weak disorders (e.g., **Figure 3b**), the states near the center of the band are quasi-Bloch characterized by the band index, j, the wave vector, $\boldsymbol{k}$, and the mean free path, which is connected to the phase $\phi(\boldsymbol{r})$ of the eigenfunction through the relation $\langle \exp[i\phi(\boldsymbol{r}) - i\phi(0)] \rangle = \exp(-r/2l')$, where the symbol $\langle\ \rangle$ denotes the average value over all random variables. The mean free path l' satisfies eqn [1] with σ_i instead of σ_{ti}. As one moves from the center of the band toward the mobility edges, the amplitude of the eigenstates develops ever increasing fluctuations, both in magnitude and spatial extent up to a maximum length, ξ; for distances, r, such that $a \ll r \lesssim \xi$, the eigenfunctions exhibit a statistically self-similar, fractal, behavior (a is the lattice spacing). As $|E| \to |E_c|$, ξ blows up, $\xi \to A/(|E_c| - |E|)^\gamma$, where γ has been estimated to be 1.58 (in the absence of a magnetic field) and 1.43 (in the presence of a magnetic field). On the localized sides of the spectrum, the eigenfunctions, $\psi(\boldsymbol{r})$, decay exponentially on the average, that is, $\langle |\psi(\boldsymbol{r})| \rangle_g \sim \exp(-r/L_c)$, as $r \to \infty$, where $\langle |\psi(\boldsymbol{r})| \rangle_g \equiv \exp[\langle \ln|\psi(\boldsymbol{r})| \rangle]$, and L_c is the so-called localization length. For distances, r, such that $a \ll r \lesssim L_c$, the eigenfunctions exhibit fluctuations of statistically self-similar, fractal nature. As $|E| \to |E_c|$, L_c blows up: $L_c \to A'/(|E| - |E_c|)^\gamma$.

The contribution of a band to the conductivity, $\sigma(T)$, is given by $\sigma(T) = \int dE \sigma(E,T)(-\partial f/\partial E)$, where $f(E)$ is the Fermi–Dirac distribution, $\sigma(E,T)$ is proportional to the DOS $\rho(E)$ and the mobility $\mu(E)$ at E. The conductivity $\sigma(E,T)$ for a finite cube of length L at $T = 0\,\text{K}$ is given by

$$\sigma(E,0) = \frac{e^2}{\hbar}\left(\frac{0.066}{\xi} + \frac{A_1}{L}\right) \qquad [3]$$

for E in the extended region (A_1 is ~ 0.05), $\sigma(E,0) \sim \exp(-2L/L_c)$ in the localized regime, and for $L_c \ll L$. Actually, L, besides the length of the specimen, can be of any number of upper cut-off lengths such as L_ϕ or $R_B = (\hbar c/eB)^{1/2}$, the cyclotron

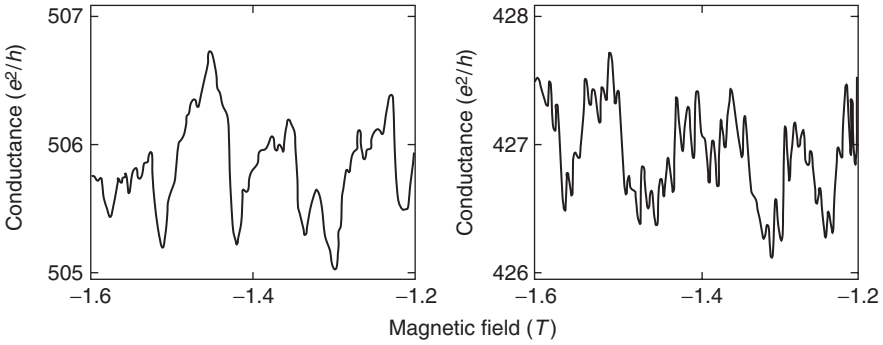


Figure 2 Electrical conductance vs. applied magnetic field for the configuration of **Figure 1d** (left) and for that of **Figure 1e** (right). (Reproduced with permission from Webb RA and Washburn S (1988) Quantum interference fluctuations in disordered metals. *Physics Today* 41: 46; © American Institute of Physics.)

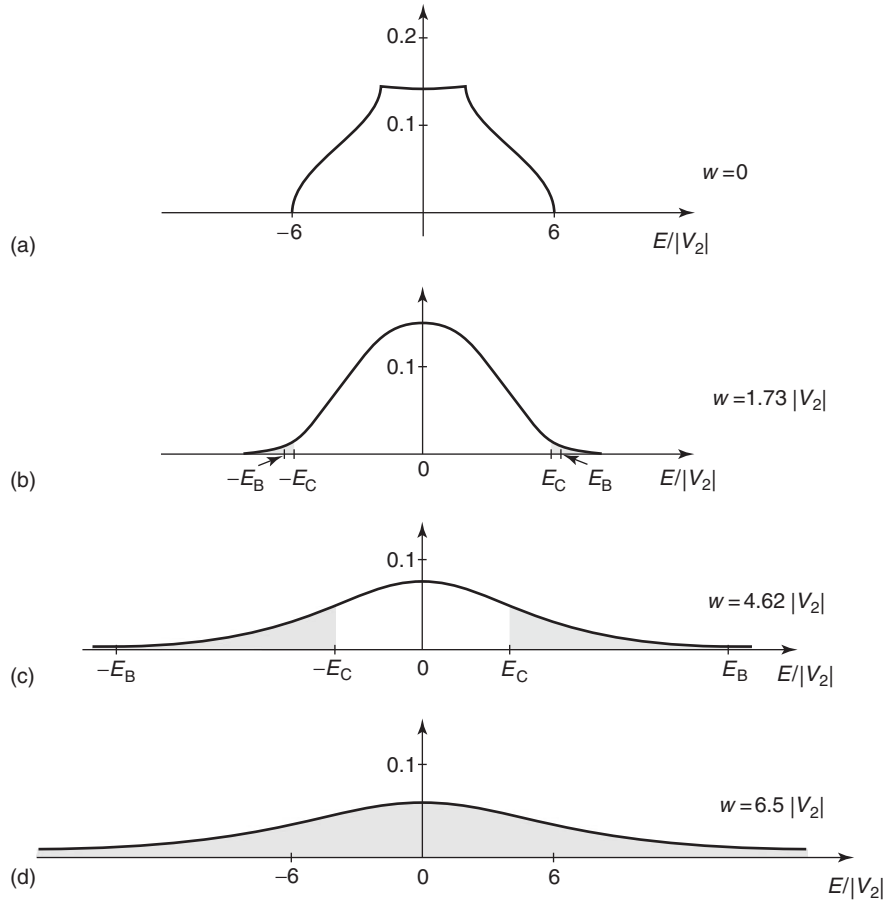


Figure 3 The density of states (DOS) of a band associated with the *s*-level of atoms placed periodically in a simple cubic lattice. The *s*-level of each atom is an independent random variable of Gaussian probability distribution with zero-average value and standard deviation equal to w, which goes from zero (case (a)) to $6.5|V_2|$ (case (d)). The bandwidth at $w = 0$ is $12|V_2|$. Shaded regions correspond to localized eigenstates.

radius in the presence of a magnetic field, or $L_\omega = (D/\omega)^{1/2}$, the diffusion length in the presence of an AC field of frequency ω, etc. In the presence of all these upper cut-off lengths, L in the above formulas should be replaced by an effective length, $\tilde{L}$, probably of the form $\tilde{L}^{-2} = c_1 L^{-2} + c_2 L_\phi^{-2} + c_3 L_\omega^{-2} + c_4 R_B^{-2} + \cdots$, where the weights c_i are of the order of unity.

Two-Dimensional Systems

2D systems, such as very thin films or electrons trapped at the interface of Si/SiO_2 or $GaAs/Al_xGa_{1-x}As$,

are borderline as far as localization properties are concerned. In the absence of magnetic forces, all eigenstates are localized, no matter how weak the disorder is (exceptions do exist). In the presence of magnetic forces, the situation is qualitatively similar to that shown in **Figure 3b–3d** with an extremely narrow region of extended states at the center of the band. This picture has been employed to interpret the so-called integral quantum hall effect (IQHE), where as the Fermi level moves in the localized regime, the Hall resistance does not change until it enters the extended regime and a rather abrupt step is exhibited.

In the absence of magnetic forces, the localization length at E is related to the mean free path $l(E) : L_c(E) \simeq 2.7l(E)\exp[S(E)l(E)/4]$, where $S(E)$ is the length of the line in $\boldsymbol{k}$-space satisfying the equation $E_k = E$. The conductivity $\sigma(E,0)$ is given by an approximate interpolation formula involving $\tilde{L}$ and L_c as well as a length $r_0 \simeq c'l$. In the limit $\tilde{L} \ll L_c$,

$$\sigma \approx \sigma_0 - \frac{e^2}{\pi^2\hbar}\ln\frac{\tilde{L}}{r_0} \qquad [4]$$

while for $L_c \ll \tilde{L}$, $\sigma = (2\sigma_0\tilde{L}/L_c)\exp(-2\tilde{L}/L_c)$. Equation [4] is similar to eqn [2b] and is known as a weak localization correction. For $\tilde{L} \simeq R_B$, eqn [4] gives $\Delta\sigma \equiv \sigma(B) - \sigma(0) = (e^2/2\pi^2\hbar)\ln[1 + (eL_\phi^2 B/1.78\hbar c)]$, that is a negative magnetoresistance, $\Delta R/R \simeq -\Delta\sigma/\sigma$, as shown in **Figure 4**. For $L \simeq L_\phi \sim T^{-1/2}$, the above formulas for the conductivity show that $\sigma \to 0$ as $T \to 0$ K. Notice, however, that electron–electron interactions produce a similar temperature dependence. Nevertheless, recent experiments at very low temperatures, in very clean interfaces $GaAs/Al_xGa_{1-x}As$ or Si/SiO_2, have shown that $\sigma(T)$ stops decreasing and starts increasing with decreasing temperatures.

One-Dimensional and Quasi-One-Dimensional Systems

For 1D systems, such as the one shown in **Figure 1b**, all eigenstates are localized, no matter how small the disorder is (exceptions do exist); the localization length is proportional to the mean free path, $L_c = cl$, where the coefficient c equals 4 for not-so-strong a disorder. The conductivity, σ, is given by $\sigma = (e^2\tilde{L}/\pi\hbar)/[\exp(2\tilde{L}/L_c) - 1]$. Thus in the limit $\tilde{L} \ll L_c$, $\sigma = (e^2L_c/2\pi\hbar) - (e^2\tilde{L}/2\pi\hbar)$; the first term gives the classical result $\sigma = 2e^2l/\pi\hbar$, since $L_c = 4l$, while the second is linear in $\tilde{L}$ as in eqn [2a]. Taking into account that the conductance G in 1D equals to $\sigma/\tilde{L}$ and that the geometrically averaged transmission coefficient, T, equals to $\exp(-2\tilde{L}/L_c)$, the formula for σ can be recast as $G = (e^2/\pi\hbar)T/(1 - T)$. However, a direct calculation of G gives $G = (e^2/\pi\hbar)T$. This discrepancy is due to the fact that the first formula gives the conductance of the wire *per se* without taking into account the two contact resistances $2R_c = \pi\hbar/e^2$, and it requires a four-probe measurement for its determination; on the contrary, the result $G = (e^2/\pi\hbar)T$ is the outcome of a two-probe measurement and it includes the contact resistances, so that $1/G = 2R_c + R_w = (\pi\hbar/e^2) + (\pi\hbar/e^2)[(1 - T)/T] = (\pi\hbar/e^2)/T$. In the presence of M coupled parallel to 1D channels in the limit where $\tilde{L} \ll l$ and $T \simeq 1$, the two-probe measurement will give $G = (e^2/\pi\hbar)M$ in agreement with the experimental results in **Figure 5**. The importance of the above formulas lies in their

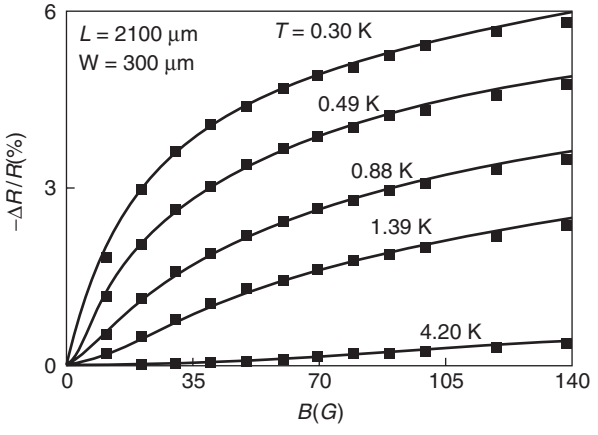


Figure 4 Experimental data (solid squares) and theoretical results (continuous lines) for the relative magnetoresistance $-\Delta R/R$, in a two-dimensional sample of electronic concentration $n = 1.6 \times 10^{11}$ cm^{-2} and mobility $\mu = 27\,000$ cm^2 V^{-1} s^{-1} (Reproduced from Choi KK *et al.* (1987) Dephasing time and one-dimensional localization of two-dimensional electrons in $GaAs/Al_xGa_1 - xAs$ heterostructures. *Physical Review B* 36: 7751.)

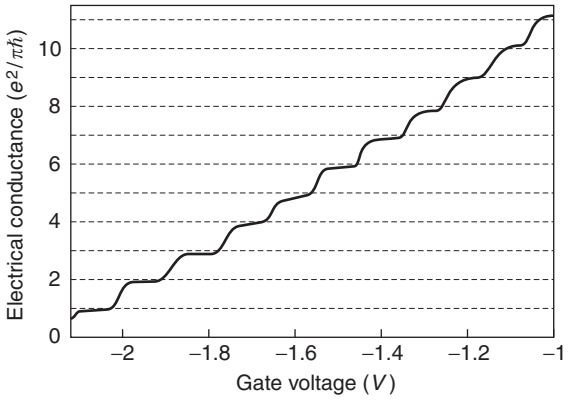


Figure 5 The electrical conductance of the configuration shown in Figure 1b vs. the gate voltage, V_g. The step-like dependence on V_g according to the formula $G = (e^2/\pi\hbar)M$, $M = 1, 2, \ldots$, is shown. For $V_g < -2V$, the 1D limit $M = 1$ has been reached (Reproduced from van Wees BJ *et al.* (1988) Quantized conductance of point contacts in a two-dimensional electron gas. *Physical Review Letters* 60: 848.)

capability to be generalized to a multiterminal system; each terminal, i, is in contact with a reservoir at electrochemical potential $\mu_i = -|e|V_i$, and feeds the system with a net current I_i

$$I_i = \frac{e^2}{\pi\hbar}\sum_i T_{i\leftarrow j}(V_i - V_j) \qquad [5]$$

Equation [5] is very useful in treating many mesoscopic systems, such as quantum dots exhibiting the phenomenon Coulomb blockade.

Justifying the Localization Picture

Several theoretical/numerical methods have been employed to obtain the various localization quantities such as L_c, ξ, E_c, $\delta\sigma$, etc.

Level Statistics

Localized eigenstates belonging to neighboring energy levels of a very large system are, in general, far away from one another and hence, they have negligible overlap and as a result, negligible level repulsion. On the contrary, extended eigenstates always overlap, and thus always exhibit level repulsion. Hence, the probability distribution of the separation of two consecutive energy levels (which can be obtained numerically) is quite different for localized and extended eigenstates. This difference is an efficient way for determining regions of localized and extended states.

Transfer Matrix Techniques

A transfer matrix, t_{12}, propagates the general solution of the wave equation from point x_1 to point x_2 along 1D and quasi-1D systems. By multiplying transfer matrices, $t_{0L} = t_{12}t_{23}\ldots t_{NL}$, the solution at one end $(x = 0)$ can be connected to the solution at the other end $(x = L)$. For disordered quasi-1D systems, the eigenvalues of t_{0L} come in pairs $(\mu_i, \bar{\mu}_i)$, such that $\mu_i \to \exp(L/L_{ci})$ and $\bar{\mu}_i \to \exp(-L/L_{ci})$ as $L \to \infty$, where $\pm L_{ci}^{-1}$ are the so-called Lyapunov exponents. The longest L_{ci}, L_c, determines the transmission coefficient, T, of the quasi-1D system of length L as $T = (-2L/L_c)$. By considering quasi-1D strips of M coupled channels arranged in a plane, or wires of square cross section with M^2 channels, $L_c(M)$ can be determined numerically; by employing a scaling property, the limit $M \to \infty$ can be taken, which determines L_c for the 2D case ($L_c = \lim L_c(M)$ as $M \to \infty$), and both L_c (when $L_c(M)$ converges as $M \to \infty$) and ξ (when $L_c(M) \sim M^2/\xi$) for the 3D case.

Scaling Approach

In the model considered in **Figure 3**, the only relevant parameter is the dimensionless ratio $w/|V_2|$. Then a cluster of p neighboring atoms of length $L_1 = p^{1/d}a$ (where a is the lattice spacing) is examined and one may attempt to define a single dimensionless ratio, Q_1, which would play for the cluster the same role as $w/|V_2|$ plays for each atom. By repeating the rescaling n times, where n eventually approaches infinity, one finds that $Q_n = (\hbar/e^2)G$, where G is the conductance of a d-dimensional "cube" of linear dimension $L = p^{n/d}a$. Since $Q_n \equiv Q(L)$ is the only relevant parameter, the occurrence or not of localization depends on whether $Q(L)$ tends to infinity or to zero respectively, as $L \to \infty$. To find this limit, the logarithmic derivative $\beta = \mathrm{d}\ln Q/\mathrm{d}\ln L$, which depends on the parameter Q, is defined. In the limit $Q \to \infty$, the classical behavior is approached, which gives that $Q \sim G \sim L^{d-2}$; hence $\beta \to d - 2$, that is, negative for $d < 2$ and positive for $d > 2$. Assuming that β is a monotonically increasing function of Q, one can conclude that in the limit $L \to \infty$, $G \to 0$ for $d < 2$. For $d > 2$, $G \to \infty$ as $L \to \infty$ (in the region $\beta > 0$), and $G \to 0$ as $L \to \infty$ (in the region $\beta < 0$).

Potential Well Analogy

More sophisticated approximate analytical approaches for the calculation of σ permit one, under certain conditions, to map the localization question to that of an effective d-dimensional potential well (in the absence of magnetic forces) of depth $V_0(E) \sim [\sigma_0(E)]^{-1}[l(E)]^{-d}$ and linear extent $a \sim l(E)$. If this effective potential well sustains a bound state of decay length $L_c(E)$, then the states at E are localized with localization length equal to $L_c(E)$. If the effective potential well exhibits a resonance in the cross section at $E_r(E)$, then the eigenstates at E are extended with a length ξ given by $E_r(E) = \hbar^2/2m\xi^2(E)$, $\xi \gg l$. Using the fact that a potential well, no matter how shallow, always sustains a bound state for $d \leqslant 2$, while it must exceed a critical value of V_0a^2 to sustain a bound state for $d = 3$, the basic picture presented before is recaptured, before including the expressions for the localization length in 1D and 2D systems. Furthermore, by employing the critical value of V_0a^2, one finds that localized states appear in 3D systems, when the product $lk \lesssim 0.85$ or $l/\lambda \lesssim 0.14$; this is the so-called Ioffe–Regel criterion.

Classical Wave Localization

In addition to the work of electronic localization presented above, the question of classical localization has received attention. This interest is due partly to the fact that classical waves, such as electromagnetic or acoustic/elastic, being subject to destructive

interference and hence to the possibility of localization, can serve as a model system for testing the theory of Anderson localization of electrons experimentally in a clean way, without the complication of strong inelastic scattering and other effects of electron–electron and electron–phonon interactions. On the other hand, it is harder to localize classical waves, mainly due to the fact that at low frequencies, the effects of disorder tend to be wiped out, whereas electrons at low energies are trapped very effectively, even by a weak random potential.

To see this, consider the simple scalar wave equation, $\nabla^2 u + (\omega^2/c^2)u = 0$, where the velocity, $c(\boldsymbol{r})$, in the medium varies from point to point between a maximum, c_M, and a minimum, c_m, value. Compare this equation with Schrödinger's equation in the presence of a potential $V(\boldsymbol{r})$ which varies between V_M and V_m. The two equations are mathematically equivalent if one makes the following correspondences:

$$\frac{\omega^2}{c_M^2} \leftrightarrow \frac{2m}{\hbar^2}(E - V_M) \qquad [6a]$$

$$\omega^2\left(\frac{1}{c_M^2} - \frac{1}{c(\boldsymbol{r})^2}\right) \leftrightarrow \frac{2m}{\hbar^2}[V(\boldsymbol{r}) - V_M] \qquad [6b]$$

Equation [6a] implies that classical waves correspond to electronic waves, but for energy above the maximum value of the potential; eqn [6b] shows that the scattering producing fluctuations in the classical wave case are multiplied by ω^2, and thus are fully ineffective for low frequencies. The conclusion is that classical waves can only be localized, if at all, at intermediate frequencies. The analogies between the classical and quantum problems indeed lead to many cross fertilizations, since solutions obtained in one field can be carried over to the other. However, considerable care has to be exercised in transforming the results of the theory of localization of electrons to the case of classical waves. The most important difference is that for classical waves, the equivalent of the particle number is not conserved. The quantity conserved here is the energy, leading to a diffusion behavior of the energy density. Another difference is that the scattering potential is frequency dependent as shown in eqn [6]. As a consequence, the energy transport velocity entering the diffusion coefficient for a strongly disordered system may be appreciably renormalized, and, as a result, diffusion coefficients can be quite small even far from the localization transition. In particular, the low values experimentally obtained for the diffusion coefficient, $D = v_E l/3$, are caused by extremely small values of the energy transport velocity, v_E, and not necessarily by the small values of l, which would signify strong localization. It is well understood that low values of v_E for classical waves are due to the Mie resonances of the scatterers. Near resonances, a lot of energy is temporarily stored inside the dielectric scatterer or, equivalently, the wave spends a lot of time (dwell time) inside the dielectric scatterer.

The outstanding problem in classical wave localization is to find the optimal conditions for its realization. It has been suggested that an intermediate frequency window of localized states separates the low-frequency extended states, characterized by Rayleigh scattering, from the high-frequency extended states, described by geometric optics. Theories based on weak scattering limit and on the coherent potential approximation (CPA) predict frequency intervals within which localization should be observed. These predictions are based on extrapolation of results obtained in the weak disorder regime. Wiersma *et al.*, in 1997, reported experiments at the near-infrared in GaAs powders, where the transmission coefficient, T, was measured in the extended, critical, and localized regimes. In the critical regime, T should vary with the inverse of the square of the sample thickness and in the localized regime, it should vary exponentially with the thickness. F Scheffold *et al.*, in 1999, argued that the results can be interpreted on the basis of the classical diffusion theory if absorption is included. One way to separate the effects of absorption and localization is to measure fluctuations of the transmission T, in addition to more conventional transport properties. This can be done easier at microwave frequencies, where one can measure both the amplitude and the phase of the electric field.

To facilitate the solution of the classical wave localization problem, it was proposed to construct first a periodic medium which, hopefully, will produce a spectral gap, and then by a disordering process, to create tails of localized states in this gap. The first part of this proposal produced much more than what one was looking for initially: it led to the creation of the so-called photonic (and phononic) band gap (PBG) artificial materials, and to the emergence of a new technological field.

See also: Disordered Magnetic Systems; Disordered Solids and Glasses, Electronic Structure of.

PACS: 71.23.−k; 71.23.An; 71.30.+h; 72.10.−d; 72.15.−v; 61.43.−j; 41.20.Jb; 05.60.Gg

Further Reading

Abrahams E, Kravchenko S, and Sarachik MP (2001) Metallic behaviour and related phenomena in two dimensions. *Reviews of Modern Physics* 73: 251.

Brandes T and Kettemann S (eds.) (2003) *Anderson Localization and its Ramifications*. Berlin: Springer.
Datta S (1995) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.
Economou EN (1983) *Green's Functions in Quantum Physics*, 2nd edn. Berlin: Springer.
Imry Y (2002) *Introduction to Mesoscopic Physics*, 2nd edn. Oxford: Oxford University Press.
Joannopoulos JD, Meade RD, and Winn JN (1995) *Photonic Crystals Molding the Flow of Light*. Princeton, NJ: Princeton University Press.
John S (1997) Frozen light. *Nature* 390: 661–662.
Kushwaha MS (1999) Bandgap engineering in phononic crystals. *Recent Research Developments in Applied Physics*. 2: 743–855.
Lifshits IM, Gredelkul SA, and Pastur LA (1998) *Introduction to the Theory of Disordered Systems*. New York: Wiley.
Sakoda K (2004) *Optical Properties of Photonic Crystals*, 2nd edn. Berlin: Springer.
Scheffold F, Lenke R, Tweer D, and Maret G (1999) Localization or classical diffusion of light? *Nature* 398: 206–207.
Soukoulis CM (ed.) (2001) *Photonic Crystals and Light Localization in the 21st Century*. Dordrecht: Kluwer.
Taylor PL and Heinonen O (2002) *A Quantum Approach to Condensed Matter Physics*, pp. 315–336. Cambridge: Cambridge University Press.
Wiersma DS, Bartolini P, Lagendijk A, and Righini R (1997) Localization of light in a disordered medium. *Nature* 390: 671–673.

Disordered Magnetic Systems

P Nordblad, Uppsala University, Uppsala, Sweden

Introduction

Disordered magnetic systems, when defined widely, comprises the whole field of applied magnetic materials – ranging from the extremely soft permalloy to the hardest permanent magnet as well as the physics of model systems of strongly disordered magnetic materials, spin glasses, and random-field Ising systems.

Magnetic ordering requires that the electron system forms local or itinerant magnetic moments and that there is interaction between these moments. Such interacting magnetic moments will eventually, as the temperature is lowered, become correlated and in certain cases form a long-ranged ordered phase at a critical temperature. The simplest magnetic order in crystalline systems arises when the leading interaction is ferromagnetic and there is only one type of magnetic ions in the material. The low-temperature phase then shows a regular pattern with the magnetic moments all aligned in parallel. Magnetic disorder in an ideal system of this type is enforced by thermal energy, and above the critical ordering temperature, a random paramagnetic phase forms. However, intrinsic quenched disorder, due to lattice dislocations, impurities, and other defects, always appears in a real magnetic system, also in the magnetically ordered phase. This kind of disorder is the key to the macroscopic magnetic properties of an ordered ferromagnet. The disorder and the microstructure of the system then determine if a weakly anisotropic magnet becomes a good soft ferromagnet or if a strongly anisotropic system attains applicable permanent magnetic properties. It is thus disorder, on nano- and micrometer length scales, that governs the quality of applied magnetic materials, whether soft or hard.

Models

A random distribution of the atomic magnetic elements in alloys and compounds introduces strong disorder and deviations from a regular interaction pattern. The interaction between the magnetic constituents becomes random in size and, in some cases, also in sign. This causes the appearance of new magnetic phases and phenomena. An interacting magnetic Ising system can be described by its spin Hamiltonian

$$H = -\sum_{i,j} J_{i,j} S_i S_j \qquad [1]$$

where $J_{i,j}$ is the exchange interaction strength between spin S_i and spin S_j: spins that in an Ising system point either up or down. A random-exchange ferromagnet is described if all $J_{i,j}$ are positive but of different magnitude, and a spin glass is described if the sign of $J_{i,j}$ varies but the mean value is zero. Equation [1] can be used to describe systems of any dimensionality and any range of interaction. Another important class of disordered magnetic Ising systems is the random field system

$$H = -\sum_{i,j} J_{i,j} S_i S_j + \sum_i h_i S_i \qquad [2]$$

where $J_{i,j}$ are positive interaction constants and h_i is a local field of random size and sign with zero mean that acts on spin number S_i. This system does not have an experimental realization, but a dilute Ising antiferromagnet in an applied homogeneous magnetic field has been shown to describe an equivalent problem:

$$H = -\sum_{i,j} J \varepsilon_i \varepsilon_j S_i S_j + H_a \sum_i \varepsilon_i S_i \qquad [3]$$

where J is a negative exchange interaction constant, and ε_i is 1 if site i is occupied by a magnetic ion, and 0 if the site contains a nonmagnetic element. H_a represents a homogeneous applied magnetic field, and the equation shows that the random field is proportional to this field and thus arbitrarily tunable.

Spin Glasses

An alloy with magnetic ions randomly dispersed in a nonmagnetic host metal forms an archetypical spin glass material at low magnetic ion concentrations. The interaction in the system is random in size and sign due to the Ruderman–Kittel–Kasaya–Yoshida (RKKY) nature of the interaction in a metal. The RKKY-interaction is mediated by the conduction electrons, and is oscillatory in size and decays with the cube of the distance between the magnetic constituents. The signature and the initial cause of interest in this type of dilute magnetic alloys was the observation of a cusp in the in-phase component of the AC-susceptibility by Cannella and Mydosh in the early 1970s. This finding was a clear indication of a phase transition to a new kind of magnetic order, and this phase, later named the spin glass phase, has been found to have surprising magnetic properties and has created a theoretical problem of much greater complexity than what only the occurrence of a cusp in the AC susceptibility promises. Spin glasses provide a useful introduction to the physics of strongly disordered magnetic systems, in general, and are therefore described in some detail below.

A spin glass problem can be defined using eqn [1] by setting the spatial dimension of the system as well as the distribution and range of the interaction $J_{i,j}$. In the Sherrington–Kirkpatrick (SK) model, the sum is taken over all pairs (i,j) of N spins, and a Gaussian distribution of $J_{i,j}$ with zero mean is used. The famous Parisi solution of this problem using replica symmetry breaking (RSB) forms one school of models describing spin glasses: the RSB or mean-field picture. An alternative model to describe finite dimensional systems is the droplet picture, which is based on the scaling hypothesis and yields a spin glass with properties, remarkably different in parts than the mean-field predictions. A main difference concerns the nature of the spin glass in a magnetic field, where the droplet picture prescribes that there is no equilibrium spin glass phase, whereas the RSB model yields a phase line in the field-temperature (H–T) plane (the Almeida–Thouless line) that separates the paramagnetic from the spin glass phase. Alternatively, in zero magnetic fields, both models anticipate a low-temperature phase with relaxation processes that lasts over exceedingly wide timescales and a nonstationary nature of a system that has been quenched to a temperature below its spin glass temperature (T_g).

There is an extensive literature on the static and dynamic magnetic properties of three-dimensional spin glass materials as observed in experiments and Monte Carlo (MC) simulations on three-dimensional Ising spin glass models. A clear empirical picture as to the magnetic properties of spin glasses can be extracted from this literature. The behavior is, however, complex and has allowed interpretations in terms of both the RSB and the droplet scaling model or combinations of the two pictures.

It is firmly established that there is a second-order phase transition to a low-temperature spin glass phase in the zero field. This transition is evidenced from the dynamics through analyses of the critical slowing down on approaching the phase transition temperature (spin glass temperature) and from the static susceptibility through a divergence of the nonlinear susceptibility on approaching T_g. In **Figure 1**, AC-susceptibility data at different frequencies measured on the Ising spin glass $Fe_{0.5}Mn_{0.5}TiO_3$ are shown; the original spin glass signature, a (frequency-dependent) cusp in the AC susceptibility is clearly seen. The temperature dependence of the relaxation time τ, that can be derived from the frequency dependence of the cusp (or freezing) temperature T_f, obeys critical slowing down and implies a phase transition with a relaxation time that diverges according to critical slowing down: $\tau = \tau_0 t^{-z\nu}$, where τ_0 is an atomic relaxation time $\sim 10^{-13}$ s, $z\nu \approx 10$ is the dynamic critical exponent, and t is the reduced

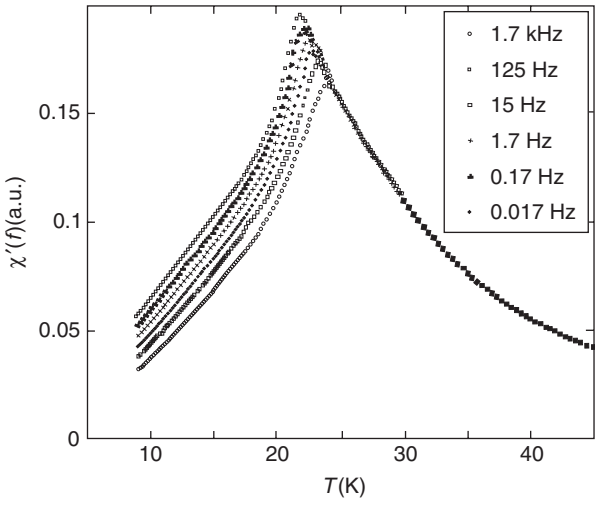


Figure 1 The AC susceptibility, $\chi'(f,T)$, measured along the c-axis of the Ising spin glass, $Fe_{0.5}Mn_{0.5}TiO_3$. The frequency (f) of the weak AC field is indicated in the figure. (Data from the theses of Mattsson J (1994) Uppsala University and Jonsson T (1998) Uppsala University.)

temperature, $t = T/T_{g-1}$. Critical slowing down and the existence of a low-temperature spin glass phase have been convincingly established from experiments and simulations on three-dimensional Ising spin glasses. This is one side of the phase-transition problem; a key question, which distinguishes the RSB- and droplet-model predictions, is whether the phase transition persists in an applied magnetic field or not. This seemingly simple question has not been easy to answer from either experiments or simulations. However, analyses of the in-field dynamics of the Ising spin glass $Fe_{0.5}Mn_{0.5}TiO_3$ show that the slowing down of the relaxation time is significantly altered in an applied field compared to the zero field and suggest that there is no phase transition in a magnetic field in a three-dimensional Ising spin glass.

The response of a spin glass to the application of a weak magnetic field in the spin glass phase extends from the atomic relaxation time (τ_0) to exceedingly long times well beyond the timescale of any experimental probe. That is, in common with other glassy systems, the relaxation times are exceedingly long in the low-temperature phase. Also, in common with other glasses, the low-temperature spin glass falls out of equilibrium as the system is cooled through the spin glass transition temperature. Once the system is kept at constant temperature in the low-temperature phase, it spontaneously and continuously reorganizes the spin structure – the system ages! Aging has a very characteristic influence on the response function of spin glasses. The relaxation of the magnetization $M(t)$ attains an inflection point at an observation time (t) of the order of the wait time (t_w) at constant temperature before the weak magnetic field was applied in a plot of $M(t)$ versus $\log(t)$. **Figure 2a** shows the relaxation of the magnetization $M(t)$ of an Ag(11 at.% Mn) spin glass measured after the different wait times indicated in the figure; also plotted in the figure is the relaxation rate, $S(t) = \partial M(t)/h\partial \log(t)$, that

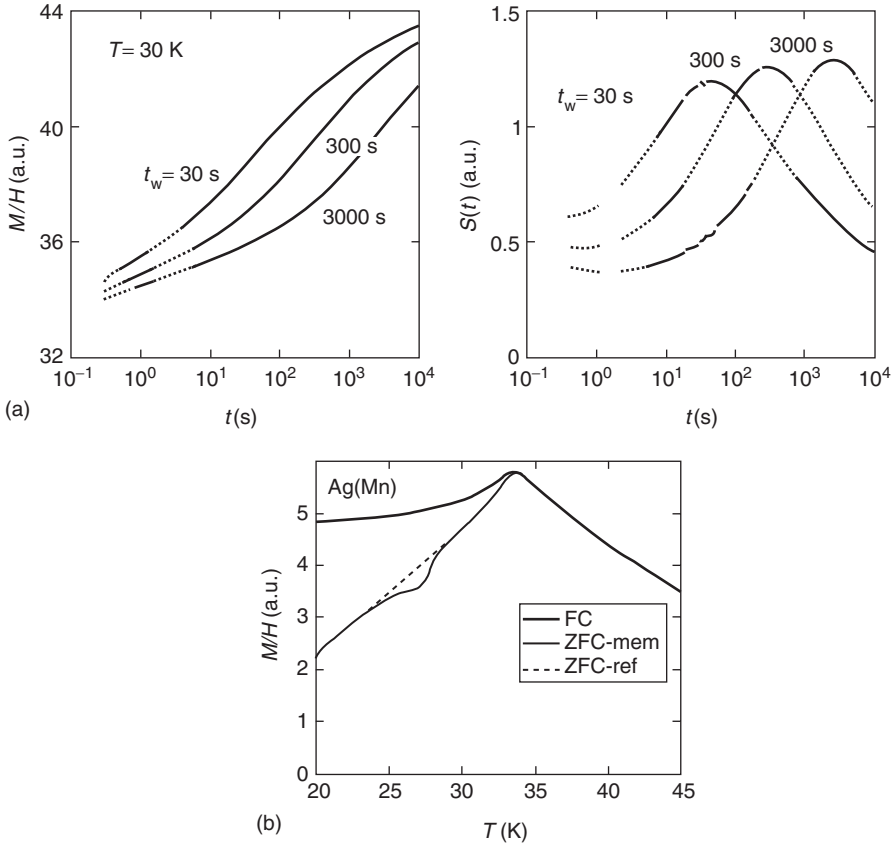


Figure 2 (a) The left figure shows the relaxation of the ZFC magnetization $M(t)$ of the spin glass Ag(11 at.% Mn) at $T = 30$ K. The magnetization has been measured after different wait times at the measurement temperature as indicated in the figure. The right figure shows the relaxation rate $S(t) = \partial M(t)/h\partial \log(t)$ of the magnetization relaxation curves. $T_g \approx 32$ K, $h = 40$ A m^{-1}. (Data from the thesis of Jönsson P (2002) Uppsala University.) (b) ZFC and FC magnetization vs. temperature ($M(T)/H$) for the spin glass Ag(11 at.% Mn). The curve labeled ZFC-ref shows the ZFC-magnetization measured after continuous cooling to low temperature, whereas the curve labeled ZFC-mem shows the corresponding curve measured after the sample has been kept at 27 K for 10^4 s during otherwise continuous cooling. $H = 8$ A m^{-1}. (Data from the thesis of Mathieu R (2002) Uppsala University.)

is, the local slope of the relaxation curves; this quantity exhibits a pronounced maximum at an observation time equal to the wait time. Results from MC simulations of the dynamics of three-dimensional Ising spin glass models yield similar results and confirm the existence of an aging phenomenon. The main conclusion from the aging (nonequilibrium) behavior is that the spin glass continuously evolves toward an equilibrium state at constant temperature in the spin glass phase and that this equilibrium state is different at different temperatures (the system appears chaotic).

An additional and remarkable feature of the nonequilibrium spin glass phase is that it incorporates a memory phenomenon. This phenomenon is most simply revealed by measurements according to a standard protocol to investigate the temperature dependence and irreversibility of the magnetization through the phase-transition temperature of a magnetic system. The zero-field-cooled (ZFC) and field-cooled (FC) magnetization, which are measured by cooling the sample in zero field, apply a weak field at a low temperature and record the magnetization (ZFC) while continuously heating the system to a temperature above T_g and then recool the magnetization in the applied field while recording the magnetization (FC). Such magnetization curves are shown for the archetypical spin glass Ag(11 at.% Mn) in **Figure 2b**. In addition, the figure shows a magnetization curve where the cooling of the sample has been interrupted for some hours at 27 K, and after reaching the low temperature, the magnetization was recorded on continuous heating in the same way as the ZFC-reference curve. This ZFC-mem curve shows a pronounced dip at the temperature for the stop. The system keeps a memory of the cooling history – and the memory is recalled when recording the ZFC magnetization on reheating the spin glass. The observed dip in the ZFC-mem curve is reminiscent of the aging that occurs in the system when it is kept at constant temperature, which results in a slower relaxation with increasing age of the system as was illustrated in **Figure 2a**. The memory-aging behavior are manifestations of some crucial concepts – aging, rejuvenation, and chaos – that characterize the spin glass phase and are of key importance for the modeling of spin glasses.

Spin glass behavior is not only confined to the model systems with atomic magnetic moments described above; there are several other disordered magnetic systems that have been found to show spin-glass-like magnetic behavior, for example, interacting magnetic nanoparticle systems, granular magnetic systems, some manganites and coboltites, and certain high-temperature superconductors.

Re-Entrant Spin Glasses

A spin glass is formed when there is disorder and frustration due to an equal amount of positive and negative interactions in the system, that is, the mean value of the interactions is close to zero. Such systems can simply be modeled by using an equal amount of interactions of the same magnitude $\pm J$ in the system or by using a Gaussian distribution of the interactions with zero mean. If the random interactions have a nonzero mean value, ferro $(+)$ or antiferromagnetic $(-)$, the system can show re-entrant behavior. The system first orders according to the dominant interaction at a high temperature and forms a long-ranged ordered random ferromagnetic or antiferromagnetic phase. However, on further lowering the temperature, the system seems to re-enter into a disordered phase, a spin glass phase. Re-entrant systems are quite common; they appear, for example, at higher concentration of magnetic atoms in dilute magnetic alloys and amorphous metals. Re-entrant ferromagnets show some remarkable properties: both the re-entrant spin glass phase and the long-range ordered phase show an aging phenomenon, and on cooling into the low-temperature spin glass phase, the ferromagnetic long-range order seems to persist although macroscopic AC-susceptibility data indicate critical slowing down and a true re-entrance into a spin glass phase. Many unanswered questions remain as to the physics of re-entrant magnetic systems.

The Random-Field Ising Model

A random-field Ising system is described by eqn [2]. The properties of that Hamiltonian have been extensively studied theoretically and by simulations and have been found to show a behavior that is very different from a random exchange or pure ferromagnetic system with respect to critical exponents, critical dimensions, and magnetic properties. Experimentally, there is no real material that provides local random fields as prescribed by the random-field Ising model. However, the random-field Hamiltonian is mapped by a dilute antiferromagnet in a homogeneous magnetic field: eqn [3] above – and such materials are readily fabricated in single crystalline form and provide systems that allow a continuously tunable strength of the random field only by varying the strength of an applied homogeneous magnetic field. A major part of the experimental work on three-dimensional random-field magnetic systems has been performed on the model Ising antiferromagnetic system FeF_2 diluted with nonmagnetic zinc atoms: $Fe_{1-x}Zn_xF_2$. The study of random-field systems is made more complicated by nonequilibrium effects that

appear when the system is cooled to low temperatures and hinder the system to reach thermodynamic equilibrium. A quite consistent picture as to some thermodynamic equilibrium properties of random-field systems of different spatial dimensionality has, however, been empirically attained.

The two-dimensional Ising system forms a low-temperature ordered phase via a second-order phase transition, the only second-order phase transition that has been given an exact solution (Onsager). This transition is, both in theory and experiments, found to be destroyed by finite random fields, that is, the lower critical dimension of the random-field Ising system is 2 or larger. In 3d random-field Ising systems, after many years of experimental and theoretical controversies, it has been confirmed that the second-order phase transition survives in random fields, and both theory and experiments agree that the lower critical dimension is equal to 2. The critical exponents describing the asymptotic behavior at the phase transition are significantly different from those of the pure three-dimensional Ising and also different from the random-exchange Ising model values.

Nonequilibrium effects appear as the random-field system is cooled into the ordered phase; these effects give rise to domain states and relaxation phenomena that are not yet fully satisfactorily described by theory and experiments. However, a quite high level of empirical understanding of the physics of random-field Ising model systems has been realized by combining results from macroscopic (susceptibility, magnetization, and specific heat) measurements with those from microscopic probes (neutron scattering and X-ray scattering experiments).

Amorphous Magnetic Systems

Crystallographically disordered materials with an exchange interaction strength that is dependent on the distance between the magnetic entities would naturally form disordered magnetic systems that mimic a random exchange or more complicated model system. A realization of such systems is amorphous metals fabricated by rapid quenching methods and stabilized by introducing glass-forming elements in the melt. Magnetic metallic glasses were first fabricated in the late 1960s and early 1970s and have provided excellent model systems for disordered magnetic systems and the influence of random exchange on the magnetic properties. There are outstanding possibilities to tailor materials using the rapid-quenching method, since almost any composition of different metals can be mixed without a need to consider phase segregation or consult complicated phase diagrams. It was earlier found that amorphous magnetic metals can form a variety of magnetic structures ranging from simple ferromagnets to complex ferrimagnetic spiral structures and spin glasses. However, a huge challenge and obstacle for the theoretical description of these metals is, of course, the lack of a periodic lattice as a backbone for the description of the electronic structure.

Amorphous metals are trapped in a metastable state and the structure has to be stabilized against crystallization, which severely limits the possible production processes. However, an important part of the interest in amorphous magnetic metals is due to their potential as magnetic materials in many applications. The possibility of fabricating materials with marginal magnetocrystalline anisotropy, very low magnetostriction, and high resistivity compared to their crystalline counterparts promises materials with excellent soft magnetic properties for medium- and high-frequency applications. Soft amorphous magnetic metal systems have also found important applications in, for example, small-scale transformers, motors, and generators. The rapid-quenching method does not only allow amorphous metals to form, it also yields possibilities of fabricating nanocrystalline systems. These systems, at certain microstructure sizes, form extremely hard magnetic materials with exceedingly high coercivity (which are useful building blocks for good permanent magnets), whereas under other compositions and nanocrystalline sizes, provide materials with excellent soft magnetic properties.

Magnetic Materials

Disorder is the physical parameter that governs extrinsic magnetic properties, such as coercivity, remanence, and energy product, of an intrinsically ferromagnetic or ferrimagnetic material. The producer of magnetic materials is thus forced to understand disorder to be able to effectively design and fabricate both soft and hard magnetic materials. It is, in fact, the ability to master disorder and find new material combinations that underlies the fast improvement of the technical performance that magnetic materials have been subjected to during the last few decades. Maybe, the prime example of such progress is found in the rapidly increasing areal storage density in magnetic information-storage media.

Summary

Theoretical and experimental studies of disordered magnetic model systems, such as the Ising spin glass and the random-field Ising model, have provided new insights and understanding of the influence of disorder on the properties of not only magnetic but also

other physical systems. In addition, the models and methods that have been developed in this research have found applications in many different fields of science that deal with complex and disordered systems. Examples are found in the physics of ordinary glasses and in optimization problems as well as in studies of stock markets and descriptions of social systems.

See also: Ferromagnetism; Magnetic Interactions; Magnetic Order; Paramagnetism.

PACS: 75.10.Nr; 75.30.Et; 75.30.Gw; 75.40.Cx; 75.40.Gb; 75.40.Mg; 75.50.Lk; 75.50.Tt; 75.50.Ww

Further Reading

Belanger DP (2000) Experimental characterization of the Ising model in disordered antiferromagnets. *Brazilian Journal of Physics* 30: 682–692.

Belanger DP and Young AP (1991) The random field Ising model. *Journal of Magnetism and Magnetic Materials* 100: 272–291.

Fisher KH and Hertz JA (1991) *Spin Glasses*. Cambridge: Cambridge University Press.

Hansen P (1991) Magnetic amorphous alloys. In: Buschow KHJ (ed.) *Handbook of Magnetic Materials*, vol. 16, pp. 292–452. Amsterdam: Elsevier.

Hasegawa R (1991) Amorphous magnetic materials – a history. *Journal of Magnetism and Magnetic Materials* 100: 1–12.

Herzer G (1997) Nanocrystalline soft magnetic alloys. In: Buschow KHJ (ed.) *Handbook of Magnetic Materials*, vol. 10, pp. 415–462. Amsterdam: Elsevier.

Luborsky FE (1980) Amorphous ferromagnets. In: Wohlfarth EP (ed.) *Ferromagnetic Materials*, vol. 1, pp. 451–530. Amsterdam: Elsevier.

Mezard M, Parisi G, and Virasoro MA (1987) *Spin Glass Theory and Beyond*. Singapore: World Scientific.

Mydosh JA (1993) *Spin Glasses: An Experimental Introduction*. London: Taylor and Francis.

Young AP (ed.) (1998) *Spin Glasses and Random Fields*. Singapore: World Scientific.

Nomenclature

f	frequency (Hz)
H	magnetic field ($\mathrm{A\,m^{-1}}$)
J	exchange interaction
M	magnetization ($\mathrm{A\,m^{-1}}$)
t	reduced temperature, time (s)
t_w	wait time (s)
T_f	freezing temperature (K)
T	temperature (K)
T_g	spin glass temperature (K)
τ	relaxation time (s)
χ	magnetic susceptibility ($\mathrm{H\,m^{-1}}$)

Disordered Solids and Glasses, Electronic Structure of

S R Elliott, University of Cambridge, Cambridge, UK

Introduction

In this article, the term "electronic structure" is taken to mean a description of the details of the solution, in its totality, of the Schrödinger equation for electrons in solids. Two quantities characterize the Schrödinger eigenequation solution, namely the eigenvalue (or electron energy, E) and eigenvector (or electron wave function, $\psi(\boldsymbol{r})$).

In the simple case of electron states in perfect single-crystalline materials, that is, those considered in most textbooks on the electronic theory of solids, the term electronic structure is often restricted to a description solely of the electronic energy since, for translationally periodic systems, the electronic states themselves always have the same character and generic functional form. The electronic wave function is always spatially extended throughout the volume of the crystalline sample, and has the functional form of the "Bloch" function, that is, a product of a spatially extended plane-wave state and a function, $u_{\boldsymbol{k}}(\boldsymbol{r})$, having the periodicity of the underlying crystal lattice:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \qquad [1]$$

where $\boldsymbol{k}$ is the (quantized) wave vector of the plane-wave component, and is used as a label for the eigenstate. (This is not the case for disordered systems, as will be seen.)

Thus, for perfect single crystals, the electronic structure, characteristic of a particular material, is taken to refer to either the four-dimensional (4D) quantity, the electronic band structure, $E(\boldsymbol{k})$, or its 2D derivative, the electronic density of states (EDOS), $g(E)$, where generally

$$g(E) = \sum_i \delta(E - E_i) \qquad [2]$$

The EDOS in this representation is simply a sum of Dirac delta functions at all allowed, quantized electron-energy eigenvalues, E_i; it is the number of electron states at a particular energy in unit energy interval, and $g(E)$, therefore, has the units of inverse

energy. Alternatively, for a single crystal, the EDOS can be obtained as a surface integral in $\boldsymbol{k}$-space over the Brillouin zone of a constant-energy surface S_E of the electronic band structure, that is,

$$D(E) = \frac{2}{(2\pi)^3} \int_{E(\boldsymbol{k})=\text{const}} \frac{\mathrm{d}S_\mathrm{E}}{\nabla_k E} \qquad [3]$$

Here, $D(E) = g(E)/V$, where V is the volume of the sample, and hence this form of the EDOS has units of inverse energy times volume.

One special feature of the EDOS of crystalline solids is the presence of van Hove singularities, that is, sharp features in the density of states corresponding to flat regions of the band structure, for which $\nabla_k E = 0$ (cf. eqn [3]), that is, regions where there are a large number of electron states with different $\boldsymbol{k}$-values in a very small energy interval.

It will be seen that the introduction of disorder in a solid can have dramatic effects on the above simple picture of electron states in single-crystal materials.

Nature of Disorder

The translational periodicity characteristic of perfect single crystals can be disrupted in a number of ways. Although, in a given disordered material, a glass for instance, several different types of disorder may occur together, nevertheless it is instructive to consider such possible forms of disorder separately. Perhaps, the simplest form is substitutional disorder in which a different type of atom is substituted for another in the crystalline lattice of the host material. The simplest example of substitutional disorder is the inclusion of one type of atom in the structure consisting entirely of another type, for example, the electrical dopants, *n*-type phosphorus or *p*-type boron, in crystalline silicon (**Figure 1a**). A related substitutional defect is the so-called antisite defect, for example, in a binary crystalline compound, AB, where only heteropolar (A–B) bonding occurs in the completely ordered case. The replacement of an A atom, for example, by a B atom leads to the formation of homopolar (B–B in this case) nearest-neighbor bonds (**Figure 1b**). A related defect is the atomic vacancy, where an atom is missing at a lattice-related site (**Figure 1c**). This type of disorder involves point (or zero-dimensional) defects.

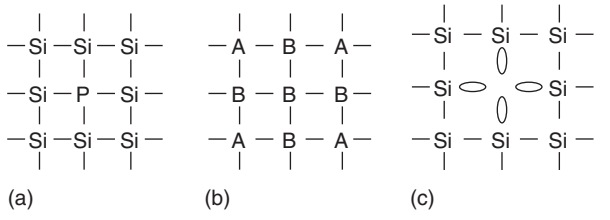


Figure 1 Schematic illustration of different types of substitutional disorder associated with point defects in solid (crystalline) materials: (a) a phosphorus donor atom in crystalline silicon; (b) an antisite defect in a binary crystal, AB; and (c) an atomic vacancy (e.g., in crystalline silicon, showing four dangling-bond orbitals pointing into the vacancy).

A somewhat related 2D defect is the so-called antiphase boundary in a binary crystal, AB, where an entire plane of atoms is missing (i.e., a stacking fault exists): in this case, like atoms face each other across the antiphase boundary (**Figure 2**).

Another type of disorder prevalent in crystalline materials is positional disorder, characteristic of vibrational motion: atoms are instantaneously displaced randomly from their equilibrium, lattice-related crystallographic sites, according to a Gaussian (normal) distribution in the case of thermal vibrational disorder (**Figure 3**). Positional disorder in crystals due to vibrational motion is dynamic in nature: the time-averaged atomic positions are simply the equilibrium, lattice-related sites. However, for the case of noncrystalline (e.g., amorphous or glassy) materials, the positional disorder is static in nature, that is, the equilibrium atomic positions do not correspond to a crystalline lattice.

What really distinguishes an amorphous material from its crystalline counterpart, however, is the presence of topological disorder, for example, involving the presence of noncrystalline ring orders in (covalent) network materials (**Figure 4**). In this case, the disorder is more profound, and there is no simple displacement relationship between noncrystalline and crystalline structures, as there is for positional disorder, for example.

The presence of topological disorder also permits the existence of other types of defective structures,

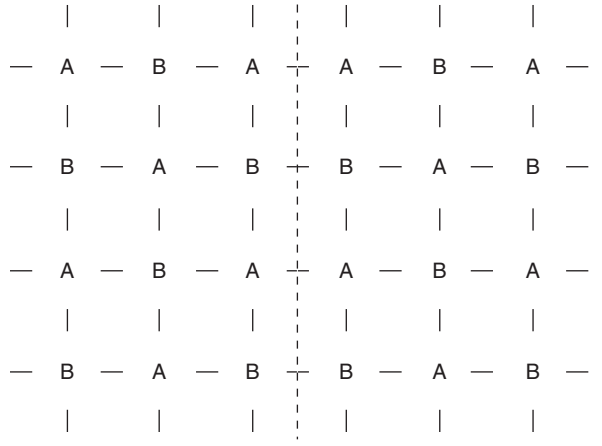


Figure 2 Schematic illustration of a two-dimensional stacking fault (missing layer) or antiphase boundary (dashed line) in a binary crystal, AB.

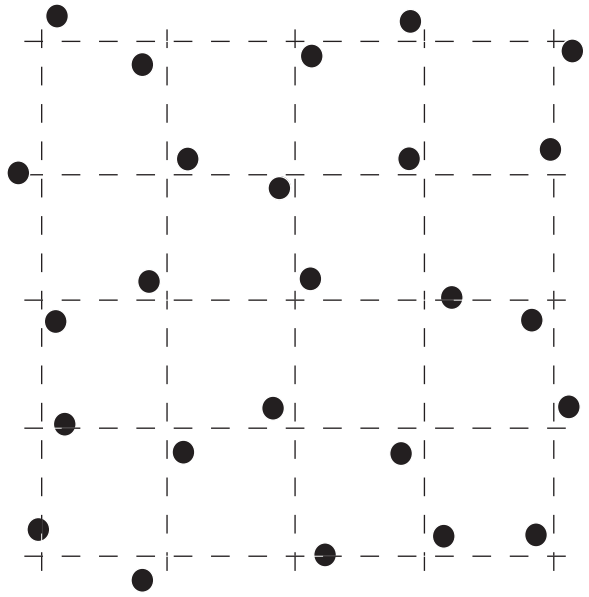


Figure 3 Schematic illustration of positional (e.g., vibrational) disorder.

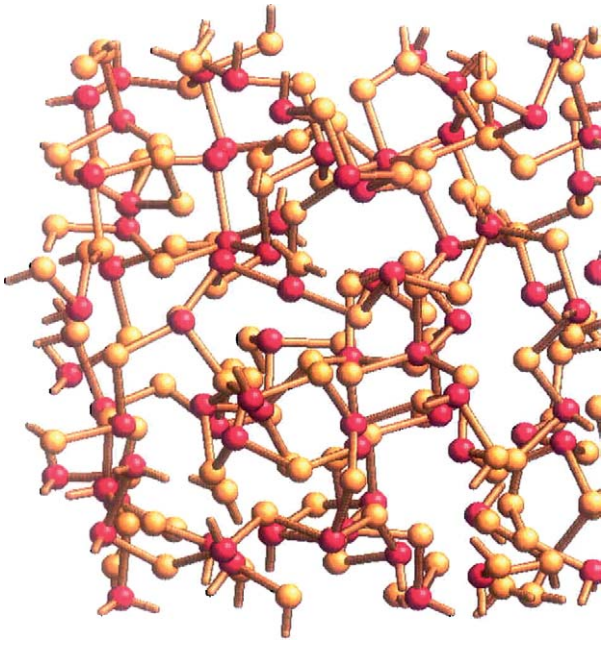


Figure 4 Illustration of topological disorder in a computer-simulated model of amorphous As_2S_3, created by *ab initio* molecular dynamics: As atoms are colored pink and S atoms are yellow. (S Simdyankin and SR Elliott, unpublished data.)

unique to the amorphous phase and not found in the crystalline state. For instance, in materials with strongly directed covalent bonds, isolated dangling-bond (under-coordinated) coordination defects can occur instead of the clusters of dangling bonds that always occur at the site of atomic vacancies in crystals of such materials, for example, the four (unreconstructed) dangling bonds pointing into the vacancy in the tetrahedral crystalline material, c-Si (**Figure 1c**). In addition, isolated homopolar-bond defects can occur in amorphous compound materials (where in the stoichiometric composition, only heteropolar bonds would otherwise be expected), instead of the clusters of homopolar-bond defects that exist around an antisite defect (**Figure 1b**) or along an antiphase boundary (**Figure 2**) in the corresponding crystalline materials.

Effects of Disorder on the Electronic Structure

General Aspects

In the case of a perfect crystal, the allowed values of the quantized electron wave vector, $\boldsymbol{k}$, are set by the periodicity of the lattice: for example, for a 1D crystal with a unit-cell parameter a, and with N atoms in the sample, there are correspondingly N discrete $\boldsymbol{k}$-states, equally spread in $k(=|\boldsymbol{k}|)$ in the first Brillouin zone, that is, for the range of $\boldsymbol{k}$-values $-\pi/a < k < \pi/a$. Each electron eigenvalue (energy) is associated with a particular $\boldsymbol{k}$-value, which shows that the electronic band structure, $E(\boldsymbol{k})$, is a well-defined quantity (see **Figure 5**).

The presence of a single defect (or a surface) is sufficient to break the infinitely ranged translational invariance characteristic of ideal single crystals. In the case of a disordered material with no translational periodicity, there is no crystalline unit cell and hence the electron wave vector becomes an ill-defined quantity and no longer is useful as a label for electron states. Although at low electron energies corresponding to small k (the electron wavelength being long compared to characteristic atomic-structural distances, for example, the nearest-neighbor bond length), the uncertainty in k is rather small, it rapidly increases with increasing energy (**Figure 5**). In other words, the electron wave function for a disordered material can no longer be expressed in terms of a single plane wave, characterized by one particular $\boldsymbol{k}$-value, as in the case for Bloch states in single-crystalline materials (eqn [1]). Instead, in a disordered material, electronic wave functions can be regarded as comprising an increasing number of Fourier components with increasing energy.

In the light of the above discussion, evidently the electronic band structure, $E(\boldsymbol{k})$, is not an appropriate description of electron states for disordered materials. However, since the electron energy remains a well-defined quantity, even in the presence of strong disorder, the EDOS (as defined as a sum of delta functions by eqn [2], but not in terms of the band

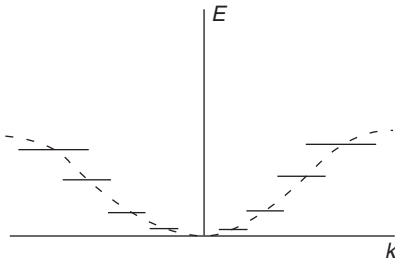


Figure 5 Schematic illustration of a one-dimensional electronic band structure for a 1D crystal (dashed line) and an amorphous solid (solid lines, illustrating the uncertainty in electron wave vector $\boldsymbol{k}$ for a given eigenenergy).

structure, eqn [3]) remains a valid form of description for the electron states of disordered materials.

The Electronic Density of States

The effect of disorder on the EDOS is varied: some changes are striking, others are more subtle. Perhaps, the most striking effect of the loss of translational symmetry is in the destruction (broadening) of the van Hove singularities in the EDOS that are a signature of the presence of flat electron-energy surfaces in $\boldsymbol{k}$-space, characteristic of crystals (see **Figure 6**).

However, van Hove singularities apart, the EDOS of many amorphous materials is often similar to that of their crystalline counterparts (see **Figure 6** for the case of silicon). This is particularly true for covalently bonded semiconductors/insulators, where the similarity in the electronic structure arises from marked similarities in the short-range atomic structural order between the two phases. In the case of amorphous and crystalline silicon, for example, the short-range structural motif is the $SiSi_4$ tetrahedron in both cases: the average nearest-neighbor bond length and bond angle are practically the same, although the amorphous phase is characterized by having an $\sim \pm 10\%$ variation in bond angle from the average tetrahedral value, $\theta = 109°28'$. (The translational periodicity of the crystalline phase is destroyed in the amorphous material, not so much by this bond-angle disorder as by the disorder in dihedral angles, that is, the angle of twist between pairs of corner-sharing connected tetrahedra.)

The reason for the overall similarity in EDOS between amorphous and crystalline states, for example, of silicon, can be seen from a simple tight-binding description of the origin of the electron states, and the similarity in the short-range order in both cases. Only two types of local electronic interactions are assumed to be operative, namely a nearest-neighbor intersite interaction, V_2, between orbitals emanating from pairs of nearest-neighbor atoms and forming

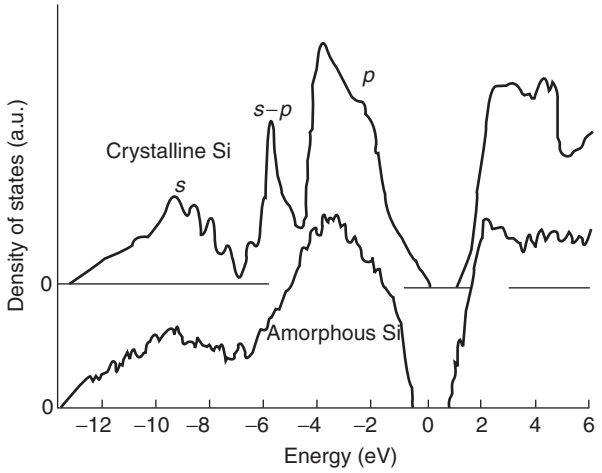


Figure 6 Calculated EDOS for amorphous and crystalline silicon, for a 216 atom model. (Reproduced with permission from Street RA (1991) *Hydrogenated Amorphous Silicon*. Cambridge: Cambridge University Press.)

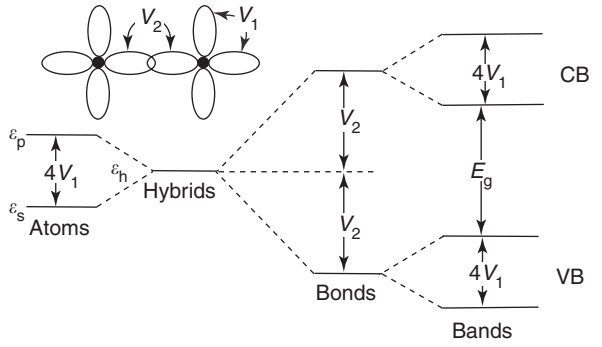


Figure 7 Illustration of the origin of the bandgap between valence and conduction bands for a tetrahedrally coordinated material, such as silicon. The successive sections show the atomic *s*- and *p*-levels, the sp^3-hybrid level, the bonding–antibonding splitting resulting from the intersite electron interaction, V_2, and the formation of bands resulting from the intrasite interaction, V_1. The origins of the interactions V_1 and V_2 are shown in the inset.

common bonds, and an intrasite interaction, V_1 between orbitals emanating from the same atom (**Figure 7**). In the case of silicon, for example, atomic *s*- and *p*-orbitals hybridize to give the sp^3-hybrid basis orbitals. The intersite interaction, V_2, is responsible for a gap of magnitude $2V_2$ occurring between occupied and unoccupied electron states (i.e., the bonding–antibonding splitting). The intrasite interaction, V_1, in this model, is responsible for the broadening of bonding and antibonding levels into, respectively, the valence and conduction bands, each of width $4V_1$ (**Figure 7**), since it couples orbitals together in such a way that, together with the intersite interaction, an electron may hop from any given orbital on one atom to any other orbital on any other atom, thereby becoming delocalized (extended)

throughout the sample volume. It should be noted that both V_1 and V_2 are negative quantities. In this model, a bandgap occurs between filled valence and empty conduction bands as long as the bonding–antibonding splitting is sufficiently large to overcome the opposing effect of the band broadening, that is, for

$$|V_2| > 2|V_1| \qquad [4]$$

and the bandgap, therefore, has the value

$$E_g = 2|V_2 - 2V_1| \qquad [5]$$

Since the average nearest-neighbor bond lengths and angles in amorphous and crystalline silicon are very similar, the magnitude of the electron-hopping interactions V_1 and V_2 are also expected to be very similar, and hence, for example, the width and shape of the valence band, and the magnitude of the bandgap between valence and conduction bands, is expected to be very similar for the two phases (see **Figure 6** for the case of Si).

However, subtle changes in the EDOS can be caused by disorder. Since the bandwidth in this tight-binding picture is governed by the intrasite interaction, V_1, and this interaction (as well as the degree of orbital hybridization) depends on the angle, θ, between (hybrid) orbitals at a site (**Figure 7**), the fluctuations in θ will lead correspondingly to variations in the bandwidth, that is, a tailing of states at the band edges into the gap region (see **Figure 8**). It appears, on the basis of experimental data, that these band tails in amorphous semiconductors generally have an exponential energy dependence:

$$g(E) = \frac{N_0}{E_0} \exp(E/E_0) \qquad [6]$$

where N_0 is the total number of tail states, and E_0 is the energy decay parameter.

Substitutional disorder, for example, phosphorus donor dopants in silicon, that is, Si:P (**Figure 1a**), also gives rise to electron states lying outside the bands of the perfect crystal; in the case of Si:P, these are donor levels just below the conduction-band edge. (Acceptor states associated with substitutional acceptors, for example, boron in silicon, Si:B, correspondingly lie just above the valence-band edge.) Substitutional defects comprising atoms having a greater electronegativity

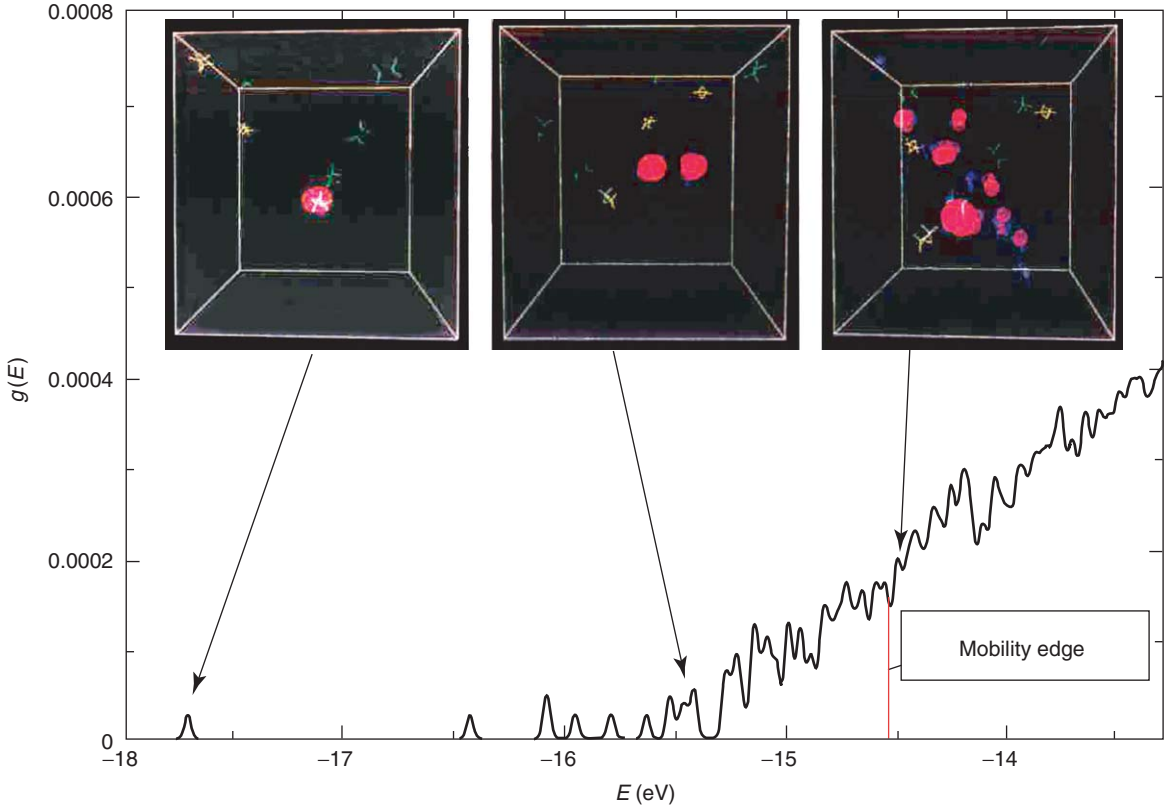


Figure 8 The density of states for the (lower-energy) valence-band tail calculated for a model of amorphous silicon, showing the computed position of the LD transition. Visualizations of three representative localized electron eigenstates are also shown: a highly localized near-mid-gap state associated with a dangling bond; a localized tail state and a critical state at the LD transition, each showing constituent island states, the number of which increase on approaching the LD transition. (J Ludlam and SR Elliott, unpublished data.)

difference with the atoms in the host material correspondingly give rise to electron states lying deeper within the bandgap.

Although relatively small levels of disorder, for example, bond-angle deviations in covalent amorphous semiconductors, produce relatively small changes in the EDOS, that is, moderate band tailing into the bandgap, larger degrees of disorder can cause correspondingly larger changes in the EDOS. For instance, electron states associated with gross instances of disorder (e.g., coordination defects such as dangling bonds in amorphous network materials) have energy levels deeper in the bandgap than the bond-angle-fluctuation-induced band tails. The reason why (singly occupied) dangling-bond orbitals should have an electron energy lying near midgap in the case of a-Si can be seen from the simple tight-binding hybrid–orbital picture illustrated in **Figure 7**: the dangling-bond orbital is a nonbonding state and hence should lie at the level corresponding to that of the sp^3-hybrids (i.e., at midgap in this model). Such singly occupied, electrically neutral dangling-bond defects in silicon can be represented by the symbol Si_3^0, where the subscript denotes the coordination number at the defect, and the superscript denotes the electrical charge state of the defect. Such Si_3^0 defects are electron-spin resonance (ESR) active because of the unpaired electron spin in the dangling-bond orbital.

Not all dangling-bond defects in amorphous semiconductors are paramagnetic, as they are in a-Si or a-Ge, however. In chalcogenide glasses, which contain the chalcogen (C) elements, S, Se, or Te, alloyed with neighboring elements (e.g., As, Ge, Sb, and B), bond reconstruction can make electron pairing at dangling-bond defects energetically favorable. An electron, transferred between two C_1^0 paramagnetic chalcogen dangling-bond defects to create a C_1^+, C_1^- defect pair, can be energetically stabilized by the formation of a dative bond between the empty p-orbital on C_1^+ and the filled nonbonding, lone-pair p–π orbital on a neighboring, normally bonded chalcogen (i.e., C_2^0), thereby forming an over-coordinated defect, C_3^+, with the overall defect reaction being

$$2C_2^0 \rightarrow C_3^+ + C_1^- \qquad [7]$$

The so-called valence-alternation pair (VAP) of defects, C_3^+, C_1^-, is spin-paired and hence diamagnetic, and has energy levels closer to the band edges than the near mid-gap position of C_1^0.

The Electronic Wave Function

Disorder in a material can cause a qualitative change in the nature of the electron wave functions compared with the Bloch form (eqn [1]) characteristic of perfectly crystalline solids. Instead of all wave functions being spatially extended (delocalized) as in the case of single crystals, the presence of disorder can cause some electron states instead to become spatially localized: the envelope of the magnitude of the wave function then decays exponentially with distance away from the center of localization (or the site of maximum magnitude of the electron charge density, $\rho_e(r)$, or $|\psi(r)|^2$) as

$$\psi(r) = Ae^{-\alpha r} \qquad [8]$$

where α^{-1} is the localization length.

An isolated substitutional impurity, for example, a P donor in c-Si (**Figure 1a**), also shows clearly the disorder-induced localized nature of the associated electron wave function: the isolated donor state can be represented reasonably accurately by a 1s-like hydrogenic orbital having an exponential envelope, the extra donor electron being bound to the positively charged P^+ donor atom by the Coulomb interaction, screened by the dielectric constant of the Si host, resulting in the effective Bohr radius of the electron orbit (equivalent to the localization length) being several tens of unit-cell lengths in size.

The isolated dangling-bond electron state lying near mid-gap in topologically disordered amorphous semiconductors is more strongly exponentially localized, with the localization length being comparable to the nearest-neighbor bond length, $\alpha^{-1} \cong r_1$. It is generally true that, the deeper the energy level of an electron state in the bandgap, away from the valence- or conduction-band edges, the stronger the degree of spatial localization (i.e., the smaller the value of α^{-1}) of the associated electron wave function.

A remarkable feature of the nature of electron states in disordered materials, such as amorphous solids, is that there are critical energies in the tails of the bands, at which the character of the eigenstates changes from being spatially extended (but non-Bloch-like, and having a dispersion of $\boldsymbol{k}$-values), for states nearer the band center, to being spatially (exponentially) localized, for states nearer the gap. The localization length, α^{-1}, of the envelope of the localized states diverges at the localization–delocalization (LD) transition. **Figure 8** shows this LD transition energy, sometimes called the (electron) mobility edge, because the electron transport changes there from being high-mobility diffusive for extended states to being low mobility involving thermal hopping for localized states, calculated for a structural model of a-Si with a realistic tight-binding electron-interaction Hamiltonian. Also shown in this figure are visualizations of the computed localized electron wave function for different representative energies in

the gap and the band tail. A dangling-bond state deep in the bandgap is very strongly localized at a coordination defect, the localized wave function consisting of a single region of charge density having an exponential envelope. However, localized electron states in the band tail, especially with energies approaching the LD transition, do not consist of single uniform regions of charge density with ever-increasing localization length, α^{-1}. Instead, discrete "islands" of charge density occur within the overall exponential envelope of the localized wave function; these individual islands are also exponentially localized, but their localization length, α_I^{-1}, increases only very gradually, and the number of islands also increases significantly, as their energy moves toward the LD transition. The electron eigenstate at the LD transition itself is a multifractal state consisting of a dense collection of the individual island states (see **Figure 8**).

Conclusions

The effects of disorder on electron states in solids are fourfold: (1) the electron wave vector is no longer a good quantum number, (2) sharp features in the density of states are destroyed, (3) bands broaden, creating band tails having an exponential energy dependence, and (4) electron states in band tails and in the bandgap region, lying below a critical energy and corresponding to very disordered sites, are exponentially spatially localized.

See also: Disorder and Localization Theory.

PACS: 71.23. – k; 71.2.An; 71.23.Cq; 71.55.Jv

Further Reading

Elliott SR (1990) *Physics of Amorphous Materials*, 2nd edn. Harlow: Longman.

Elliott SR (1998) *The Physics and Chemistry of Solids*. Chichester: Wiley.

Kamimura H and Aoki H (1989) *The Physics of Interacting Electrons in Disordered Systems*. Oxford: Clarendon Press.

Kramer B and MacKinnon A (1993) Localization: theory and experiment. *Reports of Progress in Physics* 56: 1469.

Mirlin AD (2000) Statistics of energy levels and eigenfunctions in disordered systems. *Physics Reports* 326: 259.

Morigaki K (1999) *Physics of Amorphous Semiconductors*. London: Imperial College Press.

Singh J and Shimakawa K (2003) *Advances in Amorphous Semiconductors*. London: Taylor and Francis.

Nomenclature

a	unit-cell parameter
C_i^{α}	chalcogen atom with atomic coordination i $(=1-3)$ and electric charge $\alpha(-,0,+)$
$D(E)$	EDOS (per energy per volume) $= g(E)/V$
E	electron energy
E_g	bandgap energy
$g(E)$	electronic density of states (per energy) $= D(E) \times V$
k	electron wave vector magnitude
$\boldsymbol{k}$	electron wave vector
$\boldsymbol{r}$	real space position vector
S_E	constant-energy surface of electron states in $\boldsymbol{k}$-space
$u_k(r)$	function having periodicity of crystal lattice (in Bloch function)
V	volume
V_1	intrasite electron interaction
V_2	intersite electron interaction
α	inverse localization length of envelope of a localized state
α_I	inverse localization length of island state within a localized state
$\delta(E)$	Dirac delta function (of energy)
$\nabla_k E$	gradient of electron energy with respect to wave vector
θ	bond angle
$\rho_e(r)$	electron charge density
$\psi(\boldsymbol{r})$	electron wave function

DNA and RNA, Biophysical Aspects

M D Frank-Kamenetskii, Boston University, Boston, MA, USA

Introduction

DNA and RNA together are known as nucleic acids. Their full names are deoxyribonucleic acid and ribonucleic acid, respectively. They play a crucial role in all living organisms because they are the key molecules responsible for storage, duplication, and realization of genetic information. They are both heteropolymeric molecules consisting of residues (nucleotides) of four types. Four RNA nucleotides are shown in **Figure 1**. They consist of the phosphate group, sugar (called ribose) (these two elements are identical for all four nucleotides), and nitrous bases,

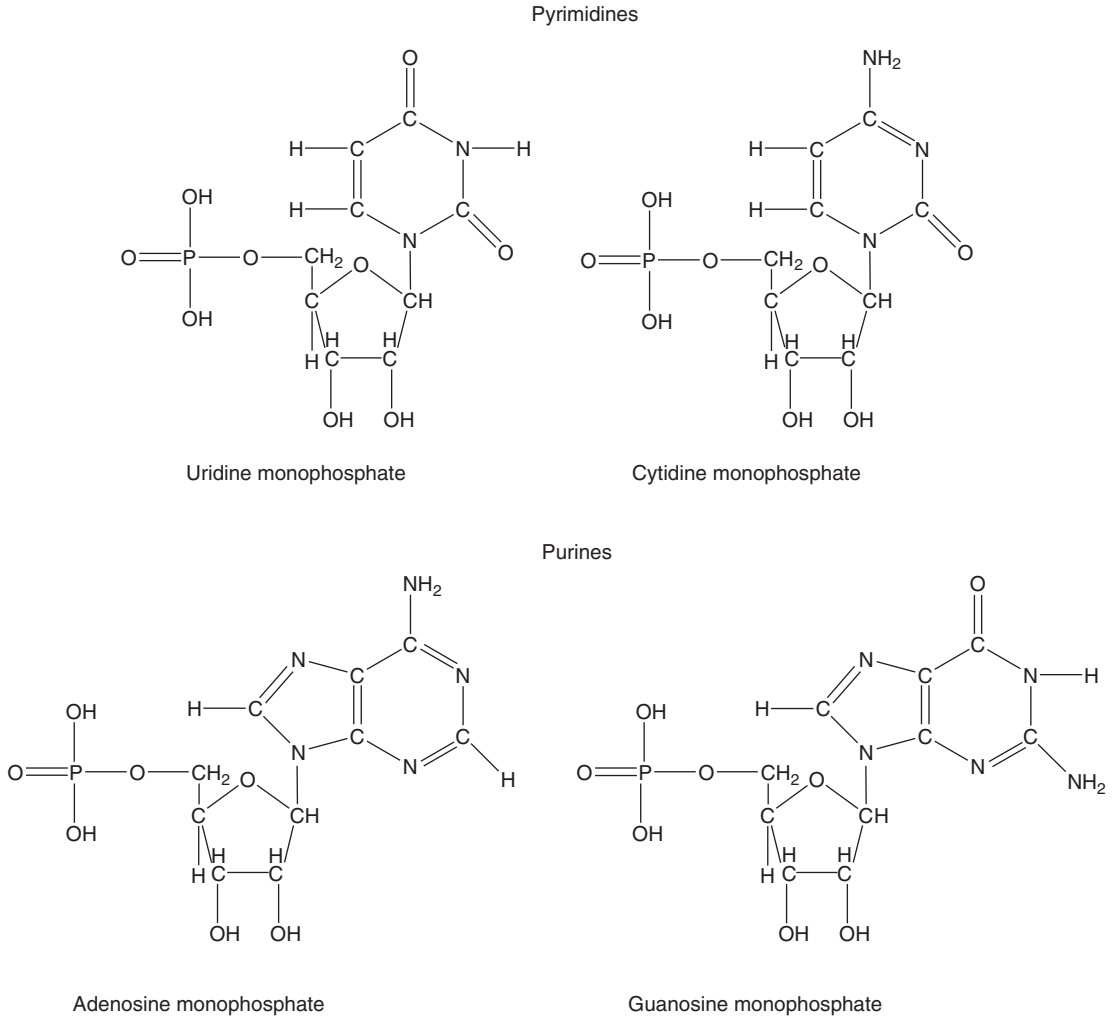


Figure 1 Chemical formulas of residues of RNA, i.e., of nucleotides. At the top are the pyrimidine nucleotides (U and C), and below, the purine nucleotides (A and G). Nucleotides within DNA differ in that, instead of the right-hand lower OH group, they simply have H. In addition, DNA, instead of the uridine nucleotide, includes the thymidine nucleotide whose top C–H group in the ring is replaced by the C–CH_3 group.

which determine the type of the nucleotide. DNA nucleotides are very similar to RNA nucleotides. The main difference is that the right-handed OH group in sugar is replaced with just H (making deoxyribose). Three bases (A, G, and C) are the same for DNA and RNA. Only the fourth base is different: instead of uracil (U), DNA carries thymine (T) (see **Figure 2**). The DNA chain is shown in **Figure 3**.

In living nature, as is observed today, DNA plays the most important role because genetic information is stored and duplicated in the form of DNA in all living cells and organisms without any exceptions. Only in some viruses, which are not living creatures because they cannot reproduce themselves outside cells, is genetic information carried in the form of RNA molecules. A notable class of such viruses is retroviruses, which include many cancer-inducing viruses as well as HIV, the AIDS virus.

A genetic message is "written down" in the form of a continuous text consisting of four letters (DNA nucleotides A, G, T, and C). This continuous text, however, is subdivided, in its biological meaning, into sections. The most significant sections are genes, parts of DNA, which carry information about the amino acid sequence of proteins. Whereas DNA stores and duplicates genetic information, RNA plays a key role in its realization, that is, in the process of synthesis of a protein molecule in accordance with the nucleotide sequence of the corresponding gene.

This process consists of two major steps. First an RNA copy of the gene is made with the help of a special enzyme, RNA polymerase. The process is called transcription and the RNA copy of the gene is

called messenger RNA (mRNA). The mRNA is an exact copy of the gene, in which all thymines (T) are replaced with uracils (U). In the second step, called translation, mRNA is used as a template for protein synthesis in a special cell machinery, the ribosome. Ribosome is a complex aggregate of special ribosomal RNA (rRNA) and a number of proteins. It translates the nucleic acid text into amino acid language. In so doing, it uses a special dictionary, the genetic code. In the process of translation, a very important class of additional small RNA molecules plays a crucial role. These RNA molecules consist of about hundred nucleotides and are called transfer RNA (tRNA).

In this article, an overview of the present knowledge of physical structures of nucleic acids is provided. The most important structural transitions, such as DNA melting and B–Z transition, are explained. Special attention is paid to the topological properties of DNA, which play a significant role in its functioning within the cell. Theoretical models, which are most popular in the field of biophysics of DNA and RNA are also overviewed. Recently emerged areas of DNA and RNA biophysics, such as single-molecule experiments and DNA/RNA nanostructures are also considered.

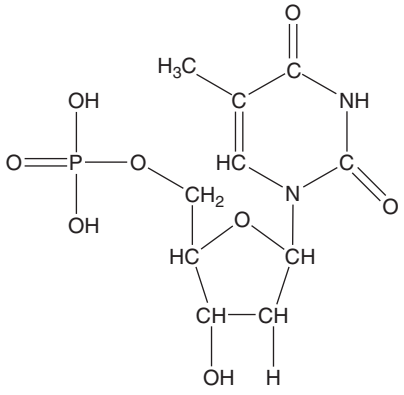


Figure 2 Thymidine monophosphate is a thymine nucleotide that is part of DNA. The remaining three DNA nucleotides have similar structure, but each has a nitrous base of its own (the top group). These three bases (adenine, guanine, and cytosine) are identical in DNA and RNA (see **Figure 1**).

Structures of DNA and RNA

Traditionally, the structural versatility of nucleic acids had been considered as much narrower than that of proteins. This viewpoint fitted well "the division of power" between nucleic acids and proteins in present-day living creatures. The major function of nucleic acids is to store and reproduce genetic information while proteins perform an innumerable variety of reactions within the cell and are also its main construction blocks. The discovery of ribozymes (RNA molecules capable of catalytic functions) and of the remarkable versatility of the RNA structure completely changed the attitude toward the pivotal question of which type of major biopolymers, proteins or nucleic acids, had been the prebiotic molecule, forefather of the living cell. Earlier, everybody believed that only protein could fulfill this function. Now, the common view is that this should be the RNA molecule (the RNA World concept).

However, in the present-day cell, RNA plays an important but mostly an auxiliary role. The central player is DNA.

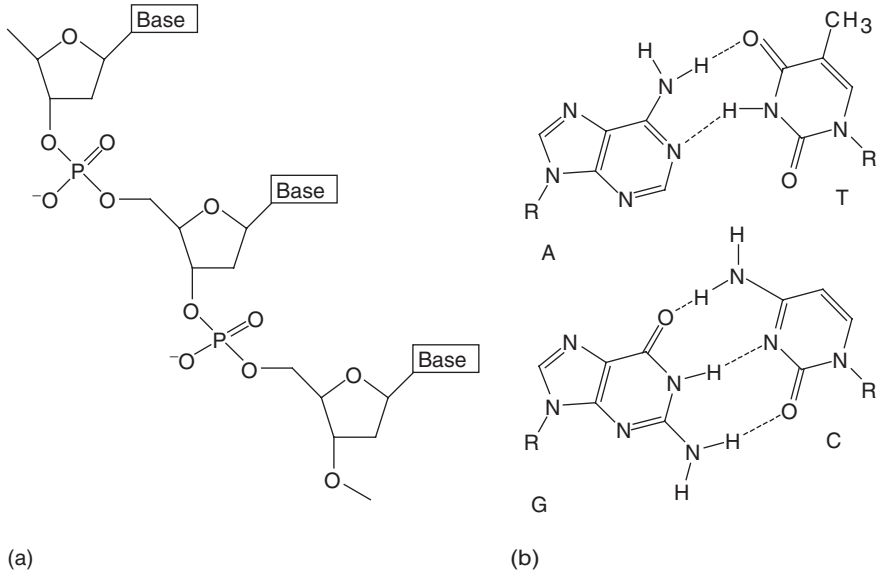


Figure 3 (a) The DNA single strand and (b) the Watson–Crick complementary base pairs.

Major Helical Structures of DNA

In spite of enormous versatility of living creatures, and, accordingly, variability of genetic texts which DNA molecules in different organisms carry, they all have virtually an identical physical spatial structure: the double-helical B form discovered by Watson and Crick in 1953. Sequences of the two strands of the double helix obey the complementary principle. This principle is the most important law in the field of nucleic acids, and, probably, the most important law of living nature. It declares that in the double helix, A always opposes T and vice versa, whereas G always opposes C and vice versa.

B-DNA It (see **Figure 4a**) consists of two helically twisted sugar-phosphate backbones stuffed with base pairs of two types, AT and GC. The helix is right-handed with 10 bp per turn. The base pairs are isomorphous (see **Figure 3**): the distances between glycosidic bonds, which attach bases to sugar, are virtually identical for AT and GC pairs. Due to this isomorphism, the regular double helix is formed for an arbitrary sequence of nucleotides and the fact that DNA should form a double helix imposes no limitations on DNA texts. The surface of the double helix is by no means cylindrical. It has two very distinct grooves: the major groove and the minor groove. These grooves are extremely important for DNA functioning because in the cell, numerous proteins recognize specific sites on the DNA via binding with the grooves.

Each nucleotide has direction: if one treats sugar as a ring lying in the plane, then the CH_2 group will be above the plane (see **Figure 1**). This carbon is designated as 5′ whereas the carbon atom connecting the sugar to the next nucleotide in the polynucleotide chain is designated as 3′. In the DNA double helix, the two strands have opposite directions.

In B-DNA, base pairs are planar and perpendicular to the axis of the double helix.

Under normal conditions in solution, often referred to as "physiological" (neutral pH, room temperature, ~200 mM NaCl), DNA adopts the B form. All available data indicate that the same is true for the totality of DNA within the cell. It does not exclude, however, the possibility that separate stretches of DNA carrying special nucleotide sequences would adopt other conformations.

B′-DNA Up to now only one such conformation has been demonstrated, beyond any doubts, to exist under physiological conditions. When three or more A residues in one strand (and, accordingly, T's in the other DNA strand) are met, they adopt the B′ form. In many respects, the B′ form is similar to the classical B form. There are two main differences: base pairs in B′-DNA are not planar (they form a kind of propeller with the propeller twist of 20°) and the minor groove in B′-DNA is twice as narrow as in B-DNA.

A-DNA Similar to B-DNA, the A form can be adopted by an arbitrary sequence of nucleotides. As in B-DNA, the two complementary strands in A-DNA are antiparallel and form right-handed helices. Normal DNA undergoes transition from the B to A form under drying. In A-DNA, the base pairs are planar but their planes form a considerable angle with the axis of the double helix. In so doing, the base pairs shift from the center of the duplex forming a channel in the center.

Z-DNA It presents the most striking example of how different from the B form the DNA double helix can be (**Figure 4b**). Although in Z-DNA the complementary strands are antiparallel as in B-DNA, unlike in B-DNA they form left-handed, rather than right-handed, helices. There are many other dramatic differences between Z- and B-DNA.

Not any sequence can adopt the Z form. To adopt the Z form, the regular alternation of purines (A or G) and pyrimidines (T or C) along one strand is strongly preferred. However, even this is not enough for Z-DNA to be formed under physiological conditions. Nevertheless, Z-DNA can be adopted by DNA stretches in cell due to DNA supercoiling (see below). The biological significance of Z-DNA, however, remains to be elucidated.

ps-DNA The complementary strands in DNA duplex can be parallel. Such parallel-stranded (ps) DNA is formed most readily if both strands carry only adenines and thymines and their sequence excludes formation of the ordinary antiparallel duplex. If these requirements are met, the parallel duplex is formed under quite normal conditions. It is right-handed but the AT pairs are not regular Watson–Crick ones, but rather so-called reverse Watson–Crick.

Some other sequences also can adopt parallel duplexes. For instance, at acidic conditions two strands carrying only C residues form parallel duplex consisting of hemiprotonated CC^+ base pairs (see the section "Quadruplexes"). However, any possible biological role of psDNA remains obscure.

RNA Structure

Quite naturally, similarity between DNA and RNA in their chemical nature entails significant similarity in structures they adopt. RNA can also form the double helix with complementary base pairs AU and

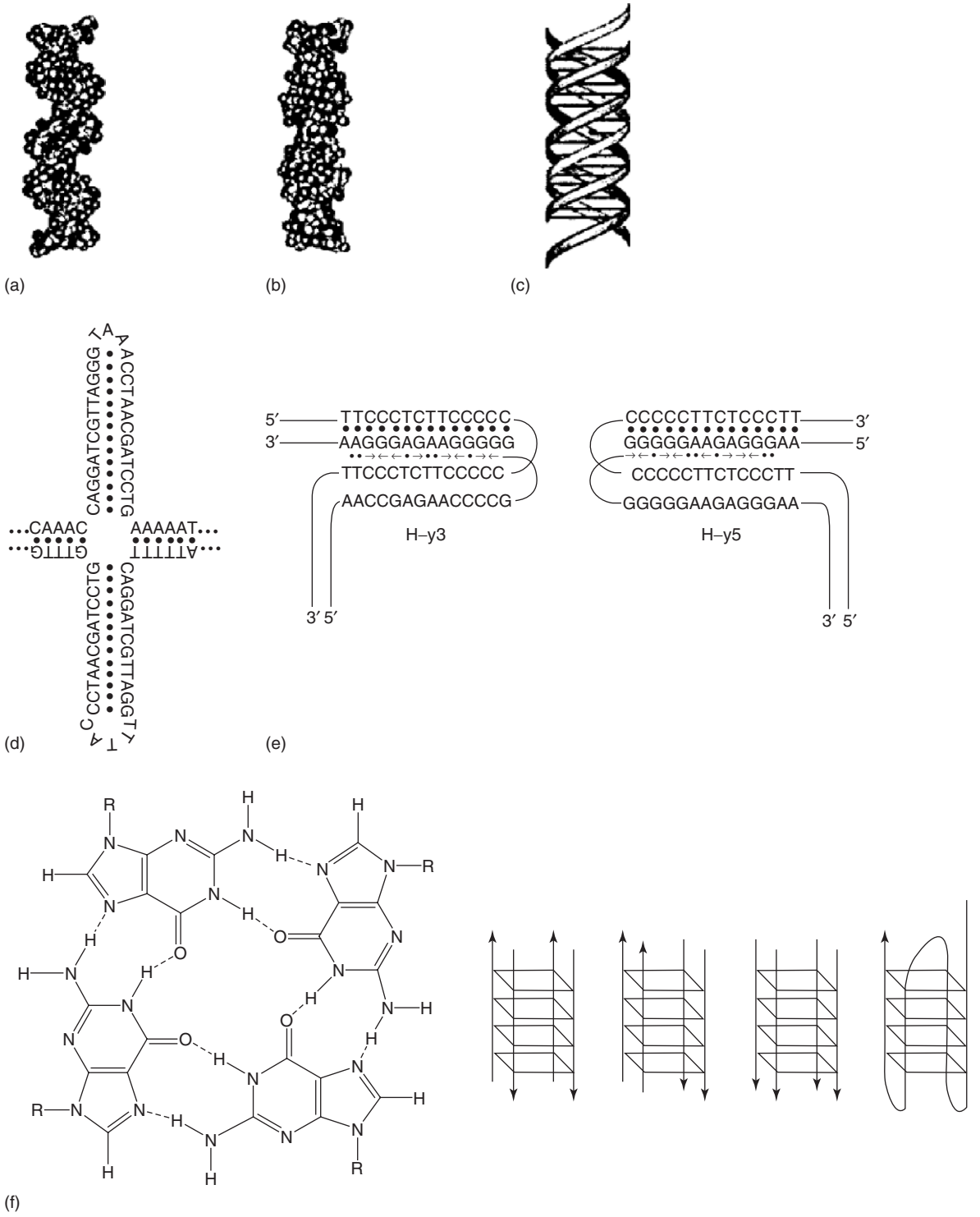


Figure 4 Schematics of (a) B-DNA and some unusual structures of DNA: (b) Z-DNA, (c) intermolecular triplex, (d) cruciform, (e) H-DNA, (f) G-quadruplex and its various folding modes.

GC. However, the major RNA duplex conformation remains A-form rather than B-form.

In general, due to bulkier OH group in the sugar ring, the versatility of RNA duplexes is rather narrower than those of DNA.

In a cell, different types of RNA molecules (mRNA, rRNA, tRNA, etc.) exist as single strands. Of course, these strands are not unfolded, but fold into complex spatial structures. Mutually complementary stretches of the same molecules form numerous short duplexes,

which are a major structural motif of RNA molecules. Single-stranded regions form loops of different types. Sometimes very complex folding patterns, called pseudoknots, are met.

Spatial structure plays an important role in RNA functioning. For instance, rRNA organized in complex spatial structure forms a scaffolding to which numerous ribosomal proteins attach to form the functional ribosome. The elucidation of the full 3D structure of the ribosome by X-ray crystallography has been one of the major achievements of biophysics and structural biology. Spatial structure at tRNA molecules are specially designed to permit them to fulfill the central role in the realization of the genetic code. One part of tRNA carries the anticodon, a trinucleotide complementary to a codon (a trinucleotide corresponding to one amino acid in the genetic code), whereas the amino acid, which corresponds to the codon in the genetic code, is attached to one terminus of the molecule.

DNA Topology

DNA very often operates within the cell in the form of a closed circular molecule in which both strands form closed circles (**Figure 5**). The physical properties and physiological behavior of closed circular (cc) DNA are different in many respects from those of linear DNA molecules. These differences stem from the fact that ccDNA acquires, as compared to linear DNA, a new fundamental feature, topology. There are two levels of DNA topology.

Knots of DNA

The double helix, as a whole, may form a simple circle (i.e., a trivial knot) or be knotted (see **Figure 6**).

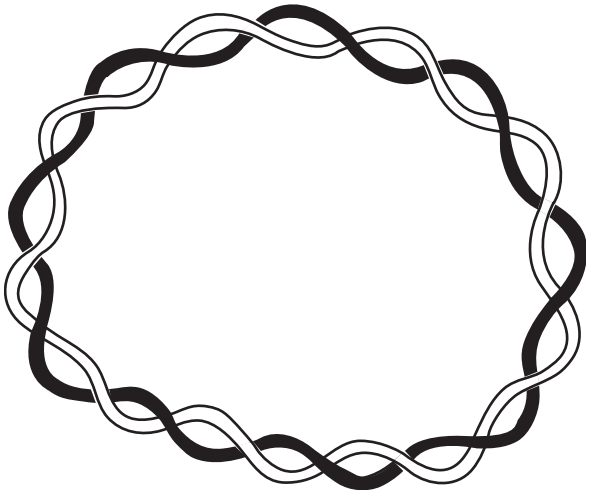


Figure 5 In a ccDNA, two complementary strands form a linkage of a high order.

Knotting of DNA may be conducted by random closure of linear molecules with "cohesive" ends (mutually complementary single-stranded overhangs). There are also different enzymes which do the job. The most important class of such enzymes is known as DNA topoisomerases II. The type of knot is the topological invariant, that is, it cannot be changed (converted to another type of knot) without breakage of both DNA strands. As a result, identical ccDNA molecules belonging to different knot types may be separated in gel.

DNA Supercoiling

The two complementary strands in ccDNA are topologically linked with one another (see **Figure 5**). The degree of linkage, which is designated as *Lk*, the DNA linking number, is a topological invariant, which cannot be changed with any deformations of the DNA strands without strand breaks. The *Lk* value is easily calculated as the number of times one strand pierces the surface placed on the other strand. If N is the number of base pairs in DNA and γ_o is the

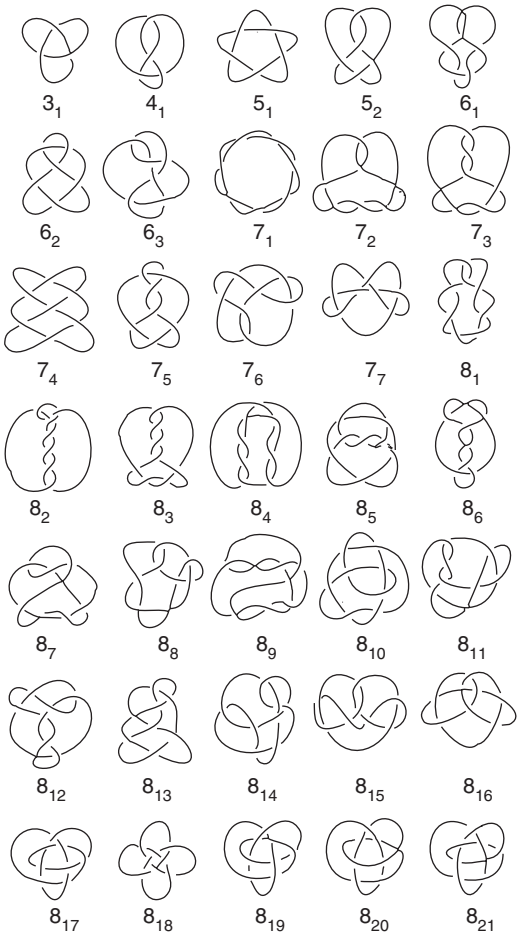


Figure 6 Knots.

number of base pairs per one turn of the double helix under given ambient conditions, then the difference:

$$\tau = Lk - N/\gamma_o \qquad [1]$$

is the number of superhelical turns in ccDNA under given conditions. If $\tau \neq 0$, then ccDNA cannot form a planar ring with the helical twist corresponding the γ_o value. Either the helical twist should change or the DNA molecule as a whole should writhe in space. Actually, with $\tau \neq 0$ both things happen: the DNA adopts the writhe, and its twist changes. The distribution of energy associated with DNA supercoiling (i.e., with $\tau \neq 0$) between twist and writhe is determined by the torsional and bending rigidities, which have been measured with good accuracy. Supercoiling is also significantly affected by replication and transcription. DNA topology plays an extremely important role in DNA functioning.

Structural Transitions and Unusual Structures

If negative supercoiling in ccDNA becomes too high, the bending and torsional degrees of freedom may become insufficient to accommodate the superhelical energy. Under these circumstances, the integrity of the DNA double helix breaks down at its most "weak" sites. These sites, which have the potential to adopt structures significantly different from B-DNA, can undergo structural transitions into unusual (non-B-DNA) structures.

DNA Melting

On heating or under superhelical stress, the DNA complementary strands separate forming single-stranded loops. In case of linear duplex molecules, the process ends up, at sufficiently high temperature, with complete separation of the two strands. The process has been studied comprehensively both experimentally and theoretically and is known as DNA melting (the terms denaturation or helix-coil transition also exist in literature). AT pairs have significantly lower stability than GC pairs.

With increasing temperature, long DNA sections melt out. Melting of each such section is a cooperative, all-or-none phenomenon. However, melting of a long enough molecule becomes a series of such cooperative processes, each occurring at its own temperature (**Figure 7**). Roughly speaking, sections more enriched with GC pairs melt out at higher temperatures. Thus, DNA melting is not a classical phase transition because two different "phases," helical and melted, coexist with each other in one and the same molecule.

When two mutually complementary single-stranded DNA molecules are mixed, the process reverse to melting occurs. It is called renaturation, annealing or hybridization.

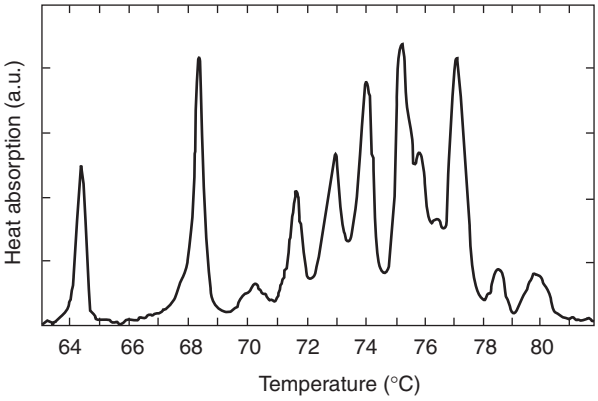


Figure 7 Melting of DNA. This curve is also often called the differential melting curve. The curve was obtained for DNA that has the code name of ColE1 and contains ~6500 nucleotide pairs.

DNA melting is an exceptionally important phenomenon. It occurs in the cell when DNA replicates and when DNA polymerase "reads" the mRNA copy of the gene. Cycles of melting and annealing are utilized in the PCR machines, the most remarkable tools of the current biotechnology revolution.

B–Z Transition

Negative supercoiling favors the formation of Z-DNA most because, in this case, the maximal release of superhelical stress per 1 bp adopting a non-B-DNA structure is achieved. As a result, although under physiological conditions the Z form is energetically very unfavorable as compared with B-DNA, it is easily adopted in negatively supercoiled DNA by appropriate DNA sequences (with alternating purines and pyrimidines).

Cruciforms

Another structure readily formed under negative supercoiling is cruciform, which requires palindromic regions (see **Figure 4d**). To form cruciform, a palindromic region should be long enough. For example, 6 bp long palindromes recognized by restriction enzymes do not form cruciforms under any conditions.

Triplexes

Intramolecular triplex or H-DNA forms a special class of unusual structures, which are adopted under superhelical stress by sequences carrying purines (A and G) in one strand and pyrimidines (T and C) in the other, that is, homopurine–homopyrimidine sequences. The major element of H-DNA is triplex formed by a half of the insert adopting the H form and by one of two strands of the second half of the

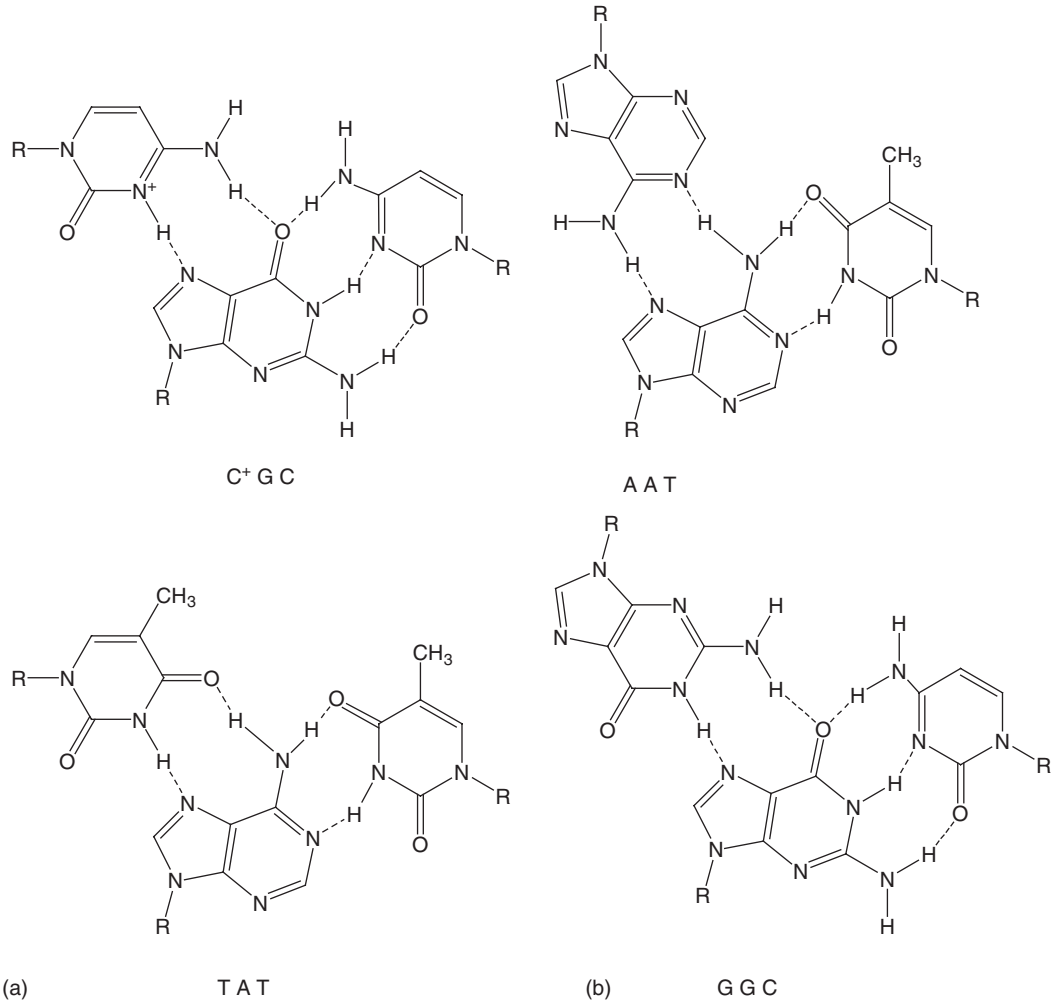


Figure 8 The structure of (a) pyrimidine (TAT and C^+GC), and (b) purine (AAT and GGC) base triads of which the DNA triple helix is made.

insert (**Figure 4e**). Two major classes of triplexes are known: pyrimidine–purine–pyrimidine (PyPuPy) and pyrimidine–purine–purine (PyPuPu). **Figure 8** shows the canonical base-triads entering these triplexes.

Always two isomeric forms of H-DNA are possible, which are designated as H-y3, H-y5, H-r3, and H-r5 depending on which kind of triplex is formed and which half of the insert forms the triplex (see **Figure 4e**).

Discovery of H-DNA stimulated studies of intermolecular triplexes, which may be formed between homopurine–homopyrimidine regions of duplex DNA and corresponding pyrimidine or purine oligonucleotides (**Figure 4c**).

Quadruplexes

Of all nucleotides, guanines are most versatile in forming different structures. They may form GG pairs but the most stable structure, which is formed in the presence of monovalent cations (especially potassium), is G4 quadruplex (see **Figure 4f**). G-quadruplexes may exist in a variety of modifications: all-parallel, all-antiparallel, and others. As a result, G-quadruplexes are easily formed both inter- and intramolecularly, again with a variety of modifications. A remarkable variety of quadruplex structure was discovered for protonated cytosines. This structure, called i-motif, consists of two pairs of parallel stranded duplexes consisting of hemiprotonated CC^+ base pairs. In the final structure, two duplexes and mutually antiparallel and CC^+ base pairs from one ps duplex alternate with CC^+ base pairs from the other ps duplex. The structure is stable only at acidic pH because protonation is necessary.

Bent DNA

For an arbitrary sequence, the minimum energy conformation of the DNA double helix corresponds to the straight DNA axis. However, there is a notable exception to this rule. If three or more A nucleotides

are located in a row in one of the DNA strands, the corresponding regions adopt the B′ structure. As a result, the DNA axis experiences a bend of 20° in such a region. DNA bending plays an important role in DNA functioning.

RNA Unusual Structures

Most DNA unusual structures have their analogs in the RNA world. Specifically, formation of triplexes was first demonstrated for artificial RNA chains. Solitary RNA base-triads are often met in tRNA structures. Triplexes are formed by mixed DNA–RNA hybrids and some of them are rather stable.

However, in general, studies of RNA unusual structures lag behind corresponding DNA structures because methods to study unusual structures are much better developed in the case of DNA than in the case of RNA.

Nanostructures and Nanodevices from DNA and RNA

One of the most rapidly developing areas in DNA and RNA biophysics consists in using the remarkable ability of nucleic acid to self-assembly for creation of nanostructures and nanodevices. In so doing, virtually all unusual structures described above are extensively used: triplexes, quadruplexes, Z-DNA, etc., as well as the duplex. Specifically, a family of topological and pseudotopological nanostuctures has been assembled as is shown in **Figures 9a–9c**. An example of a nanodevice (nanoactuator) is shown in **Figure 9d**. By adding specially designed single-stranded DNA chains, the nanoactuator is transformed from the relaxed form shown in **Figure 9d** to a straightened form in which the A strand is extended into a straight double helix.

Biophysical Methods to Study DNA and RNA

The whole arsenal of methods, which is traditionally used to study molecular structure, is applied to DNA and RNA. A leading position is occupied by X-ray crystallography, which provided us with a knowledge of details of the atomic structure of DNA and RNA. X-ray crystallography made it possible to solve the structure of the nucleosome, the main building block of chromosomes consisting of DNA and hystone proteins, and of a major cellular machinery, ribosome, consisting of RNA and many proteins. The next method is proton NMR, which is especially valuable when biological molecules resist crystallization, a not uncommon problem with nucleic acids. An enormous body of data on the structure of DNA triplexes, quadruplexes, and various ribozymes was delivered by NMR. Electron microscopy (EM), crio-EM and atomic force microscopy (AFM) have become indispensable in studying the DNA and its complexes with proteins. All kinds of spectroscopy methods, UV, IR, Raman, and CD are very useful in studying conformational transitions in DNA and RNA. A special role is played by various fluorescence methods, which find more and more applications not only in biophysical studies, but also in biotechnology. Below, specific methods, which have been specially developed in the field of DNA and RNA biophysics are discussed.

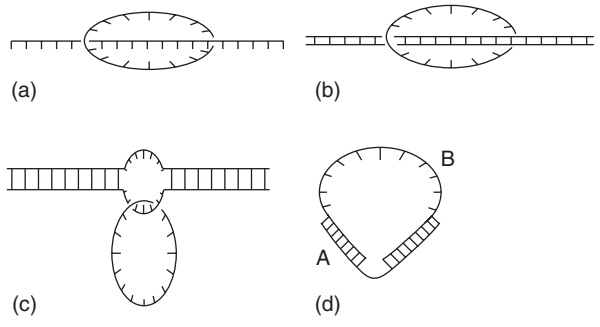


Figure 9 Examples of DNA nanostructures: (a) padlock probe (pseudororaxane), (b) sliding clamp (pseudororaxane), (c) earring probe (pseudocatenane), and (d) nanoactuator.

Theoretical Models

As in the study of any important physical object, a number of simplified theoretical models of nucleic acids exist, different models being used to analyze different properties. **Figure 10** presents schematics of some of these models.

The DNA double helix may be treated as an isotropic elastic rod (**Figure 10a**). In the framework of this model, the DNA molecule is described by only three parameters: bending, torsional rigidities, and the diameter. The model has proved to be extremely useful for analyzing hydrodynamics and other properties of linear DNA, when it behaves as a usual polymeric molecule. It has also provided a comprehensive theoretical treatment of DNA topology of both levels: knotting and supercoiling. More recently, the model has been extensively used for quantitative analysis of the force-extension curves for DNA in single-molecule experiments (see below).

A quite different, but also very successful, model treats the DNA double helix as consisting of base pairs of two types: closed and open (**Figure 10b**). This is the helix-coil model, which explains all major features of DNA melting quantitatively.

The polyelectrolyte model (**Figure 10c**) treats DNA itself just as a charged cylinder but allows for mobile counter-ions surrounding the double helix.

To predict the RNA structure, a simplified model of RNA folding is widely used (**Figure 10d**).

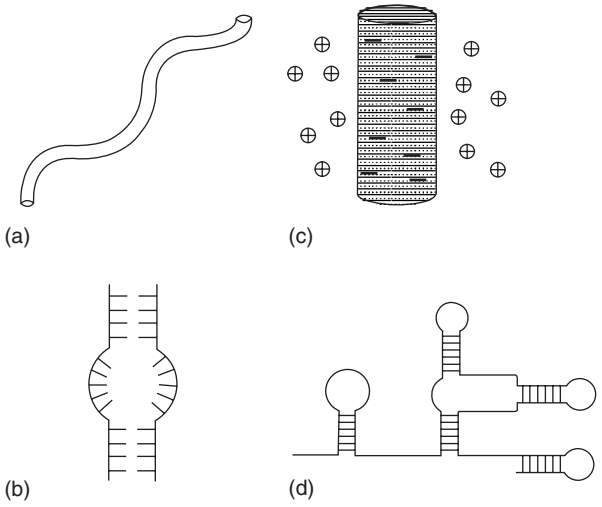


Figure 10 Theoretical models of nucleic acids: (a) elastic-rod model, (b) helix-coil model, (c) polyelectrolyte model, and (d) model of RNA folding.

Gel Electrophoresis

Gel electrophoresis is a simple technique, introduced in the early 1970s, which truly revolutionized the biophysical studies of DNA and RNA and, subsequently, the whole field of molecular biology. All developments in this field in the past 30 years are connected, directly or indirectly, with the gel electrophoresis method. Gel electrophoresis has pushed aside ultracentrifugation as the method to separate nucleic acids.

Gel electrophoresis differs from electrophoresis in solution only in the nature of the medium in which molecules are separated in the electric field. In case of gel electrophoresis the medium is the gel, a polymer network. The most popular in the field of nucleic acids are gels made of polyacrylamide or agarose. Originally, the great advantage of gels in the separation of nucleic acids was discovered purely empirically. The understanding came later after some ideas of De Grennes P-J were borrowed from polymer physics, namely the notion of reptation of polymer molecules. Gel electrophoresis is indispensable in all methods of DNA sequencing which have revolutionized the entire field of biology. The DNA sequencing techniques have been developed to perfection, based on different conceptual approaches. The most fertile approach proved to be that of Sanger who developed a method of DNA sequencing, which uses enzymes working on DNA.

Single-Molecule Experiments with DNA and RNA

In recent years, remarkable progress in the application of the single-molecule techniques to DNA and RNA studies has been achieved. There are many different formats for single-molecule experiments.

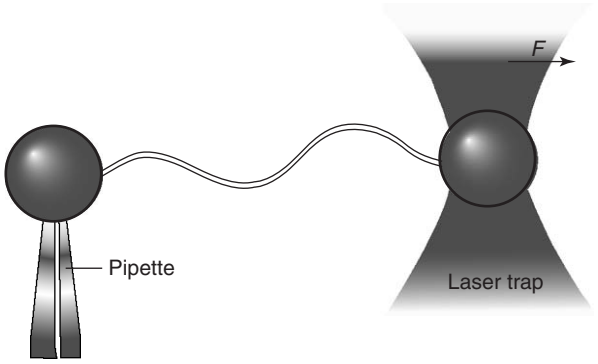


Figure 11 Stretching a single DNA molecule with optical tweezers.

For example, a DNA molecule consisting of about 50 kbp is attached via one terminus to a pipette tip equipped with a streptavidin-covered polymer bead. Streptavidin is a protein, which binds very strongly to a small molecule, biotin. The DNA terminus is covalently tagged with biotin to be attached to the streptavidin-covered bead. The other terminus is also attached to the polymer bead, but the bead is not fixed. Then, optical tweezers are used to drag the free bead with the DNA attached to it. As a result, a single DNA molecule can be stretched in a fully controllable manner and the force created in the process can be accurately measured (see **Figure 11**).

These experiments have allowed, for the first time, measurement of the force–extension curves of ds and ss DNAs and comparison of the curves with theoretical predictions. One of the important directions of these studies include isothermal melting of DNA and RNA duplexes under the external force. The influence of the force on the performance of major enzymes working on DNA and RNA (DNA and RNA polymerases) has also been studied. A very important feature of single-molecule experiments consists in the fact that they introduce a new variable on which DNA and RNA structure critically depend, the external force.

See also: Biomolecules, Scanning Probe Microscopy of; Protein Folding, Engineering of.

PACS: 87.15

Further Reading

Bloomfield VA, Crothers DM, and Tinoco I (2000) *Nucleic Acids: Structures, Properties, and Functions*. Sausalito, CA: University Science Books.

Frank-Kamenetskii MD (1997) *Unraveling DNA*. Cambridge, MA: Perseus Publishing.

Sinden RR (1994) *DNA Structure and Function*. San Diego, CA: Academic Press.

E

Effective Masses

A L Wasserman, Oregon State University, Corvallis, OR, USA

Introduction

Electrons in condensed matter interact strongly with their environment. This environment generally consists of a crystalline lattice, phonons, other electrons and an assortment of localized "defect" potentials, as well as applied external fields. As a result of these interactions, free, single particle energies ε_0, where

$$\varepsilon_0(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m} \qquad [1]$$

with $\boldsymbol{k}$ the electron wave vector and m the "bare" electron mass, are usually altered in significant ways, which can include simple energy shifts as well as imparting finite lifetimes to the single particle states.

It is often possible and convenient, both for experimentalists and theoreticians, to regard the particles plus interactions as a new fermionic entity, referred to as a "quasiparticle" with energy

$$\varepsilon(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m_{\text{eff}}} \qquad [2]$$

where an effective mass m_{eff} can parametrize important features of the interacting environment. There is, in fact, a considerable range of experimental phenomena whose results are usefully parametrized in this surprisingly simple way. The extent to which some (or any) of them relate to first-principles microscopic interactions are briefly surveyed here.

In an attempt to bring some uniformity to the calculation and interpretation of effective masses, one can present a general approach to the problem by deriving a single formulation for effective mass calculations. This formulation is first adapted to semiconductor and insulator problems where effective mass parameters are attributed, in the first order, to interactions with the periodic lattice and, in particular, to electrons whose occupation is confined to the edges of the conduction and valence bands. The formulation is then modified for dense Fermi systems in which particle behavior at the Fermi surface is the crucial determinant of electronic properties, including effective masses.

Consequences of Interactions

A general expression which describes the effects of interactions is

$$\varepsilon(\boldsymbol{k}) = \varepsilon_0(\boldsymbol{k}) + \Gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})] \qquad [3]$$

where $\varepsilon(\boldsymbol{k})$ is the fully interaction-modified "quasiparticle" energy, $\varepsilon_0(\boldsymbol{k})$ is the noninteracting energy, as given in eqn [1], and $\Gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})]$ is the quantitative consequence of all the environmental perturbations to the free electron system. It should be noted that Γ depends on the wave vector $\boldsymbol{k}$, but may also depend, self-consistently, on $\varepsilon(\boldsymbol{k})$, the quasiparticle energy of the state itself. Moreover, $\Gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})]$ is, in general, a complex quantity:

$$\Gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})] = \Delta[\boldsymbol{k}, \varepsilon(\boldsymbol{k})] + \mathrm{i}\gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})] \qquad [4]$$

the imaginary part usually being responsible for quasiparticle lifetime effects. Although $\gamma \neq 0$ is often the case, this possibility does not explicitly contribute to the following discussion. The apparent abstractness of eqns [3] and [4] notwithstanding, one finds that a quasiparticle effective mass m_{eff} is a useful and intuitive concept.

An Effective Mass Formulation

In formulating the general case for an effective mass theory which can embrace simple interactions, such as band structure in semiconductors, as well as phonons and the more subtle electron–electron interactions, one begins with eqns [3] and [4] and assumes that the imaginary part $\gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})]$ has a negligible effect in determining m_{eff}.

The initial step in obtaining an explicit form for the "quasiparticle" energy $\varepsilon(\boldsymbol{k})$ is most simply done

by applying Newton's iterative method to eqn [3], which is first rewritten in the form

$$\varepsilon(\boldsymbol{k}) - \varepsilon_0 - \Delta[\boldsymbol{k}, \varepsilon(\boldsymbol{k})] = 0 \qquad [5]$$

Newton's method consists of first expanding eqn [5] to linear order about some initial value $\varepsilon^{(0)}$ (which will be chosen below) to give

$$\varepsilon^{(0)} - \varepsilon_0(\boldsymbol{k}) - \Delta(\boldsymbol{k}, \varepsilon^{(0)}) + \left[1 - \frac{\partial\Delta(\boldsymbol{k}, \varepsilon)}{\partial\varepsilon}\right]_{\varepsilon=\varepsilon^{(0)}} [\varepsilon(\boldsymbol{k}) - \varepsilon^{(0)}] = 0 \qquad [6]$$

Now solving for ε, one finds

$$\varepsilon(\boldsymbol{k}) = \varepsilon^{(0)} - \frac{[\varepsilon^{(0)} - \varepsilon_0(\boldsymbol{k}) - \Delta(\boldsymbol{k}, \varepsilon^{(0)})]}{[1 - (\partial\Delta(\boldsymbol{k}, \varepsilon)/\partial\varepsilon)]_{\varepsilon=\varepsilon^{(0)}}} \qquad [7]$$

Temporarily assuming, for simplicity, that $\boldsymbol{k}$ is a scalar and expanding to the first order in k^2 about $k = 0$, one has

$$\varepsilon(\boldsymbol{k}) = \varepsilon^{(0)} - \frac{[\varepsilon^{(0)} - \varepsilon_0(k) - \Delta(0, \varepsilon^{(0)}) - (\partial\Delta(k, \varepsilon^{(0)})/\partial(k^2))|_{k=0} k^2]}{[1 - (\partial\Delta(0, \varepsilon)/\partial\varepsilon)]_{\varepsilon=\varepsilon^{(0)}}} \qquad [8]$$

Finally, starting the iterations from $\varepsilon^{(0)} = 0$, after a single iteration one obtains

$$\varepsilon(k) = \frac{[\varepsilon_0(k) + \Delta(0,0) + (\partial\Delta(k,0)/\partial(k^2))|_{k=0} k^2]}{[1 - \partial\Delta(0,\varepsilon)/\partial\varepsilon]_{\varepsilon=0}} \qquad [9]$$

Assuming the constant energy shift $\Delta(0,0) = 0$ and from the above equation $\varepsilon_0(\boldsymbol{k}) = \hbar^2 k^2/2m$, one can obtain

$$\varepsilon(k) = \frac{\hbar^2}{2m}\left\{\frac{1 + (2m/\hbar^2)[\partial\Delta(k,0)/\partial(k^2)]_{k=0}}{(1 - (\partial\Delta(0,\varepsilon)/\partial\varepsilon))_{\varepsilon=0}}\right\} k^2 \qquad [10]$$

Therefore, an effective mass can be defined by

$$\frac{m}{m_{\text{eff}}} = \left\{\frac{1 + (2m/\hbar^2)[\partial\Delta(k,0)/\partial(k^2)]_{k=0}}{(1 - (\partial\Delta(0,\varepsilon)/\partial\varepsilon))_{\varepsilon=0}}\right\} \qquad [11]$$

In the case that k is not a scalar, one should expand instead to the second order and similarly find an effective mass tensor

$$\left(\frac{m}{m_{\text{eff}}}\right)_{\mu\nu} = \left\{\frac{\delta_{\mu\nu} + (m/\hbar^2)[\partial^2\Delta(\boldsymbol{k},0)/\partial k_\mu \partial k_\nu]_{k=0}}{(1 - (\partial\Delta(0,\varepsilon)/\partial\varepsilon))_{\varepsilon=0}}\right\} \qquad [12]$$

In the general semiconductor band case,

$$\varepsilon_0(\boldsymbol{k}) + \Delta(\boldsymbol{k}, \omega) \rightarrow E(\boldsymbol{k}, 0) \qquad [13]$$

where $E(\boldsymbol{k})$ is the band dispersion relation and the $\boldsymbol{k}$-expansion of eqn [8] is taken about any band extremal point $\boldsymbol{k}_m$ (the Fermi gas case is discussed below). Once the relevant perturbing terms are known, eqn [11] or [12] may be straightforwardly applied.

Band-Structure Effective Mass

The most common application of effective mass theories is in parametrizing the features of a band structure. This is especially practical in semiconductors and insulators where most experimental features arise from electron occupation near the extremal points of band maxima (valence bands) and band minima (conduction bands). In the energy band approximation, it is first noted that $\partial\Delta(0,\varepsilon)/\partial\varepsilon = 0$. Recasting eqn [12] to reflect the energy band approximation (see eqn [13]), one obtains

$$\left(\frac{1}{m_{\text{eff}}}\right)_{\mu\nu} = \frac{1}{\hbar^2}\left[\frac{\partial^2 E(\boldsymbol{k}, 0)}{\partial k_\mu \partial k_\nu}\right]_{\boldsymbol{k}=\boldsymbol{k}_m} \qquad [14]$$

where $E(\boldsymbol{k},0)$ is the electron energy band dispersion relation and $\boldsymbol{k}_m$ are wave vectors of any conduction-band minima or valence-band maxima. Since the derivative in eqn [14] is related to the "curvature" of the dispersion relation, bands of large curvature correspond to small effective masses (light "quasiparticles") while flatter bands, that is, bands with small curvature, are "heavier" quasiparticles. Moreover, since valence bands have maxima at $\boldsymbol{k}_m$, their curvature will have negative values and, therefore, be quasiparticles with negative effective masses. This inadmissible condition is resolved by assigning to these quasiparticles a positive charge, referring to them as holes and thereby restoring proper signs to transport properties.

Effective Mass of a Degenerate Fermi System

In the case of a degenerate Fermi system, Fermi–Dirac statistics usually restrict dominant behavior to those electrons at the Fermi surface where $\boldsymbol{k}_F$, the Fermi wave vector, is neither a band maximum nor a minimum. Therefore, the noninteracting energy band relation $\varepsilon_0(\boldsymbol{k})$ is expressed as an expansion about $k = k_F$. Since k_F is not an extremal point, one can obtain an expression which has a term linear in $(k - k_F)$:

$$\varepsilon_0(k) = \varepsilon_0(k_F) + \frac{\hbar^2 k_F}{m}(k - k_F) + \cdots \qquad [15]$$

Similarly, eqn [7] is expanded about $k = k_F$ instead of $k = 0$ giving, to linear order,

$$\varepsilon(k) = \varepsilon^{(0)} - \{[1 - (\partial\Delta(k_F, \varepsilon)/\partial\varepsilon)]_{\varepsilon=\varepsilon^{(0)}}\}^{-1}[\varepsilon^{(0)} - \varepsilon_0(k_F) - \Delta(k_F, \varepsilon^{(0)}) - (\hbar^2 k_F/m)([1 + (m/\hbar^2 k_F) \times (\partial\Delta(k, \varepsilon^{(0)})/\partial k)|_{k=k_F}](k - k_F) + \cdots] \quad [16]$$

Iterating once again, starting with $\varepsilon^{(0)} = 0$, one obtains

$$\varepsilon(k) = \{[1 - (\partial\Delta(k_F, \varepsilon)/\partial\varepsilon)]_{\varepsilon=0}\}^{-1}[\varepsilon_0(k_F) + \Delta(k_F, \varepsilon^{(0)}) + (\hbar^2 k_F/m)[1 + (m/\hbar^2 k_F) \times (\partial\Delta(k, \varepsilon^{(0)})/\partial k)|_{k=k_F}](k - k_F) + \cdots] \quad [17]$$

Comparing this result with eqn [15], an expression for an effective mass is given by

$$\frac{1}{m_{\text{eff}}} = \left[\frac{1}{m} + \frac{1}{\hbar^2 k_F}\frac{\partial\Delta(k, \varepsilon^{(0)})}{\partial k}\bigg|_{k=k_F}\right] \times \left\{\left[1 - \frac{\partial\Delta(k_F, \varepsilon)}{\partial\varepsilon}\right]_{\varepsilon=0}\right\}^{-1} \quad [18]$$

In the case where the band structure is of dominant interest in interpreting an effective mass, one can write (see eqn [14])

$$\frac{1}{m_{\text{eff}}} = \left[\frac{1}{\hbar^2 k_F}\frac{\partial E(k, 0)}{\partial k}\bigg|_{k=k_F}\right]\left\{\left[1 - \frac{\partial\Delta(k_F, \varepsilon)}{\partial\varepsilon}\right]_{\varepsilon=0}\right\}^{-1} \quad [19]$$

which is inversely proportional to the "slope" of $E(k,0)$, the electron dispersion at the Fermi surface. Here, $\Delta(k_F, \varepsilon)$ represents all perturbing terms beyond the band approximation (phonons, etc.). This should be compared to the semiconductor case in which proportionality to the inverse "curvature" at the band extrema determines the effective mass.

Electron Correlations: Hartree–Fock

The simplest case for calculating an effective mass arising from electron–electron interactions is to consider the contribution to $\Gamma[\boldsymbol{k}, \varepsilon(\boldsymbol{k})]$ in the Hartree–Fock approximation. The widely known Hartree–Fock result has only a real part

$$\Delta(k, 0) = -\frac{\hbar^2 e^2 k_F}{2\pi}\left[2 + \frac{(k_F^2 - k^2)}{k k_F}\ln\left|\frac{k + k_F}{k - k_F}\right|\right] \quad [20]$$

This approximation, although far from an adequate picture of electrons in metals, provides an example to which one can apply eqn [18]. Differentiating eqn [20] as required gives the nonsensical result:

$$m_{\text{eff}} = 0 \quad [21]$$

which can be interpreted as an indication of just how flawed the Hartree–Fock approximation is for the degenerate Fermi gas.

In the case of exotic heavy fermion alloys, for example, CeB_6, $CeCu_2$, which have effective masses 10–100 times those of bare electron masses, the mixture of strong on-site *f*-state correlations and hybridization between *s*-states and *f*-states requires a description in terms of strongly hybridized bands. These hybridized bands have extremely large "slopes" at the Fermi surface due to the large *f*-state interactions having forced open a small energy gap just above the Fermi energy. This results in flattening the occupied conduction band to an almost horizontal conformation at k_F, so that according to eqn [19], an unusually large effective mass is to be expected.

Density-of-States Effective Mass

Although optical experiments are capable of probing a band structure with considerable detail and can even determine effective mass tensor components, thermodynamic experiments (e.g., specific heat) do not have this capability. For example, the low-temperature specific heat in Fermi systems is linear in temperature (its signature):

$$C_V = \frac{\pi^2}{3}k_B^2 \mathscr{D}(E_F)T \quad [22]$$

and proportional to $\mathscr{D}(E_F)$, the "density of electronic states" at the Fermi energy E_F, where T is temperature in K and k_B is the Boltzmann constant. The density of states is able to make no distinctions about the detailed geometry of the electron energy bands and thereby presents an effective mass interpretation which, at best, averages the mass tensor components in a specific way.

The density of electronic states is determined by "sifting" through all the electrons in a band and counting those that have a particular energy. The sifting and counting process is repeated for all bands. This process is succinctly described by the expression

$$\mathscr{D}(\varepsilon) = 2\sum_{\boldsymbol{k}}\delta[\varepsilon - E(\boldsymbol{k})] \quad [23]$$

Note how the Dirac δ-function combs through the band electrons and registers a count when it encounters one with energy ε. It then continues to

comb and count. The factor of 2 accounts for spin degeneracy. This is the simple meaning of the density of states.

Now, to see how the effective mass enters into the density of states, evaluate eqn [23] for a single, simple isotropic band such as represented by eqn [2]. Using the fact that in three dimensions

$$\sum_{\boldsymbol{k}} \rightarrow \frac{V}{(2\pi)^3}\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{\infty} \mathrm{d}k\, k^2 \qquad [24]$$

eqn [23] becomes

$$\mathscr{D}_{\mathrm{c}}(\varepsilon) = \frac{2V}{(2\pi)^3}\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{\infty} \mathrm{d}k \times k^2 \delta\left[\varepsilon - \left(E_{\gamma \mathrm{c}} + \frac{\hbar^2 k^2}{2m_{\mathrm{eff}}}\right)\right] \qquad [25]$$

for the choice of an isotropic conduction band in an effective mass approximation,

$$E_{\mathrm{c}}(\boldsymbol{k}, 0) = E_{\gamma \mathrm{c}} + \frac{\hbar^2 k^2}{2m_{\mathrm{eff}}} \qquad [26]$$

where $E_{\gamma \mathrm{c}}$ is a bandgap energy.

Using the Dirac identity

$$\delta[f(k)] = \sum_{k_{\mathrm{R}}} \frac{\delta(k - k_{\mathrm{R}})}{\left[\left|\frac{\partial f(k)}{\partial k}\right|_{k=k_{\mathrm{R}}}\right]} \qquad [27]$$

where the values of k_{R} are determined by the solutions of $f(k_{\mathrm{R}}) = 0$, one gets the density of states per unit volume:

$$\frac{\mathscr{D}_{\mathrm{c}}(\varepsilon)}{V} = \frac{m_{\mathrm{eff}}^{3/2}\sqrt{2}}{\pi^2\hbar^3}\sqrt{\varepsilon - E_{\gamma \mathrm{c}}}, \quad \varepsilon > E_{\gamma \mathrm{c}} \qquad [28]$$

This isotropic case does not, of course, expose any of the averaging that will take place in the anisotropic case. To understand better what kind of averaging this is, the ellipsoidal band case is taken as

$$E(\boldsymbol{k}) = E_{\gamma} + \frac{\hbar^2}{2}\left(\frac{k_x^2 + k_y^2}{m_{\perp}} + \frac{k_z^2}{m_{\mathrm{II}}}\right) \qquad [29]$$

Following eqn [23] but this time integrating in cylindrical coordinates, one gets, for a conduction band,

$$\frac{\mathscr{D}_{\mathrm{c}}(\varepsilon)}{V} = \frac{(m_{\perp} m_{\perp} m_{\mathrm{II}})^{1/3}\sqrt{2}}{\pi^2\hbar^3}\sqrt{\varepsilon - E_{\gamma \mathrm{c}}}, \quad \varepsilon > E_{\gamma \mathrm{c}} \qquad [30]$$

which defines the averaging of effective mass tensor components imposed by the density-of-states effective mass. For an ellipsoidal valence band, one obtains

$$\frac{\mathscr{D}_{v}(\varepsilon)}{V} = \sqrt{2}\frac{(m_{\perp} m_{\perp} m_{\mathrm{II}})^{1/3}}{\pi^2\hbar^3}\sqrt{-\varepsilon}, \quad \varepsilon < 0 \qquad [31]$$

It is now interesting to note that the thermodynamic implication of the Hartree–Fock result of eqn [23] is that C_V (Hartree–Fock) $= 0$, which continues to point to the inadequacy of that approximation.

Electron–Phonon Interactions

After explicitly evaluating the density of states for an isotropic band structure, the constant volume specific heat for a Fermi gas, eqn [22], becomes

$$C_V = \frac{k_{\mathrm{B}}^2 m_{\mathrm{eff}} k_{\mathrm{F}}}{\hbar^2} T \qquad [32]$$

As is well known, measured specific heats generally exceed the free electron values even for simple metals in the alkali series, which have nearly free electron bands. By determining an effective mass which arises from electron–phonon interactions, it can be seen that a fraction of this excess specific heat can be attributed to electron–phonon interactions. One can use eqns [11] and [12] to explore this case by referring to the well-known result that for electron–phonon interactions in a Fermi gas,

$$\Delta_{\mathrm{e-p}}(k_{\mathrm{F}}, \varepsilon) = -\lambda\varepsilon \qquad [33]$$

where ε is the self-consistent "quasiparticle" energy that is generally characteristic of $\Delta(\boldsymbol{k}, \varepsilon)$ and

$$\lambda = 2\int_0^{\infty} \mathrm{d}\omega \frac{\alpha^2(\omega) F(\omega)}{\omega} \qquad [34]$$

is a measure of the electron–phonon coupling strength, with $\alpha^2(\omega)$ the electron–phonon coupling function and $F(\omega)$ the phonon density of states.

Returning to eqn [18] (ignoring band structure effects),

$$\varepsilon(k) = \frac{\varepsilon_0(k_{\mathrm{F}})}{[1 - (\partial\Delta_{\mathrm{e-p}}(k_{\mathrm{F}}, \varepsilon)/\partial\varepsilon)]_{\varepsilon=0}} = \frac{\hbar^2 k_{\mathrm{F}}^2}{2m(1+\lambda)} \qquad [35]$$

so that an electron–phonon enhanced effective mass may be identified as

$$m_{\mathrm{eff}} = m(1+\lambda) \qquad [36]$$

Other Effective Mass Parametrizations

An effective mass is used to parametrize interacting electron behavior in other measurements. For example,

1. Electrical conductivity

$$\sigma = \frac{ne^2\tau}{m_{\text{eff}}} \qquad [37]$$

   where τ is a transport relaxation time.
2. The theory of excitons and interband optical absorption, in which a reduced effective mass μ is defined:

$$\frac{1}{\mu} = \frac{1}{m_{\text{eff,c}}} + \frac{1}{m_{\text{eff,h}}} \qquad [38]$$

   in order to parametrize the excitonic Rydberg states. Here, $m_{\text{eff,c}}$ is a conduction band effective mass and $m_{\text{eff,h}}$ is a valence band effective mass.
3. The de Haas–van Alphen magneto-oscillations, where the temperature-dependent amplitude of the rth harmonic $r = 1, 2, \ldots$, is

$$A_r = \left\{ \sinh\left[\frac{2\pi^2 r k_{\text{B}} m_{\text{eff}} T}{ehB} \right] \right\}^{-1} \qquad [39]$$

The de Haas–van Alphen effective mass is almost identical to the density-of-states effective mass which appears in the specific heat:

$$m_{\text{eff}}(C_V) = m(1 + \lambda_{\text{e-p}} + \cdots) \qquad [40]$$

except that the de Haas effective mass refers to only one extremal areal slice of the Fermi surface perpendicular to the magnetic field direction, whereas the specific heat effective mass is averaged over the entire Fermi surface.

There are experiments for which some kind of effective mass description seems appropriate, but for which there is no rigorously applicable or meaningful theory. For example, the plasma frequency

$$\omega_{\text{p}}^2 = \frac{ne^2}{m_{\text{eff}}\varepsilon_0} \qquad [41]$$

which can be determined from optical or electron stopping power experiments (and a Hall measurement to determine the carrier concentration n) is parametrized by an effective mass. But the plasma frequency is a collective excitation of all available mobile electrons so that effective mass calculations of the type described above have little relevance. In those cases, phenomenology rules.

Comparing effective masses from different types of experiments may have to be done with great care and with the knowledge of the theoretical origins of the parameter in each specific case.

See also: Electron–Phonon Interactions and the Response of Polarons.

PACS: 71.18.+y

Further Reading

Abrikosov AG, Gor'kov LP, and Dzyaloshinskii IY (1965) *Quantum Field Theoretical Methods in Statistical Physics*. Oxford: Pergamon.
Ibach H and Luth H (1990) *Solid-State Physics*. Berlin: Springer.
Kittel C (1987) *Quantum Theory of Solids*. New York: Wiley.
Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.
Knox RS (1963) *Theory of Excitons*. New York: Academic Press.
Marder MP (2000) *Condensed Matter Physics*. New York: Wiley.
Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.
Wasserman A and Springford M (1996) *Advances in Physics* 45: 471–503.
Ziman J (1964) *Principles of the Theory of Solids*. Cambridge: Cambridge University Press.

Elastic Behavior *See* Mechanical Properties: Elastic Behavior.

Elasticity in Lattices *See* Lattice Dynamics: Structural Instability and Soft Modes.

Electric and Magnetic Fields in Cells and Tissues

J Gimsa and L Haberland, Universität Rostock, Rostock, Germany

Introduction

Electromagnetic phenomena play an important role at all hierarchical and structural levels of biological organisms, and even in interactions among organisms and their orientation within the biosphere. Electromagnetic fields in cells and tissues may be inherent to the biological system or the result of natural or man-made exposure to such fields. Fields may interact with electrically passive or active biological structures. In both cases, the response by specific mechanisms will depend on field frequency and intensity. Whereas external fields may range from electrostatic fields up to Gamma-irradiation, the generation of fields in cells is limited in frequency range and intensity. The frequencies of physiologically generated electromagnetic fields range from 0 Hz for static (DC) fields up to light frequencies for fluorescence effects driven by captured photons or ATP, as is the case with lightning bugs.

In this article, the electric and magnetic properties of biological matter for frequencies ranging from DC up to GHz are considered. The main focus is on electric phenomena, since the interaction with magnetic fields is usually weak for the generally low-magnetic susceptibility of biological matter. Impedance and AC electrokinetic techniques as methods that apply fields for exploring the passive properties of tissues and suspended cells are also considered. To focus on tissues and cells, specific mechanisms discussed in relation to speculations of possible harmful effects of electromagnetic fields are not considered. Neither are fields beyond GHz in the far infrared or visible light range considered.

Mechanisms and Frequencies of Field Generation and Interaction

At the atomic and molecular levels, static electric fields associated with ions and dipole molecules are important for intermolecular forces. At the next higher structural level, static fields are found in electric double layers, such as at the interfaces of membranes with aqueous electrolytes. Whereas these fields are part of the thermodynamic (electrochemical) equilibrium at the interface, the (nonequilibrium) physiological transmembrane potential of living cells is built up by ATP-driven ion pump proteins. Other nonequilibrium processes are found in electrorheological effects as motion of fluids in or around cells and tissues, and in the deformation of tissue leading to a piezoeffect, as in bones, for instance. Even though the transmembrane potential cannot usually be measured in the external medium, in some cases, static fields can be detected around single cells or even organisms, for example, the DC potentials between various parts of an animal body or of plants. In lesions, wound potentials are found.

Cell excitation, as in nerve and muscle cells, is based on active electric cell properties. Voltage-gated, ion-channel proteins play a key role in the generation of membrane-potential spikes. Spike frequencies reach the kHz range. In excitable tissues, such as the brain, the so-called "field potentials," superimposed potentials generated by a great number of cells, are found. Such potentials are detected, for example, in medical diagnostics in electromyograms, electrocardiograms, or electroencephalograms (EMG, ECG, or EEG). Between limbs, ECG potentials of several mV are detected.

Even though mechanical oscillations in the lower radio-frequency range are common for ultrasound orientation of animals, the emission of electromagnetic fields by biological systems in the radio-frequency range is unknown. Nevertheless, the hair cells of the inner ears of these animals or the sound-detecting cells in insects must, of course, be capable of electrically processing these frequencies.

Beyond the radio-frequency range, spectroscopic methods detect molecular properties. In the microwave and far infrared regions, energy transitions (the absorption and emission of energy quanta) are particularly related to the rotation of molecules. Oscillations inherent to the molecules possess higher energies and are subsequently detected in the near infrared range. The highest energies are associated with the electrons of the molecules. The corresponding energy transitions are mediated by the quanta of the visible and ultraviolet optical ranges.

It is generally believed that molecular oscillations are not synchronous and thus generate localized fields of low strength. Nevertheless, these fields mediate inter- and intra-molecular interaction forces (van der Waals forces). Since physiological energy is steadily degraded by chemical reactions which leads to warming, living organisms radiate electromagnetic energy of the infrared-frequency range into the environment.

The exploration of the properties of biological matter has always been connected to medical and

biotech applications of electromagnetic fields. For the response of cells to electric fields, both the passive properties of the different media as well as the active properties of various membrane proteins, such as ion pumps and exchangers, are of importance. The patch-clamp technique has revolutionized the exploration of protein channels, another class of membrane proteins. Muscle and nerve stimulations as well as hyperthermia are well established in medical therapy. Electric cell transfection by foreign DNA and the production of hybridoma cells by fusion are based on membrane poration by high-induced transmembrane potentials. New biotech applications such as the development of a lab-on-chip technology are based on a combination of single-cell techniques and AC-electrokinetics.

Passive Electric Properties of Biological Matter

An electric field E passing through a medium is weakened by a factor that is given by the absolute value of its relative permittivity $|\varepsilon^*|$. The electric displacement D^* is constant:

$$D^* = \varepsilon^* \varepsilon_0 E \qquad [1]$$

ε_0 stands for the permittivity of vacuum. The asterisk denotes complex parameters. For such parameters, a phase shift may appear in between the exciting field or voltage and the resulting displacement or current. As a result, the material parameters, for example, ε^*, consist of a real (ε'-in-phase) and an imaginary (ε''-out-of-phase) part. The capacitance of a medium with a relative permittivity ε^* that is located in between two parallel electrodes of distance d and area A is

$$C^* = \frac{1}{\varepsilon^* \varepsilon_0} \frac{A}{d} \qquad [2]$$

For thin films, such as membranes, an area specific capacitance $C_m = C/A$ can be defined. When C_m values are given in the literature, in most cases, a possible frequency dependence is neglected. Often the passive electric properties of biological matter are described by its impedance Z^* that is the generalized, frequency-dependent resistance:

$$Z^* = \frac{1}{\sigma^*} \frac{d}{A} \qquad [3]$$

describes the impedance of a medium with a specific conductivity σ^* that is located in between two parallel electrodes of distance d and area A. In the simplest case, the medium is homogeneous and isotropic like aqueous electrolytes. In contrast, biological tissue is structured, for example, by cell membranes. This compartmentalization is the reason for structural dispersions.

Dispersions

The impedance of biological material is characterized by a variety of characteristic frequency-dependent changes. These changes are based on the fact that permittivity contributions of certain relaxation processes, which follow the external field at lower frequencies, disperse at higher ones. Consequently, the permittivity of biological material drops by several orders of magnitude over some frequency decades. A permittivity drop at a given frequency corresponds to a certain conductivity increase, since the field-induced movement of bound charges can no longer be distinguished from the movement of charges, such as ions, that are already free to move at DC.

The aim of the research was to assign the measured dispersions to certain biological structures on different hierarchical levels and to classify the contributing processes. For classification, two different approaches were chosen – the dispersions were either sorted according to their physical nature, for example, as Maxwell–Wagner and Debye dispersions, or according to the frequency range where they occur. While Maxwell–Wagner dispersions are based on the structuration of the material, Debye dispersions are related to the orientation of molecular dipoles. A classification according to frequencies yielded a scaling where the dispersions with increasing frequencies were assigned to the α-, β-, and γ-range (**Figure 1**). It was assumed that well-defined processes are responsible in a certain dispersion range.

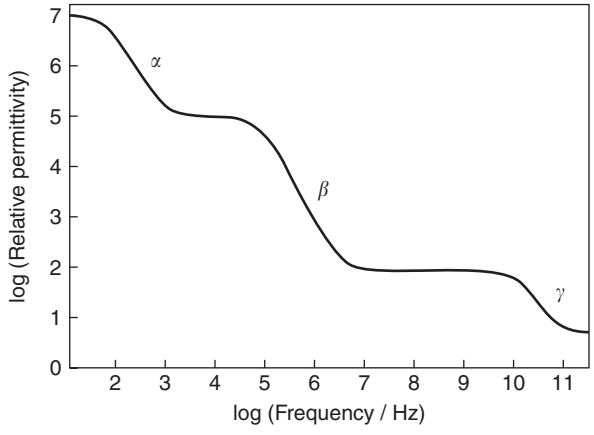


Figure 1 Schematic frequency dependence of the relative permittivity of tissue or cell suspensions according to H P Schwan (1957) Electrical properties of tissue and cell suspensions. In: Lawrence JH and Tobias CA (eds.) *Advances in Biological and Medical Physics*, pp. 147–209. New York: Academic Press.

Despite their low frequencies, analysis of α-dispersions in the sub-1 kHz frequency range is especially difficult. Many different processes may overlap, including such processes as electrode processes, hydrodynamic relaxation of electroosmotically induced convections within the measuring chamber and around the cells, membrane channels, cell orientation and deformation, as well as cell electrophoresis. The difficulty is that these phenomena are hard to separate, and concurrently influence suspension medium and cells (**Figure 2**).

In tissues and cell suspensions, the major contributions of the β-dispersion stem from structural polarizations as those of the cytoplasmic membrane or of internal membrane systems, as well as the polarization of the cytoplasm. It is to be noted that membrane and cytoplasmic dispersions are of different qualities. Whereas membrane dispersions are based on the capacitive bridging of the thin membrane film, the cytoplasmic dispersions mediate the switching from a field distribution, according to the conductivities, to a distribution according to the permittivities of the bulk

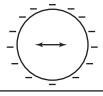
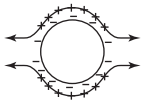
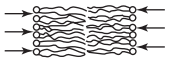
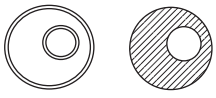
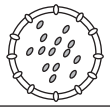
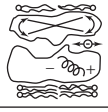
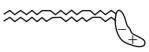
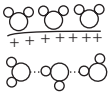
Structural level "theory"	Sketch of structure	Dispersing effect
Microscopic forces - Tissue - Cell - Membrane - Electric double layer - Intercellular fluid - Cytomechanics - Hydrodynamics - Double layer theory		- Cell deformation
		- Cell orientation
		- Cell electrophoresis
		- Electroosmosis - Chemical ion adsorption
		- Electrostriction of membranes
Structural effects - Membranes - Compartment geometry - Media - Laplace equation		- Membrane polarization - Bulk polarization
Molecular effects - Molecules - Side chains - Water cluster polarization - Bound water - Bulk water - Biochemical reactions - Dispersion theory - Debye theory - Reaction kinetics - Molecular models		- Debye polarization of membrane & cytoplasmic proteins
		- Protein transport - Protein conformation - Side chain orientation
		- Orientation & flexing of lipid head groups
		- Dissociation/association
		- Water cluster & molecule orientation - Localization of partial charge

Figure 2 Polarization processes based on various biological structures. The size of the polarized structures and the physical nature of the dispersion processes determine their characteristic times and frequency dependencies. The first column summarizes the structural levels and theoretical approaches for describing dispersion processes. The sketches in the second column present examples for processes that are designated in the third column.

media. The structural dispersions may be influenced by the dispersion of other processes, such as the surface conductance or dipole relaxation of large molecules. In cellular systems, protein dispersions usually occurring around 1 MHz are masked by the structural membrane dispersions.

In the γ-range above 100 MHz, quick dissociation–association relaxation or Debye dispersions of small molecules or charged groups must be expected. The dielectric properties of proteins are believed to be important for their physiological function such as protein-protein association or the interaction with charged ligands. Dispersions between 200 MHz and 10 GHz were assigned to bound water and small molecules (δ-dispersion). Dispersion over 10 GHz can mainly be assigned to free water.

Properties of Aqueous Media

Aqueous Electrolytes

The intracellular space and the cytoplasm can be considered to be aqueous solutions of ions and organic molecules. Up to GHz frequencies, their properties are often simplified to those of aqueous electrolytes that can be described by frequency-independent specific conductivities and permittivities. For electrolytes, the conductivity σ and permittivity ε are based on ionic conduction and the permittivity of water, respectively. Whereas the permittivity of water of 80 at 18°C shows a slight decrease with temperature and ion concentration, the conductivity may vary by orders of magnitude. For solutions of one salt, it can be calculated from salt concentration and equivalent conductivity. These relations are more complex for mixed solutions. Accordingly, the complex specific conductivity σ^* consists of the frequency-independent contribution of ion conduction and frequency-dependent displacement currents based on the water permittivity

$$\sigma^* = \sigma + \mathrm{j}\omega\,\varepsilon\,\varepsilon_0 \qquad [4]$$

where ω and j stand for the circular frequency of the field and $\sqrt{(-1)}$, respectively. After introduction of eqn [4], eqn [3] describes a system that consists of a resistor (for the ion current) in parallel to a capacitor (for the displacement current). Accordingly, the impedance of the medium steadily decreases with frequency.

Cytosolic Proteins

The cytoplasm of most cells is extremely rich in proteins influencing the electric properties. Proteins reduce the water concentration resulting in a reduced ion mobility and conductivity. The reduction factor may be higher than 2 in red blood cells. A second contribution comes from the proteins themselves. Because of their high dipole moments, they are the cause for extraordinary high permittivities at low frequencies. Nevertheless, their bulky structure is incapable of following the orientation force of an external field beyond a certain frequency. This is the reason for dielectric dispersions, that is, the decrease of ε by the dielectric decrement, and a corresponding increase in σ, which are typically in the frequency range $\sim$1 MHz. The dispersion in pure protein suspensions can be considered to be Debye dispersions which were extensively investigated. In tissues and cell suspensions, protein dispersions are masked by structural dispersions. Therefore, dispersions of cytoplasmic proteins were predominantly investigated after cell lysis or in cell-free suspensions.

Passive Membrane Properties

Area Specific Capacitance

Membranes are known to be inhomogeneous and anisotropic. For a membrane, solely composed of a lipid double layer, a specific capacitance of $\sim 7.5 \times 10^{-3}\,\mathrm{F\,m^{-2}}$ can be estimated. Nevertheless, biological membranes are rich in membrane spanning proteins. Thus, protein and lipid areas form a parallel circuit of two capacitors, resulting in a specific capacitance of about $10^{-2}\,\mathrm{F\,m^{-2}}$ at frequencies below 1 MHz. Accordingly, protein regions present areas of higher specific capacitance. The various regions of the membrane, that is, bound water layers, lipid headgroup regions, the hydrophobic lipid chain regions as well as the orthogonal segments formed by transmembrane proteins, will exhibit a different frequency-dependent dielectric behavior. The glycocalix surface layer made up by charge-carrying polysaccharides that are bound to membrane lipids and proteins is of special importance for the α-dispersion.

Lipid Region

The lipid regions of membranes have an overall thickness of 4–5 nm with roughly 0.8 nm at each side being occupied by the headgroups. The hydrophilic headgroup regions contain the strong P–N dipoles, the ester bonds of the fatty acids and the glycerol backbones. Together with the associated water molecules, dipole potentials as high as 240 mV (positive at the membrane surface) are generated. In normal direction, the relative permittivity drops from $\sim$78 in the bulk water to $\sim$2 in the fatty acid chain

region. The permittivity of approximately 30 in the headgroup region disperses with a decrement of ~28. In a tangential direction, experimental and theoretical results hint on relative permittivities as high as 300 to 400 with a dispersion that is related to the in-plane orientation of the headgroups. For both orientations, dispersion occurs ~50 MHz resulting in a high-frequency permittivity of ~2. Due to its low permittivity already at low frequencies, the fatty acid chain region is highly hydrophobic, forming the major energetic barrier for the permeation of ions and small polar molecules.

Membrane Proteins

Intercalation of proteins in the membrane relies on a microenvironment of specific dielectric properties. Nonpolar amino acids form the protein surface where it is exposed to the lipid chain region. Highly polar protein groups as loops and tails of the polypeptide chains are exposed to aqueous surfaces. They possess a higher degree of molecular freedom suggesting higher permittivities. The peptide bonds of the membrane part of most proteins are very rigid. Their highly restricted mobility results in a low permittivity. Up to low kHz frequencies, the protein mediated ion transport is also of importance for electric protein properties.

Bound Water

The thickness of the bound water layers is ~0.5 nm, with 11–16 water molecules per lipid. About 4 water molecules per lipid are tightly bound and strongly oriented. More distant water molecules show a smoothly decreasing orientation profile. The average permittivity of the bound water layer is ~60. The dispersion frequency will depend on the binding strength with higher frequencies for lower binding strengths.

Impedance and AC-Electrokinetic Phenomena

Impedance experiments provided the first proof for the existence of biological membranes. Impedance experiments analyze the frequency-dependent electric response to an applied field to explore dielectric properties of tissues and cells. The electric response is based on the polarization of the sample structures, like suspended cells. Single-cell polarization may lead to orienting, deforming, moving, or rotating in AC-electrokinetics. The different motions depend on the different spatial properties of the field determining its interaction with the polarized cell. The frequency dependence of each effect is related to the frequency dependence of the polarizability of the various constituents of the cell. All effects are exploited in cell characterization methods: electroorientation, electrodeformation, dielectrophoresis, and electrorotation. For modeling the effects, the induced-cell dipole moment is obtained by integrating over the local fields of all subcellular structures in the cell volume. As in the case of impedance, models for multishell spherical, cylindrical, and ellipsoidal geometries are readily available.

Generally, the force experienced by the cell is related to the induced dipole moment which is the volume integral of the polarization of the cell. For an ellipsoidal cell with confocal layers, for each field frequency, a Maxwellian equivalent body can be found that exhibits the same complex dipole moment m^*. The equivalent body has homogeneous medium properties and a constant local field. The induced dipole moment reduces to

$$m^* = (m_x \;\; m_y \;\; m_z) = \varepsilon_0 \varepsilon_e V f^*_{CM} E \qquad [5]$$

Its frequency-dependent part is described by the Clausius–Mossotti factor f^*_{CM}. The x-component of f^*_{CM} is given by

$$f^*_{CMx} = \frac{\varepsilon^*_i - \varepsilon^*_e}{\varepsilon^*_e + (\varepsilon^*_i - \varepsilon^*_e) n_x} \qquad [6]$$

The permittivities of the external and internal media are marked by indices e and i, respectively. n_x is the depolarizing factor in x-direction. Depending on the ellipsoidal shape, n_x can vary from 0 to 1. n_x is 1/3 for a sphere since the sum of the three factors along the three principal axes is unity.

AC-electrokinetic methods yield similar information as does impedance, though with a higher resolution. A rotating field induces a dipole moment which rotates at the angular frequency of the external field. A spatial phase shift of the external field vector and the induced dipole moment, caused by any dispersion process, gives rise to a torque generating individual cell rotation. For example, the membrane dispersion generates a cell-typical electrorotation peak. Interestingly, the torque and therefore cell rotation is at maximum if relaxation time of the dispersion process and external field frequency match ($\omega = \omega_c$, see below). For their high-resolution, dielectric single cell methods are of special value for cell characterization and the determination of the field distribution at the subcellular level. All AC-electrokinetic phenomena can also be used for cell or molecule

manipulation. They form the basis for the lab-on-chip technology.

The Induced Transmembrane Potential ($\Delta\phi$)

Electric cell transfection and cell fusion are routine biotech applications based on membrane poration by high-induced $\Delta\phi$. For poration, DC or AC fields have to be applied for a sufficiently long time or at a sufficiently low frequency, respectively. Pore induction is a strongly nonlinear process mimicking a "critical" transmembrane voltage of $\sim 1\,\mathrm{V}$ corresponding to $\sim 200\,\mathrm{MV\,m^{-1}}$ in the lipid phase. $\Delta\phi$ is proportional to the external field strength and depends on cell size, shape, orientation, and electric cell properties.

Physicists deal with the potential induced on the surface to describe the dependence of the local field deviation on the shape of the bodies or cavities. H Fricke was the first to apply the physical knowledge on the depolarizing factors of ellipsoidal bodies or cavities to express the induced transmembrane voltage $\Delta\phi$. For an ellipsoidal cell with semi-axes a, b, and c at the pole of a semi-axis, a $\Delta\phi$ for a cell of the general ellipsoidal shape is given by

$$\Delta\phi = \frac{1}{1-n} aE = a_i aE \qquad [7]$$

E, a, n, and a_i stand for the external field strength, the length of the semi-axis oriented in field direction, the depolarizing factor along semi-axis a, and the relative influential radius, respectively. The equation applies for the DC steady state and a cell with zero-membrane conductance. The parameter "influential radius" ($a_i^* a$) depends solely on the shape of the ellipsoidal body. For a vacuum body, it describes the distance of the symmetry plane of the body to that equipotential plane in the absence of the body that is touching the pole in its presence.

For a sphere, $a = b = c$, and $n = 1/3$ leading to the well-known expression: $\Delta\phi = 1.5rE$. For axisymmetric cells $a = b$. The symmetry axis is c. For such spheroidal cells limited in shape by an infinitely thin disk and an infinitely long cylinder, respectively, the shape dependence can be sufficiently good, described by

$$\Delta\phi = \frac{a+2c}{a+c} aE \qquad [8]$$

and

$$\Delta\phi = \frac{a+2c}{2} E \qquad [9]$$

for a field oriented in a-direction and c-direction, respectively. At higher frequencies, the membrane impedance decreases by capacitive bridging. The characteristic frequency ω_c at which the induced potential is decreased by $-3\,\mathrm{dB}$ is given by

$$\omega_c = \frac{1}{aC_m}\left(\frac{\sigma_e\sigma_i}{\sigma_e + (a_i - 1)\sigma_i} + ag\right) \qquad [10]$$

σ_i, σ_e, and g stand for the internal and external conductivities, and the area-specific membrane conductance in $\mathrm{S\,m^{-2}}$. The complete equation for $\Delta\phi$ at a pole of an ellipsoidal cell is

$$\Delta\phi = \frac{a_i a}{\left(1 + ag\left(\frac{1}{\sigma_i} + \frac{a_i - 1}{\sigma_e}\right)\right)\sqrt{1 + \frac{\omega^2}{\omega_c^2}}} E \qquad [11]$$

with ω standing for the field frequency. Assuming a superposition of the $\Delta\phi$ and the physiological $\Delta\psi$ (see below), is the simplest way to explain the asymmetry in cell poration often found after dielectric membrane breakdown.

Absorption

An exposed system is only influenced by that part of the electromagnetic field energy that is absorbed (principle of Grotthus and Draper). For this reason, research on the impact of electromagnetic fields must focus on the mechanisms of absorption and the identification of the absorbing structures. Generally, the energy input into a given volume is described by Ohmic heating that is proportional to the square of the local field strength and the effective, specific conductivity of a certain medium. Magnetic absorption processes, also those based on quantum effects, such as on the spins of unpaired (radical-) electrons, can usually be neglected. Dispersions lead to subcellular field redistributions. Up to GHz-range frequencies, the local field distribution and thus local absorption greatly depend on the geometric structure and compartmentalization of a cell.

The local power dissipation (Ohmic heating) per volume element P/V in an AC-field is not directly proportional to the applied field but depends on the effective local field E_{eff}

$$\frac{P}{V} = \sigma E_{\mathrm{eff}}^2 \qquad [12]$$

The equation suggests that the power dissipation is described by a frequency-independent specific-DC conductivity. In practice, the specific conductivity of structured media increases with frequency in correspondence to the permittivity decline given in **Figure 1**.

Considerations show that the field strength in the external medium and thus absorption in the vicinity of the cell strongly depends on field frequency even for the simple single-shell geometry. The picture dramatically deviates from the classic one when assuming more realistic membrane and cytoplasmic properties based on their molecular structure. It suggests a possible role of the membrane in field absorption even at frequencies far above the membrane-dispersion frequency.

The Physiological Transmembrane Potential ($\Delta\Psi$)

Nerve excitation or muscle activity generate AC fields of the Hz and kHz ranges. The physiological transmembrane potential $\Delta\Psi$ is of special importance for this phenomenon. $\Delta\Psi$ generation is based on the special electric properties of biological membranes. $\Delta\Psi$ is defined as the difference of the potentials of the internal (index i) and external (index e) bulk media:

$$\Delta\Psi = \Psi_i - \Psi_e \qquad [13]$$

Because of the surface potentials at the two membrane sides, fixed charges in the glycocalix and the dipole potentials of the lipid headgroups, the actual potential difference across the lipid phase of the membrane may strongly deviate from the $\Delta\Psi$. At specialized membranes such as the thylakoid membranes of plant chloroplasts, $\Delta\Psi$ may be higher than $-200\,\text{mV}$. Accordingly, the field strengths across the membrane under physiological conditions may be well above $10\,\text{MV}\,\text{m}^{-1}$. The membrane properties allow for the storage and transformation of electrochemical energy into chemical energy such as ATP, or in electrical energy as in nerve excitation or the electrical organs of fishes. $\Delta\Psi$ for resting neurons or muscle cells are in the range of -100 to $-40\,\text{mV}$. Action potentials are typically in the range of 30 to 80 mV.

In excitable cells of vertebrate species, sodium and potassium ions are actively pumped outwardly and inwardly, respectively, leading to a roughly inverse distribution for the two ion species. In resting cells, the membrane permeability for potassium is much higher than for sodium. Consequently, $\Delta\Psi$ is determined by the potassium distribution. Since its concentration is much higher in the cytoplasm, random diffusion across the membrane via potassium channels leads to a net potassium loss and a negative charging of the cytoplasm. Nerve cells are rich in voltage-gated ion-channel proteins. The transition from resting to action potential is driven by a sequential change in the membrane permeabilities for sodium ions and potassium ions. The relations are described by the Goldman equation:

$$\Delta\Psi = \frac{RT}{F}\ln\frac{P_{Cl}a^{i}_{Cl} + P_{K}a^{e}_{K} + P_{Na}a^{e}_{Na}}{P_{Cl}a^{e}_{Cl} + P_{K}a^{i}_{K} + P_{Na}a^{i}_{Na}} \qquad [14]$$

R, T, F, P, and a stand for the gas constant, the absolute temperature, the Faraday constant, permeability, and ion activity, respectively. Indices i, e, Na, K, and Cl designate internal and external bulk media, as well as the ion species sodium, potassium, and chloride, respectively. Please note that for $P_K \gg P_{Na},\ P_{Cl}$ and $P_{Na} \gg P_K,\ P_{Cl}$, the Goldman equation reduces to the Nernst equations for potassium and sodium, respectively. A feedback loop where the local membrane field alters the ion permeabilities of voltage-gated ion-channel proteins is the basis for the propagation of cell excitation, for example, along the axon of a nerve cell.

Passive Magnetic Properties of Biological Matter

In analogy to the electric field, a static magnetic field of strength H causes a magnetic induction B of

$$B = \mu\mu_0 H \qquad [15]$$

where μ_0 and μ are the magnetic field constant and the relative permeability of the material, respectively. Since the permeabilities of biological matter are very close to unity, it is reasonable to list susceptibilities $\chi = \mu - 1$ (**Table 1**)

The low magnetic susceptibility is the reason for the very weak interaction of static magnetic fields with biological matter. Most biological cells and tissues are diamagnetic ($\chi < 0$). Paramagnetic ($\chi > 0$) deoxygenated red blood cells are an exception.

The high structuration of biological tissues and cells is reflected in their magnetic anisotropy $\Delta\chi$, which

Table 1 Magnetic susceptibility of different matter

Material	$\chi/10^{-6}$
Air	0.264
Water	−9.04
Arterial blood	−9.3
Oxygenated red blood cells	−9.03
Deoxygenated red blood cells	3.88
Venous blood	−7.8
Lungs (air-filled to 15%)	−4.2
Muscle	−9.03
Liver	−8.26
Bone	−10
Biogenic Fe_3O_4-crystals	10^{12}

describes the difference in the magnetic susceptibilities parallel and perpendicular to characteristic biological structures. This feature is especially pronounced in lipid membranes that can even be orientated by magnetic fields at flux densities larger than 1 T.

Biogenic magnetites (Fe_3O_4-crystals) are found in magnetostatic bacteria. The susceptibility of these ferromagnetic particles is as large as 10^6. The particles are the major component of magnetosomes, cell organells that are surrounded by a membrane. They enable the bacteria to orientate themselves along the vector of the geomagnetic field. To a lesser extent, biogenic magnetites have also been found in different animals and even in the human brain. Nevertheless, it is still not clear as to whether magnetic fields can interfere with physiological processes in humans. The mechanisms of magnetic orientation of animals are not fully understood, either.

Generation and Effects of Magnetic Fields

Of course, all biological processes which involve charge movement generate endogenous magnetic fields. Such processes are the transmembrane transport of ions or charged compounds as well as ion currents related to signal transduction in nerves and muscles. The displacement of ions and charged particles in the bloodstream or the lymphatic system induces streaming potentials and electric currents. Nevertheless, in no case does the induced magnetic flux densities exceed the pT-range (compared to 20–70 μT of the geomagnetic field). Currently, no hard experimental data are available for the coupling of external with endogenous magnetic fields, subsequently influencing biologically relevant processes.

Exogenous low-frequency magnetic fields can induce electrical (eddy) currents. Nevertheless, for effective induction, structures larger than single cells are required. Candidate structures are, therefore, the blood system and cells that are interconnected via gap junctions allowing for the formation of large induction loops. It can be estimated, that a flux density of at least 1 mT is necessary to effect biological functions in the brain or of the heart. Static and low-frequency fields may deflect moving charges by the Lorentz force. However, ions and charged compounds can only be significantly influenced at flux densities higher than 10^6 T. The inverse interaction, current induction in a nerve or blood vessel crossing magnetic field lines, may cause nerve excitations when the flux density exceeds 10 mT.

Magnetic fields can influence the spins of unpaired electrons, for example, of organic-label molecules, and nuclei with a spin different from zero, for example, of hydrogen. To orientate the spins, typically static-magnetic fields of 50 mT to 6 T, and 340 mT to 1.25 T are applied in nuclear magnetic resonance (NMR) for medical diagnostics and electron spin resonance (ESR), respectively. The field introduces an energy difference for the two possible spin orientations. The transition from the low-energy orientation to the higher one can be induced by resonant absorption of a radio-frequency field in the MHz to GHz range with a frequency corresponding to the energy difference. Either the absorption or emission of the radio-frequency field induced by the relaxation of the spin distribution is detected in the methods.

Static and low-frequency magnetic fields have been shown to influence chemical reactions which involve radical-pair intermediate states in electron transfer processes. The lifetime of triplet–singlet interconversions in such reactions have been altered at magnetic flux densities of 1 mT and beyond. This mechanism may as well apply to biochemical reactions. For the interaction of magnetic fields with biological processes, other mechanisms such as cyclotron resonance in protein channels and ion parametric resonance in macromolecules have also been proposed, but could not be verified.

See also: Bioelectronics; Biomembranes; Electromagnetic Biological Effects; Electrophysiology.

PACS: 01.30.Rr; 41.20.Cv; 41.20.Gz; 66.10.−x; 72.20.−l; 75.20.−g; 75.30.Gw; 75.70.Cn; 76.60.Jx; 77.22.−d; 84.37.+q; 87.15.−v; 87.50.−a; 87.50.Jk; 87.50.Mn; 87.50.Rr; 87.90.+y

Further Reading

Able KP (1994) Magnetic orientation and magnetoreception in birds. *Progress in Neurobiology* 42: 449–473.

Adey WR and Lawrence A (1984) *Nonlinear Electrodynamics in Biological Systems*. New York: Plenum.

Delgado AV (2001) *Interfacial Electrokinetics and Electrophoresis*. New York: Dekker.

Edmonds DT (2001) *Electricity and Magnetism in Biological Systems*. Oxford: Oxford University Press.

Glaser R (2001) *Biophysics*. Berlin: Springer.

Grandolfo M, Michaelsom SM, and Rindi A (1983) *Advances in Biological Effects and Dosimetry of Low Energy Electromagnetic Fields*. New York: Plenum.

Grimnes S and Martinsen O (2000) *Bioimpedance and Bioelectricity Basic*. San Diego: Academic Press.

Jones TB (1995) *Electromechanics of Particles*. Cambridge: Cambridge University Press.

Kortüm G (1962) *Kalorimetrie Photometrie und Spektrometrie*. Berlin: Springer. (in German).

Landau LD and Lifschitz EM (1985) *Elektrodynamik der Kontinua*, vol. 8. Berlin: Akademie-Verlag (in German).

Maret G, Boccara N, and Kiepenheuer J (1986) *Biophysical Effects of Steady Magnetic Fields*. Berlin: Springer.
Neumann E, Sowers A, and Jordan C (1989) *Electroporation and Electrofusion in Cell Biology*. New York: Plenum.
Polk C and Postow E (1996) *CRC Handbook of Biological Effects of Electromagnetic Fields*, 2nd edn. Boca Raton: CRC Press.
Schütt W, Klinkmann H, Lamprecht I, and Wilson T (1991) *Physical Characterization of Biological Cells*. Berlin: Verlag Gesundheit GmbH.

Nomenclature

a	length of semi-axis a (m)
a_i	relative influential radius (m)
a_x	ion activity ($\mathrm{mol\,l^{-1}}$)
A	area ($\mathrm{m^2}$)
B	magnetic flux density (T)
C	capacitance (F)
C_m	area specific capacitance ($\mathrm{F\,m^{-2}}$)
d	distance (m)
D	electric displacement ($\mathrm{C\,m^{-2}}$)
E	electric field strength ($\mathrm{V\,m^{-1}}$)
F	Faraday constant ($9.648\,456 \times 10^4$ $\mathrm{C\,mol^{-1}}$)
f_{CM}	Clausius–Mossotti factor
g	area specific conductance ($\mathrm{S\,m^{-2}}$)
m	dipole moment (Cm)
n	depolarizing factor
P	power (W)
P_x	ion permeability ($\mathrm{m\,s^{-1}}$)
R	gas constant ($8.314\,510\,\mathrm{JK^{-1}\,mol^{-1}}$)
T	temperature (K)
V	volume ($\mathrm{m^3}$)
Z	impedance (Ω)
$\Delta\phi$	induced transmembrane potential (V)
$\Delta\chi$	magnetic anisotropy
$\Delta\Psi$	physiological transmembrane potential (V)
ε	relative permittivity
ε_0	permittivity of vacuum (8.854×10^{-12} $\mathrm{C\,V^{-1}\,m^{-1}}$)
μ	magnetic permeability
μ_0	magnetic field constant (1.256×10^{-6} $\mathrm{Vs\,A^{-1}\,m^{-1}}$)
σ	specific conductivity ($\mathrm{S\,m^{-1}}$)
χ	magnetic susceptibility
ω	circular frequency (Hz)
Asterisks (*)	indicate complex terms

Electrodynamics: Continuous Media

A C Eringen, Princeton University, Littleton, CO, USA

Introduction

Electromagnetic (E-M) interactions with material bodies give rise to a large number of interesting phenomena. E-M loads (forces, couples, and energy) applied to material bodies produce mechanical effects (deformation, stress, strain, and heat). For example, under an applied electric field, an elastic solid deforms (develops stress). This is known as "electrostriction." This is a nonlinear phenomenon. There is a parallel effect under magnetic fields, called "magnetostriction." A thin rod or plate subjected to a sufficiently large magnetic field will buckle. Under E-M fields, a body may produce heat (pyroelectric and pyromagnetic effects). E-M fields also produce torques. Conversely, mechanical deformations produce E-M fields. A deforming dielectric produces electric polarization, known as "piezoelectricity." Thus, E-M and mechanical fields are intertwined to produce many "cross effects."

Electrodynamics of continua is essentially nonlinear in nature. This stems from the fact that E-M loads are nonlinear in E-M fields and bodies may undergo large deformations and motions, producing nonlinear accelerations and finite strains. However, in many cases, linear theories can be constructed under some restrictions. In the case of piezoelectricity, the linear theory has been successful. But, in the case of magnetohydrodynamics (MHD), the joule dissipation remains finite and accelerations contain convective nonlinear terms relevant to turbulent motions.

Electromagnetic continua is not the sum (superposition) of the E-M theory and the mechanical theory of continuous media. It is an interconnected, nonlinear composition of both.

The E-M theory of continuous media is based on three fundamental notions: (1) balance laws; (2) thermodynamics of irreversible processes; and (3) constitutive equations. Balance laws are fundamental macroscopic equations governing E-M and mechanical fields. They are valid for any type of media (solids, crystals, metals, bones, fluids, polymers, blood, oil, air, etc.) irrespective of their constitutions. Associated with each balance law, there is a jump condition across a discontinuity surface σ, which may be sweeping the body with its own velocity. When σ is considered to coincide with the surface of the body, jump conditions become boundary conditions. Jump

conditions are also fundamental to the discussion of shock waves.

Material properties are brought into play through the constitutive equations. These are phenomenological relations among the response functions (dependent variables) and the independent variables. These may be called the "cause and effect relations." The response functions are the dependent variables that enter the expression of the energy balance law. The independent variables are the ones that make a given material able to respond, producing effects in the material. Both sets of variables must be frame independent (objective) and admissible by the second law of irreversible thermodynamics. Generally, different materials require different constitutive equations. For example, constitutive equations of E-M elastic solids and those of E-M viscous fluids are different. While, by means of the irreversible thermodynamics and objectivity, unified, all-inclusive constitutive equations can be constructed, for most complex materials, these are neither of practical use nor helpful in understanding the physics of the field. In this article, constitutive equations are obtained for E-M elastic solids, magnetoelasticity, E-M viscous fluids, and MHD. The method presented here for the construction of the constitutive equations can be used for other classes of more complex materials, for example, memory-dependent materials, materials with microstructures, and nonlocal continua. For these, the reader is referred to the references cited in the "Further reading" section.

Once constitutive equations are constructed, by combining these equations with the balance laws, the field equations are obtained. The field equations are, generally, partial differential equations (PDEs) that must be solved under appropriate boundary and initial conditions to determine the response of the body. In this article, example solutions are presented to demonstrate the theories of E-M elastic solids and E-M viscous fluids.

Balance Laws

A tensor is called frame-independent (objective), if it transforms as a tensor under the rigid motions of the frame-of-reference, that is, under the transformations:

$$\bar{\boldsymbol{x}} = \boldsymbol{Q}(t)\boldsymbol{x} + \boldsymbol{b}(t), \qquad \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{1}, \quad \det \boldsymbol{Q} = 1, \quad \bar{t} = t - t_0 \qquad [1]$$

where $\boldsymbol{Q}(t)$ is the time-dependent orthogonal group, $\boldsymbol{b}(t)$ is an arbitrary translation, and t_0 is a shift of time. The superscript T denotes the transpose. For constants $\boldsymbol{Q}$ and $\boldsymbol{b}$, invariance under [1] denotes the Galilean invariance.

Local balance laws can be obtained by using the invariance of the global (integral) energy balance law under transformation [1] of the frame-of-reference $\boldsymbol{x}$. It is well known that the E-M balance laws are invariant under the Lorentz group, but the mechanical balance laws are not. The E-M and mechanical balance laws considered here are Galilean invariant. E-M balance laws (Maxwell's equations) in a fixed laboratory frame R_{F}, in the Heviside–Lorentz system, are expressed by

Gauss's law

$$\boldsymbol{\nabla} \cdot \boldsymbol{D} = q_e \quad \text{in } V \qquad [2a]$$

$$\boldsymbol{n} \cdot [[\boldsymbol{D}]] = w_e \quad \text{on } \sigma \qquad [2b]$$

Faraday's law

$$\boldsymbol{\nabla} \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0 \quad \text{in } V \qquad [3a]$$

$$\boldsymbol{n} \times \left[\left[\boldsymbol{E} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}\right]\right] = 0 \quad \text{on } \sigma \qquad [3b]$$

Absence of free magnetic poles

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \quad \text{in } V \qquad [4a]$$

$$\boldsymbol{n} \cdot [[\boldsymbol{B}]] = 0 \quad \text{on } \sigma \qquad [4b]$$

Ampere's law

$$\boldsymbol{\nabla} \times \boldsymbol{H} - \frac{1}{c}\frac{\partial \boldsymbol{D}}{\partial t} = \frac{1}{c}\boldsymbol{J} \quad \text{in } V \qquad [5a]$$

$$\boldsymbol{n} \times \left[\left[\boldsymbol{H} - \frac{1}{c}\boldsymbol{v} \times \boldsymbol{D}\right]\right] = 0 \quad \text{on } \sigma \qquad [5b]$$

The law of conservation of charge is obtained by taking the divergence of eqn [5a] and using eqn [2a]:

$$\frac{\partial q_e}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0 \quad \text{in } V \qquad [6a]$$

$$\boldsymbol{n} \cdot [[\boldsymbol{J} - q_e \boldsymbol{v}]] = 0 \quad \text{on } \sigma \qquad [6b]$$

where double brackets denote the jump conditions at the discontinuity surface σ. A superposed dot or D/Dt denotes the material time derivative, for example,

$$v_{k,l} = \frac{\partial v_k}{\partial x_l}, \qquad \dot{\rho} = \mathrm{D}\rho/\mathrm{D}t = \frac{\partial \rho}{\partial t} + \rho_{,k} v_k$$

where, and throughout, repeated indices denote summation over (1, 2, 3). Indices following a comma denote the partial derivatives with respect to the Cartesian coordinates.

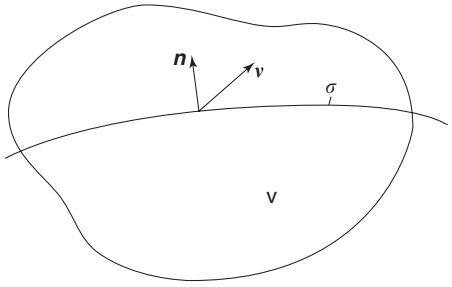


Figure 1 Discontinuity surface σ sweeping body.

In these equations, $\boldsymbol{D}, \boldsymbol{E}, \boldsymbol{B}, \boldsymbol{H}, \boldsymbol{J}, q_e, w_e$, and c are, respectively, the dielectric displacement vector, electric field vector, magnetic induction vector, magnetic field vector, current vector, volume charge density, surface charge density, and the speed of light in a vacuum. In each of eqns [2] to [6], the first set, marked by an "a" (e.g., [2a]), is valid in the volume of the body, V; and the second set, marked by a "b" (e.g., [2b]), is valid on the discontinuity surface σ, which may be sweeping V with its own velocity $\boldsymbol{v}$ in the direction of the unit normal $\boldsymbol{n}$ of σ (**Figure 1**).

$\boldsymbol{D}$ and $\boldsymbol{B}$ are decomposed as

$$\boldsymbol{D} = \boldsymbol{E} + \boldsymbol{P}, \qquad \boldsymbol{B} = \boldsymbol{H} + \boldsymbol{M} \tag{7}$$

where $\boldsymbol{P}$ and $\boldsymbol{M}$ are, respectively, polarization and magnetization vectors. They originate from the presence of the material.

For deformable bodies, it is necessary to express the E-M fields in a reference frame, R_M, co-moving with a material point of the body. These are expressed in terms of the fields in the fixed reference frame by

$$\begin{aligned}
\mathscr{E} &= \boldsymbol{E} + \frac{1}{c}\boldsymbol{v} \times \boldsymbol{B}, \qquad \mathscr{M} = \boldsymbol{M} + \frac{1}{c}\boldsymbol{v} \times \boldsymbol{P}, \\
\mathscr{J} &= \boldsymbol{J} - q_e \boldsymbol{v}, \qquad \mathscr{B} = \boldsymbol{B} - \frac{1}{c}\boldsymbol{v} \times \boldsymbol{E}, \\
\mathscr{H} &= \boldsymbol{H} - \frac{1}{c}\boldsymbol{v} \times \boldsymbol{D}
\end{aligned} \tag{8}$$

where, $\boldsymbol{v}$ is the velocity of the material point. The E-M force $\boldsymbol{F}^E$, couple $\boldsymbol{C}^E$, the power W^E, E-M stress tensor $\boldsymbol{t}^E_{kl}$, E-M momentum $\boldsymbol{G}$, and Poynting vector $\mathscr{S}$, are given by

$$\begin{aligned}
\boldsymbol{F}^E &= q_e \mathscr{E} + \frac{1}{c}(\mathscr{J} + \check{\boldsymbol{P}}) \times \boldsymbol{B} + (\boldsymbol{P} \cdot \boldsymbol{\nabla})\mathscr{E} \\
&\quad + (\boldsymbol{\nabla}\boldsymbol{B}) \cdot \mathscr{M} \\
\boldsymbol{C}^E &= \boldsymbol{P} \times \mathscr{E} + \mathscr{M} \times \boldsymbol{B} \\
W^E &= \boldsymbol{F}^E \cdot \boldsymbol{v} + \rho \mathscr{E} \cdot (\boldsymbol{P}/\rho)^{\cdot} - \mathscr{M} \cdot \dot{\boldsymbol{B}} + \mathscr{J} \cdot \mathscr{E} \\
\boldsymbol{t}^E &= \boldsymbol{P} \otimes \mathscr{E} - \boldsymbol{B} \otimes \mathscr{M} + \boldsymbol{E} \otimes \boldsymbol{E} + \boldsymbol{B} \otimes \boldsymbol{B} \\
&\quad - \tfrac{1}{2}(E^2 + B^2 - 2\mathscr{M} \cdot \boldsymbol{B})\boldsymbol{1} \\
\boldsymbol{G} &= \frac{1}{c}\boldsymbol{E} \times \boldsymbol{B}, \quad \mathscr{S} = c\mathscr{E} \times \mathscr{H}
\end{aligned} \tag{9}$$

where $\check{\boldsymbol{P}}$ is the displacement current, defined by

$$\check{\boldsymbol{P}} = \dot{\boldsymbol{P}} - (\boldsymbol{P} \cdot \boldsymbol{\nabla})\boldsymbol{v} + \boldsymbol{P}\boldsymbol{\nabla} \cdot \boldsymbol{v} \tag{10}$$

These expressions are obtained by statistical mechanical methods, from the atomic Maxwell–Lorentz equations, applied to a stable atomic particle (molecule).

Mechanical Balance Laws

Conservation of mass

$$\dot{\rho} + \rho \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0 \quad \text{in } V \tag{11a}$$

$$[[\rho(\boldsymbol{v} - \boldsymbol{v})]] \cdot \boldsymbol{n} = 0 \quad \text{on } \sigma \tag{11b}$$

Balance of momentum

$$t_{kl,k} + \rho(f_l - \dot{v}_l) + F^E_l = 0 \quad \text{in } V \tag{12a}$$

$$[[\rho v_l(v_k - v_k) - t_{kl} - t^E_{kl} - v_k G_l]]n_k = 0 \quad \text{on } \sigma \tag{12b}$$

Balance of moment of momentum

$$\varepsilon_{klm} t_{lm} + C^E_k = 0 \quad \text{in } V \tag{13}$$

Balance of energy

$$\begin{aligned}
&-\rho\dot{\varepsilon} + t_{kl}v_{l,k} + \boldsymbol{\nabla} \cdot \boldsymbol{q} + \rho h + \rho \mathscr{E} \cdot (\boldsymbol{P}/\rho)^{\cdot} \\
&\quad - \mathscr{M} \cdot \dot{\boldsymbol{B}} + \mathscr{J} \cdot \mathscr{E} = 0 \quad \text{in } V
\end{aligned} \tag{14a}$$

$$\begin{aligned}
&[[\{\rho(\varepsilon + \tfrac{1}{2}\rho v^2) + \tfrac{1}{2}(E^2 + B^2)\}(v_k - v_k) \\
&\quad -(t_{kl} + t^E_{kl} + v_k G_l)v_l - q_k + \mathscr{S}_k]]n_k = 0 \quad \text{on } \sigma
\end{aligned} \tag{14b}$$

where ρ, t_{kl}, f_l, ε, $\boldsymbol{q}$, and h are, respectively, mass density, the stress tensor, the body force density, the internal energy density, heat vector, and the heat input. ε_{klm} is the permutation symbol.

The second law of thermodynamics

$$\rho\dot{\eta} - \boldsymbol{\nabla} \cdot (\boldsymbol{q}/\theta) - \rho h/\theta \geqslant 0 \quad \text{in } V \tag{15a}$$

$$[[\rho\eta(\boldsymbol{v} - \boldsymbol{v}) - \boldsymbol{q}/\theta]] \cdot \boldsymbol{n} \geqslant \boldsymbol{0} \quad \text{on } \sigma \tag{15b}$$

where η and θ are, respectively, the entropy density and the absolute temperature.

Introducing the generalized Helmholtz's free energy, defined by

$$\psi = \varepsilon - \theta\eta - \rho^{-1}\boldsymbol{P} \cdot \mathscr{E} \tag{16}$$

into eqn [14a], and eliminating h from eqn [15a] by means of eqn [14a] gives

$$-\rho(\dot{\psi}+\eta\dot{\theta})+t_{kl}v_{l,k}+\frac{1}{\theta}\boldsymbol{q}\cdot\boldsymbol{\nabla}\theta - \boldsymbol{P}\cdot\frac{\mathrm{D}\mathscr{E}}{\mathrm{D}t}-\mathscr{M}\cdot\dot{\boldsymbol{B}}+\mathscr{J}\cdot\mathscr{E}\geqslant 0 \qquad [17]$$

This is the Clausius–Duhem (C–D) inequality which is fundamental to the construction of the constitutive equations.

E-M Thermoelastic Solids

Elastic solids deform under E-M fields and acquire stress. Conversely, applied mechanical loads produce E-M fields and electric current.

An elastic solid, at the natural state (free of all fields and undeformed and unstressed), occupies a volume V in the three-dimensional Euclidean space E^3. A material point $\boldsymbol{X}$ in V, under external loads, moves to a spacial point $\boldsymbol{x}$, at time t. The motion is expressed by

$$\boldsymbol{x}=\boldsymbol{x}(\boldsymbol{X},t)\quad\text{or}\quad x_k=x_k(X_K,t),\quad k,K=1,2,3 \qquad [18]$$

which is assumed to have the unique inverse:

$$\boldsymbol{X}=\boldsymbol{X}(\boldsymbol{x},t)\quad\text{or}\quad X_K=X_K(x_k,t) \qquad [19]$$

where X_K and x_k are, respectively, Cartesian coordinates of a material point in the material (Lagrangian) frame $\boldsymbol{X}$, and the spacial (Eulerian) frame $\boldsymbol{x}$. Deformation tensor C_{KL} and Lagrangian strain tensor E_{KL} in R_L and corresponding Eulerian tensors c_{kl}, e_{kl} in R_E are defined by

$$C_{KL}=x_{k,K}x_{k,L}=\delta_{KL}+2E_{KL}$$
$$c_{kl}=X_{K,k}X_{K,l}=\delta_{kl}-2e_{kl} \qquad [20]$$

Introducing the displacement vector by $\boldsymbol{u}=\boldsymbol{U}=\boldsymbol{x}-\boldsymbol{X}$, E_{KL} and e_{kl} are expressed as

$$E_{KL}=\tfrac{1}{2}(U_{K,L}+U_{L,K}+U_{M,K}U_{M,L})$$
$$e_{kl}=\tfrac{1}{2}(u_{k,l}+u_{l,k}-u_{m,k}u_{m,l}) \qquad [21]$$

Constitutive equations for E-M elastic solids are constructed with the help of a C–D inequality expressed in material coordinates X_K. This is necessary because specification of the boundary conditions in the spatial (current) configuration is not possible in the nonlinear theory, since the boundary of the current configuration is not known, *a priori*. From eqn [13] and the second of eqn [9], it can be seen that t_{kl} is not a symmetric tensor. A symmetric stress tensor $\bar{t}_{kl}$ and deformation tensor d_{kl} are introduced by

$$t_{kl}=\bar{t}_{kl}-P_k\mathscr{E}_l-\mathscr{M}_kB_l$$
$$d_{kl}=\tfrac{1}{2}(v_{k,l}+v_{l,k}) \qquad [22]$$

Spatial (Eulerian) tensors t_{kl}, q_k, P_k, $\mathscr{M}_k$, $\mathscr{J}_k$, and e_{kl} are related to their counterparts in material (Lagrangian) tensors by

$$\bar{t}_{kl}=\frac{\rho}{\rho_0}\bar{T}_{KL}x_{k,K}x_{l,L}$$
$$(q_k,\mathscr{E}_k,B_k)=(\mathscr{E}_K,B_K)X_{K,k}$$
$$(P_k,\mathscr{M}_k,\mathscr{J}_k)=\frac{\rho}{\rho_0}(Q_K,\Pi_K,\mathscr{M}_K,\mathscr{J}_K)x_{k,K} \qquad [23]$$
$$\dot{E}_{kl}=d_{kl}x_{k,K}x_{l,L}$$
$$\rho_0/\rho=\det x_{k,K}$$

where ρ_0 is the mass density at the natural state. With these, the C–D inequality [17] becomes

$$-\rho_0(\dot{\Psi}+\eta\dot{\theta})+\bar{T}_{KL}\dot{E}_{KL}+\frac{1}{\theta}Q_K\theta_{,K} - \Pi_K\frac{\mathrm{D}\mathscr{E}_K}{\mathrm{D}t}-\mathscr{M}_K\dot{B}_K+\mathscr{J}_K\mathscr{E}_K\geqslant 0 \qquad [24]$$

This inequality is linear in $\dot{\theta}$, $\dot{E}_{KL}$, $\mathrm{D}\mathscr{E}_k/\mathrm{D}t$, and $\dot{B}_K$, and it cannot be maintained in one sign unless

$$\eta=-\frac{\partial\Psi}{\partial\theta},\quad \bar{\boldsymbol{T}}=\rho_0\frac{\partial\Psi}{\partial\boldsymbol{E}}$$
$$\Pi=-\rho_0\frac{\partial\Psi}{\partial\mathscr{E}},\quad \mathscr{M}=-\rho_0\frac{\partial\Psi}{\partial\boldsymbol{B}} \qquad [25]$$

and

$$\frac{1}{\theta}Q_K\theta_{,K}+\mathscr{J}_K\mathscr{E}_K\geqslant 0 \qquad [26]$$

The general solution of this inequality is given by

$$\boldsymbol{Q}=\frac{\partial\Phi}{\partial(\boldsymbol{\nabla}\theta/\theta)}+\boldsymbol{U}_Q,\quad \mathscr{J}=\frac{\partial\Phi}{\partial\mathscr{E}}+\boldsymbol{U}_{\mathscr{J}} \qquad [27]$$

where constitutive residuals $\boldsymbol{U}_q$ and $\boldsymbol{U}_{\mathscr{J}}$ are subject to

$$\boldsymbol{U}_Q\cdot\frac{\nabla\theta}{\theta}+\boldsymbol{U}_{\mathscr{J}}\cdot\mathscr{E}=0 \qquad [28]$$

ψ is a function of $\Omega=(\theta,E_{KL},\mathscr{E}_K,B_K)$, Φ is a function of $\omega=(\boldsymbol{\nabla}\theta/\theta,\mathscr{E}_K)$ and the set Ω. The internal energy density ε and the dissipation potential Φ are required to be positive semidefinite, that is,

$$\varepsilon\geqslant 0,\quad \Phi(\Omega,\omega)\geqslant 0,\quad \Phi(0,\omega)=0 \qquad [29]$$

Expressions [25] and [27] are the most general constitutive equations of E-M elastic solids. Eulerian expressions can be obtained by substituting [25] and [27] into [23].

When the functions ψ and Φ are expressed in terms of their variables, and residuals selected to obey [28], the constitutive equations [25] and [27] take explicit forms. For example, ψ and Φ may be expressed as polynomials of any order in terms of their variables, to construct nonlinear constitutive equations that may be necessary when strong fields are applied to materials.

The balance laws [11a], [12a], and [14a] can be expressed in terms of the material (Lagrangian) fields by means of [23]:

$$\rho_0/\rho = \det(x_{k,K}) \qquad [30]$$

$$(\bar{T}_{KL}x_{k,L})_{,K} + \rho_0(f_k - \dot{v}_k) + F_k^E = 0 \qquad [31]$$

$$-\rho_0\theta\dot{\eta} + Q_{K,K} + \rho h + \mathscr{J}_K\mathscr{E}_K = 0 \qquad [32]$$

This permits the field equations to be obtained directly, by substituting [25] and [27] into [31] and [32].

Among various order theories, the linear theory is the most widely used in practical applications (e.g., piezoelectricity). In the linear theory, in [23], $x_{k,K}$ and $X_{K,k}$ become Dirac delta functions, and the distinction between Eulerian and Lagrangian representations disappears. Moreover, the constitutive residuals are also discarded. In this case, it proves more convenient to express constitutive equations in Eulerian forms.

By setting $\theta = T_0 + T$, where T is a temperature change from a constant ambient temperature $T_0 \gg 0$, $|T| < \theta$, and absorbing T_0 into the material moduli of Ψ and Φ, Ψ and Φ are expressed by second-degree polynomials:

$$\begin{aligned}\rho_0\Psi = &-\tfrac{1}{2}\rho_0\gamma T^2 - T(p_k\mathscr{E}_k + m_kB_k + \beta_{kl}e_{kl})\\ &+ a_{lmk}e_{lm}\mathscr{E}_k + b_{klm}e_{kl}B_m + \tfrac{1}{2}\lambda_{klmn}e_{kl}e_{mn}\\ &- \tfrac{1}{2}\chi^E_{kl}\mathscr{E}_k\mathscr{E}_l + \tfrac{1}{2}\chi^B_{kl}B_kB_l - h_{kl}\mathscr{E}_kB_l\end{aligned} \qquad [33]$$

$$\Phi = \tfrac{1}{2}\sigma_{kl}\mathscr{E}_k\mathscr{E}_l + \tau_{kl}\mathscr{E}_kT_{,l} + \tfrac{1}{2}\kappa_{kl}T_{,k}T_{,l} \qquad [34]$$

Generally, all material moduli γ, p_k, m_k, β_{kl}, a_{lmk}, b_{klm}, λ_{klmn}, χ^E_{kl}, χ^B_{kl}, h_{kl}, σ_{kl}, τ_{kl}, and κ_k depend on temperature and density.

Linear constitutive equations of E-M thermoelastic solids follow from [25] and [27]:

$$\begin{aligned}\eta &= \gamma T + \rho_0^{-1}(p_k\mathscr{E}_k + m_kB_k + \beta_{kl}e_{kl})\\ t_{kl} &= -\beta_{kl}T + \lambda_{klmn}e_{mn} + a_{klm}\mathscr{E}_m + b_{klm}B_m\\ P_k &= p_kT + \chi^E_{kl}\mathscr{E}_l - a_{lmk}e_{lm} + h_{kl}B_l\\ \mathscr{M}_k &= m_kT - \chi^B_{kl}B_l - b_{lmk}e_{lm} + h_{lk}\mathscr{E}_l\\ q_k &= \kappa_{kl}T_{,l} + \tau_{kl}\mathscr{E}_l + e(\boldsymbol{E}\times\boldsymbol{\nabla}T)_k\\ \mathscr{J}_k &= \sigma_{kl}\mathscr{E}_l + \tau_{kl}T_{,l} + h(\boldsymbol{E}\times\boldsymbol{B})_k\end{aligned} \qquad [35]$$

With these, upon using [16], the energy balance law [14a] is reduced to

$$-\rho_0T\dot{\eta} + \boldsymbol{\nabla}\cdot\boldsymbol{q} + \mathscr{J}\cdot\mathscr{E} + \rho_0h = 0 \qquad [36]$$

In order to obtain the field equations, by means of [7] to [10], and [35], $\boldsymbol{D}$, $\boldsymbol{H}$, $\boldsymbol{F}^E$, and [35] are expressed in terms of $\boldsymbol{E}$, $\boldsymbol{B}$, and $\boldsymbol{J}$, then substituted into Maxwell's equations, [12a] and [36].

The physical meaning of various moduli in [35] are γ – heat capacity, p_k – pyroelectric moduli, m_k – pyromagnetic moduli, β_{kl} – thermal stress moduli, λ_{klmn} – elastic moduli, a_{klm} – piezoelectric moduli, b_{klm} – piezomagnetic moduli, χ^E_{kl} – dielectric susceptibility, h_{kl} – magnetic polarizability, χ^B_{kl} – magnetic susceptibility, κ_{kl} – heat transfer moduli, σ_{kl} – electric conductivity, τ_{kl} – Peltier moduli for heat, Seebeck moduli for current, e – Righi – Leduc effect, and h – Hall effect. In linear theory, $e = h = 0$.

Constitutive eqns [35] contain some terms that may not be allowed. With the reversal of time (e.g., reversal of electronic spin), $\mathscr{M}_k$ and B_k change signs (i.e., they are time antisymmetric); all other fields are time symmetric. Consequently, the material symmetry group (magnetic group) must produce the same sign on both sides of the eqns [35]. Those that do not produce the same sign are not allowed. For example, pyroelectricity, represented by p_k, does not vanish in only ten noncentrosymmetric magnetic point groups: 1, 2, m, 2mm, 3, 3m, 4, 4mm, 6, 6mm. There is also a contribution to pyroelectricity from the thermal expansion (the secondary pyroelectricity). Similarly, piezoelectricity, represented by a_{klm}, exists in only 20 noncentrosymmetric groups. The one remaining noncentrosymmetric class, denoted by O (*Shönflies*) excludes piezoelectricity. Piezomagnetism (b_{klm}) exists in 11 groups. In centrosymmetric materials, piezoelectricity and piezomagnetism do not exist. Earlier, some authors, erroneously assuming that all crystals are time symmetric, concluded that pyromagnetism and piezomagnetism are forbidden effects.

This effect was observed in the antiferromagnetic crystals CoF_2 and MnF_2.

Materials are classified as dielectrics, diamagnetics, if conduction is absent and E-M fields cause polarization or magnetization. They are called ferromagnetic, if magnetization exists in the absence of the magnetic fields. Materials with positive susceptibility are called paramagnetic, and those that display electric polarization in the absence of applied electric fields are called ferroelectrics. Examples of dielectrics (also called insulators) are glasses and ceramics. Iron is a ferromagnet. Ceramic magnets are ferrites.

For experimental values of piezoelectric constants, the reader is referred to the "Further reading" section.

Some references employ the inversion of [35], e.g., expressions of e_{kl} in terms of (t_{kl}, P_k, $\mathscr{M}_k$, and T).

For isotropic elastic solids, the odd order tensor moduli vanish and the even order moduli become isotropic. Also, because of the time-antisymmetries of $\boldsymbol{B}$ and $\mathscr{M}$, h_{kl}, a_{klm}, b_{klm}, a_k, p_k, and m_k are not admissible, reducing [35] to

$$\begin{aligned} \eta &= (\gamma T + \rho_0^{-1}\beta \mathrm{tr}\, \boldsymbol{e}) \\ t_{kl} &= (-\beta T + \lambda_e \mathrm{tr}\, \boldsymbol{e})\delta_{kl} + 2\mu_e e_{kl} \\ \boldsymbol{P} &= (\varepsilon - 1)\mathscr{E} \\ \mathscr{M} &= (1 - \mu^{-1})\boldsymbol{B} \\ \boldsymbol{q} &= \tau\mathscr{E} + \kappa\boldsymbol{\nabla}T + e\boldsymbol{E}\times\boldsymbol{\nabla}T \\ \mathscr{J} &= \sigma\mathscr{E} + \tau\boldsymbol{\nabla}T + h\boldsymbol{E}\times\boldsymbol{B} \end{aligned} \quad [37]$$

where ε and μ given by

$$\varepsilon = 1 + \chi^E, \qquad \mu = (1 + \chi^B)^{-1} \quad [38]$$

are, respectively, called dielectric permeability and magnetic permeability.

It is simple to show that the internal energy ε and Φ will be positive-semidefinite, if and only if

$$\begin{aligned} &\gamma \geqslant 0,\ 3\lambda_e + 2\mu_e \geqslant 0,\ \mu_e \geqslant 0,\ \chi^E \geqslant 0, \\ &\chi^B \geqslant 0,\ \sigma \geqslant 0,\ \sigma\kappa - \tau^2 \geqslant 0 \end{aligned} \quad [39]$$

Linear-field equations of isotropic E-M thermoelastic solids are then given by

$$\begin{aligned} &-\beta\boldsymbol{\nabla}T + (\lambda_e + 2\mu_e)\boldsymbol{\nabla}\boldsymbol{\nabla}\cdot\boldsymbol{u} - \mu_e\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{u} \\ &\quad + \rho_0(\boldsymbol{f} - \ddot{\boldsymbol{u}}) + \boldsymbol{F}^E = \boldsymbol{0} \end{aligned} \quad [40]$$

$$-\rho_0 T\dot{\eta} + \boldsymbol{\nabla}\cdot\boldsymbol{q} + \mathscr{J}\cdot\mathscr{E} + \rho_0 h = 0 \quad [41]$$

$$\begin{aligned} \boldsymbol{\nabla}\cdot\left(\varepsilon\boldsymbol{E} + \frac{\varepsilon - 1}{c}\boldsymbol{v}\times\boldsymbol{B}\right) &= q_e \\ \boldsymbol{\nabla}\times\boldsymbol{E} + \frac{1}{c}\frac{\partial\boldsymbol{B}}{\partial t} = \boldsymbol{0}, \quad \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 \end{aligned} \quad [42]$$

$$\begin{aligned} &\boldsymbol{\nabla}\times\left(\mu^{-1}\boldsymbol{B} + \frac{\varepsilon - 1}{c}\boldsymbol{v}\times\boldsymbol{E}\right) - \frac{1}{c}\frac{\partial(\varepsilon\boldsymbol{E})}{\partial t} \\ &\quad - \frac{1}{c}(q_e\boldsymbol{v} + \sigma\mathscr{E} + \tau\boldsymbol{\nabla}T + h\boldsymbol{E}\times\boldsymbol{B}) = 0 \end{aligned} \quad [43]$$

$$\frac{\partial q_e}{\partial t} + \boldsymbol{\nabla}\cdot(q_e\boldsymbol{v} + \sigma\boldsymbol{E} + \tau\boldsymbol{\nabla}T + h\boldsymbol{E}\times\boldsymbol{B}) = 0 \quad [44]$$

where η, t_{kl}, $\boldsymbol{q}$, and $\mathscr{J}$ are given by eqn [37] and $\boldsymbol{F}^E$ by eqn [9].

Magnetoelasticity

The theory of magnetoelasticity is concerned with the deformation and magnetic behavior of nonferrous materials. For these materials, the effects of charge, displacement current, and polarization can be neglected. For simplicity, here, the isotropic bodies are considered. In this case, [43] reduces to

$$\boldsymbol{\nabla}\times\boldsymbol{H} - \frac{1}{c}\left[\sigma\left(\boldsymbol{E} + \frac{\mu}{c}\boldsymbol{v}\times\boldsymbol{H}\right) + \tau\boldsymbol{\nabla}T\right] = 0 \quad [45]$$

Curl of [45], upon the using the second of eqn [42], gives

$$\begin{aligned} &v_H\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{H} + \frac{\partial\boldsymbol{H}}{\partial t} - \boldsymbol{\nabla}\times(\dot{\boldsymbol{u}}\times\boldsymbol{H}) = \boldsymbol{0}, \\ &v_H = c^2/\sigma\mu \end{aligned} \quad [46]$$

Substituting the simplified expression of $\boldsymbol{F}^E$, given by [9] after neglecting q_e, $\boldsymbol{P}$, and $\breve{\boldsymbol{P}}$ into [12a] leads to

$$\begin{aligned} &(\lambda_e + 2\mu_e)\boldsymbol{\nabla}\boldsymbol{\nabla}\cdot\boldsymbol{u} - \mu_e\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{u} - \beta\boldsymbol{\nabla}T \\ &\quad + \mu_1(\boldsymbol{\nabla}\times\boldsymbol{H})\times\boldsymbol{H} + \rho(\boldsymbol{f} - \ddot{\boldsymbol{u}}) = \boldsymbol{0} \end{aligned} \quad [47]$$

Substituting constitutive eqns [37] into energy eqn [41], one obtains

$$\begin{aligned} &-\rho_0\gamma T_0\dot{T} - \beta T_0\boldsymbol{\nabla}\cdot\dot{\boldsymbol{u}} + \kappa_E\nabla^2 T \\ &\quad + (c^2/\sigma)(\boldsymbol{\nabla}\times\boldsymbol{H})\cdot(\boldsymbol{\nabla}\times\boldsymbol{H}) \\ &\quad - (c\tau/\sigma)(\boldsymbol{\nabla}\times\boldsymbol{H})\cdot\boldsymbol{\nabla}T + \rho h = 0 \end{aligned} \quad [48]$$

where

$$\mu_1 = \mu(2 - \mu), \qquad \kappa_E = \kappa - \tau^2/\sigma \quad [49]$$

Equations [46] to [48] are the field equations of magnetoelasticity. The solution of these equations under appropriate boundary and initial conditions determines $\boldsymbol{u}(x,t)$, $\boldsymbol{H}(\boldsymbol{x},t)$, and $T(\boldsymbol{x},t)$. Note that these equations are nonlinear. By linearizing [46] to [48] about a constant bias field H_0, the coupling of elastic waves with magnetic systems can be studied. These couplings are used to measure the dispersion of the elastic waves, and the coupling between the lattice and magnetic degrees of freedom.

The industrial applications of magnetic devices are many: ferrites are used in TV and sound amplifiers, computer disks, recorders, speakers, engines, and superconducting trains. Magnetostrictive transducers are used to convert E-M radiation into elastic waves and magnetoelastic delay lines.

Dissipative and Nonlocal Media

In many materials, the memory of the E-M fields and the stress tensor represents significant factors in terms of the outcome. Internal friction causes dissipation of energy. For such media, it is necessary to include the memory effects. For the conduction of high frequency and short wavelengths, nonlocality plays an important

role. In these cases, the constitutive equations involve the nonlocal effects of space and time.

Nonlocal constitutive equations in space–time generalize the Maxwell theory:

$$D_k(\boldsymbol{x},t) = \int_{-\infty}^{t} \mathrm{d}t' \int_V \varepsilon_{kl}(\boldsymbol{x}',\boldsymbol{x},t',t) \times E_l(x',t')\ \mathrm{d}v' \qquad [50]$$

where ε_{kl} is a nonlocal dielectric permeability tensor. Similar nonlocal constitutive equations exist for $\boldsymbol{B}(\boldsymbol{x},t)$ and t_{kl}, q_k, and $\mathscr{J}_k$. These permit the introduction of frequency and wave number dependents, nonlocal material moduli, such as dielectric and magnetic permeabilities $\varepsilon_{kl}(\boldsymbol{k},\omega)$, $\mu_{kl}(\boldsymbol{k},\omega)$.

The nonlocal theory provides rich grounds to discuss a variety of E-M phenomena, with realistic models, for example, media with absorption, polariton, gyroscopic media, and superconductivity.

Constitutive Equations of Rigid Solids

Rigid bodies do not deform, that is, stress and strain vanish:

$$T_{KL} = E_{KL} = 0, \quad \boldsymbol{Q} = \boldsymbol{q}, \quad \mathscr{E} = \boldsymbol{E}, \quad \mathscr{M} = \boldsymbol{M} \qquad [51]$$

Consequently, the constitutive equations of rigid solids can be obtained from those of the elastic solid by using [51].

E-M fields propagate in a vacuum as well. For a vacuum, the constitutive equations are of the form

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E}, \qquad \boldsymbol{B} = \mu_0 \boldsymbol{H} \qquad [52]$$

where the vacuum dielectric permeability ε_0 and the vacuum magnetic permeability μ_0 are dimensional constants related to the speed of light $c = (\varepsilon_0\mu_0)^{1/2}$, as predicted by the solution of Maxwell's equations. This, then, explained that light is an electromagnetic wave, a crowning achievement of the Maxwell theory. Thus, the hypothetical concept of ether (whose vibrations, historically, were considered to produce light) vanished from physics.

For an account on macroscopic electrodynamics of rigid bodies, the reader is referred to the "Further reading" section. The discussion of electrodynamics on the atomic scale requires a quantum mechanical approach.

Nonlinear E-M Theory

Discussion of nonlinear optics, electrostriction, magnetostriction, ferroelectrics, ferromagnetics, and other nonlinear phenomena requires the establishment of the nonlinear constitutive equations. These are usually obtained by expressing ψ and Φ as polynomials of a higher degree than 2, in their independent variables. The degrees of $\boldsymbol{E}$, $\mathscr{E}$, $\boldsymbol{B}$, and $\boldsymbol{\nabla}\theta/\theta$ depend on the accuracy desired. In the case of ferromagnetic materials, to represent the inner structure, a director is also necessary.

Electrostriction and Magnetostriction

Under strong E-M fields, elastic solids deform, developing stress. This is known as E-M striction. This is a nonlinear phenomena. For simplicity, as an example, consider simple nonlinear expressions for ε and Φ to demonstrate the electrostrictive effect:

$$2\rho_0\Psi = -\rho_0\gamma\theta^2 - 2\beta\theta\,\mathrm{tr}\,\boldsymbol{E} + \lambda(\mathrm{tr}\,\boldsymbol{E})^2 + 2\mu\,\mathrm{tr}\,\boldsymbol{E}^2 - [\chi^E + \alpha(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E}\cdot\mathscr{E} \qquad [53]$$

$$\rho_0\varepsilon = \rho_0\gamma\theta^2 + \tfrac{1}{2}\lambda(\mathrm{tr}\,\boldsymbol{E})^2 + \mu\,\mathrm{tr}\,\boldsymbol{E}^2 + \tfrac{1}{2}[\chi^E + \alpha(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E}\cdot\mathscr{E} \qquad [54]$$

$$2\Phi = [\sigma + \sigma^E(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E}\cdot\mathscr{E} + 2[\tau + \tau^E(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E}\cdot\frac{\boldsymbol{\nabla}\theta}{\theta} + [\kappa + \kappa^E(\mathrm{tr}\,\boldsymbol{E})^2]\frac{\boldsymbol{\nabla}\theta\cdot\boldsymbol{\nabla}\theta}{\theta^2} \qquad [55]$$

where γ, λ, μ, β, α, χ^E, σ, σ^E, τ, τ^E, κ, and κ^E are the material moduli. It can be seen that restrictions $\varepsilon \geqslant 0$ and $\Phi \geqslant 0$ are fulfilled, if and only if

$$(\gamma, \chi^E, \alpha, \sigma, \sigma^E) \geqslant 0, \qquad 3\lambda + 2\mu \geqslant 0, \quad \mu \geqslant 0, \qquad \sigma\kappa - \tau^2 \geqslant 0, \qquad \sigma^E\kappa^E - (\tau^E)^2 \geqslant 0 \qquad [56]$$

Constitutive equations follow from [25] and [27], upon using [53] and [55]:

$$\begin{aligned} \eta &= \gamma\theta + \rho_0^{-1}\beta\,\mathrm{tr}\,\boldsymbol{E} \\ T_{KL} &= [-\beta\theta + (\lambda - \alpha\mathscr{E}\cdot\mathscr{E})\mathrm{tr}\,\boldsymbol{E}]\boldsymbol{\delta}_{KL} + 2\mu E_{KL} \\ \Pi_K &= [\chi^E + \alpha(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E}_K \\ \mathscr{J} &= [\sigma + \sigma^E(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E} + [\tau + \tau^E(\mathrm{tr}\,\boldsymbol{E})^2]\frac{\boldsymbol{\nabla}\theta}{\theta} + h\boldsymbol{E}\times\boldsymbol{B} \\ \boldsymbol{Q} &= [\kappa + \kappa^E(\mathrm{tr}\,\boldsymbol{E})^2]\frac{\boldsymbol{\nabla}\theta}{\theta} + [\tau + \tau^E(\mathrm{tr}\,\boldsymbol{E})^2]\mathscr{E} + e\boldsymbol{E}\times\boldsymbol{\nabla}\theta \end{aligned} \qquad [57]$$

It is observed that the electric field produces stress. If pressure is defined by

$$p = -\tfrac{1}{3}T_{KK} = \beta\theta + [\alpha\mathscr{E}\cdot\mathscr{E} - \lambda - \tfrac{2}{3}\mu]\mathrm{tr}\,\boldsymbol{E} \qquad [58]$$

the volume of the body will change under this pressure. Electrical contribution to the pressure is given by $\alpha\mathscr{E}\cdot\mathscr{E}$. Conversely, the strain field causes polarization.

Parallel to this, magnetostriction may be illustrated similarly.

E-M Viscous Fluids

The interaction of E-M fields and fluids is observed in magnetohydrodynamics, plasma physics, sun flares, atmospheric physics, ferrofluids, electrohydrodynamics, and electrogas dynamic energy converters. Fluids do not possess natural states. Thermodynamics of fluids can be discussed by means of the spacial form [17] of the C–D inequality. The response functions (dependent variable set Z) and the independent variable set Y are given by the collections:

$$\begin{aligned} Z &= (\psi, \eta, t_{kl}, q_k, P_k, \mathscr{M}_k, \mathscr{J}_k) \\ Y &= (\rho, \theta, d_{kl}, \nabla\theta/\theta, \mathscr{E}_k, B_k) \end{aligned} \qquad [59]$$

All these quantities are objective (frame independent). Stress tensor t_{kl}, given by [22], is expressed as a sum of the static and dynamic parts, denoted, respectively, by the left-hand indices R and D:

$$t_{kl} = {}_R\,\bar{t}_{kl} + {}_D\,\bar{t}_{kl} - P_k\mathscr{E}_l - \mathscr{M}_k B_l \qquad [60]$$

The free energy ψ and dissipation potential Φ, for the linear theory, are quadratic forms expressed by the invariants of $\mathscr{E}$, $\boldsymbol{B}$ for ψ and $\boldsymbol{d}$, $\mathscr{E}$, $\nabla\theta/\theta$ for Φ:

$$\begin{aligned} \rho_0\psi &= \rho_0\Psi_0(\rho,\theta) - \tfrac{1}{2}\chi^E\mathscr{E}\cdot\mathscr{E} + \tfrac{1}{2}\chi^B\boldsymbol{B}\cdot\boldsymbol{B} \\ \Phi &= \tfrac{1}{2}\lambda_v(\operatorname{tr}\boldsymbol{d})^2 + \mu_v\operatorname{tr}\boldsymbol{d}^2 + \frac{1}{2}\kappa\frac{\nabla\theta\cdot\nabla\theta}{\theta^2} \\ &\quad + \frac{1}{2}\sigma\mathscr{E}\cdot\mathscr{E} + \tau\frac{\mathscr{E}\cdot\nabla\theta}{\theta} \end{aligned} \qquad [61]$$

The linear constitutive equations for E-M thermoviscous fluids follow from these:

$$\begin{aligned} &\eta = -\frac{\partial\Psi_0}{\partial T}, \quad t_{kl} = (-\pi + \lambda_v d_{rr})\delta_{kl} + 2\mu_v d_{kl}, \\ &\pi = \rho^2\frac{\partial\Psi_0}{\partial\rho} \\ &\boldsymbol{P} = \chi^E\mathscr{E}, \quad \mathscr{M} = -\chi^B\boldsymbol{B}, \\ &\boldsymbol{q} = \kappa\nabla T + \tau\mathscr{E}, \quad \mathscr{J} = \sigma\mathscr{E} + \tau\nabla T \end{aligned} \qquad [62]$$

where π is the thermodynamic pressure and (λ_v,μ_v) and κ are, respectively, the viscosities and heat conduction coefficient. σ is the conductivity and τ represents the current generated by the temperature gradient and the heat generated by the electric field. The field equations are obtained by substituting [62] into balance laws:

$$\nabla\cdot\left[(1+\chi^E)\boldsymbol{E} + \frac{1}{c}\chi^E\boldsymbol{v}\times\boldsymbol{B}\right] = q_e \qquad [63]$$

$$\nabla\times\boldsymbol{E} + \frac{1}{c}\frac{\partial\boldsymbol{B}}{\partial t} = 0 \qquad [64]$$

$$\nabla\cdot\boldsymbol{B} = 0 \qquad [65]$$

$$\begin{aligned} &\nabla\times\left[(1+\chi^B)\boldsymbol{B} + \frac{\chi^E}{c}\boldsymbol{v}\times\boldsymbol{E}\right] - \frac{1}{c}\frac{\partial[(1+\chi^E)\boldsymbol{E}]}{\partial t} \\ &\quad - \frac{1}{c}\left[q_e\boldsymbol{v} + \sigma\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right) + \tau\nabla T\right] = 0 \end{aligned} \qquad [66]$$

$$\dot\rho + \rho\nabla\cdot\boldsymbol{v} = 0 \qquad [67]$$

$$\begin{aligned} &-\nabla\pi + (\lambda_v + 2\mu_v)\nabla\nabla\cdot\boldsymbol{v} - \mu_v\nabla\times\nabla\times\boldsymbol{v} + \rho(\boldsymbol{f} - \dot{\boldsymbol{v}}) \\ &\quad + q_e\boldsymbol{E} + \frac{1}{c}[q_e\boldsymbol{v} + \sigma\boldsymbol{E} + \tau\nabla T + \chi^E\dot{\boldsymbol{E}} - \chi^E(\boldsymbol{E}\cdot\nabla)\boldsymbol{v} \\ &\quad + \chi^E\boldsymbol{E}\nabla\cdot\boldsymbol{v}]\times\boldsymbol{B} + \chi^E\left(\boldsymbol{E} + \frac{1}{c}(\boldsymbol{v}\times\boldsymbol{B})\right)\cdot\nabla\boldsymbol{E} \\ &\quad + \frac{1}{c}\chi^E\boldsymbol{E}\cdot\nabla(\boldsymbol{v}\times\boldsymbol{B}) - \chi^B(\nabla\boldsymbol{B})\cdot\boldsymbol{B} = 0 \end{aligned} \qquad [68]$$

$$\begin{aligned} &-C\dot T - G\nabla\cdot\boldsymbol{v} + \lambda_v(\operatorname{tr}\boldsymbol{d})^2 + 2\mu_v\operatorname{tr}\boldsymbol{d}^2 \\ &\quad + \nabla\cdot\left[\kappa\nabla T + \tau(\boldsymbol{E} + c^{-1}\boldsymbol{v}\times\boldsymbol{B})\right] \\ &\quad + (\sigma\boldsymbol{E} + \boldsymbol{\tau}\nabla T)\cdot\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right) \\ &\quad + \rho h = 0 \end{aligned} \qquad [69]$$

where

$$C = -\rho T\frac{\partial^2\Psi_0}{\partial T^2}\Big|_0, \qquad G = \rho^2 T\frac{\partial\Psi}{\partial\rho\partial T}\Big|_0 \qquad [70]$$

Magnetohydrodynamic Approximation

In MHD the motion is nonrelativistic, that is, $v/c \ll 1$, $\chi^E \ll 1$, so that $\boldsymbol{B} = \mu\boldsymbol{H}$, $\mu = (1+\chi^B)^{-1}$. Displacement current $\partial\boldsymbol{D}/\partial t$ is negligible. Also, in the expression $\boldsymbol{F}^E$, terms containing $\breve{\boldsymbol{P}}$ and $\boldsymbol{P}$ are neglected. With this, $\boldsymbol{F}^E$ is reduced to

$$\begin{aligned} \boldsymbol{F}^E &= -\nabla(\mu_1 H^2/2) + \mu(\boldsymbol{H}\cdot\nabla)\boldsymbol{H}, \\ \mu_1 &= 2\mu - \mu^2 \end{aligned} \qquad [71]$$

With these simplifications, and taking the curl of [66], the field equations of magnetohydrodynamics

follow from [66] to [69]:

$$\dot{\rho} + \rho \nabla \cdot \boldsymbol{v} = 0 \qquad [72]$$

$$-\eta^B \nabla \times \nabla \times \boldsymbol{B} + \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \frac{\partial \boldsymbol{B}}{\partial t} = 0, \quad \eta^B = c^2/\sigma\mu \qquad [73]$$

$$-\nabla\left(\pi - \frac{\mu - 1}{2\mu}B^2\right) + (\lambda_v + 2\mu_v)\nabla\nabla \cdot \boldsymbol{v} - \mu_v \nabla \times \nabla \times \boldsymbol{v} + \rho(\boldsymbol{f} - \dot{\boldsymbol{v}}) + \frac{1}{\mu}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} = 0 \qquad [74]$$

$$-C\dot{T} - G\nabla \cdot \boldsymbol{v} + \lambda_v (\mathrm{tr}\, \boldsymbol{d})^2 + 2\mu_v \mathrm{tr}\, \boldsymbol{d}^2 + \left(\kappa - \frac{\tau^2}{\sigma}\right)\nabla^2 T + \frac{c^2}{\mu^2 \sigma}(\nabla \times \boldsymbol{B})^2 - \frac{\tau c}{\sigma\mu}\nabla T \cdot (\nabla \times \boldsymbol{B}) + \rho h = 0 \qquad [75]$$

In the following sections, three example problems in the E-M theory of continuous media are discussed.

Piezoelectric Waves

t_{kl} and P_k, given by eqn [35], with $T = 0$, $\boldsymbol{B} = 0$, in R_F are written as

$$t_{kl} = \tfrac{1}{2}\lambda_{klmn}(u_{m,n} + u_{n,m}) + a_{klm}E_m \qquad [76]$$

$$P_k = -\tfrac{1}{2}a_{lmk}(u_{l,m} + u_{m,l}) + \chi^E_{kl}E_l \qquad [77]$$

A plane dielectric elastic wave propagating in the direction of a unit vector $\boldsymbol{n}$ may be represented by

$$G(\boldsymbol{n} \cdot \boldsymbol{x} \pm c_0 t), \quad \boldsymbol{u} = \boldsymbol{u}(\boldsymbol{n} \cdot \boldsymbol{x} \pm c_0 t) \qquad [78]$$

where c_0 is the phase velocity of the wave. The electric potential ϕ is introduced as

$$\boldsymbol{E} = -\nabla\phi \qquad [79]$$

Substituting [76]–[79], into balance laws [2a] and [12a], with $q_e = 0$ and $f_k = 0$,

$$\epsilon_{kl}\phi'' n_k n_l + a_{lmk}u''_l n_m n_k = 0 \qquad [80]$$

$$\lambda_{klmn}u''_m n_n n_k - a_{klm}\phi'' n_m n_k - \rho c_0^2 u''_l = 0 \qquad [81]$$

Eliminating ϕ'' from [81] by means of [80] yields

$$(A_{lm} - \rho c_0^2 \delta_{lm})u''_m = 0 \qquad [82]$$

where

$$A_{lm} = \bar{\lambda}_{klmn} n_k n_n, \quad \bar{\lambda}_{klmn} = \lambda_{klmn} + (\varepsilon_{ij} n_i n_j)^{-1} a_{kln}\, a_{mpq} n_p n_q \qquad [83]$$

Since A_{lm} is a symmetric tensor, whose proper numbers are real, the three roots of

$$\det(A_{lm} - \rho c_0^2 \delta_{lm}) = 0 \qquad [84]$$

are distinct, and the acoustical axes $\boldsymbol{u}''/|\boldsymbol{u}''|$ form an orthogonal triad at a point on the wave surface. c_0^2 must be positive. Thus, for a real wave to propagate in every direction, the acoustical tensor A_{lm} must be positive-definite. It can be shown that the necessary and sufficient condition for this is the strong ellipticity of the piezoelectrically stiffened stiffness tensor, that is,

$$\bar{\lambda}_{klmn} A_l A_n B_k B_m \geqslant 0 \qquad [85]$$

for arbitrary vectors $\boldsymbol{A}$ and $\boldsymbol{B}$. From this discussion, it is clear that the material symmetry regulations for the moduli $\lambda_{klmn}, \alpha_{klm}$, and ϵ_{kl} are important to the outcome.

Piezoelectricity has many industrial applications. A few of them include air bag sensors, audible alarms, keyless entry, computers, speakers, dialysis equipment, sonar, guidance systems, and hydrophones.

Magnetohydrodynamic Poiseulle Flow

An incompressible, electrically conducting fluid with constant viscosity μ_v and constant electrical conductivity σ flows between two parallel straight insulated walls, in the direction of x_1, parallel to the walls. The applied constant magnetic field H_0 is in the x_2-direction, normal to the walls. Flow is steady and two dimensional. Introducing nondimensional variables

$$x = \frac{x_1}{L}, \quad y = \frac{x_2}{L}, \quad v(y) = \frac{v_1}{V}, \quad v_2 = v_3 = 0$$
$$H(y) = \frac{H_1}{H_0}, \quad H_2 = H_0, \qquad [86]$$
$$H_3 = 0, \quad p_0 = p/\rho_0 V^2$$

where p is the unknown pressure, and L and V are, respectively, a characteristic length and a characteristic velocity, the equations of motion follow from [73] and [74]. Heat conduction is neglected:

$$\frac{\partial v}{\partial y} + \frac{1}{R_\mathrm{B}}\frac{\mathrm{d}^2 H}{\mathrm{d}y^2} = 0 \qquad [87]$$

$$\frac{1}{R}\frac{\mathrm{d}^2 v}{\mathrm{d}y^2} + \frac{1}{A^2}\frac{\mathrm{d}H}{\mathrm{d}y} = \frac{\partial p_0}{\partial x} \qquad [88]$$

$$\frac{1}{A^2}H\frac{\mathrm{d}H}{\mathrm{d}y} = -\frac{\mathrm{d}p_0}{\mathrm{d}y} \qquad [89]$$

where R, R_B, and A are, respectively, the Reynolds number, the magnetic Reynolds number, and the Alfvén number, defined by

$$R = \rho_0 VL/\mu_\nu, \quad R_B = VL/\eta^B,$$
$$A = (\rho V^2/\mu_\nu H_0^2)^{1/2} \qquad [90]$$

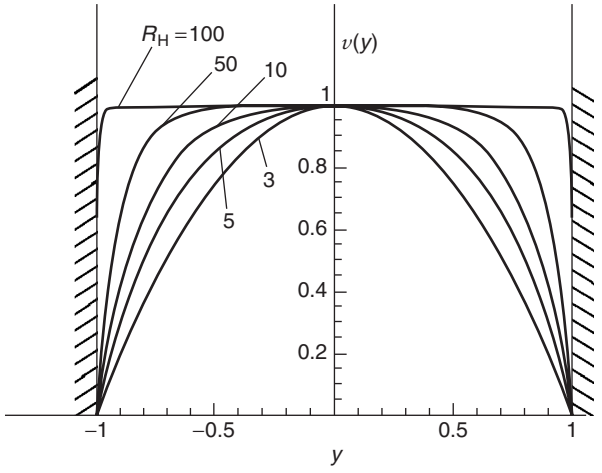


Figure 2 Velocity profile across channel for $R_H = 3, 5, 10, 50, 100$.

Boundary conditions are

$$v(\pm 1) = 0, \quad v(0) = 1,$$
$$H(\pm 1) = 0 \quad \text{(insulated walls)} \qquad [91]$$

The solution of these equations is

$$\begin{aligned} v &= [\cosh R_H - \cosh(R_H y)]/(\cosh(R_H - 1) \\ H(y) &= (R_B/R_H) \\ &\quad \times [\sinh(R_H y) - y \sinh R_H]/(\cosh R_H - 1) \\ p_0 &= -\tfrac{1}{2} A^{-2} H^2(y) + Gx + \bar{p} \end{aligned} \qquad [92]$$

where $R_H = (RR_B/A^2)^{1/2}$ is the Hartmann number and

$$G = (R_H/R)(\cosh R_H - 1)^{-1} \sinh R_H \qquad [93]$$

and $\bar{p}$ are constants. G is the pressure gradient and $\bar{p}$ is the pressure at $x = 0$. As R_H tends to zero, the velocity profile becomes a classical parabolic (Poiseulle) velocity profile

$$v = 1 - y^2 \qquad [94]$$

In the limit $R_H \to \infty$, $v = 1$ everywhere, except $y = \pm 1$, that is, the velocity profile is flat, quickly vanishing at the walls. The velocity profile and magnetic field profile, for various values of R_H, are shown in **Figures 2** and **3**, respectively.

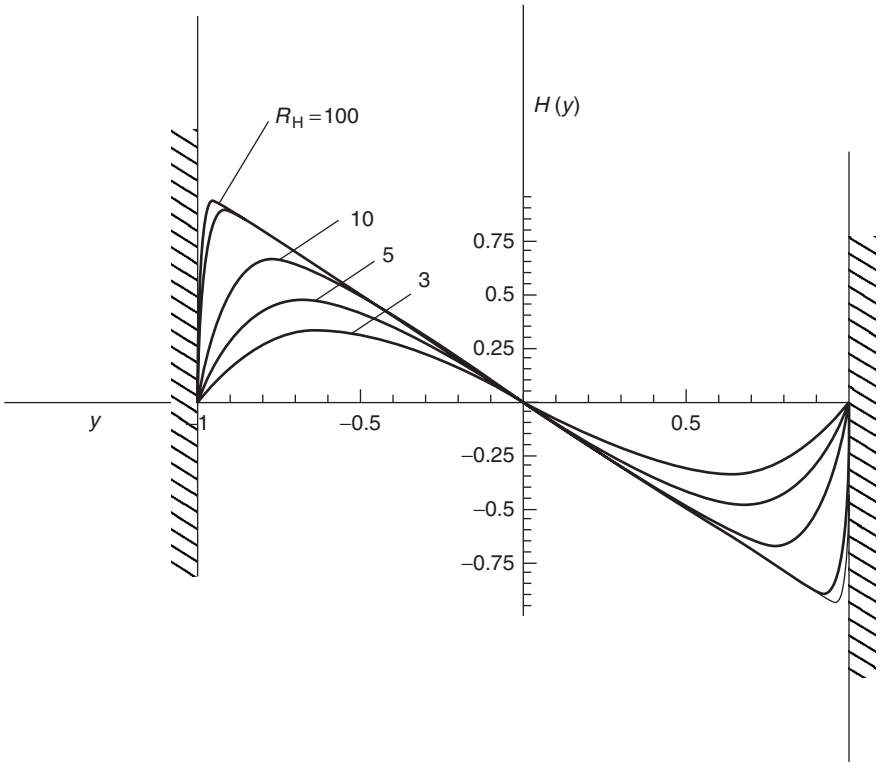


Figure 3 Magnetic field profile for $R_H = 3, 5, 10, 50, 100$.

See also: Electrodynamics: Quantum Electrodynamics.

PACS: 40; 41; 41.20.−q; 41.20.Cv; 41.20.Gz; 41.20Jb; 41.90.+e; 46; 46.05.+b; 46.70.−p; 47; 47.60.+l; 47.65.+a

Further Reading

Agranovich VM and Ginsburg VL (1984) *Crystal Optics with Spatial Dispersion and Excitons*, 2nd edn. Berlin: Springer.

American Institute of Physics (1957) *American Institute of Physics Handbook*. New York: McGraw-Hill.

Berlincourt DA, Curran DR, and Jaffe H (1964) Piezoelectric and piezomagnetic materials and their functions as transducers. In: Mason (ed.) *Physical Acoustics*, vol. IA. New York: Academic Press.

Cady WG (1964) *Piezoelectricity*, vol. I and II. New York: Dover.

Eringen AC (1967) *Mechanics of Continuous Media*, 2nd edn. Florida: Robert E. Krieger Publishing Co.

Eringen AC (2002) *Nonlocal Continuum Field Theories*. New York: Springer.

Eringen AC and Maugin GA (1989) *Electrodynamics of Continua I & II*. Springer.

Groot SR and Suttorp LG (1972) *Foundations of Electrodynamics*. Amsterdam: North-Holland.

Grot RA and Eringen AC (1966) Electromagnetic interactions with matter. *Int. J. Engng. Sci.* 4: 639–670.

Grot RA and Eringen AC (1966) Relativistic continuum mechanics. *International Journal of Engineering Science* 4: 611–638.

Jackson JD (1975) *Classical Electrodynamics*, 2nd edn. New York: Wiley.

Kiral E and Eringen AC (1980) *Constitutive Equations of Nonlinear Electromagnetic Elastic Solids*. New York: Springer.

Landau LD, Lifschitz EM, and Pitaevskii LP (1984) *Electrodynamics of Continuous Media*, 2nd edn. Oxford: Reed Educational and Professional Publishing Ltd.

Maradudin AA, Montrol EW, Weiss GH, and Ipatova IP (1971) *The Theory of Lattice Dynamics in Harmonic Approximations*, 2nd edn. New York: Academic Press.

Mason WP (1950) *Piezoelectric Crystals and Their Applications to Ultrasonics*. Princeton, NJ: Von Nostrand.

Maugin GA, Pouget J, Drout R, and Collet B (1992) *Nonlinear Electro-Mechanical Coupling*. West Sussex, UK: Wiley.

Nomenclature

a_{klm}	piezoelectric moduli ($\mathrm{V\,m^{-1}}$)
b_{klm}	piezomagnetic moduli ($\mathrm{V\,m^{-1}}$)
A	Alfvén number
$\boldsymbol{B}$, B_k	magnetic induction vector ($\mathrm{V\,m^{-1}}$)
c	speed of light in a vacuum ($\mathrm{m\,s^{-1}}$)
$\boldsymbol{C}^E$, C_k^E	E-M couple ($\mathrm{V^2\,m^{-2}}$)
$\boldsymbol{d}$, d_{kl}	deformation-rate tensor ($\mathrm{s^{-1}}$)
$\boldsymbol{D}$, D_k	dielectric displacement vector ($\mathrm{V\,m^{-1}}$)
e_{kl}	Eulerian strain tensor
$\boldsymbol{E}$, E_k	electric field vector ($\mathrm{V\,m^{-1}}$)
E_{KL}	Lagrangian strain tensor
$\boldsymbol{f}$, f_k	body force density ($\mathrm{m\,s^{-2}}$)
$\boldsymbol{F}^E$, F_k^E	E-M force vector ($\mathrm{V^2\,m^{-3}}$)
$\boldsymbol{G}$, G_k	E-M momentum ($\mathrm{V^2\,m^{-2}}$)
h	heat supply ($\mathrm{m^2\,s^{-3}}$)
h_{kl}	magnetic polarizability
$\boldsymbol{H}$, H_k	magnetic field vector ($\mathrm{V\,m^{-1}}$)
$\boldsymbol{J}$, J_k	current vector ($\mathrm{V\,m\,s^{-1}}$, A)
m_k	pyromagnetic moduli ($\mathrm{V\,m^{-1}\,^\circ C}$)
$\boldsymbol{M}$, M_k	magnetization vector ($\mathrm{V\,m^{-1}}$)
$\boldsymbol{n}$, n_k	unit vector
p_k	pyroelectric moduli ($\mathrm{V\,m^{-1}\,^\circ C}$)
$\boldsymbol{P}$, P_k	polarization vector ($\mathrm{V\,m^{-1}}$)
$\boldsymbol{q}$, q_k	heat vector ($\mathrm{kg\,m^{-1}}$)
q_e	volume charge density ($\mathrm{V\,m^{-1}}$, $\mathrm{C\,m^{-3}}$)
$\boldsymbol{Q}$, Q_K	Lagrangian heat vector ($\mathrm{kg\,m^{-1}}$)
R	Reynold's number
R_B	magnetic Reynold's number
R_H	Hartmann number
t	time (s)
$\boldsymbol{t}$, t_{kl}	stress tensor ($\mathrm{kg\,m^{-2}}$)
$\boldsymbol{T}$, T_{KL}	Lagrangian stress tensor ($\mathrm{kg\,m^{-2}}$)
T_0, T	temperature (K, °C)
$\boldsymbol{u}$, u_k	displacement vector (m)
$\boldsymbol{U}$, U_K	Lagrangian displacement vector (m)
$\boldsymbol{v}$, v_k	velocity vector ($\mathrm{m\,s^{-1}}$)
V	volume ($\mathrm{m^3}$)
w_e	surface charge density ($\mathrm{C\,m^{-2}}$)
W^E	E-M power ($\mathrm{V^2\,m^{-2}}$, W)
$\boldsymbol{x}$, x_k	Eulerian coordinates (m)
$\boldsymbol{X}$, X_K	Lagrangian coordinates (m)
β_{kl}	thermal stress moduli ($\mathrm{kg\,m^{-2}\,^\circ C}$)
γ	heat capacity ($\mathrm{m^2\,s^{-2}\,^\circ C^2}$)
δ_{kl}, δ_{KL}	Kronecker symbol
ϵ	internal energy density ($\mathrm{m^2\,s^{-2}}$)
$\mathscr{E}$, $\mathscr{E}_k$	electric field vector ($\mathrm{V\,m^{-1}}$)
η	entropy density ($\mathrm{m^2\,s^{-2}\,^\circ C}$)
$\mathscr{H}$	magnetic field vector ($\mathrm{V\,m^{-1}}$)
θ	absolute temperature (K)
$\mathscr{J}$, $\mathscr{J}_k$	Eulerian current vector (A)
$\mathscr{J}_K$	Lagrangian current vector (A)
κ_{kl}, κ	heat transfer moduli ($\mathrm{kg\,^\circ C^{-1}}$)
λ_{klmn}	elastic moduli ($\mathrm{kg\,m^{-2}}$)
λ_e, μ_e	Lame constants ($\mathrm{kg\,m^{-2}}$)
λ_v, μ_v	viscosities ($\mathrm{kg\,s\,m^{-2}}$)
$\mathscr{M}$, $\mathscr{M}_k$	Eulerian magnetic induction vector ($\mathrm{V\,m^{-1}}$)
$\boldsymbol{v}$, v_k	velocity of σ ($\mathrm{m\,s^{-1}}$)
π	thermodynamic pressure ($\mathrm{k\,m^{-2}}$)
Π	Lagrangian polarization vector ($\mathrm{V\,m^{-1}}$)
ρ, ρ_0	mass density ($\mathrm{kg\,s^2\,m^{-4}}$)
σ_{kl}, σ	electric conductivity ($\mathrm{s^{-1}}$)
τ_{kl}, τ	Peltier moduli ($\mathrm{V\,s^{-1}\,^\circ C^{-1}}$)
Φ	dissipation potential ($\mathrm{V^2\,m^{-2}\,s^{-1}}$)
ψ, Ψ	Helmholtz' free energy ($\mathrm{m^2\,s^{-2}}$)

Electrodynamics: Quantum Electrodynamics

G Scharf, Universität Zürich, Zürich, Switzerland

Introduction

Quantum electrodynamics (QED) is the quantum counterpart of classical (Maxwell's) electrodynamics in two ways. First, the electromagnetic field is quantized. This is necessary if one considers high enough frequencies or certain aspects of the interaction between radiation and matter (spontaneous emission). Second, the charged particles which are the sources of the electromagnetic fields have to be described by their correct relativistic equations of motion: the Klein–Gordon (KG) equation for integer spin and the Dirac equation for spin 1/2, in particular for electrons. This is necessary if the kinetic energy of the particles is comparable with their rest energy mc^2, or if the fine structure of the atomic spectra plays a role. However, the KG and Dirac equations are one-particle equations which have no consistent physical interpretation, because the energy is unbounded from below.

The correct procedure is to consider the KG and Dirac equations as field equations for the scalar or electron-positron field, respectively, and to quantize them (second quantization). Second quantization is the method to transform an unphysical single-particle theory into a physical many-particle theory. As such it plays a crucial role in solid-state theory, too. For this reason, the next section starts with a general discussion of second quantization. In the section "Quantized electron–positron and electromagnetic fields," this method is applied to the quantization of the free Dirac and electromagnetic fields. The interaction is introduced in the section "Causal perturbation theory," in the framework of the S-matrix theory. The S-matrix is constructed by the causal perturbation theory which avoids the nasty problems with ultraviolet divergences completely. In the last section, a review of results and an outlook to more complicated (non-Abelian) gauge theories are given of which QED is a part.

Second Quantization

Let $\mathscr{H}$ be the one-particle Hilbert space, for example, $L^2(\mathbb{R}^3)$ in case of a spinless particle. Then the n-particle space is given by the tensor product $\mathscr{H}_n^{\pm} = S_n^{\pm}\mathscr{H}^{\otimes n}$, where S_n^+ is the symmetrization operator in case of integer spin (bosons) and S_n^- the antisymmetrization operator in case of spin 1/2 (fermions). In order to describe all multiparticle states simultaneously, one introduces the Fock space

$$\mathscr{F}^{\pm} = \bigoplus_{n=0}^{\infty} \mathscr{H}_n^{\pm}$$

where the one-dimensional space $\mathscr{H}_0$ consists of the vacuum Ω. An element of $\mathscr{F}$ is an infinite sequence of states $\Phi = (\varphi_0, \varphi_1, \dots)$ with $\varphi_n \in \mathscr{H}_n$.

All operators in Fock space can be defined with the help of emission (or creation) and absorption (or annihilation) operators. The emission operator $a^+(f)$ adds a particle with wave function $f \in \mathscr{H}_1^{\pm}$ to all components φ_n in Φ:

$$a^+(f)\Omega = f$$

$$(a^+(f)\Phi)_n = \sqrt{n}S_n^{\pm}(f \otimes \varphi_{n-1}), \quad n = 1, 2, \dots$$

The absorption operator $a(f)$ is just the adjoint of $a^+(f)$. These operators are often written in distributional form in x-space or momentum space, for example,

$$a^+(f) = \int \mathrm{d}^3k\, a^+(\boldsymbol{k}) f(\boldsymbol{k})$$

Then, for the (distributional) field operators

$$(a^+(\boldsymbol{k})\Phi)_n(\boldsymbol{k}_1, \dots, \boldsymbol{k}) = \sqrt{n}S_n^{\pm}(\delta(\boldsymbol{k} - \boldsymbol{k}_1)\varphi_{n-1}(\boldsymbol{k}_2, \dots, \boldsymbol{k}_n))$$

$$(a(\boldsymbol{k})\Phi)_n(\boldsymbol{k}_1, \dots, \boldsymbol{k}) = \sqrt{n+1}\varphi_{n+1}(\boldsymbol{k}, \boldsymbol{k}_1, \dots, \boldsymbol{k}_n)$$

and $a(\boldsymbol{k})\Omega = 0$.

The most important properties of these operators are their commutation relations in case of bosons or their anticommutation relations in case of fermions

$$\begin{aligned} [a(f), a^+(g)]_{\mp} &= (f, g)_1 \\ [a(f), a(g)]_{\mp} &= 0 = [a(f)^+, a^+(g)]_{\mp} \end{aligned} \qquad [1]$$

where $(\cdot\,,\,\cdot)_1$ is the scalar product in the one-particle space. For the field operators, this gives the simple anticommutation relations

$$[a(\boldsymbol{k}), a^+(\boldsymbol{k}')]_{\mp} = \delta(\boldsymbol{k} - \boldsymbol{k}') \qquad [2]$$

and zero otherwise. For fermions, this implies that $a^+(f)a^+(f) = 0$, that is, two particles cannot be in the same state f (exclusion principle). Any operator in Fock space can be expressed by emission and absorption operators in the so-called normally (or

Wick) ordered form, for example,

$$\boldsymbol{H} = \int \mathrm{d}^3k\, a^+(\boldsymbol{k})H(\boldsymbol{k})a(\boldsymbol{k})$$

Normally ordered means that all absorption operators stand to the right of all emission operators.

In condensed matter physics, the occupation number representation is very popular. One chooses a basis f_k in $\mathscr{H}_1$ and the corresponding tensor basis

$$S_n^{\pm} f_{k_1} \otimes \cdots \otimes f_{k_n} \qquad [3]$$

in $\mathscr{H}_n$. To fix the sign in eqn [3] in case of fermions, one orders the factors with increasing indices $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_n$. If n_i is the number of k's which are equal to i, that is, the number of particles in the state f_i, then an orthonormal basis in $\mathscr{H}_n$ is given by

$$|n_1, n_2, \ldots\rangle = \sqrt{\frac{n!}{n_1!n_2!\ldots}} S_n^{\pm} f_{k_1} \otimes \cdots \otimes f_{k_n} \qquad \sum_{i=1}^{\infty} n_i = n \qquad [4]$$

The occupation numbers n_i assume all values $0, 1, 2, \ldots$ in case of bosons, but only 0, 1 for fermions. The emission operator $a_k^+ = a^+(f_k)$ can be written as

$$(a_k^{\pm}\Phi)_n = \sqrt{n_k + 1}\, |n_1, \ldots n_k + 1, \ldots\rangle$$

in case of bosons and the absorption operator is given by

$$(a_k\Phi)_n = \sqrt{n_k}\, |n_1, \ldots n_k - 1, \ldots\rangle$$

For fermions, one has to be careful with the sign.

Quantized Electron–Positron and Electromagnetic Fields

Free relativistic electrons are described by the Dirac equation

$$\mathrm{i}\hbar\gamma^\mu \partial_\mu \psi(x) = mc\psi(x) \qquad [5]$$

Here, the relativistic notation: $\mu = 0, 1, 2, 3$, $x = (x^\mu) = (ct, x^1, x^2, x^3)$, $\partial_\mu = \partial/\partial x^\mu$ is used. γ^μ are four 4×4 matrices satisfying

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$$

where $g^{\mu\nu} = \mathrm{diag}\{1, -1, -1, -1\}$. Furthermore, it is convenient to assume the following property under the Hermitian conjugation:

$$\gamma^{\mu +} = \gamma^0\gamma^\mu\gamma^0 \qquad [6]$$

The solution $\psi(x)$ of eqn [5] is a four-component object, a Dirac spinor, which has a well-defined behavior under the Lorentz transformation $x' = \Lambda x$ such that the Dirac equation [5] holds in any frame of reference.

The general solution of eqn [5] is best written as a Fourier integral

$$\psi(x) = (2\pi)^{-3/2} \int \mathrm{d}^3p [b_s(\boldsymbol{p})u_s(\boldsymbol{p})\mathrm{e}^{-\mathrm{i}px} + \mathrm{d}_s(\boldsymbol{p})^+ v_s(\boldsymbol{p})\mathrm{e}^{\mathrm{i}px}] \qquad [7]$$

Here, $px = p^0x^0 - \boldsymbol{p}\boldsymbol{x}$ with $p^0 = E = \sqrt{\boldsymbol{p}^2 + m^2}$, $(c = 1)$ and the u and v spinors are solutions of the linear algebraic equations

$$(p_\mu\gamma^\mu - m)u_s(\boldsymbol{p}) = 0$$

$$(p_\mu\gamma^\mu + m)v_s(\boldsymbol{p}) = 0$$

They are labeled by the polarization index $s = 1, 2$ and normalized by

$$u_s^+(\boldsymbol{p})u_{s'}(\boldsymbol{p}) = \delta_{ss'} = v_s^+(\boldsymbol{p})v_{s'}(\boldsymbol{p})$$

and orthogonal $u_s^+(\boldsymbol{p})v_s(-\boldsymbol{p}) = 0$.

The quantized Dirac field has the same form [7] where $b_s(\boldsymbol{p})$ are absorption operators of electrons and $d_s(\boldsymbol{p})^+$ emission operators of positrons, respectively. According to the last section, they satisfy anticommutation relations

$$\{b_s(\boldsymbol{p}), b_{s'}(\boldsymbol{p}')^+\} = \delta_{ss'}\delta^3(\boldsymbol{p} - \boldsymbol{p}') = \{d_s(\boldsymbol{p})^+, d_{s'}(\boldsymbol{p}')\} \qquad [8]$$

All other anticommutators vanish. This gives the following anticommutation rule for the total electron positron field [7] and the adjoint $\bar{\psi}(x) = \psi(x)^+\gamma^0$

$$\{\psi(x), \bar{\psi}(x)\} = -\mathrm{i}S(x - y) = (\gamma^\mu\partial_\mu^x - \mathrm{i}m)D(x - y) \qquad [9]$$

where

$$D(x) = \frac{\mathrm{i}}{(2\pi)^3} \int \mathrm{d}^4p\, \delta(p^2 - m^2)\mathrm{sgn}\, p^0 \mathrm{e}^{-\mathrm{i}px} \qquad [10]$$

is the Jordan–Pauli distribution. It has a causal support, that means it vanishes for spacelike x: $x^2 < 0$.

The first term in eqn [7] is called the absorption part $\psi^{(-)}$ (the misleading notion "positive frequency"

part is avoided), the second term is the emission part $\psi^{(+)}$. The anticommutator

$$\{\psi^{(-)}(x), \bar{\psi}^{(+)}(x)\} = -\mathrm{i}S^{(+)}(x-y) = (\gamma^\mu \partial^x_\mu - \mathrm{i}m)D^{(+)}(x-y) \quad [11]$$

where $D^{(+)}$ has the same form as eqn [9] with $\mathrm{sgn}(p^0)$ substituted by $\Theta(p^0)$ (here the meaning of positive frequency part is clear). The adjoint relation involves the negative frequency part $D^{(-)}$ with $-\Theta(-p^0)$ instead of $\mathrm{sgn}(p^0)$.

Consider the quantization of the electromagnetic field. Here, one has to quantize the four-vector potential $A^\mu(x)$, which satisfies the wave equation $\Box A^\mu(x) = 0$. Starting again from a classical solution

$$A^\mu(x) = (2\pi)^{-3/2} \int \frac{\mathrm{d}^3 k}{\sqrt{2\omega}} (a^\mu(\boldsymbol{k})\mathrm{e}^{-\mathrm{i}kx} + a^\mu(\boldsymbol{k})^* \mathrm{e}^{\mathrm{i}kx}) \quad [12]$$

where $\omega = k^0 = |\boldsymbol{k}|$. To consider the $a^\mu(\boldsymbol{k})$ as absorption operators in a photon Fock space with positive definite metric, the (bosonic) commutation relations are

$$[a^\mu(\boldsymbol{k}), a^\nu(\boldsymbol{k}')^+] = \begin{cases} \delta(\boldsymbol{k}-\boldsymbol{k}') & \text{for } \mu = \nu \\ 0 & \text{for } \mu \neq \nu \end{cases}$$

However, Lorentz covariance forces one to change the meaning of the conjugation in eqn [12] for the scalar component: $a^0(\boldsymbol{k})^* = -a^0(\boldsymbol{k})^+$. This makes A^0 a skew-adjoint operator instead of a self-adjoint one, but this is harmless because scalar photons are unphysical degrees of freedom. From eqn [12], one finds the covariant commutation rule

$$[A^\mu(x), A^\nu(y)] = g^{\mu\nu}\mathrm{i}D_0(x-y) \quad [13]$$

where D_0 is the Jordan–Pauli distribution [10] for mass zero.

It is well known that the radiation field has only two transverse physical degrees of freedom. Consequently, besides the scalar photons there are unphysical longitudinal ones which are generated by emission operators

$$a_{||}(\boldsymbol{k})^+ = \frac{k_j}{\omega} a^j(\boldsymbol{k})$$

The total Fock space is a direct sum of a physical subspace and the rest. Only states in the physical subspace can appear as incoming and outgoing states in scattering processes. However, the unphysical degrees of freedom contribute to the interaction. This is the typical situation in quantum gauge theories.

Causal Perturbation Theory

Most experimental information in QED is contained in the S-matrix, which until now is only defined by its perturbation series

$$S(g) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}^4 x_1 \ldots \mathrm{d}^4 x_n T_n(x_1, \ldots, x_n) \times g(x_1) \ldots g(x_n) \quad [14]$$

Here $g(x)$ is a Schwartz test function which cuts off the long-range part of the interaction and, therefore, serves as an infrared regulator. To compute observable quantities, one usually performs the limit $g \to e$, where e is the coupling constant which is equal to the unit of charge (assuming $\hbar = 1,\ c = 1$). The T_n are the chronological (or time-ordered) products which are expressed by the free asymptotic fields constructed in the last section. It is not necessary to consider interacting fields, if one is only interested in the S-matrix. In QED the first order is given by

$$T_1(x) = ie : \bar{\psi}(x)\gamma^\mu \psi(x) : A_\mu(x) \quad [15]$$

The double dots mean normal ordering which is essential, otherwise the product of field operators at the same space-time point x is not well defined.

The higher orders $n \geqslant 2$ can be constructed inductively from T_1 on the basis of a few basic axioms for the T_n (Bogoliubov axioms). The axioms express symmetry of T_n in $x_1, \ldots, x_n$, Lorentz and translation invariance and, most importantly, causality. Causality means that T_n factorizes

$$T_n(x_1, \ldots, x_n) = T_m(x_1, \ldots, x_m) \times T_{n-m}(x_{m+1}, \ldots, x_n) \quad [16]$$

if all $x_1, \ldots, x_m$ have time coordinates greater than all $x_{m+1}, \ldots, x_n$ in some Lorentz frame. To see the strength of this property, the construction of T_2 is shown explicitly. With the unknown T_2, consider the difference

$$T_2(x_1, x_2) - T_1(x_1)T_1(x_2) = A_2(x_1, x_2)$$

For $x_1^0 > x_2^0$, this distribution vanishes by causality (advanced distribution), similarly

$$T_2(x_2, x_1) - T_1(x_2)T_1(x_1) = R_2(x_1, x_2) \quad [17]$$

vanishes for $x_1^0 < x_2^0$ (retarded distribution). In the difference

$$D_2(x_1, x_2) = R_2 - A_2 = T_1(x_1)T_1(x_2) - T_1(x_2)T_1(x_1) \quad [18]$$

the unknown T_2 drops out, so this distribution can be computed from the first order. Since the supports of the advanced and retarded distributions have only the plane $x_1^0 = x_2^0$ in common, the splitting of D_2 at this plane gives R_2 and A_2 separately. Then T_2 is obtained from eqn [17], for example. The situation is even simpler. The commutator eqn [18] has causal support, that is, inside the double cone $(x_1 - x_2)^2 \geqslant 0$, because it involves commutators of free fields which all have causal support. Therefore, the two supports intersect in the vertex of this double cone, only, which makes the distribution splitting much simpler. In fact, the splitting which preserves the singular order of the distribution D_2 is unique up to a distribution with point support which is a sum of the derivatives of $\delta(x_1 - x_2)$, if the singular order is $\geqslant 0$.

The above construction goes through in all orders. It goes back to Epstein and Glaser (1973) and has the big advantage that all quantities are well defined, such that no ultraviolet divergences appear. The latter have plagued quantum field theorists for decades and have even been taken as a motivation to abandon the quantum field theory completely and study the string theory. On the other hand, it is pretty clear that if one starts from well-defined quantities (free fields) and makes well-defined operations only, then no divergences can appear. In the following, it is shown for the second order of QED.

To compute the causal distribution D_2 [18], one has to multiply two factors T_1 [15]. The result can be written in normally ordered form by using Wick's theorem, which for two factors simply reads:

$$A_1 A_2 = :A_1 A_2: + :\overbrace{A_1 A_2}: + \cdots \qquad [19]$$

Here the pairing symbol denotes a contraction, that is a commutator (in the case of Bose fields) or an anticommutator (in the case of Fermi fields) between the absorption and emission parts of the two contracted field operators. In QED, only the following three contractions appear:

$$\overbrace{\psi_a(x)\bar{\psi}_b(y)} = \{\psi_a^{(-)}(x), \bar{\psi}_b^{(+)}(y)\} = \frac{1}{\mathrm{i}} S_{ab}^{(+)}(x-y)$$

$$\overbrace{\bar{\psi}_a(x)\psi_b(y)} = \{\bar{\psi}_a^{(-)}(x), \psi_b^{(+)}(y)\} = \frac{1}{\mathrm{i}} S_{ba}^{(-)}(y-x)$$

$$\overbrace{A_\mu(x)A_\nu(y)} = [A_\mu^{(-)}(x), A_\nu^{(+)}(y)] = g_{\mu\nu}\mathrm{i}D_0^{(-)}(x-y)$$

The dots in eqn [19] mean the sum of all possible contractions, there can be more than one contraction between A_1 and A_2. The individual terms in eqn [19] are often represented by graphs. For the photon field operator in eqn [15], one draws a wavy line and for ψ and $\bar{\psi}$ a normal line; all three lines meet at the vertex x. For the second order, one takes two such vertices and connects as many lines as there are contractions. The uncontracted field operators remain as external lines, they represent incoming and outgoing particles in the scattering process. One obtains second-order graphs as shown in **Figure 1**.

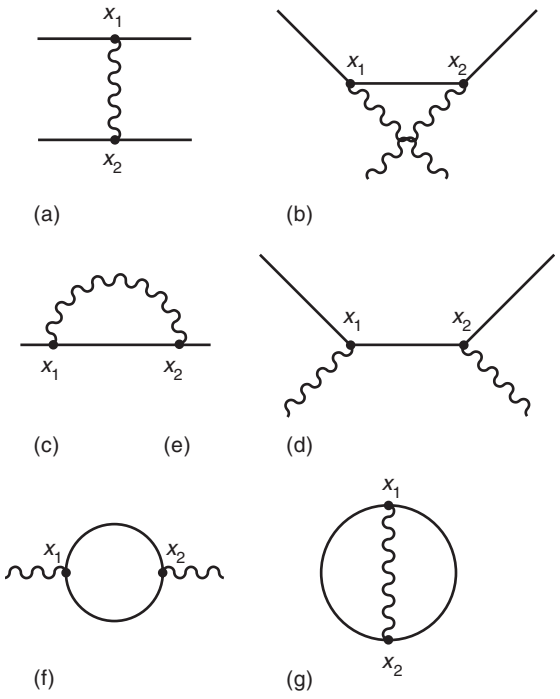


Figure 1 Second-order graphs corresponding to the terms ((a)–(g)) in eqn [20].

The corresponding analytical expression [18] is given by

$$D_2(x_1, x_2)$$

$$= e^2 \gamma_{ab}^\mu \gamma_{cd}^\nu \{: \bar{\psi}_a(x_1)\psi_b(x_1)\bar{\psi}_c(x_2)\psi_d(x_2): g_{\mu\nu}\mathrm{i}$$
$$[D_0^{(+)}(x_2 - x_1) - D_0^{(+)}(x_1 - x_2)] \qquad [20a]$$

$$- :\psi_b(x_1)\bar{\psi}_c(x_2)::A_\mu(x_1)A_\nu(x_2):$$
$$\times \frac{1}{\mathrm{i}}[S_{da}^{(+)}(x_2 - x_1) + S_{da}^{(-)}(x_2 - x_1)] \qquad [20b]$$

$$- :\psi_b(x_1)\bar{\psi}_c(x_2):[S_{da}^{(+)}(x_2 - x_1)D_0^{(+)}(x_2 - x_1)$$
$$+ S_{da}^{(-)}(x_2 - x_1)D_0^{(+)}(x_1 - x_2)]g_{\mu\nu} \qquad [20c]$$

$$- :\bar{\psi}_a(x_1)\psi_d(x_2)::A_\mu(x_1)A_\nu(x_2):$$
$$\times \frac{1}{\mathrm{i}}[S_{bc}^{(-)}(x_1 - x_2) + S_{bc}^{(+)}(x_1 - x_2)] \qquad [20d]$$

$$- :\bar{\psi}_a(x_1)\psi_d(x_2):[S_{bc}^{(-)}(x_1 - x_2)D_0^{(+)}(x_2 - x_1)$$
$$+ S_{bc}^{(+)}(x_1 - x_2)D_0^{(+)}(x_1 - x_2)]g_{\mu\nu} \qquad [20e]$$

$$+ : A_\mu(x_1)A_\nu(x_2) : [-S^{(-)}_{bc}(x_1-x_2)S^{(+)}_{da}(x_2-x_1) + S^{(+)}_{bc}(x_1-x_2)S^{(-)}_{da}(x_2-x_1)] \quad [20f]$$

$$+\frac{1}{\mathrm{i}}g_{\mu\nu}[S^{(-)}_{bc}(x_1-x_2)S^{(+)}_{da}(x_2-x_1)D^{(+)}_0(x_2-x_1) - S^{(+)}_{bc}(x_1-x_2)S^{(-)}_{da}(x_2-x_1)D^{(+)}_0(x_1-x_2)]\} \quad [20g]$$

the numbering indicates the corresponding graph in **Figure 1**.

The graphs with only one contraction (tree graphs) can be simply computed. For example, the bracket in (a) is equal to the Jordan–Pauli distribution $-D_0$ [10]. The splitting into retarded and advanced part is trivial by multiplying with $\Theta(x_1^0 - x_2^0)$. Then the retarded distribution R_2 is proportional to $D_0^{\mathrm{ret}}(x_1 - x_2)$ and the total chronological product proportional to the Feynman propagator

$$D_0^{\mathrm{ret}}(x_1-x_2) + D_0^{(+)}(x_1-x_2) = D_0^{\mathrm{F}}(x_1-x_2) \quad [21]$$

according to eqn [17]. This simple procedure leads to the Feynman rules, where each internal line is represented by a Feynman propagator. But this method does not work for loop graphs. Here the distribution splitting must be done more carefully, otherwise one gets the famous ultraviolet divergences.

For example, for the vacuum polarization graph eqn [20f] we get

$$D_2^{(6)}(x_1,x_2) = -e^2\mathrm{Tr}[\gamma^\mu S^{(-)}(y)\gamma^\nu S^{(+)}(-y) - \gamma^\mu S^{(+)}(y)\gamma^\nu S^{(-)}(-y)] : A_\mu(x_1)A_\nu(x_2) : \quad [22]$$

with $y = x_1 - x_2$. The trace herein has the form $P_1^{\mu\nu}(y) - P_1^{\mu\nu}(-y)$, where $P_1^{\mu\nu}$ is best calculated in momentum space

$$\hat{P}_1^{\mu\nu}(k) = -e^2(2\pi)^{-4}\left(\frac{k^\mu k^\nu}{k^2} - g^{\mu\nu}\right)\hat{d}_1(k) \quad [23]$$

with

$$\hat{d}(k) = \hat{d}_1(k) - \hat{d}_1(-k) = \frac{2\pi}{3}(k^2+2m^2)\sqrt{1-\frac{4m^2}{k^2}} \times \Theta(k^2-4m^2)\,\mathrm{sgn}(k_0) \quad [24]$$

This causal distribution has a singular order $\omega = 2$, then its retarded part is given by the dispersion integral

$$\hat{r}(k) = \frac{\mathrm{i}}{2\pi}\int_{-\infty}^{+\infty}\mathrm{d}t\frac{\hat{d}(tk)}{(t-\mathrm{i}0)^{\omega+1}(1-t+\mathrm{i}0)} \quad [25]$$

This integral is finite and leads to the 'renormalized' vacuum polarization tensor directly. If one does not determine the singular order correctly, the dispersion integral is ultraviolet divergent.

Results

There are two condensed matter systems which are relevant for QED because they give very precise values of the fundamental constants of nature: (1) the AC Josephson effect and (2) the quantum Hall effect. In (1), one applies a DC voltage V across the barrier of a superconducting junction and gets an alternating current of frequency $\nu = 2\,\mathrm{eV}\,h^{-1}$. In (2), one measures the Hall resistance R_H of a two-dimensional electron gas in a strong perpendicular magnetic field. R_H has discrete values

$$R_\mathrm{H} = \frac{h}{ne^2}, \quad n = 1, 2, \ldots$$

which directly give the dimensionless fine structure constant

$$\alpha = \frac{\mu_0 c e^2}{2h}, \quad \mu_0 = 4\pi\cdot 10^{-7}\,\mathrm{Hm}^{-1}$$

The fine structure constant got its name from the bound-state energies E_{nj} of the Dirac equation with a Coulomb potential Ze^2/r:

$$E_{nj} = mc^2\left[1+\frac{(Z\alpha)^2}{(n-\beta)^2}\right]^{-1/2}$$

where n is the principal quantum number, j the total angular momentum and

$$\beta = j + \tfrac{1}{2} - \sqrt{(j+\tfrac{1}{2})^2 - (Z\alpha)^2}$$

It is difficult to obtain a precise value for α from this formula because various corrections must be taken into account. One of the biggest and well calculable by QED is the vacuum-polarization correction of the Coulomb potential. It is the main contribution to the splitting of the $2S_{1/2} - 2P_{1/2}$ states, the so-called Lamb shift.

The most accurate value of α comes from the spin magnetic moment of the electron $\mu = g\mu_\mathrm{B}$, where $\mu_\mathrm{B} = e\hbar/2mc$ is Bohr's magneton and g the dimensionless g value. The Dirac equation gives $g = 2$. The difference $a = g - 2$ has been calculated in QED to better than 1 ppb by now; the result is

$$a = \frac{\alpha}{2\pi} - 0.328478965\ldots\left(\frac{\alpha}{\pi}\right)^2 + 1.181241456\ldots\left(\frac{\alpha}{\pi}\right)^3 - 1.5098(384)\left(\frac{\alpha}{\pi}\right)^4 + 4.393(27)\cdot 10^{-12}$$

The terms up to α^3 are known analytically; the α^3-term was calculated only in 1996 by Laporta and Remeddi. The α^4 term requires evaluation of 891 four-loop diagrams, one only has numerical estimates for them. The experimental value a_e for the electron has been measured to about 4 ppb. The theoretical value for the muon is less accurate because corrections due to strong and weak interactions are necessary.

In high-energy particle physics, only low-order QED results are needed. The most important scattering processes are the following: electron–electron (Møeller) scattering, electron–positron (Bhabba) scattering, photon–electron (Compton) scattering, pair production and annihilation, and the processes with muons. However, as the center of mass energy is in the GeV region, weak and strong interactions have to be taken into consideration. In fact, QED is part of a bigger theory which is now called the standard model. It has grown out of QED and is, in contrast to QED, a non-Abelian gauge theory. For the sake of completeness, its basic ideas are given.

Instead of one gauge field $A^\mu(x)$, one has $N=12$ (photon, $W^\pm$, Z and 8 gluons) denoted by $A^\mu_a(x)$, $a=1,\ldots,N$. On these asymptotic free fields, one defines a gauge variation generated by a gauge charge Q

$$[Q, A^\mu_a(x)] = \mathrm{i}\partial^\mu u_a(x)$$

The $u_a(x)$ are fermionic ghost fields. In contrast to QED, the non-Abelian gauge fields have self-coupling. Its form $T_1(x)$ is uniquely determined by quantum gauge invariance:

$$[Q, T_1(x)] = \mathrm{i}\partial_\mu T_1^\mu(x) \qquad [26]$$

The resulting coupling is of the form

$$T_1(x) = \mathrm{i}g f_{abc} : A_{\mu a}(x) A_{\nu b}(x) \partial^\mu A^\mu_c(x) : + \cdots$$

where f_{abc} are structure constants of a Lie algebra, and the dots represent couplings to ghost fields and additional fields in the theory which are all determined by gauge invariance. The above coupling is called Yang–Mills coupling. It has been found that all interactions in nature, including gravity, are quantum theories in the sense of eqn [26] above.

See also: Geometrical Optics; Quantum Hall Effect.

PACS: 11.10; 11.15; 11.20

Further Reading

Akhiezer AI and Berestetski VB (1965) *Quantum Electrodynamics*. New York: Wiley.

Bogoliubov NN and Shirkov DV (1976) *Introduction to the Theory of Quantized Fields*, 3rd edn. New York: Wiley.

Hughes VW and Kinoshita T (1999) Anomalous g values of the electron and muon. *Review of Modern Physics* 71: pp. S133–39.

Itzykson C and Zuber JB (1980) *Quantum Field Theory*. New York: McGraw-Hill.

Jauch JM and Rohrlich F (1975) *The Theory of Photons and Electrons*. Berlin: Springer.

Kaku M (1993) *Quantum Field Theory*. Oxford: Oxford University Press.

Kinishita T (1996) The fine structure constant. *Reports of Progress in Physics* 59: pp. 1459–92.

Scharf G (1995) *Finite Quantum Electrodynamics*, 2nd edn. Berlin: Springer.

Scharf G (2001) *Quantum Gauge Theories*. Chichester, NY: Wiley.

Weinberg S (1996) *The Quantum Theory of Fields*, vols. I and II. Cambridge: Cambridge University Press.

Electromagnetic Biological Effects

C Marino, ENEA Centro Ricerche Casaccia, Rome, Italy
B Veyret, Université de Bordeaux, Pessac, France

Introduction

Research on biological and health effects of electromagnetic fields (EMFs) has been driven in the last few decades by public concerns about potential detrimental effects. Therapeutic applications have not yet been developed except for microwave hyperthermia.

This article gives a review of the scientific knowledge available on the biological and health effects of EMFs in the two main frequency ranges of concern: the extremely low frequency (ELF, <300 Hz) range mainly at power frequency (50/60 Hz) and the radiofrequency (RF) range, mainly between 300 MHz and 3 GHz, where most sources emit (radio, TV, mobile telephones, wireless networks, etc.) (see **Table 1**).

Explicit distinctions must be made between the concepts of interaction, biological effect, and health hazard, consistent with the criteria used by international bodies when making health assessments: biological effects occur when fields interact to produce responses that may or may not be experienced by people. Deciding whether biological changes have health consequences depends, in part, on whether they are reversible, are within the range for which the

Table 1 Description of frequency ranges

Frequency (Hz)	*Wavelength* (m)	*Description (or radiation name)*
0		Static field
$1–100 \times 10^3$	6×10^6 at 50 Hz	Extremely low frequencies (ELF)
100×10^3–100×10^6	3×10^3–3	Radiofrequencies (RF)
10×10^9	0.03	Microwave (MW)
300×10^{12}	1×10^{-6}	Infrared (IR)
390×10^{12}	760×10^{-9}	Visible red and light
10^{15}	300×10^{-9}	Ultraviolet (UV)
10^{18}	3×10^{-10}	X-ray
10^{21}	3×10^{-13}	γ-ray

Hertz (Hz) = unit of frequency: one cycle per second.
If $f < 300$ GHz (i.e., 3×10^{11} Hz), $hf < 1.2 \times 10^{-3}$ eV → Nonionizing radiation (NIR), where as If $f \geqslant 2.4 \times 10^{15}$ Hz (UV), $hf \geqslant 10$ eV → Ionizing radiation.

body has effective compensation mechanisms, or are likely, taking into account the variability of response among individuals, to lead to unfavorable changes in health. World Health Organization (WHO) defines health as the state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity. Not all biological effects are hazardous; some may be innocuously within the normal range of biological variation and physiological compensation. Others may be beneficial under certain conditions, and the health implications of others may be simply indeterminate. In this article, a health hazard is defined as a biological effect outside the normal range of physiological compensation and adverse to well-being.

It is known that electromagnetic interference with electromedical devices, such as cardiac pacemakers, leads to adverse health consequences, but this article is only concerned with the direct effects of EMFs on living tissues.

Scientific approaches to health risk assessment related to nonionizing EMFs are the same as in the ionizing range, that is, epidemiology, experimental studies on humans, animals, and cells in culture. The experimental protocols are also similar but the choice of exposure systems and biological models are different.

Bioeffects of ELF Fields

Many effects on biological systems exposed to ELF fields have been reported. In the last twenty years, numerous experiments have been performed in laboratories worldwide to assess the biological effects of ELF magnetic and electric fields. This has led to the design and building of several types of exposure systems for *in vitro* and *in vivo* work. The goal is to expose the biological samples (cells or animals) to well-characterized fields under well-defined environmental conditions. Only recently, these have been of sufficient quality to warrant that biological findings were not due to physical artifacts.

In view of the interaction processes of fields with biological organisms (mainly induced electric fields at ELF), the effects of magnetic fields, which penetrate the body, have been the main focus of research projects. In contrast, electric fields at ELF do not penetrate the organism well because of its conductivity. This is not true for direct stimulation of nerves and other excitable cells using electrodes. Electric fields can also induce polarization of the cell membrane leading to specific effects such as electroporation. There have been many improvements in the design of exposure systems in recent years and most of the key parameters described below are well-characterized and controlled. However, standardization of these systems has only been achieved within some of the multicenter research programs.

Many mechanisms have been proposed through which living systems could interact with ELF electric and magnetic fields. However, both laboratory studies and controlled human studies have provided evidence only for electrically based phenomena that occur in the presence of ELF fields and that are present only during exposure to these fields.

Induced electric signals in tissues of the nervous and visual system of animals exposed to ELF electric and magnetic fields can produce effects that lead to field perception. The most notable effect that occurs in humans exposed to fields of moderate intensity is the induction of visual phosphenes. The neuronal circuits of the retina form a part of the central nervous system (CNS) and thus the magneto-phosphene threshold may be taken as a realistic indicator of the threshold value for functional changes induced in other parts of the CNS. Threshold values at 20 Hz were estimated at $\sim 100\,\mathrm{mV\,m^{-1}}$. Direct effects on the peripheral nervous system occur only at relatively high levels of electrical stimulation ($\sim 2\,\mathrm{V\,m^{-1}}$ and with increasing thresholds for frequencies above 1 kHz).

A variety of other potential mechanisms for interaction of ELF electric and magnetic fields with living systems have been proposed, but not demonstrated to produce effects that could significantly perturb normal biological functions. Examples of such interactions are the following: (1) magneto-chemical interactions with charge-transfer reactions which do not appear to influence biochemical pathways such as those involving electron transport, and (2) resonance interactions involving combined static and

ELF fields in specific geometric configurations which have not been shown to have significant biological effects, consistent with predictions based on well-founded physical arguments.

Cellular Models

There is a large and varied body of research literature on the ability of fields to effect changes in cellular models. Overall, the evidence for stimulation of cell division by ELF fields is at best equivocal. There is also no clear evidence that ELF magnetic fields can alter calcium ion homeostasis. Finally, the data on several aspects of intracellular signaling do not consistently support the existence of field effects on these signaling pathways. Furthermore, the types of effects reported, when they have been found, have been of a magnitude that gives little support to the conclusion that they are of consequence for human health.

Animal Models

Overall, a large number of studies have been carried out, principally investigating the possible effects of exposure to ELF electric and magnetic fields, testing many different biological endpoints in a variety of animal species.

With regard to genotoxicity, there is no clear evidence that exposure to ELF magnetic fields presents a hazard. Relatively few animal studies have been carried out, however, and the subject has been more extensively investigated at the cellular level (see above). Similarly, no convincing evidence has been found from experimental studies supporting the hypothesis that exposure to power-frequency magnetic fields increases the risk of cancer.

Many studies of the reproductive and developmental effects of exposure to ELF power-frequency electric and magnetic field effects and visual-display-frequency (VLF) magnetic fields using chick and mammalian species have been carried out. Overall, the data do not support the hypothesis that low-frequency EMF exposures result in reproductive toxicity.

Generally, the evidence for an effect of exposure to power-frequency EMFs on melatonin levels and melatonin-dependent reproductive status is mostly negative in seasonally breeding animals. In addition, no convincing effect on melatonin levels has been observed in a study of nonhuman primates chronically exposed to power-frequency EMFs. The interpretation of the outcome of studies on rat pineal and serum melatonin levels was more problematic: both positive and negative effects were reported in studies without apparent technical or analytical deficits.

With regard to possible effects on other hormones, with the possible exception of transient stress following the onset of ELF electric field exposure at levels significantly above perception thresholds, no consistent effects have been seen in levels of the stress-related hormones of the pituitary–adrenal axis, growth hormones and hormones involved in controlling metabolic activity or those associated with the control of reproduction and sexual development. There is little consistent evidence of any inhibitory effect of ELF EMF exposure to various aspects of the immune system function including those relevant to cancer. In addition, there was no compelling evidence that hematological variables were affected by exposure to ELF fields.

In general, many studies of nervous system functions were either negative, or difficult to interpret because of a weakness in the experimental design or confounding by artifact. However, several studies suggest possible EMF effects on the opioid and cholinergic systems, along with concomitant changes in analgesia, and in the acquisition and performance of spatial memory tasks. Effects on opioid and cholinergic systems should thus be further investigated.

The potential sensitivity of people with epilepsy to EMF exposure is an important factor that should be considered. Exposure of animals to EMFs before testing has been reported to have an inhibitory effect on subsequent tests of epileptic-seizure sensitivity but the type of response was not consistent among different studies. Further investigation needs to be carried out for possible effects during exposure.

There is convincing evidence that ELF electric fields can be perceived by animals, most likely as a result of surface charge effects; threshold values for both rats and baboons lie in the range 5–15 $\mathrm{kV\,m^{-1}}$. Exposures above threshold are sufficient to be mildly aversive and will result in transient arousal but mostly disappear following prolonged exposure.

Human Models

Small inconsistent changes have been described in the circulating levels of leukocytes, but as with animals, the relevance of these changes to human health is not clear. No indication of a clear-cut cytogenetic effect was observed in the blood cells of exposed subjects. The circulating hormone levels in humans, including melatonin, are not adversely affected by exposure to power-frequency electric and/or magnetic fields.

Positive results from studies on the power spectrum of different electroencephalogram (EEG) frequency bands and sleep structure have been

performed at occupational exposure levels higher than general environmental levels.

In the light of cognitive and performance studies that show a large number of positive biological effects, the need for further studies is warranted in order to clarify their significance for human health.

The majority of published cardiovascular changes stay within the range of normal biological variations; however, the study protocols and the results have not been successfully replicated in other laboratories.

As perception of electric or magnetic fields has been reasonably well characterized, skin symptoms, headache, and mood disturbance claimed by working near computer screens or in proximity of other environmental ELF sources have not reliably been linked to any electromagnetic exposure produced by these devices. Also, the studies performed do not indicate any ability of individuals with suspected electromagnetic hypersensitivity to react to any test-field situation, suggesting that other environmental or psychosocial factors should be looked for by people suffering from this pathology.

Conclusion

There is little scientific evidence suggesting that ELF EMF exposures pose a health risk. The strongest evidence for health effects comes from an association observed with childhood leukemia. In contrast, mechanistic studies and animal toxicology data fail to demonstrate any consistent pattern across studies, although sporadic findings of biological effects (including increased cancers in animals) have been reported. The lack of connection between human and experimental data (animal and mechanistic) severely complicates the interpretation of these results.

Radiofrequency Fields

Most of the research activity being devoted to the potential effects of mobile telephony equipment, only experimental findings related to these signals will be given in this section.

Mobile (or cellular) telephony has developed very rapidly over the past ten years. It is now part of the basic equipment of modern life and over 1.3 billion phones are in use worldwide. Concerns about health effects caused by exposure to the microwaves emitted by mobile telephones and base stations have increased over the last few years, becoming a major societal issue in some countries, or at least among part of the population. Stories about health risks from RF radiation (RFR) from mobile phones and base stations have become common in the media over the past five years.

Several systems are used in mobile telephony worldwide, all based on the same principle, that is, "cellular" mapping of the territory. In each cell, a base station emits toward and receives signals from the mobile telephones active in that cell (up to ~ 50). There are, for example, 30 000 base stations in France used by three networks. The carrier frequency varies from 400 to 2100 MHz and the voice or data information is coded digitally either by frequency or phase modulation. Mobile telephones are two-way radio transmitters operating in the 400–2100 MHz frequency range. In the global system for mobile communications (GSM), for example, the peak power emitted is 2 W, but the time-averaged power is always below 1/8 of this value, as power control reduces emission to the lowest level required. About half of the emitted power is absorbed by the user's head, that is, a maximum of 125 mW. Power absorption is expressed as specific absorption rate (SAR) in $\mathrm{W\,kg^{-1}}$.

Major improvements have been achieved in measuring the SAR in liquid phantoms and calculating power distribution in the head using numerical phantoms over the last ten years. Presently, the worst-case SAR associated with the average GSM phone on the market is $\sim 0.5\,\mathrm{W\,kg^{-1}}$, that is, 1/4 of the recommended local-exposure limit value. It is now known with certainty that temperature increase in the brain periphery caused by the waves emitted by mobile telephones does not exceed 0.1°C.

As a result of current changes in usage – increasing use of text and image messages and hands-free kits – mobile telephones are less frequently placed against the ear. This dramatically reduces exposure of the tissues in the head.

GSM base-station antennas have an emitting power of ~ 20 W. They are generally placed on rooftops and the emission beam is disk-shaped. Maximum exposure occurs on the ground, ~ 200 m from the base station and it is almost zero at the bottom of the building or mast on which the antenna is mounted. Exposure of the public to the RFR emitted by base stations is typically 1/10 000 of the recommended limit in terms of incident power. There is a consensus in the scientific community that base stations do not represent a health hazard.

There is much scientific evidence, based on existing research, that warrants limiting exposure to high-level RFR due to the "thermal effects" caused by heating of the tissues at SAR levels that correspond to a temperature elevation of a few degrees. However, this does not occur with mobile telephones. The search is thus for nonthermal effects and most of the research activity has been aimed at defining the thresholds for these effects, with respect to existing exposure guidelines based on acute effects known to be due to heating.

Health risk assessment associated with RFR benefits from a database spanning over 50 years: the WHO and Institute of Electrical and Electronics Engineers (IEEE) databases list ~1300 peer-reviewed publications, from biophysical theoretical analyses to human epidemiological studies. Half of these studies relating to cancer have been overwhelming in finding no evidence that RFR exposure initiates or promotes cancer. More than 350 studies are specifically related to mobile telephony.

Cellular Models

A number of replication studies that addressed some positive findings on enzyme activity, gene expression, and DNA alteration have all proven negative so far. Research is currently very actively investigating possible alterations of heat-shock proteins, seen as potential markers for RFR exposure and/or leading to physiological alterations in cells.

A wide range of short-term, low-level *in vitro* experiments have shown that exposure did not cause cell death, implying that RFR is not a toxic agent. Furthermore, the weight of evidence available at present (induction of DNA strand breaks, chromosome aberrations, micronuclei formation, DNA repair synthesis, sister chromatid exchange, and phenotypic mutation) supports the conclusion that RFR is not genotoxic. However, the synergy of RFR with chemical agents or other physical agents still needs further investigation.

WHO recommendations for short-term projects on cellular systems are: "the expression of stress (heat-shock) proteins in mammalian cells exposed to RF should be studied experimentally to follow-up recently published data. Biologically relevant hypotheses, if supported by experimental data (in particular, if related to the function of the CNS), should be tested to explore the utility of such data in risk assessment."

Animal Models

A large number of animal experiments have been performed over the past forty years, using various frequencies and modulations. It is clear from these data that the vast majority of the reported biological effects are due to heating. These effects result either from a rise in tissue or body temperature exceeding 1°C or in physiological and behavioral responses aimed at minimizing the total heat load.

Major improvements in exposure system design have made it possible to better characterize the SAR within the organism, and allow for either local exposure that mimics mobile telephone use (e.g., loop antenna, carousel) or whole-body exposure related to base stations (e.g., Ferris wheel, reverberation chamber, circular waveguide).

Results on most of the noncancer endpoints have been negative (memory, EEG, hearing, etc.) except for data on the permeability of the blood–brain barrier, which was found to be increased by two research groups but not by several others.

Therefore, most of the major ongoing studies deal with cancer models. All of the long-term bioassays or sensitized studies have given negative results except for one using transgenic mice, genetically modified to increase the background incidence of lymphomas; an increased tumor incidence was found following GSM exposure. No such finding emerged from a recent confirmation study, using a different design. While awaiting the results of a further replication study, there is no convincing evidence from animal investigations that the incidence of lymphomas and other types of tumors is influenced by lifetime daily exposure to mobile telephony RFR.

WHO recommendations for short-term projects on animals are: "Follow-up studies to immune system studies that suggest an effect of RF exposure (i.e., Russian publications from several years ago). Studies to assess the accuracy and reproducibility of published RF effects on the permeability of the blood–brain barrier and other neuropathologies (e.g., dura mater inflammation, dark neurones). Additional studies of the effect of RF exposure on sleep are recommended. More quantitative studies on the effects of heat on the development of the CNS, particularly the cortex, in the embryo and fetus using morphological and functional endpoints."

Human Models

In spite of the obvious limitations of human experiments in terms of endpoints and exposure characterization, several investigations have been performed using various models. Findings have either been negative or difficult to replicate (sleep, EEG, cognitive functions, etc.). Currently, there is no conclusive evidence from human studies of detrimental health effects of mobile telephones.

However, a report from a Dutch laboratory has drawn a lot of attention from the media and the scientific community as it reported minor effects on the well-being and cognitive functions of volunteers exposed to weak base-station signals. This work, which is not yet published, has been heavily criticized but its protocol may serve as the basis of some more refined work. However, it is not directly relevant to mobile telephone exposure.

Clearly, the main issue at present is the potential greater sensitivity of children to mobile telephone

RFR. Their lifetime exposure, the fact that their CNS is still developing, and, possibly, increased RFR absorption in the head, have led to concerns that cannot be easily resolved through laboratory investigations and numerical modeling.

WHO recommendations for short-term projects on humans are: "Replication and extension of the studies which demonstrated effects on sleep. Studies of RF exposure and headaches in a controlled laboratory setting. Studies of memory performance should be expanded to include children."

Health Risk Assessment

The process of health risk assessment by bodies such as International Commission on Non-ionizing Radiation Protection (ICNIRP), IEEE, International Agency for Research on Cancer (IARC), and WHO relies heavily on judging the quality of investigations. As stated above, the quality of exposure systems has greatly improved and can now be considered adequate. The use of well-grounded experimental protocols (sham-exposure, blinding of exposure and biological tests, positive controls) has become generalized. Moreover, it is now a common practice in the field of bioelectromagnetics to ascertain that any positive results are replicated in at least one independent laboratory. In spite of these improvements, it should be noted that only a few top-level biology laboratories have engaged in this type of research, partly due to interferences created by societal and media pressure.

Within its EMF International Project, WHO has reviewed the science and issued research recommendations. The main conclusion from these reviews is that EMF exposures below the limits recommended in the ICNIRP guidelines do not appear to have any known impact on health. However, there are still some key gaps in knowledge, requiring further research to provide definitive health risk assessments: IARC will issue a cancer classification of RFR in 2005, and WHO and ICNIRP's evaluations on RFR and health are due in 2006.

Remaining uncertainties in the science database have led to pressure to introduce precautionary measures until gaps in knowledge are filled. If precautionary measures are introduced to reduce RFR levels, it is recommended that they should be voluntary and that health-based exposure limits be mandated to protect public health.

Conclusion

RFR can cause biological effects when exposure is sufficiently intense: possible injuries include cataracts, skin burns, deep burns, heat exhaustion, and heat stroke. They are mostly due to heating. There have been scattered reports of effect, the nonthermal effects, that do not appear to be due to temperature elevation. None of these effects have been independently replicated, and most have no obvious consequence to human health. Furthermore, there are no known biophysical mechanisms which suggest that such effects could occur.

Following the very rapid development of mobile telephony, a major research effort has been carried out worldwide (tens of millions of euros per year). Europe is most active (UK, Germany, Italy, and Finland, in particular), but many research groups are contributing in Japan, US, and Australasia.

Most governments have addressed the issue of mobile telephony and health, and several international and national expert committees have written accurate summaries of current knowledge (see the "Further reading" section). Their conclusions converge toward an absence of health effects related to mobile telephones, but all encourage continuing research in some areas.

In answer to the question: "mobile telephony: is there evidence of harm?", one must conclude that the weight of scientific evidence does not support health concerns or indicate any health risk from mobile phones in normal use, nor is there any accepted mechanism for potential health effects at the low levels associated with these devices. Findings to date, including epidemiological studies and laboratory studies of animals, exposed both for a short-term and their entire lifetimes, have not provided evidence that exposure causes cancer, or affects biological tissues in a manner that might lead to, or augment, any disease. However, there are still some issues pending, in particular those related to the potentially greater sensitivity of children. The many ongoing research projects should help clarify these issues by the end of 2005.

See also: Electric and Magnetic Fields in Cells and Tissues; Metalloproteins, Electron Transfer in.

PACS: 41.20.Jb; 87.50.Jk; 87.50.Mn

Further Reading

IARC Monography no. 80 on ELF bioeffects at http://www.iarc.fr.

ICNIRP: Exposure to static and low frequency electromagnetic fields, biological effects and health consequences (0–100 kHz) – review of the scientific evidence and health consequences at www.icnirp.de

ICNIRP, International Commission on non-ionising radiation protection statement (1998) Guidelines for limiting exposure to time varying electric, magnetic, and electromagnetic fields (up to 300 GHz). *Health Physics* 74(4): 494–522.

IEGMP, Independent expert group on mobile phones and health, c/o National Radiological Protection Board, Chilton, Didcot UK ("Stewart Report" www.iegmp.org.uk), 2000.
NIEHS (1998) Niehs report on health effects of exposure to power-line frequency electric and magnetic field. www.niehs.nih.gov/emfrapid/home.htm.
NRPB (2001) *ELF Electromagnetic Fields and the Risk of Cancer*. Report of an advisory group on non-ionising radiation. Documents of the NRPB, vol. 12, # 1, National Radiological Protection Board, Chilton, Didcot ("Doll Report") UK.
NRPB at http://www.nrpb.org
WHO International EMF Project (1996) Health and environmental effects of exposure to static and time varying electric and magnetic fields, www.who.int/peh-emf
WHO research agenda at http://www.who.int/peh-emf

Electromagnetically Induced Transparency

M Artoni, University of Brescia, Brescia, Italy

Introduction

Three-level atomic and molecular systems coupled to two laser fields exhibit transparency effects that result from the cancellation of absorption at a resonance transition frequency. These effects are deliberately induced by one of the two laser fields upon modifying the medium optical response to the other field. The large degree of transparency in a medium where strong absorption and hence opacity would normally be expected, has then been termed as "electromagnetically induced transparency" (EIT). The effect is a rather subtle one. It relies on atomic states coherence and quantum interference, and the necessary background tools to understand the basic physical ideas behind EIT is presented here. A brief and preliminary discussion of coherent population trapping (CPT), a closely related effect whose observation first raised interest in induced transparency phenomena, is also presented.

From the very beginning EIT has stimulated a considerable amount of work on fundamental issues such as nonlinear optics at low-light levels, photon entanglement and entanglement of atomic ensembles, quantum information processing, quantum information storage, and enhanced acoustooptical effects, just to mention a few. In view of such a significant and rapid advancement, one may conclude by overviewing one of its most relevant aspects and prospects.

Coherence and Population Trapping

Cancellation of absorption due to the coherence between a pair of atomic states and subsequent trapping of the population is perhaps one of the earliest examples of transparency induced in an otherwise absorbing medium. Specifically, the coherent superposition of two atomic states leads, under certain conditions, to trapping of the population in one of the states and hence to transparency. This typically occurs in three-level atomic systems, or at least systems that can be adequately reduced to a three-level configuration when interaction with two electromagnetic fields is considered. A basic configuration leading to the effects of CPT is reported in **Figure 1**.

This phenomenon was first observed in Gozzini's group in Pisa and later in Stroud's group in Rochester. The Pisa group performed experiments that established a coherence between the Zeeman split lower levels of sodium atoms using a multimode laser. They employed a spatially varying magnetic field and observed a series of spatially separated dark lines, originally called "dark resonances," corresponding to the locations where the Zeeman splitting matched the frequency difference between modes of the coupling laser. The other experiment involved, instead, the hyperfine lower levels of atomic sodium and similar findings were observed.

The usual atomic dipole selection rules normally require that two pairs of levels are dipole coupled,

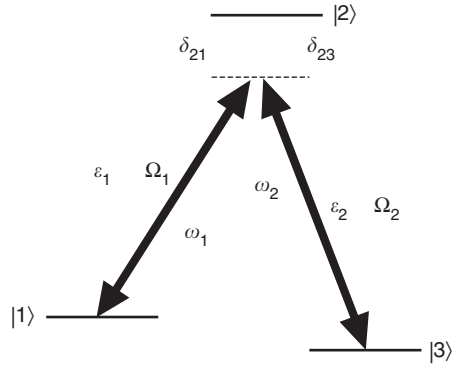


Figure 1 Lambda (Λ) scheme leading to coherent population trapping. The applied fields of amplitude $\mathscr{E}_1$ and $\mathscr{E}_2$ and with frequencies ω_1 and ω_2 need not be in single-photon resonance, but the two-photon resonance condition ($\delta = 0$) should be met as shown in the figure. Field strengths are typically of comparable magnitude ($\Omega_1 \simeq \Omega_2$). The states $|1\rangle$ and $|3\rangle$ are typically Zeeman or hyperfine sublevels of the ground state and are initially populated.

whilst the transition between the third pair is dipole forbidden. The basic scheme in **Figure 1** comprises the states $|1\rangle$, $|2\rangle$, and $|3\rangle$ coupled by two near-resonance laser fields whose strengths are defined in terms of the Rabi frequencies $\Omega_1 = \mu_{12}\mathscr{E}_1/\hbar$ and $\Omega_2 = \mu_{23}\mathscr{E}_2/\hbar$, respectively. Defining the frequency of transitions between states as $\omega_{21} = (E_2 - E_1)/\hbar$, $\omega_{23} = (E_2 - E_3)/\hbar$, and $\omega_{31} = (E_3 - E_1)/\hbar$, one can further introduce the one-photon frequency detunings $\delta_{21} = \omega_{21} - \omega_1$ and $\delta_{23} = \omega_{23} - \omega_2$ as well as the two-photon (Raman) frequency detuning $\delta = (\omega_{21} - \omega_{23}) - (\omega_1 - \omega_2)$.

In a simple treatment, the bare atom Hamiltonian $H_o = \sum_{i=1,3} E_i|i\rangle\langle i|$ should be modified to include the interactions due to the two atom-field couplings to obtain in the dipole approximation a total Hamiltonian of the form,

$$H = H_o + H_I \equiv H_o - \frac{\hbar\Omega_1}{2}|2\rangle\langle 1| - \frac{\hbar\Omega_2}{2}|2\rangle\langle 3| + \text{h.c.} \quad [1]$$

In general, the eigenstates of H turn out to be linear superpositions of the bare atomic states $|1\rangle$, $|2\rangle$, and $|3\rangle$. At exact resonance ($\delta_{21} = \delta_{23} = 0$), two of the eigenstates of H turn out to be symmetric and antisymmetric coherent superpositions of the two lower states of the bare atom basis, namely,

$$|C\rangle = \frac{\Omega_1}{\Omega}|1\rangle + \frac{\Omega_2}{\Omega}|3\rangle$$
$$|NC\rangle = \frac{\Omega_2}{\Omega}|1\rangle - \frac{\Omega_1}{\Omega}|3\rangle \quad [2]$$

where $\Omega = (\Omega_1^2 + \Omega_2^2)^{1/2}$. It should be noted that no component of the bare state $|2\rangle$ appears in these superpositions and that state $|C\rangle$ is coupled to the upper state $|2\rangle$ via the electric-dipole interaction whilst the state $|NC\rangle$ is not. This can easily be seen by evaluating the (dipole) transition matrix element $\langle 2|H_I|NC\rangle$. The two probability amplitudes that are summed to give the whole transition amplitude between $|NC\rangle$ and the atom's upper state $|2\rangle$ have in fact equal and opposite magnitudes. In the present lambda configuration, coherent trapping occurs due to the destructive quantum interference between the probability amplitudes associated with two (dipole) transitions. Conversely, the transition amplitude $\langle 2|H_I|C\rangle$ does not cancel to zero and $|C\rangle$ remains coupled to the upper state $|2\rangle$.

When the steady-state regime has been reached, the state $|NC\rangle$ will acquire all the atom's population under the combined effects of optical pumping from $|C\rangle$ to $|2\rangle$ and spontaneous emission from $|2\rangle$ to $|NC\rangle$. At last, no absorption process will depopulate the state $|NC\rangle$ where all population will remain trapped. The noncoupled state $|NC\rangle$ is often referred to as "dark" state whilst the other state $|C\rangle$, which remains coupled to the fields, is instead referred to as "bright" state.

Electromagnetic Induced Transparency

A highly opaque medium may be rendered almost transparent through EIT. Even in this case, as for CPT, cancellation of absorption relies on a process involving laser-induced coherence and quantum interference between the atomic coherences induced by the electromagnetic fields.

Formally, coherences are identified with the off-diagonal elements of the density matrix ρ. Off-diagonal elements of the density matrix play a critical role in the evolution of an atom coupled to electromagnetic fields. Many calculations on the effects of EIT in three-level systems are, therefore, developed in terms of the density matrix. Besides, this approach naturally lends itself to the inclusion of dampings that cause the decay of populations and coherences. Characteristic configurations for which EIT takes place in three-level atoms interacting with two near-resonance electromagnetic fields are shown in **Figure 2**. Atomic dipole selection rules, for example, in the lambda configuration of **Figure 2**, normally require that $|g\rangle - |e\rangle$ and $|e\rangle - |m\rangle$ transitions be dipole allowed but not the $|g\rangle - |m\rangle$ transition. Often $|m\rangle$ is a metastable state. In the Λ or ladder scheme, $|g\rangle$ is normally the ground state of the atom where the largest part of the population initially resides while states $|e\rangle$ and $|m\rangle$, which remain essentially unpopulated throughout the process, need not be ground state(s).

To understand how EIT works to reduce absorption when two laser fields interact with a three-level atom, for definiteness, the Λ configuration in **Figure 2** is examined. The states $|g\rangle$ and $|e\rangle$ are coupled by a probe field of amplitude $\mathscr{E}_p$ and frequency ω_p while the excited state $|e\rangle$ is coupled to $|m\rangle$ by a coherent field of frequency ω_c and Rabi frequency Ω_c. The one-photon detunings δ_p and δ_c as well as the two-photon (Raman) detuning δ are defined here as in the previous section. The off-diagonal decay rates for the coherences ρ_{eg}, ρ_{em}, and ρ_{gm} are denoted by γ_1, γ_2, and γ_3, respectively. The explicit form of the total Hamiltonian $H = H_0 + H_1 + H_2$ is

$$H_o = \hbar\omega_g|g\rangle\langle g| + \hbar\omega_e|e\rangle\langle e| + \hbar\omega_m|m\rangle\langle m|$$
$$H_1 = -\frac{\mu_{eg}\mathscr{E}_p}{2}e^{-i\omega_p t}|e\rangle\langle g| + \text{h.c.} \quad [3]$$
$$H_2 = -\frac{\hbar\Omega_c}{2}e^{-i\omega_c t}|e\rangle\langle m| + \text{h.c.}$$

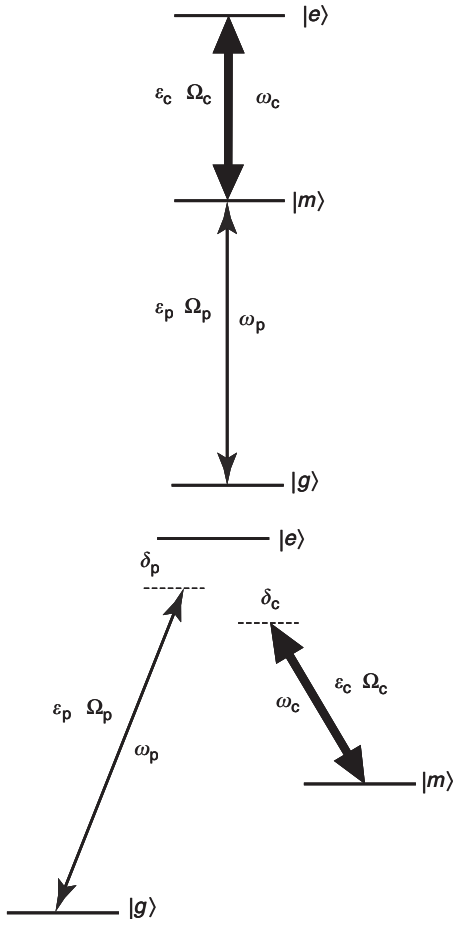


Figure 2 Basic schemes leading to electromagnetic induced transparency (EIT): (upper) a ladder scheme, (lower) a Λ scheme. In the Λ configuration levels, E_m and E_g are commonly almost degenerate. As for coherent population trapping (CPT) the two-photon (Raman) resonance condition should be met, yet the strength of one of the two fields could be appreciably larger than that of the other ($\Omega_c \ll \Omega_p$). State $|3\rangle$ in the Λ scheme need not be a ground state.

From the Liouville equation for the density matrix ρ, one obtains a set of coupled equations connecting the density matrix elements and their temporal derivatives. These equations are generally solved by numerical methods. If a steady-state limit is assumed, however, all time derivatives in the density matrix elements vanish. Under the additional assumption that the coupling field is much stronger than the probe, the former is the only field that must be retained to all orders. In these two limits, one arrives at a simplified form of the coupled density matrix equations and an exact solution can be attained. If initially the atoms are in the ground state $|g\rangle$, that is,

$$\rho_{gg}(0) \simeq 1, \quad \rho_{ee}(0) \simeq 0, \\ \rho_{mm}(0) \simeq 0, \rho_{em}(0) \simeq 0 \qquad [4]$$

the simplified set of coupled equations for a resonant coupling beam ($\delta_c = 0$) reads as

$$\dot{\tilde{\rho}}_{eg}(t) = -(\gamma_1 + i\delta_p)\tilde{\rho}_{eg}(t) + \frac{i\mu_{eg}\mathscr{E}_p}{2\hbar} + \frac{i\Omega_c}{2}\tilde{\rho}_{mg}(t) \qquad [5]$$

$$\dot{\tilde{\rho}}_{mg}(t) = -(\gamma_3 + i\delta_p)\tilde{\rho}_{mg}(t) + \frac{i\Omega_c}{2}\tilde{\rho}_{eg}(t) \qquad [6]$$

where $\delta_p = \omega_{eg} - \omega_p$ denotes the probe-field detuning while the new variables $\tilde{\rho}_{eg} = \rho_{eg}\, e^{i\omega_p t}$ and $\tilde{\rho}_{mg} = \rho_{mg}\, e^{i(\omega_p + \omega_{me})t}$ have been introduced. In the steady-state limit one has from [5] and [6],

$$\rho_{eg} = \frac{\mu_{eg}\mathscr{E}_p(\delta_p - i\gamma_3)}{(\delta_p - i\gamma_3)(\delta_p - i\gamma_1) - \Omega_c^2/4} \qquad [7]$$

The complex steady-state coherence [7] is directly proportional to the probe susceptibility χ_p and hence ρ_{eg} describes the medium optical response to the incident probe and, in particular, its absorption or transparency properties.

The expression for ρ_{eg} contains a number of terms in the various parameters that will lead to cancellation of its value, both real and imaginary parts, when two-photon resonance takes place. In this case, in fact, the real part of [7] is always zero while for appropriately strong coupling beams, or when $\Omega_c^2 \gtrsim \gamma_1\gamma_3$, the imaginary part of eqn [7] is seen to become several orders of magnitude smaller than the value it acquires when the coupling beam is absent, making the medium essentially transparent. It is likewise seen that the Lorentzian-like absorption dip obtained by expanding the imaginary part of eqn [7] around resonance exhibits a width given approximately by Ω_c^2/γ_1. Because there exists an upper bound to the coupling beam Rabi frequency Ω_c for EIT to take place, namely $\Omega_c \lesssim \gamma_1$, it follows that quenching of absorption only takes place over a small bundle of probe frequencies which is appreciably narrower than the natural line width γ_1. Thus, under the two-photon resonance condition and for Ω_c's such that

$$\sqrt{\gamma_1\gamma_3} \lesssim \Omega_c \lesssim \gamma_1 \qquad [8]$$

a probe field can propagate without absorption within a small transparency window, even though it would be strongly absorbed in the absence of the coupling beam.

Within the context of a density matrix formulation, the interference that leads to induced transparency manifests itself in the vanishing of the coherence ρ_{eg}. This stems from the existence of the coherence ρ_{mg} which is coupled to it and which only appears when the coupling laser beam is present. The

contribution from ρ_{mg} to the coherence ρ_{eg} may cancel, in fact, with the direct contribution driving this coherence due to the applied probe field (see eqns [5] and [6]).

The relevant optical response of an EIT medium is described by a linear susceptibility whose form depends, in general, on the presence of the external coupling beam. Imaginary and real parts of the susceptibility at the probe transition frequency ω_p are obtained from the macroscopic polarization $\mathscr{P}(\omega_p) = \varepsilon_o \chi(\omega_p)\mathscr{E}_p$, which can, in turn, be related to the microscopic coherence $\rho(\omega_p)$ at ω_p via the expression $\mathscr{P}(\omega_p) = (N/V)\mu_{eg}\rho(\omega_p)$. Here, N/V is the atoms density in the medium while μ_{eg} is the dipole matrix element associated with the transition. This relation holds for a medium sufficiently dilute that dipole–dipole coupling between atoms can be ignored. Besides, for the relatively large fields used in most EIT experiments, a semiclassical treatment, where the fields are treated classically (i.e., in terms of Maxwell's equations and susceptibilities) and the coherences are treated quantum mechanically with spontaneous decay added as a phenomenological damping, proves adequate. Only when atoms are coupled to cavities modes or when the statistical properties of light are important, a fully quantum approach is indeed required. In the semiclassical limit, the complex probe susceptibility for a resonant coupling beam ($\delta_c = 0$) can be written as,

$$\chi_p = 3\pi \mathscr{N}_p \frac{\Gamma_1(\delta_p - i\gamma_3)}{(\delta_p - i\gamma_3)(\delta_p - i\gamma_1) - \Omega_c^2/4} \qquad [9]$$

where $\mathscr{N}_p$ is the scaled sample average density $N\lambda\!\!\!^{-3}_p/V$ and $\lambda\!\!\!^{-}_p = \lambda_p/2\pi$ is the resonant probe reduced wavelength. The expression [9] depends upon parameters, namely detunings and laser intensities, that can all be directly controlled within an experiment. The real and imaginary parts of the refractive index $n_p \equiv \eta_p + i\kappa_p = \sqrt{1+\chi_p}$ associated with the medium dispersion and absorption, respectively, are plotted in **Figure 3** as a function of the probe detuning δ_p. The absorption vanishes at exact resonance and this is a striking result when compared with the situation in which the coupling field is absent ($\Omega_c = 0$).

Most EIT experiments have concentrated on ensembles of atoms. In these experiments, one measures the transmission of a weak probe in the presence of a strong-coupling field through an otherwise optically opaque medium. A landmark demonstration was performed by the Harris group at Stanford in two atomic systems, namely strontium and lead vapors. In both experiments, narrow-band pulsed laser radiation were used.

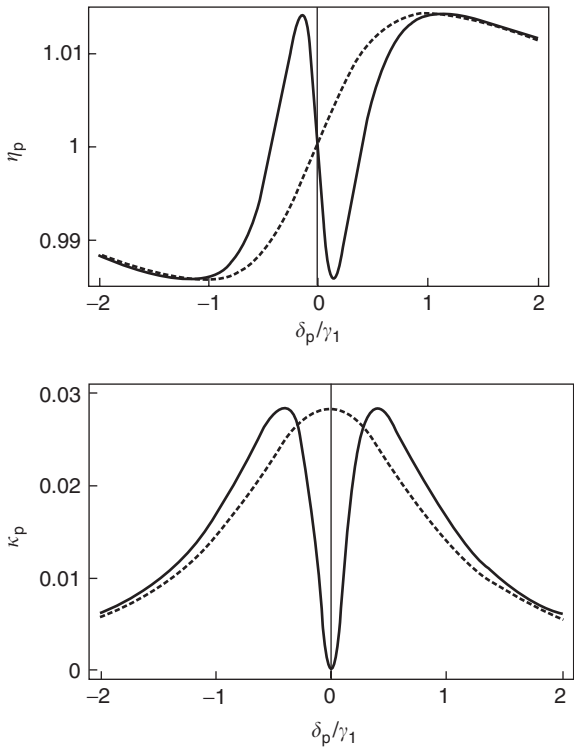


Figure 3 Real (upper) and imaginary (lower) part of the probe refractive index n_p as a function of the normalized probe detuning δ_p/γ_1 in the absence (dashed) and in the presence (solid) of a resonant coupling beam with Rabi frequency $\Omega_c = 0.8\gamma_1$. Due to destructive interference the imaginary part κ_p, associated with medium absorption, vanishes at exact resonance when the coupling field is on. The real part η_p, which determines the medium dispersion, is likewise modified in the presence of the coupling field and acquires a rather steep slope around resonance.

In the Sr experiment, the first to be reported and whose results are shown in **Figure 4**, the atoms are pumped into the $4d5d\,^1D_2$ autoionizing state starting from the $5s5p\,^1D_2$ state via a pulsed probe laser (337.1 nm) spanning the transition that was to be rendered transparent. A coupling laser (570.3 nm) was, instead, applied between this autoionizing state and the metastable bound state $4d5p\,^1D_2$. In the absence of the coupling beam, the probe field excited the system to the state $|2\rangle$, exhibiting strong absorption which made the Sr vapor completely opaque. Transmission, measured as the ratio of the transmitted to the incident intensity, was estimated to be e^{-20} at resonance. When the coupling laser was applied, instead, the resonant transmission increased dramatically to $\sim$40%.

In the Pb experiment, on the other hand, transparency was demonstrated using bound states of a collisionally broadened medium. Here, a ladder configuration was adopted where probe and coupling beams coupled the $6s^2 6p7s\ ^3P_1$ excited state respectively to the $6s^2 6p^2\ ^3P_0$ ground state and the state $6s^2 6p7p^3D_1$. These two experiments serve to

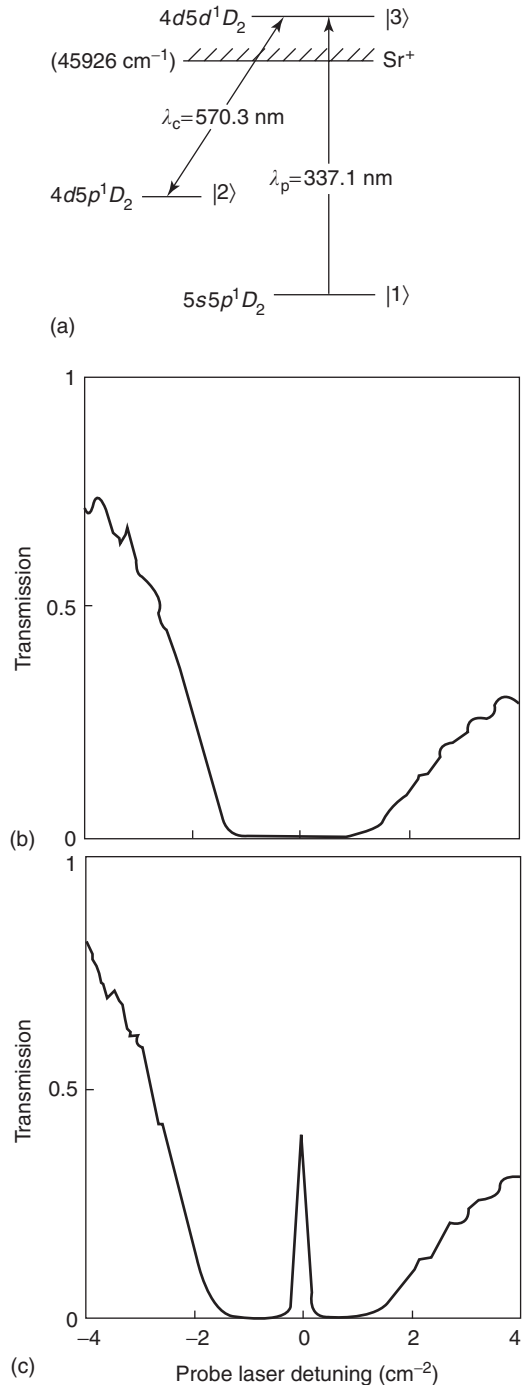


Figure 4 Electromagnetically induced transparency in strontium (Sr) vapors. (a) Sr energy levels and Λ configuration with coupling and probe beam of wavelength λ_c and λ_p, respectively. Transmission vs. probe laser detuning in the absence (b) and in the presence (c) of a coupling beam. Under the conditions of this experiment transmission goes from zero to $\sim$40%.

demonstrate the principle of EIT in three-level atomic systems. They also indicate how EIT can take place when the excited states are either autoionizing or collisionally broadened.

Fast and Slow Light

Along with induced transparency, there is a concomitant modification of the refractive properties of the medium. These are characterized by a very steep variation of the real part η of the medium complex refractive index n with frequency as shown in **Figure 3**. Here, for comparison, η is plotted both in the absence and in the presence of the coupling beam that causes the transparency.

In an ideal EIT regime, the atoms susceptibility vanishes from the light fields and at resonance, and the complex index n remains close to unity throughout the EIT transparency window. The propagation speed c/n of a monochromatic beam of light is then essentially unchanged from its speed c in vacuum. This means that the propagation velocity of the beam phase front in the medium, that is, its phase velocity, is equal to c.

Yet for a light pulse, which contains, instead, several Fourier components at different frequencies, the small variations in the phase velocity that each component experiences in the medium can add up to make an appreciable effect. Components with slightly different speeds get, in fact, out-of-phase with each other, with the net result that the velocity of the pulse envelope, that is, its group velocity, is different from c.

Consider, for instance, the superposition of two monochromatic plane waves of the form $\cos[(\omega_o \pm \delta\omega)\ t - (k_o \pm \delta k)z]$. Addition gives $2\cos[\delta\omega(t - (\delta k/\delta\omega)z)]\cos(\omega_o t - k_o z)$, which is a wave with an amplitude that retains its shape as it propagates with the velocity $\delta\omega/\delta k$. More generally, if it is superposed over many waves whose frequencies and wave numbers are distributed in a small range about the center values ω_o and k_o, one obtains a wave of the form $A(t - z/v_g)\cos(\omega_o t - k_o z)$. The intensity, averaged over a few optical periods, is proportional to $A^2(t - z/v_g)$, which describes a light pulse propagating without change of shape at a group velocity v_g defined as $d\omega/dk$.

Using the relation $k = \omega n/c$ with $n = \eta + i\kappa$, one can, for very small absorptions ($\kappa \ll 1$), express v_g in terms of the real part η of the refractive index and its derivative $d\eta/d\omega$, so that for a given frequency ω one has

$$v_g(\omega) \equiv \frac{d\omega}{dk(\omega)} = \frac{c}{\eta(\omega) + \omega\, d\eta(\omega)/d\omega} \qquad [10]$$

In the presence of EIT, $d\eta(\omega)/d\omega$ is large and positive around resonance as one can see from **Figure 3** and according to [10], very small group velocities can be attained. By substituting the expression [9] into [10], one has indeed at the resonance

frequency ω_o,

$$\frac{v_g^o}{c} \simeq \frac{\Omega_c^2}{6\pi \mathcal{N}_p \gamma_1 \omega_o} \qquad [11]$$

Note that v_g^o depends on the control field intensity, proportional to Ω_c^2, and on the atomic density N/V. Decreasing the control beam power or increasing the atomic density can make v_g^o orders of magnitude smaller than the speed of light c in vacuum.

Evidence for ultraslow light is typically provided by measuring the time delay between a pulse propagating in a medium of given length L, and a twin pulse that has instead propagated in vacuum. The leading edge of a pulse that enters the medium is, in fact, rapidly decelerated while its tail outside the medium still propagates at speed c. As a result, upon entering the medium, the spatial extent of the pulse is compressed (by the ratio c/v_g) and its energy decreased as photons are being expended to establish the coherence between the states $|g\rangle$ and $|m\rangle$, or, in other words, to change the atomic state, with any excess energy carried away by the control field. As the pulse exits the medium, its spatial extent increases again and the atoms return to their original ground state. The pulse as a whole, however, has experienced a peak-shift given approximately by

$$\Delta \simeq L\left(1 - \frac{c}{v_g^o}\right) \qquad [12]$$

with a concomitant pulse delay τ given by Δ/c.

The possibility of manipulating the group velocity in EIT media was first pointed out by Harris. The subject, however, was truly brought to focus by a remarkable experiment of Hau in which an ultracold gas of Na atoms was used to slow light pulses down to few tens of $m\,s^{-1}$. The large pulse delay τ measured in this experiment is reported in **Figure 5** for different experimental conditions. Although the first demonstration of ultraslow light used ultracold gases, shortly thereafter, two groups reported ultraslow light propagation in hot Rb gases. This work was a milestone on the way to making ultraslow light "on the cheap," that is, without expensive cold-atom traps.

Furthermore, it follows from eqn [10], that group velocities can become infinite or even negative when $d\eta(\omega)/d\omega$ is sufficiently negative, that is, when the real part of the refractive index decreases with frequency. Because in the visible, $\eta(\omega)$ normally increases with frequency, the case $d\eta(\omega)/d\omega < 0$, for historical reasons is called anomalous dispersion. Directly from **Figure 3**, one sees that in EIT media $\eta(\omega)$ acquires negative slopes slightly off-resonance. A negative group velocity means that the peak of the pulse emerging from the medium moves faster than the pulse peak that has propagated in vacuum at speed c. This, according to eqn [12], leads to a positive peak shift Δ or delay τ rather than negative ones as obtained in ultraslow light experiments. Such anomalous group velocities can occur without significant pulse distortion. An infinite group velocity, on the other hand, means that the peak of the pulse emerging from the medium occurs at the same time as the peak of the pulse entering the medium as seen again from eqn [12]. The pulse, in other words, appears to cross the medium instantaneously. Superluminal peak velocities have been observed, though with substantial pulse attenuation and pulse compression. Such a superluminal behavior does not contradict relativity as the pulse peak velocity is not, in general, the velocity at which a signal can be transmitted.

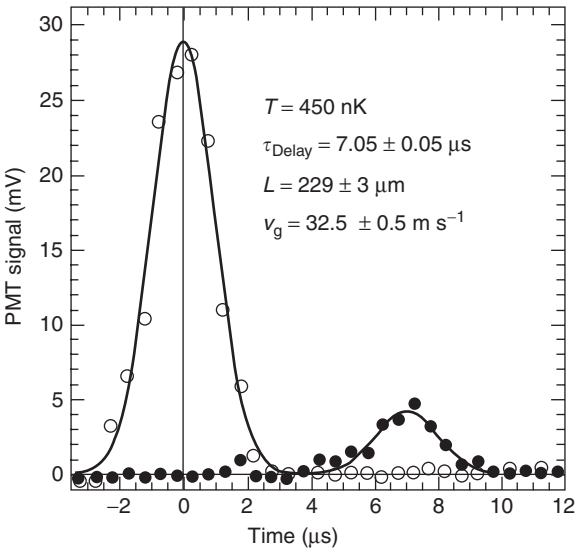


Figure 5 Pulse delay measurement. The front pulse (open circles) is a reference pulse that has propagated in vacuum. The other pulse (filled circles) is delayed by 7.05 μs in a 229-μm-long atom cloud. The corresponding light speed is $32.5\,m\,s^{-1}$. The curves represent Gaussian fits to the measured pulses.

Unlike the pioneering experiments that demonstrated anomalous group velocities of single-photons wave packets carried out nearly a decade ago by Chiao at Berkeley, the anomalous dispersion regime attained by using EIT has been shown to lead to a much larger superluminal effect. Again, these experiments in no way contradict the principle of Einstein causality, that is, the principle that no signal can propagate faster than c.

Slow Light: Some Prospects

Recent advances in quantum information science have led to many interesting new concepts such as

quantum computation and quantum memories. The practical implementation of these concepts involves storing, retrieval, and transporting quantum states between different nodes of a network. Methods for manipulating photons and their quantum information in EIT atomic media have recently been suggested. Photons are indeed the most promising carriers of quantum information: they are fast, robust, and readily available. Atoms, on the other hand, represent reliable and long-lived storage units. These methods essentially consist in trapping single-photon Gaussian pulses in atomic media by adiabatically reducing their group velocity to zero and subsequently releasing them after a suitable interval of time. This represents a reversible technique for coherent transfer of quantum information carried by light pulses to atoms and vice versa.

Photon–photon interaction strength is typically very weak and conventional nonlinear optics requires powerful laser beams tightly focused in nonlinear materials. Ultraslow light pulses in EIT media may experience, however, exceedingly large nonlinearities so that nonlinear optical processes may become efficient even at energy densities as low as a few photons per atomic cross section. Low-light-level nonlinear optics has recently become quite relevant in the context of resonant four-wave mixing, teleportation of atomic ensembles, and production of correlated photon states.

Furthermore, laser pulses slowed down to the speed of sound can produce strong coupling between acoustic waves and the electromagnetic field. This might be employed for efficient multiwave mixing as well as for novel acoustooptical EIT-based devices. Proof-of-principle experiments in support of some of these ideas have already been conducted while others are just underway.

Finally, slow and stopped light can also be attained in solid materials enabling important steps toward many potential applications. Recent attempts have indeed led to the observation of group velocities of the order of few tens of $\mathrm{m\,s^{-1}}$, for example, in ruby crystals either at cryogenic temperatures of few degrees Kelvin or at room temperature.

See also: Nonlinear Optics.

PACS: 42.50.Gy; 42.50.Hz; 32.80.t; 42.65.Ky; 42.65.Lm

Further Reading

Bigelow M, *et al.* (2002) *Physical Review Letters* 88: 023602.
Bigelow MS, Lepeshkin L, and Boyd RW (2003) Observation of ultraslow light propogation in a ruby crystal at room temperature. *Physical Reveiw Letters* 90: 113903.
Boller KJ, Imamolu A, and Harris SE (1991) *Physical Review Letters* 66: 2593.
Field JE, Hahn KH, and Harris SE (1991) *Physical Review Letters* 67: 3062.
Harris S (1997) Electromagnetically induced transparency. *Physics Today* 50(7): 36.
Hau LV, Harris SE, Dutton Z, and Behroozi CH (1999) Light speed reduction to 17 meters per second in a ultracold atomic gas. *Nature* 397: 594.
Lukin M (2003) Trapping and manipulating photon states in atomic ensembles. *Review of Modern Physics* 75: 457.
Marangos JP (1998) Electromagnetic induced transparency. *Journal of Modern Optics* 45: 471.
Matsko AB, Kocharovskaya O, Rostovtsev Y, Welch GR, Zibrov AS, *et al.* (2001) Slow, ultraslow, stored and frozen light. *Advances in Atomic Molecular and Optical Physics* 46: 191.
Milonni PW (2002) Controlling the speed of light pulses. *Journal of Physics B* 35: 471.
Scully M and Zubairy M (1997) *Quantum Optics*. Cambridge: Cambridge University Press.
Turukhin AV, Sudarshanam VS, Shahriar MS, Musser JA, Ham BS, *et al.* (2002) Observation of ultraslow and stored light pulses in a solid. *Physical Review Letters* 88: 023602.

Electron and Positron Sources

S Tazzari, Università di Roma "Tor Vergata," Rome, Italy
M Ferrario, Laboratori Nazionali di Frascati of Istituto Nazionale di Fisica Nucleare, Rome, Italy

Introduction

A particle source is a device producing particle beams with well-defined and controlled intensity, divergence, mean energy, and energy spread.

Electron beams are used in science and technology to both investigate and modify matter. Applications in science mainly concern material investigation, processing and the study of elementary particle, nuclear and atomic physics. Technological applications cover a vast number of fields, including particle accelerators and lasers, microscopy, microprocessing, communications, medical diagnostics and treatment, welding and machining, and household appliances. Positron beams, more difficult to produce, are almost exclusively used in basic material, nuclear and atomic sciences. Electrons can be directly produced using thermal or field emission processes. Positron beams

are obtained either from radioactive sources, or from primary electron or photon beams via the pair production mechanism.

Particle beams are accelerated by DC or alternating (RF) electric fields, handled and transported using electric or magnetic components equivalent to lenses and prisms. One therefore speaks of beam optics. Because particles are charged, space charge, radiation, and other possible forces between particles, including those mediated by the beam environment, have to be taken into account in addition to optical aberrations.

The main parameters characterizing a particle beam are energy E, energy spread ΔE, momentum ($\boldsymbol{P}$), current (I) and brightness (B). The beam current is defined as $I = qN\,\beta c$, q being the particle charge, βc the particle velocity, and N the number of beam particles. Brightness (see below) is a measure of the current flowing through the unit area and solid angle and ultimately of the current that, for a given energy spread and optics, can be focused onto a spot of a given size: for a given current, the higher the brightness the smaller the spot, down to the diffraction limit. A brief introduction to beam dynamics, given below, is required to precisely define the terminology and relations between the parameters.

Depending on their main application, electron sources can be loosely classified as

- low (DC) current, high brightness, industrially produced guns designed for "micro-applications" such as electron microscopy, material analysis and processing, microlithography, micromachining;
- high average (DC or pulsed) current, moderate brightness "conventional" guns to drive devices such as appliances, power tubes, klystrons, and conventional-type linear accelerators (Linacs);
- high-(pulsed) current, high-brightness, high-voltage "preinjector" sources, usually including additional acceleration stages, for high-energy accelerators, synchrotron radiation sources, and free electron lasers.

Positron sources, devoted to science, are instead best classified into low intensity, radioactive and nuclear sources; and high intensity ones generated via pair production.

Basic Beam Dynamics Elements

The equivalent of the optical axis in a beam optical transport system is the trajectory followed by an ideal reference particle having the nominal beam momentum $\boldsymbol{P}_\mathrm{o}$ and an initial position. The position and momentum of the generic beam particle are defined relative to those of the reference particle and described by a point in the six-dimensional (6D) phase space whose coordinates are (x, p_x, y, p_y, s, p_s). The transverse space and momentum coordinates, x, p_x, y and p_y, as well as p_s, are in general, functions of the longitudinal coordinate s along the trajectory. The reference system is fixed to the reference particle so that the latter always occupies the origin. Introducing the familiar from optics angular coordinates x', y', the prime indicating the derivative with respect to s, the 6D phase space reduces to [x, $\beta\gamma x'$, y, $\beta\gamma y'$, s, p_s], with γ the reference particle Lorentz factor.

Consider an ideal paraxial beam of N noninteracting particles (forces between particles negligibly small), subject only to external conservative forces (no radiation) with constant or slowly varying momentum P_o and p_x, $p_y \ll P_\mathrm{o}$. In this case, the transverse [x,$\beta\gamma x'$, y, $\beta\gamma y'$] and longitudinal [s, p_s] phase space projections become (to first order in the coordinates) uncoupled; it can then be shown (Liouville's theorem) that the volume and area respectively occupied by the beam particles in each of the two projections are constants of motion. From the invariance of the beam occupied area in [s, p_s], it follows that $\Delta E/I$ is a constant of motion. In transverse phase space, consider first a cylindrically symmetric beam with radius r and angular half aperture θ: the occupied phase space volume $(\beta\gamma)^2 \Delta S\, \Delta\Omega \approx \pi^2 (\beta\gamma \cdot r\theta)^2 \equiv \pi^2 \varepsilon_\mathrm{n}^2$, with ΔS the beam cross section and $\Delta\Omega$ the beam subtended solid angle, is invariant and so is the here-defined quantity ε_n called energy-normalized (or reduced) beam emittance. Note that the quantity $\varepsilon = \pi^2 (r\theta)^2$, is the area occupied by the beam in the physical position–angle space. Called physical emittance, it is the relevant quantity in physical space and, because the invariant quantity is $\beta\gamma \cdot r\theta$, it becomes smaller as the beam is accelerated. The normalized brightness is defined as $B_\mathrm{n} = I/\pi^2 \varepsilon_\mathrm{n}^2$ and is seen to be none other but the energy-normalized phase space current density. (Note that brightness is often defined in literature either, incorrectly, as a current density $B = \Delta I/(\Delta S.\Delta\Omega)$, or, correctly, in normalized form, as $B_{\mathrm{n}V} = I/(\Delta S \cdot \Delta\Omega \cdot V)$, qV being the reference particle kinetic energy. $B_{\mathrm{n}V}$ and our adimensional energy dependence definition B_n are related through equation $B_\mathrm{n} = (E_\mathrm{o}/q) \cdot B_{\mathrm{n}V}$, with E_o the particle rest energy.) As expected, the smaller the beam radius and angular aperture, the brighter the beam. More generally, when motions in x and y are uncoupled, the horizontal and vertical normalized emittances, $\beta\gamma \int z'\, \mathrm{d}z \equiv \pi\varepsilon_{\mathrm{n}z}$ (z stands for x or y), defined in the respective (z, p_z) phase space projections, are separately constant. (Here the factor π in the definition of $\varepsilon_{\mathrm{n}z}$ arises from the beam boundary in 2D phase space

being an ellipse when the simple approximation is made of external forces linear in the particle displacements.) The resulting normalized brightness value is $B_n = I/\pi^2\ (\varepsilon_{nx} \cdot \varepsilon_{ny})$. For a Gaussian beam with standard deviations $\sigma_z(s)$ and $\sigma_z'(s)$ (z standing for either x or y), the most used normalized emittance definition is $\varepsilon_{nz} = \beta\gamma\ \sigma_z\sigma_z'$.

When dealing with nonideal, irregular beam shapes and current densities, it becomes useful to introduce the rms (or average) normalized beam emittances

$$\pi\varepsilon_{nz}^{(rms)}(s) = \sqrt{\langle z^2 \rangle \langle (\beta\gamma \cdot z')^2 \rangle - \langle z \cdot \beta\gamma z' \rangle^2}$$

To define the actual "beam stay clear" envelope in either physical transverse space projection, most authors use the expression $4\pi\varepsilon_{nz}^{(rms)}$, the resulting rms beam brightness being $B_n^{(rms)} = I/(16 \cdot \pi^2\ \varepsilon_{nx}^{(rms)}\ \varepsilon_{ny}^{(rms)})$.

Note finally that when in real life one has to deal with nonlinear optical elements and internal forces, the rms normalized emittances are no longer invariant. They will, in general, depend on the beam local current density, tending to blow up during beam acceleration and transport, thereby spoiling $B_n^{(rms)}$ (see **Figure 4**). A great amount of theoretical and experimental work has been and is, therefore, devoted to methods and techniques to minimize the final output brightness deterioration particularly of high-current sources. Last, keeping the beam energy spread small is also important particularly to reduce chromatic aberrations; the ratio of $B_n^{(rms)}$ to the beam relative momentum (energy) spread $\Delta p/P_o$, called rms normalized spectral brightness or brilliance, B_{nS}, quantifies this aspect.

Electron Sources

General

The typical e^- source (electron gun) consists of a cathode, the primary emitter of low-energy (few eV) electrons, and of a system to accelerate the electrons and shape them into a paraxial beam. Emission from the cathode material is basically obtained lowering the material work function by either heating or by means of external electric fields or photoelectrically. The most common cathodes are metal filament or disk-shaped thermionic emitters. Field effect and photoelectric cathodes are used for special applications.

Electron Sources for Micro-Applications

The main requirement for microscopy type applications is highest space resolution in the plane of the object to be examined or processed, and consequently, highest source brightness and brilliance. Micromachining and semiconductor processing put rather more emphasis on high-power density and scanning capabilities, requiring higher source currents and somewhat lower brightness. The first element of a micro-application-type source is a "needle" cathode (**Figure 1**) producing a cylindrical symmetry beam. An accelerating and focusing column, comprising one or more electrodes and electrostatic or magnetic lenses, collects the cathode-emitted electrons and focuses them into the sample to be examined or processed. Cathode tip diameters down to the sub-μm range are used, a size limiting the maximum extracted current to typically a few tens of μA. Further, collimation down to a few nA (effective current) takes place in the column allowing for very small normalized emittances ($\varepsilon_{nz} \approx 10^{-8}$–$10^{-11}$ m rad) and nanometer size spots at the sample. The resulting normalized brightness can exceed 10^{12} A $(m^{-2}\,rad)^{-2}$ with energy spread in the 10^{-5} range. Electrons are extracted from the cathode by either thermionic emission from heating it to ≈2000–2800 K, or by cold field emission (CFE) in an electrostatic field strong enough for electrons to "tunnel" through the surface barrier at room temperature. A third technique (Schottky emission) combines thermo and field emission allowing for lowering of the operating temperature to ≈1700 K. A low operating temperature makes for longer cathode lifetime and higher brightness since beam emittance is ultimately determined by the thermal motion of the extracted electrons. With a cylindrical cathode of radius r, the thermal contribution to the normalized emittance is approximately $\varepsilon_{th}^2 \approx r^2 \cdot K_B T/E_o$, K_B being the Boltzmann constant and T the absolute cathode temperature. CFE cathodes can therefore ultimately produce the smallest emittance, but, at extracted

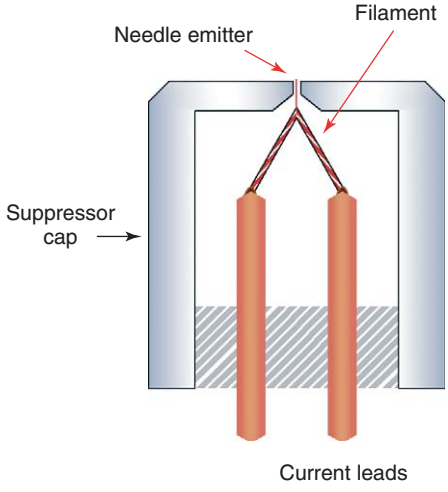


Figure 1 Schottky type cathode cross section.

Table 1 Cathodes for micro-applications: order of magnitude parameters for various cathode materials

Main parameters	*W*	LaB_6	*Cold field*	*Schottky*	*Carbon nanotubes*
Effective source size (nm)	$\geqslant 10^4$	$\geqslant 10^3$	3	10	3
Operating temperature (K)	2800	2000	300	1800	300
Useful current (nA)	1000	1000	1	10	
Normalized brightness ($A\,m^{-2}\,rad^{-2}$)	10^{10}	10^{11}	$>10^{13}$	10^{13}	10^{14}
Invariant emittance (m rad)	10^{-9}	5×10^{-9}	5×10^{-12}	10^{-11}	
Energy spread (eV)	1.0	1.0	0.5	0.5	0.3

current values comparable to those of the other two types, their lifetime is shorter. Schottky cathodes are therefore the most popular solution for highest resolution and brilliance, associated with a long lifetime.

Cathode materials vary from metal (mainly W) filaments to single crystals of various metals or of low work-function composites such as LaB_6. A typical Schottky cathode arrangement is shown in **Figure 1**: the few μm-radius single-crystal tip, mounted on a hairpin filament that heats it to the desired temperature, protrudes through a hole in a negatively biased shield (Wehnelt cap) so that only electrons emitted from the very tip can enter the column. In the column the beam is further collimated, accelerated to energies from a few hundred eV to a few tens of keV (typically ≈ 25 keV) and focused down into the sample to be studied or processed.

Schottky cathodes made of carbon nanotube bundles are also under development: very high brightness has been demonstrated in test stands.

A typical order of magnitude performance of micro-application-type sources equipped with different kinds of cathodes is presented in **Table 1**.

Figure 2 The cathode (gray disk) assembly of an S-band Linac gun. (Courtesy of Sincrotrone Trieste.)

Conventional-Type Electron Guns

Electron guns covered in this section are used to drive a great variety of devices such as power tubes, klystrons, and linear accelerators for industrial, medical, and scientific applications. The emphasis being on high current, ease of operation, lifetime and price, they are mostly rugged thermionic cathode DC guns of the diode or triode (gridded) type. The gun voltage and current specifications vary widely, depending on the application, from the $\approx$50–100 kV/$\approx$10 A of Linac guns to the $\approx$100–1000 kV/$\approx$100–1000 A of the highest power klystron guns.

Common cathodes consist of a high melting-point metal disk coated with a low work-function material and indirectly heated by a filament. The extracted current density $J_T = A\,T^2 e^{-(W/K_B T)}$ is a function of temperature and of the material work function, W. The most used material is Ba or Sr oxide coated tungsten having a work function of $\approx$1.6 eV, compared to 4.5 eV of pure tungsten, and capable of 10–20 $A\,cm^{-2}$ at temperatures of $\approx$1500 K. Because in a high-temperature and high-voltage environment the coating tends to be removed rather rapidly by desorption, sputtering, and evaporation, dispenser cathodes have been developed made of porous tungsten impregnated with oxide which slowly diffuses towards the cathode surface, continuously replacing the removed material. Operated in good vacuum ($\leqslant 10^{-7}$ hPa) and at current densities $\approx 5\,A\,cm^{-2}$, a dispenser cathode can typically reach useful lifetimes of $\sim$40 000 h. The gun operates ideally in the space-charge limited Child–Langmuir regime: a voltage applied to Pierce geometry electrodes (**Figure 2**) provides the radial focusing required to form the cathode-emitted particles into an ideally laminar flow beam. Laminarity ensures a constant current density emission at the cathode, lowest emittance, and highest brightness. Normalized emittance and brightness values lie typically in the $\varepsilon_n^2 \sim 10^{-3}$–$10^{-5}\,A\,m^{-2}\,rad^{-2}$ and $B_n \sim 10^8$–$10^{11}\,A\,m^{-2}\,rad^{-2}$ range, respectively.

The relation between applied voltage and extracted current is $I = P^* V^{3/2}$, P being a constant,

characteristic of the device, called perveance. Its measurement unit is the perv$\equiv 1\,A/V^{3/2}$. Highest brightness values are obtained with perveance values of less than 1 μperv because above it, space-charge forces make it increasingly difficult to produce laminar beams. For very high peak and average currents more complicated gun designs are therefore needed, using annular beams or multiple beam (cathode) devices (**Figure 3**). Multiple-beam klystron guns having several large cathodes (up to $\approx 100\,\mathrm{cm}^2$) and operating at $\leqslant 2.5\,\mathrm{A\,cm}^{-2}$ s are being developed for applications aimed at MW average power with 10^5 h lifetimes. The various cathode beams are combined after acceleration so as to minimize space-charge effects.

Pulsed beam configurations are obtained by various methods, such as pulsing either the anode or the grid voltage, or periodically sweeping the continuous beam across a collimator (chopping), or by imposing a periodic longitudinal velocity chirp on the beam in such a way as to make particles bunch together in a following field-free region (drift space) of appropriate length (velocity bunching). The main parameters of a typical S-band Linac gun and of DC and pulsed high-power klystron guns are listed in **Table 2**.

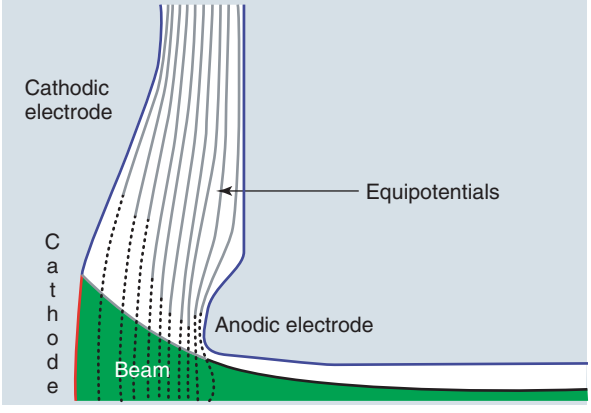


Figure 3 Pierce-type geometry scheme (upper half of cross section shown).

Electron Preinjectors

To drive free electron laser (FEL)-based, fully coherent, short-wavelength (X-ray) sources and to implement techniques such as laser and plasma acceleration for a future generation of compact electron accelerators requires the most demanding beam parameters: $\varepsilon_{nr} < 1$ μm rad, $(\Delta E/E) < 0.1\%$, $B_{nr} \sim 10^{15}$ A (m rad)$^{-2}$, $\tau \approx ps$ and peak bunch current $\hat{I}_b \sim$ kA. (Such specifications are beyond those at present routinely obtainable. Physics and technology of generating and transporting very high brightness, intense beams are therefore being actively researched and developed worldwide and advancing very fast.) Next generation colliders for particle physics also need a similar source performance and, in addition, high bunch charge, polarized, ribbon-shaped beams.

Preinjector systems (so-called because they are normally only the first part of the injector to the final accelerator) have been developed to approach the above listed specifications, consisting of a high-energy gun and of a number of auxiliary components. A several MeV electron gun is the first component needed to accelerate the source-emitted electrons as fast as possible to relativistic energies to minimize space-charge force effects, scaling like $1/\gamma^2$, which tend to blow the beam emittance up. It can be shown, by ideally slicing up the bunch like a loaf, that the emittance at the gun output has a fan-like structure in phase space (**Figure 4a**) due to space-charge forces being stronger on slices near the bunch center than on tail ones. Each bunch slice thus contributes one of the ribs of the fan, and the rms normalized emittance of the whole fan becomes much larger than that of any individual slice. The fraction of the fan-like blow-up caused by the linear components of the space-charge force can be taken out (compensated) by adding a solenoid at the gun exit followed by a drift space, and by additional accelerating sections and by proper shaping the bunch-charge distribution (should ideally be uniform). How the scheme works is seen in the simulation whose results are presented in **Figures 4** and **5**. The simulation follows a uniform charge distribution through

Table 2 Parameters of typical power thermionic electron guns: for an S-band linear accelerator injector and for power klystron tubes

Electron guns	*V* (KV)	*I* (A)	*Perv.* (μperv)	$\langle W \rangle$ (kW)	T_M (μs)	*Rep rate* (Hz)
S-band linac	100	10	0.3	1.0	10	50–60
High power	100	20	0.6	1000		DC
	500	210	0.6	1	1.5	10
	535	700	2.0	35	3	60
	800	4000	0.9	8	1	5

V: gun voltage; *I*: av. beam pulse current; $\langle W \rangle$: average beam power; T_M: pulse duration.

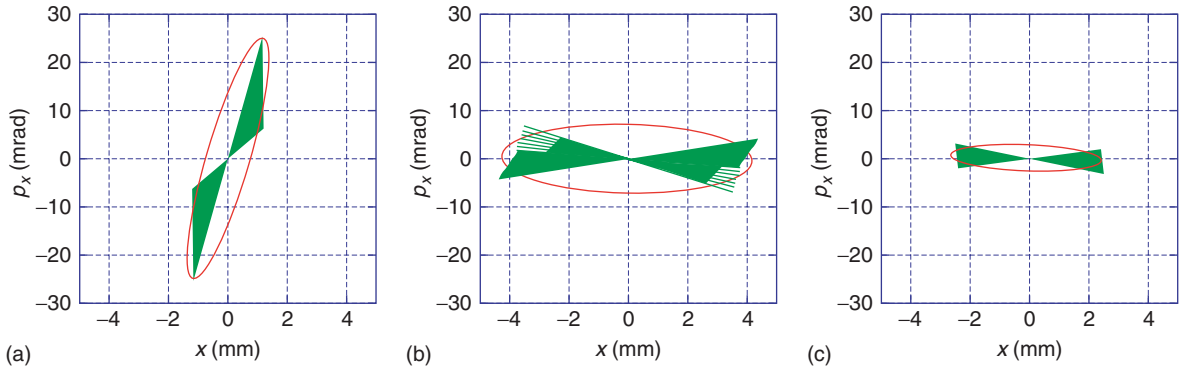


Figure 4 Fan-like beam footprint in the (x, p_x) phase plane at: (a) the gun exit; (b) the solenoid exit; (c) the drift space point where a first minimum is obtained. The beam-stay-clear footprint is shown as an ellipse whose area is four times the computed rms normalized fan area, $\pi\varepsilon_n^{(rms)}$.

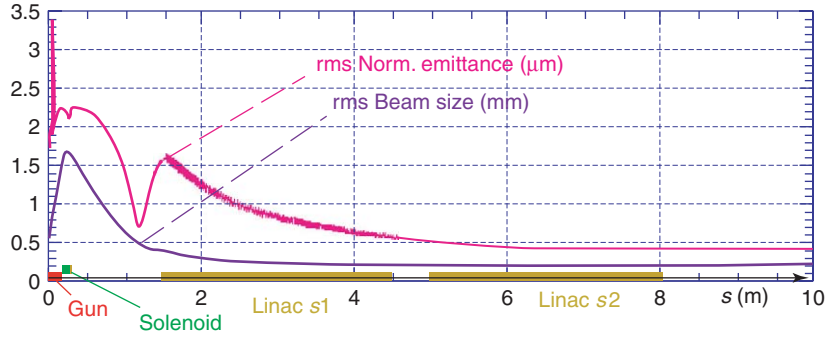


Figure 5 Computed rms normalized emittance and rms horizontal envelope evolution in the preinjector. The emittance is taken through a minimum at $s = 1.5$ m by the compensating solenoid and the following drift space. It then starts increasing again and is damped back in the next two accelerating Linac sections.

from cathode to the end of such a space-charge compensated preinjector whose two-Linac sections provide acceleration to 150 MeV.

Figure 4a shows the beam phase-space footprint at the gun output, and its beam-stay-clear equivalent ellipse having four times the area of the fan-like rms distribution. **Figures 4b** and **4c** show how the compensating solenoid and the following drift space reduce the footprint. A further reduction of the rms emittance in the final accelerating sections is shown in **Figure 5**. The end ε_n value is essentially due to thermal contributions and to nonlinear space charge force components only.

Peak-normalized brightness values measured at state-of-the-art preinjector installations are shown in **Figure 6**. Normalized emittances in the 2–4 μm rad range are routinely obtained today at peak bunch currents up to $\hat{I}_b \sim 50 - 100$ A.

To raise the peak bunch current, preinjectors are being developed using the preinjector additional accelerating sections to reduce the bunch length also: energy-chirping of the moderately relativistic (≈ 5 MeV) bunch produced by the gun and injection into the following accelerator sections far off the RF

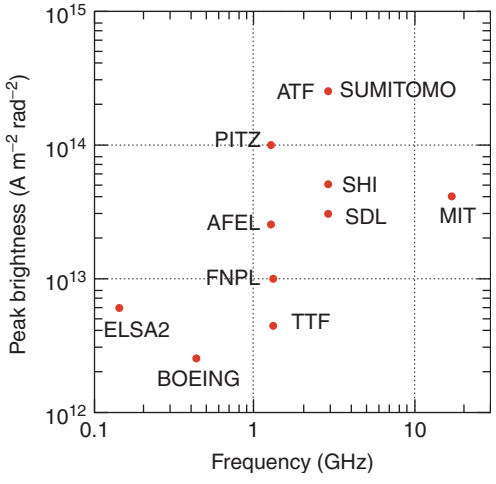


Figure 6 Peak normalized brightness of state-of-the-art preinjectors, mostly operating at L-, S-, or K-band frequencies with bunch charges of ≈ 1 nC. Emittances are in the 2–4 μm rad range.

field crest produces a velocity bunching so that the bunch itself is simultaneously compressed and accelerated. Focusing solenoids are added around the accelerating sections to prevent an emittance blow-up.

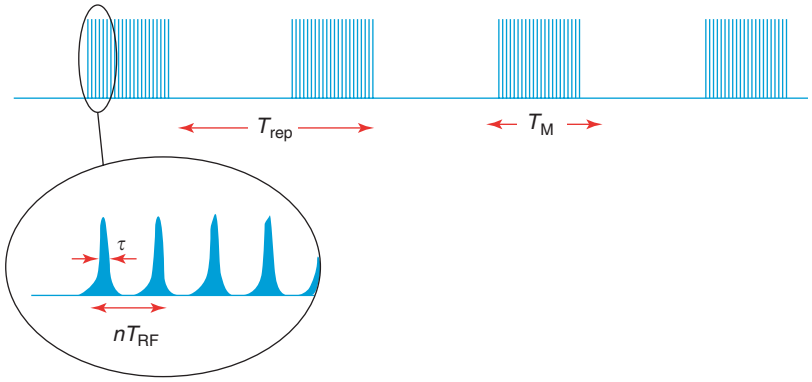


Figure 7 Time structure of a resonant pulsed linear accelerator. T_M long macropulses repeat with period T_{rep}. The microstructure repetition period is (equal to or) multiple of the accelerating field period T_{RF}.

Concerning the beam time structure, preinjectors generate bunch trains whose typical configuration is shown in **Figure 7**: a train of (macro) pulses, each containing a train of short bunches. The macropulse duration T_M and repetition frequency, $1/T_{rep}$, are usually determined by the allowed average power consumption of the particle source and of the accelerator. Duty factors range from $\sim 10^{-5}$ to $\sim 10^{-1}$. The bunch repetition frequency in the macropulse is a subharmonic of the accelerating field frequency $1/T_{RF}$, which also determines the initial bunch duration, τ.

Preinjector high-voltage guns in operation are basically of two types. The first consists of a normal thermionic DC gun followed by a high frequency, standing or traveling wave-accelerating structure. The second (RF gun), simpler and more compact, combines gun and accelerating structure in one by directly embedding the cathode in a standing-wave RF cavity. A $\approx$50–100 MV m^{-1} peak RF field in the $\approx$10 cm long cavity provides very fast acceleration to several MeV, thereby minimizing the effect of the defocusing space charge forces. The cathode of either type gun can be either thermionic or a photocathode from which electrons are photoelectrically extracted by illuminating it with a laser beam.

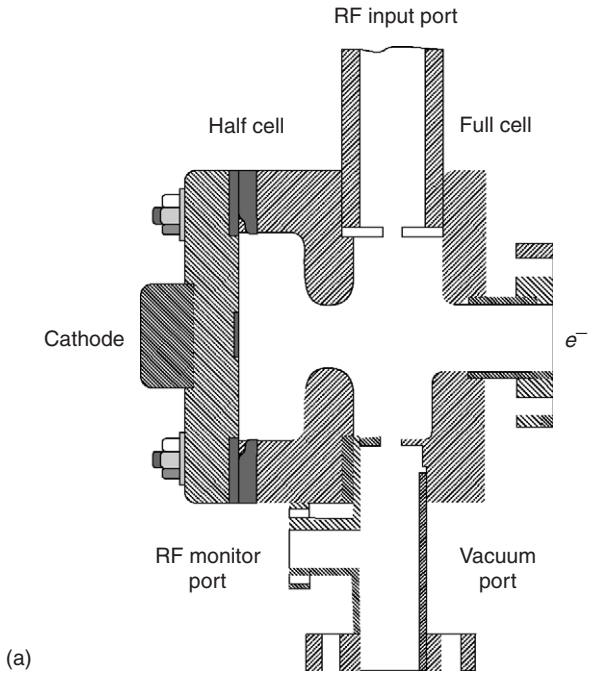
Schematic diagrams of RF gun structures are shown in **Figures 8a** and **8b**, one mounting a thermionic cathode, the other a photocathode.

The main asset of thermionic, metallic cathodes is stability of the emitted current and long lifetime. Drawing from micro-applications technology single crystal, mm-diameter LaB_6 cathodes are being proposed to achieve a high current with emittance values at the thermal limit. The required bunch microstructure must be produced by chopping or velocity-bunching their continuous macropulse current.

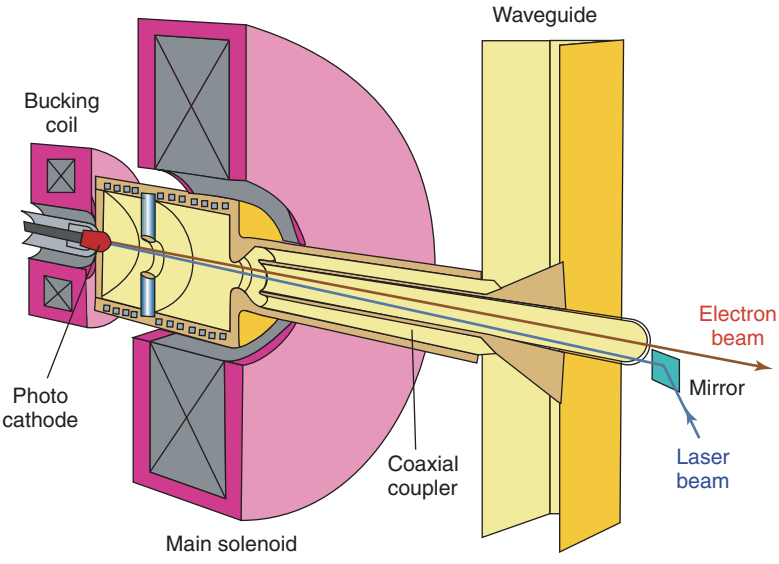
With photocathodes, the beam microstructure can be directly generated by modulating the drive laser pulse length, shape, spot size, and intensity on a pulse-to-pulse basis. On the other hand, stabilising to percent accuracy, the extracted current, which depends on the pulse-to-pulse stability of the high-power high-repetition rate driving laser, is more difficult. Nevertheless, because they offer much higher extracted current densities than thermionic devices, photocathodes are today's choice for high-peak current, bright-beam applications.

Photocathode materials such as Cu and Mg, to be illuminated with UV light, can operate in pressures of $\sim 10^{-7}$ hPa, have low, albeit constant, quantum efficiencies (QE) in the 10^{-3}–10^{-4} range. Alkali (Cs, K, Rb, RbCs) telluride semiconductor devices, illuminated with 4.5 eV UV light, require ultrahigh vacuum ($\approx 10^{-9}$ hPa) but have high efficiency and long lifetime: Cs_2Te photocathodes have been routinely operated in accelerators with initial QE of $\approx$10%, gradually dropping to a few percent over many months of useful lifetime.

Negative electron affinity (NEA) photocathodes have the advantage of operating at longer laser wavelengths. Illuminated with polarized light, NEA-GaAs devices can also produce 80–90% polarized beams; they are being developed aiming at ps long, $\approx$1 nC bunches with $\varepsilon_n \approx 5 \times 10^{-5}$ m rad. For acceptable quantum efficiency lifetime though, they must operate in $\sim 10^{-11}$ hPa vacuum, a specification not easy to meet with high-frequency RF guns because of the intrinsic low pumping conductance of the RF structure. Preinjectors equipped with easier-to-evacuate DC guns, low-frequency (larger iris) RF guns or open RF gun structures of the "plane wave transformer" type are therefore also being developed. Other photoemitters such as diamond-based, ferroelectric ceramic, and nanostructured carbon cathodes are also being studied. Parameters of the most commonly used photocathode materials are collected in **Table 3**.



(a)



(b)

Figure 8 (a) Schematic layouts of a thermionic cathode equipped RF gun. (Courtesy of the LCLS Collaboration, SLAC, USA.) (b) Schematic layout of photocathode equipped RF gun. RF power is fed coaxially from behind the cathode. (Courtesy of the TESLA Collaboration, DESY, De.)

Table 3 Parameters of various photocathode materials

Material	*Quantum efficiency* (*QE*)	*Lifetime*	*Required vacuum* (hPa)	*Laser wavelength* (nm)
Cu	2×10^{-4}	Years	10^{-7}	260
Mg	2×10^{-3}	years	10^{-7}	260
Alkali Antimonide	0.05	days	10^{-10}	532
Alkali Telluride	0.1	>1 year	10^{-9}	260
NEA-GaAs	0.05	weeks	10^{-11}	780

Gun developments include RF devices capable of generating bright, continuous microbunch trains for applications requiring the highest possible average current such as CW operation of superconducting linear accelerators.

New types of preinjectors are also being studied, such as high-brightness devices, promising to directly generate 100 fs, 0.1 nC bunches with peak currents up to $\approx 10^3$ A, by integrating the gun into the preinjector high-energy accelerating section. A photocathode is mounted only a few mm away from the back wall of a Linac section serving as anode. By generating GV m^{-1}, sub-ps pulses across the mm wide cathode-to-anode gap, photoelectrons are accelerated to several MeV in a few fs and directly injected into the preinjector final acceleration section, so that essentially no space-charge-induced emittance blow-up is expected.

More exotic schemes are being investigated for applications to further future laser–plasma accelerators. As an example, an "all-optical" electron source for a laser-driven plasma wake-field accelerator is claimed to be capable of GV cm^{-1} electric fields, micrometer spot sizes, femtosecond-long bunches, and GA cm^{-2} current densities. The scheme foresees a first-drive laser creating a plasma wake by ionizing a confined gas to below its self-trapping threshold (free electrons oscillating around ions). The transverse ponderomotive force of a second (injection) laser pulse, shining orthogonal to the first, then imparts to the plasma electrons an extra kick in the wave direction so as to inject them into the plasma wake-field with the correct phase to be trapped and accelerated.

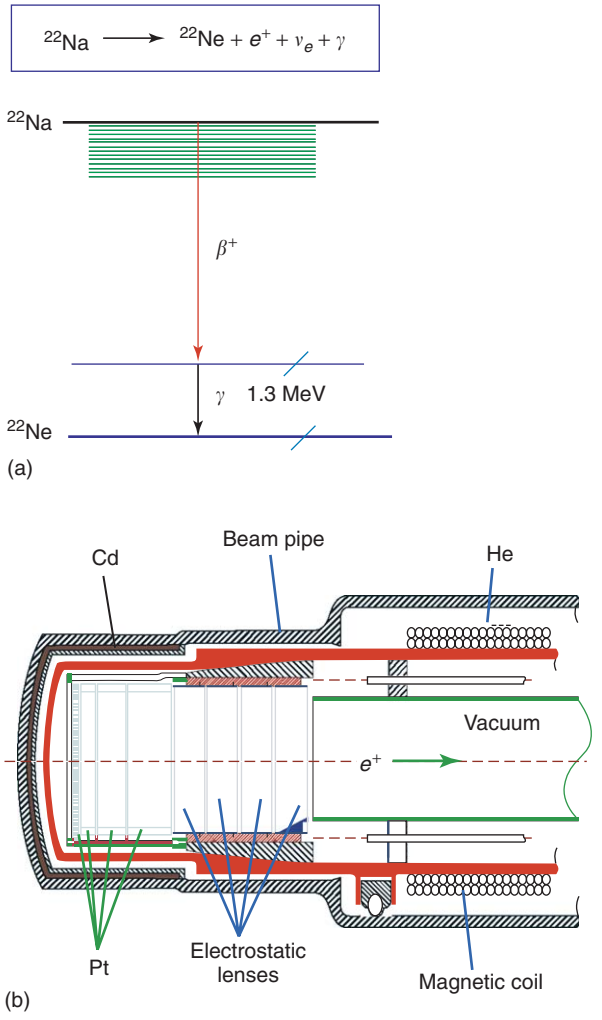


Figure 9 (a) Main decay mode of isotope ^{22}Na; (b) Slow neutron driven source. (Courtesy of Prof. C. Hugenschmidt, TUM, Garching.)

Positron Sources

Positron Production

Positron beams are produced either by β-radioactive sources or via pair production by photons. Photons are normally obtained either from neutron-induced nuclear reactions such as $^{113}\mathrm{Cd}(n,\gamma)^{114}\mathrm{Cd}$ or from electron-induced electromagnetic showers in solid targets. Radioactive and nuclear e^+ sources are mainly used to produce low intensity, quasi-monochromatic positron beams in the <1 eV to $\sim$MeV range, used in materials science for surface and bulk investigations. Radioactive sources are relatively inexpensive and suited for low-energy laboratory installations, while neutron-induced production requires reactor-grade slow n beams. For high intensities, primary high-energy electron or γ-ray beams must be used.

Radioactive and Nuclear e^+ Sources

The most popular radioactive source is the ^{22}Na isotope which decays, according to the diagram of **Figure 9a**, with 2.6 years half-life. One of its advantages is that the fast γ-ray accompanying the decay can be used to perform e^+ annihilation lifetime measurements. Positrons are emitted with a very large energy spread $\sim$200 keV, extending to 540 keV. A common method to produce a beam with the desired energy and a narrow energy spread is to first thermalize the decay positrons in a system of tungsten vanes, from which they emerge with energy of $\sim$0–2 eV, and then re-accelerate them. Using electrostatic acceleration, continuous positron beams with typical DC current values of a few pA, energies up to several MeV, and relative energy spreads of $\sim 10^{-4}$ can be obtained that, for microprobe-type applications, can be focused to spot sizes of a few μm. Higher intensities, up to $\approx$100 pA can be obtained from other, less easily handled sources such as ^{64}Cu.

The design of a much higher intensity e^+ source based on the $^{113}Cd(n,\gamma)$ ^{114}Cd neutron capture reaction is schematically shown in **Figure 9b**. Slow neutrons from a reactor are captured in a Cd blanket, each capture event producing a cascade of γ-rays. The high-energy photons are then converted into e^+ e^- pairs in a set of Pt absorbers that also serve to slow the positrons down. A system of electrical lenses and solenoids finally collects the slow particles in a continuous beam with design intensity of up to ~ 10 nA. Reactors are, in principle, capable of orders of magnitude higher e^+ production rates but only over much larger target volumes, which makes it difficult to collect the particles in a compact beam.

High-Intensity e^+ Sources

High-intensity positron beams are obtained by letting primary, high-energy, intense e^- beams shower in a converter, the shower photons producing the positrons via pair production, or by directly hitting the converter with a photon beam. The relatively low energy, large emittance, large energy spread e^+ positron beam out of the converter is captured in a system of magnetic lenses surrounding the converter, followed by an accelerating structure.

The converter thickness for maximum e^+ yield depends on the nature and energy of the primary. The number of photons in an e^- initiated shower increases approximately proportional to the primary electron energy (E_e) for $E_e \geqslant 1$ GeV, while the thickness yielding the maximum collectable e^+/photon ratio increases proportional to $\approx \ln(E_e)$. At $E_e \sim 1$ GeV, it is $\approx 3\,X_o$ (X_o being the converter material radiation length $\propto 1/Z^2$) and the e^+ effective source radius at the converter exit, which affects the positron collection efficiency, is $\approx 0.3\ X_o$ neglecting the contribution from the primary beam emittance. The positron energy spectrum exiting the optimum target thickness peaks at around a few MeV tailing out at a few tens of MeV, almost independently of the primary energy. A typical yield figure, when collecting $\approx$2–20 MeV positrons, is $\sim 5 \times 10^{-2}$ positrons per incident electron and per GeV of the primary beam.

The overall source intensity is ultimately limited by the target material sustainable peak and average thermal stresses from local, sudden temperature peaks due to the energy deposited in the converter by sub-μs electron bunches over very small volumes. Materials with short X_o, high-mechanical strength and high-melting point are thus favored, the most used being tungsten (^{74}W). Specific energy deposition in a ^{74}W converter can typically not exceed a few tens of J g^{-1}. For very high intensity e^+ sources, cooled rotating converters are used. Liquid metal converters are also being studied. Large amounts of radiation and residual activity produced in the target surroundings require strict maintenance and personnel protection measures.

Higher positron beam intensities can be produced using primary intense, bright, photon beams, such as can be obtained from free electron lasers or from high-energy e^- beams spontaneously radiating in an undulator magnet. High-power laser beams Thompson back-scattered from a multi-GeV electron beam are also candidates being studied to produce high-energy photon beams.

Because the photon pair production cross section saturates at ≈ 1 GeV, where the positron yield per photon peaks at approximately $0.3\,X_o$, low Z, high-heat capacity converter materials (typically ^{22}Ti) can be used, offering less multiple scattering and correspondingly higher collection efficiency, less energy deposition per collected e^+, and therefore, higher allowed deposited energy density. In addition, more than 40% linear or circular-polarized positrons can be obtained by using polarized primary photons.

Finally, it has recently been shown that the periodic atomic potential inside a single crystal converter can act like a very short period undulator magnet and that the positron yield from electrons or photons incident at glancing angle to the crystal axis can be two to three times larger than from an amorphous converter.

See also: Core Photoemission; Local Field Effects; Valence Photoemission.

PACS: 41.85. – p; 41.75.Fr, Ht; 29.25.Bx; 29.25. – t; 52.38.Kd

Further Reading

Artru X (2003) Advantages of axially aligned crystals used in positron production at future linear colliders. *Physical Review Accelerators and Beams* 6: 91003.

Boscolo I and Michelato P (2000) Photocathodes: The state of the art and some news. *N. I. & M. A* 445: 389.

Chao AW, Tigner M (eds.) (1999) Particle sources. In: *Accelerator Physics and Engineering*. Singapore: World Scientific.

Clendenin JE (2000) Recent advances in electron and positron sources. *American Institute of Physics Conference Proceedings* 569: 563.

Hugenschmidt C, Koegel G, Repper R, Schreckenbach K, Sperr P, *et al*. (2002) Monoenergetic positron beam at the reactor based positron source at FRM-II. *Nuclear Instruments and Methods in Physics Research B* 192: 97.

Humphries S Jr. (1990) *Charged Particle Beams*. New York: Wiley.

Jansen GH (1990) *Coulomb Interaction in Particle Beams*. Academic Press.

Levinson HJ (2001) *Principles of Lithography,* SPIE Press Monograph, vol. PM97.

Reiser M (1994) *Theory and Design of Charged Particle Beams*. New York: Wiley.

Serafini L and Ferrario M (2001) Velocity bunching in photoinjectors. *American Institute of Physics Conference Proceedings* 581: 87.

Serafini L and Rosenzweig JB (1997) Envelope analysis of intense relativistic quasilaminar beams in RF photoinjectors: A theory of emittance compensation. *Physical Review E* 55: 7565.

Tazzari S and Ferrario M (2003) Trends in high energy particle accelerators. *Reports on Progress in Physics* 66(6): 1045.

Williams DB and Carter CB (1996) *Transmission Electron Microscopy*. New York: Plenum.

Nomenclature

B, B_n	brightness, normalized brightness
c	speed of light
e	electron charge
e^-, e^+	electron, positron
E	energy
E_o	particle rest energy
E_e	electron energy
I	beam current
I_b, $\hat{I}_b$	bunch current, peak bunch current
m_o	particle rest mass
N	number of particles
$\boldsymbol{P}$, $\boldsymbol{P}_o$	momentum, nominal beam momentum
q	charge
Q	quality factor
s	longitudinal coordinate
v	velocity
V	voltage
W	work function
x, y	horizontal, vertical coordinates
X_o	radiation length
z	z stands for either x or y
Z	atomic number
β	reduced velocity v/c
γ	Lorentz factor
$\Delta\boldsymbol{P}$	relative momentum deviation $\boldsymbol{P}-\boldsymbol{P}_o$
ε, ε_n, ε_{th}	emittance, normalized emittance, thermal emittance
τ	bunch duration

Electron Gas (Theory)

M P Tosi, Scuola Normale Superiore, Pisa, Italy

Introduction

The homogeneous gas of electrons (also known as jellium) is the most important basic model in the physics of condensed matter. It was introduced by Drude in 1900 to account for the electromagnetic properties of plasmas and was used by the pioneers of quantum mechanics in the early 1930s to model the sea of interacting valence electrons in a metal. The valence electrons are the outer electrons of an atom and determine how it interacts with its surroundings, but in a metal they are released by the atoms into a collective state. In the early studies, the ions of the crystalline lattice in the metal were smeared into a rigid uniform background of positive charge, which exactly balances the negative charge associated with the distribution of outer electrons. More generally, through the development of density-functional theory the electron gas has become the basic reference system in most calculations of electronic structure not only for metallic or insulating solids and liquids, but also often applied to describe the electronic structure and the cohesion of atomic-scale systems up to biomolecules.

In metals the density of valence (or conduction) electrons is high enough that many of their properties can be understood in terms of a picture in which they move almost independently of each other, forming a paramagnetic fluid made of equal numbers of up and down spins. The success of this single-electron picture of metals rests on the work carried out in the 1950s by Landau in developing the theory of Fermi liquids. But with decreasing density the potential energy associated with the Coulomb interactions between the electrons grows to dominate over their kinetic energy and ultimately at very low density the electron gas freezes into the "Wigner crystal," in which the electrons become localized on lattice sites. This behavior is just the opposite of what happens with classical systems, where it is high density that favors a lattice structure. It is also believed that crystallization of the electron gas in the jellium model is preceded, in an intermediate range of density, by the stabilization of a spin-polarized state in which the spins spontaneously align into a ferromagnetic state of order. This possibility was already foreseen by Bloch in 1929, within the context of the Hartree–Fock approximation. This theoretical framework, when widened to include man-made semiconductor devices in which the electronic carriers are constrained to move in a plane (as in a quantum well) or on a line (as in a quantum wire) or under three-dimensional confinement (as in a quantum dot), provides much of the basic understanding that is necessary for modern condensed matter physics and nanotechnology.

It is appropriate to note at this point, before embarking on a more detailed presentation of the issues indicated above, that especially over the last few decades some more exotic properties of the valence-electron gas have come to the fore. These developments really started with the idea of Cooper pairs in the BCS theory of superconductivity in metals and have been stimulated by discoveries of novel materials, such as the cuprate high-T_c superconductors and other oxides, or of novel physical effects such as the Kondo scattering against magnetic impurities or the fractional quantum Hall effect. The quantum theory of solids is now no longer confined to the electron as the quasiparticle of a Fermi liquid, but in special situations reconstructs the system of valence electrons to produce quasiparticles that may behave like fragments of an electron, such as composite fermions, composite bosons, or holons and spinons.

Ideal Fermi Gas and Fermi Liquid Theory

Conduction electrons in normal metals under ordinary laboratory conditions form a highly degenerate Fermi fluid, in which the mean interparticle spacing a is a small fraction of the characteristic de Broglie wavelength $\lambda = h/p_{dB}$ (here h is Planck's constant and p_{dB} is the momentum of a particle of mass m having kinetic energy equal to the thermal energy $k_B T$, with k_B the Boltzmann constant and T the temperature). If for the moment one neglects all interactions, one can write the energy levels E of the gas in terms of the single-particle kinetic energies ε_k and of the number n_k of electrons in a single-particle state at energy ε_k, $E = \sum_k n_k \varepsilon_k$. Here k is a label specifying the momentum and the spin of each electron. The Pauli exclusion principle restricts the values that n_k may take for particles of half-integer spin to $n_k = 0$ or 1.

A cell of volume h^3 in phase space can thus contain at most two electrons with opposite spins, so that in the ground state at $T = 0$ the electrons must have a spread of momenta in a range up to a maximum momentum p_F (the "Fermi momentum"). For N electrons with up and down spins inside a volume V and defining $k_F = p_F/\hbar$, one can write the density $n = N/V$ as $n = k_F^3/3\pi^2$. The wave number k_F is the radius of a spherical surface in wave number space (the "Fermi surface") and the energy of an electron with momentum corresponding to the Fermi surface is the Fermi energy ε_F. This also is the lowest energy with which one may add a further electron to the gas (i.e., ε_F is the chemical potential of the gas at $T = 0$) and evidently increases with the electron density, since $k_F = (3\pi^2 n)^{1/3}$.

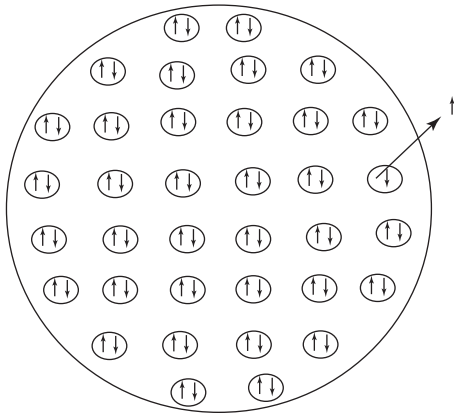


Figure 1 Schematic representation of a section of the Fermi sphere for a paramagnetic electron gas, showing also a process of excitation of an electron–hole pair.

The notion of a Fermi sphere allows a vivid picture of ground state and excitations in a normal Fermi fluid. At $T = 0$, the single-particle states inside the Fermi sphere are fully occupied by the fermions and the states outside it are empty. The momentum distribution (i.e., the occupation number of an electron state of given momentum and spin) thus takes a discontinuous change from 1 to 0 as the momentum crosses the Fermi surface. Thermal excitations at finite temperature bring fermions from states inside to states outside the Fermi surface, leaving "holes" behind and lowering the chemical potential of the gas (see the schematic picture in **Figure 1**).

It is somewhat surprising that a number of properties of normal metals are observed to behave as predicted by the ideal Fermi-gas model, in view of the electron–electron and electron–ion interactions that may be expected to lead to substantial departures from ideality. In particular, the electronic contribution to the specific heat of a metal increases linearly with temperature and the paramagnetic spin susceptibility is independent of temperature, both quantities being proportional to the density ν_F of electronic states at the Fermi surface. The explanation of these behaviors has been given by Landau: he assumed that in a Fermi liquid the net effect of the interactions would be to "dress" a particle excited out of the ground state with a cloud of surrounding particles, building up what is termed a quasiparticle. Once one deals with quasiparticles rather than with "bare" electrons, the excitations of the fluid can be described as an almost ideal gas of quasiparticles. A crucial point in Landau's theory is that the quasiparticles obey Fermi statistics, so that at low temperature the low-lying excited states in terms of the excitation of quasiparticles across the Fermi surface can be described. The correct behaviors of the electronic specific heat and spin susceptibility follow at once.

There is, however, one major difference to be emphasized between quasiparticles and bare particles, that is a quasiparticle carries a renormalized effective mass from the backflow that through its interactions it induces in the surrounding fluid as it tears through it. This implies a shift in the density of states at the Fermi surface and hence in the absolute values of the thermodynamic properties of the interacting electron gas relative to the ideal Fermi gas. The spin susceptibility is enhanced by a further factor which favors the alignment of the spins of neighboring electrons into a parallel configuration (see below).

Exchange Effects and Hartree–Fock Theory

As a start to a calculation of the ground-state energy of the interacting electron gas, one might ask what is the distribution of electron pairs in the ideal Fermi gas as a function of their relative distance r. The pair distribution function, that is written in general as $g(r)$, is defined by setting equal to the quantity $ng(r)(4\pi r^2 \mathrm{d}r)$ the average number of electrons that lie within a spherical shell centered on an electron and having radius r and thickness dr. For the electron gas with equal numbers of up and down spins it can be written as $g(r) = [g_{\uparrow\uparrow}(r) + g_{\uparrow\downarrow}(r)]/2$, the average of the values for pairs of electrons with parallel spins and antiparallel spins.

Whereas $g(r) = 1$ in a classical ideal gas, the probability of finding two electrons with the same spin at a distance r vanishes exactly as $r \to 0$. This property derives from the symmetry imposed by the Pauli exclusion principle, according to which the wave function of a Fermi fluid changes sign under the operation of exchange of two identical fermions. That is, the many-body wave function of the electron gas has a node at the mid-point between each pair of parallel-spin electrons. The property $g_{\uparrow\uparrow}(0) = 0$ is an exact result and implies that each electron of given spin induces a local depletion of the density of electrons with the same spin.

If we bear in mind that the Coulomb interactions between electrons are repulsive, it is seen that the presence of the "Pauli hole" that was introduced above is accompanied by a gain in potential energy of the electron gas. Exchange effects keep apart electrons with parallel spins and therefore reduce on average the Coulomb repulsive energy of parallel-spin pairs. This consequence of the Pauli principle is for the electron gas the analog of the so-called Hund rule for the electronic structure of atoms, according to which there is on total energy grounds a preference for the state of maximum total spin. As seen in more detail below, this same effect favors the alignment of the spins of neighboring electrons in the interacting electron gas and stabilizes a spin-polarized state at strong coupling.

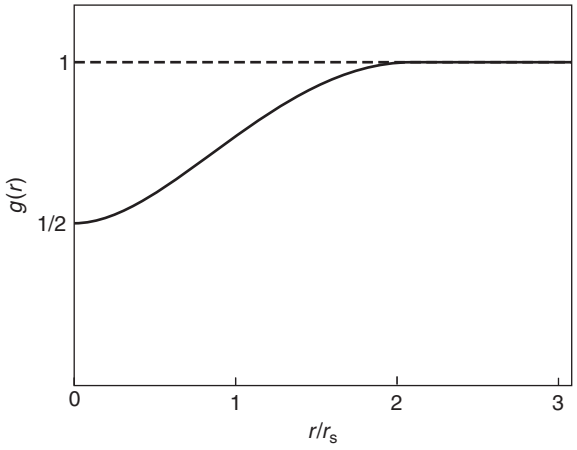


Figure 2 The Pauli hole surrounding an electron of given spin in the ideal Fermi gas.

In the ideal Fermi gas $g(r)$ takes the expression

$$g_{\mathrm{ideal}}(r) = 1 - \tfrac{9}{2}\,[j_1(k_{\mathrm{F}}r)/(k_{\mathrm{F}}r)]^2$$

where $j_1(x) = (\sin x - x\cos x)/x^2$ is a spherical Bessel function. The shape of this function, which describes the Pauli hole in the ideal Fermi gas, is shown in **Figure 2**. Of course, no correlations exist at this level between pairs of electrons with antiparallel spins, that is, $g_{\uparrow\downarrow}(r) = 1$.

The expression for $g_{\mathrm{ideal}}(r)$ can be used to evaluate the ground-state energy in first-order perturbation theory. This calculation corresponds to the energy of the electron gas in the Hartree–Fock approximation, since by translational symmetry the self-consistent Hartree–Fock single-particle orbitals must be the plane-wave orbitals of the ideal gas. The result is

$$\begin{aligned} E_{\mathrm{g}}^{\mathrm{HF}}(r_{\mathrm{s}}) &= \left(\frac{3}{5\alpha^2 r_{\mathrm{s}}^2} - \frac{3}{2\pi\alpha r_{\mathrm{s}}}\right)\mathrm{Ryd} \\ &\cong \left(\frac{2.21}{r_{\mathrm{s}}^2} - \frac{0.916}{r_{\mathrm{s}}}\right)\mathrm{Ryd} \end{aligned}$$

per electron, with $\alpha = (9\pi/4)^{-1/3}$ and in units of the Rydberg ($1\,\mathrm{Ryd} = e^2/(2a_{\mathrm{B}})$ with $a_{\mathrm{B}} = \hbar^2/(me^2)$ being the Bohr radius). Here, the density parameter r_{s} by the definition $4\pi(r_{\mathrm{s}}a_{\mathrm{B}})^3/3 = n^{-1}$ is introduced, that is, r_{s} represents the mean distance between neighboring electrons in units of the Bohr radius. The two terms in $E_{\mathrm{g}}^{\mathrm{HF}}(r_{\mathrm{s}})$ are the average kinetic energy and the average potential energy per electron, respectively, and their ratio is evidently proportional to $1/r_{\mathrm{s}}$. This shows that the role of the electron–electron

interactions becomes progressively more important as r_s increases – that is, as the density of the electron gas decreases.

Coulomb Correlations: Screening and Plasmons

It has been already remarked that no correlations between antiparallel spins are present in the ideal Fermi gas. Further depletion of the local electron density around each electron arises when the Coulomb repulsive interactions are switched on and allowed to modify the many-body wave function. The main effect is that the motions of opposite-spin electrons start to be correlated, since pairs of electrons with parallel spins are already kept apart by the Pauli principle.

Figure 3 shows the "Pauli–Coulomb hole" at increasing values of the coupling strength r_s. A property of this function is that the total local depletion of electron density around the electron at the origin corresponds to taking away exactly one electron from its neighborhood. This is a microscopic manifestation of Faraday's law, according to which an electric field cannot penetrate in the depth of a conductor. The bare electrical potential generated by each electron is on average completely screened by a local rearrangement of the other electrons, and the effective potential that an electron creates decays to zero over a microscopic distance. The notion of screening provides the microscopic basis for the Landau quasiparticles.

The screening effect has both static and dynamic manifestations. Considering first the screening of a static foreign charge, its bare potential is screened in Fourier transform by an electronic dielectric function $\varepsilon(q)$ which depends on wave number. The Thomas–Fermi theory yields $\varepsilon(q) = 1 + (q\lambda_{TF})^{-2}$ where $\lambda_{TF} = \sqrt{(\pi\hbar^2/4mk_Fe^2)}$ is the Thomas–Fermi screening length. In fact, the presence of the Fermi surface induces microscopic "Friedel oscillations" of wavelength π/k_F in the local electron density. The electron density $n(r)$ at a distance r from the foreign charge is asymptotically described by the expression

$$n(r) - n \propto \cos(2k_F r + \phi)/r^3$$

This result should be contrasted with the monotonic decay of the screened potential in a classical plasma, as found for instance in the Debye–Hückel theory of electrolytic solutions. Instead the effective potential created by an ion inserted in the electron gas, as well as the effective potential with which each electron interacts with the other electrons, are oscillating functions of distance showing both repulsive and attractive regions. These properties have crucial consequences for the electron theory of metals and for the electronic systems of interest in nanotechnology.

With regard to dynamical screening, an argument given by Langmuir in 1928 shows that the electron plasma can be driven into harmonic oscillation at the

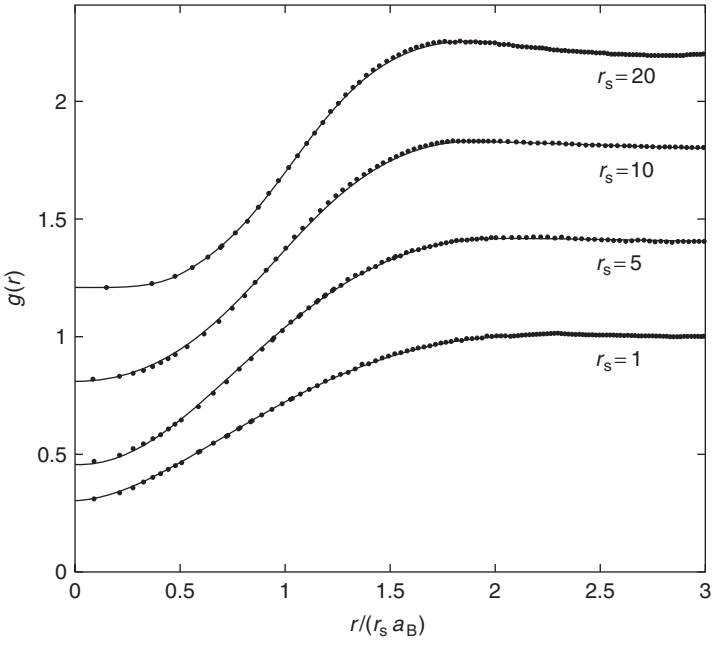


Figure 3 The Pauli–Coulomb hole in the interacting electron gas at various values of the coupling strength r_s: a comparison between theory (full curves) and quantum Monte Carlo data (dots). The values of $g(r)$ at $r_s = 5$, 10, and 20 have been rigidly shifted upwards for clarity by 0.4, 0.8, and 1.2.

"plasma frequency"

$$\omega_p = \sqrt{4\pi ne^2/m}$$

Such "plasmon excitations" are observed in inelastic scattering experiments on metals, using either a beam of fast electrons (in an EEL (electron energy loss) experiment) or a beam of X-rays.

Hubbard Model and the Para–Ferro Quantum Phase Transition

A simple way to illustrate the role of short-range correlations in driving the formation of order in the spins is provided by the model Hamiltonian introduced by Hubbard in 1963. The Hubbard model adds to the kinetic energy term a point-like repulsive interaction of magnitude U between two electrons with antiparallel spins whenever they meet on the same site. Considering here just the case of "itinerant" (nearly free) electrons, this interaction favors the alignment of the spins of neighboring electrons: the spin susceptibility in a constant magnetic field is enhanced by a factor $(1 - Uv_F/2)^{-1}$ relative to the Pauli value for the ideal Fermi gas (recall that v_F is the density of electron states on the Fermi surface).

Therefore, if U is increased in a continuous manner, it is found that the static spin susceptibility diverges when U attains the value $2/v_F$, and this corresponds to an instability of the paramagnetic state against the formation of a translationally invariant state of spin order. That is, the interacting electron gas stabilizes itself by a transition to a state having incipient ferromagnetic order. This "quantum phase transition" has been seen in quantum Monte Carlo studies carried out by Ceperley and Alder in 1980 and may correspond to the observation of magnetic (spin) polarization reported by Young *et al.* in 1999 for a dilute electron gas created in calcium hexaboride by doping with lanthanum. Of course, the electrons move in this system inside an ionic lattice, so that their states involve an additional quantum number which is absent in the electron gas model. Such "band effects" increase the effective mass of the electrons and hence the value of v_F, enhancing the natural tendency of the electron spins to align.

Wigner Crystallization

It has been already remarked that as the dimensionless electron spacing r_s increases (i.e., as the electron density decreases) the kinetic energy is becoming less and less important relative to the potential energy associated with the Coulomb repulsions. In the 1930s Wigner had noticed that an optimal value is obtained for the potential energy of jellium if the electrons are placed on the sites of a body-centered cubic lattice structure. Localization of the electrons raises their kinetic energy and an electron crystal may become favored only at very low density, where the potential energy becomes dominant. Of course, in the crystal the electrons are not strictly localized on the lattice sites but execute vibrational motions around them. Strong anharmonicity and high concentrations of lattice defects will be present in this quantum crystal near melting.

The problem of "Wigner crystallization" has become a classic in many-body physics, from both the theoretical and the experimental point of view. Quantum Monte Carlo studies have indicated that the ground state of three-dimensional (3D) jellium, after undergoing a continuous transition from a paramagnetic fluid to a fully spin-ordered ferromagnetic fluid, undergoes a first-order transition to a ferromagnetic crystal at $r_s \approx 65$. Similar studies of 2D jellium indicate a first-order transition from a paramagnetic to a ferromagnetic fluid followed by crystallization into a triangular lattice at $r_s \approx 35$. The search for Wigner crystallization in the laboratory has been addressed to quasi-2D assemblies of electronic carriers, mostly in man-made semiconductor structures.

The first unambiguous observation of Wigner crystallization was reported by Grimes and Adams in 1979 for electrons floating on top of a substrate of liquid ^{4}He in a quasi-classical regime. Metal–insulator transitions have subsequently been reported in quasi-2D systems of carriers subject to very strong magnetic fields in the fractional quantum Hall effect regime or in the presence of disorder and also in high-purity quasi-2D samples having exceptionally high carrier mobility.

Electron Gas in Reduced Dimensionalities

As it has been already seen in the introduction to this article, electron gases in reduced dimensionalities present strong fundamental interest and high technological relevance. A general consequence of reducing the dimensionality is an increasing role of the electron–electron correlations at each given value of the adimensional coupling strength r_s. This is ultimately a consequence of the modified topology of the electronic motions as the electrons are constrained to move in a plane or along a line or inside a quantum dot. This profoundly modifies the excitation spectrum of the ideal Fermi gas and enhances its susceptibility to any perturbation.

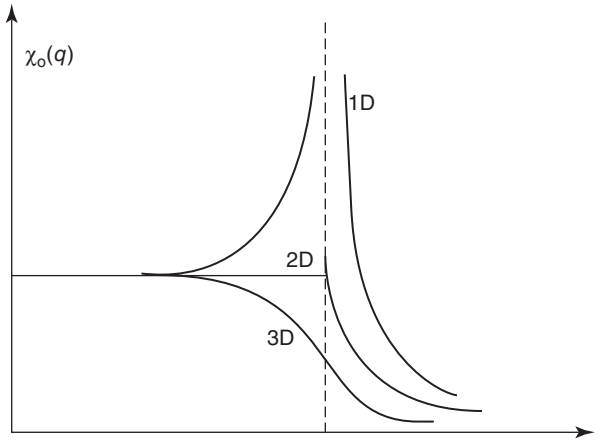


Figure 4 Static susceptibility of the Fermi gas in dimensionalities 1D, 2D, and 3D, as a function of wave number q; the vertical dashes show the position $q = 2k_F$.

Consider in particular a system in which the conduction electrons are restricted to move along a line. The Fermi surface reduces to two planes located at $k = \pm k_F$ and only forward or backward scattering processes are allowed. The susceptibility of the 1D Fermi gas to a static electrical potential varying along the line with a wave number q becomes

$$\chi_o(q) = -(2m/\pi\hbar^2 q)\ln\left|\frac{q+2k_F}{q-2k_F}\right|$$

This function presents a logarithmic divergence at $q = \pm 2k_F$, which can drive the "Peierls distortion" when the Coulomb repulsions are switched on: for example, a monoatomic 1D conductor with a half-filled conduction band (one electron per atom) may gain electronic energy by a lattice distortion opening a gap at the Fermi level through the formation of dimers.

The sketch shown in **Figure 4** illustrates the dependence of $\chi_o(q)$ on dimensionality. In all cases there is a singularity at $q = \pm 2k_F$, whose character depends on dimensionality: from a logarithmic divergence in the derivative of $\chi_o(q)$ in 3D, to a discontinuity in the derivative of $\chi_o(q)$ in 2D, and to a logarithmic divergence in $\chi_o(q)$ in 1D as seen above.

In conclusion, for conduction electrons in reduced dimensionalities (D = 1 and also in some 2D tight-binding-like models) a "Luttinger liquid" state may be induced by the electron–electron interactions. A Luttinger liquid differs from a normal Fermi liquid in two respects, that is (a) the discontinuity in the momentum distribution across the Fermi surface is replaced by another type of nonanalytic behavior, and (b) spin–charge separation occurs.

See also: Density-Functional Theory; Fermi Surface Measurements; Ferromagnetism; Hartree and Hartree–Fock Methods in Electronic Structure; Liquids, Theory of: Fermi Liquids; Metallic Bonding and Crystals; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Paramagnetism; Plasmons; Quantum Hall Effect; Scattering, Inelastic: Electron; Scattering, Inelastic: X-Ray (Methods and Applications); Semiconductor Nanostructures.

PACS: 71.10. – w

Further Reading

Abrikosov AA, Gorkov LP, and Dyzaloshinski IE (1975) *Methods of Quantum Field Theory in Statistical Physics*. New York: Dover.

Anderson PW (1997) When the electron falls apart. *Physics Today* 50: 42.

March NH (1999) *Electron Correlation in the Solid State*. London: Imperial College Press.

March NH and Tosi MP (1984) *Coulomb Liquids*. London: Academic Press.

March NH and Tosi MP (1995) Collective effects in condensed conducting phases including low-dimensional systems. *Advances in Physics* 44: 299.

Sachdev S (1999) *Quantum Phase Transitions*. Cambridge: Cambridge University Press.

Schofield AJ (1999) Non-fermi liquids. *Contemporary Physics* 40: 95.

Nomenclature

a_B	Bohr radius
$g(r)$	pair distribution function
h	Planck constant
$\hbar$	Planck constant h-crossed (h divided by 2π)
k_B	Boltzmann constant
k_F	Fermi wave number
m	electron mass
n	electron density
p_F	Fermi momentum
q	wave number
r	distance
r_s	adimensional electron–electron coupling strength
T	temperature
U	Hubbard repulsion
$\chi_0(q)$	static susceptibility
ε_F	Fermi energy
λ	de Broglie wavelength
λ_{TF}	Thomas–Fermi screening length
ν_F	density of states at the Fermi surface
ω_P	plasma frequency

Electron Spin Resonance

A-L Barra, Grenoble High Magnetic Field Laboratory, Grenoble, France
A K Hassan, Bruker Biospin AG, Fallanden, Switzerland

Introduction

Since its discovery in 1945 by Zavoisky, the electron spin resonance (ESR) technique is now commonly used in fields ranging from biology to solid-state physics and materials science. ESR, also known as electron paramagnetic resonance (EPR), is a spectroscopic method used to study the systems possessing a permanent magnetic moment of electronic origin. Most of these systems belong to the class of paramagnets, but the same experimental setup can be used for the study of ferro- and antiferromagnets. The states of magnetic materials and their corresponding properties are discussed elsewhere in this encyclopedia. The special cases of ferro- and antiferromagnets, corresponding to ferromagnetic resonance (FMR) and antiferromagnetic resonance (AFMR), respectively, is, however, not treated here.

The paramagnetic systems which have been studied extensively include free radicals, paramagnetic metal complexes (especially transition metals or rare-earth compounds), semiconductors (impurities, defects, etc.), and excited triplet states of molecules. Another important branch of ESR spectroscopy relies on the use of spin labeling, where free radicals are used to study diamagnetic hosts. As their magnetic moments are influenced by their local environment, the analysis of their ESR spectrum can give structural information such as the local symmetry, the nature of the environment, and the electronic distribution.

Usually, the ESR spectrum is obtained by irradiating the system under study at a fixed frequency and varying the applied external magnetic field until the resonance condition is fulfilled. The position of the resonance, its intensity, its lineshape, and line width are the characteristic features of the spectrum. An exact treatment of the ESR phenomenon requires the use of quantum mechanics, as the electronic magnetic moment is related to the electronic spin. However, a preliminary description can be obtained from a classical approach.

Phenomenological Approach

The magnetic moment $\boldsymbol{M}$ is related to the angular momentum $\boldsymbol{J}$ by the equation $\boldsymbol{M} = -\gamma \boldsymbol{J}$, where γ is the gyromagnetic ratio and $\boldsymbol{J}$ is in units of $\hbar$. When an external magnetic induction $\boldsymbol{B}_0$ is applied to the magnetic moment $\boldsymbol{M}$, it induces a torque $\boldsymbol{\Gamma} = \boldsymbol{M} \times \boldsymbol{B}_0$ so that the resulting equation of motion for the system is $\mathrm{d}\boldsymbol{J}/\mathrm{d}t = \boldsymbol{\Gamma}$ or $\mathrm{d}\boldsymbol{M}/\mathrm{d}t = -\gamma \boldsymbol{M} \times \boldsymbol{B}_0$. This describes the precession of the magnetic moment around the applied magnetic induction (**Figure 1**), with an angular frequency $\omega_0 = \gamma B_0$ for completely isolated magnetic moments. In order to take into account the interaction of the magnetic moments with the surroundings, a relaxation term has to be added which is assumed to be exponential. Let $\boldsymbol{M}_0$ be the equilibrium value of $\boldsymbol{M}$; it is parallel to $\boldsymbol{B}_0$ which is chosen to be along z. Two different relaxation times, the spin–lattice relaxation time T_1 (along the z-direction) and the spin–spin relaxation time T_2 (perpendicular to the z-direction), are introduced to describe the relaxation processes. T_1 corresponds to the energy transfer between the magnetic moments and the lattice via phonons, whereas T_2 describes the vanishing of the transverse components, a result of the growing incoherence in the perpendicular plane.

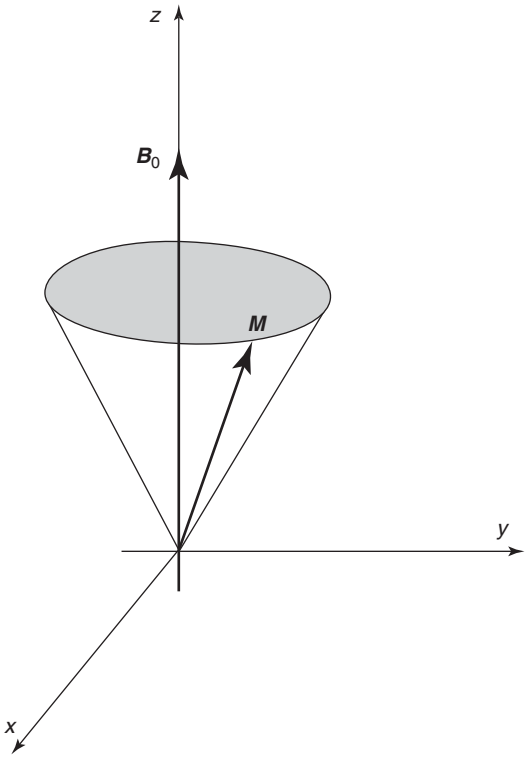


Figure 1 Magnetization precession about the applied magnetic induction $\boldsymbol{B}_0$.

Bloch Equations

When performing an ESR experiment, an oscillating magnetic induction $\boldsymbol{B}_1(t)$ is superimposed to the static

induction $\boldsymbol{B}_0$, with $\boldsymbol{B}_1$ perpendicular to $\boldsymbol{B}_0$ and $B_1 \ll B_0$. The oscillating induction usually has a linear polarization, which can be decomposed into the sum of two opposite circular polarizations. However, only one of these two circular components acts on the magnetic moment, so the other one can be neglected. Then the Bloch equations describing the evolution of a magnetic moment in the presence of static and oscillating external magnetic induction are obtained:

$$\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}t} = -\gamma \boldsymbol{M} \times \boldsymbol{B} - \begin{bmatrix} M_x/T_2 \\ M_y/T_2 \\ (M_z - M_0)/T_1 \end{bmatrix}$$

$$= -\gamma \boldsymbol{M} \times \begin{bmatrix} B_1 \cos(\omega t) \\ B_1 \sin(\omega t) \\ B_0 \end{bmatrix} - \begin{bmatrix} M_x/T_2 \\ M_y/T_2 \\ (M_z - M_0)/T_1 \end{bmatrix}$$

At this stage, it is easier to consider the evolution of the system in the rotating frame instead of the laboratory frame. The rotating frame is a Cartesian frame (x^r, y^r, z^r) whose z^r-axis coincides with the z-axis of the laboratory frame, whereas the x^r- and y^r-axis rotate around z at ω so that $\boldsymbol{B}_1$ is stationary in this frame. The x^r-axis is chosen along $\boldsymbol{B}_1$. The transformed Bloch equations describe the evolution of the magnetic moment $\boldsymbol{M}^r$ in the rotating frame:

$$\frac{\mathrm{d}\boldsymbol{M}^r}{\mathrm{d}t} = -\gamma \boldsymbol{M}^r \times \begin{bmatrix} B_1 \\ 0 \\ B_0 - \omega/\gamma \end{bmatrix} - \begin{bmatrix} M_x^r/T_2 \\ M_y^r/T_2 \\ (M_z - M_0)/T_1 \end{bmatrix}$$

The effective magnetic induction along the z-axis is reduced to $B_0 - \omega/\gamma = \Delta\omega/\gamma$, which vanishes for the resonance condition $\omega = \omega_0$. At resonance, the system sees only $\boldsymbol{B}_1$ in the rotating frame and precesses around this induction with an angular frequency $\omega_1 = \gamma B_1$.

Solution of the Bloch Equations

The steady-state solution of the Bloch equations, which describes correctly the system for time duration t that is large with respect to T_1 and T_2, is obtained easily in the rotating frame:

$$M_x^r = \frac{\omega_1 \Delta\omega T_2^2}{1 + (T_2\Delta\omega)^2 + \omega_1^2 T_1 T_2} M_0$$

$$M_y^r = \frac{-\omega_1 T_2}{1 + (T_2\Delta\omega)^2 + \omega_1^2 T_1 T_2} M_0$$

$$M_z^r = M_z = \left(1 - \frac{\omega_1^2 T_1 T_2}{1 + (T_2\Delta\omega)^2 + \omega_1^2 T_1 T_2}\right) M_0$$

where M_x^r is the response of the system in-phase with respect to $\boldsymbol{B}_1$, and M_y^r is the out-of-phase response.

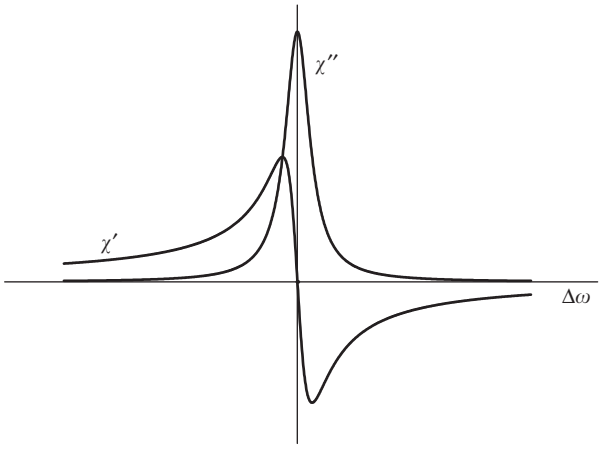


Figure 2 In-phase (χ') and out-of-phase (χ'') components of the dynamic magnetic susceptibility, as a function of the angular frequency deviation $\Delta\omega$.

These two components are associated to the dynamic susceptibility $\chi = \chi' - \mathrm{i}\chi''$ at the angular frequency ω. As $M = \chi H = \chi B/\mu_0$ (in the case where the relative permeability equals 1), and under nonsaturating conditions $(\omega_1^2 T_1 T_2 \ll 1)$,

$$\chi' = \frac{\Delta\omega\omega_0 T_2^2}{1 + (\Delta\omega T_2)^2}\,\chi_0$$

$$\chi'' = \frac{\omega_0 T_2}{1 + (\Delta\omega T_2)^2}\,\chi_0$$

where $\chi_0 = \mu_0 M_0/B_0$ is the static magnetic susceptibility. χ'' represents the power absorption and has a Lorentzian lineshape (**Figure 2**) with a half-width at half height equal to $1/T_2$, whereas χ' is the dispersion of the resonance.

Zeeman Interaction

Basis for a Free Electron

The classical picture given by the Bloch equations allows one to visualize the magnetic resonance phenomenon. However, as ESR deals with the electronic spin, it is necessary to perform a quantum mechanical treatment. The description with the Bloch equations then corresponds to the evolution of the medium value of the magnetic moment, that is, the evolution of the magnetization.

When an electron with spin 1/2 is placed in an external magnetic induction $\boldsymbol{B}_0$ applied along the z-direction, the degeneracy of the two energy levels corresponding to the two possible spin projections ($\pm 1/2$) is lifted (**Figure 3**). This is the Zeeman interaction described by the spin Hamiltonian: $H_S = -\boldsymbol{\mu} \cdot \boldsymbol{B}_0 = \gamma\hbar B_0 S_z = g_e \mu_B B_0 S_z$, where g_e the Landé factor for a free electron, is proportional to the variation of the

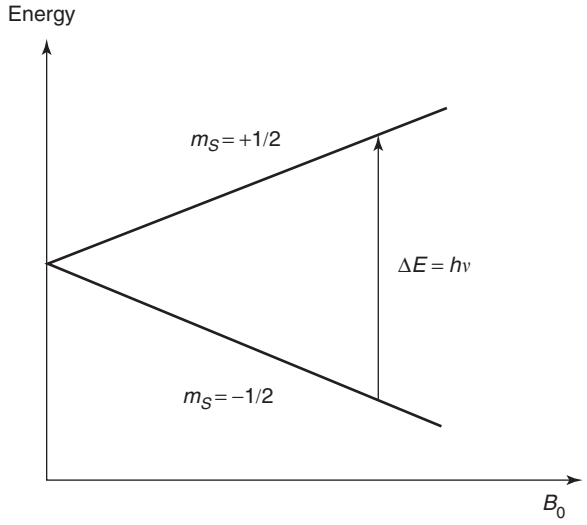


Figure 3 Zeeman effect on the energy levels for a spin $S = 1/2$.

energy levels with the applied induction; for a free electron, $g_e = 2.0023$. It is not equal to 2, as would be expected from a simple approach, due to relativistic corrections. $\mu_B = e\hbar/2m$ is the Bohr magneton, with e the absolute value of the charge, m the mass of the electron, and $\hbar$ the reduced Planck constant. S_z is the dimensionless spin projection operator along the z-direction. The eigenvectors of H_S are the eigenvectors of S_z, $|m_S = \pm 1/2\rangle$, with the corresponding eigenvalues: $E_\pm = \pm(1/2)g_e\mu_B B_0$. Thus, the resonance condition is $h\nu = E_+ - E_- = g_e\mu_B B_0$, where ν is the frequency of the exciting radiation.

Besides the resonance condition, there also exists an important selection rule resulting from the fact that the photon possesses only two spin polarizations, ± 1 (in units of $\hbar$). The conservation of the total angular momentum implies the selection rule for the paramagnetic system: $\Delta m_S = \pm 1$. Such a rule is valid when one uses a "standard" spectrometer, for which the exciting light propagates with its oscillating magnetic induction $\boldsymbol{B}_1$ perpendicular to the main magnetic induction $\boldsymbol{B}_0$.

General Case

Up to now, when looking for the resonance phenomenon only the Zeeman interaction of a free electron with the external induction has been considered. However, a spin also feels the internal magnetic field created by the system it is embedded in. As a result, its resonance will give direct information on this internal field and thus on the spin environment. A very important feature of ESR spectroscopy is to provide local properties of the system and not global ones, as does the magnetic susceptibility. This comes from the spatial dependence of the interactions creating the internal magnetic field, which decrease rapidly with increasing distances (e.g., the magnetic dipole–dipole interactions decrease like $1/r^3$). The contributions to the internal field can be divided into two classes: the ones induced by the external field and thus proportional to it, and the others which are independent of the applied field and correspond to permanent properties of the system.

The first class of interactions is taken into account within the formalism of the Landé factor, g. This means that instead of considering a total induction $\boldsymbol{B} = (1 + \alpha)\boldsymbol{B}_0$ acting on the system, the effect of the induced term is introduced via a Landé factor differing from the free electron value: $H_S = g_e\mu_B \boldsymbol{S} \cdot (\boldsymbol{B}_0 + \alpha\boldsymbol{B}_0) = g\mu_B \boldsymbol{S} \cdot \boldsymbol{B}_0$ with $g = (1 + \alpha)g_e$. In the general case of a spin 1/2 embedded in a low symmetry environment, g will transform to a second-rank matrix $\boldsymbol{g}$, not necessarily symmetric. In its principal axes (x, y, z), the $\boldsymbol{g}$-matrix reduces to

$$\begin{pmatrix} g_x & 0 & 0 \\ 0 & g_y & 0 \\ 0 & 0 & g_z \end{pmatrix}$$

and an external induction $\boldsymbol{B}_0$ can be expressed as

$$\begin{pmatrix} B_0 \sin\vartheta \cos\varphi \\ B_0 \sin\vartheta \sin\varphi \\ B_0 \cos\vartheta \end{pmatrix}$$

so that the Zeeman interaction leads to

$$\begin{aligned} E_\pm &= \pm\tfrac{1}{2} \\ &\quad \times \sqrt{g_x^2 \sin^2\vartheta \cos^2\varphi + g_y^2 \sin^2\vartheta \sin^2\varphi + g_z^2 \cos^2\vartheta} \\ &\quad \times \mu_B B_0 \\ &= \pm\tfrac{1}{2}\, g\mu_B B_0 \end{aligned}$$

From this equation, it can be seen that the g^2 value for each field orientation is found on an ellipsoid having g_x^2, g_y^2, and g_z^2 as principal values (**Figure 4**).

Three different cases should be distinguished depending on their symmetry: cubic symmetry, axial, anisotropy and rhombic symmetry. It is the local symmetry at the unpaired electron position that has to be taken into account and not the crystal symmetry. In the case of cubic symmetry, the three principal values of the $\boldsymbol{g}$-matrix are equal to each other: $g_x = g_y = g_z$; the ellipsoid reduces to a sphere. Thus, the resonance position will be the same for any orientation of the applied external induction. The powder spectrum of such a system, in which all the possible orientations of the external induction are present with the same

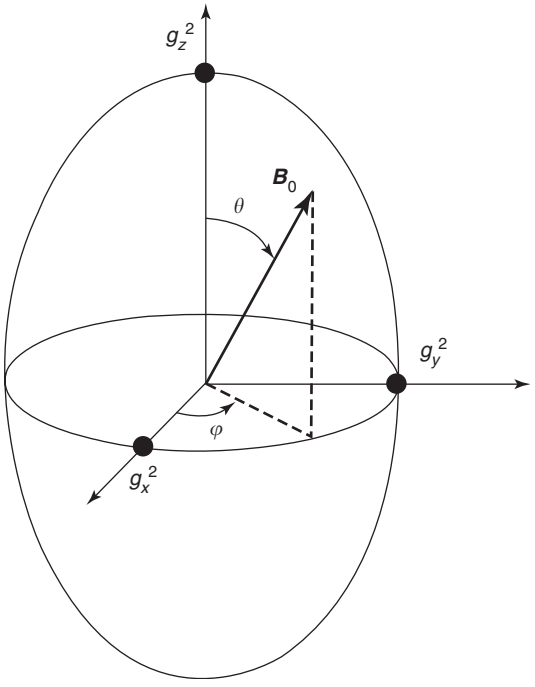


Figure 4 g^2 ellipsoid and definition of the angles θ and φ.

probability, consists of a single symmetric line, the same that one would obtain for a single crystal.

In the case of axial symmetry, the principal values are reduced to two independent values, $g_{||}$ and $g_{\perp}$, with $g_{\perp}$ doubly degenerate. The ellipsoid associated to such a system has a cylindrical symmetry. Such a case is encountered when the axis of symmetry has a symmetry higher than twofold. Then, for all the magnetic induction directions perpendicular to the symmetry axis, the resonance position will be the same; however, it changes when the angle between the external field and the axis of symmetry varies. When studying a powder sample, the resonance spectrum will have the shape shown in **Figure 5a** if $g_{||} < g_{\perp}$, and in **Figure 5b** if $g_{||} > g_{\perp}$.

In the case of rhombic symmetry, there are three different principal values and this is reflected on the ellipsoid corresponding to it (**Figure 4**). This is observed not only when the system has no symmetry elements at all but also if there are only twofold axes or inversion centers or mirrors as local symmetry elements. Then, any rotation of the external field will result in a variation of the resonance position. The shape of the spectrum obtained from a powder sample is shown in **Figure 6**.

Field-Independent Interactions

Besides the ***g***-matrix, the other important properties, which can be obtained from ESR spectroscopy, of a paramagnetic system are field independent. Here, the two most important properties, resulting from the interactions giving rise to the hyperfine and fine structures, are addressed.

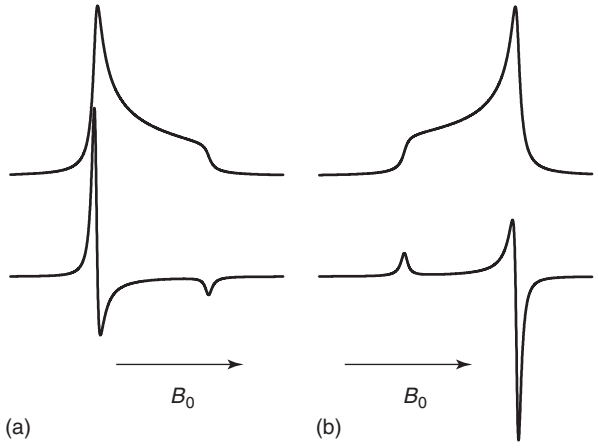


Figure 5 Powder spectrum (up: absorption and down: derivative) for an axially symmetric $S = 1/2$ system. (*a*) $g_{||} < g_{\perp}$; (*b*) $g_{||} > g_{\perp}$.

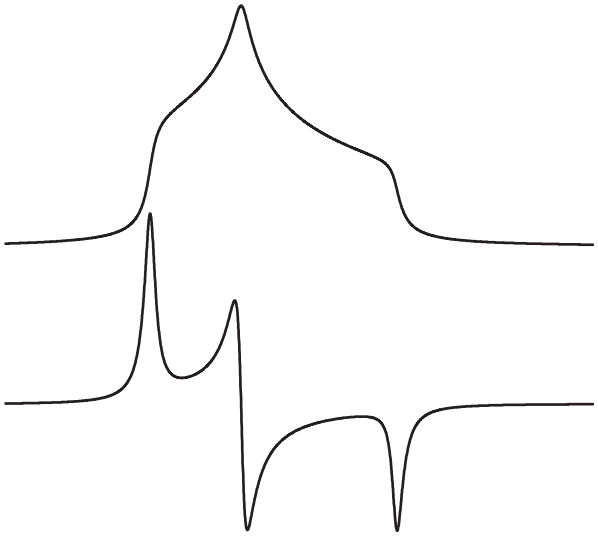


Figure 6 Powder spectrum (up: absorption and down: derivative) for a totally rhombic $S = 1/2$ system.

Hyperfine Interaction

The hyperfine structure results from the interaction of an electronic spin $\boldsymbol{S}$ with the neighboring nuclei possessing a nonzero nuclear spin $\boldsymbol{I}$. Its analysis yields very important information on the system under study, such as the determination of the chemical elements surrounding the electronic spin or the spin density of the electronic wave function. The hyperfine interaction is described by the Hamiltonian: $H_{\mathrm{hf}} = \boldsymbol{S} \cdot \bar{a} \cdot \boldsymbol{I}$, where $\bar{a}$ is the hyperfine tensor. The

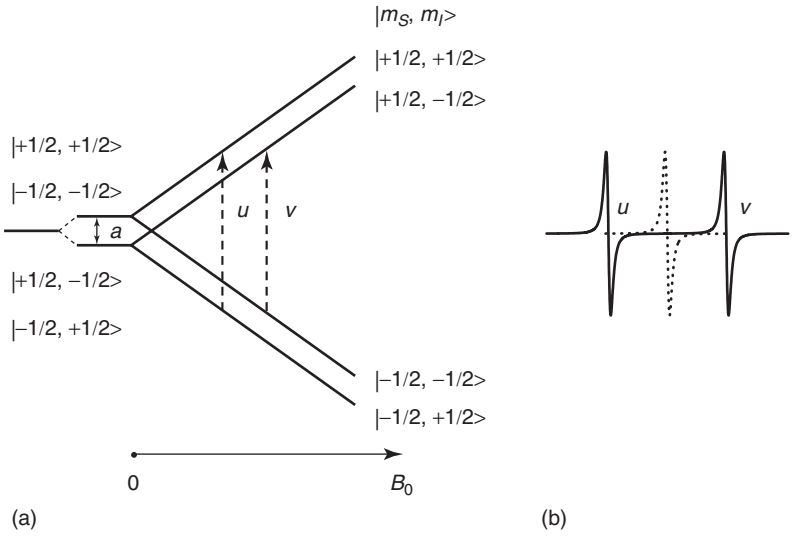


Figure 7 (a) Energy levels of a system with $S = I = 1/2$, showing first the split induced by an isotropic hyperfine interaction a. In zero field, there are two levels doubly degenerated split by the hyperfine interaction. Then the degeneracy removal from the application of the magnetic induction is shown, together with the two expected ESR transitions (labeled u and v). (b) ESR spectrum for an $S = I = 1/2$ system. The dotted line represents the resonance position for $a = 0$.

hyperfine interaction splits each of the electronic Zeeman levels into $2I+1$ levels, depending on the relative orientation of both spins, as illustrated in **Figure 7a** for $S = I = 1/2$. The existence of a hyperfine interaction adds a new selection rule for observing an ESR transition: $\Delta m_I = 0$.

There are two contributions to the hyperfine tensor: the first comes from the magnetic moments associated to the electronic and nuclear angular momenta ($\boldsymbol{S}$ and $\boldsymbol{I}$), which interact with each other via the well-known magnetic dipolar interaction. For a distance vector $\boldsymbol{r}$ between them, and neglecting the g-anisotropies, the effective Hamiltonian is

$$H_{\mathrm{dip}}(\boldsymbol{r}) = -\frac{\mu_0}{4\pi} g\mu_{\mathrm{B}} g_{\mathrm{n}} \beta_{\mathrm{n}} \left[\frac{\boldsymbol{S}\cdot\boldsymbol{I}}{r^3} - \frac{3(\boldsymbol{S}\cdot\boldsymbol{r})(\boldsymbol{I}\cdot\boldsymbol{r})}{r^5}\right]$$

where g_n and β_n are the nuclear g-factor and magneton, respectively, and μ_0 the permeability of vacuum. The dipolar contribution to the hyperfine tensor is then obtained by averaging the electron probability distribution resulting from the electronic wave function. The second contribution comes from the contact or Fermi interaction, when the dipolar treatment breaks down. It is directly proportional to the electronic spin density at the nucleus, and thus concerns electrons with wave functions containing s-type orbitals, which have a nonzero probability density at the nucleus. The symmetry of these orbitals implies that the contact interaction is isotropic. On the other hand, the dipolar interaction, being a traceless interaction, gives no isotropic term and is responsible for the anisotropic part of the hyperfine tensor.

Here, the hyperfine interaction is illustrated in the simplest case of a spin $S = 1/2$ interacting with a nuclear spin $I = 1/2$, but this treatment can be generalized to spins (S as well as I) of higher multiplicity, and considering isotropic values both for the Zeeman and the hyperfine interactions. Choosing the z-axis along the applied magnetic induction $\boldsymbol{B}_0$, and in the limit where the resulting Zeeman effect is much larger than the hyperfine interaction, the Hamiltonian describing the system reduces to $H = g\mu_{\mathrm{B}} B_0 S_z + a S_z I_z$. For the sake of simplicity, the nuclear Zeeman interaction is neglected. The eigenvectors are then $|m_S = \pm 1/2; m_I = \pm 1/2\rangle$ with the corresponding eigenvalues $E(m_S, m_I) = g\mu_{\mathrm{B}} B_0 m_S + a m_S m_I$. There exist four different energy levels (**Figure 7a**) between which two ESR transitions can be observed (**Figure 7b**) separated by $|a|/(g\mu_{\mathrm{B}})$. The center of these lines corresponds to the resonance position of the system in the absence of hyperfine interaction. A nice illustration of the simple case ($S = I = 1/2$) is found with the example (classical system) of n-type silicon doped with phosphorus, Si : P. When the doping level is low enough ($n < 10^{18}\,\mathrm{cm}^{-3}$), the phosphorus atoms can be treated as isolated. Two resonance lines are observed at low temperatures with a splitting of 42 G (**Figure 8**), as a result of the hyperfine coupling of the donor electron spin and the phosphorus nuclear spin (^{31}P with 100% natural abundance).

It should be pointed out that an ESR study cannot determine the sign of the hyperfine parameter. It can be inferred from theoretical considerations or, from the experimental point of view, by using a double-resonance method (such as electron nuclear double

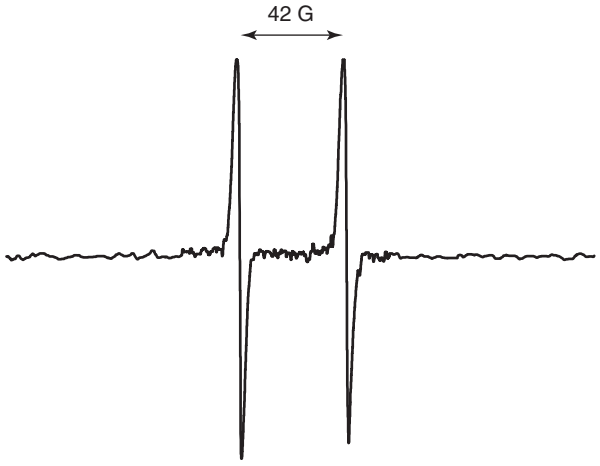


Figure 8 Low-temperature ESR spectrum of an Si : P sample (donor concentration 10^{16}) showing the double line expected from the hyperfine interaction between the ^{31}P nuclear spin $I = 1/2$ and the donor electron spin $S = 1/2$, with a hyperfine constant $a = 42$ G.

resonance (ENDOR)). ENDOR is indeed a very powerful method when dealing with complex systems with several nuclear spins interacting with the unpaired electrons, also because it allows one to take advantage of the different gyromagnetic factors of nuclei.

Fine Structure

When the system under study possesses more than one unpaired electron, the interactions between these electrons have to be taken into account. Examples of such systems are triplet-state molecules (either in the fundamental or in the excited state), biradicals and mostly systems containing transition metal or rare-earth ions. Besides the Coulomb interaction, which contributes to the energy difference (described by the exchange interaction) between levels having a different spin S and which will be ignored here, the electron–electron interaction leads to new terms in the spin Hamiltonian. These are the fine structure terms which may appear when the electronic spin is larger than 1/2. Among these terms only the second order term, the lowest possible order, is treated here.

There exist two different mechanisms giving rise to the fine structure term. One results from the magnetic dipolar interactions between the individual electronic spins $\boldsymbol{S}_1$, and $\boldsymbol{S}_2$, and is described in a way similar to the magnetic dipolar contribution to the hyperfine tensor by

$$H_{\mathrm{dip}}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} g^2 \mu_{\mathrm{B}}^2 \left\{ \frac{\boldsymbol{S}_1 \cdot \boldsymbol{S}_2}{r^3} - \frac{3(\boldsymbol{S}_1 \cdot \boldsymbol{r})(\boldsymbol{S}_2 \cdot \boldsymbol{r})}{r^5} \right\}$$

and neglecting the g-factor anisotropy, for the sake of simplicity. Integrating over the electron coordinates, it is then possible to derive the magnetic dipolar contribution to the fine structure, which is described by a Hamiltonian of the form $H_{\mathrm{f}}^{\mathrm{dip}} = \boldsymbol{S} \cdot \bar{D}_{\mathrm{dip}} \cdot \boldsymbol{S}$.

The other contribution involves the spin–orbit coupling and is usually considered within a perturbation treatment using either the crystal field interaction model or the more sophisticated ligand field approach to take into account the effect of the surrounding atoms. Here, only a simple approach giving the basic ideas is underlined. The unperturbed Hamiltonian H_0 corresponds to the electronic kinetic energy and the Coulomb interaction between the charged particles (nuclei and electrons). The perturbation Hamiltonian involves the effect of the surrounding atoms (crystal field), the spin–orbit and the Zeeman interactions. Depending on their respective strength, the effect of the crystal field is treated before or after the spin–orbit interaction. Consider the simplest case where the spin–orbit and the Zeeman interactions are the smallest contributing terms. The corresponding perturbation Hamiltonian can then be written as $H_{\mathrm{p}} = H_{\mathrm{SO}} + H_{\mathrm{Z}} = \lambda \boldsymbol{L} \cdot \boldsymbol{S} + \mu_{\mathrm{B}}(\boldsymbol{L} + g_{\mathrm{e}}\boldsymbol{S}) \cdot \boldsymbol{B}_0$, with λ the spin–orbit coupling constant and g_{e} the free electron Landé factor. In the hypothesis of an orbitally nondegenerate ground state $|0, m_S\rangle$ (here $|0\rangle$ represents the spatial part of the eigenvector), there is no energy correction to the first order of perturbation and all the $|m_S\rangle$ levels are degenerate in the absence of the external field. In the second order of perturbation, the energy correction has the form

$$E^{(2)}_{0,m_S} = \sum_{n \neq 0} \frac{\langle 0, m_S | (\mu_{\mathrm{B}}\boldsymbol{B} + \lambda \boldsymbol{S}) \cdot \boldsymbol{L} + g_{\mathrm{e}}\mu_{\mathrm{B}} \boldsymbol{B} \cdot \boldsymbol{S} | n, m'_S \rangle \langle n, m'_S | (\mu_{\mathrm{B}}\boldsymbol{B} + \lambda \boldsymbol{S}) \cdot \boldsymbol{L} + g_{\mathrm{e}}\mu_{\mathrm{B}} \boldsymbol{B} \cdot \boldsymbol{S} | 0, m_S \rangle}{E_0 - E_n}$$

with E_0 and E_n the eigenvalues of the unperturbed Hamiltonian corresponding to the states $|0\rangle$ and $|n\rangle$, respectively. Since it has no orbital operator, the term $g_{\mathrm{e}}\mu_{\mathrm{B}}\boldsymbol{B} \cdot \boldsymbol{S}$ vanishes. The other term can be factorized into two parts, one involving the orbital variables and the other involving only the spin variables. By introducing a matrix

$$\bar{\Lambda} = \sum_{n \neq 0} \frac{\langle 0 | \boldsymbol{L} | n \rangle \langle n | \boldsymbol{L} | 0 \rangle}{E_0 - E_n}$$

the energy correction can be associated to the effective spin Hamiltonian, $H^{\text{eff}} = \mu_B \boldsymbol{B} \cdot \bar{g} \cdot \boldsymbol{S} + \lambda^2 \boldsymbol{S} \cdot \bar{\Lambda} \cdot \boldsymbol{S}$, with $\bar{g} = g_e \bar{1} + 2\lambda\bar{\Lambda}$ where $\bar{1}$ is the unit matrix. This Zeeman term explains the origin of the *g*-anisotropy and, more generally, how the *g*-values can differ from the free electron *g*-factor. The second term of the effective Hamiltonian gives the spin–orbit contribution to the fine structure, so that in this simple case, $\bar{D}_{SO} = \lambda^2 \bar{\Lambda}$.

While the dipolar contribution to $\bar{D}$ is traceless, there is no such constraint for the one involving the spin–orbit effect. However, this isotropic part only contributes to a global energy shift of the whole *S* multiplet and can simply be omitted here (it gives no spectroscopic contribution). Thus, the general expression for the fine structure interaction up to the second order is $H_f = H_f^{dip} + H_f^{SO} = \boldsymbol{S} \cdot \bar{D} \cdot \boldsymbol{S}$, with D_x, D_y, and D_z the principal values of the $\bar{D}$ tensor. $\bar{D}$ is called the zero-field splitting (ZFS) tensor, as it leads to an energy difference between the different m_S levels of the *S* multiplet in the absence of a magnetic field (**Figure 9**). As $\bar{D}$ is traceless, $D_x + D_y + D_z = 0$, the three principal values reduce to only two independent parameters, *D* and *E*, defined by $D = 3D_z/2$, $E = (D_x - D_y)/2$. D_z is chosen to be the largest absolute principal value ($|D_z| > |D_x|$ and $|D_z| > |D_y|$) which results in $-1/3 \leqslant E/D \leqslant 1/3$. Thus, the fine structure term can be written as $H_f = D(S_z^2 - S(S+1)/3) + E(S_x^2 - S_y^2)$. *D* characterizes the axial anisotropy, whereas *E* describes its rhombicity ($E = 0$ for a system with axial symmetry).

The above Hamiltonian applies for all systems with spin $S > 1/2$. However, with respect to ESR spectroscopy, it is worth noting the difference between integer and half-integer spin systems associated to the fine structure. For a half-integer spin *S*, time reversal invariance implies that the levels corresponding to the same absolute value of the spin projection are degenerate ($E(m_S) = E(-m_S)$) in the absence of an external magnetic field. They form the so-called Kramer's doublets. In zero field, the existence of a zero-field splitting will induce an energy gap between the levels $|\pm 1/2\rangle$, $|\pm 3/2\rangle$, …. As the energy level $|\pm 1/2\rangle$ is at least doubly degenerate and will split with the applied magnetic field, half-integer spin systems should, in principle, always give rise to a signal (at any ESR frequency). For an integer spin *S*, there is no such restriction on the splitting of the multiplet and all the $|m_S\rangle$ levels can differ in energy in zero magnetic field, depending on the local symmetry of the system. As a result, the observation of an ESR signal is no more guaranteed for integer spin systems but depends on the spectrometer (frequency and available magnetic field range). This has led to the denomination "ESR silent" spin systems (**Figure 9b**), in order to describe integer spin systems giving no signal essentially in X-band ESR (9 GHz). However, using frequencies high enough to overcome the splitting should allow one to obtain a spectrum.

Finally, another property of $S > 1/2$ spin systems worth mentioning is the possibility to observe the so-called forbidden transitions. At high field, when the

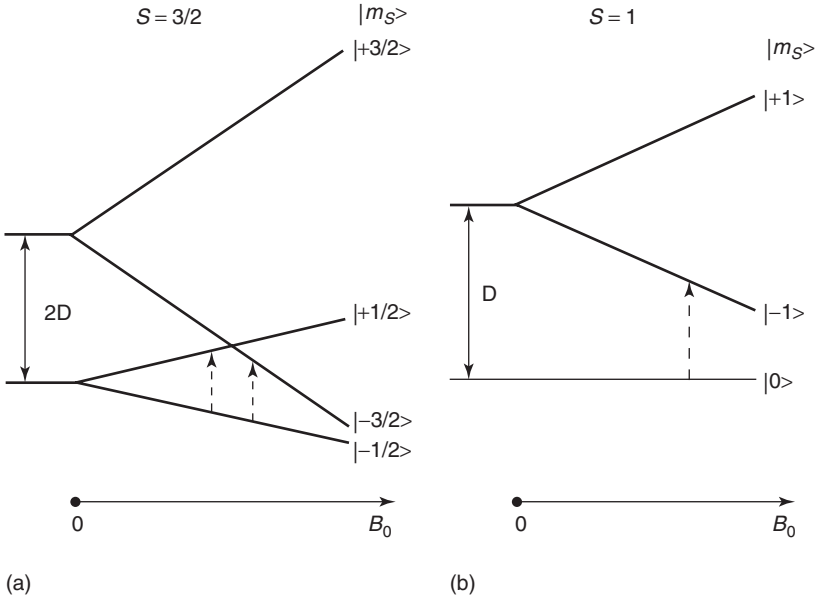


Figure 9 (a) Energy levels of an $S = 3/2$ system with axial symmetry, and their Zeeman splitting for a magnetic induction applied along the *z*-axis. (b) Idem for an $S = 1$ system. The dotted arrows depict the allowed ESR transitions.

Zeeman interaction dominates with respect to the field independent interactions, the field direction defines the quantization axis and the eigenvectors of the system are the corresponding $|m_S\rangle$ levels. Then the selection rule $\Delta m_S = \pm 1$ holds for a spectrometer operating in the perpendicular configuration. At low field, the eigenvectors are still labeled with the same $|m_S\rangle$; however, they are no more pure levels as the other interactions become comparable to, or larger than, the Zeeman term. In general, the quantization axis will not coincide with the field direction and it will be possible to observe "forbidden transitions," that is, transitions corresponding to $\Delta m_S \neq \pm 1$ with m_S defined by the field direction. For instance, for spin $S = 1$, an additional half-field signal is observed. Even if they are called "forbidden," these transitions are observed because they are actually allowed. They play an important role because, when dealing with powders, they often give rise to intense signals as they are relatively isotropic.

High-Frequency ESR

A recent evolution in ESR spectroscopy is the use of higher magnetic fields and frequencies. Usually, ESR is considered high field when performed in magnetic fields beyond the electromagnets range, that is, for fields higher than 2 T and for an exciting frequency higher than 60 GHz. The reasons for using higher fields and higher frequencies may vary. One reason is obviously the increased Zeeman interaction. Using higher fields can help in separating signals with slightly different g-values and resolving small g-anisotropies. Starting from the resonant condition $h\nu = g\mu_B B$, the separation between two lines associated with two different g-factors, g_1 and g_2, is given by

$$\Delta B = B_1 - B_2 = \frac{h\nu}{\mu_B}\left(\frac{1}{g_1} - \frac{1}{g_2}\right)$$

Thus, increasing the field, and consequently the frequency, increases the separation of the lines linearly, and leads to an increased resolution as long as the line width does not change.

Another main interest of using higher frequencies concerns spin systems with $S > 1/2$. High frequencies allow one to overcome large zero-field splittings and to observe transitions inaccessible with a lower-frequency excitation, especially in the case of integer spin systems (there are no more "EPR silent" systems). This characteristic has been used for the study of simple systems, such as mononuclear or binuclear species, for instance, single ion complexes with high spin Mn(III) ($S = 2$)or Ni(II) ($S = 1$). It has also proven very important for unraveling the striking properties of more complex compounds known as molecular spin clusters. These are molecules containing several spins, most often arising in metal ions, coupled together. Among the huge variety of these systems, a particularly attractive class is the one composed of single-molecule magnets (SMMs), in which the spin clusters behave like superparamagnets (showing, for instance, a hysteresis in the magnetization curve). SMMs are obtained for molecules with a large ground spin state associated to an Ising-type magnetic anisotropy, as observed for an octanuclear Fe(III) cluster (**Figure 10**, inset) which has an $S = 10$ ground state associated to a ZFS of $-0.2\,\mathrm{cm}^{-1}$. A single-crystal spectrum obtained on this system, at 230 GHz and 35 K with the field oriented along the principal x-axis of the SMM, illustrates the benefits of high-frequency EPR in the study of spin clusters, and more generally of systems with large ZFS (**Figure 10**). Besides the impurity signal in the center, the spectrum shows a regular pattern of 20 lines (the expected $2S$ transitions of the $S = 10$ multiplet). The distance between the lines allows one to determine the ZFS along the recording direction, as ZFS is proportional to this distance. The intensity pattern is not symmetric with respect to the center of the spectrum, because the intensity of the lines also depends on the thermal population of the spin levels.

Finally, another important property of high-frequency ESR is the possibility to reduce second-order effects, which make the spectra difficult to interpret, with the increase of the Zeeman term. More generally, when it is not possible to reach the high-field limit, operating at several frequencies allows one to distinguish the interactions independent of the magnetic field from the Zeeman term, thus obtaining a complete picture of the system under study.

Technical Aspects

Standard Spectrometer

Even if many different ESR spectrometers have been constructed, the most frequently encountered one is the cw X-band ESR spectrometer (see **Table 1** for a review of the most used bands), operating with the standard elements shown in **Figure 11**.

The 9 GHz microwaves, conveying the exciting magnetic field, are generated by a klystron. The sample is mounted in a resonant cavity placed at the center of the electromagnet generating the main magnetic field. A small coil is inserted in the main magnet, and produces a small (with respect to the line width) modulating field superimposed to the main field, resulting in the observation of a derivative signal. The microwaves reflected by the cavity are

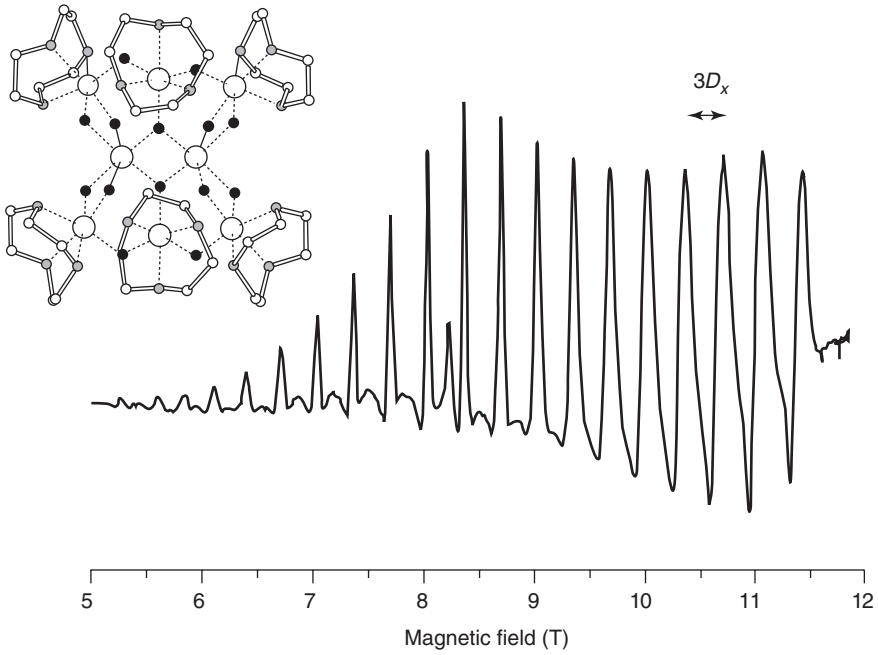


Figure 10 Inset: Schematic view of the molecular structure of the Fe_8 cluster, $[Fe_8O_2(OH)_{12}(tacn)_6]Br_8 \cdot 9H_2O$ where tacn = 1,4,7-triazacyclononane. The oxygen atoms are shown in black, nitrogen atoms in gray, carbon atoms in white (small circles) and the large white circles denote Fe(III) ions. Main: Single-crystal ESR spectrum of Fe_8 recorded at 35 K and 230 GHz along the principal axis *x*. The signals are regularly spaced, with a separation of $3D_x$, as expected when close to the high field limit. The intensity of each line results from the interplay between the transition probability and the thermal population of the involved levels. The small signal in the center of the spectrum (8.2 T) is due to an impurity.

Table 1 Main ESR frequencies and corresponding band denomination

Band	*Wavelength* (cm)	*Frequency* (GHz)	*Resonance* ($g = 2$)(T)
L	19	1.5	0.05
S	9.5	3	0.11
X	3	9.5	0.34
Q	0.8	35	1.3
W	0.3	95	3.4

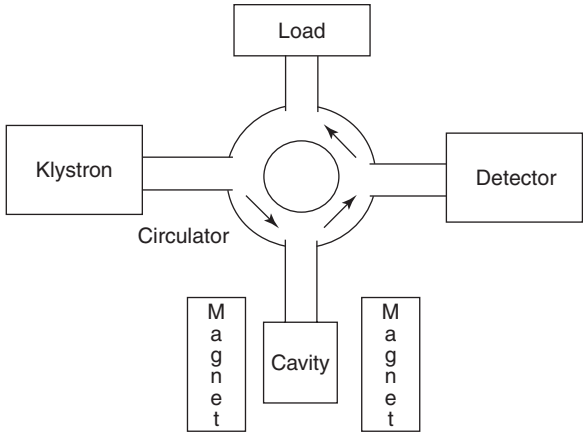


Figure 11 Block diagram showing the main components of an ESR spectrometer.

then directed to a diode detector, associated to a lock-in (detecting the first derivative signal). The parasitic microwaves reflected from the detector are directed to a load which absorbs them. A circulator is used to direct the microwaves from one part to the other.

Pulsed Techniques

Since the early 1970s, important technical advances have led to a renewal of ESR spectroscopy. One of the changes as already mentioned, deals with the development of spectrometers using higher magnetic fields and higher frequencies, whereas the other is the introduction of pulsed methods. Both of these techniques have developed enough to become, at least partially, commercially available. While NMR switched to pulsed operation and looked for higher fields very early, a similar development was not possible for ESR mostly due to the higher frequencies involved in ESR spectroscopy and, for pulsed ESR, the short relaxation times of paramagnetic species.

Basically, pulsed ESR relies on the use of short (typically ~ 1 ns) and powerful (typically $\sim 10^4$ W) microwave pulses to modify the spin distribution and relies on the detection of the time evolution of the perturbed system (free induction decay). Then, the Fourier transformation of the time response gives

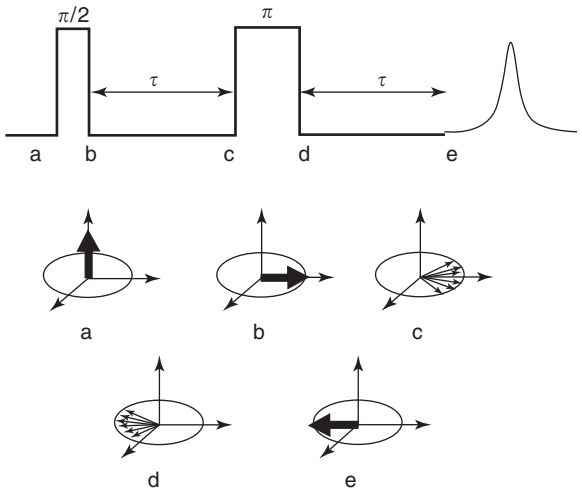


Figure 12 Echo experiment. Upper part showing a two pulse sequence and the echo signal. Lower part showing the corresponding magnetization evolution.

the frequency spectrum. Another important feature of pulsed operation is that the main magnetic field is no more swept; conversely, the microwave field becomes time dependent (pulsed). These experiments give access to time-dependent properties of the system under study, such as relaxation times. The basic experiment, already proposed in the 1960s by Mims, is a simple spin echo measurement (**Figure 12**) and has led to a specialized branch called electron spin echo spectroscopy. Another very important pulsed technique for ESR is known as the electron spin echo envelope modulation (ESEEM; also developed by Mims in 1965), which deals with the nuclear spins coupled to the paramagnetic center. Several pulse sequences have now been developed for ESEEM. More generally, in the same way NMR evolved, more and more pulse sequences are being implemented. Clearly, ESR is turning to Fourier transform ESR, even if it is still very far from the status of NMR where almost only FT-NMR exists.

See also: Crystal Field Splitting; Jahn–Teller Effect; NMR Methods for the Determination of Local Structure; Paramagnetism; Transition Metal Compounds, Electronic and Magnetic Properties of.

PACS: 76.30.−v; 75.10.Dg; 71.70.Ch; 71.70.Ej; 33.35.+r; 33.15.Kr; 33.15.Pw

Further Reading

Abragam A and Bleaney B (1986) *Electron Paramagnetic Resonance of Transition Metal Ions*. New York: Dover.

Bencini A and Gatteschi D (1990) *Electron Paramagnetic Resonance of Exchange Coupled Systems*. Berlin: Springer.

Griffith JS (1961) *The Theory of Transition-Metal Ions*. Cambridge: Cambridge University Press.

Mabbs FE and Collison D (1992) *Electron Paramagnetic Resonance of d-Transition Metal Compounds*. Amsterdam: Elsevier.

Pake GE and Estle TL (1973) *The Physical Principles of Electron Paramagnetic Resonance*. New York: Benjamin, WA.

Pilbrow JR (1990) *Transition Ion Electron Paramagnetic Resonance*. Oxford: Clarendon Press.

Poole CP Jr. (1967) *Electron Spin Resonance: A Comprehensive Treatise on Experimental Techniques*. New York: Wiley.

Schweiger A and Jeschke G (2001) *Principles of Pulse Electron Paramagnetic Resonance*. Oxford: Oxford University Press.

Smith GM and Riedi PC (2002) Progress in high field EPR. In: *Electron Paramagnetic Resonance*, vol. 18, chap. 9, pp. 254–303. The Royal Society of Chemistry

Weil JA, Bolton JR, and Wertz JE (1994) *Electron Paramagnetic Resonance*. New York: Wiley-Interscience.

Electronic Structure (Theory): Atoms

V M Ramaglia, Università degli Studi di Napoli "Federico II," Naples, Italy

Introduction

The analysis of the spectrum of the hydrogen atom has been the starting point of the development of quantum mechanics at the beginning of the twentieth century. The calculation of the energy eigenvalues and of the corresponding eigenfunctions of this one-electron system is simple enough both in nonrelativistic and in relativistic quantum mechanics. When the number of the electrons in the atom increases, an exact solution, such as for the hydrogen atom, is no more possible. A multielectron atom is a many-body system in which the electron–electron interactions together with the fermionic nature of the electrons as identical particles play a fundamental role. It is necessary to introduce some approximations. The majority of these approximations use a spherical self-consistent mean field $V_c(r)$ that includes, in an average way, all the interactions (mainly the electron–electron and electron–nucleus interactions) and in which the electrons move as independent particles. The electrons, being fermions of spin $\hbar/2$, obey the Pauli principle, and this circumstance,

as it will be shown later, allows one to justify the structure of the periodic table of the elements, that is the most important regularity property of the electronic structure of atoms.

The study of the helium atom, which has only two electrons, demonstrates the goodness of the mean-field approach because it provides a very good agreement with experiments from both a qualitative and a quantitative point of view. The choice of a suitable central mean potential depending on the atomic number Z provides a description of the periodic table of the elements. The most important ways to achieve $V_c(r)$ are the Thomas–Fermi, the Hartree, and the Hartree–Fock methods.

To improve this approximation, the nonspherical part of the electron–electron interaction has to be taken into account. The effect of this residual electrostatic interaction can be estimated as a perturbation on the Slater determinants containing the single-particle states of $V_c(r)$. The residual electrostatic interaction mixes these determinants, yielding states with a total angular momentum L and with a total spin S (the orbital angular momentum l and the spin s of the single electrons are no more good quantum numbers).

The spin–orbit coupling gives the fine structure of the levels (multiplets) that are also characterized by the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, so that the stationary states of the atom are eigenstates of $\boldsymbol{J}^2$, $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, J_z. The knowledge of the multiplets structure allows the interpretation of the complex spectra of many-electron atoms also in the presence of electrical and magnetical fields.

The Helium Atom

The Hamiltonian of the two electrons that includes only the electrostatic interactions after the separation of the center-of-mass motion, is

$$H = \sum_{i=1}^{2}\left(-\frac{\hbar^2\nabla_{r_i}^2}{2\mu} - \frac{Ze^2}{r_i}\right) + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} - \frac{\hbar^2}{M}\nabla_{r_1}\cdot\nabla_{r_2} \qquad [1]$$

The nucleus, in the origin, has mass M and charge $+Ze$, the electrons have the reduced mass $\mu = Mm/(m+M)$ (m is the electronic mass), positions $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and charge $-e$. In eqn [1], the terms $\hbar^2\nabla^2/2\mu$ are the kinetic energies of the relative electronic motion; the electron–electron interaction $e^2/|\boldsymbol{r}_1 - \boldsymbol{r}_2|$ screens the external Coulomb nuclear potential $-Ze^2/r$. The last term in eqn [1] represents the mass polarization term which, in the first approximation, will be neglected.

A mean spherical field $V_c(r)$ can be obtained by substituting Z with an effective charge Z_{eff}. This charge is estimated within a variational approach using, for the ground state, the trial wave function

$$\psi(r_1, r_2) = \frac{1}{\pi a^3}\,\mathrm{e}^{-(r_1+r_2)/a}\chi_{00}(1,2)$$

where the single-particle ground state of the potential $-Z_{\text{eff}}e^2/r$ appears ($a = \hbar^2/Z_{\text{eff}}me^2$). Here, $\chi_{00}(1, 2)$ is the singlet spin wave function of the total spin $\boldsymbol{S}$. The addition of two 1/2 spins gives a singlet state $S = 0$, $M_s = 0$ and a triplet state $S = 1$, $M_S = 0, \pm 1$

$$\begin{aligned}
&\chi_{00}(1,2) = \tfrac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \beta(1)\alpha(2))\\
&\chi_{10}(1,2) = \tfrac{1}{\sqrt{2}}(\alpha(1)\beta(2) + \beta(1)\alpha(2))\\
&\chi_{11}(1,2) = \alpha(1)\alpha(2), \qquad \chi_{1-1}(1,2) = \beta(1)\beta(2)
\end{aligned} \qquad [2]$$

where α is the spin-up state, β the spin-down state, while the numbers 1 and 2 distinguish the electrons. The fermionic wave function is totally antisymmetric, having a spatial part symmetric (exchanging the coordinates) and a spin part antisymmetric (exchanging the spins). The minimization of the energy provides the following effective charge:

$$Z_{\text{eff}} = Z - \tfrac{5}{16} \qquad [3]$$

In **Table 1**, the ground-state energies calculated with the variational method are compared with those estimated from experiments for three different two-electron systems (He, Li^+, and Be^{2+}). As it can be seen, the central-field approximation overestimates the electron–electron repulsion.

An improvement has been obtained by Hylleraas and Kinoshita using trial wave functions (depending on $r_1 + r_2$, $r_1 - r_2$, and $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$), that take into account the correlation between the electrons. The density probability of such wave functions is no more the product of two independent functions of r_1 and r_2. When the mass polarization, the relativistic, and the radiative effects are included, the agreement between theory and experiment is within the experimental error, as **Table 2** shows.

The independent-particle scheme applies to the excited states too. The lowest-lying excited states

Table 1 Values of the ground-state energy of two-electron He atom and ions in eV. Z is the atomic number, E_0 the energy without the electron–electron interaction (with only the Coulomb field of the nucleus), E_{var} the variational estimate (with $Z_{\text{eff}} = Z - 5/16$), and E_{exper} the experimental value

	Z	E_0	E_{var}	E_{exper}
He	2	−108	−76.6	−78.6
Li^+	3	−243.5	−195.6	−197.1
Be^{++}	4	−433	−368.1	−370.0

Table 2 I_p^∞ is the ionization potential (in cm^{-1}, $1\,cm^{-1} = 1.24 \times 10^{-4}$ eV) calculated including the correlation effects of the two-electron interaction without the relativistic corrections. The complete theoretical estimate I_0^{th} is obtained by adding the corrections due to the reduced mass, the mass polarization, the relativistic corrections, and the radiative quantum electrodynamics corrections. It coincides with the measured value I_0^{exp} within the experimental error

	H^-	*He*	Li^+
I_p^∞ non rel. ion. pot.	6 090.644 298	198 344.580 143 48	610 120.488 2
Reduced mass corr.	− 3.315 791	− 27.192 711	− 47.768 9
Mass pol. corr.	− 3.928	− 4.785	− 4.960
Relativistic corr.	− 0.304	− 0.562	19.69
Radiative corr.	− 0.003 7	− 1.341	− 7.83
I_0^{th}	6 083.092	198 310.699	610 079.62
I_0^{exp}	6 100 ± 100	198 310.82 ± 0.15	610 079 ± 25

correspond to the couple of single-particle states:

$$\psi_{100}(\boldsymbol{r}),\ \psi_{nlm}(\boldsymbol{r}) \qquad [4]$$

where n is the principal quantum number, $l(l+1)\hbar^2$ is the eigenvalue of the orbital angular momentum operator $\boldsymbol{l}^2$, $m\hbar$ the eigenvalue of the operator l_z. Now, a two-independent-particle state (totally antisymmetric) can have either the singlet or the triplet spin part

$$\begin{aligned}\Phi^+_{nlm} &= \tfrac{1}{\sqrt{2}}(\psi_{100}(\boldsymbol{r}_1)\psi_{nlm}(\boldsymbol{r}_2) \\ &\quad + \psi_{100}(\boldsymbol{r}_2)\psi_{nlm}(\boldsymbol{r}_1))\chi_{00} \\ \Phi^-_{nlmM_S} &= \tfrac{1}{\sqrt{2}}(\psi_{100}(\boldsymbol{r}_1)\psi_{nlm}(\boldsymbol{r}_2) \\ &\quad - \psi_{100}(\boldsymbol{r}_2)\psi_{nlm}(\boldsymbol{r}_1))\chi_1 M_S\end{aligned} \qquad [5]$$

Starting from the Coulomb nuclear potential $-Ze^2/r$, these states have the energy eigenvalues

$$E_n = -\frac{Z^2e^2}{2a_0}\left(1+\frac{1}{n^2}\right), \quad a_0 = \frac{\hbar^2}{me^2} \qquad [6]$$

on which the electron–electron interaction e^2/r_{12} has to be included. The degeneracy with respect to l and between triplet and singlet states is removed with the energy shifts

$$\begin{aligned}\varepsilon^\pm_{nl} &= \int \mathrm{d}^3r_1 \int \mathrm{d}^3r_2 |\psi_{100}(\boldsymbol{r}_1)|^2 |\psi_{nlm}(\boldsymbol{r}_2)|^2 \frac{e^2}{r_{12}} \\ &\quad \pm \int \mathrm{d}^3r_1 \int \mathrm{d}^3r_2 \psi^*_{100}(\boldsymbol{r}_1)\psi^*_{nlm}(\boldsymbol{r}_2) \\ &\quad \times \psi_{100}(\boldsymbol{r}_2)\psi_{nlm}(\boldsymbol{r}_1)\frac{e^2}{r_{12}} \\ &= I_{nl} \pm K_{nl}\end{aligned} \qquad [7]$$

where the signs + and − are relative to singlet and triplet states, respectively.

The first integral is called the Coulomb (or direct) integral. It represents the Coulomb interaction between the charge distributions of the two electrons. The integral K is known as the exchange integral and does not have a classical analog. Since the Coulomb integral is positive and being in the general case $K_{nl} > 0$, it turns out that a spin triplet state has a lower energy than the corresponding spin singlet state with the same n and l quantum numbers. Although the Hamiltonian does not contain spin-dependent terms, the purely electrostatic electron–electron interaction removes the spin degeneracy. The exchange term is much stronger than the spin-dependent terms arising from relativistic effects, such as the spin–orbit interaction. In fact, as Heisenberg first observed, the exchange term is strong enough to keep the electron spins aligned in certain solids, giving rise to the phenomenon of ferromagnetism.

Figure 1 shows the splitting of the unperturbed helium level for $n = 2$. In the electric dipole approximation, the transitions between the singlet states and the triplet states are forbidden and the most intense helium spectral lines belong to two independent families and apparently two different kinds of helium exist: the orthohelium (with $S = 1$) and parahelium (with $S = 0$).

When $n \to \infty$, then I_{nl} and K_{nl} go to zero and the first ionization limit of two-electron atoms (ions) is reached on the energy $E_\infty = -Z^2e^2/2a_0$ (see eqn [6]). The crude approximation of noninteracting electrons yields, for higher-energy excited bound states with n, $n' > 1$, the levels

$$E_{nn'} = -\frac{Z^2e^2}{2a_0}\left(\frac{1}{n'^2}+\frac{1}{n^2}\right) > E_\infty \qquad [8]$$

These levels are degenerate with the continuous energy spectrum of the ion starting from E_∞. The electron–electron interaction induces radiationless transition between the bound state and the scattering state. These doubly excited states are called autoionizing states because they can de-excite losing one of the electrons. The phenomenon of autoionization is known as the Auger effect. Not only helium-like

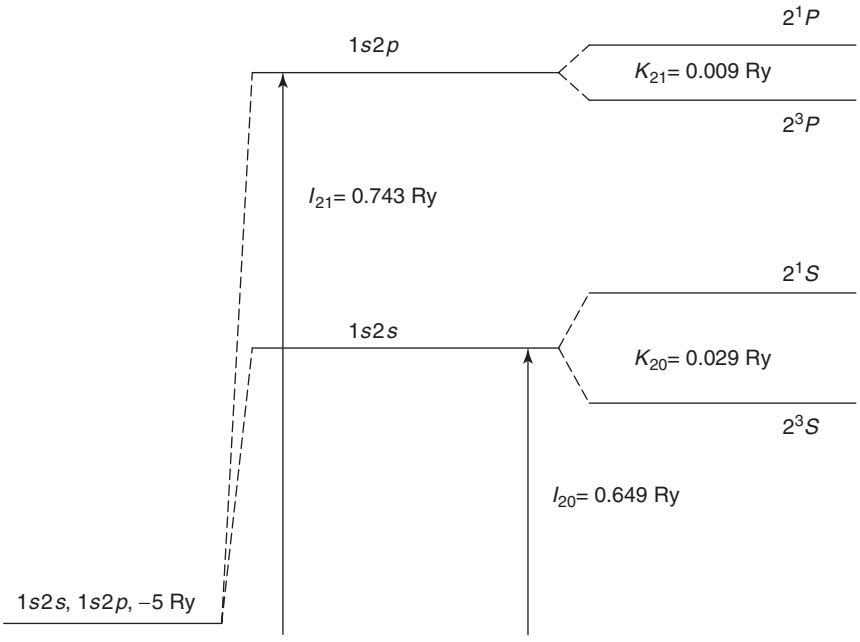


Figure 1 Splitting of the first unperturbed excited state $n=2$ of the He by the electron–electron interaction. The Coulomb and exchange integrals are given in Rydberg (1 Ry = 27.2116 eV).

(ions) but all atoms or ions with two or more electrons have such autoionizing states.

The Periodic Table of the Elements

The Hamiltonian of an atom with Z electrons without relativistic spin-dependent terms and external fields is

$$H = \sum_{i=1}^{Z}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \frac{1}{2}\sum_{i=1}^{Z}\sum_{j(\neq i)=1}^{Z}\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \quad [9]$$

The ground state and the first excited states can be described using a mean central field $V_c(r)$, in which the electrons move as independent particles. In this approximation

$$H \simeq H_c = \sum_{i=1}^{Z}\left(\frac{p_i^2}{2m} - V_c(r_i)\right) \quad [10]$$

and the overall effect of the electrons is to screen the Coulomb nuclear field more and more increasing the distance from the nucleus: near the nucleus $V_c(r)$ is of the form $-Ze^2/r$ and very far from it, $V_c(r) \simeq -e^2/r$. The screening can be represented by an effective charge depending on the distance r and the spherical mean field is no more Coulombian

$$V_c(r) = -Z_{\text{eff}}(r)e^2/r \quad [11]$$

The eigenstates of H_c are the Slater determinants $Z \times Z$ that can be formed with the single-particle states $|nlm_l m_s\rangle$ which are eigenstates of $p^2/2m + V_c(r)$, of the orbital momentum operators $\boldsymbol{l}^2$, l_z, and of spin z-component σ_z. The single-particle energy levels e_{nl} have a $2(2l+1)$ degeneracy $(-l<m_l<l,\ m_s = \pm 1/2)$. The central potential V_c changes with the atomic number Z but the order in the levels sequence e_{nl} is nearly the same for all the atoms and it is shown in **Table 3**.

The total energy is the sum of the individual energy levels and the Pauli principle allows the occupancy with only one electron of the state $|nlm_l m_s\rangle$. The distribution of the electrons with respect to the quantum numbers n and l is called the electron configuration. The states with the same n and l form a subshell, the subshells with the same n form a ;shell. The ground state of an atom has a configuration with filled (closed) subshells, each containing $n_i = 2(2l+1)$ electrons and an open subshell with $n<n_i$ electrons (called equivalent). A subshell has a degeneracy:

$$d_i = \frac{n_i!}{n!(n_i - n)!}$$

the closed subshells have $d_i = 1$.

As an example, the ground state of the carbon atom $Z=6$ has the configuration $1s^2 2s^2 2p^2$. The subshells $1s$, $2s$ are nondegenerate while the subshell $2p$ has a degeneracy:

$$\frac{6!}{2!4!} = 15$$

The total degeneracy of the ground state is $1 \times 1 \times 15 = 15$.

Table 3 The ordering of the single-particle energies of each subshell. The energy increases from bottom to top. The levels enclosed in the square brackets have very similar energies and their order can change by varying the atomic number Z

Quantum numbers n,l	*Spectroscopic notation for subshell (n,l)*	*Maximum number of electrons allowed in the subshell = 2(2l + 1)*
[6,2	[6*d*	10
5,3	5*f*	14
7,0]	7*s*]	2
6,1	6*p*	6
[5,2	[5*d*	10
4,3	4*f*	14
6,0]	6*s*]	2
5,1	5*p*	6
[4,2	[4*d*	10
5,0]	5*s*]	2
4,1	4*p*	6
[3,2	[3*d*	10
4,0]	4*s*]	2
3,1	3*p*	6
3,0	3*s*	2
2,1	2*p*	6
2,0	2*s*	2
1,0	1*s*	2

The last tightly bound electrons, which are in the subshell of the highest energy, and are in an insufficient number to form another closed subshell, are called valence electrons. They are responsible for the bondings between the atoms in the molecules.

All the information that is needed to discuss the electronic structure and the "building up" (aufbau) of atoms is on hand now.

The list of the elements begins with hydrogen that has an ionization potential of 13.6 eV. The next element is helium that has the largest ionization potential (24.59 eV). The configuration $1s^2$ corresponds to the closed shell $n = 1$. There are no valence electrons and helium is the first inert rare gas. The third element is lithium (the first alkali atom) with the ground-state configuration $1s^2 2s^1$. If the screening of the closed $n = 1$ shell were perfect, the ionization potential would be $13.6/4 = 3.4$ eV ($Z_{\text{eff}} = 1$); but the electron–electron potential reduces Z_{eff} and rises the ionization potential to 5.39 eV. Berillium has a configuration $1s^2 2s^2$. The ionization potenial is larger (9.32 eV) because of the increase of the nuclear charge. The subshell $2p$ is progressively filled starting from boron ($Z = 5$) up to neon ($Z = 10$), for which the shell $n = 2$ is full. Neon is the second inert rare gas for which the ionization potential is increased up to 21.56 eV. From $Z = 11$ (sodium, the second alkali atom) to $Z = 18$ (argon, the second rare gas), the subshells $3s$ and $3p$ are progressively filled. The filling of the shell $n = 3$ is interrupted at $Z = 19$ (potassium, third alkali element) because the level $4s$ is lower in energy than the $3d$ level that starts to be filled after the $4s$ from scandium ($Z = 21$) to zinc ($Z = 30$). There is a competition between the levels $4s$ and $3d$ that are very close in energy and the filling is not so regular as for the elements with lower atomic numbers. The elements with an incomplete $3d$ subshell belong to the first transition or iron group. The higher incomplete subshells of chromium ($Z = 24$) correspond to the state $[\text{Ar}]4s^1 3d^5$ ([Ar] stands for the noble gas argon configuration $1s^2 2s^2 3s^2 3p^6$) that is energetically more favorable to the $[\text{Ar}]4s^2 3d^4$ configuration. The high number of valence electrons allows the very rich class of molecules that can be formed with chromium. The filling competition between the subshells $5s$ and $4d$ characterizes the second transition or palladium group (from $Z = 39$ (yttrium) to $Z = 48$ (cadmium)), and the third transition or platinum group (from $Z = 71$ to $Z = 80$ (mercury)) corresponds to the irregular filling of the states $6s$ and $5d$. The list of elements ends with the actinides beginning from $Z = 89$ (actinium) in which a competition occurs between $5f$ and $6d$ states. The $Z = 103$ (lawrencium) is the element with the largest atomic number because the actinides are radioactive elements with highly unstable nuclei that decay by spontaneous fission into lighter elements.

The list of the elements can be organized in a table by arranging the elements with similar outer subshells in columns. All the alkali metals have a single weakly bound electron in a subshell s ($2s$ Li, $3s$ Na, $4s$ K, and so on). All the halogens (F, Cl, Br, I) have a subshell p^5 lacking only one electron. The alkalis and the halogens have high chemical reactivities because they can reach more stable energetic arrangements, losing or acquiring an electron in such a way as to form completely filled subshells. The last column of the table contains the rare gases (He, Ne, Ar, Kr, Xe, and Rn).

The recurrences of similar chemical properties led Mendeleev in 1869, long time before the existence of the electrons and the nuclei were known, to classify the elements in the periodic table. He placed elements with similar chemical properties in the same column. The chemical properties change progressively along the rows, which are called periods. There are seven periods and each of the periods begins with an alkali element and ends with a noble gas atom, except for the seventh period, which is incomplete.

The ionization potential increases along a period and has a sudden drop passing from a noble gas to the alkali of the next period (**Figure 2**).

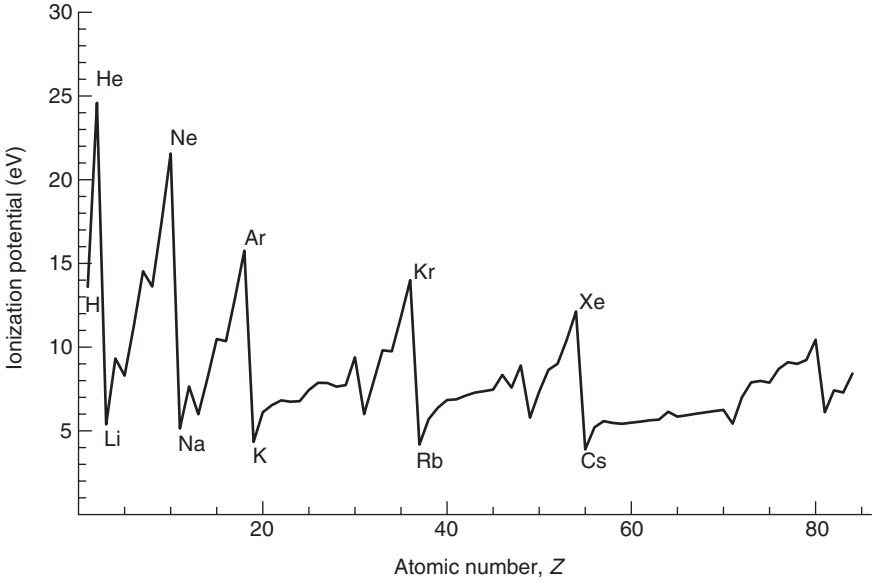


Figure 2 The first ionization potential against the atomic number.

Beyond the Central-Field Approximation

There are two main corrections to the approximate Hamiltonian H_c. The first one is the correlation due to the electron–electron interaction

$$\begin{aligned} H_1 &= H - H_c \\ &= -\sum_i \frac{(Z - Z_{\text{eff}}(r_i))e^2}{r_i} + \frac{1}{2}\sum_{i,j(j\neq i)} \frac{e^2}{r_{ij}} \\ r_{ij} &= |\boldsymbol{r}_i - \boldsymbol{r}_j| \end{aligned} \qquad [12]$$

that is called the residual electrostatic interaction and is no more spherically symmetric. The second correction is the spin–orbit term that, within the central field approximation, is

$$\begin{aligned} H_2 &= \sum_i \xi(r_i)\boldsymbol{l}_i \cdot \boldsymbol{s}_i \\ \xi(r_i) &= \frac{1}{2mc^2}\frac{1}{r_i}\frac{\mathrm{d}V_c(r_i)}{\mathrm{d}r_i} \end{aligned} \qquad [13]$$

If the spin–orbit coupling terms are neglected, the effect of the residual electrostatic interaction can be evaluated treating H_1 as a perturbation on H_c. The ground state and the first excited states of H_c are strongly degenerate and the levels of H can be obtained by diagonalizing H_1 in the subspace S_0 of the lowest eigenvalue E_0 of H_c. As an example, the subspace S_0 of the ground state of the carbon atom contains 15 Slater 6×6 determinants formed with the four states of the two closed subshell $1s^2$ and $2s^2$ and two states of the incomplete $2p$ subshell, namely,

$$\left|-1, \tfrac{1}{2}\right\rangle,\ \left|-1, -\tfrac{1}{2}\right\rangle,\ \left|0, \tfrac{1}{2}\right\rangle,$$
$$\left|0, -\tfrac{1}{2}\right\rangle,\ \left|1, \tfrac{1}{2}\right\rangle,\ \left|1, -\tfrac{1}{2}\right\rangle$$

where the first number is the eigenvalue of l_z, the second one that of s_z (in units $\hbar$). The subspace S_0 is called a multiplet. The multiplet is spanned by the eigenvalues of the angular orbital momenta $\boldsymbol{l}_i$ and of the spins $\boldsymbol{s}_i$ of the single-particle levels of the central potential $V_c(r)$. If the overall electron–electron interaction is taken into account, only the total angular momentum $\boldsymbol{L} = \sum \boldsymbol{l}_i$ commutes with H, while the single angular orbital momenta are not conserved (H is invariant only under overall rotations). The perturbation H_1 is diagonal in the multiplet basis $\{\boldsymbol{L}^2, L_z, \boldsymbol{S}^2, S_z\}$ whose vectors $|\alpha LSM_LM_S\rangle$ are eigenvectors of the total orbital momentum and the total spin, and the quantum number α labels the different vectors having the same total angular momentum and total spin. The use of this basis, that has the symmetries of the total Hamiltonian H, greatly simplifies the diagonalization of H_1

$$\begin{aligned} &\left\langle \alpha' L' S' M_L' M_S' | H_1 | \alpha L S M_L M_S \right\rangle \\ &\quad = \delta_{LL'}\delta_{SS'}\delta_{M_L M_L'}\delta_{M_S M_S'} V_{\alpha\alpha'}^{(LS)} \end{aligned} \qquad [14]$$

and one needs only to diagonalize the matrices $V_{\alpha\alpha'}^{(LS)}$ corresponding to each pair of quantum numbers LS. The complete electron–electron interaction does not remove the degeneracy with respect to M_L and M_S because H_1 has an overall spherical symmetry and it

does not have a preferential direction. The levels LS (generally denoted as ${}^{2S+1}C$, C being the capital letter S for $L=0$, P for $L=1$, D for $L=2,\ldots$) are left by the residual electrostatic interaction with a $(2L+1)(2S+1)$ degeneracy.

The vectors $|\gamma LSM_LM_S\rangle$ can be constructed by summing up the single orbital momenta $\boldsymbol{l}_i$, and the single spins $\boldsymbol{s}_i$. The Slater determinants spanning the subspace $\boldsymbol{S}_0$ have to be mixed with each other to form the eigenstates of the basis $\{\boldsymbol{L}^2, L_z, \boldsymbol{S}^2, S_z\}$. The antisymmetry of the wave function forbids some values of L and S compatible with the single-particle angular momenta addition. Returning to the example of the carbon atom ground state and setting aside the Pauli principle, the addition of the single-particle angular momenta of two $2p$ electrons gives the following spectral terms:

$$^3S, ^3P, ^3D, ^1S, ^1P, ^1D$$

that have an overall degeneracy of $3\times 1+3\times 3+3\times 5+1\times 1+1\times 3+1\times 5=36=6\times 6$, while only 15 states are available. It turns out that the terms of this list satisfying the Pauli principle are

$$^3P, ^1S, ^1D$$

as the triplet- and the singlet-spin states are, respectively, symmetrical and antisymmetrical in the exchange of the spins, and as the states S and D, and P are, respectively, symmetrical and antisymmetrical in the interchange of the orbital variables. The number of linearly independent antisymmetrical states is $9+1+5=15$, just the number of the Slater determinants in $\boldsymbol{S}_0$.

Although the perturbation H_1 does not depend on the spin S, the perturbed levels depend on S via the exchange integrals introduced by the two-body operator H_1, such as those that appear in the excited levels of the helium. The repulsive electron–electron interaction decreases with the mutual distance of the electrons, that is, when the orbital part of the wave function becomes more antisymmetrical. The spin part of the wave function is more symmetrical for larger S and, therefore, the orbital part is more antisymmetrical for larger S. The levels of a given configuration are in the order of decreasing total spin (Hund's rule). In the ground state of the carbon atom, the lowest level is the triplet one ^3P, followed first by the singlet ^1D and then by ^1S. These two levels have the same spin multiplicity, but ^1D has an higher orbital momentum and in this state the electrons are farther away and the repulsive effect of electron–electron interaction on it is less.

Subshells that contain the same number of occupied or empty states (holes in the filled subshell) are equivalent to one another. It can be seen that the two subshells with k and $4l+2-k$ equivalent electrons have the same LS terms, so the perturbative corrections due to H_1 are the same.

The addition of the spin–orbit coupling H_2 to the electrostatic correction H_1 gives a less symmetric perturbative correction to H_c: H_1+H_2 commutes with $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$, but neither with $\boldsymbol{L}$ nor with $\boldsymbol{S}$. The relative importance of H_2 increases rapidly with Z. In the light and medium atoms $|H_1|\gg|H_2|$; in heavier atoms (starting from Pb), $|H_1|$ and $|H_2|$ are of the same order of magnitude. In the first case, H_2 is treated as a small perturbation on LS levels. This scheme is called Russel–Saunders coupling or LS coupling. First, H_1 removes the degeneracy of the linear combinations of Slater determinants corresponding to the values of $\boldsymbol{L}=\sum \boldsymbol{l}_i$ and of $\boldsymbol{S}=\sum \boldsymbol{s}_i$ allowed by the Pauli principle. Then, H_2 gives a different energy eigenvalue to each possible value of J ($J=L+S, L+S-1,\ldots, |L-S|$) with a residual degeneracy of order $(2J+1)$.

The other scheme relatively simple is the jj coupling, where $|H_2|\gg|H_1|$. The Slater determinants are formed using single-particle states perturbed by a spin–orbit interaction on V_c: first, the orbital angular momentum $\boldsymbol{l}_i$ is coupled with the spin $\boldsymbol{s}_i$ as $\boldsymbol{j}_i=\boldsymbol{l}_i+\boldsymbol{s}_i$ by H_2, then H_1 removes the degeneracy between the allowed values of $\boldsymbol{J}=\sum \boldsymbol{j}_i$. The level structure of the ground state of the heavier atoms is intermediate between those given by LS and jj coupling.

It turns out that the matrix elements in the subspace of a level (αLS) with a degeneracy of the order $(2L+1)(2S+1)$ are the same as those of the operator $\gamma(\alpha, L, S)\ \boldsymbol{L}\cdot\boldsymbol{S}$, where $\gamma(\alpha, L, S)$ is a constant characteristic of the unperturbed level:

$$\begin{aligned}&\langle \alpha LSM_LM_S|H_2|\alpha LSM'_LM'_S\rangle \\ &\quad=\gamma\langle \alpha LSM_LM_S|\boldsymbol{L}\cdot\boldsymbol{S}|\alpha LSM'_LM'_S\rangle\end{aligned} \qquad [15]$$

This relationship between the matrix elements of $\sum_i \boldsymbol{l}_i\cdot\boldsymbol{s}_i$ and $L\cdot\boldsymbol{S}$ is proved in the quantum mechanics theory of angular momentum by the Wigner–Eckart theorem. The constant γ is proportional to the average value of $\xi(r)$ on the radial orbital part common to all the single-particle states spanning the subspace of the level. The spin–orbit correction on a closed subshell is null (for this level, $L=S=M_L=M_S=0$). While H_2 is not diagonal in the representation $\{\alpha LSM_LM_S\}$, as $L_xS_x+L_yS_y+L_zS_z$ does not commute with L_z and S_z, it becomes diagonal in the equivalent representation $\{\alpha LSJM_J\}$. Since

$$\boldsymbol{J}^2=\boldsymbol{L}^2+\boldsymbol{S}^2+2L\cdot\boldsymbol{S}=\tfrac{1}{2}(\boldsymbol{J}^2-\boldsymbol{L}^2+\boldsymbol{S}^2) \qquad [16]$$

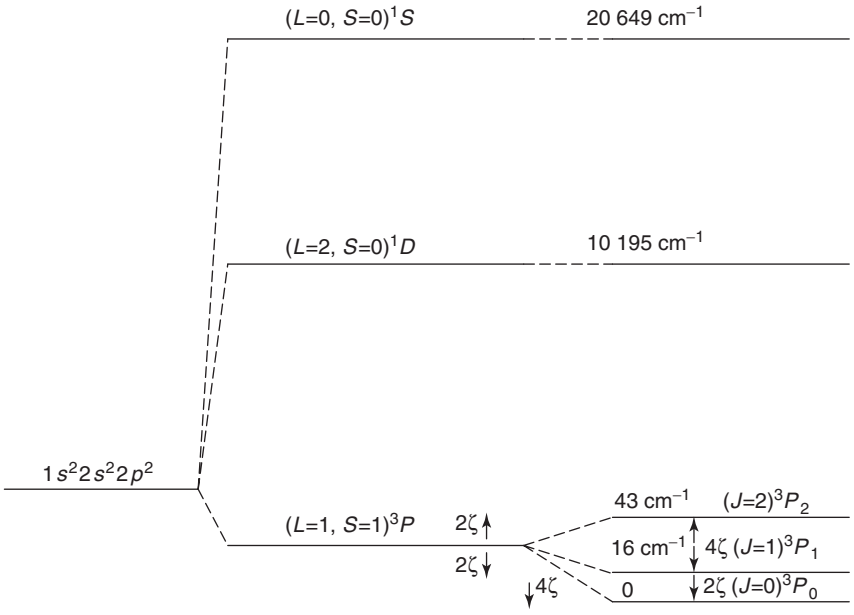


Figure 3 Splitting of the ground-state configuration of the carbon atom due to electrostatic residual interaction and to the spin–orbit coupling. Here $\zeta = \gamma(\alpha, L, S)\hbar^2/2$. The numbers near the levels (in cm^{-1}) are the experimental values of the energy levels. The Landè interval rule is approximately satisfied. The separation between $^3P_0$ and $^3P_1$ gives a value of $\zeta = 8\,cm^{-1}$ and the energy of $^3P_2$, above the ground state $^3P_0$, foreseen by the interval rule is of $6 \times 8 = 48\,cm^{-1}$, a little larger than the measured value of $43\,cm^{-1}$.

one has

$$\begin{aligned}&\left\langle \alpha LSJM_J | H_2 | \alpha LSJM'_J \right\rangle \\ &\quad = \tfrac{1}{2}\gamma\hbar^2(J(J+1) - L(L+1) - S(S+1)) \\ &\quad = E(J)\end{aligned} \qquad [17]$$

and the unperturbed level αLS splits into as many levels as the possible values of $J(= L+S, \ldots, |L-S|)$. The degeneracy of these levels is of the order $(2J+1)$ (**Figure 3**).

The energy separation between adjacent levels $E(J)$ and $E(J-1)$ is proportional to J:

$$\begin{aligned}E(J) - E(J-1) = \tfrac{1}{2}\gamma\hbar^2[&J(J+1) - L(L+1) \\ &- S(S+1) - J(J-1) \\ &+ L(L+1) + S(S+1)] \\ = \gamma\hbar^2 J\end{aligned} \qquad [18]$$

This result is known as the Landè interval rule and is well satisfied experimentally for Z large enough, when among the overall magnetic interactions between the angular momenta

$$a_{ij}\boldsymbol{l}_i \cdot \boldsymbol{s}_j,\ b_{ij}\boldsymbol{l}_i \cdot \boldsymbol{l}_j,\ c_{ij}\boldsymbol{s}_i \cdot \boldsymbol{s}_j$$

only the diagonal terms a_{ii} dominate all the others. For a single open subshell that is less than half-filled, the constant γ is positive and the lowest energy value has the smallest possible value of J. These multiplets are called regular. On the other hand, multiplets that are more than half-filled have $\gamma < 0$, the order with J is turned upside down and they are called inverted. Regular and inverted multiplets have opposite spin–orbit corrections

$$\begin{aligned}&\left| \sum_{k \in \text{filled states}} \boldsymbol{l}_k \cdot \boldsymbol{s}_k \right| \\ &\quad = \left| \left(\sum_{k \in \text{all shell}} - \sum_{k \in \text{empty states}} \right) \boldsymbol{l}_k \cdot \boldsymbol{s}_k \right| \\ &\quad = -\left| \sum_{k \in \text{empty states}} \boldsymbol{l}_k \cdot \boldsymbol{s}_k \right|\end{aligned} \qquad [19]$$

being

$$\left| \sum_{k \in \text{all shell}} \boldsymbol{l}_k \cdot \boldsymbol{s}_k \right| = 0 \qquad [20]$$

When the subshell is just half-filled, there is no multiplet splitting.

See also: Hartree and Hartree–Fock Methods in Electronic Structure.

PACS: 31.10.+z; 31.25.–v; 31.30.–i; 32.10.–f; 32.10.Hq

Further Reading

Bethe HA and Salpeter EE (1957) *Quantum Mechanics of One- and Two-Electron Atoms*. Berlin: Springer.
Bransden BH and Joachain CJ (eds.) (1983) *Physics of Atoms and Molecules*. England: Longman.
Condon EU and Shortley GH (eds.) (1953) *The Theory of Atomic Spectra*. Cambridge: Cambridge University Press.
Messiah A (ed.) (1969) *Quantum Mechanics*, vols. I and II. Amsterdam: North-Holland.
Rose ME (1957) *Elementary Theory of Angular Momentum*. New York: Wiley.
Slater JC (1968) *Quantum Theory of Matter*, 2nd edn. New York: McGraw-Hill.

Electronic Structure (Theory): Molecules

R Righini, University of Florence, Florence, Italy

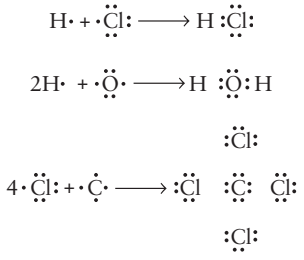
Introduction

In this article, some of the most popular and simple empirical theories of electronic structure and chemical bonding are described. The quantum mechanical treatment is then introduced in a qualitative way, with reference to valence bond and molecular orbital theories. A more detailed description of a quantitative evaluation of molecular properties is also presented.

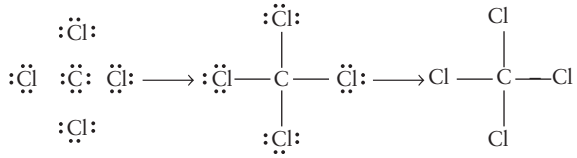
Lewis Theory of Bonding

The description of the electronic structure of molecules in terms of the bonds connecting the constituting atoms is a natural consequence of basic chemical concepts. Historically, the Lewis theory represents the first systematization of the matter.

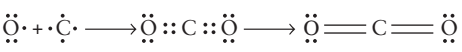
The basic idea is that chemical bonds are made by sharing one electron pair between two atoms. Thus, the electronic structure of some simple molecules can be represented (each dot represents one electron) as



The bonding electron pairs are generally indicated by a dash (the bond), while the nonbonding pairs (lone pairs) are often omitted:



In a similar way, the Lewis theory describes the formation of double and triple bonds:



In these examples, all the atoms belonging to the first row of the periodic table satisfy, the octet rule, with the exception of hydrogen, that is, they form bonds in such a way that each atom is surrounded by eight electrons. This is, in fact, the most stable electronic configuration, the so-called closed shell configuration characteristic of the noble gas atoms.

Electron Pairs and Molecular Geometry

The Lewis picture of chemical bonding and molecular structure does not account at all for the molecular shape. This is, of course, a fundamental aspect of the molecular electronic structure: it can be introduced along the same phenomenological line as that of the Lewis theory, by using the concept of valence shell electron pair repulsion (VSEPR). This approach provides a simple and efficient tool for rationalizing and predicting the geometry of molecules, starting from their Lewis electronic structure. The VSEPR method is based on a set of simple and essentially empirical rules, with the main assumption that the geometry of a molecule depends only upon the electron–electron interactions. The main points are:

- Atoms in a molecule are bound by bonding electron pairs; multiple bonds consist of more than one set of bonding pairs. The electron pairs not involved in bonds are called lone pairs.
- The general criterion for organizing bonding pairs and lone pairs around any atom in a molecule is that of minimum mutual interaction between electron pairs. This rule comes from the simple and intuitive idea that negatively charged electrons repel each other.

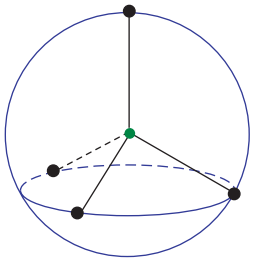
Applying the method requires that the single, double, and triple bonds binding the atoms in the molecule and the lone pairs are first determined on the basis of the Lewis theory. VSEPR then provides the criteria for positioning the bonds in space.

It can be directly and easily applied to molecules containing a central atom and just single bonds, but it can be easily extended to less simple structures.

Elementary geometrical considerations, based on the idea of maximizing the interpair angular distances around the central atom, give the following geometries for the disposition of n electron pairs (disregarding whether bonding pairs or lone pairs):

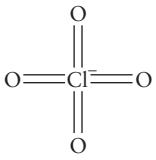
Electron pairs	*Geometry*
2	Linear
3	Trigonal
4	Tetrahedral
5	Trigonal bipyramidal
6	Octahedral

The method fully accounts for the structure of a molecule such as methane CH_4, in which all the H–$\hat{C}$–H angles equal the tetrahedral angle of 109.5°.



Also, the geometry of ammonia NH_3 is correctly predicted (the same as for methane, with one bond replaced by the lone pair sitting on the nitrogen atom). However, the experimental H–$\hat{N}$–H angle in ammonia does not coincide with the tetrahedral angle (107° instead of 109.5°). In order to complete the picture, a further VSEPR rule is introduced, stating that lone pairs require more space than bonding pairs.

In the VSEPR method double and triple bonds are treated, from a geometrical point of view, as single bonds. For instance, the Lewis structure of the perchlorate ion ClO_4^- is



The theory correctly predicts, for this ion, a tetrahedral structure (**Figure 1**), the same that would have been obtained if all the bonds were single bonds.

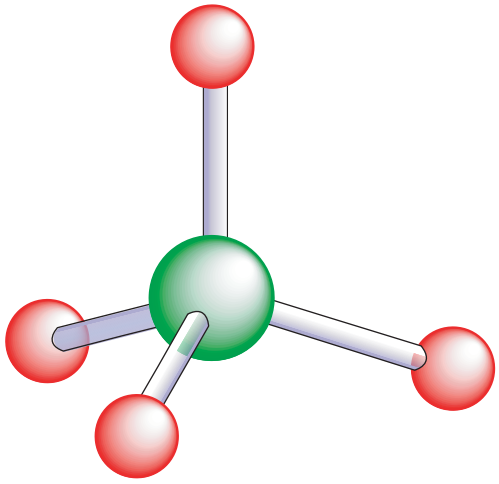
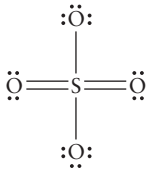


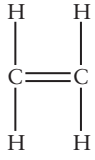
Figure 1 The tetrahedral structure of perchlorate ion.

According to VSEPR rules, the same is true in cases where both single and double bonds are present, such as in the sulphate ion SO_4^{2-}, whose Lewis structure is

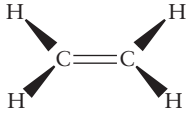


Also in this case, a tetrahedral geometry is correctly predicted.

VSEPR can also be applied to molecules without a central atom: for instance, ethylene C_2H_4 has the Lewis structure



Each carbon atom is surrounded by three bonds (two single bonds and one double bond). The model predicts that the bonds are distributed around each carbon atom according to a trigonal geometry. The overall shape of the molecule is then



with 120° angles between the bonds. Strictly speaking, the planar structure of ethylene molecule cannot be inferred from these simple arguments: its justification requires concepts concerning the nature of the double bonds that can be better understood in the quantum mechanical picture (*vide infra*). This is just one of the limitations of the VSEPR method. It has been already noticed that only approximated values are predicted

for the bond angles; in addition, the method does not tell anything about the length of the bonds. The strength of the method is its simplicity, but it clearly lacks a sound theoretical basis that can justify the overall picture and provide quantitative estimates.

Quantum Mechanical Description of Molecular Electronic Structure

The proper treatment of the electronic structure of atoms and molecules requires adopting the fully quantum mechanical approach. For a molecule, one has to consider a collection of n electrons and N nuclei interacting through the Coulomb potential. The quantum treatment requires that the Schrödinger equation

$$H\Psi = E\Psi$$

for the full system is solved, with the Hamiltonian H containing the kinetic and potential energy of all the electrons and nuclei. In principle then, the eigenfunction Ψ explicitly depends on both electronic and nuclear coordinates.

Here comes the main assumption common to all the generally adopted quantum treatments for molecules: the Born–Oppenheimer approximation. It derives from considering that, since the mass of the nuclei is much larger than that of the electrons, the nuclear coordinates can be considered fixed in the timescale typical of the very fast electronic motions. This assumption allows one to consider just an electronic Hamiltonian H_{el}, which does not contain the kinetic energy of the nuclei. In other words, the positions of nuclei are considered fixed, and the electronic Schrödinger equation $H_{el}\Psi_{el} = E\Psi_{el}$ is solved for the given nuclear geometry. By choosing different sets of nuclear positions, it is possible to follow the dependence of the electronic structure and energy on the nuclear coordinates.

The orbital description of the electronic structure of molecules follows the scheme already adopted for atoms. In the orbital approximation, each electron is assumed as independently described by a wave function, consisting of the product of a function ϕ depending on the spatial coordinates of the electron and of a spin function. The overall wave function of the molecule is then given by the product of the spin-orbitals, that is, of the one-electron wave functions, including the spin functions:

$$\Psi = \phi_1(1)\phi_2(2)\ldots\phi_n(n) \times \text{spin function}$$

There are two main approaches to the quantum mechanical description of the molecular electronic structure: the valence bond (VB) theory and the molecular orbital (MO) theory. Both were formulated in 1927; many of the most common and intuitive concepts of chemical bonding can be found in the VB theory, but the MO theory became more and more popular during the years, being definitely better suited for quantitative estimates of the molecular properties.

Valence Bond Theory

The VB theory can be seen as the translation, in quantum mechanical language, of the basic Lewis ideas on chemical bonding. The basic assumption is that of considering a molecule as formed by atomic "cores" (i.e., atomic nuclei surrounded by inner electrons) and valence (outer) electrons, which are responsible for the formation of chemical bonds. The starting point is the orbital description of the electronic structure of atoms: bonding in molecules occurs via the overlap of two half-filled orbitals of the atoms taking part in the bond formation.

In the case of the hydrogen molecule H_2, there are two nuclei a and b, and two electrons 1 and 2: one possible structure (ϕ_1) is that electron 1 is assigned to nucleus a (in the atomic orbital $1s_a$) and electron 2 to nucleus b (in the atomic orbital $1s_b$). Since the electrons are indistinguishable, the opposite assignment (ϕ_2) is also equally probable. A good wave function (Ψ) for the system is then the linear combination of the two structures ϕ_1 and ϕ_2:

$$\phi_1 = 1s_a(1)1s_b(2), \quad \phi_2 = 1s_a(2)1s_b(1)$$
$$\Psi = c_1\phi_1 + c_2\phi_2$$

By introducing this wave function in the Schrödinger equation, one obtains the two solutions $\Psi_+ = N_+(\phi_1 + \phi_2)$ and $\Psi_- = N_-(\phi_1 - \phi_2)$, where $N_\pm$ is a normalization factor. The corresponding energies are

$$E_+ = \frac{Q + J}{1 + S}, \quad E_- = \frac{Q - +J}{1 - S}$$

where Q is the Coulomb integral $\langle \phi_1|H|\phi_1 \rangle$, J is the exchange integral $\langle \phi_1|H|\phi_2 \rangle$, and S is the overlap given by the integral $S = \langle 1s_a(1)|1s_b(1) \rangle \langle 1s_a(2)|1s_b(2) \rangle$.

The overlap of the $1s$ orbitals to give the H–H bond is shown in **Figure 2**.

The $\Psi_\pm$ functions depend only on the spatial electronic coordinates: the complete wave function is obtained by multiplying the Ψ functions by the appropriate spin function. For the two electrons, there are four possible spin functions:

$$\alpha(1)\alpha(2)$$
$$\beta(1)\beta(2)$$
$$1/\sqrt{2}(\alpha(1)\beta(2) + \alpha(2)\beta(1))$$
$$1/\sqrt{2}(\alpha(1)\beta(2) - \alpha(2)\beta(1))$$

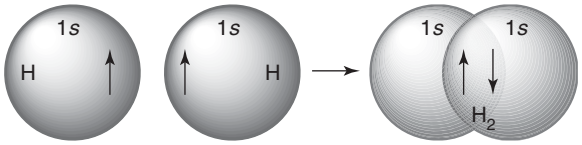


Figure 2 The H–H bond if formed by the overlap of the 1*s* atomic orbitals of hydrogen.

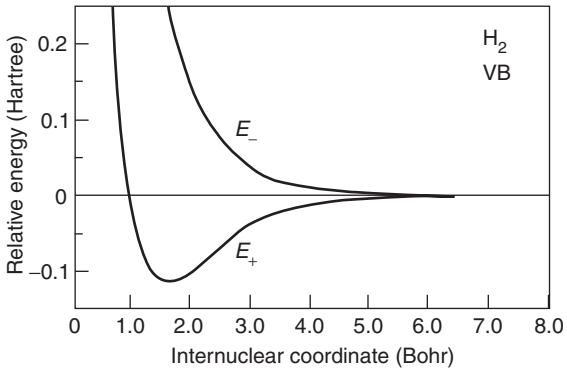


Figure 3 The Ψ_+ wave function corresponds to the bound state of molecular hydrogen: the corresponding energy E_+ has a minimum at the equilibrium internuclear distance. Ψ_- instead is a dissociative state: the E_- energy has no minimum.

The Pauli principle requires that the overall wave function is antisymmetric for the exchange of the coordinates of the electrons. Ψ_+ is clearly symmetric, and can only be multiplied by the antisymmetric spin function (the last one): its spin multiplicity is 1, and Ψ_+ corresponds to the singlet state of the bound atoms. The antysimmetric Ψ_- can be multiplied by any of the other three symmetric spin functions: the corresponding triply degenerate state is a triplet state.

Following the Born–Oppenheimer approximation, the Schrödinger equation can be solved for different interatomic distances; the results for the energies are given in **Figure 3**. Evidently Ψ_+ corresponds to the ground state of the molecule: its (monodimensional) energy surface has a minimum corresponding to the stable molecular structure. E_-, instead, has no minimum, and the corresponding Ψ_- wave function describes a dissociative state.

In the case of atoms with higher atomic number, bonding involves higher orbitals: fluorine for instance, has the atomic structure F: $(1s)^2(2s)^2(2p)^5$; the electrons in the 1*s* and 2*s* orbitals are part of the atomic core; four of the five 2*p* electrons form two nonbonding pairs that occupy two 2*p* orbitals. The third 2*p* orbital contains one unpaired electron, which is available for bonding. The bonding to H to form the hydrogen fluoride HF can be depicted as the overlap of the 1*s* orbital of hydrogen and the 2*p* orbital of fluorine containing one electron (**Figure 4**). In a similar way, the two collinear 2*p* orbitals of two

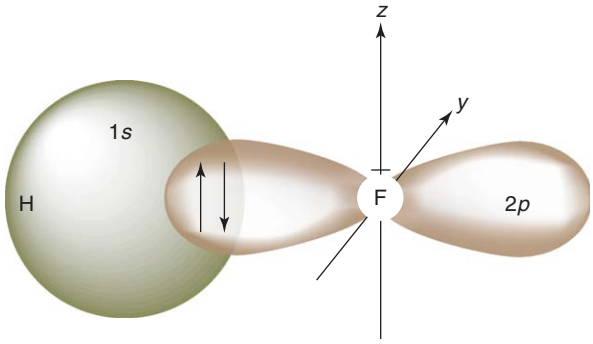


Figure 4 1*s* orbital of hydrogen and one 2*p* orbital of fluorine overlap to accommodate the bonding pair of hydrogen fluoride HF.

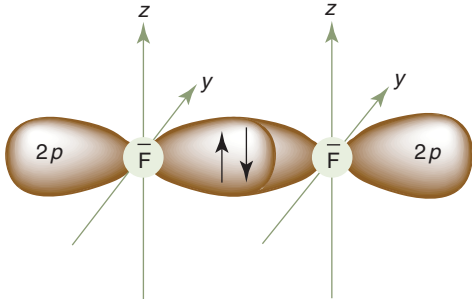


Figure 5 The F–F bond of molecular fluorine F_2 is formed by the overlap of two collinear 2*p* atomic orbitals.

fluorine atoms overlap to accommodate the bonding pair of the F_2 molecule (**Figure 5**).

Hybrid Orbitals

The *p* orbitals being mutually orthogonal, it is evident that the VB picture needs some improvement, in order to account for the trigonal, tetrahedral, octahedral, etc., molecular structures observed experimentally and predicted by the VSEPR method (**Figure 6**). This can be done by introducing the concept of orbital hybridization: the carbon atom provides a good example.

The electronic configuration of the carbon atom is $(1s)^2(2s)^2(2p)^2$; the 1*s* electrons are core electrons and do not take part in bond formation. The four valence electrons are distributed among the available orbitals according to

2*s*	$2p_x$	$2p_y$	$2p_z$
↓↑	↑	↑	

In this configuration, carbon has only two electrons available for pairing. Thus only two bonds are possible. This is obviously in contrast with the existence of compounds such as methane CH_4. The VB solution

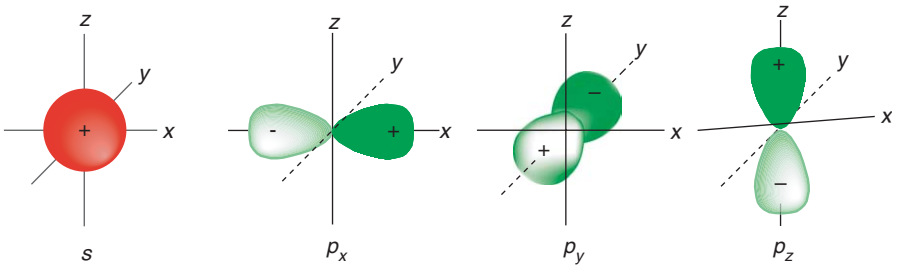


Figure 6 The 2*s* and the three orthogonal 2*p* atomic orbitals.

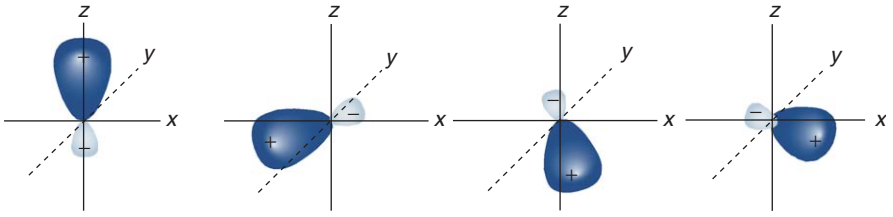


Figure 7 The four hybrid orbitals of sp^3 type, generated by the mixing of the atomic orbitals of **Figure 6**, are directed in the tetrahedral directions.

to this inconsistency is the introduction of hybrid orbitals. The 2*s* and the three 2*p* wave functions mix together yielding four iso-energetic hybrid orbitals of the sp^3 type:

sp^3 orbitals

↑	↑	↑	↑

The sp^3 orbitals are directed along the tetrahedral directions, and each one is occupied by one unpaired valence electron (**Figure 7**). The overlap of the sp^3 hybrid orbitals with the 1*s* orbital of hydrogen leads to the formation of the molecule of methane (**Figure 8**).

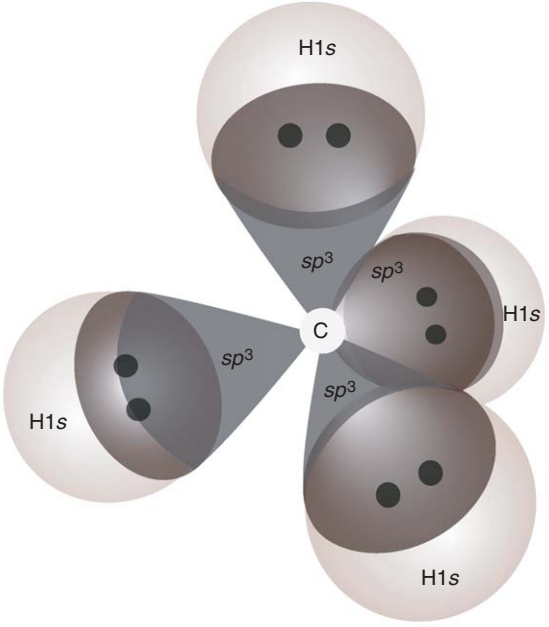


Figure 8 In the VB theory, the tetrahedral geometry of methane CH_4 is attributed to the formation of four *s* bonds involving the sp^3 hybrid orbitals of methane.

σ and π Bonds

In all the bonds considered above, the electron density has a cylindrical symmetry around the line connecting the two bound atoms. Bonds of this type are named σ bonds. There is also a different way of having two atomic orbitals overlapped. Two *p* orbitals, for instance, may overlap sideways, forming a π bond (**Figure 9**).

A good example again comes from the carbon atom. Starting from its electron configuration $(1s)^2(2s)^2(2p)^2$, if two of the three 2*p* orbitals are hybridized with 2*s*, three sp^2 hybrid orbitals are generated, whose lobes are coplanar (with angles of 120°) and directed along the trigonal directions:

2*s*	$2p_x$	$2p_y$	$2p_z$	→	sp^2 orbitals				$2p_z$
↓↑	↑	↑			↑	↑	↑	↑	↑

The three sp^2 hybrid orbitals are used to form three σ bonds per carbon atom in the *xy* plane. This is the case of ethylene C_2H_4 in **Figure 10**.

The remaining $2p_z$ orbitals, orthogonal to the σ bond plane, overlap sideways, giving rise to a bond of π character between the two carbon atoms (**Figure 11**).

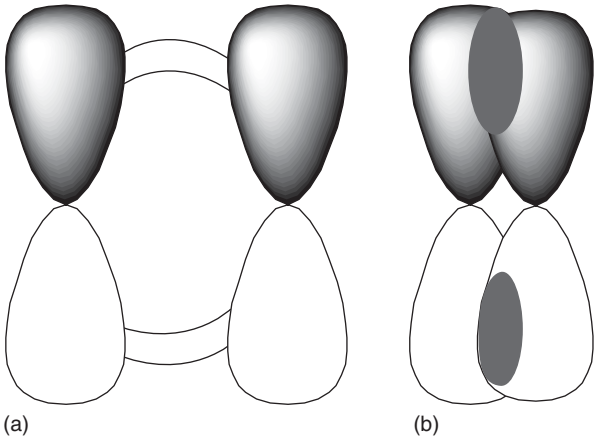


Figure 9 Two 2*p* orbitals overlap sideways to form π bond.

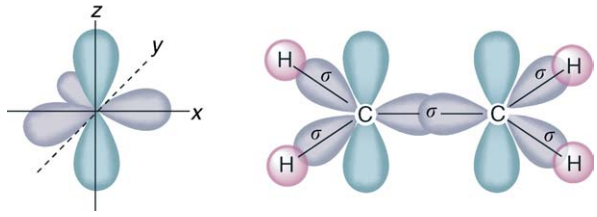


Figure 10 On the left: the sp^2 hybrid orbitals, obtained from 2*s* and two 2*p* atomic orbitals. On the right: the sp^2 hybrid orbitals participate in the formation of the σ type bonds of ethylene molecule.

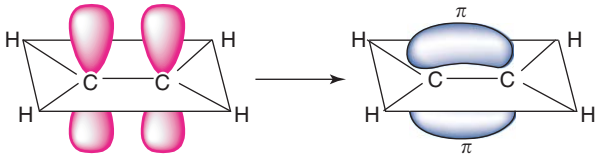


Figure 11 The π bond between the carbon atoms participates in the formation of the double C=C bond of ethylene. Its geometrical constraints force the molecule to assume a planar configuration.

The π bond is characterized by two lobes separated by a nodal plane (the *xy* plane) where the electron density is zero. It is evident that the two CH_2 groups must be coplanar so that the double bond C=C may be set up.

The VB theory translates the most basic concepts of chemical bonding in the language of quantum mechanics; it has been very successful in providing support to essentially empirical theories, such as, the Lewis theory and VSEPR, and has established a large part of the language that chemists use in describing the properties of chemical bonds. The very fast development of computers, however, demonstrated severe limits of the VB approach in applications where quantitative estimates are required: accurate calculations of molecular energies, of excited state structures and dissociation energies, of molecular multipole moments, polarizability, etc, are especially for large molecules, practically not possible in the VB framework. The alternative quantum mechanical approach, the MO theory, acquired more and more popularity for its easy implementation in computer programs: it is now practically the only method in use in the quantum chemistry community.

Molecular Orbital Theory

Differently from the VB theory, the MO theory does not base its treatment on the formation of bonds between the atoms of a molecule. Its goal is, instead, that of describing the distribution of electrons around the atomic nuclei that represent the backbone of the molecule. Each electron is then described by a wave function (the molecular orbital), extended over all the nuclei. The Born–Oppenheimer approximation, allowing for a separation of the timescales of the electron and nuclear motion, is at the basis of the treatment. The Schrödinger equation cannot be solved exactly for polyelectronic molecules: the MO theory employs approximate methods to solve the problem.

The most simple molecule to start with is the hydrogen molecular ion H_2^+. Its electronic Hamiltonian is

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r_a} - \frac{1}{r_b}$$

where the first term on the right-hand side is the kinetic energy, and the other two terms are the potential energy of the electron in the field of the two nuclei *a* and *b*.

This monoelectronic system is the only molecule for which the Schrödinger equation can be solved exactly; however, it is useful to consider its approximate treatment, since it provides the general scheme for describing the electronic structure of more complex molecules.

The first step is that of choosing a convenient form for the zero-order approximate wave function. It is easily acceptable in the MO spirit, that this function should correspond to the electron "belonging" both to nucleus *a* and to nucleus *b*: a simple form is $\phi = c_a 1s_a + c_b 1s_b$, where only the 1*s* atomic orbitals have been considered. This choice for the basic MO function takes the name of linear combination of atomic orbitals (LCAO); it is a frequent choice, although absolutely not the only one possible.

Minimization of the energy provides the values for the coefficients c_a and c_b, and the resulting orbitals are

$$\phi_1 = N_+(1s_a + 1s_b), \quad \phi_2 = N_-(1s_a(1) - 1s_b(2))$$

($N_\pm$ are normalization factors)

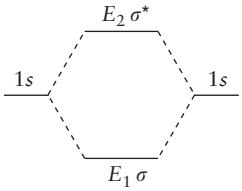
with corresponding energies

$$E_1 = \frac{H_{aa} + H_{ab}}{1 + S_{ab}}, \qquad E_2 = \frac{H_{aa} - H_{ab}}{1 - S_{ab}}$$

with the definition of the integrals

$$H_{ab} = \langle 1s_a | H | 1s_b \rangle, \qquad S_{ab} = \langle 1s_a | 1s_b \rangle$$

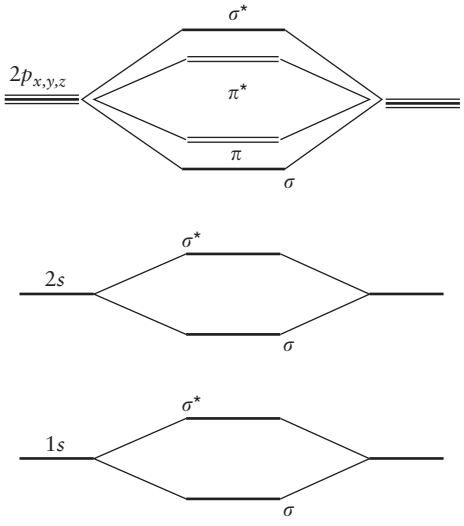
H_{ab} has a negative sign, so that the two VB energy levels are



The shape of the two molecular orbitals is shown in **Figure 12**.

The MO with energy lower than the two 1*s* atomic orbitals is said to be a bonding orbital; the one with higher energy is the antibonding orbital. Both MOs are cylindrically symmetric around the line connecting the bound atoms: they are σ molecular orbitals.

The mechanism of generating molecular orbitals can be extended to the combination of higher energy atomic orbitals 2*s*, 2*p*, etc. If *x* is the internuclear direction, the molecular energy level scheme produced from 1*s*, 2*s*, and 2*p* orbital is



The combination of the $2p_x$ atomic orbitals produces two σ MOs; the $2p_y$ and $2p_z$ atomic orbitals combine sideways, giving rise to two doubly degenerate π orbitals (**Figure 13**).

The extension of the method to the hydrogen molecule H_2 is immediate. With the Hartree assumption, the global electronic molecular wave function is the product of two one-electron wave functions, each one describing one electron in the $\sigma 1s$ bonding orbital:

$$\Psi = \sigma 1s(1)\sigma 1s(2)$$

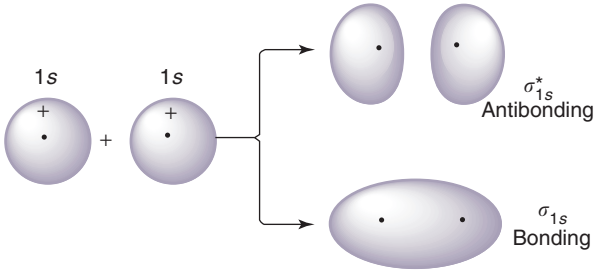


Figure 12 Bonding and antibonding molecular orbitals of H_2^+ molecular ion.

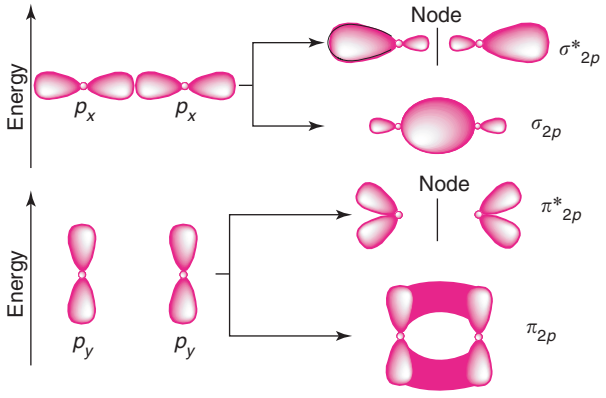
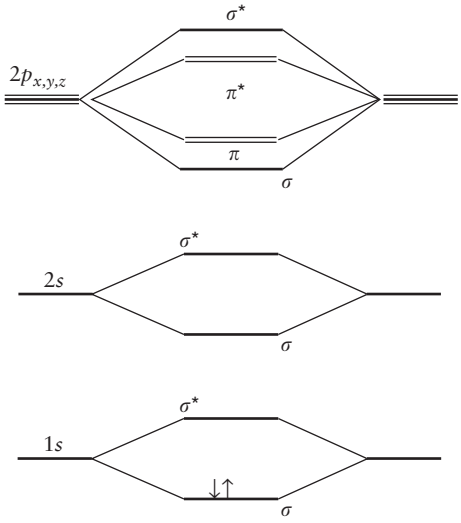


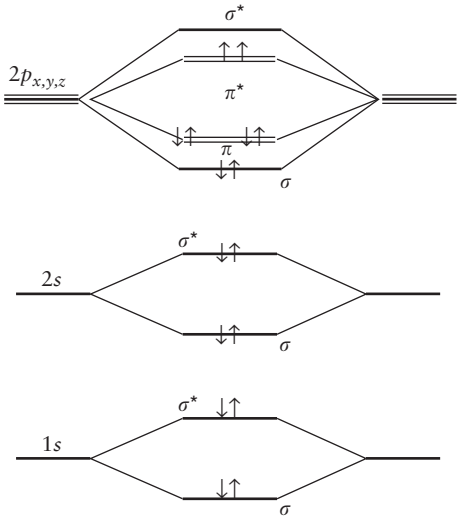
Figure 13 Bonding and antibonding molecular orbitals obtained by linear combination of 2*p* atomic orbitals. Upper figure: σ-type orbitals generated by two collinear 2*p* atomic orbitals; lower figure: π-type orbitals generated by two parallel 2*p* atomic orbitals.

The scheme of molecular orbitals developed for H_2^+ can be extended, in a qualitative way, to other diatomic molecules. Following the same procedure used for atoms, the sequence of molecular orbitals can be filled with the electrons provided by the atoms of the molecule: thus, for molecular hydrogen, one has just two electrons to allocate in the available orbitals; the ground state configuration is then



Similar to atoms, an aufbau principle holds for molecular orbitals also: the levels are filled in an increasing order of energy, respecting the Hund's rule (maximum multiplicity favored) and the Pauli principle (two electrons cannot be described by the same spin–orbital).

For instance, an oxygen atom has eight electrons: there are 16 electrons to be placed in the molecular orbitals given above. The distribution of the electrons in the MO is



The electronic configuration of molecular oxygen is then $(\sigma 1s)^2(\sigma^* 1s)^2(\sigma 2s)^2(\sigma^* 2s)^2(\sigma 2p)^2(\pi 2p)^4(\pi^* 2p)^2$. Note that the two electrons in the highest occupied doubly degenerate π^* orbital have parallel spins (as required by Hund's rule): the ground state of oxygen is then a triplet state, and the molecule is paramagnetic.

From the electronic configuration comes the definition of the bond order: it is defined as one half the number of electrons in bonding orbitals minus the number of electrons in antibonding orbitals. The bond order in an oxygen molecule is then 2.

The MO approach has been particularly successful in explaining the structure of a class of organic molecules called conjugated molecules. These are planar organic molecules consisting of a framework of carbon atoms joined by alternating single and double bonds. A typical example is benzene C_6H_6 (**Figure 14**). The six carbon atoms are at the vertices of a planar hexagon in the *xy* plane; each carbon is bound to two neighbor carbons and to one hydrogen (not shown). The planar structure originates from σ bonds involving the $2p_x$ and $2p_y$ atomic orbitals. The three double bonds in the above figure are due to π orbitals obtained by the linear combination of the $2p_z$ orbitals, orthogonal to the molecular plane.

In reality, all six C–C bonds in benzene are identical (the two bonding schemes in **Figure 14a** are equally probable) and possess properties intermediate between those for single and double bonds. The electron pairs forming the π bonds are not localized between specific carbon atoms but are delocalized over the whole ring (**Figure 14b**). A simple adaptation of the MO theory, called Hückel theory, which takes the σ bonds for granted and describes the π MOs as linear combinations of the $2p_z$ atomic orbitals on each carbon atom, provides a very good explanation of the properties of conjugated molecules.

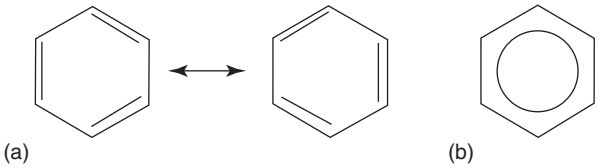


Figure 14 The actual structure of benzene molecule "resonates" between the two structures in (a). The delocalization of the three double bonds is often represented as shown in (b).

The shape of the six MOs, obtained by combining the six $2p_z$ carbon orbitals in the Hückel approximation together with the distribution of the six $2p_z$ electrons in the energy levels, is given in **Figure 15**. The three occupied MOs are bonding orbitals; no electron occupies the antibonding orbitals. The π contribution to the bond order for the six C–C bonds is then 1/2.

The highest occupied molecular orbital (HOMO), and the lowest unoccupied molecular orbital (LUMO) are widely used acronyms indicating the so-called frontier orbitals. In general, the energy difference of the two, and their energy separation from the relative zero-energy determine many spectroscopic and chemical properties of molecules.

The above discussion is very useful for understanding many aspects of the electronic structure of molecules, but is essentially qualitative. For a quantitative treatment, the full electronic Hamiltonian of the molecule must be considered, the necessary approximations introduced, and the solutions at different levels of accuracy discussed and compared.

Molecular hydrogen is the simplest case: the exact electronic Hamiltonian for H_2 is

$$H = -1/2\nabla_1^2 - 1/2\nabla_2^2 - \frac{1}{r_{a1}} - \frac{1}{r_{a2}} - \frac{1}{r_{b1}} - \frac{1}{r_{b2}} + \frac{1}{r_{12}}$$

that can be written as the sum of two one-electron Hamiltonians plus the electron–electron interaction terms $1/r_{12}$, which makes the exact solution of the Schrödinger equation impossible:

$$H = H_1^o + H_2^o + \frac{1}{r_{12}}$$

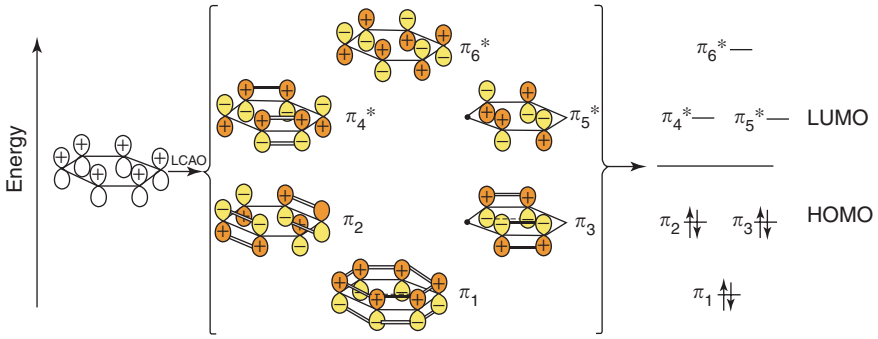


Figure 15 The π molecular orbitals of benzene as described in the Hückel picture. On the right hand side of the figure, the orbital energy levels and their occupancy by the $2p_z$ electrons are shown.

With the Hartree assumption, and following the example of H_2^+, the general form of the molecular wave function is the product of two one-electron orbitals and of an antisymmetric spin function

$$\Phi = \sigma_g 1s(1)\sigma_g 1s(2)\frac{1}{\sqrt{2}}[\alpha(1)\beta(2) - \alpha(2)\beta(1)]$$

With the LCAO assumption, its spatial part is

$$\Phi = N[1s_a(1) + 1s_b(1)][1s_a(2) + 1s_b(2)]$$

The approximate solution is obtained by minimization of the variational integral

$$\langle \Phi|H|\Phi \rangle = 2E_1 + \langle \Phi|1/r_{12}|\Phi \rangle$$

The evaluation of the integral on the right-hand side is the main obstacle. The Hartree–Fock method provides an approximate solution to the problem: it transforms the n-dimensional problem (for a molecule with n electrons) in the solution of n one-electron eigenvalue equations of the type

$$f\phi_i = \varepsilon_i \phi_i$$

where f is the Fock operator, an effective one-electron operator, which includes the mean field generated by all the $n-1$ electrons and acting on electron i.

The ϕ_i functions are generally expressed as a combination of basis functions

$$\phi_i = \sum_s \chi_s c_{si}$$

The basis functions may be atomic orbitals (for the LCAO method), or may have other functional forms (Gaussian, for instance). Generally, a larger number of basis functions ensures a higher accuracy for the results of the calculation. However, the solution of Fock equations requires the evaluation of a large number of integrals involving the basis functions; this poses a practical limitation to the size of the basis set, specially for large molecules. The full Hartree–Fock method, where all the integrals are calculated is named *ab initio* Hartree–Fock method. From it, a number of approximate methods have been developed, in which all or part of the integrals are not explicitly calculated, but are treated as semi-empirical parameters. Methods known with acronyms, such as CNDO, MNDO, MINDO, etc. are widely used, because their computational cost is very low, while providing satisfactory results in many cases.

Electron Correlation

Ab initio Hartree–Fock calculations, nowadays, can be very accurate and reliable, even for large molecules. In addition to the ground electronic structure and energy, knowledge can be gained of the molecular excited states, of the dissociation energies, of the molecular multipoles, and polarizability; the vibrational and electronic spectra can be calculated. There is, however, an important point missing in the Hartree–Fock approach, namely the electron correlation. The Coulomb repulsion between electrons makes them "avoid" each other at any time: this effect is not fully accounted for by the mean-field interelectronic potential adopted in the Hartree–Fock method. Configuration interaction (CI) is one of the most general ways to improve upon the Hartree–Fock theory by taking into account the correlations between electron motions.

The wave function for the ground electronic configuration of the molecule thus includes contributions of the wave functions corresponding to excited configurations. Even for small systems, their number is enormously large. Generally, the CI expansion is limited to the singly and doubly excited configurations, but also with this restriction the method is not applicable to large molecules.

A different approach to electron correlation has become very popular in recent years: the second-order

Møller–Plesset perturbation theory (MP2). The Hartree–Fock Hamiltonian is the unperturbed operator $H^{(0)}$; the perturbation operator $H^{(1)}$ is the difference between the Hartree–Fock and the exact Hamiltonian. The first nonzero term of the perturbation expansion is the second-order correction to the energy $E^{(2)}$; it corresponds to including the contribution of an electron pair excitation, that is, it accounts for the correlation energy between electron pairs. MP2 leads to a simple and tractable expression for the correlation energy that can be easily implemented in computer programs, with satisfactory results for large molecules also.

Density-Functional Theory

An alternative approach to solving the problem of the electronic structure of molecules, which is becoming more and more popular, is the one known as the density-functional theory (DFT). The method is particularly suited for computer implementation, and, most of all, it allows one to include the electron correlation effects in a much simpler and less costly way.

The theory is based on the Hohenberg–Kohn theorem, stating that any molecular property can be expressed in terms of the electron density $\rho(x, y, z)$. Kohn–Sham then demonstrated that the ground-state energy of an n-electrons molecule is expressed in terms of the one-electron density $\rho(1)$ as

$$E_0 = -\frac{1}{2}\sum_{i=1}^{n} \langle \psi_i(1)|\nabla_i^2|\psi_i(1)\rangle - \sum_{\alpha}\int \frac{Z_\alpha \rho(1)}{r_{1\alpha}}\,\mathrm{d}v_1 + \frac{1}{2}\int\int \frac{\rho(1)\rho(2)}{r_{12}}\,\mathrm{d}v_1\mathrm{d}v_2 + E_{\mathrm{xc}}[\rho]$$

where ψ_i are the Kohn–Sham orbitals, Z_α is the nuclear charge, and E_{xc} is the so-called exchange–correlation energy; ρ is a function, then $E_{\mathrm{xc}}[\rho]$ is a functional of the electron density.

In a broad sense, it can be included in the family of the molecular orbital techniques: the Kohn–Sham orbitals are, in fact, obtained by solving equations formally similar to the Fock equations:

$$F_{\mathrm{KS}}(1)\psi_i(1) = \varepsilon_{i,\mathrm{KS}}\psi_i(1)$$

The Kohn–Sham operator contains (like the Fock operator) a mean-field potential term that is expressed as a function of the exchange–correlation energy $E_{\mathrm{xc}}[\rho]$. One important difference is that the electronic correlation contribution can be included from the beginning in $E_{\mathrm{xc}}[\rho]$. In contrast to the Hartree–Fock method, the calculations performed in the DFT do not need any further improvement to include electron correlation, with a remarkable saving of computer time. In spite of the formal similarities, DFT substantially differs from the MO treatment: the latter, in fact, optimizes the results by improving the approximate wave function. DFT, instead, is oriented toward improving the quality of the approximate Hamiltonian, which is eventually solved exactly.

The main drawback of DFT is that the form of the $E_{\mathrm{xc}}[\rho]$ functional cannot be known *a priori*; many efforts have been devoted to formulate convenient forms for this functional, and different "recipes" are now available that have shown a very good heuristic power. The popularity of the method has been growing fast: DFT is now at the basis of *ab initio* treatments of complex systems. For instance, the implementation of DFT techniques made it possible to perform molecular dynamics simulations calculating the forces from first principles: electronic properties of molecules and condensed phases, reaction mechanisms in solution, and structure and dynamics of biomolecules and biological systems (enzymes, for instance) can be investigated at *ab initio* accuracy level.

See also: Electronic Structure (Theory): Atoms.

PACS: 31.15.Ar; 31.15.Ct; 31.15.Ew; 31.15.Rh; 31.15.Ne; 31.15.Eb; 31.15.Qm

Further Reading

Abright TA, Burdett JK, and Whangbo MH (1985) *Orbital Interactions in Chemistry*. New York: Wiley-Interscience.

Atkins PW (1983) *Molecular Quantum Mechanics*, 2nd edn. Oxford: Oxford University Press.

Hehre WJ, Radom L, Schleyer PR, and Pople JA (1986) *Ab Initio Molecular Orbital Theory*. New York: Wiley.

Hincliffe A (1995) *Modeling Molecular Structures*. New York: Wiley.

Kunc K (1985) In: Devresse JT and Van Camp P (eds.) *Electronic Structure, Dynamics and Quantum Properties of Condensed Matter*. New York: Plenum.

Levine IN (1991) *Quantum Chemistry*, 4th edn. Englewood Cliff: Prentice-Hall.

McQuarrie DA (1983) *Quantum Chemistry*. Mill Valley: University Science Books.

Pisani C (ed.) (1996) *Quantum Mechanical Ab Initio Calculation of the Properties of Crystalline Materials*, Lecture Notes in Chemistry, vol. 67. Berlin: Springer.

Pople JA and Beveridge DL (1970) *Approximate Molecular Orbital Theory*. New York: McGraw-Hill.

Richards WG and Cooper DL (1992) *Ab Initio Molecular Orbital Calculations for Chemists*, 2nd edn. Oxford: The Clarendon Press.

Schatz GC and Ratner M (2000) *Introduction to Quantum Mechanics in Chemistry*. Englewood Cliff: Prentice-Hall.

Segall MD (2002) Applications of *ab initio* atomistic simulations to biology. *Journal of Physics: Condensed Matter* 14: 2957–2973.
Szabo A and Ostlund NS (1989) *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory*. New York: McGraw-Hill.
Tse JS (2002) *Ab initio* molecular dynamics with density functional theory. *Annual Review of Physical Chemistry* 53: 249–290.
Tuckerman ME (2002) *Ab initio* molecular dynamics: basic concepts, current trends and novel application. *Journal of Physics: Condensed Matter* 14: R1297–R1355.

Electronic Structure Calculations: Plane-Wave Methods

M L Cohen, University of California at Berkeley, Berkeley, CA, USA

Introduction

Understanding the electronic structure of solids is basic to explaining and predicting solid-state properties, but what is an appropriate description of electronic states in solids?

Although it is clear that solids are made of atoms and that one can think of the gas-to-liquid-to-solid transitions as a process in which atoms just get closer together, the electronic properties of solids often differ significantly from what one would expect from a collection of isolated atoms. In cases where the interactions between the atoms are weak, as in the case of rare gas solids, it is reasonable to consider these interactions as perturbations of atomic states. However, when atoms are close to each other in a solid, sometimes they can give up their outermost electrons easily, and it is even possible to end up with a sea of nearly free electrons. In fact, a nearly free-electron model is appropriate for metals such as the alkali metals, aluminum, and many others.

It is also fortunate and useful that this nearly free-electron model applies to a wider class of solids when a few modifications are added. A totally free electron is described by a plane-wave wave function. Hence, if one assumes that the electrons describing a metal behave as if they are nearly free, then a plane-wave position-dependent wave function $\Psi(\boldsymbol{r})$ confined to a box of volume Ω, having a wave vector $\boldsymbol{k}$ which can be written as

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega}} \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$

may be a good starting point for describing the electrons in metals such as sodium. If one examines the resulting electronic density corresponding to a plane-wave wave function, it is constant since it is proportional to $\psi^*(\boldsymbol{r})\psi(\boldsymbol{r})$. Therefore, the collection of atoms with localized wave function on each atomic site and peaked electronic density yields a valence electron density which is vastly different.

What about bonds? For example, covalent bonds arise because electrons tend to concentrate in certain regions between atoms. Can the bonds and charge modulation known to exist in metals and especially in covalent semiconductors be described using plane waves as a starting point? The answer is yes for a broad class of solids. A plane wave is a solution to Schrödinger's equation when the electron–ion and electron–electron potentials are zero. However, when these potentials are included (and if they are not too strong and if the electrons are not strongly correlated), a perturbative scheme can be used to modify the free-electron picture, and this is the "nearly free-electron model" which is described in many textbooks. However, more generally, Schrödinger's equation is solved by diagonalizing a Hamiltonian matrix with appropriate potentials and a basis set composed of plane waves to describe the wave function.

What about alternatives to the nearly free-electron model? The opposite approach to using plane waves is the utilization of atomic wave functions for the basis set. There are tight-binding methods where linearized combinations of atomic orbitals form the basis. This description is also very useful especially when the solid-state wave functions are not too different from the atomic wave functions forming them. If done correctly with sufficient orbitals in the basis set, both the above methods work and are useful.

In the following discussion, the plane-wave basis is featured. In particular, a pseudopotential-density functional approach is described which has been applied successfully to a large number of solids, clusters, nanostructures, and molecules.

The Standard Model of Solids

In **Figure 1**, a schematic picture of a solid is presented where the nuclei and core electrons for the atoms are represented as positive cores and the stripped-off valence electrons are contributed to the electron sea which flows through the periodic array of cores. In this model, the solid system is viewed as having two types of particles: cores and valence electrons. The interaction between the positive cores is taken to be a standard Coulomb interaction describable by Madelung sums. The electron–core and electron–electron

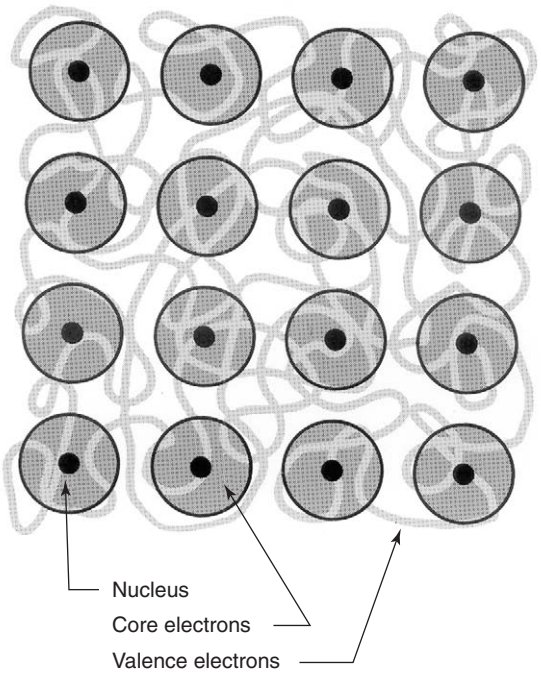


Figure 1 Schematic view of a solid with positive cores at fixed lattice positions in a periodic array and a sea of relatively free valence electrons moving through the lattice.

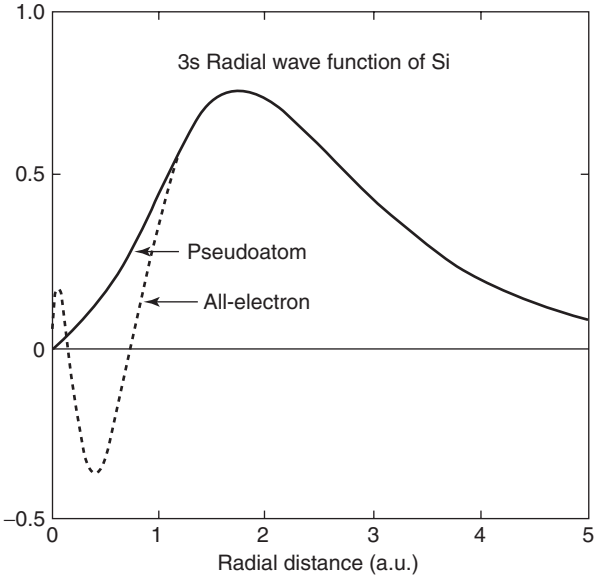


Figure 2 The solid line represents the Si 3*s* pseudo-wave-function resulting from the use of a pseudopotential. The dashed line results from an "all electron" calculation.

interactions are essential inputs to this model. As mentioned above, the form of the standard model of solids, represented by **Figure 1**, which is focused here, will model the electron–core interaction using pseudopotentials and the electron–electron interaction using the density-functional theory – in particular, the local density approximation (LDA).

Pseudopotentials

Arguments for the use and/or appropriateness of a weak electron–core potential have been made since the 1930s. In 1934, Fermi constructed a pseudopotential to describe highly excited electronic states of alkali atoms. His weak potential was designed to yield an accurate description of the outer portions of the wave function and to avoid the more difficult calculation of the properties of the oscillations of the wave function near the core. This would require a strong potential. Since solid-state effects are dominated by the outer portions of atomic wave functions, this is a viable approach for constructing pseudopotentials for solids. As shown in **Figure 2**, a "pseudoatom" wave function is generated which is identical to an "all-electron" 3*s* radial wave function for Si away from the core and smooth near the core. The weak pseudopotential which produced the pseudoatom wave function can be computed from first principles using only the atomic number as input. There are many schemes available for producing these *ab initio* pseudopotentials, which in principle can be constructed for all atoms.

The electron–electron interactions can be approximated using the electronic density. This density-functional approach represents the Hartree and exchange-correlation potentials as functionals of the electronic density. Hence, a self-consistent calculation can be done with a plane-wave basis set, pseudopotentials, and the LDA to density-functional theory. The calculated electronic density can be used to produce LDA potentials which in turn can produce updated wave functions and subsequently, new LDA potentials. This scheme and variants using pseudopotentials and plane waves are often referred to as the "plane-wave pseudopotential method" (PWPM). In general, the method has been extremely successful for determining the electronic structure and many other properties of solids.

Although most modern pseudopotential calculations use *ab initio* pseudopotentials of the type described above, the pseudopotential method went through a semi-empirical stage of development where experimental input was used to generate a potential representing the total valence electron potential (electron–core plus electron–electron). This approach, called the empirical pseudopotential method (EPM), used plane waves and a few Fourier coefficients of the potential which were fit to experimental data. In addition to explaining the origin of the optical structure in the visible and UV in terms of interband electronic transitions and many other properties of solids,

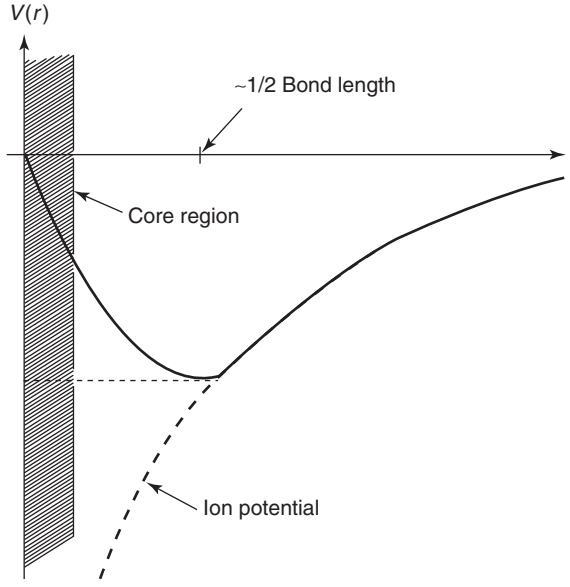


Figure 3 Schematic drawing of a pseudopotential (solid curve) compared to a Coulomb ion potential (dashed line).

the EPM set the stage for the first-principles methods and the standard model.

A schematic real space potential is illustrated in **Figure 3**. The strong ionic Coulomb potential is "cut off" in the core region. The physical origin of the cancelation of the ionic potential in the core region was attributed by Phillips and Kleinman to a repulsive potential arising because of the Pauli exclusion principle preventing the valence electrons from occupying core states. The result is again a weak net potential describable by a few Fourier form factors and a wave function which can be expanded in a plane-wave basis set. Usually, three pseudopotential form factors per atom gave highly accurate electronic band structures.

The EPM was the first successful plane-wave method for real materials with applications to metals, semiconductors, and insulators. Dozens of band structures and optical response functions were calculated with high precision for the first time. The optical spectra of semiconductors were interpreted in terms of critical points in the electronic band structure, and when the wave functions were used to compute the electronic density, the covalent and ionic bonds were pictorially displayed – again for the first time.

In **Figure 4**, a plot of the predicted electronic density for Si is displayed and compared with results based on subsequent X-ray studies. In addition to the striking agreement between theory and experiment, these plots reveal the nature of the covalent bond. It is particularly satisfying to see the local build-up of charge in the Si–Si bond region even though the basis

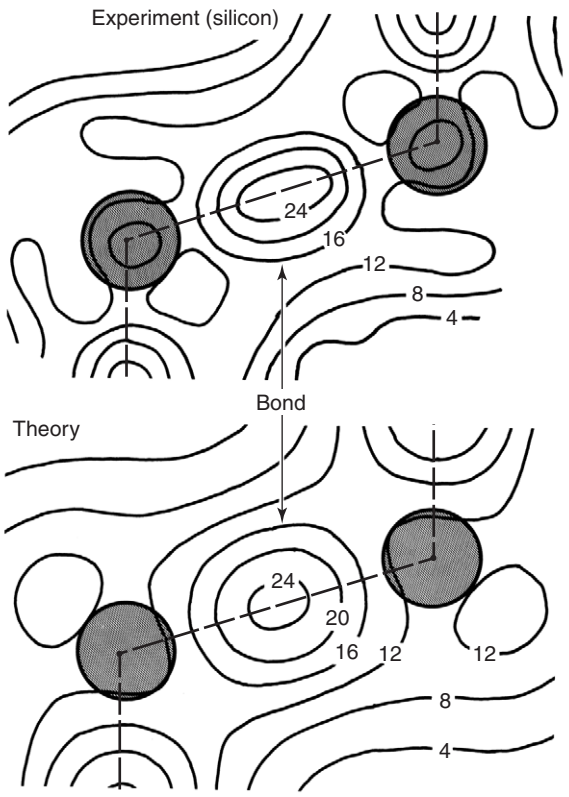


Figure 4 Valence electron density for silicon. The contour spacings are in units of electrons per unit cell volume. The shaded disks represent atomic cores.

set is composed of plane waves which yield a constant density before the potential is "turned on." There is no prior prejudice for this pile-up of charge in the bond region built into the basis set.

Total Energy Calculations and Structural Properties

The EPM calculations were constructed to deal with bulk crystals which were not terminated. Standard techniques, described in textbooks, use periodic boundary conditions to avoid having surfaces or interfaces. The EPM pseudopotential $V(\boldsymbol{r})$ has the periodicity of the lattice so it could be expressed as an expansion using reciprocal lattice vectors $\boldsymbol{G}$:

$$V(\boldsymbol{r}) = \sum_{\boldsymbol{G}} S(\boldsymbol{G}) V(\boldsymbol{G}) \mathrm{e}^{\mathrm{i}\boldsymbol{G}.\boldsymbol{r}}$$

where $S(\boldsymbol{G})$ is the structure factor that locates the atoms relative to a lattice point. It is the form factors $V(\boldsymbol{G})$ which are fit to experiment in the EPM scheme.

In order to deal with surfaces and interfaces, one has to allow charge to redistribute at a surface or interface and break translational invariance. To do

this, the EPM evolved into a scheme where the electron–core and the electron–electron potentials were separated. As in modern calculations, the electron–core potential is referred to as the "pseudopotential" and the electron–electron potential was expressed in terms of the density. In early calculations, the Slater and Wigner approximations were used. This allowed the charge density to readjust to a surface or interface and because of self-consistency, so did the potentials. The surface or interface was modeled by introducing a supercell which contained the geometric configuration of interest. For example, a (1 1 1) Si surface was modeled using a slab containing 12 atomic layers separated from the next slab by an equivalent space with no atoms. The surface of the slab represented the surface of the Si crystal and self-consistency required the rearrangement of the electronic density to correspond to the new geometry. **Figure 5** shows this rearrangement of electron density on an ideal Si (1 1 1) surface in one of the first calculations done with this method. Later calculations used supercells to examine Schottky barriers, heterojunctions, and reconstructions at surfaces and interfaces.

The evolution of the use of the LDA was straightforward since in this approach the electron–electron potentials are expressed as functionals of the electron density. Hence, surfaces, interfaces, and localized configurations in general could be dealt with using supercells and by letting the charge density adjust to the geometric constraints. The use of the PWPM and the LDA also allowed the exploration of other solid-state properties.

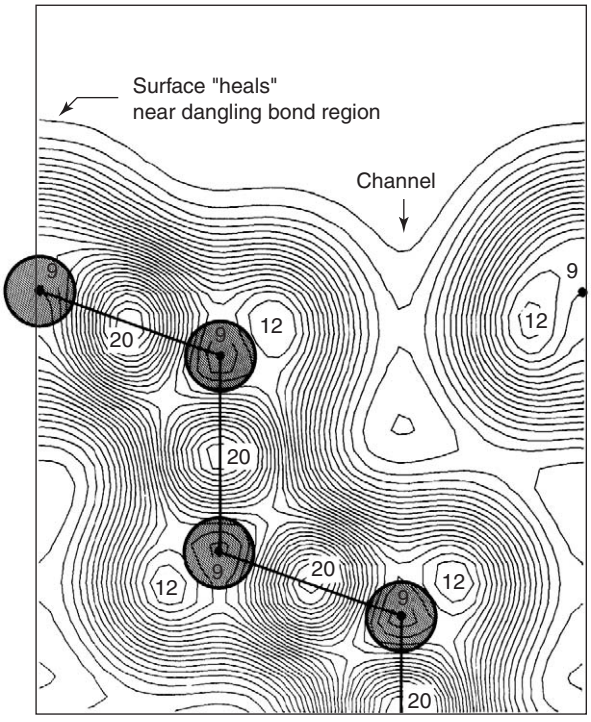


Figure 5 Valence electron density for the ideal Si (1 1 1) surface. Contours are plotted in the perpendicular (1 1 0) plane with units corresponding to one electron per primitive cell. The shaded disks represent atomic cores.

Important applications became possible once a scheme was developed to calculate the total structural energy E_S of the solids for given configurations of the atomic cores. Once E_S is evaluated for a series of candidate structures, a comparison of their structural energies immediately reveals the lowest energy structure for the group. At a given volume, this lowest energy structure would be the most stable structure if all possible structures are considered. Since it is not possible to test all structures, a limitation on this method is the possible omission of the most stable structure. However, in practice, it has often been possible to choose the most likely structures, and the success rate of the PWPM is high for predicting new structures and structural properties. Because E_S depends on volume (or lattice separations), changes induced by pressure can cause solid–solid phase transitions where crystal structures can change. Therefore, calculations of the dependence of E_S on volume can predict transition pressures and volumes for the structural phase changes. Similar calculations for lattice constant changes caused by strains or alloying can yield elastic constants and even vibrational spectra.

Using the standard model with cores and valence electrons as illustrated in **Figure 1**, the total structural energy of the solid can be expressed as

$$E_S = E^{cc} + E^{ec} + E_k^e + E_c^{ee} + E_x^{ee}$$

and all the energy components can be evaluated using the PWPM. The core–core Coulomb energy E^{cc} can be calculated from point–ion sums as discussed earlier, E^{ec} is evaluated using the pseudopotential, and both the Hartree (Coulomb) electron–electron energy E_c^{ee} and the exchange-correlation energy E_x^{ee} can be approximated using the electron density. The electron kinetic energy E_k^e is obtained once the wave function is computed. Therefore, with pseudopotentials and the LDA approximation, the total energy E_S and many ground-state properties can be determined for different structural configurations.

It is convenient to evaluate the above energy contributions to E_S using expressions for the potentials and energies in reciprocal space. An early application was the examination of the ground-state properties of Si. **Figure 6** contains the results for calculations of E_S as a function of volume for seven candidate structures. The volume is normalized by the experimental volume of the diamond phase at atmospheric pressure. Therefore, when the volume is unity, this

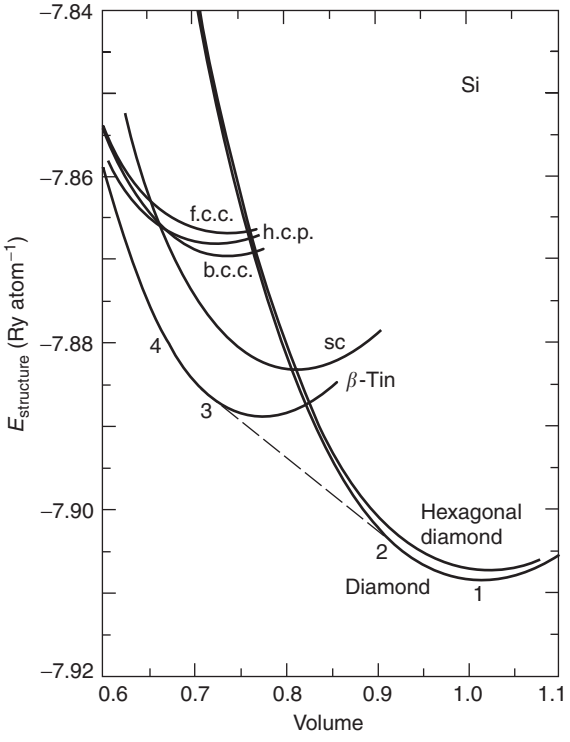


Figure 6 Structural energies for seven structures of Si as a function of volume normalized to the experimental volume at ambient pressure (solid lines). The common tangent represented by the dashed line illustrates the structural transition between the diamond and β-tin structures.

corresponds to ambient pressure and the diamond structure has the lowest energy. As the pressure increases and the volume decreases, it is noted that the hexagonal diamond-structured material is always at a higher energy than the diamond (cubic) structure; therefore, diamond is the more stable phase. At smaller volumes (higher pressures), the β-tin phase has a lower energy than diamond-structured Si, and β-tin is therefore the stable phase. If one views the transition as occurring along the dashed line (**Figure 6**) which is the common tangent between the diamond and β-tin curves, the transition volumes (points 2 and 3) can be evaluated along with the transition pressure which is the slope of the common tangent. These values are all found to be in good agreement with the measured data. In addition, the minimum energy volume for the diamond structure yields a calculation of the equilibrium lattice constant while the curvature of E_S as a function of volume near the minimum energy point yields the bulk modulus or compressibility for Si. In all the calculations, it is assumed that the temperature is zero.

Using the above method, the lattice constants and bulk moduli have been calculated for many solids. Typical accuracies are less than 1% for lattice constants and less than 5% for bulk moduli. Discrepancies between theory and experiment for transition volumes and pressures are normally in the few percent range. In addition, modifications of structures at interfaces and surfaces can also be addressed with this approach. By calculating the changes in energy with movements of specific atoms or by computing Hellmann–Feynman forces, it is possible to determine reconstructions of surfaces and interfaces. In the calculations described above, the only input is the atomic number of the constituent atoms to calculate the pseudopotential and the candidate structure.

Another important application of this method is the determination of vibrational spectra. The method often used within the PWPM is called the "frozen phonon" approach, where the crystal structure is distorted to represent the displaced atoms associated with a particular phonon mode. By comparing the energies of the undistorted and distorted structures or by directly calculating the forces on the atoms, the phonon spectrum can be determined. Here, one needs to input the atomic masses of the constituent atoms as well as their atomic numbers. Again, agreement with experiment is excellent.

Superconductors, Optical Properties, and Novel Materials

A particularly striking success of this approach is its application to superconductivity. The systems considered are expected to be "BCS electron–phonon superconductors." A variation on the frozen phonon method allows a determination of the electron–phonon coupling constants. As described above, the PWPM within the LDA yields ground-state properties with high precision. Since the superconducting transition temperature in the BCS description is extremely sensitive to the strength of the electron–phonon parameter λ and, to a lesser extent, the Coulomb parameter μ^*, it is imperative that the normal-state properties of the material studied be known with high precision. Only a few first-principles calculations of μ^* are available; however, this parameter does not vary a great deal for different materials, and it scales reasonably well with the density of states at the Fermi energy. In contrast, λ can have a much larger variation and therefore the focus for first-principles calculations for superconducting materials has been on the determination of λ.

The enhancement of the effective mass is also related to λ:

$$m^* = m_b(1 + \lambda)$$

where m_b is the band mass calculated without including the electron–phonon interactions. The strong dependence on λ for the superconducting transition temperature, T_c, can be illustrated by equations such

as the McMillan equation which have the generic form

$$T_c = AT_D \exp\left(\frac{1}{\lambda^* - \mu^*}\right)$$

where A is a constant, T_D is the Debye temperature, and $\lambda^* = \lambda/(1+\lambda)$.

The PWPM formulation for the calculation of λ had as its first application a study of high-pressure phases of Si. The results were dramatic because the existence of the two high-pressure solid metallic phases, simple (or primitive) hexagonal (s.h.) and hexagonal close packed (h.c.p.) were predicted by PWPM calculations. In addition to predicting the transition pressures needed to obtain these phases, their electronic, structural, vibrational, and superconducting properties were successfully predicted. The input information was only the atomic number and atomic mass of Si, the candidate structures, and an estimate of μ^*.

Although the PWPM has limited use for highly correlated materials, it is sometimes used as a starting point to obtain the electronic structure before correlation effects, such as a Hubbard U parameter, are considered to estimate the expected changes when correlation is considered. However, when new superconductors are found, it is common to apply the approach described above as a test of whether one is dealing with a conventional superconductor or not. Superconducting doped C_{60} materials, nanotubes, and MgB_2 are good representatives of materials in this category.

The case of MgB_2 serves as an excellent example of what can be done now with the PWPM and Eliashberg theory, which is an extension of the BCS theory. The method yielded the electronic and vibrational properties of MgB_2 and demonstrated that the strongest electron–phonon coupling was associated with a particular phonon mode. By considering phonon anharmonicity and including the anisotropy of the superconducting energy gap which was found to be important when doing the calculation, it was possible to reproduce the experimental data and predict new properties. The results were consistent with the proposal that MgB_2 is a BCS superconductor with electron pairing caused by electron–phonon coupling and that the superconducting energy gap was multivalued. The results for MgB_2 are particularly striking since this PWPM calculation had no experimental input except for the estimate of μ^* and the known structure of MgB_2.

The optical properties of solids are of particular interest when one is describing a method for computing an electronic structure. The interplay between the development of quantum mechanics and the study of atomic optical spectra was extremely important to the evolution of both fields. However, identifying peaks in optical response functions for solids appeared to be a formidable task. Unlike atomic spectra that contained very sharp peaks which could be interpreted in terms of electronic transitions between narrow electronic states, solid-state spectra were broad and featureless.

The challenge remained in essence until the 1960s, and, for the most part, it was the use of the EPM which provided the tools to deal with these spectra, particularly for semiconductors. An important feature is the fact that peaks in optical response functions, such as reflectivity, arise from interband transitions between states where the bands are parallel. For example, if an electron in a valence band n at point $\boldsymbol{k}$ in the Brillouin zone with energy $E_n(\boldsymbol{k})$ is excited to an empty state m with energy $E_m(\boldsymbol{k})$, this requires a photon having energy $\hbar\omega = E_m(\boldsymbol{k}) - E_n(\boldsymbol{k})$. This transition will result in a more prominent structure or "signature" in the reflectivity if the initial occupied and final unoccupied bands are parallel at the point $\boldsymbol{k}$. The condition is $\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k}) = \nabla_{\boldsymbol{k}} E_m(\boldsymbol{k})$, and this is the requirement for a "critical point" to occur in the band structure. It can be shown that this critical point is associated with the structure in the joint density of states between the two bands. This structure, referred to as a Van Hove singularity, in turn, appears in the optical response functions such as the reflectivity or frequency-dependent dielectric function. It was the analysis of these critical points in semiconductor spectra that yielded the energy separations used to fit the pseudopotential form factors for the EPM. A typical result is shown in **Figure 7** where the modulated or derivative reflectivity spectrum of Ge is displayed. Derivative spectroscopy is used to accentuate the Van Hove structures, and the zeros and peaks arise because of the nature of the critical points. As discussed before, the positions of these are related to the energies of the critical points. In **Figure 7**, near the 2 eV region, it is noted that although there is reasonable agreement between theory and experiment for the critical point energies, the magnitudes of the curves differ considerably. This arises because of electron–hole attraction effects which are not included in the calculation. These are discussed later.

Although the transition from the EPM to *ab initio* PWPM applications was fairly straightforward, there were particularly interesting problems associated with the calculations of energy-band separations, the minimum bandgap, optical spectra, and the role of electron–hole effects described above. The most dramatic signal that one could not just apply the LDA and the PWPM to obtain experimental band structures was the fact that researchers using this

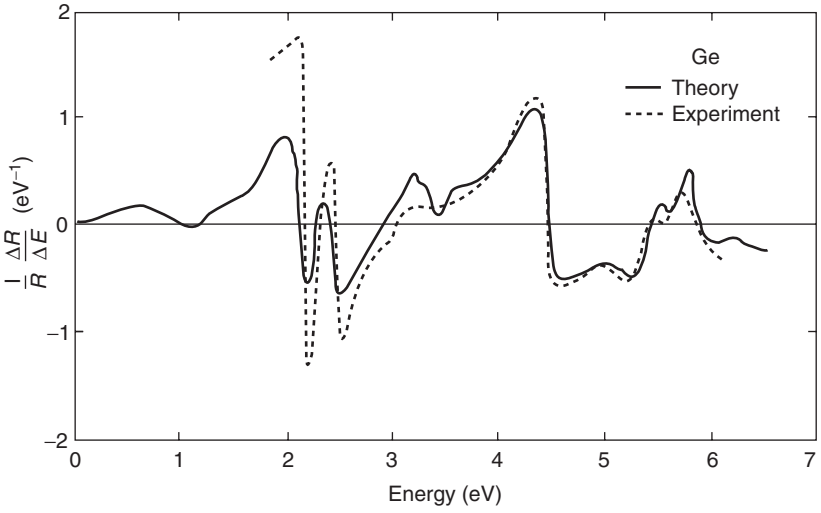


Figure 7 The modulated or derivative reflectivity spectrum for Ge. A comparison is made between the experimental spectrum and the theoretical calculation.

approach were consistently underestimating bandgaps. The specific focus was on Si where the measured and EPM values for the minimum gap were 1.1 eV, while the PWPM–LDA gap was half this value. The EPM was fit to give the correct result, but the *ab initio* method had as its input the LDA, the pseudopotential, and the plane-wave bases. It was soon found that other methods using the LDA gave similar results. Hence, it was concluded that it was the use of the LDA which resulted in calculations which underestimated the bandgap. Since it was emphasized by the inventors of the LDA that, in principle, this approximation should not be applicable for bandgap determination, this result came as no surprise. However, because of the ease of use of the LDA and its success for determining ground-state properties, it was hoped that modifications could be made to standard LDA methods to allow the computation of bandgaps and optical spectra.

A successful practical solution to this dilemma that was often referred to as "the bandgap problem" involved the use of the "GW" method. In this approach, the quasiparticle energies appropriate to excited electron and hole states are calculated starting with LDA calculations. In this scheme, the electron self-energy is computed using the Green's function G and a screened Coulomb potential W. When the resulting self-energy is included in the LDA calculation, the electronic states are renormalized and yield bandgaps in excellent agreement with experiment for a broad class of materials. An analysis of the results of applying the GW method reveals that two important physical features are mainly responsible for the success of the method. First, by using an appropriate dielectric function which includes local field effects, the variations in the charge density arising because of bonding effects are correctly included. Second, the renormalization of the electronic energies because of their interactions with plasmons and electron–hole excitations is included to produce appropriate quasiparticle states produced because of the excitations.

Hence with the inclusion of GW modifications to the LDA, the calculations employing the PWPM can yield bandgaps and band structures in agreement with experiment on a level of accuracy found for the experimentally fit EPM results. In addition, these *ab initio* calculations have been extended further and have addressed the problems of the effects on the optical response functions arising from electron–hole interactions. This problem was raised earlier when discussing the results shown in **Figure 7**. Modifications of spectra of this kind are sometimes referred to as exciton enhancements of oscillator strengths near band edges. However, even when excitons, which are bound states of electron–hole pairs, are not formed, the electron–hole interactions in the excited state can still modify the oscillator strengths and in turn the optical response functions. Applications of the Bethe–Salpeter approach for two-particle interactions to this problem with electron and hole wave functions computed using the PWPM have been successful. At this point, applications of the GW approximation including electron–hole effects have been done for several materials. The computational requirements are somewhat heavy, but the results in all cases tested yield excellent spectra.

A major and important component of condensed matter physics is the research focused on the development of new and novel materials. In the past

two or three decades, new systems have revealed new physical properties, new physics, and important application. Systems such as superlattices, fullerenes, nanotubes, and high-temperature superconductors have broadened the horizons. Often novel materials are complex, structurally or electronically, and they offer new challenges for theoretical methods such as the PWPM. In particular, the number of atoms in a unit cell or supercell is a constraint on this method. Usually, complex materials or configurations of materials can only be modeled with a large number of atoms, and this may require a very large number of plane waves in the basis set. If the electronic structure results in very localized configurations as might be appropriate for complex molecules, again many plane waves may be required. A further and interesting aspect of most novel systems on the nanoscale level is that there are effects of confinement. These effects can be dominant for lower-dimensional or very small structures, and they must be accounted for in the calculational scheme.

A typical example is the case of research on nanocrystals where there have been significant advances in recent years. Here, supercells can be employed along with the PWPM. Other techniques which start out with plane waves have had success. One such approach is to use the EPM to generate Bloch functions for a bulk crystal and then convert these itinerant states into more localized states represented by Wannier functions. If the nanocrystal is then modeled using a potential which is bulk-like inside the nanocrystal and zero outside, the Wannier functions can be used to solve the resulting Hamiltonian appropriate to the nanocrystal. One example of the application of this method is a study of the bandgap of InAs as a function of the nanocrystal radius. The results are in excellent agreement with experiment.

Many applications of the PWPM to C_{60} and nanotubes have appeared in the literature. Electronic structure calculations for C_{60} were used to interpret optical, photoemission, superconducting, and other properties. For nanotubes, the explorations have been even more extensive. For carbon nanotubes, calculations of the properties of semiconductor and metal nanotubes have revealed interesting aspects of their electronic structure for interpreting transport and optical properties. Theoretical predictions of Shottky barrier and heterojunction systems composed of semimetallic and semiconductor nanotubes have been verified. There is even evidence of superconductivity which is consistent with semi-empirical calculations based on PWPM results. No completely *ab initio* calculations of the superconducting properties of fullerenes or nanotubes as in the case of MgB_2 have as yet been done.

A dramatic success of applications of the methods described here to this area is the successful prediction of compound nanotubes. In particular, the prediction of BN nanotubes was verified. In this case, unlike the situation for carbon nanotubes, all BN nanotubes without defects or impurities are semiconducting independent of their "chirality." Theory also suggests that when doped, BN nanotubes will have interesting behavior. An example is the nature of the free electron-like wave function associated with the lowest conduction band for this system. It is predicted that when doped appropriately, the electron density for this state will reside in the center of the BN nanotube. It is expected that this system will resemble a tube with electrons filling its center. Other interesting predictions are that, for BC_2N nanotubes, doping can result in electrons traveling along a spiral path producing effects resembling a nanocoil. For BC_3N, it is expected that each tube will behave as if it were a semiconductor, but a collection of these tubes would be metallic. Although the existence of these predicted compound nanotubes has been established, in general, their transport properties have not been measured, so tests for these unusual predicted properties will have to wait until the appropriate experimental techniques become available.

Conclusion

Summarizing, the history of the use of plane-wave methods for electronic structure calculations has many components. As was discussed, the physical justifications came in different forms, and the development of the techniques required physical interpretation. For example, the major changes in the development of the PWPM were motivated by the desire to answer new questions or to broaden the application of the method. Although the advances in computer hardware had a large effect, the largest impact came from the knowledge gained from doing the calculations. It is important to know what is important, or as Linus Pauling said, "The secret to having good ideas is to have many ideas and the judgment to know which ones should be discarded."

Acknowledgments

This work was supported by National Science Foundation Grant No. DMR00-87088 and by the Director, Office of Science, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, US Department of Energy under Contract No. DE-AC03-76SF00098.

See also: Electronic Structure (Theory): Atoms; Electronic Structure (Theory): Molecules; Electronic Structure Calculations: Scattering Methods; Hartree and Hartree–Fock Methods in Electronic Structure.

PACS: 71.15.Ap; 31.10.+z

Further Reading

Balkanski M and Walis RF (2000) *Semiconductor Physics and Applications*. Oxford: Oxford University Press.

Bassani F and Pastori Parravicin G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon Press.

Cohen ML (1982) Pseudopotentials and total energy calculations. *Physica Scripta*, vol. T1: 5–10.

Cohen ML and Chelikowsky JR (1988) *Electronic Structure and Optical Properties of Semiconductors*. Berlin: Springer-Verlag.

Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.

Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press.

Yu PY and Cardona M (1996) *Fundamentals of Semiconductors*. Berlin: Springer.

Electronic Structure Calculations: Scattering Methods

J S Faulkner, Florida Atlantic University, Boca Raton, FL, USA

Introduction

Walter Kohn and his collaborators derived the density-functional theory (DFT) that justifies the replacement of the complicated Schrödinger equation for N electrons with simpler equations for the one-electron orbitals:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\psi_i(\boldsymbol{r}) = E_i\psi_i(\boldsymbol{r}) \qquad [1]$$

Methods such as the local density approximation (LDA) or the generalized gradient approximation (GGA) are used to generate the effective one-electron potential $V(\boldsymbol{r})$. In all of these approaches, $V(\boldsymbol{r})$ is a functional of the density of the electrons in the system, $\rho(\boldsymbol{r})$, which means that the one-electron equation must be solved self-consistently.

An electronic structure calculation involves finding the eigenvalues and the orbitals from the one-electron equation. Any property of the structure, be it a cluster, a molecule, or a solid, can then be calculated. For example, the charge density is

$$\rho(\boldsymbol{r}) = \sum_{\text{occ.}} \psi_i^*(\boldsymbol{r})\psi_i(\boldsymbol{r}) \qquad [2]$$

and the total energy is

$$E_{\text{total}} = \sum_{\text{occ.}} E_i - \frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' + E_{\text{xc}}[\rho] - \int \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\boldsymbol{r})}\rho(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \qquad [3]$$

The sums are over all occupied states including spin, and the exchange correlation functional $E_{\text{xc}}[\rho]$ is determined by the particular method used to calculate the effective one-electron potential.

The standard approach to solving the one-electron Schrödinger equation is the Rayleigh–Ritz variational method. The one-electron orbitals are expanded in a complete set of trial functions, and the Schrödinger equation is converted into a set of coupled equations for the expansion coefficients. A problem with the variational method is that all of the core states of the atoms have to be included in the calculation so that the valence-electron orbitals are orthogonal to them. Plane waves are convenient basis functions. The difficulty that these functions cannot be made orthogonal to the core states is circumvented by converting the actual potential into a pseudopotential that has no core states. The pseudopotential method works well for covalent crystals, for example, diamond and silicon, and for simple metals. It is difficult to apply to solids containing transition metals or rare earths because they have too many electrons outside the core, although a class of potentials called soft pseudopotentials has had some success in that area. A completely different method, which was also initiated by Lord Rayleigh, is based on the scattering theory. The orbitals corresponding to a given energy are automatically orthogonal to those for any other energy, so the core levels pose no problem. Another advantage to the multiple-scattering method is that it is the most convenient way to obtain the one-electron Green's function for a solid, which is discussed elsewhere in this encyclopedia.

Scattering Theory

In the Dirac notation, eqn [1] can be written as an inhomogeneous equation, $(E - H_0)|\psi\rangle = V|\psi\rangle$, where H_0 is the kinetic energy operator. The solution of the homogeneous equation is $\varphi(\boldsymbol{r}) = \langle \boldsymbol{r}|\varphi\rangle$, where $(E - H_0)|\varphi\rangle = 0$. A formal solution of the

inhomogeneous equation is the sum of the solution of the homogeneous equation with a particular solution of the inhomogeneous equation $|\psi\rangle = |\varphi\rangle + G_{0+}V|\psi\rangle$, where $G_{0+} = \lim_{\varepsilon\to 0}(E - H_0 + i\varepsilon)^{-1}$. This is the Lippmann–Schwinger equation, and it is the starting point for modern treatments of the scattering theory. The infinitesimal ε avoids the singularity at $E - H_0 = 0$. Letting it approach zero through positive values guarantees the boundary conditions, as will be seen later. Using the identity, $(A-B)^{-1} = A^{-1} + A^{-1}B\,(A-B)^{-1}$, the Lippmann–Schwinger equation can be written in another way,

$$|\psi\rangle = (1 + G_{0+}T)|\varphi\rangle \qquad [4]$$

where the t-matrix is defined by

$$T = V + VG_{0+}V + VG_{0+}VG_{0+}V + \cdots \qquad [5]$$

Suppose the potential V is the sum of N potentials, $V = \sum_{i=1}^{N} v_i$. The physical meaning of this is that it describes the interaction of the electron with a cluster of atoms having nuclei located at positions $\boldsymbol{R}_i$. It is assumed that, in the position representation, the potentials are localized in space and do not overlap, so that $v_i(\boldsymbol{r})v_j(\boldsymbol{r}) = 0$ for $i \neq j$ and all $\boldsymbol{r}$. Define an operator Q_i so that T can be written as a sum

$$T = \sum_{i=1}^{N} Q_i$$

Inserting the expressions for V and T into the definition of T leads to

$$\begin{aligned}\sum_i Q_i &= \sum_i v_i\left(1 + G_{0+}\sum_j Q_j\right)\\ &= \sum_i\left[v_i G_{0+}Q_i + v_i\left(1 + G_{0+}\sum_{j\neq i} Q_j\right)\right]\end{aligned}$$

so

$$Q_i = t_i\left(1 + G_{0+}\sum_{j\neq i} Q_j\right)$$

where $t_i = (1 - v_i G_{0+})^{-1} v_i$. Iterating this equation leads to

$$\begin{aligned}T = &\sum_i t_i + \sum_i t_i G_{0+}\sum_{j\neq i} t_j\\ &+ \sum_i t_i G_{0+}\sum_{j\neq i} t_j G_{0+}\sum_{k\neq j} t_k + \cdots\end{aligned}$$

In terms of the operators Q_i, eqn [4] becomes

$$\begin{aligned}|\psi\rangle &= \left(1 + G_{0+}\sum_{j\neq i} Q_j\right)|\varphi\rangle + G_{0+}Q_i|\varphi\rangle\\ &= |\varphi_i^{\text{in}}\rangle + |\varphi_i^{\text{out}}\rangle\end{aligned}$$

The site i that is chosen to focus on can be any of the sites in the cluster. The incoming wave on this site is taken to be

$$|\varphi_i^{\text{in}}\rangle = \left(1 + G_{0+}\sum_{j\neq i} Q_j\right)|\varphi\rangle = |\varphi\rangle + \sum_{j\neq i}|\varphi_j^{\text{out}}\rangle$$

and the outgoing wave from the site is defined as

$$|\varphi_i^{\text{out}}\rangle = G_{0+}Q_i|\varphi\rangle = G_{0+}t_i|\varphi_i^{\text{in}}\rangle$$

In other words, the first of these equations states that the incoming wave on site i is the sum of the wave coming onto the cluster of scatterers plus the waves scattered out from all of the other scatterers. At first sight, this seems obvious, but one might wonder about the effect of intervening scatterers. It is fundamentally a statement about the superposition of solutions of a linear differential equation. The second equation states that the outgoing wave from a site is obtained from the incoming wave by the solution of a simple scattering problem, which is obvious. To sum up, the fundamental equations of the multiple-scattering theory are

$$\begin{aligned}|\psi\rangle &= |\varphi_i^{\text{in}}\rangle + |\varphi_i^{\text{out}}\rangle\\ |\varphi_i^{\text{in}}\rangle &= |\varphi\rangle + \sum_{j\neq i}^{N}|\varphi_j^{\text{out}}\rangle\\ |\varphi_i^{\text{out}}\rangle &= G_{0+}t_i|\varphi_i^{\text{in}}\rangle\end{aligned} \qquad [6]$$

Using the lowest-level approximation to the t-matrix, $T = \sum_i t_i$, the Lippmann–Schwinger equation can be written as

$$\begin{aligned}\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \varphi_{\boldsymbol{k}}(\boldsymbol{r}') + \sum_{i=1}^{N} e^{i\boldsymbol{k}\cdot\boldsymbol{R}_i}\int G_{0+}(\boldsymbol{r}_i - \boldsymbol{r}_i')\\ \times\ t_i(\boldsymbol{r}_i', \boldsymbol{r}_i'')\varphi_{\boldsymbol{k}}(\boldsymbol{r}_i'')\,d\boldsymbol{r}_i'\,d\boldsymbol{r}_i''\end{aligned}$$

where $\boldsymbol{r}_i = \boldsymbol{r} - \boldsymbol{R}_i$, and $\boldsymbol{R}_i$ is the position of the ith nucleus. The function $t_i(\boldsymbol{r}_i', \boldsymbol{r}_i'')$ is zero outside the range of the potential $v_i(\boldsymbol{r}_i)$, which is a relatively small region surrounding $\boldsymbol{R}_i$. The incoming wave is taken to be a plane wave, which may be written as

$$\begin{aligned}\phi_{\boldsymbol{k}}(\boldsymbol{r}) &= (2\pi)^{-3/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\\ &= e^{i\boldsymbol{k}\cdot\boldsymbol{R}_i}(2\pi)^{-3/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\\ &= e^{i\boldsymbol{k}\cdot\boldsymbol{R}_i}\phi_{\boldsymbol{k}}(\boldsymbol{r}_i)\end{aligned}$$

Assuming that $r_i \gg r_i'$, the Green's function can be recast into the form

$$\begin{aligned}G_{0+}(\boldsymbol{r} - \boldsymbol{r}') = G_{0+}(\boldsymbol{r}_i - \boldsymbol{r}_i') &= -\frac{1}{4\pi}\frac{e^{ik|\boldsymbol{r}_i - \boldsymbol{r}_i'|}}{|\boldsymbol{r}_i - \boldsymbol{r}_i'|}\\ &\approx -\frac{1}{4\pi}\frac{e^{ikr_i}}{r_i}e^{-i\boldsymbol{k}_i'\cdot\boldsymbol{r}_i'}\end{aligned}$$

where $\boldsymbol{k}_i' = (\boldsymbol{r}_i/r_i)k$. The solution is then

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}}\left[e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \sum_{i=1}^{N}\frac{e^{ikr_i}}{r_i}f^i(\boldsymbol{k}_i', \boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{R}_i}\right] \quad [7]$$

where

$$\begin{aligned} f^i(\boldsymbol{k}_i', \boldsymbol{k}) &= -\frac{1}{4\pi}\int e^{-i\boldsymbol{k}_i'\cdot\boldsymbol{r}_i'}t^i(\boldsymbol{r}_i, \boldsymbol{r}_i')e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\,d\boldsymbol{r}_i\,d\boldsymbol{r}_i' \\ &= -2\pi^2\langle \boldsymbol{k}_i'|t^i|\boldsymbol{k}\rangle \end{aligned} \quad [8]$$

This form of the formula holds for positions $\boldsymbol{r}$ outside the range of any individual potential function. If the solution is to be evaluated only for points outside the cluster of atoms, then $r \gg R_i$ for any i and

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}}\left[e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{e^{ikr}}{r}\sum_{i=1}^{N} f^i(\boldsymbol{k}', \boldsymbol{k})e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i}\right] \quad [9]$$

where $\boldsymbol{k}' = (\boldsymbol{r}/r)k$.

The approximations used to derive eqn [9] are obviously very good for a standard X-ray, neutron, or electron diffractometer, where the sample size is of the order of millimeters, while the counter is a sizable fraction of a meter away. If the sample is a monatomic crystal, the scatterers are all identical atoms and they are distributed in a periodic array. It follows that the $f^i(\boldsymbol{k}, \boldsymbol{k}')$ are all the same, and $\sum_{i=1}^{N} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i} = N\sum_n \delta(\boldsymbol{k} - \boldsymbol{k}' + \boldsymbol{K}_n)$, where $\boldsymbol{K}_n$ are the reciprocal lattice vectors, which explains the occurrence of Bragg peaks. The application of the multiple-scattering equations to explain diffraction was pioneered by P P Ewald, and the Born approximation:

$$f^i(\boldsymbol{k}', \boldsymbol{k}) = -\frac{1}{4\pi}\int v_i(\boldsymbol{r}_i)\,e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_i}\,d\boldsymbol{r}_i$$

is adequate to describe the weak scattering from the atoms. More complicated effects can be treated by including higher-order terms in the t-matrix.

These multiple scatterings can be used to explain many other scattering experiments, such as the scattering of electrons from molecules, low-energy electron diffraction (LEED) from surfaces of solids, and extended X-ray absorption fine structure (EXAFS). They are also used to describe the propagation of radio waves in inhomogeneous media and the scattering of sound waves from obstacles in water or air. In 1892, Lord Rayleigh introduced the concept that evolved into the multiple-scattering theory in a treatment of the electrical conductivity of a solid with a periodic array of rods embedded in it. Jan Korringa made the conceptual leap that the multiple-scattering equations can also be used to calculate the stationary states of electrons in a solid. The computer power was not adequate to carry out useful calculations with Korringa's equations when they were published in 1947, so they did not get the attention they should have. In 1954, Walter Kohn and Norman Rostoker rederived Korringa's equations using the Kohn variational method. The resulting energy band theory is called the Korringa–Kohn–Rostoker (KKR) method.

Electronic States in Periodic Solids

The picture of a solid that is used in theoretical treatments of the stationary electronic states is an infinite array of atoms that fills all space. This makes it possible to ignore surface states, which are not of interest at this stage. The incoming wave on the cluster $|\varphi\rangle$ may be set equal to zero in the multiple-scattering equations, so eqn [7b] is replaced by

$$|\varphi_i^{\text{in}}\rangle = \sum_{j\neq i}^{\infty}|\varphi_j^{\text{out}}\rangle \quad [10]$$

A simple description of the resulting stationary states is that the electrons scatter from one atom to the other *ad infinitum*, so there is no incoming or outgoing wave.

Since the $v_i(\boldsymbol{r})$ are bounded in space and do not overlap, there is an interstitial region between them within which the potential is a constant, usually taken to be zero, $V(\boldsymbol{r}) = 0$. In the interstitial region, eqn [1] becomes

$$\lfloor \nabla^2 + \alpha^2 \rfloor\psi_i(\boldsymbol{r}) = 0 \quad [11]$$

where $\alpha = \sqrt{2mE}/\hbar$. The solutions of this equation are $Y_{lm}(\vartheta, \varphi)j_l(\alpha r)$ and $Y_{lm}(\vartheta, \varphi)n_l(\alpha r)$, where $Y_{lm}(\vartheta, \varphi)$ is the spherical harmonic, $j_l(\alpha r)$ is a Bessel function, and $n_l(\alpha r)$ is a Neumann function. The incoming wave on site i can thus be written in the position representation

$$\langle \boldsymbol{r}|\varphi_i^{\text{in}}\rangle = \varphi_i^{\text{in}}(\boldsymbol{r}_i) = \sum_{l,m} Y_{lm}(\hat{\boldsymbol{r}}_i)j_l(\alpha r_i)d_{lm}^i$$

where the d_{lm}^i are undetermined coefficients and $\boldsymbol{r}_i = \boldsymbol{r} - \boldsymbol{R}_i$. The unit vector $\hat{\boldsymbol{r}}_i$ defines the angles ϑ and φ in a coordinate system centered on the site of the ith nucleus, $\boldsymbol{R}_i$. The Green's function may be expanded in the interstitial region:

$$\begin{aligned} G_{0+}(\boldsymbol{r} - \boldsymbol{r}') &= -\frac{1}{4\pi}\frac{e^{i\alpha|\boldsymbol{r}_i - \boldsymbol{r}_i'|}}{|\boldsymbol{r}_i - \boldsymbol{r}_i'|} \\ &= -i\alpha\sum_{l',m'} Y_{l',m'}(\hat{\boldsymbol{r}}_i)h_{l'}^+(\alpha r_i)j_{l'}(\alpha r_i')Y_{l',m'}^*(\hat{\boldsymbol{r}}_i') \end{aligned}$$

where $h_l^+ = j_l + in_l$ is the outgoing Hankel function. It follows that eqn [6a] can be written in the position representation:

$$\begin{aligned} \psi(\boldsymbol{r}) = \sum_{l,m}\Bigg[& Y_{lm}(\hat{\boldsymbol{r}}_i)j_l(\alpha r_i) \\ & - i\alpha\sum_{l',m'} Y_{l'm'}(\hat{\boldsymbol{r}}_l)h_{l'}^+(\alpha r_i)t_{l'm',lm}^i(E)\Bigg]\, d_{im}^i \end{aligned} \quad [12]$$

where

$$t^i_{l'm',lm} = \int\int j_{l'}(\alpha r'_i) Y^*_{l',m'}(\hat{\boldsymbol{r}}'_i) t^i(\boldsymbol{r}'_i, \boldsymbol{r}''_i) \times Y_{l,m}(\hat{\boldsymbol{r}}''_i) j_l(\alpha r''_i)\, \mathrm{d}\boldsymbol{r}'_i\, \mathrm{d}\boldsymbol{r}''_i$$

and the outgoing wave can now be identified as

$$\langle \boldsymbol{r} | \varphi^{\text{out}}_i \rangle = \varphi^{\text{out}}_i(\boldsymbol{r}_i) = -\mathrm{i}\alpha \sum_{l,m}\sum_{l',m'} Y_{l',m'}(\hat{\boldsymbol{r}}_i) h^+_{l'}(\alpha r_i)\, t^i_{l'm',lm}(E) d^i_{lm}$$

The content of eqn [10] is that the incoming wave on site i is the sum of the outgoing waves from all of the other sites. The incoming wave is a sum of Bessel functions that are finite for all $\boldsymbol{r}$. The outgoing wave from the site j is infinite at the nuclear position $\boldsymbol{R}_j$, but is finite in the neighborhood of the ith site. The equality can be demonstrated with the help of the mathematical identity:

$$-\mathrm{i}\alpha Y_{l'm'}(\boldsymbol{r}_j) h^+_{l'}(r_j) = \sum_{l,m} Y_{lm}(\boldsymbol{r}_i) j_l(\alpha r_i) g_{lm,l'm'}(E, \boldsymbol{R}_{ij})$$

where the vector $\boldsymbol{R}_{ij}$ is the position of the ith scatterer relative to the jth, $\boldsymbol{R}_{ij} = \boldsymbol{R}_i - \boldsymbol{R}_j$, and

$$g_{lm,l'm'}(E, \boldsymbol{R}_{ij}) = -4\pi i \alpha i^{(l-l')} \sum_{l'',m''} i^{-l''} C^{l''m''}_{lm,l'm'} \times h^+_{l''}(\alpha R_{ij}) Y^*_{l''m''}(\boldsymbol{R}_{ij})$$

The $C^{l''m''}_{lm,l'm'}$ are called Gaunt factors

$$C^{l''m''}_{lm,l'm'} = \int\int Y_{l''m''}(\theta,\phi) Y^*_{lm}(\theta,\phi) Y_{l'm'}(\theta,\phi) \times \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi$$

Equation [10] thus leads to a set of homogeneous simultaneous equations that determines the unknown coefficients d^i_{lm}

$$\sum_{j,l''m''} \Big[\delta_{ij}\delta_{lm,l''m''} - (1-\delta_{ij}) \times \sum_{l'm'} g_{lm,l'm'}(E, \boldsymbol{R}_{ij}) t^j_{l'm',l''m''}(E)\Big] d^j_{l''m''} = 0 \quad [13]$$

The eigenvalues of the one-electron equation, eqn [1] for the potential $V(\boldsymbol{r}) = \sum_{i=1}^{\infty} v_i(\boldsymbol{r})$, are the energies for which the determinant of the coefficients in the square bracket in eqn [13] is zero. It is more convenient to deal with a symmetric matrix, and this can be done by defining the coefficients $c^i_{lm} = \sum_{l'm'} t^i_{lm,l'm'}(E) d^i_{l'm'}$. These coefficients satisfy the set of linear equations

$$\sum_{j,l'm'} M^{ij}_{lm,l'm'} c^j_{l'm'} = 0 \quad [14]$$

with the elements of the matrix $\boldsymbol{M}$ being

$$M^{ij}_{lm,l'm'} = m^i_{lm,l'm'}\delta_{ij} - (1-\delta_{ij}) g_{lm,l'm'}(E, \boldsymbol{R}_{ij}) \quad [15]$$

and the $m^i_{lm,l'm'}$ are the elements of the inverse of the t-matrix.

Periodic Solids, One Atom Per Unit Cell

Equations [14] and [15] can, in principle, be used to calculate the electronic states for any solid, but they require that the zeros of an infinitely large matrix must be found. The problem is simplified considerably for periodic solids in which all of the potentials $v_i(\boldsymbol{r}_i)$ are the same, and the nuclear positions $\boldsymbol{R}_i$ form a periodic array. Bloch's theorem holds for such a system because the potential $V(\boldsymbol{r})$ is invariant under the operations of the lattice translation group. Bloch's theorem states that the eigenfunctions of eqn [1] can be indexed by a quantum number $\boldsymbol{k}$, and $\psi_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}_i) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\psi_{\boldsymbol{k}}(\boldsymbol{r})$. This implies that the coefficients in eqn [14] can be written in terms of coefficients that do not depend on the site, $c^j_{l'm'} = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_j} c_{l'm'}(\boldsymbol{k})$. Premultiplying eqn [14] by $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}$ and summing leads to

$$\sum_{j,l'm'} M_{lm,l'm'}(E, \boldsymbol{k}) c_{l'm'}(\boldsymbol{k}) = 0 \quad [16]$$

where

$$M_{lm,l'm'}(E, \boldsymbol{k}) = m_{lm,l'm'}(E) - A_{lm,l'm'}(E, \boldsymbol{k}) \quad [17]$$

and

$$A_{lm,l'm'}(E, \boldsymbol{k}) = \sum_{j\neq i} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{ij}} g_{lm,l'm'}(E, \boldsymbol{R}_{ij}) \quad [18]$$

These are the KKR band theory equations for a periodic crystal with one atom in the unit cell. The Ewald summation process can be used to calculate the structure constants $A_{lm,l'm'}(E, \boldsymbol{k})$. It is easy to prove that the sum in eqn [18] does not depend on i because of the periodicity of the system.

The energy eigenvalues of the periodic solid for a particular $\boldsymbol{k}$, $E_b(\boldsymbol{k})$, are the roots of the equation $\det \boldsymbol{M}(E, \boldsymbol{k}) = 0$. The eigenfunctions are found by solving eqn [16] with $E = E_b(\boldsymbol{k})$. The dimension of these matrix equations is technically infinite, but by ignoring all contributions that correspond to an angular momentum quantum number l greater than $l_{\max}$, they have dimension $(l_{\max}+1)^2$. The justification for this approximation is that the matrix elements of the t-matrix $t_{lm,l'm'}$ are very small when l and l' are greater than $l_{\max}$, and the elements of the inverse matrix $m_{lm,l'm'}$ are very large. The choice of $l_{\max}$ depends on the specific calculation, but $l_{\max} = 3$ can be used in many applications.

Periodic Solids, Several Atoms Per Unit Cell

Periodic solids with several atoms in a unit cell can be treated using the multiple-scattering equations in the form of eqns [14] and [15], but it is convenient to change the notation. The origin of the coordinate system in the mth unit cell is $\boldsymbol{R}_m$, and the positions of the n_c particles in the unit cell are $\boldsymbol{R}_m + \boldsymbol{a}_i$. It is possible to write

$$\sum_{nj,L'} M^{mi,nj}_{L,L'} c^{nj}_{L'} = 0 \qquad [19]$$

with the elements of the matrix $\boldsymbol{M}$ being

$$M^{mi,nj}_{L,L'} = m^{i}_{L,L'}\delta_{mn}\delta_{ij} - (1-\delta_{mn}\delta_{ij})g_{L,L'}(E, \boldsymbol{R}_{mn} + \boldsymbol{a}_{ij}) \qquad [20]$$

where $\boldsymbol{R}_{mn} = \boldsymbol{R}_m - \boldsymbol{R}_n$ and $\boldsymbol{a}_{ij} = \boldsymbol{a}_i - \boldsymbol{a}_j$. In order to simplify the appearance of the equations in the face of the increasing number of indices, the abbreviation L is used for the pair of angular momentum quantum numbers l, m. This crystal is invariant under translations of entire unit cells, not atoms, so the relation between the coefficients due to Bloch's theorem is $c^{nj}_{L'} = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_n} c^{j}_{L'}(\boldsymbol{k})$. Premultiplying eqn [14] by $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_m}$ and summing leads to

$$\sum_{j=1}^{n_c}\sum_{L'} M^{ij}_{L,L'}(E,\boldsymbol{k}) c^{j}_{L'}(\boldsymbol{k}) = 0 \qquad [21]$$

where

$$M^{ij}_{L,L'}(E,\boldsymbol{k}) = m^{i}_{L,L'}(E)\delta_{ij} - A^{ij}_{L,L'}(E,\boldsymbol{k}) \qquad [22]$$

and

$$A^{ij}_{L,L'}(E,\boldsymbol{k}) = (1-\delta_{ij})g_{L,L'}(E,\boldsymbol{a}_{ij}) + \sum_{m\neq n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{mn}} g_{L,L'}(E, \boldsymbol{R}_{mn} + \boldsymbol{a}_{ij}) \qquad [23]$$

The structure constants, $A^{ij}_{L,L'}(E,\boldsymbol{k})$, are somewhat more complex than those for the monatomic case, but they can also be evaluated with the help of the Ewald summation process. If the contributions to these equations that correspond to angular momentum quantum numbers greater than l_{max} are ignored, the dimension of the matrix equations is $(l_{max}+1)^2 n_c$.

Madelung Energy

The atoms in a crystal with more than one atom per unit cell are typically charged. The exact value for the net charge q_i on the atom at the position $\boldsymbol{R}_m + \boldsymbol{a}_i$ is somewhat arbitrary, depending on the manner in which space is divided up. The sum of the charges in the unit cell must satisfy the neutrality condition, $\sum_{i=1}^{n_c} q_i = 0$. The one-electron potential at the site i is shifted by the Madelung potential V_i, and the proper calculation of these electrostatic shifts is an important part of the band theory calculation. In principle, V_i can be calculated from the usual Coulomb equation from electrostatics

$$V_i = \sum_{j\neq i=1}^{n_c} \frac{q_j}{|\boldsymbol{a}_i - \boldsymbol{a}_j|} + \sum_n \sum_{j=1}^{n_c} \frac{q_j}{|\boldsymbol{a}_i - \boldsymbol{a}_j + \boldsymbol{R}_n|}$$

but this sum is notoriously difficult to converge. In practice, the potential is calculated from the expression

$$V_i = \sum_{j=1}^{n_c} \alpha^{ij}_{M} q_j \qquad [24]$$

where the elements of the Madelung matrix, α^{ij}_{M}, are obtained by a sophisticated mathematical method, such as the Ewald summation process or the fast multipole method.

Electronic States in Nonperiodic Solids Using Supercells

It is often desirable to calculate the electronic structure of systems of atoms that are not arranged periodically. Examples are solids containing impurities or structural defects, liquid metals, amorphous solids, or solid solution alloys. An approach to these problems might be to construct a cluster of N atoms and use the multiple-scattering equations in the form of eqns [14] and [15]. This will not work for a number of reasons. The ratio of near-surface atoms to bulk atoms is so high that surface effects dominate the calculation. Because of the slow convergence of the Coulomb sums, the Madelung potentials in the cluster are quite different from those in a solid.

A way around these difficulties is to reproduce the N-atom cluster periodically to fill all space. This now looks like a periodic solid with N atoms per unit cell, and eqns [21]–[23] can be used to find the electronic structure. Such supercell calculations have been used extensively to treat solids without symmetry.

The eigenvalues $E_b(\boldsymbol{k})$ of the matrix $\boldsymbol{M}(E, \boldsymbol{k})$ from eqn [22] have a k-dependence due to the forced periodicity of the supercell. The error that is introduced is of the order of the bandwidth, $\delta E_b(\boldsymbol{k})$, which can be estimated by taking the difference between the eigenvalues at the Brillouin zone boundary and the center of the zone:

$$\delta E_b(\boldsymbol{k}) = E_b\left(\frac{\pi}{D}\right) - E_b(0) \approx \frac{\partial E_b}{\partial k}\frac{\pi}{D}$$

The way to make the error small is to make D large, where D is a dimension of the supercell and is proportional to $\sqrt[3]{N}$. The algorithms for finding the eigenvalues of any matrix require a number of floating point operations that increases as the cube of the dimension of the matrix. The dimension of the matrix

$\boldsymbol{M}(E, \boldsymbol{k})$ for the supercell is $(l_{\max}+1)^2 N$, so the computer time required to diagonalize the matrix increases as the cube of the dimension, N^3 or D^9. There is thus a competition between the goals of reducing the error in the eigenvalues $\delta E_b(\boldsymbol{k})$ and minimizing the computer time required for the calculation. A solution to this dilemma is to develop order-N methods for which the number of floating point operations increases as N instead of N^3. Two order-N methods based on the multiple-scattering approach are discussed elsewhere in this encyclopedia.

Surface Electronic States

One of the reasons that the solid was assumed to fill all space in the preceding two sections is to focus on the bulk states. There are electronic states in a real solid that are specifically associated with the surface. An understanding of the surface states is important in applications, such as chemical catalysis. The magnetic properties of atoms in the surface region are frequently different from the atoms in the bulk.

The multiple-scattering method can be used to calculate the electronic states associated with surfaces. Consider planes of atoms stretching to infinity in directions perpendicular to the z-direction. The planes that are well separated from the terminal plane at the surface will have the same arrangement of atoms and the same separation from each other as they have in the bulk solid. Closer to the surface, the distance between planes is not the same as the bulk, and the arrangement of atoms in the planes may reconstruct into another arrangement. In any case, there is periodicity in the xy-plane, and hence a 2D Bloch's theorem for each plane.

Perhaps the easiest model to visualize is a slab geometry that starts with a central plane of atoms with the same arrangement as the bulk. Similar planes are placed on either side of the central one, building a slab with symmetry about the central plane and containing an odd number of planes. These slabs are then periodically reproduced in the z-direction, separated by a large distance that contains no atoms. Making use of the symmetry in the xy-planes, the system can be modeled with long, narrow supercells that are a few lattice constants wide in the planar directions but many angstroms long in the z-direction. Equations [21]–[23] can be used to calculate the electronic structure, but the dependence on k_z can be ignored because of the length of the supercell.

After the electronic states have been calculated self-consistently for a given configuration, the atomic positions in the outer planes of atoms are allowed to relax and the calculations are repeated. The Hellmann–Feynman theorem is useful in this step, which leads to a prediction of the spacing between planes and possible reconstruction of surface atoms. Experience has shown that, at least in metals, only the first few planes nearest the surface differ significantly from the bulk.

The multiple scattering equations have been used in other ways to treat surfaces. Instead of periodically reproducing the slabs, a mirror-like scatterer with a planar surface is put outside of the slab on either side to reflect the electrons back in. Another variation is to extend the slab to infinity in the minus z-direction. This is called the layer Korringa, Kohn, Rostoker (LKKR) method, and it is more like the surface on a real solid.

Approximations and Extensions

The Muffin–Tin Approximation

Until recently, the one-electron potentials used in multiple-scattering calculations were approximated by muffin–tin (MT) potentials that are spherically symmetric within bounding spheres centered on the nuclei and are constant outside them. The spheres do not overlap, and a 2D drawing of the function is similar to a pan that is used to cook muffins.

The inverse of the scattering matrix can be written as

$$m^i_{L,L'} = \alpha \sum_{L'} c^i_{LL''} s^i_{L''L'} - i\alpha$$

where the cosine and sine matrices are defined by

$$c^i_{LL'} = \delta_{LL'} - \alpha \int n_l(\alpha r) Y_L(\boldsymbol{r}) v_i(\boldsymbol{r}) \Phi^i_{L'}(E, \boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

$$s^i_{LL'} = \alpha \int j_l(\alpha r) Y_L(\boldsymbol{r}) v_i(\boldsymbol{r}) \Phi^i_{L'}(E, \boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

In these equations, $\Phi^i_L(E, \boldsymbol{r})$ is the solution of a one-electron equation containing the single potential $v_i(\boldsymbol{r})$ that approaches the limit $\lim_{r\to 0} \Phi^i_L(E, \boldsymbol{r}) = j_l(\alpha r)\, Y_L(\boldsymbol{r})$. These equations are laborious to solve for nonspherical potentials because both $v_i(\boldsymbol{r})$ and $\Phi^i_L(E, \boldsymbol{r})$ have to be expanded in spherical harmonics. For the spherical potentials used in the MT approximation, they simplify to

$$c^i_{LL'} = -\alpha S_i^2 \delta_{LL'} \left[n_l(\alpha r) \frac{\mathrm{d}\Phi^i_l(E, r)}{\mathrm{d}r} - \frac{\mathrm{d}n_l(\alpha r)}{\mathrm{d}r} \Phi^i_l(E, r) \right]_{r=S_i}$$

$$c^i_{LL'} = \alpha S_i^2 \delta_{LL'} \left[j_l(\alpha r) \frac{\mathrm{d}\Phi^i_l(E, r)}{\mathrm{d}r} - \frac{\mathrm{d}j_l(\alpha r)}{\mathrm{d}r} \Phi^i_l(E, r) \right]_{r=S_i}$$

where S_i is the radius of the MT sphere.

The MT approximation works very well for electronic structure calculations on metals in close-packed structures. For nonperiodic structures, it is frequently desirable to calculate the forces on the atoms using the

Hellmann–Feynman theorem. Any shape approximation on the potential will cause significant errors in such calculations. The MT approximation is also not acceptable for loosely packed covalently bonded crystals. With some care, the multiple-scattering equations written in the preceding sections can be used for any nonoverlapping potential.

An atomic sphere approximation (ASA) has been proposed to simplify multiple-scattering calculations. The potential is forced into a form that is spherically symmetrical for r less than the radius of a sphere that has the same volume as the Wigner–Seitz cell for the atom. The ASA potentials overlap, and this violates the fundamental assumptions of the multiple-scattering theory.

Relativity

The multiple-scattering equations have been written in the nonrelativistic form because they are easier to visualize. The technical problems involved with a multiple-scattering theory based on the Dirac equation rather than the Schrödinger equation have been resolved, and all of the statements made above apply to the relativistic theory. Spin polarized versions of these equations have been used to discuss magnetism in metals, alloys, and surfaces.

Linearized Versions of the Multiple-Scattering Equations

The KKR equations are time-consuming to solve because all of the matrix elements depend on the energy. The eigenvalues and eigenvectors can be found much more easily for equations of the form that is obtained from a linear variational method in which the matrix $\boldsymbol{M}$ in eqn [21] is replaced by one with the elements:

$$M^{ij}_{L,L'}(E,\boldsymbol{k}) = E\delta_{ij}\delta_{LL'} - D^{ij}_{L,L'}(\boldsymbol{k})$$

because it is only necessary to diagonalize the matrix $\boldsymbol{D}$. The linear combination of MT orbitals (LMTO) approach achieves this by using trial functions suggested by the multiple-scattering theory in a Rayleigh–Ritz variational calculation.

When the potentials fill all space, the interstitial volume goes to zero. Thus, the value V_{int} of $V(\boldsymbol{r})$ when $\boldsymbol{r}$ is in the interstitial region, chosen to be zero in the preceding equations, becomes irrelevant. This fact can be used to derive a version of the multiple-scattering equations that has the eigenvalue form of the preceding equation.

The freedom to choose any value for V_{int} for space filling potentials can be used in another way to simplify the multiple-scattering equations. If V_{int} is chosen much larger than the average value of $v_i(\boldsymbol{r})$ in the region of the Wigner–Seitz cell farthest from the nucleus, the equations describe a barrier that the electron must tunnel through to propagate from one site to another. The sum that defines the structure constants in eqn [23] converges after a small number of nearest-neighbor shells. The resulting screened-structure-constant (SSC) approximation leads to multiple-scattering equations that are similar to tight-binding equations.

See also: Density-Functional Theory; Disordered Solids and Glasses, Electronic Structure of; Intermetallic Compounds, Electronic States of; Magnetoresistance Techniques Applied to Fermi Surfaces; Metals and Alloys, Impurity and Defect States in; Surfaces and Interfaces, Electronic Structure of.

PACS: 71.15.–m; 71.15.Ap; 71.15.Mb; 71.20.Be; 73.20.At; 73.21.–b

Further Reading

Antonios G and Butler WH (1999) *Multiple Scattering in Solids, Graduate Texts in Contemporary Physics*. Berlin: Springer.

Newton RG (2002) *Scattering Theory of Waves and Particles*,, 2nd edn. New York: Dover.

Electron–Phonon Interactions and the Response of Polarons

J T Devreese, Universiteit Antwerpen (UA), Antwerpen, Belgium

Developments in Polaron Physics

A conduction electron (or hole) together with its self-induced polarization in an ionic crystal or in a polar semiconductor forms a quasiparticle, which is called a polaron. The polaron concept is discussed in an article by E I Rashba in the present encyclopedia (for references about the standard polaron theory, see the "Further reading" section). A polaron is characterized by its binding (or self-) energy E_0, an effective mass m^* and by its characteristic response to external electric and magnetic fields (e.g., DC mobility and optical absorption coefficient).

If the spatial extension of a polaron is large compared to the lattice parameter of the solid, the latter can be treated as a polarizable continuum. This is the case of a "large polaron." An electron or a hole trapped by its self-induced atomic (ionic) displacement field in a region with linear dimension of the order of the lattice constant is called a "small polaron."

An all-coupling polaron theory was developed by Feynman using his path-integral formalism. He first studied the self-energy E_0 and the effective mass m^*, and later the mobility of polarons. Subsequently, the path-integral approach to the polaron problem was generalized and developed to become a tool to study optical absorption, magnetophonon resonance, and cyclotron resonance.

The present article, which is primarily devoted to the response properties of polarons, deals with the following aspects of polaron physics.

First, the polaron self-energy, the effective mass, the mobility, and the optical response are discussed. The optical absorption spectrum of a single large polaron at all electron–phonon coupling strengths has been analyzed using the path-integral method. The path-integral approach is also applied to investigate the properties of the interacting polaron gas including its optical absorption. Also Monte Carlo simulations have been used to examine the large-polaron model. Furthermore, the dynamics of the polaron formation has become a subject of investigation.

Second, polarons are studied in systems of reduced dimension and reduced dimensionality, for example, using cyclotron resonance. Magnetopolarons have been investigated in layered semiconductor compounds and quantum wells. The role of polaronic quasiparticles has been confirmed for the transport mechanism in quasi-1D conductors and for the optical properties of quantum dots. In semiconductor nanostructures, the mixing of the electron–hole or exciton states by the polaron interaction leads to the need for a nonadiabatic treatment.

Third, the polaron concept is used to study the properties of a variety of materials, for example, transition-metal oxides, conjugated polymers, perovskites, high-T_c superconductors, layered MgB_2 superconductors, and fullerenes. No consensus exists concerning the microscopic mechanism leading to the creation of Cooper pairs in high-T_c superconductors. Nevertheless, there is evidence for the existence of polarons and bipolarons in high-T_c superconductors (as manifested, for example, in their optical absorption spectra in the normal phase).

Recent Developments

Polaronic effects were observed in a 2D electron system on helium films covering a glass substrate. For thin films and for strong holding electric fields, the conditions were achieved under which the calculated binding energies are larger than k_BT, and the mobilities were shown to be in good agreement with the theoretically predicted values for the polaron mobility.

A theoretical analysis of the small-polaron and small-bipolaron formation in a cubic perovskite lattice showed that there exists a large window of electron–phonon coupling where the polaron is stable but the bipolaron decays into two polarons. An interesting development of the polaron concept was proposed in connection with recent studies in the theory of high-T_c superconductivity. The bipolaronic mechanism of superconductivity in terms of small bipolarons supposes that both electron–phonon and Coulomb interactions are short range (i.e., on-site). This model satisfactorily describes several properties of the insulating phase of the cuprates, but it cannot explain the high superconducting critical temperature. The effective mass of a bipolaron within the model with short-range interactions is very large, and the predicted critical temperature T_c is very low. Motivated by this fact, the so-called Fröhlich–Coulomb model of the high-T_c superconductivity was proposed, in which it is assumed that a long-range Fröhlich interaction, rather than a short-range Holstein interaction, should be the adequate model for the cuprates. Differently from the usual continuum Fröhlich polaron, a multipolaron Fröhlich-like lattice model is used with electrostatic forces taking into account the discreteness of the lattice, finite electron bandwidth and the quantum nature of phonons. Recently, this model was applied in order to explain various physical properties of superconducting cuprates (see the monograph by Alexandrov).

A path-integral Monte Carlo scheme was presented to solve the large-polaron model in three and two dimensions. The ground-state features of the Fröhlich polaron model were revisited numerically using a diagrammatic Quantum Monte Carlo method, and analytically using the variational Hamiltonian approach. Three aforementioned approaches demonstrate the remarkable accuracy of the Feynman path-integral approach to the polaron ground-state energy. A linear superposition of two variational wave functions, which describe the asymptotic regimes of weak and strong electron–phonon coupling, provides an estimate of the polaron ground-state energy, which is slightly lower than that obtained within the Feynman path-integral method, and compares well with the energies calculated within a diagrammatic Quantum Monte Carlo method. The α-dependence of the polaron effective mass and the other calculated ground-state properties suggest that the crossover between the two asymptotic regimes

characterizing a polaron occurs smoothly and does not suggest any sharp "self-trapping" transition. From the spectral analysis of the Green's function, it was concluded that, for all coupling strengths, in contrast to the small-polaron model, there is no stable excited state for the Fröhlich polaron in the energy gap between the ground-state energy and the incoherent continuum; however, several many-phonon unstable states are found.

The formation of a small polaron from a bare electron has been studied by Ku and Trugman. The polaron formation time is related to the dephasing time of the continuum of unbound phonon excited states in the spectral function, which depends on the phonon frequency and the electron–phonon coupling strength. As the electron–phonon coupling increases, qualitative changes in polaron formation occur when the one-phonon polaron bound states form. In the adiabatic regime, a potential barrier between the quasi-free and heavy polaron states exists in both two and three dimensions which is crucial for polaron formation dynamics.

The density-functional theory and its time-dependent extension is exploited by Bassani and co-workers to construct an appropriate effective potential for studying the properties of the interacting polaron gas beyond the mean-field theory. The main assumptions of the theory are to consider the coupled system of electrons and ions as a continuum, and to take the weak electron–phonon coupling limits into account.

The optical properties of quasi-one-dimensional metallic $LaTiO_{3.41}$ studied for the polarization along the *a* and *b* axes revealed that with decreasing temperature, modes appear along both directions suggestive of a phase transition. A pronounced midinfrared band was observed with a temperature dependence consistent with interacting polaron models. The polaronic picture was also corroborated by the presence of strong electron–phonon coupling and the temperature dependence of the DC conductivity. Band softening was interpreted in terms of large polaron models including polaron–polaron interactions. These findings suggest the general importance of polaronic quasiparticles for the transport mechanism in quasi-1D conductors.

The studies of thermal dissociation and photoionization for bound hole O^- polarons in oxide compounds MgO and corundum showed a surprising closeness of the results obtained on the basis of independent models: a continuum theory at arbitrary electron–phonon coupling and a quantum-chemical method. The bound-polaron model, applied to describe the optical properties of submonolayer CdSe insertions epitaxially grown between ZnSe layers, revealed excited-state energies equidistantly separated by the LO phonon energy as a typical signature of bound polarons.

Polaronic Response in Solids

Optical Absorption of Polarons at Arbitrary Coupling: Internal Structure of Polarons

The theoretical optical absorption spectrum of a single large polaron, at all electron–phonon coupling strengths, was obtained by Devreese and co-workers

$$\Gamma(\omega) = -\frac{1}{n\varepsilon_0 c}\,\frac{\mathrm{Im}\,\Sigma(\omega)}{[\omega - \mathrm{Re}\,\Sigma(\omega)]^2 + [\mathrm{Im}\,\Sigma(\omega)]^2} \quad [1]$$

with ω the frequency of the incident radiation expressed in units of the LO-phonon frequency in the medium ω_{LO}, ε_0 the dielectric permittivity of the vacuum, n the index of refraction of the medium, and c the velocity of light. The memory function $\Sigma(\omega)$ contains the dynamics of the polaron and depends on α and ω. This function is related to the impedance function $Z(\omega)$ of the polaron through $-iZ(\omega) = \omega - \Sigma(\omega)$. The starting point was the general eqn [1] together with the $[\mathrm{Im}\,\Sigma(\omega)]$ obtained by Feynman and co-workers for the Feynman polaron. A key ingredient is the calculation of $\mathrm{Re}\,\Sigma(\omega)$ (which is a Kramers–Kronig-type transform of $[\mathrm{Im}\,\Sigma(\omega)]$).

The resulting polaron optical absorption spectra are shown in **Figure 1** for values of α in the range from 1 to 6.

The absorption spectrum for relatively weak coupling (e.g., $\alpha = 1$) consists of a "one-phonon line," similar to the weak-coupling result obtained by Gurevich, Lang, and Firsov. The absorption peak for $\alpha = 3$ contains components with more than one free phonon in the final state. It was suggested that part of the oscillator strength in the polaron absorption for $\alpha = 3$ is due to incipient lattice relaxation in the final states, though to a much lesser extent than for $\alpha = 6$.

From [1], as applied to the Feynman model, the three different kinds of polaron excitations, which were found by Kartheuser, Evrard, and Devreese using the adiabatic strong-coupling approximation, appear in the calculated optical absorption spectra for polarons at $\alpha \geqslant 5$:

1. scattering states where, for example, one real phonon is excited (the structure starting at $\omega = \omega_{LO}$);
2. relaxed excited states (RES), the "resonance" polaron states, which result if the lattice polarization is allowed to adapt to the electronic distribution of the excited electron (which itself then adapts its wave function to the new potential etc., leading to a self-consistent final state);

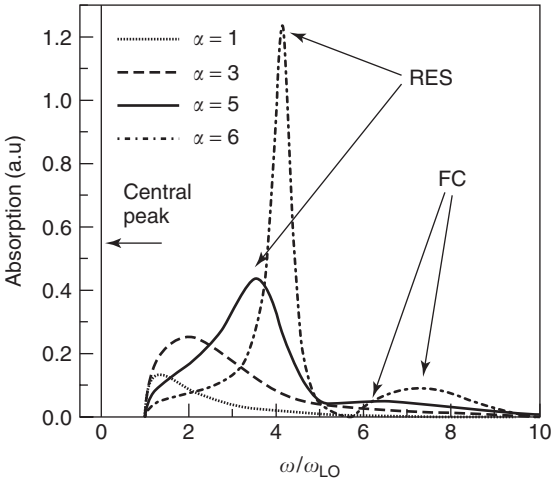


Figure 1 Polaron optical absorption spectra at zero temperature, calculated within the path-integral method according to J T Devreese and co-workers for different values of the electron–phonon coupling constant α. The peaks are labeled in the picture according to their physical origin ("RES" = relaxed excited state, "FC" = Franck–Condon state) as discussed in the text. A δ-like central peak (at $\omega = 0$) is schematically shown by a vertical line. (Reproduced with permission from Devreese JT (2003) Polarons. In: Avella A and Mancini F (eds.) *Lectures on the Physics of Highly Coordinated Electron Systems VII,* Proceedings of the 7th Training Course in the Physics of Correlated Electron Systems and High-T_c Superconductors, Vietri sul Mare, Italy, 14–16 October 2002, pp. 3–56 Melville: AIP; © American Institute of Physics.)

3. Franck–Condon (FC) states, the (unstable) polaron states, in which the lattice polarization corresponds to the electron ground state, while the electron is excited.

Furthermore, at zero temperature, the optical absorption spectrum for one polaron contains a "central peak" [$\sim\delta(\omega)$] at the origin. For nonzero temperature, this "central peak" smears out and the optical absorption spectrum consists of both a broad envelope and an "anomalous" Drude-type low-frequency component.

For example, in **Figure 1**, the main peak of the polaron optical absorption for $\alpha = 5$ at $\omega = 3.51\omega_{LO}$ is interpreted as due to transitions to an RES. A "shoulder" at the low-frequency side of the main peak is attributed to one-phonon transitions to "scattering states." The broad structure centered at about $\omega = 6.3\omega_{LO}$ is interpreted as an FC band. As seen from **Figure 1**, when increasing the electron–phonon coupling constant to $\alpha = 6$, the RES peak at $\omega = 4.3\omega_{LO}$ stabilizes and becomes very intense compared with the FC peak centered at $\omega = 7.5\omega_{LO}$.

Recent calculations of the optical conductivity for the Fröhlich polaron performed within the diagrammatic Quantum Monte Carlo method fully confirm the results of the path-integral variational approach at $\alpha \lesssim 3$. In the intermediate-coupling regime $3 < \alpha < 6$, the low-energy behavior and the position of the maximum of the optical conductivity spectrum as obtained by Mishchenko and co-workers follow well the prediction of the path-integral approach based on eqn [1]. There are the following qualitative differences between the two approaches in the intermediate and strong coupling regime: in the Quantum Monte Carlo method, the dominant peak broadens and the second peak does not develop, giving rise instead to a flat shoulder in the optical conductivity spectrum at $\alpha = 6$. This behavior can be attributed to the optical processes with the participation of two or more phonons. The nature of the excited states of a polaron needs further study.

Free-carrier polaron effects were revealed through measurements of the optical constants, of the cyclotron resonance, and of the carriers mobility in polar semiconductors and insulating photoconductors. Optical absorption experiments revealed spectra, consistent with the theoretical predictions in **Figure 1** up to $\alpha \sim 3$. The more dramatic structure of $\Gamma(\omega)$, eqn [1], characteristic of transitions to the RES for $\alpha \geqslant 5$ arises in cyclotron resonance experiments (see "Further reading" section).

Scaling relations for polarons in 2D and in 3D Several scaling relations connect the polaron self-energy, the effective mass, the impedance Z and the mobility μ in two dimensions to the same quantities in three dimensions. Those relations were obtained at the level of the Feynman model and are listed here:

$$E_{2D}(\alpha) = \frac{2}{3} E_{3D}\left(\frac{3\pi}{4}\alpha\right) \qquad [2]$$

$$\frac{m^*_{2D}(\alpha)}{m_{2D}} = \frac{m^*_{3D}(3\pi/4)\alpha}{m_{3D}} \qquad [3]$$

$$Z_{2D}(\alpha, v) = Z_{3D}\left(\frac{3\pi}{4}\alpha, v\right) \qquad [4]$$

where v is the frequency of the external electromagnetic field, and

$$\mu_{2D}(\alpha) = \mu_{3D}\left(\frac{3\pi}{4}\alpha\right) \qquad [5]$$

Expressions [2] to [5] provide a straightforward link between polaron characteristics in three dimensions with those in two dimensions.

Many-Polaron Systems

Interacting polarons In most polar semiconductors, the weak-coupling regime is realized. The perturbation theory is then adequate to describe the polaron gas in such semiconductors. In early approaches to

the problem of a high-density polaron gas, the interaction between polarons was supposed to be completely screened. The ground-state energy of a gas of interacting polarons was first derived by L F Lemmens, J T Devreese, and F Brosens using a canonical transformation extended to the case of a many-polaron system. It was shown that the minimum of the total ground-state energy per particle for a polaron gas in ZnS and in ZnO lies at a density lower than that for the electron gas.

A Wigner lattice of polarons is assumed to be the ground state of the many-polaron system in the low-density limit. The nondegenerate system of interacting polarons in polar-doped insulators is analyzed by Quimerais and Fratini using a Feynman model for the polaron, while the effect of other polarons forming a Wigner lattice is approximately described by a parabolic confinement potential. With increasing density, there is a competition between the melting of the polaron Wigner lattice and the dissociation of the polarons themselves. As a result, for sufficiently large values of the electron–phonon coupling constant α, the formation of a polaron liquid appears to be impossible.

Optical absorption coefficient of a many-polaron gas The possibility that polarons play a role in high-T_c superconductors has increased interest in the physical properties of many-polaron systems and, in particular, in their optical properties. Theoretical treatments were extended from one-polaron to many-polaron systems, and 2D as well as 3D structures were examined.

The optical absorption spectrum of a many-polaron gas has been derived starting from a many-polaron canonical transformation and from the variational many-polaron wave function. This approach to the many-polaron optical absorption allows one to include the many-body effects in terms of the dynamical structure factor of the electron (or hole) system. It is then possible, within the weak (and intermediate) coupling polaron theory, to select the level of approximation used in the treatment of the many-polaron gas by choosing an appropriate expression of the dynamical structure factor for the electron (or hole) gas.

The many-polaron optical conductivity is governed by the response of the current density to an applied electric field (here along the x-axis) with frequency ω. Introducing the standard expression for the dynamical structure factor of a system of charge carriers interacting through a Coulomb potential:

$$S(\boldsymbol{q}, v) = \int_{-\infty}^{+\infty} \left\langle \varphi_{\mathrm{el}} \left| \frac{1}{2} \sum_{j,\ell} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{r}_j(t)-\boldsymbol{r}_\ell(0))} \right| \varphi_{\mathrm{el}} \right\rangle \mathrm{e}^{\mathrm{i}vt}\, \mathrm{d}t \qquad [6]$$

the real part of the many-polaron optical conductivity can be expressed in the form

$$\mathrm{Re}[\sigma(\omega)] = \frac{n}{\hbar\omega^3} \frac{e^2}{m_{\mathrm{b}}^2} \sum_{\boldsymbol{k}} k_x^2 |V_{\boldsymbol{k}}|^2 S(\boldsymbol{k}, \omega - \omega_{\mathrm{LO}}) \qquad [7]$$

where $n = N/V$ is the density of charge carriers, $V_{\boldsymbol{k}}$ is the electron–phonon interaction amplitude and k_x is the x-component of the wave vector. Formula [7] for the optical absorption of the many-polaron system has an intuitively appealing form. In the one-polaron limit, eqn [7] reduces to the result of Gurevich and co-authors for the nondegenerate case.

Equation [7] is reminiscent of the formula of Hopfield that describes the effect of impurities on the optical absorption of metals. The earlier discussion of many-body effects related to Fermi statistics was subsequently extended by Tempere and Devreese to study the influence of plasmons and the many-body effects in polaron systems.

The optical properties of the many-polaron gas were investigated by calculating the correction to the RPA dielectric function of the electron gas due to the electron–phonon interaction. Variational parameters of a single-polaron Feynman model were exploited for the treatment of the interacting many-polaron gas. In order to find the many-polaron optical conductivity, the memory-function approach for the optical absorption of noninteracting large polarons is used. As a result, a suppression of the optical absorption with increasing density is found, as seen from **Figure 2**.

Comparison of theoretical optical-absorption spectra to the observed infrared spectra of $Nd_{2-x}Ce_xCuO_{2-y}$ and $La_{2/3}Sr_{1/3}MnO_3$ The experimental optical absorption spectrum (up to $3000\,\mathrm{cm}^{-1}$) of $Nd_2CuO_{2-\delta}$ ($\delta < 0.004$), obtained by Calvani and coworkers is shown in **Figure 3** (shaded area) together with the theoretical curve for the interacting many-polaron gas and, for reference, the one-polaron optical absorption result. At lower frequencies (600–$1000\,\mathrm{cm}^{-1}$), a marked difference between the single polaron optical absorption and the many-polaron result is manifest. The experimental "d-band" can be clearly identified, rising in intensity at about $600\,\mathrm{cm}^{-1}$, peaking around $1000\,\mathrm{cm}^{-1}$, and then decreasing in intensity above that frequency. A good agreement between the theoretical predictions and the experimental results (for a density of $n = 1.5\times 10^{17}\,\mathrm{cm}^{-3}$) is found. A background contribution, taken to be constant over the frequency range of the "d-band," is subtracted in **Figure 3**.

The material parameters characterizing polarons ($\varepsilon_0, \varepsilon_\infty, m_{\mathrm{b}}$) are not known for $Nd_2CuO_{2-\delta}$ ($\delta < 0.004$). For ω_{LO}, a value of $7.005\times 10^{-2}\,\mathrm{eV}$

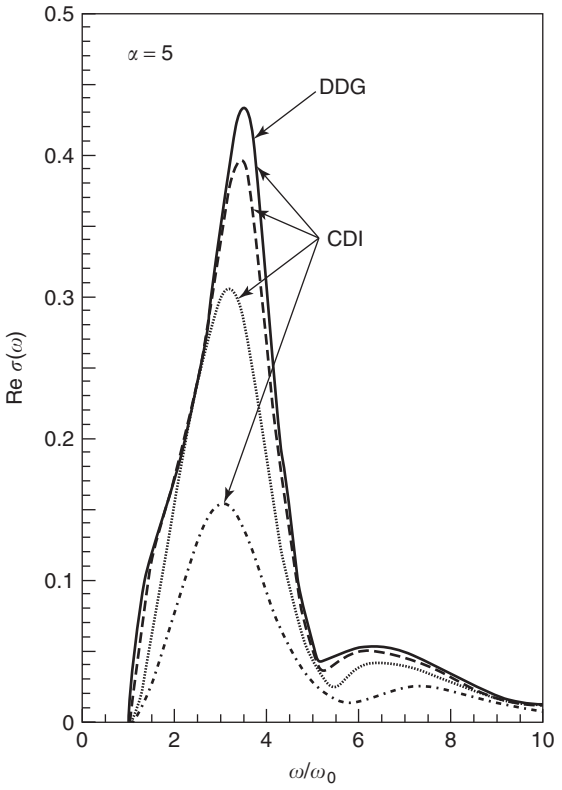


Figure 2 Optical conductivity of a polaron gas at $T = 0$ as a function of the frequency (CDI) for different values of the electron density: $n_0 = 1.4 \times 10^{-5}$ (solid curve), $n_0 = 1.4 \times 10^{-4}$ (dashed curve), $n_0 = 1.4 \times 10^{-3}$ (dotted curve), and $n_0 = 1.4 \times 10^{-2}$ (dotted–dashed curve). The value of $\varepsilon_0/\varepsilon_\infty$ is 3.4. The electron density is measured in units of R_p^{-3}, where R_p is the Fröhlich polaron radius. The solid curve practically coincides with the optical conductivity of a single polaron (DDG). (Reproduced with permission from Cataudella V, De Filippis G, and Iadonisi G (1999) *The European Physical Journal B* 12 : 17–22; © EDP Sciences, SIF, and Springer. Courtesy of G. Iadonisi.)

was derived from available experimental data. In the calculations, the values $\alpha = 2.1$ and $m_b = 0.5m_e$ were considered. However, these figures need further analysis in view of the limited available experimental data for $\varepsilon_0, \varepsilon_\infty, m_b$.

There exists also a fair agreement between the theoretical and the experimental values of the normalized first-frequency moment for the samples of $Nd_{2-x}Ce_xCuO_4$ with lowest density, which have a cerium doping content of $x < 0.12$. For $x > 0.12$, a discrepancy between the theoretically predicted first-frequency moment for unpaired polarons and the observed first-frequency moment appears. It has been observed experimentally that the weight of the low-frequency component in these samples (with $x > 0.12$) is significantly larger than the corresponding weight in samples with $x < 0.12$. This was interpreted as a consequence of an insulator-to-metal transition taking place around a cerium-doping level of $x = 0.12$. Therefore, it seems reasonable to assume that above this doping level a change in the nature of the charge carriers takes place. One could hypothesize that, as the formation of bipolarons is stabilized with increasing density of the polaron gas, bipolarons start playing a role in the optical absorption spectrum. In a variety of other cuprates and manganites, the presence of bipolarons has also been invoked to interpret a number of response-related properties.

The experimentally observed optical conductivity spectrum of $La_{2/3}Sr_{1/3}MnO_3$ is compared with different theoretical large-polaron models in **Figure 4**. The one-polaron approximations (the weak-coupling approach and the phenomenological model) lead to narrower polaron peaks than the peak with maximum at $\omega \sim 900\,\mathrm{cm}^{-1}$ given by the many-polaron approach. It is seen from **Figure 4** that the many-polaron approach describes the experimental optical conductivity better than the single-polaron methods.

Thermal conductivity of cuprates In cuprates, it is difficult to separate experimentally the electron and phonon contributions to the thermal conductivity, since they are of the same order of magnitude. Recently, a way to experimentally determine the Lorenz number has been realized by Zhang and co-authors, based on the thermal Hall conductivity. The thermal Hall effect allows for an efficient way to separate the phonon heat current from the electric heat current. The particular interest of the experiment lies in the conclusion of Alexandrov and coworkers that the Wiedemann–Franz law is violated in the cuprates in the temperature range below the temperature T^* of the BCS-Bose liquid crossover.

The Lorenz number of the electron–phonon system has been calculated by Lee, Alexandrov, and Liang in order to show that the Wiedemann–Franz law breaks down because of the interference of small-polaron and small-bipolaron contributions in the heat flow. When thermally excited polarons are included in the consideration, the bipolaron model explains the violation of the Wiedemann–Franz law in the cuprates and the behavior of the Hall–Lorenz number as seen in the experiment. As follows from **Figure 5**, the model is in good agreement with the experimental Hall–Lorenz number, which provides direct evidence for bipolarons in the cuprates. It also quantitatively reproduces the (quasi)linear in-plane resistivity and the inverse Hall ratio, as observed in the cuprates (upper inset of **Figure 5**). The lower inset of **Figure 5** shows a slightly lower Lorenz number compared to the calculated Hall–Lorenz number. Because the thermal Hall conductivity determines the Lorenz number in the framework of the model, it can be used to find the lattice contribution to the heat

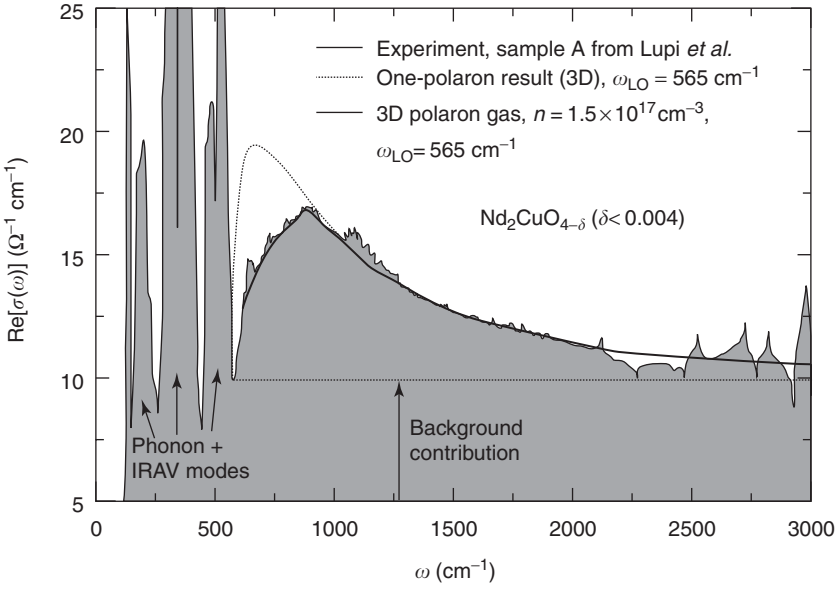


Figure 3 The infrared optical absorption of $Nd_2CuO_{2-\delta}$ ($\delta < 0.004$) as a function of frequency. The experimental results of P Calvani and co-workers are represented by the thin black curve and by the shaded area. The dotted curve shows the single polaron result. The full curve presents the theoretical results obtained for the interacting many-polaron gas with $n = 1.5 \times 10^{17}$ cm^{-3}, $\alpha = 2.1$, and $m_b = 0.5m_e$. (Recently, a numerical error was found in the fitting procedure, concerning strength of the plasmon branch, so that the parameter values given above may not be the optimal parameter values.) (Reprinted with permission from Tempere J and Devreese JT (2001) *Physical Review B* 64: 104504. 1–10; © American Physical Society.)

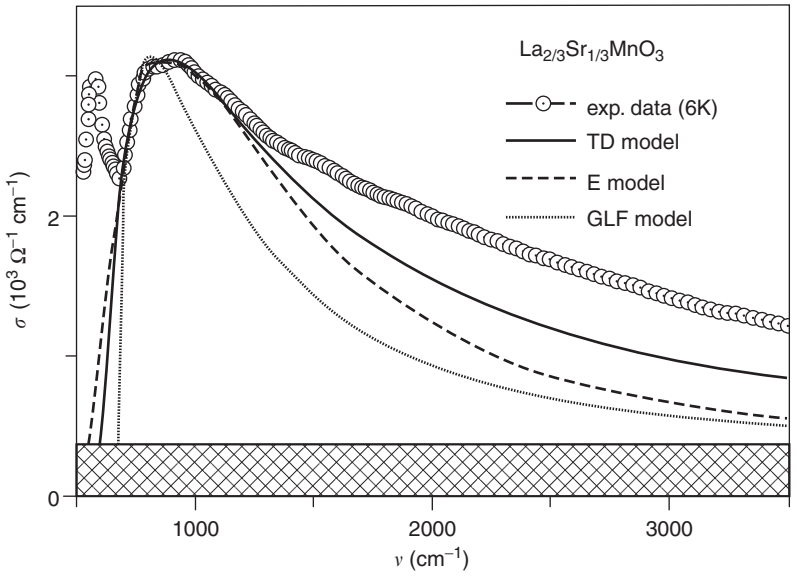


Figure 4 Comparison of the mid-infrared optical conductivity in $La_{2/3}Sr_{1/3}MnO_3$ measured by Ch Hartinger and co-authors at $T = 6$K to that given by several model calculations for $m_b = 3m_e$ and $n_0 = 6 \times 10^{21}$cm^{-3} . The solid curve refers to the many-polaron single-mode approach (TD model), the dashed curve is the result of the phenomenological approach for self-trapped polarons (E model by D Emin), the dotted curve is the weak-coupling one-polaron result (GLF model). (Reprinted with permission from Hartinger Ch, Mayr F, Deisenhofer J, Loidl A, and Kopp T (2004) *Physical Review B* 69: 100403(R). 1–4; © American Physical Society.)

flow as well. When the electronic contribution, determined by using the Lorenz number is subtracted, the lattice contribution to the diagonal heat flow appears to be much higher than is anticipated in the framework of any Fermi-liquid model.

Cyclotron Resonance of Polarons

Polaron effects in a magnetic field are sometimes referred to as "magnetopolaron" effects. These effects have been reviewed, for example, Larsen and Petrou and McCombe.

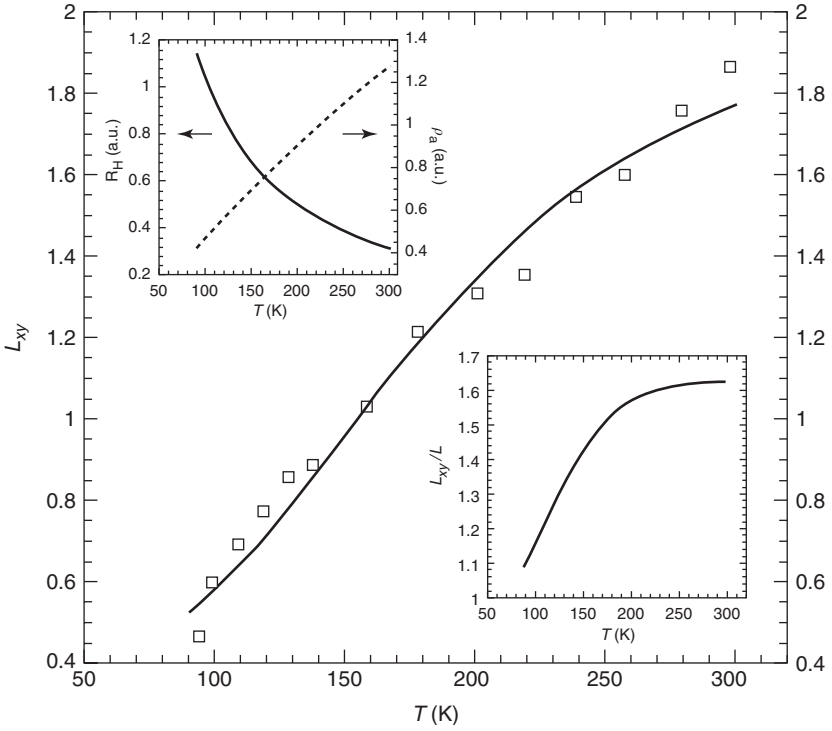


Figure 5 The experimental Hall–Lorenz number obtained by Y Zhang *et al.* in $YBa_2Cu_3O_{6.95}$ fitted by the small itinerant bipolaron theory. The upper inset shows the linear in-plane resistivity and the Hall ratio. The lower inset shows the ratio of the Hall–Lorenz number to the Lorenz number as a function of temperature. (Reprinted with permission from Lee KK, Alexandrov AS, and Liang WY (2003) *Physical Review Letters* 90: 217001; © American Physical Society.)

A powerful technique to study the electron–phonon coupling in semiconductors is cyclotron resonance. Changing the magnetic field B allows one to vary the cyclotron frequency $\omega_c = eB/m_b c$. When ω_c is brought closer to resonance with the phonon frequency ω_{LO}, a strongly renormalized "magnetopolaron" state is formed. This results in a splitting of the cyclotron resonance signal, which is a measure of the electron–phonon (polaron) coupling constant α.

The experimental observation by Johnson and Larsen of a discontinuity in the energy spectrum of a polaron in a magnetic field under the resonance condition $\omega_c = \omega_{LO}$ has stimulated numerous theoretical works aimed at the description of observed magnetoabsorption spectra. In the early works, the energy spectra of a polaron subjected to a magnetic field were studied in the weak-coupling limit, and the polaron cyclotron mass was determined from the energy differences between the polaron (Landau-) energy levels.

For an adequate description of experiment, the magnetoabsorption spectrum itself must be calculated. In the sixties and seventies, several weak-coupling approximations for the magnetoabsorption spectra of polarons were suggested.

The most direct approach to study the (magneto-) polaron mass is to calculate the magnetooptical absorption spectrum of the polaron (the quantity which is actually measured) and to define the polaron mass in the same way as the experimentalists do, from the peak positions in the spectrum. This approach was chosen by Peeters and Devreese on the grounds of the memory-function formalism and elaborated over many years as a basic theoretical tool to analyze experimental data on magnetopolaron spectra.

A direct evidence for the polaron character of charge carriers in AgBr and AgCl was obtained through CR experiments in external magnetic fields up to 16 T. The magnetoabsorption calculated within the all-coupling approach leads to the best quantitative agreement between theory and experiment for AgBr and AgCl.

Nicholas and co-workers demonstrated polaron coupling effects using cyclotron resonance measurements in a 2DEG, which naturally occurs in the polar semiconductor InSe. One clearly sees, over a wide range of magnetic fields ($B = 18-34$ T), two distinct magnetopolaron branches separated by as much as 11 meV ($\sim 0.4\omega_{LO}$) at resonance.

In quasi-2D semiconductor structures, the phonon modes are modified by the presence of the interfaces. Wang and co-authors studied the cyclotron resonance of polarons in $GaAs/Al_{0.3}Ga_{0.7}As$ multiple quantum well structures. It appeared possible to observe the anticrossing effects due to different phonon modes.

The analysis of the experimental data has shown that the observed anticrossing effects are caused by the resonant interaction of electrons with the confined GaAs slab phonon modes and the AlAs-like interface optical phonons.

The availability of very high magnetic fields allows one to achieve the resonant condition $\omega_c = \omega_{LO}$ as well as the condition $\omega_c > \omega_{LO}$ even for II–IV semiconductor compounds, where the effective mass of the conduction electrons can be considerably larger than in II–V semiconductors. Experiments on cyclotron resonance in bulk *n*-type CdS at ultrahigh magnetic fields (up to 450 T) revealed a pronounced maximum in the temperature dependence of the cyclotron mass. Polaron effects beyond the weak-coupling approximation significantly influence the magnetoabsorption even for a relatively small Fröhlich coupling constant like 0.527 in *n*-type CdS. Extending the memory-function approach to account for the elastic scattering of electrons by piezoacoustic phonons and lattice defects provides a quantitative interpretation for the experimental data of Imanaka, Mivra, and Nojiri.

Polaronic Response in Confined Structures

Electron–Phonon Interaction and Cyclotron Resonance in Quantum Wells

Numerous experiments on cyclotron resonance give clear evidence of this effect. The resonant magnetopolaron coupling manifests itself near the LO-phonon frequency for low electron densities and also for higher electron densities. Cyclotron resonance measurements performed on semiconductor quantum wells with high electron density reveal anticrossing near the TO-phonon frequency rather than near the LO-phonon frequency. This effect is interpreted by invoking mixing between magnetoplasmons and phonons and in terms of a resonant coupling of the electrons with the mixed magnetoplasmon–phonon modes.

The CR spectra for a polaron gas in a GaAs/AlAs quantum well are theoretically investigated, taking into account (1) the electron–electron interaction and the screening of the electron–phonon interaction, (2) the magnetoplasmon–phonon mixing, and (3) the electron–phonon interaction with all the phonon modes specific for the quantum well under investigation. As a result of this mixing, different magnetoplasmon–phonon modes appear in the quantum well which give contributions to the CR spectra.

It is clearly seen from **Figure 6** that, for the high-density polaron gas, anticrossing of the CR spectra occurs near the GaAs TO-phonon frequency ω_{T1} rather than near the GaAs LO-phonon frequency ω_{L1} for both the experimental and the calculated spectra. This effect is in contrast with the cyclotron resonance of a low-density polaron gas in a quantum well, where anticrossing occurs near the LO-phonon frequency. The occurrence of the anticrossing frequency close to ω_{T1} instead of ω_{L1} is due to the screening of the

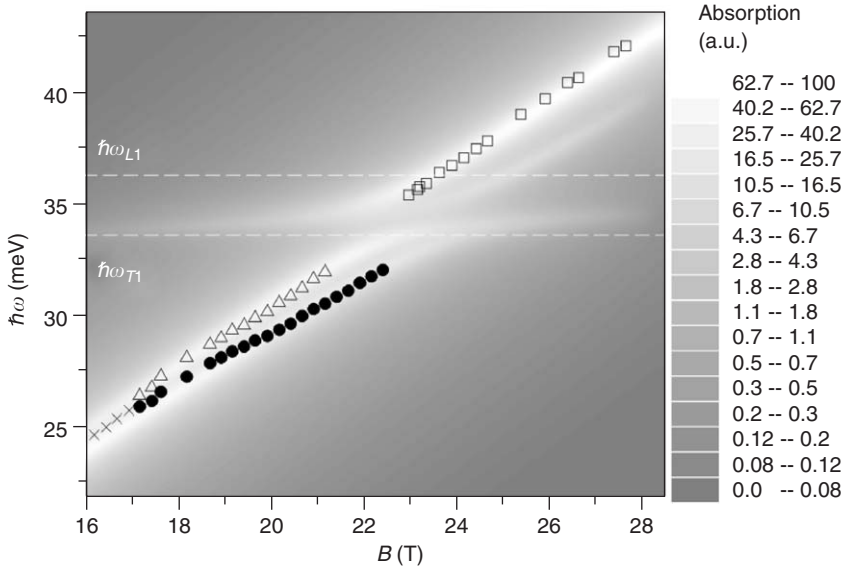


Figure 6 Density map of the calculated magnetoabsorption spectra for a 10-nm GaAs/AlAs quantum well. Symbols indicate peak positions of the experimental spectra (which are taken from figure 3 of Poulter AJL, Zeman J, Maude DK, Potemski M, Martinez G, *et al.* (2001) *Physical Review Letters* 86: 336–339.) Dashed lines show LO- and TO-phonon energies in GaAs. (Reprinted with permission from Klimin SN and Devreese JT (2003) *Physical Review B* 68: 245303. 1–8; © American Physical Society.)

longitudinal polarization by the plasma vibrations. A similar effect appears also for the magnetophonon resonance: as shown by Afonin and co-authors, the magnetoplasmon–phonon mixing leads to a shift of the resonant frequency of the magnetophonon resonance in quantum wells from $\omega \approx \omega_{L1}$ to $\omega \approx \omega_{T1}$.

Interacting Polarons in Quantum Dots

The ground-state energy and the optical conductivity spectra for a system with a finite number of interacting arbitrary-coupling large polarons in a spherical quantum dot are calculated using the path-integral formalism for identical particles. A parabolic confinement potential characterized by the confinement energy $\hbar\Omega_0$ and with a background charge is considered. Using a generalization of the Jensen–Feynman variational principle, the ground-state energy of a confined N-polaron system is analyzed as a function of N and of the electron–phonon coupling strength α.

The total spin S is analyzed as a function of the number of electrons in the quantum dot for different values of the confinement energy, of the coupling constant, and of the ratio η of the high-frequency and the static dielectric constants.

Confined few-electron systems, without electron–phonon interaction, can exist in one of the two phases: a spin-polarized state and a state obeying Hund's rule, depending on the confinement frequency. For interacting few-polaron systems with $\alpha \geqslant 3$, besides the above two phases, there may occur a third phase: the state with minimal spin.

To investigate the optical properties of a many-polaron system, the memory-function formalism is extended by Devreese and co-workers to the case of interacting polarons in a quantum dot. The shell structure for a system of interacting polarons in a quantum dot is clearly revealed when analyzing both the addition energy and the first-frequency moment of the optical conductivity. The envelope of the optical absorption spectrum for polarons in a quantum dot bears a similarity to the corresponding spectrum for an interacting polaron gas, especially at low densities.

Electron–Phonon Interaction and the Optical Response of Polaronic Excitons in Quantum Dots

The interest in the optical properties of quantum dots has been continuously growing, among others, because of the prospects of these structures for optoelectronic applications. Recent measurements of photoluminescence of self-assembled InAs/GaAs quantum dots reveal high probabilities of the phonon-assisted transitions. Attempts to interpret some of these experiments on the basis of the adiabatic theory meet with difficulties. For spherical CdSe quantum dots, for example, the Huang–Rhys factor S calculated within the adiabatic approximation takes values which are significantly (by one or two orders of magnitude) smaller than the so-called "experimental Huang–Rhys" factor, determined from the ratio of measured intensities of the phonon satellites. In the framework of the adiabatic approach, various mechanisms, which cause the separation of the electron and hole charges in space are commonly considered as a possible origin for the unexpectedly large Huang–Rhys factor.

It has been shown that nonadiabaticity of exciton–phonon systems in various quantum-dot structures drastically enhances the efficiency of the exciton–phonon interaction, especially when the exciton levels are separated with energies close to the phonon energies. Also "intrinsic" excitonic degeneracy can lead to enhanced efficiency of the exciton–phonon interaction. The effects of nonadiabaticity are important to interpret the surprisingly high intensities of the phonon "sidebands" observed in the optical absorption, the photoluminescence, and the Raman spectra of quantum dots. Considerable deviations of the oscillator strengths of the phonon-peak sidebands, observed in some experimental optical spectra from the Franck–Condon progression, which are prescribed by the commonly used adiabatic approximation, find a natural explanation within the nonadiabatic approach.

Fomin and co-authors proposed a method to calculate the optical absorption spectrum for a spherical quantum dot taking into account the nonadiabaticity of the excition–phonon system. This approach has been further refined by Gladilin and co-authors for the matrix elements of the evolution operator: a closed set of equations has been obtained using a diagrammatic technique. This set describes the effect of nonadiabaticity both on the intensities and positions of the absorption peaks.

From the comparison of the spectra obtained in the adiabatic approximation with those resulting from the nonadiabatic approach, the following effects of nonadiabaticity are revealed. First, the polaron shift of the zero-phonon lines with respect to the bare-exciton levels is larger in the nonadiabatic approach than in the adiabatic approximation. Second, there is a strong increase of the intensities of the phonon satellites as compared to those given by the adiabatic approximation. Third, in the optical absorption spectra found within the nonadiabatic approach, there appear phonon satellites related to nonactive bare-exciton states. Fourth, the optical absorption spectra demonstrate the crucial role of nonadiabatic mixing of different exciton and phonon states in quantum

dots. This results in a rich structure of the absorption spectrum of the exciton–phonon system. Similar conclusions about the pronounced influence of the exciton–phonon interaction on the optical spectra of quantum dots have been formulated by Verzelen and co-authors in terms of a strong coupling regime for excitons and LO phonons. Such a strong coupling regime is a particular case of the nonadiabatic mixing related to a (quasi-) resonance which arises when the spacing between exciton levels is close to the LO phonon energy. The large enhancement of the two-phonon sidebands in the luminescence spectra as compared to those predicted by the Huang–Rhys formula, which was explained by nonadiabaticity of the exciton–phonon system in certain quantum dots, has been reformulated in terms of the Fröhlich coupling between product states with different electron and/or hole states.

Due to nonadiabaticity, multiple absorption peaks appear in spectral ranges characteristic for phonon satellites. From the states, which correspond to these peaks, the system can rapidly relax to the lowest emitting state. Therefore, in the photoluminescence excitation (PLE) spectra of quantum dots, pronounced peaks can be expected in spectral ranges characteristic for phonon satellites. Experimental evidence of the enhanced phonon-assisted absorption due to effects of nonadiabaticity has been provided by PLE measurements on single self-assembled InAs/GaAs and InGaAs/GaAs quantum dots.

See also: Cyclotron Resonance: Semiconductors; Dielectric Function; Local Field Effects; Polarons.

PACS: 71.38. – k; 63.20.Kr; 71.38.Mx; 71.10. – w; 73.21. – b; 74.20.Mn

Further Reading

Afonin VV, Gurevich VL, and Laiho R (2000) *Physical Review B* 62: 15913–15924.

Alexandrov AS (2003) *Theory of Superconductivity. From Weak to Strong Coupling*. Bristol: IOP Publishing.

Alexandrov AS and Mott NF (1994) *High-Temperature Superconductors and Other Superfluids*. London: Taylor and Francis.

Alexandrov AS and Mott N (1996) *Polarons and Bipolarons*. Singapore: World Scientific.

Bassani FG, Cataudella V, Chiofalo ML, De Filippis G, Iadonisi G, *et al.* (2003) *Physics Status Solidi (b)* 237: 173–185.

Calvani P (2001) *Optical Properties of Polarons*. Bologna: Editrice Compositori.

Devreese JT (1972) *Polarons in Ionic Crystals and Polar Semiconductors*. North-Holland: Amsterdam.

Devreese JT (1996) Polarons. In: Trigg GL (ed.) *Encyclopedia of Applied Physics*, vol. 14. pp. 383–413. Weinheim: VCH.

Devreese JT (2001) Note of the path-integral variational approach in the many-body theory. In: Janke W, Pelster A, Bachmann M, and Schmidt HJ (eds.) *Fluctuating Paths and Fields*, pp. 289–304. Singapore: World Scientific.

Devreese JT, De Sitter J, and Goovaerts M (1972) *Physical Review B* 5: 2367–2381.

Devreese JT, Klimin SN, Fomin VM, and Brosens F (2000) *Solid State Communications* 114: 305–310.

Devreese JT, Verbist G, and Peeters F (1995) Large bipolarons and high-T_c materials. In: Salje E, Alexandrov A, and Liang W (eds.) *Polarons and Bipolarons in High-T_c Superconductors and Related Materials*, pp. 385–391. Cambridge: Cambridge University Press.

Feynman RP (1955) *Physical Review* 97: 660–665.

Feynman RP, Hellwarth RW, Iddings CK, and Platzman PM (1962) *Physical Review 127: 1004–1017.*

Fomin VM, Gladilin VN, Devreese JT, Pokatilov EP, Balaban SN, and Klimin SN (1998) *Physical Review B* 57: 2415–2425.

Gladilin VN, Klimin SN, Fomin VM, and Devreese JT (2004) *Physical Review B* 69: 155325. 1–5.

Gurevich VL, Lang IG, and Firsov YuA (1962) *Fiz. Tverd. Tela* 4: 1252–1262 (*Sov. Phys. Solid State*, 918–925 (1962)).

Hopfield JJ (1965) *Physical Review* 139: A419–A424.

Imanaka Y, Miura N, and Nojiri H (1998) *Physica B* 246–247: 328–332.

Johnson EJ and Larsen DM (1966) *Physical Review Letters* 16: 655–659.

Kartheuser E, Evrard R, and Devreese J (1969) *Physical Review Letters* 22: 148–151.

Klimin SN, Fomin VM, Brosens F, and Devreese JT (2004) *Physical Review B* 69: 235324. 1–16.

Ku LC and Trugman SA. Quantum Dynamics of Polaron Formation, cond-mat/0310226.

Kuper CG and Whitfield GD (eds.) (1963) *Polarons and Excitons*. Edinburgh: Oliver and Boyd.

Larsen DM (1972) Polaron energy levels in magnetic and Coulomb fields. In: Devreese JT (ed.) *Polarons in Ionic Crystals and Polar Semiconductors*, pp. 237–287. Amsterdam: North-Holland.

Larsen DM (1991) Polaron effects in cyclotron resonance. In: Landwehr G and Rashba EI (eds.) *Landau Level Spectroscopy* (in 2 volumes), pp. 109–130. Amsterdam: North-Holland.

Lemaître A, Ashmore AD, Finley JJ, Mowbray DJ, Skolnick MS, *et al.* (2001) *Physical Review B* 63: 161309 (R). 1–4.

Lemmens LF, Devreese JT, and Brosens F (1977) *Physics Status Solidi* 82: 439–447.

Lupi S, Maselli P, Capizzi M, Calvani P, Giura P, *et al.* (1999) *Physical Review Letters* 83: 4852–4855.

Mahan GD (1972) Polarons in heavily doped semiconductors. In: *Polarons in Ionic Crystals and Polar Semiconductors*, pp. 553–657. Amsterdam: North-Holland.

Mishchenko AS, Nagaosa N, Prokof'ev NV, Sakamoto A, and Svistunov BV (2003) *Physical Review Letters* 91: 236401.

Nicholas RJ, Watts M, Howell DF, Peeters FM, Wu X-G, Devreese JT, van Bockstal L, Herlach F, Langerak CJGM, Singleton J, and Chery A (1992) *Physical Review B* 45: 12144–12147.

Petrou A and McCombe BD (1991) Magnetospectroscopy of confined semiconductor systems. In: Landwehr G and Rashba EI (eds.) *Landau Level Spectroscopy* (in 2 volumes), pp. 679–775. Amsterdam: North-Holland.

Quémerais P and Fratini S (1998) *International Journal of Modern Physics* 12: 3131–3136.

Verzelen O, Ferreira R, and Bastard G (2002) *Physical Review Letters* 88: 146803.

Wang YJ, Nickel HA, McCombe BD, Peeters FM, Shi JM, Hai GQ, Wu X-G, Eustis TJ, and Schaff W (1997) *Physical Review Letters* 79: 3226–3229.

Zhang Y, Ong NP, Xu ZA, Krishana K, Gagnon R, and Taillefer L (2000) *Physical Review Letters* 84: 2219.

Electrons and Holes

G C La Rocca, Scuola Normale Superiore, Pisa, Italy

Introduction

The paramount role of semiconductors in technology ever since the invention of the transistor consists mainly in their use in electronic devices, particularly very large scale integrated circuitry, in which electric currents are due to mobile electrons and holes. The possibility of growing single crystals of very high purity and to control their transport properties through intentional *n*-type and *p*-type doping has been crucial to the development of the electronic industry. Silicon is presently the material of choice for electronic devices not only because of its outstanding processing properties, but also because its native oxide forms a high-quality semiconductor–insulator interface. Semiconductors have at the same time attracted a lot of attention from the viewpoint of basic science, and understanding their electronic structure, optical, and transport properties has been a great success of the modern theory of solids.

The resistivity of semiconductors is intermediate between that of metals and insulators, varying at room temperature between 10^{-4} and $10^{10}\,\Omega\,\mathrm{cm}$, and is extremely sensitive to the presence of impurities. The value of resistivity is not the best criterion defining a semiconductor, but rather the property that in pure enough samples the resistivity decreases with increasing temperature in a way that is exponential with the inverse temperature. Another property of semiconductors is that not only the electrons, but also the positive charge carriers (i.e., the holes) may be responsible for electrical conductivity, as shown by the sign of the Hall coefficient.

Semiconductors are characterized by an energy gap of the order of an electronvolt separating the highest occupied band (the valence band) from the lowest unoccupied one (the conduction band); the former may be populated by mobile holes (corresponding to missing electrons) and the latter by mobile electrons. In the following, the behavior of electrons and holes in homogeneous semiconductors will be described considering relevant electronic band structure properties, doping and equilibrium carrier concentration, mobility, and electrical conductivity.

Electron and Hole Pockets

The transport properties of semiconductors are dominated by carriers (electrons and holes introduced by doping or excited across the energy gap) which occupy the bottom of the conduction band or the top of the valence band. An accurate description of the electronic energy bands near these extrema is usually obtained employing the $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbative method. Then, the response of the carriers to external perturbations, such as electric and magnetic fields, slowly varying on the scale of the lattice constant can be dealt with by using the effective mass approximation.

The $\boldsymbol{k}\cdot\boldsymbol{p}$ method consists of a perturbative calculation of the electronic states with a Bloch wave vector $\boldsymbol{k}$ close to a given $\boldsymbol{k}_0$ point (usually of high symmetry) in the Brillouin zone (BZ) corresponding to a maximum in the valence band (hole pocket) or a minimum in the conduction band (electron pocket). For a nondegenerate band, as typical for electrons, a single-band approximation is usually employed, the interaction with all other bands being treated in perturbation theory up to second order in $(\boldsymbol{k}-\boldsymbol{k}_0)$. In this case, one gets simply

$$E(\boldsymbol{k}) = E(\boldsymbol{k}_0) + \frac{\hbar^2}{2}\,(\boldsymbol{k}-\boldsymbol{k}_0)\cdot\mathbf{m}^{-1}\cdot(\boldsymbol{k}-\boldsymbol{k}_0) \quad [1]$$

where the inverse effective mass tensor $\mathbf{m}^{-1}$ for the extremum of interest has been introduced. Finally, measuring the energy from $E(\boldsymbol{k}_0)$ and the wave vector from $\boldsymbol{k}_0$, and choosing the $\hat{1}$, $\hat{2}$, and $\hat{3}$ directions along the principal axes of $\mathbf{m}^{-1}$, the standard expression for ellipsoidal pockets is obtained:

$$E(\boldsymbol{k}) = \frac{\hbar^2}{2}\left(\frac{k_1^2}{m_1}+\frac{k_2^2}{m_2}+\frac{k_3^2}{m_3}\right) \quad [2]$$

Then, in the presence of slowly varying external perturbations, the dynamics of the electrons in a conduction band minimum can be described as that of free carriers (i.e., without considering the rapidly varying crystal potential) with a positive (anisotropic) effective mass and a negative charge $-e$. In particular, they will give rise to a current in the same direction as the applied electric field while moving in the opposite direction. Likewise, a hole pocket is associated with a maximum in the valence band and is characterized by negative values of the effective mass parameters appearing in eqn [2]. Such states, which in a pure crystal at zero temperature are completely occupied, may become empty and, upon the action of an applied electric field, give rise to a current that can be thought of as due to positively charged particles with a positive (anisotropic) effective mass. This is so because instead of considering the motion of the very large number of electrons which fill almost the entire valence band, it is convenient to focus on the dynamics of the few empty

states whose electrons are missing (i.e., the hole states). As the energy of a hole state increases when the energy of the corresponding missing electron decreases, in the hole picture used in the following, the sign of the effective mass parameters of a hole pocket will be taken as positive and at the same time the charge of a hole as $+e>0$. A hole moves in the same direction as the applied electric field and contributes to a current in the same direction.

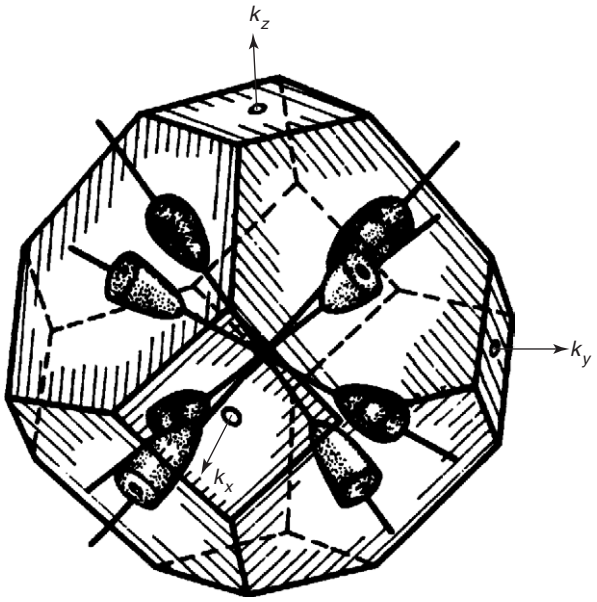


Figure 1 Electron pockets in Ge.

The symmetry of the dispersion relation of electron (or hole) pockets as well as the values of the corresponding effective mass parameters can be measured by cyclotron resonance, that is, the absorption of radio frequency radiation in the presence of an external magnetic field $\boldsymbol{B}$. The resonance frequency is $\omega_c = (eB/m^{\star}(\hat{B})c)$, with

$$m^{\star}(\hat{B}) = \sqrt{\frac{m_1 m_2 m_3}{m_1(\hat{B}\cdot\hat{1})^2 + m_2(\hat{B}\cdot\hat{1})^2 + m_3(\hat{B}\cdot\hat{3})^2}} \quad [3]$$

From the dependence of the observed absorption on the magnitude and direction of the magnetic field, the symmetry of the pockets of carriers as well as their effective mass parameters can be determined. Also magnetoresistance measurements are sensitive to the mass anisotropy and other band structure features, in particular from the strong dependence on the current and magnetic field directions with respect to the cubic axes the symmetry of the carrier pockets involved can be identified, similar to the case of cyclotron resonance measurements.

The conduction band of Ge has four electron pockets at the L point of the BZ (see **Figure 1**; the half pockets at $\boldsymbol{k}$ and $-\boldsymbol{k}$ are to be combined together); they are ellipsoids of revolution elongated along the $\langle 1\,1\,1\rangle$ directions. The values of the longitudinal mass $m_L = m_3 \simeq 1.6\,m$ and of the transverse mass $m_T = m_1 = m_2 \simeq 0.08\,m$ have been determined from cyclotron resonance measurements such as those shown in **Figure 2**: with $\boldsymbol{B}$ along

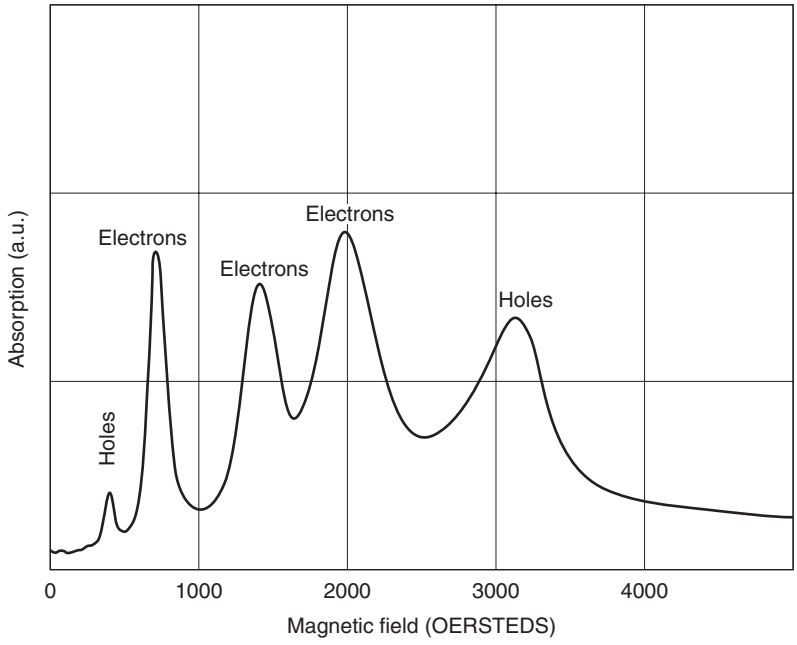


Figure 2 Cyclotron resonance in Ge (at a frequency of 24 GHz and a temperature of 4 K) with magnetic field in the (1 1 0) plane at 60° from the [0 0 1] axis. (From Dresselhaus G, Kip AF, and Kittel C (1955) *Physical Review* 98: 368.)

$[\sqrt{3}, \sqrt{3}, \sqrt{2}]/2$ the four $\langle 1\,1\,1 \rangle$ electron pockets split into three groups inequivalently oriented with respect to the magnetic field, each one of these giving rise to a distinct resonance according to eqn [3]. The conduction band of Si has six electron pockets along the $\langle 1\,0\,0 \rangle$ directions (see **Figure 3**); they are ellipsoids of revolution elongated along the same directions with a longitudinal mass $m_L = m_3 \simeq 1.0\ m$ and a transverse mass $m_T = m_1 = m_2 \simeq 0.2\ m$ (measured from cyclotron resonance; see **Figure 4**, where due to the choice of magnetic field orientation two distinct electronic resonances show up). Typical zinc blende semiconductors have a simple electron pocket at the center of the BZ (Γ point) with an isotropic effective mass; for instance, the electron effective mass of GaAs is $0.067\,m$.

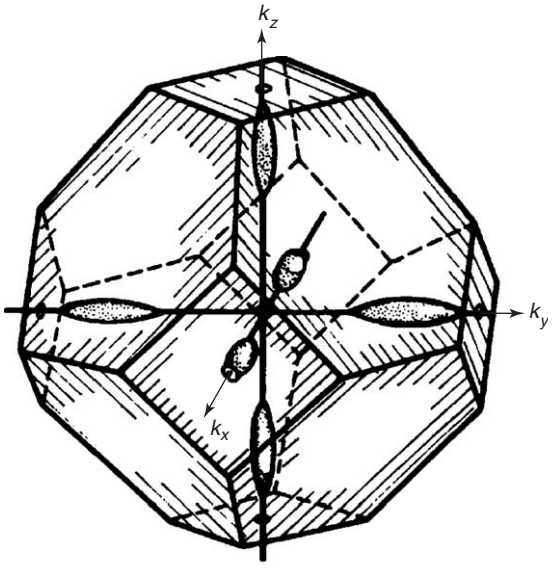


Figure 3 Electron pockets in Si.

The hole pockets in Si, Ge, and typical zinc blende semiconductors are at the center of the BZ and correspond to a threefold *p*-like degenerate band (without spin) that splits due to the spin–orbit interaction into an upper fourfold degenerate ($J = 3/2$) band and a lower twofold degenerate ($J = 1/2$) band (spin included). The former corresponds to the top of the valence band and the latter to the split-off valence band which, apart from spin, is nondegenerate. The split-off hole states can be described as a simple band with an isotropic effective mass $m_{so} \simeq 0.25\ m$ in Si, $m_{so} \simeq 0.08\ m$ in Ge, $m_{so} \simeq 0.2\ m$ in GaAs. In a rough approximation, the topmost hole states can be described as belonging to two distinct spherical pockets each characterized by an isotropic mass: the heavy hole pocket (with mass $m_{hh} \simeq 0.5\ m$ for Si, $m_{hh} \simeq 0.3\ m$ for Ge, $m_{hh} \simeq 0.5\ m$ for GaAs) and the light hole pocket (with mass $m_{lh} \simeq 0.16\ m$ for Si, $m_{lh} \simeq 0.044\ m$ for Ge, $m_{lh} \simeq 0.09\ m$ for GaAs). Also the hole masses have been determined from cyclotron resonance measurements such as those shown in **Figures 2** and **4**. A complete description of the hole states beyond this simple picture can only be given within a multiband $\boldsymbol{k} \cdot \boldsymbol{p}$ approach, in which the valence band degeneracy is fully accounted for.

Donors and Acceptors

The ability to control the content of impurities is a key issue in semiconductor technology: the concentration of undesired dopants in Si is typically kept below $10^{12}\,cm^{-3}$. A foreign atom may give rise to electronic states, more or less localized around the impurity site, with energies lying in the forbidden energy gap of the host material. The impurity levels may either be close to the conduction or valence band

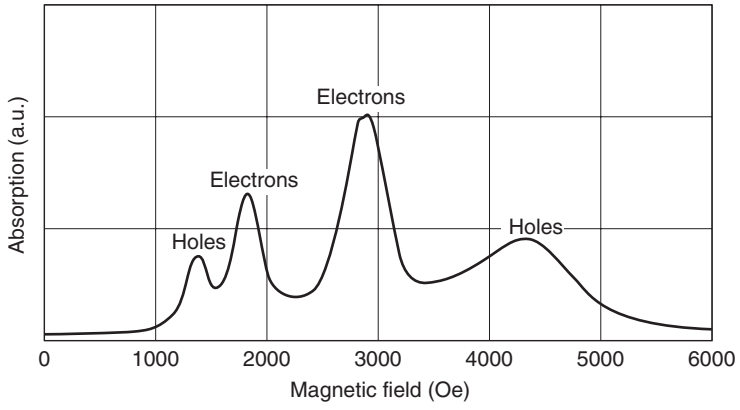


Figure 4 Cyclotron resonance in Si (at a frequency of 24 GHz and a temperature of 4 K) with magnetic field in the (1 1 0) plane at 30° from the [0 0 1] axis. (From Dresselhaus G, Kip AF, and Kittel C (1955) *Physical Review* 98: 368.)

edges (shallow donor or acceptor states) or to the middle of the gap (deep states). In the first case, they behave as dopants and dominate the transport properties providing extrinsic free carriers; in the second case, they mainly affect the recombination of free electron–hole pairs by trapping them at the same site.

Depending on their chemical valency, foreign atoms typically used as dopants may provide one extra electron or one extra hole (i.e., one missing electron) with respect to the host atoms they replace. For instance, a group V impurity, say P, substituting a group IV host atom, such as Si, provides an extra electron to the crystal (donor behavior), while a substitutional group III impurity, say B, at the same site would provide an extra hole (acceptor behavior). The extra hole or electron, with respect to the fully occupied valence band, is bound to the defect site by the long-range Coulomb potential due to the charge difference between the impurity and host nuclei. Because of the low effective masses ($m^\star \simeq 0.2\, m$) and the high dielectric constants ($\varepsilon \gtrsim 10$) typical of semiconductors, the corresponding localized states extend over many unit cells and may be described within the effective mass approximation as a solid-state analog of the hydrogen atom with scaled down Rydberg energy $R^\star = (m^\star e^4/2\hbar^2\varepsilon^2)$ and Bohr radius $a^\star = (\hbar^2\varepsilon/m^\star e^2)$ ($R^\star \simeq 20$ meV, $a^\star \simeq 50$ Å). As their binding energy is small, such impurity levels are shallow, that is, close to the bottom of the conduction band for the case of donors and close to the top of the valence band for the case of acceptors. Thus, they are easily thermally ionized providing free electrons to the conduction band (*n*-type doping) or free holes to the valence band (*p*-type doping), as discussed below.

Free Carrier Density

According to the Boltzmann distribution law (valid at typical carrier concentrations and temperatures, that is, excluding the case of quantum degeneracy), the density n of electrons in the conduction band and of holes p in the valence band can be expressed as

$$n(T,\mu) = N_c(T)\,\exp((\mu - E_c)/k_B T) \qquad [4]$$

$$p(T,\mu) = N_v(T)\,\exp((E_v - \mu)/k_B T) \qquad [5]$$

with a chemical potential μ well inside the energy gap between E_v and E_c (the top of the valence band and the bottom of the conduction band, respectively), being

$$N_c = \left(\frac{k_B T}{\pi\hbar^2}\right)^{3/2} \frac{(m_c^{(D)})^{3/2}}{\sqrt{2}}$$

$$N_v = \left(\frac{k_B T}{\pi\hbar^2}\right)^{3/2} \frac{(m_v^{(D)})^{3/2}}{\sqrt{2}}$$

where $m^{(D)}$ is the density of states effective mass given for a many-valley conduction band by an average over the equivalent valleys $m_c^{(D)} = g^{2/3}(m_T^2 m_L)^{1/3}$, g being the valley degeneracy (6 for Si, 4 for Ge), and for the valence band by an average over heavy and light holes $m_v^{(D)} = (m_{lh}^{3/2} + m_{hh}^{3/2})^{2/3}$. The densities of electrons and holes in thermal equilibrium at temperature T satisfy the law of mass action

$$np = N_c N_v e^{-E_g/k_B T} \qquad [6]$$

In a pure (intrinsic) semiconductor, the density of electrons n_i thermally excited into the conduction band equals the density of holes p_i left behind in the valence band and they are given by

$$n_i = p_i = \sqrt{N_c N_v}\,\exp(-E_g/2k_B T)$$

the intrinsic concentration of carriers is dominated by the exponential dependence on $E_g/(2k_B T)$. For instance, one obtains at room temperature, for Si, $E_g = 1.1$ eV, $N_c \simeq 3 \times 10^{19}$ cm^{-3}, $N_v \simeq 1 \times 10^{19}$ cm -3, $n_i = p_i \simeq 1.5 \times 10^{10}$ cm^{-3}; for Ge, $E_g = 0.66$ eV, $N_c \simeq 1 \times 10^{19}$ cm^{-3}, $N_v \simeq 6 \times 10^{18}$ cm^{-3}, $n_i = p_i \simeq 2.3\times 10^{13}$ cm^{-3}; for GaAs, $E_g = 1.43$ eV, $N_c \simeq 4.4\times 10^{17}$ cm^{-3}, $N_v \simeq 8 \times 10^{18}$ cm^{-3}, $n_i = p_i \simeq 2\times 10^{6}$ cm^{-3}. For the nondegenerate regime here considered, the intrinsic Fermi level (chemical potential) is given by $\mu_i = (E_c + E_v)/2 + (3/4)k_B T \ln(m_v^{(D)}/m_c^{(D)})$, and is close to the center of the gap (i.e., $(E_c + E_v)/2$).

In a doped semiconductor with a concentration of donors N_d and of acceptors N_a, if $N_d > N_a$ (*n*-type) the electron (majority carrier) density is approximately $n = N_d - N_a$ and the hole (minority carrier) density is correspondingly $p = (n_i p_i)/n$; if $N_d < N_a$ (*p*-type) similar formulas hold with the role of majority and minority carriers interchanged. In *n*-type semiconductors the chemical potential μ is shifted a bit closer to E_c with respect to μ_i, as given by

$$\frac{N_d - N_a}{2n_i} = \sinh\left(\frac{\mu - \mu_i}{k_B T}\right) \qquad [7]$$

the converse is true for *p*-type semiconductors. In the above, the temperature has been assumed to be high enough to ionize all the impurities, but not so high that the concentration of intrinsic carriers (i.e., those thermally excited across the energy gap) should be comparable to the majority carrier concentration. In this case, that is, the extrinsic regime, the carrier concentration is independent of temperature and determined by the doping level. At lower temperatures, the concentration of ionized impurities decreases because more and more electrons fall in the impurity bound states and the concentration of free carriers drops (freeze-out regime). The typical temperature dependence of the majority carrier concentration in a doped

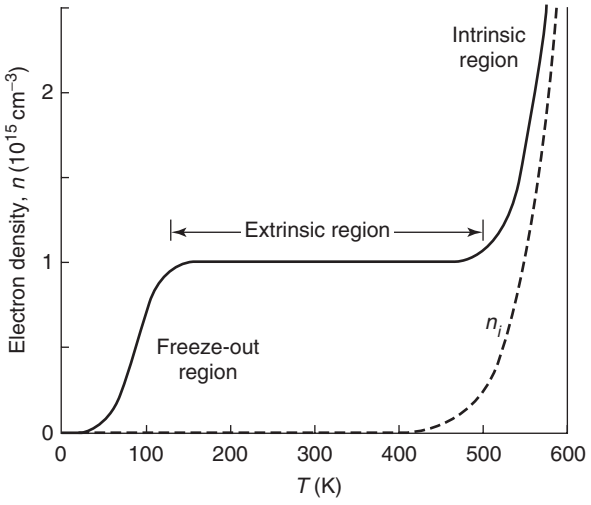


Figure 5 Temperature dependence of majority carrier concentration in *n*-doped Si ($N_d = 10^{15}$ cm^{-3}). The intrinsic carrier concentration is shown by the dashed line for comparison.

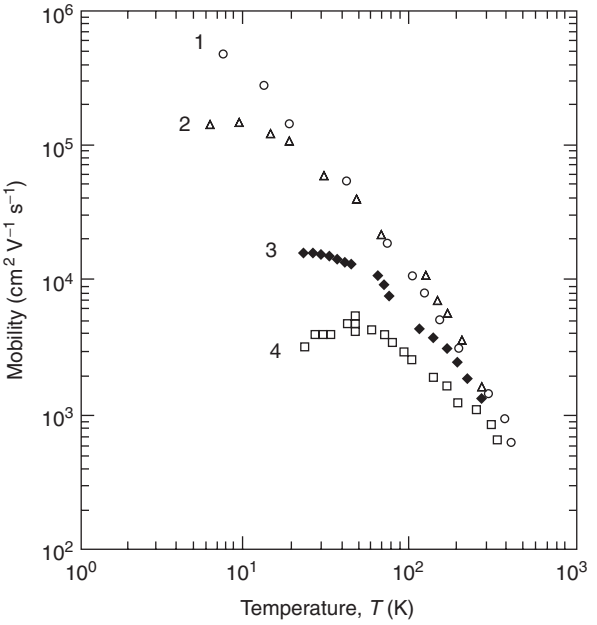


Figure 6 Temperature dependence of mobility in *n*-type Si at various doping levels: (1) $N_d < 10^{12}$ cm^{-3} (high purity), (2) $N_d < 10^{13}$ cm^{-3} (high purity), (3) $N_d = 1.6 \times 10^{16}$ cm^{-3}, (4) $N_d = 1.3 \times 10^{17}$ cm^{-3}. (From the database NSM maintained by the Ioffe Institute.)

semiconductor over a wide temperature range is depicted in **Figure 5**. The wide plateau (extrinsic regime) around room temperature shows how the carrier concentration can be suitably fixed by intentional doping, a characteristic semiconductor behavior of paramount importance for electronic applications.

Carrier Mobility and Electrical Conductivity

In a cubic semiconductor, the low-field conductivity tensor relating the current density $\boldsymbol{J}$ to the electric field $\boldsymbol{E}$ is a scalar $\boldsymbol{J} = \sigma \boldsymbol{E}$ given by

$$\sigma = e(n\mu_n + p\mu_p) \qquad [8]$$

where n and p are the densities of conduction band electrons and valence band holes, respectively, and μ_n and μ_p the corresponding mobilities, relating the average drift velocities of electrons or holes to the electric field, given by

$$\mu_{n,p} = (e/m_{n,p}^{(\sigma)})\tau_{n,p} \qquad [9]$$

where $m^{(\sigma)}$ is the appropriate effective mass and τ the appropriate momentum relaxation time as determined by the scattering mechanisms briefly discussed below. For a many-valley band extremum, such as for the electron pockets in Si and Ge, the conductivity effective mass to be used above is an average given by

$$1/m_n^{(\sigma)} = (2/m_T + 1/m_L)/3 \qquad [10]$$

and n is the total density of electrons equally distributed among the equivalent valleys. From the values of the electron transverse and longitudinal masses, one obtains, for Si, $m_n^{(\sigma)} \simeq 0.27\, m$ and, for Ge, $m_n^{(\sigma)} \simeq 0.12\, m$. When more than one electron or hole bands are populated (e.g., light and heavy hole bands) their contributions add up in the above expression for σ. The conductivity of a semiconductor, in contrast to that of a metal, varies strongly because of the dramatic dependence of the electron and hole densities on the temperature and the doping level; only a smoother variation is brought about by mobility changes. The carrier mobilities are related to the respective diffusion coefficients by the Einstein relation $eD_{n,p} = \mu_{n,p} k_B T$.

At room temperature, in pure or moderately doped elemental semiconductors, the mobility is limited mainly by phonon scattering, whereas, at lower temperatures or heavy doping levels, the ionized impurity scattering is dominant. In nonpolar semiconductors such as Si and Ge, the main electron–phonon scattering mechanism is due to longitudinal acoustic phonon deformation potential scattering with a scattering rate $1/\tau_{ac}$ depending on temperature like $T^{3/2}$. However, significant deviations from this behavior are observed, mostly due to the mixture of acoustic and optical phonon scattering. In zinc blende semiconductors such as GaAs, at room temperature the main scattering mechanism is due to polar longitudinal optical phonon scattering with a stronger dependence on temperature. The mobility values in pure materials at room temperature are, for Si,

$\mu_n \simeq 1500\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$, $\mu_p \simeq 500\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$; for Ge, $\mu_n \simeq 3900\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$, $\mu_p \simeq 1900\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$; for GaAs, $\mu_n \simeq 8500\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$, $\mu_p \simeq 400\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$. The impurity Coulomb scattering rate $1/\tau_{\text{im}}$, instead, is proportional to $T^{-3/2}N_{\text{im}}$, N_{im} being the density of ionized impurities. Experimentally, it is not easy to see the $T^{3/2}$ dependence of the impurity scattering limited mobility as the temperature range in which this mechanism is dominant is restricted at low temperatures by the carrier freeze-out and at high temperatures by the onset of phonon scattering. However, when more sources of scattering are competing the total resistivity is approximately given by the sum of the resistivities obtained from each single scattering mechanism at a time (Matthiessen's rule). As a result, the mobility of doped semiconductors may exhibit a maximum at the temperature corresponding to the crossover from impurity scattering to phonon scattering. Such a behavior is evident in **Figures 6** and **7** which show, respectively, the electron mobility in Si and GaAs versus temperature for various doping levels.

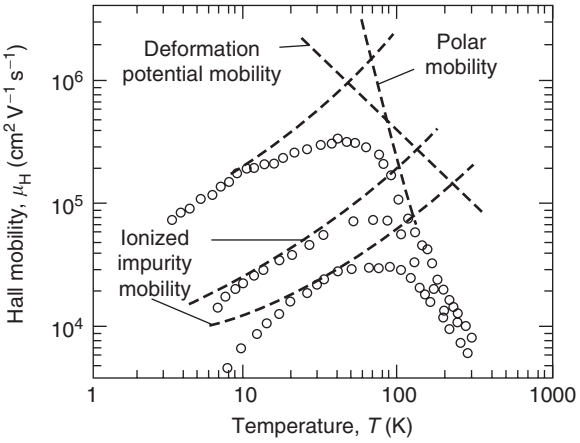


Figure 7 Temperature dependence of mobility in *n*-type GaAs at various doping levels: top curve $N_d = 5 \times 10^{14}\ \text{cm}^{-3}$, middle curve $N_d = 10^{15}\ \text{cm}^{-3}$, bottom curve $N_d = 5 \times 10^{15}\ \text{cm}^{-3}$. (From the database NSM maintained by the Ioffe Institute.)

The temperature dependence of the conductivity of Si at different doping levels is shown in **Figure 8**. The intrinsic regime is evident over most of the temperature range; the flattening of some of the curves at about room temperature corresponds to the extrinsic regime in which the conductivity is determined by

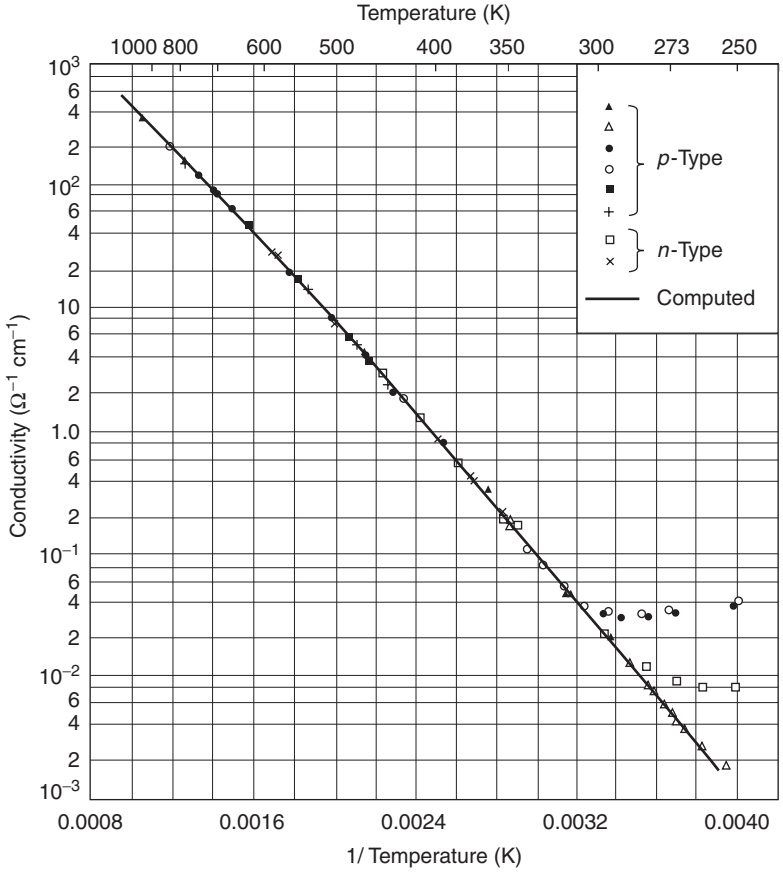


Figure 8 Temperature dependence of conductivity in Si at various doping levels. (From Morin EJ and Maita JP (1954) *Physical Review* 94: 1525.)

the doping level. The temperature dependence of σ is dominated by that of the carrier concentration in the intrinsic regime, whereas its weaker variation in the extrinsic regime is due to the temperature dependence of the mobility.

Measuring the Hall coefficient is a widely used method to characterize the concentration (and type) of free carriers; the Hall coefficient R_H is given by

$$R_H = \frac{p\mu_p^2 - n\mu_n^2}{ce(p\mu_p + n\mu_n)^2} \qquad [11]$$

In weak fields, the mobilities appearing in the numerator of the above expression are in general to be multiplied by corrective factors of order unity (i.e., the ratios of Hall mobilities to drift mobilities: μ_H/μ) which depend on the scattering mechanism and the band structure details. Notice that in *p*-type samples the sign of R_H is positive (i.e., opposite to that of *n*-type samples).

So far the low-field conductivity corresponding to the ohmic regime has been considered. With increasing field strength, the free carrier drift velocity approaches their thermal velocity and hot carrier effects become important. In particular, in Si and Ge, the high-field mobility (limited mainly by inelastic phonon scattering) eventually decreases linearly with the electric field and the drift velocity saturates to values close to $10^7\,\mathrm{cm\,s^{-1}}$ at electric fields $\sim 10\,\mathrm{kV\,cm^{-1}}$. In the presence of hot carriers, the equivalent valleys need not be equally populated and, as a consequence, even in a cubic material like *n*-type Ge, the current need not be parallel to the electric field. Moreover, also inequivalent, higher-energy valleys may become populated: for instance, dramatic changes in the Hall coefficient of *n*-type Ge at electric fields higher than $1\,\mathrm{kV\,cm^{-1}}$ are brought about by the transfer of electrons from the $\langle 1\,1\,1 \rangle$ lowest minima to the $\langle 1\,0\,0 \rangle$ higher minima, the Hall coefficient is dominated by the contribution of the $\langle 1\,1\,1 \rangle$ high-mobility valleys and starts to increase as their population starts to decrease. In several compound semiconductors such as GaAs, such a carrier transfer to inequivalent valleys is responsible for a decrease of drift velocity leading to negative differential resistance (at a field value of about $3\,\mathrm{kV\,cm^{-1}}$ in GaAs, as shown in **Figure 9**) and also to peculiar current oscillations (Gunn effect) at microwave frequencies. In the presence of hot holes, complex features of the degenerate valence band become important, such as the non-parabolicity of the hole dispersion (i.e., the variation with kinetic energy of the effective mass).

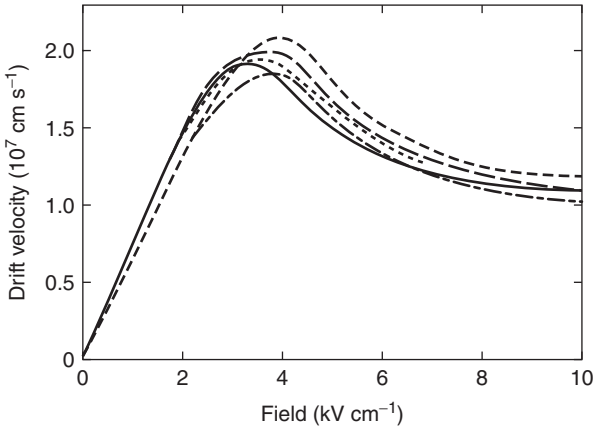


Figure 9 High-field electron drift velocity in GaAs at room temperature: the solid line is calculated, the other ones are experimental. (From the database NSM maintained by the Ioffe Institute.)

Finally, increasing the electric field strength beyond the saturation regime, a sudden "run-away" growth of the conductivity is eventually observed (breakdown). This is due to the avalanche multiplication of the number of free carriers: the field is so strong that a carrier can gain enough kinetic energy between two collisions to knock out an electron from the valence to the conduction band. Typical breakdown electric fields are $10^5\,\mathrm{V\,cm^{-1}}$ for Ge, $3 \times 10^5\,\mathrm{V\,cm^{-1}}$ for Si and $4 \times 10^5\,\mathrm{V\,cm^{-1}}$ for GaAs.

See also: Elemental Semiconductors, Electronic States of; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Devices; Transport in Two-Dimensional Semiconductors.

PACS: 71.20.−b; 71.20.Mq; 71.20.Nr; 72.20.−i; 72.80.Cw; 72.80.Ey

Further Reading

Bassani F and Pastori Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Landolt-Börnstein (ed.) (1982) *Physics of Group IV Elements and III–V Compounds*, vol. 3–17a. Berlin: Springer.

Seeger K (1991) *Semiconductor Physics*. Berlin: Springer.

Singh J (1994) *Semiconductor Devices*. New York: McGraw-Hill.

Sze SM (1985) *Semiconductor Devices*. New York: Wiley.

Yu PY and Cardona M (1995) *Fundamentals of Semiconductors*. Berlin: Springer.

Electrons in Periodic Potentials with Magnetic Fields

J Bellissard, Georgia Institute of Technology, Atlanta, GA, USA

Introduction

A magnetic field applied to the electron gas of a solid breaks the time-reversal symmetry giving access to information at the atomic scale that is inaccessible otherwise. The de Haas–van Alphen oscillations of the magnetic susceptibility or the Shubnikov–de Haas oscillations of the conductivity are among the classical experimental tools used in high-magnetic-field facilities worldwide to investigate the structure of the Fermi surface. The Hall effect gives access to the type of charge carriers, electrons, or holes. The cyclotron resonance frequency can also be measured through radiowave oscillations, which give access to the electron effective mass or various relaxation times. Finally, NMR experiments give local information on the local density of states.

In most situations, the magnetic field can be considered as uniform at least on the scale of the electron mean free path. In much the same way, it is only at frequencies in the THz range and beyond, namely, of the order of the inverse of the inelastic collision time, that the magnetic field must be considered as non-constant in time. Hence, a magnetic field of up to 500 T applied during a microsecond (explosion) can still be considered as constant in time.

The determination of the energy spectrum of independent electrons in a periodic potential and submitted to a magnetic field was immediately recognized as a difficult task. This problem was investigated more systematically from the early 1950s onward. As a consequence of the Peierls approach, Harper designed, in 1955, a simplified model which is the paradigm of this problem (see eqn [6]). The calculation of the band spectrum of this model was investigated in several papers later on, until Hofstadter, in 1976, made a numerical computation and showed that it is a remarkable Cantor set (see **Figure 1**), at a time when even the word "fractal" had not been coined by Mandelbrot. Since then, this "Hofstadter butterfly" has never ceased to fascinate physicists and mathematicians, even though it is still controversial whether it can be fully observed in realistic experiments. This butterfly has zero Lebesgue measure, namely, it cannot fly!

Low Magnetic Fields

In the simplest approximation, the electron gas in a metal can be seen as a perfect Fermi gas of independent, free electrons (or holes) with mass m_* and charge $q = \mp e$. If $\boldsymbol{R} = (R_1, R_2, R_3)$ are the position operators and $\boldsymbol{P} = (P_1, P_2, P_3)$ the momentum operators, then $[R_i, P_j] = \mathrm{i}\hbar\delta_{ij}$. The free Hamiltonian is given by $H = \boldsymbol{P}^2/2m_*$. A uniform constant magnetic field $\boldsymbol{B} = (0, 0, B)$ can be added through the substitution

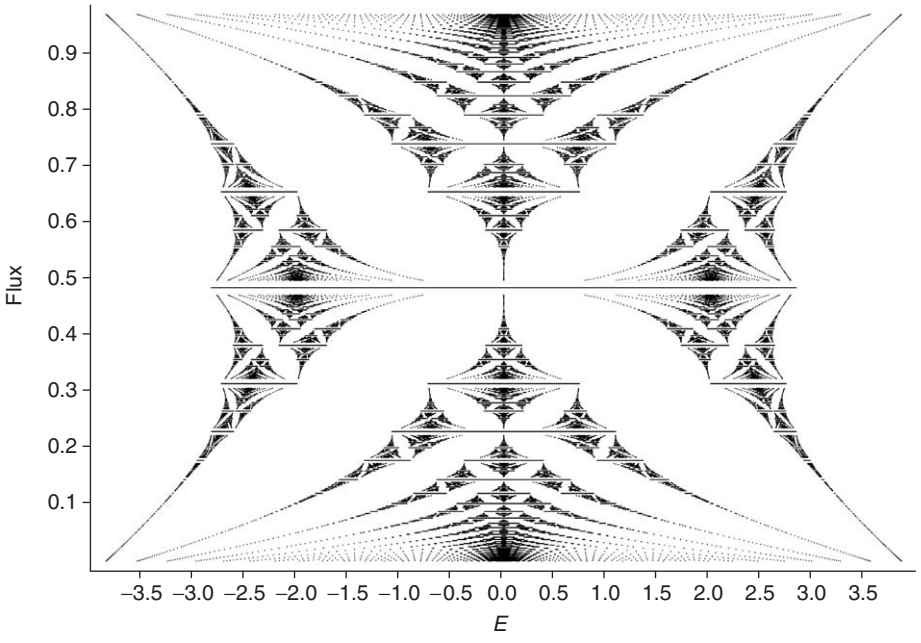


Figure 1 Band spectrum for the Harper model describing a 2D Bloch electron in a uniform magnetic field.

$\boldsymbol{P} \mapsto \boldsymbol{P} - q\boldsymbol{A}$ if $\boldsymbol{A} = (A_1, A_2, A_3)$ is the vector potential, depending upon the position and defined by

$$\partial_1 A_2 - \partial_2 A_1 = B$$
$$\partial_2 A_3 - \partial_3 A_2 = 0 = \partial_3 A_1 - \partial_1 A_3$$

The quasimomentum operator is defined by $\boldsymbol{K} = (\boldsymbol{P} - q\boldsymbol{A})/\hbar = (K_1, K_2, K_3)$ and satisfies the following commutation rules:

$$[K_1, K_2] = -\mathrm{i}\frac{qB}{\hbar}, \qquad [K_2, K_3] = [K_1, K_3] = 0 \quad [1]$$

The Hamiltonian for the free particle in a magnetic field becomes

$$\hat{H}_B = \frac{\hbar^2}{2m_*}(K_1^2 + K_2^2 + K_3^2) = H_B + \frac{\hbar^2}{2m_*}K_3^2$$

Since K_3 commutes with K_1 and K_2, it is sufficient to compute the spectrum of H_B. H_B is called the Landau Hamiltonian and represents the electronic motion in the plane perpendicular to $\boldsymbol{B}$. Thanks to the commutation rules [1], H_B can be seen as a harmonic oscillator with effective Planck's constant $\hbar_{\mathrm{eff}} = qB/\hbar$. In particular, its spectrum consists of isolated eigenvalues (called Landau levels)

$$E_n = \hbar\omega_{\mathrm{c}}(n + 1/2), \quad n = 0, 1, 2, \ldots, \quad \omega_{\mathrm{c}} = qB/m_*$$

ω_{c} is called the cyclotron frequency. Thanks to the translation invariance of the problem, each such eigenvalue has, however, an infinite multiplicity given by 2π times the number of flux quanta per unit area, namely, $eB/\hbar$.

If the electron gas is now placed in a periodic potential $V(\boldsymbol{R})$, such as what happens in a real metal, the one-particle Hamiltonian is $\hat{H}_B + V$. Then there are three vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ in $\mathbb{R}^3$ such that $V(\boldsymbol{R} + \boldsymbol{a}) = V(\boldsymbol{R})$ whenever $\boldsymbol{a} = m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2 + m_3\boldsymbol{a}_3$ and $m_i \in \mathbb{Z}$. If $B = 0$, the Bloch theory shows that the energy spectrum is made of bands described by functions $E_j(\boldsymbol{k})$ depending upon the quasimomentum $\boldsymbol{k}$ and periodic with respect to the reciprocal lattice defined as the set of $\boldsymbol{b} \in \mathbb{R}^3$ such that $\boldsymbol{a}_i \cdot \boldsymbol{b} \in 2\pi\mathbb{Z}$ for $i = 1, 2, 3$. Here j is the band index and runs in a countable set. If $B \neq 0$ is small enough, the effective Hamiltonian describing electrons can be obtained from the band functions by substituting $\boldsymbol{K}$ to $\boldsymbol{k}$ (Peierls substitution). Using the periodicity, each band function can be expanded into a Fourier series

$$E_j(\boldsymbol{k}) = \sum_{m \in \mathbb{Z}^3} E_{j,m} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2 + m_3\boldsymbol{a}_3)}$$
$$\Rightarrow H_{\mathrm{Peierls}} = \sum_{m \in \mathbb{Z}^3} E_{j,m} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot(m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2 + m_3\boldsymbol{a}_3)} \quad [2]$$

Actually this substitution is only the lowest order contribution to an expansion in powers of B, but the following theorem holds.

Theorem 1. *Let $\hat{H}_{B=0} + V$ have an interval $I \subset \mathbb{R}$ in its spectrum that is disconnected from the rest I^c of the spectrum and let I be generated by a finite family $E_1(\boldsymbol{k}), \ldots, E_N(\boldsymbol{k})$ of band functions. Then as $B \neq 0$ is small enough, the gaps separating I from I^c do not close. Moreover, the spectrum of $\hat{H}_B + V$ inside this region is described by an effective operator H_{Peierls} of the form given by eqn [2] where the Fourier coefficients are now $N \times N$ matrices depending continuously on the magnetic field.*

This theorem gives an effective Hamiltonian in terms of a noncommutative Fourier expansion using the three unitary operators

$$U_i = \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{a}_1} = (U_i^{-1})^*$$
$$U_i U_j = \mathrm{e}^{\mathrm{i}\theta_{ij}} U_j U_i$$
$$\theta_{ij} = -\theta_{ji} = 2\pi\frac{e}{h}\,\boldsymbol{B}\cdot\boldsymbol{a}_i \times \boldsymbol{a}_j$$

The angle θ_{ij} represents 2π times the ratio of the magnetic flux through the parallelogram defined by vectors $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ by the flux quantum $\phi_0 = h/e$. To complete the analogy with Fourier series, $\mathscr{T}$ may be the linear map defined on the set of polynomials in the U_i's such that

$$\mathscr{T}(U_1^{m_1} U_2^{m_2} U_3^{m_3}) = 0 \quad \text{if } |m_1| + |m_2| + |m_3| \neq 0$$
$$\mathscr{T}(\mathbf{1}) = 1 \quad [3]$$

It can be checked that if A, B are two such polynomials, then $\mathscr{T}(AB) = \mathscr{T}(BA)$ (trace property). Moreover, it can be shown that $\mathscr{T}$ coincides with the trace per unit volume. Clearly, this map generalizes the integration over $\boldsymbol{k}$.

For most values (modulus and directions) of the magnetic field $\boldsymbol{B}$, the three angles θ_{ij} are incommensurate so that the algebra generated by the U_i's becomes simple, namely, it cannot be decomposed easily, making the explicit computation of the energy spectrum a very difficult problem. However, the $\theta_{ij}/2\pi$ can be seen as the dimensionless effective Planck's constant and for a realistic crystal with $|\boldsymbol{a}_i| = 1$ Å and a magnetic field $B = 1$ T, their value is $O(10^{-5})$, thus justifying the use of semiclassical methods.

Semiclassical Methods and Magnetic Oscillations

The semiclassical methods go back to the early days of the old theory of quanta, with the Bohr–Sommerfeld quantization formula. The use of

such methods for electrons in magnetic fields was the basis of the Onsager argument to explain the de Haas–van Alphen (dHvA) and Shubnikov–de Hass (SdH) oscillations.

As explained in the section "Low magnetic fields," the magnetic field B quantizes the quasimomentum space with an effective Planck constant proportional to B. Moreover, if a periodic potential is added, the relevant dimensionless Planck constant is so small as to make the problem semiclassical. In particular, the Heisenberg equations of motions for the effective Peierls Hamiltonian [2] give

$$\frac{\mathrm{d}\boldsymbol{K}}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}[H_{\mathrm{Peierls}}, \boldsymbol{K}]$$

Since K_3 commutes with K_1, K_2, it follows that $\mathrm{d}K_3/\mathrm{d}t = 0$, so that the motion takes place in the plane perpendicular to $\boldsymbol{B}$. Moreover, the commutation rules [1] lead to the following approximate equation of motion:

$$\frac{\mathrm{d}k_1}{\mathrm{d}t} = -\frac{qB}{\hbar^2}\frac{\partial E_j}{\partial k_2}, \quad \frac{\mathrm{d}k_2}{\mathrm{d}t} = \frac{qB}{\hbar^2}\frac{\partial E_j}{\partial k_1}, \frac{\mathrm{d}k_3}{\mathrm{d}t} = 0$$

implying that $E_j(\boldsymbol{k})$ is conserved. Thus, the classical orbit associated with conduction electrons lies on the Fermi surface $E_j = E_{\mathrm{F}}$ in a plane perpendicular to the magnetic field corresponding to $k_3 = \boldsymbol{k}\cdot\boldsymbol{B} = \mathrm{const}$. Each such plane plays the role of a phase space for one degree of freedom, where k_2 is equivalent to the position and k_1 to the momentum, while the band function $E(k_1, k_2)$ plays the role of the Hamiltonian. Thanks to the Bohr–Sommerfeld quantization formula, only orbits with quantized action survive quantum interferences, leading to

$$\oint_{\mathrm{orb}} k_1\, \mathrm{d}k_2 = -2\pi e\frac{B}{\hbar}(n+v)$$
$$n\in\mathbb{Z}, \quad v = \text{Maslov index} \qquad [4]$$

where the contour integral is computed on the classical orbit. By the Stokes formula, the LHS is the area A enclosed by the classical orbit. The usual argument is now to compute the free energy F of such an electron gas. It should be noted that it takes into account these orbits with all possible values of k_3 provided $\boldsymbol{k}$ stays inside the Fermi surface. The Bohr–Sommerfeld condition [4] selects cylinders indexed by n with areas depending on the magnetic field (see **Figure 2**).

As the magnetic field increases, the radius of each such cylinder increases as well, until it gets out of the Fermi surface. Then the cylinder with index n has an area equal to the maximum area $A_{\max}$ of a section of the Fermi surface with a plane perpendicular to the magnetic field. As B increases again, the cylinder with index $n-1$ dominates, leading to oscillations of the free energy with periods given by the Onsager formula

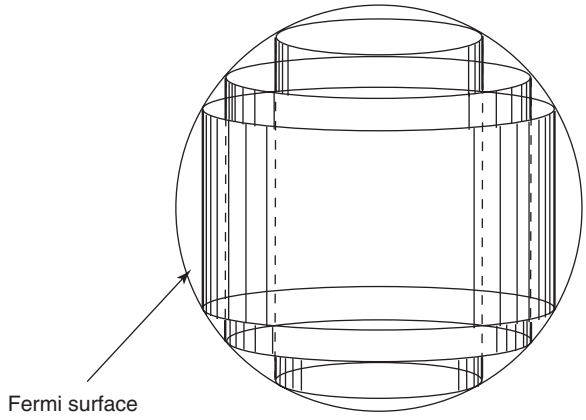


Figure 2 Landau cylinders inside the Fermi surface.

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar A_{\max}}$$

Since the magnetization is given by the thermodynamical relation $M = \mu_0(\partial F/\partial B)\mu$, it also exhibits oscillations at the same period and this is what experiments measure, giving access to the size of the Fermi surface. Extensions of this theory (the Lifschitz and Kosevich formula, and the magnetic breakdown) have been discussed by Shoenberg.

Strong Magnetic Fields in 2D

If the magnetic field is oriented along one of the crystal axes, the operator U_3 commutes with the others, so the model can be reduced to the two-dimensional (2D) plane perpendicular to $\boldsymbol{B}$. On the other hand, modern devices which are effectively 2D are currently available in labs. Such is the case for interfaces between the semiconductor and the oxide of an MOSFET or between the two parts of a heterojunction used in mesoscale or nanoscale technology.

Then, if only one band contributes, two unitaries suffice to express the effective Hamiltonian as

$$H = \sum_{m_1,m_2} h_{m_1,m_2}\mathrm{e}^{\mathrm{i}\theta m_1 m_2/2}U_1^{m_1}U_2^{m_2}$$
$$U_1U_2 = \mathrm{e}^{\mathrm{i}\theta}U_2U_1 \qquad [5]$$

where $\bar{h}_{m_1,m_2} = h_{-m_1,-m_2}$ (self-adjointness of H) are called the Fourier coefficients of H. The simplest example is the almost-Mathieu model (if $t_1 \neq t_2$) or Harper model (if $t_1 = t_2$) given by

$$H_{\mathrm{Harper}} = t_1(U_1 + U_1^{-1}) + t_2(U_2 + U_2^{-1}) \qquad [6]$$

where t_1, $t_2 > 0$ are called the hopping parameters. A numerical calculation of the band spectrum is

possible if $\theta = 2\pi p/q$, where p/q is a fraction. Here is the recipe:

1. U_1 and U_2 are replaced by $U_i(k_1, k_2) = e^{ik_i}u_i$, where

$$u_1 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{2i\pi p/q} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & e^{2i\pi(p-1)p/q} \end{bmatrix}$$

   Then H becomes a $q \times q$ Hermitian matrix depending periodically on $k = (k_1, k_2)$ with period 2π. Then k is called the quasimomentum.
2. For each value of (k_1, k_2) the eigenvalues $e_1(k_1, k_2) \leqslant \cdots \leqslant e_q(k_1, k_2)$ of $H(k_1, k_2)$ are computed. The spectrum is obtained as the union of all these values. In practice, the e_j's are computed for only a finite number of values of (k_1, k_2). Then, as in the case of the band theory, the symmetries of the model to choose the relevant (k_1, k_2)'s are used. In the Harper or in the almost-Mathieu models, the values $(k_1, k_2) = (0, 0)$ or (π, π) give the band edges.

Figure 1 represents the energy spectrum of the Harper model (see eqn [6]) and was obtained for the first time in 1976 by D R Hofstadter in his PhD thesis as a numerical calculation following the previous algorithm. The vertical axis represents the values of $\theta/2\pi$ between 0 and 1, while the horizontal one represents the values of the energy in units of t (here $t_1 = t_2$), which then varies between -4 and $+4$. The picture obtained suggests that the gap edges are continuous with respect to θ even though a closer look shows that these functions have a discontinuous derivative at each rational point. This has been proved rigorously by J Bellissard in 1994.

Theorem 2. *Let H be given by eqn [5] with Fourier coefficients depending smoothly on θ and decreasing fast enough as (m_1, m_2) increase to infinity. Then the gap edges of the spectrum of H are continuous functions of θ. Moreover, these functions are differentiable almost everywhere. At each rational value of θ, the right and left derivatives of such functions exist but are not necessarily equal. The discontinuity of the derivative at a rational point p/q at the edge of the band j, represented by the band function $e_j(k_1, k_2)$, can be computed through the following formula:*

$$\Delta = |\det \partial_\mu \partial_\nu e_j(k^{(j)})|^{1/2}$$

where $k^{(j)} = (k_1^{(j)} k_2^{(j)})$ is the value of the quasimomentum corresponding to the gap edge of the band j.

In the specific case of the almost-Mathieu operator (see eqn [6]), additional properties have been proved over the years by several mathematicians, prominent among them being B Simon, Y Last, S Jitomirskaya, and J Puig.

Theorem 3. *Let H be given by eqn [6]. Then*

1. *for all irrational values of $\theta/2\pi$ and for $t_1 \neq t_2$, the Lebesgue measure of the spectrum of H is equal to $4|t_1 - t_2|$ (Aubry's conjecture);*
2. *for almost all irrational values of $\theta/2\pi$ and for $t_1 = t_2$, the spectrum of H has zero Lebesgue measure. For a large set of such θ's, the Hausdorff dimension of the spectrum is less than or equal to* 1/2; and
3. *if $\theta/2\pi$ satisfies a* diophantine condition, *then the spectrum of H is a Cantor set (i.e., its complement is a dense open set and it has no isolated points).*

Let $\mathcal{T}$ be the trace defined in eqn [3]. Then, the integrated density of states (IDoS) $\mathcal{N}(E)$ is defined as the number of eigenvalues of H per unit volume smaller than or equal to E. The Shubin formula asserts that

$$\mathcal{N}(E) = \mathcal{T}(\chi(H \leqslant E))$$

where $\chi(H \leqslant E)$ is the projection on the space spanned by the eigenstates of H with eigenvalues smaller than or equal to E. Then the following theorem holds.

Theorem 4. *Let H be given by eqn [5]. Then*

1. *Its IDoS is a nondecreasing function of E varying from 0 to 1. It is constant of spectral gaps and its values on spectral gaps are given by the fractional part of $n\theta/2\pi$ from some integer n* (gap labeling theorem).
2. *If H is the almost-Mathieu model and if $\theta/2\pi$ is irrational. then for any integer n there is a gap with IDoS equal to the fractional part of $n\theta/2\pi$.*

The previous class of models can be seen as a tight-binding representation for the effective Hamiltonian representing the charge carrier motion on a perfect

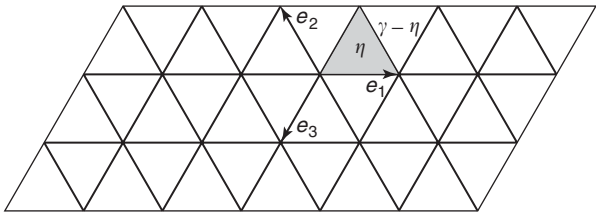


Figure 3 Triangular lattice, its basis, and two types of triangles.

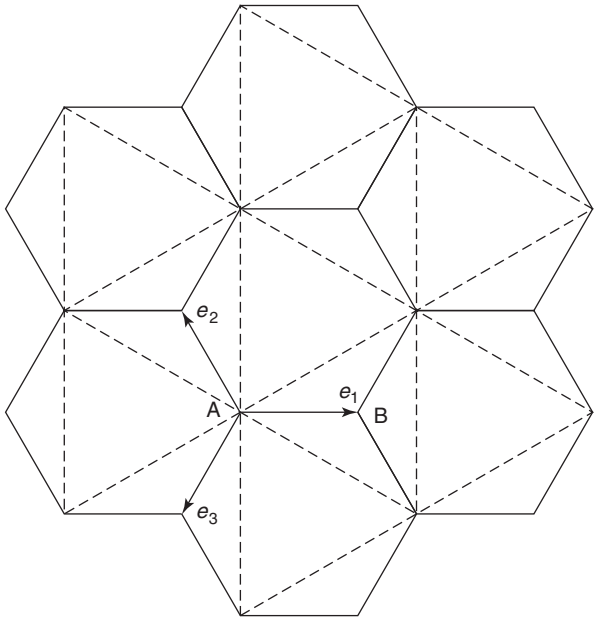


Figure 4 The honeycomb and its basis.

2D lattice. Other models can be constructed by considering that U_1 and U_2 represent translation along the basis vectors of a 2D crystal. Examples of models are given below.

1. The Harper model on a triangular lattice: there are three translations U_1, U_2, U_3 with $U_1U_2U_3 = e^{2i\pi\eta}$ and $U_1^{-1}U_2^{-1}\ U_3^{-1} = e^{2i\pi(\gamma-\eta)}$, if η and $\gamma-\eta$ are the dimensionless fluxes in the two types of triangles (see **Figure 3**). The Hamiltonian becomes

$$H_\Delta = t(U_1 + U_1^{-1} + U_2 + U_2^{-1} + U_3 + U_3^{-1})$$

2. The corresponding model for the honeycomb lattice takes into account the existence of two sublattices A and B (see **Figure 4**). Each site in A has its neighbors in B in directions $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, where these three vectors add up to zero and make an angle of 120° and vice versa. Hence, a wave packet can be seen as a spinor with components $|\psi_A\rangle$ and $|\psi_B\rangle$ and the Hamiltonian becomes a 2×2 matrix with coefficients depending on the U_i's as follows:

$$H_{\text{hexa.}} = t\begin{bmatrix} 0 & U_1+U_2+U_3 \\ U_1^{-1}+U_2^{-1}+U_3^{-1} & 0 \end{bmatrix}$$

Here, γ represents the dimensionless flux through the cell generated by $\boldsymbol{e}_1, \boldsymbol{e}_2$ namely, it is the third of the flux through the unit cell given by a hexagon and $\eta = \gamma/2$.

Various compounds admit a honeycomb structure, such as graphite monolayers built on the surface of silicon carbide SiC. More generally for quasi-2D systems, the previous construction can be used to build the effective Hamiltonian. It has been applied, for instance, to the magnetic breakdown observed in organic conductors of type (BEDT–TTF) which are mainly a stack of weakly coupled 2D layers.

Semiclassical Methods at Strong Magnetic Field

It is remarkable then that semiclassical methods could be used at strong magnetic fields as well. For indeed, the stronger the magnetic field the bigger $\hbar_{\text{eff}}$. The work by Helffer & Sjöstrand in the eighties, and by Rammal & Bellissard in the 1990s on the Harper model and its extensions showed the method.

If θ is close to a rational multiple of 2π, say $\theta = 2\pi p/q + \delta$, with $\delta > 0$ small, then the two unitaries U_1 and U_2 can be taken as $U_i = e^{ik_i}u_i \otimes \hat{U}_i$ with $\hat{U}_1\hat{U}_2 = e^{i\delta}\hat{U}_2\hat{U}_1$ implying $U_1U_2 = e^{i\theta}U_2U_1$. These unitary operators are $q\times q$ matrices with operator coefficients. The phase factor e^{ik_i} can be chosen arbitrarily, without changing the spectrum. This substitution is the rigorous version of the Peierls one. The following representation is used

$$\hat{U}_i = e^{i\sqrt{\delta}K_i}, \quad K_i = K_i^*, \ [K_1, K_2] = -i\mathbf{1}$$

Then using the Feshback method (also called Schur complement or Green function method), the $q\times q$ Hamiltonian H is transformed into a 1×1 effective one given by

$$H_{\text{eff}}(z) = \langle j|H|j\rangle + \left\langle j\middle| HQ_j\frac{1}{z - Q_jHQ_j}Q_jH\middle|j\right\rangle,$$
$$Q_j = \mathbf{1}_q - |j\rangle\langle j|$$

For $\delta = 0$, the Hamiltonian becomes a $q\times q$-matrix valued function of k denoted by $\mathscr{H}$, while for $\delta \neq 0$ its matrix elements depend on $\hat{U}_1, \hat{U}_2$. It is possible to expand the Hamiltonian in powers of $\sqrt{\delta}$, leading to a generalization of the semiclassical expansion. However here, the Hamiltonian being a $q\times q$ matrix, some

complication occurs depending upon what point of the spectrum is considered. The simplest expansion corresponds to expanding near a band edge. Then $k=(k_1, k_2)$ is chosen to correspond to this band edge in the eigenvalue e_j of the Hamiltonian at $\delta=0$. Moreover, if $|j\rangle$ is the q-dimensional eigenvector of the Hamiltonian for eigenvalue e_j at k, the eigenvalue equation is then $E|\psi\rangle = H_{\text{eff}}(E)|\psi\rangle$, which is nonlinear in E. Expanding this operator in powers of $\sqrt{\delta}$ to the lowest order gives again a harmonic oscillator as the lowest-order contribution which can be interpreted as a Landau Hamiltonian.

Using the Landau spectrum, this gives eigenvalues in the following form (see **Figures 5** and **6**):

$$E_{n,j}(\delta) = e_j(k) + |\delta|(n+1/2)(\det D^2 e_j(k))^{1/2} + \delta \frac{\partial e_j(k)}{\partial \delta} + \delta\sigma + O(\delta^{3/2})$$

where

$$\sigma = \frac{\mathrm{i}}{2}\sum_{l:\neq j} \frac{(<_j|\partial_1 \mathscr{H}|l\rangle\langle l|\partial_2 \mathscr{H}|j\rangle - \langle j|\partial_2 \mathscr{H}|l\rangle\langle l|\partial_1 \mathscr{H}|j\rangle)}{e_j - e_l}$$

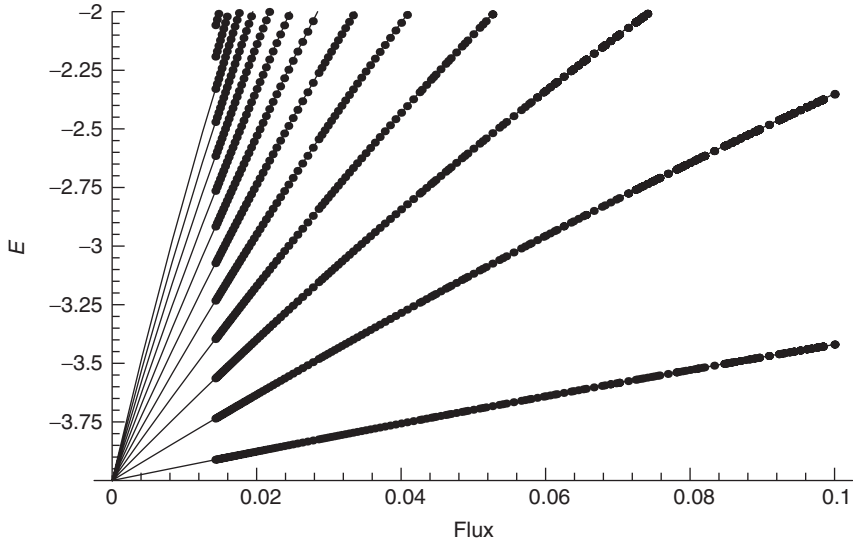


Figure 5 Comparison between semiclassical calculations (continuous lines) and exact numerical spectrum (points) for Landau levels in the Harper model [6] on the square lattice near the band edge at $\theta=0$.

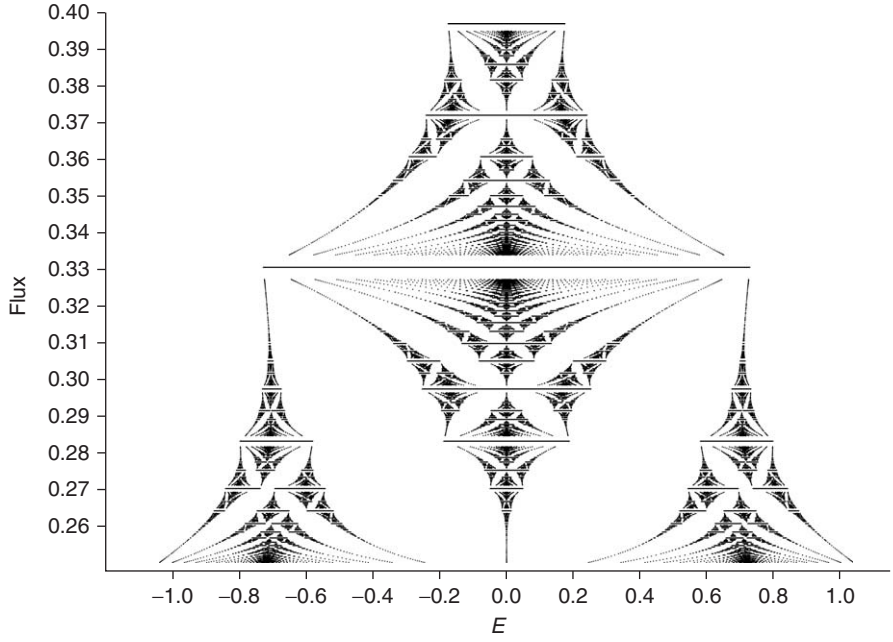


Figure 6 Discontinuity of derivatives of the band edge near the magnetic flux $\alpha=\theta/2\pi\approx 1/3$ for the Harper model on the square lattice.

The very same formula holds if $\delta < 0$ using a similar method. This formula gives rise to the discontinuity of the derivative of the band edges as stated in Theorem 2. It is remarkable that a systematic higher order expansion gives right corrections to the Landau levels, as long as the level width is ignored.

If two bands are touching (see **Figure 7**) for $\delta = 0$, the previous analysis fails. The degeneracy of the eigenvalues e_j and say, e_{j-1} at k for $\delta = 0$ requires using a degenerate perturbation theory. The lowest order in the $\sqrt{\delta}$ expansion is given by a 2D Dirac operator of the form

$$H_{\text{Dirac}} = C, \sqrt{\delta}\begin{bmatrix} 0 & K_1 + \mathrm{i}K_2 \\ K_1 - \mathrm{i}K_2 & 0 \end{bmatrix} + O(\delta)$$
$$\Rightarrow E_n^{\pm} = \pm C\sqrt{n\delta}(1 + O(\delta))$$

where C is a model-dependent constant, the $E_n^{\pm}$'s are the eigenvalues and $n = 0, 1, 2, \ldots$ is an integer (see **Figure 8**).

As can be seen in **Figures 1, 6,** and **8,** the Landau–Dirac level broadens into a band as δ gets too large. This is a general phenomenon due to tunneling in the phase space. A spectacular illustration of the tunneling effect is the braiding seen in **Figure 9** in the following model:

$$H_2 = t(U_1 + U_1^{-1} + U_2 + U_2^{-1} + \tau(U_1^2 + U_1^{-2} + U_2^2 + U_2^{-2})) \qquad [7]$$

As τ passes beyond the critical value 1/4, the minimum of energy at zero flux bifurcates to become a local maximum and gives four degenerate minima. If τ is close enough to 1/4, these four minima are so close that the tunneling effect between them becomes enormous. The semiclassical analysis for the tunneling effect consists in extending the classical motion into the complex values of (k_1, k_2). Then the classical (complex) action is given by

$$S = \oint_{\gamma} k_1 \, \mathrm{d}k_2$$

where γ is any path joining two such minima in the complex energy surface $E(k + \mathrm{i}k') = E_0$ (if E_0 is the minimum of the energy). Cohomology arguments

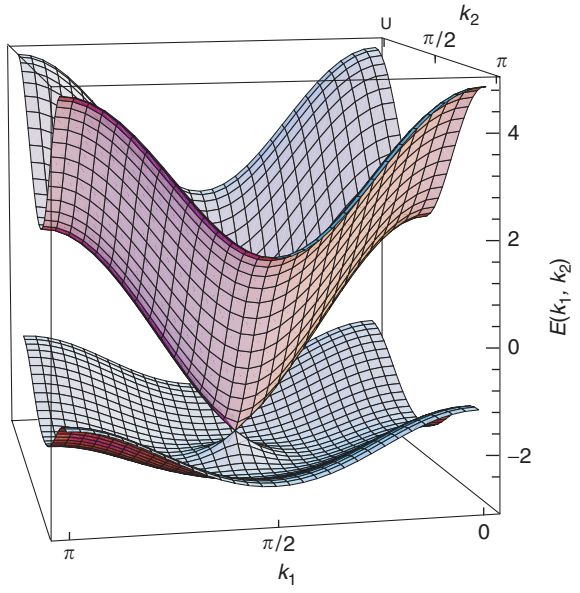


Figure 7 Bandtouching and bandoverlapping near half-flux in the model $H = U_1 + U_2 + 1/2(U_1^2 + U_2^2) + \text{h.c.}$

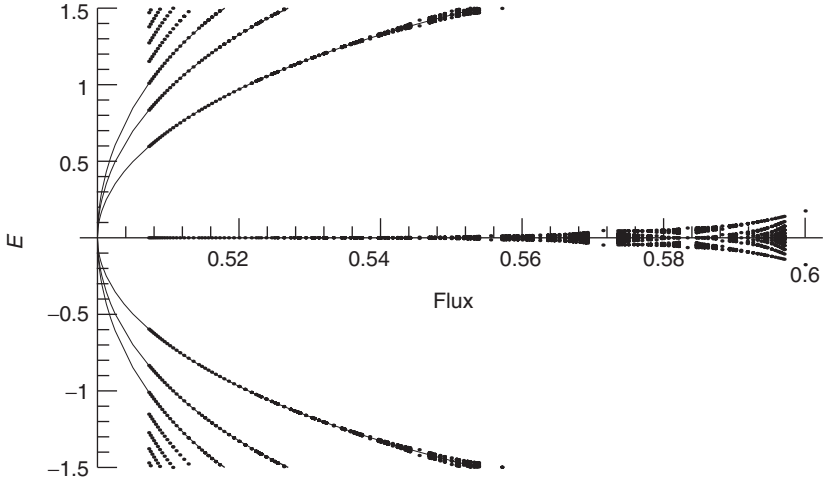


Figure 8 Comparison between semiclassical calculations (full curves) and exact numerical spectrum (points) for Dirac levels in the Harper model near half-flux.

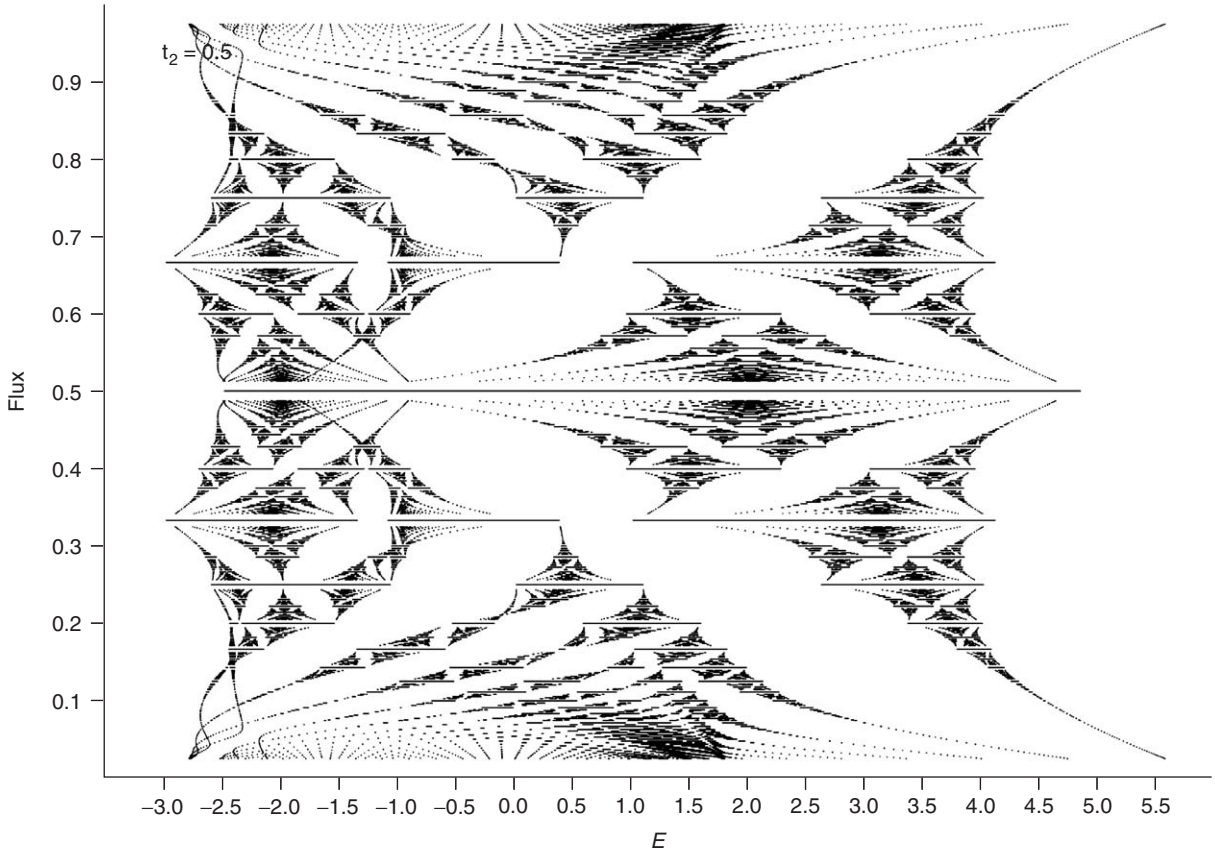


Figure 9 Spectrum of a Harper-like model [7] for $t_2 = 1/2$. In the left corners, braids are due to tunneling effect (magnetic breakdown).

(Stokes theorem) show that the value of S depends only upon the homotopy class of γ provided it stays in the energy surface. In the previous case, $S = S' + iS''$ with $S' \neq 0$ and $S'' > 0$. Then the splitting between the four levels, consecutive to tunneling between the classical minima is proportional to Re (exp $i(\phi + S/\delta)$), for some phase ϕ depending upon what splitting is considered, leading to (1) a damping of the amplitude as $\delta \downarrow 0$ given by $e^{-S''/\delta}$, and (2) an oscillation of the amplitude due to the term $e^{iS'/\delta}$. This is illustrated in **Figure 10**.

Tunneling also explains the broadening of the Landau subbands. Indeed, the classical effective Hamiltonian obtained for $\delta = 0$ (or for $\theta = 0$ when $p/q = 0$) is a periodic function of the quasi-momenta k_1, k_2. Thus, any band edge occurring at k also occurs at $k + 2\pi(m_1, m_2)$ for all $m_i \in \mathbb{Z}$. Hence, each classical orbit is infinitely degenerate in the phase space. Each such orbit represents an approximate eigenstate that can, in principle, be computed semiclassically and called a quasimode. Hence, the Hilbert space corresponding to these semiclassical eigenvalues is infinite dimensional and the tunneling effect between the quasimodes produces a broadening of the Landau or Landau–Dirac levels. As usual, with the tunneling effect, this broadening is exponentially small in the effective Planck constant (here given by δ). Such a calculation is difficult and many attempts were made in the 1960s. Eventually, Helffer and Sjöstrand carried it out for the special case of the Harper model based upon Wilkinson's analysis.

The main result of this analysis is that the Landau subbands are well described by an effective Hamiltonian of the form

$$H_{\text{subband}} \overset{\delta \downarrow 0}{=} E(\delta) + (e^{iS/\delta}(\tilde{U}_1 + \tilde{U}_2)e^{-i\bar{S}/\delta} \times (\tilde{U}_1^{-1} + \tilde{U}_2^{-1})) + O(e^{-S_1/\delta})$$

where $E(\delta)$ is the semiclassical value of the eigenvalue computed by the previous method, t is a model-dependent constant (computable by semiclassical methods), S is the classical action between the two nearest-neighbor repetitions of the corresponding band edge in the complex energy surface, $S_1 > |\text{Im } S|$ and

$$\tilde{U}_1 \tilde{U}_2 = e^{i4\pi^2/\delta} \tilde{U}_2 \tilde{U}_1$$

Hence the spectrum of the subband is itself approximately the same as the Harper spectrum, but for a renormalized value of the dimensionless flux namely

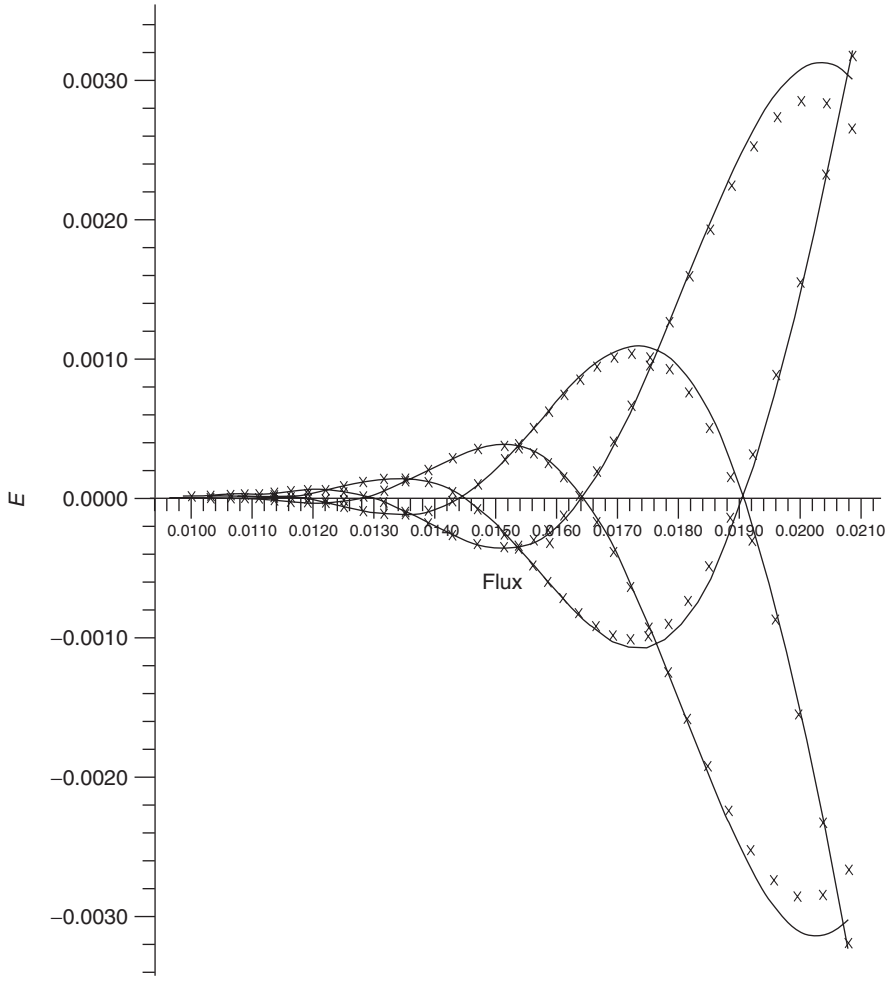


Figure 10 Comparison between semiclassical calculations (full curves) and exact numerical spectrum (crosses) for Landau braids of model [7] . (Barelli A and Kreft C (1991) Braid-structure in a Harper model as an example of phase space tunneling. *Journal of Physics I (France)* 1: 1229–1249.)

$\alpha = \delta/2\pi \mapsto 1/\alpha = 2\pi/\delta$. Since this new flux arises in a phase, adding or subtracting an integer to $1/\alpha$ does not change the phase, so that the map becomes $\alpha \in (0,1) \mapsto G(\alpha) = \{1/\alpha\} \in (0,1)$ where $\{x\}$ denotes the fractional part of x. G is called the Gauss map and is known to generate the continuous fraction expansion of the number α, namely, if $[x]$ denotes the integer part of x

$$\alpha = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_n + \cdots}}}, \quad a_n = [G^{n-1}(\alpha)]$$

This rule does not work near a band center. Then the analysis is more complicated because the classical Hamiltonian associated with the Harper equation has a saddle point. Then tunneling between periodic equivalent quasimodes is enhanced by the proximity near a saddle point. Helffer and Sjöstrand showed that depending on the subband considered, there are three types of normal forms for the effective Hamiltonian describing them. Moreover, the rule for the new flux might not be given by the Gauss map as was shown by Hofstadter. They also showed that each of them can be analyzed in the same way as the Harper equation so that at each step of the renormalization procedure, the effective subband Hamiltonian takes on one of the four normal forms previously considered: Harper or the three ones occurring at saddle points. Such an analysis explains why the Hofstadter spectrum shown in **Figure 1** (see an enlargement in **Figure 6**) looks like a fractal set, namely it is self-reproducing at any scale.

See also: Electric and Magnetic Fields in Cells and Tissues; Shubnikov–de Haas and de Haas–van Alphen Techniques.

PACS: **70; 71; 73.43.–f; 73.43.Cd; 73.43.Qt; 75.47.–m; 02; 02.30.–f; 02.30.Nw; 02.30.Tb; 02.40.Gh**

Further Reading

Barelli A and Heckinger R (1992) *Physical Review B* 46: 11559–11569.

Barelli A and Kreft C (1991) Braid-structure in a Harper model as an example of phase space tunneling. *Journal of Physics I (France)* 1: 1229–1249.

Bellissard J (1994) Noncommutative methods in semiclassical analysis. In: *Transition to Chaos in Classical and Quantum Mechanics (Montecatini Terme, 1991)*, Lecture Notes in Mathematics, vol. 1589, pp. 1–64. Berlin: Springer.

Hofstadter DR (1999) *Gödel, Escher, Bach: An Eternal Golden Brand*. Basic Books Inc. ((1979) 20th Anniversary Edition).

Fortin JY and Ziman T (1998) Frequency mixing of magnetic oscillations: beyond Falicov-Stachowiak theory. *Physical Review Letter* 80(14): 3117–3120.

Helffer B and Sjöstrand J (1988) Analyse semi-classique pour l'équation de Harper (avec application à l'ùation de Schrödinger avec champ magnétique) (French) [Semiclassical analysis of the Harper equation (with application to the Schrödinger equation with magnetic field)]. *Memoires de la Societe Mathematique de France* 34: 113.

Helffer B and Sjöstrand J (1989) Semiclassical analysis for Harper's equation. III. Cantor structure of the spectrum. *Memoires de la Societe Mathematique de France* 39: 1–24.

Helffer B and Sjöstrand J (1990) Analyse semi-classique pour l'équation de Harper. II. Comportement semi-classique pès d'un rationnel. (French) [Semiclassical analysis for Harper's equation. II. Semiclassical behavior near a rational number]. *Memoires de la Societe Mathematique de France* 40: 139.

Kohn W (1958) Theory of Bloch electrons in a magnetic field. *Proceedings of the Physical Society of London* 72: 1147–1150.

Kohn W (1959) Theory of Bloch electrons in a magnetic field: the effective Hamiltonian. *Physical Review* 115: 1460–1478.

Lifshitz IM and Kosevich AM (1956) *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* 29: 730.

Onsager L (1952) Interpretation of the de Haas-van Alphen effect. *Philosophical Magazine* 43: 1006–1008.

Osadehy D and Aoron JE (2001) Hofstadter/butterfly as quantum phase diagram. *Journal of Mathematical Physics* 42: 5665–5671.

Peierls RE (1956) *Quantum Theory of Solids*. Oxford: Oxford University Press.

Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.

Wilkinson M (1984) Critical properties of electron eigenstates in incommensurate systems. *Proceedings of the Royal Society of London A Series* 391: 305–350.

Electrophysiology

S E Pagnotta, University of Rome "Roma Tre," Rome, Italy

Introduction

Physiology can be defined as the study of the functions of living organisms and their parts, together with the physical and chemical factors and processes involved. The branch of physiology dealing with the electric phenomena associated with living bodies and involved in their functional activity is called electrophysiology.

The major target of electrophysiological studies is the central nervous system, a structure composed of a particular cell type, the neuron. Indeed, neurons are cells specialized for the integration and propagation of electrical events, and through such electrical activity they communicate with each other as well as with muscles and other end organs. Even though some concepts and techniques that are discussed in this article are general and can be applied virtually to all cell types, some are specifically referred to neurons. In the first part of this article, basic theoretical concepts are introduced, and the ionic basis of cell membrane and action potential is discussed using an equivalent circuit treatment of the membrane. In the second part, the major experimental techniques applied to electrophysiological studies are described.

Basic Theories

Cell Resting Potential

All living cells are surrounded by a plasma membrane, which separates the internal cellular compartment from the environment. Across this plasma membrane, there is an unequal distribution of ions, both organic and inorganic. Potassium ions are predominant in the intracellular fluid, together with organic anions, such as amino acids, proteins, or nucleic acids. Conversely, in the extracellular fluid the predominant ionic species are sodium and chloride. Normally, electrical phenomena arise whenever charges of opposite sign are separated or can move independently. Indeed, one of the main electrical features of all cellular types is the existence, at resting conditions, of an electric potential difference between the inner and the outer sides of the cell membrane. This potential difference, called membrane resting potential, depends on the ionic concentration gradient across the membrane and on the relative permeability of the membrane to the ions present in the extracellular and intracellular medium. Moreover, as discussed later, transient changes in the membrane potential are used by nerve cells and muscle fibers to generate and propagate electrical signals. Such potential changes are in turn due primarily to changes in the permeability of the cell membrane to the different ionic species.

To better understand the nature of the established ionic gradients in the resting cell and how such gradients come to exist, some simple observations are described. Consider a water system (like the one depicted in **Figure 1a**) with two compartments separated by a membrane permeable only to small inorganic ions but not to large organic ones. Different concentrations of an organic potassium salt $[K^+A^-]$ are placed in each compartment, and the potential difference between them is measured with a voltmeter. Initially (at time $t = 0$), no voltage difference will be observed, because both sides of the system are electrically neutral. However, potassium (K^+) ions will immediately start diffusing from compartment 1 to compartment 2, down their concentration gradient following the Fick's law: $J_{diff} = -D\partial c/\partial x$ (J_{diff} is the ion flux, D is the diffusion coefficient, c the concentration of K^+ ions, and $\partial c/\partial x$ is the concentration gradient along the x axis). This diffusion process gives an excess of positive charge to compartment 1, building up across the membrane an electric field E directed against the movement of K^+ ions. Conversely to K^+ ions, the A^- inorganic ions cannot cross the membrane, so that the charge separation persists. Thus, at $t > 0$ there are two factors governing the movement of K^+ ions: the concentration gradient and the electric field. These factors push ions toward opposite directions, until a balance between electrical and diffusional fluxes of K^+ ions is reached, the total net flux is zero and the system no longer changes (**Figure 1b**). At this moment, the concentration of K^+ ions in both compartments changes, while the concentration of A^- ions does not change. There is no more electrical neutrality in each compartment and a potential difference ΔV_{1-2} between them is established. ΔV_{1-2} can be easily calculated with the well-known Nerst equation:

$$\Delta V_{1-2} = (RT/zF)\ \ln([K^+]_2/[K^+]_1) \qquad [1]$$

where R is the gas constant, T is the temperature, z is the charge of K^+ ion, F is the Faraday constant, and $[K^+]_1$ and $[K^+]_2$ are the K concentration in the compartments 1 and 2, respectively.

A real cell membrane is permeable to several inorganic ions (K^+, Na^+, Cl^-) present in the extracellular and intracellular medium in different concentrations. Assuming that: (1) the potential difference drops only across the membrane and occurs with a uniform gradient, (2) the membrane is a homogeneous phase, and (3) the fluxes of different ions are independent, it is possible to derive the so-called Goldman–Hodgkin–Katz (GHK) voltage equation, which gives the membrane potential (called the reversal potential V_{rev}) at which no net current flows:

$$V_{rev} = (RT/F)\ \ln[(P_K[K]_o + P_{Na}[Na]_o + P_{Cl}[Cl]_i)/(P_K[K]_i + P_{Na}[Na]_i + P_{Cl}[Cl]_o)] \qquad [2]$$

where indices i and o refer to inside and outside of the cell membrane, respectively, and P_K, P_{Na}, and P_{Cl} are the membrane permeability coefficients of the corresponding ions. When the cell is in resting condition, the GHK equation yields the membrane resting potential, which is now clearly established by ion concentration gradients and permeability. If the membrane is permeable only to one ionic species, the GHK equation reduces to the Nernst equation, which gives the reversal potential for that ion.

Measuring the membrane potential of real cells, one can observe that the intracellular compartment is always at a potential value more negative than the

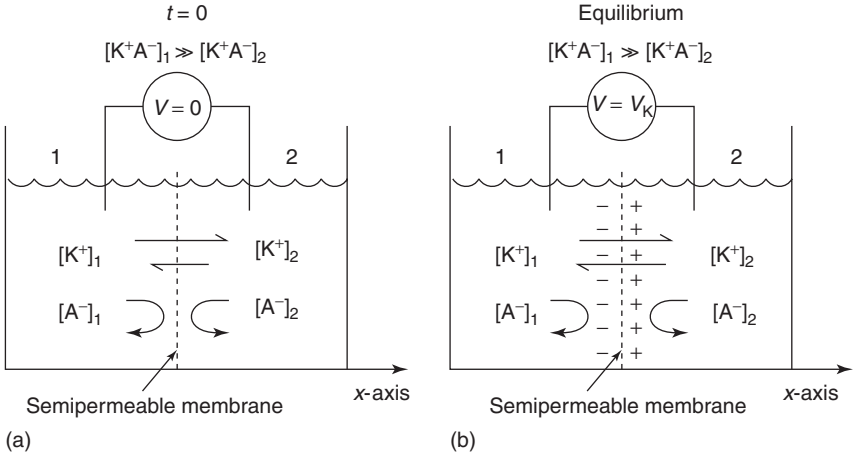


Figure 1 Ion diffusion across a semipermeable membrane.

Table 1 Ionic concentrations and reversal potentials for mammalian skeletal muscles and squid giant axon

Ion	*Extracellular concentraion* (mM)		*Intracellular concentration* (mM)		*Reversal potential* (mV)	
	Skeleton	*Squid*	*Skeleton*	*Squid*	*Skeleton*	*Squid*
Na^+	145	440	12	50	+67	+55
K^+	4	20	155	400	−98	−75
Cl^-	123	560	4.2	52	−90	−60

extracellular one. As just discussed, this is a consequence of the different concentration of ions in the two compartments. Typical values for such concentrations and for the corresponding reversal potentials are reported in **Table 1** for two kinds of cells. It is experimentally observed that membrane resting potential and potassium reversal potential values are quite similar if measured in the same cell, indicating that in resting conditions potassium gives the main contribution to membrane permeability. This is a general phenomenon observed in many cell types.

How do ions permeate the cell membrane? Particular membrane proteins, called ion channels, connect the intracellular and extracellular compartments of the cell forming a macromolecular pore in the membrane, which can be crossed by ions. A wide number of different ion channels, accounting for a variety of cellular functions, are expressed in all the cellular types. They can be selective for specific ions (sodium, potassium, etc.), and most of them are gated. This means that they fluctuate between closed and open states with characteristic open times. For a few channels, the open state predominates at resting conditions (e.g., some potassium channels responsible for the membrane resting potential), while the remainder are predominantly in the closed state, and the probability of an individual channel opening is low. An appropriate stimulus can activate or inactivate channels, sharply increasing or decreasing the frequency of openings. Some channels may respond to changes in the membrane potential, and thus are called voltage-gated ion channels, while others may respond to applied chemicals, and thus are called ligand-gated ion channels.

Electrical Properties of Cell Membranes

Depending on their electrical properties, biological membranes or cells can be subdivided into two main categories: excitable or nonexcitable. Both classes are characterized by passive electrical properties, while only excitable membranes or cells show active electrical properties and are, therefore, able to generate and propagate electrical signals.

Passive properties As far as it concerns passive properties, common to all cell types, the cell membrane can

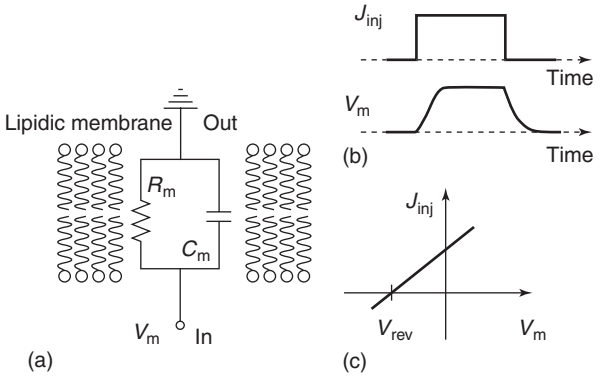


Figure 2 (a) Equivalent circuit of membrane passive properties. (b) Membrane voltage recorded in response to a current pulse injection. (c) Current–voltage relationship for a nonexcitable membrane.

be schematized with a simple RC circuit. C_m and R_m, the membrane capacitance and resistance, are independent of the membrane potential and represent the membrane dielectric or insulating properties and the ion permeability through the membrane, respectively (**Figure 2a**). If a pulse of electric current is passed through the cell membrane, its potential will rise to a maximal and sustained value and then will decay, following in both cases an exponential law:

$$V_m = V_0[1 - e^{-t/R_m C_m}] = V_0[1 - e^{-t/\tau_m}] \quad [3a]$$

for the rising phase and

$$V_m = V_0 e^{-t/R_m C_m} = V_0 e^{-t/\tau_m} \quad [3b]$$

for the decaying one (**Figure 2b**). τ_m is the membrane time constant. For the sake of the following discussion, a change in the membrane potential toward more positive values is referred to as a depolarization, while a change toward more negative values will be referred to as a hyperpolarization. The R_m values being voltage and time independent, the current–voltage relationship for the maximal sustained potential during current pulse application will be linear (**Figure 2c**), and if the membrane potential is kept at a particular voltage level V_m, the total current flow will be

$$J_m = J_c + J_i = C_m(dV_m/dt) + G_m(V_m - V_{rev}) \quad [4]$$

where $G_m = 1/R_m$ is the membrane conductance and V_{rev} is the previously defined reversal potential.

Actually, for excitable membranes C_m and R_m values are voltage-dependent (as voltage-gated ion channels are present on the membrane), and the previous schematic description holds only at resting condition or in a range of membrane potential below a particular threshold. Indeed, in these conditions the values of C_m and R_m become virtually constant, while in any other condition active properties arise.

Active properties: action potential and the Hodgkin–Huxley model Active electrical properties account for the generation of electrical signals and their propagation in excitable membranes. To this end, excitable membranes contain voltage-dependent ionic channels. Mainly, three different kinds of channels – selective for sodium, potassium, or chloride, respectively – are necessary. The electric equivalent circuit for the membrane now becomes a little bit more complicated, as depicted in **Figure 3a**. C_m is again the membrane capacitance, g_{Na}, g_K, and g_{Cl} are the voltage-dependent conductances, and V_{Na}, V_K, and V_{Cl} are the Nernst reversal potentials for sodium, potassium, and chloride, respectively. Holding the membrane potential at a value V_m, the current flowing is

$$\begin{aligned} J_m &= J_c + J_{Na} + J_K + J_{Cl} \\ &= C_m(dV_m/dt) + g_{Na}(V_m - V_{Na}) \\ &\quad + g_K(V_m - V_K) + g_{Cl}(V_m - V_{Cl}) \end{aligned} \quad [5]$$

and, in resting condition when $J_m = 0$ and $dV_m/dt = 0$, the membrane potential becomes

$$\begin{aligned} V_m &= V_{resting} \\ &= (g_{Na}V_{Na} + g_K V_K + g_{Cl}V_{Cl})/(g_{Na} + g_K + g_{Cl}) \end{aligned} \quad [6]$$

The current–voltage relationship of excitable membranes is obviously nonlinear, due to the voltage and time dependence of the ionic conductances present and, in particular, one can observe different nonlinear current–voltage relationships for the isolated potassium and sodium currents (**Figure 3b**).

But how do excitable cells, and in particular neurons, code electrical signals? As previously discussed, neurons are the basic units of the central nervous system – receiving, analyzing, and perceiving information, and communicating with one another through synapses. The single neuron codes and transmits signals using the so-called action potential (AP). The AP is a sudden and transient depolarization of the cell membrane, and it occurs as an all-or-nothing event. This means that, in the same cell, a stimulus intense enough to produce an AP will always produce APs with the same amplitude. The information about stimulus intensity will be coded in the frequency of APs. A qualitative scheme of an AP is depicted in **Figure 4a**. AP has three main phases: (1) depolarization; (2) repolarization (or hyperpolarization); and (3) after-hyperpolarization. It is mainly due to the activation of two kinds of voltage-dependent ionic channels, permeable to sodium and potassium, respectively, showing different and nonlinear current–voltage relationships. If the cell membrane is depolarized over a threshold level, sodium channels (inactive at resting conditions) are suddenly

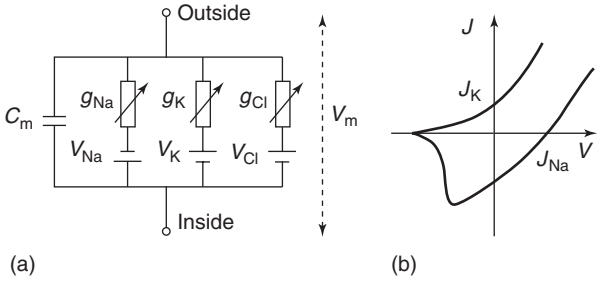


Figure 3 (a) Equivalent circuit of membrane active properties. (b) Current–voltage relationship for potassium and sodium channels.

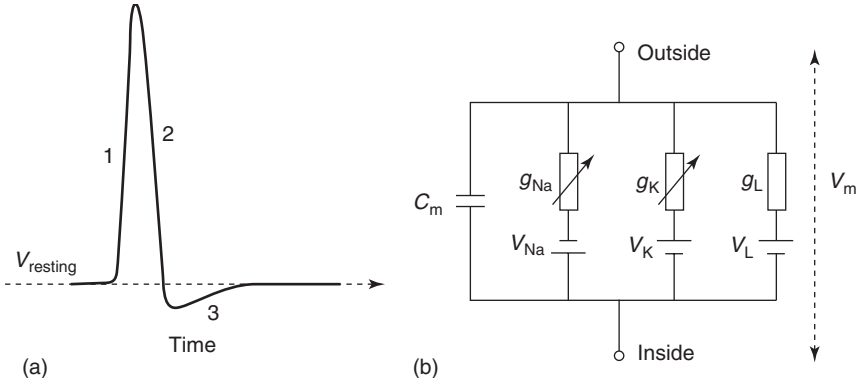


Figure 4 (a) Schematic representation of an action potential. (b) Equivalent circuit for cell membrane in the Hodgkin–Huxley model.

activated. Sodium enters into the cell and the membrane is further depolarized. When the membrane potential reaches the threshold for potassium channels activation, potassium starts to exit from the cell, repolarizing the membrane, even to values more negative that the resting one.

A fundamental step in understanding AP nature was done by A L Hodgkin and A F Huxley in the mid-twentieth century, proposing a semi-empirical model describing the AP generated in the squid giant axon. According to their model g_K and g_{Na}, potassium and sodium conductances, can be expressed as experimentally determined maximum conductances $\hat{g}_K$ and $\hat{g}_{Na}$ multiplied by coefficients, representing the fraction of the maximum conductances actually expressed. The kinetics of the conductances is considered to be driven by independent membrane-bound particles, each one with a certain probability to be in the right position to turn on the corresponding conductance. Thus,

$$\begin{aligned} g_K(V,t) &= n(V,t)^4 \hat{g}_K \\ g_{Na}(V,t) &= m(V,t)^3 h(V,t) \hat{g}_{Na} \end{aligned} \qquad [7]$$

where n, m, and h are the probabilities for the hypothetical particles to be in the permissive state. Consider the axon membrane patch represented by the equivalent circuit in **Figure 4b**, where g_L and V_L are the voltage-independent conductance and reversal potential of a leakage current J_L. Accordingly, the current flowing through this patch of the membrane will be

$$\begin{aligned} J_m = C_m(dV_m/dt) &+ n(V,t)^4 \hat{g}_K (V_m - V_K) \\ &+ m(V,t)^3 h(V,t) \hat{g}_{Na} (V_m - V_{Na}) \\ &+ g_L (V_m - V_L) \end{aligned} \qquad [8]$$

All the kinetic properties of J_m depend on the time and voltage dependences of n, m, and h. In the Hodgkin and Huxley model, these parameters can be calculated as

$$\begin{aligned} n(V,t) &= n_\infty - (n_\infty - n_0)\, e^{-t/\tau_n} \\ m(V,t) &= m_\infty - (m_\infty - m_0)\, e^{-t/\tau_m} \\ h(V,t) &= h_\infty - (h_\infty - h_0)\, e^{-t/\tau_h} \end{aligned} \qquad [9]$$

where n_∞, m_∞, h_∞, τ_n, τ_m, and τ_h are experimentally determined parameters related to the voltage-dependent rate of activation or inactivation of sodium and potassium conductances.

Equations [8] and [9] together are referred to as Hodgkin and Huxley equations. By solving them, it is possible to obtain remarkable fits between the recorded and calculated APs. Even though this model is to some extent oversimplified, it is yet sufficiently good to predict correctly the major feature of excitability, such as the AP shape and conduction velocity.

Electrical models of cells The equivalent electric circuit approach can be extended to the description of the electrical properties of the whole cell obviously with some simplifications. The simplest case is that of a spherical cell with a uniform membrane potential (isopotential cell). The current flowing across the membrane will be again due to capacitive and conductive terms: $J_m = J_i + J_c$. While considering particular cells like neurons, it should be borne in mind that these cells have extensive dendritic trees and axons, and that electrical signals, like APs, travel along them. When a neuron is active, or receives signals from other neurons, it is nonuniformly polarized and it cannot be considered isopotential. In this situation, the linear cable theory is applied. Axons or dendrites can be considered as cylindrical compartments schematized by the electrical equivalent circuit in **Figure 5**. For simplicity r_0, the extracellular space resistance, can be set equal to zero, because of the isopotentiality of such a compartment (this condition is not true only in some special cases). r_i, r_m, and c_m (the intracellular resistance, the membrane resistance, and the membrane capacitance,

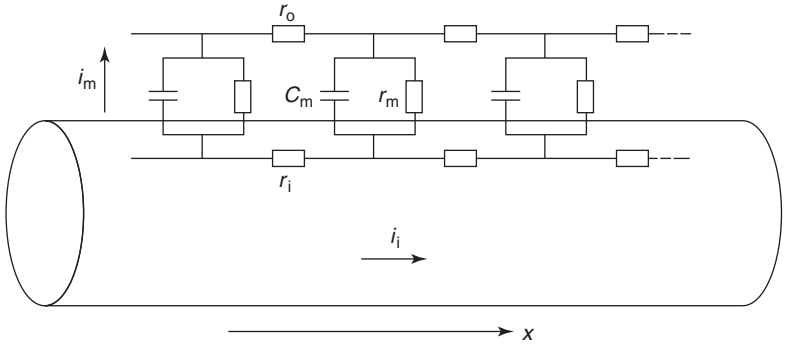


Figure 5 Equivalent circuit of an axon using the cable theory.

respectively) are considered to be uniform and constant along the entire cylinder, and to be independent of V_m. With some simple computations, it is possible to find the cable equation describing the membrane potential V_m produced in response to current injection at some point along the cylinder. This potential will be a function of time and of the distance from the point of injection:

$$\lambda^2 \partial^2 V_m(t,x)/\partial x^2 = \tau_m \partial V_m(t,x)/\partial t + V_m \qquad [10]$$

where $\lambda = \sqrt{(r_m/r_i)}$ is the space (length) constant and $\tau_m = r_m c_m$ is the membrane time constant. The solutions of the cable equation were derived for specific situations: infinite cable, finite cable, semifinite cable, and finite cable with soma. The potential V_x at a distance x for the infinite cable is given by $V_x = V_0 e^{-x/\lambda}$, with V_0 equal to the potential at the point of current injection. In this case, the conduction velocity, θ, of the potential is given by $\theta = 2\lambda/\tau_m$ and thus depends on cell membrane characteristics and on cell cytoplasm resistance.

Experimental Techniques

Several experimental techniques have been developed to understand the electrical functioning of cells, and in particular of neurons of the central nervous system. Depending on the level of analysis, different methods can be used. One can move from a subcellular approach, in which patch clamping and single-channel recording are the main techniques, to a behavioral approach where neuronal recordings in awake primates can be performed. Noninvasive electrophysiological techniques, such as electroencephalography (EEG) and event-related potential recordings, are also possible, together with methods of metabolic imaging to measure neuronal activity (e.g., the PET).

Channel Function

As discussed in previous paragraphs, electrical activity of biological systems at the cellular level consists of ion movement through highly specialized channels spanning the cell membrane. "Native" channels, normally expressed in a particular cell, can be studied. However, a deep understanding of the ion channels functioning is now greatly facilitated by the use of particular experimental systems, such as artificial and natural membranes in which specific ion channels can be selectively expressed. For natural membranes, ionic channels are newly synthesized by foreign DNA or RNA injected into a large living cell (such as the *Xenopus laevis* oocyte) via large pipettes.

Apart from the choice of the system expressing the ion channels under investigation, different techniques can be used to record their electrical activity. A fundamental method for electrophysiological studies, namely the single-cell patch clamp, was developed around 1978 by Sakmann and Neher. The starting point of this technique is the formation of a high-resistance (gigaohm) seal when the tip of a fire-polished glass micropipette is pressed against a cultured or acutely isolated cell membrane. The gigaohm seal essentially allows a high current gain, and low noise amplification, necessary for recording small, brief currents passing through single ionic channels. After this first step, the patch-clamp method can be applied in at least four configurations: cell-attached, inside-out, outside-out, and whole-cell (**Figure 6**). The first three configurations allow study of individual ion channels under different conditions.

The cell-attached configuration (**Figure 6a**) is essentially obtained with the formation of the gigaohm seal. In this mode, the electrical activity of single channels, present under the pipette tip, is recorded without disruption of the cell membrane. In the inside-out and outside-out configurations, the membrane patch is detached from the cell after the gigaohm seal is formed, and single-channel activity is recorded in isolation from the cell. In particular, in the inside-out configuration (**Figure 6c**), the patch of

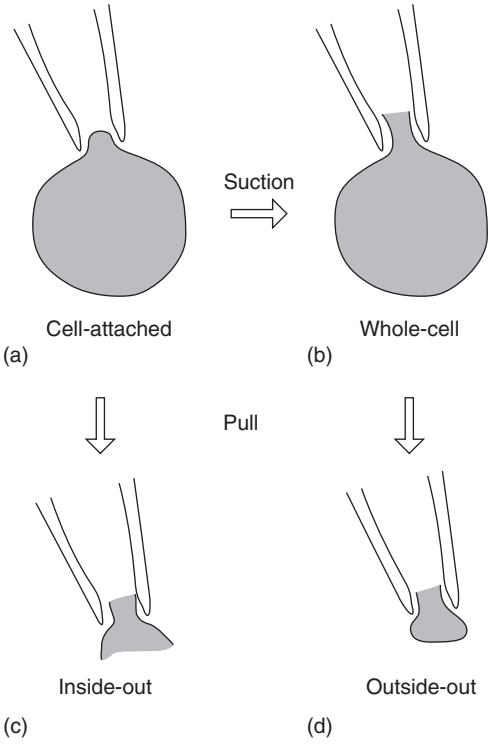


Figure 6 The four configuration of the patch-clamp technique.

membrane is gently pulled away from the cell, and the patch remains attached to the pipette with its cytoplasmic surface now exposed to the bathing solution. On the other hand, to reach the outside-out configuration (**Figure 6d**), the patch of membrane is gently pulled away from the cell only after an intermediate step (the whole-cell configuration that will be described below), in which the membrane patch under the pipette is ruptured by applying a strong suction to the recording pipette interior. Once detached, the membrane seals over the pipette tip, in favorable cases forming a membrane patch in which the extracellular membrane surface is exposed to the bathing solution.

The last configuration is the whole-cell clamp (**Figure 6b**). It is a form of the cell-attached configuration that uses the same pipette type and gigaohm seal method described above but that, by rupturing the membrane under the tip, allows recording of the "macroscopic" or summed currents flowing through all channels in the entire cellular membrane, rather than through a single channel, and for this reason is mainly used to study cellular electrical function (as it will be described in the next paragraph). In this configuration, the diffusible contents of the pipette exchange with those of the cell over time.

Cellular and Neural Circuit Function: *In Vitro* Studies

A neuron normally receives contacts from thousands of other neurons, and many of those inputs may be synchronized. To understand the way in which the cell integrates these numerous inputs and alters its own electrical activity, or how an action at one locus can lead to transmission of information to another locus, without interference from outside events, preparation of intact neurons in isolation or semi-isolation have been developed. An "acutely isolated neuron" model, prepared by mild enzymatic digestion and gentle agitation and trituration of brain slices, and chronically cultured neurons and cell lines (such as those also used for previously described single-channel studies) has been a major cellular source for electrophysiological studies, as have single neurons within brain slices. This last system can also be used in electrophysiological studies of synaptic transmission and local brain circuits. Another more recently developed model is the slice culture (or organotipic culture). In this case, cultures are prepared by long-term incubation of brain slices in a rotating drum, where the slices develop into a monolayer of neurons retaininig their local cytoarchitectonic organization.

To study the electrical activity of cellular systems, extracellular recording is probably the most straightforward electrophysiological recording situation. With this technique, the tip of a microelectrode is positioned immediately adjacent to, but outside of, a single neuron. In close proximity to the cell, current fields generated by APs in that cells are detected by the microelectrode as small voltage deflections. A multiple-cell recording configuration is also possible, with several microelectrodes inserted into one region of the system under study. By recording simultaneously from many neurons, their concerted activity can be investigated.

With intracellular recordings, membrane current or potential is measured via an electrode inserted into, or in contact with, the intracellular compartments. Two configurations are possible: the current clamp and the voltage clamp. In the current-clamp mode, the voltage difference across the cellular membrane is measured while injecting constant positive or negative current into the cell. Without current injection, the membrane resting potential can be measured. On the other hand, by injecting repetitive constant current pulses into the cell, a variety of information can be obtained, such as a relative measure of the resistance of the membrane, the current–voltage curves, and the voltage responses to activation of inhibitory or excitatory afferents. Conversely, during voltage-clamp recordings, the current required to hold the cell at a constant voltage is measured. A major advantage of this method over current-clamp recordings is that it is possible to measure ionic currents directly. In practice, voltage and current-clamp recordings involve the same sorts of micropipettes. In fact, one often penetrates a cell with the micropipette under the current-clamp mode. Then after the cell stabilizes, a series of adjustments of the recording characteristics can allow stable switching to the voltage clamp.

Neural Circuit Function and Behavioral Link

***In vivo* single-cell studies** The intact, functioning brain can be readily explored with microelectrodes in anesthetized animals. The microelectrode tip can be positioned in a desired position, so that any site within the brain can be found and cellular activity recorded. X-ray or magnetic resonance imaging methods may also be used for this purpose in human studies. In these experiments, the impulse activity of neurons is typically recorded extracellularly. There are many experimental applications of *in vivo* single-cell electrophysiology, such as iontophoresis (and other methods for local drug application), stimulation recordings (for discerning the functional effect of an afferent input to a neuron), and antidromic activation (used to confirm the projection to an area and to

determine the time required for conduction of an impulse along a projection pathway).

Behavioral electrophysiology A function at the behavioral level is a product of neuronal networks rather than simply the product of properties of individual neurons. To this end, neurons are typically associated in functionally related groups and circuits that can be specialized, for example, for sensory and motor functions or for associative activities. Many questions concerning cognition require experiments in the intact animal, because molecular and cellular experiments are unable to completely explain cognitive functions such as attention, perception, emotion, or memory. Hence, most studies in the electrophysiology of cognitive processes involve recording single neurons in behaving animals. Such studies often employ extracellular recordings from a metal microelectrode held in a miniature micropositioner on the animal's head. As with all extracellular recordings, it is important to know whether activity seen on an individual electrode is generated from one neuron only, or from several nearby neurons simultaneously recorded. Results of the latter, termed multiple-cell recording or multi-unit recording, are more difficult to interpret because neurons in the multiple-cell population may be physiologically heterogeneous. In addition, it is more difficult to ensure the stability of the recorded signal over time with multiple-cell activity.

See also: Electric and Magnetic Fields in Cells and Tissues; Ion Transport in Biological Membranes; Neuroscience.

PACS: 87.16.Dg; 87.16.Uv; 87.17.Nn; 87.19.La; 87.19.Nn

Further Reading

Hille B (1992) *Ionic Channels of Excitable Membranes*. Sunderland: Sinauer Associated.

Johnston D and Miao-Sin Wu S (1995) *Foundations of Cellular Neurophysiology*. MIT Press.

Kandell ER (2000) *Membrane potential. Local signaling. Propagated signaling*. In: Kandell ER, Schwartz JH, and Jessel TM (eds.) *Principle of Neuroscience*, 4th edn., pp. 125–170. McGraw-Hill.

Kettenmann H and Grantyn R (eds.) (1992) *Practical Electrophysiological Methods. A Guide for In Vitro Studies in Vertebrate Neurophysiology*. New York: Wiley-Liss.

Nunez PL (1981) *Electric Field of the Brain: The Neurophysics of EEG*. Oxford University Press.

Sakmann B and Neher E (eds.) (1995) *Single-Channel Recording*, 2nd edn. Plenum Publishing Corporation.

Elemental Semiconductors, Electronic States of

G C La Rocca, Scuola Normale Superiore, Pisa, Italy

Introduction

Silicon, the paradigmatic elemental semiconductor, is by far the most relevant material in the electronic industry and possibly the best-known solid both from the point of view of basic science and technology. The present understanding of virtually all the properties of a crystalline solid such as Si is based on an accurate knowledge of its electronic states, that is, its electronic band structure. Quantum mechanics predicts the energies and wave functions available to the electrons in the periodic crystal potential. In contrast to the case of atoms or molecules, the translational symmetry of the crystal leads to extended states characterized by a Bloch wave vector $\boldsymbol{k}$. Just as in atoms and molecules, the Pauli exclusion principle dictates how distinct electronic states are occupied, each by a single electron starting from the lowest energy levels. From such information, mechanical, transport, and optical properties can be inferred.

Semiconductors are characterized by an energy gap (E_g) between the highest occupied states (top of the valence band) and the lowest unoccupied ones (bottom of the conduction band), which is comparable to 1 eV. Silicon ($E_g = 1.17$ eV) and germanium ($E_g = 0.774$ eV) are the most important elemental semiconductors, while diamond with an energy gap exceeding 5 eV is an insulator and grey tin a "zero-gap semiconductor." All these elements, belonging to group IV of the periodic table, are tetrahedrally coordinated and crystallize in the diamond structure with a cubic symmetry. The focus here is on Si and Ge, and to a lesser extent on α-Sn, while other elemental semiconductors, such as boron and selenium, having a more complex crystal structure and being less representative, are not considered.

The main features of electronic bands are determined by the symmetry of the crystal more than by the detailed shape of the crystal potential, and there are strong similarities in the band structures of C, Si, Ge, and α-Sn. In the following, the symmetry properties of

the diamond crystal and the main ingredients to calculate the energy bands are first described, and then the electronic band structures of Si, Ge, and α-Sn and their direct comparison with optical and photoemission experiments.

Crystal Structure and Symmetry

The group IV elemental semiconductors, such as carbon, are characterized by four electrons in the outer shell, with an electronic configuration s^2p^2 which, in the solid, gives rise to four tetrahedrally oriented sp^3 hybrid orbitals. The resulting network of covalently bonded atoms corresponds to a face-centered cubic (f.c.c) crystal in which each atom occupies an equivalent position. Such diamond crystals belong to the $Fd3m$ (O_h^7) crystallographic space group having the point group full cubic symmetry O_h.

Its translational symmetry is described by a Bravais f.c.c. lattice of primitive translations: $\boldsymbol{\tau}_1 = (a/2)\ (0,1,1)$, $\boldsymbol{\tau}_2 = (a/2)(1,0,1)$ and $\boldsymbol{\tau}_3 = (a/2)(1,1,0)$, where a is the side of the conventional unit cell, that is, the crystal is invariant for translations by any vector $\boldsymbol{R} = n_1\boldsymbol{\tau}_1 + n_2\boldsymbol{\tau}_2 + n_3\boldsymbol{\tau}_3$, with integral values of n_1, n_2, and n_3. The lattice constants of cubic elemental semiconductors are $a = 5.43$ Å for Si, $a = 5.65$ Å for Ge, $a = 6.46$ Å for α-Sn. The corresponding reciprocal lattice is body-centered cubic (b.c.c.) with primitive vectors: $\boldsymbol{K}_1 = (2\pi/a)(\bar{1},1,1)$, $\boldsymbol{K}_2 = (2\pi/a)(1,\bar{1},1)$ and $\boldsymbol{K}_3 = (2\pi/a)(1,1,\bar{1})$; the reciprocal lattice vectors, $\boldsymbol{G}$ ($\boldsymbol{G} = m_1\boldsymbol{K}_1 + m_2\boldsymbol{K}_2 + m_3\boldsymbol{K}_3$, with m_1, m_2, and m_3 integers), have the property: $e^{i\boldsymbol{R}\cdot\boldsymbol{G}} = 1$, for any lattice translation $\boldsymbol{R}$. The primitive (symmetric) cell of the reciprocal lattice, that is, the first Brillouin zone (BZ), of the diamond structure is shown in **Figure 1**. The physical significance of the reciprocal lattice and of the BZ is that, as a consequence of the periodic translational symmetry, the extended electronic states in a crystal are described by wave functions of the form $\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u(\boldsymbol{r})$ (Bloch's theorem), where the wave vector $\boldsymbol{k}$ belongs to the BZ and the function u is invariant under lattice translations ($u(\boldsymbol{r}+\boldsymbol{R}) = u(\boldsymbol{r})$).

Beside the f.c.c. lattice translations, the rotations and reflections belonging to the full cubic group (O_h) leave the diamond structure invariant. At the microscopic level, though, half of the 48 symmetry operations of O_h are accompanied by a fractional translation such that the two interpenetrating f.c.c. lattices are exchanged (the space group of diamond being a nonsymmorphic space group with a basis of two atoms). This is the main difference with respect to the zinc-blende symmetry of compound semiconductors such as GaAs in which the two sublattices are not equivalent. In particular, an elemental semiconductor such as Si has an inversion symmetry with respect to the middle point of its homonuclear covalent bond, whereas in a compound semiconductor such as GaAs, this symmetry is lacking. The point-group symmetry determines the degeneracies of the electronic levels as a function of their wave vector $\boldsymbol{k}$ in the BZ. In particular, the electronic wave functions at points and lines of high symmetry in the BZ, those marked in **Figure 1**, are characterized by their transformation properties under point-group symmetry operations that leave their wave vector unchanged (i.e., they belong to the irreducible representations of the little group of their wave vector).

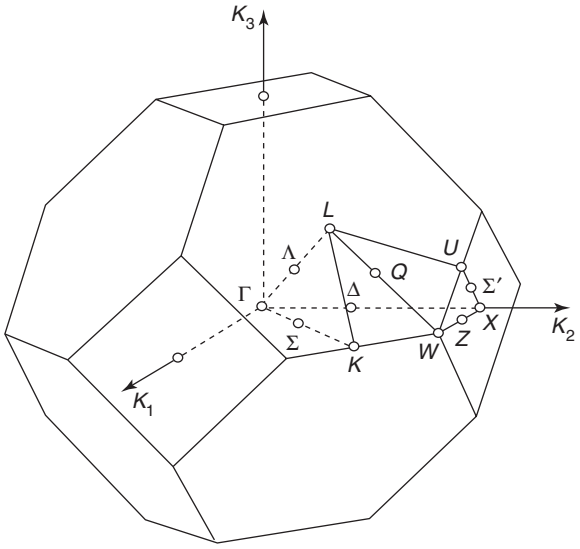


Figure 1 BZ of the diamond crystal; lines and points of high symmetry are indicated.

Besides space-group symmetry, time-reversal symmetry also affects the properties of a crystal, and it may bring about additional degeneracies of the electronic states. In particular, when the spin of the electrons is taken into account, in a crystal with inversion symmetry (such as Si) at each wave vector $\boldsymbol{k}$ the electronic states are (at least) doubly degenerate, because the state $|\boldsymbol{k}, \uparrow\rangle$ with spin up is degenerate with the state $|-\boldsymbol{k}, \uparrow\rangle$ by inversion symmetry, and the latter is degenerate with the state $|\boldsymbol{k}, \downarrow\rangle$ by time-reversal symmetry. The spin degeneracy between the states $|\boldsymbol{k}, \uparrow\rangle$ and $|\boldsymbol{k}, \downarrow\rangle$ is, in general, absent in materials lacking an inversion center (such as in GaAs). Standard group theoretical methods are employed to systematically obtain all the wealth of information that symmetry alone provides.

Energy Band Calculations

Even though translational symmetry enormously simplifies the computation of the electronic states

in a crystal, reducing the size of the problem to that of a unit cell, still a quantitative band structure calculation generally adopts a number of approximations. The electrons are considered to move in the presence of nuclei fixed at the equilibrium positions of the lattice (adiabatic approximation). The electron–electron interactions are treated at the mean field level neglecting many body correlations (one-particle approximation). Finally, the interactions with the nuclei, also including typically the core of inner electronic shells, and with all other electrons in average are lumped into a local potential $V(\boldsymbol{r})$ which has the full symmetry of the crystal (band approximation). The crystal potential $V(\boldsymbol{r})$ can be either chosen empirically or determined within a self-consistent *ab initio* scheme. The solution of the corresponding Schrödinger equation for the electronic states

$$\left[\frac{p^2}{2m} + V(\boldsymbol{r})\right]\Psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = E_n(\boldsymbol{k})\Psi_{n,\boldsymbol{k}}(\boldsymbol{r})$$

gives the (possibly degenerate) Bloch eigenfunctions Ψ, with a wave vector $\boldsymbol{k}$ in the BZ and a band index n, which correspond to the electronic energy levels $E_n(\boldsymbol{k})$, that is, the energy bands which depend continuously on $\boldsymbol{k}$. Relativistic corrections can be included in the above equation and may be important for heavy elements as discussed later on.

There are numerous and diverse methods to calculate the energy bands. In particular, some consider a crystal as a collection of atoms (localized orbital methods), while others treat it as a perturbed uniform fluid (plane-wave methods). A very popular localized orbital approach is the tight-binding method in which the wave functions are expanded in a basis of linear combination of atomic orbitals (LCAO); usually, only a limited number of atomic orbitals are included and only the interactions among a limited number of neighbors are considered. This method, at least in its simpler semiempirical versions, is transparent and economical; it has been very successful in calculating both electronic and structural properties of semiconductors and establishing their chemical trends. The minimal tight-binding model to rationalize the electronic band structure of elemental semiconductors, at least concerning their valence band states, is sketched in **Figure 2**. For each one of the two atoms in the elementary unit cell, a basis of atomic states is chosen including only one *s* and one *p* level respectively (two and six times degenerate including spin). If one imagines bringing two atoms close together as in a diatomic molecule, each level would give rise to one bonding and one antibonding molecular orbital, the former having a lower energy.

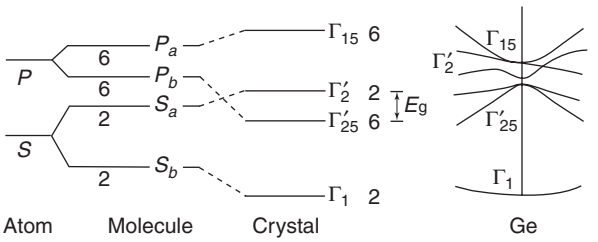


Figure 2 Scheme showing how lower-lying bonding (b) and higher-lying antibonding (a) molecular levels give rise to Bloch states of specific symmetry at the Γ point in Ge.

Table 1 Energy levels in silicon and germanium in eV, from the top of the valence band; the double-group symmetry labels including parity are indicated

	Si	*Ge*
Γ_6^+	−11.7	−12.7
Γ_7^+	−0.04	−0.29
Γ_8^+	0.00	0.00
Γ_6^-	+3.09	+3.04
Γ_8^-	+3.13	+3.37
Γ_7^-	+4.32	+0.90
X_5	−8.43	−8.84
X_5	−3.00	−3.36
X_5	+1.30	+1.12
L_6^-	−10.4	−10.7
L_6^+	−7.01	−7.63
L_6^-	−1.20	−1.37
$L_{4,5}^-$	−1.16	−1.12
L_6^+	+2.47	+0.74

From recent empirical tight-binding calculations.
Courtesy of Dr. J-M Jancu.

In the Ge crystal, for example, such states give rise at the center of the BZ at $\boldsymbol{k} = 0$ (the Γ point) to a sequence of levels having the full cubic symmetry and corresponding, in order of increasing energies, to an *s*-bonding-like Γ_1 level, a *p*-bonding-like $\Gamma_{25}^{'}$ level, an *s*-antibonding-like $\Gamma_2^{'}$ level and a *p*-antibonding-like Γ_{15} level. The eight outer shell electrons (four for each atom in the unit cell) fill the Γ_1 and $\Gamma_{25}^{'}$ levels completely, the latter being thus the top of the valence band, while the $\Gamma_2^{'}$ and Γ_{15} levels remain empty (conduction states). State-of-the-art tight-binding calculations, of course, are based on larger basis sets, possibly including also *d* states, and can also describe well the more delocalized conduction band states (see **Table 1**).

The plane-wave methods take an opposite point of view and start considering the electronic states in the presence of a very small crystal potential: in such an "empty lattice," the zeroth-order eigenfunctions of nearly free electrons are just plane waves folded back in the BZ according to the translational symmetry as sketched in **Figure 3**. For instance, looking again at

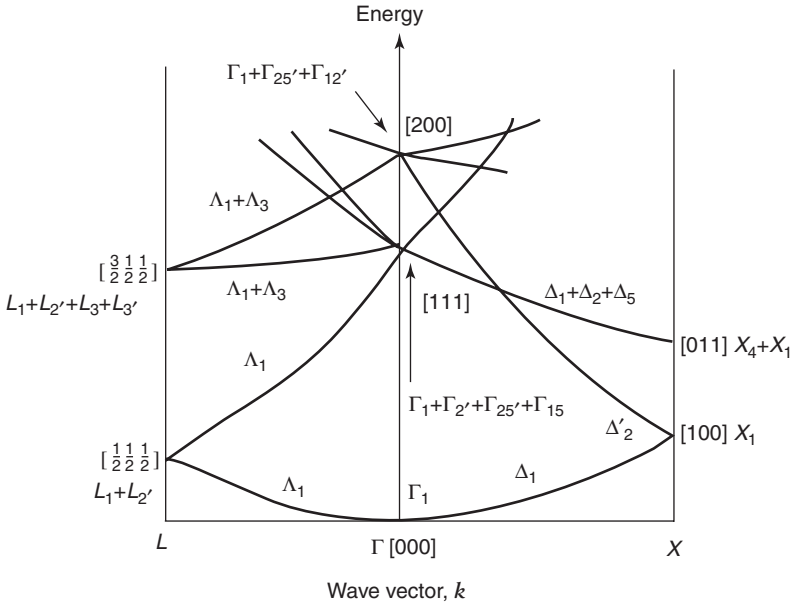


Figure 3 Empty lattice band structure for a diamond-like crystal.

the Γ point, the lowest energy level corresponds to a plane wave with a vanishing wave vector (Γ_1 state in the cubic crystal), while the next higher level includes the eight plane waves having a wave vector of the form $q=(2\pi/a)(\pm 1, \pm 1, \pm 1)$, which differ from zero by one of the shortest nonzero reciprocal lattice vectors. When the crystal potential is turned on, these eight states (sixteen including spin) split into several levels according to the cubic symmetry as indicated. However, the actual crystal potential is not small at all; the most important plane-wave method developed to circumvent this problem is the pseudopotential approach. In a crystal, the valence (or conduction) band wave functions are rapidly varying near each atomic site as they are orthogonal to those of the core states; therefore, their expansions in plane waves are very slowly convergent. It is possible, though, to substitute them with pseudo-wave-functions that are smooth (nodeless) near the atomic sites while replacing the ionic potential with a much smaller pseudopotential, in such a way that the energies $E_n(\boldsymbol{k})$ of the valence and conduction band states can be computed using a limited number of plane waves; furthermore, from the pseudo-wave-functions, the electronic charge density outside the cores can also be obtained. The pseudopotential (or rather its first few Fourier components) can be empirically chosen to reproduce a set of significant data (mostly optical transition energies and effective masses). For instance, empirical calculations using only three Fourier components of the pseudopotential can already give a reasonable description of the lowest energy bands of elemental semiconductors. A significant improvement is obtained using energy-dependent nonlocal pseudopotentials.

For state-of-the-art *ab initio* self-consistent calculations, made possible by the great increase of computational capabilities, the pseudopotential method has also been very successful in conjunction with the density-functional theory in the local-density approximation (LDA), that is, using an exchange-correlation potential dependent only on the local electronic charge density. Such an LDA is not quite satisfactory; in particular, the energy gaps of semiconductors are generally significantly underestimated (even though the shapes of both valence and conduction bands are well described). This shortcoming can be mended by going beyond the one-electron approximation and calculating the quasiparticle self-energies including (approximately) many-body effects, or adopting a time-dependent version of the density-functional theory.

Band Structures of Si, Ge, and α-Sn

The band structures of group IV elemental semiconductors show pronounced similarities and definite chemical trends. Strong analogies are expected in view of the identical configuration of the outermost electronic shell and of the identical crystalline symmetry. At the same time, going from Si to α-Sn, decreasing energy gaps accompany larger lattice constants due to the increasing size of the atomic cores.

The band structure of Si, qualitatively similar to that of diamond, is shown in **Figure 4** neglecting relativistic effects. The top of the valence band is at the

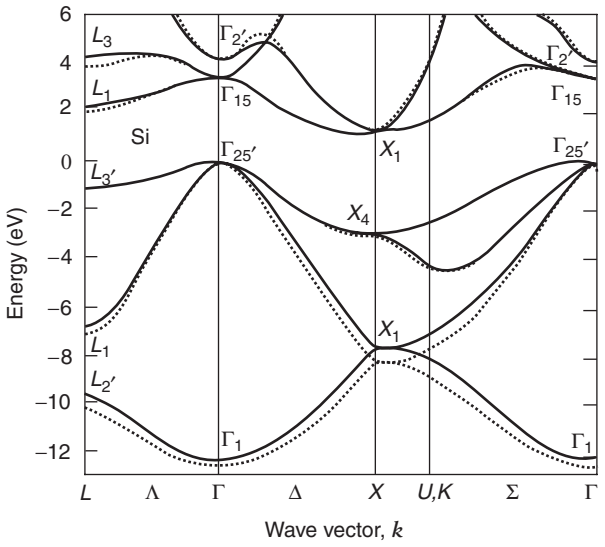


Figure 4 Band structure of Si from an empirical nonlocal (full line) or local (dotted line) pseudopotential calculation. (Reprinted figure with permission from Chelikowsky JR and Cohen ML (1976) Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Physical Review B* 14: 556; © American Physical Society.)

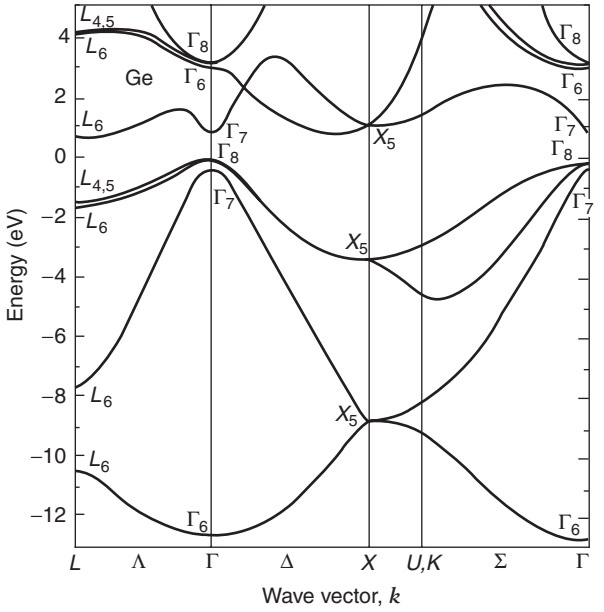


Figure 5 Band structure of Ge from an empirical nonlocal pseudopotential calculation. (Reprinted figure with permission from Chelikowsky JR and Cohen ML (1976) Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Physical Review B* 14: 556; © American Physical Society.)

Γ point and corresponds to the $\Gamma_{25}^{'}$ level (three times degenerate, spin excluded). The bottom of the conduction band is along the Δ line from the Γ to the X point at $\boldsymbol{k} \simeq (2\pi/a)(0.85, 0, 0)$ and has a Δ_1 symmetry (nondegenerate, spin excluded). Thus, the conduction electrons have six pockets centered at the six equivalent conduction band minima (multivalley structure); they are ellipsoids of revolution elongated along the $\langle 1\,0\,0 \rangle$ directions with a longitudinal mass $m_{\mathrm{L}} = \simeq 1.0m$ and a transverse mass $m_{\mathrm{T}} \simeq 0.2m$, as measured from the cyclotron resonance. A peculiarity of the band structure of silicon, showing up in its optical spectrum around 3.4 eV as discussed below, is the fact that along the Λ line from the Γ to the L point the higher valence band (of symmetry Λ_3) and the lowest conduction band (of symmetry Λ_1) are very nearly parallel.

The band structure of Ge including the spin–orbit coupling interaction is shown in **Figure 5**. This is a relativistic effect which, similar to the atomic case, can reduce the degeneracy of the levels. For instance, without spin–orbit coupling, the top of the valence band would be, just as in Si, a p-bonding-like $\Gamma_{25}^{'}$. In the presence of spin–orbit coupling, the sixfold degenerate states obtained from the $\Gamma_{25}^{'}$ ones by the inclusion of spin are split into a fourfold degenerate Γ_8^+ level and a twofold degenerate Γ_7^+ level, which play the role, respectively, of $J = 3/2$ and $J = 1/2$ total angular momentum states. The corresponding spin–orbit gap is $\Delta = 0.3$ eV (the corresponding value for Si is only 0.04 eV). Besides spin–orbit splitting, other relativistic effects become more important in the heavier Ge than in Si; in particular, there are corrections affecting primarily s-derived states which lower their energy. These effects and the presence in the core of Ge of d electrons that, together with p electrons, contribute to pushing the Γ_{15} states to higher energies (corresponding to the Γ_8^- and Γ_6^- spin–orbit split states), through orthogonalization to the core states, explain why in Ge the s-derived $\Gamma_2^{'}$ level (i.e., the Γ_7^- state, spin included) is lower than the Γ_{15} states. The interchange of the conduction band energy level sequence at the Γ point with respect to Si is accompanied by the fact that the minimum of the conduction band in Ge occurs at the L point at the BZ border (the L_6^+ level). The conduction band of Ge, therefore, leads to a multivalley structure with four electron pockets (half pocket for each of the eight L points at the BZ border); they are ellipsoids of revolution elongated along the $\langle 111 \rangle$ directions. The values of the longitudinal mass $m_{\mathrm{L}} \simeq 1.6m$ and of the transverse mass $m_{\mathrm{T}} \simeq 0.08m$ have been determined from cyclotron resonance measurements.

The top of the valence band in both Si and Ge is at the center of the BZ: $\Gamma_{25}^{'}$ without spin, Γ_8^+ and Γ_7^+ including spin, as discussed above. The Γ_7^+ split-off hole states can be described as a simple band with an effective mass $m_{\mathrm{so}} \simeq 0.25m$ in Si and $m_{\mathrm{so}} \simeq 0.08m$ in Ge. In a rough approximation, the Γ_8^+ hole states can be described as belonging to two distinct spherical

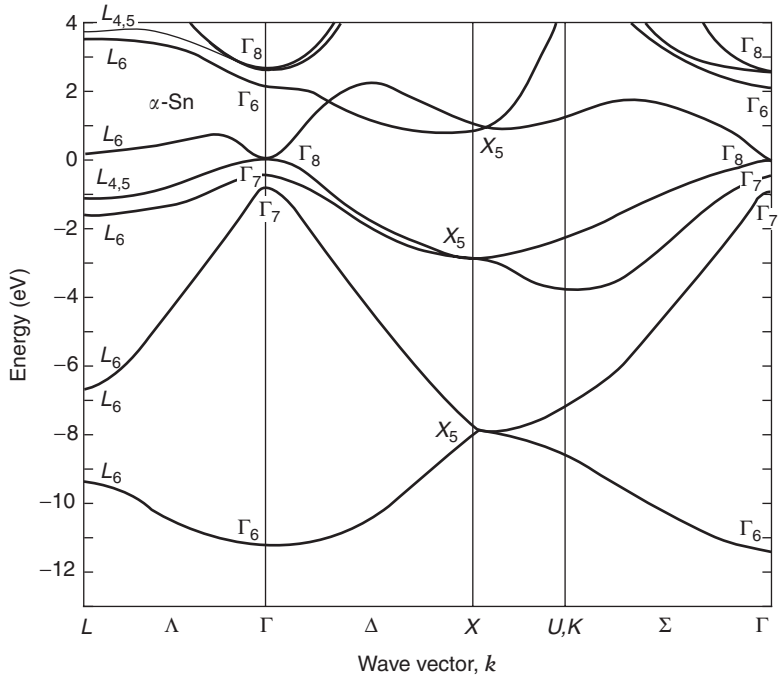


Figure 6 Band structure of α-Sn from an empirical nonlocal pseudopotential calculation. (Reprinted figure with permission from Chelikowsky JR and Cohen ML (1976) Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Physical Review B* 14: 556; © American Physical Society.)

pockets each characterized by an isotropic mass: the heavy-hole pocket (with mass $m_{\mathrm{hh}} \simeq 0.5m$ for Si and $m_{\mathrm{hh}} \simeq 0.3m$ for Ge) and the light-hole pocket (with mass $m_{\mathrm{lh}} \simeq 0.16m$ for Si and $m_{\mathrm{lh}} \simeq 0.044m$ for Ge). Also, these hole effective masses have been determined from cyclotron resonance measurements in low magnetic fields (classical regime). However, this simple picture is not quite correct: because of the valence band degeneracy, the hole bands are not isotropic, but "warped" (particularly, the heavy-hole ones). A complete description of the hole states can be given using the multiband effective mass approach, in which the valence band degeneracy is fully accounted for.

The tendency to have a lower energy *s*-derived Γ_7^- level is even stronger in gray tin where this level is positioned between the spin–orbit split Γ_8^+ and Γ_7^+ levels ("inverted gap" configuration). As a consequence, the band structure of α-Sn, shown in **Figure 6**, is characterized by the fact that the fourfold degenerate (spin included) Γ_8^+ states correspond, at the same time, to the topmost filled valence band and to the lowest empty conduction band which touch each other at the Γ point. Gray tin is thus a "zero-gap semiconductor" rather than a semiconductor proper.

Comparison with Experimental Data

Just as for atoms and molecules, even for solids, the most direct and informative experimental techniques to study their electronic structure over a wide range of energies are optical spectroscopy and photoelectron spectroscopy.

In crystalline semiconductors such as Si and Ge, the absorption due to direct interband transitions is very strong and reflectivity measurements are more popular than transmission ones. From reflectivity spectra taken over a large frequency range, it is possible to obtain the real and imaginary parts of the complex dielectric function $\tilde{\varepsilon} = \varepsilon_1 + \mathrm{i}\varepsilon_2$ (via Kramers–Kronig analysis) and, thus, all the linear optical properties. From the theoretical point of view, the imaginary part $\varepsilon_2(\omega)$ can be calculated summing over all possible transitions between filled (valence) states and empty (conduction) states corresponding to the absorption of a photon of frequency ω. As the wavelength of light is much larger than the lattice constant, such optical transitions take place between electronic states having the same wave vector within the first BZ (direct or "vertical" transitions) and differing in energy by $\hbar\omega$, thus:

$$\varepsilon_2(\omega) = \frac{1}{\pi} \sum_{v,c} \int_{\mathrm{BZ}} \mathrm{d}^3k |d_{cv}|^2 \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega)$$

where d_{cv} is the transition electric dipole matrix element between the initial and final state. It is to be noticed that, because of the energy-conserving delta function in the equation above, specific singularities in the frequency dependence of ε_2 are obtained when $\nabla_k(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k})) = 0$ (critical points). Such

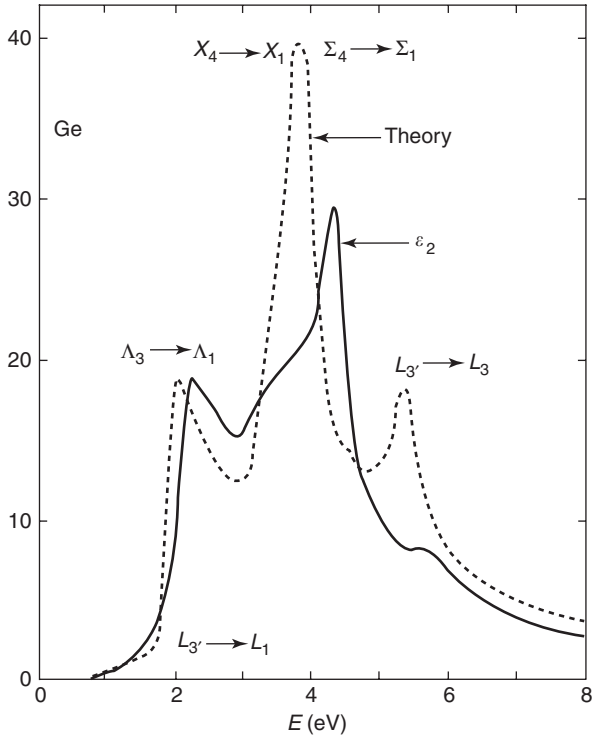


Figure 7 Imaginary part of the dielectric function of Ge: the solid line is experimental and the dashed one theoretical; optical transitions at critical points are identified. (Reprinted figure with permission from Brust D, Phillips JC, and Bassani F (1962) Critical points and ultraviolet reflectivity of semiconductors. *Physical Review Letters* 9: 94; © American Physical Society.)

structures (many of which are dictated by symmetry alone) have been found in all elemental semiconductors, and this has allowed the identification of the critical-point transition peaks and the verification of the band structure.

As a classical and pioneering example, **Figure 7** reports the optical excitation spectrum of Ge, measured as explained above, and computed from the equation above on the basis of an empirical pseudopotential band structure calculation. It is emphasized that not only have the critical points been identified, but also the full band structure has been put to test by the described analysis of the optical properties. A better resolution of the critical point behavior can be obtained by employing various kinds of modulation spectroscopy (e.g., piezoreflectance, electroreflectance) that amounts to measuring the first (or higher) derivatives of the dielectric function. As a further example, one considers a recent calculation of the optical absorption of Si shown in **Figure 8** in comparison with experimental data. The time-dependent density-functional theory includes electron–hole correlations (i.e., excitonic effects) which are crucial to reproduce the conspicuous peak at about 3.4 eV associated with interband transitions near the L point.

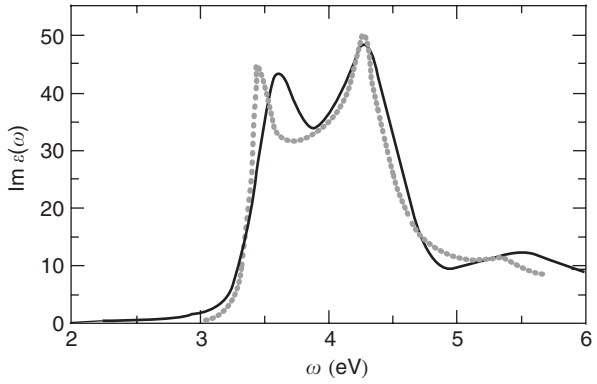


Figure 8 Imaginary part of the dielectric function of Si: the dots are experimental data and the solid line theory. (Reprinted figure with permission from Onida G, Reining L, and Rubio A (2002) Electronic excitations: density functional versus many-body Green's-function approaches. *Review of Modern Physics* 74: 601; © American Physical Society.)

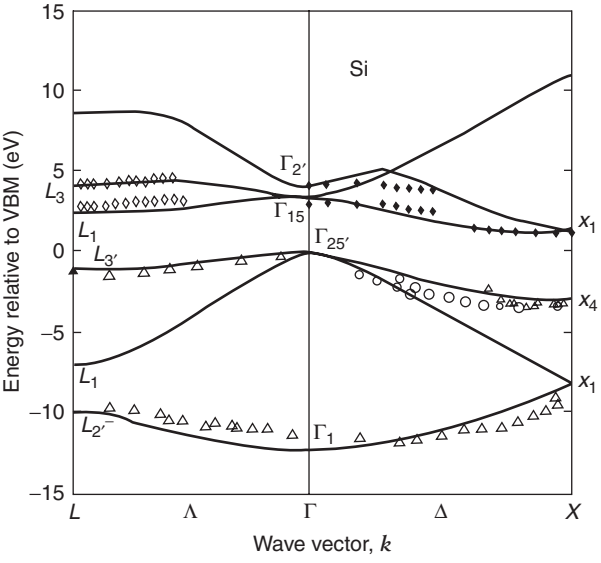


Figure 9 Conduction and valence band dispersions of Si from angle-resolved photoemission and inverse photoemission. (Reprinted figure with permission from Ortega JE and Himpsel FJ (1993) Inverse-photoemission study of Ge(100), Si(100), and GaAs(100): bulk bands and surface states. *Physical Review B* 47: 2130; © American Physical Society.)

The direct optical transitions discussed so far are usually dominant when allowed by energy conservation. It is however possible, through the absorption or emission of a phonon, to have indirect transitions in which the wave vector of the initial and final electronic states are different. These transitions are much weaker, but in the case of indirect gap semiconductors such as Si and Ge having the top of the valence band and the bottom of the conduction band at different points in the BZ, they give rise to absorption at energies below the onset of the much stronger direct transitions. The corresponding

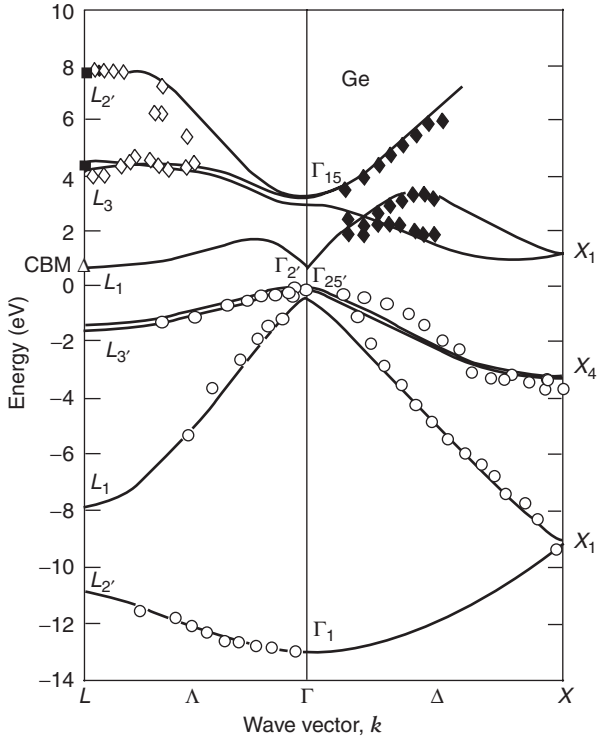


Figure 10 Conduction and valence band dispersions of Ge from angle-resolved photoemission and inverse photoemission. (Reprinted figure with permission from Ortega JE and Himpsel FJ (1993) Inverse-photoemission study of Ge(100), Si(100), and GaAs(100): bulk bands and surface states. *Physical Review B* 47: 2130; © American Physical Society.)

(temperature-dependent) characteristic features have also been observed and compared with theory providing additional evidence for the band structure determination.

Photoelectron spectroscopy has also proved to be a very useful experimental technique to measure the electronic structure of semiconductors. The photoemission process takes place through the following steps: first, upon absorption of an incoming photon, an electron is excited from a filled band to a higher empty band; then, the excited electron travels towards the surface and, finally, if energetic enough, it escapes into the vacuum to be detected. Thus, low-energy conduction states below the vacuum level are not involved in photoemission processes as they can be neither initial states (empty in the ground state) nor final states (unable to escape from the solid once excited). These levels can be investigated by inverse photoemission which is, in a sense, the time-reversed process of photoemission: an incoming electron falls in an electronic level available in the solid, emitting a photon which is then detected. In photoelectron spectroscopy, the energy, wave vector, and polarization (spin) of the photon (electron) can be selectively chosen or averaged over. The main advantage of photoemission and inverse photoemission with respect to optical spectroscopy is the capability of providing absolute energy levels rather than energy differences between levels. Besides a lower resolution and (especially for inverse photoemission) lower statistics, the main disadvantage for the study of bulk properties is that photoelectron spectroscopy is much more surface sensitive because of the short escape or penetration depths of electrons (~ 10 Å).

The detailed analysis of angle-resolved (inverse) photoemission data allows one to obtain the complete band structure of a semiconductor. In **Figures 9** and **10**, the energy bands of Si and Ge obtained by first-principle quasiparticle calculations are compared with photoemission and inverse photoemission data showing very good agreement.

In conclusion, all the main features of the band structures of elemental semiconductors are firmly established both from the theoretical and the experimental point of view. These results represent one of the most important achievements in solid state physics and provide a framework within which virtually all material properties can be understood.

See also: Electrons and Holes; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Optics.

PACS: 71.20.Mq

Further Reading

Bassani F and Pastori Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Cohen ML and Chelikowsky JR (1988) *Electronic Structure and Optical Properties of Semiconductors*. Berlin: Springer.

Jancu J-M, Scholz R, Beltram F, and Bassani F (1998) Empirical *spds** tight-binding calculation for cubic semiconductors: general method and material parameters. *Physical Reviews B* 57: 6493.

Yu PY and Cardona M (1995) *Fundamentals of Semiconductors*. Berlin: Springer.

Elementary Excitations in Biological Systems

H J M de Groot, J Matysik, and J Zaanen,
University of Leiden, Leiden, The Netherlands

Introduction

In condensed matter physics, it would be difficult to understand the fate of energy injected into a solid or a quantum liquid without invoking the notion of elementary excitation. An elementary excitation is an entity with a well-defined identity, interacting weakly with others of its kind, which is to a good approximation responsible for the dynamic evolution in the system when it is not driven too far from equilibrium. Quasiparticles are more often than not highly collective entities which have little, if anything, to do with the microscopic constituents of the system such as electrons or ions. Examples are the Landau quasiparticles in metals (electrons accompanied by an electronic screening cloud), the polarons (electrons dressed up with a phonon cloud), as well as topological excitations such as solitons, domain walls, and vortices involving a reshuffling of an infinity of microscopic degrees of freedom. In systems characterized by a degree of translational symmetry, elementary excitations are delocalized. However, when the symmetry is also broken by, for example, lattice defects, localized excitations are still considered essential to arrive at a proper theoretical description of the solid state.

While elementary excitations are nowadays central to the understanding of the physics of "hard" matter, the idea barely plays a role in the current descriptions of the biological solid state, in particular, the enzymes built up from soft condensed matter at the mesoscopic scale. While it lacks the regularity of crystalline matter, biological matter is quite orderly and it can be regarded as more similar to a solid than to a simple fluid, lacking structure. This problem is either approached in the language of chemistry with its emphasis on the microscopic scale (e.g., electrons covalently binding atoms) or from the perspective of molecular biology: protein function and the mechanisms of protein operation are linked to protein structure and the DNA sequence of a protein, according to the central dogma of molecular biology. Although still limited in scope, biophysicists have made some progress over the years in changing this situation and it appears that at least in some circumstances, the concept of elementary excitation is quite useful.

Since a biological system serves a purpose in the biological hierarchy, the energy entering the system is used in a coherent and orderly fashion through a time sequence of highly nonlinear self-organization events. For instance, enzymes involved in transport, energy transduction, or signal transduction processes operate unidirectionally, like molecular machines. Such devices rely on highly nonlinear self-organization mechanisms, a chain of elementary processes to channel thermal or external energy from one subprocess into the next. The resulting energetic pathway fulfills both the requirements set by the evolutionary constraints that apply to the system and the universal physical principle that physical changes and processes tend to take the easiest or least action path. This is possible since a protein is a biological construct; its folding is downward causal, and is controlled by the cellular hierarchy. Every protein is a small architecture, a compartment that is metastable. In this compartment, the energy pathways are part of the design properties.

Biology builds upon "stochastic machines" that are robust against noise. As illustrated with various examples, experimental evidence emerges that the notion of an elementary excitation is quite helpful in understanding the workings of these biological machines. This "mechanics" typically involves metastable states of the protein, with a well-defined identity, involving cooperative change of the equilibrium structure of the protein on a larger scale. As in the solid, this excitation is largely responsible for the dynamical evolution of the system when it is perturbed away from equilibrium. A difference with conventional solids is that the lack of symmetry renders these excitations to be quite a bit more diverse, "complex." Due to this complexity, mathematical theory is of much less use than in conventional condensed matter physics, while the standard methods of experimental physics are not quite suited to probe the time and length scales of relevance to protein dynamics. It is a blessing in disguise, however, that this complexity makes possible that these biological systems can resemble the machinery of the human world. The various examples discussed below might further illustrate the case. The discussion starts with the short-time dynamics in various photoactive systems where the analogies with condensed matter systems are quite obvious. The focus is then on the dynamics on larger lengths and timescales (ion channels, motor proteins, substrate channeling), where the situation is much less clear because of the aforementioned difficulties.

Light-Induced Electron Transfer in Photosynthetic Reaction Centers

In photosynthesis, the reaction center (RC) uses the sunlight to pump electrons to generate chemicals with a high redox potential. The O_2 synthesized by plants and trees drives the respiratory chain of many of the known life forms, and, therefore, the photosynthetic electron pumping complex reaction chain is among the most important chemical reactions for life on earth. The initial photochemical process in green plants occurs in RC protein complexes bound to the photosynthetic membrane of chloroplasts. Two RCs are working sequentially: photosystem II has a strong oxidation power, which allows the oxidation of water, and photosystem I has a strong reductive power.

In all types of RCs, light is absorbed by a (bacterio)chlorophyll aggregate, known as the primary electron donor P. In the RC of purple bacteria (**Figure 1**), P is a bacteriochlorophyll dimer, called the "special pair." In the photochemically excited state of the dimer, an electron (polaron) is emitted into one branch of cofactors (**Figure 2**). This functional asymmetry in the energy channel appears to be related to different mesoscopic Coulomb force fields in both branches. In RCs of purple bacteria, the primary electron acceptor is a bacteriopheophytin H cofactor. While the light-induced electron transfer from P to H occurs in 3 ps, the back-reaction to P in its electronic ground state occurs at the millisecond timescale. Originally, the slow-down has been explained by the high Gibbs energy, shifting the back-reaction into the inverted Marcus region. There are, however, several experimental observations, contrasting the view of a static cofactor arrangement, on which the Marcus concept is based. A change of mesoscopic dielectric properties as a consequence of the change of charge-state has been observed on bacterial RCs by optical spectroscopy and by photovoltage measurements and is described by theoretical methods. The temperature dependence of the electron transfer kinetics and photoacoustically detected volume contraction provide evidence for small structural (conformational) changes in the protein when the charge-separated state is established. Upon long illumination, radical-pairs with lifetimes of several minutes have been reported. The elementary excitation is here represented by the highly cooperative changes in the protein in response to the transfer of an electron, which leads to well-defined nonequilibrium states that are important for the function. A mesoscopic nonlinear dielectric screening phenomenon to stabilize the energy channel would help to prevent the back-reaction, and this is of interest for the construction of artificial RCs. In this sense, the operational mechanism of RCs could be described in terms of an electron pump. The piston P, shuttling between different electronic states is distinguished by its affinity to electrons. The valve avoiding electron back-transfer corresponds with the nonlinear dielectric response. Interestingly, a similar terminology was introduced in condensed matter physics by David Thouless already in 1983 for charge pumping and is now common in the field of solid-state photovoltaics.

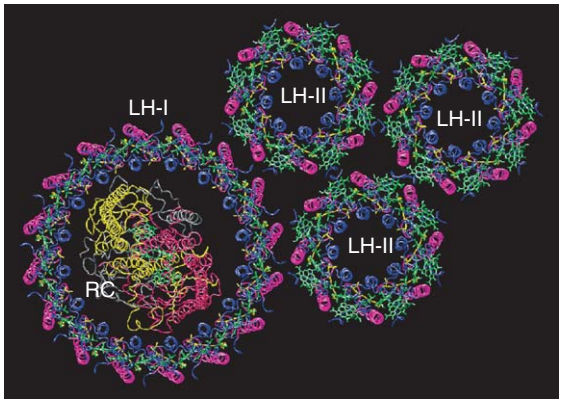


Figure 1 The photosynthetic apparatus of purple bacteria. Reproduced with permission from Hu X, Damjanovic A, Ritz T, and Schulten K (1998) Architecture and function of the light harvesting apparatus of purple bacteria. Proceedings of the National Academy of Sciences, USA, 95: 5935–5941, © 2005 National Academy of Sciences.

Excitons in Photosynthetic Antenna Complexes

In photosynthesis, light energy is used to drive chemical processes. Under average conditions, the photon

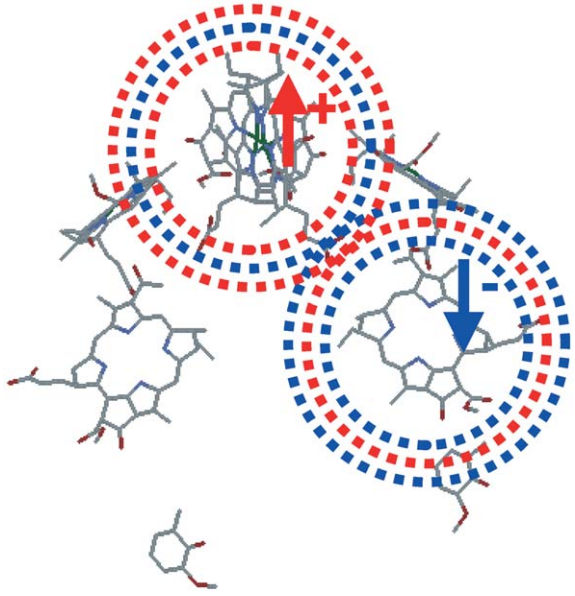


Figure 2 Charge separation in a photosynthetic reaction center. (Source: Leiden Institute of Chemistry, Leiden, NL.)

flux is limiting for photosynthesis. To improve the performance, all photosynthetic systems embed RCs in energy concentrators: arrays of light-harvesting antenna complexes that increase the photosynthetic efficiency. In most bacteria, the ratio is ~40 antennae to one RC; in plants, however, ratios of several hundreds antenna per RC are common. The structure of antenna complexes varies widely. The membrane-integrated antenna LH1 and LH2 of purple bacteria have ring-shaped structures (**Figure 1**). In LH2, 18 BChl *a* molecules (B850) are sandwiched between two polypeptides and form a continuous overlapping ring-shaped superstructure. Another ring of BChl *a* (B800) is positioned between the subunits of the outer cylinder with the bacteriochlorin rings perpendicular to the transmembrane helix axis. The entire LH2 complex has a ninefold rotational symmetry. Chlorosome antennae of green bacteria obtain a size of several micrometers and are bound to the surface of the photosynthetic membrane where they have contact with several dozens of RCs. The interior of chlorosomes is formed by self-aggregation of bacteriochlorophyll molecules, a process which can also be observed in an apolar solvent.

The elementary excitations in photosynthetic antenna systems are excitons. Despite differences in their structures, all antenna share similar functional properties: absorption of light over a broad spectral range, storage of this excitation in the form of an exciton, and excitation transfer into the RC. There is currently a discussion about the exact description of the exciton in LH2, whether it is localized on a single B850 cofactor and hopping around the superstructure, or delocalized over several B850 cofactors. Excitations in chlorosome-type aggregates have been shown to be delocalized over at least 10–15 BChl molecules and can migrate through the superstructure.

Double Bond Isomerization in Rhodopsin

The next example of an elementary excitation presented here is the polaronic-type structure that drives the isomerization process in the rhodopsin photoreceptor of vision, located in the disk membranes of the rod outer segment of the retina of higher organisms. Rhodopsin has a retinal chromophore in an 11-*cis* configuration that is bound to the protein by a covalent Schiff base bond to the nitrogen of a lysine residue L-296 (upper panel in **Figure 3**). The nitrogen is protonated and stabilized by a negative counterion. The proton positive charge is delocalized into the polyene chain of the retinal. The charge in the electronic ground state can be observed with NMR. This is illustrated in the lower panel of **Figure 3**. The color-coded spheres over the carbons represent an

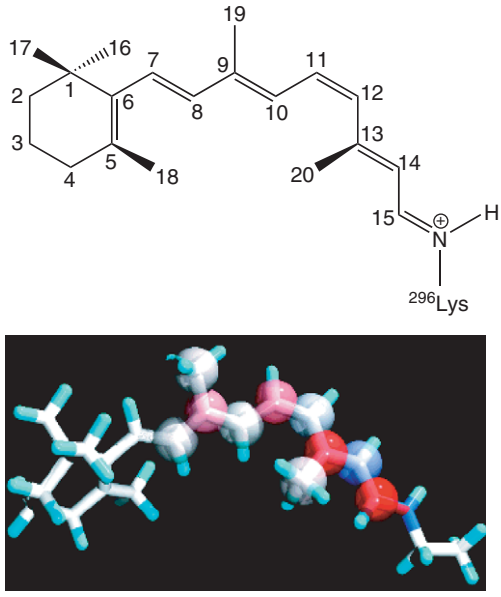


Figure 3 Retinal chromophore.

NMR "snapshot" of the chemical shifts relative to a neutral polyene. These shifts translate into charge, from excess negative charge (blue) to strong positive charge (red). The alternating charge pattern over several bonds in **Figure 3** corresponds with a trapped positively charged polaron at the odd-numbered lattice sites with a half-width of several bonds. This example shows how the self-organization by nature produces biological constructs that are physically in a metastable state, far out of equilibrium. The rhodopsin can adapt to environmental constraints. For instance, the electronic structure of the extended π-system and the polaron can be tuned by the properties of the protein pocket by variation of Coulomb fields, which cause stabilization or destabilization of the different electronic states (Opsin shift). In this way, variations in the protein can affect the HOMO–LUMO gap and the action spectrum of vision.

Upon excitation with light, the 11-*cis* double bond isomerizes to *trans* within 200 fs and this chemical reaction triggers the visual signal transduction cascade. The precise energy channel is difficult to assess experimentally, but *ab initio* molecular dynamics suggests that a light-induced flip of the *cis*-hydrogen bound to *C*-12 to the corresponding *trans*-position relative to *C*-11 yields a strained *trans*-chromophore in the 11-*cis* binding pocket. Upon excitation with light, the positive polaronic charge in **Figure 3** is transferred to the left side of the polyene, at the even lattice sites. This initiates the motion of the charge along the polyene backbone in the excited state, where it is attracted by the negative counterion charge at the Schiff base end. Since the electronic

charge is present in the form of a polaron, it increases its effective mass by two orders of magnitude, by autodressing with phonons on a very short timescale, $\sim$20 fs. Hence, the electronic charge is a very heavy particle on the 200 fs timescale of the isomerization. This leads to a classically coherent channel in the excited state that keeps the energy focused. The potential energy stored in the excitation is partly transferred into kinetic energy of the atoms and for another part into molecular strain by the isomerization. The polaron carries a node in the electronic wave function into the isomerization region that is eliminated in a quantum nuclear tunneling event, and leads to a spring-loaded molecule that activates the receptor protein in a series of first-order dark reactions: the visual cascade. Every subsequent step in this cascade takes more time, from picoseconds to nanoseconds, microseconds, milliseconds, and seconds. At every intermediate, the system partially relaxes and prepares itself for the next step, on a longer time and length scale. Hence, a least action path is effectuated by a compartmentalization in space and time, where every self-organizing process in a chain of events helps to stabilize the energy in the corresponding subsystem.

Proton Transport in Bacteriorhodopsin

Another retinal protein, bacteriorhodopsin, is a light-driven proton pump found in halophilic archaebacteria. During the photocycle, a proton is pumped through this protein. The difference in pH and electric potential associated with the proton gradient is used for the synthesis of the cellular "fuel" adenosine triphosphate (ATP). Similar to the rhodopsin, the retinal Schiff base in bacteriorhodopsin is protonated and stabilized by a complex counterion that is part of the proton conduction pathway. A schematic representation of the proton pathway is presented in **Figure 4**. A complex counterion in a biological solid is an ice-like structure and can support elementary excitations in the form of ionic defects and Bjerrum defects that participate in a Grotthus proton transfer mechanism. The figure shows the proton background potential for the idealized case of a one-dimensional channel of hydrogen-bonded atoms. The potential energy is highest at the position of the heavy atoms, and the two minima between each pair of atoms are separated by a barrier u and correspond with symmetric hydrogen bonds. An excess proton in the chain forms an ionic defect extending over $d \sim 3$ atoms. Such defects are thermodynamically stable if the proton chemical potential in the protein μ_p exceeds the defect creation energy μ_c. To first order, the $\mu_c \sim j/d + u \times d$, that is, the energy cost associated with improper charge balance between adjacent atoms due to the Coulombic interactions between the protons (measured by j), and the energy cost from the energy barriers u between the minima in each hydrogen bond. Stability requires that $\partial\mu_c/\partial d = 0$ and $\mu_c \sim 2\sqrt{j \times u}$ with $d \sim \sqrt{j/u}$. This illustrates how a simple self-organization process can give rise to an elementary excitation in a protein water channel. It has been experimentally verified that such excitations can migrate. The H-bonding network, required for such Grotthus/Bjerrum proton transfer mechanism, has been found by X-ray structure analysis in many cases, for instance in bacteriorhodopsin and in photosynthetic RC electron-pumping proteins. In cytochrome *c*-oxidase, it was proposed that a H-bonded unit may operate as a temporary storage device for multiple elementary excitations, to accumulate energy for its electron pump mechanism. Mechanistic details about structural changes of the protein during the proton pumping in the bacteriorhodopsin protein are currently under discussion.

Conformational Changes during Molecular Transport

The broad definition of an elementary excitation as a well-defined entity that is responsible for the dynamic evolution of a system encompasses a wide variety of biological machines that operate with mechanical changes. In this context, the elementary excitations are the few unique conformations of the protein that it cycles through during its function. For instance, ion channels are membrane-spanning proteins forming a

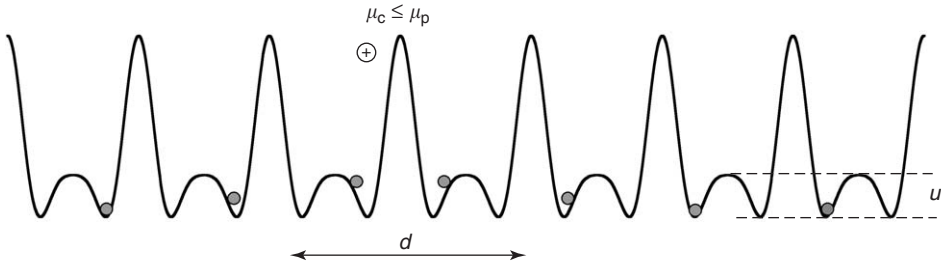


Figure 4 A one-dimensional proton wire.

pathway for the flow of inorganic ions across cell membranes. Like enzymes, ion channels have their specific substrates as potassium, sodium, calcium, and chlorine ions. Potassium channels exhibit a remarkable ability to discriminate between potassium and sodium by a factor of $\sim 10\,000$, even though these ions are similar in size. Without such selectivity, electrical signal production, which is related to many fundamental biological processes as electrical signaling in the nervous system, would be impossible. The flow of ions occurs down their electrochemical gradient. Although the high selectivity suggests strong energetic interaction between ion and protein, the transport rate approaches the diffusion limit. The flow of ions can be gated by external triggers as ligand binding or membrane polarization. Potassium channels are tetramers of identical subunits forming a pore in the center. Selectivity and gating are the main features of this molecular machine. Cation selectivity is an intrinsic property of the pore architecture originating from a special arrangement of cation-attractive α-helices. This specific sequence of amino acids forming the selectivity filter is conserved in all potassium channels. In potassium channels opening in response to intracellular calcium, the receptor binding leads to structural changes opening the gating ring at the intracellular membrane surface. Voltage-dependent potassium channels are switched by "voltage-sensor paddles" on the outer perimeter of the channel surface. These paddles positioned inside the membrane contain positively charged arginine residues. The angle of the paddles relative to the axis of the pore changes with the polarization of the membrane. This movement controls the gate of the pore.

Molecular motors or "mechanochemical proteins" convert chemical energy, usually in the form of ATP, or the potential energy of an ion gradient into mechanical forces and motion. The motion can be translational or rotational. Molecular motor proteins that carry cargo directly along a cytoskeletal track are myosins (along actin) and kinesins and dyneins (along microtubules). Rotary motors are bacterial flagelar motors and F_1F_0 ATP synthase (**Figure 5**). The molecular structures of these motors show a large variation. In the past, it was generally assumed that motor proteins generate forces by rectifying thermal fluctuations. According to such fluctuation ratchets models, motion is not produced directly by the chemical reaction or diffusive process. Feynman had shown that Brownian ratchets cannot produce a net-transport at equilibrium conditions but require a gradient. Under nonequilibrium conditions, transport can occur due to fluctuations correlated to the surroundings. Driving by such colored noise is called stochastic resonance. This means, that random walk moves the

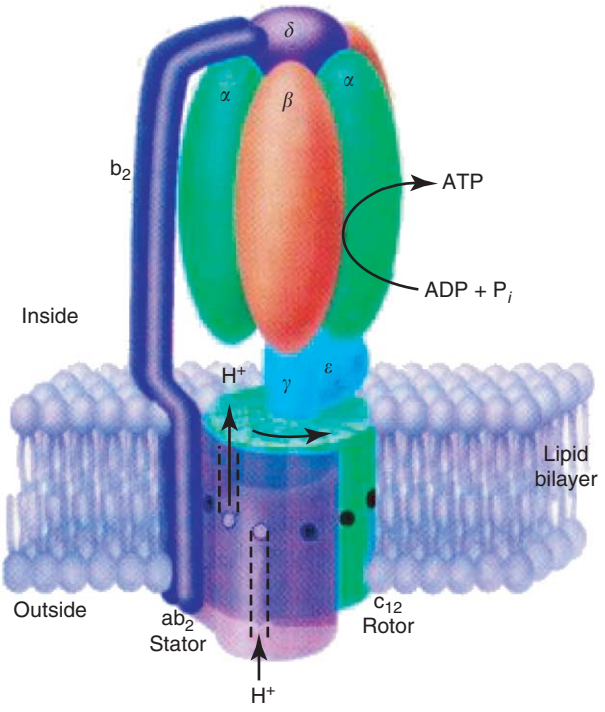


Figure 5 ATP synthase.

motor, and a linked chemical or diffusional process biases the walk such that forward steps become more likely than back steps. However, recent experiments have shown that motor proteins "walk" over a biological rail by a series of conformational changes. Dyneins, for example, are ATP-driven microtubule motors assembled from three parts: stalk, head, and stem. The stalk is linked to the track, and the stalk and stem are connected to the head similar as the two hands of a clock. During the chemomechanical cycle, the positions of the hand change. Here, the elementary excitation is the coherent dynamics in the form of a stepping motion involving conformational changes when it burns ATP.

Finally, direct transfer of metabolites from one protein to another or from one active site to another in the same protein is called substrate channeling. Such direct link decreases transport time of metabolites, prevents loss by diffusion, protects reactive intermediates from solvent interaction, and avoids interference with competing metabolic pathways. The longest channeling pathway discussed currently is that connecting the three active sites of carbamoyl phosphate synthase (CBS) separated by a linear distance of ~ 100 Å. One of the best-studied examples is the intramolecular hydrophobic tunnel in tryptophan synthase connecting two active sites of ~ 25 Å distance. As a channeling mechanism, tunneling by Brownian dynamics as well as electrostatic attraction

of the metabolite are considered. Although conformational changes are not included in the traditional definition of elementary excitations in condensed matter physics, they do generally fit into the broader definition. It is a challenge for condensed matter biophysicists to deal with these elementary excitations experimentally, and to understand the dynamic evolution theoretically. One can only speculate at present how far this can go, from the molecular level upwards, to the cellular level, and beyond.

See also: Bioelectronics; Biomolecules, Scanning Probe Microscopy of.

PACS: 87.15.–v; 87.64; 5.45.Yv; 87.14.Ee; 87.16.Nn

Further Reading

Astumian RD (1997) Thermodynamics and kinetics of a Brownian motor. *Science* 276: 917–922.

Bennett CH (1988) Notes on the history of reversible computation. *IBM Journal of Research and Development* 32: 16–23.

Buda F, de Groot HJM, and Bifone A (1996) A charge localization and dynamics in rhodopsin. *Physical Review Letters* 77: 4474–4477.

Feynman RP, Leighton RB, Sands M (1963) In: Ratchet and Pawl (eds.) *The Feynman Lectures on Physics*, vol. I, chap. 46. Reading: Addison-Wesley.

Keller D and Bustamante C (2000) The mechanochemistry of moleculr motors. *Biophysical Journal* 78: 541–556.

Laughlin PRB, Pines D, Schmalian J, Stojkovic BP, and Wolynes P (2000) The midle way. *Proceedings of the National Academy of Sciences of the United States of America* 97: 32–37.

Ouporov IV, Knull HR, Huber A, and Thomasson KA (2001) Brownian dynamics simulations of aldolase binding glyceraldehyde-3-phosphate dehydrogenase and the possibility of substrate channeling. *Biophysical Journal* 80: 2527–2535.

Prokhorenko VI, Holzwarth AR, Müller MG, Schaffner K, Miyatake T, and Tamiaki H (2002) Energy transfer in supramolecular artificial antennae units of synthetic zinc chlorins and coaggregated energy traps. A time-resolved fluorescence study. *Journal of Physical Chemistry B* 106: 5761–5768.

Rammelsberg R, Huhn G, and Lübben MK (1998) Bacteriorhodopsin's intramolecular proton-release pathway consists of a hydrogen-bonded network. *Biochemistry* 37: 5001–5009.

Rubin AB, Kononenko AA, Shaitan KV, Paschenko VZ, and Riznichenko GYu (1994) Electron-transport in photosynthesis. *Biophysics* 39: 173–195.

Sham YY, Muegge I, and Warshel A (1998) The effect of protein relaxation on charge–charge interactions and dielectric constants of proteins. *Biophysical Journal* 74: 1744–1753.

Steffen MA, Lao KQ, and Boxer SG (1994) Dielectric asymetry in the photosynthetic reaction-center. *Science* 264: 810–816.

Thouless DJ (1983) Quantization of particle transport. *Physical Review* B27: 6083–6087.

Trissl HW, Bernhardt K, and Lapin M (2001) Evidence for protein dielectric relaxations in reaction centers associated with the primary charge separation detected from Rhodospirillum rubrum chromatophores by combined photovoltage and absorption measurements in the 1–15 ns time range. *Biochemistry* 40: 5290–5298.

van Oijen AM, Ketelaars M, Köhler J, Aartsma TJ, and Schmidt J (1999) Unraveling the electronic structure of individual photosynthetic pigment-protein complexes. *Science* 285: 400–402.

Environmental Science, Physical Principles and Applications

A Speranza and V Lucarini, University of Camerino, Camerino, Italy

Introduction

Environmental science almost invariably proposes problems of extreme complexity, typically characterized by strongly nonlinear evolution dynamics. The systems under investigation have many degrees of freedom – which makes them complicated – and feature nonlinear interactions of several different components taking place on a vast range of time–space scales – which makes them complex. Such systems evolve under the action of macroscopic driving (typically the solar heating) and modulating (e.g., the Earth's rotation and gravitation) agents. The most comprehensive example is the entire climatic system. In its most rigorous definition, the climatic system is constituted by four intimately interconnected subsystems: atmosphere, hydrosphere, cryosphere, and biosphere (**Figure 1**), and is powered by the electromagnetic radiation emitted by the Sun (**Figure 2**). These subsystems are by themselves complicated and complex and interact nonlinearly with each other on various time–space scales.

Moreover, environmental sciences are not usually provided with real laboratories where theories can be tested against experiments, since phenomena often take place only once and cannot be reproduced. Theories can often be tested only against observational data from the past, which may feature problems of various degrees of criticality, essentially because of the physical extension of the systems under analysis. In many cases, the available observations feature a relatively low degree of reciprocal synchronic coherence and individually present problems of diachronic coherence, due to changes in the strategies of data

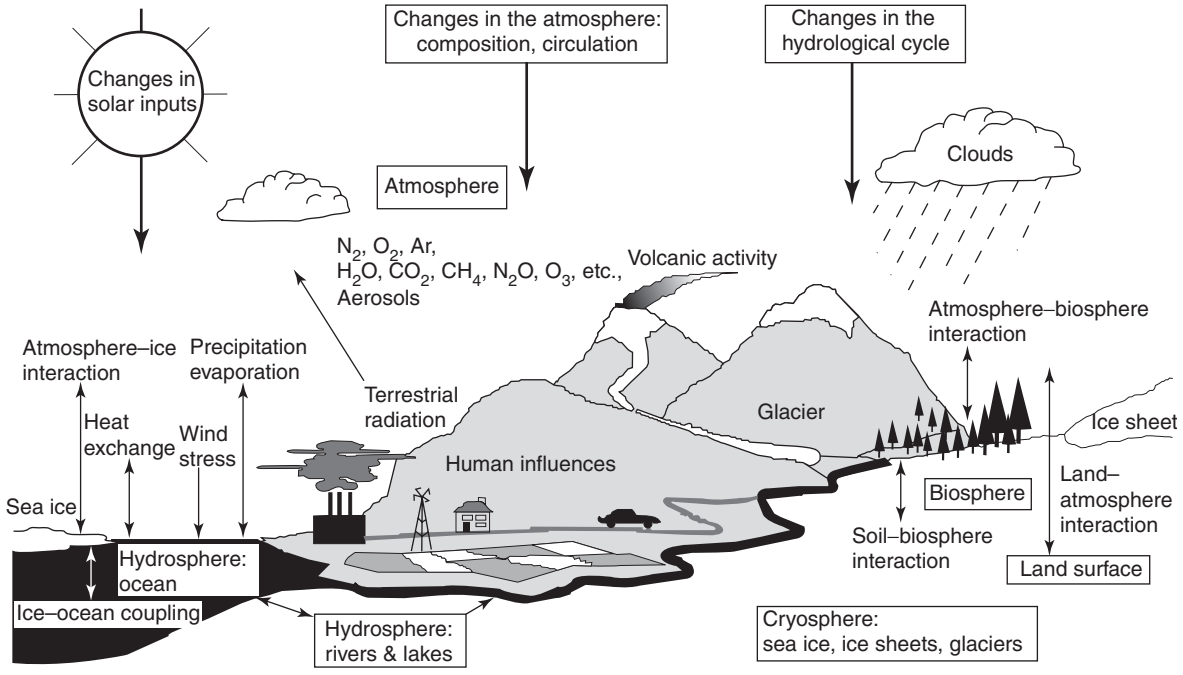


Figure 1 Components of the climatic system; the main interactions are also indicated. (Reproduced with permission from IPCC (2001) *Climate Change 2001: The Scientific Basis*, Contribution of Working Group I to the 3rd Assessment Report of the IPCC. Cambridge: Cambridge University Press.)

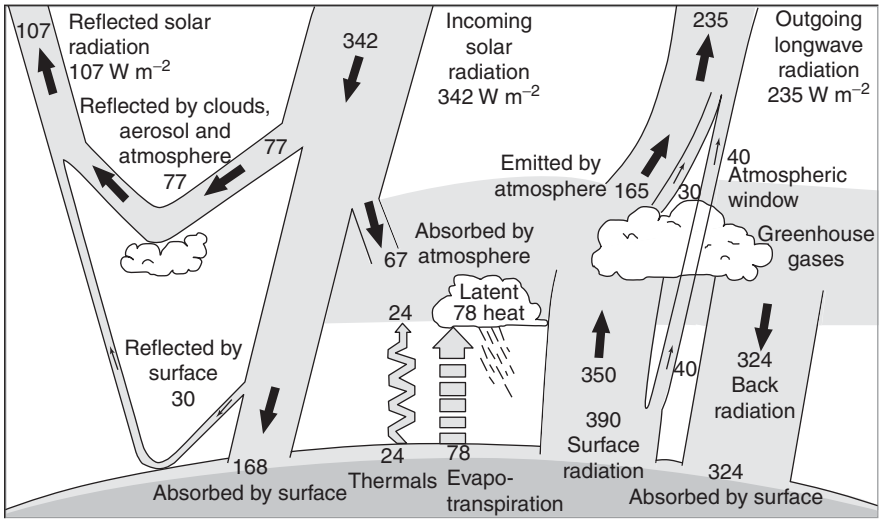


Figure 2 Energy fluxes (in units of W m^{-2}) through the atmosphere. (Reproduced with permission from IPCC (2001) *Climate Change 2001: The Scientific Basis*, Contribution of Working Group I to the 3rd Assessment Report of the IPCC. Cambridge: Cambridge University Press.)

gathering with time. The presence of strong variability of both the outputs of theoretical models and of the real systems contributes to blur the line between a failed and a passed test, in terms of model reliability.

It is important to emphasize that the theoretical models unavoidably suffer from two distinct kinds of uncertainties. The uncertainties in the initial conditions of the systems, termed "uncertainties of the first kind," are problematic because of the effects of chaos. These uncertainties may be partially dealt with using well-suited Monte Carlo approaches on the initial conditions.

The "uncertainties of the second kind" are structural uncertainties, due to the simplifications adopted in the models for the description of the processes and of the feedbacks characterizing the system under

investigation. These latter uncertainties seem to be somehow more fundamental – the dynamic framework is uncertain – and harder to deal with.

Therefore, it is conceptually incorrect to expect that the essentially phenomenological theories which have been developed for the environmental sciences could provide answers having comparable precision and similar structure to those provided by theories relevant for the study of systems having lower degree of complexity.

In view of the above difficulties, it is not surprising that the historical development of "environmental science" took place essentially along two different main paths: on the one side the descriptive, basically qualitative, approach of naturalistic tradition, and on the other the, quantitative formulation of physical–mathematical tradition. Linneo's classification of living beings and Laplace tidal equations are classical examples. The synthesis between the two approaches took place in the nineteenth and twentieth centuries mostly through applications of fluid dynamics and/or thermodynamics to specific systems. However, a certain separation (sometimes even contradiction) between the two approaches occasionally still emerges today. Meteorological maps, interpreted on the basis of the analogies of the so-called "synoptic" meteorology, are examples of the modern evolution in the descriptive tradition, while numerical weather forecasting exemplifies the quantitative one.

The description of the macroscopic dynamics of environmental systems is based on the systematic use of dominant "balances" derived on a phenomenological basis in order to specialize the dynamical equations. Such balances are suitable classes of approximate solutions of the evolution equations which represent a reasonably good approximation to the actual observed fields when sufficiently large spatial or temporal averages are considered. Actually, different balances have to be considered depending on the time and space scales that are being focused on. Such an approach reflects the fundamentally heuristic-inductive nature of the scientific research in environmental sciences, where the traditional reductionistic scientific attitude is not always effective. In order to exemplify this procedure, the very relevant case of the motion of the fluids that permit the existence of life on Earth, air and water: the so-called geophysical fluids is considered.

Geophysical Fluids and Phenomenological Balances

Geophysical fluid systems are very complex in microphysical structure and composition and evolve under the action of macroscopic driving (solar heating) and modulating (Earth's rotation and gravitation) agents. The complexity essentially depends on the presence of nonlinear interactions and feedbacks between the various parts of the system, which couple very different time and space scales. In many cases, the dynamics of such systems is chaotic (the self-correlation of the fields vanishes within a finite time domain) and is characterized by a large natural variability on different timescales.

The dynamics of a geophysical fluid for an observer in a uniformly rotating frame of reference is described by the Navier–Stokes equation:

$$\rho\frac{\mathrm{d}u}{\mathrm{d}t} + 2\rho\Omega \times \boldsymbol{u} = -\boldsymbol{\nabla}p - \rho\nabla\Phi + F \qquad [1]$$

where ρ is the density of the fluid, $\boldsymbol{u} = (u, v, w)$ is the velocity vector, Ω is the Earth's rotation vector, p is the pressure, Φ is the geopotential (essentially coincident with the gravitational potential), F is the frictional force per unit mass, and the total derivative is expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \qquad [2]$$

thus including the nonlinear advection term.

Through the use of suitable approximations, the equations of motion can be focused upon the desired components of the flow. The filtering process is the introduction of a set of mathematical approximations into the Navier–Stokes equation having the scope of filtering out (excluding) solutions corresponding to the physical processes that are heuristically assumed to contribute only negligibly to the dynamics of the system at the time and space scale under examination. The magnitudes of various terms in the governing equations for a particular type of motion are estimated using the so-called scale analysis technique.

Hydrostatic Balance

As a relevant example of these procedures, the "hydrostatic approximation" is considered here. In a local Cartesian coordinate system where z is perpendicular to the surface, an obvious fixed point (stationary solution) of the Navier–Stokes equation can be derived from the time-independent classical "hydrostatic equilibrium" equations:

$$\begin{aligned} \boldsymbol{u} &= (0, 0, 0) \\ -\rho_{\mathrm{h}} g - \frac{\partial p_{\mathrm{h}}}{\partial z} &= 0 \end{aligned} \qquad [3]$$

where the functions $\rho_{\mathrm{h}}(x, y, z, t)$ and $p_{\mathrm{h}}(x, y, z, t)$ satisfying the above balance (the subscript h stands for "hydrostatic") depend only on the spatial variable z.

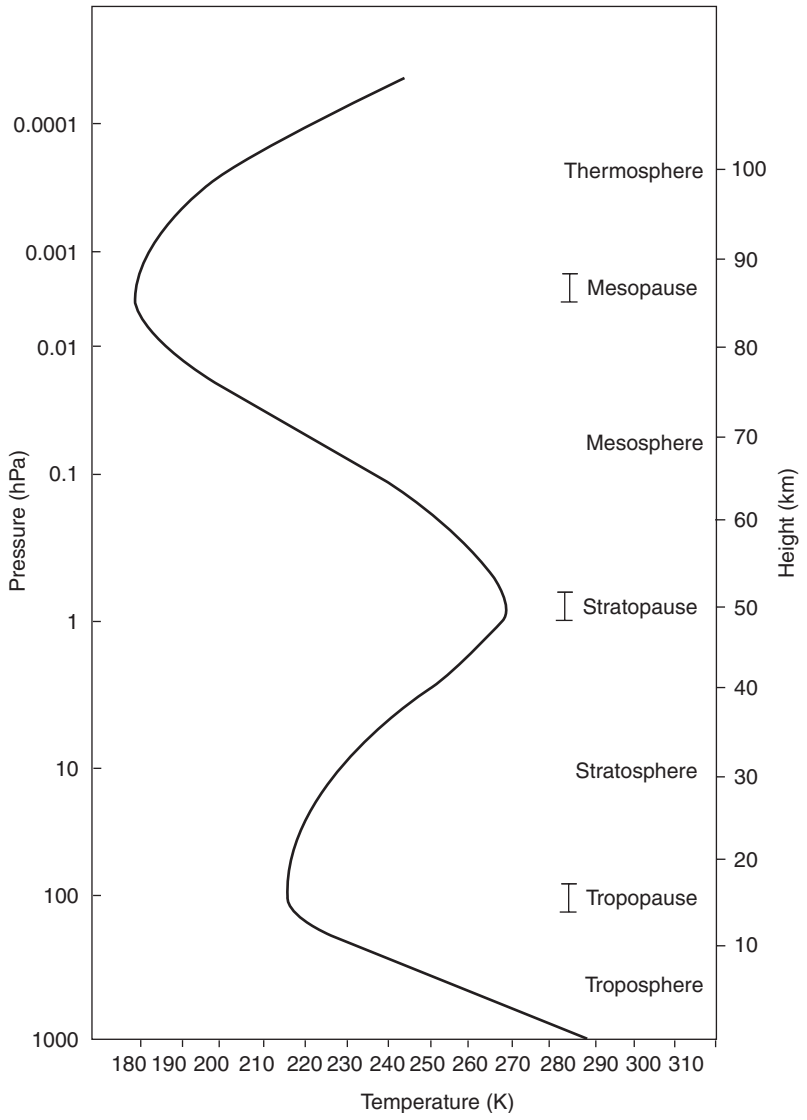


Figure 3 Typical temperature vertical profile of the atmosphere. The main sections of the atmosphere are indicated and the small extent of the quasi-constant temperature sections is emphasized.

Three points related to the above-described balance are worth mentioning explicitly:

1. the fluid velocity is assumed negligible everywhere in the system;
2. the condition of hydrostatic equilibrium does not determine univocally the thermodynamic fields $\rho_h(z)$ and $p_h(z)$ but only their relationship; other information are requested in order to determine univocally the profiles of density, pressure, and temperature; and
3. the balance is observable (and practically useful!) only if it is stable.

Now, the interesting point is that quite different geophysical flows (atmosphere, ocean, surface, and ground water) can be represented as being in near hydrostatic equilibrium when suitable spatial and temporal averaging process is considered (**Figure 3**).

In general terms, hydrostatic equilibrium is realized when the portions of a fluid with lower specific entropy are positioned below those with higher specific entropy, where directions are defined by the local gravity force vector. When this condition is broken because of an external forcing, the stratification is not stable and the fluid rearranges so that hydrostatic condition is re-established. Therefore, violations of hydrostatic equilibrium exist only on short time and space scales and often are not easily observable. Typical models are formulated in such a way that local nonhydrostatic conditions are quickly eliminated and the equilibrium condition that is

recovered is parametrized in terms of variables explicitly represented on the numerical grids. In a later section, the general scope of the parametrization procedures is briefly described.

As a matter of fact, nonhydrostatic models are nowadays available and are currently used. However, for any practical purpose, sufficiently detailed initial and boundary conditions are not available and the above-expressed methodological problems reappear in connection with their definition.

Geostrophic Balance

Another instructive example is that of time-independent purely horizontal balanced flows where the horizontal components of the pressure gradient force and the Coriolis force cancel out in the Navier–Stokes equation. Such flows are termed geostrophic, which, etymologically, means Earth-turning. The structure parameter determining qualitatively the goodness of such approximation is the Rossby number $Ro = U/f_0L$, where U is a characteristic horizontal velocity of the fluid, L is a characteristic horizontal extension of the fluid, and f_0 is twice the value of the projection of Earth angular velocity vector on the plane tangent to the surface of the Earth at the considered latitude. This approximation holds for relatively large-scale extratropical flows in the regions where friction is not important – that is, away from the boundaries of the fluid. Geostrophic equations are obtained as a zeroth-order Ro expansion of the Navier–Stokes equation. The fluid motion is introduced by considering small perturbations breaking the translation symmetry of the baseline purely hydrostatic density and pressure fields. This is achieved by assuming that the actual density and pressure fields are given by $\rho(x,y,z) = \rho_g(x,y,z) + \rho_h(z)$ and $p(x,y,z) = p_g(x,y,z) + p_h(z)$, respectively. Geostrophic equations thus result to be the following:

$$\begin{aligned} \rho_h u f_0 &= -\frac{\partial p_g}{\partial y} \\ \rho_h v f_0 &= -\frac{\partial p_g}{\partial y} \\ g\rho_g &= -\frac{\partial p_g}{\partial z} \end{aligned} \qquad [4]$$

Geostrophic and hydrostatic balance constrain atmospheric motions so that the fluid flows are uniquely determined by the pressure field, since the currents are parallel (rather than perpendicular as in nonrotating fluids) to the isobars at a given geopotential height (**Figure 4**). The time-independent nature of these equations implies that the nongeostrophic terms, although relatively small, are important for the time evolution of the flow.

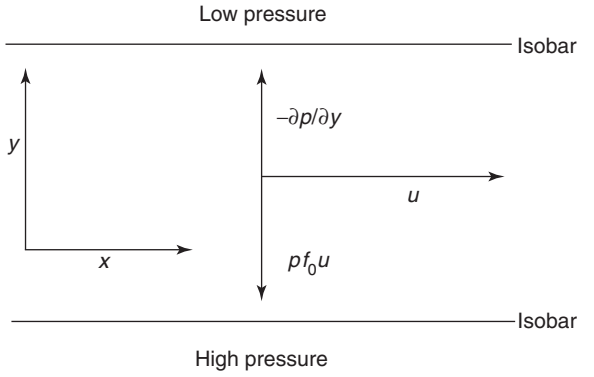


Figure 4 Geostrophic balanced flow; pressure gradient force (upper arrow), and Coriolis force (lower arrow) cancel out and the flow (horizontal arrow) is parallel to the isobars.

A system or flow that evolves slowly in time compared to $1/f_0$ can be described using the so-called "quasi-geostrophic" theory. This is a perturbative theory obtained as a first-order Ro filtering of the Navier–Stokes equation and analyzes departures from geostrophic balance. The quasi-geostrophic approximation is used in the analysis of extratropical systems, in which currents can be closely approximated by their geostrophic values. The use of the quasi-geostrophic approximation effectively filters out solutions corresponding to high-speed atmospheric gravity waves. This approximation is not accurate in situations in which non-geostrophic currents play an important advective role, for example, around the frontal discontinuities. Although modern computers allow a fully nongeostrophic description of most geophysical fluids, quasi-geostrophic approximation remains a fundamental tool of theoretical research and is used for practical applications in everyday life: every time we read an atmospheric isobaric map in terms of wind flowing along the isobars we are, in fact, using the notion of geostrophic balance.

Applications of Geophysical Fluid Dynamics: Weather and Climate

Given the nature and the impact on human society of their research, environmental scientists have had access to powerful computers since the early days of numerical computing. Actually, in the late 1940s, the first large-scale application of automatic computing consisted in the first numerical weather forecast, based on greatly simplified equations, which was proposed by Von Neumann and mainly devised by Charney. Since the late 1950s, the US technical services have been using computer-assisted numerical integration of relatively accurate equations descriptive of the physics of the atmosphere to routinely produce publicly available weather forecasts.

Numerical models of geophysical fluids are usually divided into limited area and global models, the fundamental difference being that the former require lateral boundary conditions (at the boundaries of the integration domain) that have to be continuously provided by either observations or, more typically, larger-scale models, while the latter essentially perform quasi-autonomous integrations – obviously, still there is external forcing. In geophysical fluids model, the three-dimensional fields are discretized on a lattice, and the temporal evolution is discretized through the adoption of a time-step. The adoption of a finite spatial and temporal resolution implies that the physical processes occurring on a space and/or timescale smaller than the actual resolution of the model can be taken care of only with approximate parametrizations relating such processes to coarser grain information. The parametrization of the so-called subgrid processes is usually heuristic and devised *ad hoc* by statistical interpretation of observed or model-generated data.

Numerical modeling options strongly rely on the available computer power, so that the continuous improvements in both software and hardware have permitted a large increase in the performances of the models and at the same time an impressive widening of their horizons. On one side, the adoption of finer and finer resolutions has allowed a more detailed description of the large-scale features of the dynamics, and, more critically, a more direct physical description of a larger set of processes, thus limiting the need for parametrization procedures, which, where needed, have become more accurate. On the other side, it has been possible to implement and then refine the coupling between models pertaining to different systems having a common boundary, such as the atmosphere and the ocean. The main purpose of such procedure is to study in detail the time–space scales of the interactions between the different systems and reframe such interactions in the more intrinsic terms of internal feedbacks of the compounded system.

Today, computer simulations probably constitute the best laboratory instruments for environmental scientists, since on one side permit the testing of theories, at least in the previously exposed weak sense, and on the other side provide stimulations for formulating new hypotheses. It is notable that the present most powerful computer has been engineered in order to simulate geophysical fluid dynamics with the particular scope of providing the most detailed simulation of the present and past climate.

When considering complex systems like those investigated by the environmental sciences, there is a clear distinction between providing efficient descriptions of the local and of the global properties in the phase space of the system. The local properties of the phase space are addressed by considering dynamic equations able to represent the short-time deterministic evolution of the system. The study of the global properties of the phase space entails considering dynamic equations whose goal is to represent the statistical properties of the system. The efficient description of such properties requires very long integrations of the equations. In order to clarify these very distinct – albeit intersecting – perspectives, the problems of weather forecast and of climate simulation are specifically discussed.

Weather Forecast

The task of weather forecast is to provide the best description of the short-time evolution of the relevant fields descriptive of the state of the atmosphere – wind, temperature, pressure, humidity, and precipitation – as represented on a 3D lattice. Weather forecast is addressed to describing and predicting the instantaneous values of the main atmospheric variables. Weather forecast models use in the equations quite a number of *ad hoc* assumptions and parametrizations that have empirically proved their reliability for that purpose. Since the aim of these models is to be as precise as possible in the local sense, they do not necessarily need to obey global constraints, for example, energy conservation, which are obviously fundamental in terms of physics, but not necessarily relevant when a small neighborhood of the phase space is considered.

The time range of such forecasts extends up to the deterministic predictability horizon of the atmosphere, which can be estimated in $\sim$7–10 days. It is important to emphasize that the actual spread of the model predictions can be estimated by using Monte Carlo techniques on the initial state of the system. Such an approach is termed "ensemble forecasting" and relies on sampling the phase space of a system by integrating forward in time a set of initial conditions. The initial conditions are suitably selected in order to represent the unavoidable uncertainty on the initial state of the system (**Figure 5**).

Climate Simulation

The climate consists of a set of statistical properties – in the time and/or space domains – of quantities that describe the structure and behavior of the various parts of the climatic system. The goal of climate modeling is to represent such statistical properties, thus aiming at a precise picture of the main properties of the phase space attractor of the system (**Figure 6**). In the case of past or present climate change studies, models try to assess how such statistical properties change with time. In the case of climate

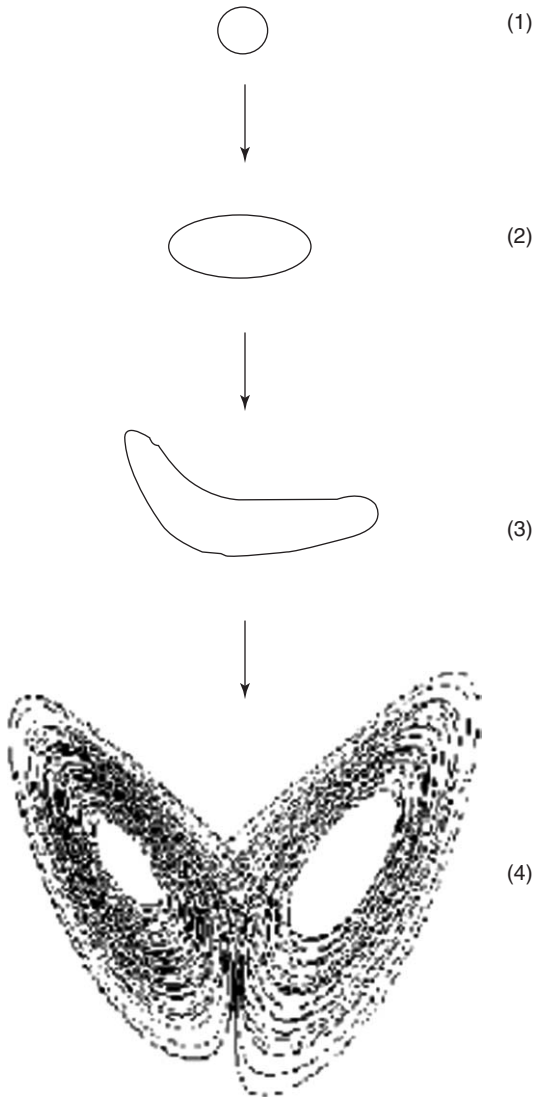


Figure 5 Scheme of the loss of predictability for ensemble forecasting. The initial Gaussian distribution of the ensemble (1) is distorted (2), and then loses its Gaussian shape (3), eventually the ensemble points populate the attractor of the system (4).

modeling, the relevant uncertainties are not related to the initial conditions of the system but rather to the very structure of the model. Since *ab initio* formulations based on first principles are not possible and since anyway the relatively coarse-grain discretization is being used (the best available models have average resolutions of 100 km in the horizontal and 200 m in the vertical directions), the presence of structural uncertainties, which are equivalent to uncertainties in the attractor's shape, is unavoidable.

A critical point in climate modeling is that, depending on the timescale of interest and on the problem under investigation, the relevant active degrees of freedom, which need the most careful modelization, change dramatically. For relatively short timescales (1–100 years) the atmospheric degrees of freedom are active while the other subsystems can be considered essentially frozen. For longer timescales (100–1000 years) the ocean dominates the dynamics of climate, while for even longer timescales (1000–10 000 years) the ice sheet changes are the most relevant factors of variability. Therefore, the very structure of climate models is gauged with respect to the purpose of the specific study and, depending on this, the models may be formulated in totally different ways.

Lately, a relevant strategy, which may be termed "statistical ensemble forecasting," has been proposed for dealing with the delicate issue of structural uncertainties, especially in the context of the study of the future projections of climate change, even if in principle this is not its only application. The main idea is to consider a climate model and apply Monte Carlo techniques not only on the initial conditions, but also on the value of some key uncertain parameters characterizing the climate machine. The resulting statistics (and change of the statistics) is expressed as a set of probability distributions rather than as a set of values.

Statistical Inference

Another prominent approach to the problem of providing an efficient description of the main properties of the environmental systems consists in trying to infer information directly from observations. Such an approach constitutes a natural and necessary complement to the theoretical and model-based analyses, which unavoidably suffer from the above-discussed uncertainties. A classical way of gathering information on the investigated system from the actual measurements consists in applying techniques that fall into the wide chapter of statistical inference. Environmental science is rich in such applications, some of which are quite well known to a wide public. In the last two decades, the outstanding example is that of global warming. In statistical terms, the problem consists in trying to detect signs of anthropogenic changes (expressed in terms of trends) of local as well as global measures of relevant climatic variables. In particular, the average global surface temperature is – somewhat anthropocentrically – usually considered as the variable descriptive of such change. The general problem is quite abundantly documented in the reports of the International Panel on Climate Change (IPCC).

What is the concrete appearance of surface temperature time-series? The presence of variability on a vast range of timescales, or, in other terms, of a rather composite spectrum, implies that suitable filtering

Figure 6 Evolution of the complexity of climate models in terms of representation of the main processes; the red lines indicate the integration of specific models into climate models. (Reproduced with permission from IPCC (2001) *Climate Change 2001: The Scientific Basis*, Contribution of Working Group I to the 3rd Assessment Report of the IPCC. Cambridge: Cambridge University Press.)

procedures are needed in order to detect changes with time of the signal. An example of how statistical inference can be applied to characterize the climatology of a system is discussed below.

Figure 7 shows the 1951–2000 daily maximum temperature records for two synthetic stations representative of northern and southern Italy. These data have been obtained after a suitable homogenization process of few dozens of real meteorological stations in each of the two parts of Italy. Apart from the signature of the seasonal cycle, the signal features an evident noisy structure. When the higher-frequency components are removed and the inter-annual components only are retained, one may observe an increasing trend for both records, which is consistent with the large-scale picture for the last 50 years. The observed warming trend can be proved to be statistically significant. But, instead of focusing on the long-term changes in the value of the signal, we are often interested in detecting changes in the properties of the seasonal cycle, defined in terms of phase and amplitude of the $(1\,\text{year})^{-1}$ component of the spectrum. Such a problem may be approached by providing a time-dependent estimate of the spectrum, obtained with a "shifting window" Fourier analysis, and then performing a statistical analysis of the various estimates of the locally defined (in time) seasonal signal.

In the case of the records shown in **Figure 7**, the amplitude and the phase of the signal, albeit characterized by a relatively large variability, do not have statistically significant trends. In **Figure 8**, the results for the amplitude of the signal are provided: the linear trend does not have a statistically well-defined sign for both records. Similar situation occurs for the phase of the signals. From these results, one can infer that the seasonal cycle has not undergone relevant long-term changes. Apart from the analysis of the variability of the signals under examination, one can infer from the average properties some basic climatic differences between the two systems, which in this case are defined in terms of geography only. First, quite trivially, northern Italy is colder, since its average temperature is notably lower (**Figure 7**). This is mostly due to latitudinal effects. Second, northern Italy has a more continental climate, since its seasonal

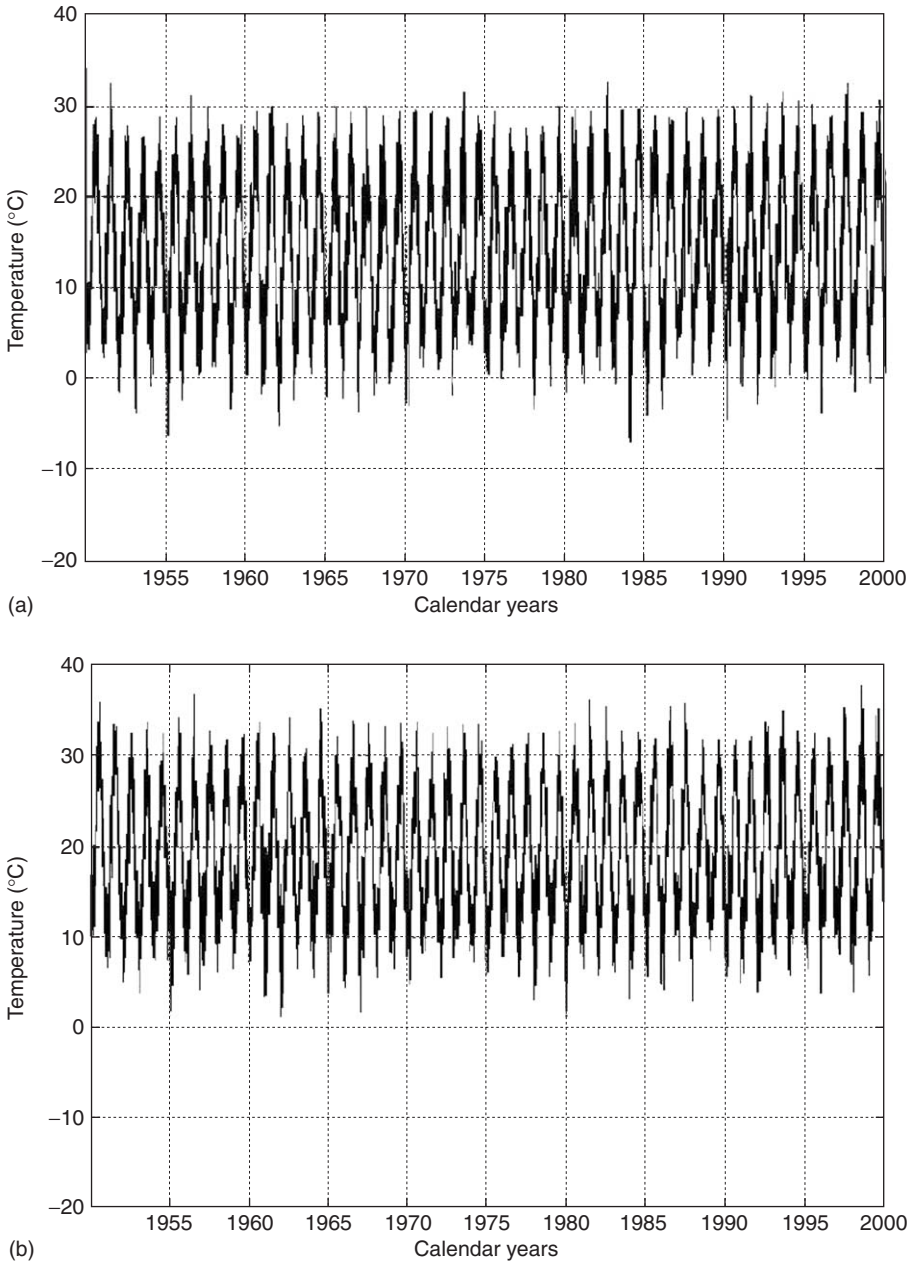


Figure 7 Daily records (1951–2000) of maximum temperature for northern (a) and southern (b) Italy. The signature of the seasonal cycle is evident.

cycle has larger average amplitude (**Figure 8**). As a first approximation, one can interpret this property as an effect of the smaller effectiveness of the air–sea thermal exchanges in smoothing out the effect of the annual changes in the intensity of the incoming solar radiation.

Diagnostics

Another relevant area of matching between quantitative and qualitative approaches to the study of environmental systems is that of diagnostics of the complex space–time structure of fields based on indicators derived from measure theory. It is well known that, for example, the definition of the structure of precipitation, land cover, etc., can be given in terms of estimators of the fractal measure derived from observations. In fact, the very concept of "fractal dimension" has been elaborated also in the effort of measuring territorial systems, with the classical example given by Mandelbrot for the evaluation of the length of the southern coast of England.

Consider standard precipitation, which is defined as the quantity of atmospheric water reaching the

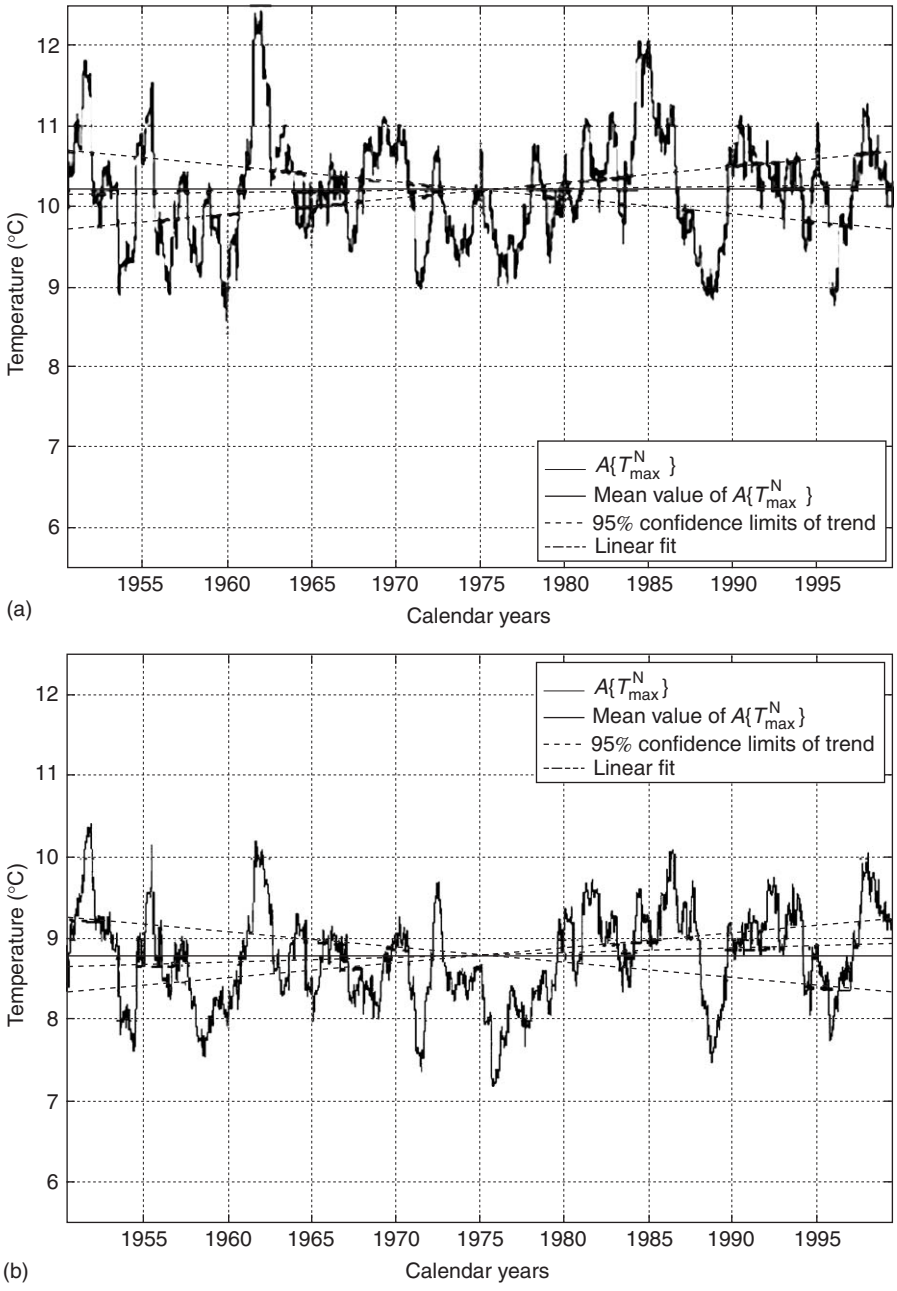


Figure 8 Amplitude of the seasonal signal $A\{T^{N/S}_{max}\}$ for the maximum temperature of northern (a) and southern (b) Italy. No statistically significant trend can be detected in either case.

Earth's surface in unit time. Such a quantity is the final result of an extremely long and complex chain of events, virtually covering all the space scales, from molecular water to the global scale of cyclonic areas, and the timescales from the fractions of second of enucleation to the lifetime of weather perturbations (days). Of course, other modulations exist at longer timescales (like the seasonal oscillation), but these are to be considered variation of external forcing rather than part of the precipitation process in itself. Therefore, any series of precipitation observations shows structure virtually at all scales. One way of measuring such structure is to make use of the so-called "box counting," which is very instructive since it makes reference to a basic definition of dimension in terms of the statistics of occupation of the considered parameter space. When investigating the statistical properties of homogeneous, stationary, and isotropic rainfall, a more appropriate definition of structure functions can be given through the moments of the integral measures of precipitation. Consider a positive random field $P(x)$, $P(x) > 0$, defined in

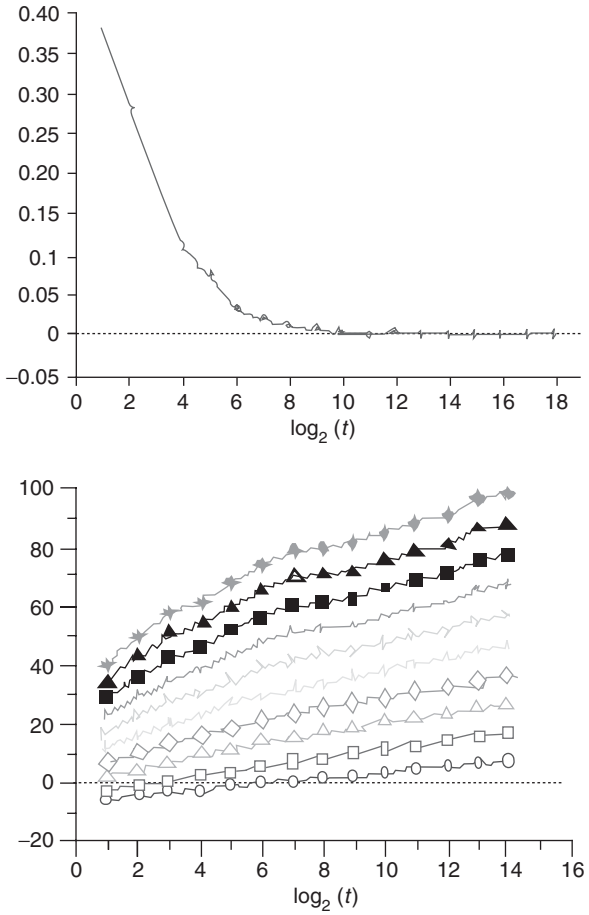


Figure 9 Time-structure of the precipitation at a rain gauge in the area of Rome (1 min interval of sampling from 1992 to 1998). Autocorrelation (top); and sequence of moments in increasing order from bottom up (bottom). The timescale in abscissa is in powers of 2 min. Ordinates are expressed in nondimensional units.

the set $x \in [0, L]$. Without loss of generality, $P(x)$ is normalized to 1, that is,

$$\int_0^L P(y)\, \mathrm{d}y = 1 \qquad [5]$$

for any realization. Our primary interest is to discuss the scaling properties of the structure function:

$$\int_x^{x+r} P(y)\, \mathrm{d}y \equiv \mu_x(r) \qquad [6]$$

$P(x)$ is said to display anomalous scaling properties if

$$\langle [\mu_x(r)]^q \rangle \sim r^{\xi(q)} \qquad [7]$$

where $\xi(q)$ is a nonlinear function of q. The scaling exponents $\xi(q)$ are also referred to as multifractal exponents of $P(x)$. The random field is assumed to be ergodic, so that in the previous equation $\langle\ \rangle$ stands both for x-average (x being either a space or time coordinate) and ensemble average. It follows that one must require also stationarity and homogeneity. Furthermore, if the field is embedded in a space of dimension $d > 1$, the field is also assumed to be isotropic, that is, the scaling properties depend only on the modulus of r.

In **Figure 9**, some statistical results for the precipitation in Rome are presented. It is learnt from correlation (the station of Rome is quite typical in this respect) that precipitation is characterized by a limited (with respect to other atmospheric variables like, for example, the surface temperature considered above) autocorrelation. The scale of decay of correlation is of the order of hours. A simple fractal structure would be represented in the logarithmic diagram of moments by a straight line. Apart from modulations due to sampling, it is seen that the higher moments show exponent dependence on scale and, therefore, indicate a multifractal distribution. The meaning of the inflection $\sim 2^8$ min (~ 4 h) is that organized dynamical processes (like the formation of rain bands) are operating at larger timescales, while the incoherent, micrometeorological processes dominate at smaller scales. This kind of analysis can be performed also in space (although this requires an extensive observational coverage that is rarely available) and produces similar results. Other statistical approaches are possible, in particular for what concerns the parametric distributions that are used in order to represent the above shown type of behavior. As seen earlier, such an analysis can identify statistical features revealing aspects of the structure and the dynamics of the systems in question. Another use is for operations of upscaling and downscaling the observed fields.

See also: Computer Simulation Techniques in Condensed Matter Physics; Nonlinear Dynamics and Nonlinear Dynamical Systems; Numerical Approximation and Analysis.

PACS: 02.50. – r; 02.60. – x; 02.70. – c; 05.45. – a; 47.27. – l; 47.32. – y; 92.10. – c; 92.40. – t; 92.60. – e; 92.70. – j

Further Reading

Holton JR (1992) *An Introduction to Dynamic Meteorology*. San Diego: Academic Press.

Lorenz EN (1967) *The Nature and Theory of the General Circulation of the Atmosphere*. Geneva: World Meteorological Organization.

Mandelbrot B (1977) *Fractal Geometry of Nature*. New York: WH Freeman.

Pedlosky J (1998) *Geophysical Fluid Dynamics*. New York: Springer.

Peixoto JP and Oort AH (1992) *Physics of Climate*. New York: American Institute of Physics.

Saltzmann B (2002) *Dynamic Paleoclimatology – Generalized Theory of Global Climate Change*. San Diego: Academic Press.

Epitaxy

L Miglio and A Sassella, Università degli Studi di Milano – Bicocca, Milan, Italy

Introduction

Epitaxy is the growth of a crystalline film suitably oriented onto a substrate: the first advantage in pursuing such a condition is to force the crystal in exposing free surfaces, which are not naturally obtained by cleavage or by conventional bulk growth. Moreover, the structural and compositional quality obtained by recent epitaxial growth techniques is better than the one obtained in the bulk: this quality can be controlled by very sensitive and precise methods, which are particularly suited for thin films. Easy and spatially resolved doping, along with monolithic integration on silicon wafers, have fostered this technique for microelectronic applications. Moreover, the stacking of different semiconductors in quantum wells and superlattices with sharp interfaces is a building block of modern optoelectronics. Finally, the strain control obtained by governing the lattice mismatch between the substrate and the film allows for intriguing modifications of the relevant physical properties in the latter, which in turn can be used to tailor device performances. Due to these exciting and recent applications, the focus here is on epitaxial semiconductors, both inorganic and organic, the latter ones being a rapidly growing field, which contains interesting basic issues and promising device developments. Epitaxial structures by metallic materials have been historically and conceptually important, particularly relevant for microstructural applications, for example, the elastic stiffening in superlattices produced by the long debated supermodulus effect, and for magnetic applications, more recently, in the field of spintronics. Still, for the sake of synthesis, a satisfactory review of this domain cannot be included, and the interested reader is referred to the "Further reading" section.

What is an Epitaxial Structure?

The concept of epitaxy was introduced by the mineralogist L Royer in 1928 to describe the growth of crystalline nuclei on various solid substrates. Since then, it has been widely extended to any case when some crystallographic relationship holds between a film and its substrate. A general definition of epitaxy comes from the Greek origin of the term itself, composed of the prefix 'επί (= on, over, above) and the word τάξις (= ordered arrangement), so that one can propose: EPITAXY – ordered growth of a crystalline or, at least, highly oriented material over another one, acting as substrate.

Epitaxy is therefore a fairly general concept, indicating rather different relationships between the substrate and the film, depending on their structural features and on the strength of their interactions, as is seen in the organic case. In particular, the surface unit vectors of the substrate, $a_{1\mathrm{sub}}$ and $a_{2\mathrm{sub}}$, are to be compared to the ones of the film, $a_{1\mathrm{film}}$ and $a_{2\mathrm{film}}$. When the lattice constants match, that is, $a_{\mathrm{sub}} = a_{\mathrm{film}}$, the interface is coherent; when a relation such as $ma_{\mathrm{sub}} = na_{\mathrm{film}}$ exists, with m and n integers (and the lengths ma_{sub} and na_{film} not too long on the lattice length scale), the interface is commensurate; otherwise it is incommensurate, although some weaker relationship, typically an orientational one, may hold, especially when the intralayer or molecular interactions are stronger than substrate-layer interactions. In the case of incommensurability, real epitaxial growth and perfect matching can still be obtained upon deformation of the film structure. Indeed, for sufficiently thin layers and for moderate deformations, well within the linear elastic regime (i.e., some percentage of the lattice parameter), the incommensurate film can be fully accommodated on the substrate. To this end, an important parameter is the geometric misfit f, expressing (usually in percentage) the amount of mismatch between the two structures, along one surface vector or the other one:

$$f = |a_{\mathrm{film}} - a_{\mathrm{sub}}|/a_{\mathrm{sub}} \qquad [1]$$

When a film is epitaxially grown on a substrate, crystallographic information on the former is usually expressed in terms of the ones of the latter following the notation introduced by E A Wood. An example can be α-Fe on GaAs (1 1 0), $1 \times 1/2$, where the indication of the substrate plane, (1 1 0), is followed by the expression $j \times k$, representing the number of unit vectors of the substrate plane, j along one direction and k along the other one, needed for matching the ones of the overlayer. In case the unit vectors of the film are not aligned with respect to the ones of the substrate, the rotation by XX° necessary to obtain alignment is indicated at the end by R XX°. Simple examples of (2×1) and $(\sqrt{2} \times 2\sqrt{2})$ R45° interfaces between substrate and epitaxial film are illustrated in **Figure 1**.

Homoepitaxy, Heteroepitaxy, van der Waals Epitaxy, and Quasi-Epitaxy

The increase of a solid by addition of new atoms or molecules of the same material on top of its surface is

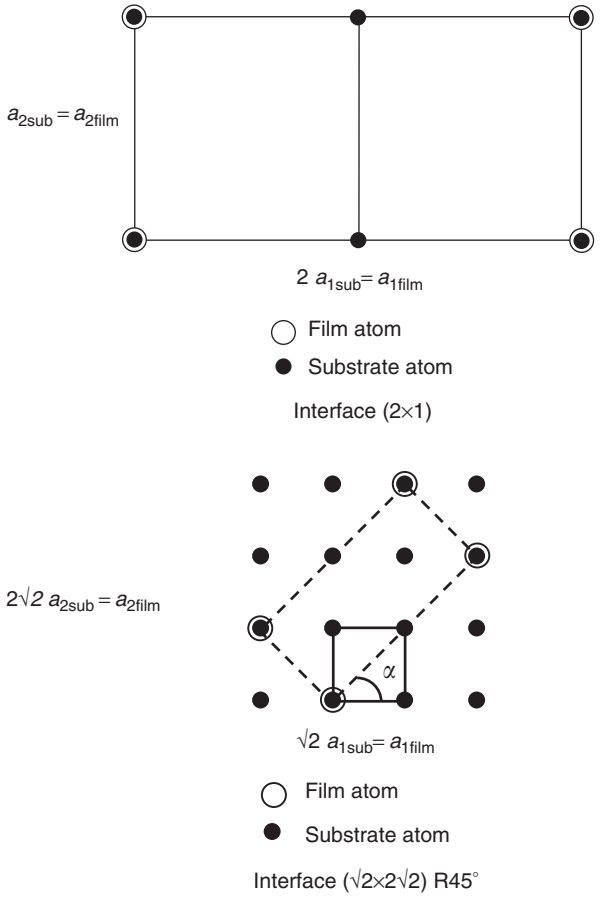


Figure 1 Examples of epitaxial relations between the film and the substrate: (2×1) and $(\sqrt{2} \times 2\sqrt{2})$ R45°.

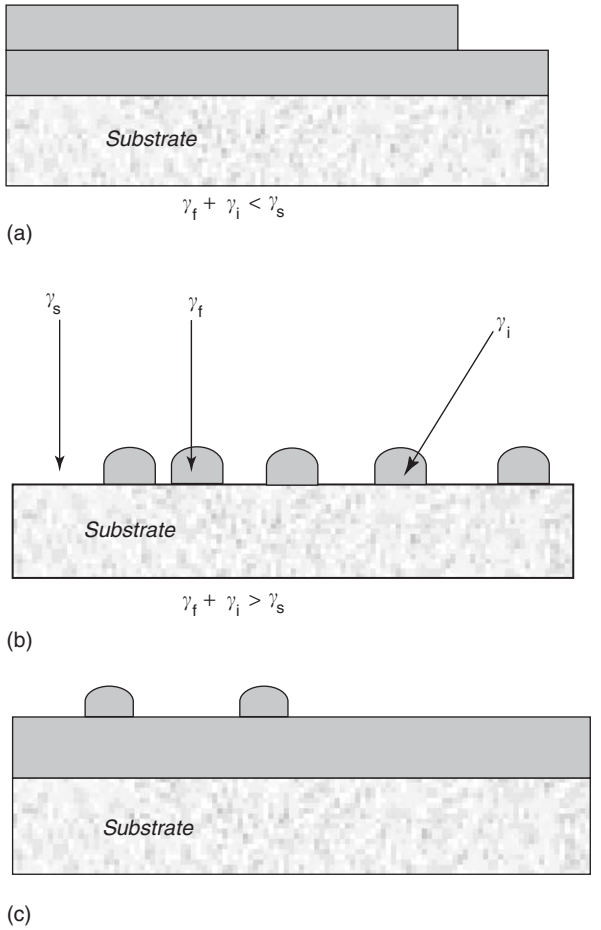


Figure 2 Different growth modes: (a) Frank–van der Merwe, (b) Volmer–Weber, and (c) Stranski–Krastanov; γ_f, γ_s, and γ_i are the free energies of the film surface, the substrate surface, and the film/substrate interface, respectively.

called "homoepitaxy" or "isoepitaxy" ('όμοιος = same, 'ι΄σος = equal). Although the process does not seem to need a particular attention, being in principle close to the usual growth of a bulk crystal, homoepitaxy is very important, since it is a fully controlled process, leading to structurally perfect and chemically clean surface layers. The most important example of homoepitaxy comes from silicon technology, where the (0 0 1)-oriented wafers are usually refined by an epitaxial Si deposition, whose quality is superior to that of the original substrate surface. In addition, such homoepitaxial layers can be properly doped for particular applications, independently of the substrate doping.

When the oriented or single-crystalline growth of a layer occurs on a substrate of a different material, "heteroepitaxy" is considered ('έτερος = different); in the case of a liquid substrate, one has rheotaxy. Heteroepitaxy is widely employed in semiconductor technology, where, through the choice of proper materials and alloys, artificial structures, such as quantum wells and superlattices, can be grown with precisely tuned properties. The first criterion proposed for a prediction of heteroepitaxial growth was the structural one, requiring a small misfit value. In the general case, however, this is not essential, since heteroepitaxial interfaces between materials with large misfit, even exceeding 15%, can also be grown and have actually been reported. In this respect, more complete theories identify the nucleation of crystalline islands, two- or three-dimensional (2D or 3D), as the essential process driving epitaxy, so that the balance between the surface and the volume contributions is critical, as usual. Heteroepitaxial films can indeed grow in the following three modes, schematically illustrated in **Figure 2**. The Frank–van der Merwe (FM), or layer-by-layer mode, which leads to smooth surfaces, is rather rare (classical examples are rare gases on graphite, incommensurate, and GaAs on Ge, forming a coherent interface). It is characterized by the fact that each monolayer starts growing by 2D islands, after the previous one is completed. FM growth is favored when the misfit is very low and when the surface free energy of the film is lower than

the sum of the one of the substrate and of the interface free energy (full wettability), $\gamma_s > \gamma_f + \gamma_i$. The Volmer–Weber (VW), or island mode is the opposite one, which leads to the formation of 3D islands nucleating over the substrate surface. In terms of surface free energy, VW mode occurs when $\gamma_s < \gamma_f + \gamma_i$, that is, when the surface energy of the film is larger than the ones of the substrate and interface, as it is frequently the case when the elastic energy produced by a sizeable misfit is included in γ_f. An intermediate situation is called Stranski–Krastanov (SK), or layer-plus-island mode, and it is characterized by the layer-by-layer growth of a few monolayers (the so-called wetting layer), followed by the nucleation of 3D crystalline islands on top of them. This process is the most frequent one: it is produced by an increasing unbalance with thickness between the tendency to wet the substrate surface and the accumulation of elastic energy by misfit in the film.

Another type of heteroepitaxy, named van der Waals (vdW) epitaxy, was recently suggested by the successful growth of organic thin films on inorganic substrates, occurring when the film and substrate interact only through van der Waals forces. The organic molecules are just physisorbed on the substrate surface, without forming strong chemical bonds, as in the case of inorganic films: the prototypical example is the growth of phtalocyanines on alkali halides. This type of epitaxy holds also when, strictly speaking, a huge misfit (above 100%) exists, since the lattice parameters for common organic crystals are easily larger than 10–20 Å. In this case, the lattice parameters of the organic material may be perfectly matched with a (very large) multiple of the ones of the inorganic substrate, so that commensurability and proper epitaxy is recovered. When the growth of an organic material on an organic substrate is concerned, a small misfit may be expected and heteroepitaxy occurs similarly as for the inorganic case, still at larger lattice parameters.

Finally, in the case of the weakest interaction, a sort of vdW epitaxy may hold between incommensurate structures, and the term quasi-epitaxy (or orientational epitaxy) is used. In such a situation, azimuthal order has to be attained at least, originating from a well-defined orientational relationship between the crystalline structure of the organic film and the substrate. Since the organic structures are usually fairly deformable, quasi-epitaxy may also lead to the growth of a strained organic film, which closely resembles a vdW epitaxial layer (indeed, quasi-epitaxy is also called strained vdW-epitaxy).

Organic films usually grow as crystalline 3D islands, whose size strongly depends on the substrate and growth conditions. Some cases of FM growth mode have also been reported, for example, pentacene on silicon. Among organic materials, polymers can also grow epitaxially, but this is often limited by the difficulty in setting proper crystallographic interactions between the substrate, either inorganic or organic, and the polymer film: quasi-epitaxy occurs in most of these cases. It is, nonetheless, a very interesting phenomenon, since it provides a tailored change of the physical properties: for example, it improves the adhesion of the two materials.

Analysis of the Epitaxial Film

The epitaxial growth process can be monitored *in situ* by reflection high energy electron diffraction (RHEED), which is sensitive to the roughness of the surface and, in turn, to the crystallographic completion of the surface layer. Another *in situ* technique which is able to identify the growth mode in heteroepitaxy is the Auger electron spectroscopy (AES) by a progressive monitoring of the peak intensity for the substrate and the film. For organic films, electron diffraction techniques can be destructive, so that they can be used in test samples, which permit the process tuning for further depositions. Optical monitoring, for example, by ellipsometry or by reflectance anisotropy spectroscopy (RAS) is usually to be preferred in such cases, even if the interpretation of the results is not straightforward and needs some modeling.

Assessing the quality of an epitaxial film, mostly *ex situ*, is an important branch of applied physics: as it happens for bulk materials, the analysis can be divided into structural (i.e., crystallographic) issues, microstructural (i.e., defect-related) features, and morphologic aspects of the surface, which are particularly relevant in a thin film. The first one deals with the crystal phase obtained during the growth, which may not be the one corresponding to the bulk phase (see next section). Moreover, the determination of the surface orientation and of the surface reconstruction, along with possible domain patterns, is a routine characterization of the film structure. This is performed by RHEED, together with other diffraction techniques, such as low-energy electron diffraction (LEED), transmission electron diffraction (TED), and X-ray diffraction (XRD), including recent developments, for example, Bragg rod scattering and coherent Bragg rod analysis, which are also suitable for checking the crystallographic relationships between the film and substrate, eventually determining whether the film is epitaxial or not. Finally, when high-resolution transmission electron microscopy (HR-TEM) is used, a complete characterization of the interface in cross section is obtained.

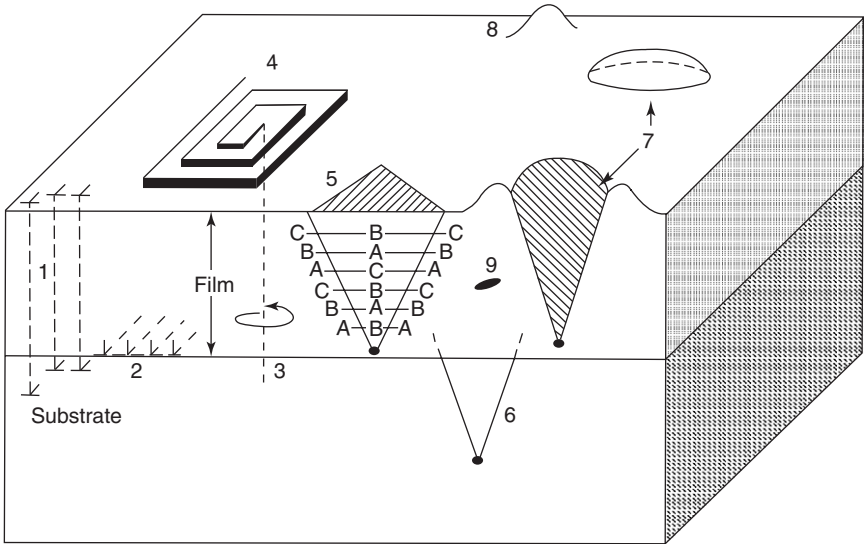


Figure 3 Schematic representation of crystal defects in epitaxial films: (1) threading edge dislocations, (2) interfacial misfit dislocation, (3) threading screw dislocation, (4) growth spiral, (5) stacking fault in film, (6) stacking fault in substrate, (7) oval defect, (8) hillock, (9) precipitate or void. (Reprinted with permission from Ohring M (1992) *The Materials Science of Thin Films*, 320pp. Academic Press.)

The microstructural analysis mostly considers the interior of the film and the interface with the substrate, which is likely to be a source of defects propagating in the film. They can be extrinsic point defects, including the intermixing at the interface in heteroepitaxy; intrinsic point defects, such as interstitials, vacancies, and clusters of them, which are quite frequent in the case of low-temperature growth; grain boundaries, as a final consequence of island ripening and coalescence; dislocations induced by the process of strain release in sufficiently thick layers (see next section), which are usually composed by misfit segments at the interface with the substrate and threading arms crossing the film up to the surface; twins, staking faults, and related "oval defects," which are particularly observed during the growth of compound semiconductors by molecular beam epitaxy (MBE). Classical TEM micrographs, along with indirect analysis of photo luminescence spectra, and other optical spectroscopies (recently, also micro-probe mapping techniques, such as micro-Raman) can be used to get a quantitative estimate of defect position and concentration. In **Figure 3** a schematic representation of crystal defects is reported, while the reader is referred to the "Further reading" section for a more detailed discussion. Here, it may be mentioned that controlling the density of such defects, in particular the threading dislocations and the stacking faults, is a major goal of present semiconductor epitaxy, since they critically degrade the performances of microelectronic devices by altering the carrier concentration and mobility. Acting as nonradiative recombination centers, they also reduce the minority carrier lifetime and the quantum efficiency of photonic devices.

The surface morphology is the last important feature to be considered: large-scale modulations on the micrometer scale or 3D islanding, from tens to hundreds of nanometers in width, generate a surface roughness, which is the by-product of the elastic strain release in competition/collaboration with the plastic process involved by dislocations in large misfit heteroepitaxy. Therefore, a morphologically flat surface is the first indication that strain relaxation has been obtained, and for a few decades this has been the primary goal of epitaxial growers. Nowadays, quantum dot formation by strain is a key issue of novel nanoelectronic and photonic devices, and strained layer governance is also required in present microelectronics, for example, in Si–Ge field effect transistors. Therefore, the full and quantitative characterization of the surface morphology by microscopy techniques (such as TEM, scanning tunnel microscopy (STM), and atomic force microscopy (AFM) (also used for evaluating the areal density of emerging threading dislocations after etching)) is presently an unavoidable feedback in epitaxial growth experiments. Some examples of epitaxial growth for inorganic and organic films are reported in **Figure 4**.

Heteroepitaxy and Strain Control in Semiconductors

Small Misfit Systems: Plastic Relaxation by Misfit Dislocations

This section considers systems of small misfits, typically of a few percent. In the following, the possibility of SK or VW growth is neglected and the

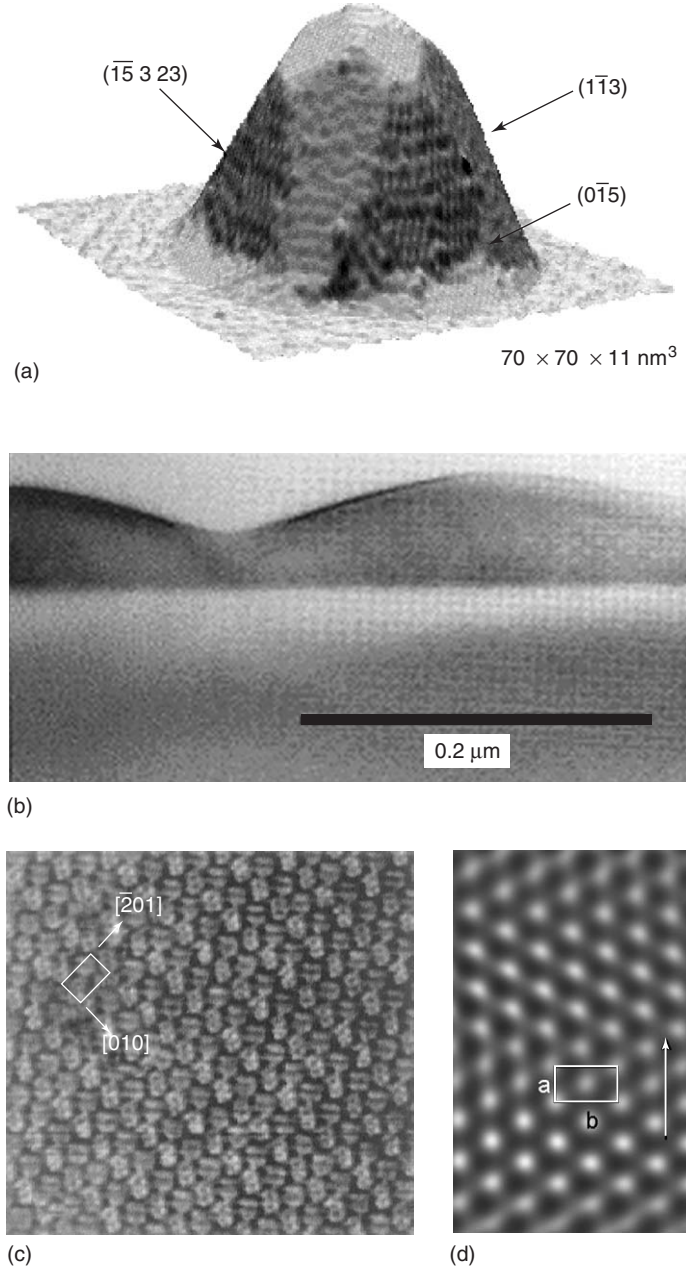


Figure 4 (a) High-resolution STM topographic 3D image of a Ge dome on Si(0 0 1), showing the facet orientations and the size of the dot. (Courtesy of A Rastelli, MPI-FKF Stuttgart.) (b) Cross-sectional [1 0 0] TEM image of an uncapped $Si_{0.81}Ge_{0.19}$ layer on Si, showing surface modulations. (From Cullis AG (1996) Strain-induced modulations in the surface morphology of heteroepitaxial layers. *MRS Bulletin* 21(4): 21–26 (Figure 4). Reproduced by permission of MRS Bulletin.) (c) 20 nm × 20 nm STM image of a substituted perylene (PTCDA) deposited on pyrolythic graphite by OMBD, showing the surface crystal structure, with two molecules per unit cell. (Reprinted with permission from Kendrick C *et al.* (1996) STM study of the organic semiconductor PTCDA on highly-oriented pyrolytic graphite. *Applied Surface Science* 104–105: 586.) (d) 5 nm × 5 nm filtered AFM image at molecular resolution of a 4 T film surface; the surface unit cell is evidenced together with the (0 1 0) crystal direction of the potassium phthalate substrate. (Courtesy of M Campione, University of Milano – Bicocca, Milano.)

surface of the epitaxial films is simply assumed to be flat. These are the conditions for which a sufficiently thin film has the same in-plane lattice parameter as the substrate (**Figure 5a**). Since the stress imposed by the substrate, $\sigma_{||}$, is 2D and the component perpendicular to the surface, $\sigma_{\perp}$, is zero, the epitaxial film is free to rearrange its lattice parameter perpendicular to the surface, in such a way that its internal energy is

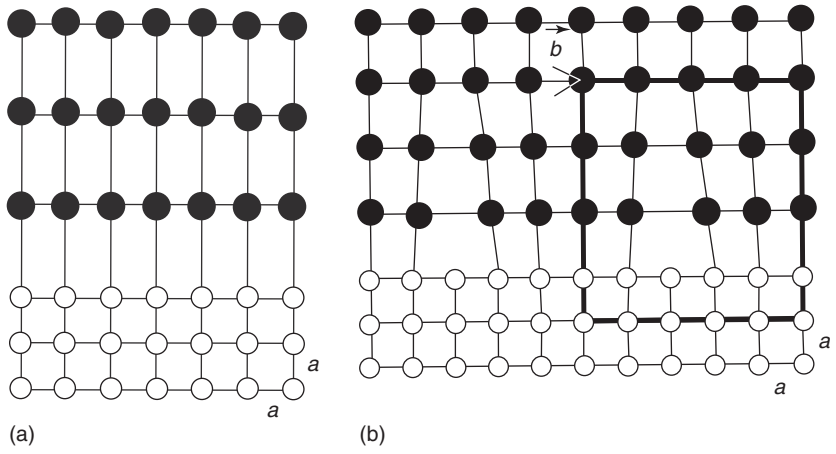


Figure 5 Sketch of (a) coherent and (b) incoherent interface; the film lattice parameters are indicated in relation to the one of the substrates; in (b) in the Burgers vector ***b*** is also defined. (Reproduced with permission from von Känel H, Onda N, and Miglio L (1995) *Science and Technology of Thin Films*, p. 31. Singapore: World Scientific.)

minimized. The resulting strain components are $\varepsilon_{||}$ and $\varepsilon_{\perp}$, parallel and perpendicular to the surface, respectively. They are opposite in sign, since the natural tendency of the strained structure is to preserve the volume of the unit cell. In fact, they can be expressed in terms of stress components, via the elastic equations involving the Poisson ratio v and the Shear modulus μ, which are in turn related to the bulk modulus B, that is, the stiffness of the material:

$$\begin{pmatrix} \varepsilon_{||} \\ \varepsilon_{\perp} \end{pmatrix} = \frac{1}{2\mu(1+v)} \begin{pmatrix} 1-v & -v \\ -2v & 1 \end{pmatrix} \begin{pmatrix} \sigma_{||} \\ \sigma_{\perp} \end{pmatrix} \quad [2]$$

Since $\sigma_{\perp} = 0$, $\varepsilon_{\perp}$ and $\sigma_{||}$ can be expressed in terms of $\varepsilon_{||}$:

$$\sigma_{||} = 2\mu\left(\frac{1+v}{1-v}\right)\varepsilon_{||} \quad [3]$$

$$\varepsilon_{\perp} = -\left(\frac{2v}{1-v}\right)\varepsilon_{||} \quad [4]$$

The elastic energy per unit area stored in a homogeneously strained film of thickness h is therefore given by

$$\begin{aligned} u_{\text{el}} &= \tfrac{1}{2}\, h(2\sigma_{||}\varepsilon_{||} + \sigma_{\perp}\varepsilon_{\perp}) \\ &= 2\mu\left(\frac{1+v}{1-v}\right)h\varepsilon_{||}^2 = 3B\left(\frac{1-2v}{1-v}\right)h\varepsilon_{||}^2 \end{aligned} \quad [5]$$

It may be noted that the elastic energy is a quadratic form of the strain $\varepsilon_{||}$, as in the case of a hydrostatic deformation, where $\varepsilon_{||} = \varepsilon_{\perp}$:

$$u_{\text{hyd}} = \frac{9}{2}\, Bh\varepsilon_{||}^2 \quad [6]$$

The energy in eqn [5] is obviously smaller, since relaxation in the perpendicular direction occurs, and it correctly vanishes in the limit case of volume preservation, when $v = 0.5$. Actually, the latter ranges, for most materials, between 0.25 and 0.35.

As long as the film is strained in such a way that its in-plane lattice constants match those of the substrate, the modulus of the parallel strain equals the misfit, that is, $|\varepsilon_{||}| = f$. As the thickness h increases, the elastic energy increases correspondingly, up to a limit h_c above which the uniformly strained state is no longer energetically favorable. The system can lower its total free energy instead, by generating so-called misfit dislocations (**Figure 5b**), in order to relieve part of the strain.

Actually, the calculation of the critical thickness h_c is not an easy task, since several contributions enter the energy balance. In the simplest approach, one assumes thermodynamic equilibrium and takes into account just two elastic energy contributions:

1. the energy due to the residual homogeneous strain $|\varepsilon_{||}| < f$, and
2. the energy stored in the inhomogeneous strain field of a square grid formed by two independent sets of parallel dislocation lines.

In order to evaluate the first contribution, one has to consider the misfit f_d taken up by an array of dislocations with linear density p_d, and with the component $b_{||}$ of the Burgers vector lying in the interface plane and perpendicular to the dislocation line. This somewhat complicated definition stems from the fact that in many relevant cases the Burgers vector is neither perpendicular to the dislocation line nor lying in

the interface plane, as in fact it is defined in **Figure 5b** for simplicity.

In the general case, the misfit taken up by a set of parallel dislocations spaced at a distance of $1/p_d$ is given by

$$f_d = p_d b_{||} \equiv p_d b \cos \lambda \qquad [7]$$

where λ is the angle between the Burgers vector and the direction in the interface plane, which is orthogonal to the dislocation line. The magnitude of the reduced strain is proportional to the modulus of the reduced misfit:

$$|\varepsilon_{||}| = |f - f_d| = |f - p_b b_{||}| \qquad [8]$$

and finally, eqn [5] provides the reduced elastic energy as a function of dislocation density:

$$u_{el}(p_d) = 2\mu\left(\frac{1+\nu}{1-\nu}\right)h(f - p_d b_{||})^2 \qquad [9]$$

The second contribution to elastic energy stems from the energy per unit length required for the bond breaking (at the core) and from the long-range strain field surrounding a misfit dislocation. The region of interest is a semicylindrical crust, extending from the dislocation core (conventional radius $= b/4$) at the interface to the free surface of the film. By a classical integration procedure, one finds

$$u_{dis} = \frac{\mu b^2}{4\pi}\left(\frac{1 - \nu \cos^2 \beta}{1 - \nu}\right)\ln\left(\frac{4h}{b}\right) \qquad [10]$$

where β is the angle between the Burgers vector and the dislocation line. This dislocation energy correctly displays a quadratic dependence on b, to which the related strains are proportional.

The energy per unit area of a square grid of independent dislocations is simply twice the linear density times the dislocation energy. The total elastic energy is therefore a sum of two contributions, that is, one quadratic (eqn [9]) and one linear in the dislocation density. Up to a critical thickness h_c, the minimum total energy is found at $p_d = 0$, that is, the film is homogeneously strained, and there are no dislocations. Beyond that limit p_d is finite, so that partial strain relief is obtained. An implicit expression for h_c at a given misfit f is obtained by setting to zero the first derivative of the total elastic energy, at zero dislocation density:

$$h_c = \frac{b}{8\pi f \cos \lambda}\left(\frac{1 - \nu \cos^2 \beta}{1 + \nu}\right)\ln\left(\frac{4h_c}{b}\right) \qquad [11]$$

Figure 6 shows the dependence of the critical thickness with the misfit, as predicted by eqn [11] for SiGe/Si(0 0 1).

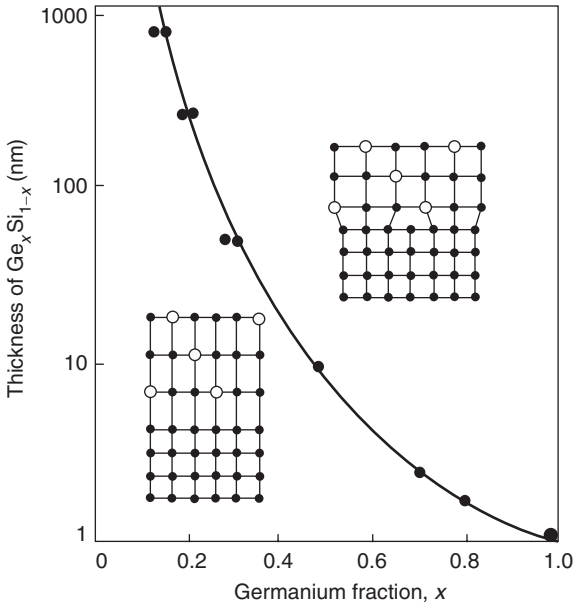


Figure 6 Experimental dependence of the critical thickness h_c with the misfit f for $Si_{1-x}Ge_x$ layers on Si(0 0 1), where the misfit is proportional to the Ge fraction. (Reproduced with permission from Bean JC (1986) The growth of novel silicon materials. *Physics Today* 39(10): 36; © American Institute of Physics.)

Intermediate Misfit Systems: Elastic Relaxation by Surface Corrugation and 3D Dots

If the misfit is some percents, consider the prototypical case of Ge on Si(0 0 1) with $f = 4.2\%$, elastic relaxation takes place before the critical thickness for dislocation nucleation is reached. In this case, $\varepsilon_{||}$ is lowered via surface roughening, which allows for local lateral expansion of the lattice parameter at the edges of 2D islands, or at the facets of 3D aggregates, including the case of surface modulations (see **Figures 4a** and **4b**). In particular, coherent clusters as large as tens or hundreds of nanometers appear in the SK or VW growth modes, with different shapes from rectangular-based pyramids (huts) to square-based pyramids, to larger domes with increasing aspect ratio (vertical size/lateral size). Indeed, as the aspect ratio increases the strain relaxation increases, especially close to the top, as reported in **Figure 7**, which is obtained by numerical simulations at the atomic scale. But for kinetic effects, which are important at growth temperatures below some 450°C, the different shapes are essentially produced by a balance between the elastic energy release in volume, depending on the aspect ratio, and the surface energy, which is related to the orientation of the facets and on their reconstruction. In Ge/Si(0 0 1), the (1 0 5) facets are particularly low in energy, and they appear both in the pyramids and in the surface modulations. As the size of the larger domes increases

with deposition, and the amount of residual strain with it, plastic relaxation by dislocations within the dots may take place, leading to very interesting structures where plastic and elastic relaxation phenomena coexist in a nanometric region.

Large Misfit Systems and Symmetry Mismatch: Growth of Epitaxially Stabilized Metastable Phases

As the misfit gets larger, say 8–10%, it becomes increasingly likely that, apart from the effects associated with the growth mode and with strain relaxation, other phenomena come into play. Actually, with a misfit in this range existing, it is no longer possible to strain even a single monolayer into coherency with the substrate. On the other hand, the symmetry of the film structure could be lower than the one of the substrate, with large differences in misfit depending on the orientations, so that a patchwork of domains is the best epitaxial result one can obtain. Still, it may exist in another phase, with a cohesion energy not much larger than that of the bulk phase and a symmetry closer to that of the substrate. By a suitable orientation, it may display equilibrium lattice parameters close to one of the substrates. In this case, the growth of a coherent film with such metastable (in the bulk) structure may be energetically favored, and it becomes an epitaxially stabilized phase. An appropriate term for such a phase would really be pseudomorphic, although this expression is misused in place of coherent in much of the literature.

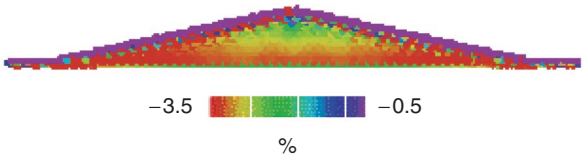


Figure 7 Compressive strain release inside a Ge pyramid on Si(0 0 1), as large as 27 nm; in a color scale (red = compression), the simulated strain component parallel to the surface is mapped, showing an expansion close to the top of the pyramid. (Courtesy of R Marchetti, University of Milano – Bicocca, Milano.)

The situation leading to the epitaxial stabilization of an unstable bulk phase is indicated schematically in **Figure 8**. The left panel shows the total energy curves of the two phases in question, under hydrostatic conditions and with the tetragonal relaxation included (see eqns [5] and [6]), as a function of the substrate lattice parameter, with the misfit of the bulk stable phase assumed to be larger. The important parameter, as far as phase stability is concerned, is the value of the energy at the substrate lattice constant a_{sub}. Evidently, the energy of the bulk stable phase strained into coherency with the substrate exceeds the energy of the coherent bulk unstable phase. This shows that the term epitaxial stabilization is actually a misnomer since it is rather the bulk stable phase that is destabilized under epitaxial conditions. It is clear that the pseudomorphic phase cannot remain stable to arbitrary film thicknesses. The energy of the stable bulk phase will start to decrease at the

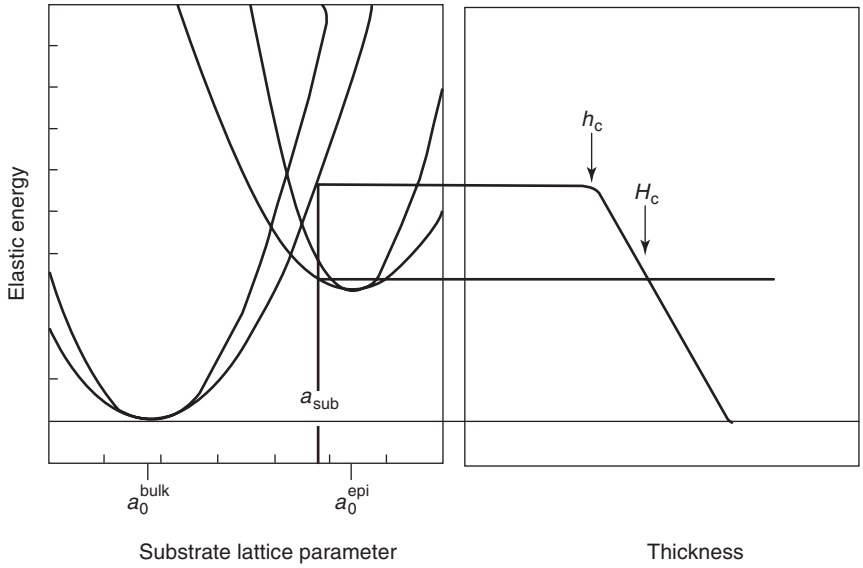


Figure 8 Left side: elastic energy of the two competing phases under hydrostatic stress (light curves) and after tetragonal relaxation (bold curves), as a function of the substrate lattice parameter. Right side: energy variation with film thickness; here, h_c is the critical thickness value at which strain relaxation sets in and the energy of the bulk phase decreases; H_c is the critical thickness value at which the bulk phase becomes favorite. (Reproduced with permission from von Känel H, Onda N, and Miglio L (1995) *Science and Technology of Thin Films*, p. 37. Singapore: World Scientific.)

film thickness h_c where plastic relaxation sets in, as is indicated schematically in the right panel of **Figure 8**. As the film thickness grows further, the two total energy curves eventually cross, giving rise to a new critical thickness H_c, at which the bulk stable phase becomes energetically favorable again. The epitaxial stabilization process is very interesting, since it virtually allows for the growth of new material phases with different properties with respect to the bulk ones. Still, it occurs just in case several polymorphic phases are possible, such as the case of metals or transition metal silicides, which have been studied very intensively in the last 20 years in relation to metallization and interconnects of integrated circuits.

Kinetic Issues

At small misfits, especially in semiconductor epitaxy, eqn [11] predicts a lower value for h_c than is often found experimentally. Films with thickness $h > h_c$ are thus frequently found in a metastable, strained state: in Si–Ge heterostructures, for instance, h can be 10 times h_c. However, by increasing the growth temperature (see **Figure 9**), or by subsequent annealing, h is reduced to h_c. Such a thermally activated behavior is usually governed by kinetics rather than by equilibrium thermodynamics. Indeed, the activation energies associated with the nucleation of new misfit dislocations at the interface and with the displacement or bending of the existing ones, in order to position the Burgers vector for the maximum strain relief, are not considered in the current thermodynamic model outlined above. Early attempts to include the kinetics of strain relief are due to Matthews (see "Further reading" section), but most of the work is still to be done. In the case of large misfits, where metastable phases may be stabilized by the substrate, the kinetics of the evolution with thickness is complicated by the fact that, in addition to the activation energies required to nucleate and move misfit dislocations, the activation energy for the structural transition from one phase into the other has to be taken into account. In particular, it has to be considered which kind of atomic motion is involved in the particular transition (martensitic, displacitive, massive, etc.), where the new phase nucleates, and how it grows. This is a formidable task, both from the experimental and the theoretical points of view. Finally, kinetic aspects are also very important when 3D islands provide an elastic release of the misfit strain. In fact, the growth in size and the morphological evolution to higher aspect ratios with deposition are surely directed toward an elastic energy lowering, but the way this is accomplished fully relies on the surface diffusion of the deposited atoms. Activation energies related to the diffusion of single atoms (or molecules) onto the wetting layer, along the dot facets, and across the multiple steps shaping the dot morphologies are key issues in understanding why rather different results are experimentally obtained for different deposition rates and temperatures. Probably, this is the most affordable, still complex, path in addressing the kinetic issues of the epitaxial growth, a subject still relying more on speculations than on predictions.

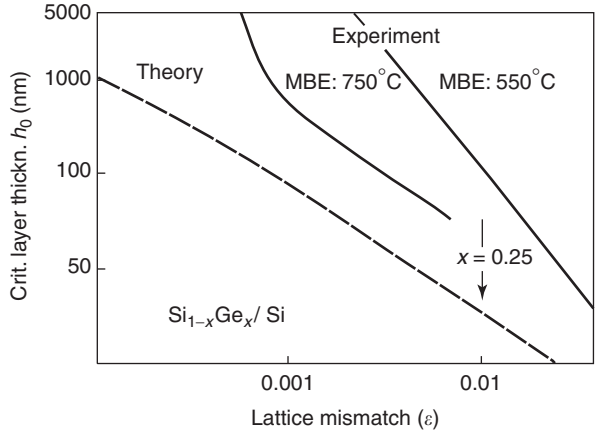


Figure 9 Critical thickness h_c vs. lattice misfit f for $Si_{1-x}Ge_x$ films on an Si substrate; the experimental data refer to samples deposited by MBE at different substrate temperatures. (Reproduced with permission from Kasper E (1987) Silicon germanium heterostructures on silicon substrates. *Festkoerperprobleme* 27: 265.)

See also: Crystal Growth, Bulk: Methods; Film Growth and Epitaxy: Methods; Integrated Circuits; Light Emitting Diodes; Mechanical Properties: Elastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Organic Semiconductors; Quantum Devices of Reduced Dimensionality; Scanning Probe Microscopy; Transistors; Transmission Electron Microscopy.

PACS: 61.72.–y; 68.35.–p; 68.43.–h; 68.55.–a; 81.15.–z; 61.10.–l; 61.14.–x; 68.37.–d; 68.49.–h; 68.60.–p; 85.30.–z; 85.35.–p; 85.60.–q; 85.65.–h

Further Reading

Bimberg D, Grundmann M, and Ledentsov NN (1999) *Quantum Dot Heterostructures*. New York: Wiley.

Forrest SR (1997) Ultrathin organic films grown by organic molecular beam deposition and related techniques. *Chemical Reviews* 97: 1793.

Hooks DE, Fritz T, and Ward MD (2001) Epitaxy and molecular organization on solid substrates. *Advanced Materials* 13: 227.

Koma A (1995) Molecular beam epitaxial growth of organic thin films. *Progress in Crystal Growth and Characterization* 30: 129.

Luth H (1995) *Surfaces and Interfaces of Solid Materials*, 94 p. Berlin: Springer.

Matthews JW (ed.) (1975) *Epitaxial Growth*. New York: Academic Press.

Schowalter LJ (ed.) (1996) Heteroepitaxy and strain. *Materials Research Society Bulletin* 21(4).
Stoneham AM and Jain SC (eds.) (1995) *GeSi Strained Layers and Their Applications*. Bristol: IOP Publishing.
Stringfellow GB (1982) Epitaxy. *Reports on Progress in Physics* 45: 469.
Venables JA (2000) *Introduction to Surface and Thin Film Processes*. Cambridge: Cambridge University Press.
Wittman JC and Lotz B (1990) Epitaxial crystallization of polymers on organic and polymeric substrates. *Progress in Polymer Science* 15: 909.

Nomenclature

a_{film}	lattice parameter of the film (m)
a_{sub}	lattice parameter of the substrate (m)
b	Burgers vector (m)
$b_{\|\|}$	component of Burgers vector lying in the interface plane (m)
$b_{\perp}$	component of Burgers vector perpendicular to the dislocation line (m)
B	bulk modulus ($N\,m^{-2}$)
f	misfit
f_d	misfit taken up by an array of dislocations
h	film thickness (m)
h_c, H_c	critical film thickness (m)
p_d	linear density of dislocations (m^{-1})
u_{el}	elastic energy per unit area stored in a strained film ($J\,m^{-2}$)
ν	Poisson's ratio
β	angle between the Burgers vector and the dislocation line
γ_f	surface free energy of the film ($J\,m^{-2}$)
γ_i	surface free energy of the interface ($J\,m^{-2}$)
γ_s	surface free energy of the substrate ($J\,m^{-2}$)
ε	film strain
$\varepsilon_{\|\|}$	component of strain parallel to the substrate
$\varepsilon_{\perp}$	component of strain perpendicular to the substrate
λ	angle between the Burgers vector and its component in the interface plane
μ	shear modulus ($N\,m^{-2}$)
σ	stress of the film ($N\,m^{-2}$)
$\sigma_{\|\|}$	component of stress parallel to the substrate ($N\,m^{-2}$)
$\sigma_{\perp}$	component of stress perpendicular to the substrate ($N\,m^{-2}$)

Exciton Complexes

B P Zakharchenya[†] **and S A Permogorov**, Ioffe Institute, St. Petersburg, Russia

Introduction

Exciton complexes are the electronic excitations of the crystal, the existence of which is caused by the presence of impurities or defects or by the interaction of more than two free carriers of different signs. In most cases, the exciton complex can be considered as an electron–hole pair bound to the impurity or defect. It does not move in the crystal and does not transfer the energy. As a result, the alternative description of exciton complexes as excited states of defects is possible. However, since the energies of such electronic excitations are normally close to that of free excitons, in most cases such complexes are designated as "bound" or "localized" excitons, as they are referred to further.

Exciton complexes were first detected in alkali–halide crystals, where the so-called α and β absorption bands downshifted in energy by some tenths of eV below the main exciton peaks were observed. In 1957, using the method of adiabatic cycles, F Bassani and N Inchauspe showed that the observed shifts are compatible with the model, assuming the exciton localization to be at the anion (α) or at the F-center (β). Later, the exciton complexes were observed and studied in many ionic, molecular, and semiconductor crystals.

The interest in the study of exciton complexes is based on its large role in photoluminescence. An optical excitation creates free excitons which migrate along the crystal and localize on defects forming the exciton complexes, which are very stable at low temperatures. As a result, most of the emission comes from the bound exciton states. The better the crystal quality, the longer is the exciton diffusion length, and higher is the intensity of bound exciton emission in relation to the free exciton emission. The particular lines of bound excitons can serve as fingerprints for the presence of certain defects and impurities.

Besides the localized exciton complexes, the mobile exciton complexes can be formed. An example of such complexes is the biexciton, which represents a bound pair of two Frenkel excitons in molecular crystals, or two holes bound with two electrons in semiconductors. Another kind of mobile complexes (trion) has been observed in recent times in semiconductor quantum wells (QWs). A trion consists of two electrons or holes bound with one carrier of the opposite sign (hole or electron).

[†]Deceased.

Exciton Complexes in Semiconductor Crystals

In good quality crystals, the absorption and emission lines of immobile exciton complexes are extremely sharp (halfwidth ~ 0.1 meV), and the strict resonance of absorption and emission exists at the energy, corresponding to the total energy of the exciton complex, even in semiconductors with indirect bandgaps. The momentum conservation for complexes is relaxed due to the presence of heavy impurities accommodating any excessive momentum. The high intensity of the bound exciton absorption is caused by the giant oscillator strength of the corresponding transitions. The oscillator strength of the bound exciton transition per one impurity f_b exceeds, by far (up to 10^4 times), that for the free exciton f_{exc} per unit cell. The reason is that a large crystal volume (of the order of, or exceeding a free exciton volume) is involved in this transition. As it was predicted by E Rashba in 1957, the ratio of oscillator strengths is $f_b/f_{exc} \approx (E_b/\Delta E_{exc})^{3/2}$, where E_b is the binding energy of the exciton complex, representing the energy difference between the free and bound exciton states, and ΔE_{exc} is the energy width of the free exciton band. This relation has been proved experimentally in the study of luminescence decay times in the range 10^{-9}–10^{-10} s for bound excitons with different binding energies in CdS crystals. The sharpness and a high oscillator strength of bound exciton transitions make them very suitable for studies of magnetooptic effects allowing one to determine the nature, symmetry, and the values of the carrier *g*-factors in exciton complexes.

Complexes of Excitons with Shallow Donor or Acceptor Impurities: Quantum Chemistry of Bound Excitons

The analogy of a free exciton in semiconductor crystals with a hydrogen atom has proved to be very instructive. In 1958, M A Lampert generalized this analogy to exciton complexes with shallow donors and acceptors comparing their electronic structure and binding energy with hydrogen molecules, molecular ions, or positronium aggregates. In such an approach, the binding energy of a bound exciton E_b can be expressed as $E_b \approx R_{ex} E_{mol}/R_H$, where R_{ex} is the exciton Rydberg in the crystal, E_{mol} is the binding energy of the corresponding molecular analog, and R_H is the Rydberg energy of the hydrogen atom (13.6 eV). The binding energies of hydrogen-like molecular analogs are well known from the molecular spectroscopy of gases.

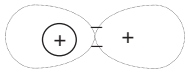
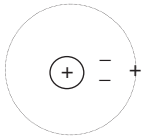
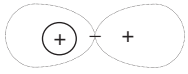
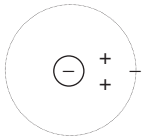
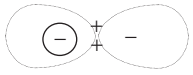
Donors and acceptors in semiconductors, due to the central cell corrections, are not exactly hydrogen-like. To account for this fact, the exciton Rydberg R_{ex} in the above relation should be substituted by $E_{A,D}$, the ionization energy of the particular acceptor or donor, so that $E_b \approx E_{A,D} E_{mol}/R_H$. It appears that the type of molecular analog and the magnitude of E_b strongly depend on the ratio of electron and hole effective masses m_e^*/m_h^*. The theoretical estimations of binding energies for some exciton complexes are presented in **Table 1**.

In the limit $m_h^*/m_e^* \gg 1$, the complex of an exciton with a neutral donor is analogous to the hydrogen molecule H_2. Two light electrons with antiparallel spins move on the orbit enveloping the positively charged core of the donor and the hole with the maximum of the electron density situated between two positive centers. The binding energy of the H_2 molecule is 4.6 eV and the Rydberg energy of the H atom is 13.6 eV, which leads to the value of 0.33 E_D for the binding energy of the complex. In the opposite case of $m_h^*/m_e^* \ll 1$, the character of the carrier orbital motion is quite different. Two relatively heavy electrons move around the positively charged donor core forming the center analogs to the negatively charged H^- ion with a net charge (-1). The motion of the hole around this center has a hydrogen-like character. Using the method of adiabatic cycles, W Kohn in 1960, showed that the energy necessary to remove the free exciton from such a complex approximately corresponds to the dissociation energy of H^- ion and can be represented as 0.055 E_D. The exciton bound to the charged donor in the limit $m_h^*/m_e^* \gg 1$ has an analogy with the molecular ion of hydrogen H_2^- (dissociation energy 2.6 eV), and its binding energy can be estimated as 0.19 E_D. The complex becomes unstable at $m_h^*/m_e^* < 1.4$. The above considerations can be applied also to the acceptor-based exciton complexes taking into account the reversal of the sign for all charges.

The realistic values of m_h^*/m_e^* ratio are 4–5 in II–VI and ~ 10 in III–V semiconductors. So, the real binding energies of exciton complexes should have intermediate values within the limits indicated in the table. The complexes both with the donors and acceptors can be expected for neutral centers, whereas for ionized centers only the donor-based complex is stable.

The emission spectra of excitons bound to neutral donors and acceptors were first observed by J R Haynes in 1960 in a study of low-temperature luminescence of high-quality Si intentionally doped by Sb, P, As, and Bi group-V donors, and by B, Al, Ga, and In group-III acceptors. Very sharp and intensive lines were observed and attributed to the emission of excitons bound to neutral impurities. The binding energies of complexes E_B were measured as energy shifts of emission lines from the exciton band bottom.

Table 1 Binding energies E_b of some types of exciton complexes in the units of donor or acceptor ionization energies E_D and E_A in dependence on the carrier effective mass ratio

Type of complex	*Effective mass ratio* (m_h^*/m_e^*)	*Model*[a]	*Molecular analog*	*Binding energy* (E_b)
Exciton bound to neutral donor	$\gg 1$		H_2	0.33 E_D
	$\ll 1$		$H^- + e^+$ See text	0.055 E_D[b]
Exciton bound to charged donor	$\gg 1$		H_2^+	0.19 E_D
	< 1.4			Unstable
Exciton bound to neutral acceptor	$\gg 1$		$H^- + e^+$ See text	0.055 E_A[b]
	$\ll 1$		H_2	0.33 E_A

In column 4, the molecular analogs of the complexes are indicated.

[a] In model drawings, the symbol " $-$ " represents the electron in conduction band with effective mass m_e^* and the symbol " $+$ " the hole in valence band with effective mass m_h^*. Symbols $\oplus$ and $\ominus$ represent the charged (ionized) donors and acceptors, respectively. e^+ stands for positron. Thin dotted lines show the character of carrier orbital motion in the complexes.

[b] Estimated by W Kohn in private communication to: Haynes JR (1960) *Physical Review Letters* 4(7): 361–363.

In agreement with theoretical predictions, it has been found that both for donor- and acceptor-bound complexes $E_B \approx 0.1\, E_i$, where E_i is the ionization energy of impurity. However, a slight but distinct difference between the two cases can be detected (**Figure 1**).

Similar regularity has been established for II–VI semiconductors also. However, for these compounds, in distinction to the case of Si, it has been found that $E_B \approx 0.2\, E_D$ and $E_B \approx 0.1\, E_A$ for complexes with neutral donors and acceptors, respectively. Besides that, complexes with charged donors were observed. In uniaxial II–VI compounds with a wurtzite structure, where the valence band is split by a crystal field into three components A, B, and C, the complexes of the $n = 1$ excitons of the two uppermost valence subbands A and B were detected. In CdSe crystals, the complexes of the excited state $n = 2$ of the A-exciton with the neutral donor were observed. The molecular analogs of such complexes are the excited states of the hydrogen molecule H_2^*. The fine structure of the spectrum and the binding energies (normalized on the binding energy of the complex with the ground exciton state $n = 1$) were in reasonable agreement with those for the three lowest stable states of H_2^* ($2s\sigma$, $2p\sigma$, and $2p\pi$). The identification of the types of exciton complexes in II–VI compounds was carried out mostly on the basis of magnetooptical measurements. The energies of bound excitons in different II–VI compounds are ranging from 3 to 7 meV for donor complexes, and from 6 to 60 meV for acceptor complexes.

Biexcitons in Semiconductors

A biexciton in semiconductors is a bound state of two electron–hole pairs. The hydrogen-like analog of a biexciton is a hydrogen molecule H_2. However, this analogy cannot be used for the estimation of the biexciton binding energy since the effective masses of the electron and hole are usually comparable in magnitude. More elaborate variational calculations predict that the binding energy of biexciton complexes should increase from $\approx 0.03\, R_{ex}$ to $\approx 0.15\, R_{ex}$ (R_{ex} is the exciton Rydberg), when the mass ratio m_h^*/m_e^* increases from 1 to 10. So, the expected values of the biexciton binding energy are rather small.

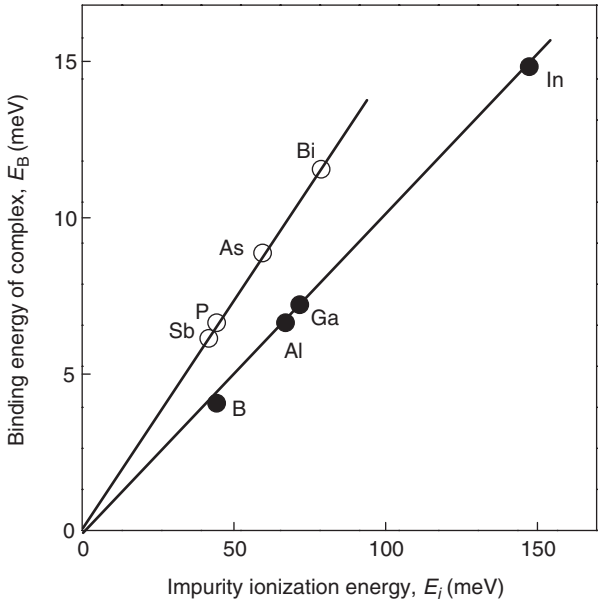


Figure 1 Dependence of the binding energy E_B on the ionization energy of impurity E_i for exciton complexes bound to Sb, P, As, and Bi neutral donors (open circles) and B, Al, Ga, and In neutral acceptors (filled circles) in high-quality Si crystals at $T = 2$ K. (Modified from Haynes J R (1960) *Physical Review Letters* 4(7): 361–363.)

Experimental observation of biexciton optical transitions in conventional one-photon spectroscopy is very difficult. The absorption process, corresponding to creation of the biexciton in a bound state has an energy approximately twice the transitions into the exciton states, and falls in the region of strong interband absorption. So, only the emission processes, when one of the excitons annihilates radiatively leaving the other in the ground or excited state, can be observed. The biexciton population is proportional to the square of the exciton concentration, and observation of biexcitons needs high-excitation density. In such conditions, the biexciton transitions are overlapped with some other emission processes, corresponding to elastic and inelastic exciton collisions which makes their identification difficult. Moreover, the increase of the excitation density leads to the formation of the electron–hole plasma, which screens both the biexcitons and free excitons. As a result, the biexciton emission can be observed only in a certain range of excitation intensities. Nevertheless, the observation of biexciton luminescence has been claimed for several direct-gap semiconductors (Cu_2O, CuCl, CuBr, ZnSe, HgI_2, etc.). In most cases, the difference between the measured and expected binding energies of biexcitons has been observed.

A reliable observation of biexciton luminescence has been made in Si crystals uniaxially stressed in different crystallographic directions. The effect of stress is to lift the degeneracy of both the valence and conduction bands, which increases the effective mass ratio m_h^*/m_e^* for the lowest exciton state. In conditions of maximum lifting of degeneracy (stress along $\langle 1\,0\,0 \rangle$ direction) at the increase of the excitation intensity, a new band appears in the luminescence spectrum. On the basis of the line position, its quadratic dependence on the excitation intensity, the temperature dependence of line intensity, and the theoretical analysis of the lineshape, a new band was assigned to the luminescence of biexcitons with a binding energy ≈ 2 meV.

The large progress in the detection of the biexciton states in direct-gap semiconductors has been achieved by means of two-photon nonlinear spectroscopy. Basically, three different experimental methods can be used.

1. Two-photon absorption (TPA), corresponding to direct excitation of biexciton states at simultaneous absorption of two photons. For observation of TPA, a narrow-band intensive laser pulse and a time-synchronized light source with a broad spectrum are normally used. The process of TPA is observed as an appearance of a new absorption line in a broadband spectrum when the sample is excited by both pulses simultaneously. The position of this line changes with the tuning of the narrow-band laser frequency so that the sum of the energy positions of the laser and the new absorption line is constant. This summary energy corresponds to the energy of the biexciton state.
2. If the exciton lifetime in the sample is sufficiently long, the stationary exciton population is formed at excitation with a properly tuned narrow-band laser. In such a case another new line can appear, corresponding to the optical transitions from the exciton into biexciton states. In comparison to the case of TPA, the position of this line does not change on tuning the laser frequency. This two-step process is called the induced biexciton absorption.
3. Resonant hyper-Raman (HR) scattering through the biexciton states. This process represents a version of the four-wave mixing (FWM) of polariton states in the crystal. Two photons of a narrow-band tunable laser with energies $\hbar\omega_l$ and the wave vector $\boldsymbol{q}_l$ and $\boldsymbol{q}_2$ (with some angle between them) excite a real or virtual biexciton state with energy $2\hbar\omega_l$ and wave vector $\boldsymbol{K} = \boldsymbol{q}_l + \boldsymbol{q}_2$. This state decays with the creation of two polaritons. One of them $E_1(q)$ is photon-like and the other $E_2(k)$ is the exciton-like polariton or the longitudinal exciton. The energy and momentum conservation laws require that $\boldsymbol{q} + \boldsymbol{k} = \boldsymbol{K}$ and $E_1 + E_2 = 2\hbar\omega_l$. As a result, the produced photon-like polaritons can

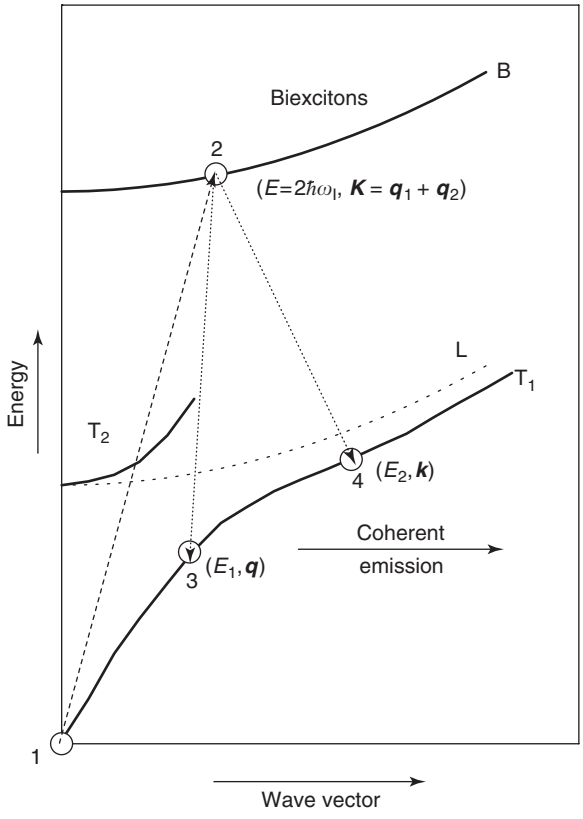


Figure 2 Schematic representation of hyper-Raman scattering in the region of biexciton resonance. Laser light produces a two-photon transition from crystal ground state 1 into biexciton state 2, which decays into photon-like polariton 3 and exciton-like polariton 4. Energy and wave vector for states 2, 3, and 4 are indicated in brackets. T_1 and T_2 are lower and upper polariton branches, L and B are the bands of longitudinal excitons and biexcitons. Only one of the possible processes is shown.

propagate only along some selected directions. These allowed directions and the energy of photon-like polaritons $E_1(q)$ depend on the detailed character of the polariton dispersion curves. At the crystal surface, the photon-like polaritons are converted into spatially resolved coherent light beams representing the output of HR scattering. The cross section of the HR scattering shows a giant enhancement at the exact resonance of the energy of two exciting photons with the real biexciton state. The process of HR scattering is presented schematically in **Figure 2**.

The biexciton states were studied in detail by means of two-photon spectroscopy in CuCl, CuBr, and CdS semiconductors.

Multiexciton Complexes in Indirect-Gap Semiconductors

A study of the low-temperature ($T = 2\,\mathrm{K}$) luminescence of Si crystals intentionally doped with shallow donor or acceptor impurities has shown that on increasing the excitation intensity, in addition to the lines of bound excitons (BE) reported by J R Haynes in 1970 (see above), a series of new narrow emission lines appears at lower energies. The series includes $\sim$10 lines and the distances between the lines have an order of 2–3 meV. The intensities of all lines grow with the increase of the excitation density: the larger is the line red-shift, the stronger is the intensity dependence. A similar emission has also been observed in doped Ge crystals.

The new type of emission has been attributed to complexes formed by the neutral donor or acceptor and several excitons. The possibility of formation of such complexes is connected with the complicated band structure of Si and Ge, and is caused by high degeneracy of both the valence and conduction bands in these semiconductors. With the account for the spin, the valence band is fourfold degenerate. The conduction band in Si has six valleys at Δ-points of the Brillouin zone. In Ge, there are four electron valleys at the L-points. Correspondingly, the degeneracy of conduction bands is 12 and 8 for Si and Ge, respectively, whereas in direct-gap semiconductors the conduction band degeneracy is only 2. As a result, in multivalley semiconductors several electrons can occupy the lowest 1*s* state.

A complete understanding of the electronic structure of multiexciton complexes has been achieved with the help of the "shell" model proposed in 1977 by G Kirczenow. An explicit consideration of the donor electronic spectrum, taking into account the local symmetry and valley-orbit splitting, has shown that it can be represented as a system of "shells" (3 for Si – Γ_1, Γ_3, and Γ_5; 2 for Ge – Γ_1 and Γ_5), which are gradually filled by electrons in accordance with the Pauli principle when the number of excitons in the complex is increased. This model has allowed one to get a full understanding of line positions, optical selection rules, and the number and polarization of split components in an external magnetic field and under uniaxial stress.

The impurity-bound multiexciton complexes were also observed in other indirect-gap materials (GaP, SiC, AlAs, AlSb). Presently, the possibility of observation of free multiexciton complexes (triexcitons, tetraexcitons, or even complexes of higher order) is discussed theoretically.

Trions in Semiconductor Heterostructures with Quantum Wells

Trions are complexes of an electron–hole pair (e–h) with one more electron (2e–h, X^-) or hole (e–2h, X^+). The stability of trions in semiconductors was discussed first by M Lampert in 1958. For realistic

effective mass ratio, very small binding energies were expected, much less than the 0.1 of exciton Rydberg. Experimentally, trions were not observed in bulk semiconductors.

The situation has changed drastically with the development of exciton spectroscopy of low-dimensional semiconductor heterostructures. The reduction of the dimensionality from 3D to 2D leads to an increase of the exciton binding energy up to four times, so that the excitons in some QW systems remain stable up to room temperatures. A similar effect of dimensionality is also expected in the case of trions. The first observation of stable trions was performed for CdTe/CdZnTe in 1993.

Afterwards, both negatively (X^-) and positively (X^+) charged trions were observed in many other systems with QWs (GaAs/GaAlAs, CdTe/CdMgTe, ZnSe/ZnMgSSe, ZnSe/ZnBeSe, and others).

Due to giant oscillator strength, the optical transitions to and from trion states were observed both in reflection and luminescence. The control of the charge sign of trion states is usually performed by providing excess carriers in the QW by δ-doping of the adjacent regions of the barrier layers either by donors or acceptors. It has been found that the binding energy of a trion has an order of several meV and strongly depends on the parameters of the QW, first of all on its width. The decrease of the width of the QW can increase the binding energy of the trion up to five times.

A trion is a charged particle, in which the spins of two carriers are compensated by one another. Such an electronic structure opens wide possibilities for electrooptical and magnetooptical studies of trion states. Also, the present state of QW engineering allows one to design many new experiments with trion states. As a result, the study of trion properties in different types of structures with QWs forms an interesting and a fast developing domain of semiconductor physics.

See also: Effective Masses; Electrons and Holes; Elemental Semiconductors, Electronic States of; Excitons in Crystals; Luminescence; Nanostructures, Electronic Structure of.

PACS: 71.35. – y; 71.35.Aa; 71.35.Cc; 71.35.; 71.35.Ji; 71.55.Fr; 78.20. – e; 78.20.Bh; 78.20.Ls; 78.40. – q; 78.40.Fy; 78.40.Me; 78.55. – m

Further Reading

Grun JB, Hönerlage B, and Levy R (1982) Biexcitons in CuCl and related systems. In: Rashba EI and Sturge MD (eds.) *Excitons*, pp. 459–504. In Series: Agranovich VM and Maradudin AA (eds.) Modern Problems in Condensed Matter Sciences, vol. 2. Amsterdam, New York: North-Holland.

Halsted RE (1967) Radiative recombination in the region of fundamental absorption edge. In: Aven M and Prener JS (eds.) *Physics and Chemistry of II–VI Compounds*. Amsterdam, New York: North-Holland.

Kulakovskii VD, Pikus GE, and Timofeev VB (1981) Multiexciton complexes in semiconductors. *Soviet Physics Uspekhi* 24(10): 815–840.

Ossau WJ and Suris R (eds.) (2003) *Optical Properties of 2D Systems with Interacting Electrons*, NATO Science Series II. Mathematics, Physics and Chemistry, vol. 119. Dordrecht, Boston, London: Kluwer Academic.

Rogachev AA (2001) Exciton molecules in multi-valley semiconductors. *Progress in Quantum Electronics* 25(4): 141–164.

Thewalt MLW (1982) Bound multiexciton complexes. In: Rashba EI and Sturge MD (eds.) *Excitons*, pp. 393–458. In Series: Agranovich VM and Maradudin AA (eds.) Modern Problems in Condensed Matter Sciences, vol. 2. Amsterdam, New York: North-Holland.

Timofeev VB (1982) Free many particle electron–hole complexes in an indirect gap semiconductor. In: Rashba EI and Sturge MD (eds.) *Excitons*, pp. 349–392. In Series: Agranovich VM and Maradudin AA (eds.) Modern Problems in Condensed Matter Sciences, vol. 2. Amsterdam, New York: North-Holland.

Excitons in Crystals

B P Zakharchenya[†] **and S A Permogorov**, Ioffe Institute, St. Petersburg, Russia

Introduction

Excitons represent a wide class of intrinsic electronic excitations in crystals of semiconductors and dielectrics. A characteristic feature of excitons is that their formation (e.g., at optical excitation) does not lead to the separation of carriers, so that the excitons are electrically neutral excitations. Originally, the concept of excitons was introduced by Ya I Frenkel in 1931 to explain the light absorption in crystals which does not lead to photoconductivity. In the Frenkel model, the exciton is considered as an electronic excitation of one crystal site with the energy close to, but a bit smaller than that necessary for the excitation of a free electron. Due to the translation symmetry of the crystal, the exciton can move along the lattice sites transferring the energy to the electrically

[†]Deceased.

active or luminescence centers. So, the excitons play a significant role in most of the photoconductivity and light emission processes.

The character of exciton motion depends on the strength of the exciton interaction with phonons. In the limits of a weak interaction, the exciton can be considered as a free particle with a well-defined wave vector and kinetic energy. In the opposite case, the exciton strongly polarizes the crystal lattice, and self-trapping of the exciton occurs. There is a variety of intermediate cases, when the exciton motion can be described either as a hopping process or by a strong scattering regime.

The Frenkel model is a good approximation for molecular crystals with van der Waals bonding, where the exciton is well localized within one molecule. For semiconductor crystals with covalent or slightly ionic bonding, a better description of excitons is given by another model proposed by G Wannier and N Mott. In this model, the exciton is considered as an electron–hole pair bound by a Coulomb interaction. In the case of weak electron–phonon interaction, both the electron and the hole can be considered as free particles which results in the hydrogen-like spectrum of exciton energies with a large number of excited states. However, in strongly ionic crystals, such as alkali halides, a self-trapping of holes at exciton formation takes place which leads to an essential deviation from the hydrogen-like model.

Excitons in Semiconductor Crystals

Energy Spectra of Excitons

In semiconductor crystals, the exciton can be considered as a hydrogen-like quasi-atom in which the electron and the hole are bound by Coulomb attraction. Since the distance between the electron and the hole in the exciton in most of the semiconductors is much larger than the lattice constant, the joint influence of all the crystal atoms on the exciton can be accounted for by using the band values of effective masses m_e and m_h, and the macroscopic value of the static dielectric constant ε_0. The exciton states, as in the case of hydrogen, can be characterized by quantum numbers n, l, and m. The binding energy of the exciton with the principal quantum number n is given by

$$R_n = \mu e^4/2\hbar^2\varepsilon_0^2 n^2$$

where e is the electron charge, μ is the exciton reduced mass $\mu = m_e m_h/(m_e + m_h)$, and ε_0 is the static dielectric constant. Since the exciton reduced mass is usually much smaller than a free electron mass ($\mu \cong 0.1 m_0$) and the typical value of the static dielectric constant ε_0, is ~ 10, the exciton binding energy is

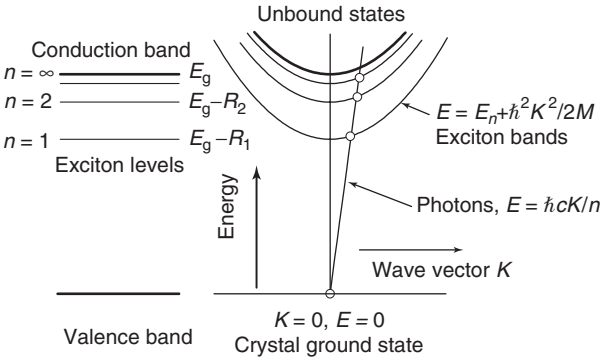


Figure 1 Exciton levels in the band scheme of semiconductor (left). Right part shows the exciton bands taking into account the exciton kinetic energy. Open circles indicate the crystal ground state and the states in the exciton bands to which the optical transitions occur.

$\sim 10^3$ times smaller than that of a hydrogen atom and ranges from 0.1–5 meV in III–V semiconductors to 10–30 meV in partly ionic II–VI semiconductors. Due to the possibility of a free motion, every exciton state can be characterized by a wave vector K and the kinetic energy $E_{kin} = \hbar^2K^2/2M$, where M is the exciton translation mass $M = m_e + m_h$. The full energy of the exciton state is given by $E_n = E_g - R_n + E_{kin}$ and the exciton energy spectrum represents a series of bands, the bottoms of which are downshifted below the fundamental gap E_g by the corresponding binding energies R_n, as it is shown in **Figure 1**.

Optical Transitions to Exciton States

At optical excitation below the fundamental gap E_g, the excitons can be directly created by light. The first experimental observation of an exciton optical absorption spectrum was performed by E F Gross in 1951 for Cu_2O crystals, where the hydrogen-like series with up to 6 lines was observed at $T = 77$ K. Further measurements at helium temperatures allowed one to resolve up to 11 members of this series. The important difference between the optical spectra of excitons and atoms is that in the case of the hydrogen atom, the photons induce the transitions from the ground state $n = 1$ to higher excited states $n = 2, 3, \ldots$, whereas for excitons the optical transitions take place from the "crystal ground state" ($E = 0$, $K = 0$) to the exciton states $n = 1, 2, 3, \ldots$. In pure crystals, in accordance with the energy and momentum conservation laws, the energy and momentum of the exciting photon has to be completely transferred to the newly produced exciton quasi-atom, so that $K_{ex} = K_{\hbar\omega}$ and $E_{ex} = \hbar c K_{ex}/n$. In the range of typical exciton energies (1–5 eV), the wave vectors of photons are much smaller than the limiting values at the edge of the Brillouin zone. As it can be seen from **Figure 1**, the conservation laws allow only

the creation of excitons with very small wave vectors $K_{ex} = K_{\hbar\omega} \approx 0$. The kinetic energies of such states are negligibly small $E_n(K_{\hbar\omega}) = E_n + \hbar^2 K_{\hbar\omega}^2/2M \approx E_n$. As a result, in the absorption spectrum of direct-gap materials appears a series of narrow (~ 1 meV) lines, with the positions corresponding to E_n ($K \approx 0$) energies. An example of such a spectrum for Cu_2O crystals is given in **Figure 2**. The oscillator strength for the transitions into the excited states decreases proportionally to $1/n^3$.

The exciton states can be formed also by conduction and valence bands with the extrema situated at different points of the Brillouin zone, like it is the case in GaP, Ge, or Si. The direct optical transitions into the exciton states in this case are forbidden. However, the optical transition into the exciton states is possible with simultaneous absorption or emission of the lattice phonon, which compensates the difference in K – positions of the band extrema. In the framework of the perturbation theory, such a process can be considered as a direct optical transition from the crystal ground state into the virtual state at $K \approx 0$ with subsequent scattering by phonons into the final state (**Figure 3**). Phonon-assisted transitions can excite the exciton states with any value of a wave vector (or kinetic energy). As a result, the indirect absorption leads to the appearance of the continuous "steps," the low-energy edges of which are shifted in position from the exciton band bottom by $\pm\hbar\omega_{MC}$, the energy of the momentum conserving phonon. The intensity of transitions with phonon absorption is proportional to phonon occupation number N and strongly depends on temperature, whereas the processes with phonon emission depend on temperature only as $(N+1)$. The indirect or phonon-assisted optical transitions also take place in the direct-gap semiconductors. However, the oscillator strength of such transitions is much smaller than that for direct transitions.

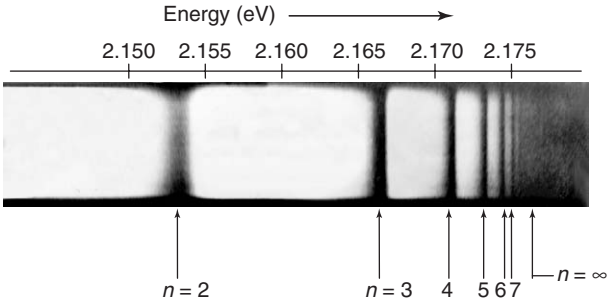


Figure 2 Part of the absorption spectrum of high-quality Cu_2O crystal at $T = 2$ K. The arrows with numbers indicate the optical transitions to the states with the corresponding principal numbers n. (BP Zakharchenya, unpublished data.)

Exciton Optical Spectra and Band Structure

Exciton optical spectra are closely related to the band structure of the crystal. Exciton transitions usually

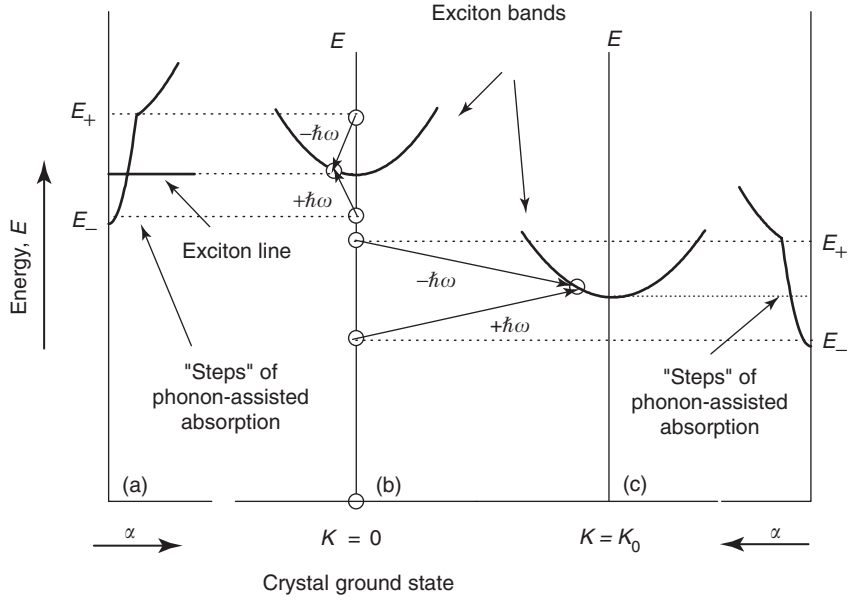


Figure 3 Schematic illustration of phonon-assisted exciton absorption in direct-gap and indirect-gap semiconductors (central part b). Open circles indicate crystal ground state, virtual intermediate states, and the final states in which the excitons are created. Arrows $-\hbar\omega$ and $+\hbar\omega$ show the scattering processes with the emission and absorption of the momentum conserving phonons, respectively. Left and right parts (a and c) depict schematically the spectra of phonon-assisted absorption. The absorption, corresponding to the processes shown in central part, starts at the energies E_- and E_+.

have high oscillator strength, narrow bandwidth, and strict polarization selection rules. This provides a unique tool for the study of the crystal band structure and determination of its parameters, such as the bandgap energies, the symmetry properties of the bands, and the carrier parameters. Each pair of valence/conduction bands produces a separate exciton series. In cubic semiconductors with zinc blende structure (GaAs, InP, ZnSe, ZnTe, CdTe, etc.), the valence band is split by a spin–orbital interaction into two components. The upper band, in turn, consists of two degenerated at $K = 0$ bands of light and heavy holes. As a result, two exciton series are expected in the exciton spectra. However, the series originating from the spin–orbit split valence band falls into the energy region of strong continuous interband absorption from the upper valence band and is strongly damped. In some zinc blende semiconductors (GaP, AlAs) and crystals of elemental semiconductors with a diamond-like structure (Ge, Si), the bandgap corresponds to transitions from the $K = 0$ states of the valence band to the conduction band minima situated at the edge of the Brillouin zone. In these cases, the transitions into the exciton states are possible only with the phonon participation, and have a form of continuous "steps."

The brightest example is given by the cubic Cu_2O crystals (cuprite structure), where the two hydrogen-like series (yellow and green) arising from the valence band splitting were observed. In the case of the yellow series, the wave functions of the carriers in the valence and conduction bands had the same parity. As a result, the optical transitions to the s states ($l = 0$) are forbidden in the dipole approximation, and the $n = 1s$ line is observed only as a weak quadrupole transition. The allowed (dipole) transitions starts $n = 2p$. The exciton lines up to $n = 11p$ were observed for the absorption spectrum of the "yellow" exciton series at $T = 4.2$ K. The transitions into the s and d states in Cu_2O were studied in detail using the two-photon absorption and excitation spectra of Raman scattering. In very thin Cu_2O platelets, two additional exciton series ("blue" and "violet") were observed. These series correspond to transitions from two upper valence subbands to the next conduction band with higher energy.

In uniaxial semiconductors with the wurtzite structure (CdS, CdSe, ZnS, ZnO, GaN), the valence band is split into three components, which give rise to three exciton series (A, B, and C) with different degrees of polarization of the dipole optical transitions with respect to the direction of the crystal axis c.

The precise determination of the bandgap energy is possible when at least two lines of the exciton series are detected. However, if even only one exciton line is observed, the analysis of the exciton spectra complimented by the independent information on the effective masses of carriers and the value of the dielectric constant allows a good estimate of the bandgap.

Fine Structure of Exciton Levels

The analogy of the exciton energy spectrum with that of the hydrogen atom is valid only in a first approximation. The important distinctions, leading to the essential and specific features of the exciton spectra, should be noted.

1. The screening of the Coulomb interaction between the electron and the hole cannot be described by the unique value of the dielectric constant. This screening arises due to the displacement of many atoms or ions in the crystal lattice, and depends on the character of internal motion of the electron and the hole, in particular, exciton state. As a result, the position of the $n = 1s$ exciton state, which is characterized by the fastest relative motion, is usually up-shifted with respect to the position predicted by the series convergence. Also, for the excited states, in distinction to the hydrogen atom, the degeneracy of s, p, and d levels in the exciton is usually lifted. In addition to the different character of the carrier motion, the splitting of s and p excited states arises also from the anisotropy of the hole effective masses, resulting from the complex character of the $E(K)$ dispersion in the valence bands.
2. The band states of electrons in the conduction band are characterized by the spin value $s = \pm 1/2$, and the states of the holes in the doubly degenerated valence band by the values of the angular momentum $J = \pm 1/2;\ \pm 3/2$. As a result, the exciton states with the magnetic quantum numbers $m = 0,\ \pm 1,\ \pm 2$ can be formed. As it follows from the conservation of the angular momentum, the one-photon (dipole) transitions are allowed only into the $m = \pm 1$ states ("bright" excitons). The energy levels of the exciton states with different m are split by exchange interaction. The magnitude of this splitting is usually small (0.1–1 meV for $n = 1s$), and decreases for the excited states in proportion to the oscillator strength. The exciton states with $m = 0,\ \pm 2$ ("dark" excitons), for which the dipole transitions are forbidden, nevertheless play a significant role in exciton relaxation and are very important for an understanding of the polarization properties and recombination kinetics of exciton emission.
3. For excitons, as for other intrinsic dipole excitations, one should distinguish the transverse states (T, with the exciton dipole momentum directed

perpendicular to the exciton wave vector K) and the longitudinal states (L, with the dipole momentum along the exciton propagation direction). The macroscopic electric field associated with the exciton oscillating polarization splits the L and T states by an energy proportional to the oscillator strength of the corresponding exciton transition. The magnitude of L–T splitting for $n = 1$ states in direct-gap semiconductors is ~ 0.5–2 meV. Due to the transversal character of the electromagnetic field of light, only the transverse excitons can be excited at the normal incidence of light on the crystal surface.

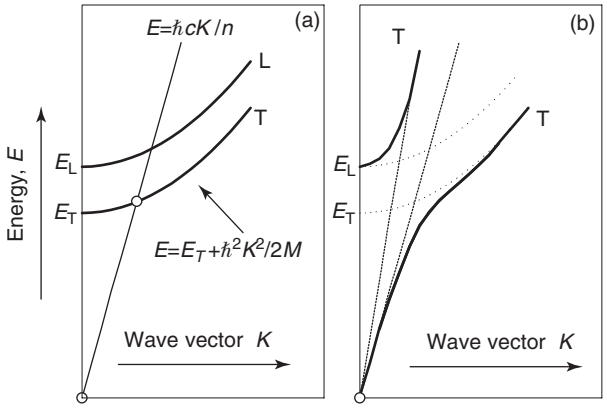


Figure 4 (a) Splitting of transversal T and longitudinal L excitons bands in mechanical approximation. Open circles indicate the unique exciton state which can be excited by light. (b) Dispersion curves of transverse polariton branches T_1 and T_2 arising due to the exciton–photon coupling.

The transversal waves of exciton polarization in the region of the exciton resonance are strongly mixed with the transversal waves of photon polarization, like it is a case for transversal optical phonons. As a result, in the resonance region, corresponding to the transition energy into the exciton state, the intrinsic excitations of the crystal can be no more considered as excitons or photons but represent a mixed excitation with a specific dispersion curve called the polaritons (**Figure 4**). The relative part of the exciton and photon properties smoothly changes along this dispersion curve, and the rather extended energy region (of the order of L–T splitting) is involved in the processes of resonant light absorption. In the region of the exciton resonance at the same energy, two waves transfer the energy inside the crystal. This situation was first designated by S I Pekar as an appearance of the additional waves. The oscillating structure, arising due to the interference of these two waves, was clearly detected in the transmission and reflection spectra of CdSe and CdS thin films (**Figure 5**). The dielectric constant in this region is a function not only of the energy, but also of the wave vector which represents a clear manifestation of the spatial dispersion. The account for the polariton nature of the exciton resonance is very important in the analysis of the

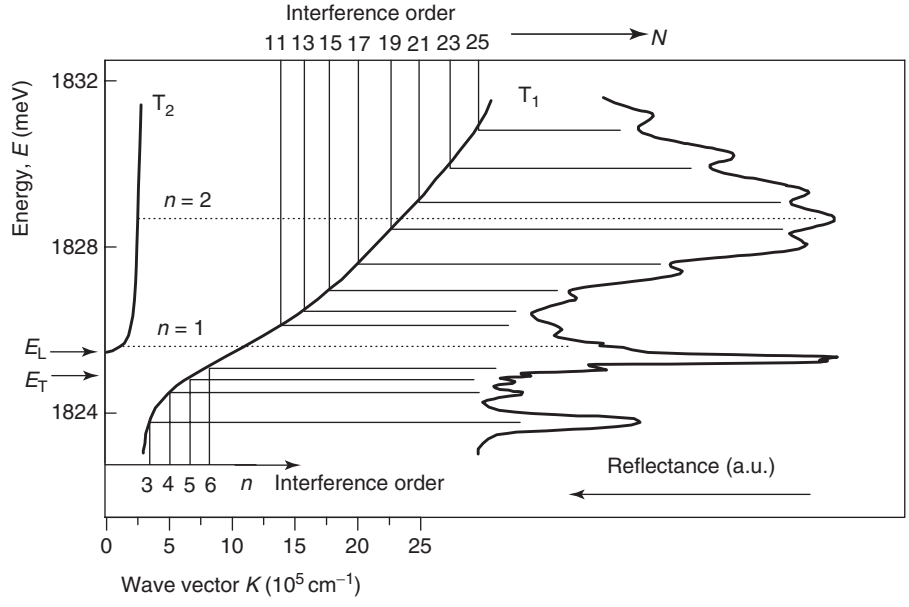


Figure 5 Interference pattern in the reflection spectrum in the region of $n=1s$ exciton resonance for thin CdSe platelet (thickness 0.24 μm, $T=4.2$ K) (right part). Left part shows the branches of polariton dispersion T_1 and T_2. The reflection minima with orders $n=3$–6 correspond to Fabri–Peroit interference of branch T_1 below the resonance. Dotted lines $n=1$ and $n=2$ indicate the minima of Fabri–Peroit interference for branch T_2 above the resonance. Reflection maxima with orders $N=11$–25 correspond to the interference between the branches T_1 and T_2. (Modified from Kiselev VA, Razbirin BS, Uraltsev IN (1973) *Pis'ma JETP* 18(7): 504 (in Russian).)

optical reflection spectra in the region of the exciton transitions.

Exciton Luminescence

At optical excitation below the fundamental gap E_g, the excitons can be created by light through the direct transitions to the E_n states. At optical excitation above E_g, the excitons can be formed through the Coulomb attraction of free electrons and holes generated in the process of interband absorption. The binding of free carriers into the excitons takes place after the carrier cooling on phonons and takes less than 10^{-9} s. However, in the process of cooling, an essential part of the free carriers can be trapped by donor or acceptor impurities.

The study of the excitation spectra of the exciton luminescence in partly ionic semiconductors has shown that a very efficient mechanism of exciton formation at the optical excitation above E_g is the creation of excitons in the process of phonon-assisted absorption with the simultaneous emission of LO phonons. This process takes part in resonance with the continuum exciton spectrum with many intermediate states available. As a result, the process of exciton formation through this mechanism is very fast ($\sim 10^{-12}$ s) and can involve emission of several LO phonons (up to 5–6), so that the spectral region for which it is important has an extent of ~ 0.2 meV.

The character of exciton luminescence in semiconductor crystals essentially depends on the purity of the crystal and the exciton lifetime. The translation motion of excitons in the crystal lattice provides an efficient transport of the energy of optical excitation to the impurities or defects where the excitons can be trapped or ionized. As a result, the luminescence spectra of semiconductors are usually dominated by the impurity-related emission, and the detectable luminescence of free excitons can be observed only in the most pure samples. For such crystals, the stationary population of excitons in the exciton bands can be formed. If the exciton lifetime is sufficiently long, the distribution of excitons on the kinetic energy will be in thermal equilibrium with the lattice. The excitons are quasiparticles with an integer spin and should obey the Bose statistics. However, for not very large exciton concentrations and not very low temperatures, the distribution of exciton kinetic energies can be described by the Boltzmann statistics. The resulting distribution can be written as

$$\frac{dN}{dE} \propto E_{kin}^{1/2} \exp\left(-\frac{E_{kin}}{kT}\right)$$

where E_{kin} is the exciton kinetic energy and T is the temperature. This distribution can be visualized through the process of phonon-assisted exciton annihilation which is a reciprocal one to the phonon-assisted exciton absorption. The exciton wave vector is transferred to the emitted phonons. As an example, **Figure 6** shows the luminescence spectrum of free excitons in thick ZnSe/GaAs epitaxial films, where the bands with the threshold energies down-shifted from the exciton resonance E_1 by the energy of 1LO and 2LO phonons are observed. At elevated temperature ($T = 30$ K), the shape of these bands demonstrates the Maxwell–Boltzmann distribution of excitons on the kinetic energy. The bands of phonon-assisted exciton annihilation were observed in many semiconductors, and they provide an excellent instrument for the monitoring of exciton stationary populations and the temperature of electronic subsystems.

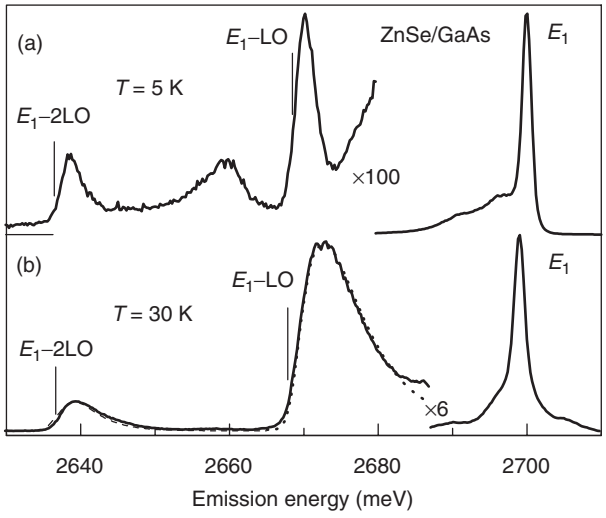


Figure 6 Exciton luminescence spectra of thick (4 μm) epitaxial ZnSe layer on GaAs at (a) $T = 5$ K and (b) $T = 30$ K. E_1 is the line of resonance luminescence of $n = 1s$ exciton state, bands $E_1 - \text{LO}$ and $E_1 - 2\text{LO}$ represent phonon-assisted exciton annihilation with simultaneous emission of one or two LO phonons. Dotted lines are theoretical fits of band shapes. (SA Permogorov, unpublished data.)

The theoretical description of the band shape and intensity of the zero-phonon line of exciton annihilation is more complicated due to the effects of resonant reabsorption. In this case, the model, simultaneously taking into account the exciton diffusion inside the crystal and the polariton structure of dispersion curves in the region of exciton resonance, should be used.

Exciton Spectra in External Fields

The apparent analogy between the exciton and atomic states has stimulated extensive studies of the influence of the external fields on the exciton spectra. The application of the uniaxial stress to the crystal

causes the splitting of exciton lines into several polarized components. The number and polarization of these components depends on the orientation of stress direction and the exciton wave vector with respect to the crystallographic axes, and allows one to obtain information on the symmetry of the exciton states, and hence, the bands responsible for exciton formation.

The study of the influence of an electric field on the exciton states is very difficult, both experimentally and theoretically. The experimental difficulties are connected with a strong heating of the sample by photocurrent arising at the sample illumination. Another difficulty is connected with the strong inhomogeneity of the electric field inside the sample caused by the formation of depletion regions. However, several experiments have been performed on low-conductivity Cu_2O samples and in compensated Ge.

The electric field causes a strong mixing of exciton states with different orbital numbers within one exciton series as well as between the states of different series. The main effect is the appearance of new lines forbidden without the electric field. A strong interaction of exciton states produces a complicated pattern of splitting and a nonlinear shift of exciton lines. In a strong electric field, the progressive broadening and disappearance of the exciton lines (starting from the higher members of the series) occur due to the exciton ionization (Franz–Keldysh effect).

The influence of a magnetic field on exciton spectra has been studied extensively. At low fields, a linear splitting of lines, very similar to the Zeeman effect in atoms, is observed. Additionally, a shift to the higher energies, quadratic in the magnetic field, (diamagnetic shift) takes place for all exciton lines. This purely relativistic effect is caused by the cyclotron acceleration of internal motion of the carriers in the exciton. In atoms, such a shift has been observed only for highly excited states in very strong fields. However, due to the large radii of exciton orbits, the diamagnetic shift in exciton spectra can be observed even at a few tesla.

A strong magnetic field modifies the internal structure of excitons. Due to the cyclotron acceleration of carrier motion, the exciton dimensions in the plane perpendicular to the magnetic field shrink and the binding energies of the exciton states increase. These modified states strongly elongated in the direction of the magnetic fields are called diamagnetic excitons. Further increase of magnetic fields leads to the formation of Landau subbands, both in the conduction and valence bands. The optical transitions between the Landau subbands lead to the appearance of oscillations in the spectrum of interband absorption above the fundamental edge. It has been found that these oscillations have a fine structure corresponding to the optical transitions in the states of diamagnetic excitons associated with every pair of the Landau subbands. In high magnetic fields, the shift of the lines of diamagnetic excitons follows that of the Landau subbands and is linear in the magnetic field.

Another analogy between the excitons and atoms is demonstrated by the possibility of exciton optical orientation at resonant pumping by circularly polarized light, or of the optical alignment of exciton dipole momenta at linearly polarized excitation. At resonant excitation, the resonant exciton emission demonstrates a very high degree of polarization of corresponding sense (up to 100%). A study of the dependence of this polarization on the strength and direction of the external magnetic field provides detailed information on the processes of energy and spin relaxation of excitons.

Excitons in Low-Dimensional Structures

In recent time, research efforts are concentrated mostly on the study of exciton states in heterostructures involving ultrathin "quantum wells" (QWs) formed by materials with bandgaps smaller than that of surrounding barriers. The confinement of carriers in QWs strongly influences both the internal motion of electrons and holes in the exciton and the translation motion of the exciton as a whole. The shrinkage of the exciton dimensions in the direction perpendicular to the plane of the QWs (z) increases the exciton binding energy up to four times, which is a theoretical limit for the ideal two-dimensional QWs. As a result, the exciton states can be stable and observable in the spectra up to room temperatures. This makes such heterostructures very attractive for practical applications in optoelectronics.

The exciton wave vector K_z is quantized in the z-direction, and the free translation exciton motion is possible only along the xy-plane, which essentially modifies the energy distribution of the density of states in the exciton bands.

The study of exciton spectra in QWs provides vast information on the properties of QWs, such as geometrical parameters, composition of forming material, the band offsets relative to the barrier, and the strain conditions of the QWs.

Excitons in Molecular Crystals

Energy Spectra and Band Properties of Excitons

In molecular crystals, the excitons can be very well described by the Frenkel model. Since the interaction between the atoms in the molecule is much stronger

than that between the atoms in different molecules, most of the crystal excitations are well localized within the single molecule and are closely connected with the molecular excitations in the gas phase. However, even relatively weak intermolecular interaction produces two very important features of the exciton. First, due to translation symmetry of the crystal lattice, the excitons can propagate in the crystal with a definite wave vector. The width of exciton bands is relatively small and ranges from 0.01 to 0.1 eV for dipole states. As a result, the exciton bands are strongly nonparabolic and the exciton velocities are rather small. Another consequence of intermolecular interaction is the splitting of nondegenerate levels into several components, the number of which is equal to or less than the number of molecules in the elementary cell of the crystal (so-called Davydov splitting). Optical transitions into the exciton states in molecular crystals, as in semiconductors, due to the momentum conservation are allowed only into the states with $K \approx 0$. The components of Davydov splitting are strictly polarized along the crystallographic axes.

Classical objects for the study of molecular excitons are crystals of benzene, naphthalene, anthracene, and other aromatic molecules, the lowest electronic excitations of which are situated in the visible or near-UV region of the spectrum. The crystals of aromatic molecules have a low symmetry (rhombic D_{2h}^{15} for benzene with four molecules in the unit cell, and monoclinic C_{2h}^{5} for naphthalene and anthracene with two molecules in the unit cell).

Three different kinds of excitons – electronic, vibrational, and vibronic – can exist in molecular crystals. Electronic excitons correspond to internal molecular electronic transitions into singlet and triplet excited states. The transitions into singlet states are allowed and produce strong absorption lines. The radiative times for such transitions have an order of several nanoseconds, whereas for triplet states the radiative times can reach tens of milliseconds.

Vibrational excitons originate from the high-frequency intramolecular vibrations and in many aspects are similar to the electronic excitons. They are characterized by narrow bands with a well-defined wave vector and are split on strictly polarized Davydov components. The intermolecular vibrations have much smaller frequencies and form a usual phonon spectrum of a crystal.

Since the electronic transition inside the molecule decreases the frequencies of intramolecular vibrations, the bound state of electronic and vibration excitations can be formed. Such new excitations can propagate in a crystal with a common wave vector, and are called vibronic excitons. However, in most cases the dissociated pairs of electronic and vibronic excitons are produced by optical excitation. This process is completely analogous to the phonon-assisted absorption in semiconductors and leads to the appearance of broad bands. Since the widths of some exciton bands (electronic or vibronic) are very small, the shape analysis of such bands allows one to restore the distribution of density of states in the exciton bands.

Exciton Luminescence in Molecular Crystals

As in semiconductor crystals, the luminescence spectra of molecular crystals strongly depend on the crystal purity and exciton lifetime. In crystals of moderate quality, most of the luminescence is associated with defects or impurities, since the excitons provide a good transport of energy to these centers. The luminescence of free excitons can be observed only in extremely pure and perfect crystals.

A very interesting phenomenon has been observed in the decay kinetics of free exciton luminescence in molecular crystals. The allowed singlet transitions have radiative lifetimes of several nanoseconds and demonstrate very fast decay at pulsed excitation. However, it has been found that the corresponding emission can last for tens of milliseconds. Such a phenomenon has been called a delayed fluorescence. It occurs if the long-living states of triplet excitons have lower energy than the singlet states. A fast relaxation of optical excitation to the triplet states creates a pool of stored energy. The radiative annihilation takes place at the collision of two triplet excitons with the formation of a singlet state. The delayed fluorescence in molecular crystals represents one of the most clear examples of cooperative radiation processes in solids.

See also: Crystal Field Splitting; Effective Masses; Electrodynamics: Continuous Media; Electrons and Holes; Exciton Complexes; Excitons: Theory; Organic Semiconductors; Polaritons; Polarons.

PACS: 71.35. – y; 71.35.Aa; 71.35.Cc; 71.35; 71.35.Ji; 78.20. – e; 78.20.Bh; 78.20.Ls; 78.40. – q; 78.40.Fy; 78.40.Me; 78.55. – m

Further Reading

Agranovich VL and Galanin MD (1982) In: Agranovich VM and Maradudin AA (eds.) *Electronic Excitation Energy Transfer in Condensed Matter*, Modern Problems in Condensed Matter Sciences, vol. 3. Amsterdam, New York: North-Holland.

Agranovich VM and Ginzburg VL (1984) In: Cardona M (ed.) *Crystal Optics with Spatial Dispersion and Exciton*, Springer

Series in Solid State Sciences, vol. 42. Berlin, Heidelberg, New York: Springer.
Bir GL and Pikus GE (1974) *Symmetry and Strain-Induced Effects in Semiconductors*. New York: Wiley.
Broude VL, Rashba EI, and Sheka EF (1985) *Spectroskopy of Molecular Excitons*, Springer Series in Chemical Physics, vol. 16. Berlin, Heidelberg, New York: Springer.
Davydov AS (1962) *Theory of Molecular Excitons*. New York: McGraw-Hill.
Ivchenko EL and Pikus GE (1997) In: Cardona M (ed.) *Superlattices and Other Heterostructures. Symmetry and Optical Phenomena*, Springer Series in Solid State Sciences, vol. 110. Berlin: Springer.
Knox RS (1963) *Theory of Excitons*. Rochester, New York: Academic Press.
Meier F and Zakharchenya BP (eds.) (1984) In: Agranovich VM and Maradudin AA (ed.) *Optical Orientation*, Modern Problems in Condensed Matter Sciences, vol. 8. Amsterdam, New York: North-Holland.
Pope M and Swenberg CE (1982) *Electronic Processes in Organic Crystals*. Oxford, New York: Oxford University Press.
Rashba EI and Sturge MD (eds.) (1982) *Excitons*. In Series: Agranovich VM and Maradudin AA (eds.) Modern Problems in Condensed Matter Sciences, vol. 2. Amsterdam, New York: North-Holland.
Ueta M, Kanzaki H, Kobayashi K, Toyazawa Y, and Hanamuraa E (1984) In: Cardona M (ed.) *Excitonic Processes in Solids*, Springer Series in Solid State Sciences, vol. 60. Berlin, Heidelberg, New York: Springer.

Excitons: Theory

P B Littlewood, University of Cambridge, Cambridge, UK

Introduction and Basic Ideas

The electronic band structure in a solid gives a formal accounting of the excitation energy $E_n(\boldsymbol{k})$ for adding to the ground state of the system a single carrier in the state of momentum $\boldsymbol{k}$. If the energy is greater than the chemical potential – that is, an empty state – this refers to an electron; if the energy is below the chemical potential – a filled state, the excitation is a hole. In a semiconductor or insulator there is a gap between the creation of electrons and holes. Of course, adding one electron and one hole in different bands or with different momenta creates an excited state of the system – and if the electrons and holes did not interact, one could of course enumerate the energies of the excited states just from the one-particle bands. However, the electron and hole are oppositely charged and therefore attract – the minimum energy required to create an electron–hole pair is therefore always less than to create either particle separately. The attraction between these two oppositely charged single-particle excitations creates a bound state called an exciton – just as the Coulomb attraction between a proton and an electron creates the bound states of the hydrogen atom (**Figure 1**).

Excitons recapitulate within a solid much of the physics of solids themselves, but with the added richness provided by the complexity of the bandstructure on which they are based. There are varieties of excitons, just as there are varieties of excited atomic states. Excitons themselves interact: there exist

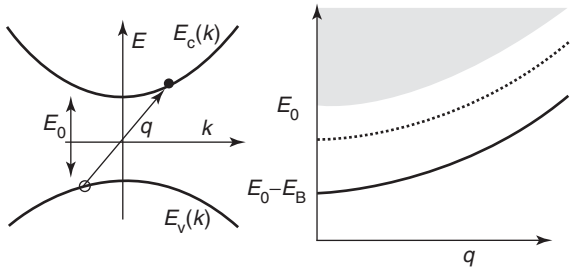


Figure 1 On the left is a sketch of the lowest conduction $E_c(k)$ and highest valence $E_v(k)$ bands for a simple direct gap semiconductor with a quadratic dispersion and masses for electron and hole of m_e, m_h. If these bands described noninteracting particles, the lowest excitation energy of total wave vector q is then $E_o + \hbar^2 q^2/2M$, where $M = m_e + m_h$, which marks the edge of a continuum of excitations of free electrons and free holes, shown by the grey hatched region on the right. Allowing for the binding of the electron and hole, the lowest excitations are now excitons – a series of discrete states that disperse approximately parallel to the bottom of the continuum, but shifted downward by the binding energy E_B.

excitonic molecules, excitonic liquids, and possibly even coherent condensed states akin to the Bose–Einstein condensation of cold atoms – though these have not yet been unequivocally observed. In some materials, excitons are heavy and easily localized, whereas in others they are very light and readily propagate throughout the solid. Excitons are composite bosons (formed from two fermions) and have spin degrees of freedom. They may trap on defects and disorder, and may also have strong interactions with the lattice leading to the formation of excitonic polarons. And most importantly, excitons often interact strongly with light, so that they are often the dominant feature in optical spectra of solids. Excitons may be observed in such a case by their absorption profile, but more frequently by their luminescence – the recombination emission of light

which is therefore, at the total energy of the electron–hole pair. Even if they do not decay radiatively, excitons are a transient phenomena – they have a finite lifetime, which can however be very long in some materials at low temperature.

Binding of Excitons

Frenkel Excitons

Conceptually, the simplest picture of an exciton emerges for crystals made of tightly bound electron states, such as rare gases or highly ionic solids. In these materials the electronic bands are narrow, and excitations within the solid are thus nearly localized. An atomic excited state, embedded within a lattice of unexcited atoms, is a cartoon for a Frenkel exciton – but of course since all the atoms in the solid are identical, this excitation hops from atom to atom throughout the solid to form a weakly dispersing band with a large mass. While the original concept of a Frenkel exciton was based on a transition occurring in a lattice of neutral atoms, highly localized excitons may also involve charge transfer between orbitals on nearby atoms or groups of atoms. The latter is often the dominant situation in polymers, and organic materials. Often the correct description is a mixture of Frenkel and charge transfer exciton, as one can see in **Figure 2** by the shift of absorption and the development of new structure between an isolated molecule and a crystal.

Generally, these highly localized excitons require detailed microscopic theory to analyze their spectral structure. Being localized, they couple strongly to vibrational degrees of freedom of the molecule, often producing extra "vibronic" structure.

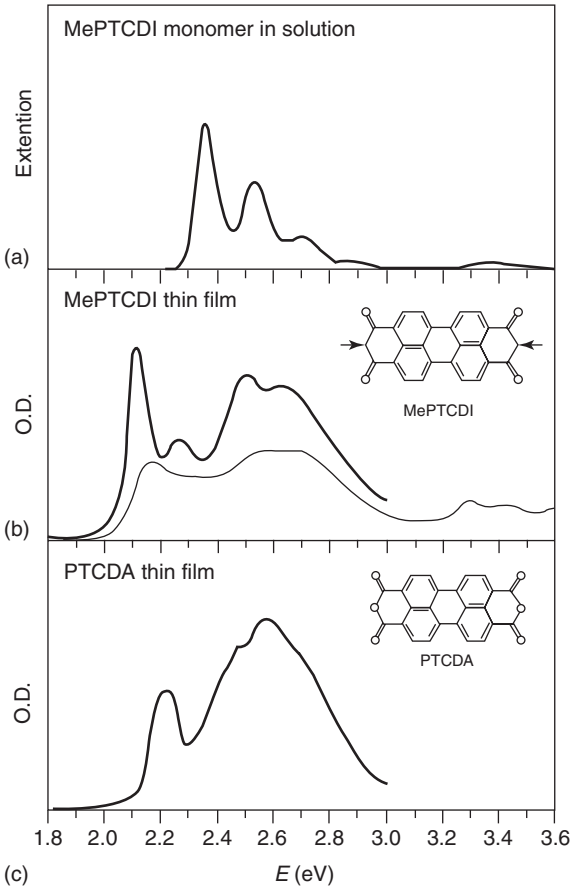


Figure 2 A comparison of absorption spectra of monomers with crystals where the excitons are highly localized. (a) The absorption of isolated molecules in solution, (b) The optical density of a polycrystalline film of the same material. The thick curve is at 5 K in *p*-polarization and the thin curve at 293 K in *s*-polarization. (c) A crystalline film of a similar compound at 5 K. (From Hoffmann M, Schmidt K, Fritz T, Hasche T, Agranovich VM, and Leo K (2000) *Chemical Physics* 258: 73.)

Hydrogenic Excitons

The opposing limit of weakly bound excitons is much more straightforward to analyze, and this produces the excitons familiar in direct gap semiconductors such as GaAs. Here the band dispersion is large, so that the energy cost of occupying states of large momentum $k \approx \pi/a$ (where a is the lattice constant) is large. Hence, the excitonic bound state is a wave packet made up from electron and hole states near $k \approx 0$. The valence and conduction band states are Bloch states, of the form

$$\psi_{nk}(\boldsymbol{r}) = u_{nk}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}k\cdot r} \qquad [1]$$

where u is a periodic function on the crystal lattice, and $n = c, v$ will refer to the conduction and valence bands respectively. If the momentum dependence of u_{nk} is neglected, the wave function for an exciton becomes separable,

$$\Psi(\boldsymbol{r}_\mathrm{e}, \boldsymbol{r}_\mathrm{h}) = u_{c\boldsymbol{k}=0}(\boldsymbol{r}_\mathrm{e}) \times u_{v\boldsymbol{k}=0}(\boldsymbol{r}_\mathrm{h}) \times f(\boldsymbol{r}_\mathrm{e} - \boldsymbol{r}_\mathrm{h}) \times \mathrm{e}^{\mathrm{i}q\cdot R} \qquad [2]$$

Here R is the center of mass coordinate, and q the corresponding momentum (see **Figure 1**). $f(\boldsymbol{r})$ is the exciton wave function in relative coordinates, and satisfies the Schrödinger equation

$$\left[\frac{-\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{4\pi\varepsilon\varepsilon_\mathrm{o} r}\right] f(\boldsymbol{r}) = (E - E_0) f(\boldsymbol{r}) \qquad [3]$$

where $\mu = m_\mathrm{e} m_\mathrm{h}/(m_\mathrm{e} + m_\mathrm{h})$ is the reduced mass, ε the relative dielectric constant, and E_0 the single-particle gap.

This Wannier exciton then has the spectrum of a hydrogen atom, with the difference being that, in

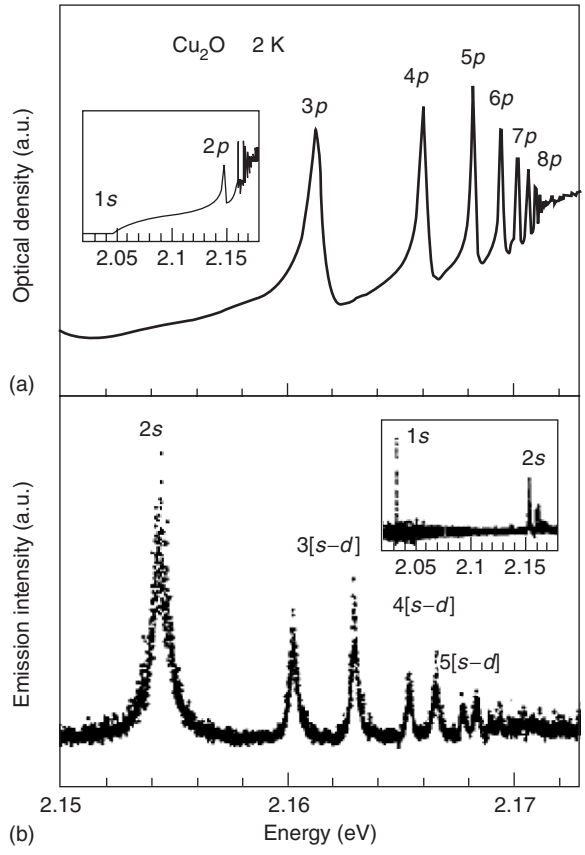


Figure 3 Exciton spectrum in Cu_2O. The upper panel is derived from one-photon absorption (and thus begins with the 2*p* state) and the lower panel shows the two-photon absorption, showing thus *s*- and *d*-series. The insets show the full spectrum. (From Matsumoto H, Saito K, Hasuo M, Kono S, and Nagasawa N (1996) *Solid State Communications* 97: 125.)

direct gap semiconductors such as GaAs, the effective mass is small ($\mu/m \approx 0.1$) and the dielectric constant is large ($\varepsilon \approx 10$). This means that the binding energy

$$E_B = \frac{\mu e^4}{8\varepsilon^2 \varepsilon_o^2 h^2} = \frac{\mu}{m}\frac{1}{\varepsilon^2}\,\text{Ryd} \qquad [4]$$

may be a small fraction of the Rydberg (Ryd = 13.6 eV). The size of the exciton is given by an effective Bohr radius

$$a^* = \frac{\varepsilon \varepsilon_o h^2}{\pi \mu e^2} = \varepsilon \frac{m}{\mu} a_o \qquad [5]$$

much larger than the hydrogen Bohr radius $a_o =$ 0.05 nm. Thus, the wave function extends over several unit cells, which justifies the approximate expression made for the wave function in eqn [2] (**Figure 3**).

Indirect Excitons

It is not necessary for exciton states to correspond to vertical band transitions (i.e., $q = 0$ in **Figure 1**) because a wave packet can be constructed about any band extremum. So, for example, in Ge the lowest excited state is made from wave packets of holes at $k = 0$ and electrons at the equivalent conduction band minima along the [1 1 1] direction at the Brillouin zone boundary. A similar situation pertains for Si, except that the electrons are confined to six pockets in the [1 0 0] directions (inside the Brillouin zone). These excitons are called "indirect" because they require a nonzero momentum to connect the electron and hole. Momentum conservation then does not allow the exciton to directly decay into a photon, and one requires either a process involving recombination also with emission of a phonon, or scattering from an impurity, a surface, or other confinements. This is typically very inefficient, which is why Si is not used for optical devices.

Confined Excitons: Quantum Wells, Wires, and Dots

The engineering of quantum wells in semiconductors can also be used to manipulate excitonic states. In $GaAs/Ga_{1-x}Al_xAs$ quantum wells (the canonical system, though many of the III–V and II–VI series are used), one may confine both electron and hole to a narrow sheet, comparable or smaller in size than the effective Bohr radius a^*. It is also possible to create more complicated structures including quantum wires and quantum dots. There are two principal effects. First, the extra confinement brings the electron and hole closer together and increases (sometimes substantially) the binding energy producing a shift in the energy of absorption and luminescence in comparison to bulk material. The second effect arises because of the structure of the top of the valence band which in a bulk system has a degeneracy between the light and heavy hole bands at $k = 0$ (the states are *p*-like, which therefore have a threefold orbital degeneracy under cubic symmetry; with spin–orbit coupling the pair of bands are the fourfold-degenerate $J = 3/2$). This degeneracy is broken by the quantum well, which fixes a local axis, and splits the light and heavy hole bands, leading to a further separation of the two types of excitons.

Since an applied field will pull the electron and hole apart from each other the exciton energy and its binding are affected by applied electric fields. In a bulk system quite small fields can ionize the exciton, but if the exciton is confined either to a well or a quantum dot (see **Figure 4**) the electron and hole will be trapped and very large shifts of the exciton luminescence with electric fields can be produced. This is the "quantum-confined Stark effect," a principle used in semiconductor laser modulators.

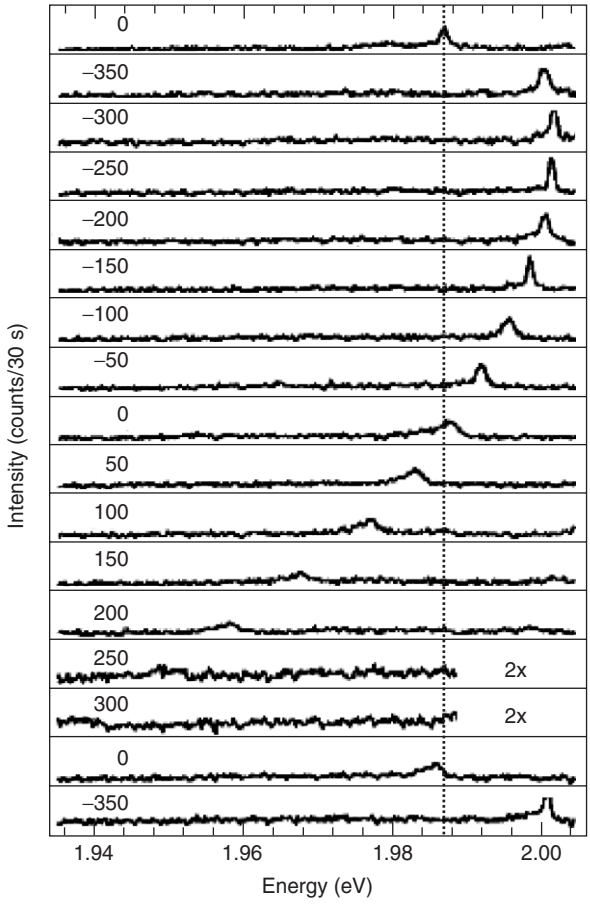


Figure 4 Luminescence of a 7.5 nm CdSe nanocrystal shows the quantum-confined Stark effect. The position of the exciton line shifts under a range of electric fields – values in kV cm^{-1} are shown for each trace. (From Empedocles SA and Bawendi MG (1997) *Science* 278: 2114.)

Interactions between Excitons

Excitonic Molecules and Complexes

Like atoms, excitons interact and may bind to form molecules and other complexes. Atomic H binds to form H_2 in an exothermic reaction that generates about 1/3 Ryd per molecule. The analog of H_2 is the biexciton (sometimes labeled as X_2) – which is however typically less bound than H_2 because of the finite hole mass. In a bulk semiconductor with equal mass electrons and holes, the binding energy of a (hydrogenic) biexciton is about 0.03 E_B, proportionately one order of magnitude smaller than for H_2; but in GaAs, where the hole is proportionately heavy, the binding energy can be substantial (see **Figure 5**).

Stable species can also be made when excitons bind to electrons and holes (trions) – a situation which can arise upon optical excitation of a low-density electron or hole gas in a semiconductor. In general, many exciton complexes have been discovered.

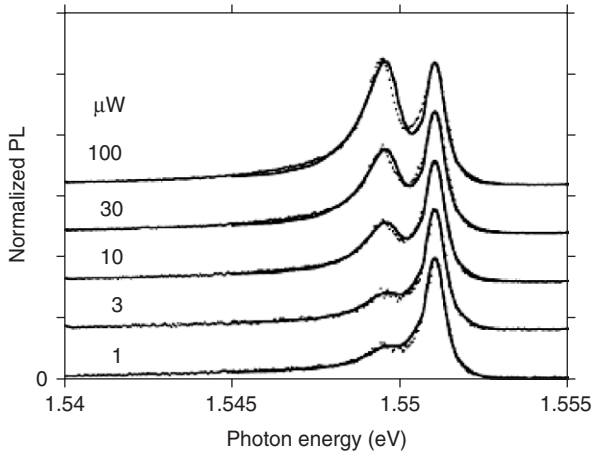


Figure 5 Optical emission from excitons and biexcitons in GaAs, normalized to the intensity of the excitonic peak and resulting from optical pumping at an energy well above the exciton emission. At increasing intensities of illumination, the biexciton emission grows because two excitons must collide and bind within the timescale for recombination. The solid line is a fit to the data. (From Phillips RT, Lovering DJ, and Denton GJ (1992) *Physical Review B* 45: 4308.)

Electron–Hole Liquid

If the density of excitons is increased so that they overlap, they lose their individual character as atoms, as the electrons and holes become unbound. The high-density two-component plasma so formed is then an electrical conductor – with one fermi surface each for electrons and holes – similar to an equilibrium semimetal. The transition between the excitonic insulator and the plasma state occurs at the density $n(a^*)^d \approx 1$ for the Mott metal–insulator transition. This has for many years been a focus of study in semiconductors because the large effective Bohr radius means that this regime can be reached at moderate excitation levels. A typical semiconductor laser is in this density range, though also usually at a temperature well exceeding the exciton binding energy.

In Si and Ge, it turns out that the dense plasma phase is lower in energy than to distribute the electron–hole pairs as a dilute gas of excitons or biexcitons. The origin of this phenomenon is the multivalley nature of the conduction band in these materials – in the plasma phase the kinetic energy cost of the excitons is distributed over many valleys so its stability is correspondingly enhanced. In consequence, high-density electron–hole droplets (**Figure 6**) form spontaneously and can also be manipulated by applied elastic strain.

Coherence and Bose–Einstein Condensation

Because excitons are light-mass bosons (albeit composite), there has been a long-standing interest to

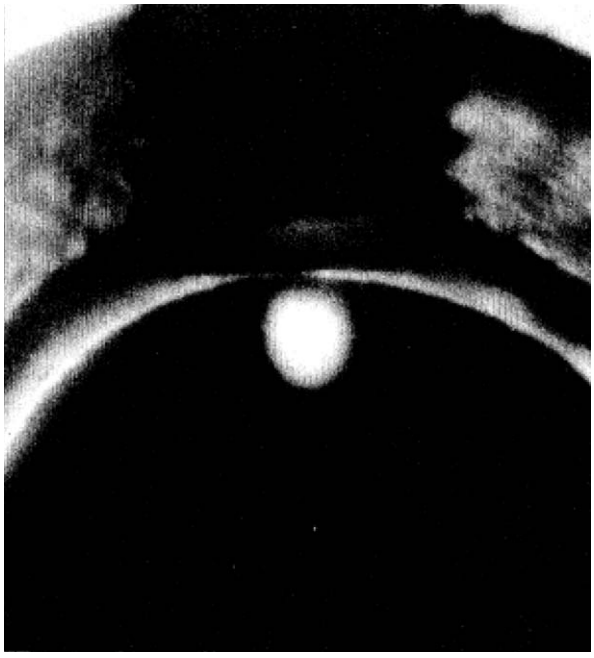


Figure 6 Electron–hole liquid in Ge. A photograph of a droplet of electron–hole liquid at 2 K in a 4 nm diameter Ge disk, stressed by a nylon screw (top) along (110) and viewed through an (001) face. The central bright region is the *e–h* recombination luminescence from the liquid confined in a potential well of maximum shear stress. The bright ring around the sample is recombination luminescence scattered from the surface of the crystal. (From Markiewicz RS, Wolfe JP, and Jeffries CD (1977) *Physical Review B* 15: 1988.)

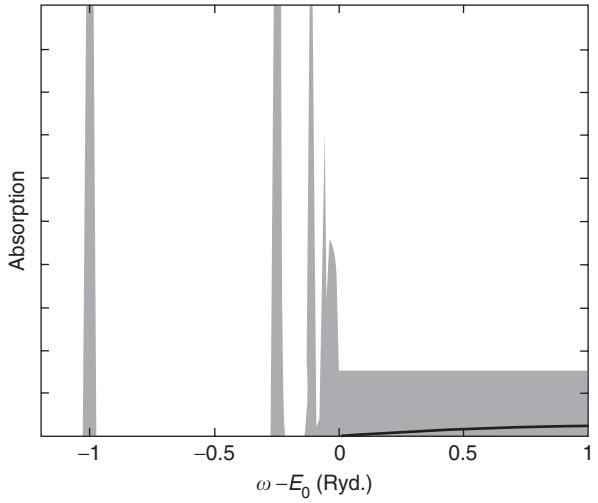


Figure 7 Absorption predicted by the Elliott formula of eqn [7]. Notice that even above the continuum edge at E_0 the absorption (gray) is enhanced over that predicted in the absence of Coulomb interactions (solid line).

observe Bose–Einstein condensation of excitons. In order to explore such physics, one needs to lower the temperature below the point where the thermal de Broglie wavelength exceeds the interparticle spacing, that is, a temperature $k_B T_o = \hbar^2 n^{2/3}/M$. Amongst the systems studied are Cu_2O (which has long-lived excitons that are dipole forbidden), CuCl (which has very stable biexcitons), and two-dimensional quantum wells (in which the electrons and holes can be physically separated). Such a state would be interesting not just for its statistical physics, but also because of the presence of spontaneous phase coherence, as in superfluid He, or cold atomic condensates.

While it has proved possible to enforce or encourage coherence through optical means, evidence for Bose–Einstein condensation of excitons in thermal equilibrium is still meagre.

Optical Properties

As discussed above, one of the principal reasons that excitons are so evident is because of their strong effects on the optical properties of solids. For very localized Frenkel excitons, it is largely correct to think of excitons as being atomic or molecular species, localized on specific sites in the solid. In general, it is not simple to calculate optical properties of excitons, except for the specific (and fortunately rather common) case of the hydrogenic Wannier exciton.

Dipole Coupling

Here, only the common case where the conduction and valence bands are of relative odd parity is considered (as in zinc blende II–VI and III–V semiconductors) so that the dipole matrix element

$$p_{cv} = \langle \psi_{ck} | \boldsymbol{e} \cdot \hat{\boldsymbol{p}} | \psi_{vk} \rangle_{k=0} \qquad [6]$$

is nonzero and in a cubic system independent of the polarization direction $\boldsymbol{e}$. The probability P per unit time and per unit volume that an electromagnetic wave with a vector potential $A\boldsymbol{e}\, e^{i\omega t - \boldsymbol{q}\cdot\boldsymbol{r}}$ will be absorbed to create an exciton is given by the Elliott formula

$$P = \frac{4\pi}{\hbar}\left(\frac{eA}{m}\right)^2 |p_{cv}|^2 f(0)^2 \delta(E_{ex} - E_0 - \hbar\omega) \qquad [7]$$

where the wavelength of light is assumed to be long enough so that $\boldsymbol{q} = 0$.

The answer is a sequence of lines at the energies $\hbar\omega_n = E_0 - \mathrm{Ryd}^*/n^2$ corresponding to the principal quantum numbers $n = 1, 2, 3, \ldots$, their intensities depending on the relative wave function $f(r)$ evaluated at the origin (when the electron and hole are coincident). Thus, the absorption is largest if the

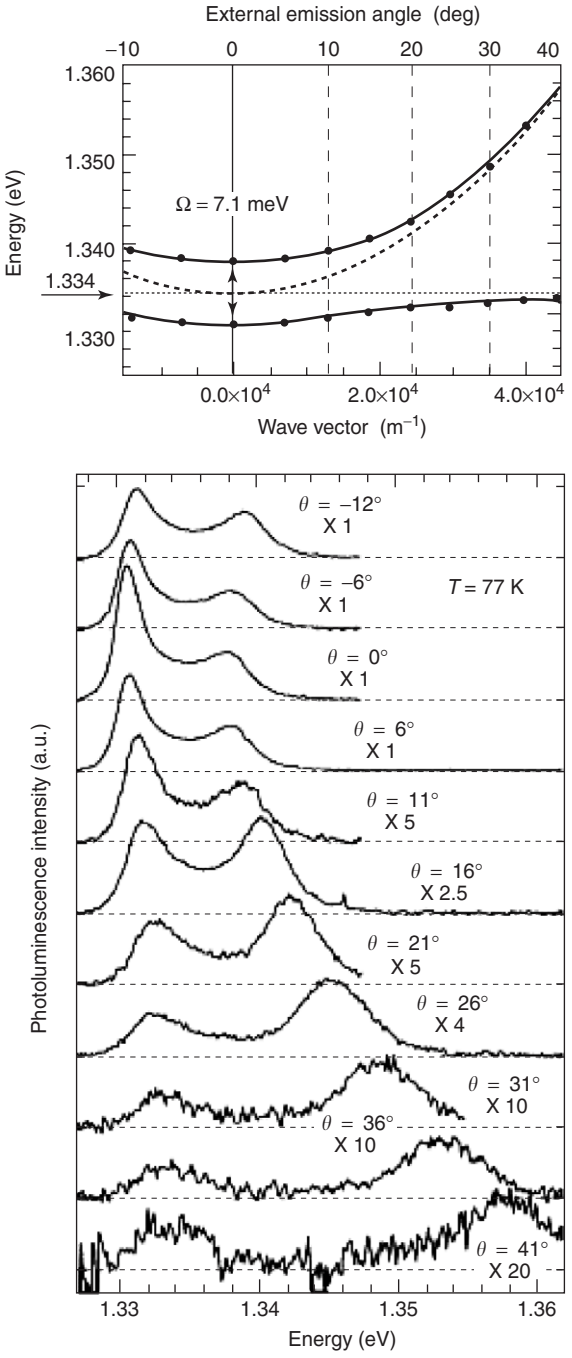


Figure 8 Dispersion of two-dimensional microcavity polaritons. Light is confined between two planar mirrors so as to fix the perpendicular momentum and to give a dispersion $\omega = \sqrt{\omega_0^2 + c^2k^2}$, where k is the momentum parallel to the plane, $\omega_0 \approx c\pi/L$, where L is the separation between the mirrors, and this can be chosen to be resonant with the exciton energy as in the figure (a). On the scale of the wavelength of light, the exciton dispersion is negligible. The mixing of the two modes gives rise to level repulsion, and the dispersion marked by the solid points. This dispersion is derived from the lower figure, which shows the luminescence from the microcavity as a function of angle, which can be directly converted into the parallel momentum k. (From Houdre R, Weisbuch C, Stanley RP, Oesterle U, Pellandini P, and Ilegems M (1994) *Physics Review Letters* 73: 2043.)

exciton is tightly bound (since $f(0)^2 = 1/\pi(a^*n)^3$) and also if the basis states supporting the Bloch states are strongly dipole active. Equation [7] gives only the Rydberg series below the continuum, but above and close to the edge of the continuum absorption, the excitonic effects enhance the absorption above that to be expected in free electron theory, as shown in **Figure 7**.

Exciton Polaritons

A peak in the absorption such as shown in **Figure 7** corresponds to a pole in the real part of the dielectric function, so below the gap,

$$\varepsilon(\omega) \approx \varepsilon_\infty + \frac{A_n}{\omega_n - \omega} \qquad [8]$$

where the amplitudes A_n are proportional to the coefficients in eqn [7]. A pole in the dielectric response is a free mode of oscillation of the polarization; when light travels in a polarizable medium such as this, the coupled modes are called "polaritons." These modes can be made visible beautifully by confining the light between mirrors to form a planar microcavity (**Figure 8**).

For excitons that are strongly localized and thus subject to disorder and incoherent fluctuations, one rarely observes fully coherent modes. But nevertheless the resonant process that transfers an exciton from one localized state to a resonant one nearby can be important, in a process called Förster transfer.

See also: Excitons in Crystals; Excitons: Theory; Organic Semiconductors; Polaritons; Semiconductors, General Properties; Semiconductors, Optical and Infrared Properties of.

PACS: 42.50.Ct; 71.35. – y

Further Reading

Cho K (1979) *Excitons*. Berlin: Springer.
Chuang SL (1995) *Physics of Optoelectronic Devices*. NY: Wiley.
Davydov AS (1971) *Theory of Molecular Excitons*. NY: Plenum.
Elliott RJ (1957) Intensity of optical absorption by exciton. *Physical Review* 108: 1384.
Fox M (2002) *Optical Properties of Solids*. Oxford: Oxford University Press.
Haug H and Koch SW (1993) *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2nd edn. Singapore: World Scientific.
Jeffries CD and Keldysh LV (eds.) (1983) *Electron–Hole Droplets in Semiconductors*. New York: Elsevier.
Knox RS (1963) *Theory of Excitons*. New York: Academic Press.
Moskalenko SA and Snoke DW (2000) *Bose-Einstein Condensation of Excitons and Biexcitons and Coherent Nonlinear Optics with Excitons*. Cambridge: Cambridge University Press.

F

Fatigue *See* Mechanical Properties: Fatigue.

Fermi Surface Measurements

C Bergemann, University of Cambridge, Cambridge, UK

Introduction

The Fermi surface is the envelope of the filled electronic states of a metal in momentum space. In the simplest and radically approximate view of a metal as free electrons in a box, for example, the electronic states are plane waves with an energy depending quadratically on the momentum. The Pauli exclusion principle prevents the electrons from all simultaneously occupying the lowest energy, zero momentum state, and instead the overall ground state is formed by the electrons occupying distinct states such that the total energy is minimized: they are bunched around the origin of momentum space in a sphere. In this case, the Fermi surface is simply the surface of that sphere.

Surprisingly, the basic notions of this picture survive even when the free-electron approximation is not applicable, for example, in the presence of substantial crystal lattice potentials or strong electron–electron correlations. The deep insight gained by Fermi-liquid theory is that the elementary excitations in most metals still possess a well-defined wave vector, and that there is still a Fermi surface separating filled from empty states. The main differences to the free-electron case are (1) that surface is, in general, no longer spherical, and (2) the effective electron mass can deviate from its bare value.

The physical properties of any system at low temperatures are determined by the low-energy excitations, and for a metal these are the states near the Fermi surface. This means that knowledge of the shape of the Fermi surface, and of the energy dispersion in its immediate vicinity, are of paramount importance in order to understand the properties of a metal.

Correspondingly, there are many direct as well as indirect ways of determining Fermi surface properties from experiment. On the most basic level, the DC conductivity is proportional to the overall Fermi surface area, for example, while the Hall effect can indicate whether the surface is mostly electron-like (enclosing filled states) or hole-like (enclosing empty states). However, a small number of experimental methods have proved to be extraordinarily sensitive and suitable to give a full account of the Fermi surface geometry, and this article attempts to give a brief introduction into these methods, with a particular focus on the de Haas–van Alphen (dHvA) effect and angle-resolved photoemission spectroscopy (ARPES).

New Materials and New Questions

Fermi surface measurements date back to the seminal discovery of the dHvA effect in bismuth in 1930. From that time until the 1970s, both measurement techniques and theoretical insight were gradually refined until the physics of most ordinary metals, in particular the elements, was very precisely and satisfactorily understood. This "golden age of fermiology" was superseded in the 1980s by a new era where novel materials broadly denoted as correlated electron systems have posed new challenges, both theoretical and experimental, to our understanding of metals. Examples of such new materials are the heavy fermion compounds, organic molecular crystals, high-T_c and other unconventional superconductors, and quantum critical systems.

These challenges have been met both with further refinements of existing techniques, like dHvA, and the arrival of new ones such as ARPES. At the same time, the experimental scope has shifted slightly, in that Fermi surface probes are increasingly used to study not merely the Fermi surface geometry itself, but also the influence of correlations on the electronic structure. Such experimental information is very valuable since a realistic and reliable quantitative prediction of the effect of correlations, and especially of the electron mass enhancement, is

presently well beyond the capabilities of theoretical band structure calculations.

One further development, less obvious at first sight, is that many correlated materials of contemporary interest possess a reduced effective dimensionality. In particular, the layered transition metal oxides and a number of organic molecular metals can be very anisotropic: quasi-2D or even quasi-1D. This has led to the development of new techniques specifically tailored for such materials of reduced dimensionality.

Quantum Oscillations – Physical Principles

The dHvA effect is the traditional workhorse of Fermi surface measurements. Broadly speaking, it relies on the interplay between the Fermi surface and orbital quantization in a magnetic field. This interplay is only visible under conditions that require significant experimental efforts: very low temperatures ($\sim$1 K or less), very large magnetic fields ($\sim$10 T or more), and for very pure materials, typically with mean free paths $\sim$1000 Å or longer, corresponding to low-temperature residual resistivities $\sim 1\,\mu\Omega$ cm or less. These efforts are rewarded with precise information about both the Fermi surface geometry and the thermodynamic effective masses.

To understand the origin of the effect, we need to consider the behavior of electrons in a magnetic field B. Classically, a freely moving electron in this case undergoes helical motion in real space; in momentum space, the electron traverses a circular orbit. In quantum mechanics, the solutions ψ_n of the Schrödinger equation in the presence of a magnetic field reflect this cyclotron motion, with a radial part peaked at a distance $r_n^2 = (2n+1)\hbar/eB$ from the center of the motion. The orbital radius and area in momentum space are $k_n = eBr_n$ and $A_{k,n} = (n+1/2)2\pi eB/\hbar$, respectively. Thus, the cyclotron orbits in k-space are confined to so-called Landau tubes with an area that is quantized in units proportional to B.

Consider a 2D gas of free, noninteracting electrons, at low temperature and in zero magnetic field. The filled plane wave eigenstates are inside the Fermi contour with radius k_F around the origin in momentum space. If a magnetic field is switched on, the electrons redistribute into the orbital eigenstates, and instead of having all k-states up to $k < k_F$ filled, one now has all the orbital states up to $k_n < k_F$ occupied. (Taking k_F, and with it the chemical potential, to be constant – this is valid in most 3D metals, but in materials of reduced electronic dimensionality this assumption can break down, necessitating a more intricate treatment than set out here.)

As the magnetic field is increased, the orbital areas $A_{k,n}$ grow proportional to B, and one by one the orbital states will pass through the Fermi contour. As a state does so, it abruptly becomes vacant. Since the degeneracy of the orbital states is extremely high, this means that the system is abruptly shedding a macroscopic fraction of its electrons. This occurs every time an orbit crosses the Fermi contour: $A_{k,n} = A_F$, where A_F is the Fermi contour area, or $\hbar A_F/eB = 2\pi(n+1/2)$. This implies that the electron number has an oscillatory component proportional to $\sin(2\pi F/B)$ as a function of inverse magnetic field, where F is the so-called dHvA frequency: $F = \hbar A_F/2\pi e$ (which has the units of magnetic field).

In real metals, which are non-2D, the basic concept remains the same: in the presence of a magnetic field, the quasiparticle eigenstates of a Fermi liquid redistribute analogously into orbital states in k-space along an equal energy contour of the electronic band dispersion perpendicular to the applied field. The geometrical shape of these orbits can vary, but provided one deals with closed orbits, their area is still quantized in units of $2\pi eB/\hbar$ (this is a consequence of the Bohr–Sommerfeld quantization rule). Orbital states still become vacant as they cross the Fermi level, but the major difference to the idealized 2D case is that now the area A_F of the equal energy contour at the Fermi level is not constant any more. The respective local dHvA frequency $F = \hbar A_F/2\pi e$ therefore varies in k-space, so that there is a continuous band of oscillations $\sin(2\pi F/B)$ which has to be integrated.

This integration can usually be simplified in the saddle-point approximation. The integrand $\sin(2\pi F/B)$ is rapidly varying, and only the extremal cyclotron orbits areas lead to a macroscopic magnetization. Therefore, the observed oscillation frequencies give information about the "maximum" and "minimum" cross-sections of the Fermi surface perpendicular to the magnetic field direction (**Figure 1**).

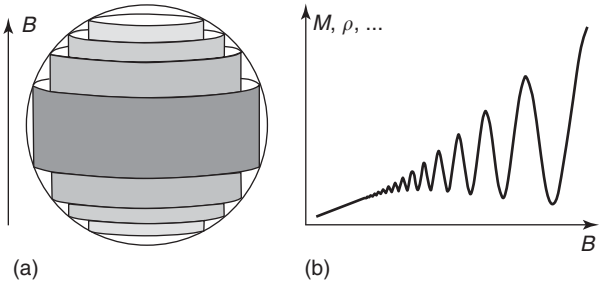


Figure 1 (a) Landau tubes and the Fermi surface in momentum space. As the magnetic field increases, the tubes expand and move through the Fermi surface, leading to (b) quantum oscillations. They are sketched here with typical impurity damping, leading to rapid loss of signal at low fields.

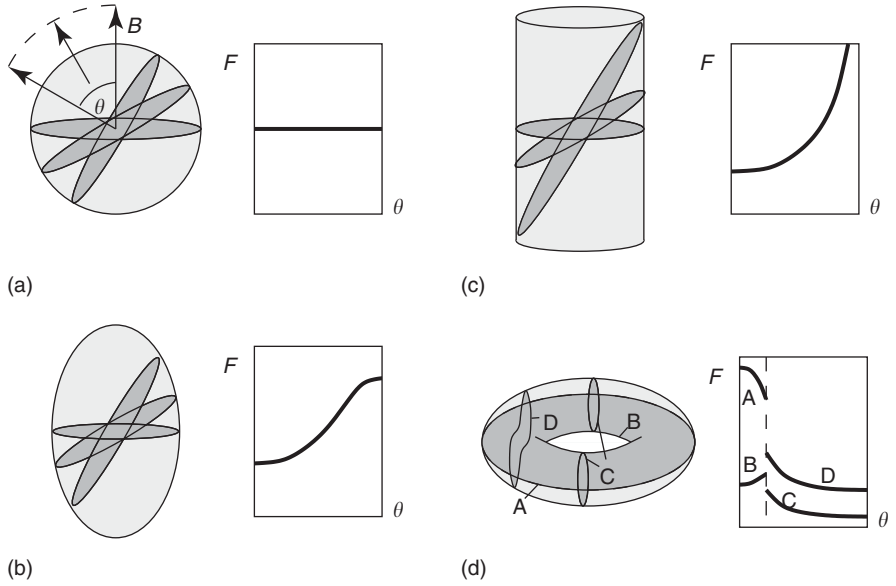


Figure 2 Fermi surface tracing using dHvA frequency information. (a) A spherical Fermi surface has direction-independent dHvA frequencies. (b) The cross-sectional areas, and dHvA frequency, of prolate ellipsoids increase for off-axis field directions, and (c) even diverge in the extreme case of a cylindrical Fermi surface. (d) More complicated features are possible, in particular for multiply connected surfaces with hole orbits.

In order to establish the overall shape of the Fermi surface one has to solve the inverse problem: given the observed dHvA frequencies along several field directions, what is the likely shape of the surface? This is usually done with the guidance of band structure calculations, but **Figure 2** gives a flavor of the relation between Fermi surface shape and the angular frequency behavior.

Quantum Oscillations – Experimental Approach

These quantum oscillations can, under favorable conditions, be observed in most thermodynamic and transport material properties. Experimentally, the best signal-to-noise level is usually achieved by probing quantum oscillations in the susceptibility (dHvA effect) or in the resistivity (Shubnikov–de Haas, SdH effect), the latter being the method of choice in restricted geometries without sufficient space for susceptibility coils, such as pressure cells.

The material in question needs to be of sufficient purity that, broadly speaking, the scattering time is longer than the cyclotron orbital period. This requirement translates, for typical Fermi velocities and Fermi surface geometries, into an impurity-limited low-temperature resistivity of $1\,\mu\Omega\,\mathrm{cm}$, or a mean free path of 1000 Å. These values constitute serious challenges to metallurgists and materials scientists, in particular for the binary and ternary compounds of contemporary interest. The purity requirement becomes less stringent in larger fields; so many dHvA experiments are carried out at dedicated DC or pulsed high magnetic field facilities.

Temperature Damping and Effective Masses

Finally, observation of the dHvA effect necessitates very low sample temperatures – but this temperature sensitivity turns out to be not just a vice, but also a virtue of the technique. The energy difference between two adjacent Landau tubes is $\hbar\omega_c^* = \hbar eB/m^*$, and if the thermal energy $k_B T$ exceeds this value, the quantum oscillations will be rapidly attenuated. The characteristic frequency ω_c^* and mass m^* are the fully renormalized thermodynamic values which – in most cases studied so far – add up to the α-coefficient in the low-temperature specific heat. ω_c^* and m^* usually differ from the values probed in a cyclotron resonance experiment.

An exact quantitative treatment shows that the oscillations will be damped by a factor $X/\sinh X$, known as the Lifshitz–Kosevich formula, where $X = 2\pi^2 k_B T m^*/\hbar eB$ is related to the ratio of the two energy scales above. The characteristic temperature scale is $\sim 1\,\mathrm{K}$ for carriers with bare electron mass m, but only 10 mK for carriers in a heavy fermion system with masses of order $m^* = 100m$. Conversely, the experimental determination of the temperature dependence of the dHvA amplitude can identify the electron masses with high accuracy.

Angle-Resolved Photoemission Spectroscopy

Photoemission spectroscopy provides, in principle, a direct way to measure the electronic density of states of a material. It uses the photoelectric effect: an incoming photon, typically in the ultraviolet, transfers its energy onto an electron and expels it from the material (**Figure 3a**). The energy difference between the photon and the electron corresponds to the binding energy of the electron relative to the vacuum level, and the energy distribution of the outgoing electrons hence mimics the electronic density of states and their occupancy inside the material (**Figure 3b**).

Photoemission experiments are conducted either at synchrotron sources, tapping into their vast supply of intense UV rays with variable energy, or in a more standard laboratory setup, typically using a He gas discharge lamp at a fixed photon energy of 21.2 eV. With modern detectors and at low temperatures (down to $\sim 10\,\mathrm{K}$), electron energy resolutions on $\sim 1\,\mathrm{meV}$ can be achieved. One problematic aspect of photoemission spectroscopy is the short escape depth of the photoelectrons, typically only a few Å for UV energies. Therefore, the technique is sensitive to a thin surface layer only, which necessitates careful sample preparation and treatment under UHV conditions. Even then, subtraction techniques using different photon energies often have to be employed to separate the surface from the "bulk" signal.

Importantly, for correlated materials the "density of states" probed by photoemission is really the spectral function $A(\omega)$ which can exhibit many-body features like Hubbard bands or superconducting, pseudo-, and Mott gaps. Photoemission spectroscopy has therefore become one of the most powerful modern probes of electronic structure.

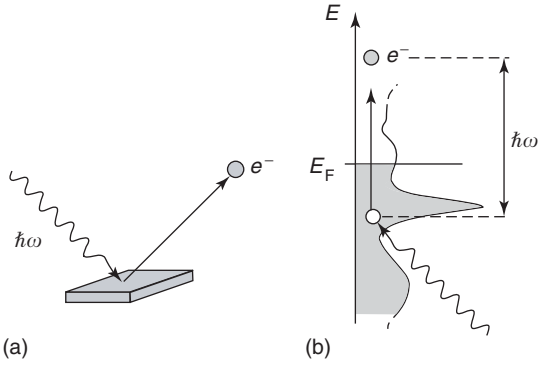


Figure 3 (a) Schematics of a photoemission experiment. (b) The energy distribution of the emitted electrons mimics the occupied density of states (i.e., $A(\omega)f(\hbar\omega - E_F)$; gray shaded area) inside the material.

The real power of photoemission is unleashed in angle-resolved spectroscopy (ARPES) on single-crystalline samples, where the momentum of the photoelectron is measured along with its energy. Within limits – discussed below – this photoelectron momentum can be related to the crystal momentum k of the initial electronic state. Within common approximations, the energy distribution curve measured at the detector for a particular angle, or momentum, is proportional to $A(k, \omega)f(\hbar\omega - E_F)$, the momentum-resolved spectral function times the Fermi function. This product in turn encodes the complete band structure and Fermi surface of the material since the spectral function $A(k, \omega)$ peaks whenever $\hbar\omega = E(k)$, while $f(\hbar\omega - E_F)$ demarcates the filled states in k space (**Figure 4**). For noninteracting electrons, the peak in $A(k, \omega)$ is δ-like, while for interacting systems the peak broadens out and is supplemented by an incoherent background contribution to $A(k, \omega)$.

Much of the quantitative analysis relies on what happens to the crystal momentum k when the electron is excited by the photon, refracted at the surface, and expelled from the material, to arrive at the detector with a momentum K. The momentum component parallel to the surface is usually taken to be conserved, apart from a reciprocal lattice vector, $k_{=} = K_{=} + G_{=}$. In contrast, the perpendicular component $k_{\perp}$ is usually not conserved during the refraction process. The analysis then either requires

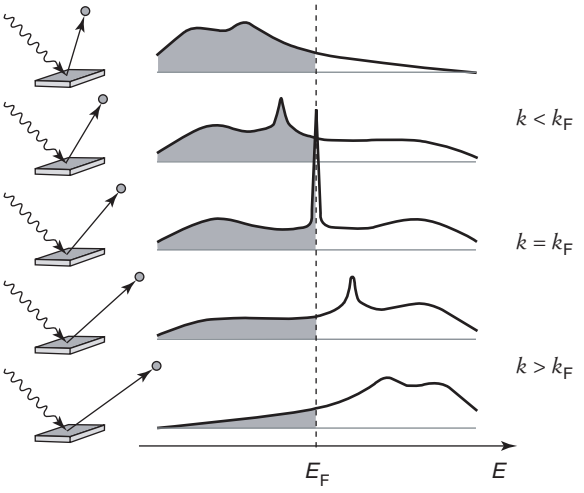


Figure 4 Angle-resolved photoemission spectroscopy (ARPES). For a particular emission direction, the momentum-resolved occupied spectral function $A(k, \omega)f(\hbar\omega - E_F)$ is probed (gray shaded). For Fermi liquids, $A(k, \omega)$ possesses a quasiparticle peak which indicates the band dispersion. The peak narrows when k approaches the Fermi surface from the inside and gets depopulated outside it. In this way, ARPES can probe band structure, Fermi surface, and many-body effects.

some assumptions about the (nearly free electron) nature of the excited state inside the crystal, or comparison of extensive data sets from different crystal faces, to determine $k_\perp$.

In the special case of materials with a quasi-2D electronic structure, the spectral function only depends weakly on $k_\perp$ anyway, and it is in these materials that ARPES is at its most powerful. Quasi-2D materials offer a few more experimental simplifications – for example, they often cleave easily, exposing flat crystalline faces with little surface reconstruction. This is true for the high-T_c superconductors, especially the Bi-based compounds, where ARPES has made major contributions to the understanding of the electronic structure of both the normal and superconducting states.

Angle-Dependent Magnetoresistance Oscillations

In metals with reduced effective dimensionality of the electronic structure (quasi-1D and quasi-2D), the magnetoresistivity for currents along the low conductivity axes shows a strong dependence on the direction of the applied magnetic field. This angle-dependent magnetoresistance oscillation (AMRO) effect was first observed in the late 1980s in a number of organic metals and subsequently explained as being due to "magic angles" in the Fermi surface geometry. In this way, the exact form of the AMROs constitutes a subtle probe of the shape of the Fermi surface.

One can understand AMROs from the viewpoint of semiclassical transport theory. If an electric field is applied to drive a current in a metal, this field acts as a force on the electrons and changes their momentum and hence their k value. Depending on the relative orientation of the electric field and the Fermi velocity, this amounts to an acceleration or deceleration of the electron, and hence to a positive or negative change in electron energy, respectively. This energy change induces an occupancy redistribution that deviates from the ground state, which in turn leads to a net current.

In the presence of a magnetic field one technically has to deal with orbital states again, but it is sufficient to view the problem in the semiclassical approximation in which the wave vector k is taken to be an almost good quantum number which slowly evolves along the cyclotron trajectory: along the Fermi surface and perpendicular to the applied magnetic field.

For some directions of the applied magnetic field, known as magic angles and to be determined below, the Fermi velocity component along some crystallographic direction can average to zero over every cyclotron trajectory. In this situation the time-averaged energy shift induced by a c-axis electric field is very small: over the course of each cyclotron trajectory, the electric field sometimes accelerates the electron, sometimes decelerates it, with little net effect. This means that the occupancy redistribution, and the resulting current, are both small – the magnetoresistance is large.

The two best-known magic angle situations are so-called "Yamaji oscillations" for quasi-2D metals, and "Danner–Chaikin oscillations" for the quasi-1D case. Both cases are illustrated schematically in **Figure 5**. Yamaji oscillations are visible in the interlayer resistivity as a function of the polar angle of the magnetic field, while Danner–Chaikin oscillations can be observed in the resistivity along the lowest conductivity axis as the magnetic field is rotated from that axis onto the highest conductivity direction.

The exact positions of the Yamaji magic angles depend on the details of the geometry of the Fermi surface. From **Figure 5a**, it is evident that, if the Brillouin zone height remains the same, larger Fermi surfaces will have smaller Yamaji angles. In fact, it can be shown that in the simplest case and in the clean limit the magnetoresistance behaves like $\rho_c = \rho_0 / J_0^2(ck_F \tan\theta)$ where c is the interlayer spacing, θ is the polar angle of the applied magnetic field, and J_0 is the Bessel function. In this way, the oscillation period (in $\tan\theta$) yields the Fermi wave vector, so that Yamaji oscillations constitute a "caliper" measurement of the Fermi surface. This periodicity also applies in more general cases although the exact shape of the AMROs can be different, corresponding to more unusual interlayer transport behavior.

For Danner–Chaikin oscillations, the magnetoresistance behaves like $\rho_c = \rho_0 / J_0^2(2t_b c \tan\theta / \hbar v_F)$ and gives information about the relative strength of the corrugation of the Fermi surface, and in particular the tight-binding transfer integral t_b along the intermediate conductivity direction.

In practice, the analysis of AMROs works best if the Fermi surface possesses one sheet only, and if a reasonable first guess at its geometry is available from theory or from other techniques. This is because quite often only the first few oscillations are visible, so that fitting procedures are required, and it is correspondingly difficult to disentangle the contributions from several bands.

Sufficient Fermi velocity averaging over the whole cyclotron orbit is only possible if the scattering rate is small and the magnetic field is large ($\omega_c \tau \geqslant 1$). AMROs are therefore subject to similar impurity damping as dHvA oscillations, but to much weaker

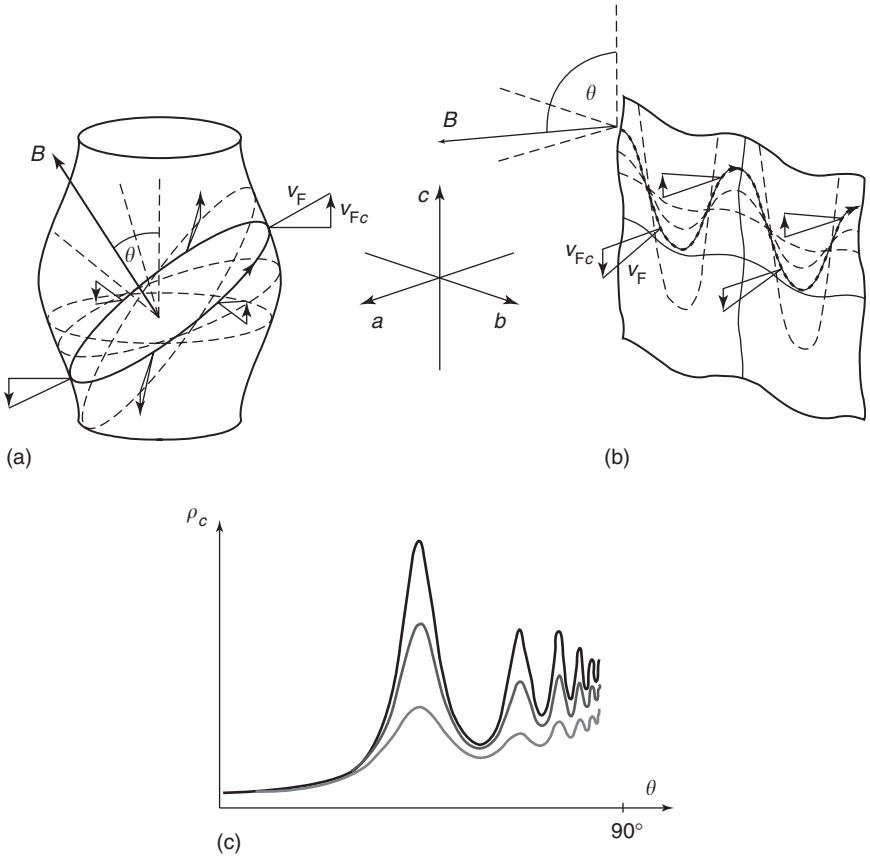


Figure 5 Geometry for (a) Yamaji and (b) Danner–Chaikin oscillations. The bold solid trajectory is at the magic angle where the *c*-component v_{Fc} of the Fermi velocity always averages to zero. (c) The magnetoresistance is similar in both cases and shows peaks at the magic angles – quasi-periodic in $\tan\theta$ – that grow stronger with increasing field.

thermal damping since AMROs are a semiclassical effect independent of orbital quantization.

Advantages and Disadvantages

Depending on the material and the situation, the three experimental methods discussed above each have their own relative merits and disadvantages – in this section some of the key aspects are covered for guidance.

Sensitivity to Correlations

DHvA probes the effects of correlations on the low-energy scale (<0.1 meV), through the effective electron masses. In contrast, electron correlations are manifest in photoemission and ARPES on a higher energy scale (typically 1 eV), in Hubbard bands and in the incoherent part of $A(k, \omega)$. In this sense, dHvA and photoemission provide very complementary information. Unfortunately, AMROs are largely independent of many-body renormalization effects and do not therefore constitute a suitable experimental probe for electronic correlations.

Purity Requirements

DHvA and AMROs have similar crystal purity and homogeneity requirements ($\omega_c\tau \geqslant 1$); ARPES is much less stringent as long as the surface preparation is adequate. This is one of the key reasons why ARPES has been so useful for investigating the high-T_c cuprates.

Temperature Requirements

DHvA is most demanding ($k_BT < \hbar\omega_c^*$); AMROs remain visible until the temperature-induced extra scattering invalidates the condition $\omega_c\tau \geqslant 1$; the ARPES signal remains observable at much higher temperatures (with an energy resolution slightly broadened by k_BT). This means that typically AMRO and dHvA studies are restricted to the low-temperature ground state of a material, while ARPES can often probe electronic phase transitions at elevated temperatures.

Dimensionality

AMROs are only visible in quasi-1D and quasi-2D materials. A reduced dimensionality is also helpful for the interpretation of ARPES and boosts the signal size in dHvA, but both techniques still retain much of their power for 3D materials – dHvA perhaps even more so than ARPES.

k-Space Resolution

In favorable situations, dHvA can resolve k-space features ~ 10 ppm of the Brillouin zone; more typically, it can achieve $\sim 0.1\%$ resolution with accurate magnetic field calibration. While significant improvements have been made in ARPES, state-of-the-art detectors discriminate by $\sim 0.2^\circ$ in angle, and Fermi surface determination down to within perhaps 1% of the Brillouin zone is often possible. The resolution offered by AMROs depends strongly on the number of oscillations visible in the magnetoresistance – $\sim 1\%$ precision in k_F is realistic for Yamaji oscillations in good crystals.

Bulk Sensitivity

DHvA and AMROs probe the bulk of the material, while ARPES is restricted to a thin surface layer. The vicinity and influence of the surface can on occasion distort the electronic structure, especially when the exposed surface undergoes a structural and/or electronic reconstruction. Photons with variable energies and escape depths can be used to separate surface from bulk contributions, but this separation is often not reliable, and cases where a variation of photon energy itself drastically modifies the observed apparent electronic structure have also been reported.

Other Fermi Surface Probes

The three experimental methods discussed above are probably the most common Fermi surface measurement techniques today. There are a number of other probes, some of them occupying specialist niche roles, others mostly of historical interest. These are briefly presented here for completeness.

Positron Annihilation

If positrons from a radioactive source are injected into a solid, they thermalize in a very short time, losing most of their momentum in the process. After about a few hundred picoseconds, the positron and one of the electrons in the material annihilate in a flash of two γ-rays. The γ-ray trajectories are almost back-to-back, but angular corrections can be observed and attributed to the initial crystal momentum of the electron. This maps out the filled states of the material in k-space. The advantage of the technique is that it is a rare bulk probe that also works at elevated temperatures, or for inhomogeneous materials. On the downside, it only offers filling information that is integrated along lines in k-space; its resolution is less than spectacular (typically 5% of the Brillouin zone); and the effects of positron trapping in lattice vacancies can distort the observed bulk electronic structure.

Cyclotron Resonance

The cyclotron motion of the electrons of a metal in a magnetic field causes resonant absorption of an incident microwave field if its frequency coincides with the cyclotron frequency $\omega_c = eB/m_c$, or – in the so-called Azbel'–Kaner geometry – integer multiples thereof. Primarily, the technique does not measure the Fermi surface geometry itself but the dynamical cyclotron mass m_c. Importantly, m_c can differ from both the bare band mass and from the effective mass m^* measured in dHvA, and those differences give clues about the electronic interactions in the material. In addition, recent measurements on quasi-1D and -2D systems have uncovered geometrical resonances (known as Fermi surface traversal and periodic orbit resonances, respectively) which can be more closely related to the Fermi surface geometry itself.

Magnetoacoustic Effect

Rarely used today, this method relies on resonant attenuation effects between cyclotron motion in real space and propagating sound waves in a metal every time the cyclotron orbit diameter $2r_c$ and the sound wavelength λ form a half-integer ratio, $2r_c/\lambda = n + 1/2$.

Anomalous Skin Effect

Famous for its role in the experimental determination of the Fermi surface of copper, this technique is also very rarely used today. It exploits a subtle relation, valid only if the mean free path is much longer than the skin depth, between the magnitude of the microwave surface impedance and the Fermi surface curvature near those points where v_F is parallel to the sample surface.

See also: Effective Masses; Electrons in Periodic Potentials with Magnetic Fields; Liquids, Theory of: Fermi Liquids; Magnetoresistance Techniques Applied to Fermi Surfaces; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Valence Photoemission.

PACS: 71.18.+y; 71.20.−b; 71.27.+a; 72.15.Gd; 72.55.+s; 78.70.Bj; 78.70.Gq; 79.60.−i

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*, chs 12–15. Philadelphia: Saunders College Publishing.

Damascelli A (2003) *Probing the Low-Energy Electronic Structure of Complex Systems by ARPES*, cond-mat/0307085.

Hüfner S (1995) *Photoelectron Spectroscopy: Principles and Applications*. Berlin: Springer.

Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.

Singleton J (2001) *Band Theory and Electronic Properties of Solids*. Oxford: Oxford University Press.

Singleton J, Blundell SJ, and Harrison N *Magnetic Oscillations in Strongly Correlated Metals*. Oxford: Oxford University Press (to appear).

Springford M (ed.) (1980) *Electrons at the Fermi Surface*. Cambridge: Cambridge University Press.

Wosnitza J (1996) *Fermi Surfaces of Low-Dimensional Organic Metals and Superconductors*. Berlin: Springer.

Ferroelectricity

A P Levanyuk, Madrid Autonomous University, Madrid, Spain
B A Strukov, Moscow State University, Moscow, Russia

Introduction

Ferroelectricity refers to many related phenomena specific to ferroelectrics. Properties of these materials have made them important for various applications, some of which are mentioned in this article. Most ferroelectrics have small conductivity and are considered to be a special class of insulators (dielectrics). Some ferroelectrics are semiconductors. However, considering them as dielectrics is an adequate starting point. Ferroelectrics exhibit unconventional pyroelectricity and piezoelectricity. The article begins with a classification of electrical properties of dielectrics with a view to identify the position of ferroelectrics among related classes of materials. A plethora of specific features of ferroelectrics is connected to their "structural softness," which is an indication of the closeness of a second-order phase transition. The meaning of this "closeness" is explained below; considering this transition is the key to understanding ferroelectricity. The transition is included in the formulation in a way that is model-independent and general, but somewhat formal. In order to illustrate this approach and also give general ideas on the microscopic mechanisms of the transition, this article employs a simple model that nonetheless offers a physically transparent treatment.

Ferroelectrics, Pyroelectrics, and Piezoelectrics

Properties of a dielectric can be derived from its equation of state, which describes the dependence of the polarization P, on parameters that define the state of the material. The most important among these are voltage, elastic strains, and temperature. In fact, it is not the polarization that is probed experimentally, but either the charge of the capacitor containing the dielectric (**Figure 1**) or the current in the circuit containing the capacitor (**Figure 2**). According to electrostatics, the density of charge at the capacitor plate is equal to the electric displacement $D = \varepsilon_0 E + P$, where $E = V/L$ is the electric field, (L being the distance between the electrodes and V the voltage), and ε_0 is the permittivity of vacuum. (It should be noted that in normal experimental conditions, $\boldsymbol{D}$ is along the perpendicular to the plate. The use of vectors is thus avoided in this article by assuming that all vectors are directed along the same line.)

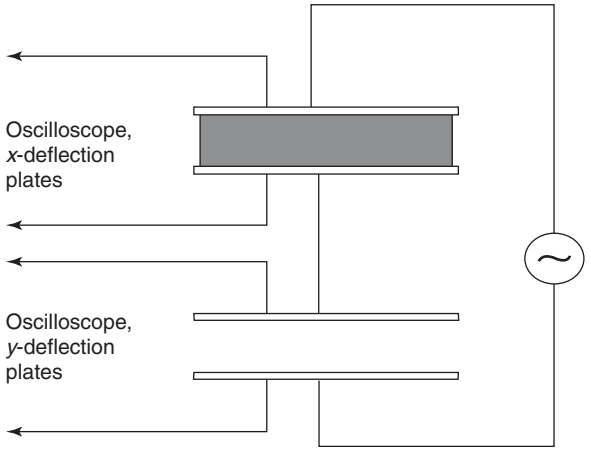


Figure 1 Probing the dependence of polarization on the electric field by measurements with alternating voltage. The y-deflection is proportional to the charge Q of the reference (empty) capacitor, which is the same as the charge of the capacitor filled by dielectric. The x-deflection is proportional to the voltage on the dielectric plate. Since $D = Q/S$, where S is the area of the plate, the experiment provides $D(E)$ and the dependence $P(E)$ can be then calculated. In ferroelectrics $\varepsilon_0 E \ll P$, so that $D \sim P$.

By controlling the voltage on the dielectric plate and the charge of the reference capacitor (**Figure 1**), one can obtain the dependence $P(E)$. For most dielectrics, P depends linearly on E, up to dieletric breakdown fields (**Figure 3a**). In ferroelectrics, however, $P(E)$ has a different form (**Figure 3b**). The measurement frequency is so low (see caption for **Figure 3**) that the dependence $P(E)$ in **Figure 3b** is considered as "static." Remarkably, in ferroelectrics $P(E)$ is nonlinear and even multivalued. There are two values of P for $E = 0$. After switching the voltage off, the ferroelectric finds itself in one of the two states that correspond to the remanent ("spontaneous") polarization of opposite signs. By applying an electric field in the direction opposite to the spontaneous polarization and exceeding the "coercive field" (see **Figure 3b**), the spontaneous polarization can be inverted (switched). This property makes ferroelectrics useful in memory and logic devices. This was proposed before the 1950s but has become commercially significant only recently. Another application of ferroelectrics is connected with the fact that their dielectric response is large: note the difference in scales of the P-axis in **Figures 3a** and **3b**. If the amplitude of the change in the field is much less than the coercive field, the ferroelectric behaves as an ordinary dielectric, but with a very high dielectric constant $\varepsilon = 1 + dP/dE$. This explains why ferroelectrics are used in $\sim 50\%$ of the total number of the commercially available capacitors: high capacitance can be achieved in a small volume. The name "ferroelectrics" is due to the similarity of the hysteresis curves of the ferroelectrics (**Figure 3b**) and the ferromagnetics.

Another type of experiment is presented in **Figure 2**, where there is no voltage source and one probes only the current. In most dielectrics, no current appears but in some of them a transient current is observed to flow. If it flows because of a change in the temperature of the dielectric, one observes the pyroelectric effect. If it does because of a deformation, one observes the piezoelectric effect. The dielectrics in these two cases are, respectively, pyroelectrics and piezoelectrics. These effects are not completely independent. Indeed, there are two types of deformations: those lowering the symmetry of the crystal (e.g., unilateral stretching of a cubic crystal), and those leaving it unchanged (e.g., deformation due to hydrostatic pressure). Since the deformations due to thermal expansion are symmetry-conserving, the piezoelectric effect is a contribution to the pyroelectric effect, that is, the two effects take place simultaneously. Conversely, if the symmetry-conserving strains do not provoke the piezoelectric effect, the crystal is not a pyroelectric. Symmetry-lowering strains may not necessarily lead to the piezoelectric effect. Symmetry arguments of this type are very effective in predicting the crystal properties. It is the crystal symmetry that "allows" or "forbids" these properties. Later this point is illustrated for the case of a simple model, but here it suffices to mention that the pyroelectric effect is possible in structures

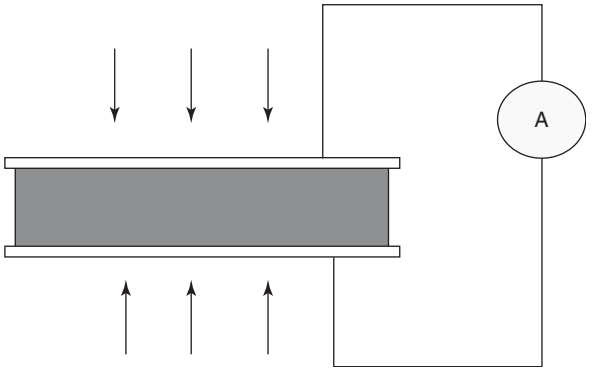


Figure 2 Probing the polarization change by measurement of transient current provoked by change of temperature (pyroelectric effect) or by deformation (piezoelectric effect). The current $I = S\ dP/dt$ in a circuit without resistance. The arrows indicate either heat flow or mechanical stress.

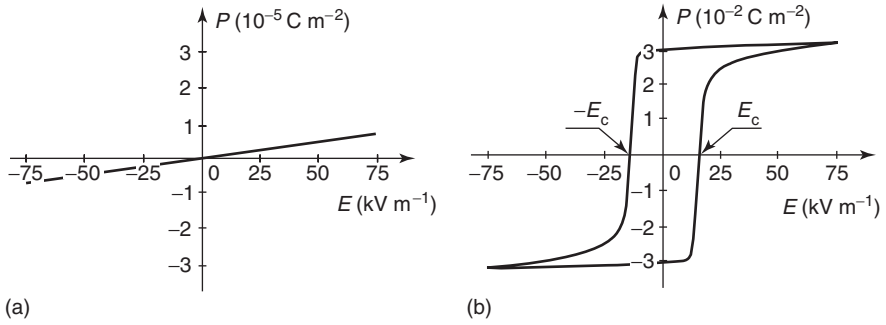


Figure 3 Dependence $P(E)$ (a) for a linear dielectric with $\varepsilon = 10$ and (b) obtained experimentally for ferroelectric triglycine sulfate (TGS) (see **Table 1**) at ambient temperature and frequency 0.1 Hz. The coercive field $E_c = 15\,\mathrm{kV\,m^{-1}}$, the spontaneous polarization is $3.1 \times 10^{-2}\,\mathrm{C\,m^{-2}}$.

whose point group does not contain transformations that change a vector. Such structures are generally called polar.

Ferroelectrics have (at least) two polar structures: the pyroelectric current has opposite directions in the two states with opposite spontaneous polarization, that is, the temperature influences the modulus of the spontaneous polarization but not its direction. The same is valid for the piezoelectric effect, because of the symmetry-conserving strains. This is not the only specific feature of a ferroelectric as a special case of pyro- and piezoelectric. As noted below, the magnitude of the two effects can be much higher in these materials than in ordinary pyro- and piezoelectrics. There are large scale practical applications of ferroelectrics as pyroelectric thermal detectors and, especially, as piezoelectric transducers.

Ferroelectric Phase Transition

An important property of ferroelectrics is that they change their structure to become nonpolar at relatively moderate changes of their external conditions, for example, by heating to temperatures that are much less than the melting temperature or by applying a low enough pressure that does not change the density significantly. This is another aspect of the polar state "softness," which is a characteristic of the dielectric response of ferroelectrics. The reason of this softness is that the polar structure of a ferroelectric is a result of small structural changes of a nonpolar structure; here, "small" means, in particular, that the average positions of the ions are shifted by values much less than the typical interatomic distances.

The observation that some structural changes though arbitrarily small lead to a change in the crystal symmetry, that is, to a phase transition is the basis of the Landau theory of second-order (continuous) phase transitions. This theory allows one to predict specific properties ("anomalies") of a crystal close to a second-order transition. The term "close" has the same meaning as the above "small." In this sense, most ferroelectrics are almost always "close" to the transition. This is not surprising if one takes into account that the temperatures of ferroelectric (nonpolar–polar) phase transitions are typically several hundred kelvin, whereas the characteristic scale of the temperature dependence of crystal constants is the so-called "atomic temperature" (T_{at}) estimated as 10^4–10^5 K. The estimation is made by comparing the kinetic energy of thermal motion of an ion and the binding energy, that is, the work necessary to remove it from the crystal. Therefore, the Landau theory is useful in understanding the properties of a ferroelectric, say, at ambient pressure and temperature. This was demonstrated by Ginzburg and Devonshire, and thus the Landau theory applied to ferroelectrics is usually called Landau–Ginzburg–Devonshire (LGD) theory.

Ferroelectric phase transitions *per se* are also of great interest. In the vicinity of a transition, the crystal has unusual properties, which are discussed below. It is significant to note that the Landau theory is not applicable in some neighborhood of a second-order phase transition. This theory is applicable "close" but not "too close" to the transition. For a long time, it was difficult to develop a theory valid in the immediate neighborhood of the transition. This development is frequently characterized as the main achievement of statistical physics of the twentieth century, which has had an important impact on other branches of physics. However, it seems very difficult to apply this theory to interpret experimental data regarding ferroelectrics. First, this immediate neighborhood ("scaling region") proves to be fairly narrow in ferroelectrics. Second, and even more important, various and poorly controllable crystalline defects become important in this region where the phase-transition anomalies are changed and often the transition is "smeared." Beyond the scaling region, the observed deviations from the Landau theory are small and should be interpreted within another theory, which is more simple and more informative, than the strict theory, but, is inapplicable in the immediate neighborhood of the transition. Beyond this immediate vicinity, the LGD theory can well account for the main properties of ferroelectrics.

An important advantage of this theory is its generality: it only uses symmetry considerations and some other macroscopic arguments, that is, it is model-independent. This theory provides the shortest way to understand the physical background of ferroelectricity; it establishes relations between different experimental data, but it cannot provide numerical values of the relevant physical quantities without the use of other experimental data. It cannot predict, in particular, the composition of crystals that could have, especially, advantageous ferroelectric properties. This is why theories that start from a basic quantum mechanical approach are now being intensively developed. At the intermediate level, various types of microscopic model theories are considered that help in understanding the underlying physical mechanisms specific for a given class of crystals.

One-Ion Model

For pedagogic convenience, the basic ideas of the LGD approach are presented. It has been mentioned

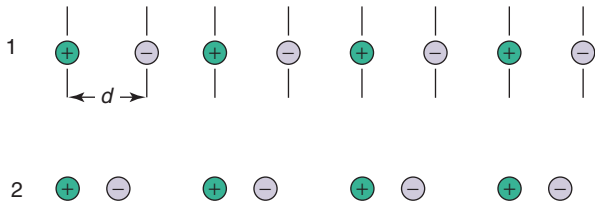


Figure 4 Two periodic ionic chains ("one-dimensional crystals"). Chain 1 has symmetry planes perpendicular to the page and crossing the ions. Chain 2 is a result of displacement of anions of chain 1 by the same distance. The symmetry planes are lost because of the displacement ("breaking of symmetry").

already that the structures of polar and nonpolar phases of a ferroelectric are very similar. An example of two structurally similar one-dimensional "crystals" is presented in **Figure 4**. The structures have different symmetries: the symmetry planes in structure 1 are perpendicular to the plane of the page and pass through the ions. In structure 2, they are absent.

Pyroelectric Effect and Spontaneous Polarization

Imagine a crystal consisting of parallel chains with electrodes on the surfaces perpendicular to the chains and short-circuited (**Figure 2**). The pyroelectric effect means that a change in temperature leads to a transfer of charge from one electrode to the other. Since $E = 0$, the density of charge at the electrodes is proportional to P. It will now be checked whether the polarization changes with temperature.

Structures 1 and 2 both have a dipolar moment. Still, the more symmetric structure does not exhibit the pyroelectric effect. Indeed, the dipole moment of a pair of neighboring negative (anion) and positive (cation) ions is qd, where q is the ion charge and d is the distance between the ions. The dipole moment of the whole crystal is qdN_1N_2, where N_1 is the number of anions (cations) in a chain and N_2 is the number of chains. The volume of the crystal is $v_f = S \times L = S \times (2N_1 - 1)d$, where S is the area of the capacitor electrode and L the distance between electrodes of a capacitor, and the polarization (dipole moment per unit volume) is $qN_2/2S$. The charge at the electrode is $qN_2/2$. This charge is independent of temperature. For structure 2 (**Figure 4**), there is a contribution to the polarization which is associated with a difference in distances between, say, an anion and two neighboring cations and not with the unit cell length $2d$. This contribution is dependent on temperature.

Frequently, pyroelectricity is ascribed to materials having "spontaneous polarization," normally understood as the dipole moment of the sample divided by its volume. However, this can be confusing.

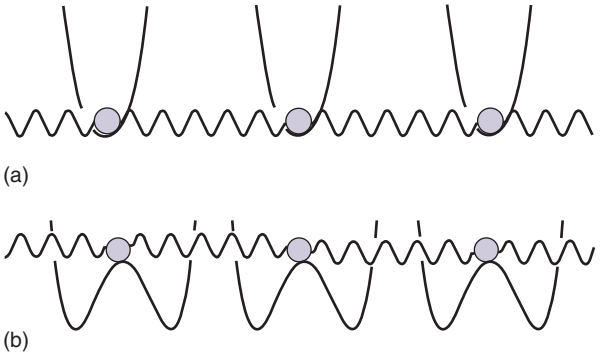


Figure 5 Diagrams illustrating the model: (a) $p > p_c$, stable symmetric configuration and (b) $p < p_c$, unstable symmetric configuration, the anions will shift to the right or to the left.

Symmetry does not prohibit the existence of the so-defined spontaneous polarization in a finite crystal, although it prohibits the pyroelectric effect. But if the term spontaneous polarization is understood as the difference between the polarization of a given structure and that of the "reference" (more symmetric) structure, just as in **Figure 4**, the term becomes meaningful. It does not contradict the above empirical definition, and one can associate the pyroelectric effect with the spontaneous polarization defined in this manner.

Symmetry Breaking

To illustrate a nonpolar–polar phase transition, consider a model in which the cations are immobile and their role is to create potential energy for the anions ("one-ion model"). This energy does not originate because of the Coulomb interaction between the ions which is neglected for now. A short-range interaction of an ion with only its nearest neighbors is considered. The interaction between anions in different cells is also short-ranged and is modeled with springs (**Figure 5**). The ions are heavy enough to be considered as classical and not quantum particles ("quantum fluctuations" are neglected). The phase transition occurs because of a change in the pressure p. Suppose that at $p > p_c$ the potential energy has only one minimum (**Figure 5a**). The anions are, of course, in this minimum. If, for $p < p_c$, the potential energy is a two-well one (**Figure 5b**), the anions shift either to the right or to the left, forming a less symmetric structure with the loss of structural equivalence of two neighboring ions. The springs do not deform: all the anions shift by the same distance. Only one variable is needed to describe different configurations of the systems: the shift of any anion. The role of the springs is to force them to shift collectively.

LGD Theory for the Model

For models with classical ions at $T=0$, the Landau theory is exact. What this theory means for the above model is discussed below.

The Landau theory describes lowering of the symmetry as a result of some internal deformation, which nowadays is called the "order parameter." In this case, the order parameter can be identified with the shift of anions or with a difference in polarizations of structures with shifted and nonshifted anions. Here it is termed "polarization" and designated as P. It has been seen earlier that it is the only variable needed. Choosing the simplest function to describe the evolution of potential energy of an anion, the potential energy per unit volume U is obtained in the form

$$U = N\left(a\frac{P^2}{2} + b\frac{P^4}{4}\right) \qquad [1]$$

where N is the number of anions (cations) in a unit volume, $a = a'(p - p_c)$ and $b = \text{const}$.

Even if the Coulomb interaction can be neglected when considering the interaction between the ions ("at microscopic level"), the shift of the anions leads to a macroscopic effect. Let a capacitor with the ferroelectric (FC) (**Figure 6**) be connected to a voltage source capacitor (SC) with $C \to \infty$ (C is the capacitance and voltage V). For $P=0$, one has $U=0$ and the total energy is due to the electric field:

$$\int_{v_f} \frac{\varepsilon_0 E^2}{2}\,dV + \int_{v_s} \frac{\varepsilon_0 E^2}{2}\,dV$$

where v_f, v_s are the volumes of FC and SC, respectively. The second integral is $Q^2/2C$, where Q is the charge of SC. For $P \neq 0$, the electric field energy within FC remains the same but the energy of SC changes: a part of the charge passes to FC. If P and E have the same sign, the charge of the source diminishes. The surface density of the additional charge is P and the energy of SC changes by

$$\frac{(Q-PS)^2}{2C} - \frac{Q^2}{2C} = \frac{PSQ}{C} + \frac{(PS)^2}{2C}$$

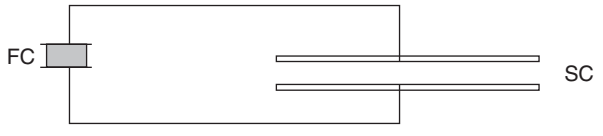


Figure 6 For the calculation of energy of the system consisting of a ferroelectric capacitor (FC) and the voltage source ("large capacitor," SC).

Since $Q \gg PS$, the second term can be neglected and the difference in energies is equal to $-PSV$ or $-PEv_f$. The total energy per unit volume is

$$U(P) = \frac{\alpha P^2}{2} + \frac{\beta P^4}{4} - PE \qquad [2]$$

where $\alpha = aN$, $\beta = bN$. The next step is to minimize $U(P)$ with respect to P to obtain the equation of state: $P = P(E, p)$. However, consider first the generalization of the treatment to the case $T \neq 0$.

Generalization

Experimentally, ferroelectric phase transitions are mainly due to change in temperature, not in pressure. Recall that the equilibrium state at $T \neq 0$ is that of maximum probability. Instead of the probability of states with different values of P ($w(P)$), one can consider a "nonequilibrium thermodynamic potential" $F(P) = kT \ln w(P)$, where k is the Boltzmann constant. The minimum of $F(P)$ corresponds to the equilibrium just as the minimum of $U(P)$ does at $T=0$. Naturally, F is a different function of P at different T and p. Therefore, in eqn [2], U can be replaced by F, and α (and, in principle, β) can be considered as functions of T, p. At $p = \text{const.}$, the phase-transition temperature (T_c) is given by the condition $\alpha(T_c) = 0$. Landau assumed that $\alpha = \alpha'(T - T_c)$ close to T_c. Nontrivially, this natural assumption proved to be incorrect very close to T_c. But as already mentioned, this is of no practical importance for ferroelectrics.

The form of $F(P; E=0, T, p)$ can be obtained without referring to any model but from symmetry arguments. Considering the nonpolar phase, F can be expanded in the Maclaurin series in P taking into account that some symmetry elements (in this model: reflection in the plane) of this phase change the sign of P. The form of the function has to remain unchanged under symmetry transformations, so the series can contain the powers of P^2 only. Incorporating then the voltage source in the same way as above, one obtains

$$F(P; E, T, p) = F_0 + \frac{\alpha(T,p)P^2}{2} + \frac{\beta(T,p)P^4}{4} - PE + \cdots \qquad [3]$$

where the higher-order terms are omitted. For a nonpolar phase, $\alpha > 0$ and changes its sign at phase transition into a polar phase.

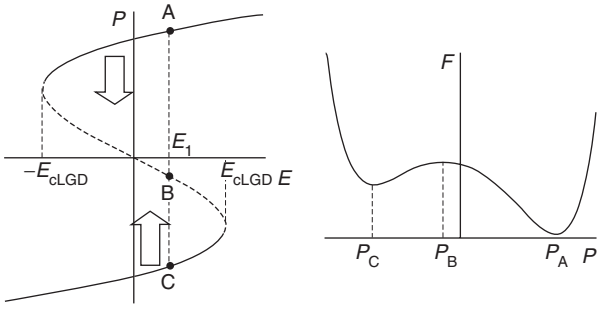


Figure 7 (a) Dependence $P(E)$ according to eqn [4]. The points of the broken line correspond to maxima, not to minima, of the free energy as shown in (b). They present unstable, not realizable states. The thin lines present metastable states, the bold lines present equilibrium ones. This is illustrated in (b), where $F(P, E=E_1)$; E_1 is shown in (a). Points A, B, C correspond to equilibrium, nonstable, and metastable states. Vertical arrows indicate transitions between metastable and equilibrium states.

Equation of State and Switching

Minimizing F given by eqn [3] with respect to the order parameter P, one obtains $P(E)$ as

$$\alpha P + \beta P^3 = E \qquad [4]$$

For $T<T_c$ ($\alpha<0$), $P(E)$ is shown in **Figure 7a**. Transitions from metastable to equilibrium states shown in the figure prevent the system from reaching the "coercive field," E_{cLGD} calculated with the help of eqn [4]. The barrier for these transitions disappears at $E=E_{cLGD}$ but, at any finite temperature, the transitions become possible before disappearance of the barrier. It is not surprising, therefore, that the experimental coercive fields are less than E_{cLGD}. But it is surprising that this difference is very high: two or three orders of magnitude. Escapes from metastable states normally occur locally, similar to the liquefaction of under-cooled vapor. It proceeds because of nucleation, that is, formation of very small drops of water, and subsequent growth of the nuclei. The nuclei arise either because of thermal fluctuations or due to the assistance of some "nucleation centers" as those provided by a jet liner, leading to the formation of the trace visible from the surface. It is natural to expect that a change of direction ("switching") of the polarization occurs in a similar manner: first, small regions (small "domains") of opposite polarization arise and then these regions grow and merge to complete the switching. Some small domains in partially switched crystals were repeatedly observed, but it remains unclear how they do arise. However, what is clear is that only large enough nuclei can grow and not collapse, both in the under cooled vapor and the ferroelectric. The estimations of the "critical dimensions" provide very high values both for the dimensions and the energy of "critical nuclei" in ferroelectrics. It is unimaginable that they arise because of thermal fluctuations at electric fields much lower than E_{cLGD}. A natural assumption is that the nucleation is because of some centers, just as in the case of the jet liner. Unfortunately, one has no idea what these centers are as no model of such a "nucleation-assisting" center has yet been developed. Although the "nucleation and growth model" helps in understanding many experimental data, it has no solid base yet.

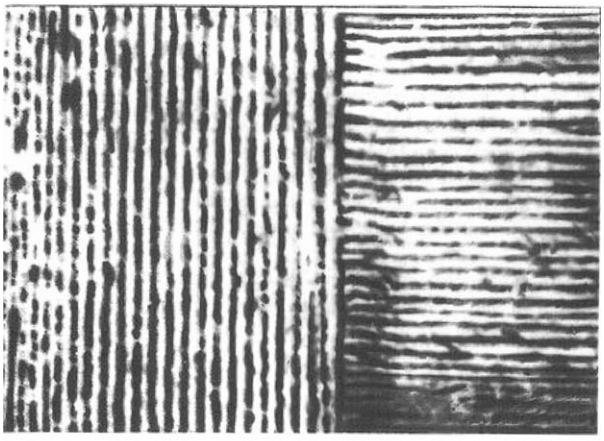
Domains

Domains or regions with different directions of the spontaneous polarization do not arise only in the process of switching. Ferroelectrics normally appear to consist of domains when cooled through the transition temperature (**Figure 8**). Despite the domain structures being diverse, even for the same crystal, there are some regularities, especially in the orientations of the boundaries between the domains – domain walls. Their energy strongly depends on their orientations. In particular, this prevents the walls from being "charged": the spontaneous polarizations in the adjacent domains tend to have equal components perpendicular to the wall. A domain wall normally has positive energy, that is, the formation of domain walls is energetically unprofitable for the crystal. The domain structures present either metastable states, or there are energetic gains associated with their formation. The most evident gain is diminishing of the energy of the electric field associated with any inhomogeneity of the polarization along a direction that is not perpendicular to the polarization. It might, in particular, be due to the boundary of the sample. In a short-circuited plate whose plane is perpendicular to the spontaneous polarization, the domains are not expected to appear if the electrodes are ideal and the sample is completely homogeneous. Experimentally, this is hardly possible, and it is not surprising that the domain structures are almost always observed.

Phase-Transition Anomalies

Phase-transition anomalies or peculiarities of crystal properties near a second-order phase transition have already been mentioned. However, only anomalies obtainable from eqn [3] are touched upon here. For

(a)



(b)

Figure 8 Domain structures of (a) TGS and (b) KH_2PO_4 as observed in a microscope. In both cases, the cross section is perpendicular to the direction of P_s. In the case (a), the crystal surface is covered by a liquid crystal to make the visualization possible. In case (b), the water steam was used as a contrast reagent. ((a) Courtesy of Natalia A Tikhomirova, Universidad Autonoma de Madrid, Madrid.)

$E = 0$, one finds that

$$P_s = 0 \quad \text{for } T > T_c$$

and

$$P_s = \pm(-\alpha/\beta)^{1/2} \quad \text{for } T < T_c$$

This means that $P_s \propto (T_c - T)^{1/2}$.

For small values of E, one can linearize the equation in both the phases. For $T > T_c$, one finds $P = E/\alpha$, that is, for dielectric permeability in a small electric field, $\varepsilon = 1 + dP/\varepsilon_0 dE$, one obtains the Curie–Weiss law, $\varepsilon = C/(T - T_c)$, where $C = 1/\alpha'\varepsilon_0$ is the Curie–Weiss constant. The value of C is an important parameter of a ferroelectric material. Similar temperature dependence, $\varepsilon \propto 1/|T - T_c|$ with another Curie–Weiss constant, takes place for $T < T_c$. One also finds that the pyroelectric coefficient $\gamma = dP_s/dT \propto (T_c - T)^{-1/2}$, so it also tends to infinity as $T \to T_c$ but with a different power behavior. The same is valid for the piezoelectric coefficient $r = dP_s/d\sigma$, where σ is a symmetry-conserving strain.

Dynamic Properties

Consider again the model for $T = 0$. Since P is proportional to the collective shift of ions B, the equation of motion for P in the case of small oscillations in the nonpolar phase reads:

$$\frac{Md^2P}{dt^2} + \alpha P = E(t) \qquad [5]$$

where $M = mN$, and m is the mass of an anion. This is the equation of a harmonic oscillator whose natural frequency is $(\alpha/M)^{1/2}$. The oscillator presents one of the modes of the optic branch corresponding to the vibrations of the anion sublattice with respect to the cation one. The frequency of the mode becomes zero at the phase transition. This is why it is called the "soft mode." Vibrations corresponding to other modes of the branch involve deformations of the springs, and their frequencies remain positive at the transition.

Generalization of eqn [5] to the case $T \neq 0$ is a difficult problem, which has not been solved satisfactorily. A natural and minimalist assumption is to take into account the damping of the oscillator:

$$\frac{Md^2P}{dt^2} + \frac{\xi dP}{dt} + \alpha P = E(t) \qquad [6]$$

where ξ is the viscosity coefficient. If the damping constant ($\Gamma = \xi/2M$) is much larger than the natural frequency, the oscillator converts into a "relaxator," that is, the equation of motion takes the approximate form

$$\frac{\xi dP}{dt} + \alpha P = E(t) \qquad [7]$$

Instead of the frequency, it is the reciprocal relaxation time α/ξ, that goes to zero at the transition. The dynamics described by eqn [6] is considered as typical for the displacive systems, while that described by eqn [7] is typical for the order–disorder ones. The motivation for these terms is explained in the following. Since the two equations represent two extreme cases, it is quite natural that there are many cases where none of the equations is applicable.

This is especially valid for eqn [5]. In fact, a soft mode whose frequency goes to zero at the transition has never been observed. Equation [5] is relevant for high enough frequencies but not close to the transition, if at all. The dynamics of "displacive"

systems close to the transition is still poorly understood.

Displacive and Order–Disorder Systems

These systems can be illustrated as limiting cases of the simple model. At $T=0$, the only possibility for the transition is due to the change of the potential exerted by the cations on the anions. The transition consists in the displacement of the anion sublattice and is of the displacive type. At $T\neq 0$, the phase transition can take place even if the potential does not change at all and remains two-well one. Indeed, at high T, the atoms have high kinetic energy, which allows them to jump between the wells if the springs are relatively soft. For high enough T, the anions find themselves in the left or the right well with the same probability, that is, they become disordered. The average position of an anion is symmetric, that is, the system is in the nonpolar phase. This "order–disorder" transition occurs when the kinetic energy of the ion (kT) becomes comparable with the typical change of the potential energy of the spring when an ion changes the well. The depth of the well is still much larger than kT. A nonequilibrium population of wells in such a system approaches equilibrium in a relaxation process similar to that described by eqn [7]. The phase-transition temperature T_c is the only temperature scale for α, that is, $\alpha' \sim \alpha_{at}/T_c$, where α_{at} is the "atomic" value of α, that is, its typical value far from the phase transition or in an "ordinary" nonferroelectric material. Since the dielectric constants of such materials are of the order of magnitude of unity, one finds that $\alpha_{at} \sim \varepsilon_0^{-1}$ and $C \sim T_c$.

For very rigid springs, such a phase transition is impossible. But if the one-ion potential changes with temperature and becomes a single-well one, the transition will take place once again, as at $T=0$, due to displacement of the anion sublattice. This again is a displacive transition. Of course, the springs should be infinitely rigid to have only displacement and no disorder. Therefore, any real displacive system has an admixture of "order–disorder component," and there exists no boundary between the two types of systems.

The above example of displacive phase transition is good to illustrate the dynamics, but it does not correspond to any real case. A characteristic temperature scale for changes of atomic potentials is T_{at}, and this temperature is much higher than the melting temperatures. It may seem that a displacive transition cannot result by a change in T. However, recall that the Coulomb interaction between the ions has been neglected. The difference between the Coulomb interaction energies in polar and nonpolar structures, called the dipole–dipole interaction energy, is now of much interest. Calculation of this part is fairly involved. What is important here is that the dipole–dipole interaction can be destabilizing for a nonpolar state.

In the model at $T=0$, it can well happen that, because of the dipole–dipole interaction, the polar state is realized even for single-well one-ion potential. In other words, the coefficient α in eqns [2] and [3] now has two contributions: the short-range interactions and the dipole–dipole interactions, $\alpha=\alpha_s+\alpha_d$; $\alpha_s>0$, $\alpha_d<0$. The two contributions are, in principle, temperature-dependent and the scale of this dependence is, of course, T_{at}. If $\alpha(T=0)<0$ but $|\alpha(T=0)| \ll |\alpha_s|, |\alpha_d|$, that is, the effects of the dipole–dipole and short-range interactions are nearly compensated, it can well happen that, at some $T \ll T_{at}$, the balance changes its sign, and the nonpolar state becomes the equilibrium one. Otherwise, the polar state conserves up to the melting temperature, and the crystal is a pyroelectric but not a ferroelectric. Therefore, a system with "rigid springs" can also exhibit polar–nonpolar phase transition because of a change in T. Although the temperature of such a transition is much less than T_{at}, the latter conserves its role as the scale for the temperature changes, that is, in this case $C \sim T_{at}$. This is very different from the order–disorder systems. Note that no dynamics was referred to in the discussion of reasons for the phase transition. It is not surprising, therefore, that ferroelectrics can be fairly well divided into two classes with different orders of magnitude of the Curie–Weiss constant (see **Table 1**), while their division into the order–disorder and displacive systems with respect to their dynamics is far less definitive.

First-Order Phase Transitions: Perovskites

In most ferroelectrics, the phase transition is discontinuous, that is, it is of the first order. In particular, this is the case of ferroelectrics with the perovskite structure (perovskite is a mineral, $CaTiO_3$). Another specific feature of this group of ferroelectrics, which is very important for applications, is that they are cubic in the nonpolar phase. This means that the order parameter of the ferroelectric transition is a three-component one. This can be treated within the Landau theory (as well as the discontinuity of the transition) as long as the structural changes are relatively small. The generalization is straightforward but it is beyond the scope of this article.

Table 1 Some parameters of some ferroelectrics

Substance	*Chemical formula*	*Transition temperature*	*Curie–Weiss constant C* (K)
Potassium niobate	$KNbO_3$	435°C	2.9×10^5
Barium titanate	$BaTiO_3$	130°C	1.5×10^5
Lead titanate	$PbTiO_3$	490°C	1.1×10^5
Antimony sulphide iodide	SbSI	20°C	0.5×10^5
Sodium nitrite	$NaNO_2$	163°C	5.0×10^3
Potassium dihydrogen phosphate	KH_2PO_4	123 K	3.3×10^3
Triglycine sulfate	$(CH_2NH_2COOH)_3 \cdot H_2SO_4$	49°C	3.2×10^3
Rochelle salt	$NaKC_4H_4O_6 \cdot 4H_2O$	−18, 24°C	2.2×10^3

In most perovskite ferroelectrics, the spontaneous polarization arising at transition between nonpolar and polar phases is directed along one of the cubic axes. The crystal symmetry changes from cubic to tetragonal. As a result, the unit cell expands (or shrinks) along the polarization direction losing its cubic form. This symmetry-breaking strain is another aspect of the structural changes at the transition. Without polarization, it would lead to a lowering of the cubic symmetry but less radical than the polarization: a prism with a quadratic base has, as a symmetry element, a mirror plane parallel to the base at half the height of the prism. This symmetry element is absent if there is a polarization perpendicular to the base. This means that the symmetry-breaking strain is a secondary effect of a more profound change of symmetry.

Improper, Liquid Crystalline, and Incommensurate Ferroelectrics

Spontaneous polarization can also be a result of another, more profound change of symmetry. Such ferroelectrics are called improper. They exhibit quite different phase-transition anomalies than the usual ("proper") ferroelectrics, for example, their dielectric constant changes little near the transition and has low values, similar to those of ordinary dielectrics.

In some liquid crystals, the so-called chiral smectics, there exists spontaneous polarization that changes its direction ("winds") if one translates along a special direction in these systems. The winding is periodic with a macroscopic pitch: 5–10 μm. The polarization can be made homogeneous by applying a relatively small electric field.

Another type of spatially inhomogeneous polarization, with changing magnitude rather than direction, is observed in some ordinary (solid) crystals. The period of the inhomogeneity is larger only by one to two orders of magnitude than typical interatomic distances. It changes with temperature independently of the period of structural features that do not contribute to the polarization ("basic lattice"). The crystal has two independent ("incommensurate") periodicities. At some temperature, the polarization becomes homogeneous (within macroscopic domains, as usual), that is, there is a phase transition between the incommensurate and the ordinary ("commensurate") ferroelectric phase. This phase transition belongs to a special class of transitions ("lock-in transition"), which take place not only in ferroelectric but in many other systems, for example, in systems with charge density waves. In ferroelectrics, this transition is usually accompanied by a strong increase of the dielectric constant.

See also: Crystal Symmetry; Dielectric Function; Ferromagnetism; Incommensurate Phases; Lattice Dynamics: Vibrational Modes; Liquid Crystals; Local Field Effects; Piezoelectricity; Point Groups; Pyroelectricity.

PACS: 77.80.Bh; 77.80.Dj; 77.80.Fm; 64.60.Cn; 64.70.Kb; 64.70.Rh

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.

Burfoot JC and Taylor GW (1979) *Polar Dielectrics and Their Applications*. London: MacMillan Press.

Burns G (1985) *Solid State Physics*. Boston: Academic Press.

Kittel C (1996) *Introduction of Solid State Physics*, 7th edn. Chichester: Wiley.

Lines ME and Glass AM (1977) *Principles and Application of Ferroelectrics and Related Materials*. Oxford: Oxford University Press.

Smolenskii GA, *et al.* (1984) *Physics of Ferroelectric Phenomena*. New York: Gordon and Breach.

Strukov BA and Levanyuk AP (1998) *Ferroelectric Phenomena in Crystals*. Berlin: Springer.

Nomenclature

a'	coefficient in the dependence of a on pressure
a, b	coefficients of expansion of the potential energy U in a series of powers of P^2
C	capacitance

$2d$	unit cell length
D	electric displacement
E	electric field
E_c	coercive field
E_{cLGD}	coercive field within the LGD theory without accounting for domains
F	free energy
I	current
k	Boltzmann constant
L	distance between electrodes of a capacitor
m	anion mass (within the model)
M	anion mass density (within the model)
N	number of anions (cations) in a unit volume
N_1	number of anions (cations) in the atomic chain
N_2	number of the atomic chains in a crystal
p	pressure
p_c	pressure of change of the form of the short-range potential energy of anions within the model
P	polarization
P_s	spontaneous polarization
q	anion (cation) charge (within the model)
Q	charge of a capacitor
S	area of the capacitor electrode
T_{at}	atomic temperature
T_c	phase-transition temperature
U	potential energy of an anion due to its short-range interaction with the cations (within the model)
V	voltage
w	polarization probability density distribution
α'	coefficient in the dependence of α on the temperature
α, β	coefficients of expansion of the potential energy U and the free energy F in series of powers of P^2
α_{at}	atomic value of α
α_d	contribution to α due to dipole–dipole interactions
α_s	contribution to α due to short-range interactions
ε_0	permittivity of vacuum
v_f	volume of ferroelectric
v_s	volume of the voltage source considered as a large capacitor
ξ	viscosity coefficient

Ferromagnetism

N Magnani, Università di Parma, Parma, Italy

Introduction

As far as isolated atoms are concerned, only two different magnetic behaviors exist: diamagnetism and paramagnetism. The latter property is only present if the considered atoms possess a nonzero permanent magnetic dipole, and is responsible for the weak attractive force experienced by a paramagnetic substance in an applied magnetic field. However, the most interesting magnetic properties of condensed matter arise when the interaction between the individual elementary moments of each atom is considered. In particular, ferromagnetism concerns the situation where the moments have the tendency to align, one parallel to another, even in the absence of any external magnetic field, thus leading to the possibility of obtaining very large values of the magnetization with very little values of the applied magnetic field (spontaneous magnetization). Only a few elements in the periodic table (Fe, Co, Ni, Gd, Dy, and Tb) order ferromagnetically, but an extremely large number of compounds and alloys do. From the practical point of view, apart from their high values of the magnetic susceptibility and relative permeability, one of the most interesting properties of ferromagnetic compounds is the possibility of retaining a large amount of the induced magnetization when the external field is completely removed. This makes it possible to develop devices such as permanent magnets (i.e., magnetized bodies which produce a constant magnetic field in a given volume of space without the need to continuously supply electrical or chemical energy), memory storage devices, magnetic circuits, etc.

The Molecular Field Approach to Ferromagnetism

At the beginning of the twentieth century, Pierre Weiss interpreted the tendency of ferromagnets to spontaneous magnetization by introducing a "molecular field" responsible for the ordering of the elementary magnetic moments, and he assumed that this field is linearly proportional to the bulk

magnetization (total magnetic moment per volume unit). It may be recalled that the magnetization of a set of identical noninteracting ions, each having total angular momentum J, in a magnetic field H can be expressed as

$$M = ng\mu_B J B_J\left(\frac{g\mu_B JH}{k_B T}\right) \qquad [1]$$

where g is the Landé factor and

$$B_J(x) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}x\right) - \frac{1}{2J}\coth\left(\frac{x}{2J}\right) \qquad [2]$$

is the Brillouin function; the magnetization of a ferromagnet can then be obtained by substituting in eqn [1] the applied magnetic field H with $H + \lambda M$, leading to the implicit equation

$$\frac{M}{ng\mu_B J} = B_J\left(\frac{g\mu_B J}{k_B T}H + \frac{g\mu_B J}{k_B T}\lambda M\right) \qquad [3]$$

In order to investigate the possibility of spontaneous magnetization, the $H = 0$ case is considered; eqn [3] can be rewritten as

$$m = tx = B_J(x) \qquad [4]$$

with

$$x = \frac{g\mu_B J\lambda M}{k_B T}; \quad t = \frac{k_B T}{n\lambda(g\mu_B J)^2} \qquad [5]$$

A graphical solution of eqn [4] is given in **Figure 1**; if $t \geqslant t_C = (J+1)/3J$, only the trivial solution $x = 0$ is present, while otherwise there is also a possible solution with $x \neq 0$. This means that there exists a critical temperature T_C (Curie temperature), below which a ferromagnetic compound can have a nonzero magnetization in the absence of an applied magnetic field. The Curie temperature and saturation magnetization for several ferromagnetic elements are listed in **Table 1**. It is worth noting that eqn [4] can be expressed in terms of J, $m = M(T)/M(0)$ and $t = (J+1)/(3J)\ T/T_C$ only. In the classical limit ($J \to \infty$), $B_J(x)$ is substituted by the Langevin function $\coth x - 1/x$, and the function $m(t)$ is the same for all ferromagnetic materials; the latter statement is often referred to as the law of corresponding states.

Figure 2 shows the temperature dependence of the magnetization measured for nickel (Ni), together with the curve derived from eqn [4] considering a pure-spin moment $J = S = 1/2$. The overall agreement is fairly good, and can be further improved

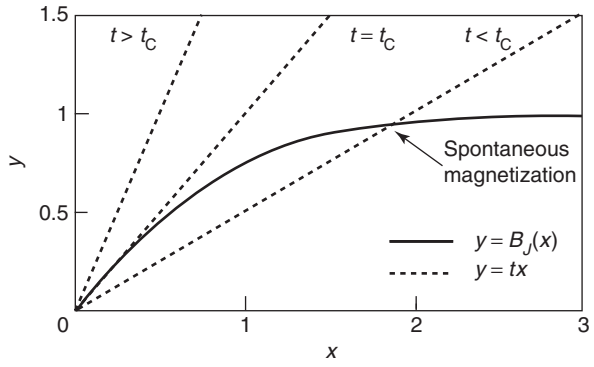


Figure 1 Graphical solution of eqn [4] (see text for details).

Table 1 Curie temperature, saturation magnetization at room-temperature, and magnetic moment per formula unit (at 0 K) of ferromagnetic elements

Element	T_C (K)	M_S (10^6 A m^{-1})	μ_B/*formula unit*
Cobalt	1388	1.42	1.72
Iron	1043	1.71	2.22
Nickel	627	0.48	0.61
Gadolinium	293		7.10
Terbium	221		9.34
Dysprosium	85		10.0

Data from: (1) Kittel C (1962) *Introduction to Solid State Physics*, 7th edn. London: Wiley; (2) Jiles D (1991) *Introduction to Magnetism and Magnetic Materials*. London: Chapman and Hall; (3) Elliott JF, Legvold S, and Spedding FH (1954) Some magnetic properties of dy metal. *Physical Review* 94: 1143; (4) Hegland DE, Legvold S, and Spedding FH (1963) Magnetization and electrical resistivity of terbium single crystals. *Physical Review* 131: 158; (5) Heller P (1967) Experimental investigations of critical phenomena. *Reports on Progress in Physics* 30: 731.

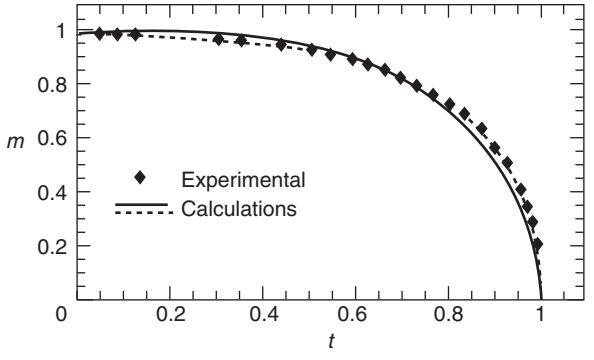


Figure 2 Experimental (diamonds) and calculated (lines) value of $m = M(T)/M(0)$ vs. $t = T/T_C$ for nickel ($J = S = 1/2$). Experimental data are taken from Weiss P and Forrer R (1926) Aimantation et phenomene magnetocalorique du nickel. *Ann. Phys.* 5: 153. The full black line is calculated by means of eqn [4]; dashed lines take into account the spin-wave excitations at low T and the critical behavior near the Curie point.

by the use of more accurate models (the low-temperature decrease due to spin waves and the critical behavior near T_C are shown as dashed lines for comparison).

Above T_C, the sample is in a paramagnetic phase (disordered moments), and its magnetic susceptibility can be calculated by noting that $B_J(x) \simeq (J+1)\,x/(3J)$ for small values of x, leading to

$$\chi \equiv \frac{M}{H} = \frac{C}{T - \lambda C} \qquad [6]$$

where C can be identified with the paramagnetic Curie constant ($\chi = C/T$ if no ferromagnetic order exists, i.e., $\lambda = 0$) and $T_C = \lambda C$. This very simple model gives a fairly accurate description of the experimental magnetic susceptibilities of ferromagnetic compounds at high temperatures.

It must be noted that λM can be as large as several times the maximum magnetic field which can be produced in a standard laboratory, and is about four orders of magnitude larger than the dipolar interaction due to the other magnetic ions in the crystal. The physical origin of this huge molecular field was not clear until the development of quantum mechanics made it possible to point out the role of the exchange interaction (a purely quantum effect and a direct consequence of Pauli's principle) in determining the magnetic spin alignment. The exchange interactions between a collection of N spins can be described by the Heisenberg–Dirac Hamiltonian, which has the form

$$H_{\text{H–D}} = -\sum_{i,j} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \qquad [7]$$

where the sign of the exchange constants J_{ij} depends on whether the two spins labeled i and j are coupled ferromagnetically (positive) or antiferromagnetically (negative), and the sum usually involves only the z nearest neighbors since the interaction strength drops very rapidly as the distance increases. In the presence of an applied magnetic field $\boldsymbol{H}$, the Zeeman term

$$H_Z = -2\mu_B \sum_i \boldsymbol{H} \cdot \boldsymbol{S}_i \qquad [8]$$

must also be considered (a gyromagnetic factor of 2 was taken, in the hypothesis that pure spin moments are being dealt with). The molecular field approximation consists in reducing the N-body Hamiltonian $H_{\text{H–D}} + H_Z$ to N one-body Hamiltonians H_i, by replacing all the spin operators except $\boldsymbol{S}_i$ with their average values:

$$H_i = -\boldsymbol{S}_i \cdot \left[\sum_j J_{ij} \langle \boldsymbol{S}_j \rangle + 2\mu_B \boldsymbol{H} \right] \qquad [9]$$

The case now considered is that of a ferromagnetic body, with all the relevant exchange constants $J_{ij} = J > 0$ for simplicity. Since

$$\boldsymbol{M} = 2\mu_B \sum_l \langle \boldsymbol{S}_l \rangle \qquad [10]$$

is the total magnetization of the sample, eqn [9] can be rewritten as

$$H_i = -2\mu_B \boldsymbol{S}_i \cdot [\boldsymbol{H} + \lambda \boldsymbol{M}] \qquad [11]$$

with

$$\lambda = \frac{zJ}{4\mu_B^2 N} \qquad [12]$$

Equation [11] immediately shows that the mean-field analysis based on the Heisenberg–Dirac exchange Hamiltonian provides an *a posteriori* justification to Weiss' molecular-field hypothesis; as one might expect, the accuracy of this approximation is higher when z and J are larger, since in this case the role of the fluctuations $\boldsymbol{S}_i - \langle \boldsymbol{S}_i \rangle$ is practically negligible and the substitutions made in eqns [7] and [8] are well-grounded.

One word of caution is required before the end of this section. The existing theories of magnetism may roughly be divided into two groups, dealing respectively with localized and itinerant magnetic moments. The Heisenberg approach which was mentioned before belongs to the former; to give a meaning to eqn [7], it must be explicitly assumed that each ion in the crystal carries on it a well-defined magnetic moment $\boldsymbol{S}_i$ due to the electrons which are bound to it. Although this approach is very useful when dealing with certain substances and alloys and for studying critical magnetic behaviors, its validity for metals is questionable due to the presence of conduction electrons, not tied to a specific ion but belonging to the whole crystal. On the other hand, band models considering the role of itinerant electrons have managed to solve several flaws of localized theories, such as the noninteger values of the magnetic moment per atom experimentally measured in most ferromagnets (**Table 1**), which could not be explained within a localized-moment framework.

Magnetostatic Energy and Demagnetizing Field

One may consider a sample of a magnetic substance, composed of n magnetic atoms per unit volume, each carrying an elementary moment $\boldsymbol{\mu}$, and imagine applying a magnetic field which is strong enough to align all these dipoles along the same directions. As a result of this process, the sample will display a saturation magnetization

$$M_S = n|\boldsymbol{\mu}| \qquad [13]$$

(it may be recalled that the magnetization is defined as the total magnetic moment per volume unit).

The energy required to obtain this configuration is

$$E = -\mu_0 \int \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{M} \qquad [14]$$

where $\boldsymbol{H}$ indicates the magnetic field and $\boldsymbol{M}$ is the magnetization vector. In the absence of dissipative processes, an equivalent amount of magnetic potential energy is stored within the sample.

In turn, any magnetized body produces a magnetic field in its surrounding space, as one can immediately witness by means of a compass needle. In close analogy with the charge polarization of a dielectric material, one can describe this field as being generated by "free" magnetic poles (i.e., positive poles not neutralized by the presence of a negative pole in their immediate neighborhood and vice versa) on the surface; in addition to this, the presence of a magnetic field generated by an external source is usually considered. Outside the sample, the magnetic induction $\boldsymbol{B}$ and the magnetic field $\boldsymbol{H}$ are always proportional (as $\boldsymbol{B} = \mu_0[\boldsymbol{H} + \boldsymbol{M}]$ and $\boldsymbol{M} = 0$) and their flux lines are oriented from the north (N) to the south (S) poles. On the other hand, the total magnetic field $\boldsymbol{H}_{\mathrm{int}}$ inside the sample is found by summing the external field, which is parallel to the magnetization, and the contribution from the free poles, which is antiparallel to it (in fact, the flux lines of $\boldsymbol{H}$ always begin on N poles and end on S poles except when the magnetic field is due to electric currents, in which case they are closed and continuous). The free-poles contribution is also called demagnetizing field, and can be represented as

$$\boldsymbol{H}_{\mathrm{d}} = -\underline{N}_{\mathrm{d}} \cdot \boldsymbol{M} \qquad [15]$$

where $\underline{N}_{\mathrm{d}}$ is, in general, a second-order tensor whose components depend on the geometry of the studied system; in some particular cases and/or for special directions of the magnetic field, $\boldsymbol{M}$ and $\boldsymbol{H}$ are collinear, and one can write a simpler relation where N_{d} is a dimensionless number called demagnetizing factor. For example, a homogeneous spherical sample (**Figure 3**) has $N_{\mathrm{d}} = 4\pi/3$, hence

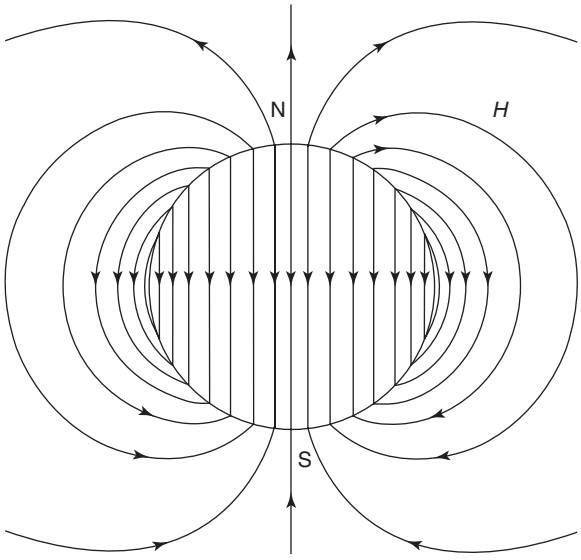


Figure 3 Magnetic field generated inside and outside a uniformly magnetized sphere. The magnetization vector $\boldsymbol{M}$ is directed from the S to the N pole; the resulting demagnetizing field (inside the sphere) is antiparallel to $\boldsymbol{M}$.

$$\boldsymbol{H}_{\mathrm{d}} = -\frac{4\pi}{3}\boldsymbol{M} \qquad [16]$$

Demagnetizing factors for samples of various geometries and for different field directions can be found in specialized textbooks.

Magnetic Anisotropy

Another contribution to the energy balance, that is, magnetic anisotropy, will be dealt with now. While the meaning of such a locution is quite simple (the magnetic behavior of a material depends on the direction), it is not easy to understand the physical reasons behind its very existence. In fact, magnetic anisotropy may be produced by several different mechanisms, which include (but are not limited to) magnetocrystalline (or crystal-field) anisotropy, shape anisotropy, stress anisotropy, annealing in a magnetic field, etc. For the sake of simplicity, the considerations made in the following will be referred to a uniformly magnetized single-crystal sample.

From a phenomenological point of view, the effect of magnetic anisotropy may be taken into account by adding to the energy balance an extra term E_{A} (anisotropy energy), which depends on the relative orientation of the magnetization vector with respect to the crystallographic axes. As shown by the Russian

physicist Akulov, the free anisotropy energy per unit volume can always be expressed by an infinite series

$$\begin{aligned} E_A &= \sum_{n=1}^{+\infty}\left(\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_n=1}^{3} c_{i_1,i_2,\ldots,i_n}\alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_n}\right) \\ &= \sum_{i=1}^{3} c_i\alpha_i + \sum_{i=1}^{3}\sum_{j=1}^{3} c_{i,j}\alpha_i\alpha_j \\ &\quad + \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} c_{i,j,k}\alpha_i\alpha_j\alpha_k \\ &\quad + \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{i,j,k,l}\alpha_i\alpha_j\alpha_k\alpha_l + \cdots \end{aligned} \quad [17]$$

where α_1, α_2, and α_3 are the direction cosines of the magnetization vector with respect to the Cartesian axes x, y, and z, and several coefficients may equal zero for symmetry reasons; for example, all the terms of the sum with odd n vanish if eqn [17] is invariant under inversion of the magnetization. In most relevant cases, eqn [17] can be rewritten as

$$E_A = \sum_{n=1}^{+\infty} K_n A_n(\alpha_1, \alpha_2, \alpha_3) \quad [18]$$

for example, a ferromagnetic crystal with cubic lattice symmetry has $A_0(\alpha_1, \alpha_2, \alpha_3) = 1$; $A_1(\alpha_1, \alpha_2, \alpha_3) = \alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3^2$; $A_2(\alpha_1, \alpha_2, \alpha_3) = \alpha_1^2\alpha_2^2\alpha_3^2$; etc. The coefficients K_n, named anisotropy constants, can be determined by magnetization measurements; a knowledge of a reliable set of anisotropy constants for a given sample can lead to a direct phenomenological interpretation of the magnetic processes and is useful to make comparisons between different systems.

For lattices with uniaxial symmetry (e.g., cylindrical, hexagonal, and tetragonal), compact expressions for the anisotropy energy can be derived as a function of the polar and azimuthal angles θ and φ, which define, respectively, the angle between the magnetization vector $\boldsymbol{M}$ and the z-axis (coincident with the symmetry axis), and the angle between the projection of $\boldsymbol{M}$ on the xy-plane and the x-axis. In this case,

$$E_A = \sum_{n=0}^{+\infty} \sin^{2n}\theta \sum_{0 \leqslant l < n/m} K_n^{(l)} \cos(lm\varphi) \quad [19]$$

As for the cubic phase, the K_0 term is an isotropic energy shift, and will therefore be dropped in the following. The value of m depends on the particular symmetry, leading to

$$\begin{aligned} E_A &= K_1 \sin^2\theta + K_2 \sin^4\theta + K_3 \sin^6\theta \\ &\quad + K_3' \sin^6\theta \cos 6\varphi + \cdots \end{aligned} \quad [20]$$

($m = 6$) for a hexagonal lattice and

$$\begin{aligned} E_A &= K_1 \sin^2\theta + K_2 \sin^4\theta + K_2' \sin^4\theta \cos 4\varphi \\ &\quad + K_3 \sin^6\theta + K_3' \sin^6\theta \cos 4\varphi + \cdots \end{aligned} \quad [21]$$

($m = 4$) for a tetragonal system. For a perfectly cylindrical system ($m \to \infty$), the anisotropy on the xy-plane vanishes and one is left with

$$E_A = \sum_{n=0}^{+\infty} K_n \sin^{2n}\theta \quad [22]$$

If no applied field is present, the sample magnetization will have the tendency to stay along a direction where the anisotropy energy is lowest; this is called in brief "easy magnetization direction" (EMD). This feature will be discussed in detail later, but it may be immediately inferred that the larger the anisotropy, the larger the magnetic field required to rotate the magnetization direction. Therefore, anisotropy is a key property to take into account when designing or choosing suitable magnetic compounds for particular applications.

In the following section, the most important physical mechanisms giving rise to magnetic anisotropy are discussed in brief.

Magnetocrystalline Anisotropy

Exchange interaction in a ferromagnet can be seen as a spin–spin coupling which aligns the spin moments along the same direction. On the other hand, the orbital moments of any magnetic ion in the lattice preferably align along given crystallographic directions due to the crystal-field potential, which reflects the local crystal symmetry. Magnetocrystalline anisotropy is generated by the spin–orbit interaction, which favors mutual alignment between the spin and orbital moments; if this coupling is relatively strong, as for rare-earths and actinides, moment directions which are farther from the crystallographic easy-axis will have a higher cost in energy. Although several types of anisotropy exist, only magnetocrystalline anisotropy is intrinsic to the material.

In terms of anisotropy constants, subsequent terms of eqn [18] will contain increasing powers of the ratio of the spin–orbit energy to that of the crystal field. As the spin–orbit interaction is weak for $3d$ metals, the series is generally truncated after the second or fourth order, and only terms in K_1 and K_2 are included. In principle, this is not possible for $4f$ elements; however, in the case of intermetallic rare-earth–transition-metal alloys, approximate expressions can be found as linear combinations of generalized Brillouin functions, with appropriate

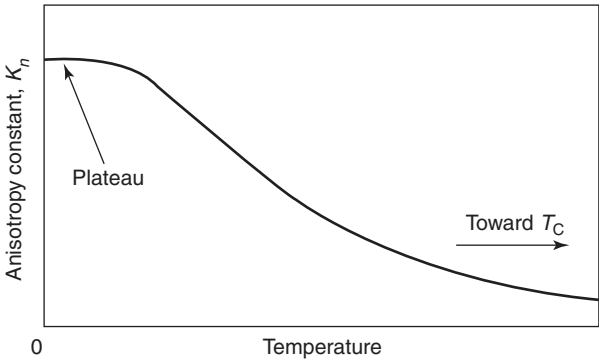


Figure 4 Typical temperature behavior of anisotropy constants (arbitrary units are used).

coefficients related to the decomposition in terms of spherical harmonics of the crystal-field potential at the rare-earth site. On these grounds, it is usually possible to consider $K_n = 0$ for $n \geqslant 4$. The typical temperature dependence of anisotropy constants is shown in **Figure 4**. On changing the temperature, the anisotropy constants of different order can have different relative variations; as a result of this, the EMD of a sample may spontaneously change, giving rise to a spin reorientation transition (SRT) induced by temperature even in the absence of a magnetic field.

Shape Anisotropy

The free poles on the surface of a perfectly spherical and homogeneous sample give rise to a demagnetizing field which is collinear and linearly proportional to the magnetization vector, as described by eqn [16]. In more general terms, however, one must refer to eqn [15] to find the correct value and direction of the demagnetizing field corresponding to a given magnetization vector. The relatively simple case of an elongated rotation ellipsoid with two axes of equal length is considered; for symmetry reasons, one has

$$\underline{N}_{\mathrm{d}} = \begin{pmatrix} N_\perp & 0 & 0 \\ 0 & N_\perp & 0 \\ 0 & 0 & N_\parallel \end{pmatrix} = \begin{pmatrix} N_\perp & 0 & 0 \\ 0 & N_\perp & 0 \\ 0 & 0 & N_\perp + \Delta N \end{pmatrix} \qquad [23]$$

where $\Delta N = N_\parallel - N_\perp$ According to eqn [15] and knowing that all homogeneous bodies delimited by second-order surfaces magnetize uniformly, one can write

$$\boldsymbol{H}_{\mathrm{d}} = -\underline{N}_{\mathrm{d}}\boldsymbol{M} = -N_\perp \boldsymbol{M} - \Delta N(\boldsymbol{M}\cdot\hat{\boldsymbol{z}})\hat{\boldsymbol{z}} \qquad [24]$$

where $\hat{\boldsymbol{z}}$ is the unit vector along the direction of the elongated axis of the ellipsoid. The associated energy can be calculated as

$$E = -\mu_0 \int \boldsymbol{H}_{\mathrm{d}} \cdot \mathrm{d}\boldsymbol{M} = \mu_0 M^2 \left(\frac{N_\perp}{2} + \frac{\Delta N}{2}\cos^2\theta\right) = \mu_0 M^2 \left(\frac{N_\parallel}{2} - \frac{\Delta N}{2}\sin^2\theta\right) \qquad [25]$$

which corresponds to eqn [22], the only nonzero anisotropy constants being $K_0 = \mu_0 M^2 N_\parallel/2$ and $K_1 = -\mu_0 M^2 \Delta N/2$. The extrinsic character of shape anisotropy can be immediately noticed, as the calculated K_1 strongly depends on the shape and bulk magnetization of the sample. While shape anisotropy is usually smaller than crystal-field anisotropy for ferromagnetic rare-earth elements and alloys, they may be of the same order of magnitude for transition metals such as iron and cobalt.

Stress Anisotropy

Applying mechanical stress treatments such as lamination, rolling, or extrusion to a magnetic sample can lead to significant crystal deformations. Stress anisotropy may then be generated due to the magnetoelastic coupling between the lattice and the magnetic moments present in the system (magnetostriction).

Anisotropy Induced by Magnetic Annealing

Heat treatments performed under the Curie temperature of the sample in the presence of an applied magnetic field ("magnetic annealing") may result in the generation of uniaxial anisotropy with respect to the field direction. The induced anisotropy is, in general, rather small, although larger effects may be obtained when a structural transition takes place during the annealing.

Magnetic Domains

As has been mentioned in the introduction, the situation of complete alignment of the elementary magnetic dipoles present in a ferromagnetic substance is energetically favorable due to the exchange interaction. However, it is easy to verify that a piece of iron, taken off the shelves and kept below its Curie temperature, is basically unmagnetized and does not produce any magnetic field. Thermal disorder should obviously not be expected to affect the net magnetization significantly, since the exchange interaction responsible for the alignment of the moments is several orders of magnitude larger than the dipole–dipole interaction in a paramagnet. On the other

hand, if one sticks to the molecular-field approach, one must be aware that the solution $M = 0$, $H = 0$ satisfies eqn [3] both above and below T_C.

In order to gain some physical insight on this point, the energy of a magnetized body, under the effect of an external magnetic field $\boldsymbol{H}$, is considered:

$$E = -\mu_0 \int (\boldsymbol{H} + \boldsymbol{H}_d) \cdot d\boldsymbol{M} \qquad [26]$$

In the simple case that the magnetization is uniform all over the sample and that the demagnetizing field can be expressed as $\boldsymbol{H}_d = -N_d\boldsymbol{M}$, eqn [26] becomes

$$\begin{aligned} E &= -\mu_0 \boldsymbol{H} \cdot \int d\boldsymbol{M} + \mu_0 N_d \int \boldsymbol{M} \cdot d\boldsymbol{M} \\ &= -\mu_0 \boldsymbol{H} \cdot \boldsymbol{M} + \mu_0 N_d \frac{M^2}{2} \end{aligned} \qquad [27]$$

The first term on the right-hand side is the familiar expression for the Zeeman energy, and can be minimized by aligning the magnetization vector with the external field and maximizing its modulus. The second term is the self-energy of the magnetized body due to the free poles, and can be minimized by minimizing the net magnetization. On the other hand, this term also represents the amount of energy which is stored in the external magnetic field generated by the sample; therefore, if the applied field $\boldsymbol{H}$ equals zero, it can be inferred that in the lowest energy state, the magnetic field generated by the sample must be as small as possible. From the microscopic point of view, the energy balance results from a competition between exchange coupling (which lowers the total energy if two spins are parallel) and the dipole–dipole interaction (which raises the total energy if two spins are parallel). As a result, the sample is divided in several macroscopic regions, each one having uniform magnetization, called domains. Within a single domain, the spins are parallel to one another; in turn, spins belonging to two different domains can have very different directions (**Figure 5**) in order to minimize the bulk magnetization.

One may naively expect that the subdivision in domains occurs as in **Figure 6a** (i.e., all the spins belonging to domains A and B are arranged along the easy-axis direction), and that spins belonging to different neighboring domains are antiparallel to one another. This configuration certainly minimizes the anisotropy energy, but at a large cost of exchange energy for the two antiparallel spins situated at the interface. In general, it is less expensive to have the spins arranged as in **Figure 6b**; the 180° spin rotation is performed in several different steps within a region between domains A and B, known as "Bloch wall."

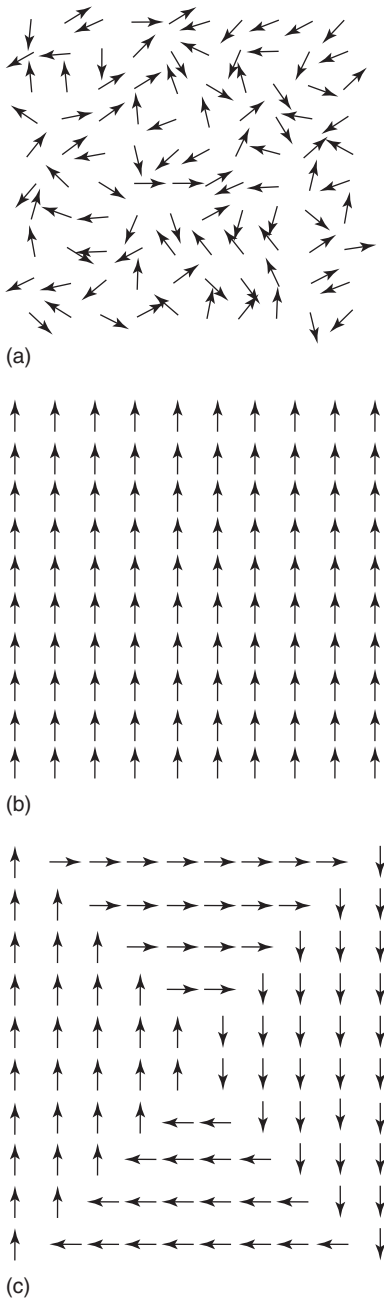


Figure 5 Schematic view of the dipole arrangements in (a) a paramagnet (or ferromagnet above T_C), (b) a magnetized (saturated) ferromagnet, and (c) an unmagnetized ferromagnet (divided into domains). While the moments in the paramagnetic phase are disordered at the atomic level, the domain size is macroscopic (even several hundred micrometers).

For the sake of clarity, consider the case of a simple cubic crystal with lattice constant a, with a nearest-neighbor exchange constant J and EMD along [1 0 0]. The energy balance may be estimated by assuming that the rotation is performed within the xz-plane in N equal steps, each of an angle π/N; then the total exchange energy per unit area of the Bloch

wall is

$$E_{ex} = -J \sum_{l=0}^{N} \boldsymbol{S}_l \cdot \boldsymbol{S}_{l+1} = -J \sum_{l=0}^{N} S^2 \cos\frac{\pi}{N}$$
$$= -JS^2 \frac{(N+1)}{a^2} \cos\frac{\pi}{N} \qquad [28]$$

while, according to eqn [18] with $\alpha_y = 1$, $\alpha_z = \cos\theta$, and $\alpha_x = \cos(\pi/2 - \theta) = \sin\theta$, the anisotropy energy per area unit is

$$E_A = a \sum_{l=0}^{N} K_1 \sin^2\frac{l\pi}{N} \cos^2\frac{l\pi}{N} = \frac{aK_1}{8} N \qquad [29]$$

(strictly valid for $N > k$). E_{ex} and E_A are opposite in sign, and the absolute value of both grows with N; the actual size of the Bloch wall then results from a competition of the two, being smaller when the anisotropy dominates and larger when the exchange interaction dominates. In the large-N limit $E_{ex} \simeq JS^2\pi^2/(Na^2)$, and minimizing the total energy gives an approximate domain wall width of

$$\ell = Na \simeq 2\pi S \sqrt{\frac{2J}{aK_1}} \qquad [30]$$

typically a few hundred lattice constants for iron.

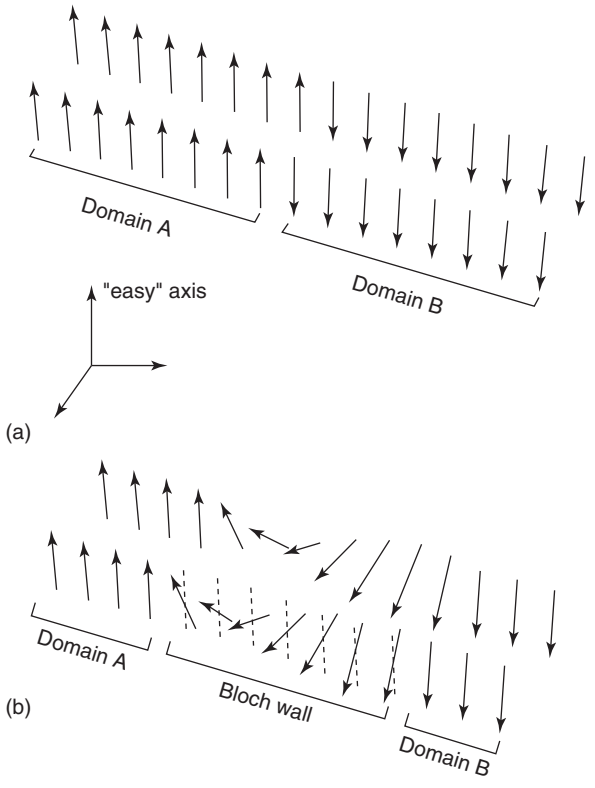


Figure 6 Schematic view of the dipole arrangements at the boundary between two domains (Bloch wall).

The Magnetization Curve

From the macroscopic point of view, ferromagnets may be considered at a first glance as magnetic materials with extremely large susceptibility and permeability. In practice, as both of them are strongly dependent not only on the temperature, but also on the applied field and on the past history of the sample, they are not very useful parameters and it is quite simple to describe the physical properties of a ferromagnet by means of $\boldsymbol{M}$ versus $\boldsymbol{H}$ (or $\boldsymbol{B}$ versus $\boldsymbol{H}$, being $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M})$) plots, whose main characteristics will be discussed in the following sections.

The Hysteresis Cycle

When an external magnetic field is applied to a ferromagnetic sample in a demagnetized state in order to bring its resulting magnetization to saturation, one obtains the magnetization curve shown in **Figure 7**. Two distinct and coexistent processes must be considered in order to understand this (**Figure 8**): boundary motion and domain rotation. In the former case, those domains which are favorably oriented with respect to the applied field grow in size at the expense of the others, without changing their overall magnetization direction; this process is dominant at low fields. At intermediate fields, the process of sudden rotation of the magnetization of unfavorably oriented domains to the easy-axis direction(s) nearest to that of the applied field also becomes significant. At a given value of the magnetic field (labeled H_S), saturation is finally achieved ($M = M_S$, defined in eqn [13]), mainly by coherent rotation of the domain magnetization toward the applied field direction. For $H > H_S$, the $M(H)$ curve is flat since the applied field is already strong enough

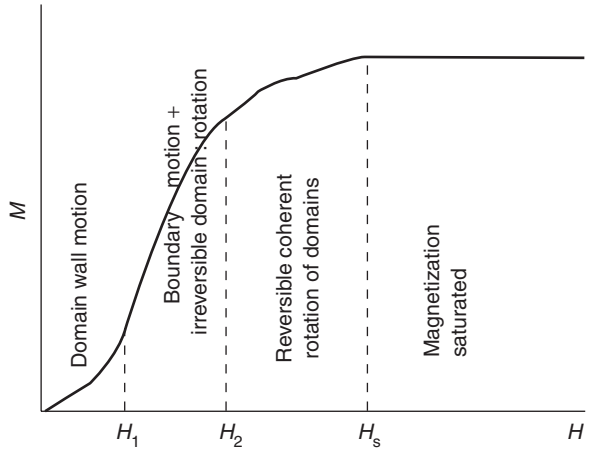


Figure 7 "Virgin" magnetization curve of a ferromagnet. The main mechanisms by which the magnetization process advances for several applied field values are indicated.

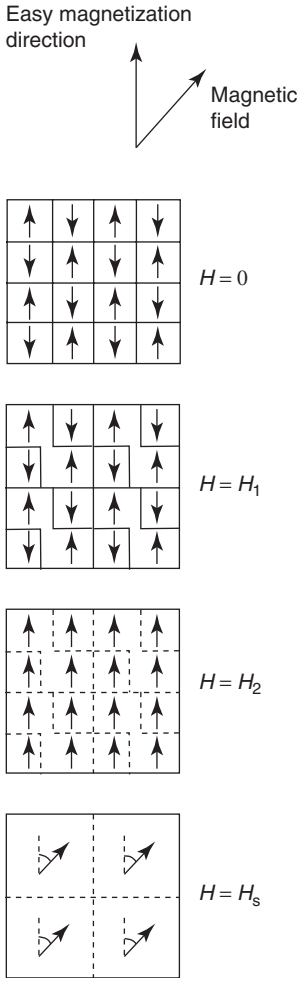


Figure 8 Schematic view of the main mechanisms by which the magnetization process advances for several applied field values.

to align all the elementary dipoles in the sample along its direction.

If the value of the applied field is now reduced to zero, it is noticed that the obtained $M(H)$ curve does not coincide with the previous one apart from a small region near H_S. In particular, the magnetization at zero applied field does not go back to zero (retentivity), and its value M_R is called remanence or remanent magnetization. In order to force the magnetization to zero, it is necessary to apply a magnetic field H_C in the opposite direction; this value is called coercivity or coercive field. (Some authors make use of a distinction between remanence/coercivity (indicating the sample behavior after it has reached saturation) and remanent magnetization/coercive field (relative to a magnetization curve reaching an arbitrary value $M<M_S$).) Raising the reverse field strength to H_S leads to the saturation of the magnetization in the opposite direction as before. Reversing the field once again, one obtains the so-called hysteresis loop (**Figure 9**).

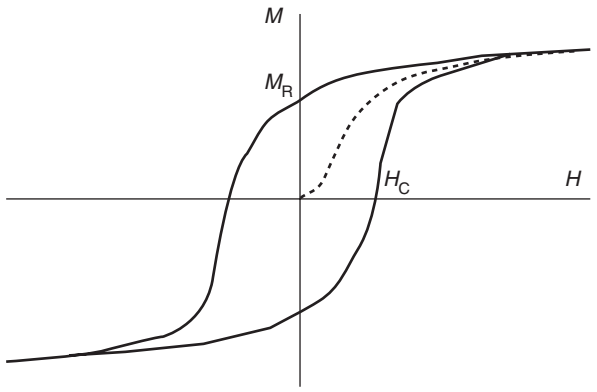


Figure 9 Hysteresis cycle of a ferromagnet. Remanence (M_R) and coercivity (H_C) are indicated on the graph axes.

While "hard" magnetic materials (i.e., those displaying a large coercivity) are generally useful as permanent magnets or for memory-storage purposes (in order to avoid demagnetization due to stray magnetic fields), "soft" ones may be good candidates for the realization of power transformers or electromagnets. In the former case, apart from retentivity, coercivity, and Curie temperature, another figure of merit which is often used in practice is the maximum energy product BH_{max}, which is the maximum value of the product $\boldsymbol{B}\cdot\boldsymbol{H}$ in the demagnetizing quadrant ($M>0$, $H<0$) of the hysteresis curve. This corresponds to the energy stored in the considered material within a magnetic circuit operating at an optimized workpoint, and should not be confused with the energy lost due to the irreversible processes during one hysteresis loop, that is,

$$\oint \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{H} \qquad [31]$$

Magnetization Rotation and Anisotropy

Most of the energy loss during the hysteresis cycle occurs at low fields as a result of irreversible domain boundary motion, due to inhomogeneous microstrains (dislocations) which obstructs the rotation of the magnetic moments. On the other hand, processes involving domain rotation are governed by a competition between the anisotropy and the Zeeman energy $E_Z = -\mu_0 \boldsymbol{H}\cdot\boldsymbol{M}$. To study the role of anisotropy in determining the shape of the magnetization curve, assume that the boundary motion of domain walls can be achieved with a negligibly small magnetic field. In the case of a uniaxial sample, $M(H)$ can be found by minimizing

$$E_A + E_Z = \sum_n K_n \sin^{2n}\theta - \mu_0 \boldsymbol{H}\cdot\boldsymbol{M} \qquad [32]$$

Let θ be the angle between the magnetization vector and the c-axis, which will be considered also as the EMD, for simplicity. If $\boldsymbol{H}$ is applied along the c-axis, both E_A and E_Z are minimized with $\boldsymbol{M} \| \boldsymbol{H}$ and saturation is reached immediately when the external applied field is equal to the demagnetizing field. If $\boldsymbol{H}$ is applied perpendicularly to the easy-axis, differentiating eqn [32] and knowing that $M = M_S \sin\theta$ leads to

$$HM_S = \sum_n 2nK_n \left(\frac{M}{M_S}\right)^{2n-1} \quad [33]$$

an implicit expression for the magnetization curve. By putting $M = M_S$, one can immediately find

$$H_S = \frac{2}{M_S} \sum_n nK_n \quad [34]$$

and

$$\left.\frac{dM}{dH}\right|_{H=0} = \frac{M_S^2}{2K_1} \quad [35]$$

Similar calculations can be made for any other orientations of the field and of the EMD, leading to the remarkable result that, in principle, complete saturation cannot be achieved if $\boldsymbol{H}$ is not parallel to an "extremal" direction (i.e., one which minimizes or maximizes the anisotropy energy).

See also: Diamagnetism; Magnetic Domains; Magnetic Interactions; Magnetic Materials and Applications; Paramagnetism.

PACS: 75.50. – y

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Saunders College.

Bozorth RM (1951) *Ferromagnetism*. Princeton, NJ: Van Nostrand.

Brailsford F (1966) *Physical Principles of Magnetism*. London: Van Nostrand.

Cullity BD (1972) *Introduction to Magnetic Materials*. Reading, MA: Addison-Wesley.

Hubert A and Schaefer R (1998) *Magnetic Domains*. Berlin: Springer.

Mattis DC (1965) *The Theory of Magnetism*. New York: Harper and Row.

Nomenclature

B	magnetic induction (T)
H	magnetic field ($A\,m^{-1}$)
M	magnetization ($A\,m^{-1}$)
T	temperature (K)
μ	magnetic moment (μ_B)

Film Growth and Epitaxy: Methods

F Lévy, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Introduction

The growth of thin films is achieved through many methods resulting from the combination of various steps and devices as:

- source material and its handling,
- transport from the source to the substrate,
- growth substrate and its conditioning, and
- suitable environment in a reactor chamber.

The methods of deposition and growth can be classified according to the source and transport. In the most frequent cases where the material is transported through the vapor phase, two main groups are distinguished: physical vapor deposition (PVD) and chemical vapor deposition (CVD). Besides vapor transport, films can grow from liquid or solid phases.

The key process parameters determine the source, transport, and growth steps as well as the properties of the deposited films. The substrate temperature is the most important parameter. Depending on the method, the pressure, the energy delivered to the growing film (applied power density, electrical bias, bombardment, etc.), the chemistry and the environment considerably influence the growth and properties of the films. The functionality of the film meets the requirements of surface engineering or of sophisticated physical and chemical exigencies in microelectronics and photonics, for example. It leads to the distinction between normal polycrystalline thin films and epitaxial single crystalline films.

The growth of epitaxial thin films by PVD and CVD needs special techniques and equipment. These methods are therefore described in separate sections. The thicknesses of the functional thin films and coatings grown by thin-film techniques, range from a few

nanometers to a few micrometers. Thicker films are generally called thick films, as obtained, for example, by serigraphy or powder technology.

Several characterization and analysis techniques are used to control the film during the growth process such as thickness measurement, growth velocity, structural analysis, chemical composition, etc. The impressive variety and number of deposition methods express the absolute necessity of tailoring the surface properties of a material to meet the requirements of functionality.

Film Growth by Physical Vapor Deposition

The common feature of the PVD methods is the vapor-phase transport of the source material which constitutes the film material. Reactive processes can induce chemical reactions during the transport or at the surface of the growing film.

Evaporation

Thermal evaporation methods include the simple case of emission at the heated source of the material which condenses on the cold substrate. The source can be a refractive wire or a crucible electrically heated and containing the material to be evaporated (**Figure 1**). The evaporation rate is determined by the vapor pressure of the source material at the evaporation temperature. The Joule heating of typical tungsten or molybdenum crucible allows reaching 2800 K, which provides appreciable vapor pressures for most metals. A deposition chamber under high vacuum ($<10^{-2}$ Pa) is necessary in order to avoid oxidation of the source. To prevent the harmful reaction between source and the crucible, the heating can be done rapidly and the process is known as flash evaporation. The properties of the deposited films are usually different if thermally or flash evaporated. Inductive heating of the crucible is also possible.

The thermal evaporation of highly sensitive materials, in particular in ultrahigh vacuum (UHV) ($<10^{-6}$ Pa), is achieved by electron-beam (e-beam) heating (**Figure 1c**). The main advantages of e-gun heating is to prevent contamination by the crucible and the high evaporation temperature reachable. Electron-beam sources are generated either by thermoionic emission or by field effect cathodes.

Every evaporation process can be assisted by plasma, microwave, or electron resonance, in order to influence the deposition rate or the properties of the species condensing on the substrate.

Sputtering

The sputter deposition method includes a large number of particular techniques, from classical cathode sputtering to more sophisticated reactive magnetron sputtering, for example. The material is extracted from the source (cathode) by bombardment with accelerated positive ions of inert gas such as argon, transmitting energy and momentum to the target atoms. In the reactor, the ions are created in a DC or RF plasma sustained at pressures between 0.1 and 10 Pa. For the diode-sputtering reactor, the target and the substrate face each other (**Figure 2**). Other configurations are possible, such as confocal multitargets, ion beam sputtering or triode sputtering, each with specific advantages. Reactive sputtering takes place when a chemically active gas is mixed in the plasma, such as nitrogen or oxygen.

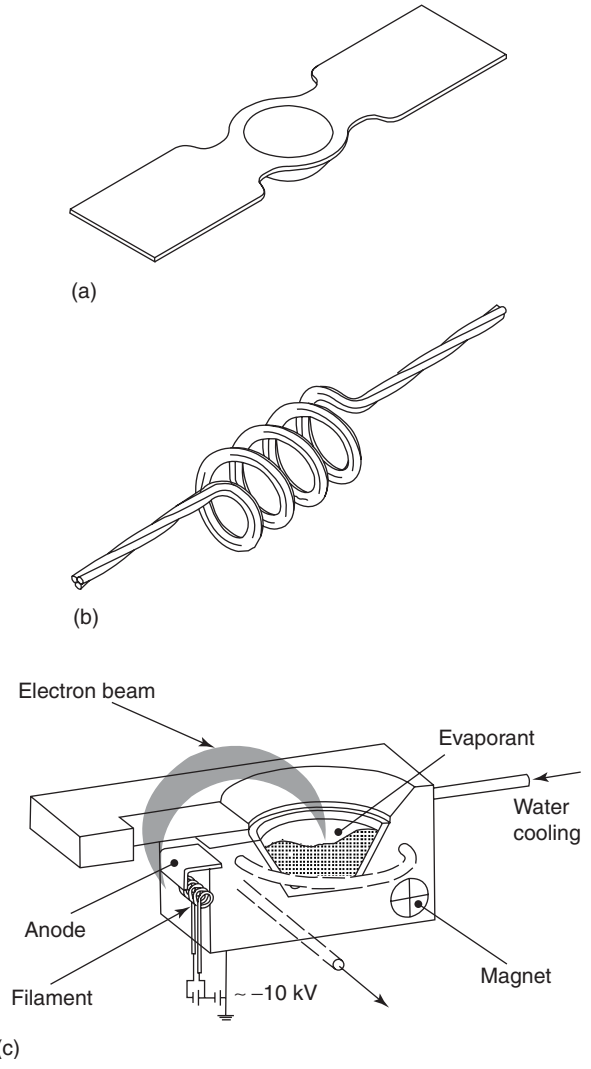


Figure 1 Thermal evaporation: basic design of evaporation sources; (a) boat crucible, (b) twisted wire, and (c) electron beam evaporator.

Magnetron sputtering One of the most powerful sputtering techniques includes permanent magnets or

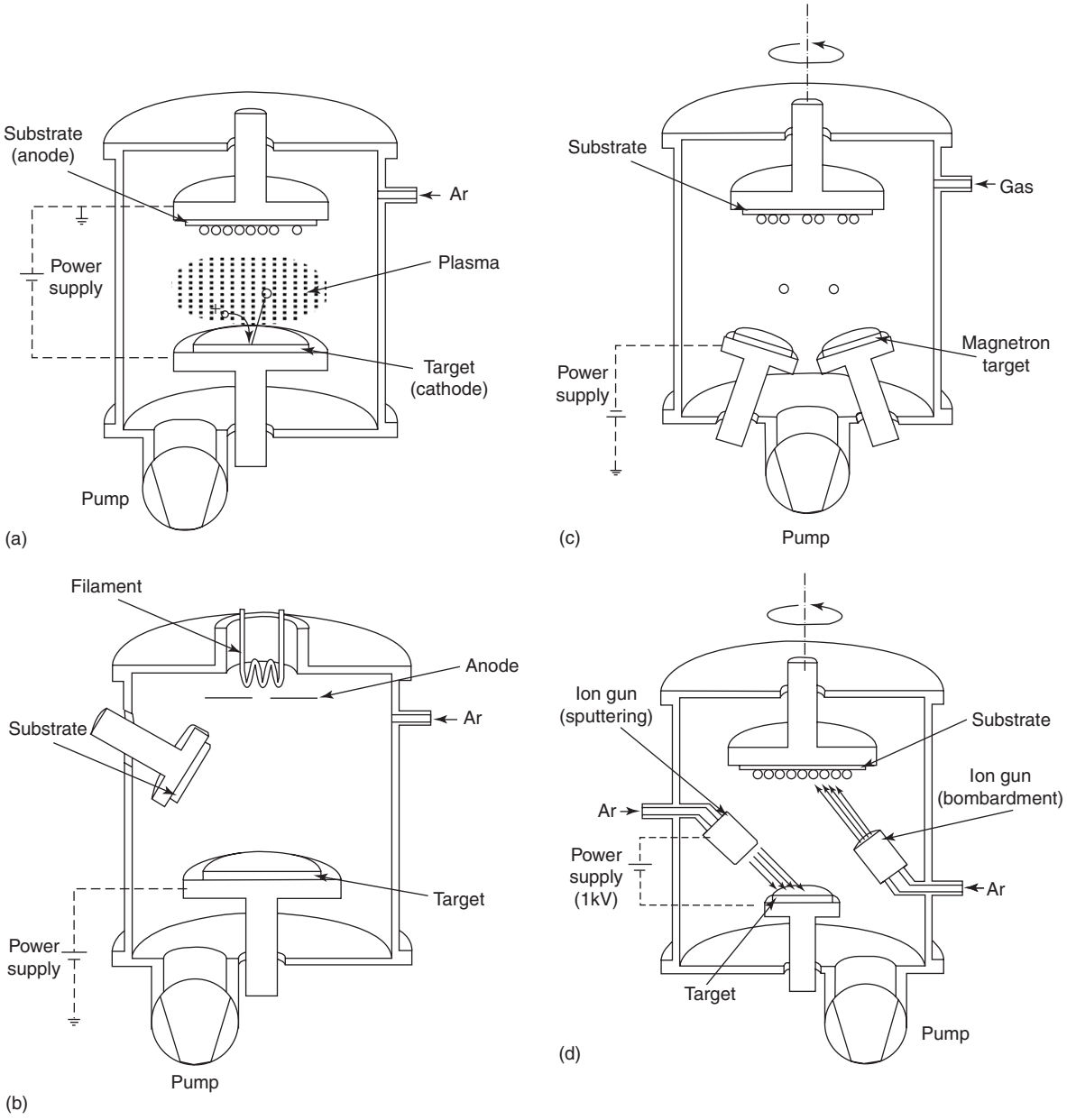


Figure 2 Reactors for PVD; (a) planar diode sputtering, (b) triode sputtering, (c) Two-source confocal magnetron sputtering, and (d) ion beam sputtering or IBAD.

coils in the target holder. The resulting magnetic field bends the secondary electron paths and concentrates the ion sputtering on specific domains on the target. The increased ionization rate results in high-deposition rates, large-deposition areas. The low substrate heating impressively extends the application ranges of magnetron sputtering to substrates with low melting temperature. Shaped cathodes allow coating the surface of the most irregular substrate, in particular, hollow vessels and tubes.

In the normal planar diode magnetron configuration, the magnetic confinement encloses the discharge near to the cathode. There is no current on the substrates. The discharge potential remains lower ($\sim 500\,\mathrm{V}$) and the work pressure of $\sim 0.1\,\mathrm{Pa}$ increases the discharge current and the deposition velocity. Several high-performance magnetron sources have been developed.

It can be profitable to keep the magnetic field open, including the substrate. This process of unbalanced magnetron sputtering allows increasing the plasma density near the substrate. Resulting currents on the substrate are then significantly higher than the normal current density $\sim 1\,\mathrm{mA\,cm^{-2}}$ produced by a

substrate bias of -30 to -100 V. The process becomes ion assisted similar to ion plating.

Cathodic-Arc Deposition

The cathodic arc deposition takes place between two metallic electrodes in vacuum (≈ 1 Pa). A cathodic arc is a low-voltage, high-current plasma discharge. The process can be made reactive by introducing a reactive gas, for example, nitrogen, in the argon plasma. The arc current is concentrated at the cathode at discrete sites called cathode spots ~ 1–$10\,\mu$m in size. The current carried by a cathode spot reaches 1–10 A. A typical arc discharge current of a few 100 A gives rise to a dense plasma of cathode material. The equipment extends from a few millimeter cathodes to typically large plants of a meter size with suitable power supplies (**Figure 3**).

The extracted material can be in the form of atoms or molecules, or metallic droplets rapidly solidified, 0.1–10 μm in diameter. This process is one of the most industrially used techniques and is suitable, for example, to TiN coatings.

Ion Plating

The definition of ion plating does not specify the source. It refers to the substrate and the growing film originally negatively polarized. The process applies in evaporation or sputtering to improve the film adhesion or influence the morphology. It provides a high energy flux to the surface, increasing the surface temperature which enhances the surface diffusion and the possible chemical reactions. It can also increase the defect concentration, at the interface and on the surface.

A variation of the technique is the "ion beam assisted deposition" where the bombardment originates from an independent ion source (ion or plasma gun) in the reactor (**Figure 2d**).

Laser Beam Ablation

Laser beam ablation or pulse laser deposition (PLD) results from the interaction of laser light with a solid. The energy of the absorbed light turns into heat and melts the solid at the surface, which evaporates (**Figure 4**). The beam light through the vapor is absorbed producing a plasma. The plasma ions can reach energies up to 1000 eV. Filters between the substrate and the source are necessary to avoid the splashing effect of liquid droplets reaching the substrate. Compounds of complicated composition such as superconducting or ferroelectric oxides, and incongruently melting solids can be deposited by laser ablation with the correct original chemical composition.

Thermal-Spraying Processes

Thermal spraying consists in melting and projecting particles (droplets of diameter $\sim 10\,\mu$m) onto a substrate, where they spread out and form a thick coating (several 10 μm thick) by progressive stacking. The material in the form of wire or powder has to melt congruently. The necessary source of energy must be powerful enough to melt, or, at least, soften the material (heat transfer) and accelerate the droplets (momentum transfer). Depending on the spraying energy source, several processes exist (**Table 1**).

- In the flame-spraying process, the powdered metal is melted in a combustion flame which expands and sprays the droplets (**Figure 5a**)

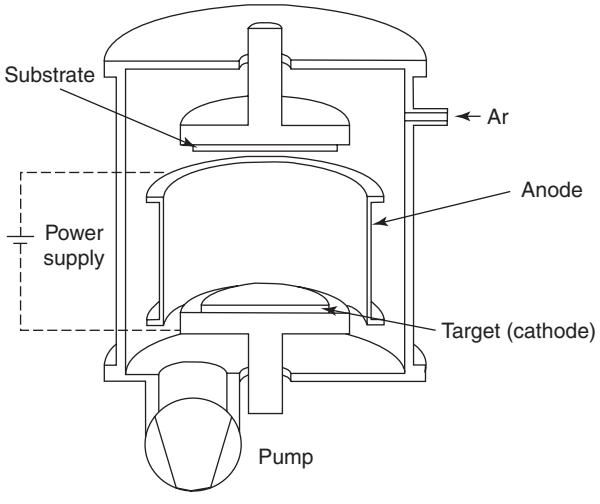


Figure 3 Reactor for cathodic-arc deposition.

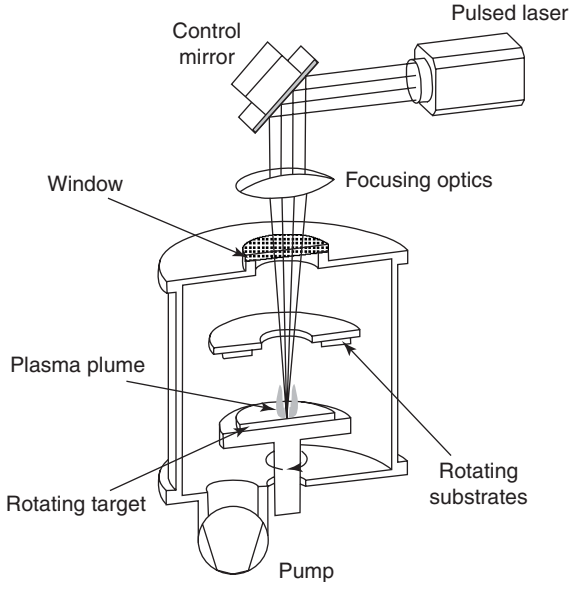


Figure 4 Reactor for laser beam ablation or PLD.

Table 1 Comparisons of different methods of thermal spray

Process	*Flame of plasma exit temperature* (°C)	*Substrate temperature* (°C)	*Particle impact velocity* ($m\,s^{-1}$)	*Porosity* (%)	*Adhesion (bond strength)*
Detonation gun	3000	20–150 max	800–1000	0.1–1	Extremely high
Plasma jet (HVOF)	2500–3100	500–700	500–800	1–10	Very high
Plasma spraying	5500–8300	700–1000	200–600	1–10	Very high
Wire arc	4000–6000	500–800	240	10–20	High
Flame spraying	2500–3000	500–700	30–180	10–30	Low

After A-Flame corp. Cincinnati, Detonation Coating System.

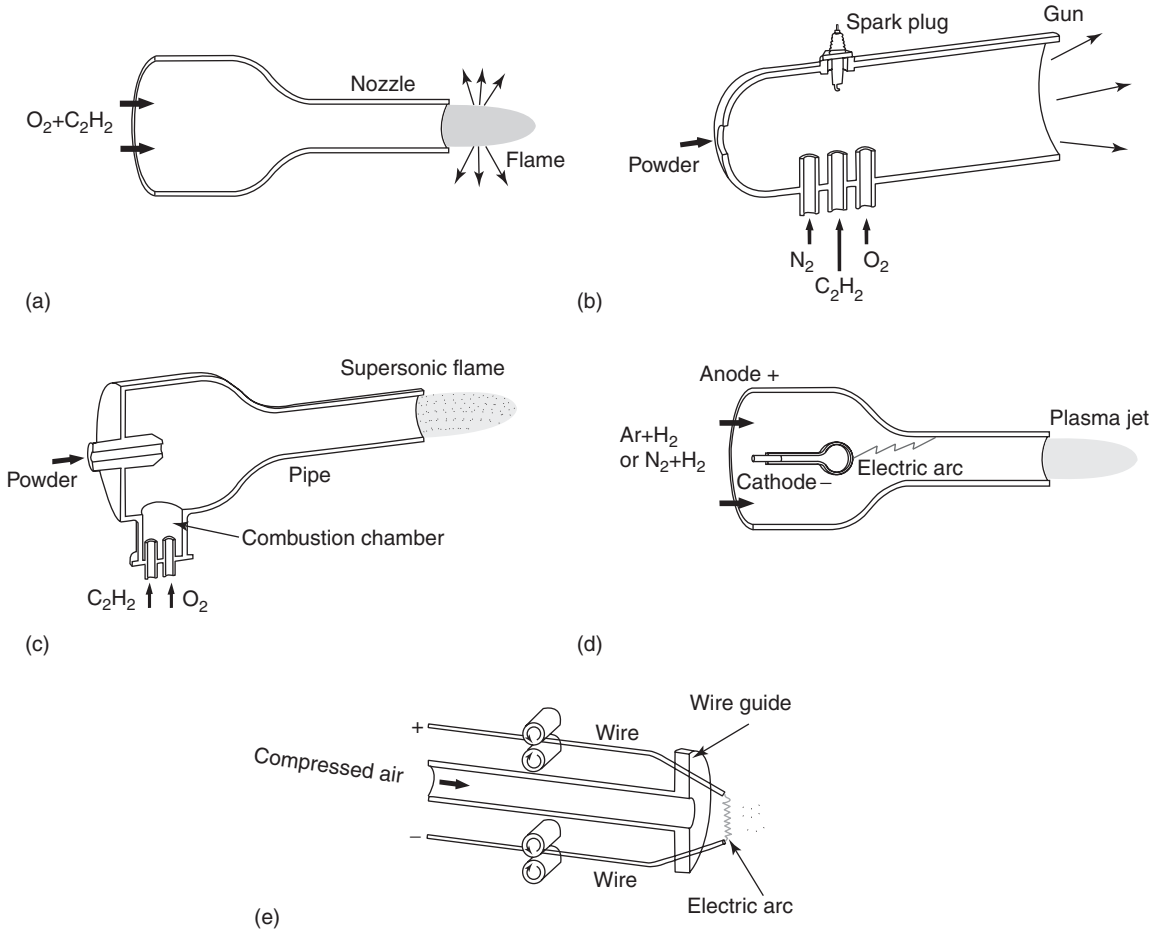


Figure 5 Thermal-spraying processes: (a) flame spraying, (b) detonation gun spraying, (c) HVOF spraying, (d) arc plasma spraying, and (e) wire arc or electric-arc spraying.

- In the gas-detonation process, the powder is injected into a gun with a gas mixture (oxygen and acetylene), which is ignited by a sparkle at a frequency of $\sim$10 Hz (**Figure 5b**). This process can maintain the substrate below 150°C.
- The high-velocity oxy-fuel (HVOF) process is a continuous combustion flame spraying under pressure. The gas temperature and, especially, the jet velocity are increased (**Figure 5c**).
- In arc plasma spraying, the powder is introduced in a low-voltage, high-current plasma. High temperature is reachable so that refractory materials can be deposited (**Figure 5d**).
- The electric-arc spraying applies to metallic wire electrodes, which are molten by the arc with the help of a compressed-air jet and projected on the substrate. The wires are continuously fed near to their intersection point (**Figure 5e**).

Thick coatings deposited by thermal-spraying processes are usually dense (up to 99.5% of theoretical density), hard, with low porosity and high adhesion.

Film Growth by Chemical Vapor Deposition

In CVD processes, the source materials generally in the form of vapors or gases are projected onto the heated substrate and react chemically to produce the expected deposit. In principle, the chemical reaction is irreversible. In order to avoid too high a deposition temperature, complementary activation processes are applied such as plasma, light, and microwave. The properties of the films depend on the precursor and the activation process.

Pyrolysis

Polycrystalline silicon (polysilicon) is deposited by pyrolyzing silane (SiH_4), for example, at 600–650°C. The deposition temperature depends on the precursor. Originally, the use of trichlorosilane ($SiHCl_3 + H_2 \rightarrow Si + 3HCl$) needed 1000–1200°C. Such a temperature is compatible with the silicon growth on hot silicon slim rods. It is less suitable to epitaxial growth or microelectronics technology. Pyrolysis also applies to the deposition of refractory metals such as tungsten and tantalum.

Chemical Vapor Deposition

The process applies to the deposition and the growth of compounds such as SiO_2, Si_3N_4, TiN, and TiCl:

$$\text{Oxides:} \quad SiCl_2H_2 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2HCl$$
$$\text{or} \quad SiH_4 + O_2 \rightarrow SiO_2 + 2H_2$$

$$\text{Nitrides:} \quad 3SiH_4 + 4NH_3 \rightarrow Si_3N_4 + 12H_2$$
$$\text{or} \quad TiCl_4 + N_2 + 4H_2 + Ar \rightarrow 2TiN + 8HCl + Ar$$

$$\text{Carbides:} \quad TiCl_4 + CCl_4 + 4H_2 + Ar \rightarrow TiC + 8HCl + Ar$$

The processes evolve at atmospheric pressure or low pressure in various reactors (**Figure 6**). One of the most impressive applications of CVD is the

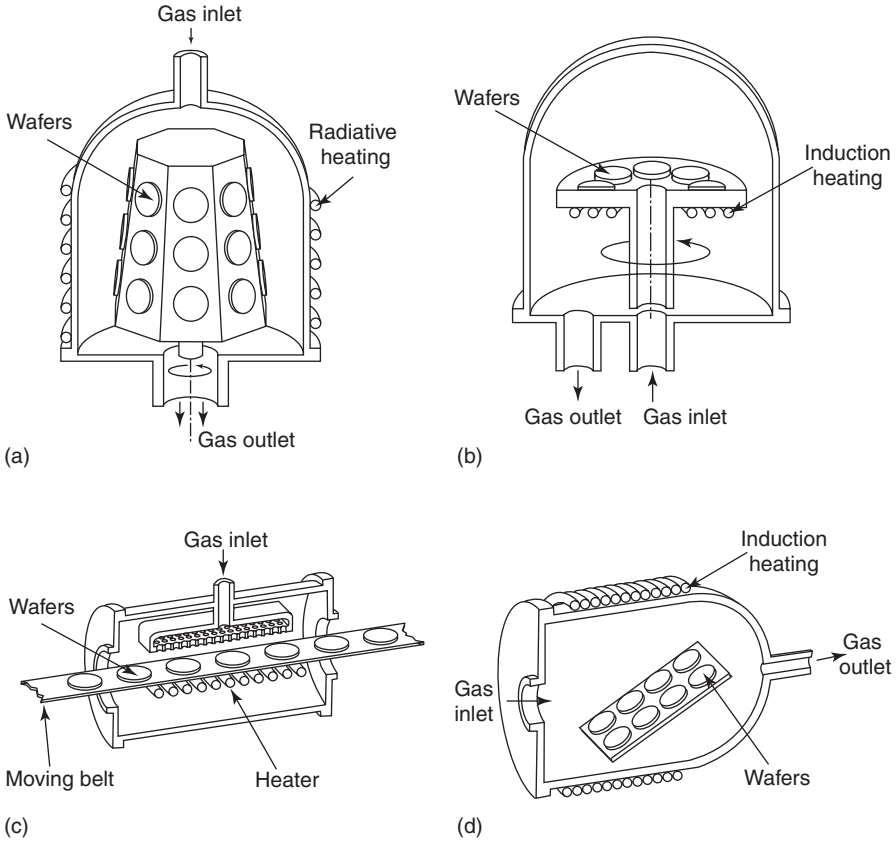


Figure 6 Reactors for vapor-phase epitaxy; (a) barrel reactor, (b) bell jar reactor radial flow, (c) continuous CVD reactor, and (d) horizontal lateral flow reactor.

vapor-phase epitaxy from $SiCl_4$, of silicon layers on Si wafers for microelectronics.

The thermal activation of the chemical reactions up to 1400°C imposes limitation to the heated substrate. Because of the temperature homogeneity in the reactor, the deposited films are also homogeneous and conformal. Geometries with the gas flow perpendicular to the surface of the substrate generally provide faster growth velocities (**Figure 6b**).

Organometallic Chemical Vapor Deposition

The thermal instability of organometallic compounds is advantageous for low-temperature decomposition and reaction. Used as precursors, these organometallics extend the chemical deposition process to many substrates. Care has to be taken regarding the toxicity of most of them. Besides, some properties (e.g., adhesion and stability) of the deposited coating can change. For example, the deposition of gallium arsenide from trimethylgallium and arsine precursors involves the reaction

$$(CH_3)_3Ga + AsH_3 \underset{H_2}{\rightarrow} GaAs + 3CH_3$$

The method is principally used for epitaxial growth, as described later.

Laser Chemical Deposition

The thermal activation of the chemical reaction can be achieved locally by irradiation with a laser beam (continuous wave or pulsed). The light beam can also excite the precursor molecules and increase the reactivity of the gas species.

Plasma-Enhanced Chemical Vapor Deposition

Low-temperature deposition ($\sim$25–400°C) is possible with a plasma process induced by microwaves or RF fields. Active species are produced in the gas phases either in the reactor or at the surface of the substrate. Besides the extension to nearly every kind of substrates, the deposition velocity increases.

Microwave-Assisted Deposition

Microwave energy (at industrial frequency of 2.45 GHz) can be beamed into a deposition reactor through a dielectric window. With a magnetic field ***B*** parallel to the direction of the microwave beam, the electrons describe spiral trajectories and gain energy due to the phenomenon of electron cyclotron resonance (ECR) (**Figure 7**). This ECR-activation contributes to sustaining any plasma by increasing the local plasma density by typically two orders of magnitude. Increased, directed, and focused deposition also results with suitable geometries and ***B*** values.

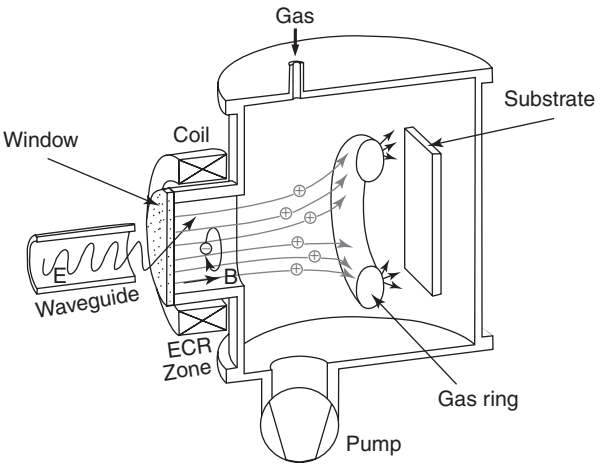


Figure 7 Microwave-enhanced plasma deposition with an ECR zone.

Microwave assistance also applies to every physical evaporation source (PVD) instead of gas supplies.

Hot Wire Chemical Vapor Deposition – Catalytic Chemical Vapor Deposition

The technique of catalytic chemical vapor deposition (Cat-CVD) is based on the decomposition of precursor process gases by a catalytic reaction at the surface of a hot catalyser wire. It applies to large-area substrates (1 m size) and high-deposition rates (2–30 nm s^{-1}). Silicon-based thin films have been deposited such as Si:H, poly-Si, SiN_x from silane or Teflon-like films from hexafluoropropylene oxide (HFPO), for example. The process gas is introduced in a showerhead equipped with catalyzer wires heated at 1300–1800°C (**Figure 8**). The wires or ribbons arrays can be tungsten or nichrome filaments. The deposition generally occurs in the pressure range 1–10 Pa.

Thermoionic Arc Ion Plating

Anodic arc deposition techniques (**Figure 9**) are considered among CVD methods since they are often operated reactively. They allow evaporating refractory metals with high melting points and low vapor pressure such as titanium. An arc current of a few 100 A (50 V) melts the metal in the crucible (anodic source). The ionized metallic vapor, in the reactor, progressively substitutes for the argon in the anodic arc and serves as a deposition source. Reactive gases introduced in the reactor involve chemical reactions, plasma or thermally (heated substrates) activated. The process is known as activated reactive ion plating. The anodic arc deposition becomes a true CVD technique, if the source is just the high-current DC plasma constituted by precursor reactive gases

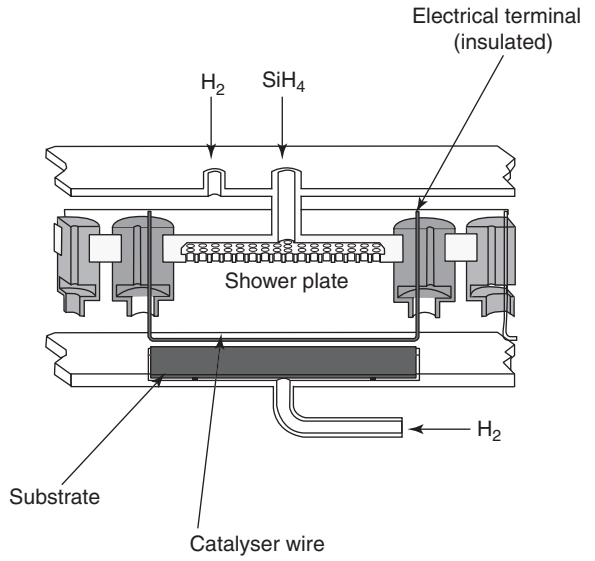


Figure 8 Cat-CVD.

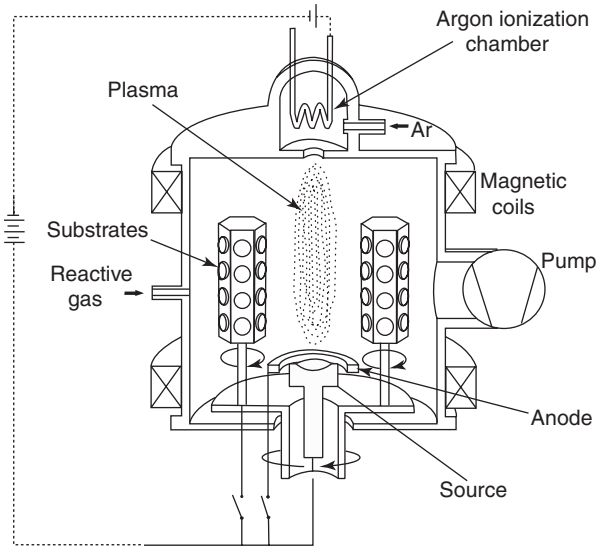


Figure 9 Thermoionic arc ion plating.

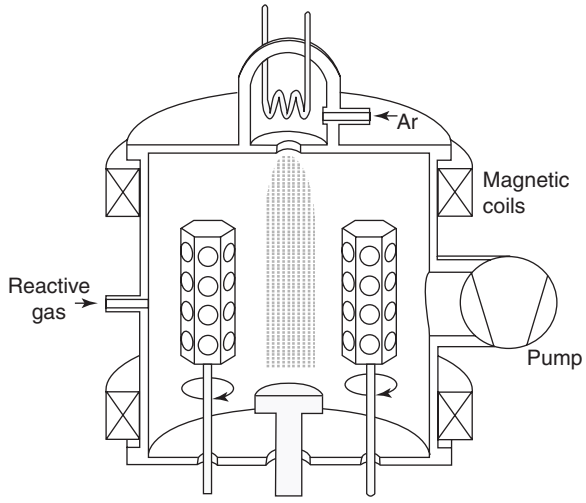


Figure 10 Reactor with a high-current DC plasma source.

(**Figure 10**). Using argon, hydrogen, and methane, this low-pressure process appears particularly efficient for the deposition of diamond coatings. Ammonia offers the possibility of nitriding, in particular with diborane and silane, to produce boron or silicon nitride, respectively.

Film Growth by Condensed-Phase Reaction

Sol–Gel Deposition

The sol–gel deposition process includes synthesis techniques to produce a large quantity of nanosized material with modeled and controlled particle size, morphology, orientation, and crystal structure as well as optimized physical and chemical properties. It applies to the fabrication of metal oxide thin-films including chemical reaction, nucleation, growth, and aging, as an aqueous thin-film processing method (aqueous chemical growth). Tuning the thermodynamics and kinetics of nucleation, growth, and aging, and the precipitation conditions provides a simple and efficient control of the particle size and its distribution. The requirement of minimum surface energy drives nanoparticles to develop spherical morphology. However, in suitable synthesis (precursors) and dispersion conditions, the shape can be driven by the crystal symmetry and the chemical environment. Most process parameters influence the interfacial tensions which appear to play a prominent role. The control of the crystal structure is related with the nucleation kinetics. The more stable phase appears to be more sensitive to secondary growth after heteronucleation, such as the anatase–rutile transition.

The low-cost technique consists of heating the aqueous solution of metal salt or complex at given pH and temperature ($<100^{\circ}C$) in a closed bottle including substrates. Organic solvents or surfactants can also be used with the disadvantages of safety hazards, toxicity, and purity of the deposited film. For example, coatings of crystalline nanoparticulate of various textures have been grown with magnetite (Fe_3O_4), hematite (Fe_2O_3), CrO_3, Cr_2O_3, ZnO, Mn_2O_4, or RuO_2.

Langmuir–Blodgett Films

The film-deposition method known as Langmuir–Blodgett (LB) consists in transferring a floating

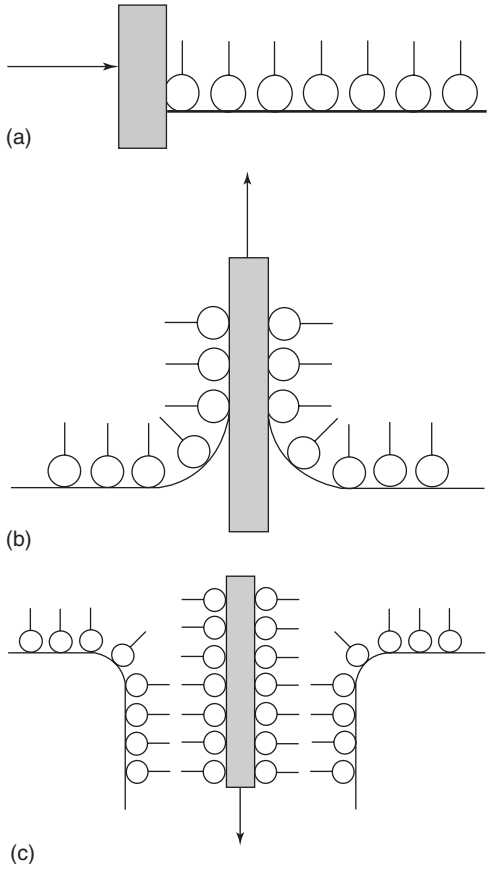


Figure 11 Schematic LB film deposition: pulling and dipping; (a) compression of the monolayer, (b) raising (hydrophilic substrate), and (c) dipping (second layer).

monolayer on water, to a solid plate (substrate). It applies to molecular structures characterized by a hydrophilic head group in contrast with a hydrophobic tail. Stearic acid ($C_{17}H_{35}$ CO_2H) is a classical monolayer forming molecule with (–CO_2H) which is easily soluble in water and a long alkyl chain ($C_{17}H_{35}$–) which is hydrophobic or oleophilic. Various molecular structures (fatty acids, polymerizable materials, aromatic compounds) have been developed with incorporation of metal ions sometimes. The principle of the transfer is illustrated by the extraction or dipping of the substrate through the surface monolayer on the water (**Figure 11**). Simple and sophisticated troughs have been developed to achieve every inventive deposition (**Figure 12**), including surface pressure measurements and displacements control.

Chemical-Bath Deposition

The chemical-bath deposition process uses a controlled chemical reaction to lead the deposition of a thin film by precipitation. In practice, the substrates are

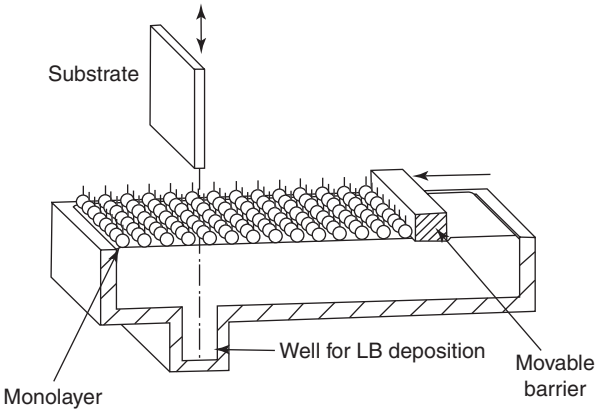


Figure 12 Example of LB trough.

immersed in an alkaline solution containing, for example, the chalcogen source, the metal ion added base, and a complexing agent. In the chemical-bath deposition of ZnS thin films, ammonia and hydrazine have been used as the complexing agents, for example.

Electrodeposition

The essential characteristic of the electroplating technique consists in the passage of an electric current between an anode (positive electrode) and a cathode (negative electrode) immersed in an electrolytic bath. The cathode is the part to be plated. At the anode, the metal is dissolved and then transported to the cathode. If the metal is insoluble in the plating solution, a soluble compound in solution constitutes the electrolyte, such as copper sulphate. Dissolved in an aqueous medium, $CuSO_4 \cdot 5H_2O$ ionizes. The copper cations are attracted to the cathode, whereas the negatively charged sulphate radicals reach the anode. The presence of sulfuric acid increases the conductivity of the electrolyte and permits the decrease of the electrode voltage of the order of a few 10 V. The growth velocity of the deposited film depends on the current density, the plating time, the cathode efficiency, the distance between anode and cathode. The important metals electrodeposited from aqueous baths are chromium, copper, gold, silver, nickel, tin, and zinc, including alloys. Besides metals, ceramics, epoxy resins, composite functional coatings are deposited on conductive substrates. The application fields include printed circuit boards, automobile bodies, rotating shafts, jewelry or electrical contacts.

The equipment include DC power supplies, eventually pulsed generators, and plating tanks either in polyvinyl chloride or lined steel. Automatic continuous processes are common for strip and large-objects plating. Small pieces are electroplated in perforated tumbling barrels.

The related "anodizing" process refers to the deposition of an oxide film on a metal polarized as an anode in an electrolytic aqueous solution. It is mainly used for aluminum. The oxide results from the reaction at the anode of dissolved aluminum, with oxygen: $2Al + 3O^{2-} \rightarrow 6e^- + Al_2O_3$, whereas hydrogen is produced at the cathode. The film consists in a thin barrier below a porous film which supports the electric current by the electrolyte in the pores.

Electroless plating or chemical deposition does not need the passage of an electric current. It is based on a chemical exchange reaction, for example, in immersion tin flashing of aluminum. The oxidation reaction of aluminum, dipped in a $NaS_2SnO_3 . 3H_2O$ aqueous solution, provides three electrons which reduce tin in $HSnO_3^-$. This reduction consumes three electrons as at a cathode. The global reaction leading to a tin deposit is

$$3Na_2SnO_3 + 3H_2O + 6Al \rightarrow 6NaAlO_2 + 3H_2 + 3Sn$$

It is improved by heating (at $\sim 100^\circ C$).

Another example is electroless nickel (Ni–P films) for which wide variations in the properties are reported, related with chemical composition and structure. The properties of the deposited films depend on the operating conditions of solutions. Increasing the metal-ion concentration, the temperature, and the agitation result in an increase of the grain size. In contrast, additive agents and increase in current density or polarization decrease the grain size. The use of additives in aqueous electroplating includes various benefits such as morphology (porosity, texture, grain size, etc.), surface (roughness, optical, and so on), stress reduction, wider range of current density, etc.

Interdiffusion and Solid-State Reaction

The reaction at the interface of two solid phases can lead to the growth of a functional film. Transition metal silicides are grown on silicon single crystals by interdiffusion at the metal–silicon interface which progressively moves as Si is consumed to form, for example, $TiSi_2$ with epitaxial relationships on the Si lattice. The first step consists in the deposition of the material on the substrate in clean conditions (PVD or CVD). The diffusion between the film and the substrate, as well as the chemical reaction is thermally activated. Optimized conditions allow building interfaces as perfect as possible, chemically and structurally. Since rather high temperatures are necessary, the method is less suitable to device production in microelectronic than in interface research.

Epitaxial Growth

The basic property of epitaxial growth is the single-crystalline structure of the deposited film with close orientation relationships with the substrate which generally is a single crystal also. Usually, the substrate and the growing film have the same composition, such as silicon epitaxial layers deposited on silicon wafers (homoepitaxy). Nevertheless, often the film is deposited with changes in the chemical composition like in $GaAs - Al_xGa_{1-x}As$ superstructures, or in lead zirconate titanate (PZT) films grown on $SrTiO_3$ (heteroepitaxy). In this case, the structural relationships require a lattice matching. The lattice constants that mismatch $\Delta a/a$ should be within a few 10^{-3}. Moreover, since the thermal expansion coefficients of the substrate (e.g., InP) and of the epitaxial layer (e.g., GaInAsP) are different, the layer can be lattice-matched at the temperature of growth and not at room temperature. Depending on the thickness, the epi-layer can nevertheless admit some degree of stress. By using complementary controlled sources, all the techniques described below are well adapted for doping.

The simplest methods are known as atomic layer deposition (ALD) or atomic layer epitaxy (ALE) and can use rather simple equipment even with stringent vacuum requirements. The choice of the precursors is crucial for the chemical reactions and the process temperatures. The halogen-based metal precursors react with atomic hydrogen near room temperature to form volatile HCl, whereas the CVD process with H_2 takes place above 800°C (**Figure 13**).

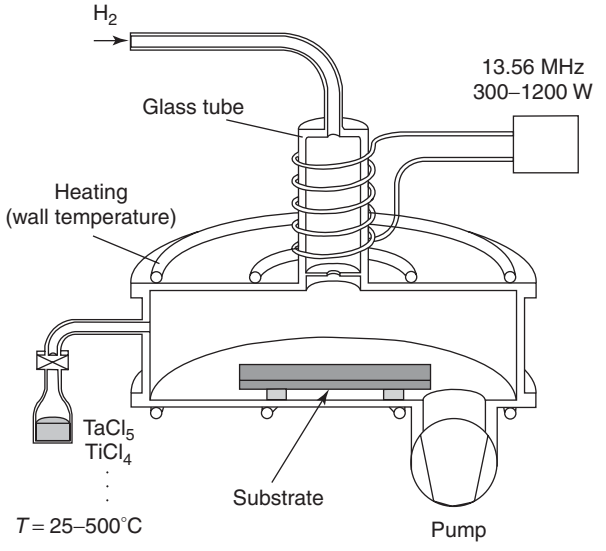


Figure 13 Schematic simple reactor for ALD.

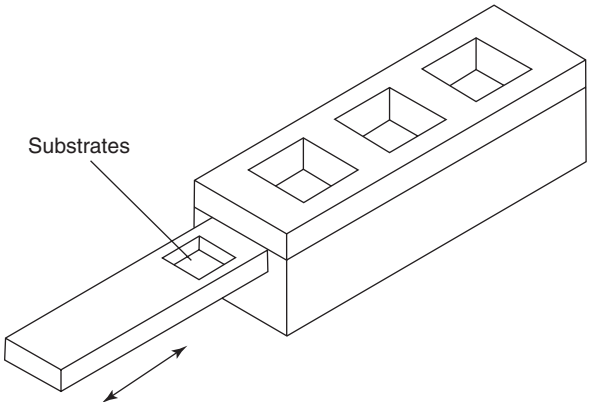


Figure 14 Growth of epitaxial layers from the liquid phase in a multiple-well slider boat.

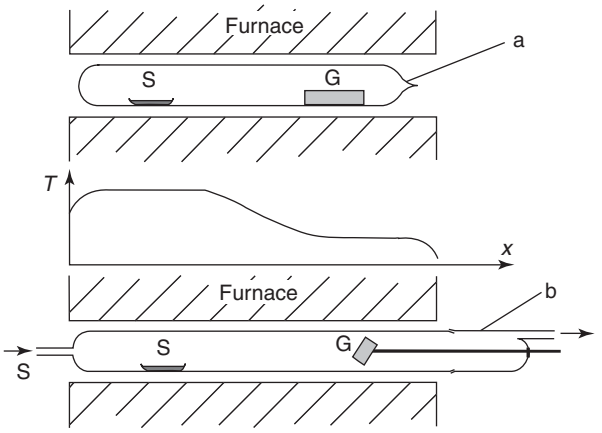


Figure 15 Vapor-phase epitaxy: (a) sealed quartz tube, (b) simple open tube reactor; S: source material (nonvolatile in a carrier gas and volatile), G: seed substrate, T: temperature profile (high-temperature source zone and low-temperature growth zone).

Liquid-Phase Epitaxy

The first technique developed to grow epitaxial layers of compound semiconductors was liquid-phase epitaxy of GaAs, AlGaAs, and then GaInAsP. Single-layer and heterostructure devices have been prepared by this method. The simplest technique is to fix the substrate at the extremity of a tube containing the melt or the solution, in a tube furnace that can be tipped. The source material is melted and after tipping, the substrate is dipped in the melt or in the sursaturated solution. If suitable temperature and pressure conditions are reached, epitaxial growth occurs. In more sophisticated equipment, the compounds are melted in a graphite boat, whose bottom includes a slider containing the single-crystalline substrates (**Figure 14**) which are pushed in contact with the liquid phase. A barrel slider has been developed to grow GaP-epitaxial layers. The growth boat is protected in a growth apparatus with rigorous temperature and atmosphere control.

Vapor-Phase Epitaxy

A primitive technique of vapor-phase epitaxy is based on the chemical transport reactions realized in a closed evacuated crucible in a tube furnace (**Figure 15a**). The source material, in the hot zone, reacts with a transport agent (generally I_2). In a suitable temperature gradient, it is transported to the growth zone where a reverse reaction occurs, and the growth takes place on a seed at lower temperature. The technique has been applied to the growth of III-V semiconductors but has been rapidly replaced by a continuous-flow reactor (**Figure 15b**) with several process parameters to control the growth: source materials (volatile or nonvolatile), temperature profile, and gas flow rates. For As or P, the trichloride XCl_3 or better the hydride XH_3 is transported by the carrier gas H_2. It reacts with the heated metal source Ga or In. If HCl is used as a carrier gas, the metal chloride forms and mixes with the hydride. The reacted compound is then epitaxially deposited if correct thermodynamic conditions are secured.

Actually, the CVD processes described above, can be run under such conditions that epitaxial growth occurs on the appropriate single-crystalline substrate.

Metal Organic Chemical Vapor Deposition

The complexity of new compound semiconductors and heterojunction devices as well as the sophisticated engineering requirements in ultrathin layers, graded layers, or abrupt interfaces has stimulated the development of metal organic chemical vapor deposition (MOCVD), which utilizes volatile organometallic compounds as the source of metallic atoms. These organometallics are mixed with transport agents of the other elements, most often hydrides, and give rise to chemical reactions of the types, for example,

$$(CH_3)_3Ga + (CH_3)_3Sb \rightarrow GaSb + 6CH_4$$

$$(C_2H_5)_2Zn + H_2Se \rightarrow ZnSe + 2C_2H_6$$

The organometallics are transported to the heated substrate by a carrier gas (nitrogen or hydrogen) charged by passing through a bubbler vessel. All the temperatures, gas flows, and pressures can be controlled. The method is particularly efficient for the deposition and growth of mixed compounds. The growth temperatures are within the range 700–900°C. Depending on the process parameters, the typical growth rate are a few tens of nanometers per minute.

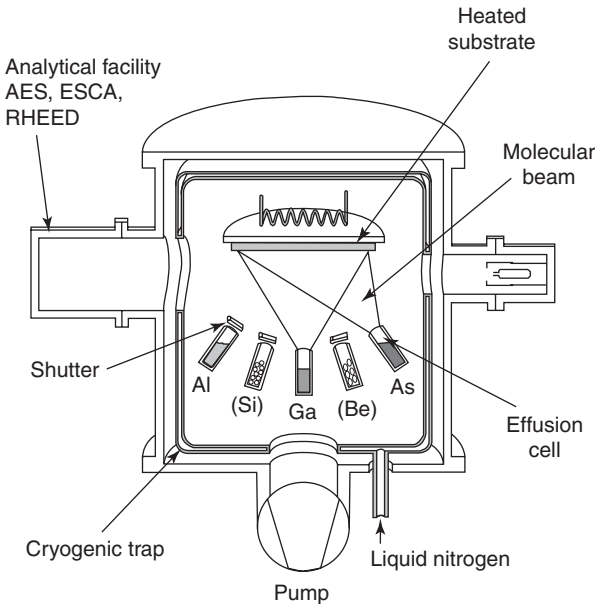


Figure 16 Schematic reactor for MBE.

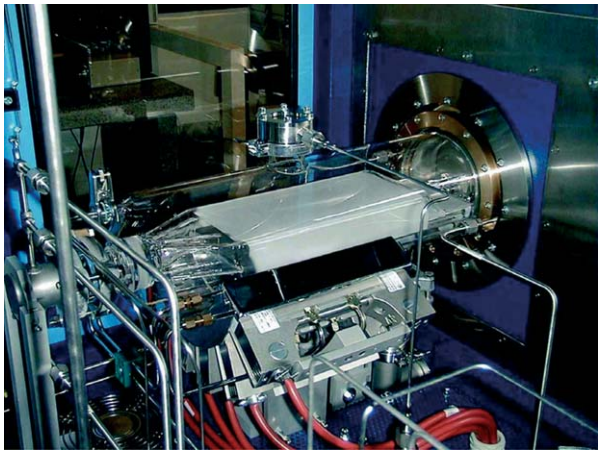


Figure 17 Double reactor for MOVCD epitaxial growth of GaAs-InP semiconductors; in the quartz reactor, the graphite substrate susceptor is radiatively heated; gas inlets from the left hand side. (Courtesy from IPEQ-SB-EPFL.)

The reactors are similar to those developed in the CVD technology (**Figure 16**). The equipment to deliver and mix the gaseous sources is important. The carrier gases, hydrides, and other gases are obtained in high-pressure cylinders. Stainless steel or quartz bubblers are used for the organometallic compounds and other liquid sources. Filtering, purification, drying, mixing are essential steps, avoiding dead spaces and leak. Automated control systems are necessary for the switching of the valves and the control of the flow rates, temperatures, and growth.

Molecular Beam Epitaxy

The method of molecular beam epitaxy (MBE) is the most improved deposition technique to control the chemical composition and doping, the thickness and smoothness of the grown epitaxial layers. The growth actually occurs layer by layer and the growth rate is determined by the flux of the high-temperature melting element (Al, Ga, and In), independent of the substrate temperature. The UHV condition of the reactor avoids external contamination. The materials sources are the elements heated and evaporated in effusion cells controlled by automated shutters (**Figure 16**).

The UHV in the reactor is compatible with surface analytical tool-like reflection high-energy electron diffraction (RHEED), electron spectroscopy for chemical analysis (ESCA), and Auger spectroscopy used *in situ*. The growth of high-purity materials necessitates unbroken UHV in the deposition chamber (10^{-8}–10^{-9} Pa), which is accessible through a sample exchange chamber.

A molecular beam is generated by heating (resistive or electron-beam heating) the source material in a cylindrical crucible (Knudsen cell). The effusion beam is directed through an orifice oriented toward the substrate. The vapor pressure of the heated materials strongly depends on the temperature. It requires a precise control of the effusion cell temperature. For the group-III elements (Al, Ga, and In), organometallic gases are also well adapted for supplying the metal (**Figure 17**). Transported to the heated substrate, the pyrolytic reaction takes place and generates the metal by removing the alkyls. For the group-V elements, the standard method is to feed arsine (AsH_3) or phosphine (PH_3) into a high-temperature "cracker" cell, where it decomposes and supplies As_2 and P_2 beams for the reaction and growth of the deposited film. The reproducible, uniform, and controlled heating of the substrate is a peremptory requirement for the growth of epitaxial films.

Acknowledgments

Sébastien Tanniger is warmly acknowledged for the original line drawings of the figures.

See also: Electron and Positron Sources; Thin Films, Mechanical Behavior of; Vapor Transport Processing: Mechanisms.

PACS: 81.15; 81.16.Mk; 81.20.Fw; 81.10.Bk; 82.45.Mp; 82.45.Qr; 85.40.Ls; 85.40.Sz; 52.77

Further Reading

Audisio S, Caillet M, Galerie A, and Mazille H (1999) *Revêtements et traitements de surface. Fonctionnalités, durabilités, procédés*. Lausanne Presses polytechniques et universitaires romandes.

Bunshah RF (1982) *Deposition Technologies for Films and Coatings*. Park Ridge: Noyes Publications.

Foord JS, Davies GJ, and Tsabg WT (1997) *Chemical Beam Expitaxy and Related Techniques*. Chichester: Wiley.

Glocker DA and Shah SL (eds.) (1995) *Handbook of Thin Film Process Technology*. Bristol: Institute of Physics Publishing.

Mahan JE (2000) *Physical Vapor Deposition of Thin Films*. New York: Wiley.

Maissel LI and Glang R (eds.) (1970) *Handbook of Thin Film Technologies*. Boston: McGraw-Hill.

Ohring M (1992) *The Materials Science of Thin Films*. San Diego: Academic Press.

Poate JM, Tu KN, and Mayer JW (1978) *Thin Film – Interdiffusion and Reaction*. New York: Wiley.

Roberts G (ed.) (1990) *Langmuir-Blodgett Films*. New York: Plenum.

Schnegraf KK (ed.) (1988) *Handbook of Thin Film Deposition Processes and Techniques: Principles, Methods, Equipments and Applications*. Park Ridge, NJ: Noyes Publications.

Smith DL (1995) *Thin-Film Deposition: Principle and Practice*. Boston: McGraw-Hill.

Wasa K and Hayakawa S (1992) *Handbook of Sputter Deposition Technology: Principles, Technology and Applications*. Park Ridge, NJ: Noyes Publications.

Fluorescent Biomolecules

F Lenci, G Checcucci, and A Sgarbossa, CNR-Istituto di Biofisica, Pisa, Italy

M M Martin and P Plaza, Ecole Normale Superieure and CNRS, Paris, France

N Angelini, CNR-Istituto Processi Chimico Fisici, Messina, Italy

The study of the fluorescence emission of biological molecules provides information on quite a number of their intrinsic properties as well as a variety of structural and functional features of macromolecules or subcellular/multicellular systems which they are associated with.

The first section of this chapter briefly discusses the main aspects of the fluorescence phenomenon in polyatomic organic molecules (in condensed phase, at room temperature), the second section concisely describes some of the principal and most advanced experimental techniques, the third section reports the key fluorescence properties of some of the main natural biological fluorophores, and the final section shows, through a few examples, how the study of fluorescence can provide crucial data to (1) elucidate structural properties of proteins and other macromolecules, (2) investigate biological membrane properties and functions, and (3) clarify the basic mechanisms responsible for biophysical/biological processes.

Fluorescence Basics

Fluorescence is the emission of light that occurs, after the absorption of a photon, from an electronically excited state which has the same spin multiplicity of the ground state. The different processes taking place between the absorption and the emission of light in organic molecules are schematized in the Perrin–Jablonski diagram, shown as **Figure 1**.

The ground state of a polyatomic molecule is generally a singlet state (S_0) with zero total spin (an exception to this rule is molecular oxygen, $^3O_2$, which is in a triplet state (T_0) in the ground state, with a total spin value equal to 1).

At room temperature, thermal energy ($kT \sim 200\,\mathrm{cm}^{-1}$) is too low to significantly populate the excited states of bond stretching or bending vibrational modes. The absorption typically occurs from molecules in the lowest vibrational level of the ground electronic state (S_{00}). After the absorption of a photon, the molecule is promoted to an excited singlet state (S_{ij}, where i refers to the electronic level and j to the vibrational level), with a probability (molecular cross section or molar extinction coefficient) that depends on the nature of the electronic charge distributions of both states (i.e., nature of the involved orbitals) and the molecular structure (in particular symmetry). The probability to directly excite the molecule into a triplet state (with a spin change) is extremely low and generally not observed. The absorption transition has an "instantaneous" nature: in solution, it occurs in times of the order of femtoseconds ($1\,\mathrm{fs} = 10^{-15}\,\mathrm{s}$, about one period of a UV–visible light cycle) and too short a time for significant displacements of the nuclei to occur (Franck–Condon principle). The $S_{00} \rightarrow S_{ij}$ vibronic transition probability (shape of the absorption spectrum) is dictated by the overlap of the initial and final vibrational wave functions in the ground-state configuration of the nuclei (Franck–Condon factors). The probability for the subsequent fluorescence, emitted from the lowest vibrational level of S_1 to any vibrational level of the ground state, is directly

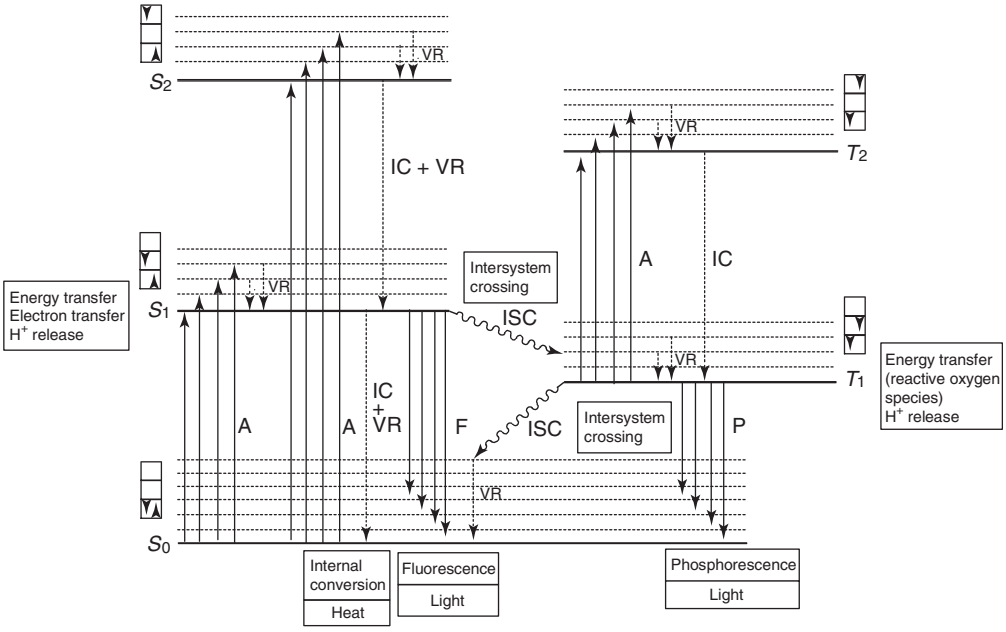


Figure 1 Perrin–Jablonski diagram of electronic states (*S* for singlet and *T* for triplet states) and of intramolecular transitions of polyatomic molecules. A: absorption; VR: vibrational relaxation; IC: internal conversion; ISC: intersystem crossing; F: fluorescence; P: phosphorescence (see A Krasnovsky's comment, in C Lambert and L Jones "Basic Photophysics" Digital Photobiology Compendium, D. Valenzeno Editor). Aleksander Jablonski did not propose a triplet manifold of states to explain phosphorescence (delayed fluorescence). Perrin was the first to claim the existence of a metastable state between the ground and fluorescent states of dyes (Perrin F (1929) *Annals of Physics* (Paris) 12: 169). A few years later Jablonski proposed that radiative transition from the metastable to the ground state is a reason for the low-temperature phosphorescence of dyes (Jablonski A (1933) *Nature* 131: 839 and (1935) *Zeitschrift für Physik*. 94: 38). The idea that the metastable state is a triplet state was first proposed by Terenin AN ((1943) *Acta Physico-Chimica USSR* 18: 210–241) and by Lewis GN and Kasha M ((1944) *Journal of American Chemical Society* 66: 2100–2116). The diagram referred to as a Jablonski diagram, is really a Perrin–Jablonski diagram modified by Terenin and Lewis–Kasha.

related to that of the corresponding $S_0 - S_1$ absorption transition.

A fluorophore excited to some high-lying vibrational level of S_1 or to higher singlet excited states rapidly relaxes (10^{-12} s or less), in condensed phase, to the lowest vibrational level of S_1 (Kasha's rule), by internal conversion and vibrational relaxation with energy transfer to the solvent. Several processes, both radiative and nonradiative, are known to depopulate the first excited singlet state, all of them influencing the characteristics of the fluorescence emission. Nonradiative de-excitations are given as follows:

1. internal conversion between same-spin electronic states (the lowest vibrational level of the highest electronic state overlaps the vibrational levels of the lowest electronic state that have the same energy);
2. intersystem crossing between states with different spin multiplicity (typically from S_1 to T_1); this transition is forbidden by the spin selection rule, but occurs because the real molecular electronic states are a superposition of singlet and triplet states due to the spin–orbit interaction; the molecule in the lowest triplet state can then return to the ground state by phosphorescence emission, but the probability of this transition is very low and phosphorescence is generally not observed at room temperature; and
3. any de-excitation process resulting from collisions with solute molecules or any reaction such as charge or energy transfer and functional reactions.

Fluorescence Spectral Distribution

As mentioned above, Kasha's rule states that the fluorescence emission of a molecule in condensed phase generally occurs from the lowest vibrational level of the first excited singlet state, S_{10}. Consequently:

1. As clearly appears from the Perrin–Jablonski diagram, the energy of emission is typically less than that of absorption (Stokes' shift), that is, the fluorescence emission spectrum is almost completely localized at longer wavelengths with respect to the absorption spectrum. In addition, fluorophores can exhibit an increased Stokes' shift due to solvent reorganization, excited-state reactions (such as photocomplexation) or energy transfer.
2. The emission spectral distribution (measured upon excitation at a single constant wavelength) is

usually independent of the selected excitation wavelength. This is intuitively understood: upon excitation to higher electronic and vibrational levels, the excess energy is quickly dissipated into the solvent, leaving the fluorophore in the lowest S_{10}, from which fluorescence occurs. In this way, the fluorescence spectrum is always the same set of the allowed $S_{10} \rightarrow S_{0n}$ transitions, irrespective of the excitation wavelength.

3. The emission spectrum is normally the mirror image of the $S_0 \rightarrow S_1$ absorption band, not of the whole absorption spectrum. In fact, the probability of a particular vibronic absorption transition ($S_{00} \rightarrow S_{1n}$) – between the zeroth vibrational level of S_0 and the nth vibrational level of S_1 – is close to one of the "corresponding" emission transition ($S_{10} \rightarrow S_{0n}$, with nearly identical overlap of the initial and final vibrational wave functions).

Fluorescence Lifetime and Quantum Yields

Fluorescence lifetimes and quantum yields, together with the spectral distribution, are the most important characteristics of a fluorophore. The fluorescence quantum yield (Φ_F) is the ratio of number of emitted photons (per time and volume unit) relative to the number of absorbed ones ($\Phi_F \leqslant 1$). Hence, Φ_F is the ratio of the intrinsic radiative rate constant (k_{rad}) to the sum of the rate constants of all the processes that contribute to depopulate the first excited singlet state:

$$\Phi_F = \frac{k_{rad}}{k_{rad} + k_{IC} + k_{ISC} + k_{react}} = \frac{k_{rad}}{k_{rad} + k_{nr}} \quad [1]$$

Such processes are: internal conversion (k_{IC}), intersystem crossing (k_{ISC}), and quenching or photophysical/photochemical reactions of various types (k_{react}). Thus, the fluorescence quantum yield is close to unity, only if the sum of the radiationless decay rates is much smaller than the rate of radiative decay, that is, $k_{nr} \ll k_{rad}$.

It is worthwhile emphasizing that, due to the many nonradiative processes, the first singlet excited state lifetime (in other words, the observed fluorescence lifetime τ_F) is, in general, shorter than the radiative lifetime (the reciprocal of the radiative rate). On the other hand, the time-dependent fluorescence intensity is proportional to the excited-state population. The kinetic equation, describing the decay of the singlet excited-state population $[S_1]$, after the excitation has been switched off, is constructed by adding up all parallel de-excitation pathways:

$$\frac{d[S_1]}{dt} = (k_{rad} + k_{nr})[S_1] \quad [2]$$

The solution of this equation reads

$$[S_1](t) = [S_1](0)e^{-t/\tau_F} \quad [3]$$

where $[S_1](0)$ is the excited-state population at time zero (defined any moment after the end of the excitation), and τ_F is the observed fluorescence lifetime:

$$\tau_F = (k_{rad} + k_{IC} + k_{ISC} + k_{react})^{-1} = (k_{rad} + k_{nr})^{-1} \quad [4]$$

In practice, when employing eqn [3] for describing the observed decay of the fluorescence intensity, one chooses the origin of time at the end of the excitation, that is, at the maximum of the measured fluorescence signal.

Combining eqns [1] and [4], one obtains

$$\Phi_F = \tau_F k_{rad} = \tau_F / \tau_{rad} \quad [5]$$

Fluorescence quantum yields and lifetimes, likewise fluorescence emission spectra, generally do not depend on the excitation wavelength. For an ideal instrument, the fluorescence excitation spectrum of a molecule (defined as the fluorescence intensity measured at a fixed wavelength as a function of the excitation wavelength), therefore, entirely overlaps the absorption spectrum. Discrepancy between absorption and excitation spectra is a clear indication of excitation-wavelength-dependent photoinduced reactions.

Time-Resolved Fluorescence

As absorption occurs in $\sim 10^{-15}$ s, steady-state absorption spectra are not sensitive to molecular dynamics and only provide information on the ground state itself and on the ground-state environment of the chromophore. Steady-state fluorescence is more informative on excited-state processes, but some information is unavoidably lost during the time-averaging process. For example, macromolecules can exist in more than one conformation without much change of the emission spectrum. The decay time, on the other hand, may strongly depend on the conformation because it reflects the interactions of the fluorophore with the solvent or with other molecules and macromolecules of its close environment. Time-resolved fluorescence then appears as a valuable tool to discriminate these conformations and, in general, to understand the dynamical photoinduced processes in molecules and macromolecules. In practice, a fluorescence decay profile can exhibit several lifetimes that one has the possibility to assign to specific reactions or species.

Fluorescence Quenching

In photoinduced biological phenomena originating from the first excited singlet state, the very primary events (electron, proton, or energy transfer, photoisomerization, and, in general, any kind of photophysical or photochemical reaction) are alternative nonradiative pathways for the excited-state decay. They yield a decreased fluorescence intensity, called "quenching," and reduced fluorescence quantum yield and lifetime (see **Figure 2**).

Quite a number of very different processes can be responsible for fluorescence quenching. In the following, some aspects of collisional quenching and of energy transfer are concisely discussed.

Collisional quenching takes place when the excited-state fluorophore is deactivated upon contact with some other molecules in solution (quencher). The mechanism of quenching varies with the fluorophore–quencher pair. For instance, it can be due to an electron transfer from the fluorophore to the quencher, which does not occur in the ground state. Alternatively, in the case of oxygen, halogens, and heavy atoms, it can be due to spin–orbit coupling and intersystem crossing (triplet-state formation). For collisional quenching, the decrease in fluorescence yield and lifetime is described by the Stern–Volmer equation (the fluorescence yield can be probed by the fluorescence intensity F, as long as the experimental conditions are kept identical):

$$\frac{F_0}{F} = \frac{\tau_{F_0}}{\tau_F} = 1 + K[Q] = 1 + k_q \tau_{F_0}[Q] \qquad [6]$$

where K is the Stern–Volmer quenching constant, k_q is the bimolecular quenching constant, τ_F and τ_{F_0} are the quenched and the unquenched fluorescence lifetimes respectively, F and F_0 are the quenched and the unquenched fluorescence intensities respectively, measured under the same experimental conditions, and $[Q]$ is the quencher concentration. The value of k_q, which can be obtained simply by plotting F_0/F or τ_{F_0}/τ_F versus $[Q]$, for a free fluorophore in solution is, at most, $\sim 10^{10}\,M^{-1}s^{-1}$, that is, the rate at which the fluorophore and the quencher molecules diffuse through the solvent and encounter. However, k_q can be smaller than the diffusion rate, if the reaction occurring in the encounter complex has a low probability. When chromophores are bound to or incorporated into macromolecules, k_q sometimes is substantially smaller, because of the shielding effects of the macromolecules structure.

Another important quenching process is the fluorescence resonance energy transfer (FRET). This process occurs whenever the emission spectrum of the fluorophore, called the donor, overlaps the absorption spectrum of another molecule, called acceptor. It is very important to note that FRET does not involve emission of light by the donor and consequent absorption by the acceptor. The donor, in its excited state, and the acceptor, in the ground state, are coupled by dipole–dipole interaction, resulting in the deactivation of the donor excited state and the excitation of the acceptor. This process is also called long-range coulombic or Förster-type energy transfer. Another type of nonradiative excitation energy transfer is the Dexter-type energy transfer, by electron exchange in an encounter complex, which is not discussed here. The rate constant of the Förster resonance energy transfer $k_T(r)$ strongly depends on the distance r between the donor and the acceptor:

$$k_T(r) = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \qquad [7]$$

where τ_D is the donor fluorescence lifetime in the absence of energy transfer and R_0 the Förster radius or critical distance which depends on the donor fluorescence quantum yield, the acceptor absorption transition probability, the spectral overlap mentioned above, and the relative orientation of the two interacting dipoles. On the basis of eqn [7], when the donor–acceptor distance is equal to the Förster distance ($r = R_0$), the rate of transfer is equal to the decay rate of the donor in the absence of the acceptor

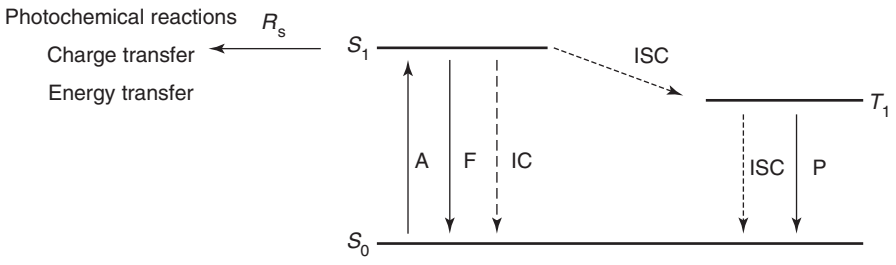


Figure 2 Nonradiative deactivation pathways originating from the first excited singlet state resulting in "functional" processes (e.g., photosynthesis (energy transfer) and photoperception (photoisomerization, photoinduced charge transfer)).

and, consequently, the transfer efficiency is 50%. The transfer quantum efficiency (Φ_{FRET}), defined as the r-dependent fraction of photons absorbed by the donor that are transferred to the acceptor, can be obtained directly by eqn [7]:

$$\Phi_{FRET} = \frac{k_T}{k_T + (1/\tau_D)} = \frac{R_0^6}{R_0^6 + r^6} = 1 - \frac{\tau_{DA}}{\tau_D} = 1 - \frac{F_{DA}}{F_D} \quad [8]$$

The transfer efficiency is typically measured using the relative fluorescence yield (or intensity under the same experimental conditions) or lifetime in the absence (F_D and τ_D) and presence (F_{DA} and τ_{DA}) of the acceptor, provided that the donor–acceptor pairs are separated by a fixed distance during the lifetime of the donor. On the basis of eqn [8], fluorescence resonance energy transfer provides an opportunity to measure the distance between sites on macromolecules, provided r is not too different from the Förster distance R_0: typically, in the range of 15–60 Å, which is comparable to the diameter of many proteins and the thickness of a membrane.

A completely different cause of fluorescence intensity decrease is photobleaching, due to a photoinduced degradation or damage of the fluorophore, normally caused by focused laser beams or very high intensity excitation light.

Fluorescence Anisotropy

Fluorescence anisotropy measurements, commonly used in a variety of biochemical applications, provide information on the size and shape of macromolecules or the rigidity of various molecular environments. Anisotropy measurements are based on the principle that linearly polarized excitation light selectively excites the fluorophores, the absorption transition dipole of which is parallel to the electric vector of the excitation. In an isotropic solution, the fluorophores are randomly oriented, so that this selective excitation results in a partially oriented population of excited fluorophores, provided the excitation pulse duration is shorter or of the same order of magnitude as the fluorophore reorientation time (e.g., with femtosecond pulses). The fluorescence emission, in turn, is polarized along the direction of the emission transition dipole of the fluorescent molecule. The anisotropy (r) of the emission is defined as follows:

$$r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp} \quad [9]$$

where $I_\parallel$ and $I_\perp$ are the fluorescence intensities, respectively, measured parallel and perpendicular to the polarization direction of the excitation light. The maximum attainable anisotropy (r_0) depends on the relative angle (β) between the absorption and emission transition moments. For an initially randomly oriented sample, it reads

$$r_0 = \frac{2}{5}\left(\frac{3\cos^2\beta - 1}{2}\right) \quad [10a]$$

so that the "maximum maximorum" value of (r_0) is 0.4, with $\beta = 0$. Several phenomena can decrease the measured anisotropy to values lower than this maximum. The most obvious cause is rotational diffusion that, during the lifetime of the excited state, induces a randomization of the orientation of the fluorophore, namely its emission dipole moment. In fluid solution, most molecules rotate in 50–100 ps (10^{-12} s). Hence, fluorescent molecules can rotate many times during their 1–10 ns excited-state lifetime and, consequently, the emission is completely depolarized (the steady-state anisotropy is close to zero). Nevertheless, provided that the experiments are done with a sufficiently short time resolution, the molecular reorientation can be followed by measuring the time-resolved anisotropy $r(t)$:

$$r(t) = \frac{I_\parallel(t) - I_\perp(t)}{I_\parallel(t) + 2I_\perp(t)} = r_0 e^{-6D_r t} \quad [10b]$$

where D_r is the rotational diffusion coefficient and $(6D_r)^{-1}$ the rotational correlation time (θ).

For macromolecules and proteins, the rotational correlation time is much longer ($\sim$1–100 ns) and it is possible to use fluorescence anisotropy for probing the mobility of any fluorophore and determine how rigidly it is bound to some macromolecule. Assuming no other processes result in fluorescence depolarization, the expected steady-state anisotropy (r) is given by the Perrin equation:

$$r = \frac{r_0}{1 + (\tau_F/\theta)} \quad [11]$$

where r_0 is maximum anisotropy (given by eqn [10a]), τ_F is the fluorescence lifetime, and θ is the rotational correlation time for the diffusion process, which is given by the Debye–Stokes–Einstein relation

$$\theta = \frac{\eta V}{kT} \quad [12]$$

where η is the medium viscosity (in poise, P), V the molecular volume, k is the Boltzmann constant, and T is the temperature in kelvins.

Main Instrumentation

Steady-State Fluorimetry

Figure 3 shows a simplified scheme of a typical multipurpose fluorimeter. The excitation light from a xenon (high pressure) lamp is used to excite the sample at a wavelength selected by a motorized monochromator (UV–visible). The fluorescence is collected at right angle with respect to the excitation beam direction and sent onto a photomultiplier through a second motorized monochromator. Computer-controlled scanning of the latter at a fixed excitation wavelength allows measurements of the fluorescence spectra, whereas scanning of the excitation monochromator at a fixed emission wavelength allows measurements of the excitation spectra. A small part of the excitation beam is sent to a reference detector, the signal of which is used for compensation of the excitation intensity fluctuations and correction of the excitation spectral distribution when recording excitation spectra. Fluorescence emission spectra are also corrected for instrumental distortions. For anisotropy measurements, polarizers are added on the excitation and emission beams as shown in the scheme.

Time-Resolved Fluorescence Techniques

Different techniques can be used for measuring fluorescence decays. The most straightforward is to excite the sample by a short light pulse $E(t)$ (lamp or laser flash) and to record the fluorescence signal $F(t)$ as a function of time. The time resolution is then limited by the excitation pulse width and the temporal response of the detection device. The measured signal $F(t)$ is the convolution product (eqn [13]) of the apparatus response function $R(t)$ (equal to $E(t)$, if detection is not a limit) and the purely molecular fluorescence function $I_F(t)$, expected for an infinitely short excitation pulse, that is, the relevant decay function that describes the excited-state dynamics.

$$F(t) = R(t) \otimes I_F(t) = \int_{-\infty}^{t} R(t')I_{\mathrm{F}}(t - t')\, \mathrm{d}t' \quad [13]$$

Provided a short enough excitation source is available, nanosecond (10^{-9} s) and subnanosecond decay times can be easily measured by collecting the sample emission onto a photodiode or a photomultiplier coupled to an oscilloscope. To date, equipments even allow to readily achieve the picosecond regime. For this time range, streak cameras are, however, more suitable, although they have an intrinsic limitation due to their low dynamic range (low saturation level).

Alternatively, the time-correlated single-photon counting technique is a powerful method quite often used for reliable measurements in time ranges that extend from a few tens of picoseconds to a few nanoseconds. For longer lifetimes, care must be taken when choosing the excitation-pulse repetition rate, since the fluorescence signal must have decayed when the next excitation pulse comes in. On the other hand, decay components of $I_{\mathrm{F}}(t)$ shorter than the width of the response function $R(t)$ can be determined by deconvolution of the fluorescence signal $F(t)$, according to eqn [13].

A typical single-photon counting setup is shown in **Figure 4**. The principle consists in measuring a

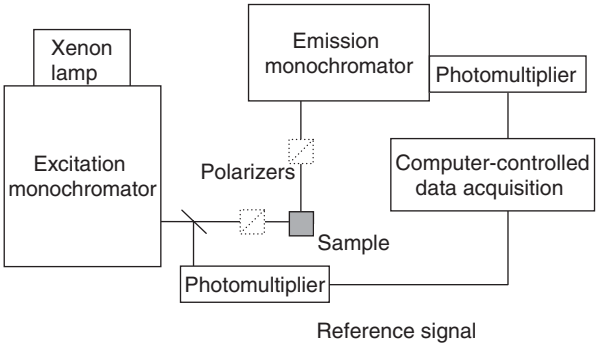


Figure 3 Scheme of a fluorimeter. The sample is excited with a xenon lamp at a wavelength selected by the excitation monochromator. Its fluorescence is collected at right angle and sent onto a photomultiplier through the emission monochromator. Scanning of the latter allows measurements of fluorescence spectra. A small part of the excitation beam is used to produce a reference signal which is used for compensation of the excitation-intensity fluctuations. Excitation spectra can be measured with the same setup by fixing the emission wavelength and scanning the excitation monochromator. Polarizers can be inserted in the beams for anisotropy measurements.

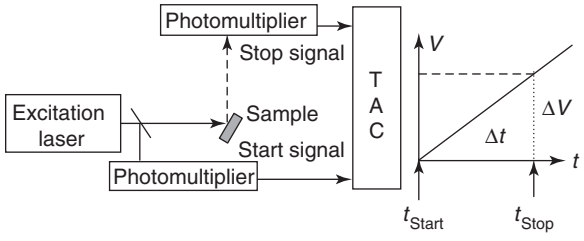


Figure 4 Simplified scheme of a single-photon counting setup. The sample is excited with a low fluence, high repetition rate, and excitation pulse. An electrical pulse associated to the optical pulse starts the deflection ramp of a time-to-amplitude converter (TAC) while the sample is excited. The fluorescence is collected onto a photomultiplier operating in a single-photon counting mode providing a signal that stops the deflection ramp when a single photon is detected. The voltage at which the deflection ramp is stopped is proportional to the delay between the start and stop signals. Measurements are repeated until the histogram of delays builds up. The histogram, recorded with a multichannel analyzer, reproduces the rise and decay $F(t)$ of the sample fluorescence.

histogram of the time delays at which a single photon of fluorescence is emitted after excitation. Without going into details, the excitation pulse is split into two pulses, one of which is used to excite the sample. The second pulse is used to generate an electrical pulse that triggers (start signal) the deflection ramp of a time-to-amplitude converter (TAC) while the sample is excited. The excitation fluence is fixed at such a low level that less than one photon of fluorescence is recorded per excitation pulse. The emitted fluorescence is collected onto a photomultiplier operating in a single-photon counting mode so that, when a single photon is detected, the photomultiplier provides a signal that stops the deflection ramp of the TAC. The voltage at which the deflection ramp is stopped is proportional to the delay between the start and stop signals. Measurements are repeated until the histogram of delays builds up. This histogram, recorded with a multichannel analyzer, reproduces the rise and decay $F(t)$ of the sample fluorescence.

Several other time-resolved fluorescence techniques, such as phase-modulation fluorimetry, Kerr-shutter fluorimetry, and fluorescence up-conversion, exist. The up-conversion technique is the only technique that allows direct measurements of femtosecond fluorescence decays, that is, the fluorescence from quasi nonfluorescent molecules, but is out of the scope of this review. In the Kerr-shutter fluorimetry, the electric field of an intense laser pulse is used to induce a temporary anisotropy (Kerr effect) in a Kerr sample placed between two crossed polarizers so that light can be temporarily transmitted by this arrangement. By adjusting the delay between the time at which a fluorescent sample is excited and the time at which the Kerr-shutter is opened, one can record the transmitted fluorescence intensity, or spectrum, as a function of time. In phase-modulation fluorimetry, the sample is excited with a sinusoidally modulated light, at high frequency. The fluorescence response is sinusoidally modulated at the same frequency, but delayed in phase and partially demodulated. These effects depend on the modulation frequency of the excitation light and on the fluorescence decay time. The experiments allow measurements of the Fourier transform of the fluorescence kinetics. Kerr-shutter and phase-modulation fluorimetry are suitable techniques for lifetime measurements in the nanosecond and picosecond regimes.

Fluorescence Correlation Spectroscopy

In fluorescence correlation spectroscopy (FCS), one measures the fluorescence intensity of a very small sample volume ($\sim 10^{-15}$), for example, under a confocal microscope, and records its fluctuations in time, $\delta F(t)$, around an equilibrium value, $\langle F \rangle$. The result is quantified in terms of the normalized autocorrelation function $G(\tau)$:

$$G(\tau) = \frac{\langle F(t+\tau)F(t) \rangle}{\langle F \rangle^2} = \frac{\langle \delta F(t+\tau)\delta F(t) \rangle}{\langle F \rangle^2} + 1 \quad [14]$$

where $\delta F(t) = F(t) - \langle F \rangle$. The main source of fluctuation comes from the molecules diffusing in and out of the examined volume. The mean diffusion time and thus the size of the molecule can be calculated on the basis of an analysis of the G function. The photophysics (intersystem conversion) and photochemistry (photobleaching, excited-state reactivity) of the fluorophores may also come into play. The high sensitivity of the detectors even allows the measurement of the fluorescence fluctuations of a single molecule. Fluctuation analysis has proved to be a very versatile and powerful tool for the detection and dynamics study of biomolecules at ultralow concentrations on surfaces, in solutions, and in living cells.

In practice, a parallel laser beam is concentrated through a high aperture microscope lens on a focal point of $< 0.5\,\mu m$ in diameter. This focal point forms the open measuring volume in a highly diluted solution in which the fluorescing molecules diffuse. The focal point is confocally imaged on a pinhole behind which a single photon detector is placed. In this way, only fluorescent biomolecules stimulated by the laser in the focal point are detected. Although the route of the molecule through the focal point is random, the mean diffusion time and thus the size of the molecule can be calculated on the basis of a temporal correlation of the signal of many single molecules.

Main Natural Biological Fluorophores

In **Figure 5**, the formulas and the optical absorption and emission spectra of the three aromatic amino acids are reported. Tyrosine and phenylalanine have lower molar extinction coefficients ε ($\varepsilon = A$ (absorbance)/(C (molar concentration in M) $\times$ l (optical path in cm))) with respect to tryptophan, the molar extinction coefficient of which, however, is rather low when compared to that of other absorbing molecules (e.g., for tryptophan, ε is $\sim 5600\,M^{-1}\,cm^{-1}$, whereas for chlorophyll, ε is $\sim 120\,000\,M^{-1}\,cm^{-1}$). As a matter of fact, tryptophan is the most commonly and easily studied amino acid as an intrinsic probe for protein microenvironment and bulk properties. Usually, tryptophan is selectively excited at 295 nm because, exciting at 280 nm, a significant contribution to fluorescence by tyrosine can also be observed.

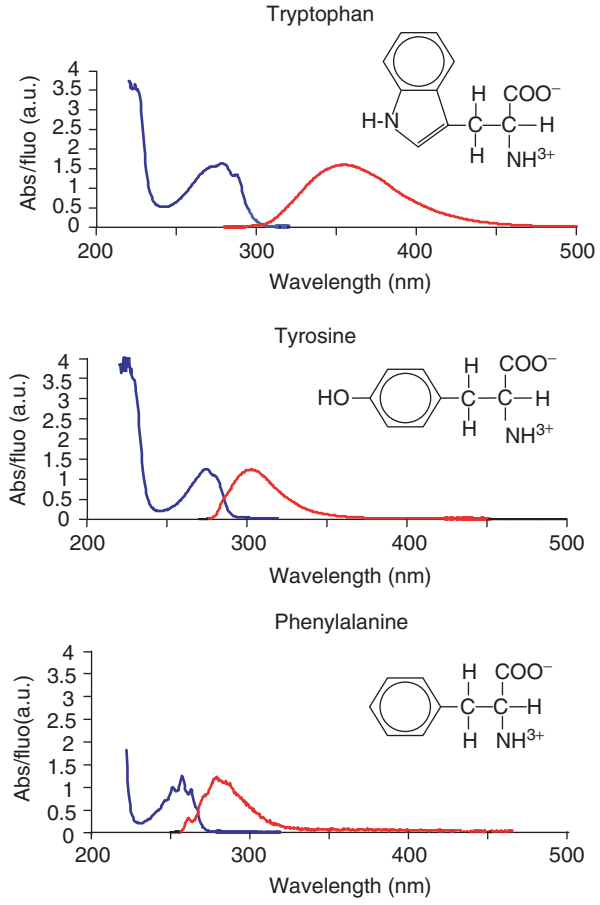


Figure 5 Formulas and absorption and fluorescence spectra of aromatic amino acids.

Figures 6–9 show the formulas and the spectra of some of the most widely distributed biological fluorophores: chlorophyll (see **Figure 6**), which play a key role in photosynthesis, flavins (riboflavin, flavin mono- and dinucleotide) (see **Figure 7**), which play a key role in quite a number of biological processes (see the next section), phytochrome (see **Figure 8**), a most important photosensing system for plant photomorphogenesis, and hypericin (see **Figure 9**), a typical natural pigment with photodynamic properties.

Fluorescent Biomolecules for Structural and Functional Studies

When a fluorescent biomolecule is used for structural or functional studies, one usually measures the main fluorescence parameters described in the first section which explains the fluorescence excitation and emission spectra, the fluorescence quantum yield, lifetime, and anisotropy.

If the biological system under investigation does not contain any intrinsic fluorophore or if its intrinsic fluorophores are of limited use, smart synthetic fluorescing molecules can be used as probes and labels in areas such as membrane biophysics, cell sorting, ion transport, and immunoassay, to mention only a few. In this article, the case of these numerous extrinsic fluorescent probes are not discussed: their different, specific, and dedicated applications can be easily found in a number of commercial catalogs. The focus is rather on "intrinsic" probes, such as tryptophan, flavins, and the other fluorophores described in the previous section.

In the following, the comparison of the fluorescent properties of these fluorophores in different molecular environments highlights how and why fluorescence is a valuable tool to probe and monitor molecular or supramolecular structures and chromophore–protein interactions.

Fluorescence for Probing Molecular Environments

The chemical and physical properties of the solvent, such as pH and polarity, can strongly affect the fluorescence emission spectrum as well as the fluorescence quantum yield and lifetime of a free fluorophore. These changes may alter its fluorescence anisotropy, if the nature of the optical transitions is changed (see eqns [9] and [10]), by a different excited-state pathway for example, or through a change of lifetime (eqn [11]). If it is assumed that this is not the case (same transitions and fluorescence lifetime), then the fluorescence anisotropy depends solely on the solvent viscosity as eqns [11] and [12] express.

In flavin adenine dinucleotide (FAD), for example, the strength of the ground-state stacking interactions between the flavin and the adenine moieties, which are responsible for the flavin moiety fluorescence quenching with respect to riboflavin or flavin mononucleotide (FMN), depends on pH. The unfolding of the complex at pH 3 is reflected by an increase in the flavin fluorescence quantum yield.

In isolated fluorophores, solvent polarity can influence the fluorescence spectra (solvatochromic shift) and high viscosity can enhance the fluorescence anisotropy. The vibronic structure of the electronic transitions can also be affected by the solvent nature, depending on the type and strength of fluorophore–solvent interaction. Quite similar effects can be observed, when the fluorophore binds to (or associates with) a protein or a macromolecular structure.

As mentioned before, selective excitation of the tryptophan fluorescence has been extensively used as an endogenous probe of the amino acid location. Because of their particular interactions with their molecular microenvironments, the insertion of

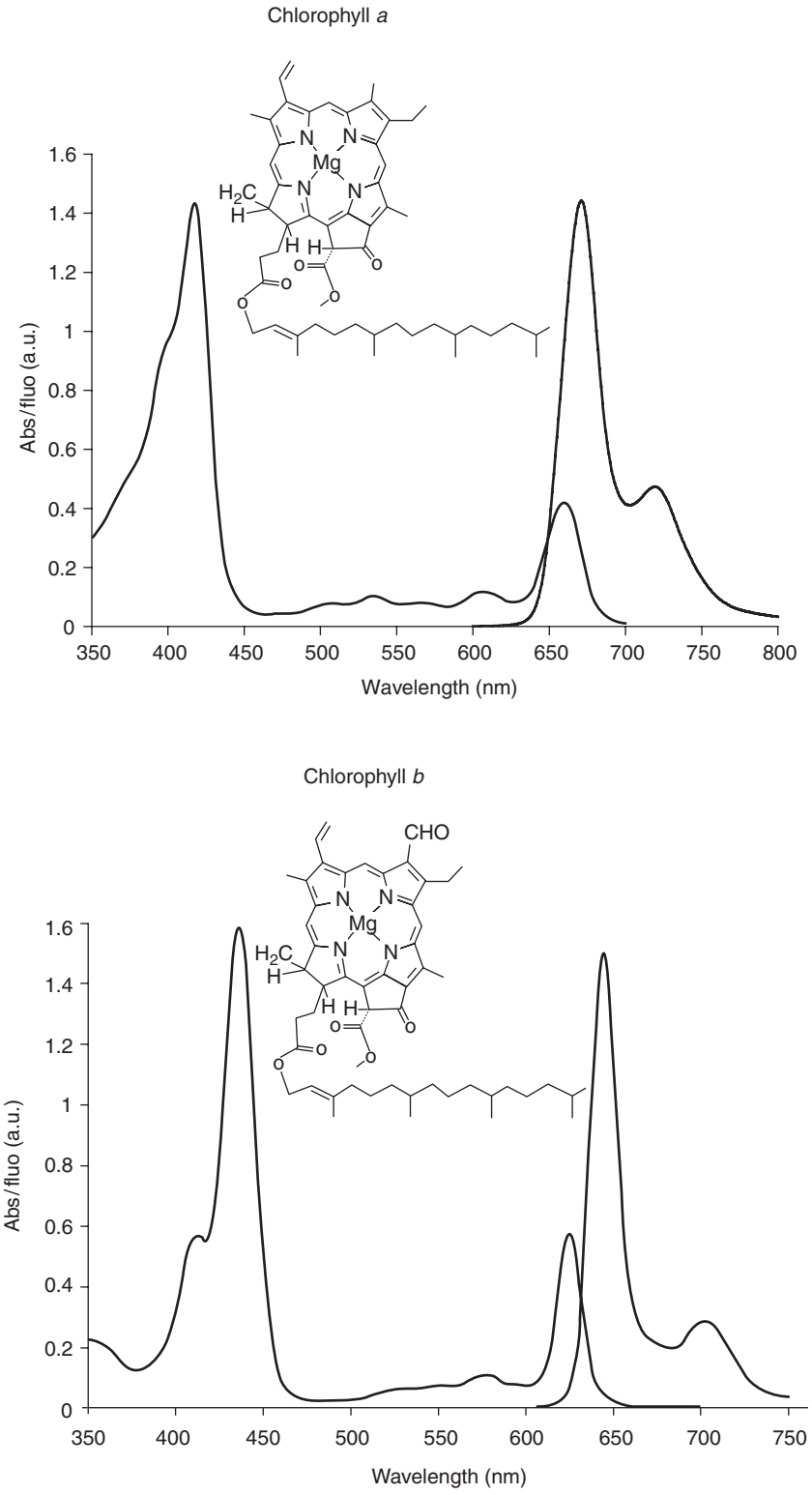


Figure 6 Formulas and absorption and fluorescence spectra of chlorophyll *a* and *b*.

tryptophan molecules into a proteic framework causes marked changes in its spectroscopic properties. Among these characteristic effects of the surrounding environment is hypochromism, a phenomenon that affects both absorption and emission and has several causes. For example, the presence of ordered layers of chromophores favors the screening of light, thus effectively reducing the molar extinction coefficient

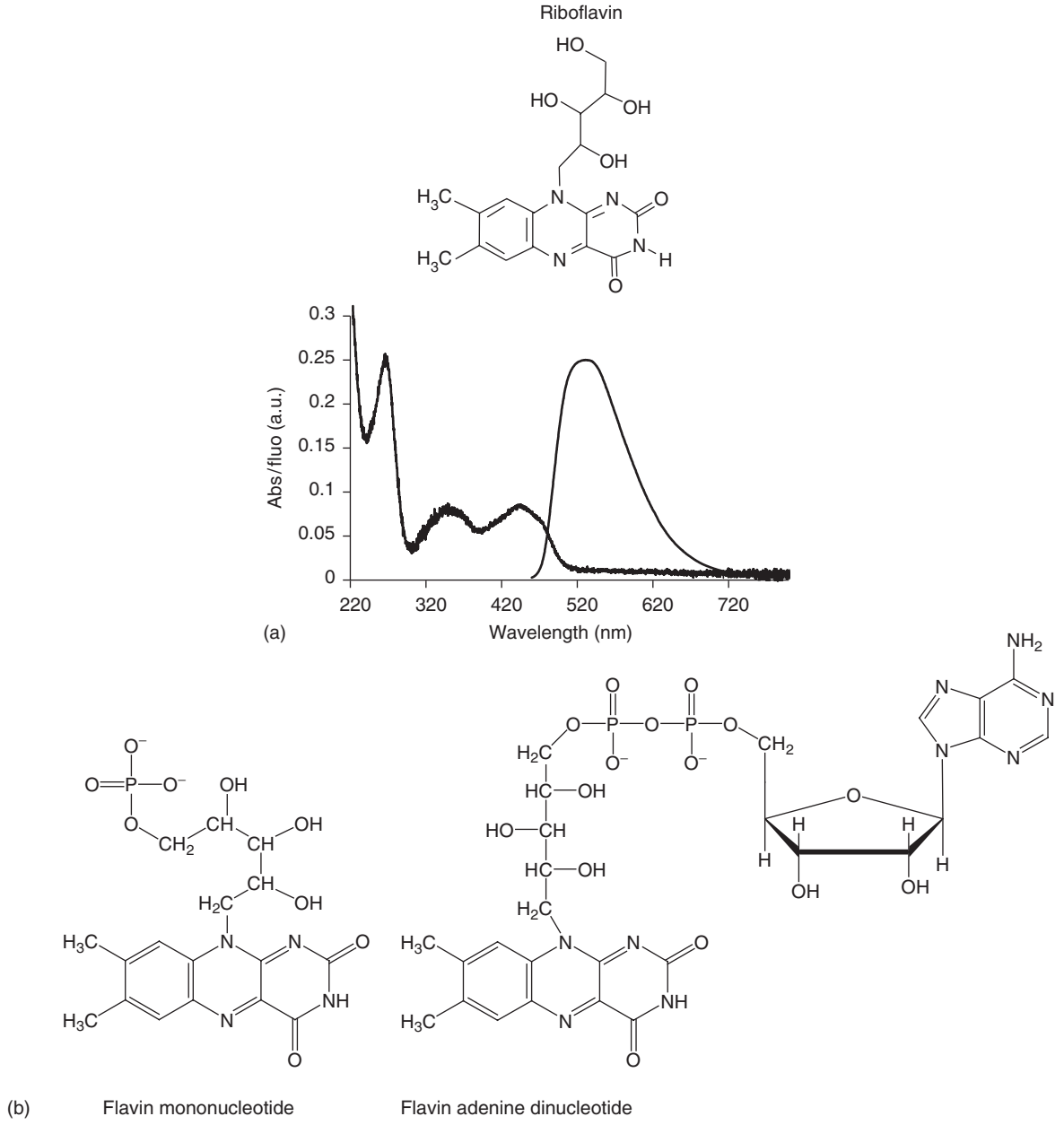
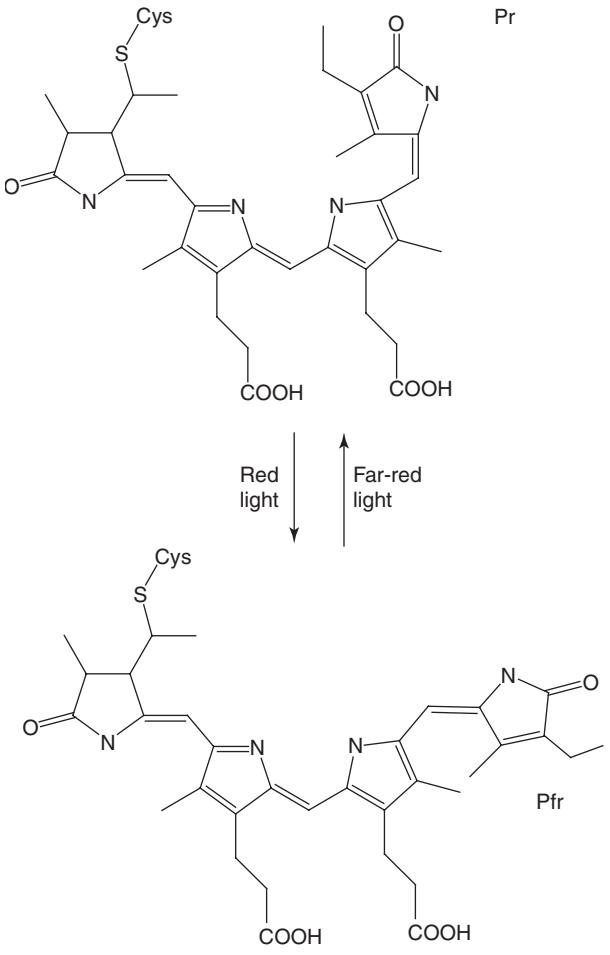


Figure 7 (a) Formulas, and absorption and fluorescence spectra of riboflavin. (b) Formulas of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Their absorption and fluorescence spectra are not reported because substantially indistinguishable from those of riboflavin.

of the amino acid (self- or hetero-screening hypochromism); the screening can decrease the molar extinction coefficient to half, depending on the parallelism of the chromophores dipoles (the higher the parallelism, the lower the coefficient). Also non-absorbing groups can affect the emission properties of this amino acid. In fact, in a protein, several groups can quench the fluorescence because their presence favors a photoinduced intramolecular charge transfer process. For example, a proton transfer is at the basis of tryptophan fluorescence quenching caused by carboxyl and thiolate groups, whereas amide, peptide, ammonium, ammine, histidine, and other sulfur-containing groups quench tryptophan fluorescence by means of an electron transfer. Changes in the structure of the fluorescence emission spectrum have also been observed. In tryptophan up to five different emission maxima can be identified, depending on how buried is the amino acid in the protein. When the amino acid is close to a hydrophilic environment (i.e., near the external surface of the protein), the fluorescence maximum is similar to that shown by the amino acid in aqueous solution (peak at 353 nm), while the maximum shifts



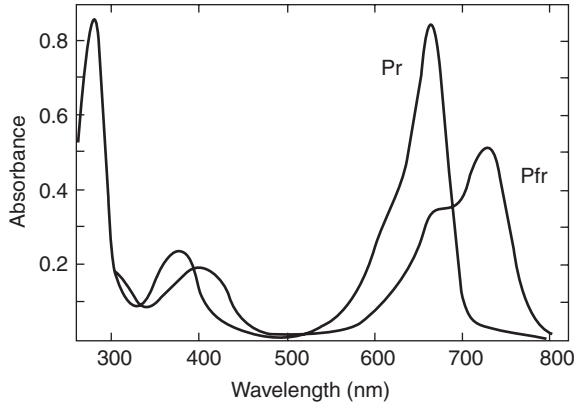
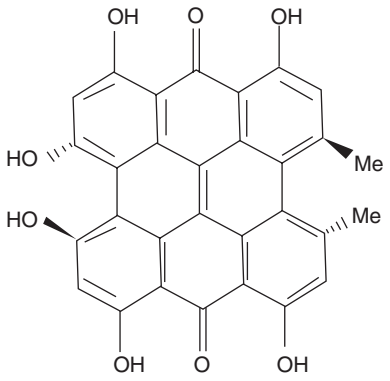


Figure 8 Formulas and absorption spectra of red (Pr) and far-red (Pfr) forms of phytochrome.



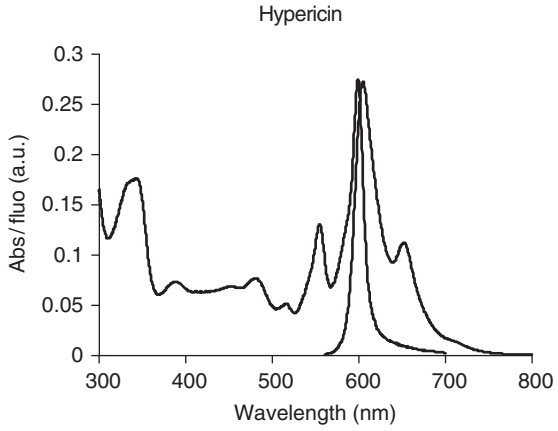


Figure 9 Formula, and absorption and fluorescence spectra of hypericin in ethanol.

toward the blue (peak at 307 nm) when tryptophan is buried in the inner hydrophobic parts of the protein. Also fluorescence polarization measurements (eqns [9] and [11]) can provide information on the degree of rigidity of the tryptophan specific molecular environment, both in physiological and denaturating conditions. Shielding effects by the macromolecular structure which tryptophan is associated with can cause a decrease of the fluorescence collisional quenching constant k_q (see Stern–Volmer equation, eqn [6]). Measurements of Φ_F and τ_F can, therefore, contribute to evaluate the accessibility of the residue to the solvent and dissolved quenchers, such as iodide or acrylamide.

Fluorescence as a Tool for Probing Biological Photoprocesses

Among important enzymes, flavoproteins constitute an important class of redox systems. The study of flavin fluorescence has provided key contributions to the study of their dynamic behavior as well as detailed information on both their catalytic mechanisms and atomic structure.

As light irradiation can elicit transient variations, and subsequent dark-regeneration, of flavin fluorescence emission, light-induced fluorescence changes (LIFCs) can be used as a spectrophotometric assay for photochemical primary reactions in light-modulated flavoproteins. Fluorescence quenching of flavin by photoinduced electron transfer can be used as a probe for angstrom-scale structural changes in the

proteic framework. Probing the fluorescence lifetime of flavins also allows observing the variation of the flavin–amino acid residue distance over time. Correlation analysis can finally reveal conformational fluctuations at multiple timescales due to the existence of multiple interconverting conformers related to the fluctuating catalytic reactivity.

The case of chlorophyll fluorescence as a "spy" of an efficiency loss in the photosynthetic process is dealt with elsewhere in this encyclopedia. Here it is to be recalled that a damage to the photosynthetic apparatus, due to any kind of environmental stress, can lower the yield of Förster energy transfer to the reaction centre (see **Figure 2**) and can impair the electron transport chain between PhotosystemII and PhotosystemI (Z scheme), thus giving rise to a more efficient radiative transition. Chlorophyll fluorescence can, therefore, be used, also by means of space-based remote-sensing systems, to monitor the healthiness of vegetation in terrestrial and aquatic ecosystems.

Plants and even microorganisms have evolved a range of sophisticated photosensing and phototransducing devices to adapt and respond to their natural habitat. A number of light-absorbing molecules enable organisms to respond to changes in the natural light environment. The responses range from the regulation of the organism motile behavior to changes in structure and form, such as seed germination, leaf expansion, stem elongation, flower initiation, and pigment synthesis. Several families of photoreceptors are known to mediate the effects of light on plant development. These include the phytochromes (phy), which monitor the red (600–700 nm) and far-red (700–750 nm) regions of the solar spectrum and two classes of blue light receptors: the cryptochromes and the phototropins.

The phytochromes, a family of soluble homodimeric biliproteins responsible for photomorphogenic responses, have a number of spectroscopic properties which make them potentially good candidates for fluorescent probe development. Particularly interesting are their long wavelength absorption maxima, high molar extinction coefficients and the ability of phytochrome apoproteins to, spontaneously and covalently, assemble with the linear tetrapyrrole chromophore precursors, phytochromobilin (PFB) and phycocyanobilin (PCB). Owing to photoisomerization, native phytochrome can adopt two photo-interconvertible configurations, termed the Pr (red light absorbing) and Pfr (far red light absorbing) forms. A direct consequence of the high yield of photo-interconversion is the very low fluorescence quantum yield for both Pr and Pfr forms ($<10^{-3}$ at room temperature). When the natural bilin prosthetic group of phytochrome is substituted with phycoerythrobilin (PEB), the phytochrome is converted into a highly fluorescent biliprotein, termed "phytofluore" (Φ_F of the order of 0.7), photostable, stable over a wide pH range, which can be reconstituted in living cells. Phytofluores represent a resourceful new family of fluorescent protein probes which can give important contributions to understand how plants perceive and transduce environmental light signals to control their growth and development.

In phototropins (phot1 and phot2), the light-activated molecular switches that regulate important plant responses to blue light (such as phototropism, blue light-induced chloroplast relocation, and blue light-induced stomatal opening), the light-absorbing chromophore, noncovalently bound, is FMN. Phototropins undergo a self-contained photocycle (see **Figure 10**), characterized by fluorescence and absorption changes induced by blue light. The generation of the signalling state, a covalent protein–FMN photo-adduct, can be monitored by fluorescence spectroscopy. A decrease in fluorescence intensity, in fact, is measured upon formation of the FMN-cysteinyl adduct, whereas during subsequent dark incubation, the fluorescence recovers completely.

DNA Photolyase (involved in the repair of DNA damaged by pyrimidine dimers formation), photoactivated adenylyl cyclase (PAC, the photoreceptor for some photomotile responses of the flagellated microalga *Euglena gracilis)* and Cryptochromes (cry1 and cry2, the photoreceptors for circadian rhythms in plants and mammals) use FAD as

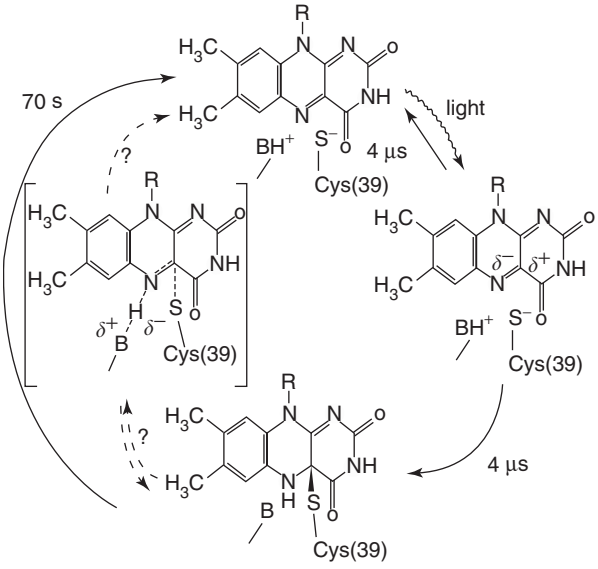


Figure 10 Scheme of FMN photoadduction reaction in phototropins.

chromophoric group. Also in these cases, FAD fluorescence dependence on pH and denaturation (in native PAC, for instance, the FAD fluorescence is strongly reduced) can valuably be used to test the photoreception process.

The fluorescence of the single tryptophan residue of the photoactive yellow protein (PYP), the photoreceptor pigment of the photomotile bacterium *Ectothiorhodospira* (*Halorhodospira*) *halophila,* efficiently monitors the structural changes of PYP which are responsible for the signalling state formation. The fluorescence spectral distribution and quantum yield of this tryptophan, in fact, depend on the local polarity, which changes during the PYP photocycle and accompanies the formation of the signalling state.

FRET is of distinctive efficiency for studying membrane model systems, such as liposomes and bilayers, biological membrane structure and processes, and protein–protein and protein-membrane interactions. FRET has been extensively used to study interactions between soluble proteins or membrane proteins and soluble protein binding partners and, also, more recently successfully employed to investigate the interaction of membrane proteins in their native environment.

Molecular interactions between natural pigments and proteins and other cellular components can be studied by means of fluorescence spectroscopy, thus clarifying if and how the pigment binds to the proteins. These data can be of great importance, if the pigments under investigation are photosensitizers generating radicals and reactive oxygen species following illumination. Wide varieties of natural and selectively artificially modified dyes are, in fact, currently utilized in the photodynamic therapy of tumors and bacterial and viral infections. A detailed study of the different fluorescence parameters (fluorescence excitation and emission spectra, fluorescence quantum yield, fluorescence lifetime, and fluorescence anisotropy) can reveal the occurrence and elucidate the characteristics of the chromophore binding to the target biological macromolecule and its localization in the target tissue/cellular compartment, thus providing key information for assessing its therapeutic efficacy or undesired phototoxicity. Hydrophobic and amphiphilic chromophores, like hypericin, for example, which give rise to non fluorescent polydispersed aggregates in aqueous solutions, monomerize and fluoresce when inserted in suitable molecular pockets.

Concluding Remarks

This article briefly: (1) presents the fundamental aspects of the fluorescence phenomenon in polyatomic organic molecules in condensed phase, (2) schematically describes the basic instrumentation for measuring fluorescence parameters (spectral distribution of the intensity, polarization, yield, and lifetime), and (3) concisely discusses a few typical examples of the use of fluorescence spectroscopy as a tool for probing structural and functional properties of biomolecules involving fluorophores such as tryptophan, flavins, chlorophyll, phytochrome, phototropins, and hypericin. The "Further reading" section will help the interested reader to find further information on the fluorescence spectroscopy and deeper and more detailed presentations of the different specialized topics.

Acknowledgments

This work was performed in the framework of a CNR–CNRS Cooperation agreement between CNR Istituto di Biofisica, Pisa, Italy and CNRS-ENS, Département de Chimie, Paris, France. F Lenci gratefully acknowledges ENS, Paris for a visiting professor grant.

See also: Fluorescent Proteins; Photosynthesis, Physics of; Single-Molecule Methods in Biophysics.

PACS: 87.64.Ni

Further Reading

Christie JM and Briggs WR (2001) Blue light sensing in higher plants. *Journal of Biological Chemistry* 276: 11457–11460.

Gensch T, Hendriks J, and Hellingwerf KJ Tryptophan fluorescence monitors structural changes accompanying signalling state formation in the photocycle of photoactive yellow protein. *Photochem. Photobiol. Sci.* (in press).

Horspool W and Lenci F (eds.) (2003) *CRC Handbook of Organic Photochemistry and Photobiology*, 2nd edn. Boca Raton: CRC Press.

Lakowicz JR (1999) *Principles of Fluorescence Spectroscopy*, 2nd edn. New York: Kluwer Academic/Plenum.

Murphy JT and Lagarias JC (1997) The phytofluors: A new class of fluorescent protein probes. *Current Biology* 12: 1–7.

Selvin PR (2000) The renaissance of fluorescence resonant energy transfer. *Nature Structural Biology* 7: 730–734.

Stryer L (1978) Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemistry* 47: 819–846.

Taroni P, Valentini G, and Marcu L (eds.) (2004) Fluorescence spectroscopy and imaging. In: *Lasers and Current Optical Techniques in Biology*, Comprehensive Series in Photochemical & Photobiological Sciences. Cambridge: Royal Society of Chemistry (in press).

Valenzeno D (ed.) Digital photobiology compendium (http://classes.kumc.edu/grants/dpc/develop/photophysics.asp).

Valeur B (2002) *Molecular Fluorescence. Principles and Applications*. Weinheim: Wiley-VCH.

Valeur B and Brochon J-C (eds.) (2001) *New Trends in Fluorescence Spectroscopy. Applications to Chemical and Life Sciences. Springer Series on Fluorescence Methods and Applications*. New York: Springer.

Vekshin NL (2002) *Photonics of Biopolymers*. Berlin: Springer.

Yan Y and Marriott G (2004) Analysis of protein interactions using fluorescence technologies. *Drug Discovery Today* 9(Suppl.N.2): S27–S31.

Fluorescent Proteins

R Nifosí, V Tozzini, and F Beltram, Scuola Normale Superiore, Pisa, Italy

Introduction

Among the 20 natural amino acids, three (tryptophan, tyrosine, and phenylalanine) contribute to the intrinsic fluorescence of a protein. This UV intrinsic fluorescence is commonly used to monitor the dynamics of protein folding, since the excitation/emission wavelength of fluorescent amino acids (mainly that of tryptophan, the one with strongest fluorescence) is sensitive to the local molecular environment. Fluorescent proteins can be engineered by covalently linking various fluorophores or by insertion of synthetic fluorescent amino acids. Extrinsic fluorescence is also achieved upon binding of fluorescent molecules, such as specific cofactors.

The focus here is on a special case of protein fluorescence, the green fluorescent protein (GFP). The intrinsic fluorescence of GFP, brighter than common protein fluorescence and red-shifted to the visible spectrum, stems from the presence of a chromophore fragment, spontaneously formed starting from an internal three-amino-acid sequence. Such a reaction is completely autocatalytic, so that the protein scaffold acts as the enzyme for chromophore formation and also provides the original fragment from which chromophore formation proceeds. Thanks to this unique feature, GFP can be functionally expressed in host cells and used to create fluorescent chimeras by fusion technology allowing noninvasive fluorescent labeling of proteins in living cells. The application of GFP in molecular biology experiments has had and is still having a tremendous impact. Such a relevance stimulated an intense effort in understanding GFP properties and tuning them by molecular engineering. It also stimulated the search for new fluorescent proteins and the subsequent discovery of several GFP homologs emitting in other regions of the spectrum. Together with the artificial GFP mutants, they constitute a broad class of intrinsic fluorescent proteins (FPs) virtually covering the entire visible spectrum.

Investigation of FPs requires a highly interdisciplinary approach, which involves biology, chemistry, and physics tools and concepts. From the perspective of condensed matter physics, on one side, visualization of fluorescent proteins down to the single molecule level stimulates progress in optical spectroscopy techniques; on the other side, the complex and intriguing photophysics of GFP is a challenge to understand the details of protein internal dynamics. These studies hold the promise of enabling novel applications that exploit these naturally evolved structures for nanoscale biomolecular photonic devices.

Natural Sources of Fluorescent Proteins

GFP was discovered in the luminescent jellyfish *Aequorea Victoria*. *Aequorea* glowing originates from Aequorin, a protein that in the presence of calcium ions catalyzes the oxidation and subsequent blue-light emission of the cofactor coelenterazine. The luminescence of *Aequorea* is however green, pointing to the existence of an energy-shifting mechanism, identified in radiationless energy transfer from Aequorin to GFP.

GFPs exist in other luminescent coelenterates, among which are hydroid *Obelia* and sea pansy *Renilla*. There, as in *Aequorea*, they convert the blue chemiluminescence of coelenterazine into green light. As recently discovered, GFP-homologous proteins are present in a variety of nonluminescent animals of the Anthozoa class, such as reef corals and sea anemones. At present, $\sim$30 members of this family have been cloned and found to share a sequence homology $\sim$20–30% with GFP. Not all members of this family are fluorescent: some are chromoproteins (CPs). Whereas they are not functionally linked to bioluminescence, these homologs are responsible for the blue-purple colorations and act as photoprotectors. Photoprotection might indeed have been the function of GFP ancestors, fluorescence being only an accessory function of the protein. Unless otherwise stated, in this article GFP refers to the original

Aequorea protein and GFPs indicate wild-type GFP (wtGFP) and its mutants.

GFP Structure

GFP contains 238 amino acids (**Figure 1a**) and weighs ~27 kDa. As discovered by X-ray studies of GFP crystals, GFP sequence folds in a compact cylindrical form, which measures 42 Å in height and 24 Å in diameter and contains several secondary structure motifs (**Figures 1b** and **1c**). This rather unique fold is called β-can or β-barrel, since its lateral wall is an 11-stranded β-sheet.

The β-can is capped on both ends by short α-helical sections and run through by an α-helix, which contains the three-amino-acid sequence from which chromophore formation originates. As a result, the chromophore is positioned right in the middle of the protein, rigidly held within the barrel and screened from the outer solvent and fluorescence quenchers such as molecular oxygen and halides. All structurally solved FPs and CPs share the same fold as GFP.

Whereas *Aequorea* GFP can exist as a monomer at diluted concentrations, *Renilla* GFP exists only in the dimeric form; DsRed is an obligate tetramer and all other homologs appear to be oligomeric. This has somehow hindered their applications in biology experiments and great effort is currently being devoted to obtain bright monomeric mutants.

Chromophores of Fluorescent Proteins

Figure 2 shows the chromophores of some proteins within the GFP family. Besides GFP and the already cited DsRed, the figure shows the chromophore of EqFP611 (a far-red FP from the sea anemone *Enctamea quadricolor*), AsCP (a red-light absorbing CP

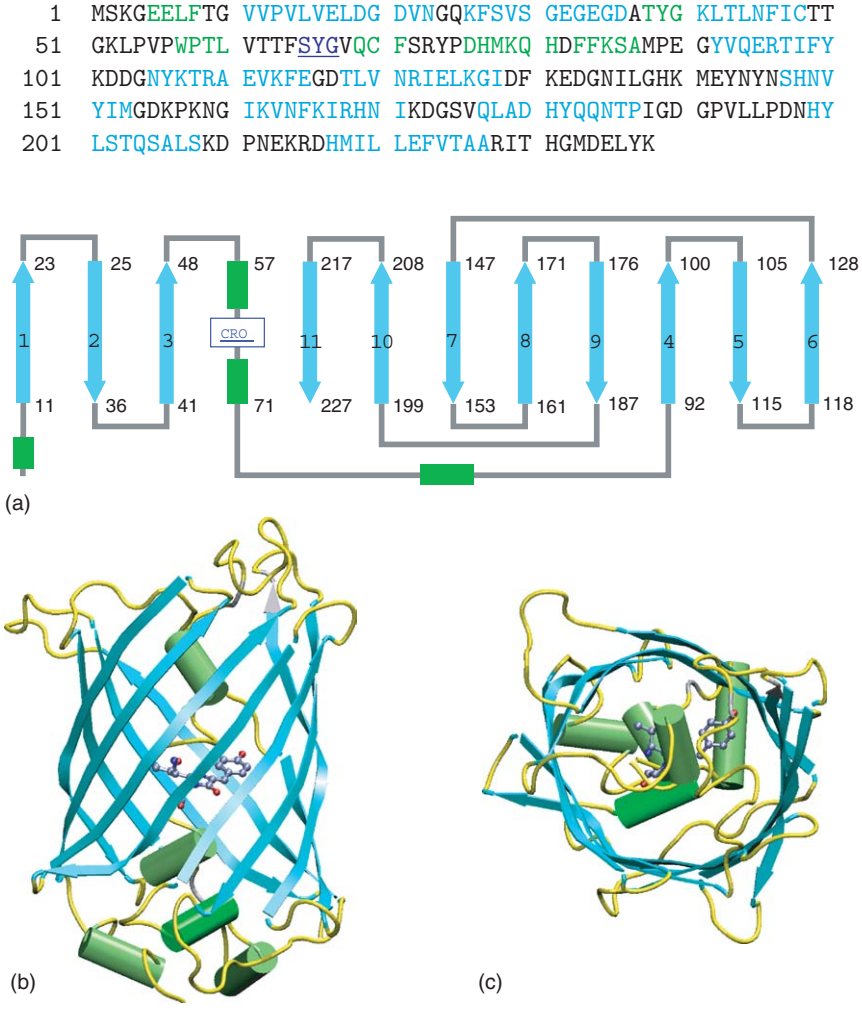


Figure 1 (a) Primary, (b) secondary, and (c) tertiary structure of GFP. In (c) the usual cartoon representation is employed, where α-helices are cylinders and β-sheets arrows. The chromophore is shown in a ball-and-sticks representation with the standard coloration for atom elements (i.e., gray for carbon, blue for nitrogen, and red for oxygen).

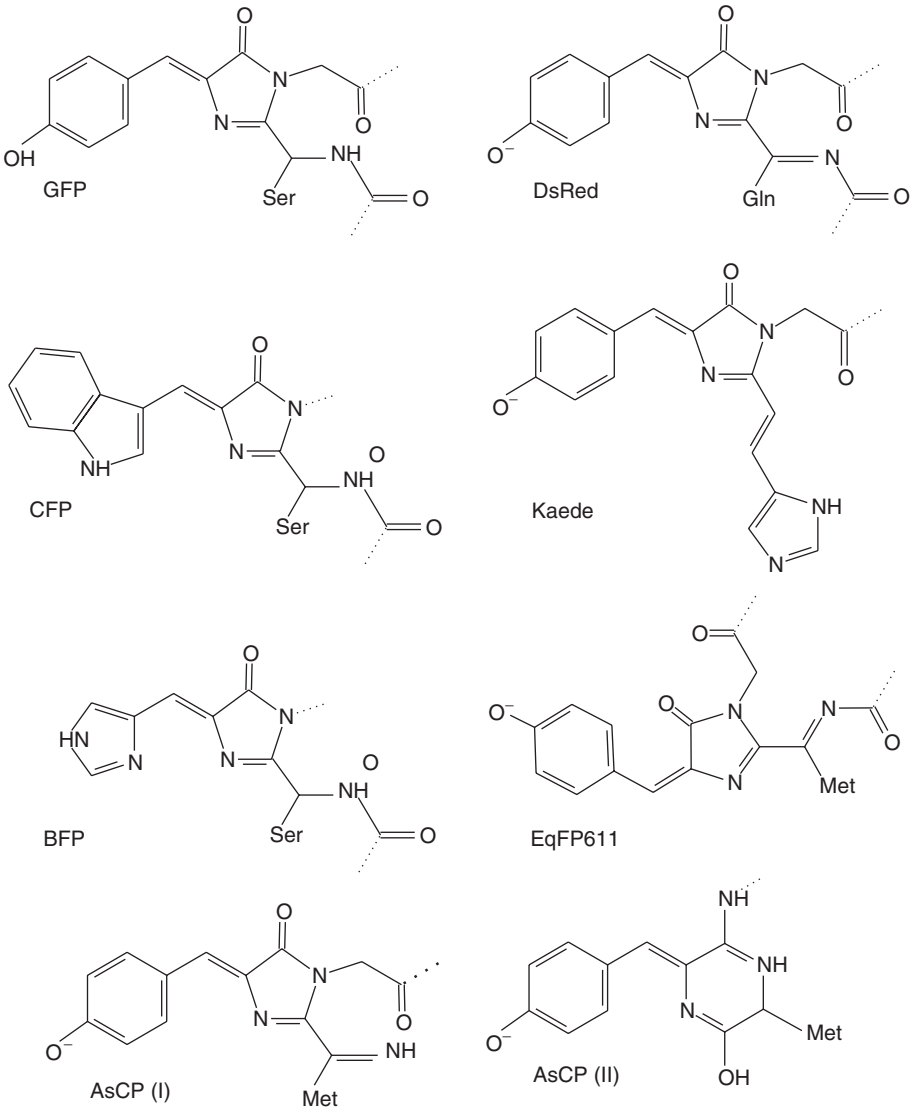


Figure 2 Structure of chromophores of fluorescent proteins. Each chromophore is labeled with the name of the corresponding FP.

from another sea anemone, *Anemonia sulcata*), Kaede (a red FP from stony coral *Trachyphyllia Geoffroyi*), and blue and cyan fluorescent proteins (BFP and CFP). The last ones are obtained by mutations of amino acid at position 66 of wtGFP, from the original tyrosine into histidine and tryptophan, respectively. EqFP611 chromophore is similar to that of DsRed but in trans-conformation. In the protein, it yields a far-red emission peaking at 611 nm, to date the most red-shifted fluorescence among the natural GFP homologs. Since no X-ray structures of AsCP and Kaede are available as yet, the structures of their chromophore are deduced only from indirect observations (in the case of ASCP, **Figure 2** reports the two hypotheses).

The different excitation/emission energies (see **Table 1**) can be explained in terms of the extension of the π-conjugated system and of the presence of electron–donor groups that compensate the electron-withdrawing carbonyl of the imidazolidinone (where present). Histidine has a moderately electron-donating NH, so excited-state transfer of electron density over the carbonyl requires higher energy with respect to tryptophan, which is rather electron rich. The phenolate anion of tyrosine in wtGFP is the most electron-donating group. The chromophore of DsRed gives even larger excitation wavelength, because the increased alternation of single and double bonds supports a more extended π-conjugated system.

The currently accepted mechanism of GFP and DsRed chromophore formation is shown in **Figure 3**. An alternative cyclization mechanism was proposed for AsCP, which leads to a cleavage of the protein in

Table 1 Optical properties of selected GFP mutants[a] together with FPs

Name	*Mutations*[b]	λ_{ex}[c] (ε)[d]	λ_{em}[e] (Φ_F)[f]
Blue (*BFPs*)			
BFP	y66h	382 (21)	448 (0.24)
EBFP	f641/y66h/y145f	383 (26.6)	447 (0.26)
Cyan (*CFPs*)			
CFP	y66w	430	480
ECFP	f641/s65t/y66w/ n146i/m153t/ v163a/	434 (33.9)	476 (0.40)
		452	505
Green			
wtGFP	None or q80r	395 (25–30)	508 (0.79)
		475 (9–14)	503
	f641	396 (18.3)	508 (0.85)
		473 (4.0)	
H9	s202f/t203i	399 (20)	511 (0.6)
	s65t	394	
		489 (52–48)	511 (0.64)
EGFP	f641/s65t	400	
		488 (55.9)	507 (0.60)
	s65g	476	510
	e222q	478	506
	t203v/e222q	496	513
Yellow (*YFPs*)			
	t203h	403	
		506	517
	t203y	404	
		510	524
E^2GFP	f641/s65t/t203y	415	515 (<0.1)
		516	525
EYFP	s65g/s72a/t203y	514 (84.0)	527
Citrine	s65g/v681/q69m/ s72a/t203y	516 (77)	529 (0.76)
Topaz	s65g/s72a/k79r/ t203y	514 (94.5)	527 (0.60)
GFP homologs[g]			
Renilla GFP	None	498 (133)	509 (0.8)
DsRed	None	558 (75)	584 (0.79)
AsCP	None	568	595 ($\ll 0.001$)
	a148s	572 (15)	597 (0.012)
mcavFP	None	508	520
		572	580
rfloRFP	None	506	517
		566	574
Kaede	None	508 (98.8)	518 (0.80)
uvKaede[h]	None	572 (60.4)	582 (0.33)
		533,508	

[a] Data from Zimmer M (2002) Green fluorescent protein (GFP): applications, structure and related photophysical behavior. *Chemical Reviews* 102: 759–781. Tozzini V, Pellegrini V, and Beltram F (2003) Green fluorescent proteins and their applications in cell biology and bioelectronics. In: Horsphool WM and Lenci F (eds.) *Handbook of Organic Photochemistry and Photobiology*, ch. 139. Washington, DC: CRC Press.

[b] Mutations are indicated with the usual one-letter code for amino acids.

[c] Peak wavelength of excitation spectrum in nm.

[d] Peak wavelength of emission spectrum in nm.

[e] Extinction coefficient.

[f] Quantum yield of fluorescence.

[g] Data taken from Labas YA *et al.* (2002) Diversity and evolution of the green fluorescent. *PNAS* 99: 4256–4261.

[h] uvKaede refers to Kaede after UV irradiation. (From Ando R *et al.* (2002) An optical maker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *PNAS* 99: 12651–12656.)

two chains that remain bound together. As demonstrated by the fact that expression of FP genes in other organisms leads to fluorescence, chromophore synthesis does not require any organism-specific enzyme. It requires, however, exogenous oxygen, in the absence of which FPs do not develop fluorescence.

The dynamics of GFP maturation and subsequent onset of fluorescence is rather slow, taking 90 min to 4 h after protein synthesis. The bottleneck of the process is chromophore dehydrogenation by molecular oxygen, which requires up to 2 h. DsRed maturation occurs in two steps: in the initial fluorescent stage, the proteins emit green light. Because of the additional dehydrogenation, formation of the mature red chromophore takes several hours and is incomplete.

As shown in **Figure 4**, inside GFP the chromophore participates to a structured network of hydrogen bonds, being surrounded by four entrapped water molecules and a surprising number of charged and polar residues, among which are Gln69, Gln94, Arg96, His148, Thr203, Ser205, and Glu222. Analogous hydrogen-bond (H-bond) interactions are found in other FP structures. These networks crucially determine the excited-state dynamics of the chromophore. Indeed, isolated model chromophores in solution are nonfluorescent at room temperature but become fluorescent when the temperature is decreased to 77 K. Accordingly, even after the chromophore is formed, denaturation of the protein results in loss of fluorescence, whereas renaturation restores the fluorescent state, indicating that in order to emit light the chromophore must be embedded in the folded β-can. Because of this property, GFP is also a useful system for the study of folding dynamics in proteins with extended β-motives.

Spectral Properties of GFP

The excitation energy of the chromophore depends on its protonation state, so that proton (H^+) equilibrium between the chromophore and the surrounding protein matrix strongly influences GFP spectral properties. In addition, H-bonds to and from the chromophore finely tune the excitation energy because they drive the electronic structure toward one or the other of the possible resonant structures.

Figure 5 shows the room-temperature absorption and emission spectra of wtGFP. In addition to the 278 nm band, which stems from aromatic amino acids (one tryptophan and ten tyrosine residues), wtGFP exhibits a broad absorption band at 398 nm that is roughly three times higher than a minor band at 475 nm. These two lower-energy bands uniquely attribute to absorption of the chromophore. Their

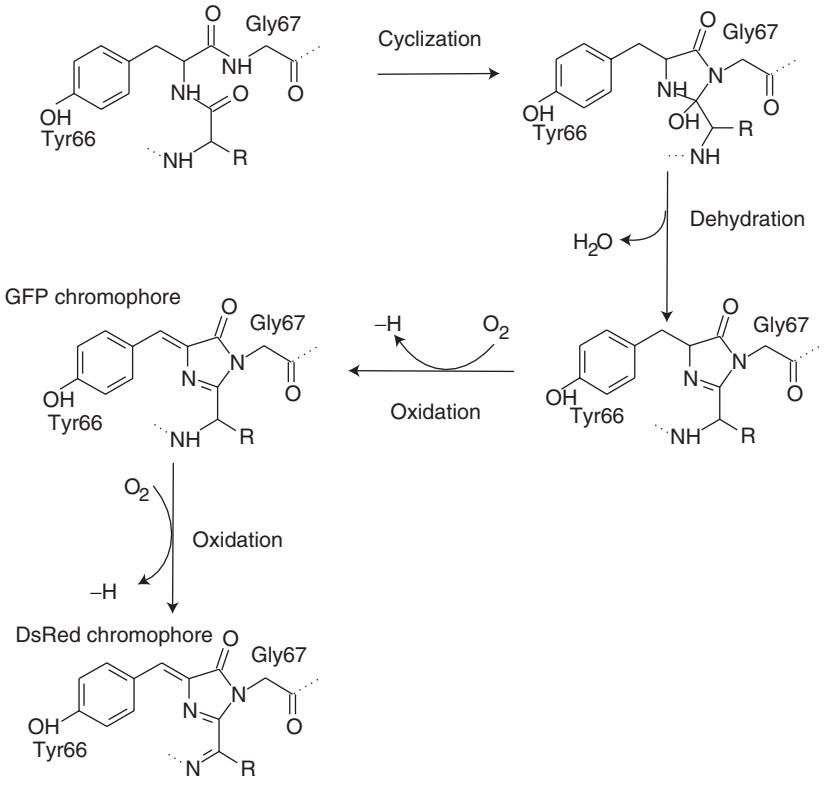


Figure 3 Accepted mechanism of formation of the GFP chromophore. Once GFP folds in a nearly native conformation, the imidazolidinone (the five-membered heterocyclic ring) is formed by cyclization, through nucleophilic attack of Gly67 on the carbonyl of residue 65, and subsequent dehydration. Then, molecular oxygen dehydrogenates the β-carbon of Tyr66, and leads to formation of the conjugated π-electronic system responsible for near-UV/visible light absorption and fluorescence. The final structure consists of two rings: the phenol, coming from the side chain of Tyr66, and the imidazolidinone, coming from the cyclization of the backbone. In the case of the DsRed chromophore, a second dehydrogenation occurs at the α-carbon of Gln66 resulting in additional extension of the π-conjugated system.

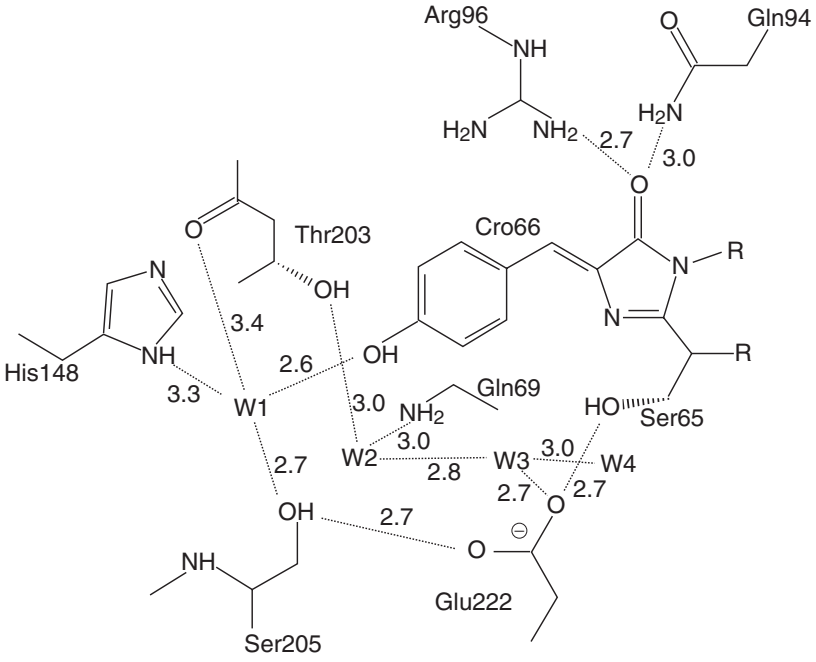


Figure 4 Two-dimensional scheme of amino acid side chains and water molecules in the immediate vicinity of the chromophore in wtGFP. The relevant distances (given in Å) are taken from the X-ray structure. (Yang F *et al.* (1996) The molecular structure of green fluorescent protein. *Nature Biotechnology* 14: 1246.)

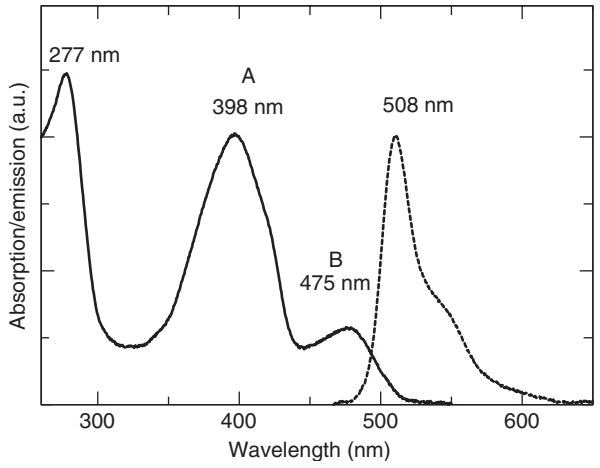


Figure 5 Absorption (solid line) and emission (dashed line) spectrum of wtGFP at room temperature and pH = 7.

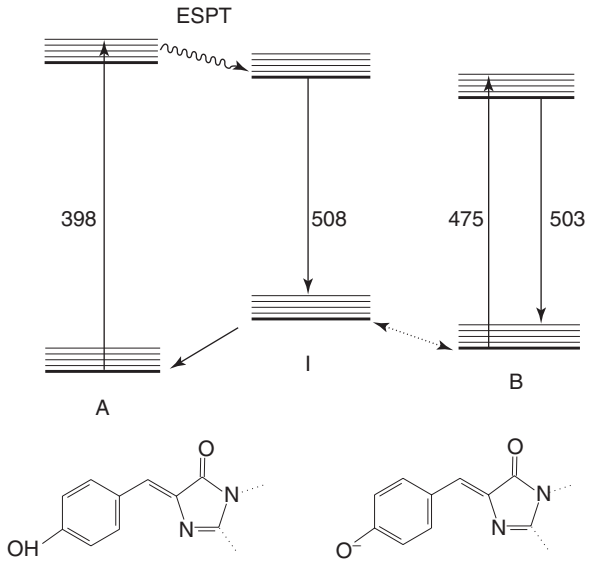


Figure 6 Fluorescence mechanism of wtGFP. The spectroscopic states are reported together with the corresponding excitation and emission wavelengths. Upon excitation of A, corresponding to the neutral chromophore, ESPT efficiently occurs to the excited I state (with anionic chromophore), which after photon emission reconverts to A. Excitation of state B (with anionic chromophore) leads instead to direct fluorescence. The equilibrium between B and I is believed to be relatively slow, because it involves the conformational rearrangement of Thr203.

relative height depends on proton concentration: at increasing pH, the minor lower-energy band increases at the expense of the higher-energy band.

This behavior arises from the ground-state equilibrium between two states of the chromophore, differing in protonation of the phenolic group from Tyr66. The phenolic oxygen of the chromophore is protonated in the state absorbing at 398 nm (state A) and deprotonated in the state absorbing at 475 nm (state B). It is largely accepted that the two other possible protonation sites (i.e., the nitrogen and the carbonyl oxygen of the imidazolidinone) in the chromophore are deprotonated in both the absorbing states, thereby giving an overall neutral chromophore in state A and anionic in state B.

Excitation of state A by irradiation at 398 nm leads to the fluorescence spectrum of **Figure 5**, peaking at 508 nm and with a rather high quantum yield of 0.79. State B (475 nm) yields a similar fluorescence spectrum, slightly blue-shifted and peaking at 503 nm. In both cases, fluorescence comes from emission of the singlet excited state of the anionic chromophore. This is suggested by a well-known tendency of phenols to become acidic in their excited state. Thus, while excitation of the anionic chromophore yields direct fluorescence, excitation of the neutral chromophore results in ultrafast (4 ps) excited-state proton transfer (ESPT) and subsequent emission of the anionic form (see **Figure 6**). ESPT is thus the origin of the large Stokes shift observed when exciting at 398 nm.

Though the anionic chromophore is the emitting species in both states, the configuration of the surrounding residues is different and the decay time of the excited state (few nanoseconds) is too short to allow equilibration, resulting in the observed slight shift of emission peak (508–503 nm). The intermediate state resulting from ESPT, called I, thus contains the anionic chromophore in an unrelaxed environment, where the hydroxyl of Thr203 points away from the chromophore phenolate, as in the state A.

Effects of Mutations on the Optical Properties

About 60 mutants of GFP have been characterized so far. Several mutation sites were selected that yield to improved maturation of GFP or to changes in the optical properties, either by shifting the absorption/emission wavelengths or by affecting brightness (extinction coefficient and/or quantum yield) or photostability. The interested reader is referred to the reviews cited in the "Further reading" section for a detailed analysis of the various mutants. Here, some significant examples of the effect of mutations on the spectral properties are given. The optical properties of selected GFP mutants are reported in **Table 1**, together with those of other FPs.

Mutations that modify the chromophore Altering the covalent structure of the chromophore results in sizable changes of the excitation/emission spectrum. In addition to mutations Tyr66His and Tyr66Trp,

which lead to the already mentioned BFPs and CFPs, synthetic amino acids such as amino-, methoxy-, iodo-, and bromo-phenylalanine were recently tried at position 66. The resulting fluorescence is blue-shifted and the quantum yield is lower than that of wtGFP ($\sim$0.4 for amino- and methoxy-phenylalanine and less than 0.02 for iodo- and bromo-phenylalanine). Red-shifted mutants were instead obtained by changing the substituent group on the CFP chromophore. Inclusion of a 4-amino-tryptophan at position 66 leads to an excitation peak at $\sim$460 nm and an emission peak at 560 nm, the most red-shifted spectra among the GFP mutants existing to date.

CFPs and BFPs are suitable partners for fluorescent resonant energy transfer (FRET) experiments with complementary GFPs. In both cases, however, additional mutations than those at position 66 are required in order to increase folding efficiency and brightness, such as in the enhanced blue and cyan variants (EBFP and ECFP). Even with those improvements, CFPs and BFPs suffer from a low quantum yield (0.24–0.42) and are prone to fast photobleaching.

Mutations that red-shift the fluorescence On the other side of the spectrum, longer emission wavelengths are obtained by replacing threonine at 203 with aromatic residues. Indeed, thanks to their delocalized π-electron system, aromatic groups are more polarizable than the aliphatic side chain of threonine and provide a more responsive medium to dipole-moment changes due to electron excitation. As a result, the emission wavelengths of mutants with Thr203Tyr are red-shifted to 525–527 nm. The crystal structures of Thr203Tyr mutants show that Tyr203 phenolic ring is stacked right below the phenolic ring of the chromophore, supporting the increased polarizability as the main mechanism for red shifting. Though the fluorescence peak is still in the green region, mutants containing Tyr203 are known as yellow because the emission appears yellowish to the eye, thanks to the long tail of the fluorescence spectrum. Yellow fluorescent proteins (YFPs) are generally brighter and more resistant to photobleaching than BFPs and CFPs.

Mutations that change the relative population of A and B: acid–base equilibrium The ground-state equilibrium between states A and B can be affected by modifying the environment of the chromophore phenolic ring (Tyr66) or of Glu222. Glu222 is commonly believed to be the other titrable group that influences GFP fluorescence (the first one being Tyr66). The titration curve of fluorescence by excitation in the 475 nm band (the one of the anionic chromophore) exhibits two abrupt increases, one at pH $\sim$ 5 and the other at pH $\sim$ 13. At pH $<$ 4 both sites are protonated, at pH between 6 and 11 only one is protonated, and at pH $>$ 13 both are deprotonated (at pH $>$ 12 the protein is denaturated and fluorescence is lost). In the intermediate region, an equilibrium is established between the two protonation states Glu222$^-$/Tyr66 (corresponding to state A) and Glu222/Tyr66$^-$ (state B). Such an equilibrium can be affected by mutations of and around the two sites.

For example, substitution of Glu222 with glycine promotes complete deprotonation of the chromophore, because glycine is necessarily neutral. The same effect is achieved when nonpolar residues are inserted at position 65, because Ser65 H-bond to anionic Glu222 is subtracted. Also mutation of Ser65 into threonine results in destabilization of state A, though threonine still possesses a H-bond donor. Presumably, because of the extra methyl, threonine side chain solvates Glu222 less efficiently than serine.

Mutations of Thr203 into nonpolar residues suppress state B, because they cannot form any H-bond with the anionic phenolate of the chromophore, which is instead H-bonded with a hydroxyl group when Thr203 is present. By contrast, state B can be favored by addition of H-bond donors close to Tyr66, such as in the Ile167Thr mutation.

Mutants with a predominant state B are preferred for applications because their red-shifted excitation wavelengths lead to a decreased background from cell autofluorescence. They are also more photostable, since the A–B photoconversion observed in wtGFP is absent, the B state being already the stable form in these mutants. A popular mutant of this class is the enhanced GFP (EGFP) with mutations Ser65Thr and Phe64Leu. The latter is a folding mutation that improves protein maturation at higher temperatures (i.e., 37°).

Spectral Properties of GFP Homologs

Less is known about the spectral properties and fluorescence mechanisms of the GFP homologs. The absorption spectrum of DsRed reveals only one band, corresponding to the anionic form of the chromophore. FRET between the chromophores in the tetramer was invoked to explain the high extinction coefficient and quantum yield. Tetramerization might also be responsible for the reduced photobleaching.

Whereas the X-ray structure of eqFP611 shows interactions between the chromophore and the protein matrix that are similar to those of GFP, the trans-conformation of the chromophore drastically

changes its environment, inserting it in a π-stacking "sandwich" between a histidine and a phenylalanine. This might be the source of the large red shift (~ 30 nm) in the absorption wavelength.

The color transitions and activation phenomena occurring in several GFP homologs are also mentioned here. Apart from the already-mentioned green-to-red transition of DsRed upon maturation, several other homologs – such as Kaede, mcavFP, and rfloRFP – were reported to undergo green-to-red transition upon UV illumination. AsCP and some of its mutants can be "kindled," that is, they become fluorescent in the red in response to intense green-light illumination, and are quenched back in the chromo form by blue-light illumination. The activation is accompanied by a 10–15 nm red shift in absorption. Stable fluorescent forms were obtained by site-directed mutagenesis. Chromophore cis–trans isomerization was proposed as the mechanism of kindling.

Fluorescence Dynamics and Dark States

Because of the interplay between the chromophore and the surrounding residues, several configurational substrates couple to the possible protonation forms of the chromophore, giving rise to many intriguing photophysical and photochemical processes. ESPT, already mentioned above, is an example (actually, the first) of the several effects discovered by studying the fluorescence dynamics of FPs. A new class of phenomena was found by applying fluorescent correlation spectroscopy (FCS) to FPs in solution and single-molecule microscopy to individual proteins immobilized in aqueous polymer gels. These studies revealed intensity fluctuations, blinking, and fluorescence switching, which were completely hidden in ensemble measurements, thus providing further insight into GFP photophysics.

Fluorescence Correlation Spectroscopy of GFP and DsRed

FCS is based on the analysis of fluctuations in fluorescence emission. In the case of free fluorophores in solution, fluctuations arise either from the change in the number of emitting molecules as they randomly diffuse in and out of the focal volume (order of 10^{-15} l $= 1\,\mu m^3$) or from intramolecular transitions between different states. From FCS experiments, it is possible to extract the decay times associated with diffusion and with the transitions from fluorescent to nonfluorescent (dark) states. The diffusional relaxation limits the timescale of detectable processes to submillisecond scale. In this range, wtGFP and other investigated mutants display "fast" fluctuations named flickering. Flickering at the 100 μs timescale probably stems from transitions between the anionic and the neutral state of the chromophore. Such transitions can occur either by protonation from the external solvent (pH-dependent component of flickering) or from the internal residues (light-driven component). In addition, intersystem crossing to the triplet state might play a role for the fastest fluctuation $\sim 50\,\mu s$. Also DsRed does exhibit several flickering components, though none seems to be dependent on pH, supporting the view that DsRed chromophore is always in the deprotonated form. Understanding the molecular mechanisms of DsRed flickering is challenging since energy transfer among the units of the tetramer is involved in the process.

Single-molecule Behavior

GFP fluorescence is bright enough to be observed at the single-molecule level. Observation of single GFP molecules dated to 1997 and was achieved independently by Moerner and co-workers and by Vale and co-workers. Sudden bursts of fluorescence were observed separated by dark intervals of the order of some seconds. Unlike the dark times, the bright times depend upon irradiation intensity (they become shorter at increasing laser power), so that fluorescence termination is light driven, while conversion back to the bright state is spontaneous. DsRed

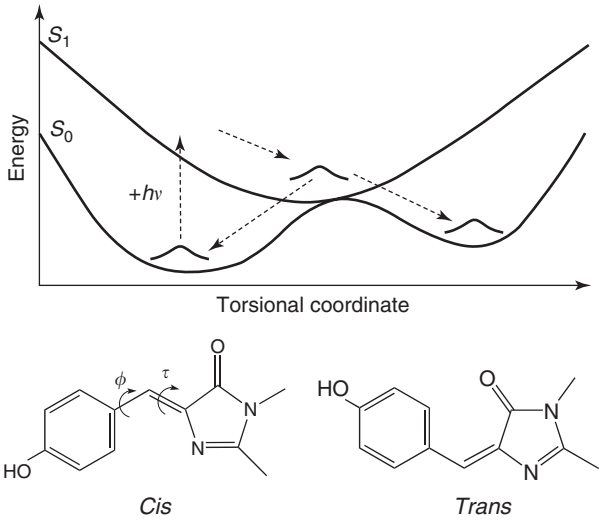


Figure 7 Scheme of a photoisomerization process. The absorption of a photon of energy $h\nu$ brings the (nuclear) wave packet to the S_1 state PES. Upon dynamics on S_1 PES, it encounters a conical intersection from which it relaxes to the ground-state energy surface, either in the same configuration as the starting one or in a different isomer. Energy curves are plotted qualitatively against a hypothetical reaction coordinate with main contribution from the torsion around the exocyclic bond. Cis- and trans-configurations of the chromophore are shown in the bottom part of the figure with the two possible torsion angles τ and ϕ.

single-molecule behavior also includes blinking events. Each single molecule is a tetramer, and within each tetramer all combinations of mature red and immature green monomers are possible. Energy effectively transfers between the various units, and emission proceeds from any of the four fluorophores.

After few cycles of blinking, all GFP molecules undergo photobleaching, that is, irreversible transition to a dark state. The relatively short lifetime of these fluorophores limits single-molecule applications in biology experiments that require detection limits longer than few tens of seconds. A very promising property of some YFP mutants is their ability to be photoconverted back to the bright state after seemingly irreversible photobleaching. The E^2GFP mutant (see **Table 1**) can be "switched on" with 350 nm light irradiation, once the reversible dark state is reached upon intense 476 nm irradiation.

It is challenging to determine the nature of the dark states associated with blinking and "on/off" switching, and the molecular pathways for bright-to-dark transition and vice versa. Switching seems to occur only in mutants with Thr203Tyr and Thr203Phe. Variants that embody Thr203Tyr mutation have a poor quantum yield of fluorescence upon excitation of the state A, so that both intrinsic fluorescence and ESPT are suppressed by some competing nonradiative decay mechanism. One such mechanism for the GFP chromophore is cis–trans photoisomerization, an ultrafast excited-state process that involves torsion around the bonds connecting the two rings of the chromophore. This dynamics (see **Figure 7**) involves crossing from the excited state to the ground-state potential energy surface (PES) through a conical intersection, resulting in complete quenching of photon emission. Cis–trans photoisomerization is known to play an important role in the retinal of rhodopsin and is probably the origin of the extremely poor quantum yield of the free GFP chromophore in solution. In this line of thought, it is tempting to consider cis–trans photoisomerization as the molecular mechanism of E^2GFP switching, where the reversible dark state of the protein is attributed to the trans-configuration of the chromophore.

Applications of FPs

High-resolution optical-spectroscopy techniques applied to the imaging of proteins labeled with fluorescent tags offer new tools for investigating protein trafficking and interactions inside the living cell. In this context, GFP fusion technology offers an

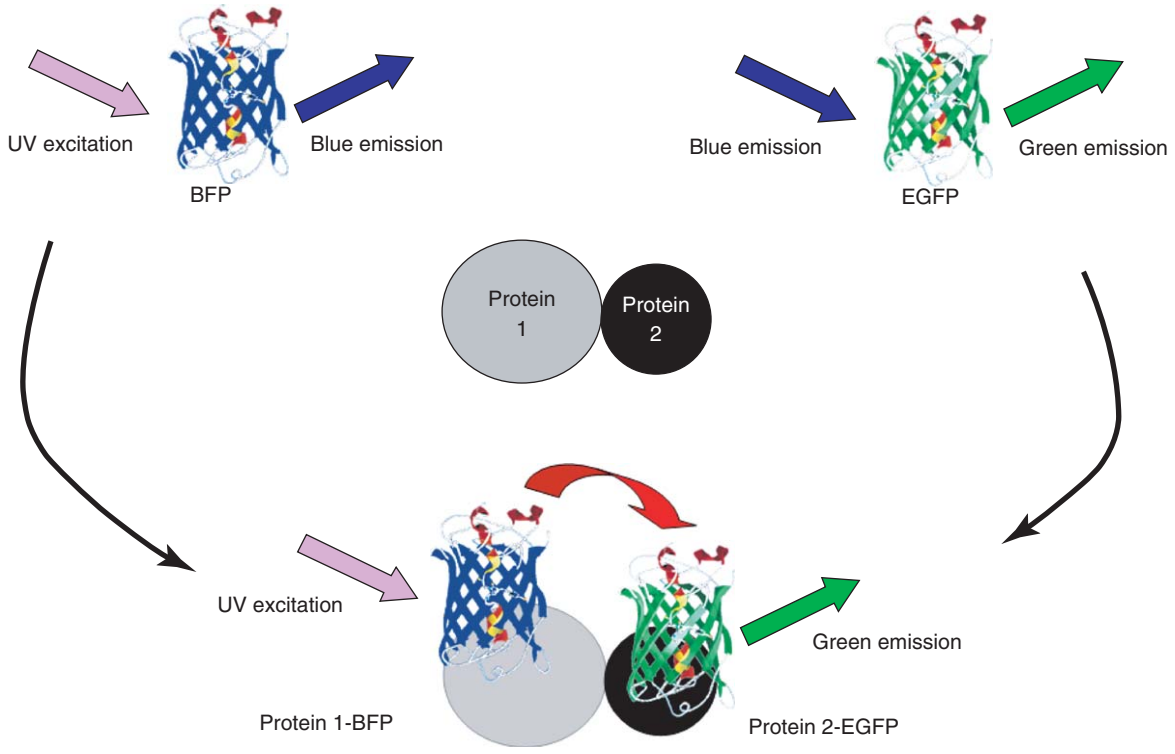


Figure 8 A typical FRET experiment with FPs would employ, for example, EGFP and EBFP as FRET partners (EYFP and ECFP is also a common pair), fusing them with the proteins of interest. Interaction between the two proteins is then detected by comparing the EGFP emission at 520 nm excited at (direct EGFP fluorescence) with emission by excitation of EBFP at 360 nm (EGFP fluorescence via FRET). Comparable levels of fluorescence in the two cases prove that the pair is interacting.

inexpensive way to mark one protein with one-to-one protein-tag correspondence, avoiding toxicity associated with exogenous fluorescent dyes. In addition, the availability of a large set of GFP mutants is precious in that it offers several pairs of fluorophores with suitable overlap between absorption and emission spectra, ideal for FRET experiments. Detection of FRET between donor and acceptor fluorophores previously fused to a pair of proteins indicates that the pair is actually interacting at distances of the order of the nanometer (see **Figure 8**). Indeed, thanks to FRET sensitivity to the pair distance, this technique discriminates between interaction and simple colocalization in cellular compartments.

FRET between a GFP pair interposed by a peptide is also exploited to monitor conformational rearrangements, either in protein folding studies or for sensing calcium concentration. GFP-based calcium reporters, the so-called Cameleons, are composed of a GFP pair linked by a peptide that changes conformation in the presence of calcium. Other sensing applications exploit the fact that GFP fluorescence is (or can be made by mutagenesis) sensitive to pH and is quenched upon binding of halides or metal ions.

Photobleaching can be used to investigate protein dynamics. The mobility of GFP chimeras is determined by high-intensity illumination of a defined cellular area and subsequent monitoring of the fluorescence recovery after photobleaching (FRAP) within the same area. Fluorescence loss in photobleaching (FLIP) works in the opposite way, that is, by monitoring the loss of fluorescence outside the repeatedly bleached area.

There exist also potential applications of GFPs outside the cell. As mentioned above, it is possible to toggle the molecular states of certain GFP mutants by purely optical means (on/off switching). Such photoinduced bistability can be the basis for information storage down to the single molecule level. Thanks to molecular recognition, FPs are easily targeted to specific sites of a patterned substrate for building ultradense optical memory devices.

Applications of the other FPs are still at the embryo stage. Tuning the optical properties of these new proteins by mutagenesis will extend the spectral range of biological fluorophores available for applications. New building blocks for photonic devices might emerge from the investigation of their largely unexplored photophysics and the increasing ability to manipulate protein structure by mutagenesis and insertion of synthetic amino acids.

See also: Bioelectronics; Fluorescent Biomolecules; Rhodopsin and the First Step in Vision; Single-Molecule Methods in Biophysics.

PACS: 33.50. – j; 82.37.Ps; 82.37.Vb; 87.15.He; 87.15.Mi

Further Reading

Chalfie M and Kain S (eds.) (1998) *Green Fluorescent Protein: Properties, Applications and Protocols*. New York: Wiley-Liss.

Conn PM (ed.) (1999) *Methods in Enzymology, Volume 302: Green Fluorescent Protein*. New York: Academic Press.

Moerner WE (2002) A dozen years of single-molecule spectroscopy in physics, chemistry and biophysics. *Journal of Physical Chemistry B* 106: 910–927.

Rothschild KJ, Gite SS, Mamaev S, and Olejnik J (2003) Building photonic proteins. In: Horsphool WM and Lenci F (eds.) *Handbook of Organic Photochemistry and Photobiology*, ch. 133. Washington, DC: CRC Press.

Sullivan K and Kay S (eds.) (1999) *Methods in Cell Biology, Volume 58: Green Fluorescent Proteins*. New York: Academic Press.

Tsien RY (1998) The green fluorescent protein. *Annual Review of Biochemistry* 67: 509–544.

Verkhusha VV and Lukyanov KA (2004) The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nature Biotechnology* 22: 289–296.

Zumbusch A and Jung G (2000) Single molecule spectroscopy of the green fluorescent protein: A critical assessment. *Single Molecules* 1: 261–270.

Folding and Sequence Analysis

A Tramontano, University of Rome "La Sapienza," Rome, Italy

Introduction

A genome is a string made of four different characters, the nucleotides. The human genome consists of a collection of three billion of these nucleotides, and it is believed to code for at least one hundred thousand different functions. Since *Homo sapiens* is only one of the many species whose genomic sequence is known, it is clear that the amount of available data is impressive. However, handling such a large amount of data is not a problem; many other sciences need to cope with equivalently large data sets, and powerful techniques and computers are available to handle them. The problem rather arises from the peculiarities of biological data, which are of different types,

ranging from sequences, that is, linear arrays of monomers, to three-dimensional structures to their interactions to functional assignments. The real hurdle is indeed to try and extract functional information by correlating the many diverse sets of information. This is not only an intellectual challenge, but is also of enormous practical importance, since it is the basic requirement to try and interfere rationally with these functions for medical and biotechnological applications.

A function is the result of the action of interacting atoms positioned in a specific relative position, so that they can, for example, bind a ligand, catalyze a biochemical reaction, recognize foreign substances, regulate the transcription or translation of other proteins, and build scaffolds to achieve cellular organization and compartmentalization. These mechanisms can achieve an exquisite specificity and sensitivity in response to internal and external stimuli, allowing a living organism to survive and evolve.

The large majority of these functions is performed by proteins, linear polymers of 20 different amino acids whose sequence is specified by the genetic material. Proteins fold to a unique three-dimensional structure, which determines the relative position of their atoms and, therefore, their function. For example, the biological function of an enzyme is to catalyze a biochemical reaction by lowering the free energy of the activation barrier. This is achieved through the stabilization of the reaction intermediately brought about by a few specific amino acids located in a very precisely defined relative position in the protein (**Figure 1**).

The structure of a protein is by and large only dictated by the specific linear sequence of its building blocks as first demonstrated by Anfinsen in a historical experiment.

He showed that a protein (in this particular case, ribonuclease S) once denatured, that is, unfolded, *in vitro*, recovers its native conformation when the

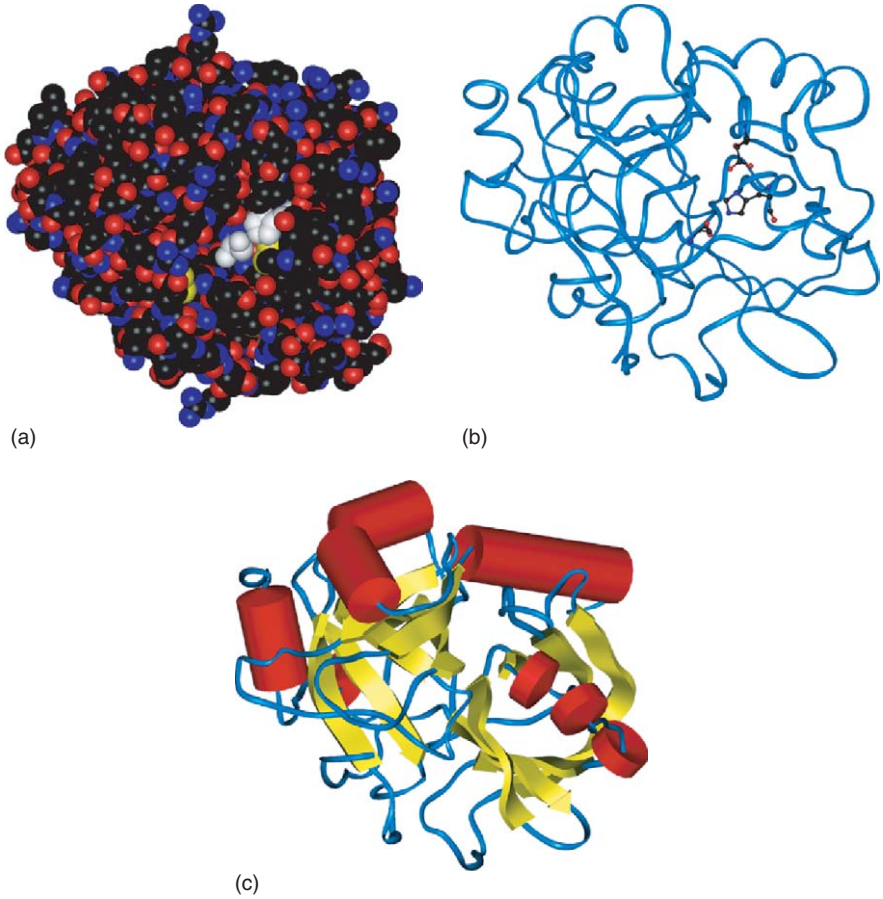


Figure 1 Different representations of the three-dimensional structure of the enzyme thrombin. (a) All atoms of the enzyme are shown as spheres. The white spheres indicate the atoms involved in the catalysis. (b) The enzyme is shown as a ribbon connecting its amino acids and the three amino acids involved in the catalytic mechanism are shown in ball and stick representation. They are distant in the amino acid sequence, but come close together upon folding of the protein, (c) Regular elements of secondary structures are shown as cylinders (α-helices) or arrows (β-strands).

denaturing agents are removed from the test tube. This implies that the information about the three-dimensional structure of a protein is contained in its amino acid sequence.

The subsequent discovery of cellular mechanisms that catalyze folding of some proteins has left this view unchallenged, since these systems accelerate the folding process, but do not affect the structure of the final native state.

Therefore, a function is directly linked to the genetic material via the encoded proteins and their three-dimensional structure.

One needs to understand how a given genomic sequence codes for a protein, how the protein folds into its defined three-dimensional structure, and how this structure is able to carry on a function either alone or through its interactions with other partners.

In higher organisms, the detection of the segments of a genome that code for a given protein (genes) is not straightforward. In humans, only three percent of the genome codes for proteins and the portions of the genes coding for the corresponding protein sequences are not contiguous in the DNA sequence. The gene-finding problem is not discussed here; it suffices to say that methods to identify the coding regions are still being developed and that the problem is not yet satisfactorily solved.

The Protein Folding Problem

Once the amino acid sequence of a protein is known, one faces the problem of determining its three-dimensional structure, a problem that can be addressed with experimental techniques, such as X-ray crystallography or nuclear magnetic resonance. However, it is unthinkable that it is possible to determine experimentally the structure of all the proteins of the universe. Therefore, one of the most important problems of modern biology, often referred to as "the holy grail" of computational biology, is to be able to infer (or predict, as we erroneously say) the three-dimensional structure of a protein from its amino acid sequence.

A protein folds if the difference in free energy between its disordered, unfolded state and its structured native state is negative. Upon folding, a protein loses the entropy associated with the many alternative conformations that it can assume when disordered, and gains internal interactions and entropy. Interactions in a protein structure are weak, mainly van der Waals and electrostatic interactions (the latter are usually few and quite weak, since most proteins are embedded in a polar solvent). Entropy gain is due to the burial of the hydrophobic side chains of amino acids. The main problem is that the difference between the free energy of the unfolded state and that of the structured state derives from the difference of large terms, but is a very small number, of the order of a few kcal/mol. In other words, proteins are only marginally stable and one would need to calculate the energetic contribution of each interaction with a precision that is not achievable today. Therefore, the search for the free-energy minimum conformation of a protein, that is, the *ab initio* solution of the folding problem, is beyond one's present capabilities.

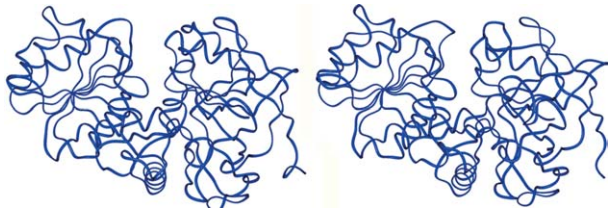


Figure 2 Structure of two evolutionary related proteins (human and cod alcohol dehydrogenase). Their sequences share 60% identity but their main-chain structure is virtually identical.

Computational biology is, therefore, trying to find a heuristic solution that, although not satisfactory from the point of view of an intellectual understanding of the problem, is able to provide practically useful solutions.

One of the most interesting approaches widely used today is to address the problem in its differential form. In other words, rather than asking the question of which is the structure and function of a protein, the following question can be asked: given a protein of known structure and function, how do changes in its sequence affect its structure and its function(s)?

Evolution provides one with many examples of proteins that descend from a common ancestral protein and whose structure and function have been modified via a process of sequence changes and selection. One can, therefore, use these solved examples to try and infer the rules relating the relationship among sequence, structure, and function changes (**Figure 2**).

The Protein Sequence Space

Two protein sequences are linear arrays of basic building blocks, the amino acids, and their distance can be measured by counting the number of evolutionary events (amino acid substitutions, for example) that can transform one into the other.

It is reasonable to use a maximum parsimony approach by calculating the correspondence between the two linear arrays of amino acids that minimizes the number of differences between them, a problem

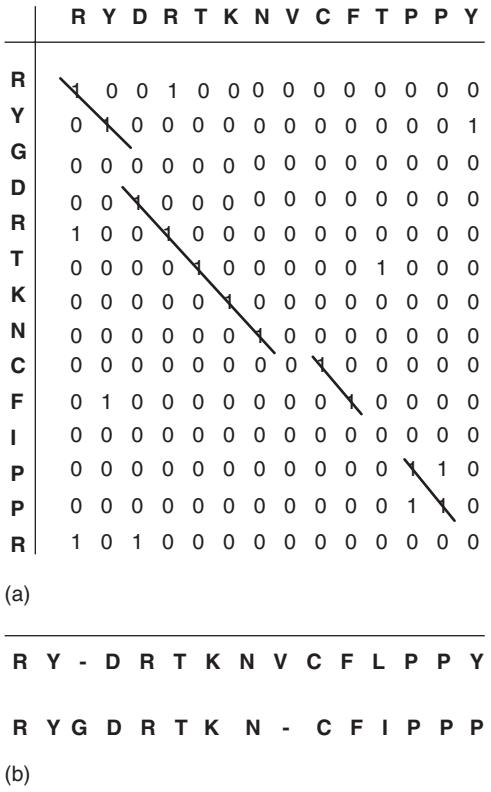


Figure 3 (a) Example of a matrix indicating the identity between two protein sequences. The maximum likelihood aligned regions between the two protein sequences are represented by the segments and the corresponding optimal overall alignment is shown in (b).

that can be solved exactly, for example, with dynamic programing.

Figure 3 illustrates the method. A matrix is used, where each column represents one amino acid of the first sequence, and each row an amino acid of the second. One element of the matrix is set to 1 if the amino acid in the column and that in the row are the same, and to zero, otherwise. The problem of finding the correspondence between the amino acids of the proteins is equivalent to finding the path that goes from the upper left corner to the lower right corner, such that the sum of the content of the cells it connects is maximal.

In order to obtain a biologically sensible alignment of two protein sequences, one more level of sophistication is required. Two amino acids are not simply the same or different, they can be more or less similar according to their physicochemical properties. The probability that a charged amino acid is replaced by another charged amino acid is expected, and found to be higher than the probability that it is replaced by a small hydrophobic amino acid. These probabilities can be estimated by analyzing families of evolutionary related proteins, and counting how frequently a given amino acid substitution is observed with respect to a background distribution. These numbers are derived from closely related proteins or well-conserved regions of proteins, where changes can be detected without ambiguity and subsequently, one would need to extrapolate their values for distantly related proteins. This can be done in different ways but in all cases, one then faces the problem of having to assign probabilities also to some phenomena, which are rarely observed in closely related proteins or in well-conserved regions, such as insertions or deletions of residues. This is usually done using empirical methods.

Improvement in the quality of sequence alignments, that is, in their ability to reproduce the course of evolution, can be achieved by taking advantage of the availability of more sequences from the same evolutionary family. It can be said that, with some limitations, one knows how to measure the distance between two evolutionary related protein sequences.

Given what has been said before, the accuracy is higher for closely related proteins and it can become very low when the distance between the two proteins is very large.

The Protein Structure Space

Given two protein structures and a correspondence table between their atoms, the definition of their distance in structural space is trivial; one needs to find the rigid body translation and rotation that minimize the distance (usually the root mean square distance (RMSD)) between corresponding atoms.

The problem becomes far more complex when the correspondence between the atoms is unknown and/or incomplete (that is, when some regions of one protein might not correspond to any region in the other), and several methods have been developed for this purpose.

First of all, it should be clear that the solution to the problem is not unique, since different superpositions can include different numbers of atoms and result in different values of RMSD (**Figure 4**).

The most commonly used methods rely on either substructure superpositions or on distance matrix analysis. The latter is an interesting, and often very useful, representation of a protein structure where the protein sequence is written both in the first column and first row of the matrix, and a cell of the matrix is filled if the two corresponding amino acids are within a certain threshold distance. Distances can be calculated between Cα atoms, Cβ atoms, between the center of mass of the side chains, or even between the midpoints of secondary structure elements. The matrices corresponding to two protein structures can

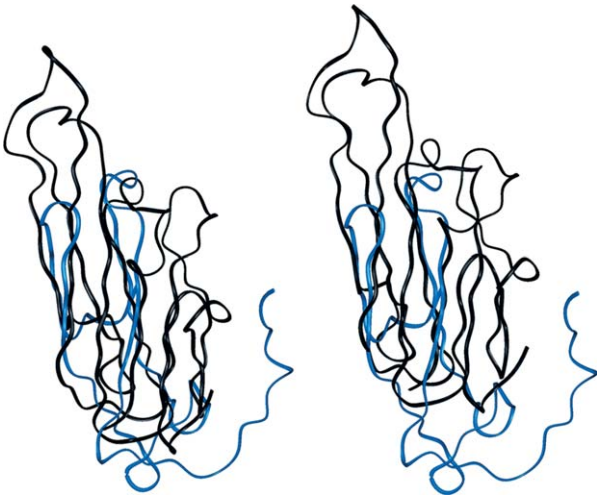


Figure 4 Two alternative structural superpositions of comparable quality of tumor necrosis factor-α (in black) and of the middle component of the bean pod mottle virus (in blue).

be compared to find the optimal superposition between their structures, for example, with a "divide and conquer" algorithm.

The problem of the nonuniqueness of the computation of the superposition, and consequently, the uncertainty in the measure of the distance between two protein structures, becomes clearly more relevant as the more different the two protein structures the more distantly related the two proteins are.

The Protein Function Space

In the previous paragraphs, the problem of defining distances between well-defined entities, sequences, and structures was faced. In function space the problem becomes much more complex, because it is very difficult to univocally define function.

Not only can the coding of a function be ambiguous, but the protein function is, by nature, multidimensional. A molecule has a molecular, a biological, and a cellular function, and each of them can be defined at different levels of detail.

For example the protein "thrombin" can be functionally defined as:

- an enzyme, that is, a biological catalyzer;
- a hydrolase, that is, an enzyme catalyzing the hydrolysis of a chemical bond;
- a protease, that is, a hydrolase that hydrolyzes the peptide bond of proteins;
- a serine protease, that is, a protease where the catalytic reactive group is the side chain of the amino acid serine;
- a trypsin-like serine protease, that is, with a specific topology that allows the clustering of the catalytic site residues; and
- thrombin, that is, a specific trypsin-like serine protease.

At the same time, it has the biological function of being part of the blood coagulation cascade and the cellular function of being an extracellular protein.

How does one then define the distance between the thrombin function and, say, the function of polymerase, the enzyme that catalyzes the replication of a nucleic acid molecule?

Finding an exact solution to the problem is probably unfeasible, but at least there are now attempts to systematically classify the functions through the production of controlled vocabularies that can be used to describe gene product attributes, such as the system under development by the gene ontology (GO) consortium.

The Relationship between Sequence Space and Structure Space

As mentioned, a native protein structure is stabilized by a large number of weak interactions; therefore, the probability that small accepted changes, such as those happening during evolution, can change the overall structure substantially is very low. This implies that evolutionarily related proteins, that is, proteins that derive from a common ancestral protein via the stepwise accumulation of small changes will have similar structure. The buildup of the effects of these small changes might change the details of the structure quite substantially, but it is not expected to modify the overall topology of the protein chain.

On the other hand, if two proteins derive from a common ancestor via small changes, their sequences might preserve a similarity sufficiently high to highlight their evolutionary relationship.

One can derive that proteins with similar sequences have similar structures. This observation, that has been quantitatively analyzed by several authors, forms the basis of one method for protein-structure prediction, called "modeling by homology" or "comparative modeling." Blind tests of the method, evaluating the predictions of a protein structure before its experimental determination, demonstrate that the accuracy of the method can be rather high for evolutionarily closely related proteins.

It should be mentioned that the relationship between sequence and structure in proteins is more complex; often proteins with seemingly unrelated sequences have similar topologies. This might be due to the fact that some topologies are favored and, therefore, many protein sequences have structurally converged toward them, or to the fact that they are indeed evolutionarily related, but the number of

accumulated changes is too high for the relationship to be detectable.

This observation has led to a method for protein-structure prediction, called "fold recognition" whereby a fitness value is calculated for the compatibility of a given protein sequence with each of the previously known topologies.

The accuracy of the predictions is somewhat lower for this method than for comparative modeling, but in both cases the question is, which level of accuracy does one need to be able to use the prediction to define the function.

The Relationship between Sequence Space and Function Space

The most common assumption in relating sequence to function is that proteins evolving from a common ancestor preserve their function, implying that functional information can be transferred from one protein to evolutionarily related ones. However, complications arise since two homologous proteins can be diverged after a speciation event and, therefore, preserve the same function (orthologous), or after a duplication event (paralogous) could have developed new functions. The problem, therefore, is to understand how their distance in sequence space can be used to discriminate between the two cases or, in other words, which minimum level of sequence identity between two sequences is required to safely conclude that the function is preserved. Several studies have addressed this problem concluding that, in general, a sequence identity of 50% or higher is strongly correlated with function conservation.

The Relationship between Structure Space and Function Space

Until a few years ago, the determination of the structure of a protein was carried out by groups interested in the specific protein and, therefore, on proteins that had been extensively biochemically characterized.

More recently, there are several worldwide projects, the so-called structural genomics projects, devoted at solving experimentally the structure of as many proteins as needed to "cover" the structure space, that is, at solving preferentially the structure of proteins unrelated to any protein of known structure. Furthermore, the genomic projects have made available a large number of putative protein sequences.

This implies that more and more often, structural information on proteins of unknown function is available and the question arises as to whether the latter can be deduced from the former.

The problem is very complex, partly because a measure of "distance" between functions cannot be defined and partly because proteins develop novel functions not only through the evolution of an existing function (which is often achieved after the genetic material encoding for the original function is duplicated, and can therefore evolve without the constraints of preserving it), but also via recruitment of structural modules and mixing of pre-existing protein domains (structurally compact segments of proteins).

There are several examples that illustrate the difficulty of the problem. For example, the same identical protein is used in the duck as the eye lens protein and as an enzyme that hydrolyzes lactose in the same organism, but in a different cellular environment. Clearly, the knowledge of the sequence and structure of the protein, identical in the two cases, would be of no help to deduce the function.

Furthermore, it is as yet unclear how unique a protein function is; in other words, how often the same protein can perform different functions in different cellular compartments or in the presence of different ligands.

Present approaches to structure-based function prediction rely on methods to detect common substructures in proteins that can be associated with specific activities, in the hypothesis that important functional features are structurally conserved during evolution. On the experimental side, the development of whole-organism investigations permits other approaches to function prediction. Novel high-throughput experiments, namely transcriptomics and proteomics experiments, allow the detection of proteins that are coregulated in response to an external stimulus or that interact with one another.

These data, combined with the recognition of conserved sequence and/or structural features in coregulated proteins and of correlations between occurrences of related proteins in different organisms, can give strong indications about functional properties and are, therefore, slowly but steadily enlarging our catalogue of sequence, structure, and function relationships.

See also: Fluorescent Biomolecules; Fluorescent Proteins; Luminescence.

PACS: 87.15. – v

Further Reading

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr., Brice MD, *et al.* (1977) The protein data bank: a computer-based archival

file for macromolecular structures. *Journal of Molecular Biology* 112: 535–542.

Branden C and Tooze J (1999) *Introduction to Protein Structure*, 2nd edn. New York: Garland Publishing.

Finkelstein AV and Ptitsyn OB (2002) *Protein Physics*. London: Academic Press.

Goldsmith-Fischman S and Honig B (2003) Structural genomics: computational methods for structure analysis. *Protein Structure* 12: 1813–1821.

Jones DT (2000) Protein structure prediction in the post genomic era. *Current Opinions in Structural Biology* 10: 371–379.

Lander ES, Linton LM, Birren B, Nasbaum C, Zody MC, *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.

Lesk AM (2001) *Introduction to Protein Architecture*. Oxford: Oxford University Press.

Lesk AM (2002) *Introduction to Bioinformatics*. Oxford: Oxford University Press.

Onuchic JN and Wolynes PG (2004) Theory of protein folding. *Current Opinions in Structural Biology* 14: 70–75.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, *et al.* (2001) The sequence of the human genome. *Science* 291: 1304–1351.

G

Genetic Algorithms for Biological Systems

K Sneppen, Niels Bohr Institutet, Copenhagen, Denmark
S Maslov, Brookhaven National Laboratory, Upton, NY, USA

Introduction

Cells are controlled by the action of molecules upon molecules. Receptor proteins in the outer cell membrane sense the environment and may subsequently induce changes in the states of specific proteins inside the cell. These proteins then interact again and convey the signal further to other proteins and so forth, until some appropriate action is taken. The final results of such signaling may be a transcription regulation, thereby making more of some kinds of proteins.

The reader is reminded that proteins are produced from DNA through a two-step process, where first the DNA code is transcribed into mRNA by an RNA polymerase, and subsequently, the mRNA is translated into a protein by a ribosome. This is illustrated in **Figure 1**. The simplest way that one protein regulates the production rate of another protein is illustrated in **Figure 2**. The regulated protein may in turn regulate other proteins, and thereby be part of the transcription regulatory network. Regulatory genetic networks are essential for epigenetics and thus multicellular life, but are not essential for life. In fact, there exist prokaryotes with nearly no genetic regulation.

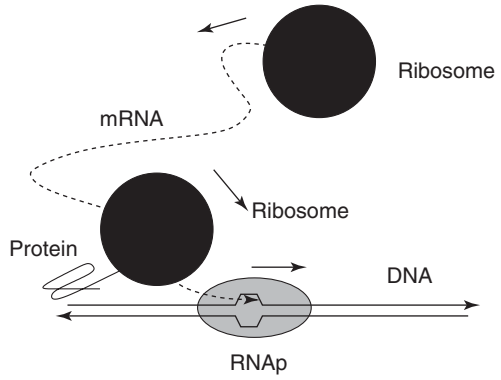


Figure 1 RNA polymerase (RNAp) moves along the DNA while transcribing it into an mRNA string. This mRNA is subsequently translated into a protein by a ribosome, here shown in black. The same mRNA may easily be translated into several proteins.

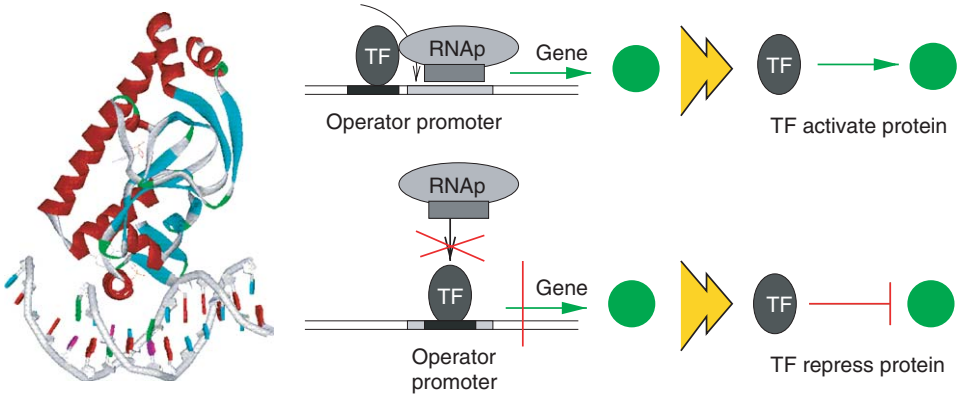


Figure 2 On left panel an example of a transcription factor is shown, the protein CAP that binds to a piece of DNA. On right panel is illustrated respective positive and negative regulation by a transcription factor (TF) on a protein shown as a green filled circle. The TF is a protein that binds to a region on the DNA called an operator (dark region on the DNA). The positive regulation is through a binding between the TF and the RNAp, that increases the chance that RNAp binds to the promoter that is shown as a medium dark region on the DNA strand. Negative regulation occurs when the operator is placed such that the bound TF prevents the RNAp from binding to the promoter. On the rightmost panels is shown how one typically draws the elementary regulation as an arrow in a regulatory network.

An interesting overall observation dealing with the architecture of genetic regulatory networks is that the fraction of proteins that regulate other proteins N_{reg}/N increases with total number of proteins N. In other words, the relative size of the bureaucracy increases with system size. In fact for prokaryotes, the fraction of regulators increases linearly with the system size, reaching $\sim$10% for the prokaryote *Pseudomonas auriginosa* with its $\sim$6000 genes. If a living cell could be understood as an essential core plus a number of modules (genes regulated together), each, for example, associated to respond to a corresponding external situation, then the fraction of regulators would be independent of the number of genes N. The fact that N_{reg}/N grows linearly with N indicates that each added gene or module should be regulated with respect to all other gene modules. Thus, already prokaryotic organisms show features of a highly integrated computational machine. Networks are indeed important, from the smallest hubs with feedback to the whole integrated circuitry.

Topology of Genetic Regulatory Networks

In **Figure 3**, the known regulatory network for the single-celled eukaryote, the yeast *Saccromyces cerevisia* is shown. One feature of regulatory networks is the wide distribution of directed arrows from individual proteins. There are many proteins which control only few other proteins, but there also

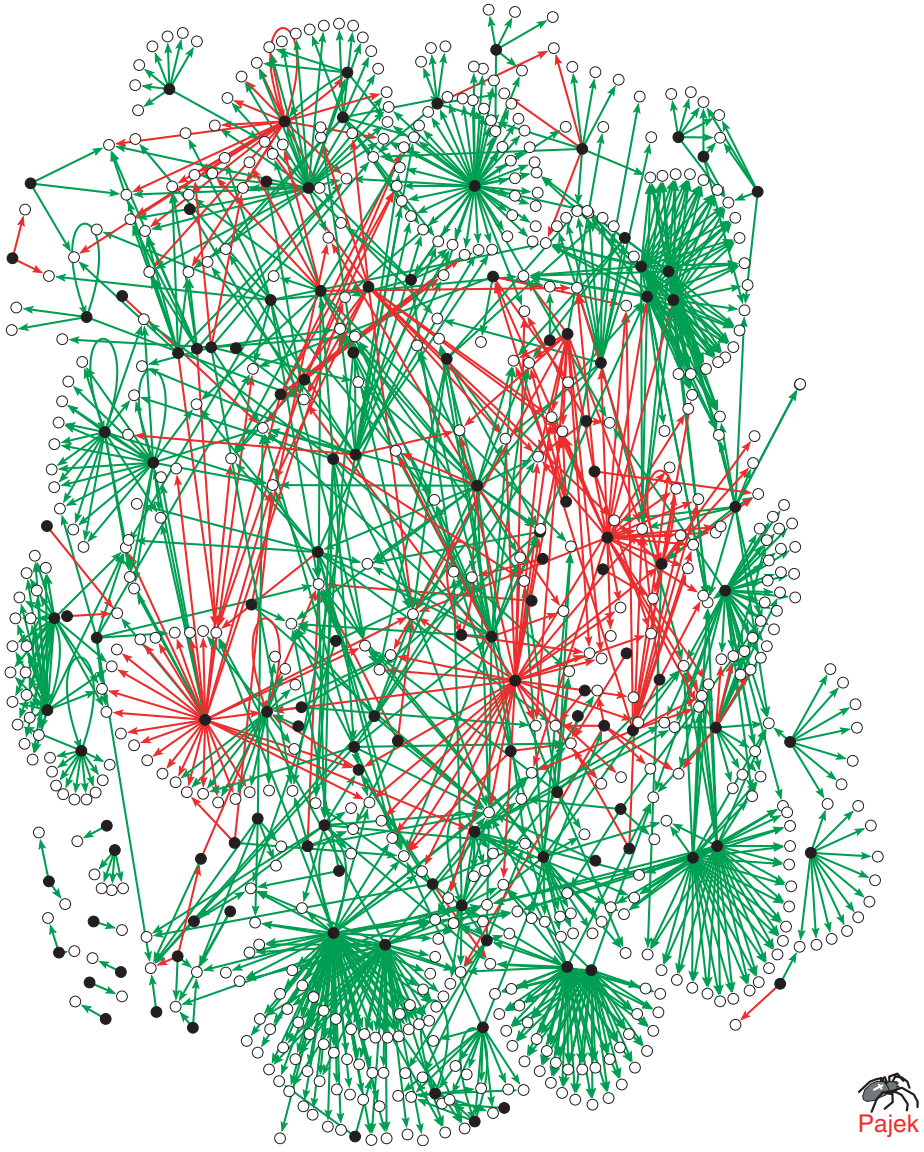


Figure 3 Presently known transcription regulations in yeast (*Saccromyces cerevisia*), with green arrows indicating positive regulation (activation) and red arrows indicating repression.

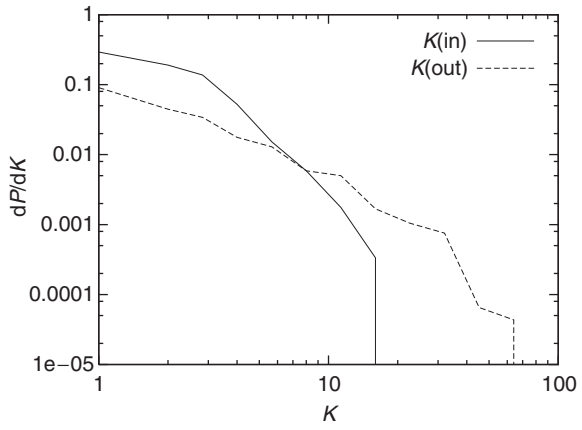


Figure 4 Statistics of regulation in the yeast *Saccromyces cerevisia*. (Data from Proteome.)

exist some proteins which control the expression level of many other proteins. In fact, the distribution of proteins with a given number of neighbors (connectivity) K may (very crudely) be approximated by a power law

$$N(K) \propto 1/K^{\gamma} \qquad [1]$$

with exponent $\gamma \sim 1.5 \pm 0.5$ for the "out-degree" distribution of transcription regulators (see **Figure 4**). It is also to be noticed that the distribution of the number of proteins in which a given protein regulates, the "out-degree" differs from the narrower distribution of "in-degrees." Overall however, both of these distributions are broad, and one would like to develop an intuition to find possible reasons why life organizes its regulation in this way.

One aspect of a wide distribution of connectivity in signaling networks is the possible amplification of signals in the network. Thus on an average, a signal entering a node could be transmitted to a number of other nodes given by the amplification factor:

$$\mathscr{A} = \frac{\langle K_{\text{in}} K_{\text{out}}(\text{given } K_{\text{in}}) \rangle}{\langle K_{\text{in}} \rangle} = \frac{\langle K_{\text{in}} K_{\text{out}} \rangle}{\langle K_{\text{in}} \rangle} \qquad [2]$$

the last equality implies that the output connectivity is independent of the input connectivity. The above formula follows from following a signal that enters a node with probability proportional to K_{in}, and leaves along any of the K_{out} links. The equation is thus the average amplification factor in a randomly wired directed network. When $\mathscr{A} > 1$, signals tend to be exponentially amplified and thus, most signals will influence signaling over the entire network. For networks with degree distribution which is a power law, $\mathscr{A}$ typically depends on the most highly connected node. Thus, highly connected nodes will be effective in transmitting a signal very fast to a substantial part of the whole system. This has obvious advantages, and may be one of the reasons why one also observes very highly connected nodes in both genetic and other signaling networks in living cells.

Power-law networks have also been proposed to be due to statistical processes, with little regard to functionality. One very widespread suggestion is the preferential attachment idea, that uses the fact that, if a network is growing by a mechanism where each new node is attached to nodes selected as being in the ends of already existing links in the network, then one obtains a steadily growing scale-free network. Another suggestion is the observation that even a nongrowing network can develop into a steady-state attractor with a scale-free degree distribution, if one assumes that a major element in the network rewiring dynamics is merging of old vertices and creation of new vertices. In the perspective of signaling networks, the latter model represents an on-going attempt to minimize the length of signaling pathways, competing with an on-going pressure to increase the protein diversity. In any case, traces of the actual evolution of networks can be quantified using the rather frequent gene duplication and associated protein homologs that one finds in any presently living organism. The overall lesson of such studies is that the genome of any organism is dynamic, and that reorganizations indeed are possible if they are functionally desirable.

The real reason for the observed broad distribution of edge degrees in biological networks should probably be seen in the perspective of where especially the highly-connected nodes sit in the network. To analyze this, a discussion on how to analyze network topologies beyond the degree distribution is taken up now. The key idea is to generate a random network, and then to compare this with the real network.

As was pointed out in the general context of complex scale-free networks, a broad distribution of degrees indicates that the degree itself is an important individual characteristic of a node and as such, it should be preserved in the randomized null-model network. In addition to degrees, one may choose to preserve some other low-level topological properties of the network in question. Any measurable topological quantity, such as, the total number of edges connecting pairs of nodes with given degrees, the number of loops of a certain type, the number and sizes of components, the diameter of the network, can then be measured in the real complex network and separately in its randomized version. One then

concentrates only on those topological properties of the real network that significantly deviate from their null-model counterpart.

An algorithm giving rise to a random network with the same set of individual node degrees, as in a given complex network, was proposed by Maslov S and Sneppen K in 2002. It consists of multiple repetitions of the following simple switch move (elementary rewiring step) illustrated in **Figure 5**: "Randomly select a pair of edges A→B and C→D and rewire them in such a way that A becomes connected to D, while C to B."

To prevent the appearance of multiple edges connecting the same pair of nodes, the rewiring step is aborted and a new pair of edges is selected if one or two of the new edges already exist in the network.

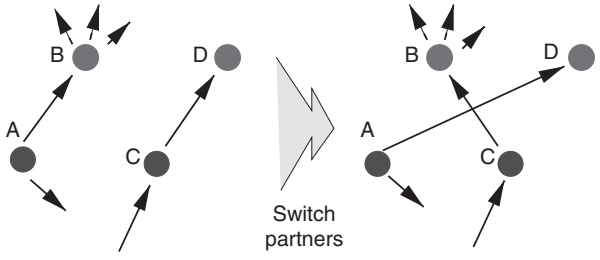


Figure 5 One step of the local rewiring algorithm. A pair of directed edges A→B and C→D. The two edges then switch connections in such a way that A becomes linked to D, while C to B, provided that none of these edges already exist in the network, in which case the move is aborted and a new pair of edges is selected. An independent random network is obtained when this is repeated a large number of times, exceeding the total number of edges in the system. This algorithm conserves both the in- and out-connectivity of each individual node. (Reproduced from Maslov S and Sneppen K (2002) Specificity and stability in topology of protein networks. *Science* 296: 910–913; American Association for the Advancement of Science.)

A repeated application of the above rewiring step leads to a randomized version of the original network. The set of MATLAB programs generating such a randomized version of any complex network can be downloaded from the internet.

Sometimes it is desirable that the null-model random network in addition to degrees of nodes, conserves some other topological quantity of the real network. In this case, one could supplement the random rewiring algorithm described above with the Metropolis acceptance/rejection criterion.

The "correlation profile" of any large complex network quantifies correlations between degrees of its neighboring nodes. The topological property of the network giving rise to its correlation profile is the number of edges $N(K_0, K_1)$ connecting pairs of nodes with degrees K_0 and K_1. To find out if in a given complex network the degrees of interacting nodes are correlated, $N(K_0, K_1)$ should be compared to its value $N_r(K_0, K_1) \pm \Delta N_r(K_0, K_1)$ in a randomized network, generated by the edge rewiring algorithm. When normalized by the total number of edges E, $N(K_0, K_1)$ defines the joint probability distribution $P(K_0, K_1) = N(K_0, K_1)/E$ of degrees of interacting nodes. Any correlations would manifest themselves as systematic deviations of the ratio

$$R(K_0, K_1) = P(K_0, K_1)/P_r(K_0, K_1) \qquad [3]$$

away from 1. The correlation profile for the yeast regulatory network is shown in **Figure 6**. The statistical significance of such deviations can, in principle, be quantified by their Z-score

$$Z(K_0, K_1) = (P(K_0, K_1) - P_r(K_0, K_1))/\sigma_r(K_0, K_1) \qquad [4]$$

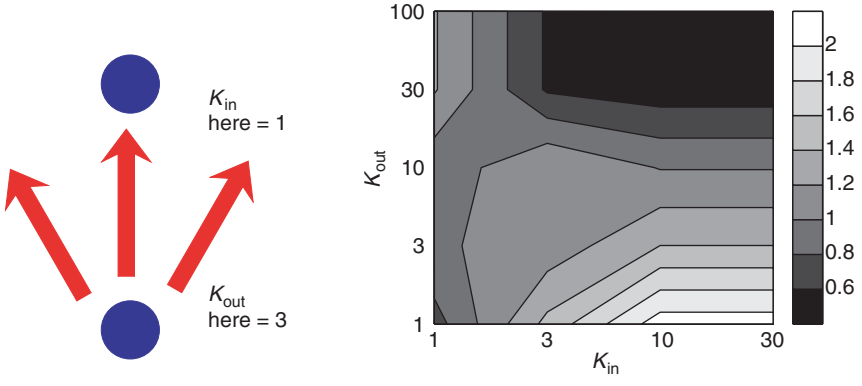


Figure 6 Correlation profile showing correlation between connected proteins in the regulatory network of yeast, quantified in terms of Z scores. The single output module is reflected in the abundance of high K_{out} controlling single K_{in} proteins. The dense overlapping regulons correspond to the abundance of connections between $K_{out} \sim 10$ and $K_{in} \sim 3$ proteins. Finally, the suppression of connections between highly connected proteins show that these tend to be on the periphery of the network. (Maslov S and Sneppen K (2002) Specificity and stability in topology of protein networks. *Science* 296: 910–913.)

where $\sigma_r(K_0, K_1) = \Delta N_r(K_0, K_1)/N$ is the standard deviation of $P_r(K_0, K_1)$ in an ensemble of randomized networks.

The result of the above analysis is that there is a significant suppression of links between hubs, both for regulatory networks analyzed in **Figure 6** and also for the protein networks measured in large-scale two-hybrid experiments on yeast. Similarly, by defining higher-order occurrences of various local patterns of control, Shen-Orr and co-workers suggested some frequent motifs of gene control in 2002, as illustrated in **Figure 7**.

Finally, it may be mentioned that the tendency of highly connected proteins to be at the periphery of regulatory networks in itself teaches one something about the origin of broad connectivity distributions. To see this, one could consider the fact that a protein is a fairly simple unit, that presumably gives essentially the same information to each of its downstream targets. Thus, the feature that a hub protein sits on the periphery of the network and does not link to other hub proteins, presumably reflects a one-hub-one-function structure. As a corollary of this, one may suspect that the broad distribution of molecular networks is not an intrinsic property of the network, but rather reflects the broad distribution of external requirements that the surroundings of an organism puts on it. Some functions simply require many proteins, whereas many functions require only a few proteins. This scenario is also supported by the observation that there is essentially no correlation between connectivity of a regulatory protein and its importance measured by its chance to be essential in knock-out experiments. If highly connected proteins were involved in many functions, their likelihood of being essential would grow linearly with their connectivity. This is not the case, for the yeast transcription network: the chance that a transcription factor is lethal is independent of how many proteins it regulates!

To summarize the topological properties of regulatory and signaling networks, the broad connectivity distributions may very well reflect the widely different needs associated to widely different functions that a living cell needs to cope with as it changes its environment.

Combinatorics in Genetic Networks, and an Evolution Model

In addition to the topological properties, the regulation of genes also have a combinatorical part. This means that the regulation of a given gene may depend on the combination of the concentration of its regulators. An inspiring way to model such a combinatorical regulation was proposed already in 1969 by S Kauffman. In this very simplified approach, each gene was assigned a binary number, 0 or 1, that counts whether the gene was off or on. Each gene i was assumed to be on or off depending on some logical function of the state of the gene that regulated it. A very simple version of this Boolean rule would be the threshold network where the state of the gene i at time t is given by

$$\sigma(i,t) = \Theta\left(\sum A_{ij}\sigma(j,t-1)\right) \qquad [5]$$

where $\Theta(x)$ is zero for $x<0$ and $=1$ otherwise. The A_{ij} is nonzero for genes j that regulate gene i. A_{ij} is positive for activation, and negative for repression. For a large regulatory network, the above update defines a cellular automata in a geometry defined by the connectivity matrix. Starting in an arbitrary initial condition, the system will move through a sequence of states that at some point is bound to repeat itself. This is because the state space is finite and the update deterministic. Thus, much effort has been used to characterize the behavior of these attractors as a function of the connectivity pattern of the system.

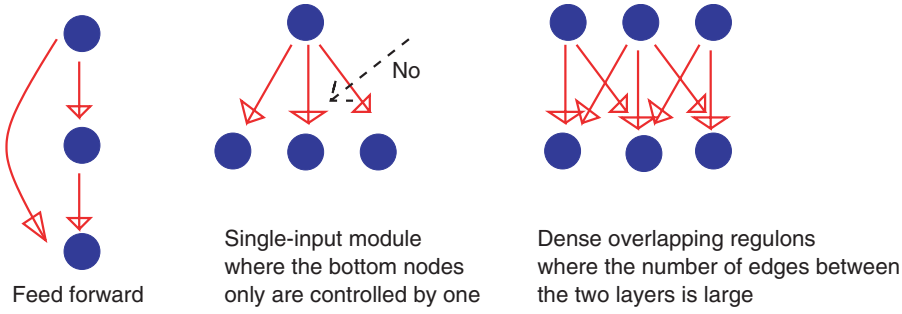


Figure 7 Genetic control motifs which are found to be over represented in regulatory networks of both *E. coli* and yeast. (Milo R, Shenn-Orr S, Itzkovitz S, Kashtan N, and Chklovskii D (2002) Network motifs: simple building blocks of complex networks. *Science* 298: 824–827.)

One simple result is that the system undergoes a percolation phase transition when the average connectivity is increased beyond a critical out-connectivity $\langle K_{\text{out}} \rangle \sim 2$. Above this transition, the attractors become very long (exponentially growing with system size) and the behavior is called chaotic.

The Boolean paradigm allows for modeling large-scale expression patterns of idealized regulatory networks. Thus, it also allows one to address how long-time evolution of rewiring and changes of combinatorical rules may constrain the evolution of such networks. In this regard, it is especially interesting to consider robustness as an evolutionary principle. One way to implement this is to "Evolve a new network from an old network by accepting rewiring mutations with a probability determined by the expression overlap."

This minimal-constraint scenario has no outside fitness imposed. Also, it ignores competition between networks, as it only compares a network with its predecessor. However, the model naturally selects networks which have a high overlap with neighbor mutant networks. This feature is associated to robustness, defined as the requirement that not only should the present network work, but also mutations of the networks should work. In terms of network topology, this means a change in the wiring $\{A_{ij}\} \rightarrow \{A'_{ij}\}$ that takes place on a much slower timescale than the $\{\sigma_i\}$ updating using the Boolean dynamics itself.

The system that is evolved is the set of couplings A_{ij} in a single network. One evolutionary time-step of the network is:

1. Create a daughter network by (a) adding, (b) removing, or (c) adding and removing a weight in the coupling matrix A_{ij} at random, each option occurring with probability $p = 1/3$. This means turning an $A_{ij} = 0$ to a randomly chosen ± 1 or vice versa.
2. Select a random input state $\{\sigma_i\}$. Iterate simultaneously both the mother and the daughter system from this state until they either have reached and completed the same attractor cycle, or until a time where $\{\sigma_i\}$ differs between the two networks. In case their dynamics is identical, then replace the mother with the daughter network. In case their dynamics differs, keep the mother network.

The result of this type of modeling is an intermittent evolutionary pattern: occasionally the evolved network is trapped in states with few evolution paths away from it; sometimes it is, instead, in highly active regions of "network space," thus allowing fast readjustments of genome architecture. More important, the obtained network exhibits a less chaotic behavior than random ones. The computational structure thus emerging has a simplicity of form consistent with that obtained with real molecular networks by placing highly connected nodes on the periphery of the network. In the Boolean model networks explored here, the attractors for the networks are shorter, and there are more frozen regulators in a random network at the same average connectivity.

See also: Environmental Science, Physical Principles and Applications.

PACS: 87.10. + e; 87.16. – b; 87.23. – n; 89.75. – k

Further Reading

Barabasi A-L and Albert R (1999) Emergence of scaling in random networks. *Science* 286: 509.

Bornholdt S and Sneppen K (2000) Robustness as an evolutionary principle. *Proceedings of Royal Society of London B* 267: 2281–2286.

Davidson *et al.*, A genomic regulatory network for development. *Science* 295(5560): 1669–2002.

Derrida B and Pommeau Y (1986) Random networks of automata: a simple annealed approximation. *Europhysics Letters* 1: 45–49.

deS Price DJ (1976) A general theory of bibliometric and other cumulative advantage processes. *Journal of American Society Information Science* 27: 292.

Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* 22: 437–467.

Kim BJ, Trusina A, Minnhagen P, and Sneppen K Self organized scale-free networks from merging and regeneration. nlin.AO/0403006.

Maslov S, Sneppen K, and Eriksen KA (2004) Upstream plasticity and downstream robustness in evolution of molecular networks. *BMC Evolutionary Biology* 4: 9.

Maslov S, Sneppen K, and Zaliznyak A (2004) Pattern detection in complex networks: correlation profile of the internet. *Physica A* 333: 529–540.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, and Teller E (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21: 1087.

Newman MEJ, Strogatz SH, and Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Physical Review E* 64: 026118, 1.

Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, *et al.* (2000) Complete genome sequence of *Pseudomonas Aeruginosa PA01*, an opportunistic pathogen. *Nature* 406: 959.

The set of MATLAB programs can be downloaded at http://www.cmth.bnl.gov/maslov/matlab.htm

Geometrical Optics

G Molesini, Istituto Nazionale di Ottica Applicata, Florence, Italy

Introduction

Geometrical optics is a branch of optics where light is described by rays. Light rays are conceived as geometrical lines originating from sources, extending through media, and being revealed by detectors; their directions account for paths along which light flows.

Although it was originated and developed as a discipline on its own, geometrical optics can be traced to Maxwell's equations within the general theory of electromagnetism. The derivation evidences the approximations that are made, and defines the realm of validity for geometrical optics.

Light paths are intrinsically related to the refractive index n of the media, where light propagates (n is the ratio between the speed of light in vacuum and in the medium). Ray equations can be written in general form, assuming n is a smooth continuous function of the position. In particular, if n is constant, rays are straight lines. Discontinuities in the refractive index due to interfaces between different media result in abrupt deviation of rays, named "refraction" and accounted for by simple equations.

Based on such equations, and a proper selection of the nature, shape, and location of the interfaces, rays can be guided to flow in an orderly manner through selected media ("lenses"), and eventually reproduce the source features ("image"). The basic properties of optical imaging systems are described by the first-order approximation of the ray equations. The resulting scheme is classically the core of geometrical optics.

Derivation from Maxwell's Equations

The major steps from electromagnetism to geometrical optics are summarized in **Figure 1**.

In a dielectric medium free of charges and currents, Maxwell's equations result in the wave equation

$$\nabla^2 \boldsymbol{E} = \varepsilon\mu \frac{\partial^2}{\partial t^2} \boldsymbol{E} \qquad [1]$$

where $\boldsymbol{E}$ is the electric field, t is time, and ε, μ are the electric permittivity and the permeability of the particular medium, respectively; in vacuum, $\varepsilon_0 = 8.8542 \times 10^{-12}\,\mathrm{F\,m^{-1}}$ and $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.

The velocity of light in vacuum is $c = 1/\sqrt{\varepsilon_0\mu_0}$, and $v = 1/\sqrt{\varepsilon\mu}$ in general. The equation for the magnetic induction $\boldsymbol{B}$ is identical to that for $\boldsymbol{E}$.

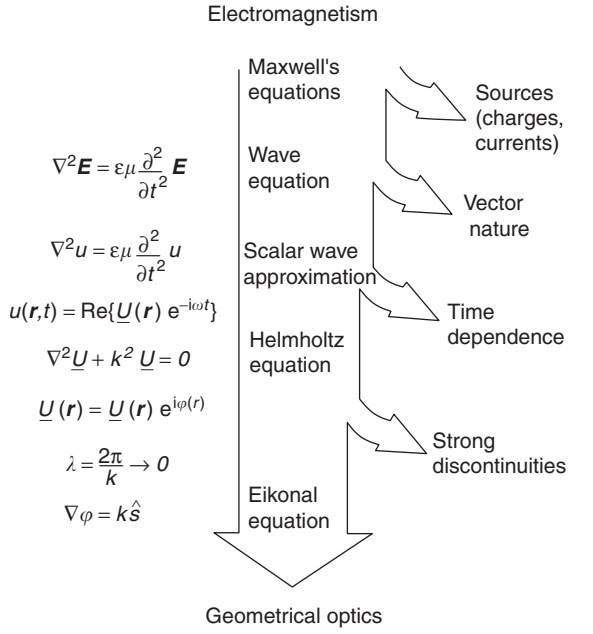


Figure 1 Schematic derivation of geometrical optics from Maxwell's equation.

Several classes of optic phenomena are satisfactorily described by separate components of the electromagnetic field. The resulting approach is the scalar approximation, where the vector nature of $\boldsymbol{E}$, $\boldsymbol{B}$ is dropped; eqn [1] is then written for a single (scalar) component u, it being a function of position $\boldsymbol{r}$ and time t. Monochromatic fields, that is, single harmonics at angular frequency ω, are sought as solutions. The complex notation $\underline{U}$ is also adopted, intending that the physical quantity u is the real part of $\underline{U}$:

$$u(\boldsymbol{r}, t) = \mathrm{Re}\{\underline{U}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\omega t}\} \qquad [2]$$

Substitution in the wave equation leads to the Helmholtz equation

$$\nabla^2 \underline{U} + k^2 \underline{U} = 0 \qquad [3]$$

which is independent of time; the quantity k is termed the wave number and is given by

$$k = \frac{\omega}{c} \qquad [4]$$

The wavelength λ is also introduced:

$$\lambda = \frac{2\pi}{k} \qquad [5]$$

The generic solution to the Helmholtz equation is written as

$$\underline{U}(\boldsymbol{r}) = U(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\varphi(\boldsymbol{r})} \qquad [6]$$

with $U(\boldsymbol{r})$ and $\varphi(\boldsymbol{r})$ the amplitude and phase of the field, respectively. Substitution into eqn [3] provides

$$\nabla^2 U - U\boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\varphi + k^2 U = 0 \qquad [7]$$

Now, the basic assumption is made that U does not vary significantly over distances of the order of λ, and particularly

$$\frac{\nabla^2 U}{U} \ll k^2 \qquad [8]$$

differently stated, the limiting case of $\lambda \to 0$ is studied. Under the above assumption, eqn [7] leads to the so-called eikonal equation:

$$\boldsymbol{\nabla}\varphi = k\,\hat{s} \qquad [9]$$

where $\hat{s}$ is the unit vector in the direction of $\boldsymbol{\nabla}\varphi$. Surfaces $\varphi = \text{const.}$ are geometrical wave fronts; $\hat{s}$ is, at every point, orthogonal to such surfaces. Since in isotropic media, the Poynting vector has that same direction, energy flows along lines singled out by $\hat{s}$ above. Such lines are the rays of geometrical optics.

The Ray Equation

Equation [9] can be written as

$$\boldsymbol{\nabla}\varphi = nk_0 \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} \qquad [10]$$

where

$$n = \frac{k}{k_0} = \frac{\lambda_0}{\lambda} = \frac{c}{v}$$

is the refractive index. Deriving eqn [10] with respect to s, and noting $\mathrm{d}\varphi/\mathrm{d}s = nk_0$ yields

$$\boldsymbol{\nabla}n = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s}\right) \qquad [11]$$

which is the differential form for the ray equation. The above equation is also the analytical expression for the Fermat's principle, according to which the light, going from point P_1 to point P_2, traverses a trajectory such that the quantity optical path length (OPL) defined by

$$\mathrm{OPL} = \int_{P_1}^{P_2} n(s)\,\mathrm{d}s \qquad [12]$$

is stationary with respect to neighboring trajectories. As a physical property, the optical path is reversible.

The curvature vector $\mathrm{d}\hat{s}/\mathrm{d}s$ of a ray is defined as

$$\frac{\mathrm{d}\hat{s}}{\mathrm{d}s} = \frac{1}{\rho}\,\hat{\sigma} \qquad [13]$$

where ρ is the radius of curvature, and $\hat{\sigma}$ is the unit principal normal, so that

$$\frac{1}{\rho} = \hat{\sigma} \cdot \boldsymbol{\nabla}(\log n) \qquad [14]$$

Equation [14] shows that a ray bends toward the region where n is higher. Overall, the ray equation provides the light path through continuous media assuming that $n(\boldsymbol{r})$ is known (e.g., bending of the rays in the Earth's atmosphere). As a particular case, in a homogeneous medium it is $n = \text{const.}$, and the ray path is a straight line ($\rho = \infty$).

Laws of Reflection and Refraction

Resting on homogeneous media, interfaces are defined as surfaces of separation between contiguous media with different n. A plane interface separating media with refractive indices n, n', and a single ray impinging on the surface at an angle i to the surface normal is considered first. The ray and the surface normal at the point of incidence define the plane of incidence. In most conditions, the incident ray is split into two secondary rays, one reflected and the other refracted (**Figure 2**), such that:

- both the reflected and the refracted rays belong to the plane of incidence,
- the reflected ray remains in the medium where the incident ray is from, at an angle i' symmetrical with respect to the surface normal, that is,

$$i' = -i \qquad [15]$$

- the refracted ray enters the facing medium, forming an angle i' to the surface normal so that

$$n \sin i = n' \sin i' \quad \text{(Snell's law)} \qquad [16]$$

(by widespread convention, primed symbols refer to quantities after refraction or reflection).

If i is fixed and light is analyzed in terms of single harmonics (i.e., pure colors, or wavelengths), the angle of reflection, $-i$, is the same for all colors, while the angle of refraction i' from eqn [16] depends on the color (**Figure 3**). This phenomenon is called dispersion and is described in terms of physical optics as a dependence of the refractive index of each medium on the wavelength.

Although formulas of geometrical optics are quite general, it is customary to refer to the various wavelength intervals as follows: 400–700 nm, where the eye is sensitive – visible spectrum; and the close

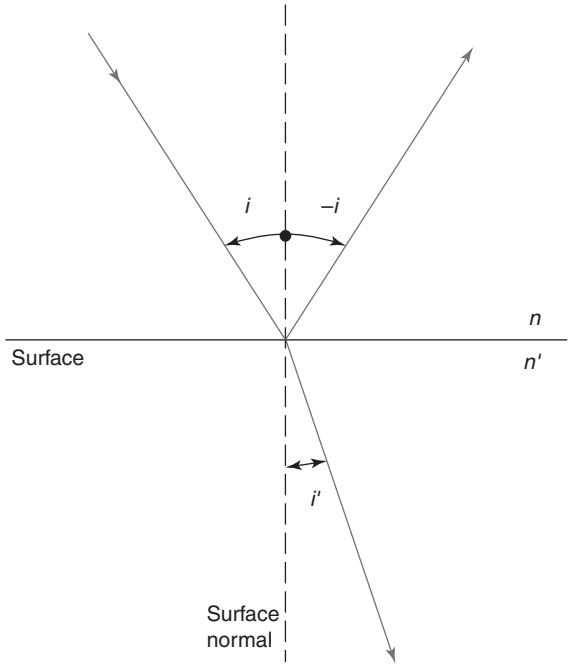


Figure 2 Ray reflection and refraction at a plane interface.

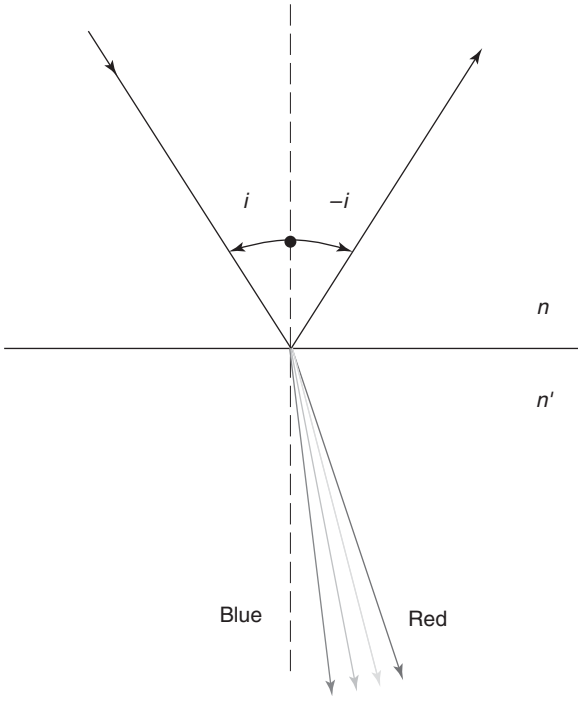


Figure 3 Fan spreading of refracted rays due to dispersion.

neighboring ranges, shorter wavelengths (<400 nm) – UV (ultraviolet), and longer wavelengths (>700 nm) – IR (infrared). For most optical materials, such as glasses, the dispersion equation for $n(\lambda)$ commonly referred to is

$$n^2(\lambda) - 1 = \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3} \quad [17]$$

(Sellmeier equation). A particular glass is specified by a set of B_1, B_2, B_3, C_1, C_2, C_3 coefficients. For easy reference, the refractive index at spectral lines listed in **Table 1** is usually quoted.

An example of the curve $n(\lambda)$ for a typical glass is given in **Figure 4**. By convention, optical materials are characterized by two parameters, indicating the central height of the curve and its slope: the refractive index n_d at the line d, and the so-called V-number V_d defined as

$$V_d = \frac{n_d - 1}{n_F - n_C} \quad [18]$$

According to eqn [18], more dispersive materials exhibit smaller V-number. For optical glasses, n_d and V_d range from 1.45 to 1.90, and from 85 to 20, respectively. Due to historical reasons, glasses with $n_d > 1.60$, $V_d > 50$, and $n_d < 1.60$, $V_d > 55$, are named crowns; the others are flints.

Ray Tracing and Image Formation

Sources are luminous objects emitting rays in directions around. In a simple case, a single point on an extended object can be modeled as a point-like source. An image is obtained by making the emitted rays to reconverge, or to focus, to a further ray concentration, where detection can take place or a new propagation process can begin. The image may be real, in the sense that it can be cast on a true screen, or a virtual one, if it only appears to diverge from a nonexistent ray concentration by virtue of reflection or refraction effects. The focusing action can be provided by a single surface, either refracting or reflecting, or by a sequence thereof. For reasons of fabrication, such surfaces are generally spherical. In particular, in refraction the ray-bending mechanism is produced by the application of Snell's law to each incident ray (**Figure 5**), with reference to the interface given by the osculating plane at the point of incidence. An extended object is considered to be made of many point sources, all processed in parallel, so that the final image results in many elementary images of the object points.

Modeling the refraction mechanism, as depicted in **Figure 6**, requires a consistent set of conventions to be established. Although it is not unique, the set here adopted is as follows:

- light travels from left to right, surfaces are numbered in a sequence, according to the order they are encountered by light;

Table 1 Reference spectral lines commonly used in geometrical optics

Line	*H*	*g*	*F'*	*F*	*e*	*d*	*C'*	*C*	*r*
Element	Hg	Hg	Cd	H	Hg	He	Cd	H	He
λ (nm)	404.7	435.8	480.0	486.1	546.1	587.6	643.8	656.3	706.5

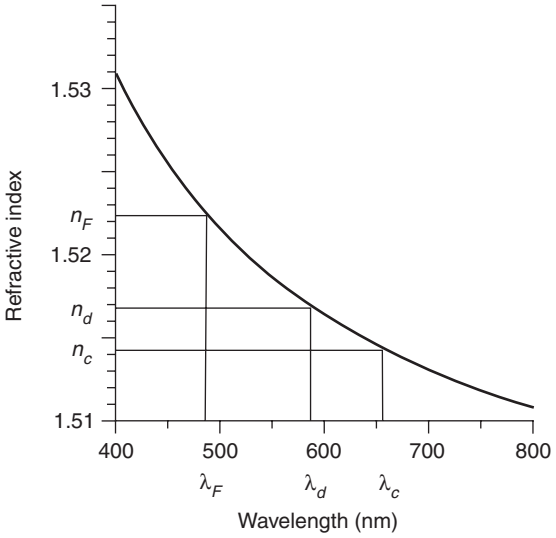


Figure 4 Curve of the refractive index as a function of the wavelength for a typical glass.

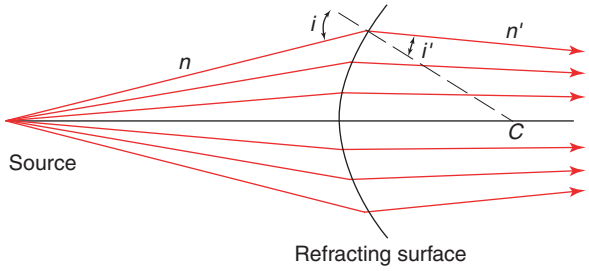


Figure 5 Bending of rays emitted by a point-like source at a spherical interface.

- the point source and the surface center define the optical axis, the origin is located at the surface vertex *V*, distances to the right and above the origin are positive, and those to the left and below are negative;
- the surface radius of curvature *r* is positive if the center of curvature *C* lies to the right of the vertex *V*, and negative otherwise;
- the angle *u* (and *u'* as well) formed between a ray and the optical axis is positive if an anticlockwise rotation smaller than 180° leads the oriented axis to the oriented ray, and negative otherwise;
- the angles of incidence and of refraction *i*, *i'* follow the same rule as *u*, *u'*, considering the surface normal in place of the axis.

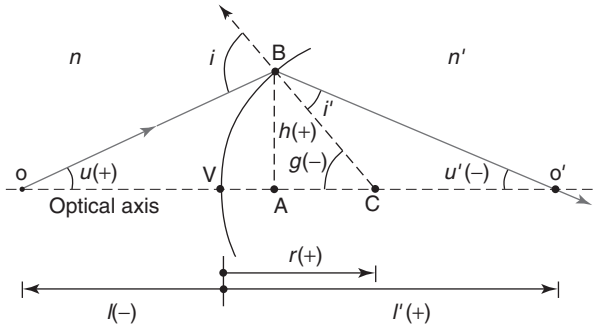


Figure 6 Refraction at a single spherical interface.

Table 2 Ray tracing through a refracting surface

Quantity	*Formula*
Incidence angle	$i = \arcsin\left[(r-l)\dfrac{\sin u}{r}\right]$
Refraction angle	$i' = \arcsin\left(\dfrac{n}{n'}\sin i\right)$
Angle at the center	$g = u - i$
Image side angle	$u' = i' + g$
Image distance	$l' = r\left(1 - \dfrac{\sin i'}{\sin u'}\right)$
Height at the surface	$h = -r \sin g$

Where the incident ray intercepts the axis is the object point O; the distance $\overline{VO}$ in **Figure 6** is $l<0$. Similarly, after refraction in B, the ray intercepts anew the axis at the image point O'; the distance $\overline{VO'}$ is $l'>0$.

The values for *n*, *n'*, *r*, *l*, and *u* are assigned; the mathematical expression for the quantities *i*, *i'* *g*, *u'*, *l'*, *h* is given in **Table 2**.

The ray so traced lies entirely on a plane containing the axis, and is thereupon named axial (or meridional). The plane itself is called meridional (or tangential). Rays which do not intercept the axis are called skew rays. Tracing skew rays requires more elaborate computations, and is not treated here.

Optical systems are made of a sequence of optical surfaces, encountered by the ray while it progresses from left to right. Here it is assumed, for simplicity, that all the surface centers are aligned along the same axis. A generic surface of an ordering number *j* is specified by a set of three data: the radius of curvature r_j, the axial distance d_j to the next surface, and

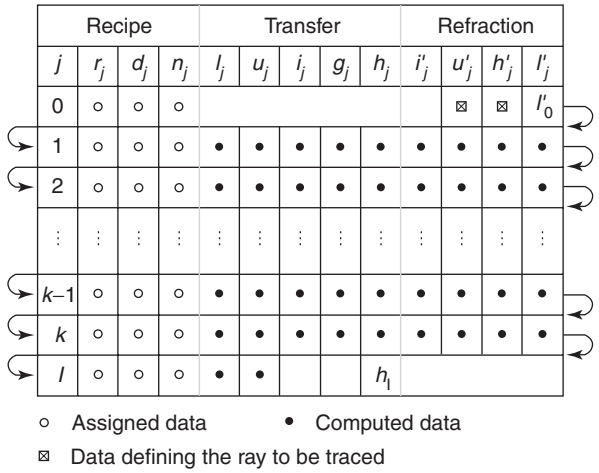


Figure 7 Scheme for recursive ray tracing through an optical system.

Table 3 Recursive ray tracing

Transfer equations	*Refraction equations*
$l_j = l'_{j-1} - d_{j-1}$	$i' = \arcsin\left(\frac{n_{j-1}}{n_j} \sin i_j\right)$
$u_j = u'_{j-1}$	$u'_j = i'_j + g_j$
$i_j = \arcsin\left[(r_j - l_j)\frac{\sin u_j}{r_j}\right]$	$h'_j = h_j$
$g_j = u_j - i_j$	$l'_j = r_j\left(1 - \frac{\sin i'_j}{\sin u'_j}\right)$
$h_j = -r_j \sin g_j$	

the refractive index n_j of the medium in between. The overall sequence of r_j, d_j, n_j sets makes up the optical prescription (or recipe) of the system. Starting from an object at an assigned distance from the first surface, a ray to be traced is identified by its slope angle u and height h above the axis. Tracing the ray implies transferring from each surface to the next, applying refraction there and so forth, proceeding from left to right until the last surface, say of ordering number k, and finally closing with the image surface. Computations can then be carried out by recursive methods according to the flow scheme of **Figure 7** and **Table 3**, to end up with the slope and height of the output ray on the image plane assigned by the recipe.

The use of electronic computers and specific software has speeded up the ray tracing task enormously. Ideally, neglecting diffraction, the image of a point-like source is a point-like image; tracing a high number of evenly spaced rays gives an idea of the departure of the actual system from the ideal case. The ray intercept so produced on the image plane is the spot diagram. Programs are available that evaluate the extent of the spot diagram, and can vary the recipe data with optimization cycles so as to minimize the spot size.

Paraxial Optics

Ray tracing is made easier in the case of rays approaching the optical axis. In that case, the argument α of sine and tangent functions is $\alpha \ll 1$ rad; expanding in power series and retaining only the first significant term, it is written as

$$\sin\alpha \approx \tan\alpha \approx \alpha \qquad [19]$$

The domain of such approximation, to be adopted from now on, is paraxial (or first-order, or Gaussian) optics. Ray tracing is accounted by simple algebraic equations, where a key parameter is the optical power ϕ. For a single refracting surface, the optical power is

$$\phi = \frac{n' - n}{r} \qquad [20]$$

Tracing a ray from O to O′ is described by

$$\frac{n'}{l'} - \frac{n}{l} = \phi \qquad [21]$$

showing that l' is independent of the particular ray that has been traced. The distances l, l' are said to be conjugate.

Paraxial formulas are also written approximating those given in **Tables 2** and **3**. As a special approach, two-element column vectors are used to represent rays, and 2×2 matrices to implement transfer and refraction. The usual algebra of matrix operators is found to apply. The main properties of optical systems are then derived from the resulting matrix elements.

Lenses and Lens Systems

When it comes to fabricating interfaces, optical surfaces can only be made in pairs, enclosing some transmissive material. The result is a lens. Lenses can be given different shapes, as indicated in **Figure 8**. The behavior of a lens with respect to image formation is first studied in cases when its thickness can be neglected as compared to l, l' (thin lens approximation). Equation [21] still holds, provided that the optical power ϕ assumes the form

$$\phi = \phi_1 + \phi_2 \qquad [22]$$

with

$$\phi_1 = \frac{n_L - n}{r_1}, \qquad \phi_2 = \frac{n' - n_L}{r_2} \qquad [23]$$

n_L, n, n' are the refractive indices of the lens material and the media on the source and image side,

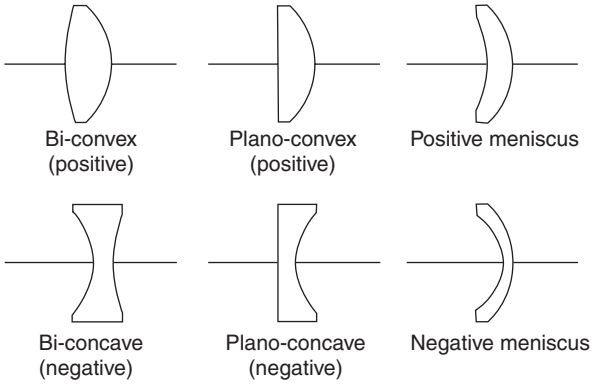


Figure 8 Basic layout of simple lenses.

Table 4 Truth table for eqn [21] to hold

$l<0$	Real source
$l>0$	Virtual source
$l'>0$	Real image
$l'<0$	Virtual image

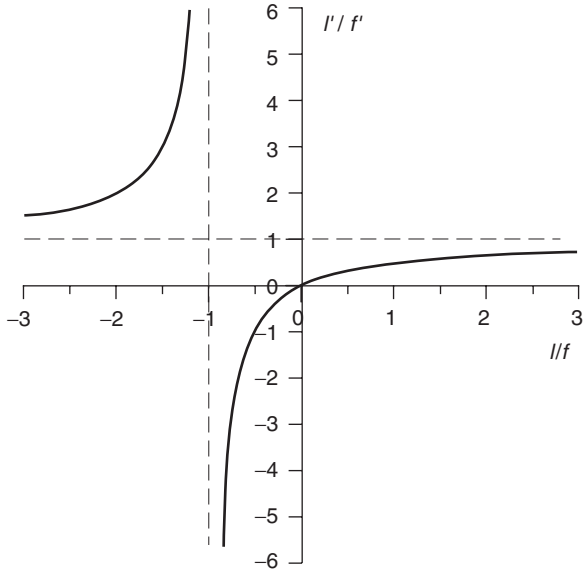


Figure 9 General plot of the conjugation law.

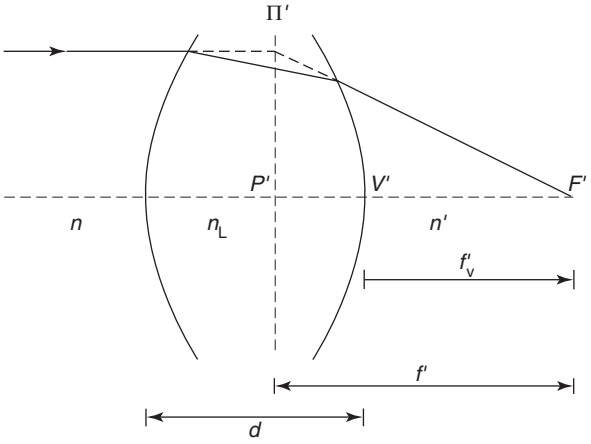


Figure 10 Definition of principal plane Π' and principal point P'.

respectively, and r_1, r_2 are the radii of curvature of the two surfaces; the domain of validity of eqn [21] is also bound to the truth table detailed in **Table 4**. Lenses with $\phi>0$ are said to be positive, or convergent, and are graphically represented with a segment with a pair of outward pointing arrows; otherwise, lenses are negative, or divergent, and are drawn with inward pointing arrows. It is customary to introduce the focal lengths f, f' according to

$$f = -\frac{n}{\phi}, \qquad f' = \frac{n'}{\phi} \qquad [24]$$

As it appears, f' locates the point F' where rays intersect the axis when the object goes to infinity ($l \to -\infty$), while f locates the point F where the object should be placed to produce an image at infinity ($l' \to \infty$). F, F' are the focal points. To a graphical representation of the general eqn [21], l' is normalized to f', and l to f. The resulting plot is presented in **Figure 9**. The plot is valid both for positive and negative lenses, paying due attention to the sign of the quantities involved.

Next is taking into account the actual thickness of the lens (thick lens approximation). Considering the scheme of **Figure 10**, a ray from infinity on the left is drawn. The ray undergoes two refractions, and crosses the axis at the focal point F'. Extending the in and out rays within the lens (dashed lines) shows that the overall deflection can be thought of as produced by a thin lens located at the intersection of the ray extensions. The location of such a thin lens is the principal plane Π'; its axial intercept is the principal point P'. In a similar manner, but with a ray going to infinity on the right, the principal plane Π and the principal point P are defined. The resulting scheme for the thick lens is presented in **Figure 11**. Again, the conjugation law is provided by eqn [21], but the optical power ϕ is now

$$\phi = \phi_1 + \phi_2 - \frac{d}{n_L}\phi_1\phi_2 \qquad [25]$$

d being the axial thickness of the lens; in addition, distances l, l' have to be measured from P, P', respectively. Also the focal lengths, still defined by eqn [24], are taken from P, P', that is, $f = \overline{PF}$, $f' = \overline{P'F'}$. Further quantities of interest are the front focal length

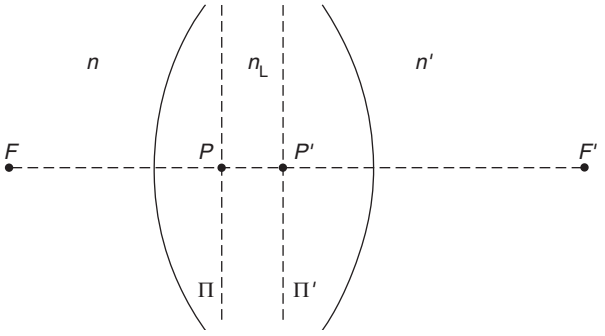


Figure 11 Paraxial scheme for a thick lens.

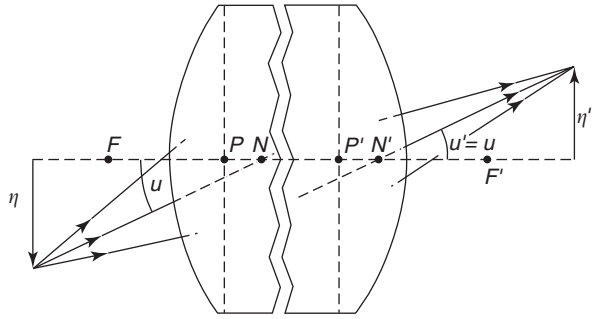


Figure 12 Nodal point construction for a thick lens.

Table 5 Thick lens parameters

	Unprimed quantity	*Primed quantity*
Front and back focal lengths	$f_v = \frac{n((d/n_L)\phi_2 - 1)}{\phi}$	$f'_v = \frac{n'(1-(d/n_L)\phi_1)}{\phi}$
Front and back vertex power	$\Omega = \frac{\phi}{1-(d/n_L)\phi_2}$	$\Omega' = \frac{\phi}{1-(d/n_L)\phi_1}$
Location of principal points	$\overline{VP} = \frac{nd\phi_2}{n_L\phi}$	$\overline{V'P'} = -\frac{n'd\phi_1}{n_L\phi}$
Location of nodal points	$\overline{PN} = \frac{n'-n}{\phi}$	$\overline{P'N'} = \frac{n'-n}{\phi}$

$f_v = \overline{VF}$, the back focal length $f'_v = \overline{V'F'}$, the front vertex power $\Omega = -n/f_v$, and the back vertex power (or ophthalmic power) $\Omega' = n'/f'_v$. Expressions for the above quantities are given in **Table 5**, along with formulas locating P, P' with respect to V, V'.

The case of a system of two thin lenses with powers ϕ_1, ϕ_2 separated by a distance d can be analyzed as if it were made of a single thick lens, with surface powers ϕ_1, ϕ_2, thickness d, and refractive index n_L equal to that of the medium in between. Formulas are available also for the case of three thin lenses. In general, however, the relevant quantities are computed by ray tracing.

So far, only on-axis point sources have been considered. Tracing rays from an off-axis source, it is noticed that just one ray exits the optical system without any angular deviation. Extending the in and out portions of that ray to cross the axis, the nodal points N, N' are singled out (**Figure 12**). Their positions, also given in **Table 5**, are shifted by the same amount with respect to the principal points, and coincide with P, P' for $n' = n$. The six points F, F', P, P', N, N' are named cardinal points, and thoroughly define the optical system within the paraxial approximation.

In the case of an off-axis source, the image is off-axis as well. Indicating with η the height of the source above the axis, and with η' that of the image, the lateral (or transverse) magnification m_L is defined as

$$m_L = \frac{\eta'}{\eta} \qquad [26]$$

In paraxial conditions, it is found to be

$$m_L = \frac{nl'}{n'l} \qquad [27]$$

A negative value of m_L indicates that the image is reversed upside down with respect to the object; $|m_L| < 1$ means that the image size is smaller than that of the object.

Considering infinitesimal increments of l, l', the axial (or longitudinal) magnification m_A is defined as

$$m_A = \frac{\delta l'}{\delta l} \qquad [28]$$

Differentiating eqn [21] yields

$$m_A = \frac{n'}{n} m_L^2 \qquad [29]$$

Being positive, m_A shows that increments δl and $\delta l'$ are pointing in the same direction.

As stated, conjugate distances l, l' are measured from the principal points P, P', which are difficult to locate. Measuring instead axial distances from the foci F, F', which are more easily identified, and posing $z = \overline{FO}$, $z' = \overline{F'O'}$ for the conjugate distances, it is found that

$$zz' = -\frac{nn'}{\phi^2} \quad \text{(Newton's equation)} \qquad [30]$$

Alternative expressions for the lateral and axial magnification can be written as well.

Mirrors

Formally, reflective optical surfaces may be treated with the same equations as refracting surfaces by posing

$$n' = -n \qquad [31]$$

After reflection, light travels backwards, so that in optical recipes the axial distance d that follows should be taken as negative. If a second reflection occurs, the refractive index returns positive, light proceeds forward, and d is positive again.

The power ϕ of a mirror is

$$\phi = -\frac{2n}{r} \qquad [32]$$

with r the radius of curvature of the reflecting surface; the sign convention for r is the same as that of refractive surfaces. The focal points F, F' are coincident; the focal lengths f, f' become

$$f = f' = \frac{r}{2} \qquad [33]$$

It follows that concave mirrors have positive power and negative focal length; such mirrors behave as convergent lenses. The opposite occurs with convex mirrors. The principal points P, P' are also coincident; their location is at the surface vertex. The nodal points N, N' coincide as well, and are located at the center of curvature of the reflecting surface. The cardinal points are indicated in **Figure 13**, both for a concave and a convex mirror.

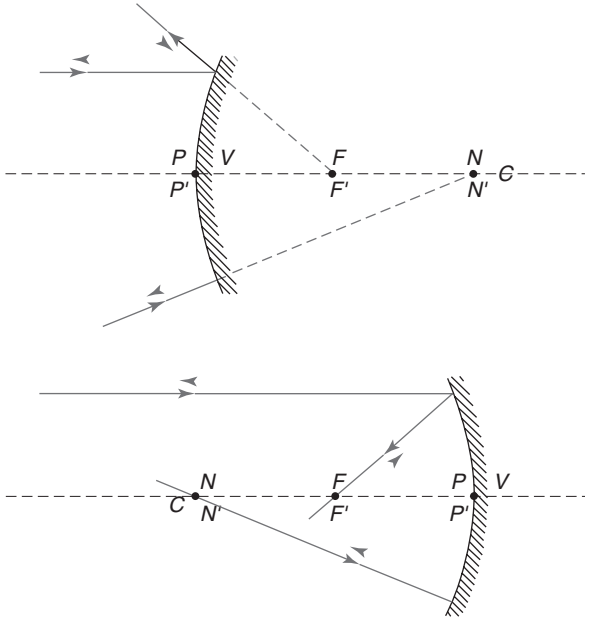


Figure 13 Cardinal points of concave and convex mirrors.

The conjugation law reads

$$\frac{1}{l'} + \frac{1}{l} = \frac{2}{r} \qquad [34]$$

The truth table for eqn [34] to hold is given in **Table 6**. The lateral magnification m_L becomes

$$m_L = -\frac{l'}{l} \qquad [35]$$

and the axial magnification m_A is

$$m_A = -m_L^2 \qquad [36]$$

Simple equations can be written also for the case of two mirrors, separated by a given distance. Ray tracing mechanisms do apply as well, provided that the appropriate conventions are observed.

Stops and Pupils

In a strict sense, the paraxial approximation is only valid at infinitesimal height from the axis. However, to convey an effective amount of light, the actual opening of optical components should be increased to finite values. Any element, be it the edge of a lens or a suitable diaphragm, that sets a maximum limit to the amount of light emitted by an axial source which is accepted by the optical system, is the aperture stop, or simply the stop. The ray that touches the edge of the stop is the axial marginal ray.

The stop physically divides the optical system into two parts. The part in front forms an image of the stop that is seen from the source side; such an image is the entrance pupil of the system. In a similar manner, the part following the stop forms an image of it that is seen from the image side and constitutes the exit pupil. Rays traced from the object toward the entrance pupil pass through the stop and are relayed to the exit pupil. Considering an off-axis point source, the ray that passes through the center of the stop (and the center of the pupils as well) is called the principal ray.

The extent of the object, either angular or linear, whose image is to be formed constitutes the field. The aperture that defines the actual size of the field accepted by the system is the field stop. Such an aperture is usually located at or near an intermediate image, or at the final image plane. As a general constraint, to

Table 6 Truth table for eqn [34] to hold

$l<0$	Real source
$l>0$	Virtual source
$l'>0$	Virtual image
$l'<0$	Real image

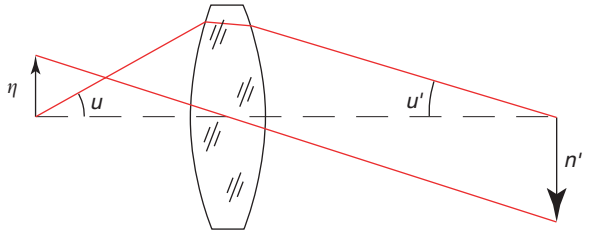


Figure 14 Paraxial quantities related to the definition of the optical invariant.

avoid having a system with infinite aperture, the axial position of the aperture stop is not allowed to coincide with the axial position of field planes, neither the field itself nor its images.

In order to obtain the relevant paraxial properties of an optical system, only two rays need to be traced. A very useful choice for such rays is the pair made of the axial-marginal ray and the principal ray. Referring to **Figure 14**, quantities related to the axial-marginal ray are the angle u and the height h. Similarly, the same quantities belonging to the principal ray are noted as $\bar{u}$, $\bar{h}$. The quantity H defined as

$$H = n(\bar{u}h - u\bar{h}) \qquad [37]$$

is named the optical (or Lagrange) invariant; the value of H does not change throughout the optical system. In particular, considering the object plane, where $h = 0$, $\bar{h} = \eta$, and the image plane, where $h = 0$, $\bar{h} = \eta'$, it is written as

$$H = -nu\eta = -n'u'\eta' \qquad [38]$$

In other words, u and η are inversely proportional. This is a basic relationship that has its counterpart in radiometry ("conservation of the radiance") and in optical information processing ("conservation of the space-bandwidth product").

The data about the maximum u-value $u_{\max}$ accepted by an optical system are conventionally expressed by the numerical aperture NA, defined as

$$\mathrm{NA} = n \sin u_{\max} \qquad [39]$$

Such a parameter expresses the ability of the system to resolve the object details and is typically given for optics that work with an object conjugate shorter than the image conjugate, as it happens with microscope lenses. In the opposite case, the f-number, $f_{\#}$, is preferably given and defined as

$$f_{\#} = \frac{\text{Focal length } f'}{\text{Entrance pupil diameter}} \qquad [40]$$

Typically, this is the parameter referred to in photographic lenses. The $f_{\#}$ also relates to luminosity and resolution of the optics. In principle, small f-numbers denote bright images and fine resolution.

Optical instruments are generally given a first description in terms of paraxial optics. These include visual instruments such as the eye itself, the magnifier, the eyepiece, the telescope, and the microscope. Other basic devices are the camera, the projector, relay systems, and more.

Aberrations

With a few exceptions, even if the lenses were perfect in terms of the spherical shape of the surfaces and centering on the optical axis, the image is affected by basic defects known as aberrations. Aberrations are intrinsic to the mechanism of image formation by refraction or reflection, and become significant as the aperture and the field depart more and more from infinitesimal values. After the name of the scientist who first investigated this subject, the primary image defects are usually referred to as the Seidel aberrations. In addition, due to the fact that the refractive index is a function of the wavelength, chromatic aberrations are produced. Here, only the noticeable effects of these aberrations are mentioned. Although an effective mathematical account can be given in terms of geometrical wave fronts, rays and spot diagrams are used.

There are five Seidel aberrations. Three of them – spherical aberration, coma, and astigmatism – cause basic deterioration of the image quality, making it blurred. The remaining two – Petzval field curvature and distortion – alter the image geometry.

- For a point-like source, the effect of spherical aberration is a symmetrical blurring of the image (**Figure 15**). Under ideal conditions, this is the only defect that can affect the image near the center of the field. Spherical aberration occurs because of a variation of the focal length along the diameter of the lens: in a single positive lens, the outer rays intersect closer to the lens than the inner rays (**Figure 16**).
- Coma can be defined as the variation of magnification with the aperture. When a bundle of rays is incident on a lens with coma, the rays passing through the peripheral part are imaged at heights different from those passing through the central part. A typical spot diagram is shown in **Figure 17**. The name of this aberration derives from the comet-shaped flare of the image.
- Astigmatism is produced because the incident cone of rays from an off-axis point-like source encounters different curvatures of the optical surfaces in their principal sections. In the presence of

astigmatism, the image of a point is not a point but takes the form of two axially separated lines. Between these lines, the spot evolves to an elliptical and circular blur (**Figure 18**).

- Every optical system has, associated with it, a basic field curvature, called Petzval curvature, which is a function of the index of refraction of the lens elements and their surface curvatures. As a consequence, the sharpest image is formed on a curved focal surface rather than a flat focal plane. If the detector geometry does not conform to this focal surface, image blurring will result locally.
- Distortion is a variation of the magnification with the field angle. In the case of a square object, distortion may produce a "pincushion" image or a "barrel" image.

Chromatic defects are longitudinal chromatic aberration and lateral color. Longitudinal chromatic aberration is caused by dispersion: in polychromatic light, the focal region of a lens is actually an axial

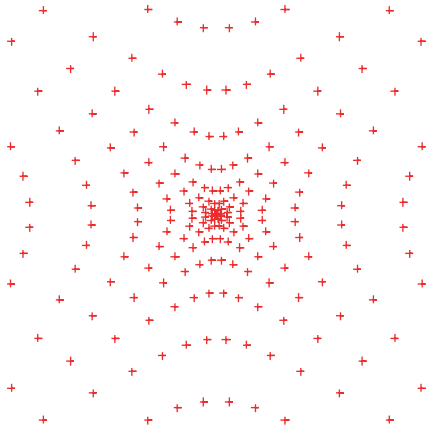


Figure 15 Typical spot diagram in the presence of spherical aberration.

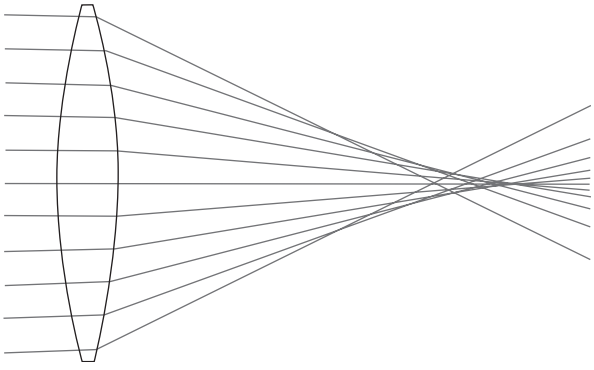


Figure 16 Ray tracing through a lens affected by spherical aberrations.

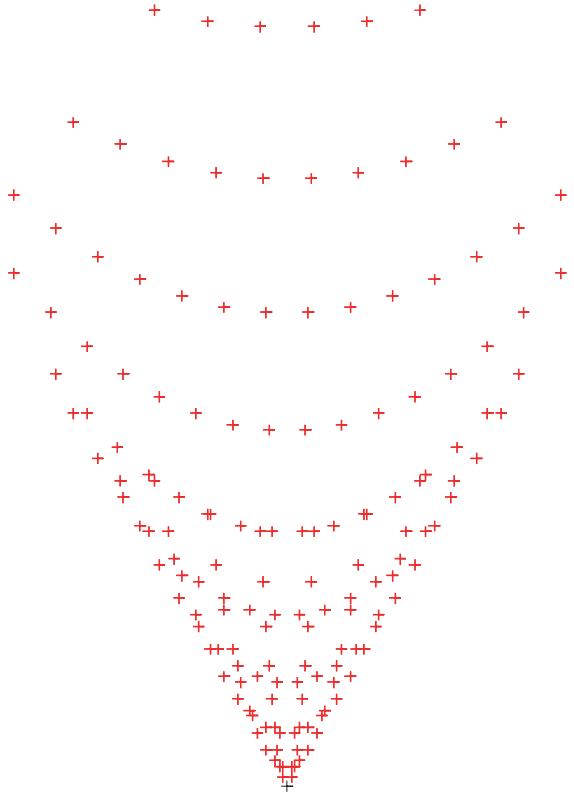


Figure 17 Typical spot diagram in the presence of coma.

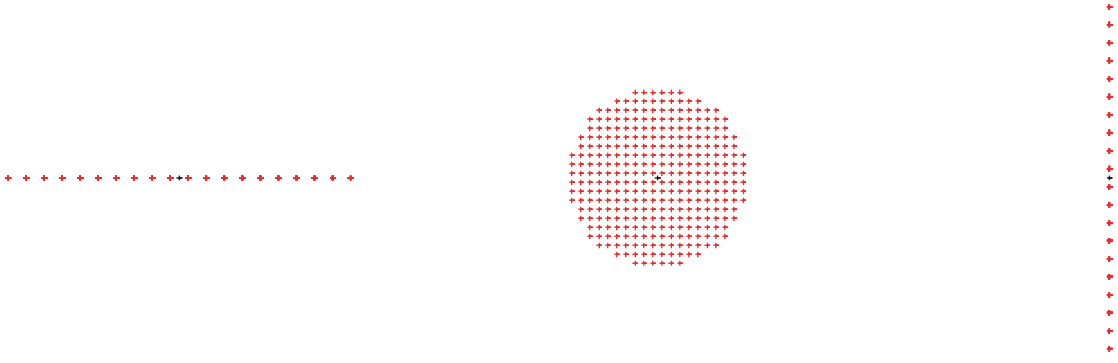


Figure 18 Typical through-focus spot diagrams in the presence of astigmatism.

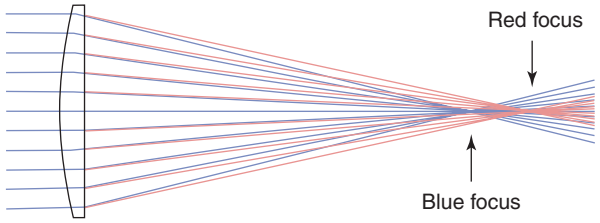


Figure 19 Longitudinal chromatic aberration of a simple positive lens.

sequence of foci of the different colors of the rainbow (**Figure 19**). As a result, the image of a point is a bright dot surrounded by a variously colored halo. Lateral color occurs in off-axis images because of the varying magnification with wavelength. The result is a lateral spread of rainbow colors and manifests itself as edge color fringes.

Although discussed separately, in general, the aberrations are simultaneously present in differing degrees. Combined, they tend to degrade the image quality; efforts are then made to develop optical recipes such that aberrations are minimized. This usually leads to optical systems of increased complexity.

See also: Electrodynamics: Continuous Media; Imaging and Optical Processing; Optical Instruments; Optical Microscopy.

PACS: 42.15. – i; 42.15.Dp; 42.15.Fr

Further Reading

Born M and Wolf E (1989) *Principles of Optics*, 6th (corrected) edn. Oxford: Pergamon.
Gerrard A and Burch JM (1975) *Introduction to Matrix Methods in Optics*. London: Wiley.
Goodman DS (1995) Geometric optics. In: Bass M (ed.) *Handbook of Optics*, 2nd edn., vol. I, pp. 1.3–1.109. New York: McGraw-Hill.
Hecht E (2002) *Optics*, 4th edn. San Francisco: Benjamin Cummings.
Jenkins FA and White HE (2001) *Fundamentals of Optics*, 4th edn. New York: McGraw-Hill.
Kidger MJ (2002) *Fundamental Optical Design*. Bellingham: SPIE Press.
Kingslake R (1978) *Lens Design Fundamentals*. New York: Academic Press.
Longhurst RS (1973) *Geometrical and Physical Optics*, 3rd edn. London: Longman.
Meyer-Arendt JR (1995) *Introduction to Classical and Modern Optics*, 4th edn. Englewood Cliffs: Prentice-Hall.
O'Shea DC (1985) *Elements of Modern Optical Design*. New York: Wiley.
Smith WJ (1990) *Modern Optical Engineering*, 2nd edn. New York: McGraw-Hill.
Smith G and Atchison DA (1997) *The Eye and Visual Optical Instruments*. Cambridge: Cambridge University Press.
Welford WT (1986) *Aberrations of Optical Systems*. Bristol: Adam Hilger.

Geometrical Phase and Polarization in Solids

R Resta, University of Trieste, Trieste, Italy

Introduction

The macroscopic polarization $\boldsymbol{P}$ is the most essential concept in any phenomenological description of dielectric media: it is an intensive vector quantity that intuitively carries the meaning of electric dipole moment per unit volume. It is thus tempting to try to define $\boldsymbol{P}$ as the dipole of the macroscopic sample divided by its volume, that is,

$$\boldsymbol{P} = \frac{1}{\Omega_{\text{sample}}} \int_{\Omega_{\text{sample}}} \mathrm{d}\boldsymbol{r}\, \boldsymbol{r}\rho(\boldsymbol{r}) \qquad [1]$$

where ρ is the charge density. Notice, however, that the integral has contributions from both the surface and the bulk regions, which are not easily disentangled: in fact, the value of [1] can be modified by just varying the surface charge of the sample, without varying its bulk charge density while instead, bulk polarization must be unambiguously defined by something "happening" in the bulk. For a crystalline solid, where the bulk charge density is lattice-periodical, a second tempting definition would be

$$\boldsymbol{P} = \frac{1}{\Omega_{\text{cell}}} \int_{\Omega_{\text{cell}}} \mathrm{d}\boldsymbol{r}\, \boldsymbol{r}\rho(\boldsymbol{r}) \qquad [2]$$

with Ω_{cell}, a crystal cell. However, this definition – appearing in many textbooks – is also flawed, because the result of [2] depends on the shape and location of the unit cell.

The clue to understanding polarization comes from a closer look at the experiments. Most measurements of bulk macroscopic polarization $\boldsymbol{P}$ do not address its absolute value, but only its derivatives, which are expressed as Cartesian tensors. By definition, permittivity is the derivative of $\boldsymbol{P}$ with respect to an electric field, pyroelectricity with respect to temperature, piezoelectricity with respect to macroscopic strain, and Born charges (also known as "dynamical" or

"infrared" charges) with respect to zone-center phonon amplitudes. Even when the "spontaneous" polarization of a ferroelectric material is measured, the genuine observable bulk property is the finite difference of polarization between two states of the same solid (polarization reversal, discussed below). The modern theory concerns polarization differences $\Delta \boldsymbol{P}$ as well, and avoids defining absolute polarization at all.

At this point, it is convenient to clearly distinguish between two cases, which prove to be very different from a theoretical standpoint. Case one concerns polarization in an electric field, and is, of course, needed for defining the dielectric constant (or tensor). Case two concerns, instead, polarization in a zero field, under which circumstances are most normally defined by the piezoelectric tensor, the Born effective charge tensor, and the spontaneous polarization in ferroelectric materials. In both cases, the polarization difference $\Delta \boldsymbol{P}$ induced by a given perturbation coincides with the transient current flowing through the bulk of the sample while the perturbation is switched on:

$$\Delta \boldsymbol{P} = \boldsymbol{P}(\Delta t) - \boldsymbol{P}(0) = \int_0^{\Delta t} \mathrm{d}t\, \boldsymbol{j}(t) \qquad [3]$$

where $\boldsymbol{j}$ is the macroscopic (i.e., cell-averaged) current density. The theory addresses precisely this same current, but the typical methods for dealing with the two cases are very different.

The problem of evaluating the dielectric constant of a solid has been solved in 1962. This solution relies on the linear-response theory; first-principle implementations date from the mid-1980s. A modern formulation of linear-response theory in crystalline solids, also known as density-functional perturbation theory is outlined. Next, an outline of the so-called "modern theory of polarization," developed in the 1990s is given, which allows one to evaluate a polarization difference $\Delta \boldsymbol{P}$ between two states of a given solid in a zero field. This theory is based on a Berry phase, also called geometric phase. Preliminarily, a brief treatment of the Berry phases in a general setting beyond the polarization problem is provided.

Throughout, a density-functional framework is adopted: the single-particle orbitals ψ and potential V, unless otherwise noted, are the Kohn–Sham ones.

Macroscopic Fields and Dielectric Constants

The discussion starts by observing that the problem of an extended solid in a macroscopic field is a very difficult one, given the fact that the electrons are in a nonperiodic potential, and therefore, the orbitals no longer have the Bloch form. Even worse, the Hamiltonian is no longer bounded from below. Despite such major problems for finite fields, it is possible to perform electronic-structure calculations which are correct to the first-order in the field magnitude, without trading away the major simplification of dealing with pure Bloch states.

Consider a finite macroscopic sample, having discrete single-particle orbitals ψ_i, which vanish at infinity: the thermodynamic limit will be taken only at the end. The electronic density is written as

$$n(\boldsymbol{r}) = \sum_i f_i |\psi_i(\boldsymbol{r})|^2 \qquad [4]$$

where f_i is an occupancy factor (either 2 or 0) selecting the doubly occupied orbitals. The crystalline observable of interest is the macroscopic dielectric tensor ε_∞, defined as

$$\varepsilon_\infty = 1 + 4\pi \frac{\partial \boldsymbol{P}_{\mathrm{el}}}{\partial \boldsymbol{E}} \qquad [5]$$

where $\boldsymbol{P}_{\mathrm{el}}$ is the macroscopic electronic polarization linearly induced by the (screened) field $\boldsymbol{E}$; in a cubic material, such a tensor obviously reduces to a constant. Using [4], one may evaluate

$$\frac{\partial \boldsymbol{P}_{\mathrm{el}}}{\partial \boldsymbol{E}} = \frac{q_{\mathrm{e}}}{\Omega_{\mathrm{sample}}} \sum_i f_i \left\langle \psi_i | \boldsymbol{r} | \frac{\partial \psi_i}{\partial \boldsymbol{E}} \right\rangle + \mathrm{cc} \qquad [6]$$

where q_{e} indicates the electron charge and cc stays for complex conjugate. The dipole matrix elements are then transformed using the identity:

$$\langle \psi_i | \boldsymbol{r} | \psi_j \rangle = -\mathrm{i}\hbar \frac{\langle \psi_i | \boldsymbol{v} | \psi_j \rangle}{\varepsilon_i - \varepsilon_j}, \quad i \neq j \qquad [7]$$

where ε_i are the single-particle eigenvalues, and the velocity operator $\boldsymbol{v}$ coincides with $\boldsymbol{p}/m_e$ in the simple case where all potentials are local.

It may be emphasized that the current appears in the formulation precisely at this point. If one is actually dealing with a finite system, both members of [7] can be equivalently used. But eventually, the thermodynamic limit is to be performed, in which case the left-hand member of [7] becomes ill-defined: for an extended system in the thermodynamic limit, it is then mandatory to use the velocity formula, hence the current.

Straightforward manipulations transform [6] into a Kubo formula:

$$\frac{\partial \boldsymbol{P}_{\mathrm{el}}}{\partial \boldsymbol{E}} = -\frac{\mathrm{i}q_{\mathrm{e}}}{\Omega_{\mathrm{sample}}} \sum_{i,j} \frac{f_j - f_i}{\varepsilon_j - \varepsilon_i} \langle \psi_i | \boldsymbol{v} | \psi_j \rangle \left\langle \psi_j \middle| \frac{\partial \psi_i}{\partial \boldsymbol{E}} \right\rangle \qquad [8]$$

where the fact that $|\partial \psi_i / \partial \boldsymbol{E}\rangle$ has been exploited may be chosen as orthogonal to $|\psi_i\rangle$.

In order to proceed further, a self-consistent scheme providing the derivatives of the orbitals and of the potential with respect to the perturbation is needed. The first step consists in writing the first-order corrections to the orbitals as

$$|\Delta\psi_i\rangle = \sum_{j\neq i} |\psi_j\rangle \frac{\langle\psi_j|\Delta V|\psi_i\rangle}{\varepsilon_i - \varepsilon_j} \qquad [9]$$

As a second step, one looks for an independent relationship providing, instead, ΔV in terms of the orbital variations $|\Delta\psi_i\rangle$: iterating over these two steps, the self-consistency goal is reached.

The self-consistent perturbation potential ΔV is written as

$$\Delta V = \Delta V_{\text{ext}} + \Delta V_{\text{H}} + \Delta V_{\text{xc}} \qquad [10]$$

where ΔV_{ext} is the bare perturbation (in the present case the potential of the unscreened field), and the remaining two terms are the Hartree and exchange-correlation contributions, respectively. The Hartree term is linear in the induced charge density:

$$\Delta V_{\text{H}}(\boldsymbol{r}) = q_{\text{e}}^2 \int \mathrm{d}\boldsymbol{r}' \Delta n(\boldsymbol{r})/|\boldsymbol{r} - \boldsymbol{r}'| \qquad [11]$$

or in shorthand

$$\Delta V_{\text{H}} = v_{\text{c}} \Delta n \qquad [12]$$

while the exchange–correlation contribution is linearized as

$$\Delta V_{\text{xc}} \simeq f_{\text{xc}} \Delta n \qquad [13]$$

The functional derivative defining this linear term is

$$f_{\text{xc}}(\boldsymbol{r}, \boldsymbol{r}') = \frac{\delta V_{\text{xc}}(\boldsymbol{r})}{\delta n(\boldsymbol{r}')} = \frac{\delta^2 E_{\text{xc}}[n]}{\delta n(\boldsymbol{r}) \delta n(\boldsymbol{r}')} \qquad [14]$$

where E_{xc} is the exchange–correlation energy functional.

When a macroscopic field $\boldsymbol{E}$ acts on the solid, it proves useful to write the self-consistent perturbation potential ΔV in a slightly different form from [10], separating the macroscopic field from the microscopic one (also called local field). The potential of the screened macroscopic field – due to both the bare and the Hartree terms in [10] – is written as $-q_{\text{e}}\boldsymbol{E}\cdot\boldsymbol{r}$, while the remaining microscopic term is lattice-periodical in the thermodynamic limit. Therefore [10] is replaced with

$$\Delta V = -q_{\text{e}}\boldsymbol{E}\cdot\boldsymbol{r} + v_{\text{c}}\Delta n + f_{\text{xc}}\Delta n \qquad [15]$$

where the term $v_{\text{c}}\Delta n$ amounts to solving Poisson's equation.

The self-consistent loop is performed cycling over [15] and [9], while the value of the screened field $\boldsymbol{E}$ is kept constant during the iteration. Whenever [15] is inserted in [9], the matrix elements of $\boldsymbol{r}$ are transformed to the velocity form again using [7]. Once self-consistency is reached, one gets the orbital variations $|\Delta\psi_i\rangle$ induced by a given macroscopic field $\boldsymbol{E}$, which are exact to linear order in the field magnitude: whence the field derivative to be used in [8] and in ε_∞, eqn [5].

All of the above equations have a simple and well-defined expression for the infinite periodic crystal in the thermodynamic limit, where the index i is identified with the band index and the Bloch vector altogether. Essentially, this is the formulation of the linear-response theory, which is nowadays implemented in first-principle calculations (see the "Further reading" section for a complete account).

Piezoelectricity and Ferroelectricity

Next to be addressed are the phenomena where the macroscopic polarization is induced by a source other than an electric field. Even in this case, the polarization may (or may not) be accompanied by a field, depending on the boundary conditions chosen for the macroscopic sample.

Two possible realizations of the piezoelectric effect are schematically shown in **Figure 1**. In (a) the crystal is uniaxially strained (along a piezoelectric axis) while kept in a shorted capacitor; in (b) the sample is strained while kept isolated. Focusing on (a), one may notice that the phenomenon manifests itself as a bulk current traversing the sample, while nothing peculiar happens at the surfaces. Indeed, the current flowing across the shorting wire is the quantity which is actually measured in experiments, and the modern theory addresses precisely this same current. Since it is the phase of the wave function which carries

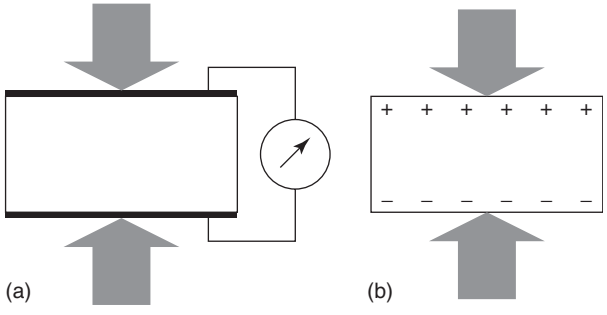


Figure 1 Two possible realizations of the piezoelectric effect: (a) The sample is in a shorted capacitor, and the current is measured. (b) The sample is isolated.

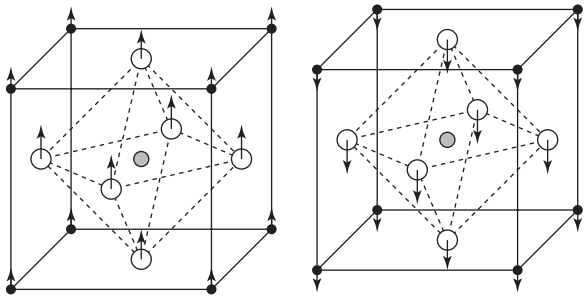


Figure 2 Tetragonal structure of $KNbO_3$. Solid, shaded, and empty circles represent K, Nb, and O atoms, respectively. The internal displacements (magnified by a factor 4) are indicated by arrows for two ($\bar{A}$ and $\bar{B}$) enantiomorphous ferroelectric structures.

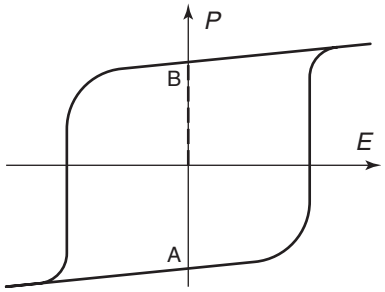


Figure 3 A typical hysteresis loop; the magnitude of the spontaneous polarization P_{spont} is also shown (vertical dashed segment). Notice that P_{spont} is defined as a zero-field property.

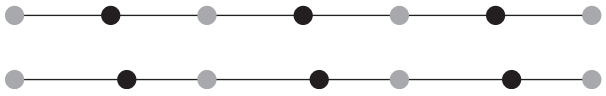


Figure 4 A one-dimensional chain in the centrosymmetric structure (upper sketch) and in the broken symmetry ferroelectric structure (lower sketch). Black and gray circles schematize anions and cations.

information about the current, it is not surprising that a phase (Berry phase) occurs in the theoretical description of electronic polarization.

Now a discussion on the phenomenon of spontaneous polarization in ferroelectrics is taken up. The most studied ferroelectric materials are in the class of perovskite oxides: they have a cubic prototype phase above the Curie temperature, and display a series of structural transitions to low-symmetry ferroelectric phases when the temperature is lowered. Typically, the first transition is to a tetragonal phase, which is characterized by microscopic displacements of the ions out of their high-symmetry sites. These determine a preferred polarity of the tetragonal axis, and are responsible for the occurrence of spontaneous polarization. This is illustrated in **Figure 2** for the tetragonal structure of $KNbO_3$, which was the very first case study for the modern theory of polarization in 1993.

A typical measurement of the spontaneous polarization, performed through polarization reversal, is schematically shown in **Figure 3**. The two enantiomorphous ferroelectric structures A and B are characterized by opposite values of the displacements, and the transition between them is driven by an applied electric field: the experimental setup typically measures the integrated macroscopic current flowing through the sample during the hysteresis cycle, shown in **Figure 3**. One half of the difference $P_B - P_A$ defines the magnitude P_{spont} of the spontaneous polarization; from **Figure 3**, it is also clear that P_{spont} can also be defined as the polarization difference ΔP between the ferroelectric B structure and the centrosymmetric structure (where the displacements are set to zero). Notice that ΔP can be evaluated along the vertical axis, where the macroscopic field is identically zero.

The phenomenon is schematized with a toy model in one dimension, **Figure 4**, meant to represent the central vertical line of Nb and O atoms in **Figure 2**. By definition, the spontaneous polarization of the ferroelectric structure is the integrated current, which flows through the linear chain while the ions are continuously displaced from the centrosymmetric (top sketch) to the ferroelectric (bottom sketch) structures, keeping the electric field vanishing.

Berry Phases

A digression about the Berry phase on most general grounds is provided; its relevance in the polarization phenomena is discussed in the next section.

One may start with a quantum Hamiltonian having a parametric dependence

$$H(\xi)|\psi(\xi)\rangle = E(\xi)|\psi(\xi)\rangle \qquad [16]$$

where ξ is defined in a suitable domain: a two-dimensional real ξ has been chosen for display in **Figure 5**. For the time being, there is no need to specify which quantum system is described by this Hamiltonian, or which is the physical meaning of the parameter; the special case of polarization is addressed below. It is also assumed that $|\psi(\xi)\rangle$ is the ground state, nondegenerate for any ξ. The phase difference between the ground eigenstates at two different ξ points is defined in the most natural

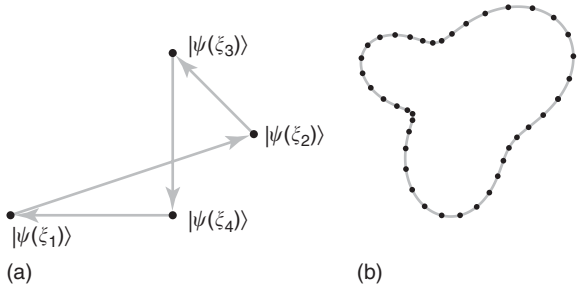


Figure 5 (a) Four state vectors in parameter space, and a closed path connecting them. (b) A smooth closed curve C in parameter space, and its discretization.

way:

$$e^{-i\Delta\varphi_{12}} = \frac{\langle\psi(\xi_1)|\psi(\xi_2)\rangle}{|\langle\psi(\xi_1)|\psi(\xi_2)\rangle|} \quad [17]$$

$$\Delta\varphi_{12} = -\mathrm{Im}\log\langle\psi(\xi_1)|\psi(\xi_2)\rangle \quad [18]$$

This phase, defined modulo 2π, cannot have any physical meaning. In fact, any quantum-mechanical state vector is arbitrary by a constant phase factor: one may refer to choosing this phase as to the choice of the gauge. A simple change of gauge at ξ_1 and/or at ξ_2 will give to $\Delta\varphi_{12}$, the most arbitrary value. Despite this fact, when one considers the total phase difference along the closed path in **Figure 5a**:

$$\begin{aligned}\gamma &= \Delta\varphi_{12} + \Delta\varphi_{23} + \Delta\varphi_{34} + \Delta\varphi_{41} \\ &= -\mathrm{Im}\log\langle\psi(\xi_1)|\psi(\xi_2)\rangle\langle\psi(\xi_2)|\psi(\xi_3)\rangle \\ &\quad\times\langle\psi(\xi_3)|\psi(\xi_4)\rangle\langle\psi(\xi_4)|\psi(\xi_1)\rangle \end{aligned} \quad [19]$$

one easily realizes that γ is gauge invariant, since all the gauge-arbitrary phases cancel in pairs.

The Berry phase is generally defined in the continuum limit. Suppose a smooth closed curve C in the parameter domain, such as in **Figure 5b**, is discretized with a set of points on it: it is assumed that the gauge is so chosen that the phase is single-valued and varies in a differentiable way along the path. The phase difference between two contiguous points is

$$e^{-i\Delta\varphi} = \frac{\langle\psi(\xi)|\psi(\xi+\Delta\xi)\rangle}{|\langle\psi(\xi)|\psi(\xi+\Delta\xi)\rangle|} \quad [20]$$

$$-i\Delta\varphi \simeq \langle\psi(\xi)|\nabla_\xi\psi(\xi)\rangle\Delta\xi \quad [21]$$

The total phase difference γ converges to the circuit integral of a real linear differential form, called the Berry connection:

$$\gamma = \sum_{s=1}^{M}\Delta\varphi_{s,s+1} \to i\oint_C\langle\psi(\xi)|\nabla_\xi\psi(\xi)\rangle d\xi \quad [22]$$

The above simple-minded algebra leads to a result of overwhelming physical importance: in fact, a gauge-invariant quantity is potentially a physical observable. Traditionally, one is accustomed to regard any observable effect in quantum mechanics as the eigenvalue of some operator. But in 1984, in a milestone paper, M V Berry pointed out that there may be observable effects of a completely different nature: γ cannot be expressed in terms of the eigenvalues of any Hermitian operator, whereas it is, by definition, a gauge-invariant phase of the state vector. Having understood this, one must reexamine the initial assumption of a parametric Hamiltonian, and discuss its most fundamental meaning. In general, a quantum system having a parametric dependence in its Hamiltonian is not isolated: the parameter schematizes a kind of coupling with other variables not included in the Hilbert space, or more generally with "the rest of the Universe," to use Berry's words. The parametric Hamiltonian allows one to deal with a part of a larger system as if it were isolated: as a trade-off, some observable effects may occur as gauge-invariant phases. As far as polarization is concerned, the coupling to the rest of the universe is perspicuous in **Figures 1a** and **3**, which refer to ideal measurements closely related to the theoretical approach outlined below.

How the Modern Theory of Polarization Works

One may show how the polarization problem fits into the general Berry-phase setting: what the parameter ξ is; what the Berry connection of the problem is; which is the closed path C allowing one to evaluate ΔP as a Berry phase. The article is limited to giving a flavor of how the theory works, using the specific example of the toy model of **Figure 4** and without providing any proof (see the "Further reading" section for more details).

A scalar parameter λ is first introduced, which measures the amplitude of the ferroelectric distortion in dimensionless units: $\lambda = 0$ for the paraelectric structure (**Figure 4**, top), and $\lambda = 1$ for the ferroelectric one (**Figure 4**, bottom). Since the electric field is zero (see also **Figure 3**), for any value of λ the electrons move in a lattice-periodical potential $V^{(\lambda)}(x)$, having period a. In the simplest model insulator, there are only two electrons per cell, hence one doubly occupied band, whose orbitals have the form:

$$\psi^{(\lambda)}(k,x) = e^{ikx}u_k^{(\lambda)}(x) \quad [23]$$

where k is the Bloch vector in the one-dimensional Brillouin zone $[-\pi/a, \pi/a]$, and $u_k^{(\lambda)}(x)$ is a periodic function of x with period a. The Schrödinger

equation written in terms of u reads

$$\left[\frac{\hbar^2}{2m_e}\left(-i\frac{d}{dx}+k\right)^2+V^{(\lambda)}(x)\right]u_k^{(\lambda)}(x) = \varepsilon_k^{(\lambda)}u_k^{(\lambda)}(x) \quad [24]$$

In order to apply the above general Berry-phase formulation to this Hamiltonian, it is enough to define a two-dimensional parameter $\boldsymbol{\xi}$ whose two components are k and λ:

$$H(\boldsymbol{\xi}) \equiv \frac{1}{2m_e}(p+\hbar k)^2 + V^{(\lambda)}; \quad |\psi(\boldsymbol{\xi})\rangle \equiv |u_k^{(\lambda)}\rangle \quad [25]$$

The Berry connection of the problem is

$$i\langle\psi(\boldsymbol{\xi})|\nabla_{\boldsymbol{\xi}}\psi(\boldsymbol{\xi})\rangle\cdot d\boldsymbol{\xi} = i\left\langle u_k^{(\lambda)}\middle|\frac{\partial}{\partial k}u_k^{(\lambda)}\right\rangle dk + i\left\langle u_k^{(\lambda)}\middle|\frac{\partial}{\partial\lambda}u_k^{(\lambda)}\right\rangle d\lambda \quad [26]$$

The Berry phase [22] is evaluated, where the curve C is chosen as the contour of the rectangle shown in **Figure 6**. According to the modern theory, the polarization difference ΔP between the ferroelectric and paraelectric structures is

$$\Delta P = \frac{q_e}{\pi}\gamma + \Delta P_{ion} \quad [27]$$

where ΔP_{ion} is the trivial contribution of the classical ions.

The three-dimensional, multiband analog of eqn [27] is nowadays implemented – at the first-principle level and using various technical ingredients – in most publicly available electronic-structure codes. Applications have concerned piezoelectricity, lattice dynamics, and ferroelectricity in real materials. In the case of ferroelectric polarization, where no quantum-mechanical theory was available prior to 1993, the numerical results are in spectacular agreement with the experimental data and provide an insight into the physics of ferroelectricity.

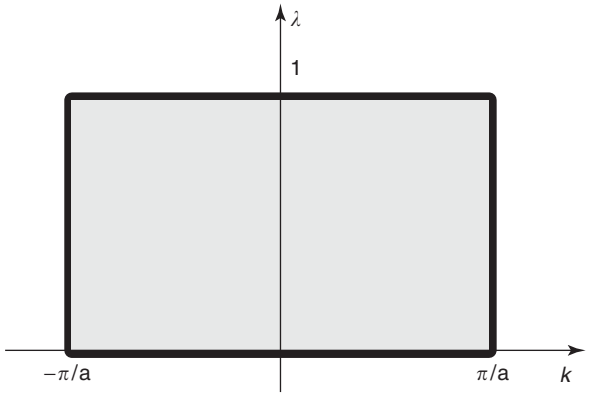


Figure 6 The Berry phase γ, which defines the electronic contribution to ΔP is evaluated along the contour of the rectangular domain in the (k, λ) plane.

Conclusions

The modern theory of polarization provides an unambiguous definition of what the physical observable really is: an integrated current, whose electronic contribution can be evaluated either from the linear-response theory, or as a gauge-invariant phase of the occupied crystalline orbitals. Macroscopic polarization bears no relationship to the periodic charge density of the polarized crystal: the latter, in fact, only depends on the square modulus of the occupied orbitals, where any phase information is deleted. Besides providing profound insight into the polarization phenomenon, the modern theory is also an extremely powerful tool for calculations in real materials.

Acknowledgment

This work is partially supported by the US Office of Naval Research through grant N00014-03-1-0570.

See also: Ferroelectricity; Piezoelectricity.

PACS: 42.25.Ja; 42.68.Mj; 42.81.Gs; 77.22.Ej

Further Reading

Baroni S, Giannozzi P, and Testa A (1987) Green's-function approach to linear response in solids. *Physical Review Letters* 58: 1861.

Baroni S, de Gironcoli S, Dal Corso A, and Giannozzi P (2001) Phonons and related crystal properties from density-functional perturbation theory. *Reviews of Modern Physics* 73: 515.

Berry MV (1984) Quantal phase factors accompanying adiabatic changes. *Proceeding of the Royal Society of London A* 392: 45.

Dreizler RM and Gross EKU (1990) *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Berlin: Springer.

King-Smith RD and Vanderbilt D (1993) Theory of polarization of crystalline solids. *Physical Review B* 47: 1651.

Resta R (1994) Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Review of Modern Physics* 66: 899.

Resta R (1998) Quantum mechanical position operator in extended systems. *Physical Review Letters* 80: 1800.

Resta R (2000) Manifestations of Berry's phase in molecules and in condensed matter. *Journal of Physics: Condensed Matter* 12: R107.

Resta R (2002) Why are insulators insulating and metals conducting? *Journal of Physics: Condensed Matter* 14: R625.

Resta R (2003) *Ab-inito* simulation of the properties of ferroelectric materials. *Modelling and Simulation in Materials Science and Engineering* 11: R69.

Resta R, Posternak M, and Baldereschi A (1993) Towards a quantum theory of polarization in ferroelectrics: the case of $KNbO_3$. *Physical Review Letters* 70: 1010.

Shapere A and Wilczek F (eds.) (1989) *Geometric Phases in Physics*. Berlin: Springer.

Souza I, Íñiguez J, and Vanderbilt D (2003) First-principles approach to insulators in finite electric fields. *Physical Review Letters* 89: 117602.

Vanderbilt D (1997) First-principles based modelling of ferroelectrics. *Current Opinions in Solid State and Materials Science* 2: 701.

Nomenclature

$\boldsymbol{E}$	macroscopic electric field (inside the material)
m_e	electron mass
$\boldsymbol{p}$	canonical momentum (equal to $-i\hbar\nabla$ within the Schrödinger representation)
$\boldsymbol{P}$	macroscopic dielectric polarization
$\boldsymbol{P}_{el}$	electronic contribution to the macroscopic dielectric polarization
$\boldsymbol{P}_{spont}$	spontaneous polarization in ferroelectrics
q_e	electron charge ($= -\|e\|$)
ε_∞	"static high-frequency" dielectric constant (or tensor)

Glasses

J M Parker, Sheffield University, Sheffield, UK

Introduction

Glasses feature widely in everyday life and also in science. Their transparency and chemical durability, combined with ease of fabrication, ensure a role in windows, containers, tables, and laboratory ware; they are key components in optical instrumentation, for example, microscopes, cameras, and optical fibers. Some devices utilize specific interactions between light and glass, for example, lasers. Other applications use their electrical and magnetic behavior, strength, and thermal characteristics, including crystallization and phase separation. The range and flexibility of known glass compositions allow their properties to be independently tailored. Combining glasses with other materials or modifying their behavior by coating has further extended their use. Understanding glass structure and its consequences is also challenging the boundaries of condensed matter physics.

What Is a Glass?

Glasses have been defined as inorganic products of fusion cooled to a solid without crystallizing. This encapsulates key properties: glasses have a liquid-like structure, but cooling has caused a transition from liquid to solid behavior. It also implies features that are now disputed: organic glasses exist, and melting is not the only route for glass making. A characteristic implicit in the definition is that crystallization rates are low; this is often associated with high melt viscosities below the liquidus (T_L). Glasses differ from amorphous solids, in that they can be reheated to the liquid state; the latter crystallize on heating before melting.

How Are Glasses Made?

Often glasses are made from solids by heating until fully melted, bubbles have dissipated and the melt is homogeneous. Commercially, temperatures may reach 1600°C and melting may take days. Laboratory melts need a few hours. Density separation of batch components, differential melting rates, contamination by the container, and volatilization losses can limit the glass quality. High melt viscosities limit diffusion and militate against homogeneity; mechanical or convection-driven stirring may be necessary. The oxidation state of any multivalent ions depends on the atmosphere around the batch during initial melting; subsequently, equilibrium with the furnace atmosphere is achieved only slowly. Rapid quenching may be needed, for example, roller quenching, splat cooling onto a rotating copper wheel, or fiber pulling.

Alternative production routes include: condensation of vapor phase reaction products on cold substrates; hydrolysis of organic derivatives of silicon (e.g., $Si(OC_2H_5)_4$) or other components to give a gel that is dried and sintered to a fully dense product; and extended grinding or applying pressure to crystals.

Glass Structure

Short-Range Order

Glass structures have been studied by X-ray or neutron diffraction, various spectroscopic techniques, and computer modeling. Their short-range structure often mirrors that of corresponding crystalline compounds, particularly higher temperature phases. Zachariasen postulated that single-component glass-forming systems have cations with low coordination numbers (3 or 4) surrounded by bridging anions, each linking two coordination polyhedra at corners

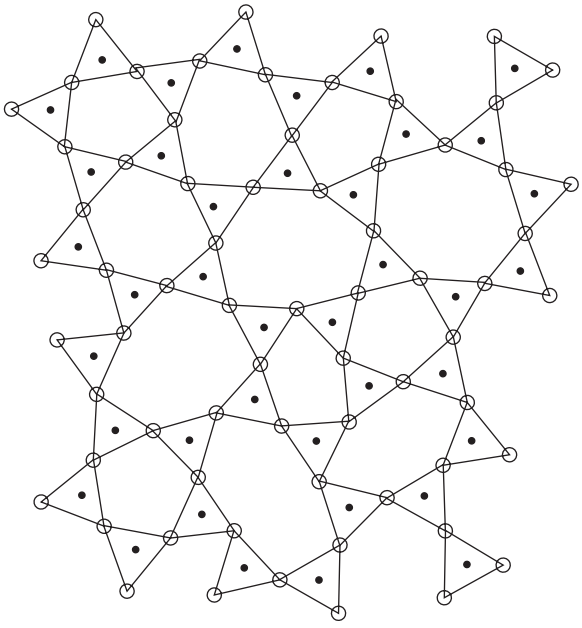


Figure 1 A schematic representation of the structure of vitreous silica. The tetrahedral SiO_4 units in silica are represented by triangular units.

rather than edges. SiO_2 consists of a matrix of SiO_4 tetrahedra (**Figure 1**) that defines its low thermal expansion coefficient (TEC), high viscosity at T_L, and high melting point. Other single-component glass-forming systems include GeO_2, BeF_2, B_2O_3, and As_2O_3.

A second component can extend glass formation, for example, to 58 mol.% Na_2O in the Na_2O–SiO_2 system. Each oxygen added breaks one corner linkage in the network, creating two nonbridging oxygens (**Figure 2**). Silicons are called "network formers" and sodiums, which enter interstices surrounded by bridging and nonbridging oxygens, are termed "network modifiers." A silicon linked to three bridging and one nonbridging oxygen is denoted by Q_3; the spread of Q_n species is generally narrow. Adding soda (Na_2O) increases the TEC and reduces viscosity by depolymerizing the network, while increasing density by filling interstices. Various cations adopt a network-forming role (e.g., Ge, P, Pb). The melt structure, being flexible, can accommodate many modifier ions in spite of different radii and preferred coordination, for example, Li, K, Rb, Mg, Ca, Fe^{2+}. Intermediate ions (Al^{3+}, Fe^{3+}) can take either role in the structure. Some oxide combinations form glasses even though the components do not – the network-forming ion is called a "conditional" glass former, for example, Al in CaO–Al_2O_3 glasses. The oxygens can also be (partially) replaced by other anions, for example, F, S, N.

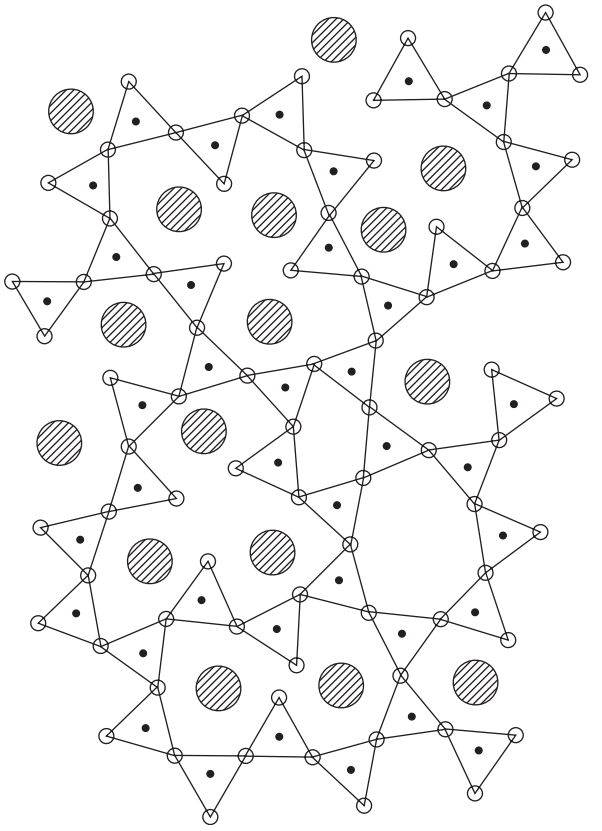


Figure 2 A schematic representation of the structure of an Na_2O–SiO_2 glass illustrating the break-up of the network by the introduction of nonbridging oxygens (linked to only one Si), and the sites taken by the sodium ions.

Properties can often be modeled using expressions that are linear functions of composition and the viscosity of industrial glasses is usually calculated rather than measured. Difficulties occur if a cation changes role with composition (e.g., B coordination depends on alkali content); the TECs of M_2O–B_2O_3 glasses display a minimum with composition (the "boric oxide anomaly"), requiring a $[B_2O_3]^2$ term in composition models. Liquidus temperatures, which are discontinuous across phase boundaries, are best modeled using the underlying free energy functions. Components can also interact: Na_2O decreases network connectivity, while Al_2O_3 is a network former at low concentrations and, being oxygen deficient, removes nonbridging oxygens. If $Na_2O{:}Al_2O_3 = 1$, a highly connected, viscous system results; excess Na_2O or Al_2O_3 decreases viscosity, in the latter case because Al coordination changes from [4] (former) to [6] (modifier).

Medium-Range Order

Diffraction experiments give structural data in the form of radial distribution functions, representing the

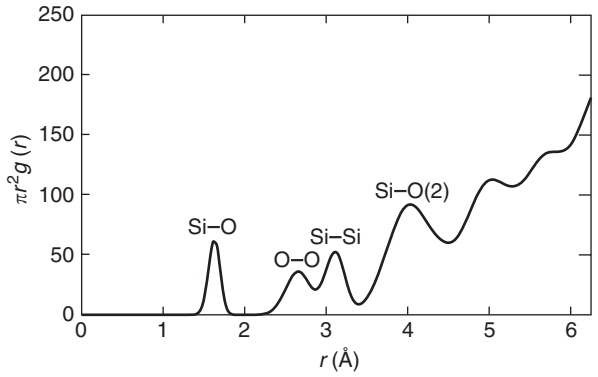


Figure 3 An idealized radial distribution function for SiO_2. The first four peaks give the distribution of nearest neighbor Os around Si, the Si–Si distribution, the O–O distribution, and the distribution of second nearest neighbor oxygens around Si. The peak areas give coordination numbers and the breadths indicate thermal and structural disorder. A random distribution would give a parabolic curve. Experimental curves show some ripple at low r caused by data termination errors. Many authors now plot correlation functions ($\pi r\, g(r)$) rather than r.d.f.s.

summed distribution of atoms with distance around all species present (**Figure 3**); network-forming ions sit in remarkably regular sites but longer distance correlations are harder to resolve because of peak overlap. Extended X-ray absorption fine structure (EXAFS) experiments allow study of the environment of individual species and suggest that network modifiers also sit within polyhedra of notable regularity, that is, all cations optimize their local structure.

Such distribution functions show that the network-forming polyhedra can also form extended structures, for example, vitreous B_2O_3 consists of BO_3 triangles, that form planar groupings, termed boroxol rings (B_3O_6), not found in B_2O_3 crystals. The modified random network model for alkali silicate glasses is based on segregation of network-modifying atoms into layers surrounding SiO_2-rich regions (**Figure 4**). The diffraction patterns for many glasses have a sharp low-angle first diffraction peak, consistent with extended ordering, as confirmed by modeling and electron microscopy.

IR spectroscopy supports such structural models by accounting for the network vibrational modes. In Raman and Brillouin scattering experiments, glasses have, as a common feature, a boson peak at $\sim 100\,\mathrm{cm}^{-1}$ that provides additional insight into nanometer-scale inhomogeneities. Models incorporating medium-range order are distinct from older microcrystallite theories of glass, discounted by the lack of small angle X-ray or neutron scattering beyond that predicted from incipient random fluctuations in density.

These conclusions underpin the thermodynamic modeling of glass properties based on contributions

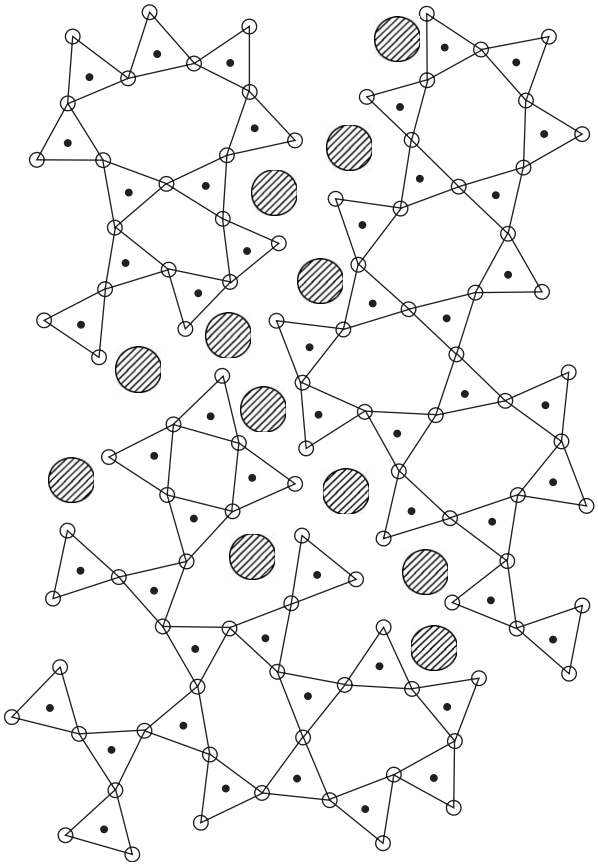


Figure 4 A schematic representation of the modified random network model for an Na_2O–SiO_2 glass showing the tendency for modifier ions to segregate.

from components with appropriate compositions in defined proportions.

Defects

Glasses can contain bonding and electronic defects induced by radiation or high temperatures. Their concentrations are low but they influence optical properties by introducing visible or ultraviolet (UV) absorptions and causing refractive index (r.i.) changes. Defects may involve unpaired electrons (*) and include: the nonbridging oxygen hole center, $\equiv Si-O^*$; E′ center, $\equiv Si^*$; peroxy radical, $\equiv Si-O-O^*$; or species such as Zr^{3+}, F_2^-. Some defects can be eliminated by optical or thermal bleaching, while others survive for years under normal conditions.

Nonsilicate Glass-Forming Systems

Silicate glasses have three-dimensional (3D) network structures with intermediate ionic/covalent bonds. Alkali phosphate glasses have similar bonding but are often better described as chains. Chalcogenide glasses are largely covalent and based on networks,

sheets, chains, and rings. Hydrogen-bonded molecular glasses are common in organic systems (e.g., sugars), while some highly ionic or metallic systems adopt almost random close-packed structures. Various salts (nitrates, sulfates, acetates) also form glasses. Even crystals can show glassy behavior. Spin glasses may form in magnetic systems, based on disordered arrangements of unpaired electron spins, or molecules may have random orientations on a regular lattice as in KBr–KCN. Polymers usually consist of covalently bonded chains, which may have irregular monomer unit sequences or cross-link randomly along their lengths, giving glassy materials at low temperatures that cannot untangle into crystalline forms. In extreme cases, oxide glass melts have been described with two or more configurations that transform only slowly from one to the other, or may coexist in equilibrium.

Crystal Growth and Phase Separation

Once a melt is cooled below T_L, crystal growth is possible. The driving force (difference in free energy between crystalline and liquid state) increases with undercooling and initially growth rates rise as temperature falls (**Figure 5**). Subsequently kinetic factors, for example, diffusion through the melt to the growing crystal and atomic rearrangement at the interface, limit growth. In stable glass-forming systems, growth rates may be only $\mu m\,h^{-1}$ but may reach $m\,s^{-1}$ in glassy metals. For large undercoolings the fastest growing phase is not necessarily the primary phase defined by the phase diagram; metastable phases also often form.

Crystallization requires a nucleation step. Homogeneous nucleation is an intrinsic property. The free energy gain on producing a crystalline nucleus must balance the energy penalty of generating its surface. Surface effects dominate for small particles and so nuclei must exceed a critical size before they are viable. Small undercoolings require a large critical nucleus whose spontaneous formation is improbable, so melts can exist metastably for long periods just below T_L. At lower temperatures, the critical nucleus size decreases and nucleation rates rise rapidly to a maximum just above T_g, subsequently falling as kinetic factors dominate.

Homogeneous nucleation is, however, rare. Nucleation starts more easily from pre-existing discontinuities, for example, "dirty" surfaces or container walls. Nucleation can be engineered by dissolving at high temperatures a component that becomes insoluble on cooling and precipitates (by homogeneous nucleation or phase separation). Examples include halides (NaF, CaF_2, CuCl), oxides (TiO_2, ZrO_2, P_2O_5), or metals (Ag, Au).

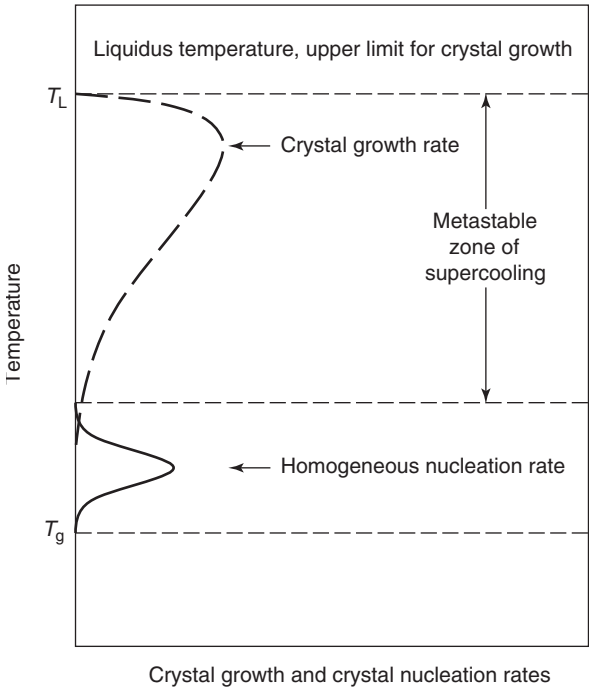


Figure 5 A schematic representation of the crystal growth and nucleation curves for a single phase precipitating from a glass-forming melt. In reality other phases may also appear at different temperatures with their own growth rate curves.

If homogeneous nucleation rates are high, two-step heat treatment in the nucleation and growth regions can give a finely crystalline structure. The resulting "glass ceramics" are made by standard melting and forming technology but, once crystallized, can have enhanced properties, while lacking the porosity intrinsic to conventional ceramics. Glass ceramics can remain transparent if the crystals are smaller than the wavelength of light and r.i. differences are minimized.

Phase Separation

Phase separation is the chemical segregation of a melt into two liquids, driven by an associated reduction in free energy, for example, Na_2O–SiO_2 melts, with $Na_2O < 20$ mol.%, separate into SiO_2 and sodium silicate; the SiO_2 structure is relatively intolerant to impurities. The minor phase may occur as droplets or a continuously connected phase, whose scale depends on heat treatment time and whose microstructure can be preserved by quenching. The initial step may involve homogeneous nucleation or spinodal decomposition (no energy barrier). Borosilicate glasses also phase separate; the connected phase can be removed by solution, allowing manufacture of finely porous SiO_2.

Glass Formation

First, the flexibility of linkages in network systems results in only small energy differences between glassy and crystalline states (cf. the many SiO_2 polymorphs). This is a key thermodynamic factor in glass formation. Second, the network connectivity and bond strength cause the high melt viscosity and slow crystallization rates. Finally, the high melting temperatures destroy, by reaction or solution, foreign particles that might otherwise catalyze crystallization.

Mathematical models, based on crystal growth and nucleation behavior, and their degree of overlap with temperature, can predict critical cooling rates for glass formation. They suggest that many systems, including some metal alloys, can form glasses. Cooling rates of $10^{6\circ}C\,s^{-1}$ may be needed, so heat transfer limits their production to thin artifacts, for example, ribbons or wires. New glass-forming systems have been identified, based on TeO_2, Al_2O_3, Ga_2O_3, ZrF_4, InF_3, by using faster quenching, smaller samples, and purer materials melted under clean conditions to eliminate nucleating agents.

Glasses are often associated with eutectics where liquidus temperatures are low and the melts differ most in composition (and structure) from crystalline compounds. Adding more components can also improve glass-forming ability – the so-called "confusion" or "garbage-can" principle.

Viscosity

Melt viscosities (η) typically follow the Vogel–Tammann–Fulcher equation (VTF) with temperature: $\log \eta = A + B/(T - T_0)$ with A, B, and T_0 as fitted constants. At the temperature where $\eta = 10^{12}\,Pa\,s^{-1}$, deformation rates are slow, the material is effectively solid over short times ($\sim 10^2$ s), and this provides one definition of the "glass transition temperature, T_g." By the strain point ($\eta = 10^{13.6}\,Pa\,s^{-1}$), stress release takes hours. First-order thermodynamic properties such as volume are continuous through the glass transition, in contrast to phase changes such as crystallization, but second-order thermodynamic properties (specific heat (C_p) or TEC) are stepped (**Figure 6**). T_g can therefore also be measured from C_p or length changes using a differential scanning calorimeter or a dilatometer, but the values depend on the heating rate used, since the transition is kinetically defined, and are not necessarily equal for different properties.

Some have concluded that glass flows at room temperature, it being a liquid, and have perpetuated a myth that old windows are thicker at their base. Neither measurements nor calculations substantiate such conclusions, flow having become immeasurably slow. Where old panes are thicker at one edge, it is a manufacturing defect.

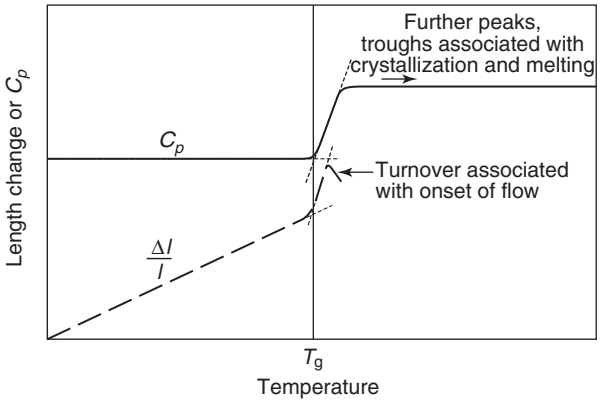


Figure 6 An idealized graph of the specific heat of a glass as a function of temperature during a constant heating rate experiment and the length of a glass specimen as a function of temperature under similar heating conditions. There is a step in C_p at T_g and a change of slope (corresponding to a step in the TEC) in the specimen length.

A glass artifact, cooled quickly, will solidify with large temperature gradients present; the differential contraction on reaching room temperature can induce significant stresses. These stresses depend on cooling rate, TEC, elastic properties, and thickness; bulk articles are therefore annealed, by equilibration above T_g, followed by slow cooling to the strain point.

Angell has compared glass-forming systems using normalized plots of $\log \eta$ against T_g/T; those displaying Arrhenius behavior most closely he termed "strong" glasses while those with large deviations he termed "fragile." Fragility implies cooperative breakdown of the melt structure into increasingly small units above T_g. The VFT formalism predicts $\eta = \infty$ at $T \to T_0$. Many studies have attached physical significance to T_0, relating it to the temperature where extrapolation of experimental data predicts that the entropy of a glass becomes equal to that of the corresponding crystal, that is, the configurational entropy falls to zero, a concept underpinning the "Kauzmann paradox."

Processing

High melt viscosities, along with gravity and surface tension, aid shaping as the melt cools. Sheet glass achieves optical flatness by gravity-driven flow, while floating on a molten tin bath with a 7 mm equilibrium thickness defined by surface tension and density. Optical fibers maintain their r.i. profiles during drawing. Containers and laboratory ware are blown and pressed by complex machinery to give the desired glass distribution. Strain rates may be high.

Some glasses display non-Newtonian flow by structural alignment during forming.

Transformation Range Effects

Glass structure does not depend only on composition but also on thermal history; densities may vary by 1%. Fast quenching freezes in a structural state corresponding to equilibrium at a higher temperature than slow cooling. The effective quenched-in equilibrium temperature is called the "fictive temperature." Control of cooling through the transformation range allows fine tuning of an optical glass r.i., and low cooling rates ensure the whole sample has the same thermal history (and minimum stress). Optical glass manufacturers recommend that their products should not be heated above $T_g - 150°C$ because of potential property changes. Similar considerations apply to the processing of glass sheets for LCD displays. The electronic circuitry used to address individual pixels is printed at high temperatures in several layers; precise registry demands dimensional stability.

A glass heated in the transformation range may, after an initial instantaneous change, relax exponentially to a new equilibrium state (Maxwell spring and dashpot model). Stress release rates can be measured and accurate data as $f(t, T)$ are needed for high-precision fabrication. Detailed analysis suggests a stretched exponential formulation: $S_t = S_0 \exp(-[t/\tau]^n)$ where S_t is the stress at time t, S_0 the initial stress, and τ is proportional to viscosity, n ranges typically from 0.5 near T_g to 1. Many other properties behave similarly although not necessarily with the same constants for a given glass. Study of relaxation behavior and structural dynamics has given fruitful insights into the nature of the glass transition.

At lower temperatures, bulk flow ceases but atoms can slide into new relative positions corresponding to different local energy minima, associated with the disordered structure. Such effects are manifest in internal friction experiments, for example, measurement of the damping of a pendulum suspended from a glass fiber under torsion. Damping involves transfer of mechanical energy to thermal energy and is observed down to room temperature in silicate glasses. Different mechanisms, termed α and β relaxations, dominate as temperature rises. These are attributed to mechanisms from modifier ion motion and relative movement of nonbridging oxygens, to motion of whole structural units.

A related phenomenon is semipermanent densification of v-SiO_2 at high pressures (GPa). Deformations of several percent do not fully relax on return to normal atmospheric conditions. Kinetic studies of relaxation suggest a low activation energy process, which is therefore not bond breakage; Raman spectroscopy confirms this.

Early mercury-in-glass thermometers also displayed delayed elasticity. After calibration at 100°C, the mercury failed to return immediately to its 0°C calibration point on cooling. This relaxation behavior was minimized in mixed alkali glasses. Similar effects have been observed in rare-earth doped glasses; optical pumping into particular excited states, and the consequent energy transfer to the lattice, has excited structural changes that only recover at raised temperatures. The associated r.i. changes allow holographic image storage.

Thermal Properties

Specific heats are consistent with Dulong and Petit's law near T_g but show a positive step at T_g, the size of which indicates the level of structural breakdown occurring.

Thermal conductivities are lower than for equivalent crystalline phases and decrease with falling temperature, consistent with the short phonon mean free paths associated with structural disorder. At high temperatures, heat transport by photons becomes significant – glass is semitransparent in the near infrared allowing absorption and re-emission as a heat transport mechanism, termed radiation conductivity. Radiation conductivity varies as T^3 and is significant in commercial glass melting.

Mechanical Behavior

Glasses are brittle, with strengths dominated by their surface state. In microhardness tests, they exhibit limited flow under compression, attributed to densification and faulting. No flow occurs under tensile stress however; surface flaws remain atomically sharp at high loads and are effective stress raisers (stress at crack tip is enhanced by $(r/\rho)^{1/2}$ where r is crack length and ρ tip radius).

Flaws are induced by impact, friction, adhering particles of different TEC, and devitrification. Strengths of several GPa in undamaged fibers can fall to < 50 MPa. Water attack at flaw tips under stress causes static and dynamic fatigue, which is maximized in glasses of poor chemical durability; conversely, v-SiO_2 is unusually strong because of its excellent durability. The stress intensity factor, K_{1c}, allows isolation of compositional factors by eliminating the effects of flaw size. Variations between glasses are small and related to bond strengths or atomic packing densities.

Strengths are improved by protective coating or surface compressive stresses induced by rapid cooling or ion exchange, where large cations replace smaller

cations. Many glass ceramics have enhanced strengths, the crystals impeding crack growth.

Optical Properties

Glasses are intrinsically transparent because they lack internal structure (e.g., grain boundaries) where r.i. changes would scatter light. Transparency also requires the absence of absorption mechanisms. UV photons can excite bonding electrons into empty molecular orbitals, limiting transparency. v-SiO_2 and phosphate glasses have strongly bound electrons and are transparent to <200 nm, but nonbridging oxygens, with their more weakly bound electrons, reduce transmission. IR photons can excite atomic vibrations and SiO_2 is opaque beyond 4.5 μm. OH absorptions usually limit transmission beyond 3 μm. Transmission can be extended by reducing the natural vibrational frequencies using heavier ions and weaker bonding. Fluoride glasses based on singly charged anions and heavy cations (Zr, La, In, Hf) will transmit to 7.5 μm and chalcogenide glasses (based on S, Se, Te) can transmit to 20 μm.

The UV and IR absorptions in glass are responsible for its r.i., n, and wavelength dispersion, v, as implicit in the semi-empirical fitting formula of Helmholtz–Ketteler: $n^2 - 1 = A + \sum_i c_i \lambda^2/(\lambda_{0i}^2 - \lambda^2)$ with $\lambda_{0i} =$ wavelength of absorption bands and c_i a weighting factor related to band strength. SiO_2 has a low n because its strong bonding gives short λ_{0i} values, whereas PbO-rich glasses have longer λ_{0i}, and correspondingly higher n and v values. The UV absorption is associated with the $6s \rightarrow 6p$ transition of the lone pair electrons in Pb^{2+}. These weakly bonded, highly polarizable electrons also increase the second-order coefficient of optical nonlinearity (n_2). $d^2n/d\lambda^2 = 0$ near 1.5 μm for SiO_2 because of the competing effects of UV and IR absorptions, minimizing materials-based signal dispersion in fibers.

Transition metal and rare-earth ions introduce weak absorptions, which are broader than in crystals because of the inherent structural disorder. Such species color glass. The matrix composition can strongly influence excited state lifetimes, and affects absorption and emission peak shapes. Halide and chalcogenide glasses are often preferred hosts because their low phonon energy lattices increase excited state lifetimes. Crystallization, where dopant ions segregate into the crystal phase, can offer similar advantages but with more tightly defined dopant ion environments. Both glassy and transparent glass–ceramic materials make efficient lasers (in fiber form for telecommunications).

Color can result from precipitated particles, for example, Au, Cu, and Ag color glass red or yellow and Cu(Cl,Br), Cd(S,Se), or PbTe give sharp absorption edges in the UV, visible, or IR, depending on their bandgap. If particles exceed 20 nm, light scattering can become excessive. The electronic energy levels of these precipitates are influenced by particle size via quantum confinement effects, giving rise to enhanced optical nonlinearities.

Such materials require dissolution of an appropriate compound at high temperatures that will precipitate on cooling. Metals are dissolved as ions, reduced to atoms on cooling by a redox process (e.g., $Au^+ + Ce^{3+} \rightarrow Au^0 + Ce^{4+}$), and subsequently agglomerate. Electron transfer can be catalyzed by UV photons. Thus regions of high metallic particle concentrations, capable of nucleating bulk crystallization, can be engineered. The crystallized regions can display differential HF etching rates, and finely structured products for fluidics or replacement bones for the ear can be made.

Glasses are rarely birefringent being intrinsically isotropic, but anisotropy can be induced temporarily by stress, or permanently by shearing melts with chain-like structures or showing phase separation, while cooling through T_g.

The Faraday effect involves rotation of the plane of polarization of a light beam by a magnetic field. Applications include optical devices requiring rapid switching or protection against back-reflected light. The plane of polarization is rotated in one sense by diamagnetic materials and in the reverse by paramagnetic materials. Large diamagnetic effects occur in $PbO–Bi_2O_3–Ga_2O_3$ glasses because of the polarizable Pb and Bi ions, while typical paramagnetic materials contain rare-earth ions with unpaired $4f$ electrons, for example, Tb^{3+}, Pr^{3+}, and Dy^{3+}. The equivalent effect for electrical fields is named after Kerr.

Energetic photons can form point defects or rearrange bonds, causing density and r.i. changes that may be used for laser writing; Bragg reflecting gratings, for telecoms signal demultiplexing or stress sensors, are written in Ge-doped SiO_2 optical fibers by two interfering beams from an excimer laser. Light-induced differential etching rates in chalcogenide glasses allow manufacture of photodiffractive elements.

Diffusion and Electrical Behavior

SiO_2 has an open structure allowing solution of small molecules (e.g., He, H_2) and significant diffusion rates, even at 20°C. Modifier ions block the diffusion paths but, being weakly bonded, may themselves diffuse below T_g. Some surface modification technologies, for example, chemical toughening or planar

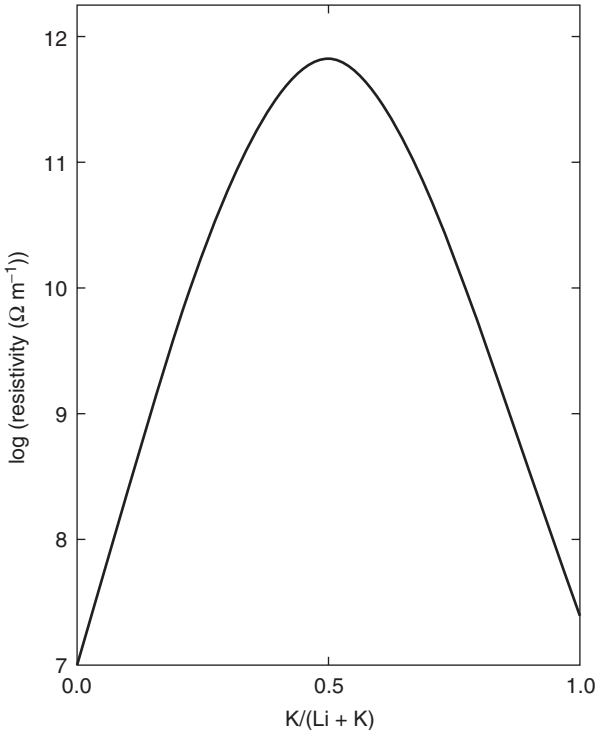


Figure 7 An idealized graph of resistivity against the concentration of K as a fraction of total alkali content [K/(Li + K)] for an M_2O–$2SiO_2$ glass. The peak in resistivity is a result of the mixed alkali effect.

waveguide manufacture, use this. At room temperature, most glasses are insulators but ion motion can cause significant AC dielectric loss; consequently glass ceramics are often preferred substrates in the electronics industry. Glasses become ionic conductors at higher temperatures ($\sim$100°C for silicate glasses with significant modifier ion content; larger ions and higher charge reduce conductivity). The phenomenon of fast ion conduction, observed in some crystalline ceramics, also occurs in glasses.

The mixed alkali effect is a characteristic of ionic conduction and diffusion. **Figure 7** plots the conductivity of glasses containing Na and K against concentration; a pronounced minimum occurs where Na:K = 50:50, explained by the different sites the two ions occupy. Sodiums hop easily between equivalent sodium sites but mixed sites create an energy barrier. The loss of vibrational coupling between adjacent ions may also reduce coordinated hopping. The mixed alkali effect is most pronounced below T_g, the cation sites being unable to relax.

Both crystalline and glassy semiconductors display valence and conduction bands, and a bandgap. Interband states exist even in undoped glasses however, particularly near the band edge. Chalcogenide glasses, based on S, Se, Te with, for example, Ge, Si, As, Sb, Ga, and glasses with high concentrations of transition metals in mixed valence states (e.g., V, Fe) can be semiconducting, electrons hopping between adjacent ions in different oxidation states in the latter case. Chalcogenide glasses have typical semiconductor characteristics, for example, photoconductivity and xerography; a-Si is a common component of solar cells. They also display switching behavior. Low voltages give small currents but higher voltages can induce crystalline filaments of higher conductivity. If crystallization is reversible, a memory switch results.

See also: Ceramics, History of; Disordered Solids and Glasses, Electronic Structure of; Insulators, Optical Properties of; Ionic and Mixed Conductivity in Condensed Phases; Optical Instruments; Optical Properties of Materials; Phases and Phase Equilibrium.

PACS: 42.70.Ce; 61.43.Fs; 62.20.Mk; 64.70.Pf; 65.60.+a; 66.30.Hs; 71.23.Cq; 71.55.Jv; 77.22.Gm; 78.20.Ci; 78.30.Ly; 78.40.Pg; 81.05.Gc; 81.05.Kf; 81.05.Pj; 81.20.Fw

Further Reading

Babcock CL (1977) *Silicate Glass Technology Methods*. New York: Wiley.
Bach H (1995) *Low Thermal Expansion Glass Ceramics*. Berlin: Springer.
Bamford CR (1977) *Colour Generation and Control in Glass*. Amsterdam: Elsevier.
Brawer S (1985) *Relaxation in Viscous Liquids and Glasses*. Columbus: American Ceramic Society.
Cable M and Parker JM (1992) *High Performance Glasses*. Glasgow: Blackie.
Doremus RH (1994) *Glass Science*. New York: Wiley.
France PW, Drexhage MG, Parker JM, Moore MW, Carter SF, *et al.* (1989) *Fluoride Glass Optical Fibres*. Glasgow: Blackie.
Lewis MH (1989) *Glass and Glass-Ceramics*. London: Chapman and Hall.
Ossi PM (2003) *Disordered Materials*. Berlin: Springer.
Paul A (1990) *Chemistry of Glass*. London: Chapman and Hall.
Pulker HK (1984) *Coatings on Glass*. Amsterdam: Elsevier.
Rawson H (1991) *Glasses and Their Applications*. London: Institute of Materials, Minerals and Mining.
Ropp RC (1992) *Inorganic Polymeric Glasses*. Amsterdam: Elsevier.
Scholze H (1988) *Glas: Natur, Struktur und Eigenschaften*. Berlin: Springer.
Uhlman DR and Kreidl NJ (1990) *Glass Science and Technology. Vol 4A: Structure, Microstructure and Properties*. Boston: Academic Press.
Varshneya AK (1994) *Fundamentals of Inorganic Glasses*. San Diego: Academic Press.
Zarzycki J (1991) *Glasses and the Vitreous State*. Cambridge: Cambridge University Press.

Group Theory

R Dirl and P Weinberger, Technische Universitat Wien, Vienna, Austria

Introduction

The theory of groups can be developed as an entity independent of its applications as it avoids confusion concerning what is proper to group theory itself and what is special to its applications. The mixture of theory and applications tends, in many cases, to obscure the boundary between them. From this viewpoint, perhaps the most important aspect of group-theoretical applications to physics or chemistry is to consistently disentangle the various concepts. The three major issues are the algebraic structure of the group in question, its corresponding representation theory usually referring to the field of complex numbers $\mathbb{C}$ (as commonly accepted in the quantum theory) and finally, what is essential to the actual physical application.

Properties of Groups

Although some basic knowledge of the group theory is assumed, basic properties are repeated here, thereby also establishing notation and conventions. Let $G = \{e, g, g', g'', \ldots\}$ be a group which is assumed to be finite (e.g., point groups), countable infinite (e.g., crystallographic space groups), or compact continuous (e.g., the three-dimensional proper rotation group $SO(3, \mathbb{R})$ whose corresponding universal covering group SU(2) plays a fundamental role for the spin degree of freedom). Recall that a group G is defined by a set of elements endowed with a certain algebraic structure. Its composition law

$$gg' = g'' \qquad [1]$$

must satisfy the associative law for ordered triplet $g, g', g'' \in G$ of group elements. Likewise, the existence of the unit element and uniquely defined inverses for each $g \in G$ must be guaranteed. Group–subgroup relations of the type $H \subset G$ imply that the group H forms a subset of G, which itself satisfies the group properties and hence is called a subgroup of the group G. A subgroup $N \subset G$ is called a normal subgroup if the conjugation

$$gNg^{-1} = N \qquad [2]$$

maps the subgroup N onto itself for all $g \in G$. Normal subgroups play a fundamental role not only when dealing with images of groups but also when studying features, such as faithfulness, of group representations.

The order of a group, which, loosely speaking, is equal to the number of its elements, is usually denoted by the symbol $|G|$. This symbol implies, for finite groups, the number of its elements, in the case of countable abelian groups, the group volume of their dual groups, and in the case of compact continuous groups, their group volume where the latter are endowed with appropriate Haar measures.

Specific relationships between different groups can be established by means of mappings of the type $\varphi : G \to H$, where G and H are two groups and $\varphi(G) = H$ defines the mapping in detail. Once the mappings satisfy the composition law of the image group H,

$$\varphi(g)\varphi(g') = \varphi(g'') \qquad [3]$$

then such mappings are called homomorphisms, and the associated kernel, denoted by $\ker_{\varphi(G)}$, defines a nontrivial normal subgroup of the pre-image group G. In mathematical terms,

$$\ker_{\varphi(G)} = \{g \in G \mid \varphi(g) = e \in H\} \lhd G \qquad [4]$$

$$G/\ker_{\varphi(G)} \sim H \qquad [5]$$

where the entry $G/\ker_{\varphi(G)}$ denotes the factor group and the special symbol $\lhd$ emphasizes that the subgroup $\ker_{\varphi(G)}$ must form a normal subgroup of the group G. A homomorphism is said to be an isomorphism if and only if the $\ker_{\varphi(G)}$ is trivial, that is, it consists of the identity element of G only. The mapping $\varphi : G \to H$ is called automorphism if and only if the group G is mapped onto itself. For more details on algebraic properties of groups and more advanced concepts, such as extensions of groups, the reader is referred to the "Further reading" section.

Properties of Representations

Every homomorphism $\varphi : G \to T(G)$ of a given group G into a group of nonsingular transformations $T(G) = \{T(g) \mid g \in G\}$ that map an underlying linear n-dimensional vector space $V^n_{\mathbb{K}}$ over the field $\mathbb{K}$ onto itself is called a representation of the group G if and only if the corresponding homomorphism property

$$T(g)T(g') = T(gg') = T(g'') \qquad [6]$$

is satisfied for all group elements $g, g' \in G$. Such a general representation is called a linear representation of the group G if and only if the linearity condition is satisfied. To be strict, one assumes that

$$T(g)(a\boldsymbol{v} + b\boldsymbol{w}) = aT(g)\boldsymbol{v} + bT(g)\boldsymbol{w} \qquad [7]$$

for all vectors $\boldsymbol{v}, \boldsymbol{w} \in V_{\mathbb{K}}^{n}$ and all $a, b \in \mathbb{K}$. In many applications, especially in quantum mechanics, one assumes $\mathbb{K} = \mathbb{C}$, sometimes one restricts $\mathbb{K} = \mathbb{R}$, whereas for some applications in crystallography, one even restricts $\mathbb{K} = \mathbb{Z}$, where the symbol $\mathbb{Z}$ denotes the ring of integers. For instance, one is inevitably led to the ring of integers when discussing the automorphisms of Bravais lattices. The restriction to the elements of the full rotation group $O(3, \mathbb{R}) = \mathrm{SO}(3, \mathbb{R}) \times \mathscr{C}_2$, where the symbol $\mathscr{C}_2 = \{E, I\}$ denotes that the inversion group does not influence the statement. It merely restricts the elements of $\mathrm{GL}(3, \mathbb{Z})$ to the elements of point groups $\mathscr{P}(T_{\text{Bravais}})$, which are finite subgroups of $O(3, \mathbb{R})$.

Matrix Representations

Every homomorphism $\vartheta : G \to D(G)$ of a given group G into a group of nonsingular matrices represented by $D(G) = \{D(g) \mid g \in G\}$, which are assumed to be finite-dimensional, is called a matrix representation of the group G if and only if the corresponding homomorphism property

$$D(g)D(g') = D(gg') = D(g'') \qquad [8]$$

is satisfied for all group elements $g, g' \in G$. Every matrix representation $\vartheta(G) = D(G)$ of the group G is called faithful if $\ker_{\varphi(G)}$ of the homomorphism is trivial. Every matrix representation of any group G is called unitary if and only if

$$D(g)^{-1} = D(g)^{\dagger} \qquad [9]$$

for all group elements $g \in G$. In the case of finite, countable, or compact continuous groups, it is assumed that every nonsingular matrix representation can be transformed by means of suitable intertwining matrices into unitary ones, if $\mathbb{K} = \mathbb{C}$ is assumed. However, this statement does not make sense if, for instance, $\mathbb{K} = \mathbb{Z}$, the ring of integers, is chosen. In this context, it should be noted that matrix groups are entities in their own right, which need not have any carrier space in order to be able to define them.

Unitary Operator Representations

A homomorphism $\varepsilon : G \to U(G)$ of a given group G into a group of unitary operators, represented by $U(G) = \{U(g) \mid g \in G\}$, that maps an underlying Hilbert space $\mathscr{H}$ over the field $\mathbb{C}$ onto itself is called a unitary operator representation of the group G if and only if the corresponding homomorphism property

$$U(g)U(g') = U(gg') = U(g'') \qquad [10]$$

$$\langle \phi, \psi \rangle_{\mathscr{H}} = \langle U(g)\phi, U(g)\psi \rangle_{\mathscr{H}} \qquad [11]$$

is satisfied for all group elements $g, g' \in G$, together with $U(g^{-1}) = U(g)^{\dagger}$ for all $g \in G$. Here the symbol $\langle ., . \rangle_{\mathscr{H}}$ denotes the scalar product of the $\mathscr{H}$. The most prominent application of this concept concerns the quantum mechanical problems, where the symmetry operations of any Hamiltonian H have to be realized as unitary operators in order to conform to the fundamental rules of quantum mechanics, such as the invariance of the expectation values of operators with respect to unitary conjugation operations.

Unitary Irreducible Matrix Representations

The so-called unitary irreducible matrix representations play the most important role in the vast majority of group-theoretical applications. Hereafter, these representations are called unirreps. Their basic properties are

$$D^{\xi}(g)D^{\xi}(g') = D^{\xi}(gg') \qquad [12]$$

$$D^{\xi}(g^{-1}) = D^{\xi}(g)^{\dagger} \qquad [13]$$

$$\frac{1}{|G|} \sum_{g \in G} D^{\xi}_{ab}(g) D^{\lambda}_{rs}(g)^{*} = \frac{1}{n(\xi)} \delta_{\xi\lambda} \delta_{ar} \delta_{bs} \qquad [14]$$

$$\sum_{\xi \in \mathscr{A}(G)} \sum_{m,n=1}^{n(\xi)} D^{\xi}_{mn}(g) D^{\xi}_{mn}(g')^{*} = |G| \delta_{gg} \qquad [15]$$

where the following notations have been introduced and further conventions have been adopted. Let $D^{\xi}(G) = \{D^{\xi}(g) \mid g \in G\}$ denote an $n(\xi)$-dimensional G unirrep. Moreover, [12] represents the composition law, [13] represents the unitarity condition, [14] illustrates the orthogonality relations, and [15] the completeness relations of G unirreps. Finally, the symbol $\mathscr{A}(G) = \{\xi, \lambda, \ldots\}$ denotes the set of G irrep labels, whilst the subscripts a, b, and likewise r, s, or eventually m, n label the rows and columns of the corresponding G unirreps.

Schur's lemma *Let G be a group and $D^{\alpha}(G)$ with $D^{\beta}(G)$ two given G unirreps. Moreover, let S be a matrix such that*

$$D^{\alpha}(g)S = SD^{\beta}(g) \qquad [16]$$

holds true for all group elements $g \in G$. Now, by applying the orthogonality relations [14] *of the G unirreps and assuming that equivalent G unirreps are chosen to be, by definition, always identical, one arrives at the following well-known result*:

$$S = \begin{cases} O_{\alpha,\beta} & \text{for } \alpha \nsim \beta \\ \delta 1_{\alpha,\alpha} & \text{for } \alpha = \beta \end{cases} \qquad [17]$$

where, in particular, the notation $\alpha \sim \beta$ signifies inequivalent G unirreps and alternatively $\alpha = \beta$ equivalent G unirreps, respectively.

Here the entry $O_{\alpha,\beta}$ denotes the rectangular $n(\alpha) \times n(\beta)$ null matrix, and the entity $1_{\alpha,\alpha}$ the square $n(\alpha) \times n(\alpha)$ unit matrix. Here, only the factor $\delta \in \mathbb{C}$ can be chosen arbitrarily. Schur's lemma, as a special case of the so-called Wigner–Eckart theorem, is one of the most frequently used group-theoretical tools in quantum mechanical applications.

Reducible Representations

Let $D(G) = \{D(g) \mid g \in G\}$ be an n-dimensional matrix representation of the group G. This implies that the composition law [8] is satisfied. Every unitary similarity transformation of the G matrix representation, $D(G)$, defined by some n-dimensional unitary matrix $W \in U(n)$ leads to an equivalent matrix representation of the group G:

$$F(g) = WD(g)W^{\dagger} \qquad [18]$$

$$F(g)F(g') = F(gg') \qquad [19]$$

What is common to these equivalent matrix representations $D(G) \sim F(G) \sim \cdots$ are the so-called characters of the corresponding equivalent matrix representations, which by construction are identical:

$$\chi(g) = \text{tr } D(g) = \text{tr } F(g) = \cdots \qquad [20]$$

Accordingly, every set of characters $\{\chi(g) \mid g \in G\}$ is an invariant and loosely speaking, the unique fingerprint of the G matrix representations. Sets of characters, say $\{\chi(g) \mid g \in G\}$ and $\{\tilde{\chi}(g) \mid g \in G\}$ of any given group G are either identical (and, hence, the corresponding matrix representations are equivalent), or different (and the corresponding matrix representations are inequivalent). Thus, this provides an important and simple tool to distinguish matrix representations of any given group G uniquely. Apart from this, one may write

$$\text{tr } D^{\xi}(g) = \chi^{\xi}(g) \qquad [21]$$

$$\frac{1}{|G|} \sum_{g \in G} \chi^{\xi}(g)\chi^{\lambda}(g)^{*} = \delta_{\xi\lambda} \qquad [22]$$

which represent the orthogonality relations of the so-called simple characters that are uniquely assigned to the corresponding G unirreps. These follow immediately from the orthogonality relations [14] by simply carrying out the trace operation for the G unirreps.

A G matrix representation $D(G)$ is called reducible if and only if this matrix representation can be decomposed by means of suitable unitary similarity matrices into a direct sum of small dimensional G matrix representations. Conversely, a G matrix representation $D(G)$ is called irreducible if this matrix representation cannot be decomposed into a direct sum of smaller-dimensional G matrix representations. Note, in passing, that due to the restriction to consider merely finite, countable, and compact continuous Lie groups, one need not distinguish between reducible and completely reducible G matrix representations, since reducibility implies complete reducibility provided that the field $\mathbb{K} = \mathbb{C}$ coincides with the field of complex numbers.

Irreducibility criterion Let $D(G)$ be an m-dimensional G matrix representation. Its corresponding set of characters $\{\chi(g) \mid g \in G\}$, where $\text{tr } D(g) = \chi(g)$ is assumed, has to satisfy

$$\frac{1}{|G|} \sum_{g \in G} |\chi(g)|^2 = \begin{cases} 1, & D(G) \text{ is irreducible} \\ M > 1, & D(G) \text{ is reducible} \end{cases} \qquad [23]$$

where the integer $M > 1$ is uniquely defined by the G matrix representation in question. This criterion is simply applicable and widely used in many practical applications of group-theoretical methods.

Decomposition of reducible matrix representations The decomposition of a reducible m-dimensional G matrix representation $D^{\text{red}}(G)$ into a direct sum of its irreducible constituents is a significant problem. The consistent sequence of the corresponding decomposition formulas may be summarized as follows:

$$D^{\text{red}}(g) = \sum_{\xi \in \mathscr{A}(G)} \oplus m(\text{red}|\xi) D^{\xi}(g) \qquad [24]$$

$$m(\text{red} \mid \xi) = \frac{1}{|G|} \sum_{g \in G} \chi^{\text{red}}(g)\chi^{\xi}(g)^{*} \qquad [25]$$

$$\frac{1}{|G|} \sum_{g \in G} |\chi(g)|^2 = \sum_{\xi \in \mathscr{A}(G)} m(\text{red}|\xi) n(\xi) = M \in \mathbb{Z}^{+} \qquad [26]$$

$$W^{\dagger} D^{\text{red}}(g) W = \sum_{\xi \in \mathscr{A}(G)} \oplus m(\text{red}|\xi) D^{\xi}(g) \qquad [27]$$

$$\{W\}_{j;\xi wa} = W_{j;\xi wa} = \{\boldsymbol{W}_{\boldsymbol{a}}^{\xi \boldsymbol{w}}\}_{\boldsymbol{j}} \qquad [28]$$

$$D^{\text{red}}(g) \boldsymbol{W}_{\boldsymbol{a}}^{\xi \boldsymbol{w}} = \sum_{\boldsymbol{b}=\boldsymbol{1}}^{\boldsymbol{n}(\xi)} \boldsymbol{D}_{\boldsymbol{ba}}^{\xi}(\boldsymbol{g}) \boldsymbol{W}_{\boldsymbol{b}}^{\xi \boldsymbol{w}} \qquad [29]$$

Several remarks should be made on this set of formulas. First, it should be noted that the unitary m-dimensional similarity matrix W is nonsymmetrically indexed. The row index of W is given by $j = 1, 2, \ldots, m$, whereas its column index is fixed by the ordered triplets (ξ, w, a), where the G irrep label $\xi \in \mathscr{A}(G)$ together with the so-called multiplicity index $w = 1, 2, \ldots, m(\text{red}|\xi)$, and finally the row index $a = 1, 2, \ldots, n(\xi)$, of the G unirrep $D^{\xi}(G)$ have to be taken into account. This nonsymmetrical indexing is common to all types of similarity matrices, for instance, subduction matrices, or Clebsch–Gordan matrices as a special case of subduction matrices. Second, it should be realized that the columns $\{\boldsymbol{W}_{a}^{\xi w}\}$ of any similarity matrix may be regarded as symmetry adapted vectors, since the transformation formulas [29] represent the prototype of symmetrized states. The basic features of symmetrized states are that they are mutually orthonormal and they transform according to G unirreps. The computation of similarity matrices can be done by applying the so-called projection method, which is briefly discussed later, since the transformation formulas [29] suggest this method. Finally, it should be noted that due to Schur's lemma the similarity matrices W cannot be unique and, hence, have to be fixed by adopting appropriate conventions. The structure of unitary similarity matrices describing this ambiguity can be simply deduced by the following statement.

Schur's lemma applied to reducible representations Let G be a group and $\Phi \in \mathrm{GL}(m, \mathbb{C})$ an m-dimensional nonsingular but otherwise arbitrary matrix. Moreover, let $D^{\mathrm{red}\oplus}(G)$, which symbolizes the RHS of [27], be an m-dimensional reducible unitary G matrix representation that is already decomposed into a direct sum of its irreducible constituents. Assume that

$$[\Phi, D^{\mathrm{red}\oplus}(g)] = O \qquad [30]$$

for all group elements $g \in G$. By applying Schur's Lemma to this general situation, one immediately arrives at the following result:

$$\Phi = \sum_{\xi \in \mathscr{A}(G)} \oplus (1_{\alpha,\alpha} \otimes \Phi^{\xi}(m(\text{red}|\xi))) \qquad [31]$$

where the submatrices $\Phi^{\xi}(m(\text{red}|\xi)) \in \mathrm{GL}(m\ (\text{red}|\xi), \mathbb{C})$ are otherwise arbitrary. Of course, if $\Phi \in U(m)$, then $\Phi^{\xi}(m(\text{red}|\xi)) \in U(m(\text{red}|\xi), \mathbb{C})$ must be satisfied. The latter condition reflects the ambiguity of similarity matrices, since W and its counterpart $W\Phi$ are equally well suited to describe identically the decomposition [27], where the phase matrices of the type [31] can be chosen arbitrarily.

The second type of ambiguities emerges from the nonuniqueness of each G unirrep whose dimension is greater than 1. Since every similarity transformation of the type

$$F^{\xi}(g) = W^{\xi\dagger} D^{\xi}(g) W^{\xi} \qquad [32]$$

by means of an arbitray unitary matrix $W^{\xi} \in U(n(\xi))$, yields equivalent G unirreps, they are equally well suited to be used in any application. One then arrives at the following formulas:

$$\bar{W}^{\dagger} D^{\mathrm{red}}(g) \bar{W} = \sum_{\xi \in \mathscr{A}(G)} \oplus m(\text{red} \mid \xi) F^{\xi}(g) \qquad [33]$$

$$\bar{W} = W\Psi \qquad [34]$$

$$\Psi = \sum_{\xi \in \mathscr{A}(G)} \oplus (1_{\alpha,\alpha} \otimes \Phi^{\xi}(m(\text{red}|\xi))) \qquad [35]$$

where, for the sake of brevity, $m(\text{red} \mid \xi) = m(\alpha)$ has been employed; for $m(\alpha) = 0$, the corresponding terms vanish. The combination of both types of ambiguities, namely the occurrence of phase matrices due to Schur's lemma and the transfer to equivalent G unirreps, have led to many unnecessary controversies in the literature. Accordingly, it is generally advised that when comparing results emerging from different sources, one should check whether the intertwining matrices W can be identified by means of phase matrices of the type [31]. If this fails, then the results of at least one source are wrong.

Subduced Representations

Let $H \subset G$ be a given group–subgroup relation, where, for the sake of simplicity, the index $s \in \mathbb{Z}^{+}$ of the subgroup G in the supergroup G is assumed to be finite. Note that in the case of group–subgroup relations between space groups and sectional layer or penetration rod groups, which are briefly discussed later, this simplifying constraint would be violated. To unify the discussion, let formally

$$G = \sum_{j=1}^{s} a_j H \qquad [36]$$

be the left coset decomposition of the group G with respect to the subgroup H, where, for the sake of distinction, coset representatives $a_j \in G$ are underlined. Cosets $a_j H$ are unique, whereas coset representatives $a_j \in G$ are never unique unless the subgroup is trivial.

By definition, the restriction of any given G matrix representation, say $D(G) = \{D(g) \mid g \in G\}$, to the elements of the subgroup H is called subduced matrix representation. Symbolically,

$$D(G)\downarrow H = \{D(h) \mid h \in H\} \qquad [37]$$

which means that only the elements of the subgroup $h \in H$ have to be taken into account. Of course, if $D(G) = D^{\xi}(G)$ forms a G unirrep, then the subduced matrix representation $D^{\xi}(G)\downarrow H$ will in general be reducible. Hence, if a complete set of H unirreps, say $\{L^{\lambda}(H) \mid \lambda \in \mathscr{A}(H)\}$, are known, the corresponding decomposition of the H matrix representations $D^{\xi}(G)\downarrow H$ can be carried out along the lines discussed previously:

$$D^{\xi}(g)\downarrow H \sim \sum_{\lambda \in \mathscr{A}(H)} \oplus m(\xi|\lambda) L^{\lambda}(h) \qquad [38]$$

$$m(\xi|\lambda) = \frac{1}{|H|} \sum_{h \in H} \chi^{\xi}(h) \chi^{\lambda}(h)^{*} \qquad [39]$$

$$W^{\xi\dagger} D^{\xi}(h) W^{\xi} = \sum_{\lambda \in \mathscr{A}(H)} \oplus m(\xi|\lambda) L^{\lambda}(h) \qquad [40]$$

The generalization of these decomposition formulas to the more general situation of reducible G matrix representations subduced to H matrix representations, say $D(G)\downarrow H$, can be readily deduced from formulas [38], [39], and [40].

Clebsch–Gordan Coefficients

The problem of computing for any group G with a fixed complete set of G unirreps, say $\{D^{\xi}(G) \mid \xi \in \mathscr{A}(G)\}$, corresponding Clebsch–Gordan matrices may be viewed as a special case of constructing, for subduced matrix representations, suitable similarity matrices that transform the subduced representation into a direct sum of its subgroup constituents.

General outer direct-product groups General outer direct-product groups are groups of the type $G_1 \otimes G_2$ where the given groups G_1 and G_2 are assumed not to be isomorphic. Its group elements are, by definition, the ordered pairs (g_1, g_2), where $g_1 \in G_1$ and $g_2 \in G_2$ are chosen independently. In mathematical terms,

$$G_1 \otimes G_2 = \{(g_1, g_2) \mid g_1 \in G_1;\ g_2 \in G_2\} \qquad [41]$$

$$(g_1, g_2) * (g_1', g_2') = (g_1 g_1', g_2 g_2') \qquad [42]$$

where its composition law, given by [42], is an immediate consequence of the definition of the outer direct-product set. Moreover, the order $|G_1 \otimes G_2|$ of the outer direct-product group $G_1 \otimes G_2$ is given by the product of the orders $|G_1|$ and $|G_2|$ of the respective groups G_1 and G_2.

Special outer direct-product groups Let G be a group and the specific construction

$$G \otimes G = \{(g, g') \mid g, g' \in G)\} = G^2 \qquad [43]$$

its corresponding outer direct-product group, whose group elements are again the ordered pairs (g, g'), where, analogously to the general situation, the group elements $g, g' \in G$ are to be chosen independently. Its composition law is again given by the rule

$$(g, g') * (\tilde{g}, \tilde{g}') = (g\tilde{g}, g'\tilde{g}') \qquad [44]$$

and shows, amongst others, that for the order of the outer direct-product group, $|G \otimes G| = |G|^2$ must be valid. Obviously, the subset

$$G \boxtimes G = \{(g, g) \mid g \in G\} = G^{[2]} \sim G \qquad [45]$$

must form a subgroup of the outer direct-product group $G \otimes G$ which is isomorphic to the original group G. It is common to denote the special group $G \boxtimes G$ as a twofold Kronecker product group. The generalization to n-fold outer direct-product groups G^n and n-fold Kronecker product groups $G^{[n]}$ is immediate.

Unirreps of special outer direct-product groups The G^2 unirreps of the special outer direct-product group $G \otimes G = G^2$ are simply obtained by the construction of ordered pairs of direct matrix products of G unirreps, say $D^{\xi}(G)$ and $D^{\eta}(G)$, where $\xi, \eta \in \mathscr{A}(G)$. Accordingly, one can summarize that

$$D^{\xi \otimes \eta}(g, g') = D^{\xi}(g) \otimes D^{\eta}(g') \qquad [46]$$

$$(\xi \otimes \eta) \in \mathscr{A}(G^2) \qquad [47]$$

$$\dim D^{\xi \otimes \eta}(g, g') = n(\xi) n(\eta) \qquad [48]$$

for all $(g, g') \in G^2$, where especially the elements $(\xi, \eta) = (\xi \otimes \eta)$ of the index set $\mathscr{A}(G^2)$ of the outer direct-product group G^2 are denoted by the symbols $(\xi \otimes \eta)$ to emphasize the outer direct-product group structure. Now, if one restricts G^2 to its Kronecker product subgroup $G^{[2]}$,

$$\begin{aligned} D^{\xi \otimes \eta}(G^2)\downarrow G^{[2]} &= D^{\xi,\eta}(G) \\ &= \{D^{\xi,\eta}(g) = D^{\xi}(g) \otimes D^{\eta}(g) \mid g \in G\} \end{aligned} \qquad [49]$$

then one arrives at $n(\xi)n(\eta)$-dimensional G matrix representations, which in general are reducible.

Clebsch–Gordan matrices By definition, the corresponding similarity matrices are nothing but the unitary Clebsch–Gordan matrices whose matrix elements are the Clebsch–Gordan coefficients. In

mathematical terms

$$D^{\xi,\eta}(g) \sim \sum_{\zeta \in \mathscr{A}(G)} \oplus m(\xi,\eta|\zeta) D^{\zeta}(g) \qquad [50]$$

$$m(\xi,\eta \mid \zeta) = \frac{1}{|G|} \sum_{g \in G} \chi^{\xi}(g) \chi^{\eta}(g) \chi^{\zeta}(g)^{*} \qquad [51]$$

$$C^{\xi,\eta\dagger} D^{\xi,\eta}(g) C^{\xi,\eta} = \sum_{\zeta \in \mathscr{A}(G)} \oplus m(\xi,\eta|\zeta) D^{\zeta}(g) \qquad [52]$$

$$\{C^{\xi,\eta}\}_{jk;\zeta wa} = C^{\xi,\eta}_{jk;\zeta wa} = \left(\begin{array}{cc|c} \xi & \eta & \zeta w \\ j & k & a \end{array}\right) \qquad [53]$$

where, as already pointed out, the corresponding Clebsch–Gordan coefficients $C^{\xi,\eta}_{jk;\zeta wa}$ are nonsymmetrically indexed. The row index of Clebsch–Gordan matrices $C^{\xi,\eta}$ are given by the ordered pairs (j,k), where $j = 1,2,\ldots,n(\xi)$ and $k = 1,2,\ldots,n(\eta)$, whereas their column indices are fixed by the ordered triplets (ζ,w,a), where the G irrep label $\zeta \in \mathscr{A}(G)$ together with the so-called multiplicity index $w = 1,2,\ldots,m(\xi,\eta|\zeta)$, and finally the row index $a = 1,2,\ldots,n(\zeta)$ of the relevant G unirreps $D^{\zeta}(G)$ have to be taken into account. The round bracket symbol occurring on the RHS of [53] is frequently used to symbolize Clebsch–Gordan coefficients but many other types of symbols may appear in applications.

Needless to repeat that Clebsch–Gordan matrices are not unique, which is not only due to Schur's lemma but also to their representation dependence. Apart from this, it is worth stressing the possible appearance of the so-called multiplicity index w in the case of twofold Kronecker product decompositions. In fact, the appearance of the multiplicity index w depends on the structure of the group G in question. A group G is called simple reducible if and only if the multiplicities $m(\xi,\eta|\zeta)$ occurring in the so-called Clebsch–Gordan series [52] are at most 1. Accordingly, a group G is called simple reducible if and only if

$$0 \leqslant m(\xi,\eta|\zeta) \leqslant 1 \qquad [54]$$

is satisfied for all G irrep labels $\xi,\eta,\zeta \in \mathscr{A}(G)$. Groups of this type are, for instance, SU(2) or its homomorphic image $SO(3,\mathbb{R})$, and certainly abelian groups, whereas in particular three-dimensional crystallographic space groups $\mathscr{G}$ are nonsimple reducible. This has as an important consequence that an additional ambiguity arises in the computation of Clebsch–Gordan coefficients and their potential applications.

Induced Representations

By definition, induced matrix representations are obtained by assuming a given group–subgroup relation, say $H \subset G$ with [36] as its left coset decomposition, and extending by means of the so-called induction procedure a given H matrix representation $D(H)$ to an induced G matrix representation $D^{\uparrow G}(G)$. To be more strict, induced G matrix representations are defined by the following expressions:

$$D^{\uparrow G}_{\underline{a},\underline{b}}(g) = \delta_{\underline{a}H,g\underline{b}H} D(\underline{a}^{-1} g \underline{b}) \qquad [55]$$

$$\delta_{\underline{a}H,g\underline{b}H} = \left\{ \begin{array}{ll} 1, & \underline{a}^{-1} g \underline{b} \in H \\ 0, & \text{otherwise} \end{array} \right\} \qquad [56]$$

$$\dim D^{\uparrow G}(G) = |G:H| \dim D(H) \qquad [57]$$

where, in part, a matrix representation has been employed and the notation $|G:H| = s$ has been introduced to stress the fact that the dimensions of the induced G matrix representations increase by the index of H in G. Moreover, if $D(H)$ defines a reducible H matrix representation, then the corresponding induced G matrix representation $D^{\uparrow G}(G)$ must be reducible in any case. Conversely, if $D(H) = D^{\lambda}(H)$ defines an H unirrep, then the corresponding induced G matrix representation $D^{\lambda\uparrow G}(G)$ still need not necessarily be a G unirrep. On the contrary, induced G matrix representations of the type $D^{\lambda\uparrow G}(G)$ are in general reducible and may be decomposed into direct sums of G unirreps in the same manner as before. The corresponding decomposition formulas read

$$D^{\lambda\uparrow G}(g) \sim \sum_{\zeta \in \mathscr{A}(G)} m(\lambda\uparrow G|\zeta) D^{\zeta}(g) \qquad [58]$$

$$m(\lambda\uparrow G|\zeta) = \frac{1}{|G|} \sum_{g \in G} \chi^{\lambda\uparrow G}(g) \chi^{\zeta}(g)^{*} \qquad [59]$$

$$m(\lambda\uparrow G|\zeta) = m(\zeta\downarrow H|\lambda) \qquad [60]$$

$$S^{\lambda\uparrow G\dagger} D^{\lambda\uparrow G}(g) S^{\lambda\uparrow G} = \sum_{\zeta \in \mathscr{A}(G)} m(\lambda\uparrow G|\zeta) D^{\zeta}(g) \qquad [61]$$

$$\{S^{\lambda\uparrow G}\}_{\underline{a}k;\zeta wa} = S^{\lambda\uparrow G}_{\underline{a}k;\zeta wa} \qquad [62]$$

where the matrix elements $S^{\lambda\uparrow G}_{\underline{a}k;\zeta wa}$ of the subduction matrices $S^{\lambda\uparrow G}$ are nonsymmetrically indexed. The row indices of $S^{\lambda\uparrow G}$ are given by the ordered pairs $(\underline{a},\ k)$ where $\underline{a} \in G:H$ and $k = 1,2,\ldots,n(\lambda)$, whereas their column indices are fixed by the ordered triplets (ζ,w,a), where the G irrep label $\zeta \in \mathscr{A}(G)$ together with the so-called multiplicity index

$w = 1, 2, \ldots, m(\lambda \uparrow G|\zeta)$, and finally the row index $a = 1, 2, \ldots, n(\zeta)$ of the relevant G unirreps $D^{\zeta}(G)$ have to be taken into account. Finally, the relations [60] describe the Frobenius reciprocity law, which, for instance, is extensively discussed in several textbooks.

Group-Theoretical Methods in Physics

Up to now, exclusively group-theoretical aspects, such as their algebraic structure and associated properties, and representation-theoretic aspects with special emphasis on the decomposition of reducible representations into direct sums of their irreducible constituents have been dealt with. This section discusses how group-theoretical methods can be applied in physical applications. The physical applications cover a wide range of, not only quantum mechanics but also classical applications (e.g., classical mechanics or hydrodynamics are possible candidates for these methods). Here the discussions are confined to standard applications in quantum mechanics.

Symmetry Groups of Hamiltonians

The first task is to construct a unitary operator representation of a given group, say G, with respect to the Hilbert space $\mathscr{H}$, which is the carrier space for the considered Hamiltonian H in question. Accordingly, it is assumed that there exits a homomorphism $\varepsilon : G \to U(G)$ of the given group G into a group of unitary operators which are represented by $U(G) = \{U(g) \,|\, g \in G\}$ that map the underlying Hilbert space $\mathscr{H}$ over the field $\mathbb{C}$ onto itself.

The group G being represented by the corresponding unitary operator group $U(G) = \{U(g) \,|\, g \in G\}$ is called a symmetry group of the Hamiltonian H if and only if the Hamiltonian commutes with all unitary operators:

$$[H, U(g)] = O \quad \forall g \in G \tag{63}$$

Moreover, assume that the eigenvalue problem for the Hamiltonian H has been solved, which implies that not only the corresponding eigenvalues, denoted by $\lambda \in \mathbb{R}$, but also suitably orthonormalized eigenstates $\{\Phi_j^{\lambda} \,|\, j = 1, 2, \ldots, \deg \lambda\}$ for each eigenvalue are known. The existence of a symmetry group G of the Hamiltonian H implies

$$\begin{aligned} & H\Phi_j^{\lambda} = \lambda \Phi_j^{\lambda} \\ \Leftrightarrow\, & H(U(g)\Phi_j^{\lambda}) = \lambda(U(g)\Phi_j^{\lambda}) \end{aligned} \tag{64}$$

that the corresponding eigenspaces $\mathscr{H}^{\lambda} \subset \mathscr{H}$, which are uniquely associated with the eigenvalues $\lambda \in \mathbb{R}$ the Hamiltonian H, have to be G-invariant subspaces of the original Hilbert space $\mathscr{H}$. Clearly, the "bigger" and the "more noncommutative" the symmetry group G, the more useful the symmetry group G. Once the so-called symmetry adapted states, which are discussed below, are known, the application of Schur's lemma leads at least to a partial solution of the eigenvalue problem. Obviously, if every G-invariant subspace $\mathscr{H}^{\lambda}$ turns out to be irreducible, then the eigenvalue problem for H is completely solved, and the degeneracy of the eigenvalues is sometimes said to be generic with respect to the group G. Conversely, if some of the subspaces $\mathscr{H}^{\lambda}$ are reducible, then the degeneracy of the corresponding eigenvalues is sometimes said to be nongeneric with respect to the group G.

Symmetrization of States

One of the most popular group-theoretical applications in quantum mechanics consists of symmetry, adapting given sets of states, say $\{\Phi\} = \{\Phi_1, \Phi_2, \ldots, \Phi_n\}$. In mathematical terms, it implies that, out of the given set of states $\{\Phi\}$, some new sets of states be constructed systematically whose elements transform according to given G unirreps and are mutually orthonormal. Such bases are usually called G-symmetrized states or simply symmetry-adapted states. The standard method to achieve the symmetry adaptation of states is the so-called projection method. The set of operators

$$\{E_{jk}^{\xi}\} = \{E_{jk}^{\xi} \,|\, \xi \in \mathscr{A}(G);\ j, k = 1, 2, \ldots, n(\xi)\} \tag{65}$$

$$E_{jk}^{\xi} = \frac{n(\xi)}{|G|} \sum_{g \in G} D_{jk}^{\xi}(g)^* U(g) \tag{66}$$

are the so-called units of the corresponding group algebra $\mathscr{A}(G)$ whose general elements are arbitrary linear combinations of the group elements. Clearly, if and only if G is finite, then $|\{E_{jk}^{\xi}\}| = |G|$, which, loosely speaking, remains valid if the group G is countable or even compact continuous but has to be refined correspondingly, since the group volumes of their dual groups have to be taken into account. In fact, this represents a topological subtlety. Apart from this, one has

$$\{E_{jk}^{\xi}\}^{\dagger} = E_{kj}^{\xi} \tag{67}$$

$$E_{jk}^{\xi} E_{mn}^{\eta} = \delta_{\xi\eta} \delta_{km} E_{jn}^{\xi} \tag{68}$$

$$U(g) E_{jk}^{\xi} = \sum_{\ell=1}^{n(\xi)} D_{\ell j}^{\xi}(g) E_{\ell k}^{\xi} \tag{69}$$

$$\sum_{\xi \in \mathscr{A}(G)} \sum_{k=1}^{n(\xi)} E_{kk}^{\xi} = 1_{\mathscr{H}} \tag{70}$$

where the extra symbol $1_{\mathscr{H}}$ denotes the unit operator of the underlying Hilbert space $\mathscr{H}$. It follows from [67] and [68] that the operators $\{E^{\xi}_{kk} \mid \xi \in \mathscr{A}(G);\ k = 1, 2, \ldots, n(\xi)\}$ are projection operators, whereas the remaining operators are shift operators that intertwine mutually orthogonal subspaces that belong to any given $\xi \in \mathscr{A}(G)$ but different $k = 1, 2, \ldots, n(\xi)$.

Step 1 First, one proves that the n-dimensional subspace which is the linear hull of the set of functions $\{\Phi_1, \Phi_2, \ldots, \Phi_n\}$ denoted by $\mathscr{H}_n$ is in fact a G-invariant subspace. For simplicity, assume that the set of functions $\{\Phi\}$ forms an orthonormal basis. Accordingly, $\langle \Phi_j, \Phi_k \rangle_{\mathscr{H}} = \delta_{jk}$ for all pairs Φ_j, Φ_k of basis functions. By virtue of the transformation law given by the expressions

$$U(g)\Phi_j = \sum_{k=1}^{n} D^{\Phi}_{kj}(g)\Phi_k \qquad [71]$$

$$D^{\Phi}(g) \sim \sum_{\xi \in \mathscr{A}(G)} m(D^{\Phi}(G)|\xi) D^{\xi}(g) \qquad [72]$$

$$\begin{aligned} \dim D^{\Phi}(G) &= n \\ &= \sum_{\xi \in \mathscr{A}(G)} m(D^{\Phi}(G)|\xi) n(\xi) \end{aligned} \qquad [73]$$

one defines an n-dimensional G matrix representation, which in general forms a reducible representation. Relation [73] presents the obvious dimension check to be carried out from the outset.

Step 2 Let $m(D^{\Phi}(G)|\xi) > 0$, then one knows that the corresponding subspaces constructed by means of the projection operators

$$\mathscr{H}^{\xi,k}_n = E^{\xi}_{kk}\mathscr{H}_n \qquad [74]$$

$$\dim \mathscr{H}^{\xi,k}_n = m(D^{\Phi}(G) \mid \xi) \qquad [75]$$

have the same dimension for each $k = 1, 2, \ldots, n(\xi)$, as otherwise the projection method would contain inconsistencies.

Step 3 Next, one constructs for the fixed G irrep label $\xi \in \mathscr{A}(G)$ (where $m(D^{\Phi}(G)|\xi) > 0$ is assumed) and for a fixed row index, say $k = k_0 = 1$, an orthonormalized basis of the $m(D^{\Phi}(G)|\xi)$-dimensional subspace $\mathscr{H}^{\xi,k_0}_n$ by applying the Gram–Schmidt orthonormalization procedure. Symbolically, as a result of the projection method,

$$E^{\xi}_{k_0k_0}\{\Phi\} \to \{\Psi^{\xi,w}_{k_0} \mid w = 1, 2, \ldots, m(D^{\Phi}(G)|\xi)\} \qquad [76]$$

$$\langle \Psi^{\xi,w}_{k_0}, \Psi^{\xi,v}_{k_0} \rangle_{\mathscr{H}} = \delta_{wv} \qquad [77]$$

Step 4 In order to obtain for the fixed G irrep label $\xi \in \mathscr{A}(G)$ (where $m(D^{\Phi}(G)|\xi) > 0$ is assumed) the remaining partner functions, one has to employ the corresponding shift operators:

$$\Psi^{\xi,w}_j = E^{\xi}_{jk_0}\Psi^{\xi,w}_{k_0}, \quad j = 1, 2, \ldots, n(\xi) \qquad [78]$$

$$\langle \Psi^{\xi,w}_j, \Psi^{\xi,v}_{\ell} \rangle_{\mathscr{H}} = \delta_{wv}\delta_{j\ell} \qquad [79]$$

$$U(g)\Psi^{\xi,w}_j = \sum_{\ell=1}^{n(\xi)} D^{\xi}_{\ell j}(g)\Psi^{\xi,w}_{\ell} \qquad [80]$$

Step 5 Finally, in order to obtain a complete basis of the n-dimensional G-invariant subspace $\mathscr{H}_n$ that consists of G-symmetrized states which are mutually orthonormal and which satisfy the transformation law [80], one has to repeat the steps described previously for all G irrep labels $\xi \in \mathscr{A}(G)$, where the corresponding multiplicities $m(D^{\Phi}(G)|\xi)$ are different from zero.

Coupling of Product States

Another important application of group-theoretical methods in quantum mechanics concerns the coupling of states, where the constituents are assumed to transform according to G unirreps, and where the coupled product states likewise should possess specific transformation properties. Let $\Psi^{\xi}_j \in \mathscr{H}_1$ and, equivalently, $\Phi^{\eta}_k \in \mathscr{H}_2$ be two sets of G-symmetrized states, where states of the type

$$\Lambda^{\xi\otimes\eta}_{jk} = \Psi^{\xi}_j \otimes \Phi^{\eta}_k \in \mathscr{H}_1 \otimes \mathscr{H}_2 \qquad [81]$$

should be transformed into such linear combinations, so that the new states transform according to G unirreps of the twofold Kronecker product group $G^{[2]}$ which is isomorphic to G. Starting from the transformation law

$$\begin{aligned} &U(g, g')\Lambda^{\xi\otimes\eta}_{jk} \\ &\quad = \sum_{j'=1}^{n(\xi)} \sum_{k'=1}^{n(\eta)} \{D^{\xi\otimes\eta}(g)\}_{j'k',jk}\Lambda^{\xi\otimes\eta}_{j'k'} \end{aligned} \qquad [82]$$

which refers to the special outer direct-product group $G \otimes G$, one readily infers that the product states [81] transform according to G^2 unirreps. The aim is to find such linear combinations of the states [81] such that they transform according to G unirreps. The desired transformation coefficients are just the

corresponding Clebsch–Gordan coefficients:

$$\Theta_a^{\xi\otimes\eta;\zeta,w} = \sum_{j=1}^{n(\xi)} \sum_{k=1}^{n(\eta)} \left(\begin{matrix} \xi & \eta \\ j & k \end{matrix} \middle| \begin{matrix} \zeta w \\ a \end{matrix} \right) \Lambda_{jk}^{\xi\otimes\eta} \quad [83]$$

$$U(g,g)\Theta_a^{\xi\otimes\eta;\zeta,w} = \sum_{b=1}^{n(\zeta)} D_{ba}^{\zeta}(g)\Theta_b^{\xi\otimes\eta;\zeta,w} \quad [84]$$

This type of coupling of product states is frequently used in describing atomic spectra, where the orbital angular momenta are coupled and independently, the spin degree of freedom. The corresponding scheme of coupled states is well known as LS coupling.

Selection Rules – Wigner – Eckart Theorem

The most important, and certainly the most frequently used, application of group-theoretical methods in quantum mechanics concerns the well-known Wigner–Eckart theorem. In order to formulate the Wigner–Eckart theorem, one has to define the so-called irreducible tensor operators with respect to the group G in question and a complete set of G-symmetrized states. The latter are formally abbreviated as $\{\Psi_\ell^{\xi,n}\}$, where the extra index $n \in \mathbb{N}$ subsumes the additional quantum numbers that are not correlated to the G irrep labels.

Irreducible tensor operators A set of operators denoted by the symbol $T^\beta = \{T_k^\beta \mid \beta \in \mathscr{A}(G);\ k = 1, 2, \ldots, n(\beta)\}$, which consists of $n(\beta)$ tensor components T_k^β, is called irreducible tensor operator of rank β with respect to the group G if and only if the following transformation law is satisfied for all group elements $g \in G$:

$$U(g)T_k^\beta U(g)^\dagger = \sum_{j=1}^{n(\beta)} D_{jk}^\beta(g)T_j^\beta \quad [85]$$

Note that an arbitrary operator cannot be identified with a component of an irreducible tensor operator with respect to the group G in question. The decomposition of an arbitrary operator into its irreducible tensor components with respect to the group G can be done along the lines of the symmetrization procedure discussed previously for arbitrary states of the Hilbert space $\mathscr{H}$.

Wigner–Eckart theorem Let G be a group, $\{D^\xi(G)\}$ a complete set of G unirreps, $\{\Psi_\ell^{\xi,n}\}$ a complete set of G–symmetrized states of the underlying Hilbert space $\mathscr{H}$, and finally T^β a given irreducible tensor operator of rank $\beta \in \mathscr{A}(G)$ with respect to the given group G. The crucial point of the Wigner–Eckart theorem consists of the fact that the matrix elements of T_k^β with respect to G-symmetrized states can be factorized in a specific manner, namely

$$\langle \Psi_\ell^{\xi,n}, T_k^\beta \Psi_j^{\zeta,m} \rangle_{\mathscr{H}} = \sum_{w=1}^{m(\beta,\zeta|\xi)} \left(\begin{matrix} \beta & \zeta \\ k & j \end{matrix} \middle| \begin{matrix} \xi w \\ \ell \end{matrix} \right)^* \langle \Psi^{\xi,n} || T^\beta || \Psi^{\zeta,m} \rangle_w \quad [86]$$

where the second factors $\langle \Psi^{\xi,n} || T^\beta || \Psi^{\zeta,m} \rangle_w$ occurring on the RHS of [86] are called reduced matrix elements of the G irreducible tensor operator. The reduced matrix elements $\langle \Psi^{\xi,n} || T^\beta || \Psi^{\zeta,m} \rangle_w$ by definition must not depend on the row indices k, j, ℓ of the corresponding G unirreps. Moreover, the reduced matrix elements $\langle \Psi^{\xi,n} || T^\beta || \Psi^{\zeta,m} \rangle_w$ contain the "physics," whereas the Clebsch–Gordan coefficients are pure group-theoretical entities which have nothing to do with the physics of the problem considered, and hence are sometimes called geometrical factors. The proof of the Wigner–Eckart theorem relies, among others, upon the use of the orthogonality relations [14] of the G unirreps and of the Clebsch–Gordan series [52]. In the case of simple reducible groups (e.g., SU(2)), the corresponding Wigner–Eckart theorem reads

$$\langle \Psi_m^{j,n}, T_M^J \Psi_{m'}^{j',n'} \rangle_{\mathscr{H}} = \left(\begin{matrix} J & j' \\ M & m' \end{matrix} \middle| \begin{matrix} j \\ m \end{matrix} \right) \langle \Psi^{j,n} || T^J || \Psi^{j',n'} \rangle \quad [87]$$

where, in addition, the fact has been exploited that the corresponding SU(2) Clebsch–Gordan coefficients are real. The simplified version of the Wigner–Eckart theorem for simple reducible groups is likewise utilized to simplify the computation of the reduced matrix elements of irreducible tensor operators by applying the method of the so-called equivalent operators in the crystal field theory. The specifically simple form of the Wigner–Eckart theorem for the group SU(2) might not reflect the general complexity of the method; in general for nonsimple reducible groups, such as SU(3), or especially the three-dimensional crystallographic space groups, nontrivial multiplicities $m(\beta, \zeta|\xi) > 1$ occur and hence complicate the problem of the factorization of the considered matrix elements into Clebsch–Gordan coefficients and reduced matrix elements that explicitly depend on the multiplicity index w.

Selection rules Apart from this, the merits of Wigner–Eckart's theorem primarily rely upon the property that if the corresponding Clebsch–Gordan coefficients vanish, which exclusively emerges from the structure of the group G and the chosen G unirreps, the considered matrix elements also vanish. Thus, this conclusion follows exclusively from the Clebsch–Gordan coefficients and not from the actual form of the G-irreducible tensor operator; hence, in

all such cases, one need not compute the corresponding reduced matrix elements. Not surprisingly, this is an important statement when investigating selection rules, such as the Stark effect for the spectral properties of the hydrogen atom, or when dealing with electron–phonon interactions in solids. However, if the respective Clebsch–Gordan coefficients are nonzero, then one has to compute the corresponding reduced matrix elements, but which might be much more complicated than to simply compute directly the original matrix elements.

See also: Group Theory in Materials Science, Applications; Magnetic Point Groups and Space Groups.

PACS: 02.20.; 03.65.; 61.50

Further Reading

Altmann SL (1977) *Induced Representations in Crystals and Molecules: Point, Space, and Non-rigid Molecule Groups*. London: Academic Press.

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Bradley CJ and Cracknell AP (1972) *The Mathematical Theory of Symmetry in Solids*. Oxford: Clarendon.

Condon EU and Shortley GH (1935) *The Theory of Atomic Spectra*. Cambridge: Cambridge University Press.

Jansen L and Boon M (1976) *The Theory of Finite Groups. Applications to Physics*. Amsterdam: North-Holland.

Wigner EP (1959) *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. New York: Academic.

Group Theory in Materials Science, Applications

R Dirl and P Weinberger, Vienna University of Technology, Vienna, Austria

Applications in Materials Science

The potential applications of group-theoretical methods in solid-state physics – synonymous for materials science – depend on the various groups that are specific for this field. For instance, the compact continuous groups SU(2) or its homomorphic image $SO(3,\mathbb{R})$ are widely used in multiple scattering theory. Proper point groups as specific finite subgroups of $SO(3,\mathbb{R})$ are employed when investigating macroscopic properties of materials that are described by tensors of certain ranks. Crystallographic space groups containing Bravais vector lattices as countable subgroups are taken into account when describing bulk properties of materials. Moreover, sectional layer groups might turn out to be of decisive importance when discussing surface properties of materials, since they describe the symmetries of crystallographic planes. Finally, penetration rod groups may eventually become important when investigating one-dimensional defects in some materials.

Point Groups – Macroscopic Properties

Proper crystallographic point groups $\mathscr{P}(\mathscr{T}(\mathscr{L}))$ are, by definition, intersection groups of the type $SO(3,\mathbb{R}) \cap SL(3,\mathbb{Z})$ where the symbol $SL(3,\mathbb{Z})$ denotes the group of all integral 3×3 matrices with $\det \mathbb{Z}(z) = +1$ such that, in addition, a certain Bravais lattice $\mathscr{T}(\mathscr{L})$ is mapped by these matrices onto itself. Proper noncrystallographic point groups $\mathscr{P}$ are, by definition, finite subgroups of the proper rotation group $SO(3,\mathbb{R})$ which are not constrained by the invariance condition of Bravais vector lattices. Macroscopic properties of materials, such as stress or conductivity, are described by tensors of certain ranks which in the presence of symmetry are invariant with respect to the symmetry group G in question. For instance, the invariance condition implies, for a second rank tensor $\{T_{jk} \mid j,k = 1,2,3\}$,

$$\sum_{j,k=1}^{3} T_{jk} R_{jm}(R)\, R_{kn}(R) = T_{mn} \quad \forall R \in \mathscr{P} \qquad [1]$$

that, depending on the symmetry group $\mathscr{P}$, certain tensor components vanish for symmetry reasons. For instance, in cubic symmetry tensors of rank two are proportional to the three-dimensional unit matrix.

Crystallographic Space Groups

A proper description of bulk properties in solids requires the introduction of three-dimensional crystallographic space groups which, loosely speaking, can be seen as the invariance groups of three-dimensional periodic structures. The basic constituents of space groups are their underlying Bravais vector lattices $\mathscr{T}(\mathscr{L})$ and admissible crystallographic point groups $\mathscr{P}$ that leave the Bravais vector lattice $\mathscr{T}(\mathscr{L})$ invariant. Possible structures of space groups and their classification schemes are discussed later.

Bravais vector lattices Let $\{\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3\}$ be a set of three linearly independent (noncoplanar) vectors of the Euclidean vector space $\boldsymbol{E}^3_{\mathbb{R}}$ where the latter is

regarded as the vector space component of the underlying Euclidean point space $\mathsf{E}^3_{\mathbb{R}}$. To simplify the notation, let $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3)$ be formally denoted as row vector and $\underline{n} \in \mathbb{Z}^3$ as column vectors with integral coefficients. Then, by definition,

$$\mathscr{T}(\mathscr{L}) = \{\boldsymbol{t} = \boldsymbol{t} \cdot \underline{n} \mid \underline{n} \in \mathbb{Z}^3\} \qquad [2]$$

$$\boldsymbol{t} = \boldsymbol{e}\, \mathbb{D}\,(\mathscr{L}) \qquad [3]$$

where the countable set of vectors denoted by $\mathscr{T}(\mathscr{L})$ determines uniquely a Bravais vector lattice which can be traced back by means of an appropriate three-dimensional nonsingular matrix $\mathbb{D}\,(\mathscr{L}) \in \mathrm{SL}(3, \mathbb{R})$ with $\det \mathbb{D}\,(\mathscr{L}) > 0$ to an orthonormalized basis $\boldsymbol{e}$ of the associated Euclidean vector space $\boldsymbol{E}^3_{\mathbb{R}}$. It is worth noting that the positive sign of the determinant guarantees the right-handedness of the associated coordinate system. Apart from this, the symbol $\boldsymbol{t} \cdot \underline{n}$ should be seen as formal scalar product and likewise $\boldsymbol{e}\, \mathbb{D}\,(\mathscr{L})$ are the formal multiplication of a row vector times a 3×3 matrix. Moreover, one can define

$$\boldsymbol{t}(z) = \boldsymbol{t}\mathbb{Z}(z) \quad \forall \mathbb{Z}(z) \in \mathrm{SL}(3, \mathbb{Z}) \qquad [4]$$

countable infinitely many vector bases for one and the same Bravais vector lattice $\mathscr{T}(\mathscr{L})$, since the group $\mathrm{SL}(3, \mathbb{Z})$ is countably infinite. Accordingly, the corresponding vector bases $\boldsymbol{t}(z)$ are equally well suited to describe one and the same Bravais vector lattice. Hence, the choice of standardized vector bases, such as vectors of the shortest lengths, is subject to certain conventions to achieve this goal.

Automorphisms of vector lattices – point groups To understand the structure of crystallographic space groups, one first has to understand the structure of crystallographic point groups $\mathscr{P}(\mathscr{T})$ that leave a certain Bravais vector lattice $\mathscr{T}(\mathscr{L})$ invariant. Obviously,

$$\mathrm{Aut}(\mathscr{T}) = \mathrm{GL}(3, \mathbb{Z}) \qquad [5]$$

which implies that the "automorphism group" $\mathrm{Aut}(\mathscr{T})$ is the most general group that leaves the Bravais vector lattice $\mathscr{T}$ invariant, where $\mathrm{GL}(3, \mathbb{Z})$ contains $\mathrm{SL}(3, \mathbb{Z})$ as normal subgroup of index two. To assure that the point group $\mathscr{P}(\mathscr{T})$ contains only orthogonal transformations, one demands the restricting condition

$$\mathscr{P}(\mathscr{T}) = \mathrm{O}(3, \mathbb{R}) \cap \mathrm{GL}(3, \mathbb{Z}) \qquad [6]$$

in order to guarantee a maximal set of orthogonal symmetry operations that leave the Bravais vector lattice $\mathscr{T}$ invariant. The action of the point group operations $R \in \mathscr{P}(\mathscr{T})$ onto the elements of the primitive vector basis $\boldsymbol{t}$ is given by

$$R\boldsymbol{t}_j = \sum_{k=1}^{3} Z_{kj}(R)\boldsymbol{t}_k \qquad [7]$$

where the additional condition $Z(R) \in \mathrm{GL}(3, \mathbb{Z})$ must be satisfied for all group elements $R \in \mathscr{P}(\mathscr{T})$ of the crystallographic point group. The actual form of the matrix representation $\{Z(R)\}$ is characteristic of the corresponding crystal class, once the associated crystal system has been determined for the Bravais vector lattice $\mathscr{T}$ in question.

Space groups – gross classification The easiest access to crystallographic space groups is to start from the locally compact continuous "Euclidean group" $E(3, \mathbb{R})$ which is composed of the locally compacts three-dimensional continuous translation group $T(3, \mathbb{R})$ and the full rotation group $O(3, \mathbb{R})$:

$$E(3, \mathbb{R}) = T(3, \mathbb{R}) \circledS O(3, \mathbb{R}) \qquad [8]$$

which symbolizes a semidirect product group between the translation group and the full rotation group. The Euclidean group $E(3, \mathbb{R})$ is the most general motion group whose corresponding transformations map the Euclidean vector space $\boldsymbol{E}^3_{\mathbb{R}}$ onto itself, such that not only the distance between two vectors, but also the angle between them remains invariant. The nonsingular transformations $\mathscr{M}(S \mid \boldsymbol{v})$ that are uniquely assigned to the groups elements $(S \mid \boldsymbol{v}), (X \mid \boldsymbol{w}) \in E(3, \mathbb{R})$ are given by the following expressions:

$$\mathscr{M}(S \mid \boldsymbol{v})\boldsymbol{x} = S\boldsymbol{x} + \boldsymbol{v} = \boldsymbol{x}' \in \boldsymbol{E}^3_{\mathbb{R}} \qquad [9]$$

$$\mathscr{M}(S \mid \boldsymbol{v})\mathscr{M}(X \mid \boldsymbol{w}) = \mathscr{M}(SX \mid \boldsymbol{v} + S\boldsymbol{w}) \qquad [10]$$

where the so-called Wigner–Seitz symbols are used to denote the group elements of the Euclidean group. In order to combine consistently the various notations, one may express an arbitrary element of the Euclidean vector space $\boldsymbol{E}^3_{\mathbb{R}}$ as follows $\boldsymbol{x} = \boldsymbol{e} \cdot \underline{x}$, where $\underline{x} \in \mathbb{R}^3$ are the corresponding column vectors with real-valued components. Finally, the composition law of the Euclidean group $E(3, \mathbb{R})$ is given by [10] in terms of the nonsingular transformations.

By definition, space groups (hereafter sometimes denoted by the shorthand notation $\mathscr{G}$) are countable subgroups of the Euclidean group $E(3, \mathbb{R})$ where the locally compact continuous translation group $T(3, \mathbb{R})$ is reduced to one of the noncountable infinitely many Bravais vector lattices $\mathscr{T}$ and where the full rotation group $O(3, \mathbb{R})$ is reduced to the utmost maximal crystallographic point group $\mathscr{P}(\mathscr{T})$ that leaves the underlying Bravais vector lattice $\mathscr{T}$ by its orthogonal transformations invariant. In mathematical

terms, one writes

$$\mathcal{G} = \{\mathcal{T}, \mathcal{P} \mid \mathcal{O}, \boldsymbol{w}_{\mathcal{P}}\} \qquad [11]$$

to symbolize space groups. Here, the entry $\mathcal{T}$ denotes the Bravais vector lattice and $\mathcal{P}$ its corresponding crystallographic point group which satisfies the group–subgroup relation $\mathcal{P}_{\min} \subseteq \mathcal{P} \subseteq \mathcal{P}_{\max}$, where $\mathcal{P}_{\min}$ and its corresponding counterpart $\mathcal{P}_{\max}$ are determined by the associated arithmetic class of space group types. Moreover, the symbol $\mathcal{O}$ denotes the chosen origin of the associated coordinate system and $\boldsymbol{w}_{\mathcal{P}} \in \boldsymbol{E}^3_{\mathbb{R}}$, the set of nonprimitive translations which are likewise called "fractional translations." They are uniquely assigned to the point group elements $R \in \mathcal{P}$. Accordingly,

$$\boldsymbol{w}_{\mathcal{P}} = \{\boldsymbol{w}_R = \boldsymbol{t} \cdot \underline{w}_R \mid R \in \mathcal{P}\} \qquad [12]$$

where the components $\underline{w}_R \in \mathbb{Q}^3$ of the fractional translations become rational numbers if the origin $\mathcal{O}$ of the space group G is suitably chosen. The corresponding space group elements are denoted by $(R \mid \boldsymbol{w}_R + \boldsymbol{t}) \in \mathcal{G}$, where again the Wigner–Seitz notation is employed. The composition law of each space group $\mathcal{G}$ must follow automatically from the composition law [10] the Euclidean group $E(3, \mathbb{R})$ due to their group–subgroup relation. To be more specific

$$\begin{aligned}(R \mid \boldsymbol{w}_R + \boldsymbol{t}) * (S \mid \boldsymbol{w}_S + \boldsymbol{v}) \\ = (RS \mid \boldsymbol{w}_{RS} + \boldsymbol{t}_{R,S} + \boldsymbol{t} + R\boldsymbol{v})\end{aligned} \qquad [13]$$

$$\boldsymbol{t}_{R,S} = \boldsymbol{w}_R + R\boldsymbol{w}_S - \boldsymbol{w}_{RS} \qquad [14]$$

where the closure relation for the space group $\mathcal{G}$ implies $\boldsymbol{t}_{R,S} \in \mathcal{T}$ for all point group elements $R, S \in \mathcal{P}$. This allows one to classify the space groups into two significantly different types, namely "symmorphic" and "nonsymmorphic" space group types. The criterion reads

$$\text{Symmorphic } \mathcal{G}: \quad \boldsymbol{t}_{R,S} = 0 \quad \forall R, S \in \mathcal{P} \qquad [15]$$

$$\text{Nonsymmorphic } \mathcal{G}: \quad \boldsymbol{t}_{R,S} \neq 0 \quad \exists R, S \in \mathcal{P} \qquad [16]$$

which simply shows that symmorphic space groups can be seen as semidirect product groups, whereas nonsymmorphic space groups are more general extensions. As a matter of fact, there exist 230 space group types which split into 73 mutually disjoint arithmetic classes, where each arithmetic class is led by a symmorphic space group type.

$\mathcal{G}$-Invariant atomic arrangements The simplest model for a crystal is the static arrangement of atoms (molecules, clusters, etc.) where it suffices to fix some atomic positions within one primitive cell $\mathcal{P}(\mathcal{T})$ and to exploit its space group symmetry to obtain an infinitely extended periodic structure. Here the symbol $\mathcal{P}(\mathcal{T})$ denotes the primitive cell of the Bravais vector lattice $\mathcal{T}$ that is nothing but the parallelepiped spanned by the primitive basis vectors $\boldsymbol{t}$. Atomic positions in $\mathcal{P}(\mathcal{T})$ are denoted by $\boldsymbol{x} \in \mathcal{P}(\mathcal{T})$ and are classified by the so-called "Wyckoff positions." Here, for simplicity $G = (R \mid \boldsymbol{w}_R + \boldsymbol{t}) \in \mathcal{G}$ a shorthand notation for the space group elements is introduced as

$$\mathcal{G}(\boldsymbol{x}_j) = \{G \in \mathcal{G} \mid \mathcal{M}(G)\boldsymbol{x}_j = \boldsymbol{x}_j\} \qquad [17]$$

$$\{\boldsymbol{x}_j(\mathcal{G})\} = \{\boldsymbol{x}_j(G) = \mathcal{M}(G)\boldsymbol{x}_j \mid G \in \mathcal{G} : \mathcal{G}(\boldsymbol{x}_j)\} \qquad [18]$$

Groups of the type $\mathcal{G}(\boldsymbol{x}_j)$ defined by [17] are called site groups. Infinite sets $\{(\boldsymbol{x}_j)\mathcal{G}\}$ defined by [18] are called single site atomic arrangements. Two positions, say $\boldsymbol{x}_j$ and $\boldsymbol{x}_k$, are called $\mathcal{G}$ equivalent if and only if, there exists at least one space group element $G \in \mathcal{G}$ such that $\mathcal{M}(G)\boldsymbol{x}_j = \boldsymbol{x}_k$. Their site groups are then conjugate subgroups with respect to the space group $\mathcal{G}$ in question. However, if one cannot find a space group element $G \in \mathcal{G}$ such that the previous condition holds, then the two positions are called $\mathcal{G}$ inequivalent. The union of mutually disjoint single site atomic arrangements defined by

$$\begin{aligned}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n, \mathcal{G}\} \\ = \{\boldsymbol{x}_1(\mathcal{G})\} \cup \{\boldsymbol{x}_2(\mathcal{G})\} \cup \cdots \{\boldsymbol{x}_n(\mathcal{G})\}\end{aligned} \qquad [19]$$

is regarded as a multiple-site atomic arrangement which describes a periodic structure that possesses the space group $\mathcal{G}$ as its symmetry group. The notion of Wyckoff position is synonymous for connected subsets of the primitive cell $\mathcal{P}(\mathcal{T})$, whose points (vectors) possess image groups $\mathcal{P}(\boldsymbol{x}_j) \sim \mathcal{G}(\boldsymbol{x}_j)$ that are conjugated subgroups with respect to its crystallographic point group $\mathcal{P}$. Wyckoff positions may consist of isolated points, lines, planes, or compact subsets of $\mathcal{P}(\mathcal{T})$. This specific classification of points $\boldsymbol{x} \in \mathcal{P}(\mathcal{T})$ is of importance in physical applications. Since different elements or any fixed Wyckoff position must have the same point group symmetry, it may or may not lead to a change of energy when the atoms vary their positions within one Wyckoff position.

Space group–subgroup relations Group–subgroup relations between crystallographic space groups play a decisive role in structural phase transitions which are accompanied by changes in the symmetry. Usually, one assumes that the space group $\mathcal{H}$ of the distorted phase is a proper subgroup (of finite

index) of the space group $\mathcal{G}$ of the parent phase which implies that $\mathcal{H} \subset \mathcal{G}$. Due to the reduction of the symmetry at the phase transition, the distorted phase can appear in several homogeneous simultaneously coexisting states which have the same structure but different orientations and/or locations in space. Let $\mathcal{G} = \{\mathcal{T}, \mathcal{P} \mid \mathcal{O}, \boldsymbol{w}_{\mathcal{P}}\}$ be the superspace group and $\mathcal{H} = \{\mathcal{S}, \mathcal{Q} \mid \mathcal{O}', \boldsymbol{w}_{\mathcal{Q}}\}$ be the subspace group of the former. In fact, three significant different types of group–subgroup relations are possible, namely

$$\text{translationsgleiche}: \quad \mathcal{S} = \mathcal{T}, \ \ \mathcal{Q} \subset \mathcal{P} \qquad [20]$$

$$\text{klassengleiche}: \quad \mathcal{S} \subset \mathcal{T}, \ \ \mathcal{Q} = \mathcal{P} \qquad [21]$$

$$\text{general}: \quad \mathcal{S} \subset \mathcal{T}, \ \ \mathcal{Q} \subset \mathcal{P} \qquad [22]$$

where by means of "Hermann's theorem," it can be shown that every general group–subgroup relation between space groups can be traced back to various chains of maximal *translationsgleiche* or *klassengleiche* subspace groups. Group–subgroup relations of infinite index are given by V Kopsky and D B Litvin in 2002 in the International Tables for Crystallography, Volume E, Subperiodic Groups, where sectional layer groups and penetration rod groups are discussed as subgroups of three-dimensional crystallographic space groups.

Space Group Unirreps

This section focuses on the representation theory of space groups with special emphasis on their unirreps. It is well known that space group unirreps can be constructed systematically by applying Mackey's induction procedure, which is exhaustively discussed by C J Bradley and A P Cracknell in 1972 in The Mathematical Theory of Symmetry in Solids. Mackey's induction procedure relies, among others, on the assumption that the given group G possesses a normal subgroup N whose unirreps are known. Specifying these assumptions to space groups, it implies that one should start from the well-known one-dimensional $\mathcal{T}$ unirreps, since every crystallographic space group $\mathcal{G}$ contains a Bravais vector lattice $\mathcal{T}$ as its natural normal subgroup.

Reciprocal vector lattices In order to be able to define the $\mathcal{T}$ unirreps, one has to define first the corresponding reciprocal vector lattice $\mathcal{T}^\star$ together with its corresponding "Brillouin zone" $\mathcal{B}(\mathcal{T}^\star)$

$$\mathcal{T}^\star = \{\boldsymbol{K} = \boldsymbol{K} \cdot \underline{n} \mid \underline{n} \in \mathbb{Z}^3\} \qquad [23]$$

$$\boldsymbol{t}_j \cdot \boldsymbol{K}_l = 2\pi\delta_{jl} \qquad [24]$$

By definition, the vectors $\boldsymbol{K} = \{\boldsymbol{K}_1, \boldsymbol{K}_2, \boldsymbol{K}_3\}$, formally written as row vectors, form the basis of the corresponding reciprocal lattice $\mathcal{T}^\star$, where $\mathcal{T}$ is the underlying direct vector lattice. The Brillouin zone $\mathcal{B}(\mathcal{T}^\star)$ is the counterpart of the Wigner–Seitz cell of the underlying direct vector lattice. The primitive cell $\mathcal{P}(\mathcal{T}^\star)$ is spanned as parallelepiped by the basis vectors $\boldsymbol{K}$. Recall that $\mathcal{P}(\mathcal{T}^\star)$ and its symmetrical counterpart $\mathcal{B}(\mathcal{T}^\star)$ are equally well suited to describe the basic domains of the reciprocal vector lattice $\mathcal{T}^\star$.

$\mathcal{T}$ Unirreps Since $\mathcal{T}$ forms an abelian group, its unirreps over the field $\mathbb{C}$ of complex numbers are one-dimensional, which implies that they are unimodular numbers. In fact,

$$D^{\boldsymbol{k}}(E|\boldsymbol{t}) = \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{t}} \qquad [25]$$

$$\boldsymbol{k} = \alpha\boldsymbol{K}_1 + \beta\boldsymbol{K}_2 + \gamma\boldsymbol{K}_3 \qquad [26]$$

where not only $\boldsymbol{k} \in \mathcal{B}(\mathcal{T}^\star)$ but also $\boldsymbol{t} \in \mathcal{T}$ should be taken into account. It is to be noted that the $\mathcal{T}$ irrep label $\boldsymbol{k} \in \mathcal{B}(\mathcal{T}^\star)$ varies continuously over the Brillouin zone. For instance, in the case of primitive cubic Bravais vector lattices, the corresponding continuous parameters have $-1/2 < \alpha, \beta, \gamma \leqslant +1/2$ their domains of definitions to avoid double counting of eventually superfluous $\mathcal{T}$ irrep labels.

Periodic boundary conditions Once the so-called "periodic boundary conditions" are imposed to any given Bravais vector lattice $\mathcal{T}$, it implies that only finite homomorphic image of the originally countable Bravais vector lattice are considered. This reduces the originally compact continuous Brillouin zone $\mathcal{B}(\mathcal{T}^\star)$ to a finite set of vectors. Hence, the original continuous parameters α, β, γ become discrete and lose their continuity properties. Usually, periodic boundary conditions are introduced to avoid mathematical difficulties that arise from the fact that the $\mathcal{T}$ unnirreps (see eqn [25]) are normalized to delta functions. On the other hand, they reduce to ordinary Kronecker delta functions if periodic boundary conditions are imposed. Closely related to these difficulties is the fact that Bloch functions, which are associated with infinitely extended crystals, cannot be normalized to unity, since their ℓ_2 norms do not exist for fundamental reasons.

Little group unirreps The next step in Mackey's induction procedure consists of determining the corresponding little group $\mathcal{G}(\boldsymbol{k}) \subseteq \mathcal{G}$ for each $\mathcal{T}$ irrep label $\boldsymbol{k} \in \mathcal{B}(\mathcal{T}^\star)$. Once this is done, the corresponding

$\mathscr{G}(\boldsymbol{k})$ unirreps are computed. To summarize,

$$\mathscr{G}(\boldsymbol{k}) = \{G \in \mathscr{G} \mid D^{\boldsymbol{k}}(G^{-1}(E|\boldsymbol{t})G) = D^{\boldsymbol{k}}(E|\boldsymbol{t})\} \quad [27]$$

$$\mathscr{P}(\boldsymbol{k}) \sim \mathscr{G}(\boldsymbol{k})/\mathscr{T} \subseteq \mathscr{P} \quad [28]$$

$$\mathscr{P}(\boldsymbol{k}) = \{R \in \mathscr{P} \mid R\ \boldsymbol{k} = \boldsymbol{k} + \boldsymbol{K}(\boldsymbol{k}, R)\} \quad [29]$$

where, in particular, it is assumed that $\boldsymbol{K}(\boldsymbol{k}, R) \in \mathscr{T}^{\star}$ for all point group elements $R \in \mathscr{P}$. The homomorphic images $\mathscr{P}(\boldsymbol{k})$ of the little groups $\mathscr{G}(\boldsymbol{k})$ are usually called "little co-groups." It is to be noted that

$$\boldsymbol{K}(\boldsymbol{k}, RS) = \boldsymbol{K}(\boldsymbol{k}, R) + R\boldsymbol{K}(\boldsymbol{k}, S) \quad [30]$$

must be valid for all point group elements $R, S \in \mathscr{P}(\boldsymbol{k})$ as, otherwise, the closure condition for the little co-group $\mathscr{P}(\boldsymbol{k})$ would be violated. The closure condition [30] plays an important role when symmetrizing plane waves or any other type of wave functions. It is seen in this context that the reciprocal lattice vectors $\boldsymbol{K}(\boldsymbol{k}, R) \in \mathscr{T}^{\star}$ may only occur if and only if, the associated $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^{\star})$ belongs to the surface of the Brillouin zone $\mathscr{B}(\mathscr{T}^{\star})$ as otherwise they are zero.

Apart from this, corresponding $\mathscr{G}(\boldsymbol{k})$ unirreps are usually obtained by computing suitable $\mathscr{P}(\boldsymbol{k})$ unirreps. Provided the underlying space group G is symmorphic, suitable $\mathscr{P}(\boldsymbol{k})$ unirreps have to be ordinary vector unirreps. However, if the underlying space group $\mathscr{G}$ is nonsymmorphic, suitable $\mathscr{P}(\boldsymbol{k})$ unirreps are the so-called projective $\mathscr{P}(\boldsymbol{k})$ unnirreps. Nevertheless, the use of projective representations is merely an auxiliary tool which simplifies the construction of $\mathscr{P}(\boldsymbol{k})$ unirreps significantly. Accordingly, suitable $\mathscr{P}(\boldsymbol{k})$ unirreps are constrained by the following condition:

$$D^{\xi}(R)D^{\xi}(R) = \mathscr{F}^{\boldsymbol{k}}(R, S)D^{\xi}(RS) \quad [31]$$

$$\mathscr{F}^{\boldsymbol{k}}(R, S) = \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{t}_{R,S}) \quad [32]$$

for all point group elements $R, S \in \mathscr{P}(\boldsymbol{k})$ and where the special translations $\boldsymbol{t}_{\underline{R},\underline{S}} \in \mathscr{T}$ are defined by [14]. As previously noted, these vectors vanish, if the space group $\mathscr{G}$ is symmorphic, and hence the corresponding $\mathscr{P}(\boldsymbol{k})$ unirreps are ordinary vector representations. However, if the space group $\mathscr{G}$ is nonsymmorphic, the corresponding $\mathscr{P}(\boldsymbol{k})$ unirreps may be projective representations, since some $\boldsymbol{t}_{\underline{R},\underline{S}} \in \mathscr{T}$ may be nonzero. By definition, the constructions

$$D^{\xi,\boldsymbol{k}}(R|\boldsymbol{w}_R + \boldsymbol{t}) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}}D^{\xi}(R) \quad [33]$$

$$\dim D^{\xi,\boldsymbol{k}}(\mathscr{G}(\boldsymbol{k})) = n(\xi) \quad [34]$$

$$D^{\xi,\boldsymbol{k}}(R|\boldsymbol{w}_R + \boldsymbol{t})D^{\xi,\boldsymbol{k}}(S|\boldsymbol{w}_S + \boldsymbol{v}) = D^{\xi,\boldsymbol{k}}(RS|\boldsymbol{w}_{RS} + \boldsymbol{t}_{R,S} + \boldsymbol{t} + R\boldsymbol{v}) \quad [35]$$

define ordinary vector $\mathscr{G}(\boldsymbol{k})$ unirreps, provided that suitable $\mathscr{P}(\boldsymbol{k})$ unirreps are constructed. Note in particular that $\xi \in \mathscr{A}(\mathscr{P}(\boldsymbol{k}))$ define the complete sets of all $\mathscr{P}(\boldsymbol{k})$ irrep labels. In summary, the set of ordered pairs

$$\mathscr{A}(\mathscr{G}(\boldsymbol{k})) = \{(\boldsymbol{k}, \xi) | \xi \in \mathscr{A}(\mathscr{P}(\boldsymbol{k}))\} \quad [36]$$

define complete sets of $\mathscr{G}(\boldsymbol{k})$ unirreps for each $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^{\star})$.

Full space group unirreps The final step is, among others, to restrict the Brillouin zone $\mathscr{B}(\mathscr{T}^{\star})$ to a suitable subset, sometimes called "representation domain," $\Delta\mathscr{B}(\mathscr{T}^{\star}, \mathscr{P})$, in order to avoid double counting of equivalent $\mathscr{G}$ unirreps. This is done by taking only one representative (arm) from each $\boldsymbol{k}$-vector star $S(\boldsymbol{k})$ defined by

$$S(\boldsymbol{k}) = \{\boldsymbol{k}_R \in \mathscr{B}(\mathscr{T}^{\star}, \mathscr{P}) \mid \boldsymbol{k}_R = R\boldsymbol{k};\ R \in \mathscr{P}\} \quad [37]$$

$$|S(\boldsymbol{k})| = |\mathscr{P} : \mathscr{P}(\boldsymbol{k})| \quad [38]$$

It is worth noting that the order $|S(\boldsymbol{k})|$ of any star is a divisor of the order $|\mathscr{P}|$ of the corresponding crystallographic point group $\mathscr{P}$. The union set of all these representative $\boldsymbol{k}$ vectors defines the representation domain $\Delta\mathscr{B}(\mathscr{T}^{\star}, \mathscr{P})$. As commonly accepted in physics, one takes a simply connected continuous subset $\Delta\mathscr{B}(\mathscr{T}^{\star}, \mathscr{P})$ of the original Brillouin zone $\mathscr{B}(\mathscr{T}^{\star})$ in order to avoid unnecessary difficulties when dealing, for instance, with the so-called "compatibility relations," whichever have been discussed for the first time by L P Bouckaert *et al.*, in 1936.

The actual final step in Mackey's induction procedure consists in inducing the full $\mathscr{G}$ representations from the $\mathscr{G}(\boldsymbol{k})$ unirreps given by [33]. A general theorem guarantees that the corresponding induced $\mathscr{G}$ representations are automatically irreducible. The induction formulas read

$$D^{(\boldsymbol{k},\xi)\uparrow\mathscr{G}}_{\underline{R},\underline{S}}(G) = \delta_{\underline{R}\mathscr{P}(\boldsymbol{k}),R\underline{S}\mathscr{P}(\boldsymbol{k})}D^{\boldsymbol{k},\xi}(\underline{A}^{-1}G\underline{B}) \quad [39]$$

$$G = (R \mid \boldsymbol{w}_R + \boldsymbol{t}) \quad [40]$$

$$\underline{A} = (\underline{R}|\boldsymbol{w}_{\underline{R}}) \quad [41]$$

$$\underline{B} = (\underline{S}|\boldsymbol{w}_{\underline{S}}) \quad [42]$$

$$\delta_{\underline{R}\mathscr{P}(\boldsymbol{k}),R\underline{S}\mathscr{P}(\boldsymbol{k})} = \begin{cases} 1, & \underline{R}^{-1}R\underline{S} \in \mathscr{P}(\boldsymbol{k}) \\ 0, & \text{otherwise} \end{cases} \quad [43]$$

$$\dim D^{(\boldsymbol{k},\xi)\uparrow\mathscr{G}}(\mathscr{G}) = |\mathscr{P} : \mathscr{P}(\boldsymbol{k})| \cdot n(\xi) \quad [44]$$

where the specific space group elements $\underline{A} = (\underline{R}|\boldsymbol{w}_{\underline{R}})$ and $\underline{B} = (\underline{S}|\boldsymbol{w}_{\underline{S}})$ are coset representatives of the space group $\mathscr{G}$ with respect to the corresponding little

group $\mathscr{G}(\boldsymbol{k})$. Specifying the conjugation $\underline{A}^{-1}G\underline{B}\in\mathscr{G}$ by taking into account [33] and [43], one immediately arrives at the following final formulas for $\mathscr{G}$ unirreps:

$$D^{(\boldsymbol{k},\xi)\uparrow\mathscr{G}}_{\underline{R}a,\underline{S}b}((R|\boldsymbol{w}_R+\boldsymbol{t})) = \delta_{\underline{R}\mathscr{P}(\boldsymbol{k}),R\underline{S}\mathscr{P}(\boldsymbol{k})}\mathrm{e}^{-\mathrm{i}\underline{R}\boldsymbol{k}\cdot\boldsymbol{t}}\Phi^{\boldsymbol{k}}_{\underline{R},\underline{S}}(R)D^{\xi}_{ab}(\underline{R}^{-1}R\underline{S}) \quad [45]$$

$$\Phi^{\boldsymbol{k}}_{\underline{R},\underline{S}}(R) = \exp\bigl(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_{\underline{R},\underline{S}}(R)\bigr) \quad [46]$$

$$\boldsymbol{t}_{\underline{R},\underline{S}}(R) = \boldsymbol{t}_{\underline{R}^{-1},R} + \boldsymbol{t}_{\underline{R}^{-1}R,\underline{S}} - \underline{R}^{-1}\boldsymbol{t}_{\underline{R},\underline{R}^{-1}}\in\mathscr{T} \quad [47]$$

Clearly, if $\mathscr{G}$ defines a symmorphic space group, then the corresponding phase factors $\Phi^{\boldsymbol{k}}_{\underline{R},\underline{S}}(R)$ defined by [45] reduce to 1, since the special translation vectors $\boldsymbol{t}_{R,S}$ are zero-vectors. However, if $\mathscr{G}$ defines a nonsymmorphic space group, then some of the phase factors defined by [45] are nontrivial and lead to quite different expressions for its corresponding $\mathscr{G}$ unirreps.

Operator Representations of Space Groups

Usually, space group symmetry and space group representations are applied in the one-particle approximation to Hamiltonians that describe the motion of an electron in periodic structures quantum mechanically. The first step consists of defining a homomorphism $\varepsilon:\mathscr{G}\to U(\mathscr{G})$ of the given space group $\mathscr{G}$ into a group of unitary operators $U(\mathscr{G})$ that map the underlying Hilbert space $\mathscr{H}=L^2(\mathbb{R}^3)$ onto itself (here the spin degree of freedom has been neglected):

$$U(\mathscr{G}) = \{U(G)|G\in\mathscr{G}\} \quad [48]$$

$$[U(G)\Phi](\boldsymbol{x}) = \phi(\mathscr{M}(G^{-1})\boldsymbol{x}) \quad [49]$$

$$U(G)U(G') = U(GG') \quad [50]$$

In this context, it is worth noting that every mapping $\mathscr{M}(G):\mathbb{R}^3\to\mathbb{R}^3$ is nonsingular but likewise nonlinear, if and only if the translational part $\boldsymbol{t}\in\mathscr{T}$ contained in $G=(R|\boldsymbol{w}_R+\boldsymbol{t})$ is nonzero, whereas the uniquely associated unitary operator representation is linear with respect to the Hilbert space $L^2(\mathbb{R}^3)$ in any case. The one-particle Hamiltonian reads

$$H = \frac{1}{2m}P^2 + V(X) \quad [51]$$

where the potential is responsible for the symmetry of the problem. Apart from this, space groups and their representation theory are likewise utilized in order to simplify systematically the diagonalization of the force matrix in the "harmonic approximation" when dealing with lattice vibrations in classical point mechanics.

Bloch Theorem – Energy Bands

The eigenvalue problem of periodic one-particle Hamiltonians, where the spin degree of freedom is neglected, presents one of the most popular examples where group-theoretical methods are applied. Taking into account that every space group $\mathscr{G}$, whether $\mathscr{G}$ is symmorphic or nonsymmorphic, contains a countable translational group $\mathscr{T}$ as a normal subgroup, one may use, as a first step, the translational symmetry of the Hamiltonian represented by the vanishing commutators

$$[H, U(E|\boldsymbol{t})] = 0 \quad \forall\boldsymbol{t}\in\mathscr{T} \quad [52]$$

to simplify the eigenvalue problem. According to a general theorem of functional analysis, the countable set $U(\mathscr{T})=\{U(E|\boldsymbol{t})|\boldsymbol{t}\in\mathscr{T}\}$ of mutually commuting unitary operators and the Hamiltonian H can be diagonalized simultaneously.

Bloch theorem Bloch's theorem predicts partly the form of the common eigenfunctions of the periodic Hamiltonian. It leads to the following well-known and extensively used statement:

$$\Psi^{\boldsymbol{k}}(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}w(\boldsymbol{k},\boldsymbol{x}) \quad [53]$$

$$w(\boldsymbol{k},\boldsymbol{x}) = w(\boldsymbol{k},\boldsymbol{x}+\boldsymbol{t}) \quad \forall\boldsymbol{t}\in\mathscr{T} \quad [54]$$

These eigenfunctions are called "Bloch functions," of the Hamiltonian, and the unitary translational operators have the form given by [53] and [54]. The translational symmetry has been utilized for this factorization. The eigenvalue problem reads

$$H\Psi^{\boldsymbol{k}} = E(\boldsymbol{k})\Psi^{\boldsymbol{k}} \quad [55]$$

where the eigenfunctions are assumed to be Bloch functions. What remains to be done in practical applications is the determination of the periodic Bloch factors

$$w(\boldsymbol{k}):\mathscr{P}(\mathscr{T})\to\mathbb{C} \quad [56]$$

for every $\boldsymbol{k}\in\mathscr{B}(\mathscr{T}^{\bigstar})$, which is a problem in its own right. Simple manipulations yield the following noncountable infinite set of decoupled partial differential equations for the unknown Bloch factors:

$$H(\boldsymbol{k}) = \frac{1}{2m}(\boldsymbol{Q}+\hbar\boldsymbol{k})^2 + V(X) \quad [57]$$

$$H(\boldsymbol{k})w(\boldsymbol{k}) = E(\boldsymbol{k})w(\boldsymbol{k}) \quad [58]$$

for all $\boldsymbol{k}\in\mathscr{B}(\mathscr{T}^{\bigstar})$ where, in particular, the operator $\boldsymbol{Q}=-\mathrm{i}\hbar\nabla$ is the usual differential operator but its domain of definition is the Hilbert space $L^2(\mathscr{P}(\mathscr{T}))$

together with periodic boundary conditions which give rise to a pure point spectrum of the operator $\boldsymbol{Q}$ in contrast to $\boldsymbol{P} = -\mathrm{i}\hbar\nabla$ that possesses a continuous spectrum with respect to $L^2(\mathbb{R}^3)$. To recapitulate, if one does not introduce periodic boundary conditions for the original eigenvalue problem, admissible vectors $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^\star)$ vary continuously over the Brillouin zone, whereas the spectrum of the momentum operator $\boldsymbol{Q}$ is discrete and coincides with the corresponding reciprocal lattice $\boldsymbol{K} \in \mathscr{T}^\star$. Assuming that [58] has been solved for all $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^\star)$, then [55] reads

$$H\Psi_s^{\boldsymbol{k},n} = E_n(\boldsymbol{k})\Psi_s^{\boldsymbol{k},n} \qquad [59]$$

where $n \in \mathbb{N}$ is usually called the "band index," where $s = 1, 2, \ldots, \deg E_n(\boldsymbol{k})$ describes possible degeneracies of the eigenvalues $E_n(\boldsymbol{k})$. The band index n is a countable index, since the Hamiltonians $H(\boldsymbol{k})$ given by [57] possess pure point spectra for fixed $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^\star)$.

Energy bands To define energy bands in terms of functions of the following type $E_n : \mathscr{B}(\mathscr{T}^\star) \to \mathbb{R}$, one has, in principle, two different possibilities. Either one defines these functions by assuming the strict ordering

$$1.\quad E_1(\boldsymbol{k}) < E_2(\boldsymbol{k}) < \cdots < E_n(\boldsymbol{k}) < E_{n+1}(\boldsymbol{k}) < \cdots \qquad [60]$$

$$2.\quad E_n(\boldsymbol{k}) = \text{analytic functions} \qquad [61]$$

or one demands that these functions be analytic functions. The second possibility is usually preferred, since it leads to smooth functions with respect to the variable $\boldsymbol{k} \in \mathscr{B}(\mathscr{T}^\star)$, whereas the first possibility may lead to cusps with band contacts and hence accidental degeneracies. Obviously, the second possibility may likewise lead to band crossings, which indicate the same type of accidental degeneracies.

Wannier functions Assume that the eigenfunctions $\{\Psi_s^{\boldsymbol{k},n}\}$ are orthonormal with respect to the scalar product of the underlying Hilbert space $\mathscr{H}$:

$$\langle \Psi_s^{\boldsymbol{k},n}, \Psi_s^{\boldsymbol{k}',n'} \rangle_{\mathscr{H}} = \delta(\boldsymbol{k} - \boldsymbol{k}')\delta_{nn'}\delta_{ss'} \qquad [62]$$

where the entries $\delta(\boldsymbol{k} - \boldsymbol{k}')$ define the delta functions. This implies that for the infinite system, that is, system without periodic boundary conditions, Bloch functions cannot be normalized to unity. This is one of the main reasons for the introduction of periodic boundary conditions, since then the delta-functions reduce to ordinary Kronecker deltas, which means that the corresponding Bloch functions become square-integrable functions. In the case of infinite systems, Wannier functions are defined as

$$\Omega_s^{n,\boldsymbol{t}} = \frac{1}{\sqrt{|\mathscr{B}(\mathscr{T}^\star)|}} \int_{\mathscr{B}(\mathscr{T}^\star)} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}} \Psi_s^{\boldsymbol{k},n} \qquad [63]$$

$$\langle \Omega_s^{n,\boldsymbol{t}}, \Omega_{s'}^{n',\boldsymbol{t}'} \rangle_{\mathscr{H}} = \delta_{\boldsymbol{t},\boldsymbol{t}'}\delta_{nn'}\delta_{ss'} \qquad [64]$$

which are square integrable and orthonormal with respect to all labels $\boldsymbol{t}$, n, s respectively. These functions are localized around the positions $\boldsymbol{t} \in \mathscr{T}$ but have infinitely many oscillations to achieve their mutual orthogonality. Whether the Bloch functions $\Psi_s^{\boldsymbol{k},n}$ are directly taken to construct [63] or specific linear combinations (U matrix approach) of $\{\Psi_s^{\boldsymbol{k},n} \mid s = 1, 2, \ldots, \deg E_n(\boldsymbol{k})\}$ are considered, turns out to be problem of its own right and is sensitively influenced by the physical problem.

Energy bands – full space group symmetry More information is gained regarding the degeneracy of the eigenvalues $E_n(\boldsymbol{k})$ and their corresponding eigenfunctions $\{\Psi_s^{\boldsymbol{k},n}\}$, if the full space group symmetry is exploited. The correspondingly refined eigenvalue equation reads

$$H\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n} = E_n^\xi(\boldsymbol{k})\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n} \qquad [65]$$

$$E_n^\xi(\boldsymbol{k}) = E_n^\xi(\underline{R}\boldsymbol{k}) \quad \forall \underline{R} \in \mathscr{P} : \mathscr{P}(\boldsymbol{k}) \qquad [66]$$

$$\deg E_n^\xi(\boldsymbol{k}) = |\mathscr{P} : \mathscr{P}(\boldsymbol{k})| \cdot n(\xi) \qquad [67]$$

where the eigenfunctions $\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}$ are written as $\mathscr{G}$-symmetry-adapted functions, the eigenvalues $E_n^\xi(\boldsymbol{k})$ are labeled by the band index n and the $\mathscr{G}$ irrep labels. Apart from this, the formula [67] for the degeneracies of the eigenvalues is valid, if and only if there is no accidental degeneracy caused by band crossings. The symmetry properties (see eqn [66]) of the eigenvalues explain why energy band calculations are restricted to subsets of $\boldsymbol{k}$ vectors, namely $\boldsymbol{k} \in \Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$, where the symbol $\Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$ denotes the so-called representation domain of the Brillouin zone $\mathscr{B}(\mathscr{T}^\star)$. Note that $\Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$ presents, loosely speaking, the $|\mathscr{P}|$th part (simply connected subwedge) of the Brillouin zone. The restriction to $\Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$ guarantees the uniqueness and completeness of the corresponding $\mathscr{G}$-irrep label set $\mathscr{A}(\mathscr{G})$:

$$\mathscr{A}(\mathscr{G}) = \{(\boldsymbol{k}, \xi) \mid \boldsymbol{k} \in \Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P});\ \xi \in \mathscr{A}(\mathscr{P}(\boldsymbol{k}))\} \qquad [68]$$

Compatibility relations It was in the famous article of L P Bouckaert *et al*. in 1936 where for the first time continuity properties of $\mathscr{G}$-symmetry-adapted eigenfunctions were discussed. The basic idea was

to change smoothly the continuous parameters $\boldsymbol{k}\in\Delta\mathscr{B}(\mathscr{T}^{\bigstar},\mathscr{P})$ in the vicinity of some $\boldsymbol{k}$-vectors and study the transformation properties of the corresponding $\mathscr{G}$-symmetry-adapted eigenfunctions $\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}$ at different $\boldsymbol{k}$-positions, say $\boldsymbol{k}$ and $\boldsymbol{k}'=\boldsymbol{k}+\varepsilon$. In addition, the authors studied whether the assigned eigenvalues, say $E_n^{\xi}(\boldsymbol{k})$ and $E_n^{\xi}(\boldsymbol{k}+\varepsilon)$ may or may not split, depending on the relationship of the associated little co-groups $\mathscr{P}(\boldsymbol{k})$ and $\mathscr{P}(\boldsymbol{k}+\varepsilon)$. The important result culminated in the so-called compatibility relations which allow one to make some predictions on the splitting of energy bands. To summarize, if and only if $\mathscr{P}(\boldsymbol{k})=\mathscr{P}(\boldsymbol{k}+\varepsilon)$, then, because of $D^{\boldsymbol{k},\xi}(\mathscr{P}(\boldsymbol{k}))\downarrow\mathscr{P}(\boldsymbol{k}+\varepsilon)=D^{\boldsymbol{k},\xi}(\mathscr{P}(\boldsymbol{k}))$ it follows that $E_n^{\xi}(\boldsymbol{k}+\varepsilon)$ must not split. Conversely, if and only if, $\mathscr{P}(\boldsymbol{k})\subset\mathscr{P}(\boldsymbol{k}+\varepsilon)$, and if, in addition, $D^{\boldsymbol{k},\xi}(\mathscr{P}(\boldsymbol{k}))\downarrow\mathscr{P}(\boldsymbol{k}+\varepsilon)$ becomes reducible, then the corresponding eigenvalue $E_n^{\xi}(\boldsymbol{k}+\varepsilon)$ may split, unless accidental degeneracies occur. The compatibility relations are an indispensable tool when calculating energy bands or phonon dispersion relations, since in many situations the numerical accuracy may be insufficient to decide whether eigenvalues degenerate or not.

Space Group Symmetrized States

As already noted, the $\mathscr{G}$-symmetrized eigenfunctions $\{\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}\}$, which are normalized to unity, are mutual orthogonal and, by definition, transform according to the $\mathscr{G}$ unirreps. One simply proves that the eigenfunctions

$$\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}(\boldsymbol{x})=\mathrm{e}^{\mathrm{i}\underline{R}\boldsymbol{k}\cdot\boldsymbol{x}}w_a^{\xi,n}(\underline{R}\boldsymbol{k},\boldsymbol{x}) \qquad [69]$$

retain their structure to form Bloch functions, where the lattice periodic Bloch factors carry the transformation properties with respect to the correspondingly conjugated little co-groups $\mathscr{P}(\underline{R}\boldsymbol{k})=\underline{R}\mathscr{P}(\boldsymbol{k})\underline{R}^{-1}$. Turning back to the penultimate step of the induction procedure, one has the forms

$$\Psi_{\underline{E}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}(\boldsymbol{x})=\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}w_a^{\xi,n}(\boldsymbol{k},\boldsymbol{x}) \qquad [70]$$

$$U(R\,|\,\boldsymbol{w}_R)\Psi_{\underline{E}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}=\sum_{b=1}^{n(\xi)}D_{ba}^{\boldsymbol{k},\xi}(R)\Psi_{\underline{E}b}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n} \qquad [71]$$

which represents the special case $\underline{R}=E$ where, especially, $G=(R|\boldsymbol{w}_R)\in\mathscr{G}(\boldsymbol{k})$ with $R\in\mathscr{P}(\boldsymbol{k})$ and $\boldsymbol{t}=0$. This yields the subset of the eigenfunctions that must transform according to the corresponding $\mathscr{G}(\boldsymbol{k})$ unirreps. Accordingly, one can define

$$[U(R\,|\,\boldsymbol{w}_R)\Psi_{\underline{E}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}](\boldsymbol{x})$$
$$=[\Psi_{\underline{E}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}](\mathscr{M}(R\,|\,\boldsymbol{w}_R)^{-1}\boldsymbol{x}) \qquad [72]$$

$$=\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}[V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)w_a^{\xi,n}(\boldsymbol{k})](\boldsymbol{x}) \qquad [73]$$

where $R\in\mathscr{P}(\boldsymbol{k})$ is assumed. It is worth noting that definition [72] is used to assign uniquely via definition [73] to each unitary operator $U(R|\boldsymbol{w}_R)$ for every $\boldsymbol{k}$-vector belonging to $\Delta\mathscr{B}(\mathscr{T}^{\bigstar},\mathscr{P})$ an isometrical operator $V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)$ where, especially, the constraints $R\in\mathscr{P}(\boldsymbol{k})$ must not be forgotten. To summarize,

$$[V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)w_a^{\xi,n}(\boldsymbol{k})](\boldsymbol{x})$$
$$=\mathrm{e}^{-\mathrm{i}R\boldsymbol{k}\cdot\boldsymbol{w}_R}\mathrm{e}^{\mathrm{i}\boldsymbol{K}(\boldsymbol{k},R)\cdot\boldsymbol{x}}w_a^{\xi,n}(\boldsymbol{k},R^{-1}(\boldsymbol{x}-\boldsymbol{w}_R)) \qquad [74]$$

where every set of operators $V^{\boldsymbol{k}}(\mathscr{P}(\boldsymbol{k}))=\{V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)|R\in\mathscr{P}(\boldsymbol{k})\}$ forms a generalized unitary operator representation with respect to the Hilbert space $L^2(\mathscr{P}(\mathscr{T}))$, since these operators leave the $L^2(\mathscr{P}(\mathscr{T}))$ scalar product invariant. Rather straightforward manipulations yield

$$V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)V^{\boldsymbol{k}}(S|\boldsymbol{w}_S)=\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_{R,S}}V^{\boldsymbol{k}}(RS|\boldsymbol{w}_{RS}) \qquad [75]$$

for all point group elements $R,S\in\mathscr{P}(\boldsymbol{k})$. In fact, the operator representation $V^{\boldsymbol{k}}(\mathscr{P}(\boldsymbol{k}))$ defines the so-called "projective multiplier" representations, since the occurrence of the factor system [32] on the RHS of [75] shows, among others, that they must be projective ones. The notion multiplier representation comes from the extra factor $\mathrm{e}^{-\mathrm{i}\boldsymbol{K}(\boldsymbol{k},R)\cdot\boldsymbol{x}}$ on the RHS of [74], where the special reciprocal vectors $\boldsymbol{K}(\boldsymbol{k},R)\in\mathscr{T}^{\bigstar}$ may occur, if and only if the vectors $\boldsymbol{k}\in\Delta\mathscr{B}(\mathscr{T}^{\bigstar},\mathscr{P})$ belong to the surface of the corresponding Brillouin zone $\mathscr{B}(\mathscr{T}^{\bigstar})$. Accordingly,

$$V^{\boldsymbol{k}}(R|\boldsymbol{w}_R)w_a^{\xi,n}(\boldsymbol{k})=\sum_{b=1}^{n(\xi)}D_{ba}^{\boldsymbol{k},\xi}(R)w_b^{\xi,n}(\boldsymbol{k}) \qquad [76]$$

which represents the transformation properties of the lattice-periodic $\mathscr{P}(\boldsymbol{k})$-symmetrized Bloch factors. It is to be noted that even in the case of symmorphic space groups the extra factors $\mathrm{e}^{-\mathrm{i}\boldsymbol{K}(\boldsymbol{k},R)\cdot\boldsymbol{x}}$ on the RHS of [74] occur and hence crucially enter into the discussions when constructing $\mathscr{G}(\boldsymbol{k})$-symmetrized functions as the penultimate step in the induction procedure. Point symmetrized states have been extensively discussed by S L Altmann and P Herzig in 1994.

Symmetrized plane waves A brief comment on the construction of symmetrized plane waves is made here. By virtue of the previous discussions, it is immediately clear that the construction of all $\mathscr{P}(\boldsymbol{k})$-symmetrized bases of the image Hilbert space $L^2(\mathscr{P}(\mathscr{T}))$ is sufficient to obtain symmetrized plane waves, since the last step in the induction procedure consists in inducing from $\mathscr{G}(\boldsymbol{k})$-symmetrized bases, the $\mathscr{G}$-symmetrized bases. As noted before, an orthonormalized basis of the Hilbert space $L^2(\mathscr{P}(\mathscr{T}))$ is

defined by the plane wave basis $\{\Phi^K \mid K \in \mathscr{T}^\star\}$:

$$\Phi^K(\boldsymbol{x}) = \frac{1}{\sqrt{|\mathscr{P}(\mathscr{T})|}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{x}} \qquad [77]$$

$$\langle \Phi^K, \Phi^{K'} \rangle_{L^2(\mathscr{P}(\mathscr{T}))} = \delta_{K,K'} \qquad [78]$$

Now, let Φ^K be a fixed plane wave. To achieve $\mathscr{P}(\boldsymbol{k})$-symmetrization, one must apply for the chosen $\boldsymbol{k} \in \Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$, the corresponding set of unitary operators $V^k(\mathscr{P}(\boldsymbol{k}))$ to the fixed plane wave Φ^K which induces a $\mathscr{P}(\boldsymbol{k})$ matrix representation, which in general should be decomposed into a direct sum of its irreducible constituents. The definition [74] yields

$$\begin{aligned}[V^k(R|\boldsymbol{w}_R)\Phi^K](\boldsymbol{x}) &= \mathrm{e}^{-\mathrm{i}R\boldsymbol{k}\cdot\boldsymbol{w}_R}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}(\boldsymbol{k},R)\cdot\boldsymbol{x}}\Phi^K(R^{-1}(\boldsymbol{x}-\boldsymbol{w}_R)) \\ &= \mathrm{e}^{-\mathrm{i}R(\boldsymbol{k}+\boldsymbol{K})}\Phi^{R\boldsymbol{K}+\boldsymbol{K}(\boldsymbol{k},R)}(\boldsymbol{x}) \end{aligned} \qquad [79]$$

which illustrates among others, the importance of the specific reciprocal lattice vectors $\boldsymbol{K}(\boldsymbol{k}, R) \in \mathscr{T}^\star$, where the latter inevitably occur if the corresponding $\boldsymbol{k}$-vector $\boldsymbol{k} \in \Delta\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$ belongs to the surface of the Brillouin zone $\mathscr{B}(\mathscr{T}^\star)$. To summarize, even in the case of symmorphic space groups, nontrivial multiplier representations may be realized. Once the $\mathscr{P}(\boldsymbol{k})$ matrix representations are decomposed into their irreducible constituents, the last step of the induction procedure consists in passing over from the corresponding $\mathscr{G}(\boldsymbol{k})$ unirreps to $\mathscr{G}$ unirreps in order to construct the correspondingly symmetrized plane waves.

Selection rules for space groups Let $\mathscr{G}$ be a space group and $\{\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}\}$ be some $\mathscr{G}$-symmetrized eigenfunctions of the Hamiltonian H in question. Moreover, let $T^{(\boldsymbol{k}_o,\lambda_o)\uparrow\mathscr{G}} = \{T_{\underline{S}s}^{(\boldsymbol{k}_o,\lambda_o)\uparrow\mathscr{G}} \mid \underline{S} \in \mathscr{P} : \mathscr{P}(\boldsymbol{k}); s = 1, 2, \ldots, n(\lambda_o)\}$, an irreducible $\mathscr{G}$-tensor operator of rank $(\boldsymbol{k}_o, \lambda_o)\uparrow\mathscr{G}$. According to the Wigner–Eckart theorem, one expects

$$\begin{aligned}&\langle \Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}, T_{\underline{S}s}^{(\boldsymbol{k}_o,\lambda_o)\uparrow\mathscr{G}}\Psi_{\underline{R}'a'}^{(\boldsymbol{k}',\xi')\uparrow\mathscr{G},n'} \rangle \\ &= \sum_{w=1}^{m((\boldsymbol{k}_o,\lambda_o),(\boldsymbol{k}',\xi')|(\boldsymbol{k},\xi))} \left(\begin{matrix} (\boldsymbol{k}_o,\lambda_o) & (\boldsymbol{k}',\xi') \\ \underline{S}s & \underline{R}'a' \end{matrix} \middle| \begin{matrix} (\boldsymbol{k},\xi)w \\ \underline{R}a \end{matrix} \right)^* \\ &\quad \times \langle \Psi^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n} \| T^{(\boldsymbol{k}_o,\lambda_o)\uparrow\mathscr{G}} \| \Psi^{(\boldsymbol{k}',\xi')\uparrow\mathscr{G},n'} \rangle_w \end{aligned} \qquad [80]$$

where the absolute square of the matrix elements describes, apart from a time dependent factor, "Fermi's golden rule," namely the transition probability from the initial state $\Psi_{\underline{R}'a'}^{(\boldsymbol{k}',\xi')\uparrow\mathscr{G},n'}$ to the final state $\Psi_{\underline{R}a}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}$, provided that the interaction operator is described by the irreducible $\mathscr{G}$ tensor operator component $T_{\underline{S}s}^{(\boldsymbol{k}_o,\lambda_o)\uparrow\mathscr{G}}$. The appearance of nontrivial multiplicities

$$m((\boldsymbol{k}_o,\lambda_o),(\boldsymbol{k}',\xi')|(\boldsymbol{k},\xi)) > 1$$

causes additional difficulties, since not only the determination of suitable Clebsch–Gordan coefficients becomes more complicated, but also the dependence of the reduced matrix elements causes additional problems. These problems might be the main reason why the application of the Wigner–Eckart theorem in solid-state physics has not been popular. Basic material concerning space group unirreps and Clebsch–Gordan coefficients for space groups were cited by A P Cracknell *et al.*, in 1979 and 1980. Important applications, such as infrared absorptions or Raman spectroscopy, of this topic are discussed by J L Birman and J F Cornwell in 1974 and 1984, respectively.

What Group Theory Provides

Finally, the usefulness of group-theoretical methods, especially in solid state physics, when applied to simplify certain tasks, is summarized. Here, exclusive comments on group-theoretical tools and methods which refer to space groups, subgroups of space groups, and eventually certain homomorphic images of space groups are made. Possible applications of so-called magnetic space groups, spin groups, and other groups to describe the symmetries of more complex periodic, or even quasiperiodic structures, such as the symmetry properties of quasicrystals, are not discussed here.

- *Bravais vector lattices and sublattices*. Group theory presents an effective tool to identify vector lattices and their sublattices, which are useful for determining not only their crystal system and their crystal class, but also their correlations to standardized lattices. Standardized lattices have basis vectors of the shortest length and specific orientation.
- *Deformed and distorted lattices*. By definition, deformed lattices are obtained by continuously varying the lattice parameters, whereas distorted lattices have lattice matrix that is post-multiplied by a nonsingular matrix which describes the distortion. The proper identification of such lattices is of vital importance in practical applications, for instance, in relaxation processes.
- *Subgroups of space groups of finite index*. Group–subgroup relations of space groups of finite index are of interest in structural phase transitions that are accompanied with changes in symmetry. The knowledge of admissible chains of subgroups from the determination of all intermediate groups is vital in such discussions, since they allow one to

predict the preferred chains of admissible structures of the intermediate phases.

- *Coset and double cosets of space groups*. In phase transitions (with changing structures) which are accompanied with the loss of symmetries, it turns out that simultaneously coexisting domain states are generated whose positions and orientations in space are directly related to the coset decomposition of the parent space group with respect to the subspace group (and correspondingly conjugated subspace groups) which describes the symmetry of the domain states. Likewise, double coset decompositions are used to describe the so-called ordered domain states.
- *Sectional layer and penetration rod groups*. The systematic determination of the symmetries of crystallographic planes, which are subsumed under the notion of sectional layer groups, as subgroups of three-dimensional space groups, plays an essential role in the analysis of the physical properties of two-dimensional surfaces. Likewise, the symmetries of crystallographic lines denoted as penetration rod groups are of interest when describing one-dimensional defects.
- *Point group symmetries of tensors*. Macroscopic properties, such as electrical conductivity, are usually described by tensors of certain ranks. In the presence of some symmetries, for example, point group symmetries, the invariance of such tensors with respect to point groups leads to restrictions as regards their nonzero components.
- *Space group unirreps*. The computation of space group unirreps by means of Mackey's induction procedure can be carried out systematically for every space group in question. Likewise, this approach allows one to gain more insight into the constructions of space–group-symmetrized states.
- *Space group Clebsch–Gordan coefficients*. Knowledge of sets of space group Clebsch–Gordan coefficients allows one not only to study selection rules when calculating transition probabilities, but also to construct systematically product states that transform according to the Kronecker product of the space group in question.
- *Bloch theorem*. One of the most prominent applications of group-theoretical methods in solid-state physics is presented by the Bloch theorem. This theorem allows one to predict the functional form of the common eigenfunctions of the mutually commuting translational operators which represent the Bravais vector lattice and the periodic Hamiltonian, where the latter is assumed to commute with the translational group.
- *Spectral properties of periodic Hamiltonians*. Once, the entire space group symmetry of the periodic Hamiltonian is taken into account, more information gained not only as regards the point group symmetries of the energy eigenvalues, but also systematic construction of the corresponding eigenstates.
- *Compatibility relations*. The continuity behavior of the energy eigenvalues and their associated eigenfunctions in the vicinity of some $\boldsymbol{k}$-vectors can be described by some specific subduction rules of the corresponding $\mathscr{P}(\boldsymbol{k})$ unirreps when passing over continuously to some adjacent $\boldsymbol{k}$-vectors with lower symmetry. Thus, group theory provides a necessary condition which decides whether the eigenvalues do not split or alternatively may eventually split.
- *Transformation properties of states*. An important task in applications, for instance energy band calculations, concerns determining according to which representation the lattice periodic Bloch factors transform. Every careful treatment of this problem inevitably leads to multiplier or even projective multiplier representations of the corresponding little co-groups provided that in the latter case the underlying space group is nonsymmorphic.
- *Symmetrized states*. The systematic construction of space group symmetry-adapted functions follows closely along the lines of Mackey's induction procedure for space group unirreps. For instance, symmetrized plane waves may be useful when calculating matrix elements of some interaction operators.
- *Symmetrized operation*. In order to apply the Wigner–Eckart theorem to space groups, irreducible space group tensors are required. However, in many cases the interaction operators do not represent irreducible space group tensors. Nevertheless, group theory allows one to decompose any operator into a sum of its irreducible space group tensor components. The corresponding method closely resembles the symmetry adaptation of states, apart from some minor modifications.
- *Wigner–Eckart theorem*. The application of the Wigner–Eckart theorem to problems which possess space group symmetry requires not only the, knowledge of space group symmetry adapted states and irreducible space group tensors, but also the corresponding space group Clebsch–Gordan coefficients. Due to the fact that space groups are nonsimple reducible groups, additional difficulties may arise because of the occurrence of nontrivial multiplicities.
- *Magnetic groups – co-representations*. The inclusion of the so-called "antiunitary" operators, for instance, the time reversal operator, requires, substantial extension of the concept of groups to the so-called magnetic groups, and their representations to co-representations.

Applications of Group Theory

Group theory was originally used extensively in evaluating the so-called energy band structures along special symmetry directions, that is for computing the eigenvalues corresponding to a special $\boldsymbol{k}$-vector from the three-dimensional Brillouin zone of a particular bulk system. The reasons were quite clear, namely the dimension of the secular matrix could be reduced substantially and therefore the computing times for inversion or diagonalization procedures were drastically shortened. Furthermore, respective compatibility relations guaranteed an unambiguous classification of energy bands along the rays in the Brillouin zone. The usefulness of such classifications was, in particular, evident in photoemission experiments and studies involving the Fermi surfaces.

With increasing speed of computers, in particular, with the innovation of architectures utilizing vectorizations, efforts at directly deriving the physical properties grew. This implied, for example, in most cases summations over all occupied states, that is, the use of Brillouin zone integrals by means of sophisticated sampling techniques. Thus, no longer were individual $\boldsymbol{k}$-states of interest, but so was the sum over all these states. Consequently, the incentive for using group theory, shifted from projecting particular states to theoretical means of reducing the computational efforts in evaluating Brillouin zone integrals.

Nowadays, most theoretical descriptions of problems in materials science, nanoscience, etc., are based on Kohn's density-functional theory, that is, on the use of effective (single-particle) Hamiltonians (H_{eff}, Kohn–Sham Hamiltonians). In particular, since the interest in solid-state physics shifted to systems with surfaces and interfaces, mostly group-theoretical means for two-dimensional translationally invariant systems are applied. It turned out that in two kinds of applications, namely in Brillouin zone integrations and in reduction of angular momentum coupling constants, the use of group theory is very valuable indeed. These two applications are discussed below and are meant to illustrate the contemporary use of group theory in solid-state physics.

Brillouin Zone Integrations

Taking the density of states as the simplest physical observable,

$$\begin{aligned} n(\varepsilon) &= -\frac{1}{\pi}\operatorname{Im}\mathcal{G}(\boldsymbol{r},\boldsymbol{r},\varepsilon) \\ &= -\frac{1}{\pi}|\mathcal{T}|^{-1}\sum_{\boldsymbol{k}}\operatorname{Im}\mathcal{G}(\boldsymbol{k};\varepsilon) \qquad [81] \end{aligned}$$

$$= -\frac{1}{\pi}|\mathcal{B}(\mathcal{T}^{\bigstar})|^{-1}\int \operatorname{Im}\mathcal{G}(\boldsymbol{k};\varepsilon)\,\mathrm{d}\boldsymbol{k} \qquad [82]$$

where the Green's function $\mathcal{G}(\boldsymbol{r},\boldsymbol{r},\varepsilon)$ is the diagonal configuration space representation of the resolvent of H_{eff}, $(z-H_{\text{eff}})^{-1}$ and $\mathcal{G}(\boldsymbol{k};\varepsilon)$ the corresponding lattice Fourier transformed Green's function, and $|\mathcal{T}|$ the order of the translational group. It should be noted that in eqn [81] a symmorphic space group is assumed and $|\mathcal{B}(\mathcal{T}^{\bigstar})|$ refers to the volume of the corresponding Brillouin zone $\mathcal{B}(\mathcal{T}^{\bigstar})$.

Suppose now that an appropriate matrix representation is used for $\mathcal{G}(\boldsymbol{r},\boldsymbol{r},\varepsilon)$ and $\mathcal{G}(\boldsymbol{k};\varepsilon)$, which, of course, depends largely on the applied quantum mechanical approach: Korringa–Kohn–Rostoker method (KKR), linear combination of muffin-tin orbitals (LMTO), or by using directly the effective Hamiltonain in the formulation of $n(\varepsilon)$ by means of a pseudopotential approach, or the linearized augmented plane wave method), then eqn [81] can be rewritten as

$$\begin{aligned} n(\varepsilon) &= -\frac{1}{\pi}|\mathcal{B}(\mathcal{T}^{\bigstar})|^{-1}\int \operatorname{Im}\operatorname{tr}\left[\underline{\mathcal{G}}(\boldsymbol{k};\varepsilon)\,\mathrm{d}\boldsymbol{k}\right] \\ &= -\frac{1}{\pi}|\mathcal{B}(\mathcal{T}^{\bigstar})|^{-1}\operatorname{Im}\operatorname{tr}\left[\int \underline{\mathcal{G}}(\boldsymbol{k};\varepsilon)\,\mathrm{d}\boldsymbol{k}\right] \qquad [83] \end{aligned}$$

where $\underline{\mathcal{G}}(\boldsymbol{k};\varepsilon)$ refers to the matrix representation of the Green's function and tr denotes the trace. In most of the methods mentioned above, the matrix representations are related in essence to angular momentum representations, that is they are expressed in terms of spherical harmonics.

Let $\mathcal{P}$ be the point group of the underlying (three-dimensional) lattice $\mathcal{T}$ and suppose $\underline{\mathbb{D}}(S)$ contains blockwise the irreducible representations of $S\in\mathcal{P}$. If $\Delta\mathcal{B}_E\ (\mathcal{T}^{\bigstar},\mathcal{P}) = \Delta\mathcal{B}_E$ as shorthand notation denotes an irreducible wedge of the Brillouin zone of volume $|\Delta\mathcal{B}_E(\mathcal{T}^{\bigstar},\mathcal{P})| = |\Delta\mathcal{B}_E|$, then any other wedge $\Delta\mathcal{B}_S(\mathcal{T}^{\bigstar},\mathcal{P})$ of the Brillouin zone is defined by

$$\Delta\mathcal{B}_S(\mathcal{T}^{\bigstar},\mathcal{P}) = \{S\boldsymbol{k}\,|\,\boldsymbol{k}\in\Delta\mathcal{B}_E|\} \qquad [84]$$

where $S\in\mathcal{P} : \mathcal{P}(\boldsymbol{k})$ is sufficient such that the union of the mutually disjoint subsets, (formally written as a sum here) can be written as follows:

$$\mathcal{B}(\mathcal{T}^{\bigstar}) = \sum_{S\in\mathcal{P}}\Delta\mathcal{B}_S(\mathcal{T}^{\bigstar},\mathcal{P}) \qquad [85]$$

It is seen from eqn [83] that

$$\underline{\mathcal{G}}(S^{-1}\boldsymbol{k};\varepsilon) = \underline{\mathbb{D}}(S)^{\dagger}\underline{\mathcal{G}}(\boldsymbol{k};\varepsilon)\underline{\mathbb{D}}(S) \qquad [86]$$

that is for a rotated $\boldsymbol{k}$ vector, the corresponding Green's function matrix is related to the original one by means of a similarly transformation.

The Brillouin zone integral in eqn [83] can, therefore, be expressed as

$$\frac{1}{|\mathscr{B}(\mathscr{T}^{\bigstar})|}\int \underline{\mathscr{G}}(\boldsymbol{k};\varepsilon)\,\mathrm{d}\boldsymbol{k} = \frac{1}{|\Delta\mathscr{B}_E|}\sum_{S\in\mathscr{P}}\mathbb{D}(S)^{\dagger}\left[\int_{\Delta\mathscr{B}_E}\underline{\mathscr{G}}(\boldsymbol{k};\varepsilon)\,\mathrm{d}\boldsymbol{k}\right]\mathbb{D}(S) \quad [87]$$

which, of course, results in a very large reduction of computing time, even if the order of the point group is rather small. In the case of a relativistic description, (Kohn–Sham–Dirac Hamiltonians) in the presence of an effective magnetization, the proper rotational invariance group of the corresponding Hamilton should be used. This is one of the rare cases when antiunitary operators are practically important. Of course, the same type of procedure as in eqn [87] can be applied when only two-dimensional translational symmetry is applied.

Reduction of Angular Momentum Coupling: The Point Group Symmetry of Single-Particle Densities

Let $\bar{\rho}(\boldsymbol{r})$ be the (shape-truncated) charge density in the cell $\mathscr{W}(\mathscr{T})$ (which may denote the corresponding Weigner–Seitz cell or any other type of primitive cell) centered at a particular origin $\boldsymbol{R}_0$ of given lattice or of any particular site in an arbitrary ensemble of scatterers,

$$\bar{\rho}(\boldsymbol{r}) = \rho(\boldsymbol{r})\sigma(\boldsymbol{r}) = \sum_L \bar{\rho}_L(r)Y_L^*(\hat{\boldsymbol{r}}) \quad [88]$$

where $\sigma(\boldsymbol{r})$ is the so-called shape function given by

$$\sigma(\boldsymbol{r}) = \left\{\begin{array}{ll} 1, & \boldsymbol{r}\in\mathscr{W}(\mathscr{T}) \\ 0, & \boldsymbol{r}\notin\mathscr{W}(\mathscr{T}) \end{array}\right\} \quad [89]$$

which usually is expanded into spherical harmonics as

$$\sigma(\boldsymbol{r}) = \sum_L \sigma_L(r)Y_L(\hat{\boldsymbol{r}}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sigma_{lm}(r)Y_{lm}(\hat{\boldsymbol{r}}) \quad [90]$$

where the expansion coefficients are determined by

$$\sigma_L(r) = \int_{\mathscr{W}(\mathscr{T})}\mathrm{d}\hat{\boldsymbol{r}}\,\sigma(\hat{\boldsymbol{r}})Y_L^*(\hat{\boldsymbol{r}}) \quad [91]$$

It should be noted that the $\bar{\rho}_L(r)$ are the coefficients of the shape-truncated charge density, that is,

$$\bar{\rho}_L(r) = \sum_{L',L''}C_{L''L}^{L'}\rho_{L'}(r)\sigma_{L''}(r) \quad [92]$$

$$C_{L,L''}^{L'} = \int \mathrm{d}\Omega Y_L(\Omega)Y_{L'}^*(\Omega)Y_{L''}(\Omega) \quad [93]$$

Table 1 Table of nonvanishing azimuthal quantum number terms

Plane	*n*	*Two-dimensional lattice*
(1 0 0)	4	Quadratic lattice
(1 1 0)	2	Rectangular lattice
(1 1 1)	3	Hexagonal lattice

Since $\rho(\boldsymbol{r})$ is a real function, in eqn [88] one needs to evaluate only terms $\bar{\rho}_{l,-m}(r)$ for $m\geqslant 0$ and then make use of the relation

$$\bar{\rho}_{l,-m}(r) = (-1)^m\bar{\rho}_{l,m}(r)^* \quad [94]$$

Furthermore, for a system with inversion symmetry with respect to $\boldsymbol{R}_0$,

$$\bar{\rho}_{lm}(r) = 0, \quad \forall l = \text{odd} \quad [95]$$

If the z-axis of the coordinate system is an n-fold rotational symmetry axis, the selection rule

$$m = \ldots, -n, 0, n, \ldots \quad [96]$$

applies, which in turn implies for example, that for simple cubic systems, there are only very few nonvanishing terms, namely those shown in **Table 1**.

Reviewing eqn [93], it is obvious that even very simple group theory can help substantially in reducing the number of terms to be evaluated. This kind of application is a very transparent example of reducing angular momentum coupling constants and is desperately needed whenever the so-called full-potential approaches are applied.

Much more advanced procedures are used whenever spectroscopic intensities are to be calculated, since then 3j-, 6j- or even higher coupling constants occur, depending, of course, on the type of excitation to be investigated.

See also: Group Theory.

PACS: 02.20; 03.65; 61.50

Further Reading

Altmann SL and Herzig P (1994) *Point-Group Theory Tables*. Oxford: Clarendon.

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Birman JL (1974) Theory of crystal space groups and infrared and Raman lattice processes of insulated crystals. In: Genzel L (ed.) *Handbuch der Physik*, XXV/26. New York: Springer.

Bouckaert LP, Smoluchowski R, and Wigner E (1936) Theory of Brillouin zones and symmetry properties of wave functions in crystals. *Physical Review* 50: 58–67.

Bradley CJ and Cracknell AP (1972) *The Mathematical Theory of Symmetry in Solids*. Oxford: Clarendon.

Cornwell JF (1984) *Group Theory in Physics*, *Volume 1: Fundamental Concepts and Applications in Molecular and Solid State*. New York: Academic.
Cracknell AP, Davies BL, Miller SC, and Love WF (1979, 1980) *Kronecker Product Tables*, *Volumes 1–4*. New York: Plenum.
Hahn T (2002) *International Tables for Crystallography, Volume A. Space Group Symmetry*. London: Kluwers.
Jansen L and Boon M (1976) *The Theory of Finite Group. Applications to Physics*. Amsterdam: North-Holland.
Kopsky V and Litvin DB (2002) *International Tables for Crystallography, Volume E. Subperiodic Groups*. London: Kluwers.
Landau LD and Lifshitz EM (1980) *Statistical Physics*, 3rd edn. New York: Pergamon.
Nye JF (1985) *Physical Properties of Crystals*. Oxford: Clarendon.
Toledano JC and Toledano P (1987) *The Landau Theory of Phase Transitions*. Singapore: World Scientific.

Nomenclature

$\mathscr{B}(\mathscr{T}^\star)$	Brillouin zone
$\mathbf{\Delta}\mathscr{B}(\mathscr{T}^\star, \mathscr{P})$	irreducible subwedge of $\mathscr{B}(\mathscr{T}^\star)$
$C^{L'}_{L,L''}$	Gaunt coefficient
$\boldsymbol{D}^{\boldsymbol{k}}(\boldsymbol{E}\|\boldsymbol{t}) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}}$	$\mathscr{T}(\mathscr{L})$-unirreps
$\boldsymbol{D}^{(\boldsymbol{k},\mathscr{E})\uparrow\mathscr{G}}(\mathscr{G})$	space group unirrep
$\boldsymbol{D}^{(\boldsymbol{k},\mathscr{E})\uparrow\mathscr{G}}$	matrix element of space group element
$(R\|\boldsymbol{w}_R + \boldsymbol{t})$	of space group unirrep
$\mathscr{G} =$	shorthand and extended space group
$\{\mathscr{T}, \mathscr{P}\|_o, \boldsymbol{w}_\mathscr{P}\}$	symbol
$\mathscr{G}(\boldsymbol{k}) \subseteq \mathscr{G}$	little group
$\boldsymbol{H}$	periodic one-particle Hamiltonian
$\mathscr{P}(\boldsymbol{k}) \sim \mathscr{G}$	little co-group
$(\boldsymbol{k})/\mathscr{T} \subseteq \mathscr{P}$	
$\mathscr{P}(\mathscr{T})$	point group of $\mathscr{T}(\mathscr{I})$
$\mathscr{P}(\mathscr{T})$	primitive cell of $\mathscr{T}(\mathscr{I})$
$(R\|\boldsymbol{w}_R + \boldsymbol{t})$	Wigner–Seitz symbol for space group element
$\boldsymbol{t} = (t_1, t_2, t_3)$	vector basis of Bravais vector lattice
$\mathscr{T}^\star$	reciprocal vector lattice
$\mathscr{T}(\mathscr{L})$	Bravais vector lattice
$\boldsymbol{w}_\mathscr{P} =$	set of fractional translations
$\{\boldsymbol{w}_R \| R \in \mathscr{P}\}$	
$\mathscr{W}(\mathscr{T})$	Wigner–Seitz cell of $\mathscr{T}(\mathscr{L})$
$\mathbb{Z}(z) \in \boldsymbol{SL}(3, \mathbb{Z})$	integral 3×3-matrix with determinant 1
$\mathbf{\Psi}^{\boldsymbol{k}}(\boldsymbol{x}) =$	Bloch function
$\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} w(\boldsymbol{k}, \boldsymbol{x})$	
$\mathbf{\Psi}^{\boldsymbol{k},n}_s$	Bloch eigenfunction of periodic Hamiltonian
$\mathbf{\Psi}^{(\boldsymbol{k},\xi)\uparrow\mathscr{G},n}_{\underline{R}_a}$	$\mathscr{G}$-symmetrized eigenfunctions of periodic Hamiltonian
$\Phi^{\boldsymbol{k}}(\boldsymbol{x}) =$	plane wave basis of Hilbert space $L^2(\mathscr{P}(\mathscr{T}))$
$\dfrac{1}{\sqrt{\lvert\mathscr{P}(\mathscr{T})\rvert}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{x}}$	
$\mathbf{\Omega}^{n,\boldsymbol{t}}_s$	Wannier functions

H

Hardness *See* Mechanical Properties: Anelasticity; Mechanical Properties: Creep; Mechanical Properties: Elastic Behavior; Mechanical Properties: Fatigue; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Mechanical Properties: Tensile Properties; Recovery, Recrystallization, and Grain Growth; Thin Films, Mechanical Behavior of.

Harmonic Generation Frequency Conversion

M Bellini, Istituto Nazionale di Ottica Applicata, Florence, Italy

Introduction

The ability to manipulate the frequency of laser light emission is essential in order to extend the range of possible spectroscopic investigations of matter in its different forms. Although the number of available laser lines is continuously growing, thanks to the introduction of new active materials operating in different spectral ranges, the extension toward the short-wavelength region of the electromagnetic spectrum would have been seriously limited without the advent of harmonic generation techniques, which rely on the nonlinear interaction between matter and intense laser light to generate new optical frequencies at integer multiples of the pump frequency.

With ordinary light sources, the electric fields associated to radiation are normally small and can be safely neglected when compared to the typical local fields (of atomic, molecular, or crystal origin) inside materials. The electrons in atoms or the nuclei in molecules just explore the bottom of their potential wells (where the shape is approximately parabolic) under the influence of the external optical field, and their motion is harmonic with the same frequency of the forcing light (**Figure 1**). The oscillating dipoles in the material are then the sources of reemitted radiation, whose spectral content is the same as the incident one. Many properties of light propagation in matter (such as refraction, reflection, dispersion, scattering, and birefringence) can be readily explained in this regime of linear optics. However, when the light field is so intense that it becomes a non-negligible fraction of the internal field, the linear approximation breaks down and the electrons or nuclei start exploring the anharmonic parts of their potential wells.

In this regime, their motion is no longer harmonic and can be expanded in a Fourier series containing frequencies at integer multiples of the forcing frequency. Such new frequencies are also present in the radiation field emitted by the material and constitute the harmonics of the fundamental driving field.

In this article, the basics of harmonic generation are presented concisely, with an attempt to illustrate the fundamental physical mechanisms which are responsible for the processes at play. A few examples are given: from the simple case of second harmonic generation (SHG), dating back to the dawn of the laser era, to the production of high-order harmonics in the extreme ultraviolet (XUV) and soft X-rays, which relies on the high-peak intensities available with the current pulsed laser systems.

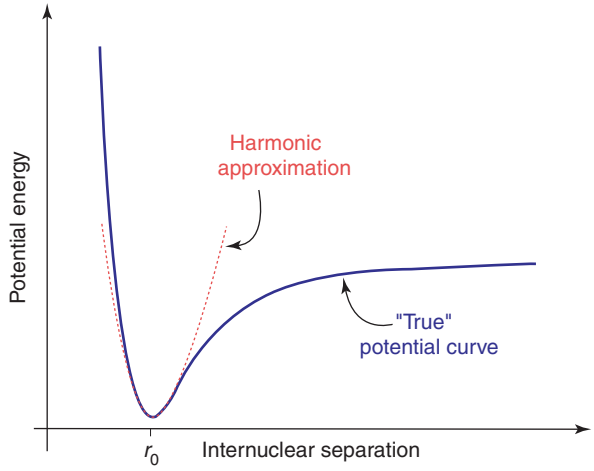


Figure 1 Potential curve illustrating the vibrational motion of nuclei in a molecule. The bottom of the curve can be well approximated by a parabolic potential and the motion is harmonic if only small amplitude oscillations are induced by an external field.

Nonlinear Polarization and Susceptibilities

Consider the case where light is sufficiently off resonance such that its interaction with matter does not induce any change in the population of the energy eigenstates of the material system (i.e., no absorption takes place), but only provokes a perturbation of the motion of electric charges within the atoms or molecules. In this nonresonant process, the material is transparent to the incident radiation, but the change in charge distribution due to the oscillating light field results in an induced electric dipole moment which, in turn, acts as a new source to emit secondary electromagnetic waves: the induced polarization $\boldsymbol{P}$ is the source term, which has to be introduced in Maxwell equations leading to the inhomogeneous wave equation:

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E} = \mu_0 \frac{\partial^2}{\partial t^2}\boldsymbol{P}$$

in order to describe the subsequent field evolution. The general form for the component at frequency ω of the polarization vector $\boldsymbol{P}$ can be written as

$$\boldsymbol{P}(\omega) = \sum_n \boldsymbol{P}^{(n)}(\omega)$$

where the nth-order polarization has the following expression:

$$\boldsymbol{P}^{(n)}(\omega) = \varepsilon_0 \chi^{(n)}(\omega_1, \omega_2, \ldots, \omega_n)\boldsymbol{E}(\omega_1)\boldsymbol{E}(\omega_2)\cdots\boldsymbol{E}(\omega_n)$$

and

$$\omega = \sum_{i=1}^{n} \omega_i$$

The quantities $\chi^{(n)}(\omega_1, \omega_2, \ldots, \omega_n)$ are known as the nonlinear optical susceptibilities of order n of the medium, and they are $(n+1)$th-order tensors that relate an n-fold product of vector components E_j to a certain component of the nth-order polarization $\boldsymbol{P}^{(n)}(\omega)$.

In the ordinary linear regime, the Fourier component of the polarization for a medium reduces to

$$\boldsymbol{P}^{(1)}(\omega) = \varepsilon_0 \chi^{(1)}(\omega)\boldsymbol{E}(\omega)$$

with the obvious meaning that, in this first-order approximation, the polarization at a given frequency is linearly proportional to the field component of the same frequency: a monochromatic incident light can only induce a monochromatic secondary wave at the same frequency. For anisotropic media, $\chi^{(1)}(\omega)$ is a 3×3-element second-order tensor, it reduces to a complex function of the frequency for isotropic materials.

In the general case, the expression for $\boldsymbol{P}^{(n)}(\omega)$ shows that, thanks to nth-order nonlinearities, it is possible to mix n waves at different frequencies to obtain an induced polarization and reemitted radiation at a different frequency: sums and differences among the frequencies of the incoming beams can be performed this way. In the particular case where only one field at frequency ω enters the material, a susceptibility of order n can produce new radiation at $n\omega$, the nth harmonic of the incident field.

The ratio of two successive orders of polarization can be roughly estimated by

$$\left|\frac{P^{(n+1)}}{P^{(n)}}\right| = \frac{|\chi^{(n+1)}||E|}{|\chi^{(n)}|} \approx \left|\frac{E}{E_{\mathrm{m}}}\right|$$

where E is the amplitude of the applied optical field, and E_{m} is the characteristic value of the internal electric field in the material. If the major mechanism contributing to the nonlinear polarization is the distortion of the electronic cloud of an atomic or molecular system (as in the case of optical harmonic generation), then E_{m} is the average magnitude of the binding field experienced by the electrons ($\sim 10^{11}\,\mathrm{Vm}^{-1}$ for the hydrogen atom). For ordinary light sources, the ratio $|E/E_{\mathrm{m}}|$ is normally so small that nonlinear polarization terms can be safely neglected, but when a laser light of high spectral intensity is used, the laser field may become a significant fraction of the material field, and the second, third, and higher nonlinear order terms start playing an important role, causing part of the incident light to be converted in the corresponding harmonic. From the above expression, one can also deduce that the nth-order polarization, hence the field emitted at $n\omega$, is proportional to the nth power of the ratio between E and E_{m} (normally much smaller than 1), so that the probability of generating higher harmonics is expected to drop exponentially with the corresponding order in this perturbative approach.

From the fact that the susceptibility tensors must remain unchanged upon the symmetry operations allowed for the medium, one can simply conclude that for all isotropic media and centrosymmetric crystals, the even-order susceptibilities ($\chi^{(2)}, \chi^{(4)}, \ldots, \chi^{(2n)}$) must be zero. Consequently, the generation of even-order harmonics is only possible in materials which lack inversion symmetry, while odd-order nonlinear effects can be exhibited by any media, although at different degrees of efficiency.

If the field amplitudes do not change much in a propagation length shorter than a wavelength (the so-called slowly varying amplitude approximation) and if nonlinearities are not too strong (so that the field amplitude at the pump frequency stays almost

constant throughout the interaction), then one can find a simple expression for the wave equation of the plane-wave field and the induced nonlinear polarization P^{NL} oscillating at the new frequency ω':

$$\frac{\partial}{\partial z} E(z,\omega') = -\frac{\mathrm{i}\mu_0 \omega'^2}{2k'} P^{\mathrm{NL}}(z,\omega')\ \mathrm{e}^{\mathrm{i}\Delta k z}$$

where one assumes linear polarizations (to reduce to a scalar equation) and collinear propagation along the z direction. The phase term $\mathrm{e}^{\mathrm{i}\Delta k z}$ accounts for the phase drift between the induced polarization and the generated electric field at ω', and depends on the mismatch in their propagation vectors

$$\Delta k = k' - k_{\mathrm{p}}$$

where $k_{\mathrm{P}} = \sum_i \omega_i n(\omega_i)/c$ and $k' = \omega' n(\omega')/c$, with $n(\omega)$ the material refraction index at frequency ω.

In the case of harmonic generation of order q, this factor reduces to

$$\Delta k = k(q\omega) - qk(\omega)$$

Second Harmonic Generation

As a simple example, consider the case of second-order nonlinearity for the generation of radiation at twice the frequency of the incoming light. As seen before, symmetry considerations impose the process to take place in noncentrosymmetric crystals only.

The polarization vector

$$\boldsymbol{P}^{(2)}(2\omega) = \varepsilon_0 \chi^{(2)}(\omega,\omega)\boldsymbol{E}_1(\omega)\boldsymbol{E}_1(\omega)$$

is the source term that has to be inserted in the wave equation and leads to the following expression for the variation of the field amplitude at the second harmonic $E_2(z)$:

$$\frac{\partial}{\partial z} E_2(z) = -\frac{\mathrm{i}\omega d}{cn(2\omega)} E_1^2(z)\ \mathrm{e}^{\mathrm{i}\Delta k z}$$

where d corresponds to the component of the $\chi^{(2)}(\omega,\omega)$ tensor appropriate for the particular geometry of the laser-crystal interaction and

$$\Delta k = k(2\omega) - 2k(\omega) = \frac{4\pi}{\lambda}\left[n(2\omega) - n(\omega)\right]$$

is the phase mismatch factor, which takes into account the different phase velocities of the waves at ω and 2ω. If a negligible absorption and little production of the wave at 2ω are assumed, so that the pump field amplitude E_1 at frequency ω can be considered as almost constant, then the coupled wave equation can be integrated straightforwardly:

$$E_2(z) = -\frac{\mathrm{i}\omega d}{cn(2\omega)} E_1^2(0) \int \mathrm{e}^{\mathrm{i}\Delta k z}\,\mathrm{d}z$$

where the integration is done over the length of the medium between 0 and L. The integration yields, assuming that $E_2(0) = 0$:

$$E_2(L) = -\frac{\omega d}{cn(2\omega)} E_1^2(0)\,\frac{\mathrm{e}^{\mathrm{i}\Delta k L} - 1}{\Delta k}$$

and the output intensity of the second harmonic is proportional to

$$I_2(L) \propto d^2 I_1^2(0) L^2\,\frac{\sin^2(\Delta k L/2)}{(\Delta k L/2)^2}$$

Some simple observations can be made about this result: the total amount of generated light at the second harmonic is proportional to the square of the intensity at the fundamental frequency; the efficiency is proportional to $|\chi^{(2)}|^2$ and to L^2 but also contains a sinc^2 term involving L and Δk.

The dependence of the emitted intensity at the harmonic frequency on the propagation distance and phase-mismatch is quite general and does not depend on the particular harmonic order considered. As seen above, passing to higher harmonic orders normally involves a rapidly decreasing value of the nonlinear coefficient d and a different power dependence on the incident intensity I_1, but both the L^2 and the sinc^2 terms remain unchanged and play an important role in the harmonic generation of any order.

Only in the optimal condition of phase-matching, with $\Delta k = 0$, will the efficiency really depend quadratically on the medium length, and longer crystals will produce more second harmonic. Note, however, that in these conditions the coupling between the fundamental wave and the second harmonic may become so strong that the depletion of E_1 while it propagates along z has to be taken into account, and some of the assumptions made above are no longer valid. In this so-called depleted-pump regime, a set of coupled equations for the pump field at ω and for the second harmonic field at 2ω has to be solved, and one finds that a complete transfer of the pump energy into the second harmonic field can be achieved.

Phase-Matching

In the general case where $\Delta k \neq 0$, the intensity of the second harmonic field generated in a nonlinear

crystal varies periodically along the z-axis with a period $2\pi/\Delta k$, and at a propagation distance of

$$L_C = \frac{\pi}{\Delta k} = \frac{\lambda}{4}\frac{1}{n(2\omega) - n(\omega)}$$

it reaches the first maximum, at a peak efficiency inversely proportional to Δk^2 (**Figure 2**).

The physical explanation for this effect is connected to the fact that, due to dispersion, the incident wave at frequency ω and the second harmonic wave at 2ω propagate in the medium with different phase velocities ($v_1 = c/n(\omega)$ for the fundamental and $v_2 = c/n(2\omega)$ for the second harmonic). Consequently, the two waves fall out of step with increasing distance in the crystal and the second harmonic waves generated at different places do no longer superpose in phase: after a distance L_C, destructive interference starts developing and it becomes complete, with the full disappearance of the wave at 2ω, at a distance of $2L_C$. L_C is the coherence length for SHG and defines the medium thickness for optimized conversion efficiency under the conditions of $\Delta k \neq 0$. Large Δk values lead to small SHG peak intensities and rapid variations with the distance, so that efficiency is severely limited for crystals longer than L_C.

The condition of $\Delta k = 0$ is referred to as the phase-matching condition and is achieved when $n(2\omega) = n(\omega)$. For ordinary waves in a medium with positive dispersion:

$$n(2\omega) > n(\omega)$$

so one always gets $\Delta k \neq 0$, unless special conditions are met.

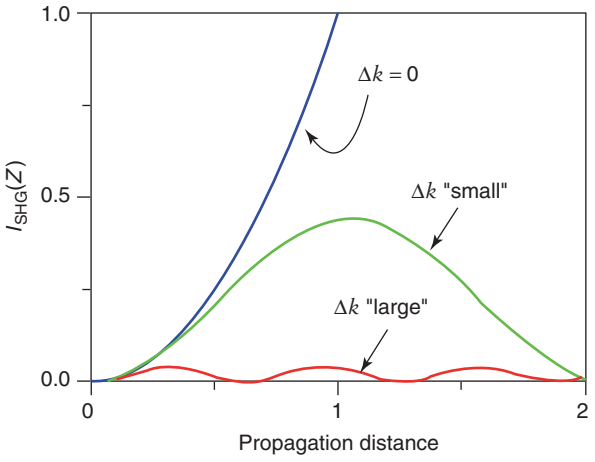


Figure 2 Intensity of the second harmonic field vs. propagation distance in the nonlinear medium. A quadratic growth is obtained for perfect phase-matching conditions ($\Delta k = 0$), while oscillations of decreasing amplitude and period arise for increasing Δk.

In a birefringent material, different polarizations see different refractive indices. "Ordinary" and "extraordinary" refractive indices can be different by up to ~ 0.1 for SHG crystals, so one can satisfy the phase-matching condition by choosing the extraordinary polarization for one wave and the ordinary polarization for the other. For a negative uniaxial crystal ($n_o > n_e$), one can achieve phase-matching at a particular frequency ω if the relation $n_o(\omega) = n_e(2\omega)$ is satisfied (**Figure 3**).

Furthermore, the extraordinary index n_e also depends on the propagation angle θ with respect to the optical axis of the crystal, so one can phase-match the fundamental and second harmonic waves at different ω just by adjusting the angle of incidence on the crystal: the phase-matching condition in this case becomes $n_o(\omega) = n_e(2\omega, \theta_m)$, where θ_m is the phase-matching angle (**Figure 4**). A geometric representation of this situation is obtained by plotting the index curves for the ordinary and extraordinary waves at the fundamental and second harmonic frequencies: the refraction index n_e or n_o for each wave can be found at the crossing of the curves with the k-vector. In the negatively uniaxial crystal case, the point of intersection between the n_o circle at ω with the n_e ellipse at 2ω indicates the phase-matching angle θ_m.

Note that in the quantum mechanical view of SHG, two photons of energy $\hbar\omega$ are absorbed during the excitation of the medium to a virtual state, and a

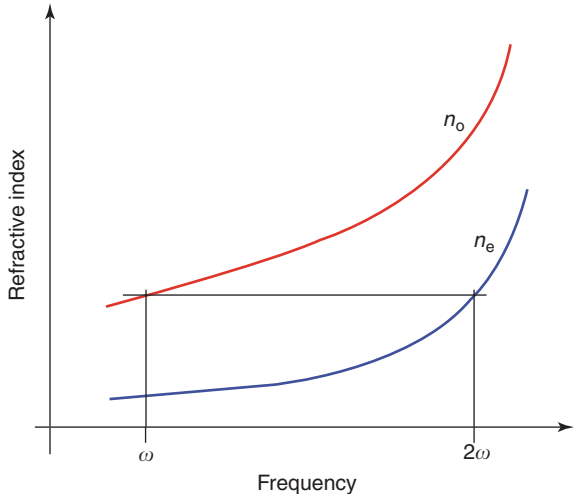


Figure 3 Phase-matching condition for a negative uniaxial crystal occurring for a particular frequency ω. In reality, the two curves for n_o and n_e represent the maximum and the minimum attainable index of refraction in the crystal, and the whole range in between the two curves can be covered by the extraordinary index n_e by changing the angle between the propagation vector and the optical axis of the medium. So, by choosing the appropriate angle, phase-matching can be obtained for different frequencies.

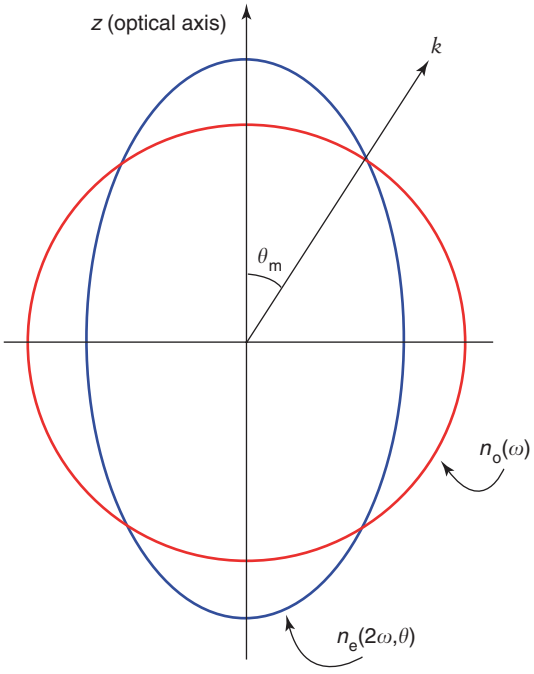


Figure 4 Geometric construction for the determination of the phase-matching angle θ_m in a negative uniaxial crystal.

single photon with an energy $2\hbar\omega$ is immediately re-emitted when the medium de-excites toward the ground state. In this framework, the phase-matching condition has the simple meaning of photon momentum conservation

$$\boldsymbol{k}_1 + \boldsymbol{k}_1 = \boldsymbol{k}_2$$

between the initial and the final states of the system.

Although it was discussed in the particular case of SHG, phase-matching is a much more general issue and is always present in all nonlinear optics phenomena where waves at new frequencies are to be generated with the highest possible efficiency. In the most general case of n waves with wave vectors $\boldsymbol{k}_i$ participating in the nonlinear process, one has to make sure that the total photon momentum is conserved in the interaction by making

$$\sum_{1}^{n} \boldsymbol{k}_i = 0$$

Quasi Phase-Matching

As seen before, when perfect phase-matching is not achieved, the fundamental and second harmonic fields slip out of phase after a propagation distance L_C and the conversion efficiency oscillates without growing monotonically with L^2 as it should do in the case of $\Delta k = 0$ (**Figure 5**). An alternative idea to improve the

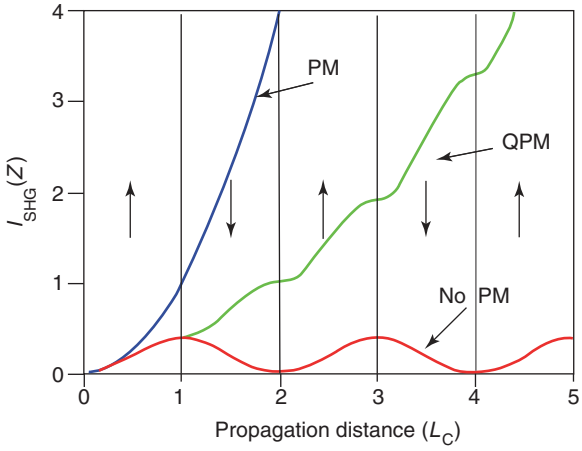


Figure 5 Intensity of the second harmonic field vs. propagation distance in the conditions of quasi-phase-matching. The sign of the nonlinear coefficient of the crystal is reversed for every coherence length L_C in order to get a monotonous increase of the second harmonic intensity. The perfect phase-matched condition and the mismatched situation are also shown.

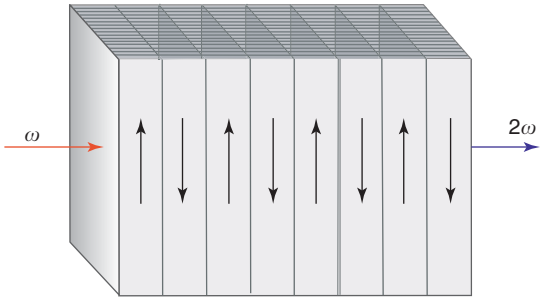


Figure 6 Schematic of SHG in a periodically poled crystal with several layers of alternatively oriented domains.

conversion efficiency is to adjust the phase of the nonlinear polarization appropriately after each coherence length: the best conditions are obtained by changing the sign of the nonlinear coefficient (and thus the sign of the polarization) with a periodicity L_C.

Under these circumstances, the nonlinear intensity starts growing monotonically, although less rapidly than in the case of perfect phase-matching, and long interaction lengths can be used to boost the conversion.

The condition of quasi-phase-matching (QPM) can be achieved in a so-called periodically-poled material, generally a ferroelectric crystal ($LiNbO_3$ is the currently most used one) where regions of periodically reversed polarization domains are permanently written by applying spatially shaped electric fields (**Figure 6**).

If technological limitations do not allow one to realize the right spacing between the inverted domains (it is currently difficult to obtain poling periods smaller than $\sim 5\,\mu m$), a higher order QPM can be

obtained, although with a lower conversion efficiency with increasing order, by flipping the nonlinear coefficients every 3 (third-order QPM), 5 (fifth-order QPM), or any odd number of coherence lengths.

High-Order Harmonic Generation

Current amplified pulsed laser systems can reach peak intensities $\sim 10^{15}$–$10^{20}\,\mathrm{W\,cm^{-2}}$ and, in these situations, the perturbative approach used above to model the nonlinear response of materials to radiation is no longer adequate. As a matter of fact, at a peak intensity of $3.5 \times 10^{16}\,\mathrm{W\,cm^{-2}}$, the electric field of the laser corresponds to the field which binds an electron to the proton in the ground state of a hydrogen atom: the assumption that successive terms in the expansion of the medium polarization in powers of the incident field get progressively smaller with the order cannot hold.

The most impressive consequence of the breakdown of the perturbative approach is the deviation from the expected exponential decay of the intensity of successive harmonic orders (**Figure 7**). The appearance of a so-called plateau, a region of the spectrum where several harmonics are generated with almost constant efficiency, is the characteristic feature of high-order harmonic generation (HHG) in gases.

Here, short and intense laser pulses are typically focused in a pulsed gas jet to produce coherent radiation in the extreme ultraviolet (XUV) and soft X-rays, a wavelength range (1–100 nm) where the lack of coherent sources has greatly limited the possibility of spectroscopic investigation so far (**Figure 8**).

As a consequence of the isotropy of the gaseous medium, only odd-order harmonics are generated, and one finds that the region of efficient production of high harmonics is limited by a so-called cutoff. The most energetic photons generated in the process possess a cutoff energy E_{cutoff}, which is experimentally found to follow the simple law:

$$E_{\mathrm{cutoff}} \approx I_{\mathrm{P}} + 3.2U_{\mathrm{P}}$$

where I_{P} is the ionization potential of the atoms in the gas jet, and U_{P} is the ponderomotive energy, which corresponds to the mean kinetic energy acquired by an electron in one cycle of oscillation in the laser field:

$$U_{\mathrm{P}} = \frac{e^2E^2}{4m\omega^2}$$

with e and m the electron charge and mass, respectively, E the amplitude of the laser field and ω its frequency.

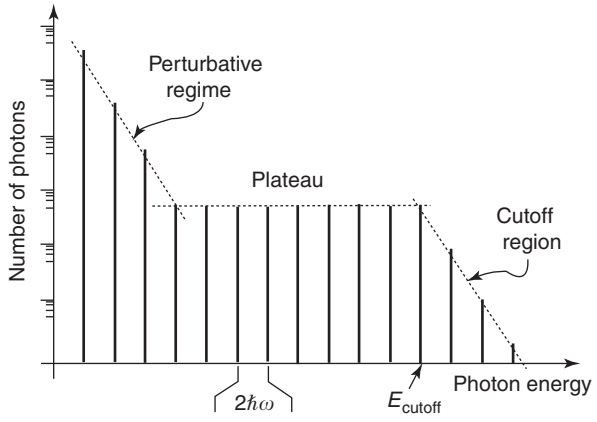


Figure 7 Schematic view of the typical behavior of high-order harmonic spectra. Only odd-order harmonic photons are generated with successive harmonics separated by twice the energy of the pump laser. The first harmonics follow the predictions of the perturbative approach, with intensities exponentially decreasing with the order. Then a plateau is formed and extends up to E_{cutoff}.

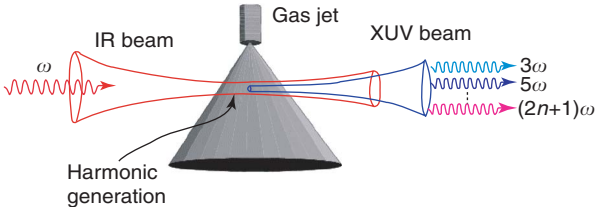


Figure 8 HHG in a gas jet.

It is evident that high laser intensities and high ionization potentials are required in order to reach very short wavelengths. Indeed, rare gas atoms and ultrashort pulses of very high peak intensity are normally used to generate high harmonics. A further advantage of using hard-to-ionize gases and short pulses is given by the requirement that neutral atoms survive long enough to experience the full peak pulse intensity instead of being completely ionized at the leading edge of the pulse itself. If $W(I)$ is the atomic ionization rate and τ_{p} the pulse duration, then the maximum effective intensity useful to produce harmonic photons is the so-called saturation intensity I_{s}, approximately given by $W(I_{\mathrm{s}})\tau_{\mathrm{p}} = 1$; $W(I)$ being a monotonously increasing function of I, higher saturation intensities can be reached with shorter laser pulses and, of course, with light noble gases having higher ionization potentials.

Three-Step Model

The standard picture of the HHG process can be easily visualized from a semiclassical perspective in a single-atom approach: in every half optical cycle of

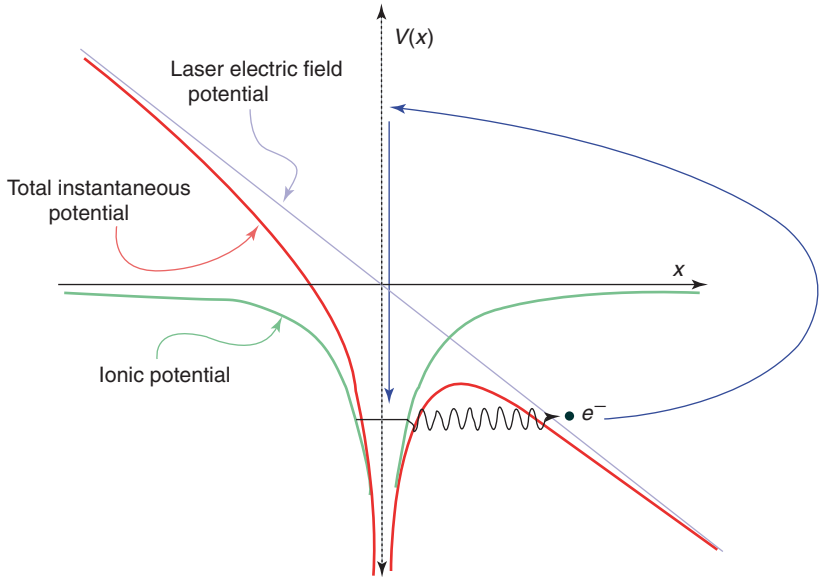


Figure 9 The three-step model for HHG. Electrons are first tunnel-ionized by the intense laser field, then they are accelerated during an optical cycle and finally recombine with the ion, thus releasing the accumulated kinetic energy in the form of harmonic photons.

the laser pulse, electrons undergo tunnel ionization through the potential barrier formed by the atomic potential and by the electric field potential of the laser itself; after being accelerated in the ionization continuum by the field, they may come back to the ion core and recombine to emit harmonic photons that release the accumulated kinetic and ionization energy (**Figure 9**).

Some of the most relevant features of harmonic radiation can be simply obtained by using this single-atom picture and solving the classical equations of motion for an electron, freed with zero velocity at different moments in an oscillating electric field. Depending on the phase of the field at the moment of ionization, the electron can either escape from the parent ion or oscillate in the field, and come back to the original position after a fraction of the optical period T. Only electrons ionized within $T/4$ after each field maximum contribute to the harmonic emission by recombining with the ion and converting their kinetic and potential energy into photons (**Figure 10**).

The experimentally found cutoff law can be simply obtained by noting that the maximum return kinetic energy (brought back by electrons which escape slightly after each field maximum every half optical cycle) corresponds to $\sim 3.17\ U_p$ (**Figure 11**). For the generation of harmonics with photon energies below the cutoff limit, there are always two different trajectories (corresponding to two different release times within each half cycle) that the electron can follow in the continuum, such that it returns to the ion core with the correct energy. These two trajectories, although giving rise to photons with the same energy, may correspond to quite different permanence times of the electrons under the influence of the laser field, and some of the spatial and temporal characteristics (phase-matching relations, defocusing, and frequency modulation) of the emitted harmonic radiation will thus be strongly affected.

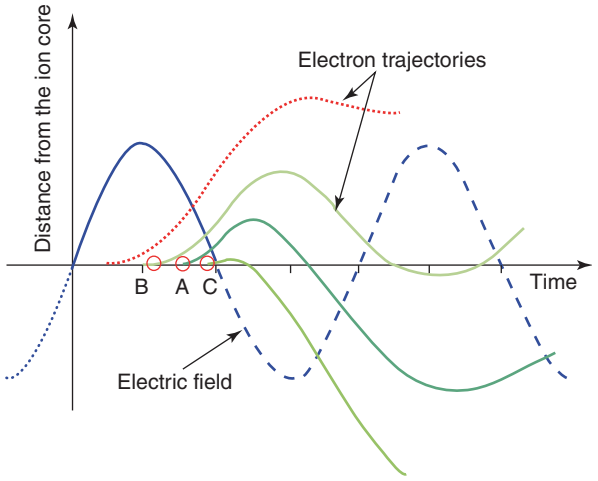


Figure 10 Schematic representation of the main electron trajectories involved in the HHG process (only the first half optical cycle is considered for simplicity, but things repeat with a $T/2$ periodicity). Electrons which are ionized when the absolute value of laser field amplitude is growing just escape from the ion and do not contribute to the process, while electrons released after the peak can come back to the core position and recombine to emit harmonic photons. Electrons ionized in A return with the maximum kinetic energy and generate the most energetic photons in the cutoff region. Electrons released in B and C follow different trajectories in the continuum but return with the same kinetic energy, thus both contribute to the generation of the same harmonic.

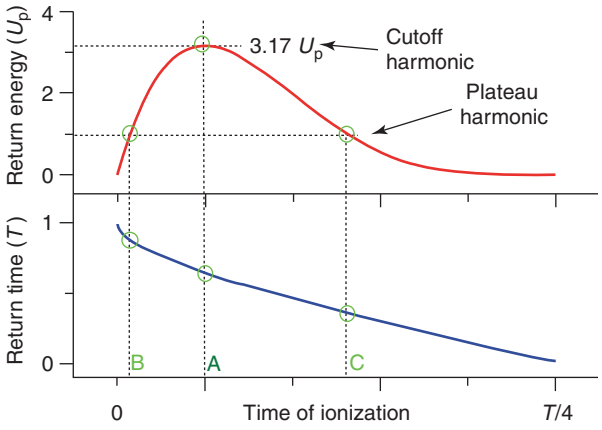


Figure 11 Upper curve: plot of the return kinetic energy for electrons ionized at different times during the optical cycle. The origin of the time axis is fixed at the peak of the field amplitude for each half optical cycle. A single trajectory, corresponding to electrons emitted in A, is responsible for the emission of a cutoff harmonic, while two distinct trajectories (originating in B and C contribute to the generation of a harmonic in the plateau. Lower curve: time (here indicated as the "return time") that the electrons spend oscillating in the continuum before re-combining with the ion. Plateau harmonics mainly come from two classes of electrons with distinct return times.

Harmonic emission during a single pump pulse is thus seen to be composed of a train of XUV photon bunches, essentially released twice every optical cycle of the laser field. This temporal structure translates into the characteristic spectral structure of high-order harmonics, with peaks corresponding to the odd orders, separated by twice the laser frequency.

The semiclassical model also offers a simple explanation for other experimentally observed characteristics of harmonic radiation. A linear polarization of the laser is needed to make sure that the electron oscillating in the field may "hit" the ion and recombine with it when it comes back after the oscillation in the continuum; experiments have shown that just a small degree of ellipticity is sufficient to completely inhibit the generation of harmonics. It is also clear that, although the expressions for E_{cutoff} and U_p seem to favor long-wavelength pump pulses for the generation of high harmonics, actually, the longer time that it takes for the electron to return to the ion (proportional to the pump laser wavelength) implies a significant spreading of the electron wave packet during its oscillation and a final lower overlap with the ion wave function, which in turns causes the probability of recombination to decrease.

Phase-Matching in High-Order Harmonic Generation

The three-step model is a good way to understand the single-atom response of the gas medium to the intense electric fields of the laser. However, when the global response of the system is considered, propagation effects have to be included. Phase-matching is again an essential condition for efficient generation of high-order harmonics to take place. In this case, the fundamental and the harmonic field can fall out of phase for several reasons: besides the usual dispersion in the neutral gaseous medium (which may, however, contain both linear and nonlinear contributions, due to the high intensities involved), the geometrical dephasing due to the focusing of the Gaussian laser beam can introduce an additional phase shift, as well as the dispersion introduced by the generated free electrons, and the so-called dipole-phase variation. In order to achieve sufficiently high conversion efficiency, the phase shift between the polarization of the qth harmonic and the generated harmonic field has to be kept as constant as possible over all the interaction region.

As for the dispersion by the neutral medium, the corresponding phase shift on a propagation distance z has the usual form

$$\Delta\varphi_{\text{ND}} = \frac{2\pi q z}{\lambda}\left(n(q\omega) - n(\omega)\right)$$

and the refraction indices at the two frequencies clearly depend both on the gas density in the jet and on the fraction of neutral atoms left after the partial ionization by the laser pulse.

The dispersion caused by free electrons produced in the interaction becomes important as soon as the intensity approaches the saturation intensity I_s and has to be accounted for by introducing the index of refraction

$$n_{\text{el}}(\omega) = \sqrt{1 - \frac{\omega_P^2}{\omega^2}}$$

where $\omega_P(z,t)$ is the so-called plasma frequency, which is proportional to the local free-electron density: the corresponding phase shift is approximately given by

$$\Delta\varphi_{\text{el}} \approx \frac{z\omega_P^2}{2cq\omega}\left(q^2 - 1\right)$$

The on-axis phase shift across the focus of a Gaussian beam is $\varphi(z) = \tan^{-1}(2z/b)$, where b is the laser confocal parameter. The phase shift due to this geometrical effect is then

$$\Delta\varphi_{\text{G}} = (1-q)\tan^{-1}\left(\frac{2z}{b}\right)$$

Finally, one finds that the laser-induced polarization has an intrinsic intensity-dependent phase which arises

from the phase of the wave function of the electrons which generate a particular harmonic. The two main electronic trajectories contributing to the dipole moment of a given plateau harmonic induce a phase which is proportional to the intensity of the laser driving field, and to the amount of time the electron has spent in the continuum. For each trajectory (labeled by j), this leads to an intensity-dependent variation of the phase

$$\Delta\varphi_{\mathrm{D},j} = -\alpha_j I(z,t)$$

which sum up to a total dipole phase:

$$\Delta\varphi_{\mathrm{D}} = \sum_j \Delta\varphi_{\mathrm{D},j}$$

All these contributions to the phase slip between the induced polarization and the generated harmonic field have to be combined to give a total phase-mismatch

$$\Delta\varphi = \Delta\varphi_{\mathrm{ND}} + \Delta\varphi_{\mathrm{el}} + \Delta\varphi_{\mathrm{G}} + \Delta\varphi_{\mathrm{D}}$$

which should not vary much in the medium length. Unfortunately, most of the above terms have a complicated dependence on the detailed spatio-temporal dynamics of the atom–laser interaction, and the task of keeping $\Delta\varphi$ fixed is normally a very difficult one, unless special simplifications can be made in order to neglect one or more of the contributions to the total phase-mismatch. Using collimated laser beams or propagation in hollow fibers can null the geometrical phase shift, while keeping a low ionization level can eliminate the free-electron contribution to dispersion. When the experimental conditions are achieved so that positive and negative terms partially cancel out in the expression for $\Delta\varphi$, a strong enhancement in the conversion efficiency of a particular high-order harmonic can be obtained.

See also: Nonlinear Optics; Optical Sum Rules and Kramers–Kronig Relations.

PACS: 42.65.Ky; 42.65. – k; 42.65.An; 42.79.Nv

Further Reading

Agrawal GP (1989) *Nonlinear Fiber Optics*. New York: Academic Press.

Bloembergen N (1965) *Nonlinear Optics*. New York: Benjamin.

Salières P (2001) High-order harmonic generation. In: Batani D, Joachain CJ, Martellucci S, and Chester AN (eds.) *Atoms, Solids, and Plasmas in Super-Intense Laser Fields*, pp. 83–97. New York: Kluwer Academic/Plenum.

Shen YR (1984) *The Principles of Nonlinear Optics*. New York: Wiley.

Hartree and Hartree–Fock Methods in Electronic Structure

G C Strinati, Università degli Studi di Camerino, Camerino (MC), Italy

Introduction

The Hartree and Hartree–Fock methods represent good starting points for the calculation of eigenvalues and eigenfunctions of complex atoms, molecules, and solids. They provide a conceptually simple and practical prescription to deal with many-body interactions in these systems, although their implementation may not be sufficient to get an appreciable agreement with experimental data under many circumstances.

The success of these methods rests essentially on the very reasonable physical intuition on which they are based, and also their mathematical foundation is sound. In addition, in spite of their limitations (to cure which more sophisticated methods have been devised), the Hartree and Hartree–Fock methods have also been useful in establishing calculation procedures (such as the very idea of "self-consistency") that have been later applied to other widely used methods (such as the local density approximation to the density-functional theory). Besides, the "mean-field" approach (with which the Hartree–Fock method is sometimes synonymous) is regarded as being nonperturbative, in the sense that it does not treat the interaction term of the many-body Hamiltonian as a perturbation to the kinetic term. This fact acquires importance both for atoms and condensed matter systems at the relevant densities, whereby the average kinetic and interaction energies are of the same order.

In a sense, the Hartree and Hartree–Fock methods can be regarded as reference methods for many-body calculations in complex systems. In many cases, in fact, their utility as reference methods is far more appreciated than the practical results which can be directly produced by the methods themselves. Finally, extensions of the methods have been successfully

applied to include broken-symmetry situations (for instance, in magnetic and superconducting systems).

Need for Nonperturbative Approximations

For atoms with more than one electron, the Schrödinger equation cannot be solved in practice without recourse to numerical approximations. It is then important to develop approximation methods for calculating the energy eigenvalues and wave functions of stationary states.

The Hamiltonian H of a system of N electrons in atoms, molecules, and solids contains two parts, namely, the kinetic energy of the particles plus the interaction with the nuclei (H_1), and their mutual interaction energy (H_2):

$$H = H_1 + H_2 \qquad [1]$$

with

$$H_1 = \sum_{i=1}^{N} h(q_i) \quad \text{and} \quad H_2 = \frac{1}{2}\sum_{i\neq j}^{N} v(q_i, q_j) \qquad [2]$$

Here, the indices (i, j) label the particles, $q = (\boldsymbol{r}, \sigma)$ is a collective coordinate including spatial position $\boldsymbol{r}$ and spin σ, $h(q)$ is the Hamiltonian relative to each particle, and $v(q, q')$ is the mutual interaction energy between two particles (which is usually taken to be the repulsive Coulomb potential).

Perturbative methods are normally constructed to deal with (at least part of) the interaction energy as the perturbation. It turns out, however, that the perturbation theory in the interaction part of the Hamiltonian is nonanalytic. This is also expected on physical grounds, from the fact that the behavior of the system should completely change if the interaction $v(q, q')$ changes its sign. It is for these reasons that nonperturbative methods (such as the Hartree and Hartree–Fock methods) acquire great value. Furthermore, for a neutral atom (where N equals the atomic number Z) with large enough Z, the Z dependence of the two-body Hamiltonian H_2 (namely, $ZN(N-1)/2 \approx Z^3$) becomes comparable with the corresponding dependence of the one-body Hamiltonian H_1 (namely, $Z^2N = Z^3$). This is an additional indication that the perturbation theory in the two-body Hamiltonian H_2 is bound to fail even in practice, at least for large enough systems.

Note that, in the absence of the electron–electron interaction H_2, the many-body problem decouples into a one-body problem for each electron in the external potential due to the nuclei. The Hartree and Hartree–Fock methods succeed specifically in reducing the many-body problem to a one-body problem, while including at least part of the effects of the electron–electron interaction.

Hartree Self-Consistent Field

The original motivation of the Hartree and Hartree–Fock methods was to obtain an average (central) field in atomic problems. A similar requirement is also present in crystals for the one-electron periodic potential, with the purpose of calculating band structures and obtaining the corresponding Bloch wave functions.

Hartree's idea (originally devised on physical considerations) was to reduce the many-electron Schrödinger equation

$$H\Phi(q_1, q_2, \ldots, q_N) = E\Phi(q_1, q_2, \ldots, q_N) \qquad [3]$$

(E being the total energy) to the equation for only one electron in its own central field, obtained from the nuclear potential and the wave functions of the other electrons. The one-electron wave functions obtained in this way are then made consistent with the fields from which they are calculated. The resulting self-consistent field thus represents, as closely as possible, the average effect of the other electrons on one electron.

The Hartree method rests on writing the many-electron wave function of eqn [3] in the form

$$\Phi(q_1, q_2, \ldots, q_N) = u_{\alpha_1}(q_1)u_{\alpha_2}(q_2)\cdots u_{\alpha_N}(q_N) \qquad [4]$$

Here, the single-electron wave functions $u_{\alpha_i}(q)(i = 1, \ldots, N)$ are assumed to be different from each other and normalized as follows:

$$\langle u_{\alpha_i}(q)|u_{\alpha_i}(q)\rangle = 1 \qquad [5]$$

with the Dirac's convention for the matrix elements. Note that the requirement for the wave function to be antisymmetric in all the electrons is not taken into account in the Hartree wave function [4].

The single-electron wave functions $u_{\alpha_i}(q)$ can be obtained by an optimization procedure based on the variational principle (as recognized only later by Slater and Fock). One thus calculates the average values

$$\langle \Phi|H_1|\Phi\rangle = \sum_{i=1}^{N} \langle u_{\alpha_i}(q)|h(q)|u_{\alpha_i}(q)\rangle \qquad [6]$$

and

$$\langle \Phi|H_2|\Phi\rangle = \frac{1}{2}\sum_{i\neq j}^{N} \langle u_{\alpha_i}(q)u_{\alpha_j}(q')|v(q, q')|u_{\alpha_i}(q)u_{\alpha_j}(q')\rangle \qquad [7]$$

in terms of the wave function [4]. Minimization of $\langle\Phi|H_1|\Phi\rangle + \langle\Phi|H_2|\Phi\rangle$ with respect to alternative $u_\alpha(q)$ and subject to the normalization constraint [5] then yields the Hartree differential equation(s)

$$h(q)u_{\alpha_i}(q) + \sum_{j\neq i}\langle u_{\alpha_j}(q')|v(q,q')|u_{\alpha_j}(q')\rangle u_{\alpha_i}(q) = \varepsilon_i u_{\alpha_i}(q) \qquad [8]$$

where ε_i are Lagrange multipliers (with the dimension of energy) introduced to enforce the constraints [5].

Equation [8] has the form of the Schrödinger equation for one electron moving under the influence of the nuclear potential and the density distribution $\sum_{j\neq i} u^*_{\alpha_j}(q')u_{\alpha_j}(q')$ of the remaining electrons. This density distribution has to be obtained self-consistently by solving the same equation. Implementation of the method thus requires recourse to successive approximations. The ensuing iterative procedure starts with a first "guess" of the single-electron wave functions with which the effective Hamiltonian of eqn [8] is computed. The Hartree equation is then solved numerically to obtain new approximations to the single-electron wave functions. Starting from these new wave functions, the effective Hamiltonian of eqn [8] is recalculated and the eigenvalue equation solved once again. This iteration is continued until the effective Hamiltonian computed from the new single-electron wave functions at the end of the cycle agrees with the one of the previous cycle within a given accuracy, thereby terminating the iteration.

Recourse to successive approximations makes the Hartree method nonlinear and renders its solution effectively of infinite order in the interaction Hamiltonian, thus overcoming the problems mentioned above with the (finite-order) perturbation theory.

Satisfying Pauli Principle

The Hartree wave function [4] satisfies the Pauli principle only in a partial way, in the sense that the single-electron wave functions are required to be all different from each other, thereby preventing two electrons from occupying the same single-particle state. However, the Pauli principle requires more stringently the many-electron wave function to be antisymmetric with respect to the interchange of any pair of electrons. The required antisymmetrization is achieved in a reasonably compact form for any number of electrons with a Slater-determinant formulation by

$$\Phi(q_1, q_2, \ldots, q_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} u_{\alpha_1}(q_1) & u_{\alpha_2}(q_1) & \cdots & u_{\alpha_N}(q_1) \\ u_{\alpha_1}(q_2) & u_{\alpha 2}(q_2) & \cdots & u_{\alpha_N}(q_2) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_1}(q_N) & u_{\alpha_2}(q_N) & \cdots & u_{\alpha_N}(q_N) \end{vmatrix} \qquad [9]$$

in the place of eqn [4] (the factor $\sqrt{N!}$ has been introduced to preserve the normalization). The Hartree–Fock wave function [9] thus contains $N!$ terms, of which the Hartree wave function [4] represents the diagonal one.

Hartree–Fock Equations

The averages [6] and [7] need to be recalculated at this point with the wave function [9] that obeys the Pauli principle. With the same variational procedure that has led to eqn [8], one now obtains the Hartree–Fock integro-differential equation(s):

$$h(q)u_{\alpha_i}(q) + \sum_j[\langle u_{\alpha_j}(q')|v(q,q')|u_{\alpha_j}(q')\rangle u_{\alpha_i}(q) - \langle u_{\alpha_j}(q')|v(q,q')|u_{\alpha_i}(q')\rangle u_{\alpha_j}(q)] = \varepsilon_i u_{\alpha_i}(q) \qquad [10]$$

The modification due to Fock (originating from the Pauli principle) has introduced the so-called exchange term on the left-hand side of eqn [10]. Correspondingly, the operator acting on the single-particle wave function $u_{\alpha_i}(q)$ is no longer purely differential as in eqn [8], but now also contains an integral part. Note that the exchange term of the Hartree–Fock operator, being of purely quantum-mechanical origin, lacks the intuitive interpretation of the (direct) Hartree term based on classical physics. Note further that the restriction $j\neq i$ has been removed from the sum in eqn [10], since the two terms with $i=j$ within brackets cancel each other. Despite the presence of these additional features, however, the Hartree–Fock equation still fails to include a correlation between electrons of either spin resulting from the Coulomb interaction.

From the above considerations, one anticipates that the Hartree approximation (whereby the exchange term in eqn [10] is neglected), being based on a classical self-consistent picture, should work better when classical considerations can give a good answer. This typically occurs when macroscopic effects are dominant, such as when the forces are slowly varying. Consistent with this expectation, one readily verifies from eqn [10] that for a short-range interaction $v(q, q')$ (as in nuclear physics or with the Hubbard Hamiltonian in a solid), the direct (Hartree) and exchange (Fock) terms are comparable in magnitude.

For a long-range interparticle potential (as in atomic physics), on the other hand, the exchange contribution is usually much smaller than the direct one, so that it is often neglected in determining the self-consistent energy levels of atoms.

The Hartree–Fock eqn [10] obviously shares with the Hartree eqn [8], the nonlinearity and self-consistency of the solution, and thus the need for iterative methods.

The total energy E of the original Schrödinger eqn [3] can be obtained from the solution of the Hartree–Fock eqn [10] by projecting on alternative wave functions $u_{\alpha_i}(q)$ and summing over i. One obtains

$$\sum_{i=1}^{N} \varepsilon_i = \langle \Phi|H_1|\Phi \rangle + 2\langle \Phi|H_2|\Phi \rangle = E + \langle \Phi|H_2|\Phi \rangle \qquad [11]$$

where $\langle \Phi|H_2|\Phi \rangle$ now contains the exchange besides the Hartree contribution [7]. Note that the total energy E is not merely the sum of the single-particle eigenvalues ε_i of eqn [10]. The average interaction energy $\langle \Phi|H_2|\Phi \rangle$ is, in fact, subtracted from this sum to avoid double counting of the interaction energy. In addition, the ground-state energy is obtained by considering the lowest N single-particle eigenvalues ε_i in eqn [11] (with occupancy consistent with the Pauli principle). Correspondingly, the remaining single-electron wave functions with higher energies represent unoccupied or "virtual" single-particle states. The value of the total energy E obtained in this way for the ground state constitutes an upper bound for the true ground state of the many-electron system, in accordance with the variational principle from which the Hartree–Fock eqn [10] has been obtained.

It remains to give a physical interpretation to the eigenvalues ε_i, which were originally introduced as Lagrange multipliers to enforce the (ortho) normalization constraints on the single-electron wave functions.

Koopmans' Theorem

Koopmans' theorem provides an interpretation to the Hartree–Fock eigenvalues ε_i as ionization potentials, in the sense that energies of the occupied states in the single-determinant approximation for the many-electron ground state represent the energy (with opposite sign) required to remove an electron from that state.

The central hypothesis of Koopmans' theorem is that one removes a particular single-electron wave function, say $u_{\alpha_{\tilde{i}}}(q)$, from the ground-state configuration [9] for N electrons, while leaving all other $(N-1)$ single-electron wave functions unchanged. By taking the difference between the total energies [11] corresponding to these N and $(N-1)$ electron configurations, one obtains

$$E(N) - E(N-1) = \varepsilon_{\tilde{i}} \qquad [12]$$

In physical terms, one says that Koopmans' theorem, by assuming a "frozen orbital" approximation, neglects the relaxation of the wave functions in the $(N-1)$ electron state. In practice, this neglect of relaxation results in very positive ionization potentials, even though Koopmans' ionization potentials constitute, in general, a reasonable first approximation to the experimental values.

The approximation of frozen orbitals is expected to be more appropriate when the electron is removed from an "outer" than from an "inner" level. By the same token, the approximation is expected to be a poor one in solids when screening (relaxation) effects are important. The valence and conduction bands of diamond constitute a notable example of a situation where "correlation effects" beyond the Hartree–Fock approximation produce important corrections to Koopmans' theorem results. As shown in **Figure 1**, the sets of valence and conduction bands of diamond get considerably closer to each other when including correlation effects on top of the Hartree–Fock results. In particular, the bandgap is reduced in this way from 15.0 to 7.4 eV, a value quite close to the experimental 7.3 eV.

Another famous example when screening (not taken into account by the Hartree–Fock–Koopman scheme) is important to account for the experimental data is provided by the ΔSCF (self-consistent field) calculations of Bagus and Schaefer for a deep excitation in a diatomic molecule. In a ΔSCF calculation, the "frozen orbital" approximation is relaxed in such a way that, after removal of one electron from a particular single-electron wave function (hole), the Hartreee–Fock single-electron wave functions for the remaining $(N-1)$ electrons are recalculated self-consistently. Bagus and Schaefer have shown that results in good agreement with the experimental data can be obtained, when one requires the deep-hole wave function to be localized at one of the two atoms of the diatomic molecule O_2, while it fails when the hole wave function is delocalized spreading over the whole molecule (as it is usually assumed for Hartree–Fock calculations with molecules).

Results for Atoms

The first numerical calculations for atoms were made direcly by Hartree and his students. Quite generally,

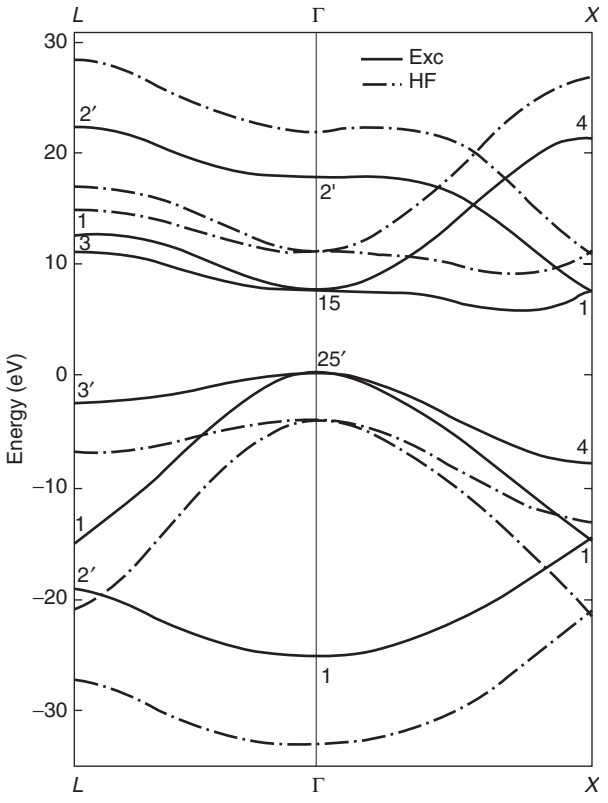


Figure 1 Valence and conduction bands of diamond including correlation effects (full lines) on top of Hartree–Fock results (dashed-dotted lines), along two symmetry directions in the Brillouin zone. (Reprinted with permission from Strinati G, Mattausch HJ, and Hanke W (1982) Dynamical aspects of correlation corrections in a covalent crystal. *Physical Review B* 25: 2867–2888; © American Physical Society.)

the first problem one encounters, when applying the Hartree or Hartree–Fock method to atoms, is that the potential due to the other electrons is not spherically symmetric, due to departures of the charge distribution from spherical symmetry. Including departures from spherical symmetry in atomic calculations would be rather difficult and probably not very useful, as it would not correspond to an improvement of the final results. Accordingly, Hartree did not use the actual nonspherically symmetric potential field but an average of this field over all directions in space, with the result that each electron moves in an effective central field.

The Hartree method provides good radial wave functions as a first approximation to the atomic structure problem and as the starting point for more refined calculations. With the inclusion of the Fock term, it turns out that the typical error of the Hartree–Fock energy for an atom (ion) is only $\sim 1\%$. The Hartree–Fock wave functions, on the other hand, may not be so accurate in some regions of configuration space which, however, do not play an important role in the variational integral from which the energy is obtained. As a consequence, matrix elements (required, for instance, to calculate transition rates) may sometimes have a substantial error when calculated in the Hartree–Fock approximation.

Standard tabulations of the Hartree–Fock results exist (see Clementi and Roetti), which gather the best results obtained for the ground state (and even for some excited states) of atoms. From these tables, one can get both total and single-particle (orbital) energies, as well as the form of the single-particle wave functions. Usually, Slater-type radial orbitals of the form $\chi(r) = Ar^{\nu-1}e^{-\eta r}$ (where $r = |\mathbf{r}|$, ν is a positive integer, η a positive number, and A a normalization constant) are chosen as a basis for these calculations, thereby determining the coefficients c_p of the expansion $u(r) = \sum_p c_p \chi_p(r)$ for the radial part $u(r)$ of the single-particle wave functions.

Results for Molecules

The Hartree–Fock approximation plays an especially important role in quantum chemistry, where it is also used as the starting point for more elaborate (configuration interaction) calculations that include the effects of electron correlations.

For molecules, the Hartee–Fock approximation is applied to the electronic equation that depends parametrically on the nuclear coordinates in the spirit of the Born–Oppenheimer approximation. In practice, for molecules it would be rather difficult to solve the Hartree–Fock integro-differential equations as one does for atoms, due to the reduced symmetry. Instead, for molecules one prefers (Roothaan) to introduce a finite set of known basis functions for the expansion of the single-particle wave functions, thus converting the Hartree–Fock equations into a set of algebraic equations for the expansion coefficients. These equations are then solved by standard matrix techniques using an iterative procedure. The basis functions are usually taken to spread over the whole molecule, with their symmetry properties classified in accordance with the symmetry group of the molecule.

The basis functions for this expansion, although linearly independent, are in general not orthogonal to each other, which results in the presence of an overlap matrix in the matrix equation. In particular, two types of basis functions are commonly used, the Slater-type and Gaussian-type functions. The use of Slater functions would in principle be preferable, since they more correctly describe the qualitative features of molecular orbitals. However, the calculation of two-electron integrals employing functions centered at the positions of different nuclei is easier to evaluate with Gaussian functions.

In literature, the results of a large number of Hartree–Fock calculations for molecules are available, especially as far as their total energy is concerned. For specific molecules (such as H_2), there also occurs good agreement between Koopmans' and experimental values of the ionization potential, although this agreement originates from a fortuitous cancellation of correlation and relaxation effects (both neglected by the Koopmans' approximation). For other molecules (such as N_2), however, the Hartree–Fock approximation is not sufficiently accurate even for a qualitative understanding of the ionization phenomena. Restricted Hartree–Fock calculations provide, instead, reasonable approximations to the experimental values for the equilibrium geometry of molecules, whose prediction is one of the most common uses of electronic structure calculations. ("Restricted" ("unrestricted") spin orbitals are constrained to have the same (different) spatial function(s) for up and down spins.)

In the case of unrestricted molecular orbitals, the equations to be solved (called the Pople–Nesbet equations) are analogous to the Roothaan's equations for the case of restricted molecular orbitals, but entail two coupled matrix eigenvalue problems to be solved simultaneously. A well-known example, when unrestricted Hartree–Fock calculations are preferred to the restricted ones, is the case of molecular dissociation. In this case, a lower energy is often obtained by allowing each doubly occupied molecular orbital to split into two spin orbitals with different spatial factors under a bond-breaking deformation of the molecule.

Results for the Homogeneous Electron Gas

Before discussing the Hartree–Fock calculations for solids, it is worth mentioning the results for the homogeneous electron gas (with which some metals such as Na and K are sometimes associated, at least to a first approximation).

When the potential contained in the single-particle Hamiltonian h of eqn [2] is constant, the Hartree–Fock eqn [10] can be solved exactly with plane waves $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ (where $\boldsymbol{k}$ is a wave vector), without the need for a self-consistent procedure (even though the corresponding energy eigenvalues $\varepsilon(\boldsymbol{k})$ are not free-electron eigenvalues). For a Coulomb interparticle potential (which requires the inclusion of a uniform positive compensating background), one then obtains, for the average (ground-state) energy per particle, the expression

$$\frac{E}{N} = \frac{2.21}{r_s^2} - \frac{0.916}{r_s} \qquad [13]$$

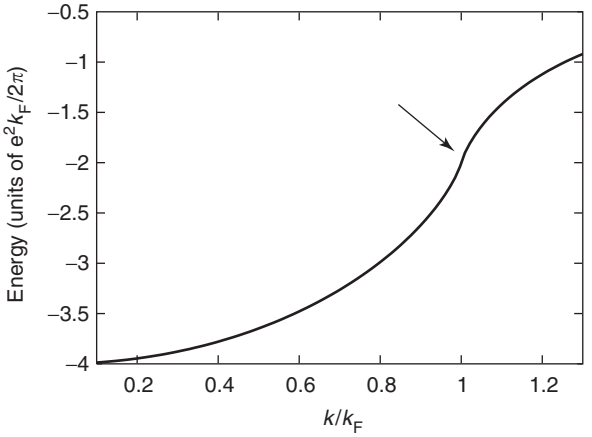


Figure 2 Hartree–Fock part $\varepsilon(\boldsymbol{k}) - \boldsymbol{k}^2/(2m)$ of the single-particle eigenvalues for a homogeneous electron gas vs. the magnitude $k = |\boldsymbol{k}|$ of the wave vector. The singularity at the Fermi wave vector k_F is evidenced by the arrow. (Adapted with permission from Fuchs HD, *et al.* (1992) *Solid State Communications* 82: 225.)

in units of Rydbergs. Here $r_s = r_0/a_0$ with $4\pi r_0^3 n/3 = 1$ (n being the density and a_0 the Bohr radius). In expression [13], the competition between the (positive) kinetic energy and the (negative) exchange energy contributions results in a minimum occurring for negative values of the total energy. For values of r_s corresponding to metals (typically in the range from 2 to 6), the exchange energy in eqn [13] is comparable in magnitude to the kinetic energy, thus confirming the need for nonperturbative calculations in condensed matter systems.

In spite of the simplified procedure, a shortcoming occurs with the Hartree–Fock result of the homogeneous electron gas. The Hartree–Fock part $\varepsilon(\boldsymbol{k}) - \boldsymbol{k}^2/(2m)$ of the eigenvalues (m being the electron mass) shown in **Figure 2** versus $k = |\boldsymbol{k}|$ has, in fact, a log-singularity in the first derivative at the Fermi wave vector k_F (related to the density via $k_F = (3\pi^2 n)^{1/3}$). This singularity, in turn, implies that the corresponding density of states vanishes at the Fermi energy, resulting for instance, in the vanishing of the electronic specific heat.

It is actually the neglect of screening (not taken into account in the Hartree–Fock approximation) that is responsible for the failure of the Hartree–Fock approximation with the homogeneous electron gas.

Results for Solids

In crystalline solids, where the atomic positions are related by translational symmetry, the solutions of the Hartree–Fock eqn [10] satisfy the Bloch theorem, and are thus labeled by a wave vector $\boldsymbol{k}$ (restricted to the first Brilluoin zone) and a band index.

In analogy with atoms, whereby the Hartree–Fock method is exploited to determine an average central potential, a simplified version of the Hartree–Fock method for solids (the X_α method) has been proposed by Slater, by setting up an average potential field in which all electrons move. In this way, the nonlocal exchange potential is replaced by a local one that depends on the local density, like constant $\times\, n(\boldsymbol{r})^{1/3}$. In addition, in analogy with molecules, the basis functions for full Hartree–Fock calculations are taken of the form of Bloch sums in terms of orbitals centered at the lattice sites.

Owing to its failure for the homogeneous electron gas, the Hartree–Fock approximation has long been viewed as a poor approximation for metals. *Ab initio* calculations of energy bands in solids with the Hartree–Fock method have thus been mostly performed for insulators with the Fermi level lying within the bandgap (the example of the energy bands of diamond has already been shown in **Figure 1**). Only recently the Hartree–Fock method has been applied to 3D transition metals, by truncating the long-range part of the Coulomb potential and thus effectively mimicking correlation effects.

In spite of these improvements, band structure calculations for solids have preferably been performed by alternative methods, such as the local density approximation to the density-functional theory of Hohenberg, Kohn, and Sham, and more recently by introducing correlation effects with the so-called GW approximation. It is nevertheless remarkable that both these alternative methods rely on the idea of a self-consistent field originally developed for the Hartree approach.

Successes and Limitations

It is worth summarizing the achievements of the Hartree–Fock method by spelling out its successes and limitations.

Among its successes, one should mention that: (1) The Hartree–Fock method is based on the variational principle, and, as such, has a sound physical–mathematical justification. (2) The method provides an approximate separation of variables of the original interacting N – electron problem, in that it enables one to solve N coupled one-electron problems in the place of one problem with N coupled electrons. (3) Even from a numerical point of view, the method provides reasonable order-of-magnitude estimates of single-particle eigenvalues.

Limitations of the Hartree–Fock methods concern instead: (1) The error in the single-particle eigenvalues, interpreted in conjunction with Koopman's theorem, which can admittedly be too large to allow for a detailed comparison with the experimental data, especially for systems where relaxation effects (screening) are relevant. (2) The first-principle extension of the method to excited states, which is not obvious since the excited states should (by construction) be orthogonal to the ground state.

Unrestricted Hartree–Fock Calculations

When discussing the results for molecules, it has already been mentioned that unrestricted Hartree–Fock calculations are sometimes preferred to the restricted ones for dealing with specific problems.

More generally, the technical term "unrestricted" refers to the fact that one allows for self-consistent field wave functions of a lower symmetry than that of the Hamiltonian. For instance, in spin-symmetry unrestricted self-consistent field solutions, the electrons on the average can be kept further apart than in the restricted solutions, thereby reducing their mutual Coulomb interaction.

It is sometimes said that "unrestricted" self-consistent field calculations enable one to incorporate part of the electron correlations within a self-consistent field scheme. This is due to the fact that a broken-symmetry ansatz effectively introduces nodes in the Hartree–Fock wave function, which amounts, in turn, to the introduction of some sort of correlation over and above the Hartree–Fock approach.

A notable example is the Hartree–Fock calculation for the Hubbard model in a crystal. In this case, the ordinary Hartree–Fock approach (with a spin-independent occupancy of the single-particle levels) merely produces the constant shift $nU/2$ of the independent-electron band eigenvalues (obtained in the absence of the electron–electron interaction). Here, n ($0 \leqslant n \leqslant 2$) is the average density per site, and U is the site interaction characteristic of the Hubbard Hamiltonian. At the same time, the Hartree–Fock wave function coincides with the independent-electron case, corresponding to the paramagnetic phase of the system. On physical grounds, these results are expected to be correct only for small values of U (such that $U \ll t$, where t is the hopping matrix element between nearest-neighbor sites). If one, instead, enforces a broken-symmetry situation to the Hartee–Fock approach (making the average spin occupancy different on alternative sites of a bipartite lattice), the approach proves its ability to describe antiferromagnetic situations which are (at least qualitatively) correct also in the large-U limit.

Generalized Mean-Field Approaches

The Hartree–Fock method is sometimes synonymous with mean field methods. As such, the method can be

readily generalized to finite temperatures within a grand canonical ensemble, whereby the Fermi–Dirac distribution contains the self-consistently determined single-particle energy eigenvalues of the Hartree–Fock equation. Cases of a broken symmetry can further be described with suitable Hartree–Fock approaches. For instance, the Stoner criterion for ferromagnetism is obtained within the Hartree–Fock approximation to the Hubbard Hamiltonian. The Bardeen–Cooper–Schrieffer (BCS) theory of superconductivity can also be regarded as a generalized Hartree–Fock decoupling of the attractive interparticle interaction. In this case, "anomalous" averages associated with the broken gauge symmetry are considered besides the usual Hartree (direct) and Fock (exchange) averages.

See also: Electronic Structure Calculations: Plane-Wave Methods; Quantum Mechanics: Atoms; Quantum Mechanics: Foundations; Quantum Mechanics: Methods.

PACS: 71.10. – w; 71.45.Gm; 71.15. – m; 71.20. – b

Further Reading

Bagus PS and Schaefer HF III (1972) Localized and delocalized 1*s* hole states of the O_2^+ molecular ion. *Journal of Chemical Physics* 56: 224–226.

Bransden BH and Joachain CJ (1983) *Physics of Atoms and Molecules*. New York: Longman.

Clementi E and Roetti C (1974) *Atomic Data and Nuclear Data Tables*, vol. 14, nos. 3–4. New York: Academic Press.

Condon EU and Shortley G (1963) *Theory of Atomic Spectra*. Cambridge: Cambridge University Press.

Fetter AL and Walecka JD (1971) *Quantum Theory of Many-Particle Systems*. New York: McGraw-Hill.

Fulde P (1991) *Electron Correlations in Molecules and Solids*. Berlin: Springer.

Mauger A and Lannoo M (1977) New method for calculating Hartree–Fock energy-band structures in solids using a linear combination of atomic orbitals basis: application to diamond. *Physical Review B* 15: 2324–2336.

McWeeny R (1992) *Methods of Molecular Quantum Mechanics*. New York: Academic Press.

Negele JW and Orland H (1988) *Quantum Many-Particle Systems*. Reading: Addison-Wesley.

Slater JC (1951) A simplification of the Hartree–Fock method. *Physical Review* 81: 385–390.

Strinati GC, Mattausch HJ, and Hanke W (1982) Dynamical aspects of correlation corrections in a covalent crystal. *Physical Review B* 25: 2867–2888.

Szabo A and Ostlund NS (1989) *Modern Quantum Chemistry*. New York: McGraw-Hill.

Nomenclature

E	total many-electron energy
h	one-electron Hamiltonian
H	many-electron Hamiltonian
k	wave vector
k_F	Fermi wave vector
m	electron mass
n	average electron density
$n(r)$	local electron density
N	total number of electrons
r	spatial position of an electron
t	hopping matrix element of the Hubbard model
u	one-electron wave functions
U	site interaction energy of the Hubbard model
Z	atomic number
α	label of the one-electron wave functions
ε	Hartree and Hartree–Fock single-particle eigenvalues
σ	spin of an electron
Φ	many-electron wave function

Highly Porous Metals and Ceramics

A E Markaki, University of Cambridge, Cambridge, UK
P Colombo, University of Bologna, Bologna, Italy
T W Clyne, University of Cambridge, Cambridge, UK

Introduction

Porous cellular metals and ceramics are now being intensively studied. This has arisen partly from a growing appreciation that biomaterials must be thoroughly understood, not only to exploit them effectively, but also to develop biomimetic materials with attractive properties and biocompatible materials and structures. Since many natural and biological materials are cellular, often with relatively high-porosity levels (although sometimes these pores are, at least, partly filled with a fluid or a soft and compliant solid in the natural state), study of manufactured porous materials has become an important branch of materials science. In fact, various types of

porous materials, quite different from any biological system, are also being heavily investigated. Furthermore, certain well-established areas of porous material usage, such as filters, catalyst supports, and heat exchangers, are being subjected to renewed scrutiny, as novel porous materials and structures are developed.

Interest extends to both processing developments and to relationships between microstructure and properties. The concept of microstructure should be defined with care in highly porous materials, since it is sometimes used to refer only to the cell structure (i.e., the size, volume fraction, shape, etc. of the pores), whereas in practice, the microstructure of material constituting the cell wall may also be important. Distinctions can be drawn between different types of porous materials, based on their cell structure. The most significant classification is that between open-cell and closed-cell materials. An obvious difference is that open-cell foams are permeable to fluids, which has various implications, but there are also some rather systematic differences between the two types of materials, relating to various thermophysical and mechanical properties, and also to the applications for which they might be suitable.

However, further classification is probably needed. Some porous materials exhibit distinct, separate cells, but with some connectivity – as, for example, might result from rupturing of some cell walls. Such configurations differ noticeably from open structures having such high connectivity that it is difficult to delineate any individual cells. It might be argued that the latter should be termed "open-network materials," rather than "cellular materials." Open-network materials can, for example, be produced by assembling a set of fibers or fine wires. Such assemblies can be isotropic or they may be highly oriented – commonly either in-plane or uniaxially aligned – and hence exhibit pronounced anisotropy in both structures and properties. Ropes, cables, braided wires, etc. are effectively open-network materials of this type, as are various types of insulation blankets, filters, etc. There are many such products in commercial use, based on metallic, ceramic, or polymeric fibers, with or without strong interfiber bonding.

Classifying ropes, or bonded assemblies of short struts, as (open-network) materials brings into sharp focus the issue of whether such things should really be termed materials or structures. In fact, many highly porous materials should probably be treated as "quasimaterials," since they are in many ways genuinely intermediate between the two. An important feature of quasimaterials, and of highly porous materials in general, is that the "pore architecture" introduces extra degrees of freedom into microstructural design. In thinking about terminology and classification, it is instructive to consider the combinations of pore content, connectivity, and anisotropy exhibited by various actual and putative materials and quasimaterials. This is illustrated graphically in **Figure 1**, which presents data for a number of well-known products and natural materials. Void content is simply evaluated, while anisotropy can be characterized via observed variations with direction of a suitable property, such as Young's modulus E (see the section "Elastic constants and stiffness-critical lightweight structures"). Intercell connectivity is more difficult to quantify. In constructing **Figure 1**, the ease of permeation of fluids through the material (see the section "Fluid permeability and specific surface area") has been used as the indicator, represented by the specific permeability κ (in m^2), which is independent of the fluid used in the measurement. Good intercell connectivity leads to high permeability. However, the permeability drops as the scale of the structure becomes finer, so κ has been multiplied by the specific surface area (in m^{-1}), which rises as this occurs, to give a connectivity parameter designed to be scale independent, at least to a first approximation. **Figure 1** highlights the wide range of cell architectures obtainable. Of course, the scale of the structure is a further important variable.

Highly porous materials are also sometimes classified according to whether their cell structures are regular or stochastic (i.e., exhibiting random variations in cell size, shape, and distribution). This can affect certain characteristics, particularly mechanical properties, since it influences stress concentration effects and crack propagation paths. The distinction is not, however, a very rigid one, since porous materials which do not, strictly speaking, have regular or periodic cell structures may nevertheless exhibit a relatively narrow distribution of cell size and shape. Finally, some porous materials tend to exhibit bimodal distributions of pore size, so that, for example, the cell walls might themselves be porous. This can influence a wide range of properties. An illustration is presented in **Figure 2** of cell structures in four different types of highly porous materials, exhibiting a wide range of regularity, anisotropy, connectivity, and scale.

In this article, a brief outline is given of how various properties of highly porous materials are predicted and controlled. It is sometimes possible to apply conventional composite theory, with one of the constituents taken as having zero stiffness or strength or conductivity, etc. However, models of this type often breakdown for high porosity levels, particularly as the pores become contiguous (open-celled), and, in general, rather different approaches are needed in this regime, taking proper account of the cell architecture.

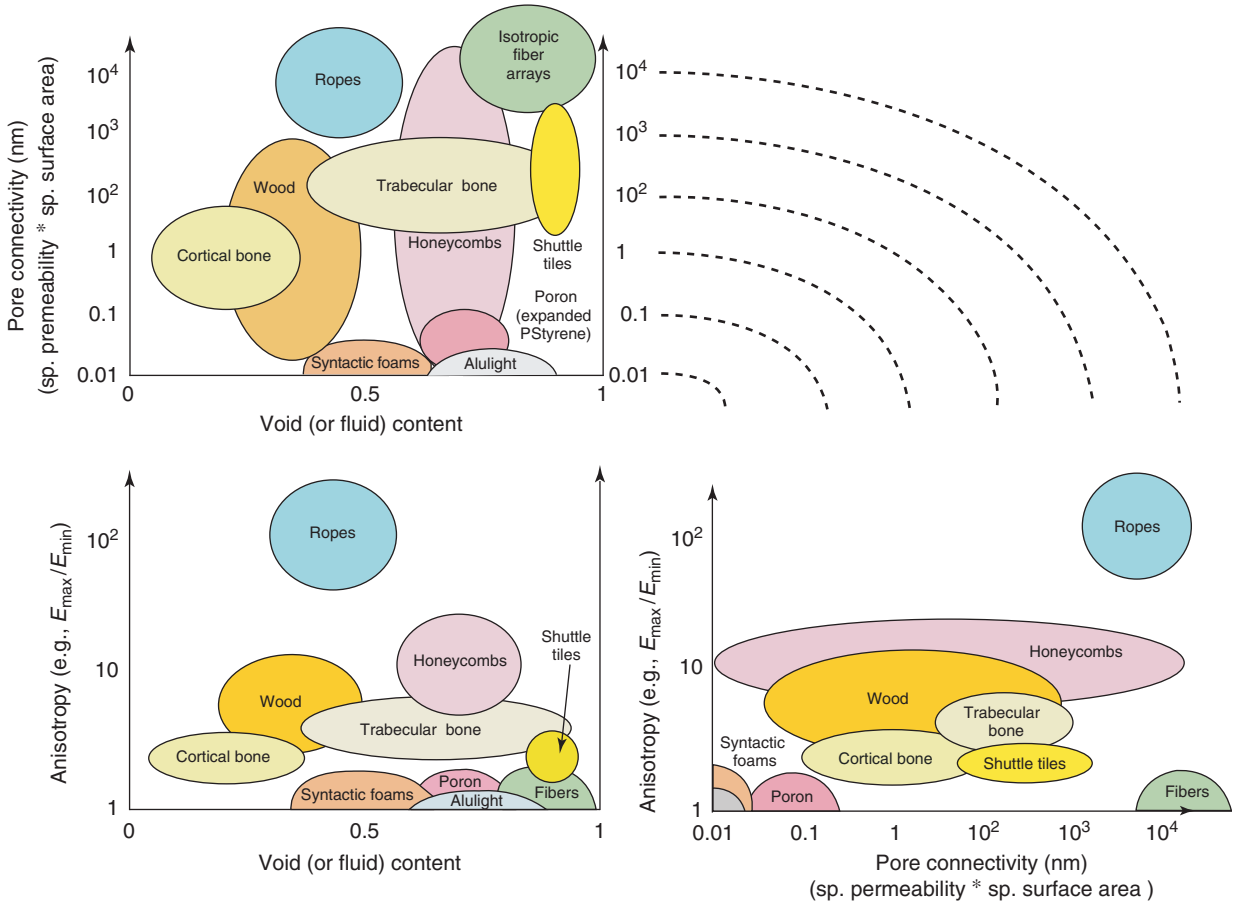


Figure 1 Semiquantitative map designed to aid classification according to pore architecture, showing the combinations of void content, anisotropy, and pore connectivity exhibited by a variety of highly porous materials. ("Alulight" is a closed-cell aluminum foam made via a gas generation route.)

Furthermore, porous materials may have interesting transport properties or potential for actuation characteristics and these often require rather specialized analysis. In outlining these structure–property relationships, reference is periodically made to the applications for which particular properties are most relevant. Of course, minimization of weight is a common motivation for using porous materials, and this is a key objective for many applications, but transport properties such as permeability and conductivity may also be of critical importance. Overall, the coverage is of necessity highly abbreviated and the bibliography provided will need to be consulted for any detailed information.

Mechanical Properties and Performance Requirements

Elastic Constants and Stiffness-Critical Lightweight Structures

The elastic constants of porous materials can be predicted using conventional composite theory models, with the "reinforcement"-ascribed zero stiffness (and zero Poisson ratio). For example, **Figure 3** shows how axial and transverse Young's modulus and Poisson ratio are predicted to vary with void content, and with void aspect ratio, according to the Eshelby method, which is based on assuming that ellipsoid-shaped inclusions (pores) are dispersed throughout the matrix. The following equation gives the elastic constants

$$\boldsymbol{C} = [\boldsymbol{C}_{\mathrm{m}}^{-1} - f\{(\boldsymbol{C}_{\mathrm{i}} - \boldsymbol{C}_{\mathrm{m}})[\boldsymbol{S} - f(\boldsymbol{S} - \boldsymbol{I})] + \boldsymbol{C}_{\mathrm{m}}\}^{-1}(\boldsymbol{C}_{\mathrm{i}} - \boldsymbol{C}_{\mathrm{m}})\boldsymbol{C}_{\mathrm{m}}^{-1}]^{-1} \quad [1]$$

where C is the stiffness tensor, the subscripts m and i refer to matrix and inclusions (pores), S is the Eshelby tensor (dependent on inclusion aspect ratio s), I is the identity tensor and f is the volume fraction of inclusions (pores). The model is known to be unreliable for high inclusion contents, but it, nevertheless, serves to illustrate a few simple points.

First, the specific stiffness of porous materials differs little from that of the matrix, since the Young's

modulus falls off in a broadly similar way to the density (rule of mixtures line). In fact, apart from the axial stiffness of a material with aligned, highly elongated pores, which approximates closely to the rule of mixtures line, the stiffness falls off more sharply than the density, so the specific stiffness of porous materials is, in general, lower than that of corresponding monolithic materials.

Second, porous materials can be quite significantly anisotropic in stiffness, if the pores are aligned and elongated. This is the case in many natural cellular materials. In fact, it is clear from the plot that the pores do not need to be highly elongated for this to be the case, since the $s = 3$ curves are fairly close to those for infinite aspect ratio. However, it is also worth noting that the plot suggests that elastic anisotropy is unlikely to be dramatic (greater than a factor of 2 or so). Some assumptions, such as those of the equal stress model, lead to much greater predicted anisotropy, but these are known to be very unreliable, at least for cylindrical inclusions (pores). On the other hand, experimental data for transverse stiffness often suggest that it is much lower than in the axial direction. For example, typical woods (with a void content of $\sim$20–40%) are often quoted as having an axial (parallel to grain) stiffness of $\sim$10 GPa, but a transverse value of only $\sim$1 GPa. For bone, the quoted anisotropy is usually lower, but is still substantial in some cases. Partial explanation for such discrepancies is often related to difficulties in measuring a true transverse stiffness. The strength (and creep resistance) is often dramatically lower under transverse loading (see below), so that inelastic straining commonly occurs at low applied loads

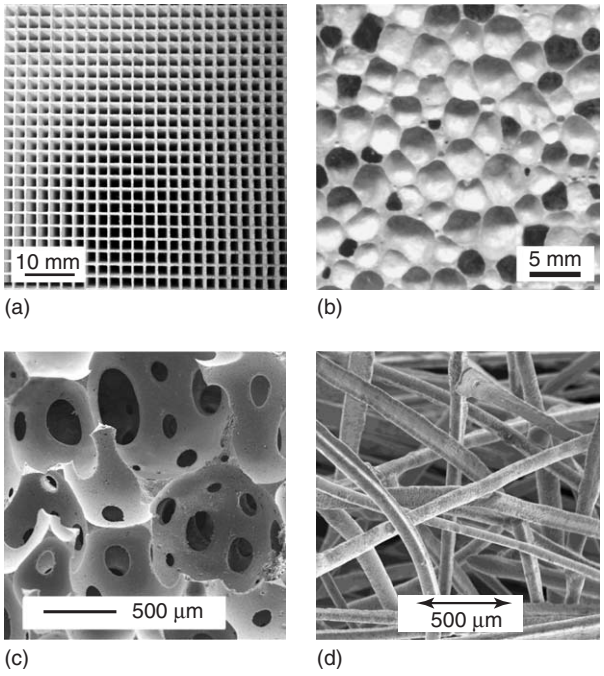


Figure 2 Microstructures of (a) a ceramic honeycomb, (b) a closed-cell metallic foam, (c) an open-cell SiOC foam, and (d) an open-network structure made by bonding metallic fibers together.

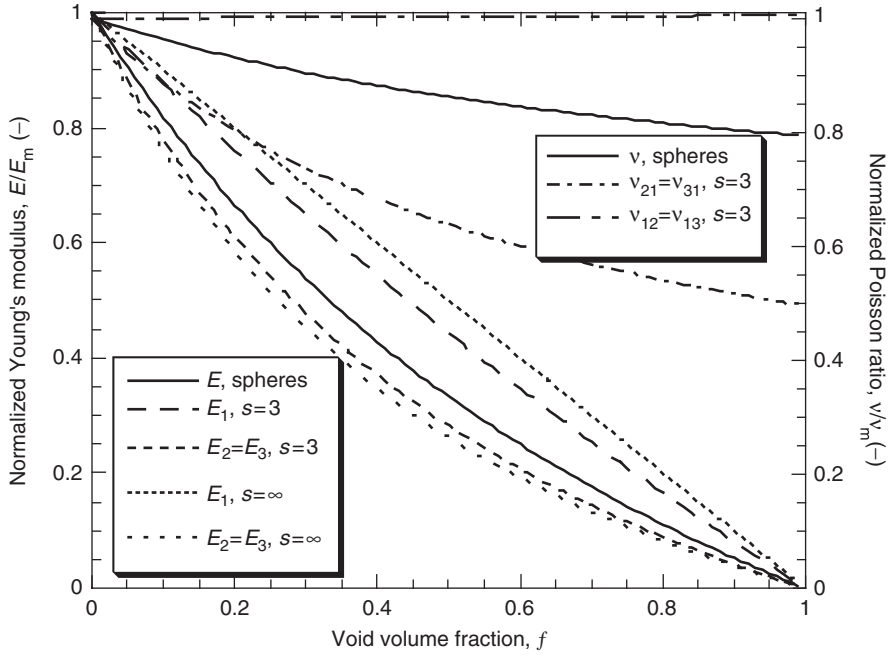


Figure 3 Predicted dependence on void content f, and void aspect ratio s, of the Young's modulus and Poisson ratio, obtained using the Eshelby method (eqn [1]).

during stiffness measurement. However, it should also be recognized that this type of model, which takes no account of the distribution of voids, and assumes that they have no contiguity (i.e., that the cells are all closed), may be very unreliable for high-void contents and for open-cell materials. Finally, it must be noted that these models assume the base material to be isotropic. For many natural materials, such as wood, the cell walls themselves exhibit high degrees of structural alignment, and hence are highly anisotropic, both elastically and in terms of strength. Obviously, such effects should be taken into account when predicting the effect of the presence of pores.

It can also be seen in **Figure 3** that the Poisson ratio may vary significantly when voids are introduced, although the relative changes are predicted to be smaller than those for stiffness. However, it should again be emphasized that, particularly at high-void contents, and with open-cell structures, geometrical aspects can assume much greater importance, and the "mean field" approach of the Eshelby method is certainly not well-suited for prediction of elastic constants at high-void contents. For highly porous materials of this type, geometrically specific models may be more appropriate. For example, consider the plots shown in **Figure 4**. These relate to porous materials composed of an isotropic (random) array of slender cylindrical fibers, bonded together at junction points – see **Figure 2d**. A model has been developed for such material, based on assuming that deformation arises solely from elastic-bending deflections of the fiber segments between junctions. The following simple equation is derived for the normalized Young's modulus

$$\frac{E}{E_{\mathrm{m}}} = \frac{9(1-f)}{32(L/D)^2} \qquad [2]$$

where L/D is the aspect ratio of fiber segments between joints. These stiffnesses, both measured and predicted, are well below those predicted by the Eshelby method for the void contents concerned. In general, open-cell materials are less stiff than closed-cell ones with similar void contents, while open-network materials are the least stiff. These discrepancies highlight some of the limitations of applying conventional composite theory approaches to highly porous materials, particularly if they have open-cell or open-network structures. It might also be noted that the model used to produce the plots in **Figure 4** predicts a Poisson ratio value of ~ 0.32, independent of void-content and fiber segment aspect ratio.

Finally, it should be emphasized that, while porous materials do not have superior specific stiffness to monolithic materials under uniaxial loading, they can exhibit dramatic advantages when subjected to bending moments or torques. The reason for this is

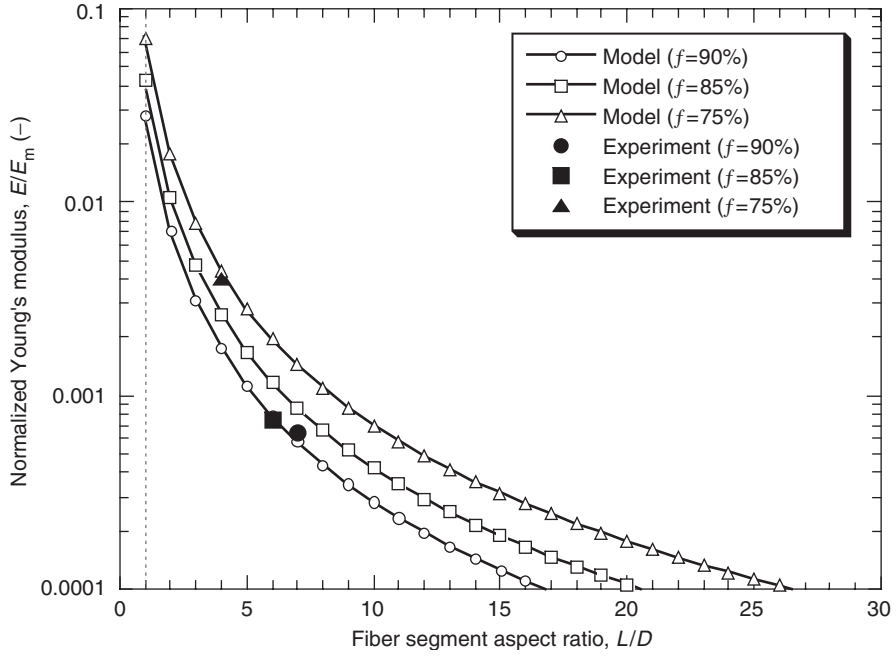


Figure 4 Predicted Young's modulus of an open-network porous material comprising an isotropic array of bonded fibers, as a function of the aspect ratio L/D, of fiber segments between junctions, according to eqn [2]. Curves are plotted for several void contents (f) and some corresponding experimental data are also shown.

that "bending stiffness" and "torsional stiffness" are raised, not only by the material having a higher Young's modulus, but also by it being located further from the neutral axis or neutral plane. Depending on the exact boundary conditions, the appropriate figure of merit for stiffness-critical weight minimization (E/ρ for uniaxial loading) is commonly $E^{1/2}/\rho$ or $E^{1/3}/\rho$ for bending, which can make highly porous materials attractive for construction of many types of stiff, lightweight structures. A simple example of this is sandwich panels, for which optimizing the design of highly porous core material is an important issue.

Strength, Fracture Mechanics, and Energy-Absorbing Structures

In general, pores are expected to cause both stress concentration effects and easier crack paths, leading to reductions in strength. In principle, a conventional fracture mechanics calculation should allow estimation of the effect of a dispersed set of (closed cell) pores on the strength and toughness. The passage of a crack through pores reduces the need for material to be fractured and hence reduces the fracture energy. Neglecting any effects of preferential crack paths or reductions in constraint near pores, the fracture energy should thus be proportional to the relative density ($1–f$), with no dependence on pore size. The stiffness falls off with increasing void content in a broadly similar manner (see the section "Elastic constants and stiffness-critical lightweight structures"). Pores will also act as crack-initiating flaws, reducing the stress required for the onset of fast fracture. In this role, the pore size is significant. To a zeroth approximation, the strength of a porous material could thus be written as a conventional fracture mechanics strength

$$\sigma_* = \left(\frac{G_c E}{\pi c}\right)^{1/2} \approx (1-f)\left(\frac{G_{cm} E_m}{\pi R}\right)^{1/2} \qquad [3]$$

where c is the flaw size, G_{cm} is the fracture energy of the base material and R is the radius of the crack-initiating (largest) pore, assumed to be spherical. This equation is plotted in **Figure 5**, which illustrates how high-porosity contents and large pores lead to low-fracture strengths. However, it should be recognized that experimental strengths (and fracture energies) tend to be very low for most highly porous metals and ceramics, even if the pores are relatively fine, and, in general, these predictions are over-estimates, particularly for metals. This is thought to be largely due to a tendency for cracks to preferentially follow high-porosity paths and for ductile tearing of ligaments between pores to occur readily as stress concentration effects operate. This can substantially reduce the effective toughness. While tensile strengths are higher if the pore (cell) structure is fine and regular, most experimentally measured values for highly porous materials are no higher than a few MPa, although some reported values for materials based on metal fibers are in the range of several tens of MPa.

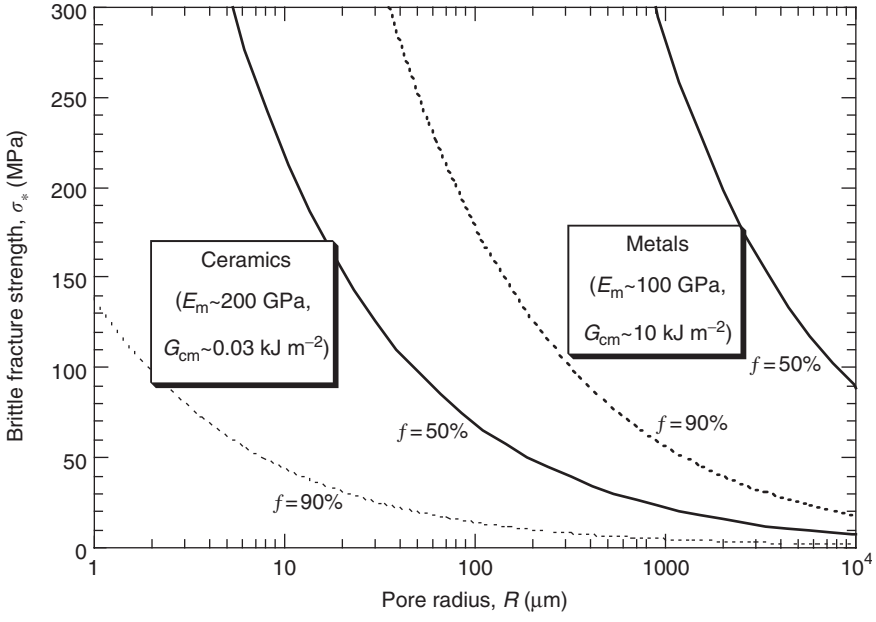


Figure 5 Predicted dependence of the fracture strength of porous metals and ceramics on void content and size, according to eqn [3].

Where highly porous materials are used in load-bearing situations, attempts are often made to ensure that they experience predominantly compressive loads (although it may be noted that this is not possible under bending and twisting moments, which is when their elastic properties can be most effectively exploited – see the section "Elastic constants and stiffness-critical lightweight structures"). Under uniaxial compression, highly porous materials tend to yield at relatively low stresses (from a few MPa up to a few tens of MPa) and then undergo quite large strains without the stress level rising much. This strain is associated with local plastic deformation and/or fracture and crushing events. Such deformation can be useful when substantial energy absorption is required without high stresses being generated, as in cushioning of impact events or redistribution of static or dynamic loading configurations. Of course, polymeric foams have long been used for such purposes, but they do not have high energy-absorbing potential and they also suffer from other drawbacks, such as susceptibility to fire. Metals offer particular promise in this regard, since very high local plastic strains can, in principle, be generated, due to the reduced constraint on individual cell walls or network members. This potential has not yet been fully exploited and further work is required on optimization of the energy-absorbing characteristics of highly porous metals. Some such material incorporates significant levels of ceramic in the cell walls (introduced to facilitate the processing) and this certainly impairs plasticity and energy absorption. On the other hand, there may be some situations in which predominance of brittle crushing processes over plastic deformation may be advantageous, since it will tend to result in reduced stress transfer to the component being protected.

Strength can be markedly anisotropic in porous materials. In highly aligned cellular materials, such as wood, factors of 10 between axial and transverse strengths are commonly observed. In materials such as wood and bone, this is partly attributable to the alignment of microstructural constituents (cellulose microfibrils, collagen, etc.) within the cell walls, as well as to the anisotropic cell geometry. However, even in the absence of such effects, strength anisotropy from cell or network geometry can be dramatic, with ropes providing a graphic example. The nature of the bonding in transverse directions is clearly relevant to the magnitude of the anisotropy and in some cases this can be very weak indeed.

Finally, it should be emphasized that the toughness and strength of highly porous materials is not a simple topic. While the presence of pores tends, in general, to promote premature yielding and fracture, there are situations in which the compressibility of porous materials represents a substantial advantage. For example, while the outstanding toughness and damage tolerance of wood arises partly from specific energy-absorbing mechanisms, notably fiber pullout, it is also partly due to the compressibility conferred by the presence of pores. This allows intrusive events, such as the introduction of screws, to take place without the large stress buildup and consequent macroscopic distortion or crack propagation which would tend to result with most fully dense materials, even those with relatively high toughness. Bone is also able to withstand screwing and similar fixation operations quite well, although they can sometimes cause cracking. It might also be argued that the damage tolerance and general mechanical versatility of ropes arises largely from their compressibility, or, at least, from the relative independence of individual strands, which is clearly associated with the presence of the free space. Unfortunately, most manufactured porous materials do not exhibit toughening mechanisms analogous to those operating in wood and bone, which are attributable to their complex, multi-scaled microstructures. Furthermore, the damage tolerance and flexibility of rope-like materials is achieved at the cost of very high compliance and low strength in transverse directions. The degree to which most manufactured porous materials benefit mechanically from their high compressibility has thus been rather limited, although it has been noted in some cases that they are suitable for screwing etc. Nevertheless, open-network quasimaterials of various types have excellent potential for offering attractive combinations of mechanical properties and their further development seems likely.

Transport Properties and Multifunctional Applications

Electrical and Thermal Insulation Characteristics

Since there is substantial resistance to the passage of both heat and electrical charge through pores (gaseous or vacuum), porous materials have long been of interest as insulators. However, there are important differences between the two cases for porous materials. In particular, pore connectivity is often an important issue for heat flow. The main reason for this is that convective heat transfer can be significant. There is a scale effect here, since natural thermal convection within closed cells tends to be significant only when the cells are relatively large (typically, $> \sim 10\,\text{mm}$). However, in open-cell materials natural convection can occur even with finer structures and forced convection can be an important heat transfer

mechanism (to such an extent that open-pore materials are often used as heat exchangers – see the section "Fluid permeability and specific surface area"). In the electrical case, on the other hand, the focus is entirely on current paths through the matrix, assuming that the breakdown field for ionization of gas in the pores is not reached. In this section, attention is concentrated on conductive transfer.

From a mathematical point of view, conduction of heat and of electrical charge are indistinguishable. The basic equations of unidirectional heat and current flow may be written as

$$q = -K\frac{\partial T}{\partial x} \qquad [4]$$

$$j = -\sigma\frac{\partial V}{\partial x} \qquad [5]$$

where q is the heat flux ($\mathrm{W\,m^{-2}}$) arising from a thermal gradient $\partial T/\partial x$ ($\mathrm{K\,m^{-1}}$) in a material of thermal conductivity K ($\mathrm{W\,m^{-1}\,K^{-1}}$), while j is the current density ($\mathrm{A\,m^{-2}}$), σ is the electrical conductivity ($\Omega^{-1}\,\mathrm{m^{-1}}$ or $\mathrm{S\,m^{-1}}$) and $\partial V/\partial x$ is the electric field ($\mathrm{V\,m^{-1}}$). It is common to use the resistivity ρ ($\Omega\,\mathrm{m}$), which is the reciprocal of the electrical conductivity. Both conductivities are proportional to the concentration of carriers, and to their mobility. For thermal conduction, either electrons or phonons can be effective carriers. Good thermal conduction is conferred by the presence of free electrons (metals) or by high phonon velocities (light, stiff materials) and large phonon mean free paths (few scattering centers, i.e., low levels of defects, grain boundaries, etc.). In contrast to this, only electrons or ions can act as carriers of electric charge. Since electrons are much more mobile than ions, materials with free electrons (metals) have vastly greater electrical conductivities than all other materials. Differences in thermal conductivity between different types of material are much less dramatic.

It therefore follows that, if the objective is electrical insulation, then most ceramics are effective anyway and there is little incentive to make them porous, unless a large volume is beneficial – for example, in increasing the thickness of a ceramic insulating layer while keeping its mass low. The electrical resistance of such a porous ceramic layer would probably be similar to a dense ceramic layer of the same thickness, although other electrical properties, such as dielectric characteristics, will often be substantially different. For metals, on the other hand, electrical conductivity will fall as the porosity is increased, but even highly porous metals are unlikely to be suitable for electrical insulation purposes. In addition, there may be interest in maximizing the electrical conductivity of porous metals, so as to facilitate electrical connections or allow resistance welding.

For thermal insulation purposes, however, there is interest in quantifying the conductivity of both ceramics and metals, as a function of porosity level and pore architecture. If the pores are treated as isolated ellipsoids, then the Eshelby method can again be employed, leading to the following equation for the conductivity (electrical or thermal)

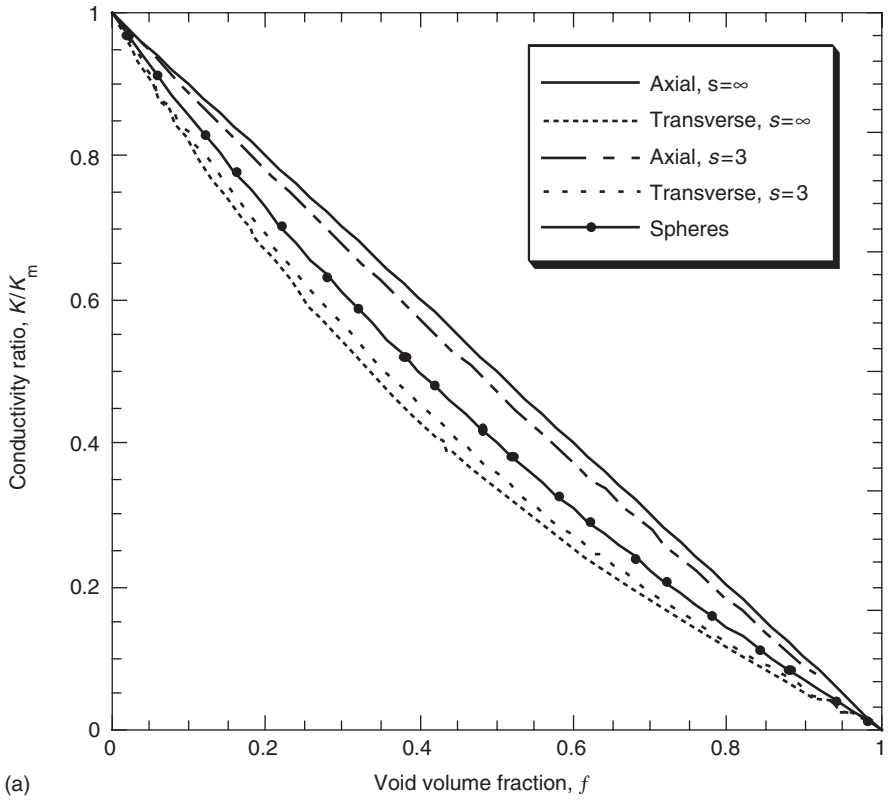
$$K = [K_{\mathrm{m}}^{-1} + f\{(K_{\mathrm{m}} - K_{\mathrm{I}})[S - f(S - I)] - K_{\mathrm{m}}\}^{-1}(K_{\mathrm{i}} - K_{\mathrm{m}})K_{\mathrm{m}}^{-1}]^{-1} \qquad [6]$$

which reduces for the case of spheres to the familiar Maxwell equation

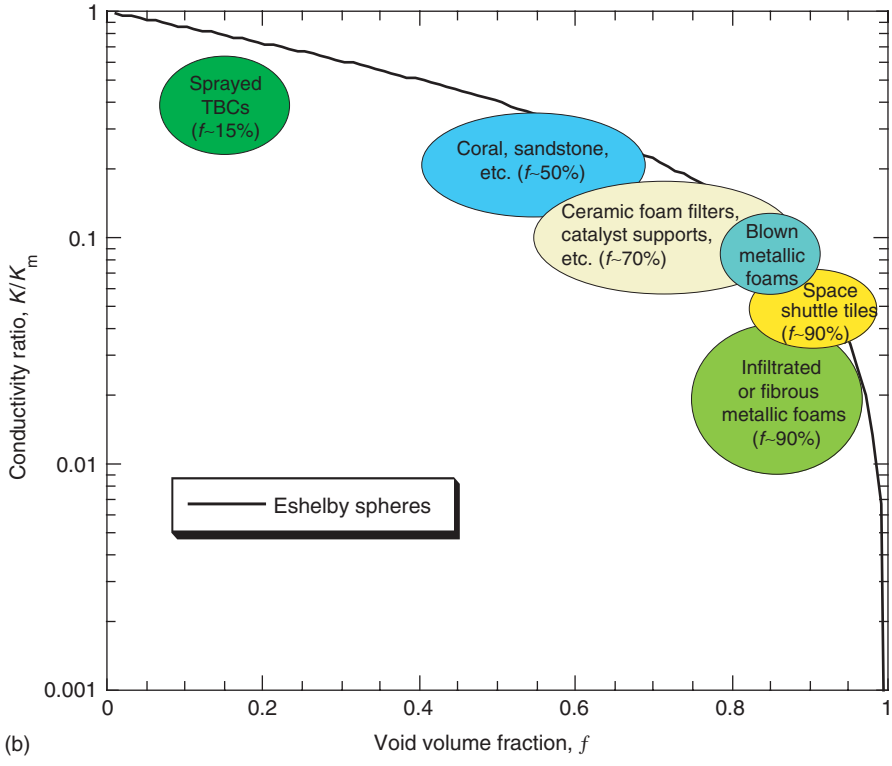
$$\frac{K}{K_{\mathrm{m}}} = \frac{2(1-f) + (K_{\mathrm{I}}/K_{\mathrm{m}})\,(1+2f)}{(2+f) + (K_{\mathrm{I}}/K_{\mathrm{m}})\,(1-f)} \qquad [7]$$

This Eshelby expression is very similar to that for the elastic constants (eqn [1]), but these are now second-rank tensors, rather than fourth rank. Predictions from eqn [6] are shown in **Figure 6a**, with the void conductivity taken as zero ($K_{\mathrm{air}} \sim 0.025\,\mathrm{W\,m^{-1}\,K^{-1}}$, compared with typical values of ~10–200 for metals and 1–150 for ceramics). As with the elastic constants, the reliability is rather poor at high-void contents ($> \sim 50\%$), when the structure will not in most cases approximate well to a matrix containing a set of isolated ellipsoidal voids and the pore architecture should be taken into account. In general, conductivities fall significantly below the Eshelby-predicted levels at high-void contents, as the conduction paths through the matrix become longer and more convoluted, particularly if the pore connectivity is high. This is illustrated by **Figure 6b**, which presents a normalized plot of thermal conductivity data for various porous metals and ceramics, with structures ranging from a set of layered splats to open three-dimensional fibrous networks. These include the silica fiber-based space shuttle tiles and plasma-sprayed zirconia thermal barrier coatings (TBCs). The latter exhibit normalized conductivities well below the Eshelby predictions, even though the void content is relatively low. This is a consequence of their layered structure, with many poorly bonded interfaces normal to the heat flow direction. (Note that the blown metallic foams have closed-cell structures, while the infiltrated and fibrous ones are open cell.)

With particular pore architectures, the behavior can often be predicted quite reliably using simple geometrical models. For example, electrical conduction through a bonded metallic-fiber array has been



(a)



(b)

Figure 6 (a) Predicted dependence of the conductivity ratio on void content, *f*, and aspect ratio, *s*, obtained using the Eshelby method (eqn [6]), and (b) typical experimental ranges for the thermal conductivity ratio of various porous metallic and ceramic materials, with the Eshelby prediction for spheres included for reference purposes.

represented using the model shown in **Figure 7a**, assuming symmetry about the current flow axis, with all fibers equally inclined to it. It is easily shown that the resistivity of such material is related to that of the fibers by

$$\rho = \frac{\rho_{\text{fib}}}{(1-f)\cos^2\theta} \qquad [8]$$

where θ is the inclination angle. A comparison is presented in **Figure 7b** between predictions from this equation and some experimental data. The resistivity of such material tends be much higher than that of pore-free metal, by several orders of magnitude, even when the porosity content is only $\sim 80\%$, and it can be seen that the model captures this effect. That the internal architecture is important in this case is clear from the predicted sensitivity to the fiber inclination angle.

However, it is important to recognize that such simple models are only expected to be reliable if the transport process is dominated by conventional conduction through an isotropic matrix. This is quite likely for electrical current flow through a porous metal, but for heat flow through metals or ceramics there may also be significant convective and/or radiative contributions, while heat conduction through gas in the pores can be relevant and may be affected by pore dimensions, gas pressure, etc. Prediction of the thermal conductivity values exhibited by the materials represented in **Figure 6b** thus requires modeling that takes into account both structural architecture and details of the heat transfer conditions.

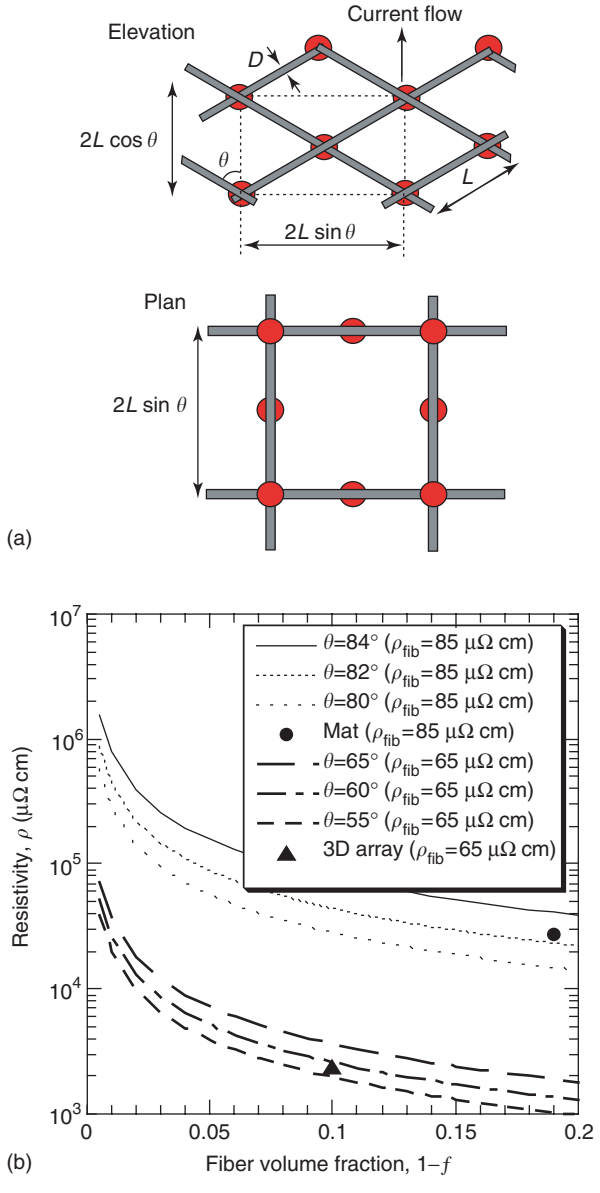


Figure 7 (a) Unit cell for a simple geometrical model of a bonded fiber array and (b) predicted resistivities compared with experimental data for two brazed materials – a mat made from long austenitic fibers (85 μΩ cm) and a 3D array material made from short ferritic fibers (65 μΩ cm).

Fluid Permeability and Specific Surface Area

The permeability of an open-celled porous material is an important property for several types of application. Permeation of fluids through such materials is a key issue for heat exchangers, filters, fuel cells, batteries, catalyst platforms, fuel tank baffles, fluid diffusers/mixers, and various types of textiles, as well as in many natural systems such as soils and rocks. The permeation process is clearly facilitated by high cell connectivity. The scale of the structure, which can be characterized by the specific surface area

$$S = \frac{A_{\text{sur}}}{V_{\text{tot}}} \qquad [9]$$

is also of importance. The resistance to fluid flow increases as the scale is refined, but a high surface area may be beneficial in some applications, including many electrochemical devices and combustion-related systems, such as solid fuel rocket propellants. For heat exchangers, there is expected to be an optimal specific surface area, allowing both rapid fluid permeation and efficient local heat exchange. This will also be true for filters, since the scale needs to be sufficiently fine to entrap the target bodies in the fluid stream, but not so fine that the entrapped bodies rapidly cause the permeability to fall as clogging occurs. Finally, it may be noted that encouragement of permeation is desirable in certain biomedical applications, such as porous surfaces of prosthetic implants into which bone growth is to be promoted.

Clearly, the permeation velocities of interest may cover a very wide range. Some example materials, designed for particular applications in which permeation and surface area are relevant, are shown in **Figure 8**.

Laminar flow of a fluid through a porous medium commonly conforms to Darcy's law, which may be written as

$$\frac{Q}{A} = \frac{\kappa}{\eta}\frac{\partial P}{\partial x} \qquad [10]$$

where Q is the volumetric flow rate ($m^3\,s^{-1}$), A is the sectional area (m^2), $\partial P/\partial x$ is the pressure gradient ($Pa\,m^{-1}$), η is the dynamic viscosity of the fluid (Pa s) and κ is the specific permeability (m^2). Permeability can thus be measured experimentally by recording the flow rate under a known pressure gradient. It should have the same value for all fluids.

Accurate prediction of permeability as a function of pore architecture is a relatively complex problem, but some simple models have been developed. The value of $\sqrt{\kappa}$ has dimensions of length and one approach involves equating this to some appropriate distance, such as the pore diameter or the hydraulic radius of the structure, defined as

$$R_h = \frac{f}{S} \qquad [11]$$

Such a rationale leads to the so-called Kozeny equations, of which the best known is probably the Blake–Kozeny equation. This is based on semi-empirical correlations for flow through packed beds and may be written as

$$\kappa \approx \frac{f^3}{4.2\,S^2} \qquad [12]$$

The specific surface area S, can be measured directly or can be estimated via a suitable geometrical

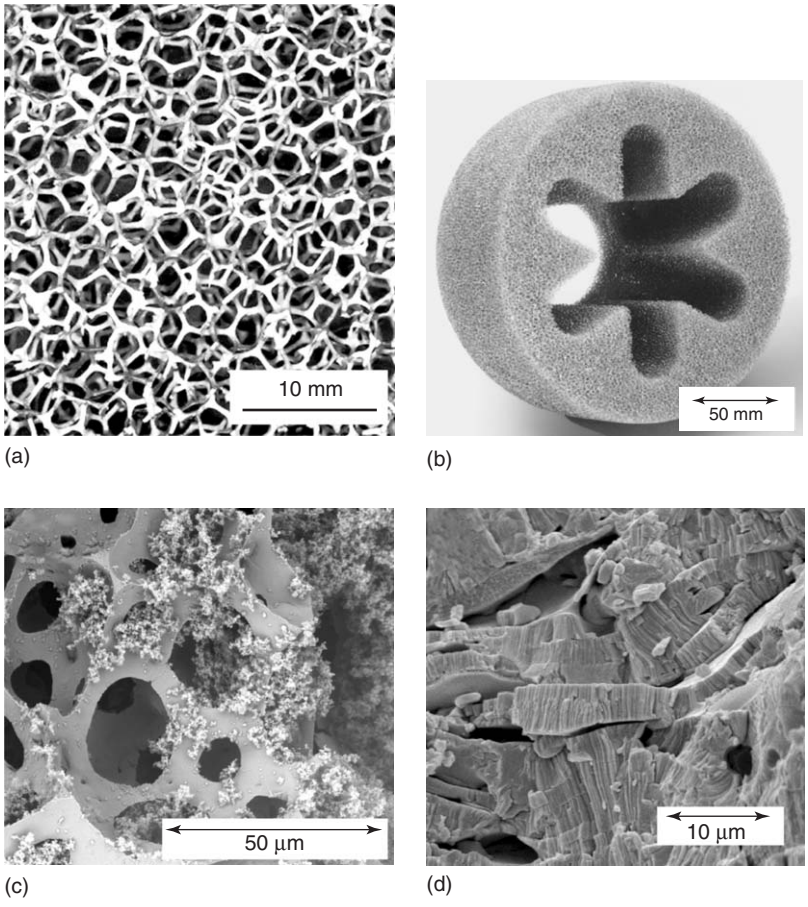


Figure 8 Examples of porous materials for specific applications: (a) microstructure of "Duocel" Al foam ($f \sim 90\%$), (b) rocket propellant component made of Duocel, (c) alumina filter ($f \sim 90\%$) with entrapped soot particles, and (d) plasma-sprayed zirconia thermal barrier coating (TBC), used in gas turbine engines ($f \sim 15\%$). ((c) From Green DJ and Colombo P (2003) Cellular ceramics: intriguing structures, novel properties, and innovative applications. *MRS Bulletin* 28(4): 296–300. Reproduced by permission of MRS Bulletin.)

model for the pore architecture, giving relationships such as the following

$$S \approx \frac{3.84}{D} \quad (f \approx 0.36) \quad \text{dense random packing of solid spheres (diameter } D\text{)} \qquad [13]$$

$$S = \frac{3\pi}{D(L/D)^2} \quad \left(f = 1 - \frac{3\pi}{4(L/D)^2}\right) \quad \text{cubic array of cylinders (diameter } D\text{, length } L\text{)} \qquad [14]$$

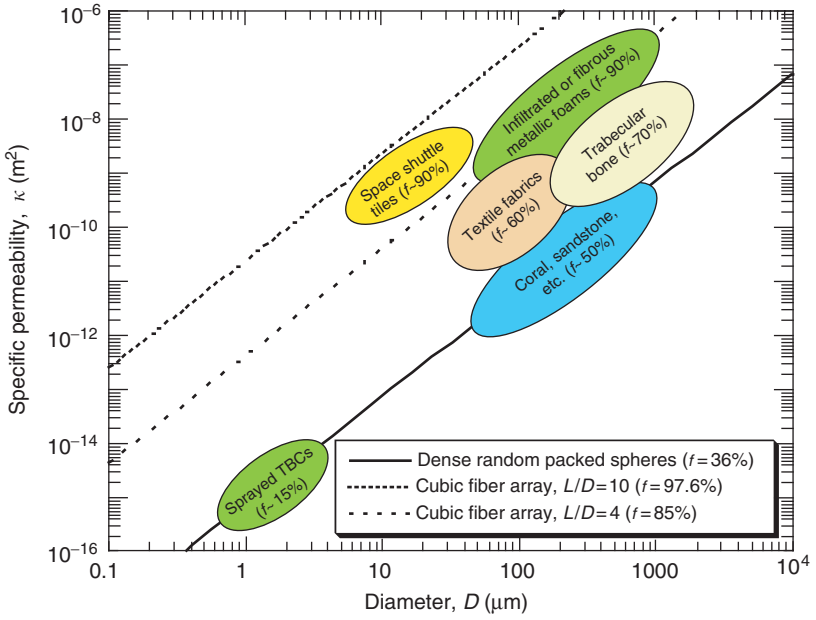


Figure 9 Plot of specific permeability against the structural length scale showing predictions from the Blake–Kozeny equation (eqn [12]), for a packed bed of spheres (eqn [13]) and for a cubic array of cylindrical fibers (eqn [14]), with two different aspect ratios between the joints. Also shown are the approximate ranges of experimental data obtained with various types of porous material.

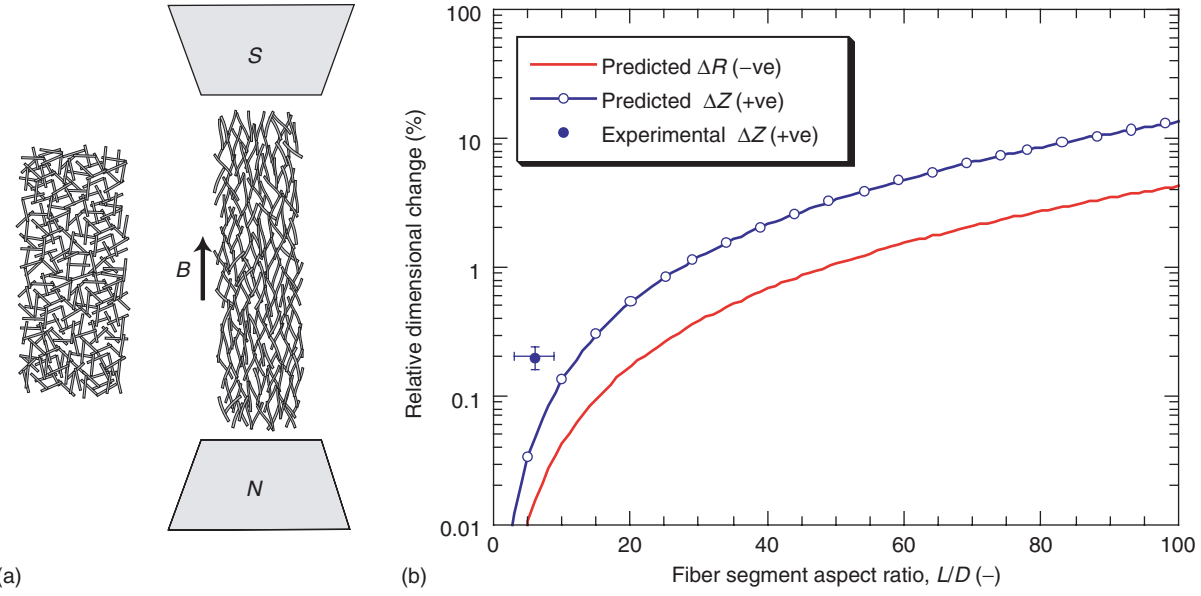
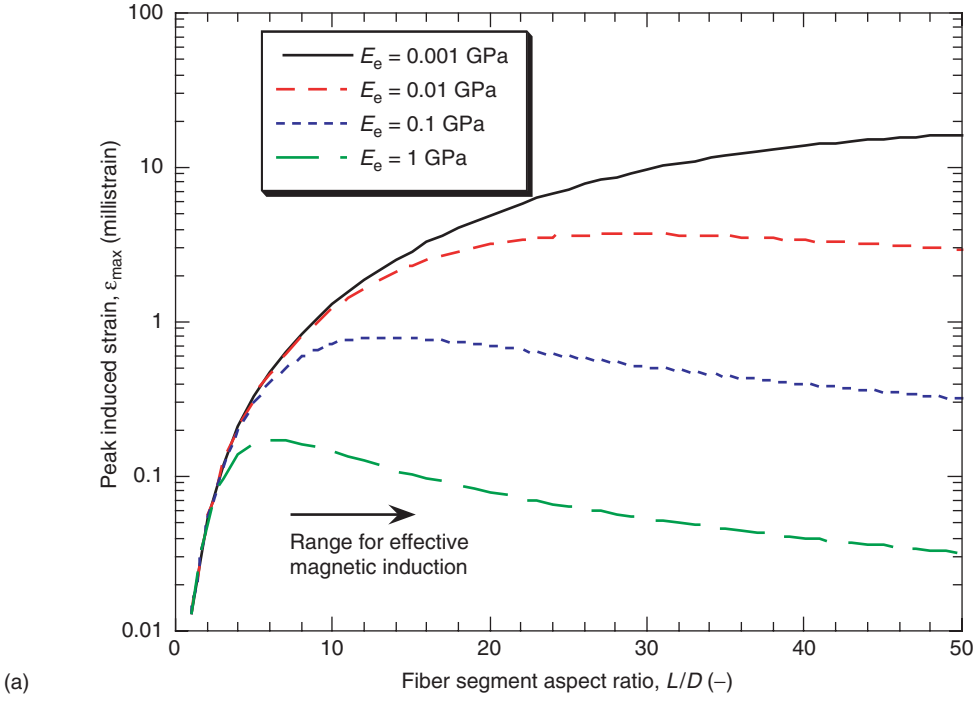
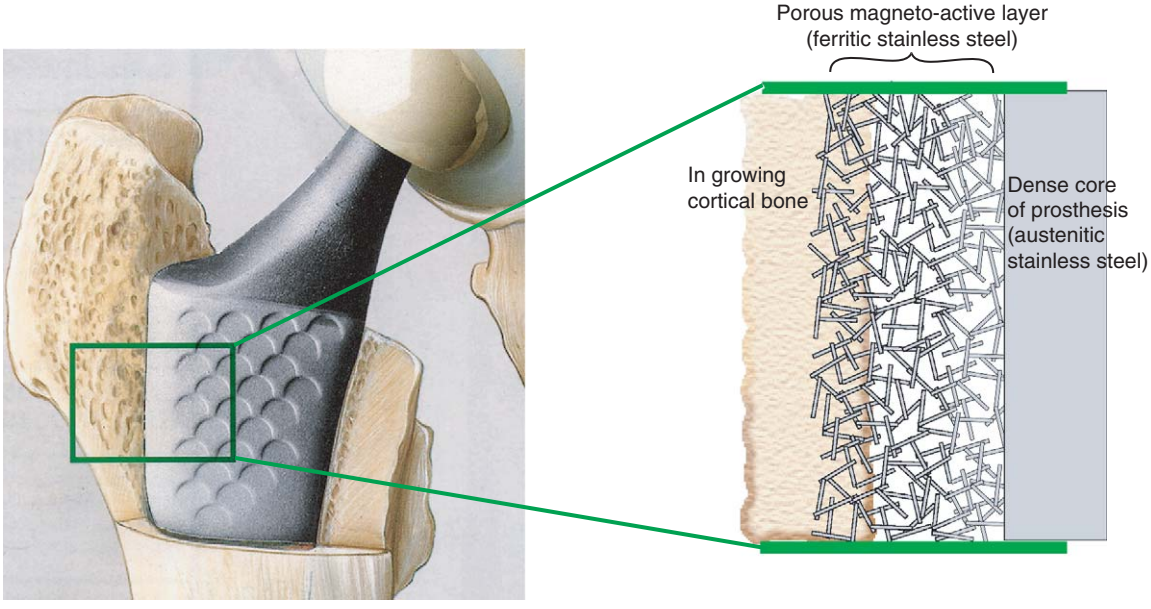


Figure 10 (a) Schematic depiction of how an isotropic porous material made by bonding of ferromagnetic fibers will tend to deform elastically in an applied magnetic field and (b) predicted (eqns [15] and [16]) and measured shape changes ($M_s = 1.6\,\text{MJ}\,\text{T}^{-1}\,\text{m}^{-3}$; $E_f = 210\,\text{GPa}$; $B = 1\,\text{T}$).

Relationships can also be derived for fiber bundles, planar random fiber arrays, woven fabrics, overlapping (poorly bonded) pancakes, etc. Predictions from the Blake–Kozeny equation, for the geometries of eqns [13] and [14], are shown in **Figure 9**, together with representative ranges of experimental data for various types of porous material. It can be seen that these data are, at least, broadly consistent with this simple model, with the permeability falling as the structural scale is refined, and as the void content is reduced. The product of permeability and specific surface area can be taken as a measure of pore



(a)



(b)

Figure 11 Estimated dependence (eqn [17]) on fiber segment aspect ratio, of the maximum strains induced in a compliant surrounding matrix when a ferromagnetic fiber array is subjected to an applied magnetic field and (b) an integrated prosthesis design, incorporating a porous surface layer in which mechanical straining could be induced in this way ($B = 1.5\,\mathrm{T}$; $M_s = 1.6\,\mathrm{MJ\ T^{-1}\ m^{-3}}$; $\theta = 30^\circ$; $E_f = 210\,\mathrm{GPa}$).

connectivity, as plotted in **Figure 1**, but it may be noted that all of the materials represented in **Figure 9** have a relatively high connectivity.

Correlations can be established between permeability and the efficiency of convective heat transfer, but such treatments are beyond the scope of the present article. However, it is worth noting that, in practice, the choice of materials and the design of pore architecture, for applications such as heat exchangers and fuel cell elements are often driven, not solely by transport phenomena optimization, but also by mechanical performance and high temperature stability requirements.

Magnetic Actuation and Magnetomechanical Devices

Porous materials have high potential for usage in various types of actuators, partly because of their relatively low stiffness. A high permeability may also be beneficial in some circumstances – for example, it might facilitate rapid temperature change, perhaps leading to shape changes via differential thermal contraction or the shape memory effect. There is also scope for magnetic actuation with certain types of pore architecture, particularly those containing slender members of ferromagnetic materials, which tend to align with an applied magnetic field. This concept is illustrated in **Figure 10a**, which shows how a material composed of bonded magnetic fibers would deform in such a field. Deformation of this type has recently been analyzed. The shape changes illustrated in **Figure 10a** are predicted to conform to the following equations

$$\frac{\Delta Z}{Z} = \left(\frac{16M_sB}{9E_f}\right)\left(\frac{L}{D}\right)^2 \qquad [15]$$

$$\frac{\Delta R}{R} = \left(\frac{-16M_sB}{9\pi E_f}\right)\left(\frac{L}{D}\right)^2 \qquad [16]$$

where M_s is the saturation magnetization of the fiber material, B is the applied magnetic field, E_f is the Young's modulus of the fiber, and L/D is the aspect ratio of the fiber segments between joints. Predictions are compared with experiment in **Figure 10b**. It can be seen that quite substantial macroscopic strains could be generated, particularly with relatively high fiber segment aspect ratios.

Estimates have also been made of the peak strains that could be induced in this way in a surrounding matrix, such as an embryonic network of bone cells growing into the fiber array. The analysis leads to the following equation for such strains

$$\varepsilon_{max} \approx \frac{12\pi M_sB \sin\theta(L/D)^2}{(9\pi E_f \tan\theta + 28E_e(L/D)^3)} \qquad [17]$$

Predictions are shown in **Figure 11a** for several matrix stiffness levels. Since such a bone network would have a stiffness of the order of 0.01–0.1 GPa, and since strains of at least around 1 millistrain are known to be required for effective mechanical stimulation of bone growth, it can be seen that this model suggests that such stimulation could be generated with a suitable external magnetic field ($B \sim 1.5$ T), provided the fiber segment aspect ratio was at least ~ 5–10. A possible prosthesis design allowing such an effect to be exploited is shown in **Figure 11b**.

See also: Biomedical Materials; Composites: Carbon Matrix; Composites: Overview; Composites: Polymer–Polymer; Conductivity, Electrical; Conductivity, Thermal; Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior.

PACS: 62.20. – x; 65.40. – b

Further Reading

Ashby MF, Evans AG, Fleck NA, Gibson LJ, Hutchinson JW, and Wadley HNG (2000) *Metal Foams: A Design Guide*. Boston: Butterworth-Heinemann.

Banhart J (2001) Manufacture, characterisation and application of cellular metals and metal foams. *Progress in Materials Science* 46: 559–632.

Bardhan P (1997) Ceramic honeycomb filters and catalysts. *Current Opinion in Solid State & Materials Science* 2: 577–583.

Buckley JD, Strouhal G, and Gangler JJ (1981) Early development of ceramic fiber insulation for the space shuttle. *American Ceramic Society Bulletin* 60: 1196–1200.

Fox AC and Clyne TW (2004) Oxygen transport through the zirconia layer in plasma sprayed thermal barrier coatings. *Surface & Coatings Technology* 184: 311–321.

Freyman TM, Yannas IV, and Gibson LJ (2001) Cellular materials as porous scaffolds for tissue engineering. *Progress in Materials Science* 46: 273–282.

Gauckler LJ, Waeber MM, Conti C, and Jacob-Duliere M (1985) Industrial application of open pore ceramic foam for molten metal filtration. *Light Metals* 1261–1283.

Gibson LJ (2000) Properties and applications of metallic foams. In: Clyne TW (ed.) *Comprehensive Composite Materials*, vol. 3, pp. 821–832. Amsterdam: Elsevier.

Gibson LJ and Ashby MF (1999) *Cellular Solids, Structure and Properties*. Cambridge: Cambridge University Press.

Hull D and Clyne TW (1996) *Introduction to Composite Materials*. Cambridge: Cambridge University Press.

Markaki AE and Clyne TW (2001) The effect of cell wall microstructure on the deformation and fracture of aluminium-based foams. *Acta Materialia* 49: 1677–1686.

Markaki AE and Clyne TW (2004) Magneto-mechanical stimulation of bone growth in a bonded array of ferromagnetic fibres. *Biomaterials* 25: 4805–4815.

Richardson JT, Remue D, and Hung J-K (2003) Properties of ceramic foam catalyst supports: mass and heat transfer. *Applied Catalysis A* 250: 319–329.

Saracco G and Specchia V (1998) Catalytic filters for flue-gas cleaning. In: Cybulski A and Moulijn JA (eds.) *Structured Catalysts and Reactors*, pp. 417–434. New York: Dekker.

Scheffler M and Colombo P (eds.) (2005) *Cellular Ceramics. Structure, Manufacturing, Properties and Applications*. Weinheim: Wiley-VCH.

Scheidegger AE (1974) *The Physics of Flow through Porous Media*. Toronto: University of Toronto Press.

Vandeperre LJ, Wang J, and Clegg WJ (2004) Effects of porosity on the measured fracture energy of brittle materials. *Philosophical Magazine* 84: 3689–3704.

Holography

P Ambs, Université de Haute Alsace, Mulhouse, France
J-P Huignard and B Loiseaux, Thales Research and Technology, Orsay, France

Introduction

Holography was discovered in 1948 by Denis Gabor who proposed to use this technique to improve the resolution in electron microscopy. Denis Gabor received the Nobel prize for this work in 1971. The word holography comes from Greek words meaning "total recording," and it is a technique for "wave front reconstruction" where both amplitude and phase of a wave front are recorded. The development of holography started in the beginning of the 1960s when coherent light sources became available with the invention of the laser. In 1962, E Leith and J Upatnieks made a major breakthrough with the invention of the off-axis reference beam optical holography. A few years later, in 1967, A Lohmann presented the first computer-generated hologram. The most spectacular application is the three-dimensional (3D) imaging, but there are now numerous applications where holography is even more useful for nondestructive testing, security, data storage, or for new optical elements, etc.

This article presents the basis of the holographic principles, and then the most commonly used optical setup to record and reconstruct holograms. A short review of the different recording materials and the applications, follows. The two last sections are dedicated to computer-generated holography and to the recent field of digital holography.

Principle of Holography

The purpose of holography is a wave front reconstruction. It is a two-step process, which involves first, the recording of the amplitude and phase of the wave front, and second, its reconstruction. All the available recording materials (photographic plates, photopolymers, photorefractive crystals, etc.) or recording devices (CCD or CMOS cameras) are only sensitive to the intensity of the wave front, and therefore all the phase information is lost. Since the interferences between two coherent waves reflect their phase differences, the idea of recording the whole wave is to record the intensity of the interferences between the desired wave front (object beam) and a known reference wave (reference beam). This can be easily illustrated by the following equation:

$O(x,y)$ called the object beam, is the wave to be recorded:

$$O(x,y) = |O(x,y)| \exp j\Phi(x,y)$$

$R(x,y)$ called the reference beam is a known wave that interferes with the object beam:

$$R(x,y) = |R(x,y)| \exp j\Psi(x,y)$$

$I(x,y)$ is the intensity of the interference pattern between the object and reference beams ($*$ indicates the complex conjugate), and using the simplified notations O and R, one obtains:

$$I(x,y) = |O+R|^2$$

$$I(x,y) = |O|^2 + |R|^2 + OR^* + O^*R$$

It is assumed that the recording medium is an ideal holographic material in which the amplitude transmittance $t(x,y)$ is proportional to the recorded intensity $I(x,y)$. To reconstruct the wave front, the hologram is illuminated by a wave identical to the reference beam and this is given by the following equation:

$$R(x,y)t(x,y) = |O|^2R + |R|^2R + O|R|^2 + O^*R^2$$

These last two terms of this development show that 3D imaging is possible since the complete amplitude and phase of the original wave front $O(x,y)$ is reconstructed. It can also be noted that due to the recording process, each point of the object contributes to the whole surface of the hologram. Therefore, even

a partially illuminated hologram reconstructs the image of the whole object, but the quality of the image is affected.

Principal Holographic Setups

In the previous section, the basic principle of holography was presented and in this section, the three principal setups that are used for optically recording and reconstructing holograms are detailed.

The In-line Hologram

The in-line configuration shown in **Figure 1** is also named the Gabor hologram since D Gabor first proposed this optical setup to demonstrate holography. The object is mostly transparent and may contain opaque particles. It is illuminated by a uniform coherent parallel beam and the transmitted wave is composed of two parts:

1. The wave directly transmitted by the object, noted as R, is constant and forms the reference.
2. The wave scattered by the particles, noted as O, and forms the object beam.

It is assumed that the intensity distribution $I(x, y)$ due to interference of O and R is recorded on a photographic plate whose transmittance $t(x, y)$ is proportional to the intensity and exhibits a constant offset t_0. Then, one has the following relation:

$$t(x, y) = t_0 + kI(x, y)$$

where k is a constant including the exposure time and the slope of the recording material.

To reconstruct the Gabor hologram, it is illuminated with a coherent parallel plane wave R (see **Figure 2**), and the wave transmitted by the hologram is given by

$$Rt(x, y) = R(t_0 + k|R|^2) + kR|O|^2 + k|R|^2O + kR^2O^*$$

The three terms of the reconstruction are presented in **Figure 2**. (1) $R(t_0 + k|R|^2 + k|O|^2)$ is the transmitted beam. (2) $k|R|^2O$ is the original object wave multiplied by a constant. It generates a reconstruction of the object located at its original position behind the hologram. It is a virtual image. (3) kR^2O^* is the complex conjugate of the original object. It generates a real image of the object in front of the hologram.

Application of the Gabor Holograms

The main application of the Gabor holograms is in the particle analysis. In conventional imaging systems, the combination of a large depth of field with a high resolution is not possible. This bottleneck can be

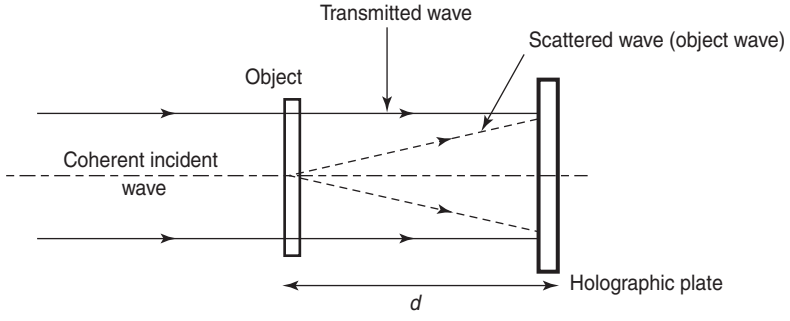


Figure 1 Recording a Gabor hologram.

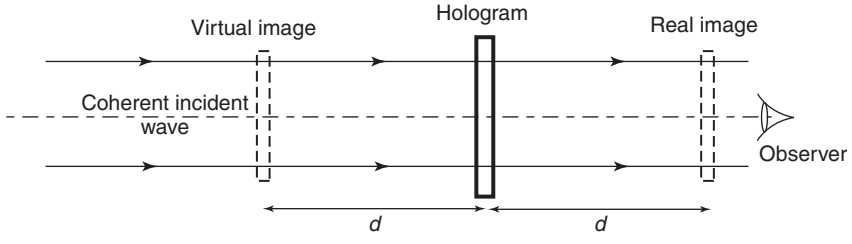


Figure 2 Reconstructing a Gabor hologram.

solved using a holographic recording that can store a high-resolution 3D image. The in-line hologram of the volume of moving particles is recorded using a triggered pulsed laser and the real 3D image can be analyzed by a conventional camera, which images the different planes of the volume. As an example of application, this setup was used to analyze the size and shape of droplets generated by injectors.

The Off-Axis Hologram

With the Gabor hologram, the three reconstructed wave fronts are on the same axis and they cannot be spatially separated. To solve this problem, E Leith and J Upatnieks proposed in 1962 the off-axis recording setup. In this case, the reference beam $R(x,y)$ and the object beam $O(x,y)$ are separated and interfering on the holographic plate under an angle α (see **Figure 3**). The object beam and the off-axis plane wave reference beam are respectively expressed by

$$O(x,y) = |O(x,y)| \exp j\Phi(x,y)$$

$$R(x,y) = |R(x,y)| \exp j2\pi\xi x$$

where

$$\xi = \frac{\sin \alpha}{\lambda}$$

As in previous relations $I(x,y) = |R+O|^2$ is the intensity of the interference fringes.

Following the above formalism, it is clear that when reconstructing this type of hologram with the off-axis reference beam as shown in **Figure 4**, it generates two spatially separated waves $W_1(x,y)$ and $W_2(x,y)$, respectively expressed by

$$W_1(x,y) = k|R|^2 O$$

where $W_1(x,y)$ is the original object wave multiplied by a constant factor and generates a virtual image of the object and $W_2(x,y)$

$$W_2(x,y) = kR^2 O^* \exp(j4\pi\xi x)$$

is the complex conjugate of the original object wave and generates a conjugate real image.

The three waves R, W_1, and W_2 are separated angularly and do not overlap as shown in the Gabor holograms. The off-axis setup is now the most common holographic recording configuration and it is used for reconstructing 3D images.

The Fourier Hologram

In this case, the object is planar and the purpose is to record the amplitude and the phase of its Fourier transform (FT). The holographic technique is used to record this complex wave front. The hologram is formed by the interference pattern between the FT of the object $O(u,v)$ and a tilted wave that is

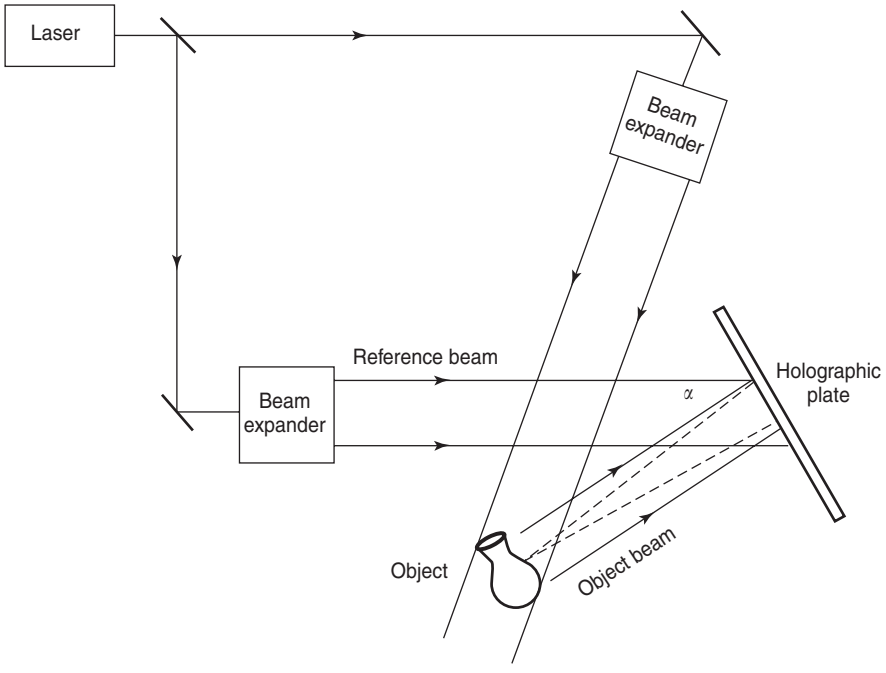


Figure 3 Recording an off-axis hologram.

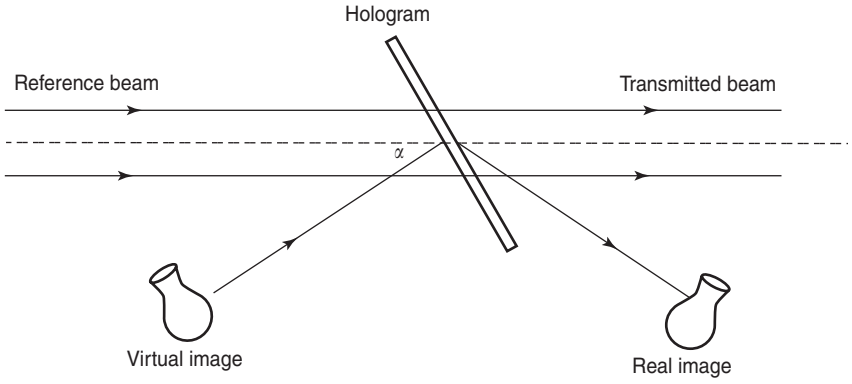


Figure 4 Reconstructing an off-axis hologram.

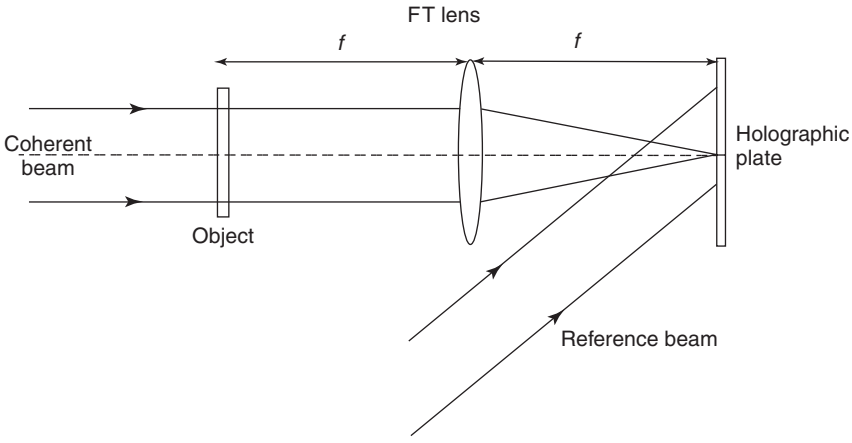


Figure 5 Recording a Fourier hologram.

the reference beam $R(u,v)$ (see **Figure 5**)

$$O(u,v) = FT[O(x,y)]$$
$$R(u,v) = \exp(j2\pi ua)$$

where a is a constant.

$I(u,v)$ is the intensity of the interferences between the object and the reference beams

$$I(u,v) = 1 + |O(u,v)|^2 + O(u,v)\exp(-j2\pi ua) + O^*(u,v)\exp(j2\pi ua)$$

For the reconstruction, the hologram is illuminated by a parallel plane wave (of amplitude unity to simplify) and the wave transmitted by the hologram is

$$U(u,v) = t_0 + kI(u,v)$$

The hologram is placed in the front focal plane of a lens (see **Figure 6**) and therefore the wave in the focal plane is the FT of $U(u,v)$

where

$$\begin{aligned} U(u,v) &= FT[U(u,v)] \\ &= (t_0 + k)\delta(x,y) \\ &\quad + kO(x,y) \otimes O(x,y) \\ &\quad + kO(x-a,y) + kO^*(-x+a,-y) \end{aligned}$$

where $\otimes$ represents the correlation operator.

This diffraction pattern is composed by a central order plus two terms:

1. $kO(x-a,\ y)$ is the $+1$ diffraction order that represents the original object beam separated by a distance a from the central order, and
2. $kO^*(-x+a,-y)$ is the -1 diffraction order that represents the complex conjugate of the original object beam separated by a distance $-a$ from the central order.

These two images of the object transparency are real.

Fourier holograms are not designed for 3D imaging; they have more specific applications such as optical correlation for pattern recognition, high-capacity

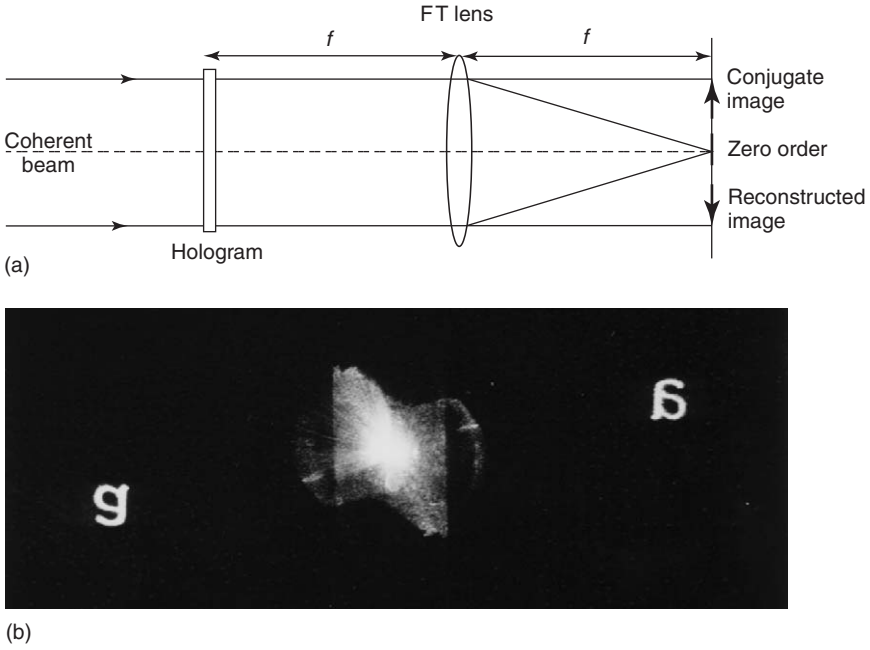


Figure 6 Reconstructing a Fourier hologram (a) the optical setup, and (b) the optical reconstruction of a Fourier hologram.

holographic data storage, 2D imaging, diffractive optical elements, and optical interconnections.

Hologram Characteristics and Recording Procedure

An important parameter of the holographic reconstruction is the diffraction efficiency. It is defined as the ratio between the light beam intensity of the desired reconstructed image and the intensity of the incident beam illuminated on the hologram during the reconstruction.

Depending on the recording material, the hologram can introduce an amplitude or a phase modulation. A spatial phase modulation is often preferred since the diffraction efficiency is higher and can reach hundred percent. If the recording material thickness is small compared to the fringe spacing, a thin grating is recorded and the Raman–Nath diffraction regime applies. The reconstruction produces a large number of diffraction orders and the efficiency gets reduced. On the other hand, a thick volume hologram is recorded if the thickness of the recording material is large compared to the fringe spacing. In such a case, a 3D grating is obtained and the Bragg diffraction regime applies. Phase volume holograms have two very attractive properties: (1) when illuminated by the reference beam, it reconstructs only one beam, and a diffraction efficiency of 100% can be theoretically reached, and (2) angular multiplexing of holograms is possible. In this case, the angle of incidence of the reference beam is different for each of these holograms, which are recorded in the same volume.

Each hologram is then addressed and reconstructed when illuminated by the reference beam with the corresponding incidence angle. It is also possible to multiplex holograms using different wavelengths. Holograms, which can be sequentially stored in the same volume location by varying the reference beam angle at each recording step, provide the capability of high-capacity data storage of digital pages of information.

Also, holograms recorded with the optical setup of **Figure 3** are of the transmission type. However, if a hologram can be recorded by two waves traveling in opposite directions and illuminating respectively the front face and the rear face of the recording medium, the fringes inside the recording medium will be parallel to its faces; this is a reflective type of hologram.

The requirements for optically recording holograms are imposed by the contrast and stability of the interference fringes. The illuminating laser source must be coherent with an adequate coherence length, and the optical setup is stable and isolated from the vibrations. Ideally, the recording media must not only have a high spatial resolution, typically higher than 1000 lines per mm, but also a linear response, a sensitivity range which can match the available laser wavelengths, and a low noise.

Most of the hologram recording materials are nonerasable, such as the commonly used materials listed below:

1. Silver halide emulsions have been used since the beginning of holography. After processing, the light intensity is then transformed into an amplitude modulation. However, it is possible to change this modulation into a phase modulation by bleaching the plate thus allowing for to reach a higher diffraction efficiency to be reached. The use of photographic emulsions implies a chemical wet processing. Silver halide emulsions fulfill most of the requirements for holographic recording media – they are relatively sensitive, have a large spectral response, and a high spatial resolution. Holographic emulsions were easily commercially available until all the large photographic material companies stopped their production several years ago. Now, only a few small companies are still producing holographic plates.
2. Dichromated gelatin is sensitized by ammonium dichromate imbedded into the gelatin, and it produces pure phase holograms. It is a very good material for holography with high resolution ($>10\,000$ lines per mm), a large change of refractive index modulation ($\Delta n > 0.1$), low absorption, and scattering. The use of dichromated gelatin is mastered at the industrial level with well-controlled processes to achieve $\sim 100\%$ diffraction efficiency.
3. Photopolymers are organic materials that transform the light intensity into a refractive index modulation through a photopolymerization process. This is a self-processing photochemical reaction and the hologram is recorded by illuminating the photopolymer, and no further processing is required. Of late, several companies have been selling their photopolymers which exhibit excellent characteristics for high-efficiency holographic components ($\Delta n > 5 \times 10^{-2}$).
4. Photoresists are also organic materials, which are well developed for microelectronic technologies. The illuminating light gives a relief modulation after wet development. Photoresists are used for the fabrication of the masters and for replicating the holograms by embossing.

Erasable holographic materials are suitable for dynamic holography, and the most commonly used materials in this category are photorefractive crystals. In this class of materials, the phase volume holographic grating is due to a photoinduced space charge field which modulates the crystal refractive index through the electrooptic effect. The index modulation can be erased by illuminating the crystal with a uniform beam intensity. Most of the photorefractive crystals also exhibit a dark storage time constant. A significant effort of research is still being done to investigate the phenomena involved in efficient hologram recording in the visible and near infrared spectral range. Photorefractive crystals permit the recording of dynamic holograms with time constants varying from nanoseconds to several seconds, and consequently, the diffracted wave front is probed in real time. Moreover, due to the nonlinear optical behavior of the material, there is an energy redistribution between the two recording beams. In other words, a low-intensity object beam can be amplified through an energy transfer from the reference beam. This unique property of the photorefractive crystals, due to self-diffraction mechanisms, opens up its applications to parametric image amplification, self-induced optical cavities, and phase conjugation with gain. Many other attractive applications can be realized based on the specific characteristics of photorefractive effects in a variety of crystals. Presently, the most important materials used in experiments are: $LiNbO_3$ iron doped; $BaTiO_3$; $KnbO_3$; Bi_{12}(Si, Ge, Ti)O_{20}; for visible and GaAs; InP or CdTe for the near IR. Besides the spectral sensitivity, they differ in terms of speed, dark storage time, diffraction efficiency, and a two-wave mixing gain coefficient.

Applications of Optical Holography

Holography has a large range of applications although it is not as commonly used as it was expected to be in its early stages in the 1960s. Several applications are listed below:

1. Display is the most spectacular and public-oriented application of holography. Large transmissive and reflective holograms have been recorded, and 3D images with a large depth of field have been reconstructed. Different varieties have been developed, particularly in hologram reconstruction with white light, such as the rainbow or color holograms. For color imaging, three holograms of the object are recorded successively on the same plate by three lasers emitting blue, red, and green light respectively. The recording media must have a high resolution and a broad spectral sensitivity range. Such color reflective holograms, reconstruct a 3D color object when illuminated by white light. Since holograms are difficult to reproduce, they are now used as a security feature, on banknotes for example. In this case, the original hologram is recorded optically on a photoresist and used to fabricate a master. The

embossing of a sheet of a reflective material produces the final hologram.

2. Holography also has industrial applications in nondestructive testing, thanks to holographic interferometry. The purpose of classical interferometry is to compare two wave fronts, though most of the time one wave front is tested with a reference one. The principle of the double exposure holographic interferometry is, first, to record holographically the wave front coming from an object at rest. Second, some constraints or stress (mechanical, thermal, etc.) are applied to the object and the new wave front coming from the object is also recorded on the same support. When the resultant hologram is illuminated, it reconstructs the two wave fronts that interfere together and the resultant fringe pattern reveals the modifications of the object with a slight change in its shape. Presently, the principle of holographic interferometry is still used, but the recording and processing are done electronically.
3. Holographic optical elements (HOEs) are also becoming an important industrial application. These elements are recorded optically, mainly on dichromated gelatin, and their purpose is to convert one wave front into another. They can implement, even simultaneously, complex optical functions such as focusing, deflecting, multiplexing, and spectral filtering. One of the applications of HOEs is the head-up display that is used in airplanes and recently in cars.
4. Holographic memories are under study in research laboratories, and an attempt to develop systems are now considered by a few companies. The data pages are written as holograms into a thick layer of photopolymer coated onto a rotating disk. The digital data can be recovered in parallel by reconstructing the holograms page-by-page. Therefore a high data rate can be achieved. The demonstrated capacity of such a 12 cm disk is presently 200 GB, but holographic memories of 1 TB could be expected in the future.

Computer-Generated Holograms and Diffractive Optics

The purpose here is the same as for optically recorded holography. It is to reconstruct a desired wave front, but the process is different since no physical recording process is involved. The wave front is entirely computed and this brings flexibility and versatility.

In the beginning, the word computer-generated hologram (CGH) was used, but presently, it is often referred to as a diffractive optical element (DOE). A Lohmann proposed the first CGH in 1967, and since then computation algorithms and fabrication methods have evolved following the progress of computers as well as the development of new fabrication technologies based on microlithography techniques. Presently, the DOEs can be classified into two classes: analytical-type DOEs and numerical-type DOEs.

Analytical-Type DOEs

The data composing such a DOE are in an analytical form and this is the case for diffractive lenses or gratings. These DOEs are pure phase elements and are fabricated accordingly. There are different approaches for the design of analytical-type DOEs, the direct DOE computation, the interferogram approach, and also the use of optical-design software.

Numerical-Type DOEs

The wave front to be reconstructed cannot be described by an analytical equation but must be computed using the propagation law. If the CGH illuminated by a plane wave reconstructs the desired wave front in the focal plane of a converging lens, the CGH will contain the inverse FT of the desired construction. If the CGH reconstructs without a lens at a given distance, then it will contain the inverse Fresnel transform of the desired reconstruction. The result of the Fourier or Fresnel transform is a 2D array of complex values. As for optically recorded holograms, the recording media are pure amplitude (mostly binary) or pure phase media, and therefore the complex data must be encoded to satisfy the requirements of the fabrication techniques. There are two types of encoding: first one is the cell-oriented encoding, and the second is the point-oriented encoding.

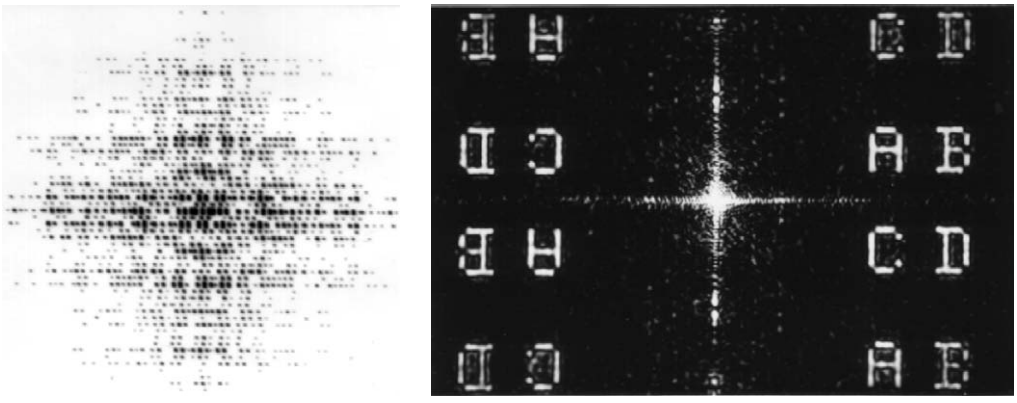
The first CGH presented by A Lohmann used a cell-oriented encoding that he called phase-detour encoding. The hologram is a 2D array of cells representing the array of complex values, and an aperture in each cell represents each complex value. The surface of this aperture is proportional to the amplitude of the complex value to be encoded and its position in relation to the center of the cell is proportional to the phase. An adequate ratio between the aperture width and the cell size introduces the right carrier frequency that allows the separation of the central order from the two reconstructed orders (see **Figure 7**). This encoding, known now as Lohmann's encoding, has given rise to several variations such as the encoding presented by W H Lee with two apertures per cell. The computation of these CGHs does not require powerful computers and is

not time-consuming since only one Fourier or Fresnel transform is calculated. Originally, these CGHs were first drawn on a sheet of paper with a computer-controlled pen plotter and then photoreduced on a high-resolution photographic film.

In the 1980s, thanks to the progress of computers in terms of power and memory capacity, several point-oriented encoding methods were developed. The principle of the three main methods is described below. For these algorithms, the desired reconstruction window is placed into a larger 2D array, therefore the hologram and also the reconstruction plane is larger than the reconstruction window. This extra space allows a better convergence of the computation algorithm due to the phase and amplitude freedom outside the reconstruction window. Amplitude CGHs construct the desired reconstruction, its complex conjugate, and a zero order. In order to separate these terms, the reconstruction window should be placed at some distance from the center of the reconstruction plane. For phase CGHs with more than three phase levels, only the desired reconstruction window appears in the reconstruction plane with no zero and no complex conjugate orders. However, this happens only if the computed phase levels are perfectly reproduced by the CGH.

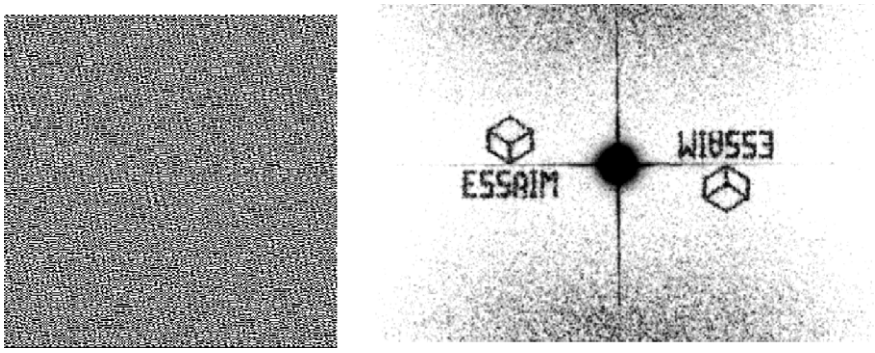
The error diffusion algorithm consists of using a pulse width modulation technique for binarizing the hologram. In the case of binary amplitude CGHs, the real part of the Fourier or Fresnel transform is made positive by adding an offset, and then normalized. Each of these sampled real values are binarized and the error, that is, the difference between the binary value and the analog value, is divided and distributed with a weight to the unprocessed neighbors before proceeding to the next point. By choosing the weight and the direction of the error diffusion carefully, it is possible to separate the quantization noise from the desired reconstruction. This algorithm has given rise to a lot of related algorithms, such as the complex error diffusion algorithm (see **Figure 8**).

The next two encoding methods that are presented are iterative methods, they give better results in terms



(a) (b)

Figure 7 Lohmann hologram: (a) the negative of the hologram and (b) optical reconstruction.



(a) (b)

Figure 8 Error diffusion encoding (a) the CGH and (b) the optical reconstruction.

of signal to noise ratio (SNR) but their computation is more time-consuming.

The direct binary search (DBS) algorithm presented by Seldowitz and co-workers in 1987 is based on the optimization of an error criterion by the use of a Monte Carlo technique. To begin with, a random hologram is generated, and the error criterion between the desired image and the reconstruction of this hologram is computed. The random CGH is then scanned, and a new error is computed after inversion of a new pixel. Pixel modification is retained only if the error decreases, and the error criterion is thereby minimized. This cycle is stopped when the error stops decreasing. The error criterion can be defined for amplitude, intensity, or a complex image, depending on the application using the CGH. For some applications, the CGH must reconstruct a wave front with a specified amplitude and phase, and the error must then be based on a complex-amplitude error criterion. However, in many other applications (display, memory, interconnect, etc.) the CGH reconstruction is detected by a quadratic detector, leaving the phase as a free parameter, and the error criterion can be based on amplitude or intensity errors. Since these errors use the phase freedom, the optimization algorithms converge better in this case. The original DBS algorithm was designed for binary CGHs; however, it can be extended to multilevel amplitude or phase CGHs at the expense of the computation time, since for each pixel, all levels must be tested.

The iterative Fourier transform algorithm (IFTA) presented in 1988 by F Wyrowski, is an iterative technique based on the Gerchberg–Saxton algorithm. Schematically, the principle of this method consists in going successively from the object plane to the spectrum plane by the application of constraints on the CGH (see **Figure 9**). To be recorded, the CGHs must be quantized and often binarized, but IFTA can also be used to compute multilevel amplitude or phase CGHs. In the Fourier domain, the spectrum of the object is thus constrained to be a binary function. The object constraint must correct the quantization noise by injection of the desired image intensity in the signal windows. The inverse FT of this spectrum gives back an object plane in which the object constraint is thus applied to match the desired image intensity in the signal windows. However, to avoid stagnation of the algorithm the quantization of the spectral plane must be progressive, as is illustrated in **Figure 9**, where three different steps of the binarization process are shown. In order to make its spectrum as uniform as possible, an equalizer can be applied to the desired signal. The simplest equalizer adds a random diffuser to the object plane.

Quality DOEs imply a very precise manufacturing with a perfect match between the computed phase and the actual phase of the component. To achieve these requirements, microlithography techniques are used to etch very precisely, in shape and depth, the chosen optical substrate such as quartz. Embossing techniques make cheaper DOEs, but the quality of their reconstruction is lower, particularly as the phase mismatch generates a zero order.

CGHs have a broad range of applications such as beam shaping, Fourier filtering, 2D or 3D display, optical interconnects, and complex optical functions generation (**Figure 10**).

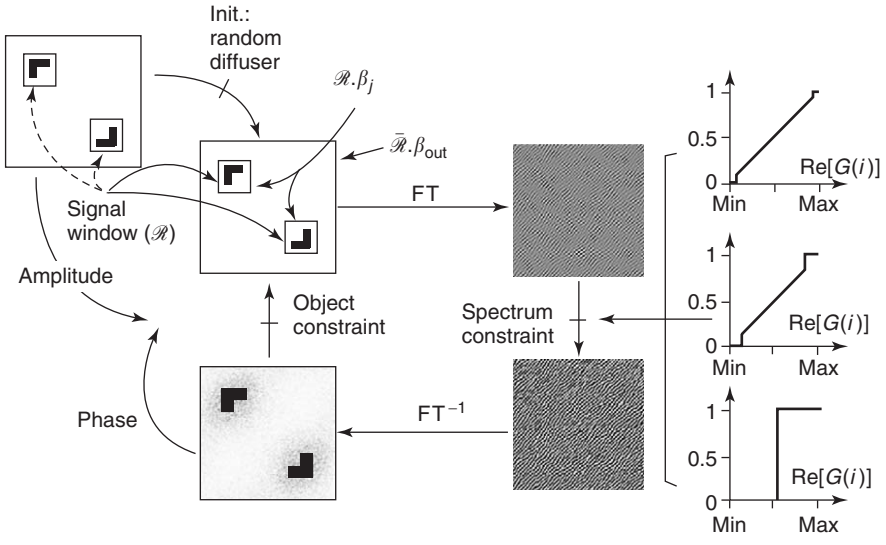


Figure 9 IFTA for computing a binary amplitude CGH. (After Wyrowski.)

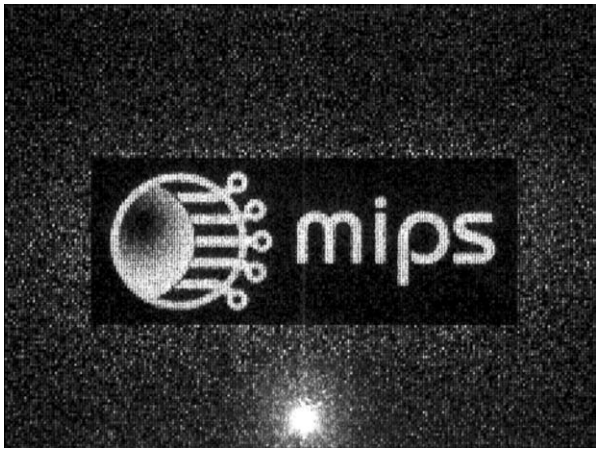


Figure 10 The optical reconstruction of an IFTA encoded binary amplitude CGH. Only the first reconstruction order is shown along with the zero order. The reconstructed image is 153×60 pixels, and a 256×256 pixel hologram was computed and replicated to fill an 800×600 array. This replication reduces the speckle noise into the reconstruction.

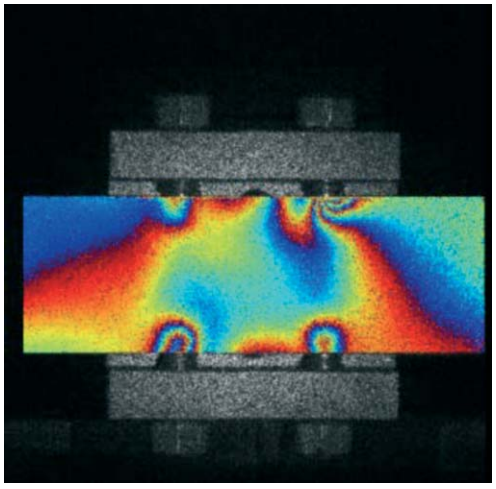


Figure 11 Applications of digital holography; electronic speckle pattern interferometry for the measurement of the deformation of a rubber piece. (Source: Thales RT.)

Digital Holography

In digital holography, the recording media is replaced by a high-resolution electronic camera. Using data from the recorded interference pattern, the object is reconstructed by a digital computer. Digital holography is a fast growing field, because it suppresses all the difficulties related to the processing of the traditional holographic recording media and brings in a lot of versatility and flexibility. Digital holography is now used in industrial applications requiring a quantitative measurement of a wave front where it is successfully replacing the classical setups. **Figures 11** and **12** show, respectively, the measurement of deformations and of vibrations using digital holography.

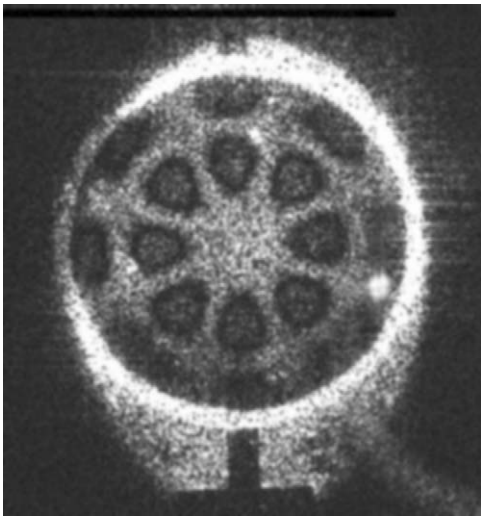


Figure 12 Applications of digital holography; MEMS membrane vibration measurement using time averaging interferometry (scale: 1 fringe $= 0.3\,\mu$m, excitation 830 kHz). (Source: Thales RT.)

See also: Electrodynamics: Continuous Media; Electrodynamics: Quantum Electrodynamics; Geometrical Optics; Optical Absorption and Reflectance; Optical Instruments.

PACS: 42.40.Jv; 42.40.Pa; 42.40.Eq; 42.40.−i; 42.70.Ln; 42.25.Fx; 42.40.Kw; 42.65.Hw; 42.70.Nq

Further Reading

Collier RJ, Burckhardt CB, and Lin LH (1971) *Optical Holography*. New York: Academic Press.

Coufal HJ, Psaltis D, and Sincerbox GT (2000) *Holographic Data Storage*. Berlin: Springer.

Goodman JW (1968) *Introduction to Fourier Optics*. New York: McGraw-Hill.

Hariharan P (1984) *Optical Holography, Principles, Techniques, Applications*. Cambridge: Cambridge University Press.

Hecht E (2002) *Optics*. San Francisco: Addison Wesley.

IIzuka K (1987) *Engineering Optics*. Berlin: Springer.

Kress B and Meyrueis P (2000) *Digital Diffractive Optics, An Introduction to Planar Diffractive Optics and Related Technology*. Chichester: Wiley.

Saxby G (1988) *Practical Holography*. New York: Prentice-Hall.

Smigielski P (1994) *Holographie Industrielle*. Toulouse: Teknea.

Turunen J and Wyrowski F (1997) *Diffractive Optics for Industrial and Commercial Applications*. Berlin: Akademie Verlag.

Imaging and Optical Processing

F Gori, University of Rome "Roma" Tre, Rome, Italy

Introduction

The most elementary treatment of imaging is based on geometrical optics. In this realm, the propagation of light is described through rays, the lines along which luminous energy is carried. Rays are rectilinear in homogeneous media, whereas they are bent toward the higher refractive index regions in nonhomogeneous media. A geometrical analysis gives a wealth of useful results but conceals the most subtle effects connected to wave propagation and diffraction. On analyzing these effects, it is found that an optical system acts as a low-pass filter with respect to the spatial frequency spectrum of the input. This, in turn, suggests that other filtering operations can be realized, thus introducing the idea of optical processing.

The basic elements of the geometrical analysis are considered first, using a matrix form that, to a certain extent, can be transferred into the ondulatory treatment. Then, the wave behavior and its implications for optical processing are examined.

Ray Matrices

The most common optical systems are endowed with a cylindrical symmetry. At the simplest level, rays are assumed to travel near the symmetry axis, say the z-axis, with little inclination with respect to it. This is known as the paraxial hypothesis. Attention is focused on the so-called meridional rays, that is, those contained in a plane passing through the z-axis.

At a typical plane orthogonal to the z-axis, a meridional ray can be specified (**Figure 1**) by a two-element matrix of the form

$$\begin{pmatrix} r \\ n\vartheta \end{pmatrix} \qquad [1]$$

where r is the distance between the intersection point and the z-axis, ϑ is the angle between the ray and the z-axis, and n is the local refractive index. The introduction of n into the matrix simplifies the formulas to follow.

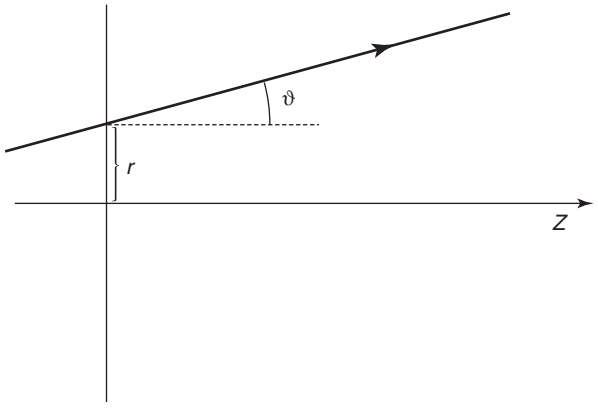


Figure 1 Specifying a meridional ray at a typical plane orthogonal to the symmetry axis (z-axis).

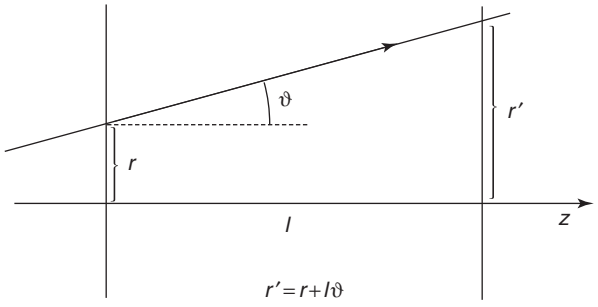


Figure 2 Effect of free propagation in a homogeneous slab of length l. The distance r varies, while the inclination θ remains unchanged.

An ordinary optical system is a sequence of homogeneous media (generally air and glasses) separated by spherical surfaces whose centers lie on the z-axis. Consider a meridional ray traveling within such a system. In the homogeneous sections (**Figure 2**), there is a continuous change of r (except for $\vartheta = 0$) without any change of ϑ. On the contrary, when the ray intersects a spherical surface between media with different refractive indices (**Figure 3**), the angle ϑ changes while r remains unaltered. These modifications are easily evaluated by means of geometrical considerations and by taking into account the

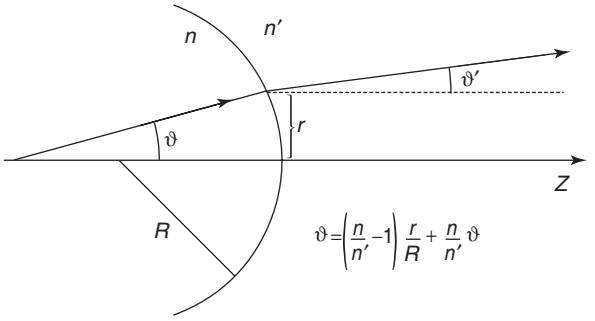


Figure 3 Ray passing through a spherical surface of radius R dividing two media with different refractive indices (n, n'). The inclination θ varies, while the distance r remains unchanged.

Snell–Cartesius law of refraction. Since the small-angle approximation can be used, the ray parameters change according to a projective transformation of the form

$$\begin{aligned} r' &= Ar + Bn\vartheta \\ n'\vartheta' &= Cr + Dn\vartheta \end{aligned} \qquad [2]$$

where the quantities A, B, C, and D depend on the optical system. This leads one to represent the relation between input and output rays in the matrix form

$$\begin{pmatrix} r' \\ n'\vartheta' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r \\ n\vartheta \end{pmatrix} \qquad [3]$$

The so-called $ABCD$ matrix of a system can be derived starting from two basic ones. First, there is the matrix relating to free propagation in a homogeneous section of length l and refractive index n. This is denoted by M_{fp}. Second, there is the matrix, say M_{sph}, that describes the passage through a spherical surface dividing two media with refractive indices n (input side) and n' (output side). Let R be the radius of curvature of the surface. The usual convention is to take R as positive if the ray impinges on the convex side of the surface. Matrices M_{fp} and M_{sph} are easily evaluated and turn out to be

$$M_{\text{fp}} = \begin{pmatrix} 1 & l/n \\ 0 & 1 \end{pmatrix}, \qquad M_{\text{sph}} = \begin{pmatrix} 1 & 0 \\ (n-n')/R & 1 \end{pmatrix} \qquad [4]$$

Since in a typical optical system, there is a sequence of the above basic processes, the overall $ABCD$ matrix is obtained by multiplying the matrices pertaining to the single elements in the right order.

It is to be noted that the determinants of both M_{fp} and M_{sph} are equal to 1. The same holds true for any of their products, so that one can express, in a general way

$$AD - BC = 1 \qquad [5]$$

As a simple example, consider a lens with refractive index n in air (whose refractive index is ~ 1). Denote by R_1 and R_2, the radii of the first and the second limiting surfaces of the lens. On assuming the lens to be thin, the propagation from the first to the second limiting surface can be neglected. Using eqns [4], the $ABCD$ matrix, say M_{l}, of the lens is found to be

$$\begin{aligned} M_{\text{l}} &= \begin{pmatrix} 1 & 0 \\ -\left(\dfrac{1}{R_1} - \dfrac{1}{R_2}\right)(n-1) & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \end{aligned} \qquad [6]$$

where f is the focal length of the lens.

The $ABCD$ formalism can be extended to optical systems including more sophisticated elements. For example, a GRIN (graded index) lens is made by a cylindrically symmetric medium in which the refractive index decreases parabolically as a function of the distance from the axis, according to a law of the form

$$n(r) = n_0 - \frac{n_2}{2} r^2 \qquad [7]$$

where n_0 and n_2 are positive constants. Equation [7] is assumed to hold for a certain range of values of r. It is further assumed that all rays of interest travel within such a range. Letting $\gamma^2 = n_2/n_0$, it can be shown that a GRIN lens of length l is characterized by a matrix of the form

$$M_{\text{GRIN}} = \begin{pmatrix} \cos \gamma l & (n_0\gamma)^{-1} \sin \gamma l \\ -(n_0\gamma) \sin \gamma l & \cos \gamma l \end{pmatrix} \qquad [8]$$

It should be noted that the present formalism does not account for the presence of any transversal limitation to the ray path (stops, pupils). Similar limitations are taken into account at a more refined stage of analysis. Furthermore, the behavior of rays whose inclinations with respect to the optical axis are not small can be studied by the so-called ray tracing process, in which the path of a ray is followed without introducing the small angle approximation. Ray tracing can also be applied to nonmeridional or skew rays. Powerful computer programs are available for this purpose.

Geometrical Imaging

The imaging condition requires that all rays starting from a fixed point of the input plane are led to pass through one and the same point at the output plane. Limiting to the meridional rays, this amounts to saying that the position r' has to be independent of the angle ϑ. The first of eqns [2] then gives the imaging condition

$$B = 0 \qquad [9]$$

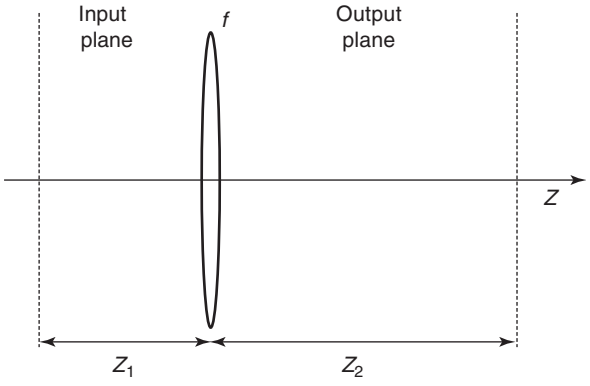


Figure 4 Geometrical imaging through a thin converging lens with focal length *f*.

If this condition is satisfied, then the relation between r and r' becomes

$$r' = Ar \qquad [10]$$

so that A is the magnification of the image with respect to the object.

A simple example is worked out by considering the case (**Figure 4**) in which three processes take place between the input and output planes: (1) free propagation in air for a length z_1, (2) passage through a lens with focal length f, and (3) free propagation in air for a length z_2. Using eqns [4] and [6], the $ABCD$ matrix for the system is evaluated as

$$M = \begin{pmatrix} 1 & z_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix} \qquad [11]$$

where, as before, the refractive index of air has been set ~ 1. On expanding the product,

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 - z_2/f & z_1 + z_2 - z_1 z_2/f \\ -1/f & 1 - z_1/f \end{pmatrix} \qquad [12]$$

The imaging condition [9] now reads

$$z_1 + z_2 - z_1 z_2/f = 0 \rightarrow 1/z_1 + 1/z_2 = 1/f \qquad [13]$$

and this is nothing else but the well-known thin lens equation. According to eqns [12] and [13], the magnification is

$$A = 1 - z_2/f = -z_2/z_1 \qquad [14]$$

Wave Analysis

The question now is how the treatment has to be modified when one passes from a geometrical to a wave analysis. The coherent, monochromatic case is considered first. In a scalar description, the field is specified by a single function of the spatial coordinates. It is understood that the time dependence is of the form $\exp(-i\omega t)$, ω being the angular frequency. Again, using a paraxial hypothesis, the Fresnel approximation can be adopted. Basically, in the Fresnel approach, an approximate expression for spherical waves is introduced. More explicitly, consider the field distribution produced across a plane $z = \text{const.}$ by a point-like source placed at the origin of the reference frame. Denoting by R the distance from the origin of a typical point of such a plane, the field is proportional to $\exp(ikR)/R$, where k is the wave number, that is, $2\pi/\lambda$, λ being the wavelength. Let $R = \sqrt{z^2 + \rho^2}$, where ρ is the distance from the z-axis. For small ρ, that is, in a paraxial neighborhood of the z-axis, the following approximation can be used: $\exp(ikR)/R = \exp[ikz + ik\rho^2/(2z)]/z$. This gives rise to the celebrated Fresnel diffraction formula (see "Further reading" section). In most textbooks, such a formula is derived for free propagation only. Yet, it can also be used for propagation through an optical system described by an $ABCD$ matrix, as was shown by Collins long ago.

Denoting the position vectors at the input and output planes by $\boldsymbol{\rho}$ and $\boldsymbol{r}$, the Collins formula reads

$$V(\boldsymbol{r}) = -\frac{ie^{ikz_t}}{\lambda B} \exp\left(\frac{ikD}{2B} r^2\right) \times \int\int V_0(\boldsymbol{\rho}) \exp\left[\frac{ik}{2B}(A\rho^2 - 2\boldsymbol{r}\cdot\boldsymbol{\rho})\right] d^2\rho \qquad [15]$$

Here, V_0 and V are the (scalar) field distributions across the input and output planes respectively. Finally, z_t is the overall optical path between the two planes.

Fourier Transformation by Lenses

A fundamental property of lenses is their ability to perform Fourier transformations. To see this, reconsider the example of a single thin lens between two layers of air. Suppose $z_2 = f$. Then eqn [12] becomes

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 1 - z_1/f \end{pmatrix} \qquad [16]$$

Accordingly, eqn [15] gives

$$V(\boldsymbol{r}) = -\frac{ie^{ik(z_1+f)}}{\lambda f} \exp\left[\frac{ik}{2f}\left(1 - \frac{z_1}{f}\right) r^2\right] \times \int\int V_0(\boldsymbol{\rho}) \exp\left(-2\pi i \frac{\boldsymbol{r}}{\lambda f}\cdot\boldsymbol{\rho}\right) d^2\rho \qquad [17]$$

This can be written

$$V(\boldsymbol{r}) = -\frac{ie^{ik(z_1+f)}}{\lambda f} \exp\left[\frac{ik}{2f}\left(1 - \frac{z_1}{f}\right) r^2\right] \tilde{V}_0\left(\frac{\boldsymbol{r}}{\lambda f}\right) \qquad [18]$$

where the tilde denotes the Fourier transformation. It can then be seen that, whatever the value of z_1, in the back focal plane of the lens the field distribution contains a factor representing the Fourier transform of the input field. The frequency at which such a transform is evaluated, the so-called spatial frequency, is $\boldsymbol{r}/(\lambda f)$. The other function of r appearing on the right-hand side of eqn [18] represents, in the paraxial approximation, a spherical wave. Therefore, one can say that the field at the back focal plane of the lens is a spherical wave that is transversally modulated by the Fourier transform of the input field. It can be observed that the spherical wave term is irrelevant if only the intensity is of interest, because in that case the phase term disappears (the intensity being proportional to the squared modulus of the field).

It is seen from eqn [18] that a pure Fourier transform is performed by the lens, up to a known constant complex factor, if the input field is fed at the front focal plane of the lens ($z_1 = f$). This has important implications for optical processing. Note that in this case the $ABCD$ matrix becomes

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 0 \end{pmatrix} \qquad [19]$$

Wave Imaging

The treatment based on geometrical optics gives the imaging condition $B = 0$. There is no *a priori* guarantee that such a condition holds true when the imaging process is analyzed in terms of waves. This point is now discussed.

It is seen from eqn [15] that the case $B = 0$ has to be dealt with by means of a suitable limiting procedure. For the moment, B is assumed to be different from zero. On completing the square in the exponential within the integral, one can write eqn [15] in the form

$$V(\boldsymbol{r}) = -\frac{\mathrm{i}e^{\mathrm{i}kz_t}}{\lambda B}\exp\left[\frac{\mathrm{i}\pi}{\lambda B}(D - 1/A)r^2\right] \times \int\int V_0(\boldsymbol{\rho})\exp\left[\frac{\mathrm{i}\pi A}{\lambda B}(\boldsymbol{\rho} - \boldsymbol{r}/A)^2\right]\mathrm{d}^2\rho \qquad [20]$$

Thanks to eqn [5]

$$\frac{1}{B}\left(D - \frac{1}{A}\right) = \frac{C}{A} \qquad [21]$$

so that eqn [20] becomes

$$V(\boldsymbol{r}) = -\frac{\mathrm{i}e^{\mathrm{i}kz_t}}{\lambda B}\exp\left(\frac{\mathrm{i}\pi C}{\lambda A}r^2\right)\int\int V_0(\boldsymbol{\rho}) \times \exp\left[\frac{\mathrm{i}\pi A}{\lambda B}(\boldsymbol{\rho} - \boldsymbol{r}/A)^2\right]\mathrm{d}^2\rho \qquad [22]$$

Recall that one of the definitions of the (two-dimensional) Dirac delta function reads

$$\lim_{\alpha\to 0}\frac{-\mathrm{i}}{\alpha}\exp\left[\frac{\mathrm{i}\pi}{\alpha}(\boldsymbol{\rho} - \boldsymbol{\rho}_0)^2\right] = \delta(\boldsymbol{\rho} - \boldsymbol{\rho}_0) \qquad [23]$$

One can use this relation on passing to the limit for $B \to 0$ in eqn [22]. This gives

$$\lim_{B\to 0} V(\boldsymbol{r}) = -\frac{e^{\mathrm{i}kz_t}}{A}\exp\left(\frac{\mathrm{i}\pi C}{\lambda A}r^2\right)V_0\left(\frac{\boldsymbol{r}}{A}\right) \qquad [24]$$

It is interesting to compare the present result with the geometrical prediction. The field specified by eqn [24] accounts for the same magnification factor A predicted by the geometrical analysis. In general, however, it cannot be said to be an image of V_0 at the field level because it is multiplied by a quadratic exponential term that describes a spherical wavefront. The only exception occurs when $C = 0$. On the other hand, if the intensity only is of interest, the presence of the exponential term is immaterial.

It is not difficult to devise an optical system that gives rise to a perfect image at the field level. It has been seen that the passage from the front to the back focal plane of a single thin lens is associated to an anti-diagonal $ABCD$ matrix (see eqn [19]). Now, it is well known that the product of two anti-diagonal matrices leads to a diagonal one. This ensures that the condition $C = 0$ is met. Therefore, one considers the sequence of two thin lenses (**Figure 5**), whose focal lengths are denoted by f_1 and f_2, in which the back focal plane of the first lens coincides with the front focal plane of the second. Such a configuration is called telescopic. The overall $ABCD$ matrix of the system is

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & f_2 \\ -1/f_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & f_1 \\ -1/f_1 & 0 \end{pmatrix} = \begin{pmatrix} -f_2/f_1 & 0 \\ 0 & -f_1/f_2 \end{pmatrix} \qquad [25]$$

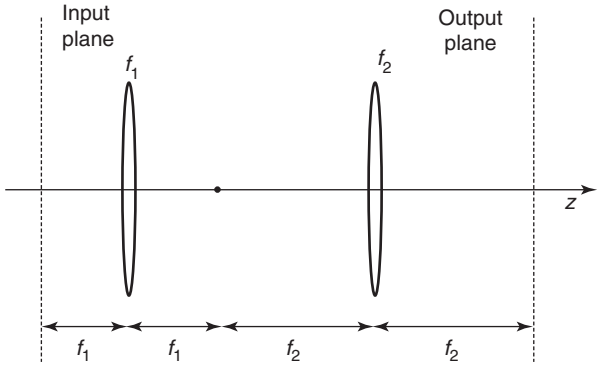


Figure 5 Wave imaging through a telescopic system.

This gives rise to a perfect image with magnification $-f_2/f_1$. Essentially, the same conclusion can be reached by noting that two cascaded Fourier transforms amount to replicating the input field up to a reversal of the coordinate axes.

Pupils

Up to now, it has been assumed that no transversal limitations were present. Obviously, in any real optical system there are diaphragms of various types. The procedure to take into account their effect is essentially to trace the paths of the rays in order to find the diaphragm that imposes the most severe limitation. The image of such a diaphragm formed by that part of the optical system that lies beyond it (along the light propagation direction) is called the exit pupil. Consider a point source at the entrance plane now. In the ideal case, its image would be a delta-like field distribution at the output plane. In other words, there would be a spherical wave converging to a certain image point with a position vector, say $\boldsymbol{r}_0$. In the actual case, such a wave is limited by the exit pupil. Denote by $H(\boldsymbol{r}, \boldsymbol{r}_0)$, the field across the output plane. Taking into account the first of eqns [4] and letting $n = 1$, eqn [15] becomes

$$H(\boldsymbol{r},\boldsymbol{r}_0) = -\frac{\mathrm{i}e^{\mathrm{i}k[d+r^2/(2d)]}}{\lambda d}\int\int p(\boldsymbol{\rho}) \times \exp\left[-\frac{\mathrm{i}k}{2d}(\boldsymbol{\rho}-\boldsymbol{r}_0)^2\right] \times \exp\left[\frac{\mathrm{i}k}{2d}(\rho^2 - 2\boldsymbol{r}\cdot\boldsymbol{\rho})\right]\mathrm{d}^2\rho \qquad [26]$$

where d is the distance between the exit pupil and the image plane. The first quadratic exponential function within the integral describes a spherical wave that converges toward the point $\boldsymbol{r}_0$. The function $p(\boldsymbol{\rho})$ is known as the transmission function of the exit pupil. For any field impinging on the pupil, it represents the pointwise ratio between the emerging and the incident fields. This function is either 1 (within the clear region of the pupil) or zero. Nonetheless, in most sophisticated cases, it could be a complex function whose modulus changes between zero and 1 in a continuous way. Rearranging terms in eqn [26], one obtains

$$H(\boldsymbol{r},\boldsymbol{r}_0) = \exp\left[\frac{\mathrm{i}k}{2d}(r^2 - r_0^2)\right]\int\int p(\boldsymbol{\rho}) \times \exp\left[-\frac{\mathrm{i}k}{d}\boldsymbol{\rho}\cdot(\boldsymbol{r}-\boldsymbol{r}_0)\right]\mathrm{d}^2\rho \qquad [27]$$

Since H gives the field at the output when the input is a delta-like function, it is called a coherent impulse response. For any reasonably good optical system, the integral is appreciable only in a small neighborhood of $\boldsymbol{r}_0$. Accordingly, the exponential in front of the integral can be approximated by 1 and eqn [27] can be written as

$$H(\boldsymbol{r},\boldsymbol{r}_0) = K\tilde{p}\left(\frac{\boldsymbol{r}-\boldsymbol{r}_0}{\lambda d}\right) \qquad [28]$$

where K is a constant. This means that the coherent impulse response is proportional to the Fourier transform of the pupil transmission function. Furthermore, it is shift-invariant, in the sense that it depends only on $\boldsymbol{r}-\boldsymbol{r}_0$.

It is now possible to write the image formation law for any coherent input field distribution. In fact, in a linear system endowed with a shift-invariant impulse response, the relation between input and output is simply a convolution. As a consequence, denoting again the input and output fields with V_0 and V, the image formation law reads

$$V(\boldsymbol{r}) = \int\int V_0(\boldsymbol{r}_0)H(\boldsymbol{r}-\boldsymbol{r}_0)\,\mathrm{d}^2r_0 \qquad [29]$$

Taking into account the convolution theorem for the Fourier transforms, it is seen that the optical system acts as a filter on the spatial frequency spectrum of the input field. The associated transfer function, which is the Fourier transform of H, is known as the coherent transfer function. Up to a multiplicative constant and a coordinate scale factor, it is given by the pupil function itself. Typically, the system behaves as a low-pass filter.

Incoherent Imaging

In many instances, the object to be imaged behaves as a spatially incoherent source. This means that any point-like element of the object radiates in an uncorrelated way with respect to the others. As a consequence, the image is obtained by summing up the intensity contributions due to the various elements. In other words, the image formation law is formulated as a relation between the intensities across the object and the image, instead of a relation between the associated fields. As for the temporal behavior of the radiation field, a quasi-monochromaticity hypothesis is adopted (a strictly monochromatic field would be spatially coherent). The so-called incoherent impulse response is then the squared modulus of the coherent impulse response [28] and the image formation process is described by the law

$$|V(\boldsymbol{r})|^2 = \int\int |V_0(\boldsymbol{r}_0)|^2|H(\boldsymbol{r}-\boldsymbol{r}_0)|^2\mathrm{d}^2r_0 \qquad [30]$$

This convolution relation implies, as in the coherent case, a filtering process. The corresponding incoherent

transfer function is the Fourier transform of $|H|^2$. Therefore, it is proportional to the autocorrelation of the pupil function. It is to be borne in mind that the quantity by which the process is described is the intensity, not the field. As a consequence, a comparison with the coherent case is difficult to establish.

Optical Processing

The basic ideas underlying the techniques of optical processing can be grasped by means of what has already been discussed with reference to the optical systems. For the sake of simplicity, refer to the telescopic setup examined in the section "Wave imaging" (**Figure 5**), assuming coherent monochromatic illumination. Furthermore, the magnification is supposed to be -1. Consider first the ideal case in which transversal limitations can be neglected. Then, through two cascaded Fourier transformations, the input field is replicated at the output plane up to a reversal of the coordinate axes. Suppose now that a circular diaphragm is centered at the origin of the intermediate plane, where the Fourier transform of the input field is displayed. This may be called the Fourier plane. The effect of the diaphragm is to cut off from the object spectrum, spatial frequencies above a certain maximum. In other words, the diaphragm acts as a low-pass filter. Much in the same way, one can perform other types of filtering operations by inserting suitable filters in the Fourier plane. Some examples are examined now.

Visualization of Phase Objects

Certain objects, when traversed by a coherent light field, alter essentially the transverse phase distribution of the field leaving the amplitude pattern unchanged except for a transversally uniform attenuation factor due to reflection. Simple examples are given by dielectric plates whose thickness changes from one point to another, by biological specimens in which the refractive index is spatially varying, or by turbulent fluid layers. Objects with this property are briefly called phase objects. Suppose an object of this type is present at the input plane of the telescopic system discussed above. Suppose further that the object is orthogonally illuminated by a coherent plane wave. Obviously, if one looks at the output plane, one sees a uniform illumination. The problem then arises as to how to transform the phase distribution into an intensity pattern, so that the phase changes can be detected by an intensity sensitive detector (visual system, photographic emulsions, TV-cameras, etc.). This problem, called the phase visualization problem, can be solved by introducing suitable filters in the Fourier plane. Typically, the phase variations across the objects to be visualized are much smaller than 2π. Consequently, the object spectrum seen at the Fourier plane has a strong peak centered at the origin. In addition, some energy is spread across the plane because of diffraction by the phase modulations. Then, a very simple way to visualize the phase object is to absorb the central peak by means of a small opaque disk. Only the diffracted light caused by the phase variations passes around the disk and reaches the output plane. There, an image of the phase distribution across the object is formed. This is known as the central stop method. From the experimental point of view, an even simpler procedure is to insert a knife edge in the Fourier plane to intercept the central spot as well as half the object spectrum. One then speaks of the Schlieren method.

A more sophisticated solution is the phase-contrast technique invented by Zernike and widely used in microscopy. In order to explain this technique, one may express the field emerging from a phase object under illumination by an orthogonal monochromatic plane wave. Such a field has the form

$$V_0(\boldsymbol{r}) = A\,\exp[\mathrm{i}\varphi(\boldsymbol{r})] \qquad [31]$$

where A is a constant and φ describes the phase structure. The associated intensity, namely,

$$I_0(\boldsymbol{r}) = |V_0(\boldsymbol{r})|^2 = |A|^2 \qquad [32]$$

does not contain any information about φ. One useful way to interpret this lack of information is obtained by observing that, for small values of φ, the field [31] can be written as

$$V_0(\boldsymbol{r}) = A[1 + \mathrm{i}\varphi(\boldsymbol{r})] \qquad [33]$$

so that the intensity reads as

$$I_0(\boldsymbol{r}) = |A|^2[1 + \varphi^2(\boldsymbol{r})] \simeq |A|^2 \qquad [34]$$

It can be said that the two terms appearing in the right-hand side of eqn [33] do not interfere with each other because they are in quadrature (dephased by $\pi/2$). Consider then the Fourier transform of V_0,

$$\tilde{V}_0(\boldsymbol{v}) = A[\delta(\boldsymbol{v}) + \mathrm{i}\tilde{\varphi}(\boldsymbol{v})] \qquad [35]$$

where $\boldsymbol{v}$ is the spatial frequency variable. Suppose now that, through the insertion of a suitable filter in the Fourier plane, the delta function is multiplied by a constant of the form $\mathrm{i}a$, where a is a real number whose modulus is not greater than 1. At the same time, it is required that $\tilde{\varphi}$ remains approximatively unchanged on passing through the filter. These two requirements can be satisfied if a small disk of a suitable material is put at the origin of the Fourier

plane. The transmission function of the corresponding filter is ia in a small region around the origin (the disk region), while it has the value 1 in the remaining part of the plane. As a consequence, the filter acts on the delta function, multiplying it by ia, whereas it leaves essentially unaltered $\tilde{\varphi}$, whose significant values are taken for frequencies different from zero. The output field can now be written as

$$V(\boldsymbol{r}) = \mathrm{i}A[a + \varphi(\boldsymbol{r})] \qquad [36]$$

and the associated intensity becomes

$$I(\boldsymbol{r}) = |A|^2[a^2 + 2a\varphi(\boldsymbol{r}) + \varphi^2(\boldsymbol{r})] \simeq |A|^2 \times [a^2 + 2a\varphi(\boldsymbol{r})] \qquad [37]$$

It is seen that the variations of the intensity are proportional to those of the phase. Therefore, one has a linear mapping of the phase structure. On reducing the value of $|a|$, the contrast of the image (pointwise ratio between the second and the first term in square brackets) is increased while the luminosity diminishes. A limiting case is obtained for $a = 0$. This corresponds to the central stop case. Instead of eqn [37], one obtains

$$I(\boldsymbol{r}) = |A|^2\varphi^2(\boldsymbol{r}) \qquad [38]$$

so that the mapping refers to the square of the phase.

Boundary Detection

In a number of instances, the problem faced is that f recognizing a certain pattern within an image. This task is made easier if the boundaries of the pattern are put into evidence. Such an operation can be performed through some process of mathematical derivation of the image. This, in turn, can be achieved by multiplicative operations on the spectrum, thanks to the derivative theorems for the Fourier transform. As a simple example, refer to a one-dimensional case in which the first derivative is sought. A filter whose transmission function is proportional to the spatial frequency is required

$$\tau(v) = \alpha v \ \mathrm{rect}\left(\frac{v}{v_{\mathrm{M}}}\right) \qquad [39]$$

where α and v_{M} are constants. The presence of the function rect, which equals 1 for $|v| \leqslant v_{\mathrm{M}}/2$ and zero otherwise, is due to the fact that, whatever the value of α, the linear dependence law can hold for a limited interval of values of v only. When such a filter is introduced in the Fourier plane, the output field is the convolution of the derivative of the input signal with a $\sin(\cdot)/(\cdot)$ function produced by Fourier transformation of the rect.

In the two-dimensional case, other types of operations can be performed, for example, the application of the Laplace operator. This modification too can be realized through a suitable filtering operation.

Holographic Filters

The previous examples show that in order to perform filtering operations, one needs to construct transparencies that alter, in a prescribed way, both the amplitude and the phase patterns across the Fourier transform of the input field. The amplitude can be controlled by a careful use of photographic emulsions. As for the phase, the thickness of a dielectric plate can, in principle, be tailored in such a way that the plate introduces, point by point, the required phase changes. This, however, is cumbersome from the experimental point of view. A more viable solution is to use holographic methods. As is well known, holography is a technique by which both the amplitude and the phase of a coherent field can be recorded in a medium, such as a photographic emulsion, which is sensitive to the intensity of the field. Furthermore, powerful methods of digital holography have been developed. In other words, the hologram can be synthesized with the aid of a computer. This makes it possible to design and realize a wide variety of filters.

Advantages and Limits of Optical Processing

Whenever the information to be processed is carried by a coherent (or nearly coherent) field, optical processing is invaluable. A typical example is the visualization of phase objects. As seen, such an operation can be performed with rather simple optical devices that work in real time. On the other hand, standard optical processors are affected by a certain lack of flexibility. Nonetheless, some reconfigurable filters (spatial light modulators) have been realized using, for example, liquid crystal valves. At present, the spatial resolution afforded by these devices is not very high, but the related technology is steadily advancing.

Images formed with incoherent light could also be treated by optical processors by first recording the image on a suitable transparency. In this case, however, the information is contained in the intensity distribution only. As a consequence, the processing can be done with a computer in a more convenient way.

See also: Geometrical Optics; Holography; Optical Instruments.

PACS: 42.30. – d

Further Reading

Born M and Wolf E (1999) *Principle of Optics*, 7th edn. Cambridge: Cambridge University Press.
Collins SA Jr. (1970) *Journal of the Optical Society of America* 60: 1168.
Goodman JW (1996) *Introduction to Fourier Optics*, 2nd edn. New York: McGraw-Hill.
Siegman AE (1986) *Lasers*. University Science Books, Sansalito.
Zernike F (1935) *Z. Tech. Phys*. 16: 454.

Impedance Spectroscopy and Mobility Spectra

R A Gerhardt, Georgia Institute of Technology, Atlanta, GA, USA

Introduction

Impedance spectroscopy is a technique that has seen revolutionary changes in the last 30 years as a result of the development of computer automation, design of better frequency-response analyzers, and the ability to acquire data over many different frequency ranges. Most commercially available equipment today can collect real and imaginary impedance data from frequencies as low as a few microhertz and as high as a few gigahertz. Due to a large variance in material resistivities (e.g., $\rho > 10^{12}$ ohm cm^{-1} for insulators and $< 10^{-3}$ ohm cm^{-1} for semiconductors), different manufacturers specialize in the characterization of specific material types and the equipment specifications can vary widely.

Impedance spectroscopy, also referred to as IS/DS because of its close relationship to dielectric spectroscopy, involves the measurement of the current (I), voltage (V), and phase angle (θ) over a wide frequency range so that the impedance vector (Z^*) may be defined as follows:

$$Z^*(\omega) = V(\omega)/I(\omega) \qquad [1]$$

where $V(\omega) = V_m \sin(\omega t)$, $I(\omega) = I_m \sin(\omega t + \theta)$, $\omega = 2\pi f$, and θ is the phase difference between the current and the voltage.

Measurements can also be carried out in the time domain and then Fourier transformed into the frequency domain. Each measurement mode has advantages and disadvantages. The frequency domain is ideally suited for determining the steady-state response of the object under test, whereas the time domain is better suited for detecting time-dependent phenomena.

Equivalent Circuits

It is customary to analyze the measured impedance response in terms of the possible equivalent circuit components that can give rise to the frequency response observed. The circuit elements of interest are capacitors represented by C, resistors represented by R, and inductors represented by L. Most dielectric materials can be represented by a parallel RC circuit, whereas others are better represented by a series RC circuit. The presence of magnetic components introduces inductive components, and one may expect that series or parallel RL circuits may represent such materials.

It is useful to remember that if elements are in series, it is easier to write the impedance response. On the other hand, if the circuit elements are in parallel, it is more convenient to write the admittance and then convert it to the impedance. For example, the impedance of a series RC circuit is written as

$$Z^* = R + [1/j\omega C] \qquad [2]$$

whereas the impedance of a parallel RC circuit is given by

$$Z^* = 1/Y^* \qquad [3]$$

where $Y^* = (1/R) + j\omega C$.

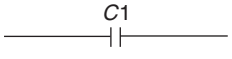
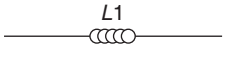
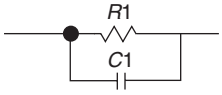
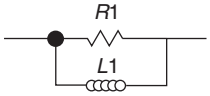
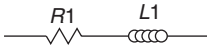
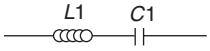
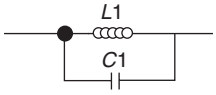
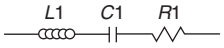
Most modern equipment include some version of software that can be used to fit the most appropriate equivalent circuit to the experimentally measured results. However, the reader is urged to consider the physical mechanisms that may have contributed to the experimental data before concluding that a derived R or C is representative of the bulk of the material or device being measured (see the section "Data analysis" for more details).

Table 1 lists the simplest equivalent circuits and their corresponding impedance equations. Data acquired can be plotted in terms of frequency-explicit and frequency-implicit impedance plots. Frequency-explicit plots are often referred to as Bode plots. They are just simply real (Z') or imaginary impedance (Z'') plotted versus frequency. The magnitude of the impedance vector, which is equal to

$$|Z^*| = [Z'^2 + Z''^2]^{1/2} \qquad [4]$$

and the phase angle θ are also plotted versus the measured frequency. The magnitude of the impedance

Table 1 Simplest equivalent circuits and their impedance equations

Resistor		$Z^* = R1$
Capacitor		$Z^* = 1/j\omega C1$
Inductor		$Z^* = j\omega L1$
RC parallel		$Z^* = 1/[1/R1 + j\omega C1]$
RC series		$Z^* = R1 + (1/j\omega C1)$
RL parallel		$Z^* = 1/[1/R1 + 1/j\omega L1]$
RL series		$Z^* = R1 + j\omega L1$
LC series		$Z^* = j\omega L1 + 1/j\omega C1$
LC parallel		$Z^* = 1/[1/j\omega L1 + j\omega C1]$
LCR series		$Z^* = j\omega L1 + (1/j\omega C1) + R$

vector is useful for quickly assessing the presence of resistive components (flat with frequency), capacitive components (decrease with increasing frequency), and inductive components (increase with increasing frequency) (see **Figure 1** for representative frequency-explicit plots for a perfect resistor, a perfect capacitor, and a perfect inductor).

Frequency-implicit plots are better known as complex plots because the imaginary impedance is plotted versus the real impedance. These plots are often referred to as Nyquist plots or Cole–Cole plots and result in the formation of semicircles only under certain conditions. As shown in **Figure 2**, a semicircle will only appear on the complex impedance plots when the measured object can be represented by a parallel RC circuit or a parallel RL circuit. It should be noted that the imaginary part of these two circuits carries opposite signs. By established convention, a parallel RC circuit will display a semicircle in the first quadrant if the imaginary axis is labeled as negative (**Figure 2a**). In contrast, a parallel RL circuit will display an upside down semicircle in the fourth quadrant (**Figure 2d**). It may be added that a series RC circuit or a series RL circuit do not display semicircles in the complex impedance plane (**Figures 2b** and **2c**).

Microscopic Interpretation

If one is interested in determining the electrical process which has given rise to a particular set of impedance spectra, it is useful to remember that all materials are made up of charged entities at various length scales (see **Figure 3**). At the atomic level, one has protons and electrons. At the molecular level, there may be anions, cations, or charged molecules. At the nano-, meso-, and micro-structural levels within a single material, there may also be charged interfaces. The interface between the contacting electrode and the material surface often results in a space-charge polarization. Hence, the impedance response measured is dictated by the total number of impediments to the current flow that an AC signal encounters on its way from the high- to the low-voltage electrode. The impedance response measured is then a function of the frequency of the applied field, how that field interacts with each polarization process present in the material, and whether it is commensurate with the dimensions of the permanent or temporary dipoles formed and the strength of the signal.

Polarization Mechanisms

Figure 4 schematically displays the real and imaginary polarization response of a material as a function of the applied frequency. Upon the application of an electric field, atomic or electronic polarization is easily detectable in the optical frequency range, and this results in the resonance response displayed. The polarization strength is proportional to the number of electrons present and is governed by the equation below:

$$P = Zqd \qquad [5]$$

where Z is the atomic number, q is the charge constant $= 1.609 \times 10^{-19}$ coulombs m^{-1}, and d is the distance.

At the molecular level, dipole polarization can be due to temporary or permanent dipoles present in a material. Permanent dipoles are present in those materials that possess a center of positive charge different from that of a negative charge. This includes all ferroelectrics below their Curie temperature, all polar polymer molecules, and all polar liquids such as water. It may also occur by aliovalent doping of compounds. Temporary dipoles may occur in weakly bound polymers or in the electrostatic attachment of nanowires to patterned substrates. Molecular temporary dipoles also occur in ionically bonded materials, such as many oxides and fluorides.

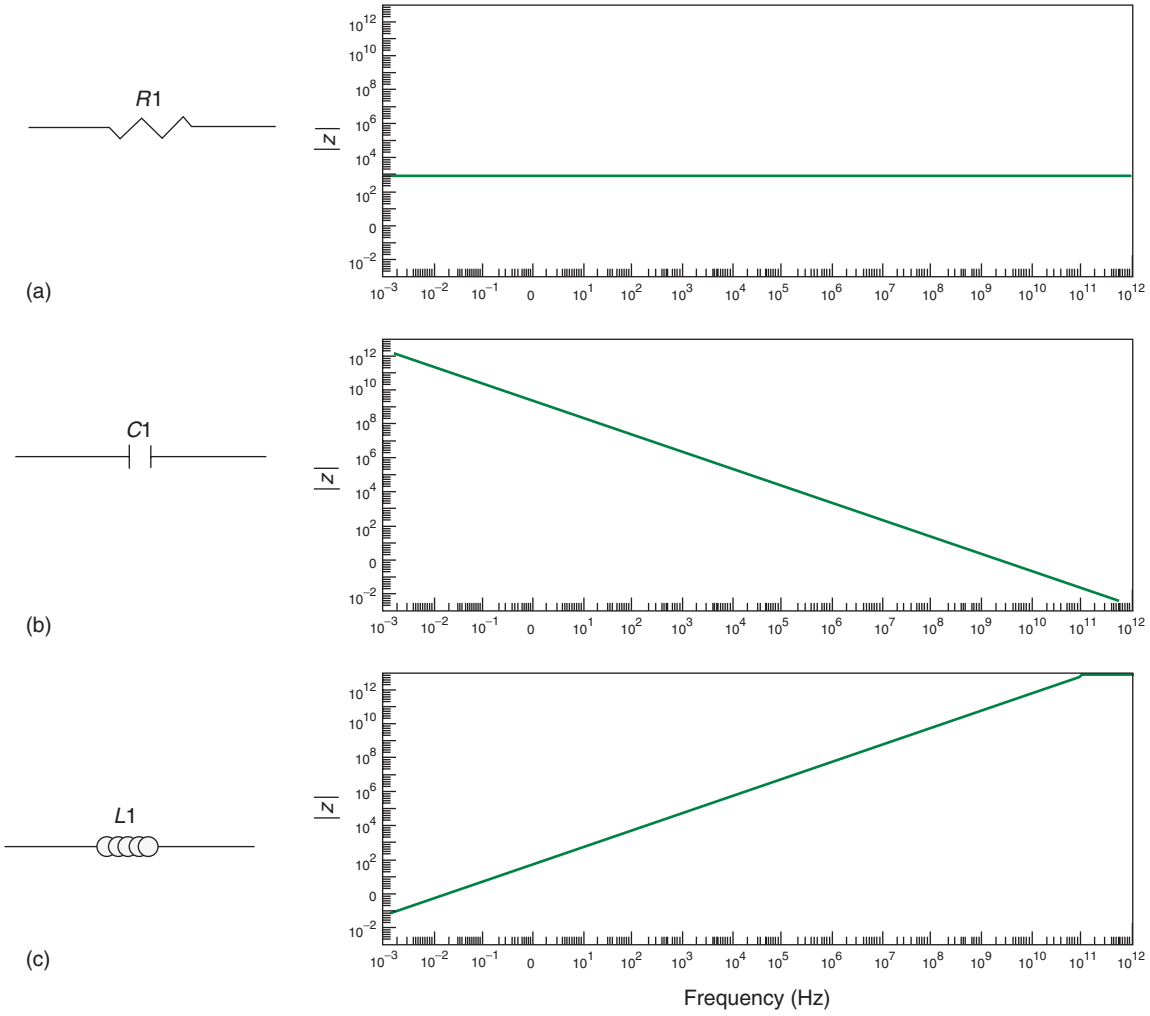


Figure 1 Representative impedance vector response for (a) a perfect resistor, (b) a perfect capacitor, and (c) a perfect inductor as a function of frequency.

Space-charge polarization occurs whenever the electrical response at an interface is dominated by a pile-up of charge carriers at that interface. This commonly occurs at the interface between the contacting electrode and the material under test, and is often assumed to occur only at very low frequencies. If the electrode contact is nonohmic or if the charge-transfer resistance is high, the space-charge polarization will overshadow any other processes that may be present in the material. Space-charge polarization may also occur at the interface between grains of phases with very different electrical properties or at the grain boundaries of a single-phase material, if the defect concentration is high or if there is segregation of impurities into or away from the boundaries. Space-charge polarization at the interface between two different materials is generally expressed in terms of the Maxwell–Wagner two-layer condenser model. These equations are discussed in the section, "Multiple relaxation."

Space-charge polarization at electrode–material interfaces is often diffusive in nature and can be represented by the Warburg impedance. This gives rise to a linear dependence of the conductivity on the frequency of the measurement, as described in the following equation:

$$Z^* = [A/(D\omega)^s](1 - \mathrm{j}) \qquad [6]$$

where A is a constant related to the valence state of the diffusing species and the concentration gradient, and s represents the diffusion exponent. For a perfect diffusion-controlled response, $s = 1/2$. This is often seen as a straight line at 45° from the real impedance axis.

Relaxation and Resonance Frequency Spectra

The frequency dispersion that allows transformation of the dielectric response from one polarization

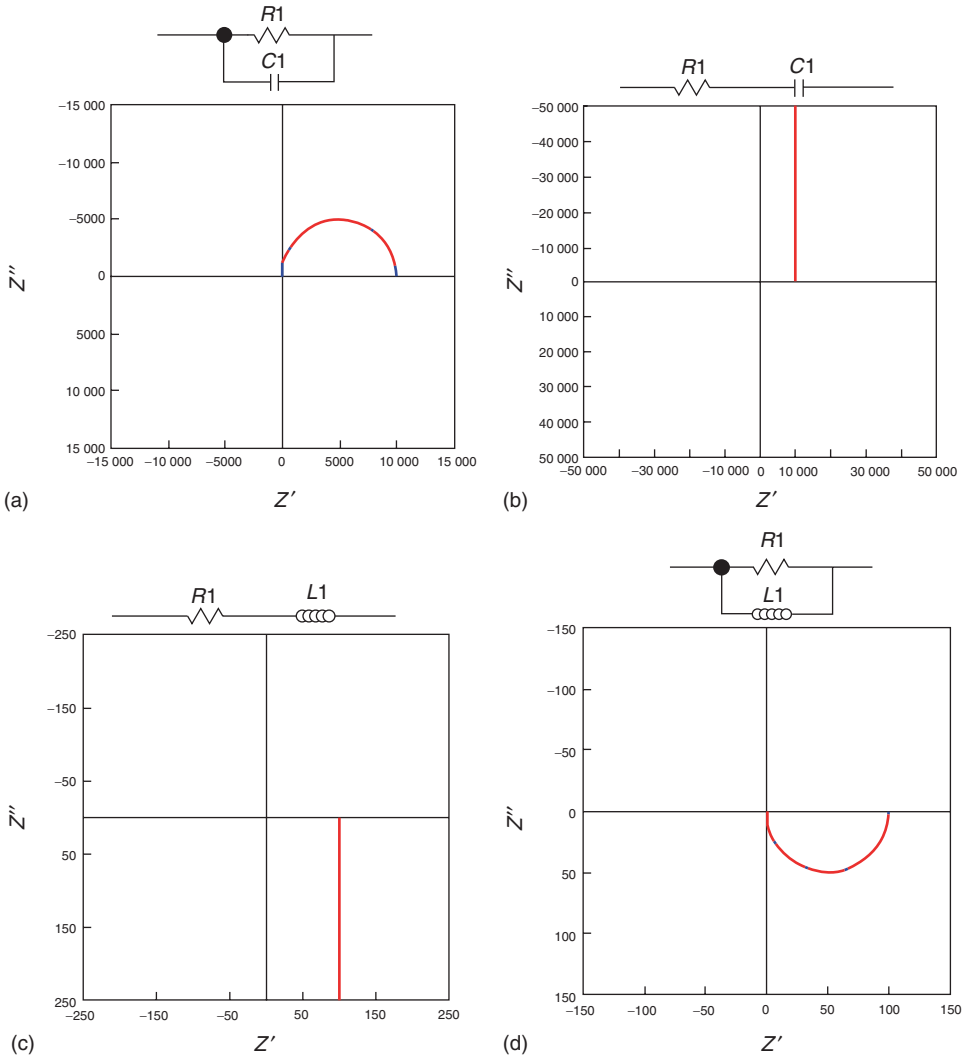


Figure 2 Equivalent circuits and representative complex impedance plots for (a) a parallel RC circuit, (b) a series RC circuit, (c) a series RL circuit, and (d) a parallel RL circuit.

mechanism to another is known as dielectric relaxation. The Debye equation describes this for essentially all relaxation processes even though it was originally developed to represent dielectric relaxation in polar liquids:

$$\varepsilon^* = \varepsilon_\infty + (\varepsilon_s - \varepsilon_\infty)/[1 + j\omega\tau] \qquad [7]$$

where ε_∞ represents the high-frequency dielectric permittivity, ε_s is the lowest frequency dielectric permittivity, ω is the angular frequency, and τ the relaxation time for the process. Cole–Cole and others have modified the Debye equation to account for deviations from the ideal Debye response. Cole and Cole introduced the α parameter as follows:

$$\varepsilon^* = \varepsilon_\infty + (\varepsilon_s - \varepsilon_\infty)/[1 + (j\omega\tau)^{1-\alpha}] \qquad [8]$$

Davidson and Cole introduced the β parameter, which accounts for deviations from ideality in the highest frequency region of the Debye relaxation. This modification is written as follows:

$$\varepsilon^* = \varepsilon_\infty + (\varepsilon_s - \varepsilon_\infty)/[1 + (j\omega\tau)]^\beta \qquad [9]$$

To account for both spreading of the relaxation time and deviations from ideality at high frequency, Havriliak and Negami combined the two parameters, as shown below:

$$\varepsilon^* = \varepsilon_\infty + (\varepsilon_s - \varepsilon_\infty)/[1 + (j\omega\tau)^{1-\alpha}]^\beta \qquad [10]$$

The differences between the Debye, Cole–Cole, and Davidson–Cole relaxation equations are depicted in **Figure 5**.

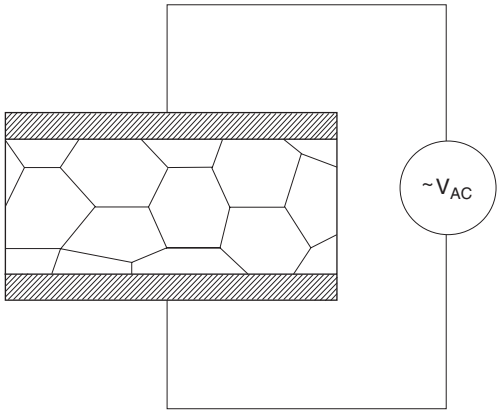


Figure 3 Schematic of experimental setup for a two probe parallel plate AC impedance measurement. The specimen cross section shows the presence of electrode–specimen interfaces and grain-to-grain interfaces which may have different response than the bulk grains of the material, any permanent dipoles present, any ionic polarization present, or any electronic polarization present.

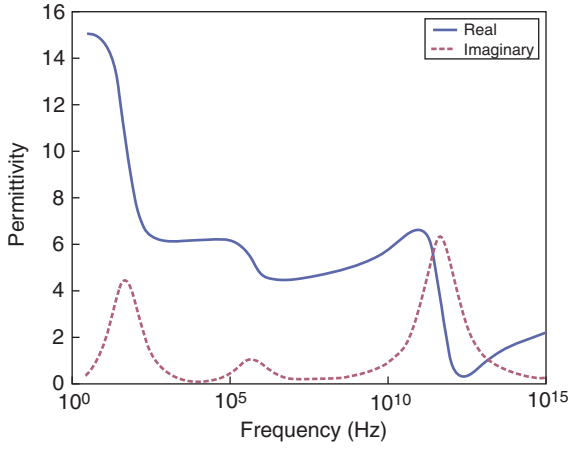


Figure 4 Schematic of the real and imaginary permittivity curves for a material that contains two dielectric relaxation processes and a single resonance process. Note that change in the real part results in the formation of a peak in the imaginary part.

Resonance spectra are hardly ever discussed in the context of IS. This is because ionic and electronic resonance generally occur at frequencies near the optical range and are best detected with optical spectroscopic methods such as IR, UV, and visible spectroscopy methods. NMR and EPR methods are also superb at detecting ionic resonances. Nevertheless, it is possible to observe resonance phenomena while making impedance measurements. Resonance behavior that occurs at the lower frequencies of impedance measurements is normally associated with parasitics in the measurement circuit, but may also be due to a nonlinear response of materials such as ferroelectrics and ferromagnetics. An equation that

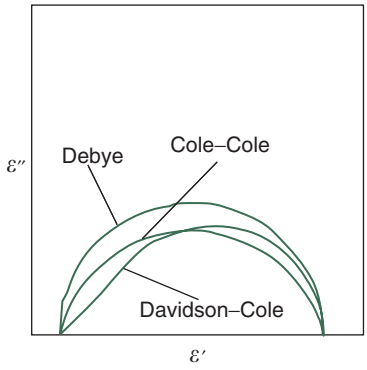


Figure 5 Simulated complex permittivity plots for representative Debye, Cole–Cole, and Davidson–Cole relaxations described by eqns [7]–[9].

can generate a resonance response is given by

$$\varepsilon^* = \varepsilon'_\infty + 1/2(\varepsilon'_s - \varepsilon'_\infty)/[(1 - (\omega/\omega_{res})) + (\beta/\omega_{res})] \quad [11]$$

where ε'_s and ε'_∞ are the permittivities at frequencies much lower and much higher than the resonant frequency ω_{res}, and β represents the strength of the restoring force that gives rise to resonance. **Figure 6** displays the frequency-explicit and frequency-implicit graphs of the dielectric resonance equations.

Various Representations of the Dielectric Function

A fact not often recognized is that the dielectric relaxation equation, originally developed by Debye for relaxation in polar liquids, provides another way of interpreting impedance data. This is possible because dielectric permittivity is inversely related to impedance, as shown below:

$$Z^* = 1/j\omega C_0 \varepsilon^* \quad [12]$$

Therefore, it follows that measured impedance data can be converted into permittivity and vice-versa, provided that the geometric capacitance, C_0, of the sample measured is known. Recalling that admittance, Y^*, is the inverse of impedance and that electric modulus, M^*, is the inverse of permittivity, it then becomes possible to analyze the same measured data in four different formalisms without the use of any adjustable parameters. **Table 2** lists all the important relationships between these four dielectric functions. Performing the conversions mentioned is advantageous because polarization or conductivity processes that may be hidden in the impedance, for example, may become fully detectable in the

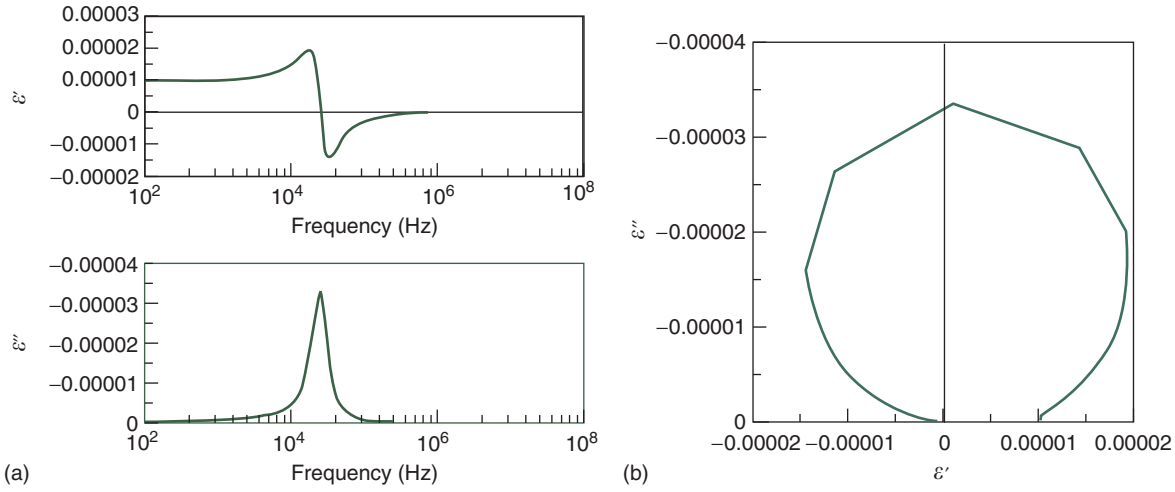


Figure 6 (a) Frequency explicit real and imaginary permittivity plots for a resonance process. (b) Frequency implicit complex permittivity plot for a resonance process.

Table 2 Relationships between dielectric functions

$Z^* = 1/j\omega C_0\varepsilon^*$	$\varepsilon^* = 1/j\omega C_0 Z^*$
$Y^* = j\omega C_0\varepsilon^*$	$Y^* = 1/Z^*$
$M^* = 1/\varepsilon^*$	$M^* = j\omega C_0 Z^*$
$\tan\delta = \varepsilon''/\varepsilon'$	$\tan\delta = Z''/Z'$

admittance plane. **Figure 7** displays the impedance and admittance of a series RC circuit. It can be seen that it does not show a semicircle in the impedance plane but shows one in the admittance plane, although this one appears in the fourth quadrant if one keeps the same sign convention as for the impedance. Likewise, a series RL circuit also displays a semicircle in the admittance plane but not in the impedance plane (see **Figure 2c**) but this semicircle will appear in the first quadrant. It should then become clear that using a multiplane analysis can be extremely useful when more than one relaxation process may be active within a measured frequency range. It should be added that mathematical transformations need to be done with caution, as the data on either extreme of the frequency range of a given impedance analyzer tend to carry a higher error. Hence, taking the inverse can amplify the size of the errors. Nevertheless, it is an extremely useful tool for determining the most appropriate model that can help identify the relaxation process more easily. Moreover, once the relaxation process has been identified, the data can then be used to obtain the activation energy for that particular process.

Temperature Dependence

Most relaxation phenomena are thermally activated. As a result, the relaxation time can be expressed in terms of an Arrhenius-type temperature dependence:

$$\tau = \tau_0 \exp(E/kT) \qquad [13]$$

where τ_0 is the relaxation frequency at infinite temperature, E is the activation energy for the relaxation process, k is Boltzmann's constant, and T is temperature in degrees kelvin. In addition, when a relaxation process occurs, $\omega_{max}\tau = 1$; therefore, one can use the frequency at which a peak occurs at a given temperature as a way of determining the activation energy of the process. **Figure 8a** schematically displays a dielectric loss process that changes peak frequency as temperature is changed. The linearized version of eqn [13] is used to obtain the activation energy as depicted in **Figure 8b**.

Similarly, conductivity processes may display peaks in the imaginary impedance or imaginary admittance, which could then be used to estimate the activation energy for that process in the same fashion as described above. More often, the activation energy for conduction, in a given system, is obtained by taking the intercept of the impedance semicircle with the real axis, as displayed in **Figure 9**, and converting that to conductivity and plotting the calculated conductivity versus inverse temperature. Either type of analysis is only accurate when the resistance and capacitance of the relaxation processes being measured are both independent of frequency. A truly bulk response will always have constant R and C at a given temperature. This is not the case for space-charge polarization under biasing conditions or diffusion-controlled processes. In both of these cases, the number of mobile charge carriers will change the values of R and may also affect the value of C.

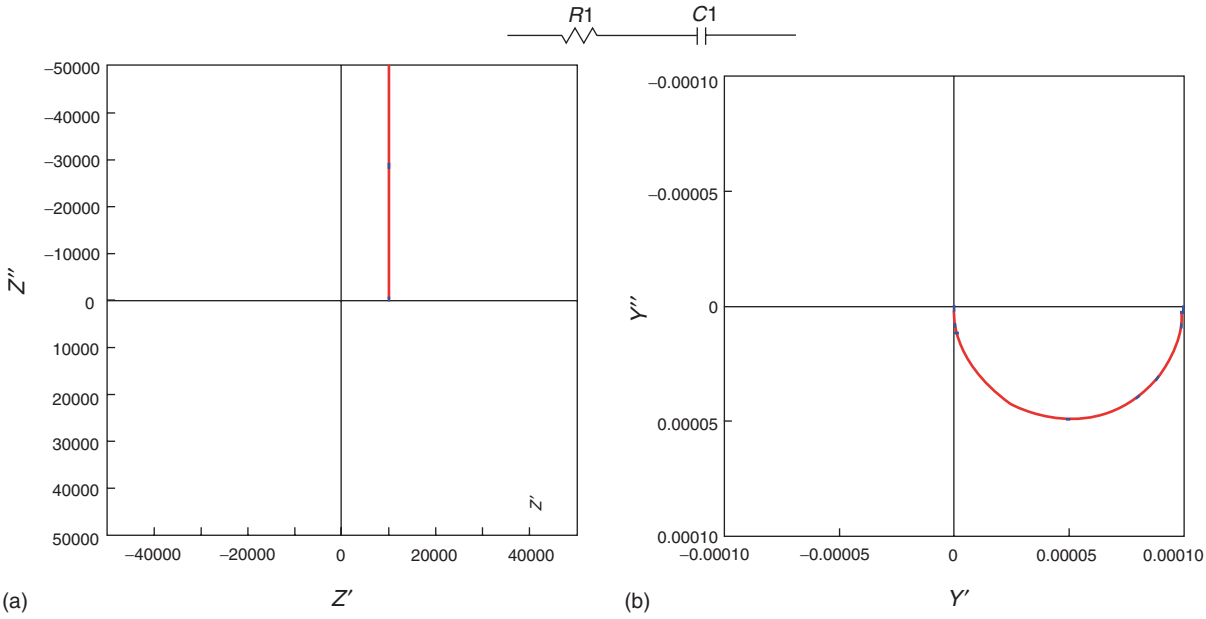


Figure 7 Complex impedance (Z'' vs. Z') and complex admittance (Y'' vs. Y') plots for a series RC circuit.

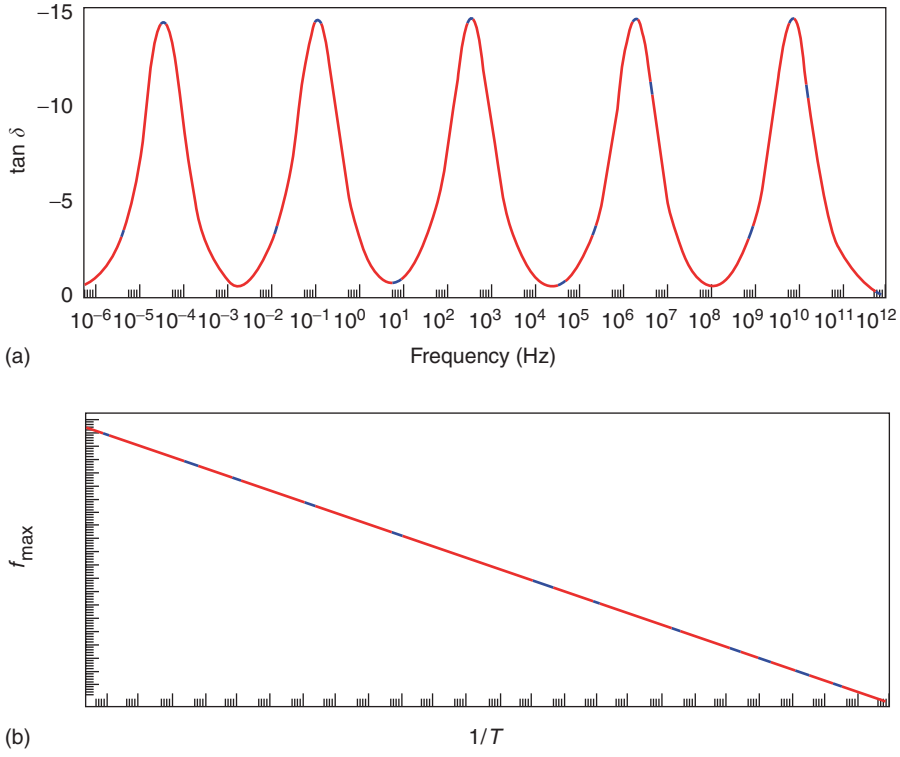


Figure 8 (a) Schematic of a dielectric loss ($\tan\delta$) process at different temperatures. Loss peaks shift to higher frequency (f_{max}) as the temperature is increased. (b) Plot of f_{max} versus the reciprocal temperature allows determination of the activation energy for the dielectric loss process.

Multiple Relaxations

If one refers back to **Figure 4**, it should not be surprising to realize that most experimentally measured spectra will be made up of more than one relaxation process. Examples include: (1) a Debye-type dielectric polarization process and a parallel leakage current, (2) a solid-state ionic material with different response at the grain boundaries than

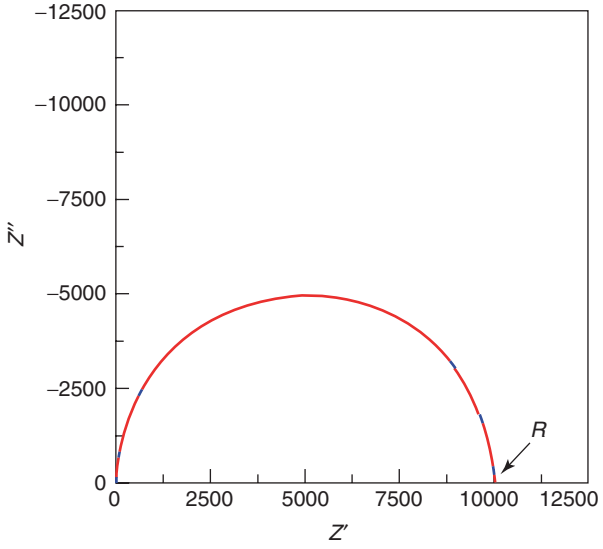


Figure 9 The intercept of the complex impedance semicircle with the real Z' axis is used to determine the resistivity of a conduction process. This must be done for different temperatures in order to obtain the activation energy for the conductivity process.

within the bulk grains, (3) a semiconducting material with space-charge polarization at the electrode/material interface, and (4) bulk response and interfacial responses of multiphase systems.

For the defect dipole case, the easiest way to represent this is to modify the Debye equation by adding a leakage current, σ, as follows:

$$\varepsilon^* = \varepsilon_\infty + \{(\varepsilon_s - \varepsilon_\infty)/[1 + j\omega\tau]\} - j\sigma/\omega \quad [14]$$

The corresponding equivalent circuit will contain a series RC circuit in parallel with a leakage resistor, which is also in parallel with the bulk capacitance. The different dielectric functions for this situation are displayed in **Figure 10**. **Figure 10a** shows that the leakage path can totally dominate the impedance spectra, while dielelctric permittivity plane (shown in **Figures 10b** and **10d**) hints at the presence of the dielectric relaxation, which is even better displayed in the modulus plane as depicted in **Figure 10c**.

In all of the remaining cases, the simplest equivalent circuit to represent them contains two parallel RC circuits in series, as shown in **Figure 11a**. The physical meaning of each of these circuit elements is, however, quite different for each of these cases. **Table 3** lists some of the possible physical processes that may be represented in each of these cases.

Irrespective of the source of the frequency dependence, when the difference between the relaxation times of the two processes present is large, that is, more than one or two orders of magnitude, or either the R or C is very small, one of the processes may be overshadowed by the other in the impedance plane. This is when a multiplane analysis should be used, as different processes can be best displayed in different dielectric functions, according to their R,C,L values. To illustrate this point, **Figures 11–13** are presented.

Figures 11b–11f display the complex impedance spectra for a bulk ionic conductor with $C1 = 1$ pF, which may have different electrical responses at the grain boundaries (with $C2 = 1$ nF) depending on the ratio of the bulk resistance $R1 = R_g$ and the effective resistance at the grain boundaries labeled $R2 = R_{gb}$. The following cases are presented: (1) $R_g = R_{gb}$, (2) $R_g < R_{gb}$, (3) $R_g \ll R_{gb}$, (4) $R_g > R_{gb}$, and (5) $R_g \gg R_{gb}$. These figures illustrate that only looking at the impedance plane may prevent identification of the presence of a second process. Taking the impedance graphs from **Figures 11d** and **11f** which correspond to cases 3 and 5 and converting them to admittance, dielectric permittivity, and electric modulus clearly reveals the existence of more than one process (see **Figures 12a–12d**). Another way, which may also reveal the presence of more than one process when the time constants are quite different, is to use logarithmic scales, as suggested by Jonscher.

When making measurements of a semiconducting material, it is important to pay attention to whether the electrode contact is ohmic or not. If the electrode contact is nonohmic or the charge transfer between the semiconductor and the electrode is inefficient, then an additional semicircle may be observed, which may be interpreted to be related to the space-charge polarization at that interface. **Figures 13a** and **13b** show the complex impedance graphs for a semiconductor in contact with a blocking electrode. The image on the left shows the response for a given size electrode whereas the image on the right displays the effect of changing the electrode size, which affects the semicircle due to the electrode but not the response of the bulk material.

The situation is again different for the impedance response of a multiphase material. The two-layer condenser, first proposed by Maxwell and Wagner to represent a layered material, contains layers of a material with conductivity σ_1 and dielectric permittivity ε_1 that are placed with alternating layers of a material with conductivity σ_2 and dielectric permittivity ε_2, and thicknesses of d_1 and d_2. This gives rise to the Maxwell–Wagner dielectric relaxation equations given below:

$$\varepsilon'_{M-W} = (\varepsilon_\infty)^{M-W}\{1 + [k/(1 + \omega^2\tau^2)]\} \quad [15]$$

and

$$\varepsilon''_{\mathrm{M-W}} = (\varepsilon_\infty)^{\mathrm{M-W}}\{[\tau/\omega\tau_1\tau_2] + [k\,\omega\tau/(1+\omega^2\tau^2)]\} \quad [16]$$

where

$$k = [(\varepsilon_s)^{\mathrm{M-W}} - (\varepsilon_\infty)^{\mathrm{M-W}}]/(\varepsilon_\infty)^{\mathrm{M-W}}$$

$$\tau = [\varepsilon_1 d_2 + \varepsilon_2 d_1]/[\sigma_1 d_2 + \sigma_2 d_1]$$

It goes without saying that quite different spectra will be obtained depending on what the values of σ_1, ε_1, σ_2, ε_2 are. The basic interpretation is that the response at the highest frequencies represents the "effective" response of the bulk of the material, whereas the response at the lowest frequencies takes into account the interfaces between the two materials. It should be added that it may be necessary to consider additional processes such as defect relaxation within the more insulating type material and/or additional

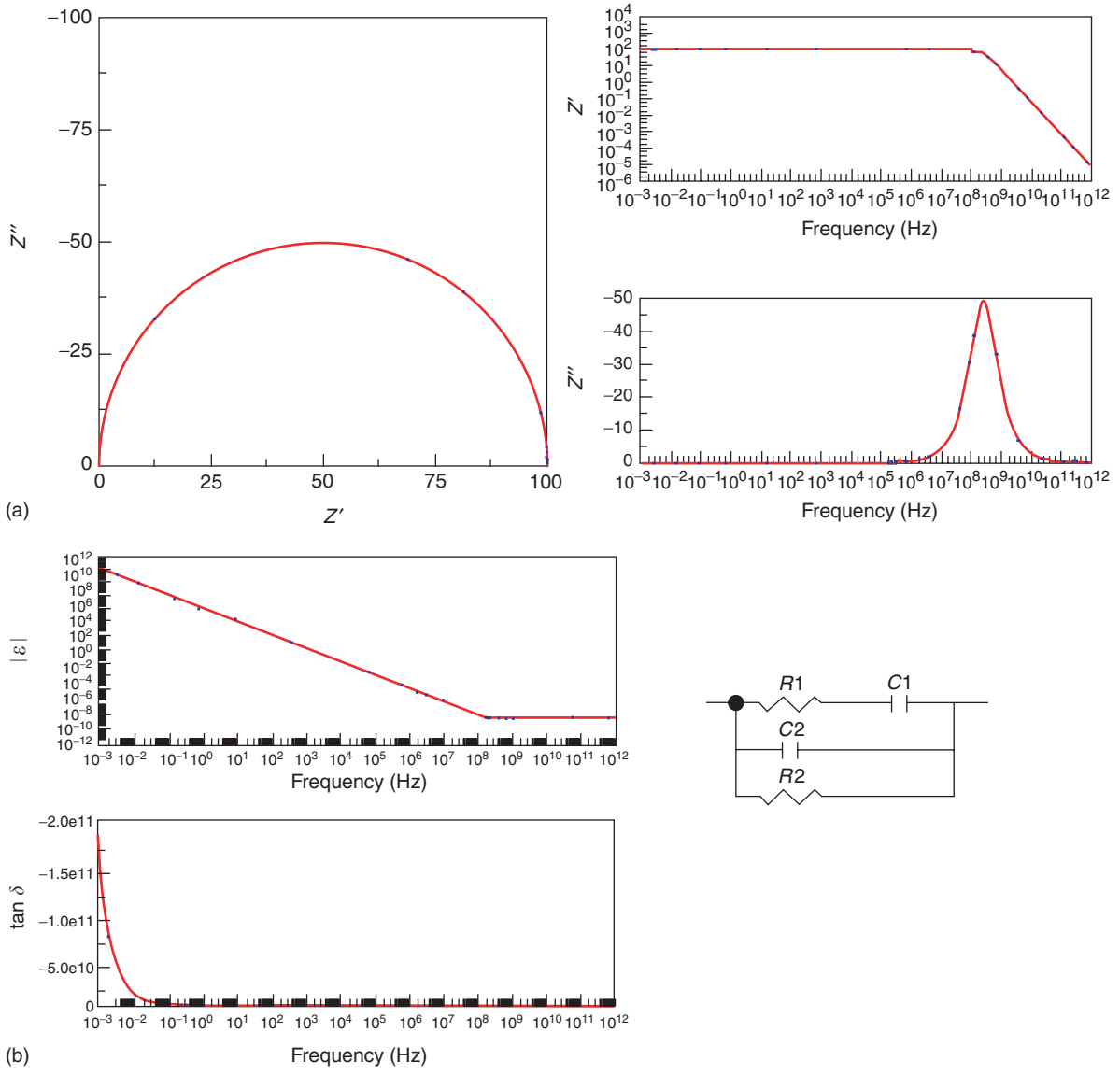


Figure 10 (a) Impedance plane response for a Debye relaxation with a measurable parallel leakage current process. The frequency implicit complex plot is dominated by the current leakage process; however, the frequency explicit real part shows both processes. (b) Frequency explicit dielectric permittivity and dissipation factor plots for a dielectric relaxation with a parallel leakage current. The equivalent circuit for such a combination of processes is also shown where *R*1 and *C*1 represent the leakage process and *R*2 and *C*2 represent the Debye relaxation. (c) Electric modulus plane plots for a Debye relxation with a parallel leakage current process. (d) Simulated frequency implicit (ε'' vs. ε') and frequency explicit dielectric permittivity plots [(ε' vs. log *f*) and (ε'' vs. log *f*)] for a material containing a Debye relaxation process with a measurable parallel leakage current. Note how the imaginary permittivity plot is dominated by the current leakage while the real permittivity plot only displays the Debye relaxation. (e) Admittance plane response for a material containing a Debye relaxation and a parallel leakage current process.

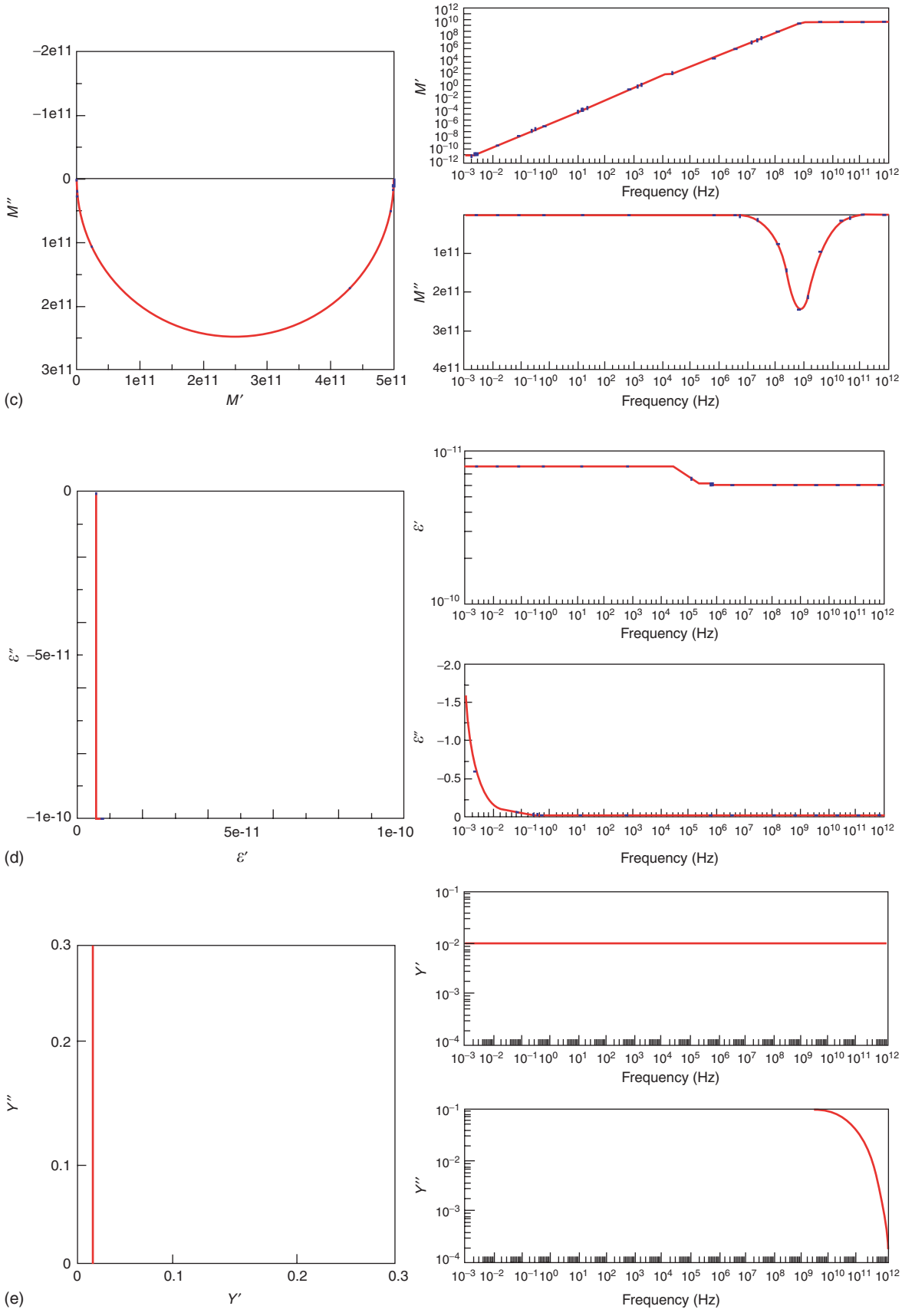


Figure 10 (Continued).

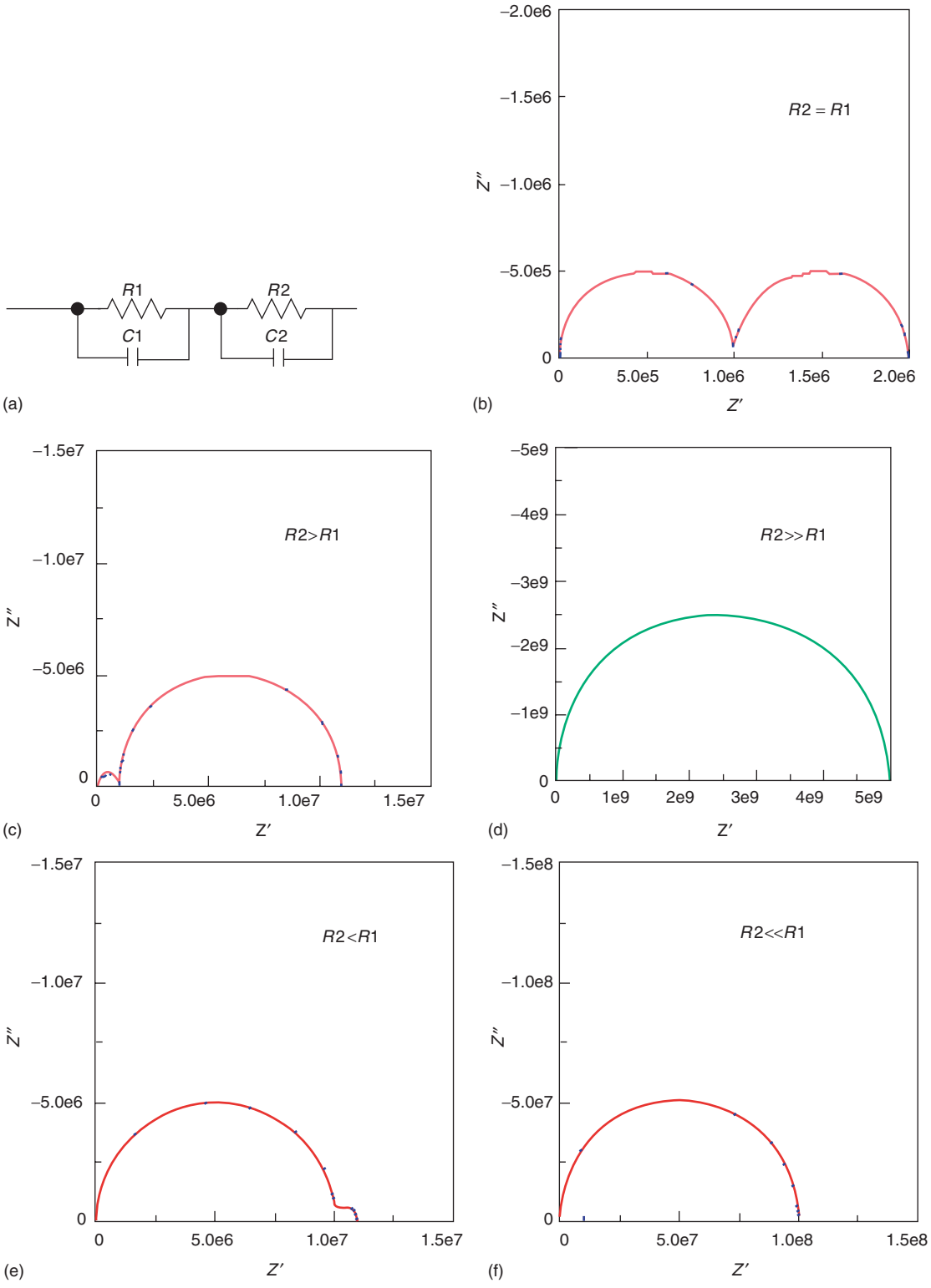


Figure 11 (a) Simplest equivalent circuit needed to represent the various two physical process systems described in **Table 3**. (b) Complex impedance plot for a two-process system where $R1 = R2$ and $C1 = 1$ pF and $C2 = 1$ nF. (c) Complex impedance plot for a two-process system where $R2 > R1$ and $C1 = 1$ pF and $C2 = 1$ nF. (d) Complex impedance plot for a two-process system where $R2 \gg R1$ and $C1 = 1$ pF and $C2 = 1$ nF. Note that $R1$ is not easily detected in this case and that the impedance plot is dominated by $R2$. (e) Complex impedance plot for a two-process system where $R2 < R1$ and $C1 = 1$ pF and $C2 = 1$ nF. (f) Complex impedance plot for a two-process system where $R2 \gg R1$ and $C1 = 1$ pF and $C2 = 1$ nF. Note that $R2$ is not easily detected in this case and that the impedance plot is dominated by $R1$.

Table 3 Possible physical processes for two parallel RC circuits in series

	*R*1	*C*1	*R*2	*C*2
Ionic conductor	Bulk	Bulk	Grain to grain interface	Grain to grain interface
Semiconductor	Bulk	Bulk	Electrode interface	Electrode interface
Multiphase 1	Bulk	Bulk	Interphase boundary	Interphase boundary
Multiphase 2	Main phase	Main phase	Second phase	Second phase

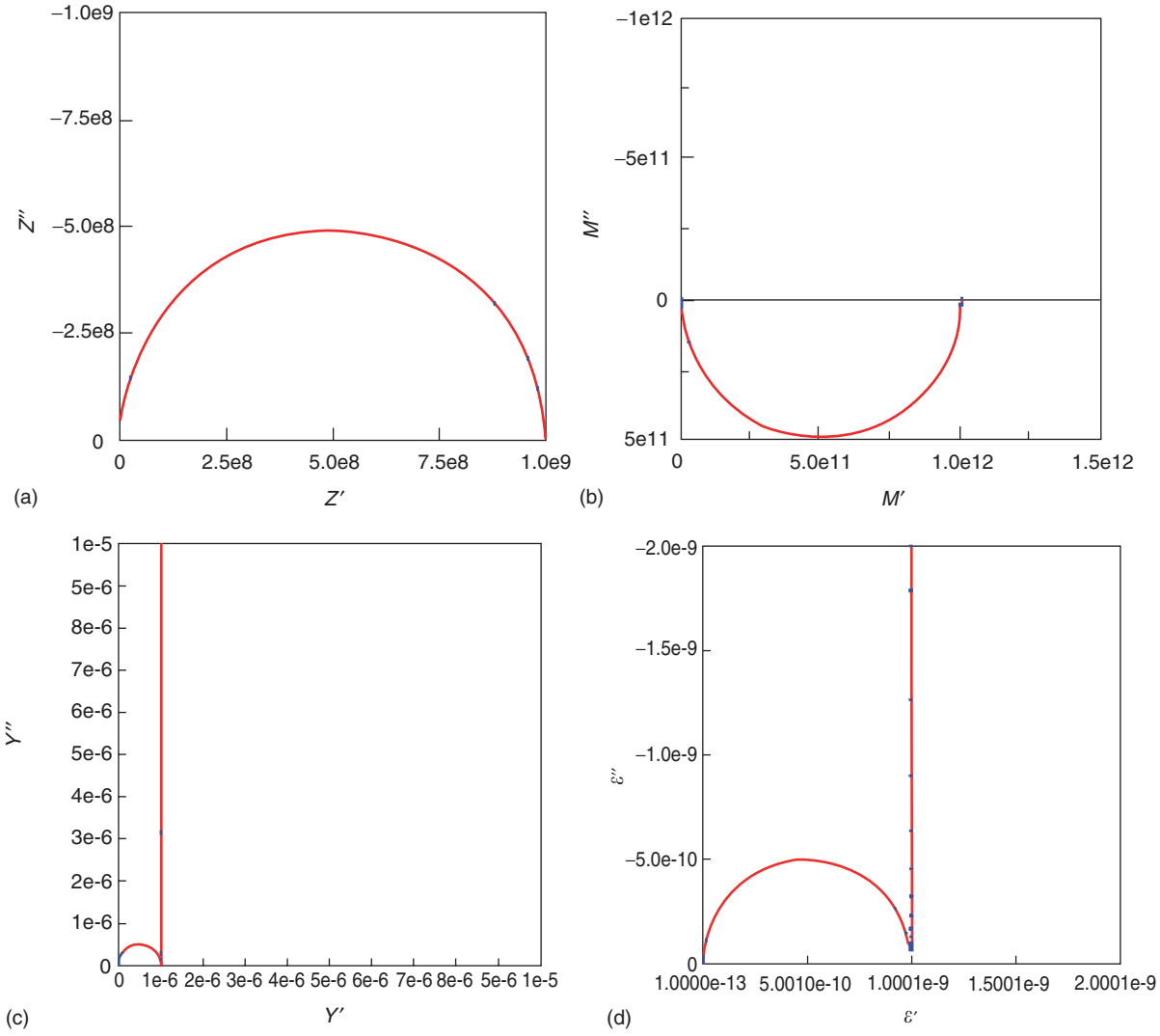


Figure 12 (a) Complex impedance and (b) complex electric modulus for a two-process system similar to that displayed in **Figures 11d** or **11f** where $R1 \ll R2$ or $R2 \ll R1$. (c) Complex admittance (Y'' vs. Y') and complex permittivity (ε'' vs. ε') plots for the same case displayed in **Figures 12a** and **12b**. Note that conversion to admittance and permittivity allows detection of both processes present in this case.

space-charge polarization processes at the interfaces of the more conducting material.

Data Analysis

Some guidelines have already been provided, but additional details will be given here. The most important concept is that of the geometric capacitance. The correct way of obtaining the capacitance of a conducting process is to use the intercept of the impedance curve with the real axis to obtain R for that process. To obtain the capacitance, the peak frequency must be identified (as was done in **Figure 8a**) so that the following equation can be used:

$$C = 1/\omega R \qquad [17]$$

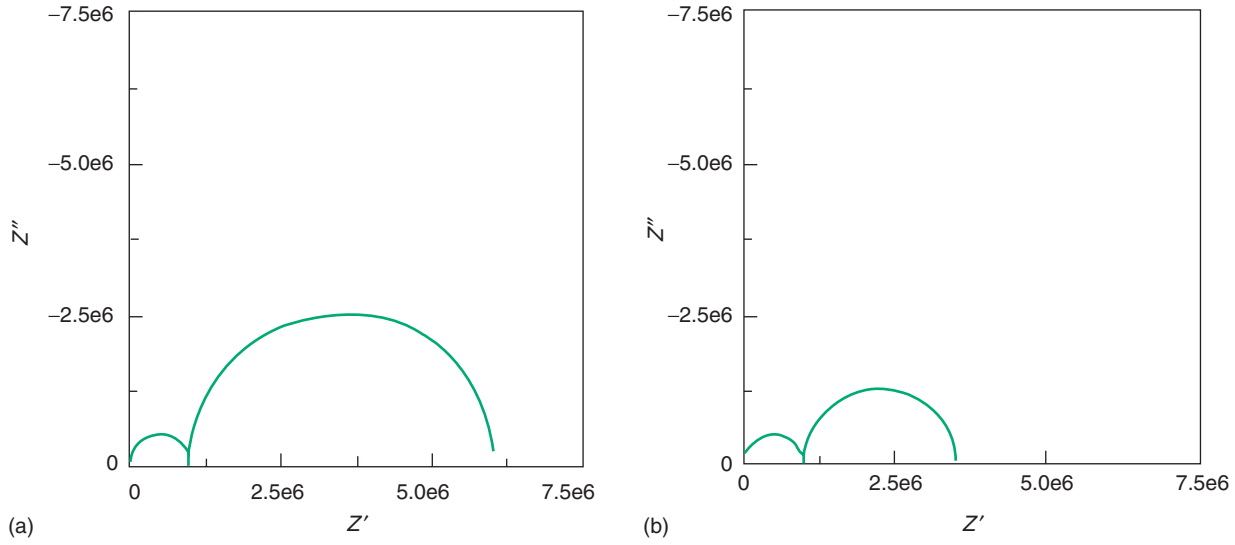


Figure 13 Complex impedance plots for an arbitrary semiconducting material system with a nonohmic or inefficient charge transfer electrodes. The only difference between (a) and (b) is the size of the electrodes used to measure the impedance. Note that while the high-frequency semicircle (due to a bulk process) remains unchanged, the low-frequency semicircle changes in proportion to the interfacial geometry.

Table 4 Effect of physical dimensions on measured capacitance

Dielectric constant	*Physical dimensions t*(m) *with* $A = 5 \times 10^{-5}\,m^2$	*Measured capacitance* (F)
2	1×10^{-3}	8.85×10^{-13}
2	1×10^{-6}	8.85×10^{-10}
2	1×10^{-9}	8.85×10^{-7}
10	1×10^{-3}	4.43×10^{-12}
10	1×10^{-6}	4.43×10^{-9}
10	1×10^{-9}	4.43×10^{-6}
100	1×10^{-3}	4.43×10^{-11}
100	1×10^{-6}	4.43×10^{-8}
100	1×10^{-9}	4.43×10^{-5}
1000	1×10^{-3}	4.43×10^{-10}
1000	1×10^{-6}	4.43×10^{-7}
1000	1×10^{-9}	4.43×10^{-4}

If one is interested in characterizing the bulk electrical properties of a given material or device, the measured capacitance should be of the same order of magnitude as that expected for the physical dimensions of the sample or device measured. This means that if such an analysis gives a capacitance that is not in agreement with the expected capacitance, then the measured data is not related to the bulk of the material or the device and is most probably related to some type of space-charge polarization mechanism. **Table 4** lists the expected capacitances for materials with different dielectric constants and physical dimensions. Note that changing the physical dimensions can change the value of the geometric capacitance drastically. Additional confirmation, that a measured process is related to a bulk process, can be obtained by varying the applied voltage or changing the size of the electrode contacts. The response for a bulk process should be the same, irrespective of the values of these. Instead, if they scale with these variables, it may be possible to separate electrode contact effects from internal interface effects. In cases where the impedance is so large that the intercept with the real axis is not available, the capacitance of the material will only be able to be obtained from either the real dielectric permittivity or the real electric modulus spectra. In most cases, the value is almost constant as a function of frequency. The only time that a semicircle will appear in both of these functions is when a dipolar-type relaxation is present.

Separating individual processes present in the bulk becomes a little bit more difficult. For example, separating the dipolar relaxation from the ionic conductivity in a good ionic conductor is not possible at temperatures where the ionic conductivity is high. However, carrying out measurements at lower temperatures suppresses the ionic conductivity and allows the detection of the defect dipole relaxation. Likewise, if a material has a large fraction of electronic conductivity and only a small fraction of ionic conductivity, detecting the latter is almost impossible within a single set of experiments. In such cases, the usage of blocking electrodes and/or controlled atmospheres to enhance or diminish a particular type of charge carrier will aid in the interpretation.

Electromagnetic Simulation Software

In recent years, a large number of commercially available software packages have become available. Many of the software packages come complete with electrostatic and electromagnetic modules. They are used mostly to model the performance of devices, but physical arrangements of structures (at various length scales) are also possible. Most commercially available simulation software is finite element based (FEA) or finite time difference based (FTDT). Some progress has been made on implementing these packages to help simulate impedance data at the microstructural level, but much more work needs to be done before they become commonplace.

Recent Advances (What to Expect in the Future)

IS of solid-state materials continues to be used in a wide variety of fields and the number of applications continues to grow. Besides being useful as a way of determining the bulk electrical properties of materials and devices, it can also be used as a way to determine: (1) the sensitivity and selectivity of gas sensors, (2) the degree of biocompatibility of new biomedical implants, (3) methods of quality control, and (4) nondestructive characterization of microstructures.

More recent developments include magneto-impedance measurements, photostimulated-impedance measurements, and bioimpedance measurements of tissue. These new methods will be a great help to the burgeoning fields of spintronics, giant magnetoresistance, electrooptical devices, and biomedical engineering. In addition, impedance measurements can be taken at all length scales; localized measurements have been acquired using macro-, micro-, and nano-sized electrodes for a number of years. However, simultaneous impedance and topographical data acquisition in an AFM/STM has the potential to revolutionize the usage of IS in the future. Single frequency impedance amplitude and phase image scans have been collected using the tapping mode or the conductive AFM mode, and are being related to topographical images in order to understand what role different features play. Recent experiments have considered Schottky barriers in single-crystal semiconductors, ionic conductivity in mixed conductors, and electron transfer in molecular wires and carbon nanotubes. It is anticipated that, with further developments in probe tips, many more exciting scanning impedance microscopy experiments will be conducted.

Concluding Remarks

Attempts to simplify the many facets of IS have been made throughout this article. The author has purposely excluded discussion of more complex equivalent circuits, as these only modify the basic concepts described here. Electrochemical impedance spectroscopy (EIS), which is mainly used to study corrosion of metallic materials, was also excluded, as there are many excellent reviews available in the literature. Readers are urged to carefully review the extensive literature in the subject matter of their interest, keeping in mind some of the guidelines presented here.

See also: Dielectric Function; Conductivity, Electrical; Ferroelectricity; Insulators, Impurity and Defect States in; Ionic and Mixed Conductivity in Condensed Phases; Mass Transport; Polarizabilities; Polymers and Organic Materials, Electronic States of; Scanning Probe Microscopy; Semiconductors, General Properties.

PACS: 81.70. – q; 84.37. + q; 72; 73; 77.22. – d; 82.20. – w

Further Reading

Gerhardt RA, Taylor SR, and Garboczi EJ (eds.) (1996) *Electrically Based Microstructural Characterization, MRS Symposium Proceedings*, vol. 411. Warrendale: Materials Research Society.

Gerhardt RA, Alim MA, and Taylor SR (eds.) (1998) *Electrically Based Microstructural Characterization II, MRS Symposium Proceedings*, vol. 500. Warrendale: Materials Research Society.

Gerhardt RA, Washabaugh AP, Alim MA, and Choi GM (eds.) (2002) *Electrically Based Microstructural Characterization III, MRS Symposium Proceedings*, vol. 699. Warrendale: Materials Research Society.

Grigas J (1996) *Microwave Dielectric Spectroscopy of Ferroelectrics and Related Materials*. Amsterdam: Gordon and Breach.

Havriliak SJ and Havriliak S Jr. (1996) *Dielectric and Mechanical Relaxation in Materials: Analysis, Interpretation, and Application to Polymers*. New York: Hanser-Gardner Publications.

Jonscher AK (1983) *Dielectric Relaxation in Solids*. London: Chelsea Dielectrics Press.

Kelly RG, Scully JR, Shoesmith DW, and Buchheit RG (2003) *Electrochemical Techniques in Corrosion Engineering*. New York: Dekker.

Macdonald JR (ed.) (1987) *Impedance Spectroscopy: Emphasizing Solid Materials and Systems*. New York: Wiley.

Nowick AS and Berry BS (1972) *Anelastic Relaxation in Crystalline Solids*. New York: Academic Press.

Von Hippel AR (1954) *Dielectrics and Waves*. New York: Wiley.

Incommensurate Phases

J M Perez-Mato, Universidad del Pais Vasco, Bilbao, Spain

Introduction

Incommensurate phases are thermodynamically stable phases that do not have lattice periodicity although they possess a long-range order. Their diffraction patterns exhibit sharp Bragg peaks but, in contrast with conventional (commensurate) crystalline solid phases, their indexation requires more than three Miller indices. They can be classified into two groups according to their structural properties: modulated structures and composite structures. The first ones can be described as a distortion of an underlying average crystal structure with conventional lattice periodicity.

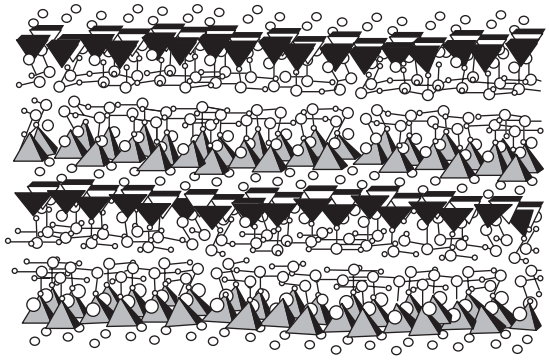


Figure 1 Incommensurate structure of $Bi_2Sr_2[Sr,Ca]Cu_2O_{8+x}$. (Reproduced from Calestani G, Rizzoli C, Francesconi MG, and Andreetti GD (1989) *Physica C* 161: 598–606, with permission from Elsevier.)

This distortion or modulation may involve any atomic parameter (atomic positions, atomic occupation probabilities, etc.), its periodicity being incommensurate with the lattice of the basic structure. This formally means that, at least in one direction, the ratio of the period of the modulation with respect to that of the basic structure is an irrational number. As an example, **Figure 1** depicts an approximate representation of the incommensurate modulated structure of the high-T_c superconductor $Bi_2Sr_2[Sr, Ca]Cu_2O_{8+x}$. In this material, some atomic displacements from the reference average structure are extremely large (more than 1 Å in the case of the oxygen atoms within the Bi_2O_2 layers). In general, atomic displacements due to the modulation are usually ~ 0.1 Å.

Incommensurate composite structures, on the other hand, are built by several crystalline subsystems whose basic lattice periodicities are incommensurate with one another. In the simplest and most common case, the number of subsystems is two. In general, their mutual interaction causes, in each of the subsystems, a modulation with the periodicity of the other one. In this sense, a composite structure is a set of modulated incommensurate structures forming a single phase. As an example, the basic features of a urea inclusion compound are shown in **Figure 2**. The urea forms a lattice with large hexagonal channels where a second subsystem of organic chains accommodates. The chains are ordered along the channels according to a basic periodicity, which is independent of the one of urea. The set of chains can be considered as a guest within the host urea. But both host and guest form a coherent single system, and a non-negligible incommensurate modulation in each subsystem, with the average periodicity of the other one, exists.

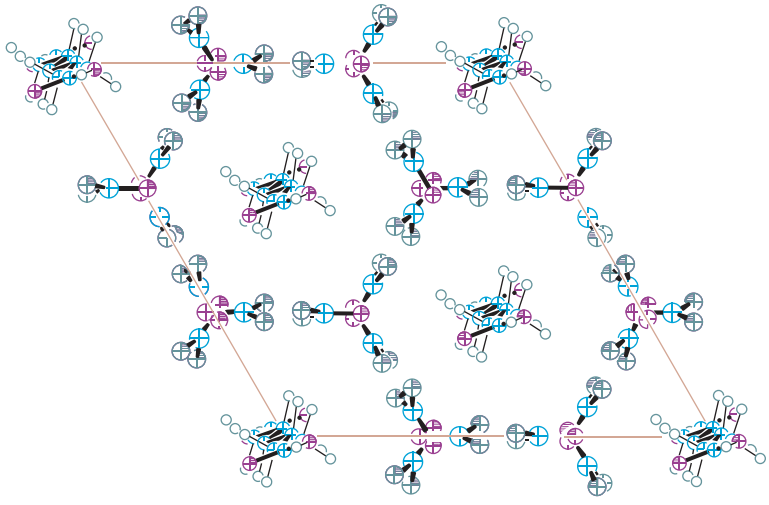


Figure 2 View of a urea inclusion compound with octanedioic acid as guest subsystem. (Courtesy of G Madariaga and I Peral.)

Strictly speaking, the definition given above for incommensurate phases would also include quasicrystals. These latter are usually distinguished by the fact that they exhibit, in their diffraction diagram, a noncrystallographic point group. However, it is now generally accepted that noncrystallographic symmetry is not essential for the concept of a quasicrystalline state, and for instance, aperiodic quasicrystalline tilings with crystallographic point groups can be designed. In fact, the conceptual borders between these three types of incommensurate phases (modulated, composites, and quasicrystals) have become progressively blurred. Nevertheless, if quasicrystals are also included within the same category, one rather speaks of aperiodic crystals instead of incommensurate phases or incommensurate crystals.

In this article, the fundamental crystallographic and physical properties of incommensurate phases in the restricted sense explained above, that is, disregarding quasicrystals, are briefly summarized.

Crystallography of Incommensurate Phases

An example of the diffraction diagram of a simple incommensurate modulated phase is shown in **Figure 3**. One can distinguish a lattice of intense Bragg reflections, the so-called main reflections, which can be indexed in a conventional way with a set of three reciprocal unit cell vectors. There are, however, additional weaker reflections, the so-called satellite reflections, separated from a neighboring main reflection by vectors of type mq, with m an integer

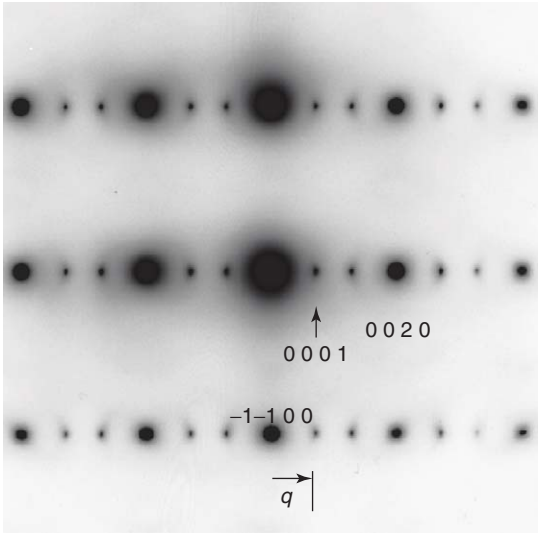


Figure 3 Electron diffraction diagram ($\langle -1, 1, 0 \rangle$ zone axis) of the incommensurate modulated $Ni_{2.667}In_{0.667}Te_2$. (Courtesy of RL Withers.)

and q a unique specific vector. These satellite reflections are the signature of a distortion with periodicity $2\pi/|q|$. Hence, the indexation of the diffraction vectors H associated to the observed Bragg reflections requires four vectors:

$$H = ha^* + kb^* + lc^* + mq;\ h,k,l,m \text{ integers} \quad [1]$$

For a given choice of the four indexation vectors, if q is rationally independent of the basis $\{a^*, b^*, c^*\}$, the indexation [1] is unique and unambiguous, and the system is incommensurate. Obviously, one could always find, within experimental resolution, a much smaller reciprocal unit cell of three vectors capable of indexing all reflections. This, however, would be impractical as the resulting cell in direct space would be too large, and most of the Bragg reflections associated with this lattice would be absent. Hence, the incommensurability of q is more a matter of efficiency and simplicity of the description, rather than a rigorous mathematical property. However, in many cases, the formal incommensurability of q is clearly evidenced by its temperature variation.

Note that, strictly speaking, the set of points described by eqn [1] is not discrete. The reason why one can talk of an unambiguous indexation and a discrete set of Bragg peaks lies on the empirical fact that the set of observed diffraction vectors H is limited to small indices (h,k,l,m).

As an example of an incommensurate composite, a simple case is considered where the two subsystems have two lattice vectors in common, a and b. The incommensurability is then restricted to the z direction. The basic structures of the two subsystems have two different lattice parameters, c_1 and c_2, so that c_1/c_2 is irrational in the practical sense mentioned above. In the diffraction diagram, as shown in the example of **Figure 4**, one can distinguish the main reflections of the two subsystems which can be indexed with their respective reciprocal unit cells (a^*, b^*, c_1^*) and (a^*, b^*, c_2^*). Since both subsystems are modulated with the periodicity of the other one, additional satellite reflections are also present. The vectors c_2^* and c_1^* act as modulation wave vectors of the subsystem 1 and 2, respectively. If q is denoted as the vector c_2^*, then again all Bragg reflections H in the diffraction diagram can be indexed in the form of eqn [1], as shown in **Figure 4**. Satellites with index l and m different from zero cannot be explained but by the presence of the mutual modulation of both subsystems.

In general, an incommensurate phase may have several independent modulation wave vectors $\{q_j, j = 1, \ldots, d\}$. The indexation of the diffraction vectors then requires $(3 + d)$ indices and one speaks

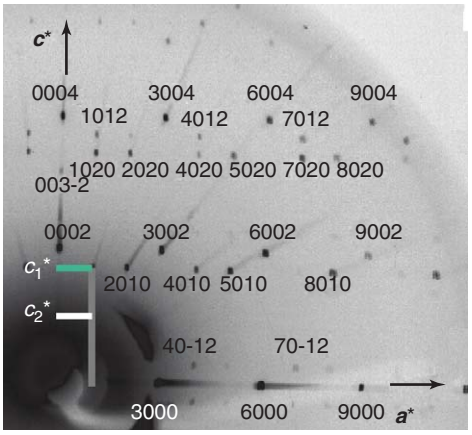


Figure 4 X-ray diffraction diagram of the incommensurate composite $Ba_{1+x}(Cu_xRh_{1-x})O_3$ with $x \approx 0.161$. The two incommensurate $\boldsymbol{c}^*$ parameters and the four integer indexation of the Bragg peaks are shown. (Courtesy of J Darriet.)

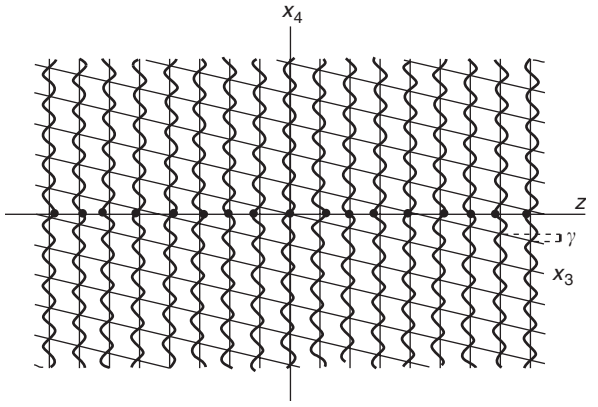


Figure 5 Projection on the plane x_3x_4 of the superspace density of a monoatomic structure with an incommensurate longitudinal sinusoidal modulation along the z-axis with wave vector $\boldsymbol{q} = \gamma \boldsymbol{c}^*$. Resulting aperiodic atomic positions are indicated by filled circles. (Courtesy of L Elcoro.)

of an incommensurate phase with a Fourier module of rank $(3+d)$.

The quantitative description of incommensurate phases cannot be done using the tools of conventional crystallography. Instead, a very efficient methodology has been developed introducing the concept of superspace. The real aperiodic structure is embedded within a $(3+d)$-dimensional space, the so-called superspace, as a section of a $(3+d)$-dimensional periodic structure. This allows the recovery of periodic crystallography, albeit in $(3+d)$ dimensions, and a straightforward description of the symmetry of these phases as a space group defined in superspace.

The superspace construction for the description of incommensurate phases is very easy to understand if one considers the Fourier spectrum of such systems. In general, one can say that their scattering density $\rho(\boldsymbol{r})$ can be developed in discrete Fourier series $\rho(\boldsymbol{r}) = \sum_H F(\boldsymbol{H})\exp(-\mathrm{i}2\pi \boldsymbol{H}\cdot\boldsymbol{r})$ with the set of diffraction vectors given by $\boldsymbol{H} = \sum_{i=1}^{n} h_i \boldsymbol{k}_i$, where $n = 3 + d$ (for simplicity, a notation is used that does not distinguish among the indexation basis vectors). A function in n variables can then be defined as

$$\rho_s(x_1, x_2, \ldots, x_n) = \sum_H F(\boldsymbol{H})\exp(-\mathrm{i}2\pi \sum_{i=1}^{n} h_i x_i) \quad [2]$$

where $F(\boldsymbol{H})$ are the Fourier amplitudes of the real 3D structure $\rho(\boldsymbol{r})$, and h_i the n indices of the corresponding vector $\boldsymbol{H}$. The n "phases" $(x_1, \ldots, x_n)$ define an n-dimensional superspace. By definition, the superspace density $\rho_s(x_1, \ldots, x_n)$ is periodic for each of its n variables with period 1, and directly defines the structure in real space $\rho(\boldsymbol{r})$ through the simple equation

$$\rho(\boldsymbol{r}) = \rho_s(\boldsymbol{k}_1\cdot\boldsymbol{r}, \ldots, \boldsymbol{k}_n\cdot\boldsymbol{r}) \quad [3]$$

Hence, a knowledge of the periodic superspace density within a single n-dimensional "primitive unit cell" and of the vector basis used in the indexation is sufficient for obtaining the aperiodic density in real space.

Note that the superspace is, in principle, an abstract space without a predefined metric. However, traditionally a particular metric is introduced which allows one to interpret eqn [3] from a geometrical viewpoint, so that the section of the superspace density $\rho_s(x_1, \ldots, x_n)$ given by eqn [3] does not only represent the real structure, but it does keeping its real space metric. In this way, the real space is imagined as a subspace of the defined superspace. In the following, this convention will be maintained. The Fourier spectrum of the incommensurate structure can then be interpreted as the projection in real space of the Fourier spectrum of the superspace density. Each vector $(h_1, \ldots, h_n)$ belonging to its reciprocal lattice has a component in real space that coincides with the real space diffraction vector $\boldsymbol{H} = \sum_{i=1}^{n} h_i \boldsymbol{k}_i$ and a perpendicular component in the complementary subspace, usually called internal space. For 1D modulated structures, the convention is to choose the metric so that the coordinates $x_4, \ldots, x_n$ corresponding to the modulation wave vectors $\boldsymbol{q}_i$, $i = 1, \ldots, d$ are associated to the internal space.

In **Figures 5** and **6**, two examples of the superspace description of incommensurate structures are depicted. They represent the projection on the plane x_3x_4 of a monoatomic structure with a sinusoidal and a

sawtooth longitudinal modulation along the z-axis, respectively. In both cases, the modulation wave vector can be expressed as $\boldsymbol{q} = \gamma \boldsymbol{c}^*$. The z-coordinates of the aperiodic point-like atomic sites along the z-axis in real space are represented by black dots, and are defined by the intersection with the real space z-axis of the sinusoidal or sawtooth lines representing the periodic superspace density. These two graphic examples illustrate the basic principle of the superspace construction. The section corresponding to the real space is irrationally oriented with respect to the superspace periodic lattice, making the real structure aperiodic. However, there is a one to one correspondence between all atomic sites in real space and points within a single unit cell in superspace. Each atomic position in real space is equivalent, by a lattice translation in superspace, to a single point within a primitive unit cell of the superspace density. The infinite set of all these points within a single n-dimensional unit cell defines the set of occupied atomic positions in real space. The superspace formalism would not be of any use if these points within the superspace unit cell would be more or less random and one should give the coordinates of all of them to define the structure. The crucial property of the superspace description (implicit in the discreteness of the diffraction diagram) is that, in general, these points form, in superspace, dense sets that can be described with few parameters. These sets are called atomic domains or atomic surfaces. In many cases, these atomic domains are continuous along the internal space, as the example of **Figure 5**. However, discontinuous atomic domains may also exist as in **Figure 6**, where the presence of a discontinuous sawtooth modulation yields in real space only two different interatomic distances along the modulation direction. More complex situations have also been hypothesized as, for instance, atomic domains with fractal character.

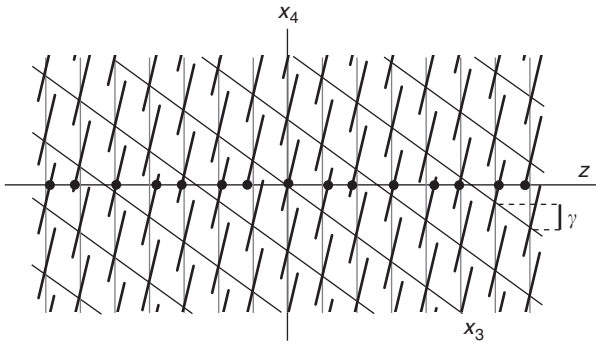


Figure 6 Projection on the plane x_3x_4 of the superspace density of a monoatomic structure with an incommensurate longitudinal sawtooth modulation along the z-axis with wave vector $\boldsymbol{q} = \gamma \boldsymbol{c}^*$. The resulting interatomic distances along z are restricted to two possible values. (Courtesy of L Elcoro.)

Once the phase problem has been solved and phases can be assigned to the experimental structure factor moduli $|F(\boldsymbol{H})|$, the superpace density can be retrieved from experimental diffraction data by means of so-called Fourier maps given by eqn [2]. **Figure 7** represents a section of the superspace density on the plane x_2x_4 corresponding to the bismuth and oxygen atoms of the Bi_2O_2 layers in $Bi_2Sr_2[Sr, Ca]Cu_2O_{8+x}$ (see **Figure 1**) as obtained from neutron diffraction experiments. The highly anharmonic modulations of both cations can be clearly seen. The strong oxygen modulation is discontinuous and the atomic domains form a sawtooth function. The flexibility of oxygen content in this compound is reflected in a variable size along the internal coordinate x_4 of these oxygen domains.

The symmetry associated to an incommensurate phase is given by a so-called superspace group. It can

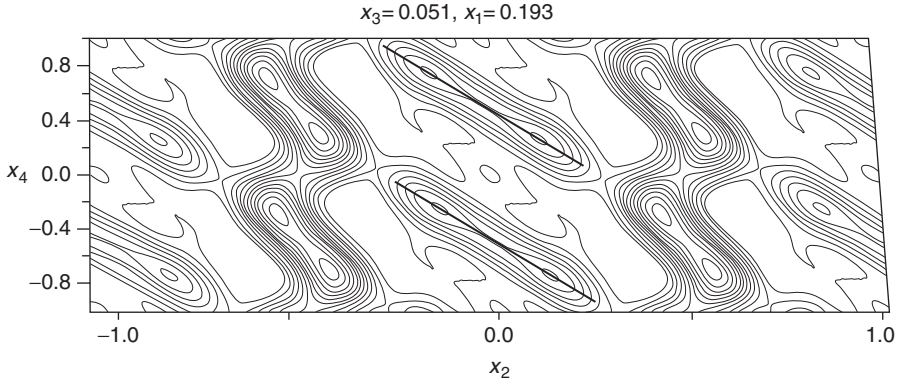


Figure 7 Fourier map corresponding to an x_2x_4 section showing the modulations of bismuth and oxygen in the Bi_2O_2 layers of $Bi_2Sr_2[Sr,Ca]Cu_2O_{8+x}$.The oxygen atomic domains are also drawn. (Reproduced with permission from Perez-Mato JM, Etrillard J, Kiat JM, Liang B, and Li CT (2003) Competition between composite and modulated configurations in $Bi_2Sr_2CaCu_2O_{8+x}$ and its relation to oxygen stoichiometry. *Physical Review B* 67: 024504; © American Physical Society.)

be easily derived from the properties of the corresponding Fourier spectrum, but it is difficult to visualize in direct space. It is described by operations that transform the real space structure in a new atomic configuration, different from the original one but physically indistinguishable. In superspace, these operations have a simple interpretation as space group operations in the n-dimensional superspace describing the symmetry of the superspace density $\rho_s(x_1, \ldots, x_n)$. However, not all n-dimensional space groups make sense as superspace groups. By definition, their rotational part should maintain the real space invariant. It is important to note that the superspace groups are not fortuitous geometrical relations, but they are thermodynamically stable in the sense that they are maintained within the whole range of a thermodynamic phase. They may only be broken via a phase transition, and in this sense, they are well-defined symmetries.

The Physical Origin of Incommensurate Phases

The systematic quantitative analysis of incommensurate phases started in 1976 with the study of Na_2CO_3 by de Wolff and collaborators. At this time, the existence of incommensurate modulated phases seemed an exceptional phenomenon. Four decades later, it is known that incommensurate phases are nearly ubiquitous and may appear in the phase diagram of systems of any type. During these years, the structure of more than 100 modulated crystals and about 50 modulated composites have been quantitatively characterized using the superspace approach, while the observation of modulated phases has been reported in many more. Incommensurate phases have been found in ionic, covalent, and molecular crystals, and also in alloys and metallic systems. Even numerous single chemical elements are now known to exhibit incommensurate phases at high pressures, either single modulated or composite. Incommensurate distortions are also present in important materials as high-T_c superconductors or ferroelectric relaxors.

An incommensurate configuration may be not only thermally stabilized, but it can also be the actual ground state of some systems. The mechanisms capable of producing the stabilization of modulated phases are quite various, and this variety of possible causes may be at the origin of the abundance of these phases in all types of materials. Incommensurate phases may in fact be stabilized through temperature, pressure, or composition.

The modulated structures of some metallic systems are attributed to a charge-density wave (CDW) mechanism, analogous to the Peierls transition. In a 1D metal, for instance, it may be energetically favorable to introduce a lattice modulation with a period π/k_F, k_F being the Fermi wave vector. This distortion opens up a gap just at the Fermi level, which decreases the electronic energy and may successfully compete with the increase of elastic energy. In general, for an arbitrary band filling, k_F is not commensurate with the underlying lattice and the modulation is incommensurate. In three dimensions, the phenomenon requires some additional (difficult to fulfil) topological conditions. In order to produce a significant gap at the Fermi level through the presence of a distortion with one or a few primary wave vectors, the topology of the Fermi surface must be such that these wave vectors connect significant dense sets of points at the Fermi surface. This is the so-called nesting condition and can only be satisfied by very peculiar Fermi surface topologies (see **Figure 8**). The nesting condition is easier to happen in systems closer to lower dimensionalities, that is, strongly anisotropic materials. Thus, typical CDW systems have chain or layer structures. Examples are transition-metal dichalcogenides, such as $NbSe_3$, NbS_3. It must be noted that this CDW–Peierls mechanism does not imply, in general, that the resulting incommensurate phase would be a semiconductor or insulator, as the gap may not be fully realized.

In some compounds, the incommensurate modulation is the result of the ordering of a minority component or of some vacancies. This modulation is then mainly occupational, that is, the occupation probability of some atomic sites varies along the modulation. In many cases, it may result in discontinuous atomic domains. This situation is rather common in nonstoichiometric compounds.

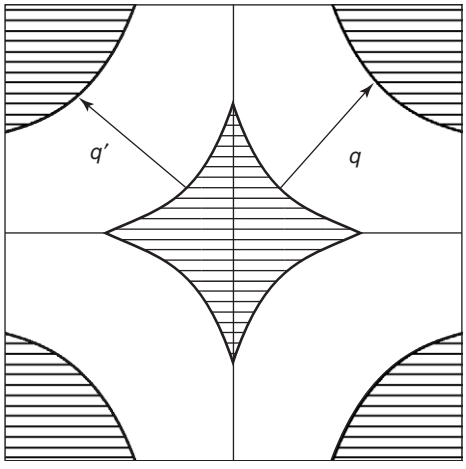


Figure 8 Scheme of a 2D Fermi surface fulfilling nesting conditions for two wave vectors.

Valence considerations and the bond valences sum rules, developed by ID Brown, have also been used to explain the stabilization of many incommensurate phases. In the case of metals such as calaverite ($Au_{1-p}Ag_pTe_2$), the need of a fluctuating valence for Au has been argued to be at the origin of the modulation. In other cases, as in Na_2CO_3, however, the requirement is of a single optimal valence for a particular cation, which is the main cause of the modulation. The average nonmodulated configuration can lead to chemically unfavorable sites for specific cations, yielding bond valence sums with significant deviations from the ideal values. As the ideal valences can be achieved by means of quite different coordination geometries, a modulation involving a correlated distortion of the coordination polyhedra corresponding to the problematic atomic sites may optimize the local chemistry of the structure. This has been shown to happen in many incommensurate insulators.

In some cases, the existence of a modulated phase can be understood using simple steric arguments and can be considered to be a purely lattice phenomenon with no direct involvement of the electronic degrees of freedom. For instance, in many materials rather rigid atomic polyhedra share vertices or edges and form an infinite framework. These framework structures possess, in many cases, specific distorting modes that are compatible with the rigidity of the polyhedral units forming the framework. These so-called rigid unit modes (RUMs) are candidates for being low frequency and even unstable modes of the system. They are usually constrained to specific modulation wave vectors within the Brillouin zone, that is, isolated points, lines, or surfaces. If the strongest instability of the systems corresponds to one RUM with a wave vector not fully fixed by symmetry, its freezing on the structure yields, in general, an incommensurate modulated phase of lower energy.

Similarly, in other cases, such as the family of orthorhombic compounds A_2BX_4, with K_2SeO_4 as an emblematic example, the optimal packing of some rigid ionic groups yields interstices, which are too large to accommodate the remaining atoms (cations). A modulation, either commensurate or incommensurate, can then palliate the problem optimizing the coordination of these interstitial sites. All A_2BX_4 compounds have a conspicuous soft phonon branch, which can be attributed to this steric effect. **Figure 9** schematizes the calculated variation of this phonon branch for different effective sizes of one of the two independent A cations. The branch is stabilized for larger A cations, while for small cation sizes, as those of potassium or rubidium, it softens and may even become unstable. In this latter case, an incommensurate frozen modulation of the wave vector around

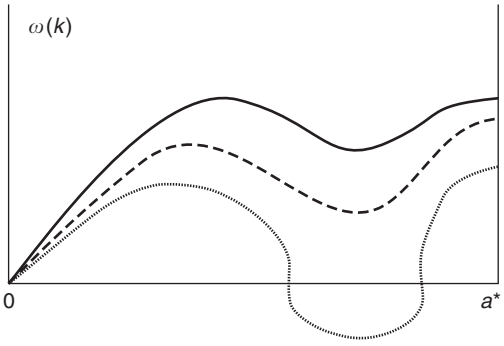


Figure 9 Scheme of the variation of the soft-mode branch in orthorhombic A_2BX_4 compounds as a function of the effective size of one of the independent A cations after the calculations in Etxebarria I, Perez-Mato JM, and Madariaga G (1992) *Physical Review B* 46: 2764. The branch lowers and becomes unstable at an incommensurate wave vector as the size increases. Imaginary frequency values are represented as negative.

$(2/3)\,a^*$ minimizes the energy. Clearly, a mismatch between the A cation size and the corresponding atomic site produced by the optimal accommodation of the more voluminous ionic units of the structure is at the origin of the incommensurate instability in these materials.

When atomic size effects are mainly involved in the incommensurate instability, thermal effects are very important as temperature can significantly change the effective size of atoms and the incommensurate phase is limited to some temperature range. For instance, in the case of A_2BX_4 compounds, the incommensurate modulation appears at a thermal phase transition as the temperature is lowered and the effective volume of the A cations decreases. In fact, the experimental soft phonon branch of K_2SeO_4 changes as the temperature is lowered in a form similar to the one given in **Figure 9**, which corresponds to a change of the size of the A cation.

It has become commonplace to talk about the existence of competing interactions as the ultimate origin of incommensurate phases. As practically all cohesive properties of any system could be traced back to the presence of competing interactions (repulsion–attraction), this does not seem very informative. In fact, the term "competing interactions" is used here in the sense that the basic interactions in these systems favor not one, but several incompatible and competing periodicities. The system is then frustrated and adopts a compromise configuration given by the incommensurate modulated structure. This simple concept is underlying the microscopic models developed to study the basic properties of the incommensurate phases. Among them, the so-called ANNNI model (axial next nearest neighbor Ising model) is the one that has been more thoroughly studied. This is essentially a 3D Ising model with

ferromagnetic nearest-neighbor coupling on one plane, while in the perpendicular direction, a ferro and antiferromagnetic coupling, J_1 and J_2, for nearest and next nearest neighbor, respectively, compete and favor different periodicities along this axis. **Figure 10** shows the approximate topology of the resulting mean-field phase diagram as a function of temperature and the relative strength of the ferromagnetic and antiferromagnetic interactions. The ground state is either ferromagnetic or commensurate modulated with a fourfold unit cell. At high temperatures the full disordered paramagnetic phase is stable, but as temperature is lowered, the system exhibits numerous phase transitions between different modulated phases including incommensurate ones. Similar results are obtained for more complex models as the so-called discrete frustrated Φ^4 model (DIFFOUR) where spins are substituted by more realistic local continuous variables on double wells.

Although the ANNNI model was initially designed to reproduce the properties of modulated magnetic materials, it has been successfully applied to understand the properties of purely structural modulated compounds, where the spin is associated to a local mode described as a pseudospin variable. In fact, a ferroelectric material, betaine calcium chloride dihydrate (BCCD), exhibits properties which strikingly reproduce those associated to the ANNNI model and some of its derivatives. On decreasing the temperature, the material passes through more than ten intermediate modulated phases, including two incommensurate ones, as the modulation wave vector locks into decreasing commensurate values before the system acquires a final ferroelectric nonmodulated configuration. This step-like decrease of the modulation wave vector, as the temperature is lowered, is given, under certain mathematical conditions, the name "Devil's staircase."

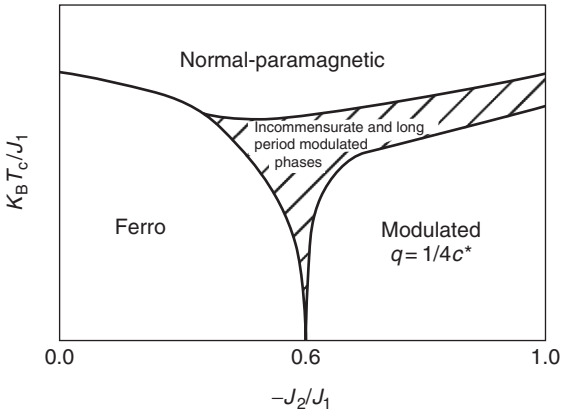


Figure 10 Schematic phase diagram of the ANNNI model within the mean field approximation.

The stabilization mechanism of incommensurate composites can also be quite varied. There are composites, as the inclusion compounds, where the interactions between the subsystems are rather weak, limited to molecular forces. Their mutual modulations are, therefore, in this case, rather small. In other compounds, the interaction is much stronger and there may even exist chemical bonding between atoms belonging to different subsystems, as in $(LaS)_{1.13}TaS_2$.

One could expect that by their own definition, nonstoichiometry is an essential component of composites. Two incommensurate atomic subsystems yield necessarily a noninteger ratio between their respective compositions. However, the interaction between both subsystems can be so strong that the actual modulation may include chemical substitution, yielding a perfectly stoichiometric (but incommensurate) composite. The family of commensurate and incommensurate trigonal compounds of general formula $A_{1+x}[A'_xB_{1-x}]O_3$ is a nice example, where this type of chemical interaction between the subsystems plays a dominant role. **Figure 4** above shows a typical diffraction pattern of one of these compounds, and **Figure 11** illustrates their main

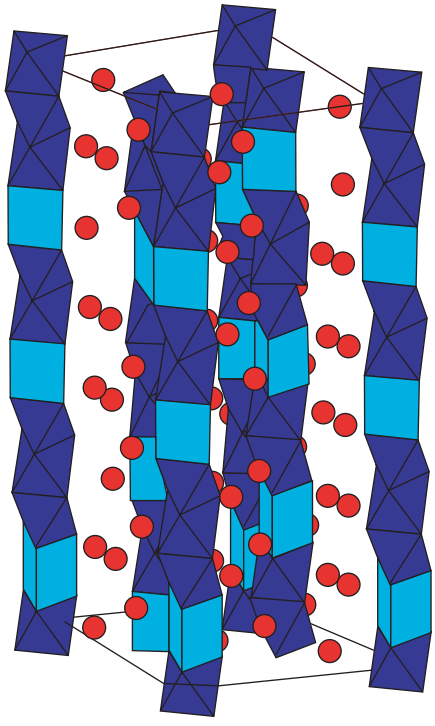


Figure 11 Approximate structure of a composite of type $A_{1+x}[A'_xB_{1-x}]O_3$ with $x \approx 3/11$.The sequence of octahedra and prisms along the $[A'_xB_{1-x}]O_3$ columns is aperiodic but locally follows the sequence $\langle 3\,1\,3\,1\,2\,1 \rangle$ corresponding to the exact commensurate composition $x = 3/11$. The periodicity and this commensurate composition are lost by "faults" in this sequence.

structural features. These trigonal materials are closely related to the 2H hexagonal perovskites ABO_3. The groups $A'O_3$ and BO_3 forming columns along the trigonal axis constitute the first subsystem, while the second subsystem is given by the cations A forming chains among these columns. The mismatch between the two subsystems is produced by the excess of A cations, in a ratio $1+x$ with respect to the number of $[A', B]O_3$ units. However, stoichiometry is maintained through the substitution of B cations by divalent A′ cations, the composition of the second subsystem being $[A'_xB_{1-x}]O_3$ (A and B are typically divalent and tetravalent, respectively). Hence, the presence of the A′ cations within the columns $[A', B]O_3$, visible in **Figure 11** through their trigonal prismatic oxygen coordination, can be considered the fundamental modulation of the $[A', B]O_3$ subsystem. In some cases, A′ and B cations can be the same element in two oxidation states and then one can talk of a valence modulation. Thus, stoichiometry is ensured, but the system is incommensurate, as the sequence of octahedra and prisms along the $[A', B]O_3$ columns is, in practice, aperiodic for x values not assimilable to simple fractions.

Landau Phenomenology

There are a significant number of compounds that, after becoming incommensurate at a thermal phase transition, remain incommensurate up to the lowest accessible temperatures (for instance PbO). In most cases, however, the range of the incommensurate phases stabilized by temperature is usually limited by two commensurate phases, a high-temperature or parent (nonmodulated) phase, and a low-temperature "lock-in" phase, where the modulation wave vector "locks" into a certain rational value $\boldsymbol{q}_c$. In the generalized Landau theory of phase transitions, extended to include incommensurate phases, both the lock-in and the incommensurate phase are usually described using the same order parameter. It is homogeneous in the first case and modulated in the second one.

One of the simplest and most common situations occurs when the lock-in wave vector lies within a symmetry line of the Brillouin zone and the associated order parameter is 2D. The order parameter can then be considered a complex number, $Q = \rho\exp(i\theta)$, representing the complex amplitude of the primary distorting mode in the lock-in phase. The incommensurate phase is then described by a space-dependent order parameter $Q(x)$ along the direction of the wave vector. This is the case of the A_2BX_4 compounds and many other incommensurate insulators. The corresponding Landau free energy includes two essential terms: a Lifshitz invariant of type $(Q\partial Q^* - Q^*\partial Q)$ (where ∂ means a derivative with respect to x), and an umklapp term, $Q^p + Q^{*p}$, of relatively low order p. These terms favor inhomogeneous and homogeneous configurations of the order parameter, respectively. The Lifshitz invariant is the essential cause of the stability of the incommensurate phase, while the second one favors the lock-in phase. The properties predicted by this generalized Landau potential agree rather well with the experimental results. As temperature is lowered, after a second-order transition from the parent to the incommensurate phase, the competition between the Lifshitz and the umklapp terms drives the modulation from an initial sinusoidal configuration into a strong anharmonic regime, the so-called soliton regime, previous to the lock-in phase transition. In this soliton regime, the modulation is soliton-like with regions, periodic along the modulation direction, where the order parameter is approximately constant (i.e., regions approximately commensurate corresponding to the lock-in phase) separated by coherent periodic discommensurations, where the phase $\theta(x)$ of the complex order parameter changes rapidly to the next "constant" value. In the simplest approximation, this step-like behavior of the phase $\theta(x)$ satisfies the sine-Gordon equation, a well-known equation in nonlinear physics. The density of the discommensurations, the so-called soliton density n_s, can be taken as an order parameter with respect to the lock-in phase. A zero soliton density would correspond to the lock-in phase while $n_s = 1$ is associated to the perfect sinusoidal modulation. Hence, the system approaches the lock-in phase in two rather different and complementary ways: (1) approaching the modulation wave vector to the commensurate value $\boldsymbol{q}_c$, and (2) transforming progressively the structural modulation into a step function corresponding to $n_s = 0$. Experimentally, the lock-in transition has a discontinuous character and takes place before either the modulation wave vector or n_s can reach their limit values. However, soliton regimes in the incommensurate phase with n_s as low as 0.2 have been detected.

Figure 12 shows the measured diffracted intensity around the satellite $(6\ 0\ 1\ -1)$ of Rb_2ZnCl_4 at 6 K above the lock-in transition, well within the soliton regime. The modulation wave vector in this case is about $(1/3-0.018)\boldsymbol{c}^*$. The signature of the strong anharmonic step-like atomic modulations is the presence of very high-order satellites around the stronger first-order $(m = -1)$ satellite. Satellites up to $m = 7$ can be observed. The rather relative strong intensity of the satellites with $m = 5$ and 7 is due to symmetry restrictions on the atomic modulations that favor Fourier components of order $6n \pm 1$. Note that at

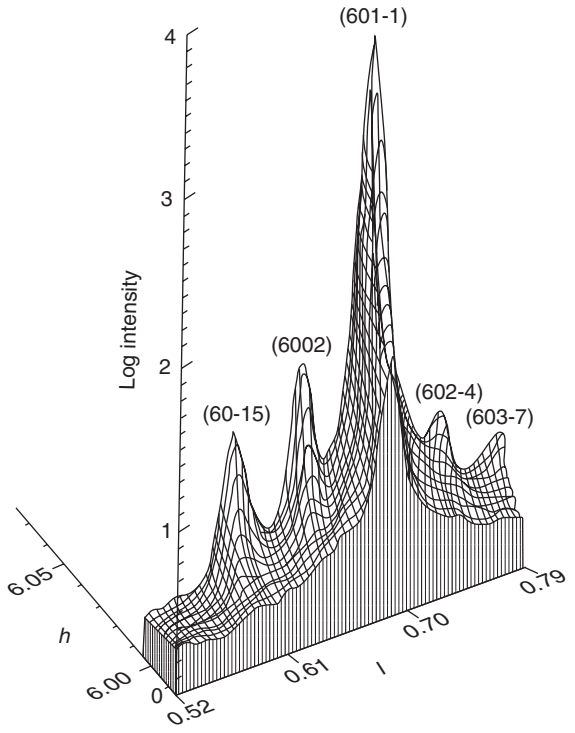


Figure 12 X-ray intensity diffracted by R_2ZnCl_4 around the satellite (6 0 1 $-$1) at 198 K showing intense high-order satellites close to the lock-in phase transition. (Reproduced with permission from Aramburu I, Madariaga G, Grebille D, Perez-Mato JM, and Breczewski T (1997) *Journal of Physics I France* 7: 371–383; © EDP Sciences.)

the lock-in phase transition where the modulation wave vector takes the commensurate value $(1/3)\boldsymbol{c}^*$, all the satellites observed in the figure will collapse into a single superstructure reflection centered at $6\boldsymbol{a}^* + (2/3)\boldsymbol{c}^*$.

The incommensurate phases that include a Lifshitz invariant in the Landau free-energy density are usually called type I, and type II otherwise. The simplest cases of type II phases are those for which the lock-in commensurate structure is associated with a 1D order parameter whose wave vector is null. The modulated order parameter in the incommensurate phase is, therefore, also usually taken as 1D. The free-energy density lacks a Lifshitz invariant, but the incommensurate modulation of the order parameter can be explained in this case by the coupling of the order parameter with a "secondary" mode. If $\eta(x)$ is called the x-dependent order parameter and $\xi(x)$ another local degree of freedom, a term of the type $(\partial\eta\xi - \eta\partial\xi)$ in the free-energy density, if allowed, favors a modulated configuration. It plays a similar role as a Lifshitz invariant and in fact, it can lead, as in thiourea, to a soliton regime similar to those observed in type I materials. The incommensurate phase will be given in this case by a modulation of

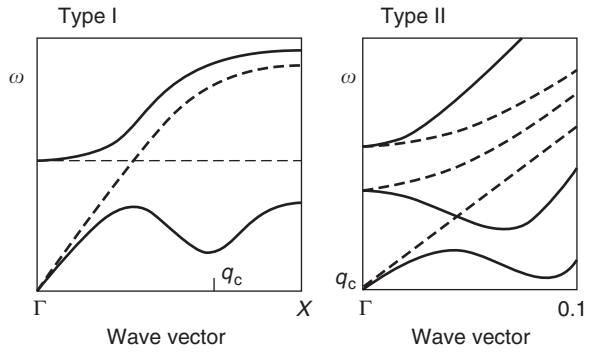


Figure 13 Idealized scheme of the soft-mode branch for the simplest cases of type I and type II incommensurate systems above the normal-incommensurate phase transition, including its possible coupling with an acoustic branch.

not only the nominal order parameter $\eta(x)$ but also of the secondary mode $\xi(x)$.

From a lattice dynamics viewpoint, a Landau free energy of type I describes a thermal soft-mode mechanism with a lock-in wave vector having no additional symmetry compared with the neighboring incommensurate wave vector, while a type II describes a soft-mode having the lock-in value $\boldsymbol{q}_c$ at a point of higher symmetry. This difference is schematically represented in **Figure 13**. In both cases, either the Lifshitz invariant or the alternative mixed coupling term places the minimum of the soft-mode branch off the commensurate value $\boldsymbol{q}_c$. While in type I, one can neglect the change of the mode eigenvector as the modulation wave vector changes and locks into $\boldsymbol{q}_c$, in type II a strong change of the modulation eigenvector is an essential part of the approach to the lock-in phase. The hybridization of the soft-mode branch with the secondary mode necessarily disappears at the lock-in into $\boldsymbol{q}_c = 0$. A strong coupling with an acoustic branch may also play a fundamental role in the phase-transition mechanism. This possibility has also been included in the schemes of **Figure 13**. In type I crystals as K_2SeO_4, in contrast with the earlier models, the anticrossing with the acoustic branch of the soft mode branch is an essential part of the physics of the transition and can even explain the incommensurate minimum of the hybridized soft-phonon branch.

Specific Physical Properties

Incommensurate phases possess, in principle, low-energy lattice dynamical excitations (zero-gap branches) in addition to the acoustic phonons. They are usually termed phasons and correspond in the superspace description to fluctuations of the superstructure along the internal subspace. The incommensurability

of the structure ensures that the phase of the structural modulation associated to each independent incommensurate vector $q_i, i = l, \ldots, d$ can be shifted without changing the structure globally and, therefore, without energy cost. In the superspace, this means that any section parallel to the one corresponding to the real space, but shifted along the internal space, describes a physically equivalent structure with the same energy. This is the cause of these peculiar phason degrees of freedom. Phason branches have been detected experimentally in several materials although they may be overdamped in many cases. One of the main discussions has been to assess the existence of a gap in these phason branches. In principle, for an ideal incommensurate system it should be zero, but defects or the underlying average lattice may produce some pinning of the modulation and the opening of a gap. A pinning of the modulation also exists if the modulation is not analytical, that is, when the atomic domains are discontinuous along the internal space. This "break of analyticity" may happen, for instance, in the soliton regime as the modulation becomes step-like when approaching the lock-in phase. It also exists in strong interacting composites. On the other hand, in the sinusoidal regime, close to the second-order normal-incommensurate transition, pinning effects of any sort are expected to be minimal.

The strong pinning effects expected in the soliton regime are considered to be the cause of special thermal hysteresis effects observed around the lock-in phase transition and termed global hysteresis. Also so-called memory effects associated to annealings within the incommensurate phase have been attributed to pinning effects produced by mobile lattice defects.

It has been argued that the existence of the phason degrees of freedom in the incommensurate phase can increase, compared with other structural phase transitions, the size of the critical region around the normal-incommensurate phase transition where universal critical behavior could be observed. Although the reasoning behind this expectation has been the object of much controversy, it is indeed an empirical fact that the normal-incommensurate phase transitions are among the few structural phase transitions, if not the only ones, where fully consistent fluctuation-driven asymptotic universal critical behavior has been measured. A thoroughly studied example is the continuous normal-incommensurate phase transition in Rb_2ZnCl_4 at about 304 K. Results of NMR, synchrotron X-ray scattering, specific heat measurements, and other techniques are all fully consistent and have demonstrated an asymptotic critical behavior corresponding to the universal class associated to the 3D XY model in a temperature interval as large as 15 K below the transition temperature (see **Figure 14**).

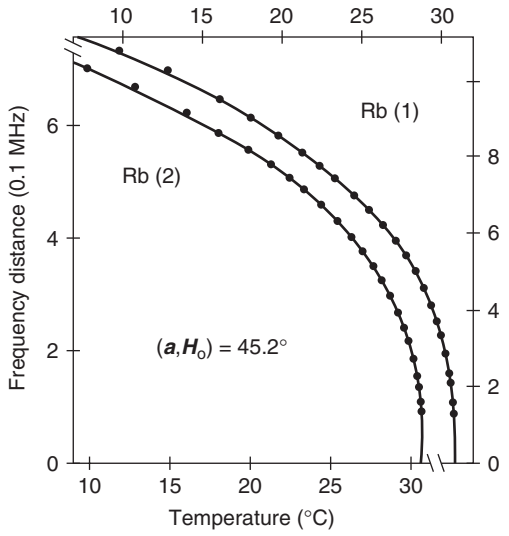


Figure 14 Temperature dependence of the width of the continuous distribution of quadrupolar perturbed NMR frequencies of the two symmetry-independent Rb atoms in the incommensurate phase of Rb_2ZnCl_4. This width is proportional to the amplitude of the order parameter. The fit curves correspond in both cases to a power law with $\beta = 0.35$. (Reproduced with permission from Walisch R, Perez-Mato JM, and Petersson J (1989) NMR determination of the nonclassical critical exponents *beta* and *beta-bar* in incommensurate Rb_2ZnCl_4. *Physical Review B* 40: 10747–10752; © American Physical Society.)

See also: Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Structural Instability and Soft Modes; Periodicity and Lattices; Quasicrystals.

PACS: 61.44.Fw; 64.70.Rh; 61.50.Ah; 61.50.Ks; 65.90.+i

Further Reading

Blinc R and Levanyuk (eds.) (1986) *Incommensurate Phases in Dielectrics*, vol. 1 and 2. Amsterdam: North-Holland.

Bruce AD (1980) Structural phase transitions II. Static critical behaviour. *Advances in Physics* 29: 111–217.

Cummins HZ (1990) Experimental studies of structurally incommensurate crystal phases. *Physics Reports* 185: 211–411.

Janssen T and Janner A (1987) Incommensurability in crystals. *Advances in Physics* 36: 519–624.

Janssen T, Janner A, Vos-Looijenga, and de Wolff PM (1995) Incommensurate and commensurate modulated structures. In: Wilson AJC (ed.) *International Tables for Crystallography*, vol. C, sect. 9.8. Dordrecht: Kluwer Academic.

Perez-Mato JM, Zuñiga FJ, and Madariaga G (eds.) (1991) *Methods of Structural Analysis of Modulated Structures and Quasicrystals*. Singapore: World Scientific.

Selke W (1998) The ANNNI model. Theoretical analysis and experimental application. *Physics Reports* 170: 213–264.

Tolédano JC and Tolédano P (1987) *The Landau Theory of Phase Transitions*. Singapore: World Scientific.

Van Smaalen S (1995) Incommensurate crystal structures. *Crystallography Reviews* 4: 79–202.

Withers RL, Schmid S, and Thompson JG (1998) Compositionally and/or displacively flexible systems and their underlying crystal chemistry. *Progress in Solid State Chemistry* 26: 1–96.

Nomenclature

$\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*$	basis of reciprocal unit cell associated to the average periodic structure of a modulated structure
$F(\boldsymbol{H}), \lvert F(\boldsymbol{H})\rvert$	structure factor and its modulus for the diffraction vector $\boldsymbol{H}$
$\boldsymbol{H}$	diffraction vector of a Bragg reflection
(h,k,l,m)	set of four integers indexing a Bragg reflection of an incommensurate phase of rank $(3+1)$
$\boldsymbol{q}$	primary wave vector of a modulation
$\{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_d\}$	set of primary modulation wave vectors for an incommensurate modulated structure of rank $(3+d)$
$\boldsymbol{r}$	position vector in real space
$\rho(\boldsymbol{r})$	scattering density in real space (electronic or nuclear, depending on the radiation source) corresponding to an incommensurate phase
$\{\boldsymbol{k}_1,\boldsymbol{k}_2,\ldots,\boldsymbol{k}_n\}$	basis of vectors indexing the diffraction diagram of an incommensurate phase of rank n
$(x_1,x_2,\ldots,x_n)$	position vector in superspace
$\rho_s(x_1,x_2,\ldots,x_n)$	superspace density
$(h_1,h_2,\ldots,h_n)$	set of n integers indexing a diffraction vector of an incommensurate phase of rank n
k_F	Fermi wave vector
$\boldsymbol{q}_c$	modulation wave vector of a lock-in phase
$Q(x)$	local complex order parameter in a simple type I incommensurate phase as a function of the x coordinate along the modulation direction
$\rho(x)$	amplitude of $Q(x)$
$\theta(x)$	phase of $Q(x)$
n_s	soliton density
∂	partial derivative with respect to x
$\eta(x)$	one-dimensional order parameter of a type II incommensurate phase as a function of the x coordinate along the modulation direction
$\xi(x)$	real amplitude of a secondary mode in a type II incommensurate phase as a function of the x coordinate along the modulation direction

Insulators, Electronic States of

F Casula and G Mula, Cagliari University, Cagliari, Italy

Introduction

Materials are traditionally classified according to their conduction properties: they are called metals if their conductivity persists even to the lowest temperatures, insulators if they do not. In the former case electrons flow freely under an applied DC field, in the latter they do not, but undergo static polarization.

In quantum mechanics, however, this distinction is somewhat blurred, since materials are usually understood in terms of a band theory picture, that is, in terms of the solution of a many-electron problem within the one-electron approximation. Within this approximation, electrons behave as independent particles embedded in a mean field which includes their interactions. Theoretical calculations thus make no distinction between metals and insulators: solids in which all occupied energy bands happen to be completely filled are insulators, and solids in which at least one occupied energy band happens to be only partly filled are metals. It is on the basis of this picture that it has been possible, at least since the 1980s, to deal on a first-principles basis with any kind of materials, irrespective of their conduction properties, including, of course, the ever more relevant case of semiconductors.

This quantum approach has led to reliable quantitative predictions for a wide range of materials and of materials properties. The decisive breakthrough has been achieved essentially, thanks to the advent of density-functional methods. These methods are based on a theorem, formulated by Hohenberg and Kohn in 1964, which states that all the ground-state properties of a system of interacting electrons, possibly embedded in an external potential $v(\boldsymbol{r})$, are completely determined by the electronic charge density $\rho(\boldsymbol{r})$. This theorem leads to an extraordinary simplification of the many-electron problem, since one needs to deal with only three spatial variables instead of the $3N$ variables needed to specify the ground-state wave function of a system of N interacting electrons. Hohenberg and Kohn also proved that

the ground-state energy of an interacting electron gas is a unique functional of $\rho(\boldsymbol{r})$, thereby paving the way for the subsequent advances in 1965 by Kohn and Sham (KS), who were able to build self-consistent equations for a fictitious system of noninteracting electrons with the same electron density as the interacting one. Solving these equations also gives, in principle, the exact solution to the full problem.

Thanks to this approach, the band theory has been shown to be able to successfully explain even the limiting cases of insulating behavior, namely carbon, solid noble gases, and alkali halides. The validity of the one-electron picture has been thus assessed independently of the type of chemical bond, whether covalent, ionic or molecular, and some of these band structure results for insulators are discussed in the following, while referring to other articles in this encyclopedia for more information on band-theory density-functional methods.

Successful as it can be, however, this theoretical framework cannot be the whole story, since in many cases, for example, when the insulating behavior is dominated by disorder (the so-called Anderson insulators), or by electron correlation (the so-called Mott insulators), the band approximation fails to give a consistent account of the ground-state properties. In particular, it fails to explain polarization and localization, the two basic phenomenological features that characterize insulators. The following sections present very recent theoretical developments which have made it possible to understand localization and polarization as two aspects of the same phenomenon, within a unified approach that deals on the same basis with all insulators, irrespective of what determines the insulating behavior – independent or correlated electrons, crystalline characteristics or disorder.

Band Structure Results for Insulators

Basically, energy band calculations for insulators are done according to the same methods which are used for semiconductors. In principle, accurate values for equilibrium electronic properties in condensed matter (such as, e.g., volume, bulk modulus, elastic constants), as well as for linear response (such as phonon spectra) and other thermodynamical properties, can be obtained by solving KS equations within the local density approximation (LDA), independently of whether the material studied is metallic, semiconductor or insulating.

Excitation properties, however, are often inaccurately determined both in semiconductors and insulators, due to the fact that LDA works well only for the ground state. The optical gap, in particular, is usually severely underestimated, as has been shown in the early 1980s by the extensive LDA calculations on group IV elements performed by M L Cohen and co-workers and published in *The Physical Review*. Nevertheless, besides obtaining very accurate results for the cohesive energies and the equilibrium structures, they were able to state unambiguously the insulator or semiconductor quality of both C and Si respectively, in the diamond structure, in spite of the somewhat poor results for their band gaps (4.1 and 0.5 eV for C and Si, respectively, against the experimental values of 5.49 and 1.17 eV). In order to consider the achievements of the band structure calculations in the field, it may be recalled that the semimetallic character again of carbon, but in graphite structure, was already obtained in the sixties by F Bassani and G Pastori and confirmed ever since. A further confirmation of the band picture of the electronic properties of these materials is the overall success of the energy gap corrections introduced via Green's function approaches (GW approximation).

Similar considerations hold for rare-gas solids, whose experimental gaps range from 9.3 eV in Xe to 21.7 eV in Ne. By taking into account the closed-shell character of the constituents atoms, it is not difficult to understand in a band picture that conduction states will result from free-electron like bands, widely separated from valence bands for orthogonality reasons. Again, LDA estimation of the energy gap can be as poor as 11.5 eV for Ne, to be improved up to 22.2 eV after the introduction of self-interaction corrections – as was done by M R Norman and J P Perdew in 1983 – but the insulating behavior cannot be questioned.

A third class of wide-gap crystals can also be considered, that is, alkali halides, which are characterized by huge values of the Madelung potential, and, therefore, by large energy separation between valence and conduction bands. Experimental bandgaps can be as large as 13.6 eV in LiF and smaller but still very large in iodides, where typical values are $\sim$6.0 eV. From the theoretical point of view, for example, considered the results for NaCl by Norman and Perdew, quoted above, which obtained for the gap a clear-cut value of 9.2 eV, to be compared with the experimental values of 8.5–9.0 eV. Despite the successes of the band theory in explaining the electronic properties of insulators, some problems remain, mainly connected to the description of the ground state. In a band theory framework, the difference in DC conductivity between metals and insulators is explained with reference to the spectrum of the electronic excitations, that is, to the availability of low-lying unoccupied energy levels. However, it makes no mention of localization, that is, of the

qualitative difference in the ground-state electrons organization which has to account for the fact that electrons in metals are able to flow freely under an applied DC field, while electrons in insulators are not. Neither can the band theory readily explain why finite macroscopic metallic samples subject to an electric field only show a polarization due to surface phenomena (screening by free carriers), while in analogous situations insulator samples show a material-dependent polarization which is a true bulk phenomenon.

These difficulties stem from the fact that electron states, being eigenfunctions of Hamiltonian operators subject to periodic Born–von Karman boundary conditions, are delocalized, and qualitatively rather similar, in both metals and insulators, so that it is a real challenge to detect differences between these two classes of materials just by looking at their one-electron ground-state wave functions.

Localization and Polarization in Insulators and Metals

W Kohn was the first, in 1964, to answer this challenge by pointing out that an essential characteristic of insulators is the disconnectedness of their ground-state many-body wave function, meaning its breaking up into a sum of functions which are localized in disconnected regions. This feature forbids the free flow of charge in insulators under the application of a DC field, so that electrons can be polarizable while remaining bound. Kohn demonstrated this result in a crystalline system of independent electrons, but argued that the same kind of disconnectedness should also hold quite generally, in any insulator.

Kohn's theory correctly points out that electronic localization in insulators is a feature of the ground-state wave functions, which, for an N-particle system, are disconnected in a high-dimensional configuration space. Therefore, localization cannot be seen by looking at the charge density, which is a function in three-dimensional space. His approach, however, leaves alone the problem of macroscopic polarization, which until 1992 even failed to have a proper definition as a bulk property independent of surface termination. Since 1992, a new theory of polarization has been developed which builds on the concept that polarization must be understood as a global property of the many-body ground-state wave function as a whole. According to this viewpoint, localization is just another aspect of polarization, and can be rightly regarded as the cause of insulating behavior since it is a physical observable completely independent of the periodic charge distribution of the valence electrons.

The 1992–93 Breakthrough

What happened around 1992 is that people came to realize that the classic definition of macroscopic polarization as the dipole of a macroscopic sample, divided by its volume, simply does not hold when periodic boundary conditions are applied (though polarization derivatives had been successfully addressed for many years and for many materials in the framework of linear-response theory). Under these conditions, which are almost mandatory in most computational condensed matter physics, the dipole of a macroscopic charge distribution cannot be evaluated since its expression for finite systems, while perfectly valid, does not have a well-defined thermodynamic limit independent of surface termination. The clue to a solution of this problem came from realizing that nobody ever measures the "absolute" macroscopic polarization of a material as a bulk property, but only polarization differences. Therefore, the modern theory of polarization only deals with finite polarization differences, avoiding any definition of absolute polarization.

Consider for example, a possible experimental arrangement for measuring the macroscopic polarization in a piezoelectric material, in which the crystal is uniaxially strained while being short-circuited in a capacitor. The induced macroscopic polarization is obtained by integrating the transient macroscopic current flowing through the sample. It must be stressed that polarization in a zero field is considered here, as is mandatory if one wants to preserve crystal periodicity.

This approach is also relevant for the analysis, at zero field, of the effective charges of lattice dynamics, as well as of ferroelectric polarization. Basically, all these situations lead to considering macroscopic polarization differences, given expressions like

$$\Delta \boldsymbol{P}_{\text{bulk}} = \int \mathrm{d}\lambda \, \frac{\delta \boldsymbol{P}^{(\lambda)}_{\text{bulk}}}{\delta \lambda} \qquad [1]$$

where λ may be thought of as an adiabatic time in the piezoelectric case, or as the amplitude of the distortion for ferroelectric materials. In all cases, $\delta \boldsymbol{P}^{(\lambda)}_{\text{bulk}}/\delta\lambda$ is just the spatially averaged adiabatic current flowing through the bulk.

Focusing on the current instead of the charge is a crucial step for a proper definition of macroscopic polarization as a physical observable in extended systems. The introduction of an integration over λ in the expression for $\Delta \boldsymbol{P}_{\text{bulk}}$ is equivalent to assuming a parametric dependence in the system Hamiltonian. This feature allows the introduction of a Berry's phase formulation, and leads to a satisfactory solution of the problem.

Though a parametric Hamiltonian could not properly be regarded as isolated, since the presence of a parameter is an expression of the coupling with "the rest of the universe" (to use Michael Berry's expression), this trick allows calculations to be done assuming just this isolation. As a trade-off it is observed that the macroscopic polarization can indeed be regarded as a proper physical observable, but only when defined as a Berry phase, that is as a gauge-invariant geometric phase which is observable even if it cannot be expressed in terms of the eigenvalues of any operator.

Berry Phase

Consider a generic parametric Hamiltonian $H(\xi)$ such that

$$H(\xi)|\Psi(\xi)\rangle = E(\xi)|\Psi(\xi)\rangle \qquad [2]$$

where ξ defined in a suitable domain and $|\Psi(\xi)\rangle$ is the ground state of the system, nondegenerate for any ξ. The phase difference $\Delta\phi_{ij}$ between ground eigenstates belonging to two different values of ξ can then be defined as

$$e^{-\Delta\phi_{ij}} = \frac{\langle\Psi(\xi_i)|\Psi(\xi_j)\rangle}{|\langle\Psi(\xi_i)|\Psi(\xi_j)\rangle|} \qquad [3]$$

$$\Delta\phi_{ij} = -\mathrm{Im}\log\langle\Psi(\xi_i)|\Psi(\xi_j)\rangle \qquad [4]$$

Since the phase of any quantum-mechanical state vector is arbitrary, this phase difference cannot have any physical meaning. However, when one considers a discretized closed path in the space of the parameter ξ, one immediately realizes that the total phase difference $\gamma = \sum_{\mathrm{cycle}} \Delta\phi_{i,i+1}$ along the path must be independent from the (arbitrary) phase choices at each point in the path, since in the end all of them cancel out. γ is called a Berry phase because it is a gauge-invariant quantity, meaning that it is a physical observable, albeit a rather anomalous one, since it obviously cannot be expressed in terms of the eigenvalues of Hermitian operators.

It can be seen that the Berry phase is relevant to the modern theory of polarization because it substitutes an ill-defined (when periodic boundary conditions are applied) dipole expression with a computable gauge-invariant quantity. An alternative, of course, would be to define macroscopic polarization for finite systems and to proceed to the elimination of surface effects through suitable limiting procedures as the system size goes to infinity.

Berry Phase Theory of Polarization

Evaluating polarization for finite systems is, in principle, fairly straightforward: for instance, one can introduce the position operator as $\boldsymbol{R} = \sum \boldsymbol{r}_i$ and compute the electric dipole as $\boldsymbol{d} = q_e\langle\Psi_0|\boldsymbol{R}|\Psi_0\rangle$. This expression, however, becomes meaningless when periodic boundary conditions are imposed, since in the resulting Hilbert space $\boldsymbol{R}$ becomes a forbidden operator. The problem, obviously, is that when one applies $\boldsymbol{R}$ to a function obeying these conditions, that is to a periodic function, one gets a new function which is no longer periodic and, therefore, does not belong to the given Hilbert space.

A way out of these difficulties was found by Resta in 1992, when he introduced expression [1] to calculate finite changes in bulk polarization. According to classical electrodynamics, the polarization current $\delta\boldsymbol{P}^{(\lambda)}_{\mathrm{bulk}}/\delta\lambda$ in eqn [1] can be expressed as

$$\frac{\delta\boldsymbol{P}^{(\lambda)}_{\mathrm{bulk}}}{\delta\lambda} = \frac{1}{V}\int_V \mathrm{d}\boldsymbol{r}\, \boldsymbol{j}^{(\lambda)}_{\mathrm{bulk}} = \boldsymbol{J}^{(\lambda)} \qquad [5]$$

where V is the volume of the system, $\boldsymbol{j}^{(\lambda)}_{\mathrm{bulk}}$ is the current density in the bulk, with λ playing the role of time. The right-hand side could be expressed as the adiabatic limit of a Kubo formula, which usually involves a summation over the excited eigenstates. It is more convenient, however, to arrive at a form which only depends on the ground-state wave function. To this end, one imposes the following boundary conditions:

$$\Psi^{(\lambda)}_{\boldsymbol{k}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i + \boldsymbol{L}, \ldots, \boldsymbol{x}_N) = e^{i\boldsymbol{k}\cdot\boldsymbol{L}}\Psi^{(\lambda)}_{\boldsymbol{k}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_N) \qquad [6]$$

to label the N-body wave function $\Psi^{(\lambda)}_{\boldsymbol{k}}$. (This way of introducing boundary conditions has recently come to be called "twisted," to avoid confusion with all the other usages of the word "phase" that appear in quantum and statistical physics.) One also introduces the wave function $|\Phi^{(\lambda)}_{\boldsymbol{k}}\rangle = e^{-i\boldsymbol{k}\cdot\boldsymbol{X}}|\Psi^{(\lambda)}_{\boldsymbol{k}}\rangle$ (it is none other than the many-body analog of the cell-periodic part of the single-electron Bloch function). If $\Psi^{(\lambda)}_{\boldsymbol{k}}$ is the ground-state of Hamiltonian $\boldsymbol{H}^{(\lambda)}$, then $\Phi^{(\lambda)}_{\boldsymbol{k}}$ is the ground-state of the Hamiltonian

$$\boldsymbol{H}^{(\lambda)}(\boldsymbol{k}) = e^{-i\boldsymbol{k}\cdot\boldsymbol{X}}\boldsymbol{H}^{(\lambda)}e^{i\boldsymbol{k}\cdot\boldsymbol{X}} \qquad [7]$$

which can be obtained from $\boldsymbol{H}^{(\lambda)}$ by performing a gauge transformation $\boldsymbol{p}_j \to \boldsymbol{p}_j + \hbar\boldsymbol{k}$ on the momentum of each particle. The $\boldsymbol{k}$-dependence has thus been transferred from the boundary conditions to the Hamiltonian. Using functions $\Phi^{(\lambda)}_{\boldsymbol{k}}$, the Kubo formula for the adiabatic current $\boldsymbol{J}^{(\lambda)}_{\boldsymbol{k}}$ can indeed be expressed, for

strictly periodic boundary conditions, in a form which depends only on the ground-state wave function:

$$\boldsymbol{J}_{\boldsymbol{k}}^{(\lambda)} = 2q_{\mathrm{e}}\mathrm{Im}\langle \partial_{\boldsymbol{k}}\Phi_{\boldsymbol{k}}^{(\lambda)}|\partial_{\lambda}\Phi_{\boldsymbol{k}}^{(\lambda)}\rangle \qquad [8]$$

which, substituted in eqn [1] through eqn [5], finally gives

$$(\Delta \boldsymbol{P}_{\mathrm{el}})_i = \frac{\mathrm{i}q_{\mathrm{e}}}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}\int_0^1 \mathrm{d}\lambda\Big[\langle \partial_{\lambda}\Phi_{\boldsymbol{k}}^{(\lambda)}|\partial_{\boldsymbol{k}_i}\Phi_{\boldsymbol{k}}^{(\lambda)}\rangle - \langle \partial_{\boldsymbol{k}_i}\Phi_{\boldsymbol{k}}^{(\lambda)}|\partial_{\lambda}\Phi_{\boldsymbol{k}}^{(\lambda)}\rangle\Big] \qquad [9]$$

Equation [9], where the $\boldsymbol{k}$-integration is over all the twisted boundary conditions, gives the net polarization change along the path parametrized by λ. Through a further manipulation, one can obtain the following expression, which only depends on the end points $\lambda = 0$ and $\lambda = 1$,

$$\Delta \boldsymbol{P}_{\mathrm{el}} = \boldsymbol{P}_{\mathrm{el}}^{(1)} - \boldsymbol{P}_{\mathrm{el}}^{(0)} \qquad [10]$$

with

$$\boldsymbol{P}_{\mathrm{el}}^{(\lambda)} = \frac{\mathrm{i}q_{\mathrm{e}}}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}\ \langle \Phi_{\boldsymbol{k}}^{(\lambda)}|\partial_{\boldsymbol{k}}\Phi_{\boldsymbol{k}}^{(\lambda)}\rangle \qquad [11]$$

The last three equations give the many-body generalization, as given in 1994 by Ortiz and Martin, of the polarization theory originally formulated in 1993 by King-Smith and Vanderbilt for independent electrons. In the *tour-de-force* leading to eqn [11], it is assumed that the ground-state is isolated from the excited states by a finite energy gap, and that there are no long-range correlations. Besides, it should be pointed out that while eqn [9] gives the exact change in polarization along the chosen path, eqns [10] and [11] only give it modulo a "quantum," whose value can be taken as a multiple of $|q_{\mathrm{e}}|L/V = |q_{\mathrm{e}}|/L^2$. This quantization of charge transport for an insulating system with periodic boundary conditions only occurs in the thermodynamic limit; for finite sizes and periodic boundary conditions, there are exponentially small corrections which destroy the exact quantization. In some way, the difference between eqns [8] and [10] can be regarded as a trade-off for getting an expression which only depends on the end points $\lambda = 0$ and $\lambda = 1$.

Resta, in 1998, gave an alternative formulation of the modern polarization theory which makes use of the unitary many-body phase operator

$$\hat{U}(\boldsymbol{k}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} \qquad [12]$$

where $\boldsymbol{k}$ is an arbitrary vector and $\boldsymbol{R}$ is the usual position operator. At variance with $\boldsymbol{R}$, $\hat{U}(\boldsymbol{k})$ is a perfectly admissible operator for Hilbert spaces defined by periodic boundary conditions, provided $\boldsymbol{k}$ components are restricted to integer multiples of $2\pi/L$, with L the imposed periodicity. One may introduce the special notations $\boldsymbol{\kappa}^{(\alpha)}$ for three useful vectors $\boldsymbol{k}$ in this class (which are chosen to be proportional to the three Cartesian unit vectors)

$$\boldsymbol{\kappa}_{\beta}^{(\alpha)} = \frac{2\pi}{L}\delta_{\alpha\beta} \qquad [13]$$

and for the three related ground-state expectation values:

$$z_N^{(\alpha)} = \langle \Psi_0|\hat{U}(\boldsymbol{\kappa}^{(\alpha)})|\Psi_0\rangle \qquad [14]$$

From these expectation values one can get, in the usual way, their phases. Explicitly,

$$z_N^{(\alpha)} = |z_N^{(\alpha)}|\mathrm{e}^{\mathrm{i}\gamma_N^{(\alpha)}} \qquad [15]$$

$$\gamma_N^{(\alpha)} = \mathrm{Im}\log z_N^{(\alpha)} \qquad [16]$$

Resta showed that the above phases are, in the thermodynamic limit, just the components of macroscopic polarization. One has

$$\boldsymbol{P}_{\alpha} = \frac{q_{\mathrm{e}}\gamma_N^{(\alpha)}}{2\pi L^2} = \frac{q_{\mathrm{e}}}{2\pi L^2}\mathrm{Im}\log z_N^{(\alpha)} \qquad [17]$$

It is worth remembering at this point that $\hat{U}(\boldsymbol{k})$ is not a one-body operator written in a many-body notation: since the integrations in eqn [14] require the full many-body wave function, a one-body density matrix would not be enough. It may be surprising that polarization, a quantity intuitively connected to the average of the electric dipole over the charge density, is not expressible in terms of the one-body density matrix. Again, here one has a trade-off between the elimination of size effects (obtained through the use of periodic boundary conditions) and the fact that macroscopic polarization, in extended systems under periodic boundary conditions, is no longer a function of the electron density. This is why there is a need to extract the information about macroscopic polarization from the phase of the many-body wave function instead of getting it from the charge density.

Another feature worth mentioning is that in eqn [17] there is no $\boldsymbol{k}$-integration, at variance with eqn [11]. Expressions like eqn [17] are usually called "single-point" formulas.

Berry Phase Theory of Localization

As has been already noted, W Kohn was the first to emphasize that even within quantum mechanics,

electron localization must be the cause for insulating behavior. His solution, the disconnectedness of the ground-state many-body wave function, remained unchallenged for ~ 35 years, and still must be regarded as an important part of a correct understanding of the problem. In 1999, however, Resta and Sorella (RS) reconsidered the problem and provided a solution which, besides being in many respects simpler than the previous one, was revolutionary in the sense that it allowed one to understand macroscopic polarization and electron localization as two aspects of the same phenomenon.

To summarize this development, it is convenient to refer to the formal tool that has been introduced in the preceding pages, namely the many-body phase operator. Its ground-state expectation value is a complex number which vanishes in metals and is finite in insulators. It has already been noted that its phase is a measure of macroscopic polarization. RS, in addition, showed that a measure of the localization length in 1D insulators can be given by the expression

$$\xi = \frac{L\sqrt{D}}{2\pi N} \qquad [18]$$

where D is a dimensionless parameter given by

$$D = -\lim_{N\to\infty} N \log|z_N|^2 \qquad [19]$$

RS showed analytically that, for a one-dimensional independent electron model of an insulator, ξ is finite, whereas for metals it diverges (even before taking the limit $N\to\infty$). They also argued that the same should hold for interacting electrons, though the divergent behavior for metals should occur only after taking the limit.

Since the same parameter, $z_N^{(\alpha)}$ as given by eqn [14], gives a measure, through eqn [17], of macroscopic polarization and, through eqns [18]–[19], of electron localization, it is apparent that polarization and localization are just two sides of the same coin. Maybe it is worth stressing again that $z_N^{(\alpha)}$ is a pure ground-state property, whose phase gives, in suitable units, a measure of the (α) component of the polarization, and whose modulus leads to a definition of localization length quite able to make a sharp distinction, in the thermodynamic limit, between metals and insulators.

To complete the picture, it must be pointed out that the following year Souza, Wilkens, and Martin (SWM) extended the above analysis to the case of many dimensions and of correlated electron systems. They also proved that the new definition of localization length corresponds to a well-defined ground-state physical observable, and suggested experimental measurements that could actually be performed. In this process, they were also able to connect the localization length to DC conductivity, that is, to the physical property the vanishing of which signals the onset of an insulating phase.

SWM introduced a statistical approach based on the definition of cumulants as the series-expansion coefficients of $\log C(\alpha)$, the characteristic function being

$$\mathbf{C}(\alpha) = \int \mathrm{e}^{-\mathrm{i}\alpha \cdot \mathbf{R}} p(\boldsymbol{R})\, \mathrm{d}\boldsymbol{R} \qquad [20]$$

where $\boldsymbol{R}$ are the electron coordinates and $p(\boldsymbol{R})$ is the probability distribution. Through this very powerful technique, they were able to recover the Berry phase definition of the electronic polarization as the first cumulant $\langle r_i \rangle_c$, where "c" stands for "cumulant," as well as the square of localization length as the second cumulant $\langle r_i r_i \rangle_c = \langle (r_i)^2 \rangle - \langle r_i \rangle^2$. A detailed illustration of such a generalization is not given here; however, the link with RS formalism is pointed out:

$$\langle r_i r_i \rangle_c = -\frac{1}{N}\left(\frac{L}{2\pi}\right)^2 \log|z_N^{(i)}|^2 \qquad [21]$$

valid in the thermodynamic limit. Moreover, SWM obtained a fundamental connection to the electrical conductivity

$$\langle r_i r_j \rangle_c = \frac{\hbar L^3}{\pi q_e^2 N} \int_0^\infty \frac{\mathrm{d}\omega}{\omega} \mathrm{Re}\ \sigma_{ij}(\omega) \qquad [22]$$

namely a fluctuation–dissipation relation, which allows one to estimate the localization length, that is, a ground-state property, by a measurement over the excited states, such as conductivity. Clearly in insulators, where $\lim_{\omega\to 0} \mathrm{Re}\ \sigma_{ii}(\omega) = 0$, the localization length has a finite asymptotic value, as expected, at variance with conductors where $\lim_{\omega\to 0} \mathrm{Re}\ \sigma_{ii}(\omega) > 0$ and, therefore, the localization length turns out infinite. Another interesting feature can be obtained by observing that, in insulators, there is no absorption below the optical gap ε_g, and therefore,

$$\langle r_i r_i \rangle_c < \frac{\hbar^2 L^3}{\pi q_e^2 N \varepsilon_g} \int_0^\infty \mathrm{d}\omega\ \mathrm{Re}\ \sigma_{ii}(\omega) = \frac{\hbar^2}{2 m_e \varepsilon_g} \qquad [23]$$

where the sum rule for the oscillator strengths has been taken into account. This is obviously the result usually stated in the band theory, that the larger the gap the stronger the insulating character of a crystal.

Conclusions

The modern theory of the electronic states in insulators heavily rests upon the same first-principle density-functional approach which works for metals and semiconductors. Most electronic properties of insulators are well described by this approach, and even the very distinction between insulators and semiconductors can be related to their bandgap values. However, basic ground-state properties such as localization and polarization remained in an unclear state until 1964, when a fundamental paper by W Kohn showed that electron localization depends on the disconnectedness of the ground-state wave function. Substantial advances came only many years later, first through the modern theory of polarization, based on a Berry's phase, then from Resta–Sorella's discovery that the use of the many-body phase operator eqn [12] allows one to view macroscopic polarization and electron localization as two aspects of the same phenomenon. The fact that this operator extracts, in a very compact and effective way, the connectedness properties of the many-body wave function, can be regarded as a direct connection between Kohn's insight and the modern approach.

Future developments in this field appear to depend on advances in the density-functional theory (DFT), for instance, by settling the basic problem that, though in DFT, by definition, the density of the fictitious noninteracting KS system equals the density of the interacting one, the KS system and interacting one can have, in principle, different polarizations. The current understanding of this problem is that, at least in weakly correlated systems, $z_N^{(\alpha)}$ for the noninteracting KS system is a good approximation for the $z_N^{(\alpha)}$ to the interacting system. This assumption seems to be well supported, up to now, from the available experimental evidence.

A further open issue is whether the theory can be extended to discriminate between normal metals and superconductors, an extension which would probably amount to introducing a suitably defined new kind of "anomalous" density. Finally, there is the need for more computational studies to check the predictions of the theory and its extension to disordered systems.

See also: Density-Functional Theory; Geometrical Phase and Polarization in Solids; Insulators, Optical Properties of.

PACS: 03.65.Ca; 03.65.Vf; 71.15.Mb; 71.23.An; 72.80.Sk; 77.22.Ej

Further Reading

Kohn W (1964) Theory of the insulating state. *Physical Review* 133: A171–A181.

Resta R (2002) Why are insulators insulating and metals conducting? *Journal of Physics: Condensed Matter* 14: R625–R656.

Resta R and Sorella S (1999) Electron localization in the insulating state. *Physical Review Letters* 82: 370–373.

Souza I, Iñiguez J, and Vanderbilt D (2002) First-principles approach to insulators in finite electric fields. *Physical Review Letters* 89: 117602.

Souza I, Wilkens T, and Martin RM (2000) Polarization and localization in insulators: generating function approach. *Physical Review B* 62: 1666–1683.

Insulators, Impurity and Defect States in

A M Stoneham and A L Shluger, University College London, London, UK

Introduction

Serious studies of materials are often serious studies of defects, for control of properties of materials implies control of defects or impurities. Understanding the electronic structure of defects in insulators has helped the development of microelectronic devices, the photographic process, solid electrolytes, and the applied surface physics of sensors and catalyst substrates. This understanding is based on ideas founded on a synthesis of theories, modeling, and experiment.

Defining the Issues

When a new material is created, which defects are important? Will there be vacancy-interstitial pairs (Schottky defects) or just vacancies of several types (Frenkel defects)? Will there be nonstandard ionic charge states, such as Fe^{3+} ions in FeO? Could one species move on to the sites of another, for example, As antisites on Ga sites in GaAs? Can the material be doped with electrically active impurities? Are there understandable trends of properties from material to material in some class, such as II–VI compounds? Electronic structure is subtle and varied, with the key ideas identified by a mix of theory and experiment.

In most insulators Coulomb energies dominate, so antisite defects are not energetically easy, and charge

transfer across sublattices is difficult, instead nonstoichiometry is common. Thus vacant lattice sites (vacancies), ions placed at normally unoccupied sites (interstitial ions), foreign ions present as impurity or dopant, and ions with charges different from those of the host ions are prevalent defect types. Electronic defects may arise either as ions with charges deviating from those on lattice ions, or from excitation of electrons from filled valence bands to empty conduction bands. When a valence-band electron is missing, one has an electron hole (hole) and its behavior can be compared with that of an electron in a normally empty conduction band. The electron and hole together can form a transient neutral species, the "exciton." In the absence of macroscopic electric fields, ionic lattices must be electrically neutral overall, implying that one charged defect must be compensated by another charged defect or defects to satisfy the electroneutrality condition locally or over the whole sample. The charges of defects and of the regular lattice particles, defined with respect to the neutral ideal lattice, are called "effective charges."

Complexity is the norm for advanced materials. When solid-state lasers degrade and mechanical properties change under irradiation, defect processes occur in parallel on timescales from picoseconds upwards. In a composite material, defect phenomena may occur in the matrix, the fiber, or the interface. When a metal is oxidized, interfacial, diffusional, and impurity processes coexist. Separating these processes by experiment alone may not be feasible. The proper use of electronic structure theory establishes priorities, and is used to organize massive amounts of data to identify what is novel and point to significant issues.

Which Electronic Structure Parameters Are of Interest?

Defects control the properties and performance of materials and devices in various ways. For luminescence, whether of rare earths in oxides or sulfides or CsI:Na, the spectroscopic properties of impurities, optical absorption, and luminescence energies are of the outmost importance. Bragg grating formation in silica fibers is associated with the optical absorption of induced defects. Defect species in gate dielectrics, such as those listed in **Table 1**, influence device performance. These defects may initiate degradation, trap charge, or create fixed charge, which scatters carriers in the Si substrate. Their electron- and hole-trapping energies and cross sections are of main interest. Many radiation-induced processes in insulators involve forming self-trapped excitons and holes, so the electron coupling to lattice deformation plays a crucial role.

The raw data of defect studies describe spectroscopic observations and their dependence on time and other parameters, such as dopant concentration, temperature, pressure, irradiation power, and dose. Carefully obtained data provide the basis for formulating defect models. **Figure 1** shows several such models for major paramagnetic defects in irradiated fused silica. The most reliable models knit together the best-available experimental data from many sources with solid-state physics theory, often aided by detailed calculations. Some of the key ideas underpinning modern calculations of the electronic structure of defects in insulators are reviewed here. The article starts with outlining some of the theoretical methods used to study defect electronic structures. Then it reviews some of the concepts necessary for understanding the electronic properties of point defects in crystalline and disordered solids and at interfaces.

Theoretical Methods for Defect Electronic Structure

Even simple answers can be useful answers. At one level, electrons and holes can be treated only implicitly. Born's model of ionic solids as polarizable point ions, interacting through transferable interatomic

Table 1 Possible defect types and their impact on device performance

Origin of defect	*Defect type*	*Possible role*
Imprecise stoichiometry	Vacancies or interstitials, annealing-induced	Fixed charge, SILC; motion under applied field
Growth-related impurities from fabrication or diffusion	Precursor related: e.g., Cl, H	Fixed charge; time dependence of operating device
	Source/drain related: e.g., B, P	
Features of polycrystallinity	Grain boundaries and associated defects, e.g., vacancies	Charge, diffusion
Amorphous structure	Structure-induced traps	Leakage current
Interface	Wrong bonds, relaxation with associated dipoles, roughness	Dipole, charge, electric field

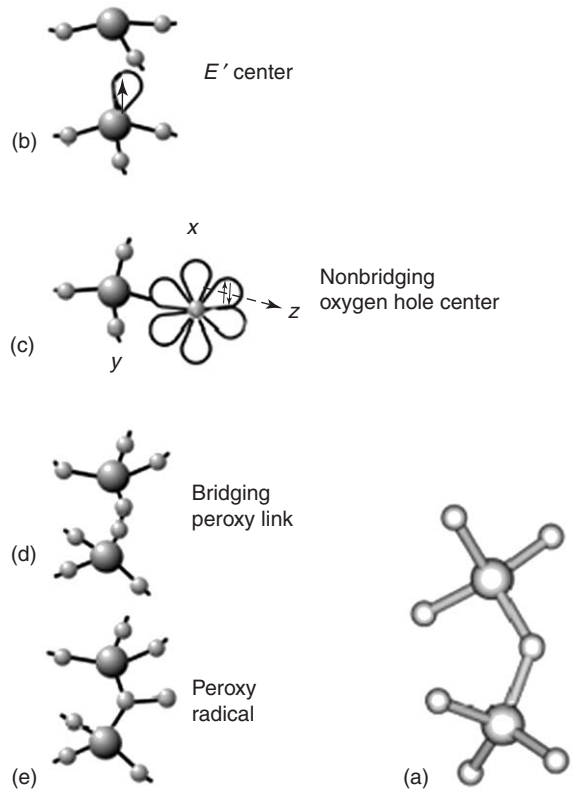


Figure 1 Schematics of defect models developed for amorphous silica. (a) A fragment of perfect SiO_2 lattice showing two tetrahedral SiO_4 units connected via a common oxygen ion; smaller balls are oxygen ions. (b) An oxygen vacancy with an unpaired electron on a dangling bond of a three-coordinated Si atom. (c) A nonbridging oxygen atom trapping a hole; the partially filled *p*-orbitals of the oxygen atom are shown. (d) A bridging peroxy linkage center formed by incorporation of an oxygen atom at a regular Si–O–Si bond. (e) A peroxy radical center formed by attachment of an oxygen ion to a regular oxygen site.

potentials, was the basis for describing cohesion, dielectric, and vibrational properties of perfect solids. The large Coulomb energies are the reason why accurate defect studies were first practical for ionic crystals. Nevertheless, polarization energies are large too (for an ion of charge $Z+2$ substituting for a Mg^{2+} ion in MgO, Madelung terms are $\sim 24 \times Z\,\text{eV}$ and polarization $\sim 6 \times Z^2\,\text{eV}$) and explain why so many charge states of impurities can be stable. Mott and Littleton's systematic way of calculating the Coulomb and polarization contributions to defect formation energies in ionic solids forms the basis of many current state-of-the-art codes.

Modeling Simple One-Electron Systems

For nonpolar semiconductors, "effective mass theory" is superb, often working well beyond its justification using empirical bulk properties. For ionic crystals, point ions are a surprisingly good start, but "pseudopotential methods" were necessary to explain why, and to give the "ion-size corrections" needed for accurate trends over the whole series of alkali halides. The F center (electron at an anion vacancy) was the first deep trap to be predicted well over a very wide range of halide and oxide hosts, and the same techniques were successful in modeling states of self-trapped excitons.

For most systems, however, the "many-electron treatment" is essential. Here both first-principles and empirical theories are often used. Empirical methods help most in new situations where what is important is not certain, for more complex systems, where computer power and technique is yet to be sufficient for the best methods, or for extrapolation across a broad range of materials. Many of the problems addressed empirically today will be solved by far better methods tomorrow.

By choosing first-principles methods, one is not reliant on parametrized theories or fitting possibly inaccurate experiments. These theories can be applied more reliably to hypothetical materials, defects for which data are not available, or for extreme temperature, pressure, and irradiation conditions. Computational techniques for modeling defects in solids exist in a form of computer codes, implementing a model of a solid with a defect, and methods for calculation of defect structure and properties.

Most first-principles methods are based on Hartree–Fock and density-functional theory (DFT). DFT states that the ground-state properties of a system are given by the charge density, and state-of-the-art methods allows one to compute the charge density and energy self-consistently, using various effective exchange-correlation potentials that account for the quantum mechanical interactions between electrons. The local density approximation (LDA) takes the exchange-correlation potential from the uniform electron gas theory at the density for each point in the material. The generalized gradient approximation (GGA) includes the effects of local gradients in the density.

The main computer models of a point defect in a crystal include a periodic model, an embedded cluster, and molecular cluster models. They differ by the boundary conditions imposed on a system wave function. The periodic model was first developed for calculating the electronic structure and properties of ideal crystals, whereas the cluster model evolved from molecular calculations (**Table 2**). Within the periodic model, a unit cell with a defect is periodically translated. It is ideal if one is interested in the ground-state properties of neutral defects, which weakly perturb the surrounding lattice. Many molecular cluster models treat a defect with surrounding

Table 2 Comparison of computational methods used in defect calculations

Models	*Methods*	*Properties*
Molecular cluster: molecule represents local defect environment in a solid	Density-functional theory and *ab initio* and semi-empirical Hartree–Fock theory in localized basis sets	Can predict local defect structure, vibrational, optical, and EPR properties providing intermediate and long-range order are not important.
Embedded cluster: local defect environment is treated quantum-mechanically and the rest of the solid structure treated classically	Density-functional theory and *ab initio* and semi-empirical Hartree–Fock theory in localized basis sets	Can predict local defect structure, ionization energies, electron and hole affinities, vibrational, optical, and EPR properties, and diffusion parameters. Can treat infinite amorphous structures.
Periodic: defects are repeated periodically in a lattice	Density-functional theory, Hartree–Fock theory, tight-binding methods in plane wave and localized basis sets	Can predict local defect structure, electrical levels, vibrational and EPR properties, and diffusion parameters providing the interaction between periodically translated defects is small. Treats amorphous structures as periodic arrangement of disordered cells.

atoms simply as a molecule or cluster. They disregard the effects of the rest of the atoms outside the cluster, but apply semi-empirical fixes to the atoms at the cluster surface to diminish their adverse effect. This major problem is addressed and partially solved in "embedded cluster" models, where a cluster with a defect is embedded in an infinite solid. This allows one to self-consistently include the effect of the polarizable environment into the cluster treatment of a defect, which is particularly important for charged defects in insulators.

Both static Hartree–Fock and DFT methods are ground-state theories, in the form of an effective one-particle Schrödinger equation (the Kohn–Sham equation in the case of DFT). The eigenvalues of these equations are often interpreted as electron energies and their differences as optical excitation energies. This interpretation is linked to an intuitive picture of independent electrons occupying well-defined energy levels. However, this is a gross simplification: the electrons are not independent, and when, for example, an electron is removed to form a hole, all remaining electrons respond by changing their electronic states. This electron relaxation and other quantum-mechanical contributions must be taken into account for energy differences to be calculated correctly. One can retain the picture of one-particle energy levels, but these particles are quasi-electrons and quasi-holes, which embody the effects of interaction with all the other particles.

The concept of quasiparticles is often used to represent the electronic states of an insulator with a defect schematically, as shown in **Figure 2**. Such energy

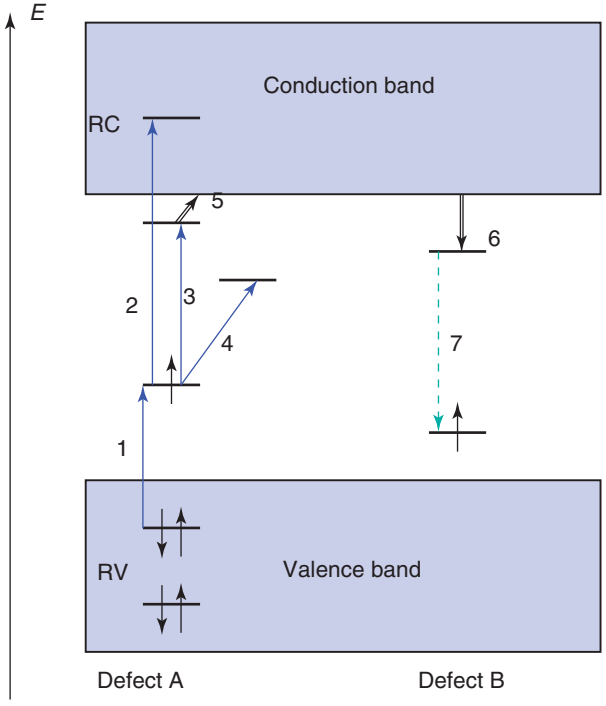


Figure 2 The quasiparticle energy diagram representing total energies of a solid with two defects A and B. The defect A produces local single-occupied and unoccupied states within the bandgap, and resonant states (denoted RV and RC) in the valence and conduction band, respectively. The defect B is characterized only by two local states in the gap. The arrows 1–4 correspond to optical transitions between states in the defect A. The arrow 5 corresponds to thermal ionization of an electron from the excited state of the defect A into the conduction band. Process 6 is nonradiative trapping of an electron from the conduction band on an empty state of the defect B. Process 7 is the de-excitation of the defect B, which may be either radiative or nonradiative.

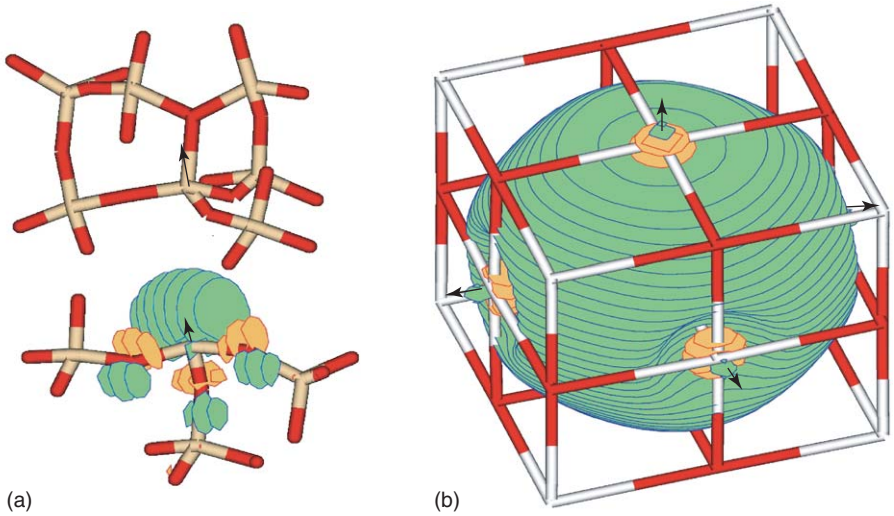


Figure 3 The geometric configurations and electronic states of positively charged oxygen vacancies in oxides. (a) The relaxed configuration of the positively charged oxygen vacancy, E_1' center, in α-quartz; red are oxygen ions. The envelope of the density of the unpaired electron mostly localized on a *p*-orbital of one Si ion is shown pointing into the vacancy. The arrows indicate the direction and relative values of the displacements of the two Si ions surrounding the vacancy. (b) The relaxed configuration of the positively charged oxygen vacancy, F^+ center, in MgO; red are oxygen ions. The envelope of the density of the unpaired electron localized on an *s*-orbital inside the vacancy is shown. Note that the electron density of the unpaired electron propagates much further out from the vacancy. Depressions in the envelope show that the unpaired electron avoids the core electrons of surrounding Mg ions. The arrows indicate the symmetric displacements of the six nearest-neighbor Mg ions surrounding the vacancy. This radial displacement of the six Mg ions is often used as a configuration coordinate, Q, in considering the spectroscopic properties of the F^+ center.

diagrams usually include the highest valence band and the lowest conduction band of the system. Point defects and impurities may induce occupied and unoccupied states in the gap and so-called resonant states in the valence and conduction bands of the system. The wave functions of localized states are often strongly localized in the defect area (**Figure 3**) and the corresponding quasi-electron energies are located in the bandgap of the solid. The resonant states differ from the ordinary band states in that they are localized in the defect region with the major wave function component decaying exponentially from the defect, but their quasi-electron energies are located within the occupied or unoccupied bands. Quasiparticle energy diagrams similar to that shown in **Figure 2** are used to summarize pictorially the electronic structure and processes involving optical excitation, luminescence, ionization, and energy transfer between defects.

Coupling of Electrons and Lattice Deformation

One of the issues crucial for the understanding of defect properties is how defect electrons couple to the motion of nearby ions or, alternatively, the way defect electrons apply forces to host ions. The static polarization and distortion of the host lattice affects defect formation and transition energies; the dynamics leads to transitions between electronic states and to defect diffusive motion. Optical spectra change qualitatively: the optical line width, the existence of a zero-phonon line, and the Stokes shift, the energy difference between optical absorption and emission energies, give direct measures of this coupling. Although any of a large number of atomic displacements might influence a defect in a solid, often a very specific combination of displacements is especially important. For example, the radial displacement of the nearest neighbors to the defect can be represented by a single configuration coordinate Q, as illustrated in **Figure 3**. It is noted that it is most unlikely that the coordinate Q will be one of the dynamically independent normal vibrational modes q. It is almost certain that normal modes of many different frequencies will be needed to make up Q. Moreover, the precise mode combinations making up Q will depend on the defect electronic state (it will be different in the excited state of defects shown in **Figure 3**). Yet, this configuration coordinate model is surprisingly effective, giving an invaluable tool for showing what happens after electronic excitation, and a simple language for describing defect processes. An example of such a diagram is given in **Figure 4**.

A central issue for optical and electronic materials is how they recover after excitation. A good X-ray phosphor emits the most visible light after excitation and degrades least from defect production. A

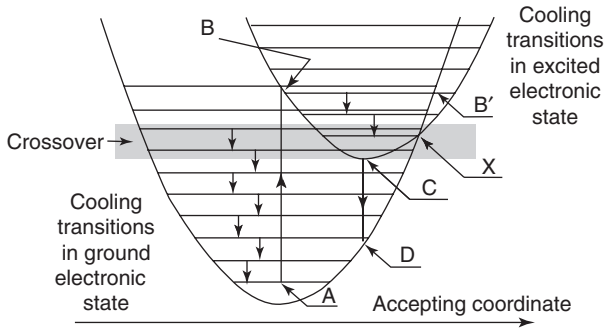


Figure 4 Configuration coordinate diagram. The figure shows the main characteristic energies and transitions, both radiative and nonradiative. Accepting coordinate represents one of the configuration coordinates effective in vibrational energy dissipation. The vertical arrow A–B corresponds to a vertical (so-called Franck–Condon) transition between the ground and excited states of the system represented by two harmonic energy surfaces. Horizontal lines indicate vibrational energy levels with small arrows showing transitions between these levels (cooling transitions) dissipating the excess vibrational energy (relaxation energy). The transition C–D between the lowest vibrational state in the excited state back to the ground-state energy surface corresponds to the Franck–Condon deexcitation with photon emission (luminescence or phosphorescence). It is accompanied by further cooling transitions in ground electronic state. The area near the crossing point X is where the system can cross nonradiatively from the excited to the ground state. In this case the relaxation energy, which is the energy difference between X and A or between C and A, is dissipated into lattice vibrations. The zero-phonon transition energy is the energy difference between C and A.

solid-state laser needs both population inversion and a lack of wasteful competing processes; carrier capture in semiconductors determines what defects and materials are acceptable; photographic (not to mention the photosynthetic) processes have their own demands. In such cases, nonradiative transitions are central, a role evident from over half a century of theory.

In nonradiative decay, excitation energy is transformed into heat, defect production, or to excitation of other electrons, the so-called Auger processes which are especially important at higher carrier densities. The most common manifestation of nonradiative decay of an excited state concerns suppression of luminescence. Can one tell, from optical absorption data, whether luminescence will occur from the absorption data alone or not? Huang and Rhys's seminal study showed one possible criterion: the larger the excitation energy, the faster the optical transition and the slower the nonradiative decay, since larger numbers of phonons would be needed (see **Figure 4**). Dexter, Klick, and Russell developed other criteria relating to the final state energy achieved by optical excitation and the energy at which the ground and excited electronic states intersected on the configuration coordinate diagram: the energy surfaces, rather than some matrix element, controlled the result. This led Bartram and Stoneham to show that simple expressions could be given for branching ratios at critical steps, and that these ideas – which could be applied to much more complex cases – correctly predicted behavior over a wide range of systems.

Self-Trapping and Localization

Landau made the remarkable suggestion of self-trapping: an electron could be so strongly coupled to the lattice that, even without a defect, it would lead to lattice distortion and be effectively immobilized. The suggestion was the basis of the important ideas of small polarons (incoherent hopping of localized carriers, as in some transition metal oxides, rather than the long mean-free-path motion of carriers in Si or metals), and even of the recently observed barrier to self-trapping. The observed small polarons (e.g., self-trapped holes in halides in **Figure 5**) show both the soundness of Landau's idea and (since naive belief in the band theory appears to be violated) the importance of theorists learning from experiment too. Distortion and polarization dominate too in negative-U behavior, where a charge disproportionation is exothermic (so the positive vacancies in Si decay from $2\,V^{+}$ to $V^{0}+V^{2+}$; for all free atoms, disproportionation is endothermic).

Localization can correspond to carrier or exciton trapping by an isolated defect (extrinsic trapping). If there is an attractive perturbation A at a defect site, then, for large enough A, a bound state emerges for the extra electron, hole, or exciton. Both impurity-induced potentials and electron–phonon interaction play crucial roles in this localization. Extrinsic electron trapping near defects is quite rare in insulators because electron polarons are not easily formed in these materials. However, exciton trapping by impurities is very common in most insulators.

As well-established examples of extrinsic hole trapping, one can consider holes trapped near cation vacancies in MgO (so-called V^{o} and V^{-} centers) and holes trapped near Ge and Al impurities in α-quartz. In particular, cation vacancies in MgO have an effective negative charge of $-2e$ with respect to the lattice and, therefore, attract positive holes. When one hole is trapped, a paramagnetic V^{-} center is formed, which is well studied experimentally using electron paramagnetic resonance and optical spectroscopy. Similar centers occur in CaO, SrO, and BaO, in the fourfold-coordinated oxides such as BeO and ZnO and in related sulfides, selenides, and tellurides. In MgO, at low temperatures the hole is

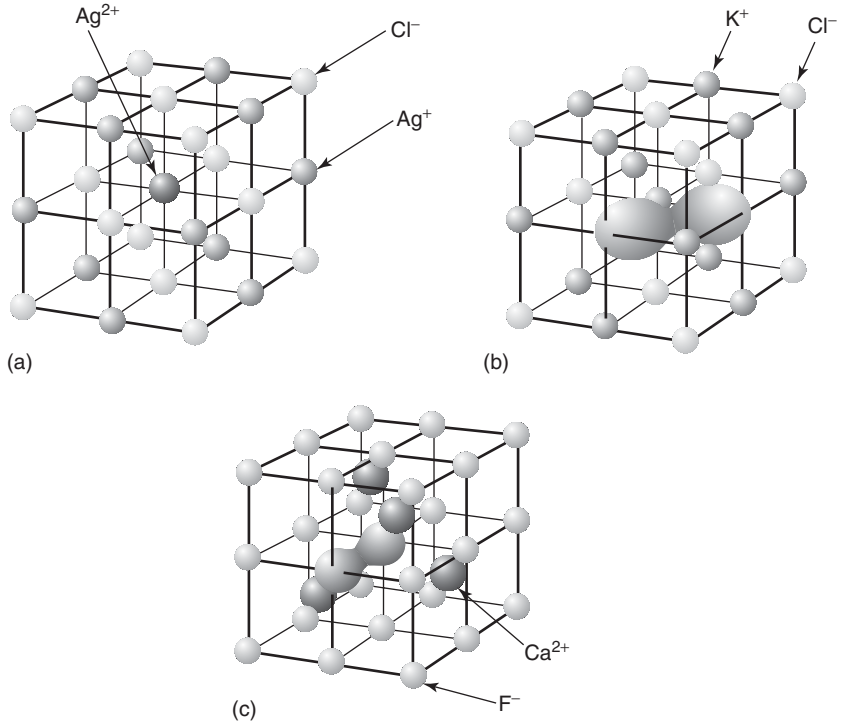


Figure 5 Atomic structure of the self-trapped hole centers in halides: (a) AgCl – the hole is localized on a silver ion; (b) alkali halides of NaCl structure – the hole is localized on two halogen ions forming a Cl_2^- molecule; (c) CaF_2 – the hole is localized on two fluorine ions forming a F_2^- molecule. Note that in (b)–(c) the X_2^- (X = F, Cl) molecular ion occupies two nearest anion lattice sites. In all cases, the hole localization is accompanied by a strong distortion of the surrounding lattice (not shown in the figure).

strongly localized on one of the six oxygen ions surrounding the vacancy. As the temperature rises, the hole moves more and more rapidly among the equivalent oxygens until it appears delocalized at room temperature. The nature of optical transitions in this center is characteristic to many other hole centers in oxides: there are two types of optical transitions. One is internal to the O^- ion on which the hole is localized. This is a crystal-field transition, since the transition energy is largely determined by the electric field at the O^- ion due to the Mg^{2+} vacancy. The other, more important, transitions involve charge transfer of the hole to other neighbors of the vacancy. These are analogous to electron excitation from resonant states in the valence band into the unoccupied hole state shown in **Figure 2**. Similar charge transfers are important in determining the color of gemstones.

Extensive experimental and theoretical studies show that excitons self-trap in both α-quartz (the most stable crystalline polymorph of SiO_2) and amorphous SiO_2. However, EPR studies show that holes and electrons do not self-trap in α-quartz. Yet, in high-purity Ge-doped quartz, low-temperature irradiation leads to the formation of a hole center localized at an oxygen site next to Ge. This formation of the localized hole centers associated with Ge impurity can be described as "extrinsic self-trapping," where both impurity-induced potential and electron–phonon interaction play crucial roles in the hole localization. In amorphous SiO_2, two types of localized hole centers have been determined that are free from pre-existing lattice defects such as vacancies, interstitials, or impurities, and are trapped by some precursor states induced by the disorder of amorphous structure. This example demonstrates that the mechanisms of localization processes in amorphous solids are more complicated than in crystalline materials due to coexistence of electron–phonon interaction, structural defects, and static potential fluctuations. It is not straightforward to distinguish the polaron self-trapping experimentally due to strong electron–phonon interaction from charge trapping by structural defects in an amorphous structure.

Defects in Disordered Materials

Defect structures and processes in amorphous materials have been studied extensively, partly as fundamental research on model amorphous materials, and partly to understand their extremely important role in technology. For example, the mechanisms and

kinetics of relaxation of electronic excitations in SiO_2 are crucial in micro- and opto-electronics applications. There are at least three basic fallacies concerning amorphous systems. First, it is wrong to assume there is only a single amorphous structure for a given composition, since the method of preparation can have significant effects. Second, it is wrong to think that the mean energies of defect formation or of trap ionization in an amorphous structure are sufficient to understand the behavior: the form and tails of the distribution are equally important. Third, one should not assume that crystals and amorphous systems of the same composition have the same values for defect and trap energies.

Structural disorder in amorphous materials generally means that all sites are different. In most experimental studies, the effect is hidden in broadening of the spectral features, and only rarely has disorder been implicated in the formation of different structural types of the same defect. Predicting defect properties and relative abundance of different defect configurations generally requires considering not one or several, as in crystals, but a statistical ensemble of structural sites. Additional factors include the sample history and the mechanism of defect formation. This provides new challenges for both theoretical and experimental analyses of defects in amorphous materials.

Surfaces and Interfaces

No discussion of defects could be complete without mention of interfaces. Their structures are rarely, if ever, the ideal terminations of perfect crystal structures, and the variety of reconstructions is becoming clearer as new microscopies (e.g., the atomic force microscope) develop. Defects or impurities at surfaces and interfaces are especially important for most applications. It is the impurities close to grain boundaries, which are most strongly involved in the operation of solid-state gas sensors, and the framework of the Debye–Hückel theory is central. The space charge in silver halides relates to their photographic effectiveness; the Mott and Gurney theory provides the basic framework. Defects at the interface between silicon and gate dielectrics control the reliability of microelectronic devices. Ceramists recognize the importance of segregation in oxides, where experiment shows that Mg in alumina gives dramatic improvements in sintering.

Defects also influence adhesion. What controls the variation from one oxide to another of the adhesion of a nonreactive metal? Why (a related question) does liquid Cu wet some oxides (NiO, urania, chromia) but not other similar oxides (MgO, thoria, alumina)? The answer lies largely in the defect populations and basic electrostatics: charged defects close to the highly polarizable metal have lower energy; those close to the unpolarizable vacuum have higher energy. The operation of gas sensors relies on electron transfer across interfaces, as does the operation of contacts to semiconductors (whether conventional or organic) and electrodes.

Yet, another issue, specific for microelectronics applications, is that the dielectric of interest can be in contact with the electrode(s) – semiconducting (e.g., silicon) or metallic – which provides a source of extra electrons. Also, a dielectric film can be placed in the electric field between the two electrodes, which may alter the defect charge states. Charged defects can cause substantial random fields, which may alter breakdown thresholds locally. Typically, for 100 ppm of charged defects, there will be random fields of order $10^5\,V\,cm^{-1}$. Defects can be involved in stress-induced leakage currents (SILC) and dielectric breakdown. In such cases, the defect offers a channel for electron transfer across the dielectric: an electron tunnels from one electrode into an oxide defect; there are then relaxation processes or a sequence of defect processes, and the electron can then tunnel to the other electrode. What can one say about the defects responsible, their electrical energy levels and relaxation energies (differences in energy for electron capture and emission)? Can one give useful estimates of electron and hole trapping cross sections? Will there be any field-stimulated diffusion of defects?

To define which charge state of an intrinsic defect (e.g., vacancy) or incorporated defect species is favored, consider the common case of an oxide film grown on silicon, the source or sink of electrons being the bottom of the silicon conduction band at the Si/SiO_2 interface. Hence, the chemical potential of the electron can be chosen at the bottom of the Si conduction band or at the silicon mid-gap energy. Electrons are assumed to be able to tunnel elastically or inelastically from/to these states to/from defect states in the oxide and create charged species. Most calculations use experimental information to estimate the energy of an electron at the bottom of the conduction band of Si with respect to the defect states.

The key point about defect states in a dielectric from the device point of view is to control the dependence of the defect charge state on the position of the system Fermi energy. It is useful to distinguish two charge-state levels: the thermodynamic level, E_{th}, and the transient level, E_{tr}. The thermodynamic charge-state level corresponds to the system Fermi energy for which a defect changes its charge in

thermal equilibrium. E_{th} is defined as a difference of total energies of systems of N and $N+1$ electrons, each in its fully relaxed ground state: $E_{th} = E(N+1) - E(N)$. Thus, if E_F is the energy required to remove an electron from some electron reservoir (e.g., Si at the Si/SiO_2 interface), $E_{th} - E_F$ is the energy required to add an electron from the reservoir to the lowest unoccupied charge-state level.

Taking into account that the electrical measurements and the electronic processes are fast compared with the time needed for the system to come to equilibrium, another transient level is introduced, which is consistent with the Frank–Condon principle. E_{tr} is defined as the energy required for trapping an electron or a hole in the geometry of the system as it was prior to the trapping. Thus, E_{tr} are obtained as the differences of the total energies of the system with N and $N+1$ electron, but in the local atomic geometries, which correspond to one of these systems. For example, $E_{tr}(0/-) = E(-) - E(0)$ denotes the energy required to add an electron to the neutral defect calculated using the neutral defect geometry.

Once the charge-state levels are calculated, the dependence of the defect charge state on the electrode Fermi energy can be introduced via a correction term $q \times E_F$, where E_F is varied in the energy range from that of the maximum of the last valence band of the oxide to the bottom of the first conduction band. The results are presented as plots of E_{th} versus E_F or E_{tr} versus E_F calculated for different values of E_F. The crossing points of the corresponding lines determine the regions where one charge state of a defect is more favorable than the other. Although the Fermi energy remains unknown in the experiment, such diagrams allow one to identify some critical cases and help in the analysis of the experimental data. The same approach is used to determine the dependence of the defect formation energies on the position of the Fermi energy.

See also: Ceramic Materials; Density-Functional Theory; Electronic Structure Calculations: Plane-Wave Methods; Electrons and Holes; Excitons: Theory; Glasses; Hartree and Hartree–Fock Methods in Electronic Structure; Insulators, Electronic States of; Insulators, Optical Properties of; Lattice Dynamics: Vibrational Modes; Luminescence; Optical Absorption and Reflectance; Point Defects; Polarons; Pseudopotential Method; Semi-Empirical and Empirical Calculations; Surfaces and Interfaces, Electronic Structure of.

PACS: 61.72.Ji; 61.80.Az; 71.35.Aa; 71.55.Ht; 71.55.Jv; 73.20.Hb; 73.20.Mf; 78.20.Bh; 78.55.Fv; 78.55.Hx

Further Reading

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. New York: Pergamon.

Dexter D, Klick CC, and Russell GA (1956) Criterion for the occurrence of luminescence. *Physical Review* 100: 603.

Hayes W and Stoneham AM (1985) Theory of light absorption and non-radiative transitions in F-centres. *Defects and Defect Processes in Non-Metallic Solids*. New York: Wiley.

Huang K and Rhys A (1950) *Proceedings of the Royal Society* A204: 406.

Itoh N and Stoneham AM (2001) *Materials Modification by Electronic Excitation*. Cambridge: Cambridge University Press.

Lannoo M and Bourgoin J (1981) *Point Defects in Semiconductors, I Theoretical Aspects*. Berlin: Springer.

Mott NF and Davis EA (1979) *Non-Crystalline Materials*. Oxford: Oxford University Press.

Song KS and Williams RT (1993) *Self-Trapped Excitons*. Berlin: Springer.

Stoneham AM (1975) *Theory of Defects in Solids*. Oxford: Oxford University Press.

Toyozawa Y (2003) *Optical Processes in Solids*. Cambridge: Cambridge University Press.

Insulators, Optical Properties of

Y Toyozawa, University of Tokyo, Tokyo, Japan

Band Model and Optical Properties

Insulators are characterized, and contradistinguished from metals, by the existence of a forbidden energy band of finite width, a bandgap, between the highest filled band (the valence band) and the lowest empty band (conduction band). The existence of a bandgap can be confirmed by the optical transparency, that is, a vanishing absorption constant, up to a finite energy a little below the bandgap ε_g, except in the lower energy region where the absorption due to lattice vibrations may arise.

The energy difference between the first absorption peak and the bandgap ε_g, from where the continuous spectrum of band-to-band transitions starts, is due to the Coulomb binding energy between the excited electron in the conduction band and the positive hole left behind in the valence band, a typical example of "configuration interaction" – a step beyond the

one-electron approximation of the band model – or of "final state interaction" in the optical spectrum. This bound electron–hole pair is called "exciton," which means a quantum of excitation.

Exciton Model and the Optical Spectrum Near the Absorption Edge

The relative motion of the electron and the positive hole with orbital radius larger than the interatomic distance can be described in the effective mass approximation, and if their effective masses m_e and m_h and hence the reduced mass $m_r = (m_e^{-1} + m_h^{-1})^{-1}$ are isotropic, the binding energies form a hydrogen-like series: $B_n = Ry_{ex}/n^2 (n = 1, 2, \ldots)$, $Ry_{ex} = (m_r/m)(\varepsilon_0/\varepsilon_e)^2 Ry$ where $Ry = 13.6\,\text{eV}$ is the Rydberg energy, m the electron mass in vacuum, ε_0 the dielectric constant of vacuum, and ε_e the electronic dielectric constant $\varepsilon(\omega)$ at $\hbar\omega \sim O(Ry_{ex}) \ll \varepsilon_g$ which is contributed only from electronic transitions except when B_n is as small as the phonon energies. The exciton radius is given by $a_n = a_{ex} n^2$, $a_{ex} = (m/m_r)(\varepsilon_e/\varepsilon_0) a_H$ where $a_H = 0.052$ nm is the orbital radius of the electron in the ground state of a hydrogen atom. This Wannier model of an exciton is valid for insulators with relatively small bandgap (less than a few eV) and hence with large values of $\varepsilon_e/\varepsilon_0$ and a_{ex}/a_H of the order of 10. In insulators with a larger bandgap, the exciton radius is smaller than or comparable with the interatomic (intermolecular) distance, and the Frenkel model of intra-atomic (intramolecular) excitation (as in rare gas solids and aromatic molecular crystals), or the model of charge transfer to neighboring atoms (from an anion to neighboring cations as in ionic crystals) is more appropriate at least for an exciton in its ground state ($n = 1$).

In the Wannier model, the matrix element of the optical transition is given as a product of that for the interband transition and the amplitude of the hydrogen-like wave function $F_n(\boldsymbol{R}_m)$ of the relative motion at the origin $\boldsymbol{R}_m = 0$ (probability amplitude that the electron and the hole are found on the same site), which is nonvanishing only for the s-like states ($\ell = 0$ where ℓ is the angular momentum). Thus, the intensity of the ns exciton line at $\hbar\omega = E_n = \varepsilon_g - B_n$ in the absorption spectrum is proportional to the square of the above amplitude, namely to n^{-3} ($n = 1, 2, \ldots$) as known in the hydrogen atom. Noting that the product of the intensity and the number density of states $|dE_n/dn|^{-1} \propto n^3$ is independent of n and hence of energy $\hbar\omega$, one finds that the absorption spectra consisting of higher exciton lines overlapping with each other (due to their finite widths) tends to a plateau of constant height as $n \to \infty$. It can also be shown analytically that this asymptotic height is in exact accordance with that of the continuous spectrum of band-to-band transition with the Coulomb-enhancement factor $(\hbar\omega - \varepsilon_g)^{-1/2}$ which cancels out the density of states effect $(\hbar\omega - \varepsilon_g)^{+1/2}$. The resulting absorption spectrum around the band edge is shown schematically in **Figure 1a** where the n-dependence of the width γ_n of discrete lines does not affect the above mentioned continuity.

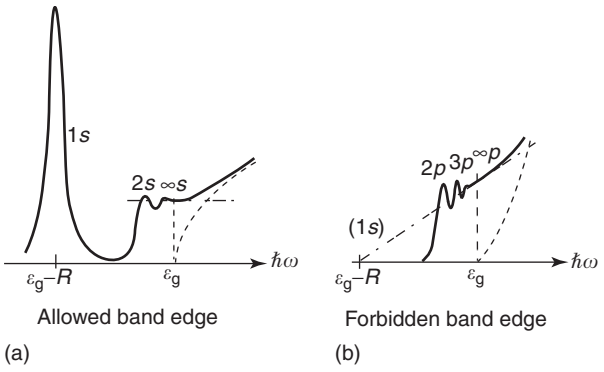


Figure 1 The effect of Coulomb interaction on the fundamental absorption edge. The broken lines represent the interband absorption edge without the Coulomb effect. (Reproduced with permission from Toyozawa Y (2003) *Optical Processes in Solids.* Cambridge: Cambridge University Press; © Cambridge University Press.)

In the case that the matrix element of band-to-band transition vanishes at the band edge ε_g due to a reason of symmetries of band wave functions, one has to consider its values above the edge. Correspondingly, one has to use the gradient $\nabla_R F_n(\boldsymbol{R}_m)$ of the hydrogen wave function at the origin, which is nonvanishing only for the p-like states ($\ell = 1$). Thus, np exciton lines ($n = 2, 3, \ldots$) appear at $\hbar\omega = E_n = \varepsilon_g - B_n$ of the absorption spectrum with intensities proportional to $(n^{-3} - n^{-5})$, with $n = 1$ missing. It can be shown, as before, that the broadened exciton lines asymptotically merge smoothly (inclusive of the finite slope extrapolating to the $n = 1$ position) to the continuous spectra of the Coulomb enhanced band-to-band transition around ε_g. The absorption spectra in this case of the forbidden band edge is schematically shown in **Figure 1b**. The intensity of the absorption spectra is smaller than that in the allowed band edge case mentioned above by a factor $(a/a_{ex})^2$, where a is the interatomic distance. Absorption spectra in alkali halides and cuprous oxides shown in **Figures 2a** and **2b** belong to the cases (a) and (b), respectively, of **Figure 1**.

Insulators with small bandgaps become intrinsic semiconductors except at low temperatures. The existence of thermally excited free carriers causes a metallic screening of a long-range electron–hole

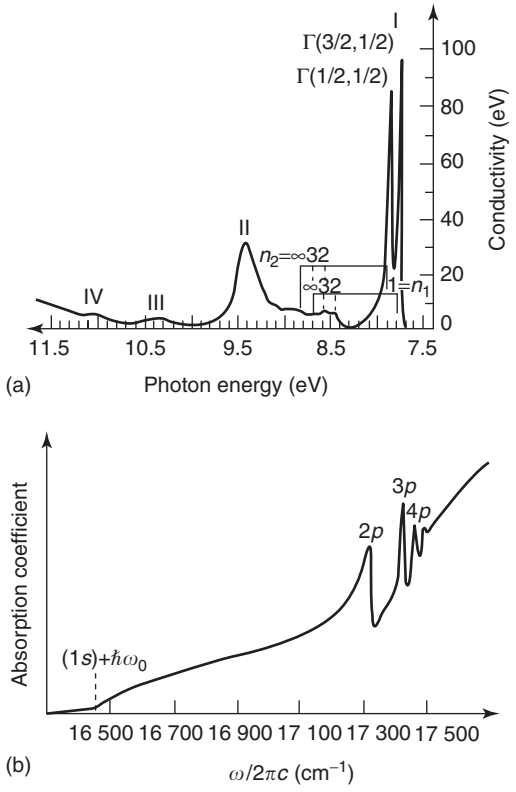


Figure 2 (a) Fundamental absorption spectrum of KCl crystal 10 K. (Reproduced with permission from Tomiki T (1969) *Journal of the Physical Society of Japan* 26: 738.) (b) The yellow series excitons at the fundamental absorption edge of Cu_2O crystal at 4.2 K. (Reproduced from Nikitine S, Grun JB, and Sieskind MJ (1960) *Journal of Physics and Chemistry of Solids* 17: 296, with permission from Elsevier.)

Coulomb interaction into a short-range one, thus affecting the interband optical spectra in such a way that the discrete exciton lines merge into the continuous spectra of interband transition as temperature rises. Moreover, the free carriers give rise to their own optical absorption (intraband optical transition) in the infrared region.

Dipole–Dipole Interaction Contributing to the Translational Motion of an Exciton

As a composite particle consisting of an electron in the conduction band with effective mass m_e and a positive hole in the valence band with effective mass m_h, the exciton has effective mass $m_{ex} = m_e + m_h$ and translational kinetic energy $\hbar^2 K^2/2m_{ex}$ near the bottom of the exciton band. Finite bandwidths of the valence, conduction, and hence exciton bands are due to the overlaps of wave functions between neighboring atoms.

The width of the exciton band is contributed by another type of interaction energy, the dipole–dipole interaction. Take the Frenkel exciton in which this interaction predominates, and consider how the intra-atomic excitation at the nth atom is transferred to the mth site. Consider the state $\Phi_n = \phi_p(\boldsymbol{r}_n - \boldsymbol{R}_n)\prod_{\ell(\neq n)}\phi_s(\boldsymbol{r}_\ell - \boldsymbol{R}_\ell)$ where all atoms are in the s-like ground state except the nth atom which is in a p-like excited state. The excitation transfers to the mth atom through the matrix element

$$H_{nm} \equiv (\Phi_n, H\Phi_m) = \int\int \mathrm{d}\boldsymbol{r}_n \mathrm{d}\boldsymbol{r}_m \phi_p(\boldsymbol{r}_n - \boldsymbol{R}_n)\phi_s(\boldsymbol{r}_m - \boldsymbol{R}_m) \times v_{nm}\phi_s(\boldsymbol{r}_n - \boldsymbol{R}_n)\phi_p(\boldsymbol{r}_m - \boldsymbol{R}_m)$$

where H denotes the total Hamiltonian of the system consisting of N atoms with one electron per atom. Note that due to the orthogonality of atomic wave functions ϕ_s and ϕ_p and the neglect of overlaps of wave functions between different atoms, only the Coulomb interaction $v_{nm} = e^2/4\pi\varepsilon_0|r_n - r_m|$ remains among the total Hamiltonian H. Due to the same reasons, one can take only the lowest dipole–dipole terms in the multipole expansion of the long-range Coulomb interaction, namely $H_{nm} = (\boldsymbol{m}^2/R_{nm}^3 - 3(\boldsymbol{m}\cdot\boldsymbol{R}_{nm})^2/R_{nm}^5)/\ 4\pi\varepsilon_0 \equiv D\ (\boldsymbol{R}_{nm})$, where $\boldsymbol{m} \equiv \int \mathrm{d}\boldsymbol{r}\phi_p\ (\boldsymbol{r})(-e\boldsymbol{r})\phi_s(\boldsymbol{r})$ is the transition dipole moment. The Hamiltonian H is diagonalized with the wave function of the Frenkel exciton given by $\Psi_K^{(e)} = N^{-1/2}\sum_n \exp(\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n)\Phi_n$, namely $H_{KK'} = \delta_{KK'}E_K^{(e)}$, with diagonal energy being $E_K^{(e)} = (\Psi_K^{(e)},\ H\Psi_K^{(e)}) = E^{(g)} + \varepsilon + D_K$ where $E^{(g)}$ is the ground-state energy of the entire system, $\varepsilon = \varepsilon_p - \varepsilon_s$ the atomic excitation energy in the crystal and $D_K \equiv \sum_{n(\neq 0)} D(\boldsymbol{R}_n)\exp(-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n)$. Thus, the dipole–dipole interaction contributes to the K-dependence of the exciton energy.

One must be careful in summing up the long-range dipole–dipole interaction $D(\boldsymbol{R}_n) \propto R_n^{-3}$ because of the slow convergence, especially when K is much smaller than the reciprocal of the interatomic distance a. Dividing the summation into the inner region $R_n < R_c$ and the outer region $R_n > R_c$ where R_c is chosen so as $a \ll R_c \ll K^{-1}$, and replacing the latter summation by integration, one obtains $D_K \equiv \sum_{n(\neq 0)}^{R_n<R_c} D(\boldsymbol{R}_n) - (N_0/3\varepsilon_0)[m^2 - 3(\boldsymbol{m}\cdot\boldsymbol{K})^2/K^2]$, which is singular at $\boldsymbol{K} = 0$ since it tends to different values depending on the direction in which $\boldsymbol{K}$ approaches 0. In a typical case that each atom is situated in the environment of cubic symmetry, the first summation vanishes. One has triply degenerate atomic p-states p_x, p_y and p_z so that one has a 3×3 matrix of the $\boldsymbol{K}$-dependent part: $(D'_K)_{ij} = -(N_0\mu^2/3\varepsilon_0)[\delta_{ij} - 3K_iK_j/K^2]$. It is diagonalized by taking the coordinate system with the

z-axis along the direction of $\boldsymbol{K}$. In fact, one obtains a longitudinal wave $\boldsymbol{m}\|\boldsymbol{K}$ with energy $(D'_K)_{33} = (2/3)(N_0\mu^2/\varepsilon_0)$, and two transverse waves $\boldsymbol{m}\perp\boldsymbol{K}$ with energy $(D'_K)_{11} = (D'_K)_{22} = -(1/3)(N_0\mu^2/\varepsilon_0)$. Their energy difference, $N_0\mu^2/\varepsilon_0$, agrees with $\hbar$ times the longitudinal–transverse splitting $D_{\mathrm{lt}} = \omega_1 - \omega_{\mathrm{t}}$ of polarization waves which originates from the depolarizing field. Under the coexistence of many oscillators, ε_0, is to be replaced by the residual dielectric constant ε'_j which is contributed by the oscillators other than the one (j) under consideration. Note that this dielectric constant is different from ε_{e}, which governed the electron–hole relative motion.

The Effect of Spin and the Interplay of Exchange and Spin–Orbit Interactions

To be more realistic, one must consider the electron spin in addition to its orbital motion. Each atomic state characterized by (n, ℓ, m) is doubly degenerate with up and down spins, $s = \pm 1/2$. The simplest closed shell atom consists of two ns electrons with configuration $(ns)^2$, and its optically allowed excited state consists of the configuration $(ns)(np)$. The two electrons in the excited state can have either antiparallel (as in the ground state) or parallel spins, with energy difference $2\Delta_{\mathrm{exch}}(>0)$ where $\Delta_{\mathrm{exch}} = \int\int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}'\phi_p(\boldsymbol{r})\phi_s(\boldsymbol{r})(\mathrm{e}^2/4\pi\varepsilon_0)|\boldsymbol{r}-\boldsymbol{r}'|^{-1}\phi_s(\boldsymbol{r}')\phi_p(\boldsymbol{r}')$ is the exchange energy. The singlet state is above the triplet state because one is concerned with the e–h (instead of the e–e) system. The optical transition to the triplet excited state is forbidden because of its total spin state $S = 1$ $(\boldsymbol{S} \equiv \boldsymbol{s}_1 + \boldsymbol{s}_2)$ different from that of the ground state $S = 0$. Noting that Δ_{exch} is the special case $(n = m)$ of H_{nm} in the preceding section and that the dipole–dipole interactions vanish for the triplet state but are to be multiplied by 2 for the singlet state (because two electrons participate), one can incorporate the exchange and dipolar energy into a single form, $2\delta_{S,0}|F_n(0)|^2[\Delta_{\mathrm{exch}} + D_k]$, where multiplication by the squared amplitude of the wave function of e–h relative motion at the origin is required only where the interaction remains in the Wannier exciton. The total energy of the Wannier exciton with the principal quantum number n, spin quantum number S, and translational wave vector $\boldsymbol{K}$ is then given by

$$E_{n,S}(\boldsymbol{K}) = \varepsilon_{\mathrm{g}} - B_n + \hbar^2\boldsymbol{K}^2/2m_{\mathrm{ex}} + 2\delta_{S,0}|F_n(0)|^2[\Delta_{\mathrm{exch}} + D_K]$$

Another term to be considered is the spin–orbit interaction $H_{\mathrm{s.o.}} = \lambda\boldsymbol{l}\cdot\boldsymbol{s}$. One can diagonalize it together with the total angular momentum $\boldsymbol{j} \equiv \boldsymbol{l} + \boldsymbol{s}$ by making use of the identity $2\boldsymbol{l}\cdot\boldsymbol{s} = (\boldsymbol{j}^2 - \boldsymbol{l}^2 - \boldsymbol{s}^2) = j(j+1) - 1\cdot 2 - (1/2)\cdot(3/2)$. Thus, the atomic state with $\ell = 1$ are split into $j = 3/2$ and $j = 1/2$ with spin–orbit energies $+\lambda/3$ and $-2\lambda/3$, respectively. According to this quasi-atomic model, the valence band of alkali halide which consists of p-orbitals of halogen is split to the upper $(j_{\mathrm{v}} = 3/2)$ and lower $(j_{\mathrm{v}} = 1/2)$ branches while the conduction band consisting of s-orbital of alkali is not split ($j_{\mathrm{c}} = 1/2$ only). Hence, the absorption edge of band-to-band transition appears at $\varepsilon_{\mathrm{g}} - \lambda/3$ and $\varepsilon_{\mathrm{g}} + 2\lambda/3$ with the intensity ratio 2:1 reflecting the degeneracy $2j_{\mathrm{v}} + 1$. This $\boldsymbol{j}$–$\boldsymbol{j}$ coupling scheme is no more valid for exciton lines because of finite exchange-dipolar interaction, $\Delta_n \equiv 2\delta_{S,0}|F_n(0)|^2[\Delta_{\mathrm{exch}} + D_0]$ mentioned above. As Δ_n/λ increases the intensity ratio decreases, tending to 0 since the lower exciton state tends to a pure spin-triplet ($S = 1$) which is optically forbidden ($\boldsymbol{L}$–$\boldsymbol{S}$ coupling). In materials with small spin–orbit coupling constants as are realized with light atom constituents, the triplet exciton is optically almost forbidden. The continuous change from the $\boldsymbol{j}$–$\boldsymbol{j}$ to $\boldsymbol{L}$–$\boldsymbol{S}$ and again to $\boldsymbol{j}$–$\boldsymbol{j}$ coupling schemes has been observed in the alloy system $CuCl_{1-x}Br_x$ in which λ varies continuously from negative to positive values with Δ_{exch} staying always small as composition x varies from 0 to 1, as is shown in **Figure 3**. The $\boldsymbol{L}$–$\boldsymbol{S}$ coupling scheme is realized only around the composition x where λ changes sign, the $\boldsymbol{j}$–$\boldsymbol{j}$ coupling scheme with intensity ratio 1:2 and 2:1 is observed on its left and right sides, respectively.

The quasi-atomic model mentioned above has been improved to the crystalline field model in which the symmetry of the crystal is fully taken into account. The anisotropic exchange energy introduced thereby as an independent material parameter has in fact been observed in some materials.

Participation of Phonons in the Electronic Excitation

The wave vector of a photon with energy enough for the electronic excitation across a bandgap of 10 eV is $\sim 10^8\,\mathrm{m}^{-1}$ which is much smaller than the reciprocal lattice $\sim 10^{10}\,\mathrm{m}^{-1}$. One can thus assume that the wave vector (pseudo or crystal momentum divided by $\hbar$) is conserved in the optical transition, namely $\boldsymbol{k}_{\mathrm{c}} - \boldsymbol{k}_{\mathrm{v}} = 0$ or $\boldsymbol{K}_{\mathrm{ex}} = 0$. In addition, it has been tacitly assumed so far that the bottom of the conduction band and the top of the valence band are at the origin, $\boldsymbol{k}^m_{\mathrm{c}} = 0$ and $\boldsymbol{k}^m_{\mathrm{v}} = 0$, which, however, is not always the case. If one or both of them are not at the origin, the absorption edge of band-to band transition appears around the indirect

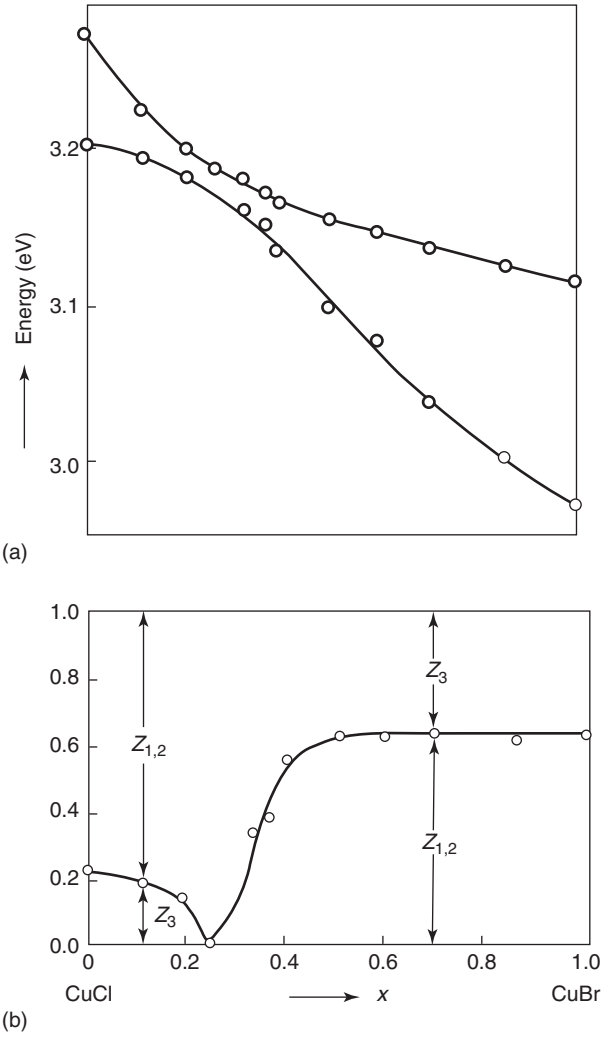


Figure 3 (a) Observed energies, and (b) intensity ratio of $Z_{1,2}$ and Z_3 excitons in mixed crystal: $CuCl_{1-x}Br_x$ as functions of the concentration x. (Reproduced with permission from Kato Y, Yu CI, and Goto T (1970) *Journal of the Physical Society of Japan* 28: 104.)

gap $\varepsilon_g = \varepsilon_c(\boldsymbol{k}_c^m) - \varepsilon_v(\boldsymbol{k}_v^m)$, but with the participation of phonons with wave vector $\boldsymbol{k}_{id} \equiv \boldsymbol{k}_c^m - \boldsymbol{k}_v^m$ to satisfy the conservation of the wave vector. Hence, the exact absorption edge appears at $\varepsilon_g - B_n' \pm \hbar\omega_s(\pm\boldsymbol{k}_{id})$ according as a phonon is absorbed or emitted, where B_1' is the binding energy of the exciton formed at $\boldsymbol{K} = \boldsymbol{k}_{id}$ and $\omega_s(\boldsymbol{k})$ is the circular frequency of the sth mode lattice vibration with wave vector $\boldsymbol{k}$. The absorption spectrum rises with the square root of energy from each edge reflecting the density of states of the exciton band, with the intensity proportional to the square of the indirect exciton–phonon interaction matrix element and the temperature dependent factor $\bar{n}_s$ and $\bar{n}_s + 1$ for absorption and emission of a phonon, respectively, where $\bar{n}_s = [\exp\{\hbar\omega_s(\pm\boldsymbol{k}_{id})/\ k_B T\} - 1]^{-1}$ is the thermal average of the number of phonons. Such indirect exciton edges are observed in germanium and silicon ($\boldsymbol{k}_c^m \neq 0$), silver halides ($\boldsymbol{k}_v^m \neq 0$), and thallous halides ($0 \neq \boldsymbol{k}_c^m \neq \boldsymbol{k}_v^m \neq 0$). Similar phonon-mediated exciton edge arises even when $\boldsymbol{k}_c^m = \boldsymbol{k}_v^m = 0$, provided the exciton transition matrix element vanishes at $\boldsymbol{K}_{ex} = 0$, as is realized with the 1s exciton in Cu_2O (see **Figure 2b**).

The widths γ of direct exciton absorption peaks, introduced phenomenologically as damping constants of oscillators, are microscopically attributed to the rate τ^{-1} of scattering of the optically created exciton $\boldsymbol{K}_{ex} = 0$ to $\boldsymbol{K}_{ex} \neq 0$ by absorbing or emitting a wave number conserving phonon, which is of the same origin as the indirect exciton transition mentioned above. An indirect exciton absorption band is, therefore, supposed to be the tail part of the broadened peaks of a direct exciton. As a second-order process, the $\boldsymbol{K} = 0$ exciton is the intermediate state which is resonant and nonresonant to the incident photon in the direct and indirect transitions, respectively.

In fact, one can incorporate both aspects in a unified expression for a multicomponent exciton absorption spectrum. For a multiband scheme of excitons with energy $E_{\lambda K}$, one can write the spectral line shape as the sum of asymmetric Lorentzians: $F(E) = \sum_\lambda \bar{f}_\lambda \pi^{-1}[\bar{\Gamma}_{\lambda 0} + \bar{A}_\lambda(E - \bar{E}_{\lambda 0})]/[(E - \bar{E}_{\lambda 0})^2 + \bar{\Gamma}_{\lambda 0}^2]$ which is in accordance with the result of the classical model of damped oscillators ($\hbar\gamma_\lambda \leftrightarrow \bar{\Gamma}_{\lambda 0}$) for the direct exciton peaks, except for the asymmetry terms with $\bar{A}_\lambda$ (the degree of asymmetry) in the numerator, and the fact that the quantities crowned with bars are the renormalized (in the phonon field) quantities dependent on a variable E, the photon energy (though not shown explicitly). The shape of the indirect absorption edge is contained implicitly in the functional form of the width $\bar{\Gamma}_{\lambda 0}(E)$, which is proportional to the combined density of states of excitons and phonons in the final state of indirect transition. The asymmetry term $\bar{A}_\lambda(E - \bar{E}_{\lambda 0})$ in the numerator originates from the interference of the direct exciton peak λ with its background continuum of indirect exciton transitions to other exciton bands $\lambda'(\neq\lambda)$ – a solid-state version of the Fano effect well known in atomic spectroscopy. Such multicomponent asymmetric Lorentzian peaks have, in fact, been observed as shown in **Figure 2b**. The asymmetry here is due to the interference of the direct exciton peak with the background continuum of phonon-assisted optical transition to the 1s exciton band to which the direct transition is forbidden. Phonon sidebands of the direct exciton peak (as are observed in some ionic crystals) are also contained implicitly in the functional form of $\bar{\Gamma}_{\lambda 0}(E)$.

Absorption Bands of Extrinsic Origin below the Bandgap

While (chemically) pure and (physically) perfect insulators are transparent in the spectral region below the first exciton peak (in the direct bandgap case), impure or imperfect insulators have various absorption bands due to the electronic transitions at impurity (guest) atoms or point defects of a lattice. The localized electron(s) is subject to a strong interaction with phonons (stronger than the mobile exciton is) particularly when its orbital radius is comparable with the lattice constant, and once excited, it causes significant displacements of nearby atoms to their new equilibrium positions (lattice relaxation). From this "relaxed excited state," the electron returns to the ground state emitting a photon (luminescence) or nonradiatively. The peak of the emission band is displaced from that of the absorption band (Stokes shift), both bands being more like Gaussian shapes as can be described by the configuration coordinate model. Examples are Tl-like atoms in alkali halides which are known as phosphorescence centers of high efficiency, and the F-centers (an electron captured at an anion vacancy) in alkali halides which are subject to a variety of photochemical reactions resulting in the color changes of the crystal.

The impurities and imperfections sometimes give rise to bound excitons as can be confirmed by the optical absorption bands just below the intrinsic exciton absorption peaks. For example, the α and β absorption bands in alkali halides correspond to exciton formation around an anion vacancy and an F-center, respectively. Shallow bound excitons, as are realized at donors in direct gap semiconductors (such as CdS), have large oscillator strengths of the order of the number of host atoms covered by their extended wave functions, the effect known as "giant oscillator strength."

Intrinsic Photoluminescence and Its Correlation with the Urbach Rule

The optically created exciton will be thermalized around the bottom of the exciton band from where it will emit a photon with energy smaller (in the indirect case) than or equal (in the direct case) to the absorption edge. In the direct case, the exciton is gradually converted into a photon during its thermalization through the polariton bottleneck. In any case, the emission spectrum from this "mobile exciton" is a sharp line.

In some materials, the exciton becomes self-trapped by the lattice distortion caused by itself. The emission spectrum of this self-trapped exciton is a broad Gaussian with a significant Stokes shift from the exciton absorption peak (direct gap case) or edge (indirect case), similar to the localized center mentioned before. In the configuration coordinate model, the self-trapped state is separated by a potential barrier from the free or mobile state, and which of them is more stable depends on whether the exciton–phonon coupling constant g, a material constant, is greater or smaller than a critical value g_c.

In contrast to different behaviors of the emission spectrum in the weak and strong coupling cases, the line shape of the exciton absorption band varies continuously as a function of g. For instance, the low energy tail of the exciton absorption spectrum in many insulators has been found to obey the Urbach rule: $\exp[-\sigma(E_0' - E)/k_B T]$ except at low temperatures where E_0' is the convergence point of semi-logarithmically plotted line contours at various temperatures T, and the steepness coefficient σ is the material constant. This exponential tail has been ascribed to momentarily localized exciton states under a fluctuating field of lattice vibrations, with σ being given by a constant (different values for direct and indirect gap cases) times g^{-1}. While empirical values of σ and g derived therefrom are rather evenly distributed over a variety of insulating materials, the observed emission spectrum of the exciton is distinctly different according as g is smaller ($\rightarrow$ sharp line almost resonant to the absorption peak in the direct gap case) or larger ($\rightarrow$ broadband with a significant Stokes shift) than the critical value g_c in accordance with the theoretical prediction from the observed value of σ.

Optical Excitation of Inner Shell Electrons

Synchrotron radiation has facilitated the spectroscopic studies over a vast energy range extending up to the X-ray, corresponding to electronic transitions from different inner atomic states (much deeper than the outermost orbital state responsible for the valence band) to the empty conduction band. Below the ionization continuum of each inner state, there appear exciton lines as in the valence band excitation. However, the exciton here can be viewed as a localized excitation since the width of the energy band formed by the inner atomic state is negligibly small. Moreover, the positive hole in the inner state is subject to decay processes such as (1) radiative recombination, and (2) Auger recombination with an electron in the upper filled band (inclusive of the valence band) which are the main causes of a lifetime broadening of the exciton line, except in shallow inner bands (next below the valence band) in some

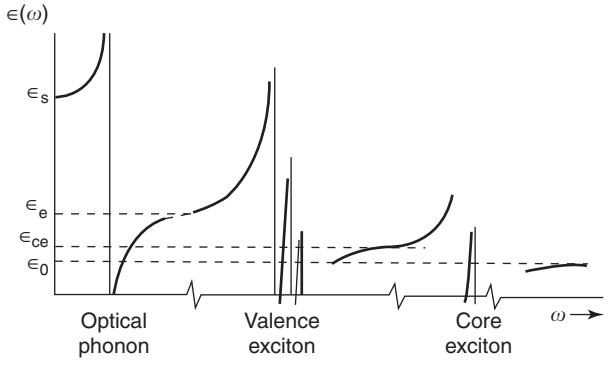


Figure 4 Typical dispersion of dielectric constant $\varepsilon(\omega)$ of an ionic crystal. At frequencies far higher than those of optical phonons but a little below the bandgap, dispersion appears due to excitons associated with the valence band hole. At even higher frequencies, dispersions appear due to excitons associated with the inner shell holes. (Reproduced with permission from Toyozawa Y (2003) *Optical Processes in Solids*. Cambridge: Cambridge University Press; © Cambridge University Press.)

materials in which (1) is insignificant and (2) is impossible because of insufficient energy of electronic recombination against the ionization of valence electrons. Whether the optical spectrum belongs to the band edge allowed or forbidden type, depends on the symmetry of the inner atomic state. The electron–hole exchange energy with its possible interplay with spin–orbit interaction appears here in the same way as before.

Typical dielectric dispersion in the entire ω-region of an ionic crystal is shown schematically in **Figure 4**, in the idealistic situation of no damping (vanishing line width).

See also: Defect Centers in Insulating Solids, Optical Properties of; Effective Masses; Electrons and Holes; Excitons: Theory; Insulators, Impurity and Defect States in; Optical Absorption and Reflectance; Polymers and Organic Compounds, Optical Properties of; Synchrotron Radiation.

PACS: 77.22; 78.20; 78.40; 78.55

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Saunders College.

Broude VL, Rashba EI, and Sheka EF (1985) *Spectroscopy of Molecular Excitons*. Berlin: Springer.

Cho K (1979) Internal structure of excitons. In: Cho K (ed.) *Excitons*, ch.2. Berlin: Springer.

Elliott RJ (1963) Theory of excitons – I (optical spectra of excitons). In: Kuper CG and Whitfield GD (eds.) *Polarons and Excitons*, pp. 269–293. Edinburgh, London: Oliver and Boyd.

Fowler WB (ed.) (1968) *Physics of Color Centers*. New York: Academic Press.

Haken H (1963) Theory of excitons – II (polaron effect of excitons). In: Kuper CG and Whitfield GD (eds.) *Polarons and Excitons*, pp. 295–324. Edinburgh, London: Oliver and Boyd.

Knox RS (1953). Theory of excitons. In: Seitz F and Turnbull (eds.) *Solid State Physics (suppl)*., vol. 5. New York: Academic Press.

Kotani A and Toyozawa Y (1979) Theoretical aspects of inner level spectroscopy. In: Kunz C (ed.) *Synchrotron Radiation, Technique and Applications*. Berlin: Springer.

Rashba EI and Sturge MD (eds.) (1982) *Excitons*. Amsterdam: North-Holland.

Song KS and Williams RT (1993) *Self-Trapped Excitons*. Berlin: Springer.

Toyozawa Y (2003) *Optical Processes in Solids*. Cambridge: Cambridge University Press.

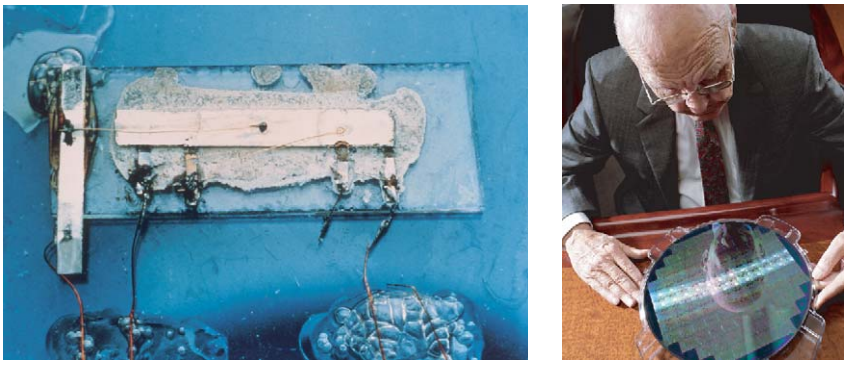
Ueta M, Kanzaki H, Kobayashi K, Toyozawa Y, and Hanamura E (1986) *Excitonic Processes in Solids*. Berlin: Springer.

Integrated Circuits

M Van Rossum, IMEC, Leuven, Belgium

Introduction

The performance of a complex electronic circuit usually scales with the number of its components. In the early days of electronics, circuits were bulky assemblies of vacuum valves (invented by Lee de Forest in 1906), resistors, and capacitors fixed on a metal frame, and connected with soldered copper wires. The invention of the transistor by Bardeen, Brattain, and Shockley in 1947 already contained the seed of further circuit miniaturization, but it would take more than 10 years before this potential was fully recognized. In 1959, Jack Kilby from Texas Instruments and Robert Noyce from Fairchild independently filed patents for an electronic circuit whose components were fabricated on a single piece of semiconductor material (germanium for Kilby and silicon for Noyce). Kilby's circuit measured 0.2×0.8 inches and had 12 components, connected together with tiny gold wires. To quote his own words, "what we didn't realize then was that the integrated circuit would reduce the cost of electronic functions by a factor of a million to one, nothing had ever done that for anything before." Today, integrated circuits (ICs) may contain tens of millions of transistors per square centimeter. Jack Kilby received the Nobel prize for his invention in 2000 (**Figure 1**).



(a) (b)

Figure 1 (a) The first integrated circuit. (b) Jack Kilby contemplating a 300 mm silicon wafer with state-of-the-art ICs. (Courtesy of Texas Instruments.)

IC Materials

The vast majority of ICs are fabricated on single crystal silicon. Although not the best semiconductor material, silicon is still unrivaled thanks to its crystalline perfection, its excellent combination of electrical and thermomechanical properties, as well as the outstanding quality of its natural oxide, SiO_2. For some low-cost applications where data speed is not an issue, polycrystalline silicon films deposited on glass can also be used. The most recent IC technology is silicon-on-insulator (SOI), which consists of a thin monocrystalline silicon film, isolated from the bulk silicon substrate by an intercalated SiO_2 layer.

Over the years, other substrate materials have found a role in niche technologies, targeting applications for which silicon is not the optimal choice. The most important IC technology after silicon is based on III–V compound semiconductors. Prominent among them is gallium arsenide (GaAs), which is a better semiconductor than silicon although less robust from a manufacturing point of view. Due to its high electron mobility and high bulk resistivity, it is the substrate of choice for high-frequency circuits. As a direct bandgap material, gallium arsenide is also an efficient light emitter and therefore well suited for optoelectronic devices, a domain where silicon cannot compete since its own bandgap is indirect. GaAs can be combined with indium, phosphorous, or antimony to tune the band structure for specific applications. Indium phosphide (InP), another III–V compound, is a convenient substrate for many epitaxial films made from ternary or even quaternary III–V alloys. A new player in this area is gallium nitride (GaN), well-suited for power electronics and optoelectronics. GaN is not available as a bulk substrate and must be grown on sapphire or silicon carbide (SiC). Together with SiC and diamond, it belongs to a special class of wide bandgap semiconductors, which are fit for high-voltage device fabrication. However, their present materials quality is lagging far behind silicon and the commercial future of these technologies is still unclear.

IC substrates are usually produced as circular wafers, a few hundred micrometers in thickness, cut from as-grown cylindrical ingots of various diameters. Silicon wafers are commercially available with diameters up to 300 mm, whereas other semiconductor substrates are significantly smaller. Silicon ICs can be classified into two groups based on the type of transistors they contain. Bipolar circuits use bipolar junction transistors as their principle elements. Metal oxide semiconductor ICs contain mainly *n*- and *p*-type metal-oxide semiconductor field-effect transistor (MOSFET) as their active components and are the dominant technology for logic circuits (i.e., processors and memories, also called digital circuits). Today, the most common architecture for digital electronics is complementary MOS (CMOS), which uses a combination of a *p*-MOSFET and *n*-MOSFET as its basic device to minimize the power consumption during switching.

IC Technology

The making of an IC typically proceeds through three standard stages: (1) fabricating individual devices, (2) wiring the devices together to form the circuit, and (3) encapsulating the IC in a package. The full sequence of the fabrication steps required to produce a complete circuit is called the process flow. Different IC technologies therefore require different process flows. In the following, the focus will be on CMOS technology because of its dominant position in the IC market.

Device Fabrication

Starting from a clean substrate, the first stage of IC fabrication – commonly called front-end process – is to form individual devices. The devices are produced through several levels of patterning, each level consisting of a well-defined combination of photolithography, deposition and etching steps. For a MOSFET, the typical front-end sequence is the following:

1. doping and isolation of the transistor wells and junctions
2. growth of the gate insulator
3. patterning of the gate electrode
4. formation of the metal contacts

In total, a CMOS process may require several hundred individual steps. The most important ones are outlined below.

The transistor well The well is the region of the substrate containing the transistor body. Since current must pass through the device, the well must be doped to provide charge carriers (electrons or holes). After the device area has been delineated by photolithography and isolated from its surroundings by thin grooves (so-called shallow trenches), etched into the substrate and filled with a dielectric material, dopants are introduced into the well region. For *n*-type MOSFETs, the well must be *p*-doped whereas *p*-MOSFETs need *n*-doped wells. In the early days of MOS technology, diffusion of impurities from a gaseous or solid-state source was the standard doping technique. However, with present deep-submicron device dimensions, this method has been replaced by ion implantation, which offers far better flexibility, precision and depth control. Ions of boron (for *p*-type), phosphorous or arsenic (for *n*-type) are accelerated by an electric field and gain a kinetic energy of a few tens of kiloelectronvolts, allowing them to penetrate some hundred nanometers into the substrate. Several implantation steps are usually required to reach the desired dopant profile. The as-implanted region has high electrical resistance due to the radiation-induced lattice damage, which can range from isolated point defects to complete amorphization. The high defect density prevents the dopants from occupying substitutional lattice sites, thereby leaving them electrically inactive. In order to activate the dopants, damage must be removed by a high-temperature annealing step (typically between 800 and 1000°C). This can be done in a conventional furnace, but a short thermal pulse produced by halogen lamps (the so-called rapid thermal process or RTP) is often preferred in order to minimize the motion of the implanted species during annealing.

The gate insulator The thinnest feature of the transistor is the gate insulator, which consists of a few nm of silicon dioxide (SiO_2) deposited directly on top of the transistor well. The gate oxide is formed by thermal oxidation of the silicon surface, which involves heating the substrate above 900°C in an oxygen atmosphere (dry oxidation) or a mixture of oxygen and water (wet oxidation). Trace amounts of other gases are often added to improve the electrical quality of the oxide. Both the temperature gradients and the gas flow pattern in the furnace must be accurately controlled to obtain a very thin and uniform oxide. Short thermal pulses produced by rapid thermal processing can also be used. Although the as-grown oxide layer is structurally amorphous, the high quality of its chemical bonds with the silicon surface limits the density of interface defects to very low levels (less than $10^{10}\,\mathrm{cm}^{-2}$). The quality of the gate oxide and of the Si/SiO_2 interface is a crucial factor to ensure the good electrical behavior of the MOSFET. However, there are now indications that further shrinking of the devices will push SiO_2 toward its application limit. For very thin oxides, direct carrier tunneling through the gate insulator produces a sizable current leakage and unwanted power dissipation. This problem can, in principle, be overcome by introducing "high-*k* insulators," that is, gate dielectrics with a higher polarizability than SiO_2. These materials would allow the physical thickness of the insulator to increase (thereby keeping tunneling currents at an acceptable level), while maintaining a strong capacitive coupling between the gate and the channel. At this time, refractory metal oxides such as HfO_2 as well as some ferroelectrics are potential candidates for replacing silicon dioxide in future ICs.

Gate, source, and drain The gate electrode sits on top of the oxide and regulates the current inside the transistor channel. It is made of doped polycrystalline silicon (usually designated as "poly"), which is a reasonably good conductor and can be patterned into very narrow lines. Today, the most advanced commercial transistors have a physical gate length of $\sim 50\,\mathrm{nm}$. Since the gate is the narrowest feature on any IC, its formation involves the most demanding steps of the front-end process flow. First, a blanket layer of poly is deposited by chemical vapor deposition (CVD), after which the gate fingers are defined by a lithography step and patterned by dry etching. In order to achieve the right threshold voltage for the transistor, the poly-gate on the NMOS is *n*-doped whereas on PMOS it is *p*-doped. Both doping steps

are performed by ion implantation and annealing. In a MOSFET, the current input and output electrodes are called source and drain. They are defined by local implantation of complementary doping species (*n* for *p*-well and *p* for *n*-well) very close to the surface, thereby forming shallow *p–n* or *n–p* junctions, depending on the transistor type. Fine-tuning of the junction profiles requires additional implantation steps followed by annealing.

Contacts Metals are needed for electrical contacts to the three device electrodes. The contact material must exhibit low electrical resistance and be chemically compatible with silicon in order to avoid interface degradation over time. For many years, metal silicides have been used extensively on the source and the drain: first titanium disilicide ($TiSi_2$), and more recently cobalt disilicide ($CoSi_2$) and nickel monosilicide (NiSi). The silicide layers are formed by solid-state reaction of a deposited metal film with the underlying silicon; therefore, it is important that the reaction should not consume too much silicon. In the same way, a polycrystalline silicide layer or polycide is formed on top of the poly in order to contact the gate electrode and to reduce the gate series resistance (**Figures 2** and **3**).

Interconnects

Connecting the individual transistors is an essential step in the production of an IC. The process flow needed for interconnect fabrication is called the back-end sequence. Since the early days of IC technology, aluminum-based alloys have been used as the interconnect metal. This choice was dictated by the low resistivity of aluminum, its high ductility, easy etching behavior, and good coverage of nonplanar geometries. The latter property is particularly important, since complex ICs normally require many levels of wiring stacked on top of each other, giving rise to a multilayer structure called the multilevel interconnect scheme. To improve the electrical reliability of the wires, particularly with respect to electromigration, aluminum–silicon and aluminum–silicon–copper alloys are preferred over pure aluminum. However, in the late nineties the series resistance of aluminum-based lines became a bottleneck for circuit performance. After more than 30 years of use, the aluminum wires were finally abandoned in favor of a copper-based metallization scheme. Copper conducts about 60% better than aluminum, but patterning it into submicron wires is difficult because traditional plasma etch techniques do not work on copper. A novel process called "damascene" had to be introduced, due to its similarity with the well-known metal inlaid technique that goes back to the Middle Ages. Following this process, the conductor wires are embedded in a dielectric film deposited on the silicon substrate or on a previous interconnect layer. Copper lines are formed by electroplating the metal inside preformed trenches, which have been etched into the dielectric. The formation of a multilevel damascene scheme requires special polishing techniques (called chemical mechanical polishing or CMP) to planarize each interconnect layer by removing the excess metal outside the trenches (**Figure 4**).

Packaging

Packaging the IC serves a twofold purpose: to protect the chip against mechanical damage and corrosion, and to provide electrical connections with the outside circuitry. An IC package consists of a sealed box which is either from ceramic material (for hermetic sealing) or from a plastic mold (for low-cost packaging). Metallic pins are attached to the base plate of the package, to be inserted in the printed circuit board. After completing the back-end process, the backside of the chip is fixed to the package base plate using a conductive paste or metal solder. Next, the bonding pads of the chip are connected to the package pins. The most common technique is wire bonding: thin aluminum or gold wires are thermally bonded to the IC pads and to the pins, sometimes with the help of ultrasonic vibrations. Wire bonding is cheap and reliable, but cannot handle the increasing density of bonding pads in the most recent IC generations. Flip-chip bonding is now an increasingly popular alternative. Microscopic solder balls are first deposited on the IC bonding pads. The chip is then positioned face down to the package base plate, so that the solder balls can be attached directly to the pin pads. Since this method does not use wires, it offers both higher connection density and better reliability than the traditional bonding schemes. After wiring, the upper lid of the package is sealed or molded to the base plate for final encapsulation (**Figure 5**).

Processing Techniques

Photolithography

Photolithography uses a beam of photons to transfer a pattern written on an optical mask to the substrate surface. In a complex IC process, a wafer will go through the photolithographic step in the order of 20–30 times. Photolithography is, therefore, the key technology of IC fabrication and an essential driver behind the miniaturization trend commonly known

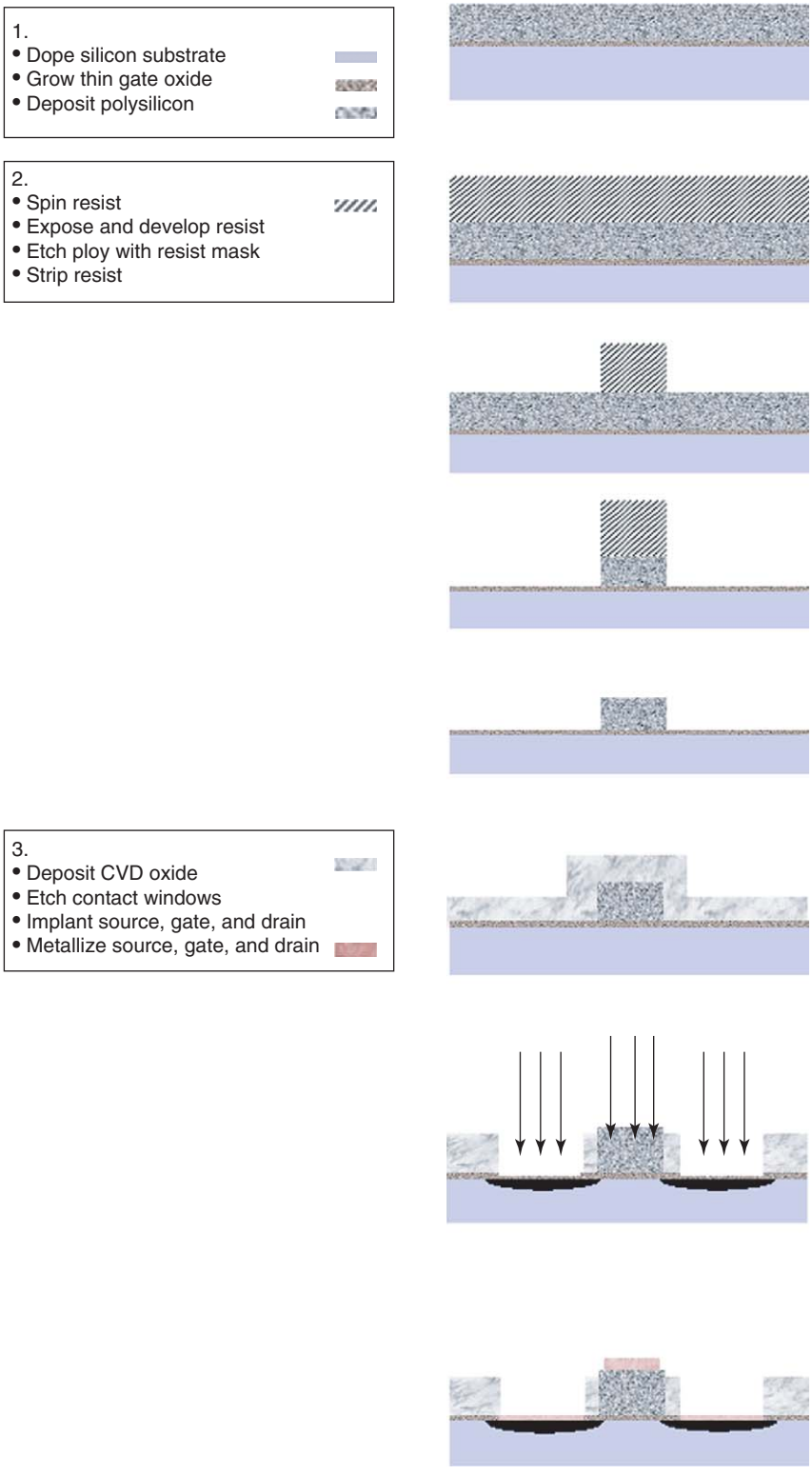


Figure 2 Schematic process flow for an MOS transistor.

as Moore's law. It was introduced almost simultaneously with the invention of the IC in 1959, but in the meantime, the minimum dimensions on the masks have decreased by several orders of magnitude, forcing continuous improvement in lithography tools and processes.

A standard lithography sequence involves a large number of steps. The wafer is initially cleaned and

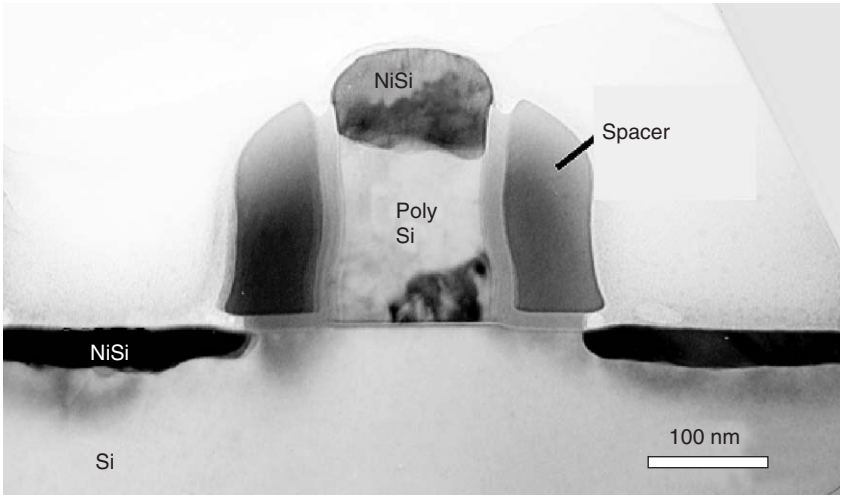
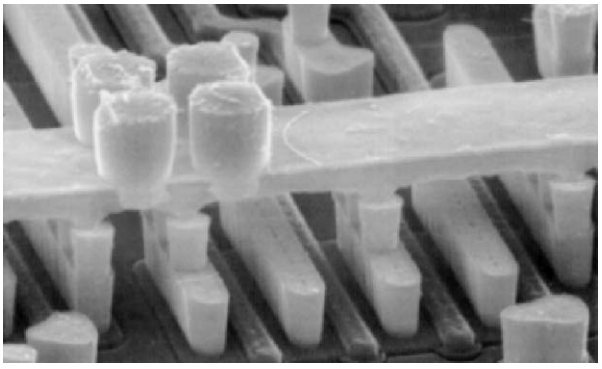
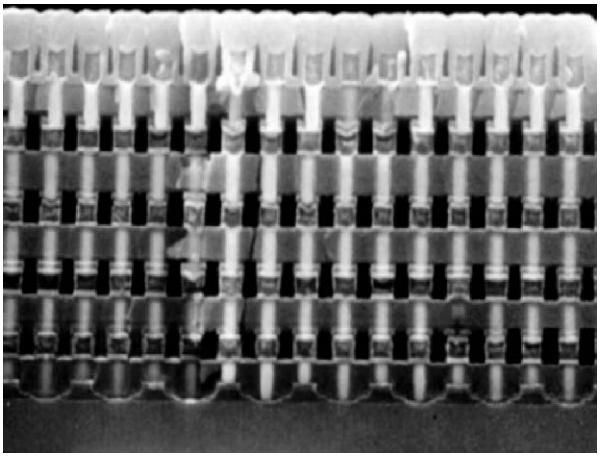


Figure 3 Cross-sectional scanning electron microscopy picture of an integrated MOS transistor. (Courtesy of IMEC.)



(a)



(b)

Figure 4 (a) Top view, and (b) cross section picture of a multilevel interconnect structure. (Courtesy of IMEC.)

Figure 5 IC in package. (Courtesy of IMEC.)

heated to remove any moisture from its surface. After deposition of an adhesion promotor, a liquid polymer called photoresist is spun on the surface while the wafer is kept rotating at high speed. After a thermal bake-out, the substrate is introduced in an optical projection tool called stepper/scanner. A light beam passing through the mask is focused on the wafer through a reduction lens system to produce the desired image in the photoresist. In order to achieve high resolution, only a small portion of the mask is imaged once, but the small image field is repeated over the surface of the wafer by a combination of stepping and scanning operations. During exposure, the photoresist material undergoes some light-triggered chemical reactions, which cause the illuminated regions to be either more or less acidic. If they become more acidic, the material is called a positive

photoresist, while in the reverse case it is a negative photoresist. The resist is then developed in an alkali solution such as sodium hydroxide (NaOH), which removes either the exposed (positive photoresist) or the unexposed (negative photoresist) polymer. After development, the wafer is finally "hardbaked" at high temperature in order to solidify the remaining photoresist. The resist can then be used as a patterning mask for etching, deposition or implantation. After patterning, the resist is stripped from the wafer with appropriate solvents.

The optical resolution of the lithography system is the main factor determining the smallest device size that can be fabricated at any time. The ability to form a clear image of a very small feature is limited by the wavelength of the light and the ability of the lens system to capture enough diffraction orders of the illuminated mask. Although theoretically the Rayleigh diffraction limit determines the maximal achievable resolution, in practice, it is possible to use special (but costly) tricks called resolution enhancement techniques that improve resolution beyond the Rayleigh limit by a factor of 2 or more. Current state-of-the-art photolithography tools use deep ultraviolet light (DUV) with wavelengths of 248 and 193 nm, allowing feature sizes below 100 nm to be printed with good yield. Extreme ultraviolet (EUV) systems using 13 nm radiation are presently under development for introduction by the end of this decade. Because such short wavelength radiation cannot be diffracted by lens systems or absorbed by bulk masks, the EUV technology will require a transition from diffractive to fully reflective optics.

Dry Etching

The aim of the etching process is to remove material from the IC surface in the areas left uncovered after photoresist development. For most patterning steps, dry (plasma) etching is now the standard procedure, while "wet" etching using liquid chemicals is still used to clean wafers. The main advantage of dry etching is its ability to achieve sharp etch profiles with nearly vertical sidewalls. This creates features with a high aspect ratio, as required by the relentless miniaturization of the circuit elements. A hot plasma formed by RF excitation of a neutral gas species is directed on the wafer by a static electric field. The radicals in the gas react with the substrate material to form volatile byproducts which are either evacuated from the etch chamber or deposited on the chamber walls. The mixture normally consists of a hydrogenated hydrocarbon diluted in an inert gas. However, the gas composition depends on the material to be etched, and will be different for silicon, metals or SiO_2. The strength of the DC field also plays a role in the etch process. A strong field increases the kinetic energy of the ions, and therefore enhances the amount of physical etching, which depends on mechanical sputtering. On the other hand, low-energy ions will act essentially by chemical etching, which proceeds through surface reactions resulting in smoother patterns. In this way, the etching mechanism can be tuned to optimize the shape of the patterned structures.

Deposition Techniques

PVD In standard silicon processing, physical vapor deposition (PVD) uses target sputtering rather than target evaporation because of its better step coverage. For aluminum interconnect layers, plasma-induced sputtering is the most convenient technique. The sputtered metal lands on the wafer surface where it forms a thin metal film, which can be further etched into wires. The film thickness is proportional to the width of the conductor lines, and usually lies in the range of a few hundred nanometers. Plasma sputtering can also be employed to deposit nonmetallic layers, especially insulators, although for the latter, CVD is the preferred method.

CVD Chemical vapor deposition is a process by which a thin film is deposited on a substrate by chemical reaction in a gas at an elevated temperature. If required, the reaction temperature can be lower by using plasma-enhanced CVD (PECVD). Over the years, CVD has become very popular due to its versatility, speed of deposition, and excellent step coverage. CVD can be used for metal, semiconductor, and insulating films. The structure of the deposited film can be monocrystalline, polycrystalline, or amorphous. The lattice structure depends primarily on the substrate, the reaction rate, and the temperature. Selective deposition is also possible, by choosing precursor gases that react preferentially with specific surfaces. CVD is the preferred method to deposit polycrystalline silicon, as well as various types of oxide films used for isolation of the interconnect layers. These interlayer dielectrics are grown by CVD at low pressure (LPCVD). Thin metallic layers such as titanium nitride (TiN), which are used as seed layers for copper electroplating or as a seal between the interlayer dielectrics, are also deposited by CVD.

Epitaxy Epitaxy is a process by which a deposited film is forced into a high degree of crystallographic alignment with the substrate lattice. Several epitaxy techniques are now available, such as molecular beam epitaxy (MBE), epitaxial CVD, or atomic layer

epitaxy (ALE). In all cases, the deposition process must be slow enough to allow the atoms to rearrange themselves on the surface according to the lattice orientation of the substrate. Contamination of the surface by impurities must be kept minimal, since it will disturb the epitaxial alignment. Epitaxial layers are now widely used in IC technology and the number of applications is still growing. In the early days, the epitaxy process was limited to deposition of silicon on silicon (homoepitaxy). Today, an increasing number of devices require thin film heteroepitaxy, for example, growing a binary alloy such as silicon–germanium (SiGe) epitaxially on the silicon wafer. For layers with high Ge concentration, the epitaxy is made difficult by the lattice mismatch between the substrate and the thin film. However, using appropriate techniques, the epitaxial layer can be grown pseudomorphically, meaning that the distance between the film atoms will adapt itself spontaneously to the bond length of the substrate material.

Physics of Integrated Circuits

Integration Density: Moore's Law

Moore's original statement, made in 1965, was modestly presented as an educated guess at the expected development of ICs over the next ten years. Or, to put it in his own words: "With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65.000 components on a single silicon chip" (see Further Reading Section). Almost four decades later, unit cost is still falling with the number of components, and as long as this favorable trend persists, the "law" will remain firmly in place. The continuation of Moore's law is made possible by the relentless shrinking of the MOSFET dimensions. Now, spanning more than two orders of magnitude, the downscaling of the MOSFET is a truly unique occurrence in the history of technology. The physical base for this miniaturization is a set of scaling rules defined some years after Moore's initial paper. In essence, the scaling algorithms tell us that the transistor performance (mainly its gate delay time and overall power dissipation) will improve with decreasing dimensions, provided the rules are properly applied. However, a point has been reached where the presence of nonscaling quantities such as the thermal voltage kT/q and the semiconductor bandgap (together with some others) can no longer be neglected. These nonscaling parameters are giving rise to so-called "short-channel" effects, such as threshold voltage shifts and many others, which tend to degrade the transistor switching speed and lower its power efficiency. These disturbing effects can be acted upon with specific counter-measures, partly technological and partly design-related. Up to now, device engineers have been remarkably successful in alleviating the negative impact of short-channel effects. However, it appears increasingly difficult to define a universal approach similar to the early scaling laws, and a certain amount of specialization in the circuit technologies (such as distinguishing between high-speed and low-power applications) is becoming unavoidable.

Speed and Power

Power and speed issues are strongly linked by the scaling rules. The width, length, and gate oxide thickness all scale roughly equally by a common factor. Therefore, the primary effect of the device scaling is to reduce all the capacitances, which provides a proportional decrease in power consumption and RC delays. However, not all the vertical dimensions of the circuit scale down by the same factor. In particular, the thickness of the interconnect lines cannot change as fast due to fundamental technology limitations. This increases the fringe capacitance from the metal side to the substrate, as well as the capacitance between adjacent line segments. Compounding this problem, the decreasing wire width raises the series resistance of the lines. As a result, RC delays in the lines tend to become more important with respect to transistor delays. This can be counteracted by optimizing the interconnect layout, and by introducing low-k dielectrics as the insulating layers between the copper wires.

The scaling rules also impact on the power consumption of the IC. There are two main sources of power dissipation: dynamic switching power due to the charging and discharging circuit capacitances, and static dissipation from leakage currents. When CMOS devices switch, the output is either charged up to the transistor bias voltage, or discharged down to the ground. The power dissipated during switching is therefore proportional to the switching speed and to the capacitive load. Decreasing the clock frequency of the circuit is, therefore, an efficient way to lower the dynamic power consumption.

Static power dissipation is linked with two types of leakage currents: reverse-bias diode leakage on the transistor drains, and subthreshold leakage through the channel when the transistor is turned off. Diode leakage must be tackled through process optimization, mainly by improving the quality of the junctions. Subthreshold current is a more complex issue. The process parameter that predominantly affects its value is the threshold voltage, since reducing the latter exponentially increases the subthreshold current.

Static power dissipation can, however, be reduced by shrinking the transistor size, and by lowering the supply voltage.

Applications

Most ICs belong to one of three main categories: memories, microprocessors, and custom ICs.

Memories

Random-access memory (RAM) A type of computer memory that can be accessed randomly; that is, any byte of memory can be accessed without disturbing the other bytes. RAM is the most common type of memory found in computers and other information processing devices. There are two basic types of RAM: dynamic RAM (DRAM) and static RAM (SRAM). Both are volatile memories, meaning that they lose their contents when the electrical power is turned off. RAM circuits are organized in cells, with each cell storing a binary 1 or 0. DRAM cells are very simple, consisting basically of one transistor and one capacitor for charge storage. Because this capacitor leaks, a DRAM needs to be refreshed thousands of times per second. This makes DRAM relatively slow and power-hungry. However, due to the small cell size, DRAM is best-suited for high-density storage (**Figure 6**). SRAMs are more complex, with 4–6 transistors per cell, providing a stable storage, which does not need to be refreshed. SRAM is very fast and power-efficient, but the complexity of its cell makes it much more expensive than DRAM. Due to its high cost, SRAM is used mainly as a memory cache.

Read-only memory (ROM) Computers almost always contain a read-only memory that holds instructions to boot the operating system and perform diagnostics. ROM is nonvolatile, that is, it does not need electrical power to keep its data stored. Like RAM, ROM memories are random-access, but they cannot change their content during the computer operation. However, some types of ROM allow changing the content by special procedures. The most common ones are erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM) and FLASH EEPROM. FLASH (not an acronym) memories, which store the information bits in a second transistor gate called the floating gate, have become increasingly popular in recent years. They can be erased and reprogrammed in blocks instead of one byte at a time, which is faster and less power consuming than other EEPROMs. PCs have their basic input-output system (BIOS) stored on a flash memory chip so that it can easily be updated if necessary. Flash memory is also very common in communication appliances and portable devices such as cellular phones and smartcards.

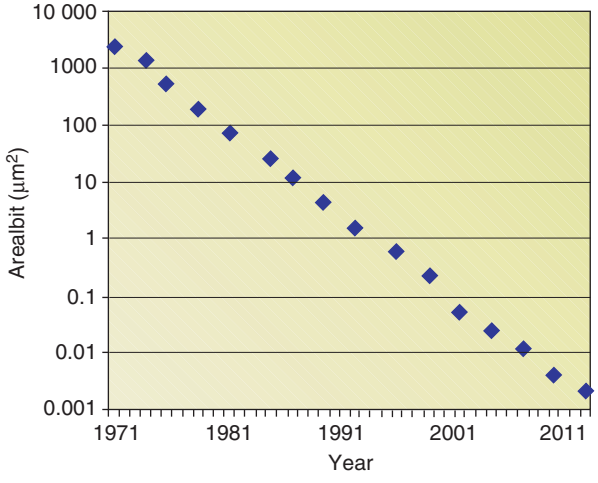


Figure 6 Moore's law in action: evolution of the chip area / bit for a standard DRAM memory cell.

Microprocessors and Custom ICs

Microprocessors Also called central processing unit (CPU), microprocessors are ICs that can execute instructions by running a software program. Although computer microprocessors have become a benchmark for the IC industry, they only represent a small part (~2% in unit sales) of the total CPU market. By far, the largest, by number, are the embedded microprocessors (also called microcontrollers), which can be found in many consumer products, ranging from toys to automobiles.

Custom ICs In contrast to standard memories and microprocessors, custom ICs are specially designed for specific product applications. The main types are:

- *Application-specific ICs (ASICs)*. They are meant to achieve optimal performance in terms of speed and/or power consumption. ASICs are custom-designed for a specific application, either full custom (meaning that the design is made from scratch) or starting from standard cells or gate arrays, where the designers can take advantage of pre-existing blocks.
- *Field-programmable chips*. These are special chips that can be reconfigured by the customer. The most popular ones are field-programmable gate arrays (FPGAs). Their circuits are diced into regular arrays, which can be disconnected by blowing microscopic fuses. In this way, the desired functionality is "burned into" the IC. This process is reversible, meaning that the original connections can be

restored and a new pattern burned. Although FPGAs are usually slower and less power-efficient than ASICs, they have become very popular due to their flexibility and ease of programming.

- *Analog/RF ICs*. Even if digital signal processing today represents the largest market for the semiconductor industry, there is also a wide variety of ICs handling analog and high-frequency signals. Some chips process both digital and analog information and are called mixed-mode ICs. Typical applications are amplifiers, mixers, multiplexers, and analog-to-digital converters. Wireless communication devices make heavy use of these components, which can be found in every cellular phone set.

See also: Semiconductor Devices; Semiconductors, General Properties; Sensors; Silicon, History of.

PACS: 85.30.–z; 85.40.–e; 85.40.Hp; 85.40.Ls; 85.40.Qx; 85.40.Ry; 85.40.Sz

Further Reading

Moore GE (1965) Cramming more components onto integrated circuits. *Electronics* 38(8): 114–117.

Sze SM and Chang CY (1996) *ULSI Technology*. New York: Mc Graw-Hill.

Turley J (2003) *The Essential Guide to Semiconductors*. Upper Saddle River: Prentice-Hall.

Wolf S and Tauber RN (2000) *Silicon Processing for the VLSI ERA*, 2nd edn. Sunset Beach: Lattice Press.

Interaction of Light and Matter

G Nienhuis, Universiteit Leiden, Leiden, The Netherlands

Introduction

Light is electromagnetic radiation in the visible part of the spectrum. Radiation and matter are the main physical objects in the world. Their behavior is described by electromagnetism and mechanics, with Maxwell's equations and Newton's laws as basic ingredients. Radiation is the only means by which one can obtain information on the remote parts of the universe. In general, without interaction between light and matter, the world would not only be invisible, but dead as well. Radiation from the Sun is by far the main source of energy on Earth, and serves as the basis of life. This energy source is tapped by the process of photosynthesis, which starts out as the interaction of photons with individual molecules. The interaction of light and matter is discussed both at the microscopic level of atoms and molecules, and at the macroscopic level, in terms of the optical properties of materials.

Transition Driven by Broadband Light

Atoms and molecules have discrete energy levels, as a result of the quantum-mechanical nature of the motion of the electrons. Consider atoms with a lower level 1 and an upper level 2, with an energy separation $E_2 - E_1 = \hbar\omega_0$, with ω_0 the transition frequency. The atoms are driven by a light field that is nearly resonant with the transition $1 \leftrightarrow 2$, which means that the frequency ω of the field is close to the transition frequency. Then the two levels are coupled by the light field. When the atom is initially in state 1, it will start oscillating in a linear superposition of the states 1 and 2. For a conventional light source, with a broad spectrum, Einstein argued that the atom can make three types of transitions between the states, each one at a fixed rate. An atom in the lower state 1 can become excited while absorbing a photon, at a rate BIn_1 that is proportional to the intensity per unit frequency I (also called the spectral intensity). An atom in the upper state 2 can go spontaneously to the lower state at a rate An_2, while emitting a photon. However, using arguments of thermal equilibrium, Einstein showed that emission of a photon must also occur in a stimulated way, at a rate BIn_2 that is also proportional to the spectral intensity. The emitted photon will have the same properties (frequency and propagation direction) as the photons in the incident beam. The coefficient A is called the Einstein coefficient for spontaneous emission, and B is the Einstein coefficient for stimulated transitions. These coefficients are properties of the atom, and also depend on the specific transition. The transitions in the energy level scheme are sketched in **Figure 1**.

As a result of these considerations, the number of atoms n_1 and n_2 in the states 1 and 2 must obey the rate equations

$$\frac{\mathrm{d}}{\mathrm{d}t}n_2 = -\frac{\mathrm{d}}{\mathrm{d}t}n_1 = BI(n_1 - n_2) - An_2 \qquad [1]$$

The three terms on the right-hand side describe absorption, stimulated emission, and spontaneous

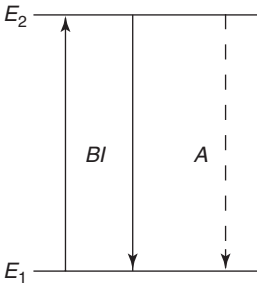


Figure 1 Energy level scheme of two states of an atom. The indicated transitions are absorption and stimulated emission (solid lines) at a rate BI, and spontaneous emission (dashed line) at rate A.

emission. Radiative transitions correspond to the creation or annihilation of photons. At absorption, a photon disappears from the light beam, whereas stimulated emission adds a photon to the beam. Therefore, the net rate of absorption of photons from the beam (direct absorption minus stimulated emission) is equal to $N_A = BI(n_1 - n_2)$. Spontaneously emitted photons can go in all directions, at a rate $N_E = An_2$. The balance between the number of excited atoms and the number of photons is best seen when the above equation is written as $dn_2/dt = N_A - N_E$. The efficiency of absorption as compared to spontaneous emission is measured by the ratio $s = BI/A$, which is called the saturation parameter.

When initially all atoms are in the lower state, the time-dependent solution for the populations of the two states are

$$\begin{aligned} n_1(t) &= \frac{n}{1+2s}[1 + s(1 + e^{-(1+2s)At)})] \\ n_2(t) &= \frac{ns}{1+2s}[1 - e^{-A(1+2s)t}] \end{aligned} \qquad [2]$$

with $n = n_1 + n_2$ the total number of atoms. It is seen that the populations approach their steady-state value exponentially. These expressions also hold for a single atom. Then the populations n_1 and n_2 must be interpreted as the probabilities that the atom is in the state 1 or 2. In this case, $n = n_1 + n_2 = 1$. Then the steady-state values are given by $n_2 = 1 - n_1 = s/(1 + 2s)$, which approaches the maximal value 1/2 for high intensity.

Transition Driven by Monochromatic Light

In modern optics, broadband light sources are commonly replaced by lasers, with a coherent and practically monochromatic output. Then the dynamics of the driven transition is quite different. The strength of the atom–field coupling is indicated by the product of the amplitude of the oscillating electric field of the light and the electric dipole moment of the transition. This product (which is an energy) is expressed as $\hbar\Omega$, so that Ω is a frequency, commonly called the Rabi frequency. Quantum mechanically, the state of a single two-level system is described by a wave function ψ, which is a linear superposition $\alpha_1\psi_1 + \alpha_2\psi_2$ of the wave functions ψ_1 and ψ_2 corresponding to the two states. The populations are $n_1 = \alpha_1\alpha_1^*$ and $n_2 = \alpha_2\alpha_2^*$. The product $\sigma_+ = \alpha_2\alpha_1^* = \sigma_-^*$, which is called the optical coherence of the atom, also has a direct physical significance, since it is proportional to the oscillating electric dipole induced by the field. The coupling between this dipole and the electric field determines the work that the field exerts on the atom, which is related to the net absorbed power. For monochromatic light, the behavior of the population and the coherence is given by the coupled equations

$$\begin{aligned} \frac{d}{dt}n_2 &= -\frac{d}{dt}n_1 = -An_2 + \frac{i}{2}\Omega(\sigma_+ - \sigma_-) \\ \frac{d}{dt}\sigma_+ &= -\left(\frac{1}{2}A - i\Delta\right)\sigma_+ + \frac{i}{2}\Omega(n_1 - n_2) \end{aligned} \qquad [3]$$

where $\Delta = \omega - \omega_0$ is the frequency detuning of the light from atomic resonance.

For monochromatic driving light, these equations [3] replace the rate equations [1] above. They closely resemble the Bloch equations for the behavior of a magnetic dipole in an oscillating magnetic field, as it appears in the theory of (nuclear or electronic) magnetic resonance. In analogy, the present equations are often called the optical Bloch equations. The steady state for the populations obeys the same expressions $n_2 = 1 - n_1 = s/(1 + 2s)$ as in the broadband case, provided that the saturation parameter is now defined as $s = \Omega^2/(A^2 + 4\Delta^2)$. Since s is proportional to Ω^2, it is also proportional to the square of the field amplitude, and thereby proportional to the intensity I. The variation of this parameter with the frequency ω reflects the absorption profile of the atom, with its maximum at atomic resonance. Even though the steady-state populations have the same form as in the broadband case, the way in which n_1 and n_2 approach their steady-state value is quite different in these two cases. For monochromatic radiation, the exponential behavior of eqn [2] is replaced by a damped oscillatory behavior (Rabi oscillations). A comparison of the approach to the steady state for monochromatic and broadband excitation is shown in **Figure 2**. When $\Omega \gg A$, that is, for strong saturation, the excited-state population is well approximated by the expression

$$n_2(t) = \tfrac{1}{2}[1 - e^{-3At/4}\cos(\Omega t)]$$

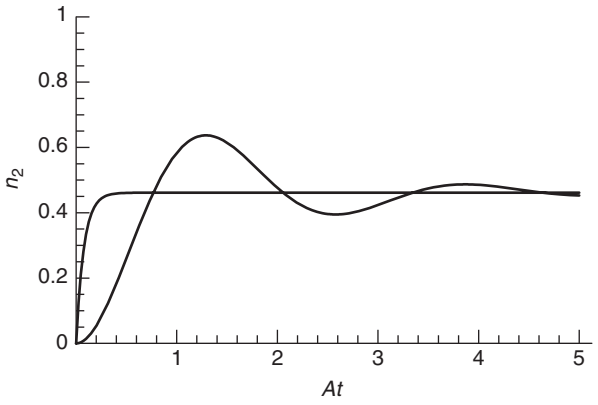


Figure 2 Time dependence of excited-state population in a broadband field (monotonic curve) and a monochromatic field (oscillating curve) at resonance. The saturation parameter amounts to $s = 6$ in both cases.

whereas for broadband radiation one finds from eqn [2] that $n_2(t) = [1 - \exp(-2Ast)]/2$.

The equations so far hold for a closed transition, where no other states than these two get populated. In practice, the population of the upper level can get transferred to many different states, in particular, for systems more complex than atoms. This leads to all kinds of radiative or radiation-free decay processes, with the final result that the absorbed light energy gets transformed in other forms of energy. But also for atoms the simple picture given above is a simplification, since an energy level is usually composed of a number of substates, which can be selectively addressed by light with a specific polarization. Moreover, at high intensities, transitions can occur under the absorption of two or more light quanta at the same time, so that light with a frequency ω can excite a transition with a transition frequency in the neighborhood of 2ω.

Propagation of Light in Gas Medium

When light with a frequency ω propagates through a dielectric medium, the electric field at each position induces a density of electric dipole moment, which is called the dielectric polarization $\boldsymbol{p}$. For not too high intensities, saturation can be ignored, and the polarization is proportional to the electric field. Then the dielectric medium is said to be linear. However, as seen above, the proportionality constant can depend strongly on the frequency ω, in particular, when this is near an atomic resonance frequency ω_0. When the electric field is expressed in the form of a plane wave as $\boldsymbol{E}(z,t) = 2\ \mathrm{Re}\ \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}kz - \mathrm{i}\omega t}$, and the polarization as $\boldsymbol{P}(z,t) = 2\ \mathrm{Re}\ \boldsymbol{P}_0 \mathrm{e}^{\mathrm{i}kz - \mathrm{i}\omega t}$, the response of the medium to the field is expressed as $\boldsymbol{P}_0 = \varepsilon_0 \chi(\omega) \boldsymbol{E}_0$, with χ the frequency-dependent susceptibility of the medium. For a dilute gas, the polarization is just the sum of the induced dipoles of the individual atoms. From the optical Bloch equations, one finds the electric dipole moment of each atom, and multiplication with the atom density ρ gives the polarization $\boldsymbol{P}$. This gives the expression

$$\chi(\omega) = -\varrho|\mu|^2 \frac{1}{\varepsilon_0 \hbar} \frac{1}{\Delta + iA/2}$$

with μ the strength of the atomic transition dipole. This expression is reliable near resonance, when the detuning Δ is small compared with the frequency ω. The susceptibility can be separated in its real and imaginary part, as $\chi = \chi_1 + \mathrm{i}\chi_2$. The presence of an imaginary part indicates that $\boldsymbol{E}$ and $\boldsymbol{P}$ are not in phase, so that the field exerts work on the medium, and the light intensity cannot be constant. Maxwell's equations show that the frequency ω and the wave number k are related by the dispersion relation $c^2k^2 = \omega^2(1 + \chi(\omega))$, with c the velocity of light in vacuum. The factor $1 + \chi = (n + \mathrm{i}\kappa)^2$ serves as the square of a complex refractive index. For a real frequency ω, the wave number $k = (n + \mathrm{i}\kappa)\omega/c$ is complex, and the field $\boldsymbol{E}$ decays as a function of z as $\exp(-\omega\kappa z/c)$. An incident beam loses energy while propagating through the medium, and the intensity I (which is proportional to the absolute square of the field amplitude) decreases as $\exp(-2\omega\kappa z/c)$. This is the solution of Beer's law of absorption $\mathrm{d}I/\mathrm{d}z = -\alpha_0 I$, with $\alpha_0 = 2\omega\kappa/c$ the unsaturated absorption coefficient.

For an atomic vapor, the loss factor κ is proportional to the absorption profile of the atoms. The absorbed energy is reemitted in the form of fluorescence radiation. For more complex media, the absorbed energy can be dissipated by all kinds of radiative or nonradiative ways, and turned into heat, or stored in some form of chemical energy.

When the light drives a transition where the population of the upper state n_2 is higher than n_1, the work exerted on the medium by the field is negative, so that the field actually gains energy. This is the situation of population inversion, which cannot occur in thermal equilibrium. It requires a pump mechanism to feed atoms in the upper state 2. The result is a negative value of χ_2, and thereby of κ, and the beam grows in intensity during propagation. This is the basis of laser action, where stimulated emission on the transition outweighs absorption.

Linear Optics in Continuous Media

Matter interacting with light can often be described as a continuum. This is true provided that the

wavelength of the light is large compared with the interatomic distance. The optical properties of the medium are then characterized by macroscopic coefficients, and the electric and magnetic fields in the medium obey a version of Maxwell's equations that contain these coefficients. Dielectric materials in general have a susceptibility $\chi(\omega)$ that specifies the dielectric polarization $\boldsymbol{P}$ for a given electric field $\boldsymbol{E}$. The frequency dependence of χ indicates that the response of the medium to the field is not instantaneous, but that its time dependence can be expressed as an integral over the past

$$\boldsymbol{P}(t) = \int_0^\infty \mathrm{d}\tau\, f(\tau)\, \boldsymbol{E}(t-\tau)$$

in terms of a memory function $f(\tau)$. When a monochromatic field $\boldsymbol{E}$ is substituted, a monochromatic polarization results with the same frequency, and the susceptibility is found as

$$\varepsilon_0\chi(\omega) = \int_0^\infty \mathrm{d}\tau\, f(\tau)\mathrm{e}^{\mathrm{i}\omega\tau}$$

in terms of the Fourier transform of the memory function. The principle of causality is expressed by stating that the response of a system to an external cause cannot precede the cause in time. In the case of a dielectric response, this is expressed by the fact that the response function $f(\tau)$ vanishes for negative values of the time delay τ. Applying the elementary theory of complex functions then leads to relations between the real and imaginary part of the susceptibility, in the form

$$\chi_2(\omega) = -\frac{1}{\pi}\mathscr{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{\chi_1(\omega')}{\omega'-\omega},$$
$$\chi_1(\omega) = \frac{1}{\pi}\mathscr{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{\chi_2(\omega')}{\omega'-\omega}$$

with $\mathscr{P}$ indicating the principal value of the integration. The Kramers–Kronig relations, which hold for all cases of linear response, show that the real and imaginary part of the susceptibility are determined by one another. Moreover, since the response function f is real, it is easy to show that $\chi_1(-\omega) = \chi_1(\omega)$, and $\chi_2(-\omega) = -\chi_2(\omega)$, so that $\chi_1(\omega)$ is an even function of the frequency, whereas $\chi_2(\omega)$ is odd.

Due to the frequency-dependence of the susceptibility, the velocity of a light wave becomes a slightly ambiguous concept. In a wave with frequency ω and wave number k, the crests and the troughs propagate with the velocity $v_{\mathrm{ph}} = \omega/k$, which is called the phase velocity. In a medium with refractive index n, the phase velocity is $v_{\mathrm{ph}} = c/n(\omega)$. A wave packet with a finite spatial extent is not monochromatic, and it must be a superposition of waves with at least slightly varying frequencies. When the frequency distribution is centered in the neighborhood of the frequency ω_0, the packet propagates with the group velocity $v_{\mathrm{gr}} = \mathrm{d}\omega/\mathrm{d}k$, where the derivative is taken at the value ω_0. The dispersion relation $\omega n(\omega) = ck$ leads to the expression for the group velocity $v_{\mathrm{gr}} = c/n_{\mathrm{gr}}$, with $n_{\mathrm{gr}} = n + \omega \mathrm{d}n/\mathrm{d}\omega$ the group refractive index. In specific ranges of the frequency, the phase velocity can be larger than the speed of light c, but the group velocity, which is the velocity of a physical signal, must be smaller than c.

For a monochromatic field, the dielectric properties of a material are expressed in the linear relation $\boldsymbol{D} = \varepsilon(\omega)\boldsymbol{E}$, where the dielectric displacement is defined by $\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}$. The coefficient $\varepsilon(\omega)$ is called the dielectric permittivity (or dielectric constant) of the medium, which is related to the susceptibility as $\varepsilon(\omega) = \varepsilon_0(1 + \chi(\omega))$. The previous section indicates how the absorption depends on the imaginary part of the susceptibility. When the dielectric permittivity is separated into its real and imaginary parts as $\varepsilon(\omega) = \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega)$, the Kramers–Kronig relations give

$$\varepsilon_1(\omega) = \varepsilon_0 + \frac{1}{\pi}\mathscr{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{\varepsilon_2(\omega')}{\omega'-\omega}$$

A conducting medium, such as a metal, can in some respects also be described by a permittivity. The current density $\boldsymbol{j}$ is proportional to the local electric field, so that $\boldsymbol{j} = \sigma\boldsymbol{E}$, with σ the electric conductivity. Both the current and the displacement appear in the Ampère–Maxwell equation

$$\boldsymbol{\nabla} \times \boldsymbol{H} = \boldsymbol{j} + \frac{\partial \boldsymbol{D}}{\partial t}$$

with $\boldsymbol{H}$ the magnetic field. For monochromatic fields at frequency ω, the right-hand side takes the form $[\sigma - \mathrm{i}\omega\varepsilon(\omega)]\boldsymbol{E}$, so that the conductivity is equivalent to a contribution $\mathrm{i}\sigma/\omega$ to the dielectric constant. This contribution has a pole at zero frequency.

For high frequencies, well above the resonances in the material, $\varepsilon(\omega)$ must approach the vacuum value ε_0, since the electrons in the material can no longer follow the rapid oscillations. The asymptotic behavior can be estimated by noting that an electron (with charge $-e$) in a rapidly oscillating field $2\,\mathrm{Re}\,\boldsymbol{E}\mathrm{e}^{-\mathrm{i}\omega\mathrm{t}}$ will undergo an oscillating force $-2e\,\mathrm{Re}\,\boldsymbol{E}\mathrm{e}^{-\mathrm{i}\omega\mathrm{t}}$. This in turn produces an oscillation in position $\boldsymbol{r}_0 + 2\,\mathrm{Re}\,\boldsymbol{b}\mathrm{e}^{-\mathrm{i}\omega\mathrm{t}}$, with the amplitude $\boldsymbol{b} = -e^2\boldsymbol{E}/m\omega^2$, and m the mass of the electron. This oscillating charge corresponds to a dielectric polarization. When the effective number density of electrons is

given as N, the effective susceptibility of the medium is $\chi(\omega) = -e^2N/(m\omega^2\varepsilon_0)$. In practice, only the loosely bound electrons have to be taken into account here. This estimates the high-frequency limit of the susceptibility, and thereby the approach of the dielectric constant to ε_0.

In nonisotropic media, such as crystals with a low symmetry, the direction of the dielectric polarization $\boldsymbol{P}$ is not necessarily parallel to the direction of the electric field $\boldsymbol{E}$. Such a situation can be described by the linear relation $\boldsymbol{P} = \varepsilon_0\tilde{\chi}^{(1)}(\omega) \times \boldsymbol{E}$, where now $\tilde{\chi}^{(1)}(\omega)$ is a susceptibility tensor that takes the form of a 3×3 matrix. The symmetry properties of this matrix reflect the crystal symmetry. In this case, for each direction of the wave vector $\boldsymbol{k}$ of a plane wave, the medium has two different refractive indices, each one corresponding to a specific polarization of the light. The material is then birefringent: light incident on such a crystal can be split into two refracted beams, each one with a different polarization.

Nonlinear Optics

When an intense light beam traverses a gaseous medium, the dipole polarization is affected by saturation. The dipole polarization is then given by the relation $\boldsymbol{P}_0 = \varepsilon_0\chi_{\text{sat}}(\omega)\boldsymbol{E}_0$, with

$$\chi_{\text{sat}}(\omega) = -\varrho|\mu|^2\frac{1}{\varepsilon_0\hbar}\frac{1}{\Delta + \mathrm{i}A/2}\frac{1}{1+2s}$$

The presence of the saturation parameter s, which is proportional to the intensity, causes the susceptibility to decrease with the intensity. Saturation intensity I_{sat} is defined by rewriting the saturation parameter as $s = I/2I_{\text{sat}}$. The saturation factor in the susceptibility also modifies the absorption law, which now takes the form

$$\frac{\mathrm{d}I}{\mathrm{d}z} = -\frac{\alpha_0 I}{1 + I/I_{\text{sat}}}$$

The effective absorption coefficient decreases for increasing intensity, so that the medium becomes more transparent at higher intensities. When $I \ll I_{\text{sat}}$, Beer's law is recovered, with the linear absorption coefficient α_0, while for high intensities $I \gg I_{\text{sat}}$, the intensity decreases linearly with z, according to the rule $I(z) \simeq -(z - z_0)\alpha_0 I_{\text{sat}}$. By the same token, the effective refractive index at high intensity deviates less from unity than at intensities well below the saturation intensity.

Characteristic for saturation is that the response of the medium (the dielectric polarization) is no longer linear in the electric field. In many cases of nonlinear response, it is convenient to consider an expansion of the response in powers of the field, of the type

$$P = \varepsilon_0(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \cdots)$$

with $\chi^{(n)}$ higher-order susceptibilities. For simplicity, for the moment the vector character of the field is ignored, thereby ignoring polarization effects. More complex effects can arise when a medium is irradiated by several light beams at different frequencies. Due to the different frequencies of oscillation, components of P with various frequencies will arise.

As an example, consider a light field that is composed of two fields with different frequencies

$$E(t) = 2\ \mathrm{Re}(E_1\mathrm{e}^{-\mathrm{i}\omega_1 t} + E_2\mathrm{e}^{-\mathrm{i}\omega_2 t})$$

so that its square contains terms proportional to all combinations $\exp(\pm\mathrm{i}\omega_i t \pm \mathrm{i}\omega_j t)$, with $i, j = 1, 2$. Each of these terms contributes to the second-order dielectric polarization, which is expressed as the sum of terms $2\ \mathrm{Re}\ P(\omega)\exp(-\mathrm{i}\omega t)$, where the oscillation frequency ω can attain all values that arise from addition or subtraction of the two field frequencies ω_1 and ω_2. Five possible terms are found. Two terms $P(2\omega_i) = \varepsilon_0\chi^{(2)}(2\omega_i)E_i^2$ for $i = 1$ or 2 express second-harmonic generation, where the dielectric polarization oscillates at twice the frequency of a field component. One term $P(\omega_1 + \omega_2) = 2\varepsilon_0\chi^{(2)}(\omega_1 + \omega_2)E_1E_2$ describes a polarization component oscillating at the sum frequency, and another term $P(\omega_1 - \omega_2) = 2\varepsilon_0\chi^{(2)}(\omega_1 - \omega_2)E_1E_2^*$ oscillates at the frequency difference. The physical processes are called sum-frequency and difference-frequency generation. Finally, there is a nonoscillating contribution $P(0) = \varepsilon_0\chi^{(2)}(0)(E_1E_1^* + E_2E_2^*)$. The process by which an optical field can create a constant dielectric polarization is referred to as optical rectification. In **Figures 3** and **4**, the processes of sum-frequency and difference-frequency generation are illustrated in terms of energy-level diagrams of the medium.

Each one of the oscillating polarization components can cause radiation. This radiation will be efficient only in directions where the contributions from different parts in space are in phase, so that constructive interference occurs. This is the condition of phase matching. For plane waves, this means that the wave vectors of the beams obey the same sum rule as the frequencies. For instance, for sum-frequency generation, the field at frequency $\omega_3 = \omega_1 + \omega_2$ will be emitted in the direction for which the wave vector is $\boldsymbol{k}_3 = \boldsymbol{k}_1 + \boldsymbol{k}_2$. The detailed study of these conditions, combined with the polarization directions of the driving and the emitted radiation is the subject of the field of nonlinear optics. It will be

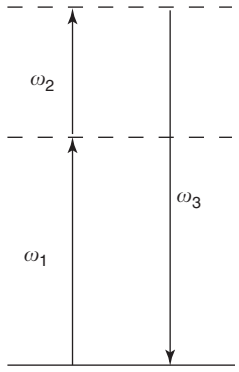


Figure 3 Process of sum-frequency generation in an energy-level diagram. Absorption of two photons with frequencies ω_1 and ω_2 leads to emission of a single photon with frequency $\omega_3 = \omega_1 + \omega_2$.

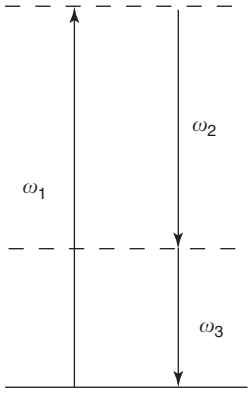


Figure 4 Process of difference-frequency generation in an energy-level diagram. Absorption of a photon with frequency ω_1 and stimulated emission of a photon with frequency ω_2 leads to emission of a photon with frequency $\omega_3 = \omega_1 - \omega_2$.

obvious that the number of combinations increases rapidly with the order of nonlinearity and the number of driving fields. For a medium with inversion symmetry, which can be mapped onto itself by space inversion, the second-order susceptibility $\chi^{(2)}$ must be zero for symmetry reasons. To see this, it is sufficient to consider the system consisting of the medium and the two input fields. Inversion of the entire system changes the sign of the input fields, while the medium does not change when it is inversion symmetric. The second-order expressions show that the output field (the created polarization) remains the same, whereas it changes sign under the inversion. Since the output is equal to its inverse, it must vanish.

Any nonlinear optical process can also take place in a spontaneous version. The description of these processes requires a quantum mechanical version of the corresponding classical theory. In the photon picture in a nonlinear crystal, a pump photon at frequency ω_1 can give rise to the emission of two photons with frequencies ω_2 and ω_3, with $\omega_1 = \omega_2 + \omega_3$, even when no beam at these frequencies is prepared. The process is called spontaneous parametric downconversion. It will be efficient in directions where phase matching is respected. This process is important in the emerging field of quantum information, since under proper circumstances the created pair of photons can be in a highly entangled state. These states display strong correlations that violate Bell's inequalities, and that cannot be explained in terms of classical physics.

Photonic Crystals and Optical Lattices

Photonic crystals are materials in which the refractive index is a periodic function of space, where the period is of the order of a wavelength. They can be designed and fabricated by modern lithographic or mechanical techniques. The optical properties of these materials can be compared with the electronic properties of crystal lattices. The dispersion relation between frequency and wavenumber is strongly affected by the structure. This is not easy to realize in the optical domain, and the first realizations concerned the microwave region. Situations of special interest occur when the periodicity gives rise to forbidden regions in frequency, where the density of modes vanishes. These regions are analogous to energy gaps in the electronic density of states. These optical bandgaps give rise to special effects, since light in that frequency region cannot propagate through the medium. As regarded from the outside, this implies that the medium should be highly reflecting in that region, whereas an excited atom placed inside cannot decay spontaneously when its resonance frequency falls within the gap. The analogy between electromagnetic waves in photonic material and electron propagation in crystals creates an interesting playground, where expertise from one field can be applied in the other one.

A comparable analogy between optics and condensed matter physics is found in the area of optical lattices. They are formed when atoms are trapped in the periodic optical potential wells that arise when pairs of light beams with different propagation directions interfere. These potentials describe the dipole forces that light with an intensity or a polarization gradient exerts on atoms. When one atom is trapped in each one of the periodic array of potential wells, an optical lattice is formed. In practice, it is not easy to obtain high atomic filling factors. Since the well depths are modest, and the optical forces are weak, the kinetic energy of the atoms must be quite low. The long-range periodicity is imposed by the light field,

and the interatomic interactions play no role. In this way, lattices are formed that are free from defects, and with a lattice symmetry that can be easily tailored by varying the number or the propagation direction of the composing light beams.

Photonic crystals and optical lattices display an attractive duality between light and matter. The periodic structure is provided by the material structure in the case of photonic crystals, and by the light beams for optical lattices. The waves in these structures are matter waves for an optical lattice, and light waves in a photonic crystal. A comparison between the developments in both fields gives a beautiful illustration of the use of analogies between neighboring domains of physics.

See also: Dielectric Function; Electrodynamics: Continuous Media; Nonlinear Optics; Optical Absorption and Reflectance; Optical Properties of Materials; Optical Sum Rules and Kramers–Kronig Relations.

PACS: 32.80. – t; 42.65. – k; 77.22. – d

Further Reading

Allen L and Eberly JH (1975) *Optical Resonance and Two-Level Atoms*. New York: Wiley.
Boyd RW (1992) *Nonlinear Optics*. Boston: Academic Press.
Jackson JD (1962) *Principles Classical Electrodynamics*. New York: Wiley.
Joannopoulos JD, Meade RD, and Winn JN (1995) *Photonic Crystals*. Princeton: Princeton University Press.
Landau LD and Lifshitz EM (1960) *Electrodynamics of Continuous Media*. Oxford: Pergamon.
Letokhov VS and Chebotaev VP (1977) *Principles of Nonlinear Laser Spectroscopy*. Berlin: Springer.
Mandel L and Wolf E (1995) *Optical Coherence and Quantum Optics*. Cambridge: Cambridge University Press.

Intermetallic Compounds, Electronic States of

O Gourdon, D Gout, and G J Miller, Iowa State University, Ames, IA, USA

Introduction

Intermetallic compounds provide a rich and diverse resource to study the relationships among chemical composition, atomic structure, electronic structure, and physical properties. Since the majority of chemical elements are "metals," even with thermodynamic constraints, the numbers, types, and complexity of intermetallic compounds are immense, which is readily apparent in *Pearson's Handbook of Intermetallic Phases*. An understanding of the electronic states of an intermetallic compound can provide insights into the chemical bonding factors influencing composition and structure as well as the electronic underpinnings of its electrical and magnetic behavior. Experimental determination of electronic states, largely through photoelectron spectroscopy, demands pure samples, which can be a challenge to obtain because most metals react readily with oxygen. On the other hand, there is a great deal of effort on quantum mechanical calculations of the electronic structure. This article provides an overview of a comprehensive general picture of electronic states in intermetallic compounds, including brief descriptions of specific examples.

What Are Intermetallic Compounds?

At first glance, the answer seems obvious: intermetallic compounds are chemical substances formed when two or more metallic elements combine with a definite composition. They typically have small heats of formation (ca. -50 to $-10\,\mathrm{kJ\,mol^{-1}}$) when compared to salts and polymeric compounds (ca. -500 to $-100\,\mathrm{kJ\,mol^{-1}}$), but frequently show finite ranges in composition for a single phase, called homogeneity widths. Some melt congruently without decomposition into different phases, while many decompose peritectically into a liquid phase and another solid with a different composition. Concerning their structures, the various metal constituents are usually ordered in different sublattices, each with its own distinct population of atoms. In this way, elements of differing sizes or electronegativities can combine to give new compounds. Often, the corresponding structures are distinct from those of the component elements, but this is not necessary.

One important aspect of the definition is the combination of "metallic" elements. Chemical substances are classified as either metals or nonmetals, and many periodic tables of the elements include a zigzag line, with metals on the left and nonmetals on the right. A closer examination of the elements near this line reveals numerous unusual trends in structural and physical properties, such that in addition to the classification as normal metals, semiconductors and insulators, there are also "metametals" and "semimetals."

The different classification schemes for "metals" show the following characteristics.

1. Normal metals (alkali, alkaline earth, noble (Cu, Ag, Au) and many "transition metals"):
   (a) Ductile materials based on close-packed structures with high coordination numbers (12, 8 + 6) and high symmetry (cubic or hexagonal).
   (b) Typically electropositive elements with low first ionization potentials and low electron affinities.
   (c) Increasing electrical resistivities with increasing temperature; values $\sim 1\,\mu\Omega\,\text{cm}$.
   (d) Temperature-independent paramagnetism. Examples include the alkali, alkaline earth, noble (Cu, Ag, Au), and many "transition metals."
2. Metametals (Zn, Cd, Hg, Ga, In, Tl, Pb, and β-Sn as assigned by W Klemm in the 1950s):
   (a) Ductile materials based on nearly close-packed structures with moderately high symmetry (hexagonal and orthorhombic) and coordination numbers (6 + 6; 4 + 8).
   (b) Higher electronegativities than normal metals (except Pt and Au).
   (c) Increasing electrical resistivities with increasing temperature; values ~ 10–$100\,\mu\Omega\,\text{cm}$.
   (d) Temperature-independent diamagnetism. W. Klemm in the 1950s assigned Zn, Cd, Hg, Ga, In, Tl, Pb, and β-Sn to this category.
3. Semimetals (Sb, Bi, Te, Po):
   (a) Brittle materials based on network structures with low densities.
   (b) Larger electronegativities than most normal metals and metametals.
   (c) Decreasing electrical resistivities with increasing temperature; values $\sim m\Omega\,\text{cm}$.
   (d) Temperature-independent diamagnetism.

These classifications depend largely on their electrical properties and, therefore, are linked closely with their electronic states, which are depicted by their positions in the periodic table, shown in **Figure 1**. The electronic behavior of an element is related to the valence atomic orbitals, which are utilized for chemical bonding, and leads to three primary categories of metallic elements (*P*# = *Pettifor number*, see below):

1. A (*s* or *sd*): alkali and alkaline-earth elements, excluding Be and Mg (*P*#: 7–16), and A′ (*fsd*): rare-earth and actinide metals (*P*#: 17–48);
2. T (*ds*): early half of the transition metal series (*P*#: 49–∼60), and T′ (*dsp*): late half of the transition metal series (*P*#: ∼58–72); and
3. M (*sp* or *spd*): post-transition elements including Be, Mg, and Al (*P*#: 73–83), and M′ (*sp* or *spd*): semimetals (*P*#: 84–92).

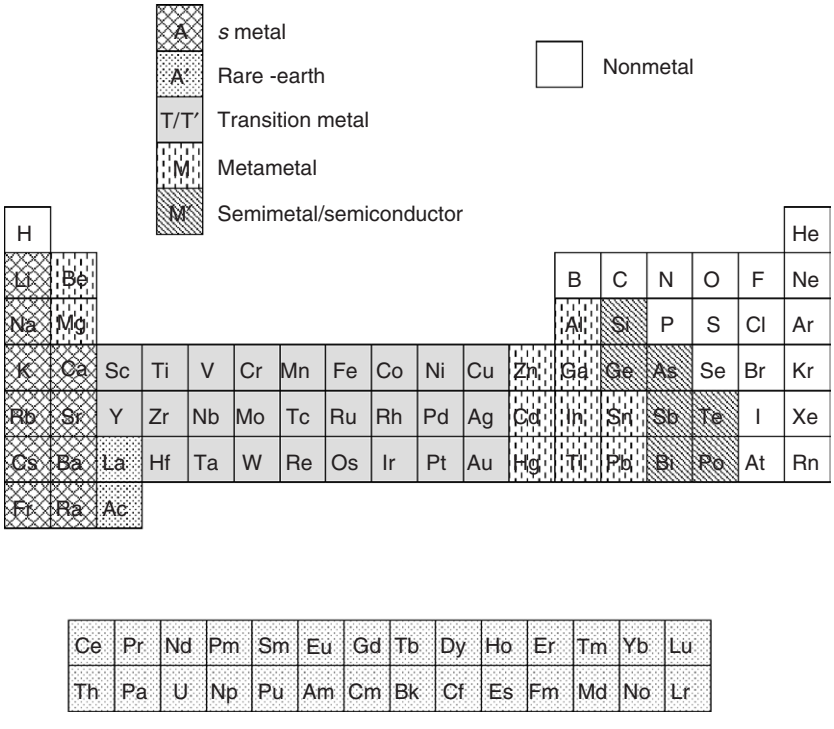


Figure 1 The periodic table of elements showing the different categories of metals based on the nature of the valence orbitals.

Periodic trends are especially important – whether the focus is on the main group (*sp*) or transition (*ds*) elements, the valence energy levels (E_{ns}, E_{np}, E_{nd}) drop linearly with the atomic number across a horizontal series. Among the main group elements, the valence *s* and *p* orbitals become less tightly bound as the valence period increases (except between 4*s* and 3*s* due to the presence of 3*d* orbitals). Furthermore, from left to right across a horizontal series, $E_{np} - E_{ns}$ increases so that *sp* hybridization is more important for lower nuclear charges. Within the transition series, the valence $(n+1)s$ orbital is less tightly bound than the *nd* orbitals, while the energies of 4*d* are lower than those of 3*d* orbitals. Thus, $E_{(n+1)s} - E_{nd}$ increases from 3*d* to 4*d* elements, but then drops going to 5*d* elements due to relativistic effects. These variations in valence state energies affect numerous properties, for example, the colors of Cu, Ag, Au; the onset of semiconducting behavior at high pressures in Ca and Sr; as well as the distortion of the hexagonal close-packed structures of Zn and Cd. To represent the general as well as subtle changes in atomic electronic structure, Pettifor developed a phenomenological numerical coordinate by placing the chemical elements in an order representative of their atomic properties and giving each element the number in the sequence – the Pettifor number (*P*#). This scheme has proved quite useful in separating different structural classes – in the periodic table, it nicely separates the different classifications of metals.

Furthermore, the term "intermetallic compound" is occasionally used interchangeably with "alloys," but they do differ from conventional metal alloys in a number of important ways. Conventional alloys consist basically of a disordered solid solution of two or more metallic elements with similar sizes and chemical characteristics. They do not have any particular chemical formula, and are best described as consisting of a base material to which certain percentages of other elements have been added. For example, type 304, a popular grade of stainless steel, has the composition Fe–18%Cr–8%Ni. An intermetallic compound, on the other hand, is based upon a definite atomic ratio, with a fixed or narrow range of chemical composition. **Figure 2** shows the Al–Ni phase diagram as an example of a binary system showing solid solutions and intermetallic compounds.

Elements such as B, Si, Ge, P, As, and Se, themselves nonmetals, as well as C, when combined with metallic elements are considered to form "intermetallic compounds" because their structures resemble those of other intermetallics, and they form with definite compositions. The elements H, N, O, S, and halogens are not considered to give intermetallic

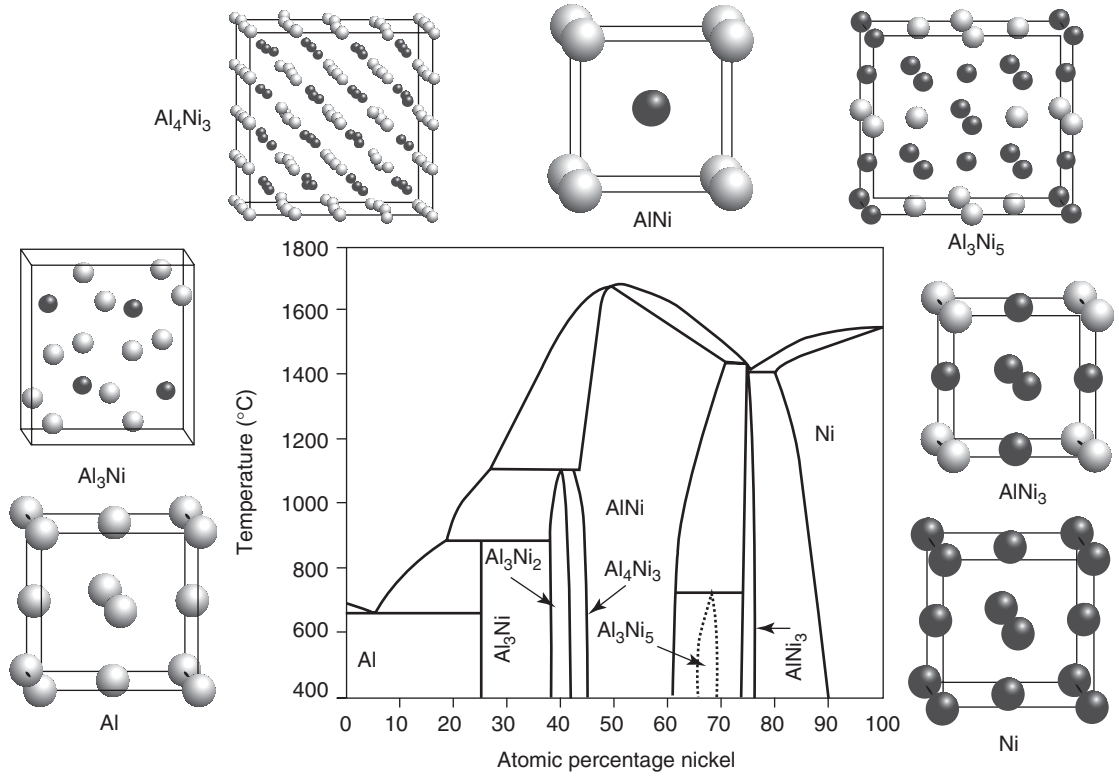


Figure 2 Al–Ni phase diagram and structures of various intermetallic compounds in this binary system. Light-colored spheres are Al; dark-colored spheres are Ni.

compounds, although numerous sulfides and some halides show a metallic character. The smaller elements, for example, H, B, C, N, and O can often occupy interstitial sites in intermetallic compounds, which greatly affect their mechanical, chemical, and physical properties.

The atoms in conventional alloys are linked through relatively weak interatomic interactions, often called "metallic bonds," with the atomic nuclei placed in a "gas" of valence electrons that are able to move relatively freely, which explains their ductility, electrical conductivity, and temperature-independent paramagnetism. In contrast, chemical bonding in intermetallic compounds may show significant polar covalent character, which can produce strong interatomic interactions, brittle mechanical behavior, transport properties ranging from semiconducting to superconducting, and various magnetic interactions leading to complex magnetic structures. However, in intermetallic compounds whose mechanical and electronic properties resemble those of alloys, the individual elements tend to occupy different positions within the crystal lattice. This condition, which is referred to as "ordering," often leads to abrupt changes in the physical properties of the material.

Classifying Electronic States in Intermetallic Compounds

One useful tool for assigning compounds into different categories is a van Arkel–Ketelaar triangle, which maps compounds according to two related chemical coordinates using an electronegativity scale based upon the average one-electron valence shell energy of the ground-state free atom, called the configuration energy (CE). In this map, shown in **Figure 3**, the horizontal coordinate is the average CE of the constituents, while the vertical coordinate is the difference, so that metallic substances lie to the left. The transition from metallic to nonmetallic behavior involves either increasing electronegativity in general, or increasing electronegativity difference between constituents. Since CE represents the average energy of a valence electron, CE correlates with the separation between the energies of the valence *s* and *p* orbitals, and so influences the way atoms will interact with one another. Low CE values give small $E_{np} - E_{ns}$ values and broad, frequently overlapping energy bands. In many cases, the electronic states for these systems behave like nearly free electron states. Moreover, valence *d* orbitals are important for the properties of transition elements and valence *f* orbitals for the rare-earth and actinides – changes in the band filling of these orbitals affect the structure, cohesive energies, thermodynamic properties, and magnetic character. Unlike valence *s* and *p* orbitals, the valence *d* and *f* orbitals give narrower energy bands, and have more atomic-like character in compounds. The 4*f* orbitals in rare-earth elements and their compounds are generally localized, while the 5*f* orbitals in actinides have a greater role in chemical bonding.

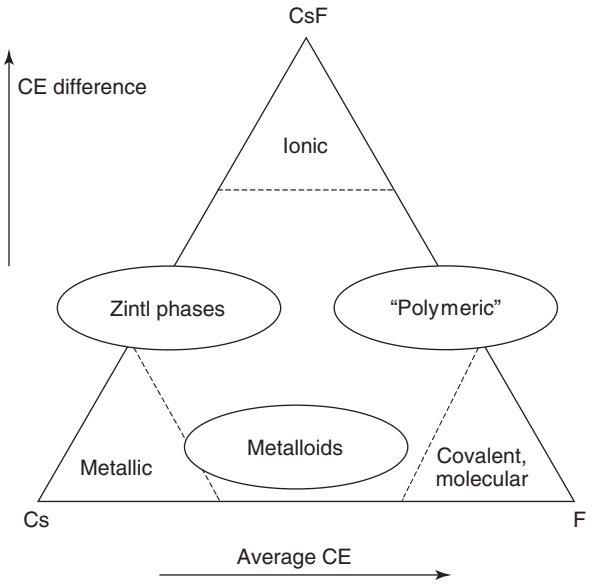


Figure 3 The van Arkel–Ketelaar triangle based on configuration energies of the atomic components.

This qualitative picture implies a variation in the description of states from a nearly free electron (NFE) model to a tight-binding (TB) model. In NFE, the electronic states are treated as linear combinations of plane waves (the eigenstates for a free electron gas). The arrangement of positively charged atomic nuclei creates a potential that perturbs the free electron states. The extension of this approach is toward pseudopotentials, which makes the valence states orthogonal to the core atomic states. In TB, the electronic states are treated as linear combinations of atomic orbitals (the one-electron eigenstates of free atoms). Pseudopotential methods as well as TB methods have been used to study metals and semiconductors – the TB methods lead to useful chemical insights by virtue of starting from atoms, and have provided valuable understanding even for systems in which NFE would be entirely appropriate.

To calculate the electronic densities of states (DOS), multi-electron terms in the total electronic energy may be treated either in the local density approximation (LDA) or in the local spin density approximation (LSDA) and the one-electron terms can be treated scalar relativistically. For the heavier elements, spin–orbit

coupling should also be considered. For many ground-state properties (structure, cohesive energy, magnetism, and electronic transport), the LDA is quite effective. Numerous recipes exist for calculating the exchange and correlation terms. An especially valuable analysis of the DOS is through the crystal orbital Hamilton population (COHP) analysis, which allows states to be identified according to the bonding, nonbonding, or antibonding character of the interatomic interactions (COHP values are, respectively, less than, equal to, or greater than zero).

Experimental determination of electronic states can be carried out by photoelectron spectroscopy, which is a surface-sensitive technique. de Haas–van Alphen effect gives information about the Fermi surface in reciprocal space, but demands quite pure samples. Hall effect measurements provide information about the charge carriers. Other methods, such as specific heat and magnetic susceptibility studies on metals, can give values of the DOS at the Fermi level.

Electronic Interactions between Different Metallic Types

An examination of Pearson's *Handbook of Intermetallic Phases* shows that the predominant classes of intermetallic compounds involve the following combinations of metallic elements using the designations in **Figure 1**: (1) A′–T′; (2) A–M and A–M′; (3) T–T′; (4) (T, T′)–(M, M′); and (5) (A, A′)–(T,T′)–(M,M′). Very few examples exist for (A,A′)–(A,A′), (M,M′)–(M,M′), (A,A′)–T, T–T, and T′–T′. This observation guides our emphasis to the important electronic interactions: (1) *d–d*; (2) *d–sp*; and (3) *sp–sp*.

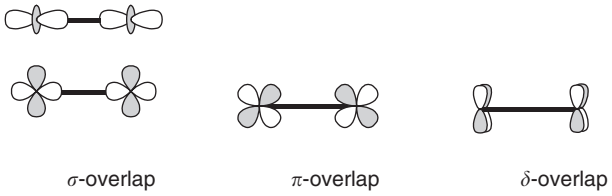
The two most important factors influencing the electronic DOS are the relative energies of the atomic orbitals of the different constituents and the strengths of the interatomic orbital interactions leading to bandwidths, that is, band dispersion. Energy gaps, pseudogaps, and peaks in the DOS arise from these two fundamental characteristics. Since the majority of intermetallic compounds show metallic properties, the "perfect" screening of conduction electrons means that there is essentially no charge transfer between different elements. Therefore, the "common band" model is most appropriate for analyzing the DOS. Additional factors include composition and the coordination environments surrounding each element in a structure.

d–d Interactions

Intermetallic compounds involving just transition metals are technologically important due to their mechanical, chemical, and physical properties. Some of the strongest chemical bonds in condensed phases are found among these compounds, as measured by their cohesive energies. The electronic states of transition metals exhibit a partially occupied, narrow valence *d* band that lies near the bottom of a broad valence *s* or *sp* band. Since the valence *d* electrons remain relatively tightly bound to their parent atoms, the appropriate method to treat the *d–d* electronic interactions is the TB model. When different transition metals combine to give an intermetallic compound, the primary effect is the way the different *d* orbitals interact – the valence *s* or *sp* band remains broad, and contributes to the chemistry and physics of these materials by adjusting the extent of filling of the *d* bands. The *d–d* orbital interactions break down into σ, π, and δ types of orbital overlap, with σ overlap the greatest and δ overlap the lowest (**Scheme 1**).

The *d–d* π interactions are important for the structural variation of the transition metals in the 4*d* and 5*d* periods (h.c.p.–b.c.c.–h.c.p.–f.c.c.) (b.c.c. – body-centered cubic; f.c.c. – face-centered cubic; h.c.p. – hexagonal close-packed). These overlaps affect the shape of the *d*-band DOS – see **Figure 4**, which compares the DOS curves for a b.c.c. and f.c.c. 4*d* transition metal. The local environments in the two structures both have cubic symmetry, and the local *d* orbitals are split into e_g and t_{2g} levels. The local coordination environment of the b.c.c. structure creates a region of low, nonzero DOS values close to six valence *s* and *d* electrons between two regions of high DOS values. This region of low DOS is called "quasigap" or "pseudogap." COHP analysis of the DOS shows that *d–d* bonding states are below the pseudogap while antibonding states lie above. On either side of six valence *s* and *d* electrons, close packed (h.c.p., f.c.c.) structures are preferred. Trends in effective nuclear charge and occupation of bonding versus antibonding states help explain the trend that across a period of transition metals, the *d* bandwidth slightly increases and then sharply decreases while the *d* band center steadily decreases. These trends have profound influences on the states of transition metal intermetallics.

When two different transition metals form an intermetallic compound, the shape of the DOS will



Scheme 1

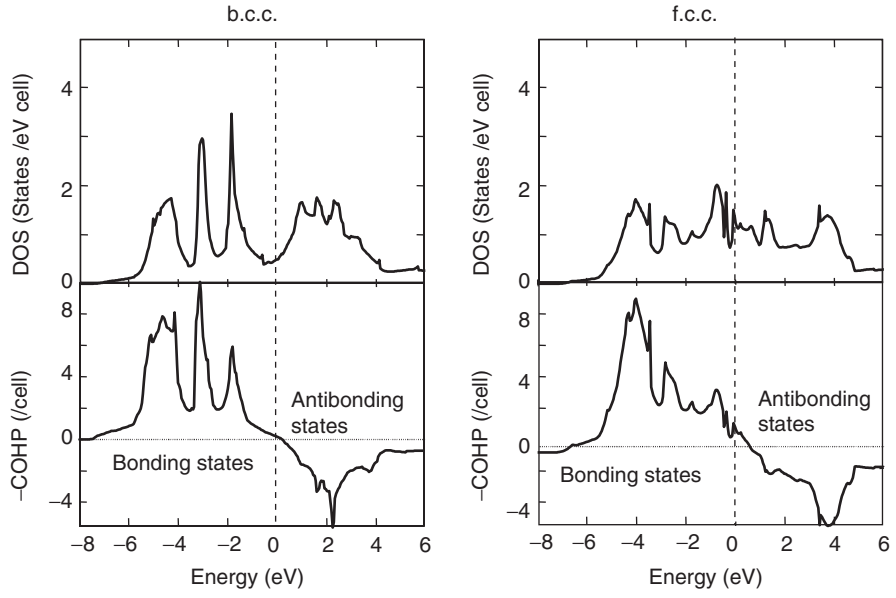
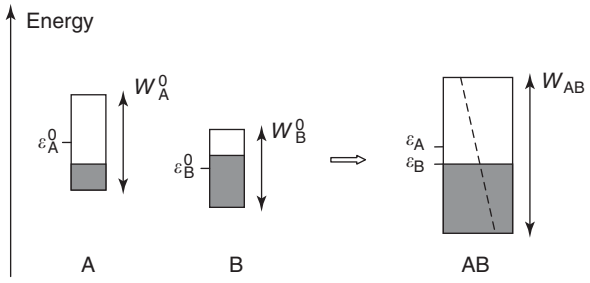


Figure 4 DOS and nearest-neighbor COHP analysis for the b.c.c. and f.c.c. structures of a typical 4*d* transition metal. The energy zero corresponds to six valence *s* and *d* electrons. In the COHP curves, metal–metal bonding states occur for –COHP > 0; antibonding states for –COHP < 0.



Scheme 2

depend on both the structure and the electronic characteristics of the component elements. In the case of disordered alloys AB, the "common band" model is appropriate for understanding heats of formation and trends in cohesive energies with respect to metals A and B (see **Scheme 2**). In this model, the valence *d* band of each transition metal is characterized by an energetic center (ε_M^0; M = A, B), a bandwidth (W_M^0; M = A, B), and band filling (N_M electrons; $0 \leqslant N_M \leqslant 10$; A, B; indicated by the shaded regions in **Scheme 2**). The resulting *d* band for the disordered alloy AB has a bandwidth given by $W_{AB} = \bar{W}[1 + 3(\Delta\varepsilon/\bar{W})^2]^{1/2}$ ($\bar{W} = (W_A^{AB} + W_B^{AB})/2$; $\Delta\varepsilon = \varepsilon_A - \varepsilon_B \langle \varepsilon_A^0 - \varepsilon_B^0 \rangle$, band filling of $\bar{N} = (N_A + N_B)/2$. To construct this *d* band for the alloy AB, there is "renormalization" of the average band energies ($\varepsilon_M^0 \rightarrow \varepsilon_M$; M = A, B) to keep local charge neutrality (i.e., no effective charge transfer) and the bandwidths ($W_M^0 \rightarrow W_M^{AB}$; M = A, B) due to changes in atomic volumes from the element to the alloy. In this *d* band, states below the band center $\bar{\varepsilon}$ are mostly from metal B, while those above $\bar{\varepsilon}$ are mostly from metal A (as indicated by the dashed line for the AB bnad). Therefore, electrons flow from A to B because B is more electronegative than A ($\varepsilon_A^0 < \varepsilon_A^0$), but also flow back from B to A through orbital overlap in alloy AB.

The common band model predicts that the most stable intermetallic compounds and alloys would arise from combinations of transition metals at opposite ends of the transition metal series and that the highest cohesive energies exist for AB systems with $\bar{N} : 5$. Specifically, the cohesive energy for AB, $U^{AB} : -k\bar{N}(10 - \bar{N})W_{AB} + cW_{AB}^2$ (*k*, *c* are positive constants) has an attractive part related to band filling and a repulsive part related to the number of nearest neighbors, which is proportional to the square of the bandwidth. The expression for the heat of formation of AB is more complicated, but takes the form $\Delta H_f^{AB} : f(\bar{N}(10 - \bar{N}))\,(\Delta N)^2$ where $\Delta N = N_B - N_A$. Negative heats of formation occur for $\bar{N}$ between 3.33 and 6.67 electrons per atom. For an intermetallic compound AB, with an ordering of A and B atoms, this model is appropriate for adjacent elements (small $\Delta\varepsilon$ values), but is less appropriate as this energy difference increases. For larger $\Delta\varepsilon$ values, a pseudogap can open in the DOS, as seen in **Figure 5** for TiFe, which adopts the CsCl-structure type. The DOS curves of Ti and nonmagnetic Fe show features as expected – a lower band center and a narrower

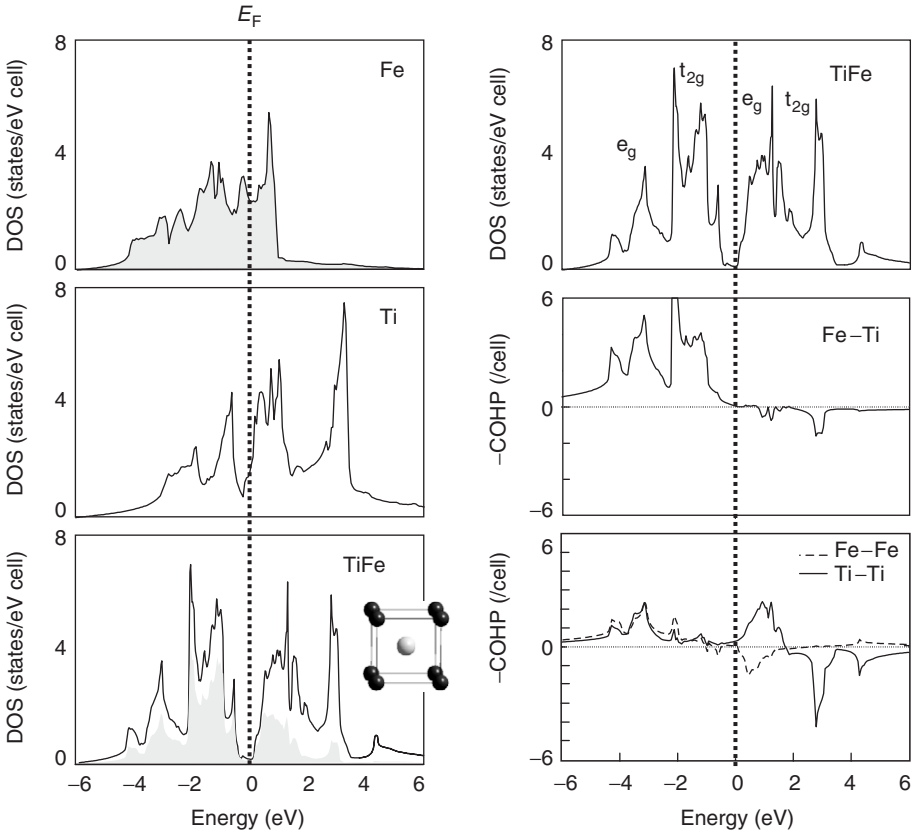


Figure 5 Left: DOS curves of nonmagnetic Fe, Ti, and TiFe. The shaded regions in the DOS of TiFe are Fe states. Right: Assignment of the peaks in the DOS of TiFe into local e_g and t_{2g} levels as well as COHP analyses for Ti–Fe, Ti–Ti, and Fe–Fe near-neighbor interactions.

bandwidth for Fe. In TiFe, the *d* bandwidth is wider than that for either element, but the shape shows a pseudogap at the center of the *d*-band and the Fermi level falls in this pseudogap. States below this pseudogap are mostly Fe-orbitals, while those above are mostly Ti-orbitals. Nevertheless, TiFe is not held together by "ionic" forces because the perfect screening of the valence electrons negates any charge transfer – there is a "renormalization" of the energies of each transition metal. The DOS shows aspects of the local coordination, but the principle factor creating the pseudogap in TiFe is the energy difference between the band centers of Ti and Fe. COHP analysis of the DOS of TiFe shows that states below the Fermi level are Ti–Fe bonding; those above are Ti–Fe antibonding. The homonuclear Fe–Fe and Ti–Ti interactions are weaker (longer distances) and bonding below the Fermi level, but the crossover from bonding to antibonding interactions does not occur at the pseudogap. Therefore, another aspect of the pseudogap in intermetallic compounds is the creation of bonding states below the gap and antibonding states above. Therefore, many stable intermetallic compounds show low DOS values at the Fermi level, which can be accounted for by certain metal–metal interactions through a COHP analysis of the DOS.

Trends in the features of DOS curves depend on the types of interactions, the relative magnitudes of orbital overlaps and atomic electronegativities. **Figure 6** shows the trends in DOS and COHP curves for TiRu, TiRh, and TiPd, which follows the structural trend b.c.c.–f.c.c.–h.c.p., but with ordered arrangements of metals. A pseudogap is noticeable in all cases for six states per formula, although it is less clear in the TiRh and TiPd cases. TiRu adopts the CsCl structure with eight nearest-neighbor Ti–Ru interactions and 12 second nearest-neighbor homonuclear (Ti–Ti and Ru–Ru) contacts. In TiRh and TiPd, there are six heteroatomic (Ti–Rh or Ti–Pd) and six homoatomic (Ti–Ti, Rh–Rh, Pd–Pd) nearest-neighbor contacts, which interferes with the pseudogap. Additionally, the bandwidths become narrower as the 4*d* metal becomes more electronegative. Again, there is no charge transfer between Ti and the 4*d* metal, in accordance with the common band model.

If the DOS value at the Fermi level is sufficiently high, then the structure is susceptible to an electronic distortion, which will perturb the electronic states

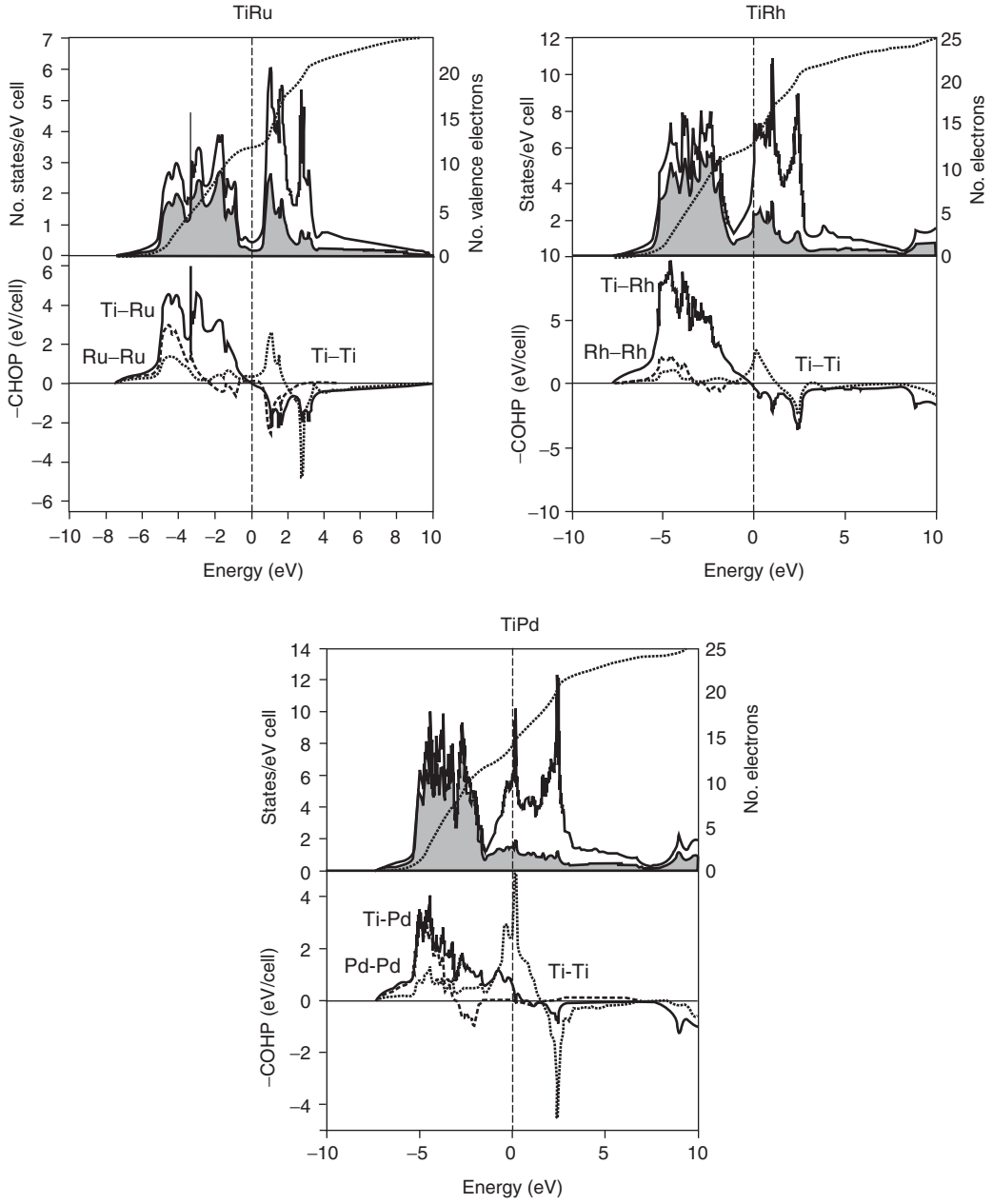


Figure 6 DOS and COHP analyses of the 3*d*–4*d* intermetallic compounds, TiRu, TiRh, and TiPd. Shaded regions of the DOS are states from the electronegative metal (Ru, Rh, Pd). Fermi levels are indicated by the dashed lines.

near the Fermi level and lower the total electronic energy. This effect can happen in one of two ways: (1) a structural distortion, as in VIr (which distorts away from the cubic CsCl-structure type); or (2) ferromagnetism. For ferromagnetism to occur, the local exchange energy must also be sufficiently large, and this occurs for some of the 3*d* elements, which have the smallest spatial extent. The Stoner criterion states that ferromagnetism is preferred when $IN(E_F) > 1$, where I is the intra-atomic exchange energy (chemists will think about the so-called "pairing energy" – the energy penalty in putting two electrons into the same atomic orbital) and $N(E_F)$ is the value of the DOS at the Fermi level. Since $N(E)$ is typically inversely proportional to the bandwidth W, when either intra-atomic or inter-atomic interactions lead to narrow bands, and the local exchange is sufficiently large, ferromagnetism frequently occurs – the 3*d* elements Fe, Co, and Ni as well as compounds involving Mn are especially susceptible.

d–(*s*)*p* Interactions

Intermetallic compounds between transition metals and main group metals display interesting magnetic, superconducting, mechanical, and structural properties. In most cases, the electronic states near the Fermi level of these compounds are dominated by the *d* orbitals of the transition metal component. Their electronic structures arise from two dominant factors: (1) the expansion of the transition metal lattice due to the insertion of the main group (*sp*) element; and (2) the interaction between the valence *d* bands on the transition metal with the *sp* bands of the main group element. Three categories occur (see **Figure 7**): (1) where the *d* levels lie much deeper in energy than the main group atom (e.g., PdLi); (2) where the opposite is true (e.g., "PdF"); and (3) the intermediate case where the two are of similar energy (e.g., PdB). In the first two cases, there is a rather weak interaction between the two sets of orbitals. The transition metal *d* band is narrower than in the pure metal, as a result of stretching the metal lattice to accommodate the nonmetal. These two cases do differ, however: (a) PdF is quite "ionic," involving charge transfer from metal *d* states to the low-lying fluorine-located bands – note that the Fermi level drops from the top of the *d* band; (b) PdLi utilizes the common band picture, but the Fermi level rises above the top of the Pd 4*d* band – note the significant contribution from Li valence states in the region of the Pd 4*d* band. A similar example is the case of CsAu, which is an unusual example of a red, transparent, semiconducting intermetallic compound – its photoelectron spectrum shows a band gap of 2.6 eV. The case of PdB is most interesting because strong mixing occurs between the two sets of orbitals. These bands can be separated into bonding, nonbonding, and antibonding regions between the metal and nonmetal. Changes to the DOS will occur on composition – the strongest interactions will typically involve transition metal–main group metal, but the DOS will depend on the concentration of different orbitals. In transition metal-rich compounds, the metal *d* band will dominate and be broadened by T–T interactions; in main group metal-rich compounds, the metal *sp* bands may now become apparent.

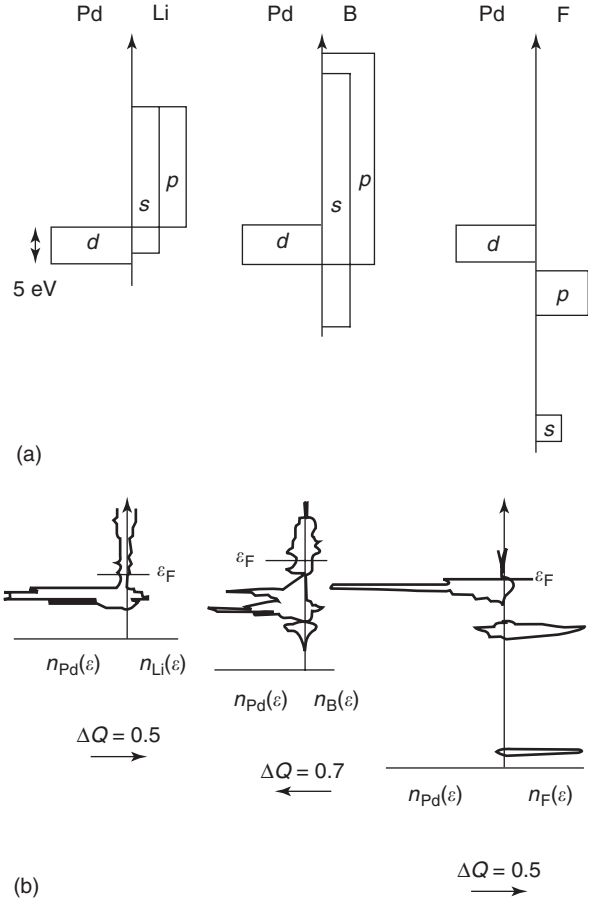


Figure 7 Schematic (a) and calculated (b) partial densities of states for the atomic components in the three examples: PdLi, PdB, and PdF. (Reprinted with permission from Gelatt CD, Jr., Williams AR, and Moruzzi VL (1983) Theory of bonding of transition metals to nontransition metals. *Physical Review B* 27: 2005–2013; © American Physical Society.)

One factor governing the atomic arrangements in intermetallic compounds is the relative strengths and concentrations of heteroatomic (between different elements) and homoatomic (between identical elements) interactions. As the molar concentration of the transition metal increases, the *d* band will become wider and may adjust its position depending on the relative electronegativity of the main group metal. For example, in the Ni–Al system, one can see the influence of composition on the DOS in **Figure 8** for three different examples. As there is an increased dilution of Ni in Al, the Ni 3*d* bands become narrower and pull away below the Fermi level.

Main group elements show extensive structural diversity as the number of valence *s* and *p* electrons changes. For less than the half-filled band situation, for example, among group 13 elements like Al and Ga, clusters based on octahedra, square pyramids, and tetrahedra are common. Such clusters can be used to visualize structures through condensation of vertices, edges, or faces and create new understanding of electronic and bonding characteristics, as in transition metal trialuminides, MAl_3 (M = rare-earth element, Ti, Zr, Hf, V, Nb, Ta). As the number of valence electrons increases in this family, the network of Al-based polyhedra changes from octahedra to square pyramids. The DOS shows a pseudogap at the Fermi level (see **Figure 9** for $TaAl_3$ as an example), which is largely influenced by the bonding states of the Al polyhedra. Nevertheless, M–Al orbital

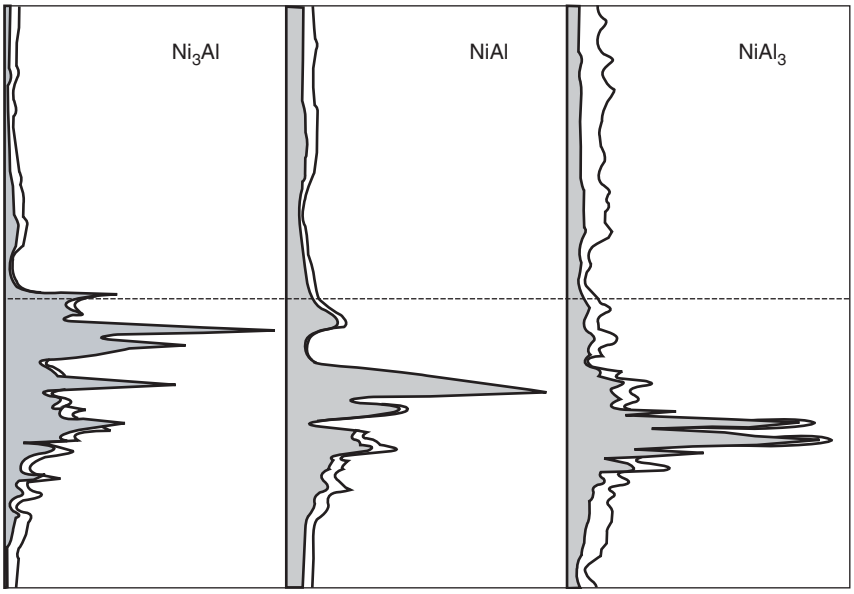


Figure 8 DOS curves for Ni_3Al, NiAl, and $NiAl_3$ presented with the Fermi level set as the energy reference. The shaded regions indicate the Ni 3*d* and 4*s* components to the DOS and Ni 3*d* states are labeled.

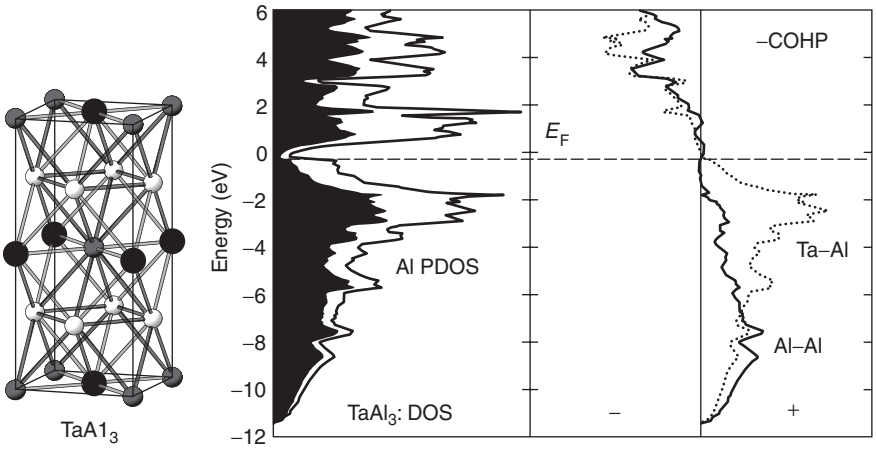


Figure 9 Structure, DOS, and COHP analysis for $TaAl_3$. Large, dark spheres are Ta; small shaded spheres are Al. The shaded regions in the DOS are Al states, and the Fermi level is marked by a dashed line. In the COHP curves, the solid line shows Al–Al interactions; the dotted line shows Ta–Al interactions.

interactions also split the 5 M *d* levels into two sets (e_g below t_{2g}), which also contributes to the observed pseudogaps in the DOS curves as well as the preferred structures and electronic properties. In fact, such compounds are called *d–p* covalent intermetallics – COHP analysis of the DOS in $TaAl_3$ shows maximum bonding character for both Al–Al and Ta–Al interactions.

Pseudogaps occur in intermetallics for reasons based on structure and atomic electronegativities. Another important growing class of compounds is the set of half-metallic ferromagnets, of which the "semi-Heusler alloys" represent significant examples ("semi-Heusler" alloys, ABX, consist of cubic closest packed X atoms with B atoms in all octahedral holes and A atoms ordered in one-half of the tetrahedral holes of the close packed array of X "Heusler" alloys, A_2BX, have all tetrahedral holes filled). These compounds are cubic intermetallic phases combining main group metals (Sn and Sb) with two different transition metals (one early T – Ti, Zr, Hf, V, Nb; one late T′ – Fe, Co, Ir, Ni, Pt). Their electronic properties, and therefore, their electronic states, depend strongly on the number of valence electrons: 18-electron examples such as TiCoSb are diamagnetic semiconductors, 17-electron examples are metallic and

Pauli paramagnetic, whereas 19- or 22-electron examples show half-metallic ferromagnetic behavior. Half-metallic ferromagnetism originates when the nonmagnetic electronic structure is susceptible to splitting of the majority and minority spin densities of states, and the Fermi level falls in an energy gap for either spin, as seen in the LSDA DOS for VCoSb, a 19-electron example, in **Figure 10** (VCoSb is a weak ferromagnet because there is an energy gap at nine states for both spin states; a half-metallic ferromagnet shows an energy gap just in the minority spin band, as seen in NiMnSb). The nonmagnetic DOS at the Fermi level for 19 electrons (VCoSb) is large, and is susceptible to ferromagnetism within the Stoner criterion. The lowest energy valence orbitals are from Sb, then from Co, finally from V since the atomic electronegativities increase from V to Co to Sb. The gap in the DOS at 18 valence electrons originates from these valence orbitals, the symmetry of the local coordination environments and metal–metal interactions: (1) Sb has four valence orbitals and the five Co 3*d* orbitals make nine low-lying orbitals to be filled for 18 electrons; (2) the tetrahedral coordination of Sb around Co and octahedral coordination around V keep the Co 3*d* levels low and split by a small amount; (3) V–Co orbital interactions help to keep a band gap at 18 electrons, which accounts for the properties of TiCoSb. In these examples, the gap opens through *d*–*d* interactions between the two different transition metals – note the strength of early-late *d*–*d* interactions, just as mentioned in an earlier section.

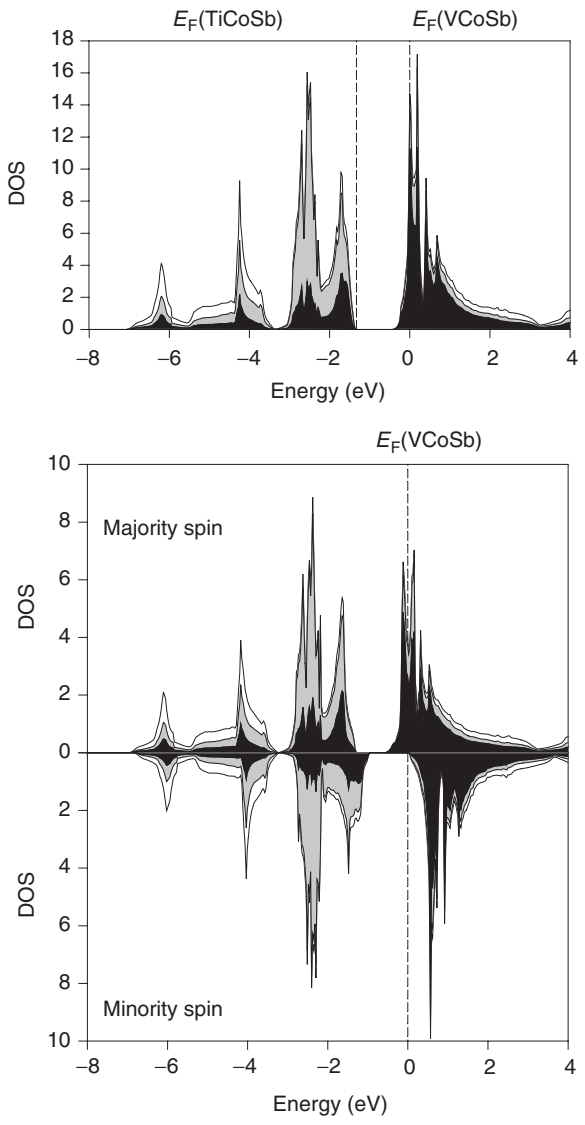


Figure 10 Nonmagnetic (top) and LSDA (bottom) DOS for VCoSb, a semi-Heusler alloy. The dashed line indicates the Fermi level – note the bandgap in the minority spin states. Black regions are V orbitals; gray regions are Co orbitals; white regions are Sb orbitals.

sp–sp Interactions

Due to the widely different electronegativities of these elements, theoretical determinations of their electronic states range from NFE through pseudopotentials to TB models. There are three principle types of interactions (see **Figure 1**): A–A, A–M, and M–M. A–A interactions are best modeled by the NFE because the active metals have the lowest $E_{np} - E_{ns}$ values and diffuse valence atomic orbitals. The more electronegative main group metals (M) require pseudopotential or TB models because $E_{np} - E_{ns}$ increases and the valence atomic orbitals become contracted. In intermetallic structures, coordination environments surrounding A-type elements involve several atoms and often resemble densely packed structures, whereas those around M-type elements show lower coordination numbers and some directional bonding character. As the electronegativity difference between main group elements increases, the DOS develops a structure, that is, peaks and pseudogaps (or gaps), as seen for the three hypothetical compounds: "NaMg," "NaAl," and "MgAl" in **Figure 11**. A–M compounds constitute an important class of compounds called "polar intermetallics," in which orbital interactions within the network of electronegative M atoms typically maximizes bonding interactions. If the local structure around the M atoms obey classical valence rules, then these compounds are called "Zintl phases." Subtle electronic effects can arise from valence *d* orbitals: for alkaline earth elements, there can be contributions of $(n-1)d$ orbitals to *ns* bands, especially under pressure, whereas for Cu, Ag, Au, Zn, Cd, and Hg, $(n-1)d$ orbitals can hybridize with the valence *ns* and *np* bands.

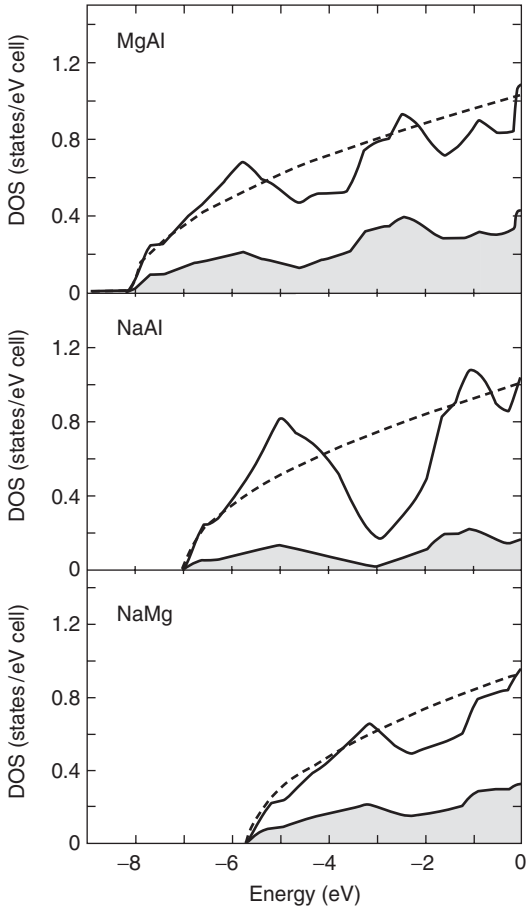


Figure 11 DOS curves for NaMg, NaAl, and MgAl, and the shaded regions are states from the more electropositive element. The dashed line shows a free-electron type DOS.

Hume–Rothery "electron compounds" are formed between noble metals and *sp* metals – in these systems, the noble metal *d* band is filled, so that the dominant orbital interactions will be *sp*–*sp* interactions. There are five phases that Hume–Rothery identified as controlled by electron count: (1) h.c.p. (3 subcategories: ζ ($c/a = 1.633$), ε ($c/a = 1.55$–1.58), and η ($c/a = 1.77$–1.78)); (2) f.c.c.; (3) b.c.c./CsCl-type (β-brass); (4) β-manganese; and (5) γ-brass. TB models account for the structural variations very well, which attests to the importance of orbital interactions influencing these structures. Analysis of the DOS and band structure for CuZn reveal, however, that optimized metal–metal bonding is not the driving force here (**Figure 12**). As Pettifor and co-workers have shown, the structural trend is driven by van Hove singularities in the DOS (places where the band structure has zero slope, which occur at the bottom or top of energy gaps at Brillouin zone boundaries). The band structure shows that the valence *d* bands are well below the Fermi level, but, nevertheless, hybridize with the NFE-like *sp* bands. Numerous quasicrystalline and approximant systems (Li–Cu–Al, Mg–Zn–Al) have been interpreted as similar "electron compounds," but the NFE gives little structural chemical information.

Zintl phases involve combinations of electropositive and electronegative main group metals, for example, LiAl, in which the electronegative element forms a structure based on the octet rule if one assumes that the electropositive element donates all of its valence electrons. The octet rule demands that

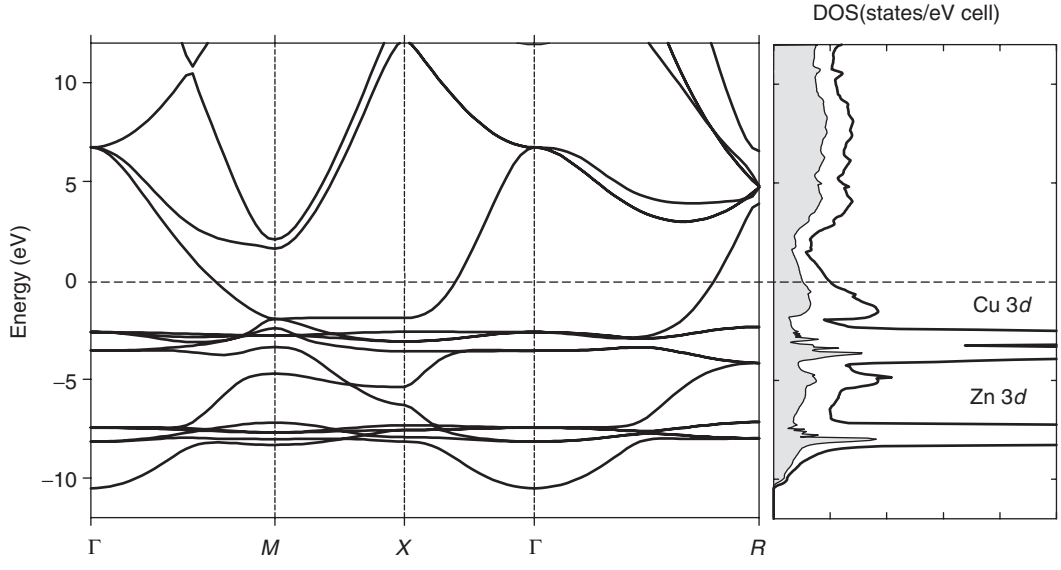


Figure 12 Energy bands and DOS for CuZn (β-brass structure). Shaded region in the DOS comes from valence *s* and *p* orbitals of Cu.

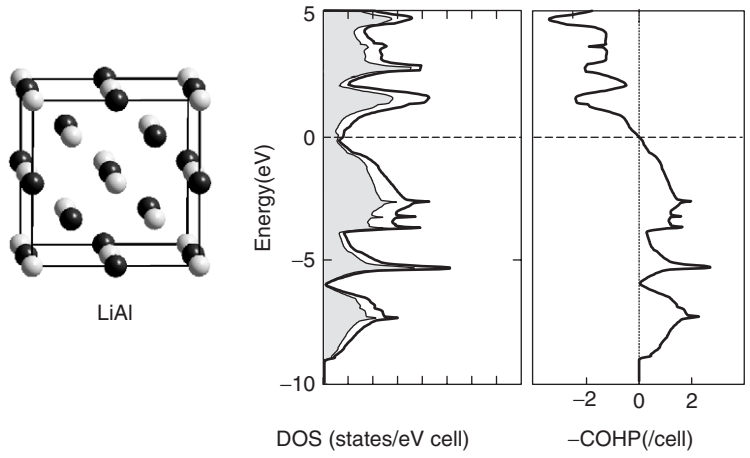


Figure 13 Structure, DOS, and Al–Al COHP analysis for LiAl. Shaded region is Al orbitals.

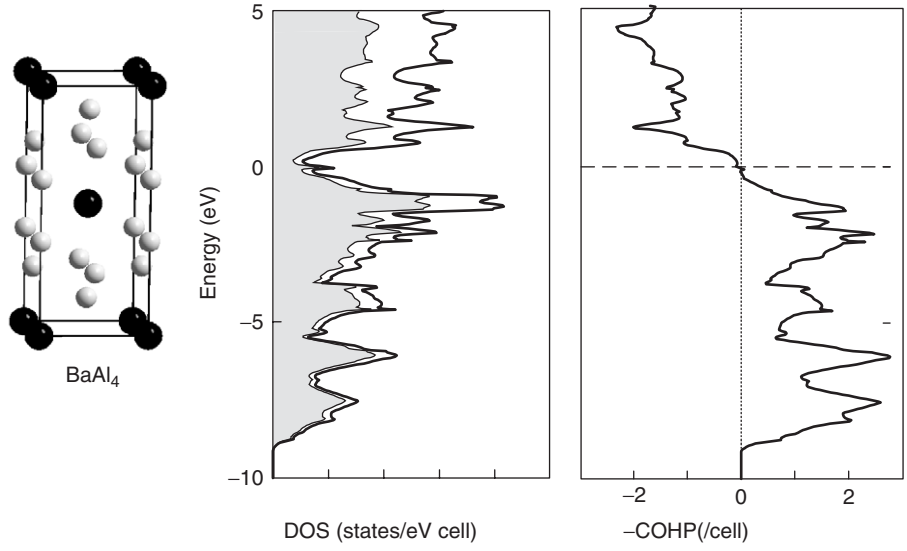


Figure 14 Structure, DOS, and Al–Al COHP analysis for $BaAl_4$. Shaded region is Al orbitals.

eight bonding or nonbonding valence electrons surround main group (*sp*) elements in stable structures. Thus, the structure of LiAl has a diamond network of four-bonded Al atoms, which is formulated as "Al^-," isoelectronic with Si. The DOS of Zintl phases show either a gap or pseudogap at the Fermi level with bonding states within the electronegative metal network fully occupied (**Figure 13**). So, "classical" Zintl phases are defined as follows: They are compounds that (1) contain an alkali or alkaline earth metal and an electronegative main group element that is a metal, semimetal, or small-gap semiconductor; (2) are electronically balanced or close-shell compounds, that is, the number of valence electrons provided by the constituents equals the number of electrons needed for two-center, two-electron covalent bonds in the structure; (3) are semiconductors or poor conductors; (4) are either diamagnetic or show a very weak temperature-independent paramagnetism; and (5) are typically brittle.

Among the *sp* class of intermetallic compounds, the Hume–Rothery phases behave as normal metals while Zintl phases are typically semiconductors. "Polar intermetallics" represent an intermediate class of compounds, formed between active metals with electronegative metals, but the DOS shows a pseudogap near the Fermi level. COHP analysis of the DOS indicates optimized bonding in the electronegative component and the active metal behaves like a cation, but typically does not donate its full complement of valence electrons. The structure of the electronegative component, however, cannot be explained by the octet rule – there is substantial delocalized bonding. The most prolific example of this

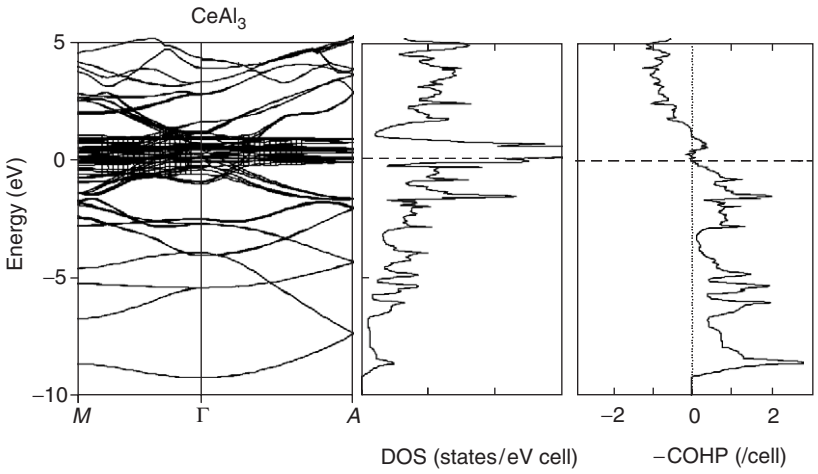


Figure 15 Band structure, DOS, and Al–Al COHP analysis for $CeAl_3$.

structural class is the $BaAl_4$-type, whose structure is adopted by hundreds of examples involving transition metals as well (**Figure 14**). There are numerous subcategories based on how different elements are arranged: $ThCr_2Si_2$, $CaBe_2Ge_2$, $BaNiSn_3$ are three common examples.

In polar intermetallics where the active metal is among the rare-earth elements, the predominant interatomic interactions involve 6*s* and 5*d* valence orbitals of the rare-earth with the valence *s* and *p* orbitals of the main group element. If there is significant hybridization between the valence 4*f* orbitals and the *sp* orbitals of the main group metals, then interesting properties may arise, for example, "heavy fermion" behavior. Valence fluctuations in the rare-earth element are also important for this situation, while a COHP analysis of the *sp*–*sp* interactions shows nearly maximum bonding within the main group framework. This is frequently observed for Ce, but is also possible with Yb, with an *f* band that mixes with states near the crossover from bonding to antibonding characteristics (**Figure 15**).

Summary

Describing the electronic states of intermetallic compounds requires an understanding of the distinct electronic structures of the various types of metallic elements. These elements combine in many different ways, and the nature of these interactions can be addressed using different quantum mechanical approximations, varying from the NFE model to the TB theory. Fundamental factors controlling the electronic states include (1) composition, (2) relative electronegativities of elements, which control the relative valence orbital energies, and (3) coordination environments (structure and atomic arrangements). Depending on the nature of these interactions, conducting properties can vary from metallic to semiconducting while magnetic properties can range from itinerant ferromagnetism to diamagnetic. Valence fluctuations and heavy fermion behavior are possible for rare-earth compounds, and 5*f* orbitals contribute to the chemical bonding interactions in actinide compounds.

See also: Density-Functional Theory; Electronic Structure (Theory): Atoms; Electronic Structure Calculations: Plane-Wave Methods; Electronic Structure Calculations: Scattering Methods; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Pseudopotential Method; Quasicrystals, Electronic Structure of; Recursive Methods for Electronic States; Tight-Binding Method in Electronic Structure.

PACS: 61.44.Br; 61.50.Lt; 61.66.Bi; 61.66.Dk; 61.66.Fn; 75.15.−m; 71.20.Be; 71.20.Be; 71.20.Dg; 71.20.Eh; 71.20.Gj; 71.20.Lp

Further Reading

Burdett JK (1995) *Chemical Bonding in Solids*. New York: Oxford University Press.

Pettifor D (1995) *Bonding and Structure of Molecules and Solids*. New York: Oxford University Press.

Westbrook JH and Fleischer RL (1995) *Intermetallic Compounds: Principle and Practice*. New York: Wiley.

Ion Beam Analysis

T L Alford, Arizona State University, Tempe, AZ, USA
D Adams, University of Western Cape, Bellville, South Africa
J W Mayer, Arizona State University, Tempe, AZ, USA

Introduction

Originally the advent of microelectronics gave rise to the need for rapid and accurate analysis of thin film structures. Generally, these planar structures were formed in silicon or germanium by either energetic ion implantation of dopants to create electrically active regions and/or by thermal reactions between thin metal films which were deposited onto them. Ion-beam modification (e.g., ion implantation) was a new technique in the early sixties, and interactions between metal films and silicon required analysis. For example, the number of ions implanted per square centimeter (ion dose) and the thicknesses of metal layers required careful control to meet the specifications of integrated circuit technology. Rutherford backscattering spectrometry (RBS) and MeV ion-beam analyses were developed in response to the needs of the microelectronics and forensic disciplines. In turn, integrated circuit technology provided the electronic sophistication used in the instrumentation in ion-beam analysis. It was a synergistic development of analytical tools and the fabrication of integrated circuits.

Rutherford backscattering analysis is the measurement of the energies of ions scattered back from the surface and near the surface, the outer micron, of a sample. **Figure 1** shows a depiction of a general ion-scattering chamber where samples are typically analyzed. Monoenergetic helium ions or protons with energies ranging from 1 to 5 MeV are used to analyze thin film materials ranging from integrated circuits to biomedical materials. Only moderate vacuum levels ($\sim 10^{-6}$ torr) are required so that sample exchange is rapid; it allows the analysis to be done in a relatively brief period of time (~ 20 min or less). Under unique conditions, inorganic materials can be analyzed with minimum damage to the sample. In this case, RBS is considered nondestructive. This is in contrast to surface sensitive techniques such as Auger electron spectroscopy (AES), where surface erosion by sputtering is required for depth analysis. Another salient feature of RBS is that the scattering cross sections are well known, and this results in an analytical technique that is extremely quantitative. The scattering cross section is a measure of the probability of a detectable

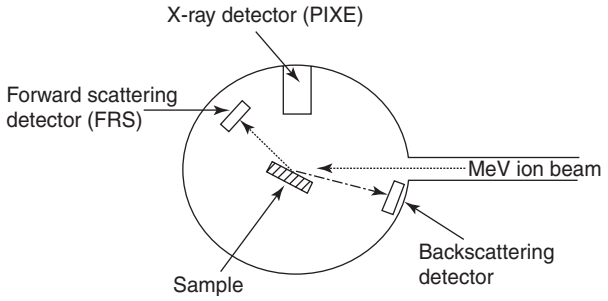


Figure 1 Schematic diagram of the target chamber and detectors used in ion-beam analysis. The backscattering detector is mounted close to the incident beam and the forward scattering detector is mounted so that when the target is tilted, ion recoils can be detected at angles of $\sim 30^{\circ}$ from the beam direction. The X-ray detector faces the sample and receives X-rays emitted from the sample.

interaction between the incident beam and the material under investigation. In other analytical techniques, such as secondary ion mass spectroscopy (SIMS), the cross sections are not well defined. The relative ion yields can vary over three orders of magnitude depending on the nature of the surface. Rutherford backscattering has been a convenient way to calibrate SIMS, which in turn, is more sensitive to the detection of trace elements than backscattering. Hence, these two techniques are complementary.

Ion-beam analysis grew out of early nuclear physics research on cross sections and reaction products involved in atomic collisions. In this work, million volt accelerators were developed and used extensively. As the energies of the incident particles increased, the lower energy accelerators became available for use in analytical applications. The early nuclear physics research used magnetic spectrometers to measure the energies of the particles. This analytical procedure was time consuming, and the advent of the semiconductor nuclear particle detector allowed simultaneous detection of all particle energies. The semiconductor detector is a Schottky barrier (typically, a gold film on silicon) or shallow diffused p–n junction with the active region defined by the high-electric field in the depletion layer. The active region extends tens of microns below the surface of the detector, so that in almost every application, the penetration of the energetic particles is confined within the active region. The response of the detector is linear with the energy of the particles providing a true particle energy spectrometer.

Analysis of ion-implanted layers and metal–silicon interactions (i.e., silicide formation) was carried out with Rutherford backscattering at 2.0 MeV energies and with semiconductor nuclear particle detectors

for decades. Rutherford backscattering became well established and was utilized in materials analysis in national, university, and eventually industrial laboratories across the world.

The recent interest in the use of implanted hydrogen to induce cleaving of thin layers of Si, has renewed the interest and importance of quantifying the amounts of hydrogen in Si and other single-crystalline substrates. Other areas such as solid-state chemistry of hydrogen production and storage have also required quantifying the amounts of hydrogen in thin films. Both these areas have led to the development of forward scattering, in which one quantifies the light-element content by measuring the energy of the recoiling hydrogen atom. The helium ion is heavier than that of hydrogen. In this case, the sample is tilted in a manner that the emerging hydrogen recoil energy can be measured with the use of a nuclear particle detector. The only modification to the RBS target chamber geometry is the tilt of the target and the position of the detector (see **Figure 1**). These forward recoil techniques have, of course, become more sophisticated with the use of heavy incident ions and detectors, which measure both the energy and the mass of the recoiling particles.

The analysis of thin film layers on species implanted into silicon is an almost ideal experimental situation for RBS because the masses of the thin films and/or atoms of implanted species and metal layers most often exceed the mass of Si. In Rutherford backscattering, the mass of the atom must be greater than that of the silicon target to separate the energy signals from the target atom from those of the silicon. However, there are few exceptions, such as carbon, nitrogen, and oxygen. The resonance analysis of these species is carried out in the same experimental chamber as used in RBS, but the energy of the incident helium ions is increased to energies where there are resonances in the backscattering cross sections. These resonances increase the yield of the scattered particle by nearly two orders of magnitude and provide high sensitivity to the analysis of oxygen and carbon in silicon. The use of these high energies, 4.3, 3.05, and 3.7 MeV for the helium–carbon, helium–oxygen, helium–nitrogen resonances, respectively, is called resonance scattering or non-Rutherford scattering. Insertion of a semiconductor X-ray detector into the analysis chamber allows the measurement of particle-induced X-rays. Particle-induced X-ray emission (PIXE) is a technique that can be used for nondestructive, simultaneous elemental analysis of bulk solids, thin films, liquids, and low specific volume gases. Like other fluorescence spectroscopies, PIXE is based on the emission of characteristic X-rays from an element. This reaction, however, is a direct result of the bombardment by MeV/amu ions (most often, protons or alpha particles). The use of proton and alpha beams as the probe-beam source offers several advantages over other fluorescence techniques in that it provides (1) a higher rate of data accumulation across the entire spectrum, and (2) better overall sensitivities, especially for the lower atomic number elements. This is due to a lower background noise.

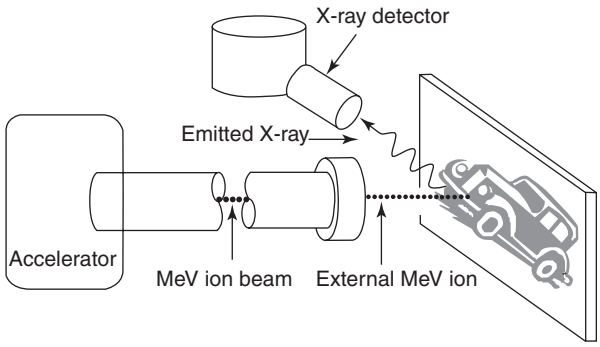


Figure 2 Schematic diagram of the setup for proton-induced X-ray emission. Protons of 3.5–5 MeV are produced in an accelerator, and travel in vacuum to the thin window at the end of the beam line. The protons have sufficient energy to pass through the window and travel in air to the painting. The emitted X-rays are detected by the X-ray detector.

Often the occasion arises such that the sample to be analyzed is too fragile or too large to fit into the ion-scattering chamber. A good example of this is the analysis of art work and archaeological artifacts. This situation requires the use of external PIXE analysis where the proton or alpha particles travel in vacuum to the thin window at the end of the beam line. The particles have sufficient energy (normally 3.5–5 MeV) to exit through the window (**Figure 2**) and to travel in air to the sample under investigation. For example, a 4 MeV proton can travel $\sim$25 cm in air before it loses all its energy in collisions with air molecules along its path. Consequently, the protons can be taken out into the air for external beam analysis.

The remaining portion of this article reviews the major techniques used in materials analysis that involve the use of MeV ion beams: Rutherford backscattering, channeling, resonance scattering, forward recoil scattering, and PIXE.

Rutherford Backscattering Spectrometry

In this section, the development of RBS analysis starts with an overview of the experimental target chamber. This is followed by a discussion of the particle kinematics, which determine mass identification and depth resolution, and then concludes with examples of the RBS analysis.

Target Chamber

In a typical scattering chamber (**Figure 1**), the sample is located so that its surface is on an axis of rotation of a goniometer such that the beam position does not shift across the sample as the sample is tilted with respect to the incident ion beam. The backscattering detector is mounted as close to the incident beam as possible such that the average backscattering angle, ϑ, is close to 180°, typically 170°, with a detector solid angle of 5 msr. In some cases, annular detectors are used with the incident beam passing through the center of the detector aperture in order to provide larger analysis solid angles. The sample can also be rotated into a glancing angle geometry when the forward scattering detector is used. In this case, a thin foil is placed in front of the detector to block the helium ions while allowing the hydrogen ions to pass through with only minimal energy loss. The stopping power (energy loss) of MeV helium ions is 10 times that of the recoiling hydrogen ions. In addition, the X-ray detector is placed such that the detector window (i.e., the entrance to the active region) is in full view of the sample surface bombarded with the incident ions. The vacuum requirements in the target chamber are comparable to those in the accelerator beam lines. Enhanced vacuum levels reduce the probability that the ion beam will lose energy along its path to the sample, and also minimizes deposition of contaminants and hydrocarbons on the surface during analysis.

In traditional backscattering spectrometry using helium ions, the energy resolution of the solid-state particle detector is typically 17 keV. This resolution can be improved to 10 keV with special detectors and detector cooling. The output signal, which is typically millivolts in pulse height is processed by silicon integrated circuit electronics and provides an energy spectrum in terms of number of particles versus energy. A multichannel analyzer records the number of backscattered particles versus a given energy in a specific channel. Hence, the energy scale is divided into channel numbers. To make the conversion from channel number to energy requires a simple energy calibration.

Scattering Kinematics

During ion-beam analysis, the incident particle penetrates into the target and undergoes inelastic collisions, predominantly with electrons, and loses energy as it penetrates to the end of its range. The range of 2.5 MeV helium ions is ~10 μm in a target of silicon atoms, the range of comparable energy protons is ~10 times that of the helium ions.

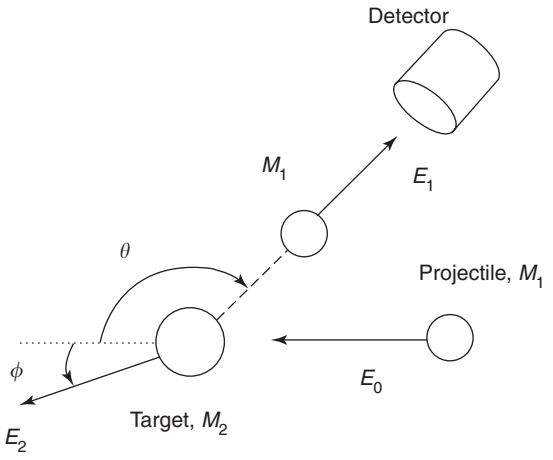


Figure 3 A schematic representation of an elastic collision between a particle of mass M_1 and initial energy E_0 and a target atom of mass M_2. After the collision, the projectile and target atoms have energies of E_1 and E_2, respectively. The angles θ and ϕ are positive as shown. All quantities refer to the laboratory frame of reference.

During the penetration of the helium ions, a small fraction undergoes elastic collisions with the target atom, which defines the backscattering signal. **Figure 3** shows a schematic representation of the geometry of an elastic collision between a projectile of mass M_1 and energy E_0 with a target atom of mass M_2 initially at rest. After collision, the incident ion is scattered back through an angle ϑ and emerges from the sample with an energy E_1. The target atom, after collision, has a recoil energy E_2. There is no change in target mass, because nuclear reactions are not involved and energies are nonrelativistic. The ratio of the projectile energies for $M_1 < M_2$ is given by

$$K = \frac{E_1}{E_0} = \left[\frac{(M_2^2 - M_1^2 \sin^2 \vartheta)^{1/2} + M_1 \cos \vartheta}{M_2 + M_1}\right]^2 \qquad [1]$$

The energy ratio $K = E_1/E_0$, called the kinematic factor, shows how the energy of the backscattered particle is a function of the incident particle and target atom masses, the scattering angle, and incident energy.

For a direct backscattering through 180°, the energy ratio has its lowest value given by

$$\frac{E_1}{E_0} = \left(\frac{M_2 - M_1}{M_2 + M_1}\right)^2 \qquad [2]$$

For incident helium ions $(M_1 = 4)$ at $E_0 =$ 2.8 MeV, the energy E_1 of the backscattered particle for silicon $(M_2 = 28)$ is 1.58 MeV and for silver $(M_2 = 108)$, the energy is 2.41 MeV. The energy E_2 transferred to the target atom has a general relation

given by

$$\frac{E_2}{E_0} = \frac{4M_1M_2}{(M_1 + M_2)^2}\cos^2\phi \qquad [3]$$

For scattering through an angle of $\vartheta = 180°$, the energy E_2 transferred to the target atom has its maximum value of

$$\frac{E_2}{E_0} = \frac{4M_1M_2}{(M_1 + M_2)^2} \qquad [4]$$

The ability to identify different mass species depends on the energy resolution of the detector which is typically 17 keV full width at half maximum (FWHM). For example, Ag has a mass $M_2 = 108$ and In has a mass $M_2 = 115$. The difference between $K_{Ag} = 0.862$ and $K_{In} = 0.870$ is 0.008. For 2.8 MeV helium ions, the difference in backscattering energy is 22 keV which is larger than the detector-system resolution, indicating that signals from Ag and In on the surface can be resolved. The difference between gold and tungsten K values, however, is 0.005, and hence at 2.8 MeV energies, it is not possible to resolve the signals between gold and tungsten. Conventional RBS systems using a 17 keV detector system can resolve signals from, and identify the elements of masses up to ~90. These systems can also resolve isotopes up to a mass of 55. For example, all of the silicon isotopes can be identified.

For the examples above, the detectors were assumed to be pristine and the system electronics to have negligible effects on the detectors. This is generally not the case. The system's electronics can contribute slightly to the value of the resolution. With an extended use and age, the worthiness of the detector resolution deteriorates and hence degrades the overall detector–system resolution.

Scattering Cross Section

The identity of target elements is established by the energy of the scattered particles after an elastic collision. This is done by measuring the yield Y, the number of backscattered particles for a given value of incident particles Q. The detector's solid angle is given as Ω. The areal density, the number of atoms per unit area, N_S determined from the scattering cross section $\sigma(\vartheta)$ by

$$N_S = \frac{Y}{\sigma(\vartheta)Q\ \mathrm{d}\Omega} \qquad [5]$$

This is shown schematically in **Figure 4**. A narrow beam of fast particles impinges on a thin uniform target that is wider than the beam. At a scattering angle θ from the direction of incidence, an ideal detector is located that counts each particle scattered in the differential solid angle $\mathrm{d}\Omega$. In the simplest approximation, the scattering cross section is given by

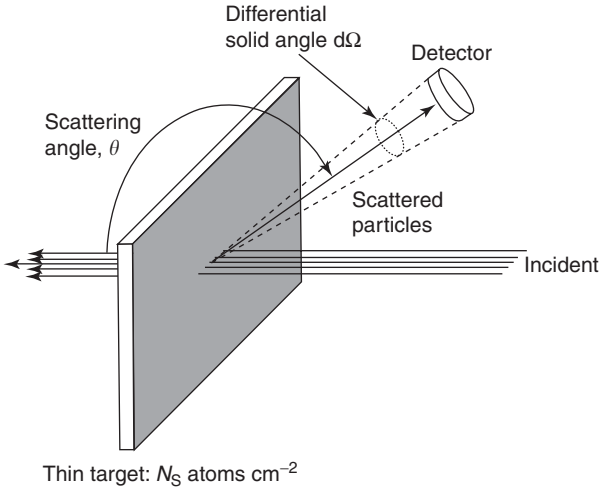


Figure 4 Layout of a scattering experiment. Only primary particles that are scattered within the solid angle dΩ spanned by the solid-state detector are counted. (Adapted and redrawn from Feldman LC and Mayer JW (1986) *Fundamentals of Surface and Thin-Film Analysis*. New York: North-Holland.)

$$\sigma(\vartheta) = \left(\frac{Z_1Z_2e^2}{4E}\right)^2 \frac{1}{\sin^4(\vartheta/2)} \qquad [6]$$

which is the scattering cross section originally derived by Rutherford. For 2 MeV helium ions incident on silver, $Z_2 = 47$ at 180°, the cross section is 2.9×10^{-24} cm^2 or 2.9 barns (where 1 barn = 10^{-24} cm^2). The distance of closest approach is $\sim 7 \times 10^{-3}$ nm which is smaller than the K-shell radius of silver (10^{-1} nm). This means that the incident helium ion penetrates well within the innermost radius of the electrons, so that one can use an unscreened Coulombic potential for the scattering. The distance of closest approach is sufficiently large so that penetration into the nuclear core does not occur and hence, nuclear reactions need not be considered.

The cross sections are relatively large such that one can detect sub-monolayers of most heavy-mass elements on silicon. For example, the yield of 2.0 MeV helium ions from 10^{14} cm^{-2} silver atoms (~1/10 of a monolayer) is 800 counts for a current of 100 nA for 15 min and detector area of 5 msr. This represents a large signal for a small amount of atoms on the surface.

Depth Scale

Light ions, such as helium, lose energy through inelastic collision with atomic electrons. In backscattering

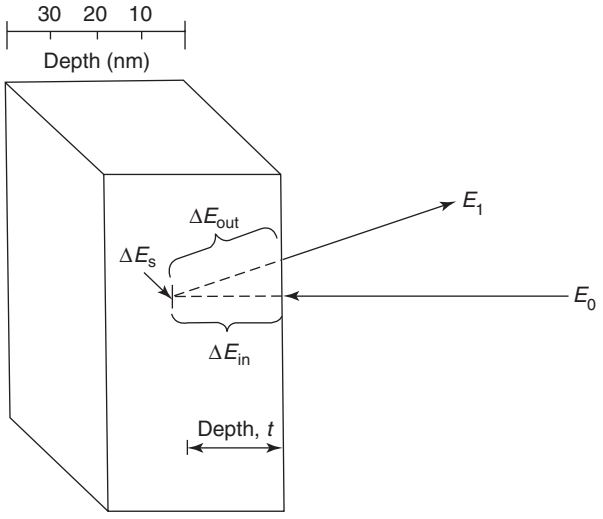


Figure 5 Energy loss components for a projectile that scatters from depth t. The particle loses energy ΔE_{in} via inelastic collisions with electrons along the inward path. There is energy loss ΔE_S in the elastic scattering process at depth t. There is energy lost to inelastic collisions ΔE_{out} along the outward path. For an incident energy E_0, the energy of the exiting particle is $E_1 = E_0 - \Delta E_{in} - \Delta E_S - \Delta E_{out}$.

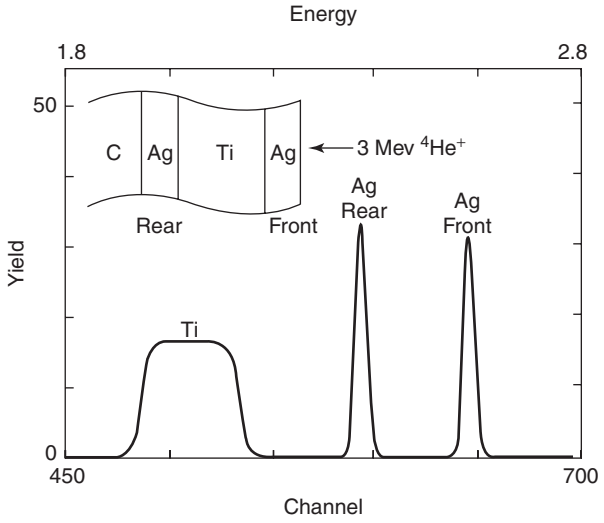


Figure 6 Backscattering spectrum of a Ti film (150 nm) with thin layers of Ag (3 nm) on the front and back surfaces of the titanium. The signals from the front and back layers of Ag are shown and are well separated in energy from each other by nearly the same energy width as the Ti signal.

spectrometry, where the elastic collision takes place at depth t below the surface, one considers the energy loss along the inward path and on the outward path as shown in **Figure 5**. The energy loss on the way in is weighted by the kinematic factor, and the total is given by the relationship

$$\Delta E = \Delta t\left(K\frac{\mathrm{d}E}{\mathrm{d}x}\bigg|_{\mathrm{in}} + \frac{1}{|\cos\vartheta|}\cdot\frac{\mathrm{d}E}{\mathrm{d}x}\bigg|_{\mathrm{out}}\right) = \Delta t[S] \quad [7]$$

where $\mathrm{d}E/\mathrm{d}x$ is the rate of energy loss with distance and $[S]$ is the energy loss factor. The particle loses energy ΔE_{in} via inelastic collisions with electrons along the inward path. There is energy loss ΔE_S in the elastic scattering process at depth t. There is energy loss due to inelastic collisions ΔE_{out} along the outward path. For an incident energy E_0, the energy of the exiting particle is $E_1 = E_0 - \Delta E_{in} - \Delta E_S - \Delta E_{out}$.

An example illustrating the influence of depth on analysis is given in **Figure 6**, which shows two thin silver layers on the front and back of a titanium film. The scattering from silver at the surface is clearly separated from Ag at the back layer. The energy width between the Ag signals, is closely equal to that of the energy width of the Ti signal. This signal is nearly square shaped because Ti exists from the front to the back surface. The depth scales are determined from energy loss values, which are given in a tabular form as a function of energy. It is often expressed as a stopping cross section in terms of $[S]$, the energy loss factor, which is given in units of $\mathrm{eV\,cm^{-1}}$. The depth resolution is given by dividing the detector resolution by the energy loss factor. For 2.0 MeV helium in silicon, one might expect a depth resolution of $\sim$20 nm for 180° scattering geometries. This can be reduced to values of $\sim$5 nm for glancing incident and exit angles. The resolution degrades as the particle transverses into the sample.

A typical use of RBS is to determine how the composition varies as a function of depth. The variation of composition with depth is used to characterize thin films and thin film reactions. For example, RBS is used to characterize the migration of Ti through a thin film of Ag after annealing in an NH_3 ambient. **Figure 7** (dashed line) shows the as-deposited spectrum for an Ag/Ti bilayer structure on a layer of SiO_2 on a silicon substrate. After a thermal anneal at 500°C in an ammonia ambient for 30 min (**Figure 7**, solid line), the RBS signal clearly shows how the composition varies as a function of depth. A computer simulation of the spectrum in **Figure 7** shows that the Ag layer contains a residual Ti concentration of <1 at.%. Rutherford backscattering spectra can be analyzed with the use of available analysis programs (e.g., Rutherford Universal Manipulation Program (RUMP). This software provides a layer-by-layer signal analysis of thin film materials and it gives the operator the ability to input: tilt angle, detector solid angle, detector resolution, individual isotopes, enhanced cross section for light elements, such as carbon and oxygen. RUMP has the ability to be used to analyze forward recoil spectra (see section "Elastic recoil spectrometry").

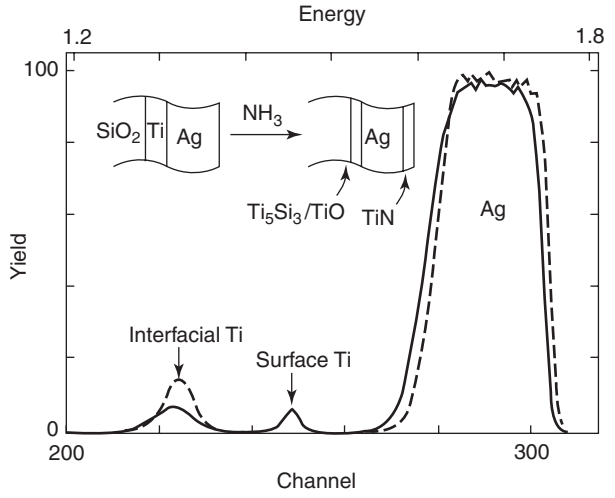


Figure 7 Ag encapsulation on SiO_2 prepared from Ag(120 nm)/Ti(22 nm) bilayers. RBS spectra show the depth distribution of Ag and Ti before and after annealing at 500°C in an ammonia ambient for 30 min. The spectra were obtained using a 2.0 MeV He^{+2} beam and a scattering angle of 170°.

Rutherford backscattering analysis is most often done in unison with one or more analytical techniques. **Figure 8** shows the backscattering spectrum from a single crystalline Si wafer implanted with 9×10^{15} Au^+ ions cm^{-2} at 3.6 MeV and the corresponding cross-sectional transmission electron microscopy (X-TEM) micrograph. The RUMP simulation of the Au RBS spectrum confirms the common depth scale. Two distinct regions of Au precipitates (dark spots) are present, coinciding in position with the two peaks in the Au backscattering profile. This example also shows the complementary nature of RBS and X-TEM analysis.

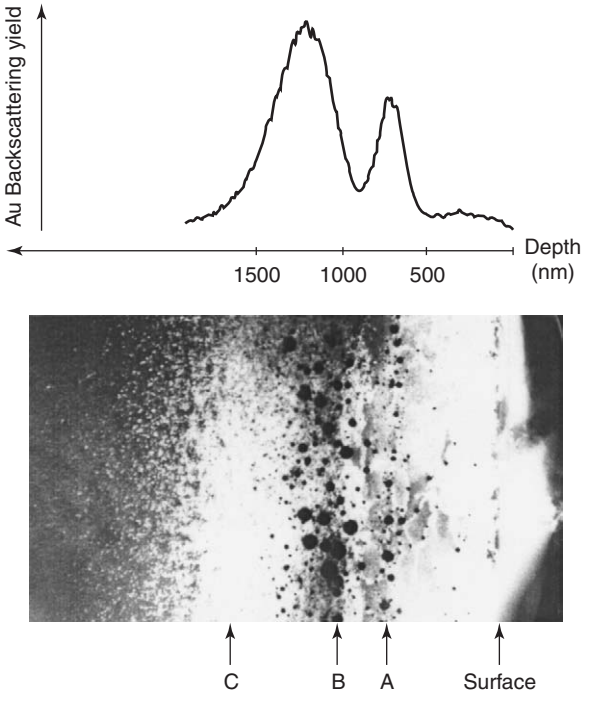


Figure 8 X-TEM micrograph of the 3.6 MeV Au^+ 8.6×10^{16} ions cm^{-1} implanted sample displaying the multiple precipitate layers which corresponds with the two Au peaks in the RBS spectrum. The depth scale is applicable to both X-TEM micrograph and the RBS spectrum. (From PhD dissertation of Alford TL.)

Ion Channeling

When a single-crystal sample is mounted on a goniometer such that a major crystallographic axis of the sample is aligned within ~0.1 or 0.5° of the incident beam, the crystal lattice can steer the trajectories of the incident ions penetrating into the crystal. It is this steering of the incident energetic beam that is known as ion channeling; the atomic rows and planes are guides that steer the energetic ions along the channels between rows and planes. The channeled ions do not approach close enough to the lattice atoms to be backscattered. Hence, the range (the total distance that an ion penetrates a solid) is increased several fold. In comparison, the reduced probability of scattering results in a two-orders of magnitude reduction in the yield.

Figure 9 shows schematically a random and aligned spectrum for MeV helium ions incident on

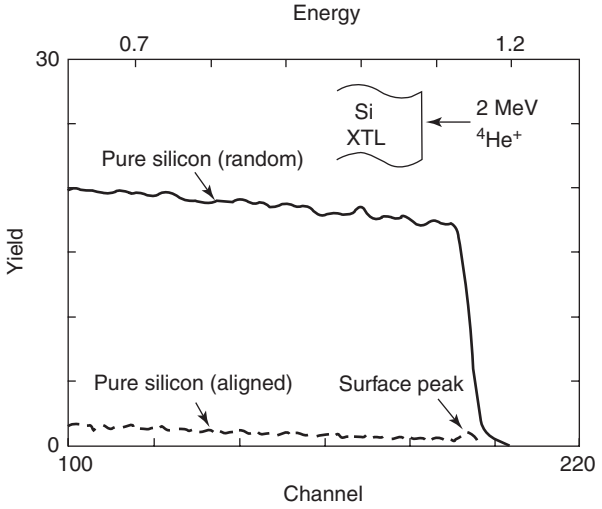


Figure 9 Random and aligned (channeled) backscattering spectrum from a single-crystal sample of silicon. The aligned spectrum has a peak at the high-energy end (near channel 195) of the Si signal. This peak, called the surface peak, represents helium ions scattered from the outer layers of Si that are exposed to the incident beam. The yield behind the peak is 1/40 of the random yield because the Si atoms are shielded from close encounter elastic collisions from the ion beam that is channeled along the axial rows of the Si crystal.

silicon. The characteristic feature of the aligned spectrum is the peak (near channel 195) at the high-energy end of the spectrum. This peak is a result of ions scattered from the outermost layer of atoms directly exposed to the incident beam. This peak is referred to as the surface peak. Behind the surface peak, at lower energies, the aligned spectrum drops to a value of 1/40 of the silicon random spectrum indicating that ~97% of the incident ions are channeled and do not make close-impact collisions with the lattice atoms. The rise in the aligned spectrum at lower energies represent the ions that are dechanneled, deflected from the steering by the lattice atoms, which can then collide in close impact collisions with the lattice atoms and hence directly contribute to the backscattering spectra. The ratio between the aligned minimum yield and random yield for the same channel number gives the χ_{min} value. The crystalline quality is given by the values of χ_{min} at a specific energy, for example, which equals 0.028 for Si when using 2.0 MeV helium ions.

Channeling phenomena is also important in the areas of ion implantation and epitaxial growth. In each case, the incident ions can be steered along the lattice planes and rows. Channeling measurements can determine the amount of lattice disorder when displaced atoms are located within the channels, they can measure the number of impurity atoms that are located sufficiently far from substitutional lattice sites so that they are accessible to backscattering from the channeled ions. Channeling also provides angular relationships between substrates and epitaxial layers.

During ion implantation, channeling leads to a deeper penetration of the incident ions when compared to the depths typically incurred for nonchanneled ions. The resulting implanted distribution of the channeled ions is a skewed Gaussian and not the typical Gaussian depth distributions characterized by nonchanneled energetic ions. **Figure 10** shows a plot of range of channeled ion and randomly aligned ions incident on a crystalline solid substrate.

The application of channeling to RBS is used to determine the amount of damage in ion-implanted single crystal silicon and the lattice location of ion-implanted dopant atoms. One important example of the contribution of channeling to integrated circuit technology is the analysis of damage evolution during ion implantation. **Figure 11** shows the channeling backscattering spectra obtained from H^+-implanted Si at different implantation doses. The peaks in the spectra of the implanted samples indicate the lattice damage induced by the hydrogen implantation. The difference in height of these peaks corresponds to the different H^+-implantation doses. At

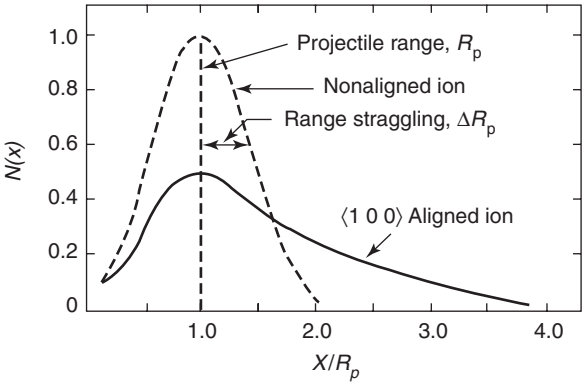


Figure 10 Distributions for channeled ions implanted along the ⟨1 0 0⟩ axis of Si. The dashed line shows the Gaussian distributions for incident ions aligned away from any channeling direction. (Adapted from Mayer W and Lau SS (1990) *Electronic Materials Science*. New York: Macmillan.)

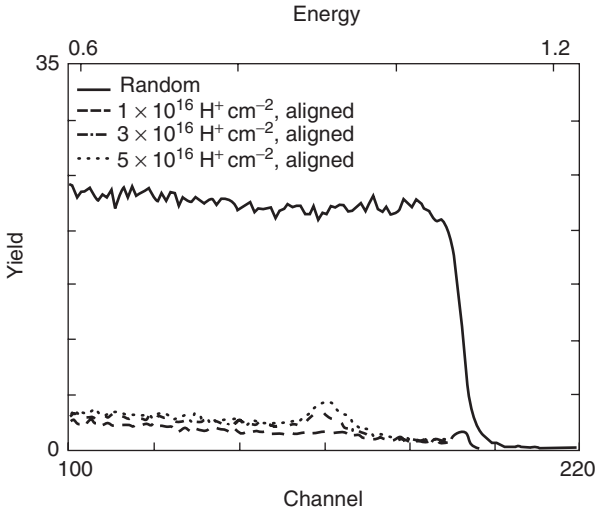


Figure 11 2 MeV He^{2+} RBS channeling (aligned) spectra of the ⟨1 0 0⟩ Si wafer, implanted with 40 keV H^+ ions at different doses. Also shown for comparison is a random spectrum.

low H^+-implantation dose, no damage peak is observed. The backscattering yield of the peaks at higher energies is quite low ($4\% < \chi_{min} < 9.5\%$ of the random signal), revealing the high crystal quality of the silicon layer above the implantation damage region. The creation of a submicron damage layer by H^+ implantation facilitates the slicing of thin high-quality silicon layers, which is essential for the production of silicon-on-insulator (SOI) and flexible displays.

Another example of channeling is the characterization of extended damage and eventual amorphization. **Figure 12** shows random and channeling spectra from silicon single-crystal samples implanted to doses of 0.4, 1.1, 1.8, and $2.3 \times 10^{15}\,cm^{-2}$ using 2.4 MeV Au^+ ions. As the implant dose increases,

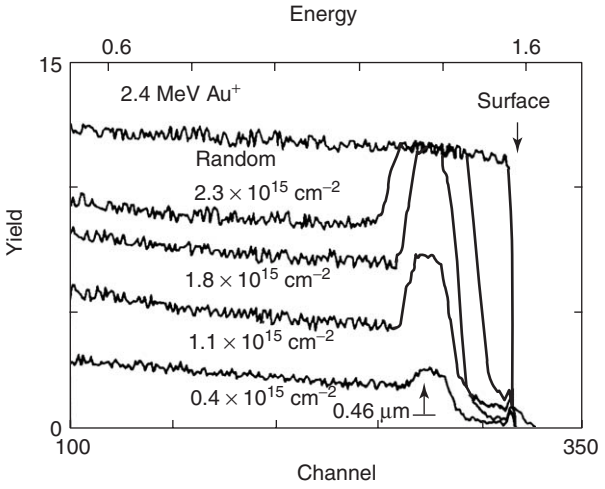


Figure 12 Random and aligned channeling ⟨100⟩ spectra from Si single crystals implanted with 2.4 MeV Au^+ to fluences of 0.4, 1.1, 1.8, and 2.3×10^{15} ions cm^{-2}. Note that only the Si signals are shown. (From PhD dissertation of Alford TL.)

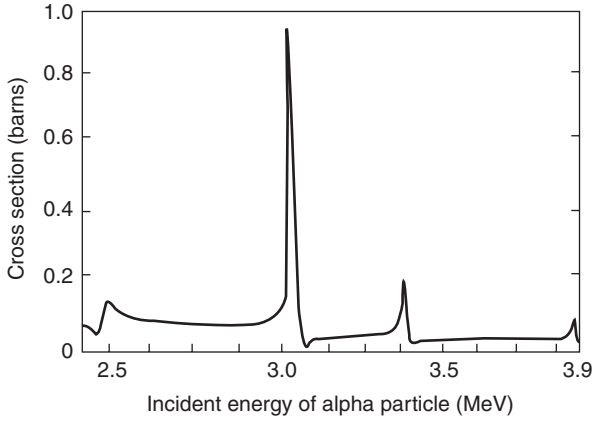


Figure 13 Elastic cross section of helium ions scattered from oxygen atoms. The pronounced peak in the spectrum ~3.04 MeV represents the resonance scattering cross section that is often used in the detection of oxygen. (Adapted from Tesmer JR and Nastasi M (eds.) (1995) *Handbook of Modern Ion Beam Materials Analysis*. Pittsburgh: Materials Research Society).

more silicon matrix atoms are displaced from the lattice sites; hence, the resulting damage profile increases as well. For the high doses, the damage-profile reaches the yield of the random spectrum. This is an indication of amorphous layer formation. Comparison of **Figures 11** and **12** reveals the usefulness of ion channeling to characterize both low and high amounts of damage in single crystals.

Ion Resonances

At energies of a few million electronvolts, nuclear reactions and strong deviations from Rutherford scattering should not play a role in the backscattering analyses. One of the exceptions is the strong increase (resonance) in the scattering cross section at 3.04 MeV for ^4He ions incident on ^{16}O, as shown in **Figure 13**. This reaction can be used to increase the sensitivity for the detection of oxygen as well as other light elements such as carbon and nitrogen.

The oxygen resonance cross section versus energy is shown in **Figure 13**. The resonance that occurs at 3.04 MeV shows a strong peak. In order to evaluate the amount of oxygen in a thin film of $TiAl_xN_yO_z$ on SiO_2/Si substrate, the oxygen resonance technique using 3.05 MeV $^4He^{2+}$ ion beam was employed (**Figure 14**). The RUMP simulation is overlapped with the real spectrum. The enhanced oxygen peak near channel 200 is a direct consequence of O resonance at 3.05 MeV and corresponds to oxygen atoms present in the thin film.

Carbon also has a resonance in its cross section leading to 100-fold increase in the backscattering signal. This resonance has been very convenient for

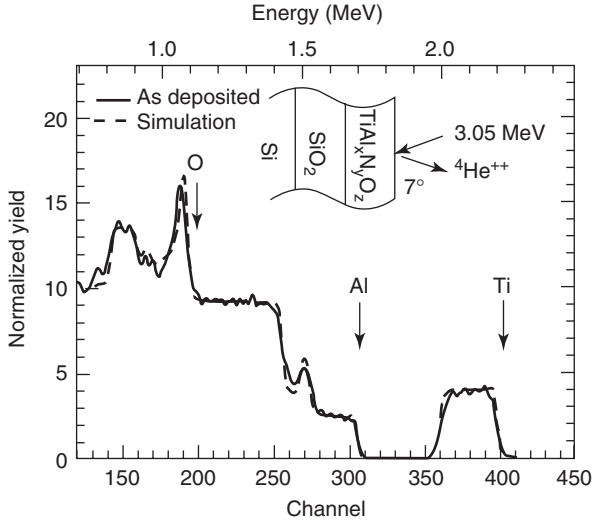


Figure 14 Backscattering spectrum (3.05 MeV He^{2+}, 7° tilt) and simulation of as-deposited $TiAl_xN_yO_z$ thin film on SiO_2/Si substrate.

analyzing 1% carbon in silicon–germanium films. The resonance for nitrogen is not as pronounced, but has been used extensively.

The use of elastic nuclear resonances provides a useful means of enhancing the sensitivity toward light elements using the same experimental setup as for Rutherford backscattering. Quantitative information about light-element concentrations is only obtainable under certain conditions. By using resonance near the peak value of the cross section, one may enhance sensitivity; however, this may result in a loss of precision, defined as day-to-day repeatability of the measurement. Conversely, using resonance in

an energy regime in which the cross section varies less rapidly may more accurately predict the actual composition than standard RBS, but prove less useful for low concentrations. Russell and co-workers in 1996, not only gave guidelines to the application of resonance analysis importance of film thickness and composition, and the variation in the incident beam energy toward the selection of the appropriate resonance regime, but also presented common sources of error. An excellent account of the effects of detector resolution, straggling, and film thickness was given by DeCoster and co-workers in 1992.

Elastic Recoil Spectrometry

The quantification of low Z elements is typically problematic due to the dependence of the scattering cross section σ on Z^2, when scattering occurs under Rutherford conditions. This has motivated the use of (α,α) resonances to enhance sensitivity in light-element profiling, requiring in turn that the necessary cross section data exist for the backscattering geometries found in different analytical laboratories. Elastic recoil spectroscopy detection (ERD) is another technique that is also similar to RBS in that the incident ion scatters from atoms in the matrix. What is unique about this technique is that the detected particles are not the backscattered particles, but light atomic weight species being forward scattered. ERD utilizes energy loss (ΔE) in the filter as a means to discriminate between signals from different atomic species present, for instance hydrogen, deuterium, and tritium in a single sample. The two techniques RBS and ERD are complementary techniques. For a given incident ion, ERD is sensitive to elements lighter than the incident beam, whereas RBS is sensitive to elements heavier than that of the incident beam.

The detection of H is not attainable with other techniques, such as SIMS. As with RBS, the depth resolution of ERD spectroscopy is based on the fact that the incident ion is slowed as it transverses the solid on both the inward and the outward paths. In a similar manner, the forward scattered ion also loses energy as it exits the solid. Hence, depth information is also obtainable for the forward scattered particle.

ERD uses the same relevant parameters (tilt angle, detector solid angle, detector resolution, and individual isotopes) as used in RBS. In the case of ERD, an energetic, higher atomic number particle impinges onto a target of a lower atomic number or a substrate containing lower atomic number atoms. Naturally, the lower-Z particle can only be recoiled in the forward direction, and, therefore, thick targets have to be bombarded at an oblique angle to allow detection of the recoil. The thin targets allow the transmission of the recoil through the target material.

In the case of thin targets, the incident ion also has a finite probability of undergoing a forward scattering event from the target into the detector. To minimize the effect of the forward scattered incident ion, a filter is typically placed in front of the detector to filter out scattered incident particles. The filter thickness ($\sim$25–100 μm of Mylar) is selected such that forward scattered, higher atomic number particles can be stopped while allowing the lower atomic number particle to recoil through the filter and into the detector. Due to kinematic reasons and straggling in this foil, the depth resolution is somewhat worse than that obtainable with RBS.

A more advanced application of the ERD technique is the use of high-energy heavy ions, such as 26 MeV Si^+. **Figure 15** shows a schematic of a typical high-energy scattering geometry. In this case, a number of target atoms lighter than the incident ion can be recoiled and detected. A salient feature of the use of high-energy ions requires that the detector has the capability to discriminate between particles of different atomic number. In this case, the foil that screens forward recoiled ions must be properly selected to minimize overlap between signals due to recoiling light elements. **Figure 16** shows the ERD spectrum for the as-deposited Cu–Ti alloy and spectra from the same film annealed in an ammonia ambient at 450°C, 550°C, and 650°C. The surface positions corresponding to O and N are indicated by the arrows. Approximately 1×10^{15} cm^{-2} nitrogen or

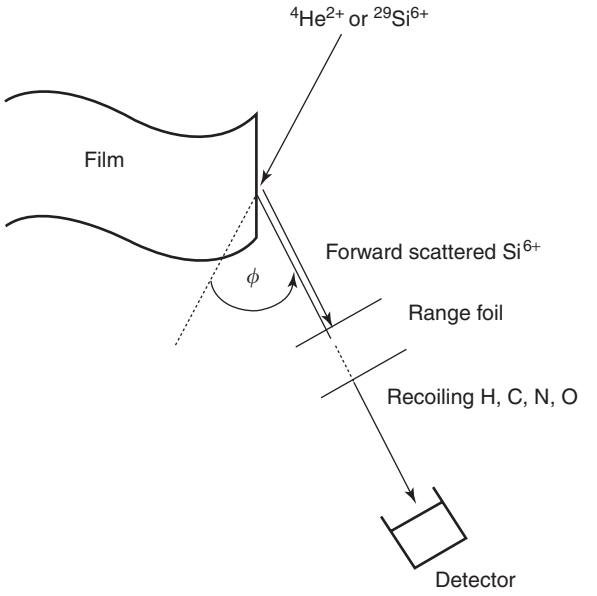


Figure 15 Schematic diagram of the elastic recoil detection process.

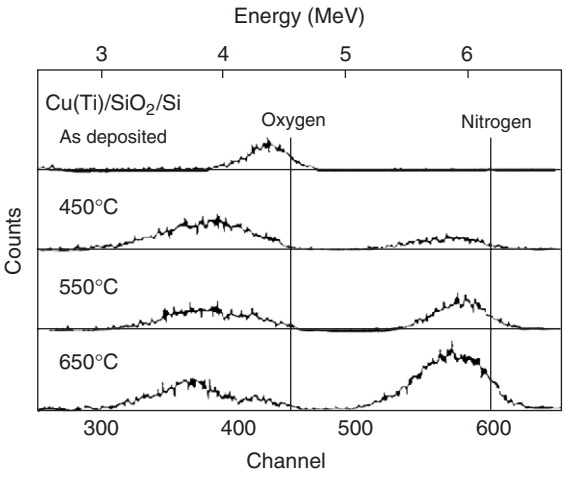


Figure 16 ERD spectra of N and O for as-deposited Cu(Ti) and for films annealed at 450°C to 650°C.

oxygen is observed in the as-deposited film by this technique. The nitrogen signal increases in area with increasing anneal temperature, indicating that the amount of nitrogen in the near-surface region of these films is steadily increasing. The N signal at 450°C is broad and flat, suggesting that a distinct N-containing layer does not exist. The height of the N signal increases up to 650°C but remains constant upon further annealing to 700°C. Also shown in the figures are signals due to oxygen. Note that the area of the major O signal remains approximately constant with annealing temperature, but shifts to lower energies. This corresponds to the oxygen-rich region being located deeper in the sample. A small step in the O signal is apparent at the two highest temperatures, with the step having its high-energy edge at the appropriate energy corresponding to the surface position of oxygen. This indicates that oxygen exists in two layers, one with a relatively small amount of oxygen located at the surface and the other with a large amount of oxygen located at a greater depth within the film.

Particle-Induced-X-Ray Emission

The PIXE analysis is based on the spectrometry of the radiation released during the filling of vacancies of inner atomic levels. These vacancies are produced by bombarding a sample with energetic (a few million electronvolts) ions. The binding energies of the electrons in the outer layers of the electron shell of an atom are of the order of some electronvolts, and radiation produced from the rearrangement of electrons in these levels is around the visible wavelength region. On the other hand, the binding energies of the inner levels are of the order of some kilo electronvolts, and, therefore, any radiation produced from processes involving these levels will typically be in the X-ray region. More importantly, as the electron energy levels of each element are quantized and unique, the measurement of the X-ray energy offers the possibility of determining the presence of a specific element in the bombarded sample. Furthermore, the number of X-rays measured of that specific energy is proportional to the amount of that element in the sample material.

The relative simplicity of the method and the penetrative nature of the X-rays yield a technique that is sensitive to most elements (for $Z > 10$) down to a few parts per million (ppm) and can be performed quantitatively from first principles. The databases for PIXE-analysis programs are typically so well developed as to include accurate, fundamental parameters, allowing the absolute precision of the technique to be $\sim 3\%$ for major elements and 10–20% for trace elements. A major factor in applying the PIXE technique is that the bombarding energy of the projectiles is a few orders of magnitude more than that of the binding energies of the electrons in the atom, and as the X-rays are produced from the innermost levels, no chemical information is obtained in the process. The advantage of this is that the technique is also not matrix dependent, and offers quantitative information regardless of the chemical states of the atoms in the sample. The major application of the technique is to determine trace-element concentrations. The nondestructive nature of this technique makes it attractive in the analysis of fragile artifacts or precious art.

In short, the technique is described schematically in **Figure 17**. A beam of energetic ions interrogates the sample. Ion–electron collisions result in the ejection of inner-shell electrons from atoms in a sample. This unstable condition of the atom cannot be maintained, and these vacancies are filled by outer-shell electrons. This means that the electrons make a transition in energy in moving from a higher energy level to a lower energy level, and this difference in energy can be released in the form of characteristic X-rays, of which the energy identifies that particular atom. In a competing process, called Auger electron emission, this energy can also be transferred to another electron that is ejected from the atom and can be detected by an electron detector. Therefore, the step from ionization to X-ray production is not 100% efficient, this efficiency being called the fluorescence yield, and has to be included in the database for quantitative measurements. The X-rays that are emitted from the sample are measured using an energy dispersive detector that has a typical energy resolution of $\sim 2.5\%$ (150 eV at 6 keV).

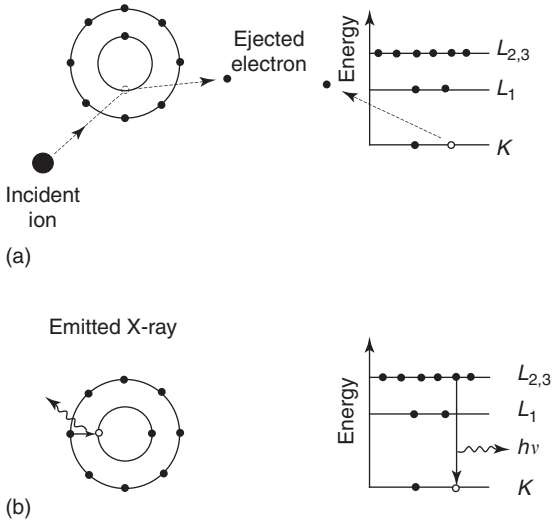


Figure 17 A schematic representation of X-ray generation by energetic particle excitation and the corresponding energy level diagrams. (a) A beam of energetic ions is used to eject inner-shell electrons from atoms in a sample. (b) These vacancies are filled by outer-shell electrons. The electrons make a transition in energy in moving from one level to another; this energy is released in the form of characteristic X-rays, of which the energy identifies that particular atom. The X-rays that are emitted from the sample are measured with an energy dispersive detector.

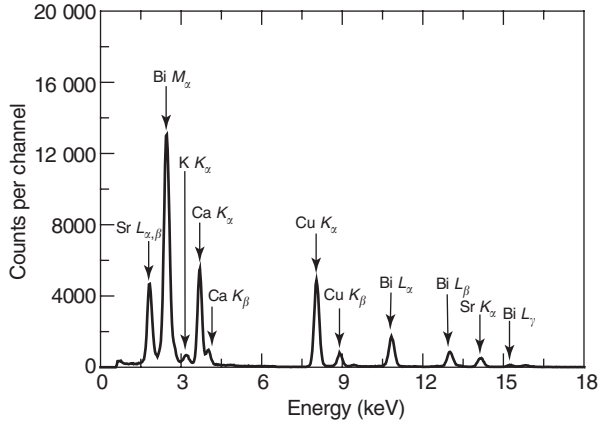


Figure 18 PIXE spectrum of a (Bi, Sr, Ca, Cu, O) superconductor taken with 1.7 MeV incident H^+ ions. (Data from K Streib at ASU.)

By convention, the transition filling vacancies in the innermost shell are called *K* X-rays, those filling the next shell are *L* X-rays, etc. The energies of the *L* X-rays are normally much lower than those of the *K* X-rays, and similarly *M* X-rays have much lower energies than the *L* X-rays. Due to the structure of the electron shells, there are naturally more possible transitions yielding *L* X-rays and even more possibilities of yielding *M* X-rays, and, therefore, it becomes more complex to measure the higher-order X-rays, and typically, the analytical method is limited to these three sets of X-rays from the elements. The limitation of detecting elements with $Z < 10$ is due to the low energies of the soft X-rays from the light elements that are absorbed before reaching the detector. This effect can also be put to good use in the technique, where the high yield of low-energy X-rays that originate from the major elements of a sample can be filtered using a selection of filters in front of the detector. Although the stopping of the bombarding ion is depth dependent, the X-ray energy measured gives no direct indication of the depth at which it was produced; therefore, the technique does not provide depth distribution information.

Typically, PIXE measurements are performed in the vacuum of a typical ion scattering chamber but it can also be performed in air with some limitations. Ion currents needed are typically around a few nanoamperes, and the current is normally not a limiting factor in applying the technique. The low-beam current also results in no or insignificant beam damage to samples during the analysis; hence, PIXE is a nondestructive analytical method sensitive to trace-element concentration levels.

The convenience of PIXE is most often compared to that of XRF (X-ray fluorescence) spectroscopy. **Figure 18** displays a spectrum taken from a bismuth strontium calcium copper oxide superconductor using PIXE analysis. This spectrum was taken with 1.7 MeV protons incident upon the sample and a 25 μm mylar filter between the sample and the detector. This filter blocks protons that are scattered from the sample from entering the detector and still allows the passage of most of the X-rays. However, the filter attenuates a small fraction of the radiation. Given that the mylar filter used in the acquisition of this spectrum is thinner, the low-energy lines are more prominent than if the filter were 50 or 100 μm. **Figure 19** displays a spectrum taken from the same sample; however, in this case, the incident radiation is molybdenum **K**α (17.48 keV) X-rays. This spectrum is taken in air and results in a small argon peak at the low-energy end of the spectrum. The radiation incident upon the sample excites some of the argon atoms present in the air before it reaches the sample. Note the presence of a molybdenum peak from the incident radiation at the high-energy end of the spectrum.

A comparison of the two spectra illustrates the advantages of PIXE over XRF. PIXE is capable of exciting X-ray lines across a wider range of energies than XRF, and is therefore capable of detecting a wider range of elements in a single spectrum. Also

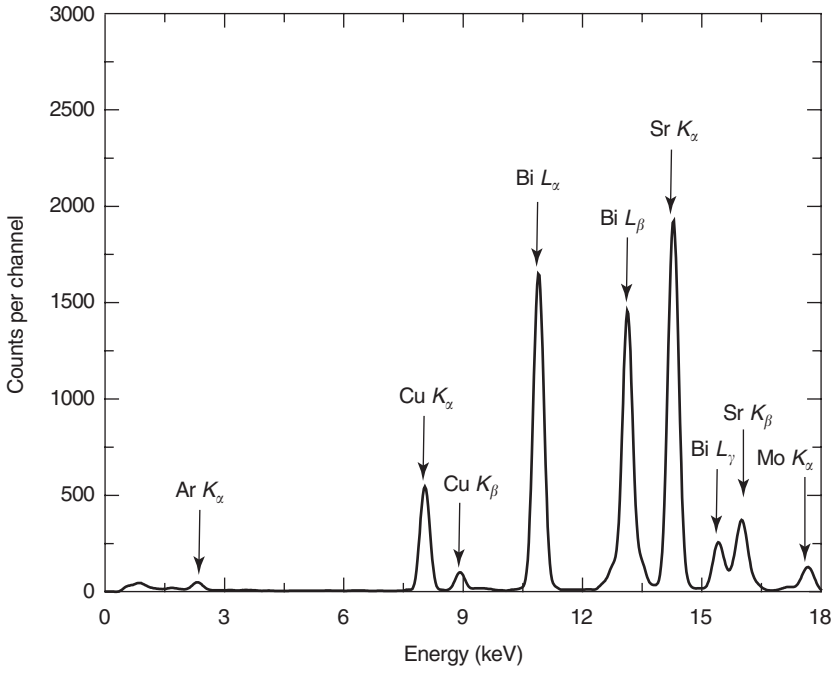


Figure 19 XRF spectrum of (Bi, Sr, Ca, Cu, O) superconductor taken with Mo K_α radiation incident on the sample. All the major peaks are indicated. The Mo peak is Mo K_α radiation scattered from the sample. The Ar K_α radiation is from air in the sample chamber (Data from K Streib at ASU.)

notice the lack of artifacts, such as argon peaks and secondary target peaks present in the XRF spectrum. This is another advantage of PIXE over XRF.

Closing Remarks

Ion-beam analysis is a mature technology that has stood the test of time. It is used in universities, national laboratories, and industry. Ion-beam analysis has also spread from semiconductor technology to fields as diverse as archaeology to the biomedical industry. References are provided to cover the basic concept and development (see the section on "Further reading").

See also: Scattering, Elastic (General).

PACS: 82.80.Yc; 78.70.En; 81.05.−t

Further Reading

Anderson HH and Ziegler JF (1977) *Hydrogen Stopping Power and Ranges in All Elements*. New York: Pergamon.

Bird JR and Williams JS (1989) *Ion Beams for Materials Analysis*. Sydney: Academic Press.

Campbell IL and Johansson SAE (1988) *A Novel Technique for Elemental Analysis*. New York: Wiley.

Chu W-K, Mayer JW, and Nicolet MA (1978) *Backscattering Spectrometry*. New York: Academic Press.

Composto RJ, Walters RM, and Genzer J (2002) Application of ion scattering techniques to characterize polymer surfaces and interfaces. *Materials Science and Engineering* R38: 107.

Cox RP, Leavitt JA, and McIntyre LC (1995) Non-Rutherford elastic backscattering cross sections. In: Tesemer JR and Nastasi M (eds.) *Handbook of Modern Ion Beam Analysis*. Pittsburg, PA: MRS.

De Coster SW, Brijs B, Goemans J, and Vandervost W (1992) Guidelines to the application of nuclear resonance to quantitative thin film analysis. *Nuclear Instruments and Methods B* 66: 128318.

Doolittle LR (1986) A semiautomatic algorithm for Rutherford backscattering analysis. *Nuclear Instruments and Methods B* 15: 227.

Feldman LC, Mayer JW, and Picraux ST (1982) *Materials Analysis by Ion Channeling: Submicron Crystallography*. New York: Academic Press.

Götz G and Gärtner K (1988) *High Energy Ion Beam Analysis of Solids*. Berlin: Academic.

Maxwell JA, Campbell JL, and Teesdale WJ (1989) The Guelph PIXE software package. *Nuclear Instruments and Methods B* 43: 218.

Mayer JW and Rimini E (1977) *Ion Handbook for Materials Analysis*. New York: Academic Press.

Russell SW, Levine TE, Bair AE, and Alford TL (1996) Guidelines to the application of nuclear resonance to quantitative thin film analysis. *Nuclear Instruments and Methods B* 118: 118.

Taft WS and Mayer JW (2000) *The Science of Paintings*. New York: Springer.

Tesmer JR and Nastasi M (eds.) (1995) *Handbook of Modern Ion Beam Materials Analysis*. Pittsburgh: Materials Research Society.

Tirira J, Serruys Y, and Trocellier P (1996) *Forward Recoil Spectrometry: Applications to Hydrogen and Determination in Solids*. New York: Plenum.

Ion Transport in Biological Membranes

M Baruscotti, Università degli Studi di Milano, Milan, Italy
G Thiel, Darmstadt University of Technology, Darmstadt, Germany
A Moroni, Università degli Studi di Milano, Milan, Italy

All living cells are surrounded by lipid membranes that are impermeable to ions and any large or charged molecules. With this, lipid barrier cells are able to separate a well-defined ionic environment inside the cell, the cytoplasm, from the extra cellular solution and also from internal compartments. The purpose of internal compartmentation is that biochemical reactions with different ionic requirements can occur simultaneously without interfering. A schematic representation of the main cellular compartments of a typical animal cell is illustrated in **Figure 1**. The maintenance of these compartments and the integrity of the membranes is an absolute requirement for cellular life.

As much as the separation of compartments by membranes is important for cell function, a tightly controlled and selective communication between the individual compartments is required. Such a communication cannot occur via simple diffusion across the lipid membrane, because this process is much too slow. Cells have therefore evolved specific transport proteins, which facilitate the diffusion of ions and solutes through the membrane in a selective and regulated fashion.

Ion channel proteins represent a very efficient mechanism of facilitated diffusion. Due to the architecture of the channel pore (see section "Structure and function of transport proteins"), these transporters allow ions to cross membranes at rates which approach diffusion in water.

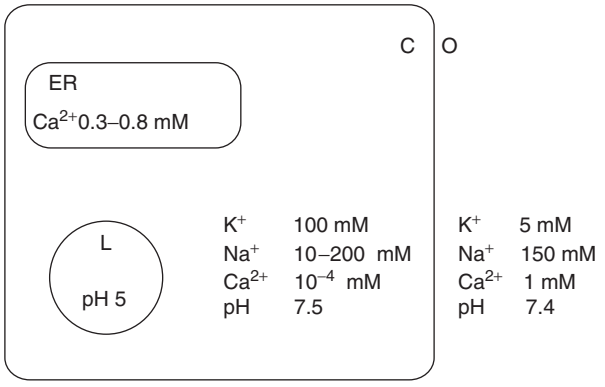


Figure 1 Concentration of main cations and pH in a typical animal cell and in some selected endo-compartments. O, external solution; C, cytoplasm; ER, endoplasmic reticulum; L, lysosome.

Diffusion is a passive process and occurs energetically only downhill. However, cells often have to concentrate ions in different compartments and in order to do this they have to move ions against their concentration gradient (see **Figure 1**). To this purpose cells utilize a class of transport proteins, termed pumps, which drive active, energy-consuming transport across the membranes. Active transport of a solute against its concentration gradient is driven by the energy released in the hydrolysis of a high-energy molecule, ATP. The pumps are therefore also called ATPases.

Transport of solutes against their gradient does not necessarily require the presence of a primary active transport system. The movement of solutes can also be coupled to the movement of other molecules down their respective electrochemical gradients (secondary active transport). In this case, the potential energy stored in ionic gradients (by the ATPases) can be utilized to transport other solutes. An example for this type of transporter is the Na^+/glucose co-transport. The concentration of sodium [Na^+] is kept low in the cytoplasm by the primary active transport system (the Na^+/K^+ ATPase). The glucose molecules are driven by secondary active transport by a Na^+/glucose co-transporter.

The best understood transporters to date, in terms of structure and function, are ion channels. This has two reasons. One reason is their key role in important physiological reactions such as nerve activity and muscle contraction. However, more significant for their understanding is their high turnover rate which allows them to be investigated at the single protein level. Two techniques, the planar lipid bilayer and the patch clamp technique, have greatly improved our understanding of ion transport through channels. The main features of the two methods are illustrated in **Figure 2**. The characteristic of the former approach is that the performance of a channel protein, even those from intracellular membranes, can be measured in a completely artificial system, in which all components such as the composition of the lipids and of the solutions are under the control of the experimenter, while the latter method allows one to measure channels in their native environment.

Both techniques report essentially the same kind of activity from ion channels. **Figure 3** illustrates a typical patch clamp recording of single-channel fluctuations in a membrane. The channel fluctuates between an open and a closed state. When it opens, it gives rise to a current, measured in picoampere (pA), with the

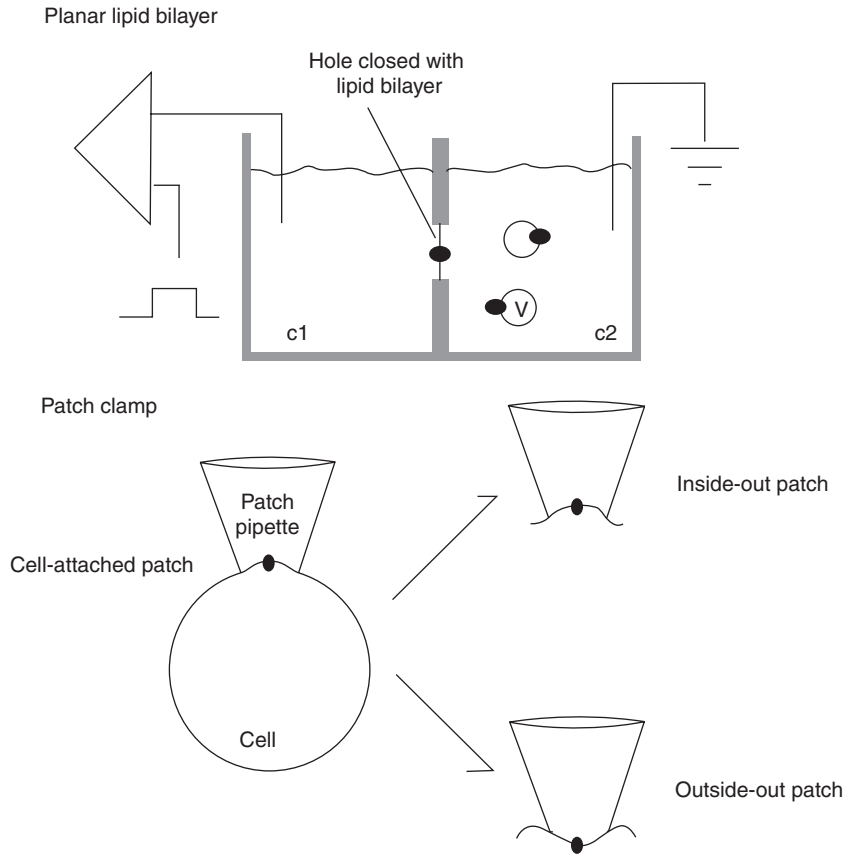


Figure 2 Sketch of main experimental techniques for single-channel recording. Planar lipid bilayer: two chambers (c1, c2) are connected by a small hole. This hole is sealed by a lipid bilayer reflecting a biological membrane. Vesicles (v) containing a channel protein (black dot) are made to fuse with this bilayer. This results in an incorporation of the channel protein into the bilayer. The activity of this channel can be recorded as a current between the two chambers. Patch clamp: A patch clamp pipette is sealed against the membrane of a cell. The current through a channel (here indicated by black dot) can be recorded. With different procedures, the membrane patch underneath the electrode can be isolated from the cell resulting in different orientations.

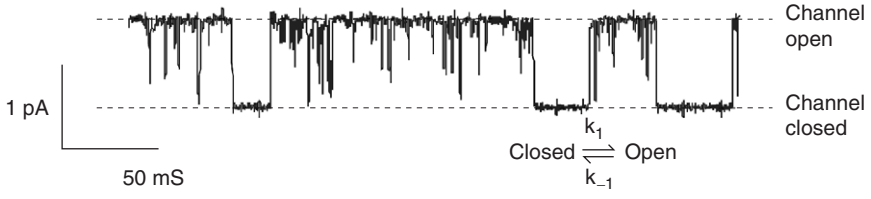


Figure 3 Example of single-channel fluctuations of K^+ channel. Opening of the channel pore allows a current flux of constant amplitude. The length which a channel spends in the open or closed state is variable. In the simplest case, the channel fluctuates in a stochastic manner with k_1 and k_{-1} between open and closed states.

characteristic open-channel amplitude. A single-channel current of 1 pA as in **Figure 3** corresponds to a flow of 6×10^6 ions per second.

In the simplest case, the conductance of an open channel behaves like an Ohmic resistor increasing linearly with the driving force. The slope of the current/voltage relation is a measure of the conductance of a channel. The conductance is characteristic of each individual channel proteins. In the case of potassium channels, they are in the range of picosiemens (pS).

The possibility of measuring and controlling the solution composition on both sides of the membrane allows examining the selectivity of these proteins. If a channel were perfectly selective for only one ion, the measured equilibrium voltage should match that predicted from the Nernst equation. Any other behavior indicates that the channel is either less selective or even, in the extreme case, not selective at all for a given ion. Channels of the latter kind have been described but they are a minority among the

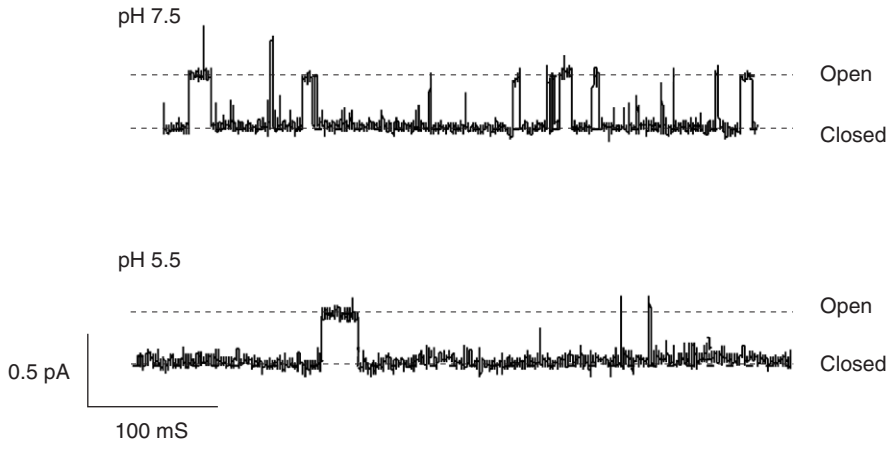


Figure 4 Effect of different pH on single K^+ channel gating. Activity of the same channel measured with a different pH on the cytoplasmic side of the membrane patch. Note that acidification increases the duration of the closed state.

physiologically relevant channels, which are in general highly selective. Experiments performed in the way indicated above show, for example, that K^+ channel are 1000 times more selective to K^+ than to Na^+. According to their selectivity, ion channels are classified as K^+, Na^+, Ca^{2+}, Cl^-, and H^+ channels.

The current flow through a channel is determined by the open-channel amplitude and by the relative time the channel stays open. As illustrated in the characteristic example in **Figure 3**, all channels fluctuate in a stochastic manner between an open and a closed state. In the simplest case, the mean dwell time of the open and the closed states are single exponentials so that the probability of a channel to be open (P_o) is given by

$$P_o = \tau_o/(\tau_o + \tau_c) \quad [1]$$

where τ_c and τ_o are the time constants for the closed and open times respectively.

While this simple kinetic model is suitable for the description of most channels, a detailed analysis of many ion channels has shown much more complex kinetic models. The analysis of the single-channel open and closed dwell times thus has led to models which include many more closed and open states organized in serial and/or parallel arrangement.

One of the important features of ion channels is that they are regulated. This guarantees that their activity is well integrated into the general physiological functions of a cell. In order to increase or decrease the current flux, channels increase or decrease their open probability, respectively. An increase in open probability can be achieved via a prolongation of the mean open dwell times or by a shortening of the mean closed dwell times or by both.

Several factors have been found to regulate ion channels in the physiological context. These include physical factors such as membrane voltage and mechanical stress. In addition, a large number of chemical modifications acting from either side of the membrane are known. **Figure 4** illustrates as an example the regulation of channel activity performed by pH on one side of the membrane. The data show that lowering of the pH results in a prolongation of the closed times, while the open times are not affected. As a result, the activity of the channels drops in a pH-dependent manner. This is mostly due to a titratable amino acid in the channel protein. As in the case of modulation by pH, most channel proteins harbor specific binding sites for their modulators.

Several different channels have been found to be regulated by chemical modification. The more general factors known to regulate channel activity in the plasma membrane of cells from the cytoplasmic side are, for example, phosphorylation/dephosphorylation, Ca^{2+}, pH, and cyclic nucleotides.

So far, the function of ion channels only at the single-protein level have been seen. From this information, one can easily extrapolate to the behavior of the ensemble of many channels of the same type in the membrane. From the aforementioned channel function, it can be deduced that the time-averaged current (I) through a channel is given by the product of the single-channel current (i) and the open probability (P_o):

$$I = i^* P_o \quad [2]$$

In this sense, the statistically weighted open-channel current is a perfect representation for the ensemble of the same kind of channels in a membrane. This only needs to be multiplied by the real number of

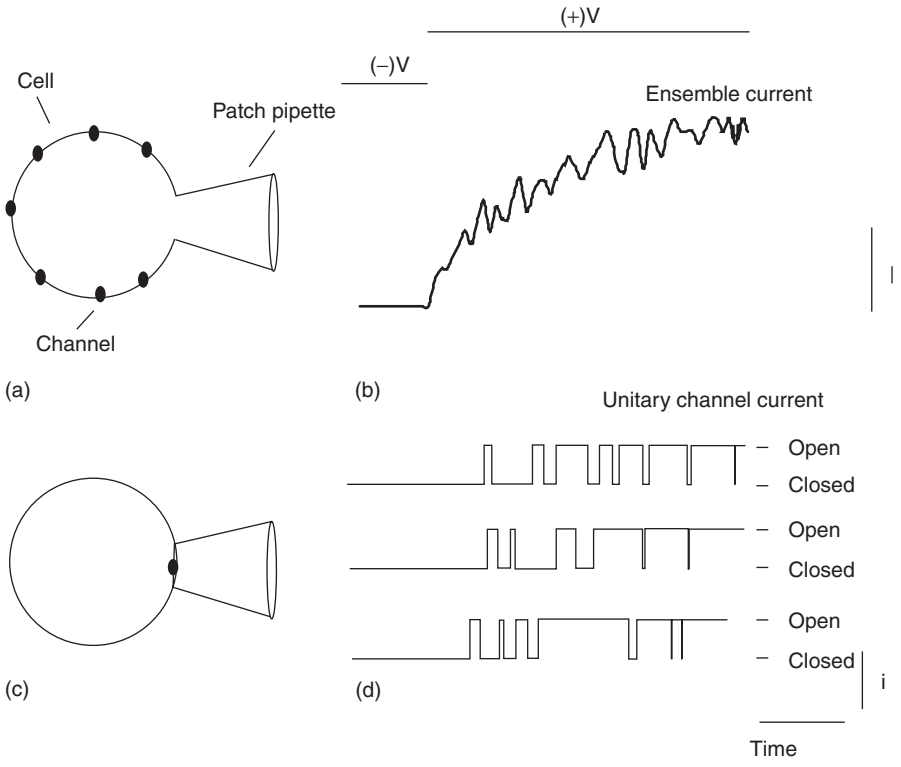


Figure 5 Relationship between single-channel activity and ensemble membrane current. Single-channel fluctuations (d) from a cell-attached recording (c, see **Figure 2**). The three simulated channel fluctuations are responses to a depolarizing voltage step (from (–)V to (+)V). The ensemble current (b) is obtained by averaging 30 sweeps as in (d). The same ensemble current (as in (b)) would be measured in the recording configuration illustrated in (a). Here the membrane underneath the patch pipette is ruptured to give a low resistance access to the cell interior.

channels in the membrane (N) to obtain the entire behavior of all channels in the membrane:

$$I = i^{*}P_{o} * N \qquad [3]$$

From an experimental point of view, it is sometimes more convenient to measure and analyze directly the ensemble currents in a membrane without any attention to the single channels. This can be achieved as illustrated in **Figure 5**. In this patch clamp configuration, referred to as "whole cell" configuration, the sum of currents through the ensemble of ion channels in the membrane is recorded. But in essence, the data between the statistical analysis of a single channel and the recording of ensemble currents are interchangeable.

Structure and Function of Transport Proteins

The first information on how membrane transport devices work came from physiological flux measurements. Physiologists traditionally divided transport mechanisms into two classes, carriers and pores. Carriers were responsible for transport which at high substrate concentration reached saturation, while pores were identified as mediating membrane transport which did not show any saturation kinetics. A carrier was essentially viewed as a molecule diffusing back and forth across the membrane while carrying small molecules or ions bound to specific sites, whereas a pore was a narrow, water-filled tunnel, permeable to few ions and molecules small enough to fit through the hole.

The numerous membrane proteins which have been purified, and the related genes which have been cloned, show that the carrier proteins are too large to diffuse or spin around at the predicted rate. The new view of carrier transport is that the protein is fixed in the membrane while exposing the transport-binding sites alternatively at the two sides of the membrane by small motions. Although the first crystal structure of a carrier appeared in the year 2000, nevertheless the specific mechanisms by which several carrier devices (transporters and pumps) work still await to be proved combining different experimental approaches, such as electrophysiology, molecular biology, and crystallography.

The view of pores as water-filled tunnels has been firmly established for a class of proteins which is now known as ion channels. The early identification of

these pores as proteins that span the membrane and the subsequent identification of the genes that codify for them have focused the work of electrophysiologists, biophysicists, and molecular biologists since the 1980s on ion channel structure–function relationships, creating a solid background of knowledge on how these channel proteins operate. The understanding of the structure and function of ion channels culminated in the landmark determination of the crystal structure of a K^+ channel, the bacterial channel KcsA, by Roderick MacKinnon. For his work on structure determination of several bacterial K^+ channels, MacKinnon received the Nobel Prize for chemistry in the year 2003.

The atomic structures of K^+ channels have especially improved the understanding of two main properties of ion channels: selectivity and gating.

Selectivity

A potassium channel can conduct K^+ ions one thousand times better than Na^+ ion. It is assumed that, in order to allow the protein channel to discriminate one ion from the other, the ions have to pass the cell dehydrated. In this situation, the atomic ray of Na^+ is smaller than that of K^+ (0.95 Å and 1.3 Å, respectively) implying that the Na^+ ions are not excluded on the basis of their dimension and that the selection process probably involves some kind of interaction with specific binding sites in the protein. The second striking characteristic of K^+ channels is that they pass K^+ ions at high speed, 10 e6, close to the diffusion limit. How can a protein interact with an ion and nevertheless allow its passage at the speed of ion diffusing in water? The atomic structure achieved for the bacterial potassium channel KcsA provides a deep understanding of K^+ selectivity and the ion conduction process.

Figure 6 reports the structure of the KcsA channel. The K^+ channel pore is formed by four identical subunits surrounding a central ion conduction pathway (two of the four subunits are shown in **Figure 6a**). Each subunit contains two transmembrane α-helices, termed inner (nearest the ion pathway) and outer helices, and a tilted shorter pore helix (in red) pointing toward the ion pathway and keeping in place the selectivity filter (in gold). The selectivity filter discriminates between potassium and other ions. It is the narrowest part of the ion permeation pathway, and is located in the extracellular third of the ion pathway. Inside the selectivity filter, the K^+ ion encounters four binding sites, which are created by evenly spaced layers of carbonyl oxygen atoms and a single layer of threonine hydroxyl oxygen atoms. At these sites a K^+ ion binds in a dehydrated

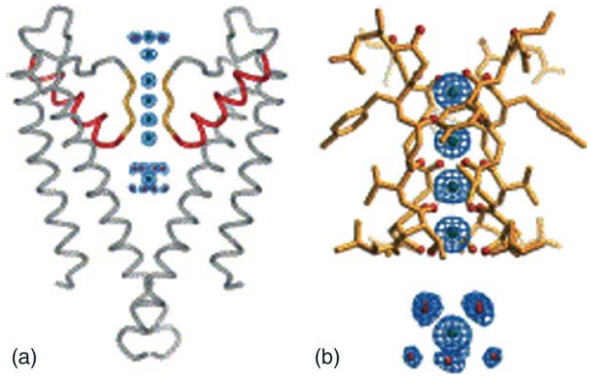


Figure 6 Ion conductance through a potassium channel pore, as deduced from the crystal structure of the bacterial channel KcsA. (a) Two of the four subunits from the pore of KcsA are shown, with the extra cellular side on top. Each subunit is formed by an outer helix that contacts the lipid membrane, an inner helix that faces the permeation pathway, a pore helix (red), and a selectivity filter (yellow). Blue mesh shows electron density for K^+ ions (green) and water molecules (red) along the pore. (b) Enlargement of the selectivity filter showing four dehydrated K^+ ions inside the filter and one hydrated K^+ ion outside (bottom). (Reprinted with permission from R MacKinnon (2003) Potassium channels. *FEBS Letters* 555: 62–65.)

state, surrounded by eight oxygen atoms from the protein, four "above" and four "below" each ion (**Figure 6b**). The arrangement of protein oxygen atoms in each binding site is very similar to the arrangement of water molecules around hydrated K^+ ion observed in the central cavity (**Figure 6b**). Potassium ions therefore diffuse from water into the selectivity filter where the energetic cost of dehydration is compensated. Sodium ions, on the other hand, do not enter the selectivity filter because the dehydrated Na^+ ions have a radius of 0.9 Å, and their coordination is precluded in this rigid structure. The precise ion coordination required for selectivity does not prevent their rapid diffusion through the pore. This is due to the fact that the selectivity filter contains more than one ion and the repulsion between closely spaced ions helps to overcome the affinity that each ion has for its binding site.

Gating

While the pore of a channel is designed to achieve minimal energy barriers to ensure a fast flow rate, other parts of the channel are responsible in determining whether the channels are open or closed, and they do so by propagating conformational changes to the ion-permeating pathway inside the pore. Potassium channels differ in relation to the various ways in which they are gated, that is, the way they change from a nonconductive to a conductive state. Some K^+ channels are ligand-gated, which means that the pore opening is energetically coupled to the binding of an

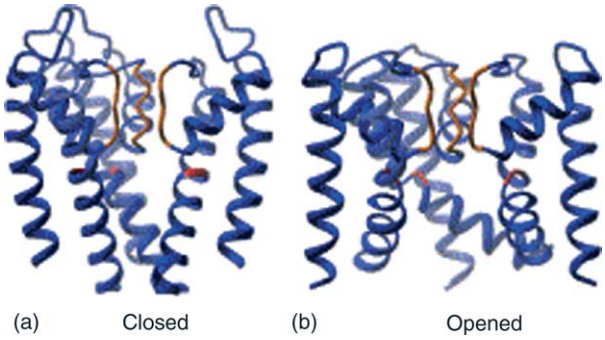


Figure 7 The pore gating of a potassium channel. Left: structure of the bacterial channel KcsA, crystallized in the closed conformation. Right: structure of the bacterial channel MthK crystallized in the open conformation. Three subunits are shown for each channel with the selectivity filter yellow and gating hinge in red. (Reprinted with permission from R MacKinnon (2003) Potassium channels. *FEBS Letters* 555: 62–65.)

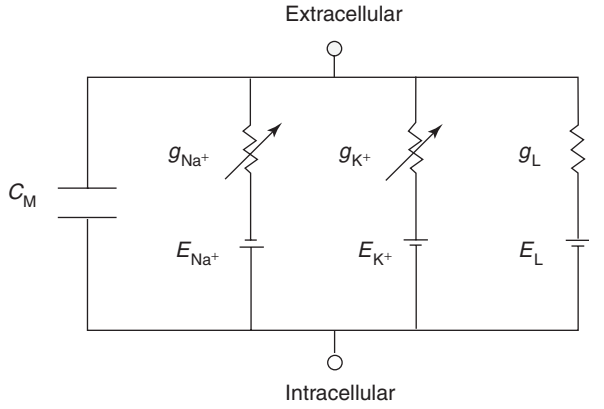


Figure 8 Electrical circuit representing the cell membrane. The lipid bilayer is modeled by a capacitance (C_M), while the different families of ionic channels are represented by conductances (g_x). The electromotive forces (E_X) generated by asymmetric ionic concentrations are included in the battery elements. Note that g_{Na} and g_k are variable conductances due to their time and voltage dependence.

ion, a small organic molecule, or even a small protein. Other K+ channels are voltage-gated, and the pore opening is energetically coupled to the movement of a charged voltage sensor within the membrane electric field. Diversity in gating is achieved by diverse structural domains attached in a modular fashion to the conserved pore domain. Binding of ligands as well as displacement of the voltage sensor by voltage changes must result in a global conformational change of the channel protein which is eventually transduced to the pore unit, causing the pore to open (and close). The recent crystallization of several K^+ channels (both in the open and closed conformation) has explained the conformational changes which underlie the pore opening in K^+ channels.

The pores of KcsA, crystallized in the closed configuration, and MthK, crystallized in the open configuration, are compared in **Figure 7**. The prominent difference between these two structures is the position of their inner helices. In KcsA they form a bundle at the cytoplasmic side of the pore that prevents free ion flow, while in MthK the inner helices are bent at a hinge point, allowing free ion flow between the cytoplasm and the selectivity filter. These structures are probably representative of the closed- and open-pore configurations of many different K^+ channels, as suggested by the amino acid conservation in the inner helices sequences of all K^+ channels (in particular of a glycine at the hingepoint) (**Figure 7**).

The Cell Membrane as an Equivalent Circuitry of Parallel Ion Channels with Variable Conductance: An Explanation for Complex Changes of Membrane Voltage

In this section, the functional roles of ion channel proteins are discussed with the aim of integrating their individual properties. This aims to explain their concerted contribution to the overall cell activity.

The concerted activity of ion channels may here be described for electrically excitable cells such as neurons and muscle cells. In general these cells can be found in two different conditions, namely at rest or excited. In the resting condition, the cell membrane voltage is stable over long periods of time; during electrical excitation, the cell undergoes a transient depolarization, a phenomenon called action potential (AP).

Both situations can be explained by a dynamic interplay of different ion channels in the plasma membrane. A physical model of the main elements in the plasma membrane of such a cell is presented in **Figure 8** where an equivalent circuit of a cell membrane built on capacitance, variable resistors, and batteries is presented. In this simple circuit, each resistor represents a specific type of ion channel, that is for K^+ or Na^+, with a variable internal resistance. The branch with the conductance g_L represents the ensemble of other conductances such as channels for Cl^- and Ca^{2+}. The variability of the resistor reflects the ability of the ion channels to undergo changes in activity. An important detail of this model is that the activity of the channels is sensitive, in a distinct way, to voltage. Hence through the voltage coupling of the channels via the membrane, the activity of each channel has an effect on the activity of the others. Due to their different distribution across cell membranes, the main ionic species involved in electrical activity (i.e., Na^+, K^+, Ca^{2+}) have different thermodynamic equilibrium (see **Figure 1**). These electrochemical gradients are a source of potential energy.

In the model they are represented by specific battery voltages, which are determined by the Nernst potential (E_X) of the respective ion. Cells consume this energy difference when ionic channels open and ions can freely diffuse. Due to its dielectric properties, the lipid component of the cell membrane is represented by a capacitor (C_M). Large and fast changes in the membrane voltage will hence cause currents to charge this capacitor.

At rest, cells have an internal potential of $-60/-80$ mV with respect to the extracellular environment. The precise value of the resting potentials is determined mainly by two factors: a relatively high K^+ (g_{K+}) and Cl^- (included in g_L) permeability of the cell membrane due to active K^+ and Cl^- channels and the respective battery voltages of E_{Cl}^- and E_{K+}. E_{Na+} does not contribute to the resting voltage since Na^+ channels are generally closed at rest ($g_{Na+} = 0$). Hence, the membrane current (I) at rest can be described according to the Ohms law by

$$I = g_{K+} * (V - E_{K+}) + g_{Cl}^- * (V - E_{Cl}^-) \qquad [4]$$

This resting state is drastically modified when cells become electrically active, that is, when they fire an AP. A simple physical description of an AP would be a wave of potential that travels at a rate of 1–100 m $*s^{-1}$ along excitable membranes, and conveys information from one point to another. The peak of the moving wave corresponds to a physical location where the membrane polarity is inverted (the intracellular part is positive up to $+40/+60$ mV with respect to the external), while the adjacent regions are still (region ahead) or have already returned (region before) to the resting hyperpolarized value.

The Hodgkin and Huxley (HH) model presented the first mechanistic explanation of the AP. This original model has now undergone major revision; however, the basic ideas can still be used to unravel how the concerted work of different ionic channels can bring about the AP. HH's interpretation of the molecular mechanisms underlying the complex event of an AP consists of a voltage-coupled interplay of two conductances: the Na^+ and K^+ channels. The elements that confer the circuit its active properties (i.e., the capacity of generating an AP) are the variable conductances (g_x) which are disposed in parallel (**Figure 8**). In the original HH model, the number of conductances is limited to three: g_{Na} and g_K that are the true, time- and voltage-dependent components, and a fixed g_L that represents a pathway for leakage currents. The last elements to consider are the source of electromotive forces E_x which complete each conductance branch with the electrochemical energy associated with the specific ionic flow and determined by the Nernst equation. The two dominant elements, the Na^+ and K^+ channels, have very distinct voltage- and time-dependent properties which can be summarized as follows: the Na^+ channels open transiently very fast upon a depolarization of the membrane. The K^+ channels in contrast open slowly upon depolarization. Because of their electrocoupling, the activity of one reflects back on the other and vice versa. Once the appropriate set of time- and voltage-dependent variables for each channel is defined, the mathematical solution of the circuit will appropriately simulate the electrical activity of the cell.

A more intuitive description of an AP can be attempted by describing the events that take place in a membrane region that is invested by an AP wave. If one could take a look at the intracellular side of the membrane, a situation would be observed where an already depolarized, positively charged region, is followed by a region in an electrically resting state of about $-70/-80$ mV. This situation would generate a depolarizing flow of positive charges toward the negative region; this event is the key element in the generation of APs. If the depolarization is not strong enough, then the resting K^+ conductance will soon buffer this perturbation and reestablish the correct resting potential, but if the depolarization reaches a threshold it will trigger a cascade of events that generates an AP. This is the reason why biophysicists describe the AP as an all-or-nothing event. It has already been seen that a great number of ion channels are gated by voltage, and this concept is crucial to the understanding of the AP mechanism. The threshold corresponds to a critical value of potential at which Na channels start opening, and thus generate an inward-going current that further depolarizes the membrane by a positive feedback mechanism that drives the membrane potential toward the equilibrium of Na^+ ions (E_{Na}). This depolarization activates two additional events: the inactivation of Na^+ channels and the activation of K^+ channels. Na^+ channels are dually gated by voltage: the positive electric field acting on these transmembrane proteins initially induces a conformational change that forces the channels to an open, conductive state, but soon after the channels relax to a nonconductive state. The voltage-gated K^+ channels involved in the AP differ from the resting ones in that the former are normally closed at resting, but open upon depolarization. K^+ channels activate with a slower kinetics compared to that of Na^+ channels, and, according to the acting electrochemical gradient, they generate an outward-going K^+ ions efflux, that repolarizes the cell to the resting state, and thus terminate the AP signaling. The electrical activity of a cell can now be summarized in three general

components: first, a trigger phase that moves the potential to the threshold; second, the activation of an inward-going current that rapidly depolarizes the cell, and finally the third part driven by a potassium outflow that terminates the electrical event. The above events are an over-simplification of a reality that is often more complex and thus adaptable; nature has evolved molecules with slightly different properties so that an array of elements is available to better accomplish the different functional requirements of excitable cells. Indeed APs can simply represent a piece of information that needs to be transferred from one point to another with high fidelity, but in many cases they must locally trigger crucial events such as a neurotransmitter release or a muscle contraction. Often, when spatial transferring of information is required, a simple system, based on limited population of channels (HH model), is perfectly fit for the function. Neuronal exocytosis and muscle contraction require the contribution of at least one more class of channels: the Ca^{2+} channel family. Ca^{2+} ions are key elements in inducing both the activation of the synaptic release machinery, and in allowing the physical interaction between the sliding filaments of muscle fibers.

Ion channels can be largely modulated by several effectors, for example calcium, second messengers such as cyclic nucleotides, phosphorylation, interaction with other proteins, and more (see above). As an example of the functional importance of these modulations, the activity of cardiac pacemaker cells are described here (**Figure 9**). These cells exhibit spontaneous voltage oscillations (**Figure 9a**) that determine the cardiac rate.

The most striking property of these cells is the slowly depolarizing or pacemaker phase (between arrows in **Figure 9a**) that, while moving the cell membrane toward the threshold (dotted line), allows for cardiac chamber refilling. This phase is generated by the intervention of several ionic channels and transporters, with the most important one being a specific current named I_f or pacemaker current. The I_f is modulated by the neural autonomic system that can increase (sympathetic branch) or decrease (parasympathetic branch) the current availability at any given potential (**Figure 9b**). The functional effect of this modulation is the modification of the slope of the diastolic depolarization that becomes steeper or shallower thus determining the autonomic-induced modulation of the heart rate (**Figure 9a**).

So far, the role of ion channels in underlying important physiological functions have been described; one can now appreciate why ion channels are important targets for drug-based therapy. Indeed in the year 2003, 15% of the top 100 best-selling drugs on the market were designed to specifically interfere with

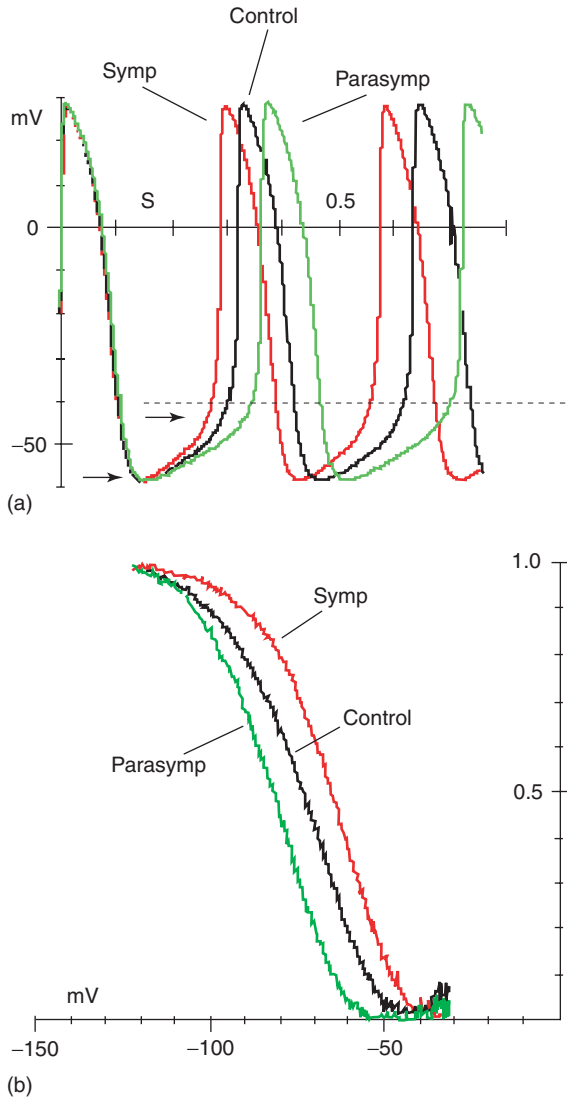


Figure 9 Heart rate modulation by autonomic system. (a) Spontaneous APs of pacemaker cells in control and during autonomic-induced modulation. (b) Fractional availability (*y*-axes) of the pacemaker current as a function of the applied voltage. The amount of available current modifies the slope of the pacemaker depolarization and thus the heart rate. (Modified with permission from Accili EA, Robinson RB, and Di Francesco D (1997) Properties and modulation of I_f in newborn versus adult cardiac SA node. *American Journal of Physiology, Heart and Circulatory Physiology* 272: H1549–H1552; and from Accili EA, Proenza C, Baruscotti M and Di Francesco D (2002) From funny current to HCN channels: 20 years of excitation. *News in Physiological Sciences* 17: 32–37; The American Physiological Society.)

ion-channel proteins. Several diseases such as epilepsy, long QT syndrome, cystic fibrosis, myopathies (but the list is in constant expansion) have been found to depend on abnormalities of genes coding for these proteins. For this reason, a new group of diseases has now been named after their molecular origin channelopathies. A still far to come, yet enticing possibility, is a strategy devised by ion-channel biologists which is

based on the "use" of ion-channel proteins as molecular tools for therapy of cardiac rhythm alteration. To date, cardiac electronic pacemakers have been widely used in the treatment of rhythm disorders. However, recent research suggests the possibility of developing therapeutic approaches based on "biological pacemakers" in which the ability of generating spontaneous cardiac electric activity could be restored by manipulating the level of ion channels expressed in a specific desired region of the heart.

See also: Biomembranes; Membrane and Protein Aggregates.

PACS: 87.16.Uv; 87.19.Nn; 87.80.Jg; 87.17.Nn; 87. 64. Gb; 87.16.Dg

Further Reading

Ashcroft FM (2000) *Ion Channels and Disease*. San Diego: Academic Press.

Armstrong CM (1975) Ionic pores, gates, and gating currents. *Quarterly Reviews of Biophysics* 7: 179–210.

Colquhoun D and Hawkes AG (1981) *On the Stochastic Properties of Single Ion Channels*. Proceedings of the Royal Society of London. Series B 211: 205–235.

DiFrancesco D (1993) Pacemaker mechanisms in cardiac tissue. *Annual Review of Physiology* 55: 455–472.

Doyle DA, Morais Cabral J, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, and MacKinnon R (1998) The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* 280: 69–77.

Hamill OP, Marty A, Neher E, Sakmann B, and Sigworth FJ (1981) Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archive: European Journal of Physiology* 391: 85–100.

Hille B (2001) *Ion Channels of Excitable Membranes*, 3rd edn. Sunderland, Massachussets, USA: Sinauer Associates, Inc.

Hille B, Armstrong CM, and MacKinnon R (1999) Ion channels: from idea to reality. *Nature Medicine* 10: 1105–1109.

Hodgkin AL and Huxley AF (1952) A quantitative description of membrane currents and its application to conduction and excitation in nerve. *Journal of Physiology, London* 117: 500–544.

Hodgkin AL and Keynes RD (1955) The potassium permeability of a giant nerve fibre. *Journal of Physiology, London* 128: 61–88.

Jiang Y, Lee A, Chen J, Ruta V, Cadene M, Chait BT, and MacKinnon R (2003) X-ray structure of a voltage-dependent K^+ channel. *Nature* 423: 33–41.

MacKinnon R (2003) Potassium channels. *FEBS Letters* 555: 62–65.

Miller C (1985) *Ion Channel Reconstitution*. New York: Plenum Press.

Neher E and Sakmann B (1992) The patch clamp technique. *Scientific American* 266: 44–51.

Sakmann B and Neher E (1995) *Single Channel Recording*. New York: Plenum.

Yellen G (1998) The moving parts of a voltage-gated ion channel. *Quarterly Reviews of Biophysics* 31: 239–295.

Ionic and Mixed Conductivity in Condensed Phases

J Maier, Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany

Introduction: Hopping Process

While a quantum mechanical particle such as a quasi-free-electron can tunnel through a barrier by interference of the states on both sides of the barrier (denoted by x and x'), that is, by seizing the hindrance in a wave-like way, the transport of a heavy particle such as an ion has to rely on the ensemble fraction with energy high enough to reach the final state (hopping). On the one hand, under favorable circumstances protons may also tunnel, while on the other hand electrons that strongly polarize their environment (large polarons) behave almost classically as ions.

In most cases of interest (consider A as an ion with charge z_A), the hopping of A from x to x' ($x' = x + \Delta x$) may be written as

$$A(x) + V'(x') \rightleftharpoons V(x) + A'(x') \qquad [1]$$

(The dash takes account of the possibility for A to have a different structural environment at x'.) In cases where A denotes an ion on a regular site in a crystal, V is the vacancy that serves as a jump partner, while in cases where A refers to an excess ion sitting on interstitial sites, V denotes the vacant interstitial site. In crystals, V (in the first case) or A (in the second case) are usually dilute with the energy levels of these sites being well defined. The concentration of the respective reaction partner is then approximately constant and the corresponding rate equation for the hopping process is pseudo-monomolecular. In amorphous solids, the sites are less defined, which also holds for the spatial energy distribution; in polymers, melts, or liquids, these sites and energy distributions are time dependent. In these cases, one would better speak of smeared-out density of states, as far as thermal excitations are concerned. The rate constant for the hopping rate equation is decisively influenced by the activation free enthalpy. The parameters k and $\Delta G^{\neq}$, for example, for the forward reaction (index $\rightarrow$) related through (prefactor k_0 is proportional to the attempt

frequency Γ_0)

$$\vec{k} = k_0 \exp\left(-\frac{\Delta \vec{G}^{\neq}}{RT}\right) \qquad [2]$$

refer to the built-in part of the forward rate constant and its activation free energy, respectively. This expression has to be complemented by the factor $\exp(-\vec{\alpha}\,\Delta\phi\, z\, F/RT)$ in order to take account of the applied electric field. The parameter $\Delta\phi$ is that part of the external voltage that drops from x to x' and the term $\vec{\alpha}\,\Delta\phi$ measures the portion of $\Delta\phi$ that is relevant for the forward reaction (distance between saddle point and initial state). The relations for the backward reaction are analogous ($\vec{\alpha} + \vec{\alpha} = 1$). This leads to the Butler–Volmer equation which describes the charge transfer from x to x' for an asymmetric free-energy profile. In the homogeneous bulk region (and also in the space charge region sufficiently far from the interface, i.e., $A = A'$), the built-in profile is symmetrical, that is, $\vec{k} = \overleftarrow{k} = k$, $\vec{\alpha} = 1/2 = \overleftarrow{\alpha}$ and $\Delta\vec{G} = \Delta\overleftarrow{G}$. The result for the rate represented by the current density i_j of carrier j in x-direction

$$\begin{aligned} i_j &= -2k_j c_j z_j F \sinh\frac{z_j F \Delta\phi}{2RT} \\ &\simeq -\sigma_j \frac{\Delta\phi}{\Delta x} \end{aligned} \qquad [3]$$

can be simplified to Ohm's law (bottom part) for $\Delta\phi \ll RT/z_j F$; the conductivity σ_j of the carrier is proportional to its mobility u_j ($u_j \propto k_j$), (molar) concentration (c_j) and (molar) charge ($z_j F$) and given as

$$\sigma_j = z_j F u_j c_j \qquad [4]$$

Allowing also for a concentration variation during the process and generalizing to three dimensions, the Nernst–Planck equation is arrived at

$$i_j = -D_j z_j F \nabla c_j - \sigma_j \nabla\phi \qquad [5]$$

where the diffusion coefficient (D) and conductivity (σ) are related through the Nernst–Einstein equation. Close to equilibrium, this equation corresponds to the result obtained by linear irreversible thermodynamics,

$$i_j = -\frac{\sigma_j}{z_j F}\nabla\tilde{\mu}_j \qquad [6]$$

which assumes the current to be proportional to the gradient in the electrochemical potential $\tilde{\mu} = \mu + zF\phi$ (chemical potential + (molar charge × electrical potential)). The latter relation is more general concerning the concentration range, while the former (as regards the diffusion term) is more general as far as the magnitude of this driving force is concerned. In cases where the current of a given species is also generated by secondary driving forces (e.g., by $\nabla\tilde{\mu}$ of a different particle), cross terms occur for which the Onsager–Casimir symmetry relations are valid. As regards the determination of the conductivity, the first task is to determine the carrier concentration as a function of the control parameters. This is feasible by solving the defect model for the system under consideration.

Equilibrium Charge Carrier Concentrations

Pure Compounds and Dilute Bulk

Figure 1 displays the fundamental charge carrier formation reaction in an "energy level" diagram for a fluid (H_2O) and an ionic crystal (AgCl). These levels refer to effective standard (electro) chemical potentials or – in the case of the "Fermi levels" that are positioned within the gap – full (electro) chemical potentials. (As long as the bulk is considered, $zF\phi$ can be neglected and one can refer to μ instead of $\tilde{\mu}$.) As long as – in pure materials – the gap remains large compared to RT, the Boltzmann form of the chemical potential of the respective charge carrier (defect) is valid which has the form

$$\mu_j = \mu_j^\circ + RT \ln\left(\frac{c_j}{c_j^\circ}\right) \qquad [7]$$

As outlined below, this relation holds largely independent of the charge carrier situation and the form of the energy level distribution. If the levels are smeared out or if there is a band of levels, μ_j° refers to an effective level (i.e., band edges in nondegenerate semiconductors or an appropriately averaged energy when dealing with fluids or disordered solids).

The coupling of the ionic level picture and the electronic level picture for AgCl (**Figure 2**) is established via the thermodynamic relation $\tilde{\mu}_{Ag^+} + \tilde{\mu}_{e^-} = \mu_{Ag}$, and hence via the component potential which reflects the precise position in the phase diagram.

The meaning of μ° becomes obvious, if for simplicity one considers an elemental crystal (with defect j). For the formation of N_j identical defects (among N regular positions), a local free enthalpy of $N_j \Delta g_j^\circ$ is required in the interaction-free case; including also the configurational contribution, the Gibbs energy of the defective crystal (G_P refers to the perfect crystal) reads

$$G = G_P + N_j \Delta g_j^\circ - k_B T \ln\left(\frac{N}{N_j}\right) \qquad [8]$$

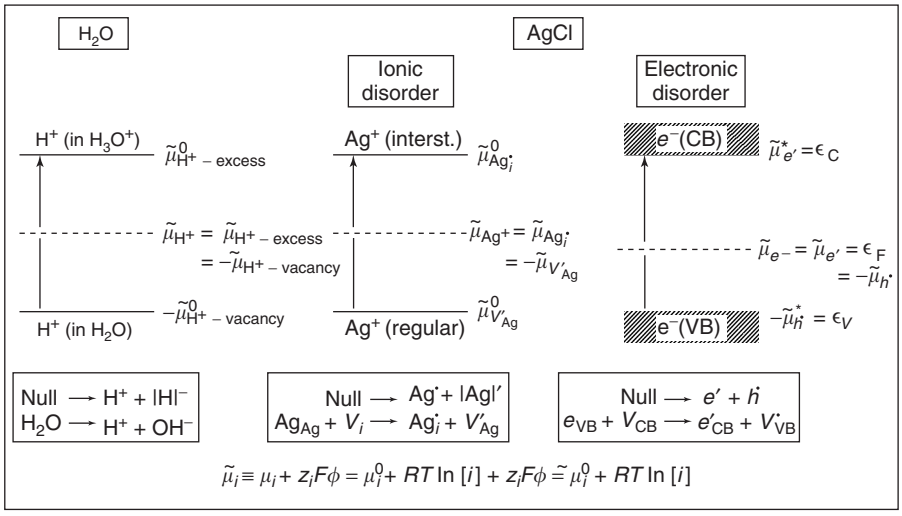


Figure 1 Electronic and ionic disorder in ionic solids and water in "physical" (top) and "chemical" language (bottom).

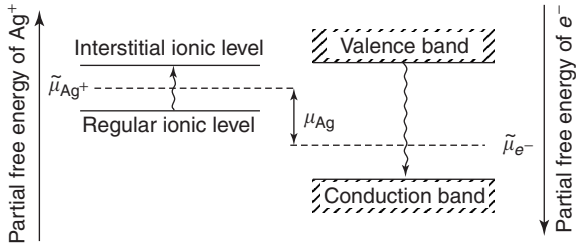


Figure 2 Coupling of ionic and electronic levels in AgCl.

The configuration term describes the number of combinations of N_j elements out of N choices without repetition. The calculation of Δg°, the free energy to form a single defect, requires an atomistic consideration. If one refers to ionic defects in crystals, it is essentially composed of lattice energy, polarization energy, and vibrational entropy terms. In more general terms, N_j is the number of defects and N the number of particles that can be made defective. Owing to the infinitely steep decrease of the configurational term with N_j, a minimum in $G(N_j)$ (see **Figure 3**) is observed in the elemental crystal, if the chemical potential of j, that is, $\mu_j \equiv \partial G/\partial(N_j/N_m)$, which follows as

$$\begin{aligned}\mu_j &= \mu_j^\circ + RT\ln\frac{N_j}{N-N_j}\\ &\simeq \mu_j^\circ + RT\ln\frac{N_j}{N}\\ &= \mu_j^\circ + RT\ln\frac{n_j}{n}\end{aligned} \qquad [9]$$

vanishes ($n \equiv N/N_m$, $n_j \equiv N_j/N_m$, N_m = Avogadro's number; $\mu_j^\circ = N_m \Delta g_j^\circ$). It is noteworthy that the strict result (top) which is formally valid for higher concentrations also is of the Fermi–Dirac type. This is due to the fact that one refers to combinations without repetition, that is, double occupancy is excluded and hence the sites are exhaustible similarly as is the case for quantum states in electronic problems. (Combinations with repetitions lead to Bose–Einstein distribution.)

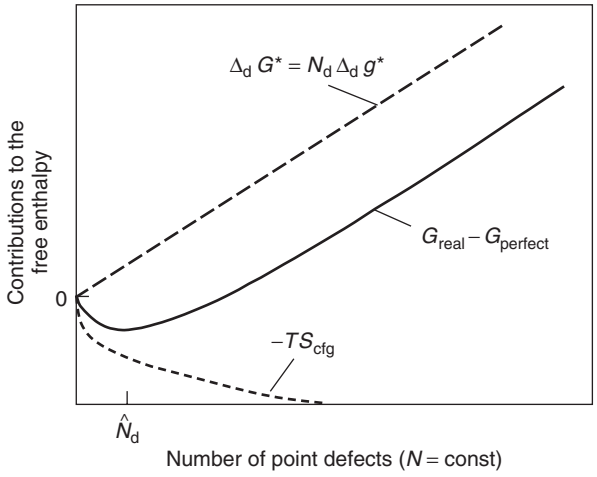


Figure 3 Contributions to the free enthalpy of the solid by defect formation with a constant total number of sites.

In a situation in which the levels are broadened to a more or less continuous zone (e.g., electrons in semiconductors and ions in fluids), the parameters N_j and μ_j° of the Fermi–Dirac form have to be attributed to an infinitely small level interval. Denoting the partial free energy level by ϵ, this integral ranges from ϵ_j to $\epsilon_j + d\epsilon_j$ and the total (molar) concentration

follows from integration:

$$n_j = \int_{\text{zone}} \frac{D(\epsilon_j)\,\mathrm{d}\epsilon_j}{1+\exp(|\epsilon_j-\mu_j|/RT)} \qquad [10]$$

whereby the molar density of states $D\ (\epsilon_j)$ appears in the integrand. At equilibrium, μ_j is independent of ϵ_j. Considering without the restriction of generality, the almost empty zone and the lower edge of this zone designated as ϵ_j, one can neglect the unity in the denominator of the above equation for all levels more distant from μ_j than ϵ_j' which leads to the Boltzmann expression $n_j \simeq \bar{n}_j(T)\exp(-|\epsilon_j'-\mu_j|/RT)$. The effective value $\bar{n}_j$ is given by

$$\bar{n}_j(T) = \int_{\text{zone}} \mathrm{d}\epsilon_j \left[D(\epsilon_j)\exp\left(-\frac{|\epsilon_j-\epsilon_j'|}{RT}\right)\right] \qquad [11]$$

$\bar{n}_j$ being usually weakly temperature dependent. It is worth noting that, owing to the high dilution, changes of the parameters with occupation have been ignored (for electrons in semiconductors, this is called the "rigid band model").

As anticipated above, the Boltzmann form of the chemical potential results with μ° representing an effective value. If n is identified with $\bar{n}_j$, the effective value μ_j° is given by ϵ_j'. As long as one refers to dilute conditions, considerable freedom of normalization is left with respect to the concentration measure.

In ionic compounds, charged defects occur pairwise, the formation of which is subject to disorder reactions (and electroneutrality conditions). Then, only the reaction combination of chemical potential vanishes, that is, $\sum_j \nu_{rj}\mu_j = 0$ (identical to $\sum_j \nu_{rj}\tilde{\mu}_j = 0$ since $\sum_j \nu_{rj}z_j = 0$), where ν_{rj} is the stoichiometric coefficient of j in the respective disorder reaction r (cf. **Figure 1**, bottom). Boltzmann expressions can be written for the individual carriers, if the formation energies and the configuration entropies are considered to be independent.

The application of chemical thermodynamics elegantly permits treatment of a variety of ionic and electronic disorder equations that simultaneously occur in a given solid. This is not only possible for the internal disorder equations (such as Schottky, Frenkel, anti-Frenkel, anti-Schottky, electron–hole formation) but also for external reactions such as the interaction of an oxide with the oxygen partial pressure and hence the stoichiometric variation. Under Boltzmann (i.e., dilute) and Brouwer (i.e., two majority carriers) conditions, an adequate expression is arrived at for any charge carrier concentration as a function of the control parameters (temperature T, doping content, component partial pressure P), namely

$$c_j(T,P) = \alpha_j^{\beta_j} P^{N_j} \prod_r K_r(T)^{\gamma_{rj}} \qquad [12]$$

($N_j, \gamma_{rj}, \alpha_j, \beta_j$ being simple rational numbers; the total pressure is assumed to be constant.) For multinary compounds (i.e., m components) at complete equilibrium $(m-1)$, component partial pressures are to be considered. This solution is only a sectional solution within a window in which the majority carriers situation (Brouwer condition) does not change. The power-law equation defines van't Hoff diagrams characterized by

$$-\frac{\partial \ln c_j}{\partial 1/RT} = \sum_r \gamma_{rj}\Delta_r H^\circ \qquad [13]$$

according to which, for not too extended T-ranges (i.e., $\Delta_r H^\circ$, the reaction enthalpy of the defect reaction γ, is constant), straight lines are observed in the ln c_j versus $1/T$ representation, as well as Kröger–Vink diagrams characterized by straight lines in the plots ln c_j versus ln P with slopes

$$\frac{\partial \ln c_j}{\partial \ln P} = N_j \qquad [14]$$

Figure 4 displays the defect chemistry within the phase width of the oxide MO. At low oxygen partial pressure (P_{O_2}), vacant oxygen ions ($V_O^{\bullet\bullet}$) and

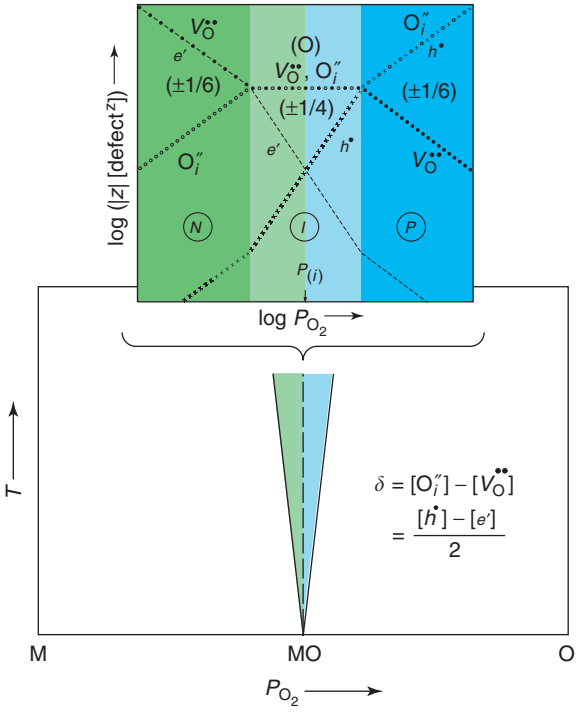


Figure 4 Internal redox and acid–base chemistry in a solid $MO_{1+\delta}$ within the phase width at a given temperature T (disorder in the M-sublattice assumed to be negligible).

reduced states (e') are in the majority ($2[V_O^{\bullet\bullet}] \simeq [e'] \gg [O_i''], [h^\bullet]$) (*N*-regime) whilst oxygen interstitials (O_i'') and oxidized states ($h^\bullet$) dominate for very high P_{O_2} ($2[O_i''] \simeq [h^\bullet] \gg [e'], [V_O^{\bullet\bullet}]$) (*P*-regime). At intermediate P_{O_2}, usually ionic disorder prevails ($[V_O^{\bullet\bullet}] \simeq [O_i''] \gg [e'], [h^\bullet]$) (*I*-regime); the alternative situation that electronic disorder predominates ($[e'] \simeq [h^\bullet] \gg [V_O^{\bullet\bullet}], [O_i'']$) is usually scarce. Exactly at the Dalton composition (i.e., at $P_{O_2}^{(i)}$ where $2[O_i''] = 2[V_O^{\bullet\bullet}] = [e'] = [h^\bullet]$), the "nonstoichiometry" δ in $MO_{1+\delta}$ (i.e., $[O_i''] - [V_O^{\bullet\bullet}] = \frac{1}{2}[h^\bullet] - \frac{1}{2}[e']$) is precisely zero. For $\delta \gg 0$, a *p*-conductor (mobility of $h^\cdot$ is usually much greater than for O_i'') is referred to, for $\delta \ll 0$, one refers (mobility of e' is usually much greater than for $V_O^{\bullet\bullet}$) to an *n*-conductor, while at $\delta \simeq 0$, mixed conduction is expected (only if the ionic concentrations are much larger than the electronic ones, pure ion conduction results). Usually, the entire diagram is not observed, as there are limits of experimentally realizing extreme P_{O_2} values as well as limits with respect to the phase stability (formation of higher oxides or lower oxides, if not of the metal).

Figure 5 gives three examples: SnO_2 as an oxygen-deficient material, La_2CuO_4 as an example of a *p*-type conduction, and PbO as an example of *I*-regime exhibiting mixed conduction. (In PbO, most probably, the counterdefect to O_i'' is $Pb_i^{\bullet\bullet}$ rather than $V_O^{\bullet\bullet}$; the results, however, are not different then.) The slopes directly follow from simple mass action considerations.

The model example of AgCl shall be considered in more detail. It also exhibits predominant ionic disorder, but now the decisive disorder reaction is the Frenkel reaction of the Ag sublattice (i.e., the formation of silver ion vacancy, V'_{Ag}, and interstitial silver ion, $Ag_i^\bullet$). The role of P_{O_2} is played by the chlorine partial pressure. The mass action law for the interaction with Cl_2 reads $[Ag_i^\bullet]P_{Cl_2}^{+1/2}[e'] = \text{constant}$, the mass action constant for the Frenkel reaction being $[Ag_i^\bullet][V'_{Ag}] = K_F$ and that for the band–band transfer $[e'][h^\bullet] = K_B$. Since $[Ag_i^\bullet] \simeq [V'_{Ag}] \gg [e'], [h^\cdot]$, it follows as solution $[V'_{Ag}] = [Ag_i^\bullet] = \sqrt{K_F}$, $[e'] \propto P_{Cl_2}^{-1/2}$, $[h^\bullet] \propto P_{Cl_2}^{+1/2}$.

Doping Effects

Dopants, as irreversibly introduced structure elements such as Cd^{2+} substituting Ag^+ to form the

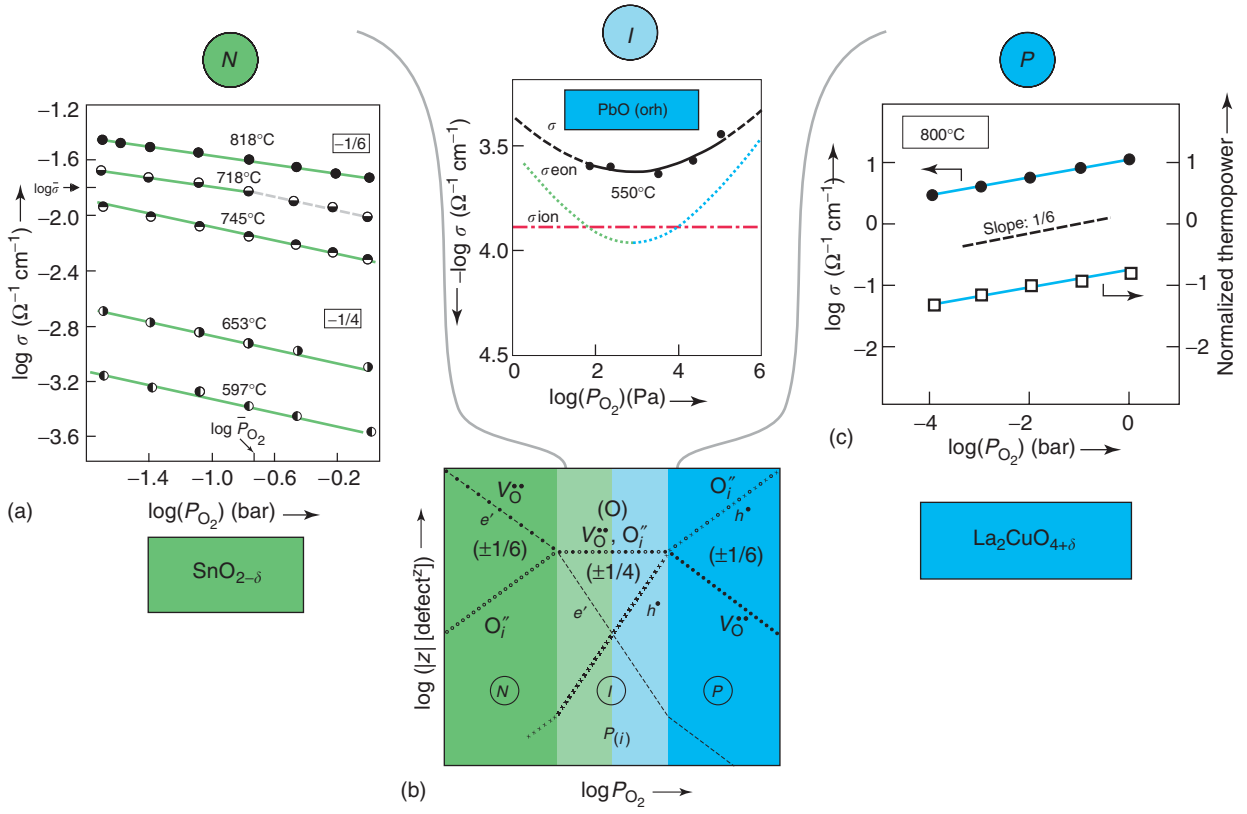


Figure 5 Three experimental examples of Kröger–Vink diagrams in a pure oxide MO with ideal defect chemistry. (a) SnO_2 as an *n*-type conductor, (b) PbO as a mixed conductor, and (c) La_2CuO_4 as a *p*-type conductor. (Reproduced with permission from Maier J (2003) Ionic and mixed conductors for electrochemical devices. *Radiation Effects and Defects in Solids* 158: 1–10; © Taylor and Francis.)

defect $Cd_{Ag}^{\bullet}$, do not significantly influence the mass action laws, but appear in the electroneutrality relation, here

$$[Ag_i^{\bullet}] + [Cd_{Ag}^{\bullet}] \simeq [V'_{Ag}] \quad [15]$$

Coupling the above relation with the Frenkel equation $[Ag_i^{\bullet}][V'_{Ag}] = K_F$ leads to

$$[V'_{Ag}] = \frac{C}{2} + \sqrt{\frac{C^2}{4} + K_F} \quad [16]$$

$$[Ag_i^{\bullet}] = -\frac{C}{2} + \sqrt{\frac{C^2}{4} + K_F} \quad [17]$$

where C stands for $[Cd_{Ag}^{\bullet}]$. Consider the Brouwer conditions again. For $C \ll \sqrt{K_F}$, the intrinsic result $[V_{Ag}^{\bullet}] = [Ag_i^{\bullet}] = \sqrt{K_F}$ is obtained. For $C \gg \sqrt{K_F}$, one arrives at

$$[V_{Ag}^{\bullet}] = C \quad [18]$$

$$[Ag_i^{\bullet}] = K_F C^{-1} \quad [19]$$

which are immediately obtained by neglecting $[Ag_i^{\bullet}]$ from the electroneutrality relation. Power-law equations are then also valid for the electronic minority carriers in AgCl.

Figure 6 displays the solutions for the ionic defect arrived at by the more accurate equations for both concentrations and conductivities. The different behaviors of defect concentrations and conductivities stem from the fact that the interstitials are more mobile than the vacancies. The introduction of the dopant increases the concentration of the oppositely charged defect and depresses the concentration of the counterdefect according to electroneutrality and mass action. This leads to a decreased ionic conductivity, before the counterdefect starts defining the overall conductivity. The conductivity minimum lies approximately at $\sqrt{K_F u_i / u_v}$.

Figures 7 and **8** display the dependence of conductivity on temperature in pure samples and on the doping concentration at given T with regard to positive and negative doping. While the response to Cd^{2+} doping ($Cd_{Ag}^{\bullet}$) follows exactly the predictions (see **Figures 7** and **8**), the effect of S^{2-} doping (S'_{Cl})

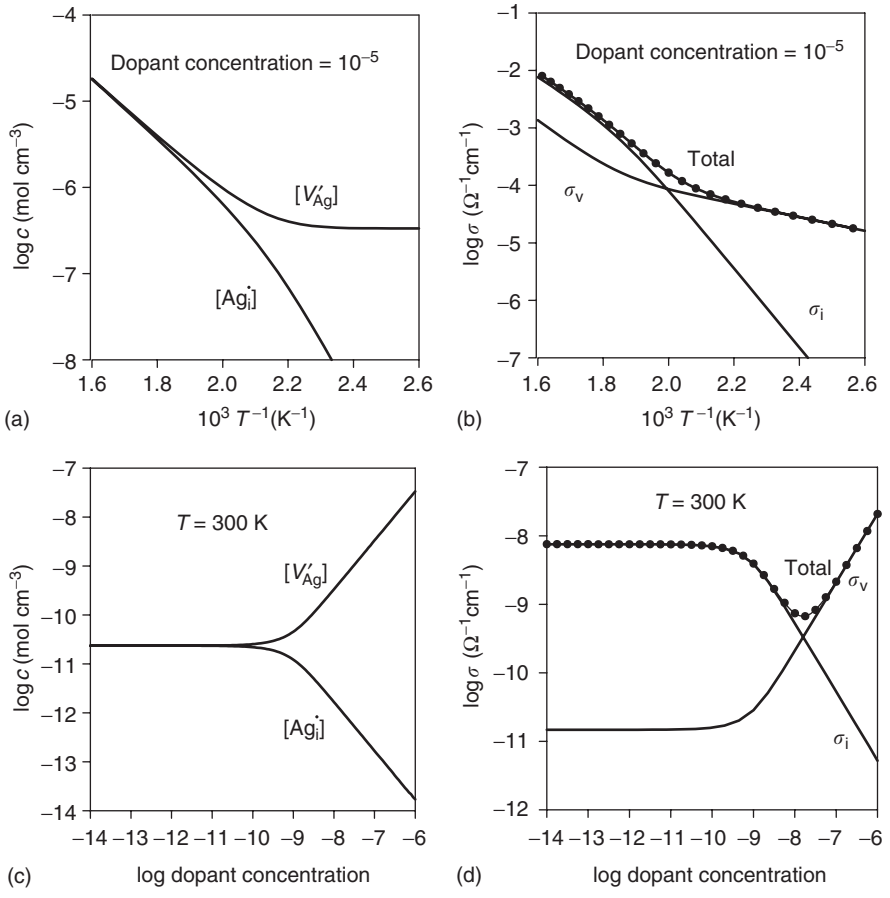


Figure 6 Defect concentrations and conductivities calculated as a function of temperature for a fixed Cd-content ((a) and (b)) and as a function of Cd-content for a fixed temperature ((c) and (d)).

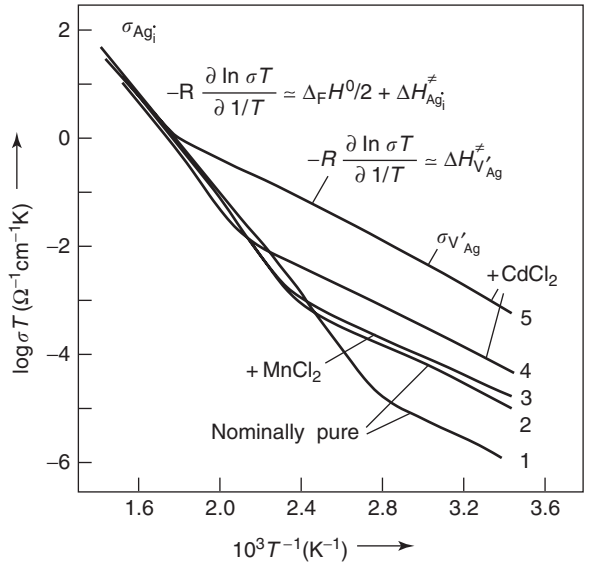


Figure 7 Experimental conductivity data for nominally pure and doped AgCl as a function of $1/T$. Here $\log(\sigma T)$ is plotted instead of $\log \sigma$ to take account of the slight T-dependence of the prefactor. However, this does not alter the slope noticeably ($\Delta_F H^0$: formation enthalpy of Frenkel defects; $\Delta H_j^{\neq}$: migration enthalpy). (Reproduced from Corish J and Jacobs PWM (1972) Ionic conductivity of silver-chloride single-crystals. *Journal of Physics and Chemistry of Solids* 33: 1799–1818, with permission from Elsevier.)

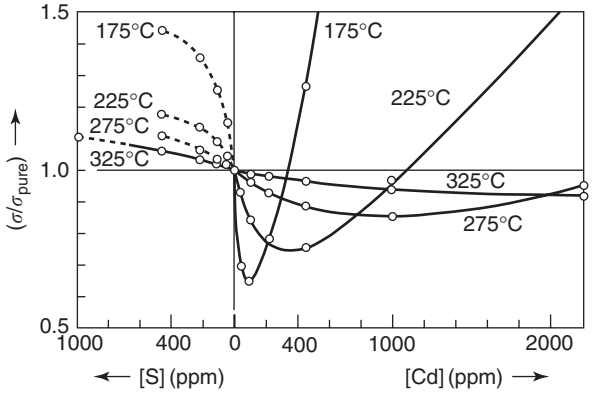


Figure 8 The dependence of the conductivity increase brought about by the impurities on the S and Cd content of AgBr. The RHS solid curves are calculated according to the ideal defect chemistry. (Reproduced with permission from Teltow J (1950) Zur Ionen-, Elektronenleitung und Fehlordnung von Silberbromid mit Zusätzen von Silber-, Cadmium- und Bleisulfid. *Zeitschrift für Physikalischs Chemie* 195: 213–224; © Oldenbourg Wissenschaftsverlag and Teltow J (1949) Zur Ionenleitung und Fehlordnung von Silberbromid mit Zusätzen zweiwertiger Kationen. I. Leitfähigkeitsmessungen und Zustandsdiagramme. *Annalen der Physik* 6: 63–70; © Wiley-VCH.)

suffers from interaction effects (see below, the absence of a minimum in **Figure 8**, LHS, is in qualitative agreement with the fact that $[Ag_i^\bullet]$ is increased).

It is clear that for doped samples under Brouwer and Boltzmann conditions, the concentration of any carrier is not only a function of P, T but also of C and reads

$$c_j(T,P,C) = \alpha_j^{\beta_j}\left(\prod_p P_p^{N_{pj}}\right) C^{M_j}\left(\prod_r K_r(T)^{\gamma_{rj}}\right) \qquad [20]$$

(P^{N_j} is replaced by $\prod_p P_p^{N_{pj}}$ in order to allow for multinary equilibria labeled by p). Since unlike (the *in situ* parameters) P and T, the variation of C requires a new high temperature preparation, C is designated as an *ex situ* parameter.

Now this also predicts sectionally constant slopes in diagrams of the type $\log C_j$ versus $\log C$ according to

$$\frac{\partial \ln c_j}{\partial \ln C} = M_j \qquad [21]$$

For a simple defect chemistry, the sign of M_j is determined by ("rule of heterogeneous doping")

$$\frac{z_j M_j}{z_d} < 0 \qquad [22]$$

which means that, for a positive (negative) doping, the concentration of all negative (positive) defects is increased and all positive (negative) defects are decreased (z_d: charge number of dopant defect). A negative doping is also the reason for change in the slope from the intrinsic value $-1/6$ to the extrinsic value $-1/4$, in **Figure 5a** for SnO_2, which directly results from the mass action law for oxygen incorporation, that is, $P_{O_2}^{1/2}[V_O^{\bullet\bullet}][e']^2 = \text{constant}$ (for $2[V_O^{\bullet\bullet}] \simeq C$, one obtains $\sigma \simeq \sigma_n \propto [e'] \propto P_{O_2}^{-1/4}$ whereas for $2[V_O^{\bullet\bullet}] = [e']$, the result is $\sigma \propto P_{O_2}^{-1/6}$).

High Charge Carrier Contributions — Interactions

The above solutions are valid for ionic and electronic carrier concentrations but require low concentrations. One way of correcting interactions is to introduce associates, that is, novel particles formed as the result of these interactions and responding differently to electrical fields or other driving forces. In this way, the scheme of defects is rescaled with the consequence of writing ideal chemical potentials (i.e., interaction free) for these to a better approximation. Examples are interactions of dopants with electronic or ionic carriers (these states appear as midgap states in **Figure 1**), associates between electronic carriers (excitons, Cooper-pairs), or between ionic carriers (such as vacancy pairs, interstitial pairs, and Frenkel associates). As an example, the association between $Cd^\bullet$ and V'_{Ag} is mentioned that leads to a thermally

activated process at low temperatures in the case of $CdCl_2$ doped AgCl (lower than those shown in **Figure 7**).

The more general approach is to include interactions via analytical corrections in the chemical potentials (μ^{ex}) of the charge carriers. (On the level of the concentration, this leads to the activity coefficient as a correction factor.) If the high carrier concentration is a consequence of the high temperatures, such a procedure is indispensable. If it is a consequence of a high doping content, it may be replaced by or better combined with the association procedure. Interactions lead to entropic corrections as well. One of them is obviously due to the exhaustibility of sites (Fermi–Dirac correction). The more detailed account of deviations from a random distribution owing to interactions leads to very complex expressions (e.g., by the Mayer theory for real gases). Fortunately in the case of weak interactions, the neglect of such entropic corrections is a tolerable approximation. Then, besides possible local vibrational entropies, only energetic nonideality corrections need to be considered. The Debye–Hückel theory (considering each central ion embedded into a diffusive cloud of counter ions, the size being characterized by λ) leads, in the first approximation, to an expression of the form

$$\mu_j^{ex} = RT \ln f_j = -\frac{z_j^2 F^2}{8\pi\varepsilon N_m \lambda} \propto c_j^{1/2} \qquad [23]$$

that reasonably corrects the chemical potential for low defect concentrations. The Debye–Hückel theory proved particularly helpful for electrolytic solutions.

At higher concentrations, the corrections become not only complicated but also individual. Here, an *ad hoc* model that is able to describe interactions in some simple binary crystals and estimate these interactions by assuming that the interaction energy (viz. the structure) can be assessed by a virtual Madelung-superlattice of defects of the mean lattice constant $\propto c^{1/3}$ is briefly discussed. The prefactor J in $\mu^{ex} = -Jc^{1/3}$ is determined by the Madelung energy of the ground lattice, the Madelung constants of lattice defect, and ground lattice as well as the dielectric constant.

The attractive interaction of oppositely charged carriers has the consequence that the formation becomes increasingly easier at high temperatures as soon as the carriers perceive each other in ionic crystals. The conductivity increases in an over-Boltzmann fashion (premelting) and eventually the avalanche effect leads to a phase transformation into the superionic (or molten) state. The order of the transition can be calculated by comparing the formation enthalpy, the formation entropy, and the interaction content.

Figure 9 shows the "thermal destiny" of a pure Frenkel disordered binary crystal from the perfect state at low T to the superionic state at high T.

Boundary Layers (Heterogeneous Doping)

At boundary layers, electroneutrality is no longer fulfilled and space charge regions (width $\sim$ Debye length) occur. The reason for this is the adjustment of structurally different regions. Not only are the electronic concentrations modified, but the mobile ionic charge carriers are also distributed according to the space charge field (both "Fermi levels" are positionally constant). Effects on the ionic conductivity can be immense.

Double layers formed by the contact of electrolyte solution to an electrode or charged membrane surfaces in electrophysiology are well-established examples.

Figure 10 depicts four prototype cases involving solid electrolytes: (1) the contact of an ionic conductor to an insulator, the surface of which is able to trap (internally absorb) ions, leading to the concept of heterogeneous doping (influence of conductivity by dispersing fine insulator particles characterized by a similar rule of heterogeneous doping as given above for homogeneous doping obtained by replacing the concentration and charge of the dopant defect by the surface charge density and its sign); (2) the contact of two ionic conductors leading to a charge transfer of ions from one conductor to the other (cf. ionic "*p–n* junctions"); (3) charge can also be stored within that core of the boundary, that is, in a grain boundary which joins two chemically identical grains and the effects can be tuned by chemical modification of the boundary or by varying the misfit; and (4) the contact to a fluid, for example, gas phase, is relevant in this context due to the fact that, acid–base active gases may cause attractive or repelling interactions with mobile ions.

Since in many mixed, even predominantly electronic conductors, ionic charge carriers may be in majority, the ionic redistribution effect may determine the overall electronic boundary concentrations (fellow-traveler effect).

Figure 11 shows the bending of ionic and electronic energy levels and their intercorrelation. The determination of the bending effect requires treatment of the chemical thermodynamics of the contact problem and hence a knowledge of energy levels in the interfacial core. In the case of a strong interfacial effect (c_o: carrier concentration in the first layer adjacent to the interface),

$$\Delta\sigma_m^{\parallel} = 2Fu\sqrt{\frac{2\varepsilon RTc_o}{L}} \qquad [24]$$

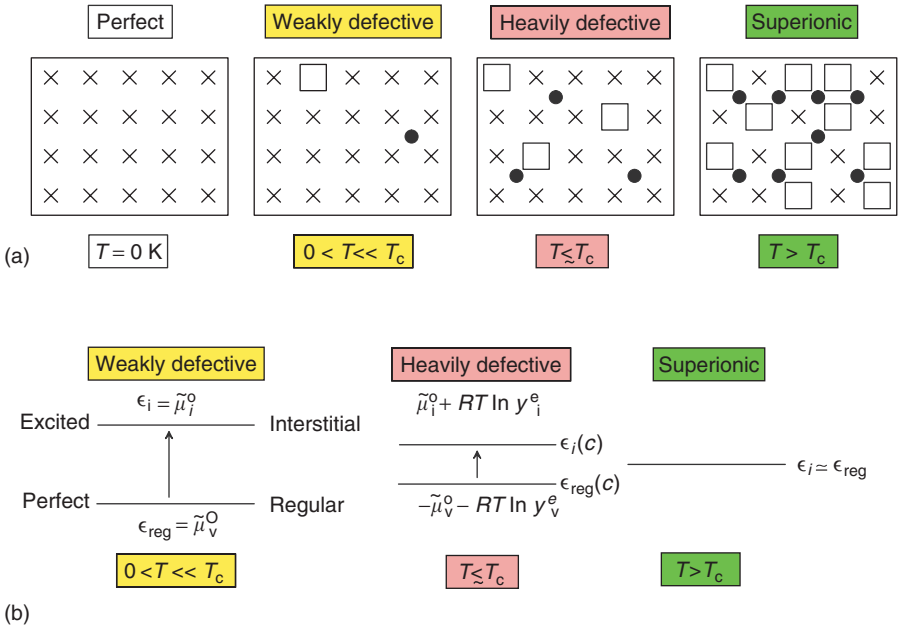


Figure 9 The "thermal destiny" of an ionic crystal above absolute zero: a progressive development from a perfect to a superionic phase via an ideally defective and a heavily defective state. (a) Crystallographic defect picture and (b) corresponding "energy level" schemes. The indices *v* and *i* refer to the vacant regular and interstitial sites. (Reproduced with permission from Maier J and Münch W (2000) Thermal destiny of an ionic crystal. *Zeitschrift für Anorganische and Allgemeine Chemie* 626: 264; © Wiley-VCH.)

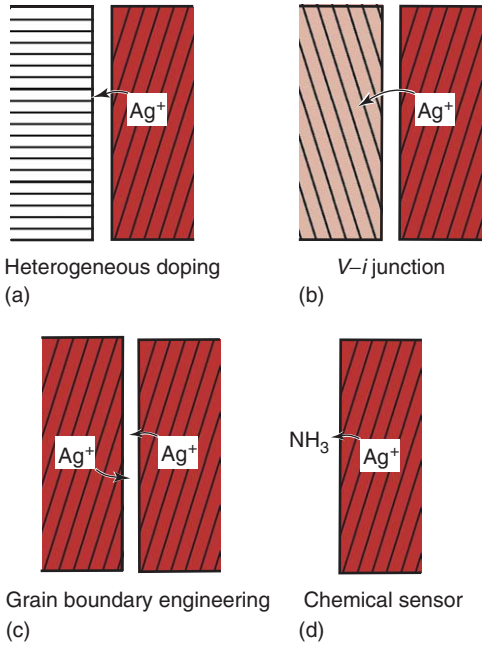


Figure 10 Four basic space charge situations involving ionic conductors (here silver ion conductor): (a) contact with an insulator, (b) contact with a second ion conductor, (c) grain boundary, and (d) contact with a fluid phase. (Reproduced with permission from Maier J (2002) Nanoionics and soft materials science. In: Knauth P and Schoonman J (eds.) *Nanocrystalline Metals and Oxides. Selected Properties and Applications,* Kluwer International Series in Electronic Materials: Science & Technology, vol. 7, pp. 81–110. Boston: Kluwer Academic; © Kluwer Academic/Plenum Publishers.)

describes the excess mean conductivity of the mobile majority defect in a Gouy–Chapman accumulation layer, and

$$\Delta\rho_{\mathrm{m}}^{\perp} = \frac{1}{L}\sqrt{\frac{\varepsilon RT}{2z_{\mathrm{d}}u^2z^3c_{\infty}F^4c_{\mathrm{o}}^2\ln(c_{\mathrm{o}}/c_{\infty})}} \qquad [25]$$

the excess mean resistivity of the mobile majority carriers across a single Mott–Schottky depletion layer situation (dopant with molar charge $z_{\mathrm{d}}F$ and concentration c_{∞}). L is the distance perpendicular to the interface over which the measurement averages.

Figure 12 shows heterolayers of CaF_2 and BaF_2 and the effect of the concentration of interfaces on the overall parallel conductance.

In a polycrystalline material, bulk and boundary conductivities have to be superimposed according to the microstructural situation. An approximate relation that takes into account the conductivities across ($\perp$) and along ($\parallel$) grain boundaries ($\hat{\sigma}$: complex conductivity; $\hat{\sigma}^{\parallel}, \hat{\sigma}^{\perp}$ refer to the local mean values) is

$$\hat{\sigma}_{\mathrm{m}} = \frac{\hat{\sigma}_{\infty}\hat{\sigma}_{\mathrm{gb}}^{\perp} + \beta_{\mathrm{gb}}^{\parallel}\varphi_{\mathrm{gb}}\hat{\sigma}_{\mathrm{gb}}^{\parallel}\sigma_{\mathrm{gb}}^{\perp}}{\hat{\sigma}_{\mathrm{gb}}^{\perp} + \beta_{\mathrm{gb}}^{\perp}\varphi_{\mathrm{gb}}\hat{\sigma}_{\infty}} \qquad [26]$$

($\beta^{\parallel}$, $\beta^{\perp}$ refer to the portions of the volume fraction φ that contribute to the path; in the simplest approach, $\beta^{\parallel} = 2/3$, $\beta^{\perp} = 1/3$.)

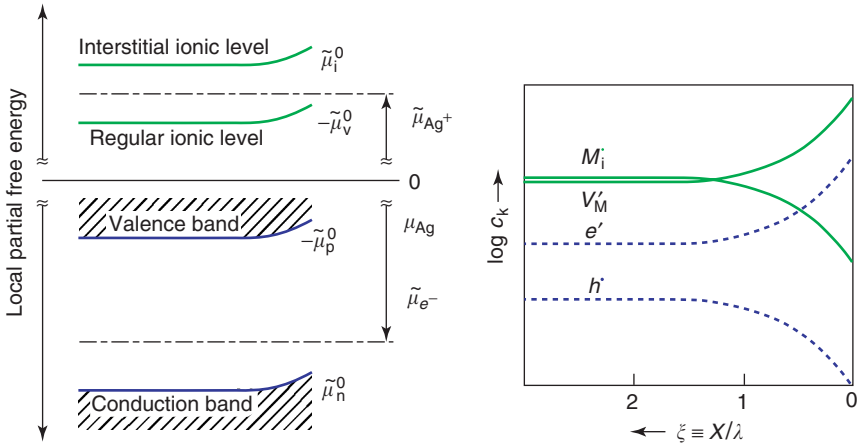


Figure 11 Bending of "energy levels" and concentration profiles in space charge zones ($\xi = 0$ refers to the interfacial edge). (Reproduced from (left) Maier J (2003) Defect chemistry and ion transport in nanostructured materials. Part II. Aspects of nanoionics. *Solid State Ionics* 157: 327 and (right) Maier J (2002) Nano-sized mixed conductors (Aspects of nano-ionics. Part III). *Solid State Ionics* 148: 367–374, with permission from Elsevier.)

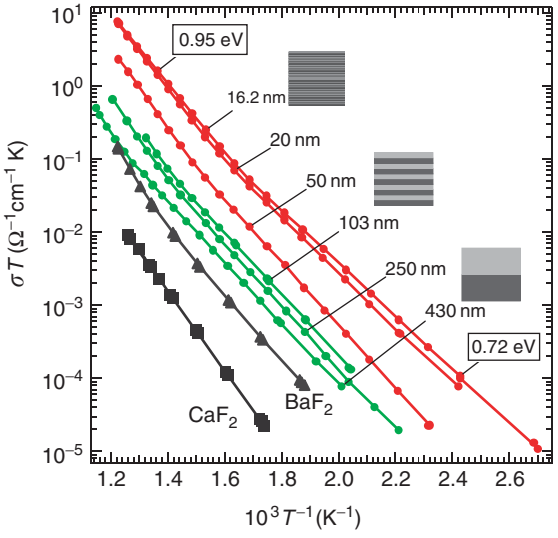


Figure 12 Variation of the parallel ionic conductivity of CaF_2–BaF_2 heterolayers as a function of temperature for different periods (spacings). (Reproduced with permission from Sata N, Eberman K, Eberl K, and Maier J (2000) Mesoscopic fast ion conduction in nanometre-scale planar heterostructures. *Nature* 408: 946–949; © Nature Publishing Group.)

Nano-Ionics

Nano-sized conductors can provide large ionic conductivity anomalies owing to the high concentration of interfaces. Moreover, the spacing of interfaces may be so narrow that they perceive each other. **Figure 13a** shows the space charge overlap occurring in materials when the distance of neighboring interfaces becomes smaller than the Debye length, with the structure being invariant (cf. also **Figure 12**). If the distance is smaller than the effective thickness of the point defect (i.e., point defect including the sphere of influence in which the structure is perceptibly modified), then the local standard chemical potential changes (**Figure 13b**). In the case of delocalized electrons, this may happen at much larger distances. These are only two examples of the nano-size effects on ion conduction. Amongst others, the effects of particle geometry or curvature are mentioned, by which approximately a Gibbs–Kelvin term ($\propto \bar{\gamma}/\bar{r}$; $\bar{\gamma}$: mean surface tension, $\bar{r}$: mean radius) is introduced to the chemical potential.

Such effects do not only affect transport properties, they are also interesting as far as mass storage is concerned.

Partial Equilibrium – Bridge between High-Temperature and Low-Temperature Situation

Arbitrary deviations from equilibrium require a detailed kinetic treatment. The situations in which some carriers behave reversibly while others are completely frozen are briefly considered here. Such considerations are helpful when the bridge has to be spanned from high-temperature equilibrium (during preparation) to a low operational temperature. In fact, such a partial equilibrium situation was already tacitly assumed when doping effects were studied. At high enough temperatures, dopants become reversible and segregation equilibria become active. Generally speaking, freezing of carriers leads to a reshuffling of *in situ* to *ex situ* parameters (i.e., from the *P*-term to the *C*-term in the power-law equation [20]). The discrepancy between the equilibrium

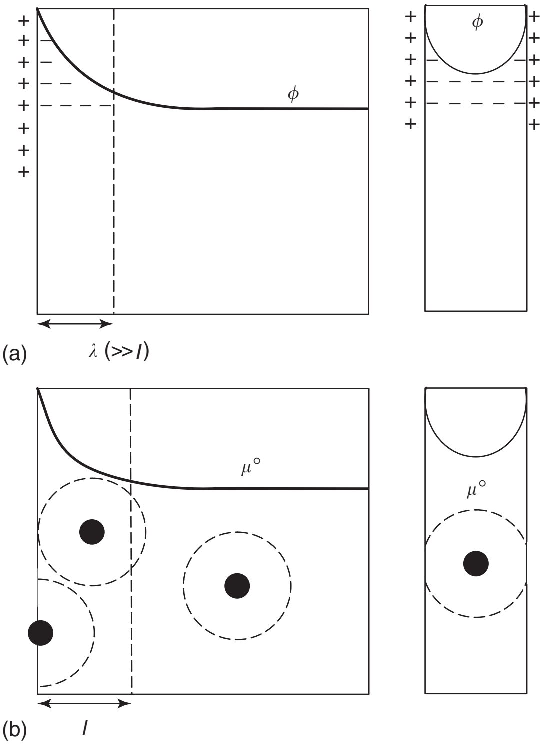


Figure 13 (a) Space charge effects decaying with λ. (b) Structural effects which decay rapidly in the bulk. Within the indicated distance (*l*); the structurally perturbed core (not shown, extension *s*) is perceived by the defects as far as μ° is concerned. (Reproduced with permission from Maier J (2003) Nano-ionics: trivial and non-trivial size effects on ion conduction in solids. *Zeitschrift für physikalische chemie* 217: 415–436; © Oldenbourg Wissenschaftsverlag.)

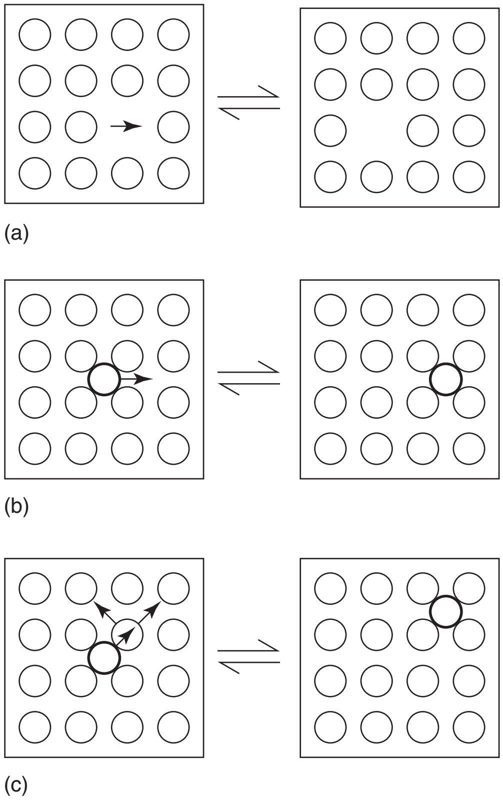


Figure 14 Elementary jump mechanisms in crystals: (a) vacancy mechanism, (b) direct interstitial mechanism, and (c) (collinear or noncollinear), indirect interstitial mechanism (interstitialcy mechanism).

concentration (usually fulfilled at high temperatures under preparation temperatures) and the partially frozen situation becomes apparent in the application of electroceramics (e.g., high-temperature superconductors and dielectrics) that are used at room temperature. Two aspects are important: (1) a reproducible preparation and (2) a detailed understanding of defect chemistry. For a systematic treatment of these complex problems, the reader is referred to the literature.

Mobility: From Low to High Concentrations

For dilute concentrations, the mobility in ionic crystals is usually taken to be independent of the concentration. According to the hopping equation ($u_j \propto k_j$), it follows as

$$u_j \propto \Gamma_{0j}(\Delta x)^2 \exp\left(+\frac{\Delta S_j^{\neq}}{R}\right)\exp\left(-\frac{\Delta H_j^{\neq}}{RT}\right) \qquad [27]$$

(Γ_0 is the attempt frequency (cf. k_0, i.e., the prefactor of k), Δx is the hopping distance; $\Delta H^{\neq}$ and $\Delta S^{\neq}$ denote activation enthalpy and entropy, respectively.)

Figure 14 displays three fundamental hopping mechanisms. While the treatment of the ionic carrier concentrations is, in essence, very similar for liquids or solids, this is less so in the case of mobility. In solids, the activation thresholds of hopping transport are usually substantial (approximately several 100 meV), while in dilute liquid electrolytes, temperature effects act on carrier mobilities mainly through changes in the viscosity. In superionic solid conductors, activation thresholds are significantly less.

At high concentrations, the jump partners in the hopping processes are important which formally leads to the appearance of a concentration factor $(1 - (c/c_{\max}))$ in the mobility. Higher concentrations, however, affect carrier transport in such an individual way that no universal relation can be given and even the decomposition into charge carrier concentration and mobility becomes questionable.

For not too strong interactions, a fairly general interaction effect is to be considered. Defect interactions lead to relaxation phenomena in that a particle

that has made a jump from x to x' does not directly create its equilibrium environment and has a higher tendency to jump back again. It is, in particular, the rearrangement of the other defects that determines the finite relaxation time. This mismatch situation causes back-jumps to be more likely, making the rate constants of the hopping equation time dependent and hence the conductivity frequency dependent.

In liquids too, a relaxation effect occurs. Since the formation of the ion cloud does not immediately follow the mobile ion under regard, the charge cloud becomes asymmetrical and acts as a brake. The relaxation time is proportional to the viscosity. The occurrence of counter transport of differently charged ions which carry along these solvation spheres leads to friction, that is, to a second effect that diminishes the mobility, the so-called electrophoretic effect. These phenomena explain the Kohlrausch's $\sqrt{c}$ law, which is well established for strong electrolytes, and reads

$$\Lambda_c = \Lambda_o - \left[\frac{A}{(\varepsilon T)^{3/2}}\Lambda_o + \frac{B}{(\varepsilon T)^{1/2}}\right]\sqrt{c} \qquad [28]$$

(A and B are constants, $\Lambda_o \equiv \Lambda_c$ $(c = 0)$, ε the dielectric constant, and Λ_c the equivalent conductivity, that is, conductivity per equivalent salt concentration.) **Figure 15** gives some examples of the above phenomena.

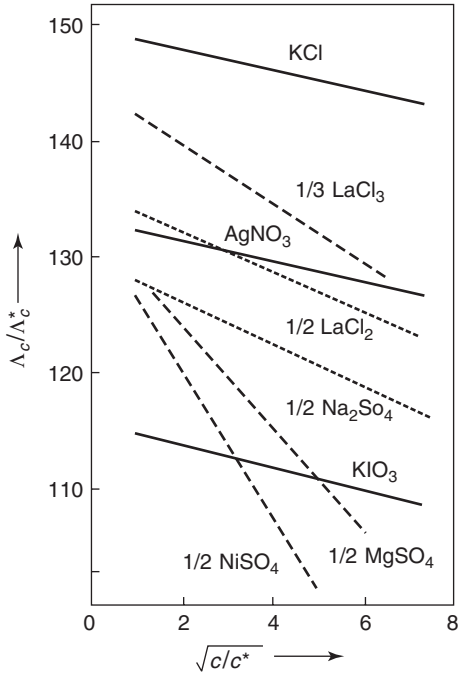


Figure 15 Equivalent conductivity for a variety of aqueous salt solutions vs. concentration (Λ_c^* and c^* denote unit values). (Reproduced with permission from Kortüm G (1972) *Lehrbuch der Elektrochemie*, 5th edn. Weinheim: VCH; © Wiley-VCH.)

Chemical Diffusion in Mixed Conductors

The simultaneous presence of ionic and electronic carriers in so-called mixed conductors leads to the phenomenon of chemical diffusion (coupled transport of ions and electrons) and the possibility of changing stoichiometry. Technologically important examples for which this phenomenon plays a decisive role are insertion electrodes, electrochemical windows, permeation membranes, and conductivity sensors; more generally, chemical diffusion is important for all electroceramics, the properties of which depend on the detailed stoichiometry (such as in dielectrics or high-temperature superconductors). Stoichiometry changes also occur, if the electrodes applied in the electrochemical measurements are not completely reversible for both ions and electrons; in this case, the partial current density is given by

$$i_j = \left(\frac{\sigma_j}{\sigma}\right) i + z_j F D^{\delta} \nabla c_j \qquad [29]$$

D^{δ} being the chemical diffusion coefficient that is determined by mobilities and concentrations of all charge carriers involved in the transport (i: total current density). If different internal reactions such as valence state changes occur for the conducting species, the equation has to be modified. The use of selectively blocking electrodes leads to the suppression of the current of one carrier type at the expense of a concentration polarization. This allows the separation of ionic and electronic conductivities and of the chemical diffusion coefficient.

A major application of ion conductors is their use as electrolytes which nullifies gradients in the electrochemical potential of ions but supports gradients in the electrochemical potential of electrons. This gives rise to a cell voltage (e.m.f.) which is consequently proportional to the integrated gradient of the neutral component (cf. distance in **Figure 2**) and corresponds to the Nernst equation. In other words, differences in the chemical component potentials can be transformed into an electrical voltage (such as exploited in batteries, fuel cells, and sensors). In mixed conductors, a partial internal short-circuit occurs, causing a nonzero gradient in the electrochemical potential of the ions, which is proportional to the gradient in the electronic Fermi level. Hence, the cell voltage is diminished by a factor that depends on the proportion of electronic conductivity and consequently allows for its determination.

See also: Conductivity, Electrical.

PACS: 72.60.+g

Further Reading

Allnatt AR and Lidiard AB (1993) *Atomic Transport in Solids*. Cambridge: Cambridge University Press.

Funke K (1993) Jump relaxation in solid electrolytes. *Progress in Solid State Chemistry* 22: 111.

Heifets E, Kotomin EA, and Maier J (2000) Semi-empirical simulations of surface relaxation for perovskite titanates. *Surface Science* 462: 19.

Kirchheim R (1988) Hydrogen solubility and diffusivity in defective and amorphous metals. *Progress in Materials Science* 32: 261.

Kortüm G (1965) *Treatise on Electrochemistry*, 2nd edn. Amsterdam: Elsevier.

Kröger FA (1964) *Chemistry of Imperfect Crystals*. Amsterdam: North-Holland.

Lidiard AB (1957) Ionic conductivity. In: Flügge S (ed.) *Handbuch der Physik*., vol. 20, p. 246. Berlin: Springer.

Maier J (1993) Defect chemistry: composition, transport, and reactions in the solid state. Part I: Thermodynamics. *Angewandte Chemie (International Edition in English)* 32: 313–528.

Maier J (1995) Ionic conduction in space charge regions. *Progress in Solid State Chemistry* 23: 171.

Maier J (2004) *Physical Chemistry of Materials – Ions and Electrons in Solids*. Chichester: Wiley. (Translation of *Festkörper – Fehler und Funktion: Prinzipien der Physikalischen Festkörperchemie* Teubner BG, Stuttgart).

Mott NF (1974) *Metal-Insulator Transitions*. London: Taylor & Francis.

Rickert H (1982) *Electrochemistry of Solids*, 1st edn. Berlin: Springer.

Sasaki K and Maier J (1999) Low-temperature defect chemistry of oxides. I. General aspects and numerical calculations; II. Analytical reasons. *Journal of Applied Physics* 86: 5422–5443.

Schmalzried H (1981) *Solid State Reactions*, 1st edn. Weinheim: VCH.

Schmalzried H (1995) *Chemical Kinetics of Solids*, 1st edn. Weinheim: VCH.

Ionic Bonding and Crystals*

M W Finnis, Queen's University Belfast, Belfast, UK

Published by Elsevier Ltd.

Introduction

One can distinguish the models of ionic materials by their increasing degrees of sophistication, which are described briefly below. All contain parameters of one sort or another that have to be fitted either to experimental data or to the best available first-principles calculations. Following this introduction there is a more detailed discussion of how ionic models, in general, can all be related to the density-functional theory within the second-order perturbation theory.

The Born model, the shell model, and the compressible ion model are three well-known types of increasing complexity. More recently, they have been joined by models including ionic polarizability up to the quadrupole level. The simplest ionic models are typified by constant nominal ionic charges. These are usually integers, although fractional charges have been used as additional fitting parameters. Variable charge models have also been developed, which allow charge transfer within a simple classical description of the equilibration of an electronic chemical potential. Finally, a class of models has been developed that includes most of the aforementioned as subclasses and also treats covalency, namely, the self-consistent tight-binding models. The scope of ionic models, especially when charge transfer can be included, is potentially very wide. Although so far they have been less developed and applied, self-consistent tight-binding models have potentially the widest scope, which embrace traditional textbook ionic materials such as NaCl, KF and all the other alkali halides, the insulating oxides with wide bandgaps such as MgO and Al_2O_3, perovskites such as $SrTiO_3$, as well as oxides which are considered as less ionic, such as TiO_2 and ZrO_2.

The simplest ionic model is the rigid ion model. It is sometimes called the Born model after its pioneer Max Born and his co-workers in Edinburgh, who over 50 years ago developed most of the mathematical description of the ionic model that is still being used (see Further reading section for more details). Within the context of the density-functional perturbation theory, one can think of it as a first-order model, as seen in the following section. The ions interact by the Coulomb interaction, in addition to which there is a short-range pairwise repulsion.

Ions are not rigid, and the shell model was first developed by Dick and Overhauser in 1958 to include their polarizability. In its simplest version, a spherical shell of charge $-q^s$ surrounds an ion of reduced charge $\Delta Z + q^s$ and is attached to it by a harmonic spring. The nominal ionic charge is ΔZ. The repulsion between ions is now modeled as a repulsion between the centers of these shells.

The compressible ion model is an extension of the shell model in which the radius of the shells is a variable. The energy of a compressible ion is a function of this radius, which adjusts according to the environment of the ion. This model is most strongly motivated by the case of oxygen, as discussed further

*Adapted by permission of Oxford University Press.

below. The traditional models were mainly constructed by fitting their parameters to experimental data. This was a disincentive to make them too sophisticated, since the number of parameters would become so large that they could fit anything without any guarantee of being physically sound. The compressible ion model was held back for this reason, since in its early days, the parameters were determined by fitting them to experimental phonon dispersion curves. The work of Madden *et al*. attempted to reduce the empirical nature of the parameters by fitting them individually, rather than altogether. They were fitted to total energy data obtained by first-principles calculation on molecules and crystals at chosen interatomic distances. This approach also made it feasible to introduce independent parameters to describe quadrupolar polarization. Even so, the most sophisticated ionic models take no explicit account of covalent effects, and this may be a significant shortcoming.

None of the previous ionic models allow for the fact that the charge associated with an ion also depends on its environment. This has been treated in an empirical way by the model introduced by Streitz and Mintmire in 1994. A more recent approach is to use a self-consistent tight-binding model, which, in principle, can be regarded as an ionic model with a degree of covalency included, as described by Finnis *et al*. in 2003. Covalency is represented in the tight-binding model by the nonvanishing elements of a Hamiltonian matrix linking atomic-like orbitals. In fact, most other models can be derived from a tight-binding description, by making specific approximations, perhaps the most drastic being the complete neglect of interatomic Hamiltonian matrix elements as in the variable charge-transfer model, or the neglect of all covalent effects as in the rigid ion model.

The Rigid Ion Model Derived

In order to derive ionic models, like other models, one starts with an input charge density $\rho^{\text{in}}(\boldsymbol{r})$. It is constructed, at least notionally, by superimposing spherical "atomic" charge densities. The atomic charge density in the sense used here is not necessarily the same as the charge density of free atoms. A judicious choice of atomic charge density could make the input charge ρ^{in} closer to the exact, self-consistent charge density ρ^{ex}, improving the accuracy of a first-order model. The usual picture of the charge in an ionic crystal is, of course, a superposition of ions rather than of atoms, and one might expect that these would be a better starting point. In any case, the variational principle suggests that the error in the energy is second-order error in the charge density, provided that the energy is formulated as a function of the charge density.

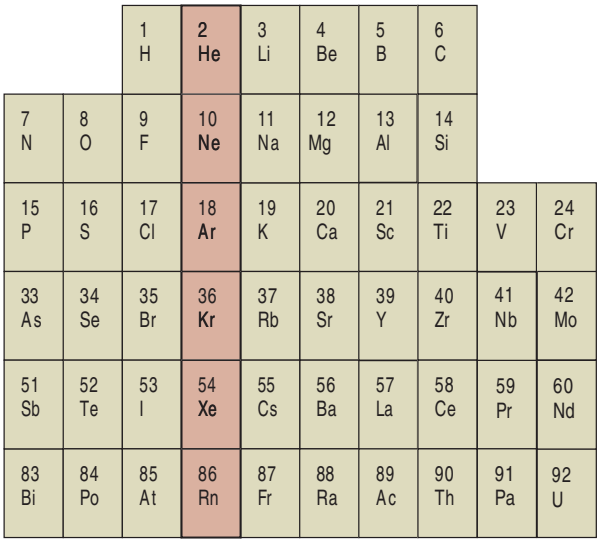


Figure 1 The periodic table centered on the noble gases.

A neat general recipe for the "atomic charge density" is to take the electron density of the nearest noble gas in the same row of the periodic table (see **Figure 1**). Thus, oxygen, fluorine, sodium, magnesium, and aluminum would all be represented by the charge density of neon atoms, chlorine and potassium by argon, and so forth.

This leads one to construct a first-order model for the total energy as follows. First, superimpose the noble gas atoms at the positions required in the material, keeping their charge densities rigid. The superimposed charge density is denoted by $\rho^{\text{in}}_{\text{ng}}$, where the individual noble gas atom on site I has a density ρ^{I}_{ng}. Thus

$$\rho^{\text{in}}_{\text{ng}}(\boldsymbol{r}) = \sum_{I} \rho^{I}_{\text{ng}}(\boldsymbol{r} - \boldsymbol{R}_I) \qquad [1]$$

The first-order energy of this system is denoted by $E^{\text{HKS}}[\rho^{\text{in}}_{\text{ng}},\ V_{\text{ext}}]$, where

$$E^{\text{HKS}}[\rho^{\text{in}}_{\text{ng}}, V_{\text{ext}}] = T_{\text{s}}[\rho^{\text{in}}_{\text{ng}}] + \int \rho^{\text{in}}_{\text{ng}}(\boldsymbol{r}) V_{\text{ext}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + E_{\text{H}}[\rho^{\text{in}}_{\text{ng}}] + E_{\text{xc}}[\rho^{\text{in}}_{\text{ng}}] + E_{ZZ} \qquad [2]$$

The notation introduced here is as follows: E^{HKS} is a density functional that describes the total energy, $T_{\text{s}}[\rho]$ is the kinetic energy of noninteracting electrons with density ρ, V_{ext} is the potential felt by the electrons from the atomic nuclei, $E_{\text{xc}}[\rho]$ is the exchange and correlation energy of the electrons with density

ρ, E_H is the electrostatic self-energy (Hartree energy) and E_{ZZ} is the nucleus–nucleus Coulomb repulsion. The energy of this system is strongly repulsive, because the equilibrium spacing of noble gas atoms is considerably larger than that of the cations and anions in ionic materials. Indeed, because one is dealing with closed shell atoms, it is expected that the first-order charge density and energy are rather good approximations to the exact, self-consistent quantities denoted by ρ_{ng} and E_{ng}. The integer charge differences ΔZ_I on the nuclei, required to create the nuclei of the actual cations and anions, results in a perturbing external potential to the electrons, given by

$$\Delta V_{ext}(\boldsymbol{r}) = -\sum_I \frac{\Delta Z_I}{|\boldsymbol{r} - \boldsymbol{R}_I|} \qquad [3]$$

To first order, this does not disturb the charge density, but it introduces a strong electrostatic cohesive energy, due to the attraction between the unlike charges, which because of their alternating arrangement in space more than compensates for the repulsion of like charges. This works as follows: the total change in the electrostatic interactions of the nuclei introduced by "transmutation" is

$$\Delta E_{ZZ} = \sum_{I \neq J} \frac{Z_I \Delta Z_J}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} + \frac{1}{2} \sum_{I \neq J} \frac{\Delta Z_I \Delta Z_J}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \qquad [4]$$

The second term is called the total Madelung energy of the particular structure, and is denoted as $N_a E_M$, where N_a is the total number of atoms and E_M is the Madelung energy per atom:

$$E_M = \frac{1}{2N_a} \sum_{I \neq J} \frac{\Delta Z_I \Delta Z_J}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \qquad [5]$$

Consider in more detail the example of a simple binary crystal structure, in which a formula unit consists of n^C cations and n^A anions of charges ΔZ^C and ΔZ^A, respectively. These charges derived according to the above definition are sometimes called formal charges, to distinguish them from actual electronic charge transfers. Generally, for any material, summing over all sites, charge neutrality requires that:

$$\sum_I \Delta Z_I = 0 \qquad [6]$$

For the binary crystal, the charge neutrality condition holds for each formula unit:

$$n^C \Delta Z^C + n^A \Delta Z^A = 0 \qquad [7]$$

Table 1 Madelung constants for ionic crystals

Structure	α_M
NaCl	1.75
CsCl	1.76
α-Al_2O_3	1.68
V_2O_5	1.49

Adapted from Johnson QC and Templeton DH (1961) Madelung constants for several structures. *Journal of Chemical Physics* 34: 2004–2007.

In this case, the Madelung energy per atom can be written as

$$E_M = \alpha_M \frac{\Delta Z^C \Delta Z^A}{2R_{CA}} \qquad [8]$$

where R_{CA} is the nearest-neighbor cation–anion distance and α_M is the Madelung constant, which is tabulated in the literature for a number of crystal structures. Examples are given in **Table 1**. For an arbitrary arrangement of ions, for example, within a supercell which may comprise thousands of atoms, or without periodic boundary conditions, there are special strategies for calculating $N_a E_M$, such as the Ewald method for periodic systems, which are well described. The first-order model for the total energy omits the second-order effects due to the deviation of both ρ_{ng} and of the final exact charge density, ρ^{ex} from ρ^{in}_{ng}:

$$\begin{aligned} E^{(1)} &= E^{HKS}[\rho^{in}_{ng}, V_{ext}] + \int \rho^{in}_{ng}(\boldsymbol{r}) \Delta V_{ext}(\boldsymbol{r})\, d\boldsymbol{r} + \Delta E_{ZZ} \\ &= E^{HKS}[\rho^{in}_{ng}, V_{ext}] + N_a E_M \\ &\quad - \sum_{I \neq J} \int \rho^I_{ng}(\boldsymbol{r} - \boldsymbol{R}_I) \frac{\Delta Z_J}{|\boldsymbol{r} - \boldsymbol{R}_J|}\, d\boldsymbol{r} + \sum_{I \neq J} \frac{Z_I \Delta Z_J}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \\ &\quad - \sum_I \int \rho^I_{ng}(\boldsymbol{r} - \boldsymbol{R}_I) \frac{\Delta Z_I}{|\boldsymbol{r} - \boldsymbol{R}_I|}\, d\boldsymbol{r} \end{aligned} \qquad [9]$$

The interaction of the atomic charge density with the perturbation on its own site has been separated out as the final term. This expression is dominated by the first two terms, namely $E^{HKS}[\rho^{in}_{ng}, V_{ext}]$ and the Madelung energy. It is a reasonable approximation to write $E^{HKS}[\rho^{in}_{ng}, V_{ext}]$ as a sum of pairwise interactions V_{IJ}, since closed-shell atoms are being dealt with. With this assumption one obtains the simplest form of the rigid ion model:

$$E_{rim} = \frac{1}{2} \sum_{I \neq J} V_{IJ} + N_a E_M \qquad [10]$$

The remaining first-order terms will be much smaller, as discussed below. If all the atoms happen

to be represented by the same noble gas, the final sum must vanish identically by the condition of charge neutrality. But in any case, it is a structure-independent term that is not relevant to interatomic forces. The third and fourth terms taken together represent the electrostatic interaction between the noble gas atomic densities, which are neutral objects, and the perturbations ΔZ_I, excluding the self-interactions. These are pairwise interactions, that are short-ranged and weak compared to the electrostatic interactions contained in $E^{\mathrm{HKS}}[\rho_{\mathrm{ng}}^{\mathrm{in}}, V_{\mathrm{ext}}]$. They fall to zero by Gauss's theorem when the neighbor J is entirely outside the charge density ρ_{ng}^{I}, which must be a good approximation beyond the nearest neighbors. Furthermore, the nearest neighbors will usually have opposite signs of ΔZ_I and ΔZ_J, so the two contributions $\rho_{\mathrm{ng}}^{I}\Delta Z_J$ and $\rho_{\mathrm{ng}}^{J}\Delta Z_I$ tend to cancel. Indeed, the cancellation is obviously exact if the noble gas is the same for I and J and if $\Delta Z_I = -\Delta Z_J$, such as in binary compounds, for example, MgO or NaF. Now comparing these terms with the pairwise electrostatic contributions to $E^{\mathrm{HKS}}[\rho_{\mathrm{ng}}^{\mathrm{in}}, V_{\mathrm{ext}}]$, one gets

$$\int\int \frac{\rho_{\mathrm{ng}}^{I}(\boldsymbol{r})\rho_{\mathrm{ng}}^{J}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' + \frac{Z_I Z_J}{R_{IJ}} - \int\left\{\frac{\rho_{\mathrm{ng}}^{I}(\boldsymbol{r})Z_J}{|\boldsymbol{r}-\boldsymbol{R}_J|} + \frac{\rho_{\mathrm{ng}}^{J}(\boldsymbol{r})Z_I}{|\boldsymbol{r}-\boldsymbol{R}_I|}\right\}\mathrm{d}\boldsymbol{r} \qquad [11]$$

This interaction will generally be much bigger. Notice that, first, it involves the full nuclear charges Z rather than ΔZ, second, it is always repulsive, because it comprises the direct Coulomb interaction between nuclei, partially screened by the electron distributions, and third, it is of longer range, being nonzero when ρ_{ng}^{I} overlaps ρ_{ng}^{J}, not merely Z_J. The terms being considered are all strictly pairwise, and so can be exactly accounted for by a relatively small correction to the V_{IJ} introduced to represent $E^{\mathrm{HKS}}[\rho_{\mathrm{ng}}^{\mathrm{in}}, V_{\mathrm{ext}}]$. Since the form of V_{IJ} in practical applications has been parametrized, the extra terms could be absorbed without explicitly evaluating them.

The rigid ion model in the form [10] is now considered again. It was first derived on the basis of overlapping noble gas atomic charge densities by Gordon and Kim in 1972 and Kim and Gordon in 1974. A completely local density-functional theory (DFT) (including the local kinetic energy functional) was used to evaluate the pair potential between noble gas atoms, with atomic charge densities overlapped but unrelaxed, and it was discovered that this could account fairly well for the lattice parameters and bulk moduli of ionic crystals. The $E^{\mathrm{HKS}}[\rho_{\mathrm{ng}}^{\mathrm{in}}, V_{\mathrm{ext}}]$ could also be fairly well fitted by the Lennard-Jones potential, the parameters of which have been used, together with the appropriate Madelung energy, to calculate bond lengths in a number of compounds, as illustrated in **Figure 2**, where the results are compared with experimental bond lengths. Because some of the crystal structures are a bit complicated, convenient simplifications have been made, so that the bond length can be predicted by minimizing the energy analytically with respect to R_{CA}. For example, only the nearest-neighbor interactions in the Lennard-Jones potential were included. Then for Al_2O_3, the sum of ionic radii for Al^{3+} and O^{2-} was taken as the "experimental" bond length, namely 0.191 nm, which is also the mean of the two close neighbor bond lengths observed in the actual corundum structure. Such ionic radii incidentally reproduce the experimental bond lengths of the other common oxides to an accuracy of 0.002 nm. Notice that in some cases, a mixture of rare gases is involved, for example, Nb_2O_5 has to be built out of Kr and Ne, because Nb in the 5+ state of ionization resembles Kr. Since a value for the Madelung constant was not listed in this case by Johnson and Templeton, it is assumed that their V_2O_5 value would do for Nb_2O_5. The agreement of calculated and observed bond lengths in such diverse materials as ZrO_2 and KCl, with no fitting of parameters to any experimental data, is quite striking. Not surprisingly, the worst candidate tested here is SiC, which is a classic covalent compound, but even in this case the crude ionic model does better than one might have expected, since the perturbations ΔZ_I are by no means small. Such is the power of the variational principle. It should be added that the

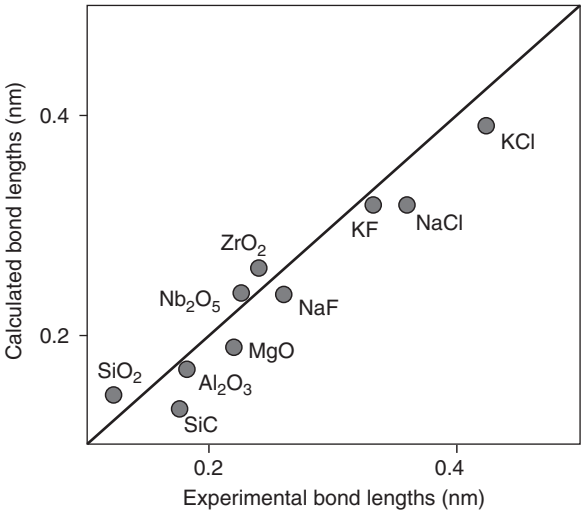


Figure 2 Bond lengths calculated from Lennard-Jones and Madelung energies.

relatively good agreement of this model breaks down when it comes to calculating the second derivative of the energy, the bulk modulus. Superficially, that is easily repaired if one allows the pair potential to be empirically fitted. Predictions of phonon frequencies and the properties of defects nevertheless generally require a higher-quality model than rigid ions can provide, even when some degree of empirical fitting is allowed.

Beyond the Rigid Ion Model

The Basic Second-Order Model

Several other models can be derived within the broad framework of a second-order functional, which can be written in terms of the reference density from the noble gas atoms as

$$\begin{aligned} E^{(2)}[\rho] = {} & E^{\mathrm{HKS}}[\rho_{\mathrm{ng}}^{\mathrm{in}}, V_{\mathrm{ext}}] + \int \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})\Delta V_{\mathrm{ext}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ & + \Delta E_{ZZ} + \int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\Delta V_{\mathrm{ext}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ & - \frac{1}{2}\int\int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\chi_{e1}^{-1}(\boldsymbol{r},\boldsymbol{r}')[\rho(\boldsymbol{r}') \\ & - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r}')]\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' \end{aligned} \quad [12]$$

The linear response function $\chi_{e1}(\boldsymbol{r},\boldsymbol{r}')$, of which the inverse operator $\chi_{e1}^{-1}(\boldsymbol{r},\boldsymbol{r}')$ appears here, measures the density induced at $\boldsymbol{r}$ by a delta function of the external potential at $\boldsymbol{r}'$. Recall that the first term represents the energy of the noble gas solid to first order, and the idea is that the above rigid ion model has taken care of this and the next two terms, that is, all terms of first order in ΔV_{ext}. The task is then to approximate the two remaining terms that depend on $[\rho - \rho_{\mathrm{ng}}^{\mathrm{in}}]$ and that are both of second order in ΔV_{ext}. The second-order energy functional in this picture is therefore

$$\begin{aligned} E^{(2)}[\rho] = {} & \frac{1}{2}\sum_{I\neq J} V_{IJ} + N_{\mathrm{a}}E_{\mathrm{M}} \\ & + \int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\Delta V_{\mathrm{ext}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ & - \frac{1}{2}\int\int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\chi_{e1}^{-1}(\boldsymbol{r},\boldsymbol{r}')[\rho(\boldsymbol{r}') \\ & - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r}')]\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' \end{aligned} \quad [13]$$

The difference between this functional and the usual second-order functional is that the one-electron energies are here subsumed into the classical pairwise potential, an approximation that is only justifiable for the closed-shell electronic structure of noble gas atoms. By this manoeuvre, the entire responsibility has been shifted for the description of any covalent bonding effects, which are of second order in $[\rho - \rho_{\mathrm{ng}}^{\mathrm{in}}]$, onto χ_{e1}. The response function χ_{e1} refers to an electron density distribution $\rho_{\mathrm{ng}}^{\mathrm{in}}$, and it can only be modeled in rather crude ways. Whatever way this is done (an example is given below), the model so generated is entirely classical in form, and requires $E^{(2)}[\rho]$ to be minimized directly with respect to ρ, without the solution of any Schrödinger equation. The resulting density will then satisfy

$$\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r}) = \int \chi_{e1}(\boldsymbol{r},\boldsymbol{r}')\Delta V_{\mathrm{ext}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \quad [14]$$

Deformable Ions

The above formalism is a way to derive the shell model, which is the simplest extension of the rigid ion model to introduce deformable ions. Originally, this model was entirely empirical, providing sufficient disposable parameters to fit phonon spectra much better than previous models did. Now from the previous analysis, one is in a better position to see how its parameters can be related to real physical quantities. The shell model is, in effect, a representation of the terms in $\rho - \rho_{\mathrm{ng}}^{\mathrm{in}}$:

$$\begin{aligned} \Delta E^{(2)}[\rho] = {} & \int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\Delta V_{\mathrm{ext}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ & - \frac{1}{2}\int\int [\rho(\boldsymbol{r}) - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r})]\chi_{e1}^{-1}(\boldsymbol{r},\boldsymbol{r}')[\rho(\boldsymbol{r}') \\ & - \rho_{\mathrm{ng}}^{\mathrm{in}}(\boldsymbol{r}')]\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' \end{aligned} \quad [15]$$

In this case, the density $\rho - \rho_{\mathrm{ng}}^{\mathrm{in}}$ of eqn [13] is modeled by the sum of displacements by $\boldsymbol{u}_I^{\mathrm{s}}$ of rigid, spherical shells of negative charge, $-\Delta q_I^{\mathrm{s}}$ originally centered about each atom. Each shell is tethered to its nucleus by a spring. At equilibrium in a perfect crystal, an undisturbed shell can be thought of as representing part of the density $\rho_{\mathrm{ng}}^{\mathrm{in}}$. The shells are assumed to be massless, since the inertia of electrons does not enter the dynamics of the system in the Born–Oppenheimer approximation, which is the framework for all the models. When the atoms are not in their equilibrium positions, the shells are subject to three kinds of forces, corresponding to three terms in the total energy. First, they are pulled toward their nuclei by the tethering spring, second, they are pulled toward other nuclei by the Coulomb attraction, and third, the shells repel each other. The first two kinds of forces are a representation of the term $\int(\rho - \rho_{\mathrm{ng}}^{\mathrm{in}})\Delta V_{\mathrm{ext}}$ in eqn [13]. In the harmonic approximation, the change in energy when the shells

are displaced by $\boldsymbol{u}_I^s$ from their nuclei has the form

$$\Delta E^{(2)} = \frac{1}{2}k u_I^{s2} + \sum_{J\neq I}\left\{-q_I^s \Delta Z_J\left[\frac{1}{|\boldsymbol{R}_I + \boldsymbol{u}_I^s - \boldsymbol{R}_J|} - \frac{1}{|\boldsymbol{R}_I - \boldsymbol{R}_J|}\right] + \frac{1}{2}q_I^s q_j^s\left[\frac{1}{|\boldsymbol{R}_I + \boldsymbol{u}_I^s - \boldsymbol{R}_J - \boldsymbol{u}_J^s|} - \frac{1}{|\boldsymbol{R}_I - \boldsymbol{R}_J|}\right]\right\} \quad [16]$$

A number of points can be made about this model. First, it captures the long-ranged Coulomb interactions between induced dipoles. The on-site elements of χ_{e1}^{-1} are represented within the first term $\frac{1}{2}ku_I^{s2}$, while the last term in the braces represents its intersite elements. The Coulomb interactions in [16] could be expanded to first order in $\boldsymbol{u}_I^s$ to display the dipole–dipole interactions explicitly, and indeed the model could be formulated in terms of variable dipoles on the sites instead of explicit shell charges and displacements. These Coulomb interactions are an important part of the inverse response function, trying to be modeled. However, they are certainly not the whole story, since the true inverse response function contains terms in the kinetic energy and the exchange and correlation energy. A short-ranged empirical potential between the shells can be introduced to model these contributions.

In a more realistic model, the shells are allowed to change their radius in response to their environment. Such an effect can be accomplished by including a kind of compressibility of the ion as a further parameter, which adds a functional dependence on $(r^s - r_0^s)$ to the potential energy, where r^s and r_0^s are the current and relaxed values of the shell radius. This is the idea behind the "breathing shell" model, or more recently the "compressible ion" model. For recent developments in the compressible ion model, the reader is referred to the "Further reading" section. The compressibility of an ion has a significant effect on properties, and is most noticeable in the case of oxygen, because the size of the O^{2-} ion is particularly sensitive to its environment. Indeed, in free space the O^{2-} ion is unstable, which can be thought of as the limiting case of the radius of the ion going to infinity as its coordination is reduced to zero. A noteworthy example of the importance of compressible oxygen is MgO. Within the rigid ion model the Cauchy relation between elastic constants, $C_{12} = C_{44}$, holds, since the potential energy is purely pairwise, yet experimentally C_{12} and C_{44} differ by $>50\%$. The shell model does not help here, because by symmetry the elastic strains do not displace the shells from their lattice sites; therefore, the Cauchy relation should still hold. The difference between C_{12} and C_{44} is only obtained when compressible oxygen ions are introduced. This is because the elastic constant C_{12} is associated with a change in volume, which induces compression or expansion of the ions if the model allows them this degree of freedom. It may help to appreciate this if C_{12} is considered as a linear combination of the bulk modulus, $B = \frac{1}{3}(C_{11} + 2C_{12})$ and the shear modulus, $C' = \frac{1}{2}(C_{11} - C_{12})$, that is, $C_{12} = \frac{1}{3}(3B - 2C')$.

There are now more sophisticated models with polarizable ions, going beyond the notion of spherical shells and harmonic springs. Closer in spirit to the density-functional framework is the approach of Ivanov and Maksimov in 1996. These authors represent $(\rho - \rho_{\mathrm{ng}}^{\mathrm{in}})$ as a superposition of atomic charge densities ρ_{ng}^I, each of which is allowed to shift rigidly with respect to its nucleus, thereby modeling the dipole moments. The associated changes in energy are calculated with a completely local density functional along the lines of Gordon and Kim. A natural extension, considered by Ivanov and Maksimov, is to include parameters that describe distortions of the atomic charge densities, such as expansion and contraction or quadrupolar distortion. Besides the local density approximations for kinetic, exchange, and correlation energies, the only limitation of this approach is the number of free parameters one is prepared to include to characterize the charge density.

Another line of attack starts as before with the assumption that the ions are polarizable as dipoles and quadrupoles, and are compressible. The strategy is then, as far as possible, to obtain the many parameters such a model requires as independently as possible, by fitting functions to an extensive set of first-principles calculations on ions in specially chosen potential wells or in clusters. It is a great improvement on the simple shell model, and can reproduce a range of strain energy data and phonon spectra. However, in spite of the extensive use of first-principles calculations, it is impossible to eliminate some arbitrary choices regarding the functional forms used, and some pairwise approximations. A detailed example of this approach was given by Marks *et al*. in 2001.

Variable Charge Transfer Models

In the models described so far in this article, a fixed charge transfer between ions is defined. The net charges on ions are fixed at the formal charges, which are specified by the electronic structure of noble gases. In reality, the occupancy of atomic orbitals varies with the local atomic environment, so it is worth thinking about how this effect can be

modeled. In fact, a model of ionic crystals that includes covalency, polarizable ions (dipoles and quadrupoles), and variable charge transfer was introduced by Finnis *et al*. under the heading of self-consistent tight binding. Self-consistent tight binding offers a way of treating pure metal atoms as well as their oxides, since, at least for *sd*-bonded metals, it models the nature of chemical bonding from the metallic to the ionic state.

Conceptually, the application of self-consistent tight binding to an ionic material is no different from its application to metals. However, an alternative input charge density might be considered. In the tight-binding models discussed previously, atomic charge densities were superimposed to create the notional input charge density, so that all the long-ranged electrostatic interactions appeared in the second-order term. However, one could also formulate the tight-binding model starting from the superimposed atomic charge densities of noble gases, in which case the Coulomb interactions would appear in the pair potential, exactly as they do in the rigid ion model. A systematic study of the consequences of one or the other starting point has not yet been made.

In spite of advantages of self-consistent tight binding in comparison to the models introduced earlier in this article, it has certain disadvantages that should be mentioned. These include the obvious one of greater computation time, but there may be others. For example, in current implementations the second-order term is represented by Coulomb interactions between point multipoles. This appears to take no account of the effect that the compressible ion model is trying to capture. The logical way to describe a compressible oxygen ion within tight binding would be, to include more orbitals of *s* and *p* character in the basis set, so that the spread of the charge density on an anion could vary by transfer of electrons between the orbitals on a site. The price to pay would be more parameters to determine and more computation time. The benefits over a first-principles tight-binding method would be diminished or eliminated.

On the other hand, notions of compressible ions and charge transfer are not absolute; a transfer of electrons from cations to anions might be rather well described in terms of the charge density by a contraction of anion densities accompanied by an expansion of cation charge densities, and vice versa. The tight-binding model predicts a violation of the Cauchy relations because of its noncentral forces and covalency, and possibly charge transfer, whereas the compressible ion model predicts the Cauchy violation from the form of its classical nonpairwise potential energy. A one-to-one mapping of these properties such as charge transfer between the compressible ion model and the tight-binding model is not unique, so, at least at the time of writing, it would be unfair to dismiss one or the other description as wrong; they may both be grasping at the same physics from different points of view.

It may be noted that the classical variable charge model is also the basis of a popular empirical model that includes charge transfers. The determination of ρ is considered as a variational problem for which the solution equalizes a classical chemical potential of the electrons everywhere. The result is a model that has some similarities to that proposed by Ivanov and Maksimov in 1996. One can think of it as a second-order ionic model with charge transfer, or as a tight-binding model in which the effect of covalency (intersite matrix elements of the Hamiltonian) is neglected. This is a model that enables large-scale molecular dynamics simulations to be carried out, giving a broad-brush description that may be adequate, if details such as the relative energies of different structures are not of interest.

See also: Crystal Structure; Electronic Structure (Theory): Molecules; van der Waals Bonding and Inert Gases.

PACS: 61.50.Ah; 61.50.Lt

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Orlando: Harcourt Brace.

Born M and Huang K (1954) *Dynamical Theory of Crystal Lattices*. Oxford: Clarendon Press.

Dick BG and Overhauser AW (1958) Theory of the dielectric constants of alkali halide crystals. *Physical Review* 112: 90–103.

Finnis M (2003) *Interatomic Forces in Condensed Matter*. Oxford: Oxford University Press.

Finnis MW, Paxton AT, Methfessel M, and van Schilfgaarde M (1998) The crystal structure of zirconia from first principles and self consistent tight binding. *Physical Review Letters* 81: 5149–5152.

Gordon RG and Kim YS (1972) Theory for the forces between closed-shell atoms and molecules. *Journal of Chemical Physics* 56: 3122–3133.

Harrison WA (1980) *Electronic Structure and the Properties of Solids*. San Francisco: W.H. Freeman.

Hohenberg P and Kohn W (1964) Inhomogeneous electron gas. *Physical Review* 136: B864–B871.

Ivanov OV and Maksimov EG (1996) Generalized variational approach to Kim–Gordon electron gas theory for ionic crystals. *Solid State Communications* 97: 163–167.

Kim YS and Gordon RG (1974) Theory of binding of ionic crystals: application to alkali-halide and alkali-earth-dihalide crystals. *Physical Review B* 9: 3548–3554.

Kittel C (1976) *Introduction to Solid State Physics*, 5th edn. New York: Wiley.

Kohn W and Sham LJ (1965) Self-consistent equations including exchange and correlation effects. *Physical Review* 140: A1133–A1138.

Marks NA, Finnis MW, Harding JH, and Pyper NC (2001) A physically transparent and transferable compressible ion model for oxides. *Journal of Chemical Physics* 114: 4406–4414.

Streitz FH and Mintmire JW (1994) Electrostatic potentials for metal-oxide surfaces and interfaces. *Physical Review B* 50: 11996–12003.

Wilson M, Huang YM, Exner M, and Finnis MW (1996) Transferable model for the atomistic simulation of Al_2O_3. *Physical Review B* 54: 15683–15689.

Ziman JM (1972) *Principles of the Theory of Solids*, 2nd edn. Cambridge: Cambridge University Press.

Ionizing Radiation Effects in Biological Materials

A Ottolenghi and F Ballarini, Università degli Studi di Pavia, Pavia, Italy

Introduction

The action of ionizing radiation in biological targets is a stochastic, multistep process involving several orders of magnitude, both in the space scale and in the time scale. A schematic representation of the main steps is reported in **Figure 1**. The main effects at the subcellular/cellular level and tissue/organ level are discussed in detail in the following sections; an overview is provided herein.

Primary energy deposition, which can be considered completed at 10^{-15} s after irradiation ("physical stage"), produces a spatial distribution of ionized and excited molecules. Although the DNA occupies only $\approx 2\%$ of the cell nucleus volume, it is the most important target for ionizing radiation. However, energy deposition in the surrounding water also plays a significant role, since during the so-called "pre-chemical stage" (from 10^{-15} to 10^{-12} s after irradiation) the dissociation of ionized and excited H_2O molecules gives rise to free radicals (e.g., the OH radical) that can attack the double helix constituents and produce DNA damage, mainly during the "chemical stage" (from 10^{-12} to 10^{-6} s after irradiation, when an intra-track chemical equilibrium can be assumed).

DNA damage, either produced by direct energy deposition in the double-helix atoms ("direct damage"), or produced by free-radical attack ("indirect damage"), is processed by specific repair enzymes. The oxygen level in the environment is an important parameter, because O_2 can also link damaged sites

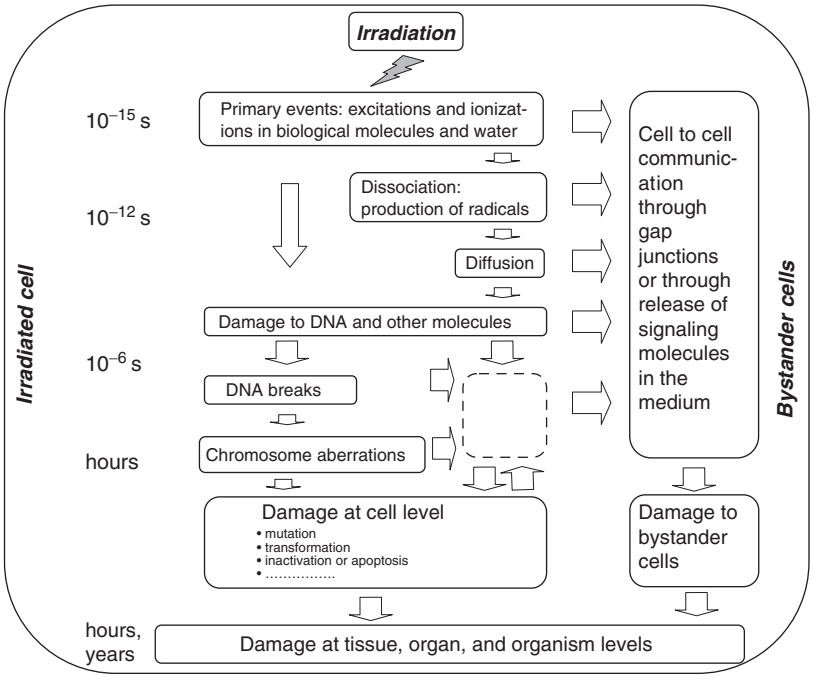


Figure 1 Schematic representation of the main steps characterizing the induction of biological damage by ionizing radiation, from initial energy deposition to different endpoints.

preventing repair and thus acting as a damage fixer. This has implications for radiation therapy, since hypoxic tumors are particularly radioresistant. In the case of incorrect DNA repair, the initial damage can evolve within a few hours into endpoints such as gene mutations and chromosome aberrations, which in turn may lead to cell death, or conversion to malignancy and possibly cancer. While cell death and conversion to malignancy can be detected *in vitro* after a few days or weeks, cancer is characterized by long latency times of the order of several years. To have a quantitative idea of the yields of the various damage types, one may consider that, for example, a cell nucleus irradiated with 1 Gy of gamma rays will be affected, on average, by about 100 000 ionizations (2000 in the DNA), 1000 DNA single-strand breaks (ssb), 40 double-strand breaks (dsb), 1 chromosome aberration, and 10^{-5} specific gene mutations. The probability of cancer development is far less than 10^{-5}.

Since the double-helix diameter is ~ 2 nm, a knowledge of the radiation track structure at the nanometer level is an essential prerequisite for understanding the mechanisms underlying the processes of interest. The expression "track structure" refers to a set of information including, for each energy deposition event (typically, an ionization or excitation), spatial coordinates of the event, event type, and amount of deposited energy. As an example, **Figure 2** shows the inelastic-scattering mean free path in water for electrons of different energies, as implemented in the track-structure code PARTRAC. The minimum value is ≈ 2 nm (that is the DNA diameter), at electron energies between 100 and 200 eV; this explains the high effectiveness of ultrasoft X-rays.

The higher-order organization characterizing mammalian DNA has to be taken into account as well. Each DNA molecule winds around globular proteins ("histones") to form structures called nucleosomes, which in turn are packed to form a 30 nm diameter chromatin fiber. The fiber is organized as a succession of loops that, for most of the cell cycle duration, are localized within distinct domains called chromosome territories, with linear dimensions $\approx 1\,\mu$m. During the 1990s, when the advances in confocal microscopy allowed more detailed investigation of nuclear architecture, it became clear that DNA damage processing mainly occurs in small channels separating neighboring domains. Therefore, the spatial distribution of the initial energy depositions needs to be considered not only at the nanometer scale (i.e., the double helix level), but also at the micrometer scale (i.e., the chromosome level).

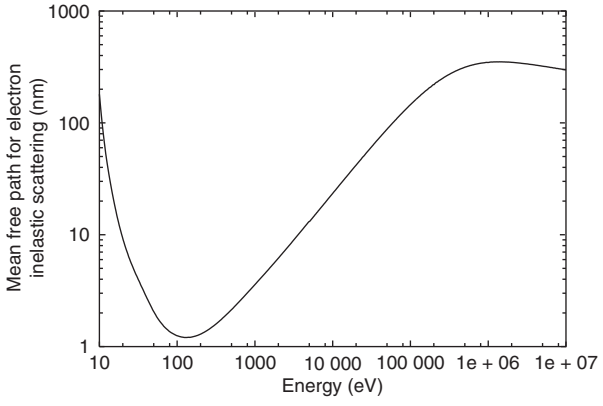


Figure 2 Mean free path for electron inelastic scattering in liquid water as a function of the initial electron energy, as implemented in track-structure code PARTRAC.

Initial DNA Damage

Ionizing radiation can induce DNA damage either by direct energy deposition, or by free-radical attack. Breaks of the DNA strands, especially dsb, are widely recognized for playing a fundamental role in the evolution of the initial damage. The sugar–phosphate moiety is the main target for the formation of strand breaks, although recent studies showed that base damage can also play a role. Energy depositions ≈ 10 eV in the sugar–phosphate can induce an ssb, either via direct dissociation, or via hydrogen loss and subsequent breakage of the sugar–phosphate bond. The sugar–phosphate is the most important target for indirect damage also, which mainly arises following an attack by OH radicals created during water radiolysis. Of the OH radicals reacting with the DNA, 80% react with a nucleobase, whereas the remaining 20% react with the sugar–phosphate. The probability that a reaction OH-DNA leads to an ssb has been estimated as 0.13.

While ssb are generally repaired efficiently, thanks to the complementarity of the two DNA strands, dsb repair is more error-prone and thus dsb are generally considered as critical initial lesions. Conventionally, a dsb is given by two ssb on opposite strands within a few base pairs (bp). A maximum separation of 10 bp is generally assumed, corresponding to a distance of 3.4 nm. Although it has been suggested that a dsb may arise also from a single energy deposition or OH attack due to charge or radical transfer, the most reliable theory on dsb formation requires two radiation-induced ssb, each of them produced either by direct energy deposition or by OH attack. The dsb yields increase linearly with dose, and 40 dsb $\mathrm{Gy}^{-1}\,\mathrm{cell}^{-1}$ is the value generally accepted for mammalian cells. The sensitivity with respect to dsb induction depends both on the target structure

and on the environment scavenging capacity. For example, plasmids are more sensitive than intact mammalian cells due to the lack of protection provided by histones and chromatin folding against OH attack. Furthermore, high concentrations of OH scavengers, which have a high probability of reacting with free radicals, lead to lower yields of ssb and dsb. Due to the clustering properties of high-LET tracks, the role of DNA structure and environment scavenging capacity becomes less important at high-LET, where indirect damage is much less important with respect to low-LET radiation. Many available experimental studies do not show significant dependence of dsb induction on the radiation LET. However, these results may be biased by the impossibility of detecting fragments smaller than a few tens kbp, and by the randomness assumption adopted in fragment analysis. Models based on track-structure simulations, which are not affected by fragment-size limitations, generally suggest a slight dsb increase with LET up to $\approx 100\,\mathrm{keV\,\mu m^{-1}}$. Recent experimental data obtained with the fragment counting technique are consistent with these expectations.

Effects at Subcellular Level

Gene Mutations

Gene mutations consist of alteration or loss of the genetic material within relatively small regions of the DNA base sequence. Generally, mutations are not lethal and result in the expression of an altered protein or in the increase/decrease in the level of a normal protein. This can lead to modifications of cellular functions including loss of proliferation/differentiation control and possibly cell conversion to malignancy. Mutations can occur either by loss of a DNA fragment ("deletion") or by change of the base sequence; changes involving only a few base pairs are called "point mutations." The simplest example of mutation is given by the alteration of a single nucleobase, which may arise from changes in the chemical structure of the base itself.

Although radiation-induced mutations are not qualitatively different from those produced by other agents (e.g., chemicals), ionizing radiation is a particularly effective mutagenic agent. Furthermore, radiation tends to produce a higher proportion of deletions relative to point mutations. While the mutation background level for a given gene is $\sim 10^{-7}$–10^{-6} per surviving cell, *in vitro* radiation-induced mutations are of the order of at least 10^{-5} per survivor. The order of magnitude of the "doubling dose," that is, the dose able to induce a doubling of the spontaneous mutation rate, has been estimated to be $\sim 1\,\mathrm{Sv}$. However, it has to be taken into account that different genes are characterized by different radiosensitivities.

Generally *in vitro* dose–response curves for mutations increase linearly up to a maximum, which is followed by a saturation region due to increased cell killing. Since mutation induction does not show significant variations with the dose rate, mutations are considered as one-hit events. Mutation yields generally increase with LET up to $\approx 100\,\mathrm{keV\,\mu m^{-1}}$, and subsequently tend to decrease. While RBE–LET relationships for mutations are qualitatively similar to those for clonogenic inactivation (see below), the relative biological effectiveness (RBE) values for mutations are higher (~ 10 at $100\,\mathrm{keV\,\mu m^{-1}}$).

Following *in vivo* exposure, mutations can affect both somatic cells and germ cells. In the former case, the consequence is generally the formation of a clone of mutated cells in the exposed individual, whereas germ cell mutations can lead to mutations in the offspring.

Chromosome Aberrations

Cells irradiated before DNA replication can show chromosome aberrations (CA), which arise from large-scale chromatin rearrangements and appear in metaphase as juxtaposition of chromatin fragments originally belonging to different chromosomes, or to different regions of the same chromosome. The product of a single unrejoined chromosome break is called "terminal deletion." Two breaks on the same chromosome can result in a "ring," whereas two breaks on distinct chromosomes can produce either a "dicentric" or a "translocation." All aberrations involving at least three breaks and two chromosomes are called "complex exchanges."

With conventional solid staining (Giemsa), all chromosomes have the same color and scoring is based on the chromosome shape and number of centromeres. Therefore, only dicentrics, rings, and a few complex exchanges, as well as excess acentric fragments, can be scored. The introduction of fluorescence *in situ* hybridization (FISH) in the early 1980s represented a fundamental turn, since it allows selective painting of specific homolog pairs. Besides dicentrics, rings, and excess acentrics, FISH allows scoring of translocations and most complex exchanges. **Figure 3** shows a translocation involving chromosomes 10 and 13 scored with multi-FISH, where each homolog pair is painted with a specific probe.

Translocations are of particular interest since they are specifically correlated with various cancer types and thus they can be used to estimate cancer risk. A typical example is the translocation involving the

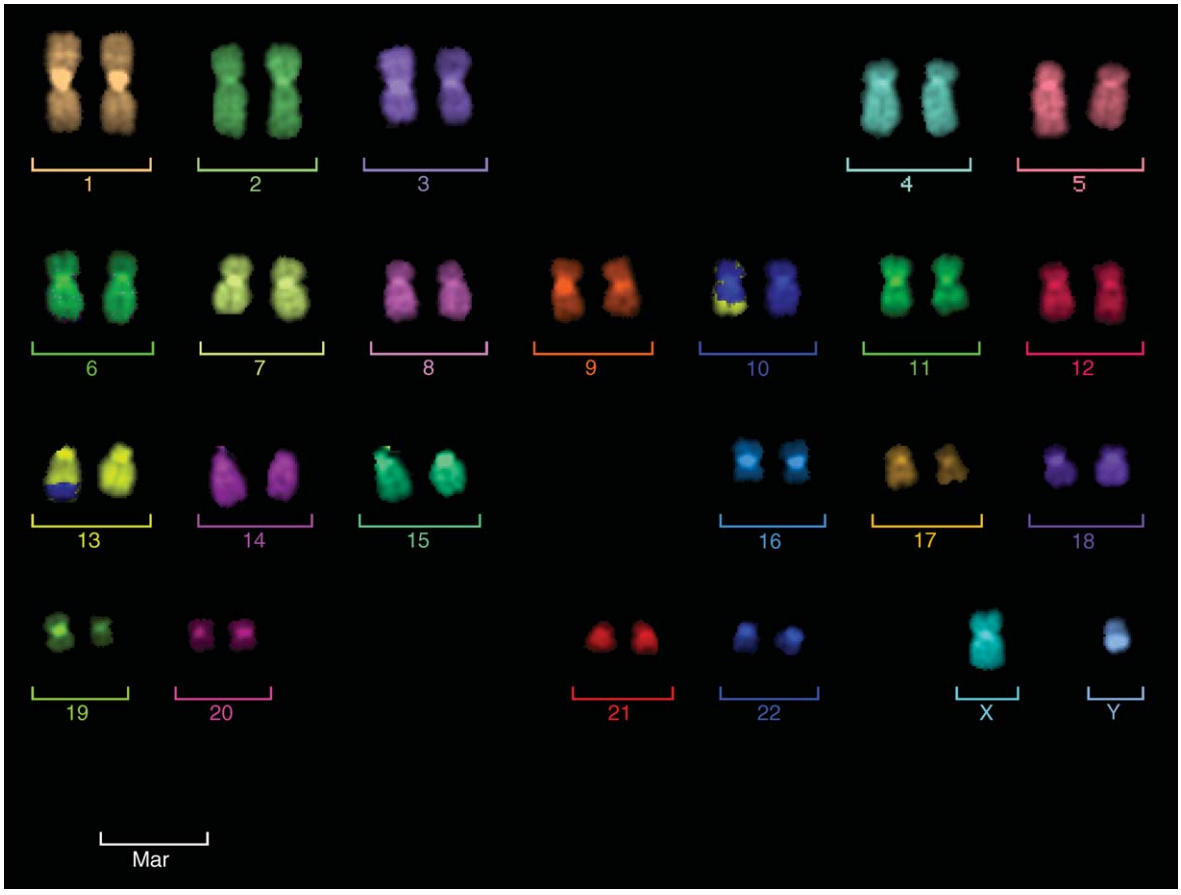


Figure 3 Translocation between one copy of chromosome 10 (painted blue) and one copy of chromosome 13 (painted yellow) following irradiation of human lymphocytes and application of the multi-FISH painting technique. (Image kindly provided by Prof. Marco Durante.)

ABL and BCR genes (on chromosomes 9 and 22, respectively), which is observed in most chronic myeloid leukemia (CML) cells. CML is one of the best-known radiation-induced cancer types and has been observed in A-bomb survivors. The interest in complex exchanges mainly relies on that they can be regarded as biomarkers of the radiation quality that can have applications in radiation protection problems such as radon exposure.

In vitro dose–response for exchange-type CA can be approximated by a linear-quadratic relationship. Low-LET dose responses show a pronounced curvature, whereas at high LET the quadratic component is generally not significant. Back in 1946, before the DNA structure was discovered, Lea hypothesized that chromosome exchanges are due to pairwise misrejoining of independent chromosome free-ends close in space and time, each free end being originated by a chromosome break ("breakage-and-reunion theory," B&R). By contrast, according to Revell's "exchange theory," all aberrations are produced by pairwise interaction of chromosome "discontinuities" arising from "unstable lesions," which can either decay or become reactive. While both these theories require two radiation-induced chromosome lesions, it has also been hypothesized that an exchange may arise from a single radiation-induced lesion, the second lesion being produced by enzymes involved in DNA repair ("one-hit hypothesis"). Although none of these theories can be disregarded at the moment, the B&R model seems to better agree with the available data with the only exception of ultrasoft X-ray data, for which the one-hit mechanism may play a non-negligible role.

The mechanisms underlying CA formation have not been fully elucidated yet. One of the key points is the nature of the initial DNA lesions that can evolve into aberrations: while many research groups think that any dsb can participate in CA formation, others assume that only clustered – and thus severe – dsb can be involved. This assumption mainly relies on those clustered dsb which show a strong dependence on the radiation quality, consistent with the dependence shown by gene mutation and cell inactivation.

Understanding the relationship of CA with other endpoints, such as cell death and conversion to malignancy, is one of the most challenging tasks in this field.

Effects at Cellular Level: Cell Death

Radiation-induced cell death occurs via two main pathways, that is interphase death and clonogenic inactivation. While the former does not allow the cell to reach the first post-irradiation mitosis, the latter consists of a loss of proliferative ability. *In vitro* clonogenic survival is generally assumed when the cell can divide at least six times. The damage produced in the case of interphase death can manifest as swelling of the nucleus and nucleoli, as well as DNA depolymerization. Different cell types can show very different radiosensitivity; lymphocytes seem to be particularly sensitive. Interphase death is closely related to apoptosis, a form of programmed cell death characterized by a specific sequence of morphologic events (detaching from neighboring cells, chromatin condensation, nucleus fragmentation, cell shrinking, and eventually separation into a number of fragments called apoptotic bodies). Apoptosis is highly cell-type dependent and is controlled by different genes including p53. Several aspects of the mechanisms underlying interphase death and apoptosis are still unclear.

Clonogenic inactivation has been studied much more extensively. A large number of *in vitro* data is available, especially on HeLa cells (human cervical carcinoma), CHO cells (hamster ovary cells), and V79 cells (hamster lung fibroblasts). These studies represent a fundamental basic knowledge for applications in radiation therapy, where the main aim consists of achieving tumor control by inactivating the cells constituting the tumoral tissue; the mean lethal dose for loss of proliferative capacity is generally lower than 2 Gy. The results of *in vitro* survival studies are usually graphed as the logarithm of the surviving fraction against the radiation dose. While survival curves obtained by irradiating viruses and bacteria are represented by straight lines, the curves relative to most mammalian cells tend to show a shoulder at low doses. Dose–response for mammalian cell survival is generally well described by a relationship of the form $S(D) = \exp(-\alpha D - \beta D^2)$. Although during the last few years deviations from this behavior have been observed, both as low-dose hypersensitivity and as increased radioresistance, the linear quadratic relationship is still considered a good description for survival fractions not smaller than $\approx 10^{-3}$.

The shoulder is modulated not only by the cell type, but also by the radiation quality. For instance, mammalian cells irradiated with high-LET radiation are generally characterized by straight survival curves. The shoulder is considered to be indicative of repair of "sublethal damage," that is, there must be an accumulation of damage before the cell loses its reproductive ability. Clonogenic inactivation correlates very well with specific chromosome aberrations (e.g., dicentrics and centric rings, see above), suggesting that the major cause of radiation-induced inactivation of mammalian cells is aberrant mitosis leading to uneven distribution of chromosomes or loss of chromosome fragments.

Most cell survival data come from *in vitro* studies. However, when techniques became available to perform *in vivo* experiments, the parameters characterizing the dose–response were similar to those found *in vitro*. In clinical radiotherapy, the radiation dose is usually delivered in a series of equal fractions separated by an interval time allowing for repair of sublethal damage. The resulting effective dose survival curve becomes an exponential function of dose, since repetition of the shoulder leads to a straight line in the lin-log scale.

Effects at the Level of Tissues and Organs

In vivo exposure, which can lead to damage at the tissue/organ and organism level, can be either acute (e.g., in accidents) or chronic, which is often the case for exposed workers and populations. The exposure can affect either the whole organism ("panirradiation") or only parts of it ("partial irradiation"). Two main types of effects are generally introduced: deterministic effects, which occur following exposure above a certain threshold dose and increase in severity by increasing the radiation dose, and stochastic effects, for which there is no threshold and the induction probability (*not* the severity) increases with dose.

Skin erythema is an example of deterministic effect, whereas cancer is the most important stochastic effect. Cataracts are generally classified as deterministic effects. However, recent studies on astronauts suggest that they are more likely to be stochastic processes. Acute total body exposure to 1.5–2 Gy causes nausea and vomiting in 50% of individuals within a few hours ("prodromal radiation sickness"), as well as depletion of granulocytes and platelets within a few weeks ("acute radiation syndrome"). With only minimal supportive care, mortality may occur following acute total body exposure of a few Gy. The acute skin dose necessary to produce erythema is ~ 6 Gy, whereas that for desquamation is in the range 15–20 Gy. In case of gonadal irradiation,

sterility can occur, either transient (following 0.5 Gy for males and ~1 Gy for females) or permanent (following doses in the range 2.5–4 Gy for males and 6–20 Gy for females).

The A-bomb survivors of Hiroshima and Nagasaki represent the most important group studied for radiation-induced cancer, since about 120 000 persons have been followed. By 1990, 6000 cancer deaths had occurred, of which about 400 were ascribed to radiation. Other examples of radiation-induced cancer in humans include leukemia in ankylosing spondylitis patients treated with spine radiotherapy, thyroid cancer in children irradiated for enlarged thymus, and breast cancer in fluoroscoped tuberculosis patients. The time interval between exposure and cancer appearance is called the latent period. While leukemia is characterized by the shortest latent period (7–12 years), solid tumors show longer latency (10–50 years).

Ionizing radiation can induce different cancer types. Acute and chronic myeloid (not lymphocytic) leukemia are the types mainly responsible for the excess cancer incidence observed in irradiated adults, for which the main data come from A-bomb survivors and ankylosing spondylitis patients. A-bomb survivors, together with Chernobyl victims, also provide the main data for thyroid cancer. Breast cancer can be induced by radiation with relatively high frequency; the corresponding data, mainly derived from female A-bomb survivors and tuberculosis patients, are generally well fitted by a straight line. Radiation-induced lung cancer has been observed following both external exposure (A-bomb survivors and ankylosing spondylitis patients) and radon inhalation (mainly underground miners; there is also some evidence from domestic exposure). The largest body of data relative to bone cancer comes from young persons employed as dial painters, who ingested radium, and from tuberculosis or ankylosing spondylitis patients injected with Ra-224. Interestingly, the dial-painter data imply that a linear extrapolation to low doses would overestimate the low-dose risk.

Current risk estimates are mainly based on A-bomb survivor data, which refer to acute exposure to relatively high doses (from 0.2 to 4 Gy) of mainly low-LET radiation. The excess incidence of cancer was assumed to be a function of dose, age at exposure, and time since exposure. For some tumors, such as breast cancer, sex dependence was considered as well. On this basis, the relationship between excess relative risk for solid tumors and equivalent dose is linear up to 3 Sv and flattens out at higher doses due to increased cell killing. The same relationship for leukemia at low doses is best fitted by a linear-quadratic relationship. However, a linear no-threshold (LNT) extrapolation to low doses has been recommended for radiation protection purposes. A "dose rate effectiveness factor" (DREF) reducing the estimated risk by a factor of 2 under conditions of low dose rate or protracted exposure has also been introduced.

According to the International Commission on Radiological Protection, the fatal cancer lifetime risk for a working population has been estimated as $0.08\,\mathrm{Sv}^{-1}$ for high doses and dose rates, and $0.04\,\mathrm{Sv}^{-1}$ for low doses and dose rates. The corresponding values for the whole population are slightly higher ($0.10\,\mathrm{Sv}^{-1}$ for high doses and dose rates, and $0.05\,\mathrm{Sv}^{-1}$ for low doses and dose rates) due to the higher sensitivity of young people. Since epidemiological studies on human exposure to high-LET radiation are not available, high-LET risk estimates are based on low-LET estimates multiplied by radiation weighting factors. The only available systematic study of the relationship between LET and tumor RBE, performed for mouse Harderian gland tumors, shows an increase in RBE with LET, with a peak of RBE ≈ 30 in the range $100–200\,\mathrm{keV\,\mu m}^{-1}$.

During the 1990s, the LNT assumption has been challenged by observations of low-dose phenomena such as bystander effect and genomic instability (see below), which suggest a supralinear low-dose response for different endpoints including mutations and cell conversion to malignancy. The question whether this may imply higher cancer risk levels than those estimated with the LNT model is still open. While a supralinear response for conversion to malignancy and nonlethal mutations might imply a supralinear response for cancer risk, death of bystander cells – especially via apoptosis – may represent a systemic response aimed to protect the tissue/organ by eliminating the most damaged cells, which would be substituted by new, undamaged cells.

"Nontargeted" Effects: Bystander Effect and Genomic Instability

That energy from radiation must be deposited in the cell nucleus to induce an effect in that cell is a long-standing dogma. However, in the 1990s, this dogma was challenged by a number of studies including bystander effects and genomic instability. "Bystander effect" refers to the induction of damage in unirradiated cells ("bystander cells"), mainly due to cellular communication. In *in vitro* studies, bystander cells are often in the proximity of hit cells, that is the case of broadbeam irradiation with low doses of light ions or irradiation with microbeams, which can deliver an exact number of particles specifically targeting selected cell nuclei. However, the term "bystander

effect" also refers to the observation of damage in unexposed cells cultured into a medium taken from irradiated cultures.

Whatever the experimental procedure, bystander effects have been observed for different endpoints, both lethal (e.g., clonogenic inactivation and apoptosis) and nonlethal (e.g., gene mutations and oncogenic transformation). Such effects start being observed at doses of the order of ≈ 0.01 Gy and tend to saturate around ≈ 0.5 Gy, where direct effects dominate. The effect is strongly dependent on factors such as cell type and degree of cell-to-cell contact, which is a key parameter. In fact, although the underlying mechanisms have not been elucidated yet, it is widely accepted that a major role is played by two main cellular communication pathways, involving either the release of signaling molecules into the medium or the transfer of small molecules and ions (e.g., Ca^{++}) through gap junctions, which are channels directly connecting the cytoplasms of adjacent cells.

"Genomic instability" refers to an increased rate of acquisition of alterations in the genome. Radiation-induced instability generally manifests in the progeny of irradiated cells as gross chromosomal rearrangements, micronuclei, and gene mutations. Similar to bystander effect, genomic instability is efficiently induced by low doses and depends both on the cell type and on the radiation quality. The prevailing hypothesis on the mechanisms is that radiation destabilizes the genome and initiates a cascade of events that increase the rate of mutations, chromosomal changes, and possibly other effects in the progeny of irradiated cells. The molecular and cellular events that initiate and perpetuate instability are still unknown. Whereas deficiencies in DNA damage response, alterations in gene expression, and perturbations in cellular homeostasis (i.e., control of activities such as proliferation and differentiation) are likely to be involved, directly induced DNA damage (e.g., dsb) probably does not play a significant role. Interestingly, instability can be induced not only following direct irradiation, but also by a bystander-type mechanism. It is, therefore, likely that both bystander effects and genomic instability are manifestations of the same processes.

Most evidence on these phenomena comes from *in vitro* experiments, and the extrapolation to *in vivo* scenarios is still an open question. However, while waiting for further *in vivo* studies, it would be prudent to take into account the possible implications of nontargeted effects when considering models of radiation-induced carcinogenesis, particularly at low doses where the LNT model might not necessarily hold.

See also: Biomedical Materials; DNA and RNA, Biophysical Aspects; Stochastic Processes in Physics and Chemistry.

PACS: 87.50. – Gi

Further Reading

Alpen EL (1998) *Radiation Biophysics*. San Diego, CA: Academic Press.

Hall EJ (2000) *Radiobiology for the Radiologist*. Philadelphia, PA: Lippincott Williams & Wilkins.

Paretzke HG (1987) Radiation track structure theory. In: Freeman GR (ed.) *Kinetics of Nonhomogeneous Processes*, pp. 89–170. New York: Wiley.

Steel GG (ed.) (1997) *Basic Clinical Radiobiology*. London: Arnold.

Turner JE (1995) *Atoms, Radiation and Radiation Protection*. New York: Wiley.

Irreversible Thermodynamics and Basic Transport Theory in Solids

J M Honig, Purdue University, West Lafayette, IN, USA

Introductory Comments

An overview of basic ideas in irreversible thermodynamics as they apply to electron transport phenomena in solids is provided; the theory is thus not presented in its full generality. The discussion is limited to systems that are at rest, remain at constant volume, undergo no chemical changes, are isotropic, and for which steady-state conditions hold. In these circumstances, the ordinary time differential operator d/dt and its partial counterpart, $\partial/\partial t$, are equivalent, and the tensorial notation can be dispensed with. Readers who desire a full introduction to irreversible phenomena should consult the listings under "Further reading" section. The discussion is also limited to the linearized regime, in which departures from equilibrium are small, such that intensive variables such as temperature, pressure, concentration, chemical potential, and the like retain their significance locally.

However, these variables are presumed to vary with position in such a manner that their gradients are small compared to their values averaged over the system.

Elementary Description of Heat and Mass Transport

Transfer of Heat between Two Different Systems

To begin with, a heuristic introduction of basic concepts characterizing relaxation processes that lead toward establishment of equilibrium conditions is given. Consider two isolated regions in space, designated by (and at) temperatures T' and T''; when brought into contact (see **Figure 1**), and on allowing the intervening partition to become slightly diathermic, heat is slowly transferred between the compartments. If the enlarged unit remains isolated, and the interfacial phenomena are ignored, the energies, E' and E'', of the two subsystems are additive, as are the entropies $S'(E',V')$ and $S''(E'',V'')$. Assume further that the volumes V' and V'' remain fixed. Then the first and second laws of thermodynamics lead to the relations

$$\mathrm{d}E' + \mathrm{d}E'' = 0 \qquad [1a]$$

$$\mathrm{d}S = (\partial S'/\partial E')_{V'}\,\mathrm{d}E' + (\partial S''/\partial E'')_{V''}\,\mathrm{d}E'' \geqslant 0 \qquad [1b]$$

On account of [1a] and with $(\partial E/\partial S)_V = T$, where T is the temperature of the bulk of the system, eqn [1b] becomes

$$\mathrm{d}S = (1/T'' - 1/T')\,\mathrm{d}E'' \geqslant 0 \qquad [2]$$

which immediately establishes the requirement $T' = T''$ as a necessary condition for thermal equilibrium. Moreover, according to the second law, as equilibrium is established, the total entropy of the isolated compound system increases toward a maximum, consistent with the remaining constraints. Hence, the time derivative of the total entropy, $\dot{S} \equiv (\mathrm{d}S/\mathrm{d}t)$, may be written as

$$\dot{S} = (1/T'' - 1/T')\,(\mathrm{d}E''/\mathrm{d}t) \geqslant 0 \qquad [3]$$

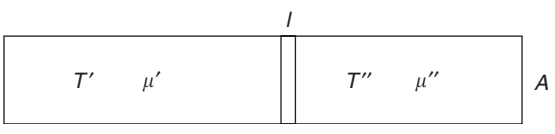


Figure 1 Two bodies of cross section A at temperatures T' and T'' and at chemical potentials μ' and μ'' are in tenuous contact. Changes in T and μ occur over a small distance l straddling the boundary. When no mass transfer occurs, the chemical potentials are irrelevant.

In the present case, no work has been performed and no material has been transferred; therefore, all energy changes involve solely a transfer of heat between the two subsystems. It is then appropriate to identify $\mathrm{d}E''/\mathrm{d}t$ with the rate of heat transfer $\dot{Q}$ across the boundary. Hence, eqn [3] is rewritten as $\dot{S} = \Delta(1/T)\dot{Q}$. Next, a heat flux is defined by $J_Q \equiv \dot{Q}/A$, where A is the cross section of the partition. In the experiment, the temperatures T' and T'' remain nearly constant within both compartments, except near the common boundary, where the change from T' to T'' occurs over a small distance l perpendicular to the partition (see **Figure 1**). The product Al roughly defines a volume V in the boundary region over which the temperature changes. One can then write $\dot{S} = \Delta(1/T)VJ_Q/l$. In the limit of small l, the ratio $\Delta(1/T)/l$ becomes the gradient of inverse temperature, $\nabla(1/T)$. The ratio $\dot{S}/V$ is the rate of entropy production per unit volume, $\dot{\theta}$, in the heat transfer region. This gives, finally,

$$\dot{\theta} = \nabla(1/T)J_Q \geqslant 0 \qquad [4]$$

Heat Transfer Equations

The above discussion is very suggestive. The rate of entropy generation is seen to be a product of two conjugate variables, namely, a heat flux, J_Q, and a generalized force, $F_T \equiv \nabla(1/T)$; hence, schematically $\dot{\theta} = F_T J_Q \geqslant 0$. Note that in eqns [3] and [4], and with the directions indicated in **Figure 1**, $\dot{\theta} > 0$ means either that $F_T > 0$, $J_Q > 0$, with $T'' < T'$, or that $F_T < 0$, $J_Q < 0$, with $T'' > T'$. In either case, heat flows from the hotter to the cooler side. When $\dot{\theta} = 0$, equilibrium prevails; J_Q and F_T both vanish. It is then plausible to claim that the heat flux is "driven" by the existence of a temperature gradient (actually, a gradient in inverse temperature).

Next, a relation between J_Q and F_T is established, which requires extrathermodynamic information, such as a microscopic transport theory or experimental considerations. Close to the equilibrium, a linear relation $J_Q = L_T F_T$ is assumed to prevail between force and flux, where L_T is a parametric function independent of J_Q or F_T. The proportionality factor is known as a phenomenological or macroscopic coefficient. The essential correctness of this very primitive approach may be checked by noting that

$$J_Q = L_T \nabla(1/T) = -(L_T/T^2)\nabla T \equiv -\kappa \nabla T \qquad [5]$$

where $\kappa \equiv L_T/T^2$ is the thermal conductivity, and one recovers the Fourier law of heat conduction.

Transfer of Energy and Mass between Two Different Regions

The above approach will now be generalized. Consider two compartments at different temperatures that contain the same material at two different mole numbers n' and n''. The two units are joined and the intervening partition is rendered slightly porous, allowing the very gradual transfer of energy and of mole numbers (**Figure 1**). As long as the compartmental volumes are unchanged, one can invoke the fundamental relation

$$\begin{aligned} \mathrm{d}S = &(\partial S'/\partial E')_{V',\,n'}\,\mathrm{d}E' + (\partial S''/\partial E'')_{V'',\,n''}\,\mathrm{d}E'' \\ &+ (\partial S/\partial n')_{V',\,E'}\,\mathrm{d}n' + (\partial S/\partial n'')_{V'',\,E''}\,\mathrm{d}n'' \end{aligned} \qquad [6]$$

Since $\mathrm{d}S = \mathrm{d}E/T - (\mu/T)\mathrm{d}n$, where μ is the chemical potential, and with $\mathrm{d}E'' = -\mathrm{d}E'$, $\mathrm{d}n'' = -\mathrm{d}n'$, the above may be recast as

$$\begin{aligned} \mathrm{d}S = &(1/T'' - 1/T')\mathrm{d}E'' \\ &- (\mu''/T'' - \mu'/T')\mathrm{d}n'' \geqslant 0 \end{aligned} \qquad [7a]$$

from which it follows that at equilibrium, $T' = T''$ and $\mu' = \mu''$, consistent with well-established thermodynamic principles. The earlier procedure can be extended by introducing two fluxes, namely $J_E \equiv \mathrm{d}E''/A\mathrm{d}t$ and $J_n \equiv \mathrm{d}n''/A\mathrm{d}t$, as well as two generalized forces, $F_T \equiv \Delta(1/T)/l \equiv \nabla(1/T)$ and $F_n \equiv \Delta(\mu/T)/l \equiv \nabla(\mu/T)$, appropriate for small distances, l, across the junction. Equation [7a] thus becomes

$$\dot{\theta} = J_E\nabla(1/T) - J_n\nabla(\mu/T) \qquad [7b]$$

The minus sign in this equation can also be rationalized via a sufficiency argument. To ensure that $\dot{\theta} > 0$, both terms in the expression need to be positive. Referring to **Figure 1**, assume that $J_E > 0$ (i.e., energy flows left to right); to keep the first term positive it is necessary that $T'' < T'$, as before. In order to render the second term positive, one requires that $(\mu''/T'' - \mu'/T')J_n < 0$. Assume that $J_n > 0$ (particles flow from left to right), which then requires the multiplier to be negative, so that $\mu'/T' > \mu''/T''$, or $\mu'/\mu'' > T'/T'' > 1$, implying that the particle flux occurs from the region of higher chemical potential to one of lower chemical potential.

Phenomenological Equations

The relation $T\mathrm{d}S = \mathrm{d}E - \mu\mathrm{d}n$ is now introduced, reformulated after time differentiation as $TJ_S = J_E - \mu J_n$, valid in the absence of any mechanical pressure–volume work (V' and V'' remain fixed). Using eqn [7b], the rate of entropy production per unit volume may be written as

$$\dot{\theta} = -\frac{1}{T}\{J_S\nabla T + J_n\nabla\mu\} \geqslant 0 \qquad [8]$$

Such a relation is of fundamental importance, in that it specifies $\dot{\theta}$ as a sum of products that involve conjugate flux–force pairs, namely, $(J_S, -T^{-1}\nabla T)$ and $(J_n, -T^{-1}\nabla\mu)$. These quantities are of great importance in the formulation of the so-called linear phenomenological equations (PEs), valid for small departures from equilibrium. These are constructed so as to relate the two fluxes J_S and J_n to the two forces $(1/T)\nabla T$ and $(1/T)\nabla\mu$ in the following linear form:

$$J_S = (L_{11}/T)\nabla T + (L_{12}/T)\nabla\mu \qquad [9a]$$

$$J_S = (L_{21}/T)\nabla T + (L_{22}/T)\nabla\mu \qquad [9b]$$

The minus signs have been absorbed in the L coefficients. One should note the implied superposition principle: both forces contribute to both fluxes. The L_{ij} are known as phenomenological or macroscopic coefficients; L_{11} and L_{22} connect the fluxes to their principal driving forces, while L_{12} and L_{21} give rise to interference effects that are studied in detail below.

Basic Principles of Mass, Energy, and Entropy Balances

Mass Balance Equation

The very heuristic approach presented so far, while leading to correct results, ought really to be supplanted by a more rigorous (though simplified) derivation of eqn [9]. One can begin with the conservation law

$$\mathrm{d}\rho_k/\mathrm{d}t = -\boldsymbol{\nabla}\cdot\boldsymbol{J}_k \qquad [10]$$

wherein the total rate change in concentration ρ_k of species k within the system (**Figure 2**) is given by the imbalance between the influx and outflow of species k across the boundaries; the minus sign arises because the normal unit vector in **Figure 2** faces outward.

Energy Balance Equations

In order to introduce relations for energy balance, one can follow the convention by changing from energy or mass per unit volume to energy per unit mass and mass fraction. This is indicated by overbars such as $\bar{u}, \bar{e}, \bar{\psi}, \bar{n}_k$, as introduced below. Further, let species k interact with a stationary external potential $\bar{\psi}_k$ (e.g., electrons in an electrostatic potential); then operating on both sides of eqn [10] with $\sum_{(k)}\bar{\psi}_k$ and

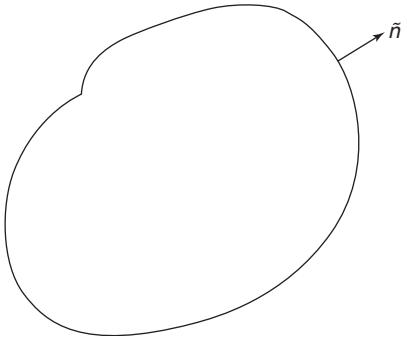


Figure 2 Boundary of a system, with the unit vector pointing opposite to the inward flow of energy, entropy, or mass.

setting $\sum_{(k)} \bar{\psi}_k \rho_k \equiv \bar{\psi}\rho$, and using the vector calculus involving a scalar function a and a vector function $\boldsymbol{v}$, $\boldsymbol{\nabla} \cdot a\boldsymbol{v} = a\boldsymbol{\nabla} \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{\nabla} a$, one obtains

$$\mathrm{d}(\rho\bar{\psi})/\mathrm{d}t = -\boldsymbol{\nabla} \cdot \sum_k \bar{\psi}_k \boldsymbol{J}_k + \sum_k \boldsymbol{J} \cdot \boldsymbol{\nabla}\bar{\psi}_k \quad [11]$$

As a third step, one introduces the conservation law for total energy density $\rho\bar{u}$ (energy cannot be generated or destroyed locally):

$$\mathrm{d}(\rho\bar{u})/\mathrm{d}t = -\boldsymbol{\nabla} \cdot \boldsymbol{J}_U \quad [12]$$

The total energy density can be separated into $\bar{u} = \bar{e} + \bar{\psi}$, where $\bar{e}$ is the internal energy per unit mass. Correspondingly,

$$\boldsymbol{J}_U = \boldsymbol{J}_Q + \sum_k \bar{\psi}_k \boldsymbol{J}_k \quad [13]$$

This nomenclature is appropriate; the net total energy flux density $\boldsymbol{\nabla} \cdot \boldsymbol{J}_U$ consistently reflects the performance of work, potential energy changes, and heat flows. The first two effects are subsumed in the term $\boldsymbol{\nabla} \cdot \sum_k \bar{\psi}_k \boldsymbol{J}_k$; the rest is contained in the net heat flux, $-\boldsymbol{\nabla} \cdot \boldsymbol{J}_Q$. Next, subtracting [11] from [12] and introducing eqn [13], this yields an expression for the overall rate of change of internal energy

$$\mathrm{d}(\rho\bar{e})/\mathrm{d}t = -\boldsymbol{\nabla} \cdot \boldsymbol{J}_Q - \sum_k \boldsymbol{J}_k \cdot \boldsymbol{\nabla}\bar{\psi}_k \quad [14]$$

This involves not only the customary divergence term but also the last term, which represents a source term that arises from the interaction of species k with the external potential $\bar{\psi}_k$.

Entropy Balance Relations

The entropy balance equations can be introduced by using the expression (for $\mathrm{d}V = 0$) $T\mathrm{d}S = \mathrm{d}E - \sum_k \mu_k \mathrm{d}n_k$ to write

$$T(\mathrm{d}(\rho\bar{s})/\mathrm{d}t) = (\mathrm{d}(\rho\bar{e})/\mathrm{d}t) - \sum_k \bar{\mu}_k(\mathrm{d}\rho_k/\mathrm{d}t) \quad [15a]$$

On insertion of eqns [14] and [10], one obtains

$$\frac{\mathrm{d}(\rho\bar{s})}{\mathrm{d}t} = \frac{1}{T}\left\{-\boldsymbol{\nabla} \cdot \boldsymbol{J}_Q - \sum_k \boldsymbol{J}_k \cdot \boldsymbol{\nabla}\bar{\psi}_k + \sum_k \bar{\mu}_k \boldsymbol{\nabla} \cdot \boldsymbol{J}_k\right\} \quad [15b]$$

The above may be rearranged by repeated application of the vector identity cited earlier; standard manipulations lead to the final result

$$\begin{aligned}\frac{\mathrm{d}(\rho\bar{s})}{\mathrm{d}t} = &-\boldsymbol{\nabla} \cdot \left\{\left(\boldsymbol{J}_Q - \sum_k \bar{\mu}_k \boldsymbol{J}_k\right)/T\right\} \\ &- (1/T^2)\boldsymbol{J}_Q \cdot \boldsymbol{\nabla}T + (1/T^2)\sum_k \bar{\mu}_k \boldsymbol{J}_k \cdot \boldsymbol{\nabla}T \\ &- (1/T)\sum_k \boldsymbol{J}_k \cdot \boldsymbol{\nabla}(\bar{\mu}_k + \bar{\psi}_k) \end{aligned} \quad [16]$$

Although this looks complicated, it is amenable to a simple interpretation which is based on the entropy balance equation in the form $\mathrm{d}(\rho\bar{s})/\mathrm{d}t = -\boldsymbol{\nabla} \cdot \boldsymbol{J}_S + \dot{\theta}$. Comparison with eqn [16] shows that it is appropriate to consider the quantity in curly brackets as a flux vector, so that

$$\boldsymbol{J}_S = \frac{1}{T}\left(\boldsymbol{J}_Q - \sum_k \bar{\mu}_k \boldsymbol{J}_k\right) \quad [17]$$

is an entropy flux vector. The remaining terms are then to be identified as source terms $\dot{\theta}$ in the entropy balance equation. This allows one to express the contribution of local events to the entropy generation rate as

$$\dot{\theta} = -(1/T)\boldsymbol{J}_S \cdot \boldsymbol{\nabla}T - (1/T)\sum_k \boldsymbol{J}_k \cdot \boldsymbol{\nabla}\bar{\zeta}_k \geqslant 0 \quad [18]$$

where the generalized chemical potential has been introduced as $\bar{\zeta} \equiv \bar{\mu} + \bar{\psi}$. For isotropic materials in the absence of external potentials, eqn [18] reduces to eqn [8]. This establishes the proper background for constructing the PEs, eqn [9], precisely in the same manner as was done earlier.

Phenomenological Equations for Thermoelectric Effects; Onsager's Reciprocity Conditions

A very important characteristic of eqns [9] is the fact that the Onsager reciprocity relations, $L_{ij} = L_{ji}$ apply

to such equations, so long as conjugate flux–force pairs are employed in formulating eqns [9]. For proofs, the reader is referred to the literature; a simpler version, applicable to processes under steady-state conditions, was provided by Tykody.

Thermoelectric Phenomena

Equations [9] can be specialized to characterize interactions between heat flow and electric currents in a conductor, known as thermoelectric effects; this topic has been treated elsewhere in great detail (see the "Further reading" section). Consider a rectangular bar clamped between two thermal reservoirs at different temperatures (**Figure 3**). The electron flow in the conductor is established by charging external condenser plates; this cumbersome method is used here (though not in actual practice) to avoid distracting complications from leads that normally connect the sample to the current source. Using the dissipation function, eqn [18] with $k = n$, substituting $\bar{\zeta} = \bar{\mu} - e\bar{\varphi}$ in place of $\bar{\mu}$, and then replacing the electron flux vector $\boldsymbol{J}_n$ by the current density $\boldsymbol{J} = (-e)\boldsymbol{J}_n$, where $-e$ is the charge on the electron and $\bar{\varphi}$ is the electrostatic potential per electron mass, the dissipation relation can be written as $\dot{\theta} = \boldsymbol{J}_S \cdot (-T^{-1}\boldsymbol{\nabla} T) + \boldsymbol{J} \cdot (-T^{-1}\boldsymbol{\nabla}(\bar{\zeta}/e))$. One can identify, from this relation, the proper fluxes and forces, so as to construct the phenomenological relations for an isotropic material (the vector notation can then be dropped):

$$J_S = -(L_{SS}/T)\nabla T + (L_{Sn}/T)\nabla(\zeta/e) \qquad [19a]$$

$$J = -(L_{nS}/T)\nabla T + (L_{nn}/T)\nabla(\zeta/e) \qquad [19b]$$

with $L_{nS} = L_{Sn}$. In doing so, one has also switched from $\bar{\zeta}$ to ζ, absorbing the mass factor and the minus sign in L_{Sn} and L_{nn}.

Thermoelectric Transport Coefficients

By imposing a variety of constraints, the above expressions can be attributed to physical significance:

(1) Set $\nabla T = 0$ and eliminate $\nabla(\zeta/e)$ between [19a] and [19b] to obtain $J_S/J = L_{Sn}/L_{nn} = -J_S/eJ_n$; since $J_S/J_n \equiv S^*$ is the total entropy carried per electron, one obtains

$$L_{Sn}/L_{nn} = -S^*/e \qquad [20]$$

Next, consider eqn [19b] with $\nabla T = 0$. For a homogeneous sample $\nabla\mu = 0$; eqn [19b] then reduces to $J = (L_{nn}/T)\nabla(-\phi) = (L_{nn}/T)E$, where E is the electrostatic field. Thus, one recovers Ohm's law $J = \sigma E$, where σ is the electrical conductivity. This leads to the identification

$$L_{nn}/T = \sigma, \quad J = \sigma\nabla(\zeta/e) \qquad (\nabla T = 0) \quad [21]$$

(2) If current flow is suppressed, set $J = 0$, solve eqn [19b] for $\nabla(\zeta/e)$, and substitute the result in eqn [19a]; this yields

$$J_S = -(1/T)[L_{SS} - L_{nS}^2/L_{nn}]\nabla T \quad (J = 0) \quad [22]$$

Now, as shown in eqn [17], in the absence of current flow the entropy flux may be replaced by J_Q/T; thus, one recovers Fourier's law of heat conduction, $J_Q = -\kappa\nabla T$, with the thermal conductivity given by

$$\kappa = L_{SS} - L_{nS}^2/L_{nn} \qquad [23]$$

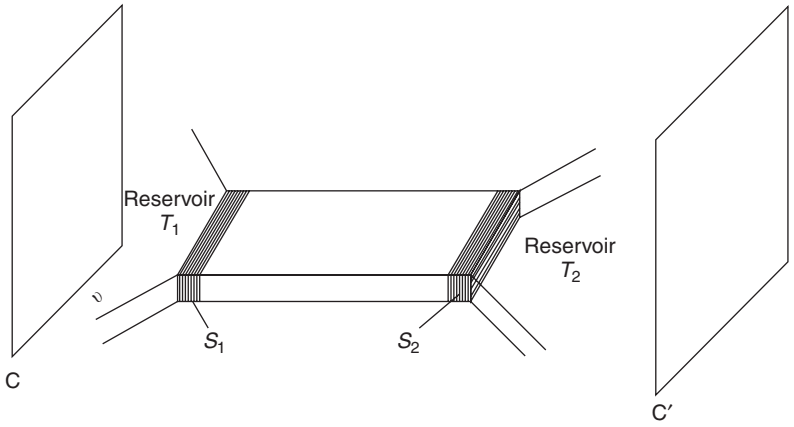


Figure 3 Idealized setup for detection of thermoelectric effects. The solid is clamped between two thermal reservoirs at temperatures T_1 and T_2; S_1 and S_2 are removable strips for thermal insulation. Current flow is activated through charging of the condenser plates C and C′.

An additional contribution to κ arises from the lattice thermal conductivity that is not included here.

(3) In the more general case when $J>0$, one may still eliminate $\nabla(\zeta/e)$ between eqns [19a] and [19b] to obtain

$$J_S = -(S^*/e)J - (\kappa/T)\nabla T = S^* J_n - (\kappa/T)\nabla T \qquad [24]$$

This shows that the total electronic entropy flux involves the entropy transport $S^* J_n$ as well as heat transport through the motion of the electrons.

(4) A second result arises by imposing the condition $J=0$ and solving eqn [19b] for

$$\nabla(\zeta/e) = (L_{nS}/L_{nn})\nabla T \equiv \alpha \nabla T \qquad [25]$$

Here, one encounters a new prediction: the imposition of a temperature gradient necessarily establishes a gradient in the electrochemical potential. The proportionality factor, here designated as α, is called the Seebeck coefficient, or less desirably, the thermoelectric power. It is measured by connecting the ends of the sample in a temperature gradient to a voltmeter under open circuit conditions, and by determining the resulting temperature difference with, for example, a thermocouple. Comparison with eqn [20] shows that $\alpha = -S^*/e = \nabla(\zeta/e)/\nabla T \approx \Delta(\zeta/e)/\Delta T$. Equation [19a] may thus be rewritten as

$$J_S = \alpha J - (\kappa/T)\nabla T \qquad [26]$$

The three phenomenological coefficients may now be solved in terms of the three experimental parameters α, κ, and σ, so as to express the PEs in their final form

$$J_S = -(\sigma\alpha^2 + \kappa/T)\nabla T + \alpha\sigma\nabla(\zeta/e) \qquad [27a]$$

$$J = -\sigma\alpha\nabla T + \sigma\nabla(\zeta/e) \qquad [27b]$$

Equation [27b] represents a further generalization of Ohm's law, showing how the current density is affected by the presence of a temperature gradient. The apparent differences in sign involving $\alpha\sigma$ arise because there is a corresponding sign difference in eqns [19a] and [19b]. Thus, Onsager's relations are indeed satisfied.

A physical explanation of the thermoelectric effect is based on the fact that electrons at the hot end have a larger thermal velocity along the downstream direction than that associated with electrons at the cold end along the upstream direction. Consequently, a net accumulation of electrons takes place at the cold end, leaving an effective positive charge at the hot end.

Hole conductors may be treated by the same approach: a rather lengthy analysis shows that one may simply replace $-e$ by $+e$; for n-type carriers $\alpha<0$, whereas for p-type materials $\alpha>0$. The results otherwise carry over without change. Thermoelectric effects thus provide a means for distinguishing between n- and p-type charge transport.

Peltier and Thomson Effects

Two other thermoelectric effects encountered in the literature are only briefly mentioned. The so-called Peltier effects arise when current is passed through an isothermal junction of two dissimilar materials (1 and 2), resulting in the production of heat. The explanation is based on eqn [13], with $k=1$, and ψ and $\mu \to \zeta$. An electrochemical potential is continuous across the junction; also, the current density is conserved. Hence, in passing from region 1 to region 2,

$$J_{U2} - J_{U1} = J_{Q2} - J_{Q1} = T(J_{S2} - J_{S1}) = T(\alpha_2 - \alpha_1)J$$

where eqns [17] and [26] and $\nabla T = 0$ have been used. The Peltier coefficient is defined as

$$\Pi \equiv (J_{Q2} - J_{Q1})/J = T(\alpha_2 - \alpha_1) \qquad [27c]$$

showing the relation between Π and α. On the other hand, if current flows through a conductor with an established temperature gradient, the heat generated in line element dx (in excess of the Joule heating effect) is proportional to the current and to the temperature gradient, as expected: $\dot{Q}/dx \propto I\nabla_x T$. On introducing a proportionality factor, τ, the so-called Thomson coefficient, one obtains

$$J_{Qx} = \tau J_x \nabla_x T \qquad [27d]$$

For an extended derivation and discussion of these effects see the "Further reading" section.

Transport Effects in Magnetic Fields

Now consider the effects that operate in the presence of a magnetic field, H, aligned with the z-axis. Provision is made in the sample for current and heat flow in the x and y directions (**Figure 4**).

Phenomenological Equations for Thermomagnetic Phenomena

A study of the thermoelectric and thermomagnetic interactions resulting in the presence of magnetic fields is useful. The effects introduced by the electric leads will be ignored. For this purpose the choice of $(J_S, \nabla T)$ and $(J, \nabla(\zeta/e))$ as conjugate variables is extended; the $1/T$ factors and minus signs are absorbed

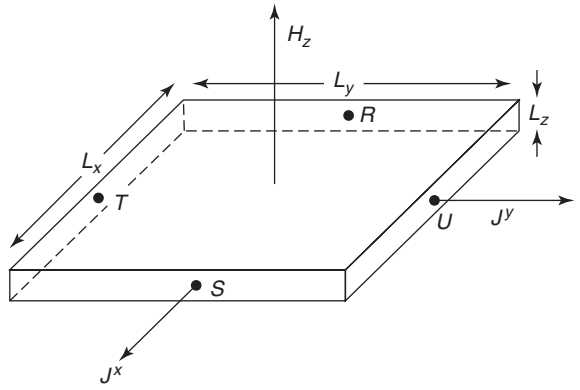


Figure 4 Sample cut in a rectangular parallelepiped geometry and placed in a magnetic field H aligned with the z-axis. Current and heat pass into the solid through leads connected to R and S along x, and to T and U along y.

into the phenomenological coefficients. Three new points must be taken into account: (1) since the observed effects occur along two dimensions, $(J_S^x, \nabla_x T)$, $(J_S^y, \nabla_y T)$, $(\nabla_x(\zeta/e), J^x)$, $(\nabla_y(\zeta/e), J^y)$ are introduced as conjugate flux–force pairs. (2) For later convenience, $J_S^x, J_S^y, \nabla_x(\zeta/e), \nabla_y(\zeta/e)$ are considered as dependent variables, so that the PEs cited below appear in partially inverted form. Such a partial interchange of dependent and independent variables is a perfectly permissible mathematical step and greatly simplifies the subsequent analysis. (3) The Onsager reciprocity conditions must now be generalized to read (Casimir–Onsager relations) $L_{ij}(H) = \pm L_{ji}(-H)$, the adopted sign depending on whether or not the forces on the charged particles change sign when the magnetic field is reversed. Without loss of generality, the sequence of signs in eqn [28a] is selected arbitrarily; the signs in the other equations are then determined by the augmented reciprocity relations. The above considerations lead to macroscopic PEs of the form

$$J_S^x = -L_{11}\nabla_x T - L_{12}\nabla_y T + L_{13}J^x + L_{14}J^y \quad [28a]$$

$$J_S^y = L_{12}\nabla_x T - L_{11}\nabla_y T - L_{14}J^x + L_{13}J^y \quad [28b]$$

$$\nabla_x(\zeta/e) = L_{13}\nabla_x T + L_{14}\nabla_y T + L_{33}J^x + L_{34}J^y \quad [28c]$$

$$\nabla_y(\zeta/e) = -L_{14}\nabla_x T + L_{13}\nabla_y T - L_{34}J^x + L_{33}J^y \quad [28d]$$

The above formulation permits a systematic investigation of the 560 possible thermoelectric and thermomagnetic effects. A selection of these is presented under a variety of headings.

Transport Coefficients for Thermomagnetic Phenomena

(1) $J^y = \nabla_x T = \nabla_y T = 0$: isothermal conditions are maintained along x and y and no current is allowed to flow along y. The PEs then reduce to

$$\begin{aligned} J_S^x &= L_{13}J^x \quad \text{(a)} \\ J_S^y &= -L_{14}J^x \quad \text{(b)} \\ \nabla_x(\zeta/e) &= L_{33}J^x \quad \text{(c)} \\ \nabla_y(\zeta/e) &= -L_{34}J^x \quad \text{(d)} \end{aligned} \quad [29]$$

Expression [29c] is simply a reformulation of Ohm's law, $J^x = \sigma_I \nabla_x(\zeta/e)$, wherein $L_{33} \equiv \rho_I$ is the isothermal resistivity of the sample. Equation [29d] describes the isothermal Hall effect: a current in the longitudinal direction induces a transverse gradient in the electrochemical potential; for convenience, the magnitude of the magnetic field is explicitly introduced: it can be seen that

$$\nabla_y(\zeta/e) = -(L_{34}/H)J^x H \equiv \Re_I J^x H \quad [30]$$

where the magnitude of the effect is characterized by the Hall coefficient $\Re \equiv -L_{34}/H$. Equations [29a] and [29b] simply relate the entropy transport along x and y to current flux along x; L_{13} and L_{14} are determined below.

(2) $J^y = \nabla_x T = J_S^y = 0$: open-circuit conditions are imposed and no heat flow is allowed along y; isothermal conditions prevail along x. The PEs now reduce to

$$\begin{aligned} J_S^x &= -L_{12}\nabla_y T + L_{13}J^x \quad \text{(e)} \\ 0 &= -L_{11}\nabla_y T - L_{14}J^x \quad \text{(f)} \\ \nabla_x(\zeta/e) &= L_{14}\nabla_y T + L_{33}J^x \quad \text{(g)} \\ \nabla_y(\zeta/e) &= L_{13}\nabla_y T - L_{34}J^x \quad \text{(h)} \end{aligned} \quad [31]$$

Equation [31f] shows that under these conditions, a longitudinal current flow produces a transverse temperature gradient; this is the so-called Ettingshausen effect $\nabla_y T = -(L_{14}/L_{11}H)J^x H$, with a coefficient specified by $\Im \equiv \nabla_y T/J^x H = -L_{14}/L_{11}H$. Insertion of [31f] into [31g] yields the expression $\nabla_x(\zeta/e) = (L_{33} - L_{14}^2/L_{11})J^x$, which is Ohm's law under "adiabatic" conditions, with a resistivity $\rho_A = (L_{33} - L_{14}^2/L_{11})$. When [31f] is used in [31h], one obtains $\nabla_y(\zeta/e) = -[(L_{13}L_{14}/L_{11} + L_{34})/H]J^x H$, which represents the "adiabatic" Hall effect, with a corresponding coefficient $\Re_A \equiv -(1/H)[L_{13}L_{14}/L_{11} + L_{34}]$.

(3) Next, set $J^x = J^y = \nabla_y T = 0$: no current flows and the temperature is uniform along y. The

PEs reduce to

$$\begin{aligned} J_S^x &= -L_{11}\nabla_x T \quad \text{(i)} \\ J_S^y &= L_{12}\nabla_x T \quad \text{(j)} \\ \nabla_x(\zeta/e) &= L_{13}\nabla_x T \quad \text{(k)} \\ \nabla_y(\zeta/e) &= -(L_{14}/H)H\nabla_x T \quad \text{(l)} \end{aligned} \qquad [32]$$

Under these conditions, TJ_S^x and TJ_S^y represent heat fluxes. Equation [32i] then specifies an "isothermal" heat flux (a contradiction in terms!) $TJ_S^x = -L_{11}T\nabla_x T$ that identifies the thermal conductivity $\kappa_I = TL_{11}$ in the absence of a transverse thermal gradient. Moreover, [32k] represents nothing other than the "isothermal" Seebeck effect, along x, for which the coefficient is given by $\alpha = L_{13}$. Furthermore, according to [32l], a T gradient along x produces a gradient in ζ along y. This is the so-called transverse Nernst effect, with a corresponding coefficient $N_I \equiv -L_{14}/H$.

(4) Another set of constraints of interest is given by $J^x = J^y = J_S^y = 0$. These reduce the PEs to

$$\begin{aligned} J_S^x &= -L_{11}\nabla_x T - L_{12}\nabla_y T \quad \text{(m)} \\ 0 &= L_{12}\nabla_x T - L_{11}\nabla_y T \quad \text{(n)} \\ \nabla_x(\zeta/e) &= L_{13}\nabla_x T + L_{14}\nabla_y T \quad \text{(o)} \\ \nabla_y(\zeta/e) &= -L_{14}\nabla_x T + L_{13}\nabla_y T \quad \text{(p)} \end{aligned} \qquad [33]$$

Equation [33n] shows that, in this case, the temperature gradient in the longitudinal direction induces one in the transverse direction; this is known as the Righi–Leduc effect, rewritten as

$$\nabla_y T = (L_{12}/L_{11}H)H\nabla_x T \equiv MH\nabla_x T \qquad [34]$$

with a corresponding coefficient $M \equiv L_{12}/L_{11}H$. A second relation of interest is found by inserting eqn [34] into [33m] and multiplying through by T; this yields $TJ_S^x = -T[L_{11} + L_{12}^2/L_{11}]\nabla_x T$, leading to the "adiabatic" thermal conductivity $\kappa = T(L_{11} + L_{12}^2/L_{11})$. Also, use of [34] in [33o] leads to $\nabla_x(\zeta/e) = (L_{13} + L_{14}L_{12}/L_{11})\nabla_x T$, which represents the adiabatic Seebeck effect, with a corresponding coefficient $\alpha_A = L_{13} + L_{14}L_{12}/L_{11}$. Lastly, combining [34] with [33p] yields $\nabla_y(\zeta/e) = (1/H)[-L_{14} + L_{13}L_{12}/L_{11}]H\nabla_x T$, which represents the adiabatic transverse Nernst effect, with a coefficient $N_A = (1/H)[-L_{14} + L_{13}L_{12}/L_{11}]$.

The above procedure may be applied repeatedly to obtain other galvano-thermomagnetic effects; these are left to the reader to explore.

Identification of Phenomenological Coefficients

The principal inference of the discussion so far is that the various parameters L_{ij} of the PEs in eqn [28] can be rewritten in terms of experimentally determined parameters. These are:

$$\begin{aligned} L_{33} &= \rho_I \quad \text{(q)}, & L_{34} &= -\Re_I H \quad \text{(r)} \\ L_{11} &= \kappa_I/T \quad \text{(s)}, & L_{13} &= \alpha_I \quad \text{(t)} \\ L_{14} &= -N_I H \quad \text{(u)}, & L_{12} &= \kappa_I M_I H/T \quad \text{(v)} \end{aligned} \qquad [35]$$

The subscript I has been introduced to indicate isothermal conditions, that is, $\nabla_y T = 0$. By inserting [35] into the PEs, a complete description of the galvano-thermomagnetic effects that are observable in the rectangular parallelepiped geometry of **Figure 4** has been obtained.

Transport in Two-Band (Electron–Hole) Conductors

In conclusion, a limited set of galvano-thermomagnetic phenomena for a material in which both holes and electrons contribute to the conduction process can be considered. For this purpose, it is necessary to use the PEs in uninverted form, so that the partial currents can be summed to form the total current. Specializing to the case $J^y = \nabla_x T = \nabla_y T = 0$, the expressions involving J_S^x and J_S^y are not needed. In an obvious notation, the PEs for electron and hole conduction become, with due regard to the Casimir–Onsager conditions,

$$J_-^x = \Gamma_1\nabla_x(\zeta/e) + \Gamma_2\nabla_y(\zeta/e) \qquad [36a]$$

$$J_-^y = -\Gamma_2\nabla_x(\zeta/e) + \Gamma_1\nabla_y(\zeta/e) \qquad [36b]$$

$$J_+^x = \Gamma_3\nabla_x(\zeta/e) + \Gamma_4\nabla_y(\zeta/e) \qquad [36c]$$

$$J_+^y = -\Gamma_4\nabla_x(\zeta/e) + \Gamma_3\nabla_y(\zeta/e) \qquad [36d]$$

Use of Constraints to Identify Transport Coefficients

Consider first the one-band, n-type semiconductor, by setting $\Gamma_3 = \Gamma_4 = 0$. If $J_-^y = 0$ as well, [36a] and [36b] can be solved for:

$$J_-^x = [(\Gamma_1^2 + \Gamma_2^2)/\Gamma_1]\nabla_x(\zeta/e) \equiv \sigma_n\nabla_x(\zeta/e) \qquad [37a]$$

which immediately identifies the n-type conductivity σ_n. One can also write

$$J_-^x = [(\Gamma_1^2 + \Gamma_2^2)/\Gamma_2]\nabla_y(\zeta/e) \qquad [37b]$$

Table 1 Galvano-thermomagnetic effects for a two-band *n–p* conductor in a magnetic field

Electric resistivity:	$\rho = [(\sigma_n + \sigma_p) + H^2\sigma_n\sigma_p(\sigma_n\Re_n^2 + \sigma_p\Re_p^2)]/D$
Hall coefficient:	$\Re = [\Re_n\sigma_n^2 + \Re_p\sigma_p^2 + (H\sigma_n\sigma_p)^2(\Re_n\Re_p)(\Re_n + \Re_p)]/D$
Seebeck coefficient:	$\alpha = [(\sigma_n + \sigma_p)(\sigma_n\alpha_n + \sigma_p\alpha_p) + (H\sigma_n\sigma_p)^2(\Re_n + \Re_p)(\Re_n\alpha_p + \Re_p\alpha_n) + H^2\sigma_n\sigma_p(N_n - N_p)(\Re_n\sigma_n - \Re_p\sigma_p)]/D$
Thermal conductivity:	$\kappa = \kappa_n + \kappa_p + (T/D)\{[\sigma_n\sigma_p(\sigma_n + \sigma_p)(\alpha_n - \alpha_p)^2 - (N_n - N_p)(\sigma_n\sigma_p H^2)] \times [(\sigma_n + \sigma_p)(N_n - N_p) - 2\sigma_n\sigma_p(\Re_n + \Re_p)(\alpha_n - \alpha_p)]\}$

$D \equiv [(\sigma_n + \sigma_p)^2 + (H\sigma_n\sigma_p)^2(\Re_n + \Re_p)^2]$ σ_n, σ_p are electrical conductivity of electrons and holes, respectively; $\Re_n$, $\Re_p$ are Hall coefficients of electrons and holes; ρ is the electrical resistivity; α_n, α_p are the Seebeck coefficient of electrons and holes, respectively; κ_n, κ_p are the thermal conductivity of electrons and holes, respectively; N_n, N_p are the transverse Nernst coefficients of electrons and holes, respectively; and H is the applied magnetic field. Note that $\alpha_n = -|\alpha_n|$, $\Re_n = -|\Re_n|$ are negative, while σ_n, σ_p, N_n, N_p are positive; hence $\Re$ and α tend to be small except in large applied magnetic fields.

which introduces the Hall coefficient for electrons as

$$\Re_n H = \Gamma_2/(\Gamma_1^2 + \Gamma_2^2) \qquad [38]$$

Equations [37a] and [38] yield the result $\Gamma_2/\Gamma_1 = \sigma_n H\Re_n \equiv X_n$.

On elimination of Γ_2 using [38], $\Gamma_1 = \sigma_n/(1 + X_n^2)$. Similarly, $\Gamma_2 = \sigma_n X_n/(1 + X_n^2)$. Analogous operations for a hole conductor yield the results $\Gamma_4/\Gamma_3 = \sigma_p H\Re_p \equiv X_p$, $\Gamma_3 = \sigma_p/(1 + X_p^2)$, $\Gamma_4 = X_p/(1 + X_p^2)$. The total current density is then given by

$$J^x = J_-^x + J_+^x = \Gamma_a\nabla_x(\zeta/e) + \Gamma_b\nabla_y(\zeta/e) \qquad [39a]$$

$$J^y = J_-^y + J_+^y = -\Gamma_b\nabla_x(\zeta/e) + \Gamma_a\nabla_y(\zeta/e) \qquad [39b]$$

with $\Gamma_a \equiv \Gamma_1 + \Gamma_3$, $\Gamma_b \equiv \Gamma_2 + \Gamma_4$. Setting $J^x = 0$ and eliminating either $\nabla_y(\zeta/e)$ or $\nabla_x(\zeta/e)$ from eqns [39]. One finds

$$J^x = \frac{\Gamma_a^2 + \Gamma_b^2}{\Gamma_a}\nabla_x(\zeta/e) \quad \text{or} \quad \sigma = \frac{\Gamma_a^2 + \Gamma_b^2}{\Gamma_a} \qquad [40a]$$

$$J^x = \frac{\Gamma_a^2 + \Gamma_b^2}{\Gamma_b}\nabla_y(\zeta/e) \quad \text{or} \quad \frac{1}{H\Re} = \frac{\Gamma_a^2 + \Gamma_b^2}{\Gamma_b} \qquad [40b]$$

The last step consists in replacing Γ_a, Γ_b by $\Gamma_1, \ldots, \Gamma_4$, and introducing X_n, X_p, σ_n, $\Re_n H$, σ_p, $\Re_p H$. The steps are elementary but tedious. The results are listed in **Tables 1** and **2**.

A similar procedure may be used to determine the Seebeck coefficient and thermal conductivity of a two-band conductor. For detailed derivations, the reader is referred to the literature. The results are again displayed in **Tables 1** and **2**.

Table 2 Galvano-thermomagnetic effects to the limit of small applied magnetic fields: $H \to 0$

$\sigma = \sigma_n + \sigma_p$,	$\alpha = \dfrac{\alpha_n\sigma_n + \alpha_p\sigma_p}{\sigma_n + \sigma_p}$
$\Re = \dfrac{\Re_n\sigma_n^2 + \Re_p\sigma_p^2}{(\sigma_n + \sigma_p)^2}$,	$\kappa = \kappa_n + \kappa_p + \dfrac{T\sigma_n\sigma_p(\alpha_n - \alpha_p)^2}{\sigma_n + \sigma_p}$

See notes in **Table 1**.

The principal points to be noted are (1) that the results are model-independent, and (2) that the contribution from electrons and holes in general are not simply additive.

Concluding Comments

This article has focused on the thermodynamic representation of irreversible effects as applied to electron transport phenomena in conductors. The advantage of this approach is its generality (aside from specific restrictions mentioned earlier); it does not depend on the adoption of a specific model for electron transport. The disadvantage is that the phenomenological parameters and transport coefficients remain unspecified. One may now invoke the microscopic transport theory to evaluate these quantities from first principles, at least in certain degrees of approximation, but this is a rather lengthy procedure beyond the scope of the present article.

Acknowledgments

The preparation of this article was supported on NSF Grant DMR 96 12130.

See also: Conductivity, Electrical; Mass Transport; Onsager Relations; Phases and Phase Equilibrium; Quantum Hall Effect; Thermodynamic Properties, General; Thermoelectric Effect.

PACS: 05.70.–a; 05.70.Ln; 05.60.Cd; 72.15.Jf; 66.30.–h; 72.10.Bg; 72.15.Gd; 72.20.Pa; 72.20.My

Further Reading

Callen HB (1985) *Thermodynamics and an Introduction to Thermostatistics*. New York: Wiley.
De Groot SR and Mazur P (1962) *Non-Equilibrium Thermodynamics*. Amsterdam: North Holland.
Denbigh KG (1951) *Thermodynamics of the Steady State*. London: Methuen.
Domenicali CA (1954) *Reviews of Modern Physics* 26: 237.
Fitts DD (1962) *Nonequilibrium Thermodynamics*. New York: McGraw-Hill.
Glansdorff P and Prigogine I (1971) *Thermodynamics of Structure Stability and Fluctuations*. New York: Wiley.
Haase R (1968) *Thermodynamics of Irreversible Processes*. Reading, MA: Addison-Wesley.
Harman TC and Honig JM (1967) *Thermoelectric and Thermomagnetic Effects and Applications*. New York: McGraw-Hill.
Honig JM (1999) *Thermodynamics*, 2nd edn. San Diego: Academic.
Jou D (1996) *Extended Irreversible Thermodynamics*. New York: Springer.
Kondepudi D and Prigogine I (1998) *Modern Thermodynamics; from Heat Engines to Dissipative Structures*. Chichester: Wiley.
Prigogine I (1955) *Introduction to Thermodynamics of Irreversible Processes*. Oxford: Blackwell.
Putley EH (1960) *The Hall Effect and Related Phenomena*. London: Butterworth.
Tykody R (1967) *Thermodynamics of the Steady State*. New York: MacMillan.
Ure RW Jr. and Heikes RR (eds.) *Thermoelectricity: Science and Engineering*. New York: Interscience.
Wisniewski S (1976) *Thermodynamics of Nonequilibrium Processes*. Dordrecht: Reidel.
Ziman JM (1960) *Electrons and Phonons*. London: Oxford Press.

Isotopic Effects in Solids

E E Haller, University of California at Berkeley, Berkeley, CA, USA
J W Ager III, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Published by Elsevier Ltd.

Introduction

Close to 80% of the elements in the periodic table are multi-isotopic. Most semiconductors contain at least one element that consists of more than one stable isotope (Al, As, and P are important mono-isotopic exceptions). For example, silicon is composed of the three isotopes $^{28}_{14}$Si, $^{29}_{14}$Si, and $^{30}_{14}$Si with abundances of 92.23%, 4.67%, and 3.10%, respectively. Gallium arsenide consists of mono-isotopic As and two Ga isotopes $^{69}_{31}$Ga (60.11%) and $^{70}_{31}$Ga (39.89%). Recent advances in isotope separation technology have led to the availability of semiconductors with controlled the isotopic composition. Many physical properties of semiconductors depend on the isotopic composition. Some of these are affected relatively weakly while others, thermal conductivity for example, can be influenced strongly. Here, the isotope-related effects are divided into three groups: (1) those that are due to atomic mass differences between the isotopes; (2) those that are due to nuclear spin differences between isotopes; and (3) those that are due to differences in thermal neutron capture cross sections.

Atomic Mass-Related Properties

Structure

The lattice constants a_0 of large, chemically pure and crystallographically highly perfect crystals can be measured with high-precision X-ray diffraction techniques. As a consequence of the zero-point motion, the atomic volumes are expected to change with isotopic composition. The resulting effect on a_0 has been carefully studied for a number of elemental semiconductors including C, Si, and Ge. A theoretical analysis of the dependence of the lattice constant on isotopic composition at low temperatures compared to the Debye temperature of the material yields the following relationship:

$$\frac{\Delta a}{a} = -\frac{C}{a^3}\frac{\Delta M}{M}\left(\gamma_0\hbar\omega_0 + \frac{3}{4}\gamma_0 k\theta\right)$$

where C is the isothermal compressibility, M is the atomic mass, γ_o and γ_a are the Grüneisen parameters for optical and acoustic phonons, respectively, ω_o is the optical phonon frequency, k is Boltzmann's constant, and θ is the Debye temperature

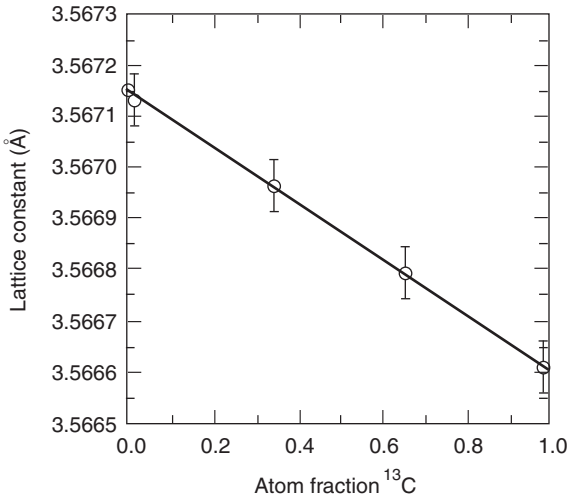


Figure 1 Lattice constant of diamond as a function of isotopic composition. (Adapted from Holloway *et al.* (1991) *Physical Review* B 44: 7123.)

(constant-mode Grüneisen parameters and a Debye frequency spectrum for the phonon modes are assumed). Evidently, the effect will be the largest for semiconductors with light atomic masses, large compressibilities, and high Debye temperatures. Experimental data for diamond are shown in **Figure 1**. The reduction of the lattice constant for ^{13}C diamond compared to ^{12}C diamond is 150 ppm, which is in qualitative agreement with the simple theory illustrated above.

Phonon Dispersion

A number of isotope effects in semiconductors are related to changes in the properties of phonons with atomic mass. A very simple harmonic analysis within the virtual crystal approximation predicts that the frequency ω of a given phonon should vary with the average mass M:

$$\omega \propto M^{-1/2}$$

As shown in **Figure 2**, this dependence is indeed observed for the zone-center optical phonon for a series of isotopically enriched single crystals of germanium with average masses between 70 and 76 amu (M of natural Ge is 72.59 amu).

Modern thin-film growth techniques have made isotopically engineered structures available to the experimentalist. Very interesting phonon effects can be observed in isotope superlattices (SLs), which consist of alternating layers with different isotopic composition. In contrast to SLs consisting of different materials such as GaAs and AlGaAs, the electronic band

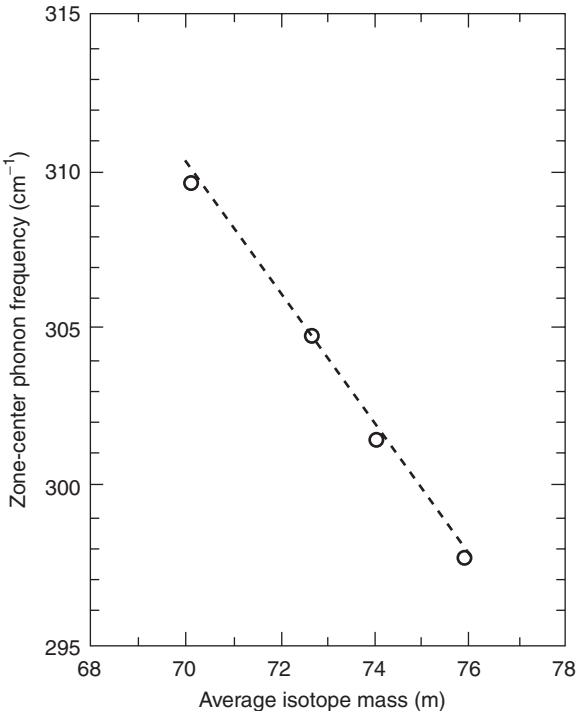


Figure 2 Zone-center phonon observed by Raman spectroscopy at 90 K for single-crystal germanium of varying isotopic composition. Dotted line shows the predicted $M^{-1/2}$ dependence. (After Fuchs HD *et al.* (1992) *Solid State Communications* 82: 225.)

structure in isotope SLs is not significantly affected by the mass differences and the material looks bulk-like for electrons throughout the crystal layers. However, the changes in the phonon properties are sufficiently large to observe confinement effects. A number of detailed experimental studies have been performed with Ge isotopes in structures consisting of a series of alternating sets of atomic planes of ^{A}Ge and ^{B}Ge where A and B are different atomic masses. Experimentally observed and theoretically calculated Raman spectra are compared in **Figure 3** for zone-center optical phonons for $^{70}\mathrm{Ge}_n/^{74}\mathrm{Ge}_n$ [0 0 1] oriented structures with n, the number of atomic layers, varying from 2 to 32. The Raman frequencies and mode mixing calculated by a simple bond-charge model, and the force constants of natural Ge and the relative peak intensities determined by a bond polarizability approach are in good agreement with the experimental data. The shortest periods studied consisted of two atomic layers each and produce a single Raman line corresponding to the zone-center optical phonon in a Ge crystal with $M = 72$ amu, the linear average of the two masses. That is, the phonon behavior is the same as in a bulk crystal with $M = 72$ amu. Confinement effects are observed at higher values of n as folded optical modes become

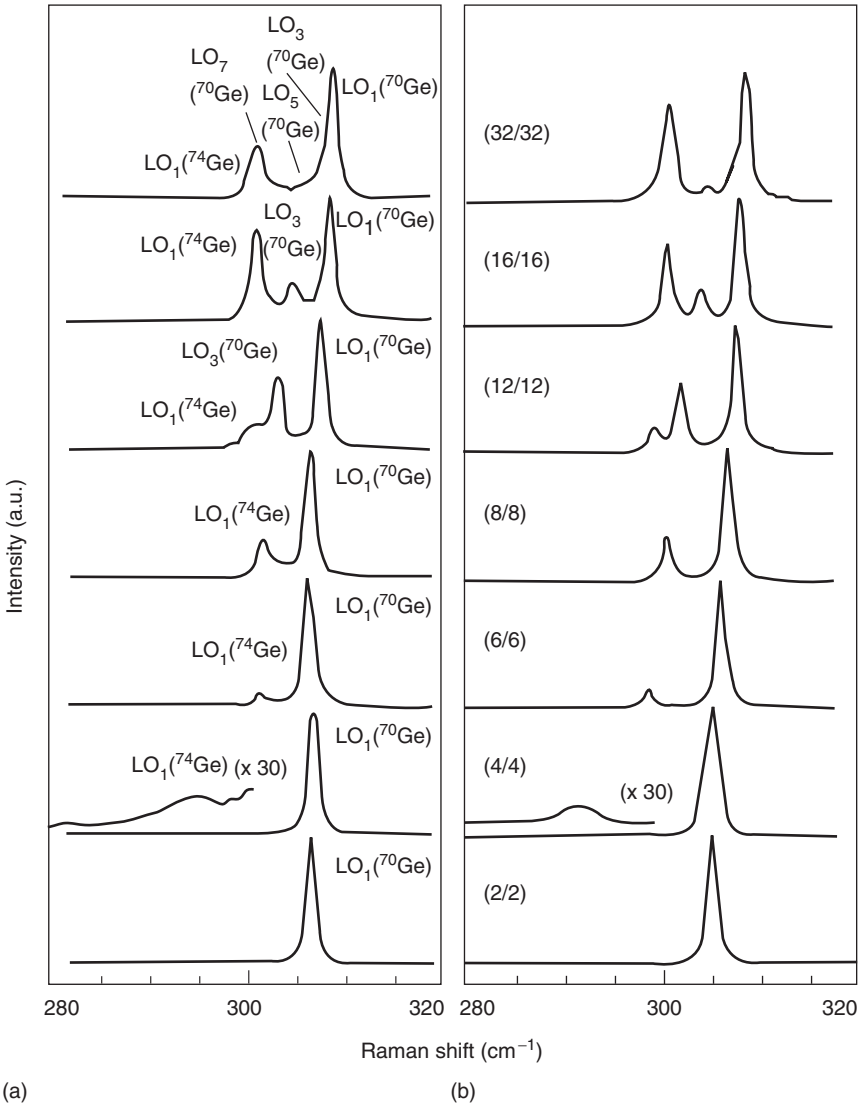


Figure 3 (a) Measured at $T = 10\,K$ and (b) calculated Raman spectra for a series of isotope superlattices $^{70}Ge_n/^{74}Ge_n$ grown along [0 0 1] with $n = 2$, 4, 6, 8, 12, 16, and 32. A number of modes confined to the ^{70}Ge and ^{74}Ge layers are evident. (Adapted from Spitzer *et al.* (1994) *Physical Review Letters* 72: 1565.)

Raman active at the zone center. For example, in the $n = 12$ superlattice at least three strong Raman lines are observed. For $n = 32$, the Raman spectra begin to approach the structure expected for two isotopically pure bulk crystals, one made up of ^{70}Ge, the other of ^{74}Ge.

Thermal Conductivity

The most dramatic phonon-related isotope effect is found for the thermal conductivity. Pomeranchuck was the first to point out in 1942 that in a multi-isotopic material, minority isotopes must be treated as mass defects and that they lead to scattering of acoustic phonons, reducing thermal conductivity with respect to a mono-isotopic material. The first experimental measurements of what can be a very strong dependence of the thermal conductivity (which depends on acoustic phonons) on small degrees of isotopic disorder were performed in the late 1950s in germanium. The most dramatic effect is found, however, in diamond. **Figure 4** illustrates the fivefold increase in thermal conductivity found near 100 K in isotopically enriched ^{12}C. (It should be pointed out that measurements of such extremely high thermal conductivities for millimeter-size samples are very difficult to perform because the samples act like close to perfect thermal short circuits.) The measured value, $410\,W\,cm^{-1}\,K^{-1}$, is the highest measured for any material above liquid nitrogen temperature (77 K). The most successful modeling of the effect of isotopic disorder on thermal con-

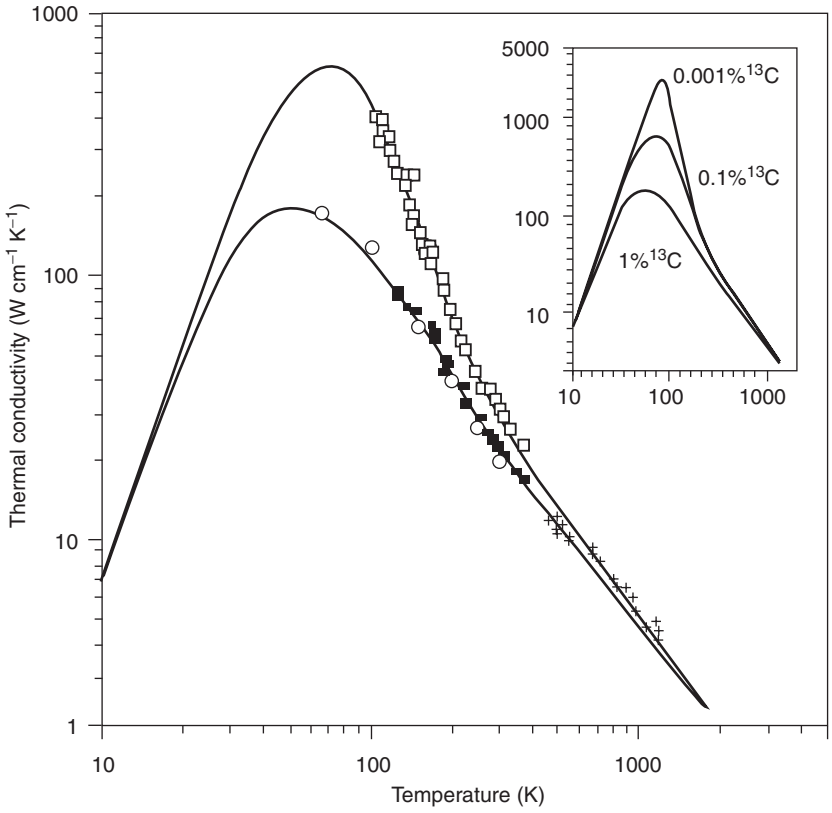


Figure 4 Thermal conductivity of natural abundance (1.1% ^{13}C) diamond (■), isotopically enriched (0.1% ^{13}C) diamond (□), together with the low-temperature data of Slack (1973), (○) and the high-temperature data of Olson *et al.* (1993) (×). The solid curves are the result of fitting the Callaway theory illustrated in the text. The inset shows the calculated thermal conductivity corresponding to 1%, 0.1%, and 0.001% ^{13}C concentrations according to the Callaway theory. (Adapted from Wei *et al.*(1993) *Physical Review Letters* 70: 3764.)

ductivity splits the thermal conductivity into four terms – normal, Umklapp, boundary, and isotope processes – which are inversely proportional to the phonon lifetime τ associated with the specific effect:

$$\tau^{-1}_{\text{Normal}} = AvT^3\lambda^{-1}$$

$$\tau^{-1}_{\text{Umklapp}} = BvT^3\lambda^{-1}\exp(-C/T)$$

$$\tau^{-1}_{\text{Boundary}} = v/D$$

$$\tau^{-1}_{\text{Isotope}} = 4\pi^3 v\lambda^{-4}V_0 x(1-x)(12+x)^{-2}$$

where v and λ are the average phonon velocity and wavelength, x is the isotope fraction of ^{13}C, V_0 is the atomic volume, and A, B, C, and D are adjustable parameters. These relations can be used not only to reproduce the observed temperature dependence but also to predict the limits of thermal conductivity for samples with very high isotopic purity. The inset in the figure shows the prediction of a thermal conductivity above 2000 W cm^{-1} K^{-1} for very highly enriched ^{12}C at 80 K.

Similar increases of the thermal conductivity at low temperature of the enriched semiconductors Ge and Si have been measured. It is important to note that at temperatures well above the thermal conductivity maximum, the isotope enrichment-related increases vanish rapidly. For example, the increase in thermal conductivity for highly enriched ^{28}Si compared to natural Si is in the range of 10–15%.

Electronic Structure

Isotope composition affects the electronic structure through the electron–phonon coupling and through the change of volume with the isotopic mass (see above section on lattice constants). The effect most often studied experimentally is the influence of the isotopic mass on the bandgap of the semiconductor. For example, using optical modulation techniques the isotope dependence of the direct and indirect bandgaps of germanium and of the indirect gap of silicon have been determined. The effect is small but satisfactorily predicted by theory. Another sensitive way to measure this dependence uses the very narrow no-phonon exciton photoluminescence (PL)

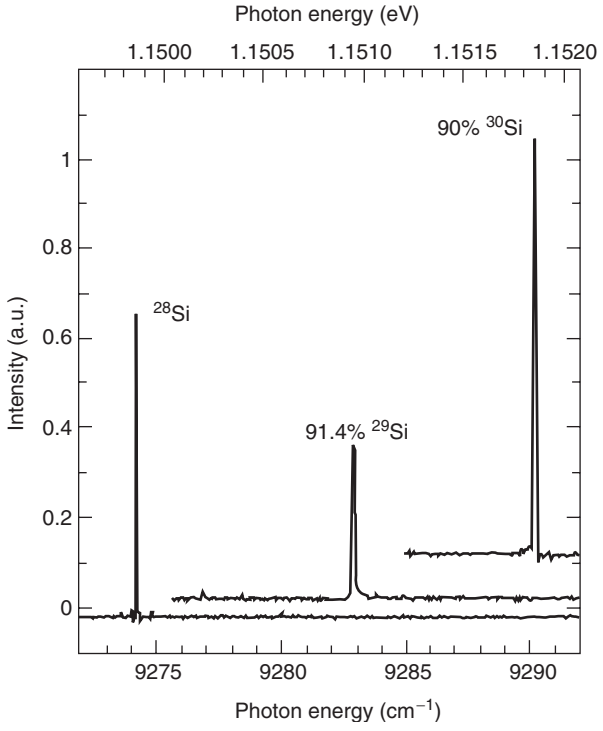


Figure 5 Effect of Si isotopic composition on the Si bandgap as observed by the shift of the P_{NP} PL line observed at 4 K as a function of isotopic enrichment. The three curves are offset for clarity.

lines found in Si and Ge. Measurements of the energies of these impurity-bound excitons allow direct determination of bandgap shifts because their radiative recombination does not require phonon participation. Furthermore, due to their large Bohr orbits, their energy depends only on the average isotopic mass and not on the isotopic disorder. **Figure 5** shows the change in the Si bandgap measured using the phosphorus no-phonon line P_{NP} for a series of samples enriched in ^{28}Si, ^{29}Si, and ^{30}Si.

In general, the contribution to the bandgap shift from the isotopically induced volume change is the smaller of the two effects and can be addressed with the methods described above for the lattice constant effects. The analysis of the electron–phonon coupling effect is similar to that used to understand the change of bandgap with temperature. The change is described by structure factors, which include electron–phonon interaction terms (Debye–Waller factors) and self-energy terms. These terms are then expanded in powers of the atomic displacements u. The leading terms are proportional to the mean-square displacements $\langle u^2 \rangle$. These, in turn, can be modeled as harmonic oscillators:

$$\langle u^2 \rangle = \hbar\left(\tfrac{1}{2} + n\right)/M\omega$$

As temperature increases, n and thus $\langle u^2 \rangle$ increase, and the bandgap is reduced. At low temperature, $n = 0$. Substituting the dependence of ω on M,

$$\langle u^2 \rangle \propto M^{-1/2}$$

For increasing M, this analysis predicts decreasing $\langle u^2 \rangle$ and a bandgap increase, in agreement with the observation.

Self- and Dopant Diffusion

The study of self- and dopant diffusion in semiconductors has been advanced significantly through the use of isotopically controlled multilayer structures. Diffusion measurements performed with radiotracer methods are limited by the half-lives of radioactive isotopes, by the damage produced during their introduction and by near-surface effects.

In contrast, the use of Secondary Ion Mass Spectrometry (SIMS) in isotope superlattice structures provides quantitative concentration profiles of the masses of the matrix as well as the dopant atoms. In the simplest case, two layers with different isotopes of a semiconductor are grown on a substrate of natural composition. The results of annealing such a structure are illustrated in **Figure 6** for a ^{nat}Ge/^{74}Ge/^{70}Ge/^{nat}Ge-substrate configuration. The ^{70}Ge concentration profile in the ^{74}Ge layer can be fitted with a simple complimentary error function (the solution to the 1D simple Fick's law diffusion equation) over a range of four and one half orders of magnitude in concentration, yielding very precise self-diffusion data. Similar structures of silicon isotopes have been used to obtain precise measurements of the Si self-diffusion coefficient shown as a function of the absolute inverse temperature (**Figure 7**).

The influence of the Fermi level (and hence of specific native defects) on self- and dopant diffusion has been studied with the aid of silicon isotope multilayers capped with an amorphous layer of natural silicon. Ion implantation into the amorphous layer was used as a source of the desired dopant and to affect Fermi level control inside the structure. Such a dopant source has been found to produce only the expected electronic effects that are due to dopant-related shifts in the Fermi level position. A representative SIMS result of silicon self- and As dopant diffusion is shown in **Figure 8**. The modeling of the silicon isotope and the arsenic concentration profiles yields the unambiguous result that the negatively charged vacancy is the major native defect involved in the diffusion process.

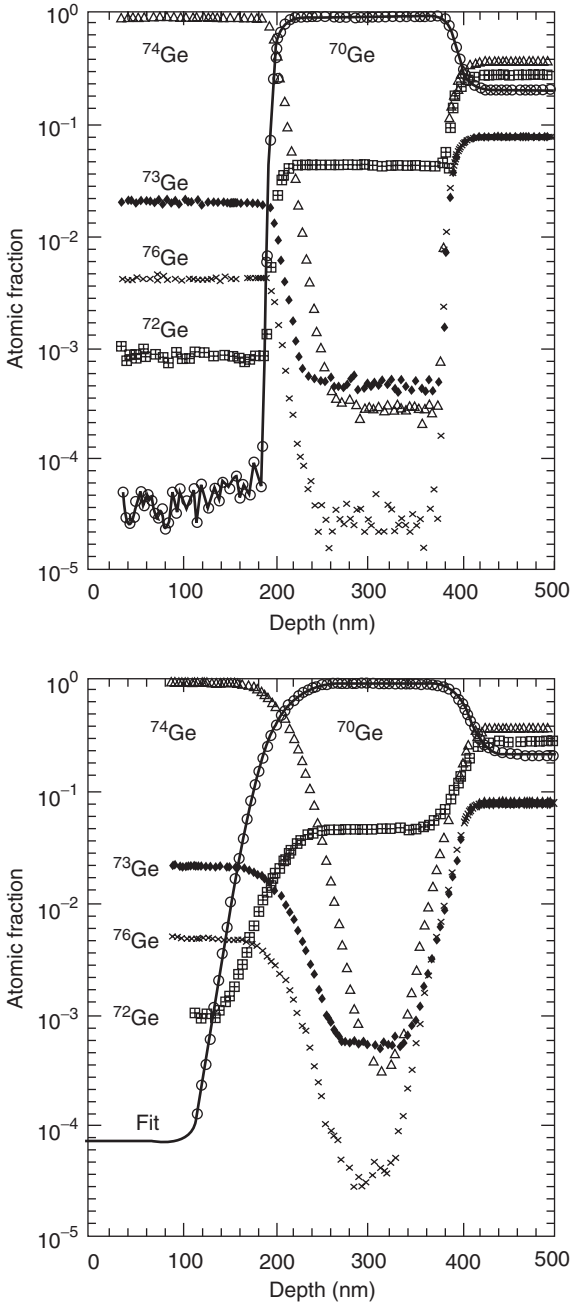


Figure 6 (a) SIMS profile of a ^{nat}Ge/^{74}Ge/^{70}Ge/^{nat}Ge-substrate structure before and (b) after annealing for 55.55 h at a temperature of 586°C . The self-diffusion of the ^{70}Ge into the ^{74}Ge layer can be modeled with a simple complimentary error function to a very high level of precision. (Adapted from Fuchs HD *et al.* (1995) *Physical Review* B 51: 16817-21.)

Isotopically controlled compound semiconductor structures offer the possibility of observing self-diffusion of each of the constituent atoms. Using this method, a very spectacular result of self-diffusion has been reported for the III–V semiconductor GaSb. Both host elements, Ga and Sb, have two isotopes occurring at comparable concentrations; a double-labeled ^{69}Ga^{121}Sb/^{71}Ga^{123}Sb heterostructure was used for the study. **Figure 9** shows the concentration profiles of the four isotopes ^{69}Ga, ^{71}Ga, ^{121}Sb, and ^{123}Sb for the as-grown structure and after thermal annealing at different times and temperatures. The Ga host atoms diffuse by a factor of $\sim$1000 faster than the Sb atoms. Even after four days at a temperature just under the bulk melting point of GaSb and long after the labeled Ga atoms have diffused into the bulk, the Sb atoms have barely moved. The large difference can be explained with the formation of Ga_{Sb} antisites that suppress the formation of Sb vacancies required for the diffusion of this constituent.

Local Vibrational Modes

The vibrations of light impurity atoms in a semiconductor can occur at much higher frequencies than the maximum single-phonon frequencies of the host lattice. In this case, the coupling to the host lattice is weak, such that this local vibrational mode (LVM) can have a long lifetime and a very narrow spectral line width as observed by infrared absorption, Raman scattering, or other techniques. Depending on the details of the lattice location of the impurity, particularly on symmetry and on how many bonds are formed with host lattice atoms, a number of normal vibrational modes are observed. Because of the simple dependence of vibrational frequencies on the inverse square root of the vibrating mass, isotope substitution of the impurity is used to obtain unambiguous proof for the association of an LVM feature with a certain impurity. The largest isotope shift caused by such a substitution of stable isotopes is close to 40% and is obtained for hydrogen substituted by deuterium.

In order to obtain a crude understanding of an LVM spectrum, it usually suffices to model the impurity and the nearest-bound neighbors as a quasimolecule embedded in the host crystal. The model can be improved substantially by taking into account next-nearest neighbor atoms in the form of an additional mass correction, the "host interaction mass" and by extending the linear force dependence on impurity displacement with higher terms (anharmonicity). When the host lattice consists of several stable isotopes with some of these present at significant atomic fractions (Ge is an example), LVM spectra can contain many closely spaced or overlapping lines because the impurity and its host lattice neighbors form quasimolecules with a number of different reduced masses. Isotope enrichment of the host material can lead to drastic simplification of the LVM spectra in such cases. This is illustrated in **Figure 10**

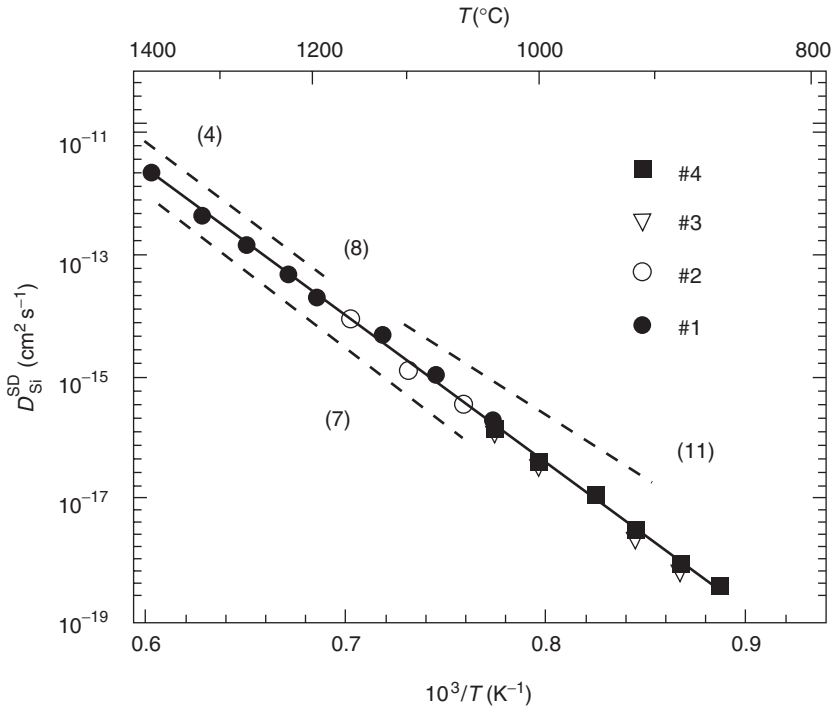


Figure 7 Self-diffusion coefficient of Si measured using multilayer structures of the stable isotopes. (The numbers in the figure refer to references in the original article.) (Adapted from Bracht H, Haller EE, and Clark-Phelps R (1998) *Physical Review Letters* 81(2): 393.)

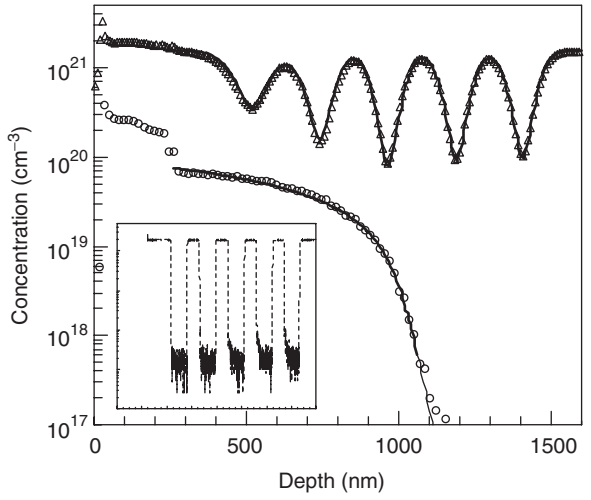


Figure 8 Depth profiles of ^{30}Si (△) and As (○) in an a-Si/(^{30}Si/$^{nat.}$Si)/$^{nat.}$Si structure with five ^{30}Si/$^{nat.}$Si periods, measured with secondary ion mass spectrometry (SIMS). Consistent modeling of both the dopant, As, and self-diffusion profiles (——) shows that the singly negatively charged self-interstitial is the mediating native defect responsible for the diffusion of both species. Inset figure shows the SIMS depth profile of the as-grown isotope structure, with the same axis limits. (Adapted from Silvestri HH, Ph.D. thesis, UC Berkeley 2004, unpublished.)

for the LVM of oxygen in Ge. Taking into account two nearest neighbors, the ^{x}Ge–O–^{y}Ge molecule has three fundamental modes. The "wag" mode (v_3) has the highest frequency. It leads to 11 lines resulting from the different mass combinations of ^{x}Ge and ^{y}Ge, each one split into several components whose intensities depend on temperature. Isotopic enrichment not only greatly simplifies the spectrum (one split line is observed instead of 11), it also allows a quantitative study of the line width and of the coupling between v_3 and the other modes as a function of temperature. In this way, it was shown that the observed splitting is a result of the nonlinear superposition of the v_2 and v_3 modes and that the temperature dependence is due to the thermal population of low-frequency v_2 mode, which can be significant, even at near liquid helium temperature.

Spin-Related Properties

A number of elements found in typical semiconductors have half-odd nuclear spins I and hence magnetic moments. Examples are ^{73}Ge ($I = 9/2$), ^{29}Si ($I = 1/2$), ^{31}P ($I = 1/2$), and both stable isotopes of Ga: ^{69}Ga ($I = 5/2$) and ^{71}Ga ($I = 5/2$). The magnitude of the magnetic moment and the spin-lattice (T_1), and transverse (decoherence, T_2) lifetimes of nonequilibrium nuclear spin populations are measured with nuclear magnetic resonance (NMR). In high-quality single-crystal material, lifetimes can be long. For example, the spin lattice relaxation time of the ^{29}Si (4.67% natural abundance) nuclear spin can be as long as several hours at room temperature in high-

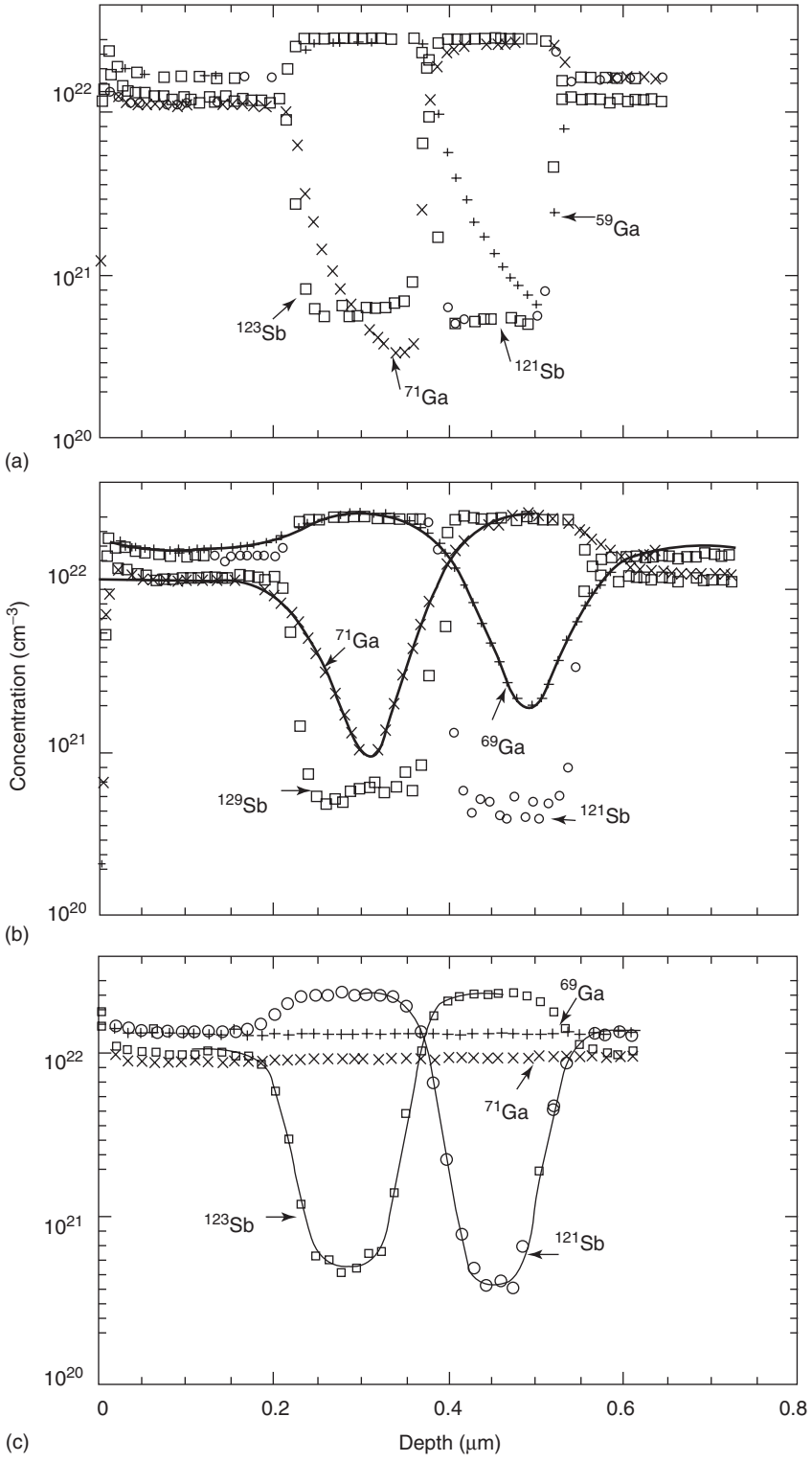


Figure 9 Concentration–depth profiles of Ga and Sb isotopes in GaSb isotope heterostructures measured with SIMS. Data are shown for ^{69}Ga (+), ^{71}Ga (×), ^{121}Sb (○) and ^{123}Sb (□). (a) Ga and Sb profiles of the as-grown ^{69}Ga^{121}Sb/^{71}Ga^{123}Sb heterostructure. Profiles after annealing the isotope structure under Sb-rich conditions at 700°C for 105 min (b) and 18 days (c). The Ga and Sb diffusion coefficients were determined by fitting the solution of Fick's law for self-diffusion across an interface to the Ga and Sb profiles measured after annealing. Solid lines in (b) and (c) show best fits. (Adapted from Bracht H, Nicols SP, Walukiewicz W, Silvera JP, and Briones F (2000) *Nature* 408(6808): 69–72.)

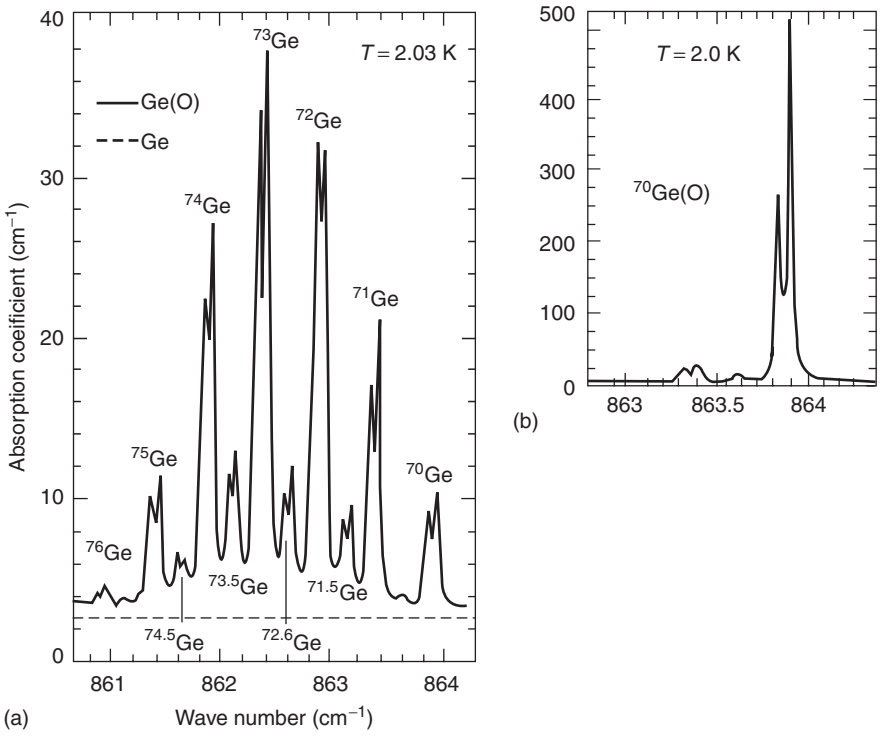


Figure 10 (a) Spectrum of the Ge_2O quasimolecules in natural Ge at 2.05 K. Eleven lines split by $\nu_2 + \nu_3$ coupling excitations correspond to the distinct isotopic mass combinations of the two Ge atoms in the quasimolecule. An oxygen-free sample produced the dashed line. (b) The spectrum of an oxygen-doped, highly enriched sample of ^{70}Ge. The lines labeled ^{71}Ge are caused by traces of ^{72}Ge in ^{70}Ge. (Adapted from Mayur *et al.* (1994) *Physical Review* B 49: 16293.)

purity single crystals; decoherence times are of the order of milliseconds (ms) at cryogenic temperatures. In the case of electron spin resonance (ESR), soon after the demonstration of the effect in the late 1950s, the effects of isotopic enrichment on the spin dynamics of donor electrons bound to ^{31}P were investigated. At that time, a doubling of T_2 from 0.2 to 0.5 ms was observed by removing the hyperfine spin–orbit coupling interaction between the electron spin and the nuclear spin of the ^{29}Si. Modern measurements on higher purity $>99\%$ ^{28}Si-enriched material have demonstrated $T_2 > 60$ ms. These and related effects are expected to be important in developing solid-state quantum computing schemes based on electron and/or nuclear spin as "qu-bits." Two references to this rapidly evolving field are given in the "Further reading" section.

Neutron Transmutation Doping

The extreme versatility of semiconductors in their wide range of applications is based on our ability to dope these solids with donors and acceptors. For device fabrication, dopants are introduced at or near the surface either by diffusion or by ion implantation. Bulk crystal doping is typically achieved by adding the appropriate minute quantities of dopant elements to the melt from which the crystal is pulled. An alternative doping technique is neutron transmutation doping (NTD).

The radioactive decay of an isotope of one element into an isotope of a different element forms the foundation of NTD of semiconductors. The NTD of Si is illustrated here. The NTD process is typically performed in a nuclear reactor. During exposure of natural silicon to thermal neutrons, all three isotopes capture neutrons with well-known capture cross sections, σ_n, and the following reactions occur:

$$^{28}_{14}\mathrm{Si} + \mathrm{n}_{th} \rightarrow {}^{29}_{14}\mathrm{Si}$$

$$^{29}_{14}\mathrm{Si} + \mathrm{n}_{th} \rightarrow {}^{30}_{14}\mathrm{Si}$$

$$^{30}_{14}\mathrm{Si} + \mathrm{n}_{th} \rightarrow {}^{31}_{14}\mathrm{Si} \overset{T_{1/2}=2.62\,h}{\longrightarrow} {}^{31}_{14}\mathrm{P} + \beta^- + \upsilon_e^-$$

The third reaction leads to the formation of the donor P. NTD differs from other dopant technologies in that very high doping uniformity can be obtained. The unmatched doping uniformity achieved by NTD is based on three factors. First, the stable isotopes are truly randomly distributed in a semiconductor of

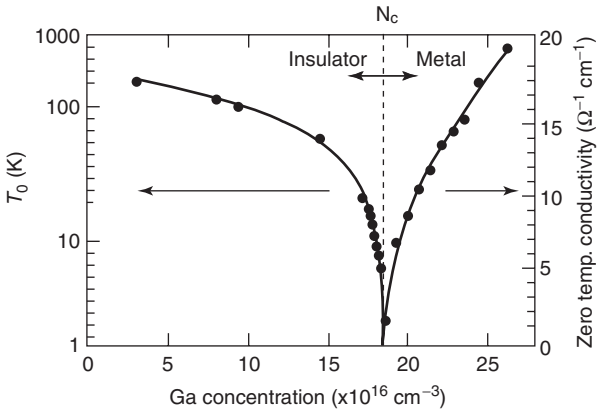


Figure 11 Extrapolated zero-temperature electrical conductivity (right axis) and characteristic temperature T_0 (left axis) for variable range hopping for 14 samples with carefully controlled concentrations of ^{71}Ga-doped ^{70}Ge made with neutron transmutation doping. The MIT is found to occur at an acceptor concentration of $1.86 \times 10^{17}\,\text{cm}^{-3}$. (Adapted from Itoh KM *et al.* (1996) *Physical Review Letters* 77: 4058.)

natural composition. Second, the source of thermal neutrons is typically larger than the semiconductor crystal to be doped. This, in turn, leads to a large homogeneous thermal neutron field guaranteeing identical exposure of all parts of the crystal. Third, the capture cross section for thermal neutrons is typically of the order of a few barns ($1\,\text{barn} = 10^{-24}\,\text{cm}^2$), which means that a few centimeters of material absorb the thermal neutron flux by significantly less than 1%, maintaining the nearly uniform neutron flux throughout the depth of the sample.

For this reason, semiconductors that are used in high-power, high-voltage applications, such as Si rectifiers (SCRs) are manufactured using NTD Si. Scientifically, NTD enables researchers to control doping levels more precisely. For example, it is possible to precisely dope Ge with Ga very close to the metal–insulator transition (MIT). As shown in **Figure 11**, it was found that the critical concentration was $1.86 \times 10^{17}\,\text{cm}^{-3}$ when the MIT was approached from both above and below. In addition, instrumentation for far infrared and submillimeter wave astronomy and astrophysics have profited from NTD. Small temperature-sensing elements of crystalline NTD germanium are assembled into large arrays in the focal planes of all modern far IR telescopes.

See also: Conductivity, Thermal; Electron Spin Resonance; Elemental Semiconductors, Electronic States of; Excitons in Crystals; Scattering, Inelastic: Raman; Semiconductor Compounds and Alloys, Electronic States of; Semiconductors, General Properties; Semiconductors, Impurity and Defect States in; Semiconductors, Optical and Infrared Properties of.

PACS: 25.40. – h; 63.20. – e; 63.20.Pw; 66.70. + f

Further Reading

Barker AS Jr. and Sievers AJ (1975) Optical studies of the vibrational properties of disordered solids. *Reviews of Modern Physics* 47(2): S1–S168.

Fahey PM, Griffin PB, and Plummer JD (1989) Point defects and dopant diffusion in silicon. *Reviews of Modern Physics* 61: 289–384.

Haller EE (1995) Isotopically engineered semiconductors. *Journal of Applied Physics* 77: 2857–2878.

Haller EE (1999) Semiconductors, isotopically engineered. In: Trigg G (ed.) *Encyclopedia of Applied Physics* pp. 257–283. New York: Wiley-VCH.

Kane BE (2000) Silicon-based quantum computation. quant-ph/0003031.

Ladd TD, Goldman JR, Yamaguchi F, Yamamoto Y, Abe E, and Itoh KM (2002) All-silicon quantum computer. *Physical Review Letters* 89: 011901–1–4.

Larabee RD (ed.) (1984) *Neutron Transmutation Doping of Semiconductor Materials*. New York: Plenum.

Lederer CM and Shirley VS (eds.) (1978) *Table of Isotopes*. 7th edn. New York: Wiley

McCluskey MD (2000) Local vibrational modes of impurities in semiconductors. *Journal of Applied Physics* 87: 3593–13617.

Pomeranchuk I (1942) On the thermal conductivity of dielectrics at temperatures lower than that of Deybe. *Journal of Physics (USSR)* 6: 237–250.

J

Jahn–Teller Effect

E E Vogel, Universidad de La Frontera, Temuco, Chile

Introduction

All the phenomena that can be explained by the coupling between vibrational modes of the material and electronic orbitals are said to be due to the Jahn–Teller effect (JTE). In its initial proposal (previous to experimental confirmation), the JTE anticipated a decrease in energy and degeneracy for nonlinear molecules with degenerate ground levels. This concept was later extended to localized electronic orbitals in crystals, ions, and vacancies. In another manifestation of the vibronic coupling, states of the same symmetry mix among themselves, changing energy intervals with respect to the states of different symmetry; such a phenomenon (which affects even singlets) is called the dynamic JTE. When the interaction is mediated by higher-energy states, it is referred to as pseudo-JTE. When the JTE effect extends through the lattice, neighboring centers cooperate causing a decrease in the ground-state energy. Thus, for instance, if one cell elongates along one direction due to JTE, a neighboring cell contracts along that direction. Such long-range ordering is reached below a critical temperature producing what is called the "cooperative JTE," which can modify several properties of the solid. Nevertheless, the basic phenomenon is always the same and can be described by means of coupling of the vibrational degrees of freedom to the electronic degrees of freedom of the system, in what generically can be called a "vibronic coupling."

The crystalline field theory can be thought of as a static electric multipole expansion of ionic and bond charges, defined by the equilibrium positions of constituent atoms. Such a symmetric field usually gives rise to a degenerate ground level or multiplet. Atoms interact due to bonding forces, which allow the propagation of vibrations through the crystal. Vibrational modes (phonons in quantum mechanical language) may raise the energy of some of the components of the electronic multiplet, while lowering the energy of other electronic states of the same multiplet. Thus, the JTE is usually associated with a splitting of the ground multiplet, which decreases the effective ground-state energy and reduces the ground-level degeneracy. In the real lattice, this can manifest itself as a local distortion (static JTE). In configuration space, this means complex functions describing observables and electronic properties (dynamic JTE).

Exceptions to the previous descriptions of the JTE are one-dimensional systems or when the degeneracy is due to a time-reversal symmetry (Kramers degeneracy). In all other cases, there is always a chance for a JTE. In some materials, properties can be explained by ignoring the JTE, which simply implies that the coupling is extremely weak. However, in other systems, a vibronic coupling can lead to structural deformations, specific heat anomalies, extra lines in optical spectra, or quenching of orbital magnetic components.

In any material, vibrations span wavelengths that range from their external dimensions to a couple of interatomic distances. Evidently, long wavelengths (elastic wave limit) do not change the relative positions of neighboring atoms, so these modes can be ignored as possible coupling modes. The best candidates for JT coupling are short wave vibrations, with large acoustic frequencies or optical modes. Thus, the JTE can be viewed as the coupling of phonons having an appropriate symmetry to which the electronic orbitals under concern are most sensitive. Due to such a variety of possibilities, it is not possible to give a general prescription to proceed with the JTE in any solid. It is necessary to examine each possible coupling to different electronic orbitals separately. Since symmetry is such an important issue in this problem, the group theory is used to designate the couplings. This is the reason why this tool is briefly described in this article; it is assumed that the reader is not a group theory expert. The particular case of a local environment with triangular symmetry is used to explain the essence of the JTE for one of the most common JT couplings ($E \times \varepsilon$), prior to which the symmetry elements for such an environment are quickly reviewed. Lastly, some examples of real manifestations of the JTE are given.

Electrons and Phonons in Triangular Environment

Two-dimensional systems are simple to understand and easy to represent on the paper. Consider a double-degenerate electronic system with states $|\theta\rangle$ and $|\varepsilon\rangle$ located at the center of a crystal with triangular symmetry in two dimensions as illustrated in **Figure 1**. There are six symmetry elements that leave this complex unaltered: E, the identity; off-plane rotations through angle π with respect to axes A, B, and C; in-plane rotations through angles $+2\pi/3$ (D) and $-2\pi/3$ (F). If the following column vector representations are used,

$$|\theta\rangle = \begin{vmatrix} 1 \\ 0 \end{vmatrix}, \qquad |\varepsilon\rangle = \begin{vmatrix} 0 \\ 1 \end{vmatrix}$$

the six symmetry elements can be represented by the following matrices:

$$E = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \qquad A = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \qquad B = \begin{vmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{vmatrix}$$

$$C = \begin{vmatrix} +\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{vmatrix}, \qquad D = \begin{vmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ +\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{vmatrix}$$

$$F = \begin{vmatrix} -\frac{1}{2} & +\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{vmatrix}$$

The cluster approximation is used, which assumes that only displacements of atoms that are close neighbors to the electronic cloud are included, which is similar to considering an isolated molecule. This is illustrated in **Figure 2** where nearest neighbors (numbered 1, 2, and 3) are considered. From the six degrees of freedom of these three atoms, two of them lead to the motion of the center of mass and they must be discarded for the present purposes. The remaining four degrees of freedom leave the center of mass fixed, allowing for interactions with the electronic orbitals at the center. With the aid of the axes x_i, y_i ($i = 1, 2, 3$) in **Figure 2**, the corresponding four normal modes can be expressed as

$$Q_B = \frac{1}{\sqrt{3}}(y_1 + y_2 + y_3)$$

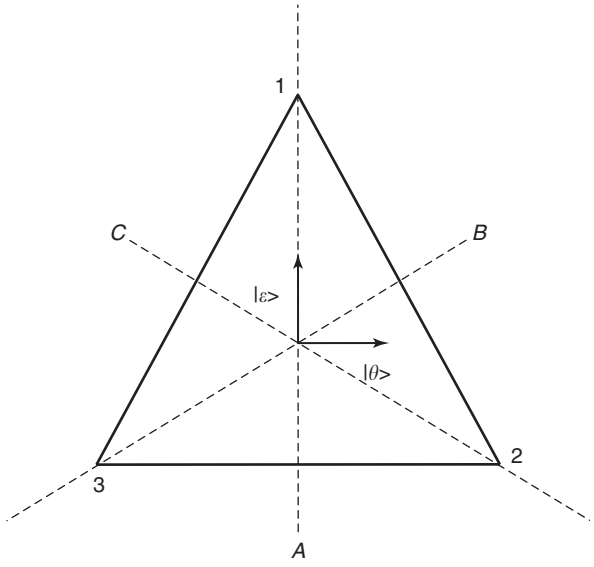


Figure 1 Symmetry elements for the triangular symmetry. Three off-plane rotations through angle π over axes A, B, and C. Additionally, operations D and F are in-plane rotations through angles $\pm 2\pi/3$.

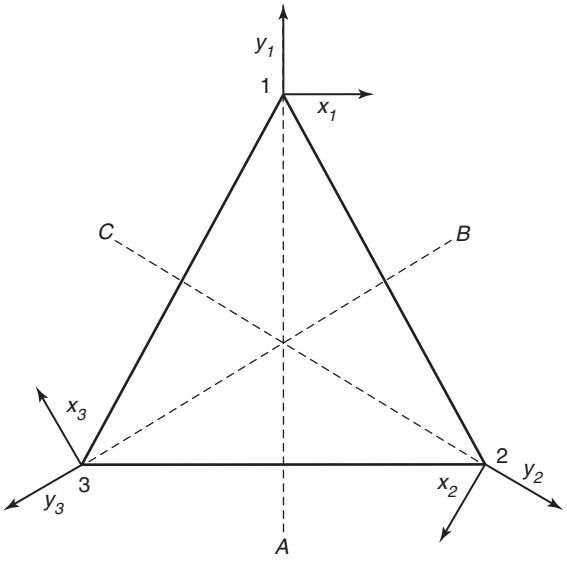


Figure 2 Basis coordinates for defining normal modes for the triangular cluster.

$$Q_R = \frac{1}{\sqrt{3}}(x_1 + x_2 + x_3)$$

$$Q_\theta = \frac{1}{\sqrt{3}}\left(x_1 - \frac{1}{2}x_2 + \frac{\sqrt{3}}{2}y_2 - \frac{1}{2}x_3 - \frac{\sqrt{3}}{2}y_3\right)$$

$$Q_\varepsilon = \frac{1}{\sqrt{3}}\left(y_1 - \frac{\sqrt{3}}{2}x_2 - \frac{1}{2}y_2 + \frac{\sqrt{3}}{2}x_3 - \frac{1}{2}y_3\right)$$

Table 1 Table of characters of group C_{3v}

Basis functions	*Irred. representation*	E	$3\sigma_v$	$2C_3$
Q_B	A_1	1	1	1
Q_R	A_2	1	−1	1
Q_θ, Q_ε	E	2	0	−1

First class E; second class $3\sigma_v$ (3 off-plane rotations A, B and C); third class $2C_3$ (2 in-plane rotations D and F).

The transformation properties of these modes can be obtained by applying symmetry operators to the basis axes x_i and y_i ($i = 1, 2, 3$). Thus, vertices (1, 2, 3) obey the following transformation laws under these operators: $E \rightarrow (1,2,3)$; $A \rightarrow (1,3,2)$; $B \rightarrow (2,1,3)$; $C \rightarrow (3,2,1)$; $D \rightarrow (2,3,1)$; and $F \rightarrow (3,1,2)$. This leads to transformation laws such as, $Ax_2 \rightarrow -x_3$, $Fy_3 \rightarrow y_1$, etc. When these transformation rules are applied to the four normal modes given by the equations above, the following properties are found:

1. E, the identity operator does not change anything (character of this operation is said to be $+1$).
2. A, B, and C leave Q_B unchanged (character $+1$), while it reverses the sense of Q_R (character -1).
3. D and F preserve both Q_B and Q_R (character $+1$) in both cases.
4. When the six operations are applied to Q_θ and Q_ε, the same transformation matrices given above for the transformation of electronic orbitals $|\theta\rangle$ and $|\varepsilon\rangle$ of **Figure 1** are obtained. The character here is the trace of the matrix (character $+2$ for E; 0 for A, B, and C; –1 for D and F).

With this information, one can build the table of characters corresponding to the triangular point group under consideration (usually designated by C_{3v}) (**Table 1**).

$E \times \varepsilon$ Coupling

Each electronic orbital couples differently to each set of vibrational modes. It is assumed here that the ground electronic level is a doublet whose wave functions transform as basis functions for the E irreducible representation of the C_{3v} point group (see **Table 1**). Such an electronic cloud sits at the center of a triangular crystal field. In the case of mode Q_R, the three atoms move simultaneously clockwise or counterclockwise, in what is a rotation or twist of the local surrounding. However, the electronic cloud at the center will not feel the atoms approaching it, so the coupling, if any, is expected to be very weak. In mode Q_B, the three neighboring atoms move outwards and inwards, in what is called the "breathing mode." This mode preserves the triangular symmetry.

The best candidates for coupling leading to a loss of symmetry are the other two modes, which involve a partial motion of atoms approaching the center. As already proved, these modes form the basis functions for the E irreducible representation of C_{3v}. However, it is customary to use Greek symbols for vibrations, so these are referred to as ε. In summary, this is the celebrated $E \times \varepsilon$ JT coupling, whose Hamiltonian can be written as

$$\boldsymbol{H} = \boldsymbol{H}_E + \boldsymbol{H}_{\mathrm{JT}} + \boldsymbol{H}_\varepsilon$$

where $\boldsymbol{H}_E$ is the component of the electronic Hamiltonian involving doublet E described by functions $|\theta\rangle$ and $|\varepsilon\rangle$; H_ε is the component of the vibrational Hamiltonian corresponding to coupling modes Q_θ and Q_ε, with vibrational quantum $\leqslant\omega$ and effective mass μ; H_{JT} is the coupling term which, in a first approximation, is given by a scalar product between a vector $\boldsymbol{G}$ of the electronic space and a vector $\boldsymbol{Q}$ of the vibrational space.

Static Case

The potential energy associated to the coupling modes is the static potential, namely

$$V = U + V_{\mathrm{JT}} = \tfrac{1}{2}\mu\omega^2\left(Q_\theta^2 + Q_\varepsilon^2\right) + K(G_\theta Q_\theta + G_\varepsilon Q_\varepsilon)$$

where the first term is the usual potential term in the harmonic approximation, while V_{JT} corresponds to the actual coupling whose strength is given by K. Normalized electronic operators G_θ and G_ε satisfy the following matrix representation on a basis formed by the vectors, $|\theta\rangle$ and $|\varepsilon\rangle$:

$$G_\theta = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix}, \qquad G_\varepsilon = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

(This can be found by studying the transformation properties of the matrix elements. For instance, under operator A, the element $\langle\theta|G_\theta|\varepsilon\rangle$ becomes $\langle+\theta|+G_\theta|-\varepsilon\rangle = -\langle\theta|G_\theta|\varepsilon\rangle = 0$.)

A general ground-state function can be written as

$$|f\rangle = \alpha|\theta\rangle + \beta|\varepsilon\rangle$$

which must be an eigenfunction of the static potential:

$$V|f\rangle = F|f\rangle, \qquad (F-U)|f\rangle = V_{\mathrm{JT}}|f\rangle$$

By letting operators G_θ and G_ε act on $|f\rangle$, it is obtained that

$$(F-U)(\alpha|\theta\rangle + \beta|\varepsilon\rangle) = K[(-Q_\theta\alpha + Q_\varepsilon\beta)|\theta\rangle + (Q_\theta\beta + Q_\varepsilon\alpha)|\varepsilon\rangle]$$

which leads to a secular equation giving the condition:

$$(F-U)^2 - K^2\left(Q_\theta^2 + Q_\varepsilon^2\right) = 0$$

A minimum energy is obtained by imposing the condition that the first derivative of the eigenvalue function F with respect to each coordinate is zero at the equilibrium position $(Q_{\theta 0},\ Q_{\varepsilon 0})$, yielding

$$F_0 = -U_0 = -\frac{\mu\omega^2}{2}\left(Q_{\theta 0}^2 + Q_{\varepsilon 0}^2\right) = -\frac{K^2}{2\mu\omega^2}$$

suggesting the parametrization,

$$Q_{\theta 0} = Q_0 \cos\varphi, \quad Q_{\varepsilon 0} = Q_0 \sin\varphi, \quad Q_0 = \frac{K}{\mu\omega^2}$$

as illustrated in the lower part of **Figure 3**.

The interpretation here is clear: new equilibrium conditions appear at a "distance" Q_0 away from the non-JT equilibrium position $(K=0)$. This is the JT distortion. The actual energy is now

$$F_0 = -\frac{\mu\omega^2}{2}\,Q_0^2 = -E_{\mathrm{JT}}$$

where E_{JT} is the so-called JT energy corresponding to the decrease from the non-JT situation. This is illustrated as the minimum energy reached by the lower curve in the upper part of **Figure 3**. Instead of the initial electronic doublet, the ground state corresponds to the singlet (lower state)

$$|f_{\mathrm{l}}\rangle = \cos\left(\frac{\varphi}{2}\right)|\theta\rangle - \sin\left(\frac{\varphi}{2}\right)|\varepsilon\rangle$$

Thus, JTE in this complex resulted in a ground level of lower energy and lower degeneracy.

The second static state is represented by the orthonormal function (upper state)

$$|f_{\mathrm{u}}\rangle = \sin\left(\frac{\varphi}{2}\right)|\theta\rangle + \cos\left(\frac{\varphi}{2}\right)|\varepsilon\rangle$$

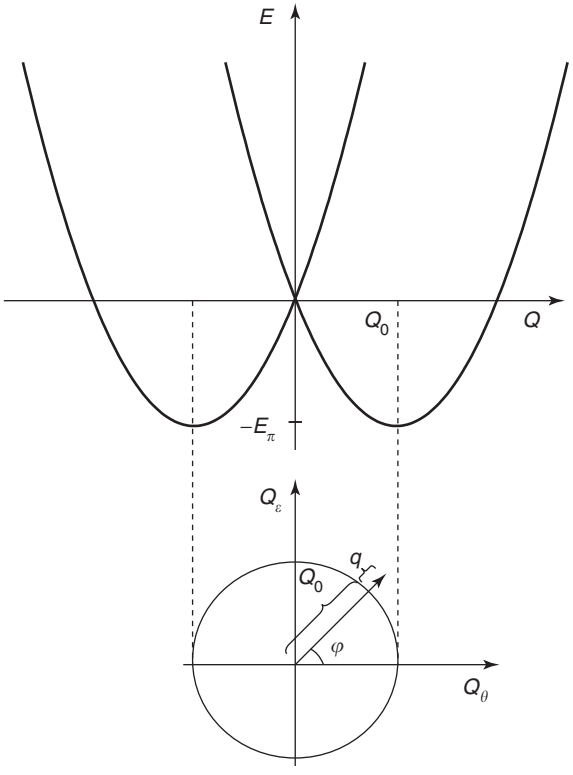


Figure 3 Parametrization of coordinates that result from solving the static problem for $E\times\varepsilon$ coupling. The coupling defines a new equilibrium condition at coordinates Q_0,φ.

Dynamic Picture

At $K=0$ (no JT), the potential can be represented by a paraboloid of revolution with its minimum at the origin of the $Q_\theta Q_\varepsilon$ plane. As K grows, the minimum energy is $-E_{\mathrm{JT}}$, occurring at $Q=Q_0$, for any φ (circle in **Figure 3**). A cut of the corresponding potential surface is illustrated in the upper part of **Figure 3**. The complete surface is obtained by revolving this curve about the E-axis; due to its shape, this potential is known as the Mexican hat.

Notice that the radial degree of freedom in the plane $Q_\theta Q_\varepsilon$ leads to oscillations, while excitations along the perpendicular direction do not find restoring forces leading to a circular motion associated to the coordinate φ. Since a transformation $\varphi\rightarrow(\varphi+2\pi)$ does not lead to the same function, a dynamic ground-state wave function must be written as

$$|f_{\mathrm{l}m}\rangle = \mathrm{e}^{\mathrm{i}m\varphi}\left(\cos\left(\frac{\varphi}{2}\right)|\theta\rangle - \sin\left(\frac{\varphi}{2}\right)|\varepsilon\rangle\right)$$

representing the family of low-energy wave functions, with m taking half-integer values. Higher energy states correspond to an admixture of oscillations $(Q\rightarrow Q_0+q)$ with occupation number n, rotational states with occupation number m and electronic basis function $|f_{\mathrm{l}}\rangle$. Such wave functions form the lower branch of vibronic levels. Similarly, states originating in static solution $|f_{\mathrm{u}}\rangle$ produce the upper branch of states.

Strength of JT Coupling

In the extremely weak coupling limit $(K=0)$, the system can be described as a doubly degenerate electronic system uncoupled to the vibrations of the surrounding, which can be represented here by a double-dimensional harmonic oscillator with occupancy quantum numbers n_θ and n_ε, energy $\hbar\omega(n_\theta+n_\varepsilon+1)$ and total vibronic degeneracy $2(n_\theta+n_\varepsilon+1)$ as represented on the left-hand side of **Figure 4**.

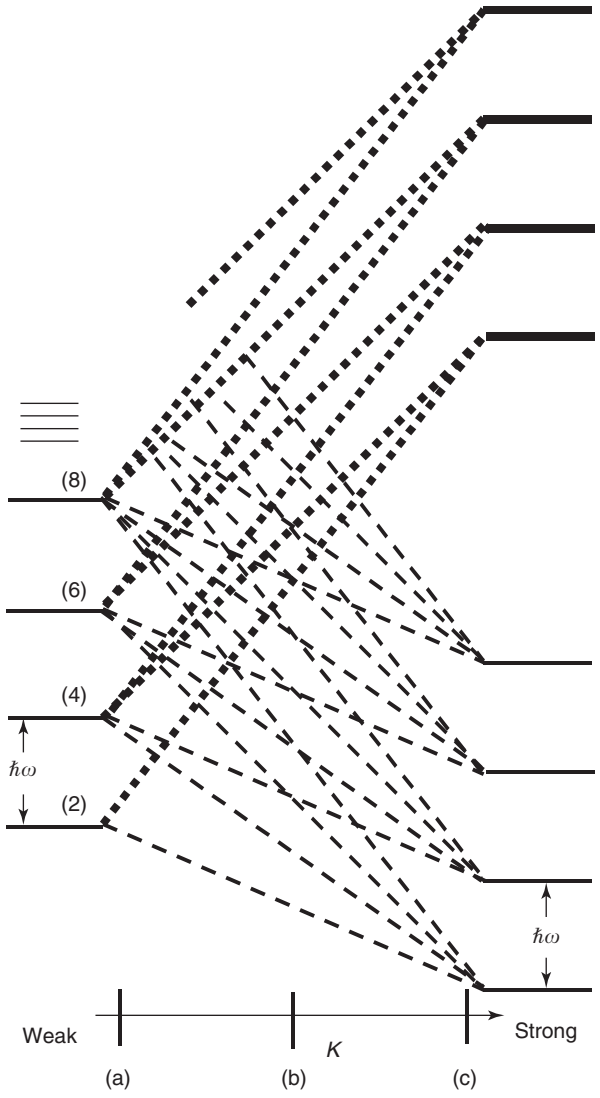


Figure 4 Schematic strength dependence of energy levels for $E \times \varepsilon$ coupling: (a) weak coupling limit, (b) intermediate coupling with admixture and crossing of levels, and (c) strong coupling limit.

On the other hand, in the extremely strong coupling limit, the system is better described by a coupled vibronic system, with states belonging to two branches. In each case, the wave function involves function $|f_l\rangle$ or $|f_u\rangle$ combined to quantum numbers n and m, producing a complex function whose actual form depends on the strength of the coupling. For large enough values of K, the energy for levels of the lower branch can be approximated by a series expansion, whose three leading terms are

$$E_{nm} = -K^2\hbar\omega + \left(n + \frac{1}{2}\right)\hbar\omega + \frac{m^2}{4k^2}\hbar\omega$$

The degeneracy of each vibronic level is infinite in the strong coupling limit ($K \to \infty$), due to the infinite possible values of m.

Figure 4 illustrates how the infinite energy levels of the weak coupling limit map onto the infinite levels of the strong coupling limit, for the $E \times \varepsilon$ coupling. All kinds of situations are possible according to the strength of the coupling. For the extremely weak coupling limit (case (a) in **Figure 4**), the JTE can be thought of as a perturbation to the adiabatic limit: no permanent distortion is to be found and physical properties show little difference with respect to the description coming from the plain crystal-field theory. On the contrary, in the strong coupling limit ((c) in **Figure 4**) large discrepancies in structure, optical and magnetic properties are possible. In such cases, the coupling is beyond perturbative methods and the system is better described by a set of displaced oscillators. Intermediate cases (such as (b) in **Figure 4**) can be extremely complicated as well as interesting, and the problem usually requires numerical methods to characterize it.

A way of measuring the strength of the coupling is the so-called Huang–Rhys factor S, defined as

$$S = \frac{E_{JT}}{\hbar\omega}$$

Normally, a weak coupling is understood for $S \leqslant 1.0$, while a strong coupling is found for $S \geqslant 10$, but this is a general guide only.

Ham Reduction Factors

The nature of the ground-wave function varies with K. In the extremely weak coupling limit, the ground-state function is a true zero-phonon wave function. However, as the strength of the coupling increases, the admixture with higher-order phonon components weakens the presence of the zero-phonon component. Thus, the contribution of purely electronic operators, such as those present in optical and magnetic properties, decreases with the strength of the vibronic coupling and it is said that the corresponding effect is quenched.

The expectation value corresponding to any given operator makes use of the vibronic functions at the appropriate K value. If JTE is ignored, expectation values make use of the purely electronic functions. When a comparison between these two results is done for the case of the ground states, reduction factors can be defined. Thus, the ratio of the first expectation value over the second one is called the Ham factor or a reduction factor for that operator. These factors give an idea of how much a physical magnitude is reduced by the JTE. For $E \times \varepsilon$ coupling,

the reduction factor $q_H(K)$ is defined as

$$q_H(K) = \frac{\langle f^K_{00\theta}|G_\theta||f^K_{00\theta}\rangle}{\langle\theta|G_\theta|\theta\rangle} = \frac{1}{2}\left(1 + \frac{1}{(2K)^4} + \cdots\right)$$

where the truncated asymptotic expression on the right-hand side is valid in the strong coupling limit only ($K \to \infty$). So, in this limit ($K \to \infty$), q tends to 0.5; then, as the coupling weakens, q grows to reach the value 1.0 in the case of no JTE ($K = 0$). Evaluation of other Ham factors can yield results in the range (0.0,1.0).

Quenching of Orbital Angular Momentum

One of the most frequently observed manifestations of the Ham effect (as it is sometimes called) is the quenching of the orbital angular momentum. The magnetic moment of ions is formed by both orbital and spin components. For free ions, both of them contribute fully to produce the final magnetic moment. In a molecule or solid, the orbital contribution depends on the zero-phonon component of the ground vibronic wave function. So, it is said that the angular momentum is quenched and the magnetic moment of the ion is different as compared with the value of the free ion. This phenomenon is directly measured by means of electron paramagnetic resonance and it is usually reported as a variation in the g factor or, more generally, as a change in the components of the g tensor.

Examples of Dynamical JT Couplings

Couplings

There is a large variety of couplings, in symmetry as well as in strength. For simplicity, the case of impurities in crystals is considered so that the previous analysis based on the point-group theory can be directly invoked. Transition metal impurities in semiconductors have been deeply investigated due to their optical properties and possible applications as sensors. One interesting system is the one defined by the presence of Fe impurities in II–VI compounds of strong ionic character such as CdTe; this case is used as an illustrative example. The crystal structure is that of zinc blende, where each cation is surrounded by four anions in a tetrahedral coordination and vice versa, leading to an overall cubic symmetry. Iron impurities substitute for the cation at a site with symmetry properties described by the operations that leave a regular tetrahedron invariant, namely point-group T_d (this is similar to the case of the operations that leave the equilateral triangle invariant in two dimensions, as discussed above).

The four anions at the vertices of the tetrahedron span 12 degrees of freedom: three translations, three rotations, and six vibrations. One of the vibrational modes is the breathing mode. There are two degenerate modes designated by ε and three degenerate modes designated by τ_2 (irreducible representations of T_d).

Iron appears in this system as a doubly ionized atom, Fe^{2+}, with a d^6 electronic configuration (within the approximation of the ionic picture). Each of the multiplets couples differently to each of the three vibrational sets of modes. Couplings to the ground multiplet ^5D are of the most importance. A crystalline field splits this level into a lower ^5E and an upper $^5T_2$ multiplet. A spin–orbit interaction further splits these multiplets as shown schematically on the left-hand side of **Figure 5**. Downward (upward) arrows on the left-hand side of this figure illustrate allowed electric-dipole emissions (absorptions) when the JTE is ignored. However, experiments show a completely different picture making it necessary to include JTE. In principle, any of the six vibrational modes can be

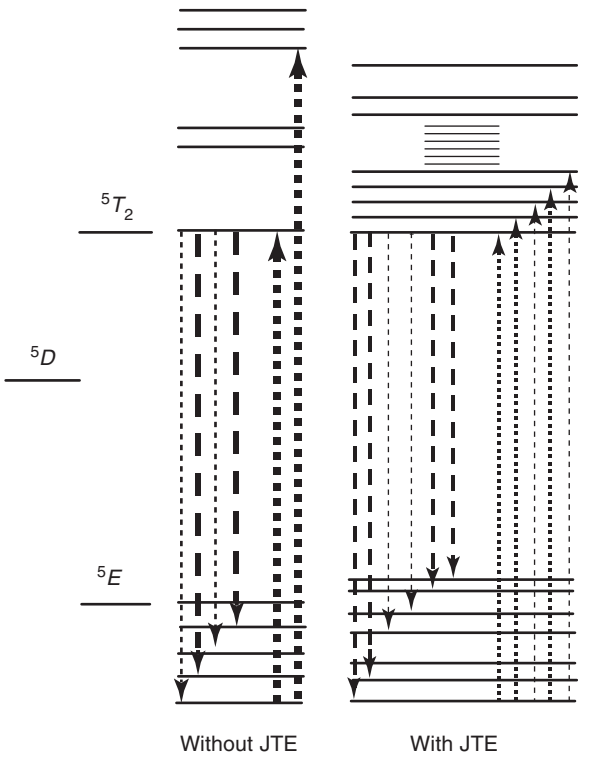


Figure 5 Consequences of the JTE on the emission and absorption spectra of Fe^{2+} in CdTe. The scheme on the left-hand side illustrates the case of no JTE, predicting few lines: four equally spaced lines in luminescence; two widely separated absorption lines. The scheme on the right-hand side shows actual emission and absorption lines after appropriate JTE is included: several unequally spaced lines in luminescence; several close lines near the absorption threshold.

responsible for this vibronic coupling. It turns out that ε modes give a good description of the main features of the emission and absorption spectra as schematically shown in the right-hand side of **Figure 5**, when appropriate values of the parameters are used. In the present case, the emphasis is on the rearrangement of energy levels corresponding to a dynamic JTE.

Emission Spectrum

Low-temperature emission lines originate in the lowest energy level of the T_2 multiplet ending in any compatible vibronic level of the E multiplet. Selection rules do not only take into account symmetry, but components must also have the same number of vibrational quanta, zero-phonon lines being the most important contributions. A good agreement with luminescent spectra is obtained after considering the vibronic coupling $E \times \varepsilon$ with the coupling phonon of $\leqslant \omega_E = 29\ \text{cm}^{-1}$ and $E_{\text{JTE}} = 3.2\ \text{cm}^{-1}$; this means a Huang–Rhys factor $S = 0.11$, evidencing a rather weak JTE. Then, this coupling can be treated as a minor perturbation, which is evident from the small deviation of vibronic levels as compared to purely electronic levels.

Absorption Spectrum

Low-temperature absorptions originate in the lowest level of the E multiplet ending in levels of the T_2 multiplet allowed by symmetry, matching components with the same vibrational quantum number. Vibronic couplings $T_2 \times \varepsilon$ and $T_2 \times \tau_2$ are possible, the former producing the energy level shown on the right-hand side of **Figure 5**, giving a good explanation of the positions and relative strengths of the absorption lines. Again, the main features of the experiment can be explained upon considering coupling to ε modes. Such modes have the same symmetry as the one considered for emission, but with a larger frequency. This mode is better suited to produce coupling to the more extended and differently oriented electronic orbitals of the upper multiplet. A good agreement with the absorption spectra is obtained for $\leqslant \omega_T = 35\ \text{cm}^{-1}$ and $E_{\text{JTT}} = 210\ \text{cm}^{-1}$; this means a Huang–Rhys factor $S = 6.0$ (~ 50 times larger than the previous one), is an important JTE, producing significant effects in the absorption spectrum. Calculations give many lines of similar intensity close to the absorption threshold (in perfect correspondence with experiments) instead of a single isolated line predicted by the plain crystal-field theory. This effect is shown in **Figure 5** by arrows pointing upward.

Cooperative JT Effect

This article so far has focused on isolated JT ions. When the crystal is formed by JT centers, local couplings can collectively combine in such a way that neighboring distortions cooperate to minimize the total energy. To visualize this effect, think of a local octahedron formed by the six anions surrounding the JT cation. If an octahedron elongates along one direction, the next octahedron along that same direction shrinks. The JT effect of each isolated ion does not need to be large, but the propagation of the phenomenon through the entire lattice makes such a cooperative JT effect (CJTE) a large effect, causing a long-range rearrangement of atoms.

One example of a system exhibiting CJTE is the perovskite $LaMnO_3$, whose ideal structure is cubic with the La at the center, one Mn at each of the eight corners and one O at the center of each of the 12 edges. The JT center (Mn ion) is surrounded by an octahedron of six O. However, the three axes of this octahedron are different due to the coupling of the Mn ions to crystal vibrations causing a JT distortion of the octahedron. The neighboring octahedra tilt in complementary ways and establish a CJTE that propagates through the crystal producing an orthorhombic lattice. This effective vibronic coupling reveals itself in other phenomena as well. Thus, the trivalent La can be substituted by bivalent ions (such as Ca), which turns some neighboring Mn ions magnetic. A vibronic coupling to these ions is important to explain the large increase of electrical resistivity of the system $La_{0.67}Ca_{0.33}MnO_3$ in the presence of a magnetic field in what is known as colossal magnetoresistance. Normally, the CJTE produces a low-symmetry structure, which appears below a certain temperature.

See also: Crystal Field Splitting; Crystal Symmetry; Group Theory; Insulators, Electronic States of; Insulators, Optical Properties of; Lattice Dynamics: Vibrational Modes; Optical Absorption and Reflectance; Point Groups; Semiconductor Compounds and Alloys, Electronic States of; Semiconductors, Impurity and Defect States in; Semiconductors, Optical and Infrared Properties of.

PACS: 31.30.Gs; 71.70.−d; 71.70.Ej

Further Reading

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Bates CA (1978) Jahn–Teller effects in paramagnetic crystals. *Physics Reports* 35: 187.

Bersuker IB (2005) *The Jahn–Teller Effect*. Cambridge: Cambridge University Press.

Bersuker IB and Polinger VZ (1989) *Vibronic Interactions in Molecules and Crystals,* Springer Series in Chemical Physics, vol. 49. New York: Springer.

Bevilacqua G, Martinelli L, and Terzi N (eds.) (1999) *Electron–Phonon Dynamics and Jahn–Teller Effect*. Singapore: World Scientific.

Grosso G and Pastori Parravicini G (2000) *Solid State Physics*. Cambridge: Academic Press.

Kaplan MD and Vekhter BG (1995) *Cooperative Phenomena in Jahn–Teller Crystals*. New York: Plenum.

Kaplan MD and Zimmerman GO (eds.) (2001) *Vibronic Interactions: Jahn–Teller Effect in Crystals and Molecules*, Nato Science Series II: Mathematics, Physics and Chemistry, vol. 39. Dordrecht: Kluwer Academic.

Schulz H-J, Scherz U, and Thurian P (eds.) (1997) *Electrons and Vibrations in Solids and Finite Systems (Jahn–Teller Effect)*. München: Verlag GmbH.

K

Kondo Effect (Theory)

B Coqblin, Université Paris-Sud, Orsay, France

Introduction to the Kondo Effect

The starting point of the long history of the "Kondo effect" is the first calculation performed by Kondo himself in 1964 to account for the observed experimental evidence of a resistivity minimum in magnetic dilute alloys such as CuMn, CuFe, YCe or LaCe. The first observed resistivity minimum in CuFe was dated from 1931.

Before discussing the Kondo effect itself in greater detail, it is important to note that the Kondo effect belongs to the wider field of the presently so-called "physics of strongly correlated electron systems," which originates primarily from the presence of an inner *d* or *f* shell in one of the atoms embedded in the considered system. The *d* electrons in transition elements and the 4*f* or 5*f* electrons in rare-earths or actinides are generally well localized and correlated. The correlations tend first to favor the existence of magnetism which is due to open inner *d* or *f* shells, as observed in pure iron, cobalt, or nickel in the 3*d* series, in rare-earth metals and in actinide metals after americium. Magnetism is also observed in many ionic, insulating, or metallic systems containing the preceding magnetic elements. But in fact, at present, the generic term of strongly correlated electron systems is mostly confined to metallic systems where there exists a strong interaction or hybridization between the inner *d* or *f* electrons and the conduction electrons. From a theoretical point of view, this type of physics started with the pioneering works of Mott for metal–insulator transition, of Friedel for the so-called "Friedel sum rule" and the theoretical concept of "virtual bound state" in 1956, of Anderson for the derivation of the hybridization "Anderson Hamiltonian" in 1961, and obviously of Kondo himself for the explanation of the resistivity minimum.

Thus, a summary of the main works on the Kondo effect is presented here, starting with the derivation of the calculation of the Kondo effect itself and the so-called "high-temperature" regime, that is, when the perturbation theory is valid; then, the "low-temperature" regime for one impurity and finally, the problems of the Kondo lattice and of the multichannel Kondo effect are described.

The "High-Temperature" Case

The first calculation of Kondo was performed in dilute alloys with magnetic transition impurities and the system is well described by the classical exchange Hamiltonian given by

$$H = -J\boldsymbol{s}_c\boldsymbol{S}_f \qquad [1]$$

In the expression [1], the localized spin $\boldsymbol{S}_f$ on the impurity interacts with the spins $\boldsymbol{s}_c$ of the conduction electrons coming from the host. The exchange constant J can be either positive (corresponding to a ferromagnetic coupling) or negative (antiferromagnetic coupling). It results that a localized spin $\boldsymbol{S}_f$ on site i interacts with another localized spin on site j by an indirect mechanism involving the conduction electrons through the exchange interaction given by [1]: this indirect interaction between two localized electron spins is called the "Rudermann–Kittel–Kasuya–Yosida" (or RKKY) interaction, which decreases as a function of the distance R between the two sites i and j as $\cos(k_F R)/(k_F R)^3$.

In 1964, Kondo computed the magnetic resistivity ρ_m up to the third order in J (or in the second-order perturbation) and found

$$\rho_m = A|J|^2 S(S+1)\left[1 + J\rho \log\frac{k_B T}{D}\right] \qquad [2]$$

where ρ is the density of states for the conduction band at the Fermi energy, A a constant value, and S the value of the spin for the localized electrons. It is immediately seen that for a negative J value, the magnetic resistivity decreases with increasing temperature. Thus, the total resistivity, which is the sum of the magnetic resistivity and the phonon one which increases with temperature, passes through a minimum. A negative J value corresponds to an antiparallel coupling between the localized and conduction electron spins, and the preceding treatment is valid

only at sufficiently high temperatures above the so-called "Kondo temperature" T_k given by

$$k_B T_k = D \exp\left(-\frac{1}{|J\rho|}\right) \quad [3]$$

where D is a cutoff corresponding roughly to the conduction bandwidth. The magnetic susceptibility decreases above T_k as $1/T$, as in a Curie law. The different transport, magnetic, and thermodynamic properties of a Kondo impurity have been extensively studied by perturbation in the so-called "high-temperature" regime (i.e., above T_k). On the other hand it is clear that, below T_k, the perturbative treatment is not valid and the low-temperature behavior was completely different from the previously described high-temperature one.

The exchange Hamiltonian [1] describes well the case of a magnetic impurity embedded in a normal matrix, when the inner d or f electrons are well localized corresponding to approximately integer values of $2S$. On the other hand, Anderson has introduced a Hamiltonian which can describe different situations, such as the nonmagnetic or the magnetic case, as well as the case of an intermediate valence or that of an almost integer valence. In the series of rare-earths, most of them have a valence of 3 and are well described by the Hamiltonian [1] with an integer $2S$ value, while systems containing anomalous rare-earths, like cerium or ytterbium, are better described by the Anderson Hamiltonian. Moreover, in the special case of a magnetic impurity with a valence close to an integer, Schrieffer and Wolff have shown that the Anderson Hamiltonian is equivalent to a Kondo Hamiltonian [1] with a negative J value given by

$$J = -\frac{|V_{kf}|^2}{|E_0 - E_F|} \quad [4]$$

where $E_0 - E_F$ is the distance between the $4f$ and the Fermi levels. Thus, this so-called Schrieffer–Wolff transformation allows one to show that the Anderson Hamiltonian in the almost ionic limit yields automatically a negative J value and, therefore, the Kondo effect.

The Anderson and the exchange Hamiltonians have been used first to describe the situation of magnetic systems with only a spin magnetic moment and a quenched orbital magnetic moment, as in the case of alloys with $3d$ transition elements. However, the physical situation of rare-earths is more complicated, because the magnetic moment has both spin and orbital components and the states are labeled according to the z-components of the total angular momentum $J = L + S$. For example, in magnetic cerium systems, the spin–orbit is sufficiently large to split the $4f^1$ configuration into a ground state $j = 5/2$ and an excited state $j = 7/2$. Then, the sixfold degenerate level $j = 5/2$ is split by the crystalline field effect which has a typical energy splitting $\sim 100\,\mathrm{K}$ and, therefore, of the same order as the usual temperature of experiment. Thus, both the orbital degeneracy and the crystalline field effect are important in the study of anomalous magnetic rare-earths such as cerium or ytterbium with $4f^1$ or $4f^{13}$ configurations. The Anderson Hamiltonian including the orbital degeneracy has been used first to explain the phase diagram of cerium metal, which provides a very fascinating example of many features of strong correlations. However, in the studied case of the so-called "Kondo limit," that is, when the number of $4f$ electrons (or holes) is close to 1, corresponding to the $4f^1$ configuration of cerium (or the $4f^{13}$ configuration of ytterbium), an effective Hamiltonian has been derived by use of the Schrieffer–Wolff transformation. Thus, the resulting so-called "Coqblin–Schrieffer" Hamiltonian including crystalline field effects has been extensively used to derive many transport and magnetic properties of cerium Kondo compounds. The magnetic resistivity, calculated by the perturbation theory at "high temperatures," that is, for $T > T_k$, presents a maximum at a temperature roughly equal to Δ, and two $\log T$ behaviors above and below Δ for the typical case of a cubic crystalline field of energy splitting Δ. The slopes of these $\log T$ behaviors are characteristic of the respective degeneracies of the levels, and these two $\log T$ behaviors separated by the maximum corresponding to the crystalline field were observed for the first time in the plot of the magnetic resistivity ρ_m of the $CeAl_2$ compound in 1972. A large effect of pressure has also been observed, because the absolute value of the $\log T$ slopes increases rapidly with pressure, which is directly accounted for by the theoretical model. These effects have been then observed in the other transport and magnetic properties and have been found in many cerium compounds.

Thus, the decrease of the magnetic resistivity in $\log T$ is now considered a clear "signature" of the Kondo effect in magnetic cerium compounds or in some related magnetic rare-earth systems, and has been in fact observed in more than a hundred compounds. The increase in the absolute value of the $\log T$ slope is also characteristic of the pressure effect in cerium Kondo compounds. Moreover, at present, the available highest pressures are much larger and there exist very nice new experiments on such cerium or ytterbium compounds.

The "Low-Temperature" Case

It has been seen that the perturbative calculation of Kondo is no more valid at low temperatures, that is, below the "Kondo temperature" given by eqn [3]. After several attempts, the theoretical problem of a single Kondo impurity was exactly solved in 1975 at $T = 0$ by Wilson using the renormalization group technique on the exchange Hamiltonian [1] with a negative J value. Wilson was the first to show numerically that the exact $T = 0$ solution is a completely compensated singlet formed by the antiparallel coupling between the conduction electron spin $s_c = 1/2$ and the localized electron spin $S_f = 1/2$. An analytical exact solution of the single Kondo impurity problem was then obtained by the Bethe–Ansatz method for both the classical exchange and the Coqblin–Schrieffer Hamiltonians. The exact solutions have given a clear description of the Kondo effect, but since it is difficult to use them for computing the different physical properties, sophisticated "mean field" methods have been introduced, such as the "slave boson" method or the method using correlators or more recently, the dynamical mean field theory. Many works may be mentioned to describe the dynamical effects or spectroscopic properties, and both the one-impurity Kondo and Anderson problems including dynamical properties have been solved by different theoretical techniques and, in particular, by the renormalization group technique.

The low-temperature (i.e., $T \ll T_k$) properties of a Kondo impurity are those of a "Fermi liquid" and can be summarized as follows:

1. The electronic specific heat constant γ and the magnetic susceptibility χ are much larger than those corresponding to a normal magnetic behavior. The first experimentally observed example for such a behavior corresponds to $CeAl_3$, where an extremely large electronic specific heat constant $\gamma = 1.62\,J\,mol^{-1}\,K^2$ was found in 1975 by Andres and co-workers. The term "Heavy fermions" originates from these high values of γ and χ.
2. The transport properties of cerium Kondo compounds are described by the Fermi liquid formalism at low temperatures. In particular, the electrical resistivity follows a T^2 law and the coefficient A of this term is particularly large in heavy fermion systems.
3. A direct evidence of the heavy fermion behavior has been obtained by de Haas–van Alphen effect experiments in many cerium compounds, such as $CeCu_6$, CeB_6, $CeRu_2Si_2$, and some uranium compounds. An effective mass up to $\sim 100\,m_e$ (where m_e is the electron mass) has been observed in cerium compounds.
4. It is finally important to notice that superconductivity was observed first in $CeCu_2Si_2$ compound by Steglich and co-workers in 1979, then in some cerium compounds at high pressure, and uranium compounds at normal pressure, which is clearly an outstanding experimental result and a fantastic theoretical challenge.

The Doniach Diagram and the Kondo Lattice Problem

Up to now, only the Kondo effect model for a single impurity has been studied and remains valid even for concentrated alloys or compounds at high temperatures above the Kondo temperature. However, when one deals with, for example, a lattice of cerium atoms, the situation becomes more difficult at very low temperatures. The competition between the Kondo effect on each atom which tends to suppress the magnetic moment with decreasing temperature and the RKKY interaction which, on the contrary, tends to give a magnetic ordering between different rare-earth atoms is well described by the well-known "Doniach diagram," where the Kondo temperature for a single impurity, and the Neel (or Curie in a few cases) temperature as a function of the exchange interaction term, $|J\rho|$ are plotted. It results in the real ordering temperature T_N increasing initially with increasing $|J\rho|$, then passes through a maximum and tends to zero at a critical value $|J\rho|_c$ corresponding to the so-called "quantum critical point" (QCP).

Such a behavior of T_N has been experimentally observed with increasing pressure in several cerium Kondo compounds, such as $CeAl_2$, CeAg, or $CeRh_2Si_2$. A maximum of the ordering temperature has been observed in some of them, for instance, CeAg, which is ferromagnetic, or $CePd_2Al_3$, but in all these compounds, the Neel temperature decreases and tends to zero at high pressures.

Above the QCP corresponding to the magnetic–nonmagnetic instability (or to the disappearance of the Neel temperature T_N), several different behaviors have been observed: a departure from the Fermi liquid (FL) behavior, called "non-Fermi liquid" (NFL), has been observed experimentally in many cerium compounds under high pressure, such as Ce_7Ni_3, which presents successively a magnetic ordering up to ~ 3 kbar, then an NFL behavior up to ~ 6–7 kbar, and finally an FL behavior above that pressure with a coefficient of the resistivity T^2 law decreasing rapidly with pressure and disappearing above 12 kbar. The NFL behavior of the Kondo lattice near the magnetic instability has been theoretically studied within the

scaling theory of critical phenomena and within the self-consistent renormalized theory of spin fluctuations. This last theory predicts a $\log T$ law for γ and a $T^{3/2}$ law for the magnetic resistivity ρ_m just at the QCP; when one goes farther from the QCP, the behavior tends back to an FL behavior, with a decrease of the $\log T$ slope of γ and a power law for ρ_m approaching the T^2 law.

In fact, the FL behavior exists above the QCP in some cerium compounds such as $CeRh_2Si_2$, and the real Kondo temperature has been found to remain roughly constant with $|J\rho|$ or with pressure in the nonmagnetic domain. A theoretical model, taking into account both the intra-site Kondo interaction and an additional inter-site Heisenberg-type interaction, can explain the observed decrease of the real Kondo temperature with respect to the one-impurity Kondo temperature, as a function of the increase of the intersite interaction and of the decrease of the number of conduction electrons; this last result corresponds to the idea of "exhaustion," which corresponds to a decrease of the Kondo effect when there are not enough conduction electrons to screen the localized electrons. The separation between the FL and the NFL behaviors above the QCP is not completely explained at present and is still an important subject of research.

There are, in fact, several different NFL behaviors, in contrast to the obviously unique FL behavior, and also several theoretical explanations for such experimentally observed NFL behaviors. In addition to the previously described NFL behavior directly connected to proximity to the QCP, a second observed NFL behavior is due to an effect of disorder. A simple model of a disordered Anderson lattice has been proposed and the disorder introduces a distribution of Kondo temperatures. This simple mechanism leads to logarithmic behaviors in thermodynamic properties and a linear temperature dependence of the resistivity, as experimentally observed in some cerium and uranium alloys. The third NFL behavior is due to the "multichannel Kondo effect," as explained in the next section.

The Multichannel Kondo Effect

An NFL behavior observed in cerium and uranium compounds arises also from the so-called "multichannel Kondo model." The regular Kondo effect corresponds to a complete screening of the localized electrons by the conduction electrons, while the multichannel Kondo effect corresponds to either an underscreening or an overscreening of the localized electrons by the conduction ones. The multichannel Kondo model for a single impurity has been completely solved analytically by the Bethe-ansatz method, and numerical solutions have been obtained to describe the different physical properties.

In the "underscreened" Kondo effect with $2S > n$ (where n is the number of channels for conduction electrons), a part $n/2$ of the impurity spin is compensated, while the part $S' = S - n/2$ is not compensated. The result is that the magnetic susceptibility is diverging in $1/T$ as for a Curie law, which is obviously different from the constant magnetic susceptibility obtained at low temperatures for the regular screened Kondo effect ($2S = n$). A ferromagnetic coupling between the resulting spin S' and the conduction electron spin is finally obtained in the underscreened Kondo model.

In the "overscreened" Kondo effect with $2S < n$, it remains $(n - 2S)$ uncoupled conduction electrons and the spin S' given by $S' = n/2 - S$ is antiparallel to S. For $S = 1/2$, $n = 2$, a logarithmic dependence of χ and γ is obtained, while for $S = 1/2$ and $n > 2$, the temperature dependence of χ and γ is given by a power law T^n. A particular model within this overcompensated case is the "quadrupolar Kondo effect," corresponding to the two-channel case: a logarithmic dependence for γ, a linear one for the magnetic resistivity, and an unusual value for the entropy (corresponding to half the value of the entropy for the regular Kondo effect) have been obtained and these theoretical results have been used to account for the NFL behavior observed in $Y_{1-x}U_xPd_3$ alloys.

Conclusions

As a present conclusion, it can be said that there have been, in the last decades, both a tremendous improvement of the experimental techniques used for the study of strongly correlated electron systems and a fantastic development of the theoretical ideas in the field of strongly correlated electron systems since the pioneering works of Mott, Friedel, Anderson, and Kondo. The Kondo effect has been understood for the case of a single impurity, while several features of the Kondo lattice problem are not clearly understood and remain interesting challenges. Several new subjects are being clearly studied at present: first, many cerium or uranium compounds become superconducting at low temperatures, generally close to the QCP and the occurrence of superconductivity in such systems, as well as the high-T_c superconductivity observed in cuprates or related compounds, remains a not-understood challenge; second, the Kondo effect has been observed for a small set of atoms in nanoscopic systems and this field is clearly a rapidly developing subject; third, a lot of new compounds have been obtained in recent years and many of these

systems present new and fascinating properties connected with magnetism and superconductivity.

See also: Liquids, Theory of: Fermi Liquids; Magnetism, History of; Magnetic Interactions.

PACS: 71.27.+a; 72.15.Qm; 75.20.Hr

Further Reading

Bickers NE (1987) Review of techniques in the large-N expansion for dilute magnetic alloys. *Reviews of Modern Physics* 59: 845–939.

Conference Proceedings of SCES: *Physica B*, volumes 230–232 (1997), volumes 259–261 (1999) volumes 281–282 (2000), volumes 312–313 (2002) and *Acta Physica Polonica B* volume 34 (2003).

Continentino MA (2001) *Quantum Scaling in Many-Body Systems*. Singapore: World Scientific.

Coqblin B (1977) *The Electronic Structure of Rare-Earth Metals and Alloys: The Magnetic Heavy Rare-Earths*. London: Academic Press.

Coqblin B (2001) Electron systems: strong correlations. In: *Encyclopedia of Materials: Science and Technology*, pp. 2591–2603. Elsevier Science Ltd.

Coqblin B, Arispe J, Bhattacharjee AK, and Evans SMM (1993) Kondo effect and heavy fermions. In: *Frontiers in Solid State Sciences, Vol. 2: Selected Topics in Magnetism*, pp. 75–105.

Friedel J (1958) Metallic alloys. *Nuovo Cimento* VII: 287–311.

Hewson AC (1992) *The Kondo Problem to Heavy Fermions*. Cambridge: Cambridge University Press.

Kondo J (1964) Resistance minimum in dilute magnetic alloys. *Progress of Theoretical Physics* 32: 37–49.

Mott NF (1974) *Metal–Insulator Transitions*. London: Taylor and Francis.

Onuki Y and Hasegawa A (1995) In: Gschneidner KA, Jr. and Eyring L (eds.) *Handbook on the Physics and Chemistry of Rare-Earths*, vol. 20, pp. 1–103. Amsterdam: Elsevier Science BV.

Wilson KG (1975) The renormalization group: critical phenomena and the Kondo problem. *Reviews of Modern Physics* 47: 773–840.

L

Landau Levels *See* Magnetoresistance Techniques Applied to Fermi Surfaces.

Laser Radiation Sources

A Agnesi, G Reali, and A Tomaselli, Università degli Studi di Pavia, Pavia, Italy

Introduction

LASER is an acronym meaning light amplification by stimulated emission of radiation. Stimulated emission can be obtained from a medium prepared in a nonthermal distribution of its energy-level populations. In a solid-state material, for example, this is achieved through the absorption of light at an appropriate pumping wavelength. In the following, for simplicity, the laser medium will mean a solid-state matrix, contributing a background refractive index η, enclosing atoms which under pumping will be all set in the same final state, so that a homogeneous lineshape of their laser transition is also understood. For the same reason, level degeneracies will not be considered. Since Nd:YAG is a widely used solid-state laser material, numerical examples will refer to it in the following. A laser can be either an amplifier or a generator of coherent light, in which case an optical feedback is provided in order to allow the device to generate the laser beam from its own internally generated optical seed.

When a laser system is properly designed and operated, it is able to generate light of extraordinary quality, in terms of spectral purity, directionality, and power. The properties of the generated beam are precisely the reason of the importance of laser devices in scientific, biomedical, and industrial applications.

The basic interactions involved in laser operation proceed via the radiative processes of (induced) absorption, spontaneous emission, and stimulated (or induced) emission.

Spontaneous emission needs a quantum treatment. In addition to providing the seed into the right cavity mode that can start the laser action, it represents an incoherent and completely uncorrelated contribution with respect to the coherent stimulated terms. Ultimately, it is also responsible for the phase and amplitude fluctuations of laser systems.

For the induced processes instead, it is possible to resort to much simpler balance equations, first introduced by Einstein, that describe how light is transported through the system. The basic idea of transport or rate equations approach is to add or subtract the quasiresonant photon energies $\hbar\omega$ from the field as it travels through the atomic population, of densities N_2, N_1 that make transitions between levels separated by $\hbar\omega_0 = (E_2 - E_1)$. This is schematically shown in **Figure 1**.

Accordingly, the incident optical intensity S (energy per unit area per unit time), that sweeps across a volume of unit cross-sectional area and thickness dz, containing N_2 dz and N_1 dz atoms in upper and lower levels interested by the interaction, will be changed as

$$\frac{1}{c}\frac{\partial S}{\partial t} + \frac{\partial S}{\partial z} = \sigma(N_2 - N_1)S \qquad [1]$$

where σ (in units of length^2) is the transition cross section, equal for both the events. The overall change in [1] will be according to the sign of $(N_2 - N_1)$: for positive sign, the medium behaves as an amplifier of light, while in the opposite case of negative population inversion, under ordinary thermal equilibrium conditions, only absorption can be realized and the optical intensity decreases in going across the medium.

In recognition of the amplifier condition, it is usual to define the gain coefficient (in units of length^{-1}) as

$$g = \sigma(N_2 - N_1) \qquad [2]$$

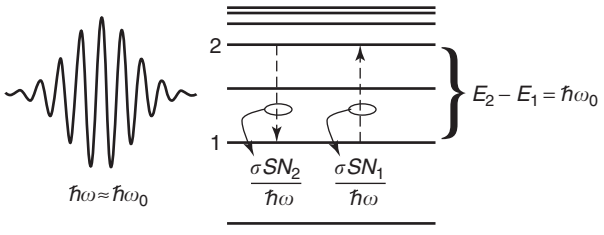


Figure 1 Schematic of the interaction of two atomic levels with a quasiresonant photon flux for an allowed transition.

so that, considering a steady-state ($\partial/\partial t = 0$) situation in [1],

$$\frac{dS}{dz} = gS \qquad [3]$$

In the case of constant gain coefficient, eqn [3] is easily solved, giving an exponential optical intensity growth

$$S(z) = e^{gz}S(0) \qquad [4]$$

An amplifier medium is used in laser oscillators, which are the most important devices. In their simplest configuration, they are set up enclosing the gain medium in a linear cavity formed by two (plane or curved) mirrors as shown in **Figure 2**, of which one, of partial reflectivity R, provides the useful output coupling.

The purpose is to trap only the almost axially emitted part of spontaneous emission and redirect it along the axis, so that it can be amplified and filtered through the combined action of the stimulated emission in the gain medium and the selective modal structure of the optical cavity. These iterated actions build up the self-sustained intense, highly monochromatic, and directional laser beam. A simple gain–loss balance analysis of the mode in the optical cavity (or resonator) immediately provides the threshold lasing condition. Following the history of gain and loss of the oscillating mode along a full round trip and requiring self-reproducibility at the end,

$$e^{2g\ell}R(1 - L_i) = 1 \qquad [5]$$

where ℓ is the gain medium length, R is the output coupling mirror reflectivity, and L_i are the double-pass passive cavity loss.

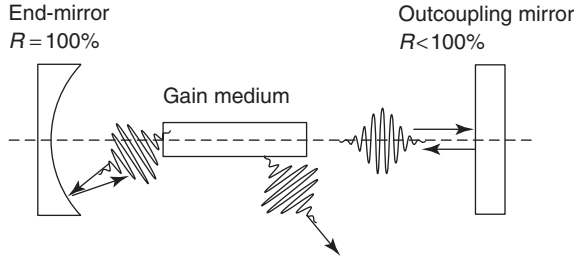


Figure 2 Optical laser cavity, with the pictorial description of its operation. Only almost axially emitted photons are trapped inside it.

This defines the threshold gain for the laser oscillator:

$$g_{th} = -\frac{\ln(R) + \ln(1 - L_i)}{2\ell} \qquad [6]$$

above which stimulated laser action initiates.

A second important condition constraints the change of phase of the oscillating mode after a round trip, and is presented later in this article.

Creating Population Inversion and Laser Beam Properties

In order to create a population inversion, the closed two-level system illustrated in **Figure 1** cannot work "as is," since the nature of up–down stimulated transitions and their relation (same cross section) prevents the possibility of using the same channel for both excitation and generation, in case of optical pumping.

The minimum number of required levels is thus three, and a more general four-level system may perform even better, depending on the relative magnitudes of the relaxation rates of the level populations and on the selective actions of the pumping system. **Figure 3** clarifies the kinematics of the level populations in a four-level system.

Typically, a four-level system is one in which upper laser level 2 is fed at a rate that almost coincides with the rate at which level 3 is fed, due to a very fast, nonradiative relaxation from level 3 to level 2; in addition, the lower laser level 1 is constantly drained off, keeping it essentially empty, by its fast decay to lower lying levels.

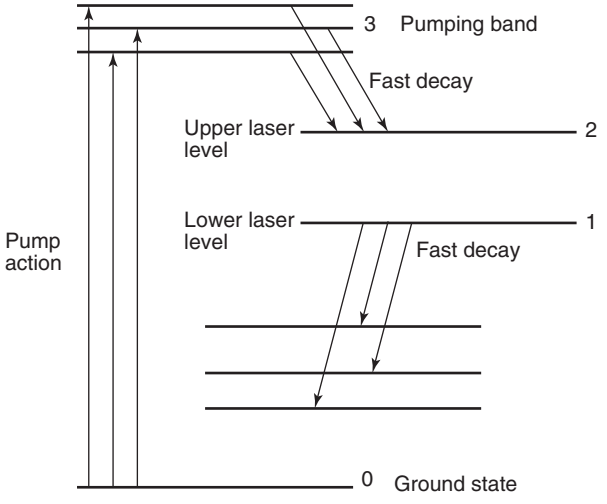


Figure 3 A "four-level" laser, with the indication of the most relevant energy exchanges: the pump, and the fast decays into the upper laser level and out of the lower laser level, which favor an efficient creation of the population inversion.

Pumping efficiency and resonator design are the most important problems in laser engineering, since the properties of the generated laser beam depend on their optimization.

Once the active medium has been chosen, and the application the laser is aimed at has suggested the operating regime (continuous wave, long pulses, short pulses, etc.), power, monochromaticity and spectral purity, coherence, and directionality are essentially controlled by pumping and by the optical resonator. An overall parameter, which can be used, for example, to compare laser light and standard generated light (sun, lamp, discharge, etc.), is the brightness which summarizes all the above properties.

As a rule, laser emission is not at a single frequency near the ω_0 of the laser transition. Because of natural lifetime and interactions of the active atoms with the surrounding environment, gain effects, and photon lifetime, there is a sizable spectral broadening of the emitted light. Still, laser light is far better than all other light sources. As an example, Nd:YAG, one of the most widely used solid-state laser materials, lases at 1064 nm, corresponding to $\omega_0 \approx 1.8 \times 10^{15}$ s^{-1}, and may have spectral purity $\Delta\omega/\omega_0 \approx 10^{-9}$, several order of magnitudes smaller than any unfiltered thermal light source, for which that figure approaches 1.

Since the stimulated emission forces all photons in exactly the same quantum state, coherence persists until the radiating atomic polarization remains temporally and spatially correlated, and is not disturbed by any spurious effects such as collisions with the environment. Collisions essentially determine the coherence time, which is connected to the degree of monochromaticity of the emitted light, $\tau_{coh} \sim 2\pi/\Delta\omega$.

To a large extent, spatial coherence is formed by the transverse modal structure of the resonator, from which the beam derives its directionality. Spatial coherence and directionality are the properties that allow the beam to be brought, with very low spill-out energy, to a tight focus, an essential requirement in most laser applications. Diffraction theory shows that the lowest order, Hermite–Gauss mode of a stable resonator diverges with an angle $\vartheta_{0x} = \lambda/(\pi w_{0x})$, where w_{0x} is the beam waist (its lowest radius) in an x-direction perpendicular to z-propagation direction; the same expression holds for the y-direction. This is the lowest divergence among beams with same beam waist. It can be shown that higher-order modes (and, more generally, a real beam made up of a mixture of these) with same waist diverge with an angle $\Theta_x = M_x^2 \vartheta_{0x}$, greater than the lowest-order mode by a factor of M_x^2, called the spatial beam quality factor; this expression can be used as an operative definition of the M_x^2 factor, which is always >1 for any real beam (a same expression holds for the y-direction).

High intensities and low divergence can be put together to define laser brightness, the power emitted per unit area per unit solid angle. The concept applies, in general, to any source of radiation; in case of lasers, a very simple form is assumed as

$$B \equiv \frac{P}{A \times \Omega} \approx \frac{P}{(w_{0x} w_{0y}) \times (\Theta_x \Theta_y)} \approx \frac{P}{\lambda^2 M_x^2 M_y^2} \qquad [7]$$

and can be used as an indicator of the laser quality.

The fact that laser brightness is directly proportional to power is one of the reasons for research in the area of short and ultrashort pulse generation; another reason is that very short pulses allow one to investigate very short phenomena, in the same way as very small spots allow one to see and work out very small details.

Interaction of Light with Matter

In order to understand the mechanisms that govern the generation of laser light, it is useful to analyze in some detail the results of the quantum treatment of light-matter interaction, using a simple two-level system as toy model.

An incident optical field $\boldsymbol{E}(\boldsymbol{r}, t)$ on a quantum system, possessing a set of stationary levels and eigenfunctions, $\{E_i, \phi_i(\boldsymbol{r})\}$, is assumed to be almost resonant with two levels 2 and 1 with energy separation $\hbar\omega_0 = (E_2 - E_1)$, so that all the other levels can be ignored. Levels 2 and 1 are also assumed to have a well-defined parity, and to be connected by a dipole-allowed transition. The time-dependent interaction is described by the second term of the Hamiltonian $H = H_0 + H_1$, where, in the dipole approximation, $H_1 = -\boldsymbol{p} \cdot \boldsymbol{E}$ and the optical field takes the form $\boldsymbol{E} = \hat{\varepsilon} E_0(t) \cos(\omega t)$, with a slowly time varying amplitude. With usual procedure, the solution is assembled in the form $\psi(\boldsymbol{r}, t) = c_1(t)\phi_1(\boldsymbol{r}) + c_2(t)\exp(-i\omega t)\phi_2(\boldsymbol{r})$. Defining the dipole matrix elements $\boldsymbol{p}_{ij} = \int_{\text{all space}} d^3 r \phi_i^*(\boldsymbol{r}) \boldsymbol{p} \phi_j(\boldsymbol{r})$, neglecting rapidly varying terms, introducing the detuning $\Delta = (\omega_0 - \omega)$, and the Rabi frequency $\Omega = (\boldsymbol{p}_{21} \cdot \hat{\varepsilon}) E_0/\hbar$, the amplitude equations result

$$\begin{aligned} i \frac{dc_1}{dt} &= -\frac{1}{2} \Omega^* c_2 \\ i \frac{dc_2}{dt} &= \Delta c_2 - \frac{1}{2} \Omega^* c_1 \end{aligned} \qquad [8]$$

These are used to form the set of equations obeyed by the density matrix

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} = \begin{bmatrix} c_1 c_1^* & c_1 c_2^* \\ c_2 c_1^* & c_2 c_2^* \end{bmatrix} \qquad [9]$$

with closure condition $\rho_{11} + \rho_{22} = |c_1|^2 + |c_2|^2 = 1$, which provides a most compact means to calculate quantum expectation values, $\langle A \rangle = \mathrm{tr}\{A\rho\}$. In their final form, the equations are

$$\begin{aligned} \frac{\mathrm{d}\rho_{11}}{\mathrm{d}t} &= -\frac{\mathrm{i}}{2}(\Omega\rho_{12} - \Omega^*\rho_{21}) \\ \frac{\mathrm{d}\rho_{22}}{\mathrm{d}t} &= -\frac{\mathrm{d}\rho_{11}}{\mathrm{d}t} \\ \frac{\mathrm{d}\rho_{12}}{\mathrm{d}t} &= \mathrm{i}\Delta\rho_{12} + \frac{\mathrm{i}}{2}\Omega(\rho_{22} - \rho_{11}) \end{aligned} \qquad [10]$$

These equations describe the evolution of the atomic densities and of the induced polarization of a material, introduced through the macroscopic variables:

$$\begin{aligned} N_1 &= N\rho_{11} \\ N_2 &= N\rho_{22} \\ \boldsymbol{P} &= N\int \mathrm{d}^3 r \psi^*(\boldsymbol{r},t)\boldsymbol{p}\psi(\boldsymbol{r},t) \\ &= N(\rho_{21}\boldsymbol{p}_{12}\exp(-\mathrm{i}\omega t) + \mathrm{c.c.}) \end{aligned} \qquad [11]$$

with N the total population density participating in the interaction, and where $\boldsymbol{p}_{ii} = 0$ has been used. Introducing the slowly varying macroscopic polarization, $\boldsymbol{P}(t) = 1/2[\hat{\varepsilon}P(t)\exp(-\mathrm{i}\omega t) + \mathrm{c.c.}]$, and comparing it with eqn [11],

$$P = 2N(\boldsymbol{p}_{12}\cdot\hat{\varepsilon})\rho_{21} \qquad [12]$$

is obtained. The macroscopic equations now result

$$\begin{aligned} \frac{\mathrm{d}N_1}{\mathrm{d}t} &= -\frac{1}{2\hbar}\Im m(E_0^* P) = -\frac{\mathrm{d}N_2}{\mathrm{d}t} \\ \frac{\mathrm{d}P}{\mathrm{d}t} &= -\mathrm{i}\Delta P - \frac{\mathrm{i}|\boldsymbol{p}_{12}\cdot\hat{\varepsilon}|^2}{\hbar}(N_2 - N_1)E_0 \end{aligned} \qquad [13]$$

These equations describe the dynamics of an ensemble of "closed" two-level systems. In this form, they do not yet reflect some very important features that extend the two-level system into a working model of laser, such as the existence of relaxation processes through collisions or radiative decays, and especially the possibility of pumping energy into the system to get population inversion, $(N_2 - N_1) > 0$. The inclusion of these physical effects is possible making a phenomenological fix of the eqn [13] based on physical considerations; this is illustrated in **Figure 4**.

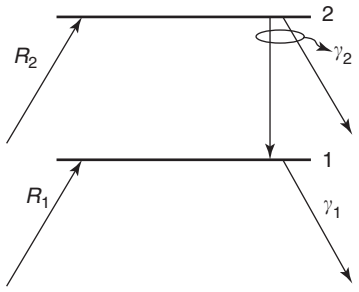


Figure 4 Pumping and decays that make the "two-level" model a "real" laser model.

The main effects to be included are: (1) pumping rates into levels 2 and 1, R_2 and R_1, respectively; (2) damping rates of levels 2 and 1 due to inelastic collisions and radiative decays, γ_2 and γ_1, respectively; and (3) elastic collisions, which only affect induced polarization, γ_{ph}. Taking all these effects into account leads to

$$\begin{aligned} \frac{\mathrm{d}N_1}{\mathrm{d}t} &= R_1 - \gamma_1 N_1 - \frac{1}{2\hbar}\Im m(E_0^* P) \\ \frac{\mathrm{d}N_2}{\mathrm{d}t} &= R_2 - \gamma_2 N_2 + \frac{1}{2\hbar}\Im m(E_0^* P) \\ \frac{\mathrm{d}P}{\mathrm{d}t} &= -(\gamma + \mathrm{i}\Delta)P - \frac{\mathrm{i}|\boldsymbol{p}_{12}\cdot\hat{\varepsilon}|^2}{\hbar}(N_2 - N_1)E_0 \end{aligned} \qquad [14]$$

where $\gamma = \gamma_{\mathrm{ph}} + (\gamma_1 + \gamma_2)/2$ is the damping rate for the induced polarization.

Rate-Equation Approximation and the Laser Amplifier

It happens quite often that in laser materials, γ is much larger than all the other rates, so that the induced polarization rapidly reaches a steady state and follows adiabatically the changes of the field and the atomic populations. In this "rate-equation approximation," $\mathrm{d}P/\mathrm{d}t = 0$ resulting in

$$P(t) = -\frac{\mathrm{i}|\boldsymbol{p}_{12}\cdot\hat{\varepsilon}|^2 E_0(t)}{\hbar(\gamma + \mathrm{i}\Delta)}(N_2 - N_1) \qquad [15]$$

This is substituted back into the population equations to give the rate equations:

$$\begin{aligned} \frac{\mathrm{d}N_1}{\mathrm{d}t} &= R_1 - \gamma_1 N_1 + W(N_2 - N_1) \\ \frac{\mathrm{d}N_2}{\mathrm{d}t} &= R_2 - \gamma_2 N_2 - W(N_2 - N_1) \\ W &= \frac{\pi|\boldsymbol{p}_{12}|^2|E_0(t)|^2}{6\hbar^2}\frac{\gamma/\pi}{\gamma^2 + \Delta^2} \end{aligned} \qquad [16]$$

A factor of 3, due to an orientational averaging performed over angles, has been included in W to

account for random local dipole disorientations with respect to the electric field:

$$\overline{|\boldsymbol{p}_{12}\cdot\hat{\varepsilon}|^2} = |\boldsymbol{p}_{12}|^2/3$$

It is recognized that the expression of W, which represents the rate at which the inversion is changed, can be written as

$$W = \frac{\sigma S}{\hbar\omega} \qquad [17]$$

using the photon flux

$$\frac{S}{\hbar\omega} = \frac{1}{\hbar\omega}\left(\frac{1}{2}\,\varepsilon_0\eta c|E_0(t)|^2\right)$$

where η is the average background refractive index of the material. From [17], the expression for the transition cross section is obtained:

$$\sigma = \frac{\pi|\boldsymbol{p}_{12}|^2\omega}{3\varepsilon_0\eta c\hbar}\,\frac{\gamma/\pi}{(\gamma^2+\Delta^2)} \qquad [18]$$

A quantum mechanical calculation of the spontaneous emission rate $A = 1/\tau_{sp}$ provides the explicit form

$$A = \frac{1}{\tau_{sp}} = \frac{\omega_0^3\eta|\boldsymbol{p}_{21}|^2}{3\pi\hbar\varepsilon_0 c^3} \qquad [19]$$

which, used in [18] to eliminate $|\boldsymbol{p}_{21}|^2$, with the consideration that $\gamma \ll \omega_0$, leads to the most popular form of the transition cross section

$$\begin{aligned}\sigma(\omega) &= \frac{\lambda_0^2}{4\eta^2\tau_{sp}}\,\frac{\gamma/\pi}{\gamma^2+(\omega-\omega_0)^2}\\ &= \frac{\lambda_0^2}{4\eta^2\tau_{sp}}L(\omega)\end{aligned} \qquad [20]$$

This expression can be used, for example, to get a numerical estimate of the cross section of Nd:YAG using available data $\tau_s = 230\ \mu s$, $\gamma = 2\pi \times 120$ GHz, $\eta = 1.82$. The computation gives $\sigma \approx 14 \times 10^{-19}$ cm^2, greater by a factor of ≈ 2 than its most quoted value, an acceptable result for this simple model.

$L(\omega)$ is the normalized Lorentzian lineshape, obtained in the hypothesis that all atoms behave in an identical manner in the interaction. Thus, the laser gain introduced in eqn [2] results

$$g(\omega) = \frac{\lambda_0^2}{4\eta^2\tau_{sp}}L(\omega)(N_2 - N_1) \qquad [21]$$

In more general situations, the lineshape has different widths and shapes, and can be largely inhomogeneous, if all the atoms do not have the same center transition frequency ω_0 (as, e.g., in Nd:glass).

From rate equations, it is immediately clear that the population inversion depends on optical intensity S, since W depends on it. In a strict four-level system, previously illustrated, in which pumping into lower laser level is suppressed, $R_1 = 0$, and the decay from the lower laser level is assumed to be very fast, these rate equations unveil this feature very clearly, when solved for the steady state. The result is

$$\begin{aligned}N_2 - N_1 &= \frac{(R_2/\gamma_2)}{1 + S/(\hbar\omega\gamma_2/\sigma)}\\ &= \frac{(N_2-N_1)_0}{1+S/S_{sat}}\end{aligned} \qquad [22]$$

where $(N_2 - N_1)_0 = R_2/\gamma_2$ is the inversion at low optical intensity, $S/S_{sat} \ll 1$, and $S_{sat} = (\hbar\omega\gamma_2/\sigma)$ is the saturation intensity. The gain thus becomes

$$\begin{aligned}g(\omega) &= \frac{\sigma(\omega)(N_2-N_1)_0}{1+S/S_{sat}}\\ &= \frac{g_0(\omega)}{1+S/S_{sat}}\end{aligned} \qquad [23]$$

where the small signal gain g_0 has been introduced with an obvious identification.

An immediate application of [23] is its use into [3] to obtain the general behavior of a single-mode amplifier. The equation

$$\frac{dS}{dz} = \frac{g_0}{1+S/S_{sat}}\,S \qquad [24]$$

has a general solution of the following form:

$$\frac{S(z)-S(0)}{S_{sat}} + \ln\left(\frac{S(z)}{S(0)}\right) = g_0 z \qquad [25]$$

where $S(0)$ is the input intensity. The exponential growth [4] is thus obtained only for small intensities compared with the saturation intensity. At the opposite limit, when the first term on the left-hand side of [25] dominates, the intensity difference grows linearly with the length as $g_0 S_{sat} z$. If multiplied by the effective area A_{eff} of the beam, and noting that $P_{ex} = g_0 S_{sat} = R_2\hbar\omega$ is the available power density from the gain medium, $g_0 S_{sat} A_{eff} z$ is seen to be the maximum extractable power from the amplifier, which is only attained working in saturation regime. In **Figure 5**, the plots of the "growth rate" $= (1/S)(dS/dz)$ and the "power extraction rate" $= dS/dz$ show that the growth rate is maximum for small S and drops to zero for large S, while an exactly specular behavior is followed by the power extraction.

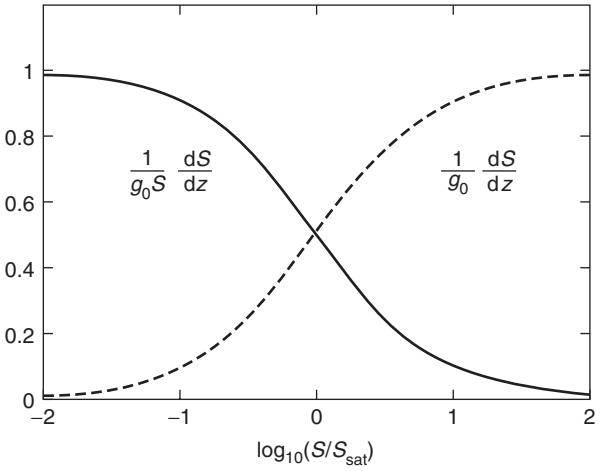


Figure 5 Growth rate and power extraction rate plots as a function of intensity normalized to the saturation intensity.

Finally, it is useful to summarize these considerations into the definition of a figure of merit for the amplifier; its internal efficiency

$$\eta_{\text{int}} = \frac{S(z) - S(0)}{R_2 \hbar\omega_{\text{p}} z} = \frac{S(z) - S(0)}{R_2 \hbar\omega z}\,\frac{\hbar\omega}{\hbar\omega_{\text{p}}} = \frac{S(z) - S(0)}{g_0 S_{\text{sat}} z}\,\eta_{\text{S}} \qquad [26]$$

where $P_{\text{p}} = R_2 \hbar\omega_{\text{p}}$ is the pump power density available in the amplifier, and $\eta_{\text{S}} = \omega/\omega_{\text{p}}$ is the Stokes efficiency.

Laser Resonators and Beam Modes

The importance and the advantages of making high-quality beams are so overwhelming that laser research has made enormous effort in studying optical resonators, which are the devices used to provide the oscillator feedback and to spatially shape the laser beams.

In a laser, the cavity is set up of (flat or curved) mirrors, surrounding the gain medium, which select only the light emitted approximately along the resonator axis, and reflect it back and forth several times, thus providing the necessary feedback that sustains the laser action. What characterizes a laser from other coherent devices at longer wavelengths (in the far-infrared and microwave ranges) is that an open resonator is needed to limit the number of possible oscillating modes that fit under the gain line width of the laser medium, and to get temporally coherent radiation.

In addition to the feedback action, the most important function of the resonator is the transverse spatial shaping of the radiated field, and the optical

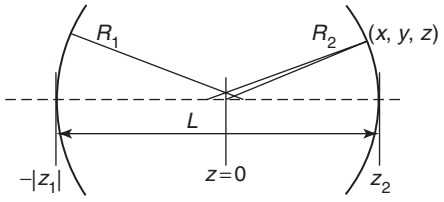


Figure 6 Empty laser resonator used to discuss the modes of a stable cavity laser. R_1 and R_2 are the curvature radii of the two confining mirrors, one of which eventually has a partial reflectivity to allow outcoupling of the laser field.

spatial quality of a beam is generated by using the modal structure of the resonator.

Optical resonators can be grouped into two classes: stable resonators are characterized by having closed ray paths, whereas unstable resonators have no closed path, and are mainly used in high gain laser systems. Only stable resonators are briefly described, while the book by Siegman provides a complete authoritative review on the subject of resonators.

The empty linear resonator shown in **Figure 6** is used to illustrate the sustained resonator modes. With this geometry, it is possible to select the fundamental TEM_{00} mode because the higher-order modes suffer significantly greater losses.

The structure of these modes can be found by solving the wave equation, which governs the propagation of beams whose transverse (i.e., (x, y)) dependence is less rapid than their z-dependence. Recognizing these different spatial rates, the complex quasimonochromatic field is written as

$$E(x, y, z, t) = E_0(t) u(x, y, z) \mathrm{e}^{\mathrm{i}(kz - \omega t)} \qquad [27]$$

where the spatial shape $u(x, y, z)$ is a slowly varying function in its variables with respect to $\mathrm{e}^{-\mathrm{i}kz}$, and satisfies the following parabolic wave equation (also known in optics as paraxial wave equation):

$$\frac{\partial u}{\partial z} + \mathrm{i}2k\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 0 \qquad [28]$$

It can be shown that, if a beam with Gaussian shape

$$u(x, y, 0) = \exp\left(-\frac{x^2 + y^2}{w_0^2}\right) \qquad [29]$$

is chosen as initial condition at its narrowest point in $z = 0$ (w_0 is correspondingly called the beam waist), then the solution for arbitrary z is given by

$$u(x, y, z) = \frac{1}{1 + \mathrm{i}(2z/kw_0^2)} \times \exp\left(-\frac{x^2 + y^2}{w_0^2 + 2z\mathrm{i}/k}\right) \qquad [30]$$

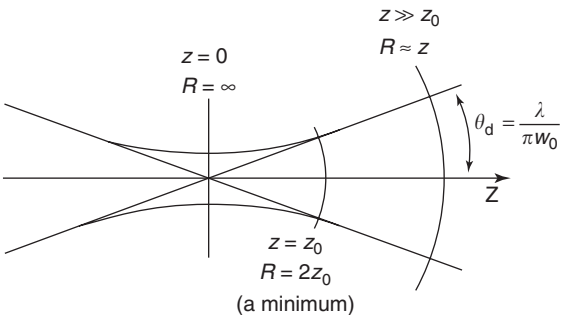


Figure 7 Geometrical meaning of the beam parameters defined in the text.

After defining new quantities (illustrated in **Figure 7**): $z_0 = kw_0^2/2$, the Rayleigh range or confocal parameter; $R(z) = z + z_0^2/z$, the beam radius of curvature; $w(z) = w_0[1 + (z/z_0)^2]^{1/2}$, the beam radius – eqn [39] can be put into the more convenient form

$$u(x,y,z) = \frac{w_0}{w(z)} \exp\left(-\frac{x^2+y^2}{w^2(z)} + \mathrm{i}\frac{k}{2}\frac{x^2+y^2}{R(z)} - \mathrm{i}\tan^{-1}\left(\frac{z}{z_0}\right)\right) \quad [31]$$

The fundamental TEM_{00} laser beam is thus

$$E_{00}(x,y,z,t) = E_0(t)\frac{w_0}{w(z)} \exp\left(-\frac{x^2+y^2}{w^2(z)} + \mathrm{i}\frac{k}{2}\frac{x^2+y^2}{R(z)} - \mathrm{i}\tan^{-1}\left(\frac{z}{z_0}\right) + \mathrm{i}(kz-\omega t)\right) \quad [32]$$

Higher-order transverse modes can also be found, and they are given in terms of the complete Hermite–Gauss set functions, so that

$$E_{mn}(x,y,z,t) = E_0(t)\frac{w_0}{w(z)} H_m \times \left(\frac{\sqrt{2}x}{w(z)}\right) H_n\left(\frac{\sqrt{2}y}{w(z)}\right) \times \exp\left(-\frac{x^2+y^2}{w^2(z)} + \mathrm{i}\frac{k}{2}\frac{x^2+y^2}{R(z)} - \mathrm{i}(m+n+1)\tan^{-1}\left(\frac{z}{z_0}\right) + \mathrm{i}(kz-\omega t)\right) \quad [33]$$

The first few transverse mode distributions are pictorially shown in **Figure 8**.

Now imagine that the resonator of **Figure 6** has mirrors at $z_1 < 0$ and $z_2 > 0$, with radii of curvature $R_1 = -R(z_1)$, $R_2 = R(z_2)$, respectively. A point (x, y, z) on mirror 2 satisfies the equation

$$R_2^2 = x^2 + y^2 + (R_2 - z_2 + z)^2 \quad [34]$$

which can be rewritten as

$$z_2 + \frac{(z-z_2)^2}{2R_2} = z + \frac{x^2+y^2}{2R_2} \quad [35]$$

The wave front of constant phase ϑ almost coincides (within lowest-order terms) with the mirror at z_2 when evaluated on the mirror surface:

$$\vartheta = k\left\{ z_2 - \frac{x^2+y^2}{2R_2} + \frac{(z-z_2)^2}{2R_2}\right\} + k\frac{x^2+y^2}{2R(z)} - (m+n+1) \times \tan^{-1}\left(\frac{z}{z_0}\right) - \omega t \quad [36]$$

and the rays strike the mirrors perpendicularly and propagate back and forth along the cavity. The losses due to phase mismatch can be eliminated by tuning the wave number $k = \omega/c$, so that the total phase change of the ray between the mirrors is an integer number l of π's:

$$l\pi = \frac{\omega L}{c} - (m+n+1) \times \left\{ \tan^{-1}\left(\frac{z_2}{z_0}\right) - \tan^{-1}\left(\frac{z_1}{z_0}\right)\right\} \quad [37]$$

where $L = (z_2 + |z_1|)$ is the cavity length.

The frequency satisfying [37] for given values of (l, m, n) is called ω_{lmn} and the corresponding field is called the TEM_{mn} transverse mode. The larger the transverse indices (m, n), the lesser the wave fronts coincide with the mirrors, which means that there is a tendency (if not otherwise managed) to have the suppression of higher-order transverse modes in favor of the lowest-order one. A change between two nearby longitudinal mode frequencies with same (m,n) is given by

$$\Delta\omega_{lmn} \equiv \omega_{l+1\,mn} - \omega_{lmn} \equiv \frac{c\pi}{L} \quad [38]$$

while varying the transverse mode indices at fixed l results

$$\delta\omega_{lmn} = \frac{c}{L}\delta(m+n) \times \left\{ \tan^{-1}\left(\frac{z_2}{z_0}\right) - \tan^{-1}\left(\frac{z_1}{z_0}\right)\right\} \quad [39]$$

where $\delta(m+n)$ is the change in $(m+n)$; it is not hard to see, for example, in the case when $z_0 \gg z_2, |z_1|$, that

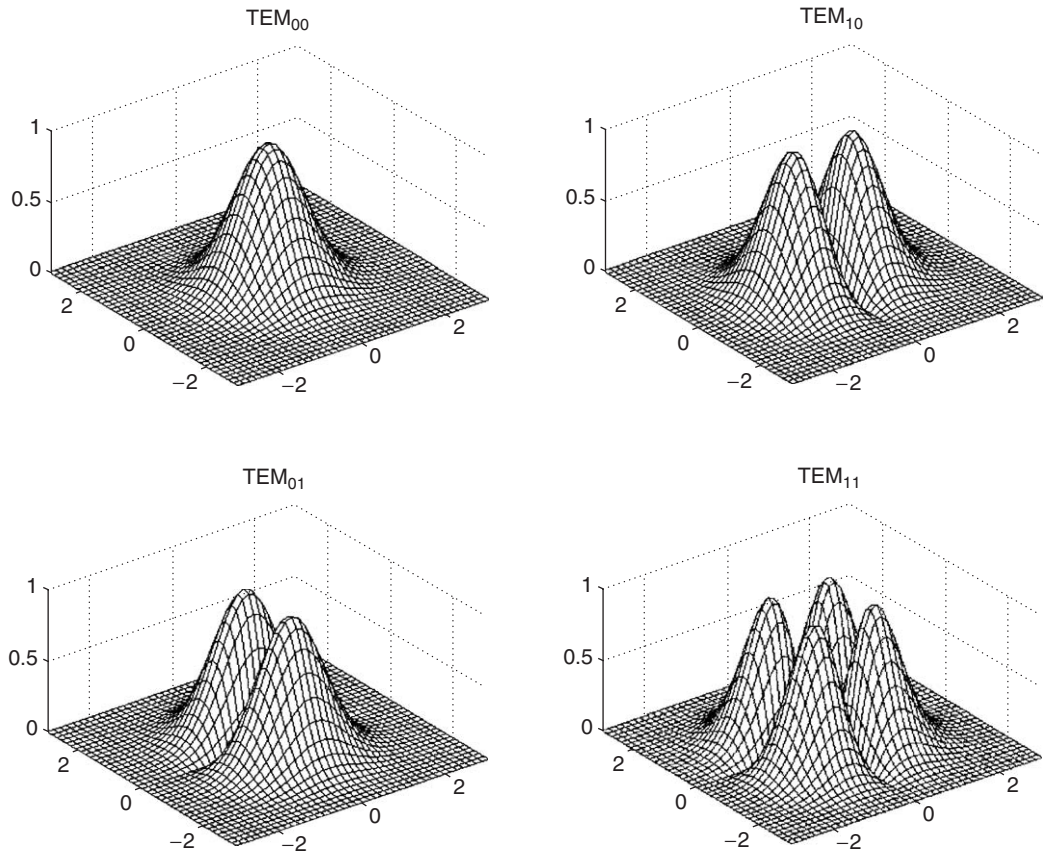


Figure 8 Patterns of first low-order Hermite–Gauss modes.

the transverse mode structure appears as a fine structure of the longitudinal mode spectrum.

It should be noted that, if the resonator is loaded with a gain medium of length ℓ, its length L must be replaced by the optical length $L' = L + (\eta - 1)\ell$, so that the phase condition [37] contains the refractive index of the gain medium $\eta(\omega,S)$, which depends on frequency as well as on laser beam intensity. This should alert that the exact longitudinal mode position is also a function of the working power of the laser.

Single Mode Continuous Laser Operation

Lasers can be operated in various regimes; some of them can be studied using the laser rate equations previously introduced:

$$\begin{aligned} \frac{dN_1}{dt} &= -\gamma_1 N_1 + \frac{\sigma S}{\hbar\omega}(N_2 - N_1) \\ \frac{dN_2}{dt} &= R_2 - \gamma_2 N_2 - \frac{\sigma S}{\hbar\omega}(N_2 - N_1) \\ \frac{1}{c}\frac{\partial S}{\partial t} + \frac{\partial S}{\partial z} &= \sigma S(N_2 - N_1) + \frac{R_{sp}}{c} - \frac{\gamma_c}{c} S \end{aligned} \qquad [40]$$

where the added spontaneous emission rate R_{sp} reminds one of the origin of the starting seed, and a loss term for the laser field has been inserted to account for both useful mirror output coupling and residual, undesirable losses due to intracavity absorption and scattering. According to [6], these are collectively given by

$$\begin{aligned} \frac{\gamma_c}{c} &= -\frac{\ln(R) + \ln(1 - L_i)}{2\ell} \\ &\approx \frac{(1 - R) + L_i}{2\ell} \end{aligned} \qquad [41]$$

where the approximation applies in the low loss limit.

In many common situations, the intensity $S(t)$ is approximately uniform along the cavity axis. Assuming this approximation, and taking into account the different lengths of the gain medium and the cavity, the final form of [40] becomes

$$\begin{aligned} \frac{dN_1}{dt} &= -\gamma_1 N_1 + \frac{\sigma S}{\hbar\omega}(N_2 - N_1) \\ \frac{dN_2}{dt} &= R_2 - \gamma_2 N_2 - \frac{\sigma S}{\hbar\omega}(N_2 - N_1) \\ \frac{dS}{dt} &= \frac{c\sigma\ell}{L} S(N_2 - N_1) + R_{sp} - \gamma_c S \end{aligned} \qquad [42]$$

which are the single-mode rate equations.

Under steady-state conditions, the rates of change of both population inversion and intensity are set to zero, and neglecting the small R_{sp} contribution, from [42] it follows

$$\begin{aligned} g(\omega, S) &= \frac{g_0(\omega)}{1 + S/S_{sat}} \\ 0 &= \left[\frac{c\ell}{L}\, g(\omega, S) - \gamma_c\right] S \end{aligned} \qquad [43]$$

where the first equation gives the saturated gain already derived in [23], with $S_{sat} = (\hbar\omega\gamma_2/\sigma)$, and $g_0 = \sigma(N_2 - N_1)_0 = \sigma R_2/\gamma_2$. Below threshold, the optical intensity remains zero until the pumping rate has grown high enough to bring the small signal gain at the threshold. Thus, g_0 grows linearly with the pump rate, and at threshold it results

$$\begin{aligned} &\frac{c\ell}{L}\, g_{th} = \frac{c\ell}{L}\, g_{0th} = \frac{c\ell}{L}\, \frac{\sigma R_{2th}}{\gamma_2} = \gamma_c \\ &R_{2th} = \frac{L}{c\ell}\, \frac{\gamma_2 \gamma_c}{\sigma} \end{aligned} \qquad [44]$$

Above threshold, $g(\omega,S)$ remains clamped to its threshold value, since $S \neq 0$ and it cannot be under the steady-state hypothesis:

$$\left[\frac{c\ell}{L}\, g(\omega, S) - \gamma_c\right] S > \left[\frac{c\ell}{L}\, g_{0th} - \gamma_c\right] S = 0 \qquad [45]$$

The result is

$$\frac{g_0}{1 + S/S_{sat}} = g_{0th} \qquad [46]$$

and for the internal intensity

$$S = S_{sat}\left(\frac{g_0}{g_{0th}} - 1\right) = S_{sat}\left(\frac{R_2}{R_{2th}} - 1\right) \qquad [47]$$

which, far above threshold, becomes $S \approx (\ell/L)(cR_2\hbar\omega/\gamma_c)$.

The resonator output intensity is, in the low loss limit of [41],

$$\begin{aligned} S_{out} &= (1 - R)S \\ &= \frac{\ell}{L}(1 - R) S_{sat}\left[\frac{2g_0\ell}{(1 - R) + L_i} - 1\right] \end{aligned} \qquad [48]$$

which is plotted in **Figure 9** as a function of the output coupling $T = (1 - R)$.

From the figure it is clear that there is an optimum working point, obtained maximizing S_{out} with respect to the output coupling at fixed small signal gain (i.e., at fixed pump rate):

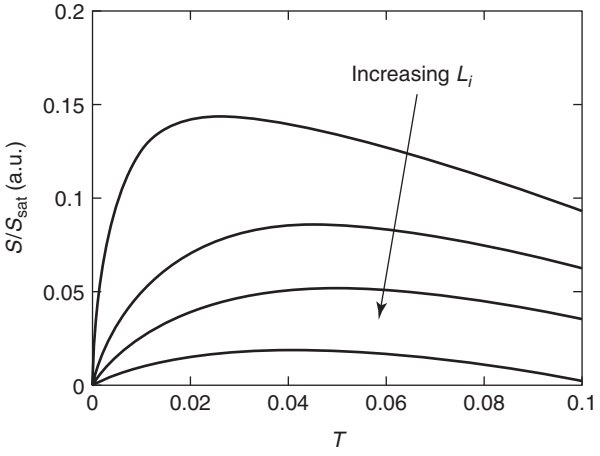


Figure 9 Output intensity as a function of outcoupling transmittance.

$$\begin{aligned} T_{opt} &= \sqrt{2g_0\ell L_i} - L_i \\ S_{out,opt} &= \frac{\ell}{L}\, S_{sat}(\sqrt{2g_0\ell} - \sqrt{L_i})^2 \end{aligned} \qquad [49]$$

which represents the working condition for well-engineered continuous wave-laser systems.

Laser Transients: Mode-Locking and Ultrashort Pulse Generation

In this section, time-dependent, transient effects are briefly discussed. There are distinctly two ways to generate powerful pulses, and which one is chosen depends on applications.

When high-power and high-energy are required, Q-switch operation is the preferred technique. This consists in keeping the Q of the resonator low at the beginning of the pump action, so that the population inversion can growth at a very high level; when this is achieved, the resonator Q is restored to its original value allowing an abrupt growth of the field to a high power concentrated in a short pulse. Q-switch can be discussed within the single-mode approximation of the previous section using eqn [44], even if fast transients always tend to bring many longitudinal modes into oscillation. Generally, this technique leads to nanosecond pulse durations, with single-pulse energies up to several hundreds of millijoules.

A different technique is used for the generation of ultrashort pulses, which find broad popularity in scientific, biomedical, and also, more recently, in industrial applications. This justifies a more in-depth description of how they are generated.

Short and ultrashort pulses generated by lasers have durations $< \sim 50$ and 1 ps, respectively. Laser media that allow such short pulses need to have wide spectral bandwidths to sustain their generation, since, according to Fourier relation, $\Delta\omega_g \tau_p \approx 2\pi$. The gain bandwidth $\Delta\omega_g$ is usually a fraction of the available spectral bandwidth. For example, the full bandwidth of the Nd:YAG allows pulses of ≈ 5 ps, while for many other laser materials this can be easily well below 1 ps.

If the number of longitudinal modes within the gain bandwidth is $N \approx \Delta\omega_g/(\pi c/L') = \Delta\omega_g T_R/2\pi$, where L' is the optical cavity length and T_R the round-trip time, an estimate from the Fourier relation gives $\tau_p \approx T_R/N$, so pulse duration depends on N. Nonetheless, this is not enough as the large number of longitudinal modes, that can be brought above threshold, tend to oscillate uncorrelated under free-running laser operation. In these conditions, the output intensity has a noisy pattern, and it is only when the modes are forced to maintain a fixed phase relationship to each other that the output as a function of time can vary in a well-defined manner. Under these conditions, the laser is said to operate in "mode-locking" regime.

Mode locking corresponds to correlating the spectral amplitudes and phases, so that the initial randomness is removed and a localized radiation pattern emerges that circulates in the laser cavity with a repetition rate of one per round trip, which appears at the output as a pulse train. The techniques that allow laser mode locking are modulation techniques implemented experimentally by placing inside the cavity either an externally driven loss or phase modulator, or a passive device which exhibits saturable absorption, which means (as for saturable gain) absorption diminishing at growing intensity. The shortest pulses, below 10 fs, have been obtained with passive techniques that are reviewed in full detail by Keller.

For the purpose of illustrating the principle, active mode locking is presented based on a time-domain model principally due to Haus, which is able to capture the main features of laser mode locking.

The model is not concerned with the starting mechanism itself, and assumes an already formed light pulse traveling inside the cavity. Looking for the conditions of self-reproducibility, let the field be

$$\boldsymbol{E} = \hat{\varepsilon} E(t) \qquad [50]$$

with

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} A(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}\omega \qquad [51]$$

$E(t)$ and $A(\omega)$ are Fourier transforms of each other, representing pulse envelopes in their domains at some position into the cavity.

Assuming a linear laser cavity of length L, containing a gain medium of length ℓ and gain coefficient with a frequency dependence $g(\omega) = g_0/[1+(\omega-\omega_0)^2/\Delta\omega_g^2]$, where g_0 is the saturated gain coefficient at the center frequency ω_0 of the gain medium, after one trip in the gain medium $A(\omega)$ becomes $\exp\{g_0\ell/[1+(\omega-\omega_0)^2/\Delta\omega_g^2]\}A(\omega)$.

The unavoidable cavity losses (either for output coupling or passive ones) can be modeled introducing a loss coefficient per unit length, so that the overall losses can be considered as $\alpha_i L_i$, with L_i an appropriate loss medium length. A passage through it produces the change $\exp(-\alpha_i L_i)A(\omega)$.

Finally, the essential element in the mode-locked laser is the modulator, for example, an active, externally driven acousto-optic device of loss coefficient α_m and length L_m, producing a time-dependent loss $\alpha_m L_m(1-\cos(\omega_m t))$ at the modulation frequency ω_m. This is best described in the time domain, affecting directly the field envelope after a passage through it, $\exp\{\alpha_m L_m(1-\cos(\omega_m t))\}E(t)$.

By taking the appropriate Fourier transforms to describe all the actions in the time domain, and assuming low net gain and low losses, so that only first-order terms in their expanded expressions are kept, after one round-trip time T_R, at the same space position,

$$E(t+T_R) = \left[1 + 2g_0\ell\left(1 + \frac{1}{\Delta\omega_g^2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right) - 2\alpha_i L_i - 2\alpha_m L_m(1-\cos(\omega_m t)\right] E(t) \qquad [52]$$

Steady-state operation requires self-reproducibility, $E(t+T_R) = E(t)$, and for this to happen it needs

$$\omega_m = 2\pi/T_R$$

Since the pulse tends to adjust itself for minimum losses, the modulation function can be expanded up to its first nonzero contributed term, getting the final form of the mode-locking equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2} - \frac{\alpha_m L_m \omega_m^2 \Delta\omega_g^2}{2g_0\ell}\, t^2 + \Delta\omega_g^2\left(1 - \frac{\alpha_i L_i}{g_0\ell}\right)\right] E(t) = 0 \qquad [53]$$

This is written in a way that shows its similarity with the Schrödinger equation of one-dimensional harmonic oscillator:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{m^2\omega^2}{\hbar^2}\, x^2 + \frac{2mE}{\hbar^2}\right]\psi(x) = 0 \qquad [54]$$

As for [54], eqn [53] has Hermite–Gauss solutions

$$E_n(t) = H_n(2\pi t/\tau_p)\exp\{-(2\pi t/\tau_p)^2/2\} \quad [55]$$

with

$$\frac{2\pi}{\tau_p} = \left(\frac{\alpha_m L_m \omega_m^2 \Delta\omega_g^2}{2g_0\ell}\right)^{1/4}$$
$$\left(1 - \frac{\alpha_i L_i}{g_0\ell}\right) = (2n+1)\left(\frac{2\pi}{\tau_p \Delta\omega_g}\right)^2 \quad [56]$$

Since as a rule the gain barely exceeds the loss, this last relation can be put into a more readable form as follows:

$$\frac{g_0\ell}{\alpha_i L_i}\left(1 - \frac{\alpha_i L_i}{g_0\ell}\right) = \left(\frac{g_0\ell - \alpha_i L}{\alpha_i L}\right)$$
$$= (2n+1)\left(\frac{2\pi}{\tau_p \Delta\omega_g}\right)^2 \frac{g_0\ell}{\alpha_i L_i}$$
$$\approx (2n+1)\left(\frac{2\pi}{\tau_p \Delta\omega_g}\right)^2 \quad [57]$$

which represents the excess gain over the threshold gain for continuous operation, needed to sustain a mode-locked Hermite–Gauss pulse of order n: higher orders need more gain.

It can be seen that gain saturation leads to the stability of lowest-order Gaussian pulse, and to the instability of higher-order solutions. If just the right net gain is supplied to excite the fundamental solution, higher-order ones do not have enough gain to start. If the supplied gain is in excess so that higher-order solutions can also start, the amplitude of the lowest order quickly grows to the point that it saturates the gain which settles it to its steady-state value, switching off the higher-order solutions.

It is thus seen from [55] that the pulse is Gaussian in shape; furthermore, from the first of [56] it is seen that the pulsewidth has a power dependence on both the modulation depth and the gain bandwidth.

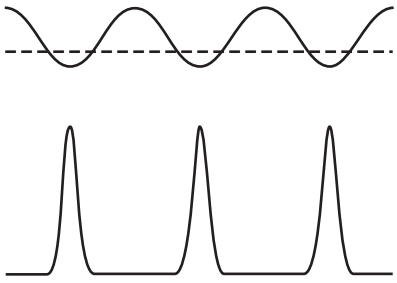


Figure 10 Gain and losses in an actively mode-locked laser, and its generated pulse train. The dotted line sets the net gain of the laser for continuous operation, and the extra, modulated loss are shown in full line above it. With periodicity of a round-trip time, a temporal window is opened in which the net gain is positive, and the pulses can go through, as shown in full line at the bottom. Pulse duration not to scale with round-trip time in the figure.

Figure 10 illustrates qualitatively the interplay of gain and losses in an actively mode-locked laser, and its Gaussian envelope pulses.

See also: Nonlinear Optics; Optical Fibers; Optical Sum Rules and Kramers–Kronig Relations; Photon Statistics and Coherence Theory.

PACS: 42.55.Ah; 42.60.Da; 42.60.Fc; 42.60.Jf; 42.60.Pk

Further Reading

Haus HA (1984) *Waves and Fields in Optoelectronics*. London: Prentice-Hall.

Keller U (1999) Semiconductor nonlinearities for solid-state laser mode-locking and Q-switching. In: Finlayson DM and Sinclair BD (eds.) *Advances in Lasers and Applications*. London: Institute of Physics.

Koechner W and Bass M (2003) *Solid-State Lasers*. New York: Springer.

Milonni PW and Eberly JH (1988) *Lasers*. New York: Wiley.

Siegman AE (1986) *Lasers*. Mill Valley: University Science Books.

Yariv A (1989) *Quantum Electronics*. New York: Wiley.

Lattice Dynamics: Anharmonic Effects

M I Katsnelson, Radboud University Nijmegen, Nijmegen, The Netherlands

Introduction

Crystal lattice dynamics is based on the concept of phonons, that is, weakly interacting waves of atomic (or ionic) vibrations and corresponding quasiparticles (according to a corpuscular-wave dualism of the quantum physics, any excitation wave in a system can also be described as a particle). Due to the smallness of typical atomic displacements, $\bar{u}$ in crystals in comparison with interatomic distances d, one can pass rigorously from a problem of strongly interacting "particles" forming the crystal (atoms, ions, or molecules) to a problem of weakly

interacting "quasiparticles" (phonons). In the leading order in the smallness parameter $\eta = \bar{u}/d$, the crystal lattice dynamics and thermodynamics can be described in terms of an ideal phonon gas (harmonic approximation). For the ground state, the harmonic approximation can be justified rigorously on the basis of the adiabatic (Born–Oppenheimer) approximation of quantum mechanics due to the smallness of the ratio of the electron mass, m to the atomic (or ionic) mass M. With the increase in temperature T, the parameter η increases as well; however, due to a semi-empirical Lindemann criterion $\eta \simeq 0.1$ at the melting point $T = T_m$, higher-order contributions to thermodynamic properties are usually small up to the melting temperature. This statement is true for most of the average characteristics. At the same time, for some peculiar modes, the harmonic approximation can be completely inadequate, especially in the vicinity of some structural transformations, such as ferroelectric phase transitions or martensitic transformations in metals (soft modes). It is not sufficient also for quantum crystals, such as solid He^3 and He^4. Even in a generic case, some phenomena can be understood only beyond the picture of the ideal phonon gas, which means beyond the harmonic approximation. All these phenomena are called anharmonic. There are anharmonic effects in the crystal lattice thermodynamics (thermal expansion, temperature dependences of elastic moduli, etc.), kinetics (phononphonon scattering processes which are responsible for the thermal conductivity of insulators), and dynamics (phonon damping and temperature dependences of the phonon frequencies measured by the inelastic neutron scattering method).

Formulation of the Problem

In the adiabatic approximation, one can split the quantum-mechanical problem of the crystal into an electronic (a solution of the Schrödinger equation for fixed coordinates of nuclei $\{\boldsymbol{r}_j\}$) and a nuclear one which is the base of the theory of crystal lattice properties. Keeping in mind the smallness of the atomic displacements $\eta \ll 1$, one can expand the energy of the nuclei $V(\{\boldsymbol{r}_j\})$ into the displacement vectors $\boldsymbol{u}_j = \boldsymbol{r}_j - \boldsymbol{R}_j$, $\boldsymbol{R}_j$ being the equilibrium lattice positions:

$$V(\{\boldsymbol{r}_j\}) = V_0 + \frac{1}{2}\sum_{j_1 j_2 \alpha_1 \alpha_2} V^{\alpha_1\alpha_2}_{j_1 j_2}\, u^{\alpha_1}_{j_1} u^{\alpha_2}_{j_2} + \frac{1}{6}\sum_{j_1 j_2 \alpha_1 \alpha_2} V^{\alpha_1\alpha_2\alpha_3}_{j_1 j_2 j_3}\, u^{\alpha_1}_{j_1} u^{\alpha_2}_{j_2} u^{\alpha_3}_{j_3} + \cdots,$$

$$V^{\alpha_1 \ldots \alpha_n}_{j_1 \ldots j_n} = \left(\frac{\partial^n V}{\partial u^{\alpha_1}_{j_1} \ldots u^{\alpha_n}_{j_n}}\right)_{u=0} \qquad [1]$$

where α_j are Cartesian indices and the linear term is absent due to the equilibrium conditions. The harmonic approximation corresponds to taking into account only the quadratic term in the expansion [1]. Then the nuclear Hamiltonian, which is the sum of the potential energy V and the kinetic energy of nuclei, can be represented as a sum of the Hamiltonians of independent oscillators (phonons) by the transformation

$$\boldsymbol{u}_j = \sum_\lambda \sqrt{\frac{\hbar}{2M_j N_0 \omega_\lambda}}\, A_\lambda \boldsymbol{e}_\lambda \exp(\mathrm{i}\boldsymbol{q}\boldsymbol{R}_j),$$
$$A_\lambda = b_\lambda + b^\dagger_{-\lambda} \qquad [2]$$

where N_0 is the number of unit cells in the crystal, M_j is the mass of the jth nucleus, $\lambda \equiv \boldsymbol{q}\xi$ are the phonon quantum numbers ($\boldsymbol{q}$ is the wave vector running the Brillouin zone and ξ is the polarization index, $-\lambda \equiv -\boldsymbol{q}\xi$), $\boldsymbol{e}_\lambda$ is the polarization vector, b_λ and $b^\dagger_\lambda$ are the annihilation and creation phonon operators. In the phonon representation, the total Hamiltonian of the crystal lattice has the form

$$H = V_0 + H_0 + \sum_{k=3}^{\infty} H^{(k)},$$
$$H_0 = \sum_\lambda \hbar\omega_\lambda \left(b^\dagger_\lambda b_\lambda + \frac{1}{2}\right),$$
$$H^{(k)} = \sum_{\lambda_1 \ldots \lambda_k} \frac{\Phi^{(k)}_{\lambda_1 \ldots \lambda_k}}{k!} A_{\lambda_1} \ldots A_{\lambda_k} \qquad [3]$$

where H_0 is the Hamiltonian of the ideal phonon gas (the harmonic approximation) and the multiphonon scattering matrix elements $\Phi^{(k)}_{\lambda_1 \ldots \lambda_k}$ are proportional to the *kth* derivatives of the potential energy and the factors $\sqrt{\hbar/2M_j\omega_{\lambda_i}}$. These matrix elements describe the processes of phonon–phonon interaction, such as a merging of two phonons into one or, vice versa, a decay of a phonon into two ($k = 3$), scattering of two phonons into two new states ($k = 4$), etc. The anharmonic effects connected with these interactions are called self-anharmonic. There is also another kind of anharmonic effect which is connected to the dependence of the phonon frequencies ω_λ on the interatomic distances in the Hamiltonian H_0. These effects are called quasiharmonic. They are characterized by the microscopic Grüneisen parameters

$$\gamma_\lambda = -\frac{\partial \ln \omega_\lambda}{\partial \ln \Omega} \qquad [4]$$

where Ω is the crystal volume. For noncubic crystals, the dependence of ω_λ on shear deformations should also be considered. The point is that for harmonic oscillators described by the quadratic potential energy and, correspondingly, by linear interatomic forces, the frequencies are not dependent either on the

amplitudes of oscillations or on the equilibrium positions. Due to nonlinear (anharmonic) effects, there are renormalizations of the frequencies. Up to the lowest order η^2, one should take into account the quasiharmonic effects and the self-anharmonic effects with $k = 3$ and 4.

Thermodynamics

Thermal Expansion

Thermal expansion, characterized by the coefficient,

$$\alpha_p = \frac{1}{\Omega}\left(\frac{\partial\Omega}{\partial T}\right)_p \qquad [5]$$

(T is the temperature and p is the pressure) equals to zero in harmonic approximation. This can be easily proved from the Gibbs distribution for the potential energy V, which is quadratic in the atomic displacements $\boldsymbol{u}_j$ (and, moreover, if it is just an arbitrary even function of $\boldsymbol{u}_j$). It can be calculated by using the known thermodynamic identity

$$\left(\frac{\partial\Omega}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T \qquad [6]$$

where S is the entropy; as a consequence, due to the third law of thermodynamics, the thermal expansion coefficient should vanish at $T \to 0$. If one calculates the entropy in harmonic approximation but with the phonon frequencies dependent on the volume (quasiharmonic approximation), one can derive the Grüneisen law

$$\alpha_p = \frac{\gamma(T) C_V(T)}{\Omega B_T} \qquad [7]$$

where $B_T = -\Omega(\partial p/\partial\Omega)_T$ is the isothermal bulk modulus,

$$\gamma(T) = \frac{\sum_\lambda \gamma_\lambda C_\lambda}{\sum_\lambda C_\lambda},$$
$$C_\lambda = \left(\frac{\hbar\omega_\lambda}{k_B T}\right)^2 \frac{\exp(\hbar\omega_\lambda/k_B T)}{[\exp(\hbar\omega_\lambda/k_B T) - 1]^2} \qquad [8]$$

is the macroscopic Grüneisen parameter (which is temperature independent assuming that $\gamma_\lambda = \text{const}$) and $C_V(T) = k_B \sum_\lambda C_\lambda$ is the constant-volume lattice heat capacity. It follows from eqn [7] that the temperature dependence of the thermal expansion coefficient at low and high temperatures is the same as for the heat capacity: $\alpha_p \sim T^3$ at $T \ll \theta_D$ (θ_D is the Debye temperature) and $\alpha_p \simeq \text{const}$ at $T \geqslant \theta_D$.

For noncubic crystals, one should introduce parameters characterizing the anisotropy of the thermal expansion $\alpha_i = \partial u_i/\partial T$, where u_i are different deformations, for example, for uniaxial crystals with c-axis different from a- and b-axes, one can introduce $du_1 = d\ln\Omega$ and $du_2 = d\ln(c/a)$. As a generalization of eqn [6], one can prove from the equilibrium conditions at finite temperatures

$$\alpha_i = \sum_j (B^{-1})_{ij}\left(\frac{\partial S}{\partial u_j}\right)_T \qquad [9]$$

where $B_{ij} = (1/\Omega)(\partial^2 F/\partial u_i \partial u_j)$ is the matrix of isothermal elastic moduli and F is the free energy. In particular, for the uniaxial crystals

$$\left(\frac{\partial \ln c}{\partial T}\right) = \frac{1}{3BB_{22}}\left[(B_{22} - 2B_{12})\frac{\partial S}{\partial u_1} + (2B_{11} - B_{12})\frac{\partial S}{\partial u_2}\right]$$
$$\left(\frac{\partial \ln a}{\partial T}\right) = \frac{1}{3BB_{22}}\left[(B_{22} + B_{12})\frac{\partial S}{\partial u_1} - (B_{11} + B_{12})\frac{\partial S}{\partial u_2}\right] \qquad [10]$$

where $B = B_{11} - B_{12}^2/B_{22}$ is the bulk modulus of the uniaxial crystal.

Temperature Dependence of Elastic Moduli

Temperature dependence of the elastic moduli B_{ij} is another important anharmonic effect. This temperature dependence in quasiharmonic approximation results from the ideal phonon gas contribution to the free energy, $F_{ph} = -k_B T \sum_\lambda \ln[2\sinh(\hbar\omega_\lambda/2k_B T)]$, and from the volume dependence of the electronic contributions to the moduli. As a result, both terms behave as $\delta B_{ij} \propto -T^4$ at $T \ll \theta_D$ and $\delta B_{ij} \propto -T$ at $T \geqslant \theta_D$. Normally, this contribution is negative (elastic moduli decrease with the temperature increase), but for some shear moduli in peculiar cases an opposite behavior sometimes takes place (acoustic soft modes), usually near the structural phase transitions. Empirically, for many cubic crystals the trigonal shear modulus B_{44} at the melting point is 55% of its value at zero temperature (Varshni melting criterion).

Lattice Heat Capacity at High Temperatures

In harmonic approximation, the molar constant-volume heat capacity at $T \gtrsim \theta_D$ is independent of both the temperature and the chemical composition of the crystal: $C_V = 3R$ (Dulong–Petit law, $R \simeq 8.3144\ \text{J mol}^{-1}\ \text{K}^{-1}$ is the gas constant). Self-anharmonic effects lead to linear temperature dependence of the heat capacity, $\delta C_V \sim R\eta^2 \sim R(k_B T/E_{coh})$

where E_{coh} is a typical energy of the chemical bonding. These terms arise from both three-phonon and four-phonon processes (the second-order perturbation effect in $H^{(3)}$ and the first-order one in $H^{(4)}$, see eqn [3]):

$$C_V^{(3)} = \frac{4k_B^2 T}{3} \sum_{\boldsymbol{k}q\xi\eta\varsigma} \frac{|\Phi^{(3)}_{\xi\boldsymbol{k};\eta q;\varsigma,-\boldsymbol{k}-q}|^2}{\hbar^3 \omega_{\xi\boldsymbol{k}} \omega_{\eta q} \omega_{\varsigma,-\boldsymbol{k}-q}} \qquad [11]$$

$$C_V^{(4)} = -k_B^2 T \sum_{\boldsymbol{k}q\xi\eta\varsigma} \frac{\Phi^{(4)}_{\xi\boldsymbol{k};\xi,-\boldsymbol{k};\eta q;\eta,-q}}{\hbar^2 \omega_{\xi\boldsymbol{k}} \omega_{\eta q}} \qquad [12]$$

The three-phonon contribution [11] is always positive (the growth of C_V with the increase in the temperature), whereas the four-phonon one [12] can be, in principle, of arbitrary sign. Experimental separation of the self-anharmonic contribution to the lattice heat capacity is a difficult problem since it has, in general, the same order of magnitude and temperature dependence as the difference

$$C_P - C_V = T\Omega B_T \alpha_p^2 \qquad [13]$$

(C_P is the experimentally measurable heat capacity at a constant pressure) and, in metals, as the electron heat capacity.

Phonon Spectra and Damping

Temperature Dependences of Phonon Frequencies

According to classical mechanics, for a generic nonlinear system, the oscillation frequencies are dependent on the oscillation amplitudes. Therefore, one can expect that anharmonic effects lead to the temperature-dependent phonon spectra, due to growth of average oscillation amplitudes with the increase in temperature. In quantum terms, the same effect can be described as an appearance of the phonon self-energy due to the phonon–phonon interaction processes. Up to the second order in the smallness parameter η, the temperature shift of the phonon frequency ω_λ is determined by the following expression:

$$\Delta\omega_\lambda = \Delta_\lambda^{(\text{qh})} + \Delta_\lambda^{(3)} + \Delta_\lambda^{(4)} \qquad [14]$$

where $\Delta_\lambda^{(\text{qh})} = -\gamma_\lambda \Delta\Omega(T)/\Omega$ is the quasiharmonic contribution due to the temperature dependence of the crystal volume $\Omega(T)$, and $\Delta_\lambda^{(3)}$ and $\Delta_\lambda^{(4)}$ are the contributions of the three-phonon and four-phonon processes, correspondingly:

$$\begin{aligned}\Delta_{\xi\boldsymbol{k}}^{(3)} = &-\frac{1}{2\hbar^2}\mathscr{P}\sum_{q\eta\varsigma} |\Phi^{(3)}_{\xi\boldsymbol{k};\eta q;\varsigma,-\boldsymbol{k}-q}|^2 \\ &\times \left(\frac{1 + N_{\eta q} + N_{\varsigma,\boldsymbol{k}+q}}{\omega_{\eta q} + \omega_{\varsigma,\boldsymbol{k}+q} + \omega_{\xi\boldsymbol{k}}} + \frac{1 + N_{\eta q} + N_{\varsigma,\boldsymbol{k}+q}}{\omega_{\eta q} + \omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\xi\boldsymbol{k}}} \right. \\ &+ \frac{N_{\eta q} - N_{\varsigma,\boldsymbol{k}+q}}{\omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\eta q} - \omega_{\xi\boldsymbol{k}}} \\ &\left. - \frac{N_{\eta q} - N_{\varsigma,\boldsymbol{k}+q}}{\omega_{\eta q} - \omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\xi\boldsymbol{k}}} \right) \end{aligned} \qquad [15]$$

$$\Delta_{\xi\boldsymbol{k}}^{(4)} = \frac{1}{2\hbar} \sum_{q\eta} \Phi^{(4)}_{\xi\boldsymbol{k};\xi,-\boldsymbol{k};\eta q;\eta,-q} (1 + 2N_{\eta q}) \qquad [16]$$

where $\mathscr{P}$ is the principal value symbol and $N_\lambda = (\exp \hbar\omega_\lambda / k_B T - 1)^{-1}$ is the Planck function. At high temperatures $T \gtrsim \theta_D$, one can use the classical asymptotic $N_\lambda \simeq k_B T/\hbar\omega_\lambda$ and see that all three contributions in eqn [14] are linear in temperature. Usually, the phonon frequencies decrease with the temperature increase; a typical behavior is shown in **Figure 1** (upper panel). However, for the soft modes, $d\omega_\lambda/dT > 0$, as it is illustrated by **Figure 2**. In the framework of the perturbation theory, this behavior is connected with the contribution $\Delta_\lambda^{(4)}$.

Phonon Damping

Beyond the harmonic approximation, the phonons cannot be considered as stable quasiparticles; for example, due to the three-phonon processes, they can decay into couples of other phonons. As a result, the phonon damping arises (which can be measured experimentally as a half-width of phonon peaks in inelastic neutron scattering spectra). In the lowest-order perturbation theory, the damping, or the inverse phonon lifetime, is equal to

$$\begin{aligned}\Gamma_{\xi\boldsymbol{k}} = &\frac{\pi}{2\hbar^2} \sum_{q\eta\varsigma} |\Phi^{(3)}_{\xi\boldsymbol{k};\eta q;\varsigma,-\boldsymbol{k}-q}|^2 \{(1 + N_{\eta q} + N_{\varsigma,\boldsymbol{k}+q}) \\ &\times \delta(\omega_{\eta q} + \omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\xi\boldsymbol{k}}) + (N_{\eta q} - N_{\varsigma,\boldsymbol{k}+q}) \\ &\times [\delta(\omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\eta q} - \omega_{\xi\boldsymbol{k}}) \\ &- \delta(\omega_{\eta q} - \omega_{\varsigma,\boldsymbol{k}+q} - \omega_{\xi\boldsymbol{k}})]\} \end{aligned} \qquad [17]$$

The delta functions in eqn [17] correspond to the energy and momentum conservation laws for the decay processes. At high temperatures $T \gtrsim \theta_D$, the damping is linear in T (see **Figure 1**, lower panel). Note that, both the damping and the frequency shift do not vanish at $T = 0$ where all $N_\lambda = 0$. These residual effects are due to quantum zero-point oscillations; they are of the order of $\sqrt{m/M}$ and small, in general (with the exception of the quantum crystals). For

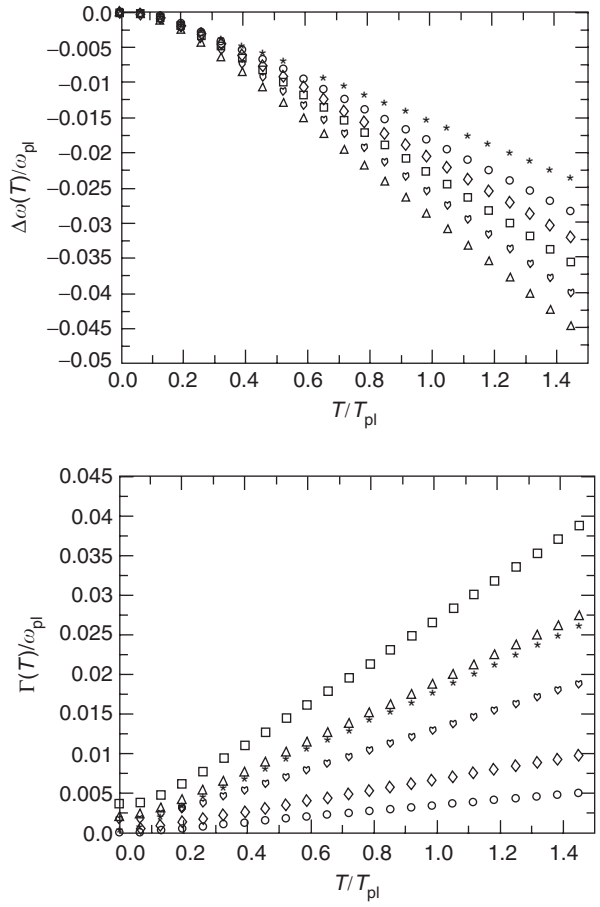


Figure 1 Temperature dependences of the frequency shifts (upper panel) and phonon damping (lower panel) for different phonon modes at the symmetric points of the Brillouin zone for face-centered cubic (f.c.c) phase of calcium. $\omega_{pl} = (4\pi Ze^2/M\Omega_0)^{1/2}$ is ionic plasma frequency, Z, M, and Ω_0 are the valency, ionic mass, and atomic volume, respectively, $T_{pl} = \hbar\omega_{pl}/k_B = 476$ K is the corresponding temperature. Asterisks and hearts label different phonon modes for L point, circles and triangles for W point, and squares and diamonds for X point. (From Katsnelson MI, Trefilov AV, Khlopkin MN, and Khromov KYu (2001) *Philosophical Magazine B* 81: 1893.)

acoustic phonons with $q \to 0$, the damping [17] is linear in q, as well as in the phonon frequency; however, this is not true for the case $ql \ll 1$, where l is the phonon mean free path. For this regime, according to general hydrodynamics consideration, the damping behaves like $q^2 l$.

For strongly anharmonic modes, the phonon damping can be comparable with the frequency; phonons are not well-defined in such a situation (high-temperature body-centered cubic (b.c.c) phases of Ti and Zr can be considered as examples). The perturbation theory reviewed here can also be insufficient under resonance conditions, where the frequency ratio for different phonons with the same wave vector is close to 1:2, 1:3, etc.

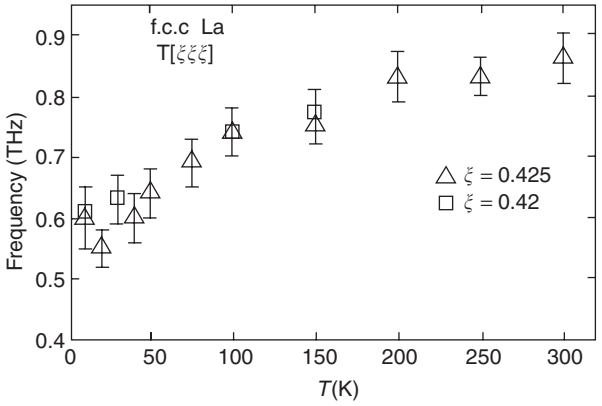


Figure 2 Temperature dependence of the phonon frequencies of the $T[\xi\xi\xi]$ phonon branch in (f.c.c) lanthanum in the soft mode region. (From Stassis C, Smith GS, Harmon BN, Ho KM, and Chen Y (1985) *Physical Review B* 31: 6298.)

Transport Properties

If one neglects phonon–phonon interactions as well as phonon scattering by any defects (impurities, different isotopes, crystal surface, etc.) and by other quasiparticles (conduction electrons in metals, spin waves in magnetic crystals, etc.), the mean free path $l = \infty$ which means, in particular, infinitely large thermal conductivity κ (any nonequilibrium distribution of the phonon momenta and energies conserves for an ideal phonon gas). This means that for the case of perfect, isotopically homogeneous, large enough insulating crystals, the anharmonic effects should determine the values and temperature dependences of both l and κ; they are connected by a simple relation:

$$\kappa = \frac{1}{3}\frac{C_V \bar{v} l}{\Omega} \qquad [18]$$

where $\bar{v}$ is a characteristic sound velocity. In principle, the same processes of phonon decay which are responsible for $\Gamma_{\xi k}$ contribute to the mean free path. However, for a continuum medium theory, any process of phonon–phonon interaction cannot lead to a finite thermal conductivity, since the momentum is conserved at any individual interaction act and any redistribution of the momenta among the phonons cannot change the energy current of the phonon gas as a whole. In crystals, Umklapp processes are possible when the momentum is conserved with the accuracy of some nonzero reciprocal lattice vector $\mathbf{g}$, and only these processes lead to the relaxation of the energy current and to finite κ. The energy conservation law for the Umklapp processes results in the appearance of some finite activation energy so that $l \propto \exp(\text{const} \times \theta_D/T)$ at $T \to 0$. For the very small temperatures, when the phonon mean free path is larger than the crystal size L, one should replace $l \to L$

in eqn [18] and thus $\kappa \propto C_V \propto T^3$. For high temperatures $T \geqslant \theta_D$, one has $\kappa \propto l \propto 1/T$.

See also: Ferroelectricity; Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Structural Instability and Soft Modes; Lattice Dynamics: Vibrational Modes; Thermodynamic Properties, General.

PACS: **63.20.Kr; 63.20.Ry; 65.40.Ba; 65.40.De**

Further Reading

Aksenov VL, Plakida NM, and Stamenkovic S (1989) *Neutron Scattering by Ferroelectrics*. Singapore: World Scientific.

Bötger H (1983) *Principles of the Theory of Lattice Dynamics*. Berlin: Academic–Verlag.

Cowley RA (1963) The lattice dynamics of an anharmonic crystal. *Advances in Physics* 12: 421.

Cowley RA (1968) Anharmonic crystals. *Reports on Progress in Physics* 31: 123.

Kaldis E, Mihailovic D, Müller KA, and Ruani G (eds.) (1995) *Anharmonic Properties of High-T_c Cuprates*. Singapore: World Scientific.

Katsnelson MI and Trefilov AV (2002) *Crystal Lattice Dynamics and Thermodynamics*. Moscow: Atomizdat.

Katsnelson MI and Trefilov AV (2004) Nonperturbative anharmonic phenomena in crystal lattice dynamics. In: Tokuyama M and Oppenheim I (eds.) *Slow Dynamics in Complex Systems: Third International Symposium*, vol. 708, p.727. AIP Conference Proceedings Series.

Leibfried G and Ludwig W (1961) Theory of anharmonic effects in crystals. In: Seitz F and Turnbull D (eds.) *Solid State Physics*, vol. 12, p. 275. New York: Academic Press.

Peierls R (2001) *Quantum Theory of Solids*. Oxford: Oxford University Press.

Vaks VG, Kravchuk SP, and Trefilov AV (1988) Microscopic theory of anharmonic effects in alkali and alkaline earth metals. *Soviet Physics – Uspekhi* 31: 474.

Varshni YP (1970) Temperature dependence of the elastic constants. *Physical Review B* 2: 3952.

Ziman JM (2001) *Electrons and Phonons: The Theory of Transport Phenomena in Solids*. Oxford: Oxford University Press.

Lattice Dynamics: Aperiodic Crystals

J-B Suck, Chemnitz University of Technology, Chemnitz, Germany

Introduction

Crystals with aperiodic structure are solids with long-range order (LRO) in the atomic positions but no translational invariance of a unit cell position. Their Fourier module consists of sharp diffraction spots, proving their LRO, which includes a long-range bond orientational order (BOO). The wave vectors Q of these diffraction spots can be represented by the sum over integer multiples of a number of rationally independent basis vectors, k, which span the Fourier module of the aperiodic crystal:

$$Q = \sum_{i=1}^{n} m_i k_i \qquad [1]$$

In contrast to periodic crystals, where the number of rationally independent basis vectors is equal to the dimension of the crystal (and of its periodic reciprocal lattice), for aperiodic crystals n exceeds the dimension of the space, in which the crystal is embedded. The minimal necessary number of independent basis vectors defines the "rank" of the Fourier module, which is finite in the quasiperiodic case. In a space of dimension n (the n-dim space), the crystal structure, aperiodic in three-dimensional (3D) Euclidean space, can be presented as a "periodic" structure with a unit cell, which usually has a relatively simple structure. Generally, the aperiodic structure can be obtained by an appropriate intersection of the n-dim periodic structure with the physical space, in which the aperiodic structure is observed. For quasicrystals (QCs), this intersection involves the intersection of the n-dim lattice with the physical space, which has an "irrational" slope relative to the orientation of the n-dim lattice. After a small change of this slope to a rational value, this intersection will lead to a crystal in physical space, which is periodic, but locally has a structure similar to the QCs. Such crystals are called "approximants" (to the corresponding QCs), the dynamics of which are discussed below as well.

Incommensurably modulated crystals, incommensurate composites, and QCs all belong to the aperiodic crystals. However, in some of their fundamental properties, they are different from each other. Therefore, the restriction to one or the other of them is mentioned explicitly as has just been done for QCs. Assuming that the physical (direct, external, or parallel) space, by which the n-dim space is intersected, is 3D, the remaining $(n-3)$-dimensional subspace is the "internal" or "perpendicular" (perp) space, which in the case of an incommensurably modulated crystal can be viewed as containing the information on the phase of the modulation function with respect to the original commensurate lattice, out of which the incommensurate phase developed. In view of the

atomic dynamics, it is important to realize that the additional dimensions belonging to the internal space can be interpreted as additional degrees of freedom not present in periodic crystals.

In n-dim space, the information about the atomic positions (density distribution) in the aperiodic structure in physical space is "stored" in $n-3$ dimensional hypersurfaces, for QCs called "atomic surfaces." The aperiodic structure is the intersection of the physical space with this periodic array of hypersurfaces. It depends on the shape of these atomic surfaces, and the consequences a shift of the physical space in the direction of the internal space will have on the aperiodic structure after the intersection: for continuous surfaces, it just results in a phase change of an incommensurate modulation. For disjunct atomic surfaces, as, for example, in the case of QCs, it leads to sudden changes of the atomic distribution, all of which, however, have the same energy. Thus, the ground state is infinitely degenerated. Likewise, a slow change of the modulation phase does not cost energy.

The Fourier module of the n-dim lattice is a periodic reciprocal n-dim lattice with sharp diffraction spots. This is the space in which inelastic scattering experiments are done to investigate the atomic dynamics. For incommensurate crystals, the 3D reciprocal space can be constructed on the basis of three reciprocal basis vectors. For QCs, which also contain rotation axes other than two-, three-, four-, and sixfold, one obtains the physical reciprocal space by a projection of the n-dim reciprocal space on the physical reciprocal space. Again, $n-3$ reciprocal lattice components belong to the reciprocal perpendicular space, and are usually called Q_{perp}. As the 3D projection plane in the n-dim reciprocal lattice has to have the same irrational slope with respect to the n-dim reciprocal lattice as the intersection in the n-dim real space had, all Bragg peaks of the n-dim reciprocal lattice will be projected into this hyperplane, as there is no mutual "shading" of different reflections because of the irrational slope. Thus, the Fourier module in this case consists of a dense set of Bragg reflections. In principle, this leads to infinitely many overlapping Brillouin zones and correspondingly to infinitely many gaps in the dispersion curves. Fortunately however, most of the Bragg peaks have an inmeasurably small intensity and can therefore be neglected, leading to a diffraction pattern with a finite number of intense Bragg (for single crystals) or Debye–Scherrer peaks (for polycrystals), the intensity of which is highest for peaks stemming from Bragg peaks in the n-dim Fourier module with smallest Q_{perp}. This fact enables the definition of pseudo-Brillouin zones (PBZs) next to reciprocal lattice points with highest Bragg peak intensity and dispersion relations for phonon and phase modes.

The dynamics of aperiodic phases has been studied most often by neutron inelastic scattering (NIS). Special questions such as the amplitudon mode and the central peak have been investigated by light scattering, and some experiments have also been done by X-ray inelastic scattering (XIS).

As aperiodic crystals usually contain more than one type of scatterer and different scatterers couple differently to the used probe, in all cases the total dynamic structure factor, $S(Q,\omega)$, was determined. For samples with n elements this is the weighted sum of the $n(n+1)/2$ partial dynamic structure factors, $S_{i,j}(Q,\omega)$, where the weights are given by the coupling strength to the probe (here NIS is used as an example in the following equation and consequently the scattering length b) and the atomic concentration c of each of the n elements:

$$\sigma S(Q,\omega) = 4\pi \sum \sum b_i b_j c_i c_j S_{i,j}(Q,\omega) + \sum \sigma_i^{\mathrm{inc}} c_i S_i^{\mathrm{s}}(Q,\omega) \quad (i,j=1,\ldots,n) \quad [2]$$

where σ are σ^{inc} are the total and the incoherent scattering cross sections and S_i^{s} is the self part of the dynamic structure factor of the element i. Alternatively, the single-particle motions can be investigated via the generalized vibrational density-of-states (GVDOS), which again is the weighted sum of the partial density-of-states of each element in the sample, weighted by the strength of the coupling of this element to the scattered probe (here for neutrons again):

$$G(\omega) = \frac{\sum w_i c_i g_i(\omega)}{\sum w_i c_i} \quad [3]$$

$$w_i = \frac{\mathrm{e}^{-2W_i}\sigma_{\mathrm{sc}}^i}{M_i} \quad [4]$$

Here, e^{-2W_i} is the Debye–Waller factor and M_i is the mass of the scatterer i. The density-of-states (DOS), as always, contains the information on the single-particle dynamics of the system, while the dynamic structure factor reflects also the collective atomic dynamics for coherent neutron scatterers, as is needed for measurements of dispersion relations.

In what follows, the vibrational modes (propagation and displacements in external space in the n-dim representation) and the phason modes (propagation in external space, displacements in internal space) of the three most often met aperiodic crystals, incommensurate modulated structures, incommensurate composites, and QCs are discussed separately.

Incommensurate Modulated Structures

Incommensurate modulated structures are crystal structures in which the atoms in the originally periodic lattice in the normal (N) phase have been displaced by a periodic (static) modulation function in the transition from the N to the incommensurate (I) phase, the period of this modulation being incommensurate with the periodicity of the original lattice. For a 3D crystal, eqn [1] can therefore be rewritten as

$$Q = ha^* + kb^* + lc^* + \sum_{i=4}^{n} m_i k_i \qquad [5]$$

where the first three terms describe the reciprocal lattice of the 3D basic (periodic) average structure, and a^*, b^*, c^*, and k_i are rationally independent. In reciprocal space, these structures (and with this also their dynamics) are characterized by a reciprocal lattice which contains the Bragg peaks of the original periodic lattice plus satellites from the incommensurate modulation. Due to the fact that incommensurate structures often form as a consequence of frustration caused by competing interactions, the incommensurate phase normally exists only within a certain temperature (and pressure) range, which is bounded by two transition temperatures, T_I and T_c. $\Delta T = T_I - T_c$ is, in some cases, only a few degrees wide, can, however, also extend down to $T_c = 0\,K$. If $T_c > 0\,K$, the wavelength of the modulation locks-in at T_c onto a rational fraction of the underlying (the main) lattice periodicity, either in a strongly discontinuous manner or by disappearance of phase defects. In contrast to this, the transition into the incommensurate phase at T_I is a second-order phase transition with a well-defined order parameter. From the viewpoint of the atomic dynamics, this transition from the periodic basis structure to the aperiodic modulation is normally associated with the softening of a phonon branch and the appearance of a Lorentzian shaped central peak at zero energy, part of which is most likely of dynamical origin also. The softening (additional (broad) minimum in the dispersion branches (see **Figure 2**)) of the dispersion occurs at a wave vector, which will become the modulation wave vector of the incommensurate phase as is shown in **Figure 1**.

For continuous modulation functions, theory (excluding the influence of defects) predicts a softening down to a zero frequency mode, with a finite phason gap for interruptions of the continuity of the modulation function by defects. Experimentally, the soft mode branch finishes at T_I with a finite frequency, which however, because of the strong overdamping

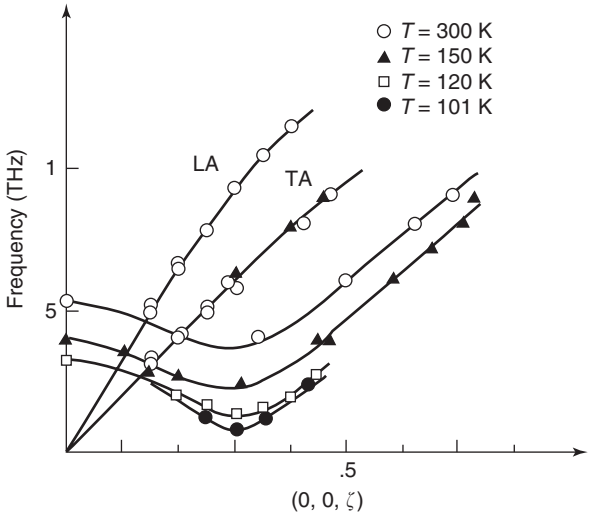


Figure 1 Mode softening in ThBr above the transition into the incommensurate modulated phase at T_I. Approaching T_I from higher T, the modes show a frequency decrease next to the modulation wave vector $q_s = 3.1\,nm^{-1}$, which arrives close to $\omega = 0$ in the I-phase.

of the soft mode, often cannot be resolved in NIS experiments. In the I-phase, this branch will then split up into two branches, the "amplitudon" (acoustic like) and the "phason" branches, the latter being related to a continuous shift of the modulation function with respect to the lattice. Near T_c, discommensurations (see below) develop which always lead to a finite gap for the phason mode, called "phason gap." The intensity of the acoustic modes is proportional to the intensity of the Bragg peak or satellite reflection at the Γ-point. This enables the identification of the phason mode, which is too intense for being an acoustic mode at the (weak) satellite reflection. In addition, the slope of the phason dispersion is lower than that of the lowest acoustic mode. The corresponding dispersion relations are shown in **Figure 2**.

While the amplitudon branch will increase with decreasing temperature (normal hardening of the acoustic phonons on temperature decrease), the phason energy may be temperature independent, as has been demonstrated for $(ClC_6D_4)_2SO_2$, for which the phason gap ($\sim 80\,GHz$) could be resolved. The temperature dependence of the width of the modes $\Gamma(T)$ is expected to be approximately the same for all three modes, which meet at the incommensurate modulation wave vector q_s.

The theoretical description of the atomic dynamics of the incommensurate phases often makes use of the periodicity of the corresponding crystals when lifted into n-dim space or applies models for the incommensurate structure in real space. Most of the

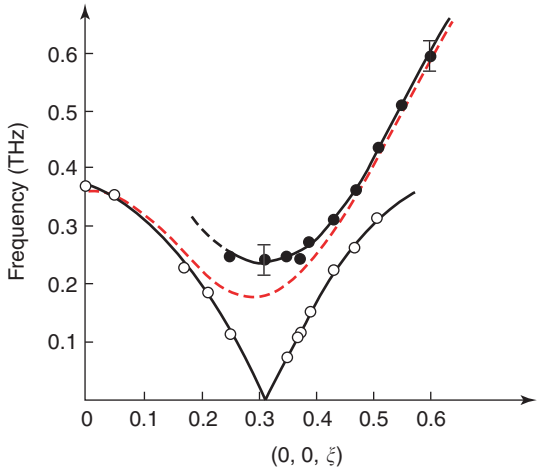


Figure 2 Phason and amplitudon branches next to the incommensurate modulation vector at 3.1 nm^{-1} in ThBr below T_I.

experimental results for incommensurate modulated crystals have been modeled very successfully using the discrete frustrated Φ^4 (DIFFOUR) model near the phase transition, where the modulation amplitude is still small. At temperatures well below T_I, the modulation amplitude may become rather large. As a consequence of this, a transition from a continuous to a discontinuous modulation function takes place, which means that the atomic surfaces are no longer continuous and smooth. This is always the case for QCs, where the atomic surfaces are generally discrete bounded objects. In the other two aperiodic crystals, modulated and composite, however, such a transition, a "discommensuration transition," may take place, resulting in domains with a structure close to a commensurate superstructure, separated by strongly incommensurate domain walls, sometimes called solitons. In this case, the description of the atomic dynamics has to also take into account an harmonic and nonlinear contributions. This situation was successfully described using the incommensurate version of the Frenkel–Kontorova model.

The transition occurs in metals and insulators for different reasons: in metals, an incommensurability of the Fermi wave vector with the basis lattice periodicity leads to a Peierls instability, which gives rise to the lattice dynamical precursor effects (local softening of the longitudinal acoustic branch) near the ordering vector, $2\boldsymbol{k}_F$. In insulators, the incommensurate structure either results from an intrinsic lattice instability, which corresponds to the "displacive" limit, or from a collective ordering mode in the "order–disorder limit." In these cases, the corresponding modulation wave vector is not predictable beforehand. Thus, the lattice dynamics is at the origin of these modulated phases.

Incommensurate Composites

Different from the crystals with incommensurate modulation, which is the only aperiodic structure, where a dominant "parent" periodic structure persists, incommensurate composites can be imagined to be made up of two (or more) interpenetrating periodic sublattices, the unit cell parameters of which are mutually incommensurate. Each of these sublattices, modeled by the incommensurate double chain model (DCM), pertain, to a large extent, to their dynamical characteristics as far as phonons are concerned. Thus, for example, two longitudinal acoustic branches have been found in $Bi_2Sr_2CaCu_2O_{8+\delta}$, which consists of two sublattices, one for each subsystem. However, as all aperiodic lattices, they also have phason degrees of freedom, which in this case concerns the phase between the two sublattices, which changes when one sublattice slides against the other, leading to "sliding modes," as predicted by the DCM. Whether or not these sliding modes have been observed experimentally is still under debate. As in the case of modulated crystals, these simplified descriptions are only valid as long as the interaction between the mutually incommensurate systems is not too strong. It is very likely that phasons are primarily "diffusive" in incommensurate composites, but considerably less experimental information exists on these systems than on the incommensurably modulated phases discussed above.

Quasicrystals

1D, 2D, and 3D QCs have been found, of which the 1D QCs consist of an aperiodic stacking (e.g., in a Fibonaci sequence) of periodic lattice planes. 2D QCs are made up of lattice planes with a quasiperiodic structure with pentagonal, octagonal, decagonal, or dodecagonal point symmetry, which are stacked periodically. Finally, 3D quasiperiodic structures have aperiodic lattices with icosahedral point symmetry.

The atomic structure of 2D and 3D QCs is based on "atomic clusters" of 1–2 nm diameter. In 2D QCs, these are cylindrical clusters made up of a stacking of the quasiperiodic lattice planes, neighboring planes often being rotated against one another, often with some "filling" atoms between the clusters. The atomic structure is dense. 3D QCs, the icosahedral ones, can be roughly divided into two classes on the basis of the dominant type of cluster characteristic of their structure: Mackay clusters and Bergman clusters, where the latter are also characteristic of the Frank–Kasper type of periodic crystals with many atoms per unit cell. As 2D QCs are often "relatives"

of the icosahedral phases, here these QCs are put in the same category as their icosahedral neighbors. Because the atomic dynamics is dominated by the type of atomic interaction (interaction potential) and the "local" structure, one expects that the type of cluster, which dominates the local environment of the vibrating atom, will influence the phonon dynamics.

Concerning the phasonic motions (displacements in internal space), one distinguishes two possible phasonic modes (apart from the static-phason strain): "phason flips," which only cause a local change in the phase of the aperiodic lattice by a (mostly collective) jump of a few neighboring atoms, leading to a "violation" of the matching rules of the perfect quasiperiodic lattice. These have to be distinguished from "phasons," which are collective diffusive modes. These are different from the incommensurate modulated phases due to the fact that QCs always have discontinuous aperiodic modulation because of the discrete bounded atomic surfaces, which cause jumps or discontinuities in the modulation function, when moving in internal space.

Of these structures, up to now, the atomic dynamics of mainly icosahedral (i-) and decagonal (d-) QCs has been investigated using NIS, XIS, and X-ray intensity correlation techniques. On single crystals, using crystal spectrometers, dispersion curves for acoustic phonons and some low-lying optic branches have been measured. On polycrystalline samples, bands at intermediate energies and the GVDOS have been determined by time-of-flight (TOF) spectroscopy for several metastable and nearly all stable quasicrystalline samples from the total dynamic structure factor, which one measures in all these cases. For i-AlCuFe also, the Fe partial density-of-state was measured using XIS and ^{57}Fe substitution. Quasielastic neutron scattering (QENS) was used on polycrystalline samples and an AlPdMn single crystal to measure phason flips and X-ray diffraction, and correlation techniques to investigate diffusing phasons in QCs. In the following, results from these four types of investigations will be discussed starting with the vibrational degrees of freedom in physical space.

Phonon Dispersions

Very characteristic for all phonon branches investigated so far is the important width of the phonon peaks in the measured spectra as soon as the modulus q of the wave vector $\boldsymbol{q}$ of the measured excitation exceeds a certain critical value, which essentially in all cases is ~ 3–$4\,\mathrm{nm}^{-1}$. The broadening of the peaks is of such an importance that the rapidly increasing width prevented measurements of excitations above energies of ~ 6–8 THz, even though the vibrational density-of-states (VDOS) often extends to twice that energy.

As an example, **Figure 3** shows some of the transverse acoustic branches measured with an AlPdMn single crystal along the twofold direction, the Γ point ($q = 0$, starting point of the dispersion) being one of the strongest Bragg peaks in this crystallographic direction. As the intensity of the inelastic structure factor is proportional to the intensity of the Bragg peak at the Γ point, and only next to a strong Bragg peak the PBZ is well defined, this choice is essential for the result of the experiment.

Figure 3 already demonstrates this very characteristic property of the lattice dynamics of QCs: the rapid damping of the lattice vibrations as soon as the wavelength of the excitation is smaller than ~ 2 nm, which in several cases is just the approximate diameter of the building unit. Corresponding results have also been obtained by computer simulations of icosahedral QCs, which show above the acoustic dispersion branches, rather dispersionless broadbands of optic modes, as they have also been found experimentally in some cases, and at even higher energies quasicontinuous spectra, which cannot be resolved in single excitations any more. This fact has limited so far the investigation of single excitations to energies below 8 THz, that is, to about half the upper energy limit of all vibrational excitations.

As is shown in **Figure 4**, for icosahedral QCs one finds "isotropic" dispersion branches for longitudinal and for transverse phonons, as these crystals, due to the high symmetry of the icosahedral structure, have – in the hydrodynamic limit – only two phonon elastic constants (plus two phason elastic constants and a phonon–phason one). This has also been proven experimentally for the sound velocity. In contrast, the sound "absorption" is anisotropic because of phonon–phason coupling.

For 2D QCs, computer simulations had predicted some anisotropies concerning lattice modes with polarization vectors in the aperiodic plane and in the periodic direction, respectively. However, experimentally the differences observed for d-AlCoNi (Mackay-type) single crystals using NIS are much less important and not far outside the error limits of the experimental results. Likewise, the gap structure of the dispersions, predicted for aperiodic lattices especially in low dimensions, has not clearly been confirmed experimentally up to now.

Above the acoustic branches, very few dispersionless branches of very broad excitations have been measured by NIS on single crystals by triple axis spectrometry (TAS) and extracted from $S(Q, \omega)$ measured on TOF spectrometers. Both results are

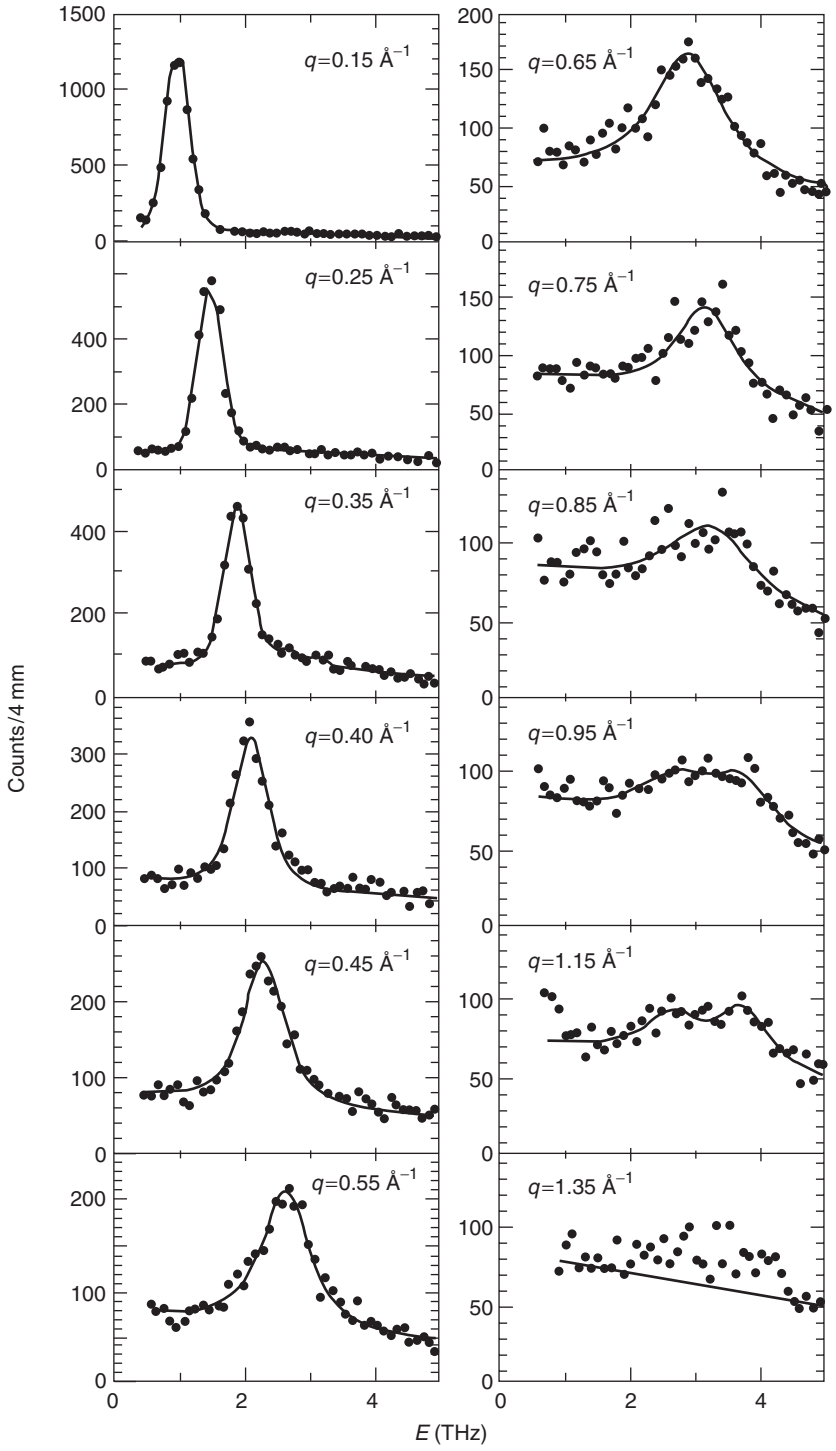


Figure 3 Spectra of transverse phonons measured by NIS at a strong Bragg peak in the twofold direction of an i-AlPdMn single crystal (Mackay type). The rapid increase of the phonon line width for $q > 4\,\mathrm{nm}^{-1}$ is obvious. The line width increases roughly proportional to q^2.

in good agreement. These excitations have been suggested to have the character of optic modes.

The phonon dispersions of QCs have been investigated in the accessible energy range (up to 4 THz in all cases, up to 6 THz for few examples, and up to 8 THz in one case) for most of the thermodynamically stable QCs using single crystals and mostly NIS and, in some cases, also XIS.

Generalized Vibrational Density-of-States

The energy range above 8 THz could be accessed by NIS up to now using only polycrystalline samples and

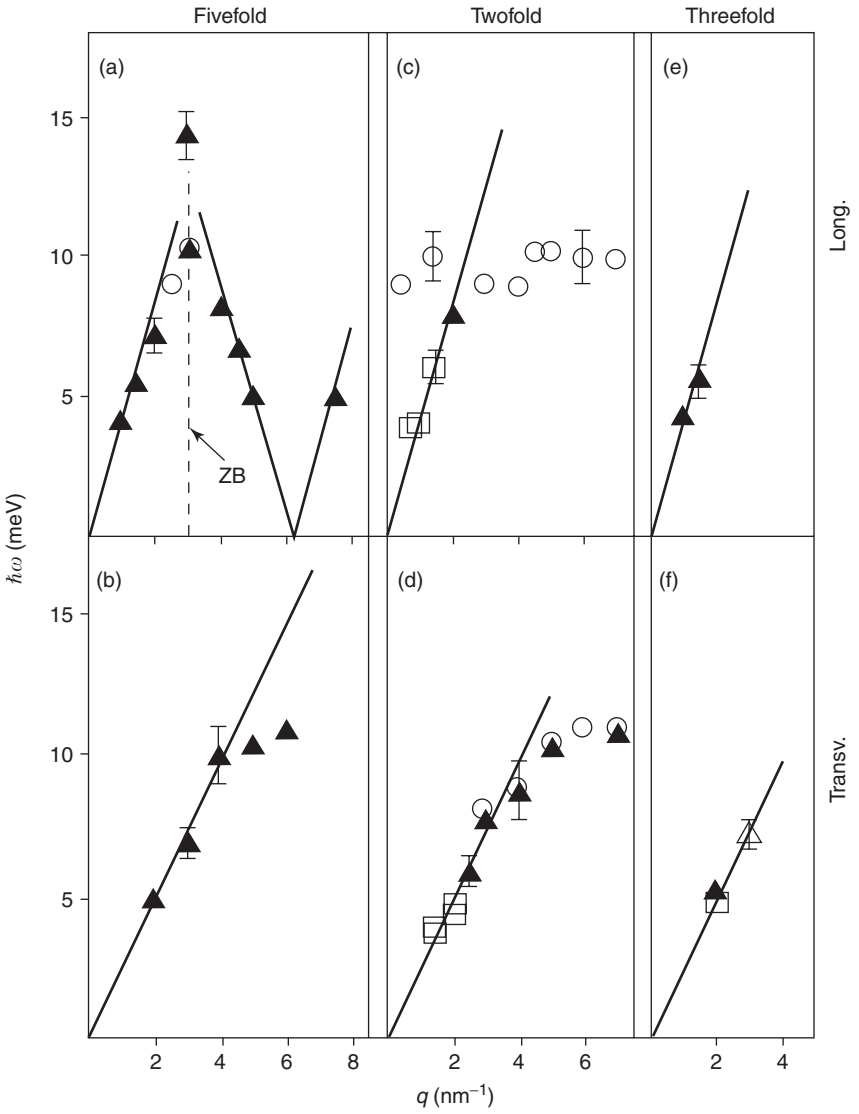


Figure 4 Dispersion of longitudinal and transverse phonons in i-AlLiCu (Frank–Kasper type) proving the isotropy of each of these excitations with respect to the three different crystallographic axes (twofold, threefold, and fivefold).

TOF spectroscopy. In this case, one obtains angular-averaged (in reciprocal space) dispersion relations from the shape of the total dynamic structure factor and the GVDOS from its intensity averaged in a reasonably large region of reciprocal space. In order to cover the full energy range, in what follows, the discussion will therefore concentrate more on the GVDOS, which have been determined for most of the thermodynamically stable and several metastable QCs. The GVDOS normally contains several energy bands, the position (energy of the center of gravity of the band) and intensity of which are usually different for different samples. Thus, it is mainly the shape (and the energy range and the slope at lowest energy (if resolved)) of the GVDOS, which can be used to characterize the atomic dynamics of a sample via the GVDOS.

One obtains considerably more detailed information, if the experimental results are combined with results from computer simulations (CSs). These simulations are based on interactions, which are either calculated directly during the simulation, as is done in the method due to Car and Parrinello, or they are modeled in a more or less realistic manner. The second basic input is a model for the atomic structure, which in the case of QCs is always a structural model for a rational approximant, because of the finite size of the computer model, and it is of similar uncertainty as is the complicated interaction. Because of these inherent uncertainties, the comparison of the results of CSs with those of experiments is "essential" for a judgment on the quality of the simulation results, that is, how realistic these results are. If they

are in fairly good agreement, nothing is proved; however, it seems reasonable to use them for the interpretation of the measured results, as they provide very detailed information, which otherwise is possible only with great effort and extensive cost of experiments. In what follows, therefore, the experimental results will be confronted with those from realistic CSs as far as they exist.

Atomic dynamics and cluster structure As mentioned above, QCs are built from different types of clusters, the structure of which should influence the dynamics as it is the local structure which has the strongest influence on the dynamics. In fact, one finds a difference in the shape of the corresponding GVDOS for the two types of QCs. QCs dominated by more or less complete Mackay icosahedra, such as i-AlPdMn, i-AlCuFe, and d-AlCoNi, seem to have very smooth, nearly structureless one- or two-band spectra, whereas QCs of the Frank–Kasper type, such as i-AlLiCu, i-ZnMGRE (RE = Ho, Er,…, and also Y), and d-ZnMgY have more structured spectra with several energy subbands. This is easily seen in a comparison of **Figure 5** with **Figure 7**.

CSs suggest that this stronger band structure is due to a stronger grouping of the local vibrational density-of-states (LVDOS), that is, the vibrational spectra of all atoms on different Wyckoff positions in this computer-approximant to the Frank–Kasper type of QCs. This is different in the case of the Mackay type of QCs, where the LVDOS are more different from one another and result in a broad smooth spectrum after summation.

Quasicrystals and their rational approximants Because one of the basic input parameters for CSs, the structural model, is based on considerable approximations in the case of QCs, quite a few simulations have also been done on periodic crystals with similar chemical compositions as the QCs have, but with a well-known structure. Compared to the atomic dynamics of QCs, the dynamics of these periodic crystals is rather similar to that of the corresponding aperiodic phase if the periodic crystal is an approximant to the QCs. For example, the GVDOS of AlLiCu in the i- (Fank–Kasper type) and the periodic R-phase (built up of the same kind of clusters) are very similar to each other, having energy bands at essentially the same energies, but with slightly different intensities. The same results have been obtained in the measurement of the phonon dispersions of the same alloy system, where also hardly any difference could be detected except for a dispersionless branch at the upper energy limit of this investigation in the case of the i-phase seen in the center of **Figure 4**. Likewise, CSs of this alloy system, providing the information on the partial (element specific) dynamical properties, showed a small difference in peak intensity but not in peak position in the calculated GVDOS of the i- and R-phase, respectively. Even more identical results have been obtained for the Mackay type of QCs d–$Al_{71.5}Co_{13.5}Ni_{15}$ and its modulated approximant $Al_{71.5}Co_{15.5}Ni_{13}$. In fact, phonon dispersion measurements done on both types of QCs suggest slightly narrower phonon groups for the rational approximant compared with the QCs in the case of the Frank–Kasper type but not in the case of QCs of the Mackay type. These nearly identical results can be understood from the fact that the atomic dynamics is most strongly determined by the local (cluster) structure and potential, which seem to be sufficiently similar for the QCs and their approximants, while their structure differs on the long-range scale. For the 2D QCs just mentioned, this striking similarity of the atomic dynamics was even found at 1000 K, which also means that the thermal expansion (see below) of the QC and its rational approximant are nearly identical. All these conclusions do not apply to periodic crystals of similar chemical composition as the QCs have, but to those which are (structurally) not approximants.

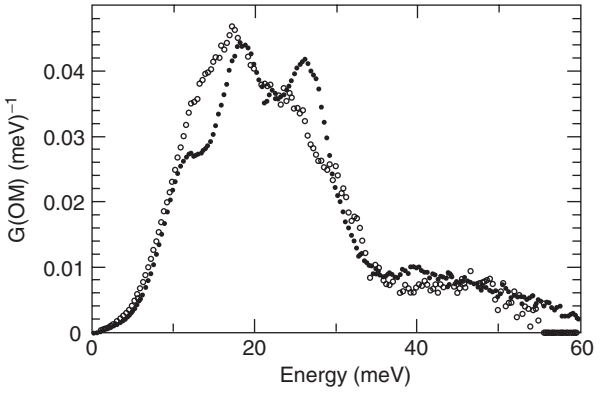


Figure 5 GVDOS of i-$Zn_{63}Mg_{26}.3gY_{10.7}$ and d-$Zn_{58}Mg_{40}Y_2$ (both Frank–Kasper type) showing the highly structured distributions with several energy bands and subbands.

Low-energy modes in quasicrystals Within the harmonic approximation, which should also apply to QCs at not too high temperatures, at low energies, one expects to find the linear dispersions $\omega = vq$, of transverse (TA) and longitudinal (LA) acoustic phonons, which join the linear dispersions of transverse and longitudinal ultrasound near $q = 0$. Here v is the sound velocity of the mode under consideration ($v = v_T$ *or* v_L) and q is the corresponding wave vector of the mode. On the basis of the Debye approximation, one therefore expects the phonon

density-of-states (PDOS) to start at the lowest energies proportional to $\omega^{(d-1)}$, where d is the dimension of the sample under consideration. For a 3D system with particle density n and a weighted mean value of the sound velocity v, one therefore expects the PDOS to start at the lowest energies proportional to the Debye DOS:

$$G_{\mathrm{Deb}}(\omega) = \frac{\omega^2}{2\pi n v^3} \qquad [6]$$

$$\frac{3}{v^3} = \frac{1}{v_{\mathrm{L}}^3} + \frac{2}{v_{\mathrm{T}}^3} \qquad [7]$$

For several – but not all – QCs (and some of their rational approximants), this general rule does not apply either for the measured GVDOS or for the CSs calculated PDOS. One finds a considerably higher intensity in the low-energy region compared to the intensity calculated by using the known sound velocities. Corresponding higher values for these QCs have also been found for the low-temperature specific heat by different groups. Likewise, the temperature dependence of the sound velocity of these QCs shows an unexpected behavior. Investigations of this kind are only possible with thermodynamically stable QCs as normal production defects have to be annealed out before.

Metastable QCs obtained by annealing of a metallic glass show additional low-energy modes stemming from the large grain boundary volume in these samples, as annealing here leads to nanocrystalline samples as has been shown for i-PdSiU and AlCuV. In this case, the intensity of the low-energy modes reaches nearly the level of a well-relaxed metallic glass.

Temperature dependence of the atomic dynamics The temperature dependence of the dynamics outside the quasielastic region has been investigated most extensively for i- and d-QCs of the Macky type. In all cases, one finds a considerable shift of the GVDOS to lower energies, that is, an anharmonic behavior of the DOS. However, on quantifying this shift with temperature, for example, with the "integrated" amount of modes shifted (e.g., integration from 0 to 16 or 25 meV) into the low-energy region, one finds a linear dependence of this transfer of modes with temperature up to 1000 K in all investigations, for d-AlCoNi even in the full range of temperatures investigated (up to 1100 K). This linear increase of the shifted intensity is parallel to the shift of the Debye–Scherrer lines in the diffraction diagrams to lower momentum transfers with increasing temperature, caused by the expansion of the lattice. Investigating the shift of lines having contributions only from the periodic axis or only from the aperiodic lattice planes, one observes, for example, in the case of d–$Al_{71.5}Co_{13.5}Ni_{15}$ that at temperatures above 750 K, the thermal expansion of the QCs in the periodic direction is stronger with increasing T than in the aperiodic plane.

Investigations of the energy dependence of the shifted intensity (by how much has one to shift the intensity of the calculated PDOS to arrive at the experimentally determined GVDOS) show a nearly constant energy shift of $\delta\omega/\omega$ of $\sim$6–10% at all temperatures analyzed from the upper limit of the GVDOS down to $\sim$12 meV. After this, one finds a strong increase of the shifted intensity, becoming more important with increasing temperature and reaching a relative shift of 20% and more. The participation ratio P ($P(\omega) = 1$ if all atoms of the model take part in the mode at frequency ω) shows, in the case of a QC or its approximant, that the smallest participation, that is, the strongest localization of the vibrations takes place at the smallest and highest frequencies. This is shown in **Figure 6**. Correlating this relative frequency shift $\delta\omega/\omega$ with the participation ratio P for the modes at these energies, CSs show that the strongest shifts correlate with the lowest participation ratio. Thus, one has to conclude that a large part of the shifted modes are most likely localized modes. The same CSs clearly demonstrate that the Al atoms contribute more strongly to these shifts than do the TM atoms, as one would expect from the structure, where the TM atoms form a rigid network in the clusters.

For i-AlPdMn, a new phenomenon comes in at temperatures above 1000 K. While one finds essentially the same linear behavior for the integrated intensity and shift of the Debye–Scherrer peaks, as for d-AlCoNi up to 1000 K, the integrated intensity suddenly increases well above this linear continuation (see **Figure 7**).

A careful investigation of the total dynamic structure factor of i-AlPdMn in the low-energy region shows that this additional intensity is due to quasielastic scattering being caused most likely by phason flips (see below), that is, a new degree of freedom in addition to the phonon degrees of freedom. As this quasielastic intensity gets considerably broader with increasing wave vector transfers Q, it overlaps with the inelastic region of the spectra and thus contributes to the intensity used to determine the GVDOS.

Phason flips As the n-dim space corresponding to the structure of the QCs is a "periodic lattice" decorated by disconnected atomic surfaces, any movement of the

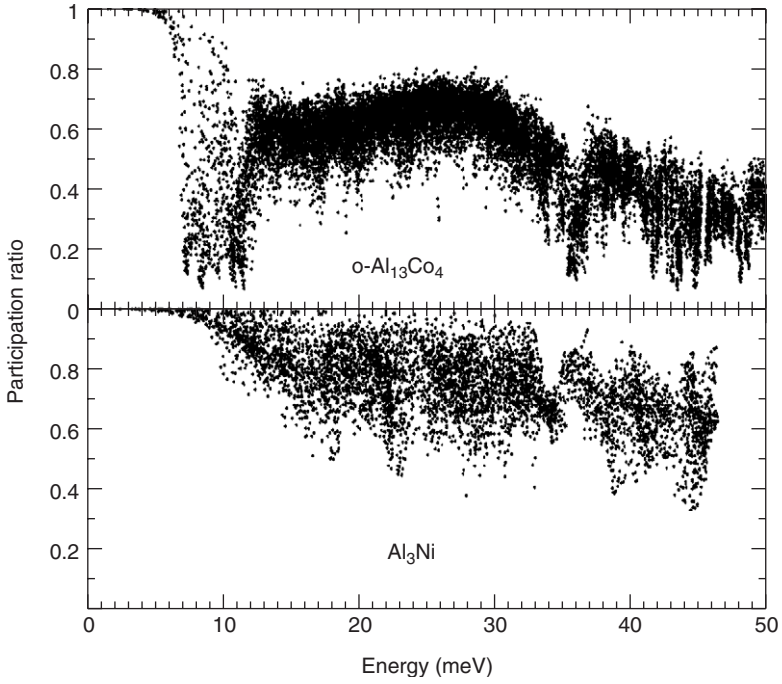


Figure 6 Participation ratio as a function of the energy of the vibrations for the orthorhombic approximant o-$Al_{13}Co_4$ to quasicrystalline d-$Al_{71.5}Co_{13.5}Ni_{15}$ in comparison with the chemically comparable but nonapproximant Al_3Ni. The strong decrease of the participation (= localization) of the modes with $\omega < 12$ meV in the case of the approximant is clearly seen.

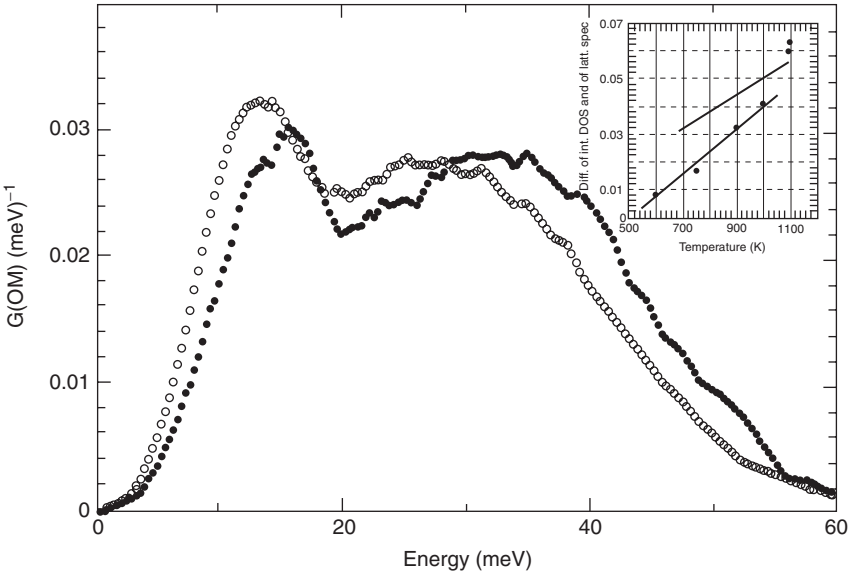


Figure 7 GVDOS of i-$Al_{63}Pd_{26.3}Mn_{10}$ (Mackay type) showing the important shift of the GVDOS to lower energies with increasing temperature. The insert demonstrates that the integrated shifted intensity has a linear behavior nearly parallel to the linear shift of the Debye–Scherrer lines in the diffraction diagrams (expansion of the alloy), except for temperatures higher than 1000 K.

physical space, with which the *n*-dim space is intersected, orthogonal to itself will lead to sudden jumps in the atomic positions instead of a continuous change as in the case of a continuous modulation function such as in perfect incommensurably modulated systems. Phason degrees of freedom will, therefore, always concern either flips of atoms in the aperiodic structure, corresponding to a local deviation of the intersection or by a diffusion kind of motion on a much longer timescale.

On the basis of calculations, one expects the phason dynamics to be frozen-in at lower temperatures even in random tilings, where phasons are principally in an unlocked state – in contrast to the locked state in energetically stabilized QCs below the unlocking transition. Only at temperatures well above room temperature, one could expect unlocking and with this, any sign from phason dynamics. Starting from a perfect (decorated) tiling at lower temperatures, phason flips will violate the matching rules and lead to a random tiling, which will be stabilized by its configurational entropy. The reverse mechanism will most likely be incomplete due to kinematic constraints. This process has been studied in several CSs, and at least in one of these investigations, correlated flips of several atoms have been observed.

Experimentally, quasielastic scattering has been found for most of the thermodynamically stable QCs by different groups at temperatures near the upper limit of the existence region of the QCs using polycrystalline samples and TOF spectroscopy. In the case of i-AlCuFe, for example, one obtains a nearly Q-independent width Γ of the Lorentzian fitted to the quasielastic spectrum, Γ being ~ 0.1 meV. This width of the spectrum would then correspond to jumps in the region of $\hbar/\Gamma \approx 7$ ps. The Q-dependence of the integrated intensity of this quasielastic signal with a maximum at $10\,\mathrm{nm}^{-1}$ suggests a process localized within a jump distance of 0.4 nm. The most detailed investigation was done at 1073 K using a single crystal of i-AlPdMn on a cold neutron TAS. While the Q-dependence of the quasielastic scattering from polycrystalline samples only allows one to determine the jump distances of the scattering atoms, experiments on single crystals allow, in addition, one to determine the orientation of the jumps. The variation of the measured intensity with the rotational angle in reciprocal space can be best fitted by a model for quasielastic scattering from simultaneous correlated jumps of several atoms in the phason flips. The quasielastic spectra can be fitted with one broad Lorentzian with $\Gamma = 0.7$ meV or (better) with two Lorentzians with $\Gamma = 0.2$ and 0.7 meV. Measurements on the twofold axis now show a maximum of the intensity near $27\,\mathrm{nm}^{-1}$, which corresponds to considerably shorter jump distances than found in the less detailed investigation of i-AlCuFe.

Phason fluctuation in quasicrystals For reasons mentioned in the previous section, one does not expect to find propagating phason modes as in the case of modulated phases. However, phasons do play a role in the elastic and dynamic properties of QCs. The elastic-free energy of a QC is given by three different terms:

$$F = F_{\mathrm{phon}}(\lambda, \mu) + F_{\mathrm{phas}}(K_1, K_2) + F_{\mathrm{cross}}(K_3) \quad [8]$$

Here λ and μ are the Lame coefficients, K_1 and K_2 the phason elastic constants, and K_3 the phonon–phason coupling term. From investigations of the diffuse scattering, assuming $K_3 = 0$, one can obtain the ratio of K_2/K_1, which at 300 K was found to be near -0.5 for i-AlPdMn. At temperatures in the range of 1000 K, a diffusive relaxation of phason strain or "phason walls" (left behind a moving dislocation in a QC) can lead to a collective diffusive phasonic excitation or long wavelength phason fluctuations. According to different theoretical investigations performed by several authors, these fluctuations will contribute to the Debye–Waller factor and to the diffuse scattering in a similar way as do phonons and defects, with the exception that the timescales of phonon and phason fluctuations differ by many orders of magnitude, lattice vibrations being in the picosecond range, while the characteristic times for long wavelength phason fluctuations is more between 0.1 and 100 s. Correspondingly, phonons can follow a lowering of the temperature instantaneously, whereas phason fluctuations fall out of equilibrium as a consequence of fast T-changes.

Experimentally, phason diffusion has been investigated in reciprocal space using high-resolution X-ray diffraction to study the phonon and the phason contribution to the Debye–Waller factor in a single crystal of i-AlCuFe in the temperature range between 300 and 1023 K, and to study diffuse scattering from a single crystal of i-AlPdMn as a function of temperature between 300 and 1043 K. Finally, X-ray intensity fluctuation spectroscopy (XIFS) has been used to investigate the time and temperature dependence of the diffuse scattering from a single crystal of i-AlPdMn at 923 K. In real space, *in situ* HRTEM was used to investigate the time-dependent structural changes in d-AlCoCu at 1123 K.

The X-ray diffraction experiment demonstrates a very strong temperature dependence of the phason contribution to the Debye–Waller factor (change by a factor of 9 between 750 and 650 K) and to the diffuse scattering, which decreases with increasing T while the intensity of large Q_{perp} reflections increases. Freezing of the phason dynamics was observed just below 900 K already. The most detailed results were obtained using highly coherent synchrotron radiation in the XIFS experiment. Here, the time evolution of the speckle pattern from diffuse scattering could be directly resolved and a time-dependent intensity correlation $g(Q, t)$ could be determined from the time dependence of the

measured intensity:

$$F_{cor}(Q,t) = [1 - \beta g(Q,t)] \quad [9]$$

where β corrects for the partial coherence of the beam. Above 923 K, the expected exponential decay $g(Q,t) = \exp(-t/\tau(Q))$ was found with $\tau(Q)$ proportional to $1/Q^2$, that is, D_{phas} proportional to λ^2, which proves the diffusive character of the long wavelength phason fluctuation observed, as the width Γ of the spectrum should be $\Gamma = DQ^2 = D4\pi^2/\tau$ for diffusive dynamics. D_{phas} was found to be $1.5 \times 10^{-16}\,m^2\,s^{-1}$ with the phason excitation wavelength λ between 50 and 100 nm. Again, a very strong temperature dependence was found, τ varying between 60 and 300 s when lowering the temperature from 923 to 873 K.

A comparable slow phason dynamics was found in the *in situ* HRTEM investigations of single crystals of d-AlCoCu at 1123 K. The duration for a change of the position of one of the columnar clusters of the decagonal structure was smaller than 10 s, the time between jumps was between 0.1 s and 10 min.

See also: Lattice Dynamics: Vibrational Modes; Point Groups; Quasicrystals.

PACS: 64.70.Rh; 61.44.+p; 63.20.Dj; 61.12.Ex; 61.10.My

Further Reading

Currat R and Janssen T (1988) Excitations in incommensurate crystal phases. *Solid State Physics* 41: 202.

Elhor H, Mihalkovic M, Rouijaa M, Scheffer M, and Suck J-B (2003) Dynamical properties of quasicrystalline alloys investigated by neutron inelastic scattering and computes simulation based on realistic potentials. In: Trebin H-R (ed.) *Quasicrystals*. 382p. Wiley-VCH.

Francoual S, Livet F, de Boissieu M, Yakhou F, Bley F, *et al.* (2003) Phason fluctuations in the AlPdMn quasicrystal. *Physical Review Letters* 91: 225–501.

Janssen T, Radulescu O, and Rubtsov AN (2002) Phasons, sliding modes and friction. *European Journal of Physics* B29: 85.

Quilichini M and Janssen T (1997) Phonon excitations in quasicrystals. *Reviews of Modern Physics* 69: 277.

Radulescu O, Janssen T, and Etrillard J (2002) Dynamics of modulated and composite aperiodic crystals. *European Journal of Physics* 385.

Lattice Dynamics: Elasticity *See* Lattice Dynamics: Vibrational Modes.

Lattice Dynamics: Structural Instability and Soft Modes

R A Cowley, University of Oxford, Oxford, UK

Introduction

The description of structural instabilities of crystals is both topical and goes back to the early days of lattice dynamics and the ideas of Born and Raman. The subject became of importance after 1960 when Cochran developed the concept of a soft mode and made specific predictions that could be experimentally verified. Born had shown that a crystal was harmonically stable if all the frequencies squared of the normal modes were positive, $\omega(qj)^2 > 0$, where q is the wave vector of the normal mode within the first Brillouin zone and j is the branch index. Clearly, if the crystal has long-range order then this result must be satisfied. Cochran discussed the case when the frequency squared was temperature dependent and made the simplest possible assumption that:

$$\omega(qj)^2 = A(T - T_c)$$

At temperatures above T_c, the frequency squared is positive and the crystal is stable, whereas below T_c, it is negative and the crystal distorts to a new crystal structure so that the frequency is no longer negative. This is the basis behind the concept of a soft mode, as a mechanical instability of the crystal against a particular normal mode.

The Classical Era

Above T_c, neutron scattering is usually the best way of studying the soft modes particularly if the wave vector of the soft mode is not at the center of the Brillouin zone. In the special case of ferrodistortive modes, the wave vector transfer is zero and studies with infrared or Raman spectroscopy may also be possible. The most studied material is $SrTiO_3$ and its structural phase transition at 105 K is illustrated in **Figure 1**. The crystal structure in the high temperature phase is cubic and in the low-temperature phase is distorted, as shown in **Figure 1**, by a rotation of the oxygen octahedra by the angle, ϕ. There are three degenerate soft modes corresponding to rotations about the three (1 0 0) axes. **Figure 2** shows the results of measurements of the soft mode frequencies. Above T_c the triply degenerate soft mode frequency squared decreases to zero in a linear fashion with all three degenerate modes. Then at T_c the degeneracy is broken and one mode increases rapidly with

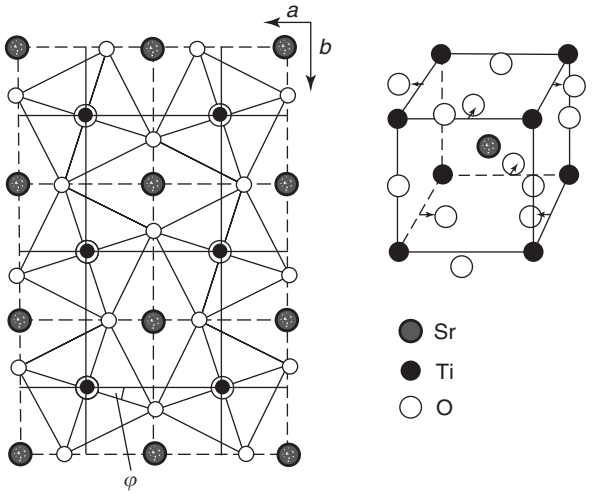


Figure 1 A projection down the cube axis of the distorted phase of $SrTiO_3$. (Reproduced from Unoki H and Sakudo T (1967) Electron spin resonance of Fe^{3+} in $SrTiO_3$ with special reference to the 110 K phase transition. *Journal of the Physical Society of Japan* 23: 546.)

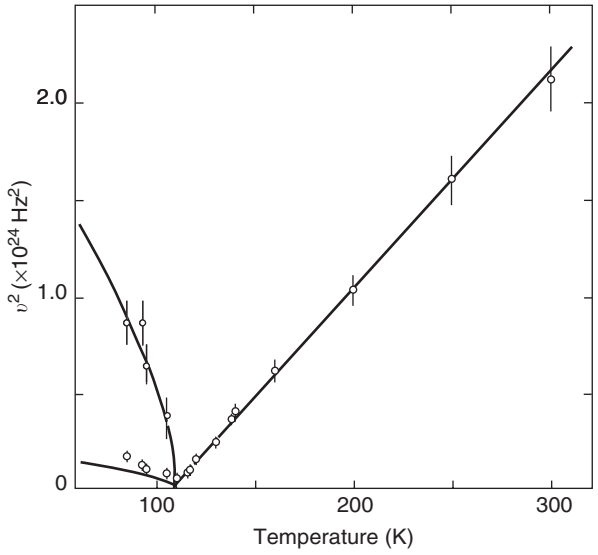


Figure 2 The temperature dependence of the frequencies of the soft mode in $SrTiO_3$, both above and below the phase transition. (Reproduced from Cowley *et al.* (1969) *Solid State Communications* 7: 181.)

decreasing temperature while the other two also increase in frequency but less rapidly.

This type of behavior above T_c has now been observed in many materials such as $BaTiO_3$, KDP, $Pb_5Ge_3O_{11}$, and others. Unfortunately, if T_c becomes comparable to typical phonon energies then the soft modes are often overdamped and it may be difficult to observe them directly.

Below T_c, the situation is more complex and depends on the degeneracy of the soft mode, j. For example consider $SrTiO_3$, there are three soft modes corresponding to the rotations of the octahedra about the x, y, or z axes and the free energy can be written as a power series in the amplitudes of the modes Q_x etc. The result is

$$F = F_0 + a(T - T_c)(Q_x^2 + Q_y^2 + Q_z^2) + d_1(Q_x^2 + Q_y^2 + Q_z^2)^2 + d_2(Q_x^4 + Q_y^4 + Q_z^4)$$

where the constants a, d_1, and d_2 are temperature-independent constants. This is a Landau expansion for the free energy of the soft mode and for $T < T_c$, the structure depends on the sign of d_2. If $d_2 > 0$ then the distorted structure has equal contributions of the x, y, and z components, as in $LaAlO_3$, whereas if $d_2 < 0$ then the structure is distorted along one of the principal axes (1, 0, 0), as in $SrTiO_3$.

Raman scattering is always a suitable tool to investigate the soft modes below T_c because at least one soft mode must have the full symmetry of the crystal and this is necessarily a Raman active mode. Nevertheless in most cases the intensity also drops to zero at T_c and that makes a precise determination of the temperature dependence difficult. Well away from T_c the higher order terms in the expansion can also play an important role so that it is difficult to test the Landau theory exactly using Raman scattering.

For some soft modes the wave vector is within the Brillouin zone and incommensurate. The theory then proceeds as discussed above, except that in the low-temperature phase the equivalent term in the free energy to d_2 is zero by symmetry. In this case one of the modes below T_c has zero frequency and it is an excellent example of a Goldstone boson that arises from the broken continuous symmetry in the direction of the order parameter. Phason is the name given to the dispersion curve of this mode and its frequency is proportional to the difference in $\boldsymbol{q}$ between the soft mode and the mode being observed. Unfortunately, there are relatively few examples of well-defined phasons that have been observed experimentally but one example is $ThBr_4$.

In some cases the phase transitions are somewhat different when the soft mode is an acoustic mode of the crystal. The acoustic modes are described by the elastic constants and their energies are proportional to the wave vector. The soft mode behavior then occurs if the slope of the acoustic modes goes to zero. Very often the symmetry of the acoustic mode is such that the soft mode undergoes a first-order transition because the free energy has a cubic invariant. An example of this is the behavior of Ce metal that undergoes a volume change when it changes from one valence state to another. In other cases, the phase

transition may be of first order as found, for example, in KDP where the ferroelectric fluctuations couple linearly with the C_{66} acoustic mode giving rise to a continuous phase transition.

Despite the undoubted success of the soft mode concept there have been relatively few examples where there is any understanding of the parameters in the Landau theory. In the case of order–disorder transitions in materials like NH_4Br, the ordering of the NH_4 molecules between the two different states is analogous to a spin vector and some progress has been made in calculating the transition temperature and determining whether the transition is ferrodistortive or antiferrodistortive. It has been suggested that the transition in KDP can be explained by the transverse-Ising model because there is a substantial difference between the transition temperatures of the hydrogenated and deuterated materials (**Figure 3**). This has however been questioned because all the atoms are involved in the ferroelectric fluctuations and the difference in the transition temperature probably arises from the change in the force constants. The most carefully studied material is $SrTiO_3$ where the temperature dependence of the soft mode arises from the anharmonicity of the lattice dynamics. Anharmonic lattice dynamics predicts that the self-energy of the normal modes depends on the frequency with which the modes are studied. Nevertheless for most materials the self-energy is only slowly varying with the applied frequency and the corrections to be applied to a temperature-dependent mode are small. In the case of $SrTiO_3$ calculations were made of the terms in the Landau expansion and other anharmonic properties and reasonable agreement was obtained with fairly realistic interatomic forces. The question then remains as to why the anharmonic effects have a much larger effect on the properties of $SrTiO_3$ than they do on, say, KBr, for example. In the high-temperature phase there is one formula unit in each cubic unit cell so that there are three ionic radii and only one lattice parameter that can be adjusted. It is therefore only good fortune if the three different ionic radii exactly fit into the available space and, in practice, the Sr and Ti ions are somewhat too large for the O ion. As a result the O ions tend to displace away from the center of the bond and the crystal structure becomes distorted. At high temperatures the motion of the atoms increases their size and there is no distortion but at low temperatures the crystal distorts to a noncubic phase. The anharmonic effects are basically no bigger than is normally the case. The unusual aspect of $SrTiO_3$ is that the harmonic frequencies are unusually low (indeed the squares of the harmonic frequencies are negative) whereas the anharmonic effects are much the same size as in other materials.

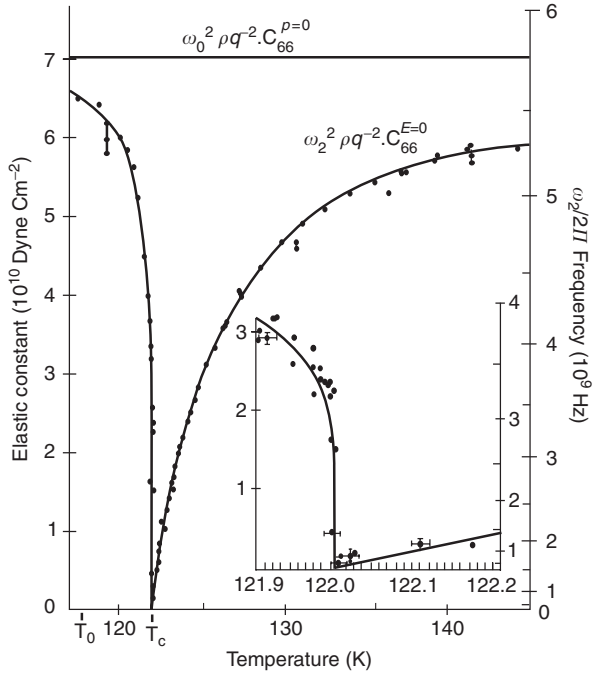


Figure 3 The elastic constants of KDP, C_{66}^E, and C_{66}^P as function of temperature. (Reproduced permission from Brody E and Cummins H (1968) Brillouin-scattering study of the ferroelectric transition in KH_2PO_4. *Physical Review Letters* 21: 1263.)

Critical Era

The theory that is outlined here has been a Landau theory that does not take account of the critical fluctuations that occur close to T_c, while the experiments have not looked carefully at the behavior close to T_c. The mean field theory neglects the interactions between the soft mode and those modes with the same branch index but with slightly different wave vectors. When these interactions are included the susceptibility and order parameter have the following forms:

$$\chi_{\pm}(T - T_c) = C|T - T_c|^{-\gamma}$$

while the order parameter is

$$\langle Q \rangle = B(T_c - T)^{\beta}$$

The exponents, β and γ, are then often very different from those predicted by mean field theory with typically γ about 1.35 and β about 0.33 for systems in three dimensions and having conventional behavior. It is generally expected that measurements of these nonclassical exponents require the reduced temperature $t = |T - T_c|/T_c$ to be less than 0.1 and to cover 2 or 3 decades of temperature. In practice it has proved surprisingly difficult to satisfy these requirements.

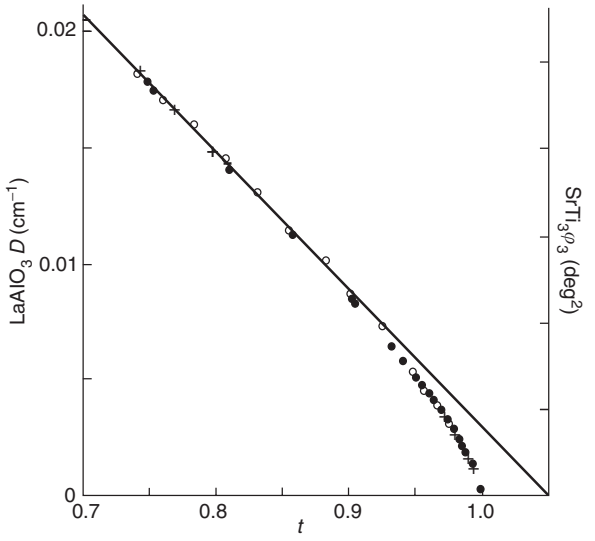


Figure 4 The temperature dependence of the order parameter of $SrTiO_3$ and $LaAlO_3$ as measured by EPR techniques. (Reproduced from Muller KA and Berlinger W (1971) Static critical exponents at structural phase transitions. *Physical Review Letters* 26: 13.)

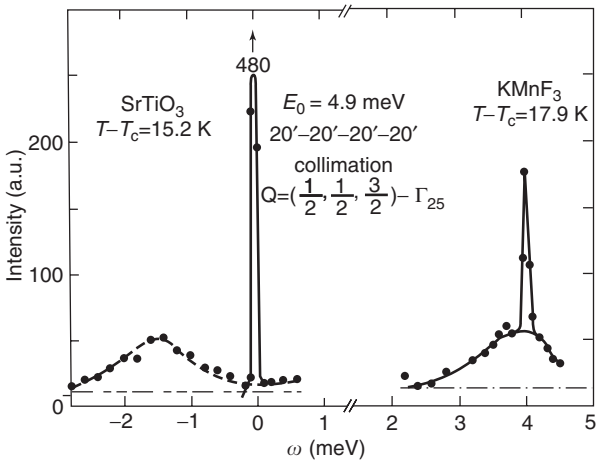


Figure 5 The scattering from the soft mode in $SrTiO_3$ and $KMnF_3$ showing the quasi-elastic scattering and the soft mode. (Reproduced from Shapiros S, *et al.* (1972) Critical neutron scattering in $SrTiO_3$ and $KMnF_3$. *Physical Review B* 6: 4332.)

The first clear evidence for nonclassical exponents came from measurements of the critical exponents of the order parameter in lightly doped Fe in $SrTiO_3$ by Muller and Berlinger, as shown in **Figure 4**. The results clearly show that the temperature dependence of the order parameter is different from classical theory and that $\beta = 0.33 \pm 0.02$. Other types of measurements have been much less successful. Raman scattering has been used below T_c but does not give reliable measurements close to the phase transition because it follows an energy singularity and is difficult to measure. Diffraction methods have been used for T above T_c but for reasons to be explained later satisfactory results have not been obtained. Many experiments have now followed the first X-ray measurements by Andrews. They have given very conflicting results that are dependent on the X-ray energy, and precise conditions for the experiment. Only one experiment has studied the incommensurate phase transition in Rb_2ZnCl_4 and that gives good agreement with the expected critical exponents.

The basic assumption of the critical theories is that close to the phase transition, the behavior is described by a single length-scale and a single timescale and that both diverge as the phase transition is approached. Experiments by Riste and co-workers in 1971 and later work by others showed that the behavior was not described by a single timescale. **Figure 5** shows the scattering observed just above T_c in $SrTiO_3$ and this has an inelastic peak and a sharp quasi-elastic peak. When the temperature is well above T_c the frequency of the mode decreases and there is little quasi-elastic scattering. Closer to the phase transition, the intensity of the quasi-elastic scattering rapidly increases while the soft mode part becomes almost independent of temperature. The scattering can be described by a response function of the form:

$$\chi^{-1}(qj) = \omega(qj)^2 - \omega^2 - 2i\omega\gamma_0 + \frac{i\omega\tau\delta_0^2}{1 - i\omega\tau}$$

in which the response is a damped harmonic oscillator plus an additional term to produce the quasi-elastic scattering and δ and τ are constants that need to be adjusted. The experiment showed that the additional relaxation was very slow and could not be easily resolved in the experiments. Similar results have now been observed at many phase transitions and indeed at most order–order phase transitions. Comparison with the results of X-ray scattering measurements has shown that the central peak needs to be included in the scattering if both X-ray and neutron scattering methods are to agree with one another and the measurements are not too close to the phase transition.

The first evidence for two length scales came from the observations of Andrews. He performed X-ray scattering measurements with Cu radiation and found that close to T_c, $|t| < 0.03$, there was additional scattering with a new and longer length scale. Subsequently, most materials that have been studied with high-resolution techniques have shown a very similar behavior, with a new long length scale that diverges at the phase transition with an exponent that is larger than that expected for the susceptibility. In summary the experimental results

for the second timescale and second length scale are as follows:

1. Two energy scales
   (a) At phase transitions between two ordered phases there is both soft mode and quasi-elastic scattering with very different energy scales for $t<1$,
   (b) The quasi-elastic energy scale is very slow, <0.01 meV,
   (c) The quasi-elastic neutron scattering is well described by a low-frequency Debye relaxation,
   (d) Both inelastic and quasi-elastic components of the neutron scattering experiments are measured by X-ray scattering techniques, and
   (e) The quasi-elastic peak is enhanced by the introducing defects but much more slowly than linearly with the defect concentration.
2. Two length scales
   (a) Two length scale components have been observed in measurements made with adequate resolution, 0.001 Å,
   (b) The narrow scattering corresponds to a length scale that is 10 times longer than the short length scale,
   (c) The critical exponents for the two length scales are different from one another with the shorter scale being broadly consistent with that expected from nonclassical fluctuations,
   (d) The long length scale scattering is broadly consistent with an isotropic Lorentzian squared, and
   (e) The scattering seems to arise from the "near," within 50 μm, surface and its intensity depends on the surface preparation.

Much effort has been put into finding the origin of the additional second timescale. It is well established that the theory of anharmonic lattice dynamics can give rise to the required behavior below T_c and also above if the mode is a special type of ferrodistortive mode. This theory cannot however explain the scattering observed from most quasi-elastic peaks and in particular cannot explain the effects observed in $SrTiO_3$. Considerable effort has been made to try to account for the results in terms of a critical theory but it is very difficult to make a clear distinction between the quasi-elastic peak and the inelastic peak.

Halperin and Varma suggested that the slow quasi-elastic peak could be accounted for by introducing defects. There is no doubt either theoretically or experimentally that defects can produce a central peak. Nevertheless, Halperin and Varma were unable to make any specific theoretical predictions that could be tested experimentally. The defects were of unknown strength and unknown concentration and it is surprising that the effects are so similar for different types of crystals containing very different

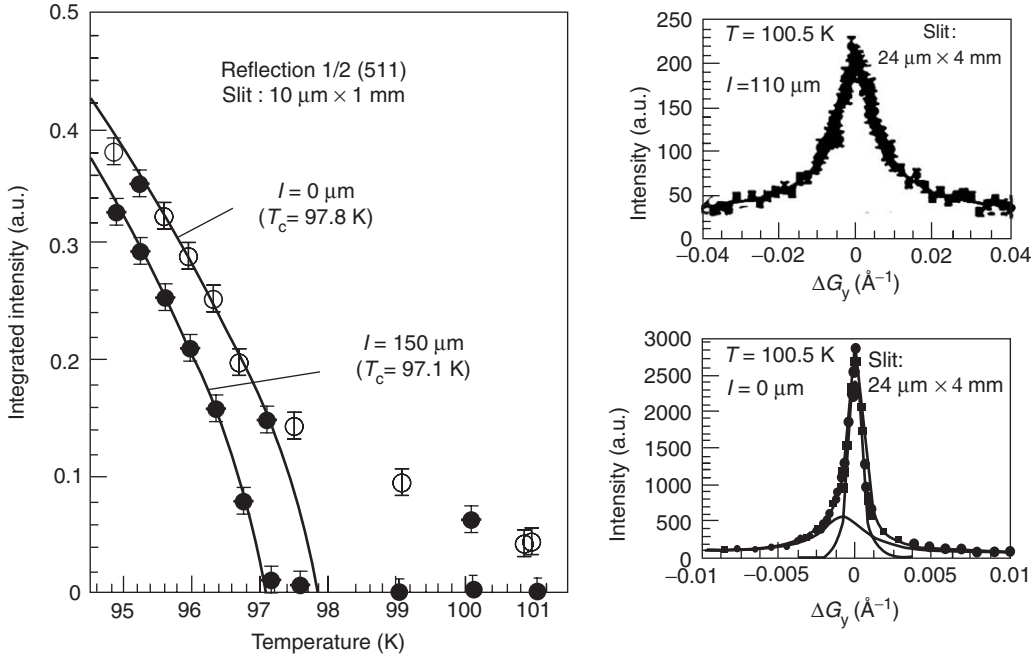


Figure 6 Critical scattering at the surface and at a depth of 110 μm from $SrTiO_3$ as well as the order parameter temperature dependence. (Reproduced permission from Rutt U, *et al.* (1997) Depth dependence of strain, mosaicity and sharp component in the critical scattering of $SrTiO_3$. *Europhysics Letters* 39: 395.)

defects. The authors conclude that the dynamic central peaks are not yet understood and that more experiments and an improved theory are needed to explain the results.

In the case of the two length scales, high-energy X-ray measurements do not show the effect because the scattering is then not dominated by the surface region. Two possible explanations have been put forward. First, Altarelli *et al*. have suggested that dislocations at the surface could alter the response in the near surface region while Harris *et al*. have suggested that random fields at the surface would be dependent on the surface preparation. There is no doubt that the preparation of the surface does make a difference to the results, but nevertheless it is difficult to understand why the defects have very similar results for very different materials and indeed for both magnetic and structural phase transitions. An alternative explanation is that the surface waves propagating at the surface are free in the sense that the surface of the crystal can respond to these strains while it cannot respond to bulk elastic strains. This can raise the temperature of the phase transition so that the distorted region steadily grows in from the surface. The new narrow component of the scattering is then a measure of the extent to which the surface distortion has grown into the crystal due to the coupling of the strain to the degree of order.

Summary

The concept of a soft mode has been reviewed as a mechanical instability of a crystal against one of the normal modes of vibration. The basic concept is correct and has been shown to very satisfactorily account for the experimental behavior in very different systems. Despite this undoubted success detailed measurements close to the phase transition are not well understood. Structural phase transitions are usually characterized by quasi-elastic scattering as well as the inelastic soft mode scattering and currently there is no satisfactory account of that scattering although it is now over 30 years since the phenomenon was discovered. The basic theory depends on the second energy scale coming from defects even though no defects have been introduced into the crystals and it is unlikely that very similar results would be obtained from mechanically very different crystals. Experiments have also been performed that have shown a second length scale in a "near" surface region. These are now understood as arising either from defects or from a part of the critical scattering that arises from the order parameter strain coupling that is dependent on the scattering from the free surface. Further experiments and theory are needed before these effects are understood (**Figure 6**).

See also: Ferroelectricity; Lattice Dynamics: Anharmonic Effects; Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Vibrational Modes; Piezoelectricity.

PACS: **63.20; 63.75; 64.70K**

Further Reading

Bruce AD and Cowley RA (1981) *Structural Phase Transitions*. London: Taylor and Francis.

Cowley RA (1996) Are there two length scales at phase transitions? *Physica Scripta* T66: 24–30.

Toledano JC and Toledano P (1987) *The Landau Theory of Phase Transitions*. World Scientific Lecture Notes Vol. 3.

Lattice Dynamics: Vibrational Modes

K Parlinski, Institute of Nuclear Physics, Kraków, Poland

Introduction

Atoms vibrate in any material. The vibrational modes describe a number of properties, like temperature dependence of crystal lattice parameters and elastic constants; they determine thermal expansion properties, heat and electric conductivity, thermo dynamic functions, and location of the critical temperature in the phase diagram. Vibrational modes enhance the diffusion processes, and play a crucial role in superconductivity, ferroelectricity, shape-memory alloys, etc. Neutron and synchrotron radiation scattering, Raman scattering, and infrared absorption methods are used to study atomic vibrations in detail. Of course, atoms vibrate in pure crystals, in crystals with substitutional and interstitial defects, atomic clusters, dislocations, or grain boundaries. Atomic vibrations on surfaces and interfaces are different compared to those in the bulk, due to differences in neighboring atoms at these locations.

Lattice Dynamics Theory

The basic theory of the lattice dynamics has been given by Born and Huang, who developed the theory

for crystal lattices with translational symmetry. The theory uses the Born–Oppenheimer adiabatic approximation, which means that the electrons, moving fast due to their lighter masses, follow the much slower motion of nuclei. This approximation allows one to write the potential energy V of the system as a function of the instantaneous position of atomic nuclei $\boldsymbol{R}(\boldsymbol{n},\mu)$:

$$V = V(\ldots, \boldsymbol{R}(\boldsymbol{n},\mu),\ \ldots)$$

where $\boldsymbol{R}(\boldsymbol{n},\mu)$ is the position of an atom vibrating close to the lattice site μ of the primitive unit cell $\boldsymbol{n}$. One defines the site position $\boldsymbol{R}^0(\boldsymbol{n},\mu)$ at the minimum of the potential energy. Crystal symmetry may fix the location $\boldsymbol{R}^0(\boldsymbol{n},\mu)$. Indeed, some intersections of two or more symmetry elements at a single point define a site symmetry, and, consequently, any crystal function, including the potential energy, may have an extremum at this point. The atomic vibrations are usually small and amount to a small percentage of the closest interatomic distances. The potential energy can formally be expanded in terms of such displacements, $\boldsymbol{u}(\boldsymbol{n},\mu) = \boldsymbol{R}(\boldsymbol{n},\mu) - \boldsymbol{R}^0(\boldsymbol{n},\mu)$:

$$\begin{aligned} V = {} & V(\ldots \boldsymbol{R}^0(\boldsymbol{n},\mu), \ldots) \\ & + \frac{1}{2} \sum_{\boldsymbol{n},\mu,\boldsymbol{m},\nu} \Phi(\boldsymbol{n},\mu;\boldsymbol{m},\nu)\boldsymbol{u}(\boldsymbol{n},\mu)\boldsymbol{u}(\boldsymbol{m},\nu) \\ & + \text{higher-order terms} \end{aligned}$$

The term $V(\ldots \boldsymbol{R}^0(\boldsymbol{n},\mu), \ldots)$ stands for the athermal crystal ground-state energy. Linear expansion terms vanish because they are the first derivatives of a continuous function at the extrema (usually minima). The second expansion term is called harmonic. The expansion coefficient

$$\Phi(\boldsymbol{n},\mu;\boldsymbol{m},\nu) = \left.\frac{\partial^2 V}{\partial \boldsymbol{R}(\boldsymbol{n},\mu)\partial \boldsymbol{R}(\boldsymbol{m},\nu)}\right|_{\mathrm{o}}$$

defines the 3×3 matrix of force constant connecting atoms $(\boldsymbol{n},\mu)$ and $(\boldsymbol{m},\nu)$. In crystals, the force constant matrices remain invariant with respect to these crystal symmetry elements, which do not alter the vector joining the two connected atoms $(\boldsymbol{n},\mu)$ and $(\boldsymbol{m},\nu)$. Third-, fourth-, and higher-order terms of the potential energy expansion are called anharmonic contributions and are neglected in the harmonic approach.

In the harmonic approximation, the equations of motion of vibrating atoms are exactly solvable. The solutions of the equations of motion for the atomic displacements are usually chosen as

$$\begin{aligned} \boldsymbol{u}(\boldsymbol{n},\mu) = {} & \frac{Q(\boldsymbol{k},j)}{\sqrt{M_\mu}}\boldsymbol{e}(\boldsymbol{k},j;\mu) \\ & \times \exp[2\pi \mathrm{i}(\boldsymbol{k}\times\boldsymbol{R}(\boldsymbol{n},\mu) - \omega(\boldsymbol{k},j)t)] \end{aligned}$$

where M_μ denotes atomic mass, $\boldsymbol{k}$ is the wave vector defining a point in the reciprocal space of the crystal lattice. The $Q(\boldsymbol{k}, j)$ is the normal-mode amplitude. The equations of motion then leads to the eigenvalue problem

$$\omega^2(\boldsymbol{k},j)\boldsymbol{E}(\boldsymbol{k},j) = \boldsymbol{D}(\boldsymbol{k})\boldsymbol{E}(\boldsymbol{k},j)$$

where $j = 1, 2, \ldots, 3r$ labels the phonon branches. The number of phonon branches can be less than $3r$, if some of them become degenerate due to crystal symmetry.

The eigenvector $\boldsymbol{E}(\boldsymbol{k}, j)$ is a column vector and the dynamical matrix forms a square matrix:

$$\boldsymbol{E}(\boldsymbol{k},j) = \begin{pmatrix} \boldsymbol{e}(\boldsymbol{k},j;1) \\ \vdots \\ \boldsymbol{e}(\boldsymbol{k},j;\mu) \\ \vdots \\ \boldsymbol{e}(\boldsymbol{k},j;3r) \end{pmatrix}$$

and

$$\boldsymbol{D}(\boldsymbol{k}) = \begin{pmatrix} \boldsymbol{D}(\boldsymbol{k};1,1), & \ldots, & \boldsymbol{D}(\boldsymbol{k};1,3r) \\ & \vdots & \\ \ldots, & \boldsymbol{D}(\boldsymbol{k};\mu,\nu), & \ldots \\ & \vdots & \\ \boldsymbol{D}(\boldsymbol{k};3r,1), & \ldots, & \boldsymbol{D}(\boldsymbol{k};3r,3r) \end{pmatrix}$$

and

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{k};\mu,\nu) = {} & \frac{1}{\sqrt{M_\mu M_\nu}} \sum_{m} \Phi(0,\mu;\boldsymbol{m},\nu) \\ & \times \exp\{-2\pi \mathrm{i}\boldsymbol{k}\cdot[\boldsymbol{R}(0,\mu) - \boldsymbol{R}(\boldsymbol{m},\nu)]\} \end{aligned}$$

$\boldsymbol{D}(\boldsymbol{k})$ is an Hermitian matrix of dimension $3r \times 3r$. Here, r is the number of atoms in the primitive unit cell. The summation $\boldsymbol{m}$ runs over unit cells of the crystal. The eigenvalue equation provides the values of squares of the eigenfrequencies, $\omega^2(\boldsymbol{k}, j)$. The eigenvectors $\boldsymbol{E}(\boldsymbol{k},j)$, also called the polarization vectors, satisfy the orthonormality conditions.

For a given wave vector $\boldsymbol{k}$, there exist $3r$ values of $\omega^2(\boldsymbol{k}, j)$ frequencies, each of which describes a normal mode vibration. In a stable crystal, all $\omega^2(\boldsymbol{k}, j)$ are real and positive. To each $\omega^2(\boldsymbol{k}, j)$ is associated an eigenvector $\boldsymbol{E}(\boldsymbol{k}, j)$ having $3r$ components, as many as

the number of degrees of freedom in the primitive unit cell. The eigenvector controls the participation and displacements of atoms in the vibrational mode. The eigenvector components are complex numbers.

Quantum Treatment

In quantum mechanics, the crystal harmonic vibrations are treated as a set of independent harmonic oscillators. The states of each quantum oscillator form equidistant energy levels, which are labeled by the quantum number n. In the classical treatment, the amplitude of oscillations can increase continuously. In the quantum approach, the amplitude of vibrations is specified by the probability of the oscillator wave functions, which spread wider for higher quantum numbers. The energy of the quantum oscillator at state n is then

$$E_n(\boldsymbol{k},j) = \hbar\omega(\boldsymbol{k},j)\left(n+\tfrac{1}{2}\right)$$

In contrast to a classical oscillator, for which the lowest energy is zero, the ground-state energy of the quantum oscillator is nonzero, $E_0(\boldsymbol{k},j) = (1/2)\,\hbar\omega(\boldsymbol{k},j)$.

The quantum of an energy difference $\Delta n = 1$ is called a phonon. Thus, to specify a crystal state, one should give the number of phonons occupying each normal mode. The amplitude of atomic vibrations, or the number of phonons, depends on the crystal temperature. In a canonical ensemble, in which the external variables are temperature T and volume V, the mean number of phonons in the vibrational normal mode $(\boldsymbol{k},j)$ is given by the Bose–Einstein statistics,

$$n(\boldsymbol{k},j) = \frac{1}{\exp[\hbar\omega(\boldsymbol{k},j)/k_{\mathrm{B}}T]-1}$$

where k_{B} is a Boltzmann constant. Any number of phonons can be inserted into the crystal. Harmonic phonons do not interact, hence a thermal equilibrium between them can be achieved only with a contribution of anharmonicity.

In the harmonic approximation, the energy E of the atomic vibrations is a sum of kinetic and potential energies, and could be represented as

$$E = \sum_{\boldsymbol{k},j} \omega^2(\boldsymbol{k},j)\langle|Q(\boldsymbol{k},j)|^2\rangle$$

where the mean square amplitude of the quantum harmonic oscillator depends on the crystal temperature, and is given by

$$\langle|Q(\boldsymbol{k},j)|^2\rangle = \frac{\hbar}{2\omega(\boldsymbol{k},j)}\,\mathrm{cth}\,\frac{\hbar\omega(\boldsymbol{k},j)}{2k_{\mathrm{B}}T}$$

Phonon Dispersion Relations

The phonon dispersion relations are defined as the $\boldsymbol{k}$ dependence of the frequencies, $\omega(\boldsymbol{k},j)$, of the normal modes for all branches j and selected directions in the crystal (**Figure 1**). The number of phonon branches, $j = 1, 2, \ldots, 3r$, is equal to the number of degrees of freedom in the primitive unit cell. Each point on the phonon dispersion curve $\omega(\boldsymbol{k},j)$ gives the frequency of a phonon, which can be visualized as a dynamical wave of length $\lambda = 1/|\boldsymbol{k}|$, propagating along the $\boldsymbol{k}/|\boldsymbol{k}|$ direction. In this wave, the atoms vibrate with frequency $\omega(\boldsymbol{k},j)$, and displace from the equilibrium positions as indicated by the polarization vectors $\boldsymbol{E}(\boldsymbol{k},j)$. Typical maximal phonon frequencies range from 10 to 30 THz, where $1\,\mathrm{THz} = 10^{12}\,\mathrm{s}^{-1}$. Other units are also used: $1\,\mathrm{THz} = 4.1357\,\mathrm{meV} = 33.356\,\mathrm{cm}^{-1}$. The amplitudes of vibrations are ~ 0.03–$0.08\,\text{Å}$.

Conventionally, the phonon dispersion relations are drawn along crystal high-symmetry axis, but other directions might also be shown. At the Γ point, which corresponds to $\boldsymbol{k} = 0$, or to a reciprocal lattice vector $\boldsymbol{k} = \boldsymbol{b}^*$, there are always three phonons with zero frequency, $\omega(\boldsymbol{k} = \Gamma, j) = 0$. The phonon

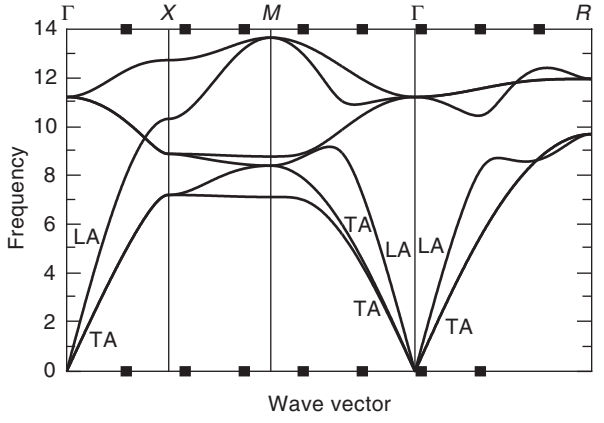


Figure 1 Schematic representation of phonon dispersion relations of a primitive cubic crystal *AB*, with atoms *A* and *B* at positions (0,0,0) and $(1/2, 1/2, 1/2)$. The high-symmetry directions Γ–*X*, *X*–*M*, *M*–Γ and Γ–*R* connect the high-symmetry points $\Gamma = (0,0,0)$, $X = (1/2,0,0)$, $M = (1/2,1/2,0)$, $R = (1/2,1/2,1/2)$ of the primitive cubic Brillouin zone. In this example the force constants for *A* and *B* atoms are assumed to be the same, while the masses are different, $M_A = 1.5\,M_{\mathrm{B}}$. Animations of normal-mode vibrations of selected phonons can be viewed at http://wolf.ifj.edu.pl. Vibrational amplitudes are much larger than in real crystals.

branches starting at $\omega(\boldsymbol{k}=\Gamma,j)=0$ are called acoustic phonon dispersion curves. Close to $\boldsymbol{k}=\Gamma$, the acoustic branches show a linear dependence, that is, $\omega\sim|\boldsymbol{k}|$, and the slope is described by a combination of the elastic constants of a crystal. The exact relationship between the slope and elastic constants depends on the crystal point group. In some high-symmetry crystals, and along high-symmetry directions, the atomic vibrations are either polarized along the propagation wave vector $\boldsymbol{k}$ (longitudinal acoustic modes, LA), or perpendicular to $\boldsymbol{k}$ (transverse acoustic modes, TA). Other phonon dispersion relations, being of nonzero frequency at Γ point $\omega(\boldsymbol{k}=\Gamma,j)\neq 0$, are called optic. Optic modes could have similar longitudinal and transverse characters. Local atomic vibrations lead usually to flat phonon dispersion relations. A strong coupling of displacements between neighboring atoms causes a large phonon dispersion, that is, a large $\boldsymbol{k}$ dependence of $\omega(\boldsymbol{k},j)$.

The frequencies of phonon modes are periodic with the reciprocal lattice vector $\boldsymbol{b}^*$,

$$\omega(\boldsymbol{k}+\boldsymbol{b}^*,j)=\omega(\boldsymbol{k},j)$$

in other words, the phonon dispersion relations in every Brillouin zone of the crystal are the same. In contrast, the polarization vectors may differ by a phase factor in going from one Brillouin zone to another zone.

In principle, the complete phonon dispersion relations can be measured by coherent inelastic neutron scattering, or by inelastic X-ray scattering, provided the phonon peaks are well resolved, and their intensities are sufficient. Some phonon frequencies $\omega(\boldsymbol{k}=\Gamma,j)$ at wave vector $\boldsymbol{k}=\Gamma$ can be measured by infrared absorption and/or Raman scattering methods. In nonmetallic crystals, the phonon modes, having symmetry which allows one to observe them by an infrared absorption method, are split into longitudinal optic (LO) and transverse optic (TO) components having different frequencies. The difference between LO and TO frequencies depends on the magnitude of the Born effective charges and the electronic part of the dielectric constant. In metals, the LO/TO splitting does not occur. In molecular crystals, the internal vibrations of molecular constituents usually present local vibrations; therefore, the phonon dispersion curves remain flat, while the intermolecular coupling provides acoustic and low-frequency lattice optic phonons of appreciable dispersion.

Realistic phonon dispersion relations for the athermal regime ($T=0$, no phonon displacements) can be calculated from first principles either by a direct method, or by a linear response perturbation method. In the direct method, the Hellmann–Feynman forces, being the first derivatives of the crystal potential energy and arising due to the atomic displacements, are calculated from first principles, and then used to derive the force constants. In the linear response method, the force constants are calculated directly as the second derivatives of the ground-state energy of the crystal.

Phonon Density of States

The phonon density of states $g(\omega)$ describes the number of phonon modes of a specific frequency $\omega(\boldsymbol{k},j)$ in a given frequency interval $(\omega-1/2(\Delta\omega),\omega+1/2(\Delta\omega))$, if the density of wave vectors $\boldsymbol{k}$ in the Brillouin zone is homogeneously distributed (**Figure 2**). Formally, $g(\omega)$ is obtained from the relation

$$g(\omega)=\frac{1}{3r\,N\,\Delta\omega}\sum_{\boldsymbol{k},j}\delta_{\Delta\omega}(\omega-\omega(\boldsymbol{k},j))$$

where the new δ function is defined as $\delta_{\Delta\omega}(x)=1$ if $\Delta\omega/2<x\leqslant\Delta\omega/2$, and $\delta_{\Delta\omega}(x)=0$ otherwise. The summation runs over wave vectors $\boldsymbol{k}$ of the first Brillouin zone, and all phonon branches. Here, N is the number of wave vectors $\boldsymbol{k}$. The phonon density of states is normalized $\int_0^\infty \mathrm{d}\omega\, g(\omega)=1$.

The phonon density of states $g(\omega)$ spreads from zero to the maximal phonon frequency existing in a given crystal. For acoustic phonons, behaving linearly close to $\boldsymbol{k}=\Gamma$, $g(\omega)$ is proportional to ω^2 close to $\omega=0$. In simple crystals, the ω^2 dependence may

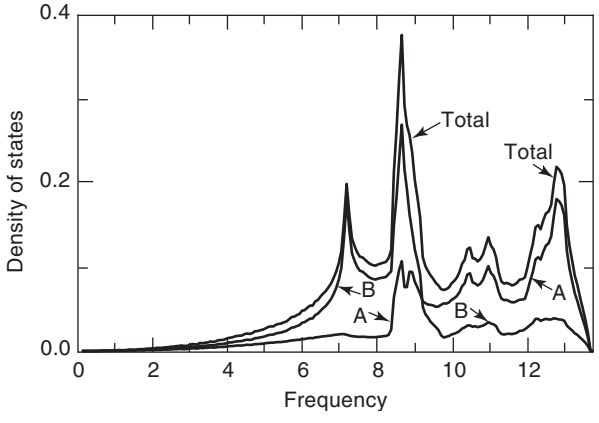


Figure 2 Schematic representation of phonon density of states of a primitive cubic crystal *AB*, with atoms *A* and *B* at positions (0,0,0) and $(1/2,1/2,1/2)$ the same as in **Figure 1**. In this example the force constants *A* and *B* atoms are assumed to be the same, while the masses are different, $M_A=1.5\,M_B$. The partial phonon density of states for *A* and *B* atoms are also shown.

cover a substantial part of the ω interval; however, in complex crystals, the quadratic dependence is limited to the close vicinity of the point $\omega = 0$. Flat phonon dispersion relations may lead to sharp peaks in $g(\omega)$. In simple crystals with a large mass difference of the constituents the vibrations could separate to a low-frequency phonon band, caused mainly by oscillations of heavy atoms, and to a high-frequency band, occupied by light atoms. The bands could be separated by a frequency gap. It should, however, be stressed that, without measurements or precise calculations, the prediction of phonon spectra is impossible.

To describe the vibrations of a specific atom μ moving along the i-direction a partial phonon density of states is introduced

$$g_{i,\mu}(\omega) = \frac{1}{3r\,N\,\Delta\omega} \sum_{k,j} |e_i(\boldsymbol{k}, j; \mu)|^2 \times \delta_{\Delta\omega}(\omega - \omega(\boldsymbol{k}, j))$$

The factor $|e_i(\boldsymbol{k}, j; \mu)|^2$ represents the abundance of a given mode. It is normalized to $\int_0^\infty \mathrm{d}\omega\, g_{i,\mu}(\omega) = 1/3r$. This density of states defines a participation of an atom μ in harmonic vibrations occurring along the selected direction i. It suggests which atom is taking part in a given peak of $g(\omega)$.

Remarks

Crystal symmetries are specified by the space groups. Possible transformations of atomic displacements are then classified by the irreducible representations of the crystal space group. An irreducible representation can be assigned to each normal mode or phonon. The classification of normal modes by irreducible representations can be done without diagonalization of the dynamical matrix, that is, without the knowledge of the phonon frequencies. These irreducible representations also describe the activity of phonon peaks in Raman scattering and infrared absorption, the possibility of LO/TO splitting, the low-symmetry space group arising at a phase transition caused by a soft mode, and the interaction of phonon branches. In particular, phonon branches labeled by the same irreducible representation cannot cross.

The energy E of the harmonic crystal shows that any phonon-like fluctuation always increases the system energy, since $\omega^2(\boldsymbol{k}, j) > 0$ and $\langle |Q(\boldsymbol{k}, j)|^2 \rangle > 0$. Thus, phonons do not destroy the crystal stability. If, however, due to external conditions, $\omega^2(\boldsymbol{k}_c, j_c)$ becomes negative, the crystal ceases to be stable. It induces a phase transition with a change of space group. The increasing vibrational amplitude is confined by anharmonic terms. The phonon with critical frequency $\omega^2(\boldsymbol{k}_c, j_c)$ is called a soft mode.

See also: Lattice Dynamics: Anharmonic Effects; Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Structural Instability and Soft Modes.

PACS: 63; 63.20.−e; 63.10.+a; 63.70.+h

Further Reading

Birman JL (1974) Space groups theory and infra-red and Raman optical processes in crystals. *Hanbuch der Physik*, 25/2b.

Born M and Huang K (eds.) (1954) *Dynamical Theory of Crystal Lattices*. Oxford: Oxford University Press.

Leibfried G and Ludwig W (1961) Theory of anharmonic effects in crystals. In: Seitz F and Turnbull D (eds.) *Solid State Physics*, vol. 12. New York: Academic Press.

Maradudin AA and Horton GK (eds.) (1974) *Dynamical Properties of Solids*. Amsterdam: North-Holland, New York: American Elsevier Publishing Company.

Maradudin AA, Montrol EW, Weiss GH, and Ipatova IP (eds.) (1971) *Theory of Lattice Dynamics in the Harmonic Approximation*. New York: Academic Press.

Maradudin AA and Vosko SH (1968) *Reviews of Modern Physics* 59: 673.

Light Emitting Diodes

E F Schubert and T Gessmann, Rensselaer Polytechnic Institute, Troy, NY, USA

Introduction

Electrons and holes in semiconductors recombine either radiatively, that is, accompanied by the emission of a photon, or nonradiatively. In light emitting devices, the former is the preferred process. Although nonradiative recombination can be minimized, it can, under practical conditions, never be reduced to zero. Thus, there is competition between radiative and nonradiative recombination. Maximization of the radiative process can be attained in a number of ways that is discussed below.

In light emitting diodes (LEDs), radiative recombination is spontaneous in nature and the recombination rate depends on the concentration of carriers.

The situation is different for laser diodes, in which stimulated recombination dominates. This process depends on the concentration of photons as well as the concentration of carriers.

The energy of emitted photons is near the bandgap energy of the active (light-emitting) material. To date, LEDs are available that cover the IR and entire visible wavelength range from 390 nm (violet) to 720 nm (red). In addition, LEDs emitting in the UV wavelength range have recently become available. For visible-spectrum LEDs, III–V compound semiconductors such as AlGaInP, GaInN, and AlGaAs are used. Their bandgap energy can be adjusted by the chemical composition, for example, the mole fraction x in $Ga_{1-x}In_xN$. III–V phosphides are used for emitters in the IR, and in the red to green wavelength range, whereas III–V nitrides are used for emitters in the green to violet range and the UV wavelength range. The emission of III–V arsenides extends from the IR to the red part of the spectrum.

Optical Properties

The probability that electrons and holes recombine radiatively is proportional to the product of electron and hole concentration, that is, $R \propto np$. The recombination rate per unit time per unit volume can be written as

$$R = -\frac{dn}{dt} = -\frac{dp}{dt} = Bnp \qquad [1]$$

where B is a proportionality constant, the bimolecular recombination coefficient, with a typical value of $10^{-10}\,cm^3\,s^{-1}$ for direct-gap III–V semiconductors.

Fundamental types of band structures and associated recombination processes are shown in **Figure 1**. During the recombination process, the electron momentum ($p = (2m^*E)^{1/2}$) cannot change significantly because momentum must be conserved and the photon momentum ($p = h/\lambda$) is negligibly small. Thus optical transitions must be "vertical" in nature, that is, electrons recombine only with holes that have the same k value as shown in **Figure 1**.

Efficient recombination occurs in direct-gap semiconductors shown in **Figure 1a**. However, the recombination probability is much lower in indirect-gap semiconductors because a phonon is required to satisfy momentum conservation, as shown in **Figure 1b**. The radiative efficiency of indirect-gap semiconductors can be increased by isoelectronic impurities, for example, N in GaP. Isoelectronic impurities can form an optically active deep level that is localized in real space (small Δx) but, according to the uncertainty relation, delocalized in k space (large Δk), so that the impurity can satisfy the momentum conservation, as indicated in **Figure 1c**.

During nonradiative recombination, the electron energy is converted to vibrational energy of lattice atoms, that is, phonons. There are several physical mechanisms by which nonradiative recombination can occur with the most common ones being point defects (impurities, vacancies, interstitials, antisite defects, and impurity complexes) and spatially extended defects (screw and edge dislocations, cluster defects). It is quite common for such defects to form one or several energy levels within the forbidden gap. The defects act as efficient recombination centers (Shockley–Read recombination centers) in particular, if the energy level is close to the middle of the gap.

The emission spectra of an AlGaInP red, a GaInN green, and a GaInN blue LED are shown in **Figure 2**. The spectral line width of spontaneous emission can be calculated to be $1.8\,kT$ and is caused by the thermal energy of carriers. However, the LEDs shown in **Figure 2** have an active region comprised of a ternary

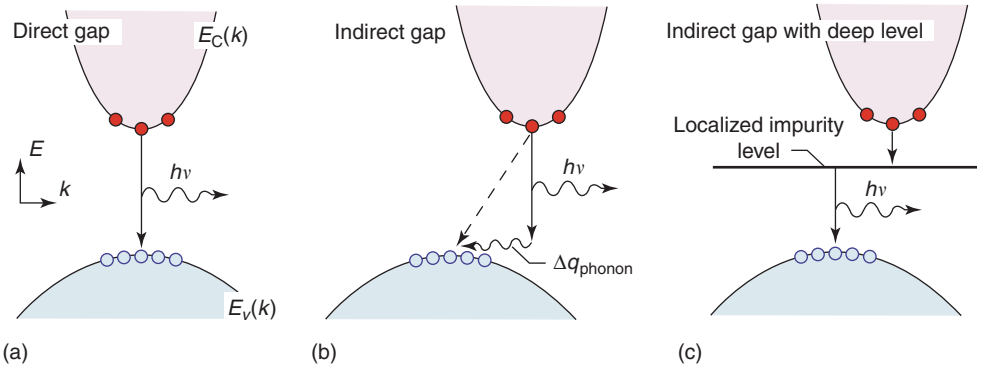


Figure 1 Semiconductor band structure for (a) direct-gap, (b) indirect-gap, and (c) indirect-gap semiconductors with deep isoelectronic impurity level. Indirect transitions have a much smaller probability due to the requirement of a phonon.

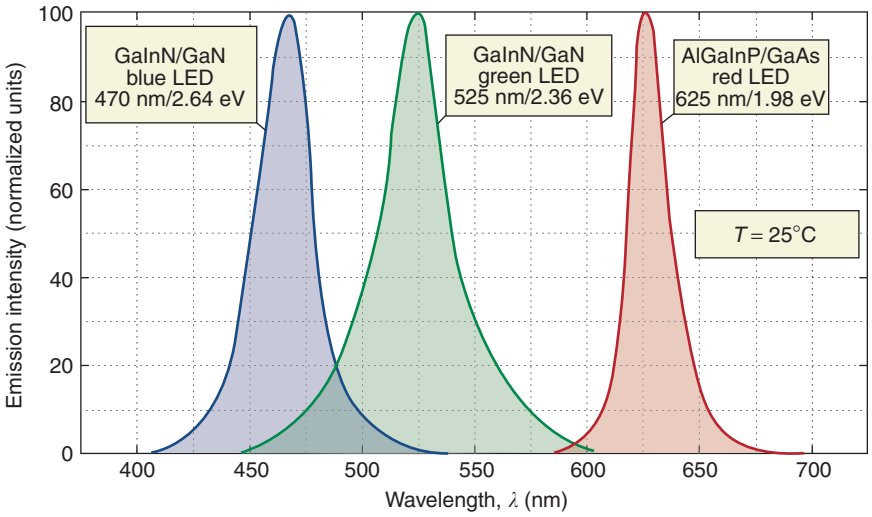


Figure 2 Typical emission spectrum of GaInN/GaN blue, GaInN/GaN green, and AlGaInP/GaAs red LEDs.

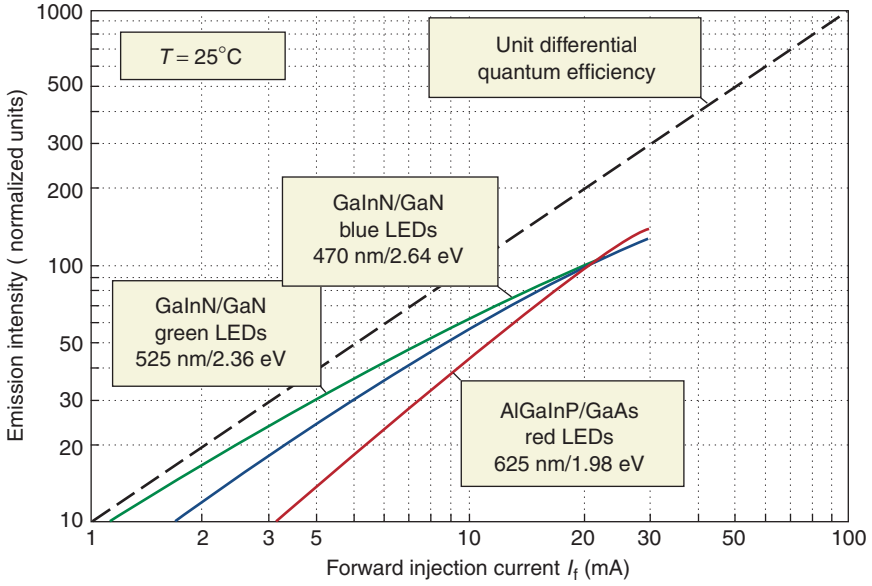


Figure 3 Typical light output power vs. injection current of GaInN/GaN blue, GaInN/GaN green, and AlGaInP/GaAs red LEDs.

or quaternary alloy, for example, $Ga_{1-x}In_xN$. Alloy broadening, caused by fluctuations of the active region chemical composition, leads to spectral broadening that goes beyond $1.8\,kT$. Alloy broadening is particularly strong in the green $Ga_{1-x}In_xN$ LEDs with line widths as wide as $10\,kT$ at room temperature.

The light-output-power versus injection-current is shown in **Figure 3** for three types of visible-spectrum LEDs. A linear dependence with unit slope (dashed line), that is, unit differential quantum efficiency, is expected for the light-versus-current curves. The mature AlGaInP LED closely follows the unit-slope line. However, the slope of the green and blue LEDs decreases with increasing current level indicating a decreasing efficiency (efficiency droop).

Electrical Properties

In the vicinity of the unbiased *p–n* junction plane shown in **Figure 4a**, electrons originating from donors on the *n*-type side diffuse to the *p*-type side where they recombine. A corresponding process occurs with holes. As a result, a region near the *p–n* junction, the depletion region, is depleted of

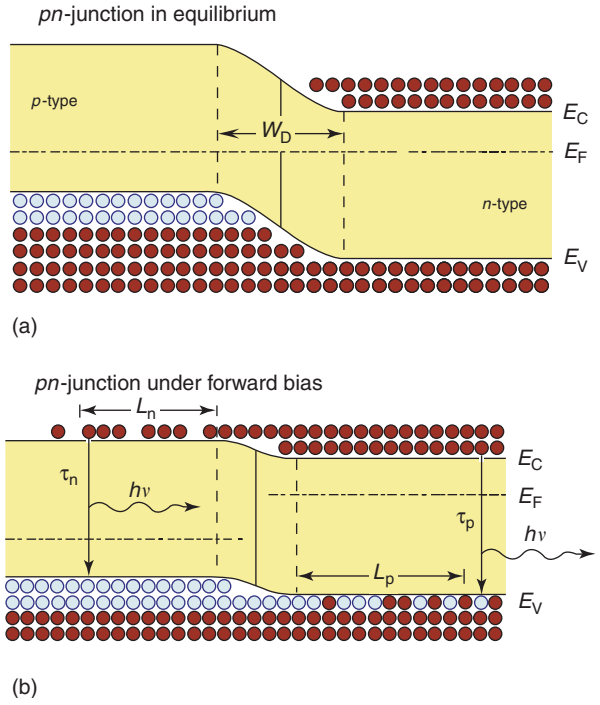


Figure 4 Band diagram and carrier distribution of a *p–n* homojunction (a) without and (b) with applied forward bias. $L_{n/p}$ and $\tau_{n/p}$ are the minority carrier diffusion lengths and lifetimes, respectively. W_D is the width of the depleted region.

free carriers. In the absence of free carriers in the depletion region, the only type of charge is ionized donor and acceptor charges. This depletion region produces a potential, the diffusion voltage, V_D, given by

$$V_D = \frac{kT}{e} \ln \frac{N_A N_D}{n_i^2} \quad [2]$$

where N_A and N_D are the acceptor and donor concentrations, respectively, and n_i is the intrinsic carrier concentration of the semiconductor. The diffusion voltage represents a barrier that free carriers must overcome in order to reach the neutral region of the opposite conductivity type.

The width of the depletion region is given by

$$W_D = \sqrt{\frac{2\varepsilon}{e}(V_D - V)\left(\frac{1}{N_A} + \frac{1}{N_D}\right)} \quad [3]$$

where $\varepsilon = \varepsilon_r \varepsilon_0$ is the dielectric permittivity of the semiconductor and V is the diode bias voltage.

An external bias applied to a *p–n* junction will drop across the depletion region, which is resistive due to the lack of free carriers. A forward bias decreases the *p–n* junction barrier causing electrons and holes to be injected into the neutral regions with opposite conductivity type. As the current increases, carriers diffuse into the regions of opposite conductivity type and recombine, thereby emitting photons.

The theory of current transport in *p–n* junctions was first developed by William Shockley in the early 1950s and the equation describing the *I–V* characteristic is known as the Shockley equation

$$I = eA\left(\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}}\frac{n_i^2}{N_A}\right)(e^{eV/kT} - 1) \quad [4]$$

where A is the junction area and $D_{n,p}$ and $\tau_{n,p}$ are the electron and hole diffusion constants and minority carrier lifetimes, respectively.

Under typical forward bias conditions, the diode voltage is $V \gg kT/e$, and thus $[\exp(eV/kT) - 1] \approx \exp(eV/kT)$. For forward bias conditions, the Shockley equation can be rewritten as

$$I = eA\left(\sqrt{\frac{D_p}{\tau_p}}N_A + \sqrt{\frac{D_n}{\tau_n}}N_D\right)e^{e(V-V_D)/kT} \quad [5]$$

The exponent in eqn [5] illustrates that the current strongly increases as the diode voltage approaches a threshold, which is about equal to the diffusion voltage, that is, $V_{th} \approx V_D$.

In an ideal diode, every electron injected into the active region will generate a photon. Thus, conservation of energy requires that the electron energy eV equals the photon energy $h\nu$, that is,

$$eV = h\nu \quad [6]$$

Beyond turn-on, the diode becomes highly conductive and the diode voltage V is about the same as the threshold voltage V_{th}. The energy of photons emitted from a semiconductor with energy gap E_g is given by

$$h\nu = E_g + (1/2)kT \approx E_g \quad [7]$$

Thus eqn [6] can be rewritten as

$$V_{th} \approx V_D \approx E_g/e \quad [8]$$

Diode I–V characteristics of AlGaInP red and GaInN green and blue LEDs are shown in **Figure 5**. The experimental threshold voltages shown in the figure and the comparison with the bandgap energy of these materials indicate that the energy gap and the threshold voltage indeed agree reasonably well.

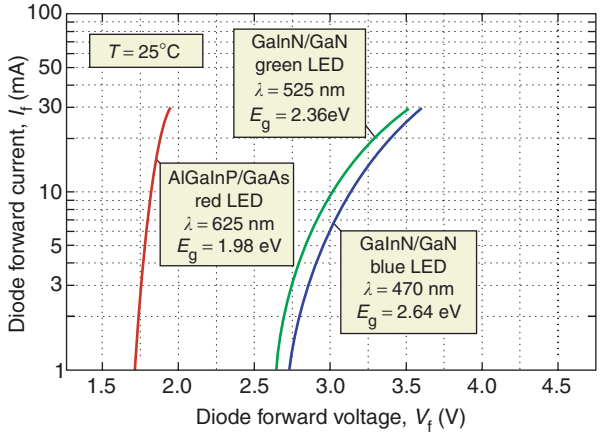


Figure 5 Forward current–voltage characteristic of GaInN/GaN blue, GaInN/GaN green, and AlGaInP/GaAs red LEDs.

Definitions of LED Efficiencies

The internal quantum efficiency of a current-injected semiconductor device is given by

$$\eta_{\text{int}} = \frac{\text{Number of photons emitted from active region per second}}{\text{Number of electrons injected into LED per second}} = \frac{P_{\text{int}}/(h\nu)}{I/e} \quad [9]$$

where P_{int} is the optical power emitted from the active region and I is the injection current. Not all photons emitted by the active region will escape from the LED semiconductor chip. The light extraction efficiency is defined as

$$\eta_{\text{extraction}} = \frac{\text{Number of photons emitted into free space per second}}{\text{Number of photons emitted from active region per second}} = \frac{P/(h\nu)}{P_{\text{int}}/(h\nu)} \quad [10]$$

where P is the optical power emitted into free space.

The external quantum efficiency is defined as

$$\eta_{\text{ext}} = \frac{\text{Number of photons emitted into free space per second}}{\text{Number of electrons injected into LED per second}} = \frac{P/(h\nu)}{I/e} = \eta_{\text{int}}\eta_{\text{extraction}} \quad [11]$$

The external quantum efficiency gives the ratio of the number of useful light particles to the number of injected charge particles.

The power efficiency is defined as

$$\eta_{\text{power}} = P/IV \quad [12]$$

where IV is the electrical power provided to the LED. Informally, the power efficiency is also called "wall-plug efficiency."

The luminous flux measures the light power of a source as perceived by the human eye in units of lumen (lm). One lumen is defined as the luminous flux of a light source emitting an optical power of (1/683) watt at 555 nm. For visible-spectrum LEDs, the luminous efficacy of radiation is defined as the luminous flux emitted by the source divided by the optical power, that is,

$$\eta_{\text{lum radiation}} = 683\,\frac{\text{lm}}{\text{W}}\int_{\lambda} V(\lambda)P(\lambda)\,d\lambda \Big/ \int_{\lambda} P(\lambda)\,d\lambda \quad [13]$$

where 683 lm/W is a normalization factor and $V(\lambda)$ is the eye sensitivity function. The luminous efficiency of the source is defined as the luminous flux emitted by the source divided by the input electrical power, that is,

$$\eta_{\text{lum source}} = 683\,\frac{\text{lm}}{\text{W}}\int_{\lambda} V(\lambda)P(\lambda)\,d\lambda/IV \quad [14]$$

where IV is the LED input power.

High Internal-Efficiency Structures

The radiative recombination rate R given in eqn [1] is proportional to the product of the carrier concentrations n and p. Therefore, a high concentration of carriers in the active region decreases the radiative carrier lifetime and the probability of nonradiative recombination.

The distribution of carriers in a p–n homojunction was shown in **Figure 4**. The minority carrier concentration decreases exponentially with distance from the p–n junction. The mean distance that a minority carrier diffuses before recombination is the diffusion length $L_{n/p}$ which is given by

$$L_n = \sqrt{D_n\tau_n} \quad \text{and} \quad L_p = \sqrt{D_p\tau_p} \quad [15]$$

where τ_n and τ_p are the electron and hole minority carrier lifetimes, respectively. In typical semiconductors, the diffusion length is of the order of micrometers. For example, the diffusion length of electrons in p-type GaAs is $L_n = (220\ \text{cm}^2\ \text{s}^{-1} \times 10^{-8}\ \text{s})^{1/2} = 15\ \mu\text{m}$. Thus, minority carriers are distributed over a region several microns thick and recombination occurs within a relatively large volume.

All high-brightness LEDs employ heterojunctions or quantum wells. A structure consisting of two barriers, that is, two large-bandgap semiconductors, is a

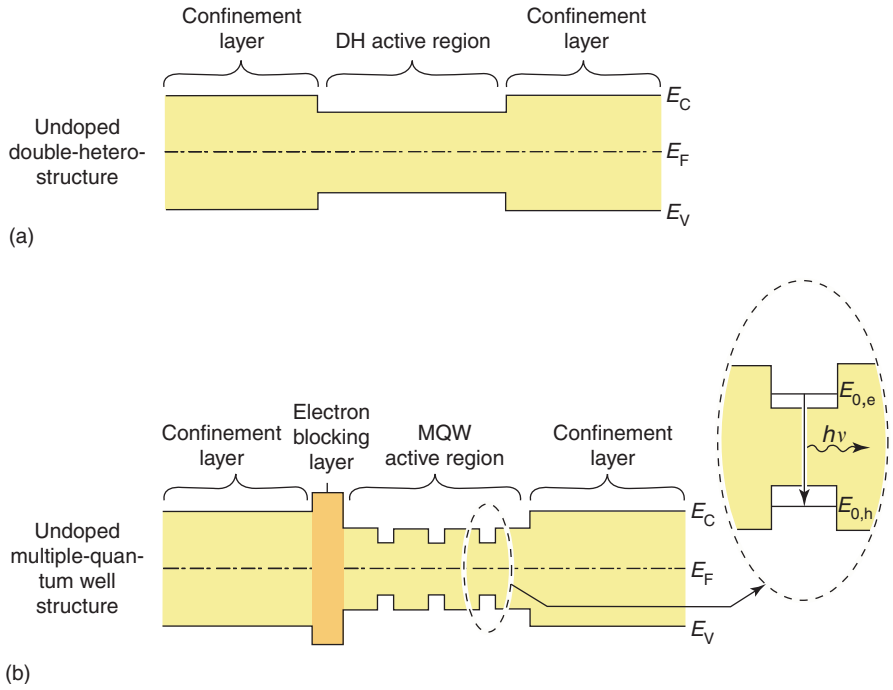


Figure 6 Band diagrams of undoped (a) double hetero (DH) and (b) multiple-quantum well (MQW) structures. The MQW structure employs an electron-blocking layer to avoid electron leakage out of the MQW active region. The inset shows the quantum well region and optical transition between quantized states.

double heterostructure (DH). A multiple quantum well structure (MQW) consists of wells made of a small-bandgap material embedded in the active region of a DH. Band diagrams of a DH and MQW structure are schematically shown in **Figure 6**. Carriers injected into the active regions of the DH or the MQW are confined to the well regions by the barriers. As a result, the thickness of the region in which carriers recombine is given by the thickness of the well rather than the diffusion length.

The consequences of this change are significant. It is assumed that the thickness of the active region is much smaller than the typical diffusion length. Whereas diffusion lengths are several micrometers, the well region of a DH ranges from 0.1 to 1.0 μm. An MQW structure can have well regions as thin as 2 nm. Thus, carriers in a DH or MQW structure have higher carrier concentrations than carriers in homojunctions and a resulting higher radiative efficiency. To further prevent leakage of carriers out of the active region, blocking layers are used, as shown in **Figure 6**. Because electrons are more mobile than holes, electron-blocking layers (rather than hole-blocking layers) are commonly used.

Although DHs and MQWs allow for impro ved LED designs, there are also problems associated with such structures. One of the problems introduced by heterostructures is the resistance caused by the formation of a barrier at the heterointerface. Carriers transferring from one semiconductor to another must overcome this barrier by either tunneling or by thermal emission. The resistance caused by heterojunctions can have a strong deleterious effect on device performance, especially in high-power devices. Fortunately, it has been shown that heterostructure band discontinuities can be completely eliminated by grading of the chemical composition of the semiconductor in the vicinity of the heterostructure.

The overflow of carriers from the active region into the confinement regions is another loss mechanism. Carrier overflow occurs at high injection current densities. As the injection current increases, the carrier concentration in the active region increases. For sufficiently high current densities, the active region is flooded with carriers and the Fermi level may rise to the top of the barrier. As a result, the optical intensity saturates. Carrier overflow of the active region is a potential problem in DH and MQW structures. In order to avoid this problem, high-current LEDs must employ thick DH active regions, or many QWs for MQW active regions.

High Extraction-Efficiency Structures

Owing to the high refractive index of semiconductors, light incident on a planar semiconductor–air interface is totally internally reflected, if the angle of

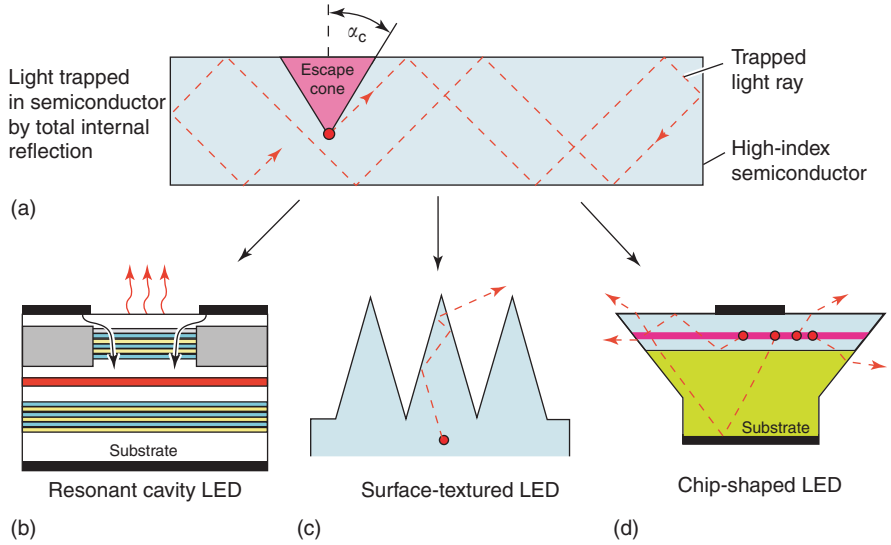


Figure 7 (a) Light rays emanating from a point-like emitter are trapped inside the semiconductor due to total internal reflection. Only light rays with propagation directions falling within the escape cone can leave the semiconductor. Strategies increasing the extraction efficiency include (b) resonant-cavity, (c) textured, and (d) chip-shaped LEDs.

incidence is sufficiently large. Snell's law gives the critical angle of total internal reflection. As a result of total internal reflection, light can be "trapped" inside the semiconductor. The occurrence of trapped light is illustrated in **Figure 7a**. In the high-index approximation, the angle of total internal reflection is given by $\alpha_c = n_s^{-1}$, where n_s is the semiconductor refractive index.

For high-index semiconductors, the critical angle is quite small. For example, for the GaAs refractive index of 3.3, the critical angle for total internal reflection is only 17°. Thus most of the light emitted by the active region is trapped inside the semiconductor. The trapped light eventually will be reabsorbed. Once absorbed, the electron–hole pair may re-emit a photon ("recycling" of a photon) or recombine nonradiatively. If re-emitted as a photon, the photon propagation direction may fall into the escape cone. Thus the exact magnitude of the active region internal quantum efficiency and the probability of a photon to be emitted into the escape cone will determine the overall quantum efficiency of a device. The magnitudes will also determine which strategy (direct light extraction or light extraction by photon recycling) will result in the most efficient device. In the limit of low and high internal efficiency, photon recycling is an ineffective and effective strategy to maximize power efficiency, respectively. Absorption of light in substrates and metallic contacts are other mechanisms lowering the extraction efficiency.

It has been known since the infancy of LED technology that the light-escape problem is strongly correlated to the geometrical shape of the LED chip. Common LED chips have the shape of a rectangular parallelepiped. Such LED chips are fabricated by cleaving the wafer along its natural cleaving planes. In a parallelepiped-shaped chip, light rays propagating within a total of six different cones (corresponding to the six surfaces of a parallelepiped) can escape from the chip. Although rectangular-parallelepiped LEDs have low extraction efficiency, a substantial advantage of such LEDs is low manufacturing cost.

There are different strategies to increase light extraction from LEDs including the resonant-cavity, surface-roughened, and chip-shaped LED shown in **Figures 7b–7d** respectively. The resonant-cavity light-emitting diode (RCLED) has a light emitting region inside an optical cavity. The optical cavity has a thickness of typically one-half or one times the wavelength of light, that is, a fraction of a micrometer for devices emitting in the visible or in the infrared. The resonance wavelength of the cavity coincides with the emission wavelength of the active region. The spontaneous emission properties from a light emitting region located inside the resonant cavity are enhanced by the change in optical mode density occurring in optical cavities. The RCLED was the first practical device making use of a spontaneous emission enhancement occurring in resonant cavities.

The placement of an active region inside a resonant cavity results in multiple improvements of the device characteristics. First, the light intensity emitted from the RCLED along the axis of the cavity, that

is, normal to the semiconductor surface, is enhanced by a typical factor of 2–10. Second, the emission spectrum of the RCLED has a higher spectral purity compared with conventional LEDs. In conventional LEDs, the spectral emission line width is determined by the thermal energy kT. However, in RCLEDs, the emission line width is determined by the quality factor (Q factor) of the optical cavity. As a result, the spectral emission width of the RCLED is a factor of about 2–10 narrower compared with conventional LEDs. Third, the emission far-field pattern of the RCLED is more directed compared with conventional LEDs. In conventional LEDs, the emission pattern is lambertian (i.e., cosine-function-like). In an RCLED, the emission pattern is directed along the optical axis of the cavity.

Other efficient ways to increase the light extraction efficiency include the use of roughened or textured semiconductor surfaces. A cone-shaped surface-roughening is shown in **Figure 7c**. Light rays emanating from the active region below the base of the cone undergo multiple reflections until they eventually have near-normal incidence at the semiconductor–air interface and then escape from the chip. For infrared GaAs-based devices, external quantum efficiencies of 40% have been demonstrated with surface-textured LEDs.

Several chip-shaped LEDs have been commercialized including the pedestal-shaped chip shown in **Figure 7d** and the truncated inverted pyramid (TIP) chip. The ray traces indicated in the figure show that light rays at the base of the pyramid can escape from the semiconductor after undergoing one or multiple internal reflections. The TIP geometry reduces the mean photon path-length within the crystal, and thus reduces internal losses. Ray tracing computer models are employed to optimize the sidewall angle thereby maximizing light extraction.

Packaging

Virtually all LED chips are encapsulated with optically transparent materials such as polymers, epoxies, or silicone. The most common encapsulant is epoxy resin, or briefly epoxy. Encapsulation is accomplished when the epoxy resin mixture is in its liquid state. The mixture is subsequently cured at high temperature to solidify the encapsulant.

The refractive index contrast between the semiconductor and air is reduced by epoxy. Epoxies have typical refractive indices of 1.5 to 1.8. A lower index contrast at the semiconductor surface increases the angle of total internal reflection thereby enlarging the LED escape cone and extraction efficiency.

Furthermore, the epoxy encapsulant commonly has the shape of a hemisphere, as shown in **Figures 8a** and **8b**, so that the angle of incidence at the epoxy–air interface is always perpendicular to the epoxy surface. As a result, total internal reflection does not occur at the epoxy–air interface.

Epoxy domes provide protection against unwanted mechanical shock, chemicals, and moisture (hermeticity). Epoxy also stabilizes the LED chip and bonding wire. Finally, epoxy provides mechanical stability to the two metal leads of the LED chip and holds them in place.

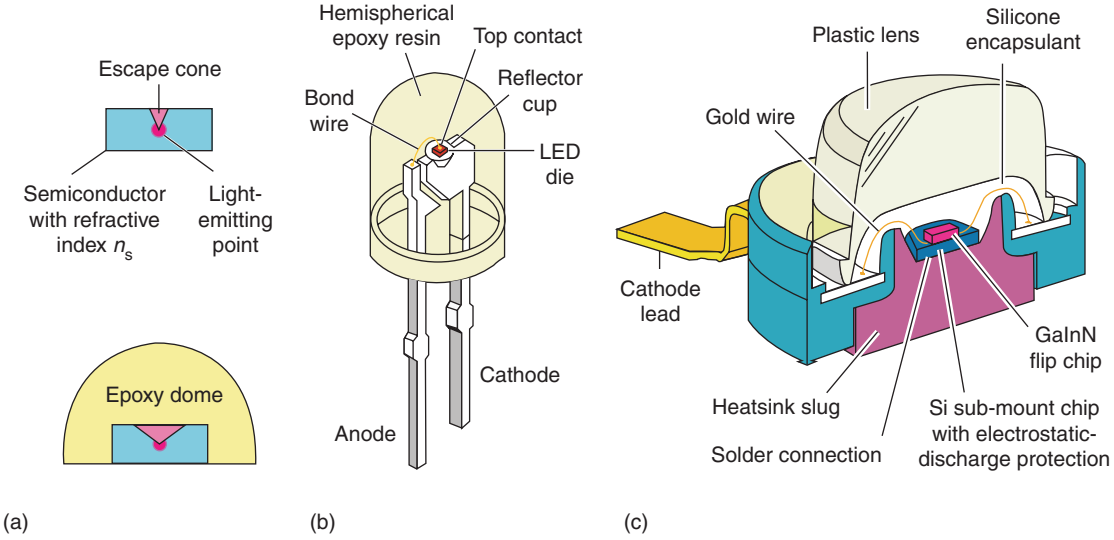


Figure 8 (a) Influence of an encapsulant on the escape cone. (b) Conventional LED in 5 mm package with hemispherical epoxy dome. (c) Package of high-power LED. (After Krames (2003) Overview of current status and recent progress of LED technology. *US Department of Energy Workshop* "*Solid State Lighting – Illuminating the Challenges*," Crystal City, VA.)

Epoxy remains transparent and does not show signs of degradation over a period of many years for long-wavelength visible-spectrum and IR LEDs. However, epoxies lose transparency in LEDs emitting at shorter wavelengths, namely in the violet and UV range. Furthermore, when exposed to temperatures exceeding 120°C for a sustained period of time, epoxies lose transparency and become yellowish.

In epi-side up GaInN LEDs grown on sapphire substrates with two side-by-side contacts, a large-area *p*-type top contact that covers the entire *p*-type mesa results in a uniform current distribution in the active region. However, such a large top contact will hinder the extraction of light. This fundamental trade-off can be avoided by flip-chip packaging where the LED is mounted with its topside facing a holder with good thermal and electrical conductivity. Flip-chip packaging is particularly advantageous for high-power devices.

A high-power LED package is shown in **Figure 8c**. The package has several advanced features including an Al or Cu heatsink slug that provides a low thermal resistivity of the package, a highly transparent silicone encapsulant that is chemically more stable than epoxy resin, and an electrostatic discharge protection (ESD) circuit integrated into a silicon submount.

The thermal resistance of LED packages together with the maximum temperature of operation determines the maximum thermal power that can be dissipated in the package. The maximum temperature of operation may be determined by reliability considerations, by the degradation of the encapsulant, and by internal-quantum-efficiency considerations. The 5 mm packages used for low-power LEDs (see **Figure 8b**) have a high thermal resistance of about $250\,\mathrm{K\,W^{-1}}$. Packages using heatsink slugs that transfer heat from the chip directly to a printed circuit board (PCB) which in turn spreads the heat, have thermal resistances of $6–12\,\mathrm{K\,W^{-1}}$.

Organic Light Emitting Diodes

Besides light emitting diodes based on inorganic materials, organic materials such as polymers and single crystals composed of organic molecules have been used as light emitters. Organic materials can display strong photo- and electroluminescence with quantum efficiencies of several percent. Organic light emitting diodes (OLEDs) have been used in active-matrix displays. With increasing efficiency, OLEDs could also become candidates for solid-state lighting applications. A schematic OLED structure consisting of a substrate, a transparent anode, organic electron and hole transport layers (ETL and HTL, respectively), and a low-workfunction metal cathode is shown in **Figure 9a**.

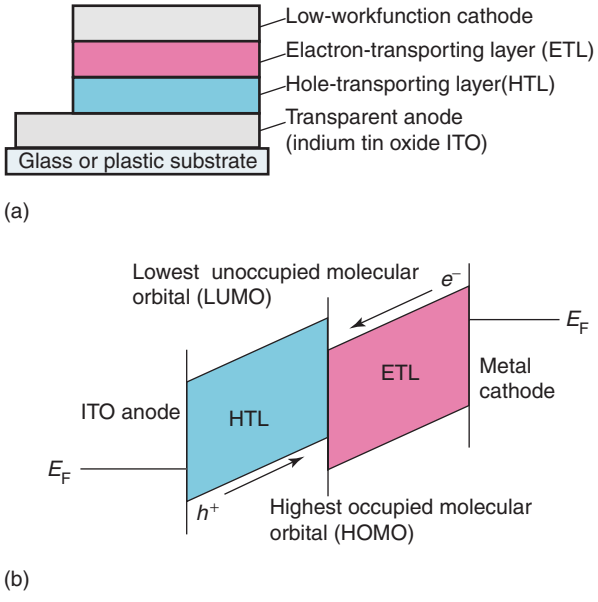


Figure 9 (a) Schematic structure of an OLED and (b) band diagram of an OLED in operation. (After Shinar (ed.) (2002) *Organic Light-Emitting Devices*, pp. 1–41. New York: Springer.)

Several organic materials are strongly luminescent such as the molecule tris-(8-hydroxy quinoline), Al (Alq_3), or the polymer poly(*p*-phenylene vinylene) (PPV). Many of the luminescent organic materials are π-conjugated compounds, in which p_z and p_y orbitals of neighboring C-atoms overlap and form π-orbitals and π-bonds. Due to the overlap, electrons in π-orbitals are delocalized along the bond axis between the two C-atoms. In organic polymers, the delocalized π-electron cloud extends along the entire length of the chain. The energy of electrons in π-orbitals is usually higher than the energies of electrons in σ orbitals. The gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is usually between 1.5 and 3 eV (see **Figure 9b**). Both HOMO and LUMO are π orbitals.

Light-emission in OLEDs is caused by radiative transitions between a singlet-exciton excited state and the singlet-exciton ground state or by a radiative decay of the triplet excitons into the singlet ground state (phosphorescence). However, competing processes such as the decay of singlet states into polarons (charge transfer) or the photooxidative formation of carbonyl groups acting as singlet-quenching centers may drastically reduce the luminescence yield.

Under the influence of an electric field, holes and electrodes are injected into the organic material from

the anode and cathode, respectively (see **Figure 9b**). In most cases, the injection of electrons from the metal contact is less efficient than hole injection through the ITO anode resulting in vastly different carrier concentrations and in reduced device efficiencies.

Transport of carriers inside the OLED structure occurs through inter-site hopping between localized states forming narrow bands around the HOMO and LUMO levels. Electrons and holes crossing the organic HTL/ETL interface into the HTL and ETL, respectively, experience a significant reduction of their mobility. This carrier deceleration results in a charge accumulation near the interface and light generation in the near-interface region. However, strong carrier-concentration gradients at the interface can also lead to local electric fields strong enough to dissociate singlet excitons and decrease the emission efficiency.

OLEDs emitting in the red-to-green wavelength region with lifetimes exceeding 10^4 h (at low current densities) and external quantum efficiencies $>4\%$ have been demonstrated. White and blue OLEDs with lifetimes around 10^3 hours and similar efficiencies have been demonstrated as well. The external quantum efficiency is expected to increase so that organic light emitters may become attractive for an increasing number of applications.

Acknowledgments

The authors gratefully acknowledge Dr. M R Krames and Lumileds Lighting, US LLC, San Jose, CA for granting permission to reproduce **Figure 8c.**

See also: Metals and Metallic Alloys, Optical Properties of; Optical Absorption and Reflectance; Semiconductor Lasers; Semiconductor Optics.

PACS: 78.20. – e; 78.60.Fi; 78.66.Fd; 85.60.Jb

Further Reading

Gessmann T and Schubert EF (2004) High-efficiency AlGaInP light-emitting diodes for solid-state lighting applications. *Journal of Applied Physics* 95: 2203.

Hack M, Kwong R, Weaver MS, Lu M, and Brown JJ (2002) Active-matrix technology for high-efficiency OLED displays. *Journal of the Society of Information Displays* 10(4): 323.

Krames MR, Amano H, Brown JJ, and Heremans PL (2002) High efficiency light-emitting diodes. *IEEE Journal on Selected Topics in Quantum Electronics* 8(2).

Nakamura S and Chichibu SF (2000) *Introduction to Nitride Semiconductor Blue Lasers and Light Emitting Diodes*. Boca Raton: Taylor and Francis.

Schubert EF (2003) *Light-Emitting Diodes*. Cambridge: Cambridge University Press (www.lightemittingdiodes.org).

Stringfellow GB and Craford MG (eds.) (1997) *High-brightness light-emitting diodes, Semiconductor and Semimetals*, vol. 48. San Diego: Academic Press.

Linearized Methods *See* Lattice Dynamics: Vibrational Modes.

Liquid Crystals

S Kumar and S-W Kang, Kent State University, Kent, OH, USA

Introduction

Liquid crystals (LCs) represent a state of matter intermediate between crystalline solids, which are characterized by a regular periodic arrangement of atoms or molecules, and isotropic liquids, which lack any type of order. They exhibit varying degree of orientational and/or translational order in specific spatial directions. For example, in the LC phase known as the smectic-A (SmA), with molecules arranged in stacked layers, one-dimensional translational order similar to a crystal exists in the stacking direction. However, within the layers, the molecular arrangement is liquid-like. Such a phase that is "liquid" in the two in-plane directions and "crystal" in the third, is one example of a large number of LC phases.

LC phases are formed by materials made of anisotropic molecules or other entities, such as aggregates of molecules in surfactant solutions known as micelles. LC phases formed by surfactant solutions are known as lyotropic liquid crystals and their discussion is beyond the scope of this article. The simplest of the LC phases, the nematic (N) phase, is formed when the molecules are oriented, on average, parallel to a common direction represented by an apolar unit vector, $\boldsymbol{n}$, referred to as the "director." The structure of the phases with degrees of order higher than the N phase is different for different molecular shapes. Different nomenclatures are introduced here to describe them. For the purpose of this article, the N phases are described first, followed by other phases exhibited by rod-like, disk-like, and then bent-core (banana) shaped molecules. Although

more than 50 LC phases are known, the focus here is only on those that are of fundamental scientific interest, commonly encountered, and/or have potential for technological applications. Classification of various phases is usually based on the concepts of symmetry relevant for liquid crystals such as orientational order, translational order, bond-orientational order, chirality, and dipolar ordering. The bond orientational order (BOO) refers to the degree to which crystal axes are parallel to each other over large distance. BOO can exist even when the lattice periodicity does not extend over large distances.

Nematic Phases

In the simplest nematic phase, rod-like molecules are oriented with their long axis on average parallel to the director, $\boldsymbol{n}$. The degree of order is described by a symmetric, diagonal, traceless, and real tensor order parameter $\boldsymbol{Q}$:

$$\boldsymbol{Q} = \begin{bmatrix} -\frac{S+\eta}{2} & 0 & 0 \\ 0 & -\frac{S-\eta}{2} & 0 \\ 0 & 0 & S \end{bmatrix}$$

When η is zero, it describes the uniaxial N phase, **Figure 1**, which may comprise of rod-like, disk-like, or bent-core molecules. In the uniaxial N phase, orientational order exists only along $\boldsymbol{n}$. The two directions orthogonal to $\boldsymbol{n}$ are degenerate. The order parameter S then measures how well the symmetry axes of the molecules are aligned. If the long axes of individual molecules are assumed to make an angle θ with respect to $\boldsymbol{n}$, then S reduces to

$$S = 1/2\langle 3\cos^2\theta - 1\rangle$$

where $\langle \cdots \rangle$ specifies thermal average. Optically, such phases are described by two indices of extraordinary and ordinary refraction, n_e and n_o in the directions parallel and perpendicular to $\boldsymbol{n}$, respectively.

Almost all LC displays used in commercial products such as laptop computers, hand-held devices, and flat-panel televisions employ the N phase. The ability to control the optic axis of this phase with the application of an electric field permits measured changes in the optical path length and the state of polarization of the transmitted light beam, and forms the basis of their operation.

If η is not zero, then $\boldsymbol{Q}$ represents a phase which has additional orientational order along a secondary director, $\boldsymbol{m}$, in the plane orthogonal to the primary director, $\boldsymbol{n}$. Such a phase, shown in **Figure 2**, is known as the biaxial nematic phase, N_{bx}, and it has three different indices of refraction along the three spatial directions. The parameter η measures the degree of biaxial order along the secondary director, $\boldsymbol{m}$.

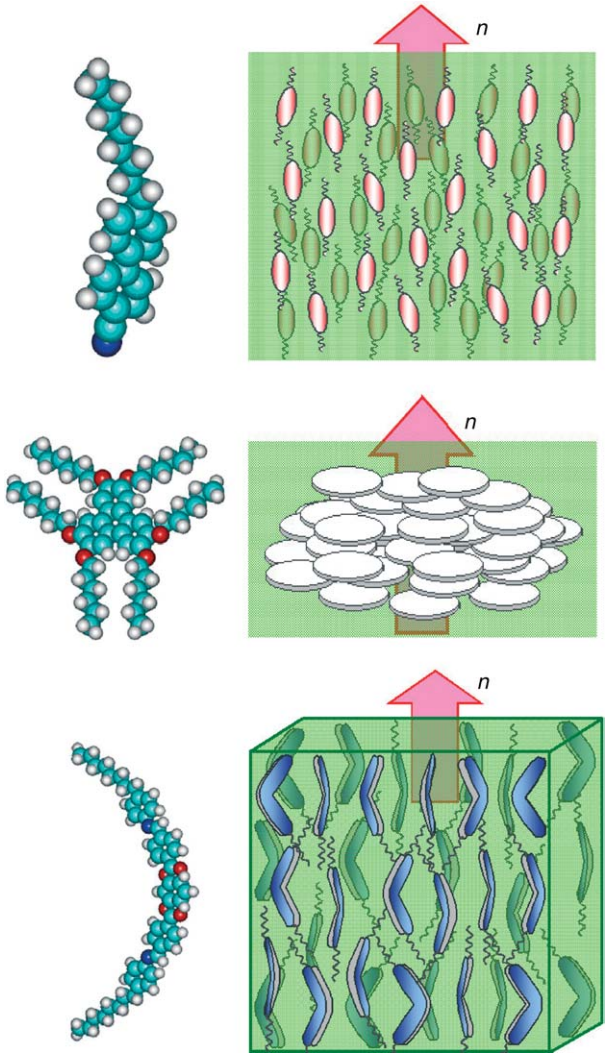


Figure 1 Schematic representation of molecular arrangement of (top to bottom) rod-like, discotic, and bent-core molecules in the uniaxial N phase.

The biaxial nematic phase was predicted on the basis of symmetry arguments in 1970 followed by a flurry of activity in theoretical and experimental areas. It was discovered in 1980 in lyotropic liquid crystals that, for a long time, have remained the only type of system to exhibit this phase. However, in monomeric (or molecular) systems, there were only false reports of its discovery. According to a group of experts who analyzed all prior reports of the existence of the N_{bx} phase in molecular LCs in 1998, the existence of this phase had not been substantiated. It is only very recently that its existence has been confirmed in the bent-core mesogens with X-ray

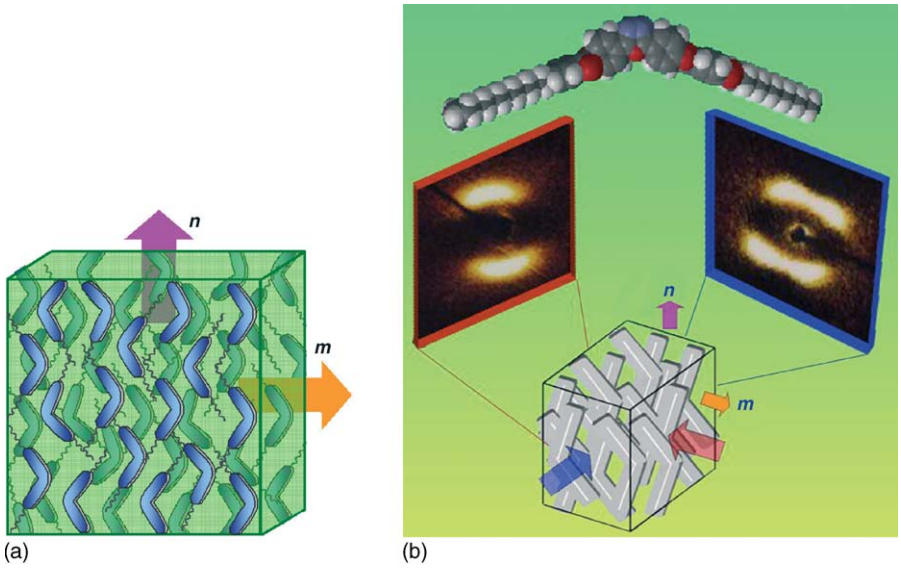


Figure 2 (a) Arrangement of bent-core molecules in the N_{bx} phase. (b) X-Ray diffraction patterns for the incidence directions (red and blue arrows) of X-rays with respect to the unit directions ***n*** and ***m***. Rubbed substrates were used to define ***n*** while an applied electric field controlled the orientation of ***m***.

diffraction on samples in which surface treatment and electric field were used to align the primary and secondary directors, respectively. The structure of this phase and experimental situations that confirmed the biaxiality is schematically shown in **Figure 2**. Subsequent to this discovery, a more careful study of the N phases formed by new bent-core materials has revealed the existence of both uniaxial and biaxial nematic phases at different temperatures in several single-component systems.

A special case of the uniaxial N phase arises when it is constituted of chiral (lacking mirror symmetry) molecules or when a small amount of chiral additive is dissolved in the N phase. This phase is known as the cholesteric phase as it was first observed in the derivatives of cholesterol. In the cholesteric phase, the unit vector ***n*** describes a helical trajectory of a well-defined pitch (**Figure 3**). Since there is no directionality to the vector ***n***, the molecular orientation repeats for every 1/2 pitch length or π-rotation. The pitch is, in general, strongly temperature dependent and often comparable to the wavelength of visible light. Cholesteric phases exhibit brilliant colors due to Bragg diffraction of light at the corresponding pitch. Thus, they show different colors of light at different temperatures. This property forms the basis of the "mood rings." (Mood rings are sold in gift shops. They change color as the wearer's body temperature (or, mood) changes.) The director for the cholesteric phase, in its component form, is written as $\boldsymbol{n} = [\cos(q_o z + \phi), \sin(q_o z + \phi), 0]$.

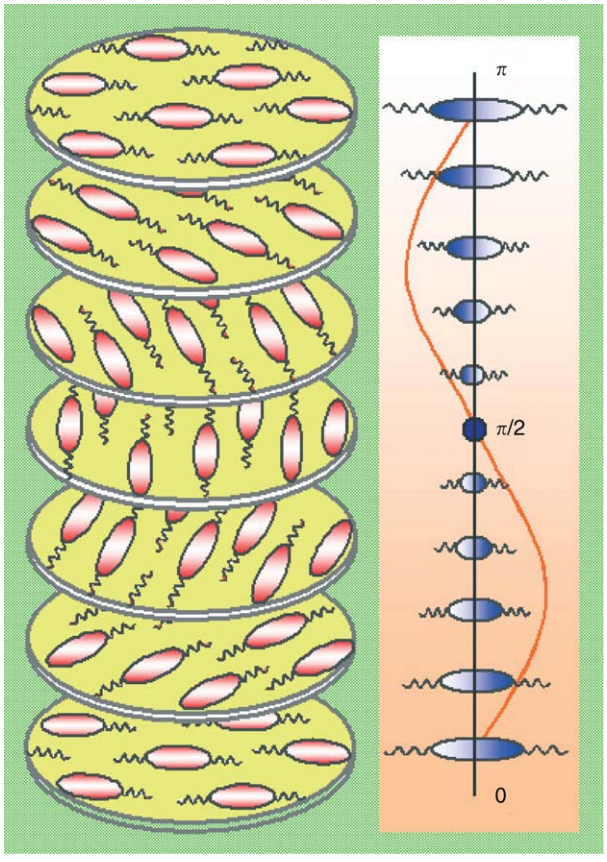


Figure 3 The orientation of molecules describes a helical trajectory in a cholesteric phase. Bragg reflections in optical regime occur for wavelengths equal to 1/2 of the pitch length.

Calamitic Mesophases

Molecules with cylindrical shape can form, in addition to the N and cholesteric phases, a class of phases that are known as the calamitic mesophases. These are layered structures referred to as the "smectic" phases. They all possess translational order in the direction perpendicular to the layers. In the two in-layer directions, the molecules can possess varying degrees of translational and bond orientational orders to form a large number of smectic phases. These phases are denoted with a different letter in the order of their discovery.

The SmA phase is characterized by a liquid-like order of molecules within the layers. As illustrated in **Figure 4**, the centers of mass of molecules are arranged in a periodic manner. The SmA phase is described by a two-component order parameter,

$$\Psi(\boldsymbol{r}) = \Psi_0 \exp(qz + \phi)$$

where Ψ_0 is the amplitude, ϕ is a phase factor, and the wave vector $\boldsymbol{q}$ is taken to be parallel to the z-direction. Within the layers, the molecules are oriented approximately parallel to each other defining the director $\boldsymbol{n}$ that is coincident with the layer normal, z.

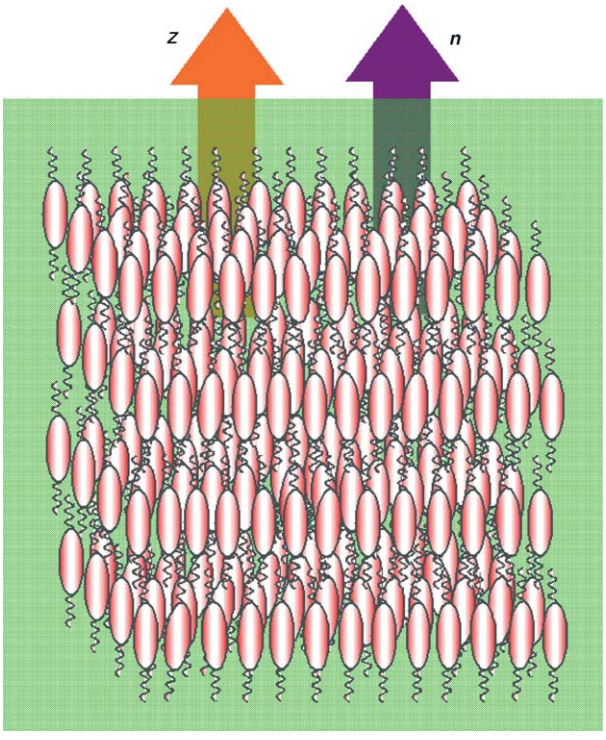


Figure 4 In the SmA phase, molecules are almost perfectly parallel to each other, $\boldsymbol{n}$, and the layer normal z. Within a smectic layer, they have no positional or bond-orientational order.

The smectic-C (SmC) phase, **Figure 5**, is very similar to the SmA phase and has layers with liquid-like molecular order. The major difference is that the molecules are tilted with respect to the layer normal. The molecular tilt in the SmC phase, or the angle between $\boldsymbol{n}$ and layer normal z, is a strong function of temperature as shown for the compound terephthal-bis-butylaniline (TBBA) in **Figure 6**.

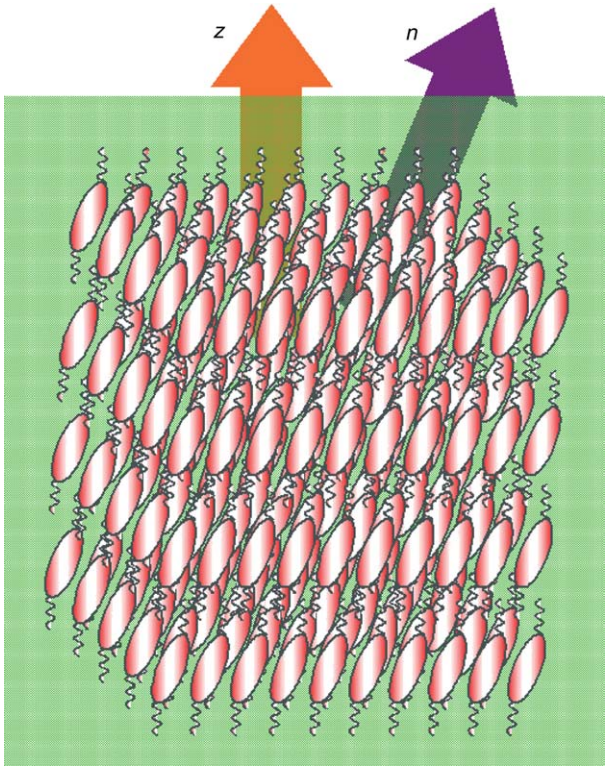


Figure 5 In the smectic-C phase, the director $\boldsymbol{n}$ along which the molecules are oriented, is oblique with respect to the layer normal z. Within a layer, the arrangement of molecules is liquid-like.

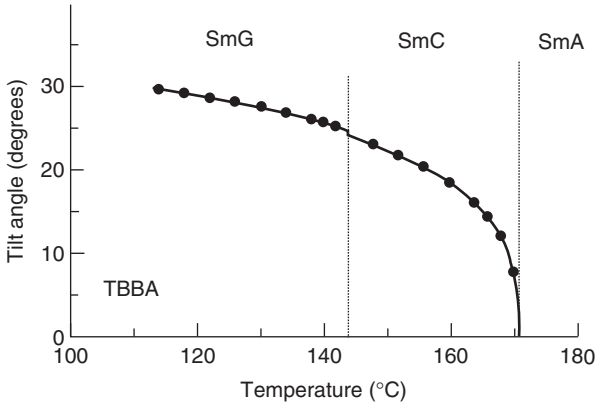


Figure 6 Temperature dependence of molecular tilt angle in the SmC and SmG phases of TBBA. The tilt continuously goes to zero as the SmA phase is approached.

Other more ordered smectic phases possess different degrees of bond-orientational and positional order within the smectic layers. Exotic calamitic, hexatic, and crystalline smectic phases with long-range bond-orientational and/or positional order are known to exist with and without molecular tilt with respect to the layers. These are not discussed in this article.

Ferroelectric and antiferroelectric smectic phases are formed by molecules lacking reflection symmetry and having a permanent molecular dipole moment directed nearly perpendicularly to the molecular long axis. These phases are very similar to the SmC phase with added chirality along the direction parallel to the layer normal. Molecular dipoles within a smectic layer are all oriented in the same direction, that is, in a ferroelectric manner. This chiral ferroelectric smectic phase is denoted as the SmC^* phase (**Figure 7a**), here "*" implies the chiral nature. The SmC^* phase possesses macroscopic spontaneous polarization density ***P***, the orientation of which can be manipulated with the help of a torque exerted on molecular dipoles by an applied electric field. Electrooptical devices based on the SmC^* phase have been shown to respond with a time constant of a few microseconds, about a thousand times faster than nematic devices.

When the dipole ordering in adjacent smectic layers is antiparallel, as shown in **Figure 7b**, the phase is known as the antiferroelectric SmC^* phase or SmC^*_A phase. The structure is still helical, but here azimuthal orientation of molecules in two layers with oppositely directed ***P*** describes the helical trajectory. Other versions of such phases are also known and can be found in the literature.

The twist-grain boundary (TGB) phase is an analog of the Abrikosov phase in a type-II superconductor, and is made of blocks of smectic layers, with no in-plane order, permeated by regions of twist deformation, **Figure 8**. Depending on the absence (SmA) or presence (SmC) of molecular tilt with respect to the layer normal, one obtains the TGB_A or

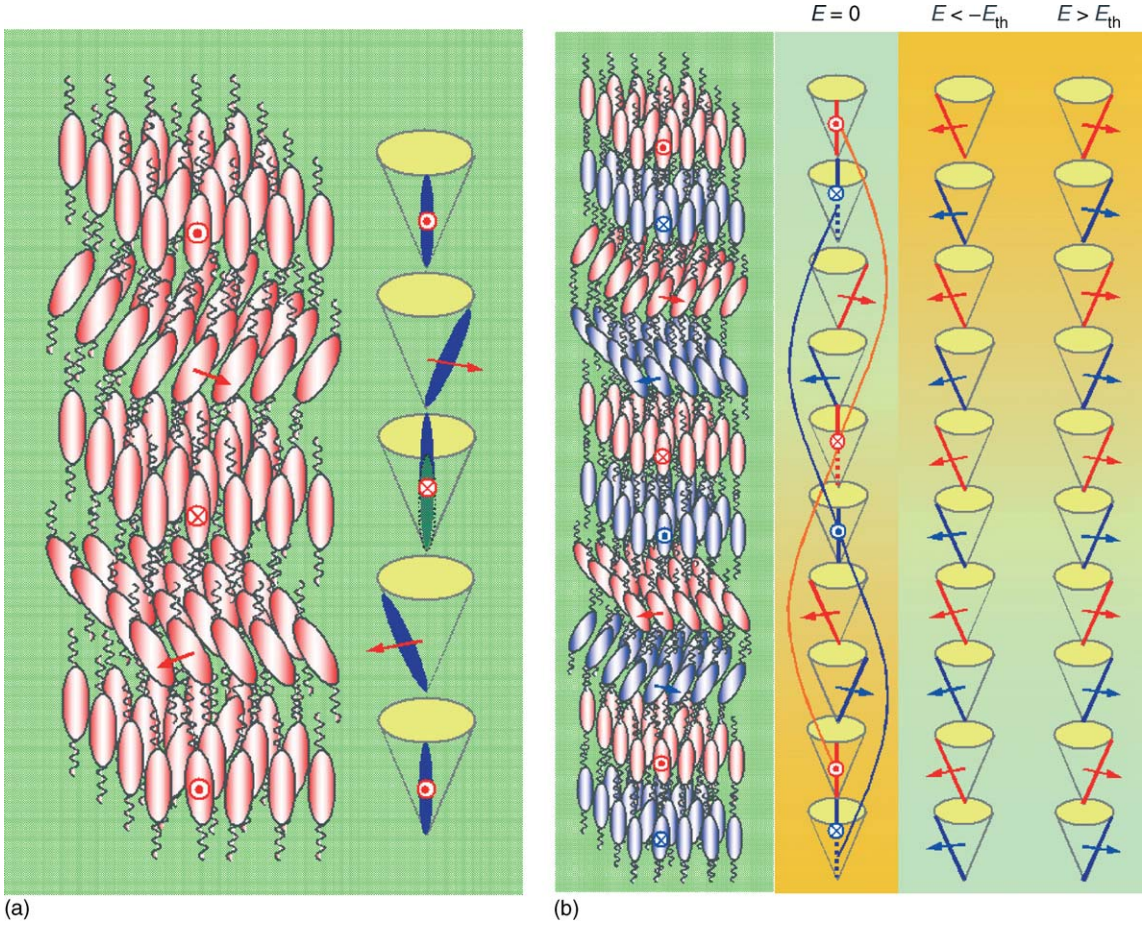


Figure 7 (a) Ferroelectric SmC^* phase. (b) Antiferroelectric SmC^*_A phase in which oppositely tilted molecules describe double helical trajectory. Red and blue arrows represent molecular dipoles. The antiferroelectric phase reverts to ferroelectric phase when a field higher than a threshold is applied. The orientation of the molecular dipoles depends on the direction of the field.

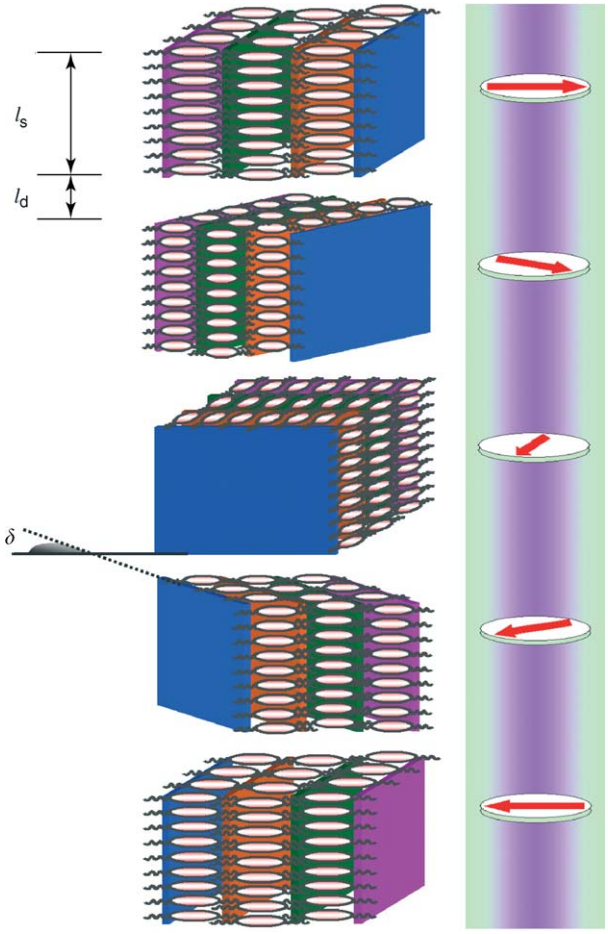


Figure 8 Schematic drawing of TGB_A phase with pitch axis vertical. The red arrows mark the direction of layer normal.

TGB_C phases. In these phases, a twist axis exists in the direction of smectic planes. Consequently, the phase appears to be a string of smectic-like slabs of a fixed dimension l_s, and arranged adjacent to each other with a narrow (of length l_d) region of twist defect. When the angle of twist, δ, between neighboring slabs is such that $2\pi/\delta$ is an integer (or, a rational number), it is referred to as the "commensurate" TGB phase and X-ray diffraction pattern of an oriented sample features a discrete set of reflections at the same magnitude of momentum transfer vector. However, if $2\pi/\delta$ is an irrational number, then the X-ray pattern consists of a continuous diffraction ring and it is referred to as the "incommensurate" TGB phase.

Discotic or Columnar Mesophases

Discotic mesophases are exhibited by somewhat larger molecules that have nearly flat disk-like cores. These molecules naturally like to pack by forming stacks (or, columns) of cores as illustrated in

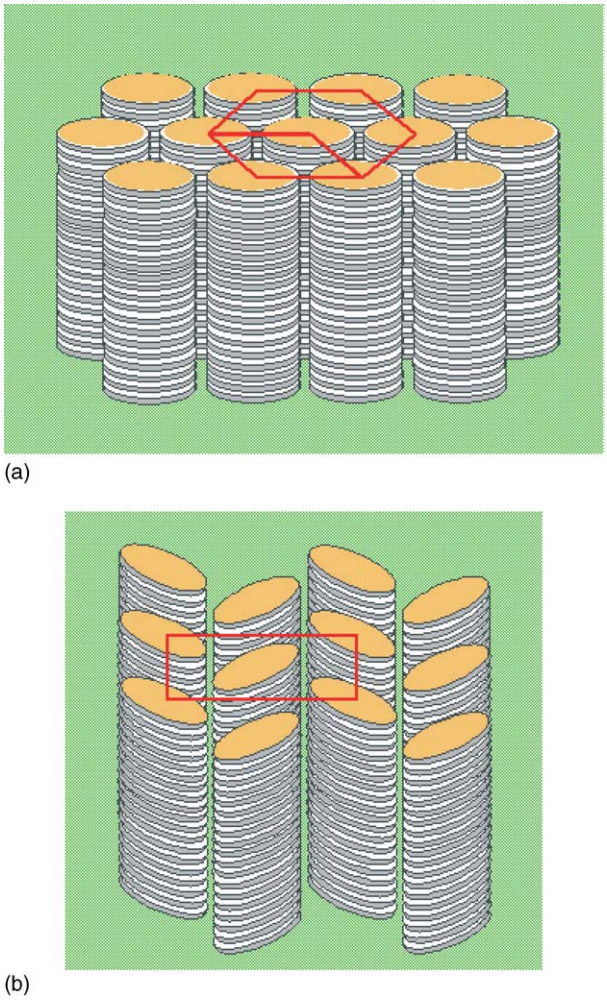


Figure 9 (a) Hexagonal D_h and (b) rectangular D_r columnar mesophases. Discotic mesophases lack positional order in the direction of column axis.

Figure 9. The alkyl chains are not shown for clarity. The columns of molecules form two-dimensional lattices depending on the details of molecular shape and intermolecular interactions. There is no long-range positional order in the direction of the column axis. They are crystal in two dimensions and liquid in one, in contrast to the calamitic mesophases.

Depending on the 2D space group corresponding to the arrangement of columns, these are given different names. In the hexagonal (D_h) phase, the disks are arranged in hexagonal close-packed manner and the plane of the molecular cores is perpendicular to the column axis. The cores can also be tilted with respect to the column axis and permit rectangular (D_r) or rectangular face-centered phase molecular arrangement. In the D_r phase, the core in a column is tilted with respect to the cores in neighboring columns, forming a herringbone arrangement. The tilted cores are shown as ellipses in **Figure 9b**.

Bent-Core Mesophases

Bent-core (or, banana) mesophases, as the name suggests, are formed by molecules resembling a boomerang (see **Figures 1** and **10**). In recent years, the discovery of novel mesophases formed by the bent-core molecules has engendered much scientific interest in them. The phases exhibited by bent-core molecules are denoted as B_i, where $i = 1,2,\ldots,8$, etc. The progression of the index i does not in any way imply a systematic evolution of the structure of the various phases. The structure of some of the phases has been determined while it remains uncertain for others. In the following, what is known at this point in time is summarized. The structure of different phases is discussed in an order to minimize confusion. This is an active and rapidly growing area of research and it is likely that the current understanding of some of these phases will undergo revision/refinements. Their names and notations are very likely to change in the near future as one learns more about them and streamlines the nomenclature on the basis of the symmetry.

Owing to their unique geometrical shape, the phases have revealed surprising structures and properties. For example, the formation of polar and chiral structures by achiral molecules is unexpected. Bent-core materials show the largest electronic second-order susceptibility $\chi^{(2)}$ of all known LCs. Until now, the formation of ferroelectric phases with macroscopic electric polarization was limited to chiral molecules. Some of the bent-core mesophases exhibit interesting ferro- and anti-ferroelectric properties.

In the B_1 phase, the molecules partially overlap along their overall long axis and they are arranged in smectic layers. Additionally, in a plane perpendicular to the smectic layers, molecules are organized in a 2D centered rectangular lattice, **Figure 10**. From studies of single domain samples, there are no indications of positional correlations in the direction perpendicular to the plane containing the 2D lattice. From a symmetry point of view, this phase resembles the columnar D_r phase. Unit cell dimensions, as expected, are related to molecular size. For example, in case of a series of bent-core azo compounds (i.e., those containing the divalent radical N:N), the rectangular cell dimensions are $a = 36.4 \pm 0.2$ Å and $c = 30.4 \pm 0.2$ Å. Two forms of the B_1 phase have been reported; in one, the molecules are orthogonal to the smectic layers while in the second form, they are tilted in a manner similar to the SmC phase. These have been assigned the names B_{1o} and B_{1t} for the orthogonal and tilted phases, respectively.

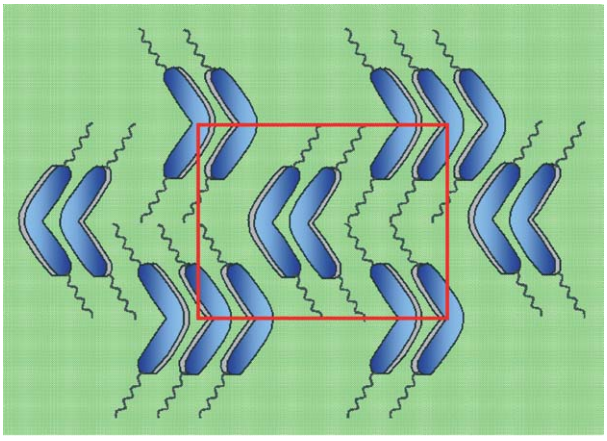


Figure 10 Molecules in the B_1 phase occupy a rectangular lattice in a plane perpendicular to smectic layers. No long-range positional correlations exist in the direction perpendicular to the rectangle.

A group of four phases which are stacked layers (or smectic) is known as the B_2 phase(s). Here, the molecules are tilted and they resemble the tilted chiral phases (SmC^*, SmC^*_A) of calamitic mesogens. The molecules forming these phases may possess a dipole p directed along their apex. Furthermore, the tilt in adjacent layers may be in the same azimuthal direction (synclinic, S) or in opposite directions (anticlinic, A). The direction of the molecular dipoles in one smectic layer may be the same as in the adjacent layers (ferroelectric, F) or antiparallel (antiferroelectric, A). These four possibilities are schematically shown in **Figure 11** and denoted as SmC_AP_F, $SmC_SP^*_F$, SmC_SP_A, and $SmC_AP^*_A$. It should be noted that the $SmC_SP^*_F$ and $SmC_AP^*_A$ phases essentially possess the same symmetry as the calamitic ferroelectric SmC^* and antiferroelectric SmC^*_A phases, respectively, and they may exist in both left- and right-handed chiral states. The other two phases, SmC_SP_A and SmC_AP_F, are antiferroelectric and ferroelectric racemates, respectively.

The B_3 and B_4 phases are actually crystalline phases. In the literature, the B_4 phase is also referred to as the blue crystal phase. This phase is chiral and exhibits circular dichroism. Only preliminary details about the structure of the B_3 and B_4 phases are available at this time. These phases are not discussed here.

The B_6 and B_8 mesophases possess smectic stacked-layer organization with no in-plane order. The B_6 and B_8 phases differ from the regular smectic phases in an important way. The smectic layer spacing is different from molecular length because in the B_6 phase, the molecules have partial lengthwise overlap at the smectic layer interface. However, unlike the B_1 phase, in the B_6 and B_8 phases, **Figure 12**, there is no other order present apart from the smectic layering. The primary layer thickness is less than the

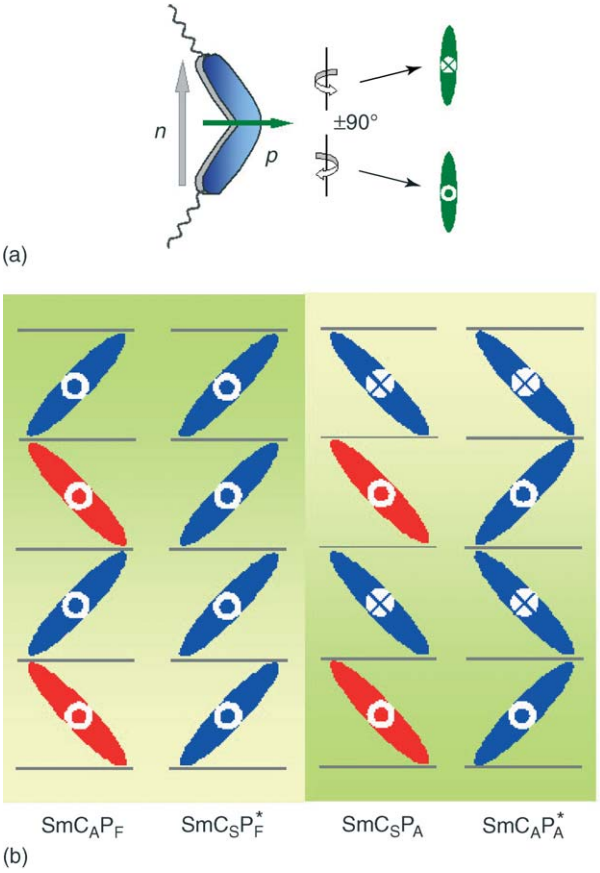


Figure 11 (a) Different views of bent-core molecules and (b) the four structures formed in the B_2 phase. In some materials, more than one structure may be simultaneously present.

molecular length. Such phases exhibit additional peaks in the X-ray diffraction pattern, corresponding to a length that is less than twice the molecular length. A tilted analog phase of the B_6 phase has also been reported. In the B_8 phase, the molecules form a weak antiparallel association with molecules in the adjacent layers and act as dimers with no overlap of the cores. The smectic layer spacing in this phase is close to twice the molecular length. The B_6 and B_8 phases can be considered as analogs of the SmA_d and SmA_2 phases in the frustrated (or polar) smectics. The B_8 phase presents an antiferroelectric arrangement of molecular dipoles. Tilted and chiral variants of the B_6 and B_8 phases are also likely to exist.

The B_7 phase exhibits chiral textures of both right and left handedness under a polarizing microscope. There are some reports of this phase responding to an applied electric field in both ferro- and antiferroelectric manner. The B_7 phase has a layered structure with either the same or different tilts in adjacent layers, which are either ferroelectric or antiferroelectric. They exhibit twisted helical filaments, high- and low-birefringence focal conics striped patterns

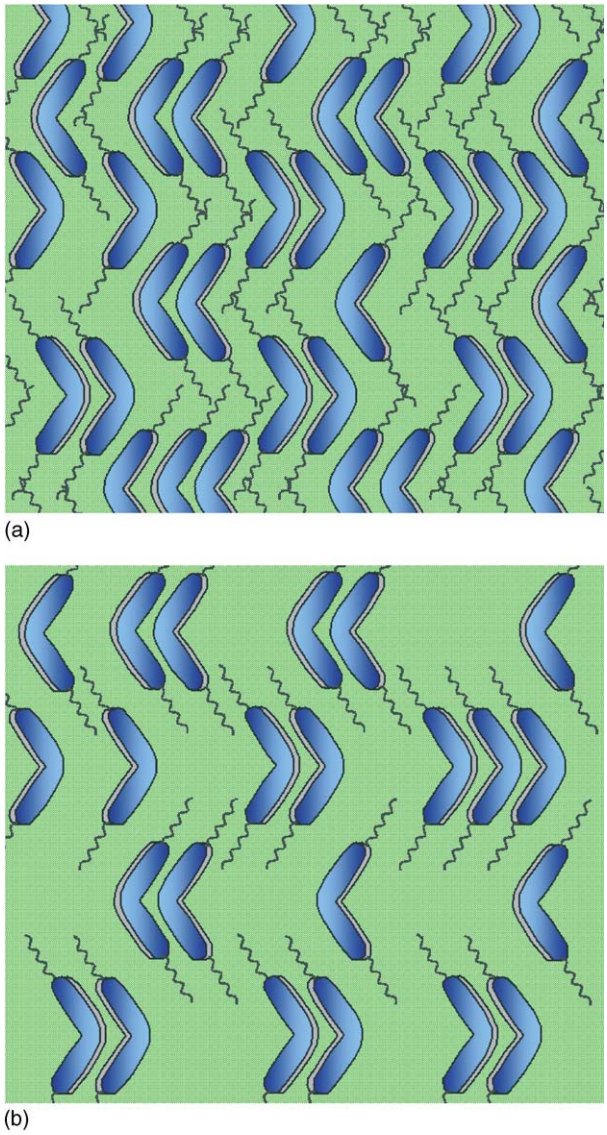


Figure 12 (a) Partial bilayer structure of the B_6 and (b) B_8 phase.

parallel or normal to the smectic layers. Unraveling the structure of this phase has been the most challenging of all bent-core mesophases. Optical, freeze fracture, and high-quality synchrotron X-ray diffraction studies of thin and bulk samples are beginning to reveal a complex 3D structure. In the B_7 phase, SmC^*-type layers are subdivided into stripes with splay deformation and defect lines shown as red and blue vertical lines in **Figure 13**. Splay deformations naturally cause smectic layer (thickness) undulations because the disorder of ***P*** and molecular tilt produces a smaller molecular tilt and thus higher layer spacing. For a specific sign of splay, there are four possible combinations of handedness and orientations of ***P*** across smectic layers. These four

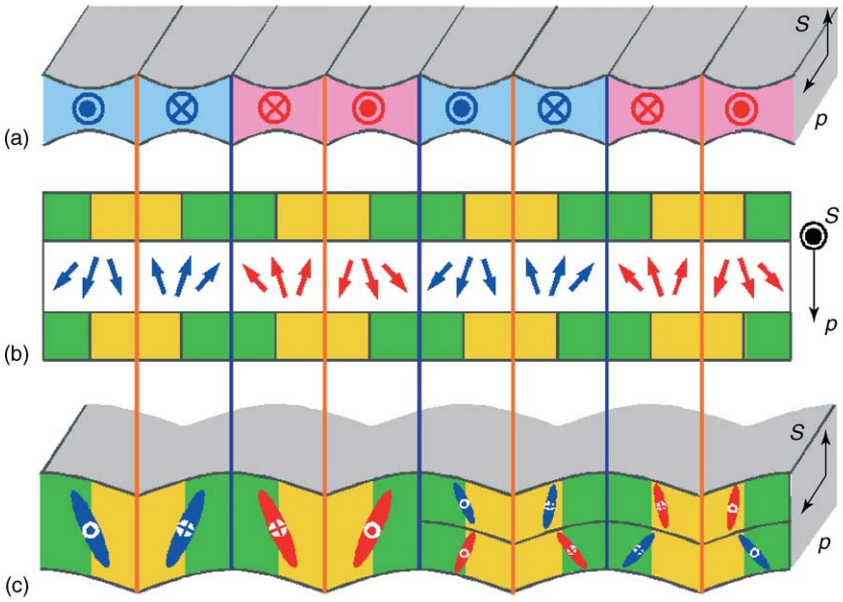


Figure 13 (a) Variation of the layer spacing due to polarization modulation induced by the formation of splay stripes shown in (b). (c) Different regions separated by defect lines possess local structures similar to the SmCSPF and SmCAPF structures of the B_2 phase. Vectors ***P*** and ***S*** represent the polarization and average layer normal directions, respectively. Arrows in (b) and vector directions in (a) and (c) represent the molecular dipoles and their splay organization.

combinations are exactly the same as already discussed for the B_2 phase (**Figure 11**). A general depiction of the essential feature of layer undulations and molecular arrangement is illustrated in **Figure 13**. The uniformity of the pitch in the layer-normal direction on a macroscopic scale requires the thinner layer regions between the defect (vertical blue and red) lines to tilt and become thicker. This results in a polarization-modulated structure with a pitch much larger than the smectic layer spacing. X-ray studies reveal that the period of modulations of the smectic layers is very well defined, giving rise to sharp Bragg peaks in the transverse direction. The undulations have a temperature- and material-dependent period ranging from $\sim$15 to 35 nm. The length over which the undulations are correlated is found to be the same as the dimension over which the smectic layers prevail, that is, undulations are long-range deformations and an integral feature of the B_7 phase.

The possibility of helical polarization-modulated filaments and twisted grain-boundary (TGB) structures with polarization modulation open the possibility of the formation of a number of unique defect structures which may be responsible for the basis of formation of a large variety of textures by the B_7 phase.

The B_5 phase is normally obtained upon cooling from the B_2 phase. The molecular arrangement in the B_5 phase is similar to the B_2 phase. The liquid-like X-ray diffraction peaks at large angles arising from the in-plane structure of the smectic layers are quite sharp. However, they remain reasonably wide to suggest that the in-plane structural order is not long range. This phase is very likely to be an analog of the orthogonal hexatic smectic (SmB_{hex}) phase of calamitic liquid crystals. The B_5 phase responds, albeit sluggishly, to an applied field confirming increased order and viscosity.

Many soft condensed matter systems and biological materials possess liquid crystalline structures. For example, rod-like objects such as tobacco mosaic virus on one hand and carbon nanotubes on the other, are known to form LC phases. With the ever expanding knowledge of naturally occurring phenomena and efforts to synthesize new molecular shapes, the list of liquid crystalline materials is bound to keep growing, revealing new structures and novel scientific phenomena.

See also: Crystal Structure Determination; Crystal Symmetry; Optical Properties of Materials; Scattering, Elastic (General).

PACS: 61.30. – v; 42.70.Df; 61.30.Gd; 83.80.Xz

Further Reading

Acharya BR, Primak A, Dingeman TJ, and Kumar S (2003) *Pramana* 61: 261.

Acharya BR, Primak A, and Kumar S (2004) *Physical Review Letters* 92: 145506.
Coleman D, Fernsler J, Chattam N, Nakata M, Takanishi Y, *et al.* (2003) *Science* 301: 1204.
de Gennes PG and Prost J (1993) *The Physics of Liquid Crystals*, 2nd edn. New York: Oxford University Press.
Eremin A, *et al.* (2002) *Liquid Crystals* 29: 775.
Freiser MJ (1970) *Physical Review Letters* 24: 1041.
Kang S-W, Reddy RA, Sadashiva BK, and Kumar S (2004) *20th International Liquid Crystal Conference*, Paper SYN-P141, Ljubljana, Slovenia, July 5.
Kumar S (2000) *Liquid Crystals – Experimental Study of Physical Properties and Phase Transitions*. Cambridge, UK: Cambridge University Press.
Link DR, Natale G, Shao R, Maclennan JE, Clark NA, *et al.* (1997) *Science* 278: 1924.
Niori T, Sekine T, Watanabe J, Furukawa T, and Takezoe H (1996) *Journal of Materials Chemistry* 6: 1231.
Pelz K, Weissflog W, Baumeister U, and Diele S (2003) *Liquid Crystals* 30: 1151.
Praefcke K, Blunk D, Singer D, Goodby JW, Toyne KJ, *et al.* (1998) *Molecular Crystals and Liquid Crystals* 323: 231.
Prasad V, Kang S-W, Qi X, and Kumar S (2004) *Journal of Materials Chemistry* 14: 1495.
Prost J (1984) *Advances in Physics* 33: 1.
Walba DM, Korblova E, Shao R, Maclennan JE, Link DR, *et al.* (2000) *Science* 288: 2181.
Yu LJ and Saupe A (1980) *Physical Review Letters* 45: 1000.

Liquids, Electronic Structure of

M P Tosi, Scuola Normale Superiore, Pisa, Italy

Introduction and Overview

Everyday experience testifies to the existence of three states of matter: solid, liquid, and gas. Solids are rigid, while liquids and gases flow under a shear stress however small. Crystals when studied by X-ray diffraction give rise to sharp Bragg reflections, while there are no sharp spots in the diffraction pattern when X-rays are scattered from a fluid or from a glass. This shows that long-range order among atoms or molecules is no longer present. However, the qualitative distinction drawn from the diffraction pattern does not carry through to distinguish between liquid and gas. Below a critical temperature and pressure these two phases can coexist at equilibrium with different densities, while above the critical point there is a single fluid phase. The difference between the two fluid phases lies in their density and this implies that the interatomic forces play a much greater role in the dense liquid than in the dilute gas.

A useful way to open an overview on the electronic structure of liquids is to recall the five types of crystal binding: covalent, hydrogen, ionic, metallic, and van der Waals bonding. Rare gas atoms such as argon are spherical and chemically saturated, so that the binding is due to van der Waals forces and the electronic structure is essentially unchanged on melting (here and in the following, liquids near the melting curve at standard pressure, except where explicitly noted, are considered). The only disorder that is possible in these fluids is related to translational motions – but when the building blocks are molecular units, the possibility of rotational disorder arises. Rotational order may disappear before melting as in plastic crystals such as N_2, or persist beyond the loss of translational order as in liquid crystals. Transitions may also occur in some molecular liquids on heating, such as a breakage of chains in Se or of eightfold rings in S.

It stands to reason that the electronic structure of a metal is essentially unchanged on melting. There are, however, a number of properties of liquid metals that intrinsically depend on the sea of conduction electrons, and these are the first item in this article. The article also describes how electrons may enter to determine the approach of a metal to the critical point, the transport properties of metals, the structure of some liquid alloys, and the phase diagram of elemental hydrogen in its various phases and of hydrogen–helium mixtures inside the giant planets.

In contrast to metals, dramatic changes in the electronic structure occur on melting in some covalent systems such as B, Si, Ge, and III–V compounds: the covalent bonds collapse into a metallic state, though with some peculiarities of its own. These elemental melts can be brought into an amorphous state by rapid quenching, just as those of other covalent compounds such as SiO_2 or $GeSe_2$ whose liquid structure is best described as a disordered covalent network.

Turning to ionic materials, their electronic structure is again quite stable across melting, although in some cases the state of aggregation of the ions may drastically change: for example, $AlCl_3$ melts from a layer structure where the Al ions are sixfold coordinated into a liquid of Al_2Cl_6 molecular dimers showing fourfold coordination of Al. The presence of chemical order in these melts can favor the acceptance of foreign electrons inside the liquid structure, as in

mixtures with metals that may be stable over a wide range of concentration.

The article ends with a brief reminder of how hydrogen bonding is formed in water and leads to the unique properties of this life-sustaining liquid.

Electron Theory of Liquid Metals

Interatomic Forces and Liquid Structure

Here, some properties of simple liquid metals with *s* and *p*-type conduction electrons, which reflect their nature as two-component liquids of ions and electrons, are reviewed. First of all, the effective interactions between pairs of ions in the metal involve screening of their bare Coulomb repulsions by the conduction electrons. For simplicity, an ion inside a fluid of interacting electrons as a static point charge Z_1e is considered. A second ion with charge Z_2e sees the potential of the first as well as the potential due to the polarization induced in the electron gas, so that the effective interaction potential of the two ions (in Fourier transform at wave number q) is

$$\phi_{12}(q) = \frac{4\pi Z_1 Z_2 e^2}{q^2 \varepsilon(q)}$$

with $\varepsilon(q)$ the dielectric function of the electron gas. The Thomas–Fermi theory affords a simple model for this function, yielding $\varepsilon(q) = 1 + (q\lambda_{TF})^{-2}$ where $\lambda_{TF}\sqrt{(\pi\hbar^2/4mk_Fe^2)}$ is the Thomas–Fermi screening length. Here $\hbar$ is Planck's constant, $k_F = (3\pi^2 n)^{1/3}$ the Fermi wave number of the electron gas, and n, m, and e the density, mass, and charge of the electrons.

As a simple example, the sound waves propagating in the metal are considered. In the absence of screening, the collective ionic motions at long wavelengths would be oscillations occurring at the ionic plasma frequency $\Omega_p = (4\pi NZ^2e^2/M)^{1/2}$. The effect of screening is to replace Z^2e^2 by $Z^2e^2/\varepsilon(q) \to Z^2e^2(q\lambda_{TF})^2$ at long wavelengths: screening thus yields an acoustic mode at frequency $\omega = c_s q$ with a value $c_s = v_F(Zm/3M)^{1/2}$ for the speed of sound. Here Z is the ionic valence, $N = n/Z$ the ionic density, M the ionic mass, and $v_F = k_F/m$ the Fermi velocity.

Classical theories of liquid structure aim to calculate the ion-pair distribution function and hence the diffraction pattern of the liquid from a given interatomic force law. Conversely, one may invoke the use of experimental diffraction data to extract an interionic pair potential by a careful analysis. **Figure 1** shows a comparison between such a "diffraction potential" from data on liquid Na and that calculated from electron theory. All major features of the diffraction potential are reproduced by the theory, though some quantitative discrepancies remain.

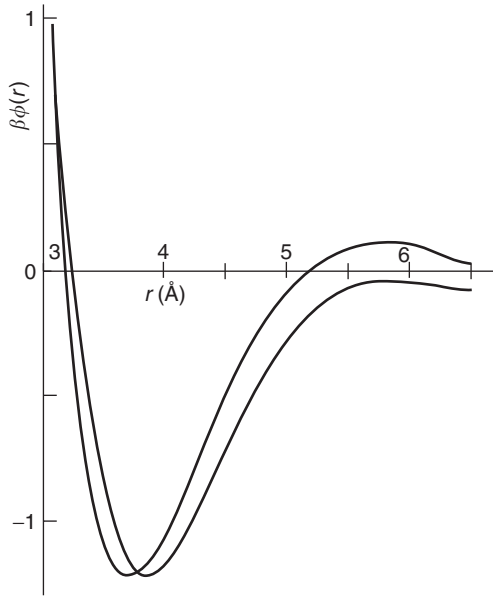


Figure 1 Effective pair potentials (multiplied by $\beta = (k_B T)^{-1}$) in liquid Na: from diffraction data (lower curve at large distance r) and from electron theory (upper curve at large r).

Expanded Fluid Cesium

Through its dependence on electron screening, the effective interionic pair potential in a liquid metal depends on the electron density in an important way. The detailed extensive work carried out in the 1980s and 1990s by Hensel and co-workers on expanded liquid alkali metals, and especially on liquid Cs, along the liquid–gas coexistence curve toward the critical point and has shown that, as the liquid density is reduced, the coordination number of each ion rapidly decreases whereas the near-neighbor distance stays remarkably constant. The coordination number of a Cs atom that is obtained from the data by extrapolation to the critical density approaches the value 2. This suggests that, as the metallic fluid is driven toward a metal–insulator transition, chain-like structures are stabilized by chemical bonding and probably coexist with small clusters such as ionized dimers. On the gas side of the critical point, one finds a large number of diamagnetic units such as Cs_2 molecules and probably larger clusters.

Transport Coefficients

The electrical resistivity ρ_e of a liquid metal in an ideally pure state is determined by the quasielastic scattering of conduction electrons on the Fermi surface against screened fluctuations in the ionic density.

The spectrum of these fluctuations as a function of wave number q and frequency ω is expressed in terms of the nucleus–nucleus dynamic structure factor $S(q,\omega)$ and, on exploiting the vast difference in the frequency scales for ionic and electronic motions, one obtains the result

$$\rho_e = \frac{3\pi m^2}{4Ne^2\hbar^3 k_F^6}\int_0^{2k_F} dq|\phi_{ie}(q)|^2 q^3 S(q)$$

Here, $\phi_{ie}(q)$ is the screened electron–ion potential and $S(q)$ is the structure factor which describes the measured diffraction pattern of the liquid metal. A comparison with the data on the electrical resistivity of liquid alkali metals as a function of temperature is shown in **Figure 2**.

The electronic contribution is also dominant in the thermal conductivity κ of metals. Indeed, the Wiedemann–Franz law asserts that the quantity $\rho_e\kappa/T$ is a universal constant, the Lorenz number $L = \pi^2 k_B^2/3e^2$ where k_B is the Boltzmann constant.

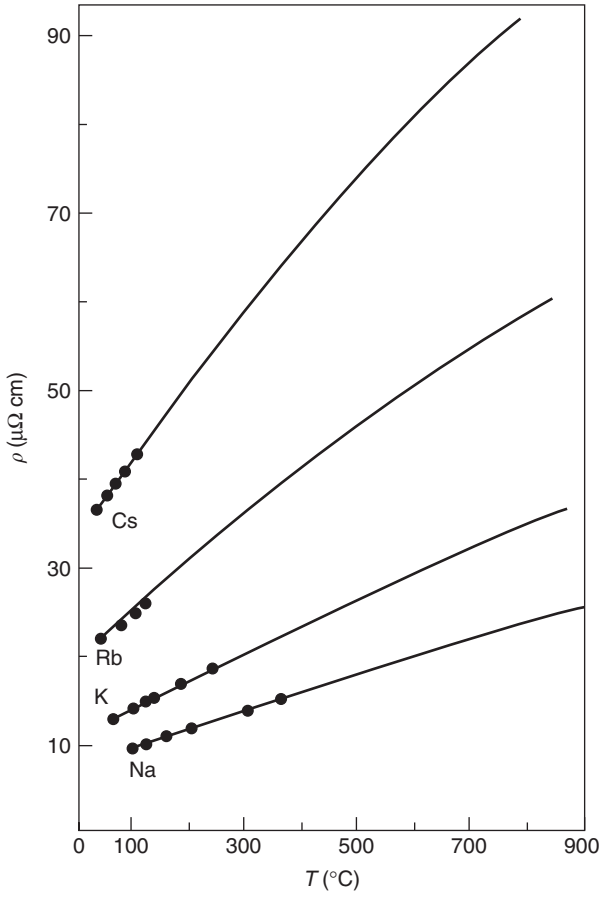


Figure 2 Electrical resistivity of liquid alkali metals as a function of temperature from electron theory (solid lines), compared with experimental data (dots).

Empirically, this law is satisfied by many liquid metals near freezing within a scatter of about $\pm 10\%$.

The electronic contribution is instead minor in the shear and bulk viscosities of simple liquid metals; The shear viscosity η is directly related to the properties of the ionic system, $\eta \propto T^{1/2}M^{1/2}N^{1/3}$, where all quantities refer to the melting point.

Chemical Short-Range Order in Some Metallic Alloys

In diffraction experiments, the melts of some alloys of metallic elements, such as Cs–Au or Li–Pb and other alloys of alkali metals with Pb or Sn, show a marked preference for unlike first neighbors near certain compositions. This behavior arises from electron transfer between the two elements, which is especially marked at such "stoichiometric" compositions. Minima in the electrical conductivity and spin susceptibility at these compositions reflect a minimum in the density of states at the Fermi level if not the opening of a gap due to full charge transfer.

The elements Cs and Au form a solid compound CsAu which is a strongly polar semiconductor with an optical bandgap of 2.6 eV. Its electrical conductivity drops on melting to a value of a few $(\Omega\,\mathrm{cm})^{-1}$, which is comparable to that of a molten salt. Electromigration experiments give evidence that in the melt, one Cs^+ ion is transported to the cathode and one Au^- ion to the anode per elementary charge to the electrodes. From X-ray diffraction, the Cs–Au first-neighbor distance at 0.36 nm can be followed away from stoichiometry up to 80% Cs, while the Cs–Cs distance of pure Cs at 0.53 nm starts emerging at 70% Cs. The electronic structure in the Cs-rich side of the phase diagram resembles that of metal–molten salt solutions (see below).

With regard to the melts of alloys of alkalis with group-IV metals, compound formation has been reported at the electronic octet composition for Li_4Pb and Li_4Sn, with chemical order extending over a range of 2 nm. Chemical order at equimolar composition in liquid alloys such as KPb and NaSn has been interpreted as due to the formation of tetrahedral Pb_4 or Sn_4 polyanions. In such a tetrahedral cluster, which is known in the solid state as a "Zintl ion," the p-type electron states of Pb, say, would be split into bonding and antibonding and the former could be filled by electron transfer from the alkali atoms.

Phase Diagram of Hydrogen and Hydrogen–Helium Mixtures in the Giant Planets

Hydrogen is the most abundant element in the Universe and has a very rich phase diagram, notwithstanding the simplest structure of the atom. In

standard laboratory conditions, cooling the gas of diatomic molecules brings it into a liquid of weakly coupled molecules and then into a crystal, in which the molecules are rotating with weak hindrance from each other. Compression at low temperature brings the crystal into a new phase, in which the molecular rotations are hindered and ultimately into a further crystal phase whose strong infrared activity indicates a mainly dipolar distortion of the electronic distribution. Molecular dissociation can occur as hydrogen is brought into the regions of high temperature and density which are relevant to astrophysics. A continuous transition from a semiconducting to a metallic state has been observed to occur at 3000 K and 140 GPa in dynamic compression experiments.

Theoretical studies of the phases of fluid hydrogen take account of several species (diatomic molecules, atoms, protons, and electrons) and examine their chemical equilibria. The results are in broad agreement with path-integral Monte Carlo studies by Ceperley and co-workers. A molecular gas is found to form spontaneously from a neutral system of protons and electrons as the temperature is lowered from 10^5 to 5000 K, with a molecular bond length which contracts with increasing density because of the stiff intermolecular repulsions. Molecular dissociation is then reported to occur as the temperature is raised or under isothermal compression. Thermally activated dissociation at high density is accompanied by decreasing pressure, signaling the presence of a first-order transition and a critical point. A proposed reason for the transition is the increase of electronic kinetic energy which is associated with bond formation: in essence, this increase derives from angular localization as the electrons leave spherical atomic-like orbitals to go into molecular bonding orbitals.

The dissociation transformation revealed by these studies, in which molecular hydrogen transforms not directly into a fully ionized plasma but first into a partially ionized atomic fluid, is also relevant to modeling the interior of giant planets such as Jupiter or Saturn. However, the situation there is more complex because hydrogen and helium coexist as main components. It is estimated that these two elements are at pressures of up to 4500 GPa and temperatures up to 24 000 K in Jupiter, and about 1000 GPa and 10 000 K in Saturn. The planet is thought to consist of three main layers: an outer layer of molecular hydrogen and atomic helium, a middle layer of metalized atomic hydrogen and helium, and a rocky core compressed to as high as 10^4 GPa at the center of Jupiter. In Saturn, an internal energy source has been proposed, from demixing and gravitational separation of the hydrogen–helium fluid at pressures below 1000 GPa.

Covalently Bonded Elements and Compounds

Molten and Amorphous Semiconductors

Semiconducting group-IV elements and polar III–V compounds crystallize in tetrahedrally coordinated open structures and melt with an increase in density into liquids, showing metallic conductivity. The first-neighbor coordination number rises on melting from four to about seven, though it remains well below that of common liquid metals which is about twelve. The elements Si and Ge can be quenched from the melt into a network-like amorphous state, in which an essentially fourfold coordination of first neighbors is restored and the semiconducting properties are regained. Melting of these solids is thus effected by a partial release of valence electrons from interatomic bonds into conducting states, and the chemical bonds are reconstructed in crystallization and in the transition to the amorphous state.

Figure 3 compares the structure factors of liquid and amorphous Ge as measured in diffraction experiments and also shows their correspondence with the location of the crystalline Bragg reflections. The shoulder to the right of the main peak in the structure factor of the liquid corresponds to the main peak of the amorphous structure factor and to the (2 2 0) and (3 1 1) Bragg spots. The interpretation of these structural features is that a remnant of the interatomic bonds is present in the melt, with some residue of the

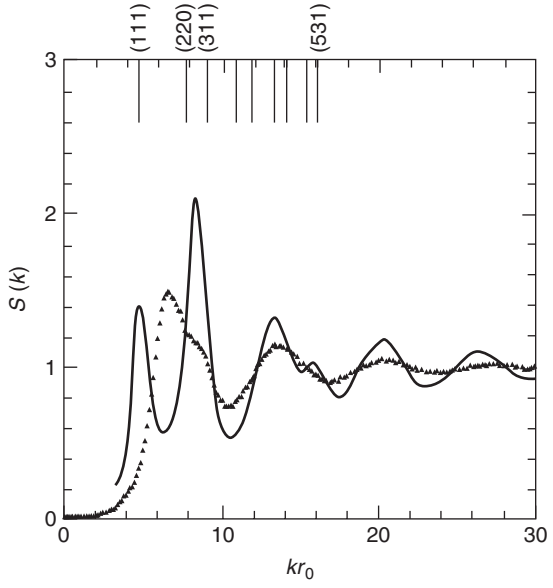


Figure 3 Structure factor of liquid Ge at 1000°C (triangles) and of amorphous Ge at room temperature (solid line). The vertical bars at the top give the location of the Bragg diffraction spots of crystalline Ge. The wave number k is scaled by the appropriate value of the first-neighbor distance r_o.

distribution of valence electrons still residing among the ionic cores, so that the transition to the crystal or to the amorphous state can be viewed as associated with "freezing of the bonds."

It is also seen in **Figure 3** that the diffraction pattern of the amorphous material shows a secondary peak to the left of the main peak. This is the "first sharp diffraction peak" (FSDP), which is a very general feature of the glassy state of matter. In Ge, its position corresponds to that of the (1 1 1) Bragg reflection from the crystal, while there is no corresponding feature in the diffraction pattern of the liquid. The FSDP comes from the reconstruction of the covalent bonds and reflects the connectivity that exists between the local tetrahedral units. The reconstruction would be frustrated if the system did not open up by lowering its density on freezing.

The presence of an FSDP is a signal of intermediate range order (IRO) in glasses. In strong glasses and liquids of the IV–VI compounds, IRO originates from the connectivity of chemically unsaturated tetrahedral units: these may join together either by pure corner sharing as in SiO_2, or by mixed corner and edge sharing as in $GeSe_2$, or by pure edge sharing as in $SiSe_2$. These connectivities give rise to networks having dimensionality $d = 3$, 2, and 1, respectively. The FSDP also behaves anomalously when compared with other diffraction peaks from disordered materials: its intensity increases with an increase in temperature and a decrease in pressure, since in both cases the available volume increases and the frustration of the network is reduced.

Crystalline, Liquid, and Amorphous Silicon Studied by the Car–Parrinello Method

The coexistence in Si and Ge of crystalline and amorphous covalent structures with a liquid metallic structure cannot receive a unified description by the standard methods of molecular dynamics (MD), which studies condensed matter systems by means of empirically fixed pair-potential models for the interatomic forces. On the other hand, density-functional theory provides a basic approach to the determination of electronic structure, but its direct use in the study of disordered systems is precluded, as it would require a full evaluation of electronic states for each of a very large number of configurations of the ionic system. The Car–Parrinello method successfully combines these two techniques in a way that allows MD calculations during which one follows the simultaneous evolution of the electronic and ionic structures and determines the interatomic forces from a parameter-free quantum mechanical approach.

In brief, the Car–Parrinello method treats variationally the total energy of a system such as liquid or amorphous Si as a function of both the electronic orbitals and the ionic positions. It uses a novel and efficient procedure of simulated annealing to carry out the minimization of the energy functional under evolving external constraints such as the volume and the strain of the sample. This is achieved by constructing a Lagrangian which determines the dynamics of orbitals, atomic positions, and constraints so as to guide the evolution of the simulation sample into a real physical system whose representative point in configuration space lies on the correct Born–Oppenheimer surface.

The three condensed states of Si indeed provided the first testing ground for the new MD approach. A number of static and dynamic properties of crystalline, amorphous, and liquid Si were calculated and successfully compared with experimental data whenever possible. The calculations yield contours of the valence charge density in correspondence to the atomic positions and allow one to visually follow its evolution in time during the atomic motions. In particular, in the case of liquid Si the persistence of some covalent chemical bonds was demonstrated from the analysis of the electronic charge density.

Solutions of Metals in Molten Salts

The liquid structure of an ionic compound such as CsI is characterized by strong short-range chemical order: each Cs^+ ion is surrounded by about five I^- ions and vice versa. A true solution is formed on a microscopic scale by adding Cs metal to molten CsI, and a continuous increase in electrical conductivity σ from the ionic to a metallic regime is observed as the concentration c of dissolved metal is increased at high-enough temperature. **Figure 4** reports the phase diagram of the CsI system together with the data showing the dependence of σ on composition. It is seen from the figure that addition of I to CsI is also possible: various intermediate compounds are formed and the system ultimately evolves into a fluid of I_2 molecules.

In the salt-rich region the added metal atoms are ionized and their valence electrons can be accommodated inside the structure of the host ionic melt. In analogy with the formation of F centers by localization of added electrons into negative-ion vacancies in the CsI crystal, optical absorption and magnetic resonance experiments show that at low c a solvated-electron state is formed in the liquid, with a residence time $\sim 10^{-12}$ s. The optical spectrum evolves from aggregation processes which start at very low metal content (< 0.01 M) and progressively reduce the

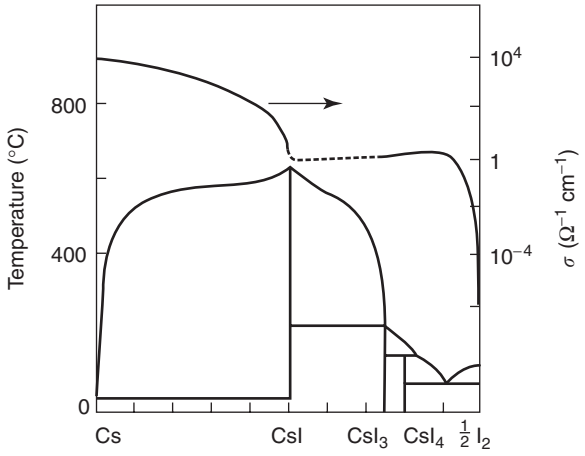


Figure 4 Phase diagram of the cesium–iodine system. The upper curve, with the scale on the right-hand side of the figure, shows the electrical conductivity σ.

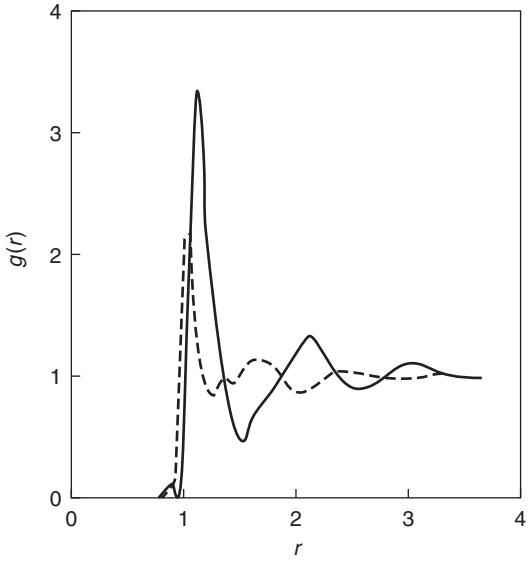


Figure 5 Pair distribution function $g(r)$ for oxygen atoms in water (dashed line) and for liquid argon (solid line). Both liquids are near freezing and the distance r is scaled by the van der Waals diameter.

fraction of electrons that are localized in F-center-like states. Electron spin resonance experiments show that spin-paired species are being formed. Finally, the transition to a metallic state is most strikingly signaled by a dielectric anomaly: according to a criterion proposed by Herzfeld in 1927, the low-frequency dielectric constant should indeed diverge at the transition as the reflectivity acquires metallic character.

Different scenarios are met on adding polyvalent metal to polyvalent metal halides. Subhalide compounds can be formed either from cationic species of intermediate valence (such as HgCl in Hg–Hg_2Cl melts) or from polyatomic cationic species (such as Bi clusters in Bi–BiI_3 at the concentration corresponding to the BiI compound). In all cases, one observes thermally activated electron transport at low c and a transition to a metallic regime shifted to the high-c region of the phase diagram by the stability of the intermediate subhalide compounds.

Hydrogen Bonding in Water and Ice

The specific nonspherical shape of the H_2O molecule and the interactions that it gives rise to have a number of important consequences. In the molecule the oxygen atom binds two hydrogen atoms by electron pairing, thereby creating a permanent electric dipole moment, and arranges its further four valence electrons in two "lone-pair" bonds. The four bonds point toward the corners of an almost perfect tetrahedron and each lone pair can interact with an electron-deficient hydrogen atom belonging to a neighboring molecule. By virtue of this hydrogen bond, each H_2O molecule in water or ice is tetrahedrally coordinated by four H_2O molecules. **Figure 5** contrasts the pair distribution function $g(r)$ of water with that of liquid argon from X-ray diffraction experiments, showing how different structural correlations build up in space for essentially tetrahedral molecules and for spherical atoms. The average number of first-neighbor oxygens is estimated from these data to be about 4.4 in water and about 10 in liquid argon. Water can be brought into a supercooled liquid state and also into a supercritical regime which is of great importance in extraction and reaction process technology.

As water crystallizes into ice at standard pressure, the local tetrahedral arrangement of the oxygens is frozen into a periodic lattice made of layers of rippled hexagons. This microscopic arrangement of the H_2O molecules in the crystalline phase is beautifully revealed by the sixfold symmetry of snowflakes. The structure of this form of ice is very open and this has two main consequences: melting is accompanied by some structural collapse leading to an increase in density, and the crystalline structure is very sensitive to pressure (some ten crystalline polymorphs of ice are known, all of them showing fourfold coordination of the oxygens).

Liquid and solid H_2O have been widely studied by diffraction techniques. X-rays are almost blind to hydrogen atoms, so that the water structure shown in **Figure 5** is in fact that taken by pairs of oxygens. Special techniques in neutron diffraction have allowed a detailed mapping of the distribution of the two atomic species in pure water and also in water around foreign cations and anions.

See also: Conductivity, Electrical; Conductivity, Thermal; Density-Functional Theory; Disordered Solids and Glasses, Electronic Structure of; Electron Gas (Theory); Glasses; Liquid Crystals; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Molecular Dynamics Calculations; Phases and Phase Equilibrium.

PACS: 71.22.+i; 71.30.+h

Further Reading

Ashcroft NW (2000) The hydrogen liquids. *Journal of Physics: Condensed Matter* 12: A129–A137.

Ball P (1999) *A Biography of Water*. London: Wiedenfeld and Nicolson.

Car R and Parrinello M (1985) Unified approach for molecular dynamics and density-functional theory. *Physics Review Letters* 55: 2471–2474.

Enderby JE and Neilson GW (1981) The structure of electrolyte solutions. *Reports on Progress in Physics* 44: 593–653.

March NH (1990) *Liquid Metals – Concepts and Theory*. Cambridge: Cambridge University Press.

March NH and Tosi MP (1995) Collective effects in condensed conducting phases including low-dimensional systems. *Advances in Physics* 44: 299–386.

March NH and Tosi MP (2002) *Introduction to Liquid State Physics*. Singapore: World Scientific.

Ross M (1985) Matter under extreme conditions of temperature and pressure. *Reports on Progress in Physics* 48: 1–169.

Rovere M and Tosi MP (1986) Structure and dynamics of molten salts. *Reports on Progress in Physics* 49: 1001–1081.

Tabor D (1993) *Gases, Liquids and Solids*. Cambridge: Cambridge University Press.

Winter R, Pilgrim WC, and Hensel F (1994) The metal-insulator transition and the static and dynamic structure factors of expanded liquid alkali metals. *Journal of Physics: Condensed Matter* 6: A245–A248.

Nomenclature

c_s	speed of sound
e	electron charge
$g(r)$	pair distribution function
$\hbar$	Planck constant h-crossed (h divided by 2π)
k_B	Boltzmann constant
k_F	Fermi wave number
L	Lorenz number
m	electron mass
M	ionic mass
n	electron density
N	ionic density
q	wave number
r	distance
T	temperature
v_F	Fermi velocity
Z	ionic valence
ε_F	Fermi energy
$\varepsilon(q)$	static dielectric function
η	shear viscosity
κ	thermal conductivity
λ_{TF}	Thomas–Fermi screening length
ρ_e	electrical resistivity
σ	electrical conductivity
$\phi_{12}(q)$	effective ion–ion potential
$\phi_{ie}(q)$	effective electron–ion potential
Ω_p	ionic plasma frequency

Liquids, Theory of: Fermi Liquids

J Spalek, Jagiellonian University, Kraków, Poland

Introduction: Quantum Liquids

Basic Characteristics

By Fermi liquid (FL), one means a quantum many-particle system of interacting fermions (i.e., particles with half-integer spin, such as electron, neutron, proton, ^{3}He atom, and also quarks), which is related to the (almost) ideal gas of fermions when the same system becomes sufficiently diluted. The FL has the same type of overall phase diagram as a classical liquid except for two specific features clearly observable for the canonical FL – condensed phase of ^{3}He atoms. First, due to the quantum nature of the particles involved, the zero-point motion prevents the solidification at temperature $T = 0$ and low pressure. Second, the FL transformations can take place to the superconducting or/and magnetic states. This feature is not touched upon here. In **Figure 1**, a schematic phase diagram for the condensed ^{3}He is provided.

First, some useful formulas for the ideal gas are provided. Namely, as the fermions obey the Pauli exclusion principle, they fill at $T = 0$ the single momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$ states up to the highest occupied level – the Fermi energy ε_F. In the case of an ideal gas with spin $s = 1/2$, the value is

$$\varepsilon_F = \frac{\hbar^2 k_F^2}{2m}, \quad \text{with } k_F = \left(3\pi^2 \frac{N}{V}\right)^{1/3} \qquad [1]$$

where $\hbar k_F$ is the Fermi momentum and N is the number of particles in volume V. For gaseous ^{3}He at pressure $p = 0$, one has the velocity of particles at the

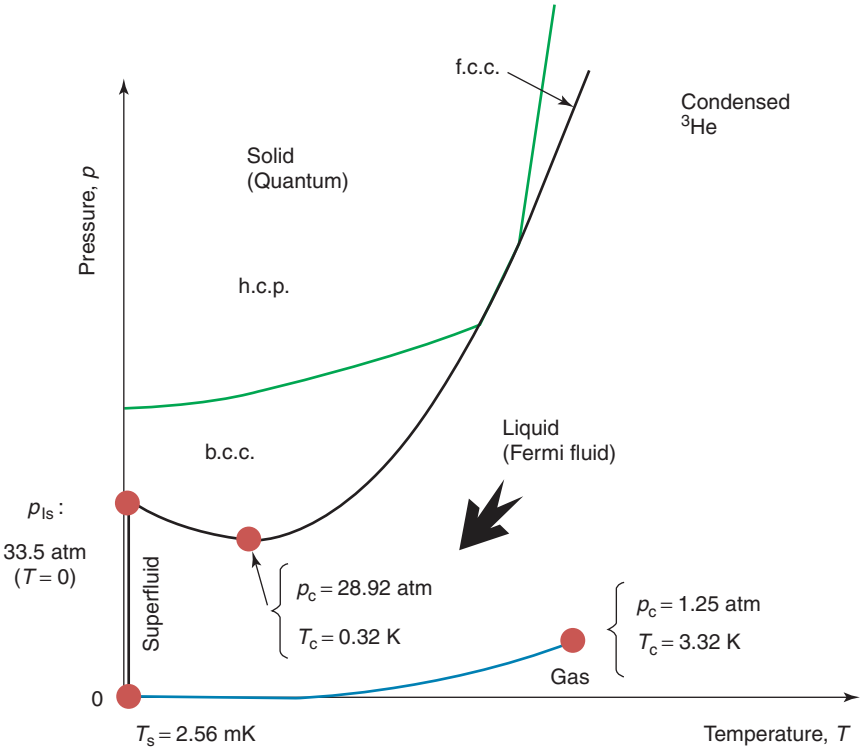


Figure 1 Schematic phase diagram of condensed ^{3}He including the critical points (solid curve).

Fermi energy $v_F = \hbar k_F/m \approx 10^4$ cm s^{-1} and the Fermi temperature $T_F = \varepsilon_F/k_B \approx 6.2$ K. The additional characteristic is the density of bare states at the Fermi level, $N_0(\varepsilon_F)$. When counted per atom per spin, it has the value

$$N_0(\varepsilon_F) = \frac{1}{8\pi^2}\frac{V}{N}\frac{2m}{\hbar^2}\left(3\pi^2\frac{N}{V}\right)^{1/3} \qquad [2]$$

In the case of electrons in a solid, one has: $\varepsilon_F \sim$ 1–10 eV, $k_F \sim (1\text{–}2)$ Å^{-1}, and $N_0(\varepsilon_F) \sim 10^{22}$ states eV^{-1}. This is the reason why the particle–energy distribution is regarded as quasicontinuous. However, the topology of the Fermi surface for electrons in metals is not spherical and the particle momentum is conserved with accuracy of the order of the inverse-lattice vector $\boldsymbol{G}$. On the contrary, the liquid ^{3}He can be regarded as a model Fermi liquid on both counts, since it is a truly translationally invariant FL system.

Basic Definitions

The Fermi liquid theory bears its name from the phenomenological approach proposed by L D Landau which aims to describe the quantum properties of liquid ^{3}He in the normal (nonsuperfluid) state at low temperatures: $k_B T \ll \varepsilon_F$. The spin $s = 1/2$ is that of ^{3}He nucleus, since the $1s^2$ filled electronic shell is spinless. The electronic liquid composed of valence electrons in metals such as Li, Na, or Cs represents a charged anisotropic FL state coupled to periodically arranged (much heavier) ions via the electron–phonon coupling. Separate classes of the FL represent the neutron stars (their external layers) and the quark–gluon plasma. The FL state is a viable concept for three-dimensional systems, whereas one-dimensional systems (fermionic chains, ladders, and nanotubes) are described as Tomonaga–Luttinger liquids (TLL). The exact nature of the two-dimensional electronic liquids (e.g., the high-temperature superconductors in the normal state or the quantum Hall liquids) is not yet clear. Also, the three-dimensional systems close to the magnetic quantum critical point represent non-Landau (non-Fermi) liquids, NFL. Therefore, by normal FL (or Landau FL), one understands that it is a three-dimensional quantum liquid composed of interacting fermions with delocalized states and without phase transition in this state.

A separate class is formed by the FLs close to the localization; these are called the almost localized Fermi liquids (ALFLs) (in the case of ^{3}He, this transition corresponds to the solidification, also included in **Figure 1**). The schematic division of the liquids of fermions, together with the corresponding examples, are listed in **Figure 2**.

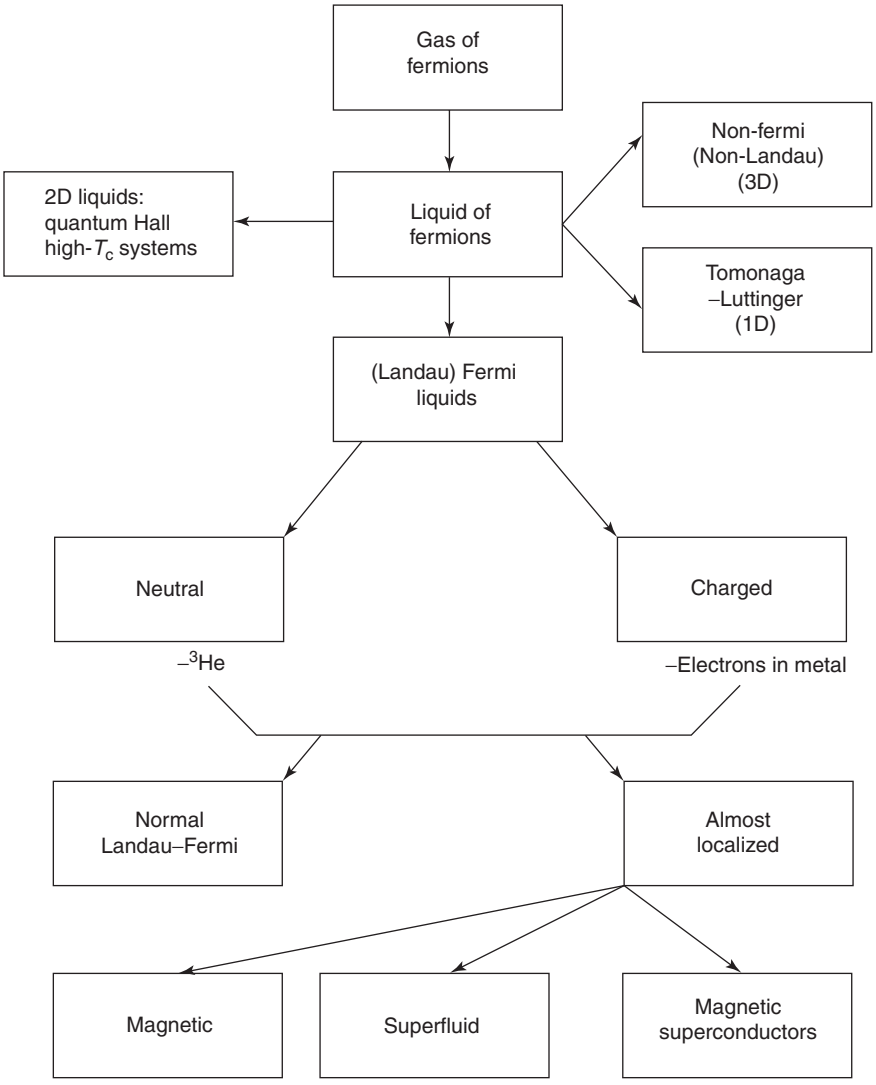


Figure 2 Division of various liquids of fermions and the subdivision of the FLs into various classes.

Signatures of the Normal FL State

The fundamental property of the FL state is the existence of a well-defined reference (gas) of single-particle states $|\boldsymbol{k}\sigma\rangle$ in the ground and first excited configurations with the energy $\varepsilon_{\boldsymbol{k}\sigma}$, where their (quasi) momentum is $\boldsymbol{p} = \hbar\boldsymbol{k}$ and the spin quantum number is $\sigma = \pm 1$ (note that $S = \sigma/2$). These reference states occupy their energy levels $\varepsilon_{\boldsymbol{k}\sigma}$ according to the Fermi distribution

$$n^0_{\boldsymbol{k}\sigma} = \frac{1}{\exp\left[\beta(\varepsilon_{\boldsymbol{k}\sigma} - \mu)\right] + 1} \qquad [3]$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature in energy units and μ is the chemical potential determined from the fact that the total number of particles is conserved, that is, $N = \sum_{\boldsymbol{k}\sigma} n^0_{\boldsymbol{k}\sigma}$. In the ground state, this distribution reduces to the step function $n_{\boldsymbol{k}\sigma} = \Theta(\mu - \varepsilon_{\boldsymbol{k}\sigma})$ expressing the Pauli exclusion principle, and μ expresses the Fermi energy ε_F. The equation $\varepsilon_{\boldsymbol{k}\sigma} = \varepsilon_F$ determines the shape of two (spin-dependent Fermi) surfaces in the $\boldsymbol{k}$ space. In the normal state and in the absence of an applied magnetic field, $\varepsilon_{\boldsymbol{k}\sigma} = \varepsilon_{\boldsymbol{k}}$, and in this circumstance, a single Fermi surface and k_F represent (angle-dependent in the general case) the Fermi wave vector.

Interaction among Particles

The FL concept is particularly useful in the low-temperature regime $T \ll T_F = \varepsilon_F/k_B$, where the interest lies only in the thermal (and dynamic) properties of weakly excited states. In this limit, one is interested in determining the system energy change δE induced by the small change $\delta n_{\boldsymbol{k}\sigma} \equiv n_{\boldsymbol{k}\sigma} - n^0_{\boldsymbol{k}\sigma} \ll 1$

(on average) and due to interaction among them, and/or by the thermal excitations (the exact knowledge of the total state energy, is difficult to calculate and even not necessary). In this situation, in the energy δE the interaction in the diluted gas of excited states can be included in the following manner:

$$\delta E = \sum_{k\sigma} (\varepsilon_k - \mu)\delta n_{k\sigma} + \frac{1}{2}\sum_{kk'\sigma\sigma'} f_{kk'}^{\sigma\sigma'} \delta n_{k\sigma}\delta n_{k'\sigma} + \cdots \quad [4]$$

where $f_{kk'}^{\sigma\sigma'}$ expresses the (spin-dependent) density–density interaction between the particles, that leads foremostly to the elastic scattering processes at the Fermi surface, since inside it the scattering processes are practically blocked out by the Pauli principle. It is also to be noted that the bare (without interaction included) energy ε_k is counted from ε_F. In this formulation, the ground state is regarded as a vacuum state.

In such an approach, there are two subtleties apart from the unknown structure of the scattering function $f_{kk'}^{\sigma\sigma'}$. The first is connected with the conservation of particles, which now takes the form $\sum_{k\sigma} \delta n_{k\sigma} = \delta N = 0$. In other words, if the single-particle excitations with $\varepsilon_k > \mu$ are regarded as particles ($\delta n_k = +1$), then those with $\varepsilon_k < \mu$ should be regarded as having $\delta n_k = -1$. In other words, the excitation from the vacuum state requires that if the charge e is associated with the state $\varepsilon_k > \mu$ and $\delta n_{k\sigma} = +1$, then the charge $-e$ has to be associated with the excitation $\varepsilon_k < \mu$ and $\delta n_{k\sigma} = -1$, so that the total charge of the system is zero. This concept bears its origin from the Dirac concept of electrons and holes, which in the present (nonrelativistic) situation does not lead to any peculiarities.

The Entropy of Interacting Particles ($T \ll T_F$)

The second feature is associated with the inclusion of the determination of δn_k at $T > 0$ or equivalently, with the determination of the system entropy. Here, the phenomenology appears in its earnest namely, one assumes that it is taken in the usual form for non-interacting fermions for the number of particles $\{\delta n_{k\sigma}\}$, that is,

$$\delta S = -k_B \sum_{k\sigma} \{\delta n_{k\sigma} \ln \delta n_{k\sigma} + (1 - \delta n_{k\sigma}) \ln(1 - \delta n_{k\sigma})\} \quad [5]$$

This assumption is justified by adopting the adiabatic principle: the single-particle (quasiparticle) states in the presence of the interparticle interaction are in a one-to-one correspondence with those for the ideal gas. In other words, the number of microconfigurations in the interacting system does not change, nor does the entropy. This theorem finds its microscopic justification through the Luttinger theorem: the volume (i.e., the number of states encompassed by k_F at $T = 0$) does not depend on the interaction magnitude as long as the system does not undergo either magnetic or superconducting or delocalization–localization transition. This theorem implies that the occupation number ($\delta n_{k\sigma}$) can change, but the Fermi ridge (discontinuity) position in $\boldsymbol{k}$ space does not.

Quasiparticles

By writing down the total energy in the form

$$\delta E = \sum_{k\sigma} \left\{ \varepsilon_{k\sigma} - \mu + \frac{1}{2}\sum_{k'\sigma'} f_{kk'}^{\sigma\sigma'} \delta n_{k'\sigma} \right\} \delta n_{k\sigma} \equiv \sum_{k\sigma} E_{k\sigma} \delta n_{k\sigma} \quad [6]$$

the energy $E_{k\sigma}$ of the quasiparticle can be defined. In effect, the distribution [3] for quasiparticles can be defined (by minimizing the free energy $\delta F = \delta E - T\delta S$), that is, with respect to $\{\delta n_{k\sigma}\}$, with the replacement $\varepsilon_{k\sigma} - \mu \to E_{k\sigma}$. **Figure 3** provides the statistical distribution function n_p for liquid ^{3}He obtained from the neutron scattering experiment of Mook. It can be seen that the Fermi distribution function describes the quasiparticle states well, but with the Fermi temperature $T_F = 1.8\,\mathrm{K}$ and the energy given by $E_k = p^2/2m^\star$, where $m^\star \approx 3m_0$ is the effective mass of ^{3}He atoms in the FL state. The Fermi temperature is also renormalized by the factor $T_F = (m_0/m^\star)T_{F0}$, where T_{F0} is the corresponding quantity for ^{3}He gas of the same density. So, the interaction changes both the mass and the Fermi energy to the same extent for all particles (i.e., not only those close to k_F).

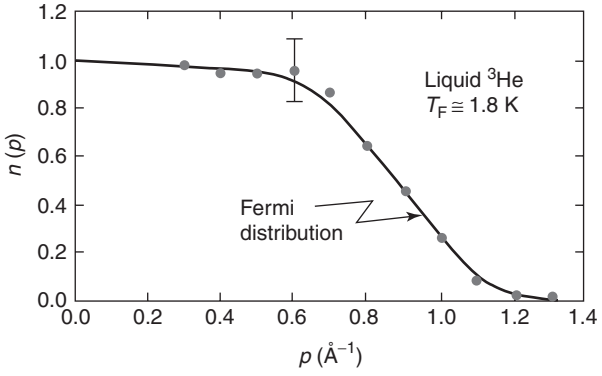


Figure 3 The Fermi distribution for liquid ^{3}He at $T = 0.37\,\mathrm{K}$ and ambient pressure. (After Mook HA (1985) Momentum distribution of ^{3}He. *Physical Review Letters* 55: 2452.)

Electrons in a Metal as Quasiparticles

The states of electrons in a metal are probed by exciting them with photons in the photoelectron emission from the solid. To describe such a process, one has to know the spectral density function $A_{k\sigma}(\omega)$ describing microscopically the form of the system energy spectrum when extracting (or adding) one electron with momentum $\hbar\boldsymbol{k}$ and the energy ω. $A_{k\sigma}(\omega)$ is, in turn, defined as

$$A_{k\sigma}(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\; G_{k\sigma}(\omega) \equiv -\frac{1}{\pi}\,\mathrm{Im}\,\frac{1}{\omega - E_{k\sigma} + \mathrm{i}\Sigma''_{\sigma}(\omega)} \qquad [7]$$

where $G_{k\sigma}(\omega)$ is the (retarded) single-particle Green function for the electron. The quantity $\Sigma''_{\sigma}(\omega)$ represents the imaginary part of the self-energy and for a three-dimensional FL is given by

$$\Sigma''_{\sigma}(\omega) = \beta\omega^2, \quad \beta = \text{const.} \qquad [8]$$

The real part of the self-energy is included in $E_{\boldsymbol{k}}$ (replacing in microscopic theory the phenomenological part $\sim f_{\boldsymbol{k}\boldsymbol{k}'}$).

The function $A_{k\sigma}(\omega)$ is of fundamental importance for the many-particle system. This is because the density of states $N(\omega)$ and the statistical distribution function $n_{k\sigma}$ are derived from it, namely,

$$N_{\sigma}(\omega) = \sum_{k} A_{k\sigma}(\omega); \quad n_{k\sigma} = \int \mathrm{d}\omega\; A_{k\sigma}(\omega) \qquad [9]$$

where the integration is over allowed values of energy of the system (e.g., the band width or the cutoff energy). In the paramagnetic systems, the spin subscript is dropped; sometimes $A_{\boldsymbol{k}}(\omega)$ is defined by summing up over $\sigma = \pm 1$. The photoemission intensity, that is, the number of electrons emitted in the direction $\boldsymbol{k}$ with the energy ω is given by

$$I_{\boldsymbol{k}}(\omega) = I_0 \int \mathrm{d}\omega' A_{\boldsymbol{k}}(\omega') n(\omega') R(\omega - \omega') \qquad [10]$$

where $R(\omega - \omega')$ is the so-called resolution (instrument) function (usually taken in the Gaussian form), $n(\omega)$ is the Fermi factor: $n(\omega) = [\exp(\beta\omega) + 1]^{-1}$, with $\omega = 0$ representing the Fermi energy (I_0 is the normalization constant). **Figure 4** provides an exemplary angle-resolved photoemission spectrum (ARPES) with the quasiparticle peak dispersing as the regime of energies is probed with $E_{\boldsymbol{k}} < \mu(= 0)$; the angle θ describes the angle at which the photoelectron is emitted, and determines the wave vector $\boldsymbol{k}_{\parallel}$ in this quasi-two-dimensional metal, the Fermi energy.

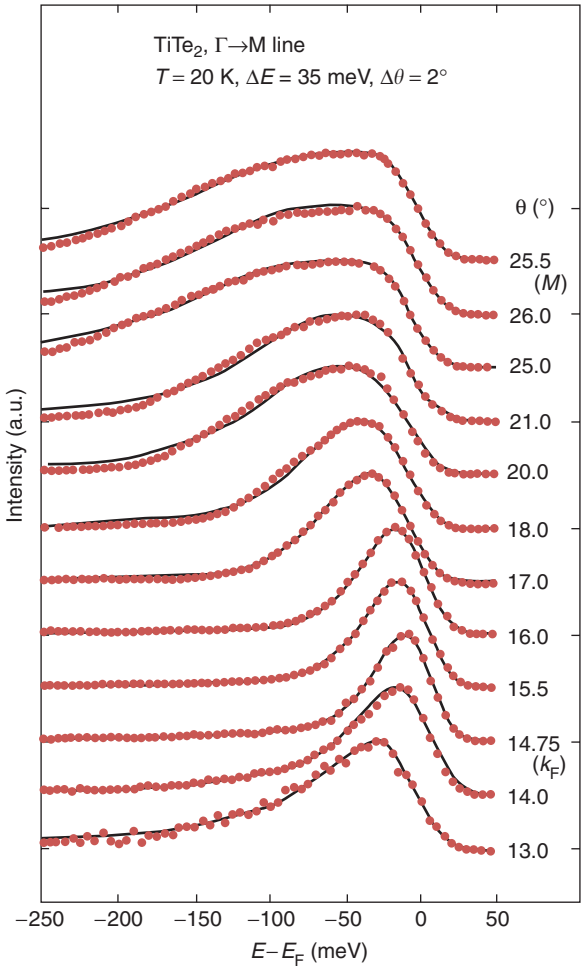


Figure 4 Photoemission spectra $TiTe_2$ along the $\Gamma \rightarrow M$ line (from bottom to top). The solid line represents the Fermi-liquid fit [9]. The spectra are normalized to the same maximum height. (After Allen JW, Gweon G-H, Claessen R, and Matho K (1995) Fermi liquids and non-Fermi liquids – the view from photoemission. *Journal of the Physics and Chemistry of Solids* 56: 1849.)

The dispersion relation for quasiparticles is linear when $E_{\boldsymbol{k}} \rightarrow \varepsilon_{\mathrm{F}}$.

Properties at Low Temperatures

A brief description of the properties at low T follows. However, for that, first the function $f^{\sigma\sigma'}_{\boldsymbol{k}\boldsymbol{k}'}$ has to be specified. Namely, if processes at the Fermi surface only are included, then one can write that $f^{\sigma\sigma'}_{\boldsymbol{k}\boldsymbol{k}'} = f^{\sigma\sigma'}_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{k}\cdot\boldsymbol{k}'/k_{\mathrm{F}}^2)$, that is, the elastic scattering processes depend only on the angle between the particles. Additionally, one can write

$$f^{\sigma\sigma'}_{\boldsymbol{k}\boldsymbol{k}'} = f^{s}(\boldsymbol{k}\cdot\boldsymbol{k}'/k_{\mathrm{F}}^2) + \sigma\cdot\sigma' f^{a}(\boldsymbol{k}\cdot\boldsymbol{k}'/k_{\mathrm{F}}^2) \qquad [11]$$

where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices. The form [11] defines the spin-symmetric

($f^{\mathrm{s}} = \frac{1}{2}(f^{\uparrow\uparrow} + f^{\uparrow\downarrow})$) and the antisymmetric ($f^{\mathrm{a}} = \frac{1}{2}(f^{\uparrow\uparrow} - f^{\uparrow\downarrow})$) functions in the situation when the spin–spin interactions are rotationally invariant (e.g., neglect the spin–orbit interaction). Additionally, introducing $\cos\theta \equiv \boldsymbol{k}\cdot\boldsymbol{k}'/k_{\mathrm{F}}^2$, the functions can be expanded in terms of the Legendre polynomials $P_l(\cos\theta)$, that is,

$$f_{\boldsymbol{k}\boldsymbol{k}'}^{\mathrm{s,a}} = \sum_{l=0}^{\infty} f_l^{\mathrm{s,a}} P_l(\cos\theta) \qquad [12]$$

where

$$f_l^{\mathrm{s,a}} = (2l+1)\int_{-1}^{+1} \mathrm{d}x P_l(x) f^{\mathrm{s,a}}(x) \qquad [13]$$

are the partial spherical-wave amplitudes. The ingenuity of this approach is rooted in the circumstance, realized only *a posteriori*, that all the principal properties are expressed in terms of a few parameters ($f_0^{\mathrm{s}}, f_1^{\mathrm{s}}$, and f_1^{a}). These properties can be expressed in terms of the dimensionless parameters $F_l^{\mathrm{s,a}} \equiv N(0)F_l^{\mathrm{s,a}}$, where $N(0)$ is the density of states at the Fermi level including both spin directions and is defined by

$$\begin{aligned} N(\mu \equiv 0) &= \sum_{k\sigma} \delta(E_{\boldsymbol{k}} - \mu) \\ &= -2(2\pi)^{-3}\int \mathrm{d}^3k(\partial/\partial E_{\boldsymbol{k}})\theta(\mu - E_{\boldsymbol{k}}) \\ &= Vp_{\mathrm{F}}(m^{\bigstar}/\hbar^3\pi^2) \end{aligned} \qquad [14]$$

Summary of Properties

1. Effective mass enhancement at $T = 0$,

$$m^{\bigstar} = m_0\left(1 + \tfrac{1}{3}F_1^{\mathrm{s}}\right) \qquad [15]$$

   where m_0 is the bare mass (band-theory mass for the case of electrons).
2. Bulk compressibility of the FL system at $T = 0$,

$$\kappa \equiv -\frac{1}{V}\frac{\partial V}{\partial p} = \frac{\rho^2}{V}\frac{\partial \mu}{\partial N} = \frac{\rho^2}{V}\frac{1+F_0^{\mathrm{s}}}{N(0)} \qquad [16]$$

   where $\rho \equiv \mathrm{d}N/\mathrm{d}V$ is the particle density at a given volume; associated with this quantity is the sound velocity V_{s}, which is $\boldsymbol{p}_{\mathrm{F}}^2/(3m_0m^{\bigstar})(1+F^{\mathrm{s}})$.
3. Paramagnetic susceptibility at $T = 0$ is enhanced in the following manner:

$$\chi(0) \equiv \left(\frac{\partial M}{\partial H_{\mathrm{a}}}\right)_0 = \frac{1}{4}(g\mu_{\mathrm{B}})^2 N(0)\frac{1}{1+F_0^{\mathrm{a}}} \qquad [17]$$

   where g is the Lande factor and $(\partial M/\partial H_{\mathrm{a}})_0$ expresses the ratio of magnetization to the magnitude of the (small) field creating it. At $T > 0$, the susceptibility takes the form $\chi(T) = \chi(0)(1 + aT^2)$, where a is a constant.
4. Electrical resistivity of the electron fluid can be represented by the formula $\rho = m^{\bigstar}/(n_{\mathrm{c}}e^2\tau)$, where n_{c} is the carrier concentration and τ is the inverse lifetime for scattering due to the particle interactions (Baber–Landau–Pomeranchuk)

$$1/\tau \sim (m^{\bigstar})T^2 \qquad [18]$$

   So the resistivity can be described by $\rho(T) = AT^2$, with $A \sim \gamma(0)^2$, where $\gamma(0)$ is the linear specific-heat coefficient $\gamma(0) = (\pi^2/3)k_{\mathrm{B}}^2 N(0)$. In the case of neutral FL, this relaxation time determines the viscosity.
5. Specific heat has both the contributions due to single-particle excitations, $\sim(\gamma(0)T)$ and that due to the collective spin fluctuations; the latter are reflected in the simplest case as a maximum in the dynamic susceptibility $\chi(\boldsymbol{q},\omega)$ at $\boldsymbol{q} = \omega = 0$. In effect,

$$C(T) = \gamma(0)T + \delta T^3 \ln(T/T_{\mathrm{sf}}) \qquad [19]$$

   where $\gamma(0) = m^{\bigstar}p_{\mathrm{F}}/(3\hbar^3)$, and δ and T_{sf} are characteristics of the spin fluctuation spectrum.
6. The Wilson ratio of the extrapolated values, $R_{\mathrm{w}} \equiv \chi(0)/\gamma(0)$, describes the relative strength of magnetic and charge excitations.

Collective Excitations

Apart from the paramagnetic spin fluctuations, there is a specific collective excitation of the FL – the zero sound. It appears for frequencies, for which $\omega \gg \tau^{-1}$, that is, in the collisionless regime for quasiparticles. The role of the restoring force during the propagation of such density fluctuations is provided by the averaged self-consistent field of the remaining particles in the system. In such a nomenclature, the hydrodynamic (collision-dominated) regime corresponds to the sound propagation for $\omega \ll \tau^{-1}$. The density oscillations in this regime correspond to first-sound propagation, which can be identified with damped oscillation, the ordinary sound. Apart from those excitations, there are electron–hole and plasmon excitations, for elaboration on which the reader is advised to consult more specialized texts (see "Further reading" section). The same remark applies to the discussion of other nonequilibrium and transport properties of Fermi liquids.

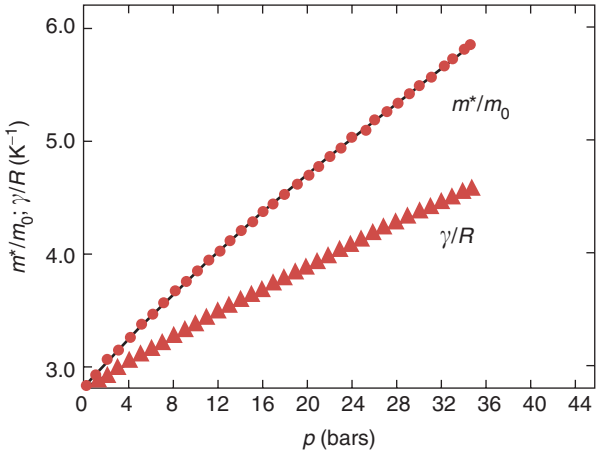


Figure 5 Pressure dependence of the linear specific-heat coefficient $\gamma(0)$ (in units of the gas constant R) and the effective mass enhancement $m^\star/m_0$ of the ^{3}He atoms in the liquid state (top), when approaching the solidification. (After Greywall DS (1986) ^{3}He specific heat and thermometry at milliKelvin temperatures. *Physical Review B* 33: 7520.)

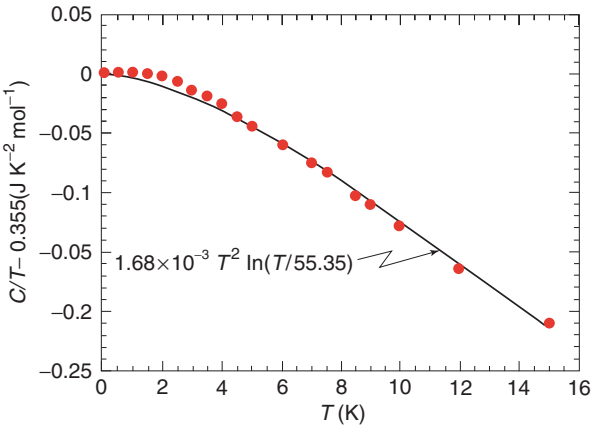


Figure 6 Temperature dependence of the (molar) specific heat of $CeCu_2Si_2$, with the linear part extracted. (Courtesy of J Flouquet group from Grenoble.)

Unconventional Fermi Liquids

A nonstandard behavior is to be expected when a normal FL undergoes a phase transition. As liquids which undergo the superconducting or superfluid transitions are discussed elsewhere in this encyclopedia, the focus here is on the solidification of ^{3}He and the Mott–Hubbard localization of an almost localized electron liquid. **Figure 5** provides the molar pressure dependence of $\gamma(0) = (\pi^2/3)k_B^2 N(0)$, as well as the mass enhancement $m^\star/m_0$ for ^{3}He calculated from the relation $\gamma(0) = \gamma_0(m^\star/m_0)$, where $\gamma_0 = (\pi^2/3)k_B^2 N_0(0)$, and $N_0(0)$ represents the corresponding density of states for an ideal ^{3}He gas (of the same density). As the liquid approaches the solidification threshold, that is, the localization of atoms into a lattice, the quasiparticle mass doubles. At the highest pressure, the system undergoes a discontinuous Mott–Hubbard transition.

Heavy-Electron Systems

A spectacular effective mass enhancement is observed in the heavy-electron systems, which compose intermetallic compounds involving itinerant $4f$ or $5f$ electrons, for example, $CeAl_3$, UPt_3, UBe_{13}, and many others. The effective masses encountered there are $\sim 10^2 m_0$. In **Figure 6**, the (molar) specific heat data for the moderately heavy-fermion compound $CeCu_2Si_2$ is presented. The extrapolated value of $\gamma(0) = 355\,\mathrm{mJ\,K^{-2}\,mol^{-1}}$ has been subtracted from the C/T data. The heavy mass of these quasielectrons is characterized by the effective Kondo (coherence) temperature $T_K = \pi^2 R/3\gamma$, which characterizes the bandwidth of these quasiparticle states. The solid curve represents the second term of the expression [19] divided by T. The heavy masses are associated, in this case, with the hybridization of $4f^1$ states of Ce^{3+} with the conduction states, and the intra-atomic Coulomb interaction for the $4f^2$ configuration is large, so the f-states can become itinerant only when the valency is slightly larger than Ce^{3+} ($4f^{1-\varepsilon}$ configuration). In other words, a small number of holes in the $4f^1$ configuration cause the almost-localized behavior when these electrons additionally hybridize with the extended valence states.

The Fermi liquid nature of the compounds for $T < T_K$ is evidenced further in the behavior of the resistivity $\rho(T) = \rho_0 + AT^2$, as well as by the Wilson ratio $R_w \sim 1$. The two observations are displayed in **Figures 7** and **8**, respectively. The straight line in **Figure 7** illustrates the relation $A \sim \gamma(0)^2$, whereas that in **Figure 8** corresponds to a Fermi gas. So, indeed, the concept of very heavy quasiparticles is applicable even in these extreme conditions.

Almost Localized Systems

The Mott–Hubbard localization takes place when the magnitude U of the short-range (intra-atomic) Coulomb repulsive interaction exceeds the average band energy per particle $(1/N)\sum_{k\sigma}\varepsilon_k \equiv \bar{\varepsilon}$. This is because the band energy represents the binding energy gained by forming an extended (Bloch) state of electron in solid, whereas the Coulomb repulsion competes with such a process as it tries to keep them as far apart as possible, that is, on their parent atoms.

To put such an argument on a quantitative basis, the renormalized band energy is represented as $q\bar{\varepsilon}$, where the band narrowing factor is $q \leqslant 1$ and vanishes at the localization threshold. The Coulomb

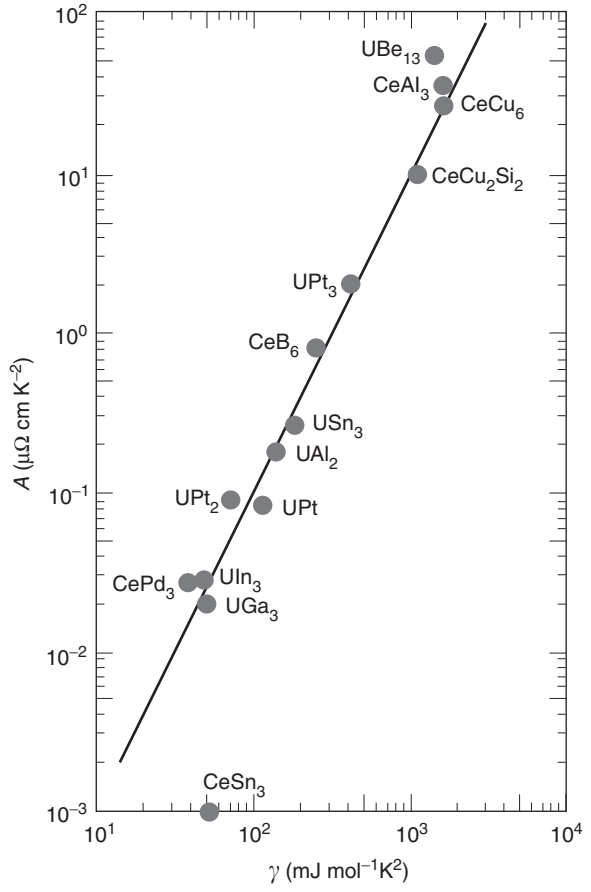


Figure 7 Universal scaling to the A coefficient in resistivity ($\rho(T) = \rho_0 + AT^2$) of various heavy fermion systems vs. the coefficient γ_0. The straight line represents the Kadowaki–Woods scaling law $A \sim \gamma(0)^2$.

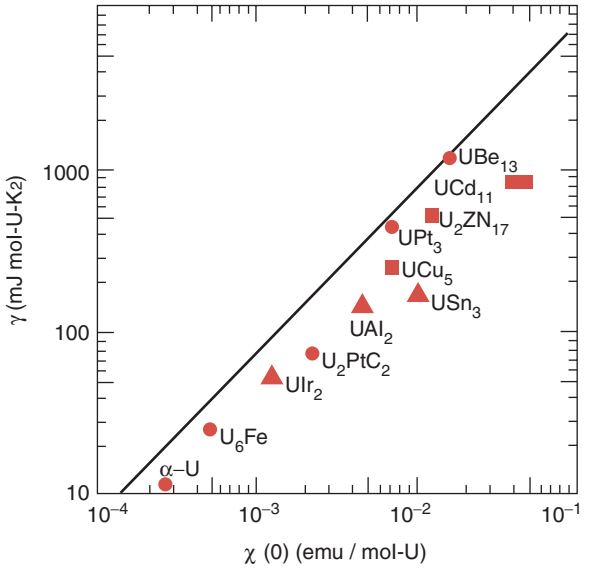


Figure 8 Universal scaling of the coefficient $\gamma(0)$ vs. the susceptibility $\chi(0)$ for various heavy fermion systems. The straight line is for the ideal electron gas.

energy (per atom) is written as $U\eta$, where $\eta = \langle n_{i\uparrow} n_{i\downarrow} \rangle$ represents the probability that the atom "i" is doubly occupied (a single state per atom is taken, so that the spins are opposite). When $\eta > 0$, the double occupancies are formed and the electrons (originally one per atom) can hop from one atom onto a neighboring one and thus conduct charge. To close the argument, one specifies $q \equiv q(\eta)$, which in the simplest formulation (the Gutzwiller ansatz) is of the form $q(\eta) = 8\eta(1 - 2\eta)$. Thus, the total energy is

$$E_G/N = q(\eta)\bar{\varepsilon} + U\eta \qquad [20]$$

In the equilibrium state, $\partial E_G/\partial\eta = 0$, $\partial^2 E_G/\partial\eta^2 > 0$. One then obtains

$$\eta \equiv \eta_0 = \tfrac{1}{4}(1 - U/U_c) \qquad [21]$$

$$q \equiv q_0 = 1 - (U/U_c)^2 \qquad [22]$$

$$E_G/N = \bar{\varepsilon}(1 - U/U_c)^2 \qquad [23]$$

where $U_c = 8|\bar{\varepsilon}|$ is the critical value of the interaction, at which $E_G = \eta_0 = 0$, that is, the energies totally compensate each other (the metallic state for $U > U_c$ is unstable). One can also calculate the magnitude of spin on the representative atom "i" and obtain

$$\langle S_i^2 \rangle = \frac{3}{4}(1 - 2\eta_0) = \frac{3}{8}\left(1 + \frac{U}{U_c}\right) \qquad [24]$$

At $U = U_c$, the spin of the particle reduces to the Pauli spin, with $\langle S_i^2 \rangle = \frac{1}{2}(\frac{1}{2} + 1)$.

Mott–Hubbard Transition as a Phase Transition

The factor $q(\eta_0)$ describes the mass enhancement also. Namely,

$$\frac{m^\star}{m_0} = \frac{\gamma(0)}{\gamma_0} = q^{-1} = [1 - (U/U_c)^2]^{-1} \qquad [25]$$

So, $m^\star$ diverges (and so does the linear specific heat) as $U \to U_c$. Additionally, one can also calculate the magnetic susceptibility $\chi(0)$, which reads

$$\chi(0) = \chi_0/q(\eta_0)\left\{1 - \frac{U}{U_c}\frac{1 + U/(2U_c)}{(1 + U/U_c)^2}\right\} \qquad [26]$$

where χ_0 is the Pauli susceptibility. Note that χ is divergent as $U \to U_c$, since $q \to 0$ in this case (the other divergence represents the renormalized Stoner criterion for the onset of ferromagnetism). Hence, the point $U \to U_c$ represents, at least within this simple approach, a quantum critical point (the Brinkman–Rice point). One can also define the parameters

F_1^s and F_0^a by comparing [25] and [26] with [15] and [17].

Localization of Electrons at Nonzero Temperature

One can generalize the meaning of the factor $q(\eta)$ by defining statistical quasiparticles with energies $E_k = q(\eta)\varepsilon_k$. Hence, the statistical distribution of quasiparticles is, as in the Landau–Fermi liquid, $f(E_k) = \{\exp[\beta(E_k - \mu)] + 1\}^{-1}$. The free energy of the metallic ALFL written for all electrons is then

$$\frac{F}{N} = \frac{1}{N}\sum_{k\sigma} E_k f(E_k) + U\eta + \frac{k_B T}{N}\sum_{k\sigma}\{f(E_k)\ln f(E_k) + [1 - f(E_k)]\ln[1 - f(E_k)]\} \quad [27]$$

In the low-T limit, the expansion to the first nontrivial order leads to

$$\frac{F}{N} = -q(\eta)\bar{\varepsilon} + U\eta - \frac{\gamma_0 T^2}{2q(\eta)} + o(T^4) \quad [28]$$

The next step is to introduce the discontinuous phase transition in the context of ALFL instability. The localized electrons (atomic $s = 1/2$ spins) are represented in the paramagnetic phase by their entropic part only, that is, their free energy is

$$\frac{F_l}{N} = -k_B T \ln 2 \quad [29]$$

where the exchange interactions between the electrons have been neglected as only paramagnetic states have been described here. Equating $F = F_l$, the coexistence of the phases is obtained.

The phase diagram for this system on the plane $T - U/U_c$ is shown in **Figure 9**, together with the "mean-field" critical points. The main feature is the reentrant (PM) metallic behavior in the high-temperature regime. Such a reentrant behavior is observed for the condensed ^{3}He (on the $p - T$ plane, cf. **Figure 1**). For the canonical Mott–Hubbard system, pure and doped V_2O_3, the upper line $T_+(U)$ represents a crossover behavior accounted for in a more refined approach.

Spin-Dependent Quasiparticle Masses

There are two specific properties of almost-localized electrons in the presence of a magnetic field. The first of them is the metamagnetism, that is, the transition in an applied magnetic field to a saturated ferromagnetic phase via a first-order transformation. The second is the appearance of a spectacular spin dependence of the quasiparticle masses, with $m^\star/m_0 \equiv m_\sigma/m_0 = 1/q_\sigma$, which appears only for a non-half-filled narrow-band situation. In the limit of strongly correlated electrons ($U \gg U_c$), the

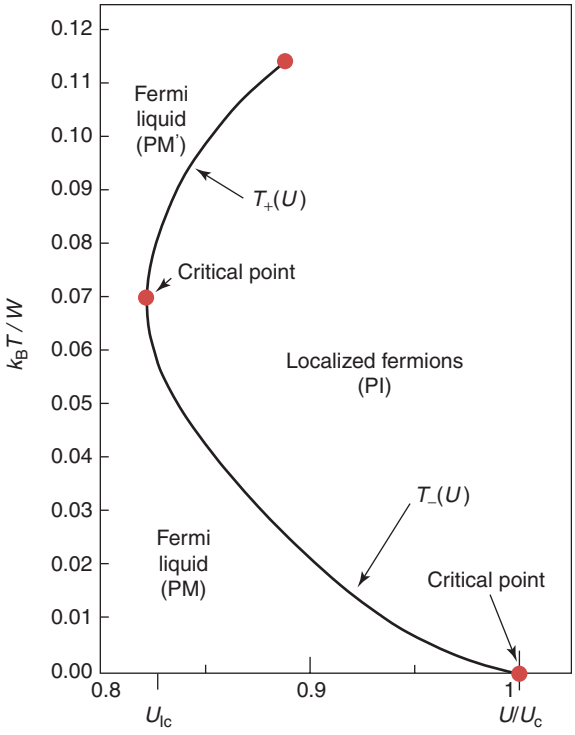


Figure 9 Schematic phase diagram for the almost localized electrons in a narrow band in the paramagnetic state.

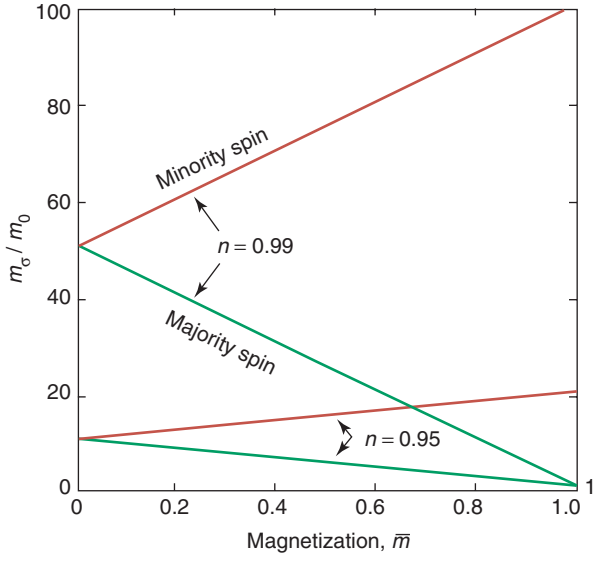


Figure 10 Spin-split effective masses for almost localized electrons as a function of magnetic polarization $\bar{m} = (n_\uparrow - n_\downarrow)/(n_\uparrow + n_\downarrow)$ for the number n of electrons per atom specified. (cf. Spalek J and Gopalan P (1990) Almost localized electrons in a magnetic field. *Physical Review Letters* 64: 2823.)

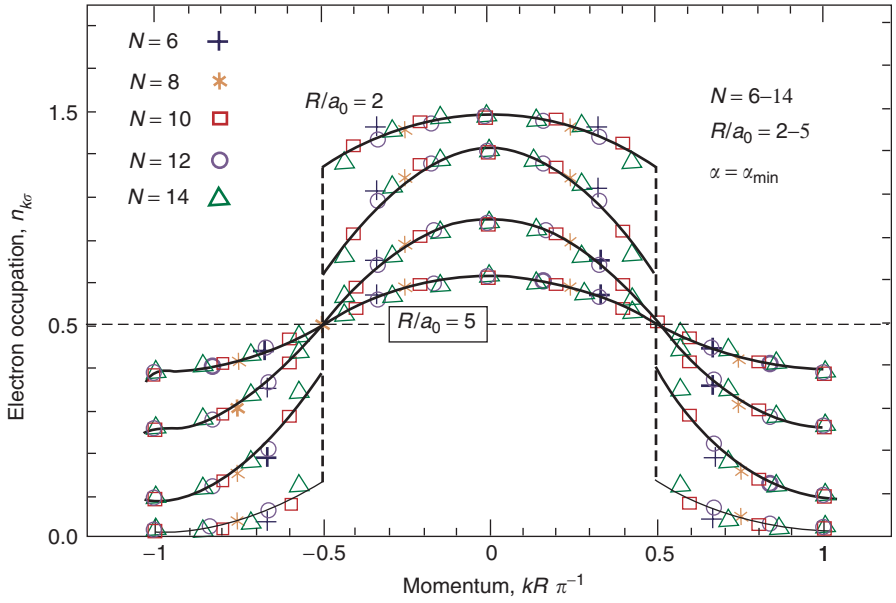


Figure 11 Momentum distribution for a nanochain of N atoms specified and for different interatomic distances R (in units of 1s Bohr orbit). (cf. Spalek J and Rycerz A (2001) Electron localization in a one-dimensional nanoscopic system: a combined exact diagonalization – an *ab initio* approach. *Physical Review B* 64: 161105.)

enhancement factor takes the form

$$1/q_\sigma = (1 - n_\sigma)/(1 - n) \qquad [30]$$

where n is the number of electrons per atom in the correlated band, with the energies of particles $E_k \equiv E_{k\sigma} = q_\sigma \varepsilon_k - \sigma \mu_B H_a$, and n_σ is the number of electrons (per atom) of spin σ (the quantity $\bar{m} \equiv (n_\uparrow - n_\downarrow)/(n_\uparrow + n_\downarrow)$ is the spin polarization). **Figure 10** presents the spin-split masses as a function of $\bar{m}$ for both the majority and the majority-spin subbands for two-band fillings close to the half-filling, where the Mott–Hubbard transition localized phase would be already stable. The spin-split masses were predicted theoretically but have not been detected as yet.

Marginal Fermi Liquid

In the case of a normal FL, one has $\mathrm{Re}\,\Sigma(\omega) \sim \omega$, whereas $\Sigma''(\omega) \sim \omega^2$, whereas in the phenomenological approach to high-temperature superconductors, Varma and co-workers assume that

$$\Sigma(\omega) \sim \lambda \left(\omega \ln \frac{\omega_c}{|\omega|} + i|\omega| \right)$$

and hence the imaginary part is comparable to the quasiparticle energies with energies $\omega \sim \varepsilon_k$ (λ is the effective coupling constant and ω_c is the so-called cutoff frequency, beyond which the marginal FL (MFL) concept does not hold). The real part of the self-energy is weakly singular and this, in conjunction with the fact that $\Sigma'' \sim |\omega|$, specifies the two-dimensional nature of those systems.

Nanosystems as Quantum Liquids

In the era of nanotechnology, a legitimate question is: how short can a quantum (monoatomic) wire be? The answer to this question can be provided by solving exactly the system of $N \sim 10$ atoms with one valence electron (e.g., a chain of N hydrogen atoms in the simplest situation). **Figure 11** shows the statistical distribution function $n_{k\sigma}$ drawn for different interatomic distances R in units of Bohr radius a_0. For $R = 2a_0$, the distribution is close to the Fermi distribution, with quasi-discontinuity (the vertical dashed line) near the Fermi point for the infinite system. The distribution is continuous $n_{k\sigma} \approx 1/2$ when the particles become localized on atoms.

Summarizing this article in one sentence, the concept of FLs introduced in 1956 is valid even today, often in a completely different context.

See also: Deterministic Aperiodic Solids, Electronic Properties of; Electrons in Periodic Potentials with Magnetic Fields; Semiconductor and Metallic Clusters, Electronic Properties of.

PACS: 71.10.Ay; 71.10.Ca; 71.27.+a; 71.28.+d; 71.30.+h; 71.90.+q; 73.21.−b

Further Reading

Abrikosov AA (1988) *Fundamentals of the Theory of Metals*, part I. Amsterdam: North-Holland.

Abrikosov AA, Gorkov LP, and Dzialoshinski JE (1963) *Methods of Quantum Field Theory in Statistical Physics*, chs. 1 and 4. New York: Dover.
Anderson PW (1984) *Basic Notions of Condensed Matter Physics*, ch. 3. Menlo Park, CA: The Benjamin Cummings.
Baym G and Pethick C (1991) *Landau Fermi-Liquid Theory*. New York: Wiley.
Coleman P (1995) Condensed matter: strongly correlated electrons Physics, World December issue, pp. 29–34.
Georges A, Kotliar G, Krauth W, and Rosenberg MJ (1996) Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Reviews of Modern Physics* 68: 13–125.
Gross EKU, Runge E, and Heinonen O (1991) *Many-Particle Theory*. Bristol: Adam Hilger.
Hewson AC (1993) *The Kondo Problem to Heavy Fermions*, ch. 5. Cambridge: Cambridge University Press.
Kuramoto Y and Kitaoka Y (2000) *Dynamics of Heavy Electrons*. Oxford: Clarendon.
Lifshitz EM and Pitaevskii LP (1999) *Statistical Physics*, part 2, ch. 1. Oxford: Butterworth-Heineman.
Mahan G (1990) *Many-Particle Physics*. New York: Plenum.
Pines D and Nozieres P (1988) *The Theory of Quantum Liquids*, 2nd edn. Redwood City: Addison-Wesley.
Spałek J (2000) Correlated Fermions: A new paradigm in physics on example of solid-state physics. *European Journal of Physics* 21: 511–534.
Spalek J, Datta A, and Honig JM (1987) Discontinuous metal–insulator transitions and Fermi liquid behavior of correlated electrons. *Physical Review Letters* 59: 728–731.
Varma CM, Littlewood PB, Schmitt-Rink S, Abrahams E, and Ruckenstein AE (1989) Phenomenology of the normal state of Cu–O high-temperature superconductors. *Physical Review Letters* 63: 1996–1999.
Vollhardt D (1984) Normal ^{3}He: an almost localized Fermi liquid. *Reviews of Modern Physics* 56: 90–120.

Liquids, Theory of: Non-Fermi Liquids *See* Liquids, Electronic Structure of.

Lithography Masks and Pattern Transfer

C Romeo and P Cantù, STMicroelectronics, Agrate Brianza, Italy

Introduction

Integrated circuits (ICs) are commonly manufactured using "planar technology." The different layers of an electronic device are fabricated with a sequence of technological processes (deposition, lithography, etching, and implantation) opportunely integrated.

Electronic performances and financial costs are essential goals in the semiconductor business and both are dependent on the ability of printing small feature sizes (transistors, connections, etc.) or more precisely small critical dimensions (CDs).

Lithography is the technique used to print these features; therefore, its role in the semiconductor industry is of paramount importance.

In a very schematic way, the process can be described as follows:

The electrical design of each layer is drawn on a lithographic photomask or reticle (both terms are used interchangeably) in a way that the circuit features are made of an absorbent material, while the remaining background is transparent to light; the reticle (which is partially transparent) is exposed to a source of light (see **Figure 1**) and its image is projected through a lens (4–5× reducing power) onto a silicon wafer previously covered with a photosensitive polymer (resist).

The light modifies the chemical nature of the exposed resist making it soluble to a specific solution. A development treatment then, with a basic solution, (0.26 N), removes the exposed resist (positive process), transferring the mask design onto the resist

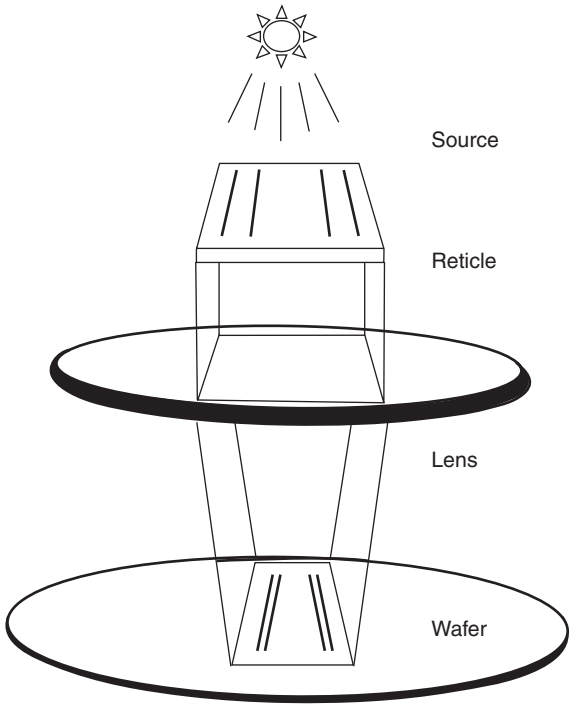


Figure 1 Schematic representation of the photomask exposure onto a silicon wafer in a lithographic process for IC fabrication.

layer that will be used to protect areas of the wafer from metal-oxide etching, implantation, etc. After the resist is removed, these process steps are repeated a number of times according to the complexity of the IC.

Currently, the semiconductor industry is mass-manufacturing ICs with a minimum CD of 130 nm (130 nm node), and the production of 90 nm electronic devices (90 nm node) has begun. Lithographic exposure tools use wavelengths of 248 and 193 nm with chemical-amplified photo-resists and phase-shifting reticles for the most critical layers.

There is a general consensus among IC makers that future developments will be the 65 nm node in 2005, the 45 nm node in 2007, and the 32 nm node in 2010.

Lithographically, this can be achieved using smaller wavelengths (13.5 nm for extreme ultraviolet lithography), a new technique called immersion lithography, lenses with higher numerical aperture, phase-shift masks (PSMs), and more advanced photo-resist processes.

The following paragraphs first describe how the reticle image is formed at the wafer level and then, a more detailed description of photomask characteristics and their manufacturing process is given.

Aerial Image Formation

Photomask or Reticle Diffraction

The interaction of electromagnetic waves with a reticle produces an image intensity profile in accordance with the diffraction theory of light. In order to describe this phenomenon, consider for simplicity, a binary reticle with equal lines and spaces; $l = s = d$ and with periodicity $P = 2d$.

For simplicity, this reticle is considered to be illuminated uniformly (Köhler's illumination) by a plane monochromatic wave λ, normally incident.

The diffracted light produces a Fraunhofer diffraction pattern with a maxima of intensity in the directions given by the angle θ:

$$\sin\theta - \sin\theta_0 = \frac{m\lambda}{nP}, \quad m = 0, \pm 1, \pm 2, \ldots \qquad [1]$$

where θ_0 is the angle of incidence, m is an integer, and n is the refraction index of the medium surrounding the mask and the lens.

If z is the distance between the reticle and the image plane (see **Figure 2**), it is known from the optical theory that in the far field approximation

$$z \gg \frac{\pi}{\lambda}(x^2 + y^2)$$

the Fraunhofer diffraction pattern in the plane (x_0, y_0) is given by the Fourier transform of the reticle transmission function calculated at the special frequencies $p = x_0/(\lambda z), q = y_0/(\lambda z)$, where z is the distance between the reticle and the image plane.

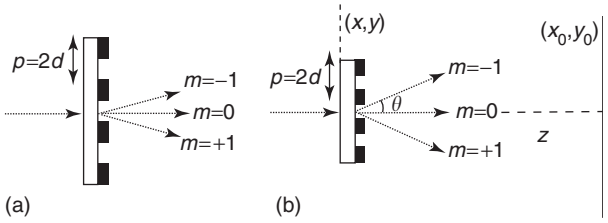


Figure 2 (a) A ray of coherent light normally incident ($\theta_0 = 0$) onto a reticle is diffracted in different directions. (b) For a smaller pitch P or larger λ, the diffraction angle θ becomes larger.

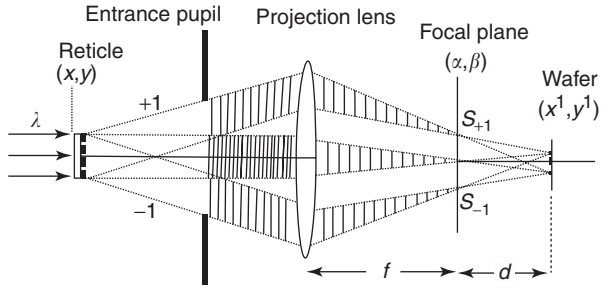


Figure 3 Reticle image formation at the wafer plane.

According to [1], the diffraction angle θ of each Fourier component is larger for a smaller pitch or for a larger λ (**Figure 2**). In other words, a lens facing a reticle with the scope of collecting the diffraction orders transmitted must have a larger diameter when the feature sizes are smaller. A diffraction-limited lens will faithfully reconstruct the image recombining all the Fourier components. However, in a real situation where the lens dimension is finite, only a few orders are collected and the image is partially reconstructed.

Image Formation

The image formation of the reticle in the image plane (wafer level) can be explained with the Abbe's theory of image formation. According to this theory, the image arises from a double diffraction process: the wave plane is diffracted first by the reticle and then by the lens (**Figure 3**).

The wave front transmitted through the reticle becomes a set of plane waves or spatial frequencies traveling in different directions (m_0, $m_{\pm 1}$, $m_{\pm 2}$, etc.). The Fraunhofer diffraction pattern obtained in the focal plane (α, β) at a distance f from the lens is proportional to the Fourier transform of the reticle

transmission function $G(x, y)$:

$$H(\alpha, \beta) = A_0 \int \int_C G(x, y) \times \exp - ik\left[\frac{\alpha}{f}x + \frac{\beta}{f}y\right] dx\ dy \qquad [2]$$

where A_0 is a constant and the integration is done over the reticle area C.

The maxima of intensity in the focal plane, S_0, $S_{\pm 1}$, $S_{\pm 2}$, etc., can be envisioned as emitters of Huygens waves that interfere in the reproducing of the image of the reticle in the image plane (x^1, y^1) of the lens.

The perfect reconstruction of the reticle image can be obtained considering that all the spatial frequencies are collected by the lens. In a real situation, the lens has a finite aperture, therefore, the highest frequencies are cut off. Only low special frequency components, corresponding to small θ in [1], passing close to the center of the lens, contribute to the image formation.

The reticle image can be reconstructed if there are at least two interfering plane waves; therefore, the lens must be large enough to capture the order $m = 0$ and $m \geqslant 1$. The minimum resolved feature is then

$$\sin\theta = \frac{\lambda}{nP} \qquad [3]$$

or

$$d = \frac{1}{2}\frac{\lambda}{n\sin\theta} = \frac{1}{2}\frac{\lambda}{NA} \qquad [4]$$

The above equation is known as Rayleigh's resolution limit for lines and spaces, for coherent illumination, where $n\sin\theta = NA$ is the numerical aperture of the lens.

It can be demonstrated that in the case of incoherent illumination, the expression can be rewritten as

$$d = \frac{1}{1+\sigma}\left(\frac{1}{2}\frac{\lambda}{n\sin\theta}\right) = \frac{1}{2}\frac{\lambda}{NA}\frac{1}{(1+\sigma)} \qquad [5]$$

where σ represents the partial coherence factor which is a measure of the physical extent of the light source. The larger the light source, the larger the σ. For a point source, $\sigma = 0$ and one has coherent illumination , for $\sigma = 1$ one has a large source and incoherent illumination.

As already said, the image of the reticle on the plane (x^1, y^1) is given by the Fraunhofer diffraction of the waves emitted by S_0, $S_{\pm 1}$, $S_{\pm 2}$, etc., and the total field distribution is

$$E(x^1, y^1) = A_1 \int \int_B H(\alpha, \beta) \times \exp - ik\left[\frac{x^1}{d}\alpha + \frac{y^1}{d}\beta\right] d\alpha\, d\beta \qquad [6]$$

where d is the distance between the focal plane and the image plane, $k = 2\pi/\lambda$, and B is the area of the aperture on the focal plane assuming that its dimension is smaller than d.

Using [2], the expression [6] can be rewritten as

$$E(x^1, y^1) = A_0 A_1 \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} P(\alpha, \beta) \times G(x, y) \exp - i\frac{k}{f}\left[\left(x - \frac{x^1}{M}\right)\alpha + \left(y - \frac{y^1}{M}\right)\beta\right] dx\ dy\ d\alpha\ d\beta \qquad [7]$$

$G(x, y)$ is assumed as zero for all points that are outside C; the pupil function $P(\alpha, \beta)$ is the transmission function in the pupil plane (equal to 1 for transmitted frequencies $\leqslant NA/\lambda$, and zero for blocked components); $f/d = -1/M$ where M is the magnification factor of the lens that for leading edges lithography equipment is 4 or 5 (reticle 4–5 times larger than its image on the wafer).

Generally, the intensity at the wafer level can be obtained by [7]:

$$I(x^1, y^1) = |E(x^1, y^1)|^2 = \left| A_0 A_1 \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} P(\alpha, \beta) \times G(x, y) \exp - i\frac{k}{f}\left[\left(x - \frac{x^1}{M}\right)\alpha + \left(y - \frac{y^1}{M}\right)\beta\right] dx\ dy\ d\alpha\ d\beta \right|^2 \qquad [8]$$

In the case of a photomask with N equidistant spaces (N large) of width s and with pitch P, and if the aperture in the focal plane is such that $-a \leqslant \alpha \leqslant a$, eqn [7] can be written in x^1-direction:

$$E(x^1) = E_0\left[1 + 2\sum_{1<m<m_0} \frac{\sin(m\pi s/P)}{m\pi s/P} \cos\frac{2\pi m x^1}{-dP/f}\right]$$

where $m_0 = aP/\lambda f$. In the case of $m_0 = \infty$, the summation can be changed with an integration and the image collected in the image plane (or wafer level) becomes similar to the object.

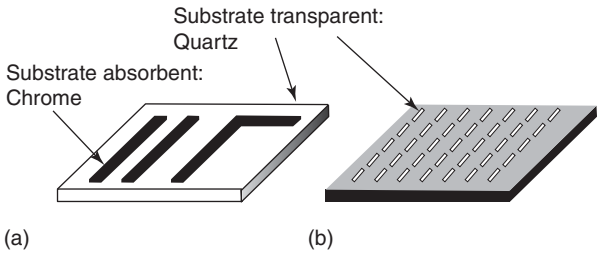


Figure 4 Schematic view of two binary masks for two different electronic layers: (a) lines and spaces (active area, poly, metal, etc) and (b) contact level.

Definition of Photomask

A mask is a transparent substrate selectively covered with a thin opaque layer, intended for the reproduction of patterns on wafers. In the case of optical photomasks, the pattern, which represents one level of an IC, is formed in a chromium layer sputtered over a fused silica substrate. Typically 20 to 30 masks, defined as a mask set, are required for imaging all layers necessary for the manufacturing of an IC. The terms mask and reticle are used interchangeably although historically, the first one was used for defining 1X patterned substrates used for exposing the whole wafer with contact printing technology, while the term reticle is used to define 2X, 4X or 5X patterned substrates used in optical projection systems. **Figure 4** represents a schematic of photomasks for two different layers: (a) lines and spaces, and (b) contact holes.

Types of Photomask

One may distinguish optical photomasks in two main categories: binary masks and PSMs. Masks for EUV applications, stencil masks or membrane masks for X-ray or electron projection lithography are not specifically discussed.

Binary Mask

A binary mask is a close replication of the circuit design pattern, with the possible addition of biases (i.e., the CD can be modified compared to the original design) in order to take into account lithography for optical proximity corrections (OPCs) or etch loading effects; its patterned area can be only clear or opaque, hence the term binary.

Phase-Shift Masks

PSMs were introduced in 1982 to improve the ultimate resolution achievable with wafer scanner exposure tools. The principle of this method consists in taking advantage of phase differences caused by different areas on the mask to obtain destructive interference and achieve better image contrast. PSMs can be separated in two categories: "weak PSMs," that include half-tone and embedded attenuated phase shift masks (EAPSMs), and "strong PSMs" that include alternating phase shift masks (APSMs). In the case of EAPSMs, the opaque region of the binary mask is substituted by a partially transmitting material that induces a 180° phase variation with respect to the bright regions. The destructive interference between clear areas and a partially transmitting background enhances the image contrast of the bright region. The improved image quality allows one to achieve a larger process latitude and depth of focus leading to a robust lithographic process. Usually 6% transmittance is used for leading edge masks; a higher transmittance allows one to achieve better contrast, due to a greater phase interaction, but is practically unusable due to the risk of unwanted side-lobe printing. Specifications related to half-tone and EAPSMs can be found in *SEMI P29-0997: Guideline for description of characteristics specific to Halftone Attenuated Phase Shift Masks and Mask Blanks*.

In strong APSMs, chrome structures are bordered with clear areas of 180° phase difference that leads to a destructive interference and thus to an enhanced image contrast. While in EAPSMs, the phase and transmission variation is obtained using appropriate materials, the most common being moly-silicide (MoSi), in strong PSMs the phase variation is generally obtained directly by etching the quartz substrate. **Figure 5** summarizes the binary mask and the PSM imaging principle.

Photomask Manufacturing Process

The manufacturing process of a photomask is quite similar to the one used for producing silicon wafers: the mask blank, which consists in a very flat piece of glass covered by a thin layer of chrome, is spun with a thin film of resist. The resist is exposed to light or to electrons in order to define the circuit geometries, and is then developed so as to form the required pattern. After development, the chrome material is etched and, finally the residual resist is removed.

In order to provide more details, the steps of photomask manufacturing can be summarized as follows (**Figure 6**): (1) mask blank preparation; (2) pattern writing; (3) processing; (4) critical dimension metrology; (5) position accuracy metrology; (6) defect inspection; (7) cleaning; (8) defect repair; (9) pellicle attachment; and (10) final inspection.

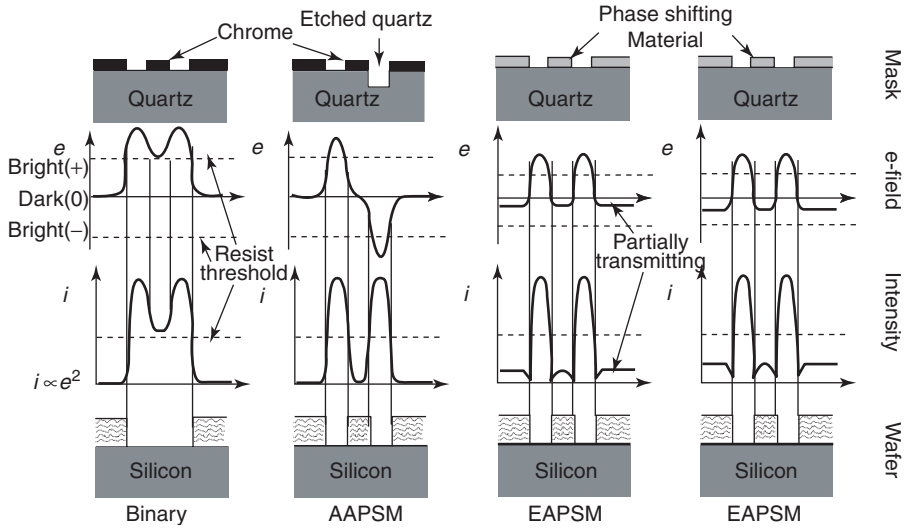


Figure 5 Comparison of binary and phase shift photomasks.

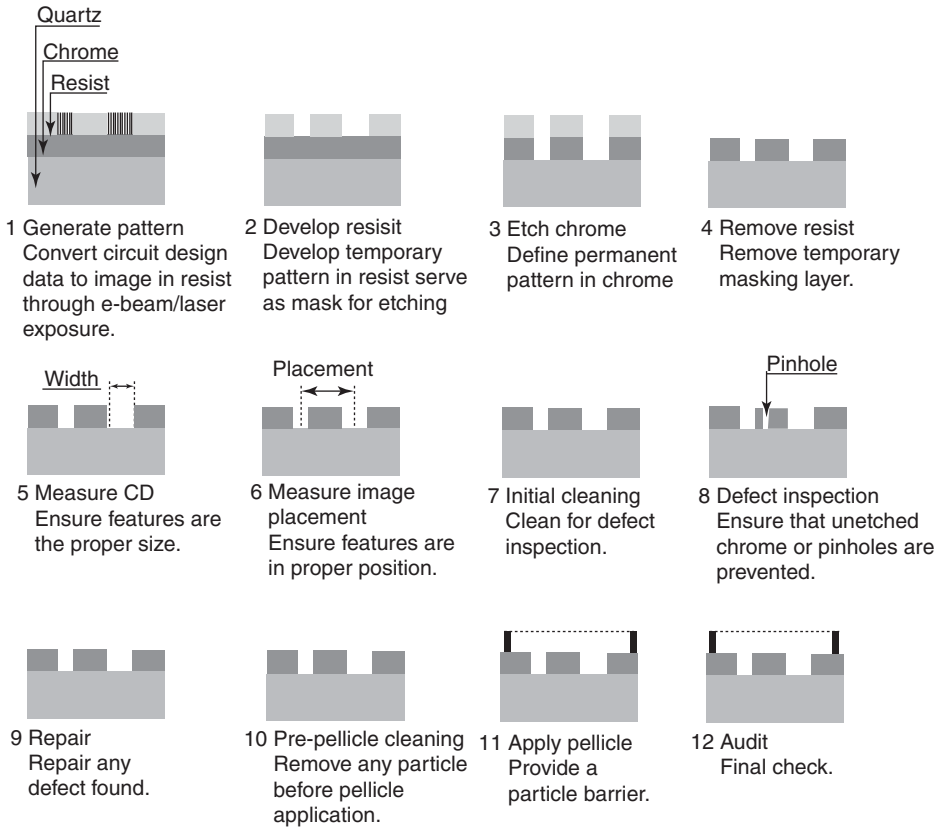


Figure 6 Photomask fabrication process.

Mask Blanks

Substrate Sizes

Photomask substrates requirements are defined in SEMI standard reticle format document (*SEMI P1-92 Re-approved P01-0299: Specification for Hard Surface Photomask Substrate*).

Actually, the standard size for leading edge masks is 15.24 cm square by 0.635 cm thick. The substrate size has been progressively increased in order to

Table 1 Optical characteristics of chrome film at different wavelengths

Wavelength (nm)	*Complex refraction index*	*Reflection from chrome glass interface*	*Transmission through* 60 nm *thick film of chromium*	*Effective transmitted light* (%)
436	1.79 *i*4.05	0.60	0.0009	0.036
365	1.39 *i*3.24	0.56	0.0012	0.055
248	0.85 *i*2.01	0.47	0.0022	0.119
193	0.84 *i*1.65	0.38	0.0016	0.098
157	0.68 *i*1.11	0.32	0.0048	0.310

accommodate larger and larger die sizes, while mask thickness has been increased from 0.09″ up to 0.25″ to minimize edge placement errors induced by the gravitational lag of the substrate as well as plate distortions caused by the clamping system of tools used for mask patterning and wafers exposure.

Substrate Materials

Mask substrate materials are chosen to comply with some mechanical and physical properties that can be summarized as follows: transparency at the exposure wavelength, thermal expansion, flatness, birefringency (only for 157 nm).

Low thermal expansion materials are required mainly for two reasons:

1. To minimize mask pattern placement errors induced by temperature variations that may happen during the mask patterning process, and
2. To minimize wafer pattern placement errors induced by mask heating that may happen when the mask is exposed to laser radiation in exposure tools.

A substantial portion of the placement error induced by temperature variations during mask patterning consists in magnification errors, which can be adequately compensated by wafer exposure tools, but residual errors caused by thermally instable material cannot be compensated.

A consideration of the depth of focus is the main driver for the tight control of reticle flatness which, for leading edge masks can be specified down to minor-equal 0.5 μm, depending on focus budget considerations.

Photomasks for leading technology use fused silica (SiO_2) substrates; this material has been chosen mainly for its low thermal expansion coefficient and for its transmission at current exposure wavelengths. For 157 nm exposure wavelength, fluorine-doped fused silica is introduced to maintain transparency at a sufficient level.

Opaque Film

The typical opaque film deposited on photomask blanks is chrome whose top surface is usually covered with a thin antireflective layer in order to minimize reflection of light reflected back by the wafer surface. The film characteristics are specified in *SEMI P2-86 Re-approved P2-0298: Specification For Chrome Thin Films For Hard surface Photomasks*. Chrome is normally deposited using a sputtering process with a thickness that may range from 50 to 110 nm. **Table 1** reports the optical characteristics of the chrome film at different wavelengths.

Several alternative materials have been tested over time; among others, a molybdenum-silicon layer has been proved to be a good alternative and is actually used for embedded phase-shift masks (EPSMs).

Mask Lithography

Mask writing tools are used to define the latent image of circuit patterns in the resist film spun over the mask substrate. Commercially available mask writing tools may be divided into two big categories:

1. electron-beam tools:
   (a) raster based (Gaussian beam);
   (b) vector based (beam shaped); and
2. laser tools.

Electron Beam Tools

Electron beam tools used for mask patterning are based on the same technology initially developed for electron microscopy. Energetic electrons induce chemical reaction in the resist, in the same way as photons do when using laser tools. One of the main advantages related to this kind of equipment, when compared to laser-based ones, is its very high resolution because electron diffraction occurs in the atomic range. Ultimate resolution is mainly limited by scattering of the electrons with the resist. To reduce this undesired effect, new tools adopt higher and higher beam energy together with complex electron proximity effect correction methodologies which are used to avoid dose variations induced by back-scattered electrons.

Electron beam tools may be divided into two subcategories: raster scan tools, which use a Gaussian beam, and vector shaped tools, where the beam shape is changed while patterning the mask.

Electron beam raster scan tools In raster scan tools, the design is divided into pixels, whose dimension, in the simplest implementation, is the one of the electron beam itself. The mask is scanned by deflecting the e-beam electromagnetically, while a high-precision stage, controlled by an interferometric system, moves the substrate. The beam is blanked or unblanked in synchronism with the deflection to form the desired circuit latent image in the resist. One should observe that the mask is progressively scanned for its whole area: even areas that should not be exposed must be scanned, leading to increased overall exposure time.

Electron shaped beam vector tools In vector tools, an electron beam is addressed just over the areas to be exposed: this approach dramatically reduces the exposure time. Additionally, these tools usually adopt a shaped beam, obtained by passing the electron beam through appropriate apertures. The pattern is obtained by exposing these elementary shapes that are butted or stitched together. Shaped beam tools have been introduced only recently, mainly due to the complexity required by their engineering, but are actually used for writing most of the leading edge masks as they allow one to achieve sharper corners and better resolution when compared to raster tools.

Laser Tool

Laser tools use an array of focused laser beam, with a roughly Gaussian profile, that is individually blanked or unblanked while scanning the mask. The light, generated by a laser, is split in a certain number of independent beams, the intensity of each beam being individually modulated using an acoustic–optic modulator. Laser beams are then scanned over the mask surface using a multifaced polygonal mirror rotating at high speed. During the exposure, the mask is clamped on a reticle stage that is moved under laser interferometric control.

Resist and Processing

Resists are radiation sensitive materials, typically composed of organic polymers and small molecules, used for enhancing lithographic performance. The exposure of resists to laser or e-beam radiation induces a change in solubility of the exposed regions compared to nonexposed ones: this property is used during the development step in order to define the pattern into the resist. The choice of the resist material is mainly related to the exposure tool environment, together with other lithographic properties such as sensitivity, contrast, and ultimate resolution achievable. Sensitivity is particularly important as it is directly related to the throughput of the tool. No material can meet sensitivity requirements for both laser and e-beam radiation, so two completely different processes are used by mask manufacturers depending on the writing tools used for patterning.

e-Beam Resists

Historically, the most common e-beam resist has been PBS poly (butane-1-sulfone), an alternating copolymer of 1-butene and sulfur dioxide, developed by Bell Laboratories. PBS is a positive resist usually developed using pentanone and other solvents such as methyl iso-amyl ketone. The need for better and better lithographic performances and the introduction of plasma etch of chrome brought about the development of alternative materials such as ZEP7000, a polymer of methylstyrene and chloromethyl acrylate from Nippon Zeon.

i-Line and DUV Resists

In general with optical ($\lambda = 365$ nm) resists, the process control can be reduced by the standing waves phenomena. A post-exposure bake (PEB) step is then usually applied in order to reduce the standing waves and improve the process control, as it happens in wafer lithography. New tools, operating at DUV wavelengths (248 nm), allowed the introduction of chemical amplified resists that obliged mask makers to start coating mask blanks themselves; historically mask blanks were pre-coated with appropriate resists by blank suppliers. PEB temperature control is quite critical and much more difficult than the one achievable with silicon wafers due to the mass of the substrate: that is the main reason why resist materials are selected with low sensitivity toward post-exposure bake temperature.

Chrome Etching

After resist development, the chrome opaque film is removed from the mask surface, using a wet-or-dry-etch process. Wet etch has been used for many years and it is based on a solution containing ceric ammonium nitrate mixed with nitric, perchloric, or acetic acid. Wet-etch processes are characterized by a non-negligible undercut that may contribute to line width variations, especially when the chrome thickness is not uniform. Additionally, patterns cannot be well defined if the undercut becomes comparable

with the dimension of the structures to be defined. For these reasons, the need for better and better pattern fidelity, together with the continuous shrinkage of CDs, pushed mask makers to move from wet-etch processes toward dry-etch processes. Typically, chrome is dry etched using a mixture of Cl and O_2 together with other gases added for reducing etch-loading effects that may bring iso-dense critical dimension variations.

Metrology

Three main parameters are measured once the mask is finished: feature dimension, pattern placement, and defect density.

Feature Dimension Control

The feature dimension control is performed to assure that geometry dimensions on the mask match the expected one. Three parameters related to CDs are usually reported: the average CD with respect to its nominal value, the CD uniformity across the reticle, generally reported as a three-sigma deviation of mask feature sizes for a single dimension, the feature size linearity down to a given dimension.

The average of a given CD with respect to its nominal value is specified as mask mean to target: which, for a given nominal CD value N (or target), is the difference between the mean of all measured CD values (CD_i) and their nominal N:

$$\text{Mask mean to target} = \frac{\sum_i (CD_i)}{n} - N$$

where n is the number of measurements.

More commonly, the CD accuracy is specified using the Max CD deviation, which, for a given nominal CD value N, is the greatest absolute value of the difference among the all measured CD values (CD_i) and their nominal N:

$$\text{Max CD deviation} = \text{MAX}|CD_i - N|$$

If more than one nominal CD is required, the Max CD deviation is the greatest among all the Max CD deviations for each nominal.

The CD uniformity across the mask is specified by the CD range which, for a given nominal CD value N (taking into account all measured CD values at that nominal) is the difference between the maximum and the minimum measured value:

$$CD_{\text{range}} = \text{MAX}(CD_i) - \text{MIN}(CD_i)$$

Finally, the linearity is specified as the difference between MAX and MIN "mask mean to target" for a range of features of the same nature, same tone, and different design sizes.

Mask CD control becomes significantly important when considering feature sizes near the diffraction limit. In such a condition, there is a change in magnitude by which feature sizes on the mask are reduced by lens reduction onto the wafer. This effect is described by the mask error factor (MEF):

$$\text{MEF} = \frac{\partial(CD_{\text{wafer}})}{\partial(CD_{\text{mask}}/M)}$$

where M is the imaging system reduction ratio, typically $M = 4$ for DUV lithography systems. For larger features, $\text{MEF} = 1$ but becomes bigger as lithography is pushed to the diffraction limit (i.e., for CDs close to the Rayleigh's resolution limit).

CD measurements have been performed for many years using optical microscopes coupled with software image analysis. Actually, CD measurements for leading edge masks are mainly performed using scanning electron microscope (SEMs) while several alternative methodologies are being explored for future technologies; these include scatterometry, Atomic force microscope (AFM), and optical immersion microscope.

Pattern placement

The pattern placement accuracy, or registration accuracy, refers to the deviation of pattern centerlines relatively to a defined reference grid. Registration is a vector quantity defined, at every point on the mask, as the difference between the vectorial position of the mask geometry and the vectorial position of the corresponding point on a reference grid. Pattern placement is measured on registration control tools that consists in a microscope, used for locating the feature to be measured, coupled with a high precision stage whose position is controlled by a laser interferometer.

Mask Defects and Inspection

A key requirement to mask makers is to produce "zero defect" reticles. A mask defect can be defined as any imperfection on the photomask pattern liable to be reproduced on a photoresist film, thus preventing the proper functioning of the microelectronic device being manufactured. One may distinguish defects in different categories:

- Macro defects: defects on mask that are visible to the naked eye, with or without the aid of a collimated spot light, for example, fingerprint left either by bare fingers or gloves;

- Micro defects: defects not visible to the naked eye, whose detection requires the use of automatic defect inspection systems or of any magnification tool.

One could also distinguish defects by their nature:

1. Soft defects, defined as defects that can be removed by a cleaning process:
   (a) particles
   (b) contaminations
   (c) process residuals
2. Hard defects, defined as defects that cannot be removed by a cleaning process
   (a) chrome protrusions
   (b) chrome intrusions
   (c) splintering: damaging of the substrate along its perimeter, glass missing
   (d) pinholes
   (e) scratches: an incision having regular or irregular shape and three dimensions: length, width, and depth where the length is much bigger than the width and the width itself is almost equal to the depth.
3. At each technological node (generation of semiconductor devices), the minimum CD is usually related to the minimum defect size allowable. In order to guarantee defect-free masks, automatic defect inspection tools have been developed. Inspection tools may work in two modes:
   - Die-to-die mode: automatic inspection of mask defectivity based on the comparison of at least two identical patterns, structured into regular arrays;
   - Die-to-database mode: automatic inspection of mask defectivity based on the comparison of the mask pattern against its original database.

The ability of an inspection tool to detect accurately all the defects present on a plate determines its defect capture efficiency that is related to many factors: inspection light wavelength, pixel size, sensors efficiency, real-time signal analysis, stage resolution, and many others. It is to be taken into account that high sensitivity requires small pixel sizes, and so higher inspection time and a higher chance to detect false defects: differences detected, even if below the defect specification requirement.

Once mask inspection is terminated, defects are normally manually reviewed by an operator who classifies all captured defects transferring coordinates of hard defects to be repaired to appropriate tools. Inspection equipment are certified by performing a mask inspection with pre-programmed defects of certified dimensions.

Repair

Mask defects may be repaired by proper tools, and several techniques have been developed over the years: laser tools are the most diffused together with focused ion beam. More recently, a mechanical method that utilizes the positional control of an AFM coupled with RAVE LLCs nanomachining head to perform material removal with nanometer level precision has been proposed.

Cleaning

The scope of cleaning is to remove resist residuals, particles, and all other possible contaminations without damaging the mask itself. Traditionally, the cleaning step has not been a major concern for the mask industry. Mask cleaning is usually performed by the use of wet methods developed by the wafer industry. In the past, immersion processes were most diffused for the cost of ownership reasons, while recently the tightening of cleaning specifications pushed mask manufacturers to introduce spin/spray processes. Several cleaning chemistries have been developed over the years: among the others, the most used are based on H_2SO_4 and H_2O_2 to remove heavy organic materials such as resist and hydrocarbons, or H_2O, H_2O_2, and NH_4OH to remove light organic residues and particles.

Pellicle

The requirement to produce "zero defect" masks has already been discussed. To avoid contamination of mask surface during its usage, pellicles are attached to the photomasks. Pellicles consist in a thin ($<1\,\mu m$) polymer film, usually made of nitrocellulose or a form of teflon, stretched on an anodized aluminum frame, that is attached to the photomask using special adhesives characterized by low outgassing and stability to exposure at actinic light. **Table 2** reports

Table 2 Pellicle materials commonly used

Materials	*Thickness* (μm)	*Transmission* (%)	*Wavelength* (nm)
Nitrocellulose	1.4	99	436 (g-line), 365 (I-line)
Cellulose	1.4	99	436 (g-line), 365 (I-line)
Fluorocarbon	1.2	99	365 (I-line), 248 (DUV)
Teflon	0.6	99	248 (DUV), 193 (EDUV)

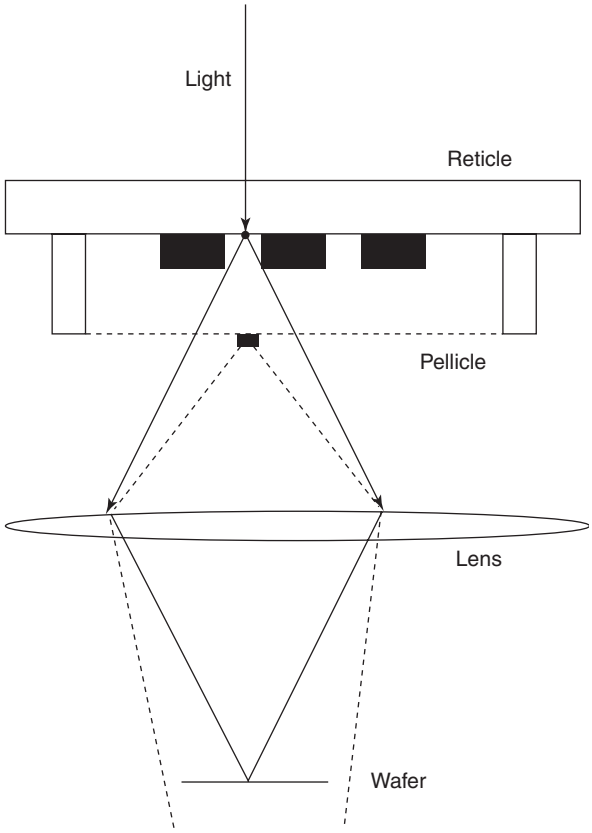


Figure 7 Cross-sectional view of a mask with pellicle. The particle accidentally deposited on the pellicle will produce an image severely blurred onto the wafer.

the most commonly used materials used for mask pellicles and their characteristics:

Characteristics of various pellicle materials are illustrated more in detail in the document *SEMI P5-94: Specification For Pellicles*. Pellicles keep contaminations far away from the chrome features that are to be printed; in that way, with a typical depth of focus $<1\,\mu m$, particles are not in the focus plane of the wafer and then are not printed on the wafer (**Figure 7**).

The typical frame height is in the range of 6 mm, small holes in the frame allow one to maintain air pressure of the space between the mask surface and the pellicle equal to the ambient pressure. The frame surface has to be quite flat so as to avoid the distortion of the mask that may induce registration and focus errors.

See also: Integrated Circuits; Semiconductors, General Properties; Semiconductors, History of; Silicon, History of.

PACS: 85.40.Hp; 42.82.Cr; 81.16.Nd; 42.50.St; 42.15.Dp; 42.30.Kq; 42.30.Lr; 42.30.Va; 42.70.Ce

Further Reading

Born M and Wolf E (1964) *Principles of Optics*, 3rd edn. Pergamon.

Choudhory PR (1997) *Handbook of Microlithography, Micromachining and Microfabrication, Microlithography*, vol. 1. Bellingham WA: SPIE Optical Engineering Press.

Gaskill J (1978) *Linear Systems, Fourier Transforms, and Optics*. New York: Wiley.

Larmore L (1995) *Introduction to Photolithographic Principles*, 2nd edn. New York: Dover.

Levinson HJ (2001) *Principles of Lithography*. Bellingham WA: SPIE Optical Engineering Press.

Wilson RJ (1994) *Fourier Series and Optical Transform Techniques in Contemporary Optics*. New York: Wiley.

Wong AKK (2001) *Resolution Enhancement Techniques in Optical Lithography*. Bellingham WA: SPIE Optical Engineering Press.

Local Field Effects

W Schattke, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

Introduction

If a dielectric material is exposed to an electric field, a dipole moment is established which is detected by the force the field exerts on it. It is a common experience to have felt the attraction acted by a charged external body as by a comb on one's hair. Such materials show induced polarization in an external electric field usually observed as surface charges disregarding the microscopic details of the interior. However, the field acts in the interior of the whole dielectric body and induces the polarization therein through the polarizability of the atomic constituents of the body. The well-known textbook Lorentz–Lorenz construction of the so-called "local field" is applied, which explicitly takes into account the presence of an induced field adding to the external field and being self-consistently determined at every point. In course of time, the Lorentz–Lorenz construction has been set on an *ab initio* fundament by considering the total many-body system, which also led to the quantitative determination by well-accepted definite rules though the whole problem cannot be solved in a closed manner.

From a macroscopic point of view, say, considering length dimensions in centimeters, the polarization is continuously spread over space, discontinuities arising at the edges of bounded bodies only. The total dipole moment causes a resulting induced field which adds to the external field usually screening it, that is, reducing it within the material. Of course, the sources for the induced fields may be inhomogeneously distributed and arbitrarily macroscopically shaped. For example, a spherical solid dielectric body introduced between the plates of a condenser expels the initially homogeneous field from the interior created by dipole charges on the surface of the sphere and situated opposite to the condenser plates. This general macroscopic situation has its origin on an atomic scale that finally gives rise to what is called "local field effects."

From a microscopic point of view, atomic polarization constitutes the source of macroscopic polarization. The atomic polarization is localized being attached to the electronic cloud around the nuclei and showing a global distribution of the dipole moments reflecting the positions of the nuclei. The local variability from granularity on an atomic scale relates to the notion of local fields. Thus, the local field considered here arises from the electric response of atomically structured material to an external field. It can be viewed as an induced granular field, which in the special case of a crystalline solid is structured by the details of the potential within the unit cell being periodically repeated. Such a local field occurs in insulators, as well as in semiconductors and metals. Crystallinity is a widely understood origin for the existence of local structure in the response fields, though the general atomic consistence of matter gives rise to local response fields also. However, the former are much easier to be dealt with than classifying or calculating, for example, amorphous or gaseous material.

The above interpretation is not to be confused with the notion of local field as applied to a theoretical extension of the random phase approximation (RPA), which is a very common and basic approximation in the calculation of dielectric properties of general condensed matter. This also traces back to the roots of the Lorentz local field, though the emphasis does not lie on the granularity and geometric structure but on the many-body influences which affect even homogeneous systems.

The first achievement along both lines may be seen in the Lorentz–Lorenz formula

$$\varepsilon_{\mathrm{M}} = 1 + \frac{\alpha/\varepsilon_0}{1 - \alpha/3\varepsilon_0} \qquad [1]$$

for the macroscopic dielectric constant ε_{M} in terms of atomic polarizability per unit volume α with vacuum permittivity ε_0 (SI units used).

Induced Field

For the theory of dielectrics, Maxwell's equations of electrostatics constitute the framework

$$\mathrm{div}\, \boldsymbol{D} = \rho, \;\; \mathrm{rot}\, \boldsymbol{E} = 0 \qquad [2]$$

to be supplemented by the appropriate material's equation, specifically

$$\boldsymbol{D} = \varepsilon\varepsilon_0 \boldsymbol{E} \qquad [3]$$

for a linear relation between the so-called dielectric displacement $\boldsymbol{D}$ and the electric field $\boldsymbol{E}$ with dielectric constant or relative permittivity ε of the medium. The charge density ρ refers only to the sources of the external field $\boldsymbol{E}_{\mathrm{ext}}$ which can be written as $\mathrm{div}\, \varepsilon_0 \boldsymbol{E}_{\mathrm{ext}} = \rho$. One usually splits off the electric field from $\boldsymbol{D}$ by defining the polarization $\boldsymbol{P}$, that is, the dipole moment per unit volume, and the external field from the microscopic field by defining the induced field $\boldsymbol{E}_{\mathrm{ind}}$

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \;\; \boldsymbol{E} = \boldsymbol{E}_{\mathrm{ext}} + \boldsymbol{E}_{\mathrm{ind}} \qquad [4]$$

Inserting eqn [4] in [2] yields

$$\mathrm{div}\, \varepsilon_0 \boldsymbol{E}_{\mathrm{ind}} = -\mathrm{div}\, \boldsymbol{P} = \rho_{\mathrm{ind}} \qquad [5]$$

This identifies $-\mathrm{div}\, \boldsymbol{P}$ as the source of the induced field associating with the induced polarization charge ρ_{ind}, an analogous to the external charge, which is a rather trivial statement from the atomic point of view.

In principle, the electrostatic relations have to be complemented in order to deal with optics. The topic of local field effects, however, allows a limitation to the discussion of the longitudinal case, that is, to the above equations of electrostatics, decoupling it from the transverse fields introduced by the time-dependent electromagnetic field. The approximation is valid as long as the length scale of the induced field is small compared to the optical wavelength. Transverse components are of course nonzero and need the full dielectric tensor which may be anisotropic. But the simplification occurs because the self-consistent generation of the total electric field from the applied external field (see the section "Macroscopic fields") derives this approximation solely from the longitudinal field, for which Maxwell's electrostatic equations are already sufficient.

The above simple relationship [3] is, however, a compact abbreviation of the most general linear relation which should be written as

$$D_i(\boldsymbol{r},t) = \sum_j \int_{-\infty}^{+\infty} \mathrm{d}t' \int \mathrm{d}^3\boldsymbol{r}' \varepsilon_{ij}(\boldsymbol{r},t;\boldsymbol{r}',t')\varepsilon_0 E_j(\boldsymbol{r}',t') \quad [6]$$

where the indices refer to the Cartesian axes of a tensor relation through ε_{ij}, the time integration sums up contributions of the electric field at various times to the dielectric displacement, and the space integration does similarly with respect to position. In physical terms, the tensor character relates an electric field in a certain direction to the displacement in any other direction with a specific weight, for example, as an anisotropic material would imply. The time integration takes into account a delay in the effect of the electric field on, for example, the polarization, which for a causal relation obviously implies zero ε for $t' > t$. Switching on a field or, in terms of Fourier transforms, of an incident optical wave which as an oscillating process, excites an oscillating polarization with a weight depending on the frequency, this is the standard case of optical spectroscopy. The dependence on position in space, is the interesting quantity in view of the effects considered here. An electric field at a specific position in principle can excite polarizations everywhere and all contributions are summed up by the integral. The above quantity ε is usually called the microscopic dielectric constant, and "constant" is often replaced by "function" to stress the functional dependence on its variables.

Macroscopic Fields

Measurements as, for example, in classical optical spectroscopy with wavelengths orders of magnitude larger than atomic distances may not resolve the atomic structure of matter, and therefore deal with spatially averaged quantities instead of the microscopic fields considered above. The averaging follows from space integration or from the zeroth Fourier component when expressing the field equations by Fourier transforms as conveniently done. Dealing with the example of a crystalline solid, the periodic repetition of the unit cell imposes this symmetry also on all observable quantities as fields, densities, coefficients of linear relations in position, space, etc., as long as the external fields are homogeneous. Then, the periodicity allows for a representation of the total field by a Fourier series

$$\boldsymbol{E}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \boldsymbol{E}_{\boldsymbol{G}} \quad [7]$$

omitting the time or respective frequency dependence in the notation. The vectors $\boldsymbol{G}$ constitute the reciprocal lattice of the crystal. A nonhomogeneous external field may vary on a macroscopic scale, then $\boldsymbol{E}_{\boldsymbol{G}}$ will also depend on $\boldsymbol{r}$. However, the length scales of this field as well as those of the additional variations of the induced fields are large. In which case, all considerations still apply locally within ranges limited by that length scale. The linear relation [3] reads, under the Fourier transformation, as

$$\boldsymbol{D}_{\boldsymbol{G}} = \sum_{\boldsymbol{G}'} \hat{\varepsilon}_{\boldsymbol{G}\boldsymbol{G}'} \boldsymbol{E}_{\boldsymbol{G}'} \quad [8]$$

where the hat is used to indicate a tensor instead of writing the Cartesian indices explicitly. If a homogeneous electric field is applied to an infinitely extended crystal, then the displacement field is constant everywhere and thus possesses only a zeroth Fourier component which is given by a suitable distribution (e.g., charged condenser plates) of its sources, the external charges. Inversion of eqn [8] then yields

$$\boldsymbol{E}_{\boldsymbol{G}} = (\varepsilon^{-1})_{\boldsymbol{G}0} \boldsymbol{D}_0 \quad [9]$$

the Fourier components of the microscopic field, that is, local field which, in principle may spatially fluctuate on each length scale within the periodicity cell. Taking only the zeroth component of eqn [9], one obtains the observable averaged electric field $\boldsymbol{E}_0$ and can write the material's equation as

$$\boldsymbol{D}_0 = \frac{1}{(\varepsilon^{-1})_{00}} \boldsymbol{E}_0 \quad [10]$$

This linear relation between averaged quantities constitutes what is called the local field effect, specifically a macroscopic permittivity ε_{M} which is the reciprocal of the (0,0) element of the inverse microscopic dielectric function matrix $\varepsilon_{\boldsymbol{G}\boldsymbol{G}'}$

$$\varepsilon_{\mathrm{M}} = \frac{1}{(\varepsilon^{-1})_{00}} \quad [11]$$

and applies to homogeneous external fields.

Furthermore, it also covers the case that the field additionally depends on $\boldsymbol{q}$ within the Brillouin zone describing larger length scales than the lattice length. Then also ε carries this additional dependence, (see the section "Dielectric response"). As indicated before, the derivation of eqn [11] is valid for the longitudinal response to longitudinal fields, so the hat was omitted because it is a scalar quantity. The transverse average fields are related through a similar reasoning. The general tensor can be written as $\hat{\varepsilon}_{\mathrm{M}}$ consisting of the bare ε and an added longitudinal correction

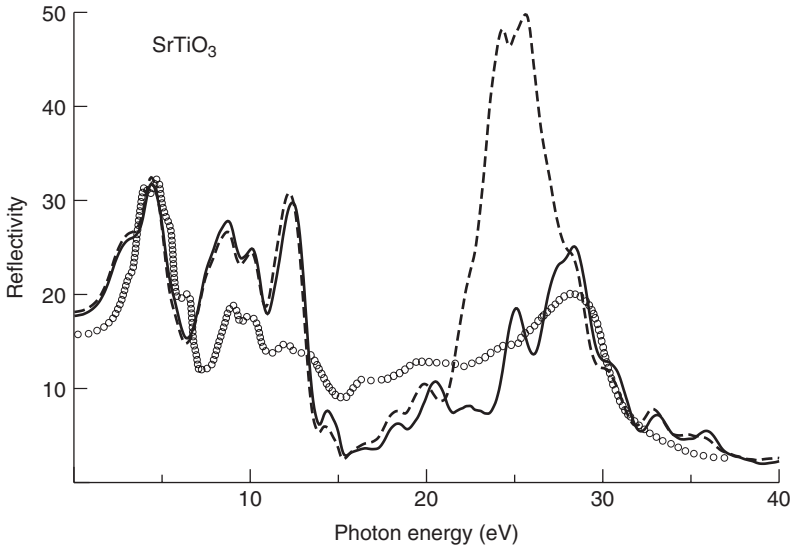


Figure 1 Calculated reflectivity of $SrTiO_3$, with local field effects included (—) and neglected (- - - -) compared to experiment (○).

which originates from the polarization by the Coulomb interaction.

$$\hat{\varepsilon}_{\mathrm{M}} = \hat{\varepsilon}_{00} - \sum_{\mathrm{GG}'(\neq 0)} \hat{\varepsilon}_{0\mathrm{G}} \, (\hat{\varepsilon}_{\mathrm{r}}^{\mathrm{LL}})^{-1}_{\mathrm{GG}'} \, \hat{\varepsilon}_{\mathrm{G}'0} \qquad [12]$$

The quantity $\hat{\varepsilon}_{\mathrm{r}}^{\mathrm{LL}}$ is the microscopic dielectric constant matrix in reciprocal space restricted (subscript r) to the subspace of nonzero reciprocal lattice vectors. Only the longitudinal component of the tensor is involved which is denoted by the upper index, that is, the tensor can be written as the product of the dyad $(\boldsymbol{q}+\boldsymbol{G})/|\boldsymbol{q}+\boldsymbol{G}| \cdot (\boldsymbol{q}+\boldsymbol{G}')/|\boldsymbol{q}+\boldsymbol{G}'|$ with a scalar $\varepsilon^{\mathrm{LL}}_{GG'}$. Taking the longitudinal component of eqn [12], algebraic manipulations lead back to eqn [11], $\varepsilon_{\mathrm{M}}^{\mathrm{LL}} = 1/(\varepsilon^{\mathrm{LL}})^{-1})_{00}$, with inversion of the whole unrestricted matrix. This generalization refers to the local field effect of optical and related spectra valid as long as the wavelength $\lambda = 2\pi/q$ is much larger than the length scale of the atomic structure. In **Figures 1** and **3**, calculated optical spectra, up to the vacuum ultraviolet, of $SrTiO_3$ and Nb with or without the local field effect are compared with experiment. The importance of the local field effect in describing the experimental result is striking. As a matrix inversion is involved, a lucid picture is absent to discuss these effects in simple terms, for example, the disappearance of the huge peak at 25 eV. The comparison of theory with experiment shows agreement with respect to the relative positions of maxima and minima, whereas the strength of the variation is smoothed by experimental broadening not considered in the theoretical calculation. It should be noted that the theoretical curves are obtained *ab initio* with the density-functional theory,

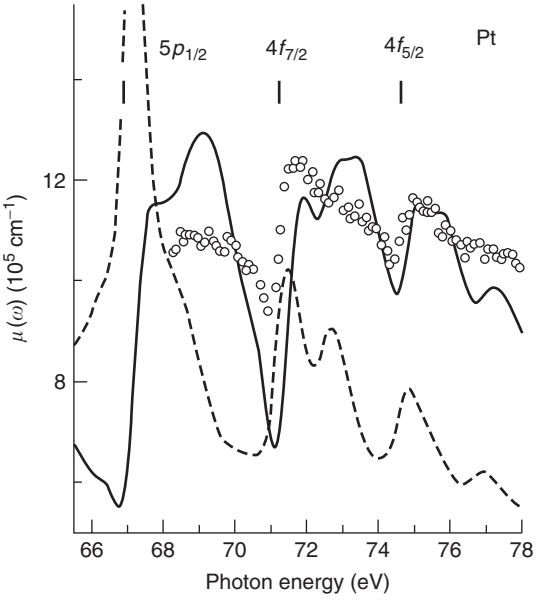


Figure 2 Theoretical absorption coefficient of Pt including local field effects (—) compared with experiment (absorption: ○, XPS energy levels: (‖)), and neglecting local field effects (- - - -).

and broadening from additional many-body corrections might also occur.

Occasionally the connection with the local polarization is obvious in the spectra as shown in **Figure 2**. A steep slope marks the onset of absorption from the localized semicore levels, $5p_{1/2}$, $4f_{7/2}$, $4f_{5/2}$, in both experiment and theory. Specifically, the $5p_{1/2}$ peak and also the relative intensities need local field effects to be included for explanation. Both 4*f* peaks are considerably enhanced by the accompanying high atomic polarization of the medium proving this to be

the origin of local field effects, even though the overall shape of a spectrum may often hide this connection, for instance, at low-energy excitations of valence states.

Dielectric Response

Maxwell's theory describes the field from its source, that is, the total charge. Conversely, the total field forces the positional adjustment of the charges through the laws of quantum mechanics, which is condensed into the dielectric constant ε. Once the bare response to a small test field is known through ε, the numerical evaluation of charge and field from eqns [2] and [10] is straightforward by using a suitable mathematical solution. Thus, one has to determine quantitatively the dielectric function. As already stated, the RPA is the basic approximation for calculating these properties *ab initio* with generally sufficient accuracy. The quantum mechanical change of the density $\delta\rho$ induced as a response to a change of the field $\delta\boldsymbol{E}$ is obtained from first-order perturbation expansion with respect to the latter. The total field must be treated here as a perturbation. Equation [2] is used in Fourier space, $i(\boldsymbol{q}+\boldsymbol{G})\delta\boldsymbol{D}(\boldsymbol{q}+\boldsymbol{G})=\delta\rho(\boldsymbol{q}+\boldsymbol{G})$, and the field in eqn [3] is replaced by its associated potential, $\delta\boldsymbol{E}(\boldsymbol{q}+\boldsymbol{G})=-i(\boldsymbol{q}+\boldsymbol{G})\delta\phi(\boldsymbol{q}+\boldsymbol{G})$. Then, the induced charge is given by the difference between total charge $\delta\rho_{\text{tot}}=\varepsilon_0|\boldsymbol{q}+\boldsymbol{G}|^2\delta\phi$ and external charge $\delta\rho$ with the result

$$\sum_{G'}(\boldsymbol{q}+\boldsymbol{G})\cdot(\hat{1}\delta_{G,G'}-\hat{\varepsilon}_{GG'}(\boldsymbol{q},\omega))\cdot(\boldsymbol{q}+\boldsymbol{G}') \times\delta\phi(\boldsymbol{q}+\boldsymbol{G}')=\delta\rho_{\text{ind}}(\boldsymbol{q}+\boldsymbol{G})/\varepsilon_0 \qquad [13]$$

This has to be compared with the expression obtained for the induced charge from the perturbation expansion, which finally leads to

$$\varepsilon^{\text{LL}}_{GG'}(\boldsymbol{q},\omega) =\delta_{GG'}+\frac{e^2}{2\pi^2\varepsilon_0|\boldsymbol{q}+\boldsymbol{G}|\,|\boldsymbol{q}+\boldsymbol{G}'|}\sum_{\lambda\lambda'}\int_{\text{BZ}}\mathrm{d}^3\mathbf{k} \times\frac{\langle\boldsymbol{k}+\boldsymbol{q}\lambda'|\mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G}')\mathbf{r}}|\mathbf{k}\lambda\rangle\langle\mathbf{k}\lambda|\mathrm{e}^{-\mathrm{i}(\mathbf{q}+\mathbf{G})\mathbf{r}}|\mathbf{k}+\boldsymbol{q}\lambda'\rangle}{E_{\lambda'}(\boldsymbol{k}+\boldsymbol{q})-E_{\lambda}(\boldsymbol{k})-\hbar\omega-i\hbar\eta} \qquad [14]$$

Here, a single matrix element of the dielectric function for reciprocal lattice vectors $\boldsymbol{G}$ and $\boldsymbol{G}'$ is represented by a sum of occupied (λ) and unoccupied (λ') states and an integral over an elementary cell (Brillouin zone) of the reciprocal lattice in an expression which, of course, reminds of the quantum mechanical perturbation theory. It contains the denominator for the energy difference between the ground and the excited states of the system including the photon energy $\hbar\omega$, and the numerator with the matrix elements for a transition between both states caused by the wave part of the Fourier decomposition of the perturbing potential. The wave vector $\boldsymbol{q}$ which is confined to the first Brillouin zone appears only once as an independent variable of ε which is diagonal with respect to it. Owing to translational symmetry, charge and field can differ in total momentum only by reciprocal lattice vectors.

One observes again that only longitudinal fields are involved in this electrostatic treatment and consequently only longitudinal components ε^{LL} of the dielectric tensor

$$\varepsilon^{\text{LL}}_{GG'}(\boldsymbol{q},\omega)=\frac{(\boldsymbol{q}+\boldsymbol{G})}{|\boldsymbol{q}+\boldsymbol{G}|}\cdot\hat{\varepsilon}_{GG'}(\boldsymbol{q},\omega)\cdot\frac{(\boldsymbol{q}+\boldsymbol{G}')}{|\boldsymbol{q}+\boldsymbol{G}'|} \qquad [15]$$

are accessed. They constitute, nevertheless, the most important contributions in the domain of optics and related fields.

Exchange and Correlation

As mentioned before, the notion of local field effects can also be applied to the many-body extensions of RPA. Because of complexity, their investigation cannot claim a success similar to the above-discussed standard effects. For simplicity, the main work is usually directed towards homogeneous systems though a few quantitative calculations also exist for solids. In **Figure 3**, such corrections for considering exchange and correlation (XC) which go beyond RPA are shown. The observed influence is rather small compared to the usual local field effects. Since the comparison of *ab initio* calculations with experimental spectra still leaves many questions open and the importance of many-body corrections is generally unknown, this is an active field of present research.

Nonlocal Dielectric Response

The term nonlocal response can be taken literally as a far reaching effect of a local perturbation. The local field effects above have been considered under the restriction of smoothly varying external fields on a macroscopic scale, that is, long wavelengths ($q=2\pi/\lambda\approx 0$). This condition is relaxed here. However, both effects interfere in the experiment and one has to be discerning in its interpretation. A change of the field at one site can result in changes of the dielectric displacement at very distant sites which, of course, is also contained in the general dependence of eqn [6]. As this dependence of the field on the charges is provided by a nonlocal long range relation,

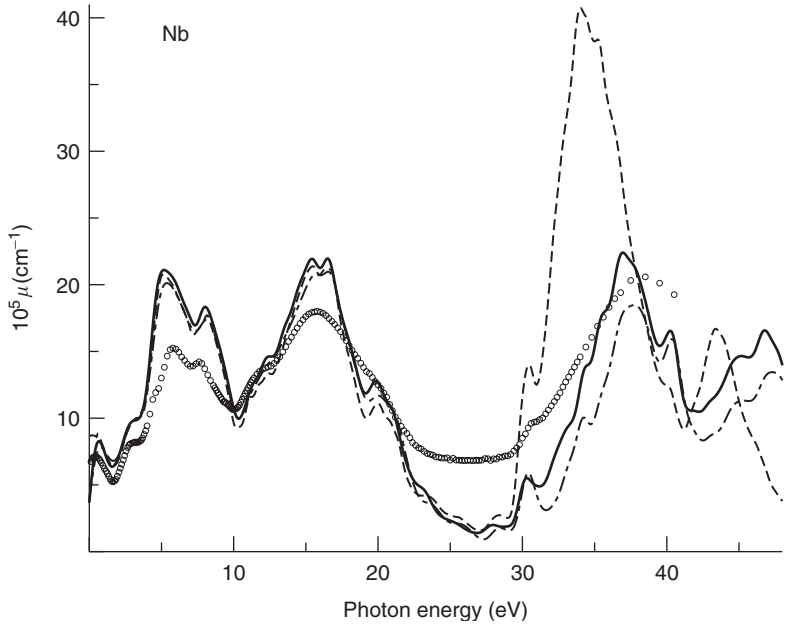


Figure 3 Calculated absorption coefficient of Nb, with local field effects included (–· –·–·–) and neglected (- - - -) compared to experiment (○); solid line includes XC corrections neglected otherwise.

Coulomb's law, each local internal polarization by the field causes induced fields and hence, dielectric displacement everywhere. This holds in the static as well as in the dynamic case where the plasmon excitation is a very prominent long-range feature.

This rather general theoretical effect will not be easily detected in measurements on homogeneous systems, it becomes visible by local perturbations as, for example, a surface imposes. There the surface may be the source for plasmon waves which travel into the bulk and modulate the field. In addition, even below the plasmon frequency multipole surface, plasmons enhance the frequency-dependent nonlocal response. The theoretical treatment, essentially governed by eqn [6], can be based on eqn [13] with the potential representing the full self-consistent field in time-dependent local density approximation (LDA)

$$\delta\phi = \phi_{\text{ext}} + \delta\phi + \delta V_{\text{XC}} \qquad [16]$$

Self-consistency is imposed by coupling to the charge density as before in RPA, in fact with Poisson's equation through $-\Delta\delta\phi = \delta\rho/\varepsilon_0$, and by using the exchange–correlation potential in LDA through $\delta V_{\text{XC}} = V'_{\text{XC}}(\rho_0)\delta\rho$. Because of the surface, translational symmetry is broken in one direction and calculation is restricted to real space in that direction utilizing an iterative process based on self-consistency. The centroid $d_\perp(\omega)$ of the charge density is suitable to characterize the surface response, that is, it yields the position of the electromagnetically effective surface by its real part and the absorbed power by its imaginary part. For a jellium model, it shows a resonance at ~ 0.8 of the plasma frequency.

The combined effect of local fields and nonlocal response shows a rather complex behavior as seen in **Figure 4**. The excitation from an external field localized just below the surface extends into the bulk with a penetration depth increasing with frequency and attains long-range behavior at the surface plasma frequency ($\hbar\omega = 6.7\,\text{eV}$) and above the bulk ($\hbar\omega = 9.2\,\text{eV}$) plasma frequency. Closely around the surface, a dipole charge distribution establishes which is replaced by a complex multipole behavior above the bulk plasma frequency. The granularity in the depth dependence reflects local field effects.

Further Application

In addition to traditional optical spectroscopy, the reflection anisotropy/difference spectroscopy (RAS/RDS) offers newer surface-sensitive optical tools for access to details of the electromagnetic response.

Local field effects are of importance not only in optical and related spectroscopy but they strongly affect electron spectroscopy as, for example, electron losses (EELS) or O photoexcited electrons of angle resolved ultraviolet photremission spectroscopy (ARUPS), too. The loss function of EELS

$$S(\boldsymbol{q},\omega) = -\text{Im}\frac{1}{\varepsilon_{\text{M}}(\boldsymbol{q},\omega)} \qquad [17]$$

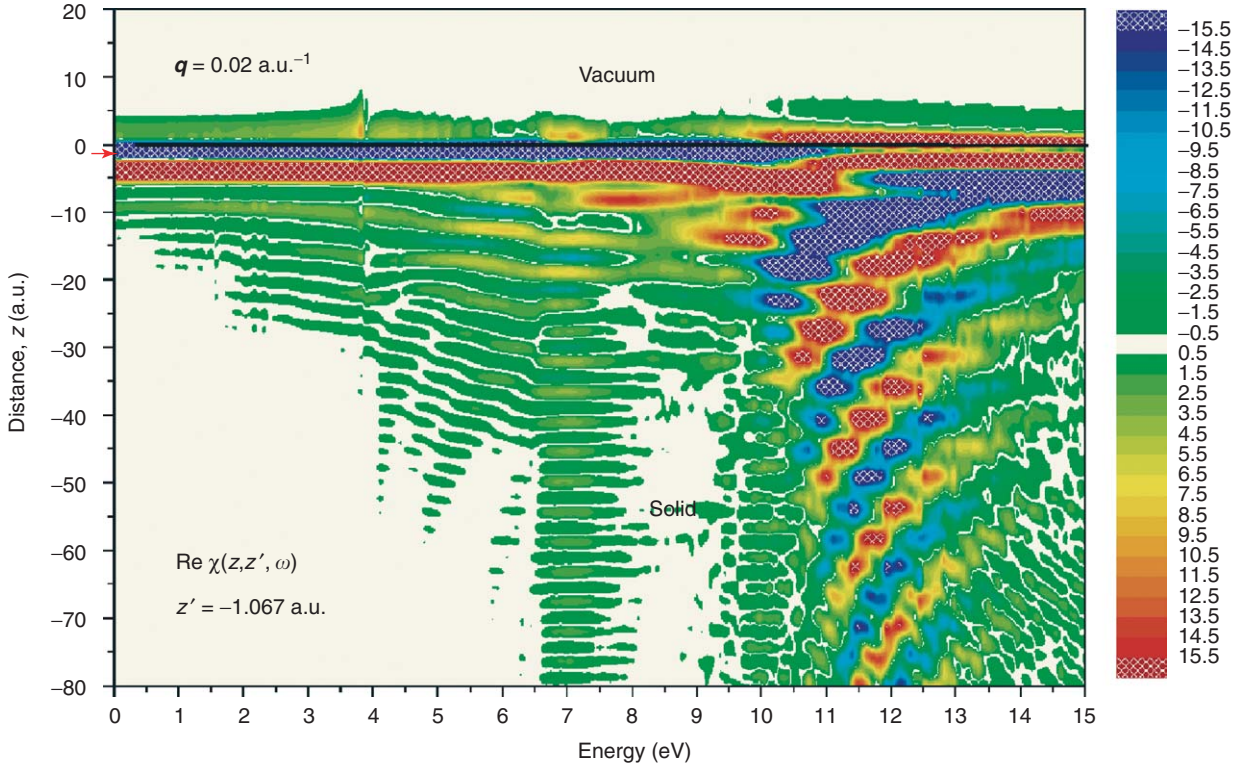


Figure 4 Susceptibility of Ag (1 1 1) slab, i.e., charge density at varying depth z from the surface at $z=0$ as response to a local excitation potential at z' with lateral wave vector $\boldsymbol{q}$, plotted vs. excitation frequency, model potential used for bulk (Courtesy of Silkin VM and Chulkov EV, Donostia International Physics Center (DIPC).)

measures the macroscopic ε_M. ARUPS reflects the local fields themselves as they cause specific photoexcitations, and also depend sensitively on the nonlocal multipole plasmon generation. Their main effects are observed near the plasmon frequency and can attain an order of magnitude change.

Recent development of photonics towards the nanoscale regime opens the door to the details of the position dependence of electromagnetic response. For example, small metallic particles close to a molecule on a surface may produce surface enhanced Raman scattering (SERS) of that molecule by many orders of magnitude, because the field considerably increases in this geometrical configuration. At sufficiently small scales, nonlocal effects will become important such that the full dielectric function will be needed for theoretical considerations there also.

Acknowledgments

The author is indebted to E E Krasovskii for providing figures and to the Deutsche Forschungsgemeinschaft for the support under FOR 353.

See also: Dielectric Function; Electrodynamics: Continuous Media; Excitons in Crystals; Optical Absorption and Reflectance; Optical Properties of Materials.

PACS: 03.50.De; 72.22.−d; 77.22.−d; 77.22.Ch; 78.20.Ci; 78.20.−e; 78.68.+m

Further Reading

Ehrenreich H (1965) *The Optical Properties of Solids*. New York: Academic.

Feibelman PJ (1982) Surface electromagnetic fields. *Progress in Surface Science* 12: 287–407.

Forstmann F and Gerhardts RR (1986) *Metal Optics Near the Plasma Frequency*, *Springer Tracts in Modern Physics*, vol. 109. Berlin: Springer.

Liebsch A (1997) *Electronic Excitations at Metal Surfaces*. New York: Plenum.

Mochán WL and Barrera RG (1985) Electromagnetic response of systems with spatial fluctuations. *Physical Review B* 32: 4984–5001.

Nie S and Emory SR (1997) Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science* 275: 1102.

Schattke W and Van Hove MA (2003) *Solid-State Photoemission and Related Methods*. Weinheim: Wiley-VCH.

Slater JC (1967) *Insulators, Semiconductors, and Metals*. New York: McGraw-Hill.

Localized and Itinerant Magnetism

H Yamada, Shinshu University, Matsumoto, Japan
T Goto, University of Tokyo, Kashiwa, Japan

Introduction

The quantum theory of magnetism has been developed from two opposite starting points. One is the localized-electron model, where the electrons remain on the atoms in the crystal and contribute to magnetism. The intra-atomic interaction between these electrons is so strong that the magnetic moment is determined on each individual atom by Hund's rules. The interatomic exchange interaction is, on the other hand, much smaller than the intra-atomic one, and competes with thermal disorder to determine the long-range magnetic ordering. This model is applied to insulating crystals and rare-earth metals. The other model of magnetism is based on the band structure of electrons. These electrons move around in the crystal and contribute to the electric conduction. The ferromagnetic state may be stabilized for materials with a strong exchange interaction and a high density of states at the Fermi level. This is the band model or itinerant-electron model, which is commonly applied to 3d transition metals, alloys, and compounds.

In this section, magnetic properties as well as magnetic ordering are discussed for both localized-electron and itinerant-electron models. A molecular field theory is extremely useful in describing the ground state and finite temperature properties. In this approximation, the focus is on how various magnetic properties can be explained, and then, how spin wave excitations and spin fluctuations at finite temperature can be studied. SI units are used in this article. The magnetic induction, $\boldsymbol{B}$, is given by $\mu_0(\boldsymbol{H}+\boldsymbol{M})$, where $\boldsymbol{H}$, $\boldsymbol{M}$, and μ_0 are the magnetic field, magnetization, and permeability in vacuum, respectively. The Bohr magneton is denoted by $\mu_B = e\hbar/2m$ in this unit system, where e and m are the charge and mass of an electron, respectively, and where $\hbar$ is the Planck constant divided by 2π. To avoid confusion, μ_B is taken to be positive, so that the direction of magnetization is the same as that of the spin.

Localized-Electron Model

The concept of exchange coupling between the spins of two or more nonsinglet atoms arose with the Heitler–London theory of chemical bonding and was applied in 1928 to the theory of ferromagnetism by Heisenberg. The spin Hamiltonian is given by $\mathscr{H} = -\sum_{i\neq j} J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j$, where J_{ij} is the interatomic exchange integral between the electrons on atoms at $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$. $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ are spin operators on these atoms. The exchange integral J_{ij} comes from the antisymmetric nature of wave functions with respect to the spatial and spin coordinates, and is given by

$$J_{ij} = \frac{e^2}{4\pi\varepsilon_0}\int\int \frac{\psi_j^*(\boldsymbol{r})\psi_i^*(\boldsymbol{r}')\psi_i(\boldsymbol{r})\psi_j(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'-\boldsymbol{R}_i+\boldsymbol{R}_j|}\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'$$

where $\psi_i(\boldsymbol{r})$ is the wave function on the atomic site $\boldsymbol{R}_i$. The value of J_{ij} becomes small obviously when $|\boldsymbol{R}_i - \boldsymbol{R}_j|$ is large, as $\psi_i(\boldsymbol{r})$ is localized on the atom at $\boldsymbol{R}_i$ in this model. Here, J_{ij} is assumed to take a constant value, J for the nearest neighbor atomic pair at $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$, and to be zero for other pairs. In this case, the Hamiltonian is written by $\mathscr{H} = -2J\sum_{\langle i,j\rangle}\boldsymbol{S}_i\cdot\boldsymbol{S}_j$, where $\langle i,j\rangle$ on the summation denotes the sum over nearest neighbor pairs at $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$.

Ferromagnetism

Before the concept of quantum mechanics had been established, Weiss proposed in 1907 a very successful description of ferromagnetism by introducing a molecular field acting on atomic magnetic moments, which is assumed to be proportional to the net magnetization. The origin of the molecular field is nowadays known to be a quantum mechanical exchange interaction. A material, which consists of the same type of atoms, is considered. When the magnetic field is applied to the z direction (quantized axis of spin), the average value (or expectation value) of the spin operator $\boldsymbol{S}_i$ has a z component $\langle S_z\rangle$ only. The magnetization is now written as $M = Ng\mu_B\langle S_z\rangle$, where N is the number of atoms in the material, $g = 2$ (g-factor for spin system), and μ_B is the Bohr magneton. The term $\boldsymbol{S}_i\cdot\boldsymbol{S}_j$ in the Hamiltonian is approximately written by $\boldsymbol{S}_i\cdot\boldsymbol{S}_j \simeq \langle S_z\rangle S_{j,z} + S_{i,z}\langle S_z\rangle - \langle S_z\rangle^2$. One gets, together with the Zeeman energy by the magnetic induction B, $\mathscr{H} \simeq -2zJ\langle S_z\rangle\sum_i S_{i,z} - g\mu_B B\sum_i S_{i,z} + \mathrm{NzJ}\langle S_z\rangle^2 = -g\mu_B B_{\mathrm{eff}}\sum_i S_{i,z} + NzJ\langle S_z\rangle^2$, where z is the number of the nearest neighbor atoms in the crystal. Here, J is taken to be positive as the ferromagnetic state is discussed. The effective magnetic field is given by $B_{\mathrm{eff}} = 2zJ\langle S_z\rangle/g\mu_B + B$. The first term $2zJ\langle S_z\rangle/g\mu_B$ corresponds to the molecular field proposed by Weiss. The molecular field coefficient ω, defined by $B_{\mathrm{eff}} = \omega M + B$, is given by $\omega = 2zJ/N(g\mu_B)^2$.

The magnetic quantum number m_S on an atom may take $S, S-1, \ldots, -S$. The average value $\langle S_z \rangle$ is evaluated by the standard statistics as $\langle S_z \rangle = SB_S(g\mu_B SB_{eff}/k_B T)$ at finite temperature T, where $B_S(x)$ is the Brillouin function defined by

$$B_S(x) = \frac{2S+1}{2S}\coth\left(\frac{2S+1}{2S}x\right) - \frac{1}{2S}\coth\left(\frac{x}{2S}\right)$$

where $x = g\mu_B SB_{eff}/k_B T$ and k_B is the Boltzmann constant. (In the limiting case of $S = \infty$, $B_S(x)$ becomes the Langevin function, which appears when $S_{i,z}$ is treated as a classical spin vector.) The average value $\langle S_z \rangle$ is obtained as a function of T and B by solving the equation for $\langle S_z \rangle$ self-consistently, as the argument in the Brillouin function contains $\langle S_z \rangle$ itself. **Figure 1** shows the calculated result of the magnetization for $S = 5/2$ as a function of T and B. T_C is the Curie temperature where $\langle S_z \rangle$ becomes zero at $B = 0$. The T-dependence of M at $B = 0$ is very weak at low T in this theory, being different from the observed results. At low T, M decreases with increasing T as $T^{3/2}$ by the thermal excitations of spin waves. This is discussed in the subsection, "Spin waves."

The susceptibility χ_p $(= \mu_0 M/B)$ above T_C is written at high T as $\chi_p = \mu_0 C/(T - T_C)$. This is called the Curie–Weiss law. Here, T_C and C are the Curie temperature and the Curie constant respectively, given by $T_C = 2zJS(S+1)/3k_B$ and $C = N(g\mu_B)^2 S(S+1)/3k_B$. This is derived by using the approximation $B_S(x) \sim x(S+1)/3S$ at small x. At $T = T_C$, M is proportional to $B^{1/3}$. Then, the susceptibility χ_p diverges as $B^{-2/3}$ when B tends to zero. The critical index δ, defined by $M \propto B^{1/\delta}$, is 3 in the molecular field theory. However, the observed result of δ for conventional ferromagnets is ~ 4. The molecular field theory breaks down at $T \simeq T_C$.

Antiferromagnetism

The crystal, which consists of two equivalent sublattices A and B, may have an antiferromagnetic spin structure when the exchange integral J_{AB} between the nearest neighbor spins on the different sublattices is negative. Sublattice moments are given by $M_A = Ng\mu_B\langle S_z^A \rangle/2$ and $M_B = Ng\mu_B\langle S_z^B \rangle/2$, where the number of atoms on each sublattice is $N/2$. Here, S_z^A and S_z^B are the z components of S on these sublattices. Magnitudes of M_A and M_B without magnetic fields are equal to M, that is, $M_A = -M_B = M$.

Effective fields acting on the A and B sublattice spins are written by $B_{eff}^A = \omega' M_A - \omega M_B + B$ and $B_{eff}^B = -\omega M_A + \omega' M_B + B$, where ω and ω' are the molecular field coefficients of inter- and intra-sublattice moments given by $\omega = 4z_{AB}|J_{AB}|/N(g\mu_B)^2$ and $\omega' = 4z_{AA}J_{AA}/N(g\mu_B)^2$. Here, J_{AA} $(= J_{BB})$ is the intra-sublattice exchange integral and z_{AB} and z_{AA} $(= z_{BB})$ are the numbers of the nearest neighbor atoms on the different sublattices and on the same sublattice. By the same procedure as used in the ferromagnetic case, the thermal averages $\langle S_z^A \rangle$ and $\langle S_z^B \rangle$ are obtained by the Brillouin function as $\langle S_z^A \rangle = SB_S(x_A)$ and $\langle S_z^B \rangle = SB_S(x_B)$, where $x_A = g\mu_B SB_{eff}^A/k_B T$ and $x_B = g\mu_B SB_{eff}^B/k_B T$. The Néel temperature T_N, where the sublattice moment at $B = 0$ vanishes, is given by $T_N = 2S(S+1)(z_{AB}|J_{AB}| + z_{AA}J_{AA})/3k_B$. Solving the simultaneous equations

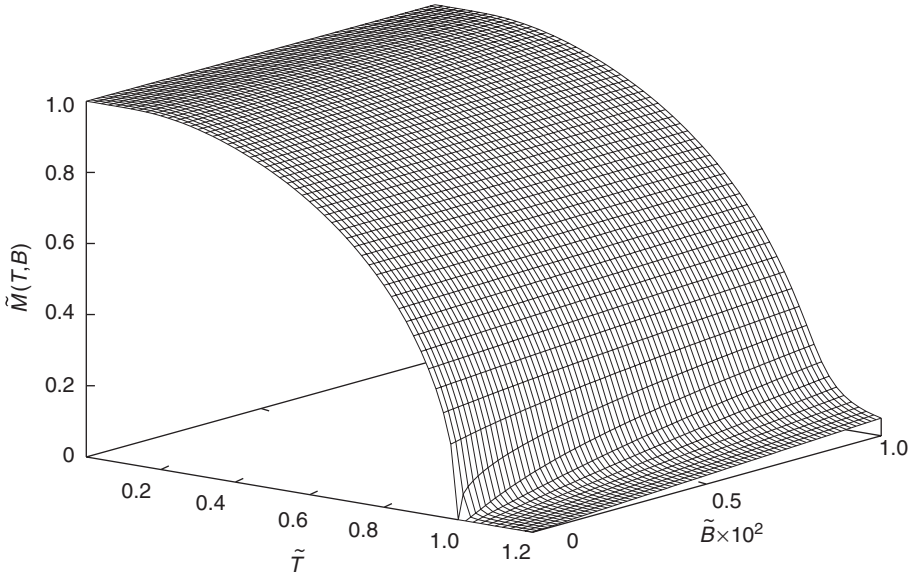


Figure 1 Temperature and field dependences of the magnetization M for $S = 5/2$. $\tilde{M} = M/Ng\mu_B S$, $\tilde{T} = T/T_C$ and $\tilde{B} = g\mu_B B/2zJS$.

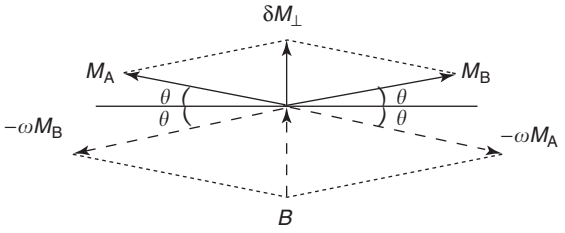


Figure 2 Magnetic induction B of the applied field and molecular fields $-\omega M_A$ and $-\omega M_B$ on the sublattice moments M_A and M_B, respectively. $\delta M_\perp$ is the induced moment.

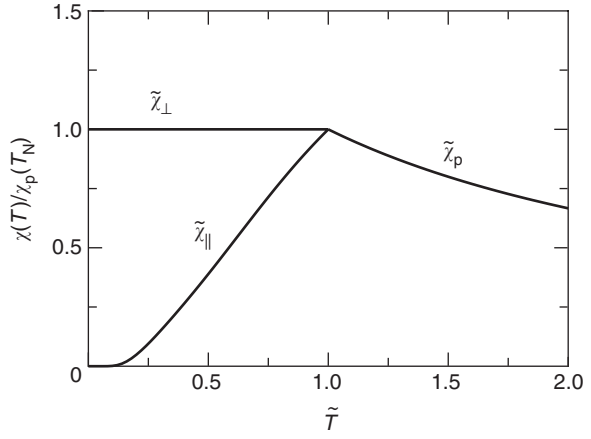


Figure 3 Parallel and perpendicular susceptibilities $\tilde{\chi}_\parallel(T)$ and $\tilde{\chi}_\perp(T)$ below T_N, and of $\tilde{\chi}_p(T)$ above T_N for $S = 5/2$ and $\omega' = 0$. $\tilde{\chi}_\parallel(T) = \chi_\parallel(T)/\chi_p(T_N)$, $\tilde{\chi}_\perp(T) = \chi_\perp(T)/\chi_p(T_N)$, $\tilde{\chi}_p(T) = \chi_p(T)/\chi_p(T_N)$, and $\tilde{T} = T/T_N$.

for M_A and M_B, the susceptibility $\chi_p(T)$ above T_N is obtained as $\chi_p(T) = \mu_0 C/(T + \Theta_p)$, where $\Theta_p = 2S(S+1)(z_{AB}|J_{AB}| - z_{AA}J_{AA})/3k_B$ and $C = N(g\mu_B)^2 S(S+1)/3k_B$. At $T = T_N$, $\chi_p(T_N)$ is given by $\chi_p(T_N) = \mu_0/\omega$. When $J_{AA} = 0$, T_N is equal to Θ_p. The observed value of T_N for conventional antiferromagnets is smaller than that of Θ_p.

Below T_N, there are two kinds of susceptibility. One is the susceptibility when the magnetic field is applied along the direction of M_A. In this case, M_A increases by $\delta M_\parallel/2$, and M_B shrinks by $\delta M_\parallel/2$. $\delta M_\parallel$ is obtained by solving simultaneous equations of M_A and M_B given by the Brillouin functions with B^A_{eff} and B^B_{eff}. Noting that $M_A = -M_B = M$ at $B = 0$, one gets the susceptibility $\chi_\parallel(T)(= \mu_0 \delta M_\parallel/B)$ parallel to the sublattice moment as $\chi_\parallel(T) = \mu_0 C/\{\Theta_p + (S+1)T/\ 3SB'_S(x)\}$, where $x = 6MT_N/Ng\mu_B(S+1)T$ and $B'_S(x)$ is the derivative of $B_S(x)$ with respect to x. At $T = 0$, $\chi_\parallel(0) = 0$ as $\lim_{T\to 0} T/B'_S(x) = \infty$. At $T = T_N$, $\chi_\parallel(T_N) = \chi_p(T_N) = \mu_0/\omega$ as $B'_S(0) = (S+1)\ /3S$. The other is the susceptibility when the magnetic field is applied in the perpendicular direction to M_A and M_B. In this case, the induced moment $\delta M_\perp/2$ appears on each sublattice in the field direction. The susceptibility $\chi_\perp(= \mu_0 \delta M_\perp/B)$ perpendicular to the sublattice moment is obtained by the balance of effective fields shown in **Figure 2**. The sublattice moments rotate by an angle θ so that $\boldsymbol{B} - \omega(\boldsymbol{M}_A + \boldsymbol{M}_B) = 0$. One gets $\chi_\perp = \mu_0/\omega$, which does not depend on T and also on the value of ω'. In **Figure 3**, $\chi_\parallel(T)$ and $\chi_\perp(T)$ below T_N, and $\chi_p(T)$ above T_N are shown as a function of T for $S = 5/2$ and $\omega' = 0$. For polycrystalline or powder samples, the antiferromagnetic moments in domains or particles will orientate in random directions, so that $\chi(T) = \chi_\parallel(T)/3 + 2\chi_\perp(T)/3$ below T_N.

Magnetizing Process of Antiferromagnets

For the magnetizing process of antiferromagnets, magnetic anisotropy plays an important role. The magnetic moment prefers to align in a certain crystallographic direction without a magnetic field. For the sake of simplicity, consider the case of uniaxial anisotropy. Taking the easy axis of magnetization along the z-axis, the sublattice moment $\boldsymbol{M}_A$ prefers to align in the positive (negative) z direction and $\boldsymbol{M}_B$ in the negative (positive) z direction. The anisotropy energy is phenomenologically given by $E_{ani} = -(K_u/2)(\cos^2\theta_A + \cos^2\theta_B)$, where θ_A and θ_B are angles between the sublattice moments and the easy axis. The coefficient K_u is called the uniaxial anisotropy constant. The anisotropy fields, B^A_{ani} and B^B_{ani} acting on the A and B sublattice moments, are defined by $B^A_{ani} = K_u M^z_A/M^2$ and $B^B_{ani} = K_u M^z_B/M^2$.

The induced moment $\delta M_\perp$ by the field applied perpendicular to the z-axis (hard axis of magnetization) is suppressed by the anisotropy field. The susceptibility $\chi_\perp$ is given by $\chi_\perp = \mu_0/(\omega + K_u/2M^2)$. This is derived in the following way. E_{ani} is rewritten as $(K_u/2)(\sin^2\theta_A + \sin^2\theta_B) - K_u$. The anisotropy field to the hard axis of magnetization is written by $-K_u \delta M_\perp/2M^2$ (negative field), and then one can get the above expression of $\chi_\perp$. On the other hand, when the magnetic field is applied along the easy axis of magnetization and when the sublattice moments align in the hard axis, the induced moment $\delta M_\perp$ in the field direction is enhanced by the anisotropy field. The susceptibility, in this case, is written by $\chi'_\perp = \mu_0/(\omega - K_u/2M^2)$.

When the magnetic field is applied along the easy axis of magnetization, the magnetic energy together with E_{ani} is approximately given by $\Delta E_\parallel = -\chi_\parallel(T)B^2/2\mu_0 - K_u$ at low fields. On the other hand, when the magnetic field is applied along the easy axis of magnetization and when the sublattice

moments align in the hard axis, the magnetic energy is given by $\Delta E_\perp = -\chi'_\perp(T)B^2/2\mu_0$. At weak magnetic fields applied along the direction of the easy axis, the sublattice moments keep to align in the easy axis by the gain of the anisotropy energy. However, under strong magnetic fields they may align in the direction perpendicular to the magnetic field, as $\chi'_\perp$ is larger than $\chi_\parallel$, and then, $\Delta E_\perp$ becomes lower than $\Delta E_\parallel$. That is, a spin flipping takes place from the parallel direction to B to the perpendicular one at a certain magnetic field. The critical value B_{sf} of the spin flipping is given by the equation $\Delta E_\parallel = \Delta E_\perp$, and one obtains $B_{sf}(T) = [2\mu_0 K_u/\{\chi'_\perp(T) - \chi_\parallel(T)\}]^{1/2}$. At $T = 0$, B_{sf} is written by $B_{sf} = \{B_{ani}(2B_{ex} - B_{ani})\}^{1/2}$, where $B_{ex} = \omega M$ and $B_{ani} = K_u/M$. When $B > 2B_{ex} - B_{ani}$, the sublattice moments become parallel to each other. The critical value is denoted by B_c $(= 2B_{ex} - B_{ani})$. The spin structures are the antiferromagnetic one at $B < B_{sf}$, the canted one at $B_{sf} < B < B_c$, and the ferromagnetic one at $B > B_c$. When T increases, B_c decreases, while B_{sf} increases slightly, as shown in **Figure 4**.

When B_{ani} is larger than B_{ex}, the canted spin state does not appear. The magnetization curve at $T = 0$ shows a jump from the antiferromagnetic state to the ferromagnetic one at $B = B_{ex}$. This is a field-induced metamagnetic transition. However, the transition at finite T becomes gradual at B_{ex}, as shown in **Figure 5**. When the magnetic field is applied to the hard axis of magnetization, the induced moment $\delta M_\perp$ increases as $\chi_\perp B/\mu_0$ up to $B'_C = 2B_{ex} + B_{ani}$ and saturates to $2M$.

Spin Waves

The low-lying energy states of spin systems coupled by exchange interactions are wave-like, as originally discussed by Bloch. This is called a spin wave or a magnon. Spin waves have been studied for all types of ordered spin arrays. In this subsection, the ferromagnetic spin wave is studied.

The energy associated with nonuniform distributions of the spin direction is given by the Heisenberg exchange energy $W = -J\sum_i \sum_\delta \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+\delta}$, where $\boldsymbol{\delta}$ is the nearest neighbor vector of $\boldsymbol{R}_i$. Here, the $\boldsymbol{S}$s are quantum mechanical spin operators. They are treated as if they were classical vectors. $\boldsymbol{S}_{i+\delta}$ is expanded in a Taylor series as $\boldsymbol{S}_i + (\boldsymbol{\delta}\cdot\nabla)\boldsymbol{S}_i + (1/2)(\boldsymbol{\delta}\cdot\nabla)^2\boldsymbol{S}_i$. For lattices with inversion symmetry, the first-order term in $\boldsymbol{\delta}$ is cancelled out by the summation over δ, and one gets $W = -J\sum_i |\boldsymbol{S}_i|^2 - (J/2)\sum_i\sum_\delta \boldsymbol{S}_i \cdot \{(\boldsymbol{\delta}\cdot\nabla)^2\boldsymbol{S}_i\}$. It is noted here that $\nabla^2(\boldsymbol{S}_i\cdot\boldsymbol{S}_i) = 0$, as $|\boldsymbol{S}_i|^2$ is a constant. Then, one can get the exchange energy for the local magnetization $\boldsymbol{M}(\boldsymbol{r}) = g\mu_B\sum_i \boldsymbol{S}_i\delta(\boldsymbol{r} - \boldsymbol{R}_i)$ as $E_{ex} = (1/2V)D\int |\nabla \boldsymbol{M}(\boldsymbol{r})|^2 \,d\boldsymbol{r}$, where $|\nabla \boldsymbol{M}(\boldsymbol{r})|^2 = \sum_\alpha\{(\partial M_\alpha/\partial x)^2 + (\partial M_\alpha/\partial y)^2 + (\partial M_\alpha/\partial z)^2\}$, $\alpha = x, y, z$, and V the volume of the sample. The coefficient D in E_{ex} is called the Landau–Lifshitz constant. The exchange field is defined by $\boldsymbol{B}_{ex}(\boldsymbol{r}) = \partial E_{ex}/\partial \boldsymbol{M}(\boldsymbol{r})$. Neglecting the boundary effects, one gets $\boldsymbol{B}_{ex}(\boldsymbol{r}) = D\nabla^2\boldsymbol{M}(\boldsymbol{r})$. Then, the torque equation is given by $\partial \boldsymbol{M}(\boldsymbol{r},t)/\partial t = \gamma D\boldsymbol{M}(\boldsymbol{r},t)\times\nabla^2\boldsymbol{M}(\boldsymbol{r},t)$, where γ is the gyromagnetic ratio $g\mu_B/\hbar$. One looks for travelling wave solutions of the form $M_x(\boldsymbol{r},t) = \delta m_x \exp\{i(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t\}$, $M_y(\boldsymbol{r},t) = \delta m_y \exp\{i(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t\}$

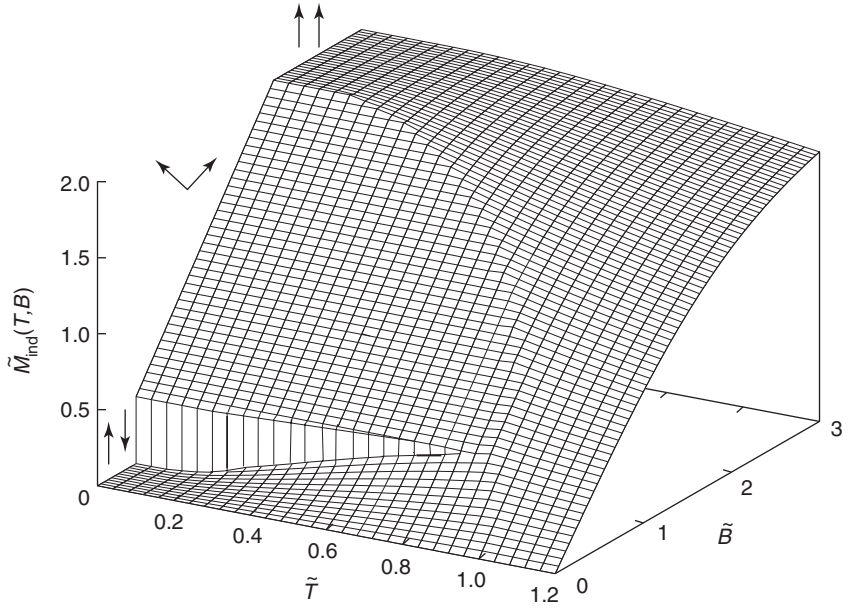


Figure 4 Temperature and field dependences of the induced moment M_{ind} for $S = 5/2$, $\omega' = 0$, and $2K_u/Nz_{AB}|J_{AB}|S^2 = 0.2$. $\tilde{M}_{ind} = 2M_{ind}/Ng\mu_B S$, $\tilde{T} = T/T_N$, and $\tilde{B} = g\mu_B B/2z_{AB}|J_{AB}|S$.

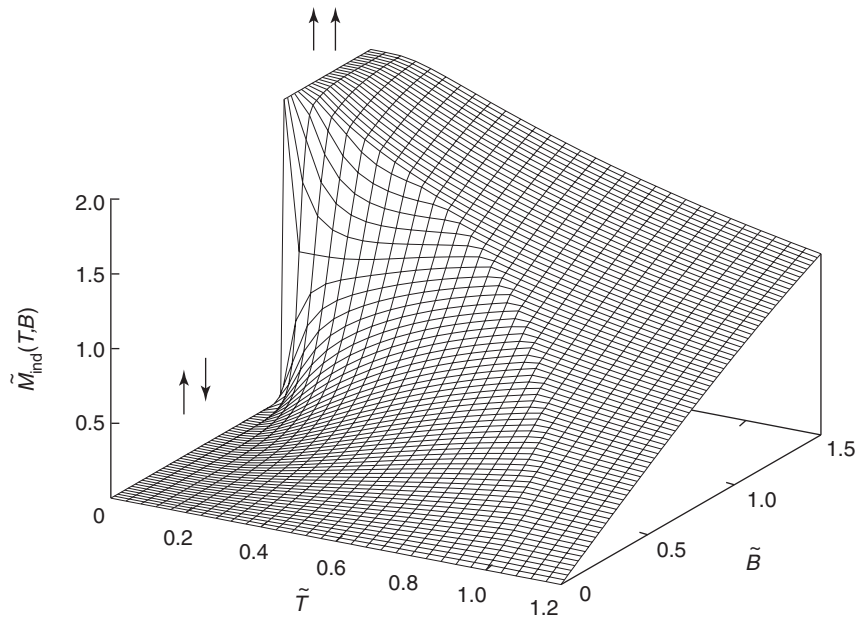


Figure 5 Temperature and field dependences of the induced moment M_{ind} for $S = 5/2$ and $\omega' = 0$ for the case of the strong anisotropic field ($B_{ani} > B_{ex}$). $\tilde{M}_{ind} = 2M_{ind}/Ng\mu_B S$, $\tilde{T} = T/T_N$, and $\tilde{B} = g\mu_B B/2z_{AB}|J_{AB}|S$.

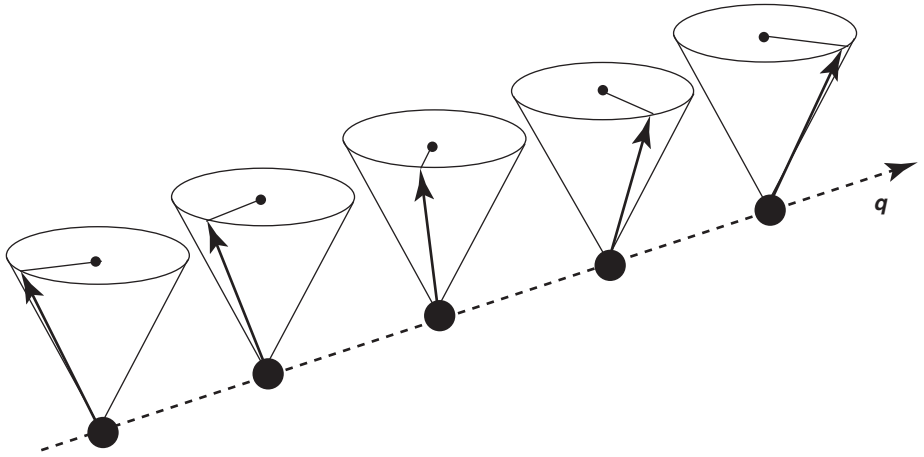


Figure 6 Schematic representation of the spin wave with wave-number vector $\boldsymbol{q}$.

and $M_z \simeq M_s$. One gets $\hbar\omega_q = g\mu_B D M_s q^2$ for small q. The coefficient of q^2 is called the exchange stiffness constant, which is observed by neutron scattering and spin wave resonance measurements. A schematic representation of $\boldsymbol{M}(\boldsymbol{r}, t)$ at a fixed time t is shown in **Figure 6**.

The wave of local magnetization, that is, the spin wave is quantized. By the analogy with phonons, the Hamiltonian for the quantized spin wave (magnon) is written by $\mathcal{H} = \sum_q (n_q + 1/2)\hbar\omega_q$, where n_q is the occupation quantum number of magnons with the wave-number vector $\boldsymbol{q}$. The temperature dependence of the magnetization is given by $M(T) = M_s - 2\mu_B \sum_q \langle n_q \rangle$, where M_s is the spontaneous magnetization and the thermal average $\langle n_q \rangle$ is a Bose distribution function $\{\exp(\hbar\omega_q/k_B T) - 1\}^{-1}$. One obtains $\Delta M(T) = M_s - M(T) \propto T^{3/2}$ at low temperature where $\hbar\omega_q$ is proportional to q^2. The theory of spin waves is also established in the itinerant-electron model of magnetism.

Itinerant-Electron Model

A quantum theory of ferromagnetism of conduction electrons in metals began with Bloch's paper in 1929. A free-electron gas with sufficiently low density was shown to be ferromagnetic. This means that

the carriers of the magnetic moment in metals also participate in electric conductions. In 1933, Stoner pointed out that nonintegral values of the observed magnetic moment per atom in 3*d* transition metals is explained if 3*d* electrons are itinerant, having a suitable density of states (DOS) and an exchange interaction parameter. Slater discussed ferromagnetism of Ni metal, using calculated results of the electronic structure and estimating the exchange couplings between electrons from spectroscopic data.

The subsequent development of itinerant-electron magnetism has proceeded mainly along two lines. One is the study on refinements of the electronic structure. The other is the study on the interaction between electrons. For both studies, electron correlations make the problem difficult. In the former, the electron correlations are taken into account in the potential of the electrons. A local spin-density functional formalism for the band calculation is established. Ferromagnetic, antiferromagnetic, and spin-density wave states are discussed for real materials by this formalism. In the latter, the interaction between electrons is treated as a many-body problem. To discuss the electron correlations, the Hubbard Hamiltonian is often made use of.

Band Theory of Ferromagnetism

The band theory (or Stoner theory) is widely used in the study of magnetism for transition metals, alloys, and compounds. In the original paper of Stoner, the DOS of electrons is assumed to be a parabolic one with respect to the energy. However, the DOS for actual transition metals is not so simple, but has a lot of peaks, as many energy levels overlap with one another. **Figure 7** denotes the calculated DOS curve for a b.c.c. Fe metal in the nonmagnetic state at the lattice constant 28.5 nm. E_F denotes the Fermi level. Electrons occupy energy levels up to E_F.

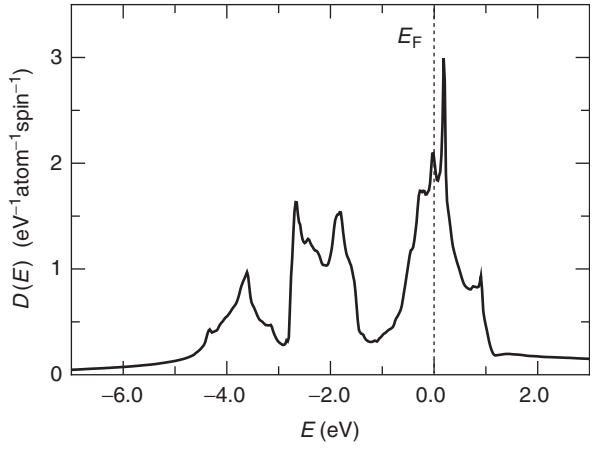


Figure 7 DOS curve for a b.c.c. Fe metal in the nonmagnetic state. E_F denotes the Fermi level. The value of D_0 is 2.05 eV^{-1} atom^{-1} spin^{-1}. The unit of energy 1 eV is 1.602×10^{-19} J.

The Pauli spin susceptibility is written in terms of the DOS curve, $D(E)$, as $\chi_0(T) = 2\mu_0\mu_B^2[D_0 - (\pi^2 k_B^2 T^2/6)\{D_0'^2/D_0 - D_0''\}]$ at low T, where D_0, D_0' and D_0'' are the values of $D(E)$ at E_F and the first and second derivatives of $D(E)$ with respect to E at E_F, respectively. The T-dependence comes from the low temperature expansion of the integrals connected with the Fermi distribution function. One takes into account the exchange interaction by the molecular field approximation. The induced magnetic moment, M, by the external field is given by $M = \chi_0(T)(B + IM)/\mu_0$, where I is the molecular field coefficient. Then, the susceptibility $\chi(T)$ $(= \mu_0 M/B)$ is obtained as $\chi(T) = \chi_0(T)/\{1 - I\chi_0(T)/\mu_0\}$. The factor $1/\{1 - I\chi_0(T)/\mu_0\}$ is the exchange-enhancement factor of the susceptibility. When the value of I is so large that the denominator in the enhancement factor becomes zero, the susceptibility $\chi(T)$ diverges. This is the magnetic instability in the paramagnetic state, which is called the Stoner condition for the appearance of ferromagnetism.

When $I\chi_0(0)/\mu_0 > 1$, the ferromagnetic state is stable at $T = 0$. In this case, the Curie temperature T_C is obtained by $I\chi_0(T_C)/\mu_0 = 1$. In the native Stoner theory, the value of I is an adjustable parameter to fit the Curie temperature. However, the value of I is nowadays obtained by the local spin-density functional formalism of the band calculation. The value of T_C estimated with the calculated value of I is much larger than the observed one, although reasonable values of the magnetic moment for 3*d* transition metals are obtained. This is because the effect of spin fluctuations at finite temperature is not taken into account in the theory.

Description of Ferromagnetism by the Landau Theory

By the fixed-spin-moment method developed recently, the magnetic energy $\Delta E(M)$ can be calculated numerically as a function of the magnetic moment M. The open circles in **Figure 8** denote the calculated results of $\Delta E(M)$ for a b.c.c. Fe metal at the lattice constant 28.5 nm. The equilibrium state is achieved at $M \sim 2.2\,\mu_B$ per Fe atom, being very close to the observed one. The calculated $\Delta E(M)$ can be fitted in the form of $\Delta E(M) = (a_0/2)M^2 + (b_0/4)M^4$, as shown by the dotted curve in the figure. This is just the Landau expansion of the magnetic energy. For some materials, the calculated $\Delta E(M)$ is expanded up to much higher order terms of M.

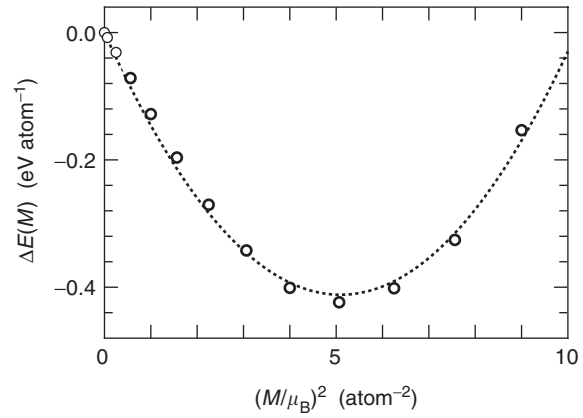


Figure 8 Magnetic energy $\Delta E(M)$ as a function of M^2 for a b.c.c. Fe metal. The unit of energy 1 eV is 1.602×10^{-19} J. The dotted curve denotes the form of $\Delta E(M) = (a_0/2)M^2 + (b_0/4)M^4$ with $a_0 = -5.59 \times 10^3$ T atom μ_B^{-1} and $b_0 = 1.10 \times 10^3$ T *atom*^{3} μ_B^{-3}.

The Landau coefficients a_0 and b_0 are also obtained by the Stoner theory. At $T = 0$, they are written by $a_0 = [1/2D_0 - I]/\mu_0\mu_B^2$, and $b_0 = [\{3D_0'^2/D_0^2 - D_0''/D_0\}/48D_0^3]/\mu_0\mu_B^4$. These expressions of a_0 and b_0 are derived by a rather crude approximation of the rigid band model, where DOS curves in the up and down spin bands are assumed not to change their shapes from those in the paramagnetic state. Nevertheless, those expressions of a_0 and b_0 are still relevant. When E_F lies near a peak of the DOS, $D_0' \sim 0$, and $D_0'' < 0$, the value of b_0 is positive at 0 K. On the other hand, when E_F lies near a minimum of the DOS, $D_0' \sim 0$, and $D_0'' > 0$, b_0 is negative. Even if D_0 is so small that the Stoner condition is not satisfied, the ferromagnetic state would be stabilized when the value of b_0 is negative and large. In the case of $a_0 > 0$ and $b_0 < 0$, a magnetic field-induced metamagnetic transition from the paramagnetic state to the ferromagnetic one may take place. This is the itinerant-electron metamagnetism. Moreover, the value of I can be estimated from the calculated values of a_0 and D_0. The evaluated value of I is 0.57 eV atom for a b.c.c. Fe metal at the lattice constant 28.5 nm.

The magnetic equation of state is given by $B = a_0M + b_0M^3$, as $B = \partial\Delta E(M)/\partial M$. The spontaneous magnetic moment M_s at $B = 0$ is given by $M_s = (|a_0|/|b_0|)^{1/2}$ if $a_0 < 0$ and $b_0 > 0$. It is noted that the magnetic moment in this theory increases with increasing B even at $T = 0$. The differential susceptibility χ_{hf} $(= \mu_0 dM/dB)$ is finite $(2\mu_0|a_0|)$ at $T = 0$. This is the high-field susceptibility in the ferromagnetic state. In the localized electron model, χ_{hf} vanishes at $T = 0$. The finite value of χ_{hf} at $T = 0$ is one of the characteristics of the itinerant-electron model.

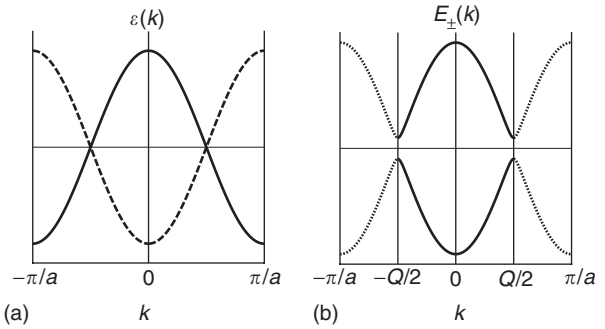


Figure 9 (a): One-electron energy $\varepsilon(k)$ in the nonmagnetic state, and (b): $E_\pm(k)$ in the antiferromagnetic state. A thin horizontal line denotes the Fermi energy.

Band Theory of Antiferromagnetism

An antiferromagnetic state in the itinerant-electron model can also be described by the electronic structure. For the sake of simplicity, consider a 1D lattice, where atoms are arranged along the x direction with a lattice constant a. The wave function $\phi_k(x)$ in the nonmagnetic state is given by a Bloch function in the periodic potential by the ions. The one-electron energy $\varepsilon(k)$ is expressed with the wave number k. The wave numbers are restricted to the first Brillouin zone $(-\pi/a < k \leqslant \pi/a)$. A schematic representation of $\varepsilon(k)$ is shown by the solid curve in **Figure 9a**.

When the antiferromagnetic moment is given by $M_Q(x) = M_0 \exp(iQx)$, where $Q = \pi/a$, the potential of electrons includes a term with the modulation of the wavelength $2a$. This modulated potential is treated by the perturbation theory of quantum mechanics. The wave function $\psi_k(x)$ in the antiferromagnetic state is given by a linear combination of the unperturbed wave functions $\phi_k(x)$ and $\phi_{k+Q}(x)$, as $\psi_k(x) = \alpha_k\phi_k(x) + \beta_k\phi_{k+Q}(x)$ with $|\alpha_k|^2 + |\beta_k|^2 = 1$. One gets two energy curves $E_\pm(k) = (1/2)[\varepsilon(k) + \varepsilon(k+Q) \pm \{(\varepsilon(k) - \varepsilon(k+Q))^2 + \Delta^2\}^{1/2}]$, where Δ is the off-diagonal matrix element of the perturbed potential. The broken curve in **Figure 9a** denotes $\varepsilon(k+Q)$. The two energy curves cross each other at $k = \pm Q/2$. A bandgap Δ appears at the crossing point, as shown in **Figure 9b**. The lower energy band $E_-(k)$ is that for electrons with up spin at the up-spin sites, or for electrons with down spin at the down-spin sites. The higher energy band $E_+(k)$ is that for electrons with opposite spin at the up and down spin sites. This is a simple mechanism for the appearance of the antiferromagnetic state in the band theory. It is called the split band model of antiferromagnetism.

In actual materials, the situation becomes more complicated. In **Figure 10**, the calculated Fermi surfaces of Cr metal in the nonmagnetic state are shown.

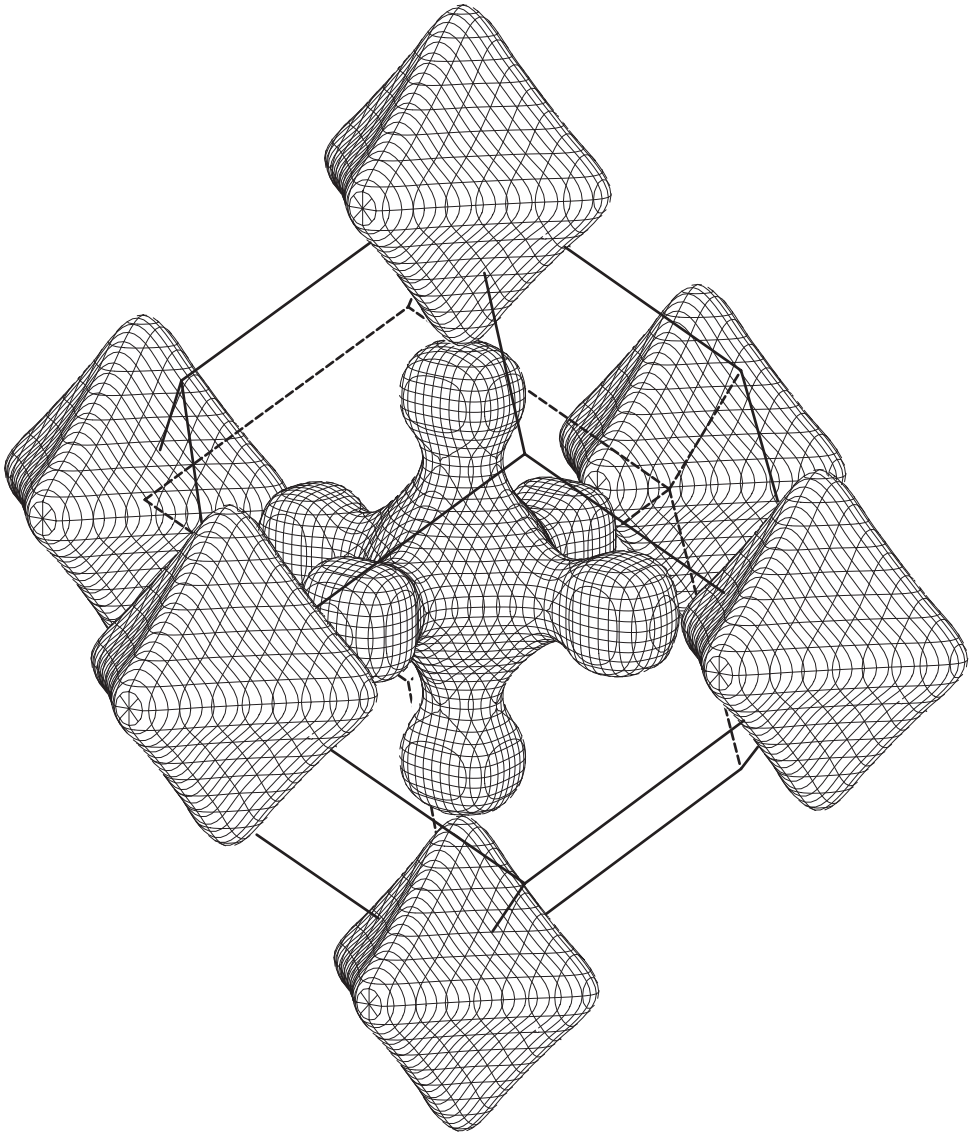


Figure 10 Main parts of the Fermi surfaces of Cr metal. The hole surface (octahedron-like) is located around H in the Brillouin zone. The surface centered at Γ (center of the Brillouin zone) is for electrons (K Nakada, unpublished work).

A Fermi surface denotes the equal energy surface in the reciprocal lattice space at the Fermi level. A hole surface has the octahedron-like shape centered at the H-point in the Brillouin zone and the electron surface is centered at the Γ-point (center of the Brillouin zone). Here, H and Γ are conventional symbols labeled to the symmetry points in the Brillouin zone. When the electron surface centered at Γ is shifted toward H, some portions of the electron and hole surfaces may overlap with one another. On these overlapping surfaces, the bandgap appears and the antiferromagnetic state is stabilized. Strictly speaking, the shift of the electron surface is not exactly the same as the length between the Γ- and H-points in this case. The antiferromagnetic wave-number vector $\boldsymbol{Q}$ is slightly different from that at the H-point. Then, the spin-density wave state is stabilized in the Cr metal. This is the nesting model of the Fermi surfaces.

Spin Fluctuations

As mentioned in the section, "Band theory of ferromagnetism," the value of T_C, estimated by the calculated value of I, becomes rather higher than the observed one. This is because the excitations of the collective modes of spins are not taken into account in the theory. The excitations of collective modes play an important role in the magnetic properties at finite temperature. In this subsection, it

is studied within the phenomenological Ginzburg–Landau theory.

The magnetic free energy ΔF_m is written by $\Delta F_m = (1/V)\int d\boldsymbol{r}\Delta f_m(\boldsymbol{r})$, where the free energy density $\Delta f_m(\boldsymbol{r})$ is given by $\Delta f_m(\boldsymbol{r}) = (a_0/2)|\boldsymbol{m}(\boldsymbol{r})|^2 + (b_0/4)|\boldsymbol{m}(\boldsymbol{r})|^4 + (D/2)|\nabla\cdot\boldsymbol{m}(\boldsymbol{r})|^2$. D is the Landau–Lifshitz constant given in the section "Spin waves." ΔF_m is rewritten by using the Fourier transformation for the ith component of the magnetization density, $m_i(\boldsymbol{r}) = M\delta_{i,z} + (1/V)\sum'_q m_i(\boldsymbol{q})\exp(i\boldsymbol{q}\cdot\boldsymbol{r})$, where M is the bulk moment in the z direction. The prime on the summation denotes the sum over q except $q = 0$. The expression of ΔF_m contains terms up to the fourth power of $m_i(\boldsymbol{q})$. Among them, only the terms with even power of $m_i(\boldsymbol{q})$ are considered. In this case, ΔF_m is given by a functional of only M and $|m_i(\boldsymbol{q})|^2$.

The equation of state, defined by $B = \langle \partial\Delta F_m/\partial M\rangle$, is written as $B = a(T)M + b(T)M^3$, where $\langle\cdots\rangle$ denotes a thermal average. The Landau coefficients $a(T)$ and $b(T)$ are given by $a(T) = a_0 + b_0\{3\langle(\delta m_\parallel)^2\rangle + 2\langle(\delta m_\perp)^2\rangle\}$ and $b(T) = b_0$. $\langle(\delta m_\parallel)^2\rangle$ and $\langle(\delta m_\perp)^2\rangle$ are mean square amplitudes of the longitudinal and transverse magnetization densities given by $\langle(\delta m_i)^2\rangle = (1/V)\sum_q \langle|m_i(\boldsymbol{q})|^2\rangle$. The T-dependence of the Landau coefficients comes from that of the mean square amplitudes of spin fluctuations $\xi(T)^2 = \langle(\delta m_\parallel)^2\rangle + 2\langle(\delta m_\perp)^2\rangle$. Isotropic spin fluctuations are assumed, that is, $\langle(\delta m_\parallel)^2\rangle = \langle(\delta m_\perp)^2\rangle$. In this case, the Landau coefficients are given by $a(T) = a_0 + (5/3)b_0\xi(T)^2$ and $b(T) = b_0$. The present T-dependence of the Landau coefficients is very different from that in the Stoner theory where it comes only from the Fermi distribution function. The degeneracy temperature for electrons estimated from the Fermi level is very high, and then the T-dependence in the Stoner theory is very weak in the conventional ferromagnets. Thus, the spin fluctuations are found to play a dominant role at finite T.

In the case of $a_0 < 0$ and $b_0 > 0$, the Curie temperature is given by $\xi(T_C)^2 = 3|a_0|/5b_0$ as the inverse of the susceptibility should be zero at T_C, that is, $a(T_C) = 0$. By classical thermodynamics, the mean square amplitude of thermal fluctuations is known to increase monotonically as T increases. The susceptibility (at $\xi(T)^2 = pT$) above T_C is written in the form of the Curie–Weiss law as $\chi(T) = \mu_0 C/(T - T_C)$, where $C = 3/5b_0p$, $T_C = 3|a_0|/5b_0p$, and p a positive constant. Such a Curie–Weiss law cannot be obtained in the Stoner model. By analogy with the localized-electron model, the square of the effective magnetic moment M_p^2 in the paramagnetic state, estimated from the susceptibility, is proportional to the Curie constant C. On the other hand, the square of the spontaneous magnetization is given by $M_s^2 = |a_0|/b_0$ at $T = 0$. Then, the value of $(M_p/M_s)^2$ is shown to be proportional to $|a_0|^{-1}$, and subsequently to T_C^{-1}. This type of behavior of M_p/M_s is observed for $3d$ transition metals, alloys, and compounds (Rhodes–Wohlfarth plot). Spin fluctuations in itinerant-electron antiferromagnets have also been studied.

See also: Density-Functional Theory; Fermi Surface Measurements; Ferromagnetism; Magnetic Order; Strongly Correlated Electron Systems.

PACS: 75.10.Jm; 75.10.Lp; 75.20.En; 75.30.−m; 75.30.Cr; 75.30.Ds; 75.30.Gw; 75.30.Kz; 75.50.Cc; 75.50.Ee; 75.60.Ej

Further Reading

Arrott A (1966) Antiferromagnetism in metals and alloys. In: Rado GT and Shul H (eds.) *Magnetism*, vol. II B, pp. 295–416. New York: Academic Press.

Bozorth RM (1951) *Ferromagnetism*. New York: Van Nostrand.

Chikazumi S (1964) *Physics of Magnetism*. New York: Wiley.

Gautier F (1982) Itinerant magnetism. In: Cyrot M (ed.) *Magnetism of Metals and Alloys*. Amsterdam: North-Holland.

Herring C (1966) Exchange interactions among itinerant electrons. In: Rado GT and Shul H (eds.) *Magnetism*, vol. IV, pp. 1–396. New York: Academic Press.

Kittel C (1962) Magnons. In: Dewitt B, Dreyfus B, and de Gennes P (eds.) *Low Temperature Physics*, pp. 441–478. New York: Gordon and Breach.

Kübler J (2000) *Theory of Itinerant Electron Magnetism*. Oxford: Clarendon Press.

Moriya T (1985) *Spin Fluctuations in Itinerant Electron Magnetism*. Berlin: Springer.

Vonsovskii SV (1974) *Magnetism*, vols. 1 and 2. New York: Wiley.

Wohlfarth EP (ed.) (1980) Iron, cobalt and nickel. *Ferromagnetic Materials*, vol. 1, pp. 1–70. Amsterdam: North-Holland.

Low-Energy Electron Microscopy

E Bauer, Arizona State University, Tempe, AZ, USA

Introduction

Modern surface science, thin film science, and in particular, nanoscience needs characterization methods with high lateral resolution. The most prominent of these methods are scanning probe methods such as scanning tunneling microscopy and atomic force microscopy that probe the topmost layer of a surface with atomic resolution. In many cases, however, a somewhat larger and somewhat tunable probing depth is required. This is provided by elastic reflection of slow electrons that has been used for many years in low-energy electron diffraction (LEED) in the structure analysis of the top several layers of crystals. Similar to the diffracted fast electrons in conventional transmission electron microscopy (TEM), diffracted low-energy electrons can also be used for imaging in low-energy electron microscopy (LEEM) though not with the high lateral resolution that scanning probe microscopy and transmission electron microscopy provide. Essential features of LEEM are not so much the tunable sampling depth but the well-defined wavelength, the high sensitivity to surface topography, and the spin-dependent interaction of slow electrons with magnetic materials that allows imaging of the magnetization distribution on surfaces and in thin films in spin-polarized LEEM (SPLEEM).

Understanding the image formation in LEEM and SPLEEM requires an understanding of the electron–specimen interaction and of the electron optics needed to form an image with the reflected electrons. The discussion begins with the electron–specimen interaction, then the electron optics, and finally illustrates the application of LEEM and SPLEEM with a number of examples.

Electron–Specimen Interaction

Electrons are scattered in matter elastically and inelastically. Elastic scattering in the backward direction determines the number of electrons available for imaging. Contrary to the fast electrons used in TEM, which are nearly completely scattered in the forward direction, low-energy electrons are also strongly scattered in the backward direction. This is illustrated in **Figure 1** for two surfaces of a W single crystal. In single crystals, the backscattered electrons are focused by diffraction into narrow beams whose

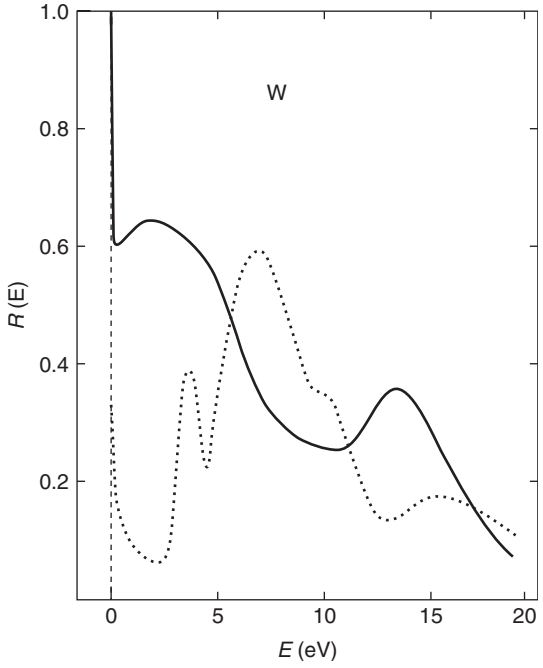


Figure 1 Specular reflectivity of two W surfaces for slow electrons. Solid line (110), dotted line (100) surface. (Reproduced with permission from Bauer E (1994) Low-energy electron microscopy. *Reports on Progress in Physics* 57: 895–938; © IOP Publishing Limited.)

intensity is determined by diffraction, that is, by the periodicity of the crystal, and by the scattering potential of its constituents. The different periodicity normal to the two surfaces shown in **Figure 1** causes the strong difference in reflected intensity and is one of the major contrast mechanisms in LEEM, called diffraction contrast. **Figure 1** shows the intensity of the specularly reflected beam ("(00)-beam"). Other (nonspecularly reflected) diffracted beams may also be used for imaging which gives contrast caused by differences of the lateral periodicity of the surface.

For imaging with the (00)-beam, the specimen does not have to be a single crystal but may consist of many crystals that have the same orientation with respect to the incident beam but otherwise random orientation. However, when the crystals are completely randomly oriented or when the specimen is amorphous, the backscattered electrons are distributed over a wide angular range. In this case, it is necessary to limit the angular acceptance of the instrument to a narrow cone around its optical axis because of the aberrations of the electron-optical system used for imaging that will be discussed below. This reduces the number of electrons available for imaging considerably which requires a considerable increase of the image acquisition time. For this

reason, LEEM and SPLEEM are mainly used for the study of specimens with strong preferred crystal orientation which allows image acquisition at video rates and thus allows the study of the dynamics of processes.

The wave nature of the electron, which is a basic requirement for diffraction, can also be used in other ways to produce image contrast. Optical path differences of the electrons reflected from the terraces adjoining a surface step produce phase contrast that can be made visible by lens aberrations and/or defocusing. This so-called step contrast is used, for example, in **Figure 2** to make the monoatomic steps on a Mo(110) surface clearly visible.

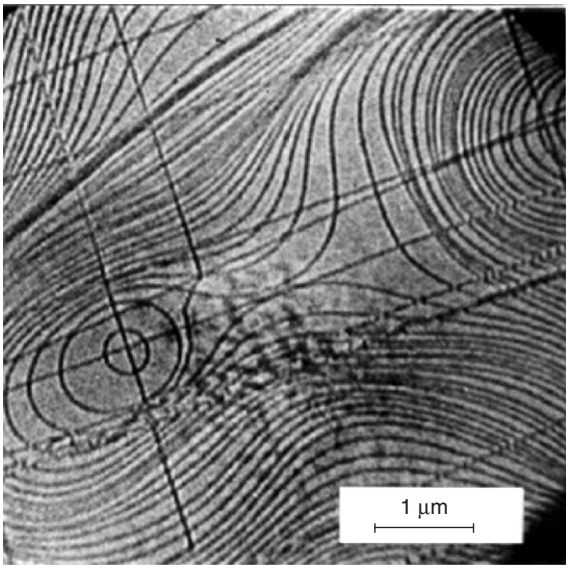


Figure 2 Monatomic steps on a Mo(110) surface. Sublimation and glide steps are visible (electron energy – 14 eV). (Reproduced from Telieps W and Bauer E (1985) An analytical reflection and emission UHV electron microscope. *Ultramicroscopy* 17: 57, with permission from Elsevier.)

Another contrast mechanism that rests on the wave nature of the electron is the so-called quantum size contrast. In a thin film with parallel boundaries and thickness t, standing electron waves can form whenever $n\lambda/2 + \varphi = t$, where λ is the wavelength and φ represents the phase shifts at the boundaries. This causes reflectivity differences between regions with different thicknesses as illustrated in **Figure 3**, similar to the interference phenomena in thin films in light optics. Finally, one contrast mechanism, which is due to the electrical field at the surface in an LEEM instrument, has to be mentioned: topographic contrast caused by the field distortion around surface asperities and pits. It is most evident when the electrons are reflected immediately in front of the surface, that is, when the instrument is operated in the mirror mode (mirror electron microscopy (MEM)), but is also noticeable in LEEM under proper focusing conditions. MEM does not require crystalline surfaces and, thus, is widely applicable, also to flat surfaces with locally varying surface potentials.

Before turning to the influence of inelastic scattering on LEEM, one important difference between the scattering of fast and slow electrons has to be mentioned. While the scattering of fast electrons by atoms can be described to a good approximation by the first Born approximation which predicts increasing scattering strength with increasing nuclear charge Z, this is not true any longer for the backscattering of slow electrons. The backscattering cross section does not vary monotonically with increasing Z or energy E. Atoms with small Z may backscatter in certain energy ranges much more than atoms with large Z. For example, ~ 50 eV Al and Si backscatter much stronger than W. Therefore, the analog to the Z contrast in TEM does not exist in LEEM.

Inelastic scattering of the electrons determines the sampling depth of LEEM unless elastic backscattering

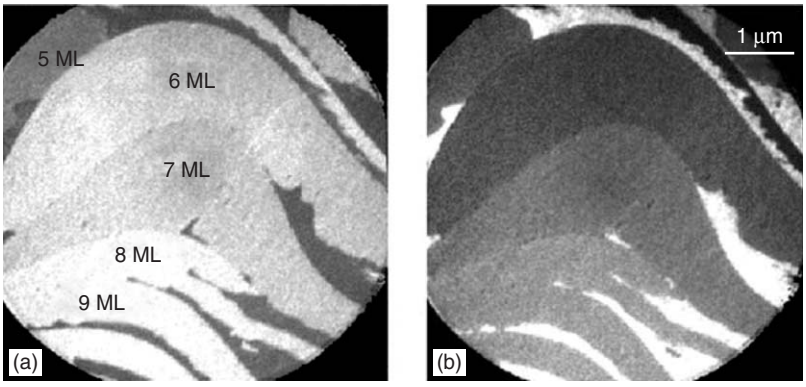


Figure 3 Quantum size contrast in an Fe layer on a W(110) surface. The film thickness varies from terrace to terrace as indicated in the image taken with 11.0 eV electrons (a). A comparison with image (b), taken with 16.4 eV electrons, illustrates the dependence of the contrast upon the wavelength $\lambda = \sqrt{1.5/E}$ (λ in nm, E in eV). (Reproduced with permission from Zdyb R and Bauer E (2002) Spin dependent quantum size effects in the electron reflectivity of ultrathin ferromagnetic crystals. *Surface Review Letters* 9: 1485.)

is very strong and inelastic scattering weak. Inelastic scattering involving small energy losses, such as excitation of lattice vibrations or spin waves in magnetic specimens, is not eliminated from the image by the energy-dispersive elements of present LEEM/SPLEEM instruments and thus does not limit the sampling depth. The sampling depth is limited by energy losses of say more than 0.5–1 eV, to give a rough limit. These electrons do not contribute to the image formation. In electron spectroscopy, the sampling depth is normally defined by the inelastic mean free path Λ that takes into account all losses ranging from single electron excitation to collective electron (plasmon) excitation to inner shell ionization. Its energy dependence is usually described by a universal curve which is a good approximation for many materials at energies above ~ 50 eV but breaks down in the energy range mainly used in LEEM ($\leqslant 20$ eV). Here, the energy loss mechanisms are strongly material-dependent. For example, in noble metals such as Au, Λ may be as large as several nanometers whereas in 3*d* metals such as Fe, it is only a few tenths of a nanometer. Surface and depth sensitivities are thus strongly material-dependent.

Instrumentation

Because of the high surface sensitivity of low-energy electrons, the instruments using them require ultrahigh vacuum (UHV) and surface cleaning facilities. For electron-optical reasons, the electrons are decelerated from the high energy as they approach the specimen. This is done in a so-called cathode lens whose light-optical analog is the immersion lens. In order to be able to observe the image of the surface, the illumination beam has to be separated from the imaging beam by a beam separator. This leads to a system configuration as schematically shown in **Figure 4a**. Lens 1 produces a demagnified image of the electron source (cathode C) – or more precisely of the "crossover" in front of C – which is imaged by lens 2 into the back focal plane of the objective lens O. In the objective lens – which also includes the specimen S – the electrons are decelerated from the high initial energy, typically 5–20 keV, to the desired low energy which is determined by the potential difference between cathode and specimen. After reflection from the surface, the electrons are accelerated again to high energy. The objective lens produces the diffraction pattern of the specimen in its back focal plane and a first intermediate image in the middle of the beam separator B. The subsequent lenses produce further intermediate images and lens 4 allows switching between imaging the specimen and its diffraction pattern, which are imaged by lens 5 into the image plane I. The brightness of the image is enhanced by a pair of channel plate multipliers so that it is high enough on the final fluorescent screen to be recorded with a CCD camera with short image acquisition times. The electron lenses are, in general, magnetic lenses but electrostatic lenses are used too. The magnetic beam separator B is designed so as to minimize the unidirectional focusing of magnetic fields. Not shown in **Figure 4a** are the deflection elements needed for beam alignment, the stigmators for correction of the astigmatism of the electron-optical components, and the apertures needed for confining the angular acceptance of the system and for selecting the specimen area contributing to the LEED pattern. The first aperture, the so-called contrast aperture, is placed in the intermediate image of the diffraction pattern produced by lens 3, the second one either in the first intermediate image of the specimen in the beam separator or into the symmetric position in the illumination beam. Various beam separators are used with deflection angles ranging from 16° to 90°.

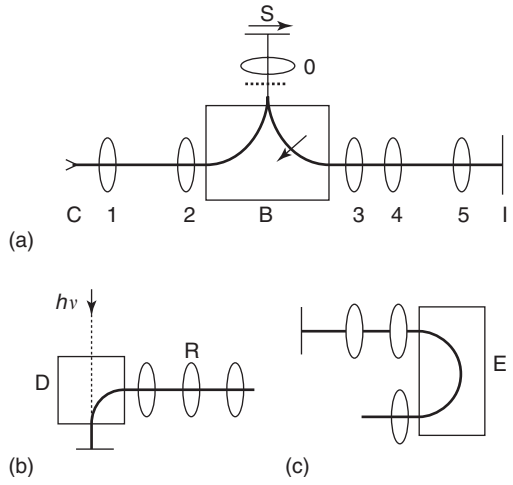


Figure 4 Schematic of an LEEM instrument. (a) Basic instrument setup, (b) illumination system of a SPLEEM instrument, and (c) imaging system for a spectroscopic instrument. The locations of the first image and diffraction planes are indicated by an arrow and dotted line, respectively.

The heart of the microscope is the objective (cathode) lens because it determines the resolution and transmission of the microscope. Cathode lenses have much larger chromatic and spherical aberrations than the lenses used in TEM. Therefore, the acceptance angle α has to be strongly limited by the "contrast aperture" in the diffraction pattern. While in TEM, the contrast aperture may accept many diffracted beams which allows one to reach resolutions in the 0.1 nm range, in LEEM only one diffracted beam, usually the (00)-beam, and its immediate environment can be used for optimum resolution.

The best resolution achieved to date is about 5 nm, using a field emission cathode that has a narrow energy distribution and, thus, minimizes the chromatic aberration, which is dominant at low energies. Most instruments use magnetic cathode lenses because of their somewhat lower aberrations and higher transmission. Resolution improvement requires aberration correction, which is possible with an electron mirror, by monochromatizing the illumination beam and energy-filtering the imaging beam. Calculations for mirror correctors predict resolutions of $\sim$1 nm and somewhat less which still have to be demonstrated with instruments presently in construction. The present routine resolutions of 10–20 nm are largely determined by the stability of the specimen and by local charging, not by the electron optics.

Extensions of the basic LEEM instrument shown schematically in **Figure 4a** are the SPLEEM instrument and the LEEM instrument equipped with an imaging energy filter, called spectroscopic photoemission and low-energy electron microscope (SPELEEM) because its main application is in spectroscopic photoemission microscopy.

SPLEEM requires a different illumination system which is schematically shown in **Figure 4b**. The electron source is now not a field emitter or a LaB_6 cathode as used in LEEM but a GaAs(100) single crystal surface which has been exposed to Cs and an electronegative gas, usually O_2, so that the lowest point of the conduction band of GaAs is slightly above the vacuum level. This surface is illuminated with circular polarized light from a laser with an energy hv corresponding to the direct bandgap of GaAs so that electrons excited from the top of the valence band into the bottom of the conduction band can be emitted. Due to optical selection rules, up to 50% of the photoemitted electrons are spin-polarized with their spin perpendicular to the surface. Changing the helicity of the light from right to left circular polarization inverts the spin direction, characterized by the polarization vector $\boldsymbol{P}$. When a magnetic surface with local magnetization $\boldsymbol{M}$ is illuminated with these electrons, then the reflected intensity not only depends upon the topography and crystal structure, but also upon $\boldsymbol{P} \cdot \boldsymbol{M}$. The difference image between two images taken with opposite directions of $\boldsymbol{P}$ eliminates all nonmagnetic contrast and shows only the distribution of $\boldsymbol{M}$.

In order to determine the direction of $\boldsymbol{M}$ completely, $\boldsymbol{P}$ must not only be inverted but adjustable in any direction. This is achieved with the 90° electrostatic/magnetic deflector D and the spin rotator lens R. Pure electrostatic deflection leaves $\boldsymbol{P}$ unchanged; in pure magnetic deflection, $\boldsymbol{P}$ rotates along with the direction of the beam, and mixed deflection allows one to cover all $\boldsymbol{P}$ directions between these two directions, that is within the plane of the drawing of **Figure 4b**. The rotator lens allows $\boldsymbol{P}$ rotation around the beam direction, that is out of this plane so that all directions are accessible. Usually the direction normal to the specimen surface and two preferred directions in it are selected. **Figure 5** illustrates the application of these possibilities.

In the SPLEEM, the imaging column is modified by inserting an energy filter and adding additional lenses at its output. In the example shown in **Figure 4c**, the energy filter is a 180° electrostatic analyzer in which the electrons are retarded to a fraction of the input energy and reaccelerated at its output to the input energy again. Energy selection is made by a slit in the dispersive plane at the exit of the energy filter. By choosing the excitation of the pre-filter lenses 4 and 5 and of a lens in the filter, either the energy-filtered image or the diffraction pattern can be obtained with the lenses behind the filter. These can also be used to image the dispersive plane and thus to do electron spectroscopy from small specimen regions selected by the field-limiting aperture. An energy resolution of $\sim$0.3 eV can be obtained without much loss of intensity in LEEM and LEED. The advantage of the energy filtering in LEEM and LEED is the elimination of secondary electrons, which can cause a strong background in specimens with large electronic band gaps. The main application of the energy filter is however in X-ray photon-excited secondary electron emission microscopy, usually called XPEEM, in which the filter allows selection of a small energy window in the wide secondary electron energy distribution and thus improves the resolution. A second application is in PEEM with photoelectrons from inner shells characteristic for the specimen composition. **Figure 6** shows the presently most widespread SPLEEM.

Applications

Clean Surfaces

Because of its importance in microelectronics, Si is the material that has been studied most extensively. The demonstration in 1985 that the "reconstruction" of the Si(111) surface with bulk periodicity into one with sevenfold periodicity at $\sim$1100 K, the Si(111)-$(1 \times 1) \rightarrow$Si(111)-(7×7) transition, is a first-order phase transition which demonstrated the power of LEEM. It also clearly showed the limitations of laterally averaging studies because surfaces with the same LEED pattern produced completely different LEEM images as shown in **Figure 7**.

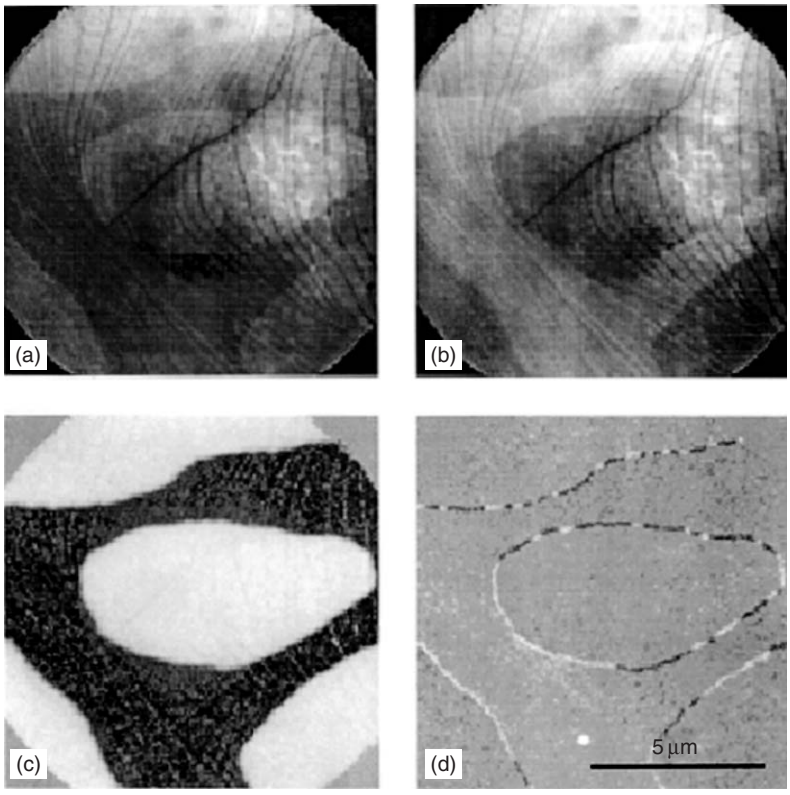


Figure 5 Imaging of the magnetization distribution in a thin Co layer on a W(110) surface with SPLEEM. Images (a) and (b) are taken with opposite direction of the polarization $\boldsymbol{P}$ parallel to $\pm\boldsymbol{M}$. Their difference image (c) shows the magnetic domains while the difference image taken with $\boldsymbol{P} \perp \pm \boldsymbol{M}$ (d) shows the domain walls (electron energy – 2 eV). (Reproduced from Bauer E *et al.* (1996) LEEM studies of the microstructure and magnetic domain structure of ultrathin films. *Journal of Magnetism and Magnetic Materials* 156: 1, with permission from Elsevier.)

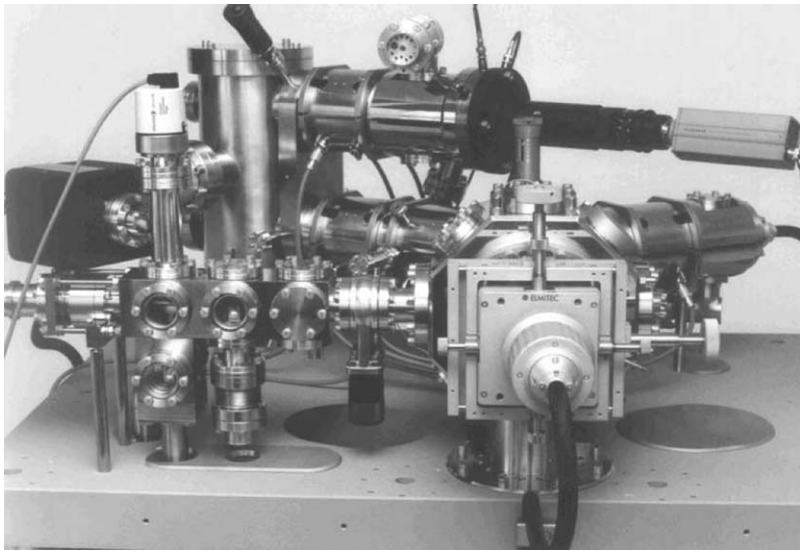


Figure 6 A SPELEEM instrument. The 60° beam separator is hidden by the specimen chamber in the center front. The illumination system is to the right, and the imaging system with the vertically mounted energy filter is on the left-hand side. In the front left side is a small preparation chamber and air lock. The specimen chamber has eight ports pointing at the specimen for mounting light sources for PEEM, evaporators and other accessories for *in situ* experiments. (Courtesy of ELMITEC GmbH.)

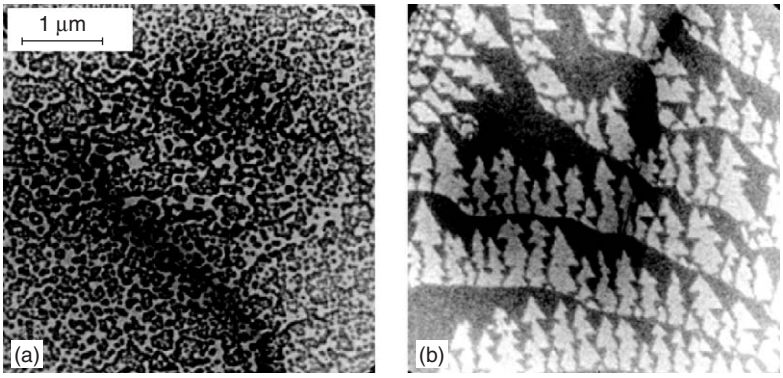


Figure 7 Coexistence of (7×7) regions (bright) and (1×1) regions (dark) on a quenched Si(111) surface. The surface in (a) was annealed for a long time at 1150 K, that in (b) was quenched directly from 1450 K (electron energy – 11 eV). (Reproduced with permission from Bauer E (1990) Low-energy electron microscopy. In: *Chemistry and Physics of Solid Surfaces VIII*, p. 267. Berlin: Springer; © Springer.)

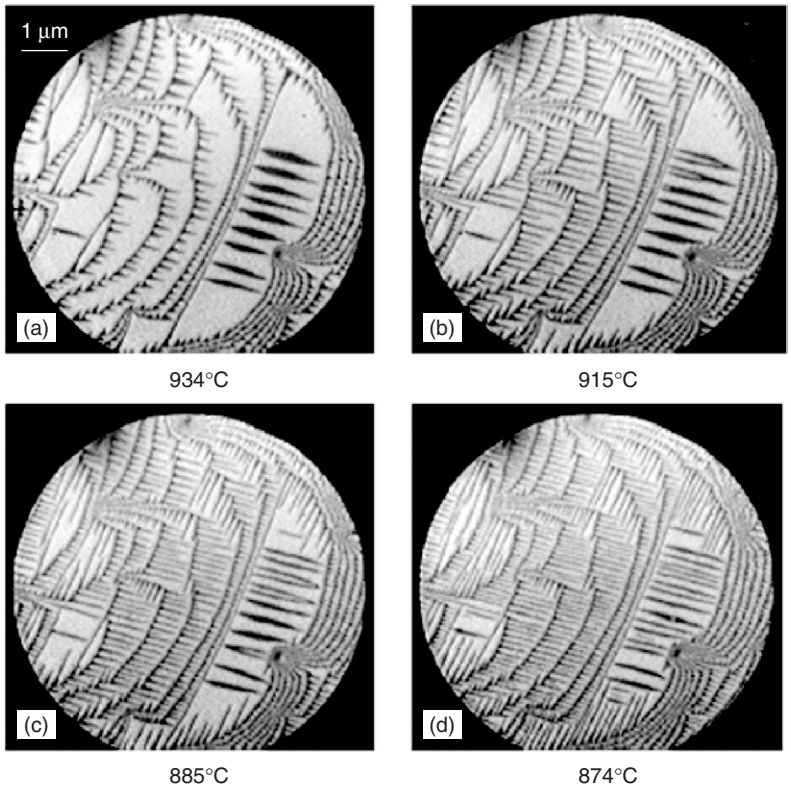


Figure 8 Step proliferation on an Si(100) surface induced by segregation of B from B-doped Si with decreasing temperature. Dark and bright regions are terraces with (2×1) and (1×2) structures that differ in height by one atomic layer (electron energy – 4.2 eV). (Reproduced with permission from Bauer E, Furukawa Y, Mori Y, and Kataoka T (eds.) (1999) Multi-method surface analysis in the 10 nm range. In: *Precision Science and Technology for Perfect Surfaces*, p. 753. Tokyo: The Japan Society for Precision Engineering.)

Extensive subsequent work on Si(111) and Si(100) surfaces has not only brought a detailed understanding of these surfaces but has contributed immensely to the understanding of processes on clean surfaces such as surface diffusion, adatom density, two-dimensional nucleation, step dynamics, and stability. An example is the step proliferation by B segregation of highly B-doped Si(100) upon cooling, illustrated in **Figure 8**, which is due to a strong increase of the anisotropy of the step energy and of the surface stress. Exchange processes between the volume and the surface also occur in clean crystals as seen, for example, in changes of the step dynamics and reconstruction on the $TiO_2(110)$ surface. Such

processes play an important role in heterogeneous catalysis, in particular, in the so-called three-way catalysts used in automotive exhaust remediation. Other frequently studied surfaces are the W(110) and Mo(110) surfaces that have become popular in the study of the growth of thin films because they can be reused nearly indefinitely as they have a high melting point and low vapor pressure. An example is shown in **Figure 2**.

Chemisorption and Chemical Reactions

The diffraction contrast in LEEM allows detailed studies of chemical processes on surfaces whenever they are connected with specific diffraction patterns. In general, chemisorption modifies the (00) intensity sufficiently so that regions covered with a specific chemisorbed species produce clear contrast. Frequently, chemisorbed phases form domains with different azimuthal orientations. These can be identified by imaging with beams in the LEED pattern (**Figure 9a**) that are characteristic for these domains as illustrated in **Figures 9c** and **9d**. The bright region in **Figure 9b** produces the hexagonal pattern in **Figure 9a**.

Another application of imaging with diffracted beams is the study of the dynamics of chemical surface reactions in which different regions of the surface have different composition and different diffraction patterns. This allows the study of reaction front propagation in oscillatory reactions as found, for example, in the reaction of NO and H_2 on Rh(110). **Figure 10** shows four images of a spiral wave taken about 1 min apart with diffracted beams from the four phases that occur in this reaction, ranging from a pure N phase (a) to a pure O phase (c). Two-dimensional alloying, domain formation, and domain dynamics in systems such as Pb on Cu(111) are other surface processes that have been elucidated by LEEM.

Thin Film Growth

LEEM is a valuable method for the study of the growth of thin films, in particular, of epitaxial growth. Si again has been one of the materials studied most extensively, ranging from homoepitaxial growth on Si(100) via the growth of Ge and Ge–Si alloy layers on Si(100) to the growth of pentacene films on Si(100). These studies have made important contributions to the development of the SiGe technology for wireless telecommunication and to the optimization of organic thin film transistors. Another example

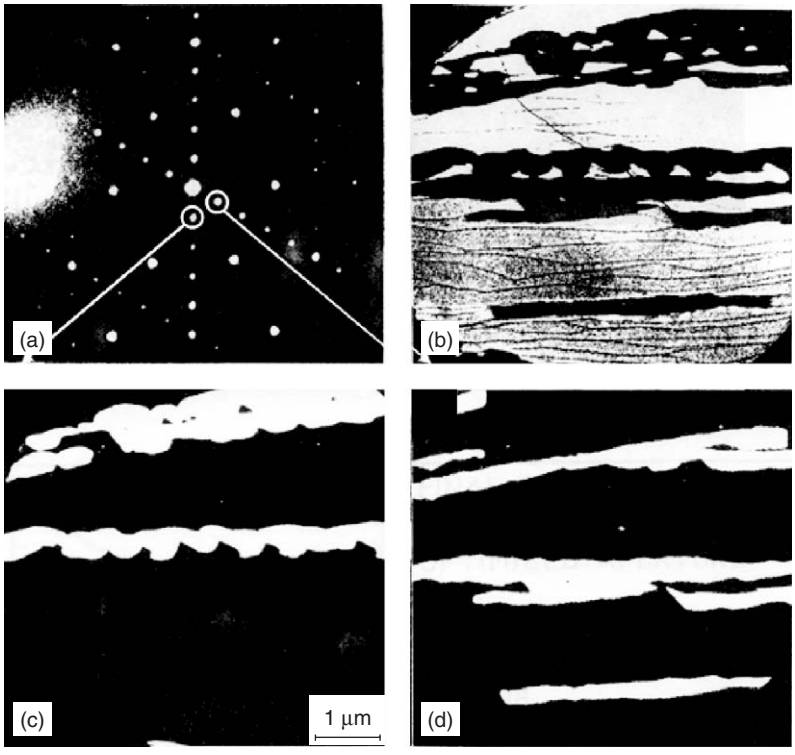


Figure 9 LEED pattern (a), taken at 30 eV, brightfield (b) and darkfield (c,d) LEEM images of an Si(111) surface covered with an Au submonolayer. Image (b) was taken with the (00) beam, images (c,d) with the diffracted beams marked by circles (electron energy – 6 eV). (Reproduced from Swiech W *et al.* (1991) A low-energy electron microscopy study of the system Si(111)-Au. *Surface Science* 253: 283, with permission from Elsevier.)

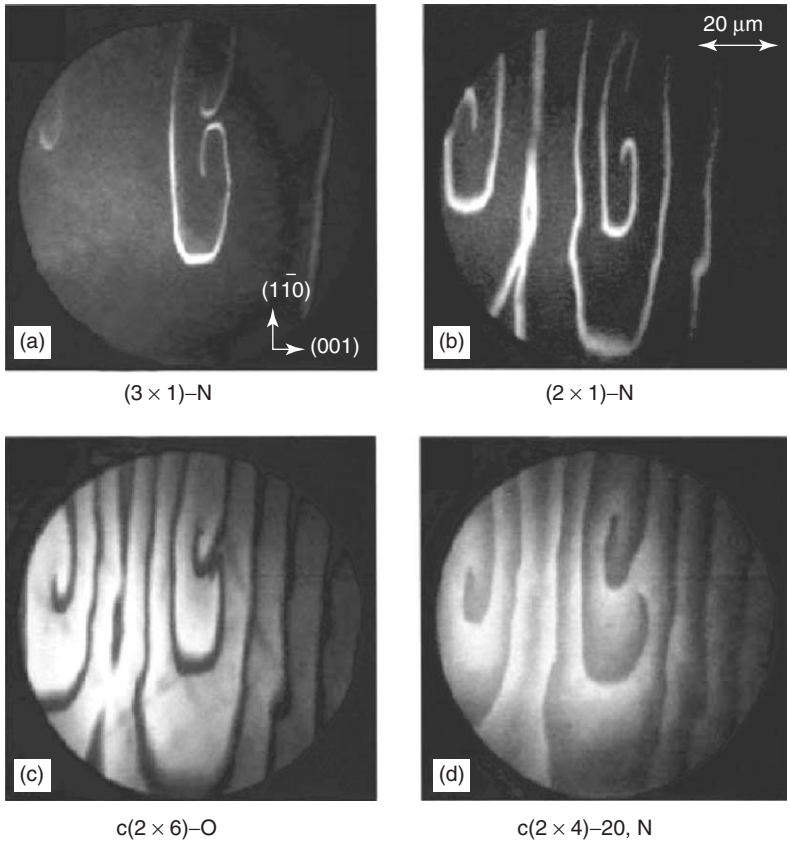


Figure 10 LEEM images of the oscillatory reaction of adsorbed NO and H_2 on an Rh(110) surface taken with diffracted beams from different ordered phases that form locally during reaction. (Reproduced from Schmidt Th *et al.* (2000) *In situ* imaging of structural changes in a chemical wave with low-energy electron microscopy: The system Rh(110)/NO + H_2. *Chemical Physics Letters* 318: 549, with permission from Elsevier.)

is the reactive growth of Si nitride on Si(111) by reaction with NH_3 whose initial stages are illustrated in **Figure 11**.

Heteroepitaxial films develop considerable roughness with increasing thickness in order to relieve misfit strain. Methods to counteract this tendency have been developed in LEEM studies, by either using so-called surfactants that tend to accumulate on top of the film or interfactants that remain at the interface between film and substrate. These methods are of increasing importance as the dimensions of many electronic and magnetic devices shrink into the nanometer range. The second method is illustrated in **Figure 12** by the influence of Au as an interfactant on the growth of Pb on Si(111). On the clean surface, the nucleation rate of Pb is small so that isolated three-dimensional islands are formed that grow individually, monolayer by monolayer, while on the Au-covered surface the two-dimensional nucleation rate is high so that, initially, a continuous monolayer is formed followed by further monolayer growth. Another result of LEEM studies of the growth of epitaxial films is that a quasiliquid Ga double layer is needed to grow smooth GaN film with the polar basal plane parallel to the substrate. Although the layer is liquid at the growth temperature, single and double layer differ sufficiently in the backscattered intensity to allow distinction between them. This result is important for the growth of GaN-based blue UV diode lasers and light emitting diodes.

Metal layer growth on W(110) and Mo(110) surfaces has been studied extensively, mainly in connection with the study of the electronic and magnetic properties of single crystal films. Two examples are shown in **Figures 3** and **5**, more can be found in the general references and below. Many metals grow below $\sim 500\,K$, quasi-monolayer by monolayer, but there are also many which form three-dimensional crystals as illustrated in **Figure 13** for the growth of Co on the W(100) surface. After initial two-dimensional pseudomorphic growth, that is with W lateral periodicity, long ribbon-shaped three-dimensional crystals develop due to the strong anisotropic misfit between the film and the substrate.

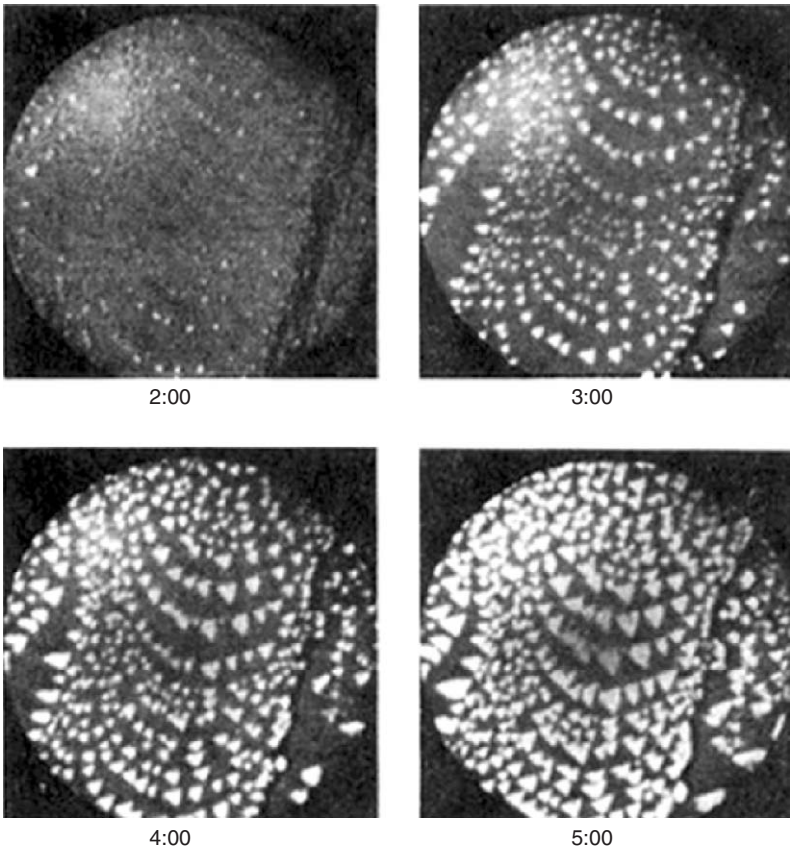


Figure 11 LEEM video frames taken during the initial stages of the reaction of an Si(111) surface with 2×10^{-7} torr NH_3 at 1180 K (electron energy – 45.4 eV). (Reproduced with permission from Bauer E *et al.* (1995) Reactive crystal growth in two dimensions: silicon nitride on Si(111). *Physical Review B* 51: 17891; © American Physical Society.)

Magnetic Properties

Because of the very short inelastic mean free path of electrons, in particular of the minority spin electrons, in ferromagnetic materials which is mainly due to excitations into the unoccupied minority spin states, SPLEEM is only sampling the topmost few layers. Together with their well-defined wavelength, this makes them uniquely suited for the study of very thin films. For example, the spin-dependence of the quantum size effect illustrated in **Figure 3** allows the determination of the spin-resolved electronic band structure above the vacuum level. For this, it is necessary to select, in the SPLEEM image, regions with well-defined thickness and measure the reflected intensity as a function of thickness and energy. The energy dependence in a six-monolayer thick film is illustrated in **Figure 14** for $\boldsymbol{P}$ parallel to the spins $\uparrow$ and $\downarrow$ of the majority and minority electrons, respectively.

The shift of the $\downarrow$ curve, caused by the exchange splitting, and its stronger damping, caused by the shorter inelastic mean free path are clearly seen. A second example is the study of the so-called spin reorientation transition in which the magnetization rotates from out-of-plane to in-plane with increasing thickness. SPLEEM showed that this transition proceeds differently in different materials. In Co films on the Au(111) surface, it starts with the broadening of domain walls; in Fe–Co alloys with Co contents up to 30%, the magnetization tilts initially somewhat into the plane of the film and abruptly switches into the plane when the initially pseudomorphic film transforms into the b.c.c. structure. SPLEEM has also contributed to the understanding of the evolution of the magnetization with thickness of the much discussed system, Fe/Cu(100).

A problem of more practical interest is the exchange interaction between two ferromagnetic layers through a nonferromagnetic spacer layer. It is the basis of the giant magnetoresistance effect that is used in the reading/writing heads of hard disk drives in computers. This interaction oscillates between ferromagnetic and antiferromagnetic coupling with increasing spacer thickness. However, biquadratic coupling, in which the magnetizations in neighboring

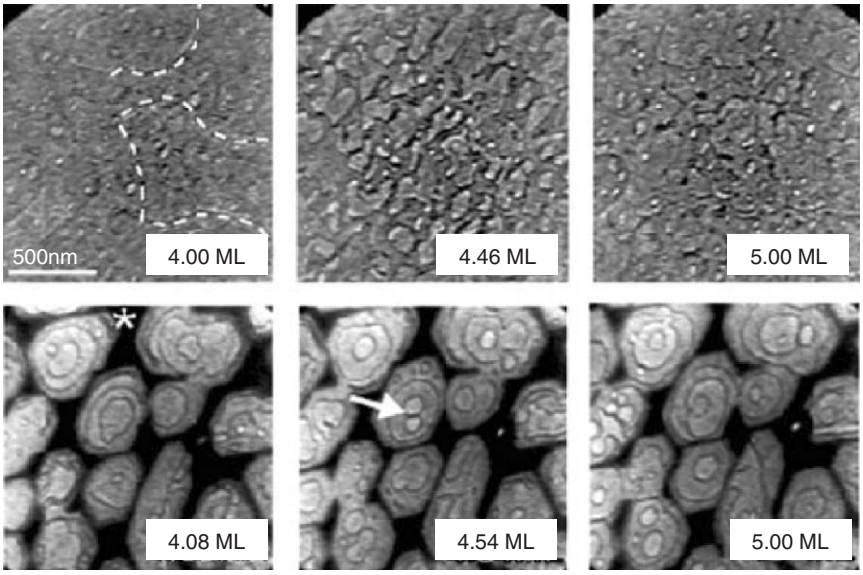


Figure 12 LEEM video frames taken during the growth of Pb from 4 to 5 monolayers on an Si(111) surface at 280 K with (top) and without (bottom) Au interfactant (electron energy – 8 eV). (Reproduced with permission from Schmidt Th *et al.* (2000) Interfactant-mediated quasi-Frank-van der Merwe growth on Si(111) *Physical Review B* 62: 15815; © American Physical Society.)

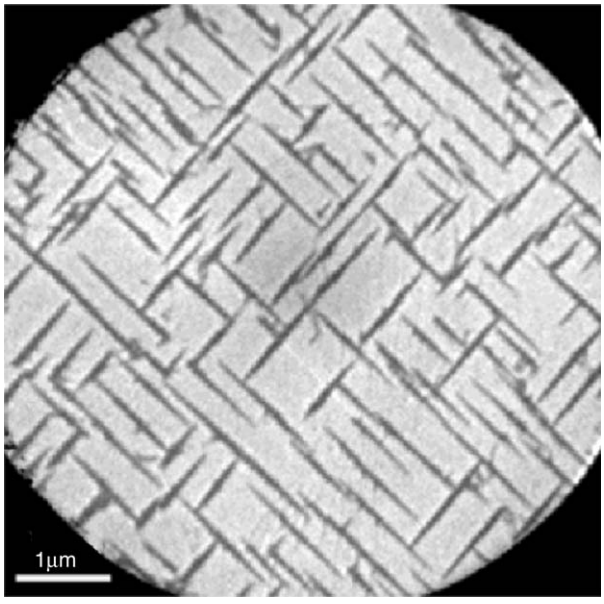


Figure 13 LEEM image of a Co layer with mean thickness of four monolayers grown at ~400 K on a W(100) surface. The Co crystals appear black because their surface is not parallel to the substrate surface (electron energy – 8.3 eV). (Reproduced with permission from Bauer E (1999) Growth of thin films. *Journal of Physics: Condensed Matter* 11: 9365; © IOP Publishing Limited.)

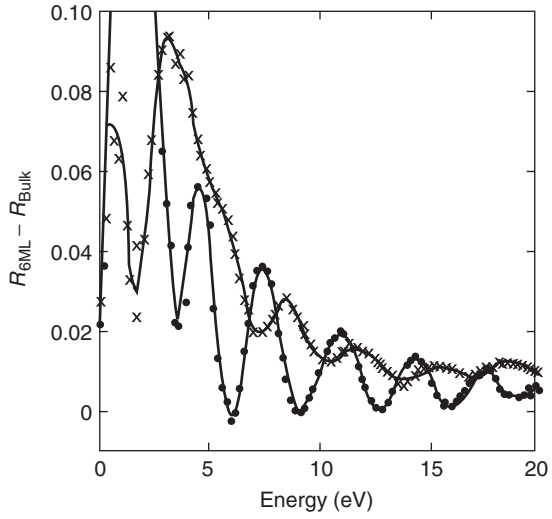


Figure 14 Quantum size oscillations of the intensity reflected from a six-monolayer thick Fe film region on a W(110) surface. The reflectivity of a thick Fe layer has been subtracted. Dots and crosses for majority and minority spins, respectively. (Reproduced with permission from Zdyb R and Bauer E (2002) Spin-resolved unoccupied electronic band structure from quantum size oscillations in the reflectivity of slow electrons from ultrathin ferromagnetic crystals. *Physical Review Letters* 88: 166403; © American Physical Society.)

layers are perpendicular to each other, is also encountered frequently. This coupling is usually attributed to interface roughness. A SPLEEM study of the system Co/Au(111)/Co, however, has shown that it can occur also with rather smooth interfaces. **Figure 15** illustrates this phenomenon for a six-monolayer thick Au spacer. The top two images are from the bottom Co layer with $\boldsymbol{P}$ along the easy ($\parallel \pm \boldsymbol{M}$) and hard ($\perp \pm \boldsymbol{M}$) magnetization axis, respectively, and the bottom-row images correspond to the seven-monolayer thick top Co layer. Ferromagnetic coupling is indicated in the left image but there is also

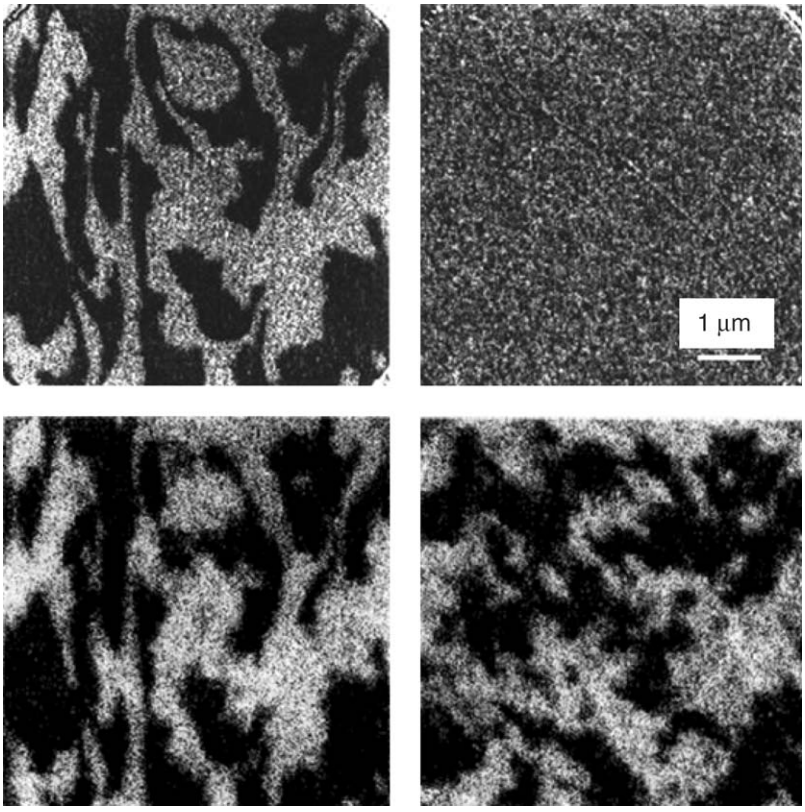


Figure 15 SPLEEM images from a Co/Au(111)/Co sandwich study. The top images are from the bottom layer, the bottom images from the top layer. For explanation see text (electron energy – 2 eV). (Reproduced with permission from Duden T and Bauer E (1999) Biquadratic exchange in ferromagnetic/nonferromagnetic sandwiches: a spin-polarized low energy electron microscopy study. *Physical Review B* 59: 474; © American Physical Society.)

strong magnetization in the right image, which is a biquadratic coupling to the bottom layer.

Concluding Remarks

This article covers only LEEM and SPLEEM. The instruments used in these techniques are equally well suited for all kinds of photoelectron emission microscopies (PEEM), secondary electron microscopy (SEEM), and Auger electron emission microscopy (AEEM), provided that they are equipped with the necessary accessories. Combining these methods allows a comprehensive surface analysis with high lateral resolution. Scanning low-energy electron microscopy, in which the electron beam is focused onto the specimen and scanned across it, has also not been discussed. This technique also uses a cathode lens in front of the specimen. When operated under ultrahigh vacuum conditions, diffraction contrast is possible too but, in general, secondary electrons are used for imaging. In this case, work function contrast or topography contrast are dominating, or – in the presence of specimen contamination – contact potential difference contrast between regions with different doping in semiconductors.

As mentioned at the beginning, LEEM and SPLEEM are microscopies for the study of surfaces and thin films. These can, of course, also be studied with many other methods. The one closely related to LEEM is reflection electron microscopy (REM) in which the surface sensitivity of fast electrons is achieved with a small grazing angle of incidence. This causes a strong foreshortening in the beam direction but has better resolution perpendicular to it. Conventional scanning electron microscopy (SEM) has the advantage that it can be used with much rougher surfaces than LEEM but lacks the surface structure contrast of LEEM. This is true both in the secondary electron emission mode, which is sensitive to the work function, and in the backscattering mode. Spin-polarized SEM, also called SEM with polarization analysis (SEMPA), is comparable in surface sensitivity to SPLEEM but like all SEM methods, is slower because of the sequential image acquisition compared to the parallel image acquisition in SPLEEM. Conventional dark field transmission electron microscopy (TEM) can also image

surface structures in very thin samples, in particular, when diffracted beams resulting from surface reconstructions are used. The resolution is much better than in LEEM but its application rather limited. Finally, there are many other surface imaging methods, foremost being the various scanning probe techniques. In general, they have better resolution than LEEM but are slower and cover smaller fields of view than LEEM. Although several of these methods have a wide overlap with LEEM/SPLEEM, they are largely complementary in their application range.

See also: Photoelectron Spectromicroscopy; Transmission Electron Microscopy.

PACS: 68.37.Nq; 61.14.Hg; 68.35.Bs; 68.55.Ac; 73.21.Fg; 75.25.+z; 75.60.Ch; 75.70.−i

Further Reading

Bauer E (1994) Surface electron microscopy. *Surface Science* 299/300: 102–115.

Bauer E (1997a) Low-energy electron microscopy, spin-polarized low-energy electron microscopy. In: Amelinckx S, van Dyck D, van Landuyt J, and van Tendeloo G (eds.) *Handbook of Microscopy*, vol. 1, pp. 487–503, vol. 2, 751–759. Weinheim: VCH Verlagsges.

Bauer E (1997b) In-situ applications of low-energy electron microscopy. In: Gai PL (ed.) *In-Situ Microscopy in Materials Research*, pp. 1–12. Boston: Kluwer.

Bauer E (1998) LEEM basics. *Surface Review Letters* 5: 1275–1286.

Bauer E (2004) SPLEEM. In: Hopster H and Oepen HP (eds.) *Magnetic Microscopy of Nanostructures*. Berlin: Springer.

Bauer E and Schmidt T (2003) Multi-method high resolution surface analysis with slow electrons. In: Ernst, F., Rühle, M. (eds.), *High Resolution Imaging and Spectroscopy of Materials*, Springer Series in Materials Science, vol. 50, pp. 363–390. Berlin: Springer.

Bauer E and Telieps W (1988) Emission and low-energy reflection electron microscopy. In: Howie A and Valdrè (eds.) *Surface and Interface Characterization by Electron Optical Methods*, NATO ASI Series, vol. B191 pp. 195–233. New York: Plenum.

Phaneuf RJ and Schmid AK (2003) Low-energy electron microscopy: imaging surface dynamics. *Physics Today* 56: 50–55.

Tromp RM (2000) Low-energy electron microscopy. *IBM Journal of Research and Development* 44: 503–516.

Tromp RM (2002) Thermodynamics of nucleation and growth. *Surface Review Letters* 9: 1565–1593.

Nomenclature

E	electron energy (eV, keV)
M	specimen magnetization (no units used)
P	electron beam polarization ($\leqslant 1$, no units used)
λ	electron wavelength (nm)

Low-Temperature Properties of Materials

T Dietl, Polish Academy of Sciences, Warsaw, Poland

Introduction

Heat and charge transport phenomena in solid-state devices are usually satisfactorily described by the semiclassical Drude–Boltzmann theory. Within this theory, the influence of quantum effects and of electron correlation are taken into account in the Born approximation for the momentum relaxation time τ and by the Fermi liquid renormalization of the effective masses m^*, respectively. However, in a number of cases such a theory fails to describe transport properties even in a qualitative way. These nonclassical aspects of electron transport show up particularly clearly at low temperatures, and exhibit a strong dependence on the system dimensionality d. The celebrated example in the two-dimensional (2D) case ($d=2$) is the integer and fractional → quantum Hall effects, but the list of important novel phenomena is much longer. Here, the Coulomb blockade in 0D quantum dots, the quantization of conductance in 1D wires, mesoscopic phenomena in small conductors, quantum localization in macroscopic systems as well as hopping conductivity in the strongly localized regime are discussed. Somewhat related, but not discussed here, is the issue of thermodynamic phenomena in small samples, such as persistent currents in nanorings and nanocylinders.

Through this article, the notion of metal refers to materials whose conductance G, the ratio of the current to the applied voltage, is nonzero in the limit of zero temperature, $G>0$ as $T\to 0$, in contrast to insulators for which $G\to 0$ as $T\to 0$.

Coulomb Blockade

In the case of standard field effect transistors, the equilibrium charge Q accumulated in the channel is determined by a competition between a positive energy of the Coulomb repulsion between the charges and a negative energy of the charge in the

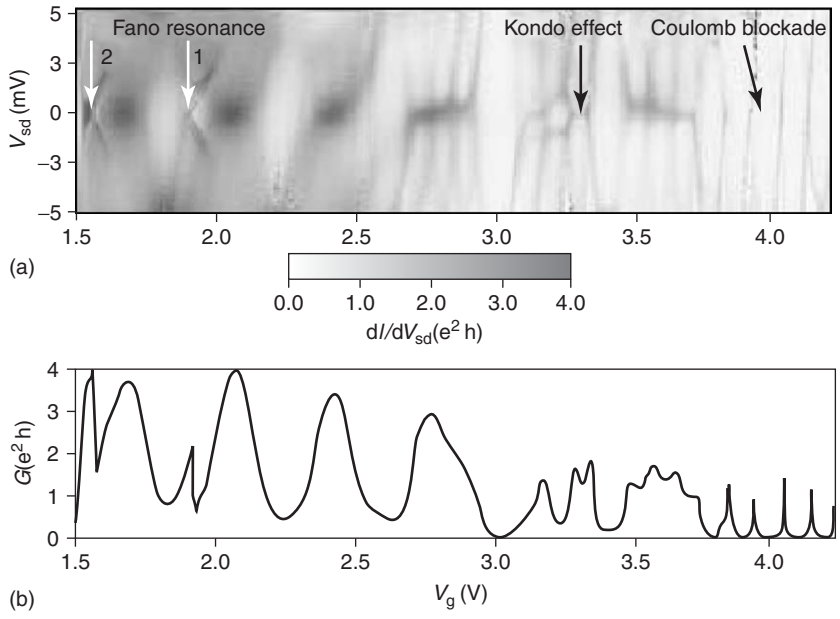


Figure 1 Two-terminal conductance at 0.3 K of a carbon nanotube as a function of the source-drain voltage V_{sd} (a) of gate voltage V_g (b) that increases both carrier density and transmittance of contacts between the sample and the current leads. Regimes corresponding to the Coulomb blockade (low V_g), Kondo effect, and Fano resonance superimposed on Fabry–Perot-like oscillations (high V_g) are visible. At high V_g, the conductance maximums attain the value expected for a metallic carbon nanotube, $G = 4e^2/h$. (Reprinted figure with permission from Babić B and Schönenberger C (2004) *Physical Review B* 70: 195408; © American Physical Society.)

gate potential V_g, $E(Q) = Q^2/2C - QV_g$. By minimizing $E(Q)$, one arrives at the standard linear dependence between Q and V_g, $Q = CV_g$. Since the channel conductance is proportional to the accumulated charge, $G \sim V_g$. If, however, channel dimensions decrease, effects of charge quantization may show up, particularly if the channel is separated from the current leads by tunneling barriers. The charging of such a quantum dot is only possible if $E[(N+1)e] = E(Ne)$. Therefore, as shown in **Figure 1**, except for $V_g, = (N + 1/2)e/C$, the current flow through the dot is blocked and $G = 0$.

The Coulomb blockade is effective as long as the corresponding maximum value of the barrier height e^2/C is greater than the thermal spread of electron energies; for example, $e^2/C = k_B T$ at 1 K for a plate capacitor with the surface $S = 1\,\mu m^2$ and the air gap $b = 4.76$ nm. The Coulomb blockade is also removed if the channel assumes the character of a quantum wire, that is, the contact transmittance is so high that the broadening Γ of the electron states, associated with a finite time the electron spends in the dot, is greater than the barrier height, $\Gamma > e^2/C$, though Fabry–Perot-like interference may lead to conductance oscillations as a function of V_g, if the contacts are not totally transparent. In the intermediate region, the Coulomb blockade is partly removed by the appearance of the Abrikosov–Suhl resonance at the Fermi level. This many-body enhancement of the density-of-states develops below the Kondo temperature whose magnitude increases with the contact transmittance. Furthermore, interference between nonresonant and resonant tunneling via dot-quantum states may lead to additional features in $G(V_g)$ dependence (**Figure 1**), known as "Fano resonances."

Several applications have been proposed for these artificial atoms–quantum dots with a controlled number of electrons. A single electron transistor (SET), electrometer, and primary thermometer constitute a few examples.

Landauer–Büttiker Formalism

A convenient approach to determine conductance of quantum structures is offered by the Landauer–Büttiker formalism, according to which the process of electron transport can be viewed as a scattering problem of fluxes of the Fermi-level electrons entering and leaving the medium through ideal leads. Within this formalism, the conductance can be expressed by eigenvalues of the familiar quantum-mechanical transmission and reflection matrices for the corresponding electron fluxes. These matrices contain information on the potential distribution in a given sample and are subject to symmetry constraints imposed, for instance, by spin degeneracy occurring if no spin-dependent perturbations are present.

Ballistic Transport

Progress in nanofabrication makes it possible to investigate samples, whose dimensions L are smaller than the mean free path of the electrons at the Fermi level, $l = v_F\tau$, so that electron motion is ballistic and G is determined by a sample geometry rather than by a random scattering potential. In a simple two-terminal wire geometry with transparent contacts, the electron spectrum consists of 1D subbands for which electron transmission probability is 1. Hence, the current I generated by the difference in the chemical potentials of the two leads, $-eV$, reads

$$I = \sum_i e^2 v_F^{(i)} N_F^{(i)} V \qquad [1]$$

where k_BT is assumed to be much smaller than the inter-subband distance, and $N_F^{(i)} = 1/hv_F^{(i)}$ is the density of states at the Fermi level of the ith occupied 1D subband. This leads to $G = se^2/h$, where s is the number of the 1D subbands (waveguide modes) crossing the Fermi level and h is the Planck constant. Hence, despite the absence of scattering, G is finite but its value is independent of the wire length and quantized in the units of e^2/h.

As shown in **Figure 2**, characteristic conductance plateau are observed as a function of the gate voltage that changes s. Similar conductance quantization is observed for junctions containing a small number of atoms or molecules. If $s \gg 1$, G can be determined rather accurately from classical "billiard ball"-type models that are particularly useful for complex sample geometries.

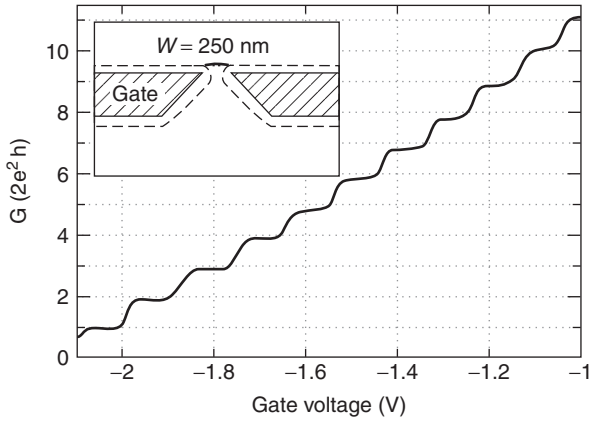


Figure 2 Quantized conductance as a function of gate voltage V_g in a split-gate modulation-doped GaAs/(Al,Ga)As:Si heterostructure. A submicrometer channel is defined in 2D electron gas by negatively biased top metallic gates that deplete carriers in the regions underneath and affect the channel width as shown in the inset. (Reprinted figure with permission from van Wees BJ, van Honten H, Beenakker CWJ, Williamson JG, Konwenhoven LP, *et al.* (1988) *Physical Review Letters* 60: 848; © American Physical Society.)

Mesoscopic Transport Phenomena

Without too much simplification, all novel quantum phenomena to be discussed in this section are caused by two effects, which to a large extent determine the spectrum of the eigenvalues of the transmission matrices in the diffusive regime, $L > l$. The first is quantum interference of transition amplitudes corresponding to different electron paths through the disordered medium. The second is a certain spectral rigidity of the eigenvalues, which has been widely discussed in the context of quantum chaos, but can be traced back to the celebrated Mott observation that in disordered system, states close in energy are far apart in real space.

Similarly, in the case of images of fringes produced by interference of light, a change of the conductance occurs over a scale corresponding to the length of the interfering waves. Such an effect was observed as irregular but reproducible conductance fluctuations as a function of the gate voltage, that is, the de Broglie wavelength of the electrons at the Fermi level. Since the magnetic field (vector potential) affects the phase of the wave function also, similar aperiodic fluctuations occur when the magnetic field is ramped, the correlation field corresponding to one quantum flux through the sample surface S, $B_c \approx h/eS$. Of course, in multiconnected geometries, a substantial number of electron trajectories encompass the same magnetic flux and the fluctuations, as a function of the magnetic field acquires a periodic component. This expectation has been confirmed in a series of beautiful experiments with nanorings, in which the Aharonov–Bohm (AB) oscillations with the period $B_{AB} = h/(e\pi r^2) \approx 4.136\,\mathrm{mT}\,\mu\mathrm{m}^2/(\pi r^2)$, where r, the ring radius has been put into evidence, as shown in **Figure 3**.

An interesting aspect of the conductance fluctuations is the universality of their amplitude, closely related to the spectral rigidity mentioned above. According to diagrammatic calculations and numerical simulations as well as numerous experimental studies, the root-mean square (RMS) value of the fluctuation amplitude of two-terminal conductance G of any metallic sample at zero temperature is given by $\mathrm{RMS}(\Delta G) = ae^2/h$. Here, the numerical value of a depends on the universality class, that is, on whether spin and/or time reversal symmetries are conserved, a typical value being $a = 1/2$. Since, however, $\mathrm{RMS}(\Delta G)$ is independent of scattering details and the system size L, the effect is known as "universal conductance fluctuations (UCF)."

As already mentioned, zero-temperature conductance contains information on the actual impurity distribution in a given sample. Actually, conductance is nonlocal and its value G depends on the impurity

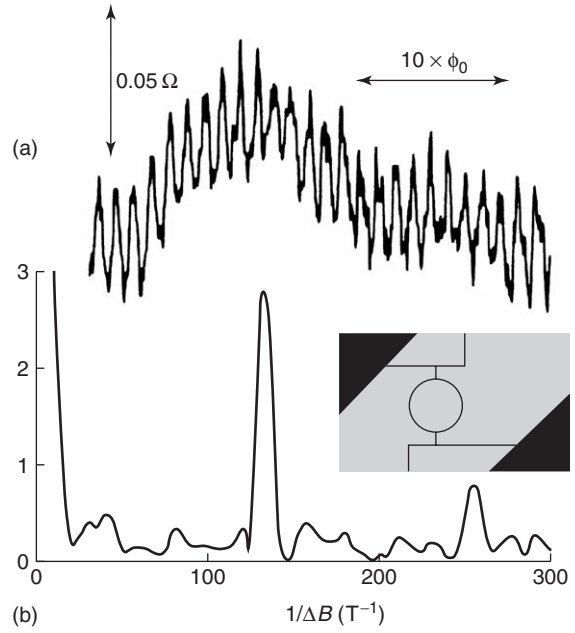


Figure 3 Oscillations of conductance at 10 mK in a gold ring of diameter 784 nm (a), which according to the Fourier transform (b), contain components corresponding to the Aharonov–Bohm period *h*/*e* and *h*/2*e*. (Reprinted figure with permission from Washburn S, Umbach CP, Laibowitz RB, Webb RA *et al.* (1985) *Physical Review Letters* 54: 2696; © American Physical Society.)

distribution even outside the region embedded by the voltage probes. Furthermore, statistical fluctuations of G for various impurity distributions are again characterized by $\mathrm{RMS}(\Delta G) = ae^2/h$. This means that conductance loses a self-averaging property at zero-temperature.

An important question then arises, why these quantum effects are not observed in the case of standard resistors or, in other words, how to model a cross-over between quantum and classical behaviors. Obviously, the coupling to the reservoir at $T > 0$ leads to various random processes such as atom diffusion, flipping of localized spins, lattice and Fermi liquid excitations. Each of these processes leads to changes in the potential landscape, and thus in G, on timescales faster than the time constant of a standard resistance meter. Hence, these processes can wash out quantum signatures in electron transport such as AB oscillations and UCF.

As an example, the effect of impurity tunneling and diffusion is considered. A mean change of G upon displacing a single scattering center by ΔR is evaluated by diagrammatic technique to be $\mathrm{RMS}(G_1) \approx \mathrm{RMS}(\Delta G)[\mathrm{C}(\Delta R k_\mathrm{F})\sigma_\mathrm{imp}/lL^{d-2}]^{1/2}$. Here $C(x \to 0) = x^2/12$, and $C(x \to \infty) = 1$; $\sigma_\mathrm{imp} = 1/ln_\mathrm{imp}$ is the cross section of the relevant scattering centers of concentration n_imp. In the case of, for example, neutral impurities in semiconductors, σ_imp is substantially smaller than that corresponding to the unitary limit $\sigma_\mathrm{imp} \approx 1/k_\mathrm{F}^{d-1}$. Hence, in general, $\mathrm{RMS}(G_1)/\mathrm{RMS}(\Delta G) \ll 1$, which means that the motion of a single impurity perturbs UCF only weakly.

However, the intensity of time-dependent processes grow with the temperature. In particular, if the fraction p of all impurities in the sample undergoes uncorrelated displacement, then the corresponding L dependent conductance variance $\mathrm{RMS}(G_p) = [p(T)L^d n_\mathrm{imp}]^{1/2}\mathrm{RMS}(G_1)$ can attain the universal value $\mathrm{RMS}(\Delta G)$. At this point, the amplitude of AB oscillations and UCF start getting reduced by random impurity motion. This means that the phase coherence length L_φ controlled by time-dependent processes has diminished to the size of the sample L. At still higher temperatures, the sample can be considered as consisting of $[L/L_\varphi(T)]^d$ uncorrelated "coherence boxes." In this regime, the classical self-averaging property of G is recovered as the sensitivity of G to, say, a magnetic field decreases with L and increases with L_φ. Here $L_\varphi = (D\tau_\varphi)^{1/2}$, where $D = k_\mathrm{F} l/dm^*$ is the diffusion coefficient and τ_φ is the phase coherence time. Usually, L_φ is controlled by Fermi liquid thermal excitations and is ~ 0.1–$1\,\mu\mathrm{m}$ at 1 K, so that $L_\varphi \gg l$, that is, a diffusing electron experiences many elastic collisions before its phase coherence is lost.

In addition to $L_\varphi(T)$, there exists one more relevant length scale at nonzero temperatures, which is associated with an averaging of the interference effects by thermal spread $\Delta\lambda$ of the de Broglie wavelengths of electrons contributing to G. This spread is of significance once $\Delta\lambda/\lambda^2$ becomes greater than the inverse of a mean length of the electron path through the sample, that is, if $L > L_\mathrm{T} = (hD/k_\mathrm{B}T)^{1/2}$. Actually, $L_\varphi(T)$ and L_T control the dimensionality also – if one of the sample dimensions, say the film width, becomes smaller than either of them, a dimensional crossover takes place.

A strong sensitivity of G to the motion of impurities has led to the suggestion that the quantum effects in question account for a surprisingly large magnitude of $1/f$ noise observed in any conductor. This is a conjecture corroborated by a change in the noise magnitude once the magnetic field lifts the spin degeneracy. **Figure 4** depicts the appearance of strong low-frequency noise and recovery of UCF in a quantum wire of n-$Cd_{0.93}Mn_{0.07}Te$:I below 300 mK, where spin-glass freezing slows down the dynamics of the Mn spins.

Quantum Localization and Hopping Conduction

The above discussion might suggest that the quantum effects are only important in mesoscopic samples L_φ,

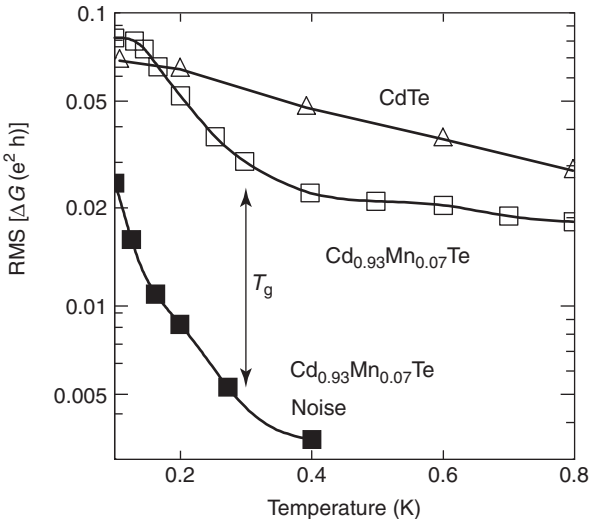


Figure 4 RMS of aperiodic conductance fluctuations as a function of the magnetic field in the vicinity of 2 kOe and RMS of conductance noise measured for a submicron wire of *n*-$Cd_{0.93}Mn_{0.07}Te$ with low-pass filters rejecting frequencies above 0.3 and 3 Hz, respectively. Slowdown of Mn spin dynamics below spin freezing temperature T_g leads to appearance of low-frequency noise and to reentrance of conductance fluctuations with the amplitude comparable to that observed in a similar wire of nonmagnetic *n*-CdTe. (Reprinted figure with permission from Jaroszyński, Wróbel J, Karczewski G, Wojtowicz T, Dietl T, *et al.* (1998) *Physical Review Letters* 80: 5635; © American Physical Society.)

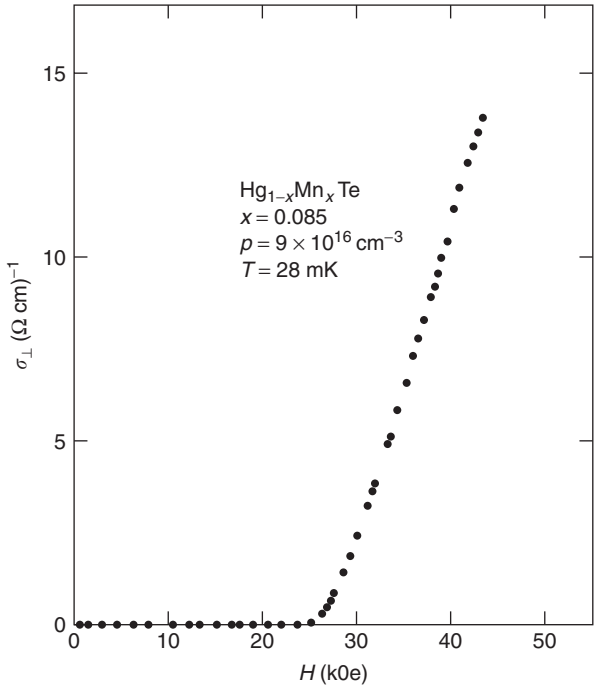


Figure 5 Low-temperature conductivity of a bulk *p*-(Hg, Mn)Te sample as a function of the magnetic field showing an insulator-to-metal transition. The transition results from the field-induced ordering of Mn spins and the corresponding increase in the hole mean free path. (After Wojtowicz T, Dietl T, Sawicki M, Plesiewicz W, Jaroszyński J, *et al.* (1986) *Physical Review Letters* 56: 2419.)

$L_T > L \gg l$, while they become totally insignificant at $L \gg L_\varphi$, L_T. Actually, there are two quantum phenomena that survive the averaging procedure and control low-temperature conductance and magnetoconductance of disordered macroscopic samples.

The first is quantum interference corresponding to the clockwise and counterclockwise electron trajectories along the same self-intersecting electron paths. Since such interference is constructive, it increases the probability of return to the starting point, and thus diminishes the conductivity. The presence of such interference is best seen in cylinders of radius $r < L_\varphi$ and length $L \gg L_\varphi$. In this and related geometries, magnetoconductance exhibits Altshuler–Aronov–Spivak oscillations characterized by the period $B_{AAS} = h/2e\pi r^2$.

The second effect stems from the earlier mentioned disorder-induced correlation between electron energies and their position in real space. Such an additional correlation introduces modifications to the standard Landau theory of Fermi liquids – electron–electron interactions cannot be longer taken into account by a qualitatively unimportant renormalization of the effective mass m^*.

These two quantum effects turn out to be controlled by diffusion poles in the particle–particle and particle–hole correlation functions of the Fermi liquid in the random potential. In the limit of weak disorder, $k_F l \gg 1$, and at low temperatures, $k_B T \ll h/\tau$, they lead to characteristic terms in conductance proportional to $T^{1/2}$ and $B^{1/2}$ in the 3D case, and to $\ln T$ and $\ln B$ in the 2D films, whose sign depends on the universality class as well as on the relative magnitude of the cyclotron and spin splittings.

Interestingly enough, these effects account for Anderson–Mott localization in disordered conductors. This localization results from scattering by many centers, so that the localization radius ξ is usually much greater than the inter-impurity distance. According to the scaling theory, a continuous disorder-driven metal-to-insulator transition occurs when $k_F l \approx 1$ in the 3D case, as shown in **Figure 5**. In doped semiconductors, this takes place when the average distance between electrons becomes 2.5 times larger than the effective Bohr radius of the relevant impurity. In contrast, in any 1D- and 2D-disordered system all states are localized, ξ growing exponentially with $k_F l$ in the 2D case. However, transport signatures of carrier localization can only be seen if $L_\varphi(T)$, L_T, and L are greater than ξ; otherwise, one is in the weakly localized regime discussed above.

In the strongly localized regime, conductance proceeds owing to phonon-assisted tunneling between occupied and empty localized states. In this case, the Mott variable range hopping sets in at low temperatures, where $G(T) \sim \exp[-(T_o/T)^{\gamma}]$, with $\gamma = (\Theta + 1)/(\Theta + 1 + d)$. Here T_o is a material constant, while Θ describes the character of the Efros–Shklovskii Coulomb gap, that is the decay of the density of localized states at the Fermi level, $N(\varepsilon) \sim |\varepsilon - \varepsilon_F|^{\Theta}$, in the energy interval important for the hopping conductivity at given temperature T.

To what extent charge and spin thermal excitations, in addition to lattice excitations, can promote hopping conductance is under active debate.

See also: Conductivity, Electrical; Conductivity, Thermal; Onsager Relations; Quantum Hall Effect; Semiconductor Nanostructures.

PACS: 73.23.–b; 72.15.Rn; 72.70.+m

Further Reading

Beenakker CWJ and van Houten H (1991) Quantum transport in semiconductor nanostructures. *Solid State Physics* 44: 1–228. New York: Academic Press.

Datta S (1995) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.

Efros AL and Pollak M (eds.) (1985) *Electron–Electron Interactions in Disordered Systems*. Amsterdam: North-Holland.

Imry Y (1997) *Introduction to Mesoscopic Physics*. Oxford: Oxford University Press.

Nomenclature

a	numerical coefficient between the magnitude of universal conductance fluctuations and the unit of conductance e^2/h
b	air gap between capacitor electrodes
B	magnetic field
B_c	correlation field in universal conductance fluctuations
C	capacitance
$C(x)$	a function describing effect of the defect displacement on the universal conductance fluctuations
d	space dimensionality; $d=1$ for one-dimensional (1D) case, $d=2$ for 2D, $d=3$ for 3D cases, respectively
D	charge carrier diffusion coefficient
e	absolute value of the magnitude of the electron charge
E	total energy of charges in a capacitor
f	noise frequency
G	electrical conductance, $G = I/V$
G_1	change of conductance for a displacement of a single scattering center
h	Planck constant
i	index of 1D electric subbands
k_B	Boltzmann constant
k_F	wave vector of the carriers at the Fermi level
l	mean free path for elastic collisions
L	sample linear dimensions
L_φ	carrier coherence length
L_T	carrier thermal length
m^*	effective mass of carriers
n_{imp}	concentration of scattering centers
p	fraction of all impurities in the sample undergoing uncorrelated displacement
Q	charge
r	radius of a ring or cylinder
S	surface
T	temperature
T_o	a temperature characterizing thermal activation of hopping conductance
V	voltage
V_g	gate voltage
v_F	Fermi velocity
γ	exponent describing character of temperature dependence of hopping conductivity
ε	carrier energy
ΔG	amplitude of conductance fluctuations
ε_F	Fermi energy
λ	de Broglie electron wavelength
σ_{imp}	cross section of electron scattering by a single impurity
Θ	exponent describing the character of the Coulomb gap in the density of localized states
ξ	localization length
τ	momentum relaxation time
τ_φ	phase coherence time

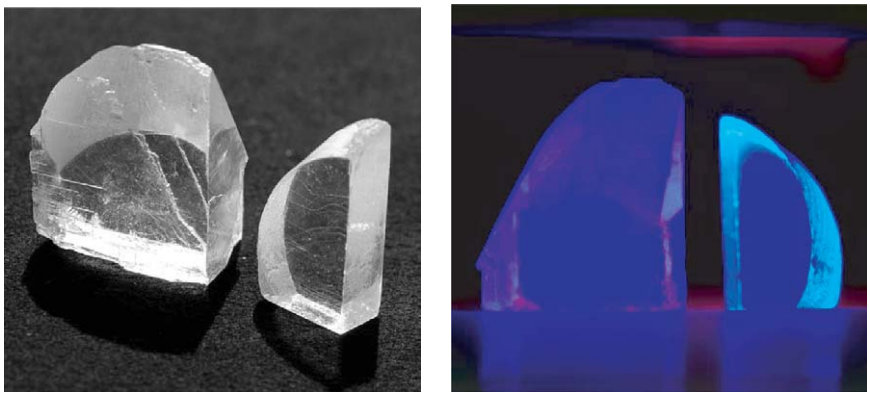
Luminescence

R Capelletti, University of Parma, Parma, Italy

Introduction

As a consequence of energy absorption at moderate temperature, a luminescent material re-emits part of it as light (IR, visible, and UV). Such a light emission was called luminescence by E Wiedemann in 1888 and it must not be confused with the "blackbody" one, which occurs at high temperatures. The energy absorption or excitation of a luminescent material can be accomplished in different ways and these are expressed in a prefix accompanying the word luminescence.

For example, the excitation energy may be supplied by sound or light waves (sonoluminescence or photoluminescence, see **Figure 1**), by moderate heating (thermoluminescence (TL)), by electric field (electroluminescence (EL)), by particles, as electrons (cathodoluminescence (CL)), by rubbing (triboluminescence), and by high-energy radiation, as X- or γ-rays, (radioluminescence). The light emission may result from chemical reactions (chemiluminescence) or physical processes, as crystallization (crystalloluminescence) or solution (lyoluminescence). The prefix may also originate from specific features of the luminescent materials, for example, bioluminescence is a form of chemiluminescence produced in living organisms and pyroluminescence is related to the radiative emission from atoms or molecules in a gas, previously excited by heating. **Table 1** provides a more detailed list of different luminescence phenomena. A further distinction comes from the time elapsed between the excitation and the light emission, that is, if the emission stops as the excitation ceases, one is dealing with fluorescence, while if it lasts for a more or less long time after the excitation, one is dealing with phosphorescence (or afterglow). Independent of this definition, a luminescent material is conventionally defined as a phosphor; on the basis of its etymology, that is, in Greek it means "to bear light," and should not be confused with the chemical element phosphorus P ($Z = 15$). Luminescence can be regarded as an ability of matter to convert various forms of energy into light; therefore, luminescent materials can be exploited and tailored for a variety of applications. In fact, the color of the emitted light depends on the intrinsic characteristics of the material, on the impurity and defect content, on the temperature, on the external applied fields (stress, magnetic, and electric), and on the particle size (quantum dots). The number of applications in science and everyday life is extremely large and continuously growing. A few examples are mentioned in the following and in **Table 1**. The solid-state laser operation (e.g., ruby, Nd:YAG) is based on photoluminescence (stimulated emission). In the common fluorescent tubes, a phosphor coating absorbs the invisible UV radiation (emitted by the de-excitation of Hg atoms) and re-emits visible light. Phosphors added to laundry detergents and paper make white textiles and paper brighter, because the sunlight UV tail excites a pale blue luminescence which adds to



(a) (b)

Figure 1 Photoluminescence excited by 253 nm UV light in KCl: Eu single crystals grown by the Bridgman method. (a) Daylight illumination; (b) the samples are illuminated only by UV light. The Eu concentration is 30 and 399 ppm in the larger and smaller samples, respectively. (The crystals were grown at Institute of Low Temperature and Structure Research, Wroclaw, Poland.) (Courtesy Prof. Maria Suszyńska. Photographs by Carlo Gardini, Parma (carlgard@inwind.it).)

Table 1 Different types of luminescence, their origin, examples and applications, based on E N Harvey

Type	*Origin*	*Examples/applications*
Candoluminescence	Emission from incandescent solids at shorter wavelengths than those expected from their blackbody emission	Rare-earth oxides in Auer lamps (outside lighting); Mg/MgO in flash-light photography
Pyroluminescence	Emission from excited atoms, ions, and molecules in flames	He, Na, Fe, Ca, etc.; Sun line emission spectrum; lamps, fireworks; I_2, NO_2 in flames
Thermoluminescence	Emission from solids previously irradiated and then moderately heated	Heavy spar; fluorite; zircon; personal and environmental dosimeters; archeological pottery dating; rock dating (e.g., stalagmites); irradiated food identification
Photoluminescence	Irradiation by UV or visible light	Impurities and defects (e.g., color centers) in crystals and minerals, phosphors: laser, counterfaction detection
Radioluminescence	Irradiation by high-energy photons (X-, γ-rays) or particles (e.g., neutrons)	Scintillators (e.g., NaI: Tl, CsI, BGO, $PbWO_4$), AgBr grains for X-ray film radiography; scintillator-based calorimeters to detect high-energy particles; Cherenkov effect
Cathodoluminescence	Excitation by electron impact on solids or gases	Color TV screen; impurity mapping, microinclusions in minerals by using electron microscopy; discharge lamps, aurora borealis in air
Anodoluminescence	Irradiation by anions	ZnS phosphors
Ionoluminescence	Excitation produced by particles (e.g., alpha particles emitted by radioactive elements)	Early clock dials; sapphire
Sonoluminescence	Ultrasonic irradiation of liquids	Small gas bubbles emit very short light flashes (a few ps long)
Galvanoluminescence	Luminescence emitted during electrolysis	Electrolysis of NaBr
Electroluminescence	Excitation by an electric field	LED; injection lasers; alpha-numeric displays; lighting, medical applications; backlight in electric clocks; flexible EL lamps; laser pointers
Crystalloluminescence	Excitation induced by rapid crystallization of certain salts from liquid solution or from molten phase	Growth of crystallites from solutions: NaCl, As_2O_3, K_2SO_4, organic compounds
Lyoluminescence	Light emitted on dissolution (e.g., in water) of pre-irradiated solids	Alkali halides; organic materials (dosimetry); irradiated food identification
Chemiluminescence	Excitation produced by chemical reactions	Luminol; lightsticks; reactions between gaseous species at high temperature; chemical species trace (e.g., iron in hemoglobin) determination
Bioluminescence	Chemiluminescence in biological matter (e.g., accompanying the oxidation of luciferin)	Bacteria, fungi, algae, glow worms, fishes, mollusks, jellyfishes, fireflies, etc.
Triboluminescence	Excitation brought about by the crushing of crystals	Uranyl nitrate, sugar, quartz, rock salt, Wint-O-Green candies

the white reflected light. Simple apparatuses exploit the UV-induced fluorescence in stamps, money (e.g., euro bills), credit cards, and traveler's checks to detect counterfeiting. Many traffic safety products utilize photoluminescence. Different color light emitting diodes (LEDs), used in a variety of displays, are based on EL. The color-TV screen phosphor operation exploits CL. The chemiluminescence of luminol reacting with hemoglobin is used in crime scene investigations to detect even tiny traces of blood, which might last for years at the scene. Scintillator applications originate from radioluminescence and are addressed to astrophysics, high-energy physics (e.g., the calorimeter on the KTeV beamline at FermiLab, see **Figure 2**), medical diagnostics, and security controls. TL is exploited in the dating of buried materials and dosimetry. Phosphorescence provides nighttime visibility to the dials of alarm clocks and watches, by using long-persistence phosphorescent materials, activated by daylight. In the earlier watches, the same purpose was accomplished by a mildly radioactive paint. In newer electric clocks, the backlight behind the dial is based on electroluminescence.

In this article, after brief hints about the historical development of luminescent phenomena, some fundamental mechanisms of emission of luminescence are described. The basic concepts and the spectroscopic tools required to enlighten such microscopic

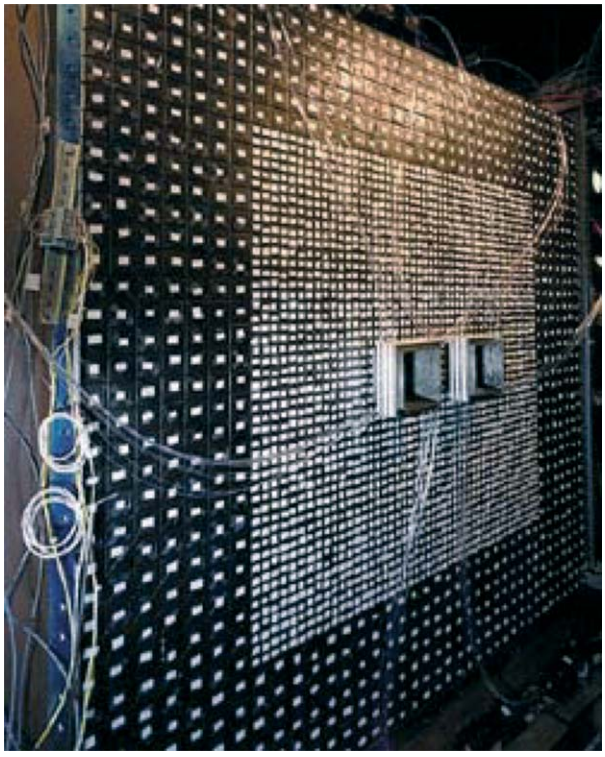


Figure 2 Scintillator-based calorimeter on the KTeV beam line at the FermiLab (Batavia, Illinois, USA). It is a 1.9 m × 1.9 m array of 3100 pure scintillating CsI crystals. The calorimeter measures with high precision energy of particles emanating from kaon decays by recording the light emitted (radioluminescence) as the particles pass through the scintillating material. (Reproduced from www.cerncourier.com with permission of Fermi National Accelerator Laboratory.)

mechanisms are illustrated mainly for photoluminescence. A few examples of thermoluminescence, electroluminescence, and cathodoluminescence are also supplied. Phenomena related to inorganic condensed matter are considered.

Historical

Luminescence-related phenomena have been observed for thousands of years, but only in the last century were they systematically studied, understood, and exploited. The bioluminescence of fireflies was reported in Chinese literature as early as three millennia ago, while Pliny, the Elder in Naturalis Historia, mentions the "cold fire" emitted in the dark by glow worms, jellyfishes, and mollusks and describes various gems, such as ruby, which shine after being exposed to the sun. In 1602, Vincenzo Casciarolo, an Italian cobbler and amateur alchemist, discovered the Bolognian Phosphorus or Litheophosphorus. This natural stone (heavy spar) became the first object for the scientific study of luminescent phenomena and raised the interest of many scientists, such as Galileo Galilei. Robert Boyle reported to the London Royal Society the glimmering light emitted by a diamond when placed on the warmest part of his body and the effects produced by the luminous bacteria which feed on decaying flesh (1672). In the same period, Francis Bacon and Lorenzo Magalotti described the light emitted in the dark by sugar, as a consequence of friction. In 1824, Friedrich Mohs observed that fluorite shone in the dark on being heated and called this property "fluorescence" after the mineral (nowadays it is named thermoluminescence). Sir David Brewster (1833) noted that a beam of sunlight, condensed by a lens into a chlorophyll solution, was marked by a bright red emission. Probably the first to explain it was Sir George Stokes, who also named the phenomenon fluorescence again after fluorite, which exhibits a strong photoluminescence under UV light, in addition to phosphorescence and thermoluminescence, already reported by Mohs. Stokes is credited with the discovery (1852) that fluorescence can be induced in certain substances by their stimulation with UV light. Wiedemann, in 1888, proposed the name luminescence (Latin, lumen – light) for this wide class of phenomena. He distinguished six types – photo-, electro-, thermo-, tribo-, crystallo-, and chemi-luminescence. An extensive study of these phenomena and of phosphorescence induced by insolation (i.e., exposure to sunlight) was also performed by Edmond Becquerel, who defined them as phosphorescence (1882–97).

Microscopic Interpretation of Some Luminescence Phenomena

Luminescence may occur in pure and perfect crystals (intrinsic luminescence), may be related to intrinsic defects (e.g., F centers in ionic crystals), or to impurity-induced defects (extrinsic luminescence). Impurities and defects are either intentionally introduced for scientific and application purposes or are already present in the material under study. The energy delivered to the crystal to excite luminescence (e.g., X-rays, high-energy particles, see **Table 1**) may also produce defects. In **Figure 3**, a few mechanisms responsible for luminescence are exemplified in the case of electronic transitions, although transitions between the vibrational levels of some molecular groups (e.g., CN in alkali halides) may also give rise to luminescence (in the IR range). An electron is raised from the valence band (VB) to the conduction band (CB) as a consequence of a transition induced by the energy absorption, that is, the excited electron, which is free in the CB, subsequently recombines with a free hole left in the VB, by emitting a photon (band to band radiative recombination, **Figure 3a**). The emission occurs in the UV, visible,

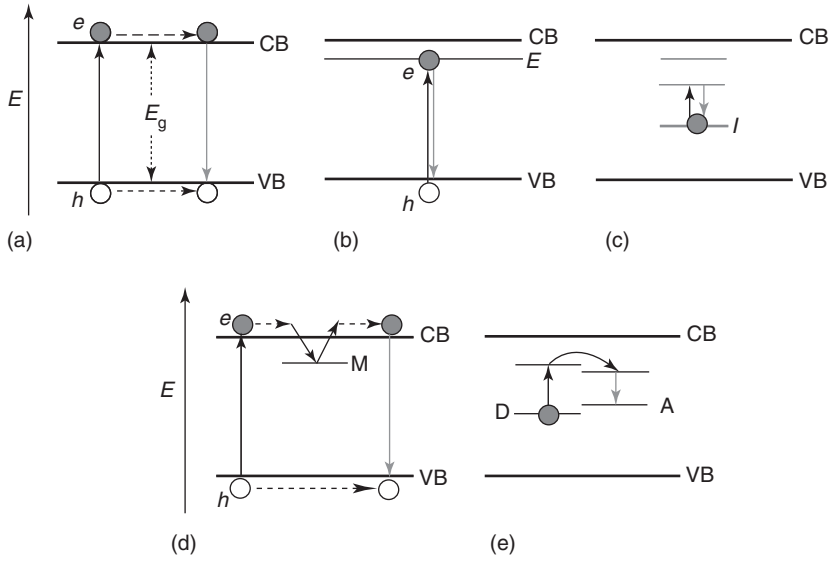


Figure 3 Examples of electronic processes responsible for luminescence: (a) Band-to-band emission; (b) excitonic emission; (c) impurity luminescence; (d) band-to-band emission, delayed by trapping and de-trapping processes; (e) energy transfer between a donor (D) and an acceptor (A). CB indicates the bottom of the conduction band, VB the top of the valence band. The gray circles indicate the electron and the open circles the hole. The vertical black arrows indicate the energy absorption, the gray ones the radiative emission. The dashed arrows indicate the electron (or hole) migration in the conduction (or valence) band. Nonvertical arrows indicate nonradiative processes.

or near IR range according to the amplitude of the energy gap E_g. The excitation can also produce an exciton, that is, an electron, which being still more or less bound to the hole left in the VB, occupies a level just below the bottom of the CB. The radiative recombination of the electron–hole pair gives rise to the excitonic luminescence (**Figure 3b**). Both processes are intrinsic. An impurity (or a defect) can introduce in the energy gap a set of localized levels: an electron, by absorbing energy, jumps from the ground to an excited level of the impurity and subsequently decays by emitting radiation (**Figure 3c**). An impurity or defect can introduce metastable or trap levels in the energy gap: an electron, raised to the CB as in **Figure 3a**, may be trapped in the metastable level M (**Figure 3d**), instead of recombining immediately with a hole in the VB. The band-to-band radiative recombination is delayed until the trapped electron gains enough energy, for example, from the lattice vibrations, to become free again in the CB. The time necessary to empty the trap may be rather long at ordinary temperatures; the delayed light emission is named phosphorescence (or afterglow). By a moderate temperature increase, the time to empty the trap can be significantly shortened, thus giving rise to thermoluminescence. The distribution of impurities (either of the same or of different kind) within the sample plays an important role. The energy acquired by an impurity (D, the donor) from the excitation source may be transferred to a neighboring impurity (A, the acceptor) which may decay radiatively. This means that D is excited and A emits its own luminescence (**Figure 3e**). The process, defined as energy transfer, occurs, for example, if the two impurities are rather close to each other (see sections "Time dependence and lifetimes" and "Concentration effects and energy transfer" and **Figures 6–8**).

Photoluminescence

The Franck–Condon Scheme

In gases, the emission from excited atoms is characterized by a narrow line lying at the same wavelength as that of the excitation (pyroluminescence). In condensed matter, due to the interaction of the emitting center with neighboring atoms, the emission is broader giving rise rather to an emission band than to an emission line. Moreover, the emission occurs, as a rule, at lower energy (longer wavelength) with respect to the excitation one. The energy shift, named Stokes shift, is explained in terms of the Franck–Condon scheme (**Figure 4**). The electron energy levels in a luminescent center depend on the average distance of the center from the neighboring atoms (that is, on the so-called configurational coordinate), which is at variance with the simplified description given in **Figure 3c**. The electron is excited from A (in the ground state) to B (in the excited state) by absorbing a photon. The transition is vertical on the basis of the adiabatic approximation. The electron gives up part of the absorbed energy to the lattice vibrations as it goes to

the relaxed excited state C. After living in C for the lifetime typical of the excited state (10^{-9}–10^{-7} s), it decays into the ground state, again by a vertical transition C–D, with a photon emission. The electron returns into the initial state A, by giving up energy to the lattice. The typical times for the vertical transitions are $\sim 10^{-15}$–10^{-16} s, only the electrons being involved. Those related to the energy transfer to the lattice are roughly three orders of magnitude longer, the more inert nuclei being involved. The excitation energy, represented by AB, is larger than the emission one CD, because part of the absorbed energy (AB) is transferred to the lattice vibrations (along BC and DA), accounting for the Stokes shift, that is, the absorption (excitation) band lies at energy higher than that of the emission band, see **Figures 4** and **5**.

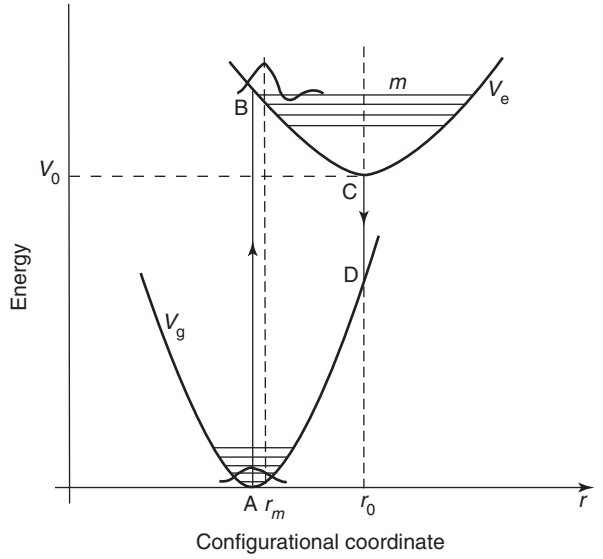


Figure 4 The Franck–Condon scheme.

Excitation and Emission Spectra

The emission intensity is related to the number $N_e(\lambda_{em})$ of emitted photons per unit time at a given wavelength λ_{em}. This depends on the number $N_a(\lambda_{exc})$ of absorbed photons per unit time at a given excitation wavelength λ_{exc} and on the luminescence quantum yield η, that is,

$$N_e(\lambda_{em}) = \eta N_a(\lambda_{exc}) = \eta N_i(\lambda_{exc})\{1 - \exp[-\mu(\lambda_{exc})x]\} \quad [1]$$

where $N_i(\lambda_{exc})$ is the number of incident photons per unit time at a given λ_{exc}, $\mu(\lambda_{exc})$ the corresponding absorption coefficient, and x the sample thickness.

For low $\mu(\lambda_{exc})x$ values, that is, for thin and/or weakly absorbing samples, eqn [1] becomes

$$N_e(\lambda_{em}) \approx \eta N_i(\lambda_{exc})\mu(\lambda_{exc})x \quad [2]$$

or

$$N_e(\lambda_{em}) \propto \eta N_i(\lambda_{exc})N_L \quad [3]$$

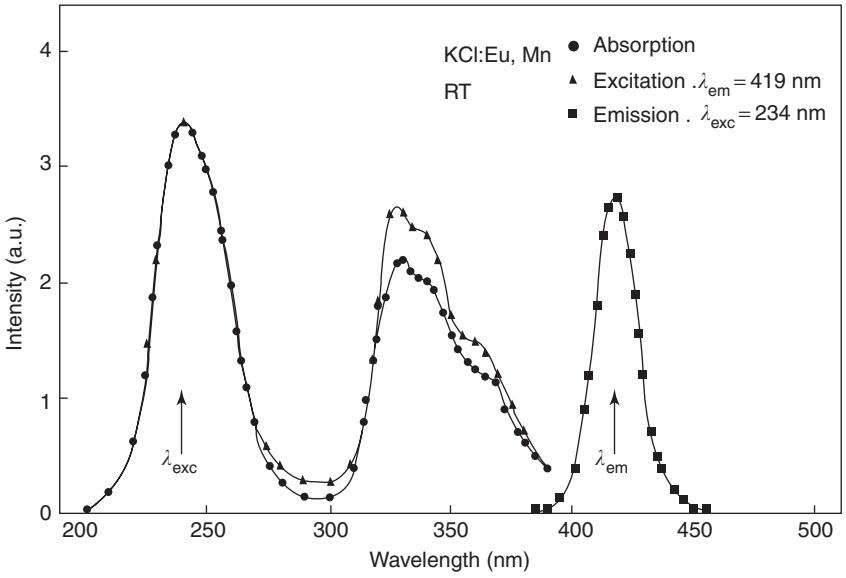


Figure 5 Example of absorption (solid circles), excitation (triangles), and emission (squares) spectra measured at room temperature on a single crystal of KCl: Eu (54 ppm), Mn (3 ppm). The sample was quenched from 500°C: in this way the impurities (Eu and Mn) are dispersed and well separated. The absorption, excitation, and emission spectra are related only to the presence of Eu. The excitation spectrum is measured by keeping constant the emission wavelength $\lambda_{em} = 419$ nm (in the blue) and its intensity is normalized to the absorption spectrum at 234 nm for the sake of comparison. The emission spectrum is measured by keeping constant the excitation wavelength $\lambda_{exc} = 234$ nm (in the UV). For the luminescence color, see **Figure 1**. (Reproduced with permission from Capelletti R, Manfredi M, Cywiński R, Damm JZ, Mugeński M, and Solzi M (1987) Energy-transfer mechanism in the KCl: Eu^{2+}, Mn^{2+} system. *Physical Review B* 36: 5124–5130.)

because μ is proportional to the number of luminescent centers N_L.

In real solids the photoluminescence, excited in a broad wavelength (energy) range, may consist of more than one emission band, due to the presence of more than one luminescent center. The analysis of the emission, excitation, and absorption spectra is a powerful tool to identify the centers. By keeping the excitation wavelength λ_{exc} fixed, an emission spectrum is obtained by measuring the emitted light intensity (or $N_e(\lambda_{em})$) as a function of the emitted wavelength λ_{em} (see **Figure 5**). On the other hand, by keeping λ_{em} fixed, an excitation spectrum is obtained by measuring the emitted light intensity as a function of λ_{exc}, see **Figure 5**. In this way it is possible to identify the λ_{exc} range which excites a given emission. The excitation band should fit to an absorption band (see eqns [1] and [2]): for example, the emission can be attributed to a band-to-band recombination or to an impurity, if the corresponding excitation spectrum coincides with the fundamental absorption or with the impurity absorption band, respectively, see **Figure 5**. The photoluminescence of a phosphor can be described comprehensively by a 3D plot in which the emission intensity is plotted versus λ_{em} and λ_{exc}. Furthermore, the analysis of the emitted light polarization supplies useful information on the symmetry of the luminescent centers.

Interaction with the Lattice Vibrations

The interaction with the lattice vibrations (phonons) is responsible for the Stokes shift and broadening of the emission (or absorption/excitation) lines related to a luminescent center in a solid, see **Figure 5**. By increasing the temperature, the lines broaden (homogeneous broadening) and weaken as a rule. The lattice disorder due to defects (e.g., dislocations, impurities) is responsible for broadening even at low temperatures (inhomogeneous broadening). This is particularly severe in glasses and amorphous solids, characterized by a disordered network. Such effects are practically absent for some color centers (e.g., R′ and M′ in LiF) and impurities (e.g., rare earths, see **Figure 10**) in some insulating crystals: very narrow lines, or zero-phonon (ZP) lines, are observed in emission spectra measured at low temperatures. This occurs when the minima of the electronic excited and ground states fall at the same configurational coordinate in **Figure 4**. A very narrow ZP line may be accompanied by a weaker, more or less structured spectrum on the lower energy side, which corresponds to electron transitions to vibrational excited states of the electronic ground state (vibronic spectrum). In other words, the electron de-excitation occurs by emitting a photon and subsequently a phonon. A similar process may take place with the emission of a photon and of a localized vibrational quantum that is related to the luminescent center itself.

Time Dependence and Lifetimes

The radiative emission from a single excited level decays as soon as the excitation is turned off, that is, the time evolution of N_e is given by

$$N_e(t) = N_e^0 \exp(-t/\tau) \qquad [4]$$

where N_e^0 is the number of emitted photons per unit time at $t = 0$ and τ is a time constant or the excited state lifetime. It does not necessarily coincide with the radiative lifetime τ_r, that is, the time the electron spends in C for an isolated center (see **Figure 4**) before decaying, but may be shortened by a number of processes. In terms of a simple model, τ may be written as

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_0} \exp \frac{-\Delta E}{kT} + \frac{1}{\tau_Q} \qquad [5]$$

where $(1/\tau_0)\exp(-\Delta E/kT)$ is the probability, at temperature T, that the electron in the excited state (e.g., C in **Figure 4**) is brought by thermal activation into a higher level from which the decay is nonradiative, ΔE is the activation energy, and $1/\tau_Q$ is the probability of any other process. Shortening of τ with respect to τ_r causes also a decrease of the luminescence quantum yield η. This is defined either as the ratio of the number $N_{L,r}$ of centers which decay by a radiative pathway to the total number N_L or as the ratio of τ and τ_r, that is,

$$\eta = \frac{N_{L,r}}{N_L} = \frac{N_{L,r}}{N_{L,r} + N_{L,nr}} = \frac{\tau}{\tau_r} \qquad [6]$$

$N_{L,nr}$ accounts for the number of centers which decay through nonradiative pathways, for example, as a consequence of heat emission or absorption, intersystem crossing, capture in metastable levels, quenching caused by other centers, and energy transfer, see **Figures 6** and **7**. Therefore, the lifetime τ measurement supplies valuable information on a variety of processes occurring in the excited state, see **Figure 8**. Since the pioneering study performed by A E Becquerel with his phosphoroscope at the end of the nineteenth century, the time analysis has been extended nowadays to times as short as 10^{-12}–10^{-15} s. The availability of femtosecond excitation pulses allows one to also watch the fast processes by which the excited electron gives up part of its energy

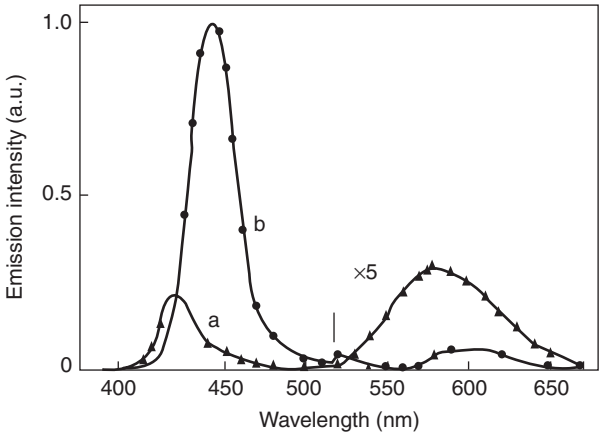


Figure 6 Example of energy transfer between Eu^{2+} (the donor) and Mn^{2+} (the acceptor) in a single crystal of KCl: Eu (54 ppm), Mn (3 ppm). The sample (the same as in **Figure 5**) was submitted to annealing at 225°C for 360 h. The treatment causes the formation of mixed Eu–Mn clusters, strongly reducing the separation between Eu^{2+} and Mn^{2+}. The emission spectra were measured at room temperature (curve a) and at 78 K (curve b). As a consequence of the excitation at 275 nm, i.e., in the Eu absorption band (see **Figure 5**), two emission bands are observed. The former in the 420–450 nm range (in blue, as in **Figure 5**) is related to the Eu emission, while the latter in the 550–620 nm range (in red, absent in **Figure 5**) is related to Mn emission, proving the energy transfer from Eu to Mn. The energy transfer is more efficient at room temperature than at 78 K, in fact the red Mn emission is stronger than the blue Eu emission. For the luminescence color, see **Figure 7**. (Reproduced with permission from Capelletti R, Manfredi M, Cywiński R, Damm JZ, Mugeński M, and Solzi M (1987) Energy-transfer mechanism in the KCl: Eu^{2+}, Mn^{2+} system. *Physical Review B* 36: 5124–5130.)

to the lattice (i.e., from B to C and from D to A in **Figure 4**). Emission spectra (see above) can be collected for different time windows along the decay and displayed as time-resolved emission spectra (TRES) in a 3D plot, which quickly reveals complexities in a sample. For example, a mixture of two luminescent centers with different lifetimes will exhibit a distinct difference in the TRES as the time window is varied.

Concentration Effects and Energy Transfer

For application purposes, one would increase the emission intensity by increasing the number N_L (or the concentration c_N) of the luminescent centers, according to eqn [3]. However, by increasing c_N the average distance among centers decreases, enhancing the energy transfer probability, and, as a consequence, $1/\tau_Q$ in eqn [5]. This causes a lifetime shortening, and, according to eqn [6], an η decrease, which might make the c_N increase ineffective, if not detrimental (concentration luminescence quenching). $1/\tau_Q$ may also increase as a result of clustering processes: the energy transfer between neighboring luminescent centers (e.g., Eu and Mn) in a cluster allows the tuning of the luminescence color, see **Figures 3e**, and **6–8**. The energy transfer to molecular groups (e.g., OH, H_2O, O_2), which decay without emitting light may cause the decrease or even the total suppression of the luminescence (impurity luminescence quenching).

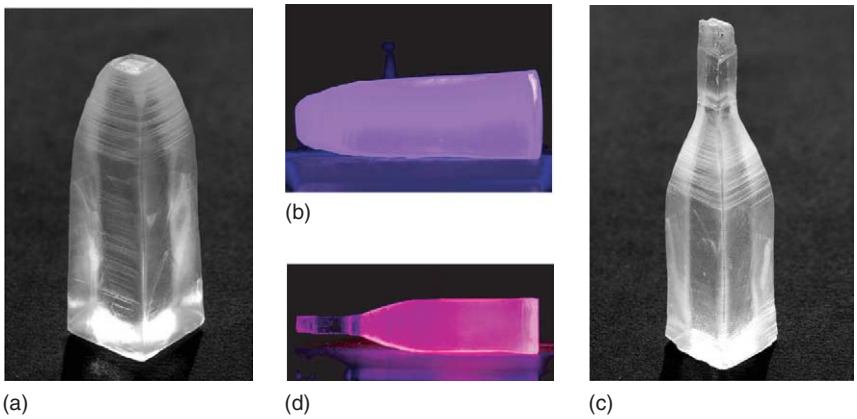


Figure 7 Photoluminescence excited by 253 nm UV light in Eu and Mn double-doped alkali halide single crystals grown by the Czochralski method and energy transfer between Eu and Mn. (a) and (b) are related to KCl: Eu (500 ppm), Mn (4000 ppm) (size: $2.2 \times 2.2 \times 6.4\,cm^3$); (c) and (d) to NaCl: Eu (100 ppm), Mn (2000 ppm) (size: $1.9 \times 1.9 \times 6.4\,cm^3$); (a) and (c) daylight illumination; (b) and (d) the samples are illuminated only by UV light. In agreement with the emission spectrum shown in **Figure 6**, the Eu→Mn energy transfer is proven by the presence of the Mn red emission which adds to the Eu blue one displayed in **Figure 1** for a KCl: Eu sample. The Eu→Mn energy transfer is more efficient (higher red emission contribution) in NaCl than in KCl due to its smaller lattice constant, which makes possible shorter distance between Eu and Mn. In (d) the luminescence is emitted from the body of the crystal which is Eu and Mn doped, but not by the seed (left end) which is undoped NaCl. The crystals were grown at Physics Department, University of Parma, Italy. (Photographs by Carlo Gardini, Parma (carlgard@inwind.it).)

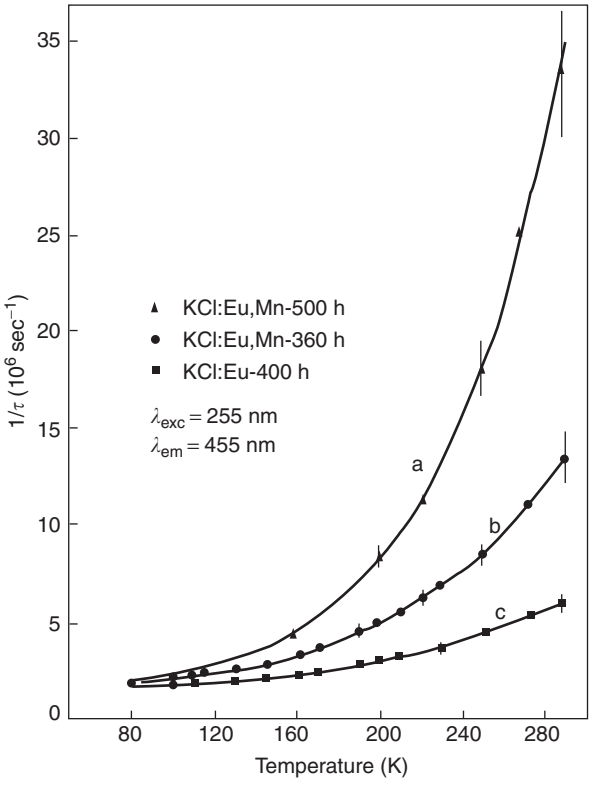


Figure 8 Eu luminescence time decay rates $1/\tau$ as a function of the temperature in a single crystal of KCl: Eu (54 ppm), Mn (3 ppm) at different steps of the clustering process. Curve (b) is related to a 360 h annealing (as in **Figure 6**) while curve (a) to a 500 h one (more advanced Eu–Mn clustering process). Curve (c) is related to a KCl:Eu (90 ppm) annealed for 400 h. For all the samples the annealing temperature is 225°C, the excitation and emission wavelengths are $\lambda_{exc} = 255$ nm and $\lambda_{em} = 419$ nm, respectively. According to eqn [6], the decay rate increases by increasing the temperature (for all the samples) and as the energy transfer from Eu to Mn takes place (compare curves (a) and (b) with curve (c)). The effect is more pronounced at a more advanced clustering step (compare curve (a) with curve (b)). (Reproduced with permission from Capelletti R, Manfredi M, Cywiński R, Damm JZ, Mugeński M, and Solzi M (1987) Energy-transfer mechanism in the KCl: Eu^{2+}, Mn^{2+} system. *Physical Review B* 36: 5124–5130.)

Size Effects

Thanks to modern technology, nanoparticles, either free standing or embedded in a host matrix, are available on a nanometer scale. Their size plays a key role in the excitation and emission spectra; an example is provided by the excitonic luminescence (see **Figure 3b**) in a semiconductor (e.g., CdSe). The particle size may be smaller than the orbit radius R ($\sim$10 nm) of the bulk exciton, regarded as a hydrogen-like system composed of a bound electron–hole pair. The exciton confinement within the nanoparticle increases the excitation and emission energies, as for a particle in a box. The exciton emission shifts from the deep red, to orange, yellow, green, and blue by gradually decreasing the nanoparticle size to $\sim$2 nm, see **Figure 9**. The effect can be exploited to tailor-tunable lasers.

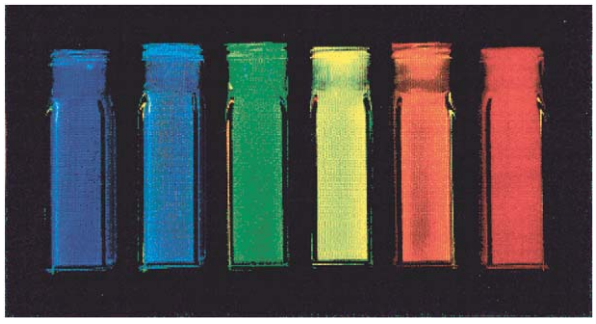


Figure 9 Color photograph demonstrating the wide spectral range of bright photoluminescence (PL) from different size samples of ZnS overcoated CdSe quantum dots, dispersed in dilute hexane solutions placed in identical quartz cuvettes and irradiated with 365 nm UV light. The sizes of the CdSe core increases from the left to the right. Their PL peaks occur at (going from left to right) 470, 480, 520, 560, 594, and 620 nm. (Photograph by F. Frankel) (From Dabbousi BO, Rodriguez-Viejo J, Mikulec FV, Heine JR, Mattoussi H, Ober R, Jensen KF, and Bawendi MG (1997) (CdSe)ZnS core-shell quantum dots: synthesis and characterization of a size series of highly luminescent nanocrystallites. *Journal of Physical Chemistry B* 101(46): 9463–9475.)

Two-Photon Processes and Anti-Stokes Emission

As a rule, the light emission from a center is excited by a single photon. With the availability of high-power light sources, two photons can be absorbed simultaneously and excite the luminescence. In the energy scheme, there is no intermediate level which can be reached by one of the two photons. The probability of such an event is very low and scales with the square of the exciting light intensity, at variance with the one-photon (OP) excited emission, which scales linearly with the incident light intensity, see eqn [3]. The information obtained from the two-photon (TP) excited luminescence spectroscopy is complementary to that obtained from the conventional OP one, because it allows one to study transitions which are forbidden if induced by an OP process, see **Figure 10**.

Generally, there is a large interest in the up-conversion processes, where more than one photon of lower energy is absorbed to cause the emission of a higher-energy photon, for example, to convert IR photons into visible ones. Two simple mechanisms are exemplified for centers with suitable level schemes. (1) In a single center the electron is enhanced by a photon to an excited state, while being there, it absorbs a second photon (excited state absorption, or ESA), reaching a still higher level, then it decays to the ground state by emitting a high-energy photon. (2) Two centers, at rather close distance, are simultaneously excited, then

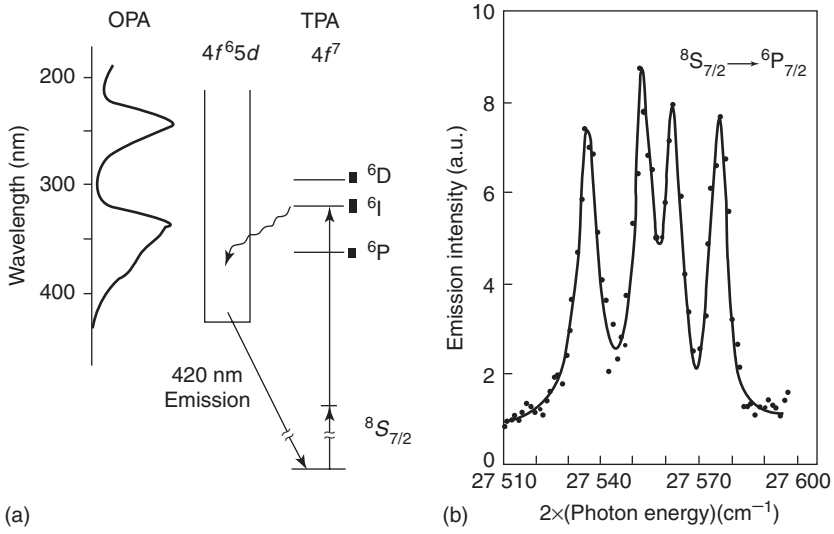


Figure 10 TP excited luminescence in a KCl: Eu single crystal. (a) Comparison between the OP and TP excitation. By the OP excitation the photon induces the allowed electron transition from the ground $4f^7$ to the excited $4f^6$ $5d$ state of Eu^{2+}. The decay to the ground state is accompanied by the blue luminescence (see **Figure 1**) at 420 nm, the emission spectrum of which is portrayed in **Figure 5**. The related excitation spectrum coincides with the OP absorption (OPA) displayed on the left side and in **Figure 5**. By the TP excitation, transitions within the $4f^7$ configuration are allowed, e.g., between the ground $^8S_{7/2}$ and excited ^{6}I manifold: on the left, the two consecutive vertical arrows indicate the TP absorption (TPA). The excited electron relaxes to the $4f^6$ $5d$ state, then decays to the ground state, by emitting the luminescence at 420 nm as above, but the excitation spectrum is different, as shown in (b). (b) TP excitation spectrum ($\lambda_{em} = 420$ nm) excited by the $^8S_{7/2} \rightarrow {}^6P_{7/2}$ transition. At variance with the broad OP excitation spectrum of **Figure 5**, it is characterized by very narrow lines, covering only 1 nm (from 363.5 to 362.5 nm). The four lines are meaningful of $^6P_{7/2}$ manifold splitting induced by cubic crystal field probed by Eu^{2+} in KCl. (From Casalboni M, Francini R, Grassano UM, and Pizzoferrato R (1986) Two-photon spectroscopy in KCl: Eu^{2+}. *Physical Review B* 34: 2936–2938.)

the former de-excites transferring its excitation to the latter, which is raised to a higher level and subsequently decays to the ground state by emitting a high-energy photon: this process is an energy transfer up-conversion. The two processes can be distinguished by analyzing the luminescence decay times (see section on "Time dependence and lifetimes").

The anti-Stokes emission occurs when a substance is excited by a photon of a given energy and emits one of slightly higher energy: seemingly, the Stokes' law is contradicted and more energy is radiated than is absorbed. After absorbing a photon, the center returns to equilibrium by absorbing phonons (heat) from the environment and then it decays to the ground state by emitting a photon, whose additional energy is subtracted from the lattice vibrations. The consequent cooling may be exploited for prospective interesting applications.

Down-conversion processes (or quantum cutting) should also be considered. A UV photon (with an energy higher than twice the energy of a visible photon) excites a center and the subsequent de-excitation takes place, for example, through a cascade emission of two visible photons. In this way, a quantum yield η, in terms of the ratio of the emitted visible photon number to that of the absorbed ones, may be >1, see eqn [6]. The process may be exploited to tailor UV-excited phosphors with enhanced emission in the visible range, for lighting applications.

Thermoluminescence

Thermoluminescence (TL) is light emission caused by moderate heating of a solid, previously exposed to ionizing radiation. It is displayed at temperatures at which the blackbody light emission is negligible. The exposure to the ionizing radiation (X- or γ-rays, high-energy particles) takes place at rather low temperatures, creates traps, and fills them (or those already present) with electrons and/or holes. The subsequent heating makes the electrons and/or holes free to migrate in the lattice, until they fall into other traps (or recombine) with a consequent photon emission (**Figure 3d**). A comprehensive description of TL is provided by 3D spectra, in which the emitted light intensity is plotted versus temperature T and emitted wavelength λ_{em}, the heating rate being, as a rule, constant, see **Figure 11**. From the analysis of such plots, the energy depth of traps can be evaluated. For many purposes, the simpler plot of the total emitted light intensity I_{TL} (regardless of the wavelength) as a function of T supplies valuable information. Materials which display one or more TL peaks which are linear functions of the absorbed radiation dose can

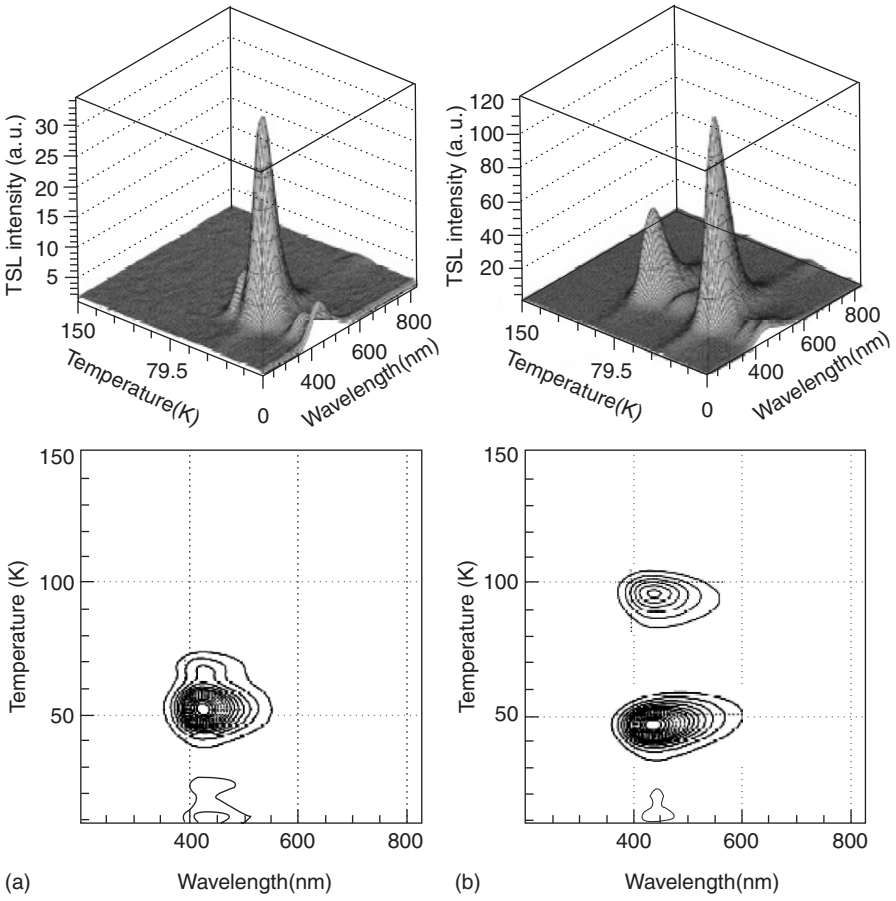


Figure 11 Example of wavelength-resolved TL of a scintillator. 3D TL measurement and contour plot of the TL in the 10–150 K range, following X-ray irradiation at 10 K of undoped (a) and La^{3+} doped $PbWO_4$ single crystal. The high density and fast radioluminescent (scintillation) response make it interesting for application in high-energy physics experiments. The comparison between (a) and (b) plots shows that the La^{3+} doping increases the emitted light intensity and introduces at least an additional TL peak. (From Martini M, Meinardi F, Spinolo G, Vedda A, Nikl M, Usuki Y (1999) Shallow traps in PWO_4 studied by means of wavelength-resolved thermally stimulated luminescence. *Physical Review B* 60: 4653–4658.)

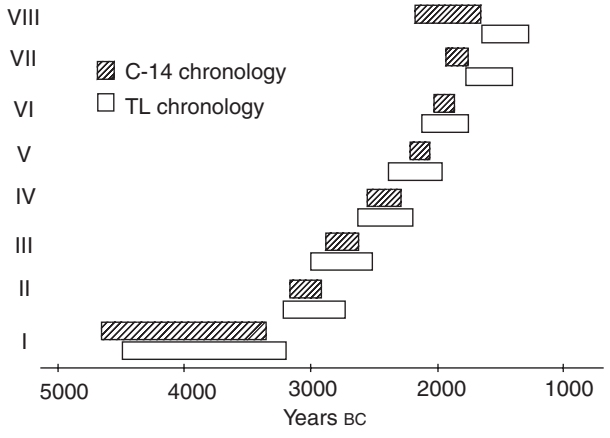


Figure 12 Example of TL application for dating. Comparison between independent absolute TL and radiocarbon (^{14}C) chronologies for dating ceramics of Valdivia culture (Ecuador). (From Martini M and Sibilia E (2001) Radiation in archeometry: archeological dating. *Radiation Physics and Chemistry* 61: 241–246.)

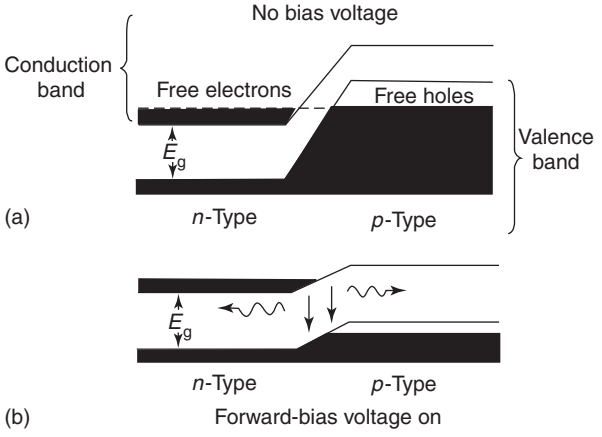


Figure 13 The EL process in an LED. A *p*–*n* junction is shown. (a) The junction is not voltage biased; (b) the junction is forward biased: the energy levels bend and, as a consequence, in the junction some free electrons in the CB recombine with free holes in the VB giving up the energy as light (EL). The mechanism is the same as in **Figure 3a**.

be exploited as dosimeters. If a calibration curve of I_{TL} versus dose is known, it is easy to evaluate the dose absorbed by a dosimeter that is, for example, worn by an individual exposed to radiation, by reading the TL emission spectrum. TL also provides a tool to date buried materials of archeological (e.g., pottery) or geological interest (e.g., stalagmites). In this case, the radiation dose is supplied by the decay of radioactive isotopes (e.g., ^{40}K, Th, and U) present in the soil. By estimating the dose/year supplied by the surrounding soil (paleodose), it is possible to evaluate the time elapsed from the last firing of a pottery item and detect possible counterfeiting. With very few exceptions, the consistency of the TL results with other absolute dating methods is found to be good, see **Figure 12**.

Electroluminescence

The application of an electric field to an insulator or semiconductor can cause the injection of electrons and/or holes. The recombination between electrons

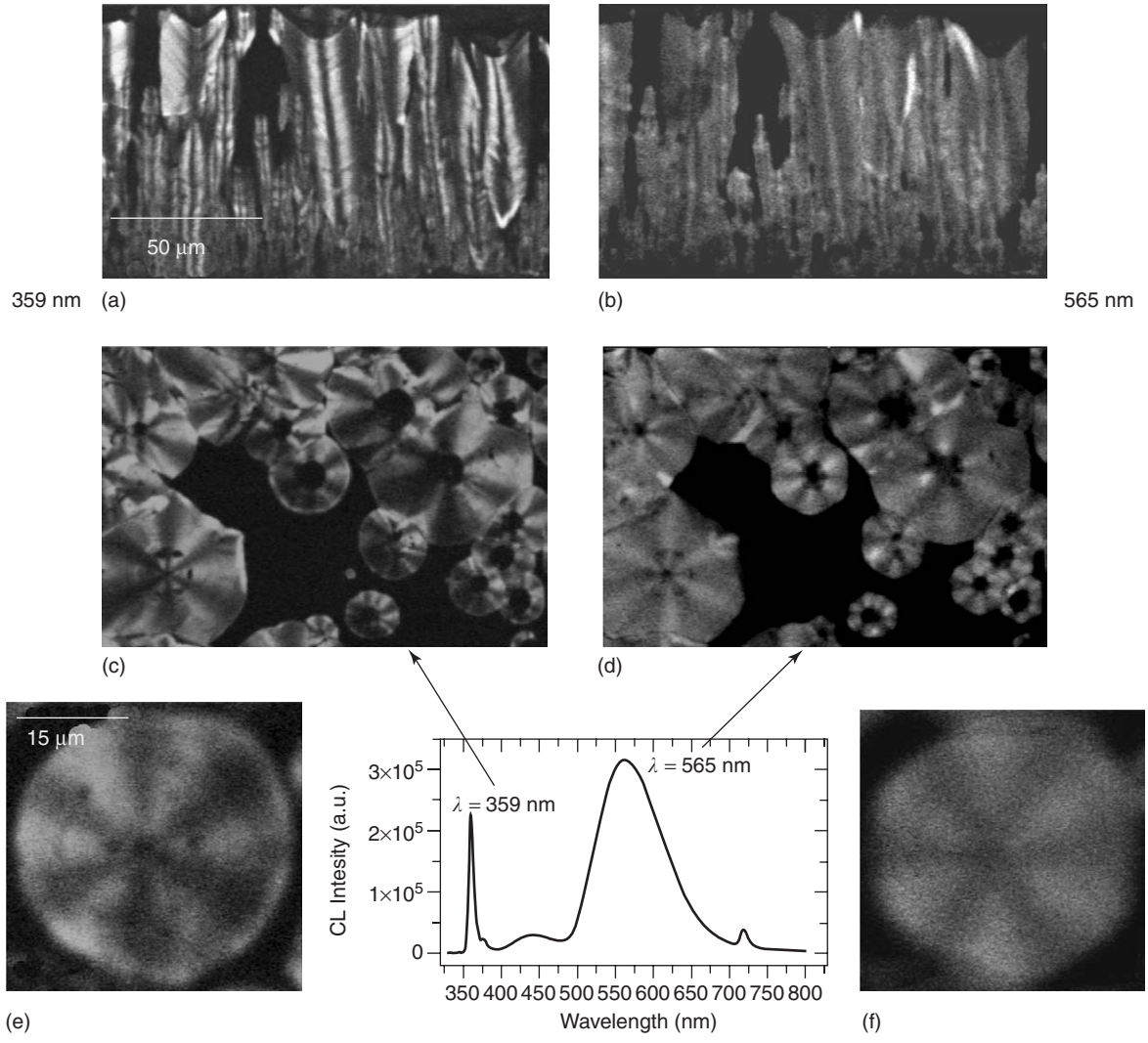


Figure 14 Example of CL excited by an SEM electron beam at 5 K in a bulk GaN sample. The emission is monitored at 359 nm (near UV) in (a), (c), and (e) images and at 565 nm (yellow) in (b), (d), and (f) images. The bright regions in both cases are those in which the related CL emission is strong. (a) and (b) are the monochromatic (at 359 and 565 nm, respectively) CL images of the same columnar structures in the sample. (c) and (d) are the CL images of the cross sections of the columnar structures, related to (a) and (b), respectively. (e) and (f) supply the magnification of one of the hexagonal structures (the same) displayed by (c) and (d). Note that the dark and bright regions are complementary. The CL spectrum, i.e., the CL intensity vs. emission wavelength λ_{em}, is displayed between the (e) and (f) images. The two main peaks at 359 and 565 nm are attributed to GaN near band edge (NBE) emission and to Si-related defects, respectively. This suggests that Si-related defects segregate along preferential directions. (From Salviati G, Albrecht M, Zanotti-Fregonara C, Armani N, Mayer M, Guzzi M, Melnik YuV, Vassilevski K, Dmitriev VA, and Strunk HP (1999) Cathodoluminescence and transmission electron microscopy study of the influence of crystal defects on optical transition in GaN. *Physics Status Solidi* (*a*) 171: 325–339.)

and holes may be accompanied by light emission, which is named electroluminescence (EL). ZnS was reported to be the first example of an electroluminescent material in 1936 by G Destriau, working in Madam Curie's laboratory. The widespread applications of LEDs are based on EL of a *p–n* junction. The charge carriers are the electrons in a CB for the *n*-type material, and holes in a VB for the *p*-type. When a forward-bias voltage is applied to the diode, the energy levels bend, see **Figure 13**: the concentration of electrons in a CB near the junction on the *n*-side increases as that of holes in the VB near the junction on the *p*-side. The electrons and holes recombine by emitting photons, whose energy is equal to that associated with the energy gap E_g. Modern semiconductor device fabrication techniques allow bandgap engineering, and junctions can be produced from semiconductors or their alloys in which E_g corresponds to almost any wavelength in the visible light spectrum, as for example, GaAs and GaN LEDs that emit in the near IR and in the blue, respectively. The junction can also work as an active medium for diode lasers (e.g., laser pointers) as a feedback is provided by the reflective ends of the laser material and most of the light is produced by stimulated emission. The use of heterojunctions improves the light confinement.

Cathodoluminescence

Fast electrons (the J J Thomson cathode rays) supply energy enough to excite a phosphor and induce light emission as cathodoluminescence (CL). The color-TV screen operation is based on CL. The electrons emitted by the electron gun in the cathode-ray tube strike little triads of different phosphor dots, which emit blue, green, and red CL. The phosphors exploit the light emission from suitable impurities, for example, Cu and Ag in ZnS (blue and green) and Eu in Y_2O_2S (red). The CL induced by the electron beam striking the sample under study in an electron microscope is used nowadays to investigate its physical properties, as for example, the impurity mapping (see **Figure 14**).

See also: Fluorescent Proteins; Defect Centers in Insulating Solids, Optical Properties of; Laser Radiation Sources; Semiconductor Nanostructures; Synchrotron Radiation; Time-Resolved Optical Spectroscopies.

PACS: 78.55. – m; 78.60, – b; 78.47. + p; 78.67. – n; 78.40. – q

Further Reading

Aitken MJ (1985) *Thermoluminescence Dating*. Oxford: Academic Press.

Aitken MJ (1990) *Science-Based Dating in Archeology*. London: Longman.

Bassani F and Pastori PG (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Bersanti M and Maccarone F (1991) Crystalloluminescence. *Rivista del Nuovo Cimento* 14: 1–67.

Birks JW (ed.) (1989) *Chemiluminescence and Photochemical Reaction: Detection in Chromatography*. New York: VCH.

Blasse G and Grabmaier BC (1994) *Luminescent Materials*. Berlin: Springer.

Chen R and McKeever SWS (1997) *Theory of Thermoluminescence and Related Phenomena*. Singapore: World Scientific.

Harvey EN (1957) *A History of Luminescence*. Philadelphia: American Philosophical Society.

http://web.mit.edu/felicef

http://www.cerncourier.com/main/article/39/2/12/1/

McKeever WS (1985) *Thermoluminescence of Solids*, Cambridge Solid State Science Series. Cambridge: Cambridge University Press.

Nakamura S and Fasol G (1997) *The Blue Laser Diode: GaN Based Light Emitters and Lasers*. Berlin: Springer.

Nakamura S, Pearton S, *et al.* (2000) *The Blue Laser Diode: The Complete Story*. Berlin: Springer.

Rodnyi PA (1997) *Physical Processes in Inorganic Scintillators*. New York: CRC Press.

M

Magnetic Domains

O Portmann, A Vaterlaus, C Stamm, and D Pescia, Laboratorium für Festkörperphysik der ETH Zürich, Zürich, Switzerland

Introduction

The Dirac equation for the relativistic quantum mechanical electron is most famous for having predicted the existence of antiparticles. One further important aspect of this equation is the explanation of the electron spin and the magnetic moment associated with it. When atoms are formed, the electron spins interact by means of the intra-atomic exchange interaction to produce atomic magnetic moments. The correct value of the atomic magnetic moments can be calculated empirically using Hund's rules and Pauli's principle. When atoms are brought together to form a body, the atomic wave functions change drastically: the atomic magnetic moments are, in general, reduced, but may remain finite in certain special cases such as Fe, Co, and Ni. Moreover, as soon as neighboring atoms are provided, the interatomic exchange interaction comes into play and may favor the parallel alignment of the neighboring magnetic moments. In this way, macroscopic regions of aligned magnetic moments are formed in the body, and the body is said to be in a ferromagnetic state.

This microscopic description of magnetism, based on the quantum mechanical exchange interaction, is not complete. Each individual magnetic moment can be viewed as an atomic current generating a magnetic field, which interacts with other magnetic moments and contributes another term to the total magnetic energy – the energy of the dipole–dipole interaction. This term is also called the magnetostatic energy, as it is described by the Maxwell equations of magnetostatics. In contrast to the exchange interaction, this term is purely classical. It is later shown that, depending for example on the shape and size of the body, this term perturbs the parallel alignment of the magnetic moments favored by the exchange interaction. Thus the parallel alignment does not extend over the whole body, but is restricted to elementary regions called magnetic domains. Within such domains, the magnetic moments are aligned, while the orientation of the magnetic moments changes upon moving to neighboring domains.

Another class of magnetic interactions arises because of the spin–orbit coupling also contained in the Dirac equation: magnetic anisotropies. A magnetic anisotropy energy determines one or more spatial directions along which the exchange-coupled magnetic moments within the domains prefer to be aligned, where "preferring" means that the total energy of the system is reduced when the moments pointf along that direction in space. Magnetic anisotropies play a crucial role in domain formation. In fact, the change of the magnetic moment direction between consecutive domains does not occur abruptly: a domain wall with finite width connects domains with different magnetic moment orientation and magnetic anisotropies stabilize this wall.

The purpose of this article is to account for the state-of-the-art knowledge on magnetic domains and the walls between them. Before going into details, it is important to estimate the energy scale of the various magnetic interactions, given by their typical strength per atom. The intra-atomic exchange interaction is of the order of eV, on a temperature scale this corresponds to 10^4 K. This energy is required to produce an excited state with no atomic magnetic moment. The interatomic exchange interaction is some 10 meV (100 K). This means that the excited state with misaligned nearest-neighbor magnetic moments is attained when the temperature exceeds some 100 K (at a temperature which is called the Curie temperature). Magnetic anisotropies range from μeV in three-dimensional solids to 0.1 meV in ultrathin films (0.01 to 1 K). The reason for this broad range is that the magnetic anisotropies are strongest when the symmetry of the system is lowered, for example, when the dimensionality is reduced. The magneto-static interaction is $\sim$0.1 meV (1 K). Clearly, the interactions relevant for domain formation are much smaller than the exchange interaction. Thus, it is not immediately evident why such small energies should have an impact in determining the total energy of the system and its ground-state magnetic moment configuration. This question is also carefully investigated in the remaining sections.

The Physical Basis of Magnetic Domains

The physical basis for the explanation of magnetic domains was laid down in a paper by Landau and Lifshitz in 1935. One can consider a body in a situation where each lattice site carries a magnetic moment vector (for simplicity, a spin vector is used, noticing that, according to Dirac's theory, the two vectors are proportional to each other). One important caveat: it will turn out that the relevant length scales intervening in domain formation are large with respect to the lattice constant. On these characteristic length scales, the number of spins is so large that one can give up the quantum description of the spin vector and consider it a classical vector. The strongest contribution to the total magnetic energy comes from the exchange interaction, which measures the energy required to turn two neighboring spins into an antiparallel configuration. If this energy is positive, the total exchange energy is minimized by a uniform spin configuration. The spin ensemble can also have a magnetic anisotropy energy contribution, which favors certain spatial directions for the spins, the so-called easy direction. In the simplest case of uniaxial anisotropy, two directions in space, say $\boldsymbol{e}$ and its negative, are energetically favored. Thus, the uniform spin configuration will be aligned either along $+\boldsymbol{e}$ or along $-\boldsymbol{e}$, as either of these directions minimizes the total energy arising from the exchange and the magnetic anisotropy. How stable is the resulting spin configuration with respect to rotations or spin nonuniformities? The magnetic anisotropy per spin is very small but it is not possible to rotate spins individually, as they are bound together by the exchange interaction. Thus, a rotation excitation would involve turning a macroscopically large amount of spins and this requires a large excitation energy which is not available at sufficiently low temperatures. If the body is divided into oppositely magnetized regions, it would not entail any lowering of the total energy; on the contrary, some energy must be introduced to form boundaries between oppositely magnetized regions, causing an increase of the total energy. A more complicated anisotropy term may have a larger symmetry, but in no case will the magnetic anisotropy lead to magnetic domains as there is no energy to be gained upon changing from one easy direction to another. The behavior is quite different when the dipolar interaction between the magnetic moments is considered. Each moment produces a magnetic field, which interacts with all the other magnetic moments by means of the Zeeman energy contribution. The most striking feature of this magnetostatic energy is that the magnetic fields arising from a magnetic moment are very long ranged, they decay only with the third power of the distance. This is in opposition to exchange and magnetic anisotropies, which are local. The energy due to the dipolar field is the most complicated (see Appendix A), but contains a distinct feature which makes it compete directly with the exchange interaction: one part of it favors an antiparallel alignment of the spins. Of course, the interaction is very weak, but its long-range character amplifies its role with respect to the competing exchange interaction. Thus, under certain circumstances, which Landau and Lifshitz showed to depend on the size and the geometry of the body, the dipolar interaction can provide enough energy gain to sustain a nonuniform spin configuration: magnetic domains arise.

The complexity of the dipolar field and the long-range character of the dipolar interaction makes it very difficult to find the spin configuration that minimizes the total energy. However, there are some rules of thumb that help in analyzing the physical situation. One of them is provided by eqn [3] in Appendix A. This equation has a simple meaning: the total magnetostatic energy of a body is the Coulomb energy of an effective magnetic charge distribution $\rho_M \doteq -\boldsymbol{\nabla} \cdot \boldsymbol{M}(\boldsymbol{x})$, where $\boldsymbol{M}(\boldsymbol{x})$ is the magnetic moment per unit volume. Its length is determined by the total magnetic moment in a sufficiently small cell centered at $\boldsymbol{x}$ divided by the volume of the cell, and its direction is given by the direction of the total magnetic moments within this cell. The size of this cell has to be chosen in such a way that it is small with respect to the smallest length occurring in domain formation, that is the size of the domain wall (see next section). The effective magnetic charge is provided by the divergence of $\boldsymbol{M}(\boldsymbol{x})$, thus minimizing the magnetostatic energy means avoiding regions where $-\boldsymbol{\nabla} \cdot \boldsymbol{M}(\boldsymbol{x})$ is too large. A sizeable divergence of $\boldsymbol{M}(\boldsymbol{x})$ exists, for instance, at the boundaries of a uniformly magnetized body. Such surface charges increase the energy of the uniform spin configuration above the value specified by the exchange interaction. To avoid such charges, the body can split into elementary regions of different magnetization. At the boundaries between such regions, the spin vectors make a finite angle so that the exchange energy is increased (at the walls) with respect to the uniform configuration. Despite this, Landau and Lifshitz have explicitly shown that by introducing domains, the total energy might effectively be lowered. In Appendix B, a simple case is worked out where the subtleties of such total-energy minimization calculations are shown.

At the end of this section, an example illustrating this "magnetic-charges" avoidance rule is discussed.

One can compare a small ferromagnetic plaquette with square geometry and a uniform spin configuration (see **Figure 1a**) to an identical plaquette with a particular nonuniform spin configuration where the magnetization vector circulates along a closed path within the plaquette (**Figure 1b**). The spin configuration in **Figure 1a** contains some effective magnetic charges appearing at the boundaries perpendicular to the magnetization vector. The spin configuration in **Figure 1b** is a very peculiar one, the magnetization vectors between two consecutive domains are orthogonal and no divergence develops at the boundary between two consecutive domains. The spin configuration in **Figure 1b** is free of magnetic charges and has no magnetostatic energy. One can show that when the plaquette is sufficiently large, the gain in magnetostatic energy overcomes the cost of introducing domain walls: the configuration in **Figure 1b** is the ground state. Experimentally, this closed path configuration is indeed observed, see **Figure 1c** as an illustration. The butterfly shape of the plaquette in **Figure 1c** is more complicated than the square plaquette just considered. Correspondingly, the experimental domain distribution (indicated by arrows), is more complex. It is also evident from **Figure 1c** that similar triangular plaquettes – the wings in the figure – have a slightly different domain distribution, showing that unresolved differences play a crucial role in determining the actual domain configuration. However, one essential element dominates the central parts of both wings: the domains are arranged in such a way that the magnetization vector circulates within the plane of the plaquette, avoiding the formation of magnetic charges, in agreement with the rule of thumb.

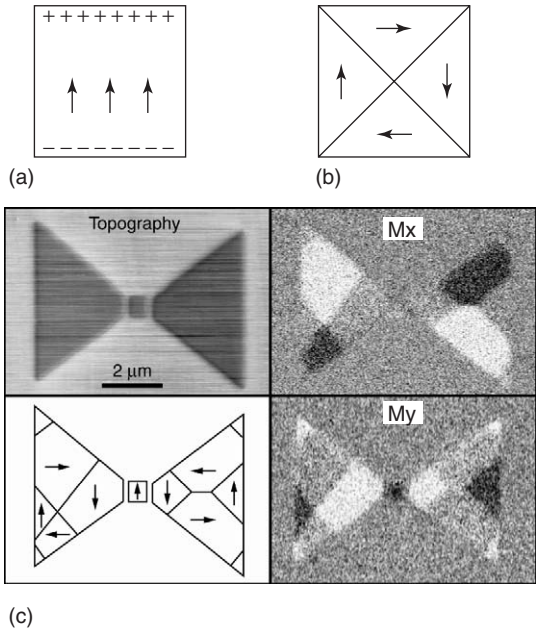


Figure 1 (a) and (b): Schematic domain distribution in a plaquette. In (a) the uniform magnetization produces magnetic surface charges, indicated in the figure. In (b) the magnetization vector follows a closed path, thus avoiding magnetic charges. (c) Image of the domain distribution in a plaquette with butterfly shape. The images were taken using scanning electron microscopy with polarization analysis (SEMPA). In SEMPA, a finely focused electron beam impinges onto the surface of a solid. Secondary electrons are ejected and used to map the topography of the surface while the beam is scanned across the surface. In the present case a $\sim$14 nm thick Co film was deposited onto a Cu(1 0 0) surface by molecular beam epitaxy (MBE) through a mask inserted between MBE source and substrate. In this mask, a butterfly-profile was carved in by focused ion beam milling, which gives rise to the laterally patterned Co film visible in the topography image. In SEMPA, not only the secondary electron current is measured, but their spin polarization is also detected. This allows mapping the surface magnetization vector. The in-plane horizontal (M_x) and vertical (M_y) components of $\boldsymbol{M}$ are reproduced in (c) (the component perpendicular to the film plane vanishes). The black-white constrast is indicative of a strong magnetization in opposite directions. The Cu-substrate is nonmagnetic and appears gray. From the contrast maps of the two in-plane components, the domain distribution can be reconstructed as schematically summarized by the arrows.

Domain Walls

Although it is the magnetostatic energy contribution that drives the rearrangement of the spin configuration into domains, it is the magnetic anisotropy that stabilizes the walls between them and makes the whole rearrangement process possible. For the sake of simplicity, the creation of a domain wall in a one-dimensional spin chain is illustrated. One can consider a chain extending along the x direction with fixed-length spin vectors $S(x)$ being allowed to rotate in the y–z plane. Thus, $S = S(0, \sin\theta(x), \cos\theta(x))$, S being the (dimensionless) length of the spin vector and θ the angle between the vector S and the y direction. The variable x is allowed to assume continuous values which is equivalent to say that, if any deviation from the parallel spin alignment occurs, it occurs over spatial length scales much larger than the lattice constant a. Assuming an exchange coupling between nearest neighbors of the simple form $-J \cdot S(x) \cdot S(x+a)$ ($J>0$ being the exchange constant), any deviation from the parallel spin alignment can be worked out to increase the exchange energy by an amount $\Delta_\Gamma = \{(\Gamma \cdot a)/2 \cdot \int \mathrm{d}x[\mathrm{d}\theta(x)/\mathrm{d}x]^2\}$ with the coupling constant $\Gamma = (J/2)S^2 z$ (z being the number of nearest neighbors, that is, 2 for a chain). A magnetic anisotropy term which favors the y direction for the spin is $\Delta_\lambda = (\lambda/(2a)) \int \mathrm{d}x \sin^2\theta(x)$ with the coupling constant λ measuring the energy required to turn the spin away from the y direction. Notice that, in accordance with the above discussion,

$\Gamma \gg \lambda$. The equilibrium $\theta(x)$ is found by setting the functional variation of $\Delta_\Gamma + \Delta_\lambda$ to zero, which produces the Euler equation of the variational problem:

$$\Gamma a \frac{d^2\theta}{dx^2} - \frac{\lambda}{a} \sin\theta \cos\theta = 0 \qquad [1]$$

One can study the stability of a wall by solving this equation with the following boundary conditions: $\theta = \pi$ (0) for $x = -\infty$ ($+\infty$), and $d\theta/dx = 0$ for $x = \pm\infty$. This establishes a spin profile with domains of opposite spins on the left- and right-hand side of $x = 0$. The solution to the Euler equation with these boundary conditions exists and was found in the work of Landau and Lifshitz:

$$\cos\theta = \tanh\frac{2x}{w} \qquad [2]$$

$w = 2a \cdot \sqrt{\Gamma/\lambda}$ being the width of the region near $x = 0$ over which the spin rotates from one domain to the domain of opposite magnetization, that is, the domain wall width. This solution is plotted in **Figure 2b** as a continuous curve. The wall energy (i.e., the energy required to establish a wall) per unit wall area can be calculated analytically by inserting the solution into the total energy expression: $\sigma_w = (1/(a^{\delta-1}))\sqrt{\lambda \cdot \Gamma}$. The prefactor contains the dimensionality of the system δ. For the interaction strengths discussed in the introduction, w is expected to be of the order of some 100 lattice constants, which is a mesoscopic scale. A direct measurement of the spin profile inside the domain wall at the surface of an Fe single crystal indeed gives ≈ 200 nm. As an example, one can consider the domain image of a head-on magnetization distribution recorded in a thin epitaxial Fe film on $W(1\,1\,0)$ (**Figure 2a**). The spatial resolution of the image is such that the continuous in-plane rotation of the magnetization vector from one direction into the opposite direction occurring at the boundary between the two domains can be detected in detail. The rotation is illustrated with a color code explained in the caption. The experimental profile of the magnetization component that changes sign, given by the circles in **Figure 2b**, follows the calculated profile (continuous curve in **Figure 2b**), obtained by fitting the experimental data with tanh $(2x/w)$.

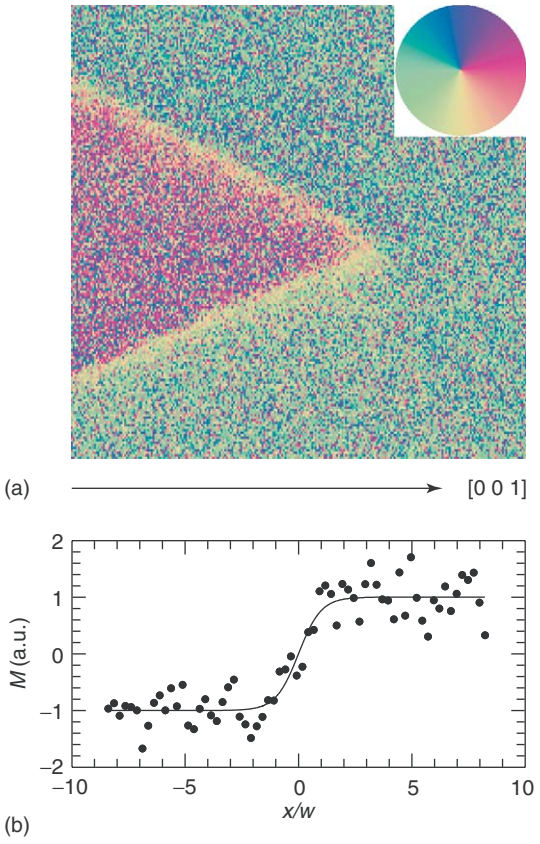


Figure 2 (a) Domain wall formation between head-on magnetized domains. The two domains, represented by different colors, are magnetized in opposite directions parallel to the horizontal axis (the [0 0 1]-direction of the cubic lattice, in this case). At the boundary between them, one recognizes that the magnetization vector rotates in the plane from one direction into the opposite one. A color code is used to represent the direction of the magnetization vector. Image size: $4.6 \times 4.6\,\mu m^2$ (b) The horizontal component of the magnetization vector (divided by its value well within the domains) is plotted along an axis x, drawn perpendicular to the domain wall (full circles). At the domain boundary, the value of the component makes a transition that can be fitted by the calculated tanh - profile (continuous curve), with a characteristic wall width $w \approx 140$ nm. To improve the statistical accuracy of the experimental profile, all experimental data points within a stripe going across the wall were averaged.

An Example of Magnetic Domains: Stripes

Having established the physical basis for domain formation, the details can be worked out, which will help to achieve a deeper understanding of the energy balance which leads to domain formation. As anticipated, magnetostatic charges have to be considered, which means that the geometry and size of the body are central to the problem. For one special geometry, the appearance of domains is illustrated (for many other geometries which have been investigated so far, see the "Further reading" section). The special case of ultrathin magnetic film extending to infinity in the x_1–x_2 plane and having a finite but small thickness d in the x_3 direction (**Figure 3a**) is worked out in detail. The magnetization vector is taken to be perpendicular to the film plane. The perpendicular magnetization in ultrathin films is a very special configuration, requiring a strong uniaxial

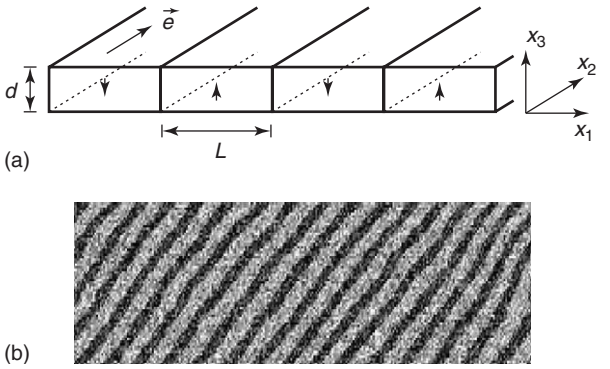


Figure 3 (a) Schematic view of an ultrathin film (thickness d) with a stripe configuration (stripe width: L) of alternating perpendicular magnetization. The geometrical elements used in Appendix B are also indicated. (b) Room temperature SEMPA image of the stripe phase in an ultrathin ($d \approx 0.3$ nm) Fe film deposited on Cu(1 0 0) by means of MBE. The image is 92.2 μm long. The component of the magnetization vector perpendicular to the film plane is shown. The black-white contrast is indicative of oppositely magnetized stripes.

magnetic anisotropy to overcome the demagnetizing effect of the dipolar interaction. This demagnetizing effect is best appreciated by comparing a uniform in-plane spin configuration with a uniform perpendicular spin configuration. The in-plane configuration has no magnetostatic energy, provided the film extends to infinity. The perpendicular configuration, on the other hand, produces opposite charges at the two boundaries of the film with vacuum. This gives rise to an electrical plane condenser geometry. Using eqn [4] in Appendix A, the magnetostatic energy (per unit area) can easily be calculated to be $2\pi M_0^2 d$ and must be overcome by a strong magnetic anisotropy favoring the perpendicular configuration in order for a perpendicular magnetization to appear.

The possibility of the existence of a nonuniform perpendicular spin configuration that might lower the energy of the system with respect to the uniform configuration in the given perpendicular magnetic configuration is examined. A possible nonuniform spin configuration with perpendicular magnetization is the one consisting of stripes of opposite perpendicular magnetization running parallel to say x_2 and repeating periodically along x_1, see **Figure 3a**. Such a configuration does certainly not introduce any extra magnetic anisotropy energy. Domain walls, however, are created and cause an increase of the exchange energy. The striped phase can only become energetically more favorable if the gain in magnetostatic energy is large enough to overcome the extra domain wall energy (this is shown to happen exactly in Appendix B). The stripes have an equilibrium width $L^* \propto w \cdot e^{\pi L_0/2d}$. L_0 is the so-called dipolar length $L_0 \doteq \sigma_w/(2\pi M_0^2)$, and is related to the characteristic dipolar energy per atom $(2\pi M_0^2 \cdot a^3)$. Using the characteristic values as discussed in the introduction, one obtains L_0 of the order of $100a$. Thus, in the ultrathin limit $d \approx a$, the equilibrium stripe width L^* can reach astronomical values at $T = 0$: the ground state will appear to be a uniform single-domain spin configuration because the stripe width might be much larger than the size of the sample. However, when the temperature is increased, the various coupling constants renormalize and the stripe phase can penetrate the sample as its width reduces to mesoscopic lengths. This situation is indeed observed to occur in several perpendicularly magnetized ultrathin films where the stripe phase has been observed to develop. An example is given in **Figure 3b** for Fe films on Cu(1 0 0).

Acknowledgments

The authors thank the Swiss National Fund and ETH Zürich for financial help. They are grateful to G M Graf for making his calculation of the equilibrium stripe width (Appendix B) available for this article.

Appendix A: Magnetostatic Energy of a Localized Distribution $M(x)$

A useful starting point for finding the magnetostatic energy of a continuous static distribution $\boldsymbol{M}(\boldsymbol{x})$, are the Maxwell equations of magnetostatics: $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ and $\boldsymbol{\nabla} \times (\boldsymbol{B} - 4\pi \boldsymbol{M}) = 0$. The magnetic field $\boldsymbol{H}$ defined as $\boldsymbol{H} \doteq \boldsymbol{B} - 4\pi \boldsymbol{M}$ can be written in terms of a magnetic potential, $\Phi_M : \boldsymbol{H} \doteq -\boldsymbol{\nabla}\Phi_M$. The first Maxwell equation now corresponds to the Poisson equation of an "effective magnetic charge" $\rho_M \doteq \boldsymbol{\nabla} \cdot \boldsymbol{M} : \boldsymbol{\nabla}^2 \Phi_M = -4\pi\rho_M$. Thus, the magnetostatic problem has been transformed into an electrostatic problem involving charges ρ_M. Their energy amounts to (Coulomb)

$$E_M = \frac{1}{2} \int \mathrm{d}^3x \, \mathrm{d}^3y \frac{\rho_M(\boldsymbol{x})\rho_M(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} \qquad [3]$$

This expression for E_M states that a possible way to minimize the magnetostatic energy is to avoid magnetic charges, that is to minimize the magnitude of $\rho_M \doteq -\boldsymbol{\nabla} \cdot \boldsymbol{M}$. A further useful expression for the energy involves the magnetic field $\boldsymbol{H}$:

$$\begin{aligned} E_M &= \frac{1}{2} \int \mathrm{d}^3x \, \mathrm{d}^3y \, \frac{\rho_M(\boldsymbol{x})\rho_M(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} = \frac{1}{2} \int \mathrm{d}^3x \rho_M(\boldsymbol{x})\Phi_M(\boldsymbol{x}) \\ &= -\frac{1}{8\pi} \int \mathrm{d}^3x \big(\nabla^2 \Phi_M(\boldsymbol{x})\big)\Phi_M(\boldsymbol{x}) \\ &= \frac{1}{8\pi} \int \mathrm{d}^3x \boldsymbol{H}^2(\boldsymbol{x}) \end{aligned} \qquad [4]$$

where the general solution of the Poisson equation $\Phi_M(\boldsymbol{x}) = \int \mathrm{d}^3y \rho_M(\boldsymbol{y})/|\boldsymbol{x}-\boldsymbol{y}|$ is used. According to eqn [4], minimizing E_M means searching for a spin configuration which avoids the formation of stray fields $\boldsymbol{H}$.

From eqn [3], an often quoted expression for E_M can be derived. After a partial integration, the surface integral vanishes provided the magnetization distribution is localized. Using $\nabla^2 1/|\boldsymbol{x}-\boldsymbol{y}| = -4\pi\delta(\boldsymbol{x}-\boldsymbol{y})$, one obtains

$$\begin{aligned} E_M &= -\frac{1}{2}\int \mathrm{d}^3x\mathrm{d}^3y \boldsymbol{M}(\boldsymbol{x})\cdot\boldsymbol{\nabla}_x \frac{\boldsymbol{\nabla}_y\cdot\boldsymbol{M}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|} \\ &= \frac{1}{2}\int \mathrm{d}^3x\mathrm{d}^3y \boldsymbol{M}(\boldsymbol{x})\cdot\boldsymbol{\nabla}_x\left(\boldsymbol{M}(\boldsymbol{y})\cdot\boldsymbol{\nabla}_y\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) \\ &= \frac{1}{2}\int \mathrm{d}^3x\mathrm{d}^3y \left(\frac{\boldsymbol{M}(\boldsymbol{x})\cdot\boldsymbol{M}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3} \right. \\ &\quad \left. -3\frac{[\boldsymbol{M}(\boldsymbol{x})\cdot(\boldsymbol{x}-\boldsymbol{y})][\boldsymbol{M}(\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y})]}{|\boldsymbol{x}-\boldsymbol{y}|^5}\right) \\ &\quad + \frac{2\pi}{3}\int \mathrm{d}x^3 \boldsymbol{M}(\boldsymbol{x})^2 \end{aligned} \qquad [5]$$

This equation states that the magnetic "dipole" $\boldsymbol{M}(\boldsymbol{x})$ interacts with a $\boldsymbol{H}$ field

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{x}) = \int \mathrm{d}y^3 \left(-\frac{\boldsymbol{M}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3}\right. \\ + 3\frac{[\boldsymbol{M}(\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y})](\boldsymbol{x}-\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^5} \\ \left. -\frac{4\pi}{3}\boldsymbol{M}(\boldsymbol{y})\delta(\boldsymbol{x}-\boldsymbol{y})\right) \end{aligned} \qquad [6]$$

produced at $\boldsymbol{x}$ by all other dipoles at $\boldsymbol{y}$ (the last term is the own field of the dipole at $\boldsymbol{x}$). This expression for E_M is often called the dipolar energy due to the dipole–dipole interaction.

A third useful equation can be derived from eqn [4] using the identity $\int \mathrm{d}^3x \boldsymbol{H}\cdot\boldsymbol{B} = 0$, holding for a localized distribution $\boldsymbol{M}(\boldsymbol{x})$:

$$E_M = 2\pi\int \mathrm{d}^3x\, \boldsymbol{M}^2(\boldsymbol{x}) - \frac{1}{8\pi}\int \mathrm{d}^3x\, \boldsymbol{B}^2 x \qquad [7]$$

Accordingly, if the length of the magnetization vector is a constant, a spin configuration with a low magnetostatic energy requires maximizing the magnetic field $\boldsymbol{B}$.

Using the same identity, one can immediately derive

$$E_M = -\frac{1}{2}\int \mathrm{d}^3x\, \boldsymbol{H}\cdot\boldsymbol{M} \qquad [8]$$

Finally, one can transform eqn [7] into the interaction energy of a system of effective currents $\boldsymbol{J}_M \doteq c\boldsymbol{\nabla}\times\boldsymbol{M}$. With $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$ and $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) = -\nabla^2\boldsymbol{A}$ (in the Coulomb gauge), one obtains $\boldsymbol{\nabla}^2\boldsymbol{A} = -4\pi/c\,\boldsymbol{J}_M$. Finally, from this Poisson equation for $\boldsymbol{A}$, one can obtain

$$E_M = 2\pi\int \mathrm{d}^3x\, \boldsymbol{M}^2(\boldsymbol{x}) - \frac{1}{2c^2}\int \mathrm{d}^3x\, \mathrm{d}^3y \frac{\boldsymbol{J}_M(\boldsymbol{x})\boldsymbol{J}_M(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|} \qquad [9]$$

Thus, minimizing the magnetostatic energy means searching for a spin configuration that maximizes the curl of $\boldsymbol{M}$.

Appendix B: The Equilibrium Stripe Width

It is very instructive for the reader to follow in detail the appearance of domains for a very specific case. One can consider perpendicularly magnetized ultrathin films of thickness d extending over a large area $\Lambda_1 \times \Lambda_2$ in the $x_1 - x_2$ plane, see **Figure 3a**. The magnetization within each stripe points along $\pm x_3$. The stripes of width L are taken to be parallel to x_2. They are joined by walls with a finite width ω with $d \ll w \ll L$. Consecutive stripes have opposite perpendicular magnetizations. The total wall energy E_w amounts to $\sigma_w(\Lambda_1/L)\Lambda_2 \cdot d$. This is just the energy per unit wall area multiplied by the length of the wall Λ_2, the film thickness d, and the number of walls within the length Λ_1. Here, the wall energy per unit film area $\sigma_w d/L$ is the main interest.

The lowering of the magnetostatic energy upon introducing stripes can immediately be seen when one considers eqn [9]. A uniform perpendicular spin configuration is curl free, so that there is no effective current J_M anywhere in space and the total energy per unit area is simply $(2\pi M_0^2 \cdot d)$. The stripe configuration, instead, develops some curl (and correspondingly some magnetic field $\boldsymbol{B}$), thus necessarily lowering the magnetostatic energy.

The following calculation shows explicitly that the energy gain is enough to overcome the wall energy introduced by the stripes. First, the effective currents are found. The stripe magnetization distribution develops a curl at the boundaries (labeled with the index n) between consecutive stripes. Along these boundaries, effective current densities $\boldsymbol{J}_n = cj_n\boldsymbol{e}$ occur where $j_n = (-1)^n 2M_0/w$ ($n = 0, \pm 1, \ldots, \pm(N-1)$) and $j_{\pm N} = (-1)^{\pm N} M_0/w$. One contribution to the Coulomb energy, eqn [9] arises from the interaction of each current with the remaining currents. There are Λ_1/L of such interaction terms and the calculation of

this part of the interaction energy reduces to the computation of the sum of elementary integrals of the type

$$\int_{-\Lambda_2/2}^{\Lambda_2/2}\int_{-\Lambda_2/2}^{\Lambda_2/2} dx_2\, dy_2 \frac{1}{\sqrt{(x_2-y_2)^2+L_n^2}} \qquad [10]$$

($L_n = \pm n$, L and $n \neq 0$). Notice that the integration along (x_3, y_3) can be approximated by multiplying the remaining two-dimensional integrals with d^2. This is possible because d is considered small with respect to L, so that one can set $x_3 = y_3 = 0$ in the integrand. The integration over (x_1, y_1) can be approximated by a factor w^2, because w is considered to be much smaller than L, allowing one to set $(x_1 - y_1)^2 = L_n^2$ in the integrand. Thus, eqn [9] reduces to the elementary integral of eqn [10]. After computing these integrals, one is left with a sum that can be solved exactly because it contains in the limit $N \to \infty$ the product of Wallis. The other contribution arises from the self energy of each current. Under the integral of the self energy, $(x_1 - y_1)$ may be smaller than $|x_2 - y_2|$. Thus, the integrations over x_1, y_1 must be performed explicitly. There are again Λ_1/L such terms and their self energy per unit area contains the integral

$$\int_0^w\int_0^w dx_1\, dy_1 \int_{-\Lambda_2/2}^{\Lambda_2/2}\int_{-\Lambda_2/2}^{\Lambda_2/2} dx_2 \times dy_2 \frac{1}{\sqrt{(x_2-y_2)^2+(x_1-y_1)^2}} \qquad [11]$$

Summarizing all contributions, the total energy per unit film area is

$$2\pi M_0^2 d \cdot \left\{1 - \left[\left(\frac{3}{\pi}+\frac{2}{\pi}\ln\frac{2}{\pi}\right) - \frac{2}{\pi}\ln\frac{w}{d}\right]\frac{d}{L} + \frac{2}{\pi}\frac{d}{L}\ln\frac{d}{L}\right\} + \sigma_w\frac{d}{L} \qquad [12]$$

The equilibrium stripe width L^* is found by minimizing the total energy per unit area with respect to the variable $x \doteq d/L$.

$$L^* = w \cdot e^{\pi(b+1)/2} \cdot e^{\pi L_0/2d} \qquad [13]$$

where the numerical factor $b = -(3/\pi + 2/\pi \ln 2/\pi)$. Notice that the equilibrium stripe width is a nonanalytic function of d for small d. It can be readily verified that the total energy of the striped phase at the equilibrium stripe width is indeed lower than the total energy of the uniform perpendicular state, by an amount $-(2/\pi)(2\pi M_0^2 d/L^*)$.

See also: Ferromagnetism; Magnetic Interactions.

PACS: 75.60.Ch; 75.70.Ak; 75.25.+z

Further Reading

Hubert A and Schäfer R (2000) *Magnetic Domains*. Berlin: Springer.
Jackson, J D *Classical Electrodynamics*, Third Edition, ch 5. New York: Wiley.
Kaplan B and Gehring GA (1993) *Journal of Magnetism and Magnetic Materials* 128: 111–116.
Rowlands G (1994) *Journal of Magnetism and Magnetic Materials* 136: 284–286.
Ter Haar D (ed.) (1967) *Collected Papers of L.D. Landau*, pp. 101–116. New York: Gordon and Breach.

Magnetic Interactions

G Hilscher and H Michor, Vienna Institute of Technology, Vienna, Austria

Introduction

Magnetism in matter is a purely quantum mechanical phenomenon according to the Bohr–van Leeuwens theorem: van Leeuwen demonstrated in 1918 that classical Boltzmann statistics applied rigorously to any dynamical system leads neither to a susceptibility nor to a magnetic moment, which means that classical mechanics cannot account for diamagnetism, and paramagnetism or any type of collective magnetic order. However, under the assumption that a molecule or atom has a finite magnetic moment $\boldsymbol{\mu}$, the diamagnetic and paramagnetic susceptibility was derived already in 1905 by Langevin using classical statistical thermodynamics and treating these moments without interactions. The assumption of a finite magnetic moment, however, requires quantized values for the spin S and angular momentum L that is a result of quantum mechanics. The coupling of spin and angular momentum to the total angular momentum J is a relativistic interaction, which can be treated as a perturbation for not too heavy elements (Russel Sounders or L–S coupling) and the ground state is

described by Hunds rules. Accordingly, magnetism arises from quantized motions of electrons giving rise to the fundamental unit of the magnetic moment, the Bohr magneton, defined by $\mu_B = eh/2m_e$ that takes the value 9.274×10^{-24} Am^2 or JT^{-1}. Thus, the magnetic moment of an atom or a molecule is given in units of μ_B (e.g., the magnetic moment of pure Fe in the metallic state is 2.2 μ_B per atom).

The magnetic properties of a system of electrons and nuclei are determined by the interactions of these particles with electromagnetic fields. These fields arise from two types of sources: first, there are fields due to sources external to the system such as applied electric and magnetic fields. Second, there are internal sources, which come from the charges and currents due to the particles of which the system is composed. Interactions with and between nuclear moments μ_N are not discussed here since μ_N is by a factor of $\sim$1800 smaller than μ_B.

A first step is to consider the magnetic moments associated with itinerant and/or localized electrons on the respective ions of a crystal to behave completely free without interacting with each other or with their surrounding. This leads to diamagnetism and paramagnetism.

In a free atom the quantum number J, the total angular momentum, is a good quantum number, which reflects the fact that a free atom has a complete rotational symmetry in the absence of external fields and its ground state is $2J + 1$ fold degenerate. In a crystal, however, the surrounding ions produce an electric field to which the charge of the central atom adjusts and the $2J + 1$ degeneracy is removed giving rise to a splitting of the ground multiplet. The object of the crystalline electric field (CEF) theory is to describe how the moments are affected by their local environment in a crystal. Finally, one has to consider the various interactions of magnetic moments with each other leading eventually to a long-range magnetic order.

Crystalline Electric Field Interactions

The charge distribution around a central ion produces an electric field with the local symmetry of the environment. This crystal field makes a contribution to the potential energy

$$V_C(\boldsymbol{r}) = \int \frac{e\rho(\boldsymbol{R})}{|\boldsymbol{r} - \boldsymbol{R}|}\,\mathrm{d}\boldsymbol{R} \approx e\sum_i \frac{q_i}{|\boldsymbol{r} - \boldsymbol{R}_i|}$$

where the charge distribution $\rho(\boldsymbol{R})$ in the point charge model is approximated by point charges q_i at the position $\boldsymbol{R}_i$. The denominator may be expanded in spherical harmonics. If V_C is small compared to the spin–orbit coupling (as in the rare-earth elements where the incomplete 4f shell is responsible for the magnetic moment), the eigenstates are adequately accounted for by the first-order perturbation theory. Provided that the ground state and the next excited states do not mix, the matrix elements of $V_C(\boldsymbol{r})$ are proportional to those obtained with operator equivalents written in terms of $\boldsymbol{J}$, which are called the Stevens operator equivalents $O_l^m(\boldsymbol{J}_i)$ and the CEF Hamiltonian reads

$$\hat{H}_{CF} = \sum_i \sum_{lm} B_l^m O_l^m(\boldsymbol{J}_i)$$

The CEF parameters B_l^m can, in principle, be calculated from the charge distribution of the environment, which yields reasonable results for insulators but is of limited success for metals.

The electrostatic interactions destroy the rotational symmetry of the ion and the electronic orbitals are linked to the symmetry of the lattice removing the $2J + 1$ degeneracy of the ground multiplet of the free ion. Schematically illustrated in **Figure 1** is the CEF splitting (e.g., Mn in the perovskite $LaMnO_3$) resulting from the interaction between the nonspherical d-orbitals and the electrostatic field of the environment, which is also nonspherical. As an example, the d_{xy}-orbital in an octahedral surrounding (a) is energetically favored in comparison to the $d_{x^2-y^2}$ orbital (b) in the same environment. Those lobes, which point to neighboring charges, have higher overlap and correspond to higher electrostatic energy than those that point between. This gives rise to the CEF splitting displayed in **Figure 1c**.

An improvement on the point charge approximation for d-transition metal ions is the ligand field theory which is an extension of the molecular orbital theory that focuses on the role of the d-orbitals of the central ion and their overlap with orbitals on surrounding ions (ligands). Also for rare-earth compounds, CEF parameters can be calculated from first principles using the density-functional theory which, however, should be checked experimentally as, for example, by inelastic neutron scattering.

It is instructive to consider two rather different cases, namely the 3d-transition metal series (Fe group) and the 4f series (rare earth), where the relative importance of the spin–orbit coupling and CEF splitting of the ground multiplet is inverted. The spin–orbit coupling which is proportional to Z^2 (atomic number) is much larger in the 4f series than in the 3d-transition metals since rare-earth elements are significantly heavier. Furthermore, the spatial extension of the 3d wave functions is much more delocalized than the 4f wave functions that are

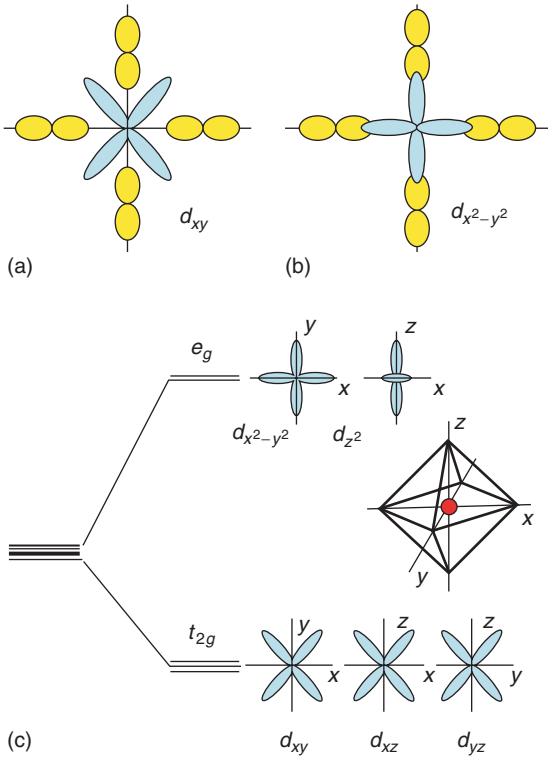


Figure 1 (a) The d_{xy}-orbital is lowered in energy with respect to the $d_{x^2-y^2}$-orbital in an octahedral environment. (b) Schematic crystal field splitting in an octahedral environment (c) The d_{z^2}-and $d_{x^2-y^2}$-orbitals are grouped together and called e_g levels (twofold degenerate). The d_{xy}, d_{xz} and d_{yz} are grouped together and called the t_{2g} levels (threefold degenerate).

additionally screened from CEF by the outer 6*s*- and 5*d*-shells. Thus, in rare-earth compounds CEF splitting is $\sim 10^2$ K and spin–orbit coupling of 10^4 K, which means that the Hunds rule ground state is usually observed both in metals and insulators and the effect of CEF interaction can be treated in terms of Stevens operators. In 3*d*-compounds, V_C is typically of the order of 10^4 K which is larger than the spin–orbit coupling (10^2 K) and V_C couples mainly to the orbital part of the wave function which lifts the orbital degeneracy of the 3*d*-states and Hunds rule ground state is usually not obeyed. If the CEF perturbation is strong enough and the symmetry low enough, the orbital degeneracy is completely removed. Thus, the orbital ground state is a singlet which is called the quenching of the orbital momentum (L is not anymore a good quantum number and virtually reduced to $L = 0$) and "spin-only" magnetic properties are observed. In 4*d*-and 5*d*-series (Pd and Pt groups), the situation is less clear-cut because the heavier ions have a larger spin–orbit splitting, and the CEF and spin–orbit interaction can be of comparable magnitude.

The impact of CEF interaction on the magnetic properties is important and in combination with spin–orbit interaction it is the principal source of magnetic anisotropy, which is the dependence of magnetic properties on the direction in the compound. Further, properties influenced by CEF are magnetoelastic phenomena and transport phenomena as well as thermodynamic properties as the specific heat.

Origin of Long-Range Magnetic Order – Exchange Interactions

In the previous section, magnetic ions embedded in the CEF environment of the crystal were considered but the interaction between the magnetic moments was neglected. The classical interaction between two magnetic dipoles $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ at a distance $\boldsymbol{r}$ between each other is known as the dipole–dipole interaction. As the magnetic energy is $E_M = -\boldsymbol{\mu}\cdot\boldsymbol{H}$, the dipolar energy reads

$$E_D = \frac{\mu_0}{4\pi r^3}\left[\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2 - \frac{3}{r^2}(\boldsymbol{\mu}_1\cdot\boldsymbol{r})(\boldsymbol{\mu}_2\cdot\boldsymbol{r})\right]$$

Thus E_D/k_B is of the order of 1 K, if one considers two moments of $1\,\mu_B$ which are separated by 1 Å. Accordingly, this interaction cannot account for Curie temperatures of 1000 K, but nevertheless can become of relevance at low and ultra-low temperatures.

Dirac and Heisenberg discovered independently that electrons are coupled by exchange interactions among which direct, indirect, itinerant exchange, and superexchange can be distinguished. Direct, indirect, and itinerant exchange occurs in metallic systems. Direct and superexchange is important in nonmetallic systems.

Origin of Exchange

In order to show how the Pauli exclusion principle together with electrostatic interactions leads to magnetic effects, consider the simplest possible case – a two-electron system – where the Hamiltonian

$$\hat{H} = \sum_{i=1}^{2}\frac{p_i^2}{2m} + V(\boldsymbol{r}_1, \boldsymbol{r}_2)$$

does not depend on the spin. In quantum mechanics one cannot distinguish between identical particles, which means that the particle coordinates q_1 and q_2 can be permuted (exchanged) in the wave function and the system remains the same physically; the new wave function differs from the original one

by a phase factor: $\psi(q_1, q_2) = e^{i\phi}\psi(q_2, q_1)$. Permuting once more gives the original wave function. Thus $e^{2i\phi} = 1$ and $\psi(q_1, q_2) = \pm\psi(q_2, q_1)$, which means the wave function must be either totally symmetric or antisymmetric under permutation. Electrons (Fermions) have totally antisymmetric wave functions; so the spin part must either be an antisymmetric singlet state $\chi_S(S = 0)$ in the case of a symmetric spatial wave function or a symmetric spin triplet state $\chi_T(S = 1)$ in the case of an antisymmetric spatial state. The overall wave function for the singlet ψ_S and the triplet state ψ_T for two electrons where both the orbital and spin part is included reads:

$$\psi_S = \frac{1}{\sqrt{2}}\chi_S[\psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2) + \psi_1(\boldsymbol{r}_2)\psi_2(\boldsymbol{r}_1)],$$

$$\text{with } \chi_S = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

$$\psi_T = \frac{1}{\sqrt{2}}\chi_T[\psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2) - \psi_1(\boldsymbol{r}_2)\psi_2(\boldsymbol{r}_1)],$$

$$\text{with } \chi_T : |\uparrow\uparrow\rangle;\ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle);\ |\downarrow\downarrow\rangle$$

If the interaction $V = e^2/r_{12}$ is absent, then these two states would have the same energy. For $V \neq 0$, the energies of the triplet and the singlet state are different

$$E_S = \int \Psi_S^* \hat{H} \Psi_S = U + J$$

$$E_T = \int \Psi_T^* \hat{H} \Psi_T = U - J$$

with

$$U = \int\int |\psi_1(\boldsymbol{r}_1)|^2 V(\boldsymbol{r}_1, \boldsymbol{r}_2) |\psi_1(\boldsymbol{r}_2)|^2 \, \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2$$

and $\Delta E = E_S - E_T = 2J$

$$J = \int\int \psi_1(\boldsymbol{r}_1)\psi_1^*(\boldsymbol{r}_2) V(\boldsymbol{r}_1, \boldsymbol{r}_2) \psi_2(\boldsymbol{r}_2)\psi_2^*(\boldsymbol{r}_1) \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2$$

where J is the exchange interaction or exchange integral, and U the Coulomb integral. The original Hamiltonian does not depend on the spin but produces an energetically favorable spin configuration for parallel spins. Thus $\hat{H}$ may now be replaced by an effective one $\hat{H}^{\text{spin}}$ acting upon the spin components only with the requirement that it should produce the same $\Delta E = E_S - E_T$. What will be the form of $\hat{H}^{\text{spin}}$?

For the two-electron system, the spin operators yield the following eigenvalues:

$$\hat{S}_i^2 = \hbar^2 S_i(S_i + 1) = \hbar^2 \tfrac{3}{4} \quad \text{and}$$

$$\hat{S}^2 = \hbar^2 S(S + 1) = \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{S}_1 \cdot \hat{S}_2 = \hbar^2 \tfrac{3}{2} + 2\hat{S}_1^2 \cdot \hat{S}_2^2$$

Accordingly, the product $\hat{S}_1^2 \cdot \hat{S}_2^2/\hbar^2 = S(S + 1)/2 - 3/4$ of both operators gives $-3/4$ for the singlet $S = 0$ and $1/4$ for triplet state $S = 1$. Hence the original spin-independent Hamiltonian can be written in the form of an effective spin Hamiltonian yielding the same eigenvalues E_S and E_T:

$$\hat{H}^{\text{spin}} = \frac{1}{4}(E_S + 3E_T) - \frac{(E_S - E_T)\hat{S}_1 \cdot \hat{S}_2}{\hbar^2}$$

$$\hat{H}^{\text{spin}} = U - J\hat{S}_1 \cdot \hat{S}_2$$

with $U = 1/4(E_S + 3E_T)$ and $J = (E_S - E_T)/\hbar^2$. If $J > 0$, $E_S - E_T > 0$ and the "parallel" coupling of the spins in the triplet state with $S = 1$ is favored, while for $J < 0$ the "antiparallel" coupling in the singlet state is favored.

The generalization of this simple relation for the two-electron system to a many-body system is far from trivial. The result is the Heisenberg Hamiltonian

$$\hat{H}_{\text{Heisenberg}} = -\sum_{i>j} J_{ij}\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j$$

where J_{ij} is the exchange interaction between the ith and jth spins. Frequently, J_{ij} is taken as a constant for the nearest neighbors (and next nearest neighbors) and zero otherwise. J_{ij} is of the order of 10^2–10^3 K for direct exchange in $3d$-compounds, 10–10^2 K for indirect exchange and superexchange. A reliable calculation of J_{ij} is a rather complex task in general.

A simple but straightforward approach to obtain a convenient expression for the exchange interaction is the mean field approximation where the exchange interaction between the ith spin and its neighbors $\sum_j J_{ij}\hat{\boldsymbol{S}}_j$ is approximated by a phenomenological molecular field H_m acting upon $\boldsymbol{S}_i$. In the Weiss molecular field model, the molecular field is set proportional to the magnetization $H_m = \lambda M$, where the molecular field constant λ is determined by the exchange interaction yielding fields $\mu_0 H_m$ of the order of 1000 T for $3d$-intermetallics such as Fe–Co alloys to account for their Curie temperatures of $\sim$1000 K ($\boldsymbol{\mu H} = k_B T_c$). Nevertheless, this phenomenological molecular field is a convenient fiction but not experienced by the electrons since exchange origins from Coulomb interaction.

Some general features of exchange interactions are noteworthy. If two electrons on an isolated atom are considered, J is usually positive and favors the parallel spin configuration (triplet state) associated with an antisymmetric orbital state which reduces the Coulomb repulsion by keeping them apart. This is consistent with Hund's first rule yielding maximum total spin under the constraint of the Pauli principle.

The situation is different for electrons on neighboring ions in molecules or in a solid. In contrast to the free electron description for a simple metal, the model of tightly bound electrons is used as an approximation not only for ionic solids but also for transition metal compounds. The total wave function is a linear combination of atomic orbitals (LCAO) centered on the respective ions (which has to be in overall antisymmetric realized by a Slater determinant.) The electrons can save kinetic energy by forming bonds because their common wave function is more extended than the original atomic orbital and comprises both ions. These molecular orbitals can be bonding (spatially symmetric) or antibonding (spatially antisymmetric) as shown schematically in **Figure 2a** for the hydrogen molecule. The latter are more localized associated with higher kinetic energy. This favors the "antiparallel" singlet state as observed in H_2.

In a d-transition metal, the atomic $5N$ d-states for spin up and down give rise to a band of $10N$ levels distributed quasicontinuously between the lower bonding and the upper antibonding levels yielding together with transfer integrals and hopping integrals, the bandwidth W (**Figure 2b**). One has to distinguish between interatomic terms of the exchange and Coulomb integrals (J_{ij} and U_{ij}, $i \neq j$) and intra-atomic terms J_{ii} and U_{ii}. According to the more or less localized character of the wave function, the interatomic terms are much weaker than the intra-atomic terms. The U terms tend to avoid two electrons with antiparallel spins in the same orbital and the exchange terms J favor electrons in different orbitals with parallel spins. The latter are about one order of magnitude smaller. Thus, one may define an average exchange energy $\bar{I}$ per pair of electrons per atom, which represents the interaction favoring "ferromagnetic" alignment. On the other hand, the transfer integrals associated with the bandwidth work in the opposite direction, since the creation of parallel spins implies the occupation of higher energies in the band (i.e., with higher kinetic energy). Following this very simplified approach, the balance between kinetic and exchange energy favors the appearance of a spontaneous magnetic order for the end of the $3d$-elements (Cr, Mn, Fe, Co, and Ni).

A proper description is performed in terms of the density-functional theory where the exchange correlation potential is added to the Coulomb correlation, which allows a straightforward determination of the exchange correlation. This treatment is called the local spin density approximation (LSDA).

To summarize, exchange can be explained in the following way: the Pauli exclusion principle keeps the electrons with parallel spins apart and so reduces their Coulomb repulsion. The difference in energy between parallel and antiparallel spin configuration is the exchange energy. This, however, is favorable for ferromagnetism if the increase in kinetic energy associated with parallel spins does not outweigh the gain in potential energy, as in the familiar examples of the H_2 molecule or the He atom. In metallic Fe,

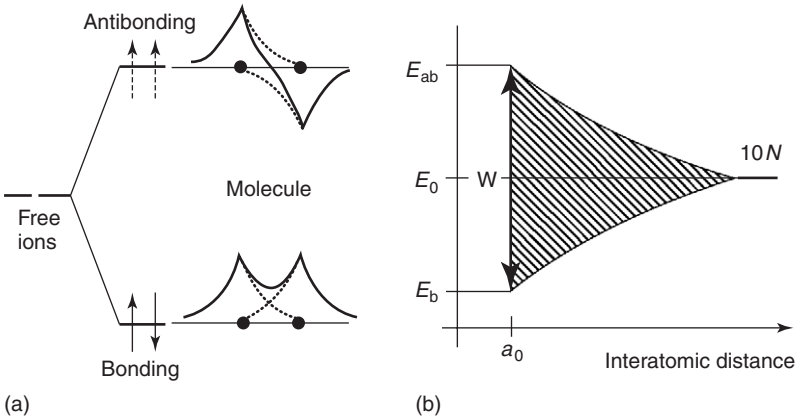


Figure 2 Schematical energy splitting of bonding (spatially symmetric) and antibonding (spatially antisymmetric) molecular orbitals in a molecule (a) yielding a singlet ground state as, e.g., in H_2 molecule. (b) Schematical formation of the bandwidth W in a solid by the quasicontinuous distribution of the $10N$ d-states between the lower and upper bonding states of the originally discrete and degenerated $5d$-atomic states for up and down spin of N transition metal atoms condensed to a crystal; E_b is the energy of the bonding state and E_{ab} of the antibonding state.

Co, and Ni, the number of parallel spins produces ferromagnetism because the cost in kinetic energy is not as great as in some other metals, where the gain in potential energy is significant yielding a paramagnetic ground state.

Direct Exchange

If there is sufficient overlap of the relevant wave functions (e.g., 3*d*-orbitals) between neighboring atoms in the lattice, the exchange interaction is called direct exchange. Although Fe, Co, and Ni are frequently referred to as typical examples for direct exchange, the situation is not that simple because of the metallic state where magnetism of itinerant electrons plays an important role. Accordingly, both the localized and band character of the electrons involved in transition metals and their intermetallics have to be taken into account as is seen in the model of covalent magnetism for transition metal intermetallics.

Indirect Exchange in Insulators – Superexchange

Superexchange is typical for insulators or materials on the border of a metal insulator transition where the ions with a magnetic moment are surrounded by ions which do not carry a moment but mediate the exchange. This type of exchange is observed in transition metal oxides. An archetypical example is MnO where two cations (Mn^{2+} in the $3d^5$ state with $S = 5/2$) are coupled via an anion (O^{2-} in the $2p^6$ state with $S = 0$); the ground state is shown schematically in **Figure 3a**. Because of the overlap of the wave functions, one of the *p*-electrons of the anion (O^{2-}) hops over to one of the cations (Mn^{2+}). The remaining unpaired *p*-electron on the anion enters into a direct exchange with the other cation. In principle, this exchange may be ferromagnetic or antiferromagnetic. The excited state shown in **Figure 3b** for an antiferromagnetic coupling is in most cases energetically in favor due to the interplay of two effects: the hopping of electrons between ligands, characterized by a hopping matrix element t being proportional to the bandwidth, and an average Coulomb interaction U between electrons on the same orbital. Superexchange involves oxygen *p*-orbitals and transition metal *d*-orbitals and is therefore a second-order process. In second-order perturbation theory, the exchange interaction is given by $J \propto -t^2/U$, and normally yields antiferromagnetic coupling.

Double Exchange

Double exchange is essentially a ferromagnetic type of superexchange in mixed-valent systems. Mixed-valency means that the transition metal ion on

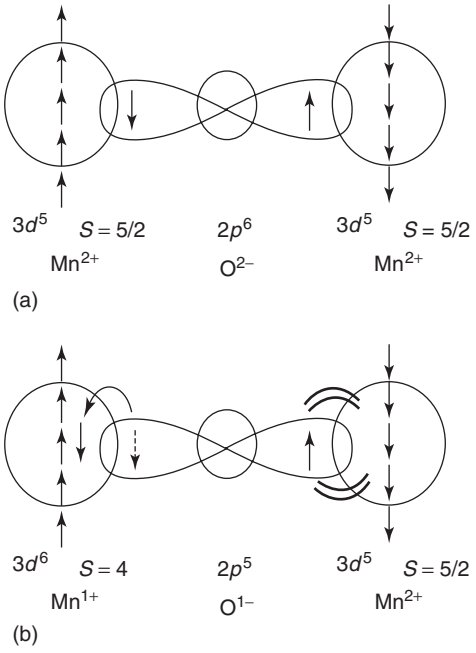


Figure 3 (a) Superexchange of a transition metal oxide as, e.g, MnO in the ground state, (b) hopping of electrons in the excited state favors antiferromagnetic coupling.

crystallographically equivalent lattice sites exhibits different ionic states, for instance, trivalent and tetravalent Mn in perovskites (La,Sr)MnO_3 or Fe^{2+} and Fe^{3+} in magnetite (Fe_2O_3). Due to the mixed valency, holes (i.e., unoccupied states) exist in the e_g states of the transition metal ion with a higher valency with respect to a cation of the same species with a lower valency. The simultaneous hopping of an electron from $Mn^{3+} \rightarrow O^{2-}$ and concomitantly from $O^{2-} \rightarrow Mn^{4+}$ is known as double or Zehner exchange. Hopping is favored if the neighboring ion is ferromagnetically aligned since kinetic energy is saved by delocalization, while it is hindered by antiferromagnetic alignment as shown schematically in **Figure 4**.

Indirect Exchange in Metals – RKKY Interaction

In transition metal alloys, intermetallic compounds, and diluted magnetic alloys where magnetic ions are dissolved in a weak paramagnetic or diamagnetic host (e.g., Fe–Au or Fe–Cu alloys), the magnetic moment carrying ions are separated by other ions and a direct overlap of wave functions responsible for the magnetic moment is hardly possible. Thus, the various types of indirect exchange interactions between magnetic moments are mediated by conduction electrons polarized by the local moment, which is known as RKKY interaction according to Ruderman, Kittel, Kasuya, and Yoshida. Under the assumption of a

spherical Fermi surface with radius k_F, the interaction is given by

$$J_{RKKY} \propto \frac{\cos(2k_F r)}{r^3}$$

The RKKY interaction is long ranged and oscillates as a function of the distance r from the polarizing local moment. Hence depending on the distance, the conduction electron polarization is positive or negative, yielding ferromagnetic or antiferromagnetic coupling with the next local moment at the position r, and is an important mechanism to describe complex magnetic structures. Archetypical examples are rare earths (R) and their intermetallics, where the localized 4*f*-electrons are responsible for magnetism without any overlap with their nearest-neighbor 4*f* wave functions. The 4*f*-moments in these R-intermetallics compare well with those of their free ion value of R^{3+} ions according to the Hund's rules and their Curie temperature are of the order of 10 to 10^2 K. In particular for systems containing *f*-electrons, the isotropic exchange interaction has to be extended by anisotropic terms. The anisotropy of the coupling is associated with the orbital moments of the electrons involved and the general exchange operator contains, besides the bilinear term $\hat{S}_i \cdot \hat{S}_j$, higher order isotropic interactions and all types of anisotropic exchange as, for example, Dzialoshinsky–Moriya exchange $\hat{S}_i \times \hat{S}_j$, which produces magnetic components perpendicular to the principal spin axis of the ferromagnetic or antiferromagnetic alignment. The 3*d*–4*f* exchange interactions, together with CEF effects, are the origin of the large magnetic anisotropy in R-3*d*-intermetallics needed for the high coercitivity of permanent magnets such as $Nd_2Fe_{14}B$ or $SmCo_5$ and Sm_2Co_{17}.

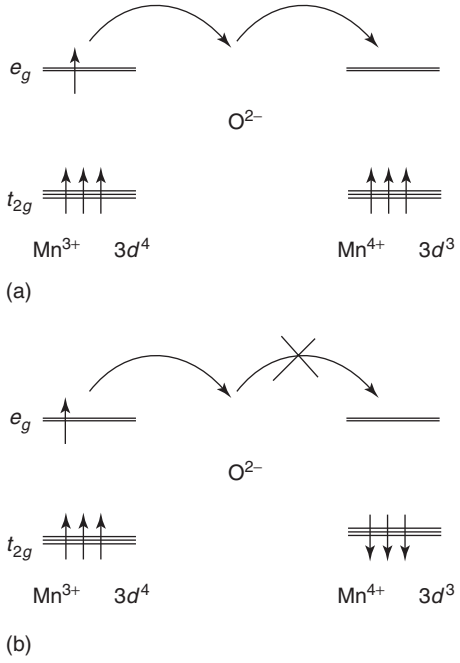


Figure 4 (a) Double exchange interaction favors hopping via an anion (O^{2-}) for ferromagnetic alignment of the transition metal ions by reducing kinetic energy, (b) while hopping is hindered by antiferromagnetic alignment.

Exchange Interactions in Free Electron Systems – Itinerant Exchange

Paramagnetism of the free noninteracting electron gas is described by the Pauli susceptibility $\chi_P = 2\mu_B^2 N(E_F)$, which relates the paramagnetic response of the system to an external field with the density of states $N(E)$ at the Fermi energy E_F. In the paramagnetic state, the number of up and down spins is equal ($n\uparrow = n\downarrow = n/2$); thus, the magnetic moment $m = \mu_B(n\uparrow - n\downarrow)$ given by the number of unpaired spins is zero; here, it is assumed that each electron carries a magnetic moment of 1 μ_B, since $S = \pm 1/2$ and gyromagnetic ratio $g = 2$ which is a good approximation for conduction electrons. If an external field H_0 is applied, the energy of an electron with spin up or down is $\pm\mu_B\mu_0 H$ depending upon the orientation of the spin with respect to the field. Thus, the external field splits the up and down spin bands with respect to each other by the energy $2\mu_B\mu_0 H$ (see **Figure 5**) yielding a small but finite magnetic moment $m = \mu_B(n\uparrow - n\downarrow)$ and a Pauli susceptibility $\chi_P = m/\mu_0 H_0$.

In the Stoner Wohlfahrth model, the susceptibility of the interacting electron gas, χ_S, is enhanced due to exchange interactions by a factor $S = (1 - I\,N(E_F))^{-1}$, known as the Stoner enhancement yielding,

$$\chi_S = \frac{2\mu_B^2 N(E_F)}{1 - IN(E_F)} = \chi_P S$$

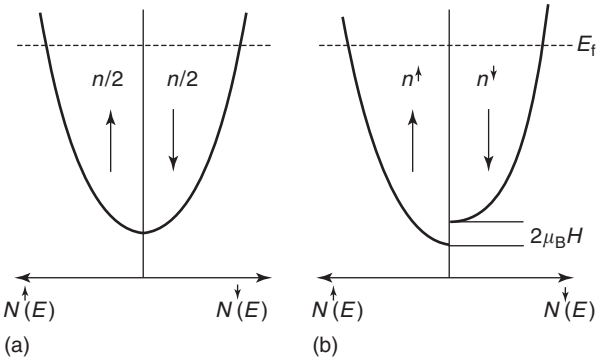


Figure 5 (a) Density of states of the free electron gas in the paramagnetic state, (b) splitting of the spin up and down bands by either an external field H_0 or by an effective exchange field H_{eff}.

The quantity I is the Stoner exchange factor describing phenomenologically, the exchange interactions of the interacting electron gas that can be interpreted as the above mentioned average exchange per pair of electrons $\bar{I}$ favoring their ferromagnetic alignment, that is, to increase $n\uparrow$ with respect to $n\downarrow$. In a more advanced treating, $\bar{I}$ corresponds to the Hubbard U which models the Coulomb interaction in the Hubbard Hamiltonian in combination with the hopping integrals t. If $IN(E_F)$ approaches 1, χ_S diverges and a second-order phase transition into a magnetically ordered state takes place. Thus $IN(E_F) \geqslant 1$ is the Stoner criterion for the onset of magnetism. That means for transition metals: a reasonably high density of states at E_F due to narrow enough d-bands together with exchange is needed for magnetic order. By analogy to the band splitting due to an external field in the paramagnetic state, the spontaneous moment in the ordered state arises from the spontaneously spin–split bands caused in a simple mean field model by an effective exchange field $H_{eff} = Im$.

See also: Diamagnetism; Disordered Magnetic Systems; Ferromagnetism; Localized and Itinerant Magnetism; Magnetic Order; Paramagnetism.

PACS: 31.10.+z; 75.10.−b

Further Reading

Barbara B, Gignoux D, and Vettier C (eds.) (1988) *Lectures on Modern Magnetism*. erlin: Spinger.

Blundells S (ed.) (2001) *Magnetism in Condensed Matter*. Oxford: Oxford University Press.

Fazekas P (1999) *Lectures on Electron Correlation and Magnetism*, Series in modern condensed matter physics, 5th edn. Singapore: World Scientific.

Jensen J and Mackintosh AR (eds.) (1991) *Rare Earth Magnetism*. Oxford: Clarendon.

Mattis DC (1981) *Theory of Magnetism I*. Solid state sciences, 17th edn. Berlin: Spinger.

Mohn P (2003) *Magnetism in the Solid State*, Solid state sciences, 134 edn. Berlin: Springer.

Richter MJ (1998) Band structure theory of magnetism in 3*d*–4*f* compounds. *Journal of Physics D: Applied Physics* 31: 1017–1048.

Magnetic Materials and Applications

H Szymczak, Polish Academy of Sciences, Warsaw, Poland

Introduction

Magnetic materials play a prominent role in modern technology. They are key components of motors, generators, and transformers. In this article, various magnetic materials of technological importance are described with an emphasis on recent developments.

Magnetic materials have a long history of usage. According to Chinese data, the compass was discovered in China more than 4500 years ago. There are also claims that a natural magnet was discovered in Magnesia (Asia Minor) more than 3500 years ago. Magnetic materials have contributed vitally to the history of civilization.

Traditionally, only those materials that exhibit ferromagnetic or ferrimagnetic properties are called "magnetic." Only nine elements are ferromagnets. All are metals, of which three (Fe, Co, Ni) are iron-group metals and the other six (Gd, Tb, Dy, Ho, Er, Tm) are rare-earth metals. The transition metals Fe and Co are the basic elements for preparation of alloys and compounds with large Curie temperature (T_C) and a large spontaneous magnetization (M_s) (see **Table 1**). Some of the intermetallic compounds with rare-earth metals are characterized by a very high value of magnetocrystalline anisotropy and magnetostriction.

The primary criterion allowing for classification of magnetic materials is coercivity, which is a measure of stability of the remanent state. Soft magnetic materials are characterized by low values of coercivity ($H_c < 10^3\,A\,m^{-1}$), while the coercivity of hard magnetic materials (usually permanent magnets) is higher than $10^4\,A\,m^{-1}$. Finally, semihard magnetic materials (mostly storage media) have coercivities in between

Table 1 Magnetic properties of 3*d* and 4*f* magnetically ordered elements

Element	*Curie temperature* (K)	*Saturation moment* (μ_B)
Fe	1043	2.22
Co	1388	1.72
Ni	635	0.61
Gd	293	7.63
Tb	220	9.34
Dy	89	10.33
Ho	20	10.34
Er	20	9.1
Tm	32	7.14

the above two values. In addition, there exist a number of sophisticated magnetic materials with unusual properties, such as giant magnetostriction, giant magnetoresistance, and giant magnetoimpedance. All these materials are important for modern technology and will be considered here. In modern applications, magnetic properties are tailored through precision tuning of the material microstructure. Thus, a central issue in the development and application of both soft and hard magnetic materials is the connection between extrinsic magnetic properties (coercivity, remanence, hysteresis loop) and microstructure. A universal relation was established between the coercivity H_c and the anisotropy constant K_1:

$$H_c = 2\alpha K_1/\mu_o M_s - N_{eff} M_s \qquad [1]$$

where the microstructural parameters α and N_{eff} describe the effects of microstructure and domain structure.

The present technology allows one to obtain a wide spectrum of microstructures ranging from amorphous, nanocrystalline to single crystals. The corresponding coercivities may vary by a factor of 10^7 to 10^8 from extremely soft to extremely hard magnetic materials.

Soft Magnetic Materials

Typical requirements for a soft magnetic material are low coercivity, high initial permeability, and low high-frequency losses. These characteristics usually go along with the following properties of the media:

- low effective magnetic anisotropy,
- high remanence,
- small magnetostriction (to suppress magnetoelastic contribution to the magnetic anisotropy), and
- low conductivity.

There was a remarkable progress in production and applications of soft magnetic materials in recent years. The static permeability, for example, has been increased to huge values in materials with near-zero values of magnetic anisotropy and magnetostriction. The best soft magnetic materials have permeabilities as large as $\mu > 10^5$. The permeability μ describes the response of magnetic materials to a weak magnetic field. For small fields, this response is linear and may be characterized by the initial permeability μ_i. Generally, the permeability is determined by two mechanisms: the growth of some magnetic domains at the expense of others (wall mechanism) and the rotation of magnetization within each domain (rotation mechanism). The same two mechanisms are also responsible for the hysteresis loop in soft magnetic materials. There are two classes of losses in these materials: hysteresis losses (determined by the area of hysteresis loop) and eddy current losses (related to the DC electrical resistivity). After a thermal demagnetization or after rapid changing of magnetization, almost all soft magnetic materials show a slow decrease of their permeability as a function of time (t), according to a law:

$$\Delta\mu = -D \ln t \qquad [2]$$

This decay of permeability with time is called "disaccommodation." The disaccommodation is ascribed to an increase in the domain wall stiffness, which originates in the presence of mobile defects interacting with magnetization.

Iron and Soft Steels

Iron is a very good and relatively cheap soft magnetic material. It has a very high saturation magnetization at room temperature and a large Curie temperature (see **Table 2**). Pure iron or soft steels (low-carbon steels, silicon steels) are used in magnetic shielding, in power and distribution transformers, in small motors, in electromagnets, and in various electromechanical devices. In AC applications, low-carbon steels are used because of their high permeability and low core loss. Silicon steels are widely used in transformers, rotating power equipment, and relays. Silicon is added to pure iron to increase resistivity and permeability and to decrease magnetostriction and magnetocrystalline anisotropy. In silicon steels, the basic magnetostriction constants approach zero at a silicon level slightly greater than 6 wt.%. Such high

Table 2 Soft magnetic materials (crystalline)

Material	*Saturation induction* B_s (T)	*Coercivity* H_c ($A\,m^{-1}$)	*Initial permeability* μ_i
Iron (ingot)	2.14	80	150
Carbon steel	2.14	100	200
Silicon iron (3% Si)			
Unoriented	2.01	60	270
Oriented	2.01	7	1 400
Permalloy (79% Ni, 5% Mo)	0.80	1	40 000
Permendur (49%Co, 2% V)	2.40	160	800
Supermendur (49%Co, 2% V)	2.40	16	1 000
Alfenol (16%Al)	0.80	3.5	4 000
MnZn ferrite	0.40	7	10 000
NiZn ferrite	0.33	80	290

silicon content alloys are usually produced by a fast solidification technique. This melt-spinning technique was initially developed for amorphous alloys. Silicon steels are divided into two main categories: Fe–Si sheets with nonoriented grain texture and grain-oriented Fe–Si sheets. In the latter case, a preferred crystallographic texture leads to a reduction of core loss and increase of permeability. Annealing of iron and steels removes stresses and improves the time stability of magnetic parameters. In order to avoid the $\alpha \rightarrow \gamma$ phase transition of iron, the temperature of annealing should not exceed 1125 K. In the case of silicon steel, this temperature is even higher since silicon stabilizes the α-Fe phase.

Nickel–Iron Alloys

Nickel–iron alloys (known as "permalloys") are used for electronic transformers and inductor applications in the 1–100 kHz range. These are characterized by high initial permeability (see **Table 2**) and good ductility for forming into very thin strips. Fe–Ni alloys with 80% of Ni are characterized by the almost simultaneous vanishing of anisotropy and magnetostriction. The best alloys have the composition $Fe_{15}Ni_{80}Mo_5$. Alloys with very high permeability are mainly used in safety devices and in magnetic field shielding.

Iron–Cobalt Alloys

The most important iron–cobalt soft magnetic alloys are "Permendur" (49%Fe–49%Co–2%V) and "Supermendur" (see **Table 2**). The technological importance of these alloys is a consequence of their high saturation magnetization and high Curie temperature. They are characterized by high permeabilities and low coercivities. Because of high cobalt content, Fe–Co alloys are relatively costly and they are used only for special purposes where cost of the magnetic material is not of great importance, for example, in airborne medium-frequency electrical engineering and in high-temperature magnetic devices.

Soft Ferrites

The term "ferrite" refers to spinel ferrites with the general formula $MO\ Fe_2O_3$, where M is a divalent cation (e.g., Mg^{2+}, Zn^{2+}, Cd^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Cu^{2+}). The spinel structure is cubic, with a face-centered cubic oxygen anion lattice. Metallic ions are distributed between tetrahedral (a) and octahedral (d) sites. In the "normal spinel" structure, tetrahedral sites are occupied by divalent cations and the octahedral sites by trivalent cations. In the "inverse spinel" structure, the tetrahedral sites are occupied by trivalent cations while the octahedral sites are occupied by divalent and trivalent cations. Magnetism of ferrites is usually determined by antiferromagnetic interactions between magnetic cations located at octahedral and tetrahedral sites. It leads to a relatively low value of saturation magnetization. The Curie temperature in most ferrites is lower than that in metals. The anisotropy constant K_1 of soft ferrites is usually small and negative with the $\langle 1\,1\,1 \rangle$ axis being the easy magnetization direction. Additions of Fe^{2+} or Co^{2+} ions to the lattice provide positive contribution to K_1 and a tendency to change the easy magnetization direction (to $\langle 1\,0\,0 \rangle$ axis). It also gives a possibility to achieve a near-zero value of the anisotropy constant. The resistivity of ferrites is in the insulating range and much larger than that of metallic alloys. The dominant mechanism for the conduction in soft ferrites is the transfer of electrons between Fe^{2+} and Fe^{3+} ions on the octahedral sites. It means that resistivity of ferrites is dependent on the ferrous ions content. Resistivities of Mn–Zn ferrites range from 10^{-2} to $10\,\Omega\,m$, while Ni–Zn ferrites have resistivities $\sim 10^7\,\Omega\,m$. The hopping Fe^{2+}–Fe^{3+} also determines the frequency dependence of μ_i.

Ferrites have some disadvantages: low saturation magnetization (see **Table 2**), relatively low Curie temperature, and inferior mechanical properties. Despite these disadvantages, ferrites are widely used in modern technology. They are used in two main fields: power electronics (mainly Mn–Zn and Ni–Zn ferrites) and low-power techniques (Ni–Zn ferrites up to 300 MHz and Mn–Zn high-permeability ferrites up to some MHz).

Ferrimagnetic Iron Garnets

The basic formula of the ferrimagnetic rare-earth iron garnets is $\{R_3^{3+}\}[Fe_2^{3+}](Fe_3^{3+})O_{12}$, where R is either a rare-earth element or yttrium. These ions occupy dodecahedral {c} sites. The iron ions are distributed between tetrahedral (d) sites and octahedral [a] sites. Garnets crystallize in the cubic system. In a ferrimagnetic garnet, the magnetic moments of the octahedral and tetrahedral sublattices are antiparallel to each other. The moments of the Gd^{3+} and heavier R^{3+} ions couple antiferromagnetically to the net moment of the Fe^{3+} sublattices. The resulting magnetization of the garnet is the sum of these three contributions:

$$M(T) = |M_a(T) - M_d(T) + M_c(T)|$$

At low temperatures the rare-earth magnetization dominates. This effect is exactly canceled out by the net magnetization of iron sublattices at the "compensation temperature" T_{comp}. Below T_{comp}, the magnetization M_c lies parallel to a saturating field,

and above T_{comp} it is antiparallel. The moments of the light rare-earth ions couple ferromagnetically to the net moment of the Fe^{3+} sublattices and no compensation points exist. In the compound $Y_3Fe_5O_{12}$ (YIG) the most attractive property is the extremely narrow ferrimagnetic resonance line width, of the order of $50\,A\,m^{-1}$. Thus, YIG with various partial substitutions for Y finds wide applications in microwave devices. Garnets find major application at lower frequencies (<5 GHz). Above 5 GHz, spinel ferrites are commonly used. YIG exhibits a room temperature saturation magnetization $\mu_o M_s \approx 0.175$ T and a magnetocrystalline anisotropy $\sim 500\,J\,m^{-3}$. The saturation magnetization of light rare-earth garnets is slightly higher while it is much lower for heavy rare-earth garnets.

Amorphous Alloys

Amorphous alloys include two large classes: transition metal-metalloids with compositions near $(Fe, Co, Ni)_{80}$ $(B, C, Si)_{20}$ and transition metals – Zr, Hf. The former type offers greater potential for technical applications. Amorphous alloys are produced mainly by rapid ($\sim 10^6\,K\,s^{-1}$) quenching from the melt or by either sputtering or evaporation. The lack of long-range atomic order in amorphous magnetic alloys is responsible for: (1) high metallic resistivity ($\sim 10^{-6}\,\Omega\,m$), (2) absence of macroscopic magnetocrystalline anisotropy, (3) absence of grain boundaries and dislocations playing the role of pinning centers in crystalline magnetic materials, (4) a broad range of compositions available, yielding a continuous spectrum of property values, (5) outstanding mechanical properties, and (6) excellent resistance to corrosion.

The first class of amorphous alloys is divided into three groups (see **Table 3**):

1. iron-based amorphous alloys with a saturation induction B_s in the 1.1–1.8 T range with excellent high-frequency (up to 100 kHz) characteristics,
2. iron–nickel based amorphous alloys with a relatively low saturation induction but at the same time a low magnetostriction, and
3. cobalt-based amorphous alloys with a near-zero magnetostriction, a very high permeability, and low losses.

In addition to the applications of Fe-based alloys in electric utility and industrial transformers, amorphous metals are increasingly used in power electronics, telecommunication equipment, sensing devices, electric power conditioning, electronic article surveillance systems, etc. Slow crystallization of materials with a high resistance to crystallization allows a decreased critical cooling and enables stable bulk metallic glass formation. In this way, Fe-based bulk amorphous alloys with good soft magnetic properties are fabricated.

In soft magnetic amorphous ribbons and wires, the "giant magnetoimpedance effect" (GMI) is observed. This effect consists of drastic changes of the complex impedance (both real and imaginary components) of amorphous alloys upon application of an external magnetic field (of a few Oe) or mechanical stresses. Sensitivity of up to 500% of relative impedance change per Oe is observed in the very low-field region. The GMI effect is strongly dependent on the frequency of the applied field and on the magnetic anisotropies present in the material. That makes GMI materials quite convenient for application in sensors of current, position, level of liquid, and pressure. The sensors are employed in technologies such as car traffic monitoring, quality control steels, detection of earthquakes, and medicine.

Nanocrystalline Alloys

Nanocrystalline alloys consist of homogeneously distributed ultrafine grains with an average size of ~ 5–20 nm embedded in an amorphous matrix. They are prepared from amorphous ribbons (with the addition of Cu and Nb), which are subjected to a recrystallization annealing. The volume ratio of the crystalline phase ranges from 50% to 80% and can be controlled by the annealing temperature. Nanocrystalline alloys present a lower coercivity field than the amorphous samples, higher initial permeability than the Co-based amorphous alloys

Table 3 Soft magnetic materials (amorphous)

Material	*Curie temperature* (K)	*Saturation induction* (T)	*Coercivity* ($A\,m^{-1}$)	*Magnetostriction constant* (10^{-6})
$Fe_{81}B_{13.5}Si_{3.5}C_2$	370	1.61	3.2	30
$Fe_{67}Co_{18}B_{14}Si$	415	1.80	4.0	35
$Fe_{56}Co_7Ni_7Zr_2Nb_8B_{20}$	508	0.71	1.7	10
$Fe_{72}Al_5Ga_2P_{11}C_6B_4$	605	1.07	5.1	2
$Fe_{40}Ni_{38}Mo_4B_{18}$	353	0.88	1.2	12
$Co_{67}Ni_3Fe_4Mo_2B_{12}Si_{12}$	340	0.72	0.4	0.5

Table 4 Soft magnetic materials (nanocrystalline alloys)

Material	*Curie temperature* (K)	*Saturation induction* (T)	*Coercivity* ($A\,m^{-1}$)	*Magnetostriction constant* (10^{-6})
$Fe_{73.5}Cu_1Nb_3Si_{13.5}B_9$	843	1.24	0.53	2.1
$Fe_{73.5}Cu_1Nb_3Si_{16.5}B_6$	833	1.18	1.1	~0
$Fe_{56}Co_7Ni_7Zr_2Ta_8B_{20}$	538	0.85	1.5	14
$Fe_{56}Co_7Ni_7Zr_2Nb_8B_{20}$	508	0.71	1.7	10
$Fe_{90}Zr_7B_3$		1.63	5.6	−1.1
$Fe_{86}Zr_7B_6Cu_1$		1.52	3.2	1
$Fe_{78}Si_9B_{13}$	688	1.56	2.4	27

(see **Table 4**), very low magnetostriction, and (due to the exchange softening) nearly zero effective magnetic anisotropy. The presence of Cu and Nb allows one to obtain well-differentiated microstructures. The resistivity of these materials is usually of the same order of magnitude as that of the amorphous ones, namely $1–1.5 \times 10^{-6}\,\Omega\,m$. The high permeability and weak energy loss of soft nanocrystalline alloys make them valuable for various sensors, safety devices, high-frequency transformers, and in many other applications.

Hard Magnetic Materials

The most important hard magnetic materials suited for permanent magnets are expected to have the following intrinsic magnetic properties:

1. *High Curie temperature*. This favors the 3*d* elements which exhibit direct exchange.
2. *High saturation magnetization at room temperature*. This again favors 3*d* elements (Fe, Co). The 3*d*–4*f* intermetallics based on the light rare earth which couple parallel with the 3*d* moment, are of special interest here.
3. *High uniaxial anisotropy*. Uniaxial anisotropy is necessary for the formation of high coercivity. The main contribution to a reasonably high anisotropy comes from the 4*f* ions. This anisotropy is of the single-ion type. To combine the best features of 3*d* and 4*f* components, a strong 3*d*–4*f* coupling is required.

Permanent magnets are characterized by an extremely wide hysteresis loop. The most important extrinsic parameter is the coercivity, H_c, connected with the total anisotropy via eqn [1].

Alnicos

Alnicos are a group of heat-treated Fe–Co–Ni–Al–Cu alloys which can be divided into two main subgroups: isotropic alloys containing 0–20 wt.% Co (alnicos 1–4) and anisotropic alloys with 22–24 wt.% Co and a titanium content of 5–8 wt.% (alnicos 5–9). The anisotropy in case of alnicos 5–9 is produced by controlled cooling or isothermal heat treatment in a saturating magnetic field. The main source of anisotropy is the shape anisotropy associated with elongated Fe–Co particles in an Ni–Al matrix aligned parallel to the magnetic field during spinodal decomposition of the alloy. Alnicos are widely used as permanent magnets (see **Table 5**). The most important applications are loudspeakers, watt-hour meters, electric motors, generators, and alternators.

Table 5 Magnetic properties of alnicos

Alloys	*Saturation induction* (T)	*Coercivity* ($A\,m^{-1}$)	*Energy product* $(BH)_{max}$ ($kJ\,m^{-3}$)
Alnico 3	0.5–0.6	54–40	10
Alnico 5	1.2–1.3	52–46	40–44
Alnico 7	0.74	85	24
Alnico 9	1.0–1.1	140–110	60–75

Hexagonal Ferrites

Hard ferrites have the general chemical formula $MFe_{12}O_{19}$, where M is Ba, Pb, or Sr. They have a hexagonal structure with Fe^{3+} ions located at five different crystallographic positions. The superexchange via bridging oxygen atoms leads to a collinear ferrimagnetic structure. The high magnetocrystalline anisotropy (see **Table 6**) has a crystal-field and dipolar origin. The intrinsic magnetic properties of hexagonal ferrites may be modified by chemical substitution. These ferrites exhibit a large coercivity ($200–300\,kA\,m^{-1}$) but a relatively low remanence (~0.4 T). They are produced by ceramic techniques and are often bonded in plastic. Owing to their low price, hexagonal ferrites are widely used in electric motors in automotive and acoustic industries (electrical equipment for vehicles, starter motors, etc.).

Cobalt–Samarium Magnets

Hard magnets based on $SmCo_5$ possess the highest uniaxial anisotropies of any class of magnets

Table 6 Intrinsic magnetic properties of $MFe_{12}O_{19}$ hexagonal ferrites

M	*Curie temperature* (K)	*Saturation magnetization* (T)	*Anisotropy constant* ($MJ\,m^{-3}$)	*Coercivity* (T)
Ba	742	0.48	0.33	1.7
Sr	750	0.48	0.35	1.8
Pb	725	0.40	0.25	1.5

Table 7 Magnetic properties of cobalt–samarium compounds

Compound	*Curie temperature* (K)	*Saturation magnetization* (T)	*Coercive field* ($kA\,m^{-1}$)	*Anisotropy constant* ($MJ\,m^{-3}$)	*Energy product* ($kJ\,m^{-3}$)
$SmCo_5$	1020	1.07	1500	17.2	160
Sm_2Co_{17}	1190	1.22	1100	3.3	260

Table 8 Magnetic properties of neodymium–iron–boron

	Curie temperature (K)	*Saturation magnetization* (T)	*Coercive field* ($kA\,m^{-1}$)	*Anisotropy constant* ($MJ\,m^{-3}$)	*Energy product* ($kJ\,m^{-3}$)
$Nd_2Fe_{14}B$	585	1.52		4.5	
Magnet (standard)			1000		260
Magnet (high remanence)			1250		400

($K \sim 10^7\,J\,m^{-3}$, see **Table 7**), while magnets based on Sm_2Co_{17} exhibit higher saturation magnetization and higher Curie temperature. Each of them has a uniaxial structure ($SmCo_5$ – hexagonal; Sm_2Co_{17} – hexagonal or rhombohedral). Chemical substitution in both compounds modifies their intrinsic and extrinsic magnetic properties. A small substitution of Cu for Co leads to precipitations of a nonmagnetic phase, which increases the coercivity. In this case, the coercivity mechanism is controlled by domain wall pinning on small $SmCo_5$ particles. A similar mechanism exists in $Sm_2(Co,Cu)_{17}$. The nucleation-controlled coercive force dominates in single-phase $SmCo_5$. The defects, which act as nucleation centers, are mostly associated with the surface of milled powder particles. The Curie temperature and saturation magnetization of Sm_2Co_{17} are substantially improved by partial Fe, Zr, and Cu substitution for Co. Fe is added to increase remanence, and Cu and Zr to form the required microstructure. In this case, the coercivity has a pinning-type character.

Neodymium–Iron–Boron Alloys

Permanent magnets based on $Nd_2Fe_{14}B$ have record values for the energy product $(BH)_{max}$. Except for the relatively low Curie temperature, the Nd–Fe–B magnets are superior to all other magnetic materials. The $Nd_2Fe_{14}B$ compound has a tetragonal structure. The Nd and Fe sublattices are aligned parallel to one another, giving a room temperature magnetization of 1.61 T, the largest of any rare-earth compound used for permanent magnets. The strong magnetocrystalline anisotropy of $Nd_2Fe_{14}B$ (see **Table 8**) results from the low symmetry of both the Nd and Fe sites. All components of $Nd_2Fe_{14}B$ can be partly substituted by metals or metalloids. Such substitution may lead to an improvement in the thermal stability of the magnetic properties above room temperature or/and to an improvement of the microstructure and performance of the magnet. For example, cobalt substitution improves the performance of $Nd_2Fe_{14}B$ by an increase of the Curie temperature, but at the same time the presence of cobalt decreases the magnetocrystalline anisotropy.

Small quantities of Dy are often substituted for Nd to improve the coercivity of Nd–Fe–B magnets. However, dysprosium reduces the magnetization of the system. Two principal fabrication routes for Nd–Fe–B magnets are: powder metallurgy (for oriented sintered magnets) and melt spinning (used mostly for bonded magnets). The optimum composition for sintered magnets was found to be $Nd_{15}Fe_{77}B_8$ which is both Nd- and B-rich with respect to the

stoichiometric $Nd_2Fe_{14}B$. The ideal two-phase microstructure in Nd–Fe–B sintered magnets consists of fully aligned grains of $Nd_2Fe_{14}B$ separated by the minimum amount of a nonmagnetic intergranular phase. Coercivity of sintered magnets is dominated by the nucleation mechanism. High remanence-oriented magnets have energy product of $\sim 400\,kJ\,m^{-3}$ and a remanence approaching 1.5 T. A different microstructure and magnetization is encountered in nanocrystalline Nd–Fe–B magnets produced by melt spinning. The origin of coercivity in melt-spun Nd–Fe–B magnets is similar to that established for sintered magnets. The nanocrystalline structure of Nd–Fe–B magnets may also be developed by mechanical alloying. Milling Fe, Nd, and B to form $Nd_2Fe_{14}B$ results in formation of α-Fe (of ~ 5 nm size) and an Nd-rich amorphous phase which should be heat treated to obtain high-coercivity material. A typical composition of these magnets is $Nd_{14}Fe_{80}B_6$. In case of fully dense isotropic nanomagnets, the characteristic energy products are $\sim 100\,kJ\,m^{-3}$ while the values of remanences are ~ 0.80 T. Mechanical alloying has also been used to synthesize other nanoscale two-phase mixtures of hard and soft magnetic phases. The exchange coupling between magnetically hard and soft phases can result in significant enhancement of the remanence. These nanocomposites are known as "exchange spring magnets." For this mechanism to be effective, the size of the soft grains should not exceed a few times the domain wall width for the hard phase ($\sim$20–30 nm). In such cases, the material behaves as if it were homogeneous, with anisotropy and magnetization being equal to the mean value of the constituent phases. Some typical examples are $Nd_2Fe_{14}B/Fe_3B$, $Nd_2Fe_{14}B/Fe$, and $Sm_2Fe_{17}N_3/Fe$.

Despite their relatively low Curie temperature and their susceptibility to corrosion, Nd–Fe–B magnets have found a great number of applications where high coercivity and energy result in design advantages. The applications include voice coil motors of hard disk drives and other small compact electronic products.

Giant Magnetostrictive Materials

Rare-earth metals have intrinsically huge magnetostriction because of the strong 4*f* electrons'–crystal lattice coupling. Although this coupling is indirect, it is huge because of the large spin–orbit coupling and the highly anisotropic localized nature of the 4*f* electrons. The relatively low magnetic ordering temperatures of rare-earth metals (R) (see **Table 1**) make use of the high value of magnetostriction at room temperature, possible only by combining them with 3*d* metals. Such a combination was realized most effectively in the Laves phase compound RFe_2, particularly in $TbFe_2$ (see **Table 9**), also known as "Terfenol." However, this compound has an intrinsically large magnetocrystalline anisotropy which limits its technical application. By addition of Dy atoms with a positive magnetocrystalline anisotropy, it is possible to compensate the negative magnetocrystalline anisotropy of terbium. The compound $Tb_{0.3}Dy_{0.7}Fe_2$ known as "Terfenol-D" maintains a giant magnetostriction (at low magnetic fields) with low magnetocrystalline anisotropy and relatively high Curie temperature. In the polycrystalline state, giant magnetostriction is possible only when the grains are aligned along the easy $\langle 1\,1\,1 \rangle$ direction and the effect is measured along this direction. This is because Terfenol-D alloys exhibit enormously large anisotropy in magnetostriction ($\lambda_{111} \gg \lambda_{100}$). These alloys are produced either by a directional solidification method, sintering, or by a polymer bonding. The last method has an advantage over the other two because including the polymer binder with a high electrical resistivity, the frequency limit of the polymer-bonded composites can be extended to hundreds of kHz. Favorable magnetostrictive properties are also observed in amorphous R–Fe alloys. The main industrial applications of "Terfenol-D" are linear actuators.

Manganites

Doped perovskite manganites $La_{1-x}A_xMnO_3$ (where A is Ba, Ca, Sr, or Pb) show a wide variety of magnetic field-induced phenomena, including the gigantic decrease of resistance under the application of a magnetic field. It is termed "colossal magnetoresistance effect (CMR)." By replacement of a trivalent rare earth by a divalent alkaline earth, the nominal valence of Mn can be continuously tuned between $3+$ (for $x = 0$) and $4+$ (for $x = 1$). The magnetic and magneto-transport properties of manganites are governed mainly by the mixed valence Mn^{3+}/Mn^{4+} ions and less affected by the oxygen ions. The *d*-electrons are subject to a variety of interactions, of which the most important are the double exchange, superexchange, and Jahn–Teller coupling. The double exchange interaction is responsible for a

Table 9 Magnetostriction constants for several rare-earth materials

Material	$T = 4.2$ K	*Room temperature*	
	λ_{111} (10^{-6})	λ_{111} (10^{-6})	*Polycrystal* (10^{-6})
$TbFe_2$	4400	2600	1750
$Tb_{0.3}Dy_{0.7}Fe_2$		1600	1200
Polymer-bonded Terfenol-D			536
$Tb_3Fe_5O_{12}$ garnet	2460		

strong correlation between magnetization and resistivity (or magnetoresistivity). The large resistance is related to the formation of small lattice polarons in the paramagnetic state. In these interactions, an important role is played by the e_g electrons on the Mn^{3+} ions. The orbital degeneracy leads to a variety of instabilities (such as magnetic, structural, and metal–insulator transitions, collapse of charge-ordered states under a magnetic field) and is responsible for a strong magnetoelastic coupling. The e_g electrons can be itinerant and hence play the role of conduction electrons, when electron vacancies or holes are created in the e_g orbital states of the crystal. The other characteristic feature of manganites is the appearance of unusual collective state of charge order, observed mostly for $x > 0.3$. This insulating state competes with two other basic phases in manganites, namely with the ferromagnetic metallic phase and the nonmetallic paramagnetic phase. A number of experiments have demonstrated that over wide ranges of compositions there can be a coexistence between the ferromagnetic metallic and charge-ordered insulating phases.

See also: Disordered Magnetic Systems; Magnetic Order; Magnetoresistance Techniques Applied to Fermi Surfaces.

PACS: 75.50.Ww; 75.50.Kj; 75.50.Dd; 75.47.Lx; 75.50.Gg; 81.01.Bc

Further Reading

Baden-Fuller A (1987) *Ferrites at Microwave Frequencies*. London: Peregrinus.
Chen CW (1977) *Magnetism and Metallurgy of Soft Magnetic Materials*. Amsterdam: North-Holland Publishing Company.
Chikazumi S (1997) *Physics of Ferromagnetism*. Oxford: Clarendon.
Coey JMD, Viret M, and von Molnar S (1999) Mixed-valence manganites. *Advances in Physics* 48: 167.
Craik DJ (ed.) (1975) *Magnetic Oxides*. London: Wiley.
Cullity BD (1972) *Introduction to Magnetic Materials*. Reading Mass. *and Magnetic Materials*. Reading, MA: Addison-Wesley.
Engdahl G (2000) *Handbook of Giant Magnetostrictive Materials*. London: Academic Press.
Gibbs MRJ (ed.) (2001) *Modern Trends in Magnetostriction Study and Application*. Dordrecht: Kluwer Academic Publishers.
Kaplan TA and Mahanti SD (1999) *Physics of Manganites*. New York: Kluwer Academic Publishers.
Moorjani K and Coey JMD (1984) *Magnetic Glasses*. Amsterdam: Elsevier.
O'Handley RC (2000) *Modern Magnetic Materials: Principles and Applications*. New York: Wiley.
Skomski R and Coey JMD (1999) *Permanent Magnets*. Bristol: Institute of Physics Publishing.
Snelling EC (1988) *Soft Ferrites: Properties and Applications*. London: Butterworth.
Tokura Y (ed.) (1997) *Colossal Magnetoresistance Oxides*. Amsterdam: Gordon and Breach.
Tokura Y and Tomioka Y (1999) Colossal magnetoresistive manganites. *Journal of Magnetism and Magnetic Materials* 200,1.
du Tremolet de Lacheisserie E, Gignoux D, and Schlenker M (2002) *Magnetism II – Materials and Application*. Dordrecht: Kluwer Academic Publishers.

Magnetic Order

D Givord, CNRS, Grenoble, France

Introduction

Magnetism is a property of electrons within matter. Strong interactions exist between the atomic magnetic moments; they are of electrostatic origin, and are termed exchange interactions.

In the simplest case, the exchange interactions favor parallelism of all magnetic moments; this leads to ferromagnetic order. The properties of ferromagnetic materials can be described within the molecular field model. However, the experiment provides evidence for the local character of the interactions which cannot be explained within this model.

More generally, the exchange interactions may favor different types of coupling, and a large variety of different magnetic structures may exist. In the antiferromagnetic structures, the moments are arranged in two groups, termed sublattices. Within each sublattice, the moments are parallel, but they are antiparallel to the moments in the other sublattice. In many instances, the magnetic structure results from a competition between different interactions which cannot be satisfied simultaneously. This is, in particular, the case of the helimagnetic and modulated structures. When the arrangement of the magnetic atoms themselves is not periodic, spin-glass arrangements are found.

The most powerful technique for the study of magnetic structures is neutron diffraction, in which the neutron magnetic moment is used to probe the arrangement of the atomic moments.

Magnetic Moment and Magnetic Field

Magnetic Moment

Within matter, magnetism is due to electron motion around the nuclei. There are two contributions to the magnetic moment of an electron:

- the orbital moment, $\boldsymbol{m}_l$, is linked to the motion of the electrons in their orbits. It may be expressed as

$\boldsymbol{m}_l = -\mu_B \boldsymbol{l}$, where $\boldsymbol{l}$ is the orbital momentum and μ_B, the Bohr magneton, is the fundamental unit of magnetic moments ($1\mu_B = 0.927\,10^{-23}\,A\,m^2$). The quantum mechanical eigenvalues of $\langle l^2 \rangle$ are equal to $l(l+1)$ where l is an integer, called the orbital quantum number. The values $l = 0$, 1, 2 and 3 are associated to electrons of the *s*, *p*, *d*, and *f* shells respectively. The eigenvalues of $\langle l_z \rangle$ are integers, denoted m, where m is the magnetic quantum number; m is restricted between two limits, $+l$ and $-l$.

- the spin moment, $\boldsymbol{m}_s = -2\,\mu_B \boldsymbol{s}$ is related to the electron spin momentum, $\boldsymbol{s}$, most often referred to as the "spin." The spin is an entity of purely quantum mechanical nature. The eigenvalues of $\langle s^2 \rangle$ are equal to $s(s+1)$, where s, the spin quantum number, equal to 1/2. The eigenvalue of $\langle s_z \rangle$ are $+1/2$ and $-1/2$.

Due to the electrostatic interactions existing between electrons located on the same electronic shell of a given atom, the spin and orbital momentum of individual electrons are coupled together, according to the so-called Hund's rules. The resulting orbital momentum of the considered shell is noted as L, and the resulting spin, S. L and S are strongly coupled together by the spin–orbit coupling. For a full electronic shell, $L = 0$ and $S = 0$; magnetism is a property of partly filled electronic shells.

In the solid state, only two series of stable elements may show spontaneous magnetism (the actinides may be magnetic as well).

1. In the 3*d* transition series, to which Fe, Co, and Ni belong, magnetism is due to the progressive filling of the 3*d* shell. The orbital momentum is almost quenched. In the case where the 3*d* electrons are localized, the maximum value of the magnetic moment is approximately equal to $-2\mu_B S$. In 3*d* metallic systems, the magnetic 3*d* electrons have a certain itinerant character. The magnetic moments, expressed in μ_B, take a fractional value, which cannot be derived by the present simple considerations.
2. In the rare-earth series, from cerium to lutetium, magnetism results from the progressive filling of the 4*f* shells. The 4*f* electrons are fully localized. The maximum value of the moment is equal to $-g_J \mu_B J$, where J is equal to $L+S$ or $L-S$, depending on the considered element and g_J is the Landé factor.

Another magnetic moment may exist within matter, associated to the constitutive elements of the nucleus. The proton moment, μ_P, and the neutron moment, μ_N, amount to $\mu_B/1836$ and $\mu_B/865$ respectively.

Magnetic Moments in a Magnetic Field – Paramagnetism

When submitted to an applied magnetic field, $\boldsymbol{H}_{app}$, a magnetic moment, $\boldsymbol{m}$, tends to align along the field direction. The corresponding energy, called Zeeman energy, e_Z, is

$$e_Z = -\mu_0 \boldsymbol{m} \cdot \boldsymbol{H}_{app} \qquad [1]$$

An assembly of moments, $\boldsymbol{m}$, may be characterized by the magnetization, M, which is the moment per unit volume measured along the field direction, z:

$$M = N \langle m_z \rangle_T \qquad [2]$$

where N is the number of moments per unit volume and $\langle m_z \rangle_T$ is the thermal averaged value of m_z, the projection of $\boldsymbol{m}$ along $\boldsymbol{H}_{app}$.

At finite temperature, moment alignment is in competition with thermal activation which tends to create disorder; this characterizes paramagnetism. For a system of classical noninteracting moments, the magnetization, M, is given by

$$M = N m_0 \mathscr{L}(x) \qquad [3]$$

In expression [3], m_0 is the value of the moment at 0 K and $\mathscr{L}(x) = \coth(\chi) - 1/\chi$ is the Langevin function. The parameter x which enters into the Langevin function is the ratio between Zeeman energy and thermal energy:

$$x = \frac{\mu_0 \boldsymbol{m} \cdot \boldsymbol{H}_{app}}{k_B T} \qquad [4]$$

where μ_0 is the permittivity of vacuum ($= 4\pi \times 10^{-7}$ SI) and k_B is the Boltzmann constant.

For small x, the Langevin function may be expanded as $\mathscr{L}(x) \approx x/3 - x^3/45 + \cdots$. The high-temperature development of [3] is

$$M = C \frac{H_{app}}{T} \qquad [5]$$

where C is the Curie constant, given by

$$C = \frac{\mu_0 N m_0^2}{3k_B} \qquad [6]$$

A useful experimental parameter is the susceptibility, $\chi = M/H_{app}$. The reciprocal susceptibility $1/\chi$ varies linearly with T, with a slope $1/C$ (**Figure 1a**). This constitutes the Curie law.

For systems of quantum moments, the Langevin function must be replaced by a Brillouin function, $\mathscr{B}(x)$. The moment m_0 in expression [3] is equal to $-2\mu_B S$ (systems in which 3*d* localized electrons are involved) or $-g\mu_B J$ (rare-earth-based systems), the

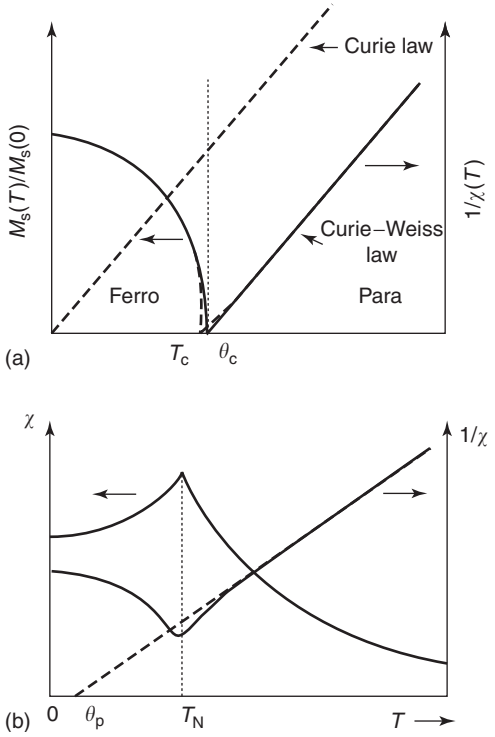


Figure 1 (a) In a system of noninteracting moments, the reciprocal susceptibility, $1/\chi(T)$, varies linearly with temperature with a slope $1/C$, where C is the Curie constant. In the presence of interactions, the molecular field model predicts a similar variation, but the representative curve is shifted towards higher temperatures. At $T = T_C$, the susceptibility becomes infinite. At $T < T_C$, a spontaneous magnetization, M_s, exists. This is ferromagnetism. The molecular field model does not properly give account for the existence of local fluctuations. These are responsible for the fact that the Curie temperature is lower than predicted and that the variation of $1/\chi(T)$ is not linear in the vicinity of T_C. (b) In an antiferromagnetic system, the susceptibility presents a maximum at the ordering temperature, T_N. In the paramagnetic state, the variation of $1/\chi(T)$ vs. T well above T_N is again linear.

maximum values of the projection of $\boldsymbol{m}$ along the field, whereas m_0 in expression [6] is replaced by μ_{eff}, the effective moment which is proportional to $\sqrt{S(S+1)}$ or to $\sqrt{J(J+1)}$.

Ferromagnetic Order

The Molecular Field Model

There is a category of magnetic materials, where the high-temperature susceptibility does not follow the Curie law, but is given by the expression

$$\chi = \frac{C}{T - T_C} \qquad [7]$$

where T_C is the Curie temperature (**Figure 1a**). Expression [7] constitutes the Curie–Weiss law.

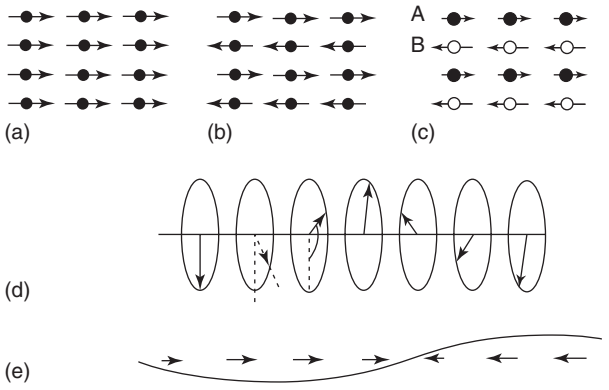


Figure 2 Schematic representations of various types of magnetic structures: (a) ferromagnetism, (b) antiferromagnetism, (c) ferrimagnetism, (d) helimagnetism, and (e) modulated structure.

It may be ascribed to interactions existing between the magnetic moments. To first approximation, these interactions are equivalent to a field, termed H_m created by the magnetic moments themselves. Within the molecular field theory of ferromagnetism, H_m is assumed to be proportional to the magnetization, M:

$$H_m = nM \qquad [8]$$

where n is the molecular field coefficient, such that

$$T_C = nC \qquad [9]$$

The value of the magnetization M may be obtained by replacing the applied field H_{app} in expression [4] by $H_{app} + H_m$ and by expressing that relations [3] and [8] must be satisfied simultaneously. Consider the special case where $H_{app} = 0$. At temperatures $T > T_C$, the only possible solution is $M = 0$; the moments are oriented at random and the system is paramagnetic, as it is in the absence of interactions. In this temperature range, the slope of the reciprocal susceptibility versus temperature is equal to $1/C$, as it is in noninteracting systems (**Figure 1a**). At $T = T_C$, the susceptibility diverges (it becomes infinite) and below T_C, a "spontaneous magnetization," M_s, exists within matter, indicating a certain degree of common alignment of all the moments. Such behavior characterizes ferromagnetic order (**Figure 2a**). The spontaneous magnetization, $M_s(T)$, progressively increases as temperature is decreased; at 0 K, it reaches the value $M_s = M_{s,0}$ ($\equiv N m_0$) of the magnetization at absolute saturation (**Figure 1a**).

Note that the magnetic energy density, associated with the formation of ferromagnetism, may be expressed as

$$E_{FM} = -\tfrac{1}{2}\mu_0 M_s H_m = -\tfrac{1}{2}\mu_0 n M_s^2 \qquad [10]$$

Beyond the Molecular Field Model: Evidence for Short-Range Order

The transition from a low-temperature ordered phase to a high-temperature disordered phase is a feature found in many physical systems. In the present case, the disordered phase is the paramagnetic phase and the ordered phase is the ferromagnetic one. First-order phase transitions are characterized by a discontinuity of thermodynamic variables such as the density or the entropy. In general, no such discontinuity exists at a paramagnetic–ferromagnetic transition. Landau has shown that "second-order" phase transitions are characterized by an abrupt symmetry change. In the paramagnetic phase, the system is invariant upon reversal of the magnetization; in the ferromagnetic phase, it is not invariant upon this reversal. In the vicinity of T_C, the magnetization is predicted to vary as

$$M_s = \sqrt{\frac{a}{2b}}(T_C - T)^{1/2} \qquad [11]$$

Here 1/2 is the critical exponent of the magnetization within the molecular field model. Experimentally, this critical exponent, denoted β, is found to be less than 0.5 (i.e., $\sim$0.3–0.4). This departure from the molecular-field-predicted value is due to the existence of local fluctuations of the magnetic state. Proper values of the critical exponents are obtained within the renormalization group theory.

Fluctuations cannot be explained within the molecular field model because H_m is proportional to M_s, a macroscopic quantity. Actually, the existence of magnetic fluctuations indicates that the magnetic interactions are short-ranged. The modification of the state for a given moment influences the state of the neighboring moments, but not that of moments which are at large distances. Magnetic disorder may proceed by keeping a certain degree of local order, while allowing the system to become disordered at the macroscopic level. The magnetic energy is reduced whereas the entropy is the same as for the case of homogeneous disorder. Due to short-range order, the experimental value of the Curie temperature is always lower than that predicted by the molecular field model. Close to T_C, the local correlations existing between moments lead to nonlinearity in the temperature dependence of the reciprocal susceptibility. At temperatures far above T_C, the influence of local fluctuations become negligible and the molecular-field-temperature dependence of the reciprocal susceptibility is recovered. The extrapolated intercept with the temperature axis gives the value of the so-called paramagnetic Curie temperature, θ_C, which is higher than T_C (**Figure 1**).

The local character of the magnetic interactions manifests itself as well in the ordered magnetic state, at low temperatures. The spontaneous magnetization decreases as $T^{3/2}$, more rapidly than predicted by the molecular field model. Specific excitations such as the spinwaves occur, which allow the local alignment of neighboring moments to be better preserved than in the case of random disorder, while allowing global disorder to develop.

Exchange Interactions

In a given ferromagnetic material, the strength of the molecular field may be derived from the value of the Curie temperature, T_C (see relation [9]). In Fe metal, the Curie temperature amounts to 1043 K, and in Co metal, it amounts to 1380 K; $H_m \approx 10^9\,\mathrm{A\,m^{-1}}$ is deduced.

From magnetostatics, it is known that a magnetic moment, $\boldsymbol{m}$, is the source of a dipolar magnetic field, $\boldsymbol{H}$, which, at point P, is expressed as

$$\boldsymbol{H} = \frac{1}{4\pi}\left[3\frac{(\boldsymbol{m}\cdot\boldsymbol{r})\boldsymbol{r}}{r^5} - \frac{\boldsymbol{m}}{r^3}\right] \qquad [12]$$

where $\boldsymbol{r}$ is the vector linking point O to point P. For $m \approx 2\,\mu_B$, which is the order of magnitude of the Fe or Co moment in metals, the corresponding magnetic field is of the order of $10^6\,\mathrm{A\,m^{-1}}$. This value is $\sim$1000 times smaller than the value derived from T_C. Except in rare cases, magnetic order cannot be attributed to dipolar interactions.

As shown by Heisenberg, the interactions at the origin of magnetic order are not magnetic in origin, but result from the repulsive electrostatic interactions existing between electrons. The Hamiltonian describing an ensemble of interacting electrons contains a term, called the exchange term, which results from electron indiscernibility. Magnetism emerges when account is taken of the Pauli exclusion principle. This stipulates that two electrons cannot occupy the same quantum state, defined by both space and spin variables. It imposes full antisymmetry of the wave function representative of the electron ensemble and, as a result, the interaction energy between electrons depends on their respective spin states.

It is often convenient to distinguish two terms in the exchange interactions. The intra-atomic interactions give rise to Hund's rules which govern the formation of atom magnetic moments; the inter-atomic exchange interactions are the source of the coupling between atom moments which allows magnetic order to be understood.

On the Connection between Exchange Interactions and Molecular Field

Within the Heisenberg theory, the exchange interaction between two spins i and j is generally

expressed as

$$e_{ij} = -2J_{ij}S_iS_j \qquad [13]$$

where J_{ij} is the exchange integral which represents the strength of the coupling between spins i and j. For an ensemble of N_T spins, the energy density is obtained by summing over all spins:

$$E_{ech} = -\frac{N_T}{V}\sum_{i,j\neq i} J_{ij}\mathbf{S}_i\mathbf{S}_j \qquad [14]$$

where V is the sample volume. Assuming that all atoms are identical, S_i and S_j may be replaced by S. Thus, the magnetic energy density can be written as

$$E_{ech} = -\frac{N_T S}{V}\left(\sum_{j\neq i} J_{ij}\right)S \qquad [15]$$

The first part of the right-hand expression in eqn [15] is proportional to the magnetization whereas the second part is proportional to the molecular field. Thus, this expression is equivalent to the molecular field expression of energy (eqn [10]); it also manifests the local character of magnetic interactions.

Magnetic Anisotropy

In ferromagnetic materials, the magnetic moments are aligned along well-defined crystallographic directions. This is the phenomenon of magnetic anisotropy. It is explained by considering that the electrons belonging to the magnetic shells interact with other charges in the environment, which constitute the so-called crystalline electric field (CEF). Due to the orbit asymmetry, the coupling energy depends on the orientation of the orbit within the CEF. A favored orientation of the orbital magnetic moment is intrinsically linked to the favored orientation of the orbit. Since the spin and orbital moments are linked, the total moment aligns along a preferred direction. This is the phenomenon of magnetic anisotropy. The anisotropy is high in rare-earth systems where a large orbital moment exists.

In uniaxial systems, there exists a unique easy direction of magnetization and a hard plane, perpendicular to the unique axis. The anisotropy field is the amplitude of the field applied in the plane and required to force the moments into this plane. In systems of higher symmetry, such as the cubic systems, there exist several easy directions and the anisotropy is weaker in general than in uniaxial systems.

Mechanisms of Magnetic Coupling

Exchange interactions constitute the general source of moment coupling. However, the details of the coupling mechanism may differ strongly from one system to another. In this section, this is illustrated by some representative examples.

Consider 3*d* electrons in iron, cobalt, nickel, or one of their alloys. The exchange-coupling existing between all electrons sitting on the same site favors a parallel coupling between their spins (Hund's rules). However, the 3*d* electrons have a certain degree of itinerant character, that is, they have a finite occupancy probability on the atom sites of the environment. An electron, arriving on another site, becomes exchange coupled to the electrons present on this site and again a parallel coupling between spins is favored. Thus, a parallel coupling between the two associated moments is favored. This mechanism explains high-temperature ferromagnetism observed in these systems.

In transition metal oxides, such as MnO, Fe_2O_3, CoO, or NiO, the fully localized 3*d* electrons are hybridized with the 2*p* electrons of oxygen. The *d* electrons, sitting on two different transition metal atoms but coupled to *p* electrons on the same oxygen atom, become indirectly coupled together. This mechanism constitutes super exchange. Often, it favors an antiparallel coupling between transition metal spins, that is, the exchange integral J is negative. Nonferromagnetic structures may result.

The 4*f* electrons of the rare-earth elements occupy an internal electronic shell, which is protected from the environment by more external shells, such as the 5*d* and 6*s* ones. The 4*f* electrons are strongly localized and bear well-defined magnetic moments. No interaction exists with 4*f* electrons on other sites. However, the 4*f* electrons are exchange coupled to the 5*d* and 6*s* electrons which become polarized. In metallic systems, the 5*d*, 6*s* electrons are itinerant. When they travel from one atom to the next, their polarization oscillates with a period which is characteristic of the band structure. Indirect interactions between the localized moments result in the RKKY type. They are the source of an indirect coupling between the 4*f* moments which is long-ranged and oscillatory.

Non-Ferromagnetic Types of Magnetic Orders

The possible existence of negative J values explains that diverse types of magnetic orders may exist, in addition to ferromagnetism.

Antiferromagnetism

Consider a system which contains one type of magnetic atoms, all bearing the same magnetic moment, but which may be separated into two groups, A and

B, called sublattices. Within the same sublattice, the exchange integrals J_{AA} and J_{BB} (coupling moments) are positive, whereas J_{AB} (the coupling moments belonging to different sublattices) are negative. The magnetic structure is such that all moments within a given sublattice are parallel to each other and antiparallel to the moments within the other sublattice. The resulting global magnetization is zero. Such a moment arrangement characterizes antiferromagnetism (**Figure 2b**). Amongst many other systems, it is found in the transition metal oxides MnO, CoO, or NiO.

Antiferromagnetism may be described within the molecular field model, by decomposing the molecular field into two separate contributions, one from each sublattice. Assuming that the two sublattices are identical, the molecular field may be expressed as

$$\boldsymbol{H}_{\mathrm{m}} = nM_{\mathrm{s}}^{\mathrm{sl}} - n'M_{\mathrm{s}}^{\mathrm{sl}} \qquad [16]$$

where $M_{\mathrm{s}}^{\mathrm{sl}}$ is the spontaneous magnetization within each sublattice, $n\,(>0)$ is the molecular field coefficient representative of the interactions within a given sublattice, and $n'\,(<0)$ is the molecular field coefficient representative of the interactions between sublattices. The coefficients n and n' are related to the J's through expressions which are similar to expression [19]. The magnetization of each sublattice is given by expression [3], but with

$$x = \frac{\mu_0 m(nM_{\mathrm{s}}^{\mathrm{sl}} - n'M_{\mathrm{s}}^{\mathrm{sl}})}{k_{\mathrm{B}}T} \qquad [17]$$

One looks for solutions of expressions [3] and [16] simultaneously. Above a critical temperature, T_{N}, called the Néel temperature, which is the equivalent of the Curie temperature for ferromagnetism, the sublattice magnetization is equal to zero; the slope of the reciprocal susceptibility versus T is the same as found in noninteracting and ferromagnetic systems, equal to $1/C$ (**Figure 1b**). At $T = T_{\mathrm{N}}$, magnetic order sets in. Below T_{N}, a nonzero magnetization exists within each sublattice.

In the ordered magnetic state, an applied magnetic field tends to align all moments along itself; it works against inter-sublattice exchange-coupling. As temperature decreases, the ratio of exchange-coupling energy to Zeeman energy increases. Thus, the magnetic susceptibility decreases. Combining this with the usual temperature decrease of the susceptibility in the paramagnetic state, it results that the magnetic susceptibility presents a maximum at T_{N}. This constitutes a common feature of antiferromagnetic-like structures, in which the global magnetization is zero.

From the point of view of quantum mechanics, Landau noted that the antiferromagnetic configuration described above, termed a Néel state, is not an eigenstate of the corresponding Hamiltonian. The actual ground state is a mixture of the two equivalent Néel states which are obtained by reversing the magnetization in the two sublattices; it is not magnetic. Due to the macroscopic nature of the antiferromagnetic domains, the Néel state is almost indistinguishable from the real ground state. Although not yet experimentally demonstrated, this is not the case in nanosystems.

Ferrimagnetism

Many transition metal oxides incorporate two different magnetic elements. The magnitude of the associated moments differs and a certain magnetization remains even when the moments are coupled antiparallel. Such ferrimagnetic structures (**Figure 2c**) are found in Fe oxides, such as γ-Fe_2O_3 and Fe_3O_4, where Fe moments in different ionic states coexist.

Ferrimagnetism is also found in rare-earth compounds with Fe, Co, or Ni, where an antiparallel coupling between the $3d$ moments and the rare-earth moments occur.

Helimagnetism and Modulated Structures

In the molecular field description of ferromagnetism, the molecular field is assumed to be a macroscopic entity. However, the local character of the magnetic interactions is established by experimental results which were mentioned above, and it emerges naturally in the Heisenberg description of the exchange interactions.

In the version of the molecular field model developed for the description of antiferromagnetism, the distinction between two different contributions gives to the molecular field a certain local character. In the spirit of a local description, the molecular field on a certain site, i, may be expressed in terms of the sum of the interactions of the moment, m_i, with all other j moments:

$$\boldsymbol{H}_i = \sum_j n_{ij} \langle m_j \rangle_T \qquad [18]$$

where the n_{ij}'s are molecular field coefficients. By analogy with the analyses described above, the magnetic moment, at temperature T, is given by

$$\langle m_i \rangle_{\mathrm{T}} = m_0 \mathscr{L}\left(\frac{\mu_0 m_0(H_i)}{k_{\mathrm{B}}T}\right) \qquad [19]$$

The high-temperature expansion of this expression is

$$T\langle m_i \rangle_T = C\sum_j n_{ij}\langle m_j \rangle_T \qquad [20]$$

with solutions of the type

$$\langle m_i \rangle_T = A\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{R}_i) \qquad [21]$$

where $\boldsymbol{R}_i$ describes the atom positions and $\boldsymbol{k}$ is a vector. One defines the quantity $\mathscr{T}(\boldsymbol{k})$ which is the Fourier transform of the exchange interactions:

$$\mathscr{T}(\boldsymbol{k}) = C\sum_j n_{ij} \exp(\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j) \qquad [22]$$

The critical temperature is given by

$$T_c = \mathscr{T}(\boldsymbol{k}_0) \qquad [23]$$

where T_c is the highest temperature for which the set of eqn [20] admits a nonzero solution, and the associated vector $\boldsymbol{k}_0$ which characterizes the magnetic structure is termed the propagation vector of the structure.

This approach may be illustrated by considering a linear chain of atoms, distant by a, such that the exchange interactions are positive between first neighbors (n_1), negative between second neighbors (n_2), and equal to zero for larger distances. For $n_1 \gg n_2$, the ferromagnetic solution is obtained (for n_1 large and negative, the structure is antiferromagnetic), but this is no more the case for $|n_2| > n_1/4$. When the magnetic moments are confined in the plane perpendicular to the chain by the magnetic anisotropy, a helimagnetic structure is found in which the moments rotate progressively from one to the next, along the chain (**Figure 2d**). This structure is found in some rare-earth metals, such as terbium or dysprosium. Reciprocally, when the moments are forced to align along the chain axis, a modulated structure occurs in which the amplitude of the moments oscillates periodically (**Figure 2e**). This occurs in thulium.

Helimagnetic and modulated structures belong to the large class of antiferromagnetic structures, which are characterized by the zero value of the spontaneous magnetization. However, in these two cases, the stable magnetic structure is the result of a compromise between the various interactions involved, and is said to be frustrated. The analysis of frustrated structures has recently been the topic of intense research, both theoretically and experimentally.

Spin-Glasses and Other Frustrated Systems

When the magnetic atoms do not form a periodic arrangement, another type of frustrated magnetic arrangement may exist. This is the case in substitutional alloys in which two or more types of atoms are randomly distributed on a given periodic lattice, or in amorphous alloys, obtained by very fast quenching from the melt. Since the magnitude and sign of interactions depend on the distance, it may be expected that the random distribution of atoms leads to a distribution of J values which is centered at 0 such that there are as many positive and negative interactions.

As long as the interactions n_{ij} are evenly distributed from positive to negative, with zero mean value, no stable magnetic order may exist. However, if eqn [20] is squared, one obtains

$$T^2 \langle m_i \rangle^2 = C^2 \left(\sum_j n_{ij} \langle m_j \rangle \right)^2 \qquad [24]$$

For a fully disordered system of infinite dimensions, the sum over the terms $n_{ij}n_{ij'} \langle m_j \rangle \langle m_{j'} \rangle$ is zero and eqn [24] can be expressed as

$$T^2 \langle m_i \rangle^2 = C^2 \left(\sum_j n_{ij}^2 \langle m_j \rangle^2 \right) \qquad [25]$$

This expression has a critical temperature given by

$$T_c = C \left(\sum_j n_{ij}^2 \right)^{1/2} \qquad [26]$$

Below T_c, the system is in a spin-glass state where the moments are randomly oriented. Equation [25] does not apply rigorously to finite-size systems. It has been a long debate to decide whether or not a real phase transition occurs in spin glasses. Experimentally, a peak is observed in the thermal variation of the magnetic susceptibility, but the temperature at which it occurs is found to depend on the time taken for the measurement (**Figure 3**). Initially, this type of measurement suggested that the spin-glass transition is not a real phase transition. Actually, it is now accepted that the spin glass transition constitutes a new class

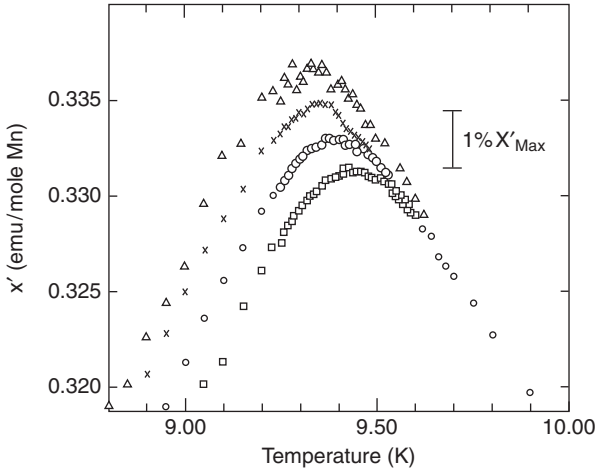


Figure 3 Thermal variation of the susceptibility in a spin glass: a Cu–Mn alloy containing 0.95% of Cu. The measurements of the susceptibility were realized under AC field of different frequencies (□ 1330 Hz, ○ 234 Hz, X 10.4 Hz, △ 2.6 Hz). The frequency dependence of the temperature at which the peak occurs in $\chi(T)$ is related to the fact that the critical slowing down of the magnetic fluctuations extends in spin glasses much farther from the value of the critical temperature than it does in other ordered magnetic systems.

of phase transitions in which the first derivative of the magnetic susceptibility versus temperature does not show any anomaly, but derivatives of higher orders do. The frequency dependence of the temperature at which the peak occurs in $\chi(T)$ is related to the fact that the critical slowing down of the magnetic fluctuations extends in spin glasses much farther from the value of the critical temperature than it does in other ordered magnetic systems.

Experimental Studies of Magnetic Structures

The macroscopic magnetic properties of magnetic materials may be derived from the magnitude of the stray field generated by the material magnetization, whether it is spontaneous or induced by the applied field. However, such measurements do not allow the moments of the various constitutive elements to be determined individually. Additionally, in the absence of applied field, antiferromagnetic-like materials do not generate any stray field. For this type of characterization, local probes such as Mössbauer spectroscopy in the case of Fe systems or NMR may be used. Neutron diffraction constitutes an invaluable tool for the study of magnetic arrangements. The neutron bears a small magnetic moment which probes the magnetism existing within matter. Neutron diffraction is the strict analog for the analysis of magnetic structures as X-ray diffraction is for the studies of crystallographic structures. Neutron diffractograms obtained with MnO are shown in **Figure 4**. In (a), the peak observed at 12° is due to the (1 1 1) reflection; it is characteristic of the face-centered cubic crystallographic structure of this compound. In (b), an additional peak is observed at ~6°. It can be indexed as 1/2, 1/2, 1/2. It is the signature of periodicity doubling. In an antiferromagnetic structure, the alternation of moment orientations leads to periodicity doubling (see scheme in the inset to **Figure 4**). The diffractograms shown in **Figure 4** have a special historical meaning. They constituted, when they were published, the first experimental proof of the existence of antiferromagnetism. Today, neutron diffraction has developed to a very sophisticated level; it allows very complex magnetic arrangements, including those found in spin glasses and other disordered systems, to be examined.

X-rays interact with the electrons within matter. Since spin–orbit coupling makes the bridge between electron distribution and magnetism, they probe magnetism as well. Although the magnetic interaction term is very small, magnetic studies using X-rays (X-ray magnetic circular dichroism and X-ray diffraction) have recently developed, thanks to the availability of very intense X-ray sources at synchrotron radiation facilities. A specific feature of X-ray studies is element selectivity, which occurs when the wavelength of the incident radiation corresponds to the characteristic energy of a given absorption edge for some element contained in the material examined.

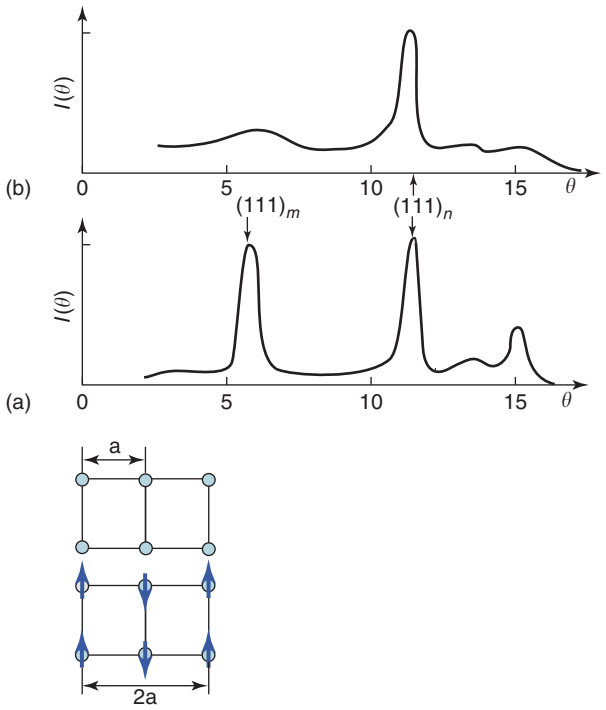


Figure 4 Neutron diffractograms obtained with MnO. (a) At 300 K, the material is paramagnetic. The observed (1 1 1) peak near $\theta = 12$ is characteristic of the cubic crystallographic structure. (b) The additional peak near $\theta = 6$ observed at 120 K reveal a doubling of the unit cell. It may be associated with the existence of an antiferromagnetic arrangement of the moments (inset).

See also: Disordered Magnetic Systems; Ferromagnetism; Magnetic Interactions.

PACS: 75.10. – b

Further Reading

Bacon GE (1962) *Neutron Diffraction*. Oxford University Press.
Chikazumi S (1997) *Physics of Ferromagnetism*. Oxford Science Publisher.
Coles BR (1984) Spin glasses. In: Pynn R and Skjeltorp A (eds.) *Multicritical Phenomena*. Plenum.
de Lacheisserie E, Gignoux D, and Schlenker M (2002) *Magnetism*, vol. 1 and 2. Kluwer Academic.
Morrish AH (1965) *Physical Principles of Magnetism*. Wiley.
Mydosh JA (1993) *Spin Glasses: An Experimental Introduction*. Taylor and Francis.

Nomenclature

$\mathscr{B}$	Brillouin function
C	Curie constant

g_J	Landé factor
$\boldsymbol{H}_{\mathrm{app}}$	applied magnetic field (vector)
H_{app}	applied magnetic field (magnitude)
H_{m}	molecular field
J	exchange integral
$\boldsymbol{J}$	total angular momentum
$\boldsymbol{k}_0$	propagation vector
l	orbital quantum number
$\boldsymbol{l}$	orbital momentum
$\boldsymbol{L}$	orbital momentum
$\mathscr{L}$	Langevin function
m	magnetic quantum number
m_0	measured magnetic moment at 0 K
$\boldsymbol{m}$	magnetic moment
$\boldsymbol{m}_l$	orbital moment
$\boldsymbol{m}_{\mathrm{s}}$	spin moment
$\boldsymbol{m}_z$	projection of m along H_{app}
M	magnetization
M_{s}	spontaneous magnetization
$M_{\mathrm{s},0}$	magnetization at absolute saturation
n	molecular field coefficient
N	number of moments per unit volume
N_{D}	demagnetizing field coefficient
N_{T}	total number of moments
S	spin or spin angular momentum
T	temperature
T_{c}	critical temperature
T_{C}	Curie temperature
T_{N}	Néel temperature
v_{at}	atom volume
x	argument of the Langevin or Brillouin function
z	number of first neighbors
χ	magnetic susceptibility
μ_{B}	Bohr magneton
μ_{eff}	effective moment
μ_0	permittivity of vacuum

Magnetic Point Groups and Space Groups

R Lifshitz, Tel Aviv University, Tel Aviv, Israel

Introduction

Magnetic groups – also known as antisymmetry groups, Shubnikov groups, Heesch groups, Opechowski–Guccione (OG) groups, as well as dichromatic, 2-color, or simply black-and-white groups – are the simplest extension to standard point group and space group theory. They allow directly to describe, classify, and study the consequences of the symmetry of crystals, characterized by having a certain property, associated with each crystal site, that can take one of two possible values. This is done by introducing a single external operation of order two that interchanges the two possible values everywhere throughout the crystal. This operation can be applied to the crystal along with any of the standard point group or space group operations, and is denoted by adding a prime to the original operation. Thus, any rotation g followed by this external operation is denoted by g'.

To start with, a few typical examples of this two-valued property are given, some of which are illustrated in **Figure 1**. In the section "Magnetic point groups" the notion of a magnetic point group is discussed, followed by a discussion on magnetic space groups in the section "Magnetic space groups". The section "Extinctions in neutron diffraction of antiferromagnetic crystals" describes one of the most

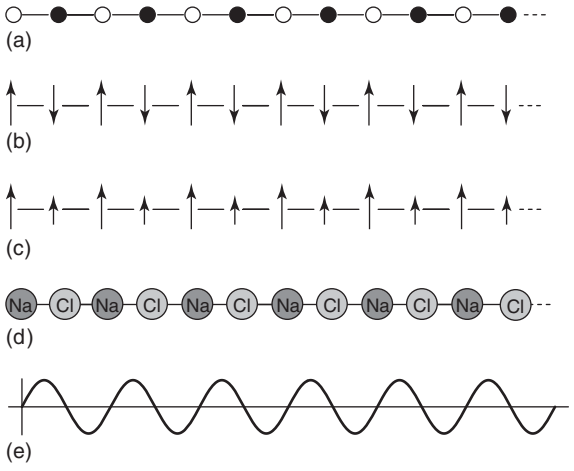


Figure 1 Several realizations of the simple 1D magnetic space group $p_p\bar{1}$. Note that the number of symmetry translations is doubled if one introduces an operation e' that interchanges the two possible values of the property, associated with each crystal site. Also note that spatial inversion $\bar{1}$ can be performed without a prime on each crystal site, and with a prime $(\bar{1}')$ between every two sites. (a) An abstract representation of the two possible values as the two colors – black and white. The operation e' is the nontrivial permutation of the two colors. (b) A simple antiferromagnetic arrangement of spins. The operation e' is time inversion which reverses the signs of the spins. (c) A ferromagnetic arrangement of two types of spins, where e' exchanges them as in the case of two colors. (d) A 1D version of salt. e' exchanges the chemical identities of the two atomic species. (e) A function $f(x)$ whose overall sign changes by application of the operation e'.

direct consequences of having magnetic symmetry in crystals which is the extinction of magnetic Bragg peaks in neutron diffraction patterns. A conclusion is given in the section "Generalizations of magnetic groups" by mentioning the generalization of magnetic groups to cases where the property associated with each crystal site is not limited to having only one of two possible values.

Consider first the structure of crystals of the cesium chloride type. In these crystals the atoms are located on the sites of a body-centered cubic (b.c.c.) lattice. Atoms of one type are located on the simple cubic lattice formed by the corners of the cubic cells, and atoms of the other type are located on the simple cubic lattice formed by the body centers of the cubic cells. The parent b.c.c. lattice is partitioned in this way into two identical sublattices, related by a body-diagonal translation. One may describe the symmetry of the cesium chloride structure as having a simple cubic lattice of ordinary translations, with a basis containing one cesium and one chlorine atom. Alternatively, one may describe the symmetry of the crystal as having twice as many translations, forming a b.c.c. lattice, half of which are primed to indicate that they are followed by the external operation that exchanges the chemical identities of the two types of atoms. A similar situation occurs in crystals whose structure is of the sodium chloride type, in which a simple cubic lattice is partitioned into two identical face-centered cubic (f.c.c.) sublattices.

Another typical example is the orientational ordering of magnetic moments (spins), electric dipole moments, or any other tensorial property associated with each site in a crystal. Adopting the language of spins, if these can take only one of two possible orientations – "up" and "down" – as in simple antiferromagnets, the same situation as above is seen, with the two spin orientations replacing the two chemical species of atoms. In the case of spins, the external prime operation is a so-called antisymmetry operation that reverses the signs of the spins. Physically, one may think of using the operation of time inversion to reverse the directions of all the spins without affecting anything else in the crystal.

Finally, consider a periodic or quasiperiodic scalar function of space $f(\boldsymbol{r})$, as in **Figure 1e**, whose average value is zero. It might be possible to extend the point group or the space group describing the symmetry of the function $f(\boldsymbol{r})$ by introducing an external prime operation that changes the sign of $f(\boldsymbol{r})$.

For the sake of clarity, the picture of up-and-down spins, and the use of time inversion to exchange them shall be adopted for the remaining of the discussion. It should be emphasized, though, that most of what is said here (except for the discussion of extinctions) applies equally to all the other two-valued properties. The magnetic crystal is described using a scalar spin density field $S(\boldsymbol{r})$ whose magnitude gives the magnitude of an atomic magnetic moment, or some coarse-grained value of the magnetic moment, at position $\boldsymbol{r}$. The sign of $S(\boldsymbol{r})$ gives the direction of the spin – a positive sign for up spins and a negative sign for down spins. Thus, the function $S(\boldsymbol{r})$ can be a discrete set of delta functions defined on the crystal sites as in **Figure 1b**, or a continuous spin density field as in **Figure 1e**.

Magnetic Point Groups

A d-dimensional magnetic point group G_M is a subgroup of $O(d) \times 1'$, where $O(d)$ is the group of d-dimensional rotations, and $1'$ is the time inversion group containing the identity e and the time inversion operation e'. Note that the term "rotation" is used to refer to proper as well as improper rotations such as mirrors.

Three cases exist: (1) all rotations in G_M appear with and without time inversion; (2) half of the rotations in G_M are followed by time inversion and the other half are not; and (3) no rotation in G_M is followed by time inversion. More formally, if G is any subgroup of $O(d)$, it can be used to form at most three types of magnetic point groups, as follows:

1. $G_M = G \times 1'$. Thus, each rotation in G appears in G_M once by itself, and once followed by time inversion. Note that in this case $e' \in G_M$.
2. $G_M = H + g'H$, where $G = H + gH$, and $g \notin H$. Thus, exactly half of the rotations in G, which belong to its subgroup H, appear in G_M by themselves, and the other half, belonging to the coset gH, are followed by time inversion. Note that in this case $e' \notin G_M$.
3. $G_M = G$. Thus, G_M contains rotations, none of which are followed by time inversion.

Enumeration of magnetic point groups is thus straightforward. Any ordinary point group G (i.e., any subgroup, usually finite, of $O(d)$) is trivially a magnetic point group of type 3, and along with the time inversion group gives another magnetic point group of type 1, denoted as $G1'$. One then lists all distinct subgroups H of index 2 in G, if there are any, to construct additional magnetic point groups of type 2. These are denoted either by the group–subgroup pair $G(H)$, or by using the International (Hermann–Mauguin) symbol for G and adding a prime to those elements in the symbol that do not belong to the subgroup H (and are therefore followed by time inversion). The set of magnetic groups (here magnetic

point groups G_M), formed in this way from a single ordinary group (here an ordinary point group G), is sometimes called the magnetic superfamily of the group G. For example, the orthorhombic point group $G = mm2 = \{e, m_x, m_y, 2_z\}$ has three subgroups of index 2 ($H_1 = \{e, m_x\}$, $H_2 = \{e, m_y\}$, and $H_3 = \{e, 2_z\}$), of which the first two are equivalent through a relabeling of the x and y axes. This yields a total of four magnetic point groups: $mm2$, $mm21'$, $m'm2'$ (equivalently expressed as $mm'2'$), and $m'm'2$.

Magnetic point groups of type 3 are equivalent to ordinary nonmagnetic point groups and are listed here only for the sake of completeness. Strictly speaking, they can be said to describe the symmetry of ferromagnetic structures with no spin–orbit coupling. Groups of this type are not considered here any further. Magnetic point groups of type 2 can be used to describe the point symmetry of finite objects, as described in the next section, as well as that of infinite crystals, as described in the section "Magnetic point groups of crystals." Magnetic point groups of type 1 can be used to describe the point symmetry of infinite crystals, but because they include e' as a symmetry element they cannot be used to describe the symmetry of finite objects, as is now defined.

Magnetic Point Groups of Finite Objects

The magnetic symmetry of a finite object, such as a molecule or a cluster containing an equal number of two types of spins, can be described by a magnetic point group. One can say that a primed or unprimed rotation from $O(d) \times 1'$ is in the magnetic point group G_M of a finite object in d dimensions, if it leaves the object invariant. Clearly, only magnetic point groups of type 2, listed above, can describe the symmetry of finite magnetically ordered structures. This is because time inversion e' changes the orientations of all the spins, and therefore cannot leave the object invariant unless it is accompanied by a nontrivial rotation. It should be mentioned that in this context, point groups of type 1 are sometimes called "gray" groups, as they describe the symmetry of "gray" objects which stay invariant under the exchange of black and white.

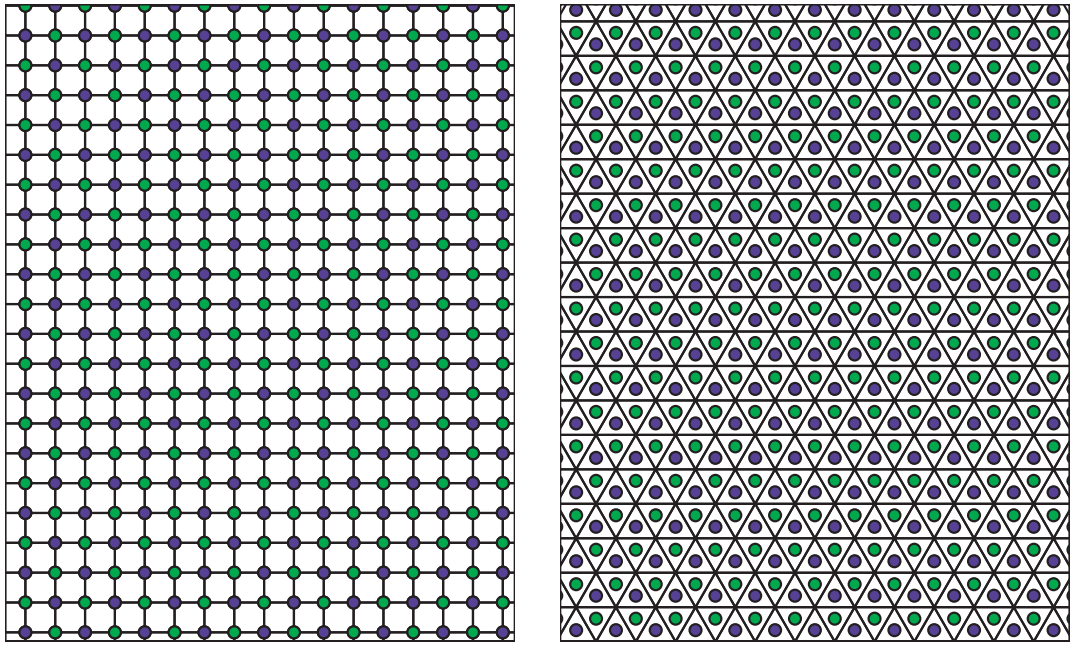
Magnetic Point Groups of Crystals

The magnetic point group of a d-dimensional magnetically ordered "periodic crystal", containing an equal number of up and down spins, is defined as the set of primed or unprimed rotations from $O(d) \times 1'$ that leave the crystal "invariant to within a translation". The magnetic point group of a crystal can be either of the first or of the second type listed above. It is of type 1 if time inversion by itself leaves the crystal invariant to within a translation. Recall that in this case, any rotation in the magnetic point group can be performed either with or without time inversion. If time inversion cannot be combined with a translation to leave the crystal invariant, the magnetic point group is of type 2, in which case half the rotations are performed without time inversion and the other half with time inversion.

Figure 2a shows an example of a magnetically ordered crystal whose magnetic point group is of type 1. This is a square crystal with magnetic point group $G_M = 4mm1'$ where time inversion can be followed by a translation to leave the crystal invariant. **Figure 2b** shows an example of a crystal whose magnetic point group is of type 2. This is a hexagonal crystal with point group $G_M = 6'mm'$. Note that all right-side-up triangles contain a blue circle (spin up), and all up-side-down triangles contain a green circle (spin down). Time inversion, exchanging the two types of spins, cannot be combined with a translation to recover the original crystal. Time inversion must be combined with an operation, such as the sixfold rotation or the horizontal mirror that interchanges the two types of triangles, to recover the original crystal. Note that the vertical mirror (the first m in the International symbol) leaves the crystal invariant without requiring time inversion, and is therefore unprimed in the symbol.

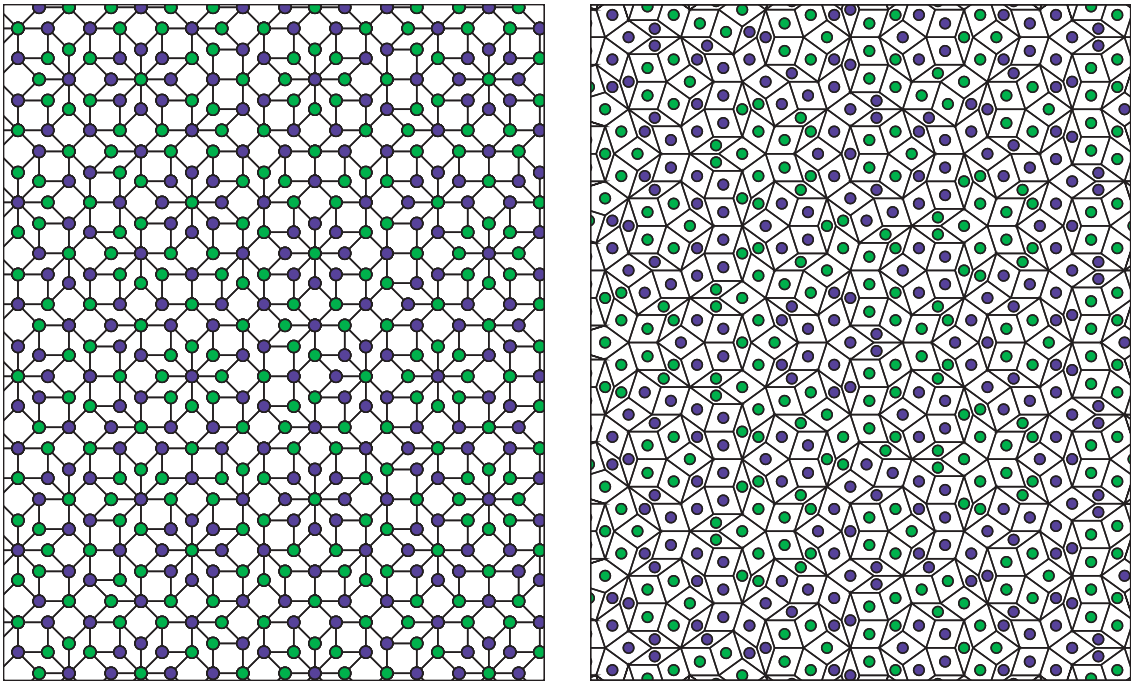
More generally, the magnetic point group of a d-dimensional magnetically ordered "quasiperiodic crystal (quasicrystal)," is defined as the set of primed or unprimed rotations from $O(d) \times 1'$ that leave the crystal "indistinguishable." This means that the rotated and unrotated crystals contain the same spatial distributions of finite clusters of spins of arbitrary size. The two are statistically the same though not necessarily identical. For the special case of periodic crystals, the requirement of indistinguishability reduces to the requirement of invariance to within a translation.

Figure 3 shows two quasiperiodic examples, analogous to the two periodic examples of **Figure 2**. **Figure 3a** shows an octagonal crystal with magnetic point group $G_M = 8mm1'$. One can see that time inversion rearranges the spin clusters in the crystal, but they all still appear in the crystal with the same spatial distribution. This is because any finite spin cluster and its time-reversed image appear in the crystal with the same spatial distribution. **Figure 3b** shows a decagonal crystal with magnetic point group $G_M = 10'm'm$. In this case, time inversion does not leave the crystal indistinguishable. It must be combined either with odd powers of the tenfold rotation, or with mirrors of the vertical type.



(a) (b)

Figure 2 Periodic antiferromagnets. (a) A square crystal with magnetic point group $G_M = 4mm1'$ and magnetic space group p_p4mm (also denoted p_c4mm), (b) A hexagonal crystal with magnetic point group $G_M = 6'mm'$ and magnetic space group $p6'mm'$.



(a) (b)

Figure 3 Quasiperiodic antiferromagnets. (a) An octagonal quasicrystal with magnetic point group $G_M = 8mm1'$ and magnetic space group p_p8mm, first appeared in an article by Niizeki (1990) *Journal of Physics A: Mathematical and General* 23: 5011. (b) A decagonal quasicrystal with magnetic point group $G_M = 10'm'm$ and magnetic space group $p10'm'm$ based on the tiling of Li, Dubois, and Kuo (1994) *Philosophical Magazine Letters* 69: 93.

Magnetic Space Groups

The full symmetry of a magnetically ordered crystal, described by a scalar spin density field $S(\boldsymbol{r})$, is given by its magnetic space group $\mathcal{G}_{\mathrm{M}}$. It was mentioned earlier that the magnetic point group G_{M} is the set of primed or unprimed rotations that leave a periodic crystal invariant to within a translation, or more generally, leave a quasiperiodic crystal indistinguishable. One still needs to specify the distinct sets of translations $\boldsymbol{t}_g$ or $\boldsymbol{t}_{g'}$ that can accompany the rotations g or g' in a periodic crystal to leave it invariant, or provide the more general characterization of the nature of the indistinguishability in the case of quasicrystals.

Periodic Crystals

As in the case of point groups, one can take any space group $\mathcal{G}$ of a periodic crystal and form its magnetic superfamily, consisting of one trivial group $\mathcal{G}_{\mathrm{M}} = \mathcal{G}$, one gray group $\mathcal{G}_{\mathrm{M}} = \mathcal{G}1'$, and possibly nontrivial magnetic groups of the form $\mathcal{G}_{\mathrm{M}} = \mathcal{H} + (g'|\boldsymbol{t}_{g'})\mathcal{H}$, where $\mathcal{H}$ is a subgroup of index 2 of $\mathcal{G}$. Only the latter, nontrivial groups are relevant as magnetic space groups consisting of operations that leave the magnetically ordered crystal "invariant." These nontrivial magnetic space groups are divided into two types depending on their magnetic point groups G_{M}:

1. *Class-equivalent magnetic space groups*. The magnetic point group G_{M} is of type 1, and therefore all rotations in G appear with and without time inversion. In this case the point groups of $\mathcal{G}$ and $\mathcal{H}$ are the same, but the Bravais lattice of $\mathcal{H}$ contains exactly half of the translations that are in the Bravais lattice of $\mathcal{G}$.
2. *Translation-equivalent magnetic space groups*. The magnetic point group G_{M} is of type 2. Here the Bravais lattices of $\mathcal{G}$ and $\mathcal{H}$ are the same, but the point group H of $\mathcal{H}$ is a subgroup of index 2 of the point group G of $\mathcal{G}$.

Enumeration of translation-equivalent magnetic space group types is very simple. Given a space group $\mathcal{G}$, all that one needs to do is to consider all subgroups H of index 2 of the point group G of $\mathcal{G}$, as explained earlier. The only additional consideration is that with an underlying Bravais lattice, there may be more than one way to orient the subgroup H relative to the lattice. For example, consider the magnetic point group $m'm2'$ which, as noted earlier, is equivalent to $mm'2'$. If the lattice is an A-centered orthorhombic lattice, then the x- and y-axes are no longer equivalent, and one needs to consider both options as distinct groups. It can be said that $Am'm2'$ and $Amm'2'$ belong to the same "magnetic geometric crystal class" (i.e., have the same magnetic point group), but to two distinct "magnetic arithmetic crystal classes." Translation-equivalent magnetic space groups are denoted by taking the International symbol for $\mathcal{G}$, and adding a prime to those elements in the symbol that do not belong to $\mathcal{H}$ and are therefore followed by time inversion.

Enumeration of class-equivalent magnetic space group types can proceed in two alternative ways. Given a space group $\mathcal{G}$, one may consider all distinct sublattices of index 2 of the Bravais lattice of $\mathcal{G}$. Alternatively, given a space group $\mathcal{H}$, one may consider all distinct superlattices of index 2 of the Bravais lattice of $\mathcal{H}$. It is also to be noted that if one prefers to look at the reciprocal lattices in Fourier space, instead of the lattices of real-space translations, the roles of superlattice and sublattice are exchanged. Although the lattice of $\mathcal{G}$ contains twice as many points as the lattice of $\mathcal{H}$, the reciprocal lattice of $\mathcal{G}$ contains exactly half of the points of the reciprocal lattice of $\mathcal{H}$.

Because of the different enumeration methods, there are conflicting notations in the literature for class-equivalent magnetic space groups. The OG notation follows the first approach (sublattices of $\mathcal{G}$), taking the International symbol for the space group $\mathcal{G}$ with a subscript on the Bravais class symbol denoting the Bravais lattice of $\mathcal{H}$. The Belov notation follows the second approach (superlattices of $\mathcal{H}$), taking the International symbol for the space group $\mathcal{H}$ with a subscript on the Bravais class symbol denoting one or more of the primed translations $\boldsymbol{t}_{e'}$ in the coset $(e'|\boldsymbol{t}_{e'})\mathcal{H}$. The Lifshitz notation, which follows a third approach (sublattices of $\mathcal{H}$ in Fourier space), resembles that of Belov but generalizes more easily to quasiperiodic crystals.

There is an additional discrepancy in the different notations which has been the cause of some confusion over the years. It is due to the fact that when $e' \in G_{\mathrm{M}}$, ordinary mirrors or rotations when unprimed may become glide planes or screw axes when primed, and vice versa. The only consistent way to avoid this confusion is to use only unprimed elements for the symbols of class-equivalent magnetic space group types. This is the approach adopted by both the Belov and the Lifshitz notations, but unfortunately not by the OG notation, having introduced errors into their list of magnetic space group types. In any case, there is no need to leave the $1'$ at the end of the symbol, as it is clear from the existence of a subscript on the lattice symbol that the magnetic point group G_{M} contains e'.

Consider, for example, all the class-equivalent magnetic space group types with point group 432. In the cubic system there are two lattice–sublattice

pairs: (1) The simple cubic lattice P has a face-centered sublattice F of index 2, and (2) The b.c.c. lattice I has a simple cubic sublattice P of index 2. Recall that the reciprocal of F is a b.c.c. lattice in Fourier space, denoted by I^*, and the reciprocal of I is an f.c.c. lattice, denoted by F^*. In the Belov notation one has F_s432, F_s4_j32, P_I432, and P_I4_j32 with $j = 1, 2, 3$. The subscript s is for "simple". In the Lifshitz notation, using Fourier space lattices, these groups become I_P^*432, $I_P^*4_j32$, $P_{F^*}432$, and $P_{F^*}4_j32$. In the OG notation these groups should become P_F432, P_F4_j32, I_P432, and I_P4_j32. Instead, they list P_F432, P_F4_232, I_P432, I_P4_132, $I_P4'32'$, and $I_P4_1'32'$, using primes inconsistently and clearly missing two groups. It is therefore recommended that one refrains from using primes when denoting class-equivalent magnetic space groups.

The reader is referred to the "Further reading" section for complete lists of magnetic space group types. However, some statistics are summarized. Starting from the 17 space group types in two dimensions, also known as the 17 plane groups, one can form a total of 80 magnetic space group types as follows: 17 trivial groups, 17 gray groups, and 46 nontrivial groups of which 18 are class-equivalent and 28 are translation equivalent. Starting from the 230 space group types in three dimensions, one can form a total of 1651 magnetic space group types as follows: 230 trivial groups, 230 gray groups, and 1191 nontrivial groups of which 517 are class-equivalent and 674 are translation equivalent. These numbers have no particular significance other than the fact that they are surprisingly large.

Quasiperiodic Crystals

Lattices and Bravais classes Magnetically-ordered quasiperiodic crystals, in general, possess no lattices of real space translations that leave them invariant, but they do have reciprocal lattices (or Fourier modules) L that can be inferred directly from their neutron diffraction diagrams. To be a bit more specific, consider spin density fields with well-defined Fourier transforms

$$S(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in L} S(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \qquad [1]$$

in which the set L contains, at most, a countable infinity of plane waves. In fact, the "reciprocal lattice" L of a magnetic crystal can be defined as the set of all integral linear combinations of wave vectors $\boldsymbol{k}$ determined from its neutron diffraction diagram, as one expects to see a magnetic Bragg peak at every linear combination of observed peaks, unless it is forbidden by symmetry, as explained in the next section. The "rank" D of L is the smallest number of vectors needed to generate L by integral linear combinations. If D is finite the crystal is quasiperiodic. If, in particular, D is equal to the number d of spatial dimensions, and L extends in all d directions, the crystal is periodic. In this special case, the magnetic lattice L is reciprocal to a lattice of translations in real space that leave the magnetic crystal invariant. The reciprocal lattices L of magnetic crystals are classified into Bravais classes, just like those of ordinary nonmagnetic crystals.

Phase functions and space group types The precise mathematical statement of indistinguishability, used earlier to define the magnetic point group, is the requirement that any symmetry operation of the magnetic crystal leaves invariant all spatially averaged nth-order autocorrelation functions of its spin-density field,

$$C^{(n)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n) = \lim_{V\to\infty} \frac{1}{V} \int_V \mathrm{d}\boldsymbol{r} S(\boldsymbol{r}_1 - \boldsymbol{r}) \cdots S(\boldsymbol{r}_n - \boldsymbol{r}) \qquad [2]$$

One can prove that two quasiperiodic spin density fields $S(\boldsymbol{r})$ and $\hat{S}(\boldsymbol{r})$ are indistinguishable in this way, if their Fourier coefficients, defined in eqn [1], are related by

$$\hat{S}(\boldsymbol{k}) = \mathrm{e}^{2\pi\mathrm{i}\chi(\mathrm{k})} S(\boldsymbol{k}) \qquad [3]$$

where χ is a real-valued linear function (modulo integers) on L, called a "gauge function". By this it is meant that for every pair of wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ in the magnetic lattice L,

$$\chi(\boldsymbol{k}_1 + \boldsymbol{k}_2) \equiv \chi(\boldsymbol{k}_1) + \chi(\boldsymbol{k}_2) \qquad [4]$$

where "$\equiv$" means equal to within adding an integer. Thus, for each element g or g' in the magnetic point group G_{M} of the crystal, that by definition leaves the crystal indistinguishable, there exists such a gauge function $\Phi_g(\boldsymbol{k})$ or $\Phi_{g'}(\boldsymbol{k})$, called a "phase function", satisfying

$$S(g\boldsymbol{k}) = \begin{cases} \mathrm{e}^{2\pi\mathrm{i}\Phi_g(\boldsymbol{k})} S(\boldsymbol{k}), & g \in G_{\mathrm{M}} \\ -\mathrm{e}^{2\pi\mathrm{i}\Phi_{g'}(\boldsymbol{k})} S(\boldsymbol{k}), & g' \in G_{\mathrm{M}} \end{cases} \qquad [5]$$

Since for any $g,h \in G$, $S([gh]\boldsymbol{k}) = S(g[h\boldsymbol{k}])$, the corresponding phase functions for elements in G_{M}, whether primed or not, must satisfy the "group compatibility condition",

$$\Phi_{g*h\dagger}(\boldsymbol{k}) \equiv \Phi_{g*}(h\boldsymbol{k}) + \Phi_{h\dagger}(\boldsymbol{k}) \qquad [6]$$

where the asterisk and the dagger denote optional primes. A "magnetic space group," describing the symmetry of a magnetic crystal, whether periodic or aperiodic, is thus given by a lattice L, a magnetic point group G_{M}, and a set of phase functions $\Phi_g(\boldsymbol{k})$

and $\Phi_{g'}(\boldsymbol{k})$, satisfying the group compatibility condition [6].

Magnetic space groups are classified in this formalism into magnetic space group types by organizing sets of phase functions satisfying the group compatibility condition [6] into equivalence classes. Two such sets, Φ and $\hat{\Phi}$, are equivalent if: (1) they describe indistinguishable spin density fields related as in [3] by a gauge function χ; or (2) they correspond to alternative descriptions of the same crystal that differ by their choices of absolute length scales and spatial orientations. In case (1), Φ and $\hat{\Phi}$ are related by a "gauge transformation,"

$$\hat{\Phi}_{g*}(\boldsymbol{k}) \equiv \Phi_{g*}(\boldsymbol{k}) + \chi(g\boldsymbol{k} - \boldsymbol{k}) \qquad [7]$$

where, again, the asterisk denotes an optional prime.

In the special case that the crystal is periodic ($D = d$), it is possible to replace each gauge function by a corresponding d-dimensional translation $\boldsymbol{t}$, satisfying $2\pi\chi(\boldsymbol{k}) = \boldsymbol{k} \cdot \boldsymbol{t}$, thereby reducing the requirement of indistinguishability to that of invariance to within a translation. Only then does the point group condition [5] become

$$S(g\boldsymbol{r}) = \begin{cases} S(\boldsymbol{r} - \boldsymbol{t}_g), & g \in G_{\mathrm{M}} \\ -S(\boldsymbol{r} - \boldsymbol{t}_{g'}), & g' \in G_{\mathrm{M}} \end{cases} \qquad [8]$$

the gauge transformation [7] becomes a mere shift of the origin, and the whole description reduces to that given in the section "Periodic crystals." The reader is referred to references in the "Further reading" section for details regarding the enumeration of magnetic space groups of quasiperiodic crystals and for complete lists of such groups.

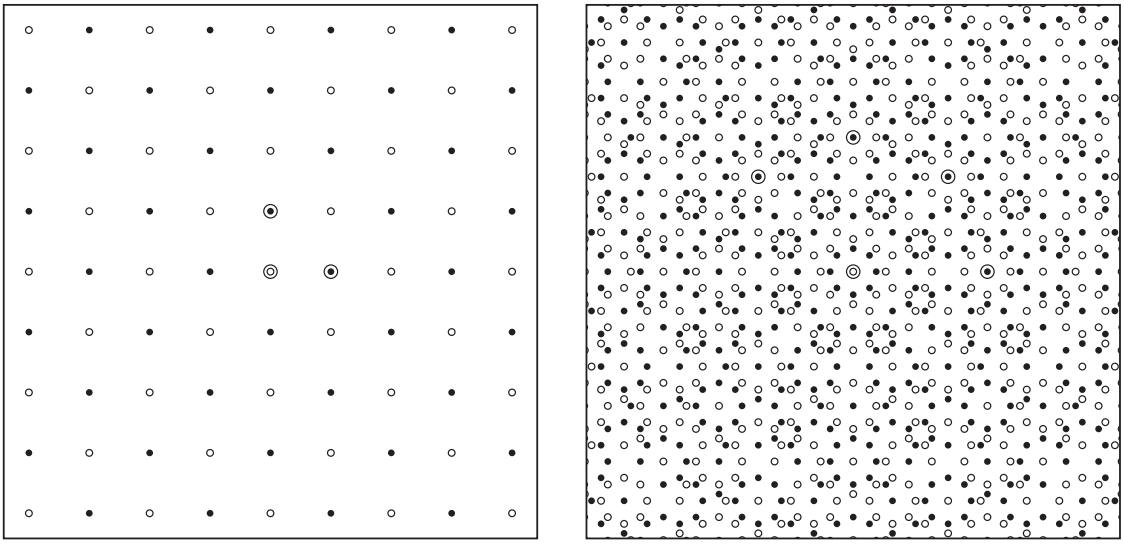
Extinctions in Neutron Diffraction of Anti-Ferromagnetic Crystals

It was said earlier that every wave vector $\boldsymbol{k}$ in the lattice L of a magnetic crystal is a candidate for a diffraction peak unless symmetry forbids it. One can now understand exactly how this happens. Given a wave vector $\boldsymbol{k} \in L$, all magnetic point-group operations g or g' for which $g\boldsymbol{k} = \boldsymbol{k}$ are examined. For such elements, the point-group condition [5] can be rewritten as

$$S(\boldsymbol{k}) = \begin{cases} \mathrm{e}^{2\pi\mathrm{i}\Phi_g(\boldsymbol{k})} S(\boldsymbol{k}), & g \in G_{\mathrm{M}} \\ -\mathrm{e}^{2\pi\mathrm{i}\Phi_{g'}(\boldsymbol{k})} S(\boldsymbol{k}), & g' \in G_{\mathrm{M}} \end{cases} \qquad [9]$$

requiring $S(\boldsymbol{k})$ to vanish unless $\Phi_g(\boldsymbol{k}) \equiv 0$, or $\Phi_{g'}(\boldsymbol{k}) \equiv 1/2$, or unless both conditions are satisfied when both g and g' are in G_{M}. It should be noted that the phase values in eqn [9], determining the extinction of $S(\boldsymbol{k})$, are independent of the choice of gauge [7], and are therefore uniquely determined by the magnetic space-group type of the crystal.

Particularly striking are the extinctions when the magnetic point group is of type 1, containing time inversion e'. The relation $(e')^2 = e$ implies – through



(a) (b)

Figure 4 Extinctions of magnetic Bragg peaks in crystals with magnetic point groups of type 1, containing time inversion e'. Both figures show the positions of expected magnetic Bragg peaks indexed by integers ranging from –4 to 4. Filled circles indicate observed peaks and open circles indicate positions of extinct peaks. The origin and the D vectors used to generate the patterns are indicated by an additional large circle. (a) corresponds to the square crystal in **Figure 2a** with magnetic space group p_p4mm, and (b) corresponds to the octagonal crystal in **Figure 3a** with magnetic space group p_p8mm.

the group compatibility condition [6] and the fact that $\Phi_e(\boldsymbol{k}) \equiv 0$ – that $\Phi_{e'}(\boldsymbol{k}) \equiv 0$ or $1/2$. It then follows from the linearity of the phase function that on exactly half of the wave vectors in L $\Phi_{e'}(\boldsymbol{k}) \equiv 1/2$, and on the remaining half, which form a sublattice L_0 of index 2 in L, $\Phi_{e'}(\boldsymbol{k}) \equiv 0$. Thus, in magnetic crystals with magnetic point groups of type 1 at least half of the diffraction peaks are missing. **Figure 4** shows the positions of the expected magnetic Bragg peaks corresponding to the square and octagonal magnetic structures shown in **Figures 2a** and **3a**, illustrating this phenomenon.

Generalizations of Magnetic Groups

There are two natural generalizations of magnetic groups. One is to color groups with more than two colors, and the other is to spin groups where the spins are viewed as classical axial vectors free to rotate continuously in any direction.

An n-color point group G_C is a subgroup of $O(d) \times S_n$, where S_n is the permutation group of n colors. Elements of the color point group are pairs (g, γ) where g is a d-dimensional (proper or improper) rotation and γ is a permutation of the n colors. As before, for (g, γ) to be in the color point group of a finite object it must leave it invariant, and for (g, γ) to be in the color point group of a crystal it must leave it indistinguishable, which in the special case of a periodic crystal reduces to invariance to within a translation. To each element $(g, \gamma) \in G_C$ corresponds a phase function $\Phi_g^\gamma(\boldsymbol{k})$, satisfying a generalized version of the group compatibility condition [6]. The color point group contains an important subgroup of elements of the form (e, γ) containing all the color permutations that leave the crystal indistinguishable without requiring any rotation g.

A spin point group G_S is a subgroup of $O(d) \times SO(d_s) \times 1'$, where $SO(d_s)$ is the group of d_s-dimensional proper rotations operating on the spins, and $1'$ is the time inversion group as before. Note that the dimension of the spins need not be equal to the dimension of space (e.g., one may consider a planar arrangement of 3D spins). Also note that because the spins are axial vectors there is no loss of generality by restricting their rotations to being proper. Elements of the spin point group are pairs (g, γ), where g is a d-dimensional (proper or improper) rotation and γ is a spin-space rotation possibly followed by time inversion. Here as well, elements of the form (e, γ) play a central role in the theory, especially in determining the symmetry constraints imposed by the corresponding phase functions $\Phi_e^\gamma(\boldsymbol{k})$ on the patterns of magnetic Bragg peaks, observed in elastic neutron diffraction experiments.

Acknowledgments

The author would like to thank David Mermin for teaching him the symmetry of crystals, and Shahar Even-Dar Mandel for joining him in his investigation of the magnetic symmetry of quasicrystals. The author's research in these matters is funded by the Israel Science Foundation through grant No. 278/00.

See also: Crystal Symmetry; Disordered Magnetic Systems; Magnetic Materials and Applications; Magnetoelasticity; Paramagnetism; Periodicity and Lattices; Point Groups; Quasicrystals; Space Groups.

PACS: 61.50.Ah; 61.12.Bt; 75.25.+z; 61.44.Br; 75.50.Kj; 61.68.+n; 02.20.Bb; 75.10.−b; 75.50.Ee

Further Reading

Lifshitz R (1996) The symmetry of quasiperiodic crystals. *Physica A* 232: 633–647.

Lifshitz R (1997) Theory of color symmetry for periodic and quasiperiodic crystals. *Reviews of Modern Physics* 69: 1181.

Lifshitz R and Even-Dar Mandel S (2004) Magnetically-ordered quasicrystals: Enumeration of spin groups and calculation of magnetic selection rules. *Acta Crystallographica A* 60: 167–178.

Litvin DB (1973) Spin translation groups and neutron diffraction analysis. *Acta Crystallographica A* 29: 651–660.

Litvin DB (1977) Spin point groups. *Acta Crystallographica A* 33: 279–287.

Litvin DB and Opechowski W (1974) Spin groups. *Physica* 76: 538–554.

Mermin ND (1992) The space groups of icosahedral quasicrystals and cubic, orthorhombic, monoclinic, and triclinic crystals. *Reviews of Modern Physics* 64: 3–49.

Opechowski W (1986) *Crystallographic and Metacrystallographic Groups*. Amsterdam: North-Holland.

Opechowski W and Guccione R (1965) Magnetic symmetry. In: Rado GT and Suhl H (eds.) *Magnetism*, vol. 2A, chap. 3. New York: Academic Press.

Schwarzenberger RLE (1984) Colour symmetry. *Bulletin of the London Mathematical Society* 16: 209–240.

Senechal M (1990) *Crystalline Symmetries: An Informal Mathematical Introduction*. Bristol: Adam Hilger.

Shubnikov AV, Belov NV, and others (1964) *Colored Symmetry*. New York: Pergamon Press.

Magnetism, History of

J M D Coey and T R Ní Mhíocháin, Trinity College, Dublin, Republic of Ireland

Introduction

Magnetism is a science with more than two millennia of recorded history. The attraction of ferrous objects to a permanent magnet across a distance has been a source of curiosity since the iron age. Feeble permanent magnets are widespread in Nature in the form of rocks known as lodestones, which are rich in the mineral magnetite (Fe_3O_4). Outcrops of these rocks can be magnetized by the huge currents in lightning strikes and such natural magnets were known and studied in ancient Greece, Egypt, China, and Mesoamerica. Investigations of magnetic phenomena led to the invention of steel magnets – needles and horseshoes – then to electromagnetics, and eventually to the panoply of hard and soft materials that underpin the modern magnetics industry. Furthermore, magnetism has played a key role in clarifying basic concepts in condensed matter physics over the course of the twentieth century.

Early History

Aristotle attributed the first reflections on the nature of magnetic attraction to Thales, who was born in Milet in Asia Minor in 624 BC. Thales, an animist, credited the magnet with a soul on account of its ability to create movement. This idea was to persist in Europe until the seventeenth century. The magnet itself is believed to be named after Magnesia, a city in Lydia, which was a source of lodestone. In the fifth century BC, Empedokles postulated the existence of four elements, earth, water, air, and fire, and related magnetism to air with special effluvia somehow passing through pores in magnetic material, a theory echoed much later by Descartes. In the first century BC, Roman poet Lucretius mentions magnetic induction, and the ability of magnets both to attract and repel. However, the Greek approach of developing philosophy into which natural observations were supposed to fit was not conducive to open-minded exploration of the natural world.

A more productive approach was followed in China, where magnetism was linked to geomancy. The art of adapting the residences of the living and the tombs of the dead to harmonize with local currents of the cosmic breath demanded a knowledge of direction. A south-pointer consisting of a carved lodestone spoon free to rotate on a polished baseplate (**Figure 1**) was in use in the first century BC, and may have been known hundreds of years earlier. An important discovery, attributed to Zheng Gongliang in 1064, was that iron could acquire a thermoremanent magnetization when rapidly cooled from red heat in the Earth's magnetic field. A short step led to the suspended compass, which was described by Shen Kua around 1088 together with declination, the deviation of a magnetic needle from a North–South axis. Floating compasses were also developed around this time. The compass appeared in Europe about a hundred years later, being referred to by Alexander Neckham in 1190. Compasses (**Figure 2**) were an enabling technology for the great voyages of discovery of the fifteenth century, leading the Ming admiral Cheng Ho to the coasts of Africa in 1433, and Christopher Columbus to America in 1492. Columbus rediscovered declination, and landed on the continent where the Olmecs may already have displayed a knowledge of magnetism in massive stone carvings of human figures and sea turtles dating from the second millennium BC.

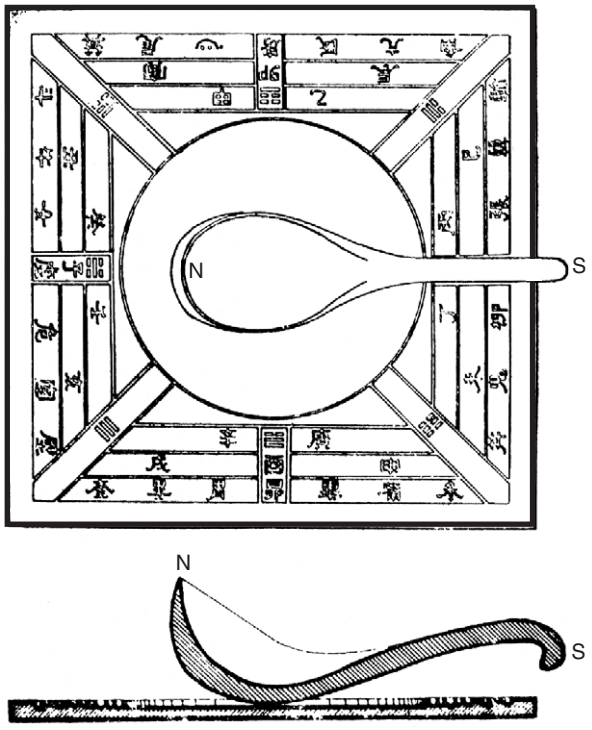


Figure 1 Baseplate and lodestone spoon of the south-pointer used in China from the first century BC. (Courtesy of Needham J (1962) *Science and Civilization in China*, vol. 4, part 1. Cambridge: Cambridge University Press.)

The first true experimentalist in Europe was the French crusader monk Petrus Peregrinus. In his paper "Epistola de Magnete," he recounted experiments

Figure 2 Fifteenth-century Portugese mariner's compass. (Courtesy of Boorstin D (1992) *Les Découvreurs*. Paris: Editions Robert Laffont.)

with floating pieces of lodestone and carved lodestone spheres called "terellae," which he had performed in Italy during the 1269 siege of Lucera. He discovered the poles of a magnet, related magnetic attraction to the celestial sphere, and included a proposal for a magnetic perpetual motion device, a common theme throughout the ages. However, credit for the inauguration of the experimental method in a recognizably modern form belongs to William Gilbert. Physician to the English queen Elizabeth I, Gilbert conducted a series of experiments on "terellae," and concluded that the Earth itself was a great magnet. The magnet itself did not align with the celestial sphere, but with the Earth's poles. He induced magnetism in iron by cooling in the Earth's field, and destroyed the magnetism by heating or hammering. He was at pains to dispel the accretion of superstition that clung to the magnet, preferring to rely on his own evidence. Gilbert's investigations are described in his masterwork "De Magnete" published in 1600. This was arguably the first scientific text.

From Navigation to Experimentation

Gilbert's theories dominated the seventeenth century until Edmund Halley's 1692 shell model for the Earth's magnetic structure, which greatly influenced compass technology and navigation. Magnetic research of the time was driven principally by naval interests. An early observation of a connection between electricity and magnetism was made in 1681, when a ship bound for Boston was struck by lightning. It was evident from the stars that "the compasses were changed; the north point was turn'd clear south," and with the compass so reversed, the ship was sailed to Boston.

The eighteenth century was marked by the professionalization of the scientist or "natural philosopher" as they were then known. Developments in magne tism were associated with improvements in naviga tion and the prestige of the great voyages of discovery. Accordingly, the natural philosopher with his mantle of theory was afforded social status, access to public funding, and credibility beyond that of artisans on the one hand, and "quacks" such as the colorful Anton Mesmer with his theories of "animal magnetism" on the other. Halley conducted magnetic surveys, producing charts of inclination and declination for seafarers (**Figure 3**). The English entrepreneur Gowin Knight, representative of this new breed of natural philosopher, greatly improved the quality of magnetic materials and compasses, coupling scientific endeavor with commercial enterprise.

The major technical breakthrough of the eighteenth century was the 1755 discovery of the horseshoe as the ideal magnet shape by the Swiss artisan Johann Dietrich. His invention, an ingenious practical solution to the problem of self-demagnetization in steel bar magnets, was enthusiastically promoted by Daniel Bernouilli.

This period also saw rapid developments in the harnessing of electricity, from the 1660 invention of the Leyden jar, culminating in Volta's 1800 invention of the Voltaic cell. A formal link between electricity and magnetism was not established until 1820 however, when during a public demonstration by Danish scientist Hans Christian Oersted, a suspended compass needle was seen to be deflected as he switched on an electric current. His report launched an experimental frenzy. Frenchman, François Arago immediately performed an experiment establishing that a current-carrying conductor behaved like a magnet. A week after Arago's report, André-Marie Ampère presented his paper to the French Academy suggesting that ferromagnetism in a body was caused by internal currents flowing perpendicular to the axis and that, by analogy, steel needles magnetized in a solenoid would show stronger magnetization. Together with Arago, he successfully demonstrated this theory in November 1820. Ten days later, the British scientist Humphrey Davy reported similar results. The electromagnet was invented by William Sturgeon in 1825, and 5 years later Joseph Henry used a powerful electromagnet in the USA to transmit the first electric telegraph.

Michael Faraday, however, revolutionized the nineteenth-century experimental science. He developed the concept of "magnetic field," working entirely by

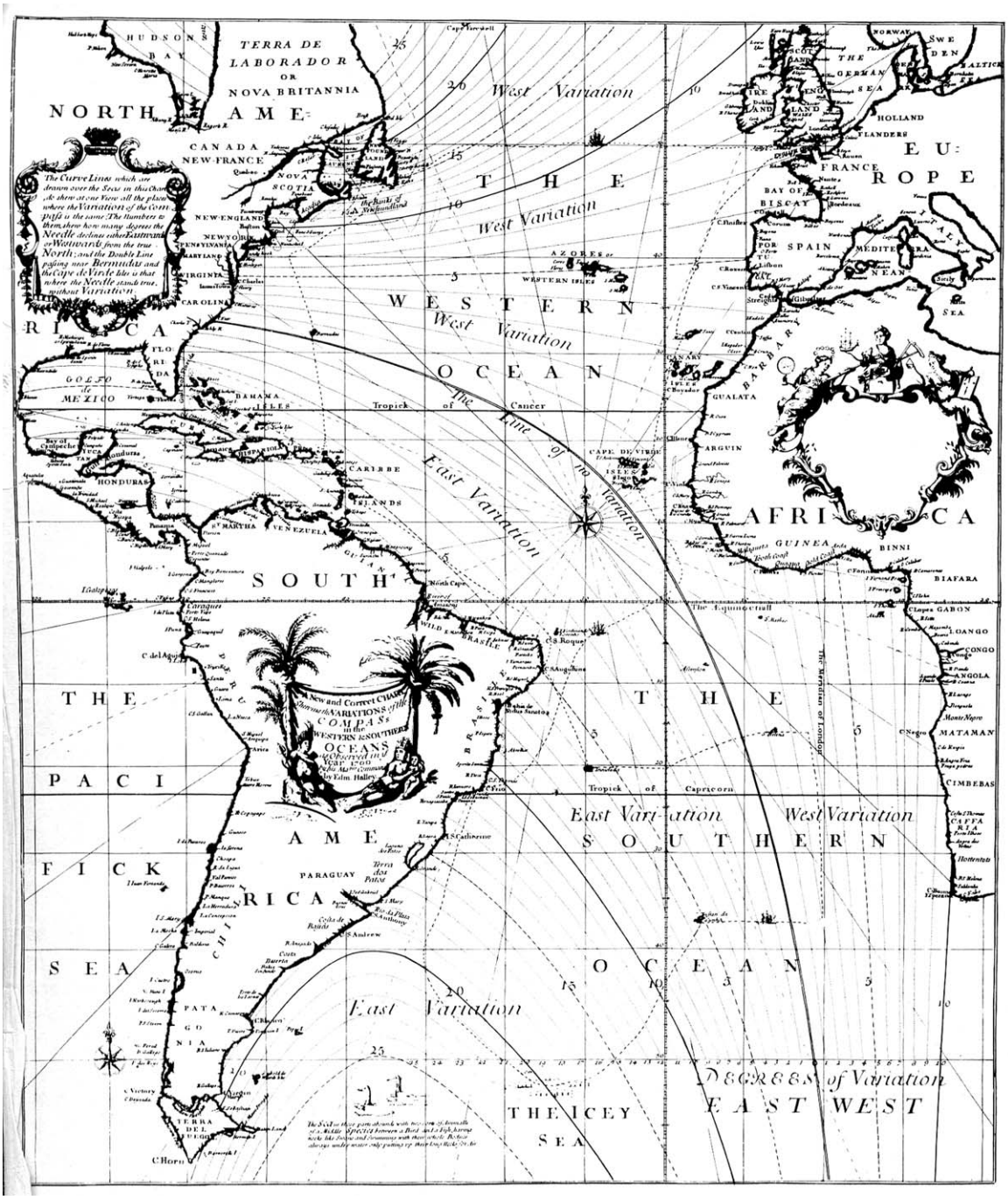


Figure 3 A section of Halley's world chart of magnetic variation published in 1700.

observation and experimentation, with no dependence on formal theory. Faraday classified materials as being either paramagnetic (substances such as iron which were easily magnetized and drawn toward stronger magnetic fields) or diamagnetic (materials which were not penetrated by magnetic fields and which were pushed away by the field). Working with an electromagnet, he discovered the phenomenon of "electromagnetic induction" which states that a flow of electricity can be induced by a changing magnetic field. His conviction that a magnetic field should have an effect on light led to his 1845 discovery of the Faraday effect which states that the plane of polarization of light is rotated upon passing through a medium in a direction parallel to its magnetization.

Figure 4 illustrates a collection of magnets from seventeenth to twentieth centuries.

Figure 4 Four centuries of magnets shown clockwise from the left: eighteenth-century horseshoe magnet; nineteenth-century electromagnet; seventeenth-century lodestone; and twentieth-century $Nd_2Fe_{14}B$.

A classical theory simultaneously emerged to account for magnetism and its relationship with electricity. Around 1824, Poisson developed a mathematical theory of "magnetostatics." Letters from Ampère to Fresnel and Faraday, found after his death, show that by 1822 he had considered the possibility, that the currents causing ferromagnetism were "molecular" rather than macroscopic in form. Weber formally presented the idea that molecules of iron were capable of movement around their centers, suggesting that they lay in different directions in an unmagnetized material, but aligned in one same direction in the presence of an applied magnetic field. This was the beginning of an explanation of hysteresis, the central phenomenon of magnetism, demonstrated experimentally by Ewing using a board of small, pivoting magnets. Ewing's achievements as a youthful professor at the University of Tokyo in the 1890s established the strong Japanese tradition of research on magnetic materials that still thrives today.

James Clerk Maxwell, the Scottish theoretician, "resolved to read no mathematics on the subject till he had first read through Faraday's 'Experimental Researches in Electricity.' " Maxwell's unprecedented equations formally defined the relationship between electricity and magnetism; that an oscillating electric charge produces an electromagnetic field. His theory considered electromagnetic waves traveling, through an all-pervading aether, at speeds close to that of light. In 1888, Heinrich Hertz showed that Maxwell's electromagnetic waves were, in fact, identical to those of light, radio, etc.

The century closed with Pierre Curie's 1895 discovery of the Curie point (the critical temperature above which materials lose their ferromagnetic behavior) and with the all-important discovery of the electron.

Units

By the middle of the nineteenth century, it became necessary to devise a standard set of units for electrical and magnetic quantities. The burgeoning telegraph industry, for example, needed a standard of electrical resistance to control the quality of electrical cables. Separate electrostatic and electromagnetic "cgs" unit systems based on the centimeter, gram, and second had come into existence. Maxwell and Jenkin proposed combining them in a coherent set of units in 1863. Their Gaussian (cgs) system was internationally adopted in 1881. Written according to this system, Maxwell's equations relating electric and magnetic fields contain factors of c, the velocity of light ($2.99792458 \times 108\,\mathrm{m\,s^{-1}}$). Maxwell also introduced the idea of dimensional analysis, based on three basic quantities of mass, length, and time. The magnetic field $\boldsymbol{H}$ and the induction $\boldsymbol{B}$ were measured in units of Oersted and Gauss, respectively, in the cgs system.

Another basic unit, of electric current, was adopted in the Système International d'Unités (SI) in 1948. The number of basic units and dimensions in any system is an arbitrary choice; SI uses four: the meter, kilogram, second, and ampere. It has been adopted worldwide for the teaching of science and engineering, employing the familiar electrical units of volt, ampere, and ohm. Maxwell's equations written in terms of two electric and two magnetic fields contain no factors of c or 4π in the SI system. The magnetic field $\boldsymbol{H}$, like the magnetization $\boldsymbol{M}$, has units of $\mathrm{A\,m^{-1}}$. The induction $\boldsymbol{B}$ is measured in tesla ($1\,\mathrm{T} \equiv 1\,\mathrm{kg\,s^2\,A^{-2}}$). Magnetic moments are measured in units of $\mathrm{A\,m^2}$, clearly indicating the origin of magnetism in electric currents, and the consequent absence of magnetic "poles" as real physical entities. The velocity of light is defined as exactly $299\,792\,458\,\mathrm{m\,s^{-1}}$. The two constants, the permeability, μ_0, and permittivity, ε_0, of free space are related by $\mu_0\varepsilon_0 = c^2$, μ_0 is defined to be $4\pi \times 10^{-7}\,\mathrm{kg\,s^2A^{-2}}$. At the present time, Gaussian cgs units remain in widespread use in magnetism, despite the manifest advantages of SI.

Fundamental Understanding – The Electron

The discovery of the electron in the closing years of the nineteenth century was a giant step toward a fundamental understanding of magnetism. The

elementary charged particle with mass $m_e = 9.11 \times 10^{-31}$ kg and charge $e = -1.60 \times 10^{-19}$ C had been named by the Irish scientist George Johnstone Stoney in 1891, several years before Jean Perrin identified negatively charged particles in a cathode ray tube, and Joseph John Thompson had measured their charge to mass ratio e/m_e by deflecting them in a magnetic field. In 1900 George Francis Fitzgerald suggested that magnetism might be due to rotation of these electrons. Electrons turned out to be the essential magnetic constituent of atoms and solids, as well as the carriers of electric current.

Thirty years were to elapse before contradictions at the heart of early twentieth century magnetism could be resolved. In these revolutionary years, classical physics and the lingering wisps of aether were blown away, and new principles of quantum mechanics and relativity were established. If, following Ampère, all magnetism is traceable to circulating electric currents, why does iron with a magnetization $M = 1.76 \times 10^6$ A m^{-1} not spontaneously melt? How can such a current be sustained indefinitely? In 1907, Pierre Weiss postulated the existence of an internal "molecular field" $\boldsymbol{H}_i$ proportional to the magnetization, $\boldsymbol{H}_i = n_W \boldsymbol{M}$, to account for the abrupt disappearance of ferromagnetism at the Curie point, T_C and paramagnetism at higher temperatures. He was at a loss, however, to explain why no trace of his enormous field ($n_W \approx 1000$) was detectable in the vicinity of a piece of fully magnetized ferromagnetic material. In the unmagnetized state, Weiss had correctly postulated the existence of domains magnetized in different directions to give no resultant moment. The "anomalous" Zeeman splitting of spectral lines in a magnetic field was a mystery. Bohr in 1911 and Van Leeuwen in 1919 had proved that at any finite temperature and in any magnetic or electric field, the net magnetization of any collection of classical electrons vanishes identically. Classical electron physics was incompatible with any kind of magnetism whatsoever!

By 1930, quantum mechanics and relativity had come to the rescue, and a new understanding of magnetism emerged in terms of the physics of Einstein, Bohr, Pauli, Dirac, Schrödinger, and Heisenberg. The source of magnetism in condensed matter was identified as being the "angular momentum" of elementary particles such as the electron. The perpetual currents in atoms were quantized in stationary states that did not decay, with angular momentum which was a multiple of Planck's constant $\hbar = 1.055 \times 10^{-34}$ J s. Weiss's "molecular field" was not a magnetic field at all, but a manifestation of electrostatic Coulomb interactions constrained by Pauli's exclusion principle, which forbade the occupancy of a quantum state by two electrons with the same spin. Spin, the intrinsic angular momentum of an electron, had been proposed by Compton in 1921; Goudsmit and Uhlenbeck demonstrated 4 years later that its value was $\hbar/2$. The corresponding electronic magnetic moment was one Bohr magneton, $\mu_B = e\hbar/2m_e$ or 9.24×10^{-24} A m^{-2}, twice as large as anticipated for orbital angular momentum. The problem of the electron's magnetism was finally resolved by Dirac in 1928 when he succeeded in writing Schrödinger's equation in a relativistically invariant form, obtaining the electron spin in terms of the 2×2 matrices previously proposed by Pauli. Together with Heisenberg, Dirac formulated the Hamiltonian $-2J\boldsymbol{S}_i \cdot \boldsymbol{S}_j$ to describe the coupling between the vector spins $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ of two many-electron atoms i and j. The value of the exchange integral J was directly proportional to Weiss's molecular field coefficient n_W.

Spin waves were found by Bloch and Slater in 1930 to be the elementary excitations of an array of atoms coupled by Heisenberg exchange interactions. The band theory of ferromagnetic metals developed by Slater and Stoner in the 1930s accounted for the nonintegral spin moments found in Fe, Co, and Ni, and their alloys.

The sixth Solvay conference, held in Brussels in October 1930 (**Figure 5**), was consecrated to magnetism. It followed 4 years of brilliant discovery in theoretical physics, which set out the modern electronic theory of matter. Yet, the immediate effect on the practical development of magnetism was slight. Dirac there made the perceptive remark "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."

Magnetic Phenomenology

A less fundamental theoretical approach was needed. Louis Néel pursued a phenomenological approach to magnetism, oblivious to the triumphs of quantum mechanics. His extension of the Weiss theory to two oppositely aligned sublattices led to the idea of antiferromagnetism in 1932. This hidden magnetic order awaited the development of neutron scattering in the 1950s before it could be directly revealed. Néel later explained the ferromagnetism of oxides such as magnetite (Fe_3O_4), the main constituent of lodestone, in terms of two unequal, antiferromagnetically coupled sublattices. The spinel structure of magnetite has an "A" sublattice of iron sites in tetrahedral oxygen coordination, and twice as many sites in octahedral coordination forming a "B" sublattice.

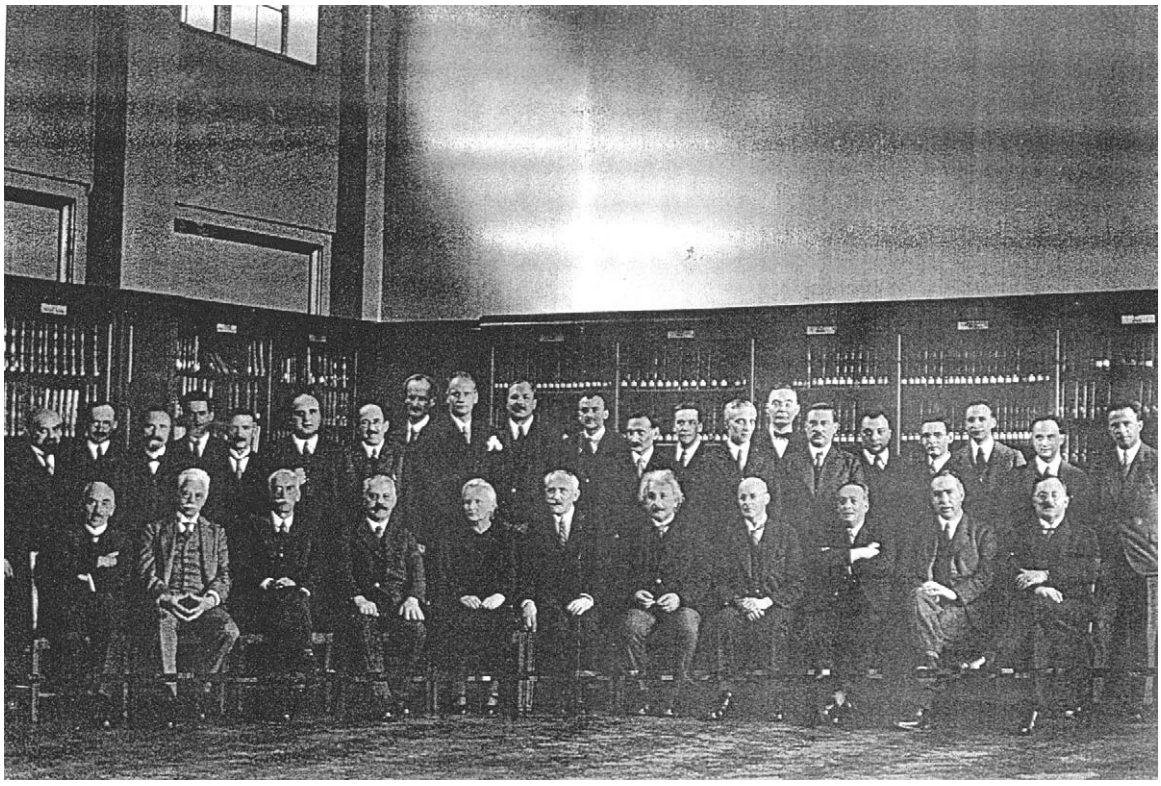


Figure 5 The 1930 Solvay Conference on Magnetism. (Courtesy of Université Libre de Bruxelles.) Back row: Herzen, Henriot, Verschaffelt, Manneback, Cotton, Errera, Stern, Piccard, Gerlach, Darwin, Dirac, Bauer, Kapitza, Brioullin, Kramers, Debye, Pauli, Dorfman, van Vleck, Fermi, Heisenberg. Front Row: de Donder, Zeeman, Weiss, Sommerfeld, Curie, Langevin, Einstein, Richardson, Cabrera, Bohr, de Haas.

The "B" sites are occupied by a mixture of Fe^{2+} and Fe^{3+}, and the "A" sites by oppositely aligned Fe^{3+}. This yields a spin moment of $4\mu_B$ per formula – a quantitative explanation of the magnetism of archetypical magnets in terms of lattice geometry and electron spin.

For many practical purposes, it is possible to ignore the atomic and electronic basis of magnetism, and simply regard the magnetization of a solid as a continuous vector. The hysteresis loop $\boldsymbol{M}(\boldsymbol{H})$ (**Figure 6**) is a record of the resultant of the metastable structure of domains and domain walls arising from the thermal and magnetic history of a particular sample. Evidence for discontinuous jumps in the Weiss domains as the magnetization was saturated was picked up by Barkhausen in 1919, and in 1931 the domains were visualized by Bitter. In 1932, Bloch introduced the idea of a domain wall as a region, where the magnetization rotates continuously from one direction to another. The exchange energy cost, written in the continuum approximation as $A(\nabla \boldsymbol{M})^2$, where $A \propto J$, is balanced by the gain in magnetostatic energy $(1/2)\int \mu_0 H^2 \, \mathrm{d}V$ associated with the magnetic field created by the ferromagnet throughout space, where V is the volume. Another term is needed to restrict the wall width; this is the anisotropy energy, which can arise intrinsically from spin–orbit coupling, a relativistic effect, or extrinsically from the sample shape. The sum of these free energy terms, together with the Zeeman energy in an external field, is minimized to yield the domain structure. This theory of "micromagnetism" was formulated by Brown in 1940.

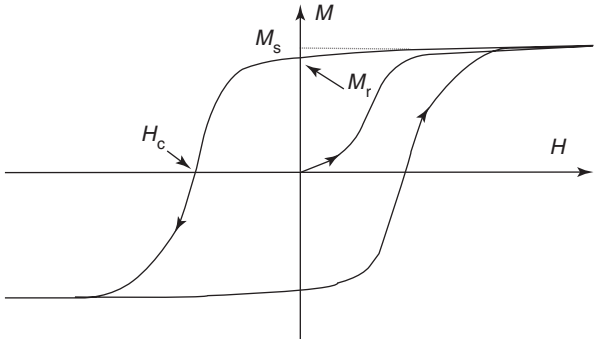


Figure 6 The hysteresis loop of $\boldsymbol{M}$ against $\boldsymbol{H}$ for a typical permanent magnet. M_s is the saturation magnetization, M_r the remanent magnetization at zero applied field, and H_c the coercive field required to reduce the magnetization to zero.

Domain walls in solids tend to be pinned by local crystal imperfections or microstructures, leading to

the metastable states represented on the hysteresis loop. The control of these has been due to as much metallurgical art as physical science. An important step was the development of a series of Co–Fe–Al–Ni alloys, the "alnicos" begun by Mishima in 1932. The mastery of coercivity gained over the course of the twentieth century (**Figure 7**) was spectacular.

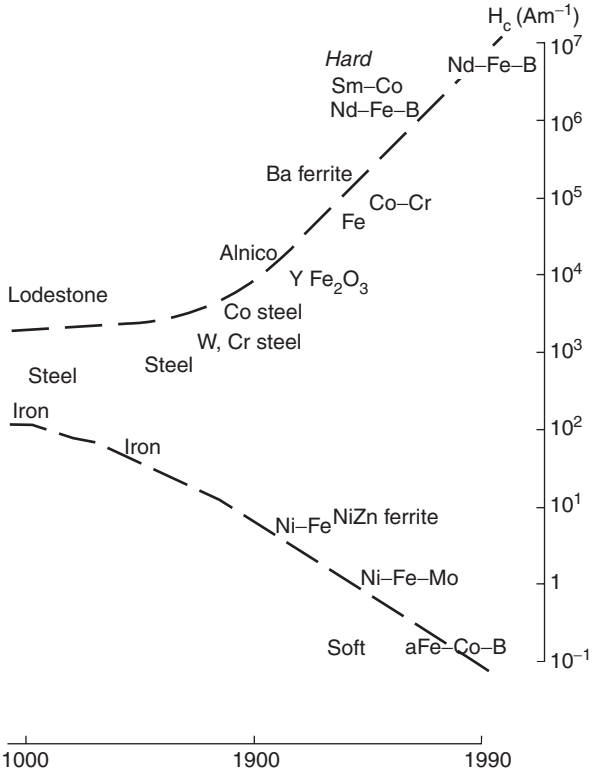


Figure 7 The development of coercivity in the twentieth-century.

Burgeoning applications of magnetism were a direct consequence. Microscopic quantum theory only began to contribute to this process in the 1970s after the advent of rare-earth permanent magnets such as $SmCo_5$ and $Nd_2Fe_{14}B$, when an understanding of the intrinsic, magnetocrystalline anisotropy helped in the design of new permanent magnet materials.

A fertile approach to the problem of hysteresis was to focus on magnetization reversal in single-domain particles. Stoner and Wolhfarth proposed an influential model for this in 1948. Meanwhile Néel, seeking to understand the remanent magnetism of baked clay, proposed a model of thermally driven "superparamagnetic" fluctuations of the magnetization of submicron-sized ferromagnetic particles the following year. The fluctuation time is dependent exponentially on the ratio of the energy barrier against magnetization reversal to the thermal energy. As a result, a component of magnetization became blocked on cooling in the Earth's magnetic field. Application of this idea of thermoremanent magnetization to cooling igneous rocks delivered a direct and convincing argument for continental drift, as rocks cooling at different times experienced fields of different polarity (**Figure 8**). This all led to the subfield of palaeomagnetism, and establishment of the theory of global plate tectonics, which had far-reaching consequences for Earth science.

Modern Developments

Magnetism since 1945 has been an area of rich and productive discovery, not least because of the enormous increase in numbers of scientists and engineers working in the field. Magnet ownership for citizens

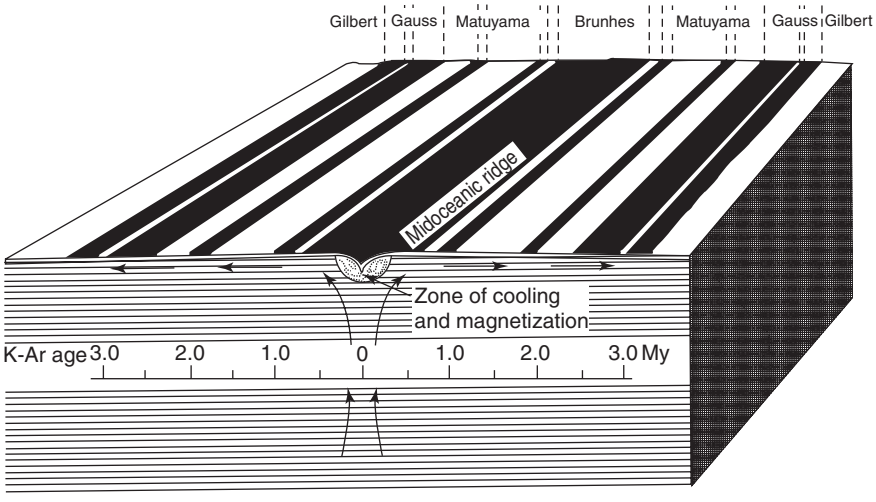


Figure 8 Sign of the thermoremanent magnetization measured across the floor of the Atlantic Ocean. Random reversals of the Earth's field, which are dated from other igneous lava flows, follow a chronological pattern which determines the rate of continental drift. (Courtesy of McElhinney MW (1973) *Palaeomagnetism and Plate Tectonics*. Cambridge: Cambridge University Press.)

of the developed world has skyrocketed from one or two magnets in 1945 to 100–200 sixty years later, or something in the order of a billion if one counts the individual magnetic bits on a hard disk in a personal computer! Countless citizens of the developing world share magnetism's bounty in the form of a cassette tape recorder.

Modern achievements can be classified as: (1) those contributing to a basic understanding of condensed matter, (2) those associated with new experimental probes, and (3) those leading to new materials and industrial technologies.

Theory

From a fundamental point of view, magnetism has proved a fruitful field for cultivating condensed matter theory. After yielding the first mean-field theory in the guise of Weiss's molecular field, magnetism has provided more sophisticated theories of phase transitions, and the data to support them. A fixed point was Onsager's 1944 solution of the two – dimensional Ising model, where spins are regarded as scalars which take values ± 1. The behavior of more complex and realistic systems such as the three-dimensional Heisenberg model was solved numerically using the renormalization group technique in the 1970s. The ability to tailor model magnetic systems having an effective spatial dimension of one or two with structures of chains or planes of magnetic ions and an effective spin dimension limited by anisotropy, has laid the foundation of the modern theory of phase transitions. The effect of disorder on antiferromagnetic interactions was studied intensively in spin glasses in the 1980s. The ground states of double chains and ladders continue to turn up surprises.

Correlations among the electrons in a metal, and their coupling to the crystal lattice, have generated a wealth of ground states and phenomena in solids such as superconductivity, heavy-fermion behavior, quantum Hall effect. A fecund line of enquiry was "Does a single impurity in a solid bear a magnetic moment?" This led to the definition, and eventually to the solution of the Kondo problem in 1980, which drew attention to the relation between magnetism and electronic transport in metals, which has been an immensely productive interface in terms of practical devices based on magnetoresistance. Long-range exchange interactions associated with the ripples of spin polarization created by a magnetic impurity in a metal led to an understanding of complex magnetic order in the rare-earth metals. Numerical methods, such as that based on spin-dependent density-functional theory, were successfully applied in the 1990s after the arrival of the cheap, powerful computers needed to calculate the magnetic ground-state properties of correlated electron systems.

Experimental Methods

The discovery of magnetic resonance, the sharp absorption of microwave or radio frequency radiation by Zeeman-split levels of an atom or nucleus in a magnetic field, was a landmark in modern magnetism. Significant mainly for the insight provided into solids and liquids at an atomic scale, electron paramagnetic resonance (EPR) and nuclear magnetic resonance (NMR) spectroscopies were complemented by that provided by the Mössbauer effect in 1958. Hyperfine interactions of nuclei and electrons provided a point probe of electric and magnetic fields at the center of the atom. In a magnetically ordered state, precession of the magnetic sublattices in a field leads to ferromagnetic resonance (FMR).

Of the new experimental probes of magnetism, neutron scattering was the most influential and generally useful. A beam of thermal neutrons from a nuclear reactor was first exploited for neutron diffraction by Shull and Wohlan in 1951, who used the magnetic scattering to reveal the antiferromagnetic order in MnO. Countless magnetic structures have been measured since, using the specially constructed facilities at Chalk River, Harwell, Brookhaven, Grenoble, and elsewhere. Excitations could be characterized by inelastic scattering of thermal neutrons, with the help of the three-axis spectrometer developed by Brockhouse. Complete spin–wave dispersion relations were used to provide a wealth of information on anisotropy and exchange.

Other beams produced at large-scale facilities that proved useful for probing magnetism were muons and synchrotron radiation. The intense, tunable ultraviolet and X-ray radiation from synchrotrons allowed the measurement of magnetic dichroism from deep atomic levels, and permitted the separate determination of spin and orbital contributions to the magnetic moment.

Increasing availability of commercial superconducting magnets from the late 1960s has led to diagnostic imaging of tissue based on NMR. Thousands of these scanners in hospitals across the world provide images of hearts, brains, and every sort of tumor. Superconducting magnets can now provide fields of over 20 T for NMR and general laboratory use. Coupled with SQUID sensors, ultrasensitive magnetometers capable of measuring magnetic moments of $10^{-11}\,\mathrm{A\,m^{-2}}$ are widely available. Larger fields of up to 30 T require expensive special installations with water-cooled Bitter magnets consuming many

megawatts of electrical power. Resistive/superconducting hybrids in Tallahassee, Grenoble, and Tokyo can generate steady fields in excess of 40 T. Still higher fields mean short pulses; the higher the field, the shorter the pulse.

Materials and Technology

A search for practical new magnetic materials led to the discovery of hard and soft ferrites in the early 1950s. The hard ferrites were the first materials whose coercive field exceeded their magnetization, making it possible for a magnet to have any desired shape, thereby breaking the centuries-old shape barrier by rendering bar and horseshoe magnets obsolete (**Figure 4**). The soft ferrites were indispensable for radiofrequency and microwave components. Rare-earth intermetallic compounds investigated from 1960 onward provided high-performance rare-earth permanent magnets, notably $Nd_2Fe_{14}B$ developed by Sagawa. Miniature electromagnetic drives that are incorporated in myriad consumer goods have been a consequence.

For many physical studies, magnetic materials were needed in special forms, especially single crystals or thin films. Crystal growers were assiduously cultivated by neutron scatterers and other condensed matter physicists. With large crystals, tensor properties such as susceptibility, magnetostriction, and magnetotransport could be measured in detail. After 1970, thin-film growth facilities (sputtering, pulsed-laser deposition, molecular beam epitaxy) began to appear in ultrahigh magnetism laboratories. Ultrahigh vacuum facilitated the study of surface magnetism at an atomic level, but much of the motivation was to investigate magnetooptics or magnetoresistance of metals. Thin films also contributed massively to magnetic storage. In recent years, nanoscale magnetic composites have extended the range of magnetic properties available in hard and soft magnets.

Magnetic recording dates from the late nineteenth century, when Poulsen recorded his voice on steel wire. It developed into analog tape recording in Germany in the 1930s. The first digital disk recording device made its appearance in the USA, launched by IBM, in 1956. A single bit then occupied $\sim 0.3\,mm^2$ on the disk surface and cost about 1¢. Fifty years later the same bit occupied $0.006\,\mu m^2$ and cost 0.000 000 005¢. This relentless miniaturization and perfection of rotating-disk technology for magnetic recording is a fascinating chapter in the history of technology (**Figure 9**). It was facilitated by a series of advances in thin film magnetic recording media with uniform crystallite size in the range 10–100 nm, and multilayer thin film read/write heads which could create and detect local magnetic fields. The read head had to pick up weak stray fields just above the disk surface. Inductive heads with coils were successively replaced by heads based on anisotropic magnetoresistance, and in 1997 by giant magnetoresistive heads. Giant magnetoresistance had been discovered by Fert and co-workers in an Fe–Cr multilayer in 1988. Reduced to its bare essentials, a sandwich of two ferromagnetic layers with a nonmagnetic metallic spacer known as a "spin valve," was needed to show the effect. Hundreds of millions were being produced annually by 2003. A variant that has been developed as a nonvolatile magnetic storage element is the magnetic tunnel junction used in magnetic random-access memory (MRAM). These magnetoresistive sensors and storage devices represented the first generation of spin electronic devices.

Future Prospects

Past record is no guarantee of future performance, so one must be wary of extrapolating the history of any science or technology very far into the future. Magnetism, like much of physics itself, has often

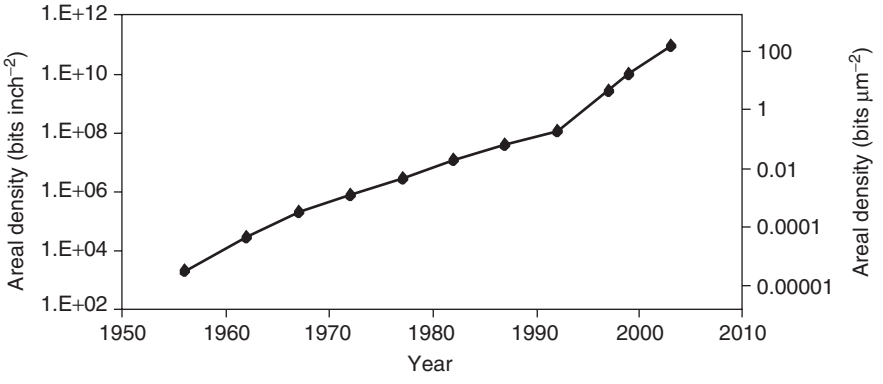


Figure 9 Exponential growth of magnetic recording density over the last 50 years.

been dismissed as "finished" in the past, only to confound the doomsayers with some new twist or unexpected development. The current focus on spin electronics and nanoscale magnetism will surely be sustained for some time to come, if only by momentum. A successful second generation of spin electronics may depend on the emergence of room temperature magnetic semiconductors, or at least effective spin injection into regular semiconductors.

Industrial applications of magnetism are based on the properties of about a dozen commercially produced magnetic materials. The discovery of just one more could animate the subject for a decade. It seems unlikely, though, that much advance will be achieved on the magnetic values on the hysteresis loop. There has been no advance in saturation magnetization in a century over that of $Fe_{35}Co_{65}$, but coercivity can be controlled, albeit with much care and effort, to almost any desired value. No Curie temperature has ever exceeded that of cobalt, which has been known since 1912. The new material will probably commend itself by a useful combination of magnetic and nonmagnetic (electrical, optical, mechanical, etc.) properties.

Magnetism has served the miniaturization of information technology and electromagnetic drives well, and may continue to do so in future. New frontiers in cell biology have to be explored. There have been half a dozen paradigm shifts – completely new ways of seeing the magnet and its field – during the 3000-year encounter of magnetism with human curiosity. Who can say there will never be another?

See also: Diamagnetism; Ferromagnetism; Magnetic Domains; Magnetic Materials and Applications; Paramagnetism.

PACS: 01.65.+g; 01.75.+m; 07.55.Db; 75; 91.25.−r; 78.20.Ls; 85.70.−w; 85.75.−d

Further Reading

Coey JMD (1996) *Rare-Earth Iron Permanent Magnets*. Oxford: Clarendon Press.

Daniel ED, Mee CD, and Clark MH (eds.) (1999) *Magnetic Recording, the First Hundred Years*. New York: IEEE Press.

Fara P (1996) *Sympathetic Attractions: Magnetic Practices, Beliefs and Symbolism in Eighteenth-Century England*. New Jersey: Princeton University Press.

Gilbert W (1958) *De Magnete* (English translation by Mottelay PF). New York: Dover Publications.

Kloss A (1994) *Geschichte des Magnetismus*. Berlin: VDE-Verlag.

Marage P and Wallenborn G (eds.) (1995) *Les Conseils Solvay et Les Débuts de la Physique Moderne*. Brussels: Université Libre de Bruxelles.

Matthis DC (1965) *Theory of Magnetism*, ch. 1. New York: Harper and Row.

Mottelay PF (1975) *Bibliographical History of Electricity and Magnetism*. New York: Arno Press.

Sin-itiro Tomonaga (1974) *The Story of Spin*. Chicago: University of Chicago Press.

Nomenclature

$\boldsymbol{B}$	magnetic induction
c	speed of light ($2.99792458 \times 10^{8}\,\mathrm{m\,s^{-1}}$)
e	charge on an electron ($1.602 \times 10^{-19}\,\mathrm{C}$)
$\hbar$	Planck's constant ($1.055 \times 10^{-34}\,\mathrm{J\,s}$)
$\boldsymbol{H}$	magnetic field
$\boldsymbol{J}$	exchange integral
m_e	mass of an electron ($9.109 \times 10^{-31}\,\mathrm{kg}$)
$\boldsymbol{M}$	magnetization
n_W	Weiss molecular field coefficient
$\boldsymbol{S}$	spin vector
T_C	Curie temperature
ε_0	electric permittivity of free space ($8.854 \times 10^{-12}\,\mathrm{C^2N^{-1}m^{-2}}$)
μ_0	magnetic permeability of free space ($4\pi \times 10^{-7}\,\mathrm{Wb\,A^{-1}m^{-1}}$)
μ_B	Bohr magneton ($9.274 \times 10^{-24}\,\mathrm{J\,T^{-1}}$)

Magnetocaloric Effect

V K Pecharsky and K A Gschneidner Jr., Iowa State University, Ames, IA, USA

Introduction

The word "magnetocaloric" is derived from Greek *Magnes* (*lithos*) (literally, stone of Magnesia, ancient city in Asia Minor), that is, magnesian (stone) or "magnetite" that has been known for thousands of years to attract articles made from iron, and Latin "calor" for heat, which evolved into French "calorique" – the name of the invisible and indestructible fluid that, according to a theory prevailing in 1700s, was thought to be responsible for virtually everything related to heat. This two-part word describes one of the most fundamental physical properties of magnetic materials: that is, when a solid is exposed to a varying magnetic field, its temperature

may be measurably increased or decreased with both the sign and the magnitude of the temperature difference between the final and the initial states of the material dependent on numerous intrinsic and extrinsic factors. The chemical composition, crystal structure, and magnetic state of a compound are among the most important intrinsic material parameters. The temperature, surrounding pressure, and sign of the magnetic field change are examples of extrinsic variables that affect these magnetic-field-induced temperature changes and, therefore, play a role in defining the magnetocaloric effect (MCE).

Inherent to every magnetic solid, the magnetocaloric effect has exceptional fundamental importance because it covers length, energy, and timescales spanning over many orders of magnitude: from quantum mechanics to micromagnetics, from statistical to macroscopic thermodynamics, and from spin dynamics to bulk heat flow and thermal conductivity. Mapping out this vast landscape is an enormously challenging task, yet even partial successes along the way facilitate a better understanding of how to design novel magnetic solids; in other words, they help in creating an environment in which a material could be tailored to exhibit a certain combination of magnetic and thermal properties. In addition to its basic significance, the magnetocaloric effect is the foundation of near-room-temperature magnetic refrigeration, which is poised for commercialization in the foreseeable future and may soon become an energy-efficient and environmentally friendly alternative to vapor-compression refrigeration technology. Practical applications of the magnetocaloric effect, therefore, have the potential to reduce the global energy consumption, and eliminate or minimize the need for ozone depleting, greenhouse, and hazardous chemicals.

Discovery and First Application of the Magnetocaloric Effect

MCE was originally discovered in iron by E Warburg, who reported the phenomenon in 1881. The nature of the MCE was explained and its usefulness to reach temperatures below 1 K by adiabatically demagnetizing a paramagnetic substance was suggested independently in the mid-1920s by P Debye and W F Giauque. Namely, after cooling a paramagnetic material in a large magnetic field to as low a temperature as possible by conventional means, such as pumping on liquid helium, the material was to be thermally isolated from its surroundings, the magnetic field removed, and the sample was predicted to cool well below 1 K due to the magnetocaloric effect.

Together with his student D P MacDougall, W F Giauque was the first to verify this prediction experimentally. As Giauque and MacDougall wrote in their 12 April 1933 letter to the editor of the *Physical Review*, "On March 19, starting at a temperature of about 3.4 K, the material cooled to 0.53 K. On April 8, starting at $\sim$2 K, a temperature of 0.34 K was reached. On April 9, starting at $\sim$1.5 K, a temperature of 0.25 K was attained." They achieved these breakthroughs by using 61 g of hydrated gadolinium sulfate, $Gd_2(SO_4)_3 \cdot 8H_2O$, and by reducing the magnetic field from 0.8 to 0 T. Giauque and MacDougal conclude their short (about 200 words) letter that was published in 1 May 1933 issue of the *Physical Review* with the following statement: "It is apparent that it will be possible to obtain much lower temperatures, especially when successive demagnetizations are utilized." This technique is known as adiabatic demagnetization refrigeration and it is successfully employed today in ultra-low-temperature research – for example, nuclear adiabatic demagnetization has been and is being used to reach micro-kelvin temperatures.

Fundamentals of the Magnetocaloric Effect

The MCE occurs due to the coupling of a magnetic sublattice with an external magnetic field, which affects the magnetic part of the total entropy of a solid. Similar to isothermal compression of a gas during which positional disorder and, therefore, the corresponding component of the entropy of a system are suppressed, isothermal magnetizing of a paramagnet near the absolute zero temperature or a ferromagnet near its spontaneous magnetic ordering temperature – the Curie temperature, T_C – greatly reduces disorder of a spin system, thus substantially lowering the magnetic part of the total entropy. In a reversible process, which resembles the expansion of a gas at constant temperature, isothermal demagnetization restores the zero field magnetic entropy of a system. The MCE, therefore, can be measured as an extensive thermodynamic quantity – the isothermal magnetic entropy change, ΔS_M, which is illustrated in **Figure 1** as a difference between the two entropy functions taken at the same temperature and is marked by a vertical arrow.

When a gas is compressed adiabatically, its total entropy remains constant whereas velocities of the constituent molecules, and therefore, the temperature of the gas both increase. Likewise, the sum of the lattice and electronic entropies of a solid must be changed by $-\Delta S_M$ as a result of adiabatically magnetizing (or demagnetizing) the material, thus resulting in an increase (decrease) of the lattice vibrations and the adiabatic temperature change, ΔT_{ad}, which is

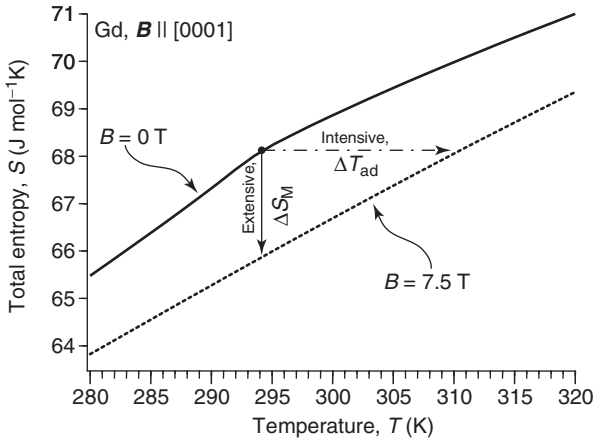


Figure 1 The total entropy of gadolinium in the vicinity of $T_C = 294$ K in 0 and 7.5 T magnetic fields. The data are for a high-purity single crystalline Gd with the magnetic field vector parallel to the [0001] crystallographic direction of the specimen. The vertical arrow illustrates ΔS_M; the horizontal arrow represents ΔT_{ad}.

an intensive thermodynamic quantity also used to measure and express the MCE. In **Figure 1**, ΔT_{ad} is illustrated as a difference between the two entropy functions taken at the same entropy and is indicated using a horizontal arrow. It is worth noting that in the case of a gas, it is a "change of pressure" that results either in the adiabatic temperature rise or drop, or the isothermal entropy change, while in the case of a magnetic solid it is a "change of the magnetic field" that brings about the MCE. Needless to say, no matter how strong the magnetic field around the sample, the MCE will remain zero as long as the field is kept constant.

For a given material at a constant pressure, the two quantitative characteristics of the magnetocaloric effect are functions of the absolute temperature (T) and the magnetic field change ($\Delta B = B_f - B_i$), where B_f and B_i are the final and initial magnetic fields, respectively, experienced by the material. The MCE can be easily computed provided the behavior of the total entropy (S) of a compound is known as a function of temperature in both the initial and final magnetic fields, e.g., see **Figure 1**:

$$\Delta S_M(T, \Delta B)_{\Delta B = B_f - B_i} = S(T, B)_{B=B_f} - S(T, B)_{B=B_i} \quad [1]$$

$$\Delta T_{ad}(T, \Delta B)_{\Delta B = B_f - B_i} = T(S, B)_{B=B_f} - T(S, B)_{B=B_i} \quad [2]$$

Equation [2], in which the entropy is the independent variable and the temperature is the dependent variable, is straightforwardly employed in direct measurements of ΔT_{ad}. Thus, the temperature of a sample is measured in both B_i and B_f, that is, before and after the magnetic field has been altered. The difference between the two temperatures yields the intensive MCE value, which is usually reported as a function of temperature for $B_i = 0$.

At equilibrium, both ΔS_M and ΔT_{ad} are correlated with the magnetization (M), magnetic flux density (B), heat capacity at constant pressure (C), and absolute temperature by one of the following fundamental Maxwell equations:

$$\Delta S_M(T, \Delta B)_{\Delta B = B_f - B_i} = \int_{B_i}^{B_f} \left(\frac{\partial M(T, B)}{\partial T} \right)_B dB \quad [3]$$

$$\Delta T_{ad}(T, \Delta B)_{\Delta B = B_f - B_i} = -\int_{B_i}^{B_f} \left(\frac{T}{C(T, B)} \times \frac{\partial M(T, B)}{\partial T} \right)_B dB \quad [4]$$

As immediately follows from eqns [1]–[4], materials whose total entropy is strongly influenced by a magnetic field and where the magnetization varies rapidly with temperature, are expected to exhibit an enhanced MCE. The latter peaks when $|(\partial M(T, B)/\partial T)_B|$ is the greatest, that is, around the T_C in a conventional ferromagnet or near the absolute zero temperature in a paramagnet. Usually, the MCE of a simple ferromagnet is gradually lowered both below and above the T_C, as is clearly seen in **Figure 2**.

Equations [3] and [4] are easily derived from general thermodynamics, yet both fail to describe the MCE in the vicinity of a truly discontinuous first-order phase transition when either or both $|[\partial M(T, B)/\partial T]_B|$ and $[T/C(T, B)]_B$ do not exist even if the phase volume change is negligibly small. (By definition, partial first derivatives of Gibbs free energy with respect to intensive thermodynamic variables, for example, T, P, or B, vary discontinuously at the first-order phase transition. As a result, the bulk magnetization is expected to undergo a discontinuous change at constant temperature; and the heat capacity is expected to be infinite during a first-order phase transformation. Thus, in theory, $[\partial M(T, B)/\partial T]_B$ and $[T/C(T, B)]_B$ do not exist at the temperature of the first-order transition. In reality, these changes occur over a few-kelvin-wide temperature range and both functions can be measured experimentally.) Equations [1] and [2], on the other hand, define the MCE regardless of the thermodynamic nature of the phase transformation that occurs, if any, in a material.

For a first-order phase transition, it is also possible to employ an approximation, which is based on the Clausius–Clapeyron equation:

$$\left(\frac{dB}{dT} \right)_{eq} = \left(\frac{\Delta S}{\Delta M} \right)_T \quad [5]$$

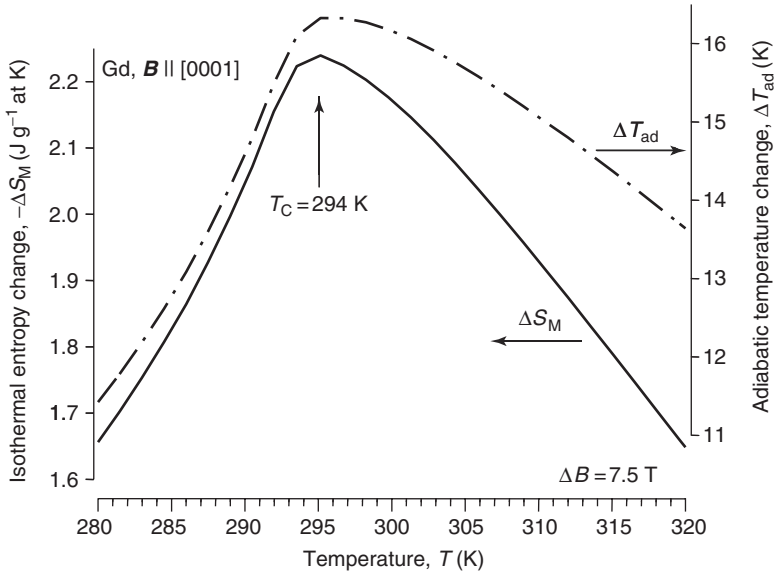


Figure 2 The extensive, ΔS_M, and intensive, ΔT_{ad}, MCE of the elemental Gd in the vicinity of its Curie temperature, $T_C = 294$ K, for a magnetic field change from 0 to 7.5 T. The data are for a high-purity single crystalline Gd with the magnetic field vector parallel to the [0001] crystallographic direction of the specimen.

In eqn [5], the left-hand-side derivative is taken under equilibrium conditions, that is, when Gibbs free energies of the two phases are identical to one another. For the right-hand side, $\Delta S = S_2 - S_1$ and $\Delta M = M_2 - M_1$, where the subscripts 1 and 2 correspond to the states of the material in the initial and final magnetic fields, respectively. Obviously, eqn [5] is only applicable when B_f is strong enough to complete the transformation from a state 1 to state 2 and when the quantity dB/dT at equilibrium is known. In other words, the B–T phase diagram for the system must be well established. By using the Clausius–Clapeyron equation, only an estimate of the extensive MCE, $\Delta S_M = \Delta S$, is possible.

Magnetic Order and the Magnetocaloric Effect

It is easy to see (Eqns [3] and [4]) that both ΔS_M and ΔT_{ad} are proportional to the derivative of bulk magnetization with respect to temperature at constant magnetic field. The ΔT_{ad} is also proportional to the absolute temperature and inversely proportional to the heat capacity at constant magnetic field. It is predictable, therefore, that any material should have the largest MCE when its magnetization is changing most rapidly with temperature. This, for example, occurs in the vicinity of a Curie temperature of a ferromagnet. (From eqn [4] it is not obvious that $\Delta T_{ad}(T)_{\Delta B}$ is at its maximum at the Curie temperature because heat capacity of a material ordering ferromagnetically also peaks around T_C. It can be shown, however, that the maximum of $\Delta T_{ad}(T)_{\Delta B}$ does indeed occur in the immediate vicinity of T_C, especially when $B_i = 0$ and $\Delta B \to 0$. By directly measuring $\Delta T_{ad}(T)_{\Delta B}$ in low magnetic fields, it is, therefore, possible to determine the Curie temperature of a ferromagnet.) The MCE will decrease below T_C, where the magnetization approaches saturation and becomes weakly dependent on temperature, and above the Curie temperature when the magnetization shows only a paramagnetic response, that is, it is proportional to T^{-1}, while its derivative maintains a proportionality to T^{-2}.

Conventional ferromagnets, therefore, typically display "caret-like MCE." This kind of behavior is easily recognizable in **Figure 2**, and it is also shown in **Figure 3a** for a $Gd_5Si_{2.5}Ge_{1.5}$ compound. The latter orders ferromagnetically via a second-order phase transformation at 312 K. As the ΔB increases, the magnitude of the MCE of a conventional ferromagnet also increases. The rate of change of the MCE with magnetic field, that is, the instantaneous derivative of the MCE with respect to ΔB is, however, the largest at the lowest B_f. The intensive measure of the MCE can vary from a fraction of 1 K T^{-1} to several K T^{-1}. Elemental Gd, for example, exhibits a $d[\Delta T(T, \Delta B)]/d(\Delta B)$ of ~ 3 K T^{-1} around T_C when B_f is on the order of 1 T. The derivative falls to ~ 1.8 K T^{-1} when B_f approaches 10 T. It is worth noting that away from absolute zero, $d[\Delta T(T, \Delta B)]/d(\Delta B)$ in conventional ferromagnetic materials is only weakly dependent on temperature (see **Figure 3a**).

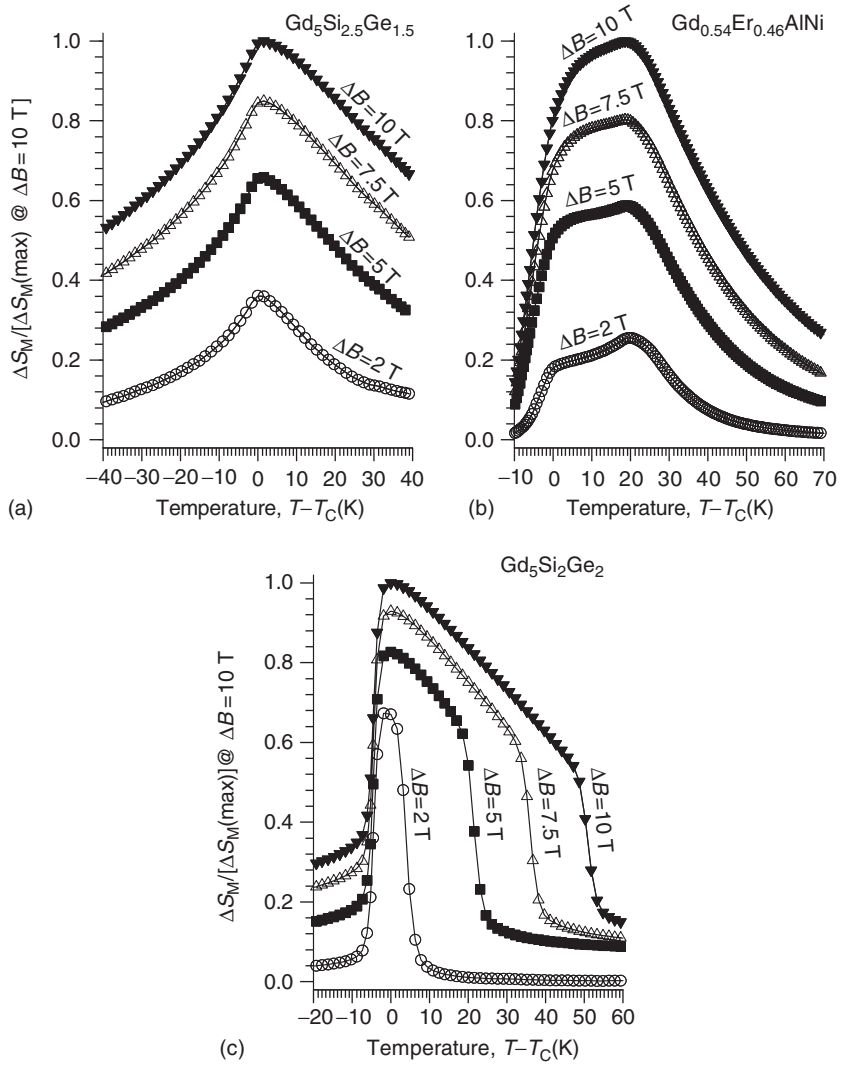


Figure 3 Normalized magnetocaloric effect of (a) $Gd_5Si_{2.5}Ge_{1.5}$ ($T_C = 312\,K$), (b) $Gd_{0.54}Er_{0.46}AlNi$ ($T_C = 14.5\,K$ (the normalization temperature) and 35 K) and (c) $Gd_5Si_2Ge_2$ ($T_C = 273\,K$) shown as a function of the reduced temperature for various magnetic field changes. In all cases, $B_i = 0$. The $\Delta S_M(max)$ values are $1.57\,J\,g^{-1}$, $0.63\,J\,g^{-1}$, and $2.61\,J\,g^{-1}$ at $\Delta B = 10\,T$ for $Gd_5Si_{2.5}Ge_{1.5}$, $Gd_{0.54}Er_{0.46}AlNi$, and $Gd_5Si_2Ge_2$, respectively.

Anomalous behaviors of the MCE are closely related to anomalous changes in the magnetic structures of solids causing unusual behaviors of the two functions, $[\partial M(T, B)/\partial T]_B$ and $[T/C(T, B)]_B$, which are carried over to both $\Delta S_M(T)_{\Delta B}$ and $\Delta T_{ad}(T)_{\Delta B}$ (see eqns [3] and [4]). One of the most commonly observed MCE anomalies is when a material undergoes two or more successive magnetic transitions in close proximity to one another. Then, instead of a conventional caret-like shape, a skewed caret, sometimes approaching a flat and an almost constant value as a function of temperature, that is, a "table-like MCE" can be observed. An example of such behavior is shown in **Figure 3b** for $Gd_{0.54}Er_{0.46}AlNi$. Both $\Delta S_M(T)_{\Delta B}$ (**Figure 3b**) and $\Delta T_{ad}(T)_{\Delta H}$ (not shown) of this material are almost constant between the two successive magnetic phase transition temperatures, which are $\sim 20\,K$ apart. Anomalous MCE behaviors may also be due to low-lying crystalline electric fields and strong quadrupolar and/or magnetoelastic interactions. Furthermore, the MCE can be adjusted over a broad range of behaviors by mixing two or more individual compounds or phases, each with a simple caret-like MCE, into a composite where the resulting isothermal magnetic entropy change will be a weighted sum of the corresponding values of the individual components.

Generally, a more complicated magnetic behavior results in an anomalous and a more complicated behavior of the MCE. For instance, upon cooling in a zero magnetic field, elemental Dy orders antiferromagnetically at $\sim 180\,K$ with a helical magnetic

structure, and then transforms from the helical antiferromagnetic state to a ferromagnet at ~90 K. When the magnetic field is between 0 and ~2 T, a few additional magnetic structures emerge in pure Dy between 90 and 180 K. Thus, when the magnetic field is low, the MCE of Dy shows a sharp step-like increase at ~90 K due to a first-order ferromagnetic–antiferromagnetic transition, and then goes through a minimum immediately followed by a weak maximum at ~180 K for $\Delta B = 1$ T (see **Figure 4**). The minimum exists because the application of a magnetic field to an antiferromagnet increases the magnetic entropy, thus inverting the sign of the MCE (in a ferromagnet the increasing field decreases magnetic entropy). When B_f is increased to 2 T, the magnetic field is strong enough to quench the first-order ferromagnetic–antiferromagnetic phase transition and it induces a noncollinear magnetic structure which yields a broad MCE maximum at ~127 K. Since a 2 T magnetic field is not strong enough to destroy this noncollinear structure, the slightly negative ΔT_{ad} is still observed at ~174 K. The latter is followed by a weak caret-type peak at ~181 K (**Figure 4**). Upon increasing the magnetic field to 5 T, it becomes strong enough to suppress all of the magnetic structures except the ferromagnetic phase and the MCE of Dy for $\Delta B = 5$ T has a single, somewhat skewed, caret-type peak at ~181 K.

Most magnetic materials on cooling undergo a second-order phase transition from a paramagnet to a ferromagnet with a conventional MCE behavior (**Figures 2** and **3a**), or from a paramagnet to an antiferromagnet with a skewed caret-like MCE if the magnetic field is large enough to destroy the antiferromagnetism and cause it to change to a ferromagnetic structure (**Figure 4**). A few materials, however, form a ferromagnetically ordered phase through a first-order magnetic phase transition, which may also be coupled with a change of the crystal lattice. One such example is illustrated in **Figure 3c**. Here, since the phase transition in $Gd_5Si_2Ge_2$ is of the first order, the $|[\partial M(T, B)/\partial T]_B|$ is larger than usual (but not infinite and undefined as could happen if the transition occurs infinitely fast at constant temperature, pressure, and magnetic field), and therefore, the MCE is also larger than usual (see eqns [3] and [4]). Technically, this behavior is known as the "giant MCE." If one compares the behavior of the giant MCE in $Gd_5Si_2Ge_2$ with that of the other two materials shown in **Figures 3a** and **3b** the most obvious difference is that at large magnetic fields the MCE versus temperature is extended considerably toward the high-temperature side of the peak, but the increase in the magnitude of the MCE with increasing field is affected to a much lesser extent than that in second-order phase transition compounds.

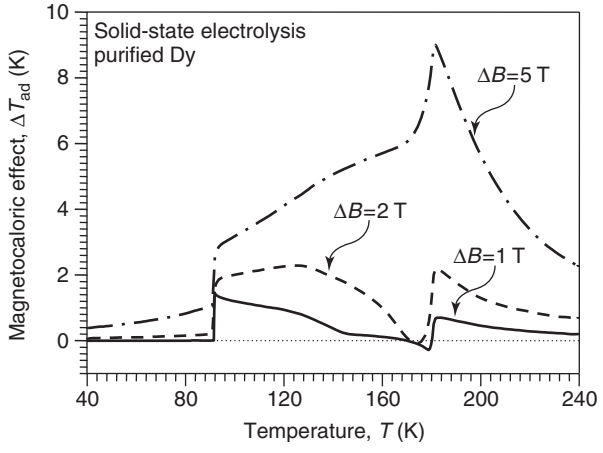


Figure 4 The magnetocaloric effect in ultrapure Dy calculated from the heat capacity data.

The MCE in first-order phase transition materials undergoing coupled magnetostructural transformations may be further enhanced by the added difference of the entropies of the two crystallographic modifications of the material, $\Delta S_{st} = \pm(S_{HF} - S_{LF})$, where HF and LF designate the high field and the low field phases, respectively:

$$\Delta S_M(T, \Delta B)_{\Delta B = B_f - B_i} = \int_{B_i}^{B_f} \left(\frac{\partial M(T, B)}{\partial T}\right)_B dB + \Delta S_{st} \qquad [6]$$

It is worth noting that unlike the first factor of the right-hand side of eqn [6], which represents the conventional contribution to the isothermal magnetic entropy change (see eqn [3]), ΔS_{st} is magnetic field independent provided the change from B_i to B_f is large enough to complete the structural transformation.

Although the last factor in eqn [6] is a hidden parameter in conventional measurements of the MCE, an estimate based on comparing the MCE exhibited by the closely related materials with and without magnetic field-induced structural transformations indicates that ΔS_{st} may account for more than a half of the total magnetic entropy change in magnetic fields below 5 T. Advanced magnetocaloric materials should, therefore, exist in solid systems where extensive structural changes are coupled with ferromagnetic ordering, and therefore, can be triggered by a magnetic field. Considering eqn [6], the strongest MCE should be found in novel compounds engineered in order to maximize the entropy of a structural transformation, ΔS_{st}. An extensive listing of ΔS_M and ΔT_{ad} values can be found in the reviews by Gschneidner and Pecharsky, and Tishin and Spichkin – see the "Further reading" section.

Active Magnetic Regenerator Cycle and Near-Room-Temperature Magnetic Refrigeration

Adiabatic demagnetization refrigeration, described at the beginning of this article, is a discontinuous process for all practical reasons because it may take as much as a few hours to repeat the refrigeration cycle and reach the target temperature. The history of continuous magnetic refrigeration can be traced to the work of S C Collins and F J Zimmerman, and C V Heer, C B Barnes and J C Daunt, who in the early 1950s built and tested two magnetic refrigerators operating between ~1 and 0.2 K by periodically magnetizing and demagnetizing iron ammonium alum. The apparatus built by Heer, Barnes, and Daunt operated at 1/120 Hz frequency (2 min per cycle) and on average extracted 12.3 $\mu J\,s^{-1}$ from the cold reservoir at 0.2 K.

It was not, however, until 1976 when G V Brown reported a near-room-temperature continuously operating magnetic refrigerator, that it became clear that magnetic refrigeration may be successfully utilized at significantly higher temperatures and achieve much larger temperature spans than the maximum observed MCE. Brown was able to attain a 47 K no-load temperature difference between the hot end (319 K) and cold end (272 K) of his unit by regenerating a column of fluid using Gd metal and a magnetic field change from 0 to 7 T and from 7 to 0 T. The achieved temperature difference was more than three times the MCE of Gd for $\Delta B = 7$ T (see **Figure 2**, where for a slightly greater $\Delta B = 7.5$ T, the maximum ΔT_{ad} of Gd is around 16 K at $T_C = 294$ K; the MCE is ~13 K and ~11 K at 319 K and 272 K, respectively).

Following the early work of Brown, the concept of active magnetic regenerator (AMR) refrigeration was introduced by W A Steyert in 1978 and developed by J A Barclay and W A Steyert in the early 1980s. It was subsequently brought to life in the later 1990s, when various magnetic refrigeration units were built in the US, Europe, and Japan. In the AMR cycle, shown schematically in **Figure 5**, a porous bed of a magnetic refrigerant material acts as both the refrigerant that produces refrigeration (i.e., temperature lift) and the regenerator for the heat transfer fluid. Assume that the bed is at a steady state with the hot heat exchanger at 18°C and the cold heat exchanger at 2°C. In **Figure 5a**, the initial temperature profile for the bed is in its demagnetized state in zero magnetic field (dashed line). When a magnetic field is applied to the refrigerant, each particle in the bed warms because of the MCE to form the final magnetized bed temperature profile (solid line). The amount each particle warms is equal to ΔT_{ad} reduced by the effect of the heat capacity of the heat transfer fluid in the pores between the particles. Next, the 2°C fluid flows through the bed from the cold end to the hot end (**Figure 5b**). The bed is cooled by the fluid, lowering the temperature profile across the bed (from the dashed line to the solid line), and the fluid in turn is warmed by the bed, emerging at a temperature close to the temperature of the bed at the warm end.

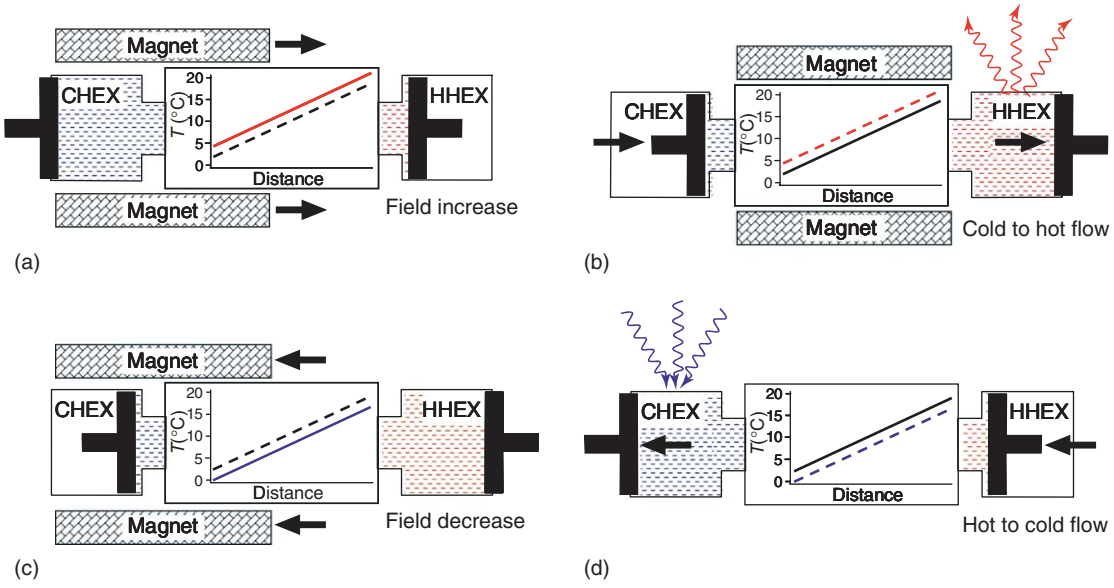


Figure 5 The four steps of the active magnetic regenerator cycle: magnetizing (a), flow from the cold heat exchanger (CHEX) to the hot heat exchanger (HHEX) (b), demagnetizing (c), and flow from HHEX to CHEX (d). This description follows that introduced by C B Zimm and A J DeGregoria in 1993.

This temperature is higher than 18°C, so heat is removed from the fluid at the hot heat sink as the fluid flows through the hot heat exchanger. After the fluid flow is stopped, the magnetic field is removed, cooling the bed by the MCE (from the dashed line to the solid line in **Figure 5c**). The refrigeration cycle is completed by forcing the 18°C fluid to flow from the hot to the cold end of the bed (**Figure 5d**). The fluid is cooled by the bed, emerging at a temperature below 2°C and removes heat from the cold sink as the fluid passes through the cold heat exchanger. After completion of this last step, the cycle is repeated beginning from step (a) (**Figure 5**).

The AMR cycle outlined above has several positive features useful for practical application in magnetic refrigeration devices. First, the temperature span of a single stage can greatly exceed that of the MCE of the magnetic refrigerant because the MCEs of the individual particles collectively change the entire temperature profile across the bed. Second, because the bed acts as its own regenerator, heat need not be transferred between two separate solid assemblies, but rather between the solid particles in a single bed via the action of a fluid. Third, the individual particles in the bed do not encounter the entire temperature span of the stage, and hence the bed may be made into layers, each containing a magnetic material with properties optimized for a particular temperature range.

The AMR refrigeration cycle has been successfully realized in a laboratory prototype magnetic refrigerator, which was recently designed and constructed by the Astronautics Corporation of America. The device is shown together with its schematic diagram in **Figure 6**. The refrigerator operates near room temperature in a magnetic field between 0 and ~1.5 T created by a permanent magnet. As the wheel spins, it brings one of its sections containing the magnetic refrigerant (the regenerator bed) into the high magnetic field volume, where the bed is heated due to the MCE (see **Figure 6a**). As long as this section of the wheel moves between the poles of the permanent magnet, it remains in the magnetized state and the heat exchange fluid (water) flows from the cold heat exchanger through the bed and into the hot heat exchanger. This corresponds to the second step of the AMR cycle shown in **Figure 5b**. While the wheel continues to spin, the same section exits from the high magnetic field volume. As the bed cools due to the MCE, the heat exchange fluid flow is reversed and it now flows from the hot heat exchanger through the bed toward the cold heat exchanger, thus completing steps (c) and (d) of the AMR cycle illustrated in **Figure 5**.

This permanent magnet-based rotating bed magnetic refrigerator was recently awarded US patent no. 6 526 759. The unit can operate at frequencies exceeding 1 Hz and it has achieved a maximum cooling power of 95 W and a maximum temperature span of 20°C. As tested by engineers at the Astronautics Corporation of America, the Carnot efficiency of this laboratory prototype system compares favorably with conventional vapor-cycle refrigeration devices, thus making magnetic refrigeration the first energy-efficient solid-state refrigeration technology.

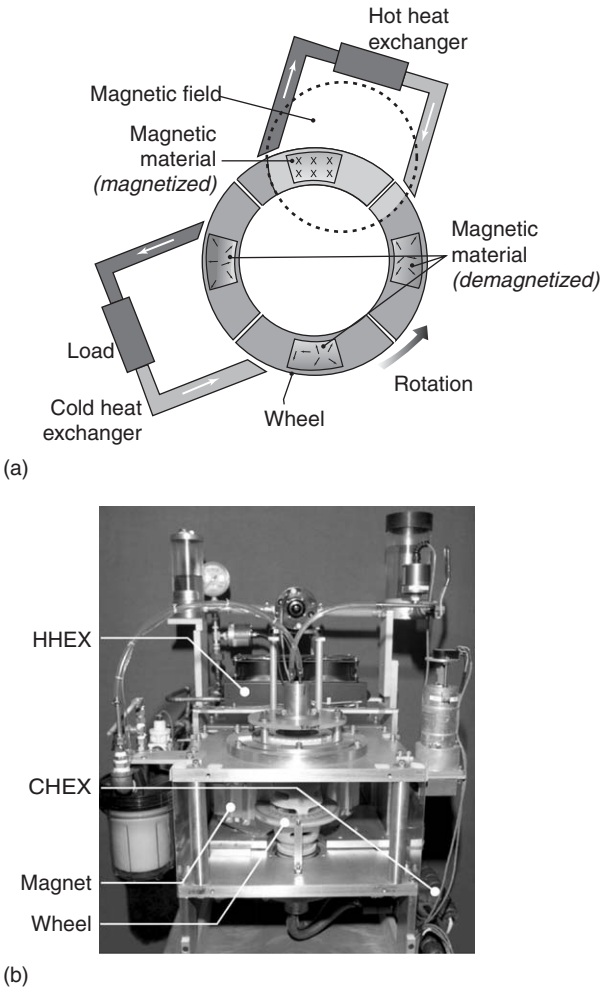


Figure 6 (a) The principal schematic diagram of the device and (b) laboratory prototype magnetic refrigerator designed and constructed by the Astronautics Corporation of America in 2001. A ~1.5 T magnetic field around the wheel filled with Gd spheres is produced by a permanent magnet. The refrigerator operates at ambient conditions with a maximum temperature span of ~20°C and maximum cooling power of 95 W. ((b) Courtesy of the Astronautics Corporation of America, 4115 N. Teutonia Avenue, Milwaukee, WI 53209.)

The Future of the Magnetocaloric Effect

Considering practical applications of the MCE, its extent is one of the most critical parameters defining

the performance of a magnetic refrigerator – the stronger the MCE the higher the efficiency of the device, all other things being equal. As is often the case with emerging technologies, numerous applications of magnetic refrigeration may stem from a reliable foundation formed by advanced magnetocaloric materials. For example, the availability of low-cost high-performance solids exhibiting enhanced MCE between ~ 250 and $\sim 350\,\mathrm{K}$ is an important requirement in order to facilitate commercialization of magnetic refrigeration for a variety of consumer uses – from home appliances to climate control in motor vehicles. In another example, when suitable magnetocaloric compounds supporting continuous magnetic cooling from ~ 300 to $\sim 20\,\mathrm{K}$ are developed, the energy penalty incurred during hydrogen liquefaction using the conventional gas-compression approach may no longer be a limiting factor preventing the widespread use of liquid hydrogen fuel in transportation. It is conceivable that many other areas where conventional refrigeration is inapplicable because of scaling difficulties and/or where thermoelectric refrigeration demands too much energy, magnetic refrigeration will eventually emerge as the cooling technology of choice.

To amplify the MCE, conventional wisdom calls for increasing the magnetic field change in addition to maximizing both the derivative, $|(\partial M(T, B)/\partial T)_B|$, and the region of magnetic fields and temperatures where the magnetization remains highly sensitive to temperature (see eqns [1]–[5]). Considering the current state-of-the-art of permanent magnet alloys and magnet design, it is quite unlikely that magnetic fields in excess of about 2 T at a reasonable cost will become common in the foreseeable future. That is why maximizing the MCE by manipulating chemical and phase compositions, atomic, microscopic, and magnetic structures of a material appears to be the most realistic option in order to reach the strongest possible MCEs in readily available magnetic fields. Although it remains a formidable challenge for basic science, a better understanding of the MCE, including the ability to control a variety of chemical, structural, and physical degrees of freedom that define the properties of solids will lead to improved existing materials and should eventually result in novel compounds exhibiting large MCE.

An especially promising area of research is the study of the extensive measure of the MCE, that is, the isothermal magnetic-field-induced entropy change that can be strongly enhanced by the added entropy of a structural transition (see eqn [6]). Provided the magnetic field completes the polymorphic transformation, this additional entropy is magnetic field independent, and therefore, may account for more than half of the observed MCE, especially in relatively weak magnetic fields. The design of novel materials exhibiting potent MCEs in weak magnetic fields is, therefore, possible by ensuring that (i) a structural transition can be triggered and completed (or nearly completed) by a weak magnetic field; (ii) the transformation has low thermal and magnetic field hystereses, that is, is reversible, and (iii) the difference between the entropies of different polymorphic modifications is maximized.

Acknowledgments

This work was supported by the Materials Sciences Division of the Office of Basic Energy Sciences of the US Department of Energy under a contract No. W-7405-ENG-82 with Iowa State University. The authors wish to thank Dr. Steven Russek and Dr. Carl Zimm of Astronautics Corporation of America for permission to use the photograph of their rotary magnetic refrigerator.

See also: Ferromagnetism; Magnetic Materials and Applications; Magnetic Order; Paramagnetism; Phase Transformation; Specific Heat; Thermodynamic Properties, General.

PACS: 75.30.Sq; 75.30.–m; 75.50.Cc; 75.50.Ee; 65.40.–b; 65.40.Gr

Further Reading

Gschneidner KA, Jr. and Pecharsky VK (2000) Magnetocaloric materials. *Annual Review of Materials Science* 30: 387–429.

Gschneidner KA, Jr. and Pecharsky VK (2002) Magnetic refrigeration. In: Westbrook JH and Fleischer RL (eds.) *Intermetallic Compounds. Principles and Practice*., vol. 3, pp. 519–539. New York: Wiley.

Morrish AH (1965) *The Physical Principles of Magnetism*. New York: Wiley.

Pecharsky VK and Gschneidner KA Jr. (1999) Magnetocaloric effect and magnetic refrigeration. *Journal of Magnetism and Magnetic Materials* 200: 44–56.

Tegus O, Brück E, Zhang L, Dagula, Buschow KHJ, and de Boer FR (2002) Magnetic phase transitions and magnetocaloric effects. *Physica B* 319: 174–192.

Tishin AM (1999) Magnetocaloric effect in the vicinity of phase transitions. In: Buschow KHJ (ed.) *Handbook of Magnetic Materials*, vol. 12, pp. 395–525. Amsterdam: Elsevier.

Tishin AM and Spichkin YI (2003) *The Magnetocaloric Effect and its Applications*. Bristol: Institute of Physics Publishing.

www.m-w.com

Magnetoelasticity

M R Ibarra García, Universidad de Zaragoza-CSIC, Zaragoza, Spain

Introduction

Joule first observed magnetostriction effect on an iron bar; further experiments performed on Fe–Ni alloys allowed Ch E Guillaume to discover the Invar effect. Since then, a vast discipline in condensed matter studies has emerged. Magnetism and symmetry together lead to a variety of phenomena which, during the twentieth century, have been a key for the understanding of the theories of solids. The earliest magnetostriction theories were based on a phenomenological model proposed by R Becker and W Döring in 1939. The most complete and quantum-mechanical theory of magnetostriction was developed by E R Callen and B Callen. As a general phenomenon in magnetism, in the magnetostriction phenomenology one has two extreme situations in which theories can closely predict and explain the behavior of the solids, that is, local-moment and itinerant approaches. The first considers the magnetic electrons identified with a given atom or ion which makes possible a theoretical treatment within the framework of the mean field crystal electric field (CEF) theory. The itinerant approach should be considered within the framework of a many-body problem and encompasses band-structure calculations. Clear examples of local-moment systems are the rare-earth inter metallics; in these compounds the anisotropic and magnetostriction properties are driven by the 4*f* electron of the rare-earth (RE) ion. The strong spin–orbital coupling gives rise in TERFENOL to the largest magnetostriction effects ever observed at room temperature. On the other hand, the 3*d* alloys represent archetypical systems in which the itinerancy and the strong CEF interaction form the basis of the magnetoelastic effects. In these systems the possible unquenched orbital moment can produce relevant anisotropic and magnetostriction effects.

Current theories rely mainly on these two approaches; nevertheless, there exists a wide phenomenology (not very well understood but fascinating) that will be briefly described in this section. These are related to the magnetic fluctuations of the magnetic moment, which in general, produces large magnetovolume effects. Carrier localization is another source of magnetostriction, as discovered recently in magnetic oxides with colossal magnetoresistance (CMR); orbital instabilities also cause large anisotropic magnetostriction in related oxides, such as the cobaltites. Isotopic effect has also been recently reported in transition metal oxides in which electron–phonon interaction plays a key role in the magnetic and magnetotransport properties. There exists another source of large magnetoelastic effect when magnetic and structural transitions occur; thermal effects produce large spontaneous magnetostriction. These transitions can also be induced by applying a magnetic field and causing large magnetostructural effects of magnetoelastic nature. Other related effects are also important, such as the change of the elastic module due to the change of magnetization – the so-called ΔE effect. The inverse magnetostriction effect, that is, a change in the magnetization originated by an induced lattice deformation, is another example.

Shape memory alloys are also very interesting when, associated with the thermoelastic difussionless transformation, a magnetic transition takes place. In this case, twin boundaries can be displaced by a magnetic field producing large magnetostriction.

Spontaneous Magnetostriction

In the absence of applied magnetic field, only certain solids exhibit magnetostriction effect: the ferrimagnetic or ferromagnetic materials. The existence of long-range magnetic order produces a lattice deformation of magnetoelastic type due to the spin–lattice coupling. Nevertheless, as the overall magnetization is zero, due to the existence of magnetic domains, the magnetostriction is also zero and each domain is magnetostrictively deformed. This is because the random distribution of deformations is, on average, canceled out. Spontaneous magnetostriction is observed when the thermal expansion of the lattice does not follow the Gruneisen law:

$$\omega_{\mathrm{ph}} = (\Delta V/V)_{\mathrm{ph}} = \int k\gamma C_v(T)\ \mathrm{d}T \qquad [1]$$

This law accounts for phonons as contributing to the thermal expansion of the lattice. Below the ordering temperature, ferromagnets experience an additional contribution to the thermal expansion of magnetoelastic origin and isotropic ω_{mag} which can be obtained from the experiment as

$$\omega_{\mathrm{mag}} = (\Delta V/V)_{\mathrm{Experimental}} - (\Delta V/V)_{\mathrm{Gruneisen}} \qquad [2]$$

This contribution is associated with the establishment of a long-range magnetic order, and can be strongly enhanced due to the existence of magnetic moment instabilities (longitudinal fluctuation of the

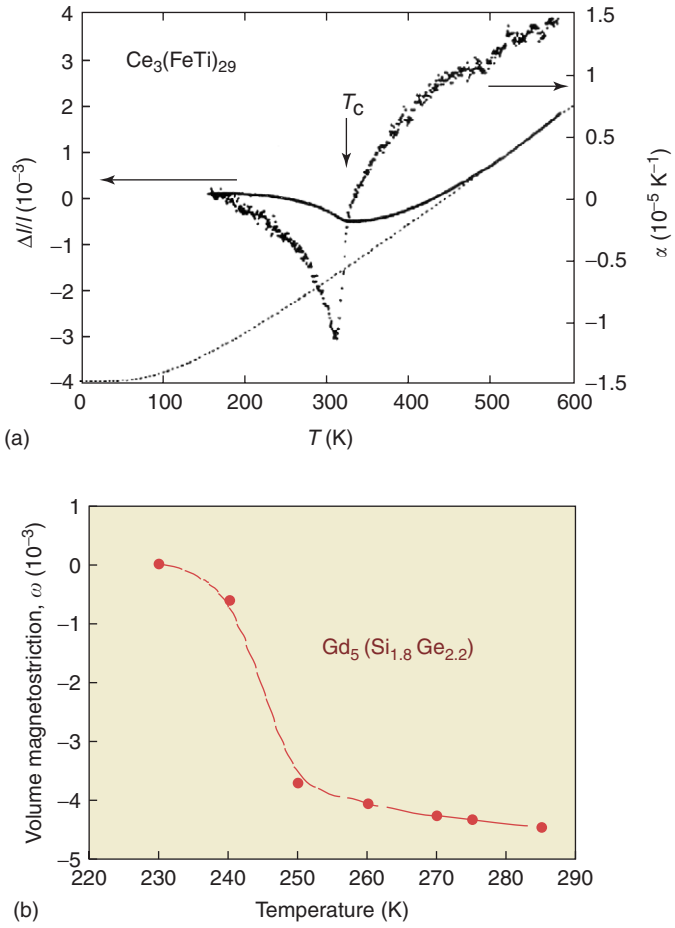


Figure 1 (a) Invar effect in rare-earth intermetallics $Ce_3(FeTi)_{29}$. Dotted line is the phonon contribution calculated using a Debye temperature of 450 K. The arrows show the curie temperature. The extra contribution to the thermal expansion is of magnetic origin. (b) Anomalous thermal expansion due to a magnetostructural transition with a large magnetocaloric effect associated.

magnetization) or structural transitions associated with the magnetic transition.

The first case can give rise to zero or negative thermal expansion coefficient – this is called Invar effect. Alloys such as FeNi, where this effect is relevant, are used in applications such as mechanical precision systems and in large cryogenic containers. Many rich intermetallic alloys present this behavior and an example is given in **Figure 1a**. In many structural transitions the magnetoelastic energy is quite large and a simultaneous structural and magnetic transition may occur; in these cases, a volume contribution also exists, and in general, large entropic content is involved in such transitions (e.g., in compounds with large magnetocaloric effect). **Figure 1b** shows the large volume anomaly observed in the magnetocaloric alloy $Gd_5Ge_2Si_2$.

Anisotropic Magnetostriction

As in magnetism, all materials have magnetostriction under an applied magnetic field. Usually, magnetostriction associated with diamagnetism and paramagnetism is quite small and is not discussed in this article. Ferromagnetic and ferrimagnetic compounds show magnetostrictive deformation when a magnetic field is applied. As already mentioned, the domains are deformed due to the magnetoelastic coupling; under the application of a magnetic field, the wall domain displacement favors an overall deformation of the lattice. To account for the anisotropic nature of the resulting strain, the deformation is measured along and perpendicular to the applied magnetic field. The strains are called parallel ($\lambda_{\parallel}$) and perpendicular ($\lambda_{\perp}$) magnetostriction, respectively. The anisotropic magnetostriction (λ_t) is defined as $\lambda_t = \lambda_{\parallel} - \lambda_{\perp}$. In crystalline materials, the strain depends on the direction of the measurements (β_i) and also the direction of the magnetization (α_i). As an example, for a cubic crystal

$$\begin{aligned}\Delta l/l[\alpha_x,\alpha_y,\alpha_z][\beta_x,\beta_y,\beta_z] &= \tfrac{3}{2}\lambda_{1\,0\,0}(\alpha_x^2\beta_x^2+\alpha_y^2\beta_y^2+\alpha_z^2\beta_z^2-\tfrac{1}{3}) \\ &\quad + 3\lambda_{1\,1\,1}(\alpha_x\alpha_y\beta_x\beta_y+\alpha_x\alpha_z\beta_x\beta_z+\alpha_y\alpha_z\beta_y\beta_z) \qquad [3]\end{aligned}$$

If the deformation is measured along the [1 0 0] axis ($\beta_x = 1, \beta_y = \beta_z = 0$), this expression becomes

$$\Delta l/l = \tfrac{3}{2}\lambda_{1\,0\,0}(\alpha_x^2 - \tfrac{1}{3}) \qquad [4]$$

As a consequence, $\lambda_{1\,0\,0}$ represents the change in length along the [1 0 0] axis when the easy magnetization direction is the [1 0 0] axis ($\alpha_x = 1$), that is, $\Delta l/l = \lambda_{1\,0\,0}$. The anisotropic magnetostriction is then

$$\begin{aligned}\lambda_{\mathrm{t}} &= (\Delta l/l)[1\,1\,1][1\,1\,1] - (\Delta l/l)[1\,1\,1][1\,0\,1] \\ &= \tfrac{3}{2}\lambda_{1\,1\,1}\end{aligned} \qquad [5]$$

Similarly, if the easy magnetization direction is along the [1 1 1] axis and the deformation is measured along this direction (i.e., $(\alpha_x = \alpha_y = \alpha_z = 1/\sqrt{3},\ \beta_x = -\beta_z = 1/\sqrt{2},\ \beta_y = 0)$,

$$\begin{aligned}\lambda_{\mathrm{t}} &= (\Delta l/l)[1\,1\,1][1\,1\,1] - (\Delta l/l)[1\,1\,1][1\,0\,1] \\ &= \tfrac{3}{2}\lambda_{1\,1\,1}\end{aligned} \qquad [6]$$

where $\lambda_{1\,0\,0}$ and $\lambda_{1\,1\,1}$ are related to the strain along the [1 0 0] and [1 1 1] directions when the magnetization is, respectively, along these directions:

$$\lambda_{\mathrm{t}} = \tfrac{3}{2}\lambda_{1\,0\,0} \quad \text{and} \quad \lambda_{\mathrm{t}} = \tfrac{3}{2}\lambda_{1\,1\,1} \qquad [7]$$

In an isotropic solid, there exists only one magnetostrictive strain parameter λ_{s}, such that the deformation along any direction is given by

$$\Delta l/l = \tfrac{3}{2}\lambda_{\mathrm{s}}(\cos^2\theta - \tfrac{1}{3}) \qquad [8]$$

where θ is the angle between the magnetization and the measurement deformation direction. Expression [8] is obtained for a cubic solid if $\lambda_{1\,0\,0} = \lambda_{1\,1\,1} = \lambda_{\mathrm{s}}$ in [3].

Single-Ion Theory of Magnetostriction

H B Callen and E Callen applied quantum mechanics and symmetry considerations to magnetostriction; with this formalism, it is possible to calculate (using microscopic models) the magnetostrictive deformations in a single crystal and, by appropriated averages, obtain information of the intrinsic magnetoelastic parameter from measurements performed in polycrystalline materials. As an example, the procedure is considered assuming cubic symmetry and the easy magnetization direction along the [1 0 0] axis.

The equation similar to expression [3], but written in terms of irreducible strains ε_i^{Γ}, is written as

$$\begin{aligned}\lambda[\alpha_1\alpha_2\alpha_3][\beta_1\beta_2\beta_3] =& \tfrac{1}{3}\varepsilon^{\alpha} + \tfrac{1}{6}\varepsilon_1^{\gamma}(2\beta_3^2 - \beta_2^2 - \beta_1^2) \\ &+ \tfrac{1}{2}\varepsilon_2^{\gamma}(\beta_1^2 - \beta_2^2) + 2\varepsilon_1^{\varepsilon}\beta_2\beta_3 \\ &+ 2\varepsilon_2^{\varepsilon}\beta_1\beta_3 + 2\varepsilon_3^{\varepsilon}\beta_1\beta_2\end{aligned} \qquad [9]$$

The definition of the ε_i^{Γ} strains in terms of the usual Cartesian ones is given in **Table 1**. From eqns [2] and [6], the tetragonal strain for the [0 0 1] direction of the magnetization is

$$h_1 = \sqrt{\tfrac{3}{2}}\varepsilon_1^{\gamma} + \tfrac{1}{\sqrt{2}}\varepsilon_2^{\gamma} \qquad [10]$$

Now, the CEF Hamiltonian referred to a frame of fourfold axes, and where the spontaneous magnetization is along the z-axis, parallel to [0 0 1], is

$$H_{\mathrm{c}} = \tilde{B}_4^0\left(\tilde{O}_4^0 + \sqrt{\tfrac{10}{7}}\tilde{O}_4^0\right) + \tilde{B}_6^0\left(\tilde{O}_6^0 + \sqrt{14}\tilde{O}_6^0\right) \qquad [11]$$

where the $\tilde{B}_4^0$ are the CEF parameters and O_n^m the Stevens DEF operators written in the Buckmaster notation (tabulated by Hutchings). The Zeeman and exchange interactions can be written as

$$H_{\mathrm{Z}} + H_{\mathrm{ex}} = -g\mu_{\mathrm{B}}(\boldsymbol{H} + \lambda\langle\mu\rangle)\boldsymbol{J} \qquad [12]$$

where $\langle\mu\rangle$ is the thermal statistical average of the magnetic moment, $\boldsymbol{J}$ the RE angular momentum and λ the mean-field exchange constant.

In the presence of magnetoelastic coupling, the corresponding Hamiltonian, up to second order terms, is

$$\begin{aligned}H_{\mathrm{ms}} =& -M_0^{\alpha}\varepsilon^{a} - M_2^{\gamma}\left(\varepsilon_1^{\gamma}\tilde{O}_2^0 + \sqrt{2}\varepsilon_2^{\gamma}\tilde{O}_2^2\right) \\ &- M_2^{\varepsilon}(\mathrm{i}\varepsilon_1^{\varepsilon}\tilde{O}_2^1 - \mathrm{i}\varepsilon_2^{\varepsilon}\hat{O}_2^1 + \varepsilon_3^{\varepsilon}\hat{O}_2^2)\end{aligned} \qquad [13]$$

where M_i^{Γ} are the magnetoelastic coupling parameters. The elastic energy is given by

$$\begin{aligned}E_{\mathrm{el}} =& \tfrac{1}{2}C^{\alpha}(\varepsilon^{\alpha})^2 + \tfrac{1}{2}C^{\gamma}\left[(\varepsilon_1^{\gamma})^2 + (\varepsilon_2^{\gamma})^2\right] \\ &+ \tfrac{1}{2}C^{\varepsilon}\left[(\varepsilon_1^{\varepsilon})^2 + (\varepsilon_2^{\varepsilon})^2 + (\varepsilon_3^{\varepsilon})^2\right]\end{aligned} \qquad [14]$$

where C^{Γ} are irreducible symmetric elastic constants. The equilibrium strains (ε_2^{γ}) are formed by minimizing the average energy $\langle H_{\mathrm{ms}}\rangle + E_{\mathrm{el}}$ with respect to the strains, and are given by

$$\begin{aligned}\bar{\varepsilon}_1^{\gamma} &= \frac{M_2^{\gamma}}{C^{\gamma}}\langle\tilde{O}_2^0\rangle, \qquad \bar{\varepsilon}_2^{\gamma} = \frac{M_2^{\gamma}}{\sqrt{2}C^{\gamma}}\langle\tilde{O}_2^0\rangle \\ \bar{\varepsilon}_2^{\varepsilon} &= \frac{M_2^{\varepsilon}}{C^{\varepsilon}}(-\mathrm{i}\langle\hat{O}_2^1\rangle), \text{ etc.}\end{aligned} \qquad [15]$$

The statistical average values $\langle\cdots\rangle$ are evaluated using the energy levels and wave functions which

Table 1 Equivalence between the irreducible ε_1^{Γ} and the Cartesian ε_{ij} for a cubic crystal

Irreducibles	*Cartesians*
ε^{α}	$\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$
ε_1^{γ}	$(1/2)(2\varepsilon_{zz} - \varepsilon_{xx} - \varepsilon_{yy})$
ε_2^{γ}	$(\sqrt{3}/2)(\varepsilon_{xx} - \varepsilon_{yy})$
$\varepsilon_1^{\varepsilon}$	ε_{xy}
$\varepsilon_2^{\varepsilon}$	ε_{xz}
$\varepsilon_3^{\varepsilon}$	ε_{yz}

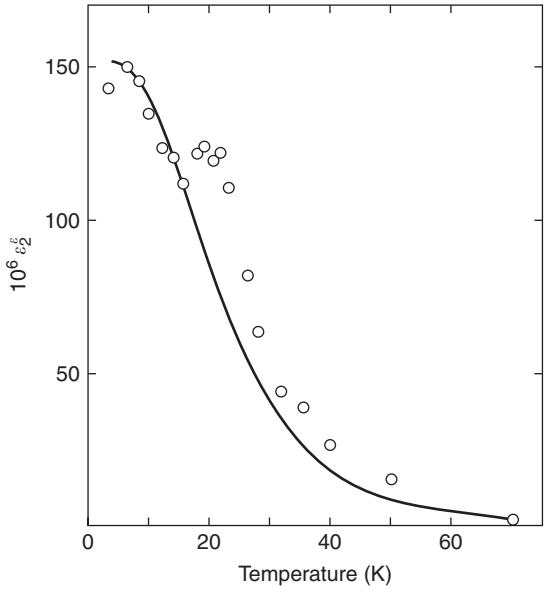


Figure 2 Thermal dependence of the rhombohedral irreducible strain, e.g., in $HoAl_2$ single crystal. The dots are experimental results and the line is the theoretical fitting using the single-ion model described in the text.

diagonalize the unperturbed Hamiltonian $H_c + H_Z + H_{ex}$. This theoretical approach allows calculating the field and thermal dependence of the magnetostriction on single crystals. As an example, **Figure 2** shows the experimental results of the irreducible deformations obtained from magnetostriction measurements in an $HoAl_2$ single crystal. The magnetoelastic coefficient is obtained from the fit of these results.

Magnetostriction in Oxides

Magnetic oxides are characterized by strong localization of the magnetic electrons on the ionic sites, which are separated by oxygen ions. The magnetic interactions in magnetic oxides are of indirect origin, such as the superexchange between ions of the same ionic valence or double exchange, in which the existence of mixed valence is necessary. Magnetostriction in magnetic oxides has been widely studied. It could be noted that the results obtained in ferrimagnetic oxides (e.g., ferrites and garnets) show only weak anisotropic magnetostriction due to the quenching of the orbital moment by the strong crystal electric field interaction. Exceptions to this low value of the magnetostriction were found in titanium, chromium, and cobalt ferrites. De Lacheisserie has summarized the magnetostriction in these compounds. There exist other sources of large spontaneous magnetovolume effect occurring in antiferromagnetic oxides, such as V_2O_3 and Fe_3O_4. In these compounds, these effects occur either due to the existence of a Mott transition, in the case of the vanadium oxides, or to the existence of the charge ordering (Verwey transition in the case of magnetite). These effects are not related to orbital degrees of freedom. Another source of magnetostriction has been found in compounds with cooperative Jahn–Teller effect transition. In this case, the removal of the orbital degeneracy brings about a lowering of the symmetry, giving rise to structural transitions. This is a thermal effect in which tunneling between energetically equivalent distortions becomes unstable as temperature changes, resulting in a new ground state where the distortions are long-range ordered (cooperative Jahn–Teller effect). Under certain conditions, it is possible to induce the transition by the application of a magnetic field. This source of magnetovolume effect does not have its origin in the spin–orbit coupling but in the electron–phonon interaction. Nevertheless, the bibliography on this topic is scarce. In fact, the magnetostriction in manganites was completely ignored until the recent discovery of the CMR phenomena. The upsurge of interest in mixed-valent manganites $R_{1-x}M_xMnO_3$ (R = rare earth, M = Ca, Sr, Ba) has attracted the interest of the scientific community. Ibarra and co-workers discovered a large and unusual magnetostriction in the paramagnetic phase of the CMR compound $La_{0.6}Y_{0.07}Ca_{0.33}MnO_3$. In this compound, the magnetic ions are Mn^{3+} and Mn^{4+}. In both ions, the relativistic spin–orbit coupling is irrelevant, given the lack of orbital moment in Mn^{4+} (t_{2g}^3) and in Mn^{3+} (t_{2g}^4) due to the Jahn–Teller distortion. This compound has a paraferromagnetic transition at $T_c = 260\,K$. Associated with this transition, there exists a metal–insulator transition with an abrupt change of resistance. The metallic phase has a low resistance, which is intrinsically related to the high mobility of the carriers in the ferromagnetic phase due to the nature of the double-exchange interaction. In the paramagnetic phase, the resistance increases as the temperature decreases, showing a semiconductor or polaronic behavior (see **Figure 3a**). Under the application of a magnetic field, the transition can be induced. This is the origin of the intrinsic CMR in this compound. As in the Invar compounds, there exists a volume anomaly at this transition, with an extra contribution over the phonon contribution to the thermal expansion as shown clearly in **Figure 3b**. In this case, such an effect is due to the carrier localization. As the Curie temperature is approached from the paramagnetic phase, the carriers are magnetic polarons (namely, the charge and a magnetic cloud). The thermal expansion effect is related to the formation of these magnetic clusters called "magnetic polarons."

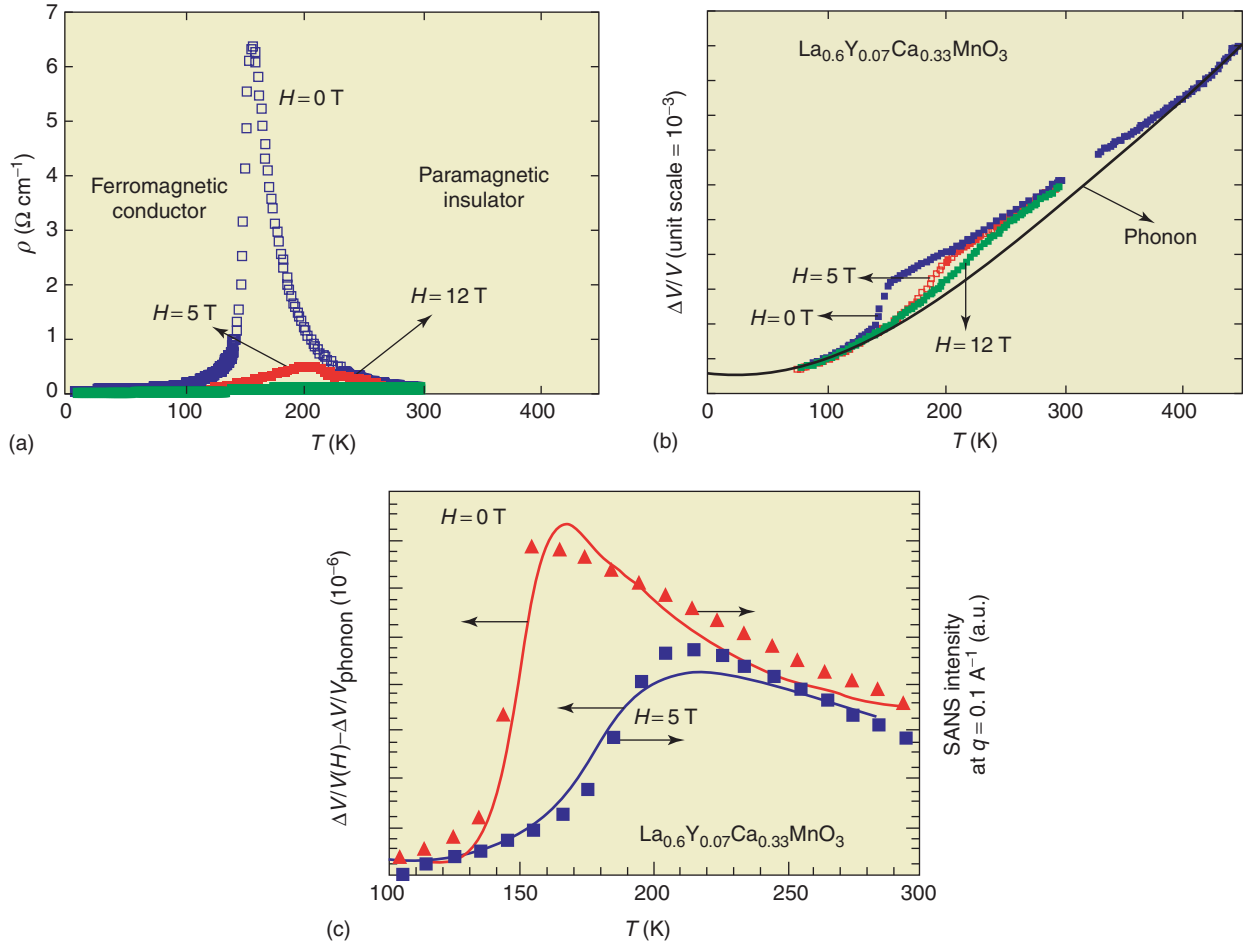


Figure 3 (a) Thermal dependence of the resistivity in the $(LaY)_{2/3}Ca_{1/3}MnO_3$ magnetic oxide. Under an applied magnetic field the resistivity drastically decreases giving rise to the colossal magnetoresistance effect. (b) Thermal expansion of $(LaY)_{2/3}Ca_{1/3}MnO_3$. The line is the phonon contribution calculated by using the Gruneisen law. (c) Anomalous thermal expansion compared with the SANS contribution in $(LaY)_{2/3}Ca_{1/3}MnO_3$.

The thermal dependence of the intensity of the small-angle neutron scattering (SANS) also showed a similar dependence, as shown in **Figure 3c**. Under the application of a magnetic field, the same dependence is kept for both magnitudes with a large change in the volume of the sample. This field dependence at one temperature is shown in **Figure 4**, where a large volume magnetostriction is observed. This large magnetostriction in magnetic oxides is a result of a growing magnetic cluster and percolation raising the ferromagnetic state and, as a consequence, the shrinkage of the lattice.

In the above-mentioned compounds, the anisotropic magnetostriction is negligible because the observed deformation is not related to an orbital effect.

There exists a family of magnetic oxides as the mixed valence cobaltites $La_{1-x}Sr_xCoO_3$ in which an orbital instability gives rise to a large magnetostriction. Unlike the manganites, the cobaltites display a large anisotropic magnetostriction while the volume

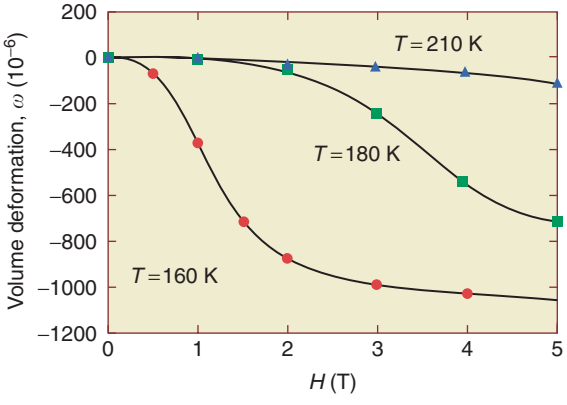


Figure 4 Volume magnetostriction (isotropic) in the paramagnetic phase of the $(LaY)_{2/3}Ca_{1/3}MnO_3$.

(isotropic) magnetostriction is negligible. The Co^{3+} ion in $LaCoO_3$, due to the comparable value of the crystal electric field (Δ_{CEF}) and intra-atomic exchange interaction (J_{ex}), can adopt different

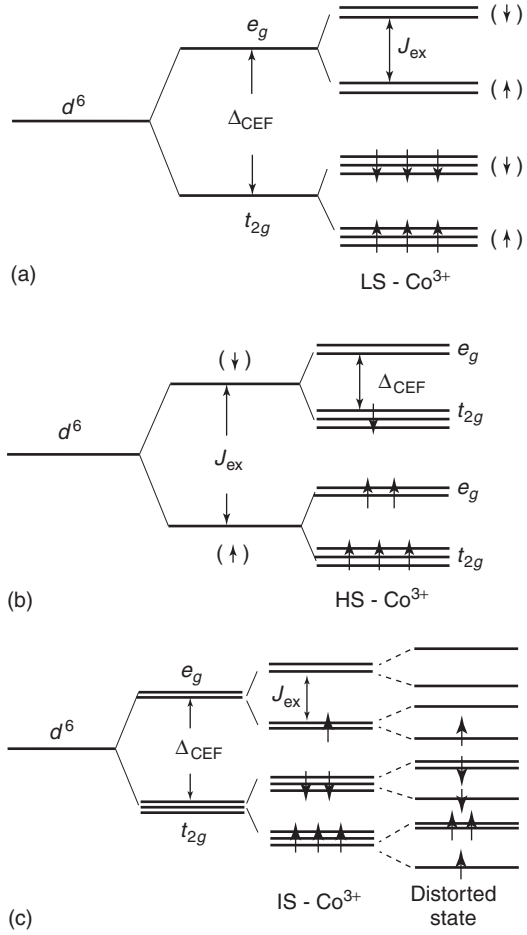


Figure 5 Energy levels splitting of the different Co^{3+} spin states in: (a) the low-spin (LS) state ($\Delta_{CEF} > J_{ex}$, t_{2g}^6, $S = 0$) is diamagnetic. (b) high-spin (HS) state ($\Delta_{CEF} > J_{ex}$, $t_{2g}^4\ e_g^2$, $S = 2$), and (c) intermediate spin (IS) state ($t_{2g}^5\ e_g^1$, $S = 1$).

spin/electronic configurations with difference in energies that can be thermally excited. This situation can be schematized as in **Figure 5**. At low temperatures, this compound is diamagnetic and it is paramagnetic above 90 K. Within the paramagnetic phase, a transition from a low-spin (LS) to an intermediate-spin (IS) state takes place at 90 K. Under hole doping in $La_{1-x}Sr_xCoO_3$, mixed-valent Co^{3+} and Co^{4+} ions appear. The interplay of double-exchange and superexchange interaction in these compounds gives rise to complex magnetic structures. In the composition $0.25 \leqslant x \leqslant 0.5$, the compounds show cluster glass behavior: short-range ferromagnetically correlated clusters embedded in a nonmagnetic matrix; no spontaneous thermally induced transition have been found between different spin states. In these compounds, the volume magnetostriction is negligible. However, the anisotropic magnetostriction is large (see **Figure 6**). This is a unique effect considering that the system is magnetically disordered and it is explained within the framework of a field-induced

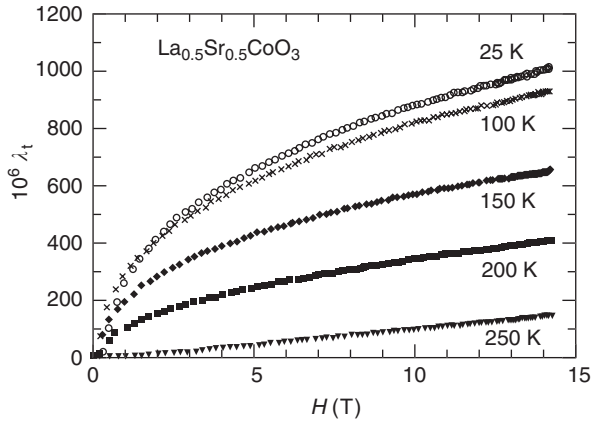


Figure 6 Huge anisotropic magnetostriction in $La_{1-x}Sr_xCoO_3$.

spin-state transition. Under the application of a magnetic field, there exists a change from the LS state to an IS state that is closer in energy. As shown in **Figure 5** the IS Co^{3+} is a Jahn–Teller ion. Under Jahn–Teller distortion, the doubly degenerated e_g splits into two singlets ($L = 0$), and the triplet t_{2g} into a singlet and a doublet ($L = 1$). The singlet t_{2g} states are occupied by two electrons with opposite spins and the doublet levels by three electrons. Due to the degeneracy of this level with nonzero angular moment, a strong intra-atomic spin–orbit coupling is created, which couples to the lattice strain giving rise to a large anisotropic magnetostriction. These results constitute a nice example of CMR arising from an orbital instability as a consequence of a transition from the LS to the IS state.

Isotopic Effect on Magnetostriction

In the polaronic regime of the manganites, the carriers have an effective bandwidth W_{eff} given by

$$W_{eff} \propto W e^{-\gamma E_b/\hbar\omega} \qquad [16]$$

Here, W is the bare conduction bandwidth, E_b the binding energy of the polaron, ω the characteristic frequency of the optical phonon, and γ a dimensionless parameter ($0 < \gamma \leqslant 1$) which depends on the ratio E_b/W in such a way that as E_b/W decreases, W_{eff} increases. In the strong Hund coupling limit (high-spin state, $J_H \gg W_{eff}$), $T_c \propto W_{eff}$. In this way the stability of the long-range ferromagnetic metallic state can be related to the phonon vibrational modes, that is, the lattice. Zhao and co-workers found an extraordinary effect in mixed-valent $La_{1-x}Ca_xMnO_3$. In $La_{0.8}Ca_{0.2}MnO_3$, the isotopic exchange of the natural ^{16}O isotope by ^{18}O gives rise to a decrease of 10 K in T_c. These results can be explained by considering the isotopic mass effect on the phonon frequency ($\omega \propto M^{-1/2}$). Therefore, an increase in the oxygen

mass will decrease ω and consequently will decrease W_{eff}. The result is, as expected, a less stable ferromagnetic metallic state, that is, one with a lower T_c. **Figure 7** shows the results of thermal expansion in isotope exchange manganite $La_{2/3}Ca_{1/2}MnO_3$. At a certain temperature (see arrows in **Figure 7a**) the same compound can be either metallic ferromagnetic (^{16}O) or insulator paramagnetic (^{18}O). This gives rise to a large spontaneous isotopic magnetovolume effect within the temperature range in which the oxygen isotopic exchange is able to destabilize the ferromagnetic metallic phase. **Figure 7b** displays the isotopic contribution to the thermal expansion. The isotopic effect can be more drastic if manganites in the composition range close to a border line separating ferromagnetic metallic and antiferromagnetic charge/orbital states are considered. In this case, the subtle balance between the electronic correlations can be strongly affected by a variation of the intensity of the electron–phonon interaction which can be driven by oxygen isotope exchange.

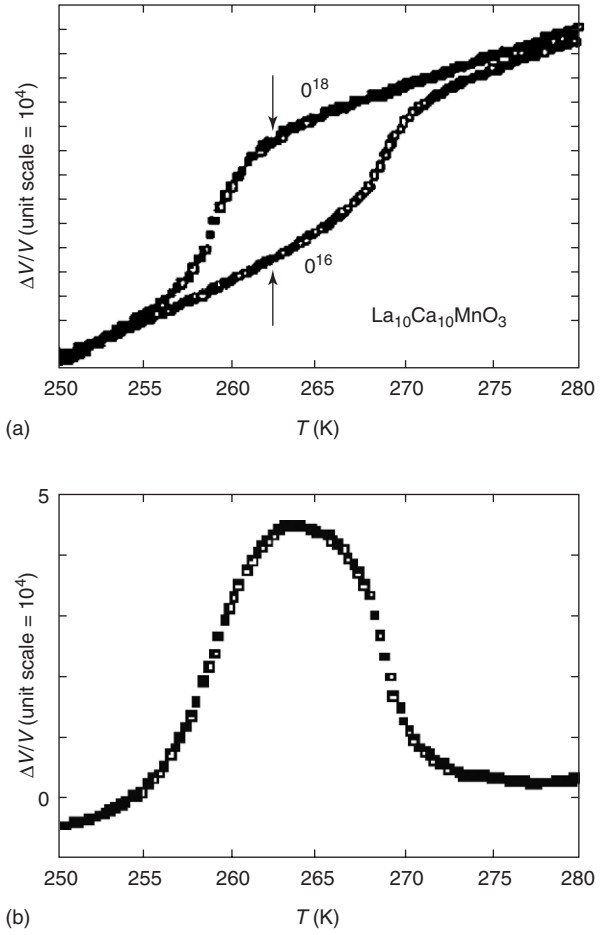


Figure 7 Volume thermal expansion of the $La_{2/3}Ca_{1/3}MnO_3$ compound: (a) comparison of the results for the ^{18}O and ^{16}O isotope compounds. Arrows show the curie temperature. (b) Differential isotope volume thermal expansion.

This is the case in $(La_{0.5}Nd_{0.5})_{2/3}Ca_{1/3}MnO_3$. The compound with ^{18}O isotope is an insulator over the whole temperature range, whereas the ^{16}O sample exhibits a metal–insulator transition at 150 K. This transition is strongly affected by external pressure or an applied magnetic field which can induce the insulator–metal transition. The volume thermal expansion results are shown in **Figure 8**. The solid line is the calculated phonon contribution using the Gruneisen law ($\theta_D = 500$ K). The thermal expansion of the ^{18}O sample follows the insulator behavior down to the lowest temperatures, in good agreement with electrical resistivity measurements. In contrast, the ^{16}O sample shows a different behavior. Above 150 K, the thermal expansion follows the behavior of an insulator, but below this temperature the behavior obtained partly corresponds to an insulator or a metal state. This suggests the existence of static phase segregation in which insulating antiferromagnetic charge/orbital region co-exists with ferromagnetic metallic regions. As is shown in **Figure 9**, large magnetostriction is present at values of a critical field at which a transition from the insulator into the metallic state takes place. For the ^{18}O sample at 5 K

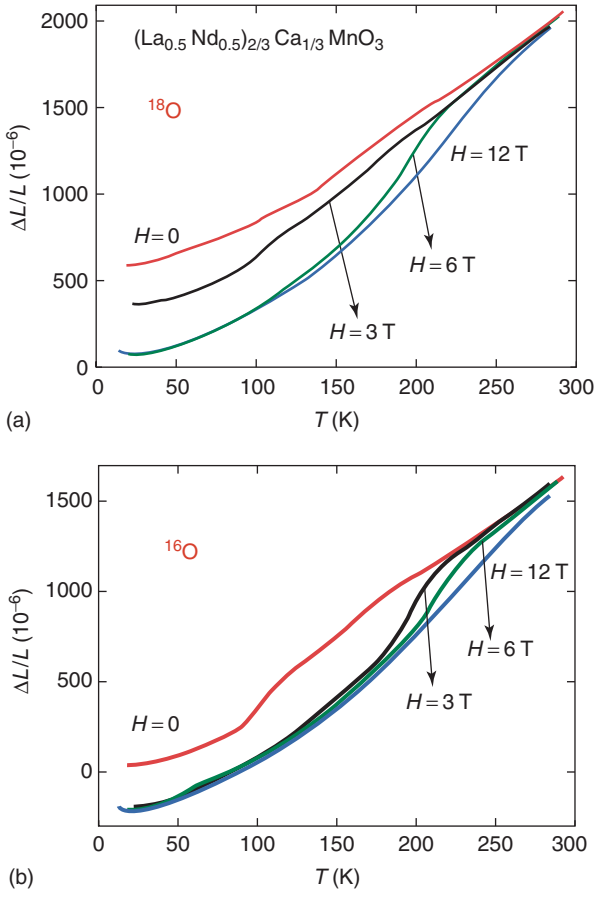


Figure 8 Volume thermal expansion of the $(La_{0.5}Nd_{0.5})_{2/3}Ca_{1/3}MnO_3$. (a) ^{18}O isotope and (b) ^{16}O isotope.

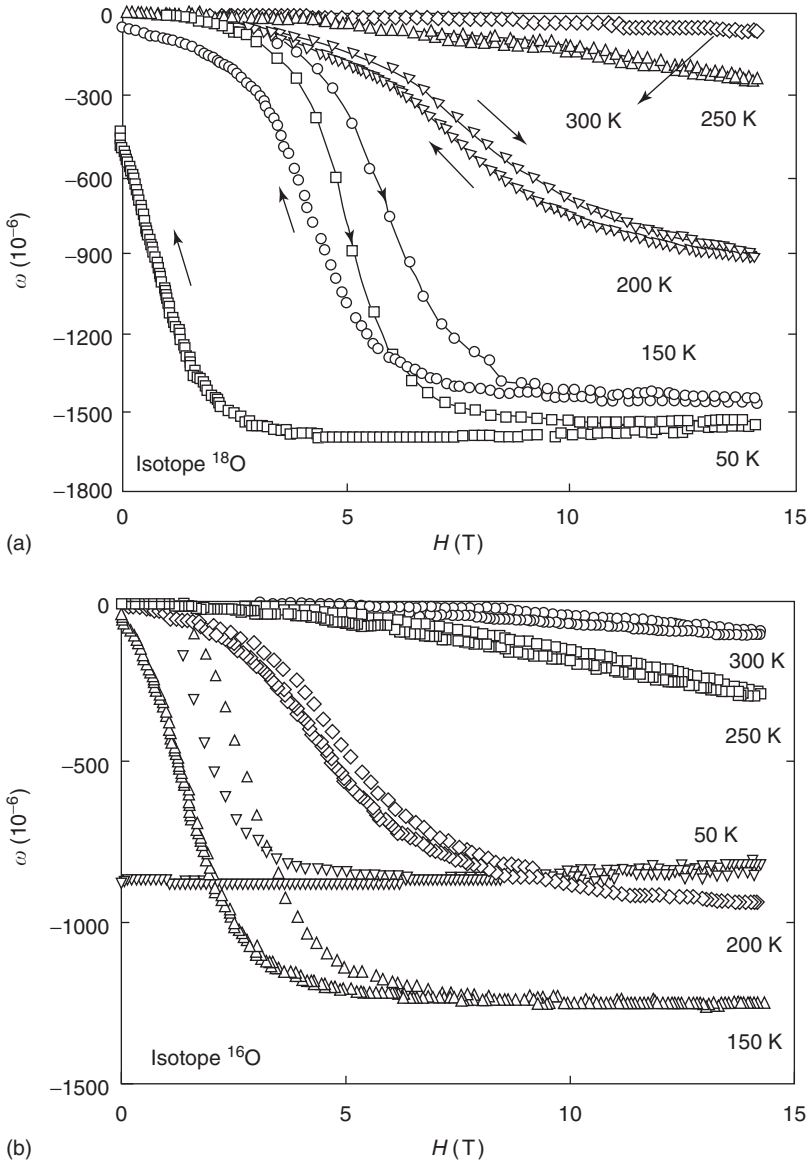


Figure 9 Volume magnetostriction of $(La_{0.5}Nd_{0.5})_{2/3}Ca_{1/3}MnO_3$: (a) ^{18}O isotope and (b) ^{16}O isotope.

the saturation magnetostriction is $\omega = 1.6 \times 10^{-3}$. However, in the ^{16}O the value obtained at the same temperature is $\omega = 0.8 \times 10^{-3}$. This constitutes a relevant observation in the field of magnetostriction in which the oxygen isotopic exchange gives rise to a magnetovolume effect of 0.8×10^{-3}.

See also: Magnetic Interactions; Magnetic Materials and Applications; Magnetic Order; Magnetism, History of.

PACS: 75.80.+q

Further Reading

Callen ER and Callen B (1965) Magnetostriction, forced magnetostriction, and anomalous thermal expansion in ferromagnets. *Physical Review* 139A: 455.

Carr WJ Jr. (1959) Magnetostriction. In: *Magnetic Properties of Metals and Alloys*, ch. 10. Cleveland, OH: American Society for Metals.

Clark AE (1980) Magnetostrictive rare earth-Fe_2 compounds. In: Wohlfarth EP (ed.) *Ferromagnetic Materials: A Handbook on the Properties of Magnetically Ordered Substances*, vol. 1, pp. 531–589. Amsterdam: Nord-Holland.

Cullen JR, *et al.* (1994) Magnetostrictive materials. In: *Materials Science and Technology*, vol. 3B, ch. 16. Weinheim: VCH.

Del Moral A *Magnetostriction: Principles and Applications* (to appear).

De Teresa JM, Ibarra MR, Algarabel PA, Ritter C, Marquina C, *et al.* (1997) Evidence for magnetic polarons in the magnetoresistive perovskites. *Nature* 386: 256.

Du Tremolet de Lacheisserie E (1993) *Magnetostriction*. Boca Raton, FL: CRC Press.

Ibarra MR, Algarabel PA, Marquina C, Blasco J, and García J (1995) *Physical Review Letters* 75: 3541.

Ibarra MR, *et al.* (2000) Magnetostriction in mixed valence magnetic oxides. In: Gibbs MRJ (ed.) *Modern Trends in Magnetostriction Study and Applications*, pp. 171–204. Netherlands: Kluwer Academic Publishers.

Morin P and Schmitt D (1990) Quadrupolar interactions and magnetoelastic effects in rare earth intermetallic compounds. In: Wohlfarth EP (ed.) *Ferromagnetic Meterials*, vol. 5, ch. 1. pp. 2–132. Amsterdam: Nord-Holland.

O'Handley RC (1999) Magnetoelastic effects. In: *Modern Magnetic Materials: Principles and Applications*. Ch. 7, pp. 218–274. Wiley–Interscience.

Szymczak H (1999) From almost zero magnetostriction to giant magnetostrictive effects: recent results. *Journal of Magnetism and Magnetic Materials* 200: 425–438.

Wassermann EF (1990) In: Buschow K and Wohlfarth EP (eds.) *Ferromagnetic Materials*, vol. 1, p. 240. Amsterdam: Nord-Holland.

Nomenclature

c^{Γ}	irreducible elastic constant (J)
g	Landi factor
H_c	CEF Hamiltonian (T)
H_{ex}	exchange Hamiltonian (T)
H_Z	Zeeman Hamiltonian (T)
O_n	Stevens operators
W_{eff}	effective polaronic bandwidth (J)
α_i	magnetization direction
β_i	measurement direction of the strain
Δ_{CEF}	crystal electric field splitting (J)
ε_i^{γ}	irreducible strains
$\bar{\bar{\varepsilon}}_i^{\gamma}$	equilibrium strains
ε_{ij_m}	Cartesian strains
λ	mean field exchange constant (T)
λ_{ijk}	strain along the (i, j, k)
λ_t	anisotropic magnetosriction
$\lambda_{\parallel}$	parallel magnetostriction
$\lambda_{\perp}$	perpendicular magnetostriction
$\langle \mu \rangle$	statistical average of magnetic moment ($J T^{-1}$)
μ_B	Bohr magneton ($J T^{-1}$)
ω	volume magnetostriction

Magnetoresistance Techniques Applied to Fermi Surfaces

S J Blundell, University of Oxford, Oxford, UK

Introduction

Magnetoresistance is the change of resistance with a magnetic field. The experimental use of magnetoresistance techniques to determine the Fermi surfaces of various metals has a long and successful history. This is partly due to the relative ease with which resistance measurements can be made, and also because there are many variables which can be adjusted. These include temperature, magnetic field strength, and the orientation with respect to the crystal axes of both the magnetic field and the current density. As described below, the presence of open and closed orbits can be easily distinguished by the field dependence of the magnetoresistance while the area of the Fermi surface pockets can be measured by the frequency of field-dependent magnetoresistance oscillations. The angle dependence of the magnetoresistance also reveals intricate features of the electron band structure.

Free Electron Gas Model – Single Carrier

In order to gain insight into the phenomenon of magnetoresistance, it is valuable to start with a very simple model. Consider a gas of charge carriers in a metal (n per unit volume, each with charge q and effective mass m^*). The forces on the charge carriers due to the applied electric and magnetic fields, $\boldsymbol{E}$ and $\boldsymbol{B}$, imply that the equation of motion for the ith carrier can be written as $m^*\dot{\boldsymbol{v}}_i = q(\boldsymbol{E} + v_i \times \boldsymbol{B})$, where $\boldsymbol{v}_i$ is the carrier's velocity. In the absence of a driving field, the current density $\boldsymbol{J} = q\sum_i \boldsymbol{v}_i$ would relax according to

$$\frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} = -\frac{\boldsymbol{J}}{\tau} \qquad [1]$$

where τ is the relaxation time. The equation of motion of the current density in the presence of electric and magnetic fields can hence be written as

$$\frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} = -\frac{\boldsymbol{J}}{\tau} + \frac{nq^2\boldsymbol{E}}{m^*} + \frac{q}{m^*}\boldsymbol{J} \times \boldsymbol{B} \qquad [2]$$

In the steady state, and in the absence of a magnetic field, this simplifies to $\boldsymbol{J} = \sigma_0 \boldsymbol{E}$, where $\sigma_0 = nq\mu = nq^2\tau/m^*$ is the conductivity and μ is the mobility. In the presence of a magnetic field, as well as in the steady state, eqn [2] can be rearranged to give

$$\boldsymbol{E} = \frac{\boldsymbol{J}}{\sigma_0} + \frac{\boldsymbol{B} \times \boldsymbol{J}}{nq} = \underline{\underline{\rho}}\boldsymbol{J} \qquad [3]$$

where $\underline{\underline{\rho}}$ is the resistivity tensor. The magnetic field, therefore, induces a transverse electric field $\boldsymbol{B} \times \boldsymbol{J}/nq$ which is perpendicular to both $\boldsymbol{B}$ and $\boldsymbol{J}$ (this is the Hall effect). If the magnetic field is along the z-axis, the resistivity tensor is given by

$$\underline{\underline{\rho}} = \frac{1}{\sigma_0}\begin{pmatrix} 1 & -\gamma & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [4]$$

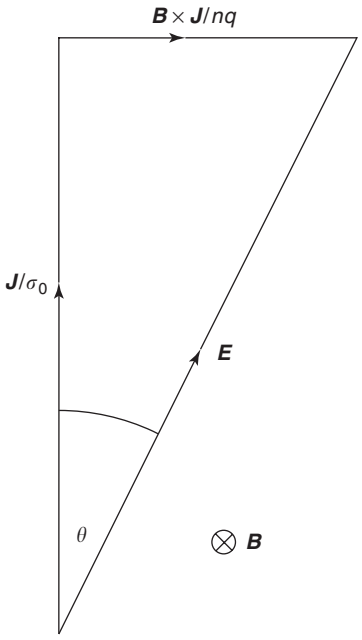


Figure 1 The current density $\boldsymbol{J}$ and electric field $\boldsymbol{E}$. The diagram also shows the Hall angle θ.

where $\gamma = \omega_c \tau = \mu B$ is the mean angle turned between collisions, and $\omega_c = qB/m^*$ is the cyclotron frequency. This predicts a Hall resistivity $|\rho_{xy}| = \gamma/\sigma_0 = 1/nq$, but the diagonal components of the resistivity tensor are all equal to $1/\sigma_0$ and do not depend on the magnetic field. This single-carrier free electron model has, therefore, failed to predict any nonzero magnetoresistance.

This can be understood by reference to **Figure 1** which shows the steady-state balance between the current density and the electric field. In an experiment, the direction of $\boldsymbol{J}$ is normally determined by the position of the electrical contacts. The magnetic field only introduces an electric field perpendicular to the current density (which is known as the Hall field) but leaves the component of the electric field parallel to $\boldsymbol{J}$ unchanged. It is this latter component which determines the resistance, and hence there is no magnetoresistance. The angle between $\boldsymbol{E}$ and $\boldsymbol{J}$ is known as the Hall angle, θ, and is related to γ by $\tan\theta = \gamma$. It turns out also that even the inclusion of some anisotropy in the effective mass does not help in predicting a nonzero magnetoresistance.

Free Electron Gas Model – Two Carriers

Nonzero magnetoresistance can be obtained by introducing more than one type of carrier into the model. In this case, an equation of motion can be written for the current density due to each type of carrier, so that

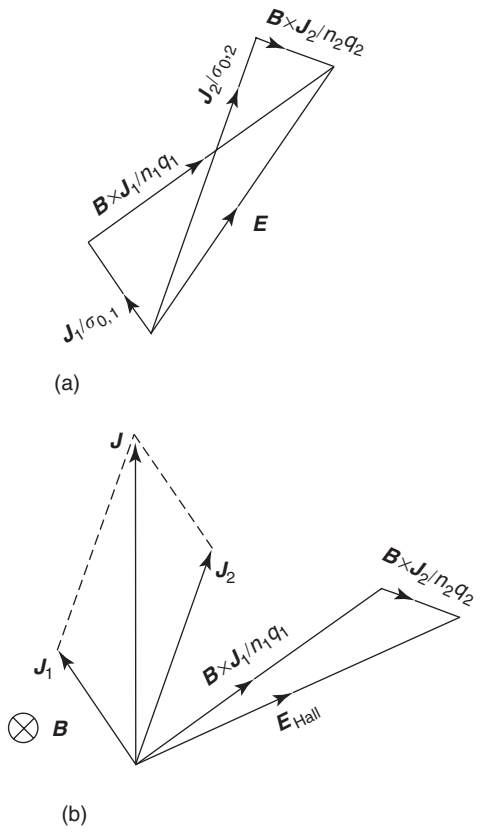


Figure 2 (a) The relationship between the current density $\boldsymbol{J}_1$, $\boldsymbol{J}_2$, and electric field $\boldsymbol{E}$ for two carriers (with charges q_1 and q_2 and number densities n_1 and n_2). The symbols $\sigma_{0,1}$ and $\sigma_{0,2}$ are the zero-field conductivities of carriers 1 and 2 respectively. (b) The total current density $\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2$. The Hall field $\boldsymbol{E}_{\text{Hall}} = \boldsymbol{B} \times (\boldsymbol{J}_1/n_1 q_1 + \boldsymbol{J}_2/n_2 q_2)$ has a component which is parallel to $\boldsymbol{J}$.

$$\frac{\mathrm{d}\boldsymbol{J}_\alpha}{\mathrm{d}t} = -\frac{\boldsymbol{J}_\alpha}{\tau_\alpha} + \frac{n_\alpha q_\alpha^2 \boldsymbol{E}}{m_\alpha^*} + \frac{q_\alpha}{m_\alpha^*}\boldsymbol{J}_\alpha \times \boldsymbol{B} \qquad [5]$$

where the subscript α labels the carrier. The total current density $\boldsymbol{J}$ is given by the sum of the contributions from each carrier, so that

$$\boldsymbol{J} = \sum_\alpha \boldsymbol{J}_\alpha = \left(\sum_\alpha \underline{\underline{\sigma_\alpha}}\right)\boldsymbol{E} \qquad [6]$$

In particular, the resistivity tensor is given by $\underline{\underline{\rho}} = (\sum_\alpha \underline{\underline{\sigma_\alpha}})^{-1}$. The effect of this can be demonstrated easily using the case of two carriers which is shown in **Figure 2**. Nonzero magnetoresistance results, since the application of a magnetic field ensures that the two current densities $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ are no longer parallel. The magnetic field can now induce a Hall electric field which has a component parallel to the total current density. Hence, the component of $\boldsymbol{J}$ which is parallel to $\boldsymbol{E}$ can be dependent upon the magnetic field.

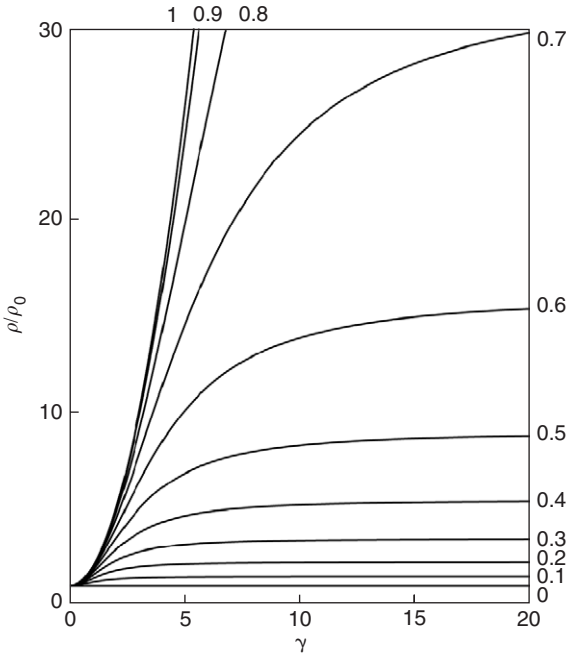


Figure 3 Calculated magnetoresistance (shown as resistivity ρ normalized to the zero-field resistivity ρ_0) for a simple two-carrier model of electrons and holes for different values of the parameter $c = n_+/n_-$.

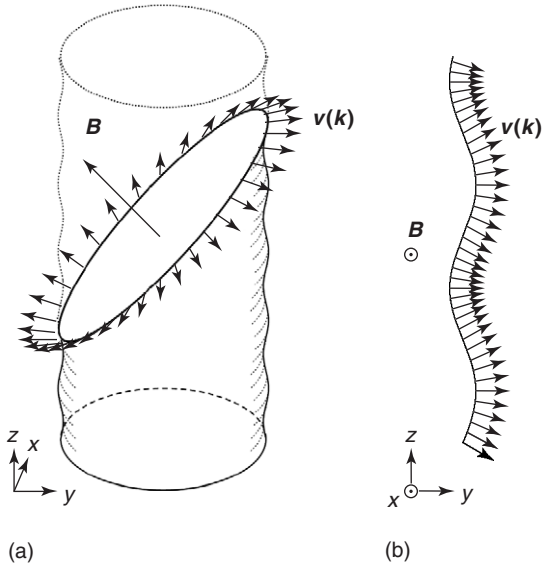


Figure 4 Example of (a) a closed orbit and (b) an open orbit. The vectors on the orbits show the real-space velocity of the electron, given by $\boldsymbol{v} = \hbar^{-1}\partial E(\boldsymbol{k})/\partial \boldsymbol{k}$, which is a vector normal to the Fermi surface. With the magnetic field $\boldsymbol{B}$ oriented as in (a), a closed orbit results. With $\boldsymbol{B}$ along the *x*-direction, an open orbit on the same Fermi surface can result, as shown in (b). (Adapted from Blundell S (2001) *Magnetism in Condensed Matter*. Oxford: Oxford University Press.)

This effect can be demonstrated in a very simple model consisting of two types of carriers, electrons and holes, with equal mobilities but opposite charges. Suppose the number density of electrons is n_- and of holes is n_+. The magnetoresistance can be easily calculated (see **Figure 3**) for different values of the parameter c given by $c = n_+/n_-$. For $c = 0$, there is no magnetoresistance because there is only one type of carrier. As c increases, there is a magnetoresistance which is initially quadratic in B (and so quadratic in γ because $\gamma \propto B$) but it saturates at high field. The magnetic field at which ρ saturates increases as c increases until c reaches unity (known as "perfect compensation" between the electrons and holes), at which point this field becomes infinite and the magnetoresistance never saturates.

Electron Orbits

The free electron model with two carriers successfully predicts nonzero magnetoresistance, and such a model can be very useful in understanding real data in simple metals and semiconductors. However, a free electron model is very often too simplistic to describe the band structures of many other real metals accurately. The electronic band structure is quantified in terms of the energy $E(\boldsymbol{k})$ as a function of the wave vector $\boldsymbol{k}$ of electrons, and this function determines the shape and size of the Fermi surface. It should be noted that magnetoresistance techniques can be used to measure many of the parameters associated with the Fermi surface. In order to understand how this comes about, it is necessary to first consider the types of orbit which may occur on the Fermi surface due to the application of a magnetic field.

Electron motion on the Fermi surface can be described semiclassically by the equation

$$\hbar\dot{\boldsymbol{k}} = -e\boldsymbol{v} \times \boldsymbol{B} \qquad [7]$$

where $\boldsymbol{k}$ is the electron wave vector on the Fermi surface, and

$$\boldsymbol{v} = \frac{1}{\hbar}\frac{\partial E(\boldsymbol{k})}{\partial \boldsymbol{k}} \qquad [8]$$

is the real-space electron velocity (which is always perpendicular to the Fermi surface, using this definition). Hence the electron motion is perpendicular to $\boldsymbol{B}$ but remains on the Fermi surface. This is because in a time dt, the change in $\boldsymbol{k}$ is $\mathrm{d}\boldsymbol{k} \propto (\partial E(\boldsymbol{k})/\partial \boldsymbol{k}) \times \boldsymbol{B}$ and hence is both normal to $\boldsymbol{B}$ and in the plane of the Fermi surface. The orbit of an electron in $\boldsymbol{k}$-space is, therefore, defined by the intersection between the Fermi surface and a plane normal to $\boldsymbol{B}$. The orbits that result can, therefore, either be "closed orbits" or "open orbits" and examples of both are shown in **Figure 4**.

Integration of the equation of motion yields

$$\hbar[\boldsymbol{k}(t) - \boldsymbol{k}(0)] = -e[\boldsymbol{R}(t) - \boldsymbol{R}(0)] \times \boldsymbol{B} \qquad [9]$$

where $\boldsymbol{R}(t)$ is the real-space position of the electron wave packet. Thus, the projection of the electron motion in real space on a plane normal to $\boldsymbol{B}$ is rotated by $\pi/2$ with respect to the $\boldsymbol{k}$-space trajectory and scaled by a factor η given by

$$\eta = \hbar/eB \qquad [10]$$

Thus as $|\boldsymbol{B}|$ increases, the real space orbits wind up more and more tightly.

Boltzmann Transport Equation

For a given orientation of the magnetic field, the conductivity σ_{ij} can be calculated using the Boltzmann transport equation

$$\sigma_{ij} = \frac{e^2}{4\pi^3}\int \mathrm{d}^3k\left[-\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})}\right]v_i(\boldsymbol{k},0) \times \int_{-\infty}^{0} v_j(\boldsymbol{k},t)e^{t/\tau}\,\mathrm{d}t \qquad [11]$$

where $f_0(\boldsymbol{k}) = (\exp([E(\boldsymbol{k})-\mu]/k_\mathrm{B}T)-1)^{-1}$ is the Fermi function and μ is the chemical potential. This expression for the conductivity tensor (sometimes known as the Chambers formula) is an integral (which at absolute zero is taken over all states at the Fermi surface) of the velocity–velocity correlation function for each Fermi surface orbit. This can change dramatically as the direction of the magnetic field is changed, because this alters the paths of all the Fermi surface orbits. The components of the conductivity tensor can be very sensitive to whether the orbits are open or closed, so that these two cases may be distinguished by differing magnetoresistance behavior.

High-Field Behavior

Figure 4a shows a closed orbit around a cylinder of Fermi surface which lies along the z-direction. The velocity $\boldsymbol{v}(\boldsymbol{k})$ is always normal to the Fermi surface so that as the magnetic field B increases, the velocity–velocity correlation function decreases as the components of the velocity are very effectively averaged. Ignoring for the moment the weak corrugation of the Fermi surface along the z-direction, and taking $\boldsymbol{B}\|z$, then in high fields ($\gamma \gg 1$), the conductivity tensor is dominated by the closed orbits and is given by

$$\underline{\underline{\sigma}} \sim \sigma_0 \begin{pmatrix} \gamma^{-2} & \gamma^{-1} & 0 \\ -\gamma^{-1} & \gamma^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [12]$$

using the Boltzmann transport equation. Thus in high fields, $\sigma_{xx} \propto B^{-2}$. The components of the resistivity tensor can be obtained by inverting the conductivity tensor (e.g., in this case $\rho_{xx} = \sigma_{yy}/(\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx})$), so that in a compensated metal (where there are an equal number of electrons and holes, so that $\sigma_{xy} = 0$) such as bismuth, $\rho_{xx}, \rho_{yy} \propto B^2$ (this occurs in magnesium and zinc also, but at very high fields in these materials, the quadratic increase is interrupted by a magnetic breakdown). In an uncompensated metal, $\sigma_{xy} \propto B^{-1}$ at high fields so that ρ_{xx} and ρ_{yy} saturate at high fields.

For an open orbit along z (see **Figure 4b**), with $\boldsymbol{B}$ along the x-direction, the component of velocity along z is similarly averaged (so that $\rho_{zz} \propto B^2$) but the component of velocity along y quickly reaches a nonzero average value as B increases (so that ρ_{yy} quickly saturates). Suppose one is trying to pass a current in the x-direction, but an open orbit is carrying current in the y-direction, giving a term $\sigma_{yy} = s\sigma_0 \gg \gamma^{-2}\sigma_0$ where s is a constant. Then

$$\underline{\underline{\sigma}} \sim \sigma_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & s & \gamma^{-1} \\ 0 & -\gamma^{-1} & \gamma^{-2} \end{pmatrix} \qquad [13]$$

In this case, $\rho_{xx} \propto \gamma^2 \propto B^2$ and a quadratically increasing field dependence is found. Suppose, however, that an open orbit carries current in the x-direction. Then

$$\underline{\underline{\sigma}} \sim \sigma_0 \begin{pmatrix} s & \gamma^{-1} & 0 \\ -\gamma^{-1} & \gamma^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [14]$$

and ρ_{xx} is found to saturate.

Angle-Dependent Magnetoresistance Oscillations

As described above, the magnetoresistance can depend quite dramatically on the direction of the applied magnetic field, and in some cases very large angle-dependent magnetoresistance oscillations at constant field can be found. Experimentally, angle-dependent magnetoresistance oscillations are measured by rotating a sample in a fixed magnetic field while monitoring the resistivity of the sample. Angle-dependent magnetoresistance oscillations can be observed at much higher temperatures and in much lower applied fields than Shubnikov–de Haas oscillations (described below in the following section). This is because Shubnikov–de Haas oscillations arise from the movement of Landau levels through the Fermi energy and, therefore, require that the temperature is low enough for the Fermi surface to be sharply defined; this restriction does not apply so stringently to angle-dependent magnetoresistance

oscillations since they do not originate from the motion of energy levels through the Fermi surface.

Consider now the angle-dependent magnetoresistance oscillations due to Fermi surface orbits around the warped cylindrical Fermi surface pocket shown in **Figure 4a**, this time including the corrugations along the z-axis. The electron dispersion can, in this case, be written as

$$E(\boldsymbol{k}) = \mathscr{E}(\boldsymbol{k}_{\parallel}) - 2t_{\perp}\cos(k_z d) \qquad [15]$$

where $t_{\perp}$ is the interlayer transfer integral, and $\boldsymbol{k}_{\parallel} = (k_x, k_y)$ and k_z are, respectively, the components of the wave vector, parallel and perpendicular to the conducting planes. This dispersion is appropriate for a highly anisotropic metal with a layered structure, so that the conductivity is much larger with the current in-plane than out-of-plane. If the magnetic field is perpendicular to the planes, both neck and belly orbits will occur around the Fermi surface. At certain inclination angles (known as Yamaji angles) of the magnetic field, all orbits will have identical area S. The average velocity in the z-direction is proportional to $\partial S/\partial K_z$ (where K_z labels the k_z position of the orbit) and thus vanishes when S is not a function of K_z. This leads to minima in the conductivity and hence angle-dependent magnetoresistance oscillation peaks at the Yamaji angles, connected with the vanishing of the electronic group velocity perpendicular to the 2D layers. The angles θ_n at which the maxima occur are given by $k_{\parallel}^{\max} d\tan(\theta_n) = \pi(n \pm 1/4) + A(\phi)$, where the signs "$-$" and "$+$" correspond to positive and negative θ_n, respectively, d is the effective interplane spacing, $k_{\parallel}^{\max}$ is the maximum Fermi wave vector projection on the plane of rotation of the field, $A(\phi)$ is a function of ϕ and $n = \pm 1, \pm 2, \ldots$. Here, positive n correspond to $\theta_n > 0$ and negative n to $\theta_n < 0$. The gradient of a plot of $\tan\theta_n$ against n may thus be used to find one of the dimensions of the Fermi surface and, if the process is repeated for several planes of rotation of the field, the complete Fermi surface may be mapped out. The function $A(\phi)$ is determined by the inclination of the plane of warping; hence, this may also be found.

Analogous angle-dependent magnetoresistance oscillations can be found for open orbits. The mechanism is similar and easier to understand. If a magnetic field is applied in the plane of a corrugated sheet in the Fermi surface, electrons will be driven along in straight lines perpendicular to the field but in the plane of the sheet. For each Fourier component of corrugation of this sheet, the velocity is more effectively averaged when electrons are not traveling along the axis of the corrugation (**Figure 5a**) than when they are (**Figure 5b**); thus, sharp resistance minima are obtained when the orbits run along a Fourier component of the corrugation (when the magnetic field lies along a so-called "magic angle").

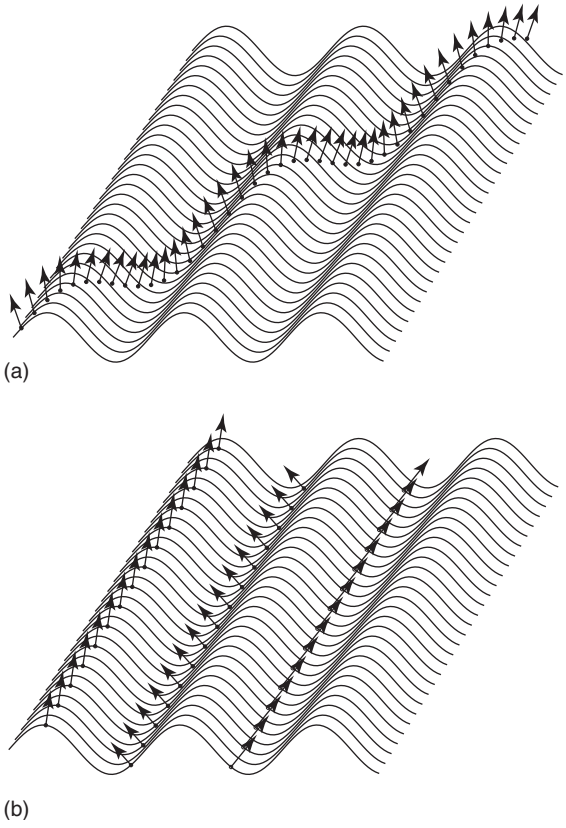


Figure 5 If a magnetic field lies in the plane of the quasi-one-dimensional sheets, all orbits become straight lines when viewed normal to that plane, along an axis which is also perpendicular to the magnetic field direction. For a given Fourier component of the corrugation, the velocity is more effectively averaged when electrons (a) are not traveling along the axis of the corrugation than (b) when they are. (Reproduced from Blundell SJ and Singleton J (1996) *Physical Review B* 53: 5609–5619.)

Other angle-dependent magnetoresistance effects are also possible, including those relating to closed orbits around small hillocks in the Fermi surface. In the semiclassical picture, all the angle-dependent magnetoresistance oscillations are caused by the degree to which the velocity components of the electrons are averaged over the series of orbits that appear at a certain inclination angle. Such an effect for a quasi-two-dimensional system is shown in **Figure 6**. In strongly correlated materials, it is also possible to see effects which are more complex in origin and reflect many-body effects associated with electron–electron correlations or special regions of k-space where the scattering rate takes a large value (Fermi surface hot spots).

Orbit Quantization

Magnetic fields also have a quantizing effect on the electron motion if the orbit is closed. The

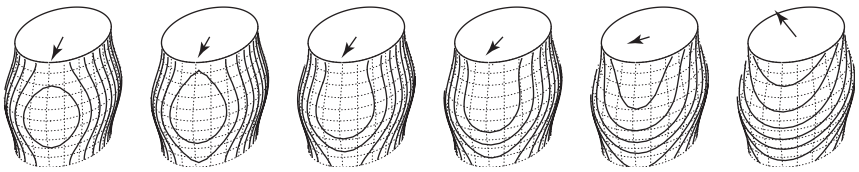


Figure 6 In a quasi-two-dimensional system, the Fermi surface may consist of a weakly warped cylinder. The figure shows how the electron orbits around this Fermi surface change as the direction of the magnetic field is rotated away from a direction lying in the quasi-two-dimensional layers. The direction of the applied magnetic field is shown by the arrow in each case. The electron orbits always lie on the Fermi surface but in a plane perpendicular to the magnetic field direction.

Bohr–Sommerfeld quantization rule $\oint \boldsymbol{p} \cdot \mathrm{d}\boldsymbol{q} = (n + \alpha)h$ (where $\boldsymbol{p}$ and $\boldsymbol{q}$ are momentum and position, respectively, n is an integer, and α is a geometrical factor which can often be taken to be 1/2) can be applied to this problem, and using $\hbar\boldsymbol{k} = \boldsymbol{p} + e\boldsymbol{A}$ yields

$$\oint (\hbar\boldsymbol{k} - e\boldsymbol{A}) \cdot \mathrm{d}\boldsymbol{R}_{\perp} = (n + \alpha)h \qquad [16]$$

where $\boldsymbol{R}_{\perp}$ is the projection of $\boldsymbol{R}$ onto a plane normal to $\boldsymbol{B}$, and $\boldsymbol{A}$ is the magnetic vector potential. Using Stokes' theorem leads to

$$\boldsymbol{B} \cdot \oint \boldsymbol{R} \times \mathrm{d}\boldsymbol{R}_{\perp} - \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = (n + \alpha)\frac{h}{e} \qquad [17]$$

and the magnetic flux through the real space orbit is given by $\Phi = \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 1/2 \oint \boldsymbol{R} \times \mathrm{d}\boldsymbol{R}_{\perp}$, this reduces to

$$\Phi = (n + \alpha)\Phi_0 \qquad [18]$$

where $\Phi_0 = h/e$ is the flux quantum. Since the area of the corresponding orbit in $\boldsymbol{k}$-space, S, is scaled up from this by a factor $\eta = \hbar/eB$, it is expected that

$$S = (n + \alpha)2\pi eB/\hbar \qquad [19]$$

so that S grows as $|\boldsymbol{B}|$ increases. When S matches S_{ext}, the extremal cross section of the Fermi surface, the free energy reaches a maximum. If the field increases further, the highest occupied Landau level becomes depleted of electrons and the free energy decreases, as does the density of states at the Fermi energy. These oscillations in the free energy and density of states as a function of magnetic field give rise to oscillations in many properties related to either the free energy or the density of states. These so-called "quantum oscillations" include not only the oscillations in the magnetization (the de Haas–van Alphen effect) but also those in the magnetoresistance (the Shubnikov–de Haas effect). Maxima in the free energy occur when $S_{\mathrm{ext}} = S$, so are governed by a periodicity

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar S_{\mathrm{ext}}} \qquad [20]$$

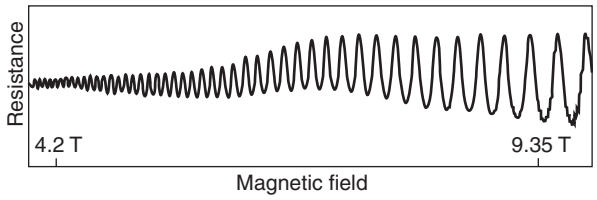


Figure 7 Shubnikov–de Haas oscillations for $AuSb_2$ recorded at 1.45 K with the current parallel to ⟨1 1 0⟩ and the magnetic field parallel to ⟨1 1 1⟩. The data show two frequencies. (Reprinted with permission from Ahn J and Sellmyer DJ (1970) *Physical Review B* 1: 1285–1297; © American Physical Society.)

or equivalently by a frequency F in reciprocal field equal to $F = S_{\mathrm{ext}}(\hbar/2\pi e)$. In fact, orbits around all cross sections of the Fermi surface (in any plane perpendicular to $\boldsymbol{B}$) give oscillatory contributions to the free energy, but the contributions from adjacent "slices" cancel out with each other so that the net effect is as if the only contribution is from extremal cross sections. Hence, measurements of these oscillations allow a precise determination of the extremal cross-sectional area of closed sections of the Fermi surface. An example of this is presented in **Figure 7** which shows Shubnikov–de Haas oscillations measured in the metallic compound $AuSb_2$. Two frequencies are clearly visible in the data, each corresponding to a different extremal cross-sectional area of the Fermi surface.

The form of the quantum oscillations can be described by the so-called Lifshitz–Kosevich formula, which is not described in detail here. Its precise form depends on whether the Fermi surface is three dimensional or two dimensional, it includes the contribution from harmonics of the fundamental frequencies, and it contains various damping factors which are discussed in the next section. This formula is valid for de Haas–van Alphen oscillations but the situation for Shubnikov–de Haas oscillations is more complex. Moreover, Shubnikov–de Haas oscillations can be harder to observe than de Haas–van Alphen oscillations in most metals because they can be swamped by the conductivity of the major portions of the Fermi surface. The Shubnikov–de Haas effect is more dramatic in semimetals or high mobility doped semiconductors where the Fermi surface is small

enough for the Shubnikov–de Haas effect to be readily seen. Two-dimensional metallic systems (whose Fermi surfaces typically contain weakly warped cylinders and sheets) are also ideal for observing Shubnikov–de Haas oscillations, since the entire Fermi surface, or at least a very large fraction of it, typically contributes to the quantum oscillations.

Phase Smearing

Shubnikov–de Haas oscillations are strongly affected by both temperature and scattering. These effects lead to phase smearing of the oscillations and a consequent reduction in the oscillation amplitude.

In the case of temperature, the blurring of the boundary between occupied and unoccupied states (governed by the Fermi function) leads to a damping factor for the oscillation amplitude equal to

$$R_{\mathrm{T}} = \pi\lambda/\sinh(\pi\lambda) \qquad [21]$$

where λ is given by

$$\lambda = 2\pi p\left(\frac{k_{\mathrm{B}}T}{\hbar\omega_{\mathrm{c}}}\right) = \frac{2\pi pm^{*}k_{\mathrm{B}}T}{e\hbar B} \qquad [22]$$

where the cyclotron frequency $\omega_{\mathrm{c}} = eB/m^{*}$ and the formula is quoted for the pth harmonic of the oscillations. When $\lambda \gg 1$, this reduces to $R_{\mathrm{T}} \propto \exp(-\pi\lambda)$, and shows that the amplitude of the oscillations becomes strongly reduced when the temperature becomes significantly larger than $\hbar\omega_{\mathrm{c}}/2\pi pk_{\mathrm{B}}$. Measurements of the temperature dependence of the oscillation amplitude can be used to extract the effective mass m^{*}.

The effect of scattering leads to a damping factor (known as the Dingle factor) $R_{\mathrm{D}} = \exp(-p\pi/\omega_{\mathrm{c}}\tau_q)$, where τ_q is the quantum lifetime of an electron in a magnetic quantum state. This can be cast in a similar form to R_{T} by rewriting it as

$$R_{\mathrm{D}} = \exp(-2\pi^2 pmk_{\mathrm{B}}T_{\mathrm{D}}/e\hbar B) \qquad [23]$$

where T_{D} is known as the Dingle temperature, and is given by

$$T_{\mathrm{D}} = \frac{\hbar}{2\pi k_{\mathrm{B}}\tau_q} \qquad [24]$$

This quantity can be extracted from the field dependence of the oscillation amplitude. The quantum lifetime τ_q is due to all scattering events, including those which result from static spatial fluctuations in the background potential of the sample. This can be different to the relaxation time τ considered earlier, which controls the mobility and low-field magnetoresistance. The relaxation time is weighted by a scattering angle factor $(1 - \cos\theta)$, where θ is the angular deviation in a particular scattering event, and is therefore much less influenced by small-angle scattering events. Thus, $1/\tau_q \propto \int W(\theta)\,\mathrm{d}\theta$ and $1/\tau \propto \int W(\theta)(1 - \cos\theta)\,\mathrm{d}\theta$, where $W(\theta)$ is the scattering cross section for scattering through an angle θ.

Landau levels (seperated in energy from each other by $\hbar\omega_{\mathrm{c}}$) are also spin split (by an amount $g\mu_{\mathrm{B}}B$) so that the energy levels are given by

$$E = \left(n + \frac{1}{2}\right)\frac{\hbar eB}{m^{*}} \pm \frac{1}{2}g\mu_{\mathrm{B}}B \qquad [25]$$

where n is the Landau level index and g is the electronic g-factor. The effect of the spin splitting of Landau levels in a magnetic field also leads to a phase smearing, and in this case it is given by $\phi = 2\pi g\mu_{\mathrm{B}}B/\hbar\omega_{\mathrm{c}}$. This results in a damping factor R_{S} given by

$$R_{\mathrm{S}} = \cos\frac{p\phi}{2} = \cos\frac{p\pi gm^{*}}{2} \qquad [26]$$

The Shubnikov–de Haas oscillations are, therefore, attenuated by a net damping factor $R_{\mathrm{T}}R_{\mathrm{D}}R_{\mathrm{S}}$.

Spin Zeros

An interesting effect can occur in the case of a quasi-two-dimensional Fermi surface, such as those which may occur in layered metals. If the magnetic field is tilted by an angle θ away from the direction normal to the layers, then the energy levels are given by

$$E = \left(n + \frac{1}{2}\right)\frac{\hbar eB\cos\theta}{m^{*}} \pm \frac{1}{2}g\mu_{\mathrm{B}}B \qquad [27]$$

Increasing the angle θ reduces the separation between Landau levels by reducing the field perpendicular to the highly conducting planes, $B\cos\theta$. When $B\cos\theta$ is such that the spin-up and spin-down sections of different Landau levels are degenerate, then the separation between successive energy levels is equal to $\hbar\omega_{\mathrm{c}}$. At this angle the Shubnikov–de Haas oscillations having the fundamental frequency F will dominate, taking their maximum amplitude. However, when $B\cos\theta$ is such that the spin-up and spin-down sections of different Landau levels are equally spaced at $(1/2)\hbar\omega_{\mathrm{c}}$, then the dominant oscillations will be those with frequency $2F$ and the amplitude of the fundamental oscillations will be at a minimum. These two situations are known as spin-maxima and spin-zeros, respectively. This effect can also be observed using magnetoresistance measurements.

An Example of Field and Angle-Dependent Magnetoresistance

The graph in **Figure 8** shows magnetoresistance data for a quasi-two-dimensional organic superconductor known as κ-(BEDT-TTF)$_2$Cu(NCS)$_2$. This material

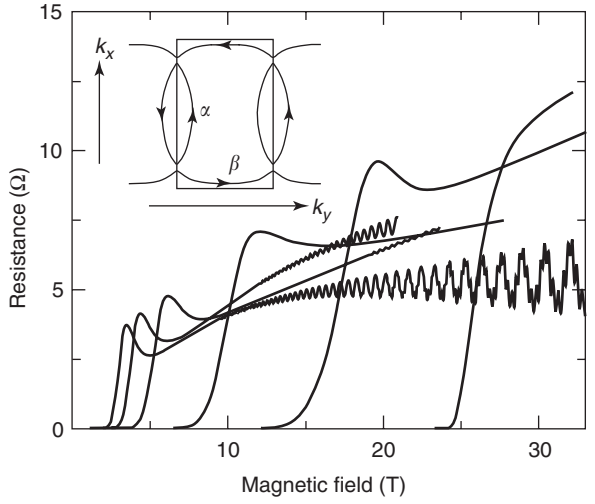


Figure 8 Magnetoresistance of the organic superconductor κ-(BEDT-TTF)$_2$Cu(NCS)$_2$ measured for various angles between the magnetic field and the conducting planes. The critical field for superconductivity is found to increase as the field is tilted into the conducting planes (right-hand trace). In some of the other traces, clear Shubnikov–de Haas oscillations are visible, due to the α-orbit. At high magnetic field, a second frequency is also visible which is the β-orbit and occurs due to a magnetic breakdown (see text). The inset shows the Fermi surface in the k_x–k_y plane and shows both the α-orbits and the open orbits, which, following magnetic breakdown, gives rise to the β-orbit. (Adapted from Nam MS, Symington JA, Singleton J, Blundell SJ, Ardavan A, *et al.* (1999) *Journal of Physics*: *Condensed Matter* 11: L477–L484.)

consists of alternating layers of the organic BEDT–TTF molecules (stacked side-to-side so that the molecular orbitals overlap) and layers of $Cu(NCS)_2^-$ anions. Two of the BEDT–TTF molecules jointly donate an electron to the anion, and the charge transfer leaves behind a hole on the BEDT–TTF molecules. This means that the bands formed by the overlap of the BEDT–TTF molecular orbitals are partially filled, leading to a metallic behavior. The transfer integrals, which parametrize the ease of hopping of electrons between BEDT–TTF molecules, are large within the BEDT–TTF planes but small in the direction perpendicular to the BEDT–TTF planes. This results in two-dimensional electronic properties. In this salt, the BEDT–TTF molecules associate into dimers, each of which collectively donates one electron to the anions, leaving behind a mobile hole. Because the dispersion is nearly isotropic in the conducting planes, the Fermi surface is approximately circular and has the same area as the first Brillouin zone. The Fermi surface cuts the Brillouin zone boundaries and at these points, a gap opens up which splits the Fermi surface into open and closed sections (see inset to **Figure 8**). The closed section has an area corresponding to $\sim$15% of the first Brillouin zone (this is known as the α pocket) but magnetic breakdown at fields above $\sim$20 T allow quasiparticles to bridge the energy gap

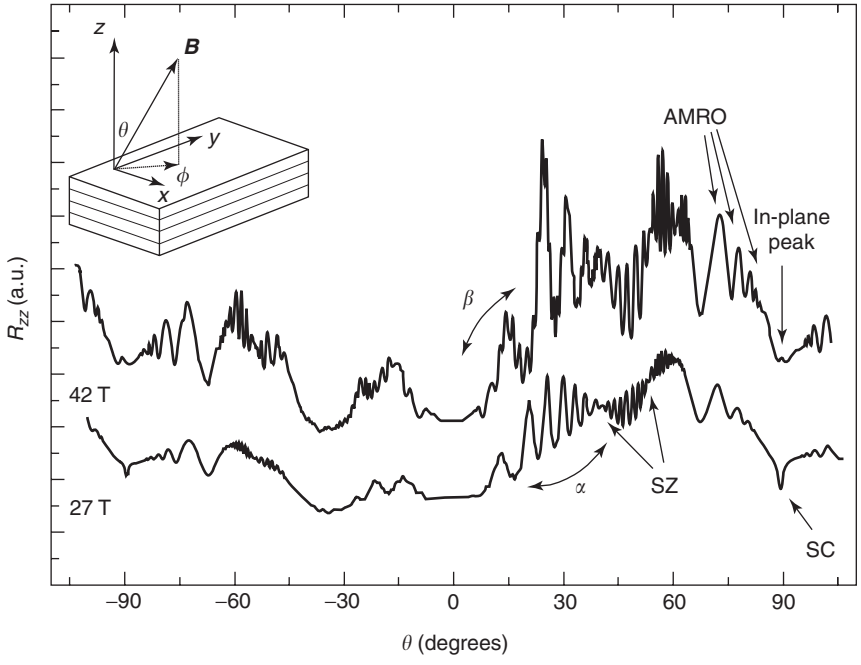


Figure 9 Typical angle dependence of the magnetoresistance of the organic superconductor κ-(BEDT-TTF)$_2$Cu(NCS)$_2$ at 0.49 K and 27 T (lower), and 42 T (upper). The data have been offset for clarity. Some representative features are indicated; Shubnikov–de Haas oscillations due to closed orbits around small pockets in the Fermi surface (α-orbits) and the breakdown orbit (β-orbit); spin-zeros in the Shubnikov–de Haas amplitudes (SZ); the onset of the superconducting transition (SC); angle-dependent magnetoresistance oscillations (AMRO), whose positions are field independent; and the resistive peak in the presence of an exactly in-plane magnetic field (in-plane peak). The inset diagram is included to illustrate the measurement geometry. (From Goddard P, Blundell SJ, Singleton J, McDonald RD, Ardavan A, *et al.* (2004) *Physical Review B* 69: 174509.)

between the two sections of the Fermi surface and make the large β-orbit, around the circular Fermi surface (see **Figure 8**).

The angle-dependent magnetoresistance shown in **Figure 9** contains a large number of features and demonstrates many of the effects described above. The plots are not symmetrical about $\theta = 0°$, reflecting the monoclinic symmetry of the crystal structure. In the upper plot, the field perpendicular to the layers around $\theta = 0°$ is sufficient for the effects of magnetic breakdown to be observed and the fast Shubnikov–de Haas oscillations due to the β-orbit are clearly seen, this time as a function of the angle (this occurs because the oscillations depend on the component of the magnetic field, $B\cos\theta$). The slower oscillations due to the α-orbit are seen in both plots and persist to higher angles. The amplitudes of these oscillations are modulated and they disappear at certain angles corresponding to spin-zeros. The positions of the features at θ-angles greater than $\sim \pm 70°$ are seen to be independent of the magnitude of the magnetic field, which reveals that they are angle-dependent magnetoresistance oscillations. Only orbits which are extremal contribute to the Shubnikov–de Haas signal, but all orbits, whether extremal or not, give rise to this angle-dependent background magnetoresistance.

In the 42 T data shown in **Figure 9**, a small peak is observed when the field lies very close to the in-plane direction, $\theta \approx 90°$. This is the in-plane peak feature (labeled SQUIT, for "suppression of quasiparticle interlayer transport"), which is due to the efficient averaging effect of closed orbits produced on the side of a corrugated cylinder in the Fermi surface when the magnetic field lies in the conducting planes (see **Figure 6**). The peak suggests a coherent nature to the interlayer transport. For highly anisotropic materials, the angular width of the in-plane peak (in radians) is approximately $2v_{\perp}^{\max}/v_{\parallel}$, where $v_{\perp}^{\max}$ is the maximum of the out-of-plane component of the quasiparticle velocity and $v_{\parallel}$ is the in-plane component parallel to the plane of rotation.

Close to $\theta = 90°$ in the lower plot, the in-plane peak is obscured by the large dip that indicates the onset of a superconducting transition. This occurs because there is a considerable anisotropy in the upper critical field of this material, and a field of 27 T is not sufficient to suppress the superconducting state when applied in a nearly in-plane direction.

These types of magnetoresistance techniques have now been applied to many systems, including superconducting cuprates, organic metals, and correlated oxides, as well as elemental metals.

See also: Fermi Surface Measurements.

PACS: 72.15.Gd; 71.18.+y; 74.70.Kn

Further Reading

Danner GM, Kang W, and Chaikin PM (1994) Measuring the Fermi surface of quasi-one-dimensional metals. *Physical Review Letters* 72: 3714–3717.

House AA, Blundell SJ, Honold MM, Singleton J, Perenboom JAAJ, *et al.* (1996) Phase boundary in the dimensionality of the angle-dependent magnetoresistance oscillations in the charge transfer salt α-(BEDT-TTF)$_{-2}$ KHg(SCN)$_{-4}$. *Journal of Physics: Condensed Matter* 8: 8829–8845.

Hussey NE, Abdel-Jawad M, Carrington A, Mackenzie AP, and Balicas L (2003) A coherent three-dimensional Fermi surface in a high-transition-temperature superconductor. *Nature* 425: 814–817.

Kartsovnik MV, Laukhin VN, Pesotskii SI, Schegolev IF, and Yakovenko VM (1992) Angular magnetoresistance oscillations and the shape of the Fermi-surface in β-$(ET)_2$ IBr_2. *Journal of Physics (France) I* 2: 89–99.

Moses P and McKenzie RH (1999) Comparison of coherent and weakly incoherent transport models for the interlayer magnetoresistance of layered Fermi liquids. *Physical Review B* 60: 7998–8011.

Pippard AB (1989) *Magnetoresistance in Metals*. Cambridge, UK: Cambridge University Press.

Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge University Press.

Singleton J (2001) *Band Theory and Electronic Properties of Solids*. Oxford University Press.

Singleton J, Goddard PA, Ardavan A, Harrison N, Blundell SJ, *et al.* (2002) Test for interlayer coherence in a quasi-two-dimensional superconductor. *Physical Review Letters* 88: 037001.

Manganites

V Z Kresin, University of California at Berkeley, Berkeley, CA, USA

Introduction

Manganites were discovered in 1950 (Jonner and van Santen). The materials are named after the manganese ion which is a key ingredient of the compounds due to its mixed-valence state. Unlike the usual ferromagnetics, the transition of manganites to ferromagnetic state is accompanied by a drastic increase in conductivity. Such transition from an insulating to a metallic and magnetic state is a remarkable fundamental feature of these materials. Another fundamental property is the colossal magnetoresistance effect

(CMR); one can observe a thousandfold (!) change in the resistance in the presence of a moderate applied magnetic field.

The chemical composition of the manganites is $A_{1-x}\,R_x\,MnO_3$; usually A≡La, Pr, Nd and R≡Sr, Ca, Ba. The compound $La_{1-x}Sr_xMnO_3$ is the most studied material.

Structure

The parent (undoped) compound $LaMnO_3$ has a perovskite structure which is a distorted cubic lattice. The ideal crystal unit is shown in **Figure 1**. It is essential that the Mn ions are caged into oxygen octahedron, so that it is surrounded by six oxygen ions (**Figure 2**). Each oxygen ion is shared by neighboring octahedra. The fivefold orbital degeneracy is split by the crystal field into two well-separated terms, t_{2g} and e_{2g} (**Figure 3**). The t_{2g}-level contains three electrons forming the so-called "*t*-core." The last *d*-electron (e_{2g}-electron) is in a loosely bound state. The e_{2g}-electron plays a crucial role in conduction and in other properties of manganites as well as in determining its magnetic order.

Doping: Phase Diagram

The phase diagram of manganites is very rich and contains many different phases. The parent compound, $LaMnO_3$, is an insulator and its transition to the conducting state is provided by doping. The doping is realized through a chemical substitution, for example $La^{3+} \rightarrow Sr^{2+}$, that is, by placing a divalent ion into the local La^{3+} position. Such substitution leads to the change in the valence of the manganese ion valence: $Mn^{3+} \rightarrow Mn^{4+}$. That is why manganites are often called "mixed-valence magnetic oxides." The four-valent state of the Mn means that the ion loses its e_{2g}-electron. The missing electron can be described as a hole which is spread over the unit cell, being shared by eight Mn ions (see **Figure 1**)

As a result, one obtains a crystal $La_{1-x}Sr_xMnO_3$, where a number of La centers are randomly substituted by the Sr-ions. Even in the presence of some holes, the crystal, at first, continues to behave as an insulator. This is due to the fact that each hole is localized and the localization corresponds to the formation of local polarons. Such insulating state is preserved with an increase in doping up to some critical value $x = x_c \approx 0.16$–0.17. At $x = x_c$, the material makes a transition into the conducting (metallic) state which persists with further doping up to $x \approx 0.5$–0.6, depending on the chosen composition.

It is remarkable that the transition at $x = x_c$ is also accompanied by the appearance of the ferromagnetic state. As was noted above, the correlation between conductivity and magnetism is the most fundamental feature of manganites. The conductivity of the best samples of the Sr-doped films at low T is of order

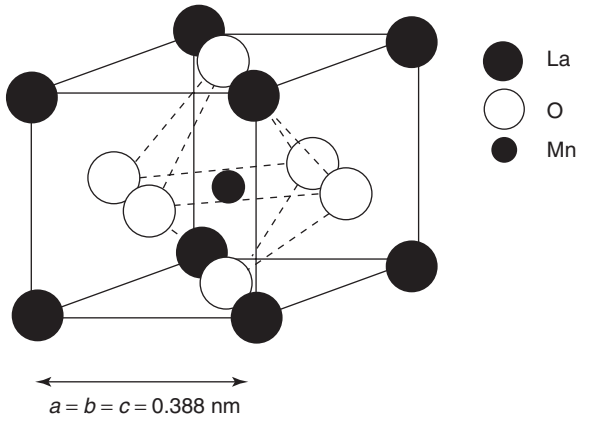


Figure 1 The ideal structure of the parent compound, $La\,MnO_3$.

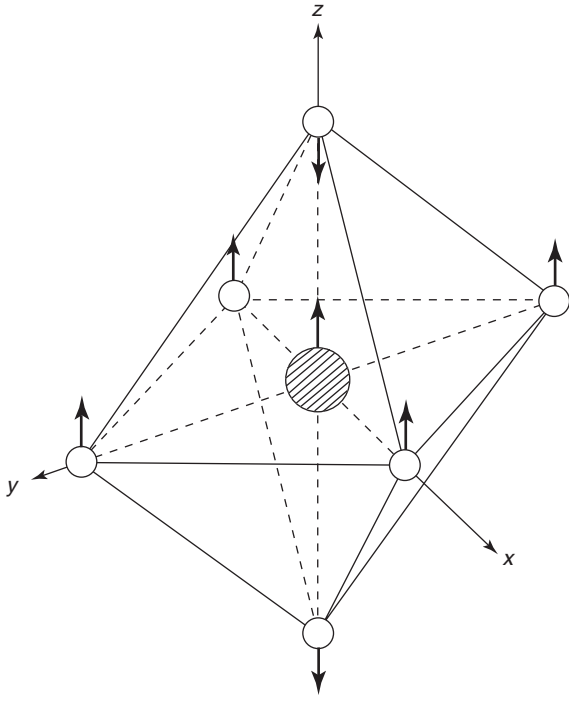


Figure 2 MnO_6 octahedron.

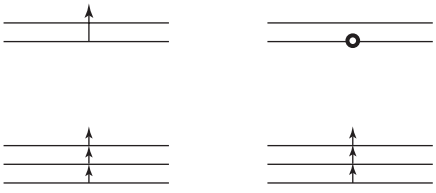


Figure 3 Electronic structures of the Mn^{3+} and Mn^{4+} ions.

$\sigma = 10^4 - 10^5\,\Omega^{-1}\,\mathrm{cm}^{-1}$, and this corresponds to a typical metallic regime.

Continuous increase in doping leads to the transition from the metallic to insulating state. For many compounds one can observe a peculiar charge-ordered state. The right end of the phase diagram (e.g., $SrMnO_3$) describes the compound which does not have the e_{2g}-electrons. Naturally, this material is an insulator.

If the compound is in the metallic ferromagnetic state (e.g., it corresponds to $0.2 < x < 0.5$ for $La_{1-x}Ca_xMnO_3$), and temperature is increased (at fixed x), then at T_C (T_C is the Curie temperature) one can observe transition to the low-conducting paramagnetic state which contains local distortions. The conductivity in this state has a hopping polaronic nature.

Main Interactions: Hamiltonian

The loosely bound e_{2g}-electron is the key player in the physics of manganites. Its Hamiltonian can be written in the form:

$$\hat{H} = \hat{H}_t + \hat{H}_{\mathrm{H}} + \hat{H}_{\mathrm{JT}} \qquad [1]$$

where

$$\hat{H}_t = \sum_{i,\delta} \hat{t}_{i,i+\delta} \qquad [2]$$

$$\hat{H}_{\mathrm{H}} = -J_H \sum_i \hat{\sigma} \cdot S_i \qquad [3]$$

$$\hat{H}_{\mathrm{JT}} = \sum_i g\hat{\tau}_i Q_i \qquad [4]$$

Here $\hat{t}$ is the two-by-two hopping matrix, $\hat{\sigma}$ is the Pauli spin matrix, S_i is the spin of the t-core, $\hat{\tau}$ is the pseudospin matrix, and Q_i is the normal mode.

The first term on the right side of eqn [1] describes the hopping of the e_{2g}-electron. The second term is the Hund's interaction which polarized the carrier's spin, and the last term corresponds to the Jahn–Teller effect which reflects the double degeneracy of the e_{2g}-term. The strength of the Hund's interaction is of the order $J_{\mathrm{H}} \approx 1$ eV; this is the largest scale of the energy. It is also essential that, because of the Hund's interaction, the conducting e_{2g}-electrons are spin-polarized. As for the hopping and the Jahn–Teller terms, they are of order $t \approx g \approx 0.1$ eV, so that $t \approx g \ll J_{\mathrm{H}}$.

The nature of the ferromagnetic spin alignment in manganites is different from that for usual ferromagnets. The concept, so-called "double exchange" mechanism was introduced in 1951 by Zener. Qualitatively, the nature of the mechanism is as follows. If one of the Mn^{3+} ions becomes four-valent (e.g., as a result of the $La^{3+} \rightarrow Sr^{2+}$ substitution, implying $Mn^{3+} \rightarrow Mn^{4+}$), a hole appears on this site. It allows for another e_{2g}-electron localized initially at the neighboring Mn^{3+} ions to jump on the new vacant place (this can be described as the hole moving in the opposite direction). But the e_{2g}-electron is spin-polarized because of the Hund's interaction with its t-core. Relative orientation of spins of the e_{2g}-electron and the "vacant" t-core drastically affects the hopping. For example, if the direction of the spin of the core for the Mn^{4+} ion is opposite to that one for the e_{2g}-electron of the neighboring Mg^{3+} ion, then the hopping is spin forbidden. At the same time, such hopping as any increase in a degree of delocalization is energetically favorable. As a result, the ground state of the ferromagnetically ordered system (all spins are polarized along one direction) energetically is below the state which is not ordered. That is why the t-cores become ferromagnetically ordered, and this, in its turn, favors "jumping" of the e_{2g}-electrons, and, therefore, conductivity. Such mechanism (it is called "double exchange") leads to ferromagnetism and to a strong correlation between ferromagnetism and conductivity.

Metallic Phase: CMR Phenomenon

Consider the sample in the metallic ferromagnetic state with a fixed carrier concentration, (e.g., $x = 0.3$) and then increase the temperature. Such ferromagnetic state persists up to the Curie temperature. For example, for the $La_{0.7}Sr_{0.3}MnO_3$ to the value of $T_C \cong 170$ K. Above this temperature the compound makes the transition into paramagnetic state. In addition, resistivity drastically increases (almost insulating state).

The conductivity of the metallic manganites is provided by the delocalized e_{2g}-electrons. It is essential that these electrons are in the double-degenerate state. This leads to the two-band picture. Presence of the two energy bands allows to explain the absence of the electron–hole symmetry for the phase diagram. In addition, this feature is important for analysis of various properties of manganites. For example, optical properties are, mainly, due to the interband transitions.

The value of the Fermi energy in the metallic manganites is small relative to ordinary metals. Its value depends on the doping level and is of the order of EF ≈ 0.2 eV. As for conductivity, its value depends strongly on the quality of the samples. As was mentioned above, for the best samples it reaches the value $\sigma \approx 10^5\ \Omega^{-1}\ \mathrm{cm}^{-1}$.

The charge transfer in the conducting ferromagnetic manganites is provided by spin-polarized electrons. Because of such spin specifics a total number

of carriers in the band of the metallic manganites is twice less than in usual metals. That is why the conducting state in manganites is called "half-metallic."

The magnetoresistance is a very important parameter of manganites. It is defined as

$$\Delta R/R_H = [R(T,H) - R(T,0)]/R(T,H) \qquad [5]$$

The magnetoresistance has a sharp peak near the Curie temperature for the manganite, and the change in resistivity caused by an applied magnetic field can be huge. Such phenomenon which is dubbed as the colossal magnetoresistance effect (CMR) has been initially observed in the ferromagnetic metallic films by Jin *et al*. in 1994. It turns out that for $La_{0.67}Ca_{0.33}MnO_3$ films $\Delta R/R_H \approx -1.3x10^3$(!); $H \approx 5$ T.

The CMR phenomenon is discussed now. Below T_C the sample is characterized by usual metallic conductivity. Contrary to it, above T_C it has a large resistivity (low conducting state with hopping polaronic transport). The presence of an external magnetic field allows to establish ferromagnetic ordering at temperatures higher than T_C ($H = 0$). The magnetic order provides metallic conductivity, because of the double exchange mechanism, and this leads to a large shift $\Delta R_H = R(T, H) - R(T, 0)$. The shift is negative, in accordance with the experimental observation and corresponds to the transition from a low conducting (almost insulating) to a metallic state.

Insulating State

The undoped material $LaMnO_3$ is an insulator and its magnetic ordering corresponds to the so-called A-structure. This structure is characterized by the presence of the ferromagnetically ordered layers with an opposite direction of the magnetization for the neighboring layers. In other words, ferromagnetism in layers is combined with the antiferromagnetic ordering in the direction perpendicular to the layers. The value of the Neel's temperature is relatively low ($\approx 10^2$ K).

The insulating behavior can be explained in the framework of the band picture. However, this picture should be modified in accordance with the Hamiltonian (eqn [1]). Unlike ordinary metals, the band is filled by spin-polarized electrons. Moreover, the degeneracy of the e_{2g}-state leads to the two-band situation. In addition, one should take into account the Jahn–Teller distortion. Because the O ion is shared by neighboring manganese octahedra, the cooperative Jahn–Teller effect is dealt with. If one includes all these factors, the energy bands appear to be filled, and the compound is the band insulator.

Doping leads to creation of holes, but it does not mean that the sample becomes a conductor. Initially holes are localized by Coulomb forces and form polarons. Such process is a favorable one for the layered system, because in the 2D case there is no energy barrier for the polaron formation. The transition to the metallic ferromagnetic state occurs at finite doping level $x = x_c$.

Percolative Transition

The doping leads to the situation when at some critical concentration of the dopant, x_c, one can observe the transition from insulating to metallic state. This transition can be described in terms of the percolation theory. Indeed, doping leads to the formation of the compound $A_{1-x}B_xMnO_3$. A position of atom B, which is a substitution for a parent atom, is completely random, and the statistical distribution of dopants leads to the percolation picture. As a result of initial doping metallic clusters are formed. These clusters are created inside the insulating matrix and eventually form the percolation "path." In accordance with the Zener's double-exchange mechanism, there is ferromagnetic order inside the clusters. The transition to the metallic ferromagnetic state corresponds to formation of the infinite cluster and, correspondingly, to an appearance of the macroscopic phase. The transition occurs at $x = x_c \approx 0.17$, and this value is, indeed, an invariant of the percolation theory. Therefore, the metal–insulator transition in manganites is a percolation phenomenon (**Figure 4**).

The percolative nature of the metal–insulator transition implies the inhomogeneity of the system, that is, one is dealing with a coexistence of two phases: metallic and insulating. This is the "phase-separation" phenomenon. This concept for the doped oxides was introduced by Gor'kov and Sokol in 1987. Manganites as well as the high-temperature superconducting cuprates are examples of such systems. Macroscopic properties at $x < x_c$ are determined by the insulating matrix, but the sample contains metallic ferromagnetic inclusions. In a similar way, the sample at $x > x_c$, being in the ferromagnetic metallic phase, contains insulating regions.

The percolative transition occurs also in the metallic sample with a fixed-level doping (e.g., $x \approx 0.3$) at $T = T_C$. The fast increase in conductivity $\sigma(T)$ at $T < T_C$ is described by the expression

$$\sigma(T < T_C) \propto (T_C - T)^{\gamma} \qquad [6]$$

where γ is a critical index.

The percolative nature of the transition at $T = T_C$ also implies the intrinsic inhomogeneities of the sample, that is, the phase separation. This means that, even below T_C, the sample, being in the metallic

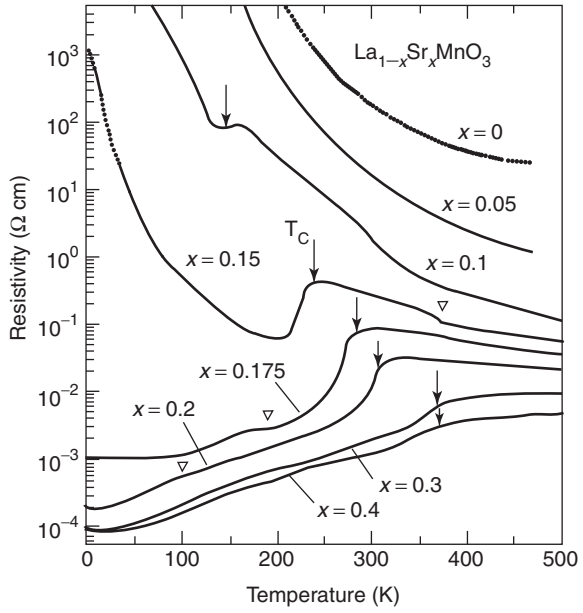


Figure 4 Temperature dependence of resistivity (for various levels of doping). The transition to the metallic state occurs at $x \approx 0.17$. (Reproduced with permission from Urishibara A, Moritomo Y, Arima T, Asamitsu A, Kido G, *et al.* (1995) Insulator–metal transition and gaint magnetoresistance in $La_{1-x}Sr_xMno_3$. *Physical Review B* 51: 14103.)

ferromagnetic state, contains inclusions of the low-conducting paramagnetic phase.

The concept of intrinsic inhomogeneity as well as the percolation picture has a strong experimental support. Various experimental techniques, such as scanning tunneling microscopy (STM), neutron scattering, X-rays absorption fine structure spectroscopy, and heat capacity data allow us to prove the picture of phase separation. For example, one can measure electronic contribution ($\propto T$) to heat capacity at $x < x_c$. The inelastic neutron scattering as well as the X-rays spectroscopy display the presence of two different bond lengths corresponding to mixed phases. An intensive study of manganites continues. Search for new materials, synthesis of high-quality samples, and a detailed theoretical and experimental analysis of all regions of the phase diagram will lead to new insights and interesting potential applications.

See also: Magnetic Materials and Applications; Magnetic Order.

PACS: 75.30.V; 81.05.Z; 75.10.Lp

Further Reading

Coey JMD, Viret M, and von Molnar S (1999) Mixed-valence manganites. *Advances in Physics* 48: 167.

Gor'Kov L and Kresin V (2004) Mixed-valence manganites: fundamentals and main properties. *Physics Reports* 400: 149.

Jin S, Tiefel TH, McCormack M, Fastnacht R, Ramesh R, and Chen LH (1994) Thousandfold change in resistivity in magnetoresistive La-Ca-Mn-O films. *Science* 264: 412.

Tokura Y (2003) Correlated-electron physics in transition-metal oxides. *Physics Today* 56: 50.

Van Tenderloo G, Lebedev O, Hervien M, and Ravean B (2004) Structure and microstructure of colossal magnetoresistant materials. *Reports on Progress in Physics* 67: 1315.

Mass Transport

R Livi, Università di Firenze, Florence, Italy

Introduction

Transport processes concern a wide range of phenomena in hydrodynamics, thermodynamics, physical chemistry, electric conduction, magnetohydrodynamics, etc. They typically occur in physical systems (gases, liquids, or solids) made of many particles (atoms or molecules) in the presence of inhomogeneities. Such a situation can result from nonequilibrium conditions (e.g., the presence of a macroscopic gradient of density, velocity, and temperature), or simply from fluctuations around an equilibrium state.

The kinetic theory of transport phenomena provides a unified description of these apparently unlike situations. It is based on the assumption that even in nonequilibrium conditions, gradients are small enough to guarantee that local equilibrium conditions still hold. In particular, the kinetic approach describes the natural tendency of the particles to transmit their properties from one region to another of the fluid by colliding with the other particles and eventually establishing global or local equilibrium conditions.

The main success of the kinetic theory is the identification of the basic mechanism underlying all the above-mentioned processes: the transport of a microscopic quantity (e.g., the mass, momentum, or energy of a particle) over a distance equal to the mean free path $\bar{\ell}$ of the particles, that is, the average free displacement of a particle after a collision with another particle. By this definition, it is implicitly assumed that the system is a fluid, where each particle is supposed to interact with each other by collisions and propagate

freely between successive collisions. Simple considerations yield the following definition

$$\bar{\ell} = \frac{1}{\sqrt{2}\sigma n} \qquad [1]$$

where n is the volumetric density of particles, which depends on the thermodynamic conditions of the system, and σ is the total cross section of collisions between particles (e.g., for the ideal case of spheric particles of radius R, $\sigma = 4\pi R^2$). For instance, in a gas at room temperature and pressure $\bar{\ell}$ is typically $\mathcal{O}(10^{-5})$cm, three orders of magnitude larger than the typical size of a particle, $\mathcal{O}(10^{-8})$cm. It is worth stressing that, reference is made to the simple case of a homogeneous isotropic system, where $\bar{\ell}$ is the same at any point and in any direction in space.

In the section "Transport in the presence of a gradient," the general case of nonequilibrium conditions are analyzed and the definition of transport coefficients is provided. In the section "Brownian motion," the process of mass transport due to fluctuations close to equilibrium is described, by relying upon the basic example of Brownian motion.

Transport in the Presence of a Gradient

Without prejudice of generality, a system is considered where a uniform gradient of the quantity $A(\vec{x})$ is established along the z-axis, and $A(x, y, z) = A(x', y', z) = A(z)$ for any x, x', y, and y'. In particular, it is assumed that $A(z)$ is a microscopic quantity, which slowly varies at constant rate along the coordinate z of an arbitrary Cartesian reference frame. Consider also a unit surface S_1 located at height z and perpendicular to the z-axis. Moving from the surface S_1 a particle will run a distance $\pm\bar{\ell}$ moving either upward or downward along the z-axis, respectively. Assuming that local equilibrium is established by collisions, the particle variable will change its original value $A(z)$ into $A(z \pm \bar{\ell})$. The net transport of the quantity $A(z)$ through S_1 amounts to the number of crossings of S_1 from each side in the unit time. Consistently with the assumption of local equilibrium, the same average velocity $\bar{v}$ is attributed to all particles crossing S_1. Isotropy and homogeneity of the system also imply that 1/3 of the particles move on average along the z-axis, half of them upward and half downward. Accordingly, S_1 is crossed along z in the unit time interval by $1/6n\bar{v}$ particles in each direction. The net amount of $A(z)$ transported through S_1 in the unit time is given by

$$\Phi(A) = \tfrac{1}{6} n\bar{v}[A(z - \bar{\ell}) - A(z + \bar{\ell})] \qquad [2]$$

Since $\bar{\ell}$ is very small compared with the macroscopic length of the fluid, one can use the approximation $A(z \pm \bar{\ell}) \simeq A(z) \pm \bar{\ell}\,\partial A/\partial z$ and rewrite eqn [2] as

$$\Phi(A) = -\frac{1}{3} n\bar{v}\bar{\ell}\frac{\partial A}{\partial z} \qquad [3]$$

The minus sign on the right-hand side of this equation indicates that A is transported in the direction opposite to the orientation of the gradient. This calculation can be performed more carefully by introducing explicitly the Maxwell distribution function of particle velocities at equilibrium. Nonetheless, one recovers the same result.

The simplest case is the presence of a density gradient along the z-axis, that is, $A(z) = n(z)$. It induces a mutual transport of particles, which diffuse through the system and

$$\Phi(n) = -\frac{1}{3} n\bar{v}\bar{\ell}\frac{\partial n}{\partial z} = D\frac{\partial n}{\partial z} \qquad [4]$$

where the quantity $D = 1/3n\bar{v}\bar{\ell}$ defines the diffusion coefficient of particles inside the fluid.

This definition can be easily extended to the case of a mixture of two different species of particles, labeled by the indices 1 and 2, which diffuse into each other. One can define the mutual diffusion coefficient as

$$D_{1,2} = \frac{n_1\bar{v}_1\bar{\ell}_1 + n_2\bar{v}_2\bar{\ell}_2}{3(n_1 + n_2)} \qquad [5]$$

The other cases of physical interest correspond to the situations where a gradient of velocity or temperature are present. In the former case, it is assumed that the fluid flows with constant macroscopic velocity $v(z)$ parallel to the (x,y) plane. In such a situation, there is a net transport of kinetic momentum $mv(z)$ (m is the mass of a particle), yielding a strain stress $\Phi(mv(z))$ between the fluid layers on the (x, y) plane:

$$\Phi(mv(z)) = -\frac{1}{3} nm\bar{v}\bar{\ell}\frac{\partial v(z)}{\partial z} = \eta\frac{\partial v(z)}{\partial z} \qquad [6]$$

where the quantity $\eta = 1/3nm\bar{v}\bar{\ell}$ defines the viscosity of the fluid. By substituting eqn [1] into the previous expression, one finds $\eta = m\bar{v}/3\sqrt{2}\sigma$. This implies that the viscosity of an ideal fluid is independent of its density, that is, of the pressure. This counterintuitive conclusion was first derived by J C Maxwell and its experimental verification sensibly contributed to establish in the scientific community a strong consensus on the "atomistic" approach of kinetic theory. It is worth stressing that such a conclusion does not hold when dealing with very dense fluids.

It remains to consider the case when $A(z)$ is the average kinetic energy of particles $\bar{\varepsilon}(z)$. At equilibrium, the energy equipartition condition yields the relation $n\bar{\varepsilon}(z) = \rho C_V T(z)$, where $\rho = mn$ is the mass

density of particles, C_V is the specific heat at constant volume, and $T(z)$ is the temperature at height z. The net flux of kinetic energy $\Phi(\bar{\varepsilon})$ can be read as the heat transported through the fluid along the z-axis

$$\Phi(\bar{\varepsilon}) = -\frac{1}{3}\rho C_V \bar{v}\bar{\ell}\frac{\partial T(z)}{\partial z} = \kappa\frac{\partial T(z)}{\partial z} \qquad [7]$$

where the quantity $\kappa = 1/3\rho C_V \bar{v}\bar{\ell} = mC_V\bar{v}/3\sqrt{2}\sigma$ defines the heat conductivity. Also κ is found to be independent of $2n$.

One can conclude that the transport coefficients, that is, the diffusion constant D, the viscosity η, and the heat conductivity κ, are closely related to each other and depend on a few basic properties of the particles, such as their mass m, their average velocity $\bar{v}$, and their mean free path $\bar{\ell}$. By comparing eqns [6] and [7] one finds the remarkable relation

$$\frac{\kappa}{\eta} = \alpha C_V \qquad [8]$$

According to the results herein reported, the value of the proportionality constant is $\alpha = 1$. In real systems this constant takes different values, which depend on specific features of particle interaction (e.g., $\alpha = 5/2$ for realistic models of monoatomic gases). The conceptual relevance of this relation is that it concerns quantities that originate from quite different conditions of matter. In fact, on the left-hand side is the ratio of two transport coefficients associated with macroscopic nonequilibrium conditions, while on the right-hand side is a typical equilibrium quantity, the specific heat at constant volume. After what has been discussed in this section, this observation is far from mysterious: by assuming that even in the presence of a macroscopic gradient of physical quantities, equilibrium conditions set in locally, the kinetic theory provides a unified theoretical approach for transport and equilibrium observables.

Brownian Motion

The basic example of transport properties emerging from fluctuations close to equilibrium is Brownian motion. This phenomenon had been observed by R Brown in 1828: small pollen particles in solution with a liquid exhibited an erratic motion, whose irregularity increased with decreasing particle size (typically, 10^{-3} cm). The long scientific debate about the origin of Brownian motion lasted over almost one century and raised strong objections to the atomistic approach of the kinetic theory (see the section, "Introduction"). In particular, the motion of the pollen particles could not be apparently consistent with an explanation based on collisions with the molecules of the liquid, subject to thermal fluctuations. In fact, pollen particles have exceedingly larger mass and size with respect to molecules. An easy calculation, based on orders of magnitude, yields the conclusion that the amount of velocity acquired by a pollen particle in a collision with a molecule is typically $10^{-6}\,\mathrm{m\,s^{-1}}$, so that the corresponding instantaneous displacement could not be practically observed, "a fortiori," with the optical devices available to R Brown in 1828. The kinetic approach implies also that the number of collisions per second of the Brownian particle with liquid molecules is gigantic and random. Thermal equilibrium conditions and the isotropy of the liquid make equally probable the sign of the small velocity variations of the pollen particle. On observable timescales, the effect of the many collisions should average to zero and the pollen particle should not acquire any net displacement from its initial position in the fluid. This way of reasoning is actually wrong, as Albert Einstein argued in his theory of Brownian motion. Einstein's methodological approach, based on the combination of simple models with basic physical principles, was very effective: he first assumed that the Brownian macroscopic particles should be considered as "big mass molecules," so that the systems composed by the Brownian particles and the solvent, where they were suspended, could be considered a mixture of two fluids at thermal equilibrium. This implies the validity of Vant'Hoff law

$$P(\boldsymbol{x}) = k_B T n(\boldsymbol{x}) \qquad [9]$$

where P is the osmotic pressure between the fluids, T is the equilibrium temperature of the solution, n is the volumetric density of Brownian particles (the "solute" of the binary mixture) and k_B is the Boltzmann constant ($k_B = 1.380658\,10^{-23}\,\mathrm{JK^{-1}}$). According to the classical kinetic theory one should not be allowed to write an equation of state like [9] for a collection of the macroscopic Brownian particles. They should be better described by the laws of dynamics, rather than by those of thermodynamics. Einstein's other point of view can be justified by considering that, despite their macroscopic dimension, the Brownian particles are not subject to standard macroscopic forces: they experience microscopic forces originated by thermal fluctuations of the fluid molecules and in this sense they can be assimilated in any respect to soluted particles. The macroscopic nature of the Brownian particles should be recovered in the timescales associated with the collision and dissipation processes inside the fluid. Einstein guessed that Brownian motion looks random on a much larger timescale than the one needed for dissipating the energy acquired by Brownian particles through collisions with molecules. In practice, this amounts

to the assumption that the dissipation mechanism should be described by the macroscopic Stokes' law

$$F = m\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -6\pi\eta R\boldsymbol{v} \qquad [10]$$

where F is the friction force proportional to the velocity v of the Brownian particle of radius R and η is the viscosity of the solvent. Accordingly, the energy dissipation process of Brownian particles has to occur on the timescale

$$t_\mathrm{d} = \frac{m}{6\pi\eta R} \qquad [11]$$

For $R \simeq 10^{-4}$ cm and for a standard solvent (e.g., water) one has $t_\mathrm{d} = \mathcal{O}(10^{-7})$s.

On a timescale much larger than t_d the Brownian particles are expected to exhibit a diffusive motion, as a consequence of the many random collisions with the solvent molecules. On the other hand, the kinetic theory associates diffusion with transport of matter in the presence of a density gradient (see the section "Introduction"). Einstein then assumed that some ideal external force should be responsible for creating such a gradient yielding a flux of Brownian particles

$$\boldsymbol{J}_\mathrm{B} = D\boldsymbol{\nabla}n \qquad [12]$$

At a microscopic level the flux of Brownian particles $\boldsymbol{J}_\mathrm{B}$ is defined as the product of their velocity $\boldsymbol{v}$ times their density n. For $t \gg t_\mathrm{d}$, the velocity $\boldsymbol{v}$ is given by eqn [10] and one can also write

$$\boldsymbol{J}_\mathrm{B} = \frac{\boldsymbol{F}n}{6\pi\eta R} \qquad [13]$$

The numerator on the right-hand side of this equation is the definition of the pressure gradient, which, according to Vant'Hoff law [9], is proportional to the density gradient. In formulas

$$\boldsymbol{F}n = \boldsymbol{\nabla}P = k_\mathrm{B}T\boldsymbol{\nabla}n \qquad [14]$$

By substituting eqn [14] into eqn [13] and comparing with eqn [12], one obtains the remarkable relation

$$D = \frac{k_\mathrm{B}T}{6\pi\eta R} \qquad [15]$$

Perrin showed that this formula provides very good quantitative agreement with experimental data.

In summary, Einstein's approach derives from the assumption that for $t < t_\mathrm{d}$ it still makes sense attributing a velocity to the Brownian particle (see eqn [10]), while for $t > t_\mathrm{d}$ the particle is assumed to diffuse, as it would in the presence of an ideal density gradient (eqn [12]). It is worth stressing that the last hypothesis is purely phenomenological, since direct inspection of the movement of Brownian particles revealed its diffusive nature. Einstein's theory of Brownian motion is based on the matching between different dynamical regimes.

The Langevin Approach

A mechanical approach taking explicitly into account both dynamical regimes was later proposed by P Langevin. For the first time deterministic and stochastic forces were introduced into the same equation. The deterministic component is the friction term $-\gamma\boldsymbol{v}$, acting on the Brownian particle: γ is the friction coefficient, which amounts to $\gamma = 6\pi\eta R$ in the Stokes regime and $\boldsymbol{v}$ is the three-dimensional velocity vector of the Brownian particles, whose components are denoted by v_i, $i = 1, 2, 3$. The stochastic component associated with the effect of collisions with the solvent molecules is a random, time-dependent force $\boldsymbol{\xi}(t)$, whose components are analogously denoted by ξ_i, $i = 1, 2, 3$. Symmetry arguments indicate that each one of these components can be assumed to be independent, isotropic, uncorrelated in time (at least for $t > t_\mathrm{d}$) and Gaussian, since it results from the combination of a very large number of collisions, which can be considered approximately as independent events. Accordingly, the average of the stochastic force is null and its time-correlation function can be approximated by a Dirac delta for $t > t_\mathrm{d}$. In the formulas:

$$\langle \xi_i(t) \rangle = 0 \qquad [16]$$

$$\langle \xi(t)_i \xi_j(t') \rangle = \Gamma\delta_{ij}\delta(t - t') \qquad [17]$$

where the brackets $\langle\ \rangle$ indicate the statistical average, Γ is a suitable dimensional constant and δ_{ij} is a Kronecker delta. Langevin equation is a Newton equation containing the stochastic force $\boldsymbol{\xi}$. For each component of the three-dimensional velocity vector it reads

$$m\frac{\mathrm{d}v_i}{\mathrm{d}t} = -\gamma v_i + \xi_i(t) \qquad [18]$$

This equation can be formally integrated, yielding the solution

$$v_i(t) = \exp\left(-\frac{\gamma}{m}t\right) \times \left[v_{0i} + \int_0^t \mathrm{d}\tau \frac{1}{m}\exp\left(\frac{\gamma}{m}\tau\right)\xi_i(\tau)\right] \qquad [19]$$

where v_{0i} is the ith component of the initial velocity. Accordingly, $v_i(t)$ is a stochastic solution which assumes properties similar to $\xi_i(t)$ (see eqns [16] and [17]), namely

$$\langle v_i(t) \rangle = v_{0i}\exp\left(-\frac{\gamma}{m}t\right) \qquad [20]$$

and

$$\langle v_i^2(t)\rangle = \exp\left(-\frac{2\gamma}{m}t\right) \times \left[v_{0i}^2 + \int_0^t \mathrm{d}\tau \frac{1}{m^2}\exp\left(\frac{2\gamma}{m}\tau\right)\Gamma\right] \quad [21]$$

$$= \frac{\Gamma}{2m\gamma}\left[1 - \exp\left(-\frac{2\gamma}{m}t\right)\right] + v_{0i}^2 \exp\left(-\frac{2\gamma}{m}t\right) \quad [22]$$

As time t grows the average velocity vanishes exponentially with the rate m/γ, while the average squared velocity approaches the value

$$\lim_{t\to\infty} \langle v_i^2(t)\rangle = \frac{\Gamma}{2m\gamma} \quad [23]$$

In order to bridge this purely mechanical approach with Einstein's theory, establishing a relation with thermodynamics remains. This is naturally obtained by attributing a thermal origin to the fluctuations of the stochastic force. In particular, if one assumes that the Brownian particles and the solvent are at thermal equilibrium with temperature T, the energy equipartition principle establishes that the average kinetic energy per degree of freedom in the limit of very large times is proportional to the temperature

$$\lim_{t\to\infty} \left\langle \tfrac{1}{2} m v_i^2(t)\right\rangle = \tfrac{1}{2} k_\mathrm{B} T \quad [24]$$

By comparison with eqn [23] one obtains the remarkable relation

$$\Gamma = 2\gamma k_\mathrm{B} T \quad [25]$$

where the amplitude of velocity fluctuations is found to depend on the frictional coefficient of the fluid, thus yielding the basic example of a fluctuation–dissipation relation.

On the other hand, this relation cannot be directly tested by experimental observations, that is, records of the erratic trajectories of Brownian particles sampled at time intervals much larger than t_d. The very predictive content of the Langevin equation can be extracted by computing the average displacement vector $\boldsymbol{r}$, whose components are related to those of $\boldsymbol{v}$ by the straightforward relation $r_i(t) = \int_0^t \mathrm{d}\tau v_i(\tau),\ \ i = 1, 2, 3$. The information of interest is contained in the expression of the squared average displacement $\langle r_i^2(t)\rangle$. Simple calculations yield the result

$$\langle r_i^2(t)\rangle = v_{0i}^2 t^2, \quad t \ll \frac{m}{\gamma} \quad [26]$$

$$\langle r_i^2(t)\rangle = \frac{2k_\mathrm{B}T}{\gamma}t, \quad t \gg \frac{m}{\gamma} \quad [27]$$

It is important to observe that this mechanical approach is able to account at the same time for the "ballistic" regime, expected for times smaller than the relaxation time $t_\mathrm{d} = m/\gamma$, and also for the asymptotic diffusive regime. The diffusion constant is expressed by the relation

$$D = \frac{k_\mathrm{B}T}{\gamma} \quad [28]$$

which is found to be equivalent to eqn [15] in the Stokes friction regime, where $\gamma = 6\pi\eta R$.

The Fokker–Planck Approach

The statistical content of the Langevin approach emerges explicitly from the properties attributed to the random force in eqn [17]. The statistical average $\langle\ \rangle$ can be interpreted as the result of the average over many different trajectories of the Brownian particle obtained by different realizations of the stochastic force components ξ_i in eqn [18]. By taking inspiration from Boltzmann kinetic theory one can get rid of mechanical trajectories and describe from the very beginning the evolution of Brownian particles by a suitable distribution function $f(\boldsymbol{x}, t) : f(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x}$ represents the probability of finding the Brownian particle at a position in between $\boldsymbol{x}$ and $\boldsymbol{x} + \mathrm{d}\boldsymbol{x}$ at time t. One can introduce the transition rate $W(\boldsymbol{x}', \boldsymbol{x})$, such that $W(\boldsymbol{x}', \boldsymbol{x})\mathrm{d}\boldsymbol{x}'\Delta t$ represents the probability that in the interval of time Δt the Brownian particle "jumps" from the neighborhood of $\boldsymbol{x}$ to $\boldsymbol{x}'$. In this approach, any reference to instantaneous collisions with the solvent molecules is lost, and one has to assume that $\Delta t \gg t_\mathrm{d}$. As proposed by Fokker and Planck, one can write the evolution equation for $f(\boldsymbol{x},\ t)$: as a "master equation," that is, a balance equation where the variation in time of $f(\boldsymbol{x},\ t)$ emerges as the result of two competing terms: a gain factor, due to jumps of particles from any position $\boldsymbol{x}'$ to $\boldsymbol{x}$, and a loss factor, due to jumps from $\boldsymbol{x}$ to any $\boldsymbol{x}'$. In formula

$$\frac{\partial f(\boldsymbol{x}, t)}{\partial t} = \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{x}' [f(\boldsymbol{x}', t) W(\boldsymbol{x}, \boldsymbol{x}') - f(\boldsymbol{x}, t) W(\boldsymbol{x}', \boldsymbol{x})] \quad [29]$$

In this equation, it has been implicitly assumed that volumes are preserved in position space by the Brownian process, in particular $\mathrm{d}\boldsymbol{x}' = \mathrm{d}\boldsymbol{x}$. On the other hand, large displacements of the Brownian particles are highly improbable and one can reasonably assume that the rate function $W(\boldsymbol{x}', \boldsymbol{x}) = W(\boldsymbol{x}'; \boldsymbol{\chi})$ is significantly different from zero only for very small values of $|\boldsymbol{\chi}| = |\boldsymbol{x}' - \boldsymbol{x}|$. This allows to introduce a formal Taylor series expansion in eqn [29] around $\boldsymbol{\chi} = 0$. By considering terms only up to the second order, one obtains the Fokker–Planck equation

$$\frac{\partial f(\boldsymbol{x}, t)}{\partial t} = -\boldsymbol{\nabla}[\boldsymbol{\mu}_1 f(\boldsymbol{x}, t)] + \frac{1}{2}\Delta[\boldsymbol{\mu}_2 f(\boldsymbol{x}, t)] \quad [30]$$

where

$$\boldsymbol{\mu}_1 = \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{\chi} W(\boldsymbol{x}';\boldsymbol{\chi})\boldsymbol{\chi} = \left\langle \frac{\delta \boldsymbol{x}}{\delta t} \right\rangle = \langle \boldsymbol{v} \rangle \quad [31]$$

$$\boldsymbol{\mu}_2 = \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{\chi} W(\boldsymbol{x}';\boldsymbol{\chi})|\boldsymbol{\chi}|^2 = \left\langle \frac{|\delta \boldsymbol{x}|^2}{\delta t} \right\rangle \quad [32]$$

The first term on the right-hand side of eqn [30] represents the contribution of viscous friction, while the second term is the diffusive contribution. Notice that $\boldsymbol{\mu}_1$ and μ_2 are both average quantities: the former amounts to the average speed of a Brownian particle, which in general results from the balance of external forces acting on the Brownian particle with the viscous friction force produced by the solvent; the latter amounts to the average squared displacement per unit time, which, in a homogeneous isotropic medium, corresponds to the definition of the diffusion constant D. Accordingly, in the absence of external forces the viscous term can be neglected and eqn [30] reduces to the standard form of a macroscopic diffusion equation

$$\frac{\partial f(\boldsymbol{x},t)}{\partial t} = D\Delta f(\boldsymbol{x},t) \quad [33]$$

The solution of this equation is

$$f(\boldsymbol{x},t) = \frac{1}{(4\pi D t)^{3/2}} \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{x}_0|^2}{4Dt}\right) \quad [34]$$

where $\boldsymbol{x}_0$ is the initial position of the Brownian particle. The average squared displacement is given by the expression

$$\langle |\boldsymbol{x}|^2 \rangle = \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{x}|\boldsymbol{x}|^2 f(\boldsymbol{x},t) = |\boldsymbol{x}_0|^2 + 6Dt \quad [35]$$

This statistical treatment of Brownian motion can be easily generalized in the presence of external forces acting on the Brownian particle. For the sake of simplicity, the case of an elastic force acting on the Brownian particle $\boldsymbol{F} = -\lambda \boldsymbol{x}$, where λ is the elastic constant, is illustrated. The mobility of Brownian particles is denoted as v (e.g., in the Stokes regime $v^{-1} = 6\pi\eta R$). The average viscous force is then given by $-\langle \boldsymbol{v} \rangle / v$ and it must be balanced by the elastic force, that is $\langle \boldsymbol{v} \rangle / v + \lambda \boldsymbol{x} = 0$. This relation implies

$$\boldsymbol{\mu}_1 = -\lambda v \boldsymbol{x} \quad [36]$$

Since eqn [35] still holds, one also obtains

$$\mu_2 = \langle |\boldsymbol{x}|^2 \rangle / t = 6D \quad [37]$$

where, without prejudice of generality, one can set $\boldsymbol{x}_0 = 0$. Equation [30] then reduces to

$$\frac{\partial f(\boldsymbol{x},t)}{\partial t} = \lambda v \boldsymbol{\nabla} \cdot \boldsymbol{x} f(\boldsymbol{x},t) + 3vk_\mathrm{B}T\Delta f(\boldsymbol{x},t) \quad [38]$$

where the generalized Einstein's relation $D = vk_\mathrm{B}T$ has been used. The solution of eqn [38] reads

$$f(\boldsymbol{x},t) = \sqrt{\frac{\lambda}{2\pi k_\mathrm{B}T(1-\mathrm{e}^{-2\lambda v t})}} \times \exp\left(-\frac{\lambda(\boldsymbol{x}-\boldsymbol{x}_0\mathrm{e}^{-\lambda v t})^2}{2k_\mathrm{B}T(1-\mathrm{e}^{-2\lambda v t})}\right) \quad [39]$$

One can easily obtain

$$\langle \boldsymbol{x} \rangle = \boldsymbol{x}_0 \mathrm{e}^{-\lambda v t} \quad [40]$$

$$\langle |\boldsymbol{x}|^2 \rangle = |\boldsymbol{x}_0|^2 \mathrm{e}^{-2\lambda v t} + \frac{k_\mathrm{B}T}{\lambda}(1-\mathrm{e}^{-2\lambda v t}) \quad [41]$$

In the limit $t \to \infty$ these relations reduce to

$$\langle \boldsymbol{x} \rangle = 0 \quad [42]$$

$$\langle |\boldsymbol{x}|^2 \rangle = \frac{k_\mathrm{B}T}{\lambda} \quad [43]$$

It can be concluded that asymptotically the distribution of Brownian particles is a Gaussian centered around the origin (which has been chosen as the equilibrium position of the elastic force), with variance $k_\mathrm{B}T/\lambda$.

See also: Irreversible Thermodynamics and Basic Transport Theory in Solids.

PACS: 83.10.Mj; 52.25.Fi; 05.40. – a; 05.40.Jc

Further Reading

Brush S (1976) *The Kind of Motion That we Call Heat*. North-Holland.

Chandler D (1987) *Introduction to Modern Statistical Mechanics*. Oxford University Press.

Einstein A (1956) *Investigations on the Theory of the Brownian Movement*. Dover.

Kubo R, Toda M, and Hashitsume N (1985) *Statistical Physics II*. Springer.

Nelson E (1976) *Dynamical Theories of Brownian Motion*. Princeton University Press.

Ornstein LS and Uhlenbeck LS (1930) *Physical Review* 36: 823.

Perrin J (1948) *Les Atomes*. Gallimard.

van Kampen NG (1962) *Fundamental Problems in Statistical Mechanics*. In: Cohen, E.G.D. (ed.) North-Holland.

Mechanical Properties: Anelasticity

W Benoit, Institut de Physique de la Matière Complexe (IPMC-FSB), Lausanne, Switzerland

Introduction

The concept of anelasticity was conceived by Zener in 1948 and published in his monograph, "Elasticity and anelasticity of metals." The ideal elastic behavior is defined by three conditions: (1) the strain response to each level of applied stress has a unique equilibrium value, (2) the equilibrium response is achieved instantaneously, and (3) the response is linear. If only the second condition is not met, a time dependence with the response appears and the behavior known as anelasticity is produced. Consequently, anelasticity implies that, in addition to an elastic response, there also exists a time-dependent nonelastic response. In this sense, anelasticity differs from the related topics of "internal friction" or "mechanical spectroscopy," which does not possess such a unifying theoretical base. The more general behavior by additionally lifting the condition (1) of complete recoverability is known as linear viscoelasticity, which thus includes anelasticity as a special case. The purpose here is to show how anelasticity is described and how internal friction measurements are interpreted to get an insight into microscopic mechanisms, which are responsible for the mechanical behavior of solids. The first part comprises the formal theory using two different viewpoints: rheological approach and thermodynamics. In the second part, mechanical spectroscopy is described and different applications, that have been used to study the behavior of point defects, dislocation dynamics and grain boundary sliding, are discussed.

Rheology

Rheology solves the problem of the deformation of solids by postulating, for every kind of material, an equation between the stress σ, the deformation ε, and their time derivatives. Moreover, in the development of the rheological equations, it is useful to build mechanical models, which simulate qualitatively the behavior described by these equations. To describe the anelastic behavior, the models can be expressed with two ideal bodies (1) the Hooke solid that is conveniently represented by a spring, with the law $\varepsilon = J\sigma$, where J is the compliance or the inverse of the Young modulus E; (2) the Newtonian liquid is represented by a dashpot with the law $\sigma = \eta\dot{\varepsilon}$, where η is the viscosity. To describe anelastic behavior, these two elements can be coupled either in parallel (Voigt model) or in series (Maxwell model). It is convenient to write the viscosity η of the dashpot as $\tau/\delta J$, where τ is a relaxation time. The model of a spring in series with a Voigt unit (**Figure 1a**) can easily be calculated to obtain the differential stress–strain equation:

$$\left.\begin{aligned}\varepsilon &= \varepsilon_1 + \varepsilon_2\\ \varepsilon_2 &= J_U\sigma\end{aligned}\right\} \Rightarrow \varepsilon_1 = \varepsilon - J_U\sigma$$

and

$$\sigma = \frac{\varepsilon_1}{\delta J} + \eta\dot{\varepsilon}_1 \Rightarrow \sigma = \frac{1}{\delta J}(\varepsilon - J_U\sigma) + \frac{\tau}{\delta J}(\dot{\varepsilon} - J_U\sigma)$$

Finally, one obtains

$$J_R\sigma + \tau J_U\dot{\sigma} = \varepsilon + \tau\dot{\varepsilon} \quad \text{with} \quad J_R = J_U + \delta J \quad [1a]$$

J_R and J_U are called the relaxed compliance and the unrelaxed compliance, respectively. This equation is the more general form of the so-called "standard anelastic solid." In order to define the relaxation parameters, the following experiment can be performed: a mechanical stress σ_0 is abruptly applied at time $t = 0$ (**Figure 2**) and held constant, while the strain ε is recorded as a function of time. Equation [1a] then has the following solution:

$$\varepsilon = \sigma_0(J_R - \delta J e^{-t/\tau})$$

Thus, the strain relaxation behavior of the standard anelastic solid is such that under a unit-applied stress, the strain increases from an instantaneous value $J_U\sigma$

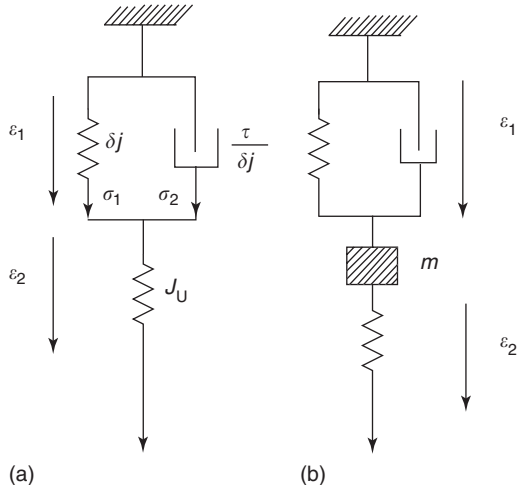


Figure 1 Rheological models: (a) standard anelastic solid: a Voigt unit (a spring and a dashpot in parallel) in series with a spring, (b) model with a mass, which may display resonance or relaxation behavior.

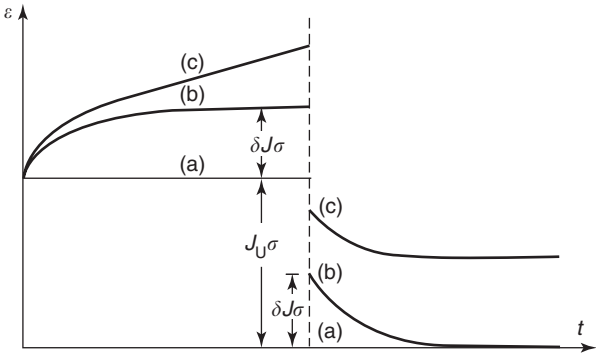


Figure 2 Creep and elastic after effect for (a) ideal elastic solid, (b) anelastic solid, and (c) linear viscoelastic solid.

to a final equilibrium value $J_R\sigma$, with a relaxation time τ. It is convenient to define a dimensionless quantity Δ, called the relaxation strength

$$\Delta = \frac{\delta J}{J_U} \text{ and then } J_R = J_U(1+\Delta)$$

The standard anelastic model can be generalized by introducing a second, third, … Voigt unit in series with the first one. The differential stress–strain equation can be calculated to obtain a linear equation with constant coefficients

$$a_0\sigma + a_1\dot{\sigma} + a_2\ddot{\sigma} + \cdots + a_n\frac{\mathrm{d}^n\sigma}{\mathrm{d}t^n} = b_0\varepsilon + b_1\dot{\varepsilon} + b_2\ddot{\varepsilon} + \cdots + b_n\frac{\mathrm{d}^n\varepsilon}{\mathrm{d}t^n}$$

This higher-order equation can also represent a resonance behavior depending on the value of the constant. Effectively, by adding a mass to the Voigt unit (**Figure 1b**), the stress–strain equation becomes

$$J_R\sigma + \tau J_U\dot{\sigma} + \delta J m J_U\ddot{\sigma} = \varepsilon + \tau\dot{\varepsilon} + \delta J m\ddot{\varepsilon} \qquad [1b]$$

This equation is also a linear differential equation of second order in ε and σ. Nevertheless, depending on the relative value of m and $\tau^2/\delta J$, this model represents a relaxation or resonant behavior. In the resonant case, the resonant frequency is given by ω_0 with $\omega_0^2 = 1/\delta J m$ (**Figure 3**). It is important to mention that, in solids, there is no example where a resonant behavior has been observed.

Thermodynamics

In this section, the basis behind the classical thermodynamic approach of anelasticity is described. The underlying idea behind such a description is that pseudo-equilibrium can be established within the solid. Thermodynamic functions, as for example, the free energy, are not only a function of the classical external state variables (stress σ, strain ε, temperature

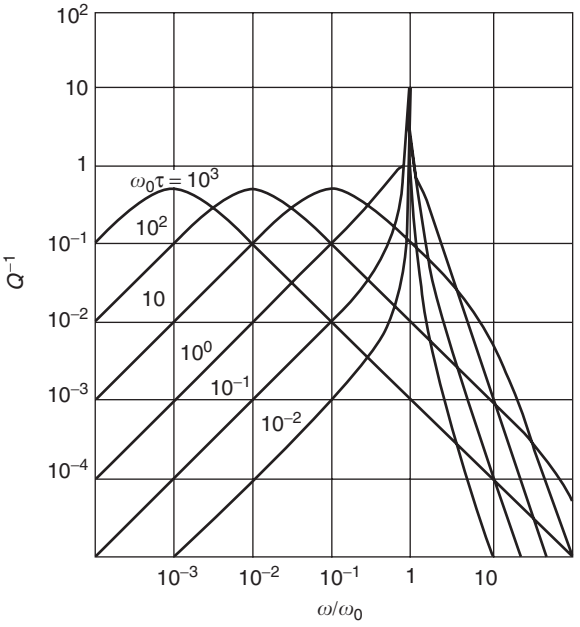


Figure 3 Frequency dependence of the internal friction for different values of the parameter $\omega_0\tau$.

T, and entropy s), but also a function of some internal variables. Such internal variables are related to microscopic mechanisms in the material and may be, for example, either a short-or a long-range order parameter, the proportion of point defects having a typical orientation, or the swept area of dislocations. In the following, only the simplest case for which the state of a solid is completely defined by the temperature, the state of stress and only one internal variable ξ is considered. Consequently, the free energy g is given by

$$g = g(\sigma, T, \xi) \quad \text{and} \quad \mathrm{d}g = -\varepsilon\mathrm{d}\sigma - s\mathrm{d}T - A\mathrm{d}\xi$$

where ε, s, and the affinity A are conjugate variables to σ, T, and ξ, respectively. The complete equilibrium at constant stress and temperature is given by $\mathrm{d}g = 0$, therefore $A = 0$. By taking as a reference state the following values of the variable, $T = T_0$ for the temperature (assuming that this value stays constant), $\sigma = 0$ for the stress, and $\xi = 0$ for the internal variable, the Gibbs function g may be expanded as a Taylor series to quadratic terms in the variables σ and ξ

$$g(\sigma, \xi, T_0) = g(0, 0, T_0) - \frac{1}{2}J_U\sigma^2 - \kappa\sigma\xi + \frac{1}{2}\beta\xi^2$$

where J_U is the unrelaxed compliance. The conjugate variables are then given by

$$\begin{aligned} \varepsilon &= J_U\sigma + \kappa\xi \\ A &= \kappa\sigma - \beta\xi \end{aligned} \qquad [2]$$

In this way, the value of ε is zero for $\sigma = 0$ and $\xi = 0$. The term $\varepsilon_{an} = \kappa\xi$ is the anelastic strain in which κ corresponds to the coupling between the internal variable ξ and the mechanical strain. When σ is assumed to be constant, but different from zero, one has to set $A = 0$ for complete equilibrium, and the equilibrium value $\bar{\xi}$ is given by

$$\bar{\xi} = \frac{\kappa}{\beta}\sigma = \alpha\sigma \qquad A = -\beta(\xi - \bar{\xi}) \qquad [3]$$

One may then consider A as a driving force toward a complete equilibrium state. From the thermodynamics of irreversible processes, it is expected that the rate $\dot{\xi}$, for small departures from equilibrium, depends linearly upon affinity,

$$\dot{\xi} = MA \text{ and consequently,}$$
$$\dot{\xi} = \frac{-1}{\tau}(\xi - \bar{\xi}) \quad \text{or} \quad \tau\dot{\xi} + \xi = \alpha\sigma \qquad [4]$$

where $\tau = 1/M\beta$ a relaxation time. Eliminating ξ and $\dot{\xi}$ in eqns [2] and [4], a differential stress–strain relation is obtained:

$$\varepsilon + \tau\dot{\varepsilon} = J_R\sigma + \tau J_U\dot{\sigma}, \quad J_R = J_U + \kappa\alpha = J_U + \delta J$$

This equation is equivalent to eqn [1] of the standard anelastic solid.

Mechanical Spectroscopy

In the first part, the relaxation parameters were measured using quasistatic experiments (**Figure 2**). Nevertheless, it is generally more convenient to use dynamic methods. In that case, an alternative stress $\sigma = \sigma_0 e^{i\omega t}$ is imposed on the system, and the phase lag δ between the strain $\varepsilon = \varepsilon_0 e^{i(\omega t - \delta)}$ and the applied stress is measured. The complex compliance is then given by

$$J^*(\omega) = J_1(\omega) - iJ_2(\omega) \quad \text{and consequently } tg\delta = \frac{J_2}{J_1}$$

It is easy to show that, in the linear case, the ratio of the energy ΔW, dissipated in the form of heat during one cycle, to the maximum stored elastic energy W is related to the loss angle δ by

$$\text{IF} = Q^{-1} = \frac{1}{2\pi}\frac{\Delta W}{W} = tg\delta$$

where Q^{-1}, the inverse of the quality factor of a resonant system, is commonly called the internal friction (IF, sometimes called damping) of a material. It generally and strongly depends on the frequency as well as on the temperature.

In order to calculate $tg\delta$ or the IF due to a relaxation mechanism, alternative stress, $\sigma = \sigma_0 e^{i\omega t}$ and strain $\varepsilon = \varepsilon_0 e^{i(\omega t - \delta)}$ are introduced in the equation of the standard anelastic solid (1), and one then obtains

$$\varepsilon + i\omega\tau\varepsilon = J_R\sigma + i\omega\tau J_U\sigma$$

The ratio between the stress and the strain gives the complex compliance $J^* = J_U + \delta J/(1 + \omega^2\tau^2) - i\delta J(\omega\tau)/(1 + \omega^2\tau^2) = J_1 - iJ_2$. Assuming $\delta J \ll J_U$, the IF is then given by classical Debye equations,

$$tg\delta = \frac{J_2}{J_1} = \Delta\frac{\omega\tau}{1 + \omega^2\tau^2} \quad \text{and} \quad \frac{\delta J_1}{J_U} = \Delta\frac{1}{1 + \omega^2\tau^2} \qquad [5]$$

where $\Delta = \delta J/J_U$. The second equation corresponds to the so-called modulus defect. In **Figure 3**, the IF is plotted as a function of the frequency for both cases of a relaxation and a resonant behavior. In the case of a relaxation behavior, a maximum is observed for $\omega\tau = 1$. For thermally activated mechanisms, an Arrhenius law is often valid, with

$$\tau = \tau_0 \exp(E/kT)$$

where τ_0 is the limit relaxation time (inverse of the attempt frequency), k the Boltzmann constant, and E the activation energy of the mechanism. The importance of such a relation is that the quantity τ may be varied over a wide range simply by changing the temperature. For the case of a Debye peak, the condition that $\ln\omega\tau = 0$ at the peak maximum gives

$$\ln\omega = \ln\omega_0 - \frac{E}{kT_P}, \quad \text{where } \omega_0\tau_0 = 1 \qquad [6]$$

If a series of peaks is measured at a number of different frequencies, a straight-line plot can be made (Arrhenius plot) of $\ln\omega$ as a function of $1/T_P$ whose slope is then E/k.

Experimental Methods

In order to measure the IF, different methods can be used, depending on the frequency range. In the very low frequency range (10^{-5}–10 Hz), a forced vibration is imposed on the system, and the phase lag between stress and strain is directly measured either as a function of the frequency or as a function of the temperature (**Figure 4**). Between 10^{-1} and 10^{+5} Hz, methods based on resonant systems vibrating at a natural frequency are used. The damping of the free vibration system is then given by $Q^{-1} = (1/n\pi)\ln(A_i/A_{i+n})$, where A_i and A_{i+n} are the ith and the $(i+n)$th amplitude of the free vibration respectively (**Figure 5**). This method is more suited to measure very low damping, but it is difficult to carry out measurements as a function of the frequency.

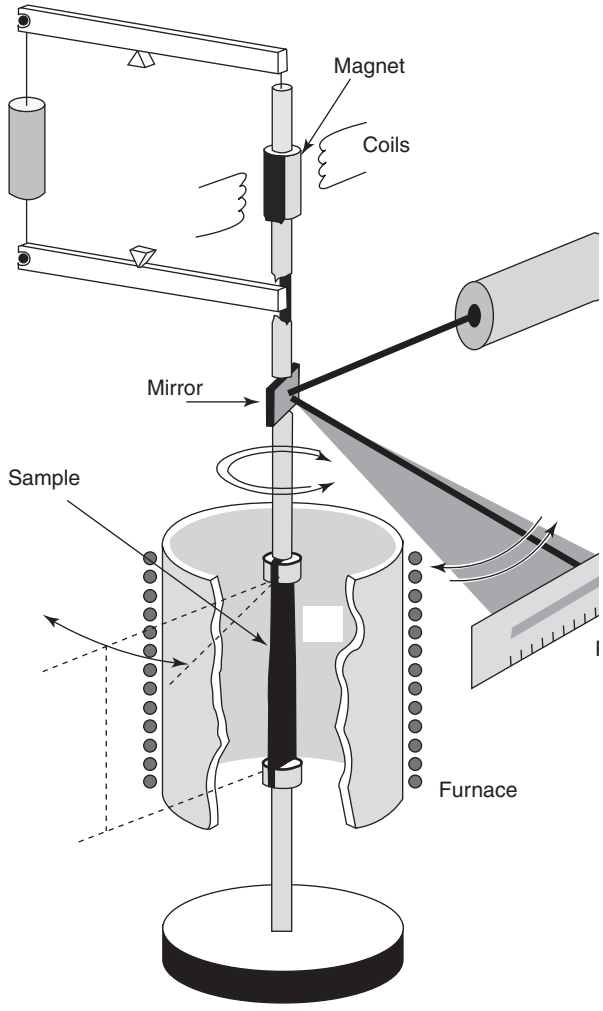


Figure 4 Inverted torsion pendulum: force vibration is imposed to the system and the phase lag δ is directly measured; frequency range: 10^{-5}–10 Hz; temperature range: 4–1800 K.

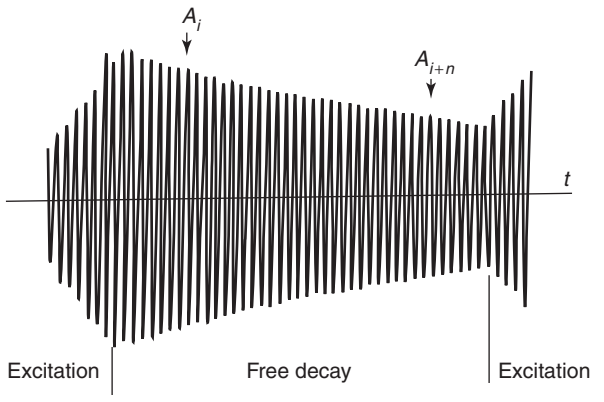


Figure 5 Free decay of natural vibrations, $Q^{-1} = (1/n\pi) \ln (A_i/A_{i+n})$.

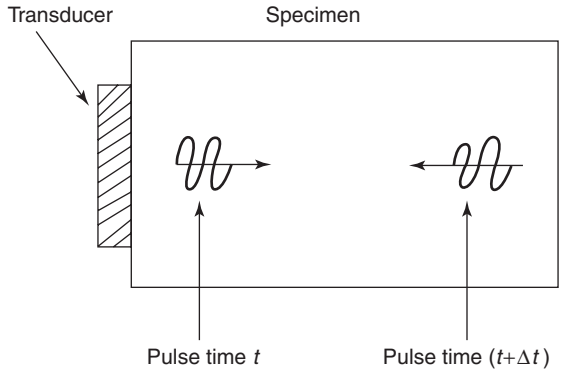


Figure 6 Ultrasonic attenuation measured by the echo technique.

In the MHz range, wave propagation methods are used (**Figure 6**). Bonding an appropriate transducer to one end of the sample generates a pulse or a wave packet, whose length is small compared to the specimen length. The velocity v and attenuation $\alpha(U = U_0 \exp - \alpha x)$ of such a pulse is determined by the echo technique: the time to return to the starting point after suffering a reflection of the free end, and the decrease in amplitude after successive reflections. The attenuation α is directly related to the IF by $Q^{-1} = \alpha\lambda/\pi$, where λ is the wavelength.

Microscopic Approach

The physical origin of the anelastic relaxation process is discussed in terms of atomic models. The cases of point defects, dislocations, and grain boundaries are considered successively. For these three cases, it is successively shown how to find the corresponding eqns [2]–[4] in order to obtain the anelastic behavior given by eqn [1].

Point Defect Relaxations

Point defects in the crystal lattice produce a local elastic distortion that is called an elastic dipole (by analogy with electric dipole), and as a result of this elastic dipole, there is an interaction between the defect and the stress applied to the crystal. It will be more easy to directly consider the particular case of the anelastic relaxation produced by interstitial solutes (e.g., C,N,O…) in dilute solution in body-centered cubic (b.c.c.) metals.

These interstitial impurities are responsible for an IF relaxation peak called Snoek relaxation. These impurities occupy the octahedral sites of the b.c.c. host lattice, as presented in **Figure 7**. So they create an elastic dipole with a tetragonal symmetry that has a single fourfold symmetry axes lying along the major axes of the octahedron. The interstitial sites may be subdivided into three groups, since the tetragonal axis may lie along each one of the cube axes

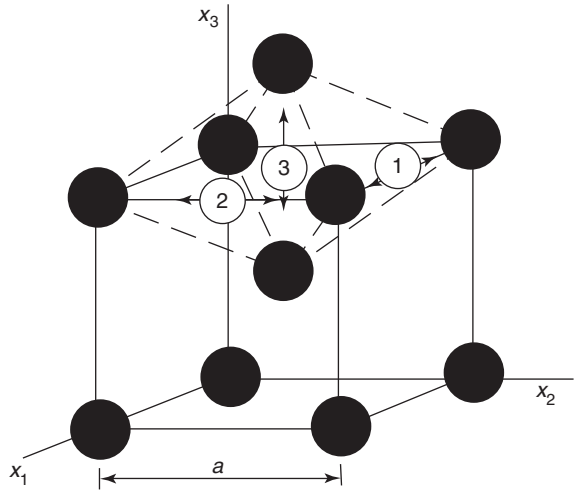


Figure 7 Octahedral sites in the b.c.c. lattice. Without external stress, all these sites are equivalent to each other.

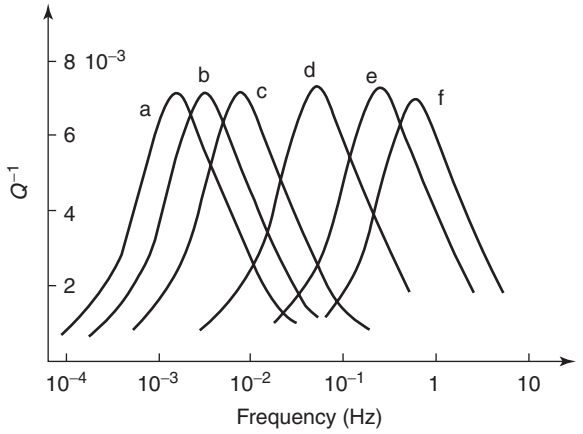


Figure 8 Snoek peak observed in Nb containing 0.15 at.% of oxygen.

x_1, x_2, x_3. It is then possible to mark the defects that have their tetragonal axis parallel to x_1, with the orientation index $p = 1$ and for x_2 and x_3, $p = 2$ and $p = 3$, respectively. If the total concentration of defects is N_0, then $N_0 = \sum_{p=1}^{3} N_p$, where N_p is the number of defects per volume unit with their tetragonal axis in the p direction. At the equilibrium, without applying an external stress all of the sites are equivalent ($N_1 = N_2 = N_3 = N_0/3$), and the extra deformation, due to the defects is the same in the three directions. When an external stress is applied in one direction, one type of defects is favored, consequently $N_1 \neq N_2 \neq N_3$. To simplify, consider only one uniaxial stress parallel to x_1 (i.e., $\sigma = \sigma_{11}$); the corresponding strain $\varepsilon = \varepsilon_{11}$ is given by

$$\varepsilon = J_U \sigma + \delta\lambda(N_1 - N_0/3)$$

where $\delta\lambda = (\partial\varepsilon/\partial N_1)_{N\text{const}}$, that is, the variation of the strain due to the migration of one defect from the position 2 or 3 to 1 and vice versa. The comparison with eqn [2] leads to $\kappa = \delta\lambda$ and $\xi = N_1 - N_0/3$. Moreover, thermodynamics of an irreversible process shows that

$$\dot{N}_1(T) = -\frac{1}{\tau}(N_1(t) - \bar{N}_1)$$

which is similar to eqn [4]. In order to obtain eqn [3], Boltzmann statistics are used and the equilibrium value $\bar{N}$ at each stress σ and temperature T is calculated. It is found that

$$\bar{\xi} = \bar{N}_1 - N_0/3 = \alpha\sigma \quad \text{with } \alpha = \frac{2N_0}{9kT}\delta\lambda$$

Finally $\delta J = \alpha\kappa = (2N_0/9kT)(\delta\lambda)^2$ and the relaxation strength is given by $\Delta = \delta J/J_U = 2N_0/(9J_U kT)(\delta\lambda)^2$. Consequently, the relaxation strength is proportional to the concentration of interstitial solutes, and the determination of Δ is an easy way to measure, with a very high accuracy, the concentration of carbon, nitrogen, or oxygen in solid solution of b.c.c. metals. Nowick and Berry also show that $\tau = a^2/36D$, where a is the side of the cube and D the diffusion coefficient of the impurities ($D = 1/2 v_D a^2 \exp(-E/kT)$, where v_D is the Debye frequency $\approx 10^{13}$ Hz). Therefore, $\tau = \tau_0 e^{E/kT}$, with E, the migration energy of the impurities and $\tau_0 \approx 10^{-14}$ s the pre-exponential factor. In conclusion, the IF and the modulus defect are given by the Debye eqn [5] with

$$\Delta = \frac{\delta J}{J_U} = \frac{2N_0}{9J_U kT}(\delta\lambda)^2$$
$$\tau = \tau_0 e^{E/kT} \approx 10^{14} e^{E/kT} \qquad [7]$$

Figure 8 presents an experimental result obtained in a force pendulum. The damping is plotted as a function of the frequency for different temperatures. The height of the peak is proportional to the concentration of the impurities and allows one to measure this concentration precisely. Plotting $\ln v_p$ (v_p being the frequency of the maximum) as a function of the temperature (**Figure 9**), the slope of the line gives E/k, that is, the migration energy of the impurity and the pre-exponential factor τ_0, which exactly corresponds to the predicted value of 10^{-14} s (**Table 1**).

When the Snoek peak is investigated for single crystals with different orientations, significant anisotropy of the relaxation strength is observed. From this type of measurements, it is possible to deduce the symmetry of the elastic dipole and to describe the position of the defect in the crystal lattice.

IF peaks observed in substitutional solid solutions are called Zener peaks and have been extensively studied for many different metallic structures (f.c.c.,

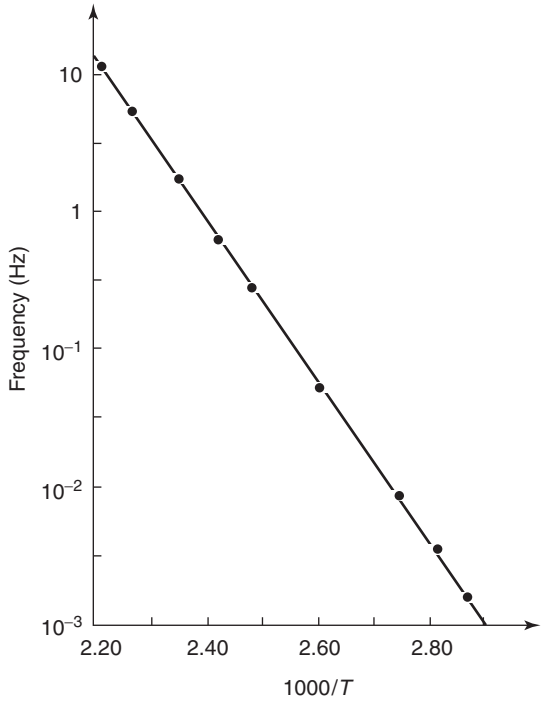


Figure 9 Arrhenius plot in case of oxygen Snoek peak (**Figure 8**).

Table 1 Snoek relaxation parameters in b.c.c. metals

System	E (eV)	τ_0^{-1} (10^{14} s^{-1})	T_P (K)
Fe–N	0.82	4.2	300
Fe–C	0.87	5.3	314
Nb–O	1.15	3.7	422
Nb–N	1.57	8.2	562S
Ta–O	1.10	1.17	420
Ta–N	1.66	2.7	615

T_p values at $f = 1$ Hz.
The activation energy corresponds to the migration energy of the defects.

h.c.p., and b.c.c.) as well as ceramics. However, in that case, the symmetry of the dipole due to one defect is similar to the symmetry of the crystal, and no stress-induced ordering mechanism is possible. Therefore, the key point in the theory is the proposal that the relaxation is due to a stress-induced reorientation of solute atom "pairs" present in the solid solution in a nearest-neighbor configuration. As a consequence, the relaxation strength Δ becomes proportional to the square of the concentration of defects, while the activation energy always corresponds to the migration energy of the substitutional atoms.

Dislocation Relaxation

It is well known that dislocation displacement gives rise to shear strain, and correspondingly, that the application of a shear stress produces dislocation motion. Consequently, a dislocation displacement meets the conditions for an internal variable that can result in an anelastic behavior. Nevertheless, this motion should be sufficiently restricted to be reversible. Such a restriction occurs when dislocations move in a potential is due to either internal stresses or obstacles (e.g., point defects) which impede a long-range dislocation motion. Equation [2] can easily be rewritten as

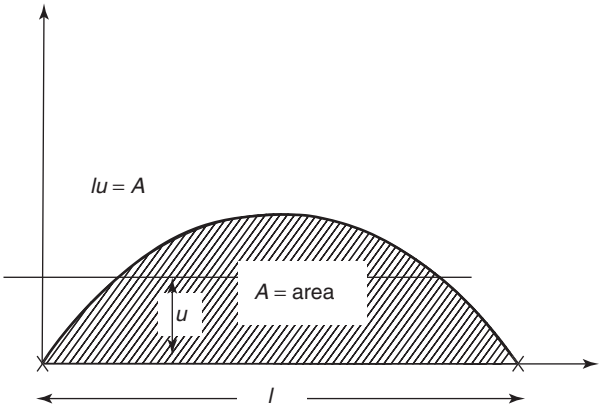


Figure 10 Dislocation segment of length l pinned at points bows into an arc. Definition of the internal variable u.

$$\varepsilon = J_U\sigma + \Lambda b u \qquad [8]$$

The second term, the anelastic strain, is due to the dislocation motion and is given by the Orowan equation where u, the internal variable, is the mean distance traveled by the dislocations (**Figure 10**), $\boldsymbol{b}$ the Burgers vector of the dislocations, and Λ the density of mobile dislocations.

To calculate the time dependence of the mean displacement u, it is necessary to write the motion equation of the dislocation. In case of small values of the external stress σ, it is assumed that dislocations move in a potential valley that gives rise to a restoring force Ku, so that the equation of movement becomes

$$m\ddot{u} = \sigma b - Ku - B\dot{u} \qquad [9]$$

where m is the inertial mass per unit length of dislocation, σb the magnitude of the force applied on the dislocation through a stress σ (Peach–Koehler force), and B a damping or drag term. The B term characterizes all the mechanisms that are able to control the motion of dislocations, for example, the interaction with phonons at high velocity (or high frequency), the motion through the Peierls valleys, the dragging of point defects, etc. Eliminating u between eqns [8] and [9] gives an equation of the form [1b], and consequently, the IF has the shape obtained in **Figure 3**. The experimental results (**Figure 11**)

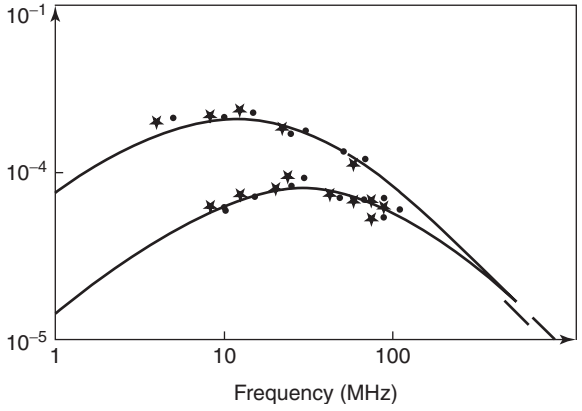


Figure 11 Dislocation relaxation measured in the MHz range for gold.

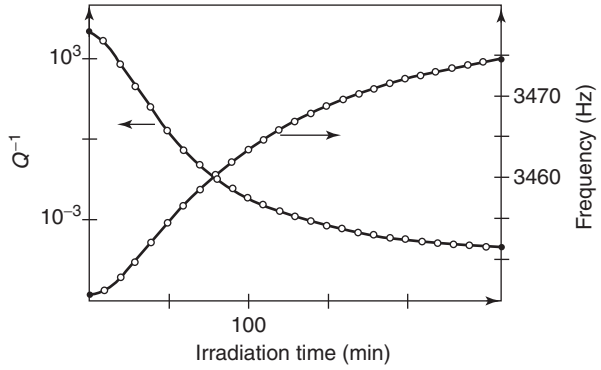


Figure 12 Effect of fast neutron irradiation on Young's modulus and internal friction of silver. Measurements in the KHz range.

prove that the dislocation motion is overdamped, that is, the inertial term $m\ddot{u}$ is negligible in eqn [9] and consequently,

$$B\dot{u} + Ku = b\sigma \qquad [10]$$

Eliminating u between eqns [8] and [10], one obtains

$$\left(\frac{\Lambda b^2}{K} + J_U\right)\sigma + \frac{B}{K}J_U\dot{\sigma} = \varepsilon + \frac{B}{K}\dot{\varepsilon}$$

an equation similar to that of the standard anelastic solid [1a], and consequently the IF as well as the modulus defects are always given by the Debye equation [5] with

$$\Delta = \frac{\Lambda b^2}{KJ_U} \quad \text{and} \quad \tau = \frac{B}{K} \qquad [11]$$

Moreover assuming that the restoring force Ku is due to the line tension γ of the dislocations with $\gamma \approx 0.5b^2/J_U$, it can be shown that, for a dislocation of length l vibrating like a string, $K = 12\gamma/l^2$ and consequently

$$\Delta = \frac{\Lambda l^2}{6} \quad \text{and} \quad \tau = \frac{Bl^2}{12\gamma} \qquad [12]$$

It has been shown that the relaxation, due to the interaction between phonon and dislocation, is responsible for the relaxation peak appearing in the MHz range (**Figure 11**). Moreover, in the low-frequency range of the Debye relaxation [5]

$$Q^{-1} = \Delta\omega\tau \propto \Lambda l^4 \quad \text{while} \quad \frac{\delta J}{J_U} = \Delta \propto \Lambda l^2 \qquad [13]$$

During neutron or electron irradiation, point defects (interstitial or vacancies), that can diffuse to the dislocation and pin them, are created. This pinning point modifies the dislocation length l, which decreases. Consequently, a strong decrease of the damping and of the modulus defect (Young modulus increases) that verifies the dependence predicted by eqn [13] (**Figure 12**) is observed. From the evolution of these two values, it is possible to follow the evolution of the dislocation network (l and Λ).

At much lower frequency, in the KHz and Hz range, other relaxation phenomenon due to the motion of dislocation are observed. For instance, mobile point defects are able to influence the damping term B. Dragging of mobile point defects exerts a retarding force, which leads to

$$\tau = \frac{B}{K} = \frac{s\frac{kT}{Gb^3}}{6\nu_D} \times \frac{l}{b}\exp\left(\frac{E_{PD}^M}{kT}\right)$$
$$= \tau_0 \exp\left(\frac{E_{PD}^M}{kT}\right)$$

where s is the number of point defects that migrate with a migration energy E_{PD}^M, and a Debye frequency ν_D. Then a relaxation peak, as a function of the temperature, is observed. Hasiguti peaks in f.c.c. metals (**Figure 13**) have been attributed to the dragging of intrinsic point defects induced either by cold-work or irradiation. The Snoek–Köster peaks, observed after cold-work of b.c.c. metals, are due to the dragging of interstitial impurities (H, C, O, N).

At low temperature, dislocations that lie in a close-packed direction of the lattice are more stable. Consequently, the potential energy of such a dislocation is a periodic function and the glide plane is similar to a corrugated iron. This periodic potential is called the Peierls potential. To minimize its energy, a dislocation, which lies in an arbitrary direction, develops a kinked structure (**Figure 14**), and the motion of dislocations is then controlled by the creation and motion of these kinks.

Several relaxation processes also characterize the kink dynamics. Considering the kink motion, the relaxation time becomes $\tau = \tau_0 \exp(H_K^M/kT)$, where

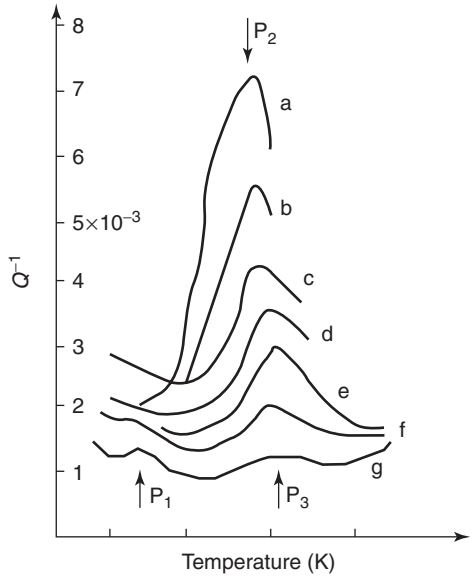


Figure 13 Hasiguti peaks observed in cold-worked gold. Curves a–g correspond to different annealing temperatures between 200 and 300 K.

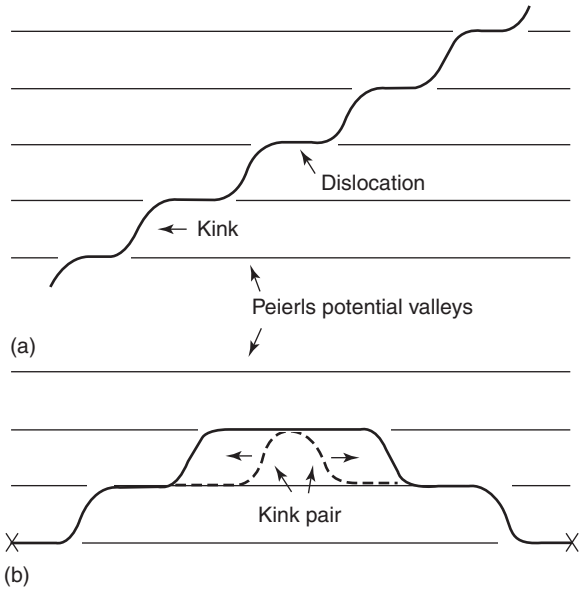


Figure 14 (a) Kinked structure of a dislocation and (b) the mechanism of kink pair formation.

H_K^M is the kink migration enthalpy. On the other hand, there exists a mechanism whereby the thermally activated kink pair formation controls the dislocation motion, which has been extensively studied. In this case, τ takes the form $\tau = \tau_0 \exp(H^{KPF}/kT)$, where H^{KPF} is the kink pair formation enthalpy. The kink dynamics has been extensively studied in f.c.c. (Bordoni relaxation), b.c.c. (α and γ relaxation), (**Figure 15**) and h.c.p. metals. From the measurement of the activation energy of these different relaxation peaks, it is possible to deduce a value of the Peierls energy or Peierls stress (**Table 2**).

The kink dynamics is now extensively studied in semiconductors, covalent crystals as well as in ceramics where it controls the mechanical properties of these materials. The experiments are difficult because it is necessary to create enough dislocations, that is, to extensively cold-work the specimen without breaking it!

Grain Boundary Relaxations in Metals

An important relaxation peak at 285°C at a frequency of 0.8 Hz in polycrystalline aluminum has been observed (**Figure 16**). This peak, called *Kê* peak, was not observed in single crystals and consequently, was attributed to grain-boundary sliding (GBS).

A simple model can be developed to explain the main characteristic of a relaxation peak due to the GBS, either in a polycrystal or in a bamboo structure. If a specimen with a bamboo structure is subjected to an applied stress σ, the boundary will shear plastically at constant rate $\dot{\varepsilon}$. The strain is then given by the sum of the elastic part and a plastic part due to the grain boundary sliding, that is

$$\varepsilon = J_U \sigma + \beta s u$$

where s is the grain-boundary density (cm^2/cm^3), u the shear displacement taking place along the grain boundary, and β a constant. The stress is given by

$$\sigma = ku + \frac{\eta}{\gamma\delta}\dot{u}$$

where the first term is due to the elastic deformation occurring in the two adjacent crystals, and the second term to the friction process due to the grain-boundary viscosity η. δ is the grain-boundary width. β and γ are geometrical parameters not too far from unity. Eliminating u between these two equations, one effectively obtains the equation of the standard anelastic solid with $\tau = \eta/\gamma\delta k$, the relaxation time, and $J_R = J_U + \beta s/k$, the relaxed compliance. Consequently, the damping and the relaxation strength are respectively given by the Debye eqns [5]

$$\text{with} \quad \Delta = \frac{\beta s}{k J_U} \quad \text{and} \quad \tau = \frac{\eta}{\gamma\delta k}$$

From the relaxation time, *Kê* deduced a value of the grain-boundary viscosity and this was considered as a very important experimental data deduced from IF measurements. In the case of a bamboo structure, k is a constant, the relaxation peak has a relaxation strength Δ that is proportional to the grain-boundary density s, and a relaxation time that is independent of s. In the case of a polycrystal, the counter stress ku is

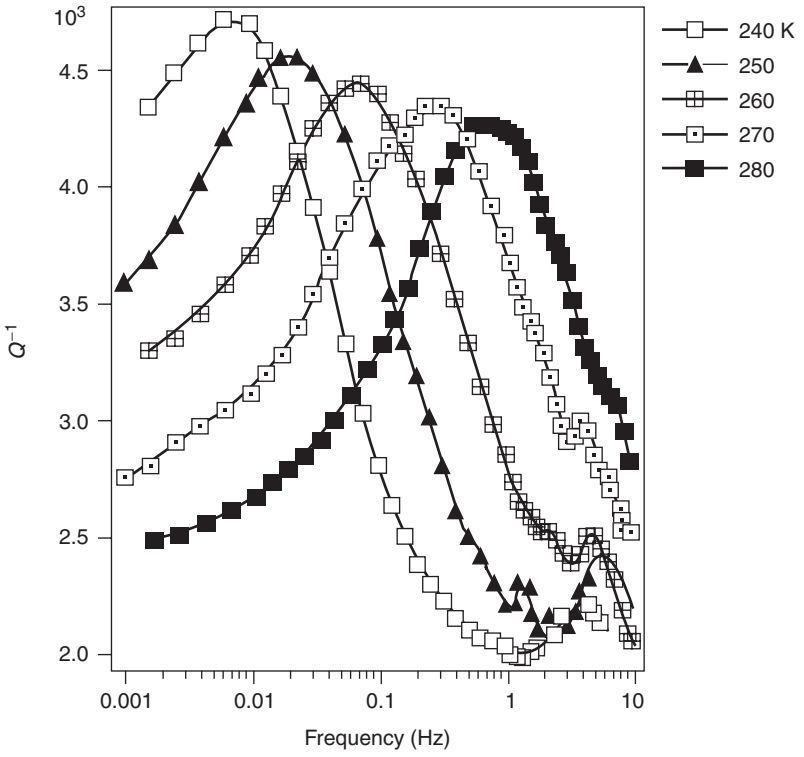


Figure 15 Relaxation peak observed in deformed Nb around room temperature. This dislocation peak (γ peak) is interpreted by a kink pair formation mechanism on screw dislocation.

Table 2 α, α', and γ are relaxation peaks observed in b.c.c. metals

Material	$T\alpha$ (K)	$E\alpha$ (eV)	$2H_K$ (edge) (eV)	$T\alpha'$ (K)	$E\alpha'$ (eV)	H_K^M (screw)	$T\gamma$	$E\gamma$ (eV)	$2H_K$ (screw)
Nb	40	0.075	0.08	40	0.07	0.07	250	0.59	0.56
Ta	31	0.057	0.062	22	0.034	0.034	350	0.71	0.74

From the relaxation parameters, the kink formation energy (H_K) and the kink migration energy (H_K^M) can be deduced for edge and screw dislocations.

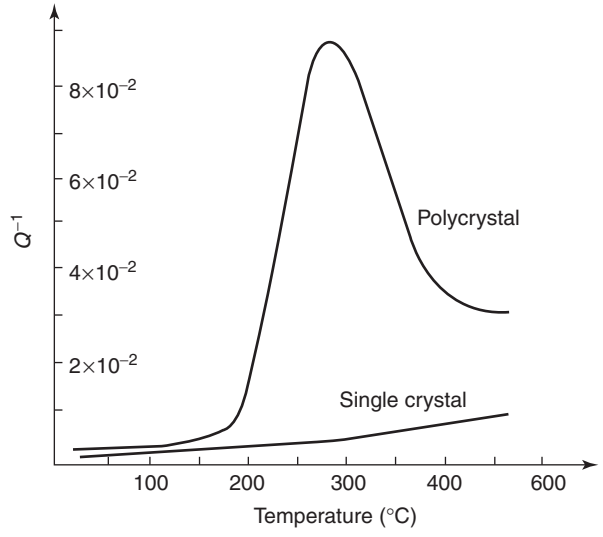


Figure 16 *Kê* peak in Al. *Kê* interpreted this peak by a grain-boundary sliding mechanism.

due to the elastic deformation of the grains between corners, and consequently, k is proportional to s and the relaxation strength does not depend on the grain size. Both effects have been observed. Assuming that the viscosity η is correlated with grain-boundary atomic self-diffusion,

$$\eta = \frac{kT}{Da} \quad \text{with}$$

$$D = D^{GB} = \frac{1}{2} va^2 = \frac{1}{2} v_D \exp\left(\frac{-E^{GB}}{kT}\right) a^2$$

where D^{GB} is the self-diffusion coefficient in the grain boundaries, a the atomic parameter, v_D the Debye frequency, and E^{GB} the self-diffusion activation energy in the boundary, that is, approximately 0.5–0.7 E^L (E^L = the self-diffusion energy in the lattice). Assuming that k and s are proportional, it is concluded that Δ should be independent of the grain size and than τ should be proportional to the grain size.

Table 3 Expected temperature of the grain boundary peak ($f = 1$ Hz) assuming $\tau_0 = 10^{-10}$ for Al, Cu, Ni, and Fe

Metal	E^{SD} (eV)	E^{GB} (eV)	T_p (K)
Al ($d = 0.02$ cm)	1.48	0.89	520
Al ($d = 0.08$ cm)			559
Cu	2.04	1.22	749
Ni	2.9	1.74	952
Fe	2.45	1.47	805

These peaks have been effectively observed at the expected temperature.

Effectively, the relaxation time τ is given by eqn [2]:

$$\tau = \frac{\eta}{\gamma\delta k} = \frac{2}{\gamma\beta}\cdot\frac{J_U kT}{va^3}\cdot\frac{\Delta}{s}$$

with $a^3/J_U = 4\,\text{eV}$, $kT = 4.31\times10^{-2}$ at $500\,\text{K}$, $\Delta \cong 0.5$ and $s \approx d^{-1}$, one obtains

$$\tau_0^{GB} = 10^{-15}\frac{d}{\delta}$$

Assuming a polycrystal of aluminum with grain size d in the range of 0.02–0.08 cm ($\frac{d}{\delta} = 0.4\times10^6 - 1.6\times10^6$), the limit relaxation time is approximately given by $\tau_0 \approx 4\times10^{-10}$–$1.6\times10^{-9}$. With this limit relaxation time and the values E^{SD} given in the literature for self-diffusion activation energy ($E^{GB} = 0.6E^{SD}$), one finds the expected temperatures of the grain-boundary peaks that correspond approximately to the measured values (**Table 3**).

Hysteretic and Nonlinear Effects

In addition to the relaxation phenomenon described above, which were attributed to anelastic effects, dislocation motion as well as the domain boundary motion in ferromagnetic or martensitic materials lead to internal friction with a hysteretic nature or vibration-amplitude-dependent damping, which strictly falls outside the scope of anelasticity. Often it is possible to consider a damping curve as composed of two parts, the first part being independent of amplitude, and the second part being dependent on amplitude. For example, it was earlier observed that the damping of high-purity annealed metals increases and becomes strongly strain amplitude dependent above strain amplitudes $\sim10^{-7}$. Such a nonlinear effect cannot be described as an anelastic mechanism. A mechanism for this effect involves breaking away from pinning points distributed along the length of dislocations. Nevertheless, depending on the temperature, dislocations can break away from pinning points (hysteretic effect at low temperature) or drag the pinning points (relaxation effect at high temperature); consequently, the separation between these different effects (anelastic, hysteretic, or strain amplitude dependent) is not evident.

In this article, it is shown that the measurement of internal friction is a very powerful method to study the dynamic behavior of point defects, dislocations, and grain boundaries in crystalline materials. Currently, it is intensively applied to study the rheologic behavior of glasses, polymers, as well as various new materials.

See also: Dislocations; Ferromagnetism; Lattice Dynamics: Structural Instability and Soft Modes; Mass Transport; Mechanical Properties: Creep; Mechanical Properties: Elastic Behavior; Point Defects.

PACS: 61.72; 61.80; 61.82; 62.20.−x; 62.30.+d; 62.40.+i; 62.80; 64.10; 81.40

Further Reading

Fantozzi G, Esnouf C, Benoit W, and Ritchie I (1982) *Progress in Materials Science*, vol. 27, p. 311. Pergamon.

Nowick AS and Berry BS (1972) *Anelastic Relaxation in Crystalline Solids*. New York and London: Academic Press.

Schaller R, Fantozzi G, and Gremaud G (eds.) (2001) *Mechanical Spectroscopy, Materials Science Forum*, vol. 360–368. Switzerland: Trans Tech Publications.

Zener C (1948) *Elasticity and Anelasticity of Metals*. Chicago, Illinois: University of Chicago Press.

Mechanical Properties: Creep

T G Langdon, University of Southern California, Los Angeles, CA, USA

Introduction

When a crystalline solid is subjected to a very high stress, it breaks catastrophically; but when it is subjected to a relatively low stress, which is insufficient to cause failure, it gradually deforms plastically over a period of time in the process known as "creep."

Creep occurs in all crystalline materials from metals to ceramics to intermetallics to rocks. It occurs in single crystals, polycrystalline materials, composites, and in all materials where the presence of an atomic lattice can deform plastically through the movement

of crystalline defects. It is an important process in a very wide range of situations – from the deformation occurring in gas turbine blades and nuclear reactors to the formation of mountains and the flow within the mantle of the Earth and other planetary bodies.

Creep is a diffusion-controlled process so that the creep rate is dependent upon the rate of diffusion. This means in practice that, although creep occurs at all temperatures above absolute zero, creep generally becomes important, and the creep rates become reasonably rapid, only in situations where the temperature is at or above approximately one-half of the absolute melting temperature of the material. Although this requirement may appear somewhat limiting, nevertheless creep occurs in many simple situations. For example, it occurs in the creep of ice flows and glaciers or in the creep of parts made from metallic lead where the low melting temperature of 600 K means that flow occurs easily even at ambient temperatures. Examples of the creep of lead can be seen by examining the distortions of old lead tiles or piping on the medieval cathedrals of Europe. Thus, creep is an important flow process that has an exceptionally wide range of applicability.

The Variation of Strain Rate with Time

A typical creep curve for a crystalline solid is shown schematically in **Figure 1** in a plot of the strain, ε, against the time, t. It is apparent that there is an initial instantaneous strain, ε_o, which occurs immediately upon application of the load and this is followed by three distinct regions of creep: the primary region (I), the secondary or steady-state region (II) and the tertiary region leading to ultimate fracture (III). In broad terms, the primary stage is a region of hardening in which a substructure develops in the material whereas the secondary region represents a period in which there is a balance between the processes of hardening and recovery so that the creep rate remains constant. The tertiary stage is a region in which the strain rate accelerates due to the accumulation of creep damage within the material: examples of creep damage are the occurrence of necking in tensile testing or the development of internal cracks or cavities. In practical situations, it is imperative to avoid the occurrence of region III since it inevitably leads to a failure of the material.

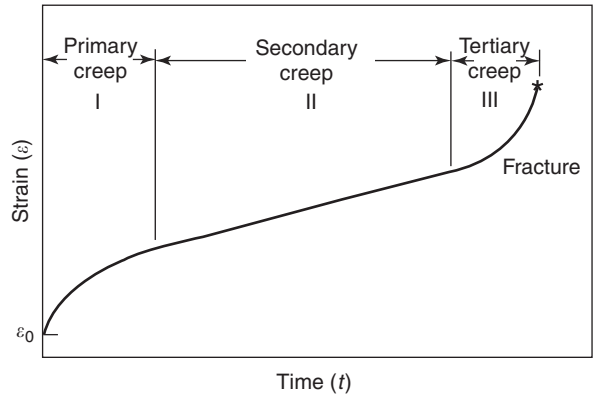


Figure 1 Typical creep curve of strain vs. time.

The steady-state or secondary region is often dominant in the creep behavior of materials in practical situations, and it is important, therefore, to develop an understanding at the atomistic level of the flow processes occurring in this region. The results from laboratory experiments generally show that the steady-state creep rate, $\dot{\varepsilon}$, varies with the applied stress, σ, through a power-law relationship of the form

$$\dot{\varepsilon} = A_1 \sigma^n \qquad [1]$$

where A_1 is a constant and n is termed the stress exponent for creep. However, when creep tests are conducted at exceptionally high applied stresses, experiments show the creep rate varies exponentially with the applied stress through a relationship of the form

$$\dot{\varepsilon} = A_2 \exp(B\sigma) \qquad [2]$$

where A_2 and B are appropriate constants. The point of transition, from power-law creep and eqn [1] at lower stresses to an exponential form of creep and eqn [2] at high stresses, is generally termed power-law breakdown.

Equation [1] is relatively simple because it relates the steady-state strain rate only to the magnitude of the applied stress. In practice, however, the measured creep rates depend also upon the testing conditions, such as the temperature, and upon microstructural conditions, such as the grain size of the material. It is usually possible to express eqn [1] in a more rigorous form by expressing the steady-state creep rate through a relationship of the form

$$\dot{\varepsilon} = \frac{ADGb}{kT}\left(\frac{b}{d}\right)^p \left(\frac{\sigma}{G}\right)^n \qquad [3]$$

where A is now a dimensionless constant, D is the diffusion coefficient, G is the shear modulus of the material, b is the Burgers vector, k is Boltzmann's constant, T is the absolute temperature, d is the grain size, and p is the exponent of the inverse grain size. The diffusion coefficient can be expressed explicitly in the form

$$D = D_0 \exp\left(-\frac{Q}{RT}\right) \qquad [4]$$

where D_0 is a frequency factor, Q is the activation energy for the rate-controlling diffusive process, and

R is the gas constant. Thus, it follows from an inspection of eqn [3] that, for any selected material, the steady-state creep rate is defined exclusively by the values of only four terms: A, Q, p, and n. This conclusion follows since all of the other terms in eqn [3] are either constants (such as b and k) or they depend on the precise testing conditions (such as G and T).

Equation [3] provides a general representation of the creep process, and all specific creep mechanisms have relationships of this form. Thus, an important objective in creep investigations is to measure the values of n, p, and Q for a material under any selected testing conditions, and then to interpret these results in terms of theoretical creep mechanisms that have been developed at the atomistic level.

The strains occurring in high-temperature creep can be broadly identified in terms of three different types of behavior. First, and probably dominant in most practical situations, plastic flow occurs within the grains of polycrystalline materials through processes, such as the intragranular glide and climb of dislocations. All of these intragranular processes may give a total strain which can be designated as ε_g. A strain may also develop through the processes of diffusion creep where strain accumulates through the flow of vacancies moving in response to the applied stress. The strain due to the flow of vacancies in diffusion creep is designated ε_{dc}. Finally, there may be an additional strain due to the occurrence of grain boundary sliding in polycrystalline materials, where the grains slide or move over each other in response to the applied stress with the movement taking place at, or very close to, the grain boundaries. The strain from this process is ε_{gbs}. Thus, the total strain accumulated in the flow process, ε_t, is given by the summation of these three separate strains through a relationship of the form

$$\varepsilon = \varepsilon_g + \varepsilon_{gbs} + \varepsilon_{dc} \qquad [5]$$

In practice, care must be taken in differentiating between ε_{gbs} and ε_{dc}, since diffusion creep requires accommodation through the movement of adjacent grains in the process termed Lifshitz sliding and grain boundary sliding, generally termed Rachinger sliding, which requires accommodation in the form of some limited intragranular deformation in order to maintain coherency at the grain boundaries.

The Development of Substructure in Creep

The primary stage of creep shown in **Figure 1** is related to the development of a substructure. For example, when creep occurs through the intragranular climb of dislocations, a substructure is formed consisting of an array of subgrain boundaries where the angles of misorientation are low (typically $\sim 2^\circ$) and arrays of dislocations within the subgrains. If creep occurs through the flow of vacancies, as in diffusion creep, no substructure develops and there is an instantaneous region of steady-state creep immediately following the application of the load.

In intragranular dislocation creep, the two important and measurable substructural quantities are the average subgrain size, λ, and the average density of dislocations within the subgrains, ρ. It is now well established, from a very large number of experimental measurements, that the values of λ and ρ are both dependent upon the magnitude of the applied stress.

If measurements are taken of the subgrain size, it is found that the value of λ varies inversely with the applied stress through a relationship of the form

$$\frac{\lambda}{b} = \zeta\left(\frac{\sigma}{G}\right)^{-1} \qquad [6]$$

where ζ is a constant having a value close to ~ 20. Similarly, if measurements are taken to determine the dislocation densities within the subgrains, where the dislocations contained within the subgrain walls are ignored, it is found that ρ varies with stress raised to a power of 2 through a relationship of the form

$$b\rho^{1/2} = \varphi\left(\frac{\sigma}{G}\right) \qquad [7]$$

where φ has a value close to ~ 1.

It is important to note that the relationships given by eqns [6] and [7] apply to both metals and ceramics, and it is reasonable to anticipate that they apply to all crystalline materials. The general trends of these two relationships are depicted schematically in **Figures 2** and **3**, where the shaded regions incorporate a large number of individual measurements taken on either metals or ceramics.

Creep in Pure Metals and Solid-Solution Alloys

There is a good understanding of the creep of pure polycrystalline metals. Two mechanisms tend to be dominant with the rate-controlling process dependent upon the level of the applied stress. At very low stresses, diffusion creep becomes rate-controlling and this can occur either through the diffusion of vacancies through the crystalline lattice in Nabarro–Herring diffusion creep or through the diffusion of vacancies along the grain boundaries in Coble diffusion creep. Both of these mechanisms require a stress exponent of $n = 1$ in eqn [3], but the processes are different because Nabarro–Herring creep requires

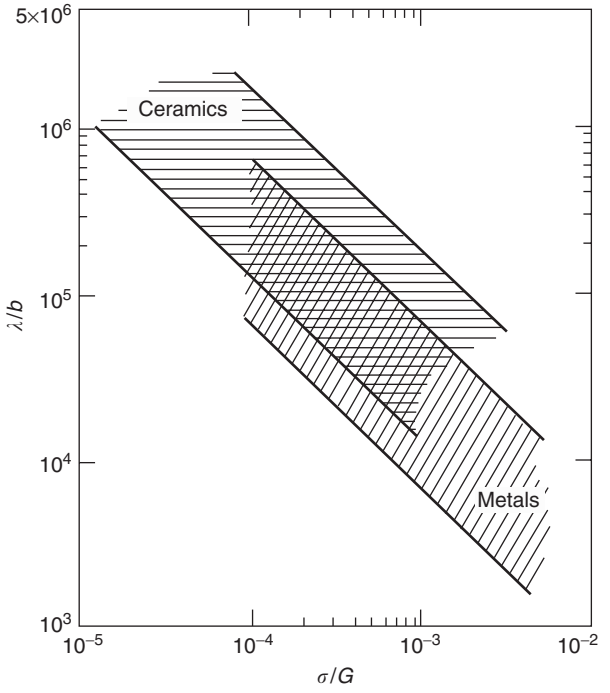


Figure 2 Representation of eqn [6] showing similar subgrain sizes in metals and ceramics.

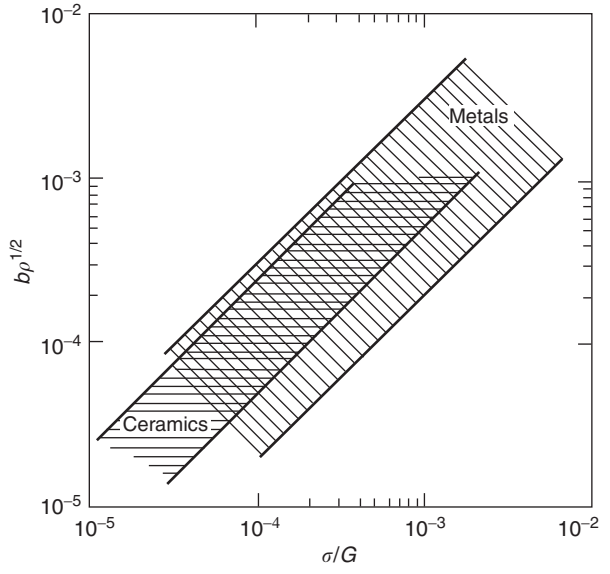


Figure 3 Representation of eqn [7] showing similar dislocation densities in metals and ceramics.

$p = 2$ and $D = D_l$, where D_l is the diffusion coefficient for lattice self-diffusion, and Coble creep requires $p = 3$ and $D = D_{gb}$, where D_{gb} is the diffusion coefficient for grain boundary diffusion. It is apparent, however, that both of these forms of diffusion creep give a stress exponent of 1, and thus they tend to become dominant at very low stresses. There is also an additional process, termed Harper–Dorn creep, which appears to be a dislocation mechanism occurring at low stresses with $n = 1$. The region at low stresses with $n = 1$ is generally designated the region of Newtonian viscous flow. Over a very wide range of intermediate stresses, creep occurs by the intragranular movement of dislocations, the climb process is rate-controlling, and the stress exponent is close to $n \approx 5$ although the precise value of n depends upon the stacking-fault energy of the metal. For pure aluminum, for example, the value of n is ~ 4.4. In experiments on high-temperature creep of pure metals, the region with $n \approx 5$ extends over a wide range of stresses with transitions to $n = 1$ at low stresses in diffusion creep, and an exponential dependence on stress at very high stresses in the region of power-law breakdown given by eqn [2]. For simplicity, the region with $n \approx 5$ where dislocation climb is the rate-controlling process, is generally designated class M or pure metal behavior.

The situation for metallic solid-solution alloys is often more complex. In solid-solution alloys, diffusion creep occurs as in pure metals but there may be a significant difference in the region of intragranular dislocation creep at intermediate stresses. Intragranular creep occurs through the movement of dislocations within the grains, where these dislocations move by both the glide and climb processes. In pure metals, the glide process occurs very quickly so that climb is always the rate-controlling step; but in solid-solution alloys, the glide process may be relatively slow because solute atom atmospheres may form around the moving dislocations and the glide process then involves the dragging of these solute atmospheres. Thus, in the solid-solution alloys there may be experimental conditions where the rate of glide, $\dot{\varepsilon}_g$, is slower than the rate of climb, $\dot{\varepsilon}_c$. Since dislocations move through both glide and climb, the formation and dragging of solute atmospheres provide a potential for introducing different transitions into the creep behavior.

An example is illustrated schematically in **Figure 4** where the creep strain rate, $\dot{\varepsilon}$, is plotted logarithmically against the applied stress, σ. At very low stresses, creep occurs by diffusion creep, and $n = 1$ in either Nabarro–Herring or Coble creep. As the stress level is increased, there is an initial region where $n \approx 5$ as in pure metals but at a well-defined stress given by σ_1 there is a transition to $n = 3$. At an even higher stress defined as σ_2, there is an increase in the creep rate and thereafter, at σ_3, the creep rate again reverts to the line for $n \approx 5$. Finally, at even higher stresses, there is evidence for an increase in the creep rate in the region of power-law breakdown. All of the regions depicted in **Figure 4** are well documented

experimentally and basically this result represents the standard behavior of pure metals (class M) but with an additional process superimposed between σ_1 and σ_3 which applies only to solid-solution alloys: this process is termed class A behavior since it occurs only in alloys.

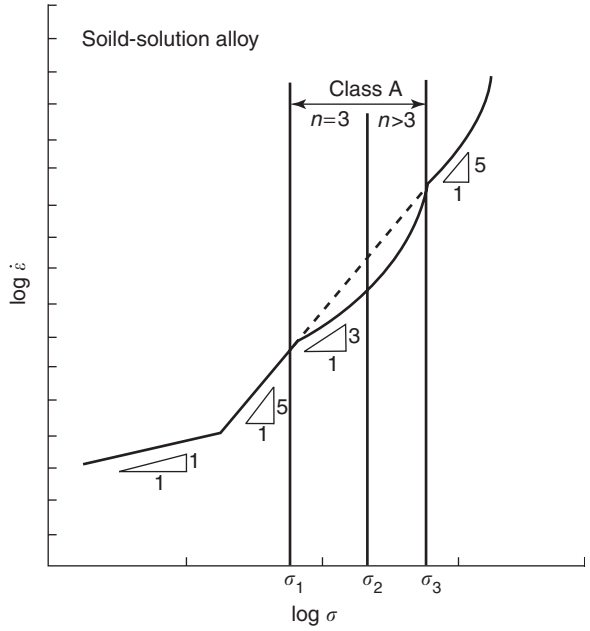


Figure 4 Strain rate vs. stress for a typical solid-solution alloy showing the different transitions in creep behavior.

The region with $n = 3$ in **Figure 4** represents a transition from climb-controlled behavior with $n \approx 5$ at the lower stresses to glide-controlled behavior and the dragging of solute atmospheres with $n = 3$ at the higher stresses. The transition from $n \approx 5$ to $n = 3$ is well represented by a simple relationship, which denotes the stress delineating the transition to a glide-controlled regime: this stress is given, in normalized form, by the following expression:

$$\frac{\sigma}{G} > \Psi\left(\frac{kT}{eGb^3}\right)\left(\frac{1}{c}\right)^{1/2}\left(\frac{\tilde{D}}{D_l}\right)^{1/2}\left(\frac{Gb}{\Gamma}\right)^{3/2} \qquad [8]$$

where e is the solute–solvent size difference, c is the concentration of the solute, $\tilde{D}$ is the coefficient for interdiffusion of the solute atoms in the matrix, D_l is the coefficient for lattice self-diffusion, Γ is the stacking-fault energy, and Ψ is a constant having a value which has been estimated as $\sim 3 \times 10^{-7}$.

The validity of eqn [8] is illustrated in **Figure 5** where the term on the right-hand side of eqn [8] is plotted against the normalized stress, σ/G, and the line at 45° represents the theoretical transition from climb-control or class M with $n \approx 5$ on the left to glide-control or class A with $n = 3$ on the right. The two solid points on the horizontal lines for the Al–3.3%Mg and Al–5.6%Mg alloys show the transition points derived experimentally and used to determine the value of $\Psi \approx 3 \times 10^{-7}$. The remaining alloys shown in **Figure 5** are plotted according to the ranges

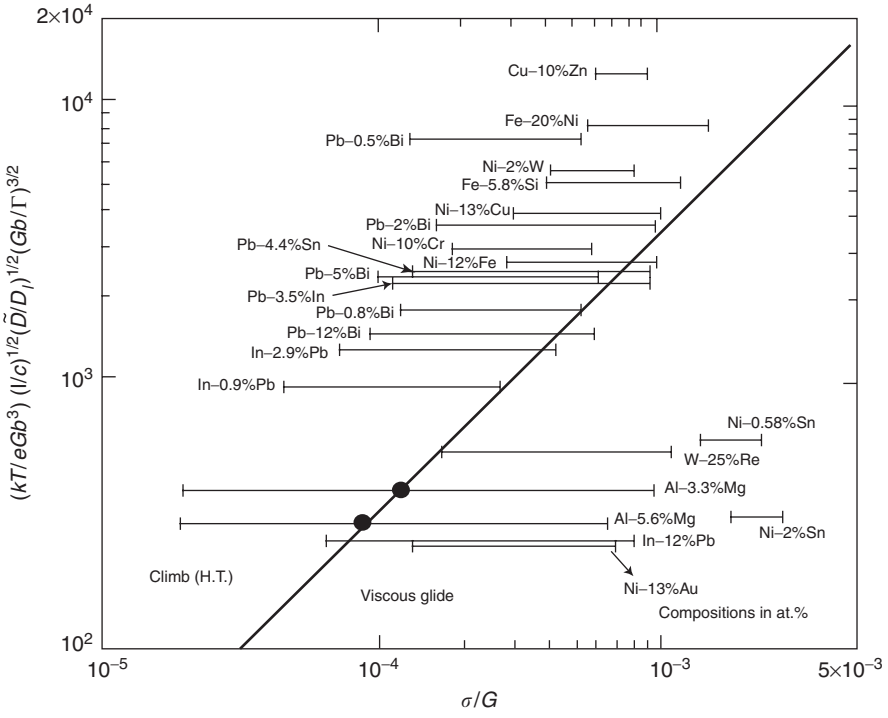


Figure 5 Representation of eqn [8] showing the transition from dislocation climb (on left) to glide (on right).

of normalized stress used experimentally, and all of these alloys lie in the correct positions relative to the published experimental data.

A second transition occurs in **Figure 4** at the point designated σ_2 where the stress exponent increases above $n = 3$ but the behavior remains within the region associated with class A. This transition occurs because at high-stress levels, the dislocations break away from their solute atmospheres so that $n > 3$. The theoretical breakaway stress, σ_b, is given by

$$\sigma_b = \frac{\alpha W_m^2 c}{b^3 kT} \qquad [9]$$

where W_m is the binding energy between the solute atom and the dislocation, and α is a dimensionless constant having a value of ~ 0.2. Thus, it follows from eqn [9] that glide is rate-controlling and $n = 3$ at stresses below σ_b, but the solute atmospheres no longer impede the dislocation mobility so that $n > 3$ at applied stresses above σ_b. Using eqn [9], the transition marking the breakaway is given in normalized form by the expression

$$\frac{\sigma}{G} > \frac{W_m^2 c}{5Gb^3 kT} \qquad [10]$$

where eqn [10] delineates the upper limit of the $n = 3$ region. Equation [10] is plotted schematically in **Figure 6** where the line at 45° represents the breakaway condition and the horizontal lines delineate experimental results plotted over the appropriate ranges of normalized stress. Each horizontal line in **Figure 6** shows a solid point denoting the approximate experimental transition from a well-defined region with $n = 3$ to a region where there appears to be a breakdown and $n > 3$. It is apparent from inspection of **Figure 6** that, despite some minor scatter, these solid points tend to be in excellent agreement with the predictions of eqn [10].

It is important to note that eqn [10] predicts a breakaway condition that depends upon the solute concentration since the normalized stress is a linear function of c. The effect of a change in the value of c is illustrated schematically in **Figure 7**, where the data are again presented in the logarithmic form of $\dot{\varepsilon}$ versus σ. This plot shows the transitions, with increasing stress, from $n = 1$ at low stresses for diffusion creep or Harper–Dorn creep to $n \approx 5$ in a climb-controlled region, and then to $n = 3$ where glide is rate-controlling, and to $n > 3$ when the dislocations break away from their atmospheres. Ultimately, the behavior reverts to $n \approx 5$ at even higher stresses when the concentration c_1 is small, but there is no additional region with $n \approx 5$ at high stresses when the concentration is high as in the alloy with a concentration of

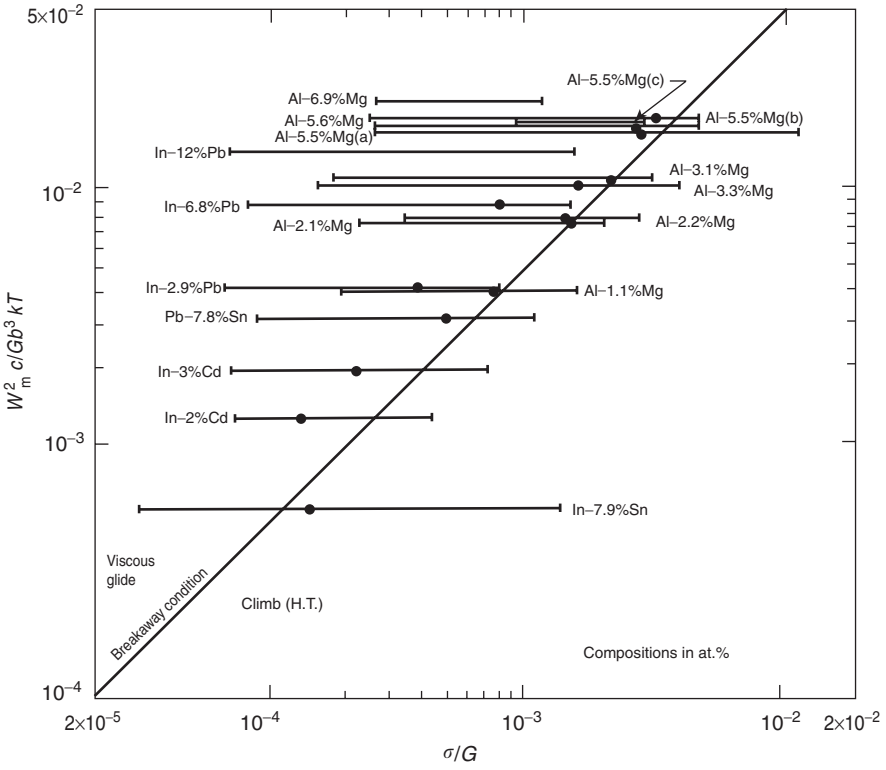


Figure 6 Representation of eqn [10] showing the breakaway of dislocations from the solute atom atmospheres.

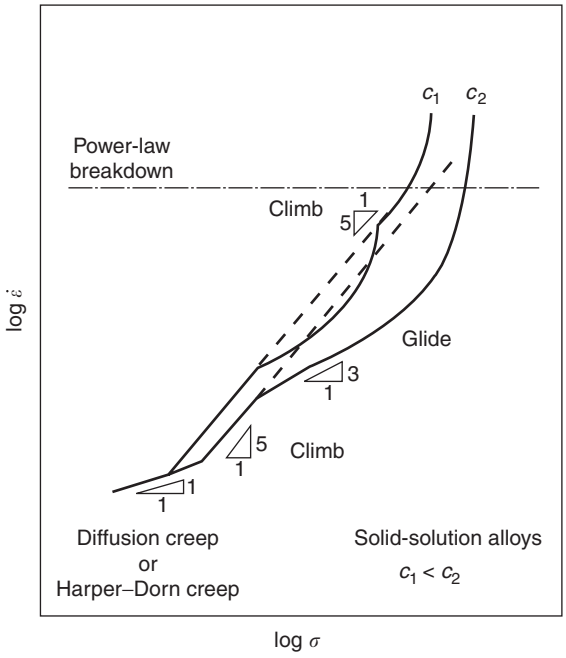


Figure 7 Schematic illustration of the effect of the solute concentration, c, on the transitions in behavior in solid-solution alloys.

c_2. At the very highest stresses, there is a conventional power-law breakdown and a transition to eqn [2].

See also: Mass Transport; Mechanical Properties: Fatigue; Mechanical Properties: Tensile Properties.

PACS: 62.20 Hg

Further Reading

Bird JE, Mukherjee AK, and Dorn JE (1969) Correlations between high-temperature creep behavior and structure. In: Brandon DG and Rosen A (eds.) *Quantitative Relation between Properties and Microstructure*, pp. 255–341. Jerusalem: Israel Universities Press.

Čadek J (1988) *Creep in Metallic Materials*. Amsterdam: Elsevier.

Cannon WR and Langdon TG (1988) Creep of ceramics: an examination of flow mechanisms. *Journal of Materials Science* 23: 1–20.

Kassner ME and Pérez-Prado MT (2000) Five-power-law creep in single phase metals and alloys. *Progress in Materials Science* 45: 1–102.

Langdon TG (1996) Grain boundary sliding revisited. In: Hondros ED and McLean M (eds.) *Structural Materials: Engineering Application through Scientific Insight*. London: The Institute of Materials.

Langdon TG (1996) Transitions in creep behaviour. *Materials Transactions JIM* 37: 359–362.

Mishra RS, Mukherjee AK, and Murty KL (eds.) (1999) *Creep Behavior of Advanced Materials for the 21st Century*. Warrendale, PA: The Minerals, Metals and Materials Society.

Mohamed FA and Langdon TG (1974) The transition from dislocation climb to viscous glide in the creep of solid solution alloys. *Acta Metallurgica* 22: 779–788.

Parker JD (ed.) (2001) *Creep and Fracture of Engineering Materials and Structures: Proceedings of the 9th International Conference*. London: The Institute of Materials.

Poirier J-P (1985) *Creep of Crystals*. Cambridge: Cambridge University Press.

Wilshire B and Evans RW (1985) *Creep Behaviour of Crystalline Solids*. Swansea: Pineridge Press.

Yavari P and Langdon TG (1982) An examination of the breakdown in creep by viscous glide in solid solution alloys at high stress levels. *Acta Metallurgica* 30: 2181–2196.

Mechanical Properties: Elastic Behavior

K Bowman, Purdue University, West Lafayette, IN, USA

Introduction

When we say that something is elastic we usually think that it is something that springs back to its original shape after we stretch, bend, twist, or compress it. The elastic properties of materials are those that govern the change in shape and the tendency of the material to return the bonds between atoms or molecules back to their original condition. Of course, we can exceed the level at which the material can completely recover its original shape resulting in permanent deformation (plastic deformation) or fracture. Even if plastic deformation or fracture intercedes, the shape change that the material which is elastic has undergone will reverse, returning the material back to its unstretched state.

The model for the elastic response of materials most commonly used is that of a spring. The response of a spring is well known for a linear relationship between force and displacement even to relatively large extensions or strains. When we apply this model to solid materials the mechanical response is approximately linear to very small strain levels. Although this nearly linear response may have been known for quite a long period in man's history, it was not until the seventeenth century that Robert Hooke defined the essential relation showing the proportionality between a tensile stress and the resulting tensile strain. That the changes in shape in all dimensions held similarly near linear responses became understood subsequently through the work of Thomas Young, S D Poisson, and others. **Figure 1** shows the linear relationships between normal stress,

σ (force divided by the cross-sectional area perpendicular to it) and normal strain, ε (the displacement normalized to the length over which it is measured).

By the twentieth century, it became apparent that the simple relations developed to describe elastic properties could vary with the direction of loading in single crystals, oriented polycrystals, and oriented molecules. This directionality is often called anisotropic elasticity. If a material does not possess this directional variation in properties, it is called isotropic. With isotropic elasticity, only two of the material properties called elastic constants are necessary to fully describe the elastic response of a material. More recently, researchers have increasingly probed the larger strain response of materials that can be decidedly nonlinear. Many biomaterials demonstrate this extended range of elasticity that is often called hyperelasticity.

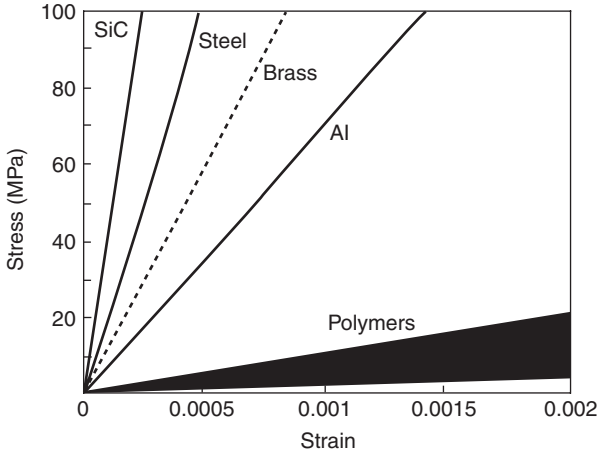


Figure 1 The stress vs. strain behavior for a number of materials measured from tensile loading experiments on commercial materials. The elastic constant called Young's modulus is the slope of each of these curves. (Adapted from Bowman KJ (2004) *An Introduction to Mechanical Behavior of Materials*. Wiley; © Wiley. This material is used by permission of John Wiley and Sons, Inc.)

Defining Elastic Constant

Isotropic Materials

Elastic constants are the constants describing mechanical response of a material when it is elastic. The isotropic elastic constants for some common materials are given in **Table 1**. If the engineering approximation that this mechanical response is linear is made, one can define a set of constants that relate any applied stress to the corresponding strain. When the material can also safely be treated as isotropic, the number of independent elastic constants required to completely describe the elastic response of a material is two. Any further elastic constants can then be defined in terms of the other two.

For a simple tension or compression test, the easiest elastic constant to define is Young's modulus, E. Young's modulus is the elastic constant defined as the proportionality constant between stress and strain:

$$E = \frac{\sigma}{\varepsilon} \qquad [1]$$

Young's modulus has the same units as stress, which is normally expressed in terms of pascals (Pa), equivalent to newtons per square meter, or pounds per square inch (psi). Strain is, of course, a dimensionless proportionality between a change in length and an original length. The Young's modulus is the slope of the linear elastic response for a number of materials as shown in **Figure 1**.

Another elastic constant that can be obtained in a tension or compression test wherein the strain along the loading direction and orthogonal to it is called Poisson's ratio, ν. Poisson's ratio is the ratio between

Table 1 Approximate elastic properties of various materials (and dilute alloys) at room temperature

Material	*Young's modulus, E* (GPa)	*Shear modulus,* μ (GPa)	*Poisson's ratio,* ν
Aluminum and aluminum alloys	69–72	24–26	0.35
Copper and copper alloys	125–135	47–50	0.34
Irons and steels	205–215	80–84	0.29
Stainless steels	190–200	75–78	0.33
Titanium and titanium alloys	115	42–44	0.32
Aluminum oxide	380–390	155–165	0.25
Silicon carbide	440–460	195–200	0.14
Glass	70–90	28–32	0.27
Polyethylene (PE)	0.2–2	[a]	0.4
Polymethylmethacrylate (PMMA)	2–3	[a]	0.4
Polystyrene (PS)	2–4	[a]	0.35
Bone[b]	5–30	3–8	0.25–0.5
Tendon[b]	0.8–1.5		

[a] These values are typically about one-third of Young's modulus.

[b] Bone and tendon are not only strongly anisotropic, but the stiffness also varies with type, position, and moisture content.

Adapted from Bowman KJ (2004) *An Introduction to Mechanical Behavior of Materials*. Wiley; © Wiley. This material is used by permission of John Wiley and Sons, Inc.

width or diameter strain, ε_1, of a tension or compression specimen and the strain along the tension or compression axis, ε_3. This dimensionless proportionality constant can be written as

$$\nu = \frac{-\varepsilon_1}{\varepsilon_3} \quad [2]$$

So far, the possibility of stresses occurring in multiple directions have not been considered. To do so, one needs to extend the idea of assigning a coordinate system to the stress state as shown in **Figure 2a**. If such a multiaxial loading condition is considered, the relations between a set of normal stresses and the corresponding strains as given by Hooke's law can be written as

$$\begin{aligned} \varepsilon_1 &= \frac{\sigma_1 - \nu(\sigma_2 + \sigma_3)}{E} \\ \varepsilon_2 &= \frac{\sigma_2 - \nu(\sigma_1 + \sigma_3)}{E} \\ \varepsilon_3 &= \frac{\sigma_3 - \nu(\sigma_2 + \sigma_1)}{E} \end{aligned} \quad [3]$$

This relation enables a set of normal stresses to be translated into corresponding strains using just the Young's modulus and Poisson's ratio.

A type of stress called a shear stress, τ, can also be defined which can produce a shear strain, γ. For shear stresses, the force is applied across the area and it results in a shearing of the material. Just as for normal stresses, shear loading can result in a mechanical response that is elastic and nearly linear, but the shape change is called a shear strain as shown in **Figure 2b**. The specific case in **Figure 2b** shows τ_4, which consists of a shear displacement in the y-direction on the z-face and in the negative y-direction on the negative z-face. To have a stable distortion, a similar pair of forces would need to be applied on the y-faces to cancel the rotational moment imposed by τ_4. The elastic constant that describes the linear relation between τ and γ is called the shear modulus, μ.

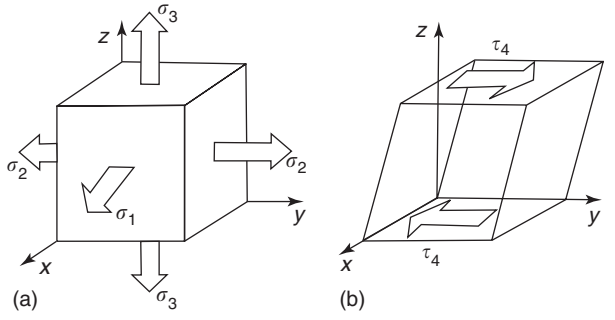


Figure 2 (a) Normal stresses referenced to an orthogonal coordinate system and (b) shear stresses and resulting distortion called shear strain.

So, one can write

$$\mu = \frac{\tau}{\gamma} \quad [4]$$

For isotropic materials, the relation between these three elastic constants is

$$E = 2\mu(1 + \nu) \quad [5]$$

Values for all three elastic constants are given for a number of materials at room temperature in **Table 1**. The values are given as ranges to represent variations expected from alloying.

Because elastic constants are related to bonding, they generally scale with melting temperature and the type of bonding. The strength and the type of bond determines the relationship between an applied force and changes in the distance between atoms. The spacing between atoms or molecules in a material that is not under stress comes about from a balance between attractive and repulsive forces that prevent the atoms or molecules from getting too close, and attractive forces that bring the atoms together. The sum of these two opposite forces is called the interatomic potential. Application of a stress to the material changes the balance point. Because the relationship between force and displacement is nearly linear, the elastic properties of materials can be described using linear relationships. As materials increase in temperature, they undergo increased thermal vibration that leads to thermal expansion. This increase in lattice constants also leads to a reduction in elastic constants with increasing temperature as indicated by **Figure 3**.

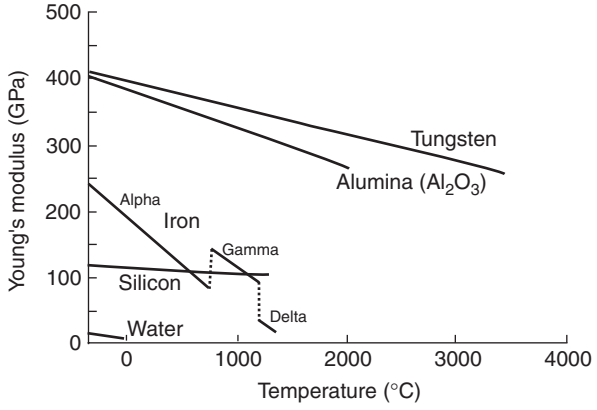


Figure 3 The dependence of Young's modulus on temperature for a number of materials. The curve for iron shows the values of Young's modulus for the three phases found in iron as a function of temperature. (Adapted from Bowman KJ (2003) *An Introduction to Mechanical Behavior of Materials*. Wiley; © Wiley. This material is used by permission of John Wiley and Sons, Inc.)

Anisotropy

When a material is anisotropic, the elastic constants vary with the direction along which the material is mechanically loaded. Even single-crystal materials have significant elastic anisotropies such that, the Young's modulus in one-crystal direction can be much different from that along another crystal direction. For example, in copper, elastically stretching the crystals along the cube axes of its unit cell requires one third as much stress as along the body-diagonal of the unit cell to produce the same strain. In iron single-crystals, the ratio between the smallest and largest Young's moduli is over two.

Because many engineering applications use materials that are comprised of many crystals or polycrystalline, the impact of directional processing (rolling, vapor deposition, extrusion, etc.) on producing elastic anisotropy was only recognized fairly recently. Developing ways to describe the relationship between oriented polycrystalline materials and the intrinsic single crystal properties has been an important area of research. Many newer applications of materials, including single crystals of superalloys used at high temperatures in jet turbines and the single crystals used in silicon computer chips, employ materials that have strongly anisotropic elastic properties.

Fiber-reinforced polymer composite materials posses directional properties that correspond to the arrangement of the fibers, which are normally stiffer than the polymer matrix material. Although models for fiber composite materials have existed for many years, their predictive capacity is often limited. The effects of the interfacial bonding, fiber arrangement and size scales, and their influence on anisotropic elasticity are not particularly well understood.

The importance of including elastic anisotropy in predictions of performance has also entered in evaluation of elastic constants of natural composite materials such as bone. Like nearly all natural materials, directional growth processes lead to elastic anisotropy. This elastic anisotropy depends on the type of bone, location on the bone, and the age of the bone. All three of these factors affect the ratio between the collagen and mineral contents of bone. The elastic properties and corresponding anisotropy have been increasingly recognized in the increasing fragility of bone with age and bone diseases, such as osteoporosis.

Defective growth patterns that occur in osteoporosis, bone diseases such as cancer or Paget's disease, can lead to bones with reduced elastic properties and strength. The mineral crystal content of bones is critical to maintaining stiff and strong bones. Bone mineral density (BMD) is measured to evaluate the condition of bone in aging patients, and provides a very potent indicator of decreasing elasticity and greater likelihood of bone fracture. One of the recently recognized limitations of bone density measurements is that it does not take into account the detailed structure of the bone mineral content, which can strongly affect the elastic properties of the bone. The morphology and preferred orientation of the bone mineral content results in strong elastic anisotropy such that the Young's modulus of a bone along its length can be nearly twice that in the radial direction. This anisotropy also plays a role in the lifetime of joint replacements. Many knee and hip replacements are principally constructed of metals or ceramics that have much higher elastic stiffnesses and lower degrees of anisotropy than does bone (see **Table 1**). The mismatch in elastic properties can lead to failure of the hip replacement.

Specification or the anisotropic elastic constants are usually facilitated through writing tensor relations between stress and strain. When the elastic constants are specified such that individual stresses are reported in terms of strains, the coefficients are called stiffnesses, c_{ij}, such that

$$\sigma_i = c_{ij}\varepsilon_j \qquad [6]$$

In eqn [6] the strain term on the right-hand side is the sum of all possible values of j.

When the elastic constants are written such that the individual strains are reported in terms of stresses, the coefficients are called compliances, s_{ij} such that

$$\varepsilon_i = s_{ij}\sigma_j \qquad [7]$$

In eqn [7], the stress term on the right-hand side is the sum of all possible values of j. The Young's modulus is the inverse of the compliance since a single applied stress can be related to the strain in the same direction, for example, for a stress applied in the x-direction.

$$\varepsilon_1 = s_{11}\sigma_1$$

Measuring Elastic Constants

For isotropic materials, direct measurements of elastic constants as shown in **Figure 1** is an effective approach to evaluate elastic constants such as Young's modulus and Poisson's ratio. Using tension and compression tests to measure anisotropic elastic constants requires testing of multiple specimens at a number of orientations. Easier approaches use other characteristics of the material that are related to elastic properties.

Probably, the easiest approach for measuring the anisotropic elastic constants of materials is to use measurements of sound velocity in a material. By

transmitting ultrasound pulses and measuring the time it takes for the pulse to propagate through the material, a quick and easy assessment of the elastic properties can often be made. When waves are propagated through a material, they interact with the mass of the atoms and the strength of the bonds between the atoms. The number of independent elastic constants required to fully describe the elasticity of materials is directly related to the symmetry of the crystals. For cubic crystals, three independent elastic constants are required to describe the elastic constants. For crystals and bulk materials with orthogonal symmetry (such that the character of the materials can be described by an orthogonal coordinate system), at least nine independent elastic constants must be determined.

Two types of waves are readily propagated through the bulk of materials, normal waves and shear waves. When normal waves are propagated through a material in the x-direction, the velocity is related to the normal stiffness as

$$v_1 = \sqrt{c_{11}/\rho} \qquad [8]$$

where ρ is the density and c_{11} is the normal stiffness. The normal stiffness c_{11} relates a stress to a strain in the same direction when the material is constrained such that, all other strains are held to zero.

For shear waves the velocity is related to the shear stiffness by the relation

$$v_4 = \sqrt{c_{44}/\rho} \qquad [9]$$

A recent development has been finding the anisotropic elastic constants using the resonant frequencies of crystals. The technique can be accomplished on a single specimen by oscillating the sample over a wide range of frequencies and finding the resonant values.

Increasing interest in nanoscale materials has led to the development of techniques to evaluate the elastic constants of thin films and particles of materials on size scales approaching a nanometer. Carbon nanotubes, which are derived from research on carbon C_{60} or "bucky balls," consist of several hundreds of carbon atoms and have diameters in the order of 1 nm. Many early predictions of their elastic properties overpredicted the Young's modulus. The planar structure of graphite leads to a very high Young's modulus within the planes of 1000 GPa and a Young's modulus across the planes that is more than a hundred times lower. The elastic constants of carbon nanotubes are also strongly anisotropic with values along the tube axis of several hundred GPa and values in the radial direction of less than 10 GPa.

Similar challenges in assessing the elastic properties are apparent in thin films as nanoindentation and scanning microscopy processes usually produce load versus displacement relations that are nonlinear due to loading geometry. When the loading behavior is modeled, it is often difficult to directly compare the elastic constants recovered to those for bulk materials. This can lead to scientists erroneously deciding that the thin film possesses an elastic behavior different from what is observed in the bulk. Thin film effects and other scaling impacts on elastic properties should always be evaluated carefully.

The Fundamental Nature of Elastic Constants

The details of the relationship between bonding and the elastic properties of materials has become more interesting as measurement techniques have improved and computer modeling has become more sophisticated. It has long been understood that elastic constants decrease almost monotonically as materials undergo greater thermal vibration at higher temperatures (see **Figure 3**). Once the melting temperature is reached the material loses its ability to support a load and the elastic constants fall to zero.

The balance between attractive forces and repulsive forces is often modeled with relatively simple expressions that can fail to accurately predict thermal expansion and elasticity and their temperature dependences. Although the approximate temperature dependence shown in **Figure 3** is useful for engineering calculations, in careful measurements, the elastic constants change rapidly as phase transformations are approached. It is apparent that the thermodynamic changes that occur at a phase transformation are signaled by strongly nonlinear changes in properties as the phase transformations are approached. The effects of alloying elements on elastic properties in solid solutions have also been shown to depend on the electronic structures of the materials, relative atomic radii, and other details of the components.

Many of the most advanced models for elastic response rely on constructing computer models that simulate the atomic arrangement in the material with assigned bonds between each of the atoms. Then virtual experiments can be conducted on a computer to simulate either the direct elastic response or the response of the lattice to wave propagation.

Extremes of Elasticity: Hyperelasticity

Materials with more than one type of fundamental bond, for example, ceramic and polymeric glasses, and materials tested under conditions at which they can sustain very high stress levels without failure, can

become decidedly nonlinear in their response. Understanding this nonlinearity and its effects of material performance has resulted in the study of what is called hyperelasticity. Hyperelastic responses are particularly observable in compressive loading with superposed hydrostatic stresses and complex localized stress states such as plane strain fracture and indentation loading. Many models for mechanical processes that rely on assumptions of linear elasticity may require substantial modification to describe the effects of hyperelasticity.

See also: Crystal Tensors: Applications; Mechanical Properties: Anelasticity; Mechanical Properties: Creep; Mechanical Properties: Fatigue; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Mechanical Properties: Tensile Properties; Recovery, Recrystallization, and Grain Growth; Thin Films, Mechanical Behavior of.

PACS: 62.20. – x; 62.20.Dc; 81.40.Jj

Further Reading

Badawi F and Villain P (2003) Stress and elastic-constant analysis by X-ray diffraction in thin films. *Journal of Applied Crystallography* 36: 869–879.

Betschart AA (1998) Soliton numbers. In: Pomeranz Y (ed.) *Optical Soliton Pulses in Fibers*, 3rd edn., vol. 2, pp. 91–130. London: Academic Press.

Buehler MJ, Abraham FF, and Gao J (2003) Hyperelasticity governs dynamic fracture at a critical length scale. *Nature* 426: 141–146.

Grigoras S, Gusev AA, Santos S, and Suter UW (2002) Evaluation of the elastic constants of nanoparticles from atomistic simulations. *Polymer* 43: 489–494.

Hartley CS (2003) Single crystal elastic moduli of disordered cubic alloys. *Acta Materialia* 51: 1373–1391.

Hurley DC, Shen K, and Jennett NM (2003) Atomic force acoustic microscopy methods to determine thin-film elastic properties. *Journal of Applied Physics* 94: 2347–2354.

Liu X, Wang X, and Niebur GL (2003) Effects of damage on the orthotropic material symmetry of bovine tibial trabecular bone. *Journal of Biomechanics* 36: 1753–1759.

Rhee M, Stolken JS, Bulatov VV, de la Rubia TD, Zbib HM, *et al.* (2001) Dislocation stress fields for dynamic codes using anisotropic elasticity: methodology and analysis. *Materials Science and Engineering* A309–310: 288–293.

Roeder RK, Sproul MM, and Turner CH (2003) Hydroxyapatite whiskers provide improved mechanical properties in reinforced polymer composites. *Journal of Biomedical Materials Research* 67A: 801–812.

Vijay A and Verma TS (2000) Analysis of temperature dependence of elastic constants and bulk modulus for ionic solids. *Physica B* 291: 373–378.

Zysset PK (2003) A review of morphology–elasticity relationships in human trabecular bone: Theories and experiments. *Journal of Biomechanics* 36: 1469–1485.

Mechanical Properties: Fatigue

F McBagonluri, University of Dayton Research Institute, Dayton, OH, USA
W Soboyejo, Princeton University, Princeton, NJ, USA

Introduction

Fatigue is a deleterious damage mechanism that occurs in advanced engineering structural components subjected to alternating loadings. Fatigue damage is driven by initiation and propagation stages. Different types of fatigue mechanisms have been identified in structural components, based primarily on loading conditions and prevailing loading environment. Most of these fatigue mechanisms manifest as fretting, dwell, thermomechanical, and multi-axial.

The result of fatigue failures is high casualty numbers and major financial losses. In fact, over 80% of mechanical failures in engineering structures and components are attributed to fatigue failures. Attempts to eliminate or reduce the incidence of fatigue failures have been the principal driving force behind the tremendous amount of research investments into the investigation of the initiation and propagation stages of fatigue.

Research in fatigue began over a century and a half ago following a series of industrial and railway accidents and the desire to reduce incidence of such failures. Much of the research was driven by the desire to understand the premature failures of components subjected to cyclic loading conditions at stress levels that were substantially below the pertinent design loads under monotonic loading.

The first systematic study of fatigue failures is usually attributed to the German engineer A Wöhler (1858–1871). In his preliminary studies, Wöhler concluded that when subject to cyclic loading conditions, the strengths of steel railway axles were lower than their static strengths. He concluded that the fatigue life was determined by the load range and not by the maximum load, as previously thought.

Ewing and Rosenhain, and Ewing and Humfrey conducted the first mechanistic studies of fatigue. In their studies conducted on Swedish iron, they identified the stages of fatigue crack initiation and propagation. In a series of investigations, they inferred, from

published optical micrographs of the surfaces of cyclically damaged specimens, that slip bands developed in the grains of polycrystalline materials subjected to alternating loads. They inferred that these slip bands increased with fatigue deformation, leading ultimately to crack formation.

Empirical models for characterizing fatigue damage evolution were developed by many researchers over the years to provide tools for the modeling of fatigue behavior. For instance, Basquin developed stress-based laws to characterize stress–life (S–N) curves of metals. He showed that a log–log plot of stress versus cycles exhibited a linear relationship over a large range of stresses.

Coffin and Manson, working independently in the early 1950s, noted that plastic strains were primarily responsible for cyclic damage. They proposed an empirical relationship between the number of load reversals to fatigue failure and the plastic strain amplitude. This formulation, known also as the Coffin–Manson relationship, is the most widely used approach for strain-based characterization of fatigue damage. Coffin and Manson also found that the plastic strain–life data can be linearized on a log–log scale. These relationships were combined to relate total strain range to life-to-failure. This relationship is called the strain–life relation. In the strain–life approach, the transition fatigue life represents the life at which the elastic and plastic strain ranges are equivalent.

The mathematical framework for the quantitative modeling of fatigue failure came later, and was primarily driven early by the stress analyses work of Inglis and later by the energy concepts of Griffith. This was followed by the work of Irwin. Irwin showed that the amplitude of the stress singularity ahead of a crack could be expressed in terms of the stress intensity factor, K. This formulation was developed based on the representation of the crack-tip stresses by the stress functions proposed originally by Westergaard, and the elasticity solutions of Love.

Paris and co-workers observed that the increment of fatigue crack advance per stress cycle, da/dN, can be related to the range of the stress intensity factor, ΔK, under constant amplitude cyclic loading. This observation set the stage for the direct application of linear elastic fracture mechanics to fatigue problems. In this approach, fatigue damage is characterized by fatigue cracks under conditions of small-scale plastic deformation at the crack tip. When gross plasticity occurs at the crack tip, the problem is considered to be one amenable to elastic–plastic or plastic analysis.

Elber provided further insights into the application of linear elastic fracture mechanics (LEFM) to fatigue. In studies of fatigue crack growth that were conducted by him as a graduate student in Australia, he noted that fatigue cracks could remain closed, even under cyclic tensile loads. He showed that the effective value of ΔK given by eqns [1] and [2]

$$\Delta K_{\text{eff}} = K_{\max} - K_{\text{cl}} \quad [1]$$

$$\Delta K_{\text{eff}} = K_{\max} - K_{\text{op}} \quad [2]$$

controls the fatigue crack growth rate, da/dN. Note that K_{cl} and K_{op} are the respective crack-closure and crack-opening stress intensity factors.

Further progress in fatigue investigations led to the breakdown of the similitude concept, in which cracked components in different dimensions were assumed to exhibit the same amount of crack growth, when subjected to the same far-field stress intensity factor range, ΔK. Following the pioneering work of Pearson, small cracks were found to exhibit anomalous crack growth behavior, with relatively fast growth rates occurring at ΔK level below the long crack fatigue threshold.

Mechanisms of Fatigue

Fatigue Crack Initiation

In most engineering structures and components, regions of high stress concentration are preferred crack initiation sites. The formation of microcracks around notches, scratches, inclusions, and along or across grain boundaries represent the onset of fatigue damage evolution. In metallic structures, microcracking is often driven by slip processes. Ewing and Humfrey referred to these slip bands as persistent slip bands (PSBs), in apparent reference to the fact that these bands were restored after electropolishing and then refatiguing. Fatigue crack initiation can occur by any one of the following identified processes:

1. From regions of localized strain hardening or softening resulting from slip step accumulation occurring at the surfaces of structures subjected to high plane strain amplitudes.
2. Formation of intrusions and extrusions on the surfaces of engineering structures resulting from sequential slip at the intersection of slip planes.
3. Microcrack formation along crystallographic planes of maximum shear stress by mode II mechanisms.
4. The coalescence of microcracks to form macrocracks, which are amenable to LEFM characterization.

Fatigue Crack Propagation

Fatigue crack propagation generally occurs in two stages, namely stage I and II. In stage I, crack

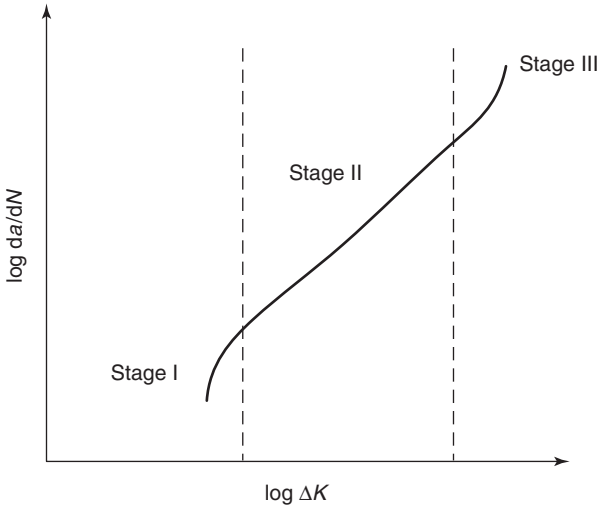


Figure 1 Schematics of the three distinct regimes of fatigue crack propagation.

propagation rate is characterized by low stress intensity factors. The crack growth behavior at this stage is associated with a threshold stress intensity value, ΔK_{th}, below which fatigue crack growth is negligible. Fatigue crack growth in the threshold is affected by stress ratio, frequency of loading, and prevailing environmental conditions.

Stage II fatigue crack growth is characterized by high ΔK and long crack lengths. In this region, crack growth rates are extremely high and most fatigue life is usually expended prior to this stage. On a plot of log da/dN versus log ΔK, the stage II region is linear (**Figure 1**). The Paris–Erdogan formulation is generally used to characterize fatigue crack growth within this regime. In variable-amplitude loading, the Forman and Walker fatigue crack propagation models are used.

Empirical Approaches

Stress-Based

The stress–life method, based on Wohler's S–N diagram, was the first quantitative approach. The stress–life approach is used in applications where the applied stress is essentially within the elastic range, and the material has a long cyclic life.

In general, the stress–life approach is used in situations where the fatigue life is in excess of 10^3 cycles. In addition, the stress–life approach excludes the true stress–strain behavior and assumes that only elastic strains are effective within the loaded structure. This assumption introduces significant computational errors, since plastic strains resulting from plastic deformation are a recognized crack initiation protocol. Thus, computational results using the stress–life approach are only informative when the degree of plastic straining is negligible, as in high cycle fatigue (HCF), where plastic strains are relatively small.

Stress–life data are usually represented by S–N curves, in which the stress range or stress amplitudes are plotted against the log of the number of cycles to failure, N. The S–N data for body-centered cubic materials generally exhibit an endurance limit, that is, a stress below which the specimens appear to have infinite life. In the case of face-centered cubic and hexagonal-closed-packed metals, a fatigue limit is often defined as a stress corresponding to a specified fatigue life, for example, 10^7 cycles.

Strain–Based

The strain–life method is one of the two classes of empirical approaches for the characterization of fatigue life in structural components. The other approach is the stress–life approach that was discussed in the preceding section. The strain–life model is based on the observation that in many structural components, the response of the material in critical locations, such as notches, is strain or deformation dependent.

Thus, the principal assumption for calculating fatigue life of a component under constant amplitude cyclic loading is that the strain range, $\Delta\varepsilon$, controls fatigue life. In the strain–life approach, the plastic strain is used directly to quantify the structural response to fatigue. However, in the stress–life approach, the plastic strain is not accounted for. The strain-based approach stipulates that at long fatigue lives, beyond the transition life, N_t, the elastic strain dominates the fatigue life, while the plastic strain is negligible with increasing life. In this case, fatigue strength (σ_f'/E) dominates fatigue.

In the case of short fatigue lives, the plastic strain is dominant and fatigue ductility (ε_f') is the controlling fatigue parameter. In general, most engineering structures are designed such that the nominal loads remain elastic. However, due to stress concentrations around notches, localized plasticity is often encountered in engineering structures and components.

One key attribute of the strain–life approach is that it disregards the presence of precracks. Fatigue-life prediction using the strain–life approach requires the following essential inputs: (1) cyclic stress–strain response and strain–life data, (2) stress–strain history obtained from critical locations such as notches, (3) damage counting techniques such as cycle counting, (4) mean stress effect models, and (5) damage summation techniques such as Miner's rule. In the strain–life approach (see **Figure 2**), the relationship between

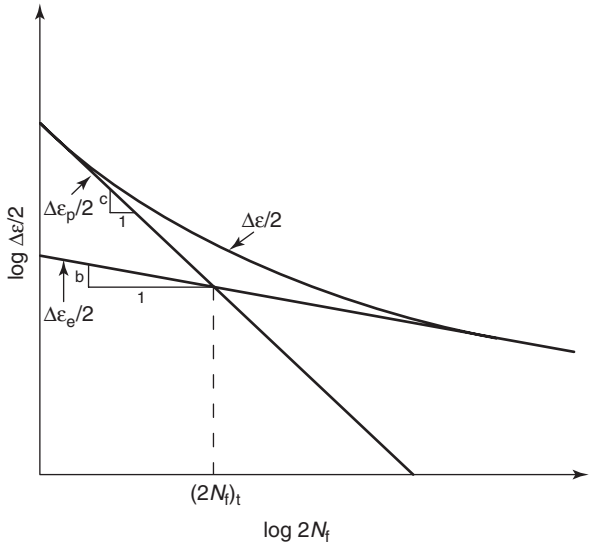


Figure 2 Schematics of the plot of elastic and plastic strain amplitude vs. life showing total strain amplitude.

the total strain amplitude and the number of load reversals to failure is given as

$$\frac{\Delta\varepsilon}{2} = \frac{\sigma'_f}{E}(2N_f)^b + \varepsilon'_f(2N_f)^c \qquad [3]$$

where ε'_f is the fatigue ductility, c is the fatigue ductility exponent, σ'_f is the fatigue strength coefficient, $\Delta\varepsilon$ is the strain amplitude range, $2N_f$ is the number of load reversals to failure, and b is the Basquin's exponent.

Variable-Amplitude Fatigue

Under service conditions, most engineering structures and components are exposed to a host of load spectra. For instance, an aircraft under turbulent conditions may be exposed to unique load spectra, compared to the conditions that are inherent to normal operating scenarios. It is, therefore, important to design for a wide range of loading conditions that can occur under unintended and uncontrollable service conditions.

The extensive literature on fatigue crack growth shows that experimental test data often differ considerably from predictions of crack growth under variable-amplitude loading conditions. Under constant-amplitude loading, the crack growth increment, Δa, is dependent only on the present crack size and the applied load. However, under variable-amplitude loading, the increment of fatigue crack growth is also dependent on the preceding cyclic loading history. In most cases, the interactions that occur under variable-amplitude loading adversely affect the fatigue lives of engineering components and structures.

LEFM Approaches

Fundamentals of Fracture Mechanics

The concept of linear elastic fractures has been applied to the design of engineering structures for nearly five decades. LEFM assumes that the material is isotropic and linearly elastic. Based on these assumptions, the stress field near the crack tip is often calculated using the theory of elasticity and Westergaard stress function. Within the framework of LEFM, crack growth occurs when the crack-tip stress intensity factor, K, exceeds the material fracture toughness, K_{Ic}. LEFM formulations are usually derived for either the plane stress or plane strain conditions associated with the three basic modes of loadings on a cracked body, namely, crack opening, sliding, and tearing.

The formulation of LEFM assumes that under applied loading conditions, crack tips produce a $1/\sqrt{r}$ singularity. The stress fields near a crack tip of an isotropic linear elastic material can be expressed as a product of $1/\sqrt{r}$ and a function of θ with a scaling factor K. The stress intensity factor plane strain formulations for mode I stress fields in Cartesian coordinates are

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \end{Bmatrix} = \frac{K_I}{\sqrt{2\pi r}}\cos\frac{\theta}{2}\begin{bmatrix} 1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2} \\ 1+\sin\frac{\theta}{2}\sin\frac{3\theta}{2} \\ \sin\frac{\theta}{2}\cos\frac{3\theta}{2} \end{bmatrix} \qquad [4]$$

In practical applications, the magnitude of the maximum stress near the crack tip and whether it exceeds the fracture toughness of a given material is usually of interest to design engineers. Hence, the stress intensity factor, K, which is a function of the loading and geometry, is expressed in terms of the "applied stresses," σ at a location expressed in polar coordinate by $r \to 0$ and $\theta = 0$ (where r and q are cylindrical coordinates of a point with respect to a crack tip). The solution of crack problems using the fracture toughness approach requires: (1) the determination of the crack geometry, (2) the calculation of the stress intensity factor, K, (3) the determination of the fracture toughness, K_C of the material (usually obtained from a material handbook), and (4) ensuring that the failure criteria $K \geqslant K_C$ is satisfied.

Fracture Mechanics and Fatigue

Paris and co-workers proposed that the increment of fatigue crack advance per stress cycle, da/dN, could be related to the range of the stress intensity factor, ΔK, under constant-amplitude cyclic loading. They

formulated the relationship given by

$$\frac{da}{dN} = f(\Delta K, K_{max}) \quad [5]$$

where

$$\Delta K = K_{max} - K_{min} = C(Y\Delta\sigma\sqrt{\pi a})^m \quad [6]$$

where K_{max} and K_{min} are the maximum and minimum stress intensity factors, respectively.

Paris and Erdogan, in their subsequent work, established that the fatigue crack advance per stress cycle could be related to the stress intensity range as follows:

$$\frac{da}{dN} = C(\Delta K)^m \quad [7]$$

where C and m are material constants.

In practical applications, two independent fracture mechanics parameters are required to characterize the state of stress at the evolving crack tip. A combination of any two of the following can be used: K_{max}, K_{min}, ΔK, and R.

Using eqn [3] and assuming a constant Y, both sides of eqn [7] can be integrated:

$$\int_{a_0}^{a_f} \frac{da}{a^{m/2}} = CY^m(\Delta\sigma)^m\pi^{m/2}\int_0^{N_f} dN \quad [8]$$

For $m > 2$:

$$N_f = \frac{2}{(m-2)CY^m(\Delta\sigma)^m\pi^{m/2}}\left[\frac{1}{(a_0)^{(m-2)/2}} - \frac{1}{(a_f)^{(m-2)/2}}\right] \quad [9]$$

For $m = 2$:

$$N_f = \frac{1}{CY^2(\Delta\sigma)^2\pi}\ln\frac{a_f}{a_0} \quad [10]$$

The constants C and m are material parameters that are usually determined experimentally. Typically, m values are in the range of 2–4 for most metals. For intermetallics, m is generally between 6 and 20. However, for ceramics, m can be as high as 100. For cases where Y depends on crack length, the above integrations are usually performed numerically. It is also important to note that the crack growth rate in the Paris regime is weakly sensitive to load ratio and is driven by ΔK.

The initial crack, a_0, can be determined using nondestructive die penetrants inspection, ultrasonics, or X-ray techniques. If no cracks are detected, one can assume a crack length to be at the resolution of the detection system. The critical crack length can be determined from the failure criterion when $K_{max} \rightarrow K_C$.

Fatigue Crack Growth Laws

Numerous fatigue propagation (see **Figure 1**) models have been developed in an attempt to characterize fatigue damage evolution in engineering materials. However, none of these models capture all aspects of the mechanisms of fatigue crack propagation. For instance, the extent of crack growth over a range of stress levels cannot be determined *a priori* from these models. However, these models do highlight the processes associated with the crack-tip opening during crack propagation.

Forsyth and Ryder proposed that fatigue crack extensions were the result of sporadic bursts of brittle and ductile fracture and that the contribution of each mechanism to crack propagation was a function of the ductility of the pertinent material. They also proposed that cracking occurs by necking, which obstructs the material until void coalescence occurs. These voids are usually created during forward cycling.

Fatigue crack propagation models include models by Paris, Forman, Walker, and Elber. The application of these models to practical engineering situations depends on the material type and loading conditions. Paris–Erdogan and Forman's models are used for constant-amplitude loading, while Walker and Elber models are used for variable-amplitude loading. These variable-amplitude loading models incorporate crack-growth retardation or acceleration, and threshold effects.

Paris' model In the mid-ΔK regime, the law shows that the crack advance per stress cycle is related to the stress intensity range. This is given by

$$\frac{da}{dN} = A(\Delta K)^m \quad [11]$$

where A and m are material constants determined experimentally.

Forman's model The Forman's equation is a form of the Paris' law that takes the effect of the applied stress ratio into consideration. It provides an empirical model for fatigue life estimation with mean stress effects. The Forman equation is given by

$$\frac{da}{dN} = \frac{c(\Delta K)^n}{K_f(1-R) - \Delta K} \quad [12]$$

where c is a pre-exponential constant, n is the power law exponent, K_f is a material constant, R is the stress ratio, and ΔK is the stress intensity factor range.

Walker's model The Walker's equation also accounts for stress ratio effects. The Walker and Forman's relationships are equivalent to Paris' law when $R = 0$, and the crack growth rate depends solely

on the ΔK:

$$\frac{da}{dN} = c[K_{max}(1-R)^m]^n \quad [13]$$

where K_{max} is the maximum stress intensity factor, R is the stress ratio, c is a pre-exponential constant, m is the power-law exponent, and n is a material constant.

Multiparameter law The multiparameter law proposes a fatigue life prediction approach using the combined effect of multiple variables on fatigue crack growth. The formulation is based on the multiple regression analysis and involves the representation of crack growth rate, da/dN, as a multiple variable statistical expression. Examples of such variables include stress intensity factor range, ΔK, crack-closure stress intensity factor, K_{cl}, and stress ratio, R.

Elber's model (crack-closure model) Elber formulated a crack-closure technique by modifying Paris' equation to include the dependence of the crack growth rate on the stress ratio. According to Elber, the plastic zone evolves around the crack tip, when the yield stress of the material is exceeded. Subsequently, a wake of plastically deformed material is developed and is encapsulated by the surrounding elastic medium as the crack grows. When the loaded component is being unloaded, the plastically deformed material causes crack surfaces to contact at zero tensile loads.

Elber concluded that crack closure retards fatigue crack growth rate by reducing the stress intensity range, ΔK. He introduced the effective stress intensity range, ΔK_{eff}, for use in Paris' law given by eqn [14]. The resulting modified Paris' law is given by eqn [16]:

$$\Delta K_{eff} = K_{max} - K_{op} = U\Delta K \quad [14]$$

where

$$U = \frac{\Delta K_{eff}}{\Delta K} \quad [15]$$

$$\frac{da}{dN} = D[\Delta K_{eff}]^P \quad [16]$$

where K_{max} is the maximum stress intensity, K_{op} is the stress intensity at which the crack opens, U is the so-called Elber closure ratio, and D is a pre-exponential constant.

Fatigue Fracture Modes

The surface morphologies of fatigue cracks are classified as follows:

- *Transgranular cleavage*. This occurs along defined planes in crystal structures. Cleavage formation is the result of high stress along the three axes where significant deformation occurs at low temperatures. Cleavage morphology is characterized by cleavage steps, feather markings, herringbone structure, and microtwins. Such morphologies are often observed in the near-threshold region.
- *Fatigue striations*. These occur generally in the mid-ΔK and high-ΔK regimes. In general, the striation spacings correspond to the crack growth rate under HCF conditions. However, under low cycle fatigne (LCF) conditions striations are sometimes not definable and hence cannot be correlated with crack growth rates.
- *Combined striations and static fracture modes*. At high-ΔK levels, a combination of fatigue striations and static fracture modes is observed, with the incidence of static modes increasing with increasing ΔK. The static modes often include secondary cracks and ductile dimples that are akin to those observed under monotonic loading. More detailed reports on fatigue fracture modes can be found in the "Further Reading" section.

The Short Crack Problem

Studies of fatigue behavior of short cracks (**Figure 3**) have attracted considerable attention since the total fatigue life of many components may be dominated by the short crack regime. These studies have altered the fatigue limit concept, from a damage limit to a critical propagation limit.

As a result of the successful application of the concepts of LEFM to the growth of long cracks, significant attempts have been made to apply similar methods to modeling the behavior of short cracks. The extent of localized plasticity, however, violates the assumptions of LEFM. This so-called short crack problem was first observed by Pearson. It has been attributed to the synergistic interactions between the effects of microstructure and localized plasticity.

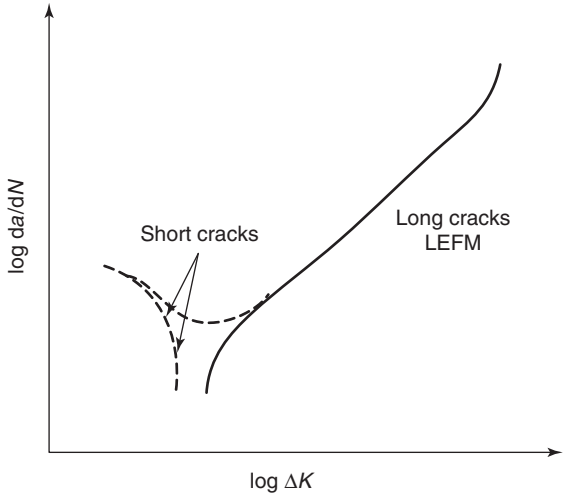


Figure 3 Schematic of typical fatigue crack growth behavior of short and long cracks.

A classification scheme for short cracks was proposed by Ritchie and Lankford by comparing their lengths to appropriate plasticity and microstructural length scales. For instance, mechanically small cracks have lengths that are comparable to their estimated plastic zone sizes; microstructurally short cracks are those that have lengths comparable to the microstructural dimensions of the pertinent continua. Physically, short cracks refer to those cracks that have dimension of $\leqslant 1$ mm.

Three common phenomena have been advanced to explain why small/short cracks can grow at higher growth rates than long cracks. These include: (1) limited crack-closure effects when the small crack has little or no wake ("physically small"), (2) inhomogeneous sampling of the microstructure when the cracks are smaller than relevant microstructural dimensions ("microstructurally small"), and (3) similarity between small crack lengths and the plastic zone size.

Summary and Concluding Remarks

The study of fatigue in mechanical components spans over a century and a half of effort, and is still an ongoing process. A basic understanding of fatigue damage stages currently finds applications in various fields such as biomedical, aerospace, and automotive. With the development of new materials and manufacturing processes, the quest to develop mechanistically based life prediction protocols should become more intense. The short crack problem, which defies the conventional LEFM approach to fatigue modeling, needs further research to develop models that include short crack effects. Recent advances in MEMS technology will provide nano- and microscopic damage information required to develop constitutive-based life prediction protocols.

Acknowledgments

The authors are grateful to Bruce Fenton and Dale Wilson of the Federal Aviation Authority (FAA) and Dr. Carmen Huber of the National Science Foundation (NSF) for their support and encouragement.

See also: Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Mechanical Properties: Tensile Properties; Periodicity and Lattices.

PACS: 61.72. – y; 62.20.Dc; 62.20.Fe; 62.20.x

Further Reading

Abelkis PR (1980) *Fatigue Loads. ASTM Standardization News*. Subcommittee Report, February 1980, pp. 18–22.

Albert WA (1838) *J. Arch. Minert* 10: 215.

Allameh P, Shrotriya AB, and Soboyejo WO (2003) Surface topography evolution and fatigue fracture in polysilicon MEMS structures. *J. MEMS* 12: 313.

Allen RJ, Booth GS, and Jutla T (1988) A review of fatigue crack growth characterization by linear elastic fracture mechanics (*LEFM*) Parts 1 and 2. *Fatigue and Fracture of Engineering Materials and Structures*, vol. 11(1), pp. 45–69 and no. 2, pp. 71–108.

Anderson TL (1995) *Fracture Mechanics: Fundamentals and Applications*, 2nd edn. Boca Raton, FL: CRC Press.

Barsom JM (1971) *Fatigue Crack Propagation in Steels of Various Yield Strengths. Trans. ASME* vol. B73, no. 4, 1190 p.; *Trans. ASME, J. Basic Eng*., vol. D89, no. 3, pp. 459–464.

Barsom JM (1973) *Fatigue Crack Growth under Variable Amplitude Loading in ASTM 514-B Steel*. ASTM STP 536, American Society for Testing and Materials (Philadelphia), pp. 147–167.

Basquin OH (1910) The exponential law of endurance test. *Proceedings of the American for Testing and Materials* 10: 625–630.

Blochwitz C (2001) Fatigue cracks: propagation of short cracks. In: Buschow KHJ, Cahn RW, Flemings MC, Ilschner B, Kramer EJ, and Mahajan S (eds.) *The Encyclopedia of Materials: Science and Technology*, pp. 2896–2906. Oxford: Elsevier.

Cottis RA (1986) The short fatigue of steels in saline environments: short cracks and crack initiation aspects. In: Ritchie RO and Lankford J (eds.) *Small Fatigue Cracks*, pp. 265–268. Warrendale, PA: TMS-AIME.

Cottrell AH (1961) *Iron and Steel Institute Spec. Rep*., 69, 281p.

Cottrell AH and Hull D (1957) Extrusion and intrusion by cyclic slip in copper. *Proceedings of the Royal Society, London* A242: 211–213.

Dowling NE (1971) *Fatigue Failure Prediction for Complicated Stress–Strain Histories*. Theoretical and Applied Mechanics Report 337, University of Illinois, Urbana.

Dowling NE (1999) *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue*, 2nd edn. Upper Saddle River, NJ: Prentice-Hall.

Elber W (1971) *The Significance of Fatigue Crack Closure*. ASTM STP 486, American Society for Testing and Materials (Philadelphia), pp. 230–242.

Ewing JA and Humfrey JC (1903) The fracture of metals under rapid alterations of stress. *Philosophical Transactions of the Royal Society* A200: 241–250.

Ewing JA and Rosenhain W (1900) Experiments in micro-metallurgy: effects of strain. Preliminary notice. *Philosophical Transactions of the Royal Society, London* A199: 85–90.

Fleck NA (1986) Fatigue crack growth – the complications. In: Smith RA (ed.) *Thirty Years of Fatigue Crack Growth Symposium*, Cambridge, UK.

Forman RG, Kearney VE, and Engle RM (1967) Numerical analysis of crack propagation in a cyclic-loaded structure. *Trans. ASME, J. Basic Eng. D* 89(3): 459–464.

Forsyth PJE (1961) *Proceedings of Crack Propagation Symposium*, vol.1, pp. 76–94. College of Aeronautics, Cranfield University, UK.

Forsyth PJE and Ryder DA (1960) Fatigue fracture. *Aircraft Engineering* 32: 96–99.

Fuchs HO and Stephens RI (1980) *Metal Fatigue in Engineering*. New York: Wiley.

Griffith AA (1921) The phenomenon of rupture and flow in solids. *Philosophical Transactions of the Royal Society* 45: 251–266.

Hertzberg IRW (1996) *Deformation and Fracture Mechanics of Engineering Materials*, 4th edn. New york: Wiley.

Inglis CE (1913) Stresses in a plate due to the presence of cracks and sharp corners. *Transactions of the Institute of Naval Architects* 55: 219–241.

Irwin GR (1957) *Journal of Applied Mechanics* 24: 361.
Laird C and Smith GC (1962) Crack propagation in high stress fatigue. *Philosophical Magazine* 8: 847–857.
Lewandowski JJ, Hippsley CA, Ellis MBD, and Knott JF (1987) Impurity effects on sustained load cracking of 2-1/4 Cr-IMo steels: crack initiation. *Acta Metallurgy of Materials* 35: 593–609.
Lindley TC, Richards CE, and Ritchie RO (1976) Mechanics and mechanisms of fatigue crack growth in metals: a review. *Metallurgia and Metal Forming* September 1976: 268–280.
Lou J, Shrotriya P, Allameh S, Buchheit T, and Soboyejo WO (2003a) A nanoindentation study on the plasticity length scale effects in LIGA Ni MEMS structures. *Journal of Materials Science* 38: 4137.
Lou J, Allameh S, Buchheit T, and Soboyejo WO (2003b) An investigation of the effects of thickness on the mechanics of LIGA Ni MEMS structures. *Journal of Materials Science* 38: 4129.
McDowell DL (1997) An engineering model for the propagation of small cracks in fatigue. *Engineering Fracture Mechanics* 56(3): 357–377.
McEvily A and Groeger J (1977) *Proc. ICF4*, Taplin (ed.), DMR2, 1293 pp.
McEvily AJ and Illg W (1958) The rate of propagation in two aluminum alloys. NASA TN 4394.
Mercer C, Soboyejo ABO, and Soboyejo WO (1999) Micromechanisms of fatigue crack growth in a single crystal Inconel 718 nickel–based superalloy. *Acta Mater* 47(9): 2727–2740.
Muhlstein CL, Brown SB, and Ritchie RO (2001) High-cycle fatigue of polycrystalline silicon thin films in laboratory air. In: deBoer M, Judy M, Kahn H, and Spearing SM (eds.) *Materials Science of Microelectromechanical Systems* (*MEMS*) *Devices III*, MRS Symposium Proceedings, vol. 657, pp. EE5.8.1–6. Pittsburgh, PA: Materials Research Society.
Paris P and Erdogan F (1963) *A Crtical Analysis of Crack Propagation Laws. Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers*. December 1963: 528–534.
Paris PC, Gomez M, and Anderson WE (1961) A rational analytic theory of fatigue. *The Trend in Engineering*, vol. 13, pp. 9–14. University of Washington.
Pearson S (1975) Initiation of fatigue cracks in commercial aluminum alloys and the subsequent propagation of very short cracks. *Engineering Fracture Mechanics* 7: 235–247.
Ritchie RO (1979) Near-threshold fatigue-crack propagation in steels. *International Metals Review* 4: 205–230.
Ritchie RO (1980) Application of fracture mechanics to fatigue, corrosion fatigue and hydrogen embrittlement. In: Shih GC (ed.) *Analytical and Experimental Fracture Mechanics*. Proceedings of the International Conference held in Rome, June 1980. Sijthoff and Nordhoff.
Ritchie RO and Knott JF (1973) *Acta Met*. 21: 639.
Ritchie RO and Lankford J (1986) *Small Fatigue Cracks*, vol. 129. Warrendale, PA: TMS.
Ritchie RO and Suresh S (1981) Some considerations on fatigue crack closure at near-threshold stress intensities due to fracture surface morphology. *Metallurgical Transactions* 13A: 937–940.
Shackelford J (1985) *Introduction to Materials Science for Engineers*, 4th edn. New York: Macmillan.
Soboyejo ABO, Shademan S, Foster M, Katsube N, and Soboyejo WO (2001) A multiparameter approach to the prediction of fatigue crack growth in metallic materials. *Fatigue and Fracture of Engineering Materials and Structures* 24(4): 225–242.
Soboyejo WO (1988) *Fatigue of Defects*. PhD thesis, University of Cambridge.
Soboyejo WO (2003) *Mechanical Properties of Engineered Materials*. New York: Dekker.
Soboyejo WO, Ni Y, Li Y, Soboyejo ABO, and Knott JF (1998) A new multiparameter approach to the prediction of fatigue crack growth. *Fatigue Fracture of Engineering Materials and Structures* 21(5): 541–555.
Stephens R, Fatemi A, Stephens RR, and Fuchs H (2000) *Metal Fatigue in Engineering*, 2nd edn. New York: Wiley.
Suresh S (1998) *Fatigue of Materials*, 2nd edn. Cambridge: Cambridge University Press.
Walker K (1970) *The Effect of Stress Ratio during Crack Propagation and Fatigue for 2024-T3 and 7075-T6 Aluminum*. ASTM STP 462, American Society for Testing and Materials (Philadelphia), 1 p.
Weibull W (1939) *Statistical Theory of the Strength of Materials*. Proceedings 151. Stockholm: Royal Swedish Academy of Engineering Sciences.
Westergaard HM (1939) *Trans. ASME* 61: A-49.
Wheeler OE (1972) Spectrum loading and crack growth. *J. Basic Eng., Trans. ASME* D94(1): 181–186.
Wohler A (1867) *Engineering* 4: 160–161.
Wood WA (1958) Formation of fatigue cracks. *Philosophical Magazine* 3: 692–699.

Mechanical Properties: Plastic Behavior

G Gottstein, M Goerdeler, and G V S S Prasad, Institut für Metallkunde und Metallphysik, Aachen, Germany

Deformation Fundamentals

Crystallographic Slip

Metals are crystalline except for very special conditions. Metals and alloys generally crystallize in a face-centered cubic (f.c.c.), body-centered cubic (b.c.c.), and hexagonal close-packed (h.c.p.) crystal structure. There is unambiguous evidence that metals retain their crystal structure during plastic deformation. This conservation principle of crystal plasticity has serious consequences for the deformation process and material properties. In a noncrystalline material, a macroscopic shape change can be accommodated on an atomistic level by a respective atomic rearrangement. If a macroscopic deformation of a crystal is also copied on an atomistic level, the crystal structure would change (**Figure 1a**); this is however contrary to observation. The crystal structure can be conserved, if the shape change on an atomistic level is accommodated by a displacement along crystallographic planes in multiples of translation vectors of the crystal lattice (**Figure 1b**). This phenomenon is referred to as crystallographic glide (or crystallographic slip). As a result, crystal plasticity proceeds by a structure

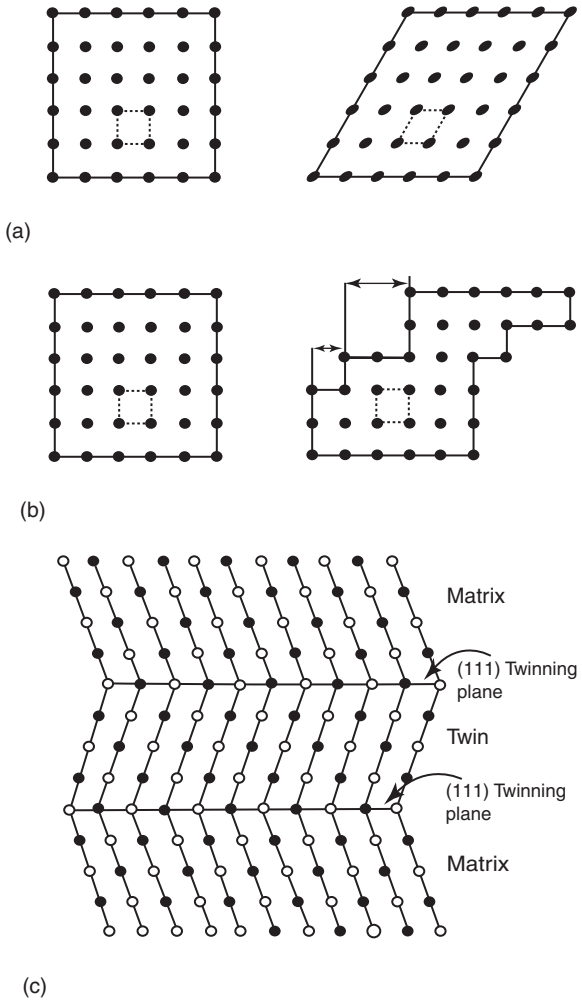


Figure 1 Plastic deformation of crystals (dotted lines = unit cell): (a) with change of crystal structure; (b) by crystallographic glide; and (c) by deformation twinning.

conserving simple shear deformation. In principle, there are an infinite number of planes and translation vectors to accomplish an imposed shape change. Due to energetic reasons, crystallographic glide is confined only to a few crystallographic planes (slip planes) and directions (slip directions), normally the most densely packed planes and directions in a crystal. A slip plane and a slip direction constitute a slip system and the most prominent slip systems of the three major crystal structures of metals are given in **Table 1**. The mechanism of crystallographic glide is related to the motion of crystal dislocations. The glide plane of a dislocation and the direction of its Burgers vector correspond to the slip plane and slip direction, respectively. On an atomistic level, therefore, plastic deformation of a metal requires the generation and motion of crystal dislocations.

Dislocations are crystal defects and, therefore, cause a distortion of the atomic arrangement, which manifests itself into a long-range stress field. Via their stress field, dislocations interact with each other. This constitutes a glide resistance, which is macroscopically felt as a yield stress and eventually causes immobilization of moving dislocations. The traveling distance (slip length) of a moving dislocation is normally much smaller than the macroscopic dimension of the deformed body; therefore, dislocations are stored in the crystal and the dislocation density grows with progressing deformation which in turn increases dislocation interaction and thus the flow stress. The relation between flow stress σ and dislocation density ρ can be represented by the Taylor equation

$$\sigma = \alpha M G \boldsymbol{b} \sqrt{\rho} \qquad [1]$$

where the dislocation density $\rho(\mathrm{m}^{-2})$ is defined as the total dislocation line length per unit volume – typically $10^{10}\mathrm{m}^{-2}$ for annealed and $10^{16}\mathrm{m}^{-2}$ for heavily deformed metals, $\boldsymbol{b}$ is the Burgers vector (slip vector), G is the shear modulus, and M is the Taylor factor which relates the applied stress to the shear stress in the glide system and depends on crystallographic texture (see below). Typically, $M \sim 3$, and the geometrical constant $\alpha \sim 0.5$.

Any change in the glide resistance of the moving dislocations will affect the flow stress. This is the reason for work hardening (increasing dislocation density), other strengthening mechanisms (grain size, solute atoms, particles, precipitates, etc.), and softening by recovery and recrystallization (decreasing dislocation density). It is important to note that the macroscopic deformation behavior can be related to the microscopic properties of the crystal dislocations. The force $\boldsymbol{K}$ on a dislocation due to a stress σ applied to a crystal is given by the Peach–Koehler equation

$$\boldsymbol{K} = \sigma \boldsymbol{b} \times \boldsymbol{s} \qquad [2a]$$

or the magnitude of the force K in the slip system due to the resolved shear stress τ

$$K = \tau \boldsymbol{b} \qquad [2b]$$

The imposed strain rate is reflected by the dislocation velocity through the Orowan equation

$$\dot{\varepsilon} = \rho_{\mathrm{m}} \boldsymbol{b} v \qquad [3]$$

where ρ_{m} is the density of the mobile dislocations moving with an average velocity v. The velocity, of course, depends on the applied stress σ. One can imagine the dislocations as rigid rods, moving with velocity v due to a force K, and this process is experienced macroscopically as a plastic deformation with strain rate $\dot{\varepsilon}$ at a flow stress σ.

Table 1a Slip systems of the major metallic crystal structures

Crystal structure	*Slip plane*	*Slip direction*	*Number of nonparallel planes*	*Slip directions per plane*	*Number of slip systems*
f.c.c.	{1 1 1}	$\langle 1\,\bar{1}\,0\rangle$	4	3	$12 = (4 \times 3)$
	{1 1 0}	$\langle \bar{1}\,1\,1\rangle$	6	2	$12 = (6 \times 2)$
b.c.c.	{1 1 2}	$\langle 1\,1\,\bar{1}\rangle$	12	1	$12 = (12 \times 1)$
	{1 2 3}	$\langle 1\,1\,\bar{1}\rangle$	24	1	$24 = (24 \times 1)$
	{0 0 0 1}	$\langle 1\,1\,\bar{2}\,0\rangle$	1	3	$3 = (1 \times 3)$
h.c.p.	$\{1\,0\,\bar{1}\,0\}$	$\langle 1\,1\,\bar{2}\,0\rangle$	3	1	$3 = (3 \times 1)$
	$\{1\,0\,\bar{1}\,1\}$	$\langle 1\,1\,\bar{2}\,0\rangle$	6	1	$6 = (6 \times 1)$

Table 1b Twinning systems of the major metallic crystal structures

Crystal structure	*Twinning plane*	*Shear direction*	*Displacement plane*	*Prototype*
f.c.c.	{1 1 1}	$\langle 1\,1\,2\rangle$	$\langle 1\,1\,0\rangle$	Ag, Cu
b.c.c.	{1 1 2}	$\langle 1\,1\,1\rangle$	$\langle 1\,1\,0\rangle$	α-Fe
h.c.p.		$\langle 1\,0\,-1\,1\rangle$		Cd, Zn

Twinning

While crystallographic glide by dislocation slip is the dominating deformation mechanism in metals, deformation twinning is also observed to accommodate a plastic strain, particularly at low temperatures in metals and many intermetallics. A deformation twin has an atomic arrangement that is the mirror image of the parent crystal. Due to the mirror symmetry, the crystallographic structure is conserved, but the crystal has a different shape (**Figure 1c**). The generation of a twin corresponds to a simple shear deformation by a displacement parallel to the twinning plane (mirror plane). In contrast to the dislocation glide, the shear strain γ is constant for a given crystal structure, for example, $\gamma = \sqrt{2}/2$ for cubic crystals. Twinning plane and displacement direction constitute a twin system in analogy to a slip system (**Table 1**).

Since for a given shear strain, the atomic displacement increases with growing distance from the twinning plane (**Figure 1c**), twins are usually thin, and their formation requires high stresses which are available only at low temperatures because the material strength increases with decreasing temperature. Deformation twinning is liable to occur if the flow stress is high and the twin boundary energy (approximately half of the stacking fault energy (SFE)) is low. The f.c.c. metals and alloys with low SFE undergo deformation twinning, such as silver at room temperature or copper at liquid nitrogen temperature. As the SFE decreases upon alloying, most alloys are prone to twinning during low-temperature deformation, for example, copper alloys, such as α brass (Cu–Zn) where twinning constitutes the major low-temperature deformation mechanism. Deformation twinning is most prominent in materials where dislocation slip cannot fully accommodate the imposed strain, for example, in h.c.p. metals and alloys with $c/a > 1.63$ owing to their low number of crystallographic slip systems. For an arbitrary deformation, a crystal needs to activate five independent slip systems to satisfy the five independent elements of the strain tensor. (A slip system of a set is independent, if its deformation cannot be accomplished by the other systems of the set.)

In h.c.p. metals with $c/a > 1.63$ (e.g., zinc), there is only slip on the basal plane and therefore, there are only two independent slip systems. This is sufficient to deform single crystals to very large strains. However, the individual grains of a polycrystal have to deform compatibly with the adjacent grains to avoid occurrence of holes or overlap. This can be accomplished only if the individual grains can undergo an arbitrary shape change at their boundaries, for which five independent slip systems are required. Therefore, polycrystals of zinc, for example, have to activate mechanical twinning for deformation. Owing to the high stresses and the less perfect adaptation to the shape change of neighboring grains in a polycrystal by twinning, the strain to fracture is smaller than in material deforming only by dislocation slip. Also, due to crystallographic effects, h.c.p. metals and alloys with $1.63 \leqslant c/a \leqslant 1.73$ cannot deform to a large degree by twinning, since twinning ceases to comply with the imposed shape change due to crystal reorientation and the fact that only one sense of twinning shear is crystallographically possible (the opposite shear $-\gamma$ does not generate a twin orientation). This is the reason why polycrystalline h.c.p. magnesium ($c/a = 1.63$) and its alloys are not ductile at ambient temperature. Note that hexagonal metals such as titanium, beryllium, or zirconium are very ductile even at low temperatures since their ratio $c/a < 1.63$, which favors nonbasal slip and thus, provides sufficient independent slip systems.

Deformation Microstructure

When gliding dislocations become immobile, they are stored in the crystal. The spatial distribution of these stored dislocations can occur at random or in distinct patterns. Most metals and alloys reveal nonstatistical dislocation distributions. Typically, dislocations arrange in cellular patterns, that is, in high dislocation density (cell) walls, which enclose volumes (cells) of considerably lower dislocation densities. For most commercial cold forming processes – such as rolling, the cellular arrangements are superimposed by deformation inhomogeneities, which typically have a band-like shape and can assume macroscopic dimensions to be recognized by the bare eye. In rolled materials, besides microscopic microbands, which appear like distinctly framed elongated dislocation cells under the microscope, there are two major types of deformation inhomogeneities namely, deformation bands and shear bands which can best be recognized in micrographs of the lateral surface. Deformation bands are elongated microstructural features that extend parallel to the rolling direction and that gradually accommodate a large misorientation to the matrix. Shear bands are deformation inhomogeneities, which extend under an angle of 30–35° to the rolling direction. Copper-type shear bands comprise clusters of elongated cells, which are confined to grain dimensions. In contrast, brass-type shear bands have (macroscopic) dimensions comparable to the sample thickness. Their internal structure is composed of very small granular features, typically globular grains of nanometer size. The origin and structure of shear bands are far from being understood, but they are generally assumed to be due to local softening (like texture softening) or enforced macroscopic shear in case of obstructed dislocation glide. Brass-type shear bands occur typically in alloys with low SFE (e.g., alpha brass). Copper-type shear bands are formed during cold rolling of f.c.c. metals and alloys with higher SFE.

Hardening Behavior

The mechanical properties of a material are typically measured in a uniaxial tension test, where the force is recorded to elongate a cylindrical specimen at constant speed. The force F divided by the initial cross section q_0 defines the engineering stress, and the length change Δl divided by the original length l_0 defines the engineering strain. If force and elongation are normalized by the instantaneous cross section q or length of the specimen l, they are referred to as true stress and true strain, respectively. The hardening curve is the dependence of the true stress on true strain. The true stress at a given strain is called the flow stress.

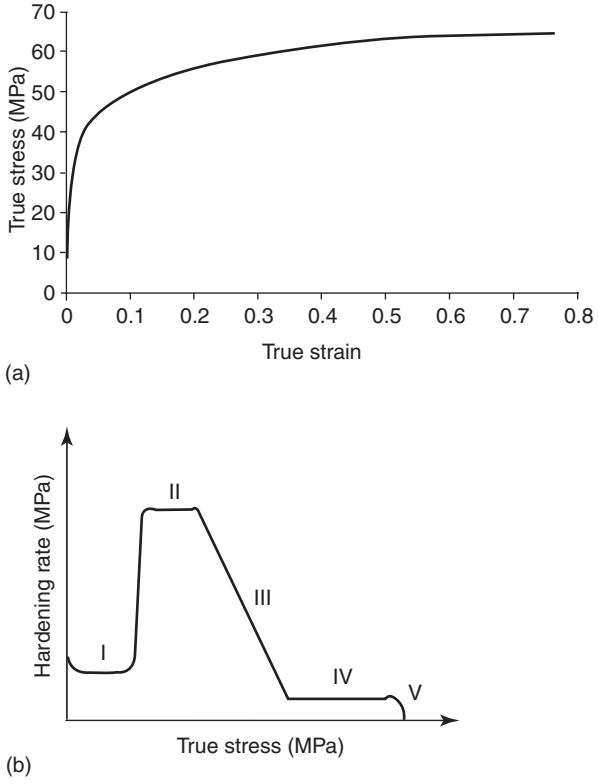


Figure 2 (a) True stress–true strain curve of an Al–1%Mn alloy (AA3103), measured at 300°C and a constant true strain rate of 0.1 s^{-1}; (b) $d\sigma/d\varepsilon$ vs. σ.

The storage of dislocations during deformation leads to an increased glide resistance for moving dislocations which is experienced as an increased flow stress with increasing strain (**Figure 2a**). The rise of stress σ with strain ε, is referred to as strain hardening or work hardening, and the slope of the hardening curve $d\sigma/d\varepsilon$ is the work-hardening coefficient or hardening rate. It is conventionally measured in uniaxial tension or compression tests. While the hardening curves of different materials or at deformation conditions vary considerably, there is a characteristic shape of the engineering stress–strain curve, namely after yielding a curve with a single maximum, although the strength level and hardening rate can be very different.

Of particular interest for forming processes is the hardening behavior for large strain deformation, for example, sheet rolling. A typical true stress–strain behavior for large strain deformation exhibits, after yielding, some transient with high but decreasing hardening rate until a low but constant rate is attained (**Figure 2b**), that is, the stress increases linearly with strain, and eventually a steady state (constant flow stress) is attained. Correspondingly, the

large strain-hardening curve exhibits five distinct stages:

1. Right after yielding, transition to athermal hardening (observed only in single crystals oriented for single slip);
2. Athermal hardening with constant high hardening rate $d\sigma/d\varepsilon \approx G/30$ (extended range only in single crystals at low temperatures);
3. Dynamic recovery range, hardening rate decreasing linearly with stress;
4. Large strain hardening with constant low hardening rate; and
5. Transition to steady state or unstable microstructure, usually not attained during cold deformation.

The important stages of the flow curves of commercial metals and alloys for industrial forming processes are stages 3 and 4. Their extent and level depends on material composition and processing and, thus, is history dependent! The hardening behavior reflects the glide resistance of the moving dislocations and, therefore, depends on the dislocation structure S, which is strongly dependent on processing history, that is, at strain rate $\dot{\varepsilon}$ and temperature T:

$$\sigma = \sigma(\dot{\varepsilon}, T, S) \qquad [4]$$

For this particular reason, the macroscopic strain is not a state parameter and, therefore, a prediction of flow stress as an empirical function of strain $\sigma(\varepsilon)$ is physically wrong although, due to simplicity, it has been a longstanding engineering practice.

Texture

A simple shear deformation is always accompanied by a rigid-body rotation, because the displacement gradient tensor is not symmetrical. Since plastic deformation in metallic materials proceeds by simple shear mechanisms (dislocation glide, mechanical twinning), the grains in a polycrystal change their crystallographic orientation during deformation – except for a few orientations which are stable under the imposed deformation conditions – and rotate toward the stable orientations. For a given forming process, such as rolling, wire drawing, or extrusion typical orientation distributions, that is, crystallographic textures develop, depending on deformation mechanism. An orientation distribution is most adequately represented in orientation space, where each orientation is represented by one point. The orientation distribution after large strain deformation is confined to a small volume in orientation space. For instance, the rolling texture of f.c.c. materials appears like a tube in orientation space (**Figure 3**) and can essentially be understood as consisting of only a few (ideal) orientations with some scatter. Crystallographic textures cause anisotropic mechanical properties, which make themselves felt during thermomechanical processing, metal forming, and in-service behavior.

Softening: Recovery and Recrystallization

During plastic deformation, dislocations are introduced to accommodate the imposed shape change. Dislocations, however, are crystal defects which are

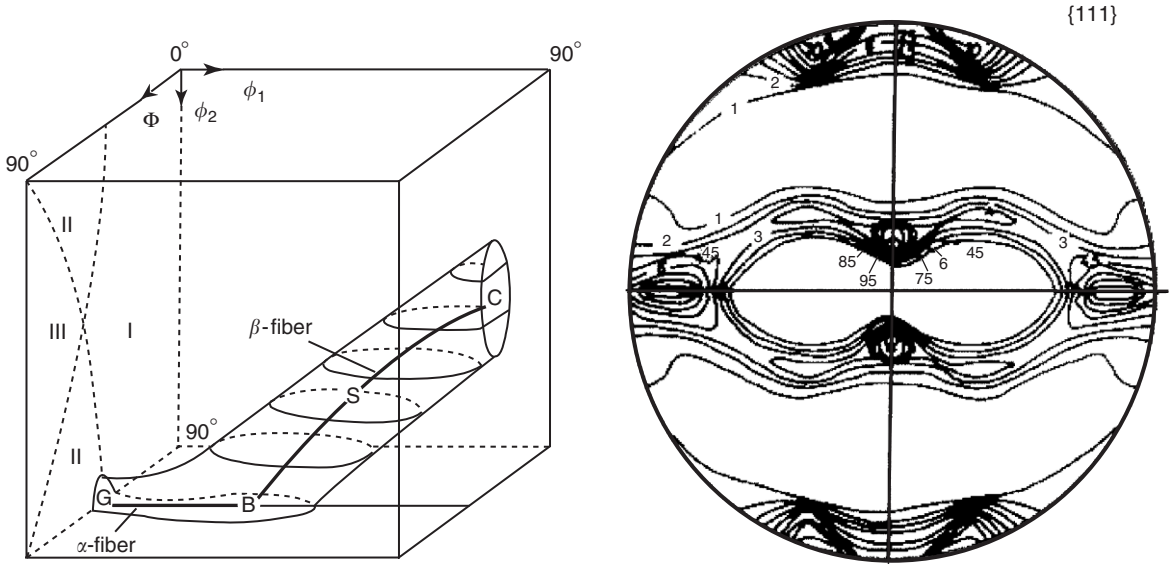


Figure 3 Rolling texture of pure copper: three-dimensional representation in orientation (Euler) space, {1 1 1} polefigure.

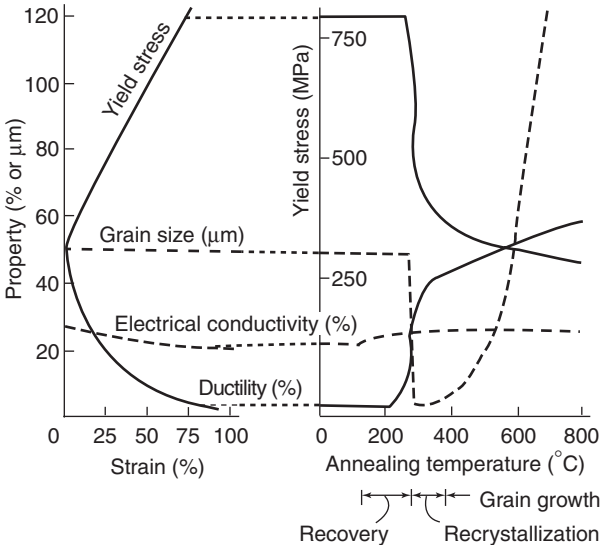


Figure 4 The effect of cold forming and annealing on the properties of a Cu–35%Zn alloy.

not stable in thermodynamic equilibrium; thus, the crystal strives to remove them. At low temperatures, the dislocation arrangement is in mechanical equilibrium and, therefore, stable. At elevated temperatures, however, thermally activated processes destabilize the dislocation structure and substantially change or even remove the deformation-induced microstructure which drastically affects microstructure, texture, and mechanical properties (**Figure 4**). There are two major restoration processes: recovery and recrystallization. They are termed static processes if they occur subsequent to deformation during annealing, or are referred to as dynamic recovery and dynamic recrystallization if they proceed during deformation. Dynamic recovery is also strain induced and is, therefore, an integral part of the strain-hardening process at any temperature. Dynamic recrystallization is an important phenomenon during hot working (see below) and liable to occur for $T > 0.5T_m$ (T_m – melting temperature). While recovery essentially consists of a dislocation rearrangement in the as-deformed structure, a completely new microstructure is formed atom by atom during recrystallization. Recrystallization proceeds by nucleation of virtually dislocation-free grains, which grow at the expense of the deformed matrix. The moving grain boundaries of the expanding nuclei are capable of absorbing the deformation-induced dislocations so that finally an almost dislocation-free granular structure is generated. Typical dislocation densities of recrystallized polycrystals are $10^{10}\,m^{-2}$ compared to up to $10^{16}\,m^{-2}$ for heavily deformed materials.

Recovery and recrystallization have different kinetics and different impact on properties. Invariably either process causes a softening of the material, and fully recovered and fully recrystallized materials have comparable yield stresses. The texture development is very different, however. The deformation texture is essentially retained during recovery or slightly increased in sharpness. During recrystallization, a completely new and different texture is generated. The recrystallization texture can be very sharp, like the cube texture in rolled and annealed aluminum, which gives rise to strong anisotropy during subsequent forming of the sheet. Typically, specific types of deformation textures, as shown in **Figure 3** for rolling, transform to specific recrystallization textures in f.c.c. and b.c.c. materials. However, small changes in chemical composition or processing history can totally alter the recrystallization texture, even if the deformation texture appears unchanged. Advanced codes for recrystallization structure and texture prediction with high spatial and temporal resolution (e.g., cellular automata or Monte Carlo simulation approaches) are now available which can be interfaced to FEM or superimposed to real deformation microstructures to account for spatial variation of the microstructure.

At temperatures above $0.5T_m$, recrystallization can also occur concurrently with deformation. This is referred to as dynamic recrystallization and makes itself felt by either a single-peak or multiple-peak flow stress behavior. A typical flow curve of a dynamically recrystallizing material and its associated microstructure are shown in **Figure 5**. Dynamic recrystallization typically occurs during the hot working of low SFE materials. Important examples are austenitic steels and copper alloys. It limits the flow stress and therefore, keeps hot working forces low and at the same time, it improves ductility (strain to fracture). The dynamically recrystallized grain size depends only on the flow stress and thus can be controlled by the strain rate and deformation temperature.

Modeling

For engineering practice, the mechanical behavior of materials must be known (e.g., the hardening curve). Most commonly, it is measured and fit to empirical functions. These functions, however, do not have any predictive power and cannot be used beyond the tested limits. For predictions of mechanical properties with the change of materials chemistry or processing conditions, microstructural models have to be used, since microstructure defines the state variables of materials properties.

Advanced work-hardening models are based on the dislocation theory. These models can either be statistical dislocation density-based models or dislocation

dynamic models (**Figure 6a**). Statistical models follow a common scheme, namely, the formulation of a structure evolution equation ($d\rho/d\varepsilon$) and a kinetic equation of state ($\sigma = f(\rho, T, \dot{\varepsilon})$) to relate the structure with the glide resistance. These relations are formulated as differential equations and hence can account for the transient microstructure evolution. The total change of dislocation density $d\rho$ in a strain increment $d\varepsilon$ reflects the structure evolution due to production of dislocations $d\rho^+$ and dynamic recovery (e.g., annihilation, dipole formation, lock formation) $d\rho^-$:

$$d\rho = d\rho^+ - d\rho^- = (f_1 - f_2)\, d\varepsilon \qquad [5]$$

The functions f_1 and f_2 have to be determined from the dislocation theory. They depend on the particular processes considered, and their actual values vary with the dislocation structure. The major problem is a proper incorporation of dislocation patterning in the theory. It is a common conception that the dislocations are arranged in a cell structure (**Figure 6b**), but different approaches have been proposed for the dominant stress and strain rate controlling processes. Due to the nonhomogeneous dislocation distribution, the flow stress will be a volume (V) weighted average of the flow stress in cell walls σ_w and cell interiors σ_i

$$\sigma = \sigma_w V_w + \sigma_i V_i \qquad [6]$$

In a dislocation dynamics model, the dislocations are modeled as line defects in a continuum and represented by their stress field. Since the continuum theory of dislocations is well established, the equations of motion can be formulated and the dislocation arrangement can be calculated from eqns [2] and [3]. However, the fundamental approach and the applicability to arbitrary deformation conditions have to be paid for by a substantial investment in computer power and computation time.

Owing to the long-range stress fields and the self-interaction of curved dislocations in three-dimensional dislocation dynamics, the problem becomes intractable for today's and tomorrow's high-power computers. Simulations with two-dimensional geometry are

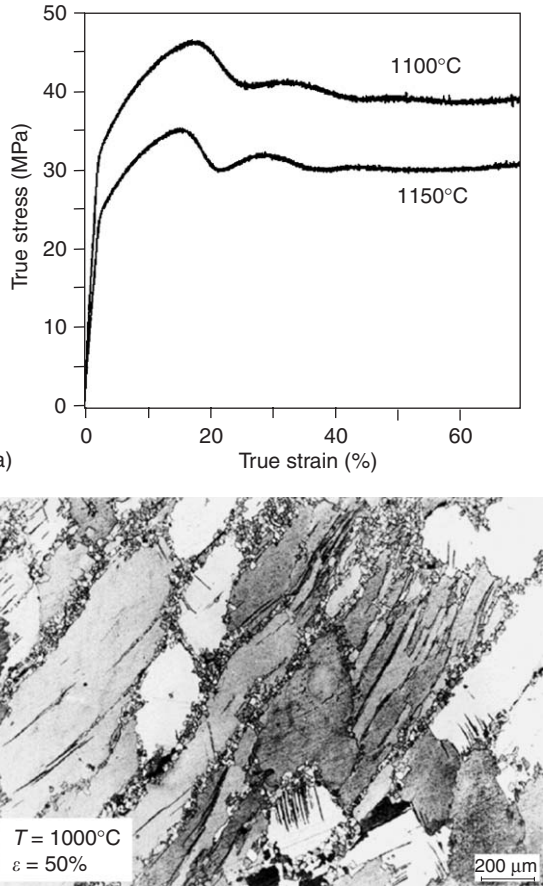


Figure 5 (a) Stress–strain curves of dynamically recrystallizing material (alloy 800 H); (b) the first recrystallized grains form predominantly at grain boundaries (necklace structure in Ni_3Al).

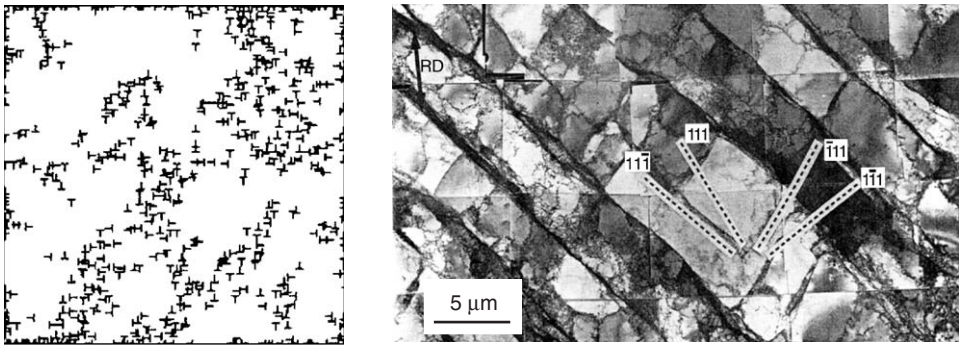


Figure 6 Deformation microstructure: (a) generated by two-dimensional dislocation dynamics; (b) observed in pure aluminum, cold rolled to a strain of 0.1. (Modified from Liu *et al.* (1998) Effect of grain orientation on deformation structure in cold rolled polycrystalline aluminum. *Acta Materialia* 46: 5819–5838, with permission from Elsevier.)

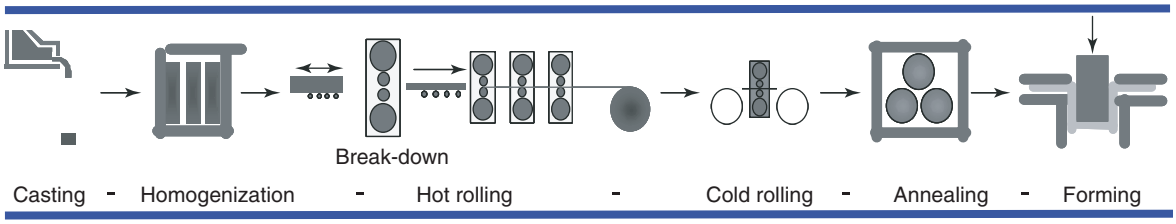


Figure 7 Production chain for sheet-forming process.

possible and have been performed but can only be applied to problems of two-dimensional geometry such as thin films. For bulk deformation, two-dimensional dislocation dynamics are of limited use.

Applications

Plasticity is the main prerequisite for metal forming, for example, rolling, forging, wire drawing, to produce semi-finished products and for the final shaping of parts, for example, by sheet forming. The excellent formability of metals is one of the reasons for their extensive industrial use. During the classical processing route of a metallic part, from the liquid state to the finished product, the material undergoes multiple deformation and heat treatment cycles. For example, consider the fabrication of an aluminum (or steel) can from a sheet (**Figure 7**).

After casting and homogenization the material is hot rolled, cold rolled, and annealed before the final part is produced from the sheet by deep drawing. While the material undergoes a shape change and experiences work hardening during rolling, it softens during interstand times and during back annealing. Since the microstructure changes during each processing step and as the microstructure determines the properties of the final part, processing has to be fine-tuned to meet terminal material specifications. These processes and the respective microstructural evolution can now be simulated by advanced computer codes.

See also: Dislocations; Orientation Texture; Recovery, Recrystallization, and Grain Growth.

PACS: 81.40.Lm; 81.40.Ef; 62.20.Fe

Further Reading

Backofen WA (1972) *Deformation Processing*. Reading: Addison-Wesley.

Cottrell AH (1961) *Dislocations and Plastic Flow in Crystals*. Oxford: Clarendon.

Courtney H (1999) *Mechanical Behavior of Materials*, 1st edn. Singapore: McGraw-Hill.

George ED (1988) *Mechanical Metallurgy*. London: McGraw-Hill.

Gottstein G (2004) *Physical Foundations of Materials Science*. Berlin: Springer.

Honeycombe RWK (1984) *The Plastic Deformation of Metals*, 2nd edn. London: Edward Arnold.

John PH and Jens L (1968) *Theory of Dislocations*. New York: McGraw-Hill.

Kocks UF, Argon AS, and Ashby MF (1975) Thermodynamics and kinetics of slip. In: *Progress in Materials Science*, vol. 19. Oxford: Pergamon.

Kocks UF and Mecking H (2003) Physics and phenomenology of strain hardening: the FCC case. *Progress in Materials Science* 48: 171–273.

Meyers MA and Chawla KK (1998) *Mechanical Behavior of Materials*. NJ: Prentice-Hall.

Nabarro FRN (ed.) (1979) *Dislocations in Solids*. Amsterdam: North-Holland.

Reed-Hill RE (1973) *Physical Metallurgy Principles*, 2nd edn. New York: Van Nostrand.

Nomenclature

$\boldsymbol{b}$	Burgers vector
F	external force
G	shear modulus
K	Peach–Koehler force (scalar)
$\boldsymbol{K}$	Peach–Koehler force (vector)
l	instantaneous length of the specimen
l_0	original length of the specimen
M	Taylor factor
q	instantaneous cross section of the specimen
q_0	initial cross section of the specimen
$\boldsymbol{s}$	dislocation line vector
S	dislocation structure
T	temperature
T_m	melting point
V_i	volume fraction of the cell interiors
V_w	volume fraction of the cell walls
α	geometric constant
γ	shear strain
Δl	change in length of the specimen
ε	strain
$\dot{\varepsilon}$	strain rate
v	average velocity of a dislocation
ρ	dislocation density
ρ_m	mobile dislocation density
σ	stress
σ_i	flow stress in the cell interiors
σ_w	flow stress in the cell walls
τ	shear stress

Mechanical Properties: Strengthening Mechanisms in Metals

A C Lund and C A Schuh, Massachusetts Institute of Technology, Cambridge, MA, USA

Introduction

The strength of metallic materials depends in considerable detail upon the structure at many length scales, from the fundamental atomic interactions at $\sim 10^{-10}$ m to interaction among macroscopic structural components as large as $\sim 10^{-3}$ m. The mechanisms responsible for strengthening metals have been the subject of considerable study for many decades, and are relatively well understood for traditional microstructural length scales. At low temperatures, the strength of crystalline metals is in large part governed by the interaction of dislocations with one another and with other features of the microstructure. Therefore, the main focus of this article is on these interactions, and the general scaling laws that relate these microscopic processes with macroscopic measurements of strength. In metals with microstructural components large enough to be regarded as continua, additional strength can be derived by sharing of the applied load between phases, which is discussed briefly. Finally, the limitations of these various strengthening mechanisms are discussed for cases where the temperature is high or the characteristic microstructural length scale is very fine.

Low-Temperature Dislocation-Based Strengthening Mechanisms

In crystalline metals, strength and other mechanical properties are governed by the prevalence and mobility of defects in the crystal structure. At low temperatures (i.e., below about one-third of the melting point), the most important defects in this regard are dislocations, which can be regarded as the "carriers" of plastic strain. Hence, the primary goal of most forms of metal strengthening is to regulate the movement of dislocations, which is generally accomplished through the introduction of obstacles. The four main strengthening mechanisms at ambient temperature, that is, work hardening, solid solution strengthening, particle hardening, and grain-boundary strengthening, are all variations upon this theme. Each of these mechanisms has been studied and modeled in considerable detail, and the reader is referred to the "Further reading" section for more exhaustive discussions. Here these concepts are explored within the framework of a relatively general scaling law that predicts the shear strength increment $\Delta\tau$ gained by incorporating obstacles of spacing L in the structure:

$$\Delta\tau \propto \frac{1}{L} \qquad [1]$$

This additional stress is required to allow dislocations to bypass the obstacles, and the proportionality constant depends in large part upon whether dislocations can cut through the obstacles ("soft" obstacles), or must bypass them by bowing ("hard" obstacles).

In what follows, how eqn [1] applies for the important mechanisms listed above is discussed broadly, assuming pure metals or dilute alloys, and neglecting details such as the crystallography of defects or the underlying crystal structure of the base metal. Furthermore, the homogeneous nucleation of dislocations is not considered to contribute to strength, since most engineering metals and alloys already contain an appreciable density of dislocations ($\sim 10^{10}$–10^{14} m^{-2}) prior to plastic flow.

Work Hardening

The principle underlying work hardening is that plastic flow induces structural changes which make subsequent plastic flow ever more difficult; the more it is deformed, the stronger a metal becomes. During plastic flow of a polycrystalline metal, work is dissipated by the motion and interaction of dislocations. Strain energy is stored through a net increase in dislocation line length through processes such as bowing around various obstacles and pinning points, and the sequential emission of new dislocations from, for example, Frank–Read sources.

As dislocations entangle, they can form sessile kinks or jogs which act as pinning points along the dislocation line, such that the obstacle spacing L is proportional to $\rho^{-1/2}$, with ρ the dislocation density. When this scaling is introduced into eqn [1], it yields the classical parabolic-hardening law used widely for polycrystalline metals:

$$\Delta\tau = \alpha\mu b\sqrt{\rho} \qquad [2]$$

where the constants α, μ, and b have now been introduced. In this formulation μ is the shear modulus, b is the magnitude of the Burger's vector, and α is a temperature-dependent proportionality constant ~ 0.2–1 in the athermal limit.

Equation [2] has been experimentally validated for many metals; an example of data compiled for copper is shown in **Figure 1**. Perhaps counter-intuitively, the

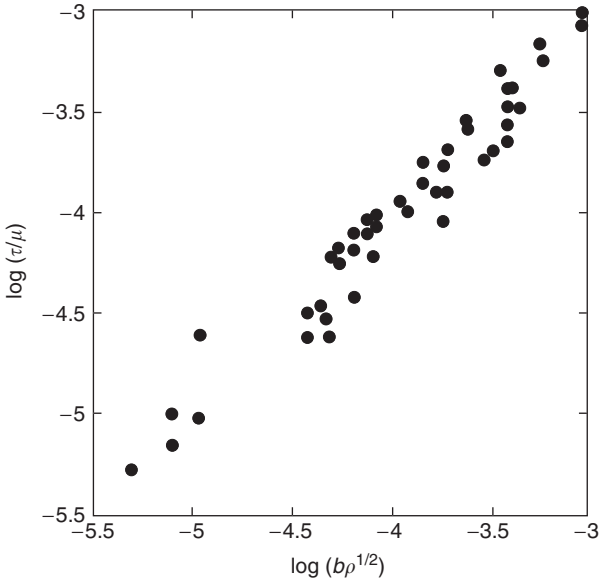


Figure 1 Data relating the shear strength of copper (normalized by the shear modulus μ) to the square root of dislocation density (normalized by the Burger's vector magnitude). The dislocation density has been obtained by chemical etching and optical microscopy on planar cross sections. (Data from Mecking H and Kocks UF (1981) Kinetics of flow and strain hardening. *Acta Metallurgica* 29: 1865–1875.)

nature of the dislocation distribution does not affect the general relevance of eqn [2]. For example, in some materials dislocations self-assemble into inhomogeneous spatial configurations. In such cases, it is possible to parametrize, for example, the average dislocation cell size, and use it to produce additional scaling laws; these ultimately reduce to either eqn [2] or other very similar relationships.

Although eqn [2] is surprisingly general in its applicability, it is important to note that $\Delta\tau$ from work hardening is fundamentally limited by the maximum value of ρ that can be sustained in a given metal. With large amounts of plastic strain, the dislocation density can increase by orders of magnitude before saturating at a temperature- and material-dependent value as high as $10^{16}\,\mathrm{m}^{-2}$. This saturation is a consequence of "dynamic recovery," wherein the production of new dislocation line length is offset by annihilation processes which become more active with increasing dislocation density. Nonetheless, strain hardening can increase a metal's strength by orders of magnitude compared with the stress required to move an isolated dislocation.

A topic of current research interest is the effect of nonuniform deformation on strain-hardening behavior, in which plastic strain gradients are necessitated by the test or microstructural geometry. Plastic strain gradients require the presence of "geometrically necessary" dislocations to accommodate them, so it has been speculated that higher total dislocation densities will result whenever such gradients exist. These effects become increasingly important at fine length scales, for example, in very small specimens or thin films, nanoindentation experiments, or for metals with fine scale microstructural features; in these cases, the strain gradients resulting from nonuniform deformation are steeper owing to the reduced dimensionality. In general, strain gradient effects are believed to occur in metals at length scales below $\sim 100\,\mu\mathrm{m}$.

Solid Solution Strengthening

For pure metals, the addition of a low concentration of a second component in solution can provide soft barriers to the motion of dislocations through the crystal. Solutes interact with dislocations through two main effects. First, solute atoms change the local values of elastic constants, which in turn change the line energy of a dislocation. Solute atoms which stiffen the matrix repel dislocations, while those which make the matrix more compliant attract them. Second, solute atoms added to a pure metal create distortions in the crystalline lattice, and the associated stress fields can interact with dislocations; again the interaction can be either attractive or repulsive, depending upon the size mismatch of the solute atoms with respect to the matrix and their position with respect to the dislocation core. Whether the solutes repel or attract dislocations, the net result is strengthening, as additional stress is required to force a dislocation to approach or depart, respectively, the vicinity of a solute atom.

Within a slip plane, the density of solute atoms is given by their concentration, c, and the spacing between solutes goes as $b \times c^{-1/2}$. Therefore, eqn [1] takes a parabolic form with respect to concentration:

$$\Delta\tau = \beta\mu b\,\frac{\sqrt{c}}{b} \qquad [3]$$

where the proportionality constant β depends upon how strongly the solute atoms interact with dislocations. In the case of substitutional alloying additions, the stress field around solute atoms is approximately spherical, which leads to rather mild interactions and values of β much less than unity. For some crystal symmetries, the strain field associated with an interstitial solute can be nonspherical, providing much harder obstacles to dislocation motion, and values of β near unity. A common important example in this regard is the interstitial sites in a body-centered cubic lattice, where the distortion accompanying a solute is tetragonal. Models are available to predict β with reasonable accuracy in each of these cases. **Figure 2** shows some classical data sets for substitutional and

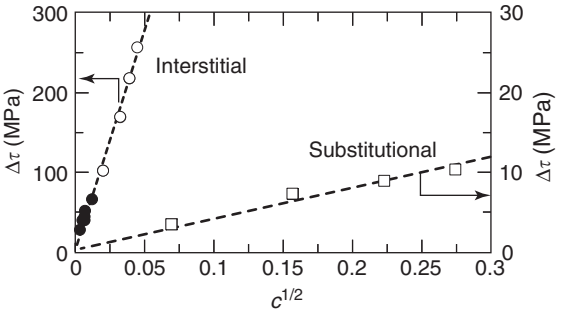


Figure 2 Classic examples of the concentration dependence of the solid solution strength increment $\Delta\tau$. Solutes with symmetric strain fields such as the substitutional case of Cu dissolved in Al (□) generally have a modest strengthening effect, while asymmetric strain fields associated with, for example, interstitial C in Fe (●) or N in Nb (○) produce more significant strengthening. (Data are from Wert CA (1950) *Journal of Metals* 188: 1242–1244, Evans PRV (1962) *Journal of the Less-Common Metals* 4: 78–91, and Koppenaal TJ and Fine ME (1962) *Transactions of the Metallurgical Society of AIME* 224: 347–353.)

interstitial alloying, where the parabolic form of eqn [3] is validated and the relative strengths of the two obstacle types are apparent.

Particle Strengthening

In the same vein as solid solution atoms, fine dispersed particles can also serve as dislocation pinning points. When second-phase particles are present by virtue of nucleation and growth processes, the metal is said to be "precipitation strengthened," whereas the extrinsic incorporation of fine particles is generally referred to as "dispersion strengthening." In the simplest case these particles can be assumed to be "hard" obstacles to dislocation motion, and their strengthening effect is based upon the spacing of the particles, L. Unlike solute atoms or dislocation locks, particles have finite sizes that can be of similar order to the interparticle spacing, λ. Therefore, the effective obstacle spacing is given by $L \approx (\lambda - 2r)$, where r is the average radius of the particles. For small particles, the spacing of obstacles in the slip plane of a dislocation goes roughly as $v_f^{1/2} r^{-1}$, so the familiar parabolic strengthening law is recovered using eqn [1]:

$$\Delta\tau = \gamma\mu b \frac{\sqrt{v_f}}{r} \qquad [4]$$

The proportionality constant γ is close to unity for hard particles, for example, for dispersed ceramic particulates or very large precipitated particles in a metallic matrix.

In many cases the particulate obstacles are, at least in principle, capable of being sheared by the passage of a dislocation. In these cases, the obstacles are said to be "soft," and the energy penalty associated with shearing them is the source of strengthening. Some notable examples by which such soft obstacles can promote strengthening are as follows:

1. Coherent particles with a lattice parameter mismatch to the matrix phase create strain fields that interact with dislocations in a manner similar to that described above for solid solution strengthening.
2. The shearing of particles creates additional interfacial area and thereby increases the total interfacial energy of the system.
3. In "order strengthening," an incident dislocation creates an antiphase boundary (with an attendant interfacial energy term) in a chemically ordered particle.
4. The difference in elastic modulus between particle and matrix changes the dislocation energy as it traverses the particle, providing an obstacle to either its entrance or exit from the particle.

All of these "soft" obstacle strengthening mechanisms, as well as some more obscure ones, tend to obey the scaling law of eqn [4], with different scaling parameters γ that can also be functions of the particle size r. However, it should be noted that some derivations for case (2) can yield a linear volume fraction dependence.

Apart from the volume fraction dependence of particle strengthening described in detail above, the absolute magnitude of the average particle dimension also has a measurable impact on the strength required to pass a dislocation. Larger average particle sizes promote strengthening by creating "harder" obstacles, because a passing dislocation must traverse a larger area of precipitate, requiring more energy via mechanisms (2) and (3) above, while the stress field of a coherent particle also increases with size and enhances strength via mechanism (1). However, a larger average particle size necessitates a proportionally wider interparticle spacing if their volume fraction is held constant; this effect leads to a weakening via eqn [1]. Therefore, for a constant v_f, increasing the average particle size gives rise to a "cutting-to-bowing" transition, with the material strength first increasing and then decreasing with average particle size. These two regimes correspond to shearing of soft obstacles for small particle sizes (cutting), and pinning of dislocations at hard obstacles for larger particle sizes (bowing). An example of such a transition is shown for the technologically relevant nickel–aluminum system in **Figure 3**, in which the particles are precipitates of the Ni_3Al phase.

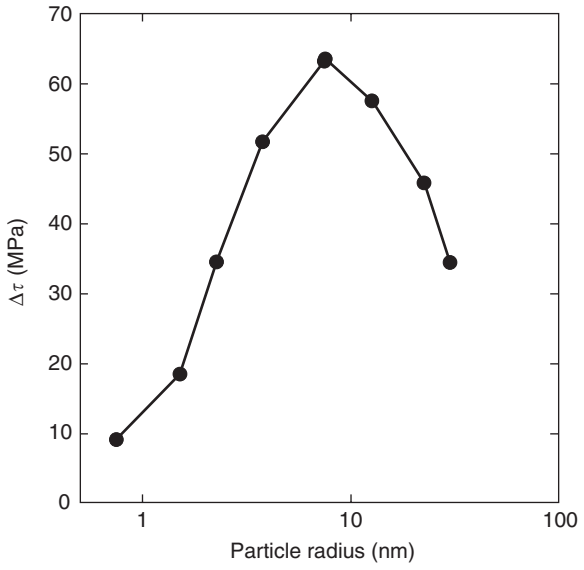


Figure 3 The strengthening of Ni due to the presence of a constant volume fraction dispersion of fine Ni_3Al precipitates with a given average radius. Over the range of radii <10 nm, increasing the precipitate size promotes strengthening, but beyond this range causes weakening. These regimes correspond to the "cutting" and "bowing" mechanisms, respectively. (Data from Brown LM and Ham RK (1971) Dislocation–particle interactions. In: Kelly A and Nicholson RB (eds.) *Strengthening Methods in Crystals*, pp. 12–135. London: Applied Science Publishers.)

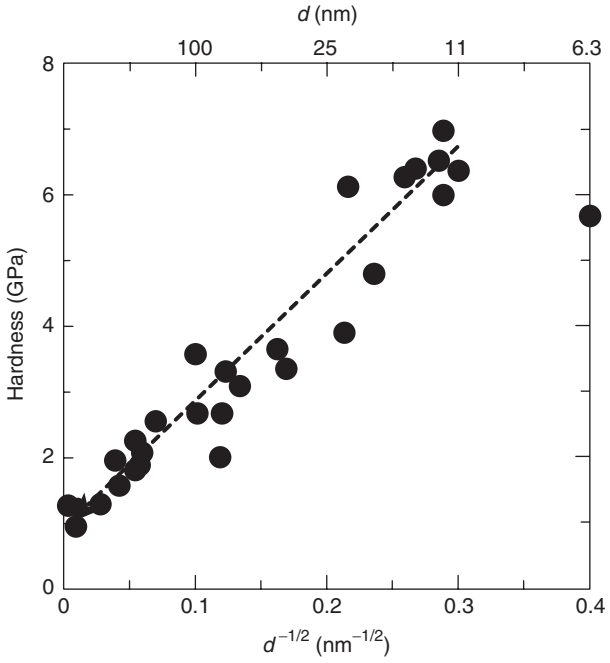


Figure 4 A typical "Hall–Petch" plot showing the hardness (which is proportional to the yield stress) of nickel as a function of the reciprocal square-root grain size, *d*. (Data are collected in Schuh CA *et al.* (2002) *Scripta Materialia* 46: 735–740, from various other sources.)

Grain-Boundary Strengthening

Because the adjacent crystals at a grain boundary have a misorientation between both the slip plane traces and the slip directions, dislocations cannot, in general, transmit through grain boundaries without an additional increment of applied strain energy. For more extreme misorientations, the stress required for transmission can be higher than that required to activate a new dislocation source in the adjacent grain. In either case, dislocations on a common slip plane in the first grain tend to pile up at the grain boundary, until the stress at the head of the pileup becomes large enough to activate slip in the second grain. In this manner, grain boundaries present very "hard" obstacles to dislocation motion, but they are not point obstacles in the same vein as solute atoms or particles. Therefore, the scaling law of eqn [1] does not simply apply to the case of boundary strengthening, although a similar parabolic hardening law does emerge from analysis of the slip transmission problem:

$$\Delta\tau = \kappa \frac{1}{\sqrt{d}} \qquad [5]$$

where d is the average grain size of the metal and κ is a material-specific constant with units of $(\text{length})^{1/2}$; eqn [5] is commonly referred to as the Hall–Petch relationship.

In **Figure 4** the hardness, which is linearly proportional to strength, is plotted for pure Ni as a function of the reciprocal square root of grain size. The data from multiple authors indicate that the strength increases linearly with $d^{-1/2}$, and eqn [5] holds over orders of magnitude in grain size, from hundreds of microns in standard engineering materials down into the increasingly important nanocrystalline regime. As evidenced by the conformity of the data from various authors working on different specimens, eqn [5] is quite robust. Potentially complicating issues such as distributions of grain size or grain-boundary misorientation tend to be subsidiary to the scaling law of eqn [5], and are subsumed in the constant κ. Far stronger variations are observed from one material to another, most particularly depending on the number of slip systems available.

A current topic of particular interest in the area of grain-boundary strengthening is the validity of the Hall–Petch relationship at grain sizes in the nanometer range, where extreme strength has been reported. However, at a grain size of zero, the Hall–Petch relationship would predict an unphysical infinite strength, so it has been widely speculated that eqn [5] will break down at nanometer length scales. Recent experimental and computational work has supported this speculation, including the experimental data for nickel shown in **Figure 4** below ~ 12 nm. This

breakdown is associated with the activation of deformation mechanisms normally associated with high temperature, as will be discussed in more detail in a later section.

Low-Temperature Continuum-Level Strengthening

On a coarser structural scale, it is possible to strengthen a base metal by adding a second phase that can bear a disproportionate fraction of the load. This is often desirable when the base metal has properties critical to the application, such as low density, high ductility, conductivity, or corrosion resistance, but lacks sufficient mechanical strength. The concept of load sharing is the basis of many technological structural materials from reinforced concrete to advanced polymeric composites. In the area of metallurgy, load sharing operates in microstructures with second phases formed *in situ*, such as steels containing coarse iron carbide or martensite phases, as well as in metal–matrix composites such as SiC-reinforced aluminum.

The load-sharing mechanism arises completely due to mismatch between the mechanical properties of the constituent phases, and, from a mechanistic perspective, can be regarded as independent of microscale mechanisms; dislocation interactions with coarse second phases are less important than the global transfer of load that results from such interactions. Accordingly, continuum mechanical approaches are generally used to derive the strengthening effect from load transfer. If the two phases or components of the microstructure are initially deformed only elastically, then the only variable controlling the strength of the dual-phase structure is the fraction of the applied stress that is carried by each. The total applied stress σ partitions such that the average stress carried by each phase (σ_1 and σ_2 for phase "1" and "2," respectively) obeys

$$\sigma = \sigma_1(1 - v_f) + \sigma_2 v_f \qquad [6]$$

where v_f is the volume fraction of phase 2. Although this rule is generally true for any dual-phase structure when the stresses are averaged over the volume of the material, eqn [6] is not always useful to predict, for example, yield strength. This is because the local values of stress may be significantly concentrated, leading to local yielding in one of the phases. Despite this limitation, eqn [6] illustrates the basic principle of load transfer, where a stronger phase can relieve the stress borne by the weaker, giving a strength intermediate between the two components. Accurate predictions of yield and fracture strengths for dual-phase materials are dependent on the details of the composite geometry, for which a vast literature exists. The reader is referred to the "Further reading" section for more detailed treatments of specific cases.

Limitations of Strengthening Mechanisms due to Thermal Activation

The strengthening mechanisms described above can all be understood on an athermal basis, and are best applied in the limit where diffusional rearrangement of matter is slow. When dislocation motion and diffusion have similar timescales, for example, at high temperatures, the above scaling laws no longer strictly apply. In these cases, new physics are required to accurately predict the strength and rheology of metals. Although it is beyond the scope of this article to treat diffusive deformation mechanisms in detail, it is important to appreciate the limitations on the mechanisms described above for cases in which (1) bulk diffusion and (2) interfacial or grain-boundary diffusion are active at the deformation temperature.

Bulk Diffusion Effects

For many of the strengthening mechanisms described above, the motion of dislocations (particularly those of edge character) is considered to be restricted to a plane. However, the climb of edge dislocations out of their slip plane can occur by the absorption or emission of vacancies at the dislocation core. In many cases, this enables an additional means for obstacle bypass, albeit one that is rate-limited by the diffusive flow of vacancies to (or from) the dislocation core. The supply of vacancies needed to enable dislocation climb can flow either through the bulk, which generally requires temperatures above about two-thirds of the melting temperature, T_m, or along the core of the dislocation itself, which can be significant at somewhat lower temperatures above about $T_m/3$.

For solid-solution strengthened materials, diffusion can also lead to mobility of the obstacles themselves. In cases where there is an attractive interaction between solute atoms and dislocations, there is a tendency for these features to move in tandem. For example, at temperatures where the solute mobility is high, the motion of dislocations becomes a viscous process as the solute atoms follow the moving dislocation. At somewhat lower temperatures, the phenomenon known as "strain-aging" can occur, whereby plastic deformation initially detaches dislocations from solute pinning points, but extended time or thermal exposure can allow solutes to redistribute and again pin the dislocations. For substitutional solid solutions, the solute can be regarded as essentially immobile at temperatures

below approximately $T_m/2$, while for interstitial solid solutions this limit can be considerably lower, down even to ambient temperature.

In multiphase microstructures, high temperatures or long durations can also lead to evolution of the size and spacing of precipitated obstacles. During this so-called "coarsening" process, the volume fraction of precipitate phase remains constant while the average precipitate size increases linearly with the cube root of time. As described above, a limited amount of such particle growth can be beneficial, but the maximum strength associated with the cutting-to-bowing transition generally occurs at very fine precipitate sizes, such that microstructural coarsening usually leads to weakening. This issue is sometimes suppressed in practice through the addition of slow diffusing elements that decelerate coarsening. Finally, coarsening is not usually an issue for dispersion-strengthened materials, which employ particles with low solubility and/or diffusivity in the matrix.

Intercrystalline Diffusion Effects

Additional time-dependent deformation mechanisms are associated with the diffusive rearrangement of atoms in intercrystalline regions such as grain and phase boundaries or triple junctions. For example, it is possible for strain to be accommodated solely by the diffusional flow of atoms along grain boundaries, from interfaces under net compression to those in net tension. For fine microstructures, grains can even become mobile relative to one another through the operation of viscous "grain-boundary sliding" and stress-induced grain rotations. Although these mechanisms do not directly influence all of the dislocation-based strengthening mechanisms described earlier, they can negate the expected scaling laws set forth above when they operate at a lower stress level than the athermal mechanisms. For example, the Hall–Petch scaling law that describes grain-boundary strengthening is known to fail at high temperatures, where finer grains actually promote weakening due to the dominance of diffusive deformation mechanisms. For these reasons it is common to develop coarse microstructures or even single crystals for high-temperature structural metals. Continuum-level load sharing can also be frustrated by interfacial diffusion and sliding between phases in a composite structure, increasing the fraction of load borne by the weaker matrix phase.

Because interfaces tend to have excess free volume, they also represent "short-circuit" diffusion paths that are more easily activated, and therefore intercrystalline diffusion can become significant at even lower temperatures than bulk diffusion. Additionally, since the net amount of matter that can be transported along interfaces scales with the interface density, these processes are also accelerated in microstructures of fine scale. For example, it is presumed that the high density of crystalline interfaces in nanocrystalline materials leads to diffusional deformation mechanisms even at room temperature, again leading to the breakdown of the Hall–Petch boundary strengthening law.

See also: Dislocations; Mechanical Properties: Creep; Mechanical Properties: Tensile Properties; Phases and Phase Equilibrium; Point Defects.

PACS: 62.20.Fe; 61.72. – y

Further Reading

Courtney TH (2000) *Mechanical Behavior of Materials*, 2nd edn. New York: McGraw-Hill.

Hirth JP and Lothe J (1982) *Theory of Dislocations*, 2nd edn. Malabar, FL: Krieger Publishing Company.

Kocks UF and Mecking H (2003) Physics and phenomenology of strain hardening: the FCC case. *Progress in Materials Science* 48: 171–273.

Martin JW (1980) *Micromechanisms in Particle-Hardened Alloys*. Cambridge: Cambridge University Press.

Mura T (1987) *Micromechanics of Defects in Solids*, 2nd edn. Dordrecht, The Netherlands: Martinus Nijhoff Publishers.

Nembach E (2000) Order strengthening: recent developments with special reference to aluminum–lithium alloys. *Progress in Materials Science* 45: 275–338.

Nomenclature

b	Burger's vector magnitude (length)
c	concentration of solute (atomic fraction)
d	grain size (length)
L	spacing between obstacles (length)
r	average particle radius (length)
T_m	melting temperature (K)
ν_f	volume fraction of particles
α	dimensionless constant for work hardening
β	dimensionless constant for solid solution strengthening
γ	dimensionless constant for particle strengthening
$\Delta\tau$	additional strength increment (force/area)
κ	constant for grain-boundary strengthening (length$^{1/2}$)
λ	interparticle spacing (length)
μ	shear modulus (force/area)
ρ	dislocation density (length^{-2})
σ	stress (force/area)

Mechanical Properties: Tensile Properties

K K Chawla, University of Alabama at Birmingham, Birmingham, AL, USA

Introduction

Tensile test is one of the most simple and important tests that can be done on a material to obtain its mechanical characteristics. It is commonly used with metals and polymers, and its importance is beginning to be appreciated even for brittle materials such as ceramics and glasses. The correct interpretation of the various types of tensile test requires a detailed knowledge of the procedure as well as the test machine. Two types of machines are used: (1) an endless screw-driven machine that has a constant speed of rotation or (2) a servo-hydraulic machine. The variables involved in a tensile test are stress, strain, and time. Generally, one of these variables is kept constant during the test. In a tensile test, one gets a uniform stress field over a relatively large effective volume of test material. These are important for evaluating inherent flaw distributions and the resultant strength distributions.

Tensile Test

In all tensile tests, a long specimen is subjected to a tensile force along its length. The material, in response to the applied force, undergoes some deformation. Commonly, the specimens are flat or round in cross section. The end of a sample is anchored to a fixed part of the machine while the other end is secured to the crosshead that can move at a constant speed. As the specimen is elongated, the applied force is measured by means of a load cell which is put in series with the specimen. It is very important to realize that all of the displacement of the crosshead is not transmitted to the gauge length section of the specimen. A part of the displacement is transmitted to the load cell, grips, coupling between specimen and the grips, and the machine frame. Thus, it is preferable to measure the specimen elongation "directly" by means of an extensometer. The output of such a test is thus a force–elongation record which can be transformed into engineering stress–engineering strain or true stress–true strain curves. The ends of the specimen are fixed in grips and a tensile force is applied by the machine such that the reduced section or gauge length is, at a given instant, under a constant force. The longitudinal deformation over the gauge length can be measured by means of an extensometer or strain gauge. Many testing systems, either screw-driven or closed-loop servo-hydraulic machines, with a variety of accessories are available. Modern machines come with a variety of options such as on-board programs to test different materials, Windows-based software, etc. **Figure 1** shows a schematic of a screw-driven machine.

The variables involved in a tensile test are stress, strain, and time. Stress is the load divided by the area while strain is the amount of deformation divided by the length over which the deformation takes place. True stress and true strain quantities can be defined. Engineering stress is the load divided by the original area of cross section of the sample while true stress is the load divided by the instantaneous area of cross section. Engineering strain is defined as

$$e = (l - l_o)/l_o$$

where l_o is the original gauge length and l is the distance between gauge marks after the force is applied. There is another type of strain, viz., true strain (also called natural or logarithmic strain) ε that describes deformation incrementally and is defined as

$$\varepsilon = \int dl/l = \ln(l/l_o)$$

It can be readily shown that for small values of engineering strain, $e = \varepsilon$. Engineering strain can be written as

$$e = (l - l_o)/l_o = l/l_o - 1, \quad l/l_o = 1 + e$$

Or, true strain, $\varepsilon = \ln(l/l_o) = \ln(1 + e)$.

If the series expansion of $\ln(1 + e)$ is examined, it is found that at low strains, indeed, $e = \varepsilon$. Commonly, one obtains a load–displacement curve by pulling on the sample at a constant rate of displacement or constant rate of force. Such a constant strain rate tensile test is very convenient to evaluate the strength and ductility of a material in a reasonable time. **Figure 2** shows the tensile stress versus strain curve for an aluminum sample: a typical stress–strain curve for a ductile metal. The initial response of a material to a force is linear elastic. This is also called Hookean behavior because the material obeys Hooke's law in this region. Hooke's law says that $E = \sigma/\varepsilon$, where E is Young's modulus, σ is the stress, and ε is the strain. In this linear region, the deformation is called elastic which means that it is reversible. When the force is removed, the material returns to its original dimensions. At some point during a tension test, a strain is reached where the deformation is no longer recoverable on unloading, that is, there is permanent or plastic deformation. The stress at which this irreversible deformation starts is called the yield stress or yield point. Frequently, it is difficult to determine exactly the yield

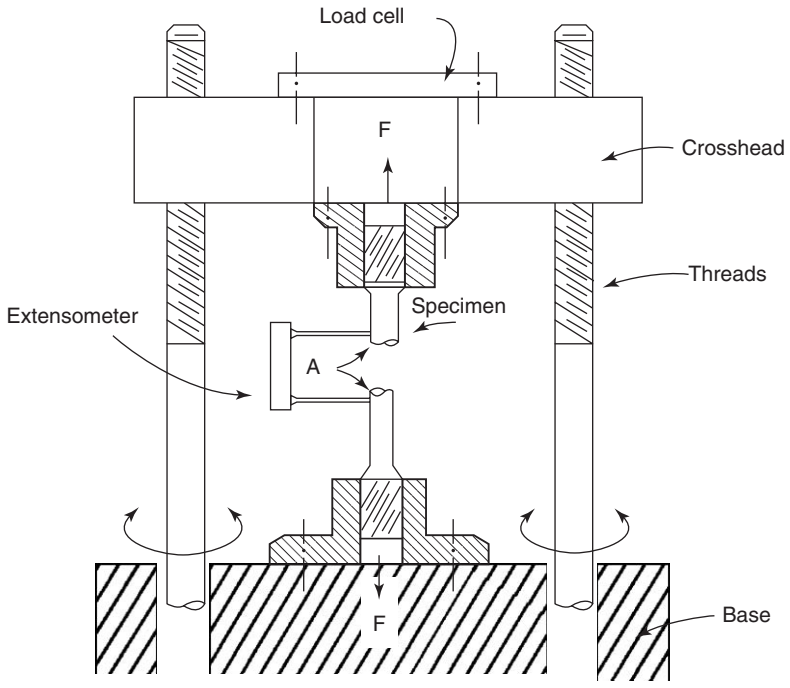


Figure 1 Schematic of a screw-driven machine.

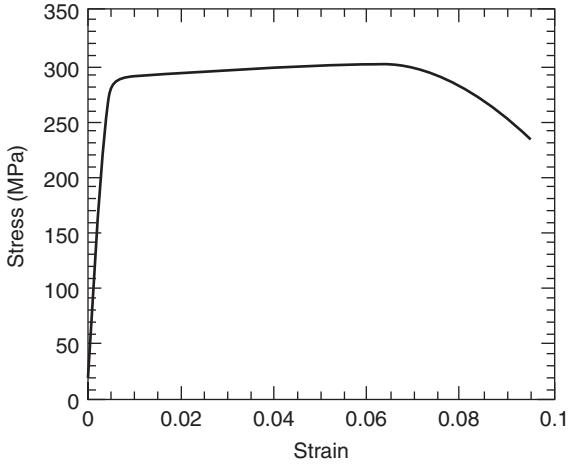


Figure 2 An engineering stress–strain curve in tension of an aluminum sample.

point. In order to avoid any ambiguity, an offset yield strength is defined. For example, 0.2% yield strength is strength obtained by drawing a line parallel to the elastic portion with an offset of 0.2% on the strain axis. Sometimes the stress–strain curve beyond the yield stress is called flow curve. For a perfectly plastic material, after attaining the yield stress, the stress will remain constant. In real material, as shown in **Figure 2** for aluminum, the flow stress increases with increasing strain. This is called "strain hardening" or "work hardening." The internal structure of material is changing during deformation, resulting in an increase in material strength. Concomitantly, the cross-sectional area of the specimen is decreasing and so the specimen is able to support a lower load. The maximum in the engineering stress–strain curve corresponds to the point state where the increase in load-carrying ability due to strain hardening is balanced by the decrease in load-carrying ability due to a decrease in the cross-sectional area. In a tensile test, this maximum in stress is called the ultimate tensile strength (UTS). Beyond UTS, the stress drops off. The extent of the stress–strain curve beyond the yield point is a measure of a material's ductility. When a material responds to a stress which is higher than its yield stress by deforming, it has some ductility. Low-ductility materials will break instead of deforming plastically. Materials showing such a behavior are called ductile materials. In ductile materials, the deformation after the maximum point is not uniform. At UTS, a unique phenomenon associated with tensile testing intervenes. It is called necking. Necking refers to a localized region where plastic deformation is concentrated in the sample. When necking occurs, the cross-sectional area is smaller than the rest of the gauge length, so the stress to continue deformation is smaller. The two phenomena, ductility and necking, are explored in the following two sections.

Ductility

A material is considered to be ductile if it is capable of undergoing a large amount of plastic deformation

before failure. The subject is of great importance in engineering because the phenomenon of ductility allows a material to redistribute localized stresses. A material that is unable to deform plastically before failure is called a brittle material. The ductility of a material can be measured easily in the tensile test. There are two quantities that can be obtained as measures of ductility. The first is the strain-to-fracture, e_f, which is defined as a percentage, as follows:

$$e_f = \frac{l_f - l_o}{l_o} 100$$

where l_f is the length at fracture and l_o is the original gauge length.

The other quantity is the reduction in area at fracture (RA). This is also given as a percentage and is defined as

$$\text{RA} = \frac{A_o - A_f}{A_o} 100$$

where A_f is the final cross-sectional area at fracture and A_o is the original area of cross section.

In a ductile material, after the beginning of necking, the load continues to fall until fracture occurs. All the deformation in this final stage (i.e., after the outset of the necking phenomenon) is concentrated in the region of the neck and its neighborhood. Thus, one may regard the total elongation of a tensile sample to consist of two components, viz.,

1. uniform elongation until the onset of necking, and
2. localized elongation after the onset of necking.

The localized strain and the uniform strain can be related by the following expression:

$$l_f - l_o = a + e_u l_o$$

where l_o is the original specimen length, l_f is the final specimen length, a is the local strain after the onset of necking, and e_u is the uniform strain. Clearly, the quantity $e_u l_o$ represents uniform elongation. The quantities a and e_u are constants that depend on the conditions of the test.

For the strain-to-fracture,

$$e_f = \frac{l_f - l_o}{l_o} = \frac{a}{l_o} + e_u$$

This equation implies that e_f is not an absolute material property for a given strain rate and temperature. It varies with the original gauge length, l_o. The uniform elongation, e_u, however, does not depend on l_o. Thus, the uniform elongation, e_u, is a better indicator of the ductility of a material.

Necking or Plastic Instability

Necking or plastic instability is a phenomenon that is unique to tensile testing. In ductile metals, necking starts at the maximum load in tension. A nonstrain-hardening, perfectly plastic material would become unstable in tension and start necking right in the beginning of the plastic yielding. Real metals, however, do show the phenomenon of strain hardening, that is, their capacity to bear load increases with strain. This effect is opposed by the gradual decrease in the cross-sectional area of the specimen with straining. Necking or localized plastic deformation starts at the maximum load, the point at which the increase in stress due to the decrease in cross-sectional area surpasses the increased load-bearing capacity due to strain hardening. This plastic instability condition is defined by $dP = 0$, that is, a small change in load equals zero. The applied load in a tensile test can be written as

$$P = \sigma A$$

where σ is the true stress and A is the instantaneous area of cross section. At necking,

$$dP = \sigma dA + A d\sigma = 0$$

or

$$-\frac{dA}{A} = \frac{d\sigma}{\sigma} \qquad [1]$$

From the condition of volume constancy during plastic deformation (i.e., the volume before plastic deformation is the same as that after the plastic deformation),

$$d\varepsilon = \frac{dl}{l} = -\frac{dA}{A} \qquad [2]$$

From eqns [1] and [2], at necking, one obtains

$$\frac{d\sigma}{d\varepsilon} = \sigma \qquad [3]$$

Thus, one can obtain the necking point corresponding to the maximum load. It is the point at which the strain-hardening rate equals the value of true stress.

Necking conditions in terms of engineering strain At necking, eqn [3] is valid. Then, denoting engineering strain by e, one can write

$$\frac{d\sigma}{d\varepsilon} = \frac{d\sigma}{de}\frac{de}{d\varepsilon} = \frac{d\sigma}{de}\frac{dl/l_o}{dl/l} = \frac{d\sigma}{de}\frac{l}{l_o} \qquad [4]$$

Now

$$\frac{l}{l_o} = \frac{l + \Delta l}{l_o} = 1 + e \qquad [5]$$

Thus, from eqns [4] and [5], one obtains

$$\frac{d\sigma}{d\varepsilon} = \frac{d\sigma}{de}(1+e) \qquad [6]$$

From eqns [3] and [6], one gets

$$\frac{d\sigma}{d\varepsilon} = \frac{\sigma}{1+e} \qquad [7]$$

Equation [7] gives what is called the "construction of Considére" for determining the maximum load point, that is, onset of necking. One can obtain the necking point corresponding to the maximum load starting from the true stress–true strain curve by finding the point on this curve at which the slope is unity or by finding the point at which the strain-hardening rate equals the value of stress.

Flow Curve of Metals

In shaping or processing of materials, large plastic strains are involved. Under such conditions, it becomes important to use true stress and true strain rather than engineering stress and engineering strain. For many metals, the flow curve, true stress–true plastic strain, can be expressed as

$$\sigma = K\varepsilon^n$$

where n is called the strain-hardening exponent and K is the strength coefficient. A log–log plot of this relationship gives slope n, and $K = \sigma$ at $\varepsilon = 1$. The strain-hardening exponent n can vary between 0 (ideally plastic) and 1 (elastic). For most metals, $0.1 < n < 0.5$. Note that the strain-hardening rate is not the same as the strain-hardening exponent. In fact,

$$n = \frac{d(\log \sigma)}{d(\log \varepsilon)} = \frac{d(\ln \sigma)}{d(\ln \varepsilon)} = \frac{d\sigma}{\sigma}\frac{\varepsilon}{d\varepsilon}$$

or

$$\frac{d\sigma}{d\varepsilon} = n\frac{\sigma}{\varepsilon}$$

At necking, $d\sigma/d\varepsilon = \sigma$, and $\varepsilon = \varepsilon_u$, the true uniform strain. Therefore,

$$n = \varepsilon_u$$

This important result shows that the true strain at the onset of necking is numerically equal to the strain-hardening exponent n. Thus, the higher the value of n, the greater the strain that the material can take before necking will intervene. This information can be very useful in material processing because it tells how much a material will deform plastically in a uniform manner before the onset of necking.

Tensile Behavior of Polymers

In metals and ceramics, one can distinguish between elastic and plastic strains rather easily. In polymers, the elastic and plastic strains are more difficult to distinguish and the stress depends both on the strain and the strain rate, more so than in metals. In metals and ceramics, for moderately high temperatures, the slope of the elastic part, that is, Young's modulus, is fairly constant. Strain rate also has a small effect. In metals, elastic deformation involves stretching of the atomic bonds; hence, it is not very time dependent. In polymers, chains can undergo uncoiling even in the elastic region, that is, elastomeric or rubbery materials can show nonlinear, elastic behavior. In thermoplastic polymers, when necking begins, the stress to continue deformation drops as in metals. However, unlike metals, the neck does not get smaller but grows along the sample; the chains become aligned and, eventually, the stress to continue the deformation increases as secondary bonds form between the chains. **Figure 3** shows schematically tensile stress–strain curves for a variety of polymeric materials. Thermoset polymers such as epoxy show a brittle behavior. Thermoplastic polymers such as polyamide (nylon) or polyethylene show ductile behavior, while elastomers or rubbers show nonlinear, elastic deformation.

Machine–Specimen Interaction

An experimental tensile stress–strain curve is obtained by causing the specimen to interact with a testing

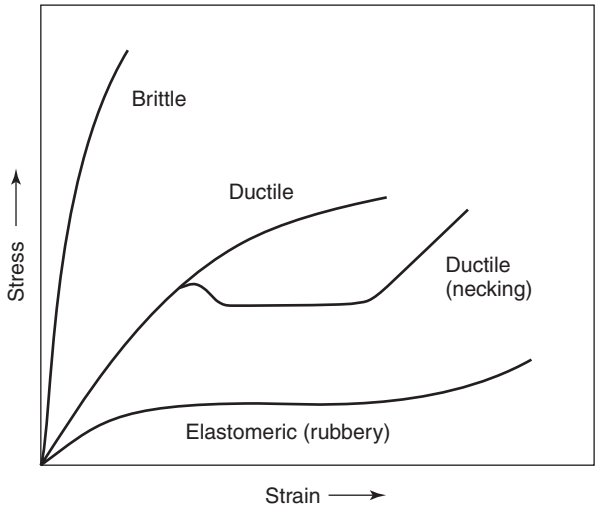


Figure 3 Stress–strain curves for a variety of polymeric materials. Thermoset polymers such as epoxy show a brittle behavior. Thermoplastic polymers such as polyamide (nylon) or polyethylene show ductile behavior, while elastomers or rubbers show nonlinear, elastic deformation.

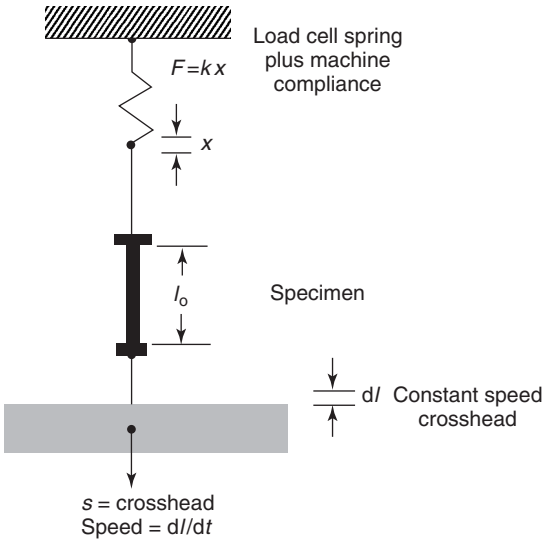


Figure 4 Machine–specimen interaction.

machine. The specimen is connected to a crosshead by means of joints, couplings, grips, etc. The crosshead moves at a constant speed, v, to deform the specimen. However, not all of the crosshead movement is translated into the deformation of the gauge length of the specimen. A part of the crosshead motion goes into the "elastic" strain of the deforming fixture and the specimen (the gauge length and the extra-gauge length portion), and the rest of the crosshead motion goes into the "plastic" deformation of the specimen. **Figure 4** shows a schematic of the specimen–machine linkage.

One can represent the elastic strain of the deforming fixture and sample by the deformation of a spring. The elastic deformation of the spring, Δx_{el}, is given as

$$\Delta x_{el} = F/K$$

where F is the applied force and K is the effective spring constant. Let Δl_p be the plastic deformation of the sample, then, the total crosshead displacement is given as

$$\Delta l_{total} = vt = \Delta x_{el} + \Delta l_p = F/K + \Delta l_p$$

where t is the total time of the crosshead motion and v is the crosshead velocity. Also,

$$\Delta l_{total} = vt = \Delta x_{sp} + \Delta x_m = F/K_m + \Delta x_{sp} \qquad [8]$$

where Δx_{sp} is the total displacement of the specimen gauge length, that is, elastic + plastic, and Δx_m is the elastic displacement of the machine and is equal to F/K_m, K_m being the machine constant. The term "machine" represents everything outside the specimen gauge length. Equation [8] can be written as

$$\Delta x_{sp} = vt - F/K_m$$

A testing machine having a high stiffness, K_m, is called a hard machine, while one having a low K_m is called a soft machine. An expression for the energy stored in the machine can be written as

$$U = \frac{K_m x^2}{2} = \frac{P^2}{2K_m}$$

In the "elastic regime," one can write

$$\Delta l_{total} = \text{machine displacement} + \text{specimen displacement}$$

A load–extension curve of a sample, obtained via crosshead displacement and not via an extensometer, would give force versus Δl_{total}. The slope of this curve in the elastic regime is E^* (the apparent Young's modulus):

$$\Delta l_{total} = vt = F/K_m + \sigma l_o/E$$

$$\frac{F}{A} = \sigma = E^* \frac{\Delta l_{total}}{l_o}$$

$$\frac{1}{E^*} = \frac{1}{\sigma}\frac{\Delta l_{total}}{l_o}$$

Now, an expression for Δl_{total} involving the real Young's modulus, E, can be written as follows:

$$\frac{\Delta l_{total}}{l_o} = \frac{F}{Kl_o} + \frac{\sigma}{E}$$

or

$$\frac{1}{\sigma}\frac{\Delta l_{total}}{l_o} = \frac{F}{Kl_o}\frac{1}{\sigma} + \frac{1}{E}$$

$$\frac{1}{E^*} = \frac{A}{Kl_o} + \frac{1}{E} \qquad [9]$$

Equation [9] is the equation of a straight line ($y = mx + c$), and it relates the real and the apparent Young's moduli. Note that as

$$l_o \to \infty, \quad A/Kl_o \to 0, \quad E^* \to E$$

A plot $1/E^*$ versus $1/l_o$, obtained by testing specimens of different gauge length, yields the true Young's modulus of the material.

Tensile Testing of Brittle Materials

In the elastic deformation of brittle materials, very small strains are involved. Any nonaxial loading or initial curvature can result in substantial error because the extensometer will give signals due to bending as well as tensile strains in the specimen. Use of two strain gauges mounted on the opposite sides of the specimen can help. Under this condition, the average change in length recorded by the two gauges

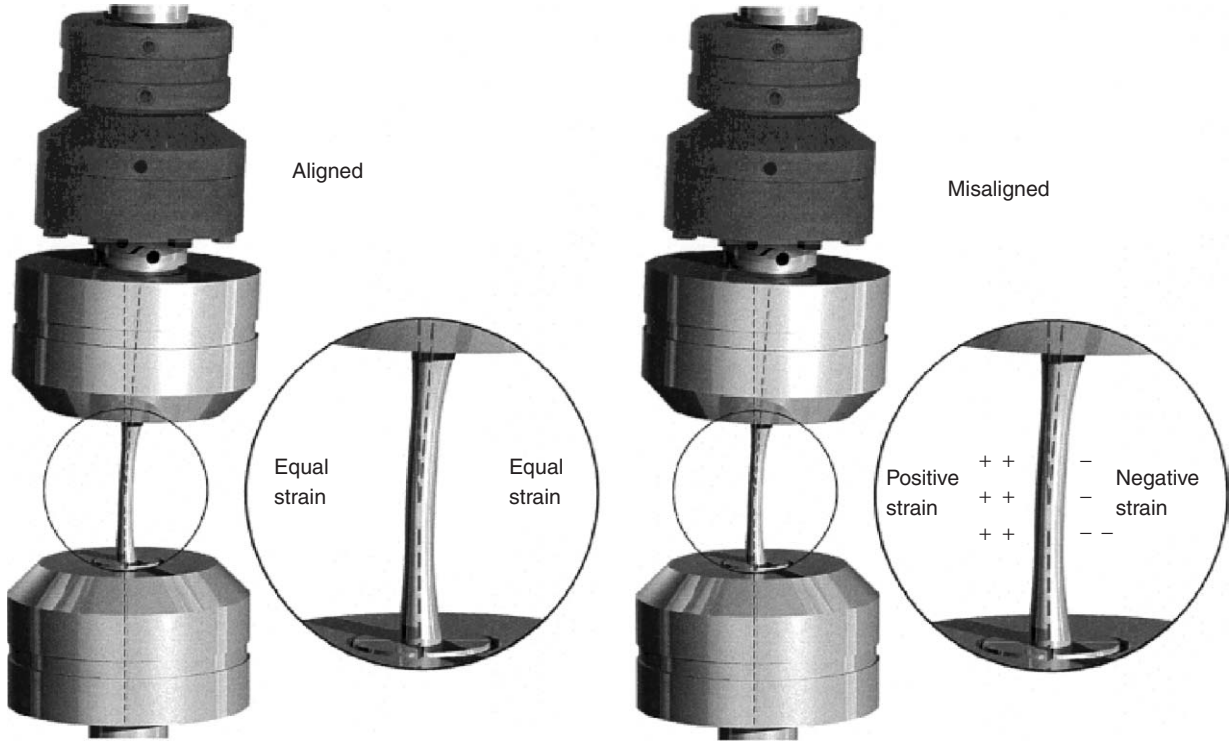


Figure 5 Aligned and misaligned grips. The misaligned grip results in parasitic bending strains in addition to the tensile strain. (Courtesy of MTS Systems Corp.)

will be independent of bending. Thus, axial alignment becomes extremely important for tensile testing of brittle materials such as alumina or silicon carbide. Any nonaxiality will lead to bending of the specimen as shown in **Figure 5**. With ductile metals, nontriaxiality is not a serious problem because the specimen will straighten after the plastic deformation begins.

Tensile testing is widely used to obtain mechanical characteristics of nonbrittle solids such as metals. Brittle materials such as ceramics and glasses are frequently tested in flexure (three-point or four-point). Although flexure tests are easy to perform, results generally do not reflect randomly occurring defects in highly stressed parts, for example, in a turbocharger rotor. In a bend test, only a small material volume of the test piece experiences the maximum load. In a tensile test, on the other hand, the whole gauge length of the specimen is subjected to a uniform stress. This is the reason that the strength values obtained from a bend test are considerably higher than those from a tensile test. In a tensile test, one gets a uniform stress field and a relatively large effective volume of the test material. These are important for evaluating inherent flaw distributions and the resultant strength distributions. One major problem in tensile testing of ceramic and ceramic composites is the bending moments that might be introduced during specimen alignment, gripping, and testing. Such accidental bending moments are sometimes referred to as "parasitic" bending moments. **Figure 5** shows the aligned and misaligned states of a specimen in the grips. Various grip designs have been developed to take care of this problem of parasitic bending moments. One such development involves a hydraulic, self-aligning grip system that results in a condition of "near-zero bending moment" during a tensile test. The specimen grip is self-aligning and has a built-in flexible coupling consisting of eight tiny hydraulic pistons. These pistons reduce bending in the specimen to near zero, $<1\%$ bending in the gauge length. With proper attention to the gripping of the sample, tensile tests on ceramics can be performed at temperatures up to 1300°C without much problem.

Gripping Devices

In tensile testing of any specimen, one of the important items is the type of gripping device to be used to transmit the applied load to the test specimen. This is especially true for tensile testing of brittle materials, which must not be nicked, damaged, or bent during gripping. Any misaligned or nonuniform loading or point and/or line contacts can produce Hertzian-type stresses leading to crack initiation and fracture of the specimen in the gripped region.

Gripping devices are classified as having active or passive grip interfaces. "Active grip interfaces" require a continuous application of a mechanical, hydraulic, or pneumatic force to transmit the load applied by the test machine to the test specimen. Generally, these types of grip interfaces require the load to be applied normal to the surface of the gripped section of the specimen. Transmission of the uniaxial load applied by the test machine is then accomplished via friction between the specimen and the grip faces. Thus, important aspects of active grip interfaces are uniform contact between the gripped section of the specimen and the grip faces, and the constant coefficient of friction over the grip/specimen interface. For cylindrical specimens, a one-piece split-collet design is commonly used. Generally, close dimensional tolerances are required for concentricity of both the grip and specimen diameter. In addition, the diameter of the gripped section of the specimen and the unclamped, open diameter of the grip faces must be within similarly close tolerances in order to promote uniform contact at the specimen/grip interface. For flat specimens, flat-face, wedge-grip faces act as the grip interface. In this case too, close tolerances are required for the flatness and parallelism as well as the wedge angle of the grip faces. In addition, the thickness, flatness, and parallelism of the gripped section of the specimen must be within similarly close tolerances in order to promote uniform contact at the specimen/grip interface. "Passive grip interfaces" transmit the applied load to the test specimen through a direct mechanical link. Generally, these mechanical links transmit the test loads to the specimen via geometrical features of the specimens such as a button-head.

See also: Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Mechanical Properties: Strengthening Mechanisms in Metals; Recovery, Recrystallization, and Grain Growth.

PACS: 51.35.+a; 62.20.−x; 68.35.Gy; 87.15.La

Further Reading

Avery DH and Findley WN (1971) *Techniques of Metals Research*, vol. 5, Pt. 1. New York: Wiley-Interscience.

Davis JR (ed.) (2004) *Tensile Testing*, 2nd edn. Materials Park, OH: ASM Intl.

Hart EW (1967) The theory of tensile test. *Acta Metallurgica* 15: 351.

McGregor Tegart WJ (1964) *Elements of Mechanical Metallurgy*. New York: Macmillan.

Meyers MA and Chawla KK (1984) *Mechanical Metallurgy*. Englewood Cliffs, NJ: Prentice-Hall.

Membrane and Protein Aggregates

O G Mouritsen, University of Southern Denmark, Odense, Denmark

Introduction

Membranes constitute Nature's preferred architecture for nano- and microencapsulation technology of living matter. All cells are bounded by one or several membranes and whereas cells of eubacteria and archaebacteria are bounded by a single plasma membrane, all eukaryotic cells have in addition internal membranes encapsulating the nucleus and the various organelles. Hence, biological membranes are the most abundant cellular structure in living matter.

Membrane assemblies are composed of lipids, proteins, and carbohydrates as illustrated in **Figure 1**. Lipids and many membrane-bound proteins are amphiphilic molecules, and the aqueous membrane system is therefore determined by self-assembly processes that are mainly controlled by the long-range hydrophobic effect. The hydrophobic effect has a substantial entropic component, and details of membrane structure and dynamics are therefore very sensitive to temperature.

The biophysics of membrane and protein assemblies is conveniently studied in model systems consisting of

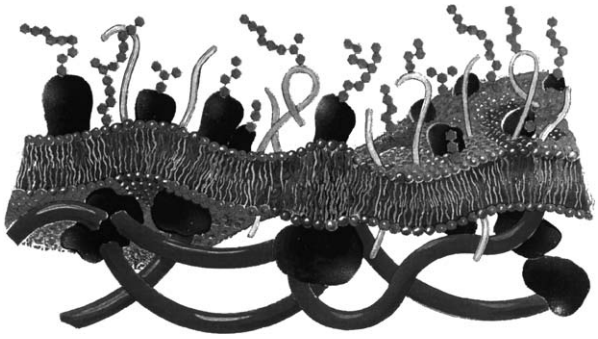


Figure 1 Schematic illustration of the plasma membrane of a eukaryotic cell with a lipid bilayer core incorporated with proteins. On the outside is a carbohydrate coat and on the inside a polymeric cytoskeleton.

bilayers of lipid molecules in which proteins as well as other molecules that interact with membranes are incorporated. Such bilayers can be in the form of free liposomes (or vesicles) in water or in the form of planar bilayers on solid supports as shown in **Figure 2**. Although lipid bilayer membranes are liquids under physiological conditions, the membranes are highly structured interfaces which belong to the class of lyotropic smectic liquid crystals.

Proteins can be associated with lipid bilayers in several different ways, some of which are illustrated in **Figure 3**. In order to understand the interactions between lipids and proteins and how membrane function is controlled by these interactions, it is important first to understand the cooperative properties of lipid bilayers and how these are manifested in terms of the trans-bilayer and the lateral structure of lipid bilayers.

Trans-Membrane Structure and Curvature Stress

The forces that are operative and which stabilize a lipid bilayer, once it is formed, have to be distributed over the $\sim d = 5\,\text{nm}$ thick bilayer sheet. The major force derives from the interfacial tension, γ, which operates between the hydrophilic and the hydrophobic parts of the bilayer as shown in **Figure 4**. The tensile force, which is $\sim 50\,\text{mN}\,\text{m}^{-1}$, is counterbalanced by two repulsive forces, the lipid head-group repulsion and the entropic chain pressure in the core of the bilayer. Since these forces are centered in different planes, a trans-bilayer lateral stress (or pressure) profile, $\pi(z)$ builds up with enormous pressure densities of the order of $2\gamma/d$, that is, typically several hundred atmospheres.

An important determinant of the lateral stress profile is the effective shape of the lipid molecules. Molecules with conical shape will exhibit propensity for forming nonlamellar and curved structures. Each monolayer of the lipid bilayer can hence display a spontaneous curvature which leads to a curvature stress field in a bilayer as illustrated in **Figure 5**. Under appropriate conditions, the bilayer cannot support the resulting strain and hence nonlamellar lipid phases are formed.

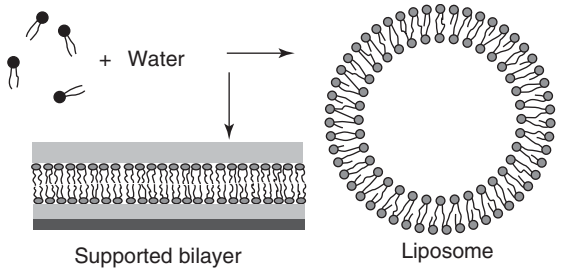


Figure 2 Spontaneous formation of lipid bilayer membranes when lipid molecules are mixed with water.

Lateral Membrane Structure

Membranes are structured laterally on different length and time scales. The structuring is determined

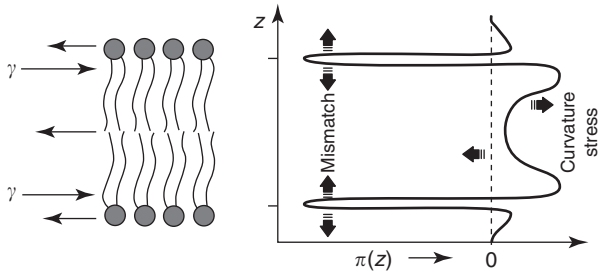


Figure 4 Left: lipid bilayer with indication of the forces that act within the layer. γ is the interfacial tension between the hydrophilic and hydrophobic parts. Right: corresponding lateral pressure profile, $\pi(z)$. The effects of hydrophobic mismatch and curvature stress on the profile are indicated.

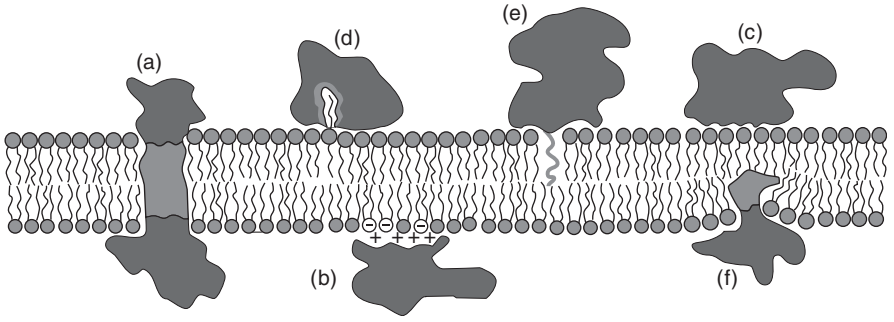


Figure 3 Schematic illustration of different modes of interaction between proteins and lipid membranes: (a) an integral membrane protein that spans the bilayer; (b) electrostatic surface binding; (c) nonspecific binding by weak physical surface forces; (d) anchoring via a lipid extended conformation; (e) anchoring by an acyl-chain anchor attached to the protein; and (f) amphiphilic protein that partially penetrates the bilayer.

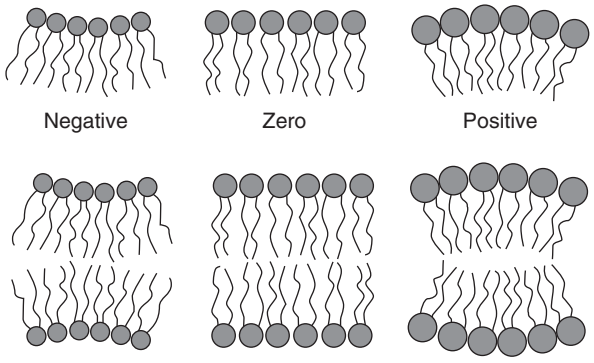


Figure 5 Illustration of the destabilization of a lipid bilayer composed of lipids with conical shapes leading to a curvature stress field, which leads to a tendency for the two monolayers to curve.

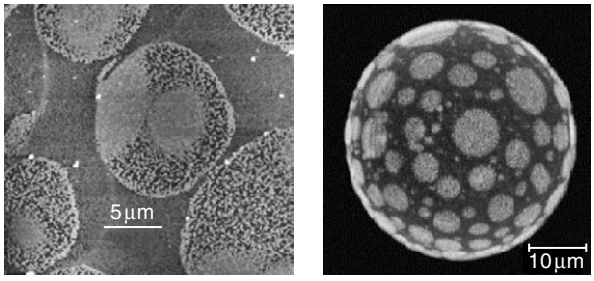


Figure 6 Examples of lateral membrane structure. Lipid domains in native pulmonary surfactant membranes consisting of lipids and proteins. Left: Atomic force microscopy image of a supported bilayer on mica. (Courtesy of Adam C Simonsen.) Right: Fluorescence microscopy image of a giant unilamellar liposome. (Courtesy of Luis Bagatolli.)

by the collective properties of the assembly and not only by the chemical properties of the individual lipids and proteins. Cooperative phenomena and phase equilibria in membranes are manifested in terms of differentiated regions or domains sometimes called rafts which persist on length scales from nanometer to micrometer. There is accumulating evidence that lipid domains support various membrane functions. **Figure 6** shows examples of lipid domains in membranes of pulmonary surfactants in a supported lipid membrane studied by atomic force microscopy and in giant liposomes studied by fluorescence microscopy. Both images refer to membranes with integral membrane proteins and it is seen that lateral structure may arise on different length scales.

The physical interactions between the molecules in a lipid membrane determine a so-called persistence (or coherence) length, ξ_L, which is a measure of the effective range over which the interactions operate and lead to structure and fluctuations in structure. The actual value of the persistence length is not only determined by the molecular interactions but it is also dependent on thermodynamic conditions and composition. The persistence length is a measure of the lipid cooperativity (or fluctuations) which in turn is the underlying source of the domain formation. Roughly, the persistence length is the average size of the domains. The range over which different proteins can "feel" each other through the lipid bilayer is also set by ξ_L. Hence, the small (nanometer) scale structure of membranes is intimately related to lipid–protein interactions. The extension, ξ_P, of the annulus of lipid molecules around a protein which is perturbed by the protein is proportional to ξ_L. Hence the lateral organization of proteins in a lipid bilayer can to some extent be modulated by altering the persistence length, for example, by changing temperature or by adding specific substances, such as drugs, that will change ξ_L.

Lipid–Protein Interactions

All integral membrane proteins have a similar structural motif with a hydrophobic domain, α-helical or β-sheet, that traverses the hydrophobic core of the lipid-bilayer membrane as illustrated in **Figure 7**. The number of membrane-spanning parts of the amino acid sequence ranges from a single one to more than 12. The hydrophobic domain contains some of the evolutionary most conserved sequences found in proteins. This suggests that the trans-bilayer coupling between lipids and proteins may be an ancient mechanism that has generic as well as fundamental implications for membrane protein interactions and hence for biological function.

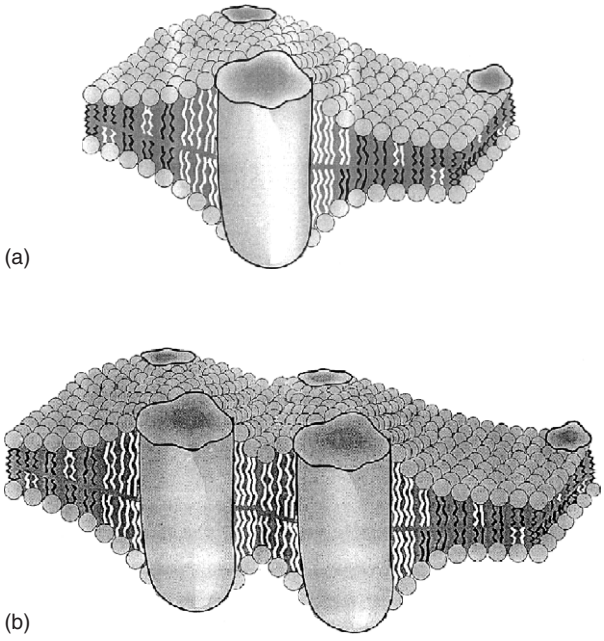
The physical constraint imposed on integral membrane proteins by the lipid bilayer thickness suggests that a mechanical hydrophobic matching principle may be operative. Hydrophobic matching means that the hydrophobic length, d_P, of the trans-membrane domain is matched to the hydrophobic thickness, d_L, of the lipid bilayer as illustrated in **Figure 8**. In order to compensate for a possible mismatch, the soft lipid bilayer yields, and the lipid molecules closest to the protein stretch out or compress in order to cover the hydrophobic core of the protein. This leads to a perturbed region around the protein over a range ξ_P.

The radial structure of the perturbed region of the lipids can be estimated by considering a single protein in a membrane and neglecting the entropy of mixing of lipids and proteins. The free energy of the system can be written in terms of a free-energy density g as

$$G = \int \mathrm{d}^d r \, g\left[d_L(r), \frac{\mathrm{d}}{\mathrm{d}_r}\{d_L(r)\}\right]$$

where r is the radial distance from the protein. Within the Gaussian approximation for g, minimization of G in the case of a relaxed boundary condition leads to

Figure 7 Schematic illustration of a trans-membrane protein (bacteriorhodopsin) imbedded in a lipid bilayer.



(a)

(b)

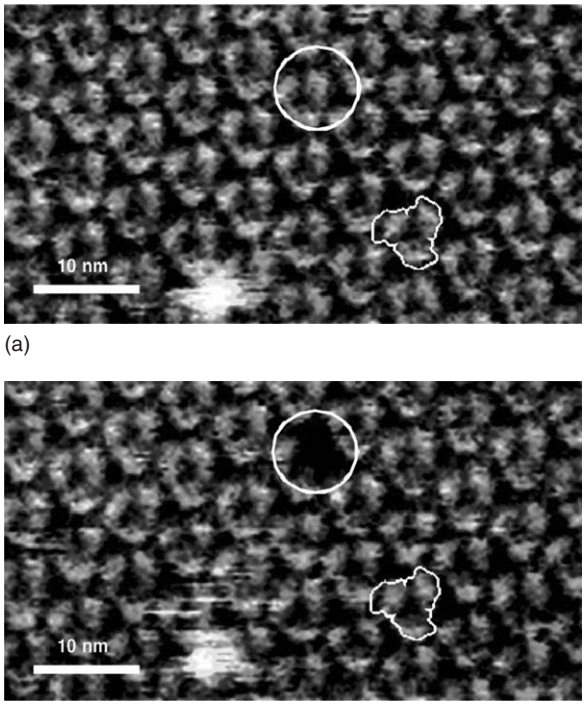
Figure 8 Schematic illustration of integral membrane proteins embedded in a lipid bilayer that is hydrophobically thinner than the hydrophobic domain of the protein. (a) Hydrophobic matching implies formation of a region of lipids around the protein that conform to the hydrophobic matching condition either by stretching the acyl chains or by recruiting the better-matched lipid species at the interface. (b) Hydrophobic matching implies a lipid-mediated mechanism for protein attraction and aggregation in the plane of the membrane.

the profile

$$d_{\mathrm{L}} = d_{\mathrm{P}} + (d_{\mathrm{L}}^{\mathrm{o}} - d_{\mathrm{L}})\exp(-r/\xi_{\mathrm{P}})$$

where $d_{\mathrm{L}}^{\mathrm{o}}$ is the hydrophobic thickness of the unperturbed lipid bilayer. A typical size of ξ_{P} lies in the range of one to a few nanometers.

Hydrophobic mismatch and the possibly imposed local curvature stress may therefore be a controlling mechanism for the way proteins interact with lipids in membranes. This mechanism could also be the basis for lipid sorting at the lipid–protein interface as illustrated in **Figure 8a** in the case of a membrane



(a)

(b)

Figure 9 Two-dimensional crystalline array of the integral membrane protein bacteriorhodopsin in a lipid-bilayer membrane. The figure illustrates how a single bacteriorhodopsin molecule (a) sitting in a two-dimensional crystal can be pulled on by the tip of an atomic force microscope leaving a hole in the crystal inside the circle in (b). (Courtesy of Daniel Müller.)

that is too thin to accommodate the protein. For a lipid bilayer with a single lipid species, this perturbed interface region in **Figure 8a** is characterized by a larger average lipid-bilayer thickness and a higher conformational chain order. Since lipid bilayers and membranes under physiological conditions are liquids, the perturbed region is a statistical entity in the sense that lipids diffuse in and out of the region. Hydrophobic matching provides a physical mechanism for protein organization in membranes. This is illustrated in **Figure 8b** in the case of two proteins. In that case, if the distance between the proteins becomes comparable to the persistence length, the annuli of the proteins overlap which tends to lower the free energy and hence facilitates an effective lipid-mediated attractive protein–protein interaction. An example of a two-dimensional crystalline aggregate formed by integral membrane proteins in a lipid membrane is shown in **Figure 9**.

Physical Triggering of Proteins

Figure 10 shows how a conformational change in an integral membrane may be triggered by hydrophobic matching condition.

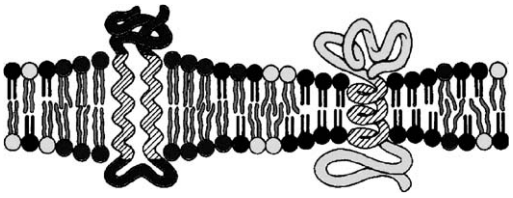


Figure 10 Schematic illustration of triggering the function of an integral membrane protein by changing the hydrophobic mismatch. (Adapted from Sackmann E (1984) Physical basis for trigger processes and membrane structures. In Chapman D (ed.) *Biological Membranes*, vol. 5, pp. 105–143. London: Academic Press, with permission.)

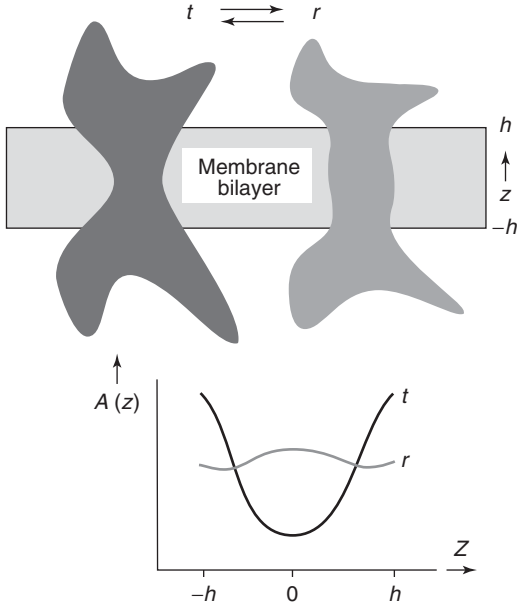


Figure 11 Schematic illustration of the change in cross-sectional area profile, $A(z)$, of an integral membrane protein that undergoes a shape change during a conformational transition from state *t* to state *s*. The area profile change couples to the lateral pressure profile, $\pi(z)$, of a lipid bilayer shown in **Figure 4**. (Courtesy of Robert S Cantor.)

It may well be that it is not the hydrophobic thickness as such that provides the direct mechanic coupling between the lipid bilayer and the protein, but rather some other membrane property which is related to the hydrophobic thickness. The lateral pressure profile shown in **Figure 4** is the most obvious candidate. Changes in this profile can lead to conformational changes in the protein as illustrated in **Figure 11**, provided the protein has a noncylindrical shape, for example, as the hourglass form of a helix bundle.

The work

$$W = -\int_z \pi(z)[A_t - A_r]\mathrm{d}z$$

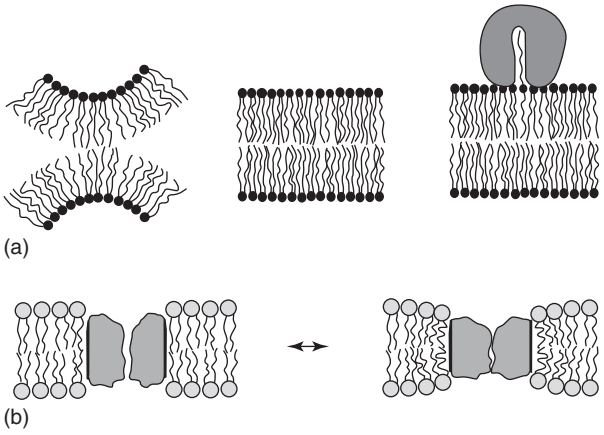


Figure 12 Schematic illustrations of the influence of lipid-bilayer curvature stress on protein function. (a) From left to right, a hypothetical situation of two relaxed lipid monolayers with intrinsic curvature leading to a lipid bilayer with built-in curvature stress which may be partially released upon binding to a peripheral protein via the extended chain-conformation mechanism. (b) The effect of curvature stress on the opening of a membrane channel. (Courtesy of Kinnunen PKJ and Andersen OS.)

that is involved in such a change can be large enough to induce a conformational change in the protein. This mechanism suggests a way of manipulating membrane proteins and their function by changing the hydrophobic thickness of a membrane or by changing the lateral pressure profile, for example, by shifting the balance of the pressure between the head-group and acyl-chain part of the bilayer. A number of drugs may have their action by acting directly on the lipids, changing hydrophobic thickness or pressure profile, and thereby indirectly couple to the protein and its function. This mechanism has been suggested to be operative in the case of general anesthetics.

Similarly, the curvature stress and the way it can be released provide a mechanism for triggering of protein binding to membrane surfaces and for the opening and closing of membrane channels as illustrated in **Figure 12b**.

Lipid-Mediated Protein Function

There is a substantial amount of experimental evidence that indirectly points to the presence of hydrophobic matching in reconstituted membrane–protein systems and in real cell membranes. The degree of matching influences the functioning of membrane channels, pumps, and transporters, as well as how integral membrane proteins are inserted into, secreted through, and folded within the membrane. For example, it has been found for a number of membrane channels, ion pumps, and sugar transporters that they, when incorporated into lipid bilayers of different thickness, function optimally for a certain

narrow range of thicknesses, where they presumably are hydrophobically well matched. Thickness alterations induced internally or by external stimuli may therefore be seen as a way of triggering these proteins to enhance or suppress their function as pictured schematically in **Figure 10**. A couple of specific examples are described below.

The first example is the integral membrane protein melibiose permease which is a cation/sugar symporter from *E. coli*. This protein catalyzes cell accumulation of α-galactosides such as melibiose. When incorporated into membranes of different hydrophobic thickness, it is found that the function of melibiose permease is at an optimum for a particular bilayer thickness. Another example refers to **Figure 13** which shows how the activity of the integral membrane protein Na,K–ATPase varies with membrane thickness. The figure shows that the activity of this ion pump is maximal for a certain type of lipid and hence for a specific membrane thickness. If cholesterol is added to the membrane, the maximum is seen to move toward lipid membranes made of shorter lipids. This can be rationalized via the hydrophobic matching principle, recalling that cholesterol tends to thicken fluid membranes which in turn will compensate for the shorter lipids.

This latter observation suggests a more general principle to be operative by which cholesterol may be used as a regulator of membrane function and the sorting and targeting of proteins, possibly via hydrophobic matching. The following serves as an illustration.

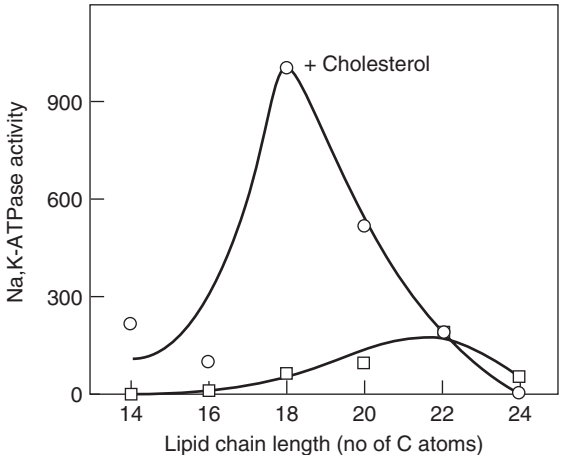


Figure 13 Schematic illustration of the dependence of protein function of the average hydrophobic length of the lipid molecules. An optimum occurs at a given length corresponding to a specific bilayer thickness. After incorporation of cholesterol, which acts as to thicken fluid lipid membranes, the optimum is shifted toward shorter lipids, rendering the actual membrane thickness at the optimum the same as without cholesterol. (Courtesy of F Cornelius.)

Proteins are synthesized at the ribosomes placed in the endoplasmatic reticulum. From there they are transported via the Golgi to the various parts of the cell where they belong, for example, in the plasma membrane. This transport, which is referred to as the secretory pathway, requires a sorting of the proteins which is performed in the Golgi. Some proteins carry specific tags that will actively target them to their destination, others will passively flow through the cell. The sorting along the secretory pathway may be performed by means of a gradient in the hydrophobic thickness of the membrane systems which the proteins have to pass on their way to their target. Indeed, the membrane contents of cholesterol and sphingomyelin, which both tend to enlarge membrane thickness, are found to increase going from the endoplasmatic reticulum, via the Golgi, to the plasma membrane.

A substantial part of the lipids found in natural membranes, for example, phosphatidylethanolamine lipids, have propensity for nonlamellar structures and therefore lead to a curvature stress and an instability which may be locally released in connection with protein binding, protein insertion, membrane fusion, and conformational changes in the cycle of protein functions. As an example, protein kinase C is one of the most important enzymes, which are involved in the signal transduction system of the cell. Upon stimulation of the cell by, for example, neurotransmitters, hormones, and growth factors, protein kinase C becomes activated upon binding to the plasma membrane, leading to a complicated cascade of biochemical signals that eventually influence cell growth, cell differentiation, as well as exocytosis. A requirement for binding to the membrane and hence for activation of the enzyme is lipids with acidic head groups, such as phosphatidylserine, and the presence of calcium ions. Calcium ions require water for solvation, and in the competition for water the membrane surface becomes dehydrated leading to a larger curvature stress. This curvature stress could be released if some lipid molecules assume an extended chain conformation as illustrated in **Figure 12a** by flipping one of the acyl chains to the outside of the membrane. This flip would normally be energetically very costly because of the hydrophobic effect. However, if the chain can be accommodated in a putative hydrophobic crevice in a protein it will both release the curvature stress of the membrane and simultaneously facilitate the membrane anchoring and hence the activation of the enzyme. Protein kinase C has such a hydrophobic crevice. A similar mechanism has been shown to facilitate binding of cytochrome *c* to membranes.

The integral membrane protein rhodopsin is a seven-helix trans-membrane protein. Rhodopsin is the light-sensitive protein in the visual pigment of our

retina which, upon activation of light, initiates the signaling pathway that eventually leads to vision. An essential stage of this process involves a certain transition between two conformational states of rhodopsin, the so-called M-I and M-II states. The M-II state is believed to correspond to a more elongated form of rhodopsin than the M-I state. The transition therefore implies a change in hydrophobic mismatch. Studies have shown that the M-II state requires the presence of lipids that have propensity for forming nonlamellar phases. The retinal rod outer segment membranes, in which rhodopsin exhibits its function, are known to have almost 50% of the polyunsaturated fatty acid docosahexaenoic acid which, due to the many double bonds, indeed supports curved structures.

There is a particular class of membrane proteins, the mechanosensitive channels, that have evolved to facilitate ion conductance in response to a stress exerted by the membrane. These proteins are nanomachines, which work as transducers of mechanical strain dissipated from the membrane. The most well-studied example is the bacterial large conductance mechanosensitive channel (MscL) from *E. coli*. Hydrophobic matching and curvature stress concepts can be used to interpret the experimental data for the conductance of MscL and how it varies when nonlamellar lipid species are incorporated into the membrane. MscL is a helix bundle protein and experiments as well as molecular dynamics simulations have supported a mechanism for channel opening that involves an iris-like expansion of the conducting pore.

Finally, evidence is available for insertion of peptides and proteins into lipid bilayers being controlled by hydrophobic matching. Proteins are more readily inserted and correctly folded into membranes to which their hydrophobic domain is hydrophobically matched. It is the same mechanism that may be in operation as the molecular sieve in the secretory pathway.

See also: Biomembranes; Ion Transport in Biological Membranes.

PACS: 87.15.Kg; 87.16.Dg; 87.14.Cc; 87.14.Ee

Further Reading

Gennis RB (1989) Biomembranes. *Molecular Structure and Function*. Berlin: Springer.

Israelachvili I (1992) *Intermolecular and Surface Forces*, 2nd edn. London: Academic Press.

Lipowsky R and Sackmann E (eds.) (1995) *Handbook of Biological Physics. vols. IA and IB: Structure and Dynamics of Membranes. From Cells to Vesicles*. Amsterdam: Elsevier

Mouritsen OG (2005) *Life – As a Matter of Fat*. Heidelberg: Springer.

Memory Devices, Nonvolatile

G Baccarani and E Gnani, University of Bologna, Bologna, Italy

Introduction

The key attribute of nonvolatile memory devices is their ability to retain the stored information for several years both under ON and OFF power conditions. Owing to this property, they have found widespread use in a variety of professional systems, such as computers, phone exchanges, measurement setups, robotic and control systems, automotive and medical electronics systems, as well as commodities and consumer products, such as cell phones, personal digital assistants (PDAs), digital video recorders and still cameras, digital TV (DTV), set top boxes, GPS-based navigation systems, toys, radio-frequency identification devices (RFIDs), and smart cards.

To date, the nonvolatile memory market has become a sizeable fraction of the whole memory market, and is largely dominated by the flash electrically erasable and programmable read-only memory (EEPROM) devices. Owing to the pervasiveness of flash memories, the problem of integration of flash arrays within logic circuits has been thoroughly investigated, and the problems associated with the harmonization of technological processes which are not fully compatible have been successfully tackled, thus making it possible to produce systems on chip with embedded flash arrays.

This article is organized as follows: the next section introduces the classification and illustrates the general features of nonvolatile memories. In the section "Read-Only Memory" (ROM), the ROM circuit functions are illustrated. The section "Electrically-Programmable Read Only Memory" (EPROM) addresses the structure and operation principles of the EPROM cell, along with its main working issue, namely: reliability. The following section is devoted to the description of the structural and functional properties of the EEPROM. In the next section, the

architectural options and the functional properties of the flash memory arrays are discussed along with reliability issues. The final section addresses a number of new memory concepts which are still within a development stage, and have not emerged as real products yet, such as the magnetic RAM (MRAM), the ferroelectric RAM (FeRAM), and phase-change memory.

Nonvolatile Memory Classification

Nonvolatile semiconductor memories are generally classified according to their functional properties with respect to the programming and erasing operations, as shown in the flow chart depicted in **Figure 1**.

ROMs are programmed at the mask level during the fabrication process by suitably interconnecting logic gates in order to generate appropriate logic functions of the input data. Therefore, the reading operation is carried out by a logic computation using combinational logic circuits, which can be manufactured by a standard CMOS process. ROMs, however, suffer a lack of flexibility, which dramatically restricts their application range.

EPROMs are based, instead, on a different physical principle, and store information by controlling the threshold voltage of a special MOSFET structure with two overlapping gates, the lower of which is electrically floating. Injecting electrons into the floating gate changes the conduction threshold of the cell. A low threshold voltage characterizes a conductive cell, and is associated with a logical "1," while a high threshold voltage represents a logical "0." The natural threshold of a virgin cell is low since the floating gate is electrically neutral; hence it automatically stores a logical "1." Writing (or programming) the cell, by which its content is turned into a logical "0," means injecting electrons into the floating gate. The physical mechanism by which electrons can be injected into the floating gate is typically channel hot-electron injection over the $Si–SiO_2$ energy barrier.

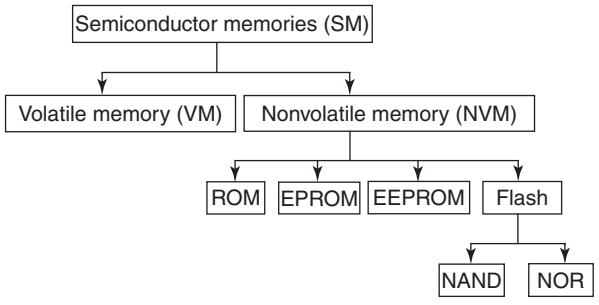


Figure 1 Nonvolatile memory classification according to functional criteria.

EPROMs can be electrically programmed on field and be optically erased by exposing the cell array to ultraviolet (UV) light. Erasing the memory, however, requires a long exposure of the cell array (30 min or so) so that electrons confined within the floating gate can be emitted by a photoelectric effect. This requires a special package with a cover transparent to UV light.

EEPROMs can be electrically programmed and erased on field, which makes them more flexible than EPROMs. This, however, occurs at the expense of a larger cell, which contains a selection transistor in addition to the EEPROM device. For a given silicon area, this larger cell size implies a smaller capacity and a higher cost per bit.

Flash EEPROMs combine, instead, the EPROM's cell size with the EEPROM's flexibility and are thus largely preferred in new applications, such as cell phones, PDAs, and smart cards. In flash memories, all the cells of a sector are erased in parallel, which justifies their "flash" name, and reprogrammed afterwards. This makes the long erasure time (several tenths of a second) acceptable at the system level.

Read-Only Memory

ROMs functionally behave as encoders, for they associate an M-bit output word, which is the memory content, to an N-bit input word representing a physical address. The ROM output can always be expressed as a sum of products of the input bits. Therefore, the reading process occurs in two steps: first, a decoder which generates the min-terms of the N-input bits, raises one of the 2^N word lines high; next, an encoder generates the output word by OR-ing the min-terms from the selection of the physical address. The output word is thus computed by activating a combinational logic circuit; hence, the content of the memory is defined at the mask level during the fabrication process, and no programming changes are possible afterwards.

In view of their functional task, the decoder is often called the AND plane and the encoder is referred to as the OR plane of the ROM. Actually, the logic gates used to generate both the OR and the AND planes are often NOR gates because of speed considerations (**Figure 2**). NAND gates can sometimes be used for small ROMs with a density advantage, but are unpractical for a number of input bits in excess of 4 to 8.

Figure 2a illustrates the typical structure of the AND plane for the simple case of a two-bit structure. The *p*-channel MOSFETs with grounded gate act as pull-up transistors of pseudo *n*-MOS logic gates. The *n*-channel MOSFETs are connected in a NOR configuration. Indicating the input bits with A and B

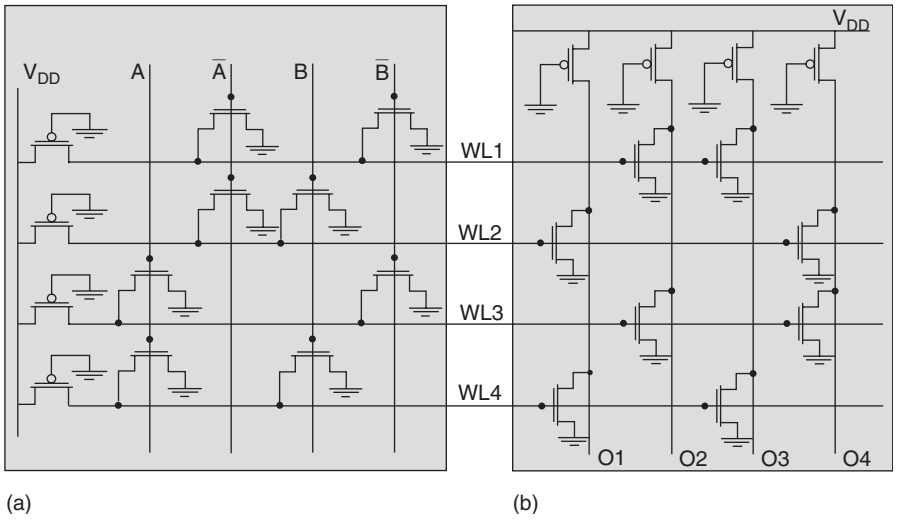


Figure 2 NOR-based implementation of a ROM: (a) AND plane; (b) OR plane.

and their complements with A′ and B′, the output nodes of the AND plane, referred to as $WL_n (n = 1, \ldots, 4)$ turn out to be

$$WL_1 = (A' + B')' = AB$$
$$WL_2 = (A' + B)' = AB'$$
$$WL_3 = (A + B')' = A'B$$
$$WL_4 = (A + B)' = A'B'$$

where De Morgan's theorem has been applied. Thus, the min-terms of the two input bits are generated by NOR operations.

The OR plane, again implemented by NOR gates possibly followed by inverters, is shown in **Figure 2b**. The latter realizes the following Boolean equations

$$O_1 = WL_1 + WL_3 = (WL_2 + WL_4)'$$
$$O_2 = WL_2 + WL_4 = (WL_1 + WL_3)'$$
$$O_3 = WL_2 + WL_3 = (WL_1 + WL_4)'$$
$$O_4 = WL_1 + WL_4 = (WL_2 + WL_3)'$$

Thus, the truth table implemented by the ROM in **Figure 2** is the following:

A	*B*	*WL*$_1$	*WL*$_2$	*WL*$_3$	*WL*$_4$	*O*$_1$	*O*$_2$	*O*$_3$	*O*$_4$
0	0	0	0	0	1	0	1	0	1
0	1	0	0	1	0	1	0	1	0
1	0	0	1	0	0	0	1	1	0
1	1	1	0	0	0	1	0	0	1

A similar procedure can be followed using a NAND implementation of the above Boolean equations.

Electrically-Programmable Read Only Memory

EPROMs rely on a cell structure with two overlapping gates, as shown in **Figure 3**. The lower gate is fully surrounded by silicon dioxide and is thus electrically floating. The upper gate is instead accessible and is referred to as control gate. Due to the need of two polycrystalline silicon levels, the EPROM fabrication technology is not fully compatible with the standard CMOS technology.

Electrons traveling from source to drain under large gate and drain voltages can possibly achieve an energy exceeding the Si–SiO_2 potential barrier, which is $\sim$3.1 eV. If such an electron undergoes a redirecting collision with a final momentum normal to the silicon surface, it can overcome the potential barrier and be injected into the oxide conduction band. If the floating-gate voltage V_{FG} exceeds the drain voltage V_D, the field in the oxide will favor electron injection into the floating gate, where it gets trapped. As more and more electrons are injected and trapped into the floating gate, its potential V_{FG} is lowered and the threshold voltage of the control gate V_{TC} increases.

The basic relationship connecting the floating-gate voltage V_{FG} with the charge injected into the floating gate Q_{inj} and the cell voltages is

$$V_{FG} = \alpha_G V_{CG} + \alpha_S V_S + \alpha_D V_D + \alpha_B (V_B + \Phi_{MS}) + Q_{inj}/C_T$$

where V_{CG} is the control-gate voltage, V_S is the source voltage, V_D is the drain voltage, V_B is the bulk voltage, and Φ_{MS} is the work function difference between the heavily doped n^+ polycrystalline silicon gate and the underlying p-doped silicon substrate. Also, α_G, α_S, α_D, and α_B are coupling coefficients

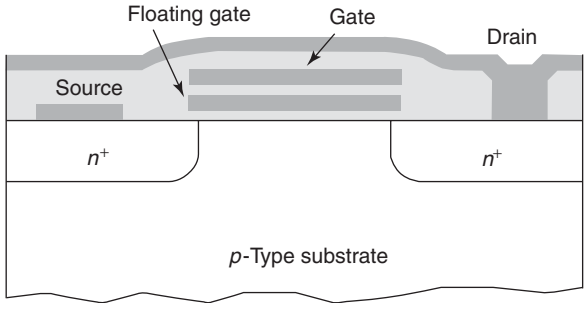


Figure 3 Cross section of a floating-gate EPROM cell.

expressed by the ratios of the gate, source, drain, and bulk capacitances C_{GG}, C_{GS}, C_{GD}, and C_{GB} with respect to the total capacitance of the floating gate $C_T = C_{GG} + C_{GS} + C_{GD} + C_{GB}$. The above equation shows that V_{FG} is a linear combination of the external voltages, with weighting factors given by the coupling coefficients, and that a negative charge Q_{inj} lowers the floating-gate voltage.

When V_{FG} equals the threshold voltage for current conduction within the EPROM cell V_{TF}, the control-gate threshold voltage V_{TC}, reads:

$$V_{TC} = (1/\alpha_G)V_{TF} - (\alpha_S/\alpha_G)V_S - (\alpha_D/\alpha_G)V_D - (\alpha_B/\alpha_G)(V_B + \Phi_{MS}) - Q_{inj}/C_{GG}$$

It may be seen from the above equation that the threshold voltage experienced by the control gate increases if electrons are injected into the floating gate. Strictly speaking, the gate-source, gate-drain, and gate-bulk capacitances are nonlinear; thus, the coupling factors are bias dependent. Quite often, however, such factors may be treated as constants. Furthermore, if $V_S = V_D = V_B = 0$, the above equation may be simplified to give

$$V_{TC} = (1/\alpha_G)V_{TF} - (\alpha_B/\alpha_G)\Phi_{MS} - Q_{inj}/C_{GG}$$

The channel hot-electron (CHE) injection methodology requires both a large gate and a large drain voltage to be applied to the EPROM cell. Thus, a large current flows while programming, with substantial power dissipation. This process is rather inefficient, since only a small fraction of the electrons moving from the source to the drain is indeed injected into the floating gate, and the required time to program the cell may be as large as a few hundred microseconds. After being injected into the floating gate, electrons lose their kinetic energy by emitting phonons, and become permanently trapped within a deep (3.1 eV) potential well.

As already pointed out, native cells are conductive when the control-gate voltage is at the high level; hence they are naturally programmed with a logical "1." Charge injection shifts the threshold voltage above 5 V, and the cell exhibits an off-state when read out. Thus, the high-voltage threshold indicates a logical "0."

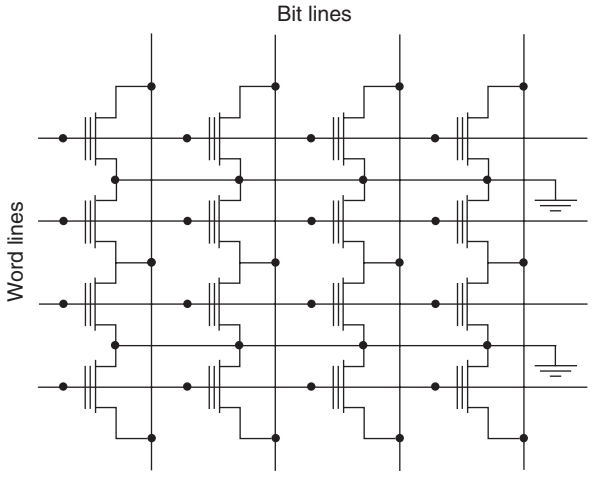


Figure 4 EPROM cell array organization.

The success of floating-gate EPROMs is largely due to the small cell size and the resulting high bit capacity, as well as to the on-field programming ability. Erasing the content of the memory can only occur by exposing the device to UV light for quite some time. This need requires a special package with a cover transparent to UV light. Electrons absorbing a high-energy photon by photoelectric effect can reach an energy exceeding the $Si–SiO_2$ potential barrier, and be ejected either into the control gate or into the silicon substrate.

The memory cell only contains a floating-gate device with its source connected to the ground, its drain connected to the bit line, and its control gate connected to the word line, as shown in **Figure 4**.

When the cell is addressed, the word line is raised high, and the bit line is set by the sense amplifier to an intermediate voltage of ~ 1 V. If a substantial current flows across the cell, the latter is detected to contain a logical "1." Otherwise, it will contain a logical "0."

EPROM memories must typically ensure a data retention of 10 years. This implies that the gate oxide and the inter-poly dielectric, which is often made by a silicon nitride (Si_3N_4) layer sandwiched by two silicon dioxide layers and is referred to as ONO, behave as ideal insulators. Besides, leakage currents must be very small even under the influence of external voltages. When a cell is programmed, all cells connected to the same word line are subject to a large positive gate voltage.

Hence, electrons can possibly tunnel into the floating gate from the channel, if the cell is in its conductive

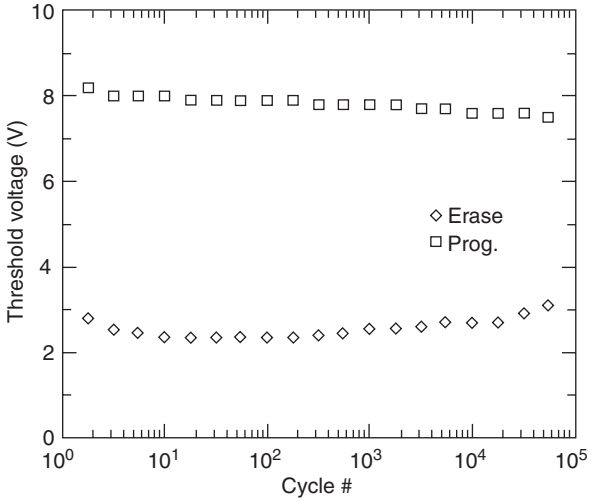


Figure 5 Aging effect on the threshold voltages of EPROM cells.

state, or into the control gate from the floating gate if the cell is in the off-state. In the former case, the gate disturbance tends to convert a logical "1" into a logical "0" while, in the latter, the opposite occurs. Likewise, all cells connected to the same bit line experience a relatively large drain voltage and, due to the low gate voltage, electron tunneling can occur from the floating gate into the drain. In this case, the disturbance generated by cell programming tends to convert a logical "0" into a logical "1."

A read disturbance additionally occurs when a cell is read out, despite the relatively smaller gate and drain voltages. The high drain voltage heats up electrons traveling from the source to the drain of the addressed cell. Hence, some hot-electron injection into the floating gate may occur in this case too. Due to the unlimited number of readings of the cell content, the drain disturbance must be carefully controlled. In order to prevent this effect, the drain is usually kept by the sense circuitry at a voltage ~ 1 V, well below the energy threshold for hot-electron injection.

The cumulative effects of gate and drain disturbance during write and read operations tend to reduce the high threshold voltage of the cells storing a logical "0," and to increase the low threshold voltage of the cells storing a logical "1." Thus, the aging effect manifests itself with a closing window between the high and low threshold voltages, as shown in **Figure 5**.

Electrically Erasable and Programmable Read-Only Memory

EEPROMs provide a further improvement over the functionality of EPROMs by allowing the information

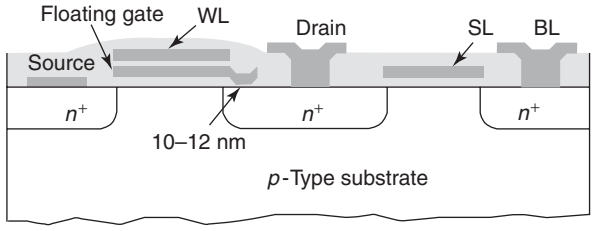


Figure 6 Schematic cross section of a floating-gate EEPROM (FLOTOX) cell.

to be electrically written and erased at the bit level. The basic cell differs from the EPROM cell in that the floating gate extends over the drain region, where a small oxide area is thinned down to 10–12 nm to allow for electron tunneling into and out of the floating gate. This device, shown in **Figure 6**, is referred to as FLOTOX.

The physical mechanism used to program and erase the EEPROM cell is electron tunneling across the thin oxide region between the floating gate and the drain. In order to let electrons tunnel into the floating gate, the control gate must be biased at a large positive voltage, while the drain is kept at zero voltage.

Likewise, cell erasing is carried out by applying a large negative voltage to the control gate, while keeping the drain at zero or positive voltage. In doing so, a large oxide field close to 10 MV cm^{-1} is applied to the tunnel oxide. The large field shapes the energy barrier to become triangular and reduces the tunneling distance. Fowler–Nordheim tunneling then occurs and either fills or depletes the floating gate of the formerly injected charge, thus raising the threshold voltage above V_{DD} or lowering it down below zero.

The poor threshold voltage control of an erased device, however, requires that the basic memory cell contains an additional MOSFET, referred to as selection transistor, connected in series with the FLOTOX device, as shown in **Figure 7**. The drain of the selection transistor is connected to the bit line BL, its gate is connected to the select line SL, and the gate of the FLOTOX device is connected to the word line WL. When the cell is not addressed, the select line SL is at zero voltage, which ensures that no leakage current flows even when the floating-gate device has a negative threshold voltage.

When the memory cell is addressed, both the word line WL and the selection line SL are raised at V_{DD}, thus allowing current to flow across the cell if the FLOTOX device has a low threshold voltage. Otherwise, no current flows, and the sense amplifier will detect a logical "0."

In order to program the floating-gate device, the word line is raised at the large positive voltage V_{PP} and the select line is raised at V_{DD}, while keeping the

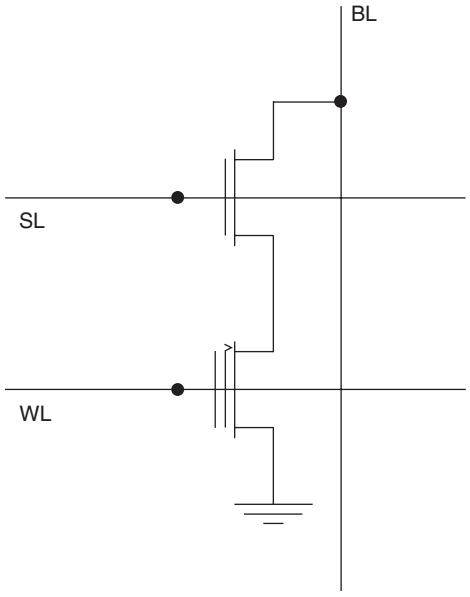


Figure 7 Circuit schematics of the EEPROM cell array.

bit line at ground potential. The V_{PP} voltage is chosen so that electron tunneling may occur into the floating gate. When the cell must be erased, the word line is negatively biased in order to generate the appropriate voltage across the tunnel oxide; the selection line is positively biased, so that the bit line can set the drain voltage of the floating-gate device at the appropriate level. Due to the small tunneling currents, several hundred microseconds may be needed to complete an erasure cycle.

The repeated electron tunneling across the thin oxide is responsible for reliability problems due to electron trapping within the oxide itself. This aging process deteriorates the electron mobility due to added Coulomb scattering and, most importantly, deteriorates the oxide integrity with resulting data retention problems.

Flash EEPROM

Flash EEPROMs are electrically erasable and programmable read-only memories. The basic flash cell is structurally similar to the EPROMs and programming can still be carried out by channel hot-electron injection into the floating gate. However, the cell can be electrically erased by Fowler–Nordheim tunneling through the gate oxide. As opposed to EEPROMs, however, erasing is only carried out in parallel at the sector level, and is followed by reprogramming each individual cell in order to set its threshold voltage at the appropriate high- or low-voltage levels. Therefore, the flash cell does not require the selection transistor for correct operation. An electron micrograph of a flash cell referred to as ETOX is shown in **Figure 8**.

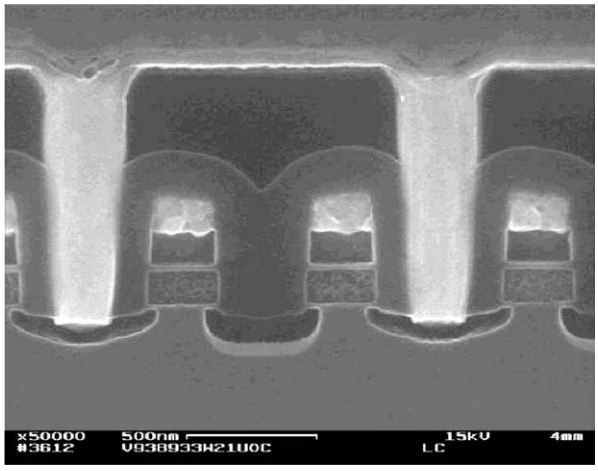


Figure 8 Scanning electron micrograph of the ETOX flash cell in 0.18 μm technology. (Courtesy of ST Microelectronics.)

The cross section is cut along the bit line, which is connected to the drain of the cell by a vertical tungsten plug. The tungsten silicide word line is perpendicular to the cross section, and appears in the micrograph as the light rectangular area overlapping the control gate of the cell. The polycrystalline silicon gates are self-aligned; the source and drain regions are shared by two neighbor mirrored cells. The lateral cell extension, measured between the source and drain mid-axes, is ~ 500 nm in this technology.

Cell programming requires a supply voltage $V_{PP} \approx 10$–12 V to be applied to the control gate, and a drain voltage $V_D \approx 5$–6 V. The former high-voltage supply is internally generated by a special circuit referred to as "charge pump." Electrons traveling from the source to the drain can eventually reach an energy exceeding the Si–SiO_2 potential barrier, and be injected into the floating gate.

Electron tunneling from the floating gate may occur either at the source side, if the latter extends laterally below the floating gate, or uniformly across the channel. In order to let it happen, it is necessary to set the control gate at a large negative voltage $V_{CG} \approx -12$ V, and the source at a large positive voltage $V_S \approx 5$ V, while keeping the drain floating. In doing so, the oxide field becomes large enough to reduce the tunneling distance down to a few nanometers. The negative voltage is again generated on a chip by a charge pump. The use of negative voltages requires a triple-well CMOS process to be worked out, with the external *p*-well biased at the most negative voltage, in order to prevent shorting of the *p*–*n* junctions with the *n*-layer biased at negative voltages.

Flash Array Architecture

From the standpoint of the cell array organization, nonvolatile memories can be based either on a NOR

or a NAND architecture. The former has a shorter access time, but the latter is denser and, thus, cheaper. A growing number of useful applications is not so critical as far as speed is concerned, while cost is always a critical issue: hence the NAND architecture is gaining momentum in the market place of flash memories.

The array organization of the NOR architecture is shown in **Figure 9a**. Cells within the same row share the word line, while cells within the same column share the bit line. Successive rows are mirrored in order to access the same common source line crossing the array parallel to the word lines. The common source lines of a sector are shorted and kept at ground potential during a read, but are raised at a positive voltage when an erasing operation takes place for the whole array sector. Because of the common source configuration of the flash cells during a read, this cell array organization is referred to as an NOR architecture.

The array organization of the NAND architecture is illustrated in **Figure 9b**. A number of cells, typically 8 or 16, are connected in series rather than in parallel, which allows for a compact layout and a higher cell density within the array. This is due to the small number of contacts to the bit line and the absence of the common source lines within the array; hence, a reduced word-line pitch becomes possible. In series with the flash cells, two select transistors are present at the top and bottom of the NAND column.

The information is stored as a charge within the floating gate. In this case, however, a logical "1" is associated with a negative threshold voltage, while a logical "0" is represented by a positive threshold voltage. Also, both programming and erasing are carried out by Fowler–Nordheim tunneling into, and out, of the floating gate.

The word lines are normally at V_{DD}. In this configuration, the unaddressed cells act as ON pass gates regardless of their actual threshold, which is typically much smaller. A read operation is carried out by lowering the word line down to 0 V, and detecting either a current, or no-current flow, on the bit line. In the former case, the threshold voltage of the addressed cell must be negative; in the latter case, instead, the threshold voltage of the addressed cell is positive. Reading a current through a series of several cells and two select transistors is, however, a slow operation: the random access time is, in fact, of the order of several microseconds.

In order to program a cell, the control gate is driven to a large positive voltage $V_{PP} \approx 15$ V, while the corresponding bit line is kept at 0 V. This ground potential is transferred to the source and drain of the addressed cell via the selection transistor and the series flash cells, with a negligible voltage drop due to the small tunneling currents flowing into the floating gate. To prevent program disturbances of the unselected cells on the same word line, all the remaining bit lines must be kept at V_{DD} and the unselected word lines at a voltage $V_m \approx 10$ V, large enough to ensure a good pass-gate function irrespective of the cell threshold voltage, yet sufficiently small to prevent electron tunneling into the floating gate of the unselected cells.

To erase a sector, electrons are injected from the floating gates to their respective channels, again by Fowler–Nordheim tunneling. This process is carried out by grounding all the word lines, and forcing a high voltage (up to 20 V) to the *p*-well of the array. In order to do so, all the peripheral circuits must be isolated from the *p*-well by lowering to ground both the select lines; otherwise the *n*–*p* junctions would become forward biased and overload the charge pumps. Compared with the NOR architecture, the erase operation is much faster, since it does not

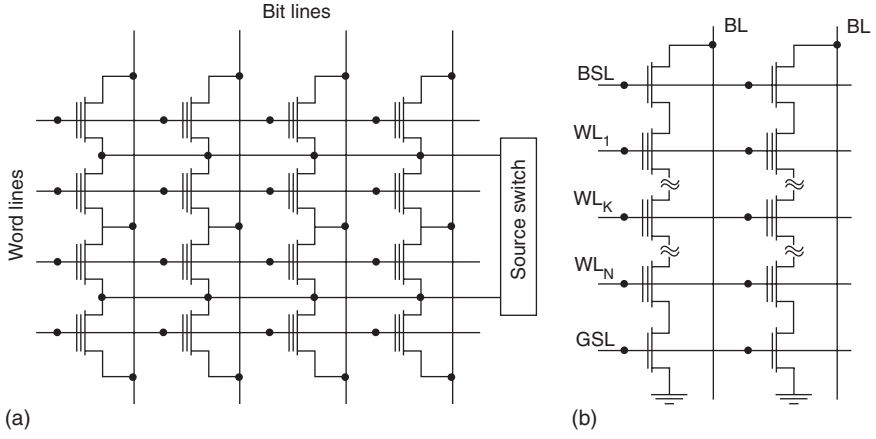


Figure 9 Flash EEPROM cell array organization: (a) NOR architecture and (b) NAND architecture.

require the program to be performed successive to "0." It should be noticed that the erase operation is made easier if the coupling coefficient α_B is small, so that the oxide field is maximized for a given *p*-well voltage.

In flash NAND architectures, the required voltage drops to ensure Fowler–Nordheim tunneling are not shared between a negative and a positive voltage generated by two distinct charge pumps. This requires a large negative voltage to be generated on chip, thus making the charge pump design more difficult. On the other hand, electron tunneling does not require high currents; hence, power consumption is reduced with respect to the NOR architecture.

Multi-Bit Memory Storage

The capacity of a flash EEPROM can be increased by storing two bits per cell, if four different threshold voltages, rather than two, can be accommodated within the cell. For a reliable operation, it is requested that the standard deviation associated with the statistical distribution of the threshold voltages be much smaller than the average separation between their central values. This may be accomplished with a tighter control of the injected charge within the floating gate, and with a more sophisticated reading procedure. More specifically, the threshold voltage must be monitored while programming, and the number and duration of writing pulses controlled by a suitable logic.

The reading procedure requires a number of steps involving two successive comparisons of the output current with three reference currents made available by suitably programmed dummy cells. In doing so, the memory capacity is doubled for a given array size, at the expense of a degraded access time.

Reliability of the Flash EEPROM

The read and write operations of the flash cell are similar to those of EPROMs: therefore, flash memories are subject to the same kind of read and program disturbances illustrated in **Figure 4**. As opposed to the EPROM, however, flash memories can undergo multiple program and erase cycles; hence their reliability problems are much more severe. The large oxide fields, which are needed for electron tunneling to occur, may be responsible for a degradation of the gate oxide.

The "endurance" of an EPROM cell refers to the property of experiencing multiple write and erase operations without affecting the memory performance. Connected with the endurance is the data retention, that is, the property of maintaining the written data unaltered for a long time (10 years) with or without any power supply. Charge loss in flash memories may be caused by: (1) defects in the gate oxide, (2) defects in the inter-poly dielectric (ONO), and (3) mobile ion contamination.

The degradation of the gate oxide typically occurs due to electron trapping during the write and erase cycles. This degradation may be responsible for premature oxide breakdown or low-voltage leakage, referred to as "stress-induced leakage current" (SILC). The latter is an anomalously large current flowing across the gate oxide in a relatively small, but significant number of cells. The SILC seems to be due to the percolation of electrons through a number of defects located within the oxide. This would justify the statistical nature of the SILC, which affects only a small percentage of the flash cells.

Experimentally, it has been found that the percentage of cells affected by the SILC is a strong function of the oxide thickness which, therefore, cannot be scaled below 8–10 nm. Similar considerations apply to the inter-poly dielectric, which has a lower limit at $\sim$10–12 nm. The nonscalability of the gate oxide and the inter-poly dielectric, in turn, does not allow the programming and erasing voltages to be reduced, in sharp contrast with the scaling trend of the CMOS technology. Besides, the generation of large programming and erasing voltages by means of charge pumps operating at small supply voltages, is a rather inefficient process, which makes the functionality of flash memories rather cumbersome.

Due to the above problems, two evolutionary approaches are currently being pursued: the nanocrystal memory and the SONOS memory. In the first approach, the floating gate would be replaced by a large number of silicon nanocrystals embedded within the gate oxide. Both programming and erasing would be carried out by electron tunneling across the gate insulator. The advantage of nanocrystals is twofold: first, if a leaky percolation path exists, only one nanocrystal would be affected and not the entire floating gate; next, the gate dielectric could be made as small as 2 nm, with a substantial advantage for the scaling properties of the cell, since electron tunneling across a thin gate insulator is known to generate less oxide defects.

The second approach, referred to as SONOS, is based on the use of an oxide–nitride–oxide gate dielectric, which improves upon a former nonvolatile memory, now obsolete, called MNOS. Electrons injected into the oxide would be permanently trapped within the nitride layer, which has a smaller barrier than the oxide, thus affecting the device threshold voltage. The trapping of each individual electron would make the tunneling across the oxide nearly impossible at low applied voltages, thus preventing

the SILC effect. Erasing would again occur by electron tunneling, by applying a large negative voltage to the gate of the SONOS cell.

New Memory Types

In addition to the semiconductor memories illustrated in the previous sections, which dominate the memory market, different device concepts are currently being investigated. Some of them, such as Magnetic RAMs (MRAMs) and Ferroelectric RAMs (FeRAMs), have been around for many years, but never succeeded in becoming truly competitive products. These memories are based on the properties of either ferromagnetic or ferroelectric materials, which are characterized by a hysteresis of their *M–H* and *P–E* relationships, respectively. Here *M* is the magnetic moment and *H* the magnetic field; *P* is the electric polarization and *E* the electric field. MRAMs, while not being competitive in density and cost with DRAMs, have comparable access times, an excellent radiation hardness, and are nonvolatile. As such, they have mainly been used in military and space applications.

Unlike MRAMs, ferroelectric memories are very appealing because of the simplicity of the basic cell, which promises a small size and a high density, as well as an easy approach to read/write operations. The major technological problems are connected with the use of ferroelectric materials, such as Pb(Zr, Ti)O_3, referred to as PZT, due to their difficult quality control, instability, and contaminating nature.

An entirely new approach is based on the use of chalcogenide materials, that is, an alloy of Ge, Sb, and Te (GST), which undergoes a phase change when heated up to high temperatures. More specifically, if a polycrystalline GST alloy is heated up to the melting temperature ($\sim$650°C), it turns into an amorphous state when it cools down. On the other hand, if an amorphous material is heated up to 500°C, it turns back into a polycrystalline state. The final state is indefinitely conserved at room temperature; hence, these materials can provide a nonvolatile memory function. Due to the large resistance change associated with a phase change, it is straightforward to read the stored information by a simple current measurement.

Magnetic RAM

Magnetic memories are based on the properties of ferromagnetic materials, such as iron (Fe), cobalt (Co), and nickel (Ni), which exhibit a strongly nonlinear and hysteretic *M–H* relationship, as shown in **Figure 10**. The residual magnetization of the above materials after the field return to zero is due to the alignment of the electron spin, which is conserved at room temperature.

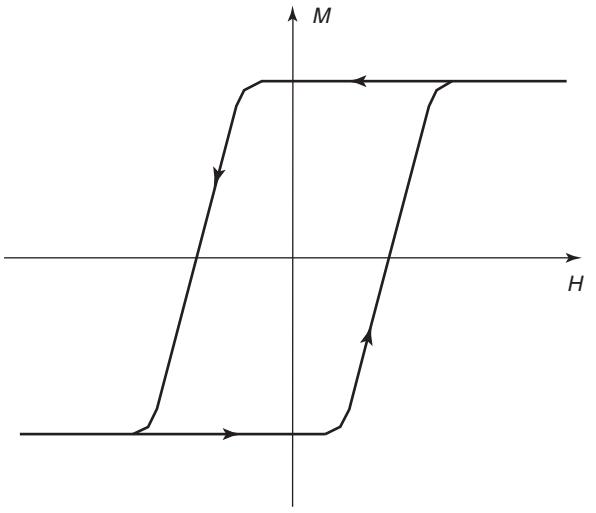


Figure 10 Magnetic moment against magnetic field in typical ferromagnetic materials.

Magnetic core memories were used in early mainframe computers, where a matrix of current carrying wires was used to program magnetic rings referred to as nuclei. Recent memory devices keep storing the information on magnetic elements, but the sensing scheme is now entirely different, and is based on the effect of magnetoresistance (MR), that is, a resistance change of a thin metallic layer due to the presence of a magnetic field. Magnetoresistance, however, is a small effect, with resistance changes of $\sim$1%.

More interesting from this standpoint is the giant magnetoresistance effect, which occurs when a thin metal layer, typically Cu, is sandwiched between two magnetic films consisting of NiFe. When the magnetic moments are parallel, the "up" electrons do not scatter, and only the "down" electrons scatter at the interface with the magnetic materials. When the magnetic moments in the two films are in opposition, instead, both type of electrons scatter and a significant resistance increase, which can be as large as 15%, typically occurs. The plot of magnetoresistance versus magnetic field is qualitatively shown in **Figure 11**. The effect according to which electrons with long mean free paths travel with low resistance when the magnetic materials have parallel alignments is called "spin valve," and the discipline which investigates the properties of materials due to the electron spin is called "Spintronics."

Several different configurations of magnetic films have been proposed for data storage in MRAMs. These include: (1) the anisotropic MRAM, (2) the spin-valve MRAM, (3) the pseudo-spin valve (PSV) MRAM, (4) the magnetic tunnel junction (MTJ), and

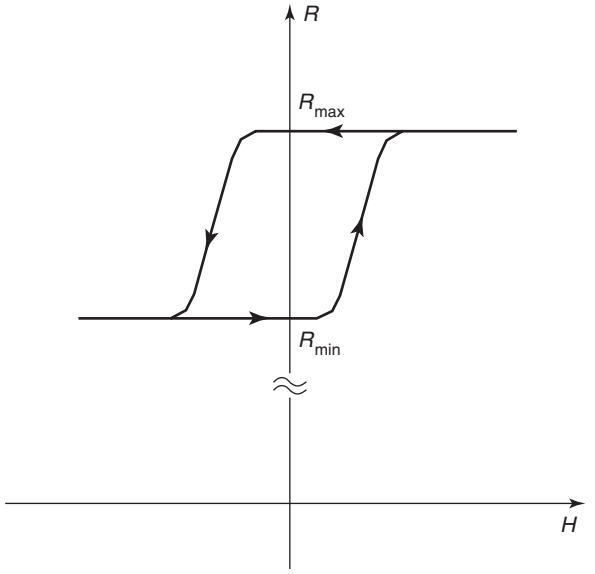


Figure 11 Plot of resistance against magnetic field showing the GMR effect.

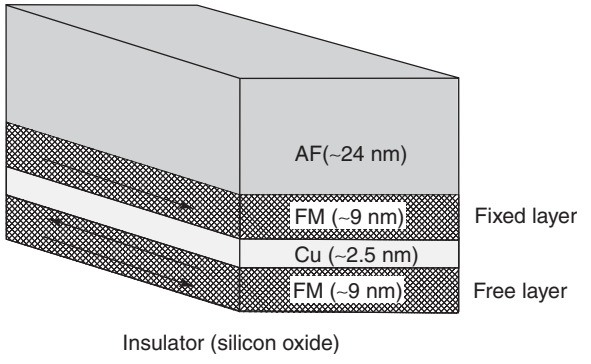


Figure 12 Schematic view of the spin-valve memory cell.

(5) the Hall effect MRAM. In what follows, only the spin-valve MRAM and the MTJ are discussed.

A schematic view of the spin valve is shown in **Figure 12.** A thin Cu layer ($\approx$2.5 nm) is sandwiched by two ferromagnetic films ($\approx$9 nm) made by a NiFe alloy. A much thicker pinning layer made by an anti-ferromagnetic material, such as FeMn or IrMn, is deposited on top of the upper magnetic layer. In spin-valve MRAMs, the magnetic moment of the magnetic layer adjacent to the anti-ferromagnetic material is pinned, while the other magnetic layer is free to switch from parallel to antiparallel orientation under the influence of an external magnetic field. When the magnetic moments of the two magnetic layers are parallel and the cell resistance is low, a logical "1" is associated with the state of the cell. When the polarizations of the two magnetic layers are antiparallel, and the cell resistance is high, a logical "0" is associated with the state of the cell.

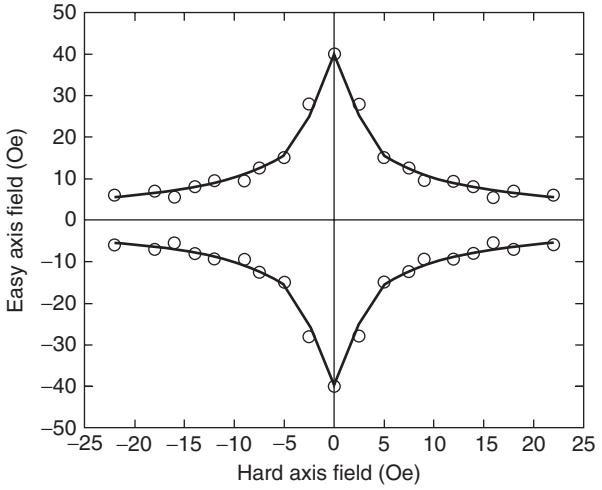


Figure 13 Switching threshold locus of the spin-valve memory cell (Asteroid curve).

Normally, a ferromagnetic material can be polarized at random directions. If, however, the magnetic film is narrower than the magnetic domain width, the film exhibits only two polarization states. This makes the switching threshold well defined.

The spin valve is located at the intersection of two conductors (word lines) whose orientations are parallel and perpendicular to the magnetization vector. For a write operation, orthogonal current pulses I_x and I_y generate longitudinal and transverse magnetic fields H_y and H_x, respectively, whose vector sum switches the magnetic moment in the free magnetic layer. The cell switch occurs in less than 2 ns. The use of two perpendicular components of the magnetic field lowers the switching threshold of the cell, as shown in **Figure 13.** This figure plots on the H_x–H_y plane the locus of the magnetic field components at which the spin-valve cell undergoes a switching. The asteroid shape of the curve indicates that combining suitable values of H_x and H_y allows much smaller magnetic fields to be used, to the advantage of the stability of the pinned-layer polarization. The same plot indicates that the normal component of the electric field (easy-axis field) switches the cell at a smaller field than the longitudinal component (hard-axis field).

To read the cell state, a specific polarization is forced on the free layer: if the cell switches, the voltage across the sense line will then change according to the resistance change. If the cell does not switch, no resistance change is sensed. In the former case, a positive voltage change indicates that the switching occurred from the parallel to the antiparallel polarization, meaning that the original cell state represented a logical "1," while a negative voltage change is the signature of a logical "0." The cell is

then written in the other direction in order to restore the original cell state. This cell can also be read by comparison of a sense current with a reference current, if an access transistor is used for each cell. In this case, the read can be nondestructive.

One of the major problems connected with MRAMs is that the resistance change of the conducting layer is still rather small for the needs of a large-scale production environment. Furthermore, the latter decreases with increasing temperature. Also, the pinning field decreases with the cell width. For submicron geometries, a spin valve that permits the polarization of both magnetic layers to switch appears to be a solution. The pseudo-spin-valve cell offers this possibility.

An alternative to the GMR spin-valve technology is the MTJ RAM. The latter is based on a thin insulating layer, typically Al_2O_3, sandwiched by two magnetic films as shown in **Figure 14** and, in this case, the current flows normal to the insulator surface. The detection mechanism is referred to as "spin-polarized tunneling," and its occurence depends on the relative orientation of the magnetization vectors in the two magnetic layers.

The resistance increase when the magnetic polarization becomes antiparallel is larger in this case, being $\sim$40% of its original value, and a higher memory density appears to be possible due to the perpendicular current flow with respect to the magnetic layers. Process uniformity is more critical, however, due to the extremely small insulator thickness, which is only a few tenths of a nanometer.

The memory organization may or may not use an access transistor, which acts as an isolation device. For a write, the access transistor is off. Current flows in orthogonal wires running above and below the cell, inducing a magnetic field, which changes the polarization of the free magnetic layer. The pinned magnetic layer is fixed. The programming wires are organized in a cross-point architecture, and the combination of the currents is designed to be just above the switching threshold of the MTJ. The read word line turns on the access transistor, as shown in **Figure 15**, and the current flowing across the MTJ is sensed on the bit lines.

A twin-cell organization improves the read reliability. In this case, the twin cells are forced to be always written in two opposite states. The data is read by detecting the resistance difference between the MTJs of the twin cells. This doubles the signal available for sensing, and improves the tracking properties with respect to a remote reference cell.

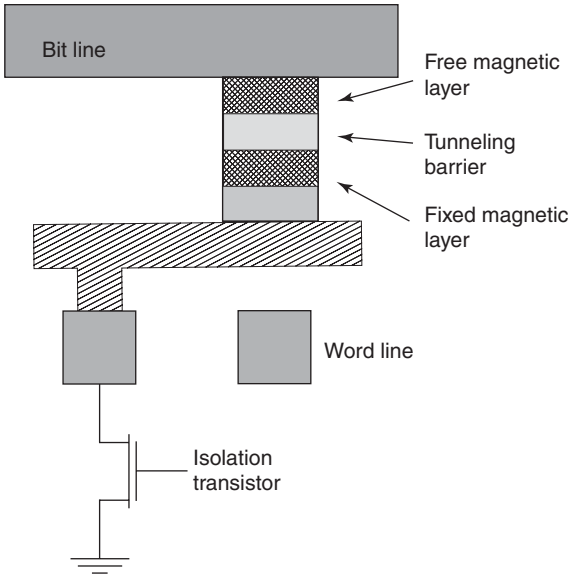


Figure 14 Schematic view of the MTJ memory cell.

Ferroelectric RAM

Ferroelectric materials exhibit a hysteretic *P–E* relationship, shown in **Figure 16**, which can be taken advantage of to fabricate memory devices.

In FeRAMs, the information is stored within the ferroelectric insulator of a capacitor in the form of an upward or a downward polarization. If a positive voltage is applied to the capacitor with an upward polarization, no change occurs in the compensating charge, and a small voltage shift is observed in the external circuit. If the polarization is downward, instead, the biased domains switch, causing a compensating charge to flow in the capacitor and the external circuit, and changing the capacitor voltage by a greater amount.

This means that the stored information can only be read by erasing the content of the memory, which

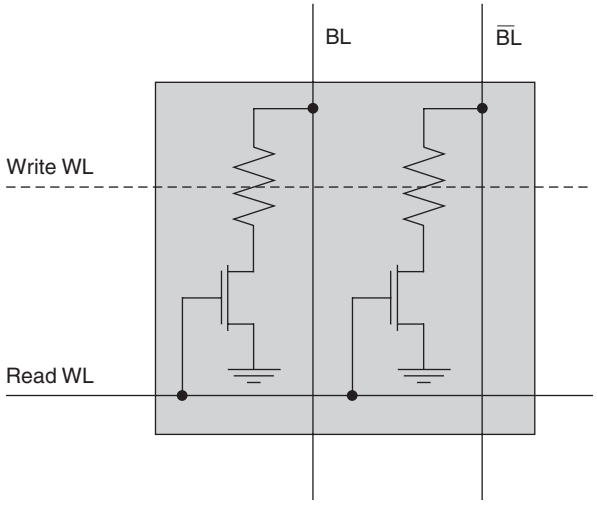


Figure 15 Cell array organization of an MTJ memory.

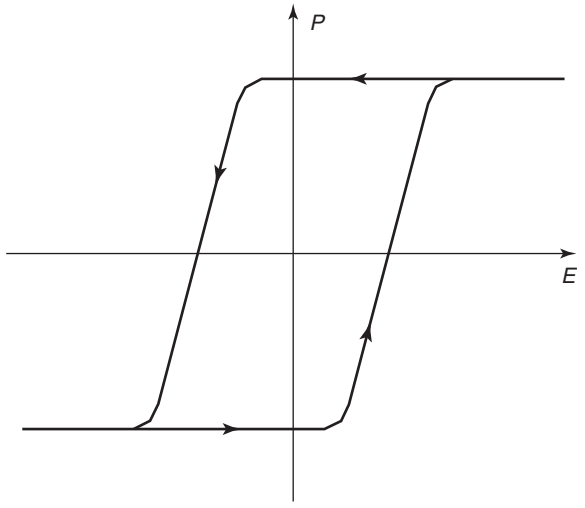


Figure 16 Plot of the electric polarization vs. electric field for a ferroelectric material.

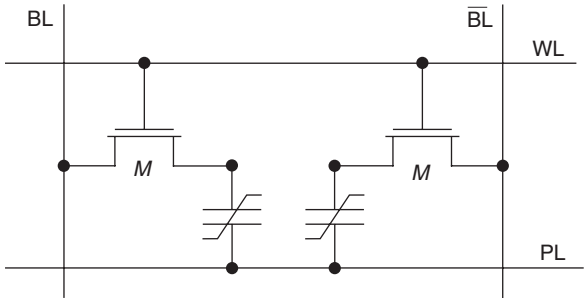


Figure 17 Circuit schematic of the 2T2C cell of a ferroelectric memory.

must then be restored to its previous state. The basic cell of a FeRAM contains either two transistors and two capacitors (2T2C), or one transistor and one capacitor (1T1C). The former cell is shown in **Figure 17**.

Two ferroelectric capacitors are connected to the bit line BL and to the bit line bar BL′ via two pass transistors driven by the same word line WL. The information is written by placing V_{DD} and 0 on BL and BL′, turning on the word line, and then cycling the plate line from 0 to V_{DD} to 0. When the plate line is at zero voltage, the ferroelectric capacitor on the bit line held at V_{DD} is polarized in a positive direction. When the plate line is pulsed to V_{DD}, the capacitor bias goes back to zero and the polarization is not affected. The other capacitor, instead, is polarized in the negative direction when the plate line goes to V_{DD}, and its state remains unmodified when it returns to zero. Therefore, the two capacitors always have opposite polarizations after cycling the plate line.

The cell is read by turning on the word line, presetting the bit lines at zero voltage in a high-impedance

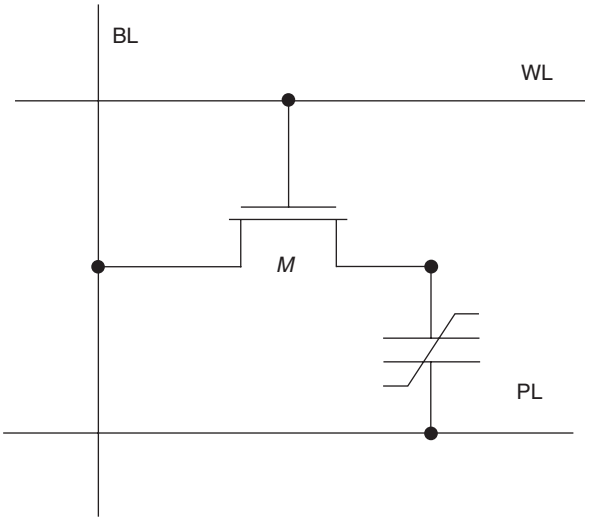


Figure 18 Circuit schematic of the 1T1C cell of a ferroelectric memory.

state, and switching the plate line to V_{DD} and back to 0. One of the ferroelectric capacitors will flip and the other will not, which generates a differential charge reading on the bit lines. The read is destructive, and the status of the cell must be restored. This may be accomplished by taking the plate line low, and letting the sense amplifier inject the read charge back into the cell. This will restore the original polarization on the flipped capacitor.

The read access time of a FeRAM is $\sim$70 ns, quite comparable with the access time of DRAMs and flash EEPROMs. The write access time is again $\sim$70 ns, that is, much faster than the write operation of a flash memory. Likewise, the write endurance of the FeRAM, which ranges between 10^8 and 10^{12} cycles depending on the ferroelectric material, is much better than the endurance of commercial nonvolatile memories.

In recent times, the 1T1C cell memory has been worked out. The cell structure is a simplified version of the 2T2C cell, as shown in **Figure 18**.

A logical "0" can be written in the cell by setting the bit line to 0, raising the word line and the plate line to V_{DD} and, after switching the ferroelectric material to a negative polarization, lowering the word line and, subsequently, the plate line down to 0. Writing a logical "1" can be accomplished instead by setting the plate line to zero, raising the word line and the bit line to V_{DD} and, after the positive polarization of the ferroelectric material, lowering the word line and, subsequently, the bit line.

Reading is accomplished by turning on the word line while setting the bit line at 0 V in a high-impedance state, and cycling the plate line from 0 to V_{DD}. For a cell in the "0" state, there will be a minimal

voltage change on the bit line, since the capacitor does not switch its polarization. For a cell in the "1" state, there will be a much larger voltage change on the bit line since the capacitor switches its polarization and the compensating charge flows in the bit line. This voltage change is compared with a reference voltage suitably set at an intermediate value between the expected low- and high-voltage levels. This switch from "1" to "0" requires a flip-back of the capacitor to restore the previous state before the word line is turned off.

The obvious advantage of the 1T1C cell with respect to the 2T2C cell is due to its smaller size, and, thus, higher density. However, the differential reading of the 2T2C cell is carried out on two capacitors which are one next to the other, and whose electrical properties are expected to track very well due to their vicinity and common history. Conversely, the reading of the 1T1C cell must be compared with a reference voltage generated elsewhere within the chip. The use of a ferroelectric dummy cell to generate the reference voltage has the potential advantage of tracking the chip-to-chip variations of the ferroelectric storage capacitors. However, the dummy cell would be switched many more times than the storage cells, thus generating a reliability problem due to fatigue; besides, the dummy cell being far away from the storage cells, additional factors such as uniformity of the ferroelectric material and temperature dependence of the *P–E* relationship come into play.

Phase-Change Memory

Phase-change memories (PCMs) are based on the peculiar properties of chalcogenide materials, that is, compounds based on sulfur, selenium, or tellurium, such as GeSe, AsS, GeAsTe, or GeSbTe (GST). The latter is the most frequently used material in PCMs. Memory switching in these materials is primarily a thermal process, which involves a phase transformation from a crystalline to an amorphous state, and vice versa, under the influence of a heat source. When the heating process stops, the material retains its new state, thus exhibiting memory. In practice, this transformation is achieved by passing a constant current through the sample for some time. The transition to an amorphous state (reset) occurs by resistive heating up to the local melting of the GST. The resistance change between the two states is $\sim$2–3 orders of magnitude, so that reading can easily be accomplished by low-bias nondestructive detection of the cell resistance.

A schematic cross section of two PCM cells is shown in **Figure 19**. The cell core is basically given by a crystalline GST layer, which is located between a vertical resistive heater and a horizontal metal line (bit line). The vertical resistor is connected to the emitter of a *p–n–p* bipolar transistor, the base of which is connected to the word line. In a process variant, the access transistor can be a MOSFET, rather than a BJT.

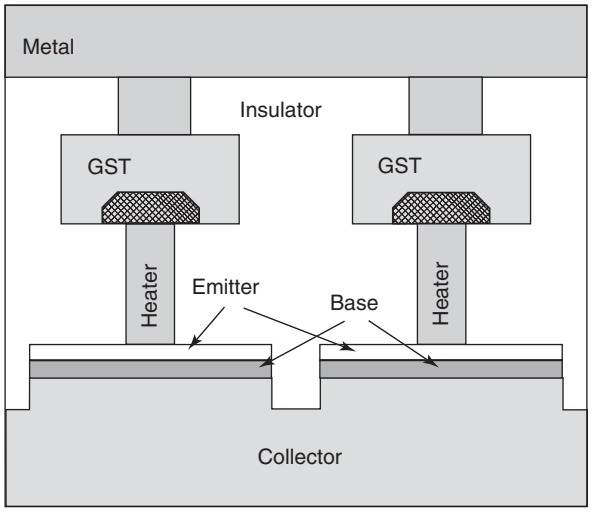


Figure 19 Schematic cross section of two PCM cells.

In order to let a phase change occur, the GST layer is heated by a current pulse. With present technologies, a 100 ns current pulse up to 0.4 mA makes the material crystalline and turns it to a low-resistance state (set). If the current pulse increases above 0.4 mA, the melting temperature of $\sim$650°C is reached at the interface of the GST film with the heater. During the cool-down process, the latter turns to a high-resistance amorphous state (reset).

An obvious concern of this technology is the thermal cross-talk between adjacent cells. In principle, the local heating of the GST layer up to the melting temperature during a reset could spread around heat and, possibly, set a neighbor cell. Thermal simulations, however, seem to indicate that this is not the case, and that no appreciable thermal cross-talk would, in fact, occur down to the 65 nm node of the ITRS. This is mainly due to the linear scalability of the reset current with the contact area for a given cell geometry. If all the physical dimensions of the cell are scaled down by a common factor k, instead, one would expect the set and reset resistances, as well as the thermal resistance, to increase by the same factor k, and the programming currents to decrease by k at constant voltage.

The scalability of the PCM cell, the elimination of large voltages required by the flash-memory technology, the reduced writing and erasing time, and the addressability of every single cell make this new

nonvolatile memory technology a promising candidate for low-voltage, high-density, low-cost, future-generation products.

See also: Conductivity, Electrical; Electrons and Holes; Integrated Circuits; Memory Devices, Volatile; Semiconductor Devices; Transistors.

PACS: 85.25.Hv; 84.30.−r; 85.30.Tv; 84.30.Bv; 84.30.Sk

Further Reading

Brown WD and Brewer JE (1997) *Nonvolatile Semiconductor Memory Technology – A Comprehensive Guide to Understanding and Using NVSM Devices*. New York: Wiley IEEE Press.
Cappelletti P, Golla C, Olivo P, and Zanoni E (1999) *Flash Memories*. London: Kluwer Academic Publishers.
Compardo G and Micheloni R (eds.) (2003) *Special Issue on: Flash Memory Technology*. Proceedings of the IEEE, vol. 91, no. 4, April.
Prince B (2002) *Emerging Memories – Technologies and Trends*. Norwell: Kluwer Academic Publishers.

Memory Devices, Volatile

G Baccarani and E Gnani, University of Bologna, Bologna, Italy

Introduction

Modern computer architectures rely on a hierarchy of memories, which differ in access times and density. The higher speed is typically achieved at the expense of a larger cell area, smaller memory capacity and, thus, higher cost per bit. At the upper levels of the hierarchy one finds registers, which are deeply rooted within the logic circuits of the processor, and three levels of cache memory, with the first two levels usually embedded within the processor chip. Caches typically employ static random access memories (SRAMs) based on either six or four transistors per cell. At the next level, the computer central memory uses dynamic random-access memories (DRAMs) based on the one-transistor memory cell. At the lower levels of the hierarchy, one finds several mass-memory devices, namely, magnetic disks optical disks, such as CDs and DVDs, and flash memory cards. Magnetic tapes are used to archive large amount of data to be seldom retrieved. As opposed to SRAMs and DRAMs, which are volatile memories, mass-memory devices are nonvolatile, that is, they retain the stored information upon power loss or shutdown. This is an essential requisite, which allows information to be safely stored and retrieved whenever necessary.

To date (2004), memory devices have found widespread use not just within computers, but also in a variety of professional systems, such as telecommunication networks and phone exchanges, robotic and control systems, automotive and medical electronics systems, as well as commodities and consumer products, such as cell phones, personal digital assistants (PDAs), digital video recorders and still cameras, digital TV (DTV), set-top boxes, GPS-based navigation systems, toys, and smart cards. Despite this wide variety of applications, which require a different throughput, latency and power consumption, and the related diversification of the memory market, the emphasis, in what follows, is on unifying factors, such as basic physical principles, and general architectural features, which allow one to treat this subject within the constraints of a tutorial presentation.

This treatment specifically addresses semiconductor memories (mass memory devices based on magnetic and optical storage are discussed elsewhere in the encyclopedia) and is split into two parts: one part is devoted to volatile memories, which are by far the most widespread devices in view of their excellent performance and density; the other part is devoted instead to nonvolatile memories, which are being used only when the nonvolatility is an essential requisite of the application.

The classification of semiconductor memories is presented in the next section. The section "Volatile memories" discusses the general features of such devices, including access time, packing density, and reliability issues. The section "Static random access memory (SRAM)" is devoted to the description of the basic cell topologies, cell-array organization, and sense-amplifier operation. The section "Dynamic random access memory (DRAM)" addresses the structural and functional properties of the one-transistor memory cell, the organization of the cell array, the structure and operation of the sense amplifier and the memory architecture. Finally, the section "Content-addressable memory (CAM)" contains a short description of the structure and operation of associative memories, as well as their basic application as a tag memory of set-associative caches.

Memory Classification

Semiconductor memories can be classified according to both structural and functional criteria. Following

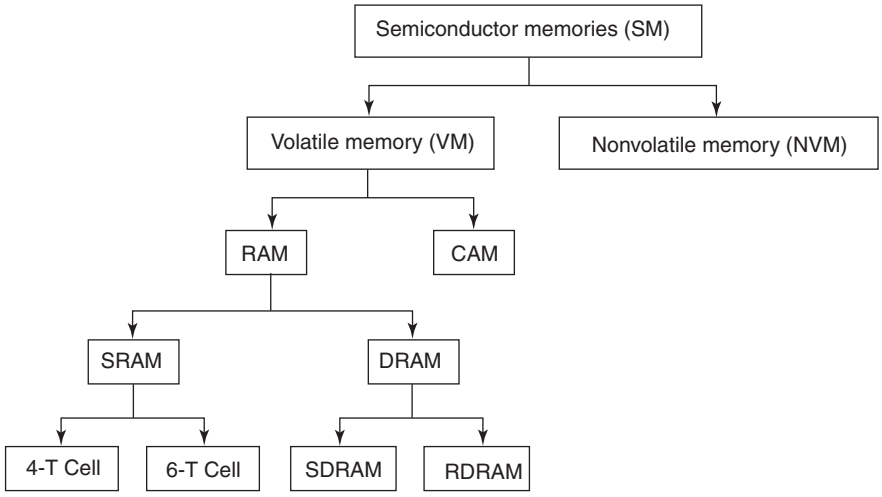


Figure 1 Semiconductor volatile memory classification according to functional criteria.

the latter approach, they are split first into volatile and nonvolatile memories, as indicated in **Figure 1**. As pointed out in the section, "Introduction," volatile memories do not retain the stored information if the power supply is turned off, while nonvolatile memories do retain the stored information in the above conditions. Also, volatile memories can be written and read on field in very short times. Nonvolatile memories, instead, can only be read on field in very short times, whereas repeated programming and erasing operations may not be possible or, if possible, typically require much longer times. Different physical approaches are used to store information in the above memories.

Most volatile memories are randomly accessed, and are thus referred to as random access memories (RAMs). Sequential memories, where bits are written and read sequentially, are rarely used, being outperformed by RAMs both in terms of speed and power consumption.

Content-addressable memories (CAM) are functionally and structurally different from RAMs. In the latter memories, every single bit, byte, or word can be accessed from the knowledge of its physical address. CAMs are, instead, addressed by an associative search using a data word as input, and provide an address as a result of this search.

Volatile memories store the information either statically (SRAM) or dynamically (DRAM). In the former case, the storing mechanism is based on a regenerative circuit characterized by two stable states. In the latter, information is stored as a charge on a capacitor.

SRAMs use a basic cell containing either six or four transistors per cell. The former type of cell is obviously larger, but is compatible with the standard CMOS technology, and is typically used for small, fast, embedded SRAMs employed in first- and second-level caches. The latter type requires a thin-film transistor (TFT) technology not compatible with the standard CMOS process. On the other hand, it provides a higher density and is usually preferred for large third-level caches outside the processor chip.

Dynamic RAMs may be characterized by different architectures. In the past, DRAMs used to be asynchronous and could start a memory cycle whenever two external commands referred to as row address strobe (RAS) and column address strobe (CAS) were successively asserted. More recently, the need to improve the cycle time of DRAMs has suggested more sophisticated architectures, which operate synchronously under the control of the system clock. To date, the synchronous DRAM (SDRAM) is the most widely used memory architecture. An improved performance is made possible by the Rambus DRAM (RDRAM), by which a careful design of the memory board allows for very large clock frequencies.

Volatile Memories

The most important parameter characterizing the memory performance is the access time, that is, the time elapsed between a read request and the availability of the requested word. The cycle time is, instead, the minimum time between successive read requests; its inverse is the memory throughput. The access time of SRAMs may vary from one to a few nanoseconds, depending on the memory size and the resulting capacitances of the word and bit lines. Therefore, these memories are used at the highest levels of the hierarchy within advanced computers and PCs.

DRAMs feature the smallest bit size and are typically used for the computer central memory. To date,

a central memory of 1 GB or more is common practice for high-end desktop and even portable PCs. Servers or mainframes typically employ several tens or hundreds of giga bytes of central memory; hence, the one-transistor DRAM cell is by far the most largely produced semiconductor device in modern industry. Under the assumption of an average main memory of 1 GB and a worldwide production of 100 million computers per year, it turns out that $\approx 10^{18}$ DRAM cells are manufactured every year.

The access time of DRAMs is $\sim$40–60 ns. Therefore, accessing data stored within the central memory of a computer requires several tens of clock cycles for a modern processor.

Volatile memories can suffer data loss due to soft errors, which are caused by the interaction of energetic nuclear particles, such as α, β, and cosmic rays, with the silicon substrate. These events may occur due to the radioactive decay of some impurity atoms contained within the package material and/or the metal layers of the memory chip. Alternatively, as for cosmic rays, these energetic particles may come from the outside space and occasionally hit the memory device. A charged energetic particle penetrating through a semiconductor substrate generates a large number of electron–hole pairs, which are then separated by the electric field within the device-active region. The generated carriers can possibly compensate the charge stored within a cell capacitor, and modify its state. Even static memories are not immune to soft errors, for an energetic particle impinging onto a bi-stable circuit can upset it and change its state. In both cases, a soft error would occur.

In order to ensure the full reliability of the system, several strategies are pursued. At the device level, the geometry of the space charge region is designed in such a way as to drain excess carriers generated by the energetic particle by reverse-biased *p–n* junctions, so that the charge stored within a cell capacitor can hardly suffer a substantial change; also, the bits of a word are not stored at consecutive physical locations, so that, at most, one bit of a word can be changed due to the occurrence of a soft error. At the system level, a few more parity bits are stored together with a 32-bit word. When a word is retrieved from the memory, the examination of the parity bits makes it possible to detect the occurrence of one or more soft errors and also correct the wrong bit.

Static Random Access Memory

The simplest bi-stable circuit is obtained by cross-connecting two CMOS inverters M_1–M_2 and M_3–M_4, as shown in **Figure 2**. Clearly, such a circuit can be biased in either of the two states, characterized by

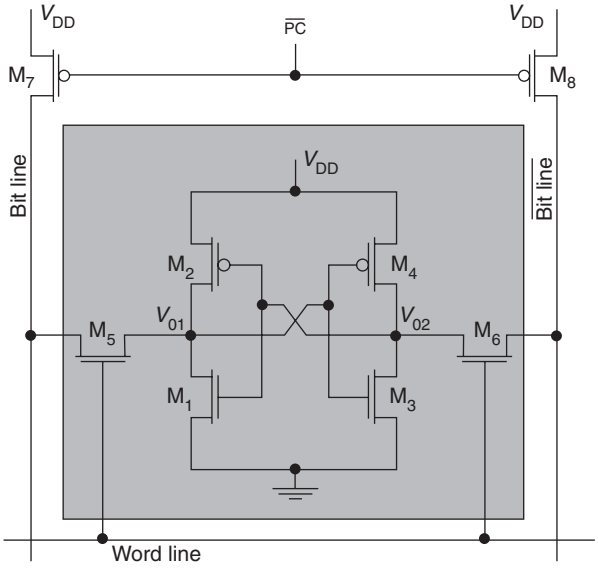


Figure 2 The six-transistor CMOS SRAM cell.

$V_{01} = 0$, $V_{02} = V_{DD}$ or $V_{01} = V_{DD}$, $V_{02} = 0$, respectively. By associating the logical value "0" to the former, and the logical value "1" to the latter bias condition, one stores one bit of information within such a circuit. So long as the power supply is on, this cell will indefinitely preserve the stored information, hence the definition of static memory.

In order to access the information, however, two more devices M_5 and M_6 are needed within the cell, as shown in the same figure. These devices act as pass transistors, which isolate the cell if the word line is low, or connect the cell to the bit lines if the word line is high. Hence, the total number of transistors per cell is six. The bit lines, which are shared by all cells within the same column, provide the readout information to the sense amplifier, that is, the circuit which generates a logic output data from a slight unbalance of the bit lines.

In four-transistor memory cells, the pullup metal-oxide-semiconductor field effect transistor (MOSFETs) M_2 and M_4 are replaced by either passive resistors, as shown in **Figure 3**, or TFTs made by polycrystalline silicon.

In the former case, their resistance R must be very large, since the standby current consumption of a cell is given by V_{DD}/R. The resistance value is thus defined according to the need of supplying the minimum current to preserve the high voltage of the node. In practice, $R \approx 1$–$10\ G\Omega$, which can be achieved using lightly doped polycrystalline silicon films. The use of TFTs as pullup devices reduces the leakage current by nearly two orders of magnitude with respect to the passive resistor cell, and improves the current-drive capability of the pullup devices in the on state.

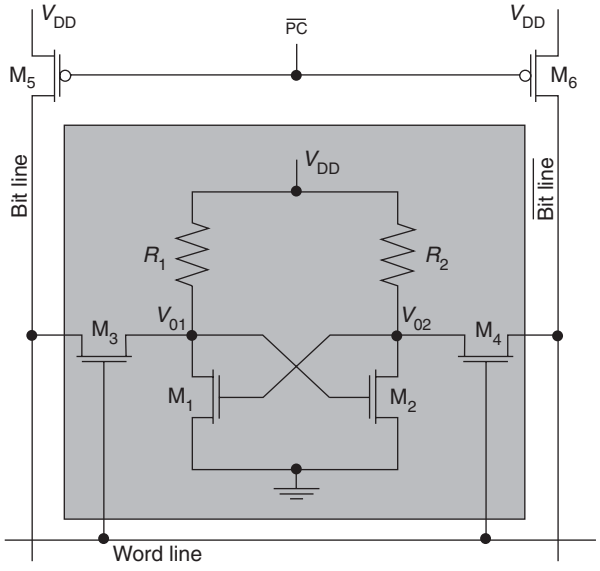


Figure 3 The four-transistor CMOS SRAM cell.

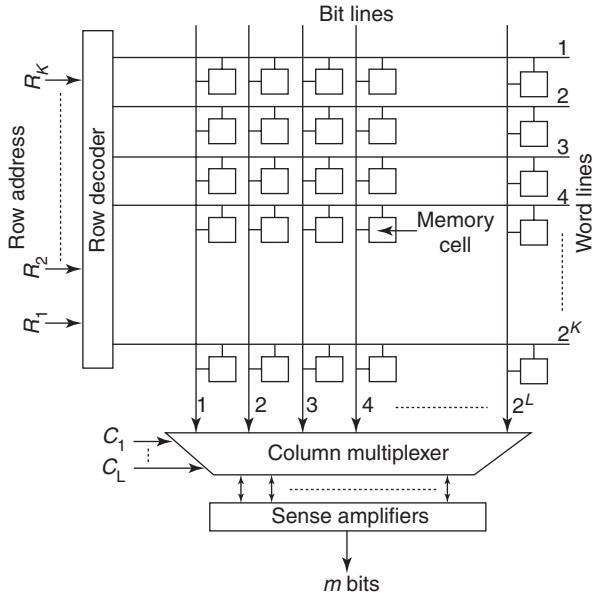


Figure 4 SRAM cell array organization.

This improvement can be achieved at the expense of a more complex technology, involving an additional masking step, and the deposition of an undoped polycrystalline silicon film.

Cell Array Organization

The cell array of an SRAM is organized as indicated in **Figure 4**. A row decoder receives the row address of the cell at the input and raises the corresponding word line high. Hence, all the cells pertaining to the above row become connected to the respective bit lines.

The output multiplexer acts as a column decoder, and selects the appropriate bit, byte, or word. The corresponding bit lines are then connected to the sense amplifier(s), which detect the stored information within the selected cells. If the depth of the column decoder is too large, it may be worth interposing the sense amplifiers between a first partial column decoder, acting on the bit lines which carry analog signals, and a second column decoder, acting on the digital outputs of the sense amplifiers. Clearly, reducing the number of sense amplifiers is an advantage from the standpoint of power consumption, but, on the other hand, interposing too many pass transistors between the bit lines and the inputs of the sense amplifier could be detrimental to the accuracy of the readings.

It should be noticed that, if the row decoder receives a K-bit address and the column decoder gets an L-bit address, then the array contains 2^K rows and 2^L columns. Thus, a total number of 2^N cells is present within the array, with $N = K + L$. The number of selected outputs m depends on the memory organization: if the SRAM is fully decoded, $m = 1$. In this case, one single bit is read from, or written into, the array. Alternatively, m can be either 8 (byte decoding), 16 (half-word decoding), or 32 (word decoding).

Cell Reading and Writing

A reading cycle proceeds as follows: (1) the bit lines, precharged at V_{DD}, are turned to a high-impedance state, (2) the word line is raised, thus connecting the cells within the addressed row to the bit lines, (3) the pull-down transistors of the addressed cells discharge one of the two bit lines, while leaving the other at V_{DD}, (4) the selected bit lines are connected to the sense amplifiers via the column decoder, (5) the sense amplifiers read the content of the addressed cells and make the bit information available to the output circuitry, (6) the row and column decoders are reset, and (7) the bit lines are reset to V_{DD}.

Likewise, a writing cycle proceeds as follows: (1) the row and column decoders are activated, thus connecting the addressed cell to the respective bit lines and the latter to the output circuitry, (2) either one of the two bit lines connected to the cell via the column decoder is discharged to zero, (3) the bit line connected to ground either confirms the status of the cell or upsets the cell in order to make it biased according to the desired input, (4) the row and column decoders are reset, and (5) the bit lines are precharged to V_{DD}.

The timing of the above operation sequence is controlled by internally generated voltages.

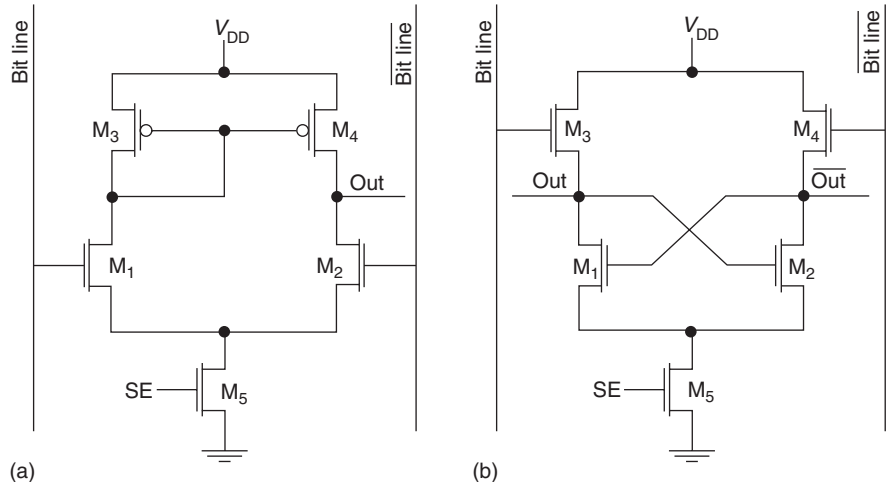


Figure 5 SRAM sense amplifier. (a) voltage amplifier, (b) level shifter.

Sense Amplifier

For static memories, the sense amplifier is often the standard one-stage CMOS differential amplifier with single-ended output, as shown in **Figure 5a**.

The function of the sense amplifier is, in this case, that of allowing for a quick reading of the differential signal on the bit lines, so that the discharging process can be stopped as soon as possible, to the advantage of speed and power consumption. One major drawback of this simple scheme is that, the bit lines being pre-charged to V_{DD}, the differential-pair M_1, M_2 operate in the triode region, with a severe degradation of the voltage gain.

A better solution is a two-stage scheme, with interposition of a differential-level shifter, shown in **Figure 5b** between the bit lines and the voltage amplifier. This circuit features a positive feedback intended to increase the differential voltage gain of the level shifter above 1. Most important, the appropriate DC bias of the differential pair optimizes the amplifier gain, resulting in an excellent performance of the sense amplifier.

Dynamic Random Access Memory

Information can be dynamically stored as a charge on a capacitor, as indicated in **Figure 6**. If the capacitor is charged, its voltage $V_c = V_{DD}$; otherwise, $V_c = 0$. Here again the MOSFET acts as a pass transistor, which can either isolate the cell or connect it to the bit line. In order to charge the cell capacitor to V_{DD}, the word line must be driven to a higher supply voltage V_{PP}. A disadvantage of this storage scheme is that, due to the transistor leakage current, the charge is not indefinitely conserved within the capacitor but, rather, leaks out in a time approximately a few milliseconds. This requires the information to be refreshed some thousand times per second; hence, the definition of dynamic memory.

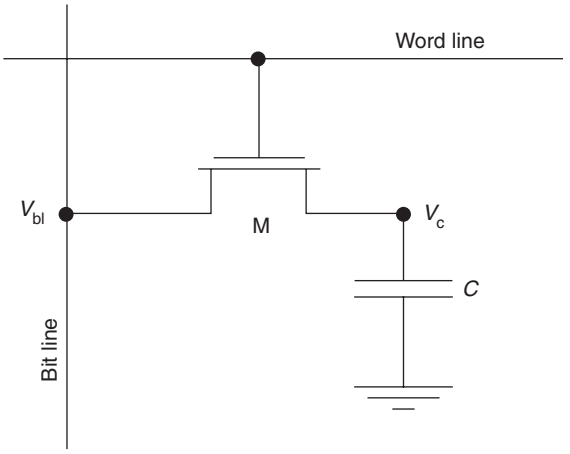


Figure 6 The one-transistor DRAM cell.

The cell capacitance must be as large as possible to ensure a long retention time as well as a safer reading of the stored information. At the same time, it must occupy a small silicon area, in order to allow for a large density and a small cost per bit. These conflicting requirements are pursued with either a trench or a stacked cell. In the former case, schematically shown in **Figure 7a**, the capacitor is placed in a deep trench engraved within the silicon substrate by reactive-ion etching. The trench surface is thermally oxidized and filled with polycrystalline silicon, which acts as the upper electrode of the capacitor and is electrically connected to the MOSFET source. The plate electrode is formed instead by the silicon substrate. In a recent implementation, even the pass transistor of the cell is fabricated along the trench sidewall and is therefore vertically oriented.

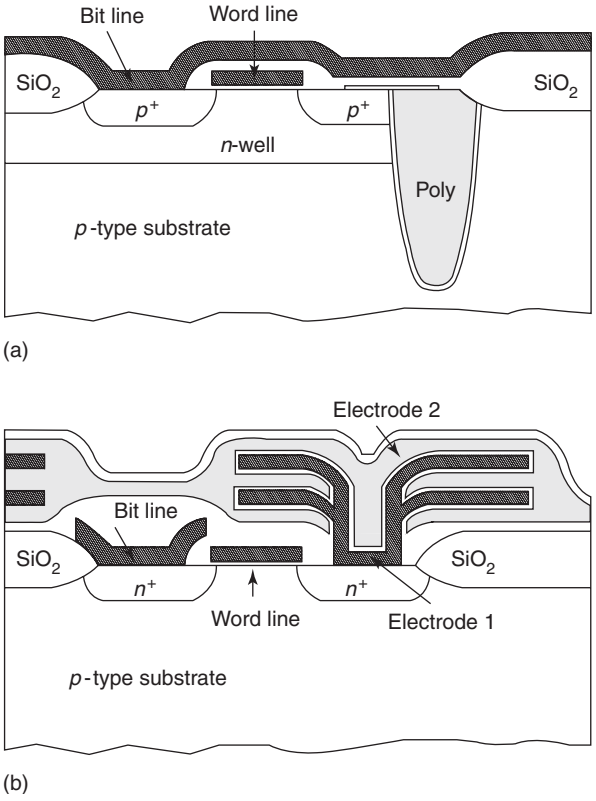


Figure 7 Schematic view of DRAM-cells. (a) DRAM-trench cell and (b) DRAM-stacked cell.

The stacked cell, shown in **Figure 7b**, aims to deploy a large capacitance by extending the surface of the upper electrode in a complex elevated structure. This is achieved by a double-fin structure, generally made by polycrystalline silicon, which is electrically insulated and surrounded by the second electrode, representing the capacitor plate.

An electron micrograph, showing a realistic image of a trench cell, is represented in **Figure 8a**. The aspect ratio of the trench cell is remarkably elongated, stressing the need for a large capacitance, typically larger than 35 fF, within a small silicon area. The polycrystalline silicon finger is electrically connected to the source of the *p*-channel transistor, whose drain is shorted to the bit line via a tungsten plug.

Figure 8b shows instead an electron micrograph of a realistic stacked cell. The double-fin structure represents the upper electrode of the cell capacitor and appears to be light as the silicide layer overlapping the transistor gate. The latter is connected to the word line by vertical W plugs shown in the upper part of the micrograph. The cell connection to the bit line is not shown in this cross section.

DRAM Cell Array Organization

The DRAM cell array may be organized as indicated in **Figure 9**. The bit line is split into two identical

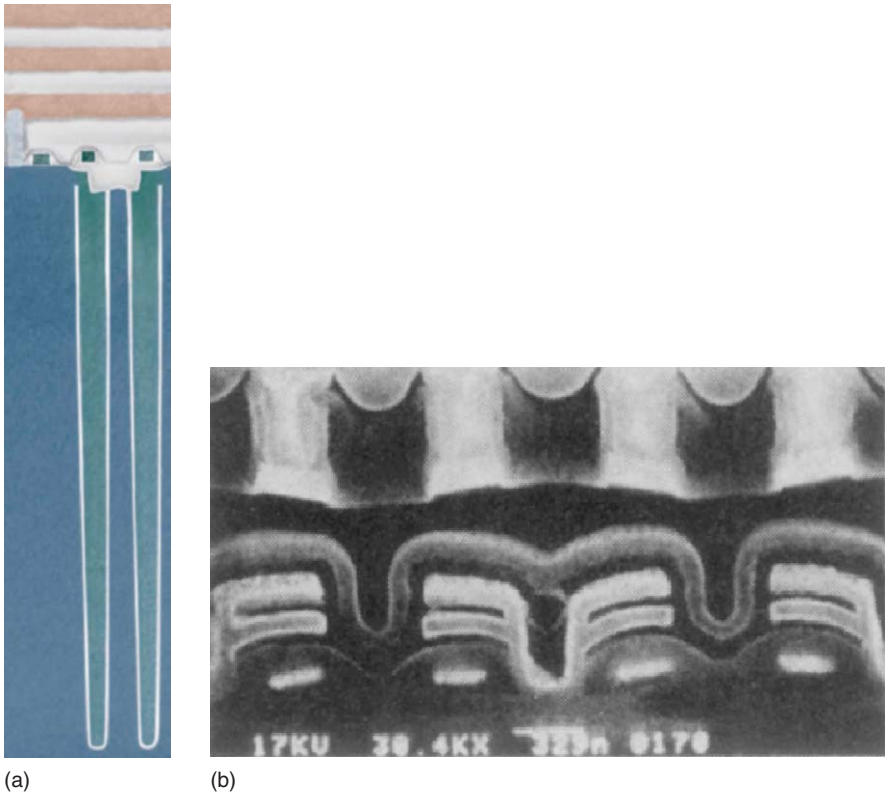


Figure 8 Electron micrographs of DRAM cells. (a) Trench cell (courtesy of IBM Corporation), and (b) stacked cell. (Courtesy of Fujitsu Ltd.)

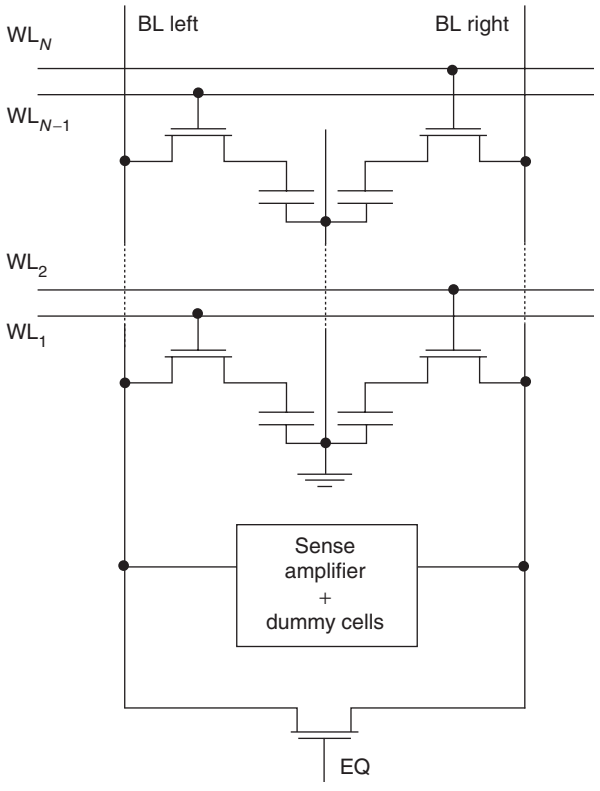


Figure 9 Folded bit-line organization of the DRAM cell array.

segments connected to the sense amplifier, which allows for a differential reading of the stored information and halves the bit-line capacitance for a given column size. In the folded bit-line architecture, shown in the same figure, both segments of the bit line are laid out on the same side of the sense amplifier. The bit-line folding makes it easier to match the layout of the sense amplifier onto the wider pitch of the two half-columns, and allows for an easy placement of the column decoders.

The bit lines are initially precharged at $V_{DD}/2$ by turning on transistor EQ, and are then kept in a high impedance state. As for SRAMs, the row decoder addresses a row of cells and turns the corresponding pass transistors on, thus connecting the cell capacitors to their respective bit lines. Due to the large capacitance ratio between the parasitic bit line and the cell capacitances C_{bl} and C_c, respectively, the bit-line voltage is only slightly affected. More specifically, the voltage change ΔV_{bl} of the bit line connected to the cell equals

$$\Delta V_{bl} = [C_c/(C_{bl} + C_c)](V_c - V_{DD}/2)$$

where V_c is the cell voltage, which may be either 0 or V_{DD}. If the ratio $C_o/C_{bl} \approx 0.1$, the voltage change ΔV_{bl} is only a few hundred millivolts. Due to an unavoidable offset of the sense amplifier, the cell capacitance must be as large as possible, while being dimensionally small to allow for a high bit density.

It is clear from the previous expression that the bit-line voltage increases with respect to $V_{DD}/2$ if the cell voltage $V_c = V_{DD}$, and decreases if $V_c = 0$. The second input of the sense amplifier is connected to a reference voltage generated by a dummy cell precharged at $V_{DD}/2$, as the bit lines. Therefore, the properties of the sense amplifier must be the following: (1) It must detect and amplify a small voltage difference between the two bit lines, eventually generating logic levels and (2) upon reading, it must refresh the content of the cell by reestablishing the original value of the charge stored within the capacitor.

DRAM Sense Amplifier

The schematic of the DRAM sense amplifier is reported in **Figure 10**. Here M_1–M_4 form a regenerative, bi-stable circuit connected to ground via M_5 and to V_{DD} via M_6. An additional pass transistor M_7 may either connect or isolate the two bit lines, thus enabling their precharge to $V_{DD}/2$. Two dummy cells are located on the opposite sides of the sense amplifier and their capacitor is precharged to $V_{DD}/2$ by a suitable circuit: their function is that of balancing the two bit lines by letting them undergo the same transient conditions before readout.

At the start of a reading, SE is low, keeping M_5 and M_6 off; the bit lines are precharged to $V_{DD}/2$ and left in a high-impedance state. Thus, the cross-coupled transistors M_1–M_4 are self-biased on their respective thresholds, and the voltages V_3 and V_4 set at $V_{DD}/2 - V_{Tn}$ and $V_{DD}/2 + |V_{Tp}|$, respectively, V_{Tn} and V_{Tp} being the threshold voltages of the *n*- and *p*-channel MOSFETs, respectively.

The reading sequence is the following: (1) the word line of the addressed cell is activated, thus connecting its capacitor to the corresponding bit line, and so is the word line connected to the dummy cell on the opposite side of the addressed cell, (2) SE is gradually raised high, allowing the sense amplifier to unbalance and reach its final state, (3) the sense amplifiers are readout via a column multiplexer, (4) the word lines are reset to their low value, (5) the signal SE is switched low, thus turning off M_5 and M_6, (6) the bit lines are shorted by raising the signal EQ, which drives them and the dummy-cell capacitors to $V_{DD}/2$, and (7) M_7 is turned off by lowering EQ and leaving the bit lines in a high-impedance state. So doing the initial reading conditions of the sense amplifier are reset.

It may be worth pointing out that, following the gradual raise of SE in step 2, M_5 and M_6 turn on lowering V_3 and raising V_4. The cross-coupled

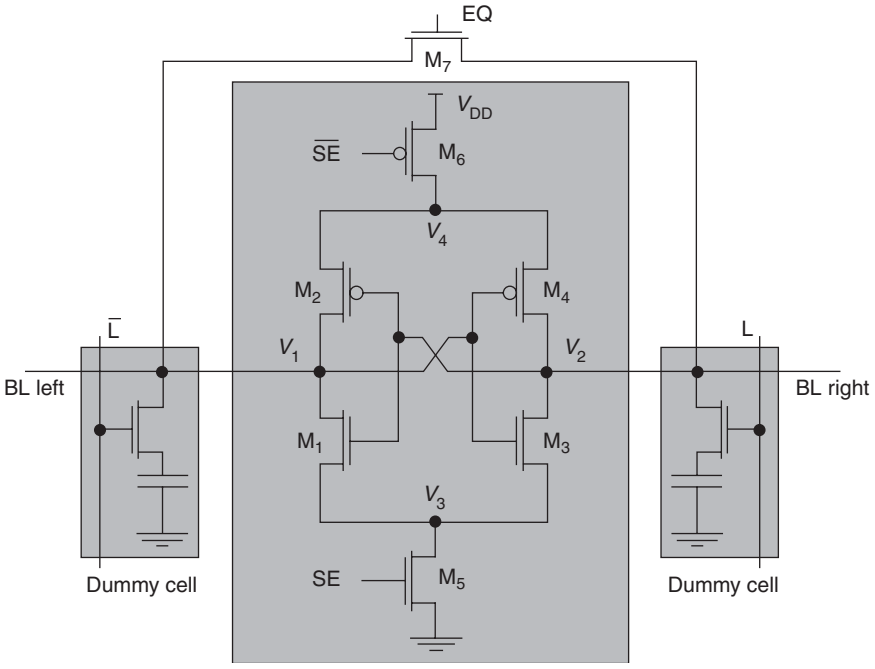


Figure 10 Schematic depiction of the DRAM sense amplifier.

transistors M_1–M_4 turn on and unbalance the sense amplifier according to the initial voltage mismatch. If the initial bit-line voltage $V_{bl}^{(left)} > V_{DD}/2$, which reflects the content of a logic "1" in the addressed cell, the transient drives it to V_{DD} while the right bit-line voltage $V_{bl}^{(right)}$ goes to zero. The opposite situation occurs if $V_{bl}^{(left)} < V_{DD}/2$, which reflects the content of a logical "0" within the same cell. Thus, reading the addressed cell restores the original voltage within the cell capacitor, which makes the reading nondestructive. Also, if the two bit-line capacitances are equal, precharge to $V_{DD}/2$ is simply achieved with no additional power consumption by connecting them via M_7.

DRAM Architecture

The block diagram of a DRAM bank is shown in **Figure 11**. The row addresses are delivered to the row decoders, which activate one word line. All the cells associated with that word line are thus connected to the bit lines, and their content is readout by the sense amplifiers. The column addresses set the column decoder (actually a simple multiplexer) and select the requested bytes or words, which are delivered to the output registers. Multiple column readings can be performed in sequence for any given row address. This improves the data rate of the DRAM. The above structure may be mirrored and replicated as many times as needed. The size of a DRAM bank is defined according to conflicting constraints: increasing the size of the cell array makes the layout more compact for a given DRAM capacity, but deteriorates the access time and, ultimately, the reading integrity, due to the larger word- and bit-line capacitances.

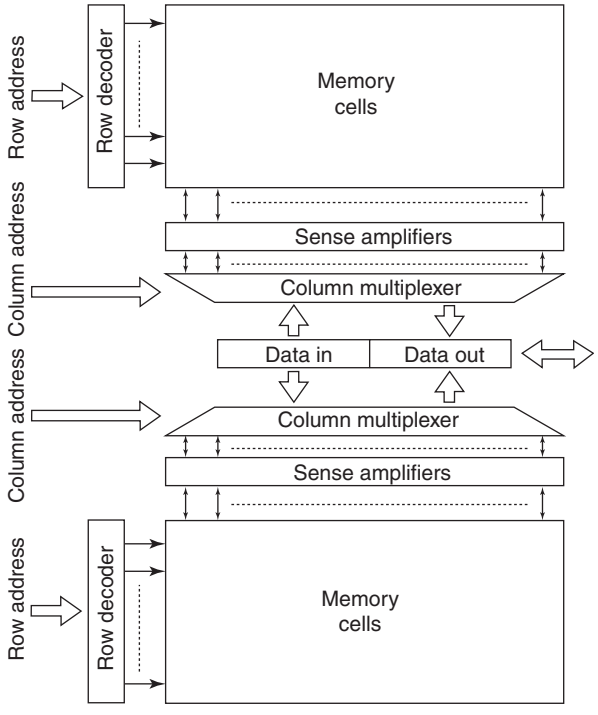


Figure 11 DRAM bank architecture.

In order to save package pins, a multiplexed addressing scheme, where the row and column addresses are presented in sequence to the same bus, is typically used. In this addressing scheme, two main control

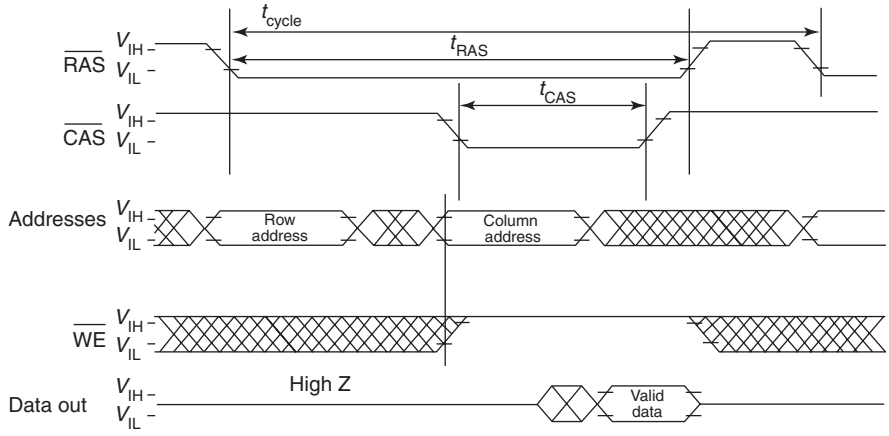


Figure 12 Timing diagram of the DRAM input signals.

signals must be provided to the DRAM, namely, RAS and CAS. Another control signal, write enable WE, indicates if the intended operation is a read or a write.

The timing diagram of a read operation is shown in **Figure 12**, where the control signals RAS′, CAS′, and WE′ are shown in complementary form. The RAS′ signal is switched low only after valid row address bits are present at the DRAM input buffer. In turn, the CAS′ signal goes low only when valid column address bits and a valid WE signal are set to their respective values. The time t_{CAS} is the minimum time that CAS′ must be kept low to generate valid output data; the time t_{RAS} is the minimum time that RAS′ must be kept low to generate valid output data; finally, the cycle time t_{cycle} is the minimum time needed between two successive RAS′ commands.

Synchronous DRAM

Early DRAM architectures used to operate asynchronously under the control of the external commands outlined in the previous section. To date, most DRAM chips are synchronous devices driven by the system clock, and are thus referred to as SDRAMs. Synchronous operation has several advantages: most notably, the possibility of a pipelined DRAM architecture with concurrent row and column addressing, and high-speed I/O operations. Taking advantage from the inherent parallelism of the memory core, this architecture improves the DRAM data rate, which is the most important performance parameter, while leaving its latency scarcely affected. Due to the hierarchical memory organization in modern computers, entire data blocks are, in fact, retrieved from the central memory when a block miss occurs at a higher level of the hierarchy. To date, SDRAM data rate may be as large as $1.6\,\mathrm{GB\,s^{-1}}$. This comes at the penalty of extra latches and buffers, as well as high-speed circuitry to support the I/O interface.

SDRAM commands, addresses, and data are latched to the rising edge of the system clock. Basic SDRAM commands are chip select (CS), RAS, CAS, WE, data mask (DM), and data strobe (DQS). The CS signal is used to let the chip know that the commands coming in over the bus are intended for it. RAS, CAS, and WE retain the usual meanings of row and column address strobe and write enable, respectively. DM masks the input data during a write; DQS is instead a bi-directional edge aligned data strobe which toggles at the same time as the output data. This is made possible by a delay locked loop (DLL), which shifts the output data in order to align DQ and DQS. The bus width is most often 64 bit.

The basic SDRAM operations are: activate (ACT), read (RD), or write (WR) followed by a precharge. SDRAM operation can be configured for CAS latency and burst length by setting the 12 bits of the load mode register (LMR). The burst length determines the amount of data transferred in consecutive cycles between the memory controller and the memory after applying one start address.

The typical delay between the RAS and CAS signals is two clock cycles, and the CAS latency is again two clock cycles; thus, the access time is four clock cycles. The SDRAM cycle time t_{cycle} depends, in this case, on the bus width and the burst length.

Rambus DRAM

The direct RDRAM architecture is based on a system approach, which maximizes the data rate on the memory board by carefully synchronizing data, addresses, and commands with the clock signal. The key elements of this technology are the following: (1) a 16-bit wide, 800 MHz channel, (2) an on-board interface that allows the memory controller talk to the RDRAM and, (3) an in-line memory module called RIMM.

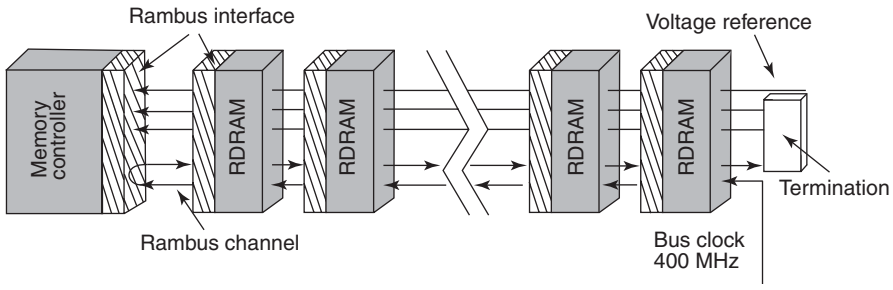


Figure 13 Direct Rambus DRAM architecture.

A Rambus channel includes a controller and a variable number of RDRAM chips connected together via a linear common bus. As shown in **Figure 13**, the controller is located at one end, and the RDRAMs are distributed along the bus, which is terminated at the opposite end with the characteristic impedance of the lines connected to the high-voltage level. Therefore, the bus driver operates in the open-drain configuration. The 400 MHz clock signal is generated at this same end; it propagates along a bus line connecting the clock generator to the controller, and turns back to the terminating end of the bus line. All commands, addresses, and data moving from the controller to the RDRAM chips and back are synchronized on the edges of the clock which propagates in the same direction. In doing so, the clock skew is minimized.

The channel uses 18 data pins, two of which for error correction code (ECC), cycling at $800\,\mathrm{MB\,s^{-1}}$ per pin to provide a bandwidth of $1600\,\mathrm{MB\,s^{-1}}$. This is achieved by latching data on both the rising and the falling edges of the clock signal.

The RDRAM has a pipelined microarchitecture, which fully supports concurrent RAS and CAS operation, as well as read/write buffering. Also, it provides 16 bytes every 10 ns on the internal bus. The Rambus interface transforms the 10 ns internal bus into an external, two-byte wide, 1.25 ns external bus.

All signal wires have an equal loading and fan-out and are routed parallel to each other on the top trace of a PCB with a ground plane located underneath. The addition of more RDRAM chips linearly increases the load and the delay, but leaves the phase relationships of the clock, data, addresses, and commands unaltered. Also, the memory granularity is simply given by the memory capacity of one chip.

Content-Addressable Memory

As already stated in the introduction, modern computers rely on a hierarchy of memories comprising three levels of cache in addition to the central memory. This complex organization is due to the design consideration that smaller memories are faster, and to the locality principle, according to which programs tend to reuse either data which have already been used in previous cycles (temporal locality), or data sets stored next to the used data (spatial locality).

Cache memories are classified according to the criteria by which the central memory is mapped onto it. In a fully associative cache, data blocks can be stored anywhere within the cache, as opposed to direct-mapped memories, where the block address uniquely defines its location. In most cases, cache memories are *n*-way set associative; this means that the cache memory is split into several sets, each containing *n* blocks: the block address uniquely defines the set where it must be stored, but the location of the block inside the set (way) is free.

In any case, a tag memory containing a list of the block addresses stored within the cache is needed. When a new instruction or a new data is requested by the processor, an associative search must be performed on the tag memory in order to find out if the block containing that instruction or data is stored within the cache and, in case it is, in what location. As opposed to standard memories, the input of the tag memory is a data and its output is the local block address within the cache. The tag memory is, therefore, a CAM. The basic cell of a CAM, shown in **Figure 14**, contains ten transistors.

The information bit is stored within a static latch made by the two cross-coupled inverters I_1 and I_2; two pass transistors M_1 and M_2 driven by the write line (WL) connect the latch outputs to the input data D and its complement D′. Two couples of pass transistors M_3–M_4, M_5–M_6, driven by the input data D and D′ and by the latch outputs as shown in the figure, connect to ground a precharged match line (ML). If the stored bit is opposite to the data D, either of the two couples of pass transistors is on and discharges the match line. Otherwise, the match line remains biased at V_{DD}. The former case indicates that no match occurs between the input and the stored data; the latter indicates the opposite condition.

The above cells are organized within an array, and all the cells of a column, sharing the same write line

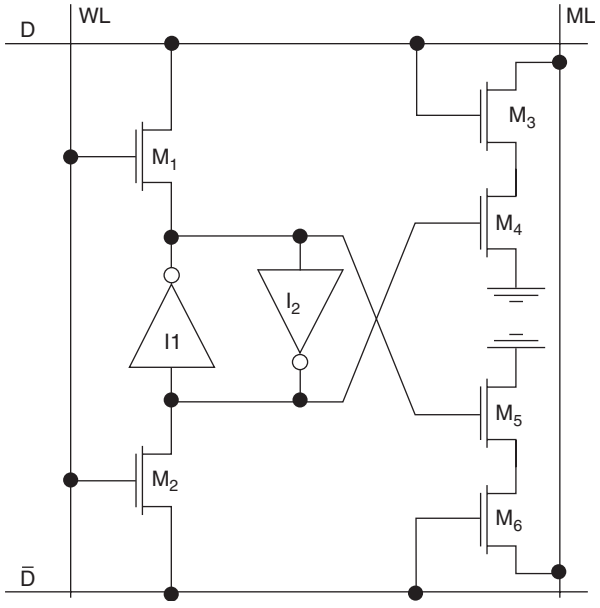


Figure 14 The ten-transistor CAM cell.

and the same match line, contain the address of one memory block. If any one of the above cells does not match its input bit *D*, it drives the match line to ground, indicating that no match occurs between the input word and the stored address. If none of the column matches the input word, this means that a block miss is occurring; all the match lines are at ground potential and an NOR gate raises the block-miss signal high.

If, on the other hand, all the input bits of a column match their respective bits stored within the cells, the match line remains high. This is an indication that the requested block is stored within the cache, and the high match line indicates the internal address of the block within the cache memory.

See also: Conductivity, Electrical; Electrons and Holes; Integrated Circuits; Memory Devices, Nonvolatile; Semiconductor Devices; Transistors.

PACS: 85.25.Hv; 84.30.−r; 85.30.Tv; 84.30.Bv; 84.30.Sk

Further Reading

Haraszti TP (2001) *CMOS Memory Circuits*. Dordrecht: Kluwer.

Prince B (1999) *High Performance Memories: New Architecture DRAMs and SRAMs – Evolution and Function*, Rev. edition. Chichester: Wiley.

Sharma AK (2002) *Advanced Semiconductor Memories: Architectures, Designs, and Applications*. Hoboken: Wiley-IEEE Press.

Meso- and Nanostructures

A Horsfield, University College London, London, UK

Introduction

Meso- and nanostructures are "small" structures. It is possible to assign a rough length scale to them (micrometers and below), but it is more instructive to define smallness in terms of the physical phenomena that it makes possible, as these are what make small structures important. In the following section, selected phenomena are briefly described, together with the principal characteristics that structures need in order to exhibit them. In the section "Contemporary structures", some contemporary structures in which these phenomena can be observed are listed. These structures are organized into categories according to the number of their dimensions which are not small. New structures are being invented constantly, so this list should be seen as illustrating possibilities rather than being comprehensive. However, some structures – such as carbon nanotubes – will remain important for the foreseeable future.

Phenomena

The following phenomena are all a consequence of smallness in one or more dimensions, though they vary in the way smallness is exploited. Many make use of the wave nature of the electron, while others (notably Coulomb blockade) have a classical origin.

Continuum Electronic States Become Discrete

In an infinite solid, the electronic states form continuous bands of states (Bloch states). When the system is large but finite, the states are discrete, but have very small energy separations, so that a continuum description remains accurate. As the size of the system is reduced, the energy spacing between levels increases roughly as $1/L^2$, where L is the size of the system (quantum confinement). When the spacing is large enough (the system size is small enough) that it exceeds any broadening (e.g., thermal), then this discreteness can be exploited. In order to obtain discrete states, the structures need to be small in three dimensions with weak coupling of the discrete states to the environment (e.g., quantum dots). Thus, nonlinear

optical effects such as photo-induced blue shifts (electron excitation leads to modification of the electronic structure, and hence to a changed response to photons) are then possible. If the system is placed between two macroscopic metallic leads, then the discrete states become resonances and resonant tunneling can be observed.

Small semiconducting crystals in solution can also be used for the direct conversion of light into chemical or electrical energy. Absorption of a photon produces both an electron in an excited state and a hole. The excited electron can be donated to a molecule (reduction), or an electron can be extracted from a molecule into the hole state (oxidation). This process can be used to create free radicals (such as $OH^{\bullet}$), and thus catalyze reactions.

Individual Electrons Become Detectable

For very small devices, the transfer of one electron from the environment onto the device can change the combined energy of the device and its environment substantially. For a charge transfer δQ, the total energy change can be approximated by $\delta E = (V_D - V_E)\delta Q + \delta Q^2/2C_D$, where V_D and V_E are potentials for the device and environment, respectively, and C_D is a capacitance for the device. If this energy change is positive, then energy must be provided from outside for the charge transfer to take place. This energy might be provided through thermal excitation, in which case the transfer will not occur unless $k_B T > \delta E$, where T is the temperature of the environment and k_B is the Boltzmann constant. If the potential of the environment can be adjusted, then the transfer of charge can always be made possible by ensuring that $V_E > V_D + \delta Q/2C_D$. The blocking of charge flow through a small device (Coulomb blockade) has been observed in both metals and semiconductors. No current flows until a large enough bias $(V_E - V_D)$ is applied to overcome the charging energy of the device. To observe this phenomenon, the device must have a sufficiently small capacitance $(\delta Q^2/2C_D > k_B T)$. Quantum dots can provide this.

Ferromagnets Have Only One Domain and a Low Anisotropy Energy

Macroscopic ferromagnets acquire domains in order to minimize the energy due to the magnetic field. When a magnetic field is applied, the atomic magnetic moments align through movement of the domain walls. If the size of the ferromagnet is smaller than that of one domain, then alignment with the magnetic field can only occur by the concerted rotation of all the moments. There is an energy penalty for this (anisotropy energy). However, the smaller the ferromagnet, the lower the anisotropy energy. Thus, if the magnetic particle is small enough, then a collection of them can behave as a paramagnet with a very large susceptibility (superparamagnetism). If the particle is elongated, it forms a single domain, while if it is circular it tends to form a vortex state.

Excitons Bind More Strongly

In semiconductors and insulators, an electron can be excited from the valence band into the conduction band, creating a hole in the valence band and leaving an electron in the conduction band. The hole and electron can bind to form a new quasiparticle (exciton). In small structures, the electrons and holes have their movement restricted, which, together with the change in electron and phonon spectra, leads to increased binding of the exciton. This means that in small structures, many excitons can be observed, even at high temperatures (such as room temperature). They have been seen in quantum wells and quantum dots.

The Wavelength of the Electron Strongly Influences Conduction

This is most easily seen in "quantum conductance." In the effective mass approximation, with coherent states, the wave functions (or, more exactly, the envelope functions) and energies for electrons in a one-dimensional conductor are $\psi_k(x) = e^{ikx}/\sqrt{L}$ and $\varepsilon_k = \hbar^2 k^2/2m^*$, respectively, where L is the length of the wire. States with $k > 0$ carry electrons from left to right, while those with $k < 0$ carry electrons from right to left. When more states with wave vectors of one sign are populated with electrons than those with the opposite sign, a net electric current is achieved. If right traveling states with energies $\varepsilon_k < \mu_R$, and left traveling states with energies $\varepsilon_k < \mu_L$ are populated, the net current is $I = 2e(\mu_R - \mu_L)/h$. As the applied bias is $V = (\mu_R - \mu_L)/e$, the quantum unit of conductance $\sigma = I/V = 2e^2/h$ is obtained. For a three-dimensional wire with square cross section of length l, the wave functions and energies are

$$\psi_{nmk}(\boldsymbol{r}) = \frac{2}{\sqrt{Ll^2}} e^{ikx} \sin\left(\frac{n\pi}{l}y\right) \sin\left(\frac{m\pi}{l}z\right)$$

$$\varepsilon_{nmk} = \frac{\hbar^2}{2m^*}\left(k^2 + \left(\frac{n\pi}{l}\right)^2 + \left(\frac{m\pi}{l}\right)^2\right)$$

where n and m are positive integers. As the chemical potential of the electrons is increased from zero, the band with $n = m = 0$ is first filled. Once the chemical potential exceeds $\hbar^2\pi^2/(2m^*l^2)$, two further bands ($n = 1$, $m = 0$ and $n = 0$, $m = 1$), and so on are filled. These partially filled bands carry a current, each

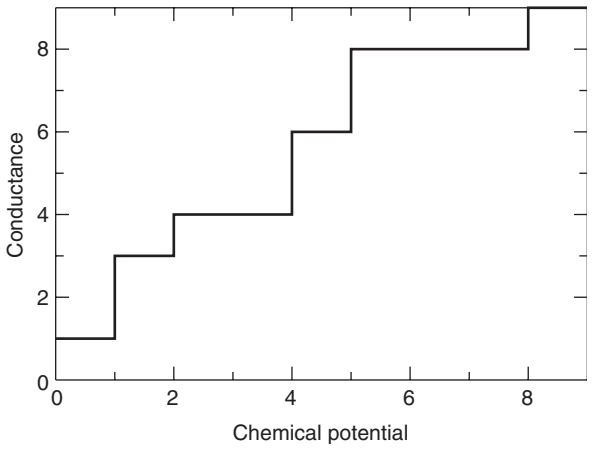


Figure 1 The change in differential conductance in units of $2e^2/h$, vs. the mean chemical potential of the electrons in units of $\hbar^2\pi^2/(2\,m^*l^2)$, for the model system described in the text.

having the same conductance. Therefore, a conductance of the form shown in **Figure 1** results. To observe the quantum conductance, the temperature of the system must be much smaller than the characteristic change in the chemical potential between the conductance steps (the narrower the channel through which the conduction occurs, the higher the temperature can be); the conduction needs to be ballistic, so the length of the conduction channel (L) needs to be less than the electron mean free path.

Electrical Resistance Is Not Proportional to the Length of the Wire

For macroscopic systems, doubling the length of a wire doubles its resistance. This is because, the scattering of electrons by defects can be considered as an incoherent process (one scattering event does not interfere with the next one). For small systems in which the electron wave function remains coherent, there can be interference between electron waves scattered off different defects, and this can have the effect of limiting the range of the wave function of the electron (localization). As a result, the resistance from two pieces of wire joined together is not just the sum of the two individual resistances, but includes an extra term which is the product of the two resistances. If, on the one hand, the wire is long compared to the localization length of the electron wave function, then the resistance grows exponentially with the length of the wire (strong localization). If, on the other hand, the wire is short compared to the localization length, then the resistance deviates from the macroscopic result by a term that is quadratic in the length of the wire (weak localization). In terms of the conductivity, the correction for weak localization is of the order of the quantum of conductance.

Quantum Interference Becomes Measurable

If two leads are attached to a small conducting ring, and a magnetic field is applied perpendicular to that ring, then the conductance of the system will vary as the magnetic field is adjusted (Aharonov–Bohm effect). The dependence of conduction on magnetic field is $G(B) = G_0(1 + \cos(eBS/\hbar + \phi))$, where e is the electronic charge, B is the magnetic field strength, S is the area enclosed by the ring, and ϕ is a phase angle. To observe this phenomenon, there must be ballistic transport throughout the ring. Thus, the ring must be small enough so that only a small amount of inelastic scattering is experienced by the electrons.

Contemporary Structures

Some contemporary meso- and nanostructures are described below. They have been selected because they are of interest at the moment. They represent the kind of structures that exhibit interesting behavior on account of their size. New structures are constantly being developed, so the following can only be considered as a snapshot.

Zero Large Dimensions

Nanostructured materials These materials are formed from a large number of very small microcrystals, and have a very fine microstructure (the grain size can be as small as 1 nm). This can produce improved mechanical properties. The nanostructured materials are, therefore, being investigated as materials for medical implants. They can be made by consolidating microcrystals with clean surfaces. This process is used to make ceramics for which the rate of densification varies inversely with particle size, leading to reduced sintering times. These structures can also be made by deposition of atoms or molecules, using chemical or physical vapor deposition, or by crystallization and precipitation. A nonchemical approach is to start from a metal with much larger grains and introduce further grain boundaries by severe plastic deformation (e.g., by ball milling).

Colloidal semiconducting nanoparticles These nanostructures are small crystals (radius < 100 nm) that are suspended in a fluid. They have been made from many semiconducting materials using chemical precipitation, hydrolysis, etching of silicon wafers, radiolysis, and sonolysis. The size of these nanocrystals can be controlled by chemical means such as the use of surface stabilizers. The crystalline structure defines the electronic properties (such as effective mass), and is influenced by the preparation conditions such as temperature and pressure. These nanoclusters contain

defects (which also depend on the method of synthesis) which can trap electrons and holes and increase their recombination rates. This can be probed using photoluminescence.

Dyes can be attached to the surface, and then pass charge to the nanocrystal once they undergo photoexcitation. This charge transfer can be extremely fast, taking as little as a few hundred femtoseconds, while giving a quantum efficiency close to 1. The addition of dyes allows low-energy excitations to take place. The selectivity and efficiency of photochemical reactions can also be improved by coating the surface with a noble metal. For example, for the photocatalytic production of hydrogen from water, a platinum coat can act as a sink for photogenerated electrons which then catalyze the production of hydrogen.

Nanocrystals can self-assemble to form superlattices with the gaps between dots being controlled by capping agents. Templates, such as liquid crystals, can be used to assist the alignment of the nanocrystals.

Strained layer islands Self-assembly allows one to create nanostructures without processing. The advantage is the absence of damage caused by processing, while the disadvantage is the limited amount of control one has over the structures that are formed. One particularly interesting form of self-assembly is the creation of small islands (or huts) during epitaxy that can act as quantum dots. They are formed during the deposition of one material (the epilayer) on another (the substrate) when the two materials have slightly different lattice constants.

In order to understand why islands form, consider the following argument. If the two materials are lattice matched so that there is no strain at the interface, then as the material is deposited to form the epilayer, there are two possible outcomes. If it is energetically favorable for the deposited material to cover the substrate, then a uniform epilayer will form (Frank–van der Merwe growth). On the other hand, if there is an energy penalty, then the substrate will remain exposed to the vacuum and the deposited material will tend to form islands (Volmer–Weber growth). If the materials have different lattice constants, there will be a strain at the interface which leads to an excess energy that increases with the thickness of the epilayer. Once one layer of material has been deposited onto the substrate, the next layer will be deposited onto this. Additional material will go down as a uniform layer until adding further material increases the strain energy so much that it exceeds any energy gain from forming a new layer. Subsequent material goes down to form islands (Stranski–Krastanov growth). This mechanism applies to materials whose lattice constants differ only by a few percent.

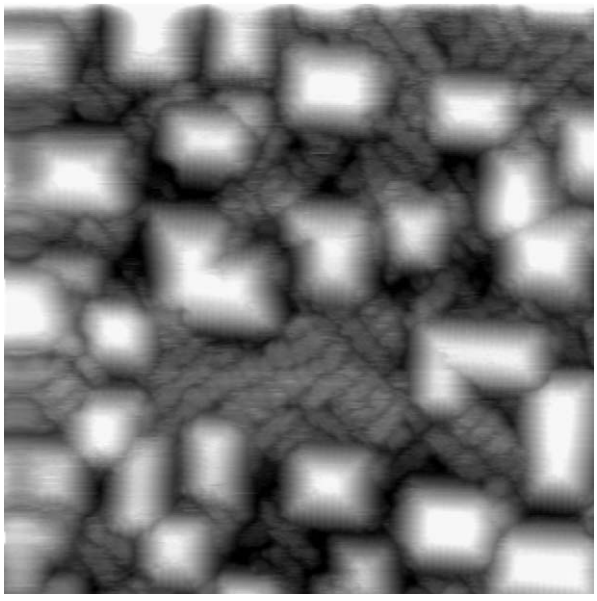


Figure 2 Germanium hut structures grow on a silicon surface. (Courtesy of J Owen.)

Because of the strain, these islands are lattice matched to the substrate. For given materials and deposition conditions, the islands have fixed shapes and sizes. The islands have flat tops and sloping sides (see **Figure 2**). As more material is added, more islands are produced. The size is governed by kinetics and so is strongly influenced by temperature and deposition rate. If material continues to be added after the maximum density of islands has been reached, then dislocations form and the islands grow in size. These large islands are no longer lattice-matched with the substrate.

If a layer of islands is covered with a buffer layer and then more of the island material is deposited on top of that, further islands will grow that lie directly above the first set of islands. This is because of strain in the buffer layer induced by the first layer of islands.

One Large Dimension

Porous silicon Porous silicon can be made by etching crystalline silicon with hydrofluoric acid. Fine tubular structures are formed, with columns having widths less than 5 nm being possible. This is a very fragile material and can degrade in $\sim$1 h under ambient conditions. These narrow columns result in confinement of the electrons, which leads to intense luminescence. Passivation with hydrogen improves the luminescence, which suggests that surface defects may play a role.

Carbon nanotubes For the purpose of this article, Buckminster fullerenes such as C_{60} are considered as

very short nanotubes. Carbon nanotubes can be considered as sheets of graphite rolled up into tubes, possibly with the ends of the tubes capped. Nanotubes are a very important form of nanostructure because of the exceptional properties they have, which are given as follows:

- they are chemically very stable;
- they are mechanically strong and are able to recover from extreme deformations;
- they provide coherent electron transport; and
- the electronic structure can be tuned by varying the helicity and diameter of the tube.

Nanotubes can be single-walled (SWNT) or multiwalled (MWNT). An SWNT consists of just one graphite sheet rolled into a tube. The typical radius of the tube is between 1 and 2 nm. Tubes tend to pack together into triangular lattices with a lattice parameter of 1.7 nm. Similarly, an MWNT consists of several coaxial graphite tubes, with the separation between the walls being 0.34 nm, which is close to the interlayer spacing in graphite.

Nanotubes can be either metallic or semiconducting, depending on the helicity and diameter of the tubes. Simple models that neglect the curvature of the tubes can give a good idea of trends, though more accurate calculations are needed to obtain precise results. Calculations (supported by experiment) show that about one-third of the tubes are metallic while about two-thirds are not, and that for semiconducting tubes, the bandgap varies inversely with the diameter of the tube.

Nanotubes are very stiff (MWNT have a Young's modulus of ~ 1.8 Tpa), and can recover reversibly after severe deformations. This makes them ideal for applications such as tips for surface probe microscopes (such as scanning tunneling microscopes and atomic force microscopes).

It is possible to insert material inside a nanotube. Strong capillary forces allow fluids to be held within. They have been filled with molten salts such as metal oxides, sulfides, and chlorides. Currently, there is interest in the so-called peapod structures in which C_{60} and other Buckminster fullerenes are inserted within nanotubes.

Carbon is not the only material that can form such tubes; some clays (which are layered silicates) can also have a tubular form (e.g., the aluminosilicate imogolite). MWNTs of WS_2 and MoS_2 in which there are alternating layers of metal and sulfur have been made. They had diameters of up to a few tens of a nanometer and had lengths of a micrometer. Both SWNTs and MWNTs have been made from boron nitride (BN). Tubes made from BN are always semiconducting (with a large gap) independent of tube diameter or helicity.

Quantum wires There are many structures that fall into this category. Here three types of quantum wires are considered.

The (0 0 1) surface of silicon has long rows of atoms that are paired up to form dimers. The atoms within the dimers are very reactive because each one is only bonded to three other atoms while it would like to be bonded to four (they each have one dangling bond). Hydrogen can be used to satisfy this bonding requirement: one hydrogen atom will bond to each dangling bond. It is possible to remove (using a scanning tunneling microscope) one hydrogen atom from each dimer, which results in a chain of dangling bonds which behave like a wire. However, the wire is a semiconductor because of the Peierls distortion. It is possible to attach metals and other atoms to the dangling bonds. The metals tend to form clusters, which grow into wires as more metal is deposited. These wires are generally of poor quality.

It is possible to grow self-assembling wires on the silicon surface. For example, if rare-earth metals are deposited and then annealed, they will form an alloy with the composition XSi_2. This alloy will have a lattice constant in one direction that matches the silicon, but in a perpendicular direction it will not match. This results in the growth of long straight wires to minimize elastic strain energy. The following metals have been used to produce wires: Dy, Er, Gd, and Sc. In a similar way, bismuth forms nanolines when it is deposited on silicon and then annealed between 570°C and 600°C (see **Figure 3**). The wires are always 1.5 nm wide, and can be several hundred nanometers long. They are never kinked or defective, and are always semiconducting.

Long single-crystal semiconductor wires can be grown using a vapor–liquid–solid technique. This requires a small catalyst (such as gold), the size of which determines the radius of the crystal wire. The catalyst is molten, and causes the material used to make the wire, which is present as a vapor, to bind to the interface between the catalyst and the substrate on which it is sitting. This causes a column of the material to grow underneath the catalyst (See **Figure 4**). This technique has been used to produce wires made of Si, Ge, Si/Ge, GaAs, GaN, InP, CdS, ZnSe, and CdSe.

Two Large Dimensions

Heterostructures Silicon-based heterostructures were originally considered as a way to overcome the problem of the indirect bandgap in silicon. To obtain efficient light emission, a direct bandgap is

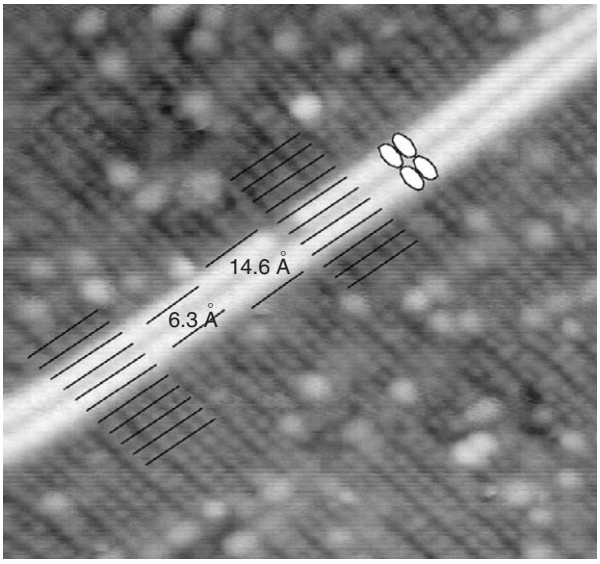


Figure 3 Lines of Bi that are self-aligned on the surface of silicon. (Courtesy of J Owen.)

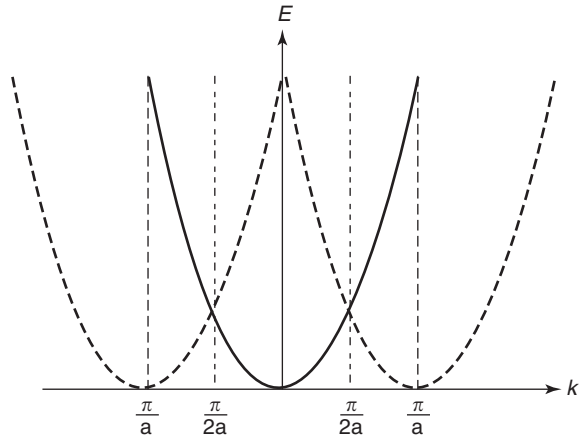


Figure 5 The hard line is the band structure for a one-dimensional system with a lattice constant *a*. Its Brillouin zone runs from $-\pi/a$ to π/a. If the cell size is doubled, then the Brillouin zone runs from $-\pi/2a$ to $\pi/2a$, and the band structure consists of both the hard line and the broken lines. Thus, at the center of the Brillouin zone states that used to be at the zone boundaries are found. This is band folding.

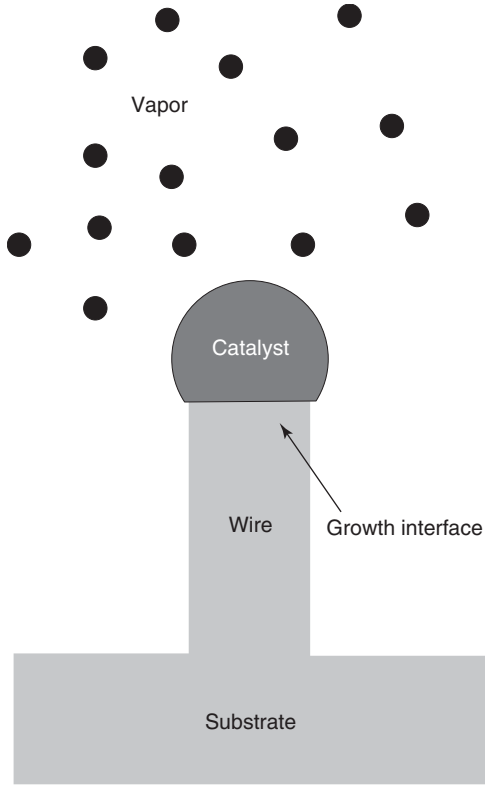


Figure 4 The geometry for the vapor–liquid–solid technique.

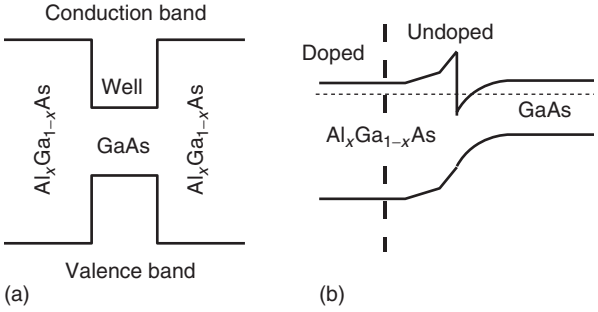


Figure 6 Quantum well band edge profiles. (a) The electron donors are in the GaAs layer which forms a well because of its reduced bandgap compared to the surrounding $Al_xGa_{1-x}As$. (b) The donors are on the left-hand side of the system, and the well is created by the realignment of the bands that takes place to ensure that the electron chemical potential (horizontal dotted line) is uniform everywhere.

required. It was thought that this might be achievable in Si by increasing the size of the unit cell in one direction and by making a layered structure with alternating layers of Si and another material (e.g., Ge). This has the effect of folding the band structure in onto itself, bringing the conduction-band minimum closer to being directly above the valence-band maximum (see **Figure 5**). Unfortunately, increasing the period of the superlattice has the effect of reducing the optical matrix elements; thus, only short period superlattices can be used. However, these heterostructures also produce quantum wells.

Quantum wells Here, the example of GaAs/$Al_xGa_{1-x}As$ is given to illustrate this type of structure. If a GaAs layer ($\sim$10 nm thick) is grown between two thicker layers of $Al_xGa_{1-x}As$, the GaAs layer acts as a two-dimensional layer for electrons and holes. This is because the GaAs layer has a bandgap of 1.4 eV, while the surrounding $Al_xGa_{1-x}As$ has a

bandgap of 1.9 eV (for $x = 0.3$), producing a well depth of 0.25 eV for both electrons and holes (see **Figure 6a**). If during the molecular beam epitaxy growth of the GaAs layer electron donor atoms are included, a two-dimensional electron gas will be produced. The ionized donors scatter the electrons very strongly. This can be remedied by using modulation doping, in which only one interface is needed, and where the donors are outside the quantum well (see **Figure 6b**). With this approach, the scattering can be made so low that the electron mean free path can exceed 10 μm. The dimensionality of the electron gas can be further reduced by attaching contacts through which a bias is applied, and which can squeeze the electrons into a wire (of variable width). This has been used to study quantum conduction.

See also: Carbon Materials, Electronic States of; Epitaxy; Excitons: Theory; Nanostructures, Electronic Structure of; Nanostructures, Optical Properties of; Porous Silicon; Quantum Devices of Reduced Dimensionality; Semiconductor and Metallic Clusters, Electronic Properties of; Semiconductor Heterojunctions, Electronic Properties of; Semiconductor Nanostructures; Small Particles and Clusters, Optical Properties of; Transport in Two-Dimensional Semiconductors.

PACS: 61.46.+w; 61.48.+c; 62.25.+g; 73.22−f; 73.23.−b; 73.63.−b

Further Reading

Bowler DR (2004) Atomic-scale nanowires: physical and electronic structure. *Journal of Physics: Condensed Matter* 16: R721–R754.

Cullis AG and Canham LT (1991) Visible light emission due to quantum size effects in highly porous crystalline silicon. *Nature* 353: 335.

Datta S (1995) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.

Nalwa HS (ed.) (2002) *Nanostructured Materials and Nanotechnology*. London: Academic Press.

Owen JHG, Miki K, Koh H, Yeom HW, and Bowler DR (2002) Stress relief as the driving force for self-assembled binanolines. *Physical Review Letters* 88: 226104.

Poole CP Jr and Owens FJ (2003) *Introduction to Nanotechnology*. Hoboken: Wiley-Interscience.

Rosei F (2004) Nonstructured surfaces: challenges and frontiers in Nanotechnology. *Journal of Physics: Condensed Matter* 16: S1373–S1436. Bristol: Institute of Physics Publishing.

Saito R, Dresselhaus G, and Dresselhaus MS (1998) *Physical Properties of Carbon Nanotubes*. London: Imperial College Press.

Tenne R, Margulis L, Genlt M, and Hodes G (1992) Polyhedral and cylindrical structures of tungsten disulphide. *Nature* 360: 444.

Thess A, Lee R, Nikolaeu P, Dai H, Petit P, *et al.* (1996) Crystalline ropes of metallic carbon nanotubes. *Science* 273: 483.

Thornton TJ (1995) Mesoscopic Devices. In: *Report on Progress in Physics*, vol. 58(3), pp. 311–364. Bristol: Institute of Physics Publishing.

Tischler MA, Collins RT, Stathis JH, and Tsang JC (1992) Luminescence degradation in porous silicon. *Applied Physics Letters* 60: 639.

Treacy MMJ, Ebbesen TW, and Gibson JM (1996) Exceptionally high Young's modulus observed for individual carbon nanotubes. *Nature* 381: 678.

Willig F (1995) In: Miller RJD, McLendon GL, Nozik AJ, Semikler W, and Willig F (eds.) *Surface Electron Transfer Processes*. New York: VCH.

Yoffe AD (2001) Semiconductor quantum dots and related systems: electronic, optical, luminescence and related properties of low dimensional systems. *Advances in Physics*, vol. 50, pp. 1–208. London: Taylor and Francis.

Metallic Bonding and Crystals

D J Willock, Cardiff University, Cardiff, UK

Metals form an important class of technological materials that are exploited widely for their ductility and high electrical and thermal conductivities. The general physical definition of the metallic state is based on these macroscopic properties. In this article, the electronic structure of metals is considered using a simple extension of the linear combination of atomic orbitals used to describe bonding in molecules applied to the solid state. This makes it possible to define the metallic state precisely using the crystal orbitals responsible for bonding and so consider the importance of the metallic state in the cohesive energy of solids and its influence on structure.

The types of bonding discussed in this section of the encyclopaedia can be broadly divided into localized effects (involving only a few atoms in the vicinity of the chemical bond) and delocalized effects (involving cooperative interactions between many atom centers). Localized interactions include covalent and hydrogen bonding while ionic and van der Waals interactions may be classed as delocalized. The concept of the metallic bond is little more difficult to define rigorously and has been the subject of much recent debate. Broadly speaking, the metallic bond can be considered as the sharing of electron density by multiple atomic centers in condensed phases. What allows one to distinguish metallic bonding from other situations is the electronic character of the bond. The classic view of the metallic bond in solids is the "sea of electrons" concept which gives the impression of delocalization

of the electrons in metallic systems. The metallic bond in this picture is an extreme example of the covalent bond and it has been suggested that the metallic bond as a concept can be subsumed into this class of bonding. However, the metallic state requires a combination of delocalized crystal orbitals and a knowledge of their occupation, and is actually a collective property of the entire electron density rather than a few specific states. In this article, these ideas are expanded and the influence of the metallic state of a solid on its stability and structure are considered.

The limit of very weak interaction between the atomic cores and the valence electrons is usually employed to build a model of the electronic states in metals using the nearly free electron approximation. In this article, the opposite approach will be used in which crystal orbitals are built up from atomic orbitals. This gives a better insight into the chemical bonding nature of the metallic state.

Band Structure

The idea of a chemical bond as the sharing of electron density between atoms can be rationalized by the formation of molecular orbitals from atomic orbitals. For example, a hydrogen atom has a $1s^1$ configuration, and when the orbitals from two atoms overlap in molecular H_2 they can combine in two ways. First, overlap between the 1*s* orbitals can be constructive, leading to a buildup of charge between the two nuclei and giving an interaction between the atoms that lowers the energy compared to the atomic state. This favors H_2 formation through the formation of a bonding orbital (σ in **Figure 1**). The alternative is destructive interference between the atomic orbitals in the overlap region. This leads to a lowering of the electron density compared to two noninteracting atoms at the same positions. If this molecular orbital is occupied, a destabilizing contribution to the energy is obtained and the total energy is higher than that of the isolated atom states. This molecular orbital is, therefore, antibonding in nature (σ^* in **Figure 1**). In molecular H_2, each atom supplies a single electron and so the bonding orbital is fully occupied (two electrons) and the antibonding orbital is not occupied; so overall the bond is favored.

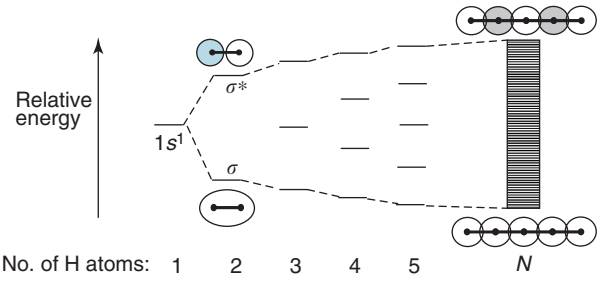


Figure 1 Schematic illustration of the buildup of a band of states from the aggregation of H atoms. *N* is the number of atoms in a bulk sample and will be of the order of the Avagadro constant ($6.022 \times 10^{23}\,mol^{-1}$). Insets show the molecular orbitals for the H_2 molecule and for the top and bottom of the condensed phase band of states.

Figure 1 also shows a schematic representation of the interactions of atomic orbitals in larger aggregations of H atoms. The difference in energy of the molecular orbitals and atomic orbitals is influenced by two main factors. First, the degree of overlap between atomic orbitals: the more the overlap, the bigger the perturbation from the atomic state. Second, the energy of the two atomic orbitals: when these are well matched, the difference between the energies of the molecular and atomic states is greater than when they are very different in energy. In the case of the H_2 molecule, the orbital energies for the atomic states are identical and so this is optimal. If a third H atom is included, the additional *s*-orbital will interact only weakly with the molecular orbitals of the H_2 molecule since the bonding and antibonding states are well separated in energy from the atomic state of the new atom. As more atoms are introduced to produce the one-dimensional (1D) solid, the same number of molecular orbitals is formed as 1*s* orbitals that are supplied, but the energy range for the set of orbitals increases only slowly with the number of atoms. Making the leap to the 1D solid with some large number of atoms (*N* in **Figure 1**), there will be *N* crystal orbitals spread over an energy range not too dissimilar to the gap between the H_2 bonding and antibonding orbitals. There are now too many crystal orbitals to illustrate, but the extreme cases of all bonding between neighboring atoms and all antibonding are shown in **Figure 1**. In general, states at the bottom of the band will be stabilizing for crystal formation from gaseous atoms since they are lower in energy than the atomic 1*s* states. The states above the band center in this situation are destabilizing since they have higher energies than the atomic 1*s* states. In this simple example, a filled band has zero contribution to the lattice energy overall.

Tight-Binding Approximation

The approach of building up condensed phase crystal orbitals based on the molecular orbital theory can be given some mathematical structure by defining the linear combinations of atomic orbitals that are present in the lD chain (**Figure 2**).

The appropriate linear combination of atomic orbitals will be

$$\psi_p = \sum_n c_{pn}\chi_n$$

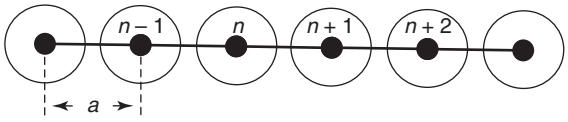


Figure 2 An idealized 1D array of atoms. Each atom contributes a single *s*-orbital to the system. The lattice parameter is shown as the separation between neighbors so that the unit cell contains a single atom.

where c_{pn} is the coefficient for the pth crystal orbital for the local atomic function, χ_n, at site n in the lattice. From the discussion above, if there are N atomic orbitals in the crystal there must be N values of p. The possible values of p can be found by considering the effect of crystal translational symmetry on the electron density.

The density of electrons at any point x along the chain is related to the crystal wave functions via

$$\rho(x) = \sum_{p}^{N_e/2} \psi_p^*(x)\,\psi_p(x)$$

where N_e is the number of electrons to be accommodated in the crystal orbitals. The density of electrons can be observed experimentally using X-ray diffraction, whereas the crystal orbitals are not experimental observables. Translational symmetry implies that the density must be the same at all equivalent points in the crystal. In the 1D lattice with repeat unit, a, the densities at a point x and at any point $x + na$ (where n is any integer) must be the same. The wave functions need not have the same property but this restriction on the density can be met if, for any crystal orbital:

$$\psi_p^*(x + na)\psi_p(x + na) = \psi_p^*(x)\psi_p(x)$$

To achieve this, the wave functions can be written in terms of Bloch states:

$$\psi_k(x) = \sum_n \exp(\mathrm{i}k\,x)\chi_n(x)$$

Here, k is the wave vector for the state, controlling the periodicity of the phase pattern of the crystal orbital along the atomic row and $\mathrm{i} = \sqrt{-1}$. The wave vector is commonly used as the label for the crystal state as has been done above, and it is related to the integer p via

$$k = \frac{2\pi p}{Na}, \quad p = -N/2, \ldots, -1, 0, 1, \ldots, N/2$$

The k-vectors defined in this way have reciprocal length units. The integer p has $N+1$ values; however, the $p = -N/2$ and $p = N/2$ cases are equivalent and so the required N orbitals ($2N$ states) are generated. **Figure 3** shows some illustrative examples of periodic wave functions with different k-vector values.

The allowed values of k in this scheme range from $-\pi/a$ to π/a. Now, another lattice or line of points could be constructed for the values of k that give Bloch functions for the lattice. This line would be related to the real space row of atomic sites but the points on it would give periodicities of waves with reciprocal length units, that is, they are points in reciprocal space. The range of allowed values of k defines the distance in reciprocal space occupied by the unique Block states which are known as the Brillouin zone.

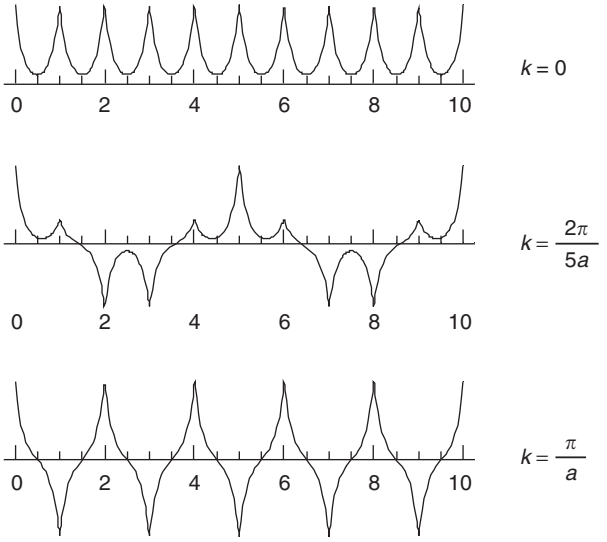


Figure 3 Plots of the real part of a periodic wave function based on Bloch functions. The radial part of each local atomic function is modulated by the periodic function with the *k*-vector indicated. The atomic positions are numbered on the axes.

Using the crystal orbitals in the Schrödinger equation gives the orbital energy as a function of the wave vector k:

$$E_k = \frac{\int_{-\infty}^{\infty} \psi_k^* H \psi_k \,\mathrm{d}x}{\int_{-\infty}^{\infty} \psi_k^* \psi_k \,\mathrm{d}x}$$

Here, H is the Hamiltonian describing the interaction of the electron in the crystal orbital with the line of atoms. The crystal orbitals in terms of Bloch states allow the numerator and denominator of this expression to be written as

$$\int_{-\infty}^{\infty} \psi_k^* H \psi_k \,\mathrm{d}x = \sum_n^N \sum_m^N \exp(\mathrm{i}k(na - ma)) \times \int_{-\infty}^{\infty} \chi_n(x) H \chi_m(x)\,\mathrm{d}x$$

and

$$\int_{-\infty}^{\infty} \psi_k^* \psi_k \,\mathrm{d}x = \sum_n^N \sum_m^N \exp(\mathrm{i}k(na - ma)) \times \int_{-\infty}^{\infty} \chi_n(x) \chi_m(x)\,\mathrm{d}x$$

in which m and n are integers identifying sites along the row of atoms, and now the integrals required are over the local atomic functions. In order to obtain a qualitative insight into the band structure, the simplifying tight-binding approximations can be used which give the integrals required for the numerator values:

$$\int_{-\infty}^{\infty} \chi_n^*(x) H \chi_m(x)\, \mathrm{d}x = \alpha, \quad m = n$$
$$\int_{-\infty}^{\infty} \chi_n^*(x) H \chi_m(x)\, \mathrm{d}x = \beta, \quad m = n \pm 1$$
$$\int_{-\infty}^{\infty} \chi_n^*(x) H \chi_m(x)\, \mathrm{d}x = 0, \quad m < n-1 \text{ or } m > n+1$$

In other words, the on-site term is given by the parameter α, the term corresponding to the nearest neighbors is given by the parameter β, and all other terms are simply ignored. This concentration on local interactions is the reason why the approximation itself is called "tight binding."

Although the interaction contributions are considered in this way, the crystal orbitals are still delocalized along the entire chain of atoms. For the denominator, the tight-binding approximation is even more severe:

$$\int_{-\infty}^{\infty} \chi_n^*(x) \chi_m(x)\, \mathrm{d}x = 1, \quad m = n$$
$$\int_{-\infty}^{\infty} \chi_n^*(x) \chi_m(x)\, \mathrm{d}x = 0, \quad m \neq n$$

Using these integral values in the energy expression gives the formula

$$E_k = \alpha + 2\beta \cos ka$$

so that the energy of a particular crystal orbital depends on the local energy terms and the wave vector of the orbital. For a given system, the local interaction integrals are simply fixed parameters. In general, the relationship between the energy E_k and the wave vector of the state, k, is termed the dispersion relation.

This model can now be applied to the band of states for the 1D chain of hydrogen atoms for which the local functions are the *s*-orbitals on each H atom. When $k = 0$, the exponential factor in the crystal orbital is 1 and so all site orbitals are in phase giving a low-energy contribution. In contrast, when $k = \pm\pi/2$, the complex exponential gives a wave that changes sign between the neighboring atoms and so all neighboring orbitals have an antibonding interaction, a high-energy contribution. **Figure 4** shows the dispersion relation for a line of *s*-orbitals from the tight-binding approximation. The low-energy states are around $k = 0$, whereas for k near its limiting values of $\pm\pi/a$, the energy of the states is higher than the energy for the isolated atomic orbitals, implying that β is negative. The overall bandwidth is 4β, so the strength of interaction between neighboring sites controls the bandwidth.

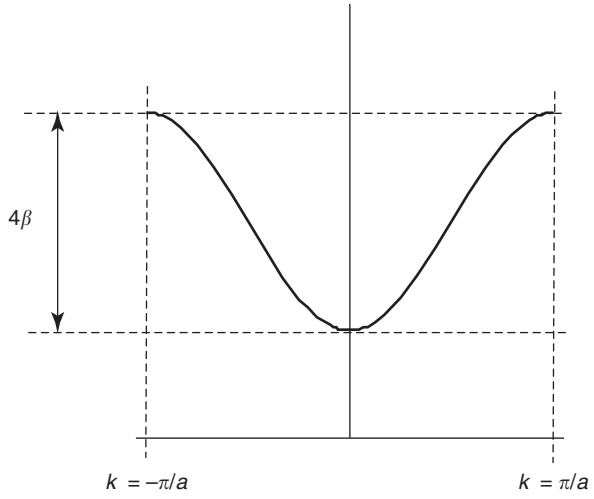


Figure 4 The dispersion relation for a line of *s*-orbitals from the tight-binding approximation.

The values of k that have been defined are evenly spaced in reciprocal space with the allowed values occurring at intervals of $2\pi/Na$. However, the relationship between the energy and the wave vector is not linear. In particular, near the Brillouin zone boundary, the dispersion curve flattens out so that many k-vectors in this region have almost the same energy. In one dimension, a similar effect is seen at the zone center. The density of states (DOS), $N(E)$, is defined as the number of energy levels within a small energy increment as a function of the state energy. In one dimension, the regions in which the dispersion curve flattens out lead to a higher DOS than in regions in which the dispersion curve is steep.

Band Filling and Metallic Behavior

The metallic state can now be defined based on the occupancy of the crystal orbitals. For the line of *s*-orbitals from hydrogen atoms, there are N-crystal orbitals, and the Pauli exclusion principle dictates that each orbital can contain a maximum of two electrons with opposite spins. The crystal orbitals, therefore, give rise to $2N$ electron states, each hydrogen atom delivers one electron and so the states are half-filled. It is assumed that states are simply filled from the lowest energy upwards until all electrons have been placed in crystal orbitals. However, since the separation of states within a band is extremely small, this strict ordering is only possible at 0 K. At any finite temperature, an electron at the top of the occupied levels can be promoted to the unoccupied

states by thermal excitation. The probability of occupancy of a state, $f(E_k)$, is controlled by the Fermi–Dirac statistics:

$$f(E_k) = \frac{1}{\exp((E_k - E_F)/k_B T) + 1}$$

Here, E_F is the Fermi energy which, in metals, corresponds to the energy of the highest occupied state at 0 K, k_B is the Boltzmann constant, and T is the temperature. More generally, E_F is defined from this formula as the energy at which the probability of occupancy is exactly one-half.

Figure 5 shows plots of the Fermi–Dirac probability function at different temperatures. At 0 K it can be seen that the low-energy states are filled up to E_F, but above E_F all states are empty. At higher temperatures, however, electrons from below the Fermi energy can be promoted to higher states, leading to a tail feature. These promotions can only occur if there are states available to receive the electrons that are excited. In the 1D H example, the band of $2N$ states is half-filled and so there are states available for the thermally excited electrons to move into. This is the factor which gives rise to electrical and thermal conduction in metals; filled and empty states are so close to each other that the promotion of electrons at energies close to E_F can occur easily. These electrons, at the interface between the occupied and unoccupied states, are active in transport giving rise to the metallic behavior of the solid. Accordingly, condensed H is expected to be a metal. The construction of the crystal orbitals also shows that the occupied states are those which are largely bonding and so solid H is favored over separated gas-phase H atoms. The liquid state of H has been achieved at low temperature and very high pressure using shock wave pressure methods and is found to have metallic conductivity. However, under most circumstances, $N/2$ diatomic molecules is an even more stable situation and so the diatomic gas is the normal state.

There are a few less exotic examples of solids with 1D aspects to their electronic structure than those for H. For example, the solid formed by salts of the $[Pt(CN)_4]^{2-}$ anion contains square planar complexes stacked in the crystallographic c-direction. In the simple $K_2[Pt(CN)_4]$ salt (**Figure 6a**), neighboring Pt centers are 3.48 Å apart. Within each complex, the Pt atom is in the $2+$ oxidation state and so has a d^8 electron configuration. The interaction with the CN ligands puts the energy of the molecular antibonding orbital involving the $d_{x^2-y^2}$ atomic orbital above the other d-orbitals, and so this is unoccupied. The dispersion curves of the bands formed between the metal centers depend on the symmetry of the orbitals with respect to the chain direction. **Figure 7** shows each orbital type and their arrangement along the line of complexes for the $k = 0$ wave vector. The d_{z^2} orbital forms a bonding arrangement similar to the s-orbitals used in the hydrogen example, and so the dispersion curve for this band will be similar to **Figure 4**. The d_{xz} and d_{xy} orbitals form a degenerate set of bands in which the $k = 0$ orbital is bonding. However, the bond has π-symmetry rather than σ (**Figure 7b**). This leads to a weaker overlap between neighbors and so a smaller value of the tight-binding β-parameter leading to a narrower range of energies in the band than for d_{z^2}. d_{xy} and $d_{x^2-y^2}$ orbitals, both have δ-symmetry bonding interactions at $k = 0$, (**Figure 7c**) so the bands are narrower still. The $6p$ orbitals of Pt are empty within the complex; however, the molecular orbital involving the Pt p_z orbital is only just higher in energy than $d_{x^2-y^2}$. The symmetry of the p_z orbital gives σ antibonding at $k = 0$ (**Figure 7d**) while at $k = \pm\pi/a$ each alternate p_z orbital is inverted, giving a bonding arrangement.

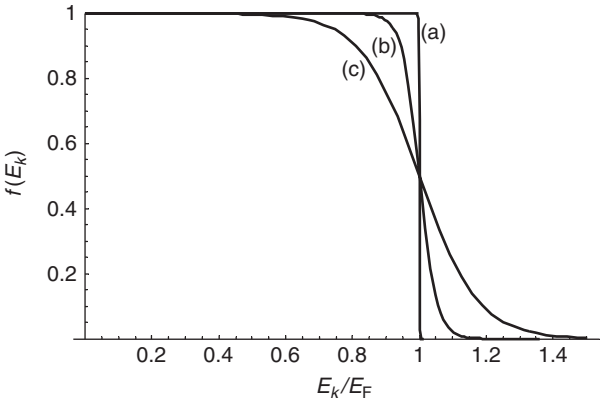


Figure 5 Plots of the Fermi–Dirac statistics function for temperatures of (a) 0 K, (b) 300 K, and (c) 1000 K.

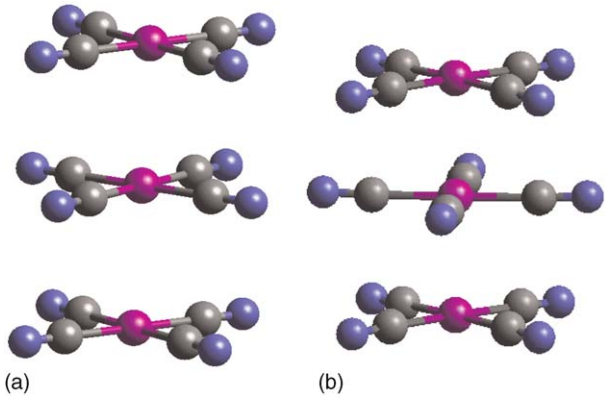


Figure 6 Chain structures formed by $[Pt(CN)_4]^{2-}$ in the solid state. (a) From $K_2[Pt(CN)_4]$, (b) from $K_2[Pt(CN)_4]Br_{0.3}$. For clarity cations, dopants, and water of crystallization have been omitted.

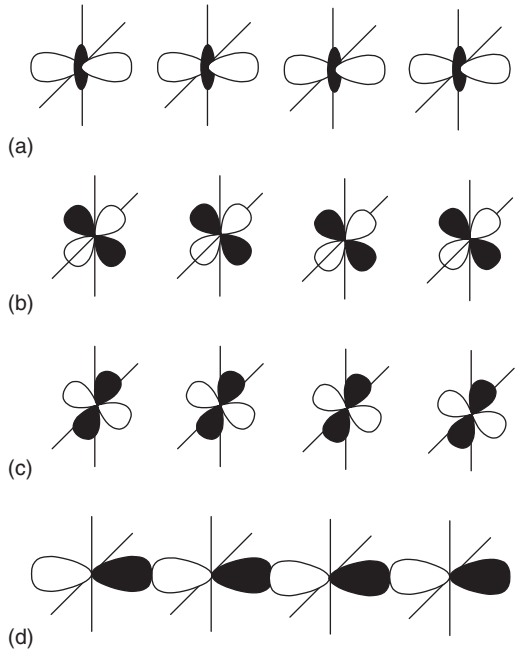


Figure 7 The arrangement of orbitals for the $k = 0$ wave vector case for a chain of atoms along the *z*-axis (which runs left to right in each diagram). (a) d_{z^2}, (b) d_{xy} or d_{yz}, (c) d_{xy}, $d_{x^2-y^2}$ have the same bonding symmetry but with orbitals rotated 45° in the *xy* plane, (d) p_z.

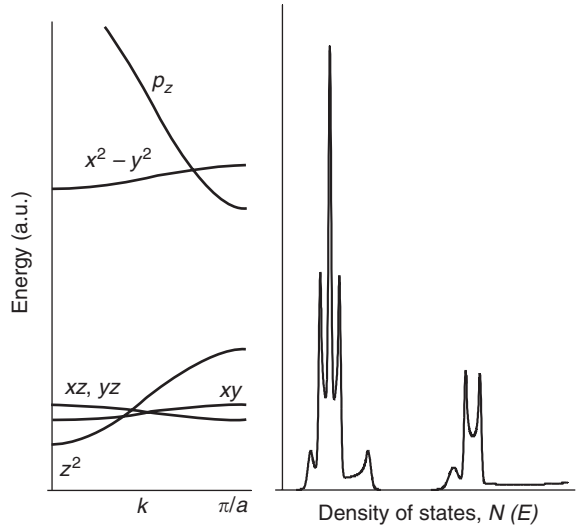


Figure 8 The dispersion curves and DOS for the bands formed along the stacking direction in $K_2[Pt(CN)_4]$ by the Pt centered orbitals.

For orbitals with this symmetry, states at the Brillouin zone boundary are lower in energy than at the center, so the dispersion relation appears inverted compared to those constructed so far. **Figure 8** shows a sketch of the dispersion curves for these bands in the Pt complex stacks. The four occupied complex orbitals involving d_{z^2}, d_{xy}, d_{xz}, and d_{yz} have similar energies and so form a continuous band of states; however, the strength of interaction between neighbors is not the same in all of them and so they have different energy ranges as discussed above. The $d_{x^2-y^2}$ and p_z orbitals form another overlapping set of bands. The p_z orbital states on the complexes are above the $d_{x^2-y^2}$; however, the interaction between neighbors is stronger in the former case and so the p_z-band is wider, leading to the lowest unoccupied states being p_z in nature.

So for $K_2[Pt(CN)_4]$, there is a bandgap between the occupied and unoccupied states, so the material is actually a semiconductor in the chain direction. However, metallic behavior can be induced by the addition of halide dopant ions to give chemical formulas such as $K_2[Pt(CN)_4]\ Br_{0.3}$. This is a level of "dopant" much greater than that used in classical semiconductors such as silicon and so the concentration of extrinsic carriers is correspondingly higher. The halide anions withdraw electrons from the Pt chains giving a free Fermi surface in the *c*-direction. Measurements confirm that the electrical conductivity is 10^4–10^5 times greater along the stack direction than perpendicular to it. The structure of the material is also strongly affected by this change to a metallic form. The Pt–Pt distance in the doped material is 2.89 Å, not much greater than in Pt metal (2.78 Å). This reduction in Pt–Pt distance can be understood in terms of the bonding character for the chain states. In the $K_2[Pt(CN)_4]$ material there is a completely filled band of states, and so both bonding and antibonding character crystal orbitals are occupied. When the dopant is added, charge is withdrawn from the high-lying antibonding states and the free Fermi surface corresponds to an overall bonding character from the crystal orbitals. The increased bonding along the chain direction leads to the observed shortening of the Pt–Pt distance. To accommodate this, there is also a rotation of alternate complexes to avoid steric clashes between CN ligands, as can be seen in **Figure 6b**.

The Density of States in Three Dimensions

In three dimensions, the wave vector has three components:

$$\boldsymbol{k} = k_x\hat{\boldsymbol{x}} + k_y\hat{\boldsymbol{y}} + k_z\hat{\boldsymbol{z}}$$

where $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ are unit vectors in their respective Cartesian directions. Each component can take any of the set of values:

$$k_i = \frac{2\pi p_i}{L}, \quad p = -L/2a, \ldots, -1, 0, 1, \ldots, L/2a$$

where L is a dimension of the sample and a is the unit cell dimension in the ith direction.

The relationship between the state energy and the wave vector now involves a vector dot product between reciprocal and real space vectors. The allowed values of k can be thought of as points in a 3D reciprocal space, the range of allowed values is defined by a reciprocal space lattice with the primitive vectors:

$$\boldsymbol{a}^* = \frac{2\pi}{V_c}\,\boldsymbol{b}\times\boldsymbol{c}, \qquad \boldsymbol{b}^* = \frac{2\pi}{V_c}\,\boldsymbol{c}\times\boldsymbol{a}, \qquad \boldsymbol{c}^* = \frac{2\pi}{V_c}\,\boldsymbol{a}\times\boldsymbol{b}$$

Now, the periodicity of the lattice will be different according to the direction considered. This means that rather than the Brillouin zone defining a single range for $|\boldsymbol{k}|$, its extent will depend on the direction taken so that the Brillouin zone is actually a volume of the reciprocal space. The first Brillouin zone is the shape formed by the Wigner–Seitz cell in this lattice. This is the primitive cell constructed by taking the intersections of planes placed at the half-way point between each reciprocal lattice point. At all points on the surface of this volume, the k-vector will have a value of $\pm\pi/a$ (a now being the repeat distance in the corresponding direction in the real space lattice) and so the cell defines the states near the set of unique k-vectors for the lattice.

The shape of the Brillouin zone is linked to the structure adopted by the metal. Examples of the Brillouin zone for two common structures found in elemental metals, body-centered cubic (b.c.c.) and face-centered cubic (f.c.c.), are shown in **Figure 9**.

In this picture, when E_F is far from the Brillouin zone edge, the k-vectors of the states at the Fermi level define the surface of a sphere centered on the origin. The DOS in 3D is now affected by the number of possible k-vectors that have equivalent state energies. For large values of $|\boldsymbol{k}|$, near the zone boundaries, there are many k-vectors with the same magnitude but with different directions in reciprocal space, leading to a high degree of degeneracy. This, coupled with the flattening of the dispersion relation, leads to a high DOS at large $|\boldsymbol{k}|$. In contrast to the 1D situation, however, at the zone center, there are only a few wave vectors of the same length, and so although the dispersion relation may flatten out there is no longer a significant increase in the DOS. Indeed at the zone center, $|\boldsymbol{k}| = 0$, and so all three components are zero and there is only one state.

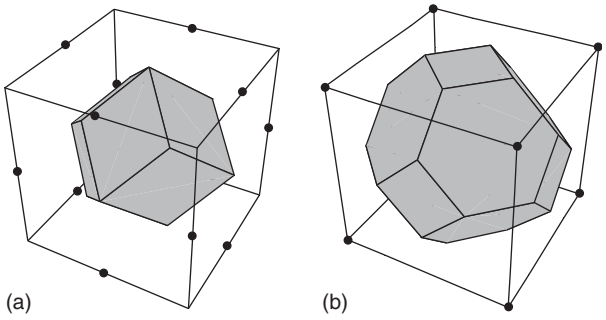


Figure 9 The first Brillouin zone for the real space lattices of: (a) b.c.c. structures and (b) f.c.c. structures. Note that the reciprocal space lattice of the b.c.c. lattice is actually f.c.c. and vice versa.

Dependence of Metallic Structure on the Density of States

The shape and volume of the Brillouin zone control the electron density at which the Fermi surface touches the Brillouin zone boundary. Near the boundary, the dispersion curve flattens off leading to an increased DOS. Differences in the reciprocal lattices between structures will affect the point at which this occurs and so different lattices will have different total electronic energies at a given density.

For example, **Figure 9** shows the shape of the Brillouin zones for the f.c.c. and b.c.c. lattices. The points in the diagrams represent the points in the reciprocal space lattice at the spacing set by the reciprocal space vectors:

$$\boldsymbol{a}^* = \frac{2\pi}{a}(-\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}+\hat{\boldsymbol{z}}), \qquad \boldsymbol{b}^* = \frac{2\pi}{a}(\hat{\boldsymbol{x}}-\hat{\boldsymbol{y}}+\hat{\boldsymbol{z}}),$$
$$\boldsymbol{c}^* = \frac{2\pi}{a}(\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}-\hat{\boldsymbol{z}}) \qquad \text{for f.c.c.}$$

and

$$\boldsymbol{a}^* = \frac{2\pi}{a}(\hat{\boldsymbol{y}}+\hat{\boldsymbol{z}}), \qquad \boldsymbol{b}^* = \frac{2\pi}{a}(\hat{\boldsymbol{x}}+\hat{\boldsymbol{z}}),$$
$$\boldsymbol{c}^* = \frac{2\pi}{a}(\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}) \qquad \text{for b.c.c.}$$

where a is the length of the cube side in the real space unit cells. From these formulas, it can be seen that the reciprocal space primitive vectors for the f.c.c. lattice are the same as the real space b.c.c. primitive vectors and vice versa.

Figure 10 shows a sketch of the DOS for the f.c.c. and b.c.c. lattices. At low electron density, all occupied states are well within the Brillouin zone boundary and so the DOS simply increases with the state energy since the number of degenerate k-vectors increases. As the wave vectors approach the Brillouin zone boundary, the dispersion curve levels out and the degeneracy of states increases giving a greater DOS. However, if the energy is increased further, the number of available states decreases again since only the sections of volume furthest from the center of the Wigner–Seitz cell are still available. This leads to a peak in the DOS in the vicinity of the Brillouin zone boundary. Because of the geometry of the reciprocal space lattices, the peak occurs at lower energy for

the f.c.c. real space lattice than for the b.c.c. case. For a metal with electron density at the f.c.c. peak, the f.c.c. structure will be preferred because the Fermi energy will be lower than in the b.c.c. lattice. However, if the density of electrons is near the b.c.c. peak, the f.c.c. DOS begins to fall dramatically and so a lower-energy electron configuration is possible in the b.c.c. structure.

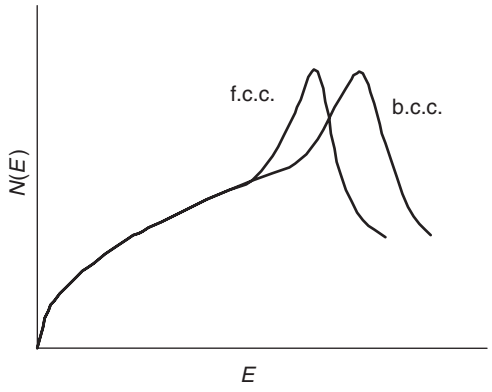


Figure 10 Schematic diagram of the DOS for the f.c.c. and b.c.c. crystal systems. The peaks correspond to the point when the *k*-vector just touches the Brillouin zone boundary.

The DOS of real metals is more complex than this illustration using a single band in f.c.c. and b.c.c. structures. However, the example does serve to show that the DOS will influence the relative total electronic energy of alternative metal lattices. Or, conversely, the electron density of a metal can have a dramatic influence on its optimal structure.

This idea has been tested using higher-level calculations on the first row transition metals to obtain the band structures and DOSs for the valence shells in b.c.c., f.c.c., and h.c.p. structural alternatives, as shown in **Figure 11**. The shape of the DOS for each structure is largely controlled by the interaction of the metal *d*-orbitals in the different crystallographic directions since the overlap of the spherical *s*-orbitals is less influenced by the lattice geometry. At low energies, the calculated DOS are all very similar. At 0.35 eV, corresponding to three electrons per atom, the b.c.c. plot shows a distinct dip and so at this point an increase in electron density with this structure must occupy higher-energy states than are required in the close packed lattices. Integration of the occupied state energies shows that in this regime the close packed lattices are indeed lower in energy than in b.c.c., and under standard conditions the early

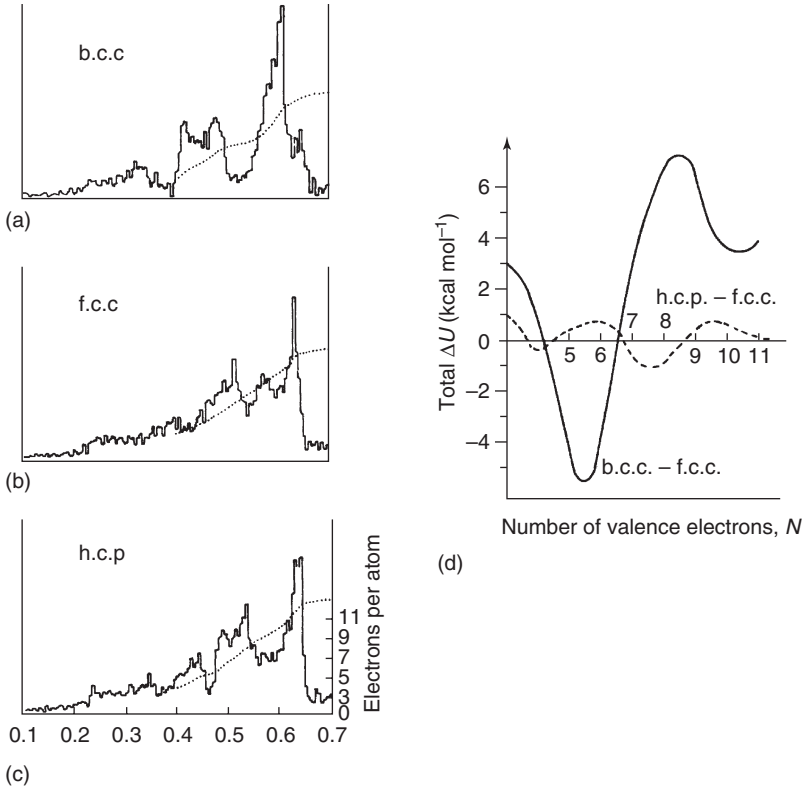


Figure 11 The DOS for (a) b.c.c., (b) f.c.c., and (c) h.c.p. structures. The dotted line in each diagram is the integrated DOS and shows the level of the Fermi energy in the metal with the number of electrons per atom indicated on diagram (c). (d) The total electronic energy of the b.c.c. and h.c.p. structures relative to that of the f.c.c. as a function of electron density. (Reproduced from Pettifor DG (1977) A physicist's view of the energetics of transition metals. *Calphad* 1(4): 305–324, with permission from Elsevier.)

transition metals, Sc ($3d^1 4s^2$) and Ti ($3d^2 4s^2$), have the h.c.p. structure suggested in **Figure 11d**. Just above 4 eV, there is a peak in the b.c.c. DOS and so the total energy at the corresponding electron densities favors the b.c.c. structure. Vanadium and chromium with five and six valence electrons, respectively, both adopt this lattice. Manganese has a complex structure which has not been included in this study, whereas Fe ($3d^6 4s^2$) has a b.c.c. structure at room temperature against the predicted f.c.c. lattice from **Figure 11d**. However, magnetic effects seem to dominate this choice of structure for Fe since above the Curie temperature, the f.c.c. structure is observed. For higher electron densities, the close packed structures are preferred, with the choice of f.c.c. or h.c.p. finely balanced. This is not surprising since the first neighbor interactions are identical in the two structures. Even so the total electronic energies indicate that at seven and eight valence electrons, h.c.p. is more stable than f.c.c. while the situation is reversed for nine and ten outer electrons. This trend is, indeed, observed with the structures of Co, Ni being h.c.p. and Cu, Zn adopting the f.c.c. structure.

See also: Ionic Bonding and Crystals; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); van der Waals Bonding and Inert Gases.

PACS: 61.50.Lt; 71.20. – b

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. USA: Saunders.
Cox PA (1987) *The Electronic Structure and Chemistry of Solids*. Oxford: Oxford University Press.
Elliott S (1998) *The Physics and Chemistry of Solids*. UK: Wiley.
Hoffmann R (1988) *Solids and Surfaces, A Chemists View of Bonding in Extended Structures*. Weinham: VCH.
Kittel C (1995) *Introduction to Solid State Physics*. UK: Wiley.
West AR (1999) *Basic Solid State Chemistry*. UK: Wiley.

Metalloproteins, Electron Transfer in

A R Bizzarri and S Cannistraro, University of Tuscia, Viterbo, Italy

Introduction

In the energy-capture machinery of biology, especially in photosynthesis and in mitochondria, there are chains of redox enzymes containing metal ions, including iron, heme-iron, copper, manganese arranged in order so as to assist electron flow and generate charge separation which represent the first step in the energy-capture process. The ability of transition metals to exist in more than one stable oxidation state makes them suitable catalysts for biological processes that require transfer of electrons. The oxidation/reduction centers are placed in order of their potentials, expressed in terms of free-energy differences, so that the directionality of the electron flow is thermodynamically determined. Along the bioenergetic pathways, electrons flow down a gradient of potential energy that spans a range of less than 1.2 eV.

However, biological material is not constituted by regular arrays of lattice pointing to allow electrons to travel over long distances as delocalized Bloch waves; therefore, the familiar concepts of metallic conduction, through partly filled conduction bands, cannot apply. Instead, biology employs redox centers which are localized potential wells, among which electron transfer (ET) occurs through a hopping mechanism from center to center. The redox centers are nearly always found buried beneath the protein surface, the protein coating serving as an insulator which protects the redox center from short-circuiting ETs. Electrons can be transported across protein-domain interfaces or along a redox chain which consists of a number of proteins (see, for instance, **Figure 1**), the related processes involving several reaction steps.

In ET intra-protein, the redox centers are held at a fixed distance and in a fixed orientation with respect to each other in order to presumably adapt optimally to physiological needs. When two redox centers that are consecutive in the ET train are located on different proteins, ET requires proteins first to form a docking or associative complex in which the partners assemble transiently through complementary contact surfaces. The mutual approach, initially, may be governed by long-range electrostatic forces reflecting the overall charges on the two partners. At shorter ranges, either hydrophobic or Coulomb interactions between opposite charge patches, or combinations of both, determine the structure of the docking complex and the relative orientation of the partners. Generally, for successful ET within a docking complex, the partners should have motional degrees of freedom that allow them to perform a rolling or sliding motion with respect to each other.

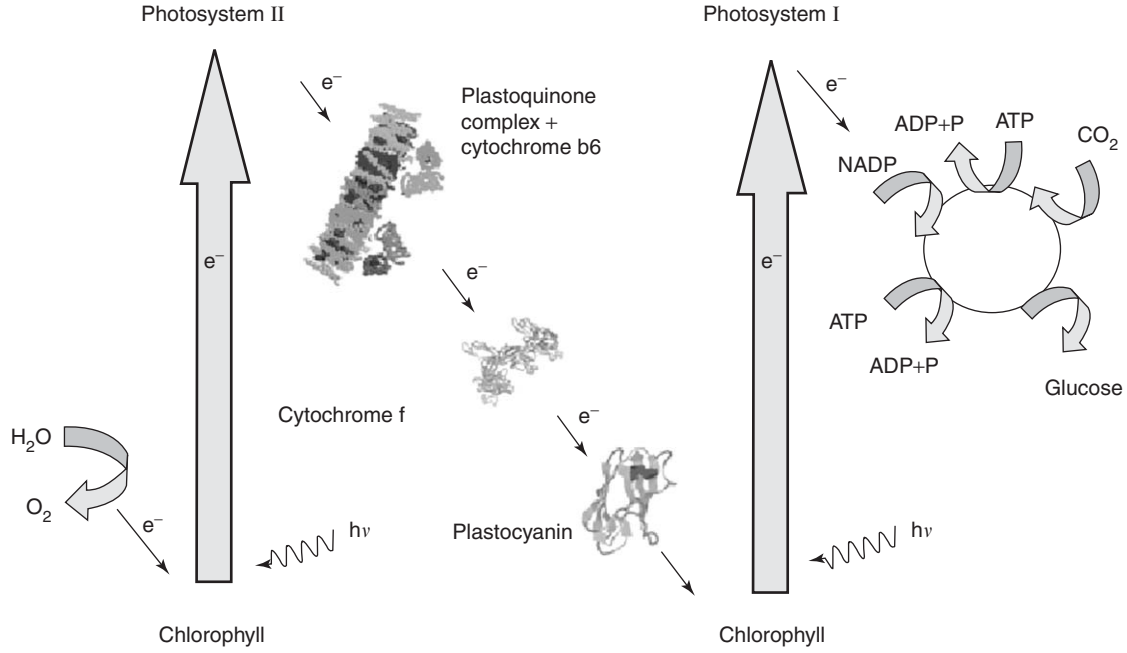


Figure 1 Photosynthesis pathways as found in green plants. Electrons are extracted from water by photosystem II and driven through a complex coupled ET pathway to the CO_2 reduction cycle by photosystem I.

The peculiarity of ET processes mediated by metalloproteins is that they occur over long distances, in a very fast, directional and efficient way, the donor and the acceptor proteins having redox centers separated by distances between 5–25 Å. In contrast to usual chemical reactions which involve the making and breaking of bonds and proceed via a well-defined reaction coordinate, the reactants and products of a protein ET are often chemically indistinguishable.

ET reactions in metalloproteins are usually termed as outer-sphere ETs, since they take place with no or very weak coupling between donor and acceptor sites of the electron.

Typical ET rate constants for biological processes lie in the order of 10^2–$10^3\,s^{-1}$. Optimization of biological ET for speed seems to be necessary to compete with back reactions, especially to prevent charge recombination to occur.

Several factors are suggested to influence the ET efficiency, such as the distance between the redox centers, the role of the protein medium, the possible existence of conducting pathways, the role of the solvent medium, and the assistance of the protein and solvent dynamics. Indeed, the dynamics of the intervening medium through which an electron passes can finely tune the ET process. In particular, a prominent role is played by low-frequency, collective vibrational modes.

Different kinds of ET metalloproteins can be found in biological processes. Among others are cytochromes in which the active site is the heme consisting of iron coordinated to the porphyrin group, and the redox potentials for the Fe^{+3}/Fe^{+2} couple ranging from 5 to 260 mV. Iron–sulfur proteins are characterized by iron–sulfur clusters centered on the iron with redox potentials covering a range, from -700 to 500 mV. Blue copper proteins, in which the copper ion directly coordinates to amino acid residues in a distorted tetrahedral arrangement, are characterized by peculiar spectroscopic properties and a redox potential for the Cu^{+2}/Cu^{+1} couple from 130 to 680 mV.

More recently, understanding the mechanism of electron transduction through biological macromolecules has assumed fundamental importance not only in increasing the knowledge of the ET process, ubiquitous in biology, but also in the development of novel, improved bioelectronic devices.

The ET Rate: A General Overview

The ET theory describes the transition of an electron from a donor D to an acceptor A, the reactant and product, DA and D^+A^-, representing the system before and after the ET process, respectively. The states $|DA\rangle$ and $|D^+A^-\rangle$ can be expressed by the combination of the wave functions for the two redox centers:

$$|DA\rangle = (\overline{\Phi}_D\Phi_A), \qquad |D^+A^-\rangle = (\Phi_D\overline{\Phi}_A)$$

where Φ_D and Φ_A are the complete wave functions describing the nuclear and electronic motions of the

two redox centers, the bar denoting the electron participating in the transition.

The traditional transition-state theory of the rates for chemical reactions involves motion along a potential energy surface in which the reactant atoms gain energy from thermal collisions, surmount an activation energy barrier to achieve a transition state, and spontaneously decay into the product. In contrast to these, usually called adiabatic reactions, where formation of the transition state leads almost inevitably to the product, the probability of long-distance ET for such a transition is small. Accordingly, a nonadiabatic description for the ET reaction is more appropriate.

For nonadiabatic ET reactions, the first-order rate constant k_{ET}

$$|DA\rangle \overset{k_{ET}}{\rightarrow} |D^+A^-\rangle$$

can be conveniently expressed by Fermi's golden rule obtained from the time-dependent quantum mechanical perturbation theory:

$$k_{ET} = \frac{2\pi}{\hbar} V_R^2 \rho \qquad [1]$$

where $\hbar$ is the Planck's constant divided by 2π, V_R^2 is the square of the quantum matrix element for electronic coupling between donor and acceptor, averaged over all possible thermal fluctuations of the system, ρ is the density of states, that is, the number of states per unity interval of energy.

The electronic coupling V_R^2 is the principal origin of the distance dependence of the ET constant rate. The simplest model, neglecting the role of the intervening medium, predicts that V_R^2 falls off exponentially with the distance R between the donor D and the acceptor A according to $e^{-\beta R}$, where β is an attenuation factor. In order to obtain ET rates in the order of 10^2–$10^3\,s^{-1}$, R must be in the range 10–20 Å (see also in the following).

Under the requirement of the Franck–Condon principle, stating that during the almost instantaneous ET process the nuclei do not change either their positions or their momenta, the ET rate can be expressed in the form

$$k_{ET} = \frac{2\pi}{\hbar} V_R^2 F_C \qquad [2]$$

where F_C is the Franck–Condon weighted density of states reflecting the overlap of the donor and acceptor nuclear and solvational wave functions.

Accordingly, ET will occur at nuclear configurations for which the total potential energy of the reactants and surrounding medium is equal to that of the products and the surrounding medium. The quantity F_C is a sum of the square of the overlap integrals $S_{v_{DA}v_{D^+A^-}}$ of the vibrational wave functions of the reactants (v_{DA}) with the corresponding ones of the products ($v_{D^+A^-}$), weighted by Boltzmann factors:

$$F_C = \sum_{v_{DA}} \sum_{v_{D^+A^-}} S^2_{v_{DA}v_{D^+A^-}} p(v_{DA}) \qquad [3]$$

with

$$S_{v_{DA}v_{D^+A^-}} = \int \chi_{v_{DA}} \chi_{v_{D^+A^-}}\, dx$$

where $\chi_{v_{DA}}$ and $\chi_{v_{D^+A^-}}$ are the wave functions for the states v_{DA} and $v_{D^+A^-}$, respectively, x being the oscillation coordinate; $p(v_{DA})$ is the equilibrium Boltzmann probability of finding the system in the vibrational state v_{DA}. The sum is over any given set of the vibrational quantum numbers $(v_{DA}, v_{D^+A^-})$ of the reactant DA and the product D^+A^-, including the solvent, such a sum being, however, limited by the fact that only a small number of states have a finite overlap and hence contributes to the final term.

The calculation of F_C can be extremely problematic and different expressions for the Frank–Condon factor can be obtained depending on the approximations done. Classical, semiclassical, and quantum mechanical approaches, according to the treatment of the nuclear motions, have been followed to work out useful expressions for F_C.

When all the vibrational frequencies are relatively small, for example, $\hbar\omega \ll k_B T$, the vibrational manifold of DA and D^+A^- can be treated as a continuum and the passage across the activation barrier can be described classically in terms of the activated complex theory, properly modified for nonadiabatic reactions. Accordingly, the ET rate is related to the free activation energy, ΔG^*, of the reaction as follows:

$$k_{ET} = k_{ET}(0) e^{-\Delta G^*/k_B T} \qquad [4]$$

where the pre-exponential factor $k_{ET}(0)$ defines the limiting rate of the reaction.

Generally, classical results are valid only at temperatures high enough so that the vibrations are fully excited. The classical approach is discussed in the next section where the expression worked out by Marcus for F_C, and hence for the ET reaction rate, is presented. When $\hbar\omega \geq k_B T$, the discrete nature of the vibrational manifold must be taken into account. The ET rate may become temperature independent and a quantum mechanical view is more appropriate. Moreover, under these conditions, electron tunneling may become important. This aspect is briefly presented in the following.

Marcus Theory of ET

The simplest theoretical treatment of the rate of ET in metalloproteins is due to Marcus using a classical

harmonic oscillator model, which generates parabolic potential energy curves. Indeed, this approach represents the most insightful and used theoretical framework to interpret experimental ET results in metalloproteins.

The ET process can be seen as an electron jump from DA to D^+A^- and requires one to match the Franck–Condon principle, which implies that the nuclear configuration and, in addition, the conservation of energy are the same immediately after the ET as before. Under these requirements, the electron jump takes place in the vicinity of the crossover, nuclear configuration C (**Figure 2**).

Thermal fluctuations and/or vibrations in some coordinates will be required for reaction to occur. Since the charge distribution of the protein matrix is different before and after the ET process, the surrounding medium will be polarized differently in the two states. Therefore, the coordinates involved in the ET reaction include vibrational coordinates of the protein, and the vibrational and orientational coordinates of the surrounding solvent. All the fluctuating nuclear coordinates relevant to the ET reaction are usually lumped together into the so-called reaction coordinate Q. The dependence of the potential energy curves of the reactant DA and the product D^+A^- on Q is assumed to be quadratic, according to the harmonic approximation, and unchanged by the ET reaction (**Figure 2a**).

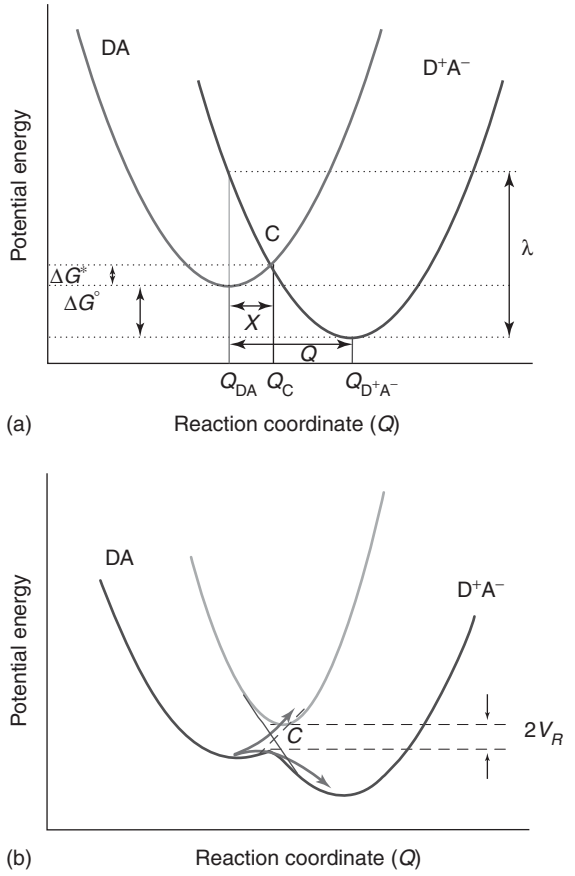


Figure 2 (a) The harmonic potential energy of the reactant (DA) and product (D^+A^-) as a function of the reaction coordinate Q. ΔG° is the standard free energy, λ is the reorganization energy, and ΔG^* is the free energy of activation. (b) Influence of the electronic coupling on the potential energy curves. In both cases, the ordinate represents the potential energy of the nuclei of the whole system: donor + acceptor + medium.

In the nonadiabatic case, the passage through C does not usually cause transition from DA to D^+A^-. Once the system reaches the intersection of the potential curves C, the probability of going from DA to D^+A^- depends on a number of factors such as the extent of coupling of the electronic orbitals of the two reactants, which in turn depends on the separation distance of the two reactants, the separation between the two potential energy curves over a vertical distance being given by $2V_R$, where V_R is the electronic coupling element between DA and D^+A^- (**Figure 2b**). If the gap is large enough, as it occurs in adiabatic reactions, the transition DA to D^+A^- will take place each time the crossing point is reached. Conversely, for small V_R as in nonadiabatic ET reactions, the electron, for most of the times it reaches the crossing point, continues its motion along the curve DA (upward-running arrow in **Figure 2b**). Only once in a while, the electron will make the transition to D^+A^- when the crossing is reached (downward-running curve in **Figure 2b**).

In the framework of the Marcus theory, the electron jumping from the equilibrium coordinate Q_{DA} of DA to the equilibrium coordinate $Q_{D^+A^-}$ of D^+A^- can be described in terms of the free energy of activation, ΔG^* which is related to two experimental observables, ΔG°, the standard free energy and λ, the reorganization energy. ΔG^* is the energy required to reach the point C from DA, overcoming the activation barrier. The standard free energy or driving force, ΔG°, is the energy difference between the ground states of DA and D^+A^-, respectively, and can be expressed as

$$\Delta G^\circ = zF(E^\circ_A - E^\circ_D) \qquad [5]$$

where z is the number of electrons transferred, F is the Faraday constant, and E°_A and E°_D are the midpoint potentials of the acceptor and donor centers.

The reorganization energy λ is the free energy required to move all the atoms from their equilibrium positions before the ET to the equilibrium positions they would have after the ET without transferring the electron. By referring to **Figure 2a**, the activation barrier ΔG^* and the reorganization energy λ are

$$\Delta G^* = \tfrac{1}{2} k_H X^2, \qquad \lambda = \tfrac{1}{2} k_H Q^2 \qquad [6]$$

where X is the distance between the equilibrium coordinate Q_{DA} and the coordinate of the crossing point C, and Q is the distance between the equilibrium coordinates Q_{DA} and $Q_{D^+A^-}$, k_H being the force constant of the harmonic potential energy corresponding to both DA and D^+A^-.

The difference between the free energy and the activation energy can be expressed as

$$\begin{aligned}\Delta G^* - \Delta G^\circ &= \tfrac{1}{2} k_H (Q - X)^2 \\ &= \tfrac{1}{2} k_H Q^2 + \tfrac{1}{2} k_H X^2 - k_H Q X \end{aligned} \quad [7]$$

By rearranging eqns [6] and [7], one gets

$$X = \frac{\lambda + \Delta G^\circ}{k_H Q} \quad [8]$$

and finally

$$\Delta G^* = \frac{1}{2} k_H \frac{(\lambda + \Delta G^\circ)^2}{k_H^2 Q^2} = \frac{(\lambda + \Delta G^\circ)^2}{4\lambda} \quad [9]$$

Therefore, from eqn [4], the ET rate can be expressed by

$$k_{ET} = k_{ET}(0) \exp\left[\frac{-(\Delta G^\circ + \lambda)^2}{4\pi\lambda k_B T}\right] \quad [10]$$

where the pre-exponential factor $k_{ET}(0)$ assumes the form

$$k_{ET}(0) = \frac{2\pi}{\hbar} V_R^2 \left(\frac{1}{4\pi\lambda k_B T}\right)^{1/2} \quad [11]$$

leading to the final expression:

$$\begin{aligned} k_{ET} = &\frac{2\pi}{\hbar} V_R^2 \left(\frac{1}{4\pi\lambda k_B T}\right)^{-1/2} \\ &\times \exp\left[\frac{-(\Delta G^\circ + \lambda)^2}{4\pi\lambda k_B T}\right] \end{aligned} \quad [12]$$

The ET rate can be modulated by thermodynamic (ΔG°) and intrinsic (λ) factors and, in addition, it varies with the temperature. Notably, the optimal rate is obtained when $-\Delta G^\circ$ matches λ, such a condition having been exploited to extract information on V_R and λ from k_{ET} values (see below).

As already mentioned, the final Marcus expression provides a sound approach to describe the ET processes in metalloproteins. However, while it takes into account the temperature dependence of the ET rate at high temperature, it fails at low temperatures, at which a deviation from an Arrhenius-like behavior, or even temperature-independence (at very low T), has been observed.

Quantum Mechanical Theories of ET

A semiclassical approach has been followed by Hopfield to derive an expression for the ET rate. He treated oscillators classically, but assumed quantized energy levels. By introducing the probability distributions $D_D(E)$ and $D_A(E')$, corresponding to the energy required to remove an electron from the donor and to give an electron to the acceptor, respectively, the density of states in eqn [1] can be expressed by

$$\rho = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} D_D(E) D_A(E') \, dE \, dE' \quad [13]$$

He assumed, moreover, that both the distributions follow a Gaussian form

$$\begin{aligned} D_D(E) &= \frac{1}{(2\pi\sigma_D^2)^{1/2}} \exp\left[\frac{-(E + E_D - \lambda_D)^2}{2\sigma_D^2}\right] \\ D_A(E') &= \frac{1}{(2\pi\sigma_A^2)^{1/2}} \exp\left[\frac{-(E' + E_A + \lambda_A)^2}{2\sigma_A^2}\right] \end{aligned} \quad [14]$$

with standard deviations σ_D and σ_A, respectively; E_D is the energy of the ground state of the reduced form of the donor and E_A is the reduced form of the acceptor; $D_D(E)$ is centered below E_D by an amount λ_D because ordinarily, the nuclear configuration in the initial reduced state will not be an equilibrium configuration for the final, oxidized state; for the same reason, the center of $D_A(E')$ is displaced upwards from E_A by an amount λ_A.

By taking into account the Franck–Condon principle and the conservation of energy in the process ($E = E'$), eqn [13] becomes

$$\begin{aligned} \rho = F_C &= \int_{-\infty}^{+\infty} D_D(E) D_A(E) \, dE \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[\frac{-(\Delta E - \lambda)^2}{2\sigma^2}\right] \end{aligned} \quad [15]$$

where $\sigma^2 = \sigma_D^2 + \sigma_A^2$, $\lambda = \lambda_D + \lambda_A$, and $\Delta E = E_A - E_D$ being the energy gap of the reaction.

By assuming that both the donor and the acceptor have a quadratic dependence on the nuclear coordinate with the same force constant k_H, and a spacing between energy levels equal to $\hbar\omega_D$ and $\hbar\omega_A$ for the donor and the acceptor respectively, it comes out that

$$\begin{aligned} \sigma_D^2 &= \hbar\omega_D \lambda_D \coth(\hbar\omega_D / 2k_H T) \\ \sigma_A^2 &= \hbar\omega_A \lambda_A \coth(\hbar\omega_A / 2k_H T) \end{aligned}$$

Therefore, the resulting ET rate becomes

$$k_{ET} = \frac{2\pi}{\hbar} V_R^2 \left(\frac{1}{2\pi\sigma^2}\right)^{-1/2} \times \exp\frac{-(\Delta E - \lambda)^2}{2\sigma^2} \quad [16]$$

Notably, in the limit of high temperatures, σ approaches $k_B T$ and eqn [16] results in being formally similar to the Marcus expression, once $-\Delta E$ is identified by ΔG^*. Conversely, at very low temperatures, the Frank–Condon factor F_C becomes temperature independent, in agreement with some experimental data.

However, the limiting expression at low temperature is not in complete agreement with experimental data. Significant improvements can be reached by a full quantum mechanical approach also taking into account the vibrations coupled to changes of the electronic state. A variety of quantum mechanical treatments with different degrees of approximations have been developed. Under the assumption that the system consists of a single group of harmonic oscillators and that there is a single prevailing mode, from eqn [3], the Levich–Degonadze–Jortner expression can be derived for F_C:

$$F_C = \frac{1}{\hbar\omega} e^{-S(2n+1)} \left(\frac{n+1}{n}\right) \times \frac{\Delta E}{2\hbar\omega} I_P\left[2S\sqrt{n(n+1)}\right] \quad [17]$$

where $\hbar\omega$ is the characteristic frequency, ΔE is the energy gap of the reaction, S is equal to $\lambda/\hbar\omega$, I_P is the modified Bessel function of the order P, and n is given by

$$n = [e^{\hbar\omega/kT} - 1]^{-1}$$

The ET rate can be then expressed by

$$k_{ET} = \frac{2\pi}{\hbar} V_R^2 \frac{1}{\hbar\omega} e^{-S(2n+1)} \left(\frac{n+1}{n}\right) \times \frac{\Delta E}{2\hbar\omega} I_P\left[2S\sqrt{n(n+1)}\right] \quad [18]$$

In the limit of very high temperature, eqn [18] again reduces to the Marcus expression. In addition, it well reproduces the experimental trend with temperature. However, it does not include the coupling of low-frequency vibrational modes of the redox center to the ET process. In order to take into account for low-frequency modes, a correction to eqn [18] has been done by including the oscillator zero-point energy considering the change of frequency, leading to an expression similar to eqn [12], containing $-\Delta G^\circ$ instead of ΔE.

Dependence of ET Rate on the Protein Matrix

A crucial aspect of the ET process is represented by the dependence of k_{ET} on the medium between the redox centers. To obtain information on the variation of k_{ET} with the nature of the medium, first it is necessary to eliminate, or at least to minimize, the dependence of k_{ET} on both the driving force and the reorganization energy. In the framework of the Marcus theory, this can be achieved by extrapolating k_{ET} when $\Delta G^\circ = -\lambda$. According to eqn [12], the exponential trend on the driving force and reorganization energy disappears. The resulting rate remains only weakly dependent on λ ($k_{ET} \sim 1/(\lambda)^{1/2}$), and its change with the medium is mainly reflected by V_R.

The dependence of V_R on the intervening medium can be cast in the form

$$V_R^2 = V_R^{o2} f_M^2 \quad [19]$$

where V_R^o represents the electronic coupling between DA and D^+A^- when the redox centers are in van der Waals contact, and f_M is a dimensionless attenuation factor which varies between 1 (van der Waals contact) and 0 (infinite distance). The dependence of f_M on the detailed structure of the medium connecting the two redox centers has been widely investigated. Only two cases are considered in the following.

In the first approach, the protein intervening medium is pictured as an organic glass, the random, disordered connections between the two redox centers constituting the overall path of ET. In this framework, the distance between the donor and the acceptor centers is the parameter governing the ET rate. The dependence of f_M^2 on the distance, R, between the center of the edge atom of the donor and that of acceptor can be expressed by

$$f_M^2 = e^{-\beta(R-R_o)} \quad [20]$$

where the exponential coefficient of decay, β, quantitatively describes the nature of the intervening medium with respect to its efficiency to mediate the ET process, for instance, through the propagation of the relevant wave functions. A variation of ET rates over 12 orders of magnitude can satisfactorily be accounted for by a distance dependence as in eqns [19] and [20], with $\beta = 1.4\,\text{Å}^{-1}$ and $R_o = 3.6\,\text{Å}$. More generally, values of β in the range 0.7–1.4 Å^{-1} are found to reproduce the experimental k_{ET} values of metalloproteins.

An alternative description of the dependence of f_M^2 with the nature of the medium invokes the so-called pathway model which is, in some sense, based on the mechanism of super-exchange, owing to the fact that the electronic coupling between the donor and

acceptor wave functions is mediated by a third center that connects the two wave functions. A pathway is defined as a combination of interacting bonds that link the donor and the acceptor. Three types of steps are distinguished, depending on whether the transfer occurs between atoms that are connected through a covalent bond (C), a hydrogen bond (H), or not connected at all (S); in the latter case, the electron must be transferred through space. The corresponding attenuation factor for a transfer path can be expressed as

$$f_{\mathrm{M}} = \prod_{i}^{N_{\mathrm{C}}} \varepsilon_i^{\mathrm{C}} \prod_{j}^{N_{\mathrm{H}}} \varepsilon_j^{\mathrm{H}} \prod_{k}^{N_{\mathrm{S}}} \varepsilon_k^{\mathrm{S}} \qquad [21]$$

where N_{C}, N_{H}, N_{S} are the number of C, H, and S paths, respectively; ε are the individual attenuation factors corresponding to a particular step in the transfer path. Semi-empirical expressions for the ε-factors are

$$\varepsilon_{\mathrm{C}} = 0.6, \qquad \varepsilon_{\mathrm{H}} = 0.36\mathrm{e}^{[-1.7(R-2.8)]},$$
$$\varepsilon_{\mathrm{S}} = 0.6\mathrm{e}^{[-1.7(R-1.4)]}$$

in which R is the distance between the two atoms exchanging the electron. Such an approach predicts that α-helices are characterized by a lower conductivity than β-sheets.

Reorganization Energy in the ET Process

When an electron is transferred in an intermolecular process, through the protein matrix, the distribution of charge is different before and after the transfer, as already mentioned. These changes have to be accommodated by the local dielectric properties contributed by polarizability, local bonds, reorientation of polar side chains, dissociation or association of protolytic groups, movements of ions in the solvent, reorientation of solvent dipoles, etc. Additionally, the local structure of the redox center might undergo changes in configuration. All these physical effects are accounted for by the reorganization energy λ.

Marcus originally divided the reorganization energy into changes occurring at the redox center (inner sphere) and those occurring in the surrounding protein/water matrix (outer sphere). Accordingly, λ can be separated into: $\lambda = \lambda_{\mathrm{i}} + \lambda_{\mathrm{o}}$ where λ_{i} is the contribution to the reorganizational energy of the inner shell of atoms, close to the redox center, while λ_{o} refers to atoms further out, generically called "solvent."

The inner-sphere reorganization energy λ_{i}, which reflects redox-dependent nuclear perturbations of the redox centers, such as changes in bond lengths and angles, can be expressed in terms of the inner shell normal vibrational modes:

$$\lambda_{\mathrm{i}} = \frac{1}{2}\sum_{j} k_{\mathrm{H}j} Q_j^2 \qquad [22]$$

where Q_j is the displacement from the equilibrium position of the jth normal coordinate caused by the ET; the constant $k_{\mathrm{H}j}$ being given by

$$k_{\mathrm{H}j} = \frac{f_j^{\mathrm{DA}} f_j^{\mathrm{D^+A^-}}}{f_j^{\mathrm{DA}} + f_j^{\mathrm{D^+A^-}}} \qquad [23]$$

where f_j^{DA} and $f_j^{\mathrm{D^+A^-}}$ are the force constants at the equilibrium for DA to $\mathrm{D^+A^-}$, respectively.

The outer-sphere reorganization energy, λ_{o}, which reflects changes in the surrounding medium, such as changes in solvent orientation, can be estimated from the polarizability of the solvent, as considered to be a continuous polar medium:

$$\lambda_{\mathrm{o}} = \frac{Ne}{4\pi\varepsilon_{\mathrm{o}}}\left[\frac{1}{2R_1} + \frac{1}{2R_2} - \frac{1}{R}\right]\left[\frac{1}{D_{\mathrm{OP}}} - \frac{1}{D_{\mathrm{S}}}\right] \qquad [24]$$

where Ne is the charge transferred from the donor to the acceptor; R_1 and R_2 are the radii of the two spherical reactants when in contact and $R = R_1 + R_2$; D_{OP} is the square of the refractive index of the medium and D_{S} is the static dielectric constant; ε_{o} is the permittivity of space.

For redox centers that are buried within a protein, λ_{o} may also include configurational changes in the protein matrix and, for interprotein reactions, in the interface between the donor and the acceptor proteins. It is to be noted that the higher the dielectric constant of the solvent, the larger is the value obtained for λ. Furthermore, for nonpolar solvent, λ_{o} vanishes and such a condition can be exploited to obtain information on λ_{i}.

In the framework of the Marcus theory (see eqn [12] and **Figure 2a**), a decrease in the driving force, for constant λ, will displace the product potential energy upward, causing ΔG^* to increase and consequently the ET rate to decrease. Similarly, at constant $-\Delta G^\circ$, an increase in λ will increase the horizontal displacement of the product from the reactant, imposing a higher activation barrier. Depending on the sign of the quantity $(\Delta G^\circ + \lambda)$, one can distinguish the so-called normal region $[(\Delta G^\circ + \lambda) > 0]$, the activationless region $[(\Delta G^\circ + \lambda) = 0]$, and the inverted region $[(\Delta G^\circ + \lambda) < 0]$. These conditions are qualitatively illustrated in **Figure 3**.

In the normal region $[(\Delta G^\circ + \lambda) > 0]$, an increase in the driving force accelerates the ET process, while it slows down in the inverted region, thus representing one of the most celebrated predictions of the Marcus expression. Notably, when $[\Delta G^\circ + \lambda = 0]$, the ET rate is least affected by variations of T and λ. For

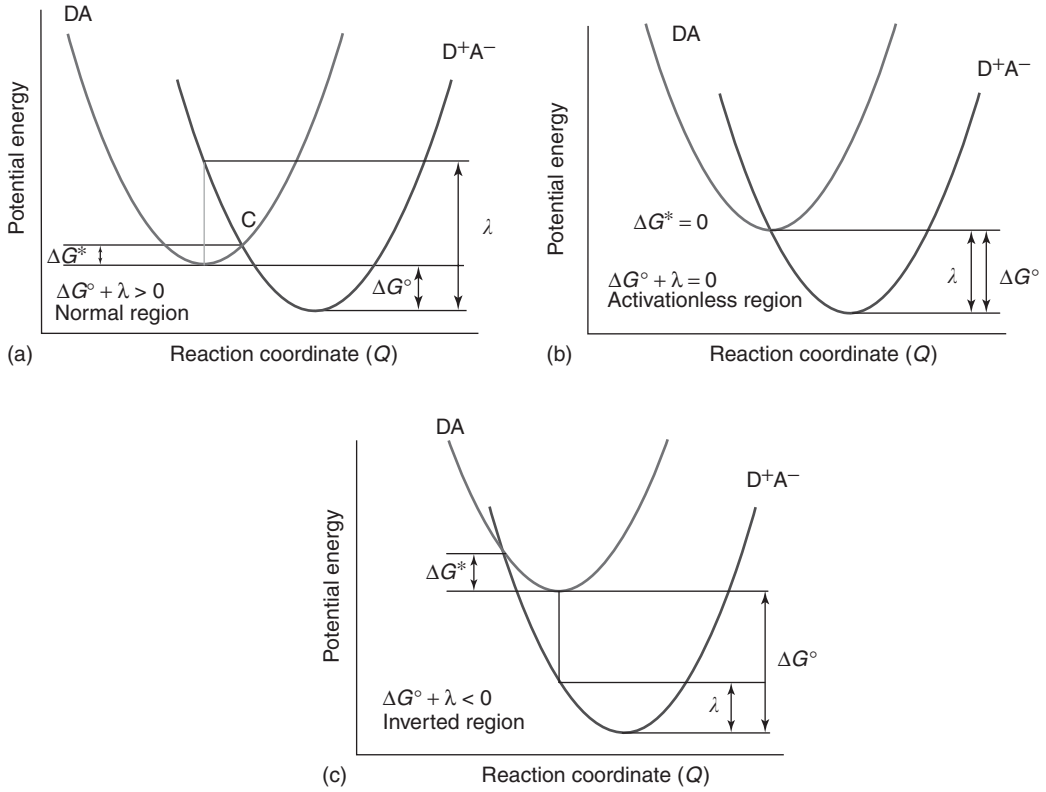


Figure 3 The harmonic potential energy of the states DA and D^+A^- as a function of the reaction coordinate Q, according to the Marcus theory, for three representative cases: $\Delta G^\circ + \lambda > 0$, normal region; $\Delta G^\circ + \lambda = 0$, activationless region; $\Delta G^\circ + \lambda < 0$, inverted region.

$[(\Delta G^\circ + \lambda) < 0]$, the slower the ET rate, the greater is the energy liberated in the reaction, such a condition being involved in the phenomenon of chemiluminescence. Furthermore, in the inverted region, the extent of electron tunneling may become more relevant.

The reorganization energy λ constitutes a crucial parameter in the ET process. Many evaluations by *ab initio* quantum mechanical calculations of λ have been done. However, despite their importance, direct and precise measurements of the reorganization energy λ are quite difficult. Generally, it can be extracted by measuring the dependence of the ET rate upon $-\Delta G^\circ$ and using the fact that, from the Marcus theory, the rate is expected to be maximal when $\Delta G^\circ = -\lambda$. Values for $\lambda \sim 0.7$ and 1 eV have been commonly found for intra and interprotein ET, respectively.

Perspectives and Final Remarks

Despite the large efforts devoted to investigate the ET process in metalloproteins, many aspects still remain far from being fully clarified. Recent developments in spectroscopic and scanning probe techniques, together with advanced computational approaches, could help in a further elucidation of the ET process in biomolecules. Laser ultrafast spectroscopy could get insights into the subtle relationships among the structure, dynamics, and functionality by also allowing one to explore the role of water dynamics in the ET process and to evaluate the reorganization energy. The mechanisms involved in intra and intermolecular ET could be deeply investigated by scanning tunneling microscopy and spectroscopy, whose role in the study of biological systems is rapidly growing. Molecular dynamics simulations, also integrated with quantum chemical calculations, can provide valuable information about possible ET pathways, elucidating at the same time, the interplay between the structural and dynamical behavior and the role played by collective motions.

From the practical side, the insights gained from these studies could lead to the introduction of new strategies to design synthetic ET systems with enhanced ET rates and efficiencies. In such a way, by exploiting their nanoscale dimensions, these molecules could be integrated into hybrid systems in the perspective to build nanodevices and nano-biosensors.

See also: Metalloproteins, Structural Determination of.

PACS: 82.30.Fi; 82.39.Jn; 82.39.Tt

Further Reading

Devault D (1980) *Quarterly Review of Biophysics* 13: 364–387.
Hopfield JJ (1974) *Proceedings of the National Academy of Sciences of the United States of America* 71: 3640–3644.
Jortner J and Bixon M (eds.) (1999) *Electron Transfer: from isolated molecules to biomolecules, Advances in Chemical Physics*. New York: Wiley.
Marcus RA and Sutin N (1985) *Biochimica et Biophysica Acta* 811: 265–322.
Moser CC, Keske JM, Warncke K, Farid RS, and Dutton PL (1992) *Nature* 366: 796–802.

Metalloproteins, Structural Determination of

C Luchinat and G Parigi, University of Florence, Florence, Italy

Introduction

About 30% of proteins in all living organisms contain a metal ion. Such metalloproteins may be diamagnetic or paramagnetic systems, depending on the nature of the metal ion and its spin state. The presence of a diamagnetic metal ion in the protein does not affect the NMR parameters appreciably, and the usual techniques for protein solution structure determination can be applied, with the additional problem of locating the metal ion. On the contrary, the presence of a paramagnetic metal ion largely perturbs the NMR parameters, as it affects both the signal line widths and the intensity of the nuclear Overhauser effects (NOEs). In fact, since the magnetic moment of an unpaired electron is 658.2 times larger than that of a proton, the major magnetic interaction of any nucleus not far from the metal ion will be provided by the dipolar coupling with the unpaired electron rather than with other closer nuclei (**Figure 1**). Such interaction sizably increases both their longitudinal and transverse relaxation rates – two parameters strongly affecting the protein NMR spectra.

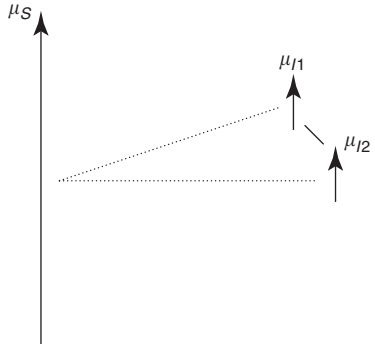


Figure 1 The magnetic interaction of a nuclear magnetic moment with the electron magnetic moment can be larger than that between two close nuclei, due to the larger electron magnetic moment.

If R_{1M} and R_{2M} indicate the paramagnetic contributions to the longitudinal and transverse relaxation rates, respectively, the increase in nuclear signal line width due to the presence of a paramagnetic metal ion will be given by

$$\Delta\nu = R_{2M}/\pi$$

and the change in NOE intensity by

$$\frac{\eta_{I(J)}^{\text{para}}}{\eta_{I(J)}^{\text{dia}}} = \frac{\rho_{I(J)} + \rho_{I(\text{other})}}{\rho_{I(J)} + \rho_{I(\text{other})} + R_{1M}}$$

where $\rho_{I(J)}$ is the contribution to longitudinal relaxation of nucleus I due to the coupling with nucleus J and $\rho_{I(\text{other})}$ includes all other contributions to relaxation of nucleus I, with the exception of the coupling with nucleus J and with the paramagnetic center, R_{1M}.

The increase in line widths and the decrease of NOE intensities make it more difficult to collect data for structure calculations. For $R_{1M} > R_{1\text{dia}}$ and $R_{2M} > R_{2\text{dia}}$, the integrals of NOESY cross-peaks decrease linearly with R_{1M}, and the scalar coupling cross-peaks decrease quadratically with R_{2M}. However, the presence of a paramagnetic metal ion makes available a new class of structural restraints: the paramagnetism-based restraints. The amount of information related to these new restraints is often able to compensate the loss of information due to the effects described above. Paramagnetism-based restraints are derived from contact shifts (CSs), pseudocontact shifts (PCSs), relaxation enhancements, and residual dipolar couplings (RDCs) arising from partial metalloprotein self-orientation, cross-correlation between Curie relaxation and dipole–dipole nuclear relaxation (CCR).

The availability of the paramagnetism-based restraints may also make it convenient to use a paramagnetic metal ion in the place of a native diamagnetic metal or as a probe to be attached to a protein.

Paramagnetism-based restraints have also the advantage of being often quicker to obtain than the classical NOE restraints, especially for proteins of large size. Therefore, the development of protocols for their efficient use is of great perspective importance in order to increase the speed in solving protein structures and to obtain structures of large proteins employing NMR.

The Magnetic Susceptibility in Diamagnetic and Paramagnetic Systems

All the paramagnetism-based effects indicated above originate from the dipolar coupling between protein nuclei and the unpaired electron(s) and from the presence of paramagnetic susceptibility, usually an anisotropic quantity, which is now introduced.

When a molecule is immersed in an external magnetic field $\boldsymbol{H}_0$, an induced magnetic moment is established along $\boldsymbol{H}_0$. A magnetization $\boldsymbol{M}$ thus arises, defined as the induced magnetic moment per unit volume, usually proportional to H_0:

$$M = \frac{\chi_V B_0}{\mu_0}$$

where μ_0 is the permeability of free space and $B_0 = \mu_0(H_0 + M) \approx \mu_0 H_0$. In systems that do not contain paramagnetic ions, χ_V is the diamagnetic susceptibility, usually independent of B_0 and is negative.

In paramagnetic systems, χ_V is dominated by the paramagnetic susceptibility per unit volume, it is independent of B_0 and is positive. An average induced magnetic moment per particle, $\langle \mu \rangle$, is, in fact, established due to the different populations of the electron energy levels, as a consequence of the fact that the magnetic field splits the S manifold of a particle with spin S according to its M_S values (**Figure 2**). The magnetic susceptibility per molecule, χ, is defined as

$$\chi = \frac{V_M \chi_V}{N_A} = \frac{V_M \mu_0 M}{N_A B_0} = \frac{\mu_0 \langle \mu \rangle}{B_0}$$

where V_M is the molar volume and N_A is the Avogadro constant.

The electron magnetic moment is related to the electron spin according to the following equation:

$$\boldsymbol{\mu} = -\mu_B g_e \boldsymbol{S}$$

where g_e is the electron g factor and μ_B is the electron Bohr magneton. The expectation value of $\boldsymbol{\mu}$ along the direction of the applied magnetic field z, is thus given by $-\mu_B g_e S_z$. Analogously, the average induced magnetic moment per particle is

$$\langle \mu \rangle = -\mu_B g_e \langle S_z \rangle$$

where $\langle S_z \rangle$, the expectation value of S_z, in the high-temperature approximation and in the limiting case where there is no contribution from the orbital magnetic moment, is given by

$$\langle S_z \rangle = -\frac{g_e \mu_B B_0}{3kT} S(S+1)$$

where k is the Boltzmann constant and T is the absolute temperature. Therefore, the Curie law is obtained:

$$\chi = \mu_B^2 g_e^2 \frac{\mu_0}{3kT} S(S+1)$$

If the orbital magnetic moment is considered, the magnetic susceptibility becomes anisotropic, and must be represented as a tensor. As a consequence, the average induced magnetic moment depends on the orientation of the $\boldsymbol{\chi}$-tensor with respect to the direction of the applied magnetic field:

$$\langle \boldsymbol{\mu} \rangle = \frac{\boldsymbol{\chi} \cdot \boldsymbol{B}_0}{\mu_0}$$

The energy of a magnetically anisotropic molecule, with average magnetic moment per particle $\langle \boldsymbol{\mu} \rangle$, in a magnetic field is given by

$$E^{\text{aniso}} = -\int_0^{\langle \mu \rangle} \boldsymbol{B}_0 \cdot d\langle \boldsymbol{\mu} \rangle = -\frac{\boldsymbol{B}_0 \cdot \boldsymbol{\chi} \cdot \boldsymbol{B}_0}{2\mu_0}$$

The fact that the energy is orientation dependent has an influence on the probability that the molecule orients along different directions.

The Dipole–Dipole Interaction

The point dipole–point dipole interaction between two particles possessing a magnetic moment is described

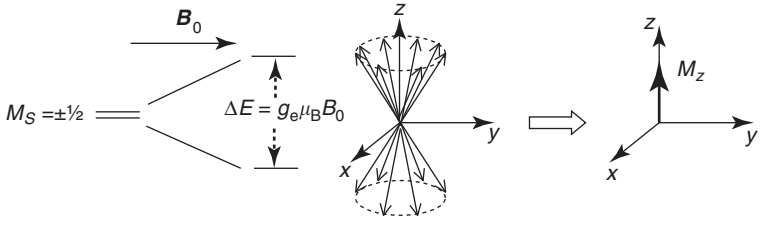


Figure 2 A magnetization arises due to the different population of the electron energy levels corresponding to $M_S = 1/2$ and $M_S = -1/2$.

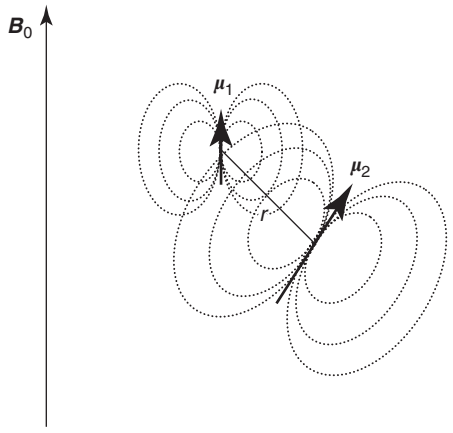


Figure 3 The dipole–dipole interaction between two magnetic moments. The nuclear magnetic moment is along $\boldsymbol{B}_0$, whereas the electron magnetic moment can be along a different direction due to the anisotropy of the magnetic susceptibility tensor.

by the Hamiltonian

$$H = -\frac{\mu_0}{4\pi}\left[\frac{3(\boldsymbol{\mu}_1\cdot\boldsymbol{r})(\boldsymbol{\mu}_2\cdot\boldsymbol{r})}{r^5} - \frac{\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2}{r^3}\right]$$

where μ_1 and μ_2 are the interacting magnetic moments and $\boldsymbol{r}$ is the vector connecting the two point dipoles (**Figure 3**). The magnetic moment of the proton is related to its spin $\boldsymbol{I}$ by the relationship $\mu_I = \hbar\gamma_I\boldsymbol{I}$, where γ_I is the proton magnetogyric ratio and $\hbar = h/2\pi$ is the Planck constant. Therefore, if the dominant interaction is between a proton and the average induced electron magnetic moment, the dipolar Hamiltonian can be written as

$$\begin{aligned} H &= -\frac{\hbar\gamma_I}{4\pi}\left[\frac{3(\boldsymbol{I}\cdot\boldsymbol{r})(\boldsymbol{r}\cdot\boldsymbol{\chi}\cdot\boldsymbol{B}_0)}{r^5} - \frac{\boldsymbol{I}\cdot\boldsymbol{\chi}\cdot\boldsymbol{B}_0}{r^3}\right] \\ &= \hbar\gamma_I\boldsymbol{I}\cdot\boldsymbol{\sigma}\cdot\boldsymbol{B}_0 \end{aligned}$$

where $\boldsymbol{\sigma}$ is called dipolar shielding tensor. In case a reference system is chosen with axes coinciding with the main directions of the $\boldsymbol{\chi}$-tensor and origin on the metal ion, the dipolar shielding tensor is

$$\sigma = -\frac{1}{4\pi r^5} \times \begin{pmatrix} (3x^2 - r^2)\chi_{xx} & 3xy\chi_{yy} & 3xz\chi_{zz} \\ 3xy\chi_{xx} & (3y^2 - r^2)\chi_{yy} & 3yz\chi_{zz} \\ 3xz\chi_{xx} & 3yz\chi_{yy} & (3z^2 - r^2)\chi_{zz} \end{pmatrix}$$

where χ_{xx}, χ_{yy}, and χ_{zz} are the principal components of the magnetic susceptibility tensor and x, y, and z are the coordinates of the nucleus in that reference system.

Pseudocontact Shift Restraints

The rotational average of the dipole–dipole interaction between the nucleus magnetic moment and the average induced electron magnetic moment causes a shift of the NMR signals, called PCS. It is calculated by taking the average of the values of the dipolar interaction energy along three principal directions, divided by the Zeeman energy $\hbar\gamma_I B_0$, and thus it is given by

$$\delta^{\text{pcs}} = -\frac{1}{3}\text{Tr}(\sigma)$$

Therefore, in the case where the reference frame coincides with the main directions of the $\boldsymbol{\chi}$-tensor, with the metal ion in the origin,

$$\delta^{\text{pcs}} = \frac{1}{12\pi r^3} \times \frac{(3x^2 - r^2)\chi_{xx} + (3y^2 - r^2)\chi_{yy} + (3z^2 - r^2)\chi_{zz}}{r^2}$$

The latter equation can be written in polar coordinates by using the ϑ and φ angles to specify the position of the nucleus in this reference frame,

$$\delta^{\text{pcs}} = \frac{1}{12\pi r^3} \times \left[\Delta\chi_{\text{ax}}(3\cos^2\vartheta - 1) + \frac{3}{2}\Delta\chi_{\text{rh}}\sin^2\vartheta\cos 2\varphi\right]$$

where $\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$ are the axial and rhombic anisotropy parameters of the magnetic susceptibility tensor of the metal,

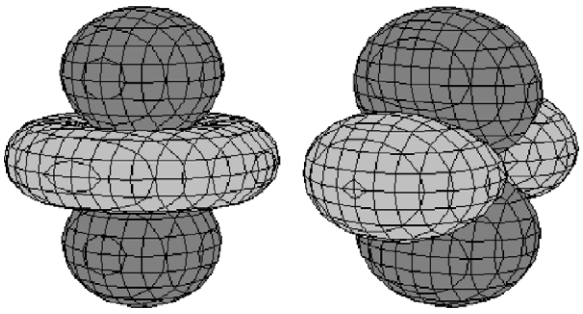
$$\Delta\chi_{\text{ax}} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2}, \quad \Delta\chi_{\text{rh}} = \chi_{xx} - \chi_{yy}$$

It should be noted that the PCS depends only on the anisotropy of $\boldsymbol{\chi}$ and not on its magnitude. If the magnetic susceptibility is isotropic, the rotational average of the dipolar energy is zero, and no PCS occurs.

In conclusion, PCS restraints depend on the coordinates of the protein atom, on the values of the magnetic susceptibility anisotropies, and on the three Euler angles defining the principal frame of the magnetic susceptibility tensor. **Figure 4** shows the form of the surface where a nucleus can stay, for a given positive (dark gray) or negative (light gray) value of the PCS and for a given magnetic susceptibility tensor, in the limiting cases of no ($\Delta\chi_{\text{rh}} = 0$) and maximal ($\Delta\chi_{\text{rh}} = 2\Delta\chi_{\text{ax}}/3$) rhombicity.

Relaxation Enhancement Restraints

The recovery to equilibrium of a nuclear spin system is characterized by two time constants, the longitudinal and the transverse relaxation times, indicated by T_1 and T_2, respectively, or by their inverse R_1 and R_2. The longitudinal relaxation time, also called spin–lattice relaxation time, is the time constant for



(a) (b)

Figure 4 Surface constraint for a nucleus with a positive (dark gray) or negative (light gray) PCS value in the case of (a) axial paramagnetic susceptibility tensor ($\Delta\chi_{rh} = 0$) or (b) maximal rhombicity ($\Delta\chi_{rh} = 2\ \Delta\chi_{ax}/3$) of the paramagnetic susceptibility tensor. The metal is located at the origin.

the evolution of the component of the nuclear magnetization along the magnetic field direction toward its equilibrium value. The more numerous and efficient the mechanisms allowing the nuclear spin to jump from one nuclear energy level to another, the shorter the longitudinal relaxation time.

The transverse, or spin–spin, relaxation time is related to the lifetime of the magnetization component in the plane perpendicular to the magnetic field direction, which is zero at equilibrium. The mechanisms causing transverse relaxation comprise those causing longitudinal relaxation, but they are more numerous/efficient, and therefore $R_2 \geqslant R_1$.

The dipolar interaction energy between the electron-spin magnetic moment and the proton magnetic moment can fluctuate in time due to changes in the metal–nucleus distance, for example, detachment of a coordinated molecule bearing the nucleus of interest, rotation of the complex with respect to the external magnetic field, or changes in the M_z or M_{xy} components of the electron-spin magnetic moment (electron relaxation). All these three processes represent mechanisms leading to nuclear relaxation, and thus causing the so-called paramagnetic enhancement of the longitudinal and transverse relaxation rates, R_{1M} and R_{2M}, respectively.

The resulting equation for R_{1M} is proportional to the average of the square of the dipolar interaction energy and to the appropriate spectral density functions:

$$R_{1M} = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 g_e^2 \mu_B^2 S(S+1)}{r^6} \times \left[\frac{\tau_c}{1+(\omega_I-\omega_S)^2\tau_c^2} + \frac{3\tau_c}{1+\omega_I^2\tau_c^2} + \frac{6\tau_c}{1+(\omega_I+\omega_S)^2\tau_c^2}\right]$$

where $\omega_I = \gamma_I B_0$ is the proton Larmor frequency, ω_S is the electron Larmor frequency ($\omega_S = 658.2\omega_I$), and τ_c is the correlation time related to the mechanisms responsible for relaxation:

$$(\tau_c^{dip})^{-1} = \tau_s^{-1} + \tau_r^{-1} + \tau_M^{-1}$$

given by the shortest among the electron relaxation time τ_s, the reorientational time τ_r, and the exchange time τ_M. In metalloproteins, τ_c is generally equal to τ_s.

Analogously, the equation for nuclear transverse relaxation rate enhancement is

$$R_{2M} = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 g_e^2 \mu_B^2 S(S+1)}{r^6} \times \left[4\tau_c + \frac{\tau_c}{1+(\omega_I-\omega_S)^2\tau_c^2} + \frac{3\tau_c}{1+\omega_I^2\tau_c^2} + \frac{6\tau_c}{1+(\omega_I+\omega_S)^2\tau_c^2} + \frac{6\tau_c}{1+\omega_S^2\tau_c^2}\right]$$

In the case of lanthanides and actinides, the term $g_J J(J+1)$ should replace $g_e S(S+1)$.

The former equations, calculated for systems without zero-field splitting and hyperfine coupling with the metal nucleus, are called Solomon equations. Such equations can be used to provide restraints for solution str7ucture determination as they contain a dependence on the distance between the resonating nucleus and the unpaired electron(s) at the sixth power, thus being very effective for locating the metal ion in the protein frame and for determining nuclear–metal distances for protons close to the paramagnetic center.

The modulation of the dipole–dipole interaction of the nuclear spin with the average induced electron magnetic moment $\langle \mu \rangle$ provides a further relaxation contribution. This relaxation mechanism is usually called magnetic susceptibility relaxation or Curie spin relaxation. Such interaction, of course, cannot be modulated by the electron relaxation, since $\langle S_z \rangle$ is already an average over the electron spin states, and therefore the relevant correlation time is determined by τ_r (in case, τ_M is longer). The contributions to R_{1M} and R_{2M} provided by this mechanism (in the approximation of isotropic magnetic susceptibility) are

$$R_{1M} = \frac{2}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\omega_I^2 g_e^4 \mu_B^4 S^2(S+1)^2}{(3kT)^2 r^6} \frac{3\tau_r}{1+\omega_I^2\tau_r^2}$$

$$R_{2M} = \frac{1}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\omega_I^2 g_e^4 \mu_B^4 S^2(S+1)^2}{(3kT)^2 r^6} \times \left(4\tau_r + \frac{3\tau_r}{1+\omega_I^2\tau_r^2}\right)$$

In the case of lanthanides and actinides, the term $g_J^4 \mu_B^4 (J(J+1))^2$, or more properly the experimental

μ_{eff}^4 value, when available, should replace $g_e^4\mu_B^4(S(S+1))^2$.

The effect of Curie relaxation on R_{2M}, and thus on line widths, is very important in paramagnetic metalloproteins, especially at high magnetic fields, due to the large value of the rotational timer τ_r, compared to the normally shorter electron relaxation time, which appears in the Solomon equation. The effect on R_{1M}, which does not contain nondispersive terms, is instead usually negligible.

Cross-Correlation between Curie and Dipole–Dipole Relaxation Restraints

A novel kind of restraint, yet to be fully exploited, is the cross-correlation arising between the dipole–dipole interaction between two nuclei (for instance, N and H in a protein amide bond) and the dipolar interaction between the magnetic moment of one of the two nuclei and the average induced electron magnetic moment, both the interactions being modulated by the rotational time. Such cross-correlation contributes to nuclear relaxation.

As for the cross-correlation present in diamagnetic systems between the dipole–dipole interaction and the chemical-shielding anisotropy, the cross-correlation between Curie and dipolar relaxation causes a difference in the line width of the two doublet components of, for example, the amide H (or the amide N). This difference in line width for the amide H, in the approximation of isotropic magnetic susceptibility, is given by

$$\Delta\nu_{1/2}^{\text{CCR}} = \frac{2}{15\pi}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{B_0\gamma_H^2\gamma_N\mu_B^2 g_e^2\hbar S(S+1)}{r^3 r_{HN}^3 kT} \times \frac{3\cos^2\theta_{MHN}-1}{2}\left(4\tau_r + \frac{3\tau_r}{1+\omega_I^2\tau_r^2}\right)$$

where the angle θ_{MHN} is that between the proton–nitrogen direction and the proton–metal ion direction and r is the proton–metal ion distance.

CCR restraints thus structurally depend on the inverse of the metal–proton distance at the third power and on the metal–proton–nitrogen angle. Their magnitude increases with the magnetic field and with the metal magnetic susceptibility, rather than with its anisotropy. Coupling between CH or CC pairs can also give rise to cross-correlation with Curie relaxation.

Self-Orientation Residual Dipolar Coupling Restraints

The dipole–dipole interaction between the magnetic moments of two unlike nuclei is given by

$$H = -\frac{\mu_0}{4\pi}\frac{\hbar^2\gamma_A\gamma_B}{r_{AB}^2} I_z^A I_z^B (3\cos^2\gamma - 1)$$

where r_{AB} is the distance between the two nuclei A and B and γ is the angle between the interspin vector $\boldsymbol{r}$ and the external magnetic field, along which both the nuclear magnetic moments are quantized. The isotropic rotational average value of the dipolar coupling is zero. However, in case the different orientations are not all equally populated, different weights must be considered for the different directions, and therefore an RDC arises:

$$\Delta\nu^{\text{rdc}} = \frac{\Delta E}{h} = -\frac{\mu_0\hbar\gamma_A\gamma_B}{4\pi^2 r_{AB}^3}\left\langle\frac{3\cos^2\gamma - 1}{2}\right\rangle$$

This can be written as a function of an orientation tensor-$\boldsymbol{S}$:

$$\Delta\nu^{\text{rdc}} = -\frac{\mu_0\hbar\gamma_A\gamma_B}{8\pi^2 r_{AB}^3}[S_{zz}(3\cos^2\Theta - 1) + (S_{xx} - S_{yy})\sin^2\Theta\cos 2\Phi]$$

where Θ is the angle between the AB vector and the z-axis of the orientation tensor, and Φ is the angle which describes the position of the projection of the AB vector on the xy-plane of the $\boldsymbol{S}$-tensor, relative to the x-axis.

As noted already, the presence of an anisotropic metal magnetic susceptibility induces PCS. It can also contribute to partial orientation of the molecule in high magnetic fields.

The degree of alignment of a molecule relative to a magnetic field can be described by the principal components of the orientation tensor $\boldsymbol{S}$, defined as

$$S_{ii} = \frac{\int \frac{3\cos^2\alpha - 1}{2}\exp\left[-\frac{E^{\text{aniso}}}{kT}(\alpha,\beta)\right]d\cos\alpha\,d\beta}{\int\exp\left[-\frac{E^{\text{aniso}}}{kT}(\alpha,\beta)\right]d\cos\alpha\,d\beta}$$

where α and β are the polar angles describing the direction of $\boldsymbol{B}_0$ in the molecular frame and $E^{\text{aniso}}(\alpha,\beta)$ is the anisotropic energy of the molecule in a field B_0. Therefore, one obtains:

$$S_{ii} = \frac{3}{2}\frac{B_0^2}{15\mu_0 kT}(\chi_{ii} - \bar{\chi})$$

The RDC due to paramagnetism can thus be expressed as a function of $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$,

$$\Delta\nu_{\text{rdc}}^{\text{para}} = -\frac{1}{4\pi}\frac{B_0^2}{15kT}\frac{\gamma_A\gamma_B\hbar}{2\pi r_{AB}^3} \times\left[\Delta\chi_{ax}(3\cos^2\Theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\Theta\cos 2\Phi\right]$$

where the angles Θ and Φ are defined in the principal frame of the paramagnetic susceptibility tensor. $\Delta\nu_{\text{rdc}}^{\text{para}}$ represents the difference in the measured ^1J

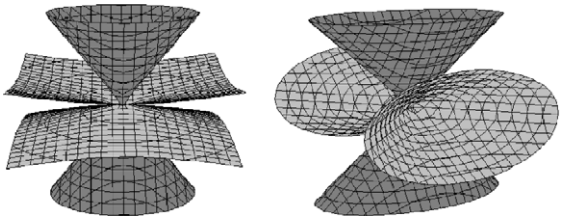


Figure 5 Surface constraint for a nucleus with a positive (dark gray) or negative (light gray) RDC value in the case of (a) axial magnetic susceptibility tensor or (b) maximal rhombicity of the magnetic susceptibility tensor. The coupled nucleus is located on the origin.

values between the paramagnetic system and a diamagnetic analog.

The functional form for PCS and RDC is the same. It is, however, important to note that the polar angles in the equation for RDC define the orientation of the coupled nuclei independently of their position with respect to the metal ion. **Figure 5** shows the surface restraint determined by a given value of the RDC and a given magnetic susceptibility tensor with no and maximal rhombicity.

Contact Shift (and Hyperfine Shift) Restraints

CSs are due to the presence of unpaired electron spin density onto the resonating nucleus, which occurs through direct spin delocalization and/or spin polarization. CSs contain structural information because they are somehow related to chemical bonds, although only in a few cases this information can be translated into structural constraints.

In the case of $Fe_4S_4^{2+}$ clusters (**Figure 6a**), it has been found that the CS of the βCH_2 protons of coordinated cysteines depends on the FeSCH dihedral angle θ as

$$\delta^c = a \sin^2 \theta + b \cos \theta + c$$

where $a \cong 10.3$, $b \cong -2.2$, and $c \cong 3.9$ ppm. The dominant dependence on $\sin^2\theta$ is ascribed to the fact that the spin density is in a *p*-orbital of the S- atom orthogonal to the Fe–S bond (**Figure 6b**). The relative values of the constant c with respect to a give an idea of the importance of the processes of electron delocalization through the σ-bond of Fe–S with respect to the π-bond of Fe–S. In these clusters, the observed hyperfine shifts are essentially contact shifts in origin.

The hyperfine shifts (given by the sum of contact and pseudocontact contributions) of heme methyl protons are given by the following heuristic equation in low-spin Fe(III) heme proteins containing a

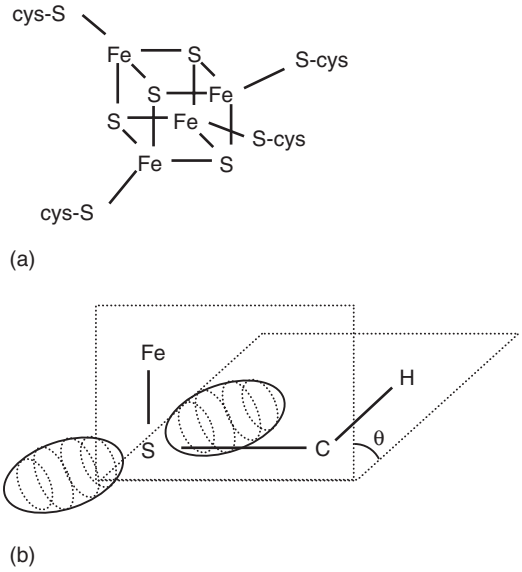


Figure 6 (a) Cybane-type Fe_4S_4 cluster. (b) Dihedral angle θ between the Fe–S–C plane and the S–C–H plane in the case of the spin density being in a *p*- orbital of the S- atom orthogonal to the Fe–S bond.

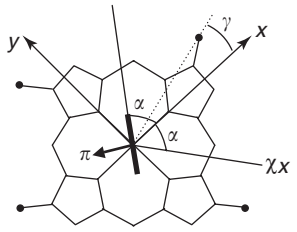


Figure 7 The γ angles for the four methyl groups (indicated with a dot) are defined as the angles between the metal–methyl *i*th vector and the *x*-axis, taken along the metal–pyrrole II direction. The angle α defines the direction of the histidine ring plane.

histidine and a cyanide as axial ligands, at 298 K,

$$\delta_i = a \sin^2(\gamma_i - \alpha) + b \cos^2(\gamma_i + \alpha) + c$$
$$a \cong 18.4, \quad b \cong -0.8, \quad c \cong 6.1$$

where γ_i is the angle between the metal–pyrrole II axis and the metal–*i*th-methyl direction, and α is the angle between the metal–pyrrole II axis and the direction of the histidine ring plane (**Figure 7**). In fact, the unpaired electron occupies one metal 3*d*-orbital that has the correct symmetry to form π-bonds with the axial ligand, and thus the contact contribution will be zero when the methyl protons lie in the nodal plane, and maximal when the methyl protons lie along the direction of the axial π-interaction, and thus $\delta_i^c \propto \sin^2(\gamma_i - \alpha)$. On the other hand, since the *z*-axis of the magnetic susceptibility tensor is always approximately perpendicular to the heme plane, the PCS is proportional to a function of the type $\delta^{PCS} \propto k \cos 2\varphi - l$, with the χ_x-axis rotated clockwise from

the metal–pyrrole direction by an angle which is the same in magnitude, but opposite in direction, to the angle α. Therefore, $\varphi = \gamma_i + \alpha$, and the equation given above is obtained.

Finally, the hyperfine shifts of methyl protons are given by the following heuristic equation in low-spin Fe(III) heme proteins containing two histidines as axial ligands:

$$\delta_i = \cos\beta[a\sin^2(\gamma_i - \alpha) + b\cos^2(\gamma_i + \alpha) + c] + d\sin\beta$$

$$a \cong 38.8, \quad b \cong -10.5, \quad c \cong -1.1, \quad d \cong 9.4$$

where α is the angle between the bisector of the angle β and the metal–pyrrole II axis, and β is the acute angle between the two histidine planes.

Similar considerations can be made for Fe(II) heme proteins with a single histidine in an axial position, when the electron configuration is high-spin and the shifts of the methyl protons are determined by a mixture of σ- and π-delocalization mechanisms.

Collecting the Paramagnetism-Based Restraints

The NMR parameters that can be obtained by performing NMR experiments are the chemical shift, the longitudinal relaxation rate, and the transverse relaxation rate. For a given nucleus, the chemical shift is defined as the resonating frequency of the nucleus with respect to a reference signal (in Hz). The chemical shifts result from the sum of two contributions: the diamagnetic shift, arising from the shielding of the observed nucleus from the external magnetic field due to electrons around, and the paramagnetic shift or hyperfine shift, which consist of the pseudocontact and the contact contributions:

$$\delta^{\text{exp}} = \delta^{\text{dia}} + \delta^{\text{c}} + \delta^{\text{pcs}}$$

As seen earlier, the contact contribution is present only for nuclei close to the paramagnetic center, since it can arise only if electron spin density can be delocalized on the investigated nuclei. Therefore, only the pseudocontact term can affect the hyperfine shift of amino acids that are not directly bonded to the metal center or H-bonded to the donor atoms.

In order to calculate δ^{pcs}, it is necessary to evaluate δ^{dia}. This can be done, to a good approximation, by performing the measurements on the analog diamagnetic protein, obtained by removing the paramagnetic metal ion, by substituting the metal ion with a diamagnetic metal with the same charge, or by reducing the paramagnetic metal to a diamagnetic state.

PCSs have been measured, among others, in many low-spin iron(III) heme containing proteins (**Figure 8** shows those measured in cytochrome b_5) and in lanthanide-substituted calcium-binding proteins (e.g., calmodulin, parvalbumin, and calbindin). As noted previously, their magnitude depends on the magnetic susceptibility anisotropy of the metal. Estimates of the latter are reported in **Table 1** for some metal ions. **Figure 9** shows the observed PCS in the protein calbindin D_{9k} after substitution of Ce(III), Yb(III) or Dy(III) in the C-terminal calcium-binding site. Metal ions characterized by a small magnetic anisotropy (e.g., Ce(III)) permit one to obtain values only for nuclei at short distances, whereas metal ions characterized by a large magnetic anisotropy (e.g., Dy(III)), although they do not provide information for nuclei close to the metal ion, due to the excessive line-broadening caused by Curie relaxation, permit the observation of values for nuclei that are much farther away from the metal.

Once the PCSs have been used to determine the magnetic susceptibility tensor and the structure is known, the contact contribution to the hyperfine shifts for atoms in amino acids close to the metal ion can be obtained. In fact, the pseudocontact contribution for such atoms can be calculated and subtracted from the hyperfine shift.

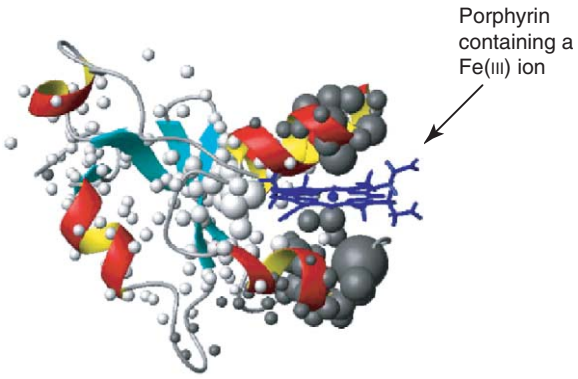


Figure 8 Observed PCS for the heme protein cytochrome b_5. The size of the balls is proportional to the magnitude of the negative (light gray) or positive (dark gray) PCS.

Table 1 Magnetic susceptibility and pseudocontact shifts for some metal ions

	χ_{av} (10^{-32} m^3)	$\Delta\chi_{ax}$ (%)	δ^{pcs} for $r = 10$ Å and $\theta = 0°$ (ppm)
Fe(III) HS	31	10	1.6
Fe(III) LS	3	80	1.3
Fe(II)	21	10	1.1
Co(II)	13	40	2.8
Ce(III)	5	35	0.9
Dy(III)	99	35	18.4
Yb(III)	15	45	3.6

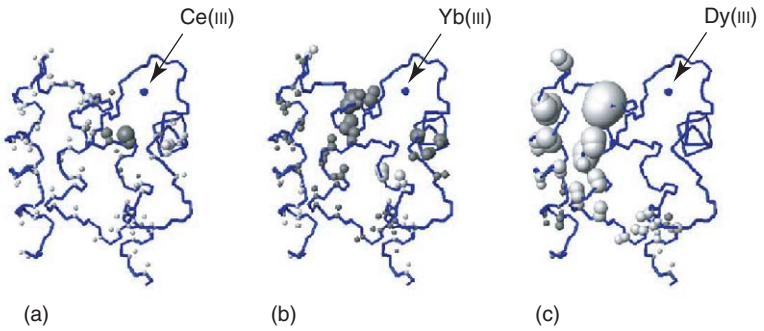


Figure 9 Observed PCS for the N and HN nuclei of the protein calbindin D_{9k} after substitution of (a) Ce(III), (b) Yb(III), or (c) Dy(III) in the C-terminal metal binding site. The size of the balls is proportional to the magnitude of the negative (light gray) or positive (dark gray) PCS.

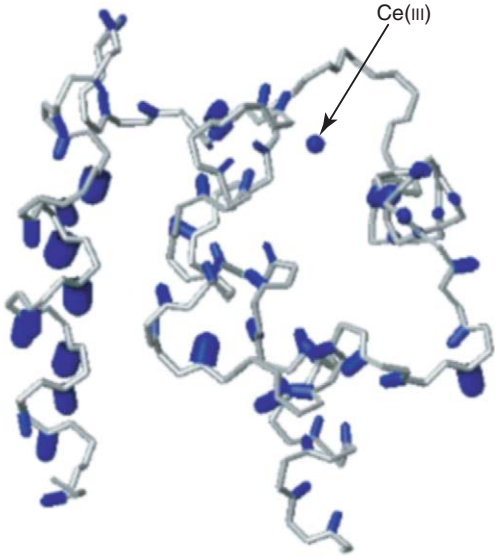


Figure 10 Observed self-orientation RDC for the HN atoms of the protein calbindin D_{9k} after substitution of Ce(III) in the C-terminal metal binding site.

Coupled nuclei experience ^1J splittings of the observed nucleus signals. Such splittings can be measured in HSQC spectra or in J-modulated experiments, and are given by the sum of the diamagnetic ^1J contribution, due to the scalar coupling and diamagnetic RDC, and the paramagnetic contribution. The values for $\Delta\nu_{rdc}^{para}$ can be obtained after subtraction of the ^1J measured at the same field and temperature on the analog diamagnetic protein, obtained by removing the paramagnetic metal ion, by substituting the metal ion with a diamagnetic metal with the same charge, or by reducing the paramagnetic metal to a diamagnetic state. Paramagnetic RDCs have been measured for many metalloproteins. Those related to the protein Ce(III)-calbindin D_{9k} are shown in **Figure 10**.

Relaxation rates are also measured as the sum of the diamagnetic and the paramagnetic contributions. Again, the diamagnetic contribution can be evaluated by performing measurements on the diamagnetic analog of the protein, by removing the paramagnetic metal ion or by substituting it with a diamagnetic one. Another way to estimate the diamagnetic contribution consists in taking an average among all the experimental relaxation rates that are below a given threshold value in the paramagnetic metalloprotein. In fact, R_{1M} is proportional to r^{-6}, and thus it becomes negligible for nuclei far from the paramagnetic center, whereas it is dominant for nuclei close to the paramagnetic center.

Transverse relaxation rates are related to signal line widths. The difference in the line width (in Hz) measured at half-height of, for example, the two components in the proton dimension of the HN doublet in the HSQC spectra of ^{15}N enriched paramagnetic metalloproteins is given by $\Delta\nu_{1/2}^{ccr}$, and it can be used as an additional restraint for protein solution structure determination. **Figure 11a** reports the observed values for amide proton cross correlations between Curie and dipolar relaxation in the Ce(III)-substituted protein calbindin D_{9k}, and **Figure 11b** reports the CCR values observed in metaquomyoglobin.

Solution Structure Determination Programs

Calculations of the solution structure of metalloproteins by employing paramagnetism-based restraints are normally performed using a modified version of the program DYANA, called PARAMAGNETIC DYANA, or the Xplor-NIH package. Both programs use molecular dynamics for energy minimization combined with a simulated annealing algorithm to calculate a family of structures starting from randomly

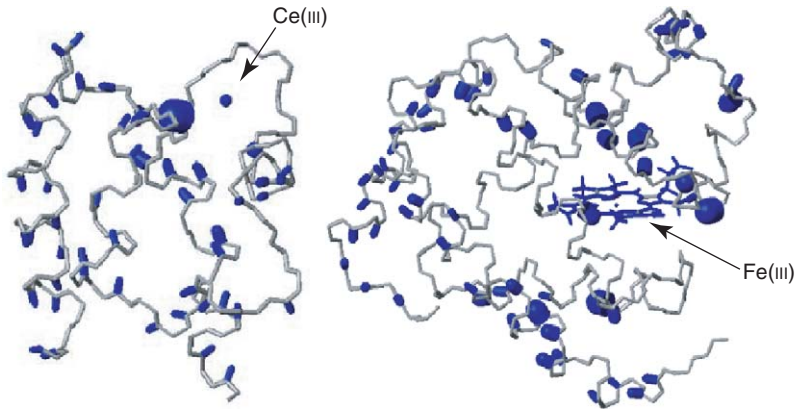


Figure 11 Observed CCR for the HN atoms (a) of the protein calbindin D_{9k} after substitution of Ce(III) in the C-terminal metal binding site, (b) of met-myoglobin, a heme protein containing high-spin Fe(III).

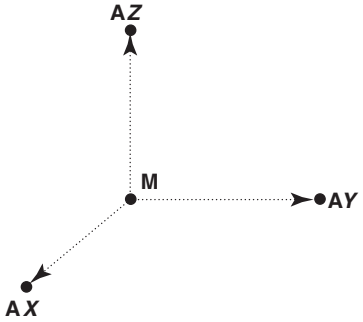


Figure 12 A pseudoresidue is defined containing pseudo-atoms needed to represent in the protein frame the position of the metal ion and the directions of the principal axes of the magnetic susceptibility anisotropy tensor.

generated conformers. Any metal ion and the relative magnetic susceptibility tensor axes are represented using pseudo-atoms as shown in **Figure 12**.

The paramagnetism-based restraints are introduced by calculating a pseudopotential E_l, for each class of new restraints

$$E_l = \sum_i w_i[\max(|X_{i,\mathrm{obs}} - X_{i,\mathrm{calc}}| - \mathrm{tol}_i, 0)]^2$$

to be added to the global target function, which is the quantity to be minimized. The index i runs over all experimental data of the considered class of restraints; tol indicates the tolerance, and w_i the weighting factor for that restraint. $X_{i,\mathrm{obs}}$ represents the experimental data and $X_{i,\mathrm{calc}}$ the calculated values in the present protein configuration according to the equations provided above.

PARAMAGNETIC DYANA and PARArestraints for Xplor-NIH use the PCS and RDC restraints by keeping fixed the values of the magnetic susceptibility anisotropies during the annealing process. These parameters ($\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$), which can be initially estimated from theoretical predictions, can be obtained by fitting the observed data on the protein structure. Therefore, both their values and the protein structure are actually obtained by a few cycles of structure calculations and fit on the protein structure, until convergence is reached.

Conclusions

It has been shown that paramagnetism-based restraints can provide relevant information for solution structure determination of paramagnetic metalloproteins, both around the metal ion and in the whole protein. CS can, in fact, identify the metal ligands and can provide dihedral angle restraints on the metal ligands. PCSs contain information on the coordinates of the metal ion and on the position of the protein atoms in the frame of the magnetic susceptibility tensor. Nuclear relaxation rates provide metal–nucleus distances. Cross-correlations contain information on distances and angles among the metal ion and coupled nuclei. Self-orientation RDCs provide information on the orientations of internuclear vectors in the magnetic susceptibility frame. All these restraints are vital for obtaining information around the metal ion, where it is usually missing, and can actually be precious to increase the accuracy of the protein structure. Furthermore, it is demonstrated that they provide enough information to be an alternative to classical diamagnetic restraints for protein solution structure determination.

See also: Metalloproteins, Electron Transfer in; Solid-State NMR Structural Studies of Proteins.

PACS: 87.15.−v; 82.56.−b; 82.56.Dj; 82.56.Pp; 82.56.Ub; 33.25.+k; 33.15.Kr; 61.18.Fs; 75.20.−g;

75.30. – m; 75.30.Gw; 75.30.Cr; 76.20. + q; 76.60. – k; 76.60.Cq; 76.60.Es

Further Reading

Banci L, Bertini I, Luchinat C, and Turano P (2000) Solution structures of hemoproteins. In: Kadish KM, Smith KM, and Guilard R (eds.) *The Porphyrin Handbook*, pp. 323–350. San Diego, CA: Academic Press.

Berliner LJ and Reuben J (eds.) (1993) *Biological Magnetic Resonance*, vol. 12, pp. 1–78. New York: Plenum.

Bertini I, Turano P, and Vila AJ (1993) NMR of paramagnetic metalloproteins. *Chemical Reviews* 93: 2833–2932.

Bertini I, Luchinat C, and Piccioli M (2001) Paramagnetic probes in metalloproteins. Turning limitations into advantages. *Methods in Enzymology* 339: 314–340.

Bertini I, Luchinat C, and Parigi G (2001) *Solution NMR of Paramagnetic Molecules*. Amsterdam: Elsevier.

Bertini I, Luchinat C, and Parigi G (2002) Magnetic susceptibility in paramagnetic NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* 40: 249–273.

Eaton DR, Fischer RD, Hawkins CJ, Horrocks WD Jr., Jesson JP, *et al.* (1973) *NMR of Paramagnetic Molecules*. New York: Academic Press.

Geraldes CFGC and Luchinat C (2003) Lanthanides as shift and relaxation agents in elucidating the structure of proteins and nucleic acids. *Metal Ions in Biological Systems* 40: 513–588.

Kurland RJ and McGarvey BR (1970) Isotropic NMR shifts in transition metal complexes: Calculation of the Fermi contact and pseudocontact terms. *Journal of Magnetic Resonance* 2: 286–301.

La Mar GN (ed.) (1995) *NMR of Paramagnetic Macromolecules*, NATO ASI series. Dordrecht: Kluwer Academic.

Markley JL, Ulrich EL, Westler WM, and Volkman BF (2003) Macromolecular structure determination by NMR spectroscopy. *Methods of Biochemical Analysis* 44: 89–113.

Ming L-J (2000) Nuclear magnetic resonance of paramagnetic metal centers in proteins and synthetic complexes. In: Que L Jr. (ed.) *Physical Methods in Bioinorganic Chemistry*, pp. 375–464. Sausalito, CA: University Science Books.

Osborne MJ, Crowe D, Davy SL, Macdonald C, and Moore GR (1997) NMR of paramagnetic proteins. *Methods in Molecular Biology* 60: 233–269.

Turner DL, Brennan L, Chamberlin SG, Louro RO, and Xavier AV (1998) Determination of solution structures of paramagnetic proteins by NMR. *European Biophysics Journal with Biophysics Letters* 27: 367–375.

Ubbink M, Worrall JA, Canters GW, Groenen EJ, and Huber M (2002) Paramagnetic resonance of biological metal centers. *Annual Review of Biophysics and Biomolecular Structure* 31: 393–422.

Nomenclature

$\boldsymbol{B}_0$	magnetic induction ($\mathrm{kg\,A^{-1}\,s^{-2}}$)
E_{S,M_S}	Zeeman energy of the level ($\mathrm{kg\,m^2\,s^{-1}}$)
g_e	electron g factor
$\hbar$	Planck constant $h/2\pi$ ($\mathrm{kg\,m^2\,s^{-1}\,rad^{-1}}$)
$\boldsymbol{H}_0$	external magnetic field ($\mathrm{A\,m^{-1}}$)
I	nuclear spin angular momentum
^1J	coupling constant ($\mathrm{s^{-1}}$)
k	Boltzmann constant ($\mathrm{kg\,m^2\,s^{-1}K^{-1}}$)
$\boldsymbol{M}$	magnetization ($\mathrm{A\,m^{-1}}$)
M_S	one-electron spin angular momentum quantum number
N_A	Avogadro constant ($\mathrm{mol^{-1}}$)
r	distance (m)
R_{1dia}	diamagnetic contribution to the longitudinal relaxation rate ($\mathrm{s^{-1}}$)
R_{2dia}	diamagnetic contribution to the transverse relaxation rate ($\mathrm{s^{-1}}$)
R_{1M}	paramagnetic contribution to the longitudinal relaxation rate ($\mathrm{s^{-1}}$)
R_{2M}	paramagnetic contribution to the transverse relaxation rate ($\mathrm{s^{-1}}$)
S	electron spin angular momentum
$\boldsymbol{S}$	orientation tensor
T	temperature (K)
V_M	molar volume ($\mathrm{m^3\,mol^{-1}}$)
γ_I	nuclear magnetogyric ratio ($\mathrm{rad\,kg\,A^{-1}\,s^{-3}}$)
δ	chemical shift
$\Delta\nu$	line width (Hz)
$\Delta\chi_{ax}$	axial magnetic susceptibility anisotropy ($\mathrm{m^3}$)
$\Delta\chi_{rh}$	rhombic magnetic susceptibility anisotropy ($\mathrm{m^3}$)
$\boldsymbol{\mu}$	magnetic moment ($\mathrm{A\,m^2}$)
μ_0	permeability of a vacuum ($\mathrm{kg\,m\,s^{-2}\,A^{-2}}$)
μ_B	electron Bohr magneton ($\mathrm{A\,m^2}$)
$\rho_{I(J)}$	contribution to longitudinal relaxation of nucleus I due to the coupling with nucleus J ($\mathrm{s^{-1}}$)
$\boldsymbol{\sigma}$	dipolar shielding tensor
τ_c	correlation time (s)
τ_M	exchange time (s)
τ_r	reorientational time (s)
τ_s	electron relaxation time (s)
$\boldsymbol{\chi}$	magnetic susceptibility per molecule ($\mathrm{m^3}$)
χ_v	susceptibility per unit volume
ω_I	nuclear Larmor frequency ($\mathrm{rad\,s^{-1}}$)
ω_S	electron Larmor frequency ($\mathrm{rad\,s^{-1}}$)

Metallurgy *See* Alloys: Aluminum; Alloys: Copper; Alloys: Iron; Alloys: Magnesium; Alloys: Overview; Alloys: Titanium; Crystalline Organic Metals.

Metals and Alloys, Electronic States of (Including Fermi Surface Calculations)

H Harima, Kobe University, Kobe, Japan

Introduction

Metals are usually or phenomenologically characterized by the physical properties of their excellent electrical and thermal conductivity, ductility, etc. From the point of view of electronic structure, crystals are clearly divided into metals and nonmetals: metals have Fermi surfaces and nonmetals do not. At zero temperature, in the Brillouin zone (which is defined as the Wigner–Seitz cell in the reciprocal space), Fermi surfaces separate two parts, in one of which the electron levels are occupied and in the other they are empty.

When the inversion symmetry is conserved in a paramagnetic crystal, as expected in many crystals, each state for electrons is doubly degenerate (the spin–orbit interaction lifts the degeneracy even in paramagnetic states, if the inversion symmetry is lost.) In this case, crystals with an odd number of electrons in the primitive unit cell must have Fermi surfaces, because the highest occupied electrons cannot fill whole of the Brillouin zone. Metals with this property are known as "uncompensated metals." On the other hand, crystals can be insulators, when the number of electrons in the primitive unit cell is even. Even in this case, they can be "compensated metals," where "compensated" means that the number of electrons equals the number of holes (holes represent the portion in the Brillouin zone unoccupied by electrons).

Most of the physical properties of metals are characterized by the Fermi surfaces, which represent the number of carriers, effective masses, anisotropy, etc. When metals undergo phase transition to superconductivity, magnetic ordering, or another ordered state, at low temperatures, the Fermi surfaces are related to the physical properties of the transition.

In early years, most of the elemental metals were comprehended based on the nearly free electron model. However, for most of the ordered alloys and compounds, numerical simulation is necessary to represent the Fermi surfaces. Otherwise, it is not possible to predict, for most of them, the metallic/nonmetallic nature. In recent years, Fermi surfaces of many crystals have been determined precisely by such simulation with highly developed computer resources and one-particle approximation, which is mostly the local-density approximation (LDA) in the density-functional theory.

Here, electronic structures and the Fermi surfaces are briefly reviewed. All of the figures presented are originally calculated and drawn by using a full-potential linear augmented plane-wave method and the LDA.

Nearly Free Electron Metals

The simplest picture for metals is given by a nearly free electron model. In this model, electrons behave as if they were noninteracting, moving through a weak periodic potential. The wave function of a Bloch state with the irreducible representation $\boldsymbol{k}$ in the translational group can be written as

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{K}} c_{\boldsymbol{k}-\boldsymbol{K}} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{K})\cdot\boldsymbol{r}}$$

where $\boldsymbol{K}$ stands for reciprocal lattice vectors. The coefficient $c_{\boldsymbol{k}-\boldsymbol{K}}$ and the energy level $E_{\boldsymbol{k}}$ are determined with the Fourier components of the crystal potential $U_{\boldsymbol{K}}$ as

$$\left[\frac{\hbar^2}{2m}(\boldsymbol{k}-\boldsymbol{K})^2 - \varepsilon_{\boldsymbol{k}}\right] c_{\boldsymbol{k}-\boldsymbol{K}} + \sum_{\boldsymbol{K}'} U_{\boldsymbol{K}'-\boldsymbol{K}} c_{\boldsymbol{k}-\boldsymbol{K}'} = 0$$

If $U_{\boldsymbol{K}} \sim 0$ for any $\boldsymbol{K}$, $\varepsilon_{\boldsymbol{k}} \sim (\hbar^2/2m)(\boldsymbol{k}-\boldsymbol{K})^2$

These electrons can be realized in some alkali and alkaline-earth metals and aluminum. The energy dispersion for Na in the first Brillouin zone almost satisfies $\varepsilon_{\boldsymbol{k}} = (\hbar^2/2m)(\boldsymbol{k}-\boldsymbol{K})^2$, as shown in **Figure 1**. The density of states $D(E)$ has the energy dependence as $D(E) \sim (m/\hbar^3\pi^2)\sqrt{2mE}$ just like free electrons, except the fine structure from the band splitting due to the effect of the weak crystal potential. (Note that the unit of $D(E)$ in figures is "states/(Ry · Primitive cell).") Electrons in Mg and Al also show such free-electron-like energy dispersions, even in other types of empty lattices, as shown in **Figures 2** and **3**.

In Na, Mg, and Al, the valence electrons actually have nodes (2 for 3*s*, 1 for 3*p*) due to the large attractive potential near the nucleus, so they are never free electrons. This situation can be understood as follows. A Bloch state is generally written as

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(\boldsymbol{r}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$$

where n stands for a quantum number inside one atom and $u_{n\boldsymbol{k}}(\boldsymbol{r})$ is a function with the periodicity of

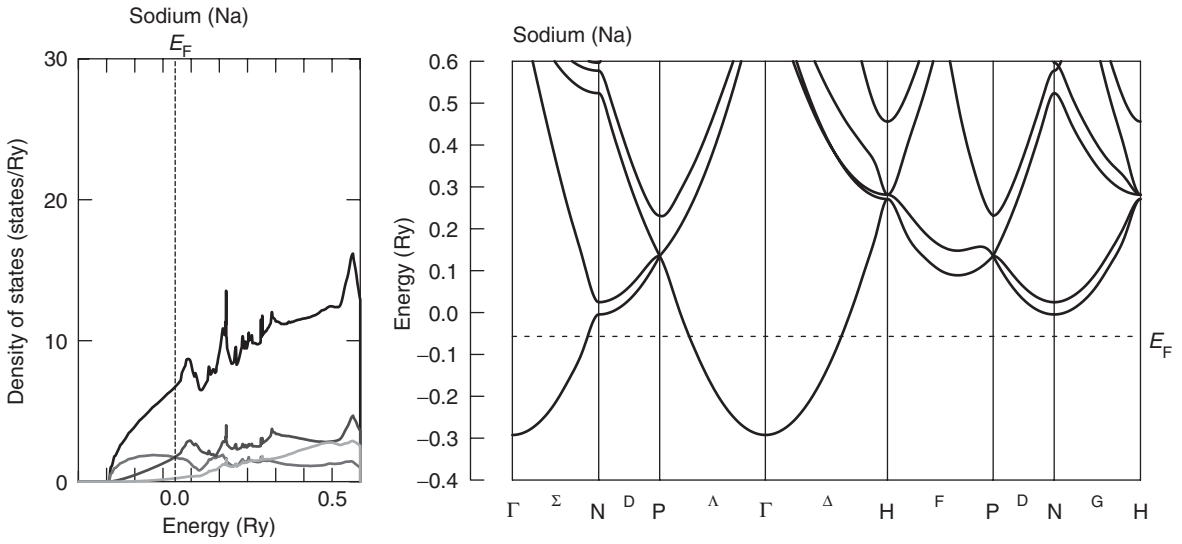


Figure 1 Energy band structure (right) and the density of states (left) of sodium in the body-centered cubic (b.c.c.) structure.

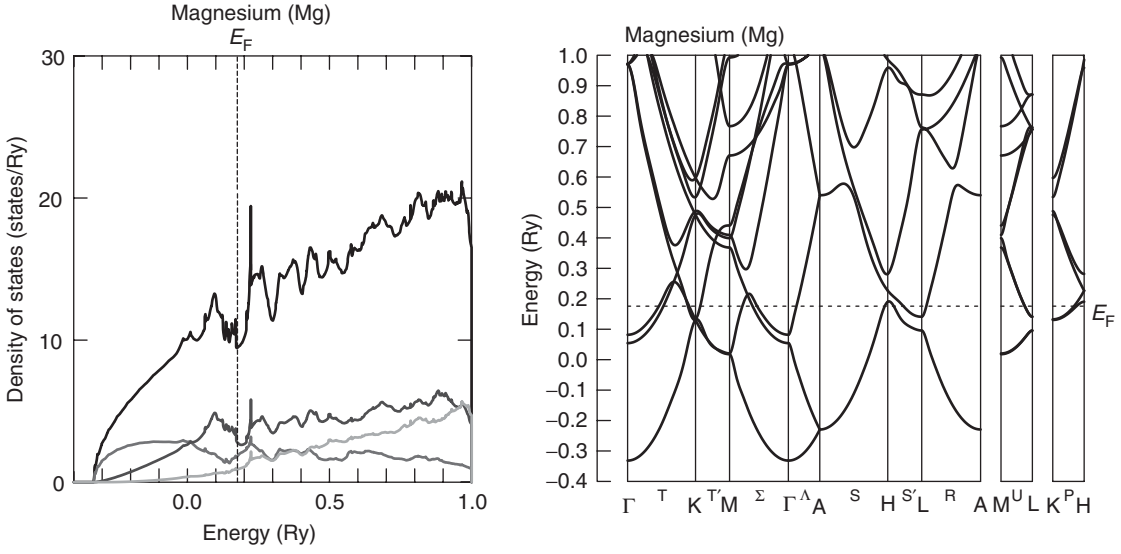


Figure 2 Energy band structure (right) and the density of states (left) of magnesium in the hexagonal structure.

the Bravais lattice:

$$u_{nk}(\boldsymbol{r}+\boldsymbol{R}) = u_{nk}(\boldsymbol{r})$$

The atomic orbitals are sometimes used for $u_{nk}(\boldsymbol{r})$.

The simplest case of $u_{nk}(\boldsymbol{r}) = \text{const.}$ obviously leads to a plane wave. Even when $u_{nk}(\boldsymbol{r})$ oscillates very near the nucleus, it can behave like free electrons, if $u_{nk}(\boldsymbol{r}) \sim \text{const.}$ far from the nuclei, that is, in the interstitial regions. To relate the plane waves with atomic orbitals, one should remember the well-known identity

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(kr) Y_l^{m*}(\hat{\boldsymbol{k}}) Y_l^m(\hat{\boldsymbol{r}})$$

where $j_l(kr)$ are spherical Bessel functions of order l, $Y_l^m(\hat{\boldsymbol{r}})$ are spherical harmonics, and l and m are the angular momentum and magnetic quantum numbers. Electrons with the radial wave function $R_l(\varepsilon, r)$, with $r < S$, can be treated as free electrons, when

$$\frac{\mathrm{d}}{\mathrm{d}r} \ln R_l(\varepsilon, r)\bigg|_S \sim \frac{\mathrm{d}}{\mathrm{d}r} \ln j_l\left(\frac{\sqrt{2m\varepsilon}}{\hbar} r\right)\bigg|_S$$

for a large S. This is realized in Na, Mg, and Al, because the positive charges of nuclei are well screened by the Ne core ($(1s)^2(2s)^2(2p)^6$) electrons.

Therefore, the plane-wave bands consist of higher-l components. In the higher-symmetry points in the Brillouin zone, the bands can be distinguished into

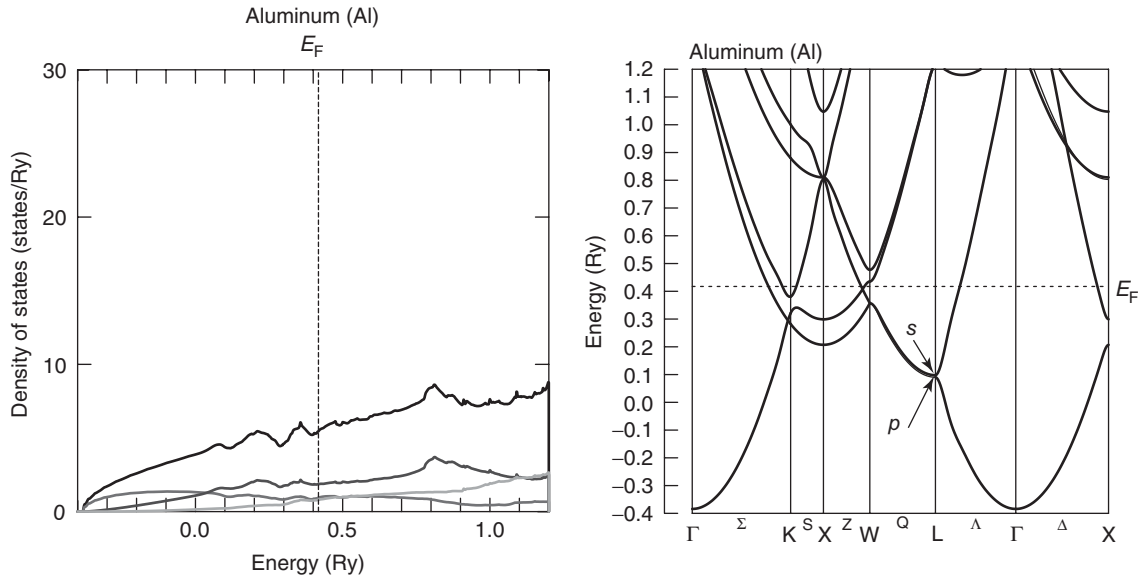


Figure 3 Energy band structure (right) and the density of states (left) of aluminum in the f.c.c. structure.

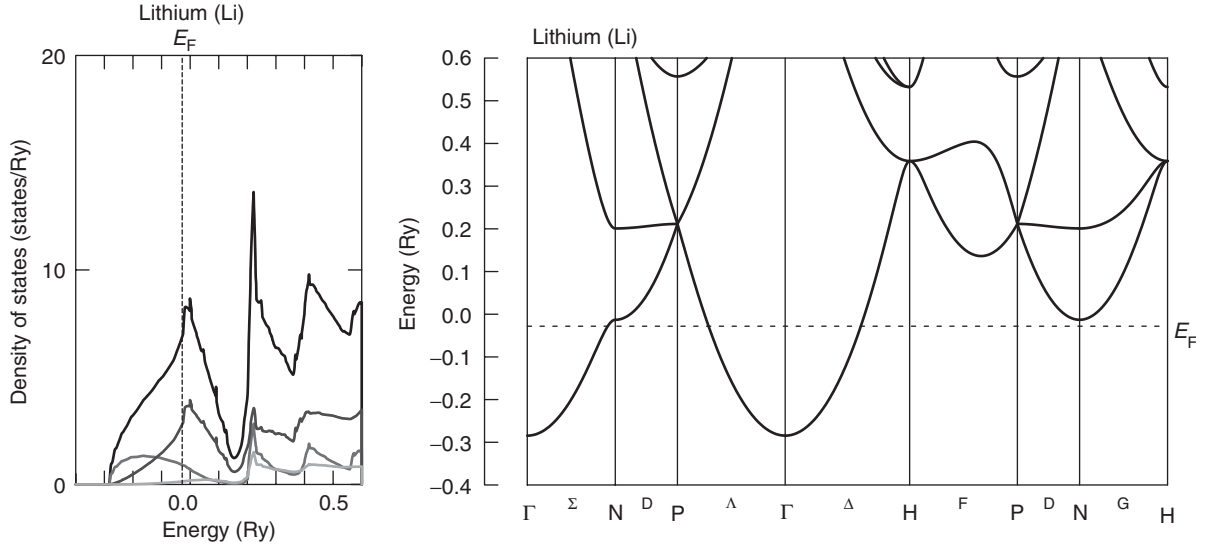


Figure 4 Energy band structure (right) and the density of states (left) of lithium in the b.c.c. structure.

the pure *s* or *p* bands, for example, the bottom contains only the *s* component. At the Brillouin zone boundary, the *s* band and *p* band are nearly degenerated, which is clearly seen at the L point in the case of Al, as in **Figure 3**.

The Fermi surfaces of Al are complicated in shape, but easily comprehended by a systematic procedure to convert a spherical Fermi surface into the first Brillouin zone, as described by Harrison in 1959.

The valence electrons in Li and Be with $(1s)^2$ core electrons do not behave as those in Na and Mg, as shown in **Figures 4** and **5**. The crystal potential affects the energy dispersions, particularly in the Brillouin zone boundary. The electronic structure of Li should be recognized as a mixture between atomic orbitals and free electrons. Actually, the depression in the density of states is clearly seen. Li with a half-filled band shows, however, a similar Fermi surface as in free electrons, while the density of states at the Fermi level in the depression in Be (four electrons in the primitive cell) is much smaller than that expected from a free-electron model.

Noble Metals (Cu, Ag, Au)

In noble metals, there are 11 valence electrons as $(nd)^{10}((n+1)s)^1$, with $n=3$ (Cu), 4 (Ag), and 5

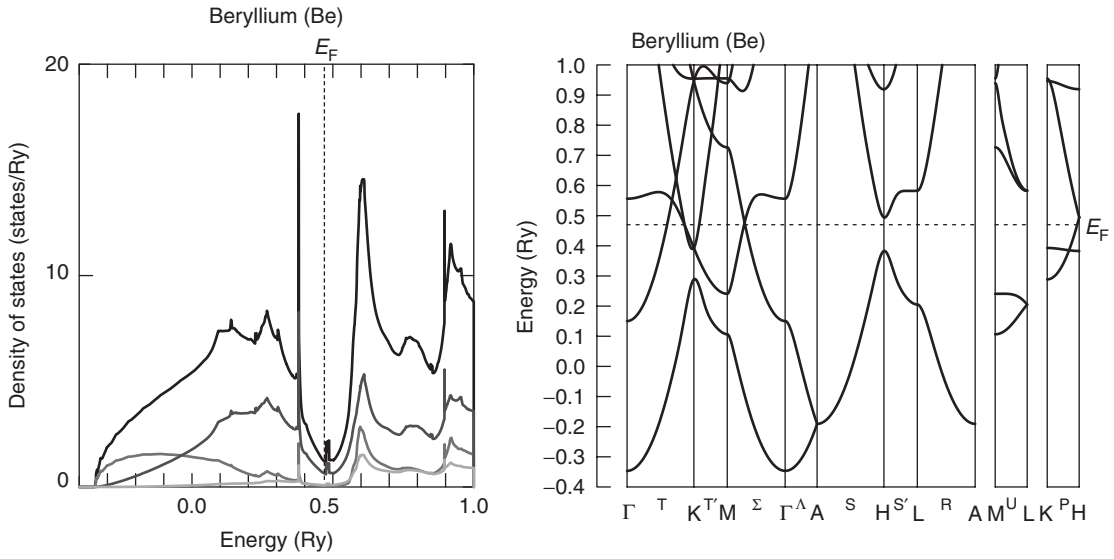


Figure 5 Energy band structure (right) and the density of states (left) of beryllium in the hexagonal structure.

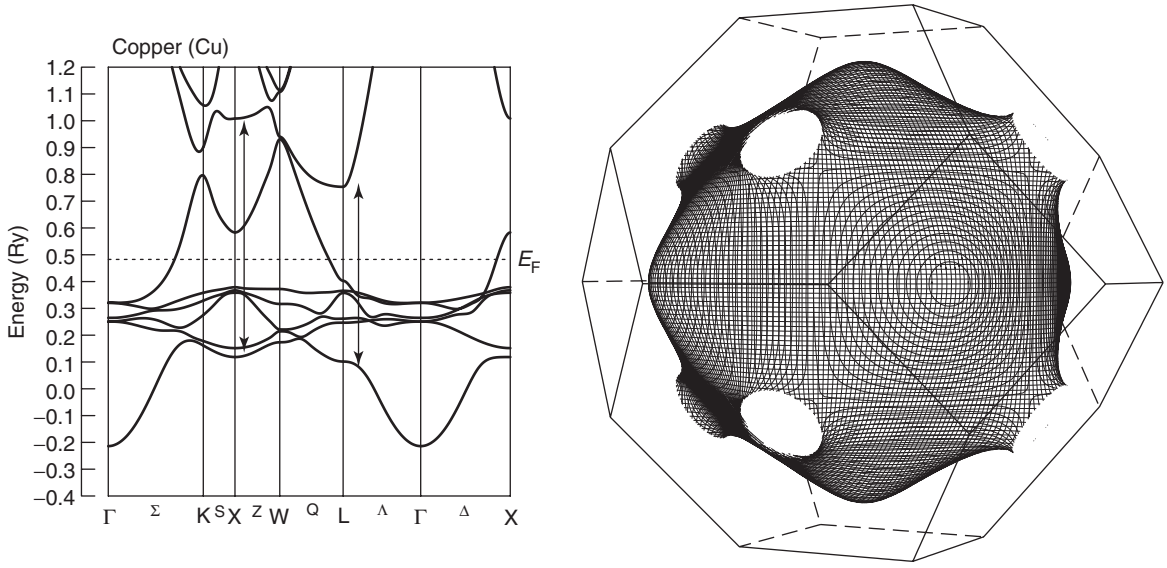


Figure 6 Energy band structure (left) and the Fermi surface (right) of copper. The arrows indicate the hybridization between the bonding and the antibonding states. Compare the band structure of Al in **Figure 3**.

(Au). $(nd)^{10}$ electrons are almost fully occupied, hence the noble metals are often thought to be monovalent metals. In fact, *s* and *p* electrons in noble metals behave as free electrons such as those in Na, while *d* electrons have the characteristics of the atomic orbitals, because the main part of the wave function is located much more inside than that of *s* and *p* electrons. Thus, the five bands with small energy dispersions for *d* electrons appear in the electronic band structure in Cu, as in **Figure 6**, where the *d* bands are located below the Fermi level. Then, the noble metals are monovalent metals. The plane-wave band with *s* character can hybridize with the *d* bands in the whole of the Brillouin zone, resulting in bonding and antibonding bands, while the plane-wave band with *p* character is left undisturbed. In the f.c.c. lattice, the *p*-character band shows the lowest energy at the L point, which is the closest point to the Γ point, on the Brillouin zone boundary.

Consequently, the conduction band is similar to the plane wave, except the low-lying part near the L point. Therefore, the Fermi surface is nearly spherical with the connected part, called the neck, centered at the L point, as shown in **Figure 6**.

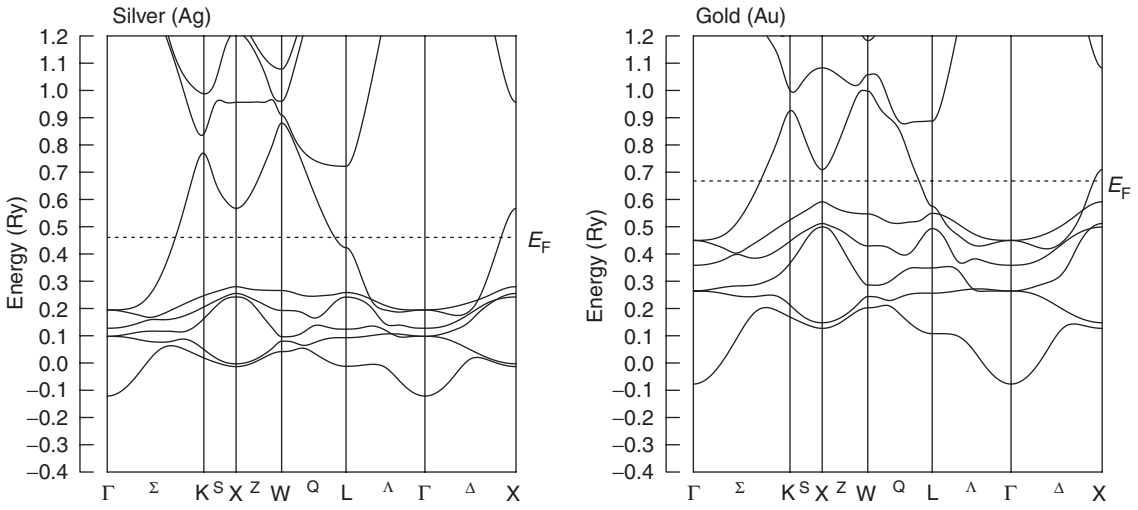


Figure 7 Energy band structure of silver (left) and gold (right).

The electronic structures of silver and gold resemble that of copper, except the location and the width of the *d* bands, as shown in **Figure 7**. The bandwidths become larger from copper to gold, because 5*d* electrons have a larger tail outside the core electrons. Meanwhile, the position of the *d* bands is the deepest in silver, so that there is no absorption from the *d* bands in the energy region of the visible spectrum. As the conduction band is less affected by the deeper *d* bands, the neck is the smallest in silver.

The core electrons screen the nucleus charge for the 4*d* electrons in silver to a lesser degree than for the 3*d* electrons in copper; the energy level of the 4*d* electrons becomes relatively deeper than that of the 3*d* electrons. One can expect much deeper *d* bands in gold; however, the relativistic effect, which is more prominent in heavier atoms, shrinks the core electrons. Then the nucleus becomes more screened in gold. The shrunk core is also noticed in the similar values of atomic radii for silver and gold. (The lattice constants of copper, silver, and gold are 3.615 Å, 4.086 Å, and 4.078 Å, respectively.)

Transition Metals, Alloys, and Compounds

The *d* electrons of the valence band in transition metals can be described within a band theory in which the effect of strong Coulomb interactions is treated as a mean field, because the underlying wide and plane-wave-like *s* and *p* electrons screen such large Coulomb interactions. The band theory explains their electrical and magnetic properties.

In some transition-metal oxides, such as high-temperature superconductors, Coulomb interactions are not negligible, because the oxygen *s* and *p* bands push up *s* and *p* electrons of the transition metal beyond the *d* bands; then the effective Coulomb interactions in *d* electrons become relatively large.

f-Electron Systems

The *f* electrons, which are valence electrons in lanthanide (4*f*) and actinide (5*f*) systems, are located well inside the outer closed shell (Xe or Rn). Even so, the 5*f* electrons are described as *d* electrons in transition metals. Particularly, the 4*f* electrons are less extended beyond the core electrons, and in many cases, are thought to be localized electrons, although the energy levels are comparable with other valence electrons. Such localized electrons, in general, do not affect electronic structures around the Fermi level. A band theory where the 4*f* electrons are treated in the Hartree–Fock approximation, for example, the LDA + *U* method, can predict the energy band structure of such compounds.

Fermi Surface Calculation

The currently available computer resources allow one to calculate energy eigenvalues at each $\boldsymbol{k}$ point very accurately, although some approximation for the exchange correlation potential must be used. When the results are compared with experiments, in many cases, the eigenvalues are interpolated by using plane waves like

$$F(\boldsymbol{k}) = \sum_{m=1}^{M} a_m \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_m}$$

where a_m should be determined under the conditions $F(\boldsymbol{k}_n) = F_n$. Here F_n is an eigenvalue at the point $\boldsymbol{k}_n$

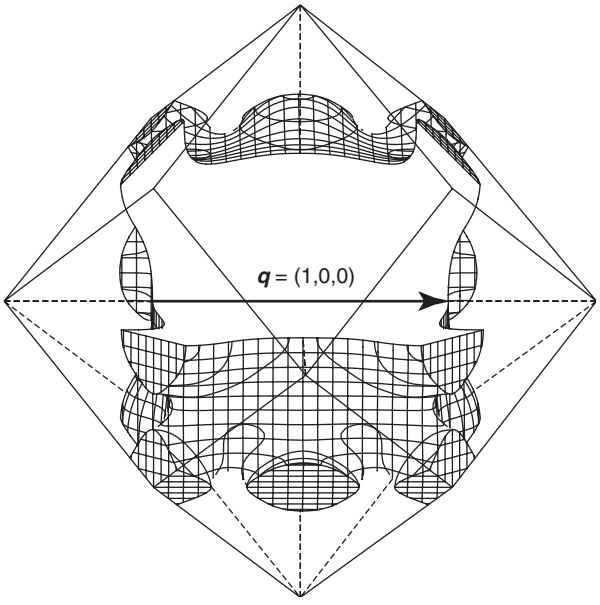


Figure 8 The Fermi surface of $PrRu_4P_{12}$ calculated by using an LDA + *U* method. One quarter is cut for the eye convenience. The volume of the Fermi surface is just a half of the Brillouin zone for the b.c.c. lattice. The arrow indicates the nesting vector.

($n = 1, N$). If $M = N$, a_m can be determined uniquely, but it is not effective. Usually, a_m is determined for $M > N$ under the conditions to minimize $\sum_m \rho_m a_m^2$, where ρ_m is a weight, for example, $1 + C_1|\boldsymbol{R}_m|^2 + C_2|\boldsymbol{R}_m|^4$, to be chosen to prevent higher components in $F(\boldsymbol{k})$. Typically, $M \sim 1.2 \times N$ for $N = 100$–500 gives sufficient accuracy to be compared with Fermi surface measurements.

Nesting Fermi Surface

The shape of the Fermi surface is related to a variety of instabilities of the metal. If the generalized susceptibility

$$\chi_0(\boldsymbol{q}) = \sum_{\boldsymbol{k}} \frac{f(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}}) - f(\varepsilon_{\boldsymbol{k}})}{\varepsilon_{\boldsymbol{k}} - \varepsilon_{\boldsymbol{k}+\boldsymbol{q}}}$$

shows a large value, instability with the $\boldsymbol{q}$ vector is realized, for example, relevant in magnetic ordering, the spin density of wave (SDW), or the charge density of wave (CDW). It happens when some part of the Fermi surface, if displaced by $\boldsymbol{q}$, coincides with another part. This is known as the nesting Fermi surface. The antiferromagnetic chromium may have an SDW state due to such a nesting mechanism.

The nesting mechanism can eliminate a part or whole of the Fermi surface to stabilize an SDW or a CDW state. If the entire Fermi surface is wiped out, it leads to a metal–insulator transition. This is possible for Fermi surfaces that resemble a checkerboard. Even in a cubic system, such a Fermi surface is realized in a skutterudite $PrRu_4P_{12}$, whose Fermi surface is a three-dimensional checkerboard, as shown in **Figure 8**. $PrRu_4P_{12}$ really undergoes a metal–insulator transition at 60 K.

See also: Alloys: Overview; Conductivity, Electrical; Fermi Surface Measurements; Insulators, Electronic States of; Intermetallic Compounds, Electronic States of; Magnetoresistance Techniques Applied to Fermi Surfaces; Metallic Bonding and Crystals; Metals and Metallic Alloys, Optical Properties of; Superconductors, Metallic; Tight-Binding Method in Electronic Structure; Transition Metal Compounds, Electronic and Magnetic Properties of.

PACS: 71.15.−m; 71.15.Rf; 71.18.+y; 71.20.−b; 71.20.Be; 71.20.Dg; 71.20.Eh; 71.30.+h; 71.45.Lr

Further Reading

Harima H, Takegahara K, Ueda K, and Curnoe SH (2003) Origin of the metal-insulator transition in $PrRu_4P_{12}$. *Acta Physica Polonica B* 34: 1189–1192.

Harrison Walter A (1959) Fermi surface in aluminum. *Physical Review* 116: 555–561.

Kübler J (2000) *Theory of Itinerant Electron Magnetism*. New York: Oxford Science Publications.

Metals and Alloys, Impurity and Defect States in

J S Faulkner, Florida Atlantic University, Boca Raton, FL, USA

Introduction

The most commonly used methods for treating the impurity and defect states in metals and alloys start with the calculation of the one-electron Green's function $G(E, \boldsymbol{r}, \boldsymbol{r}')$. The density-functional theory (DFT) makes it possible to write a simple equation for this function

$$\left[E + \frac{\hbar^2}{2m}\nabla^2 - V(\boldsymbol{r})\right] G(E, \boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \quad [1]$$

The local density approximation or the generalized gradient approximation is used to generate the effective one-electron potential $V(\boldsymbol{r})$. In any version of the DFT, $V(\boldsymbol{r})$ is a functional of the density of the electrons in the system, $\rho(\boldsymbol{r})$, which means that the one–electron equation must be solved self–consistently. The charge density is obtained from the Green's function by integration

$$\rho(\boldsymbol{r}) = -\frac{2}{\pi}\int_{-\infty}^{E_F} G(E,\boldsymbol{r},\boldsymbol{r})\ \mathrm{d}E \qquad [2]$$

where the Fermi energy E_F is defined by the requirement that the integral of $\rho(\boldsymbol{r})$ over all space is equal to the total number of electrons, and the 2 is required by the electron spins. The total energy is

$$E_{\text{total}} = \sum_{\text{occ.}} E_i - \frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' + E_{xc}[\rho] - \int \frac{\delta E_{xc}[\rho]}{\delta\rho(\boldsymbol{r})}\,\rho(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \qquad [3]$$

The sums are over all occupied states including spin, and the exchange correlation functional $E_{xc}[\rho]$ is determined by the particular method used to calculate the effective one-electron potential.

The one-electron potential for a metal or alloy can be written as a sum of potentials that describe the interaction of the electron with each of the atoms in the solid

$$V(\boldsymbol{r}) = \sum_{i=1}^{N} v_i(\boldsymbol{r}-\boldsymbol{R}_i) \qquad [4]$$

In order to eliminate surface states, it is convenient to imagine that the N atoms in the solid are in a supercell, and that the supercells are periodically reproduced to fill all space. The potential associated with the ith atom with its nucleus at $\boldsymbol{R}_i$, $v_i(\boldsymbol{r}-\boldsymbol{R}_i)$, is zero outside of the region Ω_i. The Ω_i fill the entire supercell and do not overlap. At some point, N is allowed to approach infinity and the supercell to fill all space.

The Green's Function from Multiple Scattering Theory

The multiple scattering equations for wave functions are derived and discussed elsewhere in the encyclopedia. The version of these equations that deals with Green's functions starts from the operator equation

$$G(z) = (z - H_0 - V)^{-1} = G_0(z)\,[I + VG(z)] = G_0(z) + G_0(z)\,T(z)\,G_0(z) \qquad [5]$$

where H_0 is the kinetic energy operator, z is a complex energy, and

$$G_0(z) = (z - H_0)^{-1} \qquad [6]$$

For the case that the potential is a sum of atomic potentials as in eqn [4], the scattering matrix for the system

$$T(z) = V[I + G_0(z)\,T(z)] \qquad [7]$$

can be written

$$T = \sum_{i=1}^{N}\sum_{j=1}^{N}\tau^{ij} \qquad [8]$$

where the scattering path operator τ^{ij} is

$$\tau^{ij} = t^i\delta_{ij} + t^i G_0 \sum_{k\neq i}^{N}\tau^{kj} \qquad [9]$$

and $t^i = (1 - v_i G_0)^{-1} v_i$ is the t-matrix that describes the scattering from $v_i(\boldsymbol{r}-\boldsymbol{R}_i)$ embedded in a vacuum.

A function that is very useful in analyzing experiments is the density-of-states (DOS) $\rho(E)$, which is the number of energy levels between the energies E and $E + \mathrm{d}E$. The DOS associated with just the ith atom can be calculated from

$$\rho_i(E) = -\frac{2}{\pi}\,\mathrm{Im}\lim_{z\downarrow E}\int_{\Omega_i} G(z,\boldsymbol{r},\boldsymbol{r})\ \mathrm{d}\boldsymbol{r} \qquad [10]$$

where $G(z,\boldsymbol{r},\boldsymbol{r}') = \langle \boldsymbol{r}|G|\boldsymbol{r}'\rangle$. Of course, the DOS for the supercell is $\rho(E) = \sum_{i=1}^{N}\rho_i(E)$. The Green's functions and scattering operators that are used henceforth are obtained from the above by letting z approach E from above the real axis.

When the position vectors $\boldsymbol{r}$ and $\boldsymbol{r}'$ are in the region Ω_n, the Green's function in the position representation can be shown to take the form

$$G(E,\boldsymbol{r}_n,\boldsymbol{r}'_n) = \sum_{LL'} Z^n_L(E,\boldsymbol{r}_n)\,\tau^{nn}_{LL'}(E)Z^{n^*}_{L'}(E,\boldsymbol{r}'_n) - \sum_L Z^n_L(E,\boldsymbol{r}_n)\,J^{n^*}_L(E,\boldsymbol{r}'_n) \qquad [11]$$

where

$$\tau^{nn}_{LL'} = t^n_{LL'} + \sum_{L_1,L_2,L_3,L_4} t^n_{LL_1}\sum_{i\neq n,j\neq n} g^{ni}_{L_1L_2}\tau^{ij}_{L_2L_3}g^{jn}_{L_3L_4}t^n_{L_4L'} = t^n_{LL'} + \sum_{L_1,L_2} t^n_{LL_1}\sum_{i\neq n} g^{ni}_{L_1L_2}\tau^{in}_{L_2L'} \qquad [12]$$

and $t^n_{LL'}(E)$ is the scattering matrix for the nth atom. The matrix $\boldsymbol{\tau}$ with elements $\tau^{nn}_{LL'}$ is the inverse of the

matrix $\boldsymbol{M}$ that has elements

$$M^{mn}_{LL'} = m^n_{LL'}\delta_{mn} - g^{mn}_{LL'}(1-\delta_{mn}) \qquad [13]$$

The function $Z^n_L(E,\boldsymbol{r}_n)$ is the solution of the equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v_n(\boldsymbol{r}_n) - E\right] Z^n_L(E,\boldsymbol{r}_n) = 0 \qquad [14]$$

that is regular at the origin and joins smoothly to

$$Z^n_L(E,\boldsymbol{r}_n) \rightarrow \sum_{L'} Y_{L'}(\boldsymbol{r}_n) j_{l'}(\alpha \boldsymbol{r}_n)\; m^n_{L'L}(E) - \mathrm{i}\kappa Y_L(\boldsymbol{r}_n)\; h_l(\alpha \boldsymbol{r}_n) \qquad [15]$$

when $\boldsymbol{r}$ is outside of Ω_n. As usual, $j_l(\alpha r_n)$and $h_l(\alpha r_n)$ are the spherical Bessel and Hankel functions, and $Y_L(\boldsymbol{r}) = Y_{l,m}(\vartheta,\varphi)$ is the spherical harmonic. The matrix $\boldsymbol{m}^n$ with elements $m^n_{LL'}$ is the inverse of the scattering matrix $\boldsymbol{t}^n$ whose elements are

$$t^n_{LL'}(E) = \int\int j_l(\alpha \boldsymbol{r})\; Y^*_L(\boldsymbol{r})\; \langle \boldsymbol{r}|t(E)|\; \boldsymbol{r}'\rangle\; Y_{L'}(\boldsymbol{r}') \times j_l(\alpha \boldsymbol{r}')\; \mathrm{d}\boldsymbol{r}\; \mathrm{d}\boldsymbol{r}' \qquad [16]$$

The function $J^n_L(E,\boldsymbol{r}_n)$ is a real solution of the differential eqn [14] that joins smoothly to $Y_L(\boldsymbol{r}_n)j_l(\kappa \boldsymbol{r}_n)$ when r is outside of Ω_n. The $g^{nn'}_{LL'}$ in eqn [12] can be pictured as propagators that take an electron from site n to m, and are given by

$$g^{nm}_{LL'} = -4\pi i\kappa i^{l-l'} \sum i^{-\lambda} C^{\Lambda}_{LL'} h_\lambda(\kappa R_{nm}) Y^*_\Lambda(\boldsymbol{R}_{nm}) \qquad [17]$$

where $\boldsymbol{R}_{nm} = \boldsymbol{R}_m - \boldsymbol{R}_n$ and the Gaunt factors $C^{\Lambda}_{LL'} = \int Y_\Lambda Y^*_L Y_{L'}\, \mathrm{d}\Omega$ are related to Clebsch–Gordon coefficients. This equation for the Green's function is correct for any kind of solid, but it is difficult to calculate except for pure metals, metals with impurities, and alloys using approximations.

Pure Metals

The expression for the Green's function, eqn [11], is greatly simplified for infinite periodic solids. The case of one atom per unit cell is treated because the extension to more than one atom is straightforward. The atoms are on the sites $\boldsymbol{R}_i$ of a Bravais lattice. Bloch's theorem follows from the invariance of the crystal under translations

$$\psi_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}_{ij}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{ij}}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \qquad [18]$$

There are only N Bloch vectors, $\boldsymbol{k}_i$, because of the periodic boundary conditions. It follows that the matrix $\boldsymbol{g}$ with elements $g^{ij}_{LL'}$ can be put into a block diagonal form by transforming it with the matrix $\boldsymbol{U}$ with elements

$$U^{ij}_{LL'} = \frac{1}{\sqrt{N}}\mathrm{e}^{-\mathrm{i}\boldsymbol{R}_i\cdot\boldsymbol{k}_j}\delta_{LL'} \qquad [19]$$

$$\left[\boldsymbol{U}^\dagger \boldsymbol{g} \boldsymbol{U}\right]^{ij}_{LL'} = g_{LL'}(E,\boldsymbol{k}_i)\delta_{ij} = \sum_{m=1}^{N} \mathrm{e}^{\mathrm{i}\boldsymbol{k}_i\cdot\boldsymbol{R}_{0m}} g^{0m}_{LL'}(E) \qquad [20]$$

Since the matrices $\boldsymbol{m}^i$ are the same for every cell, the transformed matrix $\boldsymbol{U}^\dagger\boldsymbol{M}\boldsymbol{U}$ is easily inverted. When the $\boldsymbol{U}$ matrix is used to transform back and then N is allowed to go to infinity, the scattering path matrix for the periodic crystal is

$$\boldsymbol{\tau}^{nn}_0(E) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}_i\cdot\boldsymbol{R}_{nn}}[\boldsymbol{m}(E) - \boldsymbol{g}(E,\boldsymbol{k}_i)]^{-1} = \frac{\Omega}{(2\pi)^3}\int \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{nn}}[\boldsymbol{m}(E) - \boldsymbol{g}(E,\boldsymbol{k})]^{-1}\, \mathrm{d}\boldsymbol{k} \qquad [21]$$

The angular momentum indices L and L' in this and succeeding equations are suppressed by the use of block matrices.

The functions $\boldsymbol{g}(E,\boldsymbol{k})$ are called structure constants in the Korringa–Kohn–Rostoker (KKR) band theory, and sophisticated methods have been developed for calculating them. The standard approximation in the application of this formula is that the elements of $\boldsymbol{m}(E)$ and $\boldsymbol{g}(E,\boldsymbol{k})$ that correspond to angular momenta such that $l > l_{\max}$ are ignored. Thus, the matrices that must be inverted in eqn [21] have a dimension $(l_{\max}+1)^2$.

Impurities in Metals

When an impurity atom replaces one of the atoms in a periodic metal, the scattering matrix $\boldsymbol{m}^i$ obviously changes on the impurity site. The scattering matrices on the adjoining sites also change because of the effect that the impurity has on its neighboring atoms. Even the Green's functions $\boldsymbol{g}^{ij}$ that connect the atoms are affected by the impurity change, because the size difference between the impurity atom and the host atoms causes displacements from the sites of the original periodic lattice. In metals, it is reasonable to assume that the influence of the impurity is screened beyond a certain number of nearest-neighbor shells to the point that it has no noticeable effect on the rest of the lattice. It follows that the $N \times N$ matrix $\boldsymbol{M}$ for the metal with an impurity can be written as $\boldsymbol{M} = \boldsymbol{M}_0 + \Delta\boldsymbol{M}$, where $\Delta\boldsymbol{M}$ is a matrix with elements

$$\Delta M^{mn}_{LL'} = (m^n_{LL'} - m^n_{0LL'})\;\delta_{mn} - (g^{mn}_{LL'} - g^{mn}_{0LL'})\;(1-\delta_{mn}) \qquad [22]$$

In this equation, $m^n_{0LL'}$ and $g^{mn}_{0LL'}$ are the matrix elements of the original periodic solid. All of the elements of $\Delta \boldsymbol{M}$ are zero except for those associated with the N_I sites of the atoms that are influenced by the impurity. Assuming that there are one or two nearest–neighbor shells, N_I is 13 or 19 for a face-centered-cubic (f.c.c) lattice and it is 9 or 15 for a body-centered cubic (b.c.c) lattice.

Consider the set of sites influenced by the impurity S_I. If $n \subset S_I$, the coefficients $\tau^{mn}_{LL'}$ that appear in the expression for the Green's function in eqn [11] are elements of the matrix

$$\boldsymbol{\tau}^I = \left[\boldsymbol{I} - \boldsymbol{\tau}^I_0 \Delta \boldsymbol{M}^I\right]^{-1} \boldsymbol{\tau}^I_0 \qquad [23]$$

The elements of $\boldsymbol{\tau}^I_0$ are the matrix elements of the scattering path operators for the host lattice $\boldsymbol{\tau}^{mn}_0$, where $m \subset S_I$ and $n \subset S_I$. The matrix $\Delta \boldsymbol{M}^I$ contains only the nonzero elements of $\Delta \boldsymbol{M}$. No angular momentum indices are included beyond $l_{\max}$, so the dimension of the matrices in eqn [26] is $N_I(l_{\max}+1)^2$.

The best way to calculate the DOS and charge density associated with the sites influenced by the impurity is

$$\rho^I_n(E) = -\frac{1}{\pi} \operatorname{Im} \sum_{LL'} \tau^{I,nn}_{LL'}(E) \int_{\Omega_n} F^n_{L'L}(E, \boldsymbol{r}, \boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \qquad [24]$$

and

$$\rho^I_n(\boldsymbol{r}) = -\frac{1}{\pi} \operatorname{Im} \sum_{LL'} \int_{-\infty}^{E_F} \tau^{I,nn}_{LL'}(E) F^n_{L'L}(E, \boldsymbol{r}, \boldsymbol{r})\, \mathrm{d}E \qquad [25]$$

where $F^n_{L'L}(E, \boldsymbol{r}'_n, \boldsymbol{r}_n) = Z^{n^*}_{L'}(E, \boldsymbol{r}'_n)\, Z^n_L(E, \boldsymbol{r}_n)$. In summary, a charge density for the host atoms and a crystal structure are assumed. The DFT is used to calculate the potential for the host atoms, $V_0(\boldsymbol{r})$, and hence the scattering matrix $\boldsymbol{t}_0(E)$. The coefficients $\tau^{nn}_{0,LL'}$ are obtained from eqn [21] and then the Green's function using eqn [11] and the charge density using eqn [2] are calculated. This process is iterated until the charges are self-consistent. At this point one has the coefficients for all the sites of the host lattice, and the elements of the matrix $\boldsymbol{\tau}^I_0$. It is now necessary to guess the charge densities for the impurity sites with $n \subset S_I$, calculate the potentials and scattering matrices, and construct the matrix $\Delta \boldsymbol{M}^I$. The coefficients and Green's functions for these sites are obtained from eqns [23] and [11]. The Fermi energy is not changed, and the new charge densities at the impurity sites are obtained from eqn [25]. The potentials must be iterated to self-consistency. These formulas have been very successful in predicting resistivities, hyperfine energies, and energies of formation, but only if the relaxation of the atoms in the nearest-neighbor shells away from the periodic lattice sites is allowed to occur. The calculations are sensitive to approximations to the shape of the self-consistent potentials, and, where relevant, spin polarization must be included.

An interesting formula for the DOS can be obtained using a theorem from the scattering theory derived by the Russian academician M G Krein. In operator language, Krein's theorem states that

$$\operatorname{Tr}[G(z) - G_0(z)] = \frac{\mathrm{i}}{2\pi} \lim_{\varepsilon \downarrow 0} \int_{-\infty}^{\infty} \frac{\ln \operatorname{Det}(I - \mathrm{i}2\sqrt{E' + \mathrm{i}\varepsilon}\, T(E' + \mathrm{i}\varepsilon))}{(E' - z)^2}\, \mathrm{d}E' \qquad [26]$$

Using the formula for the DOS in eqn [10], this equation can, with some effort, be put into the form

$$\rho(E) - \rho_0(E) = -\frac{2}{\pi} \operatorname{Im} \frac{\mathrm{d} \ln \operatorname{Det} \boldsymbol{M}(E)}{\mathrm{d}E} \qquad [27]$$

where $\rho(E)$ is the DOS for the system of scatterers, $\rho_0(E)$ is the DOS for free electrons, and $\boldsymbol{M}(E)$ is the matrix defined in eqn [13]. This is called Lloyd's formula because Lloyd derived it without the help of Krein's theorem. Krein's theorem can immediately be used for impurities by defining H_0 to be the Hamiltonian for the host lattice rather than just the kinetic energy, and a version of Lloyd's formula that gives the change in the total DOS caused by the impurity can be obtained.

The simplest theory for an impurity is to picture it as a single spherically symmetric scatterer embedded in a free-electron gas. Then Krein's theorem leads to

$$\rho(E) - \rho_0(E) = -\frac{\mathrm{i}}{\pi} \operatorname{Im} \sum_l (2l+1) \times \frac{\mathrm{d} \ln [1 - -2\alpha t_l(E)]}{\mathrm{d}E} \qquad [28]$$

For such a scatterer, the t-matrix defined in eqn [16] becomes $t_{LL'} = t_l \delta_{ll'} \delta_{mm'}$ with

$$t_l = -\frac{1}{\alpha} \mathrm{e}^{\mathrm{i}\eta_l} \sin \eta_l \qquad [29]$$

where the η_l are scattering phase shifts. From this, the Friedel sum is obtained.

$$\rho(E) - \rho_0(E) = \frac{2}{\pi} \sum_l (2l+1) \frac{\mathrm{d}\eta_l}{\mathrm{d}E} \qquad [30]$$

The Friedel formula appears in many text books on condensed matter physics, but the assumptions that

go into its derivation are very limiting. Leaving aside the fact that the electronic spectrum of the host resembles a free electron gas for very few solids, the assumption that the change in the spectrum is confined to one site is internally inconsistent. It is typical that the impurity atom loses or gains charge from the host. The charge is screened to make the system charge neutral, but the screening charge is typically distributed over one or two nearest-neighbor shells.

Coherent Potential Approximation

It should be clear that the impurity equations described in the preceding section can be generalized to deal with several interacting impurities in a metal. It is only necessary to increase the number of sites N_i influenced by the impurity atoms. However, when the concentration of impurities reaches the point at which most of the impurity atoms are within range of each other, the impurity equations are no longer useful. It is then necessary to treat the material as an alloy, and to develop a different class of theories. The problem of finding the solutions of Schrodinger's equation for a potential function that is not periodic is difficult because Bloch's theorem, eqn [18], no longer holds. Progress has been made on substitutional solid-solution alloys that are made up of two or more species of atoms distributed randomly on the sites $\boldsymbol{R}_i$ of a Bravais lattice. In the following, the concentration is on binary alloys, although the generalization to ternary, quaternary, etc. is straightforward.

The ensemble of all alloys that can be constructed out of $N_A = c_A N$ A atoms and $N_B = c_B N$ B atoms contains $M_c = N!/N_A!N_B!$ atomic arrangements. Experimentally, it is observed that all alloys with the concentrations c_A and c_B have the same electronic structure. It follows that, as $N \to \infty$, all of the arrangements have the same electronic structure with probability 1, and it can be found by averaging over the ensemble. The ensemble average of a one-electron wave function has no meaning, but, because of the linear relation between Green's function and the electronic properties, the ensemble average of Green's function is useful.

The ensemble average of the function in eqn [11] can be obtained if the assumption is made that the resulting alloy model is isomorphous. That is, it is assumed that the Green's function takes one of two forms, $G_A(E, \boldsymbol{r}_n, \boldsymbol{r}'_n)$ if there is an A atom on site n and $G_B(E, \boldsymbol{r}_n, \boldsymbol{r}'_n)$ if there is a B atom on that site. With this assumption, the ensemble averaged Green's function is

$$\langle G_n(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) \rangle = c_A G_A(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) + c_B G_B(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) \quad [31]$$

where

$$G_A(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) = \sum_{LL'} Z_L^A(E, \boldsymbol{r}_n) \langle \tau_{LL'}^{nn} \rangle_A Z_{L'}^{A^*}(E, \boldsymbol{r}'_n) - \sum_L Z_L^A(E, \boldsymbol{r}_n) J_L^{A^*}(E, \boldsymbol{r}'_n) \quad [32]$$

and

$$G_B(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) = \sum_{LL'} Z_L^B(E, \boldsymbol{r}_n) \langle \tau_{LL'}^{nn} \rangle_B Z_{L'}^{B^*}(E, \boldsymbol{r}'_n) - \sum_L Z_L^B(E, \boldsymbol{r}_n) J_L^{B^*}(E, \boldsymbol{r}'_n) \quad [33]$$

The charge density for an A atom, $\rho_A(\boldsymbol{r}_n)$, is obtained by inserting $G_A(E,\boldsymbol{r}_n,\boldsymbol{r}_n)$ into eqn [2], and similarly for the B atom. From these charge densities, the potentials $v_A(\boldsymbol{r}_n)$ and $v_B(\boldsymbol{r}_n)$ are calculated. The functions $Z_L^A(E, \boldsymbol{r}_n)$ and $J_L^A(E, \boldsymbol{r}_n)$ are obtained using $v_A(\boldsymbol{r}_n)$ in eqn [13], and the functions $Z_L^B(E, \boldsymbol{r}_n)$ and $J_L^B(E, \boldsymbol{r}_n)$ are obtained using $v_B(\boldsymbol{r}_n)$.

The averages $\langle \tau_{LL'}^{nn} \rangle_A$ and $\langle \tau_{LL'}^{nn} \rangle_B$ are the ensemble averages of the scattering path operator subject to the condition that there is an A or B atom on site n. These averages are simplified considerably by making a mean field approximation. It is assumed that, on an average, the scattering of the electrons from all the sites other than site n is described by replacing the actual potentials on those sites with a potential $v_c(\boldsymbol{r})$, called the coherent potential. The scattering from $v_c(\boldsymbol{r})$ is described by a t-matrix, $\boldsymbol{t}_c(E)$, which has an inverse $\boldsymbol{m}_c(E)$. It follows that, within the coherent potential approximation (CPA), the ensemble averages $\langle \tau_{LL'}^{nn} \rangle_A$ and $\langle \tau_{LL'}^{nn} \rangle_B$ are the elements of the matrices obtained from eqn [23]

$$\langle \boldsymbol{\tau}^{nn} \rangle_A = [\boldsymbol{I} - \boldsymbol{\tau}_c^{nn}(\boldsymbol{m}_A - \boldsymbol{m}_c)]^{-1} \boldsymbol{\tau}_c^{nn} \quad [34]$$

and

$$\langle \boldsymbol{\tau}^{nn} \rangle_B = [\boldsymbol{I} - \boldsymbol{\tau}_c^{nn}(\boldsymbol{m}_B - \boldsymbol{m}_c)]^{-1} \boldsymbol{\tau}_c^{nn} \quad [35]$$

The matrices $\boldsymbol{m}_A$ and $\boldsymbol{m}_B$ are the inverses of the t-matrices calculated from the potentials $v_A(\boldsymbol{r})$ and $v_B(\boldsymbol{r})$. From eqn [24], it can be seen that the matrix $\boldsymbol{\tau}_c^{nn}$ is

$$\boldsymbol{\tau}_c^{nn}(E) = \frac{\Omega}{(2\pi)^3} \int [\boldsymbol{m}_c(E) - \boldsymbol{g}(E, \boldsymbol{k})]^{-1} \, \mathrm{d}\boldsymbol{k} \quad [36]$$

The number of impurity sites N_I is 1, so the matrices in these equations have dimension $(l_{\max} + 1)^2$. As would be expected, none of these equations depend on the choice of a site n.

In the language of the multiple-scattering theory used here, the condition originally put forward by

P Soven that defines the mean field t-matrix $\boldsymbol{t}_c(E)$ is

$$c_A \langle \boldsymbol{\tau}^{nn} \rangle_A + c_B \langle \boldsymbol{\tau}^{nn} \rangle_B = \boldsymbol{\tau}_c^{nn} \qquad [37]$$

In words, this is the condition that, on an average, there is no scattering from A and B atoms embedded in the coherent potential host matrix. Constraints on space do not allow descriptions of all the arguments put forward to justify this CPA condition. Perhaps the most convincing arguments come from the formulation of the problem of calculating the Green's function of an alloy in terms of the infinite order perturbation theory. Diagrams are then used to analyze the contributions to the Green's functions in a manner analogous to the field theory, quantum electrodynamics, or the many-body theory. It has been argued that the CPA contains the maximum number of classes of diagrams that can be included, and still retain a Green's function that is properly analytic. The criterion that is used to specify the diagrams included in the CPA is called the single-site approximation. It has been shown by R Mills and P Ratanavararaksa and also A Mookerjee that it is possible to go beyond the single-site approximation and retain an analytic Green's function, but it is difficult.

The procedure for calculating the electronic states within the CPA is first to guess the starting values for the charge densities $\rho_A(\boldsymbol{r})$ and $\rho_B(\boldsymbol{r})$, and from them calculate $v_A(\boldsymbol{r})$ and $v_B(\boldsymbol{r})$. The t-matrices $\boldsymbol{t}_A(E)$ and $\boldsymbol{t}_B(E)$ are calculated, and a first guess to $\boldsymbol{t}_c(E)$ is just their average $\boldsymbol{t}_c(E) \approx c_A \boldsymbol{t}_A(E) + c_B \boldsymbol{t}_B(E)$. Equations [32]–[35] are then iterated to self-consistency in the charge densities and also in the coherent potential t-matrix $\boldsymbol{t}_c(E)$.

Calculations with this version of the CPA have been carried out on many alloy systems. The predicted electronic DOS agree well with the ones measured in ultraviolet photoemission spectroscopy experiments. The zero-temperature resistivities of alloys have been accurately predicted. Even the propensity of a random alloy to transform into a superlattice structure or into an incommensurately ordered structure has been successfully predicted.

A problem that is endemic to this version of the CPA is that the Coulomb effects are not treated correctly. Integrating the charge densities $\rho_A(\boldsymbol{r})$ and $\rho_B(\boldsymbol{r})$ and subtracting the nuclear charges Z_A and Z_B leads to the net charges on the A and B sites, q_A and q_B. It would seem that there should be a Madelung contribution to the potentials $v_A(\boldsymbol{r})$ and $v_B(\boldsymbol{r})$ as in ordered alloys, but the only Madelung potentials that are commensurate with the assumption of an isomorphous alloy are zero. The isomorphous model of an alloy with the charge q_A on every A atom and q_B on every B atom is not realistic. Using order-N methods that are described in the next section, it is possible to calculate the electronic structure for supercells containing thousands of atoms with no approximation other than the DFT. **Figure 1** shows the charges on 512 copper atoms and 512 zinc atoms distributed randomly on the sites of a supercell with the structure of a b.c.c. lattice. The distribution is broad, and upon reflection, is what would be expected on the basis of physical intuition. It should be compared with the prediction of the isomorphous model that all the Cu atoms would take on exactly 0.09978 electron charges, and Zn atoms would lose the same amount.

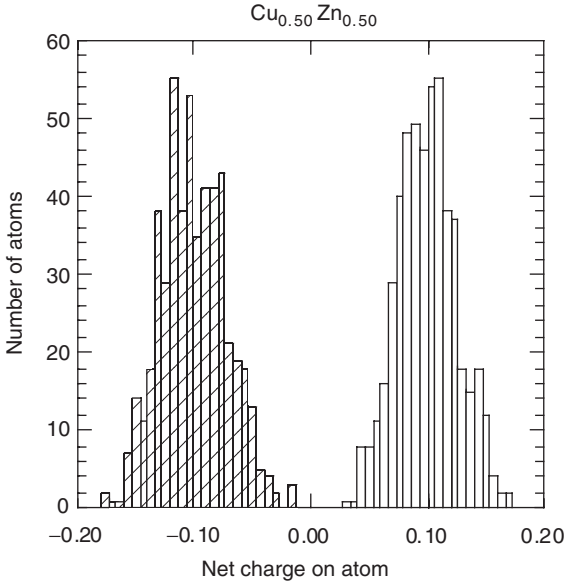


Figure 1 The distribution of charges on the sites of a 50% copper–zinc alloy modeled with a supercell containing 1024 atoms. The charges on the copper sites are shown with the white histograms, and the charges on the zinc sites are cross-hatched.

A version of the CPA can be constructed that leads to a polymorphous model of the alloy in which each atom has a different charge, but it requires a different kind of averaging. Recall that periodically reproduced supercells are considered. In the binary alloy model, N_A A atoms and N_B B atoms are distributed over the $N = N_A + N_B$ sites. The Green's function defined in eqn [11] for the alloy is different for each of the $M_c = N!/N_A!N_B!$ possible atomic arrangements, and can be labeled $G_n^i(E, \boldsymbol{r}_n, \boldsymbol{r}'_n)$ to emphasize that. Then, the ensemble average in eqn [27] can more properly be written as

$$\langle G_n(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) \rangle = \frac{1}{M_c} \sum_{i=1}^{M_c} G_n^i(E, \boldsymbol{r}_n, \boldsymbol{r}'_n) \qquad [38]$$

It is also possible to consider one supercell with the atomic arrangement i, and average over the sites

$$\{G^i(E,\boldsymbol{r},\boldsymbol{r}')\} = \frac{1}{N}\sum_{n=1}^{N} G_n^i(E,\boldsymbol{r},\boldsymbol{r}') \qquad [39]$$

It is seen that, as $N\to\infty$, $\langle G_n(E,\boldsymbol{r}_n,\boldsymbol{r}'_n)\rangle$ is independent of the site index n. It is equally clear that in the same limit, $\{G^i(E,\boldsymbol{r},\boldsymbol{r}')\}$ is independent of the arrangement index i. The equivalence of the ensemble-average in eqn [38] with the site average in eqn [39] is similar to the Ergodic theorem first proposed by Boltzmann, and can be proved by similar arguments. If the condition is imposed that the alloy must be isomorphous, the site average will lead to the same averaged potential as the ensemble average result shown in eqn [27]. However, the site average can lead to a polymorphous model.

Applying the CPA philosophy to the site-averaged Green's function leads to

$$\{G(E,\boldsymbol{r},\boldsymbol{r}')\} = \frac{1}{N}\sum_{n=1}^{N}\{G_n(E,\boldsymbol{r},\boldsymbol{r}')\} \qquad [40]$$

where

$$\{G_n(E,\boldsymbol{r},\boldsymbol{r}')\} = \sum_{LL'} Z_L^n(E,\boldsymbol{r})\{\tau_{LL'}^{nn}(E)\}Z_{L'}^n(E,\boldsymbol{r}') - \sum_L Z_L^n(E,\boldsymbol{r})\, J_L^n(E,\boldsymbol{r}') \qquad [41]$$

and

$$\{\boldsymbol{\tau}^{nn}\} = [\mathbf{I} - \boldsymbol{\tau}_{\mathrm{c}}^{nn}(\mathbf{m}_n - \mathbf{m}_{\mathrm{c}})]^{-1}\boldsymbol{\tau}_{\mathrm{c}}^{nn} \qquad [42]$$

Inserting the Green's function for an index n into eqn [2] produces a charge density, $\rho_n(\boldsymbol{r})$, and hence a one-electron potential, $v_n(\boldsymbol{r})$, for every site. From the potential, one calculates the t-matrix $\boldsymbol{t}_n$ and its inverse $\boldsymbol{m}_n$ which appears in eqn [38]. As in the isomorphous CPA, the scattering path for the coherent potential host lattice $\boldsymbol{\tau}_{\mathrm{c}}^{nn}$ is calculated from eqn [36]. The condition that defines the coherent potential is

$$\frac{1}{N}\sum_{n=1}^{N}\{\tau^{nn}\} = \tau_{\mathrm{c}}^{nn} \qquad [43]$$

instead of eqn [37].

Since the focus is on a specific supercell, the position $\boldsymbol{R}_n$ of each atom is known and the Madelung potential for the nth site is

$$v_n^{\mathrm{M}} = \sum_{m\neq n=1}^{N} M(|\boldsymbol{R}_n - \boldsymbol{R}_m|)q_n \qquad [44]$$

where

$$q_n = \int_{\Omega_n} \rho_n(\boldsymbol{r})\ \mathrm{d}\boldsymbol{r} - Z_n \qquad [45]$$

and $M(|\boldsymbol{R}_n - \boldsymbol{R}_m|)$ is the Madelung matrix. The Madelung matrix is defined so that the sum in eqn [44] gives the contribution to the Coulomb potential from all of the periodically reproduced supercells. A sophisticated method for calculating it was developed by P P Ewald, and it reduces to $1/|\boldsymbol{R}_n - \boldsymbol{R}_m|$ when $N\to\infty$. The potentials $v_n(\boldsymbol{r})$ contain the Madelung shift v_n^{M}.

The polymorphous version of the CPA has been applied to several alloy systems. For alloys that have a large amount of charge transfer, such as copper–gold or copper–palladium, it gives better results than the ones obtained with the isomorphous CPA, but the improvements are quite small for alloys such as silver–palladium that have a small charge transfer. The polymorphous CPA has the disadvantage that it requires the use of a specific supercell. This is a small price to pay for the improvement in the treatment of the Coulomb energy. It is quite easy to apply because the charge densities $\rho_A(\boldsymbol{r})$ and $\rho_B(\boldsymbol{r})$ and the effective t-matrix $\boldsymbol{t}_{\mathrm{c}}$ from the isomorphous CPA are used as the starting values in the self-consistent iterations.

Although the coherent potential $v_{\mathrm{c}}(\boldsymbol{r})$ is introduced in the discussion, it is never calculated or used. The only potentials used in eqn [14] to calculate wave functions are real potentials for real atoms, $v_n(\boldsymbol{r})$. From the analytic properties of $\boldsymbol{t}_{\mathrm{c}}$ it follows that $v_{\mathrm{c}}(\boldsymbol{r})$, if it existed, would be complex and energy dependent.

Order-*N* Methods

A brute force method for calculating the electronic states of a disordered solid is to simply model a portion of the solid in a supercell. The nuclei do not have to be placed on the sites of a Bravais lattice, and any combination of atomic species can be used. This makes it possible to approach a much wider range of problems than the CPA can deal with. Supercells containing a large number of atoms can be treated with the band-theory methods that have been developed to treat solids with N atoms per unit cell. All of these methods require an operation that is the equivalent of inverting or diagonalizing an $N\times N$ matrix, and these operations require an amount of computer time that scales as N^3. Even though computer speeds increase exponentially, this scaling always places a limit on the size of the supercell that can be used. For this reason, considerable effort has been devoted to the development of order-N methods

for electronic structure calculations for which the computer times scale as N rather than N^3.

An order-N method called the locally self-consistent multiple-scattering (LSMS) method was developed by G M Stocks and his collaborators. Each atom in the supercell is considered to be the center of a cluster of a local interaction zone (LIZ) that contains N_{LIZ} atoms. The matrix $\boldsymbol{M}^n$ with elements defined in eqn [13] is truncated so that only the scatterers in the LIZ corresponding to site n are included. This is typically the nth site and its first seven or eight nearest-neighbor shells. Truncating the angular momentum expansions at $l_{\max}$, the order of $\boldsymbol{M}^n$ is $(l_{\max}+1)^2 N_{\mathrm{LIZ}}$ and $N_{\mathrm{LIZ}} < 100$. After the inversion of $\boldsymbol{M}^n$, only the matrix elements $\tau^{nn}_{LL'}$ corresponding to the central atom are used. With these matrix elements, the Green's function $G_n(E, \boldsymbol{r}_n, \boldsymbol{r}'_n)$ is calculated from eqn [11], and hence the charge density $\rho_n(\boldsymbol{r}_n)$ and the net charge on the site q_n are obtained. The Madelung shift on the nth site is calculated using eqn [44], and the next iteration for the potential $v_n(\boldsymbol{r})$ is acquired by adding that to the DFT potential calculated from the charge density. This procedure scales with N, although the prefactor is $(l_{\max}+1)^6 N^3_{\mathrm{LIZ}}$.

The LSMS puts nothing on the sites outside the LIZ. The locally self-consistent Green's function (LSGF) method proposed by I A Abrikosov and his collaborators puts coherent potentials on those sites. The LSGF makes use of the equations for an impurity cluster in a metal, eqns [22]–[25], with the coherent potential scattering matrix $\boldsymbol{m}_{\mathrm{c}}$ replacing the host matrix $\boldsymbol{m}_0$. The matrix $\boldsymbol{\tau}^I$ in eqn [23] contains the average scattering path operators for all the scatterers in the LIZ. Only the matrix for the central site, $\{\boldsymbol{\tau}^{nn}\}$, is used in the equation that defines $\boldsymbol{m}_{\mathrm{c}}$, which is eqn [43]. The charge density $\rho_n(\boldsymbol{r}_n)$ is obtained by inserting the Green's function from eqn [41] into eqn [2]. This charge density and the Madelung shift from eqn [44] is used to calculate the new potential $v_n(\boldsymbol{r})$ in the self-consistent process.

The LSMS has the advantage that the atoms can be placed anywhere in the supercell, while the atoms should be on the sites of a Bravais lattice in the LSGF. The advantage of the LSGF is that the LIZ can be made much smaller. In fact, for $N_{\mathrm{LIZ}} = 1$, which means that it contains only the central atom, the LSGF is identical to the site-averaged CPA described above. That level of approximation is adequate for many calculations on alloys.

The LSMS and LSGF are order-N because the propagation of electrons from one atom to another is treated properly only for atoms within the LIZ. It should be noted that the range of the Coulomb interactions is not assumed to be short. The Ewald sum in eqn [44] includes contributions from all the atoms in the supercell, and indeed from the atoms in the periodically reproduced supercells that fill all space. The calculation of the Madelung shifts takes very little time in comparison with the rest of the calculation, and experience has shown that it is not a good idea to truncate it.

Extensions and Approximations

All of the equations in this article have been written in the nonrelativistic form because they are easier to visualize. They can easily be transformed to the Dirac form using the relativistic density-functional theory (RDFT) of A K Rajagopal, *et al*. The spin-polarized version of these equations has been used very successfully to explain the magnetism of metals and alloys. The calculations can be made easier by employing a shape approximation to the potentials, such as the muffin-tin (MT) or atomic sphere approximation (ASA). Instead of using the full multiple-scattering version of the equations depicted here, many authors choose to approximate them with a variational version called the linear combination of muffin-tin orbitals (LMTO). These approximations are discussed for wave functions elsewhere in this encyclopedia.

The amount of computer time required for calculations with Green's functions can be significantly reduced with the help of sophisticated techniques derived from their properties. Using the analyticity of the Green's functions, it is frequently possible to carry out integrals over the energy, such as eqn [2], by making the energy complex and integrating over a contour rather than along the real axis. It is also possible to use a finite-temperature version of the DFT that improves the convergence of some of the integrals. Finally, the range of the propagators $\boldsymbol{g}^{mn}$ that describe the motion of the electron from one site to another can be made short by shifting the zero of the potential between scatterers.

Of course, many properties of alloys are calculated using statistical methods such as Monte Carlo, molecular dynamics, and the cluster-variational method. The electronic structure appears in these calculations only in the calculation of the interatomic potentials. Different approaches to the calculation of these potentials are the embedded atom method, the generalized perturbation method, and the mixed-space cluster expansion method.

See also: Density-Functional Theory; Disordered Solids and Glasses, Electronic Structure of; Electronic Structure Calculations: Scattering Methods; Magnetoresistance Techniques Applied to Fermi Surfaces.

PACS: **71.15. – m; 71.15.Ap; 71.15.Mb; 71.20.Be; 73.20.At; 73.21. – b**

Further Reading

Elliott RJ, Krumhansl JA, and Leath PL (1974) The theory and properties of randomly disordered crystals and related physical systems. *Reviews of Modern Physics* 46: 465.

Faulkner JS (1982) The modern theory of alloys. In: Christian JW, Haasen P, and Massalski TB (eds.) *Progress in Materials Science*, vol. 27. Oxford, England: Pergamon Press.

Gonis A (1992) Disordered systems. *Studies in Mathematical Physics*, vol. 4. Amsterdam: North-Holland.

Proceedings of the 2nd Annual Meeting of the European Union (EU) funded Training and Mobility of Researchers (TMR) Network (1998) *Philosophical Magazine B* 78(5): 6.

Metals and Metallic Alloys, Optical Properties of

F Forstmann, Freie Universität Berlin, Berlin, Germany

Reflection and Transmission

Metals are characterized by large electrical conductivity and shiny surfaces, that is, intense reflectivity. Because both properties are responses to electromagnetic fields, there is a close connection.

For a transparent discussion, Maxwell's equations which govern optics are presented, with polarization, magnetization, and current shown explicitly:

$$\nabla \times \boldsymbol{B} = \mu_0 \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} + \mu_0 \left(\boldsymbol{j} + \frac{\partial \boldsymbol{P}}{\partial t} + \nabla \times \boldsymbol{M} \right) \quad [1]$$

$$\nabla \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0}(\rho - \nabla \cdot \boldsymbol{P}) \quad [2]$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial B}{\partial t} \quad [3]$$

$$\nabla \cdot \boldsymbol{B} = 0 \quad [4]$$

Here, it is clearly seen that $-\nabla \cdot \boldsymbol{P}$ acts as a source of the electric field like any charge and $\partial \boldsymbol{P}/\partial t$ (and $\nabla \times \boldsymbol{M}$) creates $\boldsymbol{B}$ like any current. This is easily understood from the definition of the polarization $\boldsymbol{P}$ as the amount of charge penetrating a plane normal to $\boldsymbol{P}$ when the polarization is created. It is conventional to distinguish between true charges ρ and apparent charges $-\nabla \cdot \boldsymbol{P}$, true currents $\boldsymbol{j}$ and apparent currents $(\partial \boldsymbol{P}/\partial t + \nabla \times \boldsymbol{M})$. $\boldsymbol{P}$ is then considered due to the shift of bound charges (such as in an oscillator-model). But it is obvious that at very high frequencies with short shifts of all charges, the distinction between free and bound charges will not be possible any more. It is impracticable in metal optics. It is only a theoretical concept. One effectively deals with a total response to the electromagnetic fields and therefore with an effective response function: a complex dielectric function ε, a complex conductivity σ, or a complex polarizability $\chi = \varepsilon - 1$.

Next, linear isotropic responses are assumed:

$$\boldsymbol{j} = \sigma \boldsymbol{E}, \quad \boldsymbol{P} = \chi_{\mathrm{p}} \varepsilon_0 \boldsymbol{E}, \quad \boldsymbol{M} = \chi_{\mathrm{m}} \boldsymbol{H} \quad [5]$$

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} = \varepsilon_0 (1 + \chi_{\mathrm{p}}) \boldsymbol{E} = \varepsilon_0 \varepsilon_{\mathrm{p}} \boldsymbol{E} \quad [6]$$

$$\begin{aligned} \boldsymbol{B} &= \mu_0 (\boldsymbol{H} + \boldsymbol{M}) = \mu_0 (1 + \chi_{\mathrm{m}}) \boldsymbol{H} \\ &= \mu_0 \mu \boldsymbol{H}, \quad \varepsilon_0 \cdot \mu_0 = \frac{1}{c^2} \end{aligned} \quad [7]$$

The relation between polarizability and conductivity can be derived by reordering of eqns [1] and [2] to the more conventional form

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{j} = \frac{\partial}{\partial t} \boldsymbol{D}_{\mathrm{eff}} \quad [8]$$

$$\nabla \cdot \boldsymbol{D} = \rho \quad [9]$$

$\boldsymbol{D}_{\mathrm{eff}}$ is easily defined for each time Fourier component. A periodicity $\mathrm{e}^{-\mathrm{i}\omega t}$ for all optical fields is introduced (the minus sign is chosen arbitrarily, but it determines many signs from now on) and one gets from [8] and [3]

$$\begin{aligned} \nabla \times \frac{1}{\mu_0 \mu} \boldsymbol{B} &= -\mathrm{i}\omega \left(\varepsilon_0 \varepsilon_{\mathrm{p}} - \frac{\sigma}{\mathrm{i}\omega} \right) \boldsymbol{E} \\ &= -\mathrm{i}\omega \boldsymbol{D}_{\mathrm{eff}} \end{aligned} \quad [10]$$

$$\nabla \times \boldsymbol{E} = +\mathrm{i}\omega \boldsymbol{B} \quad [11]$$

$$\begin{aligned} \nabla \times (\nabla \times \boldsymbol{E}) &= \nabla (\nabla \cdot \boldsymbol{E}) - \Delta \boldsymbol{E} \\ &= \frac{\omega^2}{c^2} \mu \left(\varepsilon_{\mathrm{p}} - \frac{\sigma}{\mathrm{i}\omega \varepsilon_0} \right) \boldsymbol{E} \end{aligned} \quad [12]$$

As explained at the beginning, the polarizability and the conductivity are found to be on the same footing in the wave equation [12] and therefore, the effective

complex dielectric function is defined as

$$\varepsilon = 1 + \chi_p - \frac{\sigma}{i\omega\varepsilon_0}, \quad \boldsymbol{D}_{eff} = \varepsilon_0\varepsilon\boldsymbol{E} \qquad [13]$$

In standard metal optics, it is argued that in conducting metals any charge ρ is quickly dissipated or screened and eqn [2] yields effectively $\nabla \cdot \boldsymbol{E} = 0$. An estimate of the decay time of a charge can be got by putting eqn [9] into the continuity equation

$$\dot{\rho} = -\nabla \cdot \boldsymbol{j} = -\sigma\nabla \cdot \boldsymbol{E} = -\frac{\sigma}{\varepsilon_0\varepsilon_p}\rho \qquad [14]$$

$$\rho = \rho_0 \cdot e^{-t/\tau}, \quad \text{with} \quad \tau = \frac{\varepsilon_p\varepsilon_0}{\sigma} \qquad [15]$$

For silver with $\sigma = (0.016 \times 10^{-4}\ \Omega\ \text{cm})^{-1}$ and, $\varepsilon_p \approx 1$, the decay time $\tau \approx 10^{-19}$ s and, therefore, charges decay faster than most optical fields vary. This estimate with the static conductivity is too optimistic. It is seen later that around $\omega \approx 10^{16}\ s^{-1}$ ($\hbar\omega > 3.5$ eV) charge density waves, plasma waves with $\nabla \cdot \boldsymbol{E} \neq 0$, play a role in optics. But in the regime of visible optics with $\nabla \cdot \boldsymbol{E} = 0$ and eqn [13], the level of standard optics was attained with transverse waves and the complex index of refraction

$$\tilde{n} = \sqrt{\varepsilon\mu} \qquad [16]$$

In most metals, $\mu \approx 1$ is a valid approximation. A treatment of the reflection from the surface of a metal with dielectric function ε_M facing a transparent medium (ε_1) yields the standard Fresnel formulas, which for later use are derived in a handy notation for wave vectors in the xz-plane, surface normal $\|z$ into the metal:

$$\text{in medium 1}: \quad \boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{E}_0 e^{i(k_x x + qz)} + \boldsymbol{R} e^{i(k_x x - qz)} \qquad [17]$$

$$\text{in the metal}: \quad \boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{T} e^{i(k_x x + lz)} \qquad [18]$$

$$q = \sqrt{\frac{\omega^2}{c^2}\varepsilon_1\mu_1 - k_x^2}, \quad l = \sqrt{\frac{\omega^2}{c^2}\varepsilon_M\mu_M - k_x^2} \qquad [19]$$

p- and s-polarization are distinguished and so are $\boldsymbol{E}$ parallel or normal to the plane of incidence, here the xz-plane. Because $\boldsymbol{R}$ and $\boldsymbol{T}$ are orthogonal to their respective wave vectors, only the two z-components have to be determined for p-polarization (the two amplitudes for s-polarization). Therefore, the two standard boundary conditions are used:

$$\text{continuity of } E_{tangential} \text{ or } B_{normal} \qquad [20]$$

$$\text{continuity of } D_{eff,normal} \text{ or } H_{tangential} \qquad [21]$$

Here, an obvious problem is usually brushed under the carpet – what the continuity of the normal component of $\boldsymbol{D}$ from $\nabla \cdot \boldsymbol{D} = 0$ (eqn [19] with $\rho = 0$) or of $\boldsymbol{D}_{eff}$ from $\nabla \cdot \boldsymbol{D}_{eff} = 0$ as eqn [8] or [10] tells us. Because this problem is difficult, one takes refuge in the continuity of H_{tg} which actually makes $D_{eff,normal}$ continuous as a consequence of [8] or [10]. Continuity of H_{tg} requires (from [8]) all fields to be finite, here $\boldsymbol{D}$, and no singular surface currents. But the discontinuity of E_{normal}, which goes with the continuity of $D_{eff,normal} = \varepsilon_0\varepsilon E_{normal}$ and a discontinuous ε, requires singular surface charges, from eqn [2]. Arguing with bound charges $-\nabla \cdot \boldsymbol{P}$ does not help because $\chi_p = 0$, $\boldsymbol{P} = 0$ is often a good description of the metal. Also the normal component of $\boldsymbol{j} = \sigma\boldsymbol{E}$ is discontinuous.

For frequencies in the visible approximation, [21] works. Around the eigenfrequency of the electrons, the plasma frequency ω_p, the charges, and the currents get so large that the problem behind boundary condition [21] must be solved. This is done in the section "Plasma waves in metal optics."

Boundary conditions [20] and [21] yield two linear equations for the reflection and transmission amplitudes. They are related to the reflection and transmission coefficients r and t:

$$p\text{-polarization}: \quad \frac{R_z}{E_{0z}} = \frac{\varepsilon_M q - \varepsilon_1 l}{\varepsilon_M q + \varepsilon_1 l}, \quad \frac{T_z}{E_{0z}} = \frac{2q\varepsilon_1}{\varepsilon_M q + \varepsilon_1 l} \qquad [22]$$

$$r = \left|\frac{R_z}{E_{0z}}\right|^2, \qquad t = \frac{l\varepsilon_M}{q\varepsilon_1}\left|\frac{T_z}{E_{0z}}\right|^2 \qquad [23]$$

$$s\text{-polarization}: \quad \frac{R}{E_0} = \frac{q\mu_M - l\mu_1}{q\mu_M + l\mu_1}, \quad \frac{T}{E_0} = \frac{2q\mu_M}{q\mu_M + l\mu_1} \qquad [24]$$

$$r = \left|\frac{R}{E_0}\right|^2, \qquad t = \frac{l\mu_1}{q\mu_M}\left|\frac{T}{E_0}\right|^2 \qquad [25]$$

In the infrared, the term $\sigma/(\omega\varepsilon_0)$ in eqn [13] is much larger than 1. If $\mu_M \approx \mu_1$ and $\varepsilon_1 = 1$, the reflection coefficient for normal incidence ($k_x \to 0$) can be expanded:

$$r = \left|\frac{\tilde{n}_M - 1}{\tilde{n}_M + 1}\right|^2 \approx 1 - 2\sqrt{2\frac{\omega\varepsilon_0}{\sigma}} \qquad [26]$$

This is called the Hagen–Rubens relation for the decrease of the metallic reflectivity due to the finite conductivity. It is a good approximation (with the

static conductivity) for most metals in the infrared for $\lambda > 10\,\mu m$.

For further discussion of the consequences of eqns [22]–[25], a model for the metallic response is considered, the Drude model, which approximates well the typical properties of metals. It considers only the conducting electrons. It could easily be supplemented by a Debye oscillator model for underlying bound electrons, the polarizability χ_p which could then simply be added as in eqn [13]. The Drude model is the equation of motion of the collection of electrons driven by the electric field and suffering a friction:

$$m\ddot{\boldsymbol{r}} = -e_0\boldsymbol{E} - m\gamma\dot{\boldsymbol{r}}, \quad \boldsymbol{j} = -e_0 n_0 \dot{\boldsymbol{r}} \qquad [27]$$

$$\frac{\partial}{\partial t}\boldsymbol{j} = \frac{n_0 e_0^2}{m}\boldsymbol{E} - \gamma\boldsymbol{j} = \omega_p^2 \varepsilon_0 \boldsymbol{E} - \gamma\boldsymbol{j} \qquad [28]$$

The plasma frequency is introduced here via the mean conduction electron density n_0:

$$\omega_p = \left(\frac{n_0 e_0^2}{m\varepsilon_0}\right)^{1/2} \qquad [29]$$

and the friction constant $\gamma = 1/\tau$ related to the collision time τ. The Drude model (eqn [28]) leads to the static conductivity

$$\sigma_{stat} = \frac{n_0 e^2}{m\gamma} = \frac{n_0 e^2}{m}\tau$$

but more relevantly here to the optical conductivity:

$$\sigma(\omega) = \frac{i\omega_p^2 \varepsilon_0}{\omega + i\gamma} \qquad [30]$$

and from eqn [13] ($\chi_p = 0$) to the typical metal dielectric function:

$$\varepsilon = 1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma} \qquad [31]$$

From the static conductivity, one can derive an estimate for τ or $\gamma \approx 10^{10}\ s^{-1}$, and find that at least for the visible frequencies, a good approximation is

$$\varepsilon(\omega) \approx 1 - \frac{\omega_p^2}{\omega^2} \qquad [32]$$

ε is negative for small frequencies ω, $\varepsilon = 0$ at $\omega = \omega_p$, and approaches 1 for values smaller than 1 for large ω.

The negative values of ε are responsible for the high reflectivity. In eqn [19] l is purely imaginary and $r = 1$ in [23] and [25] because the numerator and denominator are a conjugate complex. The imaginary l says that the electromagnetic waves do not travel inside the metal, but the fields decay exponentially with the penetration depth

$$d = \frac{1}{|l|} = \left|\left(\frac{\omega^2}{c^2}\,\mu\varepsilon\right)^{1/2}\right|^{-1} = \frac{\lambda_0}{2\pi|\varepsilon^{1/2}|} \qquad [33]$$

much smaller than the vacuum wavelength λ_0. At very low frequencies, ε is determined by γ (eqn [31]), varies little, and d decreases with growing frequency ("skin effect"). There is a turnover ("anomalous skin effect") when $|\varepsilon|$ decreases and approaches zero at ω_p (eqn [32]). At ω_p, l becomes real, the waves are transmitted inside the metal, and the metal becomes transparent.

This threshold of transparency, ω_p, is characterized by the density of conduction electrons in simple metals (eqn [29]). This transparency is not seen because the plasma frequencies lie in the near to far ultraviolet (for tables of plasma frequencies, see "Further reading" section).

Formally, metal optics is normally developed completely parallel to the optics of transparent media. Only the dielectric function is different. It is purely imaginary, $\sim\sigma$, in the infrared (eqn [13]), mainly real and negative in the visible (eqns [31] and [32]), has a zero at ω_p in the near ultraviolet beyond which the metal is transparent with an index of refraction smaller than 1 (eqn [32]). At frequencies below ω_p, electromagnetic fields decay exponentially inside the metal and the optical energy is reflected.

Due to the fact that ε always has some real and some imaginary part, the statements just made are approximate. Around ω_p, for instance, ε (eqn [31]) does not go exactly through zero, and there is a smooth transition from the wave vector $k = (\omega/c)\sqrt{\mu\varepsilon}$ being dominated by the imaginary part to the propagating frequencies with mainly real k, from the anomalous skin effect to transparency.

The approximation [27]–[32], essentially for nearly free electrons, does not work well for d-band transition metals or for the lanthanides. For these metals, a model must at least include a polarizable background leading to an additional χ_p as in eqn [13]. The screening by this background can shift the frequency for $\varepsilon(\omega) \approx 0$ drastically, for example, for silver, from $\hbar\omega_p = 9$ eV according to eqn [29] to the measured $\hbar\omega_p = 3.8$ eV.

In most frequency regions of interest in optics, one cannot determine the index of refraction $\tilde{n} = n(1 + i\kappa) = \sqrt{\mu\varepsilon}$ from the angle of refraction and from light attenuation (absorption). One has to use the information in reflection. The difference in reflectivity of s- and p-polarized light is exploited in ellipsometry. When linearly polarized light with an (equal) s- and

p-component is reflected from a metal surface, there generally appears a difference in phase and amplitude of the reflected components leading to elliptical polarization in the reflected beam. Measurements of the ellipse, the rotation angle of the main axis, and the difference of the two main axes, allow the determination of the real and imaginary part of $\tilde{n}$ or ε, because the reflection amplitudes eqns [22] and [24] contain the relations. The realm of variation of the phase on reflection can be easily estimated from the reflection amplitudes [22] and [24]. The transversality means $R_x = (q/k_x)R_z$ and $E_{0x} = -(q/k_x)E_{0z}$, $R_x/E_{0x} = -R_z/E_{0z}$. Then for normal incidence ($k_x \to 0$), s- and p-reflection amplitudes are the same:

$$\frac{R_x}{E_{0x}} = \frac{\sqrt{\varepsilon_1\mu_M} - \sqrt{\varepsilon_M\mu_1}}{\sqrt{\varepsilon_1\mu_M} + \sqrt{\varepsilon_M\mu_1}} \approx \frac{\tilde{n}_1 - \tilde{n}_M}{\tilde{n}_1 + \tilde{n}_M}$$
$$(\text{for } \mu_1 \approx \mu_M \approx 1 \text{ and } \tilde{n} = \sqrt{\varepsilon}) \qquad [34]$$

If the angle of incidence grows up to glancing incidence ($q \to 0$), the limits are: for p-polarization $R_x/E_{0x} \to 1$, for s-polarization $R/E_0 \to -1$, that is, a phase difference of π. Therefore, one can expect an angle θ, called "the principle angle of incidence," at which the phase difference is $\delta = \pi/2$, that is, circularly polarized light is reflected as linearly polarized. This angle for metals is reminiscent of the Brewster angle for nonmetallic reflectors, which is seen when the numerator in the p-reflection amplitude eqn [22] is zero. Strongly negative ε_M prohibits such a zero. A Brewster angle is also found for metals in the transparent region beyond ω_p, when the imaginary part of ε is small.

The measured effective dielectric functions ε or indices of refraction $\tilde{n}$ of many metals for different frequency regions can be found in several tables (see the "Further reading" section).

Because metal optics is often studied with very thin metal layers, it may be mentioned again that the distinction between nonpropagating and propagating waves is not sharp because l in eqn [19] always has a real and an imaginary part due to the complex ε. Formally, the fields inside the metal and the transmitted field are calculated as in standard optics : general solution on top, general solution inside, general solution below, and the four amplitudes (reflected, down and up inside, transmitted) determined from the four boundary conditions, [20] and [21] at each surface. This yields the exact fields everywhere. There is no use in summing an infinite series of scatterings from the top and bottom of the layer (and approximating this summation by one or two reflections).

The metal optics described so far is essentially a summary of the relevant chapters in the famous book of Born and Wolf, where extensions and details can be found (see the "Further reading" section).

Surface Plasmon Polaritons

In the region where $\varepsilon < 0$, the reflection amplitude eqn [22] has no zero in the numerator, therefore there is no Brewster angle; but the denominator can be zero (eqn [35]). For those values of ω and k_x, there can be reflected and transmitted amplitudes when the incident field vanishes: one finds eigensolutions of the system, which are called surface polaritons or surface plasmon polaritons (SPP) or surface plasmons. Their dispersion relation ω_s ($k_{\text{tangential}}$) or $k_{tg}(\omega_s)$ is found from eqns [22] and [19]:

$$\varepsilon_M(\omega_s)\sqrt{\frac{\omega_s^2}{c^2}\varepsilon_1\mu_1 - k_{tg}^2} + \varepsilon_1\sqrt{\frac{\omega_s^2}{c^2}\varepsilon_M(\omega_s)\mu_M - k_{tg}^2} = 0 \qquad [35]$$

$$k_{tg} = \frac{\omega_s}{c}\sqrt{\frac{\varepsilon_M(\omega_s)\varepsilon_1}{\varepsilon_M(\omega_s) + \varepsilon_1}} \quad (\mu_M = \mu_1 = 1) \qquad [36]$$

The SPP dispersion relations [35] and [36] are best discussed with ε approximated by eqn [32]. For small frequencies ω_s, an expansion with respect to ω_s/ω_p yields

$$k_{tg} = \frac{\omega_s}{c}\sqrt{\varepsilon_1}\left(1 + \frac{1}{2}\varepsilon_1\frac{\omega_s^2}{\omega_p^2}\right) \qquad [37]$$

while $k_{tg} \to \infty$ when the denominator in [36] goes to zero at

$$\omega_s \to \frac{\omega_p}{\sqrt{1+\varepsilon_1}} \text{ or } \omega_s \to \frac{\omega_p}{\sqrt{2}} \quad (\text{for } \varepsilon_1 = 1) \qquad [38]$$

Solutions of eqn [36] for small k_{tg} and $\omega > \omega_p$ are not eigenmodes but indicate the Brewster angle. The dispersion relation for the SPP is shown schematically in **Figure 1**.

One essential point is that $k_{tg} > (\omega_s/c)\sqrt{\varepsilon_1}$ (see eqn [37]), that is, the wavelength of the SPP is shorter than the wavelength of light at the same frequency in the medium 1 above the metal surface, the eigenmode lies "to the right of the light line $k = (\omega/c)\sqrt{\varepsilon_1}$." Therefore, according to eqn [19], the z-components q and l of the wave vector of the surface wave are imaginary; the electromagnetic fields decay exponentially toward both sides of the surface. The excitation energy in the eigenmode is bound to the surface and does not radiate away. On the other hand, this eigenmode cannot directly couple to, or be

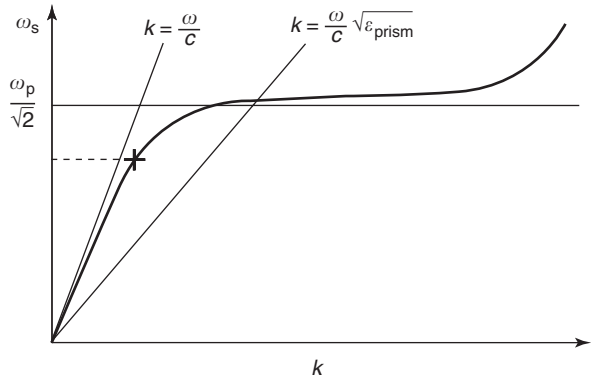


Figure 1 The surface plasmon dispersion. At "+" the wave in a prism can couple to the SPP in an ATR configuration.

excited by the light incident on the surface, because even for grazing incidence, k_{tg} of the light is too small; the periodicity parallel to the surface does not match the one of the SPP. At very large k_{tg}, the eigenfrequency rises above the limit of eqn [38] in **Figure 1**. This is a real plasma effect and is treated in the following section.

These eigenmodes were first derived by Zenneck and Sommerfeld, who explained the transmission of radio waves along the ocean surface. At metal surfaces, they were first seen as energy losses of electrons scattered from metal layers. In optics, it is possible to couple to the SPP if the incident light comes through a prism which shortens λ or increases k. In **Figure 1**, the SPP dispersion is shown for a metal surface towards vacuum or air ($\varepsilon_1 = 1$). Also indicated is the "light line" $(\omega/c)\sqrt{\varepsilon_{prism}} = k$. For angles of incidence φ with

$$k_x = \sin\varphi \frac{\omega}{c}\sqrt{\varepsilon_{prism}} > \frac{\omega}{c} \qquad [39]$$

the light is totally reflected from the prism–air interface because l (eqn [19]) in air becomes imaginary. But when this prism surface is brought in close proximity to the metal surface, at the angle where k_x from [39] fulfills the dispersion relation of the SPP [36], the eigenmode is excited, it develops large fields, soaks up energy, and attenuates the total reflection of the prism. This attenuated total reflection (ATR) method has been widely used to study the SPPs as well as the properties of layers adsorbed on metal surfaces, because adsorbed layers shift the dispersion according to $\varepsilon_1 \neq 1$ in eqn [36]. The large fields near the metal surface, when the SPP is excited, have been used to excite molecules especially brought near the surface. On the other hand, a microscope has been built on the ATR principle in order to see structures in adsorbed layers, which appear in self-organized surface reactions. The SPP can be excited by one prism, then travel along the metal surface and be coupled out by a second prism, and the metal surface between the prisms can be manipulated.

The narrow coupling condition that k_x of the exciting light must fulfill the SPP dispersion eqn [36] is true when the metal surface is translationally invariant. Surfaces with some roughness or any local obstacles can scatter light especially strong near $\omega = \omega_p/\sqrt{1+\varepsilon_1}$, where the density of eigenmodes is very high. If a periodic structure is established on the metal surface, for example, a grid of grooves, the related reciprocal lattice vectors can be added or subtracted to k_{tg} in order to overcome the restrictions for SPP excitation. Then refolding of $\omega_s(k_{tg})$ and gaps due to Bragg diffraction apply as in electron band structures, and there appear possibilities for coupling the SPPs to incident light. Also any irregularities, roughness, adsorbed molecules or metal clusters, which break the translational invariance of the surface, make excitation of surface modes possible. These often lead to extremely large fields around the imperfections. An example is surface-enhanced Raman scattering (SERS) from adsorbed molecules on silver.

Surface waves also play a role when light is transmitted through apertures smaller than the wavelength as in near-field microscopy. Recently, it was shown that light is transmitted through an array of narrow holes in a thick metal layer at a surprisingly large rate. Also, this phenomenon is explained via excitation of SPPs. The same mechanism leads to strong absorption at special frequencies in a system of grooves on a metal surface.

Generally, the excitation of SPPs leads to strong absorption of energy in the metal, to an attenuation of reflection, because the large fields of the eigenmode cause large currents and large losses in the metal. Occasionally, one finds high-energy transport at the excitation condition leading to strong scattering or even high transmission through narrow holes.

When the metal layer is thin enough, the two surfaces excite each other. Then the surface modes on both sides couple symmetrically or antisymmetrically, and the dispersion relations are changed.

Plasma Waves in Metal Optics

In 1953, Bohm and Pines discovered that the electrons in metals can sustain charge density waves with longitudinal electric fields. In an electron gas model of a metal, they found the dispersion relation for plasma waves:

$$\omega^2 = \omega_p^2 + \beta k^2 \qquad [40]$$

The Drude model (eqn [28]) leads to plasma oscillations. Taking the divergence of eqn [28] and using the continuity equation and eqn [9] yields

$$\ddot{\rho} = -\omega_p^2 \rho - \gamma \dot{\rho} \qquad [41]$$

that is, damped oscillations with the plasma frequency ω_p.

In addition to the electric field, the electrons are driven by a gradient in electron pressure, in density, in the local Fermi energy proportional to $(n_0)^{2/3}$. This pressure spreads out the electrons at metal surfaces and metal–metal interfaces leading to surface dipoles and contact potentials. In the Drude model, this additional force can be introduced by a term $-\beta\nabla\rho$:

$$\frac{\partial \boldsymbol{j}}{\partial t} = \omega_p^2 \varepsilon_0 \boldsymbol{E} - \gamma \boldsymbol{j} - \beta \nabla \rho \qquad [42]$$

The coefficient $\beta = (3/5)v_F^2$ in a free electron gas with Fermi velocity v_F^8. The expression for the high-frequency electron gas pressure was derived by Bloch. Equation [42] can be derived from a Boltzmann equation for the electron gas, and it is called the "hydrodynamic approximation" to a quantum mechanical treatment of the free electron gas. This extended Drude model describes plasma waves and makes it easy to include them in metal optics in the correct way.

First, the plasma waves are seen as eigensolutions of the system. Equation [41] is now supplemented to yield

$$\ddot{\rho} = -\omega_p^2 \rho - \gamma \dot{\rho} + \beta \nabla \cdot (\nabla \rho) \qquad [43]$$

which is solved by a wave with the plasmon dispersion [40] when the damping γ is neglected. The mode is a charge density wave (analog to a longitudinal acoustic density wave) with longitudinal electric fields $\boldsymbol{E} \| \boldsymbol{k}$. Transverse waves have necessarily $\rho = 0$ from eqn [9]. Therefore, the response, conductivity $\sigma_\perp$ and $\varepsilon_\perp$, is not affected by the new term in [42], and remains as in eqns [30] and [31]. But the longitudinal response is changed. For a plane wave pertubation $\sim e^{i(\boldsymbol{kr}-\omega t)}$, the continuity equation $(1/i\omega)\nabla \cdot \boldsymbol{j} = \rho$ replaces ρ in eqn [42], and for longitudinal fields and currents,

$$\sigma_\parallel(k,\omega) = \frac{i\omega\omega_p^2\varepsilon_0}{\omega^2 - \beta k^2 + i\omega\gamma} \qquad [44]$$

and with the general relation [13]:

$$\varepsilon_\parallel(k,\omega) = 1 - \frac{\omega_p^2}{\omega^2 - \beta k^2 + i\omega\gamma} \qquad [45]$$

The response functions are now dependent on frequency ω and wave vector $\boldsymbol{k}$. This is the so-called nonlocal response (in time and now also in space), which is familiar in nonlocal metal optics.

What do Maxwell's equations tell one about longitudinal waves? Equation [3] shows that longitudinal plasma waves have no magnetic induction field, $\boldsymbol{B} = 0$. On the other hand, the longitudinal wave is a homogeneous solution of the wave equation [12]:

$$\nabla(\nabla \cdot \boldsymbol{E}) - \Delta \boldsymbol{E} = \frac{\omega^2}{c^2} \mu \varepsilon_\parallel \boldsymbol{E} \qquad [46]$$

For a wave with $\boldsymbol{E} \| \boldsymbol{k}$, the left-hand side is zero and a nontrivial solution requires the general dispersion relation for plasmons

$$\varepsilon_\parallel(k,\omega) = 0 \qquad [47]$$

With eqn [45], neglecting the damping γ, [47] yields [40]. The hydrodynamic approximation (eqns [42], [44], and [45]) should be considered as the long wavelength, small k expansion of any sophisticated treatment of conduction electron response. For the purpose of optics, this approximation has been shown to be sufficient for the analysis of many experiments involving metal surfaces or metal layers at frequencies around the plasma frequency.

The longitudinal plasma waves are homogeneous solutions of Maxwell's equations according to eqn [46], and must therefore be included in the general solutions, which via boundary conditions yield the optical fields. Sauter was the first to point this out in 1964, and he suggested the third boundary condition, which is necessary at a free metal surface to determine the three amplitudes: of the reflected wave $\boldsymbol{R}$ (eqn [17]) and inside the metal of the transmitted transversal wave $\boldsymbol{T}$ (eqn [18]) and of the longitudinal wave $\boldsymbol{L}e^{i(k_x x + \eta z)}$, which completes the general solution [18].

Before the boundary conditions are discussed, the question of energy current is addressed. When it was brought up that the longitudinal waves should be excited on reflection at metal surfaces, it was objected that this additional channel of energy transport would have long been recognized if the suggestion was true. The wrong answer is that the plasma wave does not carry energy, because $\boldsymbol{B} = 0$ and therefore, the Poynting vector $\boldsymbol{S} = 0$. But the excitation of plasma waves becomes appreciable for frequencies of the order ω_p, which is in the UV where hardly any reflection measurements have been done. In addition, the energy current of the plasma wave is small

because of the very small group velocity of plasmons:

$$\frac{\partial\omega}{\partial k}\approx\beta\frac{k}{\omega_{\mathrm{p}}}\sim\frac{v_{\mathrm{F}}^2}{c} \qquad [48]$$

The extension of the material equation [42] leads to an extension of Poynting's theorem: By taking the dot product of eqn [3] with $\boldsymbol{H}$ and subtracting the dot product of $\boldsymbol{E}$ with eqn [3], one obtains

$$-\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{E}\cdot\boldsymbol{D}+\frac{1}{2}\boldsymbol{H}\cdot\boldsymbol{B}\right)=\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H})+\boldsymbol{E}\cdot\boldsymbol{j} \qquad [49]$$

Multiplying the material equation of motion [42] by $\boldsymbol{j}$, the terms can be rearranged with the help of the continuity equation

$$\omega_{\mathrm{p}}^2\varepsilon_0\boldsymbol{E}\cdot\boldsymbol{j}=\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{j}^2\right)+\gamma\boldsymbol{j}^2+\nabla\cdot(\beta\rho\boldsymbol{j})+\frac{\partial}{\partial t}\left(\frac{1}{2}\beta\rho^2\right) \qquad [50]$$

which by replacing $\boldsymbol{E}\cdot\boldsymbol{j}$ in eqn [49] yields Poynting's theorem with plasma waves:

$$-\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{E}\boldsymbol{D}+\frac{1}{2}\boldsymbol{H}\boldsymbol{B}+\frac{1}{2\omega_{\mathrm{p}}^2\varepsilon_0}\boldsymbol{j}^2+\frac{1}{2\omega_{\mathrm{p}}^2\varepsilon_0}\beta\rho^2\right)=\nabla\cdot\left(\boldsymbol{E}\times\boldsymbol{H}+\frac{1}{\omega_{\mathrm{p}}^2\varepsilon_0}\beta\rho\boldsymbol{j}\right)+\frac{\gamma}{\omega_{\mathrm{p}}^2\varepsilon_0}\boldsymbol{j}^2 \qquad [51]$$

On the left-hand side, the kinetic energy density of the electron current and the potential energy density connected with the density perturbation against the electron gas pressure are found. On the right-hand side, one has the energy current density in the plasma wave and the positive definite energy absorption due to the friction coefficient γ.

The problem of "additional boundary conditions" is taken up next. At first, the total response is described by ρ and $\boldsymbol{j}$, that is, formally $\chi_{\mathrm{p}}=0,\ \boldsymbol{P}=0$. It has been mentioned already after eqn [21] that the standard boundary conditions, $D_{\mathrm{eff,n}}$ and H_{tg} continuous, imply a discontinuity in E_{n} caused by singular surface charges, and from eqn [8], the denial of singular surface currents. With $\boldsymbol{j}=\sigma\cdot\boldsymbol{E}$, one also gets a finite normal component of the current up to the surface, which then disappears in a singular sink or source of charge in the surface plane. In nonlocal optics, the charges and currents are treated seriously in the plasma waves, and especially singular charges would lead to infinite pressures in eqn [42]. The model of the surface must be improved. Singular surface charges can only be avoided if the normal component of the current $\boldsymbol{j}$ is continuous (zero at a free surface). With finite charge densities everywhere, the normal component of $\boldsymbol{E}$ is also continuous. These two independent conditions replace the continuity of $D_{\mathrm{eff,n}}$, which is then a consequence. Continuous $\boldsymbol{E}$ and continuous $\boldsymbol{D}_{\mathrm{eff}}$ does not contradict eqn [13], because σ and ε are now tensors with $\varepsilon_\perp\neq\varepsilon_\parallel$.

At an interface with metal only on one side, one has one reflected wave amplitude in the dielectric, and a transverse and a longitudinal wave amplitude in the metal to be determined via three boundary conditions:

(a) continuity of $E_{\mathrm{tangential}}$ [52]

(b) continuity of E_{normal} [53]

(c) $j_{\mathrm{normal}}=\sigma_\perp E_{\perp\mathrm{normal}}+\sigma_\parallel E_{\parallel\mathrm{normal}}=0$ [54]

For the interface between two metals, a fourth amplitude via a fourth boundary condition must be determined because there are plasma waves in both metals. A reasonable requirement is that the interface should not be a singular sink or source of energy. This plausible assumption for our surface model leads to the continuity of the normal component of the energy current; so the boundary conditions at a metal–metal interface are

(a) continuity of $E_{\mathrm{tangential}}$ [55]

(b) continuity of E_{normal} [56]

(c) continuity of $j_{\mathrm{normal}}=\sigma_\perp E_{\perp\mathrm{normal}}+\sigma_\parallel E_{\parallel\mathrm{normal}}$ [57]

(d) continuity of $\frac{\beta}{\omega_{\mathrm{p}}^2\varepsilon_0}\rho=\frac{\beta}{\omega_{\mathrm{p}}^2}\nabla\cdot\boldsymbol{E}=\frac{\beta}{\omega_{\mathrm{p}}^2}\boldsymbol{k}_\parallel\cdot\boldsymbol{E}_\parallel$ [58]

The conductivity $\sigma_\perp$ for transverse fields is also from eqn [42] equal to eqn [30], and $\varepsilon_\perp$ is given by eqn [31]. With (a), (b), and (c), the Poynting vector is continuous; so the requirement [58] suffices to make the total normal energy current continuous. All boundary conditions lead to linear equations for the determination of the reflection and transmission amplitudes, therefore to no complication, in principle, compared to standard optics.

In this treatment of plasma waves in optics, the response functions $\sigma_\parallel$ and $\varepsilon_\parallel$ are dependent on frequency ω as well as on wave vector k or wavelength; hence, this kind of optics is called nonlocal optics. Equation [42] says that the current at r is not only driven by the electric field at $\boldsymbol{r}$ (local relation), but the

charge density in the neighborhood determines the other force $\sim\nabla\rho$. Charge density waves with larger gradients, larger k, oscillate faster according to eqn [40]. The standard local optics is an approximation which neglects charge densities altogether and does not treat the singular surface charges consistently which eqn [21] implies. This standard approximation is simpler and sufficient as long as the charge densities are small. At frequencies of the order of the plasma frequency, the charge densities get large due to resonant eigenmodes and must be treated consistently. The nonlocal optics can lead to results qualitatively different from standard optics and in agreement with the experiment.

A few examples of this nonlocal metal optics are discussed next. One can understand at the beginning that only when there are normal components of $\boldsymbol{E}$, normal currents leading to oscillating charge densities near the surface, a plasma wave is optically excited. Therefore, only for p-polarization with the electric field parallel to the plane of incidence, this nonlocal extension of optics is necessary, and the plasmon gets a nonzero amplitude $\boldsymbol{L}$.

In the case of reflection from a free metal surface, the extension of the treatment in eqns [17]–[23] takes the steps: the outgoing general solution [18] inside the metal is supplemented by the plasma-wave $+\boldsymbol{L}e^{i(k_x x+\eta z)}$ with η from

$$\varepsilon_{\parallel}=0=1-\frac{\omega_p^2}{\omega^2-\beta(k_x^2+\eta^2)+i\omega\gamma} \qquad [59]$$

and with

$$\eta=\sqrt{\frac{1}{\beta}[\omega(\omega+i\gamma)-\omega_p^2]-k_x^2} \qquad [60]$$

and for the longitudinal wave

$$L_x/L_z=k_x/\eta \qquad [61]$$

The three boundary conditions [52]–[54] lead to three linear equations and to the reflection amplitude

$$\frac{R_z}{E_{0z}}=\frac{\varepsilon_\perp q-l+(\varepsilon_\perp-1)k_x^2/\eta}{\varepsilon_\perp q+l-(\varepsilon_\perp-1)k_x^2/\eta} \qquad [62]$$

with $\varepsilon_\perp=\varepsilon_M$ and q, l from eqns [31], [19], and $\varepsilon_1=1,\ \mu_M=\mu_1=1$. The limit $\beta\to 0$ leads back from [42] to local [28], in [60] to $\eta\to\infty$, and from [62] back to [22]. The difference between $r=|R_z/E_{0z}|^2$ from [62] and from [22] is shown in **Figure 2**. The excitation of plasma waves really reduces the reflectivity, but only above the plasma frequency and by a small amount. This case was outlined to demonstrate that this extension of metal optics is straightforward.

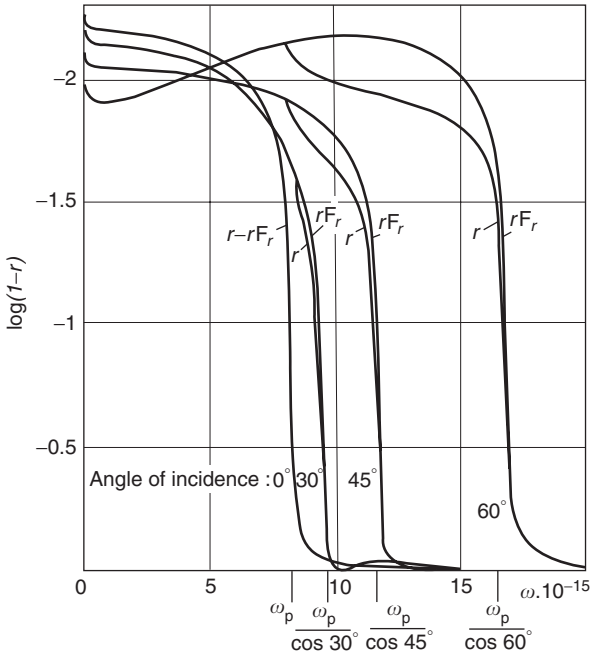


Figure 2 The reflectivity r with plasma waves, r_{Fr} standard Fresnel formula.

As in the previous section, one finds from the zero of the denominator of the reflection amplitude [62], the dispersion relation of the surface plasmon polariton

$$\varepsilon_\perp q+l-(\varepsilon_\perp-1)k^2/\eta=0 \qquad [63]$$

For low frequencies, the last term of the order of $k(v_F\omega/c\omega_p)$ can be neglected and the dispersion agrees with [35]. But for $k\approx(\omega/c)(c/2\sqrt{\beta})^{1/3}$, the frequency of the eigensolution [63] rises above $\omega_p/\sqrt{2}$, which was the limit for [35]. The approximate dispersion in this frequency range is from eqn [63]:

$$\omega_s\approx\frac{\omega_p}{\sqrt{2}}+k\frac{\sqrt{\beta}}{2} \qquad [64]$$

This linear dispersion is often extrapolated toward $k=0$. In optics, this is wrong because then, the light line is crossed and light could couple to the eigenmodes which is not the case. For larger k, the increase of ω_s is stronger than linear which is indicated in **Figure 2**. This dispersion describes the essential facts of the experiments. Only the linear rise [64] is usually suppressed to a plateau due to a skin effect before the final rise, when at large k the electron gas pressure dominates.

The dispersion relation [64] was derived when Ritchie discovered this surface mode in 1957 from an electrostatic approximation to Maxwell's equations

($c \to \infty$) and called it the surface plasmon. When the charge densities in the plasma and Maxwell's equations are treated correctly, it becomes clear that [35] and [64] are approximations to the dispersion [63] of a single kind of surface mode, which is then properly termed the "surface plasmon polariton" (SPP).

The reflection from a simple metal surface is changed so little by the inclusion of plasma waves in optics (**Figure 2**) that this is not the experiment to prove the necessity of nonlocal optics. The first relevant experiment was theoretically predicted by Melnyk and Harrison in 1968 and carried out by Lindau and Nilson in 1970. The transmittance through thin silver layers shows periodic minima when the plasma waves form standing waves in the layer and by the large amplitude, increase the absorption. It turns out that half a wavelength and even three and five halves of a wavelength fit into a layer of 100 Å. One can calculate from eqn [40] that this is the order of the plasma wavelength, while the transverse waves are more than ten times longer, so the experiment can only be explained by the optical excitation of plasma waves in the metal. Another example is shown in **Figure 3**. An electroreflection experiment measures the change in reflectivity $\Delta r/r$ when a strong (static or slowly oscillating) electric field changes the charge density at the metal surface. The optical calculations give qualitatively different results when standard local optics or nonlocal optics is employed. Only nonlocal optics describes the experiment. In the late 1970s and early 1980s, several examples were published with the same outcome (see the "Further reading" section.)

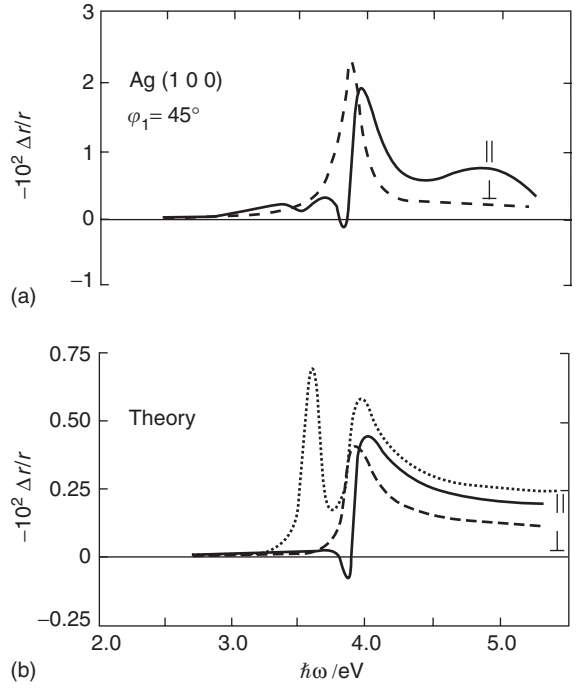


Figure 3 Electroreflection on silver. (a) Experiment for *s*-(⊥) and *p*-polarized (||) light. (b) Theory, (——) nonlocal optics, (....) standard optics.

One can draw the general conclusions: standard optics must be extended to nonlocal optics when the optical frequency ω is around the plasma frequency ($0.5\omega_p < \omega < 2\omega_p$) and when metallic layers, especially very thin ones (1–1000 Å), are involved. The nonlocal optics discussed in this article gives the right answers in these cases.

The treatment of plasma waves in optics in this section is designed for a free electron gas model of a metal, where the total response can be described as conduction electron current density and charge density. The parameters, electron density n_0 and damping γ, could be taken from a fit of a model (eqn [31]) to measure dielectric data. This is a good approximation for most metals except the *d*-band transition metals. But silver is one of the best candidates for demonstrating the effects of nonlocal optics because its plasma frequency is easily accessible just above the visible at $\hbar\omega_p = 3.8$ eV and the plasmons see little damping. The conduction electron density in silver is high and would lead to $\hbar\omega \approx 9$ eV from eqn [29]. $\varepsilon = 0$ is shifted to a lower frequency because of other "bound" electrons.

In such a situation, it has been successful to think of two responding electron systems, free and bound. The free conduction electrons follow eqn [42] while the bound ones are described by a polarizability χ_p. The total measured ε is then separated as in eqn [13]. An assumption of the conduction electron density (and the damping) allows the determination of χ_p from [13]. Then it is consistent to use in [53] and [56] the normal component of $\boldsymbol{D} = (1 + \chi_p)\varepsilon_0 \boldsymbol{E}$. This procedure was used for **Figure 3** and for many other comparisons with experiments on silver.

A final remark shall close this section: the response of a metal to electromagnetic fields is a vast field in solid-state theory where quite a number of theoreticians have invested their ingenuity. Here, a simple extension of standard optics was presented which can easily be applied by experimentalists also and has been successfully used for the analysis of several experiments. Around the plasma frequency, the predictions of nonlocal optics including plasma waves can be qualitatively different from the results of standard optics.

Magnetooptics

The response of metals as well as other materials to electromagnetic waves is determined by the forces

due to the electric field of the wave, because the Lorentz force due to the magnetic field is weaker by a factor $v/c < 1\%$ with electron velocity of the order of the Fermi velocity. Magnetooptics considers the interaction of light with external magnetic fields or strong magnetizations in ferromagnets.

A system with a magnetic field or magnetization acquires a special symmetry with respect to time inversion $(\boldsymbol{E}, \boldsymbol{B}, \boldsymbol{H}, \boldsymbol{M} \rightarrow \boldsymbol{E}, -\boldsymbol{B}, -\boldsymbol{H}, -\boldsymbol{M})$. Onsager's relations require

$$\varepsilon_{ij}(E, B, H) = \varepsilon_{ji}(E, -B, -H) \qquad [65]$$

Therefore, optical effects introduced by (proportional to) $\boldsymbol{B}$ or $\boldsymbol{M}$ manifest themselves as antisymmetric tensor components in an otherwise symmetric dielectric tensor $\hat{\varepsilon}$.

It is briefly shown that an external magnetic field leads to antisymmetric ε-components in an extension of the Drude model (eqns [27] and [28]) and that the consequence in optics is a difference in index of refraction for right and left circularly polarized waves. This leads to the Faraday rotation of the plane of polarization for linearly polarized light in transmission and to the Kerr rotation in reflection.

The Lorentz force $-e_0(\dot{\boldsymbol{r}} \times \boldsymbol{B}_0)$, in an external magnetic field $\boldsymbol{B}_0 = (0, 0, B_0)$ parallel to the z-axis, is added to the equation of motion [27]. To be more general, an elastic restoring force $-\omega_0^2 \boldsymbol{r}$ is also added (for free electrons in metals $\omega_0 \rightarrow 0$):

$$m\ddot{\boldsymbol{r}} = -e_0\boldsymbol{E} - m\gamma\dot{\boldsymbol{r}} - m\omega_0^2\boldsymbol{r} - e_0(\dot{\boldsymbol{r}} \times \boldsymbol{B}_0) \qquad [66]$$

With $\boldsymbol{r} \sim \exp(-\mathrm{i}\omega t)$ and $\boldsymbol{j} = -e_0 n_0 \dot{\boldsymbol{r}}$, one finds from [66] the conductivity tensor and obtains from eqn [13] the dielectric tensor:

$$\hat{\varepsilon} = \hat{1} + \omega_\mathrm{p}^2 \begin{pmatrix} a & \mathrm{i}b & 0 \\ -\mathrm{i}b & a & 0 \\ 0 & 0 & c \end{pmatrix}$$

$$= \varepsilon \begin{pmatrix} 1 & \mathrm{i}Q & 0 \\ -\mathrm{i}Q & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} \qquad [67]$$

$$a = \frac{\omega_0^2 - \omega(\omega + \mathrm{i}\gamma)}{(\omega_0^2 - \omega(\omega + \mathrm{i}\gamma))^2 - \omega^2\omega_\mathrm{c}^2}$$

$$b = \frac{\omega\omega_c}{(\omega_0^2 - \omega(\omega + \mathrm{i}\gamma))^2 - \omega^2\omega_\mathrm{c}^2} \qquad [68]$$

$$c = \frac{1}{\omega_0^2 - \omega(\omega + \mathrm{i}\gamma)} \qquad [69]$$

$\omega_c = eB_0/m$ is the cyclotron frequency proportional to B_0. Due to the external magnetic field, one finds an antisymmetric tensor added to an otherwise symmetric $\hat{\varepsilon}$.

The dielectric tensor is now employed on the right-hand side of the wave equation [12]. For a wave traveling in the z-direction parallel to $\boldsymbol{B}_0$

$$\left(\frac{\omega^2}{c^2}\varepsilon - k^2\right)E_x + \mathrm{i}Q\varepsilon\frac{\omega^2}{c^2}E_y = 0$$
$$-\mathrm{i}Q\varepsilon\frac{\omega^2}{c^2}E_x + \left(\frac{\omega^2}{c^2}\varepsilon - k^2\right)E_y = 0 \qquad [70]$$

$$\left(\frac{\omega^2}{c^2}\varepsilon(1 \pm Q) - k^2\right)(E_x \pm \mathrm{i}E_y) = 0 \qquad [71]$$

which yields

$$k_\mathrm{r}^2 = \frac{\omega^2}{c^2}\varepsilon(1 + Q) \quad \text{for } E_x = \mathrm{i}E_y$$
(for right circular polarized light)

$$k_\mathrm{l}^2 = \frac{\omega^2}{c^2}\,\varepsilon(1 - Q) \quad \text{for } E_x = -\mathrm{i}E_y$$
(for left circular polarized light)

$$\text{or} \quad n_\mathrm{r} = \sqrt{\varepsilon}\left(1 + \tfrac{1}{2}Q\right) \quad \text{and}$$
$$n_\mathrm{l} = \sqrt{\varepsilon}\left(1 - \tfrac{1}{2}Q\right) \qquad [72]$$

The increase of the effective ε or polarizability for right circular polarization can be understood for the simple model (eqn [66]) from the direction of the Lorentz force which increases the radius of the circular electron motion. When light travels through a piece of matter of length L, the waves accumulate a phase difference of

$$\delta = (k_\mathrm{l} - k_\mathrm{r})L = \frac{\omega}{c}L(n_\mathrm{l} - n_\mathrm{r}) = -\frac{\omega}{c}LQ \qquad [73]$$

The Faraday rotation of the plane of linear polarization is $\delta/2$. Because Q is proportional to ω due to the Lorentz force, the Faraday rotation is proportional to ω^2.

Generalizations

The dielectric tensor is more generally expressed as

$$\hat{\varepsilon} = \varepsilon \begin{pmatrix} 1 & \mathrm{i}Q_z & -\mathrm{i}Q_y \\ -\mathrm{i}Q_z & 1 & \mathrm{i}Q_x \\ \mathrm{i}Q_y & -\mathrm{i}Q_x & 1 \end{pmatrix} \qquad [74]$$

by the "Voigt vector" $\boldsymbol{Q} = (Q_x, Q_y, Q_z)$.

Then in the refraction indices (eqn [72]), Q is generalized: $Q \rightarrow \boldsymbol{Q}\boldsymbol{k}/|\boldsymbol{k}|$, to the projection of $\boldsymbol{Q}$ on

the direction of propagation $\boldsymbol{k}$. In the model and in transparent materials, Q is mainly real and one finds the Faraday rotation due to the real phase shift (eqn [73]). In ferromagnetic metals, the magnetooptical effects are larger by several orders of magnitude, but here, mainly the absorption, the imaginary part of Q is changed by the magnetization leading to elliptically polarized light in general.

Because ferromagnetism is rooted in the electron spin, a microscopic explanation of the magnetooptical effects in ferromagnets involves the spin–orbit coupling of the electron motion and the unequal number of spins parallel and antiparallel to $\boldsymbol{M}$.

For the investigation of magnetic properties of metallic layers, the magnetooptical Kerr effect (MOKE) in reflection has been developed into an efficient tool. With strong lasers, nonlinear optical response is also exploited (for detailed information see the "Further reading" section).

"Left-Handed Materials," Optics for $\varepsilon<0$ and $\mu<0$

Charge carriers exposed to oscillating electric fields show, in some frequency region, negative polarizability as a forced oscillator has a phase shift of π above resonance. Therefore, in simple metals $\varepsilon<0$ below the plasma frequency (resonance frequency $\omega_0\to 0$) (eqns [31] and [32]). Then the index of refraction $n=\sqrt{\varepsilon}$ is imaginary, electromagnetic waves decay exponentially into the metal, do not propagate at all and the reflectivity goes to 1.

In 1968, V G Veselago pointed out that the wave propagation is really determined by the product $\varepsilon\mu$ (eqns [16] and [19]) and therefore, a system with $\varepsilon<0$ and $\mu<0$ in the same frequency region would be transparent again. One has, in principle, negative magnetic polarizability in diamagnets (due to Lenz's rule) but usually the polarizabilities χ_m are extremely small ($\sim 10^{-6}$), and the susceptibility $\mu=1+\chi_m$ deviates little from 1 and does not go below 0.

Recently materials with $\mu<0$ have been manufactured. Therefore, the surprising optical properties are outlined. The term "left-handed material" (LHM) derives from the fact that Maxwell's equations (e.g., (eqn [3])) for a plane wave $\boldsymbol{k}\times\boldsymbol{E}=\omega\boldsymbol{B}$ and $\boldsymbol{H}=(1/\mu_0\mu)\boldsymbol{B}$ usually make the vectors $\boldsymbol{k},\boldsymbol{E},\boldsymbol{H}$ a right-handed tripod, while with $\mu<0$ the tripod gets left-handed. Consequently, normally the energy current $\boldsymbol{S}=\boldsymbol{E}\times\boldsymbol{H}$ is parallel to the wave vector $\boldsymbol{k}$. In LHMs, the Poynting vector is opposite to $\boldsymbol{k}$.

Refraction at an interface between material I ($\varepsilon_I>0,\ \mu_I>0$) and material II ($\varepsilon_{II}<0,\ \mu_{II}<0$) puts the refracted beam in II on the opposite side of the surface normal than usual. Therefore, diverging rays in I can converge in II which led to speculations about perfect lenses by slabs of LHMs. This uncommon refraction can be understood as follows:

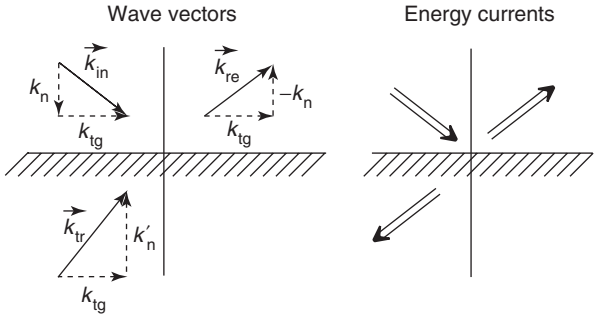


Figure 4 Refraction at a left-handed material.

$\boldsymbol{S}=\boldsymbol{E}\times\boldsymbol{H}$ is antiparallel to $\boldsymbol{k}$ in LHM II. Therefore, the correct homogeneous solution in II which carries the energy from the interface has a wave vector pointing towards the interface. The tangential components of all wave vectors are equal, so the refracted beam (i.e., the energy current) lies in II on the same side of the surface normal as the incoming beam in I (see **Figure 4**).

The subject of optics with LHMs has been revived after 2000 in connection with the investigation of the "optical" properties of conducting networks, three-dimensional wire mesh crystals. If the wavelength of the electromagnetic waves are much larger than the lattice spacing, the conducting network behaves like a metallic conductor with a low density plasma frequency and with a frequency region of negative ε according to eqn [32]. The necessary large diamagnetism is achieved by conducting rings in each cell of the lattice. In such a designed material, $\varepsilon<0$ and $\mu<0$ at frequency $\nu=10.5$ GHz, $\lambda\approx 3$ cm, was demonstrated with $\varepsilon\mu=7.28\pm i0.54$. It was shown that the refraction is indeed as in **Figure 4** (see the "Further reading" section).

See also: Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Metals and Alloys, Impurity and Defect States in.

PACS: 73.61.At; 78.66.Bz

Further Reading

Bennemann KH (ed.) (1998) *Nonlinear Optics in Metals*. Oxford: Clarendon Press.

Boardman AD (ed.) (1982) *Electromagnetic Surface Modes*. New York: Wiley.

Born M and Wolf E (1975) *Principles of Optics*, Ch. XIII. Pergamon.

Bouhelier A, Renger J, Beversluis MR, and Novotny L (2003) *Journal of Microscopy* 210: 220.

Flaetgen G, Krischer K, Pettinger B, Doblhofer K, Junkes H, *et al.* (1995) *Science* 269: 668.
Forstmann F and Gerhardts RR (1986) Metal optics near the plasma frequency. *Springer Tracts in Modern Physics* 109.
Palik ED (1985/1991) *Handbook of Optical Constants of Solids*, vols. I and II. London: Academic Press.
Raether H (1980) Excitation of plasmons and interband transitions of electrons. *Springer Tracts in Modern Physics* 88.
Raether H (1988) Surface plasmons on smooth and rough surfaces and on gratings. *Springer Tracts in Modern Physics* 111.
Shelby RA, Smith DR, and Schultz S (2001) *Science* 292: 77.

Micro- and Macrostructures

G S Rohrer, Carnegie Mellon University, Pittsburgh, PA, USA

Introduction

A crystalline solid is made up of indistinguishable groups of atoms positioned at the sites of a Bravais lattice. If the atomic structure is approximately uniform in both orientation and composition throughout the solid, the material is referred to as a single crystal. When the atomic structure is not uniform, it is known as a polycrystal. In this second and more common instance, the irregularities in the orientation, composition, and configuration of the atomic structure collectively make up what is referred to as the internal structure or microstructure. The term macrostructure is sometimes used to refer to the largest components of the internal structure. Within this article, all components of the internal structure will be referred to as the microstructure.

Polycrystals can be thought of as aggregates of much smaller crystals joined together by a network of internal interfaces, as illustrated schematically in **Figure 1**. Within the aggregate, the individual crystals are referred to as grains. The interface that separates two grains of the same composition, but different lattice orientation, is called a grain boundary and the interface that separates two grains with different compositions is called a phase boundary. The granular nature of polycrystalline matter has been recognized since at least 1775 when Grignon sketched grains observed on the surface of a piece of fractured wrought iron. This first recorded observation points to an important feature of the interfaces within the microstructure: they are frequently points of internal weakness along which failure, by fracture or corrosion, propagates through the interfacial network. During the last century, microstructural studies have flourished amid the realization that many of the macroscopic properties of polycrystalline solids are strongly influenced by the internal structure. While these studies established the solid foundation of knowledge that will be reviewed in this article, it should also be recognized that there are many unresolved questions about microstructural genesis and the microstructure–property link that continue to be actively researched. In general terms, one can say that polycrystalline microstructures result from crystal nucleation, growth, and capillarity-driven morphological changes that lead to a reduction in the interfacial energy. In this article, the basic principles behind these processes will be described and examples of prototypical microstructures will be presented. The article will focus on dense, inorganic, polycrystalline microstructures that occur in man-made materials. Single-phase microstructures are discussed first and multiphased structures are described in the second section.

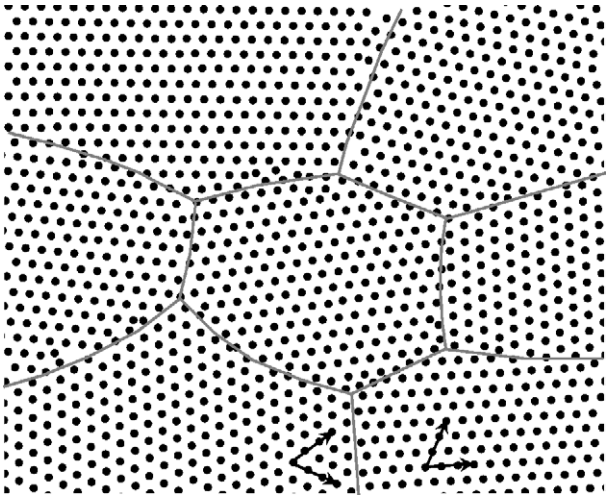


Figure 1 Schematic illustration of a polycrystalline solid, with black dots representing points on the lattice and gray lines denoting boundaries between grains with different orientations.

Single-Phase Microstructures

The term single-phase polycrystal is used to describe an aggregate of crystals that is uniform in its atomic structure and composition. Throughout a single-phase microstructure, it is only the orientation of the Bravais lattice within the individual grains that

differs. The defining elements of the microstructure are the distribution of crystal sizes, the distribution of crystal shapes, the distribution of crystal orientations, and the types of grain boundaries between the crystals. Single-phase microstructures can be formed by sintering a solid powder, by crystallizing a precursing liquid, gaseous, or amorphous phase, or by recrystallizing a deformed precursor. Microstructures resulting from each of these states are described briefly below.

Single-Phase Microstructures by Sintering

The excess free energy per unit area of a grain boundary (γ) can be thought of as the energy to create equal areas of the two free surfaces on either side of the boundary, minus a binding energy associated with the formation of bonds across the interface as the surfaces are joined. Therefore, as long as there is bonding across the interface, it is energetically preferable to replace two free surfaces with a grain boundary of the same area. This is the driving force for the process of sintering, in which individual crystalline particles change shape and bond with one another to form a dense polycrystal. Further heating after complete densification leads to grain growth, an increase in the average grain size. This too is a capillarity-driven process, because as the average grain size increases, the total grain-boundary area (and its associated energy) decreases. If the grain-boundary energy and mobility are not too anisotropic, then an equiaxed microstructure develops. In other words, each grain has approximately the same dimensions in all directions. An equiaxed microstructure is illustrated in **Figure 2**.

The shapes of the grains are determined in part by the relative rates at which the crystal grows in different directions (a kinetic effect), and in part by the requirement that the interfacial forces balance at their points of intersection (an equilibrium effect). The equilibrium equation for the point where three interfaces meet is known as the Herring condition:

$$\sum_{i=1-3} \gamma_i t_i + \frac{\partial \gamma_i}{\partial \beta_i} n_i = 0$$

The terms in this expression are described in **Figure 3**. Under the condition that the energy does not depend on orientation, this reduces to the more familiar Young equation

$$\frac{\gamma_1}{\sin \Psi_1} = \frac{\gamma_2}{\sin \Psi_2} = \frac{\gamma_3}{\sin \Psi_3}$$

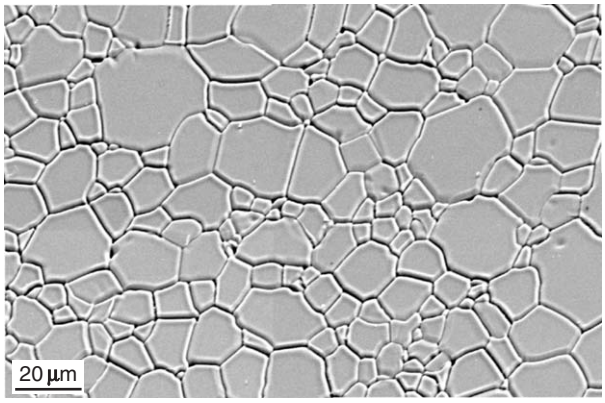


Figure 2 Montage of scanning electron microscope (SEM) images of spinel, $MgAl_2O_4$. The contrast at the grain boundaries results from thermal grooving.

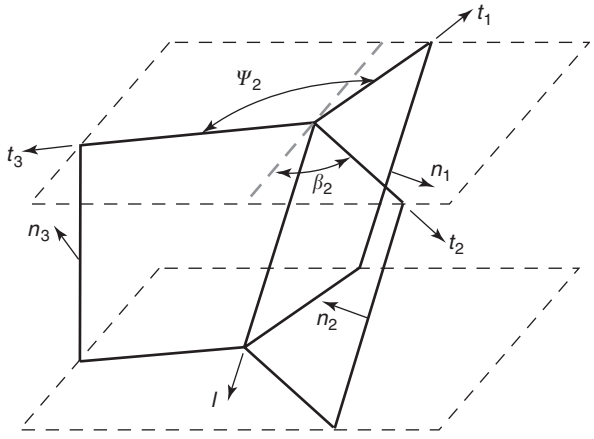


Figure 3 Schematic of the triple line (l) along which three interfaces meet, defining the terms in the Herring equation.

where the Ψ_i are the dihedral angles in **Figure 3**. If the energies of all the interfaces are assumed to be equal, then all of the dihedral angles are 120°. If this is the case, then dihedral angles at quadrajunctions where four triple lines intersect must be 109.5°. To meet these conditions and fill space, some curvature must be introduced in the boundaries and this creates a driving force for motion. Boundaries always move toward their center of curvature to reduce their total area. The situation was considered in two dimensions by von Neumann and Mullins, who showed that if equilibrium dihedral angles are to be maintained in an isotropic structure, then grains with more than six sides will always grow and grains with fewer than six sides will shrink (see **Figure 4**). While the problem is not easily solved in three dimensions, it is currently thought that grains with 13 or fewer faces shrink while those with 14 or more grow. Studies of three-dimensional shapes are consistent with this criterion. In a variety of studies, the average number of faces per grain has been measured to be

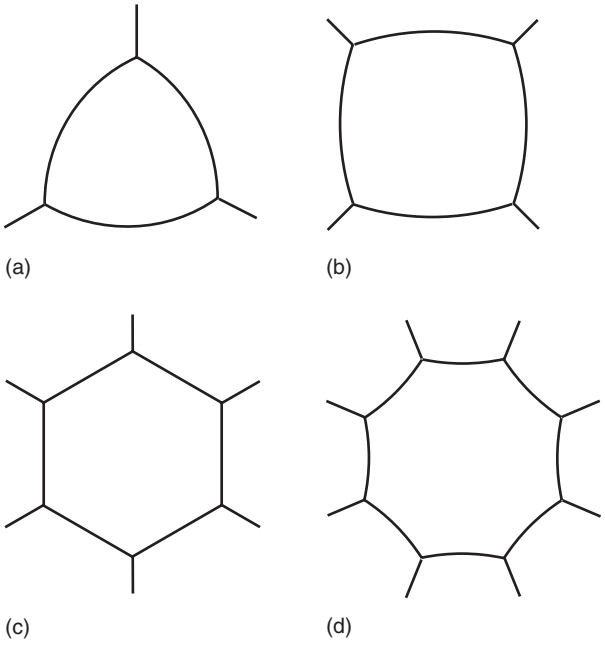


Figure 4 Schematic showing the configurations of boundaries in a two-dimensional system under the assumption that the grain boundary energy is isotropic. In this case, the boundaries must always meet with 120° dihedral angles. This means that if a grain has fewer than six sides, the boundaries are convex and will move inward ((a) and (b)), if it has exactly six sides there is no curvature and the boundaries are stationary (c) and if it has more than six sides (d) the boundaries are concave and move outward.

between 12 and 14. A sketch of a typical grain is shown in **Figure 5**.

Kinetic grain-growth theories generally predict that the average grain size increases with the square root of time (t) so that the average radius, $\langle r \rangle$, is

$$\langle r \rangle = (Kt)^n$$

where K is a positive constant and n, the grain-growth exponent, is ideally equal to 1/2. Furthermore, after the effects of the initial grain size distribution have disappeared, there is a steady-state distribution of grain sizes that is scale invariant. The form of the distribution, originally derived by Hillert, has a maximum grain size that is about 3/2 times the average size:

$$P_{\mathrm{H}}(\rho) = (2e)^3 \frac{3\rho}{(2-\rho)^5} \exp[-6/(2-\rho)]$$

In this equation, $\rho = r/r^*$, where r^* is the critical radius; grains smaller than r^* shrink while grains larger than r^* grow. The distribution is plotted in **Figure 6**. In practice, observed grain-growth exponents are usually less than 1/2, the maximum grain size greatly exceeds 3/2 times the average, and the distribution of

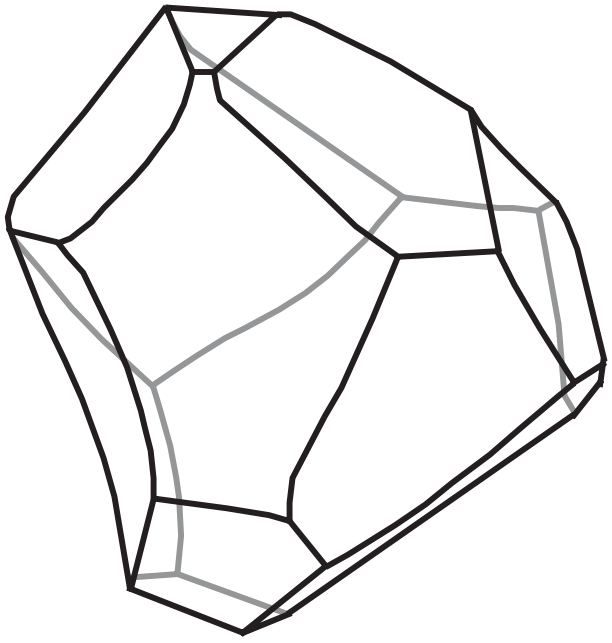


Figure 5 A schematic showing the three-dimensional shape of a grain within a polycrystal. The black lines are edges on the front of the grain and the gray lines are edges on the back of the grain.

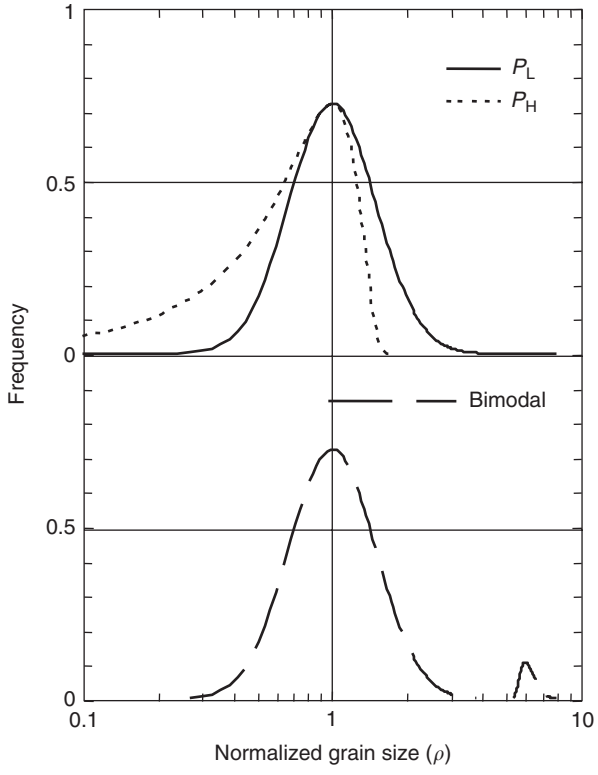


Figure 6 Grain size distributions. The most commonly observed distribution of grain sizes is lognormal (P_{L}) while the theory of capillary-driven grain growth predicts a different distribution (P_{H}). A schematic of a bimodal distribution is shown in the lower part of the figure.

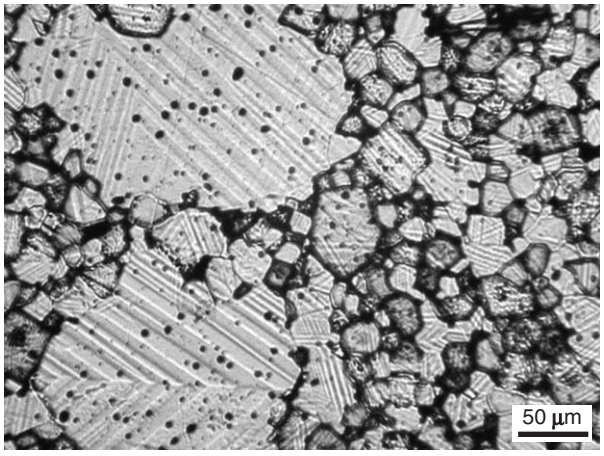


Figure 7 Visible light micrograph showing the microstructure of a polycrystalline $BaTiO_3$ specimen. Parts of two grains on the left are much larger than the cluster of average grains on the right. The regular patterns within each grain are twin plates that formed by a martinsitic transformation during cooling.

grain sizes is most frequently reported to be lognormal, where the number of grains with dimensionless size ρ is

$$P_{\mathrm{L}}(\rho) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{\rho r^*}\exp\left[\frac{-(\ln \rho)^2}{2\sigma^2}\right]$$

In this expression, σ is the width of the distribution. Microstructures that exhibit unimodal and roughly lognormal distributions of grain sizes are usually called normal. However, in some cases the distribution of grain sizes is bimodal (see **Figure 6** for comparison to normal distributions) and such distributions are said to be the result of abnormal growth. An example of a microstructure with an abnormal distribution of grain sizes is illustrated in **Figure 7**.

When the anisotropy in the grain-boundary energy and mobility is large, the grains adopt distinct shapes and are frequently bounded by flat interfaces, indicating that the orientation corresponds to a cusp or singularity in the function that describes the orientation dependence of the grain-boundary energy. For example, the microstructure in **Figure 8** shows a section through randomly oriented plate-like grains. Depending on the processing, the shapes may also exhibit preferred orientation or texture, as illustrated in **Figure 9**. In this case, the [0 0 0 1] axes of the rhombohedral lattice of Al_2O_3 are oriented in the vertical direction. The general phenomenon of nonrandom distributions of crystal orientations with respect to the external reference frame, as illustrated in **Figure 9**, is referred to as texture.

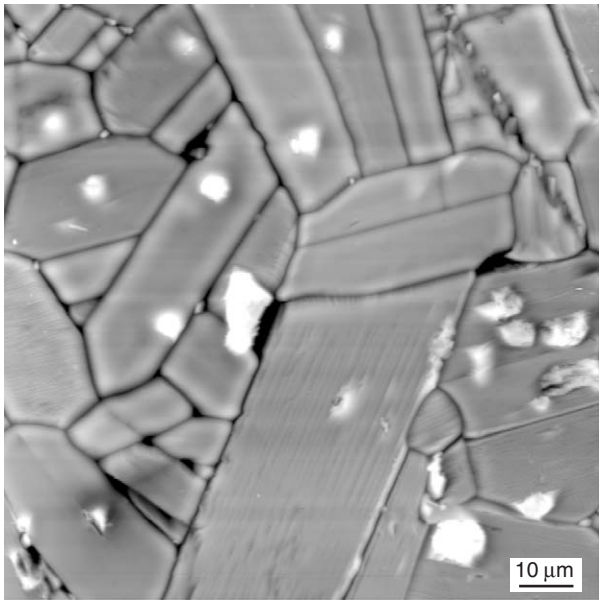


Figure 8 Atomic force microscope image of polycrystalline $Sr_2Nb_2O_7$. The orthorhombic material has platy grains elongated along the *a*- and *c*-axis.

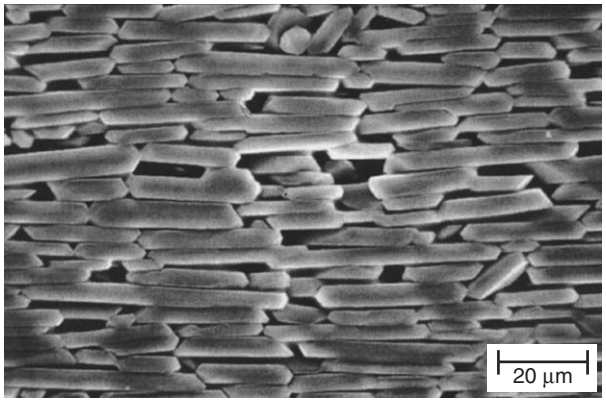


Figure 9 SEM micrograph of a tabular alumina microstructure formed by templated grain growth. The short dimension of the crystal is oriented along the [0 0 0 1] axis. (Micrograph courtesy of Gary L Messing and Matthew M Seabaugh, Department of Materials Science and Engineering, Pennsylvania State University.)

Single-Phase Microstructures by Crystallization

When a microstructure forms by crystallization from a liquid, gas, or amorphous matrix, crystal nucleation requires the formation of an interface between the crystalline and noncrystalline phase. Assuming that the interface energy (γ_{s}) is isotropic and that the change in free energy per volume upon crystallization is ΔG_{v}, then the free energy change for the formation of a nucleus with radius r is given by

$$\Delta G(r) = 4\pi r^2\gamma_{\mathrm{s}} - \tfrac{4}{3}\pi r^3\Delta G_{\mathrm{v}}$$

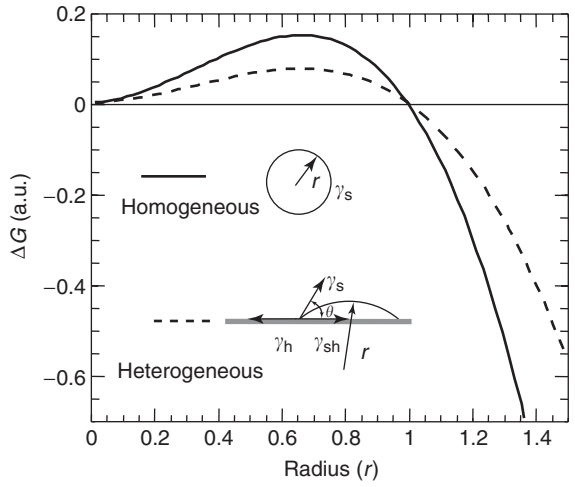


Figure 10 Plot of the free energy change as a function of nucleus radius. The change is initially positive. Additional growth reduces the energy only beyond a critical size. For heterogeneous nucleation, the barrier is lower.

As illustrated in **Figure 10**, the change in the free energy as a function r is initially positive. Only after a critical size is reached will further growth stabilize the crystal. Thus, there is an energy barrier, ΔG^*, that inhibits homogeneous nucleation. If a heterogeneous surface is present (e.g., the mould walls of a casting system or a substrate in a vapor deposition system), and there almost always is, then one not only has to consider the creation of the surface dividing the nucleus from the surrounding liquid or gas, but also has to consider the elimination of the heterogeneous surface and the creation of the interface between the nucleus and the heterogeneous surface. Using the Young equation, the relationship between the various interface energies can be expressed as a contact angle, θ, in the following way:

$$\cos\theta = \frac{(\gamma_h - \gamma_{sh})}{\gamma_s}$$

The terms in the equation above are defined in **Figure 10**, assuming the crystallization of a spherical cap on the heterogeneous surface. With this parameter, the free energy change for heterogeneous nucleation can be written in a form similar to that for homogeneous nucleation:

$$\Delta G(r) = \left[4\pi r^2\gamma_s - \tfrac{4}{3}\pi r^3 \Delta G_v\right] S(\theta)$$

where the shape factor is $S(\theta) = (2+\cos\theta)(1-\cos\theta)^2/4$. Since this term is always less than unity, the free energy change for heterogeneous nucleation is always less than that for homogeneous nucleation. Thus, crystal formation always initiates from a heterogeneity and most frequently from the walls of the container or a deliberately introduced substrate. This leads to so-called columnar microstructures, such as those illustrated in **Figures 11** and **12**. The microstructure in **Figure 11** results from nucleation and growth from the walls of a mould in a casting, and the microstructure in **Figure 12** results from the deposition of a vapor on a substrate. It should be recognized that heterogeneous nucleation can be initiated in a more or less uniform manner by introducing a finely dispersed heterogeneity and this can lead to an equiaxed microstructure. Uniform heterogeneous nucleation can also occur when a material is recrystallized from a deformed matrix.

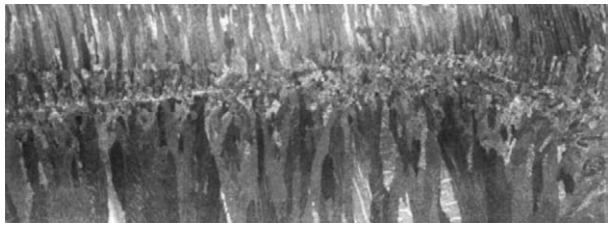


Figure 11 Section through a cylindrical casting. The elongated grains nucleated on the mould walls (top and bottom of the micrograph) and grown toward the center. (Micrograph courtesy of H P Paxton, Department of Materials Science and Engineering, Carnegie Mellon University.)

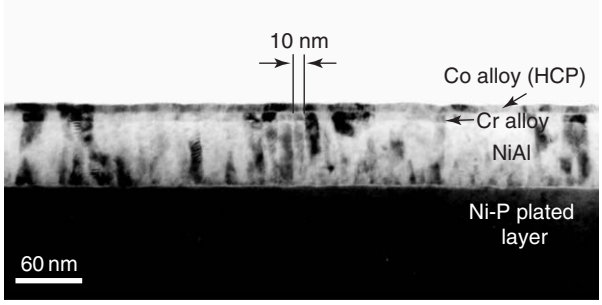


Figure 12 Columnar thin film multilayer microstructure. The NiAl grains nucleated on the Ni–P layer and the columnar structure is maintained through the Co and Cr layers. (Micrograph courtesy of David E Laughlin, Department of Materials Science and Engineering, Carnegie Mellon University.)

Grain Boundaries in Single-Phase Microstructures

The complexity of the grain-boundary networks within polycrystals arises from the number of crystallographic degrees of freedom. Three parameters are required to describe the lattice misorientation between two crystals, and two additional parameters are required to describe the orientation of

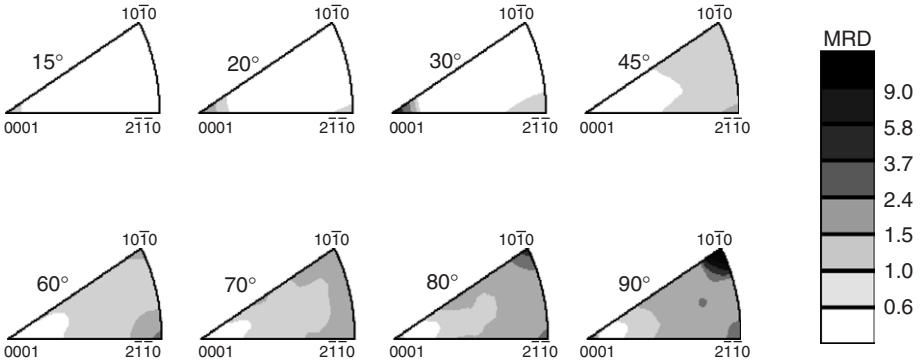


Figure 13 The grain-boundary character distribution in WC, plotted in axis-angle space. The entire range of distinguishable misorientation axes in the hexagonal system is represented in each triangle, and the individual triangles represent different rotations about these axes. The units are in multiples of a random distribution (MRD).

the grain-boundary plane. The number of distinguishable boundary types in this five-dimensional parameter space is large. For example, in a cubic material, if the five parameters are measured with 5° of resolution, then there are 5×10^4 distinguishable types of grain boundaries. In general, the distribution of grain-boundary surface area is not uniform over the five parameters. Textures have been measured both in the space of lattice misorientations and grain-boundary orientations. For example, **Figure 13** shows the distribution of lattice misorientations in WC, showing that boundaries with a 30° misorientation about [0 0 0 1] and a 90° misorientation about [1 0 $\bar{1}$ 0] occur with a higher and random frequency. These data are plotted without regard for the particular boundary plane. In **Figure 14**, the distribution of grain-boundary plane area is shown for the two highly populated misorientations identified in **Figure 13**. These data demonstrate that hexagonal WC has a preference for grain-boundary planes terminated on {0 0 0 1} and {1 0 $\bar{1}$ 0} surfaces.

While the mechanistic origin of the grain-boundary texture is not yet understood, it is known that the frequency with which different grain boundaries occur in a polycrystal is inversely related to the energy. Those boundaries with low energies or some other extraordinary property are thought to be especially influential in determining macroscopic properties. One such boundary is a twin. Twins in different materials have different crystallographic parameters, but the common feature is that they have a well-defined lattice misorientation and habit plane that allows the atoms within the interface to occupy positions that are very similar to those of the bulk. For example, in a cubic close-packed metal, a boundary formed between two crystals whose (1 1 1) planes are parallel, but rotated by 60° around the vector perpendicular to the boundary planes, is a twin. This particular boundary is only a minor disruption in the

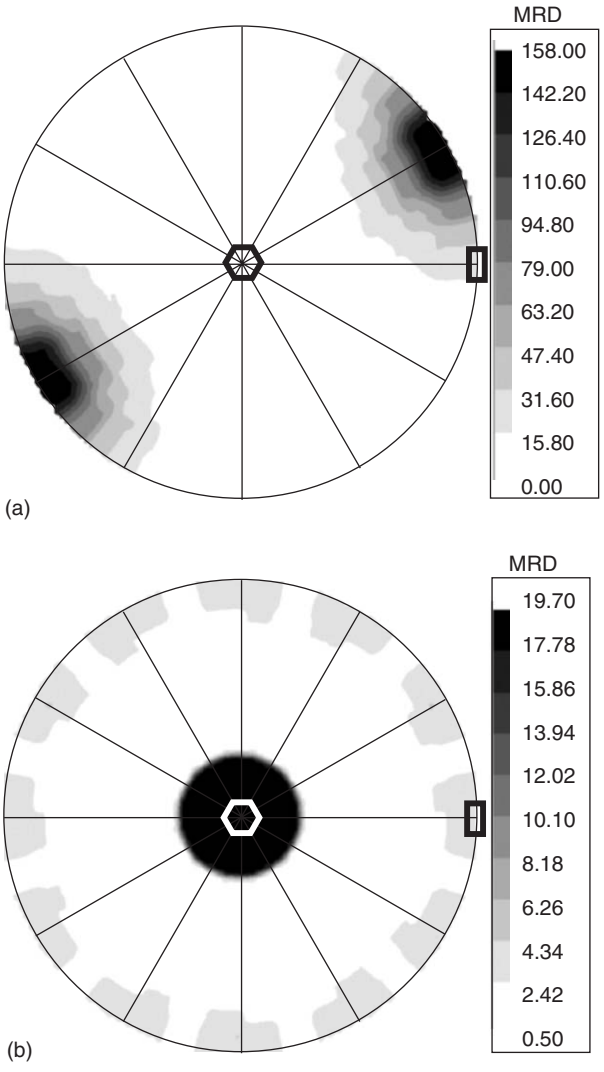


Figure 14 The distribution of grain-boundary planes at two specific misorientations: (a) for 90° rotations about the [1 0 $\bar{1}$ 0] axis and (b) 30° about the [0 0 0 1] axis. The hexagon marks the position of the [0 0 0 1] axis and the rectangle the position of the [1 0 0 0] axis. The positions of the peaks indicate that for both types of grain boundaries, the pure twist configuration is preferred.

long-range structure and is indistinguishable from a stacking fault. All of the atoms in the boundary plane still have 12 nearest neighbors in the same pattern; it is only the configuration of the next near neighbors that differs. This special condition occurs only on this one plane, so it is energetically very costly to rotate this boundary away from the (1 1 1) orientation. While the crystallography varies from system to system, this is a common feature and as a result, twins are very flat, planar boundaries without noticeable curvature. Because a twin has a low energy, the dihedral angle it makes with other more general grain boundaries is usually very close to 180°. These are the geometric characteristics that make twins easy to locate in a microstructure.

In cubic close-packed metals, twins commonly form during recrystallization. An example of a highly twined alpha brass microstructure is illustrated in **Figure 15**. In tetragonal materials, if there is only a small difference between the *a* and *c* lattice parameters, then mosaic twinning is common. A twinned $BaTiO_3$ structure is illustrated in **Figure 16** (this is a higher-resolution image of the twins that were evident in **Figure 7**). This material has a tetragonally distorted perovskite structure with a *c* to *a* ratio of 1.01. The straight boundaries all correspond to 90° rotations about a ⟨1 0 0⟩ axis with the boundary on a {1 1 0} plane.

Boundaries with misorientations that place a high fraction of lattice sites in coincidence are also distinguished from more general boundaries. These boundaries are called coincident site lattice boundaries and are assigned a coincidence number (Σ) based on the inverse of the number of coincident lattice sites. Therefore, a low Σ number signifies high coincidence. For example, illustrations of the Σ5 and Σ13 boundaries are presented in **Figure 17**. In many circumstances, boundaries of high coincidence are also referred to as "special" boundaries, although the term is not always used in a consistent manner. Note that while the coincidence relation depends only on the lattice misorientation, high coincidence is only achieved if the boundary lies on certain planes where the lattice sites overlap. Therefore, of all boundaries with a low Σ misorientation, only those with the interface planes oriented in a specific direction are expected to have special properties. The twin in the cubic close-packed system is a good example. This corresponds to a boundary with a 60° misorientation about the [1 1 1] axis, with the boundary on a (1 1 1) plane. The designation Σ3 refers to the same misorientation, but does not specify the boundary plane.

Boundaries with low Σ misorientations have recently taken on a special significance because it has

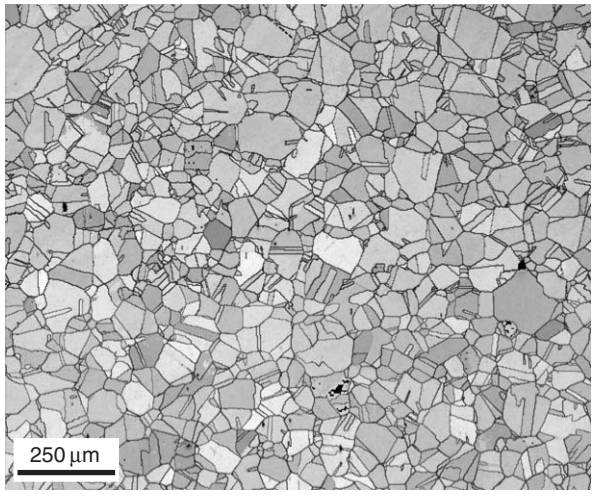


Figure 15 Image quality electron backscattered pattern map of alpha brass. The black lines indicate the positions of grain boundaries. The straight boundaries are twins. (Micrograph courtesy of Valerie Randle, University of Wales, Swansea.)

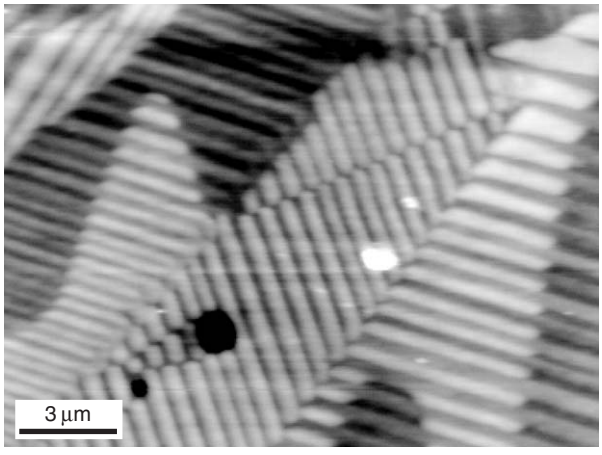


Figure 16 AFM image showing mosaic twinning in $BaTiO_3$. The sample was etched to create topographic contrast at the twin boundaries.

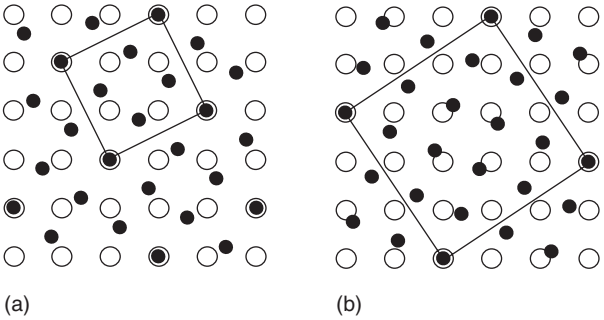


Figure 17 Schematic illustration of the (a) Σ5 and (b) Σ13 coincident site lattice geometries. The black points belong to one lattice and the open circles to the other. The squares denote the repeat unit.

been found that certain material properties are improved by increasing the fraction of these low Σ boundaries, particularly those known as $\Sigma 3^n$, including Σ3, Σ9, and Σ27. In cubic close-packed materials with low stacking fault energies, such as Cu, Ni, and Pb, the fractional population of $\Sigma 3^n$ boundaries can be increased by the process referred to as grain-boundary engineering. In this process, the materials undergo repeated cycles of mild deformation followed by recrystallization. After each cycle, the concentration of $\Sigma 3^n$ boundaries increases. The micrograph in **Figure 15**, discussed above, shows an example of a grain-boundary engineered microstructure in which approximately 45% of the boundaries are of the type $\Sigma 3^n$. As the special boundaries increase in concentration, they break up percolating clusters of general boundaries that are more susceptible to failure.

Duplex Microstructures

Duplex microstructures contain two distinct phases, which are materials that can be distinguished on the basis of structure or composition. In principle, a microstructure may contain many different phases, but the discussion here is limited to the case of two phases. The defining elements of the two-phase microstructure are the same as in the single-phase microstructure, only it is also necessary to characterize the configuration of the two phases and the volume fractions. As in the single-phase microstructures, the configuration is determined both by equilibrium phenomena and kinetic phenomena. Because of the added complexity, there is a wider range of possibilities. Selected cases are discussed in the following sections.

Duplex Microstructures Formed by Consolidation

The most direct method to create a duplex microstructure is to physically mix the two phases and then consolidate the mixture by sintering. The degree to which the two phases wet each other strongly influences the configuration. Referring once again to the Young equation, the equilibrium geometry at the points where the phases meet is a function of the interfacial energies. If one labels the phases A and B, and assumes that their interface energies are isotropic, then the dihedral angle between the two AB interfaces (see **Figure 18**) is given by

$$\cos(\Psi) = \gamma_{AA}/2\gamma_{AB}$$

Note that when $2\gamma_{AB} \ll \gamma_{AA}$, the dihedral angle goes to zero and in this case the two phases completely

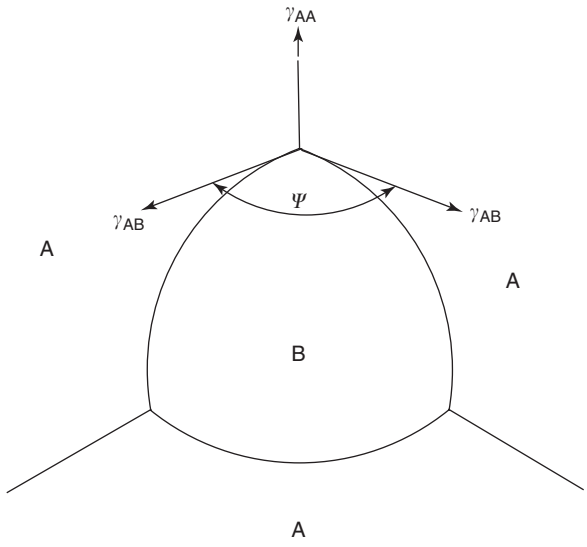


Figure 18 Schematic of the interfacial equilibrium at the point where phase B meets three grains of phase A. The dihedral angle, Ψ, is a measure of the relative interfacial energies.

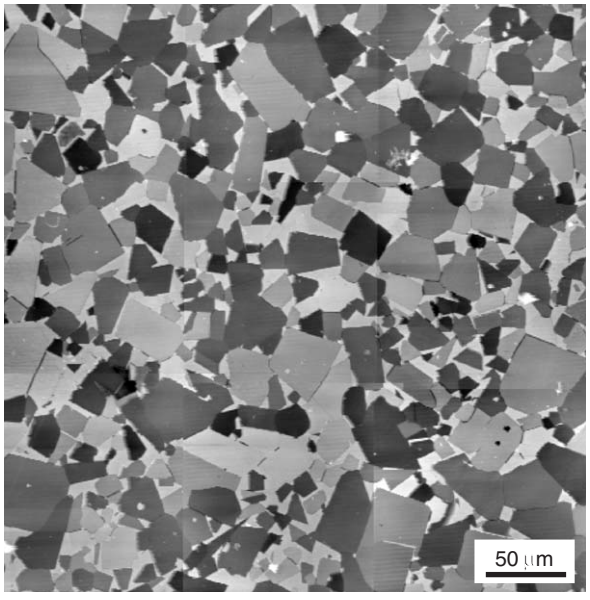


Figure 19 Montage of topographic AFM images of a WC/Co surface. The surface was etched in such a way that the WC crystals were attacked preferentially, and they appear as the darker contrast. The Co matrix has the lightest contrast.

wet each other. As γ_{AB} approached γ_{AA}, the two phases become increasingly nonwetting and the dihedral angle increases toward 120°. Two examples are shown in **Figures 19** and **20**. In **Figure 19**, Co, which is liquid at the processing temperature, completely wets the angular WC crystals. In this case, the WC crystals occupy 90% of the volume and form an interconnected, skeletal network. The aluminum

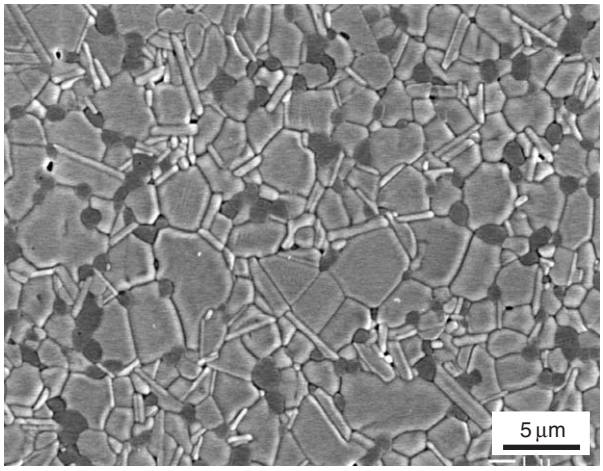


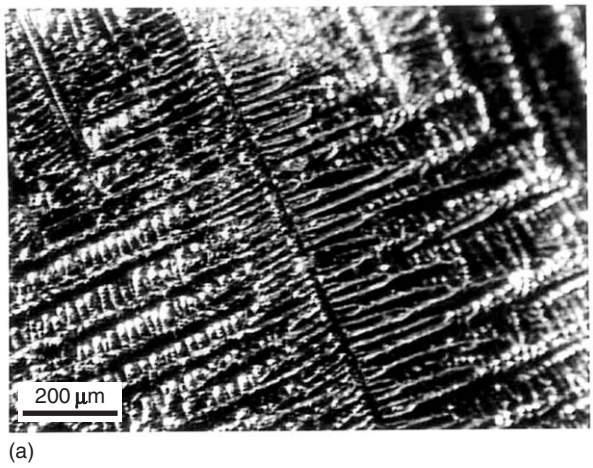
Figure 20 SEM image of Ta_2O_5 doped with 3 wt.% Al_2O_3 and 1.5 wt.% La_2O_3. The dark grains are Al-rich and the anisotropic grains are a La-rich phase. (Micrograph courtesy of Suxing Wu, Helen M Chan, and Martin P Harmer, Department of Materials Science and Engineering, Lehigh University.)

oxide–tanatalum oxide mixture in **Figure 20** provides an example of a nonwetting combination. The dark grains are the Al-rich phase which poorly wets the Ta_2O_5 boundaries.

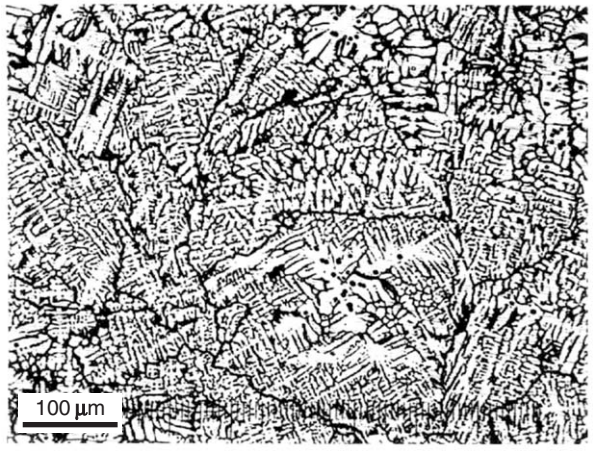
Duplex Microstructures by Solidification

When microstructures form by the solidification of a homogeneous liquid, the phases are not as uniformly dispersed as when formed by a physical mixture followed by consolidation. Interestingly, well-defined patterns can develop under these conditions. When a single-phase liquid is crystallized in a two-phase region where a solid and liquid phase coexist, one of the most common morphological patterns to develop is the dendrite, or branched structure, as illustrated by the microstructures in **Figures 21** and **22**. It should be noted that while dendrites can also form during the solidification of a pure material into a supercooled liquid, this situation is far less common than that of the binary alloy. For the single-phase case, one can imagine that protrusions in the growth front can be stabilized if the liquid is supercooled and the latent heat of crystallization is extracted through the solid. In this case, heat is removed more efficiently from the tip of the protrusion that juts out into the supercooled fluid than it is from the valleys between protrusions.

To describe the mechanism by which dendrites form in a binary alloy, consider the simple binary eutectic phase diagram illustrated in **Figure 23**. For an alloy with composition C at a temperature less than T_L, there is a range of temperatures where the solid and liquid coexist, but with different compositions.



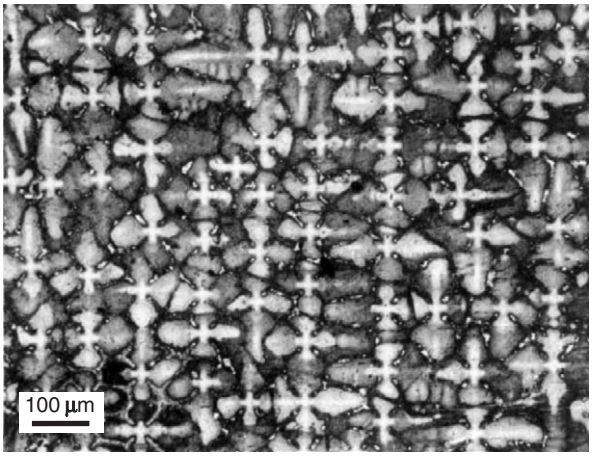
(a)



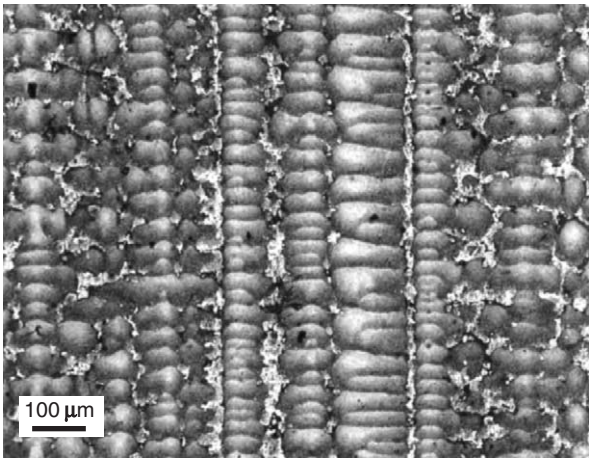
(b)

Figure 21 (a) A dendrite observed on the surface of an Al-4.5 wt.% Cu casting. (b) A section through the same casting, illustrating the appearance of random sections through the network of Al dendrites. The dark contrast is the Cu-rich phase.

The solid alloy α has composition C_α and always contains less B than the liquid, with composition C_L. In fact, one can say that the solid is always relatively more pure than the liquid from which it crystallizes. This means that as the crystal grows, the liquid at the interface is relatively enriched in B and that the enriched liquid freezes at a lower temperature, specified by the liquidus line in **Figure 23**. Therefore, a protrusion in the interface can be stabilized because it juts into a relatively more pure liquid (see **Figure 24**). In the valleys between protrusions, on the other hand, the liquid rich in B is trapped between opposing growth fronts. Thus, the growth front at the tip of the protrusion is in contact with a liquid that freezes at a higher temperature than the liquid in the valleys between the protrusions. This causes the tip to grow faster than the lateral surfaces and the result is cellular or, more commonly, a dendritic morphology.



(a)



(b)

Figure 22 (a) Transverse and (b) lateral sections through a directionally solidified Ni alloy. The transverse section shows the shape of the primary dendrite while the lateral section shows the arm spacing.

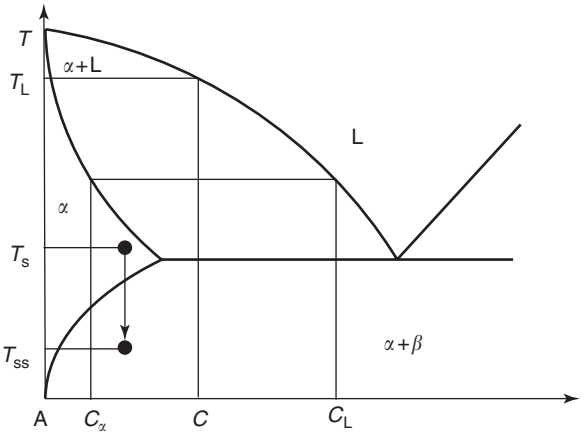
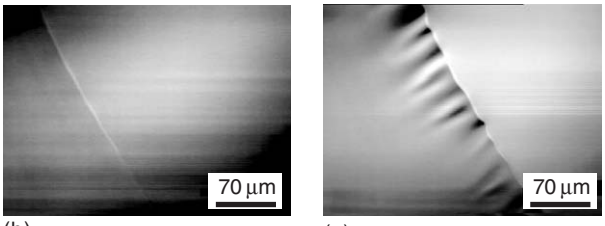


Figure 23 Schematic binary alloy phase diagram.

(a)

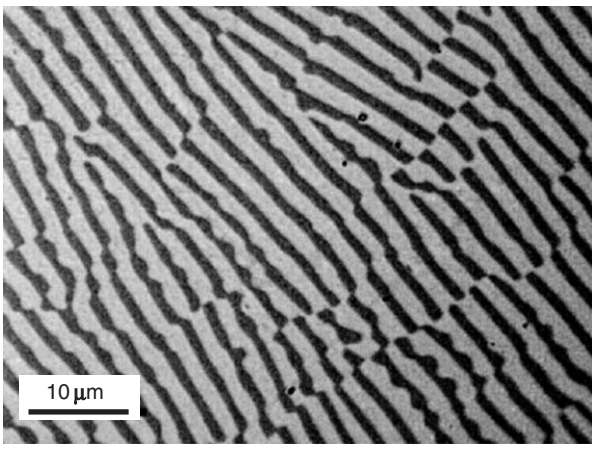


(b) (c)

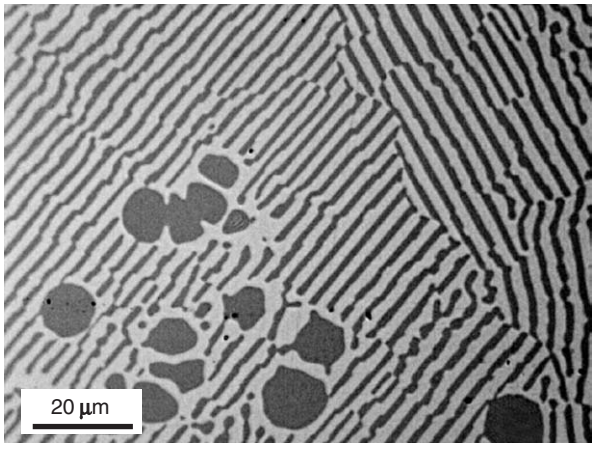
Figure 24 (a) Schematic of interface instability by constitutional supercooling. As solute is rejected from the solidifying crystal, it accumulates in the valleys between the dendritic protrusions (the solute concentration is indicated by the shading). Since the relatively more pure liquid at the ends of the protrusions freezes at a higher temperature, these portions of the solid grow faster. ((b) and (c)) *In situ* laser confocal microscope image of a solidification front in Fe–Co–Ni alloy at 1597°C. (b) The front is planar (c) a few seconds later, an instability develops. (Micrograph courtesy of Chanjoon Paek and Sridhar Seetharaman, Department of Materials Science and Engineering, Carnegie Mellon University.)

When the remaining liquid cools to the temperature T_E, two solid phases (α and β) crystallize directly from the liquid. This most frequently results in a lamellar microstructure, such as those shown in **Figure 25**, where the two phases have alternating bands of dark and light contrast. Because both solids are in equilibrium with the same liquid, but themselves have very different compositions, they must simultaneously reject solute to feed the growth of the other phase. In other words, as a layer of the α phase grows, it must reject solute (B) laterally to supply neighboring lamella of the β phase. As the β phase crystallizes, it rejects A, which sustains the growth of the α phase. This cooperative process, schematically illustrated in **Figure 26**, stabilizes a nearly planar growth front.

In most solidification processes, cooling will take the materials through a region of solid–liquid coexistence, in which a dendritic structure will grow. When the remaining liquid reaches the eutectic temperature, it will solidify in a lamellar microstructure, leading to a mixed microstructure. In general, this situation is not an equilibrium configuration. Annealing at a temperature just below the eutectic will cause the compositional concentration gradients to homogenize and the morphologies of the crystals to spheroidize and become more equiaxed.



(a)



(b)

Figure 25 Light microscope images of the microstructure of directionally solidified (a) CoO–ZrO_2 and (b) NiO–ZrO_2 eutectic microstructures. The darker phase is ZrO_2 and the lighter phase is CoO in (a) and NiO in (b). (Micrographs courtesy of Nasim Alem and Vinayak P Dravid, Department of Materials Science and Engineering, Northwestern University.)

Duplex Microstructures by Transformation in the Solid Phase

Duplex microstructures can also form by transformation in the solid phase. As an example, consider the rapid cooling of α on the phase diagram in **Figure 23**, from T_s to T_{ss}. At T_{ss}, the supersaturated α phase will decompose to $\alpha + \beta$, and this will happen entirely in the solid state. As with solidification, the characteristics of the microstructure that develop are determined by the nucleation and growth of the phase. The free energy change during the nucleation process is similar to that for heterogeneous nucleation, except that an additional term must be added to account for the volumetric strain energy associated with any expansion or contraction that might accompany the transformation. The disposition and

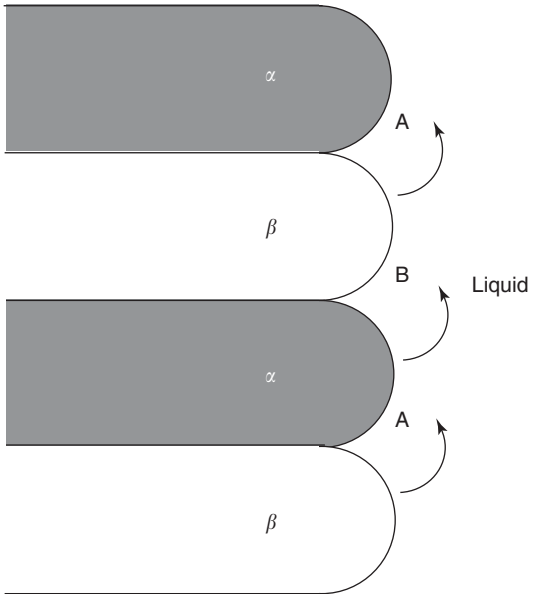


Figure 26 Schematic of the growth of a lamellar microstructure maintaining a nearly planar growth front.

shape of this phase depends on the structural relationship between the parent and daughter phases, features pre-existing in the microstructure that serve as heterogeneous nucleation sites, the degree of supersaturation during the nucleation and growth of the phase, and the mechanism of the transformation. Because of these parameters, there are a wide range of possible microstructures that can form and only a few of the characteristic examples are described in the remainder of this article.

Precipitation from a supersaturated solid solution requires the long-range diffusion of solute to the site of growth. For this reason, the distribution of a precipitated phase will depend on the degree of undercooling or supersaturation. When large, the driving force for the precipitation reaction is amplified and atomic mobility is limited. This leads to a uniform distribution of very small precipitates. For example, the microstructure in **Figure 27** shows cuboidal γ' precipitates in an $Ni_{70}Cr_{20}Al_{10}$ superalloy. In this case, the structure of precipitating phase differs little from that of the parent phase, so there is an orientation relationship between the two. In cases where the kinetics for transport are higher and the supersaturation is not as large, precipitates will nucleate at heterogeneities such as grain boundaries.

The formation of lamellar microstructures from the solidification of a eutectic has already been described. In the case of the eutectoid reaction, where a single solid phase decomposes to two new solid phases, the microstructure that is formed is similar to that formed by the eutectic reaction, but the lamellar phase typically nucleates and grows from the grain

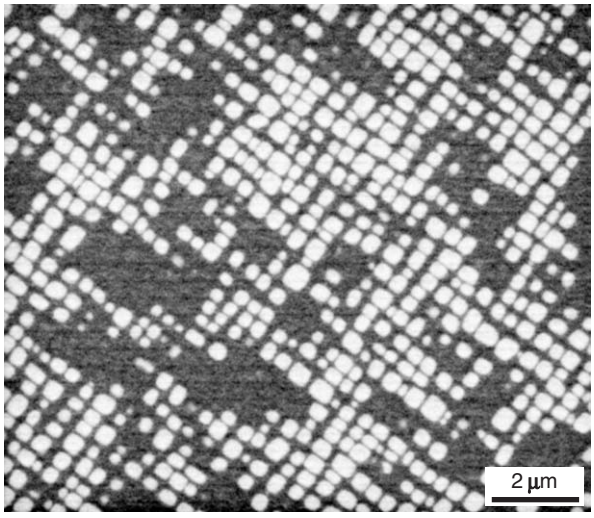


Figure 27 Focused ion beam microscope image of cuboidal γ' precipitates in a $Ni_{70}Cr_{20}Al_{10}$ superalloy. (Micrograph courtesy of Dr. Michael Uchic, Wright Patterson Air Force Base, and Marc De Graef, Department of Materials Science and Engineering, Carnegie Mellon University.)

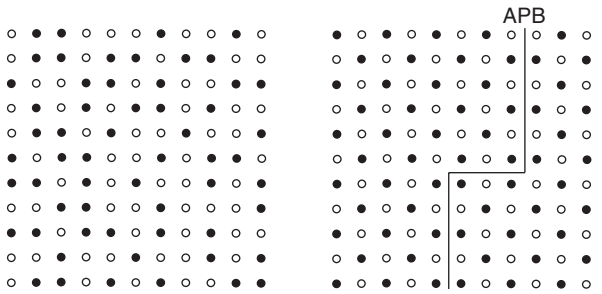


Figure 28 Schematic image illustrating the difference between (a) a disordered AB alloy and (b) ordered arrangement with an antiphase domain boundary.

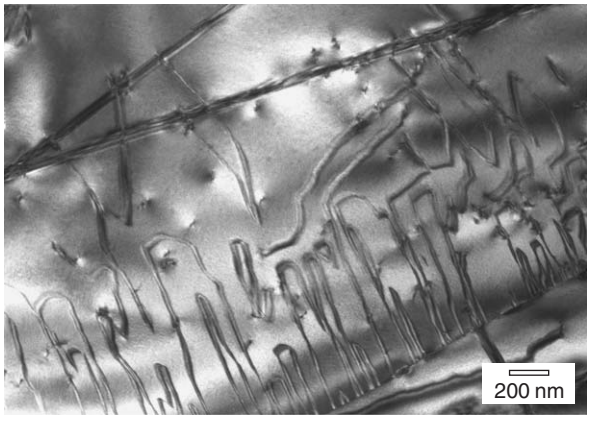


Figure 29 The transmission electron micrograph of Fe-55 at.%Pd alloy annealed at 400°C for 140 h. The micrograph shows the antiphase domains and antiphase boundaries (APBs) formed during the atomic ordering process. (Micrograph courtesy of Lisha Wang and David E Laughlin, Department of Materials Science and Engineering, Carnegie Mellon University.)

boundaries. Similar microstructures can also be generated by the cellular transformation from a supersaturated solid solution.

There are also solid to solid transformations that do not require long-range diffusion and these processes also lead to distinct microstructures. In the massive transformation, a single phase decomposes to one or more phases that have the same composition, but different crystal structures. Because it is not necessary to alter the composition, the interface can move very rapidly. The massive transformation frequently nucleates at grain boundaries and the rapid interface motion leaves an irregularly shaped precipitate. Ordering transformations can also occur with very little long-range diffusion. In this case, elements A and B originally occupy certain sites in the crystal structure at random. Below a certain temperature, they occupy the same sites, but in an ordered fashion, as illustrated schematically in **Figure 28**. Because the ordering process will begin in many places at once, impingement will create antiphase domain boundaries, where there is a shift in the ordering pattern. A characteristic microstructure with APBs is illustrated in **Figure 29**.

Martensitic transformations are diffusionless processes by which a material changes crystal structure without requiring any of the atoms to change position by more than a lattice spacing. Typically, the parent and daughter structure are related by a shear deformation process. In the resulting microstructure, lens-shaped precipitates grow quickly to the size of the grains of the parent phase, creating the characteristic microstructure, such as the one illustrated in **Figure 7**, where the twins within the grains are formed by a martensitic transformation. It has been shown that the interface between the precipitate and the parent phase can move at velocities approaching the speed of sound.

See also: Deformation Processing; Diffusionless Transformations; Mass Transport; Meso- and Nanostructures; Orientation Texture; Periodicity and Lattices; Phases and Phase Equilibrium; Phase Transformation; Powder Processing: Models and Simulations; Recovery, Recrystallization, and Grain Growth; Solidification: Models and Simulations; Thermodynamic Properties, General.

PACS: 61.72. – y; 61.72.Mm

Further Reading

ASM (1985) Metallography and microstructures. *Metals Handbook*, 9th edn. Metals Park, OH: American Society of Metals.

Bunge H-J (1982) *Texture Analysis in Materials Science*. London, UK: Butterworths (Translated by Morris PR).

Burton WK, Cabrera N, and Frank FC (1951) The growth of crystals and the equilibrium structure of their surfaces. *Philosophical Transactions of the Royal Society of London, Series A* 243: 300–358.

Herring C (1951) Surface tension as a motivation for sintering. In: Kingston WE (ed.) *The Physics of Powder Metallurgy*, pp. 143–179. New York: McGraw-Hill.

Hillert M (1965) On the theory of normal and abnormal grain growth. *Acta Metallurgica* 13: 227–238.

Mullins WW (1956) Two-dimensional motion of idealized grain boundaries. *Journal of Applied Physics* 27: 900–904.

Mullins WW and Sekerka RF (1964) Stability of a planar interface during solidification of a dilute binary alloy. *Journal of Applied Physics* 35: 444–451.

Porter DA and Easterling KE (1981) *Phase Transformations in Metals and Alloys*. London: Chapman Hall.

Randle (1996) *The Role of the Coincident Site Lattice in Grain Boundary Engineering*. London: Institute of Materials.

Rohrer GS, Saylor DM, El-Dasher BS, Adams BL, Rollett AD, and Wynblatt P (2004) The distribution of internal interfaces in polycrystals. *Zeitschrift fur Metallkunde* 95: 1–18.

Smith CS (1952) The shapes of metal grains, with some other metallurgical applications of topology. *Metal Interfaces*, p. 65. Cleveland: ASM.

Smith CS (1964) Some elementary principles of polycrystalline microstructure. *Metallurgical Reviews* 9: 1–48.

Smith CS (1981) *A Search for Structure*. Cambridge, MA: MIT Press.

Tiller WA (1991) *The Science of Crystallization: Macroscopic Phenomena and Defect Generation*. Cambridge: Cambridge University Press.

Microelectromechanical Systems

M Dragoman, National Institute for Research and Development in Microtechnology, Bucharest, Romania
D Dragoman, University of Bucharest, Bucharest, Romania

Introduction

The miniaturization, up to the submicrometer scale, of electronic and optoelectronic components grown on semiconductor substrates generated the computer and communication revolutions at the end of the twentieth century. The miniaturization of electronic components accompanied by the increase in their performances was successful due to a combination of advanced semiconductor technologies such as MBE (molecular beam epitaxy), VLSI (very large scale integration), and due to the employment of new types of semiconductors and heterostructures. However, the main material, which can be found in any computer or communication system, is still silicon.

A second revolution is born, thanks to the microelectromechanical systems (MEMSs), which are three-dimensional (3D) chips that are able to change information with the environment, and which are produced by the simultaneous miniaturization of electronic and mechanical devices. In contrast to electronic or mechanical components, which can be separately designed, fabricated, and eventually assembled in a remote place, the design of MEMSs and their fabrication is an integral process, which cannot be separated in elementary steps. MEMSs have thus the status of a system, the MEMS revolution being characterized by the miniaturization of systems and not of components.

The development of MEMS was dictated by the need for high-speed data processing in global communication systems, which require rapid reconfiguration of various communication routes. Also, the MEMS advance was boosted by the need for integration of sensing and monitoring devices in various applications that range from environment monitoring to medical surgery. Therefore, the MEMS chip is often called "smart" since it can sense the chemical, mechanical, physical, electrical, optical, etc. modifications of the environment and can exploit them in order to perform a specific task. The abilities of the MEMS chip to act and decide are expected to generate astonishing applications in electronics, communications, or biology.

The reduction of mechanical devices at the micrometer or even nanometer scale is implemented using semiconductor materials with excellent mechanical properties, such as Si, and a common semiconductor technology that exploits the already existing technological facilities. Thus, MEMSs are realized with almost the same tools as in the VLSI technology, on semiconductor substrates on which VLSI circuits can also be realized. MEMSs are not always realized using Si, other semiconductors such as GaAs, GaN, SiC, InP being used as well. Many applications also require MEMS made from metals such as gold or silver.

Although quite a large variety of micromechanical or nanomechanical devices exist, the vast majority of MEMS is implemented with only a few mechanical configurations such as cantilevers, microbridges, or membranes. In contrast to common electronic devices, MEMSs contain almost always moving parts (actuators, microrobots, switches, etc.) that are set in motion by electric, magnetic, thermal, or electromagnetic forces exerted between different parts of a MEMS.

Static MEMS (with no moving parts) can also be encountered and fabricated with specific MEMS technologies that are described below. These micro- or nanoscale mechanical devices integrated with various electronic components have succeeded in transforming the sensing, computing, microwave communications, optical communications, and optical processing at the global scale in only a couple of years due to their small sizes, low power consumption, and low costs.

Electronic or mechanical devices with at least one dimension of a few nanometers are the building blocks of nanoelectronics. The nanoelectromechanical systems (NEMSs) are nanosized systems that integrate mechanical and electronic devices in a single assembly. However, NEMSs are not simple scaled-down versions of MEMSs; they involve new concepts and design techniques, which take into account the new physical effects encountered only at the nanoscale (e.g., van der Waals forces). Also, the technological implementation of NEMSs is no longer based on standard silicon processing, but on a set of technologies specific to the nanoscale and generically termed nanotechnologies.

Main Fabrication Techniques of MEMS

Semiconductor micromachining is the main technology for MEMS implementation. Any micromachining technique removes a part of a bulk substrate or of a thin film using various etching methods. Silicon, grown either as a monocrystal or as a polycrystal, is the familiar material for micromachining, performed in this case with standard techniques for silicon processing. As can be seen from **Table 1**, Si has also very good mechanical properties, which are of paramount importance for the moving parts in an MEMS.

Two types of etching are generally used in micromachining: wet or chemical, and dry (or plasma) etching, which generate different etching forms. For example, the etching forms in **Figures 1a** and **1b** can be obtained with isotropic and anisotropic wet etching, respectively, whereas the form in **Figure 1c** is attainable with plasma etching. Anisotropic etching and plasma etching are used for Si, GaAs, or InP. In any micromachining technique, it is essential to achieve controllable, high-precision, and low-roughness shapes. Almost perfect vertical walls can only be obtained using plasma etching methods, but more complex and expensive methods such as laser etching or synchrotron radiation (called LIGA) must be employed to carve in semiconductors almost perfect shapes with low roughness.

Micromachining techniques can be divided into two classes. The first one concerns bulk micromachining, where the back of a wafer is etched to obtain a prescribed mechanical structure on its top. This technique is typical for fabrication of micromechanical membranes, as displayed in **Figure 2**, very large membranes with surfaces of $2 \times 2\,\text{cm}^2$ and thicknesses of 1 μm being achieved using chemical or plasma-etching methods.

The second class, referred to as surface micromachining, allows the fabrication of MEMS or NEMS by successive thin film depositions on a common substrate, followed by the selective etching of one of these thin deposited layers, called the sacrificial layer. The technological realization of a cantilever using surface micromachining is displayed in **Figure 3**.

Table 1 Mechanical properties of some basic materials for MEMS

Material	*Young's modulus* ($\times 10^{11}$ Pa)	*Density* (g cm^{-3})	*Thermal expansion* ($\times 10^{-6} \circ \text{C}^{-1}$)
Silicon	1.5	2.4	2.5
Carbon nanotube	10	1.4	10
Steel	2.1	8	12

Figure 1 Semiconductor shapes obtained with diverse etching techniques. (a) Isotropic etching; (b) Anisotropic etching – dependent on orientation; (c) Anisotropic etching – independent of orientation.

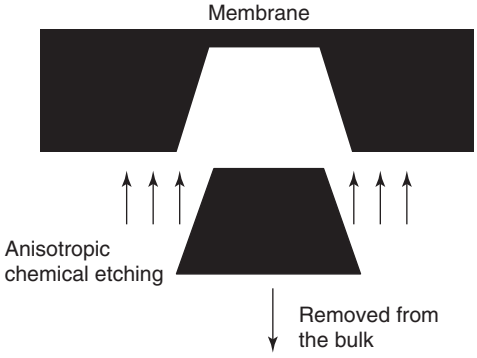


Figure 2 Membrane realization via bulk micromachining.

Basic Micromechanical Devices

The Cantilever

A basic micromechanical component in many MEMSs is the cantilever, represented in **Figure 4**,

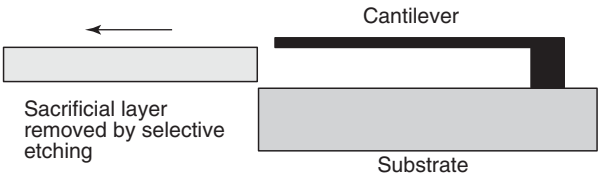


Figure 3 Cantilever realization via surface micromachining.

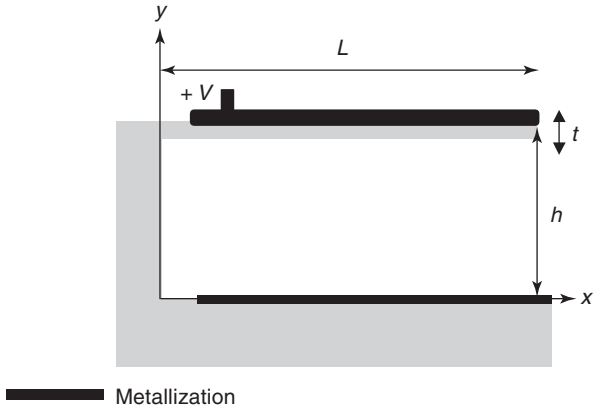


Figure 4 Electrostatic actuation of a cantilever with rectangular cross section.

which is frequently integrated with electronic and/or optoelectronic components for various purposes. Semiconductor cantilevers are mostly fabricated from Si, although very thin GaAs cantilevers were also produced via micromachining technologies. Metallic cantilevers are used mainly as adsorbers of chemical or biological substances, which produce an additional bending that can be detected in sensing applications. Typical micronic cantilevers have lengths of tens of microns and widths and thicknesses of a few microns. Nanosized cantilevers, also called nanocantilevers, have widths and thicknesses of a few nanometers and lengths of a few microns.

An electric, magnetic, or electromagnetic force bends the cantilever, the corresponding actuation being called electrostatic (toward a bottom electrode), magnetic, or optical. Heat can also actuate a cantilever. The most widespread actuation type is electrostatic (see **Figure 4**), in which attractive electrical forces are generated when a DC voltage is applied between the cantilever and the metal deposited on the substrate (the substrate electrode), the cantilever being attracted toward the substrate.

The electrical force $F(y)$ that actuates a cantilever with width W, length L, and thickness t, fabricated from a material with permittivity ε is given by

$$F(y) = -\varepsilon_0 WL \frac{V^2}{2[y + (t/\varepsilon)]^2}$$

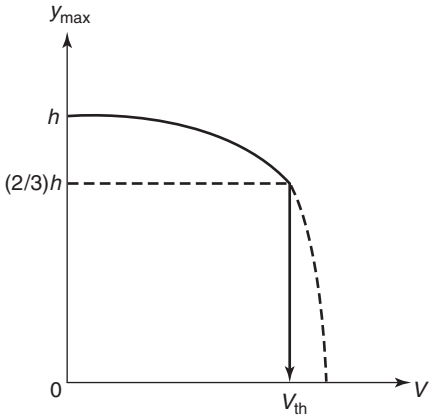


Figure 5 The dependence on voltage of the cantilever deflection.

where ε_0 is the free-space permittivity and V is the voltage applied between the cantilever and the substrate electrode. The bending of the cantilever induced by the applied actuation force can be described by the equation

$$d^2y/dx^2 = M(x)/EI$$

where I is the moment of inertia, equal to $Wt^3/12$ for rectangular cantilevers, E is the Young's modulus, and $M(x)$ is the total moment at position x, completely determined by the electric force. The maximum deflection at $y = L$ is then given by

$$y_{max} = y(L) = h - V^2 \frac{6\varepsilon_0}{Et^3[h + (t/\varepsilon)]^2} \times \left(\frac{L^2x^2}{4} - \frac{Lx^3}{6} + \frac{x^4}{24}\right)$$

The deflection of the cantilever is usually performed with optical systems or electronic detection techniques, which can be either capacitive or based on the tunneling effect. The deflection of the cantilever as a function of the applied voltage is displayed in **Figure 5**.

The dynamical equilibrium of the cantilever is ensured by the balance between the electrostatic and elastic forces, which require that

$$\frac{\varepsilon_0 WLV^2}{2[y + (t/\varepsilon)]^2} - Ky = 0$$

where the spring constant of a cantilever beam is $K = 3EI/L^3 = Et^3W/4L^3$. From **Figure 5** it follows that this equilibrium is maintained up to deflections of about one-third of the cantilever substrate height h, after which the equilibrium between the electrostatic and elastic forces is destroyed and the

cantilever collapses on the substrate electrode. The threshold voltage beyond which the cantilever collapses is given by:

$$V_{th} = \sqrt{8Kh^3/27WL\varepsilon_0}$$

The above formula dramatically changes for nanocantilevers, with nanometric h values, because in this case the equilibrium condition must take into account the van der Waals force that acts only at the nanoscale. The attractive van der Waals force is so strong that it can even pull down the nanocantilever in the absence of a DC voltage. The major influence of the van der Waals force at the nanoscale can be seen from **Figure 6**.

Although the effect of van der Waals forces can be lowered in a double-clamped NEMS, other amazing behaviors associated with the nanoscale occur. In particular, the double-clamped NEMS shows a quantized mechanical movement, the deflection varying in discrete steps every time an electron tunnels through the metallic contacts of the nanobeam (see **Figure 7**).

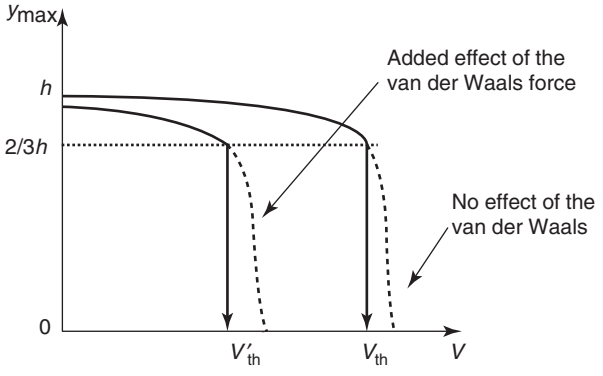


Figure 6 The van der Waals effect on the deflection of nanocantilevers.

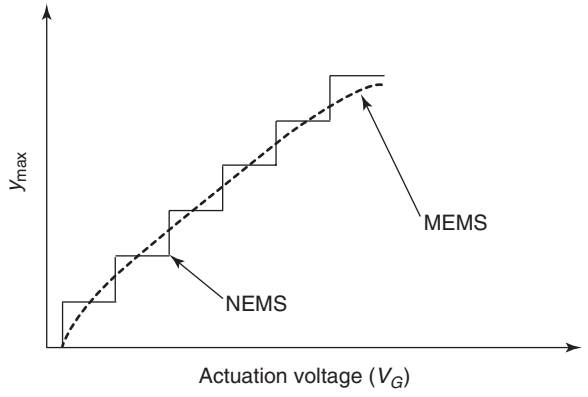


Figure 7 Double-clamped beam deflection at the micro- and nanoscale.

Another specific nanoscale force that acts on NEMS is the Casimir force, which is derived from the quantum field theory. Assuming that the vacuum is not empty but filled with virtual particles that are continuously created and annihilated, vacuum fluctuations manifest in the Casimir force that acts between two arbitrary bodies separated by a distance less than 1 μm. The Casimir force is not a classical force since it depends on geometry and the boundary conditions: it is attractive for two conducting plates separated by free space but becomes repulsive if the plates are replaced by hemispherical shells. For two metallic plates with an area of 1 cm^2 and separated at 500 nm, the Casimir force is $\sim 3\,\mu N$, a value that is easily detectable. Thus, the Casimir force is a mesoscopic force, which can be engineered via the distance between or the shapes of the interacting bodies. This engineering of the Casimir force is used to implement nonlinear mechanical oscillators and switches at the nanoscale, which consist of a metallic plate that moves under the action of an elastic force and a sphere located over the plate.

A complete analysis of the cantilever deflection requires a description of the cantilever motion in both space x and time τ coordinates. The corresponding equation is

$$EI(d^4y/dx^4) + m(d^2y/d\tau^2) = 0$$

where $A = Wt$ is the cross-sectional area of the cantilever and $m = \rho A$ is the cantilever mass per unit length, with ρ, the density of the cantilever material. Assuming a harmonic behavior of the cantilever and imposing appropriate boundary conditions, the natural resonance frequencies ω_n of the cantilever are related to its wavenumbers k_n through $k_n^4 = \omega_n^2 \rho Wt/EI$. The lowest-order wavenumbers are $k_0L = 1.875$, $k_1L = 4.694$, $k_2L = 7.855$, $k_3L = 10.996$, the first resonance frequency of the cantilever being

$$\omega_0 = 2\pi f_0 = 1.015(t/L^2)(E/\rho)^{1/2}$$

The cantilever performances are expressed by the quality factor, defined as

$$Q = 2\pi E_0/\Delta E$$

where E_0 is the stored vibrational energy and ΔE_0 is the total energy lost per cycle of vibration. The cantilever quality factor $1/Q = 1/Q_c + 1/Q_{th} + 1/Q_V + 1/Q_s$ is composed from several components: $Q_c = 2.2(L/t)^3$ originates from clamping losses and is of the order $10^3 - 10^5$, Q_{th} accounts for thermoelastic losses (which are negligible for submicron-thick cantilevers operating at kHz or MHz frequencies).

The losses due to internal friction, $Q_V = E_1/E_2$, are modeled via a complex Young's modulus $E = E_1 + iE_2$, where E_1 is the conventional Young's modulus and E_2 is a dissipation part due to friction, whereas the surface losses $Q_s = WtE_1/2\delta(3W + t)E_2$ can be modeled as a disruption of the atomic lattice having a thickness δ at the cantilever surface or as a surface contamination.

A general loss parameter $\gamma = K/\omega_0 Q$ is introduced to account for the cantilever performances, its value for a rectangular cantilever being

$$\gamma = 0.246Wt^2(E\rho)^{1/2}/LQ$$

This loss parameter is linked with the minimum detectable force F_{min}, a very important parameter in many applications. F_{min} is also related to the thermomechanical noise, which is the mechanical analog of the Johnson noise (white noise). The noise source of the cantilever is the Brownian motion around its equilibrium position resulted from the immersion of the cantilever in a thermal bath with a temperature T. The vibrational noise can be modeled by an equivalent noise force with a constant power spectrum S_F and an equivalent amplitude y. For a rectangular cantilever, the minimum detectable force (or force noise) in a bandwidth B is given by

$$F_{min} = S_F^{1/2}B^{1/2} = (Wt^2/LQ)^{1/2}(E\rho)^{1/4}(k_B TB)^{1/2}$$

or

$$F_{min} = (\gamma k_B TB/0.246)^{1/2}$$

Thus, the minimum detectable force is proportional to the temperature and losses. Nowadays, the minimum detectable force is ~ 1 attonewton (10^{-18} N) at low temperatures, and of a few pN (10^{-12} N) at room temperature.

The Membrane

Another important micromechanical device is the micromachined membrane (see **Figure 2**). The surface of the membrane is usually deformed by mechanical or electrostatic actuation, in the latter case the actuation requiring the deposition of a thin metal film on the back of the membrane. No such metallic film is necessary in applications in which the surface of the membrane is not deformed, that is, when the thin membrane with a thickness of $\sim 1\,\mu m$ is used as a propagating medium with low losses and a refractive index comparable to that of the air.

When the membrane is actuated, the surface deflection $u(x,y)$ is described by a Poisson equation

$$\nabla^2 u(x,y) = -\varepsilon_0 V^2(x,y)/Th$$

where $V(x,y)$ is the bias and T is the membrane tension. A similar Poisson equation in cylindrical coordinates is satisfied by circular membranes. The Poisson equation of rectangular and circular thin membranes has analytical solutions under certain boundary conditions.

MEMS Sensors

MEMS sensors are used in many areas ranging from cars to biology and neurosurgery. In what follows, only a brief review of a part of them is described.

Force sensors are made from a silicon beam, which encompasses a piezoelectric or a resonant strain gauge (see **Figure 8**), the resistance of which modifies in the presence of strain. One can determine the applied force by measuring the resistance.

MEMS pressure sensors are fabricated on a large scale, their applications ranging from car airbags and miniaturized microphones to complicated sensor arrays located in the airplane wings. There are many types of pressure sensors, but basically they consist of rectangular or circular MEMS membranes that vibrate and bend due to an applied pressure. Their bending is detected and transferred in an electrical signal. **Figure 9** displays a pressure sensor.

Acceleration sensors are based on a suspended proof mass linked to a fixed frame by spring elements.

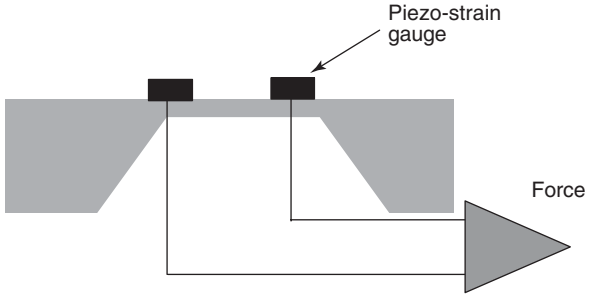


Figure 8 Schematic representation of an MEMS force sensor.

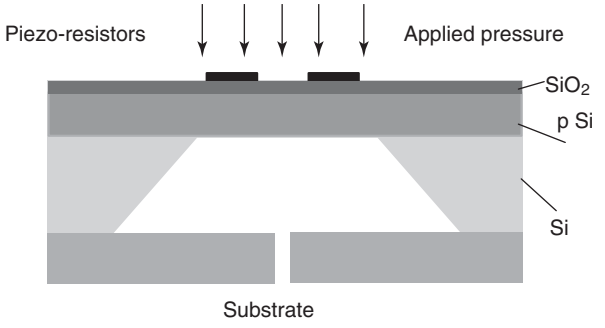


Figure 9 Schematic representation of an MEMS pressure sensor.

Any displacement of the proof mass, which is proportional with the acceleration, is detected by measuring the changes of the capacitance formed between the proof mass and an electrode or by measuring the changes of a piezo-resistor if strain gauges are integrated in the springs. A typical MEMS accelerometer is represented in **Figure 10**.

Thermal flow sensors can be divided into three categories: anemometers, calorimetric sensors, and time-of-flight sensors. In anemometers, one senses the variation of a resistance when heat is flowing over it. Practically, a thin film resistor is deposited on a thin membrane, the lateral diffusion flow of the heat over the membrane (see **Figure 11**) thermalizing the membrane at the temperature of the fluid. The flow rate can be measured with the simple flow sensor displayed in **Figure 12**.

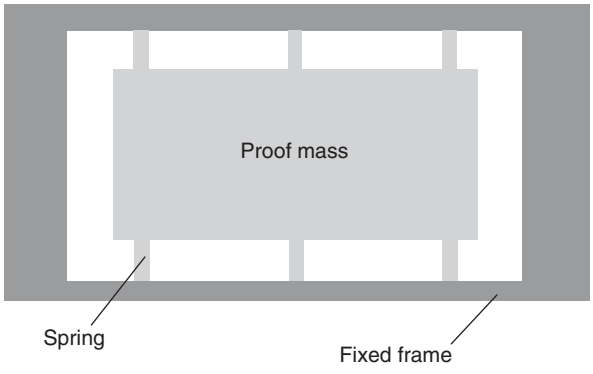


Figure 10 MEMS accelerometer.

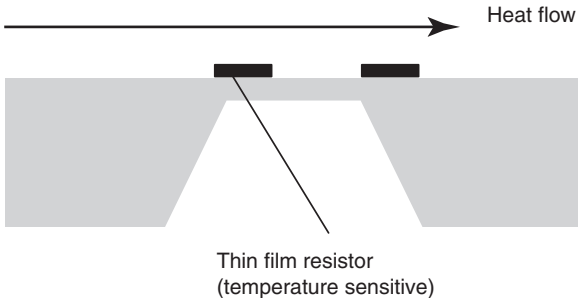


Figure 11 MEMS anemometer.

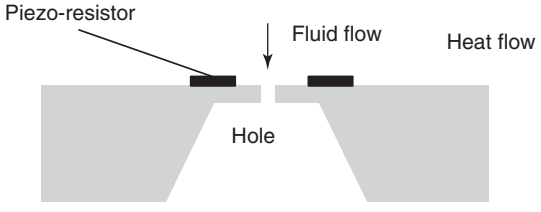


Figure 12 Schematic representation of an MEMS flow sensor.

A special category of sensors based on MEMS, with wide applications in biology and medicine, are the biological MEMS sensors or bioMEMS. For example, in surgery, catheters, pipettes, endoscopes, MEMS robotic surgery devices, including an MEMS submarine capable of traveling in human arteries, are currently fabricated using micromachining techniques; some of them contain several MEMS sensors. Minimally invasive devices able to provide *in vivo* diagnostics and therapy are currently developed using MEMS sensors and devices. Wireless MEMS needles have also been developed and implanted into the nervous system in order to monitor its activity. In the areas of drugs and genetics, MEMS microfluid devices are capable of delivering certain combinations of drugs, MEMS needles are capable of manipulating cells and genetic materials, and even smart labs-on-a-chip are created for point-of-care diagnostics.

RF-MEMS

RF-MEMS encompasses MEMS devices working in the range of microwaves, millimeter waves, and terahertz (THz) frequencies, that is, in the main frequency ranges used nowadays for communications. These frequency ranges are also used for medical imaging, space applications like teledetection, and radioastronomy.

The thriving development of RF-MEMS is caused by the ability of these devices to solve the main drawbacks of high-frequency circuits. These circuits consist of dielectric substrates (such as high resistivity Si or GaAs) over which metallic strips are patterned. The metallic strips and the substrate guide the high-frequency electromagnetic waves, but as soon as the frequency is increased beyond few tens of GHz, the losses in the metal and substrate become so high that the electromagnetic propagation is canceled.

There are two main categories of RF-MEMS devices. The first category encompasses RF-MEMS circuits supported on thin micromachined membranes of Si and GaAs, which allow the electromagnetic fields to propagate in similar conditions as in the air where the losses are minimal. The thickness of the dielectric substrate t has dramatic consequences on the performances of the device. In microwave circuits $t \ll \lambda$, a relation that guarantees the cancellation of substrate modes, which has, as a result, the reduction of the radiation losses and the dielectric losses. On the contrary, at millimeter waves, THz, and FIR frequencies, this inequality is no longer valid and very often the substrate thickness

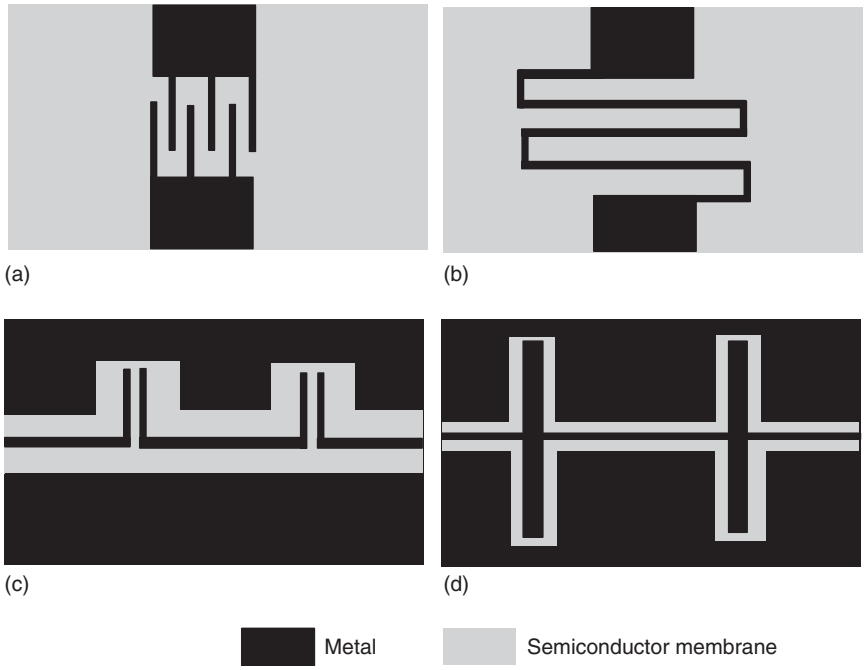


Figure 13 RF-MEMS devices: (a) capacitor, (b) inductor, (c) filter, and (d) double-slot antenna.

well exceeds the wavelength: $t \geqslant \lambda$. As a consequence, the energy generated above a critical angle is trapped in the semiconductor substrate by the substrate modes. Depending on the thickness, a large part of the radiated power (more than 90% in some situations) may become trapped into the substrate. Therefore, small losses at millimeter waves can only be obtained when propagating structures are patterned on very thin substrates. For example, the substrate thickness must be smaller than $0.04\lambda_0$ for a slot antenna and smaller than $0.01\lambda_0$ in the case of dipole antennas, with λ_0 the free-space wavelength. This means that a substrate thickness less than $2\,\mu m$ is needed, about two orders of magnitude thinner than the diameter of a hair. This amazing achievement became possible only in the last years using micromachining techniques, which are specific for MEMS. In **Figure 13**, some RF-MEMS devices realized on thin semiconductor membranes are described.

A second category of RF-MEMS devices are the reconfigurable devices. Filters, antennas, or receivers can change their characteristics in a controllable manner. This can be realized with the help of a series of RF-MEMS switches placed along some metallic paths of the circuit. RF-MEMS switches have many configurations, but in **Figure 14** only the shunt RF-MEMS switch, which is widespread in reconfigurable microwave devices, is described. This switch can also be used to commute between communication channels or communication protocols.

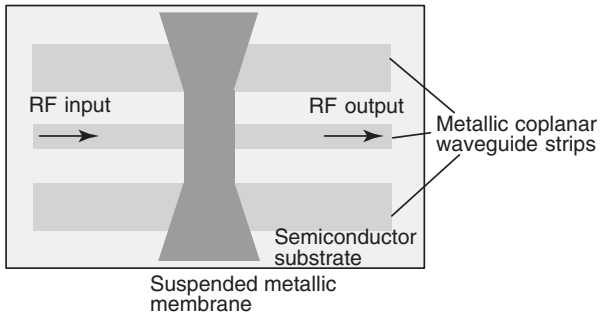


Figure 14 RF-MEMS shunt switch.

Optical MEMS (MOEMS)

Optical MEMSs integrate microoptical with micromechanical devices, the resulting microsystems being called microoptoelectromechanical systems (MOEMSs). MOEMSs are 3D integrated optical circuits that work mainly in free-space where self-calibrated and self-tested optical elements, such as microlenses, mirrors, beam-splitters, and gratings, can be moved in a controlled manner and at the precision required by optics. High-precision mechanical actuation produced by electrical, chemical, or electromagnetic forces are responsible for the controllable movement of the microoptical devices. MOEMSs have shown a huge potential for active and passive devices due to the fact that their main performances change when at least some of their components move. MOEMSs are

thus very versatile, reconfigurable, and moreover, show a high degree of integration never before met in optics and optoelectronics; more than 10^6 microoptical components can be integrated on a single MOEMS chip as in the digital display developed recently by Texas Instruments, which contains 2 million micromirrors, with a resolution of 2000×1000 pixels.

There are three main classes of optical devices based on MOEMS. The first class encompasses free-space optical devices, directly realized with micromachining techniques. The second class comprises optoelectronic devices such as lasers and photodiodes, integrated with micromechanical devices, which enhance their performances. The third class includes very complex circuits where hundreds of thousands of micromechanical devices are integrated with optical/optoelectronic devices in order to perform complex tasks such as optical signal processing, and optical data storage.

A breakthrough in MOEMS was the realization of the first microhinges and the subsequent development of an entire technology based on them, named free-space microoptical bench. In this surface micromachining technology, where selectively etching a sacrificial layer, a silicon plate linked to a hinge pin is able to turn freely around a silicon staple. The silicon plate can be positioned perpendicularly on the substrate or can be positioned at any angle due to a pair of microlatches that can also be realized using microhinges. The mechanical stability of the Si plate increases by adding V-shaped microlatches laterally to the plate. The MOEMS implementation of an integrated micro-Fresnel lens and a polarization beam splitter is displayed in **Figure 15**. In **Figure 16** a micromirror, which is a key element in many MOEMS devices, is displayed.

Optoelectronic devices can be micromechanically tuned. For example, a VCSEL laser tuned by a cantilever is displayed in **Figure 17**. The variable top DBR mirror consists of an *n*-DBR section etched into the shape of a 100 μm long cantilever, a variable air gap, and a fixed *p*-doped DBR mirror consisting of four pairs of alternating layers. Tuning the reversed bias applied on the *p*–*n* junction between the *n*-doped cantilever and the *p*-doped DBR varies the laser wavelength. If the reverse bias is increased, the cantilever is actuated toward the *p*-contact of the top DBR, reducing the air gap with ΔL and thus the laser wavelength. The VCSEL can be modeled as an active Fabry–Perot-like cavity, the change in the laser wavelength being expressed as $\Delta\lambda \cong (2/m)\, n\Delta L$, where m is the longitudinal mode number, and n the index of refraction. A very broad continuous tuning range of $\sim 20\%$ can be achieved in this manner.

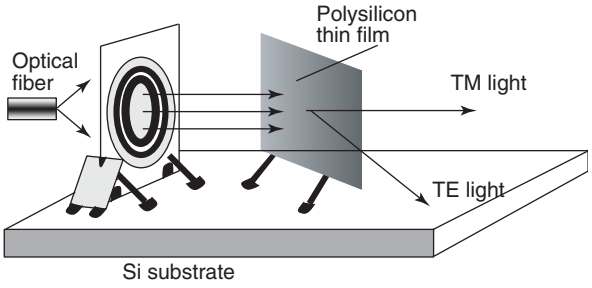


Figure 15 Micro-Fresnel lens integrated with a beam splitter.

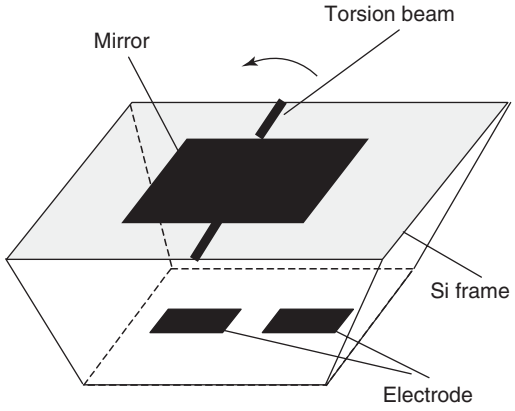


Figure 16 MOEMS micromirror.

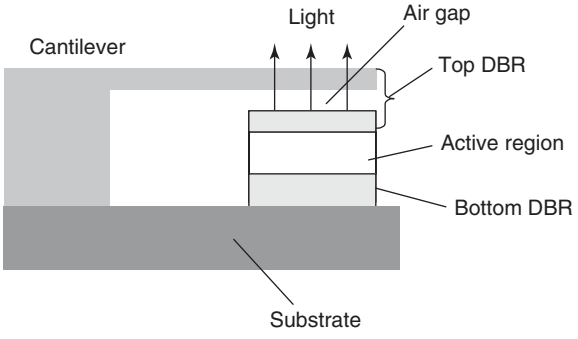


Figure 17 VCSEL tuned by a cantilever.

A more complicated MOEMS, which is an optical computer consisting of an array of cantilevers terminated with a tip, through which electrons are tunneling into a bottom electrode when the cantilever is optically actuated by incoming light, is displayed in **Figure 18**. The incoming 2D (1D) incident light beam should be sufficiently extended (eventually magnified) to cover an extended 2D (1D) array of cantilevers. The number of cantilevers in each dimension determines the step of the respective operation.

The addition $A(x, y) + B(x, y)$ of two incoherent optical beams $A(x, y)$ and $B(x, y)$ simultaneously incident on the array is performed automatically

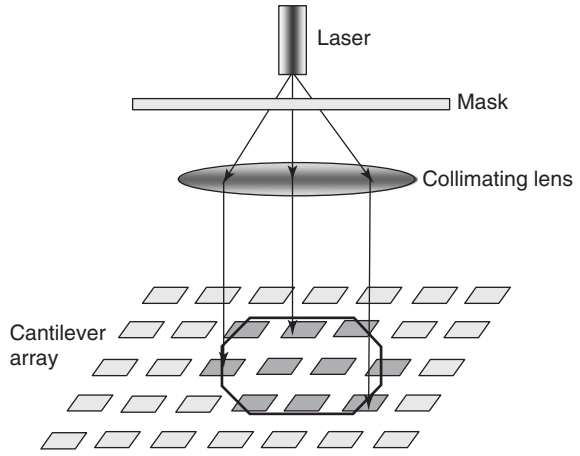


Figure 18 Schematic representation of an MEMS optical computer.

because the total deflection for each cantilever is the sum of deflections for the separate beams (the deflection is linearly proportional to the power). Moreover, since the deflection of the cantilever in the position (x_j, y_k) is proportional to $A(x_j, y_k)$, the integration of a 2D optical signal $A(x, y)$, $\int A(x,y)\mathrm{d}x\ \mathrm{d}y \approx \sum_{j,k} A(x_j, y_k)$, can be computed in one step by adding the tunneling currents of all cantilevers in the array.

Partial integration with respect to x or y is also possible by adding the tunneling currents on columns or rows, partial derivative of a 2D function being performed in a similar manner. To compute $\partial A(x, y)/\partial x$, for example, one must add the tunneling currents of two adjacent cantilevers along the x direction, with opposite signs. The opposite sign can be obtained by opposite biasing, that is, applying a voltage $-V$ instead of V. Analogously, $\partial A(x, y)/\partial y$ as well as $\mathrm{d}A(x, y)$ or even higher-order derivatives can be performed by a suitable connection of the contacts.

The method proposed above to compute integrals and derivatives of a 1D or 2D function based on MEMS devices is a breakthrough in the quest for the implementation of an optical computer since it eliminates many intermediate operations, which must be performed with conventional computing methods. The light power necessary to actuate the cantilevers is readily available from common optical sources, such as laser diodes or arrays of laser diodes.

See also: Integrated Circuits; Micromechanical Devices and Systems; Silicon, History of; Tunneling Devices.

PACS: 85.85.+j; 07.10.Cm; 07.10.Df

Further Reading

Dragoman D and Dragoman M (2001) Micro/nano-optoeletromechanical systems. *Progress in Quantum Electronics* 25: 229–290.

Dragoman D and Dragoman M (2004) Terahertz fields and their applications. *Progress in Quantum Electronics* 28: 1–66.

Elwenspoek M and Wiegerik R (2001) *Mechanical Microsensors*. Berlin: Springer.

Mantz P and Becker H (eds.) (1998) *Microsystem Technology in Chemistry and Life Sciences*. Berlin: Springer.

Plana R, Constantinidis G, and Muller A (eds.) (2003) *Radio Frequency MEMS*. Bucuresti: Editura Academiei Romane.

Rai-Choudhury P (ed.) (2000) *MEMS and MOEMS Technology and Applications*. Bellingham: SPIE Press.

Rebeiz GM (2003) *RF MEMS. Theory Design and Technology*. New York: Wiley.

Nomenclature

$A = Wt$	cross-sectional area of the cantilever
E	Young's modulus
E_0	stored vibrational energy
ΔE_0	energy lost per cycle of vibration
$F(y)$	electrical force
$F_{\min}$	minimum detectable force
h	cantilever substrate height
I	moment of inertia
k_B	Boltzmann constant
k_n	wavenumber
K	spring constant
L	cantilever length
m	cantilever mass
$M(x)$	total moment at position x
Q	quality factor
$u(x, y)$	surface deflection
V	applied voltage
V_{th}	threshold voltage
t	cantilever thickness; also thickness of dielectric substrate
T	temperature; also membrane tension
y	deflection
$y_{\max}$	maximum deflection
W	cantilever width
ε	permittivity
γ	loss parameter
λ	wavelength
ρ	density of cantilever material
ω_n	natural resonance frequencies

Micromechanical Devices and Systems

H Fujita, University of Tokyo, Tokyo, Japan

Introduction

The successful extension of semiconductor technology to fabricate mechanical parts of the sizes from 10–100 μm opened wide ranges of possibilities for micromechanical devices and systems. The fabrication technique is called micromachining. Micromachining processes are based on silicon integrated circuits (IC) technology and used to build three-dimensional (3D) structures and movable parts by the combination of lithography, etching, film deposition, and wafer bonding. The same technology base that enabled miniaturization and large-scale integration of electronics offers three distinctive features defining micromachined devices and systems: miniaturization, multiplicity, and microsystem integration. Miniaturization is clearly essential. Small parts respond fast, constitute miniature machines that work in extremely shallow spaces, add functionality to portable or wearable devices, and realize tools to investigate the nanometric world. Millions of such parts can work cooperatively to do things impossible for a single device alone. Thus, multiplicity is one key to successful micromechanical systems. The coordination of these parts is accomplished by integrating them with electrical circuits. Furthermore, optical, sensing, fluidic, and biological elements are to be integrated in multi-functional microsystems in a cost-effective manner. The research field concerning micromechanical devices is referred to as MEMS (microelectromechanical systems) in USA, Micromachine in Japan, and MST (micro system technology) in Europe.

The root of MEMS research can be found in the research of silicon sensors. A noticeable turning point from sensor research toward MEMS research was the demonstration of micromachined gears and turbines made on a silicon chip in 1987. Since then, development has continued in micromachining processes, material varieties, microactuators, and the application of MEMS.

Micromachining

MEMS generally requires complex microstructures that are thick, 3D, and movable. Therefore, many technologies have been developed on the basis of semiconductor microfabrication processes including lithography, etching, and deposition. Pre-designed patterns are transferred by exposure to resists that have sensitivity to light, X-ray, or electron beam and those that are coated on substrates. Photo lithograph through a glass-chromium mask is most commonly used. Parts of the resist remain on the substrate according to the exposed patterns by development. Using the resist pattern as a masking layer, one can selectively etch the substrate or the film on it. In another process, the resist mold pattern can be replicated by deposition. After resist removal, the next material is deposited on the patterned substrate. The material is, then, patterned by lithography and etching. These sequences are very similar to the IC fabrication.

There are special processes for micromachining to obtain 3D structures and movable parts (**Table 1**). To construct 3D microstructures, researchers have used

Table 1 Micromachining processes

Process (material)	*Feature*	*Application*
Wet anisotropic etching (single-crystal silicon)	Precise structures defined by crystal facets	Pressure sensor membrane, V-groove (channel, fiber alignment), Optical flat mirror surface
Dry anisotropic etching (silicon)	3D structure defined by mask patterns	High aspect ratio microstructure, through hole
Sacrificial layer etching (various thin films)	Integrated circuits compatible fabrication of micromovable structures	Integrated sensor, Arrayed MEMS (e.g., digital micromirror device)
Wafer bonding (Si or glass substrate)	Closed cavity formation, 3D assembly of devices on substrates	Hybrid integration of multifunctional system, Packaging and encapsulation
Electroforming (metal)	Replication of resist patterns by electroplating through the resist mask	Metallic 3D structures, Mater for injection molding
Injection molding and thermal embossing (polymer)	Low cost replication from precise master	3D microstrutures for fiber alignment or microchannels

Table 2 Selectivity of etching

Etching liquid/ gas	*Etched material*	*Resistant material*
HF (+NH_4F)	SiO_2	Si
KOH	Si	SiO_2, Si_3N_4
Hot H_3PO_4	Si_3N_4	SiO_2, Si
XeF_2 (gas)	Si	SiO_2, Si_3N_4
SF_6 (plasma)	Si	Ni, Al
O_2 (plasma)	Polymer	Al

anisotropic wet etching, deep reactive ion etching, and replication of deep lithography patterns. Another 3D fabrication technique is capable of folding micromachined plates out from the substrate by removing either a thin film under it or the substrate itself. The removal process is called sacrificial layer etching. Selective etching of the "sacrificial" material should be performed without damaging microstructures. **Table 2** shows the choice of etching solution or gas, the structural material and the sacrificial material. Even after release etching, microstructures tend to stick to the substrate. Sublimation drying, CO_2 supercritical drying, release etching by plasma or gas, and surface coating are typical methods to prevent sticking. Wafer bonding is also essential to micromachining. 3D stacking of structures, sealing of cavities and channels, hermetic or vacuum encapsulation of sensing/actuation elements and combination of devices made of different materials and processes are accomplished by bonding wafers together. Glass and silicon substrates are bonded by applying 400–500 V at 600–800 K; this is called anodic bonding. Two silicon substrates can be fusion bonded at $\sim$1370 K after careful cleaning of the surfaces to be bonded. Replication of micro molds is the inexpensive way of fabricating many microstructures. Molds can be made by deep lithography or deep etching. Electroplating of metals, chemical vapor deposition (CVD) of polysilicon films, and injection molding of polymers are commonly used for replication. Finally, direct fabrication of microstructures may be possible by beam assisted deposition or solidification by laser beam, electron beam assisted deposition, and focused ion beam (FIB) etching.

Materials for MEMS

Silicon is the most common material in micromachining because (1) the fabrication processes are well established, (2) it has excellent mechanical properties, and (3) integration with electronic circuits and sensors is possible. It also has piezoresistive properties used for strain detection. SiO_2 and SiN films serve as electrical and thermal insulators. They are also excellent masking layers for anisotropic wet etching by TMAH and KOH. Optical waveguides may be made of SiO_2. SiN and other nitride films, e.g., TiN and A1N, are very hard and used for low friction and antiwear coating.

Other materials that may not be common for IC fabrication are also useful in MEMS. Actuator materials produce force and displacement. They include piezoelectric materials (PZT, ZnO, quartz), magnetostrictive material (TbFe), shape memory alloy (TiNi), and some conductive polymers. Compound semiconductors such as GaAs, GaN and InGaAs can be micromachined in a way similar to silicon, and offer optically active and passive components. Metals can be deposited by electroplating, sputtering, and vacuum evaporation. Electroplated copper is widely used for conductors as well as Al thin films. Electroplated nickel is favored for both structural and magnetic elements. Researchers have developed batch fabrication processes for all of those materials achieving the features of MEMS mentioned above.

Microactuators

Microactuators are key devices allowing MEMS to perform physical functions. Many types of microactuators have been successfully operated. Some of them are driven by force associated with physical fields. Force can be generated in the space between stationary and moving parts using electric, magnetic, and flow fields. Some others utilize actuator materials introduced before. Thermal expansion and phase transformations, such as the shape-memory effect and bubble formation, cause shape or volume changes. Micromachining technology allows one to make structures in which a well-controlled field is generated or to deposit and pattern actuator materials. Typical sizes of microactuators are from 10 μm–1 mm. Although the physical principle that describes the motion of macroscopic objects is still applicable to microactuators, the relative importance of various forces changes in small dimensions. It is called a scaling effect. **Table 3** shows the dependence of some physical parameters and forces on the characteristic dimension of an object. The scaling effect and the compatibility with micromachining process are important issues when a microactuator is designed.

One of the most significant scaling effects is that friction dominates over inertia forces in the micro world. Researchers have tried many ways of minimizing friction by suspending movable parts with flexures, applying smooth coatings, using rotational contacts, levitating objects, and using friction drive mechanism.

Table 3 Scaling effect (L: characteristic dimension)

Parameter	*Formula*	*Scaling*	*Comment*
Gravity, f_g	Mg	L^3	g: gravity const., m: mass
Inertial force, f_i	Ma	L^4	a: acceleration
Elasticity, f_e	$eS\Delta L\_/L$	L^2	e: Yang's modulus
Surface tension, F_s	$L\gamma$	L	γ: surface wetting constant
Resonant frequency, ω	$\sqrt{K/m}$	L^{-1}	M: spring constant ($\propto L$)
Moment of inertia, I	$\alpha\, mr^2$	L^5	α: shape constant, r: radius of rotor
Reinors number, R_e	f_i/f_f	L^2	Ratio between inertia and viscosity ($\propto L^2$)
Heat conduction, Q_c	$\lambda\, \delta\, TA/\_d$	L	$\delta\ T$: temperature difference, λ: heat conductivity, A: cross-section area ($\propto L^2$)
Thermal expansion, F_r	$eS\Delta L(T)/L$	L^2	Valid for piezoelectric deformation ($\Delta\ L(E)$)
Magnetic force, F_m	$\mu\, SH^2/2$	L^4	μ: permittivity, H: magnetic field ($\propto L$)
Electrostatic force, F_e	$\varepsilon\, SE^2/2$	L^2	ε: permittivity, E: electric field (constant)

Table 4 Comparison of driving principles

	Electrostatic	*Magnetic*	*Piezoelectric*	*Shape memory alloy*	*Thermal*
Speed	Excellent	Good	Good	Fair	Poor
Force	Poor	Good	Excellent	Excellent	Excellent
Integration with circuits	Excellent	Good	Fair	Good	Excellent
Power consumption	Excellent	Fair	Good	Fair	Poor
Robustness in harsh environments	Poor	Good	Good	Excellent	Excellent

Each actuation principle has its own advantages and disadvantages (**Table 4**). The choice and optimization of an approach should be made according to the requirements of a particular application. Generally speaking, the electrostatic actuator is more suitable for performing tasks that can be completed within a chip (e.g., positioning of devices/heads/probes, sensors with servo feedback for self-test or readout, light deflection and modulation) since it is easily integrated on a chip, easily controlled, and consumes little power. On the contrary, the other types of actuators are more robust, more capable of producing larger forces, and more suitable for performing external tasks (propulsion, manipulation of objects, etc.).

Applications

Figure 1 maps the prospective applications in optics, transportation and aerospace, robotics, chemical analysis systems, biotechnologies, medical engineering, and microscopy using scanning microprobes. Most of these applications have a common feature in that only very lightweight objects such as mirrors, heads, valves, cells, and microprobes are driven, and very little physical interaction with the external environment is necessary.

Sensors

The first successful application of MEMS was pressure sensors for the precise control of an automobile engine combustion for clean exhaust in 1980s. MEMS accelerometers are currently used for air bag ignition sensors. MEMS gyroscopes, or angular rate sensors, are applied to navigation systems, active suspensions, and spin suppression systems. Precisely micromachined membranes serve as pressure sensing element; its deflection by applied pressure is detected by piezoresistive gauges made on the silicon membrane. In the accelerometer, a proof mass suspended by a flexible suspension is micromachined (**Figure 2**). When acceleration is applied to the proof mass, the initial force displaces the mass and the suspension; this motion is detected by either the strain gauge, that is, an integrated piezoresistor, or the change in capacitance between the mass and a fixed electrode. Better linearity and stability can be achieved by feedback measurement in which the inertia force is cancelled by electrostatic force; the magnitude of the force is proportional to the acceleration.

Printers

The ink jet printer head is another example of successful MEMS products. Micro nozzles and channels of 10–20 μm in size are micromachined in an array. A microheater is attached to each channel. The heater generates a bubble in the channel when a current pulse flows through it. An ink droplet is ejected from the nozzle by the pressure pulse caused by the bubble. The amount of the droplet is in the order of a few pL now. A typical head has hundreds

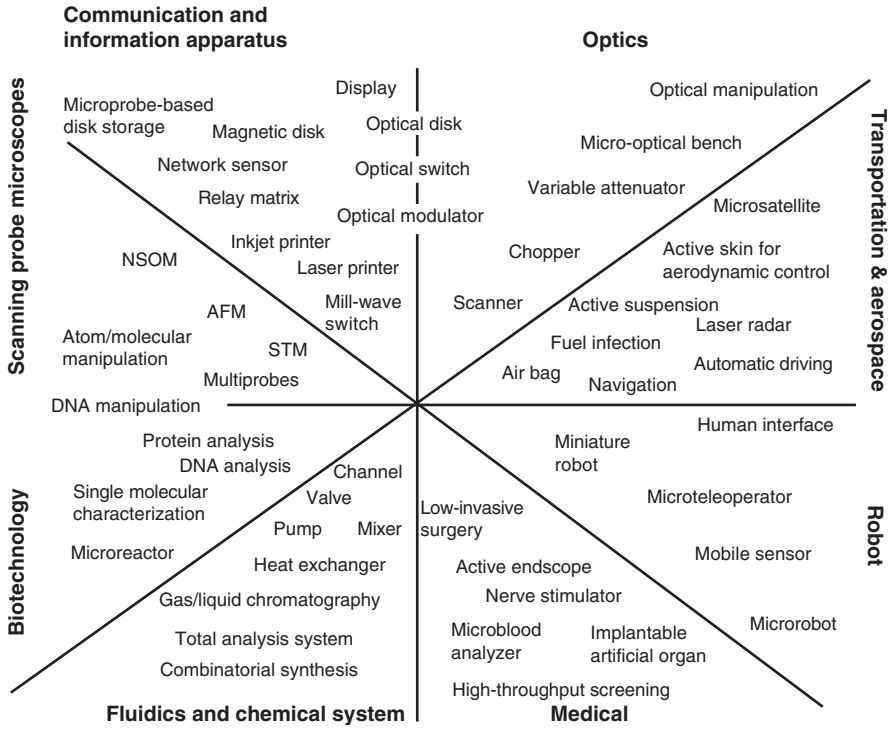


Figure 1 A map of MEMS applications.

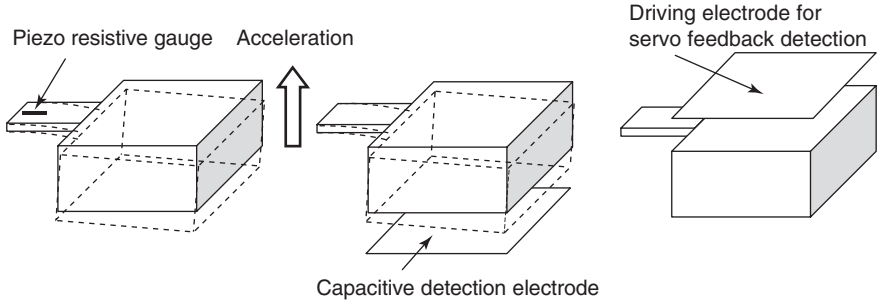


Figure 2 Micromachined accelerometer.

of nozzles and channels which operate as fast as 10 kHz. This device takes the advantage of the scaling effect, that is, the speed of heating and cooling increases with decrease in size. Thus, fine pictures and documents are printed on paper. Some heads use piezoactuators and electrostatic actuators to eject ink droplets.

Optical Devices

MEMS is very suitable to optical applications. Light beams can be controlled easily by moving mirrors, shutters, or gratings. MEMS with small feature sizes, precise motion and arraying capability can realize wide varieties of micro optical MEMS devices and systems based on interference, diffraction and refraction. MEMS technology also enables precisely aligned integration of active and passive optical devices, micromechanical elements and optical fibers. Therefore, optical MEMS devices and systems are used for optical scanners, micro-optical benches, arrayed micromirror displays, optical switches, modulators and attenuators.

The most successful example is a display. An array of movable micromirrors corresponding to pixels reflects light and projects an image on the screen. Each mirror is randomly accessible and tilted either $\pm 10^{\circ}$ by electrostatic force. The pixel becomes bright or dark depending on the angle of the mirror. The mirror moves within 0.1 ms and controls the brightness by pulse-width-modulation of reflected light. The array is micromachined on top of the integrated

circuits which allow the individual control of many mirrors within a short time.

MEMS based optical switches have been intensively studied. Optical switches, by means of mechanical movement, have favorable characteristics such as wave-length independence, polarization independence, large contrast, and small crosstalk. 2D MEMS optical switches change the optical path either straight or to a predetermined angle and connect it to one of the two output paths. The typical configuration of an N port 2D MEMS optical switch is shown in **Figure 3**. Light beams from input optical fibers are collimated by lenses. Micromirrors are inserted at appropriate positions to reflect beams and connect them to the output optical fibers. The alignment between fibers and mirrors is precisely determined by micromachined structures. Arrayed movable mirrors of several hundreds of micrometers in size are driven by microactuators. An on-off motion of the N^2 mirrors is necessary in the switch. The port count, N, is from 2–32. For port counts as large as a few hundreds, 3D MEMS switches are used. The schematic configuration is shown in **Figure 4**. Micromirrors have motion with two degrees of freedom, that is, rotation around two orthogonal axes. Two sets of arrays consisting of N such mirrors are used. Collimated input light beams hit one of the mirrors in the first array that direct the beam to one of the mirrors in the second array. The second mirror reflects the light straight into the output fiber. The analog control of mirror angles allows cross connection of a large number of signals. Those switches have application to optical cross connects, add-drop switches for wavelength divided multiplexing networks, and back-up or inspection apparatuses.

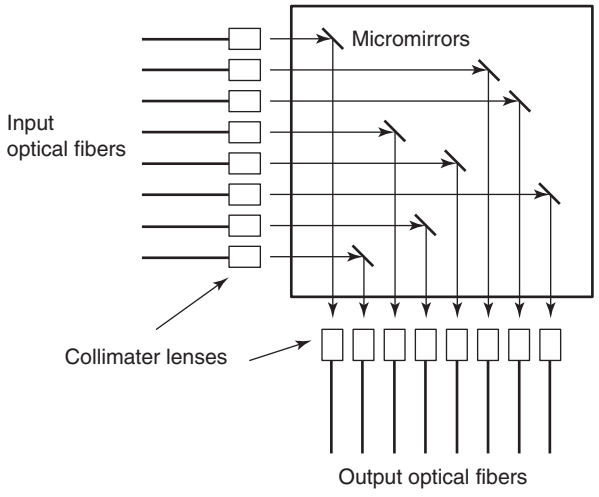


Figure 3 Principle of a 2D MEMS optical switch. Each micromirror is individually driven by a microactuator in the on-off fashion.

Optical scanners based on movable MEMS mirrors have many applications including bar-code readers, laser range finders, scanning laser microscopes, variable optical attenuators, and laser projection display. Some of them are already introduced in the market.

Chemistry and Biomedical Applications

Handling fluid in micromachined channels offers benefits in chemistry and biotechnology because of

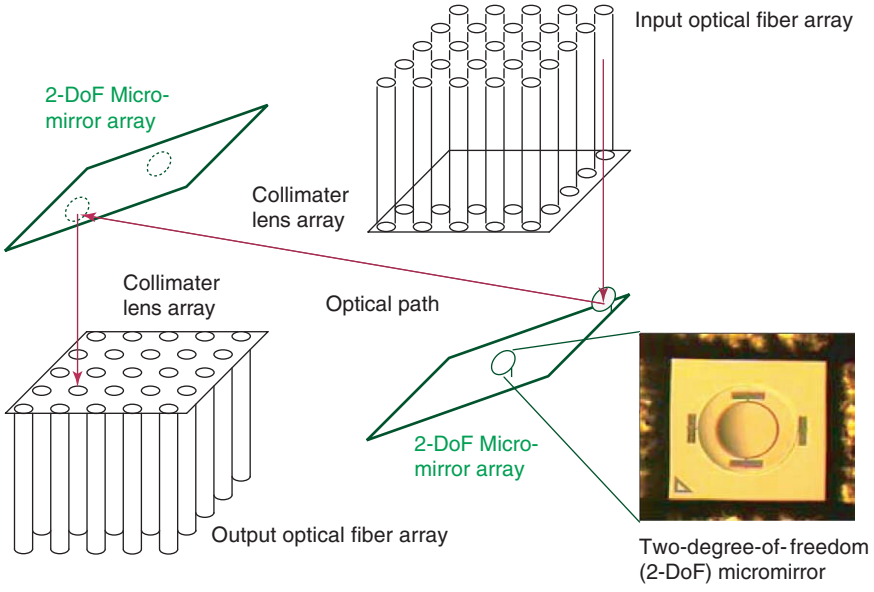


Figure 4 Principle of 3D MEMS optical switch.

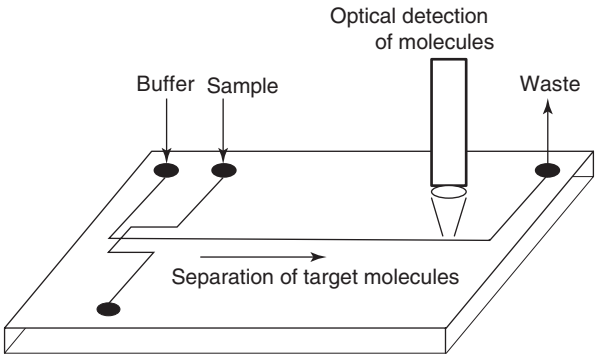


Figure 5 Principle of a chemical analysis chip based on capillary electrophoresis.

the scaling effect. The chemical reaction is completed within a few seconds due to short diffusion time in micro channels and reactors. Fast and precise temperature control is possible. Arrayed channels allow simultaneous chemical processing of a large variety of reactions. Only a small amount of samples or chemicals are required for chemical analysis and medical diagnosis. Even individual handling and observation of a cell or a molecule has been accomplished.

A typical microchemical chip is composed of one layer having grooves of 10–100 μm in dimension and another layer for sealing and connection to external pipes. Fluids flow in the channel to be mixed, heated, reacted, extracted, separated, and analyzed. Glass and PDMS (poly di-methyl siloxane) are two favored materials for the chip. Applications of microchemical chips include DNA analysis and amplification, blood analysis, environmental monitoring, combinatorial chemical synthesis, bioreactors and sensors, and production of fine chemicals.

The most successful product is the DNA electrophoresis in the micro channel. The device has two orthogonal channels (**Figure 5**). The sample solution containing DNA that is labeled by a fluorescent dye is introduced in the shorter channel by electroosmotic effect. A small amount of the sample is injected into the long separation channel by introducing a buffer solution from the end of the other channel. DNA fragments are driven by an electric field in the channel and separated due to the difference in migration speed depending on their length. Separation is completed within seconds and detected by fluorescence.

Future Development of MEMS

The future prospects of MEMS for making contributions to the future society are envisaged in three areas: (1) offering easier access to information for a wider public, (2) making human lifestyles more compatible with the environment, and (3) improving people's social welfare.

Breakthroughs to be accomplished by MEMS would be in five general areas: machine intelligence, downsizing and parallelism, biomimetics, informatics, and environment monitoring/preservation. Downsizing and parallelism will be the direct benefit from miniaturization and multiplicity. Machine intelligence and information networks can be very much improved by introducing MEMS integrated with microelectronics. These breakthroughs are likely to be realized by MEMS technologies within ten years. They may, for example, be used in healthcare for minimally invasive diagnosis and treatment.

See also: Integrated Circuits; Microelectromechanical Systems; Quantum Devices of Reduced Dimensionality.

PACS: 07.10.Cm

Further Reading

Elwenspoek M and Wiegerink R (2001) *Mechanical Microsensors*. Berlin: Springer.

Fujita H (ed.) (2003) *Micromachines as Tools for Nanotechnology*. Berlin: Springer.

Kensall DW (1998) *Special Issue on Integrated Sensors, Microactuators, and Microsystems (*MEMS*)*. Proceedings of the IEEE, 21 August 86(8): 1531–1787.

Nam-Trung N and Steven TW (2002) *Fundamentals and Applications of Microfluidics*. Boston, MA: Artech House Publishers.

Petersen KE (1982) *Silicon as a Mechanical Material*. Proceedings of the IEEE, 70: 420.

Rai-Choudhury P (ed.) (1997) Handbook of microlithography, micromachining and microfabrication. *Micromachining and Microfabrication*, vol. 2. Bellingham, WA: SPIE Optical Engineering Press.

Senturia SD (2001) *Microsystem Design*. Boston: Kluwer Academic.

Modulators, Optical

G Galzerano and P Laporta, Politecnico di Milano, Milan, Italy

Introduction

Modulated light has been used to transmit information over long distances in a fast and efficient way since the ancient times. The first modulation techniques were essentially manual, and were used to periodically obscure the light source, consisting either of reflected sunlight during the day, or fire beacons at night. Mechanical modulators, based on rotating mirrors and prisms, and different types of choppers, came into use at the beginning of the nineteenth century, producing intensity modulation of the optical source with repetition rate from a few hertz to a few kilohertz. A pulsed optical beam obtained by a rotating cog-wheel was employed for the first laboratory measurement of velocity of light performed by L Fizeau in 1849. The "photophone," invented by A G Bell in 1884, made use of a sunlight beam, the intensity of which was modulated by a mirror vibrating under the sound pressure of the voice, to transmit information. Starting from the 1930s, new modern devices were developed, based on electrooptic, acoustooptic, and magnetooptic effects, able to modulate light beams in the visible and infrared region with high efficiency and large bandwidths. Besides intensity, all relevant physical features of an optical wave, such as polarization, phase, and frequency, have been separately modulated exploiting various characteristics and physical effects, and ultrafast integrated electrooptical devices provide today modulation bandwidths up to several tens of gigahertz. At present, optical modulators are widely used with superior performance in many different fields, including optical communications, signal processing, and optical measurements (distance, velocity, etc.); in particular, they play a key role in all optical telecommunication systems. In this article, a general introduction to the most widely used optical modulators is presented, reviewing the basic principles of the theory and presenting significant results obtained with practical devices. On the basis of the different physical effects exploited, they have been subdivided into electrooptic, acoustooptic, magnetooptic, liquid crystal, and electroabsorption modulators.

Electrooptic Modulator

The term "electrooptics" mainly concerns the local effects induced by an electric field on the optical properties of a material. In particular, the Kerr and Pockels effects are related to the birefringence induced by an applied electric field to transparent materials (electrorefraction). The induced birefringence can be linearly proportional to the electric field (Pockels effect) or to the square of the electric field (Kerr effect). These effects are associated with the distortion of the electronic clouds by the applied electric field. In a quantum approach to the problem, the field modifies the wave functions of the electrons and the energy levels of the system, and consequently the dielectric response.

Electrorefraction provides, therefore, a physical method to directly modify the phase, the direction, and the polarization of a light beam in a controllable way. The Pockels and Kerr effects entered into the era of the optoelectronic technologies early in the 1950s when the use of these effects was proposed and demonstrated for the modulation and the switching of the light.

The optical properties of a material are characterized by the refractive index tensor (also called index ellipsoid), $(1/n^2)_{ij} = 1/\varepsilon_{ij}$, where ε_{ij} represents the ij element of the dielectric constant tensor, whose principal axes and eigenvalues determine the propagation of a light wave in the material. In a transparent material (lossless medium), ε_{ij} are real quantities, and for the symmetry of the tensor $\varepsilon_{ij} = \varepsilon_{ji}$, only six elements of the tensor are independent. The variations in the refractive index tensor induced by an applied electric field $\boldsymbol{E}$ are

$$\left(\frac{1}{n^2(\boldsymbol{E})} - \frac{1}{n^2(\boldsymbol{E}=0)}\right)_{ij} = \Delta\left(\frac{1}{n^2}\right)_{ij} = r_{ijk}E_k + s_{ijkl}E_kE_l + \cdots \cong r_{ijk}E_k + s_{ijkl}E_kE_l$$

where r_{ijk} and s_{ijkl} are the linear (Pockels) and quadratic (Kerr) electrooptic (EO) coefficients, respectively. The applied electric field induces a distortion of the index ellipsoid corresponding to a rotation of the ellipsoid and to a variation of the magnitude of the principal axes with regard to the zero-field condition. Due to the material dispersion and to the finite response time of the EO mechanisms (typically ranging from 10^{-12} to 10^{-14} s), the coefficients depend on the light wavelength and the frequency of the applied electric field. The Pockels effect exists only in materials that do not possess inversion symmetry (noncentrosymmetrical crystals), whereas the Kerr effect is present in all media and it is obviously the dominant term in centrosymmetrical materials (gases and liquids), where the Pockels effect is absent.

Table 1 Pockels tensors for a few noncentrosymmetrical crystallographic point groups. *N* indicates the number of nonzero independent coefficients for each group

Tetragonal

4*mm* ($N = 3$)

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$BaTiO_3$

$\bar{4}\,2\,m$ ($N = 2$)

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & r_{41} & 0 \\ 0 & 0 & r_{63} \end{bmatrix}$$

KH_2PO_4 (KDP)

Trigonal

32 ($N = 2$)

$$\begin{bmatrix} r_{11} & 0 & 0 \\ -r_{11} & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & -r_{41} & 0 \\ 0 & -r_{41} & 0 \end{bmatrix}$$

Quartz (SiO_2)

3*m* ($N = 4$)

$$\begin{bmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{bmatrix}$$

$LiNbO_3$ and $LiTaO_3$

Hexagonal

6 ($N = 4$)

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{13} \\ 0 & 0 & r_{33} \\ r_{41} & r_{51} & 0 \\ r_{51} & -r_{41} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$LiIO_3$

6*mm* ($N = 3$)

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

CdS

Axial

$\bar{4}\,3m, 23$ ($N = 1$)

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & r_{41} & 0 \\ 0 & 0 & r_{41} \end{bmatrix}$$

(GaAs, CdTe)

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The EO coefficients r are in the order of the reciprocal of the internal electric field experienced by the electrons in the material, that is, $r \approx 10^{-11}\,\mathrm{m\,V^{-1}}$. In a similar way, the s coefficients are in the order of the reciprocal of the squared internal electric field, $s \approx 10^{-22}\,\mathrm{m^2\,V^{-2}}$. The number of nonzero and linearly independent EO coefficients is determined by the point symmetry of the material. As an example, **Table 1** reports the Pockels tensors for a few commonly used crystallographic point groups.

From a device operation point of view, the key parameter is the birefringence, Δn_E, associated with the selected direction of propagation in the EO material, induced by the electric field. This corresponds to the difference between the refractive indices, $\Delta n_E = n_1 - n_2$, of the two allowed directions of polarization (eigenmodes), and determines the EO retardation, Γ, between the eigenmodes at the output plane of the material

$$\Gamma = \frac{2\pi \Delta n_E l}{\lambda} = \frac{\omega \Delta n_E l}{c}$$

where λ is the light wavelength in vacuum, l is the length of the material, ω is the angular frequency of the optical beam, and c is the speed of the light in vacuum. In a Pockels material, the retardation can be expressed as $\Gamma = \pi(V/V_\pi)$, where V is the applied voltage and V_π is the voltage corresponding to a retardation equal to π; when a Kerr material is considered the retardation is given by $\Gamma = \pi(V/V_\pi)^2$.

The electric field can be applied parallel to the direction of the light propagation (longitudinal configuration) or perpendicular to it (transverse configuration) (see **Figure 1**). For the longitudinal configuration, used only with Pockels materials, Γ linearly depends on the applied potential, V, and not on the crystal interaction length l (typical values of the V_π voltage are in the range from 1 to 10 kV for visible radiation). In the transverse configuration, Γ depends on both the applied potential (V or V^2 for Pockels or Kerr materials, respectively) and on the crystal length l. In this last case, V_π can be significantly lower than that obtained with the longitudinal configuration (in particular for waveguide structures, **Figure 1c**, where the ratio between the waveguide length and its transverse aperture, d, is much larger than 1), and, in addition, the field electrodes do not interfere with the optical beam. For these reasons, the transverse configuration is widely adopted.

EO phase modulators perform a modulation of the phase of light beams. Longitudinal and transverse configurations can be used. The (linear) polarization of the incoming optical beam has to be along a direction parallel to an externally induced principal axis. In this case, when the light propagates in the modulator, the polarization state is unmodified, but the optical field experiences a refractive index which is a function of the applied electric field. At the modulator output, the phase of the optical wave is incremented by

$$\begin{aligned} \Delta\phi(E) &= \frac{\omega}{c}[n(0) + \Delta n_E(E)]l \\ &= \Delta\phi(0) + \frac{\omega}{c}\Delta n_E(E)l \end{aligned}$$

where $n(0)$ is the unperturbed refraction index (ordinary), and Δn_E can be expressed as $\Delta n_{E,\,\mathrm{Pockels}} = n_0^3 r_{\mathrm{eff}} E/2$ or $\Delta n_{E,\,\mathrm{Kerr}} = n_0^3 s_{\mathrm{eff}} E^2/2$ for Pockels or Kerr modulators, respectively (r_{eff} and s_{eff} are the effective EO coefficients for the Pockels and Kerr

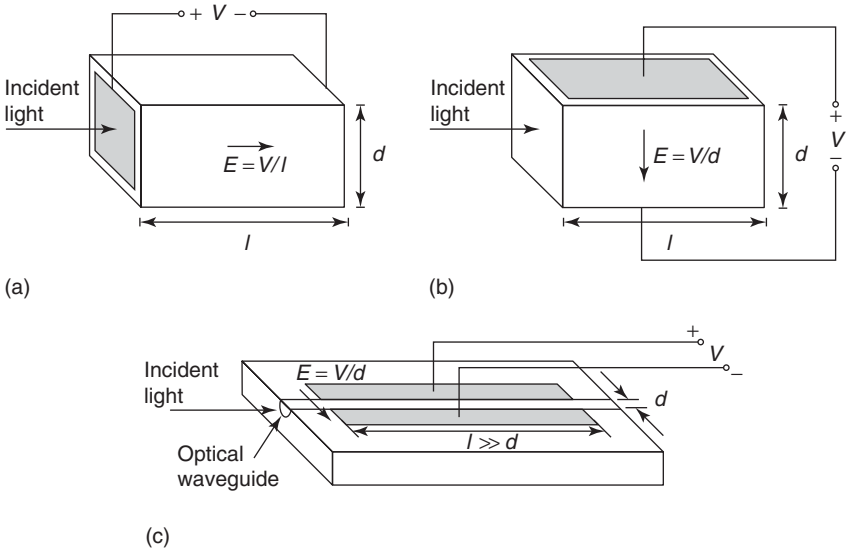


Figure 1 Electrooptic configurations: (a) longitudinal, (b) transversal in bulk crystal, and (c) transversal in an optical waveguide structure. *l* is the interaction length in the EO crystal, *d* is the aperture of the EO modulator, and *E* is the applied electric field.

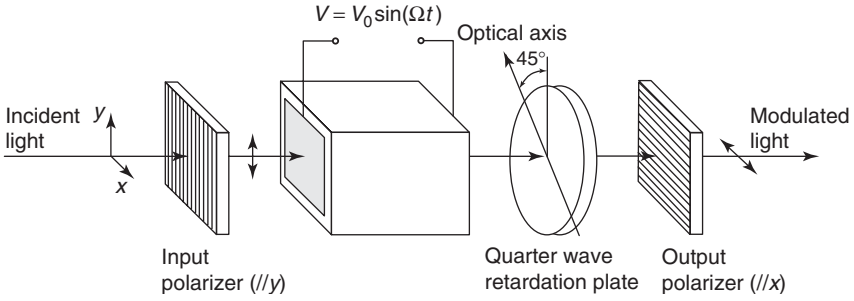


Figure 2 Schematic diagram of a typical longitudinal EO intensity modulator. The EO crystal acts as a variable retardation plate modulated at angular frequency Ω by the applied voltage *V*. The quarter-wave plate is used to obtain effective linear modulation around 50% transmission.

effects, respectively). The phase of the light beam can thus be modulated if the applied electric field is modulated (e.g., $E = E_0 \sin \Omega t$); assuming a Pockels modulator and a sinusoidal modulation of the applied electric field, the instantaneous phase of the light beam at the output of the modulator is

$$\Delta\phi(E) = n_0 \frac{\omega}{c}\left[1 + n_0^2 r_{\text{eff}} \frac{E_0 \sin \Omega t}{2}\right] l$$
$$= \Delta\phi(0) + \beta \sin \Omega t$$

where β is the phase modulation index (in unit of radiant) and Ω is the angular frequency of the modulation electric field.

In the polarization modulators, the EO crystal acts as a variable retardation plate. In this way, by means of a linear polarizer selecting the polarization of the input beam, the polarization state of the output beam is determined by the EO retardation phase Γ, which is modulated by the applied electric field. Transverse and longitudinal configurations can be used to implement an EO polarization modulator. In a general operating condition, the polarization of the transmitted light at the output of the modulator can therefore be changed with continuity between two limiting elliptical states with a frequency equal to that of the modulation signal.

To achieve intensity modulation of a light beam by means of the EO effect, it is necessary to combine a polarization modulator with a polarization analyzer placed at the output of the modulator, consisting of a second linear polarizer with crossed direction with respect to the input polarization (see **Figure 2**). In this configuration, the intensity I_t of the light transmitted by the output polarizer is related to the phase difference between the two eigenmodes of the optical beam propagating in the modulator by the relation

$$I_t = I \sin^2\left(\frac{\Phi}{2}\right) = \frac{I}{2}(1 - \cos \Phi)$$

where I is the intensity of the incident beam, and Φ is the total phase retardation given by the sum of the EO retardation Γ and the intrinsic birefringence of the material, Φ_0, which for a Kerr medium is always zero. For an efficient linear modulation, it is necessary to operate the modulator around Φ values close to $\pi/2$ (see **Figure 3**). This can be achieved either by inserting an additional quarter wave plate (for example, adjacent to the output polarizer), or by using a DC bias which produces a constant phase shift of $\pi/2$. In this case, the intensity of the transmitted light is given by

$$I_t = \frac{I}{2}[1 + \sin(\Gamma)] = \frac{I}{2}\left[1 + \sin\left(\pi \frac{V}{V_\pi}\right)\right]$$

when a Pockels modulator with $\Phi_0 = 0$ is considered. Assuming a sinusoidal modulation $V = V_0 \sin \Omega t$, and $V_0 \ll V_\pi$, the transmitted intensity is

$$I_t = \frac{I}{2}\left(1 + \pi \frac{V_0}{V_\pi} \sin \Omega t\right)$$

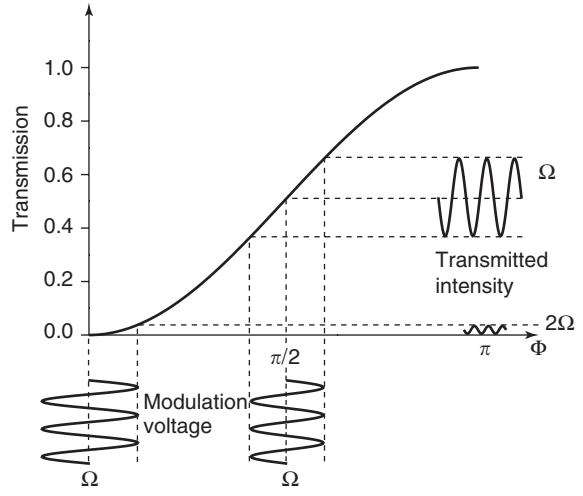


Figure 3 Transmission of the EO intensity modulator as a function of phase retardation Φ. For an operating condition around $\Phi = \pi/2$, the voltage modulation at angular frequency Ω is converted into intensity modulation. When the modulation is applied around $\Phi = 0$, the efficiency strongly decreases while the optical beam is modulated at twice the modulation frequency.

In the above description of the EO modulators, the expressions for the retardation produced by the electric field are valid in the case of low modulation frequencies. Increasing frequency, the transit time of the light through the modulator, $\tau_d = nl/c$, sets a limit, $\Omega_{max} = \pi c/2nl$, to the maximum modulation frequency (typically Ω_{max} is of the order of a few gigahertz). To overcome the transit time limitation, the modulation signal has to be applied in the form of a traveling wave by means of a suitable radiofrequency (RF) waveguide structure. In these devices, the phase velocity of the optical beam is matched to the phase velocity of the RF modulation signal. This form of modulation can be adopted only using transverse geometry (see **Figure 4**). The maximum modulation frequencies of a traveling wave modulator are in the range from 10 to 100 GHz.

EO modulators are widely used in laser and photonics fields thanks to the excellent performance over a wide wavelength spectral range, from the ultraviolet to the far infrared. Pockels modulators are much more diffused than Kerr cells because of the higher efficiency. **Table 2** shows the EO properties of some of the most commonly used Pockels materials. For laser applications, EO modulators are used both inside the laser resonator to obtain different pulsed operation regimes, such as Q-switching, cavity dumping, and mode-locking, and externally to the laser source to directly modulate the intensity, frequency, polarization, and direction of the optical beam at very high modulation frequencies (up to 100 GHz). In particular, active mode-locked lasers based on intracavity EO modulators enable the generation of very short pulses in the picosecond regime at extremely high repetition rates. A very convenient configuration for implementing the EO effect is represented by an optical waveguide structure. In this configuration, typical values of the half-wave voltages are a few volts, and low driving power can be used. The waveguide EO modulators can be realized with a single waveguide or a pair of waveguides (Mach–Zender configuration), and they can operate as phase/intensity modulators or as optical switches. These modulators are commonly fabricated using titanium indiffused $LiNbO_3$ and GaAs waveguides.

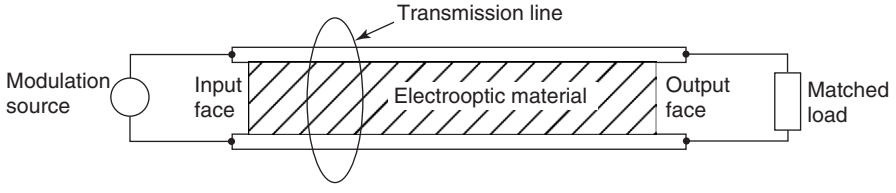


Figure 4 Schematic diagram of a traveling wave EO modulator.

Table 2 Electrooptic properties of some commonly used Pockels materials. The value of the EO coefficients and refractive indexes refer to $\lambda = 633\,\text{nm}$

Material	*Transparency range* (μm)	r_{eff} [a] (pm V^{-1})	*Ordinary refractive index* n_o	*Extraordinary refractive index* n_e
Potassium dihydrogen phosphate-KDP (KH_2PO_4)	0.2–1.5	− 10.3	1.507	1.467
Deuterium substituted-KD[a]P (KDP KD_2PO_4)	0.2–2.1	26.4	1.493	1.458
Lithium niobate ($LiNbO_3$)	0.33–4.5	26.6	2.286	2.20
Lithium tantalate ($LiTaO_3$)	0.28–4	25.1	2.176	2.18
Cadmium telluride (CdTe)	0.85–29.9	− 6.8	2.6	
Gallium arsenide (GaAs)	1.0–11	− 1.6	3.6	
Zinc sulfide (ZnS)	0.4–12.5	− 1.9	2.47	

[a] For transversal configuration.

Acoustooptic Modulators

Diffraction of light by sound was first predicted by Brillouin in 1922, and experimentally observed some 10 years later by Debye and Sears in liquids, and Raman and Nath in solids. The elastooptical effect is the basic mechanism responsible for the acoustooptic interaction. It describes the change in the refractive index of a medium caused by the mechanical strains related to an acoustic wave traveling through the material. The relationships between the change of refractive index and the strain are in general rather complicated; it can be shown that the change in refractive index is proportional to the square root of the acoustic power. By means of the elastooptical effect, the traveling acoustic wave sets up a periodic spatial modulation of the refractive index, as depicted in **Figure 5**, which, under proper conditions, may diffract the light beam into one or more directions. The interaction of the light beam with the acoustic beam can result in optical beam deflection and modulation of amplitude, frequency, phase, or polarization of the optical wave. A simplified model of the acoustooptic interaction is based on a plane wave analysis. In this model, the incident light is assumed to be a plane wave which is diffracted by a rectangular sound column into a number of plane waves propagating in different directions. The resulting set of coupled wave equations are the Raman–Nath equations. The solution of these equations can be classified in two different regimes, depending on the ratio of the interaction length, L, and a characteristic length, $L_0 = n\Lambda^2/\lambda$, where n is the refractive index of the material and Λ is the wavelength of the acoustic wave. There are two main cases of interest, namely, (1) the Raman–Nath regime, where $L \ll L_0$, and (2) the Bragg regime, corresponding to $L \gg L_0$.

(1) In the Raman–Nath regime, the acoustic diffraction grating is so thin that the diffracted light suffers no further redistribution before leaving the modulator. The light is diffracted as from a simple plane phase grating, with diffraction angles given by $\sin\theta_N = N\lambda/\Lambda$, where $N = 0, \pm 1, \pm 2, \ldots$ are the diffraction orders. A schematic picture of the interaction between the light beam and the sound wave is shown in **Figure 6**. The intensities in the various orders depend on the maximum excursion of the phase delay in the grating, which is related to the amplitude of the acoustic modulating wave. More specifically, the intensity I_N diffracted in the Nth order is given by $I_N = I_i J_N^2[(2\pi/\lambda)L\Delta n]$, where I_i is the intensity of the incident beam, J_N is the Nth order Bessel function, and Δn is the maximum refractive index variation. The largest possible diffraction efficiency in each order is the maximum value of J_N^2; for the first diffraction order this maximum is 33.8%, resulting in a relatively low efficiency of the modulator.

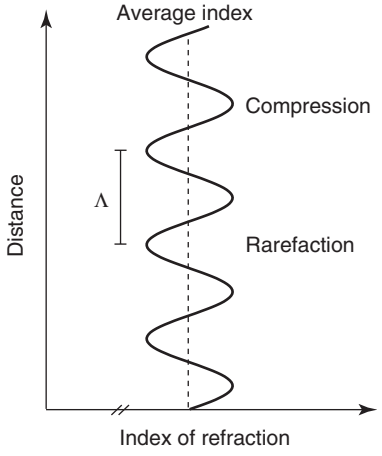


Figure 5 Schematic illustration of acousto-optic modulation. The sound wave produces alternating regions of compression and rarefaction in the medium, resulting in a spatial variation of the refractive index with period Λ equal to the acoustic wavelength.

(2) In the Bragg regime, there is a long interaction path between the sound and the incident light. Under this condition, the periodic modulation of the refractive index induced by the acoustic field behaves as a thick phase grating, and the light undergoes diffraction from a series of parallel acoustic phase fronts acting as partially reflecting mirrors. The condition for constructive interference along a direction can be found by imposing that all the points on a given mirror contribute in phase along this direction and, in addition, the diffraction from any two acoustic phase fronts adds up in phase along the direction of the reflected beam. With reference to **Figure 7**, the two conditions are satisfied when the incident angle is equal to the diffracted angle and is given by

$$\sin \theta_B = \lambda/2n\Lambda$$

where θ_B is the Bragg angle. It can be shown that in this condition the diffraction efficiency $\eta = I_d/I$ (I_d and I are the diffracted and incident intensities) is given by

$$\eta = \sin^2[(\pi/\lambda)/(L\Delta n/\cos \theta_B)]$$

and can theoretically equal 100%. Most acousto-optic devices are therefore designed to operate in the Bragg regime.

The sound wavefronts which generate the diffraction grating are moving through the medium with velocity v_s; as a consequence, the optical diffracted wave undergoes a frequency shift according to the Doppler effect. The frequency shift, $\Delta v = v_d - v$, between the frequency of the diffracted (v_d) and

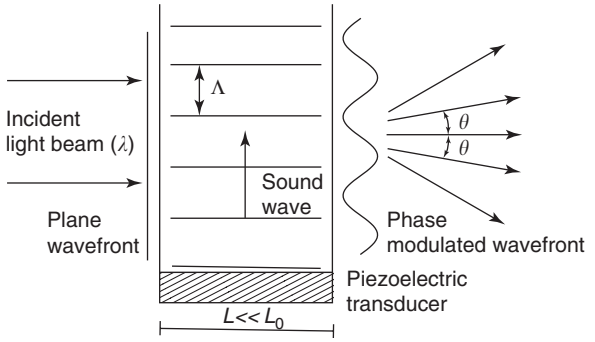


Figure 6 Diffraction of light by an acoustic wave in the Raman-Nath regime and schematic of the modulator. The interaction length $L \ll L_0$ and thus the light is diffracted as from a simple plane phase grating. The first order diffraction angle $\theta \approx \lambda/\Lambda$, where λ and Λ are, respectively, the optical and acoustic wavelength.

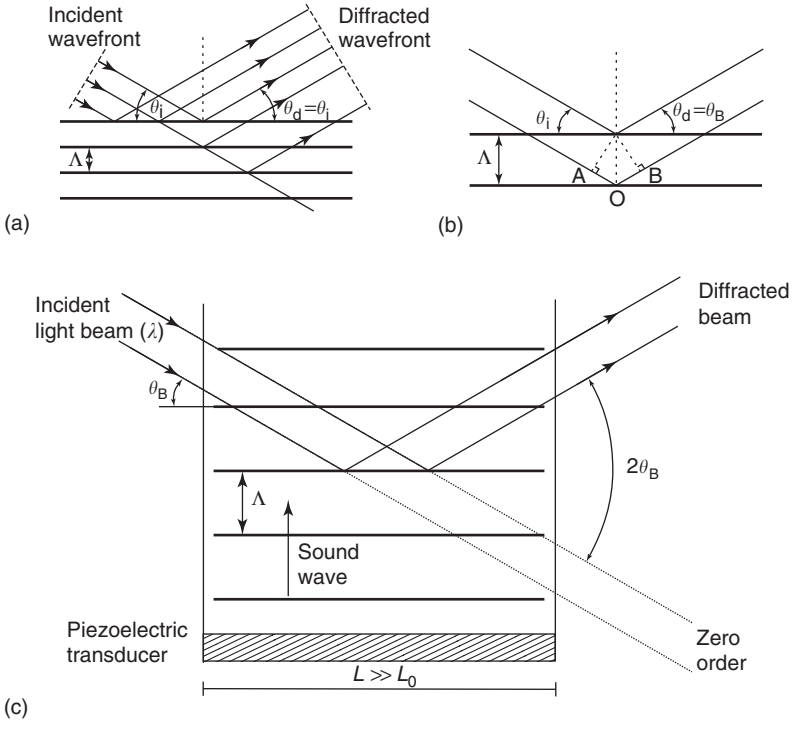


Figure 7 Diffraction of light by an acoustic wave in the Bragg regime. The interaction length $L \gg L_0$ and the light is reflected from successive layers or acoustic phase fronts of the grating. (a) All the points on a given layer contribute in phase at the new diffracted wavefront (diffraction angle θ_d equal to the incident angle θ_i); (b) the optical path difference $\overline{AO} + \overline{OB}$ is equal to one optical wavelength so that the diffraction from any two layers adds up in phase (constructive interference for diffraction angle θ_d equal to Bragg angle θ_B); (c) schematic of the acoustooptic modulator. Λ and λ are, respectively, the acoustic and optical wavelength.

incident (v) beams is given by $\Delta v = \pm 2v\, v_s \sin\theta_B\, n/c$, where the plus sign applies for a sound wave moving toward the optical beam. Substituting the Bragg condition in the above relationship, one obtains $\Delta v = \pm f$, where $f = v_s/\Lambda$ is the acoustic wave frequency. This change in frequency can be used as the basis of a frequency shifter.

Diffraction of light from acoustic waves can also be interpreted in terms of the quantum theory. Accordingly, an optical wave of angular frequency ω and wave vector $\boldsymbol{k}$ is regarded as a stream of photons with momentum $\hbar\boldsymbol{k}$ and energy $\hbar\omega$. An acoustic wave of angular frequency $\omega_s = 2\pi f$ and wave vector $\boldsymbol{k}_s$, likewise, is regarded as a stream of phonons with momentum $\hbar\boldsymbol{k}_s$ and energy $\hbar\omega_s$. Scattering of light by the sound wave can be described as series of collisions, each involving the annihilation of one incident photon and one phonon, and the simultaneous creation of a new diffracted photon with momentum $\hbar\boldsymbol{k}_d$ and energy $\hbar\omega_d$, propagating along the direction of the diffracted beam. The conservation of momentum, when the sound wave moves approaching the incident beam, requires that $\boldsymbol{k}_d = \boldsymbol{k} + \boldsymbol{k}_s$, as illustrated in **Figure 8**. The conservation of energy implies that $\omega_d = \omega + \omega_s$, which is the Doppler formula for upshifted Bragg diffraction. The conservation of momentum is equivalent to the Bragg condition, considering that the sound frequencies of interest are below 10^{10} Hz and the optical frequencies usually above 10^{14} Hz, so that $\omega_s \ll \omega$, $k_d \approx k$, and the magnitude of the sound wave vector is thus $k_s = 2k \sin\theta_B$ (see **Figure 8**).

The diffraction efficiency can be expressed in terms of acoustic intensity. In fact, the refractive index change is related to the strain s_r by $\Delta n = -n^3 p s_r/2$, where p is the photoelastic constant for an isotropic medium, and s_r is related to the acoustic intensity I_{ac} by $s_r = [2I_{ac}/(\rho v_s^3)]^{1/2}$, being ρ the mass density. Combining the above expressions with that of the diffraction efficiency, one obtains under the Bragg condition

$$\eta = \sin^2\left[\frac{\pi L}{\sqrt{2}\lambda\cos\theta_B}\sqrt{MI_{ac}}\right]$$

where $L/\cos\theta_B$ is the oblique distance of penetration of light through the acoustic wave, and $M = n^6 p^2/\rho v_s^3$ is a figure of merit relating the diffraction efficiency (the refractive index change) to the acoustic intensity. M is generally reported as a normalized quantity, referred to the value for fused silica, and turns out to be the acoustooptic figure of merit most frequently used to compare different materials, even if there are several others defined in the literature which may better apply to evaluate specific applications of various acoustooptic devices. For relatively small values, the diffraction efficiency turns out to be linearly proportional to acoustic intensity.

A property which strongly influences device performance is acoustic attenuation, mainly due to losses caused by relaxation of thermal phonon distribution toward equilibrium. In most crystals, it has been observed that the attenuation of the sound wave has a quadratic dependence on the acoustic frequency. In **Table 3**, a list of selected materials used in acoustooptic devices is reported together with their relevant properties.

Acoustooptic interaction can be exploited for temporal, spatial, and spectral modulation of light, using different devices such as intensity modulators, deflectors, tunable filters, and frequency shifters.

Intensity modulation of the optical beam is proportional to the intensity of the acoustic wave, if the sound intensity is sufficiently weak, and a linear analog modulation of light intensity can be obtained. Increasing the acoustic power, the diffraction efficiency η increases but the modulation is no longer linear, and the acoustooptic (AO) modulator serves as an optical switch. An important parameter for AO modulators is the modulation bandwidth, that is, the maximum frequency at which the modulator can efficiently operate. When the amplitude of an acoustic wave of frequency f (carrier) is varied as a function of time by a signal of bandwidth B, the acoustic wave acquires frequency components within a band $f \pm B$. The principle of operation of an AO modulator is schematically illustrated in **Figure 9**. In order to fulfill the Bragg condition and achieve modulation of the optical beam, each of these frequency components will interact with a proper plane wave of the angular spectrum of the incident light beam. The variation of the Bragg angle corresponding to a frequency

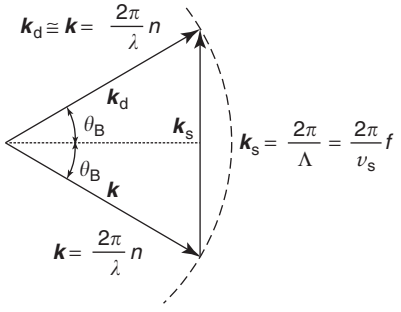


Figure 8 Momentum conservation in the interaction of an optical plane wave with wavevector $\boldsymbol{k}$ (or an incident photon with momentum $\hbar\boldsymbol{k}$) with an acoustic plane wave with wavevector $\boldsymbol{k_s}$ (or a phonon with momentum $\hbar\boldsymbol{k_s}$). The vector relation $\boldsymbol{k_d} = \boldsymbol{k} + \boldsymbol{k_s}$ is equivalent to the Bragg condition $k_s = 2\,k \sin\theta_B$. Λ and λ are, respectively, the acoustic and optical wavelength; n is the refractive index.

Table 3 Relevant properties of selected materials used in acoustooptic devices. The optical polarization is parallel to the plane of diffraction containing the acoustic and optical wave vectors and the refractive index refers to $\lambda = 633\,\text{nm}$, if not specified otherwise

Material	*Optical transmission* (μm)	*Refractive index*	*Density* ($10^3\,\text{kg m}^{-3}$)	*Acoustic polarization*[a] *(direction)*	*Acoustic velocity* ($10^3\,\text{ms}^{-1}$)	*Acoustic attenuation* ($\text{dB}\mu\text{s}^{-1}\,\text{GHz}^{-2}$)	M^{*b}
Fused silica[c]	0.2–4.4	1.46	2.2	L	5.95	7.2	1.0
GaAs[d]	1.0–11	3.37	5.34	L [1 1 0]	5.15	15.5	68.8
GaAs[c,d]	1.0–11	3.37	5.34	T [1 0 0]	3.32		30.7
GaP	0.6–10	3.31	4.13	L [1 1 0]	6.32	8.0	29.5
GaP[c]	0.6–10	3.31	4.13	T [1 0 0]	4.13	2.0	16.6
Ge	2.0–20	4.0	5.33	L [1 1 1]	5.5	16.5	482
TiO_2[c]	0.45–6.0	2.58	4.23	L [1 1 0]	7.93	1.0	6.0
$LiNbO_3$[e]	0.33–4.5	2.2	4.64	L [1 0 0]	6.57	1.0	4.6
$LiTaO_3$	0.28–4	2.18	7.45	L [1 0 0]	6.19		0.9
TeO_2[c]	0.35–5.0	2.26	5.99	L [0 0 1]	4.2	6.3	22.9
TeO_2[f]	0.4–4.5	2.26	5.99	T [1 1 0]	0.62	17.9	795
$PbMoO_4$[c]	0.42–5.5	2.39	6.95	L [0 0 1]	3.75	5.5	23.9
Hg_2Cl_2	0.38–28	2.62	7.18	L [1 0 0]	1.62		337
Hg_2Cl_2[e]	0.38–28	2.27	7.18	T [1 1 0]	0.35	8.0	703

[a] L = longitudinal, T = transverse.
[b] $M^* = M/M_{SiO_2}$ is the figure of merit normalized to that of fused silica. $M = n^6p^2/\rho v_s^3$; $M_{SiO_2} = 1.51 \times 10^{-15}(\text{s}^3\,\text{kg}^{-1})$.
[c] Optical polarization perpendicular to the plane of diffraction.
[d] Refractive index at $\lambda = 1.15\,\mu\text{m}$.
[e] Optical polarization neither parallel nor perpendicular to the plane of diffraction.
[f] Circular optical polarization.

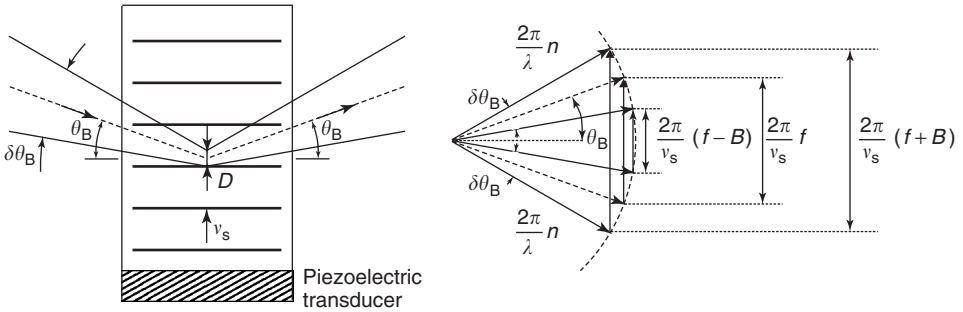


Figure 9 Principle of operation of an AO modulator. The incident optical beam of angular divergence $\delta\theta_B$ is diffracted by a plane acoustic wave with frequency spectrum extending from $f - B$ to $f + B$. Each direction (optical wave vector of a plane wave) within the angular spectrum of the optical beam is matched to a parallel acoustic wave vector with different magnitude. Increasing the modulation bandwidth B requires a corresponding increasing of the angular divergence.

spectrum of $2B$ is $\delta\theta_B = \lambda B/nv_s$ (the angle θ_B being small, $\sin\theta_B \cong \theta_B$). Assuming a plane acoustic wave, $\delta\theta_B$ has to match the divergence angle $\theta_{div} \approx \lambda/nD$, where D is the width of the optical beam inside the modulator. It follows that $B = v_s/D = 1/\tau$, where τ is the transit time of sound across the light beam width. To increase the modulation bandwidth, materials with high sound velocity should be used and the light beam should be focused to a small diameter.

AO modulators are widely used for a variety of applications and present many advantages compared to electrooptic modulators, such as low driving power, simpler electronics, and high extinction ratio; the main drawback is the lower modulation bandwidth. Devices for external laser modulation are available in the visible and in the infrared region, with rise time ranging from few tens to $\sim 2\,\text{ns}$, and efficiency 70–80%. AO devices are also used as laser intracavity modulators (thanks to the high optical quality of materials such as fused silica and $LiNbO_3$) to obtain pulsed operation in various regimes, namely, mode-locking, Q-switching, and cavity dumping. For mode-locking operation, the standing-wave AO modulator introduces a time variable loss with period T_m equal to the inverse of the longitudinal mode spacing, $T_m = 2d_l/c$, where d_l is the laser cavity length.

AO deflectors are based on the property that the direction of the diffracted beam can be varied by changing the driving frequency of the acoustic wave. In fact, starting from the Bragg condition, the deflection angle $\Delta\theta$ between the directions of incident and diffracted beam is found to be proportional to the acoustic frequency change, $\Delta\theta = \lambda\Delta f/nv_s$. To avoid tilting of the sound beam during the angular scan, an acoustic wave with angular divergence $\Delta\theta_s \geqslant \Delta\theta$ is used. A schematic picture of the principle of operation of an AO deflector is shown in **Figure 10**. The interesting parameter here is the number of resolvable spots $N = \Delta\theta/\theta_{div}$, that is, the number of non overlapping angular widths within the angular scan. It is readily obtained that $N = \Delta f(D/v_s) = \Delta f\tau$; the number of resolvable spots is therefore equal to the time–bandwidth product. To increase N, a large transit time τ should be used, but this also limits the random access time which is a measure of the speed of the deflector. There is clearly a trade-off relation between access time and resolution. It is worth noticing that the design requirement of a large τ for high-resolution AO deflector is the opposite of what is needed for large-bandwidth AO modulators.

Relevant applications of AO deflectors include laser scanners and high-resolution/wide-band Bragg cells for optical signal processing. The maximum resolution of AO deflectors is limited to a few thousand; high-performance materials, such as TeO_2 and Hg_2Cl_2, enable 2000–4000 resolvable spots with good efficiency and acceptable access time. Bandwidths of 2 GHz and peak diffraction efficiency in excess of $10\%W^{-1}$ have been demonstrated with $LiNbO_3$ devices.

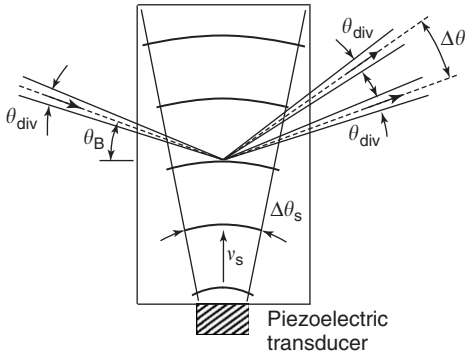


Figure 10 Principle of operation of an AO deflector. The ratio between the deflection angle $\Delta\theta$ and the divergence θ_{div} gives the number of resolvable spots *N*. The sound wave has an angular width $\Delta\theta_s \geqslant \Delta\theta$ to make momentum conservation possible without tilting the sound beam. For a better legibility of the figure, the width *D* of the optical beam inside the modulator is not highlighted.

Acoustooptic tunable filters operate on the principle of acoustooptic diffraction in anisotropic materials. These devices are designed to achieve very narrow bandwidths (from few nanometers to few angstroms) with wide tuning ranges (from few hundreds of nanometers to few microns), obtained by changing the frequency of the applied RF signal. The main advantage is the electronic tunability, but a large driving power is often needed.

AO modulators can also be realized exploiting the interaction between surface acoustic waves (SAWs) and optical modes confined in dielectric waveguides. The confinement of the acoustic power near the surface leads to extremely low modulation power. Integrated AO modulators based on SAW have been fabricated in $LiNbO_3$.

Magnetooptic Modulator

The magnetooptic effect largely used for the modulation of the light is the Faraday effect. It induces a rotation of the polarization plane of an optical beam propagating in a magnetooptic material (such as glass, quartz, zinc sulfide, and water), when a magnetic field is applied parallel to the propagation direction of the optical beam. The angle of rotation is

$$\vartheta = \upsilon \mathscr{B}\, l_m$$

where $\mathscr{B}$ is the magnitude of the magnetic flux, l_m is the length of medium, and υ is the Verdet constant that depends on the specific material and wavelength (see **Table 4** for some representative values). Clockwise or anticlockwise rotation of the polarization plane is related only to the direction of the applied magnetic field (generated by the current circulating in a solenoid) and is independent of that of the light beam. In a way similar to that already illustrated for EO devices, intensity magnetooptic modulators are implemented by inserting a Faraday rotator between two (crossed) polarizers. In this type of device, the modulation frequency is however limited to a few tens of megahertz, due to the inductance of the solenoid.

High-frequency (up to a few gigahertz) magnetooptic modulators, based on a thin film of bismuth-garnet materials (such as MnBi, Bi:YIG, and

Table 4 Typical values of the Verdet constant υ at $\lambda = 589.3$ nm

Material	υ (rad m^{-1} T^{-1})
Quartz (SiO_2)	4.0
Zinc sulfide (ZnS)	82
Glass	
Crown	6.4
Flint	23

$Bi_3Lu_3Fe_5O_{12}$), have been recently developed. The main drawback of bismuth garnets is the strong absorption in the near-infrared spectral region (from 800 nm to $\sim$25 μm), which represents a serious limitation to the diffusion of these devices.

Liquid Crystal Modulators

Liquid crystals (LCs), an intermediate phase between a liquid and a crystalline solid, were first observed in 1888 by Reinitzer. These substances, which typically consist of two benzene (or cyclohexane rings, or phenyl rings) linked with a central group, show the mechanical properties of liquids but the optical properties of crystalline solids, in the sense that the arrangement of the molecules exhibits some structural order. LCs having a liquid–crystalline mesophase variable with temperature are indicated as thermotropic; solutions of substances where the mesophase is also dependent on relative concentrations are called lyotropic. Thermotropic LCs have been widely investigated and their applications to light modulation have reached a mature stage. There are three basic types of thermotropic LCs, termed nematic, cholesteric, and smectic; nematic LCs are the most largely utilized at present.

A major characteristic of all LC compounds is the rod-like shape of their molecules. The preferred time-averaged orientation of the molecular long axes is indicated by a director $\boldsymbol{L}$, a unit vector pointing in such a direction. In the nematic state, the molecules are arranged parallel to each other in an arbitrary direction, with a casual distribution of the centers of mass of the molecules (see **Figure 11a**). In the cholesteric state, the molecules have a defined director $\boldsymbol{L}$ on a specific plane as well, but its orientation is not spatially constant. The molecular structure is characterized by equidistant parallel planes, with $\boldsymbol{L}$ rotated by a fixed angle from one plane to the next (see **Figure 11b**); after a spatial period of pitch P (of the order of the optical wavelength), $\boldsymbol{L}$ retraces the original orientation, resulting in a helicoidal ordering. From this point of view, the nematic order can be regarded as a special case of the cholesteric one with $P \to \infty$. In the smectic state, different types of stratified structures exist; for instance, smectic A has the molecules arranged in layers orthogonal to the long axis of the molecules, with casual distribution of the centers of mass in each layer (see **Figure 11c**); smectic B presents, in addition, an ordering of the centers of mass of the molecules inside each layer, similarly to two-dimensional solids (**Figure 11d**); smectic C has the long axis of the molecules tilted with respect to the layers by an angle which is a function of the temperature (see **Figure 11e**).

One important characteristic of LCs is the optical anisotropy characterized by different dielectric constants $\varepsilon_\parallel, \varepsilon_\perp$ depending on whether the polarization of the optical field is parallel to, or perpendicular to, the director $\boldsymbol{L}$. If $\varepsilon_\parallel > \varepsilon_\perp$, the material is denoted as positive, and can be treated as a positive uniaxial crystal with optical axis parallel to $\boldsymbol{L}$. The optical birefringence is rather strong, with typical value $\Delta n_B = n_\parallel - n_\perp$ ranging from 0.1 to 0.3. Another major property is the optical activity, or rotary

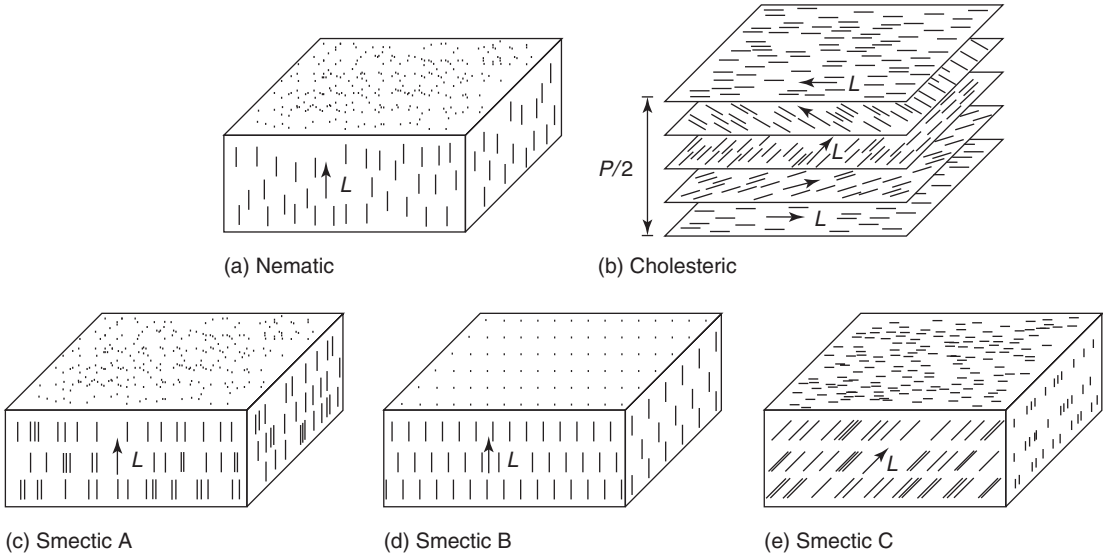


Figure 11 Schematic of the structure of thermotropic liquid crystals: (a) nematic; (b) cholesteric of pitch P; (c) smectic A; (d) smectic B; (e) smectic C.

birefringence, causing the polarization of the impinging optical field (linearly polarized) to rotate its plane of polarization, following the direction of $\boldsymbol{L}$ from the input to the output plane.

Several mechanisms for modulating light using liquid crystals have been developed, including director axis reorientation, nematic–cholesteric phase change, absorption in dichroic dye-LC systems, scattering by LC microdroplets in polymer matrix. The basic principle is related in all cases to the fact that the application of an external electric field to a positive material tends to align the long axis of the molecules parallel to the electric field in order to minimize their energy. Director reorientation is widely used in nematic LCs, where there are two different types of ordering, namely homeotropic and homogeneous, depending on whether the director of the molecules is aligned perpendicular or parallel to the liquid crystal–solid interface. These two forms are produced by a suitable treatment of the solid surface. To obtain homogeneous ordering, microgrooves are produced on the internal surface of the LC cell, usually coated with an optically transparent metallic film such as indium–tin oxide, by rubbing the surface or evaporating thin SiO_2 stripes. A thin layer of an initially homogeneously ordered LC sandwiched between two solid interfaces can be changed into a homeotropic state, when an electric field is applied perpendicularly to the liquid crystal–solid interface. The process is illustrated in **Figure 12** for increasing values of the magnitude of the electric field $\boldsymbol{E}$. If $E > E_{th}$, the molecules begin to align along the field direction; when $E \gg E_{th}$, the majority of the molecules are aligned perpendicularly to the interface.

When a linearly polarized light passes through a homogeneously ordered LC, if the polarization axis forms an angle (usually 45°) with the director, phase retardation occurs due to the birefringence of the material; for a given thickness of the LC layer, the birefringence depends on temperature, wavelength, and applied field. In particular, for $E \gg E_{th}$, birefringence and thus phase retardation sharply decrease as the field is increased, up to a saturation value. Exploiting the large phase changes that can be obtained with small voltage modulations (few volts), intensity light modulators can be fabricated by sandwiching the LC cell between two crossed polarizers. Pure phase modulators and tunable phase retardation plates can be obtained with parallel aligned LC cells as well.

In a twist nematic alignment, the opposite walls of the cell containing the positive LC material are treated to produce a homogeneous arrangement in which the directors at the interfaces are twisted with an angle. Different twist alignments are employed according to the twist angle, usually 45°, 90°, and $>90°$ (supertwist nematic, e.g., 270°). Most common LC displays are based on a 90°-twisted nematic cell. When a linearly polarized beam passes through a 90°-twisted nematic cell, the rotary birefringence causes the polarization plane to follow the twist of the LC directors and to undergo a 90° rotation. If an electric field $E \gg E_{th}$ is applied to the twisted cell, the molecular alignment becomes mainly homeotropic and the light polarization is almost not influenced by the LC layer. A schematic of the changes in a twisted nematic cell under an external electric field is shown in **Figure 13**. A light modulator can be realized by sandwiching a twisted nematic cell between two

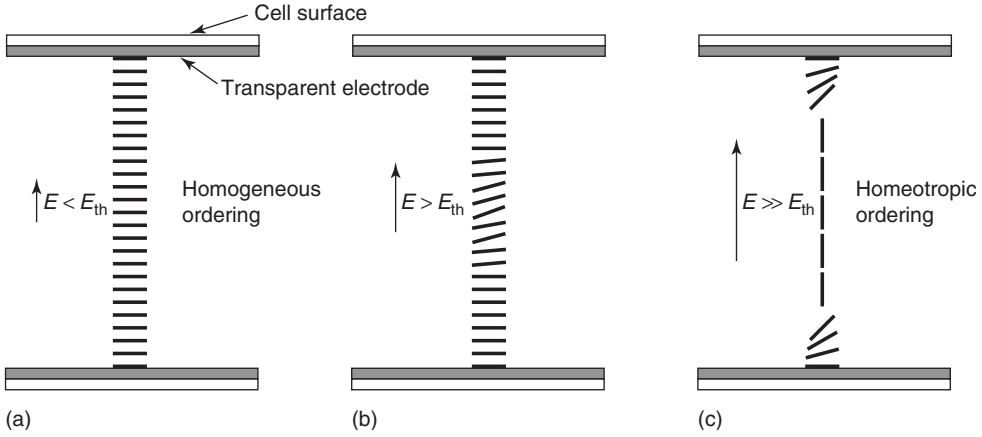


Figure 12 Illustrating director axis reorientation as an increasing electric field $\boldsymbol{E}$ is applied to an initially homogeneously ordered LC in a direction perpendicular to the liquid crystal – cell interface. (a) For $E < E_{th}$ the homogeneous ordering is not affected; (b) for $E > E_{th}$ the molecules begin to align along the field direction; (c) for $E \gg E_{th}$ the molecules are almost completely aligned in the direction of the electric field and the LC cell shows essentially homeotropic ordering.

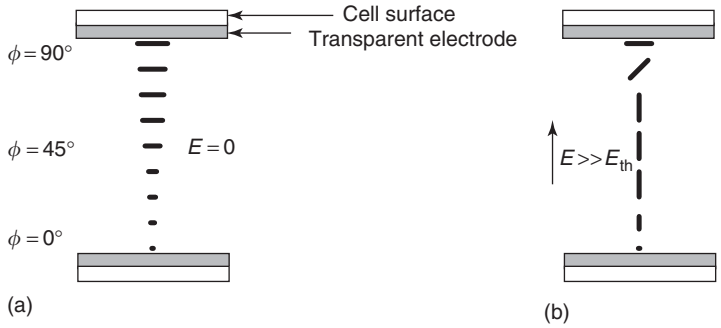


Figure 13 Director axis orientation in a 90° twisted nematic liquid crystal cell. (a) Applied field $E = 0$; (b) $E \gg E_{th}$. φ is the twist angle.

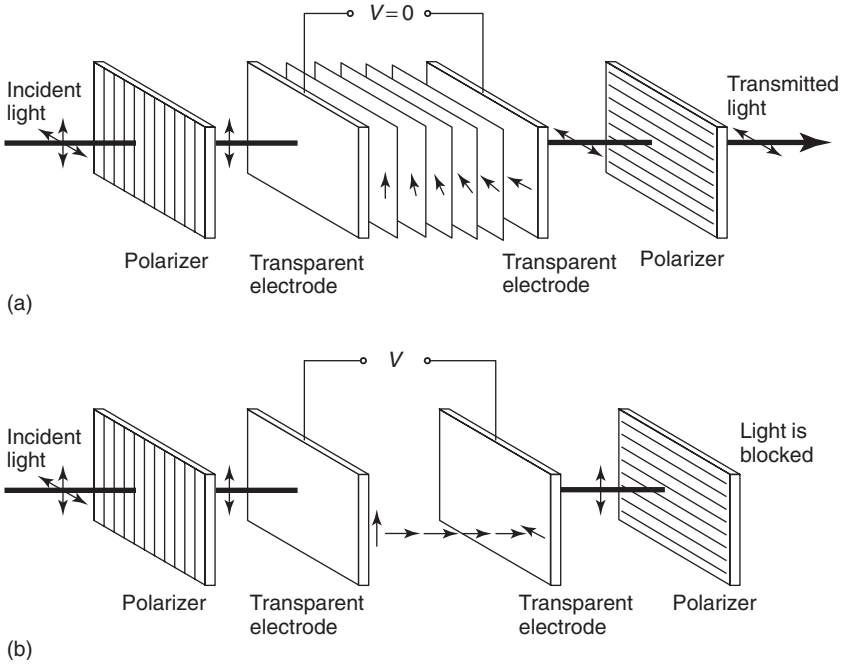


Figure 14 Operation of a twisted nematic LC cell. (a) In the absence of an electric field, the cell acts as a polarization rotator and light is transmitted; (b) in the presence of an electric field, homeotropic order of molecules does not affect the polarization, and light is blocked. V is the voltage applied to the LC cell.

crossed polarizers with axes parallel to the directors of the particular cell interface they are next to (see **Figure 14**). The typical cell thickness is $\sim 10\,\mu m$ or less. In the absence of an external electric field, the polarized beam is transmitted with the polarization plane rotated by 90°, as illustrated in **Figure 14a**; when $E \gg E_{th}$, the cell rotatory power is inhibited and the beam is blocked by the second polarizer, as shown in **Figure 14b**. Operation in the reflection mode is also possible, if a reflecting surface is placed after the second polarizer.

Dicroic dye-LC systems are formed by dissolving a suitable dye in the liquid crystal. The rod-shaped dicroic dye molecules strongly absorb the light polarized parallel to the long axis direction, showing only a weak absorption in the orthogonal direction. The type of bonding between the dye and the liquid crystal molecules tends to align the former parallel to the director and to follow its orientation. The transmission of the solution can be varied by reorienting the liquid crystal director by an external electric field, resulting in intensity modulation of the incident light.

Submicron droplets of LCs dispersed in a polymer matrix can also be used for intensity modulation. It is to be assumed that the ordinary index $n_\perp$ of the LC is equal to the refractive index n_p of the polymer (isotropic), while the extraordinary index $n_\parallel > n_p$. Due to a random orientation of $\boldsymbol{L}$ for each droplet and the consequent refractive index mismatch with the polymer, the LC droplets act as scattering centers for the incoming light, and the transmission through the cell is low. By applying a suitable electric field to the

polymer matrix, the directors of the droplets can be oriented so that the impinging optical field, parallel to the direction of the extraordinary axis (i.e., parallel to $\boldsymbol{L}$), propagates as an ordinary ray through the LC droplets, experiencing a refractive index $n_{\perp} = n_{\mathrm{p}}$; in this case, there is almost no scattering and the light beam is transmitted.

A major limitation of LC optical modulators is the limited time response, which depends on several parameters of the LC such as rotational viscosity, elastic constant, threshold electric field, on-cell geometry, and on the applied electric field (as far as the rise time is concerned). In nematic LCs, the rise and decay times are generally rather slow, of the order of milliseconds. Faster response time can be obtained using cholesteric materials; recently, smectic ferroelectric LC cells have been developed with response time from a few microseconds to a few hundreds of microseconds.

The main application of LC cells regards spatial modulation of light, and is concerned at present with display fabrication, mainly with display elements for digital watches, cellular phones, pocket calculators, flat panels, and light switches. Compared to other types of displays, such as light emitting diodes and plasma displays, the principal advantage of LC displays is their low electrical power consumption, due to the passive nature of the display. They however have a number of disadvantages, besides the relatively slow response time, including optical efficiency, limited by the use of polarizers which absorb at least 50% of the unpolarized incident light, and the angle of view limited by the reduction in the contrast of modulated light when the angle of incidence/reflection is increased. Two common LC display formats are the seven-bar-segment display used to form numbers from 0 to 9, and the 7×5 matrix display used for more complex characters.

Electroabsorption Modulator

Electroabsorption effect is related to the modification of the absorption coefficient, α, of a material induced by an external electric field. This effect permits the direct control of the intensity of a light beam passing through the material. Since the refractive index and the absorption coefficient are related by the Kramers–Kronig dispersion relation, when an electric field induces a change in the refractive index of a material (electrorefraction), the absorption coefficient also changes. The Kramers–Kronig relation,

$$n(\omega) = 1 + \frac{c}{\pi}\,\mathrm{P.V.}\int_0^{\infty} \frac{\alpha(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'$$

where P.V. stands for Cauchy principal value of the integral, is completely general and applies to any type of materials. When the photon energies of the incident light is close to the bandgap of the material, the electroabsorption effect is strongly enhanced. High-efficiency electroabsorption modulators are based on semiconductor materials (such as GaAs, InGaAs, and InP) both in bulk and in multiple quantum well structures. In the bulk structure, the applied electric field changes the effective bandgap of the semiconductor (Franz–Keldysh effect), thus increasing the absorption of the semiconductor for frequency lower than the bandgap. Quantum wells are thin (~ 10 nm or less) layers of small bandgap semiconductor material, sandwiched between larger bandgap semiconductors (barriers). A multi-quantum well (MQW) structure consists of alternating small and large bandgap layers of semiconductors. In these structures, the applied electric field shifts the energy levels, producing a similar shift of the absorption edge toward lower frequencies (quantum confined Stark effect).

To realize an intensity modulator based on the electroabsorption, the frequency of the optical beam (i.e., the energy of the incident photons) has to be just below the absorption edge (i.e., the bandgap of the semiconductor). In this way, the light is transmitted through the modulator when the applied electric field is off. When the electric field is increased, the absorption spectrum shifts toward lower frequency (red shift) and the modulator switches from transparent to highly absorbing, reducing the intensity of the transmitted light.

The main advantage of these devices is that they can be directly integrated with the conventional high-speed microelectronic components as well as with semiconductor diode lasers. These integrated optoelectronic structures constitute the key components for the realization of photonic devices needed in wide-band optical communication systems

See also: Crystal Optics; Liquid Crystals; Nonlinear Optics; Optical Fibers; Optical Instruments; Optical Microscopy; Optical Properties of Materials; Semiconductors, Optical and Infrared Properties of.

PACS: 78.20Hp; 78.20Jq; 78.20Ls; 42.79Jq; 42.79Kr; 42.79Ls; 85.70Sq; 42.70Df; 42.25Lc

Further Reading

Agulló-López F, Cabrera JM, and Agulló-López F (1994) *Electrooptics: Phenomena, Materials and Applications*. New York: Academic Press.

Chandrasekhar S (1992) *Liquid Crystals*, 2nd edn. Cambridge: Cambridge University Press.
De Gennes P-G and Prost J (1998) *Physics of Liquid Crystals*, 2nd edn. Oxford: Clarendon Press.
Ericksen JL and Kinderlehrer D (1987) *Theory and Applications of Liquid Crystals*. New York: Springer.
Gordon EI (1966) A review of acoustooptical deflection and modulation devices. *Proceedings of the IEEE* 54: 1391–1401.
Kaminow IP (1974) *An Introduction to Electrooptic Devices*. New York: Academic Press.
Korpel A (1988) *Acousto-Optics*. New York: Dekker.
Lines ME and Glass AM (1977) *Principles and Applications of Ferroelectrics and Related Materials*. Oxford: Clarendon Press.
Magdich LN and Molchanov VY (1989) *Acoustooptic Devices and Their Applications*. New York: Gordon and Breach.
Miller DAB and Wiener JS (1986) Electric-field dependence of linear optical properties of quantum well structures: waveguide electroabsorption. *IEEE Journal of Quantum Electronics QE* 22: 1816–1830.
Moss TS, Burrel GS, and Ellis B (1973) *Semiconductor Opto-Electronics*. New York: Wiley.
Pankove JI (1975) *Optical Processes in Semiconductors*. New York: Dover.
Sittig KE (1972) Elastooptic light modulation and deflection. In: Wolf E (ed.) *Progress in Optics*, vol. 10, ch. VI. Amsterdam: North-Holland.
Tsai CS (1990) *Guided-Wave Acoustooptics*. Berlin: Springer.
Yariv A (1991) *Optical Electronics*, 4th edn., ch. 9. New York: Holt, Rinehart and Winston.

Molecular Clusters

R A Broglia, University of Milan, Milan, Italy

Introduction

Nanostructured materials are made out of atoms as their more common forms, but the atoms are arranged in nanometer or sub-nanometer size units, which become the building blocks of the new materials. A prototype of these building blocks is the C_{60} fullerene. These tiny grains, atomic aggregates or clusters, in which the surface plays a paramount role, respond to external influences as light, mechanical stress, and electric fields quite differently from bulk matter. Hence, nanostructured materials display an array of novel attributes and properties.

To learn how to produce customer-tailored nanometer materials, one needs to have a thorough knowledge of the physical properties of the associated building blocks, that is, of atomic aggregates. For example, small silver particles dissolved in the stained glasses of medieval cathedrals led to the blues, the distinct ruby being produced by the presence of gold particles.

When trying to understand the color of these particles, it is not enough to refer to the optical response of either the individual atoms or the bulk metal crystals. This is because the way the properties of a solid gradually evolve as atoms are brought together to form increasingly larger units, is far from being trivial. The end points of such evolution is reasonably well understood, but the knowledge of the intermediate situations still remains poor. A sodium atom has a very simple absorption spectrum, as shown schematically in **Figure 1**. It consists of essentially one line in the visible, the well-known yellow light. In quantum physics, this phenomenon is well described as a one-electron transition from a quantal state, known as 3*s*, to an excited state, denoted by 3*p*. A sodium crystal, on the other hand, has a completely different spectrum. The absorption is strong in the infrared, goes through a minimum in the visible, and then rises again in the ultraviolet. The reason for this is that very low energy photons can excite electrons from the continuum of states just below the energy of the last occupied state to states just above the Fermi energy. The strong ultraviolet absorption is caused by interband transitions.

Suppose a small corner of a crystal is chipped off, producing a microcrystal that contains only eight atoms. Again, the absorption spectrum changes completely. A relatively broad absorption maximum displaying different lines (resonances) appears in the visible region. This absorption is due to a collective excitation of outer electrons called a plasmon excitation (collective excitation of valence electrons). It can be viewed as a collective sloshing motion of the electrons from one side of the microcrystal to the other side. Such a motion has a strong dipole moment, behaving as an antenna which gives rise to a strong absorption band in the visible region (see **Figure 2**).

The absorption of light by a sodium atom is well described within the picture of independent particle motion. In this picture, each electron moves in the atom as if it were alone. It is only when an electron tries to leave the atom that it feels that its motion is confined due to the presence of a wall produced by the pullings and pushings of all the other electrons and of the positively charged atomic nuclei. That is, in the independent particle model, the electrons move in the average potential generated by the interaction

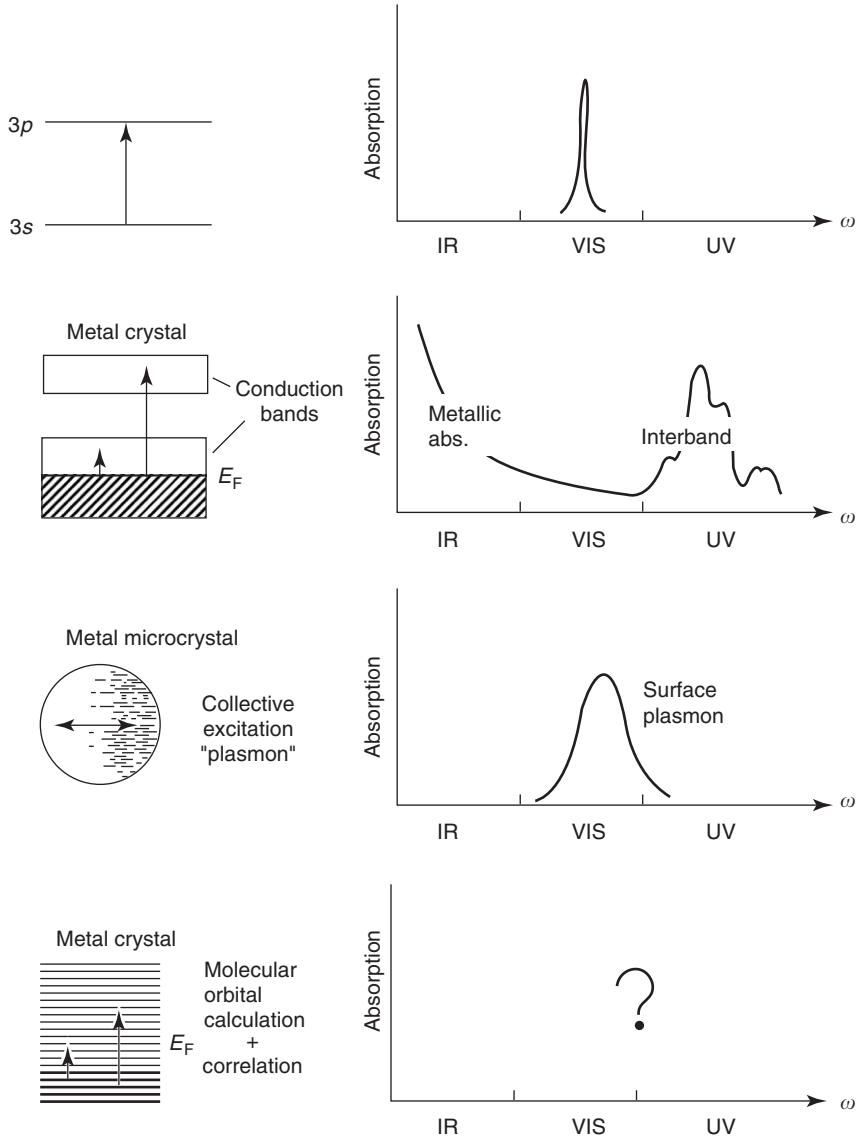


Figure 1 Absorption of light by sodium in different stages of aggregation.

with the other electrons and with the ions. This picture, however, cannot describe even qualitatively the resonance phenomenon (plasmon excitation) that dominates the optical response of an eight-atom cluster. Between one and eight atoms, one must completely change one's way of looking at the optical response of the system.

Because the response to light is quite different in Na_8 than in bulk Na, one would expect that a material made out of small sodium aggregates will display quite a color difference from that of a sodium crystal. It is likely that it will also display a number of other differences, in particular, regarding specific heat, conductivity, elasticity, etc. This is also expected to be true for other metallic aggregates (e.g., Ag- and Au-clusters).

Some of the most promising building blocks of nanophase materials are fullerenes (see **Figure 3**), molecules which have been found to maintain most of their intrinsic characteristics when placed inside an infinite crystalline lattice, whether they form van der Waals or covalent solids. Color, strength, transport properties, and other features of these materials depend on the properties of the isolated molecules, in particular, on the strength with which electrons couple to ionic vibrations (phonons). The mechanisms which are at the basis of this coupling are also important in understanding the properties of carbon nanotubes and linear carbon chains, remarkable examples of molecular quantum wires. This is in keeping with the fact that nanotubes can be obtained by bisecting a fullerene molecule at the equator,

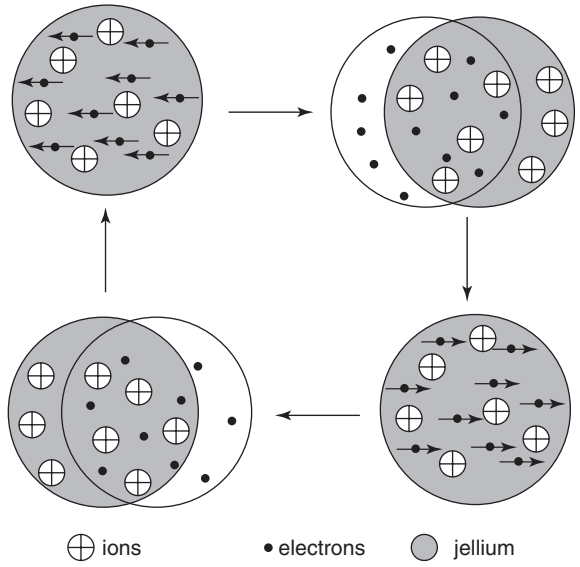


Figure 2 Schematic representation of the collective motion of valence electrons (plasmons) in a metal cluster.

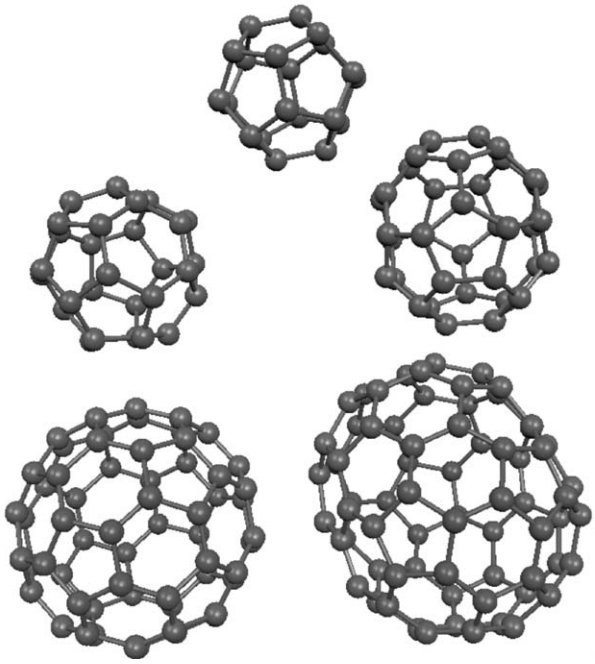


Figure 3 Selected fullerenes: molecular structure for (from top to bottom, and from left to right) C_{20}, C_{28}, C_{36}, C_{60}, and C_{70}.

joining the resulting hemispheres with a cylindrical tube made out of an arbitrary number of belts, each built out of benzoid rings. While single-wall C_{20}-derived nanotubes are the thinnest capped nanotubes which can be thought of, linear chains made out of carbon atoms are the ultimate examples of atomic quantum wires, playing a role in such different phenomena as the presence of diffuse interstellar bands (DIF) in the spectra of distant stars, and, arguably, in the field emission of open-end carbon nanotubes.

Electronic Levels

The basic starting point in the study of atomic aggregates such as metal clusters, fullerenes, nanotubes, and linear carbon chains is the mean field theory, one of the most useful approximations in all of physics and chemistry. In it, the many-particle Schrödinger equation is replaced by a single-particle Schrödinger equation.

For example, the Hartree–Fock equation

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + v_H(r)\right]\varphi_j(\boldsymbol{r}) + \int d^3r' U_x(\boldsymbol{r},\boldsymbol{r}')\varphi_j(\boldsymbol{r}') = \varepsilon_j\varphi_j(\boldsymbol{r}) \quad [1]$$

where

$$v_H(\boldsymbol{r}) = \int d^3r' \varrho(\boldsymbol{r}')v(|\boldsymbol{r}-\boldsymbol{r}'|) \quad [2]$$

with

$$\varrho(\boldsymbol{r}) = \sum_{i\in occ} |\varphi_i(\boldsymbol{r})|^2 \quad [3]$$

and

$$U_x(r,\boldsymbol{r}') = -\sum_{i\in occ} \varphi_i^*(\boldsymbol{r}')v(|\boldsymbol{r}-\boldsymbol{r}'|)\varphi_i(\boldsymbol{r}) \quad [4]$$

The functions v_H and U_x are the Hartree and the Fock (exchange) potentials, while $\varphi_i(\boldsymbol{r})$ are the single-particle wave functions. These equations can be solved self-consistently and the total energy of the system calculated. The result is however not sufficiently accurate, in that it lacks the renormalization effects of the single-particle motion arising from correlation effects. Taking these correlation effects adds an ω-dependent correlation potential to $v_H(r)$. Within the local density approximation (LDA), this potential reads

$$v_c = \frac{\partial E_{corr}(n)}{\partial n} = -0.91\ln\left(1+\frac{11.41}{r_s}\right)eV \quad [5]$$

where $E_{corr}(n)$ is the energy of the interacting electron gas, n the density functional of the system, $r_0 = r_s a_B$ the Wigner–Seitz radius while $a_B = \hbar^2/m_e e^2 = 0.529$ Å is Bohr's radius. Within the same approximation,

$$\int d^3r' U_x(\boldsymbol{r},\boldsymbol{r}')\varphi_j(\boldsymbol{r}') \approx v_x(r)\varphi_j(\boldsymbol{r})$$

with

$$v_x(r) = -e^2\frac{1}{\pi}(3\pi^2 n(\boldsymbol{r}))^{1/3} \approx -0.985e^2n^{1/3}$$

The Hamiltonian resulting from the sum of the two terms in the square brackets of eqn [1] plus v_c and v_x, as well as the external field generated by the ions, is known as the Kohn–Sham Hamiltonian. Its solution provides accurate results not only for bulk matter, but also for atomic and molecular aggregates.

What started out as a problem too complicated to solve for all but the smallest systems becomes quite manageable when one deals with one particle at a time. The many-body effects come in via the single-particle potential, which is generated from the particles themselves. The resulting self-consistent mean field theory has been enormously successful in many domains of quantum mechanics. Whenever a many-particle system exhibits single-particle behavior, which happens quite commonly, the mean field theory is likely to be useful.

In the case of metal clusters, fullerenes, nanotubes, linear carbon chains etc., there are a number of observations which testify to the validity of the independent particle picture. In fact, the dependence of quantities such as mass abundance and ionization potential vary strongly with the number of atoms (see **Figures 4** and **5**). In particular, the energy needed to free one electron from the cluster exhibits a distinct shell structure with extreme values at prescribed number of electrons. These discontinuities are strongly connected with the independent particle picture already mentioned. In fact, the orbits in which an electron moves in an average potential display a bunchiness as a function of the number of electrons which can be viewed as layers enveloping the ions, much like the configuration of an onion. These layers or shells each contain a number of orbits and thus a given number of electrons, the so-called magic numbers. Special stability is ascribed to the filling of each of these shells. In particular in the case of C_{60}, the single-particle energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is 1.8 ± 0.1 eV, a quantity much larger than room temperature (≈ 25 meV), a fact which testifies to the stability of this system (see **Figure 6**).

Another example of a closed-shell system is provided by Na_8. The lowest energy structures of this system calculated within the LDA are displayed in **Figure 7**. These isomers are characterized by the symmetries associated respectively with the groups $\mathscr{D}_{2d}$ (ground state (GS)), $\mathscr{D}_{4d}$ (89 meV), and $\mathscr{T}_d$ (95 meV) – the figures in parentheses being the total energies calculated in LDA and expressed with respect to the ground-state energy. The Kohn–Sham energy levels for the $\mathscr{D}_{2d}$ system are reported in **Table 1**.

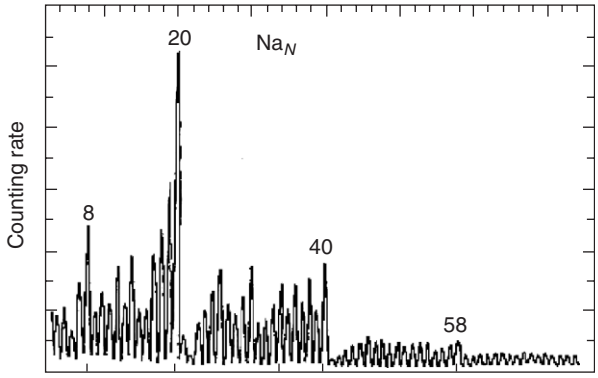


Figure 4 Na clusters abundance.

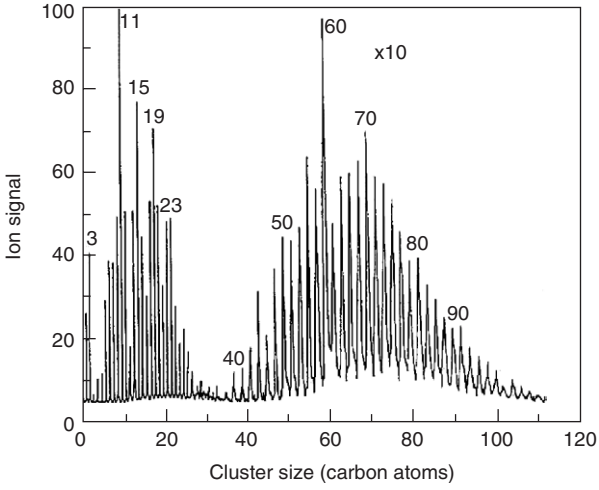


Figure 5 Carbon clusters abundance.

In the Born–Oppenheimer (BO) approximation (also known as adiabatic approximation) used in deriving the LDA equations describing the electronic motion, the equations which determine the electronic state are decoupled from those related to the ionic dynamics. In other words, the atomic motion of the system (molecule, cluster, or solid) is studied under the hypothesis that the electronic system always remains in the electronic ground state associated with the instantaneous geometrical configuration (e.g., **Figure 7**). One can then analyze the solution of the problem of atomic motion in the case of small oscillations near the equilibrium configuration (i.e., when every atom stays always near its equilibrium position, and there is no atomic diffusion). In this case, writing down a second-order expansion of the total potential felt by the ions (harmonic approximation), one can show that it is possible to reduce

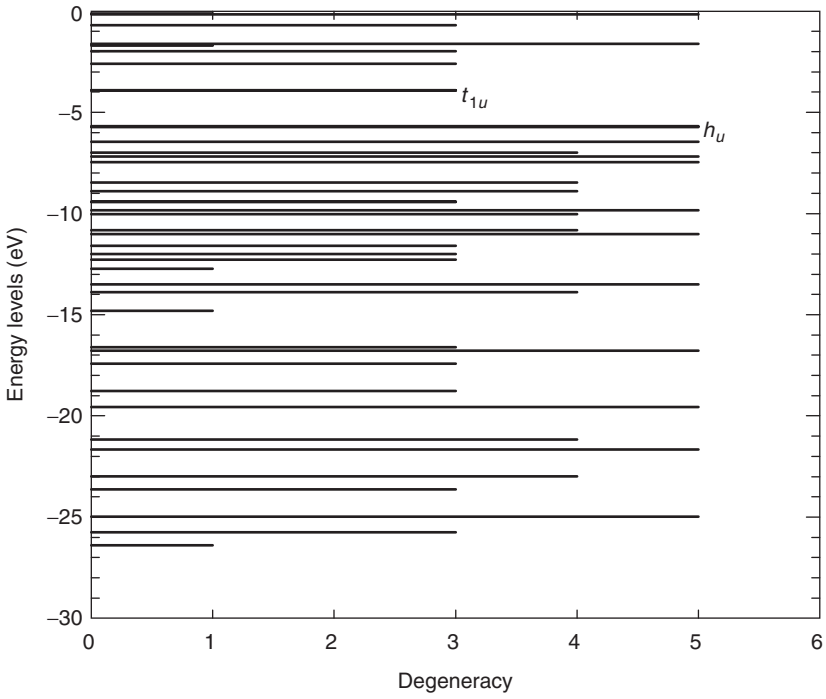


Figure 6 Energy levels of C_{60} calculated in LDA. The HOMO (h_u) and LUMO (t_{1u}) levels are indicated.

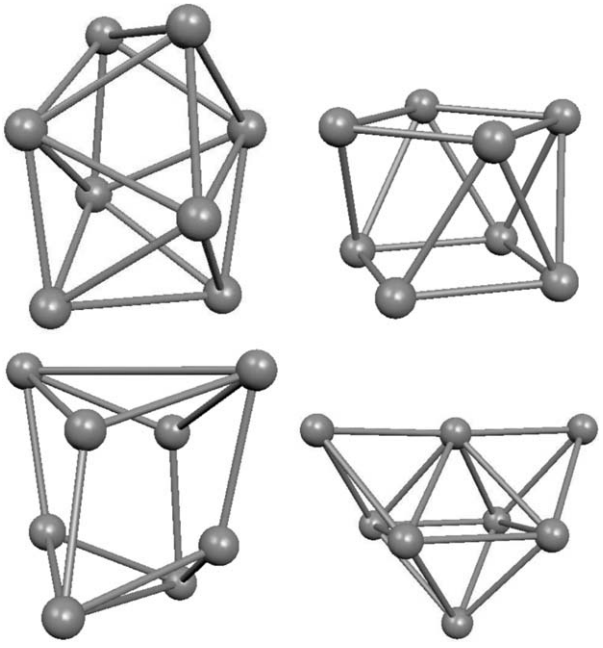


Figure 7 Ionic structure of Na_8 in its three lowest energy isomers with symmetry: $\mathscr{D}_{2d}$ (top left), $\mathscr{D}_{4d}$ (top right), $\mathscr{T}_d$ (bottom left). Also shown, the "close-packed" structure (bottom right).

Table 1 Kohn–Sham energy levels of Na_8 for its ground state configuration $\mathscr{D}_{2d}$

E (eV)	*Degeneracy*	*Occupation*
−4.75	2	2
−3.54	2	2
−3.11	4	4
−2.00	2	0
−1.84	4	0
−1.54	2	0
−1.50	2	0
−1.39	2	0
−0.65	2	0
−0.47	4	0

the problem to that of a collection of independent harmonic oscillators. The general solution is then a superposition of $3N_{at}$ normal modes of vibration, each of them having its own frequency and its own eigenvector, obtained by diagonalizing the dynamical matrix. The quantum description of this set of independent harmonic oscillators leads to the concept of phonons.

The harmonic approximation is generally a good one for solid-state systems, molecules or atomic aggregates, at sufficiently low temperature. In the case of carbon systems (graphite, diamond, fullerenes), which are characterized by a quite strong covalent bonding, the harmonic approximation is still very good at room temperature.

In the case of atomic and molecular clusters, it is possible to generalize the arguments of the previous section by discussing the general properties of the solution for the ionic motion, valid independently of the actual form of the total potential U acting on the ions: in particular, in order to determine the different

parameters of the model, namely the first and second derivatives of U, and thus the dynamical matrix.

It emerges that this matrix can be obtained from a knowledge of the derivatives, with respect to the independent $3N_{at}$ (Cartesian) atomic displacements, of the force exerted on the atoms. Such derivatives can be computed numerically, evaluating the change of the component β of the force generated on the atom m, when the atom n is displaced by a small amount in the direction α. This can be done by minimizing the DFT functional with respect to the electronic degrees of freedom and hence evaluating the forces on the ions, after having slightly displaced a given atom along a given Cartesian direction. The process must be repeated for the three components x, y, and z of every atom. Since the ionic displacements are small, the requested set of $3N_{at}$ electronic minimizations is largely facilitated if one starts them from the electronic configuration, which corresponds to the equilibrium ionic positions. Many practical applications of the above method can be found in the recent literature. A typical example from the field of carbon clusters is the case of C_{60}. Calculations confirm that good agreement with respect to the experimentally measured vibrational spectra can be obtained by using the Car–Parrinello method in the DFT–LD approximation. Similarly, the cluster C_{28} has been studied at the same level of accuracy. Its normal modes, which couple to the LUMO electronic orbital (i.e., those with A_1, E, and T_2 symmetry), are reported in **Table 2**.

Another interesting case where the *ab initio* methods discussed above can be applied is that of metal aggregates. Among them is the closed-shell Na_8 cluster. In this case, there are 14 distinct nonzero vibrational eigenvalues, four of them being doubly degenerate. Making use of the Car–Parrinello molecular dynamics (CPMD) formalism, one obtains, for the $\mathscr{D}_{2d}$ symmetry configuration, the results displayed in **Table 3**.

Figure 8 reproduces the atomic displacements associated with the highest totally symmetric mode (16.7 meV). The displacement field associated to specific modes is an important physical quantity, to which *ab initio* calculations give direct access. The displacement fields (vibrational eigenvectors) are in fact hardly accessible from the experiment; they are, however, very important objects, since they govern the size of the coupling with the electronic states (electron–phonon coupling).

Table 2 A_1, E, and T_2 normal modes of the cluster C_{28} calculated within DFT–LDA

$1/\lambda$ (cm^{-1})	
351	E
391	T_2
524	T_2
565	A_1
570	E
607	E
707	T_2
724	T_2
763	A_1
771	T_2
791	T_2
976	E
983	T_2
1093	T_2
1101	A_1
1116	E
1171	A_1
1191	T_2
1220	A_1
1260	T_2
1306	E
1381	T_2
1414	E

Table 3 Calculated frequencies for the Na_8 vibrational modes

	ω_v (meV)	$1/\lambda$ (cm^{-1})
B_1	5.84	47.07
B_2	6.85	55.26
A_1	7.93	63.92
E	8.77	70.74
E	9.36	75.52
A_1	9.48	76.45
A_2	10.2	82.15
B_2	12.3	99.50
E	14.1	113.75
B_1	14.9	119.83
A_1	15.9	128.57
A_1	16.7	134.83
E	17.5	140.95
B_2	18.6	150.00

Plasmons

If a weak external time-dependent perturbation is applied to the electronic system (e.g., a low-intensity electromagnetic field of a sufficiently high-frequency, such that one can neglect the response of the ionic degrees of freedom), the electronic density will start oscillating around the minimum energy configuration. For such weak external fields, one can, to a very good approximation, retain only the linear part of the response. In this approximation, the frequency of the induced oscillations will be determined by the curvature of the energy surface around its minimum. As the density oscillates, the effective potential changes and, in turn, induces a change in the density itself. Therefore, to describe these electron "vibrations," one must require self-consistency between

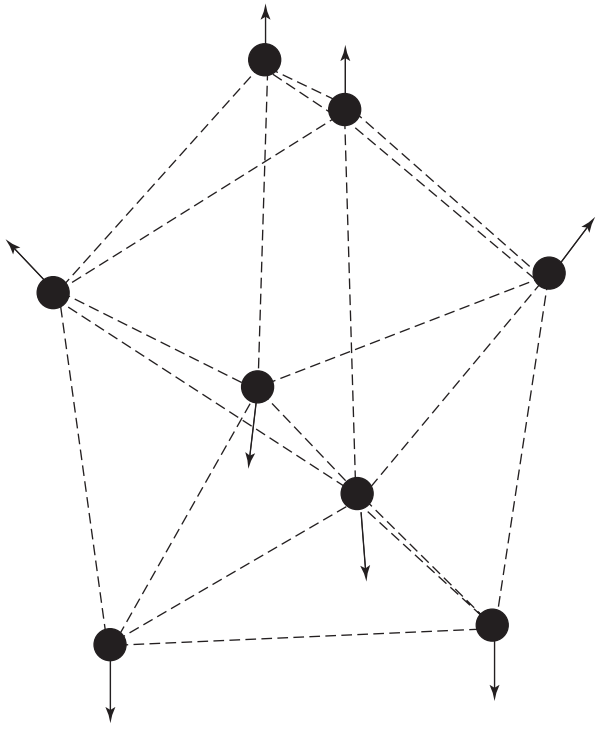


Figure 8 Displacement associated with the A1 phonon in Na_8 at energy 16.7 meV.

the variations of the density and those of the effective potential. In this way, a time-dependent theory can be obtained starting, for example, from Hartree, Hartree–Fock, or LDA energy functional within the linear response approximation. However, the first two approaches are intrinsically based on a mean-field approximation and neglect correlations beyond those arising from the Pauli principle. On the other hand, the time-dependent extension of the LDA (TDLDA) could also be derived, making a local approximation again, from an exact theorem leading to the so-called time-dependent DFT (TDDFT).

Within this scheme, the response of the system to a (weak) time-dependent field

$$V_{\text{pert}}(\boldsymbol{r},t) = V_{\text{pert}}(\boldsymbol{r})\exp(-\mathrm{i}\omega t) + \text{hc} \qquad [6]$$

is expressed in terms of the perturbed density

$$\varrho(\boldsymbol{r},t) = \varrho_0(\boldsymbol{r}) + \delta\varrho(\boldsymbol{r},\omega)\cdot\exp(-\mathrm{i}\omega t) + \text{hc} \qquad [7]$$

where ϱ_0 is the ground-state density, and where

$$\delta\varrho(\boldsymbol{r},\omega) = \int \mathrm{d}^3r'\ \Pi(\boldsymbol{r},\boldsymbol{r}',\omega)V_{\text{pert}}(\boldsymbol{r}) \qquad [8]$$

The quantity $\Pi(\boldsymbol{r},\boldsymbol{r}',\omega)$ is the so-called RPA density–density correlation function.

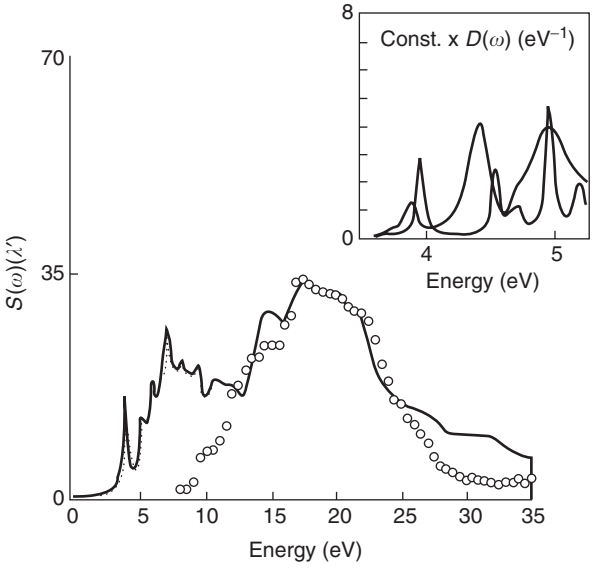


Figure 9 TDLDA response (full line) compared to the experimental points (open circles) (to be noted that the cutoff at $\approx$10 eV is due to the experimental setup). Inset: detailed spectrum at low energies compared to experimental results (broken curve).

From the induced density $\delta\varrho(\boldsymbol{r},\omega)$, one can extract physical quantities of interest such as the dynamical linear polarizability, $\alpha(\omega)$, according to

$$\alpha(\omega) = \int \mathrm{d}^3r\ \delta\varrho(\boldsymbol{r},\omega)V_{\text{pert}}(\boldsymbol{r}) \qquad [9]$$

The excitation of the system by a short-duration perturbation is conveniently described by the strength function $\mathscr{S}(E)$, defined by

$$\mathscr{S}(E) \equiv \sum_n |\langle n|V_{\text{pert}}|0\rangle|^2\delta(E-E_n) \qquad [10]$$

where $E = \hbar\omega$ is the excitation energy, and $|0\rangle$ and $|n\rangle$ denote the ground and excited states of the system, respectively.

The strength function is related to the cross section for photon absorption by the following equation (Golden rule):

$$\sigma(E) = \frac{4\pi^2}{\hbar c}E\mathscr{S}(E) = -\frac{4\pi}{\hbar c}E\ \text{Im}\,\alpha(E) \qquad [11]$$

and the total cross section is given by

$$\int \mathrm{d}E\,\sigma(E) = \frac{2\pi^2e^2\hbar}{m_e c}N \qquad [12]$$

Making use of the single-particle basis displayed in **Figure 6**, the dipole response of C_{60} has been worked out in the TDLDA, and the results are shown in **Figure 9**.

Theory provides an overall account of the experimental findings, also for the width of the plasmon (Mie) resonance located $\sim 17\,\text{eV}$ and amounting to $\sim 12\,\text{eV}$. Consequently, Landau damping, that is the decay of the collective modes into single (unperturbed) particle–hole excitations, can be viewed as the main relaxation mechanism of the above described collective plasmon mode.

The experimental data at low excitation energy correspond to the optical/UV absorption spectra of C_{60} in hexane solution, at room temperature.

The present approach yields a static polarizability of the system of $88.6\,\text{Å}^3$, which can be compared to measurements of $\alpha(0) \approx 80\,\text{Å}^3$ for a single molecule.

Coupling of Electrons to Phonons and to Plasmons

So far, electrons and ions have been considered as independent degrees of freedom. Ions have been taken either as fixed in their ground-state configuration or as moving under the action of a phonon displacement field. The electrons have been treated within the BO approximation. This means that it has been assumed that they are able to follow the evolution of the ionic configuration adiabatically. This approximation (based on the fact that typical frequencies obey $\omega_{\text{electrons}} \gg \omega_{\text{ions}}$) breaks down in a number of cases in which the ionic motion induces correlations between electronic states, in particular, in connection with the mass enhancement factor observed in specific heat measurements as well as in the case of electron–electron pairing correlations at low temperatures.

Furthermore, although the renormalization effects due to the coupling of electrons to plasmons are in principle included in the exact DFT, in practice they are incorporated only to some extent in the correlation potential of the LDA, where important effects of this coupling are left out. Moreover, the exchange of plasmons between electrons, which is the basic mechanism leading to the screening of the Coulomb field, also gives rise to an electron self-energy whose frequency dependence is essential to obtain a proper description of this phenomenon.

In other words, the Hamiltonian to be diagonalized is

$$H = H_{\text{elec}} + H_{\text{phon}} + H_{\text{el,phon}} + H_{\text{plas}} + H_{\text{el,plas}} \qquad [13]$$

where the electron, phonon, and plasmon (H_{elec}, H_{phon}, and H_{plas}, respectively) terms are already diagonal, while $H_{\text{el,phon}}$ and $H_{\text{el,plas}}$ can be diagonalized in the product basis, provided by electron and phonon states, and electron and plasmon states. The corresponding matrix elements g_ν of $H_{\text{el,phon}}$ (see **Figure 10**) lead to the partial electron–phonon coupling constants

$$\lambda_\nu = \alpha \frac{g_\nu^2}{\hbar\omega_\nu} N(0) \qquad [14]$$

where α is a statistical factor depending on the symmetry of the modes ν, $\hbar\omega_\nu$ is the corresponding energy, while $N(0)$ is the density of levels at the Fermi energy in the solid built, making use of the clusters of interest. **Tables 4** and **5** give the corresponding

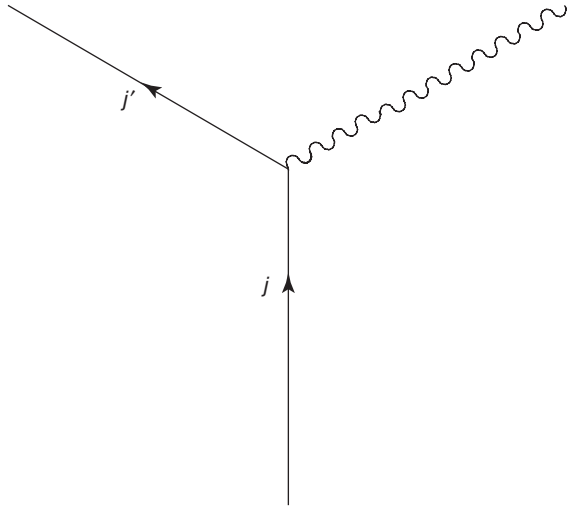


Figure 10 Graphical representation of the electron–phonon or electron–plasmon coupling, where the arrowed line represents the electron while the wavy line stands for either the phonon (ν, μ) or the plasmon (n).

Table 4 Experimental partial electron–phonon coupling constants $\lambda_\nu / N(0)$ in meV which measures the strength with which an electron moving in the t_{1u} LUMO state of C_{60}^- couples to the eight H_g and the two A_g phonons of C_{60} whose wavelengths and energies are shown in columns two and three

	Frequency		$\lambda\nu/N(0)$ (meV)
	λ^{-1} (cm^{-1})	$\hbar\omega_\nu$ (meV)	
$H_g(8)$	1575	195.3	23
$H_g(7)$	1428	177.1	17
$H_g(6)$	1250	155.0	5
$H_g(5)$	1099	136.3	12
$H_g(4)$	774	96.0	18
$H_g(3)$	710	88.0	13
$H_g(2)$	437	54.2	40
$H_g(1)$	273	33.9	19
$A_g(1)$	1470	182.3	11
$A_g(2)$	496	61.5	0
$\sum H_g$			147
$\sum(H_g + A_g)$			158

quantities for C_{60}^- and Na_8. It is seen that the partial coupling constants of Na_8 are, on an average, larger than those associated with C_{60}^-. The main reason for this result lies in the fact that the lowest Na_8 phonons have energies which are almost one order of magnitude smaller than those of C_{60}. Consequently, the square of the zero-point fluctuations $\sqrt{\hbar/2M_\lambda\omega_\lambda}$ entering the expression of $\lambda_v/N(0)$ are also larger by one order of magnitude. On the other hand, the matrix elements of the gradient of the electronic potential are comparable or smaller in Na_8 than in C_{60}^-. In this latter cluster, the delocalization of the $3p$ electrons is larger than that of the C_{60} electrons, which tend to concentrate around the bonds. The more the electrons are delocalized, and therefore far from the ion–ion bonds, the smaller are the electron–phonon matrix elements.

Table 5 Reduced electron–phonon matrix elements g_v and partial coupling constants $\lambda v/N(0)$ associated with the LUMO electron state and selected phonon states of Na_8 calculated making use of the CPMD scheme

	$\hbar\omega$ (meV)	$\lambda v/N(0)$ (meV)	g_v (meV)
A_1 (1)	7.93	36.3	12.0
A_1 (2)	9.48	173.8	28.7
A_1 (3)	15.94	56.9	21.3
A_1 (4)	16.72	12.2	10.1
Total		279.2	72.1

The matrix elements of the electron–plasmon coupling (see **Figure 10**)

$$\langle n, j'|H_{\text{el, plas}}|j\rangle = \int \mathrm{d}^3r\,\mathrm{d}^3r'\varphi^*_{j'}(\boldsymbol{r})\frac{\delta v_{\text{eff}}(\boldsymbol{r})}{\delta\varrho(\boldsymbol{r}')}\delta\varrho_n(\boldsymbol{r}';\omega_n)\varphi_j(\boldsymbol{r}) \quad [15]$$

can be calculated in terms of the functional derivative of the LDA effective potential v_{eff}, and of the matrix between the ground state $|0\rangle$ and the boson (plasmon) state $|n\rangle$, that is, $\langle n|\varrho(\boldsymbol{r}')|0\rangle \equiv \delta\varrho_n(\boldsymbol{r}';\omega_n)$.

In **Figure 11**, the results of a consistent calculation of the photoabsorbtion spectrum of the small metal cluster Na_9^+ in its lowest-energy geometry configuration $\mathscr{D}_{3h}$ calculated, taking into account both the electron–phonon and electron–plasmon couplings, is shown in comparison to the experimental findings. Both the coupling of electrons to plasmons and to phonons were considered and calculated *ab initio*. After diagonalizing the Hamiltonian, the strength function associated with the dipole operator as well as the photoabsorption cross section were calculated.

The main result of the calculation is that the electron–plasmon coupling is essential to shift the TDLDA spectrum downwards. This is shown in **Figure 11** for the case of Na_9^+. The inset of the figure, with the TDLDA peaks displayed by dashed lines and those resulting from the full calculation by solid

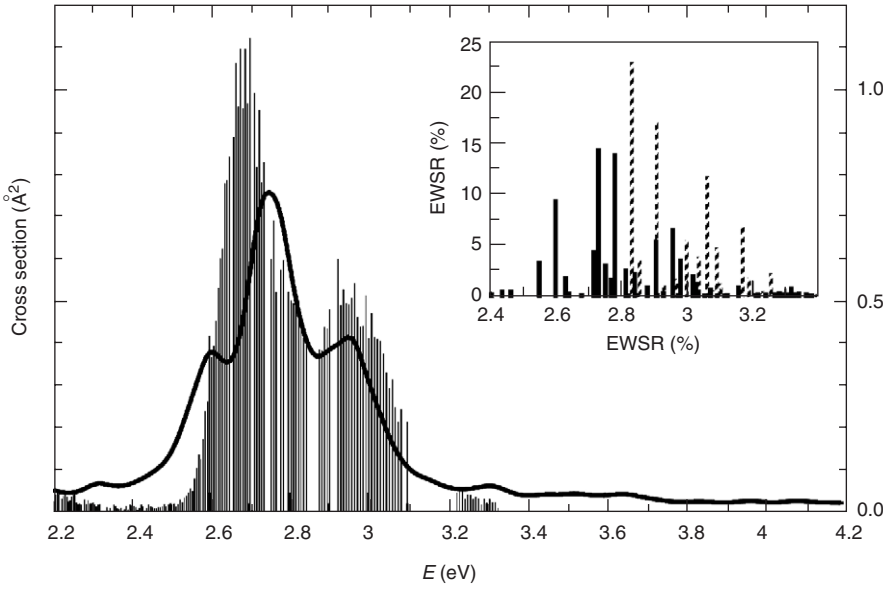


Figure 11 Photoabsorption cross section of Na_9^+ vs. energy E (continuous heavy line) calculated taking into account both the coupling to plasmons and to phonons. The experimental spectrum measured at $T \approx 39$ K is displayed by the thin vertical lines. The inset shows the percentage of the EWSR calculated making use of the TDLDA results (dashed lines), and taking into account the coupling to plasmons (continuous lines).

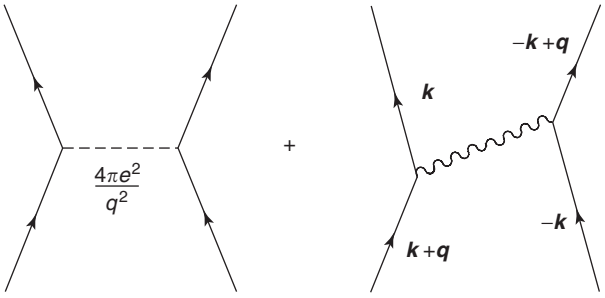


Figure 12 Scattering process of two electrons (arrowed lines) interacting via the bare Coulomb interaction (dotted line) and by exchanging a plasmon (wavy line).

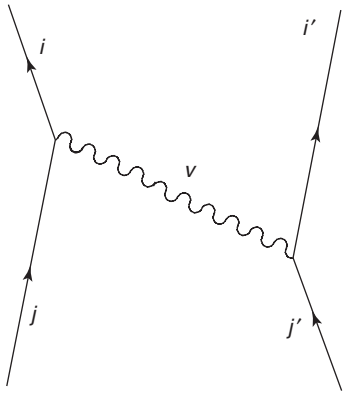


Figure 13 Graphical representation of the effective electron–electron interaction induced by the exchange of a phonon.

lines, demonstrates that the red shift of the main collective peaks is systematic and ~ 0.2 eV. This result indicates that the electron–plasmon coupling can also be of importance in other contexts where calculations of the dynamical screening are in order, for instance, in the case of the screened Coulomb potential μ (see **Figure 12**). A second result of the full calculation is that the electron–phonon coupling acts essentially as an averaging parameter ~ 0.1 eV, leading to a smoothing of the spectrum.

The Effective Electron–Electron Interaction

An electron can interact with the ions through the deformation potential and excite a phonon. This phonon can be absorbed by the same electron (self-energy process) or by a second electron. This second process gives rise to an effective electron–electron interaction (**Figure 13**) whose strength is measured by the dimensionless parameter $\lambda = \sum_v \lambda v$ (see eqn [14]). As mentioned above, the exchange of plasmons also gives rise to an electron–electron interaction (see **Figure 12**), which screens the bare Coulomb potential, an effect which is measured by the dimensionless quantity $\mu = U_e(q,\omega)N(0)$, where $U_e(q,\omega) = 4\pi e^2/q^2\varepsilon(q,\omega)$.

To approximately take into account the ω-dependence of the Coulomb pseudopotential in a way, which is consistent with the ω-dependence kept in the electron–phonon induced interaction, one can change μ to a new effective interaction μ^* obtained as

$$\mu^* = \frac{\mu}{1 + \mu \ln(B/\omega_{ph})} \qquad [16]$$

where B is a typical electron energy (half the bandwidth) and ω_{ph} is a typical phonon energy. If $B/\omega_{ph} \gg 1$, μ^* can be strongly reduced, relative to μ. In the limit $\mu \ln(B/\omega_{ph}) \gg 1$, the above expression simplifies to $\mu^* \approx 1/\ln(B/\omega_{ph})$, which may be ~ 0.1–0.2.

The (summed) effective interaction arising from the exchange of phonons and plasmons between electrons displayed in **Figures 12** and **13** can lead to an attraction between pairs of electrons moving close to the Fermi energy, giving rise to the so-called Cooper pairs, which behave as (quasi)-bosons. The associated binding energy is, in the limit of weak coupling ($\lambda \ll 1$),

$$E = 2\hbar\omega_D e^{-1/(\lambda-\mu^*)} \qquad [17]$$

where μ^* is a typical screened Coulomb potential.

Superconductivity in Fullerides

Fullerenes, for example, C_{60}, have attracted much interest since their discovery in 1985, not least because they are the only finite, and most symmetric, allotropic forms of carbon. This interest has increased drastically because methods were found to produce some of these fullerenes in large quantities, and so it was possible to make solids (fullerites). Intercalating alkali metal atoms in solid C_{60} (a van der Waals solid) leads to metallic behavior displaying a transition temperature $T_c \sim 30$ K.

Although there are still a number of open questions, several experimental findings testify to the fact that superconductivity in doped fullerides is likely to be driven by the coupling of electrons to intramolecular vibrations (phonons) of the fullerene molecules.

Making use of the density of states $N(0) \approx 8.1\ \text{eV}^{-1}$ deduced for fullerite Rb_3C_{60} as well as of the results of **Table 4** ($\lambda/N(0) = 158$ meV), one obtains, for the total electron–phonon coupling,

$$\lambda = 1.3 \qquad [18]$$

In such a situation, the weak coupling expression [17] is not adequate. A better expression is provided by the McMillan solution to Eliashberg equations:

$$T_c = \frac{\hbar\omega_{ln}}{1.2}\exp\left[-\frac{1.04(1+\lambda)}{\lambda-\mu^*(1+0.62\lambda)}\right] \qquad [19]$$

where ω_{ln} is a typical phonon frequency (logarithmic average).

In the case of C_{60},

$$\hbar\omega_{ln} = 102\,\text{meV} = 1184\,\text{K} \qquad [20]$$

The best available estimates of μ^* for alkali-doped fullerites are $0.3 \leqslant \mu^* \leqslant 0.4$. Making use of the results given in [18] and [20], and of the fact that for Rb_3C_{60}, $T_c = 29.5\,\text{K}$, one obtains $\mu^* \approx 0.34$ from [19].

The interest for using fullerenes smaller than C_{60} to produce nanometer superconducting materials can be simply understood by remembering that the critical temperature below which graphite compounds are superconducting is $T_c = 5\,\text{K}$, while the same temperature has reached $\approx 30\,\text{K}$ for C_{60}. Because C_{60} can be viewed as made out of a wrapped carbon lattice plane by introducing defects in the honeycomb lattice so as to create 12 pentagons (C_{60} is made out of these 12 pentagons and of 20 hexagons), curvature effects and electron spill-out seem to be at the basis of the increased electron–phonon coupling and screening of the Coulomb field in going from graphite to C_{60}-based materials. It is thus expected that materials made out of fullerenes (such as C_{36}, C_{28}, and C_{20}) with higher curvature and electron spill-out than those of C_{60} will display higher values of T_c.

Ab initio calculations predict values of λ for these smaller fullerenes as well as for $C_{28}H_4$, considerably larger than for C_{60}. Assuming similar values of $N(0)$ for the corresponding fullerites, one expects critical temperatures which are considerably larger than those for C_{60}-based materials. This is also true for materials based on metal clusters. Further, experimental work is needed to assess the soundness of such expectations.

See also: Excitons in Crystals; Insulators, Electronic States of; Insulators, Optical Properties of; Metals and Metallic Alloys, Optical Properties of; Nanostructures, Optical Properties of; Optical Absorption and Reflectance; Plasmons; Polarizabilities; Polymers and Organic Compounds, Optical Properties of; Small Particles and Clusters, Optical Properties of; Surfaces, Optical Properties of; Treated Surfaces, Optical Properties of.

PACS: 36.40. – C; 36.40.Vz; 36.40.Gk; 74.70.Wz

Further Reading

Broglia RA (1998) Wires of seven atoms – Feynman's very, very small world. *Contemporary Physics* 39: 371–376.

Broglia RA, Colo' G, Onida G, and Roman HE (2004) *Solid State Physics of Finite Systems*. Berlin: Springer.

Haberland H (ed.) (1994) *Clusters of Atoms and Molecules*. Berlin: Springer.

Handerson H (ed.) (1996) *Small Particles and Inorganic Clusters, ISSPIC 8*. Berlin: Springer.

Molecular Crystallography

M Wilmanns and M S Weiss, EMBL Hamburg, Hamburg, Germany

The description of X-ray crystallography as an experimental technique to determine the three-dimensional structures of molecules at atomic resolution consists of two parts. First, the relationship between the sample studied and its corresponding diffraction pattern must be derived and understood. Second, the way back from the diffraction pattern to the molecular structure must be outlined.

From Objects to Diffraction Patterns

The Scattering Equation

Electromagnetic waves interact with matter. Their electric components force charged particles to oscillate with the same frequency as the incident wave. Oscillating electrons constitute moving charges, which emit electromagnetic radiation of the same frequency into all directions of space. This phenomenon is called "elastic scattering," "Thomson scattering," or "scattering without loss of energy." Atomic nuclei interact with electromagnetic waves in the same way, but since a proton is $\sim$2000 times heavier than an electron, and consequently much less prone to move, for purposes of crystallography as a structure determination technique, scattering by nuclei can be neglected. The mathematical description of the phenomenon of scattering is illustrated in **Figure 1**.

The incoming X-ray wave (from the left) is represented by the vector $\boldsymbol{s}_{in}$ with $|\boldsymbol{s}_{in}| = 1/\lambda$, where λ is the wavelength of the incoming X-ray wave. The wave impinges upon the object and is then scattered by the whole object, in all directions of space. For

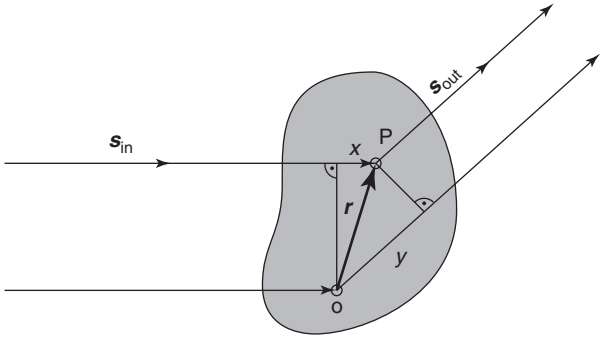


Figure 1 Scattering of electromagnetic waves by an object.

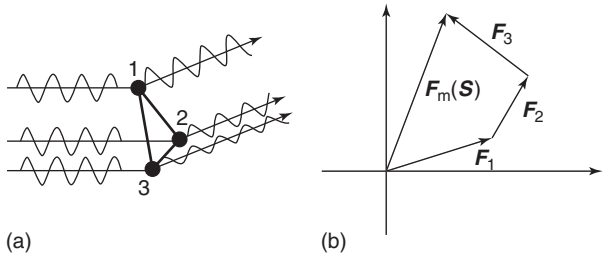


Figure 2 (a) Scattering by a three-atom molecule, with each of the three atoms containing a different number of electrons. (b) Argand diagram with the vector addition of the three individually scattered waves yielding the molecular transform.

simplicity, only the scattering at the points O and P is considered now, and only the scattering in the direction described by $\boldsymbol{s}_{out}$, where $\boldsymbol{s}_{out}$ is a vector representing the scattered wave ($|\boldsymbol{s}_{out}| = 1/\lambda$). The difference vector between $\boldsymbol{s}_{out}$ and $\boldsymbol{s}_{in}$ (i.e., $\boldsymbol{s}_{out} - \boldsymbol{s}_{in}$) is called the scattering vector $\boldsymbol{s}$. The scattering vector describes the change in direction that the X-ray wave undergoes upon scattering. The size of the object, that is, the distance between points O and P in the example given here, leads to a path difference $\Delta x = \boldsymbol{r} \cdot (\boldsymbol{s}_{out} - \boldsymbol{s}_{in})\lambda = \boldsymbol{r} \cdot \boldsymbol{s}\lambda$ between the two scattered waves shown in **Figure 1**. This path difference results in a phase difference $\Delta\alpha = 2\pi\Delta x/\lambda = 2\pi \boldsymbol{r} \cdot \boldsymbol{s}$ between the two scattered waves. The wave scattered at point P has an amplitude proportional to the electron density ρ(P) at point P and a phase of $2\pi \boldsymbol{r} \cdot \boldsymbol{s}$ relative to the wave scattered at point O. In order to compute the total wave f_T scattered by the object depicted in **Figure 1**, an integration over all volume elements (V) of the scattering object has to be performed. This leads to the general scattering equation (eqn [1]). $f_T(\boldsymbol{s})$ describes the scattering by every possible object, which in turn can be described by $\boldsymbol{r}$, $\rho(\boldsymbol{r})$, and V, into the direction defined by the vector $\boldsymbol{s}$:

$$f_T(\boldsymbol{s}) = \int_{Vol} \rho(\boldsymbol{r})\, e^{2\pi i \boldsymbol{r} \cdot \boldsymbol{s}}\, dV \qquad [1]$$

Equation [1] can also be regarded as a "Fourier analysis" of the object to be studied by the X-rays.

Scattering by Atoms and Molecules

The description of the scattering of real atoms starts from the general scattering equation (eqn [1]). The integration over the whole volume V of the atom implies that the function $f(\boldsymbol{s})$ is dependent on the form of the electron cloud of the atom. It is, therefore, also called the "atomic form factor." The atomic form factor is an intrinsic property of an atom or ion, which depends on the exact architecture of the electron cloud of the atom or the ion. Atomic form factors can be calculated using quantum mechanical approximations from the electron distribution of the atoms or ions, but they can also be measured very accurately, for instance, by X-ray scattering. Tables of the atomic form factors for each existing atom or ion are available in the International Tables for Crystallography (see the "Further reading" section).

The scattering by a molecule can now simply be calculated based on the scattering of the atoms belonging to the molecule (**Figure 2a**). The total scattered wave is the sum of the individually scattered waves in any given direction (**Figure 2b**). Using the example of a three-atom molecule, the contributions by the three atoms have to be considered: $\boldsymbol{F}_1 = f_1(\boldsymbol{s})e^{2\pi i \boldsymbol{r}_1 \cdot \boldsymbol{s}}$ for atom 1, $\boldsymbol{F}_2 = f_2(\boldsymbol{s})e^{2\pi i \boldsymbol{r}_2 \cdot \boldsymbol{s}}$ for atom 2, and $\boldsymbol{F}_3 = f_3(\boldsymbol{s})e^{2\pi i \boldsymbol{r}_3 \cdot \boldsymbol{s}}$ for atom 3. The result is shown in eqn [2]. It is also called the "molecular transform:"

$$\boldsymbol{F}_m(\boldsymbol{s}) = \sum_{j=1}^{3} f_j(\boldsymbol{s})\, e^{2\pi i \boldsymbol{r}_j \cdot \boldsymbol{s}} \qquad [2]$$

$f_j(\boldsymbol{s})$ are the atomic form factors and $\boldsymbol{r}_j$ the coordinates of the atoms j.

Scattering by a Crystal

A crystal is defined as an object, which is periodic in three dimensions. In order to build up a crystal from its smallest unit, the unit cell, one just needs to know the contents of the unit cell and the lattice translations $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$. For describing the scattering of one single unit cell of a crystal, $\boldsymbol{F}_{unit\ cell}(\boldsymbol{s})$, the contributions of all N atoms in the unit cell have to be added up (eqn [3]):

$$\boldsymbol{F}_{unit\ cell}(\boldsymbol{s}) = \sum_{j=1}^{N} f_j(\boldsymbol{s})\, e^{2\pi i \boldsymbol{r}_j \cdot \boldsymbol{s}} \qquad [3]$$

In order to consider the scattering by the whole crystal, the contributions of all unit cells of the crystal and all atoms in each unit cell have to be taken

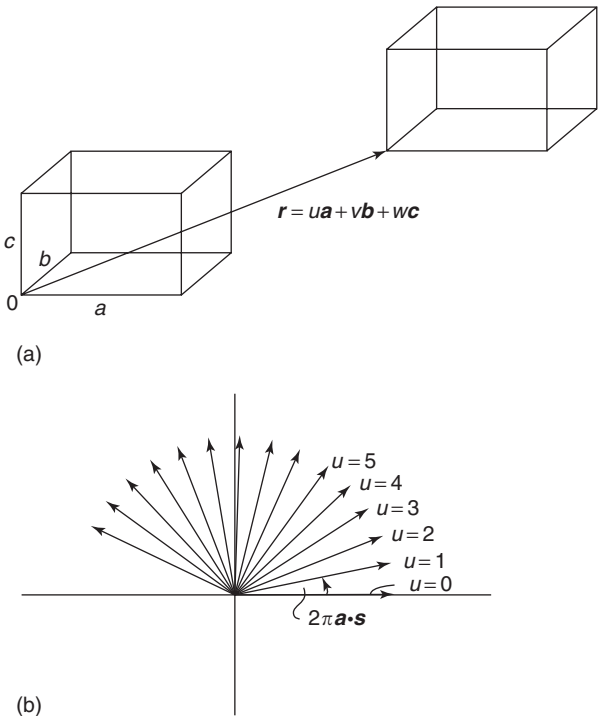


Figure 3 (a) Two unit cells of a crystal and (b) graphical representation of the Laue conditions in one dimension. In the case shown, the scalar product $\boldsymbol{a} \cdot \boldsymbol{s} \neq h$. As a result of this, the waves with different values of u exhibit phase differences different from 2π. Due to the large number of waves to be considered (one for each unit cell of the crystal), the overall sum will be zero.

into account. Let $\boldsymbol{F}_{\text{unit cell}}(\boldsymbol{s})$ be the scattering of the first unit cell, and $\boldsymbol{F}'_{\text{unit cell}}(\boldsymbol{s})$ the scattering of a second unit cell shifted by $\boldsymbol{r} = u\boldsymbol{a} + v\boldsymbol{b} + w\boldsymbol{c}$ relative to the first unit cell (u, v, and w are integer values, and $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are the lattice translations) (**Figure 3a**). It can be shown that the contribution of just the second unit cell can be expressed as the scattering of the first unit cell and the vector that relates the two unit cells:

$$\boldsymbol{F}'_{\text{unit cell}}(\boldsymbol{s}) = \boldsymbol{F}_{\text{unit cell}}(\boldsymbol{s}) \mathrm{e}^{2\pi i u \boldsymbol{a}\cdot\boldsymbol{s}} \mathrm{e}^{2\pi i v \boldsymbol{b}\cdot\boldsymbol{s}} \mathrm{e}^{2\pi i w \boldsymbol{c}\cdot\boldsymbol{s}} \quad [4]$$

To arrive at the scattering equation for the whole crystal, the sum over all unit cells should be calculated. This can be reduced to summing over all combinations of u, v, and w:

$$\boldsymbol{F}_{\text{crystal}}(\boldsymbol{s}) = \boldsymbol{F}_{\text{unit cell}}(\boldsymbol{s}) \sum_{u=0}^{n_1} \mathrm{e}^{2\pi i u \boldsymbol{a}\cdot\boldsymbol{s}} \sum_{v=0}^{n_2} \mathrm{e}^{2\pi i v \boldsymbol{b}\cdot\boldsymbol{s}} \sum_{w=0}^{n_3} \mathrm{e}^{2\pi i w \boldsymbol{c}\cdot\boldsymbol{s}} \quad [5]$$

where n_1, n_2, and n_3 are the numbers of the unit cells in $\boldsymbol{a}$-, $\boldsymbol{b}$-, and $\boldsymbol{c}$-directions. These numbers are large for typical crystals in the micrometer-size range. This means that the sums $\sum_{u=0}^{n_1} \mathrm{e}^{2\pi i u \boldsymbol{a}\cdot\boldsymbol{s}}$, $\sum_{v=0}^{n_2} \mathrm{e}^{2\pi i v \boldsymbol{b}\cdot\boldsymbol{s}}$, and $\sum_{w=0}^{n_3} \mathrm{e}^{2\pi i w \boldsymbol{c}\cdot\boldsymbol{s}}$ will be zero unless $\boldsymbol{a}\cdot\boldsymbol{s}$, $\boldsymbol{b}\cdot\boldsymbol{s}$, and $\boldsymbol{c}\cdot\boldsymbol{s}$ are integral numbers. These conditions are also known as the "Laue conditions," and they are illustrated in **Figure 3b** as vector additions in an Argand diagram. In the case of the Laue conditions being fulfilled, the product of the three sums is equal to the total number of unit cells in a given crystal.

There are two important consequences of this derivation. First, the scattering of a crystal is proportional to the scattering of just one unit cell and the number of unit cells in the crystal. Consequently, to describe the scattering by the whole crystal, it is sufficient to consider the scattering by just one unit cell. Second, the scattering of a crystal leads to a discrete diffraction pattern. Only in those directions $\boldsymbol{s}$, in which the Laue conditions are obeyed, will scattering be observed. In all other directions the sum of all scattered waves will be zero. This leads to the observation of a diffraction pattern consisting of discrete maxima, the so-called X-ray reflections. Each reflection is identified by a unique combination of the three numbers h, k, and l, the Miller indices:

$$\boldsymbol{F}_{\text{crystal}}(\boldsymbol{s}) = n\boldsymbol{F}_{\text{unit cell}}(\boldsymbol{s}) = n \sum_{j=1}^{N} \boldsymbol{f}_j(\boldsymbol{s}) \mathrm{e}^{2\pi i \boldsymbol{r}_j\cdot\boldsymbol{s}} \quad [6]$$

where n is the number of unit cells. With $\boldsymbol{r}_j = x_j\boldsymbol{a} + y_j\boldsymbol{b} + z_j\boldsymbol{c}$ (x_j, y_j, and z_j are the fractional coordinates of atom j), and with the Laue conditions $\boldsymbol{a}\cdot\boldsymbol{s} = h$, $\boldsymbol{b}\cdot\boldsymbol{s} = k$, and $\boldsymbol{c}\cdot\boldsymbol{s} = l$, eqn [6] can be transformed into eqn [7]:

$$\boldsymbol{F}(\boldsymbol{s}) = \boldsymbol{F}(hkl) = \sum_{j=1}^{N} \boldsymbol{f}_j \mathrm{e}^{2\pi \mathrm{i}(hx_j + ky_j + lz_j)} \quad [7]$$

Equation [7] is also called the "structure factor equation." Together with the "electron density equation" (see below), it constitutes one of the two central equations in X-ray crystallography. It describes the scattering of a crystal defined by the unit cell vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, and by the atoms inside the unit cells of this crystal as defined by the atomic form factors (f_j) and the fractional coordinates (x_j, y_j, and z_j).

The Diffraction Experiment

A schematic representation of an X-ray experiment is shown in **Figure 4**. The experiment itself involves the production or generation of X-rays, the selection of the X-rays for the experiment, the preparation of the crystal for the experiment, the actual data collection step, and the processing and validation of the resulting data. Of course, if one is interested only in using

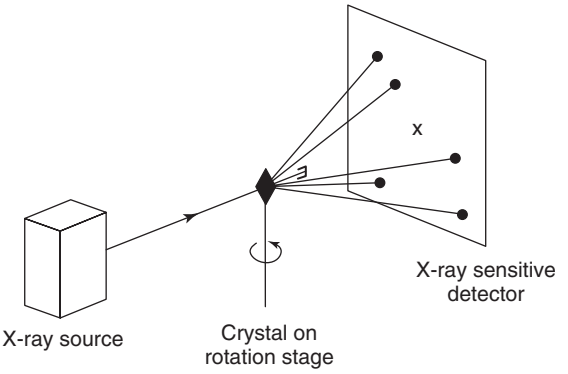


Figure 4 Schematic representation of the diffraction experiment. Depicted are the X-ray source, the object (crystal) mounted on a rotatable device, and the detector for the scattered waves.

X-rays to determine a crystal structure, how X-rays are produced and prepared for the measurement may be of little concern. However, there are some important aspects that can potentially have a great impact on the quality of the obtained diffraction data and consequently on various stages of structure determination and certainly on the information content of the resulting structure. One should always keep in mind that the quality or the information content of a crystal structure is directly determined by the quality or the information content of the diffraction data that are the basis for the determination of the structure. Furthermore, the collection of diffraction data is the last real experimental step to be carried out in the course of a crystal structure analysis. Everything after that is merely an interpretation of the collected diffraction data.

From the Diffraction Pattern to the Structure

The Electron Density Equation

The Fourier transformation of the general scattering equation (eqn [1]) leads to the electron density equation (eqn [8]), the second important equation in crystallography. Instead of integrating over the volume of the scattering object V, the integration proceeds now over diffraction space, which is sometimes also called reciprocal space, because of the reciprocal relationship to the real space with respect to its geometric parameters:

$$\rho(\boldsymbol{r}) = \int_{\mathrm{Vol}^*} f_{\mathrm{T}}(\boldsymbol{s})\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{r}\cdot\boldsymbol{s}}\, \mathrm{d}V^* \qquad [8]$$

Similar to regarding eqn [1] as a "Fourier analysis" of the object, eqn [8] can be regarded as a "Fourier synthesis." The object can be reconstructed from its diffraction pattern by taking into account all scattered waves. A simpler version of the electron density equation (eqn [9]) can be derived considering that the diffraction pattern of a crystal is discrete (see above), and that instead of integrating over the whole volume of diffraction space, one only needs to sum over the observed X-ray reflections:

$$\rho(x,y,z) = \sum_{h,k,l} |\boldsymbol{F}(hkl)|\, \mathrm{e}^{-2\pi\mathrm{i}(hx+ky+lz)+\mathrm{i}\alpha(hkl)} \qquad [9]$$

Two things become clear from eqn [9]. First, in order to calculate the value for the electron density ρ at any given point (x, y, z), all reflections should be taken into account. Second, the summation can only be carried out if both the structure factor amplitude $|\boldsymbol{F}|$ and the phase α are known for each reflection (hkl).

The Phase Problem

In a diffraction experiment, data, which are finally measured, are the intensities I of the reflections (hkl). The intensity $I(hkl)$ can be calculated by multiplying the structure factor $\boldsymbol{F}(hkl)$ with its conjugated complex structure factor $\boldsymbol{F}^*(hkl)$. The net result is that $I(hkl)$ is proportional to the square of the structure factor amplitude $|\boldsymbol{F}(hkl)|$:

$$\begin{aligned} I(hkl) &\propto |\boldsymbol{F}(hkl)|\, |\boldsymbol{F}^*(hkl)| \\ &= |\boldsymbol{F}(hkl)|\, \mathrm{e}^{\mathrm{i}\alpha(hkl)} |\boldsymbol{F}(hkl)|\, \mathrm{e}^{-\mathrm{i}\alpha(hkl)} \\ &= |\boldsymbol{F}(hkl)|^2 \end{aligned}$$

As a consequence, by measuring the intensities, one is only able to back-calculate the amplitudes $|\boldsymbol{F}(hkl)|$ of the waves. The information about the phases $\alpha(hkl)$ is lost. This is known as the so-called "phase problem" in crystallography. The phase problem is the reason why an image of the molecule to be studied cannot be calculated directly from the diffraction data. Approaches to overcome the problem are briefly discussed in the next section.

Phase Determination

Over time, a number of methods have been developed to circumvent the phase problem. Today, the decision as to which method to use seems to be governed by both the size of the molecules to be studied and the diffraction power of the respective crystals. Whereas in small-molecule crystallography (molecules of up to a few hundred atoms, and typically very good diffraction), the so-called "direct methods" are dominant and nowadays also mostly automated, in macromolecular crystallography (more than 1000 nonhydrogen atoms with no apparent upper limit,

but usually limited diffraction) the so-called "replacement methods" are the major methods to be applied. In the following, the different approaches are discussed in brief.

Direct methods The direct methods rely on the principle that the phase information is somehow included in the measured structure factor amplitudes. The underlying rationale is that structure factors are not independent of each other but that they are related through the structure. With the additional constraints that (1) the electron density can never be negative and (2) the structure to be determined consists of discrete atoms, phase sets can be constructed randomly and examined with respect to their validity.

The credibility of a given set of phases can be examined by considering probabilistic relationships between the phases of certain reflections, for example, the "triplet invariants" (eqn [10]). It can be derived that the phases of the three reflections (h_1, k_1, l_1), (h_2, k_2, l_2), and ($-h_1-h_2$, $-k_1-k_2$, $-l_1-l_2$) are related by the following equation:

$$\alpha(h_1,k_1,l_1)+\alpha(h_2,k_2,l_2) + \alpha(-h_1-h_2,-k_1-k_2,-l_1-l_2)\approx 0 \quad [10]$$

This equation is not absolutely but approximately valid, and it is normally formulated as a probability. The stronger the three reflections (h_1, k_1, l_1), (h_2, k_2, l_2), and ($-h_1-h_2$, $-k_1-k_2$, $-l_1-l_2$) are, the higher the probability that the three-phase values sum up to zero or an integral multiple of 2π.

Various measures to evaluate the quality of a given phase set have been introduced over time. The most widely known formula for this purpose is the "tangent formula" (eqn [11]). In this formula, the most probable value of the phase of one reflection $\alpha(h_1, k_1, l_1)$ is given as a function of the values of the phases of all other reflections (h_m, k_m, l_m) and ($-h_1-h_m$, $-k_1-k_m$, $-l_1-l_m$), which share a triplet invariant with (h_1, k_1, l_1). The reliability of the formula depends on the magnitude of the structure factor amplitudes of the reflections considered:

$$\tan[\alpha(h_1,k_1,l_1)] = \frac{\sum_m W_m \sin[-\alpha(h_m,k_m,l_m)-\alpha(-h_1-h_m,-k_1-k_m,-l_1-l_m)]}{\sum_m W_m \cos[-\alpha(h_m,k_m,l_m)-\alpha(-h_1-h_m,-k_1-k_m,-l_1-l_m)]} \quad [11]$$

The two sums in the numerator and the denominator run over all possible triplet invariants m, in which the reflection (h_1, k_1, l_1) is involved, and the numbers W_m describe the weighted product of the normalized structure factor amplitudes of reflections (h_m, k_m, l_m) and ($-h_1-h_m$, $-k_1-k_m$, $-l_1-l_m$).

Relationships such as the triplet invariants and the tangent formula help to reduce the complexity of the phase set. Nevertheless, because of the large number of possible phase sets, direct methods for phase determination are typically only applicable to small structures.

Isomorphous replacement One of the dominant methods in macromolecular crystallography is the method of isomorphous replacement. In this method, diffraction data sets of two isomorphous crystals are compared with each other reflection by reflection. The first crystal may be one of the native macromolecules, the second crystal may be one of the heavy-atom derivatives of the same macromolecule. The latter can be prepared by either co-crystallization or soaking a native macromolecule crystal in a heavy-atom-containing solution. Large solvent channels inside the macromolecule crystals allow the diffusion of the heavy-atom compound through the crystal. If a reactive group on the macromolecule surface is accessible to the heavy-atom compound, it can bind there without changing the structure of the rest of the protein. A detailed comparison of the two diffraction data sets will then, in a first step, yield the differences between the data sets, which are due to the heavy-atom substructure. From these differences, the substructure can be determined. With the substructure at hand, one can then calculate the structure factor amplitude and the phase of just the substructure, and then use this heavy-atom quasiphase as a reference to calculate the phases of the reflections of the macromolecule crystal. An easy way to visualize this is a vector diagram (**Figure 5**). The vector addition of the structure factor of the macromolecule $\boldsymbol{F}_P$ and the structure factor of the heavy atoms $\boldsymbol{F}_H$ will yield the structure factor of the heavy-atom derivative $\boldsymbol{F}_{PH}$. If the magnitudes of both $\boldsymbol{F}_P$ and $\boldsymbol{F}_{PH}$ (these are the structure factor amplitudes which have been measured in the two diffraction experiments) and the magnitude and direction of $\boldsymbol{F}_H$ (this can be calculated once the heavy-atom substructure is known) are now known, all other parameters of the triangle in **Figure 5** can be calculated as well.

However, this vector addition is not unambiguous. The result of this calculation will not be a unique

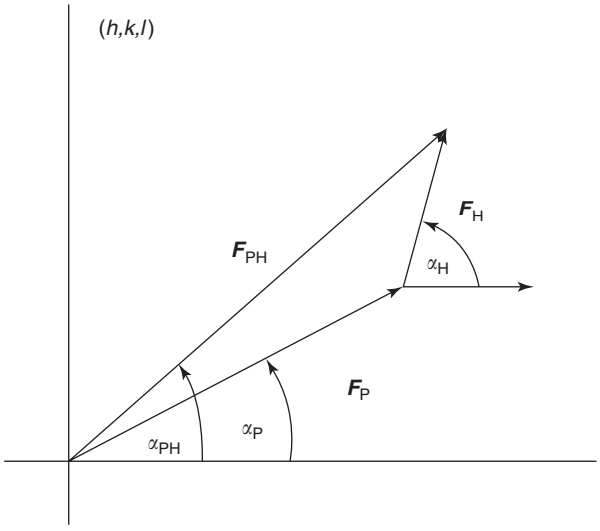


Figure 5 Vector addition of the structure factors for the native macromolecule $\boldsymbol{F}_P$, the heavy-atom substructure $\boldsymbol{F}_H$, and the heavy-atom derivative of the macromolecule $\boldsymbol{F}_{PH}$.

phase $\alpha(hkl)$ for the given reflection, but two possible solutions, which are symmetrical about the vector $\boldsymbol{F}_H$. This situation can be made unambiguous, for instance, by including a second heavy-atom derivative.

Anomalous diffraction If the structure to be determined contains anomalous scatterers (anomalous scattering occurs when the energy of the X-rays used for the diffraction experiment is close to an energy which is absorbed by an atom), a diffraction experiment can be carried out at two (or more) different X-ray energies. The differences between the diffraction patterns can then again be attributed to the anomalously scattering substructure, and in the same way as above, the substructure can be determined, and the substructure phase be used as a reference to determine the phase α_P.

Molecular replacement A conceptually completely different approach can be applied if the structure to be determined is expected to be similar to another structure which is known already. In such a case, the known structure can be placed in the unit cell of the structure to be determined, and the structure factor $\boldsymbol{F}_P$ calculated. The phase α_P of the structure factor amplitude $|\boldsymbol{F}_P|$ will of course not be correct, but it will be approximately correct and can be used as a starting value for further refinement. The initial step in the determination of the molecular replacement structure is the determination of three rotational and three translational parameters, which define, respectively, the correct orientation and the correct translation of the search model (the known structure) in the unit cell of the target structure (the structure to be determined).

Model Building and Refinement

Once initial phases have been determined by one of the methods described above, the electron density can be calculated according to eqn [8]. The electron density is nothing but an image of the structure to be determined. It shows where, in the unit cell, the electrons (i.e., the atoms) are. Since molecular structures are usually described by the position of all of their atoms, the electron density should be interpreted in terms of atomic positions and identities. Finally, the model thus obtained can be subjected to a mathematical optimization procedure called refinement, in order to optimize the fit between the structure factors that can be calculated from the model and the ones that have been observed experimentally. This constant feedback between the model and experimental data is one of the greatest strengths of the crystallographic methods.

See also: Biological Structures; Crystal Structure; Multicircle Diffractometry Methods; Polymer Structures; Scattering, Elastic (General); Structure Types and Classifications; Synchrotron Radiation; X-Ray Sources.

PACS: 61.10.Nz; 61.66.Hq; 61.68.+n

Further Reading

Blundell TL and Johnson LN (1976) *Protein Crystallography*. London: Academic Press.

Drenth J (1994) *Principles of Protein X-ray Crystallography*. New York: Springer.

Giaccovazzo C, Monaco HL, Viterbo D, Scordari F, Gilli G, *et al.* (1992) *Fundamentals of Crystallography*, 2nd edn. Oxford: Oxford University Press.

Glusker JP and Trueblood KN (1985) *Crystal Structure Analysis*. Oxford: Oxford University Press.

Hahn TH (1995) *International Tables for Crystallography, Vol. A – Space Group Symmetry*,, 4th edn. Dordrecht: Kluwer Academic.

Ladd MFC and Palmer RA (1994) *Structure Determination by X-Ray Crystallography*. New York: Plenum.

Rhodes G (1993) *Crystallography Made Crystal Clear*. San Diego: Academic Press.

Stout GH and Jensen LH (1968) *X-Ray Structure Determination*. New York: Macmillan.

Molecular Dynamics Calculations

L Colombo, University of Cagliari, Monserrato, Italy

The Basic Conceptual Framework

Consider a set of N particles with Cartesian coordinates $\{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N\}$, interacting through a potential $U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N)$. Both, atoms and molecules will be referred to as "particles." It is assumed that the interaction potential U does not depend upon particle velocities and it is represented by a well-behaved and differentiable (with respect to interparticle distances) functional.

According to classical mechanics, the total force $\boldsymbol{F}_\alpha$ acting on the αth particle is given by

$$\boldsymbol{F}_\alpha = -\nabla_\alpha U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N) \qquad [1]$$

and a mechanical trajectory $\boldsymbol{R}_\alpha = \boldsymbol{R}_\alpha(t)$ can be computed, provided that the instantaneous particle positions are computable at each time t. For condensed matter systems, the trajectories will be typically generated within a periodically repeated simulation cell, containing the full set of N particles. Therefore, the trajectory of any given particle or its periodic image is investigated, in case it migrates through cell boundary. The concept is illustrated in **Figure 1**. Periodic boundary conditions (PBCs) are introduced to minimize the absolute number of particles to deal with and to avoid the undesired free surfaces created by crude truncation at cell boundaries. Furthermore, PBCs allow to simulate ensembles with a constant number of particles, regardless of the diffusivity properties of the simulated system.

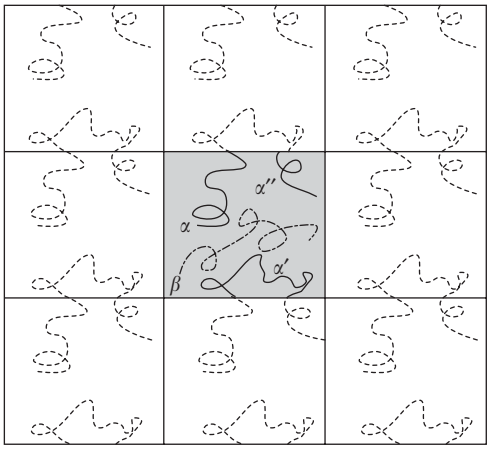


Figure 1 Two-dimensional example of periodic boundary conditions. The full-line trajectory migrates through the simulation cell (shaded area) boundaries; dashed lines represent its periodic images. In this case, the computed properties are collected along the piece-like trajectory $\alpha \rightarrow \alpha' \rightarrow \alpha''$. On the contrary, the dot-dashed trajectory β is fully contained within the simulation cell: no periodic boundary conditions need to be applied.

In order to translate the classical mechanics problem into a form suitable for digital computers, one needs to introduce the discretization of the time evolution. This is obtained by introducing a time-step δt (**Figure 2**), representing the unit of time interval, and by further assuming that the force $\boldsymbol{F}_\alpha$ is constant over δt. The resulting particle motion is linear and uniformly accelerated throughout δt. The optimal value of δt depends on the investigated system. However, in most applications it is of the order 10^{-15} s, so that $\sim$10–100 time-steps are needed to span one typical phonon period. The corresponding equations of motion are formulated according to second-order Newton dynamics

$$\begin{aligned} \boldsymbol{R}_\alpha(t+\delta t) &= \boldsymbol{R}_\alpha(t) + \dot{\boldsymbol{R}}_\alpha(t)\delta t + \tfrac{1}{2}\ddot{\boldsymbol{R}}_\alpha(t)\delta t^2 \\ \dot{\boldsymbol{R}}_\alpha(t+\delta t) &= \dot{\boldsymbol{R}}_\alpha(t) + \ddot{\boldsymbol{R}}_\alpha(t)\delta t \end{aligned} \qquad [2]$$

the particle acceleration being computed from the interaction model

$$\begin{aligned} \ddot{\boldsymbol{R}}_\alpha(t) &= \frac{1}{m_\alpha}\boldsymbol{F}_\alpha(t) \\ &= -\frac{1}{m_\alpha}\nabla_\alpha U\left(\boldsymbol{R}_1(t), \boldsymbol{R}_2(t), \ldots, \boldsymbol{R}_N(t)\right) \end{aligned} \qquad [3]$$

where m_α is the particle mass. This simple formulation allows for setting up a first rudimental iterative scheme to generate by a computer simulation the time evolution of the full system, as reported in **Figure 3**: such a procedure is referred to as molecular dynamics (MD) simulation.

By keeping the cell volume fixed, the MD simulation generates trajectories of a physical system with

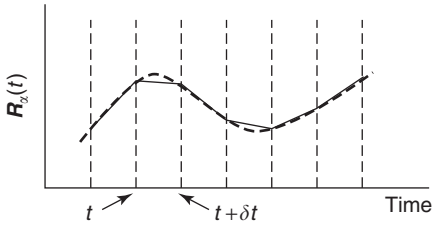


Figure 2 Time discretization of trajectories aging. Bold-dashed line represents the exact trajectory. During the numerical simulation the position $\boldsymbol{R}_\alpha$ is known only at discrete times corresponding to dashed vertical lines. The associated computer-generated trajectory is represented by the full line.

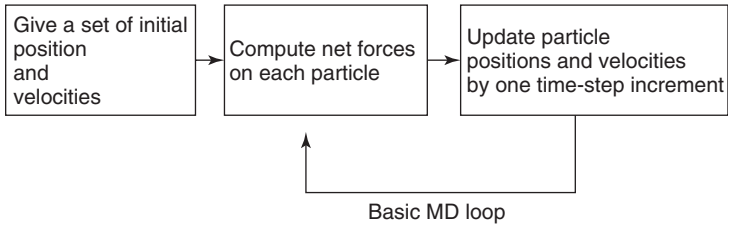


Figure 3 The simplest iterative scheme ruling MD. One iteration corresponds to increasing time by an amount δt.

constant number of particles N, constant volume V, and constant energy E (no external forces are assumed to act on the system). This corresponds to the microcanonical ensemble condition (NVE). The instantaneous kinetic temperature $T(t)$ of the system can also be computed on-the-flight during the MD run:

$$T(t) = \frac{1}{gK_B}\sum_{\alpha=1}^{N} m_\alpha \dot{\boldsymbol{R}}_\alpha^2(t) \qquad [4]$$

where K_B is the Boltzmann constant and g represents the number of degrees of freedom of the system. This relation (based on equipartition) represents the basic link between classical mechanics and thermodynamics. Indeed, one can gauge the instantaneous temperature $T(t)$ to a reference temperature T_{ref} simply by scaling at each time-step all velocities by the factor $\sqrt{T_{ref}/T(t)}$. Although this is a rather crude approach (in the next section more refined and sound algorithms are discussed), it represents a simple and effective way to control the temperature (or, more precisely, the kinetic energy of the system). By keeping T_{ref} constant in time, an approximately constant temperature is generated, that is, canonical (NVT), ensemble. Conversely, by changing the value of T_{ref} during the simulation, the system is allowed to heat or cool. It is important to stress that the present condition of constant kinetic energy is not rigorously equivalent to the condition of constant temperature, as discussed below. Nevertheless, it is widely adopted to simulate canonical systems.

Although conceptually very simple, the above iterative scheme can provide valuable information about a huge set of physical properties relevant to condensed matter physics – provided that a reliable interaction potential U is given. In particular, it is possible: (1) to elaborate structural models of bulk, low-dimensional, and molecular systems; (2) to investigate dynamical processes like film growth, surface diffusion, molecular interactions; (3) to simulate transport phenomena like bulk self-diffusion and defect mobility; (4) to model materials microstructure evolution under ion beam processing, phase transition, melting, and solid-phase epitaxy; and (5) to study the structure and topology of nano-/poly-structured, disordered, and complex materials.

The Generation of Particle Trajectories

The overall reliability of the MD method critically depends on the generation of accurate and stable particle trajectories. Special care must be devoted to selecting the best numerical algorithm for the integration of Newton's equations of motion. A widely adopted scheme is the so-called velocity-Verlet (VV) algorithm, which also represents a prototypical example of numerical integration scheme by finite difference methods. According to the VV algorithm, aging the particle trajectories corresponds to a multiple-step procedure:

1. the particle positions $\boldsymbol{R}_\alpha(t)$ are updated to time $t+\delta t$ as

$$\boldsymbol{R}_\alpha(t+\delta t) = \boldsymbol{R}_\alpha(t) + \dot{\boldsymbol{R}}_\alpha(t)\delta t + \tfrac{1}{2}\ddot{\boldsymbol{R}}_\alpha(t)\delta t^2 \qquad [5]$$

2. the atomic velocities are then computed at the mid step time $t + (1/2)\delta t$ using

$$\dot{\boldsymbol{R}}_\alpha(t+\tfrac{1}{2}\,\delta t) = \dot{\boldsymbol{R}}_\alpha(t) + \tfrac{1}{2}\ddot{\boldsymbol{R}}_\alpha(t)\delta t \qquad [6]$$

3. forces and accelerations are now computed at time $t+\delta t$ as shown in eqn [3]; and
4. particles velocities are finally evolved to time $t+\delta t$ according to

$$\dot{\boldsymbol{R}}_\alpha(t+\delta t) = \dot{\boldsymbol{R}}_\alpha(t+\tfrac{1}{2}\,\delta t) + \tfrac{1}{2}\ddot{\boldsymbol{R}}_\alpha(t+\delta t)\delta t \qquad [7]$$

This finite difference method is correct to the order δt^4.

The VV algorithm is simple and requires a straightforward numerical coding. It also easily allows for temperature control by velocities rescaling. Furthermore, it is not demanding as far as memory is required: just nine basic data (i.e., the Cartesian components of position, velocity, and force vectors) are actually needed for each particle in the system. More importantly, VV algorithm is stable and allows for the use of relatively long time-steps

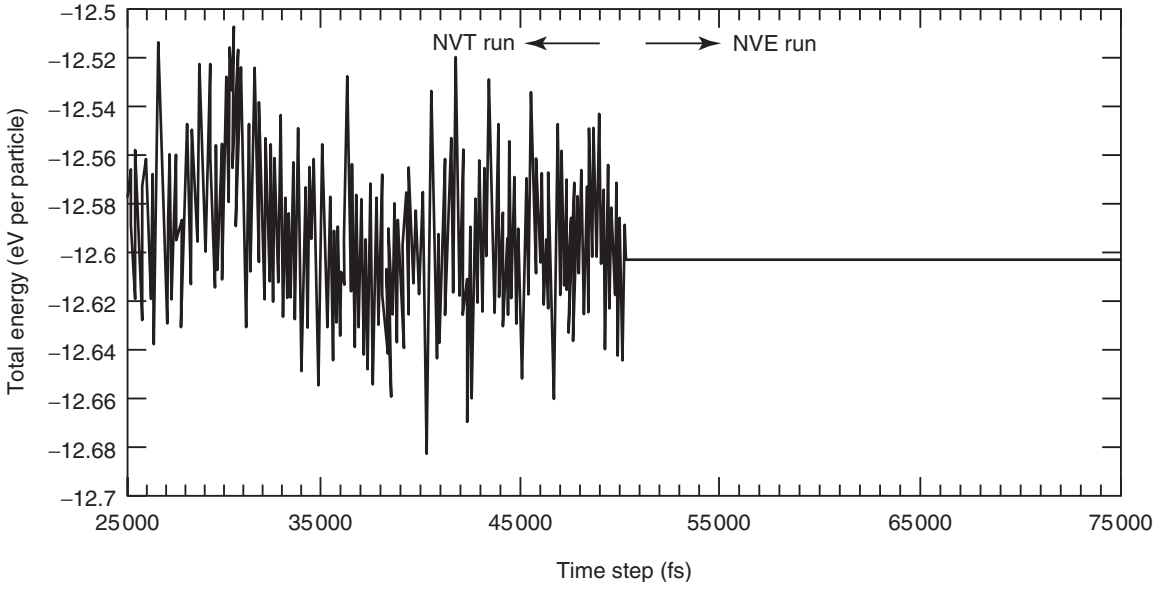


Figure 4 Time evolution of total energy per particle for a model silicon system at $T = 1200$ K. The system has been thermalized during a NVT run until time 50 000 δt. From that time step on the evolution was microcanonical (NVE): total energy is conserved to within 5×10^{-4} eV per particle.

(10^{-15} s $\leqslant \delta t \leqslant 10^{-14}$ s), still giving optimal conservation of the mechanical time-invariant quantities. **Figure 4** shows the time evolution of the total energy E for a solid-state system aged by an NVE MD run.

A common feature of any finite difference method is the calculation of the particle accelerations or, equivalently, forces. This is likely the most computationally intensive step of the MD loop, as explained in the next section. Therefore, the number of force calls at each loop should be minimized. For short-ranged potentials and for large enough systems, many particles do not contribute to the net force acting on a given particle α. It is therefore convenient to exclude those particles that do not interact with α from the expensive calculation of F_α, which, as already explained, depend upon interparticle distances. This can be obtained by means of the Verlet list algorithm, which is illustrated in **Figure 5**. Suppose that the interatomic potential U has a cutoff at d_1. Define a sphere of larger radius d_2 and put all the particles there contained in the neighboring list of particle α. Obviously, d_2 must be much smaller than the simulation cell size, in order to develop a useful algorithm. The actual particles interacting with α are now searched only within the sphere of radius d_2 (instead of within the entire simulation cell). If the difference $|d_2 - d_1|$ is properly chosen, for quite a few time-steps there is no need to refresh the neighboring list (an operation whose CPU workload scales as N^2), thus saving significant computer time. Furthermore, when the Verlet list algorithm is coupled to the linked cell method, it provides a CPU workload linearscaling in the number of particles N.

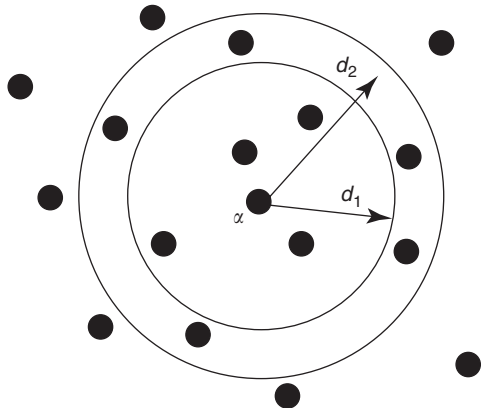


Figure 5 The Verlet neighboring list method.

A rather different approach is required when the potential energy U is long-ranged, as in polar systems where charge–charge (or even dipole–dipole) interactions must be taken into account. In this case one cannot take advantage from potential shortrangeness and it is necessary to compute the Coulomb contribution U_C to potential energy according to

$$U_C = \frac{1}{4\pi\varepsilon_0} \sum_{\alpha=1}^{N} \sum_{\beta} \frac{Z_\alpha Z_\beta}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|} \quad \text{with } \alpha \neq \beta \qquad [8]$$

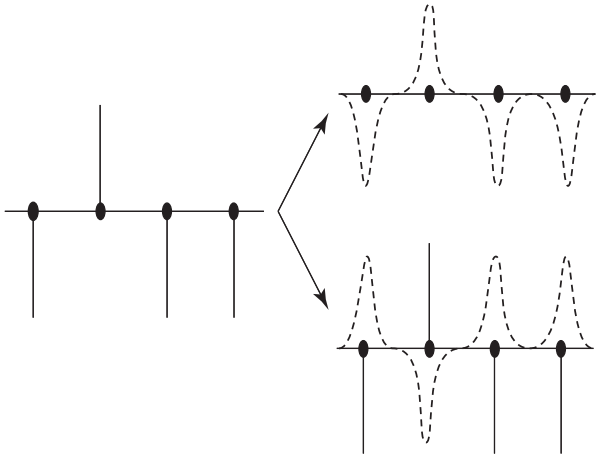


Figure 6 The Ewald sum method: a system of positive and negative point charges (left) is replaced by an equivalent system (right) of point (full line) and diffuse (dashed line) charges with screened (i.e., short-ranged) interactions.

where the index α runs over all the particles in the simulation cell, whereas the index β also labels all the (periodic) image particles. Z_α and Z_β represent the particle point charges (either positive or negative). This expression cannot be computed by a brute-force approach, since it is only conditionally convergent and, of course, cannot be simply truncated. Rather, the system of point charges is replaced with an equivalent system of point and diffuse charges where effective screening takes place. The procedure – known as Ewald sum – is illustrated in **Figure 6**: a diffuse charge of opposite sign is added to any point charge, such that the net charge at each particle position is zero (**Figure 6**, bottom right). Typically, the diffuse charge has a Gaussian-like shape. In this way, the resulting electrostatic potential due to particle α is only due to that fraction of Z_α which is not screened by the added diffuse charge and, therefore, it is short-ranged. A compensating set of diffuse charges is then added to the system to recover neutrality (**Figure 6**, top right). This latter compensating distribution is a periodic and slowly varying function and can be easily computed by a rapidly converging Fourier series.

The Control of the Thermodynamical Ensemble

The previous mechanical formulation of MD can be naturally recast into a statistical mechanics formulation. In fact, the set of instantaneous particle positions $\{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N\}$ and velocities $\{\dot{\boldsymbol{R}}_1, \dot{\boldsymbol{R}}_2, \ldots, \dot{\boldsymbol{R}}_N\}$ uniquely defines a microstate $\Gamma(t)$ in the $6N$-dimensional phase space of the system. Generating

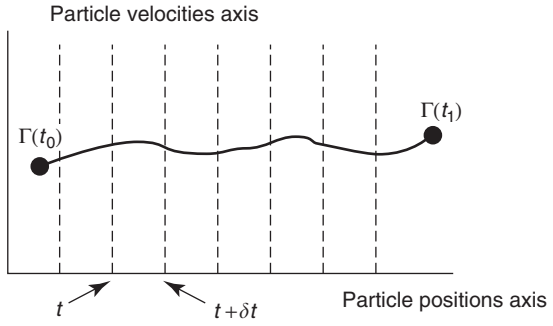


Figure 7 The time evolution of the microstate Γ representing the system under investigation.

trajectories is, therefore, equivalent to generate the time evolution of the microstate $\Gamma(t)$ in the corresponding phase space, as reported schematically in **Figure 7**. This is the basic ingredient for developing a theory based on statistical mechanics.

Consider a given system property (or observable) and define the corresponding microscopic (i.e., computable on the basis of particle positions and/or velocities) operator $\mathcal{O}$. Then, if the microstate $\Gamma(t)$ is known, the instantaneous operator value $\mathcal{O}(\Gamma(t))$ can be evaluated. In other words, one can compute the value assumed by the observable when the system is in that given microstate at time t. According to statistical mechanics, therefore, the macroscopic (i.e., the experimentally measurable) value $\mathcal{O}_{\text{macro}}$ can be defined to be the time average $\langle \mathcal{O}(t) \rangle_{t_{\text{obs}}}$

$$\mathcal{O}_{\text{macro}} = \langle \mathcal{O}(t) \rangle_{t_{\text{obs}}} = \frac{1}{t_{\text{obs}}} \int_0^{t_{\text{obs}}} \mathcal{O}(\Gamma(t))\, \mathrm{d}t \quad [9]$$

where the time integral is performed over the observation time t_{obs}. In practice, the molecular dynamics run is performed over a finite number N_{step} of time steps, so that

$$\mathcal{O}_{\text{macro}} = \langle \mathcal{O}(t) \rangle_{t_{\text{obs}}} = \frac{1}{N_{\text{step}}} \sum_{i=1}^{N_{\text{step}}} \mathcal{O}(\Gamma(t_i)) \quad [10]$$

This major conceptual step allows to refine the previous basic iterative scheme, introducing a system observation procedure as reported in **Figure 8**.

The improvement of the iterative scheme forces us to cope with the choice and the proper implementation of thermodynamical ensembles where time averages are performed. In other words, one needs to operationally select the thermodynamical state of the system so that the trajectory $\Gamma(t)$ is generated only within a subset of the phase space compatible with the imposed conditions of temperature, pressure (or stress), volume, and so on. On the other hand, this formulation widely enlarges the playground for

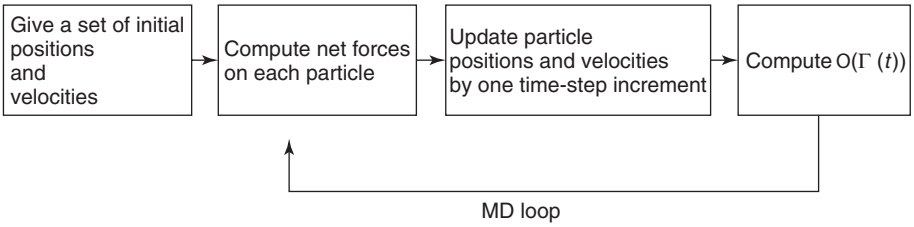


Figure 8 The refined iterative MD loop, including the evaluation of the instantaneous value $\mathcal{O}(\Gamma(t))$ of a given observable. One iteration corresponds to increasing time by an amount δt.

MD investigations of condensed matter systems to include nearly any possible thermodynamical condition.

Constant-Temperature Ensemble

According to statistical mechanics, keeping the system into the canonical ensemble means coupling it to an external thermal reservoir. Several algorithms are available to implement this concept.

In the Anderson method, the thermal bath is mimicked by stochastic collisions (impulsive forces) occasionally occurring at randomly selected particles. The system–bath coupling strength is thus controlled by the collision frequency v which, in turn, is defined by combining physical insight into the system and computational convenience. In order to generate uncorrelated collisionial events, their time distribution (i.e., the distribution of time separations between two consecutive collisions) is Poisson-like. The basic MD loop is simply modified so that, during a time interval equal to the MD time-step, a random particle undergoes a stochastic collision with probability $v\,\delta t$. If this is the case, a new velocity is assigned to that particle as extracted by the equilibrium Maxwell–Boltzmann velocity distribution corresponding to the desired temperature.

A different widely adopted algorithm is offered by the Nosé approach which, at variance with Anderson thermostat, is deterministic. The Nosé method is based on the use of an extended Lagrangian, containing fictitious coordinates and velocities additional to the particle ones. The method is not only important *per se*, but it also represents the prototypical MD approach based on extended Lagrangian formalism, which is very extensively used in modern developments. In order to mimic the coupling to a thermal reservoir, a new fictitious degree of freedom s is added to the system and a classical Lagrangian $\mathscr{L}_{\text{Nosé}}$ is written as

$$\mathscr{L}_{\text{Nosé}} = \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha s^2 \dot{R}_\alpha^2 + \frac{1}{2} m_s \dot{s}^2 - U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N) - g K_B T \ln s \qquad [11]$$

where m_s acts as the effective mass associated with the new coordinate s, and g is the number of degrees of freedom for the extended system. The Nosé Lagrangian allows for energy (i.e., heat) exchange between the physical system and the reservoir (described by s), the thermal inertia being described by the parameter m_s. The new equations of motion are easily derived for system particles

$$\ddot{\boldsymbol{R}}_\alpha = \frac{\boldsymbol{F}_\alpha}{m_\alpha s^2} - 2\dot{\boldsymbol{R}}_\alpha \frac{\dot{s}}{s} \qquad [12]$$

and for the heat reservoir coordinate s

$$m_s \ddot{s} = \sum_{\alpha=1}^{N} m_\alpha s \ddot{R}_\alpha^2 - g\frac{K_B T}{s} \qquad [13]$$

In both cases standard finite difference methods can be used. The new extended system (i.e., the N particles plus the thermal bath) is microcanonical, while it can be proved that ensemble averages performed during a Nosé MD run, reduce to canonical averages. **Figure 9** shows the thermal inertia features of the Nosé approach.

The Nosé method can be generalized into the Nosé–Hoover chains approach: the original thermostat is now coupled to a series of other fictitious thermostats whose role is basically to allow for fluctuations of the variable s. The Nosé–Hoover method has been proved to accurately generate a canonical distribution and to improve the ergodicity of the simulated system.

Constant-Pressure Ensemble

The extended Lagrangian method is useful also to simulate the constant-pressure ensemble, that is, when the system is coupled to an external piston. In this case the new additional variable is the volume V of the simulation cell, which is characterized by an effective mass m_V. Andersen has shown that by introducing the scaled variables

$$\boldsymbol{R}_{s,\alpha} = V^{-1/3}\boldsymbol{R}_\alpha \qquad [14]$$

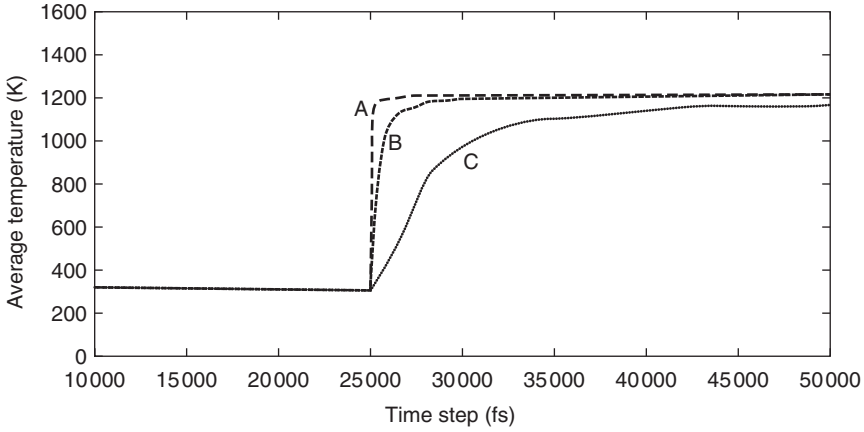


Figure 9 A crystalline silicon system previously thermalized at $T = 300\,\mathrm{K}$ is coupled to a thermostat at $T = 1200\,\mathrm{K}$ via the Nosé Lagrangian at time $25\,000\delta t$. Three different values for the thermostat effective mass m_s have been used, with ratio 1 : 10 : 100, corresponding to the curve A, B, and C, respectively. The thermal inertia effects are governed by the parameter m_s.

and velocities

$$\dot{\boldsymbol{R}}_{s,\alpha} = V^{-1/3}\dot{\boldsymbol{R}}_{\alpha} \qquad [15]$$

the new extended Lagrangian $\mathscr{L}_{\mathrm{Andersen}}$ reads as

$$\mathscr{L}_{\mathrm{Andersen}} = \frac{1}{2}V^{2/3}\sum_{\alpha=1}^{N} m_{\alpha}\dot{R}_{s,\alpha}^{2} + \frac{1}{2}m_V\dot{V}^2 - U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N) - P_{\mathrm{piston}}V \qquad [16]$$

where P_{piston} is the (hydrostatic) pressure applied by the piston to the system. The new equations of motions for scaled particle coordinates are

$$\ddot{\boldsymbol{R}}_{s,\alpha} = \frac{1}{m_{\alpha}V^{1/3}}\boldsymbol{F}_{\alpha} - \frac{2}{3}\dot{\boldsymbol{R}}_{s,\alpha}\frac{\dot{V}}{V} \qquad [17]$$

whereas the new variable V evolves according to

$$\ddot{V} = \frac{P_{\mathrm{int}} - P_{\mathrm{piston}}}{m_V} \qquad [18]$$

where P_{int} is the system internal pressure (e.g., computed by virial theorem). Once again, standard finite difference methods are applied to solve the above equations and the resulting dynamics of the extended system is microcanonical. Ensemble averages performed on unsealed trajectories are proved to be isobaric (with isotropic changes in the simulation cell volume).

The Andersen method can be generalized according to Rahman and Parrinello so as to allow for both volume and shape changes of the simulation cell. Basically, the additional variable V is now replaced by the nine components of the metric tensor $\boldsymbol{H} = (\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3)$, whose columns $\boldsymbol{h}_{1,2,3}$ are vector quantities representing the simulation cell edges. The Rahman–Parrinello Lagrangian (or its extensions) allows for investigating nonisotropic deformations under hydrostatic pressure, as well as the effect of any kind of applied stress.

The Interaction Scheme

The physics produced by an MD simulation crucially depends upon the adopted interatomic potential $U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N)$. A good interaction model should be accurate (i.e., it should provide accurate estimation of total energy and related quantities, for example, its derivatives), reliable (i.e., it should be firmly rooted into a fundamental description of interparticle interactions and it should provide deep physicochemical insight into the investigated system), and transferable (i.e., it should be able to manage the same system in arbitrary conditions of temperature, pressure/stress, state of aggregation, order/disorder, external load).

The simplest possible approach to elaborate a theory for U consists in formally expanding the total potential into a series of two-body $U^{(2)}$, three-body $U^{(3)}, \ldots,$ n-body $U^{(n)}$ terms as

$$U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N) = \sum_{\alpha>\beta} U^{(2)}(\boldsymbol{R}_{\alpha}, \boldsymbol{R}_{\beta}) + \sum_{\alpha>\beta>\gamma} U^{(3)}(\boldsymbol{R}_{\alpha}, \boldsymbol{R}_{\beta}, \boldsymbol{R}_{\gamma}) + \cdots \qquad [19]$$

The series is then arrested at a given order, as guessed by the chemical nature of the underlying interparticle bond. For instance, just a two-body term is needed to model simple metals (although at a very crude level), while both a two-body and a three-body term are

indeed required to simulate materials with directional and/or covalent bonding. Full set of interatomic potentials are now available for nearly any class of materials. Most of them are short-ranged (i.e., the sums in the above formal expansion are limited just to the nearest-neighbor particles), but long-ranged potentials (i.e., containing Coulomb interactions) are available as well for ionic and/or polar materials.

The actual expression for the $U^{(2)}$, $U^{(3)}$, …, $U^{(n)}$ functionals is either guessed by physicochemical intuition on the investigated system, or simply determined for numerical convenience. In any case, the functionals depend upon empirical parameters that need to be fitted once for all onto a suitable database of system properties. The resulting simulation is therefore classified as a model-potential MD (MPMD). The empirical character of U dramatically improves the overall numerical efficiency of the simulation, since the resulting CPU workload typically scales (for short-ranged potentials) as $m \times N$, where m is the (average) number of neighbors for each of the N system particles. Present-day massively parallel computers allow for simulations up to $N \sim 10^8$, although a more typical number is below 10^5–10^6. Unfortunately, the same empiricity also limits (sometimes even severely) the transferability of the potential itself. In other words, model potentials are computationally inexpensive, but often not accurate enough for predictive materials modeling.

In order to overcome this limitation, one can develop a theory at the more fundamental quantum level. Within the Born–Oppenheimer approximation and adopting the one-electron picture, the total potential energy of a system of ion cores (nuclei and core electrons) and (valence) electrons can be expressed as

$$U(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N) = U_{\mathrm{ee}} + U_{\mathrm{ei}} + U_{\mathrm{ii}} \qquad [20]$$

where i and e labels indicate ions and electrons, respectively. In order to compute such an expression, electron energies ε_n and wave functions Ψ_n should be known. Once they are known, for example, by solving numerically the one-electron Schrödinger equation, the total potential energy U can be written as

$$\begin{aligned} U &= U_{\mathrm{ie}} + U_{\mathrm{ii}} + U_{\mathrm{ee}} \\ &= 2\sum_{n}^{\mathrm{occ}} \varepsilon_n + U_{\mathrm{ii}} - U_{\mathrm{ee}} \end{aligned} \qquad [21]$$

where the sum is perfomed over the energies of the occupied states and the $-U_{\mathrm{ee}}$ contribution corrects the double counting of the electron–electron interactions in the first term. The factor of 2 takes into account the spin degeneracy. By applying the Hellmann–Feynman theorem (with or without Pulay's correction, depending on the choice of basis orbitals) it is straightforward to obtain interparticle forces which, therefore, contain a full quantum mechanical picture of particle interactions. The two most computationally intensive steps of this first-principles MD (FPMD) scheme are: (1) Schrödinger equation solution and (2) self-consistent evaluation of U_{ee}. The resulting CPU workload typically scales as N^3 and a large prefactor is also found, thus preventing in practice first-principles simulations with much more than $\sim 10^2$ particles.

Although the above formalism has been developed from first principles, it is convenient – in order to save computer time – to introduce some important approximation, namely the tight-binding (TB) picture. In this framework, the electron wave function Ψ_n values are expanded as linear combinations of atomic orbitals $\{\phi_{l\alpha}\}$:

$$\Psi_n = \sum_{l\alpha} c^n_{l\alpha} \phi_{l\alpha} \qquad [22]$$

where l is the quantum number index. For simplicity, the basis set is assumed to be orthogonal. The matrix elements $\langle \phi_{l'\beta} | h | \phi_{l\alpha} \rangle$ of the one-electron Hamiltonian are then treated – according to Slater–Koster theory – as disposable constants to be fitted, instead of computed, self-consistently. The fitting step is typically carried out assuming: (1) a minimal basis set, (2) short-range hoppings, and (3) two-center approximation. Furthermore, the difference $U_{\mathrm{ii}} - U_{\mathrm{ee}}$ is replaced by a totally empirical effective potential U_{eff} which is calibrated similarly to model potentials. It results a semi-empirical MD (SEMD) scheme with improved accuracy and transferability with respect to MPMD. However, the computational workload is dramaticaly reduced with respect to an FPMD approach since no self-consistency is required. However, it is still significantly heavy since the diagonalization of the TB matrix is needed at each step of the MD loop. This operation has a complexity scaling as N^3. SEMD simulations are typically performed on systems containing as many as several hundred particles.

The SEMD approach represents an interaction scheme bridging model-potential to first-principles simulations. Although it could be of great help in many cases, it nevertheless leaves unsatisfied the most demanding requirements of accuracy and transferability. There is, in other words, a real need of parameter-free MD simulations in modern condensed matter and materials physics. The best-available answer to such a need is represented by the Car–Parrinello MD (CPMD) scheme. Within the CPMD method the properties of the system are computed – according to the basic formulation of MD – as time

averages, while particle interactions are derived from the underlying electronic structure, provided by density-functional theory (DFT). The key idea consists in replacing the real quantum (ions + electrons) system with a fictitious classical one. The fictitious system is in fact described by the Car–Parrinello Lagrangian $\mathscr{L}_{\mathrm{CP}}$:

$$\begin{aligned}\mathscr{L}_{\mathrm{CP}} = \mu \sum_{j}^{\mathrm{occ}} \int |\dot{\psi}_j(\boldsymbol{r})|^2 \, \mathrm{d}\boldsymbol{r} + \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha \dot{\boldsymbol{R}}_\alpha^2 \\ - E[\{\psi_j\}, \{\boldsymbol{R}\}] \\ + \sum_{jk} \Lambda_{jk} \left(\int \psi_j(\boldsymbol{r})^* \psi_k(\boldsymbol{r}) \, \mathrm{d}r - \delta_{jk} \right) \end{aligned} \quad [23]$$

where $\boldsymbol{r}$ denotes electron coordinates and $E[\{\psi_j\}, \{\boldsymbol{R}\}]$ is the DFT potential energy functional of the coupled electron–ion system. The parameter μ plays the role of an effective mass for the electronic degrees of freedom. The Lagrangian multipliers Λ_{jk} are simple holonomic constrains imposing the orthogonality of the electron wave functions $\{\psi_j(\boldsymbol{r})\}$. As usual for extended systems, a new set of equations of motion can be derived for both the ionic degrees of freedom,

$$m_\alpha \ddot{\boldsymbol{R}}_\alpha = -\nabla_\alpha E[\{\psi_j\}, \{\boldsymbol{R}\}] \quad [24]$$

and the electronic degrees of freedom

$$\mu \ddot{\psi}_j = -\frac{\delta E[\{\psi_j\}, \{\boldsymbol{R}\}]}{\delta \psi_j^*} + \sum_{jk} \Lambda_{jk} \psi_k \quad [25]$$

In principle, the ionic trajectories generated by $\mathscr{L}_{\mathrm{CP}}$ do not coincide with the real ones, unless the functional $E[\{\psi_j\}, \{\boldsymbol{R}\}]$ is at the minimum. However, it has been proved that by a proper choice of the fictitious electronic mass μ, the electrons follow adiabatically the ionic motion, since their generated timescale is much shorter than the ionic one. When this condition is fulfilled, any sizeable deviation of the nuclear trajectories from the minimum electron energy surface corresponding to the actual ionic configuration (Born–Oppenheimer surface) is typically observed only after times much longer than a typical MD time step. Consistently, meaningful time averages can be performed. The key point of the CPMD method is that a self-consistent minimization of the total energy functional $E[\{\psi_j\}, \{\boldsymbol{R}\}]$ is not required at each step of the simulation. Rather, the equations of motion for the electronic degrees of freedom act so as to keep electrons constantly close to (ideally at) the Born–Oppenheimer surface. As usual, the MD method (e.g., control of temperature, pressure, stress) can be applied to CPMD, thus making it the most fundamental and superior simulation scheme now available for condensed matter systems.

See also: Molecular Dynamics, Car–Parrinello.

PACS: 71.15.Pd; 31.15.Qq

Further Reading

Allen MP and Tildesley DJ (1987) *Computer Simulation of Liquids*. Oxford: Clarendon Press.

Allen MP and Tildesley DJ (1993) *Computer Simulation in Chemical Physics*. Dordrecht: Kluwer Academic.

Ciccotti G and Hoover WG (eds.) (1986) *Proceedings of 97th International School of Physics "E. Fermi" on Molecular Dynamics Simulations of Statistical-Mechanics Systems*. Amsterdam: North-Holland.

Frenkel D and Smit B (1996) *Understanding Molecular Simulation*. San Diego: Academic Press.

Haile JM (1992) *Molecular Dynamics Simulation*. New York: Wiley-Interscience.

Reif F (1987) *Fundamentals of Statistical and Thermal Physics*. Singapore: McGraw-Hill.

Voter AF (ed.) (1996) *Interatomic Potentials for Atomistic Simulations*. Materials Research Society Bulletin, vol. 21(2).

Molecular Dynamics, Car–Parrinello

R Rousseau, International School for Advanced Studies, Trieste, Italy
S Scandolo, The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

Introduction

The idea of simulating the microscopic behavior of matter by directly solving Newton's equations on the computer for a collection of atoms was first proposed by E Fermi in the 1940s. The technique is generally referred to as "molecular dynamics (MD)." The exponential increase of computer power and the continuous refinement of simulation techniques have turned molecular dynamic simulations from a toy model of statistical physics into an extremely valuable and predictive tool nowadays. The application of molecular dynamics is widespread in many fields of condensed matter physics, materials science, chemistry, and molecular biology.

Evolving Newton's equations

$$\ddot{\boldsymbol{R}}_I = -\frac{1}{M_I}\frac{\partial V}{\partial \boldsymbol{R}_I}(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N) \quad I = 1, N \qquad [1]$$

for a collection of N atoms with positions $\{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N\}$ and masses $\{M_1, M_2, \ldots, M_N\}$ implies that the potential $V(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N)$ of interaction among them must be given. Determining the interatomic potential is one of the more challenging problems in molecular dynamic simulations, as discussed in detail in Colombo's contribution. In this article, the focus is on how to determine the interatomic potential V from "first principles," based on the laws of quantum mechanics, and without any empirical input. By doing so, the computational complexity of a molecular dynamic simulation increases dramatically, implying, for example, that a much smaller number of atoms can be simulated and for much shorter times than conventional molecular dynamic simulations. First-principle (or *ab initio*) molecular dynamics combines the advantages of conventional molecular dynamics in terms of statistical sampling with the accuracy of quantum mechanical methods in describing electronic structure and chemical bonding in condensed matter.

First-Principles MD

Here, the concentration is on the problem of determining a reliable interatomic potential, to be used in the MD eqn [1]. The approach that is followed starts from the "first-principles," namely from the fundamental quantum mechanical laws that govern the behavior of a collection of electrons and nuclei. Although these laws are well known – the Schrödinger equation for the coupled electrons–nuclei system – solving them on the computer without any further approximation would require an exceptional effort, many orders of magnitude more than what is available at present in terms of computer memory and speed. The first approximation to be introduced is the so-called "adiabatic," Born–Oppenheimer approximation. This approximation stems from the consideration that electrons react much faster than nuclei to external perturbations, due to their smaller mass. As a consequence, electrons can be considered to be always in their ground state, for every instantaneous nuclear configuration $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$. Therefore, the interaction potential V can be expressed as

$$V(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N) = \langle \psi_\circ | H_e(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N) | \psi_\circ \rangle \qquad [2]$$

where $\psi_\circ$ is the electronic ground-state wave function calculated for the nuclear configuration $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$, according to the Hamiltonian

$$H_e = -\sum_{i=1,N_e} \frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i=1,N_e}^{j=1,i-1} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_{i=1,N_e}^{I=1,N} \frac{Z_I e^2}{|\boldsymbol{r}_i - \boldsymbol{R}_I|} + \sum_{I=1,N}^{J=1,I-1} \frac{Z_I Z_J e^2}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \qquad [3]$$

where N_e is the number of electrons and Z_I the atomic number of atom I. Note that the Hamiltonian [3] acts on the electrons, and depends on $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$ only parametrically. This implies that the ground-state wave function ψ_0 will also depend parametrically on $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$, that is, a different wave function is obtained for every nuclear configuration. Although the Born–Oppenheimer approximation allows one to simplify the problem considerably, finding ψ_0 for the Hamiltonian [3] remains a very difficult numerical task, at least for systems with more than a few electrons, because of the many-body nature of the electron–electron Coulomb interaction (second term in the right-hand side of [3]). Among the approximations that allow one to overcome the problem, those based on the density-functional theory (DFT) are probably among the most widespread and accurate, particularly when dealing with relatively large systems (hundreds of electrons or more). DFT is an exact theory that allows one to recast the problem of finding the many-electron ground state of [3] into the problem of finding the ground state of a system of independent electrons in a potential that explicitly depends on the electron density. Although the theory is formally exact, approximations have to be made in order to derive the functional form of the potential. The most common of these approximations is the so-called local-density approximation (LDA), whereby the potential is approximated locally with the functional form of the uniform electron gas, which is known almost exactly from numerical calculations. The LDA can be systematically improved by adding contributions from the local gradient of the density (gradient corrections, or GGA).

A final approximation that is commonly made in first-principles calculations is to assume that core states are actually frozen in their atomic shape and energy. Orthogonalization of the valence states to the core states is ensured by the replacement of the electron–nuclei Coulomb potential, the third term in the right-hand side of [3], with a so-called "pseudopotential," whose scattering properties are constrained to be identical to those of the full atomic potential. When the basis set over which the wave functions are expanded consists of plane waves, a very popular choice in first-principles codes, smoother

pseudopotentials may be less demanding in terms of basis set size.

The above approximations have finally brought the problem into the category of those that can be nowadays solved numerically on a powerful computer. To summarize, a rough flow chart of how the MD time step is accomplished in a first-principles MD code consists of: (1) choosing an initial set of nuclear positions $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$, (2) solving the electron Hamiltonian [3], that is, determining the ground-state wave function for that particular nuclear configuration, (3) calculating the forces on the nuclei as the derivative of the potential [2] [N.B. this can be accomplished either numerically, by finite differences, or by exploiting the so-called Hellman–Feynman theorem, whereby the derivative of the potential can be expressed as the expectation value of the derivative of the Hamiltonian in [2], on the ground-state wave function], (4) evolving Newton's equations for all nuclei with a finite time step δt, that is, determining the new set of nuclear positions corresponding to time $t + \delta t$, and finally going back to (2) with the new set of positions. This procedure is iterated for a large number of time steps, depending on the system and on the property under study. In order to integrate Newton's equations properly, the time step δt of integration must be much smaller than the fastest period of oscillation of the nuclei. Since typical periods (vibrations, phonons, etc.) are in the range 10^{-14}–10^{-12} s, typical δt are in the range of a fraction of 10^{-16}–10^{-15} s, which means that one picosecond of simulation requires $\sim 10^4$ time steps. One of the computationally heavier steps of a first-principles MD simulation is the determination of the ground-state wave function. Considering that this step has to be repeated 10^4 times, most of the methodological effort has been focused on the search for efficient ways to update the wave function from one time step to the next one. Most of the differences between existing codes for first-principles MD [CASTEP (Cambridge), VASP (Vienna), CPMD (Stuttgart/Zürich), PARATEC (Berkeley), and FPMD (Trieste), just to name a few] can be ascribed to different techniques in the way wave functions are evolved.

The Car–Parrinello Method

A rather original and very efficient way to evolve the wave functions was proposed in 1985 by Car and Parrinello. Instead of solving the Hamiltonian [3] at each time step very accurately, they relaxed the constraint that the wave function be exactly in the ground state. The exact solution was replaced by a fictitious dynamics of the wave function about the exact ground state. In the Car–Parrinello method, the solution of the Hamiltonian [3], which is equivalent to minimizing the right-hand side of [2] with respect to the wave function, is replaced by the dynamical evolution of the following "extended" Lagrangian

$$\mathscr{L} = \frac{1}{2}\sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle + \frac{1}{2}\sum_I M_I \dot{\boldsymbol{R}}_I^2 - E[\{\psi_i\}, \{\boldsymbol{R}_I\}] + \sum_{ij} \Lambda_{ij}(\langle \psi_i | \psi_j \rangle - \delta_{ij}) \quad [4]$$

where E is the DFT functional, expressed here in terms of the wave functions and, parametrically, of the positions of the nuclei, μ is a fictitious mass that can be used to control the timescale of the electronic dynamics, and Λ_{ij} are Lagrange multipliers that ensure the orthogonality of the wave functions. The second term on the right-hand side of [4] is the kinetic energy of the ions T, while the first term is the fictitious kinetic energy of the electrons T_e. The equations of motion that can be obtained from the above Lagrangian are:

$$\mu \ddot{\psi}_i = -\frac{\delta E}{\delta \psi_i} + \sum_j \Lambda_{ij}\psi_j \quad [5]$$

$$\ddot{\boldsymbol{R}}_I = -\frac{1}{M_I}\frac{\partial E}{\partial \boldsymbol{R}_I} \quad [6]$$

Equation [5] defines the dynamics of the wave functions, which oscillates (by virtue of the second-order time derivative) around the equilibrium condition

$$\frac{\delta E}{\delta \psi_i} = \sum_j \Lambda_{ij}\psi_j \quad [7]$$

Notice that eqn [7] is the mathematical rephrasing of the DFT prescription of minimizing the DFT functional E with respect to variations of the (orthogonal) set of orbitals. Thus, by definition, $E[\{\psi_{\circ i}(\{\boldsymbol{R}_I\})\}, \{\boldsymbol{R}_I\}] = V(\{\boldsymbol{R}_I\})$. Equation [6], instead, coincides with Newton's eqn [1] as obtained from the first-principles potential [2], with the only difference that the wave functions in [6] are not at the exact ground state.

If the atoms are not allowed to move, the period of oscillation of the wave functions around the ground-state scales with $\mu^{1/2}$, which means that for a sufficiently small value of μ, the oscillations can be made much faster than the typical timescales of the dynamics of the nuclei. In this limit, the effect of the fast oscillations averages out on the timescale of the dynamics of the nuclei, and the resulting net force on the nuclei is very similar to the force that would be calculated with the electrons in their ground state.

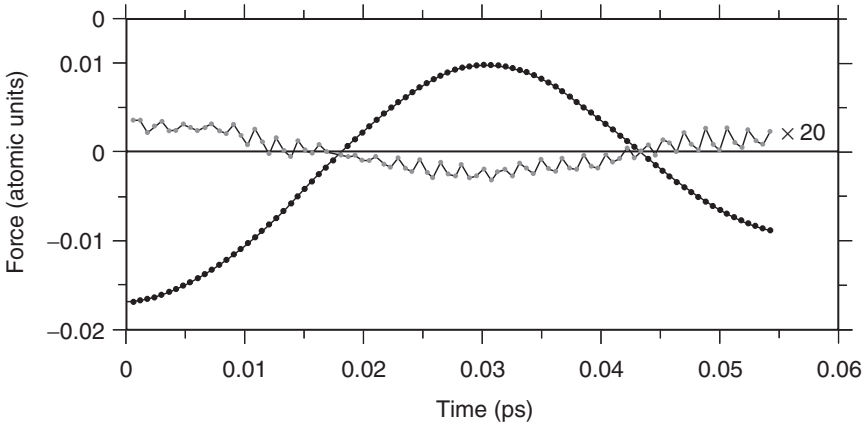


Figure 1 Time dependence of a Cartesian component of the force on a selected atom in a simulation of crystalline silicon with a cell containing 64 atoms at room temperature. The smooth curve corresponds to the force calculated along the Car–Parrinello MD trajectory (1 atomic unit corresponds to $\sim 14.4\,\mathrm{eV\,\AA^{-1}}$). Also shown is the difference (multiplied by 20) between the Car–Parrinello and Born–Oppenheimer force along the Car–Parrinello trajectory.

This is only strictly true if the atomic positions are kept fixed during the electronic dynamics. The consequences of the Car–Parrinello approximation when the atoms are allowed to evolve according to the classical equations of motion [6] are illustrated in **Figure 1**, where the force acting on a silicon atom in a Car–Parrinello simulation of crystalline silicon at room temperature is compared with the force calculated by putting the electrons in their instantaneous ground state at each step of the MD trajectory. The difference (amplified by a factor $\times 20$ in the figure) can be clearly seen to contain a fast component (much faster than the nuclear oscillation) and a slow component with frequency comparable with the nuclear motion. While the fast component is harmless as its contribution to the nuclear dynamics averages away on the timescale of the nuclear oscillations, the slow component interferes with the nuclear dynamics giving rise to a systematic bias of the nuclear trajectories, and must therefore be strictly controlled to ensure that the trajectory generated with the Car–Parrinello equations of motion retains the accuracy of a trajectory generated in the Born–Oppenheimer limit, that is, with the electrons in their instantaneous ground state. In the case of silicon, the difference is anyway very small, but in other systems, such as ionic liquids, the difference can be significant. It is important to remark that the difference can be systematically reduced by decreasing μ, which implies, however, that a shorter time step needs to be used to integrate the electronic equations of motion.

Spectral Properties from *ab initio* Molecular Dynamics

MD calculations are very useful in extracting equilibrium thermodynamics, static and dynamical information from a system. One of the most important and useful quantities that can be evaluated from an MD simulation is the time correlation function C_{AB} that correlates two time-dependent quantities A and B,

$$C_{AB}(t) = \langle A(t_0)B(t_0+t)\rangle \qquad [8]$$

where the average $\langle \ldots \rangle$ is over the initial time t_0.This function is important because it provides detailed information on the dynamics of a system. Furthermore, the Fourier transform of a time correlation function is often related to relevant experimental observables. For example, the Fourier transform of the nuclear velocity–time autocorrelation function is related, in solids, to the phonon density of states $D(\omega)$:

$$D(\omega) \propto \int_0^\infty \mathrm{e}^{\mathrm{i}\omega t} \sum_I \langle \boldsymbol{V}_I(t_0)\cdot \boldsymbol{V}_I(t_0+t)\rangle \,\mathrm{d}t \qquad [9]$$

where $\boldsymbol{V}_I$ is the velocity of atom I. The time correlation formalism was recently extended for the calculation of the infrared absorption coefficient $\alpha(\omega)$. The absolute absorption coefficient of an isotropic system can be obtained from

$$\alpha(\omega) = \frac{4\pi\omega}{3nc}\,\mathrm{Im}\sum_\gamma \chi_{\gamma\gamma} \qquad [10]$$

where n is the refractive index and c the speed of light. The imaginary part of the electronic susceptibility $\chi_{\gamma\gamma}$ is given by

$$\mathrm{Im}\,\chi_{\gamma\gamma} = \frac{V}{2\hbar}(1-\mathrm{e}^{-\hbar\omega/k_\mathrm{B}T})\int_0^\infty \mathrm{e}^{\mathrm{i}\omega t} \times \langle P_\gamma(t_0)P_\gamma(t_0+t)\rangle \,\mathrm{d}t \qquad [11]$$

where $P_\gamma(t_0)$ is the polarization vector along the γ direction. The total polarization can be evaluated from an *ab initio* (e.g., Car–Parrinello) MD simulation using the discrete Berry phase at the single-point limit.

An Example: The Water Molecule

Here, the application of Car–Parrinello MD is illustrated using a simple example – the calculation of the infrared spectrum of a H_2O molecule. To focus attention on both the strengths and weaknesses of the Car–Parrinello approach, a $T = 300$ K MD simulation in a 10 Å periodically replicated cubic box, is performed in three ways: (1) using a large, $\mu = 400\, m_e$ (m_e the electronic mass) and an MD time step $\delta t = 0.1$ fs for a duration $\Delta t = 2$ ps; (2) using smaller, $\mu = 200\, m_e, \delta t = 0.05$ fs, and $\Delta t = 1$ ps; (3) with Born–Oppenheimer forces $\delta t = 0.25$ fs and $\Delta t = 1$ ps. Note that the calculation is designed to serve only as an illustration of basic principles, but does not provide a numerically accurate representation of the vibrational properties of a gas-phase water molecule. The relevant energies for 2000 steps of simulation (1) are shown in **Figure 2**. The DFT energy E is dominated by oscillations with a period of ~ 1 fs and amplitude of 0.004 atomic units (1 atomic unit of energy corresponds to ~ 27.21 eV), resulting from the stretching motion of the O–H bonds. The sum of the nuclear kinetic (T) and DFT (E) energies, which corresponds to the total energy of a classical MD simulation and should be conserved for exact Newtonian dynamics, also oscillates with a similar frequency as the nuclear motion, but with a smaller amplitude. This is due to the fact that the Hamiltonian contains an extra term corresponding to the fictitious kinetic energy of the electrons (T_e), which when summed into $T + E + T_e$, gives a perfectly conserved quantity. As can be inferred from the above discussion, the magnitude of the oscillations in $T + E$, is a consequence of the choice of $\mu = 400\, m_e$, which induces a drag on the nuclei, hence perturbing the nuclear dynamics. The effect may be reduced by decreasing $\mu = 200\, m_e$ but this requires a simultaneous decrease in $\delta t = 0.05$ fs to retain the same degree of energy conservation. Finally, the Born–Oppenheimer forces can be used to eliminate the effect entirely, with the benefit that a larger $\delta t = 0.25$ fs may be employed. However, each MD step now requires ~ 10 self-consistent cycles to converge the wave function coefficients to 10^{-6}, which is required to obtain sufficiently converged forces for stable Newtonian dynamics. Thus, run (1) is the computationally least demanding in terms of CPU time, but also the least accurate, whereas for run (3) the converse is true.

The vibrational spectrum of water, above $2000\,\mathrm{cm}^{-1}$, is dominated by two fundamental vibrational frequencies ν_1 antisymmetric O–H stretching, and ν_2 symmetric O–H stretching, both of which may be adequately sampled on the 1–2 ps timescale.

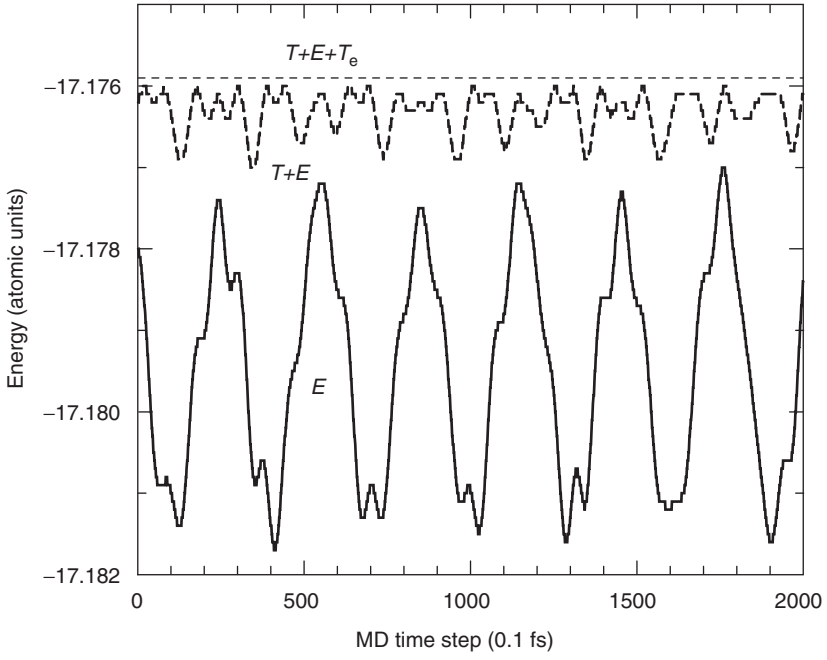


Figure 2 Time dependence of the DFT (E), nuclear kinetic (T), and total ($T + E + T_e$) energies in a Car–Parrinello MD simulation of a water molecule.

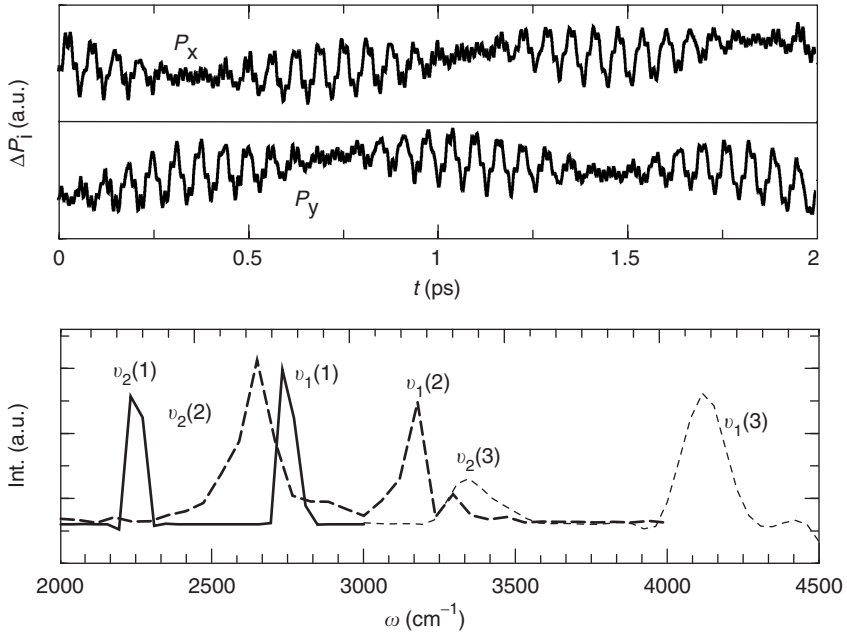


Figure 3 (Upper panel) Time dependence of the Cartesian components of the polarization vector of a water molecule lying on the *x*–*y* plane. (Lower panel) IR absorption spectrum in the region of the O–H stretching and bending modes calculated with Car–Parrinello MD with two different values of the electronic fictitious mass, as well as with Born–Oppenheimer MD.

To obtain the IR adsorption spectrum, the polarization vector is calculated at each MD time step. The water molecule is constrained to lie within the *x*,*y*-plane providing only two components of the polarization vector (P_x and P_y) which are plotted in **Figure 3** for run (1). The two components show oscillations on the femtosecond timescale due to O–H stretching, superimposed onto lower frequencies arising from the bending mode and a weak coupling to the periodic images. Upon applying eqns [10] and [11], the IR adsorption spectrum (see **Figure 3**) is obtained, which, for trajectories of a picosecond in duration, will reliably yield the peak positions but not the relative intensities or the peak shape. All three MD trajectories provide qualitatively the same spectrum, where υ_2 is red-shifted from υ_1 by $\approx 20\%$ of their average. However, the Born–Oppenheimer dynamics provides values, which are in excellent accord with the calculated harmonic frequencies of $\upsilon_1 = 3350\ \text{cm}^{-1}$ and $\upsilon_2 = 4150\ \text{cm}^{-1}$, whereas the average O–H frequency of the Car–Parrinello trajectories are red-shifted by a factor which is proportional to $\sqrt{\mu}$. Note that this rigid shift of the Car–Parrinello peaks allows one to formulate a simple correction by which the nuclear masses may be lessened to account for the electronic drag.

Conclusions

Car–Parrinello *ab initio* MD provides a computationally efficient scheme, whereby Newton's equations of motion may be integrated subject to a potential energy function derived from first-principles electronic structure theory for systems of up to several hundreds of atoms. This procedure can provide a stable nuclear dynamics, and thus allows for an effective sampling of configuration space to obtain reliable statistical distribution functions. For many systems, the Car–Parrinello approach is computationally more efficient than a procedure by which the forces required in MD simulations are obtained from exact ground-state wave functions. However, the method requires that one chooses the fictitious electronic mass μ with great care and for accurate calculation of nuclear dynamical variables, the forces must be corrected due to the drag on the nuclei by the electronic degrees of freedom.

See also: Density-Functional Theory; Electronic Structure Calculations: Plane-Wave Methods; Geometrical Phase and Polarization in Solids; Molecular Dynamics Calculations; Pseudopotential Method.

PACS: 71.15.Pd; 71.15. – m; 61.20.Ja; 61.43.Bn

Further Reading

Car R and Parrinello M (1985) Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters* 55: 2471–2744.

Marx D and Hutter J (2000) *Ab initio* molecular dynamics: theory and implementation. In: Grotendorst J (ed.) *Modern Methods*

and Algorithms of Quantum Chemistry, NIC Series, vol. 1, pp. 329–477. FZ Jülich: Germany.

Remler DK and Madden P (1990) Molecular dynamics without effective potentials via the Car–Parrinello approach. *Molecular Physics* 70: 921–966.

Scandolo S (2002) First-principles molecular dynamics simulations at high pressure. In: Chiarotti G, Hemley R, and Ulivi L (eds.) *Proceedings of the International School of Physics "E. Fermi" on High Pressure Phenomena Course CXLVII*, pp. 195–214. Amsterdam: IOS Press.

Tangney P and Scandolo S (2002) How well do Car–Parrinello simulations reproduce the Born–Oppenheimer surface? Theory and examples. *Journal of Chemical Physics* 116: 14–24.

Molecular Monolayers and Thin Films

M K Sanyal, Saha Institute of Nuclear Physics, Bidhannagar, Kolkata, India

Introduction

The last few decades has witnessed a dramatic improvement in the growth processes of thin films that involve sequential deposition of molecular monolayers on a substrate. Here, thin films and monolayers of organic molecules are discussed as the other films have been discussed elsewhere in this encyclopedia. Organic films are being studied widely as the properties of these materials can be tuned by modifying specific functional groups of the molecules and the tailor-made films thus formed are useful for various technological and biological applications. Moreover, organic films provide easy-to-form model systems to enhance the understanding regarding nanostructured materials. The size of nanostructured materials – or nanomaterials – in at least one of the three directions becomes comparable with the molecular size; as a result, the effect of molecular confinement and reduced dimensionality can alter various physical and chemical properties of these materials. For example, one can form two-dimensional (2D) nanomaterials simply by making ultrathin films, where molecules are confined between the top surface and film-substrate interface. As the sizes of organic molecules are big, even relatively thick organic films start exhibiting theoretically predicted properties of 2D systems.

Perhaps the first controlled experiments on molecular monolayers confined in air–water interface were performed by Pockels in 1891 and then, this technique of monolayer formation of amphiphilic molecules on water surface was perfected by Langmuir around 1917. One part (hydrophilic) of these molecules is soluble in water while the other part (hydrophobic) is insoluble. On water, amphiphilic molecules, which are commonly called surfactants, take a preferential orientation with the hydrophilic part (also called the head) immersed in water while the hydrophobic part (also called the tail) stays away from water. The amphiphilicity of a molecule depends on the balance between the hydrophilicity of the head group and the hydrophobicity of the tail, the latter depending on the chain length. Blodgett succeeded in transferring, monolayer by monolayer, these amphiphilic molecules from the water surface to solid (hydrophilic or hydrophobic) substrates to form multilayer organic films. The multilayer films deposited by this technique are known as Langmuir Blodgett (LB) films and the monolayers at the air–water (or any gas–liquid) interface used for this deposition has been termed as Langmuir monolayers.

Self-assembled monolayer (SAM) formation is another technique to grow thin organic films either from a solution or from a gas phase. In this technique, one also uses differential affinity of a particular portion of an organic molecule with a substrate. The term "self-assembly" signifies a spontaneous and natural process observed all around us in the formation of molecular assemblies ranging from crystals to mammals. Various materials are being fabricated using this process of self-assembly, which involves designing molecules and supramolecular entities so that shape-complementarities causes them to aggregate into desired structures. Formation of thin organic films through the process of self-assembly has a number of advantages like atomic-level modification of structure and production of structures that are relatively defect-free, self-healing, and rapid. SAM structures are generally prepared from solutions although molecules in vapor or gas phase can also be used for this purpose. The organic molecular beam deposition technique is generally used for SAM formation from the gas phase. In this technique, an organic material is evaporated in an ultrahigh vacuum chamber to get it deposited on the substrate directly as done for inorganic materials in a conventional molecular beam epitaxy (MBE) chamber.

Mentioned here briefly are two more types of organic thin films, namely, thin polymer films and liquid-crystal films. These systems are not discussed here as it has been covered elsewhere in this encyclopedia. Thin polymer films are generally deposited on

a solid substrate using spin-coating technique. In this technique, a polymer solution in a volatile solvent is spread on a substrate and then it is rotated at high speed so that the solvent leaves the surface, keeping the polymer on it. One can control the thickness of the film by altering the spinning speed and concentration of the polymer in the solution used for this purpose. Molecular layering and associated interesting mechanical and thermal properties of polymers have been observed in ultrathin films. A unique property of smectic liquid crystals is their ability to form films that are freely suspended or freestanding over an aperture in a frame. Apart from various technological applications, the availability of liquid crystalline films of variable thickness helps one to study the crossover from three- to two-dimensional behavior. 2D crystals melt through hexatic phase – a phase having long-range orientational order and short-range translational order – and both crystal-to-hexatic and hexatic-to-liquid phase transitions are theoretically predicted to be continuous in nature. Ultrathin smectic membranes (freely suspended liquid-crystal films) have been used extensively to study this melting process. It is interesting to note that one gets back conventional 3D melting with fixed melting point in thick liquid crystal films.

Experimental Techniques

Recent advances in microscopy and scattering techniques have revived the interest in organic thin films as one can now monitor the structure and orientation of molecules responsible for generating particular microscopic and macroscopic properties. Sensitivity of various spectroscopy techniques has also improved and one can now even take spectroscopic data from monolayers of organic films. Recent advances in computer simulation techniques, molecular modeling, and the phenomenological theory of phase transitions have also provided substantial amount of knowledge regarding organic monolayers and thin films.

The grazing incidence X-ray and neutron scattering techniques have provided most of the structural and morphological information regarding the organic monolayers and thin films. Three different measurements, namely specular reflectivity, diffuse scattering, and grazing incidence diffraction (GID), are generally used. In these studies, the surface of the sample is located in the x–y plane and a well-collimated incident beam of radiation is kept in the x–z plane. The scattering measurements are performed as a function of the incident angle α (starting from few milliradians) with the x–y plane, and exit angles β and ϕ with the x–y and x–z planes, respectively. The components of the wave vector transfer, q (q_x, q_y, q_z), corresponding to the incident wavelength λ are $q_x = (2\pi/\lambda)$ ($\cos\beta\cos\phi - \cos\alpha$), $q_y = (2\pi/\lambda)$ ($\cos\beta\sin\phi$), $q_z = (2\pi/\lambda)$ ($\sin\alpha + \sin\beta$). In reflectivity measurements, $\alpha = \beta = \theta$, so that $q_x = 0$ and $q_z = (4\pi/\lambda)(\sin\theta)$, the scattering vector q is perpendicular to the surface (specular direction), and one obtains information about the mean electron density profile (EDP) as a function of the depth (z). The diffuse scattering measurements are performed as a function of q_x and q_y to obtain information regarding the morphology of surfaces and interfaces through a height–height correlation in an interface and between interfaces.

One can now measure X-ray diffuse scattering data over 10 orders of magnitude in intensities at a synchrotron source, but neutron scattering studies are basically limited to reflectivity measurements due to lack of intensities. In the neutron scattering studies, the measured intensity depends on the scattering length of the nucleus, which varies considerably from one element to another and even for different isotopes of the same element. In particular, the deuterium nucleus has a large positive scattering length whereas the hydrogen nucleus has a low negative value. This property has been exploited quite extensively in the structural investigation of organic monolayers and thin films. Here, one uses this contrast variation technique to highlight part of a molecule or selected molecules or monolayers in a film by replacing hydrogen with deuterium and then investigate the film structure with neutron scattering.

In GID, one obtains in reciprocal space, rods of intensities – known as the crystal truncation rods (CTR) – corresponding to a surface structure instead of spots obtained for 3D structures. GID studies are performed by keeping the incident angle α fixed just below the critical angle, α_c, and by measuring intensity as a function of β and ϕ. This technique has been used primarily to study the structures of the different phases of the Langmuir monolayer. As α is less than α_c, X-rays are confined within a depth of a few tens of nanometers below the interfacial plane, whereas they travel in the plane (x and y directions) with a wavelength close to the free-space wavelength. This is the condition for GID that is exploited to study the structure of the interfacial region at atomic resolution, and any periodicity in the positions of the molecules gives rise to peaks of the scattered intensity. The measured intensity distribution depends primarily on two factors, namely the structure factor reflecting a 2D translational order of the molecular centers in the plane of the monolayer and the form factor of the individual molecules.

Various scanning probe microscopy (SPM) techniques have become important tools in the study of

organic films in recent years. All the SPM techniques involve scanning of a very fine probe tip over the surface of the sample in a raster pattern, measuring the interaction, and collating the information into an image of the surface. Scanning tunneling microscopy (STM) has very high lateral resolution but is limited to monolayers on conducting substrates or to conducting films. Atomic force microscopy (AFM) does not have this limitation, and can provide direct information on the surface morphology of the sample. All of the SPM techniques yield structural and morphological information about only the top surface of a film. The nature of interactions between the probe and surface can be of varied nature ranging from tunneling current, van der Waals force, to other forces such as electrostatic, frictional, capillary, or magnetic.

Optical methods such as polarized fluorescence microscopy (PFM) and Brewster-angle microscopy (BAM) are generally used more than SPM techniques to study Langmuir monolayers. These techniques have identified various phase transitions in Langmuir monolayers as a function of temperature and surface pressure. Fourier transform infrared (FTIR) spectroscopy has become a very important tool for investigating the structure of LB films and Langmuir monolayers as it is now possible to obtain satisfactory signal-to-noise ratio. To examine LB films, two types of infrared measurement are often used, namely attenuated total reflection (ATR) and grazing incidence reflection (GIR). In the ATR technique, the LB film is deposited onto both sides of an IR-transmitting crystal, for example, germanium or silicon. The radiation enters the crystal at one end so that it is incident on the sides at an angle less than the critical angle. It thus undergoes multiple reflections inside the crystal before emerging at the far end. On each reflection, the evanescent field of the IR beam penetrates the LB film and absorption may take place. GIR is performed on a highly reflective metal substrate, and *p*-polarized IR light is normally used in these measurements. This method is particularly sensitive to the orientation of the molecular groups on the substrate. In addition to FTIR measurements, several other spectroscopic measurements, such as Raman spectroscopy, are also being carried out but these techniques are not being used frequently as it demands special sample preparation. Similarly, electron microscopy and diffraction techniques are also used infrequently to characterize organic films.

Monolayers and Thin Films

Formation of Langmuir Monolayers

The Langmuir monolayer of any organic molecule is generally formed by spreading a measured quantity of this material on water by using a spreading solvent. It is desirable that the solvent be capable of dispersing the molecules of the film-forming material at the air–water interface and then evaporate completely so that the film is not contaminated. It is essential to understand the structure of Langmuir monolayers to control the growth process of LB films. Moreover, the Langmuir monolayer is an ideal model system on a smooth liquid surface for investigating ordering in two-dimensions. Monolayers of lipids or amphiphilic molecules have been studied in detail to generate a complete phase diagram as a function of two thermodynamic variables, temperature and surface pressure. The surface pressure of these monolayer films is measured by calculating the difference in surface tensions between a clean or pure liquid surface and that covered by the film. The surface tension is determined by measuring accurately the downward force on a suspending plate, known as the Wilhelmy plate – made of sandblasted platinum or any other material that is wetted by the liquid. Surface tension has also been measured using diffuse scattering technique as a function of the length scale, and interesting effects of the molecular interaction on this thermodynamic variable have been observed.

Figure 1 indicates a basic experimental setup and generally obtained surface pressure versus area isotherms of fatty acids. The area of the monolayer and hence, the average area per molecule is varied by moving barriers across the water surface in a trough at various temperatures, and the resultant surface pressure is measured to generate these isotherms – the main source of thermodynamic data. An abrupt

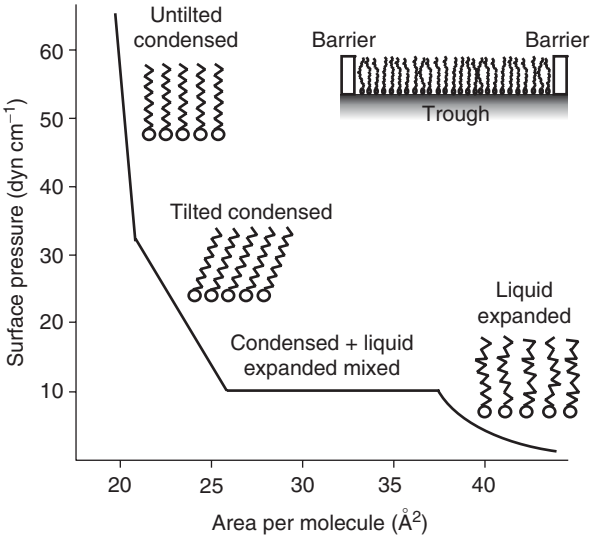


Figure 1 A typical molecular area vs. surface tension isotherm is shown with a sketch of a trough. Sketch of molecular arrangements in different phases are also indicated.

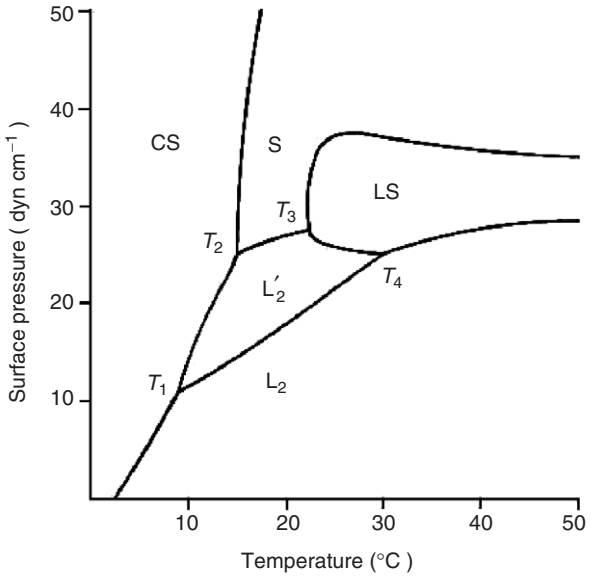


Figure 2 A typical phase diagram of monolayer of a fatty acid is shown as a function of temperature and pressure. Various phases are also marked as discussed in the text.

fracture of the monolayer occurs and 3D structures appear if the monolayer is compressed beyond a "collapse" pressure. One obtains different phases such as 2D gas (G), liquid-expanded (LE or L_1), and condensed phases with decreasing molecular area. Molecular configurations, structures of the various phases, and associated phase transitions have been identified by optical microscopy and GID techniques. The observed features have been explained in terms of the Landau theory. It should be mentioned here that these phases may not be in a thermodynamic equilibrium state, as implied by the term, but these are stable in an experimental timescale. **Figure 2** shows the phase diagram of a well-studied behenic acid as an example where T_1, T_2, T_3, and T_4 indicate phase-transition temperatures. L_2 and L_2' phases are both known as liquid-condensed (LC) and both have a centered rectangular lattice but the tails are tilted toward the nearest neighbor (NN) in the L_2 phase and in the L_2' phase, this tilt is toward the next nearest neighbor (NNN). At higher-pressure, one reaches solid (S), super-liquid (LS), or closed-packed solid (CS) phases depending upon the temperature and in all these phases, the tails become normal to the water surface without any tilt. One obtains a centered rectangular lattice in S and CS phases, but the LS phase retains a hexagonal lattice. It was observed that the temperature of a phase transition increases with the length of the hydrocarbon chain of the molecules. A significant alteration of the phase diagram of acid monolayers due to interactions between head groups has been observed as different metal ions in the subphase are added and the pH of the subphase is altered. Several other phases have been identified in Langmuir monolayers, for example, the Ov phase is a hexagonal lattice with an NNN tilt. However, the situation may become more complicated for molecules having multiple chains. For example, the biologically important class of amphiphiles and phospholipids should give rise to additional order due to the linkage between chains, but X-ray scattering techniques have not been able to detect these new phases – perhaps more sensitive measurement techniques are required.

Deposition of LB Films

To deposit an LB film, the precursor monolayer is prepared in a condensed phase and is maintained at a fixed temperature (refer **Figures 1** and **2**). During deposition, the surface pressure is held constant by moving the barriers to compensate the material transferred onto the substrate. It is well known that monolayers of salts of fatty acids are much easily transferred onto a solid support than the acids. Hence, most LB films are prepared from salts of the fatty acids. There are two different ways in which LB films of salts of long chain fatty acids, the prototypical amphiphilic molecules, may be deposited. The most common method is to spread the monolayer of the corresponding acid on a controlled subphase of the suitable inorganic salt kept at a definite pH (usually alkaline) in water. In the other method, the salt is actually prepared and spread on water at normal pH ($\sim$5.6). Although chemically, these two methods are expected to give identical results, the isotherms are found to be different for these methods. The substrates used for deposition of these films have to be treated specially for good transfer (they are made hydrophilic in most depositions). Normally, the transfer is quantified by a parameter called "transfer ratio" that is defined as the ratio of the area swept by the moving barriers of the trough during transfer to the area of film deposition on the substrate. Ideally, this should be close to unity. For hydrophilic substrates, the first stroke of deposition should be from below the water to air for efficient transfer of molecules from the Langmuir monolayer. If the substrate is made hydrophobic, then the first stroke should start from above the water. The substrate is deposited with a monomolecular layer during each subsequent up/down stroke through the Langmuir monolayer. However, it has been found by X-ray and neutron scattering techniques that substantial reordering of molecules takes place after the molecules are transferred on the substrate (**Figure 3**).

The scattering studies and AFM measurements have revealed that LB films have remarkably well-ordered vertical structure independent of the condition of deposition but in-plane morphology can change drastically depending upon substrate morphology and deposition condition. It was also observed that in-plane morphology remains unaltered at elevated temperatures. The X-ray reflectivity curve of a typical LB film and the extracted EDP by model fitting are shown in **Figure 4**. The interference (Kiessig) fringes observed in the reflectivity curve are due to the total film thickness. If Δq_z is the separation between two consecutive Kiessig fringes (refer **Figure 4**), the total film thickness (D) can be estimated as $D = 2\pi/\Delta q_z$. The separation of Bragg peaks in **Figure 4** gives a quick measure of the bilayer thickness. The extracted density profiles from a detail fitting of the reflectivity data (refer, for example, inset of **Figure 4**) provide us information about the molecular configuration. For example, from these studies on monolayer and trilayer films, one could conclude that molecules of divalent salts near a hydrophilic substrate remain in asymmetric configuration with a metal ion near the substrate, but in upper layers molecules take a natural symmetric configuration. The in-plane structure of LB films having single and multiple layers has been studied extensively. These structures have a strong similarity to structures of the condensed phases of Langmuir monolayers.

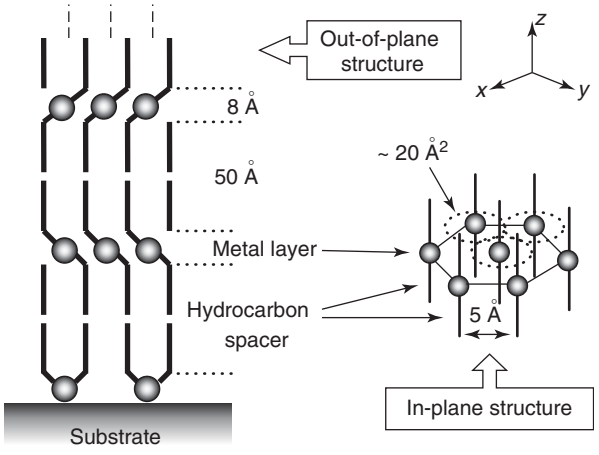


Figure 3 Model of out-of-plane and in-plane re-ordered structure of a typical Langmuir–Blodgett film of a divalent fatty acid salt is shown. The metal–metal distances in various directions are also indicated. The molecular configuration of the first monolayer near substrate is different from the upper layers as discussed in text.

Diffuse X-ray scattering studies have shown that the interfacial correlation of LB films can vary from self-affine fractal, observed in diverse physical systems, to long-range logarithmic, characteristic of capillary waves on liquid surfaces. Systematic studies of these interfacial morphologies can provide a clear basis for the theoretical and simulation studies undertaken to link the evolution of interfacial morphology with the possible growth mechanism of the LB films.

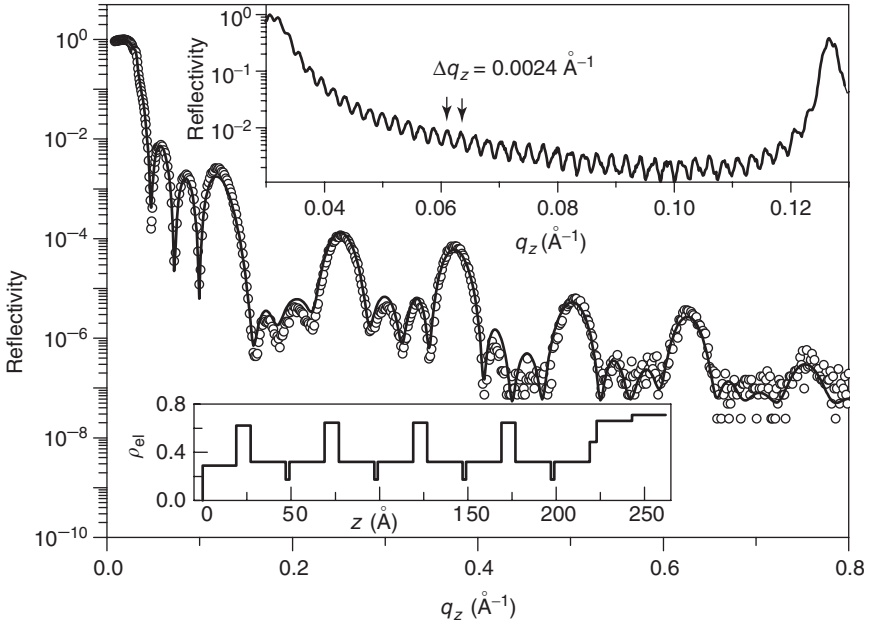


Figure 4 X-ray reflectivity profile of nine monolayer gadolinium stearate LB film (open circles) with calculated profile (solid line) for the electron density profile (shown in lower inset) corresponding to the model shown in **Figure 3**. The data of 101 monolayer film (upper inset) exhibit small separation of Kiessig fringes (marked by arrows) confirming the presence of 50 equally spaced bilayers in the LB film.

Growth of Self-Assembled Monolayers

Preparation of SAMs from a solution is a simple process – one has to only keep the substrate for a certain amount of time in a solution having the desired molecules. However, one should be very careful in cleaning the substrate surface and in keeping the solution free from contaminants. The defining feature of SAM formation is to identify the pair of the chemisorbing head-group of the molecule and the substrate. Although several systems have been found to form SAM, probably the most popular system is that of thiols on an Au(111) substrate. Organo-silicon monolayers on hydroxylated surfaces are another class of a well-studied system though these monolayers do not exhibit the same degree of long-range order observed for thiols on gold. The (111) surface of gold is the lowest energy surface and thus easy to form by evaporating gold on less-expensive substrates, such as glass, to make it more suitable for applications rather than using gold single crystals. Moreover, the relative inertness of the gold surface makes it comparatively easy to clean and SAMs of various thiol molecules form a well-defined order in a full-coverage phase. The full-coverage phase, by definition, corresponds to the highest possible packing of the molecules on a substrate surface. Early diffraction studies of the structure of alkanethiols on Au(111) indicated the existence of the full-coverage phase forming a $(\sqrt{3} \times \sqrt{3})$ structure, which was confirmed by STM. The tilt structure of the SAM was obtained directly from the analysis of the X-ray diffraction data and it was found that hydrocarbon chains are tilted by 32° from the surface normal. An IR spectroscopy study of this surface confirmed this finding. SAM can also form in a "lying down" condition on Au(111) and a lattice is obtained even in this configuration. In the "lying down" condition, molecules occupy large substrate space forming a low-coverage phase with a structure like $(23 \times \sqrt{3})$.

SAM of thiols and other molecules have also been prepared on Au(0 0 1) and other surfaces, and the obtained structure was found to be much more complicated. Even an Ag(1 1 1) surface, which is quite similar to the Au(1 1 1) surface, forms a much more complicated SAM structure of thiols including incommensurate structures. It is obvious that formation of a self-assembled lattice is a delicate process involving chain–chain interaction and headgroup–substrate interaction. Various models have been developed to understand this delicate growth process and the structural variation of SAM as a function of temperature. Another active research area in this field is to develop a heterostructure in SAM and the simplest approach is to use the end-group of SAM to adsorb other molecules. The adsorbed molecules may vary from "simple adsorbates" such as water to "complex adsorbates" such as protein or other bio-related molecules.

Specular reflectivity, diffuse scattering, and GID techniques are being used to characterize the SAM structures and the adsorbed molecules on these structures. Neutron scattering techniques have added advantage in characterizing the adsorbed layers as one can have better scattering density contrast here. Low-energy atom diffraction (LEAD) and low-energy electron diffraction (LEED) are other techniques being used for the structural characterization of SAM grown through both the solution and vapor routes. In addition to these diffraction techniques, SPM and various spectroscopy techniques are used for these films.

Applications of Organic Monolayers and Thin Films

Perhaps the first application of organic thin films was demonstrated by Franklin in calming water waves. In early 1930, coatings of amphiphilic monolayers were used to control the wetting properties of metal condenser plates in steam engines. Control of wetting, lubrication, and adhesion properties of various materials continue to be important application areas of organic films. Organic thin films are also being used to anchor capped metallic nanoclusters in the form of thin films to fabricate devices for nanotechnology. The application areas for these heterostructures range from nonlinear optics to soft-magnets. The idea of anchoring these adlayers has also been exploited to make chemically attached polymer films, which are very different from the spin-coated polymer films. Organic films are also used to link inorganic materials to biological materials for various biotechnological applications. A "dip-pen" writing technique has been developed that involves transfer of thiol molecules from an AFM tip to a gold substrate to form a patterned SAM. This technique is potentially useful for creating nanoscale devices.

Apart from all these technological applications, organic thin films are providing ideal model systems to investigate low-dimensional physics. LB films are suitable for studying 2D magnetic ordering with a fatty acid salt of a magnetic ion like gadolinium. Langmuir monolayers, on the other hand, have been used to study the interplay between hexatic ordering, chirality, and tilt in generating order in two-dimensions.

See also: Film Growth and Epitaxy: Methods; Liquid Crystals; Scanning Probe Microscopy.

PACS: 68.55. – a; 68.47.Pe; 68.37. – d; 75.70.Ak; 68.60. – p; 82.45.Mp; 61.12.Ex; 61.10.Eq; 87.64.Dz

Further Reading

Als-Nielsen J, Jacquemain D, Kjaer K, Leveiller F, Lahav M, *et al.* (1994) Principles and applications of grazing incidence X-ray and neutron scattering from ordered molecular monolayers at the air–water interface. *Physics Report* 246: 251.

Basu JK and Sanyal MK (2002) Ordering and growth of Langmuir–Blodgett films: X-ray scattering studies. *Physics Report* 363: 1.

Daillant J and Alba M (2000) High-resolution X-ray scattering measurements: I. Surfaces. *Reports on Progress in Physics* 63: 1725.

de Jeu WH, Ostrovksii BI, and Shalaginov AN (2003) Structure and phase transitions in Langmuir monolayers. *Review of Modern Physics* 75: 181.

Gompper G and Schick M (1994) *Self-Assembling Amphiphilic Systems, in Phase Transitions and Critical Phenomena*. New York: Academic Press.

Kaganer VM, Mohwald H, and Dutta P (1999) Structure and phase transitions in Langmuir monolayers. *Review of Modern Physics* 71: 779.

Mukhopadhyay MK, Sanyal MK, Mukadam MD, Yusuf SM, and Basu JK (2003) Field induced two-dimensional ferromagnetic ordering in a gadolinium stearate Langmuir-Blodgett film. *Physical Review* B68: 174427.

Penfold J (2001) Structure of the surface of pure liquids. *Reports on Progress in Physics* 64: 777.

Peng JB, Barnes GT, and Gentle IR (2001) The structures of Langmuir–Blodgett films of fatty acids and their salts. *Advances in Colloid and Interface Science* 91: 163.

Piner RD, Zhu J, Xu F, Hong S and Mirkin C (1999) Dip-Pen nanolithography. *Science* 283: 661.

Sanyal MK, Datta A, and Hazra S (2002) Morphology of nanostructured materials. *Pure and Applied Chemistry* 74: 1553.

Schreiber F (2000) Structure and growth of self-assembling monolayers. *Progress in Surface Science* 65: 151.

Schwartz DK (1997) Langmuir–Blodgett film structure. *Surface Science Report* 27: 241.

Sinha SK (1996) Surface structure reflectometry with X-rays. *Current Opinion on Solid State Material Science* 1: 645.

Smith RK, Lewis PA, and Weiss PS (2004) Patterning self-assembled monolayers. *Progress in Surface Science* 75 1.

Monocrystal *See* Crystal Growth, Bulk: Methods; Crystal Growth, Bulk: Theory and Models.

Mössbauer Spectroscopy

P Schaaf, Universität Göttingen, Göttingen, Germany

Introduction

The discovery of the recoilless nuclear resonance emission and absorption in solids is approaching its fiftieth anniversary. Although the resonance absorption of gamma rays by nuclei is analogous to the well-known optical phenomenon of resonance fluorescence and the underlying physics of the phonon theory of the lattice (Debye) and its quantum mechanical theory (Dirac, Lamb) had been well known and nuclear resonance absorption had already been predicted in 1929, it was only in 1957 that Rudolf L Mößbauer discovered this phenomenon during his doctoral studies. So it took a long time to combine nuclear physics with solid-state physics and to transfer absorption and emission resonances from the electron shell to the nucleus. This recoilless nuclear resonance absorption was named after R L Mößbauer, who was awarded the Nobel prize in 1961, and has been known as the Mössbauer effect since then. This effect is an easy way to eliminate the recoil energy loss that occurs during the emission and absorption of photons. Consequently, γ-lines of the highest possible sharpness could be obtained and applied. Soon after the discovery of the Mössbauer effect, it became obvious that this effect can be used to study various properties of materials on a microscopic scale via hyperfine interactions with an unprecedented resolution. This was the beginning of a new analytical tool – Mössbauer spectroscopy, which has now developed into a standard analytical technique in many laboratories.

The purpose of this article is to first give a short introduction to the basic physical processes and the hyperfine interactions, to sketch the typical experimental setup, and then to show how this can be applied as a unique analytical tool to many problems in solid-state physics. For more details and further reading, the reader is referred to the "Further reading" section, where the physical principles and applications of the Mössbauer effect and Mössbauer spectroscopy are described in great detail in a series of books. Nevertheless, regarding the most recent developments and applications of the method (e.g., the synchrotron-based methods), the current original literature should be referred to, because a recent review is still missing.

The Nucleus and Nuclear Radiation

When either an atom or a nucleus undergoes a transition from an excited state (E_e) to a lower or the ground state (E_g), a quantum of electromagnetic radiation (a photon) is emitted. This photon, in turn, can induce the opposite transition in an atom or nucleus of the same kind. This principle of resonant emission and absorption is sketched in **Figure 1**.

If a nucleus in an excited state with a mean lifetime of τ_N (e.g., $\tau_N = 141$ ns for ^{57}Fe) emits γ-radiation, the energy distribution $I(E)$ follows the Breit–Wigner law (Lorentzian):

$$I(E) = I_0 \times \frac{(\Gamma/2)^2}{(E - E_0)^2 + (\Gamma/2)^2} \qquad [1]$$

where E_0 is the transition energy ($E_0 = E_e - E_g$) and $\Gamma = \hbar/\tau_N$ is the resonance width (FWHM) ($E_0 = 14.401$ keV, $\Gamma = 4.7 \times 10^{-9}$ eV for ^{57}Fe). The momentum of the γ quantum is $\boldsymbol{p} = \hbar\boldsymbol{k}$.

Momentum conservation for the emission process causes the energy available for the emission process to be shared by the emitted quantum and the recoiling atom or nucleus. A free atom of mass M emitting a γ-quantum will experience a recoil with the energy $E_r = p_\gamma^2/2M$, which is generally much larger than the natural line width Γ ($E_r = 2 \times 10^{-3}$ eV is much higher than $\Gamma = 4.7 \times 10^{-9}$ eV for ^{57}Fe). As the same is true for the absorption process also, the absorption energy is increased by the recoil energy of the absorbing nucleus. Thus, the energy shifts by the recoiling nuclei are big enough to preclude resonance absorption of the emitted photon by other nuclei. This is sketched in **Figure 2** for the free atom. The shifts lead to only a very small overlap of the emission and absorption Lorentzians and thus a very low probability of resonance, illustrated in **Figure 2** by the shaded area.

If the nucleus is bound in a solid and the springs in **Figure 2** are strong enough, that is, the atoms are bound rigidly, the recoil goes to the whole lattice, and the recoiled mass M is drastically increased ($M \rightarrow 10^{16}M$). Thus, the recoil energy E_r becomes negligibly small. Since the lattice does not consist of rigid springs, in other words, the phonon system (the spring) is a quantized system, there is a certain probability that lattice vibrations (phonons) will be excited and the recoil of the atom is transferred to phonons. On the other hand, there is also a high probability that no phonons will be excited and the whole solid takes the recoil. Thus, in a solid for a certain fraction of nuclei, the emission, and in an equivalent manner the absorption, occurs without recoil energy loss. The chance that the emission takes place without inducing any phonons is called, in analogy to X-ray diffraction, the Debye–Waller factor, $f_D = \exp(-k^2 \langle x^2 \rangle)$, where $\boldsymbol{k}$ is the wave vector of the photon and $\langle x^2 \rangle$ the mean square displacement of the nuclei (atom). This factor – sometimes more accurately called Lamb–Mössbauer factor – depends on the temperature T, the energy of the γ-quantum $k = E_0/\hbar c$, and the Debye temperature Θ_D (representing the stiffness of the lattice springs) of the solid:

$$f_D(T) = \exp\left(-\frac{3\hbar^2 k^2}{4Mk_B\Theta_D}\left[1 + 4\left(\frac{T}{\Theta_D}\right)^2 \times \int_0^{\Theta_D/T} \frac{y}{\exp(y) - 1}\,dy\right]\right) \qquad [2]$$

The fraction of recoilless emission or absorption is $f_D = 0.76$ for ^{57}Fe in α-Fe with $\Theta_D = 470$ K at room temperature and thus allows fast measurement of the nuclear resonance, the Mössbauer effect. For the sake of completeness, it should be mentioned that the

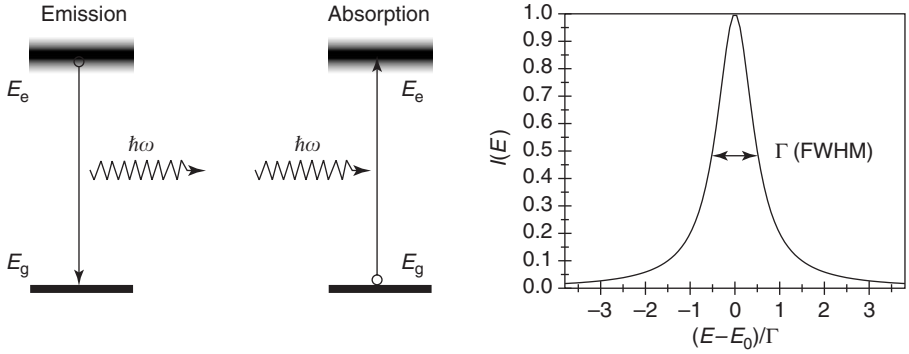


Figure 1 Schematic view of resonant emission and absorption. States with a finite lifetime are spread in energy according to the Heisenberg uncertainty principle, as indicated for the excited levels. This gives rise to a Lorentzian shape in the energy distribution of the corresponding emission (absorption) line.

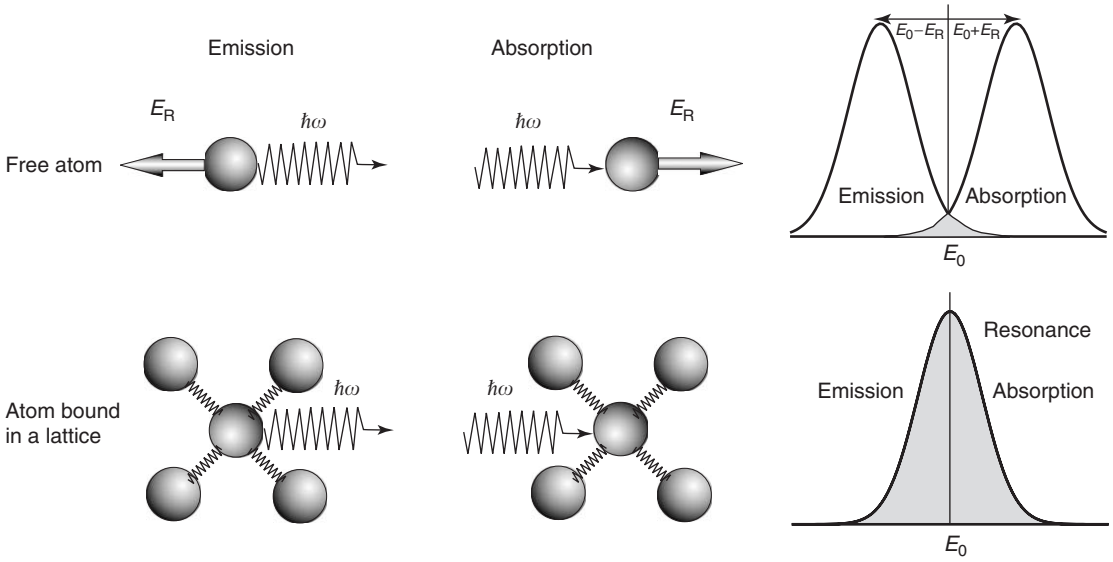


Figure 2 Principles of emission and absorption with and without recoil. In the upper part, a free atom involves recoil for emission and absorption. If the atom and thus the nucleus is bound in a lattice, the recoil is transferred to the whole lattice if the springs are tight enough. The shaded area represents the probability for resonance absorption.

Debye–Waller factors can be different in different crystallographic directions, causing the Goldanski–Karyagin effect, which leads to an anisotropy in the observed line intensities.

Another decisive factor for the feasibility and speed of the measurements is the cross section of the absorption. It determines the probability of a photon with matching energy to be absorbed by the corresponding nucleus. This cross section σ is given, together with the maximum cross section σ_0, as a Lorentzian by

$$\sigma = \sigma_0 \times \frac{(\Gamma/2)^2}{(E-E_0)^2 + (\Gamma/2)^2} \qquad [3]$$

The maximum cross section can be determined from the nuclear spin number I in the excited and ground state, the photon wavelength λ, and the internal conversion coefficient α via

$$\sigma_0 = \frac{\lambda^2}{2\pi} \frac{2I_e + 1}{2I_g + 1} \frac{1}{1+\alpha} \qquad [4]$$

The internal conversion coefficient α represents the ratio of the number of γ quanta interacting upon emission with the electron shell and thus leads to the emission of conversion electrons and conversion X-rays to the number of γ quanta with undisturbed emission, that is, the resulting probability of the γ emission is $1/(1+\alpha)$.

In order to obtain the γ quanta, one has to populate the excited state; in other words, one needs a source for the Mössbauer radiation. This is normally achieved easily via parent isotopes, which then decay to the corresponding nucleus in the excited state. Such a source is illustrated in **Figure 3**, where the decay scheme of the ^{57}Co source for ^{57}Fe Mössbauer spectroscopy is shown. The ^{57}Co isotope decays via electron capture to the excited state of ^{57}Fe. The half-life of ^{57}Co is $\sim$272 days so that with such a source, one can easily measure for a couple of years. The parent isotopes are normally produced by nuclear reactions (neutrons, deuterons, protons) in reactors or in accelerators and are commercially available. In addition, the radioactive isotopes are embedded into an appropriate matrix – normally chosen to be a cubic nonmagnetic material – in order to achieve a high Debye–Waller factor and a single line source.

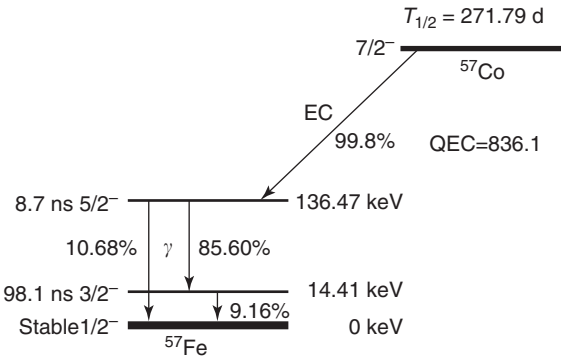


Figure 3 Schematic view of the ^{57}Co source for ^{57}Fe Mössbauer spectroscopy. The energy levels, spins, probabilities, and lifetimes are indicated.

Embedding in magnetic material can give multiline and polarized sources for certain applications. ^{57}Co is mostly embedded into the f.c.c. Rh metal as the source for Fe Mössbauer spectroscopy.

From the Effect to Spectroscopy

In the case of recoilless emission and absorption, one observes extremely sharp γ lines, the so-called Mössbauer lines. The width of these lines is often much smaller than the line shifts caused by interactions between the emitting or absorbing nuclei and their environment in the solid. For this reason, the Mössbauer effect is very suitable for resolving and measuring hyperfine interactions via a Doppler modulation of the γ energy by moving the source and the absorber relative to each other with velocity v:

$$E(v) = E_0 \times \left(1 + \frac{v}{c}\right) \qquad [5]$$

Usually, the source is moved and the velocities required for ^{57}Fe lie in the range of $10\,\mathrm{mm\,s^{-1}}$, which can easily be produced and controlled in the lab. Another Doppler effect is caused by the thermal vibrations in the lattice. The atoms and nuclei in the source and absorber have a finite temperature, have thermal energy, and are thus vibrating. As seen before, the movement causes a shift in energy, but the thermal velocities are isotropic and thus the positive and negative velocities cancel each other leading to only a small broadening of the spectral lines. Nevertheless, there are also relativistic effects, which are causing a net shift of the energies with the temperatures, which is called "second-order Doppler shift" (SOD):

$$\delta_{\mathrm{SOD}} = \frac{\bar{v}^2}{2c^2} E_0 \qquad [6]$$

The standard geometry of Mössbauer spectroscopy is the transmission geometry (transmission Mössbauer spectroscopy (TMS)), which is shown schematically in **Figure 4**.

The sample is irradiated from the source and the absorption is measured with a detector behind the sample. The source is moved by a drive system (constant acceleration, sinusoidal or constant velocity) and the emitted and Doppler-modulated photons move through an aperture toward a detector. Only those quanta are counted that have the involved transition energy E_0. Gas-filled proportional counters or solid-state detectors are most widely used for that purpose. For the measurement, the sample is placed between the source and detector, and it absorbs photons at certain energies represented by a certain source velocity. Of course, this absorption is again followed by emission, but it is not directed to the detector any more, but emitted more or less homogeneously into the full 4π solid angle (and additionally diminished by the internal conversion process). This way, one obtains a velocity-correlated intensity spectrum $I(v)$, which then contains all the

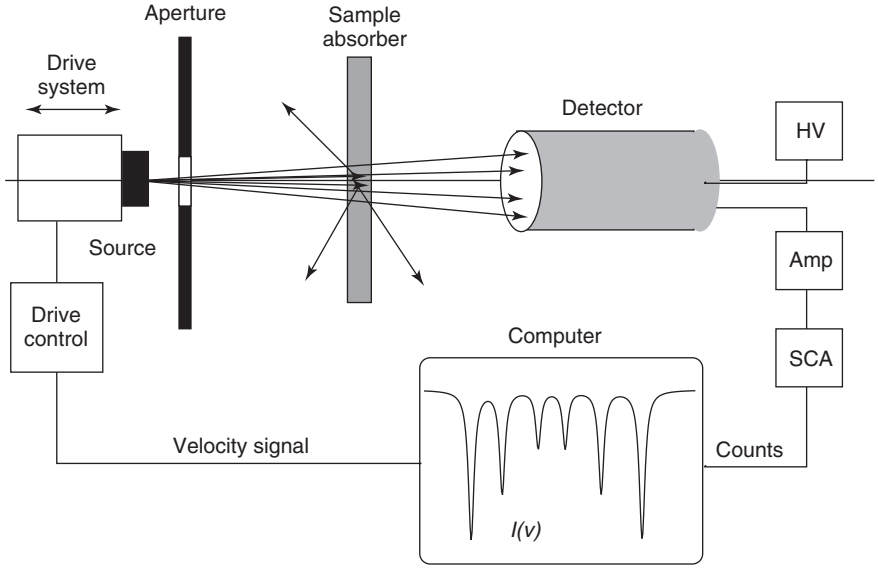


Figure 4 Schematic view of the setup for measuring a Mössbauer spectrum in transmission geometry. The source is moved in a controlled way and the photons are detected in a detector (HV – high voltage), whose signals are amplified (Amp). The single-channel analyzer (SCA) selects only those signals corresponding to the desired Mössbauer transition. These counts are stored in the computer for the respective drive velocities.

Table 1 Properties of the most frequently used Mössbauer isotopes

Isotope	a (%)	E_0 (keV)	τ_N (ns)	I_e	I_g	σ_0 (10^{-20} cm^2)	2Γ (mm s^{-1})	E_r (10^{-3} eV)
^{57}Fe	2.19	14.41300	98.1	3/2	1/2	256.6	0.1940	1.957
^{57}Fe	2.19	136.4743	8.7	5/2	1/2	4.300	0.2304	175.4
^{119}Sn	8.58	23.871	17.75	3/2	1/2	140.3	0.6456	2.571
^{151}Eu	47.82	21.64	9.7	7/2	5/2	11.42	1.303	1.665
^{125}Te	6.99	35.46	1.48	3/2	1/2	26.56	5.212	5.401
^{121}Sb	57.25	37.15	3.5	7/2	5/2	19.70	2.104	6.124
^{129}I		27.77	16.8	5/2	7/2	40.32	0.5863	3.210
^{197}Au	100	77.35	1.90	1/2	3/2	3.857	1.861	16.31
^{161}Dy	18.88	25.65	28.1	5/2	5/2	95.34	0.3795	2.194
^{161}Dy	18.88	43.84	920	7/2	5/2	28.29	0.006782	6.410
^{161}Dy	18.88	74.57	3.35	3/2	5/2	6.755	1.095	18.55

a – natural isotopic abundance, E_0 – transition energy, τ_N – half-life, I_e and I_g – nuclear spin quantum number of excited and ground states, σ_0 – maximum resonance cross section, Γ – natural line width (FWHM), E_r – recoil energy of a free atom.

information from the Mössbauer effect. These Mössbauer spectra are stored in a multichannel scaler. Sometimes, the source or the absorber (sample) is put into a cryostat, an oven, a magnetic field, etc., in order to measure temperature-dependent effects or the effects of certain physical situations. If the sample is too thick, then there is no count rate left behind it, and multiple absorption and emission processes can influence the spectra. If the sample is too thin, there are not enough nuclei for the absorption process, and not enough statistics for analysis is available (there is always nonresonant background). To be concise, one has to solve the full Hamiltonian (transmission integral) to find the absorption spectrum ($I(E) \sim I(v)$). Nevertheless, one can derive an analytical solution for the transmission integral (i.e., the folding of the emission, absorption, and re-emission over the sample), if the absorber can be treated as a thin absorber. The adimensional effective thickness of the absorber (sample) t_A is given by

$$t_A = \sigma_0 f_A N_A d_A a_A \qquad [7]$$

with the maximum cross section σ_0, the recoilless fraction (Debye–Waller factor) f_A, the atomic density N_A, the absorber thickness d_A, and the abundance of the Mössbauer isotope a_A. For ^{57}Fe at room temperature, the following values are obtained: $\sigma_0 = 256.6 \times 10^{-20}$ cm^2, $f_A \approx 0.75$, $N_A = 8.476 \times 10^{22}$ atoms cm^{-3}, and $a_A = 2.19\%$. The effective thickness should have values of $t_A \approx 1$, allowing fast measurements with the thin absorber approximation still valid. Then, there is the simple folding of the emission and the absorption spectra (with their respective energy eigenvalues E_i or v_i), resulting in a Lorentzian line again:

$$I(v) = I_0 - I_{abs} \frac{\Gamma^2}{(E - E_0)^2 + \Gamma^2} \qquad [8]$$

In this thin absorber approximation, this results in the double line width (more general by the sum of source and absorber line width: $\Gamma = \Gamma_S + \Gamma_A$). If the absorber becomes thicker, the lines become somewhat broader due to multiple re-emission and absorption effects, and for thick samples, one has to solve the full transmission integral.

A calibration of the Mössbauer spectrometer is carried out by measuring standard materials with well-known properties. For Fe, calibration is usually performed using a 10–25 μm α-Fe foil at room temperature (α-Fe: $B_{hf} = 33$ T or 10.6245 mm s^{-1}). All isomer shifts are normally related to the center of the α-Fe spectrum at room temperature. The spectra are fitted according to a least-squares routine by superimposing Lorentzian lines in simple cases or by more sophisticated theories for complicated samples.

If a sample consists of several sites or phases, the relative area of the corresponding set of Lorentzians gives the atomic fraction of the Mössbauer spy isotope in the site or phase. In order to calculate the real fraction, this area has to be corrected by the Debye–Waller factor. The latter is hardly known, so it has to be determined experimentally. As a simplification, it is often assumed that all subspectra and all phases have the same Debye–Waller factors f_D and neglecting (most of the times) smaller errors caused by that. Thus, the phase fraction can be identified with the relative fractions (areas) of the phases in the Mössbauer spectrum.

The isotopes most frequently used for Mössbauer spectroscopy are (listed according to their importance): ^{57}Fe, ^{119}Sn, ^{151}Eu, ^{121}Sb, ^{125}Te, ^{197}Au, ^{129}I, ^{155}Gd, ^{161}Dy, and some others. Their properties are summarized in **Table 1**.

Internal Conversion

The effect of internal conversion plays an important role in Mössbauer spectroscopy. It is most important

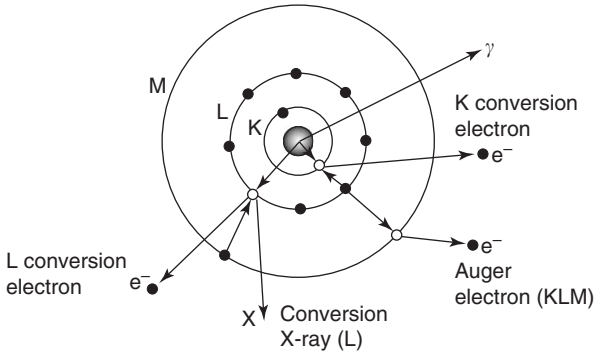


Figure 5 The internal conversion process during emission of γ-radiation. Instead of the photons conversion electrons can also be emitted leading to a cascade of Auger electrons and conversion X-rays.

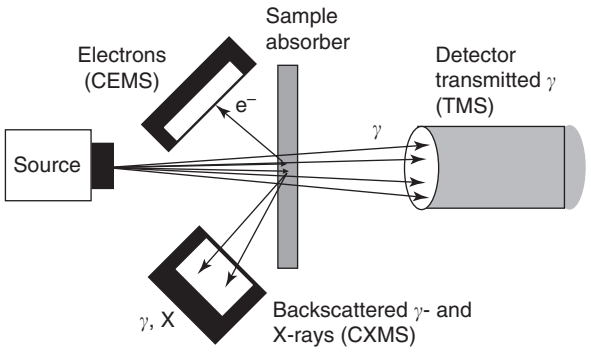


Figure 6 Mössbauer measurements in backscattering geometry leading to CEMS and CXMS. In combination with TMS, this results in STRMS.

for lower mass numbers and becomes negligible for very heavy atoms. For a certain fraction of nuclei, the emitted photon does not leave the atom, but spends its energy for kicking off an electron from the shell – a conversion electron. As a result of internal conversion, conversion electrons, and subsequently a cascade of Auger electrons and conversion X-rays are emitted. They too can be used for the measurements, since they contain the same information as the originally absorbed (emitted) photon. This internal conversion is shown in **Figure 5**.

If the electrons originating from the internal conversion process are detected for the Mössbauer measurement, this is called (conversion electron Mössbauer spectroscopy (CEMS)). Because the electrons cannot penetrate detector windows, the samples have to be mounted inside an appropriate detection system. Depending on the material and the electron energies, the electrons can only escape from the sample surface within 10–500 nm. Therefore, CEMS is especially suited for the investigation of surfaces and thin films (up to a depth of some 100 nm for Fe-Mössbauer spectroscopy) and is even able to measure the properties of a monolayer.

Also the conversion X-rays or re-emitted γ-radiation can be used to measure the Mössbauer effect in the samples under investigation. This method is named (conversion X-ray Mössbauer spectroscopy (CXMS)). The X-rays can penetrate $\sim$10–30 μm of most metals and thus the information depth of CXMS is in this order. For such backscattering measurements as visualized in **Figure 6** in reasonable times, special detectors with large solid angles have been developed.

Based on the different information depths of the different modifications of Mössbauer spectroscopy, it is especially interesting to measure these different radiations simultaneously. Simultaneous measurements achieve a discrete depth profile of the measured samples. The arrangement for the simultaneous measurement of all three modifications was named (simultaneous triple-radiation Mössbauer spectroscopy (STRMS)). It should be noted again that the sampling range in the case of iron-based materials for the CEMS measurements is $\sim$150 nm, while for CXMS, it is 10–20 μm and TMS scans the whole sample.

Hyperfine Interactions

Hyperfine interactions are based on effects caused by the interplay between the atomic nucleus and its electronic environment. As early as in 1924, Wolfgang Pauli interpreted the hyperfine structure of atomic spectral lines by the magnetic coupling between the nucleus and its electronic shell, and also predicted properties of the Zeeman and Paschen–Back splittings in the presence of a nuclear magnetic moment. So the hyperfine interactions were merely an extension of the fine structure and observed in the atomic spectra with increasing resolution of the experiments in atomic physics.

The high-energy resolution of the Mössbauer effect (as well as that of other nuclear methods such as nuclear magnetic resonance or perturbed angular correlation) allows the direct observation of the hyperfine interactions, that is, the change of the energy levels due to electric or magnetic fields acting at the Mössbauer nucleus. These interactions are mainly the electric monopole interaction, the electric quadrupole interaction, and the magnetic dipole interaction.

Isomer Shift

The electric monopole interaction is the interaction of the nuclear charge *Ze* with the electron density at the nucleus with finite size. It leads to a difference in

the energy states in the source (S) and absorber (A), when comparing to a point nucleus (infinite small):

$$E_S = E_0 + \tfrac{2}{5}\pi Ze^2|\Psi_S(0)|^2[R_e^2 - R_g^2] \qquad [9]$$

$$E_A = E_0 + \tfrac{2}{5}\pi Ze^2|\Psi_A(0)|^2[R_e^2 - R_g^2] \qquad [10]$$

where R is the radius of the nucleus in the appropriate state (g – ground state, e – excited state) and Ψ the electron wave function. $e|\Psi(0)|^2$ then gives the electron density at the nucleus. These shifts of the levels (equivalent to the isotope effect in atomic physics) lead to the isomer shift δ as sketched in **Figure 7**:

$$\begin{aligned}\delta &= E_A - E_S \\ &= \tfrac{2}{5}\pi Ze^2[|\Psi_A(0)|^2 - |\Psi_S(0)|^2][R_e^2 - R_g^2] \end{aligned} \qquad [11]$$

Thus, the lines for materials with different electron densities at the nucleus are shifted by δ with respect to the undisturbed center. Thus, charge densities in the materials can be calculated from this isomer shift.

Quadrupole Splitting

An electric quadrupole moment of the nucleus eQ, that is, a nonspherical shape of the nucleus, interacts with an electric field gradient (EFG) acting at the nucleus, which leads to a splitting of the energy levels, as displayed in **Figure 7**. Diagonalizing the EFG tensor $V_{ij} = \partial^2 V/\partial r_i \partial r_j$ (V = electrical potential) with $|V_{zz}| \geqslant |V_{yy}| \geqslant |V_{xx}|$ and $V_{xx} + V_{yy} + V_{zz} = 0$, the asymmetry parameter η is given by

$$\eta = \frac{V_{xx} - V_{yy}}{V_{zz}} \qquad [12]$$

and the energy eigenvalue for a state with spin I and magnetic quantum number m_I is

$$E_q(m_I) = \frac{eQV_{zz}}{4I(2I-1)}[3m_I^2 - I(I+1)]\sqrt{1 + \frac{1}{3}\eta^2} \qquad [13]$$

For ^{57}Fe with $I_g = 1/2$ and $I_e = 3/2$, only the excited state splits and the quadrupole splitting ΔE_q is

$$\begin{aligned}\Delta E_q &= E_q\left(m_I = \tfrac{3}{2}\right) - E_q\left(m_I = \tfrac{1}{2}\right) \\ &= \tfrac{1}{2}\, eQV_{zz}\sqrt{1 + \tfrac{1}{3}\eta^2}\end{aligned} \qquad [14]$$

This leads to a doublet in the Mössbauer spectrum, where the splitting, that is, the distance between the two lines, determines the magnitude of the quadrupole interaction and thus the field gradient. Consequently, charge distributions of atomic arrangements or defect structures can be deducted. Lattices with cubic symmetry and without any defects always result in single lines, because, due to the symmetry, no EFG can be present.

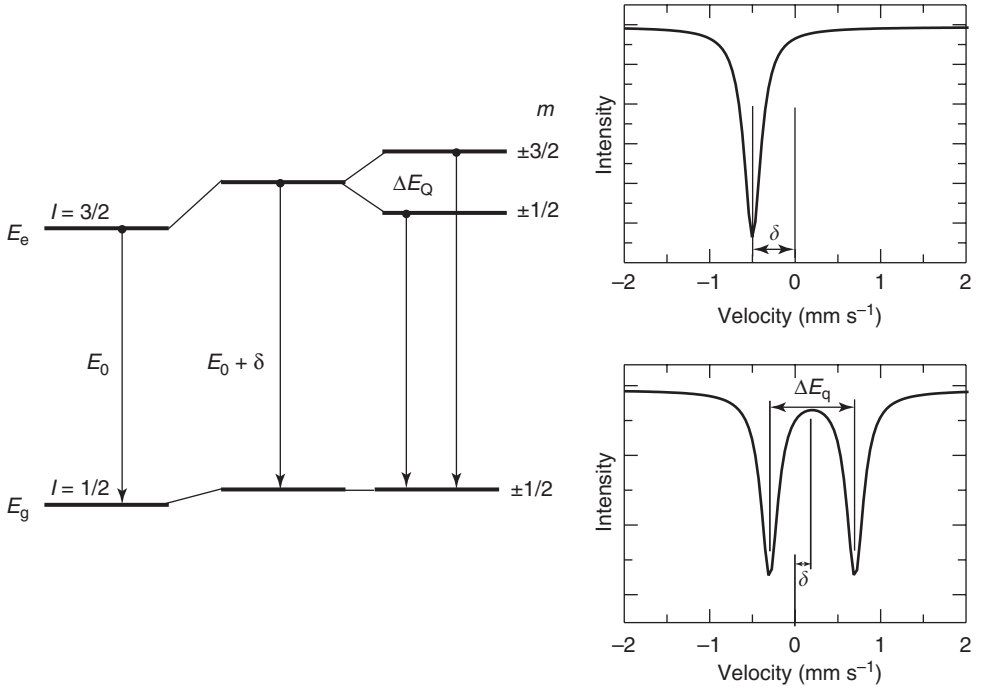


Figure 7 Electric monopole and quadrupole interactions cause a shift and a splitting of the nuclear levels, and accordingly shift and split the lines in the Mössbauer spectrum.

Magnetic Splitting

The magnetic moment μ of the nucleus interacts with a magnetic field B acting at the nucleus (nuclear Zeeman effect). This interaction gives rise to the splitting of the otherwise degenerated states. The state with spin I splits into $2I + 1$ magnetic substates with the eigenvalues:

$$E_m = -\frac{\mu B m_I}{I} = -g_N \mu_N B m_I \qquad [15]$$

where the magnetic quantum number m_I can have the values $m_I = I,\ I-1, \ldots, -(I-1),\ -I$. The magnetic moment μ is given by the Bohr magneton μ_N and the Landé factor g_N by

$$\mu = g_N \mu_N I \qquad [16]$$

Based on the selection rule for absorption (emission) of dipole radiation $\Delta m = 0, \pm 1$, only six transitions are allowed, and they represent the significant Zeeman sextet structure in the Mössbauer spectra of magnetic materials, as demonstrated in **Figure 8**.

Mixed Hyperfine Interaction

If the magnetic dipole interaction and electric quadrupole interaction act together, the eigenvalues cannot be determined as easily as for each interaction alone. Normally, perturbation methods are applied. As the magnetic interaction is usually much stronger, first the eigenvalues are determined for this magnetic interaction and the electric interaction is treated as a perturbation. For the energy states in a first-order perturbation calculation, one finds

$$E_m = -g_N \mu_N B m_I + (-1)^{|m_I|+1/2} \times \tfrac{1}{4} \times eQV_{zz} \times \tfrac{1}{2}(3\cos^2\theta - 1 + \eta \sin^2\theta \cos 2\phi) \qquad [17]$$

where θ and ϕ are the angles between the direction of the magnetic field B and the principal axis of the EFG tensor V_{zz}. For ^{57}Fe,

$$E_q = \tfrac{1}{8} eQV_{zz}(3\cos^2\theta - 1 + \eta \sin^2\theta \cos 2\phi) \qquad [18]$$

Now the line positions for the general case can be given by eqn [19]:

$$\begin{aligned} L_1 &= \delta - g_1 \mu_N B + E_q \\ L_2 &= \delta - g_2 \mu_N B - E_q \\ L_3 &= \delta - g_3 \mu_N B - E_q \\ L_4 &= \delta + g_3 \mu_N B - E_q \\ L_5 &= \delta + g_2 \mu_N B - E_q \\ L_6 &= \delta + g_1 \mu_N B + E_q \end{aligned} \qquad [19]$$

The g-factors g_i result from the Landé factors of the ground-state g_g and the excited state g_e:

$$\begin{aligned} g_1 &= \tfrac{1}{2} \times (3|g_e| + g_g) \\ g_2 &= \tfrac{1}{2} \times (|g_e| + g_g) \\ g_3 &= \tfrac{1}{2} \times (|g_e| - g_g) \end{aligned} \qquad [20]$$

The resulting shift of the energy levels and the shift of the inner four lines in the spectrum with respect to the outer lines 1 and 6 is sketched in **Figure 9**.

Angular Dependence of Absorption and Emission

The relative line intensities I_i for a sextet depend on the angle $\theta = \angle(\boldsymbol{k}, \boldsymbol{I})$ between the propagation direction $\boldsymbol{k}$ of the γ-radiation and the direction of the spin (magnetization) $\boldsymbol{I}$, as shown in **Figure 10**.

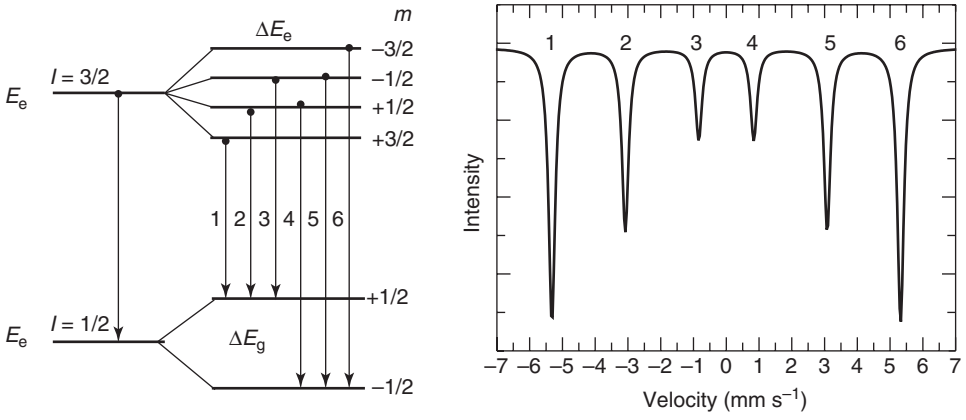


Figure 8 Magnetic dipole interaction causes a splitting of the nuclear levels, and accordingly leads to the typical Zeeman sextet in the Mössbauer spectrum.

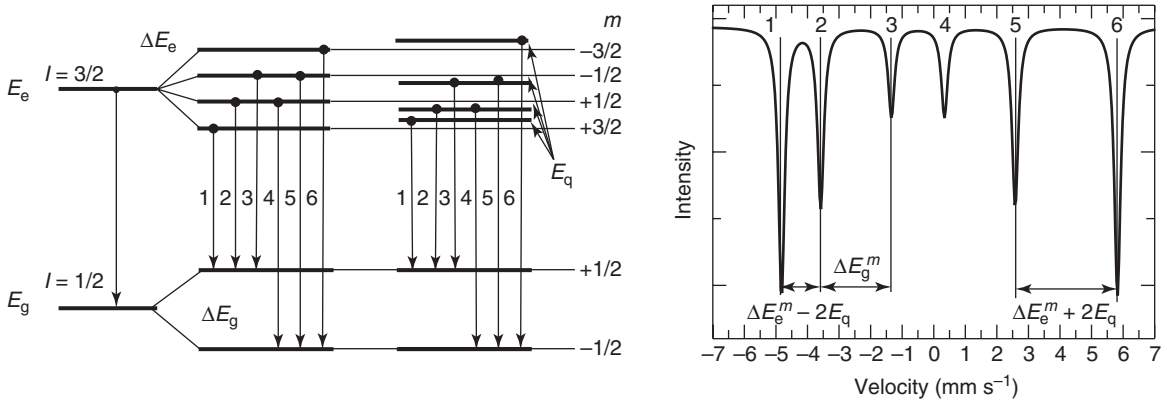


Figure 9 Mixed hyperfine interaction: magnetic dipole interaction causes a Zeeman splitting of the nuclear levels and an additional smaller electric quadrupole interaction shifts the split lines, which is also visible in the Mössbauer spectrum.

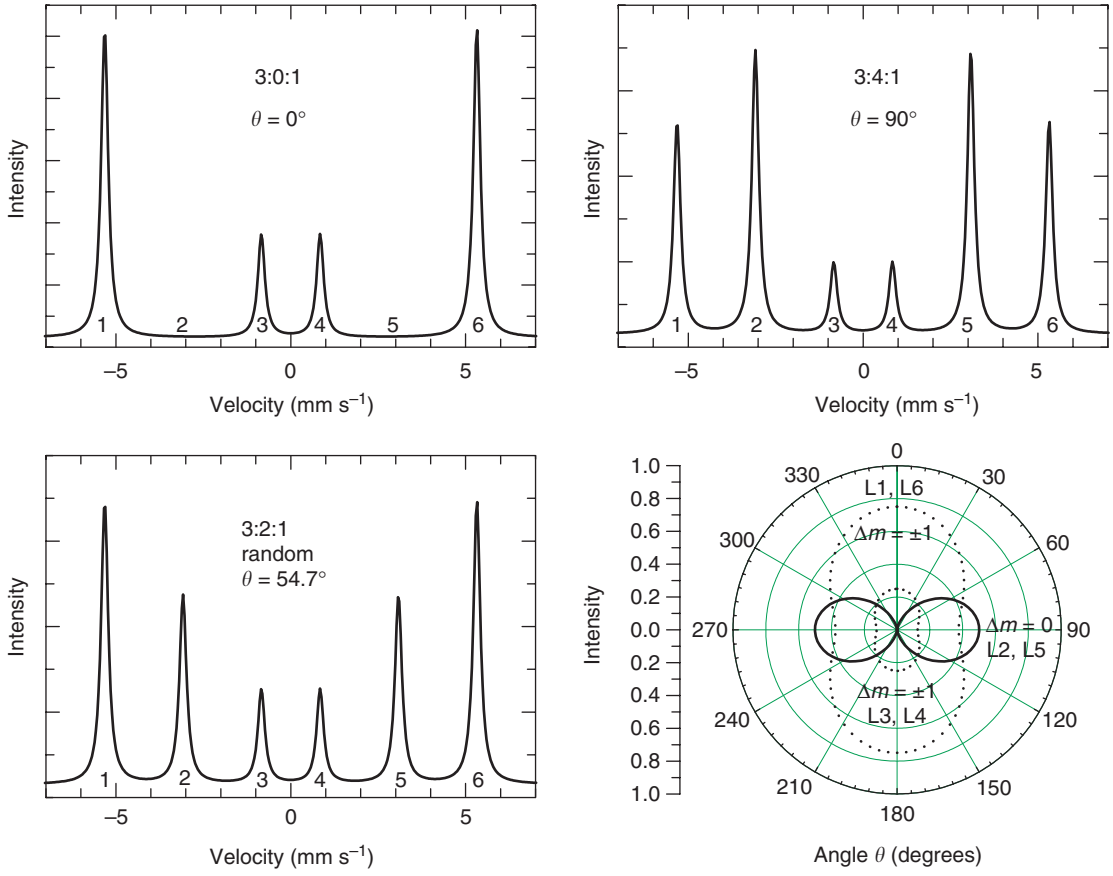


Figure 10 Angular characteristics of the dipole radiation: the angle θ is formed by the k vector of the γ radiation and the spin direction of the nucleus (i.e., the magnetization direction). The shown emission Mössbauer spectra illustrate some important cases.

For $\theta = 90°$, the intensity ratio is $I_1 : I_2 : I_3 = 3 : 4 : 1$, whereas for $\theta = 0°$, the lines 2 and 5 disappear and the ratio is $I_1 : I_2 : I_3 = 3 : 0 : 1$. If the directions of magnetization are randomly (isotropic) distributed ($\bar{\theta} = 54.7°$), the ratios are $I_1 : I_2 : I_3 = 3 : 2 : 1$.

This can be used to determine the magnetization orientation in the sample. Mössbauer polarimetry and magnetic orientation Mössbauer spectroscopy (MOMS) have evolved as specialized methods for this purpose.

Applications

One of the first applications of the Mössbauer effect was the proof of the gravitational red shift. Due to its extremely high-energy resolution, Mössbauer spectroscopy was able to verify the tiny energy shift of $\delta_{grav} = (gd/c^2)E_0$, which a γ-quantum experiences when it falls only by $d = 20$ m in the Earth's gravity g.

Phase Analysis

One of the standard applications of Mössbauer spectroscopy is the accurate phase analysis. One of the first examples was the determination of the amount of retained austenite in heat-treated steels. Here, Mössbauer spectroscopy is superior to any other method (e.g., XRD) with respect to its detection limit ($<1\%$) and accuracy ($<1\%$). In addition, by using Mössbauer spectroscopy the question of the interstitial arrangements of carbon and nitrogen in the retained austenite could be solved and it also allows an accurate calculation of the carbon or nitrogen content.

Another famous application of Mössbauer spectroscopy is the determination of Fe^{2+}/Fe^{3+} ratios in biomatter, minerals, and other oxides for medical, archeological, or industrial applications. Currently, the first Mössbauer spectrometer is doing this even on the surface of Mars.

The excellent resolution of the hyperfine interactions also allows the determination of the magnetic properties of materials, which is especially useful in alloys, and magnetic and amorphous materials. Due to the short lifetime of the nuclear states, all dynamic effects occurring in this time regime influence the Mössbauer spectra in shape and line width. This fact has been used to study diffusion and hopping mechanisms, relaxation, superparamagnetism, nanocrystalline materials, effects by RF-fields, and other time-dependent effects. More recent developments also deal with Mössbauer diffraction or the production of X-ray lasers.

Site Determination

The ability of Mössbauer spectroscopy for determining and quantifying different sites in a phase is demonstrated by a simple example, the iron nitride γ'-Fe_4N.

This material has an antiperovskite structure Pm3m, which is presented in **Figure 11**. The iron forms an f.c.c. structure and the nitrogen is located in the center of the unit cubes in an ordered manner. There are two crystallographic iron sites, Fe-I at the corners of the cube and Fe-II at the face centers, both sites having 12 iron nearest neighbors at the same distance. The Fe-I site has no nitrogen nearest neighbor (0 nn N), whereas the Fe-II sites have exactly two nitrogen nearest neighbors (2 nn N).

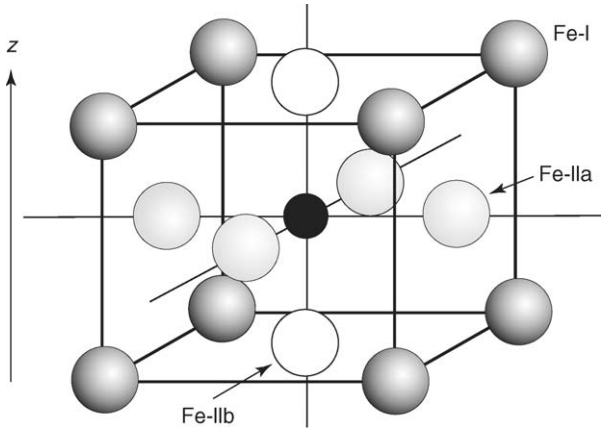


Figure 11 Crystallographic structure of the cubic γ'-Fe_4N with site assignment explained in the text.

In contrast to these two crystallographic sites, the Mössbauer spectrum of this material, as shown in **Figure 12**, has to be fitted with three subspectra corresponding to three different arrangements for the iron. The hyperfine fields are 34 T for the Fe-I sites and 21.6 T for the Fe-II sites. The difference in the two different Fe-II sites has to be explained by the magnetism of the material, which is invisible for X-ray diffraction.

The hyperfine parameters and the subspectral abundances are summarized in **Table 2**.

The f.c.c. material is ferromagnetic with a Curie temperature of $T_C = 767$ K, that is, well above the measurement temperature, which in this case is the room temperature. The magnetic easy axis is the (1 0 0) direction, which is indicated by the z-axis in **Figure 11**. There is no electric field gradient for the Fe-I site, but one for the Fe-II site having nitrogen neighbors. Thus, one has to distinguish two cases for the Fe-II site: for site Fe-IIa, the main component V_{zz} is collinear with the hyperfine field B_{hf}, that is, $\theta = 0°$. For the site Fe-IIb, EFG and z-axis are perpendicular, that is, $\theta = 90°$. For both cases, the asymmetry parameter is $\eta = 0$ and V_{zz} is identical. From eqn [18], it follows that the magnitude of E_q for the Fe-IIb site is twice as large as that for the Fe-IIa site but of opposite sign. This exactly matches the experimental findings in **Table 2**.

The subspectra areas should behave like Fe-I : Fe-IIa : Fe-IIb = 1 : 2 : 1, because 25% of the iron atoms are located at the corners, and Fe-IIa : Fe-IIb = 50% : 25% because four Fe-atoms at the face centers have $\theta = 0$ and only two have $\theta = 90°$. This theory is nicely reflected in the experimental results. This example demonstrates the ability and power of Mössbauer spectroscopy in site determination.

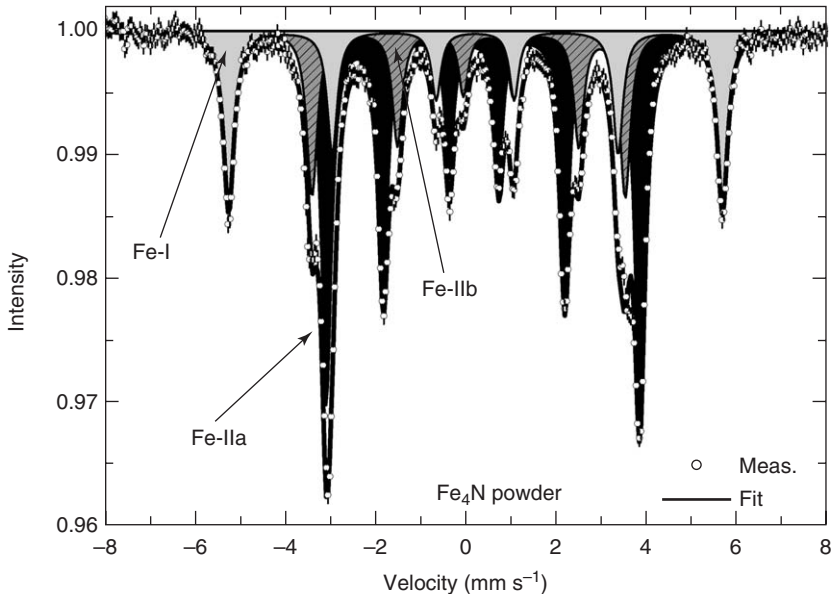


Figure 12 Mössbauer spectrum at room temperature in transmission geometry of Fe_4N powder. The spectrum is fitted by superimposing three Lorentzian sextets, corresponding to three different atomic Fe sites in the material as explained in the text.

Table 2 Hyperfine parameters as measured for Fe_4N powder at room temperature

Site	*IS* δ (mm s^{-1})	*QS* E_q (mm s^{-1})	*BHF* B_{hf} (T)	*LW* Γ (mm s^{-1})	*RA f* (%)
Fe-I	0.19(2)	0.00(1)	34.0(2)	0.28(1)	25.0(13)
Fe-IIa	0.27(2)	0.10(1)	21.6(2)	0.28(1)	51.9(09)
Fe-IIb	0.26(2)	− 0.22(1)	21.6(2)	0.28(1)	23.1(21)

IS – isomer shift, QS – quadrupole splitting, BHF – hyperfine field, LW – line width, RA – relative area.

Magnetic Orientation Mössbauer Spectroscopy

For Mössbauer polarimetry, magnetic fields and multiline Mössbauer sources (e.g., ^{57}Co in Fe gives six lines from the source and when measuring Fe with six absorption lines, this results in a spectrum of 36 lines) are used to polarize the quanta from the source and thus obtain information about the magnetic domain orientations in magnetic samples. On the other hand, with a proper modification of the traditional Mössbauer technique, with a method called magnetic orientation Mössbauer spectroscopy (MOMS), one is also able to determine the spin distribution in a sample. An important advantage of this MOMS technique in comparison to, for example, the magneto-optical Kerr effect technique (MOKE) is the absence of any external magnetic field during the measurements, which may change actual magnetization and spin distribution. Thus, it is also a magnetically nondestructive method.

In standard geometry, the photon beam is incident along the normal direction of the sample. For the MOMS experiments, the specimen normal is tilted by various angles α away from the incident γ-beam and then the spectra are measured subsequently after different sample rotations around their normal. This way, the spin distribution can be obtained by fitting the intensity ratio $I_2(\varphi)/I_3(\varphi)$ of the second and the third sextet lines as a function of the various tilt and rotation angles, yielding a "picture" of the spin distribution in the sample.

Synchrotron-Based Methods

The improvement of recent synchrotron facilities and the development of new channel-cut monochromators with superior resolution allowed the transformation from the classical Mössbauer spectroscopy in the energy domain into the time domain of nuclear resonance scattering (NRS). Many methods and applications have evolved here. Unfortunately, here this comprehensive subject cannot be treated.

Summary

In summary, it should have become obvious that Mössbauer spectroscopy is a very powerful method, which has proven its suitability for many problems in

solid-state and materials physics, chemistry, geology, archeology, environmental and space physics, and even relativistics. Mössbauer spectroscopy is based on nuclear physics, solid-state physics, and hyperfine interactions. It gives a microscopic image of the Mössbauer spy isotope and its immediate surroundings, which makes the method very powerful to study properties, short-range order and dynamic effects in many interesting materials. It is still under development and not used as widely as it should be. This may be due to the fact that "suppliers" (Mössbauer spectroscopists) and "customers" (scientists with questions) do not know enough of each other. Especially, the new developments of its synchrotron-based branch seem to lead to another push in its applications. Mössbauer spectroscopy should be seen as a powerful analytical tool, which can significantly contribute to answer many open questions for many materials. So ask your questions to a Mössbauer spectroscopist.

See also: Lattice Dynamics: Anharmonic Effects; Lattice Dynamics: Aperiodic Crystals; Lattice Dynamics: Structural Instability and Soft Modes; Lattice Dynamics: Vibrational Modes; Magnetic Interactions; NMR Methods for the Determination of Local Structure; Scattering, Nuclear Resonant.

PACS: 87.64.Pj; 76.80.+y; 33.45.+x; 61.18.Fs; 81.70.−q; 07.85.Fv; 31.30.Gs; 23.20.−g; 29.30.−h; 29.40.−n

Further Reading

Barb D (1980) *Grundlagen und Anwendungen der Mössbauerspektroskopie*. Berlin: Akademie-Verlag.

Frauenfelder H (1962) *The Mössbauer Effect*. New York: W.A. Benjamin Inc.

Gonser U (1975) Mössbauer spectroscopy, vol. 5 of *Topics in Applied Physics*. Berlin: Springer.

Gonser U (1981) Mössbauer spectroscopy II, vol. 25 of *Topics in Current Physics*. Berlin: Springer.

Greenwood NN and Gibb TC (1971) *Mössbauer Spectroscopy*. London: Chapman and Hall.

Long GJ and Grandjean F (1996) *Mössbauer Spectroscopy Applied to Magnetism and Materials Science*. New York: Plenum.

Long GJ and Stevens JG (1987) *Industrial Applications of the Mössbauer Effect*. New York: Plenum.

Mashlan M, Miglierini M, and Schaaf P (2003) Materials science in atomic scale by Mössbauer spectroscopy, vol. 94 of *NATO Science Series II: Mathematics, Physics and Chemistry*. Dordrecht: Kluwer Academic.

May L (1971) *An Introduction to Mössbauer Spectroscopy*. New York: Plenum.

Mössbauer RL (1958) Kernresonanzabsorption von Gammastrahlung in Ir^{191}. *Naturwissenschaften* 45: 538–539.

Mössbauer RL (1958) Kernresonanzfluoreszenz von Gammastrahlung Ir^{191}. *Zeitschreift Physika* 151: 124–143.

Murad E and Cashion J (2004) *Mössbauer Spectroscopy of Environmental Materials and Their Industrial Utilizations*. Dordrecht: Kluwer Academic.

Schatz G, Weidinger A, and Gardner JA (1996) *Nuclear Condensed Matter Physics (Nuclear Methods and Applications)*. Chichester: Wiley.

Stevens JG and Shenoy GK (1981) *Mössbauer Spectroscopy and Its Chemical Applications*. Advances in Chemistry Series 194, Washington: American Chemical Society.

Wertheim GK (1964) *Mössbauer Effect: principles and applications*. New York: Academic Press.

Nomenclature

a, a_A	isotopic abundance
a_A	atomic density ($1\,cm^{-3}$)
c, c_0	speed of light ($m\,s^{-1}$)
d_A	absorber thickness (m)
E_0	transition energy (eV)
E_n, E_g, E_e	energy levels of states, excited, ground state (eV)
f_D, f_A	Debye–Waller factor
h, $\hbar$	Planck constant (J s, eV s)
I	spin, nuclear spin I
I, I_0, $I(v)$, $I(E)$	intensity (counts)
M	mass, atomic mass
p	momentum ($\boldsymbol{p} = \hbar\boldsymbol{k}$)
t_A	effective thickness ($t_A = \sigma_0 f_A N_A d_A a_A$)
Z	atomic number
α	conversion coefficient
δ	isomer shift, shift (eV, $mm\,s^{-1}$)
γ	photon, radiation
Γ	line width (FWHM) (eV, $mm\,s^{-1}$)
λ	wavelength ($v \cdot \lambda = c$)
ν, ω	frequency ($\omega = 2\pi\nu$)
σ, σ_0	cross section, maximum cross section (barn)
τ_n	lifetime (s)
v	velocity ($m\,s^{-1}$)

Multicircle Diffractometry Methods

D A Walko, Argonne National Laboratory, Argonne, IL, USA

Introduction

When a monochromatic, collimated beam of radiation (X-rays or neutrons) is incident upon a stationary single crystal, the diffraction condition will probably be satisfied for few if any reflections, depending on a number of parameters such as the size of the unit cell, crystal mosaicity, and the energy spread of the radiation. It is even less likely that any particular reflection will be observed, and useful information, such as a reflection's integrated intensity, is not immediately available. Over the years, several methods of collecting scattering data have developed. The earliest X-ray experiments were performed with polychromatic radiation and a two-dimensional (2D) detector (photographic film); known as the Laue method, this technique has remained useful for quick sample alignment and has found new use in timing experiments in which an entire diffraction pattern is collected in a very short exposure. The powder-diffraction method consists of literally grinding the sample to a powder, effecting an ensemble average over all sample orientations; a powder-diffraction pattern is then collected with 1D scan of the detector. The rotation method consists of rotating a single-crystal sample through a given angle while measuring the pattern on a 2D detector with the use of modern image plates or charge-coupled device (CCD) detectors; this is the standard technique for macromolecular protein crystallography. While the above methods are practical for collecting integrated intensities of sharp, well-defined Bragg points, they may compromise on measurement of diffuse features. As an alternative, multicircle diffractometry uses computerized control of a high-precision goniometer (i.e., angle-positioning instrument) in order to hold both sample and detector at precise orientations, becoming a powerful tool to provide quantitative scattering data at well-defined positions in reciprocal space. Monochromatic radiation is used and, typically, observed by a point detector (a detector with a small, well-defined aperture).

Multicircle diffractometry is regularly applied to studies of systems in which orientational information cannot be sacrificed, such as measurement of pole figures for texture analysis or residual stress analysis; grazing-incidence or grazing-exit diffraction for studies of surfaces, thin films, and superlattices; and observation of diffuse scattering due to short-range order, defects such as twins or voids, or thermal diffuse scattering. Two examples of scattering data that require multicircle collection methods are shown in **Figure 1**. The reciprocal-space map in **Figure 1a** is from an Ni film grown on MgO(0 0 1); the peaks in the diffraction pattern are used to identify the various epitaxial orientations of the Ni crystal relative to the substrate, as well as the film's strain and mosaic spread. **Figure 1b** shows integrated intensities from two truncation rods of an epitaxial $SrTiO_3$ film on Si(0 0 1); the rods are continuous, diffuse features which can be analyzed to determine the atomic structure of the film and interface. The foremost drawback of multicircle diffractometer methods, compared to the other methods mentioned above, is that data are collected relatively slowly, often one point at a time rather than over a wide range of reciprocal space at once. The fairly involved calculations of the technique, described in some detail below, are no longer the issue that they once were, given the power and economy of modern computers; various diffractometer-control software packages are available.

In this article, several geometries of multicircle diffractometers, beginning with the standard four-circle Eulerian diffractometer, are reviewed. An overview of the transformation between reciprocal space and angle space is given, and various options in operating a diffractometer are described. Other diffractometer geometries will then be presented, including alternate four-circle geometries, then progressing to five- and six-circle geometries. Throughout the article, an attempt is made to point out potential sources of confusion, such as those caused by conflicting systems of nomenclature; correspondingly, the names of angles used here may differ from those in the literature or those implemented on various instruments. In that vein, it is to be noted that this article uses the term "multicircle diffractometers" to refer to instruments whose circles rotate the sample or detector; additional circles, which are used for polarization analysis or wavelength-dispersive analysis of the radiation, are not counted.

Four-Circle Eulerian Geometry

The basic four-circle diffractometer geometry is shown in **Figure 2**. The sample sits at the center of rotation on the ϕ circle, which rotates on the χ circle. The χ and ϕ circles together constitute a Eulerian

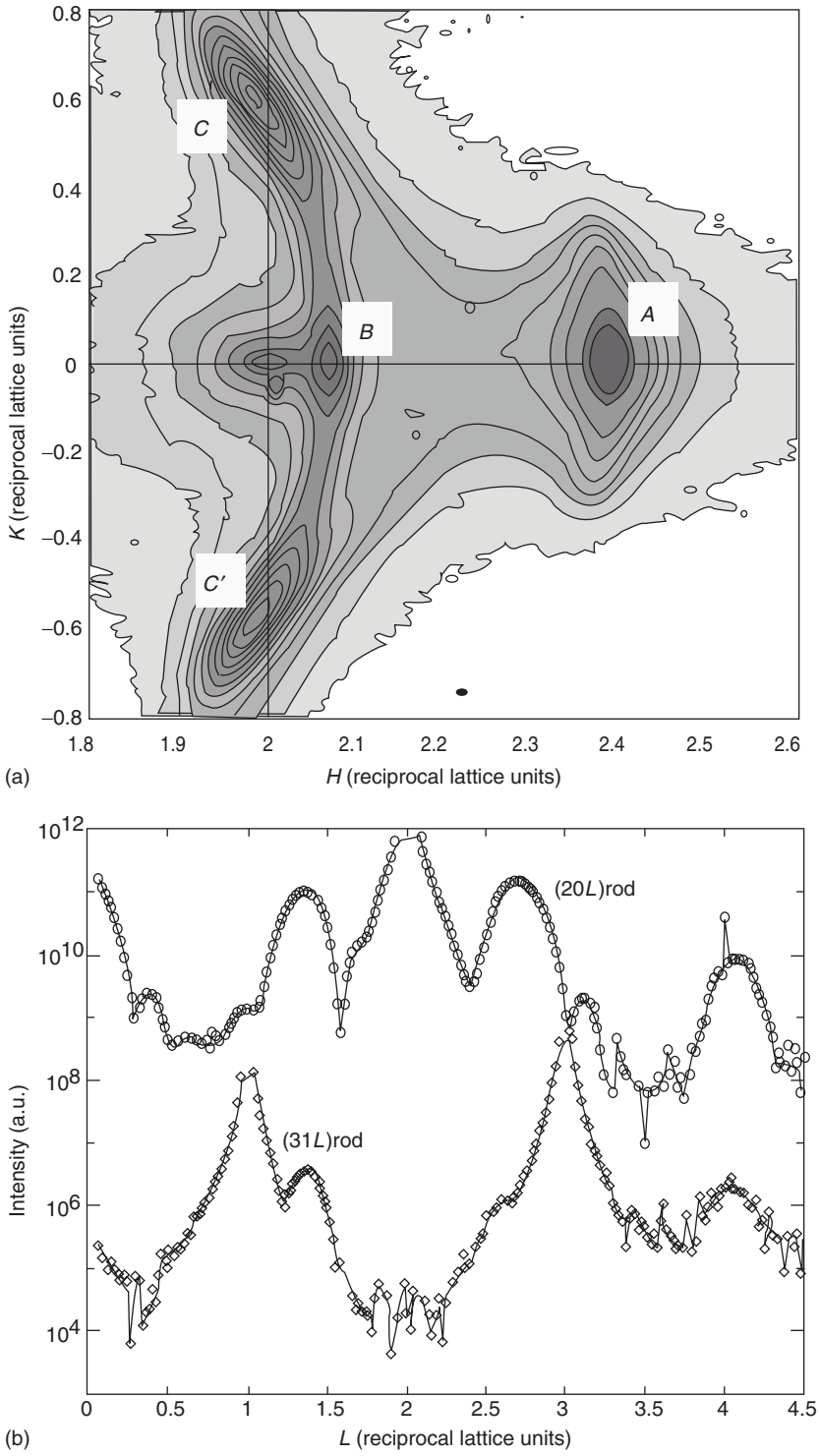


Figure 1 Two examples of thin-film diffraction data collected on a six-circle diffractometer. In both cases, the reciprocal lattice of the substrate material is used as a reference. (a) Reciprocal-space map of a 12 nm film of Ni grown at MgO(0 0 1). The map was made at $L = 0.05$ reciprocal lattice units, with the incident beam set at a grazing angle of 0.2°; the contours are spaced logarithmically. The sharp peak near the main MgO(0 0 2) peak at $H = 2$, $K = 0$ shows that the substrate is somewhat twinned. The peak marked *A* corresponds to an Ni(2 0 0) peak with relaxed cube-on-cube epitaxy. The peaks marked *B*, *C*, and *C′* are Ni(1 1 1) peaks due to regions with the epitaxial relationship (1 1 0)Ni ∥ (0 0 1)MgO. (b) Truncation-rod data of five unit cells of $SrTiO_3$ grown on Si(0 0 1), collected with a fixed 0.3° angle of incidence. Circles indicate integrated intensities along the (20*L*) rod, and diamonds represent the (31*L*) rod (vertically offset for clarity). Data points very close to the bulk Bragg points (2 0 2), (3 1 1), and (3 1 3) were not measured, since they have very high intensities and no surface sensitivity.

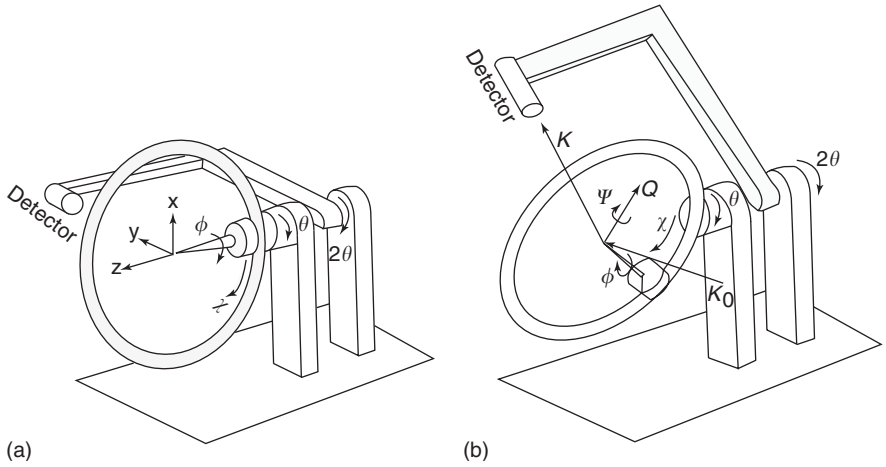


Figure 2 Diagram of a four-circle Eulerian diffractometer with (a) all angles at 0° and (b) all angles with positive displacements. The coordinate system is shown in (a), while incident wave vector $\boldsymbol{K}_0$, diffracted wave vector $\boldsymbol{K}$, scattering vector $\boldsymbol{Q}$, and azimuthal angle ψ are shown in (b).

cradle, for which this geometry is named. The Eulerian cradle then rotates on the θ circle. The detector rotates on the 2θ circle. The θ-axis is collinear with the 2θ axis; since both axes are perpendicular to the incident beam, this diffractometer is classified as normal-beam geometry. The scattering plane is defined as the plane that contains the diffracted wave vector $\boldsymbol{K}$ and the incident wave vector $\boldsymbol{K}_0$, as well as their difference, the scattering vector $\boldsymbol{Q}$, as shown in **Figure 2b**. The scattering plane is often horizontal for a laboratory-based X-ray diffractometer, so that the effects of gravity are minimized; at synchrotron sources the scattering plane is usually vertical, to take advantage of the (typically) *s*-polarization of the radiation and the higher degree of collimation in the vertical plane.

The correct operation of a diffractometer hinges on its being properly configured. **Figure 2a** shows the Eulerian four-circle diffractometer with all angles at their zero positions, in which case the direct beam travels along the axis of the χ circle, defining the zero position of θ; the detector intercepts the direct beam, defining the zero of 2θ; and the θ and ϕ axes are collinear, defining the zero of χ. The zero of ϕ is arbitrary. The sense of rotation, as shown in **Figure 2b**, is left-handed for all axes except χ.

The origin of the rather inelegant naming system is historical. In Bragg's original derivation of X-ray diffraction, the incident and exit X-ray beams both made an angle of θ with the diffracting planes, for a total scattering angle of 2θ. For normal-incidence diffractometers, the scattering angle is set by a single rotation typically labeled 2θ. The sample angle θ is not necessarily equal to one-half of 2θ, except in the special case of some three-circle diffractometers where the axes are mechanically coupled with a 1:2 gear ratio. Thus, although Bragg's law is often written as $\lambda = 2d \sin\theta$, it is better written as

$$\lambda = 2d \sin(2\theta/2) \qquad [1]$$

where λ is the wavelength and d is the lattice-plane spacing. The angle between θ and 2θ is ω, which is defined as

$$\omega = \theta - (2\theta)/2 \qquad [2]$$

The case of $\omega = 0$ is called the symmetric mode of diffraction. Unfortunately, there is no standard notation even for the most common type of diffractometer, and the θ circle is sometimes labeled ω.

Common angle scans performed on a four-circle Eulerian diffractometer include θ and $\theta/2\theta$ scans. A scan of the θ-axis probes the sample at a given magnitude of the scattering vector $|\boldsymbol{Q}|$ (defined below in eqn [9]). As an example, a θ scan might be used to probe a sample's mosaic spread; one possible source of instrumental broadening for this scan would be the beam divergence. In the $\theta/2\theta$ scan, the 2θ circle rotates at twice the rate of the θ circle. The trajectory of the $\theta/2\theta$ scan is radially outward from the origin of reciprocal space (i.e., $\boldsymbol{Q}$ increases in magnitude, with constant direction). This two-angle scan might be used to measure variations in a sample's lattice-parameter or strain; a source of instrumental broadening could be the radiation's band pass (i.e., the spread in wavelength). A $\theta/2\theta$ scan is preferable to a single-axis 2θ scan because the reciprocal-space trajectory of a $\theta/2\theta$ scan is simpler to interpret. However, the power of modern multicircle diffractometry is to move to arbitrary positions and scan in arbitrary directions of the reciprocal

space; the following sections of this article show how this is accomplished.

Angle Calculations

The Diffraction Equation

The fundamental problem in multicircle diffractometry is to map a sample's reciprocal-space coordinates to the diffractometer's angles: to measure the scattered intensity at a certain point in reciprocal space, what values should the angles have? The solution to this question was worked out by Busing and Levy for the case of a four-circle diffractometer, with extensions by many authors to other geometries. Here, the principles for the ideal four-circle case are outlined; detailed calculations for a variety of geometries can be found in the literature.

The question of mapping from reciprocal space to angle space could be restated as how to set the scattering vector $\boldsymbol{Q}$ equal to a desired reciprocal-lattice point $\boldsymbol{h}$, that is,

$$\boldsymbol{h} = \boldsymbol{Q} \qquad [3]$$

The two sides of eqn [3] must be written in the same reference frame and in the same coordinates; here one may choose to transform eqn [3] to Å^{-1} units in the laboratory frame. Beginning with the left-hand side of eqn [3], the transformation of reciprocal-lattice vector $\boldsymbol{h}$ from reciprocal-lattice coordinates ($\boldsymbol{h}$, with no subscript) to lab coordinates ($\boldsymbol{h}_\text{l}$) is achieved in a number of steps by successive application of a number of matrices. First, $\boldsymbol{h}$ in reciprocal-lattice units is transformed into Cartesian coordinates ($\boldsymbol{h}_\text{c}$, in Å^{-1} units) via the matrix $\boldsymbol{B}$:

$$\boldsymbol{h}_\text{c} = \boldsymbol{B}\boldsymbol{h} \qquad [4]$$

where $\boldsymbol{B}$ is constructed from the lattice parameters. The matrix $\boldsymbol{B}$ is most easily expressed as a combination of real-space lattice parameters a_i, α_i, and reciprocal-space lattice parameters b_i, β_i ($i = 1$–3):

$$\boldsymbol{B} = \begin{pmatrix} b_1 & b_2\cos\beta_3 & b_3\cos\beta_2 \\ 0 & b_2\sin\beta_3 & -b_3\cos\beta_2\cos\alpha_1 \\ 0 & 0 & 2\pi/a_3 \end{pmatrix} \qquad [5]$$

where $a_i \cdot b_j = 2\pi\delta_{ij}$. Generally, $\boldsymbol{B}$ is not orthonormal, but the other matrices in this section are all orthonormal rotation matrices. The next step is to transform $\boldsymbol{h}_\text{c}$ into the coordinate system of the diffractometer's ϕ axis:

$$\boldsymbol{h}_\phi = \boldsymbol{U}\boldsymbol{h}_\text{c} \qquad [6]$$

where $\boldsymbol{U}$ describes the orientation of the crystal relative to the diffractometer, and as such, is known as the orientation matrix. The determination of $\boldsymbol{U}$ is a critical step in multicircle diffractometery and is described in the next section. $\boldsymbol{h}_\phi$ is then transformed into laboratory coordinates by the successive application of the diffractometer-axes rotation matrices $\boldsymbol{R}$:

$$\boldsymbol{h}_\text{l} = \boldsymbol{R}(\theta)\boldsymbol{R}(\chi)\boldsymbol{R}(\phi)\boldsymbol{h}_\phi \qquad [7]$$

where

$$\boldsymbol{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\boldsymbol{R}(\chi) = \begin{pmatrix} \cos\chi & 0 & \sin\chi \\ 0 & 1 & 0 \\ -\sin\chi & 0 & \cos\chi \end{pmatrix} \qquad [8]$$
$$\boldsymbol{R}(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Turning to the right side of eqn [3], the scattering vector or momentum transfer, $\boldsymbol{Q}$, is defined as

$$\boldsymbol{Q} = \boldsymbol{K} - \boldsymbol{K}_0 \qquad [9]$$

where $\boldsymbol{K}$ and $\boldsymbol{K}_0$ are the diffracted and incident wave vectors, respectively. Their magnitudes are equal and defined as

$$|\boldsymbol{K}| = |\boldsymbol{K}_0| = K = 2\pi/\lambda \qquad [10]$$

The magnitude of $\boldsymbol{Q}$ is then

$$|\boldsymbol{Q}| = \frac{2\pi}{d} = \frac{4\pi\sin(2\theta/2)}{\lambda} \qquad [11]$$

It should be noted that numerous conventions exist for the definitions of eqns [9] and [10]. For example, the symbols $\boldsymbol{s}$ and $\boldsymbol{s}_0$ or $\boldsymbol{k}'$ and $\boldsymbol{k}$ may be used for the diffracted and incident wave vectors respectively. These wave vectors are often defined without the factor of 2π in eqn [10]; correspondingly, the factors of 2π are removed from eqns [5] and [11]. Other common symbols for the scattering vector include $\boldsymbol{G}$, $\boldsymbol{H}$, and $\boldsymbol{K}$.

In laboratory units, the incident wave vector is along the y-axis:

$$\boldsymbol{K}_0 = \begin{pmatrix} 0 \\ K \\ 0 \end{pmatrix} \qquad [12]$$

The diffracted wave vector is along the y-axis of the detector frame, which can be transformed into the lab frame with a rotation matrix for the 2θ axis:

$$\boldsymbol{K} = \boldsymbol{R}(2\theta)\begin{pmatrix} 0 \\ K \\ 0 \end{pmatrix},$$

$$\boldsymbol{R}(2\theta) = \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [13]$$

The scattering vector $\boldsymbol{Q}$, then, equals $[\boldsymbol{R}(2\theta) - \boldsymbol{I}]\,\boldsymbol{K}_0$, where $\boldsymbol{I}$ is the identity matrix. Substitution into eqn [3] results in the basic diffractometer equation for a four-circle diffractometer:

$$\boldsymbol{R}(\theta)\boldsymbol{R}(\chi)\boldsymbol{R}(\phi)\boldsymbol{UBh} = [\boldsymbol{R}(2\theta) - \boldsymbol{I}]\begin{pmatrix} 0 \\ K \\ 0 \end{pmatrix} \qquad [14]$$

Several possible generalizations of the above procedure are fairly straightforward to derive. For another diffractometer geometry, one simply incorporates the appropriate rotation matrices $\boldsymbol{R}$ into eqn [14]. One can also measure the directions of a diffractometer's rotation axes with, for example, an autocollimator, creating empirical $\boldsymbol{R}$ matrices rather than the idealized ones of eqns [8] and [13]. One also need not assume that the incident beam is perfectly collinear with the y-axis but can generalize the beam's direction (eqn [12]). This can be very useful in situations where the incident beam is not level (e.g., deflected upwards by a grazing-incidence mirror), but the entire diffractometer cannot be tilted. Several authors have worked out the equations for completely general diffractometer geometries.

Determining the Orientation Matrix

With the diffraction equation (eqn [14]) otherwise solved, the orientation matrix $\boldsymbol{U}$ must be determined; $\boldsymbol{B}$ must also be found if the sample's lattice parameters are not precisely known. Busing and Levy demonstrated several methods of determining $\boldsymbol{U}$ that result in a 1D set of solutions in diffractometer-angle space. To uniquely determine $\boldsymbol{U}$, a constraint must be added; selection of constraints are discussed in the following section.

If the sample's lattice parameters are known, then two noncollinear orientation reflections, $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, must be found to determine $\boldsymbol{U}$. Then, in principle, solving the following equations for $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ will determine $\boldsymbol{U}$:

$$\boldsymbol{h}_{i\phi} = \boldsymbol{UBh}_i \qquad [15]$$

where $i = 1, 2$ and the $\boldsymbol{h}_{i\phi}$ vectors are calculated from the observed angular positions of $\boldsymbol{h}_i$. But in practice, experimental uncertainties prevent one $\boldsymbol{U}$ matrix from exactly satisfying eqn [15] for both $\boldsymbol{h}_{1\phi}$ and $\boldsymbol{h}_{2\phi}$. Errors can arise from effects such as mosaic spread, refraction, beam divergence, or mechanical imperfections in the diffractometer. Therefore, the following procedure is commonly used to determine $\boldsymbol{U}$: $\boldsymbol{h}_{1\phi}$, called the primary reflection, exactly satisfies eqn [15], but $\boldsymbol{h}_{2\phi}$, as the secondary reflection, only determines the rotation of the coordinate system about $\boldsymbol{h}_{1\phi}$. Exchanging $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ generally results in slightly different solutions for $\boldsymbol{U}$. Generally, to minimize the effect of angular misalignment and achieve a more accurate solution of $\boldsymbol{U}$, the orientation reflections should be at fairly high angles (high $|\boldsymbol{Q}|$) and fairly far apart in angle (as a suggestion, the angle between $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ should be at least 40°). On the other hand, when a good alignment is not easily maintained, one could select orientation reflections that are close to each other and then probe only the nearby region of reciprocal space. Then, when another region is to be probed, a new set of orientation reflections can be chosen in the new region.

When the sample's lattice parameters are not known, the matrix $\boldsymbol{UB}$ must be determined; sometimes the $\boldsymbol{UB}$ matrix, rather than just $\boldsymbol{U}$, is known as the orientation matrix. In the unknown lattice-parameter case, three non-coplanar reflections ($\boldsymbol{h}_1$, $\boldsymbol{h}_2$, $\boldsymbol{h}_3$) with known (or assumed) Miller indices should be found that, with the selection of one constraint as discussed below, allows for a unique solution of $\boldsymbol{UB}$ in eqn [15]. With four or more reflections, a least-squares fitting method can be applied to eqn [15] to refine the lattice parameters. The resulting lattice parameters will not exactly reflect the symmetry of the crystal lattice but should be fairly close. For example, with a cubic crystal, the unit cell lengths should be nearly equal, and the angles should be within a few tenths of 90°. With a sufficiently large set of orientation reflections $\boldsymbol{h}_i$, still other parameters may also be refined, such as a displacement of the crystal from the diffractometer's center of rotation.

Operating Modes

As mentioned above, the calculation of diffractometer angles for a given reciprocal-lattice point is under-determined, since reciprocal space is three-dimensional while a four-circle diffractometer has, obviously, four degrees of freedom. An additional constraint is needed that is usually achieved by freezing one angle at a certain value. In the case of a four-circle normal-incidence diffractometer, the detector angle 2θ is determined by the magnitude of the

scattering vector $\boldsymbol{Q}$; a constraint cannot be placed on 2θ without restricting the range of accessible reciprocal space to a single magnitude of $|\boldsymbol{Q}|$. Therefore, constraints in the four-circle Eulerian geometry should be applied only to a diffractometer's sample angle or to an appropriately defined azimuthal angle.

The simplest and perhaps most common four-circle mode is to constrain ω (see eqn [2]) at a fixed value. The constraint of $\omega = 0$ is called the symmetric or bisecting mode, since the incident and diffracted beams make the same angle to the plane of the χ circle; generally, this mode permits the greatest access to reciprocal space. Another mode based on sample angles is to constrain χ or ϕ to a fixed value; such a mode might be useful when a sample environmental cell (e.g., oven, cryostat, or vacuum chamber) has a small window that must be kept oriented to the incident beam, but is otherwise fairly limiting.

Another possible operating mode is to choose a zone, a plane of reciprocal space passing through the origin, as the scattering plane. A zone can be defined by two non-collinear Bragg points; appropriate values of χ and ϕ are then selected to place the two Bragg points in the scattering plane. This mode contains two constraints, that is, it freezes two sample angles. Thus, access to reciprocal space is reduced to a single plane, but data collection is simplified since the zone can be studied with movements of only 2θ and θ.

Instead of constraining a particular sample angle to a given value, one can also constrain the azimuthal rotation of the sample about the scattering vector. As shown in **Figure 2b**, ψ is the azimuthal rotation about $\boldsymbol{Q}$. In defining ψ, a reference vector $\boldsymbol{N}$ must be chosen. $\boldsymbol{N}$ can point in any direction except parallel to $\boldsymbol{Q}$, but is often a special direction for the sample, for example, the surface normal, a magnetic field direction, or the direction of polarization in a polar material. The definition of $\psi = 0^\circ$ is when $\boldsymbol{N}$ lies in the scattering plane; $\psi = 90^\circ$ when $\boldsymbol{N}$ lies in the plane defined by $\boldsymbol{Q}$ and the diffractometer's $\theta/2\theta$ axes. The constrained-ψ mode is particularly useful in surface diffraction when $\boldsymbol{N}$ is chosen to be the surface normal. Mochrie has shown how to calculate ψ for a given incident angle or exit angle, which provide particularly convenient modes for grazing-incidence or grazing-exit studies. The constraint of $\psi = 90^\circ$ sets the incident angle equal to the exit angle.

Equivalent Angle Settings

In addition to selection of an operating mode, other choices of angle must be made in the operation of a diffractometer. Although any angle can theoretically be rotated by 360° without changing anything, the axes of most diffractometers have limited ranges of motion and cannot rotate indefinitely. The diffractometer-control software must determine, for example, whether a rotation from -170 to $+170^\circ$ will be done with a motion of -20 or $+340^\circ$. Furthermore, symmetry operations make some sets of angle combinations equivalent. For example, for a given χ and ϕ, the pair of angles $(2\theta, \omega)$ are equivalent to $(-2\theta, \omega - 180)$, but both settings may not be equally practical or accessible on a given diffractometer. In general, seven sets of angles can substitute for the set of $(2\theta, \theta, \chi, \phi)$, although not all of these sets are compatible with all operating modes. A diffractometer-control program must, implicitly or explicitly, select one of these sets. A convenient choice is to select angles that fall within the allowed range as set by software limits.

Alternative Four-Circle Geometries

Kappa Geometry

Several types of four-circle diffractometer geometries exist in addition to the Eulerian geometry described above. Perhaps the most common is the "kappa" geometry, originally developed by Enraf–Nonius. As shown in **Figure 3**, the χ circle of the Eulerian geometry is replaced by a κ-axis, which is built on an arm at an angle of α_0 to the θ_κ axis; typically, α_0 is nominally 50°. The κ-axis rotates another arm with another angle of α_0, upon which sits the ϕ_κ-axis. The "κ" subscripts on the θ_κ and ϕ_κ circles distinguish

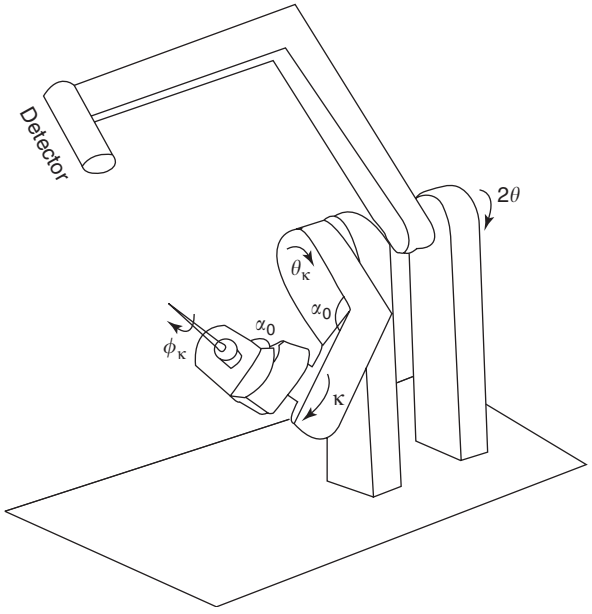


Figure 3 Diagram of a four-circle kappa diffractometer. The angles are chosen to hold a sample at the same place as the Eulerian diffractometer shown in **Figure 2b**, via the relations given eqn [16].

those circles of the kappa diffractometer from similar θ and ϕ circles of a standard Eulerian diffractometer, although not all systems of nomenclature make this explicit distinction. One could incorporate the rotation matrix of κ in deriving solutions to the diffraction equation (eqn [14]). However, it is often simpler to directly relate the kappa angles (θ_κ, κ, ϕ_κ) to the Eulerian angles (θ, χ, ϕ):

$$\begin{aligned}\theta &= \theta_\kappa - \cos(\kappa/2)/\cos(\chi/2)\\ \chi &= 2\arcsin[\sin(\kappa/2)\sin\alpha_0]\\ \phi &= \phi_\kappa - \cos(\kappa/2)/\cos(\chi/2)\end{aligned} \qquad [16]$$

The relations in eqn [16] are straightforward to be included in the diffractometer-control software after an appropriate solution to eqn [14] is computed in Eulerian geometry, with the advantage of allowing users to operate a kappa diffractometer just as the more familiar Eulerian diffractometer is. In doing so, (θ, χ, ϕ) are called "pseudoangles," whereas the physically present (θ_κ, κ, ϕ_κ) are called "real" angles.

The foremost advantage of a kappa-geometry diffractometer is the elimination of the χ circle. The primary weakness of Eulerian geometry is the limited mechanical perfection of the χ circle; any deviation from a perfect circle will, for example, increase the sphere of confusion. This issue is often exacerbated for diffractometers with an open or split χ circle. On the other hand, a closed χ circle blocks the incident or diffracted beam for some values of θ; beam blockage is less of an issue with kappa geometry. A related advantage is the relative compactness of the κ arm compared to the χ circle, making kappa geometry popular for confined spaces or crowded experimental stations. One hypothetical drawback of the kappa geometry is its limited angular range. As seen in eqn [16], a kappa diffractometer whose κ-axis has a full range of $\pm 180^\circ$ emulates a Eulerian diffractometer with a maximum theoretical range of $\chi = \pm 2\alpha_0$ (i.e., typically $\pm 100^\circ$). This issue can be dealt with by selection of a suitable angular setting, as discussed above, and should not generally pose a problem. A practical caution is that the asymmetric shape of the kappa arm, as it folds and unfolds, makes it particularly vulnerable to collisions with nearby objects (e.g., the input flight path or the detector arm); care should be taken to properly set hardware and software limit switches such that collisions cannot occur in any potential angular configuration.

Four-Circle General-Inclination Geometries

Another class of diffractometers, known as general-inclination geometries, is distinct from the normal-incidence geometries in that it provides more than one degree of freedom to the detector. Such geometries can significantly increase the range of accessible reciprocal space in situations where the angular range is limited by a sample cell. In contrast to the normal-incidence geometry, the scattering plane (containing $\boldsymbol{K}$, $\boldsymbol{K}_0$, and $\boldsymbol{Q}$) of the general-inclination geometry is no longer fixed; there is no simple one-to-one correspondence between the general-inclination and normal-incidence geometries such as that which exists between the Eulerian and kappa geometries. Since the position of the detector in a general-inclination geometry is a compound motion of several angles, the total scattering angle 2θ must be calculated, and cannot be simply read from one dial on the instrument. In this article, the convention that 2θ is the total scattering angle is followed. As such, only the normal-incidence geometry has a single circle that can be properly labeled 2θ, although many systems of nomenclature use 2θ as the label for one of the detector circles.

Several variations of four-circle inclined-geometry diffractometers have been developed, two of which are illustrated in **Figure 4**. For the "Z-axis" diffractometer (**Figure 4a**), the sample sits on the η-axis and the detector on the δ and γ-axes; all three axes then rotate about the α-axis. It may be noted that many Z-axis diffractometers call the sample axis θ (or ω); here, this axis is labeled as η to emphasize that it is not equivalent to the Eulerian θ-axis, since the η-axis is not orthogonal to the X-ray beam when $\alpha \neq 0$. The Z-axis geometry is particularly useful in holding heavy sample equipment such as vacuum chambers, as it minimizes motions that must be coupled through the sample chamber. The total scattering angle 2θ is given by

$$2\theta_{Z\text{-axis}} = \arccos(\cos\gamma\cos\delta\cos\alpha + \sin\alpha\sin\gamma) \qquad [17]$$

Another example of inclined geometry is the "S2D2" diffractometer. As shown in **Figure 4b**, the S2D2 geometry is composed of two pairs of orthogonal axes: the sample rotates on η and μ, while the detector rotates on δ and ν. Indeed, the name "S2D2" is meant to indicate that this geometry has two degrees of freedom for the sample ("S2") and two degrees of freedom for the detector ("D2"). In this nomenclature, the normal-incidence four-circle Eulerian and kappa diffractometers both have "S3D1" geometry. The Z-axis geometry does not exactly fit into this naming system, since one degree of freedom is shared between the sample and detector, but the designation $(\text{S1D2})^*1$ has been used, implying that a Z-axis diffractometer is an S1D2 (three-circle) diffractometer mounted upon a fourth circle. A practical advantage of the S2D2 geometry is

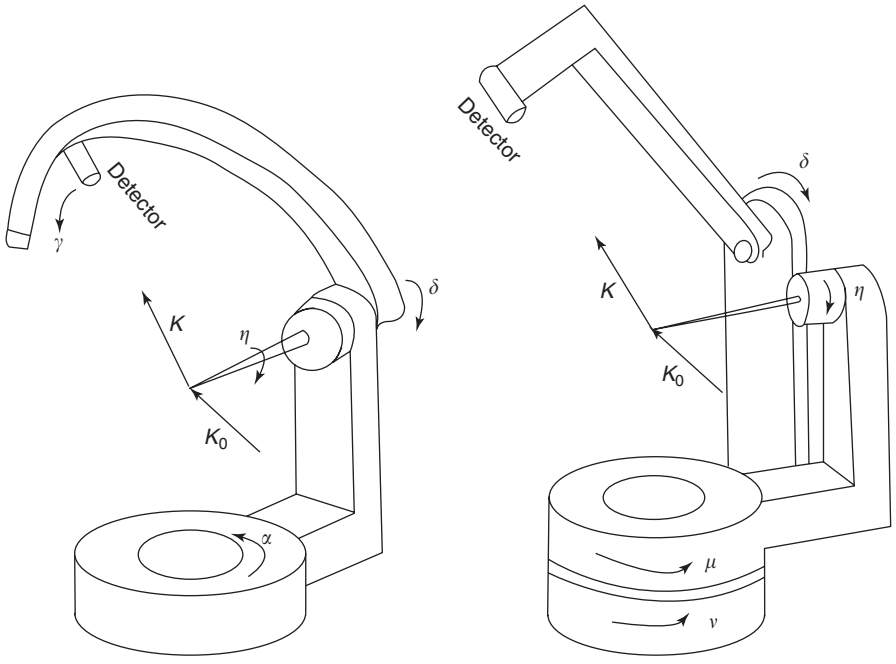


Figure 4 Two examples of general-inclination four-circle diffractometers: (a) *Z*-axis geometry; (b) S2D2 geometry.

the mechanical decoupling of the sample and detector motions, limiting the propagation of mechanical errors. The scattering angle of the S2D2 diffractometer is given by eqn [17] with γ set to zero:

$$2\theta_{S2D2} = \arccos(\cos \nu \cos \delta) \qquad [18]$$

The orientation matrix $\boldsymbol{U}$ of a four-circle inclined geometry can be calculated in the same way as for the four-circle normal-incidence geometry case, and a mode with one constraint must be selected. The diffraction equations are similar to the normal-incidence case, but the appropriate rotation matrices for the sample and detector axes must be incorporated on both sides of eqn [14].

The *Z*-axis and S2D2 geometries are particularly useful for surface scattering when the surface normal is aligned parallel to the η-axis, since the angle of incidence is given simply by the diffractometer angle α or μ, respectively. However, sufficiently precise sample alignment is often not easy, especially in a vacuum environment, which leads naturally to the desire for additional degrees of freedom for the sample. Such diffractometer geometries are described in the next section.

Five-Circle and Six-Circle Diffractometers

Representing the next level of complexity and flexibility, five- and six-circle diffractometer geometries all fall into the class of inclined geometries. In general, such diffractometers have been developed to provide better access to larger ranges of reciprocal space or to provide greater flexibility in sample positioning, given the constraints of a particular experiment.

Several varieties of five- and six-circle diffractometers have been constructed. The basic five-circle diffractometer could be characterized as a four-circle normal-incidence diffractometer (either Eulerian or kappa) with the θ-axis horizontal, mounted on a fifth rotation stage whose axis is vertical. The sample and detector motions are coupled by the fifth circle, so this five-circle geometry could be labeled $(S3D1)^*1$. This five-circle diffractometer can be converted into a six-circle instrument by the addition of another vertical axis, upon which sit the sample circles (θ, χ, and ϕ); the entire diffractometer rotates on the bottommost axis. Alternatively, one could add the χ and ϕ circles of a Eulerian cradle (or the κ and ϕ_κ circles of a kappa arm) to an S2D2 diffractometer (**Figure 4b**); this geometry decouples the sample and detector motions and would thus be labeled S4D2. Another six-circle geometry is formed by adding a Eulerian cradle to a *Z*-axis diffractometer (**Figure 4a**); the sample and detector motions of this geometry are coupled through the α-axis.

While the additional degrees of freedom in five- and six-circle diffractometers provide much greater flexibility in operation, the degree of complexity in

operation increases correspondingly. On a practical level, five- and six-circle diffractometers are generally quite large, with more (and heavier) counterweights than four-circle diffractometers. The large size and extra degrees of freedom increase the possibility of collisions and the damage that a collision can cause; hardware and software limits must be carefully set to avoid a collision in any potential angular configuration. More fundamentally, the choice of a mode becomes more complicated. Each additional degree of freedom (rotation axis) requires an additional constraint in the transformation between angle space and reciprocal space. Specifically, to uniquely determine the angle settings for a reciprocal lattice point, a five-circle or six-circle diffractometer requires, respectively, two or three constraints. The number of possible operating modes becomes correspondingly larger, and the geometry calculations supported by the diffractometer-control software also grow in complexity.

Although the modes of a five- or six-circle diffractometer require selection of several constraints, the constraints cannot be chosen arbitrarily. For example, one could not constrain two detector circles, since that would limit reciprocal-space access to a single magnitude of $|\boldsymbol{Q}|$. In the case of surface diffraction, one could not choose both the incident angle and the exit angle since that would overconstrain the azimuthal angle; one could choose to make the incident and exit angles equal but could not select a specific value for both. In the particular case of S4D2 geometry, a system was developed by You to organize mode selection by categorizing the angles which could be subject to constraints. The three categories of angles are detector, azimuthal, and sample. A detector constraint would fix one of the angles upon which the detector rotates, or would fix the angle of $\boldsymbol{Q}$ or reference vector $\boldsymbol{N}$ with respect to the horizontal plane; for example, $\boldsymbol{Q}$ or $\boldsymbol{N}$ could be fixed to lie in the horizontal or vertical plane. Azimuthal constraints include fixing the incident or exit angle, or otherwise fixing the azimuthal angle ψ. Possible sample constraints include fixing a sample angle at a certain value or constraining it relative to another angle. The key to this method is that no more than one constraint may be placed on either detector angles or azimuthal angles; the remaining constraint(s) must be applied to sample angles. Both sets of data in **Figure 1**, for example, were collected on an S4D2 diffractometer with one detector constraint (fixed angle of $\boldsymbol{N}$), one azimuthal constraint (fixed angle of incidence), and one sample constraint (η was fixed to be half of δ). One could also select one detector or azimuthal constraint and two sample constraints, or one could select three sample constraints.

The development of five- and six-circle diffractometers was driven by the need for flexibility in diffractometer operation, especially for access to the widest possible range of reciprocal space, given the experimental constraints. While some future advances in this technique may lie in the area of new diffractometer geometries, more critical needs lie in the area of improved data-collection efficiency. For example, many experiments may benefit from detectors with the sometimes-conflicting qualities of better resolution, larger dynamic ranges, or parallel data collection by use of 1D or 2D detectors. The overhead expense of motor motion can be reduced by the so-called trajectory or on-the-fly scanning: rather than collect data point by point, the diffractometer axes rotate continuously and the incoming data are binned into multichannel counters. In general, just as modern computers have made multicircle diffractometry viable as a regular research tool, increased automation will improve the efficiency in all steps of a diffraction experiment.

Acknowledgments

Data in **Figure 1** were collected at the Advanced Photon Source beamline 7ID, in collaboration with R Clarke and C Cionca (**Figure 1a**), and M J Bedzyk and D M Goodner (**Figure 1b**). Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under Contract No. W-31-109-ENG-38.

See also: Crystal Structure Determination; Crystal Structure Determination, History of; Neutrons and Neutron Scattering, History of.

PACS: 07.85.Jy; 07.85.Qe; 61.10.Nz; 61.12.Ld

Further Reading

Arndt UW and Willis BTM (1966) *Single Crystal Diffractometry*. Cambridge: Cambridge University Press.

Bloch JM (1985) Angle and index calculations for a '*z*-axis' x-ray diffractometer. *Journal of Applied Crystallography* 18: 33–36.

Busing WR and Levy HA (1967) Angle calculations for 3- and 4-circle x-ray and neutron diffractometers. *Acta Crystallographica* 22: 457–464.

Evans-Lutterodt KW and Tang M-T (1995) Angle calculations for a '2 + 2' surface x-ray diffractometer. *Journal of Applied Crystallography* 28: 318–326.

Hamilton WC (1989) Angle settings for four-circle diffractometers. In: Ibers JA and Hamilton WC (eds.) *International Tables for X-Ray Crystallography*, vol. 4, pp. 274–284. Dordrecht: Kluwer Academic.

Lohmeier M and Vlieg E (1993) Angle calculations for a six-circle x-ray diffractometer. *Journal of Applied Crystallography* 26: 706–716.

Mochrie SGJ (1988) Four-circle angle calculations for surface diffraction. *Journal of Applied Crystallography* 21: 1–3.
Sands DE (1982) *Vectors and Tensors in Crystallography*. Mineola: Dover Publications.
Vlieg E, Van der Veen JF, Macdonald JE, and Miller M (1987) Angle calculations for five-circle diffractometer used for surface x-ray diffraction. *Journal of Applied Crystallography* 20: 330–337.
You H (1999) Angle calculations for a '4S + 2D' six-circle diffractometer. *Journal of Applied Crystallography* 32: 614–623.

Nomenclature

a_i, α_i $(i = 1–3)$	real-space lattice parameters (Å,°)
b_i, β_i $(i = 1–3)$	reciprocal-space lattice parameters (Å,°)
$\boldsymbol{B}$	matrix which produces a Cartesian coordinate system for reciprocal space ($Å^{-1}$)
d	lattice-plane spacing (Å)
$\boldsymbol{h}$	reciprocal lattice vector (reciprocal lattice units)
$\boldsymbol{h}_c$	reciprocal lattice vector in Cartesian coordinates ($Å^{-1}$)
$\boldsymbol{h}_l$	reciprocal lattice vector in lab frame ($Å^{-1}$)
$\boldsymbol{h}_\phi$	reciprocal lattice vector in frame of ϕ axis ($Å^{-1}$)
$\boldsymbol{K}$	diffracted wave vector ($Å^{-1}$)
$\boldsymbol{K}_0$	incident wave vector ($Å^{-1}$)
$\boldsymbol{N}$	reference vector ($Å^{-1}$)
$\boldsymbol{Q}$	scattering vector ($Å^{-1}$)
$\boldsymbol{R}(\zeta)$	rotation matrix describing the effect of diffractometer axis ζ (unitless)
$\boldsymbol{U}$	orientation matrix (unitless)
α_0	kappa diffractometer angle (°)
γ, δ, η, α	Z-axis diffractometer angles (°)
δ, ν, η, μ	S2D2 diffractometer angles (°)
2θ	total scattering angle (°)
2θ, θ, χ, ϕ	four-circle Eulerian diffractometer angles (°)
2θ, θ_κ, κ, ϕ_κ	kappa diffractometer angles (°)
λ	wavelength (Å)
ψ	azimuthal angle (°)

Muon Spin Rotation

A Amato, Paul Scherrer Institut, Villigen, PSI, Switzerland
D Andreica, Babes-Bolyai University, Cluj-Napoca, Romania

Introduction

μSR spectroscopy uses implanted muons to probe properties of matter at the microscopic level. According to one of the earliest definitions of μSR – appeared on the cover of the first issue of the μSR Newsletter in 1974 – : "*μSR stands for Muon Spin Relaxation, Rotation, Resonance, Research or what have you*". The intention of the mnemonic acronym was to draw attention to the analogy with NMR and ESR, the range of whose applications were already well known at that time. More generally speaking, the abbreviation covers any study of the interactions of the magnetic moment of the muon with its surrounding when implanted into any kind of matter.

μSR is a relatively new nuclear method. Roughly speaking, it can be classified in between NMR and diffraction techniques. The NMR technique which is the closest parallel to μSR is "pulsed NMR," in which one observes time-dependent transverse nuclear polarization or a so-called "free induction decay" of the nuclear polarization. However, a key difference is the fact that in μSR, one uses a specifically implanted spin (the muon's) and does not rely on internal nuclear spins. In addition, and due to the specificity of the muon, the μSR technique does not require any radio-frequency technique to align the probing spin. On the other hand, a clear distinction between the μSR technique and those involving neutrons or X-rays is that scattering is not involved. Neutron diffraction techniques, for example, use the change in energy and/or momentum of a scattered neutron to deduce the sample properties. In contrast, the implanted muons are not diffracted but remain in a sample until they decay. Only a careful analysis of the decay product (i.e., a positron) provides information about the interaction between the implanted muon and its environment in the sample.

Similar to many of the other nuclear methods, μSR relies on discoveries and developments made in the field of particle physics. Following the discovery of the muon by Neddermeyer and Anderson in 1936, pioneer experiments on its properties were performed with cosmic rays. Indeed, with one muon hitting each square centimeter of the earth's surface every minute, the muons constitute the foremost constituent of cosmic rays arriving at ground level. However, μSR experiments require muon fluxes of the order of 10^4–10^5 muons per second and square centimeter. Such fluxes can only be obtained in high-energy accelerators which have been developed during the last 50 years.

Muon Production, Properties, and Implantation

Muon Production

The collision of an accelerated proton beam (typical energy 600 MeV) with the nuclei of a production target produces positive pions (π^+) via the possible reactions

$$\begin{aligned} p + p &\rightarrow p + n + \pi^+ \\ p + n &\rightarrow n + n + \pi^+ \end{aligned} \quad [1]$$

From the subsequent decay of the pions ($\tau_{\pi^+} = 26.03$ ns), positive muons (μ^+) are formed via the two-body decay

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad [2]$$

Since the neutrino has a helicity $h = -1$, both the neutrino and the μ^+ have their spin antiparallel to their momentum in the pion rest frame. According to the value of the pion momentum at the decay time, two types of μ^+-beams are available for μSR measurements.

The first type of muon beam is formed by the π^+ escaping the production target at high energies. They are collected over a certain solid angle by quadrupole magnets, and directed on to a decay section consisting of a long superconducting solenoid with a field of several tesla. If the pion momentum is not too high, a large fraction of the pions will have decayed before they reach the end of the solenoid. In the laboratory frame, the polarization of a high-energy muon beam is limited to $\sim$80%, and its energy is $\sim$40–50 MeV. Although such a high-energy beam requires the use of suitable moderators and samples with sufficient thickness, it guarantees a homogeneous implantation of the muons in the sample volume. Such beams are also used to study specimens inside recipients, for example, samples inside pressure cells.

The second type of muon beam is often called the "surface" or "Arizona" beam (recalling the pioneer works of Pfifer *et al.* from the University of Arizona). Here muons are used that arise from π^+ decaying at rest still inside, but near the surface, of the production target. Such muons, which are 100% polarized, ideally monochromatic, and have a very low momentum of 29.8 MeV c^{-1}, which corresponds to a kinetic energy of 4.1 MeV, have a range width in matter $\sim$180 mg cm^{-2}. Hence, the paramount advantage of this type of beam is the possibility of using relatively thin samples.

Finally, muon beams of even lower energy ("ultra-slow muons" with energy down to the eV–keV range) can be obtained by further reducing the energy of an Arizona beam using moderators, as a thin layer of a van der Waals gas frozen on a substrate. The tunable energy range of such muon beams corresponds to implantation depths in solids of less than a nanometer up to several hundred nanometers. Therefore, the study of magnetic properties as a function of the distance from the surface of the sample is possible.

In addition to the above-mentioned classification based on energy, muon beams are also divided according to the time structure of the particle accelerator, that is, continuous or pulsed.

For continuous muon beams, no dominating time structure is present. By selecting an appropriate muon incoming rate, muons are implanted into the sample one by one. The main advantage is that the time resolution is solely determined by the detector construction and the read-out electronics. There are two main limitations for this type of beam: (1) a non-negligible random background signal and (2) a rather limited muon incoming rate. Only the background problem can be overcome by using electrostatic deflectors to ensure that no muons enter the sample until the decay of the precedent muon has been observed.

For pulsed muon beams, protons hitting the production target are bunched into pulses which are also reflected on the secondary muon beam. The advantages of a pulsed muon beam are that, in principle, one can use the entire incoming muon intensity and there is almost no background due to accidental coincidences between the incoming muons and decay positrons. The absence of background allows the extension of the time window for measurements up to about $10 \times \tau_\mu$. The reverse of the medal is that the width of the muon pulse limits the time resolution.

μSR-Relevant Muon Properties

The main properties of the muon relevant to μSR are listed below:

- Charge: $+e$
  Since condensed matter physics or chemistry studies are carried out almost exclusively with positive muons, the following description will be limited to these muons (studies involving negative muons are primarily of interest to nuclear physicists).
- Magnetic moment: $3.183345118(89)\mu_p = 8.89059698(23)\mu_N$
- Gyromagnetic ratio: $\gamma_\mu = 2\pi \times 13.553882$ $(\pm 0.2$ ppm)kHz/gauss
  This large magnetic moment makes the muon an accurate magnetometer which can be implanted directly into a sample.
- Spin: 1/2
  The muon is free of quadrupolar interaction. Compared to NMR experiments, this simplifies

the analysis of the μSR spectra, with a reduced set of adjustable parameters.

- Mass: $m_\mu = 206.7682838(54) \times m_e = 0.112\ 609\ 526\ 9(29) \times m_p$.
  The muon can be considered as a light isotope of the proton.
- The decay of the muon into an electron and two neutrinos occurs via the weak interaction process after an average lifetime of $\tau_\mu = 2.19703(4)\ \mu$s:

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \qquad [3]$$

The parity violation in the weak interaction leads to an anisotropic distribution of the positron emission with respect to the spin direction of the μ^+ at the decay time. The positron emission probability is given by

$$W(\theta)\mathrm{d}\theta \propto (1 + \mathscr{A} \cos\theta)\mathrm{d}\theta \qquad [4]$$

where θ is the angle between the positron trajectory and the μ^+-spin, and $\mathscr{A}$ is an asymmetry parameter which depends on the polarization of the muon beam and positron energy. This anisotropic emission constitutes, in fact, the basics for the μSR technique.

The asymmetry of W is given by $\mathscr{A} = aP_\mu(0)$, where $P_\mu(0) = |\boldsymbol{P}_\mu(0)|$ is the beam polarization (~ 1), and a is an intrinsic asymmetry parameter determined by the weak decay mechanism. Theoretically, an average of $\langle a \rangle = 1/3$ is obtained if all emitted positrons are detected with the same efficiency irrespective of their energy. Practically, values of $\mathscr{A} \simeq 0.25$ are routinely obtained (**Figure 1**).

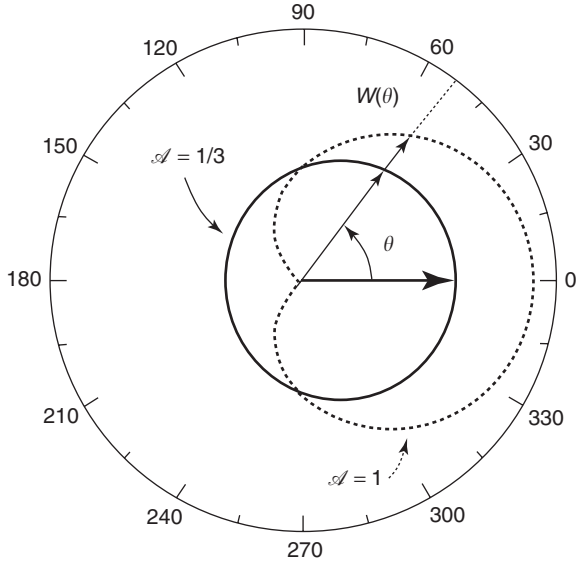


Figure 1 Polar diagram of the angular distribution $W(\theta)$ of positrons from muon decay: for the maximum positron energy (dotted line) and integrated over all positron energies.

Muon Implantation

The muons are implanted into the sample of interest where they lose energy very quickly. Fortunately, this deceleration process occurs in such a way that it does not jeopardize a μSR measurement. On one side it is very fast (~ 100 ps), which is much shorter than a typical μSR time window (up to 20 μs), and on the other side, all the processes involved during the deceleration are Coulombic (ionization of atoms, electron scattering, electron capture) in origin and do not interact with the muon spin, so that the muon is thermalized without any significant loss of polarization.

The muons invariably adopt interstitial sites of the crystallographic lattice. In metallic samples, the muon's positive charge is collectively screened by a cloud of conduction electrons. Thus, in metals, the muon is in a so-called diamagnetic state and behaves like a free muon. In insulators or semiconductors, a collective screening cannot take place and the muon will usually pick up one electron and form a so-called muonium ($\mathrm{Mu} = \mu^+ e^-$), which has size (Bohr radius), reduced mass, and ionization potential similar to the hydrogen atom.

The Technique

The Fundamentals

A schematic diagram of a μSR experiment is shown in **Figure 2**. Incoming muons trigger a clock that defines a time t_0. The trigger is defined either by the signal produced by the muon crossing a thin plastic detector located in front of the sample (continuous beam), or by the well-defined arrival time of a muon pulse (pulsed beams). If the implanted μ^+ are subject to magnetic interactions, their polarization becomes time dependent, $\boldsymbol{P}_\mu(t)$. Since the decay of the muon is asymmetric, the polarization of the muon ensemble can be deduced by measuring the positron distribution as a function of time in different detectors located around the sample. In the time-differential μSR technique, repeated measurements ($\sim 10^7$) are made of the time interval between the μ^+ implantation into the sample and the detection of the emitted positron in a particular direction, say, in the direction of the initial polarization $\boldsymbol{P}_\mu(0)$; the extension of the following discussion for other arbitrary directions of detection is straightforward.

The time histogram of the collected time intervals has the form

$$N_{e^+}(t) = B + N_0 \exp\left(-\frac{t}{\tau_\mu}\right) \times \left[1 + \mathscr{A}\,\frac{P_\mu(t)}{P_\mu(0)}\right] \qquad [5]$$

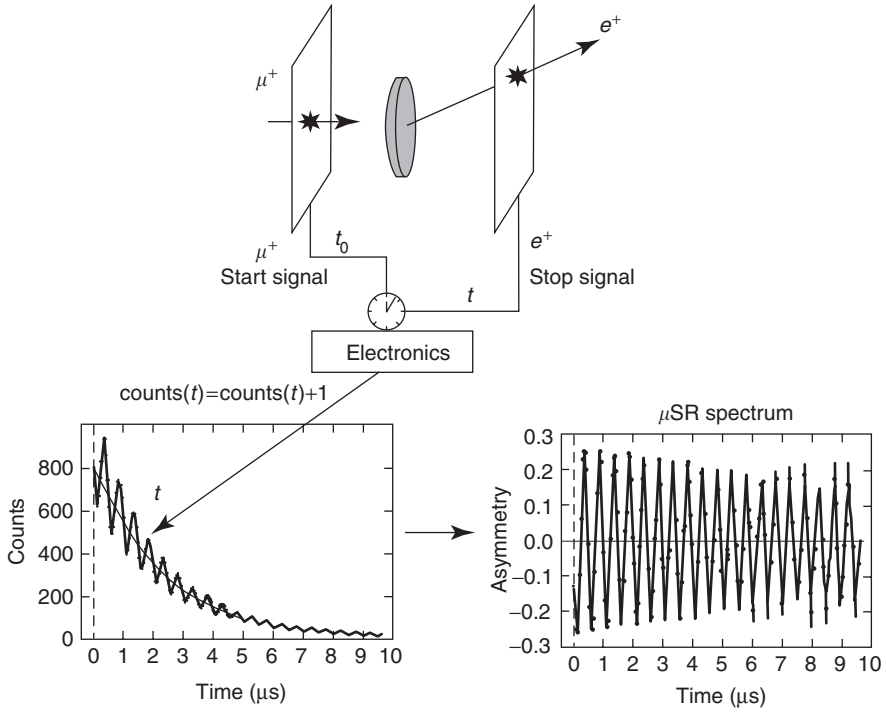


Figure 2 Schematic illustration of a μSR experiment. The arrival time t_0 of the muon is either given by a special detector (continuous beam, see figure) or by the arrival of a muon pulse (pulsed beams, not shown). The time between the muon arrival and the subsequent decay recorded by a positron emission at time t is used to build a rate vs. time histogram. By removing the exponential decay due to the muon lifetime and a possible background, one obtains the μSR signal reflecting the time dependence of the muon polarization.

where B is a time-independent background, N_0 is a normalization constant, and the exponential accounts for the decay of the μ^+. $P_\mu(t)$ is defined as the projection of $\boldsymbol{P}_\mu(t)$ along the direction of the initial polarization, that is, $P_\mu(t) = \boldsymbol{P}_\mu(t)$. $\boldsymbol{P}_\mu(0)/P_\mu(0) = G(t)P_\mu(0)$, where $G(t)$ reflects the normalized μ^+-spin auto-correlation function

$$G(t) = \frac{\langle \boldsymbol{S}(t) \cdot \boldsymbol{S}(0) \rangle}{S(0)^2} \qquad [6]$$

which depends on the average value, distribution, and time evolution of the internal fields, and therefore contains all the physics of the magnetic interactions of the μ^+ inside the sample. Equation [5] can therefore be written as

$$N_{e^+}(t) = B + N_0 \exp\left(-\frac{t}{\tau_\mu}\right)[1 + \mathscr{A}G(t)] \qquad [7]$$

where $\mathscr{A}G(t)$ is often called the μSR signal and the envelope of $G(t)$ is known as the μ^+-depolarization function. Since N_0 and B do not contain physical information about the sample investigated, usually only the μSR signal $\mathscr{A}G(t)$ is reported for a μSR measurement.

Precession and Relaxation

A first example is taken to understand the ability of the muon to study internal magnetic fields. The presence of a local internal field $\boldsymbol{B}_\mu$ (e.g., created by a magnetic state) at the muon stopping site is to be assumed. If the magnetic field is at an angle θ to the initial muon polarization, the muon will precess around a cone about the magnetic field (see **Figure 3**). The time evolution of the μ^+-polarization is therefore given by

$$G(t) = \cos^2\theta + \sin^2\theta \cos(\gamma_\mu B_\mu t) \qquad [8]$$

where $B_\mu = |\boldsymbol{B}_\mu|$ and the angular frequency $\gamma_\mu B_\mu$ is known as Larmor precession.

By assuming that the internal field is always perpendicular to the muon polarization (e.g., if an external field is applied), the first term on the right-hand side of eqn [8] disappears and a simple oscillation of the muon precession is observed (see **Figure 2**). On the other hand, if the internal field direction is entirely random (e.g., for spontaneous fields in a magnetic polycrystalline sample), then a simple averaging would yield

$$G(t) = \tfrac{1}{3} + \tfrac{2}{3}\cos(\gamma_\mu B_\mu t) \qquad [9]$$

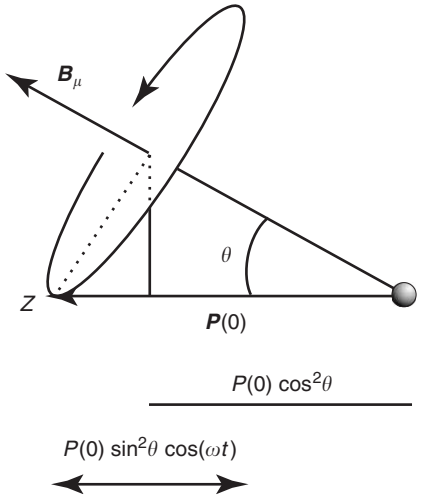


Figure 3 Muon-spin precession around an internal field B_μ at an angle θ. The wiggling and nonwiggling components of the polarization are schematically represented.

Considering the gyromagnetic ratio of the muon and a typical measurement time-window of 10–20 μs, tiny internal fields as low as 0.1 G can be detected by the μSR technique.

So far, a well-defined value of the internal field was assumed. But nature is of course more complicated and usually a distribution has to be considered. Such a distribution can arise from disorder, dynamics, or additional sources of internal field (as nuclear moments). In the case of static local magnetic fields, and assuming a magnetic field distribution function $f(\boldsymbol{B}_\mu)$, $G(t)$ is given by

$$G(t) = \int f(\boldsymbol{B}_\mu)[\cos^2\theta + \sin^2\theta\,\cos(\gamma_\mu B_\mu t)]\mathrm{d}\boldsymbol{B}_\mu \qquad [10]$$

In other words, the value of the Larmor precession observed in the μSR signal provides the value of the internal field, whereas the envelope of the signal provides information on the type and strength of the field distribution itself.

Hence, even field distributions centered around zero can be detected through a depolarization of the μSR signal. As an example, if the strength of the local magnetic field is taken from a Gaussian distribution with zero average, then one observes a signal

$$G(t) = \frac{1}{2} + \frac{2}{3}(1 - \Delta^2 t^2)\exp\left(-\frac{1}{2}\Delta^2 t^2\right) \qquad [11]$$

where Δ^2/γ_μ^2 represents the second moment of the field distribution along one Cartesian axis perpendicular to the initial μ^+-polarization, that is,

$$\Delta^2/\gamma_\mu^2 = \langle B_x^2 \rangle = \langle B_y^2 \rangle = \langle B_z^2 \rangle \qquad [12]$$

This relaxation function is often observed experimentally and reflects the field due to the neighboring nuclear dipoles, which are randomly oriented and considered static within the μSR time window.

Transverse-Field (TF) Technique

For this technique, an external field $\boldsymbol{H}_{\mathrm{ext}}$ is applied perpendicular to the initial muon polarization $\boldsymbol{P}_\mu(0)$. $\boldsymbol{P}_\mu(t)$ precesses around the total field $\boldsymbol{B}_\mu$ at the muon site, the value of which can be extracted from the oscillatory component of $G(t)$, recalling that the μ^+ frequency $\nu_\mu = \omega_\mu/2\pi = \gamma_\mu B_\mu/2\pi$ (see eqn [8]).

For a system which is not magnetically ordered, the total field at the muon site is given by

$$\boldsymbol{B}_\mu = \boldsymbol{H}_{\mathrm{ext}} + \boldsymbol{B}_{\mathrm{hf}} + (\boldsymbol{B}_{\mathrm{DF}} + \boldsymbol{B}_{\mathrm{LF}}) \qquad [13]$$

The last two terms are the demagnetization and Lorentz fields which can be easily calculated from the bulk magnetization and demagnetization tensor. $\boldsymbol{B}_{\mathrm{hf}}$ are the internal fields induced by $\boldsymbol{H}_{\mathrm{ext}}$.

The μ^+-frequency shift

$$K_\mu^* = \frac{|\boldsymbol{B}_\mu| - |\boldsymbol{H}_{\mathrm{ext}}|}{|\boldsymbol{H}_{\mathrm{ext}}|} = \frac{\omega_\mu}{\omega_{\mathrm{ext}}} - 1 \qquad [14]$$

(where $\omega_{\mathrm{ext}} = \gamma_\mu|\boldsymbol{H}_{\mathrm{ext}}| \equiv \gamma_\mu H_{\mathrm{ext}}$) can be corrected for the contributions of $\boldsymbol{B}_{\mathrm{DF}}$ and $\boldsymbol{B}_{\mathrm{LF}}$ to furnish the μ^+-Knight shift K_μ which, as in NMR, contains the relevant information about the hyperfine fields, and therefore about the local susceptibility. In addition, the knowledge of the μ^+-Knight shift provides valuable indication concerning the type of interstitial site occupied by the muon in the crystallographic lattice.

In addition, the TF technique allows one to gain more insight into the inhomogeneous field distribution at the muon site caused by static fields and into the presence of $1/T_2$ processes responsible for homogeneous line broadening.

Zero-Field (ZF) and Longitudinal-Field (LF) Techniques

With the ZF technique, one monitors the time evolution of the muon ensemble under the action of internal magnetic fields. This technique is widely utilized to measure the spontaneous μ^+-Larmor frequencies in magnetically ordered phases, providing information about the values of the static moment and the type of magnetic structures (see, e.g., eqn [9]). The observation of spontaneous μ^+-frequencies is, however, limited to systems where the presence of

static electronic moments produces well-defined local fields at the μ^+-stopping site. Hence, the μ^+-spin autocorrelation function $G(t)$ depends sensitively on details of the magnetic structure and of course on the μ^+-stopping site.

In a ZF experiment, a depolarization can occur due to the presence either of a static distribution of internal fields arising from static nuclear or electronic dipole fields (see eqn [11]), or of a time dependence of the internal fields at the muon site. This time dependence may arise from a fluctuation of the magnetic moments or from muon diffusion. For fast fluctuation (diffusion) the depolarization function will no more assume a Gaussian character at early times, but will be given by

$$G(t) = \exp(-\lambda t), \quad \lambda = 2\Delta^2\tau \qquad [15]$$

where τ represents the characteristic fluctuation (diffusion) time.

In an LF configuration (i.e., $\boldsymbol{H}_{\text{ext}}||\boldsymbol{P}_\mu(0)$), by choosing $\boldsymbol{H}_{\text{ext}}$ to be stronger than the internal fields, the muon's states "up" and "down" are eigenstates of the Zeeman Hamiltonian and any inhomogeneous (static) distribution of the internal fields will not affect the time evolution of the μ^+-polarization, which will remain constant. This behavior reflects the decoupling of the μ^+-spin from the static internal fields. Hence, to pinpoint the origin of the depolarization, decoupling experiments in a longitudinal field are routinely performed.

μSR and Condensed Matter

The aim of this section is to present some aspects of the application of μSR in a condensed matter. The interested reader is referred to more complete review articles.

Magnetism

There are several advantages in using the μSR technique to study magnetic systems.

- As said, due to the large muon magnetic moment, the technique is sensitive to extremely small internal fields, and therefore can probe local magnetic fields which can be nuclear or electronic in nature.
- The local probe character of the muon makes μSR very sensitive to spatially inhomogeneous magnetic properties. Hence, the occurrence of different phases in the sample will be reflected by different components in the μSR signal, that is,

$$\mathcal{A}G(t) = \sum_i \mathcal{A}_i G_i(t) \qquad [16]$$

  A careful analysis of the amplitude $\mathcal{A}_i$ of these components furnishes a direct measure of the fraction of the sample involved in a particular phase (with the condition $\sum_i \mathcal{A}_i = \mathcal{A}$). Similarly, μSR is a powerful tool when the magnetic order is of short range and/or random nature where neutron experiments will fail. In addition, the μSR technique can be utilized to check the coexistence of different types of ground states at the microscopic level.
- In magnetic phases, the possibility given by μSR to extract independently the value of the ordered static moment and the sample volume involved in the magnetic phase allows one to overcome a shortcoming of other techniques (such as the neutron scattering technique), for which these two parameters are strongly coupled, therefore hampering their absolute determination.

Organic magnets constitute a typical example of the application of the μSR technique to study magnetic phenomena. It is well known that neutron scattering can provide extensive information on magnets, in particular, on the magnetic structure and on the excitations. But in compounds with small magnetic moments such as the organic magnets, the μSR technique is a sensitive and useful probe: it easily yields the value of the critical temperature and information on the thermal behavior of the order parameter. Hence, the first direct observation of spontaneous magnetic order in the system para-nitrophenyl nitronyl nitroxide (*p*-NPNN) was made using the μSR technique, and has been subsequently confirmed by neutron diffraction.

Figure 4 exhibits the zero-field muon spin rotation frequency in the organic ferromagnet *p*-NPNN, which orders only at very low temperatures ($T_C = 0.67$ K. The temperature dependence of the frequency of oscillations is a direct measure of the temperature dependence of the internal magnetization, that is, of the ferromagnetic order parameter.

Superconductors

An important research area of recent μSR studies is the characterization of the superconducting state.

Meissner state For a superconducting sample in the Meissner state, diamagnetic shielding currents flowing close to the sample surface screen a static magnetic field from the interior of the body. Near the sample surface, the magnetic field falls off exponentially with distance over a characteristic length scale λ, called the "magnetic penetration depth." Inversely, the superconducting order parameter increases and is

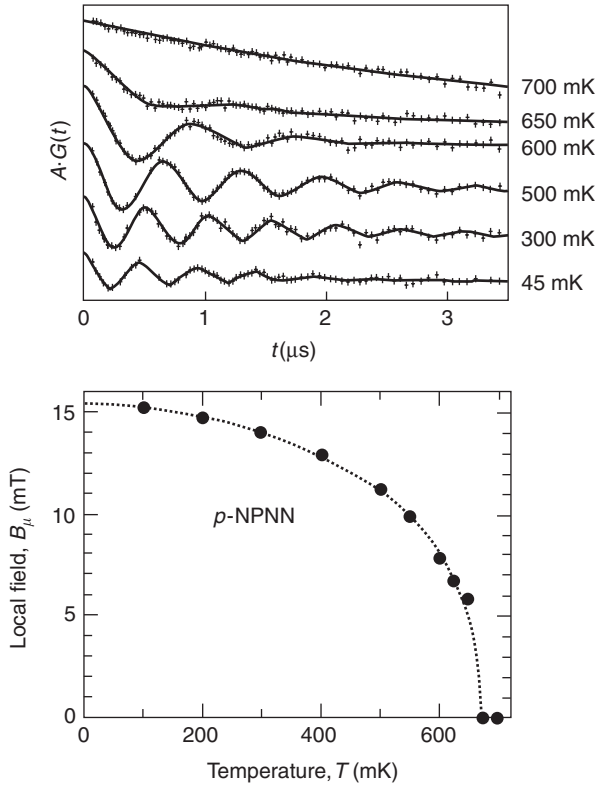


Figure 4 Zero-field muon-spin rotation frequency in the organic ferromagnet *p*-NPNN (Reproduced with permission from Blundell SJ (1999) Spin-polarized muons in condensed matter physics. *Contemporary Physics* 40: 175.) and temperature dependence of the local field at the muon site deduced from the observed frequency.

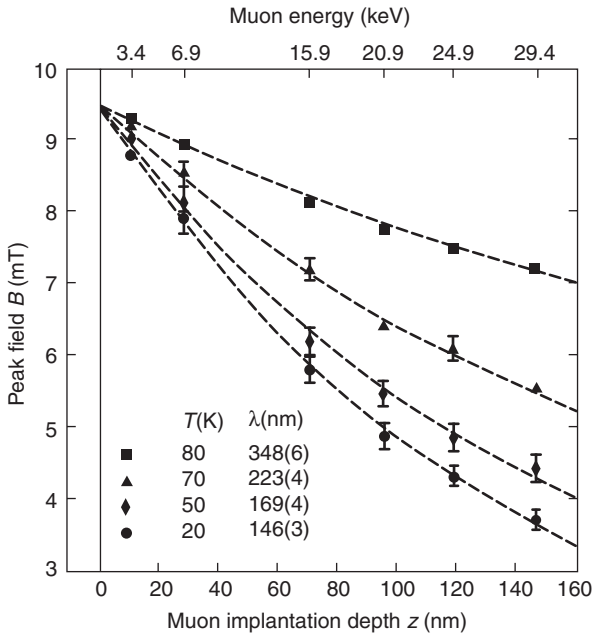


Figure 5 Values of field vs. depth measured in a high-T_c sample ($YBa_2Cu_3O_{7-\delta}$) for various values of temperature. The muon implantation depth is changed by tuning the kinetic energy of the muons. (Reproduced with permission from Jackson TJ, Riseman TM, Forgan EM, Glückler H, Prokscha T, *et al.* (2000) Depth resolved profile of the magnetic field beneath the surface of a superconductor with a few nm resolution. *Physical Review Letters* 84: 4958.)

characterized by the coherence length ξ which controls the length scale over which the superconducting order parameter can vary near the surface without undue energy cost. In the limit where the penetration depth is much larger than the coherence length ξ, and where the mean free path for scattering of the normal electrons $l \gg \xi$, a magnetic field applied parallel to the surface of a semi-infinite superconducting slab in the Meissner state decays with depth z according to the exponential decay law

$$B(z) = B \exp(-z/\lambda) \qquad [17]$$

Although generally accepted, such a law was only very recently confirmed on a high-T_c system ($YBa_2Cu_3O_{7-\delta}$) by using low-energy muons. In this experiment, muons were implanted at several different depths in the superconducting film, by adjusting their kinetic energy (**Figure 5**). Therefore, the internal field $B(z)$ could be measured directly as a function of depth, and the validity of eqn [17] could be verified experimentally for the first time. In contrast, with bulk experiments, such as magnetization, or even μSR using usual muons, the funtional form of eqn [17] is assumed when appropriate, and the magnetic penetration depth is extracted. This direct determination illustrates the power of low-energy muons for near surface studies in superconductivity and magnetism.

Vortex state It is well known that for a type-II superconductor in the mixed (or vortex) phase, an applied magnetic field $H_{c1} < H < H_{c2}$ enters the superconductor in the form of quantized flux lines or vortices arranged into a lattice, usually hexagonal. Each vortex contains one quantum of flux $\Phi_0 = h/2e$, and the distance d between vortices is, except at very high magnetic fields, much larger than the unit-cell dimensions (given for a hexagonal vortex-lattice by the relation $d^2 = 2\Phi_0/\sqrt{3}B$). Outside the core region of a vortex, the magnetic field falls off exponentially with distance over a characteristic penetration depth λ (see above). Since the implanted muons sit at certain crystallographic sites and since the vortex lattice is incommensurate with the crystal lattice, the muon ensemble will randomly sample the field distribution of the vortex lattice (see **Figure 6**). This field distribution will create a dephasing of the muon polarization, reflected by a relaxation of the μSR signal. Assuming a Gaussian

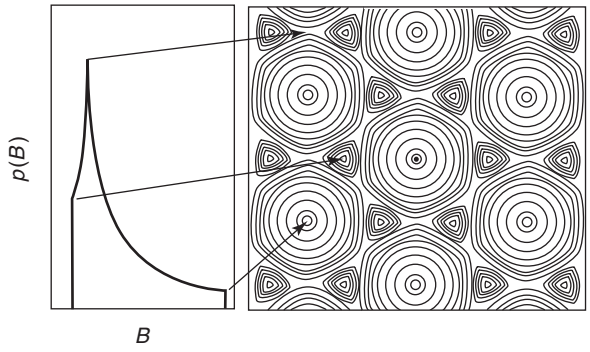


Figure 6 Schematic field distribution in a type-II superconductor in the mixed (vortex) state. The right-hand side of the figure represents a contour plot of the field *B*. (Reproduced with permission from Blundell SJ (1999) Spin-polarized muons in condensed matter physics. *Contemporary Physics* 40: 175.)

relaxation of the signal, the relaxation rate of the observed precession signal is related to the penetration depth λ by

$$\sigma = \gamma_\mu \langle B^2(\boldsymbol{r}) - \langle B(\boldsymbol{r}) \rangle^2 \rangle^{1/2} = \frac{0.0609 \gamma_\mu \Phi_0}{\lambda^2} \quad [18]$$

where $B(\boldsymbol{r})$ is the field at position **r**, and the averages are taken over all positions. Therefore, the relaxation rate σ of the μSR precession signal is a measure of the magnetic penetration depth. Its temperature dependence yields valuable information concerning the symmetry of the superconducting gap and therefore about the pairing mechanism. In addition to the determination of the magnetic penetration depth, a detailed analysis of the data provides information about the coherence length ξ.

Semiconductors

As said, implanted into semiconductors and insulators, the muon often picks up an electron and forms muonium. Muonium can exist in three charge states: Mu^0, Mu^+, and Mu^- which correspond to the three distinct charge states of isolated hydrogen: H^0, H^+, and H^-. This opens new directions of investigation as the behavior of muonium provides important information on the charge state, lattice site, and dynamics of its hydrogen analog. For example, implanted in silicon, neutral muonium has been observed at the tetrahedral site (T) and at a bond-centered (BC) site between two silicon atoms (**Figure 7**). While the muonium at the T site was found to be mobile, and rapidly diffusing between such sites, the BC muonium is stationary. The electronic state in the

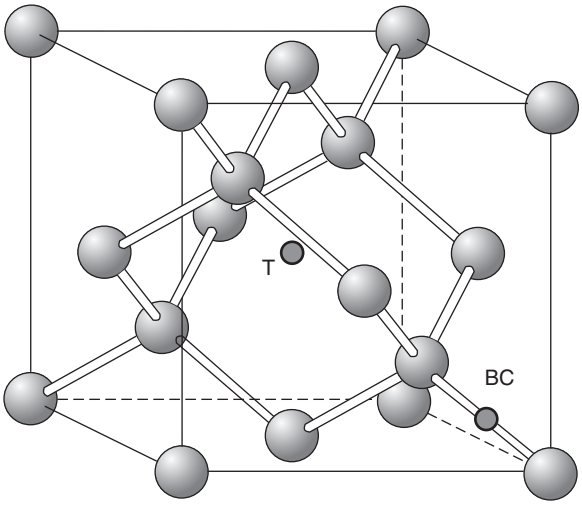


Figure 7 The crystal structure of silicon, showing the possible muonium (and therefore hydrogen) sites (T = tetrahedral site, BC = bond-center site).

T site is isotropic, whilst it is highly anisotropic for the BC site, with axial symmetry along the Si–Si bond. Such site characterization appears rather important since hydrogen is present, as an impurity, in all semiconductors but its low concentration makes it difficult to be investigated using direct spectroscopical methods. In addition, hydrogen is able to passivate dangling bounds known to be present at grain boundaries in the promising new multicrystalline Si photovoltaic cells.

In certain II–VI compounds, on the other hand, it appears that hydrogen may instead act as a shallow donor – in other words, it could be used as a dopant in its own right. This has important implications for the electronic properties of new wide-gap materials, now under development for optoelectronic applications. The most recent and compelling evidence for these shallow donor states comes from studies not of hydrogen itself but of its pseudo-isotope, muonium. Such results are particularly significant in that they confirm recent theoretical predictions.

See also: Magnetism, History of; Semiconductors, General Properties; Solid-State NMR Structural Studies of Proteins; Superconductivity: General Aspects.

PACS: 76.75.+l; 71.27.+a; 75.30.m; 75.70.Ak; 74.25.Qt; 74.72.Bk; 61.72.Ji; 61.72.Vv

Further Reading

Amato A (1997) Heavy-fermion systems studied by μSR technique. *Reviews of Modern Physics* 69: 1119.

Bakule P and Morenzoni E (2004) Generation and applications of slow polarized muons. *Contemporary Physics* 45: 203.

Dalmas de Réotier P and Yaouanc P (1997) Muon spin rotation and relaxation in magnetic materials. *Journal of Physics: Condensed Matter* 9: 9113.

Kalvius GM, Noakes DR, and Hartmann O (2001) μSR studies of rare-earth and actinide magnetic materials. In: Gschneidner Jr KA, Eyring L, and Lander GH (eds.) *Handbook on the Physics and Chemistry of Rare Earths*, vol. 32, p. 55. Amsterdam: North-Holland.

Schenck A (1985) *Muon Spin Rotation Spectroscopy*. Bristol: Adam Hilger.

Schenck A and Gygax FN (1995) Magnetic materials studied by muon spin rotation spectroscopy. In: Buschow KHJ (ed.) *Handbook of Magnetic Materials*, vol. 9, p. 57. Amsterdam: Elsevier.

Nomenclature

e	elementary electrical charge (Cb)
e^-	electron
e^+	positron
G	muon depolarization function
$\mathbf{H}_{\text{ext}}$	applied magnetic field (Oe)
Mu	muonium
n	neutron
p	proton
ν_e	electron neutrino
W	asymmetry parameter positron emission
$\mathbf{B}_\mu$	magnetic field at muon site (G)
$\mathbf{P}_\mu$	muon polarization
ν_μ	muon neutrino
Δ	Gaussian depolarization rate (μs^{-1})
γ_μ	gyromagnetic ratio of the muon ($Hz\,G^{-1}$)
λ	Lorentzian depolarization rate (μs^{-1})
λ	London magnetic penetration depth (m)
μ^+	muon
μ_N	nuclear magneton ($erg\,G^{-1}$)
μ_p	proton magnetic moment ($erg\,G^{-1}$)
ξ	superconducting coherence length (m)
π^+	pion
τ	magnetic fluctuation time (s)
τ_μ	lifetime of the muon (s)
Φ_0	magnetic flux quantum ($G\,cm^{-2}$)

N

Nanostructures, Electronic Structure of

J R Chelikowsky, University of Texas at Austin, Austin, TX, USA

Introduction

At the atomic scale, electronic states are typically characterized by discrete energy levels that are often separated by electron volts; the spatial distribution of these states is highly localized. In contrast, electronic states in crystalline matter are typically characterized by energy bands: the energy states form a quasicontinuous spectrum with delocalized spatial distributions. At the nanoscale, the distribution of energy states resides between these limits. This evolution of energy levels from an atom to large clusters to "quantum dots" is illustrated in **Figure 1**. A large cluster in this scheme may possess hundreds or thousands of atoms. When its surface is electronically passivated, such large clusters are called quantum dots. While the energy spacing between electronic levels in a quantum dot is not quasicontinuous, it is usually too small to be measured. For example, the spacing may be much smaller than the thermal energy associated with a vibrational mode.

Understanding the evolution of the electronic structure of matter at the nanoscale is a challenging task. Matter at this length scale is not especially amenable to techniques based on atomic or macroscopic descriptions, as the spatial and energetic distribution of electron states are quite disparate. Methods for calculating the electronic properties of these systems usually start from either the atomic or molecular limit or from the solid-state limit. Atomic or molecular methods are often not equipped to handle the large number of atoms or electrons in nanoscale materials whereas solid-state methods often require translational symmetry, which is not present for confined systems. Moreover, the electronic interactions and spatial extent of wave functions at the nanoscale can be remarkably different than either the atomic or macroscopic limit.

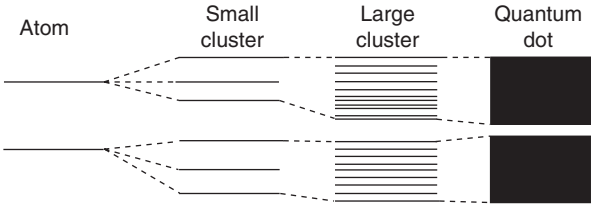


Figure 1 Schematic energy levels for an atom, a cluster, and a quantum dot.

An electron confined within a one-dimensional box of size a is considerd. The lowest energy level of this system is given in elementary textbooks as

$$E = \frac{\pi^2 \hbar^2}{2ma^2} \qquad [1]$$

where $\hbar$ is the Planck constant divided by 2π, and m is the mass of the particle, for example, the mass of an electron. If the size of the box is reduced, then E increases.

This phenomenon, that is, the energy level spacing increasing with reduced dimensionality, is called quantum confinement. Quantum confinement can be readily understood from the Heisenberg uncertainty principle. The uncertainty principle states that the uncertainty in momentum and position must be such that the product exceeds $\hbar/2$:

$$\Delta p \cdot \Delta x \geqslant \frac{\hbar}{2} \qquad [2]$$

where Δp is the uncertainty in the momentum and Δx is the uncertainty in the position of the electron. Consider the energy of a free electron with momentum p:

$$E = \frac{p^2}{2m} \qquad [3]$$

Since the uncertainty in the momentum cannot exceed the momentum itself, one can write $p > \Delta p$. As such, one has the following inequality:

$$E > \frac{\hbar^2}{8ma^2} \qquad [4]$$

If one tries to localize the position of an electron by reducing the box size, its energy must increase and diverge as the confining region vanishes.

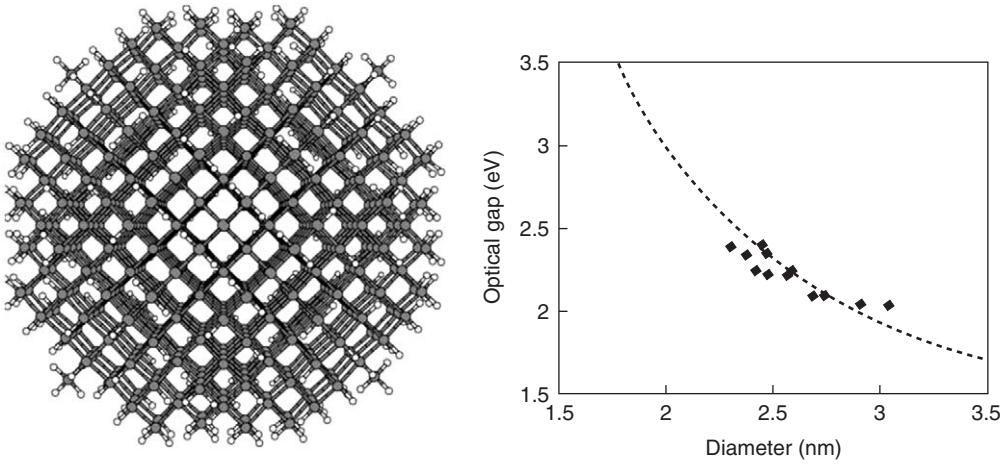


Figure 2 Ball and stick model for a quantum dot of hydrogenated silicon (left). The gray balls represent the silicon atoms, the white balls represent the hydrogens. Also shown is the corresponding optical absorption. Note the increase in gap size from the known bulk gap of silicon, which is 1.1 eV.

Although this very simple picture of quantum confinement is given for one dimension, it is also true for an electron confined in three dimensions. For example, if a particle is confined within a sphere whose radius is R, one might expect an energy level to vary with the radius according to

$$E(R) = E_\infty + \frac{\alpha}{R^2} \qquad [5]$$

where α is a constant and E_∞ is the energy as $R \to \infty$. Suppose one considers an optical gap, E_{gap}, as the difference between the high filled and lowest empty states for a system confined in a sphere of size R. In the simplest description of the gap, one might expect E_{gap} to scale as

$$E_{\text{gap}}(R) = E_{\text{gap}} + \frac{\beta}{R^2} \qquad [6]$$

where β is a constant.

Quantum confinement has been observed experimentally for nanostructures such as quantum dots of Si and CdSe. In **Figure 2**, a quantum dot of hydrogenated silicon is illustrated. The interior of the dot consists of silicon atoms in the bulk diamond structure; the surface of the dot is hydrogenated. Hydrogen removes any dangling bonds on the surface. As such, all the atoms in the system should be fully coordinated and should not contribute to the electronic properties of the quantum dot. Typically, a quantum dot is a few nanometers to tens of nanometers in size. The smallest dots contain hundreds of atoms whereas large dots may contain hundreds of thousands of atoms.

Also illustrated in **Figure 2** is the corresponding optical gap as a function of the dot size. The quantum dot gap with this size range is approximately twice the gap of crystalline silicon; silicon exhibits weak optical absorption near 1.1 eV. The absorption is weak owing to the nature of the gap in silicon. It is an indirect gap because the optical excitation couples an electron with a hole of different crystalline momentum or wave vector. Conservation of crystal momentum would require the electron and hole to have the same momentum. As such, this process can only occur by the participation of the quantized vibrational modes of the lattice or phonons, which provide the required momentum to satisfy the conservation rules. Unlike a two-body process (electron and hole), a three-body process (electron, hole, and phonon) results in a very weak interaction between light and matter.

In finite systems, crystal momentum is not well defined, owing to the lack of translation symmetry. As the system approaches nanoscale sizes, the nature of an indirect gap or direct gap becomes moot. Consequently, the optical absorption becomes stronger in quantum dots and the size of the gap increases owing to quantum confinement. An approximate fit of eqn [6] is consistent with the experimental data shown in **Figure 2**. This example shows the striking effect of size on the electronic structure of nanostructures, that is, silicon at this length scale is transformed from an optical inactive material to an optically active one.

Theory for the Electronic Structure at the Nanoscale

Quantum mechanical laws that describe the behavior of matter at the nanoscale were discovered in the first

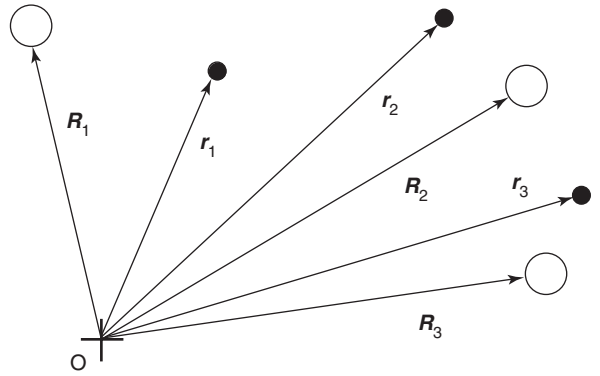


Figure 3 Simplified picture of electrons and nuclei in matter.

part of the twentieth century. Using these laws, it is possible to predict the electronic structure of matter from the atomic- to nano- to the macroscale, at least in principle. While it is relatively easy to write down the equations for interacting atoms, that is, interacting nuclei and electrons, obtaining a solution to the problem that is sufficiently accurate to make predictions is a formidable task.

Consider N nuclei of charge Z_n at positions $\{\boldsymbol{R}_n\}$ for $n = 1, \ldots, N$ and M electrons at positions $\{\boldsymbol{r}_i\}$ for $i = 1, \ldots, M$. This is shown schematically in **Figure 3**. The quantum mechanical equation for this system in its simplest form can be written as

$$\begin{aligned}\mathscr{H}(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \ldots; \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \ldots) &= \sum_{n=1}^{N} \frac{-\hbar^2 \nabla_n^2}{2\mathscr{M}_n} + \frac{1}{2} \sum_{n,m=1, n\neq m}^{N} \frac{Z_n Z_m e^2}{|\boldsymbol{R}_n - \boldsymbol{R}_m|} \\ &\quad - \sum_{i=1}^{M} \frac{\hbar^2 \nabla_i^2}{2m} - \sum_{n=1}^{N} \sum_{i=1}^{M} \frac{Z_n e^2}{|\boldsymbol{R}_n - \boldsymbol{r}_j|} \\ &\quad + \frac{1}{2} \sum_{i,j=1, i\neq j}^{M} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \end{aligned} \quad [7]$$

$\mathscr{M}_n$ is the mass of the nucleus. This expression omits some terms such as those involving relativistic interactions, but captures the essential features for nanoscale matter.

Using the Hamiltonian in eqn [7], the quantum mechanical equation known as the Schrödinger equation for the electronic structure of the system can be written as:

$$\begin{aligned}&\mathscr{H}(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \ldots; \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \ldots) \\ &\quad \times \Psi(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \ldots; \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \ldots) \\ &\quad = E\Psi(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \ldots; \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \ldots)\end{aligned} \quad [8]$$

where E is the total electronic energy of the system and Ψ is the many-body wave function.

Soon after the discovery of the Schrödinger equation, it was recognized that this equation provided the means of solving for the electronic and nuclear degrees of freedom. Using the variational principle which states that an approximate wave function will always have a less favorable energy than the true ground-state energy, one had an equation and a method to test the solution. One can estimate the energy from

$$E = \frac{\int \Psi^* \mathscr{H} \Psi \mathrm{d}^3 R_1 \mathrm{d}^3 R_2 \mathrm{d}^3 R_3 \ldots \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \mathrm{d}^3 r_3 \ldots}{\int \Psi^* \Psi \mathrm{d}^3 R_1 \mathrm{d}^3 R_2 \mathrm{d}^3 R_3 \ldots \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \mathrm{d}^3 r_3 \ldots} \quad [9]$$

However, a solution of eqn [8] for anything more complex than a few particles becomes problematic even with the most powerful computers. Obtaining an approximate solution for systems with many atoms is difficult, but considerable progress has been made since the advent of reliable computers.

Approximations to the Electronic Structure Problem

A number of highly successful approximations have been made to solve for both the ground-state and excited-state energies of matter. These approximations are constructed to remove as many unimportant degrees of freedom from the system as possible.

For example, one common approximation is to separate the nuclear and electronic degrees of freedom. Since the nuclei are considerably more massive than the electrons, the electrons will respond instantaneously to any changes in the nuclear coordinates. This approximation is called the Born–Oppenheimer or adiabatic approximation. It allows one to treat the nuclear coordinates as classical parameters. For most condensed matter systems, this assumption is highly accurate.

Another highly useful approximation is the pseudopotential approximation. The pseudopotential model of a solid has led the way in providing a workable model for computing the electronic properties of materials. Using pseudopotentials, it is possible to predict accurately the properties of complex nanosystems such as quantum dots with hundreds, if not thousands of atoms. The pseudopotential model treats matter as a "sea of valence electrons" moving in a background of ion cores (**Figure 4**). The ion cores are composed of nuclei and inert inner electrons. Within this model, many of the complexities of an all-electron calculation are avoided. Since only the valence electrons are considered, a group IV solid such as C with six electrons is no more difficult than Pb with 82 electrons as both elements have four valence electrons.

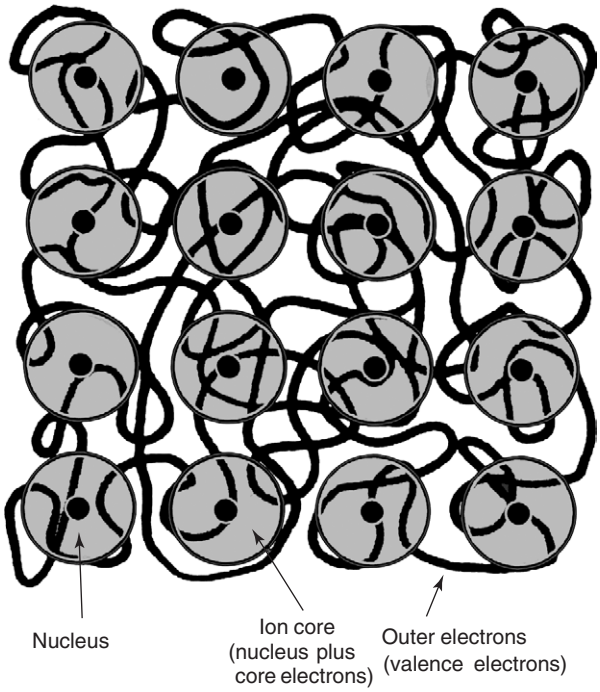


Figure 4 Schematic of the pseudopotential model of condensed matter.

Pseudopotential calculations center on replicating the wave function in the spatial region from the core, that is, within the chemically active bonding region. The smoothly varying pseudo-wave-function is taken to be identical to the all-electron wave function in the bonding regions.

A third approximation is based on the density-functional theory. Methods for understanding the electronic structure of matter based on a knowledge of the charge density have existed since the late 1920s. For example, the Thomas–Fermi description of atoms was based on a knowledge of the electronic density. This theory was one of the first attempts at a quantitative theory for the electronic structure of atoms.

Most contemporary treatments of the density-functional theory begin by considering a free electron gas of uniform charge density. The justification for this starting point comes from the observation that simple metals, such as aluminum and sodium have properties that appear to resemble those of a free electron gas. In the mid 1960s, Hohenberg, Kohn, and Sham published a classic set of papers on the density-functional theory. Their work justified the replacement of the many-body wave function by one-electron orbitals based on the electronic charge density.

A key aspect of their work is the local density approximation. Within this approximation, one can express the exchange energy as

$$E_x[\rho(\boldsymbol{r})] = \int \rho(\boldsymbol{r})\mathscr{E}_x[\rho(\boldsymbol{r})]\mathrm{d}^3r \qquad [10]$$

where $\mathscr{E}_x[\rho]$ is the exchange energy per particle of uniform gas at a density of ρ. Within this framework, the true exchange potential is replaced by a potential determined from the functional derivative of $E_x[\rho]$:

$$V_x[\rho] = \frac{\delta E_x[\rho]}{\delta\rho} \qquad [11]$$

One serious issue is the determination of the exchange energy per particle, $\mathscr{E}_x$, or the corresponding exchange potential, V_x. The exact expression for either of these quantities is unknown, save for special cases. If one assumes the exchange energy to be given by the Hartree–Fock expression for the exchange energy of the free electron gas, then one can write

$$E_x[\rho] = -\frac{3e^2}{4\pi}(3\pi^2)^{1/3}\int[\rho(\boldsymbol{r})]^{4/3}\mathrm{d}^3r \qquad [12]$$

and taking the functional derivative, one obtains

$$V_x[\rho] = -\frac{e^2}{\pi}(3\pi^2\rho(\boldsymbol{r}))^{1/3} \qquad [13]$$

In contemporary theories, numerical studies have been performed on uniform electron gases resulting in local density expressions of the form: $V_{xc}[\rho(\boldsymbol{r})] = V_x[\rho(\boldsymbol{r})] + V_c[\rho(\boldsymbol{r})]$, where V_c represents contributions to the total energy beyond the Hartree–Fock limit. It is also possible to describe the role of spin explicitly by considering the charge density for up and down spins: $\rho = \rho_\uparrow + \rho_\downarrow$. This approximation is called the local spin density approximation.

The Kohn–Sham Equation

The work of Kohn and Sham resulted in a workable framework for the electronic structure problem. Specifically, the Kohn–Sham equation for the electronic structure of matter is given by

$$\left(\frac{-\hbar^2\nabla^2}{2m} + V_p^{\mathrm{ion}}(\boldsymbol{r}) + V_H(\boldsymbol{r}) + V_{xc}[\rho(\boldsymbol{r})]\right) \times \phi_i(\boldsymbol{r}) = E_i\phi_i(\boldsymbol{r}) \qquad [14]$$

V_p^{ion} is the ion core pseudopotential, V_H is the Hartree or Coulomb potential, and V_{xc} is the exchange–correlation potential. The Hartree potential

is obtained from

$$\nabla^2 V_{\mathrm{H}}(\boldsymbol{r}) = -4\pi e\rho(\boldsymbol{r}) \quad [15]$$

where the charge density is given by

$$\rho(\boldsymbol{r}) = -e \sum_{i,\mathrm{occup}} |\phi_i(\boldsymbol{r})|^2 \quad [16]$$

The summation is over occupied states. When pseudopotentials are employed, these states correspond to the valence states of the system.

The Kohn–Sham equation is solved self-consistently. An approximate charge is assumed to estimate the exchange–correlation potential and to determine the Hartree potential from eqn [15]. These estimated potentials are inserted in the Kohn–Sham equation. Once the Kohn–Sham equation is solved, a new output charge density is determined from eqn [16]. The output charge density is used to estimate new exchange–correlation and Hartree potentials. The process is repeated until the input and output charge densities or potentials are identical to within some prescribed tolerance.

Given a solution of the Kohn–Sham equation, the total energy, E_{T}, can be computed from

$$E_{\mathrm{T}} = \sum_i^M E_i - 1/2 \int \rho(\boldsymbol{r}) V_{\mathrm{H}}(\boldsymbol{r}) \mathrm{d}^3 r + \int \rho(\boldsymbol{r})(E_{\mathrm{xc}}[\rho(\boldsymbol{r})] - V_{\mathrm{xc}}[\rho(\boldsymbol{r})])\mathrm{d}^3 r \quad [17]$$

where E_{xc} is a generalization of eqn [10], that is, the correlation energy density is included. The electronic energy as determined from E_{T} must be added to the ion–ion interactions to obtain the structural energies. This is a straightforward calculation for confined systems.

Owing to its ease of implementation and overall accuracy, the local density approximation is a popular choice for describing the electronic structure of matter. It is relatively easy to implement and surprisingly accurate. Recent developments have included the so-called generalized gradient corrections to the local density approximation. The generalized gradient approximation often yields accurate cohesive energies.

Computational Approaches for Nanostructures

Several methods exist for determining the electronic structure of nanostructures. These methods are often based on choosing a basis set for the electronic wave functions and directly solving the Kohn–Sham equation. For example, the wave functions can be expanded as

$$\Psi_n(\boldsymbol{r}) = \sum_{jk} a_{j,k}(n)\Phi_j(\boldsymbol{r} - \boldsymbol{R}_k) \quad [18]$$

where the sum is over (j, k). These indices label the atomic positions, given by $\boldsymbol{R}_k$, and a local orbital, Φ_j, associated with an (s, p, d, …) state. The local orbitals, Φ_j, can be taken to be atomic orbitals, Gaussians, exponentials, and so on. With the form of eqn [18] as the basis, the eigenvalue problem takes the form of the matrix equation:

$$\mathbb{H}^{\mathrm{KS}}\boldsymbol{\Psi}_n = E_n \mathbb{S}\boldsymbol{\Psi}_n \quad [19]$$

where the Hamiltonian matrix elements, $H_{\mathrm{lm}}^{\mathrm{KS}}$, are given by

$$H_{\mathrm{lm}}^{\mathrm{KS}} = \sum_{jkj'k'} a_{j,k}^*(l) a_{j',k'}(m) \int \Phi_j^*(\boldsymbol{r} - \boldsymbol{R}_k) \times H^{\mathrm{KS}} \Phi_{j'}(\boldsymbol{r} - \boldsymbol{R}_{k'})\mathrm{d}^3 r \quad [20]$$

and the overlap matrix elements, S_{lm}, is given by

$$S_{\mathrm{lm}} = \sum_{jkj'k'} a_{j,k}^*(l) a_{j',k'}(m) \int \Phi_j^*(\boldsymbol{r} - \boldsymbol{R}_k) \times \Phi_{j'}(\boldsymbol{r} - \boldsymbol{R}_{k'})\mathrm{d}^3 r \quad [21]$$

The Kohn–Sham Hamiltonian is given as in eqn [14]:

$$H^{\mathrm{KS}} = \frac{-\hbar^2\nabla^2}{2m} + V_{\mathrm{p}}^{\mathrm{ion}}(\boldsymbol{r}) + V_{\mathrm{H}}(\boldsymbol{r}) + V_{\mathrm{xc}}[\rho(\boldsymbol{r})] \quad [22]$$

The matrix elements can be computed directly if H^{KS} is specified. Otherwise, the matrix elements can be fit to experiment. Once the matrix elements are specified, the eigenvalue and eigenfunctions can be found using standard matrix operation packages.

Methods using local orbitals can be very efficient for nanostructures. If the local orbital expansion is restricted to a small number of orbitals, the size of the matrix in eqn [19] can be quite tractable.

Another option is to use a plane wave basis instead of a localized basis, such as a set of Gaussian orbitals. In the case of a plane wave basis, pseudopotentials play a crucial role. The pseudopotential can be expanded using a rapidly converging set of plane waves. Likewise, no cusps are present in the wave functions allowing a much smaller set of plane waves than would be sufficient to converge the all-electron wave function. The plane wave basis can be

written as

$$\Psi_{n,k'}(\boldsymbol{r}) = \sum_{G} a_G(n) \exp(\mathrm{i}(\boldsymbol{k} + \boldsymbol{G}) \cdot \boldsymbol{r}) \qquad [23]$$

$\boldsymbol{k}$ is a wave vector as for a Bloch state in a crystal and $\boldsymbol{G}$ is a reciprocal lattice vector. The reciprocal lattice vectors are defined by the size and shape of the crystalline unit cell. They are defined such that

$$\exp(\mathrm{i}\boldsymbol{G} \cdot \mathscr{R}) = 1 \qquad [24]$$

where $\mathscr{R}$ is a lattice vector.

A plane wave basis is a well-known and utilized instrument for calculating the electronic structure of crystalline matter. Methods based on a plane wave basis can be implemented for localized systems such as a quantum dot by retaining an artificial periodic system, for example, by placing the system of interest in a large unit cell, or supercell. Provided the cell size is sufficiently large, the solution of the problem results in the electronic structure of the isolated quantum dot. This method is very easy to implement and has been employed for systems of several thousand atoms. A significant complication concerns charged systems. In this case, the supercell is not neutral and the net charge on the system diverges. Compensating background charges are sometimes used to overcome this problem.

It is also possible to solve the Kohn–Sham equations without the use of an explicit basis. One approach is to solve the eigenvalue problem in real space where a grid serves the role of a basis. The real space method is highly advantageous over the plane wave method in a number of respects. Real space methods do not require supercells, and as a consequence, can be employed to examine charged systems without adding compensating backgrounds.

Real space methods for nanostructures are often based on a three-dimensional cubic grid. The Kohn–Sham equation can be discretized via a higher-order finite differencing procedure over the grid to yield

$$\begin{aligned} -\frac{\hbar^2}{2m}\Bigg[& \sum_{n_1=-M}^{M} C_{n_1}\psi_n(x_i + n_1 h, y_j, z_k) \\ & + \sum_{n_2=-M}^{M} C_{n_2}\psi_n(x_i, y_j + n_2 h, z_k) \\ & + \sum_{n_3=-M}^{M} C_{n_3}\psi_n(x_i, y_j z_k + n_3 h)\Bigg] \\ & + [V_{\mathrm{ion}}(x_i, y_j, z_k) + V_{\mathrm{H}}(x_i, y_j, z_k) \\ & + V_{\mathrm{xc}}(x_i, y_j, z_k)]\psi_n(x_i, y_j, z_k) \\ & = E_n\psi_n(x_i, y_j, z_k) \end{aligned} \qquad [25]$$

where h is the grid spacing, $C(n_1, n_2, n_3)$ are coefficients for the finite difference, the summation is over M neighboring grid points, for example, if $M = 1$, then standard finite differencing is used.

Real space methods can be utilized without a uniformly spaced cubic grid, but such implementations are more complicated, especially when used to compute interatomic forces. A uniformly spaced grid within a three dimensional is frequently used for nanostructures. Simple boundary conditions are employed, such as outside a spherical domain, the wave functions are constrained to zero. Special data structures are used to discard these points and keep only those having a nonzero value for the wave function. Iterative diagonalization methods, which can be made very efficient for sparse grids, are easy to implement for real space problems.

Calculated Electronic Levels in Nanostructures

Solutions of the Kohn–Sham equation can be used to describe the electronic structure of a variety of nanostructures such as nanowires, nanotubes, or quantum dots. In the case of nanowires or nanotubes, the structure is localized in two dimensions and often periodic in the third. In this case, Bloch wave functions can be used for the periodic structure. For quantum dots, supercells or localized domains can be utilized along with plane waves or grid methods.

In **Figure 5**, the evolution of energy levels are illustrated for GaAs quantum dots. The stoichiometry

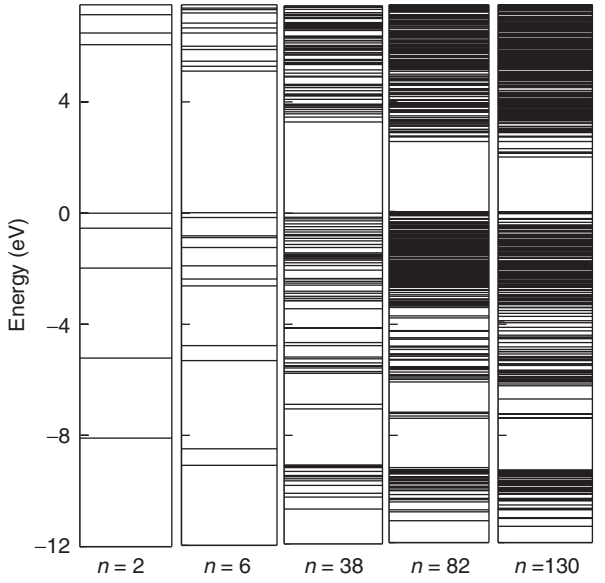


Figure 5 Energy levels of $(GaAs)_n H_m^*$ quantum dots. The number of GaAs units is indicated in the figure. The highest occupied level is taken as the zero energy reference.

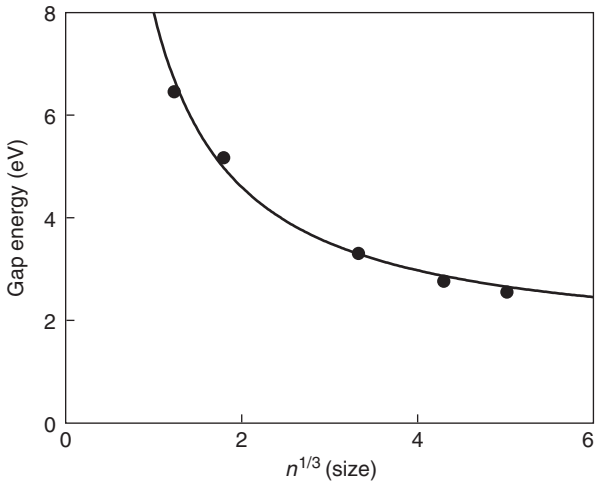


Figure 6 The energy gap in quantum dots of $(GaAs)_nH^*_m$ vs. the size of the dot. The size of the dot is estimated by the cube root of the number of GaAs units in the dot. The curve is after eqn [6] with a scaling of R^{-1}, see text.

of the dots are $(GaAs)_nH^*_m$. The H^*_m notation refers to fictitious hydrogen-like atoms used to passivate any dangling or partially occupied bonds. In this example, the structure of the dots is taken to be that of a fragment of the GaAs crystal. For small dots such as $(GaAs)_2H^*_6$, the energy levels are well separated and the gap between filled and empty states is of the order of $\sim$6 eV. Technically, this gap cannot be compared to the observed gap as these energy levels do not correspond to those created by an electron–hole pair. The Kohn–Sham eigenvalues have meaning only in solving for the total energy as in eqn [17]. However, more realistic approaches to excited state properties have validated the use of Kohn–Sham eigenvalues as a qualitative or, in some cases, a semiquantitative method of determining the optical gap.

As a function of cluster size, the size of the gap between empty and filled states decreases as the size of the cluster increases. This is illustrated in **Figure 6** where the gap size is plotted versus the cube root of the number of GaAs units present. The cube root should scale with the diameter of the quantum dot, provided a large number of atoms are present. The general trend of the gap with the size of the dot is consistent with eqn [6] with one notable exception. The scaling is not consistent with R^{-2} behavior, which would only be expected for a particle contained by an infinite well. Owing to the finite size of the quantum well in real systems, the scaling obtained from accurate calculations yields gaps that scale closer to R^{-1} as indicated in **Figure 6**.

A notable feature of the energy level spacings is the distribution of these states as a function of the dot size. For large systems, the distribution of the

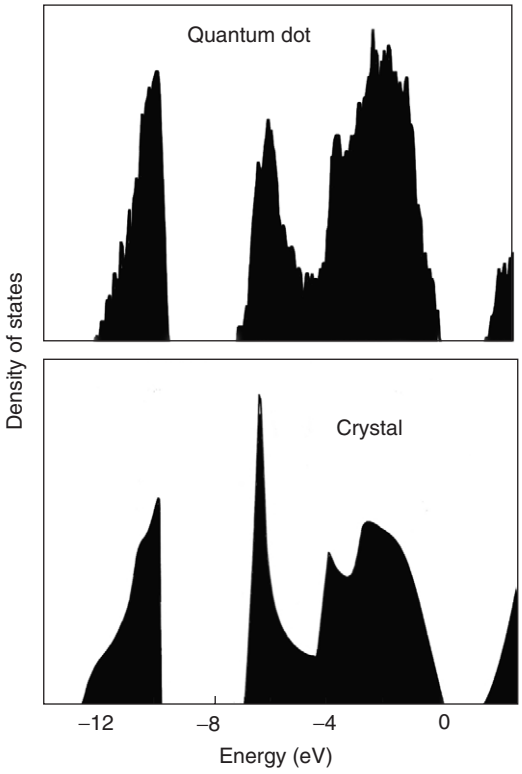


Figure 7 Densities of states for crystal gallium arsenide (bottom) and the eigenvalue distribution for $(GaAs)_{342}H^*_{192}$ quantum dot (top). Extrinsic contributions have been removed from the plot.

eigenvalues should approach the crystalline state. In **Figure 7**, the crystalline density of states is compared, that is, the number of states per unit energy, to the number of eigenvalues per unit of time for a large dot. The structure in the density of states can be attributed to the topology of energy bands. When the energy band is flat, that is, the derivative of the energy with respect to the wave function vanishes, the density of states possesses an identifiable structure. As a function of size, one would expect large clusters to possess similar structural features. The number of atoms in a dot required to reproduce bulk features of the density of states can be assessed through the direct calculation of the eigenvalue spectrum. In **Figure 7**, a comparison between the crystal and dot density of states for the dot, $(GaAs)_{342}H^*_{192}$ is illustrated. The distribution of eigenvalue states matches the bulk density of states very well with one notable exception. Some states are associated with the fictitious hydrogens; these states have been removed from **Figure 7**. More complex nanostructures can be examined using similar approaches.

See also: Density-Functional Theory; Nanostructures, Optical Properties of; Pseudopotential Method; Semiconductor Nanostructures.

PACS: 73.22. – f; 71.21.La; 71.15. – m

Further Reading

Alivisatos AP (1996) Semiconductor clusters, nanocrystals and quantum dots. *Science* 271: 933.
Beck TL (2000) Real-space mesh techniques in density functional theory. *Review of Modern Physics* 72: 1041.
Brus LE (1983) A simple model for the ionization potential, electron affinity and aqueous redox potentials of small semiconductor crystallites. *Journal of Chemical Physics* 79: 5566.
Chelikowsky JR (2000) The pseudopotential-density functional method applied to nanostructures. *Journal of Physics D* 33: R33.
Chelikowsky JR and Louie SG (eds.) (1996) *Quantum Theory of Real Materials*. Boston: Kluwer Press.
Cohen ML and Chelikowsky JR (1989) *Electronic Structure and Optical Properties of Semiconductors*, 2nd edn. New York: Springer.
Fong CY (1998) *Topics in Computational Materials Science*. River Edge: World Scientific.
Furukawa S and Miyasato T (1988) Quantum size effects on the optical band gap of microcrystalline Si: H. *Physical Review* 38: 5726.
Harrison WA (1999) *Elementary Electronic Structure*. River Edge: World Scientific.
Hohenberg P and Kohn W (1964) Inhomogeneous electron gas. *Physics Review* 136: B864.
Hummel R (1984) *Electronic Properties of Materials*. New York: Springer.
Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.
Kohn W and Sham L (1965) Self-consistent equations including exchange and correlation effects. *Physics Review* 140: A1133.
Yu P and Cardona M (1996) *Fundamentals of Semiconductors*. New York: Springer.

Nomenclature

a	box size (m)
e	electron charge (C)
E	energy (J)
E_c	correlation energy (J)
E_{gap}	optical gap (J)
E_n	energy eigenvalue (J)
E_x	exchange energy (J)
$\mathscr{E}_x$	exchange energy density (J m^{-3})
$\boldsymbol{G}$	reciprocal lattice vector (m^{-1})
h	grid spacing (m)
H	Hamiltonian operator (J)
H^{KS}	Kohn–Sham Hamiltonian (J)
$\hbar$	Planck's constant (J s)
$\boldsymbol{k}$	wave vector (m^{-1})
m	electron mass (kg)
$\mathscr{M}$	nuclear mass (kg)
p	momentum (kg m s^{-1})
$\boldsymbol{r}$	electronic position (m)
R	dot radius (m)
$\boldsymbol{R}$	nuclear position (m)
$\mathscr{R}$	lattice vector (m)
t	time (s)
V_H	Hatree potential (J)
V_p	ionic pseudopotential (J)
V_p	pseudopotential (J)
V_{xc}	exchange–correlation potential (J)
x	particle position (m)
Z	atomic number
ρ	electronic charge density (Cm^{-3})
ϕ	electronic orbital
Φ_j	localized orbital
Ψ	wave function

Nanostructures, Optical Properties of

K Cho, Osaka University, Osaka, Japan

Introduction

The reason for the optical properties of nanostructures to attract special attention is the explicit dependence of optical response on the size, shape, and internal structure of a matter, which leads to an additional possibility of controlling materials for various applicational purposes. In view of the importance of the coherence of matter excitations in the description of the optical processes in nanostructures, which has been neglected in the traditional theory for bulk systems, it is worth having a separate item for this subject.

The optical properties of matter are a consequence of the interaction between matter and light (oscillating electromagnetic (EM) field), which depends sensitively on the space–time structure of the EM field and the induced polarization of the matter. The factors that determine this dependence are the frequency or pulse structure, polarization, and propagation direction of light, and the amplitudes of the induced polarization of matter with various resonance frequencies corresponding to the quantum mechanical level structure of the matter.

The light–matter interaction is fully described by the equations of motion for the dynamical variables of light and matter. For variables, a usual choice is the electric field $E(r, t)$ and polarization $P(r, t)$, but there is another choice, that is, vector potential of light $A(r, t)$ and current density of matter $J(r, t)$. The

latter is slightly more general, because the current density may include the magnetic one (proportional to the rotation of magnetization). If this contribution is neglected, $J(r, t)$ and $P(r, t)$ are simply related as $\boldsymbol{J} = \partial \boldsymbol{P}/\partial t$, which applies to most cases of optical response of matter.

The scheme of equations of motion for A and J (or E and P) is common to quantum electrodynamics (QED) with a quantized EM field and the semiclassical treatments in terms of a classical EM field. The difference is whether the vector potential is treated as an operator or not. This difference does not always lead to a difference in the calculated physical quantities. In QED, the photons of each mode are described as a statistical ensemble, while in the semiclassical picture, the EM field of each mode is represented by a complex amplitude corresponding to an average value of the photon ensemble. The QED description is absolutely necessary when the physical quantity is related with some details of the photon ensemble, or when the optical processes involve the contribution of a noncommutative relation of the photon creation and annihilation operators. Typical examples are the squeezed light or various multitime correlation functions. However, there are many other quantities which can be correctly calculated semiclassically, that is, those depending just on the average photon number and phase (of each mode) arising from the optical processes without involving noncommutative terms of the photon operators. Examples are the light intensities arising from the linear and nonlinear processes without involving the same photon modes in the elementary steps of light–matter interaction.

Another important aspect in optical processes is the microscopic spatial structure in $E(r, t)$ and $P(r, t)$. This has been neglected in macroscopic response theories, and also in QED treating atoms and molecules. In such a treatment, the r-dependence of the EM field is assumed to be always much weaker than that of induced polarization, so that the latter is represented by its zeroth moment, that is, the electric dipole moment of the transition,

$$\int \mathrm{d}\boldsymbol{r} \, \langle \mu | \hat{\boldsymbol{P}}(\boldsymbol{r}) | v \rangle \qquad [1]$$

where μ, v are the quantum numbers of matter eigenstates. In other words, all the optical transitions are approximated as point-like electric dipoles with various eigenfrequencies and amplitudes in the long-wavelength approximation (LWA) for an EM field.

Nanostructures have attracted attention because of the size-quantized energy levels of matter. The microscopic spatial structure of induced polarization is exactly the counterpart of the energy quantization, so that it is quite consistent to consider these aspects simultaneously in the description of optical response in nanostructures. Every quantum state has its own resonant frequency and the characteristic spatial structure in its induced polarization, which depend sensitively on the size and shape of matter. Though this applies in principle to atomic and macroscopic systems also, it is only in nanostructures that one can easily observe the size θ and shape dependence of optical responses.

Optical properties of matter are usually described in terms of linear and nonlinear polarizabilities (electric susceptibility, or more simply, susceptibility). In macroscopic systems, they are considered to be intensive quantities, that is, independent of the size and/or shape of matter, which can be derived by the so-called Lorentz oscillator model based on classical mechanics (forced oscillation of harmonic oscillator with damping) under LWA. However, a quantum mechanical calculation of polarizabilities without the use of LWA gives an expression in terms of the energy eigenvalues $\{E_v\}$ and the matrix elements of the dipole density operator, a result different from that of the Lorentz model. For example, the linear susceptibility at absolute zero Kelvin is given, in units of ε_0, as

$$\chi^{(1)}_{\xi\eta}(\boldsymbol{r}, \boldsymbol{r}', \omega) = \sum_v \left\{ \frac{\langle 0 | \hat{P}_\xi(\boldsymbol{r}) | v \rangle \, \langle v | \hat{P}_\eta(\boldsymbol{r}') | 0 \rangle}{E_v - E_0 - \hbar\omega - \mathrm{i}0^+} + \frac{\langle 0 | \hat{P}_\eta(\boldsymbol{r}') | v \rangle \, \langle v | \hat{P}_\xi(\boldsymbol{r}) | 0 \rangle}{E_v - E_0 + \hbar\omega + \mathrm{i}0^+} \right\} \qquad [2]$$

where $v = 0$ stands for the ground state. The positive infinitesimal (0^+) represents the adiabatic switching of light–matter interaction. When the nonradiative damping of the matter excitations needs to be considered, the $\{0^+\}$s are replaced with the self-energies of the excitations caused by nonradiative scattering, and are often approximated by a phenomenological parameter.

This polarizability, as well as higher-order ones, is not an intensive quantity, but a position-dependent integral kernel with various resonances appearing as poles of frequency-dependent functions. It depends on the size and shape of matter through the resonant frequencies $(E_v - E_0)/\hbar$ and the coupling matrix elements of transitions, which is most conspicuously seen in the range of nanometers. One of the fundamental problems is the evolution of polarizabilities from size-dependent to size-independent regimes, which is due to the size dependence of (1) the contribution of single quantum states, and (2) the intervals of neighboring quantum states. In the case

of nonlinear polarizabilities, the size dependence is deeply related with the so-called cancellation problem.

The dependence of the resonance frequencies, that is, absorption and/or emission peaks, on the system size and shape is due to the boundary condition for the wave function of each excited state extending coherently over the whole structure. This is understood from an analogy of a particle of mass M confined in a one-dimensional well of size d. The boundary condition restricts the form of its wave function to $\sin(kx)$ and the energy eigenvalue $E = \hbar^2 k^2/2M$, with the size-quantized wave number $k = \text{integer} \times \pi/d$.

This allows one to design the optical properties of matter for various aims. Together with the technical development of nanosize fabrication, it has encouraged the studies of applicational aspects of nanostructures. At the same time, this has opened a new phase of fundamental character of matter, in particular of nanostructured materials, where the coherence of matter states plays an essential role. Due to the coherence of the excited state of matter maintained over the whole nanostructure, a full quantum mechanical treatment of polarizabilities is required. In other words, the microscopic character of the EM field and induced polarization need to be thoroughly considered, which inevitably leads to the nonlocal polarizability.

The nonlocal nature of optical response in the microscopic theory means that the cause (EM field) and the result (polarization) may occur at microscopically different spatial points, as far as the points are within the extension of the relevant wave functions. The nonlocality of response is a common feature to all excitations of a wave-like character, such as sound, light, and the quantum mechanical excitations described by wave functions.

It may seem to be a complication that polarizabilities become position-dependent integral kernels, in comparison with their constant values in the macroscopic approach. It is simply not the case, however, because the fact that they are, in general, separable kernels allows one to rewrite the (Maxwell) integral equations as simultaneous polynomial equations, which is certainly easier to handle. In this sense, nonlocal response is a key method to provide a good overview and technical ease, based on the first-principles formulation of light–matter interaction.

Scheme of Microscopic Response

For a given initial condition of light and matter, the optical response is obtained in the following way. The polarization induced by an EM field and the EM field produced by polarization are expressed, and then solved simultaneously. For nanostructures, it is necessary to do this in terms of the microscopic polarizabilities and without the use of the LWA. In view of the importance of resonant processes in nanostructures, which reflect the size and shape dependence sensitively, it is essential to solve the equations self-consistently (simultaneously).

The two sets of equations to be solved, in the case of, for instance, a linear response, are

$$P_\xi(\boldsymbol{r},\omega) = \varepsilon_0 \sum_\eta \int \mathrm{d}\boldsymbol{r}'\ \chi^{(1)}_{\xi\eta}(\boldsymbol{r},\boldsymbol{r}';\omega) E_\eta(\boldsymbol{r}';\omega) \quad [3]$$

and

$$E_\xi(\boldsymbol{r},\omega) = E^{(0)}_\xi(\boldsymbol{r},\omega) + \sum_\eta \int \mathrm{d}\boldsymbol{r}'\ G_{\xi\eta}(\boldsymbol{r},\boldsymbol{r}';\omega) P_\eta(\boldsymbol{r}';\omega) \quad [4]$$

The latter equation is the solution of the Maxwell equations for a given source polarization $\boldsymbol{P}(\boldsymbol{r},\omega)$ in terms of the radiation Green function $\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}';\omega)$ of a vacuum EM field

$$\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}';\omega) = \frac{4\pi}{V} \sum_{k\sigma} \hat{\boldsymbol{e}}_\sigma(\boldsymbol{k}) \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}}{k^2 - (q+\mathrm{i}\delta)^2} \hat{\boldsymbol{e}}_\sigma(\boldsymbol{k}) \quad [5]$$

where $q = \omega/c$ and $\hat{\boldsymbol{e}}_\sigma(\boldsymbol{k})$ $(\sigma = 1, 2)$ are two mutually orthogonal unit vectors perpendicular to k, and $\mathrm{V}\ (\to \infty)$ is the quantization volume. This Green function describes the propagation of light in vacuum. The initial condition of light (e.g., incident light) is represented by $E^{(0)}_\xi(\boldsymbol{r},\omega)$, and that of matter (usually, the ground state) is included in the susceptibilities. For a nonlinear response, the terms of nonlinear polarizations are to be added to the right-hand side of eqn [3]. The equations are a closed set for a given frequency ω in the case of a linear response, but they are not so in nonlinear cases.

The set of integral equations, [3] and [4], can be put into a set of simultaneous linear equations for the variables $X_{\mu 0}(\omega) = g_\mu(\omega) F_{\mu 0}(\omega)$, and $X_{0\mu}(\omega) = h_\mu(\omega)\ F_{0\mu}(\omega)$, where

$$g_\mu(\omega) = \frac{1}{E_\mu - E_0 - \hbar\omega - \mathrm{i}0^+} \qquad h_\mu(\omega) = \frac{1}{E_\mu - E_0 + \hbar\omega + \mathrm{i}0^+} \quad [6]$$

$$F_{\mu\nu}(\omega) = \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{E}(\boldsymbol{r},\omega) \cdot \langle \mu | \hat{\boldsymbol{P}}(\boldsymbol{r}) | \nu \rangle \quad [7]$$

The solution of such equations gives the polarization P as

$$\boldsymbol{P}(\boldsymbol{r},\omega)=\sum_{v}[X_{0v}\langle v|\hat{\boldsymbol{P}}(\boldsymbol{r})|0\rangle+X_{v0}\langle 0|\hat{\boldsymbol{P}}(\boldsymbol{r})|v\rangle] \quad [8]$$

The simplest version of the linear equations to determine $\{X\}$ is the one in the rotating wave approximation (RWA) and is given as

$$F^{(0)}_{\mu 0}(\omega)=\sum_{v}[(E_v-E_0-\hbar\omega-\mathrm{i}0^{+})\delta_{\mu v} + \mathscr{A}_{\mu 0,0v}(\omega)]X_{v0}(\omega) \quad [9]$$

where

$$F^{(0)}_{\mu 0}(\omega)=\int \mathrm{d}\boldsymbol{r}\,\boldsymbol{E}^{(0)}(\boldsymbol{r},\omega)\cdot\langle\mu|\hat{\boldsymbol{P}}(\boldsymbol{r})|0\rangle \quad [10]$$

is a known factor for a given incident light, and the radiative correction $\mathscr{A}$ is defined as

$$\mathscr{A}_{\mu 0,0v}(\omega)=-\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\langle\mu|\hat{\boldsymbol{P}}(\boldsymbol{r})|0\rangle \cdot\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\langle 0|\hat{\boldsymbol{P}}(\boldsymbol{r})|v\rangle \quad [11]$$

This is a complex quantity representing the interaction energy between two induced dipole densities $\langle 0|\hat{\boldsymbol{P}}(\boldsymbol{r})|\mu\rangle$ and $\langle 0|\hat{\boldsymbol{P}}(\boldsymbol{r})|v\rangle$ through the EM field of frequency ω. The real and imaginary parts of a diagonal term ($\mu=v$) represent the shift and width, respectively, of the matter transition energy $E_\mu-E_0$. The off-diagonal elements $A_{\mu 0,0v}$ represent the radiative interaction among different matter excitations, giving additional contribution to the radiative shift and width.

The solution of eqn [9] determines the polarization through eqn [8] (without the terms X_{0v} because of the RWA), which further determines the induced field through eqn [4]. In this way, the optical response is determined uniquely from the set of simultaneous linear equations. In the case of nonlinear response, eqn [3] contains nonlinear terms of the factors $F_{\mu v}(\omega)$ for various $\mu,\ v,$ and ω, and in this case the equations to be solved are a set of simultaneous polynomial equations of the order N, where N is the order of nonlinearity in consideration.

In this way of solution, there is obviously no requirement of the boundary conditions for an EM field at the surface/interface of matter. This is because the boundary conditions (of the matter system) are already taken into account in the expression of polarizability. An additional boundary condition (ABC) is not required when the matter excitations contain the effect of spatial dispersion. In fact, such a method is developed from an ABC-free formalism to solve confined exciton problems without using the ABC.

It should be stressed that the condition $\det|\boldsymbol{S}|=0$, where $\boldsymbol{S}=(E_v-E_0-\hbar\omega)\boldsymbol{1}+\mathscr{A}$, for the existence of nontrivial solution of eqns [9] in the absence of an external field (i.e., $F^{(0)}=0$), is exactly the resonance condition for X in the presence of the external field. The modes satisfying this condition are the eigenmodes of the coupled light–matter system. They are self-sustaining modes in the sense that they consist of finite amplitudes of E and P in the absence of external excitation. Thus, one understands that the resonance in the optical spectra appears at the frequency of a self-sustaining mode, that is, matter excitation energy corrected by a radiative shift and broadening. When there is translational symmetry in the matter system, the matrix S is diagonal with respect to the corresponding wave number, so that $\det|S|=0$ holds for each wave number separately, giving the dispersion relation $\omega=\omega(k)$. This is a most general dispersion equation for interacting light–matter systems.

The factor X_{v0} contains an energy denominator which diverges at a matter excitation energy E_v-E_0, while the factor X_{v0} itself is enhanced, not at E_v-E_0, but at the energy corrected by a radiative shift. One of the peculiar points of nanostructures is that the intervals of matter excitation energies and the radiative corrections, which are both sensitive to the size and shape of matter, are comparable in magnitude. Thus, the proper treatment of a radiative shift is essential in obtaining correct results of optical responses. An example of the size dependence of (complex) eigenmode energies with the radiative shift $\mathrm{Re}[\mathscr{A}]$ and width $\mathrm{Im}[\mathscr{A}]$ is shown in **Figure 1** for an exciton weakly confined in a dielectric sphere. The large resonant variation at a certain size for each mode is due to the radiative correction $\mathscr{A}$, which shows the central importance of this quantity in the calculation of the response spectrum.

In a macroscopic response theory, a resonance is ascribed to that of susceptibility, that is, the resonance with a matter excitation energy. However, in the microscopic treatment, the resonance is ascribed to the resonant enhancement of the internal electric field. Thus, the self-consistent treatment of response is seen to lead to a rather different picture of resonance.

Cavity Effect

Nonlocal Scheme in Terms of Cavity Green Function

Induced polarization consists of various frequency components corresponding to the transitions among

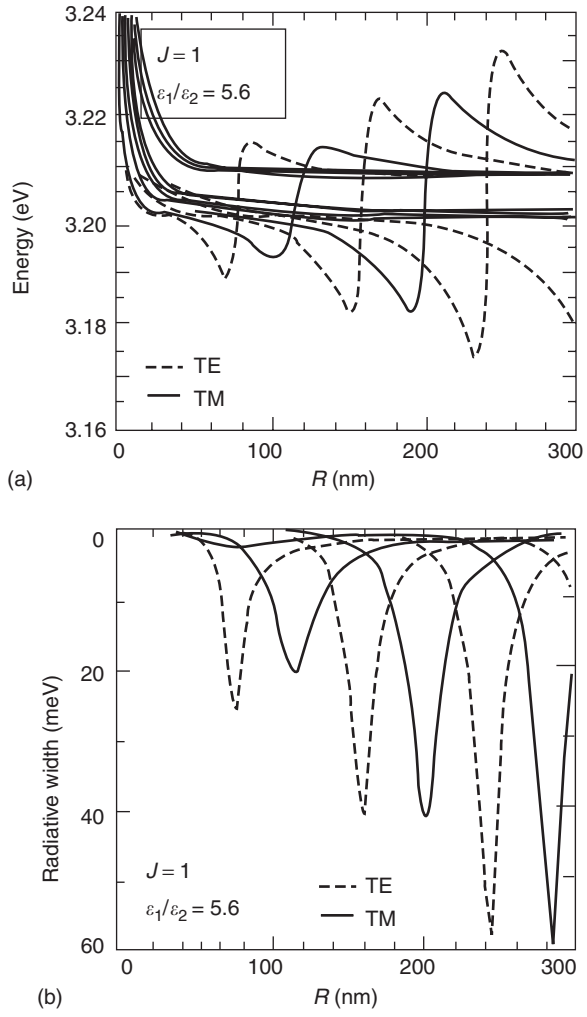


Figure 1 (a) Real and (b) imaginary parts of the eigenmode energies of an exciton in a sphere with background dielectric constant of $\varepsilon_b = 5.6$. The material parameters correspond to those of the Z_3 exciton in CuCl. Each curve is specified by total angular momentum J, its projection, TE or TM mode character, and radial quantum number for the exciton state. (from Cho K (2003) Optical Response of Nanostructures: Microscopic Nonlocal Theory. Berlin: Springer. Reproduced with permission of Dr. Hiroshi Ajiki.)

quantized levels of matter. For a given frequency of incident light, some of them are resonant and others are nonresonant. Since the number of transitions is infinite in general, it is not feasible to consider all the transitions in the above set of eqns [9]. To neglect all the nonresonant components is permissible only for optically thin systems, such as gases of atoms and molecules. For solids (and liquids, too), however, one needs to consider the effect of nonresonant (but infinitely many) components in dealing with the motions of resonant components. For appropriate configurations, these components produce "cavity modes" of an EM field, which have a completely different frequency spectrum and spatial structure from those of the vacuum modes. The interaction of cavity modes and resonant matter transitions offers very interesting problems of cavity QED, remarkable enhancement effects, and various typical situations showing the mechanism of light–matter interaction in a simplified manner.

A well-known recipe dealing with this kind of situation is to regard all the nonresonant components as an effective background medium of a given shape and size with electric susceptibility χ_b (background polarizability). Then, the Maxwell equation to determine the electric field $\boldsymbol{E}(\boldsymbol{r}, \omega)$ is written as

$$\nabla \times \nabla \times \boldsymbol{E} - q^2 \varepsilon_0 [1 + \chi_b \Theta(\boldsymbol{r})] \boldsymbol{E} = q^2 \boldsymbol{P}_{res} \quad [12]$$

where P_{res} is the resonant part of the polarization. Defining the EM Green function $\boldsymbol{G}_b(\boldsymbol{r}, \boldsymbol{r}'; \omega)$ of the background medium as

$$\nabla \times \nabla \times \boldsymbol{G}_b - q^2 \varepsilon_0 [1 + \chi_b \Theta(\boldsymbol{r})] \boldsymbol{G}_b = q^2 \delta(\boldsymbol{r} - \boldsymbol{r}') \quad [13]$$

where $\Theta(\boldsymbol{r}) = 1$ or 0 for r inside or outside the background medium, respectively, one can derive the solution of eqn [12] as

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}_b(\boldsymbol{r}, \omega) + \int d\boldsymbol{r}' \, \boldsymbol{G}_b(\boldsymbol{r}, \boldsymbol{r}'; \omega) \cdot \boldsymbol{P}_{res}(\boldsymbol{r}'; \omega) \quad [14]$$

where E_b is the field due to the background polarization alone. This field is equivalent to that of [4], expressed differently in terms of the Green function G_b.

In a nanostructure, the resonant polarization of interest will consist of a few frequency components, and is written in the form of eqn [8] with $v = 1, 2, \ldots, s$. The equations to determine the Xs are eqns [9] for $v = 1, 2, \ldots, s$, but the parameters in the equations should be changed as follows. (1) The excitation energies $\{E_v - E_0\}$ contain the interaction energy of induced polarizations, which are affected by the presence of background dielectrics. This "screening" effect can be divided into bulk-type and surface- or interface-type contributions. The former is the screened Coulomb interaction by the factor $1/\varepsilon_b$, where $\varepsilon_b = 1 + \chi_b$, and the latter is the image potential effect for the interaction among the components of induced charge density. (2) The evaluation of radiative interaction $\mathscr{A}$ should be made in terms of G_b instead of G. In terms of the modified excitation energies $\{E'_v - E_0\}$ and radiative correction $\mathscr{A}^{(b)}$, eqns [9] are solved for $v = 1, 2, \ldots, s$ to obtain the (modified) expansion coefficients X, which uniquely determine the response field and induced polarization.

A typical effect of cavity on the optical response of matter is the modification of the mode structure of an EM field and the characteristic spatial pattern of each mode. Through this change in the EM field, the coupling with matter excitations can be strongly affected, and their resonant frequencies and widths are modified rather strongly. In optical response, there appear the resonant contributions of both cavity modes and matter excitations in general, where one can distinguish two regimes, namely, the case where the cavity mode has a much smaller Q-value than that of the matter excitation, and the case where their Q-values are comparable. In the former case, only the effect on the matter excitation is appreciable as an increase in its width, while, in the latter, there occur new coupled modes through the mixing of the cavity mode and matter excitation.

The two regimes mentioned above can be described by the single scheme of the microscopic response in a cavity. The condition for the existence of a nontrivial solution of the modified equations for $\{X_{v0}\}$ in the absence of an external field leads to the eigenmodes of the matter in a cavity. The fact that this solution covers two different regimes can be seen from the structure of the eigenmode equation. The Green function G_b has poles at the energies of the cavity modes, so that the radiative correction $\mathscr{A}$ has the same poles with the strength determined by the coupling with the matter excitations. Thus, the eigenvalue equation

$$\det |(E'_v - E_0 - \hbar\omega)\mathbf{1} + \mathscr{A}^{(b)}| = 0 \qquad [15]$$

provides complex roots for the coupled cavity modes and matter excitations. Depending on the coupling strength in $\mathscr{A}^{(b)}$ and the imaginary parts of the cavity modes contained in G_b, two regimes appear as mentioned in the previous paragraph.

EM Green Function of Cavity Polariton

A typical example of matter in a cavity is a quantum well (QW) in a distributed bragg reflector (DBR). In this case, the coupled modes are called cavity polaritons, and the nonlinear response of cavity polaritons is a very popular subject of study, such as in pump–probe spectroscopy, four-wave mixing, parametric amplification, and so on. In these optical processes, probe beams are sent from outside the cavity, and the result of the nonlinear process inside the cavity is observed as an external light coming out of the cavity. A standard treatment of these processes is the so-called quasi-mode coupling scheme, where the cavity mode is assumed to be completely confined in the cavity, that is, to have an infinitely large Q-value. The well-defined cavity mode allows a simple and rigorous treatment of the light–matter coupling at various nonlinear levels, and the resulting signal light is calculated by assuming the coupling of the $Q = \infty$ cavity mode and the photon modes outside the cavity. The coupling strength is assumed to be such that it reproduces the radiative width of the cavity polaritons. This is used rather often, but it is a kind of make-shift approach. This cannot be justified in a rigorous sense, because the problem of the boundary condition at the cavity surfaces cannot be translated into that of interaction between spatially separated modes of the EM field.

An alternative way to handle such a problem rigorously is to use the Green function of the cavity polariton, which is defined by renormalizing the exciton linear polarization as well as the background polarization comprising the cavity mode, into the definition of the EM field causing the nonlinear optical interaction. The equation to define this Green function G_{cp} has a similar form as eqn [13]. The only difference is that it additionally contains, on the left-hand side, a term representing the linear polarization due to the exciton:

$$\nabla \times \nabla \times \boldsymbol{E} - q^2\varepsilon_0[1 + \chi_b\Theta(\boldsymbol{r})]\boldsymbol{E} + \varepsilon_0 \int d\boldsymbol{r}'\, \chi_x^{(1)}(\boldsymbol{r}, \boldsymbol{r}'; \omega)\boldsymbol{E}(\boldsymbol{r}') = q^2\boldsymbol{P}_{NL} \qquad [16]$$

Keeping the nonlocal character of $\chi_x^{(1)}$, one can solve this integro-differential equation, and obtain the analytical form of G_{cp} in terms of G_b alone. Since G_b is obtained in a simple form for a QW in a DBR cavity, G_{cp} can be used as a practical tool. Considering that G_{cp} describes the propagation of the EM field both inside and outside the cavity according to the detailed construction of the cavity, one understands the merit of this Green function to calculate the signal field outside the cavity caused by the nonlinear interaction in the cavity.

The EM field produced by the nonlinear polarization is given as

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}_{cp}(\boldsymbol{r}, \omega) + \int d\boldsymbol{r}\, \boldsymbol{G}_{cp}(\boldsymbol{r}, \boldsymbol{r}', \omega) \cdot \boldsymbol{P}_{NL}(\boldsymbol{r}', \omega) \qquad [17]$$

The main issue in this formalism is to give the explicit form of P_{NL} according to the problem in question. The nonlinear polarization contains several components of the EM field, so that one needs self-consistency conditions based on the above solution. It is noteworthy that the Green function for the cavity polariton is obtained in a closed form through the use of a nonlocal form of exciton linear polarizability. The use of local approximation for this polarizability

leads to a Bethe–Salpeter type equation, which is rather complicated to solve. Thus, a nonlocal description makes the solution much easier than the local one. This is another example of recommending the use of the nonlocal way of description.

See also: Interaction of Light and Matter; Optical Properties of Materials; Polarizabilities.

PACS: 78; 78.67. – n

Further Reading

Gaponenko SV (1998) *Optical Properties of Semiconductor Nanocrystals*. Cambridge: Cambridge University Press.

Halevi P (ed.) (1992) *Spatial Dispersion in Solids and Plasmas*, p. 339. Amsterdam: North-Holland.

Keller O (1996) Local fields in the electrodynamics of mesoscopic media. *Physical Report* 268: 85–262. Elsevier.

Pekar SI and Kocherga OD (1983) *Crystal Optics and Additional Light Waves* Benjamin/Cummings.

Stahl A and Balslev I (1987) *Electrodynamics of Semiconductor Band Edge*. Springer.

Takagahara T (ed.) (2003) *Quantum Coherence, Correlation and Decoherence in Semiconductor Nanostructures*. Amsterdam: Academic Press.

Nomenclature

$A(\boldsymbol{r}, \omega)$	vector potential at position r with frequency ω
$\mathscr{A}_{\mu 0,0\nu}(\omega)$	radiative interaction energy between the matter excitations $(0\rightarrow\mu)$ and $(0\rightarrow\nu)$ via the EM field of frequency ω
$\mathscr{A}^{(\mathrm{b})}$	radiative interaction energy between the components of induced polarization via the EM field of background dielectric
$\boldsymbol{E}(\boldsymbol{r}, \omega)$	electric field at position r with frequency ω
$\boldsymbol{E}^{(0)}$	solution of homogeneous differential equation, usually an incident field inducing polarization
$\boldsymbol{E}_{\mathrm{b}}$	electric field induced by the scattering of an incident field from background dielectric alone
$\boldsymbol{E}_{\nu}$	energy of the νth eigenstate of matter, $\nu = 0$ for the ground state
$\boldsymbol{E}'_{\nu}$	energy eigenvalue including the screening effect of the background dielectrics
$F_{\mu\nu}$	expansion coefficient of induced polarization
$\boldsymbol{G}$	Green function describing the propagation of an EM field
$\boldsymbol{G}_{\mathrm{b}}$	Green function describing the propagation of an EM field in the presence of a background dielectric
$\boldsymbol{G}_{\mathrm{cp}}$	Green function describing the propagation of an EM field of a cavity polariton
$\boldsymbol{J}(\boldsymbol{r}, \omega)$	current density at position r with frequency ω
$\boldsymbol{P}(\boldsymbol{r}, \omega)$	polarization at position r with frequency ω
$\hat{\boldsymbol{P}}(\boldsymbol{r})$	electric dipole density operator
$\boldsymbol{P}_{\mathrm{res}}$	resonant part of induced polarization
$\boldsymbol{P}_{\mathrm{NL}}$	nonlinear part of induced polarization
q	wave number of light in vacuum $(=\omega/c)$
Q-value	quality factor of a resonance, that is, resonant frequency divided by resonance width
$\boldsymbol{S}$	coefficient matrix of the linear equation to determine X
$X_{\mu\nu}$	another expansion coefficient of induced polarization
$\chi^{(1)}$	linear polarizability
χ_{b}	polarizability of a background dielectric
ε_0	dielectric constant of vacuum
ε_{b}	background dielectric constant
Θ	Heaviside function to describe the size and shape of a background dielectric
ξ, η	Cartesian coordinates

Neuroscience

E Boncinelli, Scientific Institute San Raffaele, Milan, Italy

Introduction

Neuroscience aims at studying, directly or indirectly, the structure and activity of the animal nervous system, including the brain. This system is responsible for receiving stimuli from the external world and producing internal states and ultimately movement in response to them, according to the scheme, perception → information processing → action. As is the case for any other biological system or apparatus, the nervous system is composed of cells, the basic elements of living beings. Many cell types are present in the nervous system but those functionally specific for it are termed "neurons." Neurons are highly specialized cells, primarily characterized by an abnormally high surface-to-volume ratio, even three orders of magnitude higher than that of other types of cells,

making them extremely well suited for exchanging information with their environment, cellular or otherwise. Neurons continually communicate with one another and are interconnected in circuits and supercircuits. The study of the nervous system requires knowledge about the nature and functions of neurons, about the way they exchange information and how this is analyzed, mostly in centralized structures, and ultimately utilized. All this is very strongly influenced by the history of its development and subsequent modifications.

General Overview of the Nervous System

From the nucleus-containing cell body, sometimes termed "soma," of a typical neuron emerge two types of protrusions: many thin and relatively short dendrites and, on the opposite side, a single thick and sometimes very long axon. Although all of them share some general features, it is appropriate to point out that neurons occur in an incredible number of different types. Neurons are strictly polarized – neural information can, in fact, flow through each one of them only in one direction, from dendrites to the cell body and from this to the axon – and tightly interconnected cells. Upstream of every neuron there is a neuron or, exceptionally, a sensory receptor, in contact with the environment, both external and internal to the body. Downstream from every neuron there is a neuron, or sometimes a muscle cell, deputed to generate movement, or internal anatomical structures able to release specific molecules in their own environment.

Every neuron on its own and the nervous system on the whole primarily work at the integrative level. On a small scale, neurons may be regarded as combined into neural circuits, some of which are clearly identifiable, whereas others may be traced only on a physiological basis. Most circuits are interconnected. On a broader scale, the entire nervous system may be regarded as composed of pathways and centers. Pathways are essentially one-dimensional sequential successions of neurons, whereas centers are three-dimensional assemblies of interconnected cell bodies. Pathways include nerves which are bundles of axons running from one location to another, often wrapped up in a myelin sheath. Pathways and centers are in turn combined into systems that serve specific functions, for example, the visual system or the auditory system. A more general distinction in the nervous system is that between the sensory and the motor systems.

From the anatomo-physiological point of view, the nervous system has been traditionally subdivided into central and peripheral components. The central nervous system comprises the brain (including the brain proper, or cerebrum, the cerebellum, and the connecting the brainstem) and the spinal cord. The peripheral nervous system includes sensory neurons, connecting the sensory receptors to the brain and spinal cord, as well as motor neurons, connecting the brain and spinal cord to muscles and glandular structures. The motor components of the peripheral nervous system are, in turn, assigned to a somatic division, which innervates skeletal muscles, and an autonomic division, innervating smooth muscles, cardiac muscles, and glands. Neurons or pathways carrying information toward the spinal cord and from this to the brain are termed "afferent," whereas those carrying information away from the brain to the spinal cord and from this to the periphery of the body are termed "efferent." A number of neurons, mostly but nonexclusively present in centers, reciprocally connect neurons and help in processing the nervous signal; they are generally termed "interneurons." Most of the processing work carried out in the brain is through interneurons.

Electrical Signals of Nerve Cells

Neural information is generated and propagated throughout the nervous system as electrical signals and neurons have evolved elaborate mechanisms for generating them, based on the flow of ions across their plasma membranes. Generated near the soma, electrical signals run along neuronal axons, which may be very long. Although capable of passively conducting electricity, axons are not good electrical conductors. In order to compensate for this deficiency, neurons have evolved a sort of a booster system allowing them to conduct electrical signals over great distances. The electrical signals produced by this booster system are called "action potentials," also referred to as "spikes" or "impulses." The production of an action potential is, of course, an energy-consuming process, but energy is also required for a neuron at rest in order to be ready to produce and conduct it. While resting, a living neuron continually generates a negative potential, termed the "resting membrane potential." The action potential abolishes the negative resting potential and makes the transmembrane potential transiently positive. Generation of both the resting potential and the action potential is best understood in terms of the nerve cell's selective permeability to different ions and the normal distribution of these ions across the cell membrane.

A difference in electrical potential is generated whenever ions flow across cell membranes. Such ion fluxes occur when the membranes are permeable

to one or more ion species, and when there is an electrochemical gradient that favors ion flow. The negative resting potential results from a net efflux of K^+ across neuronal membranes that are predominantly permeable to K^+. An action potential occurs when a transient rise in Na^+ permeability allows a net flow of Na^+ across a membrane that is now predominantly permeable to Na^+. The brief rise in Na^+ membrane permeability is followed by a secondary, transient rise in membrane K^+ permeability that repolarizes the neuronal membrane and produces a brief undershoot of the action potential. As a result of this series of molecular events, the membrane is transiently depolarized in an all-or-none fashion. When these active permeability changes subside, the membrane potential returns to its resting level because of the high resting membrane permeability to K^+. All this requires both selective membrane permeability and specific ion concentration gradients across the plasma membrane. The membrane proteins that ensure these two essential conditions are called "ion channels" and "pumps," respectively. These mechanisms appear to account for virtually all of the observed properties of all-or-none action potentials. By virtue of local current flow, action potentials can also be propagated along the length of neuronal axons, explaining how electrical signals are conveyed throughout the nervous system.

For a given axon, form, duration, and amplitude of the action potential are fixed. As a consequence, the intensity of a stimulus is encoded in the frequency of action potentials rather than in their amplitude. The velocity of signal propagation also differs from axon to axon, as the rate of the action potential propagation depends on the diameter and on the degree of electrical insulation. In this light, by acting as an electrical insulator, a myelin sheath surrounding an axon greatly speeds up its action potential conduction. Myelinated axons can conduct at velocities up to $150\,m\,s^{-1}$.

Synaptic Transmission

Communication among neurons is mediated by synapses, specific functional contact microstructures. Synapses transfer from one neuron to another (or, occasionally, to a muscle cell) the information carried by action potentials. Although there are many synaptic subtypes, they can be subdivided into two classes: electrical and chemical synapses. Electrical synapses, a minority, permit direct, passive flow of electrical current from one neuron to another. Most synapses are actually chemical and enable communication via the secretion of neurotransmitters. Chemical agents released by the presynaptic neuron produce secondary current flow in postsynaptic neurons by activating specific receptor molecules. Neurotransmitters are released from presynaptic terminals in quantal packets after the arrival of an action potential. The secretion of neurotransmitters is triggered by voltage-dependent Ca^{2+} channels, which elevate Ca^{2+} within the presynaptic terminal. The rise in Ca^{2+} concentration causes synaptic vesicles – presynaptic organelles that store neurotransmitters – to fuse with the plasma membrane and release their contents into the space, termed "synaptic cleft," between the pre- and postsynaptic cells.

The functional cycle of all neurotransmitter molecules is similar. They are synthesized and packaged into vesicles in the presynaptic cell; they are then released from the presynaptic cell and bound to receptors on one or more postsynaptic cells; once released into the synaptic cleft, they are rapidly removed or degraded. Many synapses release more than one type of neurotransmitter, and multiple transmitters are sometimes packaged in the same synaptic vesicle. The total number of neurotransmitters is not known, but is believed to be well over 100. Despite this diversity, these agents can be classified into two general categories: small-molecule neurotransmitters and neuropeptides. In general, small-molecule neurotransmitters mediate rapid reactions, whereas neuropeptides tend to modulate slower, ongoing neural functions. For this reason, neuropeptides, mostly present in the brain, are sometimes also termed neuromodulators. Glutamate is the major excitatory neurotransmitter in the brain, whereas GABA and glycine are the major inhibitory neurotransmitters.

Neurotransmitters evoke postsynaptic electrical responses by binding to members of a diverse group of proteins called neurotransmitter receptors, which, in turn, give rise to electrical signals by opening or closing ion channels in the postsynaptic membrane. The postsynaptic currents produced by the synchronous opening and closing of the ion channels change the membrane potential of the postsynaptic cell. Changes in potential that increase the probability of firing an action potential are excitatory, whereas those that decrease this probability are inhibitory. Whether the postsynaptic actions of a particular neurotransmitter are excitatory or inhibitory is determined by the class of ion channel affected by the transmitter. Because postsynaptic neurons are usually innervated by many different inputs, the integrated effect of all excitatory and inhibitory postsynaptic potentials produced in a postsynaptic cell at any moment determines whether or not the cell fires an action potential.

Neurotransmitter Receptors and Their Action

There are two major classes of receptors: those in which the receptor molecule is also an ion channel, and those in which the receptor and ion channel are separate molecules. The former are called "ionotropic receptors" or ligand-gated ion channels. They give rise to fast postsynaptic responses that typically last only a few milliseconds. The latter are called "metabotropic receptors," and they produce slower postsynaptic effects that may endure much longer. Metabotropic receptors affect postsynaptic ion channels indirectly by activating transducer molecules called G-proteins. These proteins may alter the properties of ion channels directly, or may activate intracellular second-messenger pathways that then modulate the channels.

Synapses may be distinguished according to the kind of processes that take part in them. Thus, a contact from an axon onto a cell body is termed an "axosomatic" synapse, whereas that onto a dendrite is termed "axodendritic" synapse. Similarly, a contact between two axons is termed "axoaxonic" and one between two dendrites a "dendrodendritic" synapse. Within the brain, a synapse seldom occurs in isolation; on the contrary, it is usually one of a number of very small varicosities present mostly on dendrites and termed "spines." In addition, it has to be pointed out that the synapse is not a simple mechanism for unidirectional transmission between neural cells. The complexity revealed by modern research has widened this view to include several new concepts. First, the presynaptic terminal may have receptors for its own released products; thus, the presynaptic terminal may be postsynaptic as well to its own transmitter. This can be seen as an expression of a general principle of regulation by feedback of the output. Second, the postsynaptic terminal may send retrograde signals to the presynaptic terminal. These may be rapid signals, as in the case of gaseous molecules such as NO, which mediate activity-dependent effects, or they may be slowly acting growth factors, which play roles in determining the types of neurotransmitters produced. By these actions, the synapse can be viewed as having at least a bidirectional nature. Third, the synapse is not a simple link between two neurons, but rather a complex organelle in its own right. Furthermore, it is unique among cell organelles in being constructed by more than one cell; it is, in fact, a multicellular organelle. This view allows one to appreciate the synapse as a building block of nervous circuits. It also helps one to understand how synaptogenesis, that is the establishment of new synaptic contacts (see below), is a main concern of neuronal development and evolution during life, and to what extent the synapse is adapted for information-processing functions in different nervous pathways.

Neural Signal Generation

Ultimately, the neural information is generated in sensory receptors that, through sensory components of the peripheral nervous system, provide information to the central nervous system about the internal and external environment. Here, there are specific receptor cells tuned to be sensitive to different forms of energy in the environment. These forms of energy serve as stimuli for the receptor cell that responds by generating a receptor potential, or generator potential, which ultimately leads to encoding of the response in an impulse discharge, sent to the higher centers in the brain. The receptor potential is graded smoothly and continuously in amplitude in relation to the intensity of the stimulus. This translates, in turn, within the cell, into a specific impulse frequency of the transmitted neural signals. Sensory reception thus involves the transformation from a continuously varying domain of sensory stimuli into a neural domain of all-or-nothing impulses.

In all of these processes the integrative properties of neurons prove essential, which are at the heart of the neural action and represent the essence of neural processes. Most neurons have an extensive somadendritic surface and a complex branching geometry, which provide multiple sites for local processing of synaptic and active properties. In this view, the neuron is not one node but many; it is itself an extensive computational system, equivalent in computer terms to an integrated multifunction chip. This rich complexity is, in turn, incorporated into a number of neural circuits.

Circuitry

At the first level, there is the arrangement of synapses at a local site on a cell body, dendrite, or axon terminal. These may involve simple convergence of several inputs onto that site, or simple divergence to several output sites. In addition, it may involve serial relays of information, or reciprocal interactions. In all these cases, there is a set of synaptic connections that acts as an integrative unit. These local patterns of connections have been termed microcircuits. It is very common for a particular type of microcircuit to be repeated throughout a given region or on a given cell type, thus acting as a module for a specific kind of information processing or memory storage. The microcircuits on a computer chip represent the same operating principle.

At a higher level of organization is the circuit that connects different neurons over longer distances

within a given region. This transmission may take place through a dendritic branch or dendritic trunk or it may take place through the axon of an interneuron, or the axon collateral of an output neuron. The key point about all of these pathways is that they remain within a given region. The term introduced for all these is "local circuits." The most restricted local circuits include microcircuits; the most extensive involve interlaminar and intraregional axonal connections. The function of these different types varies widely. Some provide for re-excitatory spread of activity from neuron to neuron; others provide for antagonistic interaction between neighboring integrative units within a region.

The next highest level of organization involves connections of one region with another; one may refer to this as a "projection circuit." Generally, a region receives an input from more than one other region and projects output to more than one other region. Thus, the same principles of convergence, divergence, and integration of different kinds of information operate at this level too. There are also feedback lines from one region to another. Note that feedback loops are present at all levels. The more local feedback loops can be regarded as nested within the more extensive projection feedback loops.

At a still higher level are sequences of connections through several regions. These are said to constitute a system. The function is normally to transmit information from the periphery into the central nervous system (as in a sensory system), or from the central nervous system to the periphery (as in a motor system). However, in any pathway there are often connections running in the opposite direction to provide for descending, ascending, or centrifugal control. Finally, at the highest level are sets of connections between a number of regions, which together mediate a behavior that involves, to some extent, the whole organism. These are called distributed systems. They are characteristic of higher functions of motor and sensory systems, and of many central systems.

Development and Plasticity

Nervous circuits and subtle connections are neither inborn nor established once and for all at any stage of life. On the contrary, the processes through which this working circuitry is laid down during individual development and continually modified throughout life are an integral part, if not the most characteristic part, of the study of the nervous system. Its function in fact, depends critically on its anatomy and microanatomy, and it is fundamental for its full functionality that the connectivity may be subtly adjusted throughout life. As a consequence, a major field of investigation in neuroscience is the description of the construction and modification of neural circuits. This is, by the way, a qualifying difference from human-designed artificial computing machines. Furthermore, any alteration in one of these processes is likely to cause major developmental defects or contribute to some complex genetic disorders of the nervous system. The initial development of the nervous system entails the birth of neurons, the formation of specific axonal pathways, the elaboration of vast numbers of synapses, and the emergence of increasingly complex behaviors. After birth, experiences during postnatal life shape behavioral repertoires and cognitive abilities in highly specific ways, typically within restricted temporal windows. Even in maturity, synaptic connections continue to be modified as new memories are laid down and older ones reshuffled or discarded.

Fortunately, a great deal has been learnt in the past decades about the events and principles of biological development of the animal body and, in particular, of the nervous system, and new data emerge every day. The underlying cellular and molecular mechanisms have been explored and a number of principles and developmental pathways of general validity have been discovered. Many of these appear to be evolutionarily conserved. For example, in taxa as different as vertebrates and insects, a key role in the regulation of these developmental events is played by the so-called regulatory genes, genes acting through the control of the expression of other genes placed hierarchically downstream from them and sometimes termed target genes. Regulatory genes generally code for transcription factors, that is, nuclear proteins able to recognize specific DNA sequences and modulate the level of expression of the corresponding genes.

The nervous system arises as a longitudinal tube in the dorsal midline of the embryo with a definite rostro-caudal as well as dorso-ventral polarity. Subsequently, the neural tube has to be properly segmented into the head-to-tail sequence of body regions and the appropriate cellular and functional areas have to emerge and differentiate. Uncovering the forces, and ultimately the genetic instructions behind these events has been one of the major triumphs of today's genetics and molecular biology. Within the neural tube, and to some extent away from it, neuronal development, that is neurogenesis, involves a series of steps. The first step is the proliferation of precursor cells through a sequence of cell divisions up to the birth of the neuron. After this final division, the cell is destined to become a neuron of a given type or limited range of types. Soon after, the cell begins to migrate from its site of origin to its final location and to differentiate toward its final form,

with extension of axon and dendrites and with growth in size. During this time, the expression of functional membrane properties and the formation of synapses begin. All of these mechanisms characteristically have well-defined time periods that overlap considerably. Maturation of the cell into its final form and function is a process that in many animals, especially humans, may last over considerable periods of time. In addition to production of cells and cell processes, pruning of the synaptic connections and cell death are important factors in early maturation as well as in old age. Many of these events take place in an activity-dependent fashion.

A general rule in the vertebrate nervous system is that neurons do not remain at their site of origin, but rather, migrate to their final position. This is a necessary consequence of the fact that the nervous system starts as a relatively thin tube within the embryo (the neural tube), and the final product is a much larger structure (the nervous system). In addition, the initial relations between the sites of origin of neurons may be very different from their final relations. Neurons of the peripheral nervous system, in particular, arise largely from the neural crest, so called because it is a clump of cells which first appears on the dorsal surface of the neural tube. During migration, neurons begin to elaborate the axons and dendrites that will eventually form neural pathways and networks. Axons, in particular, have to make long journeys to form their synaptic connections. At the progressing ends of the fibers, there are enlargements endowed with ameboid movements termed "growth cones." These movements enable them to push aside obstacles in their way until they reach their own destination. The study of the progression of growth cones has offered invaluable information about molecules and mechanisms favoring neuronal growth and those providing positional clues to direct the progression and pathfinding process. The essence of nervous organization is the establishment of synaptic circuits. Normal behavior requires precision in assembling these circuits, and many of the disorders of behavior that occur in humans and other animals are due to abnormalities in the development of the circuit connections. Many developmental factors contribute to keep this developmental program as near as possible to the programmed scheme.

In addition to all this, there are firm reasons to believe that at least part of learning and memory rely on the formation and reshuffling of synaptic connections in the brain, especially at the level of microwiring and establishment of dendritic spines within the cerebral cortex. In the human brain, there are not less than 10^{11} neural cells, each one possessing an average of 10^4 synaptic contacts. This makes a grand total of approximately 10^{15} synapses. Common wisdom at present maintains that the essence of our cerebral cortex unambiguously resides in the configuration of these 10^{15} synapses. The ability of neural circuits and synaptic connections to undergo major or minor changes throughout life is usually termed plasticity or neuronal plasticity.

See also: Electromagnetic Biological Effects; Elementary Excitations in Biological Systems.

PACS: 84.35.+i; 87.17.Nn; 87.19.La

Further Reading

Gilbert SF (2003) *Developmental Biology*, 7th edn. Sunderland, MA: Sinauer.

Kandel ER, Schwrtz JH, and Jessel TM (eds.) (2000) *Principles of Neural Science*, 4th edn. New York: McGraw-Hill.

Posner MI and Raichle ME (1994) *Images of Mind*. New York: Scientific American Library.

Purves D, Augustine GJ, Fitzpatrick D, Katz LC, LaMantia A-S, *et al*. (eds.) (1997) *Neuroscience*. Sunderland, MA: Sinauer.

Shepherd GM (1994) *Neurobiology*, 3rd edn. New York: Oxford University Press.

Neutron Sources

A Furrer, ETH Zürich, Villigen, Switzerland

Introduction

Neutron scattering is an important scientific and technological resource that provides essential information about the fundamental properties of condensed matter. The quality and precision of neutron scattering experiments are determined by the available neutron flux. The history of neutron sources has seen a tremendous increase of the neutron flux of fission reactors from $10^7\,\mathrm{n\,cm^{-2}\,s^{-1}}$ in 1942 to $10^{15}\,\mathrm{n\,cm^{-2}\,s^{-1}}$ in 1972; since then no further progress has been reported. Only very recently, some projects utilizing accelerator-based spallation sources

have been initiated and are expected to be realized within a few years in order to surpass the limit achieved in 1972.

This article starts with a summary of the unique character of neutron scattering experiments in the investigation of the static and dynamic properties of condensed matter. Then a survey of the historical evolution of neutron sources is presented. The basic principles of the most commonly used neutron sources based on nuclear fission and spallation reactions are discussed in greater detail. All neutron sources release neutrons in the MeV energy range; however, thermal neutrons in the MeV energy range are needed to perform scattering experiments in condensed matter research; thus, a section is devoted to the moderation of neutrons as well as to the devices that transport the thermal neutrons to the instrument positions. Finally, some prospects for future neutron sources are discussed.

Neutron Scattering: A Unique Experimental Tool for Condensed Matter Research

Modern materials research, together with the traditional scientific interest in understanding condensed matter at the atomic scale, requires a complete knowledge of the arrangement and the dynamics of the atoms and molecules, and of their magnetic properties as well. This information can be obtained by investigating the interactions of the material in question with various types of radiation, such as light, X-rays or synchrotron radiation, electrons, ions, and neutrons. Among all these tools, neutrons have outstanding properties, as simply stated in the 1994 physics Nobel prize award to Cliff Shull and Bertram Brockhouse: "Neutrons reveal where atoms are and what atoms do." The most relevant, unique character of thermal neutrons, which can hardly be matched by any other experimental technique, can be summarized as follows:

- The neutron interacts with the atomic nucleus, and not with the electrons as X-rays do. This has important consequences: the response of neutrons from light atoms (e.g., hydrogen, oxygen) is much higher than for X-rays; neutrons can easily distinguish atoms of comparable atomic number; and finally, neutrons easily distinguish isotopes which allows one, for example, by deuteration of specific parts of macromolecules (or biological substances) to focus on specific aspects of their atomic arrangement.
- For the same wavelength as hard X-rays, the neutron energy is much lower and comparable to the energy of elementary excitations in matter. Therefore, neutrons do not only allow the determination of the "static average" chemical structure, but also the investigation of the dynamic properties of atomic arrangements which are directly related to the physical properties of materials.
- By virtue of its neutrality, the neutron is rather weakly interacting with matter, which means that the properties of the investigated materials remain unchanged under neutron irradiation; furthermore, there is almost no radiation damage to living biological objects under study. Also, the rather weak interaction with matter results in a large penetration depth and therefore the bulk properties of matter can be studied. This is also important for the investigation of materials under extreme conditions such as very low and very high temperatures, high pressures, high magnetic, and electric fields, or several of these together; in such cases, the studied sample is always surrounded by numerous shields which make the use of X-rays difficult.
- The neutron carries a magnetic moment which makes it an excellent probe for the determination of the static and dynamical magnetic properties of matter (magnetic ordering phenomena, magnetic excitations, spin fluctuations).

The quality and precision of neutron scattering experiments are primarily determined by the counting rate and, therefore, by the flux of the available neutrons, normally quoted in units of neutrons per square centimeter per second ($n\,cm^{-2}\,s^{-1}$). Although about half of the world is made up of neutrons, they are tightly bound deep inside the atomic nucleus and quite difficult to set free. Facilities that accomplish this task in a manner useful for scientific and technological applications are called neutron sources.

Historical Evolution of Neutron Sources

For all neutron sources, nuclear reactions are dealt with. Some nuclear reactions with a practical potential are listed in **Table 1**. Neutrons were observed for the first time in 1932 by Chadwick who used the interaction of α-particles from decay of natural polonium with beryllium, and subsequently neutron sources based on natural α-radiation were the basis of the earliest neutron physics research. Nuclear fission reactions were the next generation of sources; however, they were initially built for research in the nuclear industry, and the consequent possibility of

Table 1 Neutron yields and deposited heat for some neutron-producing reactions

Reaction	*Energy/event*	*Yield* (n/event)	*Deposited heat* (MeV/n)
T(d,n)	0.2 MeV	8×10^{-5} n/d	2500
W(e,n)	35 MeV	1.7×10^{-2} n/e	2000
^{9}Be(d,n) 15 MeV		1.2×10^{-2} n/d	1200
^{235}U(n,f) fission		~1 n/fission	200
(T,d) fusion		~1 n/fusion	3
Pb spallation	1 GeV	~20 n/p	23
^{238}U spallation	1 GeV	~40 n/p	50

studying neutron scattering was simply an unforeseen spin off. The development started in 1942 with a thermal neutron flux of $10^{7}\,\mathrm{n\,cm^{-2}\,s^{-1}}$ (CP-1, USA) and had its climax in 1972 with the reactor at the Institute Laue-Langevin (ILL) in Grenoble, a dedicated neutron scattering facility with thermal neutron fluxes just exceeding $10^{15}\,\mathrm{n\,cm^{-2}\,s^{-1}}$.

Meanwhile, pulsed sources, starting with Alvarez's use of a radiofrequency (RF)-pulsed cyclotron, and in particular, electron accelerators that produce neutrons through Bremsstrahlung photoneutron reactions have been increasingly applied for slow-neutron research. In addition, repetitively pulsed reactors have been developed. However, all these facilities could not really compete, for instance, with the overall performance of the ILL reactor, except perhaps the pulsed reactor IBR at Dubna (Russia) which produces instantaneous thermal fluxes up to $10^{16}\,\mathrm{n\,cm^{-2}\,s^{-1}}$ in pulses several hundred microseconds long. Today the most intense pulsed neutron sources are based on proton accelerators and the production of neutrons by spallation. Currently, the worldwide leading facility is the spallation neutron source ISIS at Didcot (UK), which provides instantaneous thermal neutron fluxes over $10^{16}\,\mathrm{n\,cm^{-2}\,s^{-1}}$ with short pulse lengths $\sim 50\,\mu\mathrm{s}$. Next-generation pulsed neutron sources with fluxes $> 10^{17}\,\mathrm{n\,cm^{-2}\,s^{-1}}$ are currently under construction in USA and Japan.

Practical Requirements for Neutron Sources

The primary goal of a neutron source must be to release as many neutrons as possible in as small a volume as possible to achieve a high luminosity. Another important property of neutron sources is the heat deposition going along with the neutron release, since cooling problems are a limiting factor in practically all neutron source designs. In this respect, fusion is by far the most optimal process for neutron production as can be seen from **Table 1**, followed by spallation and fission. Fusion may be the technique of neutron production in the far future, but today the two most commonly employed reactions are thermal nuclear fission in ^{235}U and spallation by protons in the energy ~ 1 GeV, thus the discussion is restricted to the latter two reactions.

Fission and Spallation Neutron Sources

Fission Sources

Fission of the uranium isotope 235 by slow neutron capture has been the most frequently employed reaction in neutron sources till date. The reaction can be made self-sustaining because it is exothermal and releases more neutrons per fission process than are needed to initiate the process. If a slow neutron is captured by a fissionable nucleus, the resulting deformation can cause the nucleus to break into two fragments as visualized in **Figure 1**. Very often, a neutron is released directly during this process, but mostly the neutrons "evaporate" from the fragments. This is a very important feature, because a small fraction of these evaporation neutrons are released with a time delay of the order of seconds (up to minutes) and thus enable a critical arrangement to be run in a controlled fashion. The spectral distribution of the fission neutrons are well described by a Maxwellian

$$n(E) = \frac{2}{\pi^{1/2} \cdot E_{\mathrm{T}}^{3/2}} \sqrt{E} \cdot \exp\left\{-\frac{E}{E_{\mathrm{T}}}\right\} \qquad [1]$$

with a characteristic energy $E_{\mathrm{T}} = 1.29$ MeV. Fission reactors produce a continuous flux of neutrons (except IBR at Dubna which is designed as a pulsed reactor).

Spallation Sources

The term "spallation" is applied to a sequence of events that take place, if target nuclei are bombarded with particles (e.g., protons) of a de Broglie wavelength $\lambda = \sqrt{h^2/2mE}$, which is shorter than the linear dimension of the nucleus. In this case, collisions can take place with individual nuclides

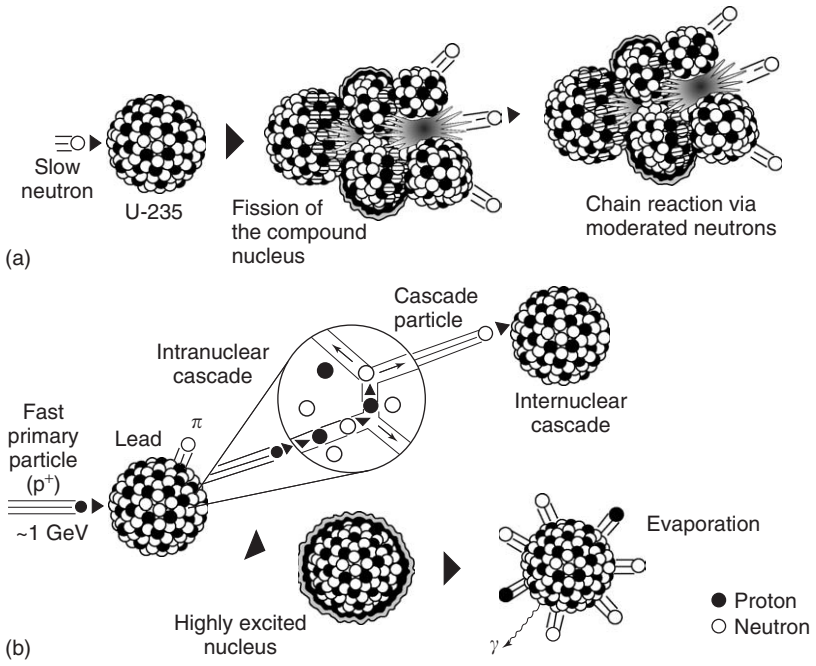


Figure 1 Schematic representation of (a) fission and (b) spallation.

inside the nucleus and large amounts of energy are transferred to the nuclides which, in turn, can hit other nuclides in the same nucleus. The net effect of this intranuclear cascade is twofold (see **Figure 1**): first, energy is more or less evenly distributed over the nucleus leaving it in a highly excited state; second, energetic particles may leave the nucleus and carry the cascade on to the other nuclei (internuclear cascade) or escape from the target. The excited nucleus left behind will start evaporating neutrons (and to a lesser extent protons). The low-energy part of the spectrum of these evaporation neutrons is quite similar to the one resulting from fission (eqn [1]), but as a consequence of the neutron escape during the intranuclear cascade, the spectrum extends to energies up to that of the incident particles (i.e., up to 1 GeV). The release of spallation neutrons takes place within less than 10^{-15} s after the nucleus was hit, so that the time distribution of spallation neutrons is exclusively determined by the time distribution of the driving particle pulse.

Modern spallation sources are generally based on a linear accelerator (linac). The process starts at the front end of the linac with the creation of negatively charged hydrogen ions by powerful ion sources. As they are electrically charged, these hydrogen ions can be accelerated in RF structures with strong fields along the linac to kinetic energies in the GeV range (which is $\sim$90% of the speed of light). When this highly energetic particle stream exits the linac, the hydrogen ions are stripped off their electrons by letting them pass through a thin carbon sieve, so that the hydrogen ions have become protons. The protons are then fed into a compressor ring which collects the protons from a large number of successive bunches fired out of the linac into a single very high-intensity proton pulse. For this purpose, an assembly of magnets bends each accelerated proton bunch into a circular orbit of such a diameter ($\sim$50–100 m) that the next bunch of protons arrives exactly when the previous has gone around once. In this way, all these bunches pile up. After $\sim$1000 revolutions, sufficient intensity is accumulated and the full proton pulse with a pulse length of $\sim$1 μs is extracted and propelled toward the target, which is normally a liquid metal (mercury or a lead–bismuth eutectic mixture) encased in special materials to take up a beam power of a few megawatts. This whole process – from the creation of the hydrogen ions to the arrival of the highly energetic protons at the target – should occur with a pulse repetition rate in the range of 10–100 Hz in order to achieve an optimal neutron economy in scattering experiments.

Neutron Energies and Neutron Properties

The energies of neutrons produced by neutron sources span many orders of magnitude. **Table 2** gives a summary of terms commonly used to characterize

Table 2 Approximate limits of neutron energy regimes classified by names

Energy range	*Classification*		*Energy range for neutron scattering*
	Nuclear physics	*Neutron scattering*	
		Ultra cold	<0.1 meV
		Very cold	0.1–0.5 meV
		Cold	0.5–5 meV
<1 keV	Slow	Thermal	5–100 meV
		Epithermal or hot	0.1–1 eV
		Resonant	1–100 eV
1 keV–0.5 MeV	Intermediate		
0.5–10 MeV	Fast		
10–50 MeV	Very fast		
50 MeV–10 GeV	High energy or ultra fast		
>10 GeV	Relativistic		

Table 3 Neutron properties and useful relations for thermal neutrons

Neutron properties	
Mass	$m = 1.675 \times 10^{-24}$ g
Charge	0
Spin quantum number	$s_N = 0.5$
Magnetic dipole moment	$\mu_N = -1.913\,\mu_K$ (μ_K = nuclear magneton)
Life time	$t_{1/2} \approx 7 \times 10^2$ s
Thermal neutrons	
Kinetic energy	$E = \frac{h^2}{2m\lambda^2} = \frac{\hbar^2 k^2}{2m} = \frac{mv^2}{2} = k_B T$
Conversion factors	$E(\text{meV}) = 81.81 \times \frac{1}{\lambda^2(\text{Å}^2)} = 2.072 \times k^2\left(\text{Å}^{-2}\right) = 5.227 \times v^2(\text{km s}^{-1})$
	$1\text{ meV} \equiv 0.242\text{ THz} \equiv 8.07\text{ cm}^{-1} \equiv 11.6\text{ K}$

different neutron energy regimes. **Table 3** summarizes the basic properties of the neutron as well as some useful relations regarding thermal neutrons.

Moderation of Neutrons

The energy spectrum of neutrons released from neutron sources is in the MeV range, whereas meV neutrons are required for scattering experiments in condensed matter research. Therefore, an energy shift of several orders of magnitude is necessary, which is accomplished by collisions with the atoms of a moderator substance. The goal in the layout of a moderator is to create the highest possible flux of moderated neutrons either in the shortest possible time (pulsed neutron sources) or in the largest possible volume (continuous neutron sources). This can be achieved by using moderators made of light atoms such as H_2O and D_2O. The time for slowing down the neutrons is $\sim 10^{-6}$ s after which the neutrons are in thermal equilibrium with the moderator kept at a constant temperature T according to the Maxwellian distribution

$$\Phi(\lambda) \propto \frac{1}{\lambda^3} \exp\left\{-\frac{h}{2k_B T m \lambda^2}\right\} \qquad [2]$$

where λ and m are the wavelength and the mass of the neutron, respectively. Moderators are generally kept at room temperature, and this is the reason why the corresponding neutrons are called thermal neutrons, with a maximum peak flux around the neutron wavelength $\lambda \approx 1$ Å, see **Figure 2**.

Scattering experiments which require cold or hot neutrons (i.e., neutrons with wavelengths considerably different from $\lambda \approx 1$ Å) would experience a tremendous flux penalty when working with the thermal neutron spectrum. However, the Maxwellian energy spectrum of the neutrons can be shifted by inserting either a cold source (e.g., a vessel containing D_2 at $T \geqslant 20$ K) or a hot source (e.g., a graphite block heated up to $T \leqslant 2000$ K) into the moderator; typical spectral shifts for cold and hot neutrons are displayed in **Figure 2**. By choosing the adequate neutron spectrum, the scattering experiments can be

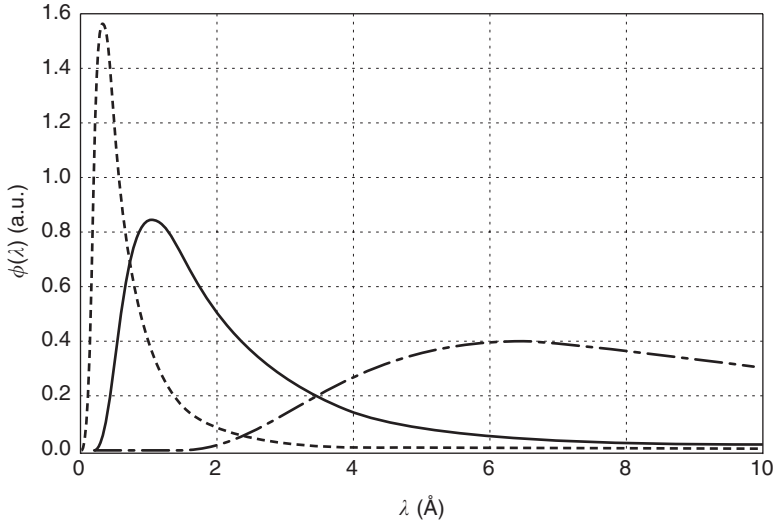


Figure 2 Wavelength distribution of neutrons from cold (dot dashed line, 50 K), thermal (solid line, 300 K), and hot (dashed line, 1000 K) moderators.

optimally tailored to the particular experimental requirements.

Neutron Beam Transport Devices

Once neutrons have been produced and moderated in the neutron source, it is still necessary to transport the neutrons to the experimental positions. Since neutron sources radiate isotropically, they must be surrounded by a heavy biological shielding. Beam tubes are inserted into the shielding to transport the neutron beam in a line-of-sight from the moderator surface to the instruments for neutron scattering. It is essential that these beam tubes are arranged tangentially to the core of the neutron source in order to reduce unwanted radiation significantly (e.g., fast neutrons). The angular acceptance of a neutron beam tube is determined strictly by the line-of-sight geometry between the source and the instrument. The beam divergence of a beam tube is normally $\sim 1^\circ$, so that the neutron flux available at the instrument position is drastically reduced by about six orders of magnitude as compared to the core flux.

The flux at the instrument position can be considerably improved by using neutron guides. A neutron guide works via total reflection of neutrons from the smooth walls of the guide material. This occurs for scattering angles less than the critical angle θ_c which is given by

$$\theta_c = 2\lambda\sqrt{\frac{\rho b}{2\pi}} \qquad [3]$$

where ρ and b are the atom number density and the coherent scattering length of the wall material, respectively. Among common materials, nickel is the best choice with critical angles $\theta_c(^\circ) = \lambda \cdot 10^{-1}$ (Å). The angular acceptance of a neutron guide can be dramatically increased by reflection from the so-called supermirrors which consist of a sequence of layers of variable thickness with alternatingly high positive and negative scattering length density. Supermirrors can be characterized by a number m defining the increase of θ_c compared to nickel. At present, $m = 3$ can be routinely achieved which results in an order of magnitude flux increase at the instrument position as compared to a conventional beam tube.

Prospects for Future Neutron Sources

An ever-growing scientific community uses neutrons in various disciplines of condensed matter research; however, the currently active neutron sources (see **Table 4**) will not be able to supply the future demand. In fact, some time between 2010 and 2020, the currently installed capacity of neutron sources serving the scientific community will decrease to a level below one-third that of today. Thus new neutron sources are indispensable to ensure the continuity and further developments in neutron scattering.

The history of the evolution of fission sources was impressive some decades ago, but progress was leveling off after the commissioning of the high-flux reactor at the ILL, Grenoble. This is clearly due to the technical difficulty of removing the heat from the reactor core. Another problem is the worldwide consciousness about the risk of installations based on nuclear fission. This makes it difficult to surpass the performance of existing high-flux reactors, so that one may conclude that fission sources do not have a serious prospect for the future.

Table 4 Currently active neutron sources offering access to the scientific community for condensed matter research (>100 users/year). The flux of pulsed sources corresponds to the peak flux

Location	*Country*	*Source*	*Type*	*Power* (MW)	*Thermal flux* (10^{14} n cm^{-2} s^{-1})
Lukas Heights	Australia	HIFAR	Fission/continuous	10	1.4
Chalk River	Canada	NRU	Fission/continuous	120	3.0
Grenoble	France	HFR	Fission/continuous	58	12.0
Saclay	France	Orphée	Fission/continuous	14	3.0
Berlin	Germany	BER-2	Fission/continuous	10	2.0
Geesthacht	Germany	FRG	Fission/continuous	5	0.8
Jülich	Germany	FRJ-2	Fission/continuous	23	2.0
Munich	Germany	FRMII	Fission/continuous	20	7
Budapest	Hungary	BNC	Fission/continuous	10	1.6
Tokai	Japan	JRR-3	Fission/continuous	20	2.0
Tsukuba	Japan	KENS	Spallation/pulsed	0.003	3
Dubna	Russia	IBR2	Fission/pulsed	2	100
Studsvik	Sweden	R-2	Fission/continuous	50	1.0
Villigen	Switzerland	SINQ	Spallation/continuous	1	2.0
Didcot	UK	ISIS	Spallation/pulsed	0.16	100
Oak Ridge	USA	HFIR	Fission/continuous	85	12.0
Gaithersburg	USA	NBSR	Fission/continuous	20	2.0
Los Alamos	USA	LANSCE	Spallation/pulsed	0.056	34
Argonne	USA	IPNS	Spallation/pulsed	0.007	5

However, future progress can be expected for spallation sources which – unlike fission sources – offer the possibility of imposing a time structure on the neutron flux, so that the source performance depends on the peak flux in the pulse rather than on the time-averaged flux. The present spallation sources and spallation source projects are largely designed such as to create rather short neutron pulses $\sim 100\,\mu s$. A further increase of the power level beyond 1–2 MW (the envisaged power level for the spallation source projects in USA and Japan) may be difficult due to target problems. Nevertheless, an order of magnitude increase of the power level seems to be feasible for increased proton pulse lengths of the order of milliseconds duration, as obtainable from a linac without pulse compression. The European 5 MW spallation source project was partly based on such a concept, and it was demonstrated that significant advantages can be drawn from long proton pulses for certain categories of neutron scattering instruments. Although target problems become increasingly severe as the power goes up, it is probably realistic to see the ultimate limit of (long-pulse) spallation sources approaching a proton power level of 20 MW (i.e., 20 mA proton current at 1 GeV).

See also: Electron and Positron Sources; Neutrons and Neutron Scattering, History of; Nuclear Fission and Fusion.

PACS: 14.20.Dh; 25.40.Sc; 28.50.Dr; 61.12.Ex

Further Reading

Bauer GR (1993) Neutron sources. In: Furrer A (ed.) *Neutron Scattering*, pp. 331–357. PSI-Proceedings No. 93-01, ISSN 1019-6447, Paul Scherrer Institute, Villigen.

Carpenter JM and Yelon WB (1986) Neutron sources. In: Sköld K and Price DL (eds.) *Methods of Experimental Physics*, vol. 23A, pp. 99–196. London: Academic Press.

Chadwick J (1932) The existence of a neutron. *Proceedings of the Royal Society* A136: 692–708.

Clausen K (2001) *Neutron sources*. Office for Official Publications of the European Communities, Luxembourg, ISBN 92-894-0037-4.

Richter D and Springer T (1998) *A Twenty Years Forward Look at Neutron Scattering Facilities in the OECD Countries and Russia*. European Science Foundation, Strasbourg, ISBN 2-912049-03-2.

Nomenclature

E	neutron energy (MeV)
h	Planck's constant ($h = 6.626 \times 10^{-27}$ erg s)
k	neutron wave number (Å^{-1})
k_B	Boltzmann constant ($k_B = 1.381 \times 10^{-16}$ erg K^{-1})
m	mass of the neutron ($m = 1.675 \times 10^{-24}$ g)
λ	neutron wavelength (Å)
v	neutron velocity (km s^{-1})
Φ	neutron flux (n cm^{-2} s^{-1})

Neutrons and Neutron Scattering, History of

F Mezei, Hahn-Meitner-Institut, Berlin, Germany

Introduction

The expression "neutron scattering" commonly refers to the study of condensed or gaseous matter by observing the scattering of a neutron beam from a sample of the material in question. If the energy of the neutrons in the beam is much higher than the typical energy of the interactions of the atoms in the sample matter (e.g., the energy of chemical binding within constituent molecules), the collision of neutrons with the nuclei of the atoms depends only on the properties of the individual nuclei and carries no information on the properties of the material. Thus, the subject matter of this article is the history of using low-energy (typically less then 1 eV) neutron radiation for exploring what happens inside common materials.

The idea that neutrons can reveal precious information about the structure of such materials – solids, liquids, and to some extent also gases – was born shortly after neutrons were discovered by Chadwick in 1932. Practical neutron scattering started as a parasitic activity on research reactors built in the 1940s for developing nuclear energy. Since the 1960s, several facilities around the world have been specifically designed, built, and dedicated primarily for neutron scattering research to serve by the end of the twentieth century a worldwide community of about 7000 users of neutrons as one of their tools in the study of a vast variety of topics in condensed matter research, including physics, chemistry, biology, geology, archeology, and engineering. In 1994, Shull and Brockhouse, the two outstanding pioneers of neutron scattering in the 1950s, were awarded the Nobel prize in physics for discovering how the neutrons show "where the atoms are and what they do."

Establishing the Conceptual Foundations

As early as in 1936, Elsasser suggested that neutrons move as waves through matter and as these waves scatter on the nuclei of individual atoms, interference may occur between the partial waves scattered from different atoms, as was known since the 1910s for X-rays. This was demonstrated in the same year by observing the strong reflection of neutrons on large magnesium oxide single crystals. In this respect, the fundamental difference that neutrons scatter on the nuclei and X-rays on the electron cloud surrounding these nuclei in the atoms was of no principal consequence.

In 1936, Bloch also made a visionary proposition, which in contrast had no analogy in X-rays. Due to the fact that the neutron possesses a magnetic moment, Bloch realized that they could also feel the magnetic fields inside magnetic materials. The magnitude of the neutron magnetic moment is $\mu_n = 1.04 \times 10^{-3}\,\mu_B$, where the Bohr magneton μ_B essentially equals the magnetic moment of a free electron. The internal average magnetic field (induction) inside ferromagnetic iron is $B = 2.2\,\mathrm{T}$ and the energy of the Zeemann interaction in this field $E = -\mu_n B = 1.32 \times 10^{-7}\,\mathrm{eV}$ acts as a potential on the neutron (attractive if the neutron magnetic moment is parallel to the field B, and repulsive for antiparallel direction), which can measurably deflect a neutron beam. Bloch pointed out that, since the internal magnetic field distribution in a magnetic material is not homogeneous but shows strong maxima centered around the position of the nuclei, the internal magnetic field distribution shows a periodic structure in crystalline magnets and therefore with neutrons, not only the atomic structure determines the Bragg diffraction peaks well known for X-rays, but there is also a contribution from the magnetism of the atoms.

It is of particular historical interest that while Bloch's visionary recognition of the great potential of neutron scattering in the study of magnetic materials proved to be of outstanding importance in the investigation of magnetic phenomena, for example, the study of high-temperature superconductors in recent decades, he proposed a mathematical expression, which was debated for nearly 15 years and finally proved to be wrong. This controversy reflected a fundamental ambiguity of electromagnetic theory, which could finally be settled by the neutron reflectivity experiments of Hughes and Burgy (1951). In electromagnetic theory, two fields are associated with magnetic materials: the so-called magnetic field H and magnetic induction B. These two are equal outside the magnetized matter, the only place where one could place a measuring device. This was going to change with the neutrons: as neutral particles, neutrons interact weakly with most materials and can penetrate deep inside heavy objects (e.g., many centimeters in steel), and as a quantum mechanical particle wave, they can theoretically spread over an unlimited volume of the sample. So magnetism could, for the first time, be observed inside magnetic materials, with the

measuring probe, the neutron occupying the same space as the sample magnetic material. Although the details did not appear to be so clear at that time, Bloch's 1936 theory implied that the interaction between the neutron magnetic moment and magnetism corresponds everywhere, including the inside of matter, to the energy

$$E = -\mu_n H \qquad [1]$$

whereas the different formulas proposed by Schwinger in 1937 were shown to correspond to the general assumption

$$E = -\mu_n B \qquad [2]$$

The experiments of Hughes and Burgy, which could only be done when a few years after World War II intense neutron beams became available for research at nuclear reactors, settled the ambiguity in favor of eqn [2]. To overcome a repulsive energy barrier of ~ 100 meV corresponding to the magnetic induction B inside iron for the neutrons with magnetic moment opposite to the magnetization of the sample, the neutrons needed to have at least $5\,\mathrm{m\,s^{-1}}$ velocity perpendicular to the surface. Below this value, the neutron beam is totally reflected. Thus for typical velocities of the neutron beams available from reactors in the range of $2000\,\mathrm{m\,s^{-1}}$, the range of total reflection will extend up to grazing incident angles of $\sim 0.1^\circ$ (the so-called critical angle). Hughes and Burgy determined the potential energy by the observation of the critical angle in magnetized Fe films. Their conclusion that $E = -\mu_n B$ was shown to imply that the magnetic moment of the neutron should be considered as a microscopic Amperian current loop, and not as a dipole built by opposing charges.

Actually, the total potential energy at a surface is the sum of the magnetic and nuclear contributions, where the latter designates the average potential presented by the interaction between the neutrons and the nuclei of the atoms in the sample, and the detailed shape of the reflectivity curve above the critical angle of total reflection depends on the detailed profile of the potential, which can show considerable structure, for example, in surfaces consisting of a number of thin layers. Decades after the pioneering work of Hughes and Burgy, neutron reflectivity became a widely-used tool for the study of surfaces and thin film structures, in particular following Felcher's important study of the penetration depth of magnetic fields at the surface of superconductors in 1979.

Neutron Diffraction

Practical neutron scattering work started more than 10 years after the theoretical suggestions by Elsasser and Bloch, when nuclear fission reactors became available for research. The energy of fission neutrons is very high ($\sim$MeV), and they first need to be slowed down before they can be used in neutron scattering work. This was achieved by the thermalization process, in which neutrons are made to collide randomly with light, small absorbing nuclei (originally C in graphite) whereby after a few collisions a neutron gas with a temperature close to the temperature of the moderator is produced. The energy of the neutron emerging from a moderator

$$E = mv^2/2 \qquad [3]$$

(where m is the neutron mass, $m = 1.67 \times 10^{-24}$ g) is on average close to the thermal energy $k_B T$ at the temperature of the moderator, that is, 25 meV at room temperature. The corresponding neutron velocity $v = 2200\,\mathrm{m\,s^{-1}}$ is equivalent in quantum wave mechanics to a wavelength of $\lambda = 0.18$ nm, as given by the de Broglie relation

$$h/\lambda = mv \qquad [4]$$

where h is Planck's constant. This wavelength is comparable to the 0.154 nm of the Kα X-ray radiation of copper, mostly used for crystallographic studies at that time. However, the wavelength distribution in the neutron beam extracted from the moderator was very broad and to build the first neutron diffractometer in 1948, Shull and Wollan at Oak Ridge used a monochromator single crystal to obtain a neutron beam with a well-defined wavelength. The next step was to measure the neutron scattering power for a large number of atomic nuclei by studying the neutron diffraction pattern for many well-known crystals. It turned out that this power changes rather randomly from one nucleus to the other, and, in particular, in neutron diffraction one can see the contributions of light elements (e.g., hydrogen) also in the presence of much heavier ones. This is an important contrast of great practical utility to X-ray diffraction, where heavy elements in the crystals tend to mask the lighter ones.

The first decisive discovery by using neutron scattering was made by Shull and Smart in 1949 at Oak Ridge: they found a new Bragg peak in the neutron diffraction pattern of MnO at low temperatures, which could not be seen in X-ray studies (**Figure 1**). It was the first experimental evidence for antiferromagnetism, showing that the magnetic moments of the Mn atoms are aligned parallel for every second

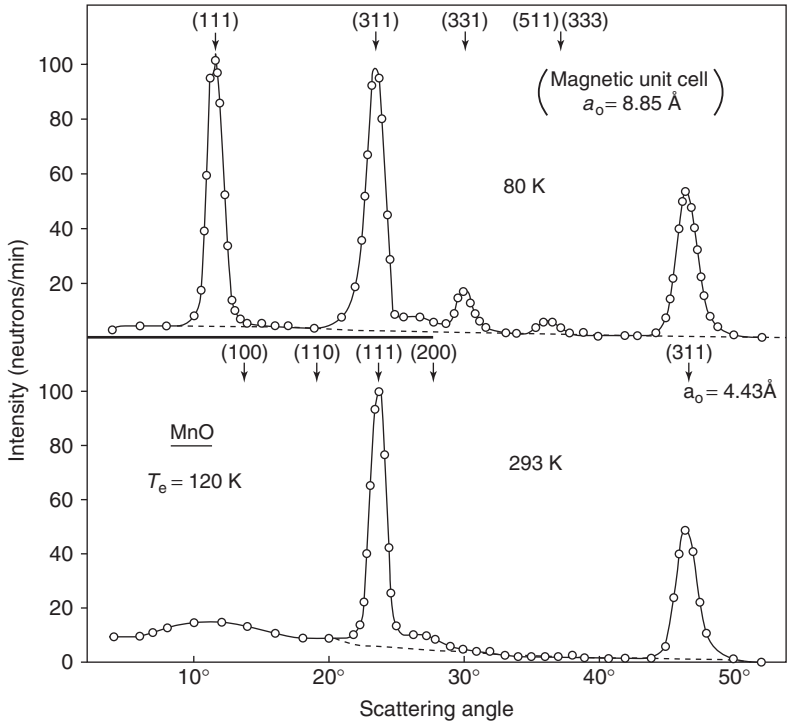


Figure 1 The first experimental evidence for antiferromagnetism by neutron diffraction. (Shull CG and Smart JS (1949) *Physical Review* 76: 1256.)

atom in certain lines and antiparallel for the atoms in-between. This breakthrough established neutron scattering as the preferred tool for the study of magnetism on the atomic scale. It is worth noting that this seminal discovery was made before the controversy on the coupling between neutron and magnetism had been settled. However, the existence of Bragg peaks in a sample without preferred magnetic orientation did not depend on the exact form of the interaction.

Neutron diffraction experiments can not only reveal crystalline long-range order, but also much less regular arrangements of atoms. In such cases the Bragg peaks are replaced by broad "diffuse scattering" structures in the scattering diagram (such as in the lower curve in **Figure 1** around 11° scattering angle), which carry information on a large variety of phenomena, such as the way in which one atom is surrounded by neighbors in a liquid or the size and arrangement of precipitates in an alloy. Information on larger structures shows up at smaller scattering angles, which gradually led (from the late 1950s) to the development of small angle neutron scattering (SANS) as a very special form of neutron diffraction.

Inelastic Scattering

The atoms and molecules inside condensed matter move around locally with energies comparable to k_BT. This leads to the appearance of the so-called inelastic scattering processes, in which the energy of the scattered particle is changed by an exchange with the kinetic energy of the atoms. Since for ordinary temperatures, the energy of the thermal motion of atoms in the sample is comparable to the energy of the neutrons from ambient moderators, the energy change of the neutron in the corresponding inelastic scattering processes can be quite substantial, and therefore readily measurable. In contrast, the quanta of X-rays with wavelength convenient for scattering work on condensed matter (i.e., in the range of atomic distances in the sample) are very high, and inelastic processes due to motion of atoms are difficult to observe. For example, the energy of the above-mentioned Kα X-ray radiation is 7.8 keV, ~300 000 times higher than energy of neutrons with similar wavelength. Since the early 1950s, several research groups have demonstrated the existence of such inelastic processes in neutron scattering. The decisive breakthrough came from Brockhouse at Chalk River by the introduction of triple-axis-spectroscopy in 1958, a technique to put these studies on firm grounds adequate for theoretical interpretation.

In 1954 van Hove developed a general theoretical framework for neutron (and other) scattering experiments in condensed matter. The particular significance of the van Hove theory for neutrons was that since the neutron interaction with most condensed

matter sample is weak (again a very important difference from X-rays), this theory applies with great precision in most cases. The theory allows us to calculate the probability for the neutron to scatter from an initial state with velocity $\boldsymbol{v}_i$ to a final state with velocity $\boldsymbol{v}_f$ for any complete model of the behavior of the sample, including both the arrangement of the atoms in space and their motion, for example, vibrations around steady equilibrium positions in solids, or the microscopic motion leading to fluidity in liquids. The van Hove scattering function $S(\boldsymbol{q}, E)$ is a function of two parameters, the momentum transfer vector $\boldsymbol{q}$ and the energy transfer E, which are defined in terms of the initial and final neutron velocities as follows:

$$\boldsymbol{q} = m\boldsymbol{v}_f/\hbar - m\boldsymbol{v}_i/\hbar \quad \text{and} \quad E = mv_f^2/2 - mv_i^2/2 \quad [5]$$

where $\hbar = h/2\pi$ and v_f and v_i are the absolute values of the velocity vectors $\boldsymbol{v}_f$ and $\boldsymbol{v}_i$. The triple-axis technique of Brockhouse allows us to select a well-defined pair of values $\boldsymbol{q}$ and E. This was a real breakthrough, since the computing power in theoretical calculations was rather limited at that time and model calculations such as the energy dispersion relation of a phonon or magnon excitation branch α in single crystals, $E_\alpha(\boldsymbol{q})$, could only be calculated for special momentum transfer vectors $\boldsymbol{q}$ in highly symmetric directions in the crystals.

The principle of triple-axis spectroscopy is very simple (**Figure 2**): both the initial and final neutron velocities can be selected with great flexibility, by a monochromator and an analyzer crystal, in order to fulfill both the equations in [5] for a wide range of desired values of the four parameters represented by the vector $\boldsymbol{q}$ and the scalar E. This requires calculating the three scattering angles for the monochromator, sample and analyzer, and automatically setting the spectrometer and the orientation of these three crystals in the right configuration. This was the key technical innovation behind the introduction of triple-axis spectroscopy. The most common mode of operation is to design scans of one of the abovementioned four parameters, for example, in "constant $\boldsymbol{q}$" scans a set of data points are taken at different E values, while the vector $\boldsymbol{q}$ is kept constant with respect to the sample. The diffractometer, or "two-axis" spectrometer, used by Shull and collaborators differs from the one in **Figure 2** only by the lack of the analyzer crystal: the detector only records the distribution of the scattered neutrons as a function of the scattering angle, without analyzing the magnitude of their velocity. This is done under the tacit and frequently valid assumption that the inelastic scattering is weak compared to the elastic one with $v_f = v_i$, and therefore it can be neglected.

An alternative approach to inelastic scattering is to produce pulses of neutrons with a constant initial velocity $\boldsymbol{v}_i$ and determine the final velocity vector $\boldsymbol{v}_f$ by recording the time of arrival of the neutrons at the immobile detector banks surrounding the sample, where the position of each detector corresponds to a given scattering angle. A 2–4 m flight path between the sample and the detectors is sufficient to determine the final neutron velocities with a reasonable precision. This "time-of-flight" (TOF) method was developed at several places in incremental steps parallel to the triple-axis technique, including early work with

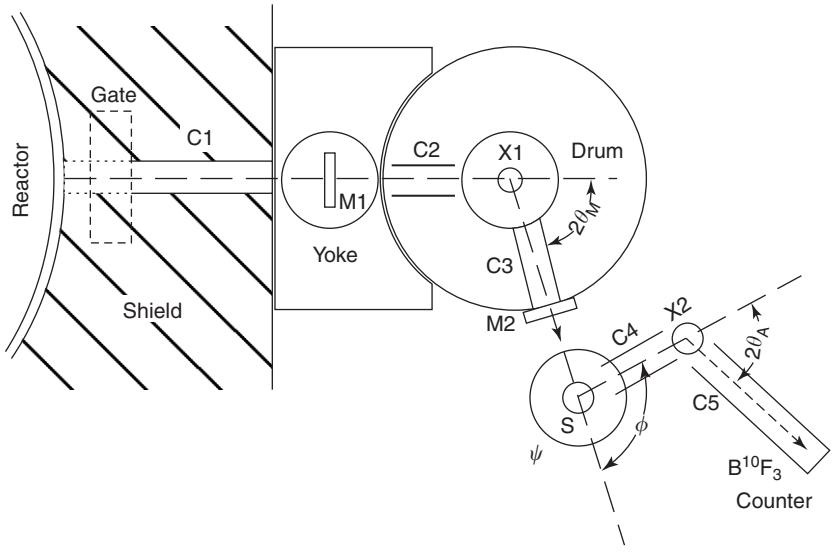


Figure 2 The basic layout of triple-axis spectrometers. The single crystals X1 and X2 are the monochromator and analyzer, respectively, and S is the sample. (Brockhouse BN and Stewart AT (1958) *Reviews of Modern Physics* 30: 236.)

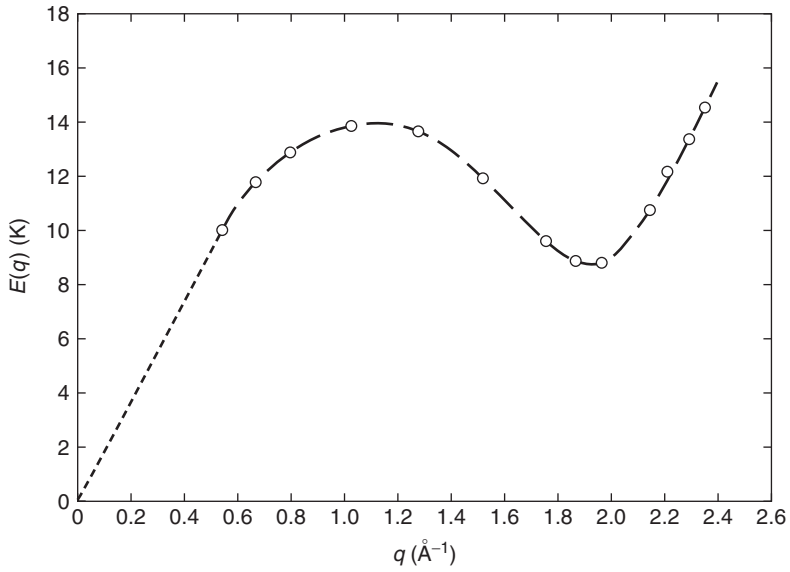


Figure 3 The dispersion relation of elementary excitations in superfluid ^{4}He observed by inelastic neutron scattering. (Yarnell JL, Arnold GA, Bendt PJ, and Kerr EC (1958) *Physical Review Letters* 1: 9.)

mechanical choppers to produce neutron pulses by Jacrot at Saclay. TOF techniques are most often used for isotropic samples, for example, liquids, where only the absolute value of $\boldsymbol{q}$ matters, not its direction compared to the orientation of the sample. **Figure 3** shows the elementary excitation spectrum of superfluid ^{4}He below 2.17 K. The observation of the minimum in the $E(q)$ dispersion relation at $q = 19.3\,\mathrm{nm}^{-1}$ provided direct evidence for the theories of Landau and Feynman.

Polarized Neutrons

The van Hove scattering function as given above applies to the simplified situation of neglecting the neutron magnetic moment. While the initial and final states of the neutrons are only fully characterized by not only keeping track of the velocity but also of the direction of the magnetic moment vector $\boldsymbol{\mu}$ (or that of the associated neutron spin with $s = 1/2$), only the velocity has been considered so far. This corresponds to the real experimental situation in "unpolarized neutron scattering," where $\boldsymbol{\mu}$ is randomly oriented in the incoming beam and its orientation is not determined in the scattered beam. In the most general case, one needs to define both the initial and final states by the five parameters ($\boldsymbol{v}_\mathrm{i}$, $\boldsymbol{\mu}_\mathrm{i}$) and ($\boldsymbol{v}_\mathrm{f}$, $\boldsymbol{\mu}_\mathrm{f}$), respectively (vector $\boldsymbol{\mu}$ represents only two free parameters, since its magnitude is given). Thus, the van Hove scattering function will take the form $S_{\alpha,\beta}(\boldsymbol{q}, E)$, describing the scattering probability from an initial neutron state with magnetic moment direction α to a final one with magnetic moment direction β.

The two basic methods used nearly exclusively until now to produce polarized neutron beams were also discovered early on, in 1951. One of them is essentially the reflection of neutrons on magnetized mirrors used by Hughes and Burgy to identify the interaction of neutrons with magnetic fields, as discussed above. The potential in eqn [2] for the internal B field in the ferromagnetic mirror (Fe or Co or their alloys) makes the critical angle of total reflection different for neutrons with magnetic moment oriented parallel or antiparallel to the direction of the field magnetizing the mirror. Thus, in the angular range between these two critical angles (referred to as "spin up" and "spin down" critical angles), the reflectivity will be substantially larger for one direction of the neutron magnetic moment, and the mirror selects from the originally unpolarized beam a preferred magnetic moment (or spin) direction. This technique is simple and inexpensive; it has the disadvantage though that the difference between the "up" and "down" critical angles is rather small, and therefore it can only select a rather limited fraction of the neutrons emerging from the reactor in all directions. In 1976 Mezei introduced the "supermirrors" to enhance the angular range of efficient polarizing neutron mirrors by complementing the range of total reflection by a range of nearly total reflection by interference on a stack of up to ~ 500 thin layers of alternating magnetic and nonmagnetic layers with carefully designed sequence of thicknesses between ~ 3 and 50 nm. Most mirror-based neutron polarizers in operation today use supermirrors.

The other basic method of polarizing neutrons is to use an adequate Bragg reflection from a single

crystal. In ferromagnetic materials, unlike the case of antiferromagnets shown in **Figure 1**, the magnetism of the atoms leads to a magnetic contribution to the Bragg peaks corresponding to the atomic crystal structure. This makes the intensity of the reflection dependent on the relative direction of the neutron magnetic moment with respect to the magnetic field polarizing the sample. For a few special Bragg peaks in various materials, the intensity of the reflection can be two orders of magnitude different for neutron spins parallel or antiparallel to the crystal magnetization. This was first observed by Shull in 1950 for a reflection in magnetite, Fe_3O_4. He also pointed out that by observing the effect of the beam polarization on the scattered beam intensity in a magnetized sample one can increase the sensitivity of neutron scattering experiments in the study of weak signals from magnetic samples by orders of magnitude, for example, for detecting ordered magnetic moments in the range of $10^{-3}\,\mu_B$ per atom.

Polarization Analysis

For the unambiguous identification of the magnetic signals in complex magnets or inelastic scattering, in particular in paramagnetic or antiferromagnetic samples which cannot be magnetized in moderate fields, one needs to resort to polarization analysis. Here the scattered beam intensity might not depend on the initial beam polarization; instead, the scattered beam polarization needs to be determined. Experimentally this can be achieved by replacing both the monochromator and analyzer crystals on a triple-axis spectrometer by a crystal with a polarizing reflection and control the beam polarization between the polarizer and analyzer crystals, for example, by maintaining it in a given direction or "flipping" it. Several researchers have made important steps in this direction, in particular, Drabkin and collaborators at Gatchina. In 1969, Koehler, Moon, and Riste established the first experimental evidence for the characteristic polarization behavior of different scattering processes. In the approach, they established (what is now called "one-dimensional polarization analysis") that both the initial and final beam polarization can be either parallel ($+$ or "up") or antiparallel ($-$ or "down") to the field direction on the sample. Thus, the choice of α and β directions in the above general form of the van Hove scattering function reduces to the combinations: "$++$" for the probability of scattering from the $+$ neutron spin direction to the same one, $S_{++}(\boldsymbol{q}, E)$, and similarly "$--$," "$+-$," and "$-+$." The first two processes are without change of spin direction ("non-spin flip") and the last two ones involve the reversal of the neutron spin ("spin flip"). Later developments (primarily at Delft, Gatchina, and ILL, Grenoble) led to generalized or vectorial (3D) polarization analysis (or polarimetry), where α and β can be arbitrary directions. A historically important example of complex vectorial behavior of the scattered beam polarization is the early prediction by Halpern and Johnson in 1942 for the magnetic scattering in paramagnetic samples, giving the direction of the scattered beam polarization $\boldsymbol{P}_f$ (which is defined as the average of the individual neutron magnetic moment directions in the beam) as a function of the initial one $\boldsymbol{P}_i$ and the unit vector parallel to the momentum transfer $\hat{\boldsymbol{u}} = \boldsymbol{q}/q$ as

$$\boldsymbol{P}_f = -\hat{\boldsymbol{u}}(\boldsymbol{P}_i\hat{\boldsymbol{u}})$$

Special cases of this relation (when $\boldsymbol{P}_i$ and $\hat{\boldsymbol{u}}$ are either parallel or perpendicular) have been demonstrated in 1969 by Koehler, Moon, and Riste, while the full vectorial relation could only be verified later.

Neutron Spin Echo

The generic scheme of neutron scattering experiments consisting of preparation of an initial beam with more or less well-defined $\boldsymbol{v}_i$ and analysis of the final distribution of $\boldsymbol{v}_f$, as it is common, for example, for triple axis (**Figure 2**), TOF, and other techniques. It implies that for the observation of small changes, the initial and final velocities must be very well defined, that is, very few neutrons will be selected from

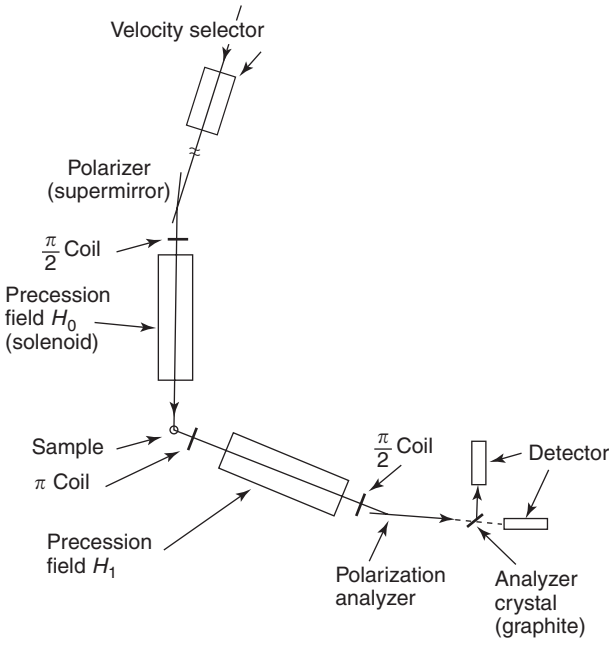


Figure 4 The layout of a neutron spin echo spectrometer. (Mezei F (ed.) (1980) *Neutron Spin Echo*. Berlin: Springer.)

a broad distribution in direction and absolute velocity. The neutron spin echo method, invented by Mezei in Budapest (1972) uses a drastically different paradigm. It consists in making each neutron effectively carry its own stopwatch and, using this, compare a selected component of v_i with another component of v_f. Thus, the determination of minute relative changes of the neutron energy in the 10^{-6} range can be measured at a rough monochromaticity of 10–20%. The Larmor precession of the magnetic moment of each neutron in well-defined magnetic fields serves as the clock attached to each neutron (**Figure 4**). Neutron spin echo thus allows to extend the upper limit of time domain 10^{-14}–10^{-9} s where atomic motions can be studied by methods of inelastic neutron scattering up to 10^{-6} s, making accessible many processes in macromolecular systems.

Steady-State and Pulsed Neutron Sources

The most recent revolution in neutron scattering is driven by the evolution of the methods of producing neutron beams. As we have seen, the need to select neutron beams with a more or less well-defined initial velocity v_i by some monochromator implies that one should throw away the vast majority of the neutrons produced with a broad velocity distribution and emitted in all directions by the moderators. Cutting out pulses of the beam by a mechanical chopper was found early on as an efficient way to select their velocity, for example, 0.2 ms pulses every 10 ms provide a suitable monochromatic beam for TOF spectroscopy over a typical flight path of 20 m. Such a monochromator only needs neutrons for a short time every 10 ms; for the rest the source could be switched off to save costs and heat production, the latter being the ultimate technical limitation. It is difficult but not impossible to operate reactors in a pulsed fashion, and this approach has been pioneered at Dubna near Moscow since the 1960s. The IBR-2 pulsed reactor source at Dubna today achieves 1500 MW instantaneous power (heat production) when switched on for about 0.3 ms, while steady-state reactors cannot practically surpass 100 MW.

Neutrons can also be produced by accelerators, that can easily operate in a pulsed mode. This approach was first pioneered by Tohuku University in the 1960s using electrons. The most efficient accelerator-based neutron production technique, the so-called "spallation," was established in Argonne near Chicago by Carpenter and collaborators in the 1970s. Here a pulsed proton beam is accelerated to typically 1 GeV energy, and each proton absorbed in a heavy-metal target sets free ~30 neutrons in a series of collisions with the nuclei. Thus, the energy needed per neutron produced (~30 MeV) is much less than the ~190 MeV heat accompanying the

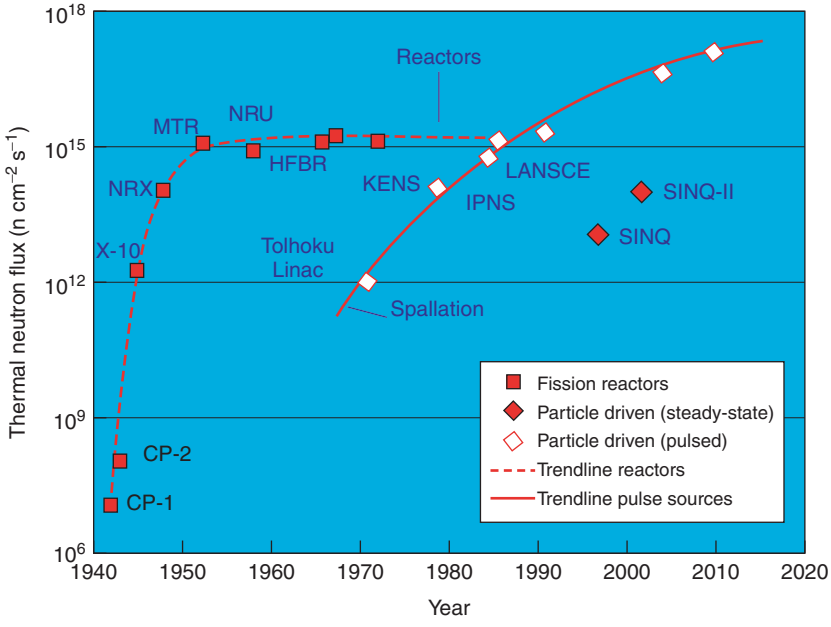


Figure 5 Evolution of the flux of neutron sources available for neutron scattering over the past decades at different facilities. The flux is defined here as the number of neutrons crossing unit area of the surface of the ambient temperature moderator in unit time. It is measured in units of $cm^{-2}s^{-1}$. For pulsed sources, the flux at the peak of the pulses is shown.

birth of a recoverable neutron in the fission process in reactors. The great new potential offered by high-power spallation sources, as illustrated in **Figure 5** by the peak instantaneous neutron flux they can produce during the pulse, is based on this double gain of efficiency: less heat is produced per usable neutron and the source is switched off for much of the time when no neutrons are needed. Although, by itself, the peak flux of a pulsed source compared to the time average flux of a steady-state reactor does not fully characterize its performance in all applications, the great potential of progress with future multi-MW spallation sources is well illustrated in **Figure 5**.

Neutron Optics

While source performance has not improved much since the 1950s in terms of number of neutrons produced, the intensity of neutron beams available in scattering experiments has steadily increased over the same period of time, in particular by the advent of Institut-Laue-Langevin (ILL) in Grenoble, the most powerful neutron scattering facility in the world operated through a European collaboration. At the 58 MW fission reactor at ILL three different moderator temperatures 25, 300, and 2000 K are available in order to provide the best-adapted neutron spectrum for the various instruments, and that has become a standard for more recent research reactors. The art of "neutron optics," the development of increasingly efficient components for transporting, shaping, monochromatizing, and analyzing neutron beams has also steadily progressed. One of the key inventions was the introduction of neutron guides by Maier-Leibniz in München in the early 1960s. Neutron guides are rectangular tubes with cross-section up to 100–200 cm^2 with the internal walls made of totally reflecting mirrors (usually highly polished glass coated with an adequate metallic layer). These guides can transport intense neutron beams over large distances with minimal losses with a beam divergence limited by the critical angle of total reflection (typically a few tenths of a degree). More recently, the divergence, and hence the intensity of the beam transported in the guides, is enhanced by using supermirror coating. The early single crystal monochromators are replaced by arrays of single crystals focusing the beam on the sample, and an increasingly larger fraction of the scattered neutrons is captured by using large position-sensitive area detectors. The combination of continuously advancing neutron optics with the enhanced neutron flux of the next-generation neutron sources means that the power of neutron scattering methods is increasing much faster than indicated by the trendline in **Figure 5** for spallation sources.

See also: Neutron Sources; Scattering, Inelastic: Brillouin; Scattering, Inelastic: Electron; Scattering, Inelastic: Raman.

PACS: 03.75.Be; 61.12.−q; 61.12.Bt; 61.12.Ex; 61.12.Ha; 61.12.Ld; 75.25.+z.

Further Reading

Bacon GE (1975) *Neutron Diffraction*. Oxford: Clarendon Press.
Mezei F (1986) La Nouvelle Vague in polarized neutron scattering. *Physica* 137B: 295–308.
Schofield P (ed.) (1983) *The Neutron and its Applications*. Bristol and London: The Institute of Physics.
Squires GL (1978) *Introduction to the Theory of Thermal Neutron Scattering*. Cambridge: Cambridge University Press.

NMR Methods for the Determination of Local Structure

R Blinc and B Zalar, J. Stefan Institute, Ljubljana, Slovenia

Introduction

Ideal crystals exhibit perfect translational symmetry. The arrangement of the atoms in a unit cell repeats itself in one, two, and three dimensions. In disordered systems on the other hand, the translational symmetry is locally broken. It is only the average structure, which is translationally invariant.

Many physical properties of disordered systems are of great interest both for basic physics as well as applications. Generally, they are not determined by the average structure, but by the local structure. Examples of such systems are, for instance, relaxors, that is, perovskite solid solutions with exceptional dielectric and electromechanical properties, which are characterized both by substitutional site and charge disorder, and the formation of polar nanoclusters. It has been shown that disorder is important also in perovskites such as $BaTiO_3$ and $SrTiO_3$, which were long considered to be classical examples of displacive soft-mode type ferroelectrics or

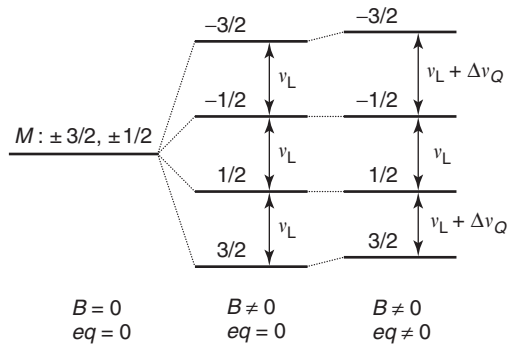


Figure 1 Shift of the Zeeman energy levels of a nucleus with a spin $I = 3/2$ in the presence of nonzero electric quadrupole coupling, $e^2qQ/h \neq 0$. Instead of a single Zeeman line now there are three lines: a central $1/2 \leftrightarrow -1/2$ transition and two satellite transitions, $3/2 \leftrightarrow -1/2$ and $-1/2 \leftrightarrow -3/2$. Here $eq = V_{zz}$ is the largest eigenvalue of the EFG tensor, e is the elementary charge, Q the nuclear quadrupole moment, and h Planck's constant.

ferroelastics, describable by anharmonic lattice dynamics. These are important in thin films and ferroelectric memory devices. The local rather than the average structure also determines the properties of hydrogen-bonded ferroelectrics such as KH_2PO_4, proton, and deuteron glasses as well as related H-bonded biomolecules and DNA. The same is true for many fullerenes as, for instance, for the purely organic ferromagnet TDAE-C_{60} or Cs_4C_{60}. Incommensurate systems, where one has perfect long-range order but no translational symmetry in one, two, or three dimensions, represent a special case. Here, the local structure is determined by the modulation wave, the periodicity of which is incommensurate to the periodicity of the basic lattice. Therefore, the translational periodicity is lost and the whole crystal is the unit cell. The small deviations from the average structure are hard to determine by X-ray or neutron scattering techniques, which depend on spatial averages. Local techniques, such as nuclear or electron magnetic resonance, on the other hand, directly reflect the local structure.

The relevant Hamiltonian is here, in the most general case, the sum of a Zeeman term, a chemical shift term, a dipole–dipole term, a nuclear quadrupole coupling term and a hyperfine coupling term

$$\mathcal{H} = \mathcal{H}_z + \mathcal{H}_{cs} + \mathcal{H}_{d-d} + \mathcal{H}_Q + \mathcal{H}_{hpf} \qquad [1]$$

The last four terms in expression [1] depend on the local structure and dynamics of the atomic surrounding of a given nucleus, and are very sensitive to small deviations from the average structure. The shifts of the Zeeman energy levels in case of a nonzero electric quadrupole coupling are schematically illustrated in **Figure 1**. The relation between the site symmetry and the quadrupole coupling term is shown in **Table 1**.

Table 1 EFG tensor and symmetry of the nuclear site

Site symmetry	*EFG tensor*
Cubic	0
Threefold, fourfold or higher symmetry rotational axis	Axially symmetric EFG tensor, asymmetry parameter $\eta = \frac{V_{xx} - V_{yy}}{V_{zz}} = 0$
Twofold axis	One principal axis of EFG tensor is parallel to the twofold axis

Disorder in Classical Perovskites

The question to be answered is whether the potential for Ti motion in the cubic phase of classical perovskites such as $BaTiO_3$ or $SrTiO_3$ exhibits a minimum at the center of the oxygen cage or whether the Ti ion is disordered between several off-center sites (**Figure 2**). Since the electric field gradient (EFG) tensor is zero by symmetry at the central position and nonzero at the off-center sites, the above problem can be solved by quadrupole perturbed Ti-nuclear magnetic resonance (NMR).

A recent NMR study indeed observed unresolved ^{47}Ti ($I = 5/2$) and ^{49}Ti ($I = 7/2$) first-order satellites indicative of nonzero quadrupole coupling in the cubic phase of $BaTiO_3$. This can only be understood if the Ti ions are off-center even in the cubic phase above the ferroelectric transition at T_c. Similar results have been obtained also in the cubic phases of $SrTiO_3$, ^{18}O enriched $SrTiO_3^{18}$, and $KTaO_3$ doped with Nb.

For the case that the EFG tensor at the Ti sites is nonzero, one expects $2I$ quadrupole perturbed NMR lines for each isotope in $BaTiO_3$. The frequency shift $\Delta\nu_m = \nu_{m-1 \leftrightarrow m} - \nu_L$ of the $m - 1 \leftrightarrow m$ transition is given by

$$\Delta\nu_m = -\nu_Q(m - 1/2)(3\cos^2\theta - 1 + \eta\sin^2\theta\cos 2\phi)/2 \qquad [2]$$

where $\nu_Q = 3K/[2I(2I-1)]$ is the quadrupole frequency, η is the asymmetry parameter of the EFG tensor, $K = e^2qQ/h$ is the electric quadrupole coupling constant, and θ and ϕ are the tilt and azimuthal angles, respectively, of the direction of the external magnetic field in the EFG tensor eigenframe xyz. Instead of sharp satellite lines, broad background components in addition to the two sharp $-1/2 \leftrightarrow 1/2$ central lines are observed, which are not affected by quadrupole coupling in the first order. The presence of the satellite background in $BaTiO_3$ and

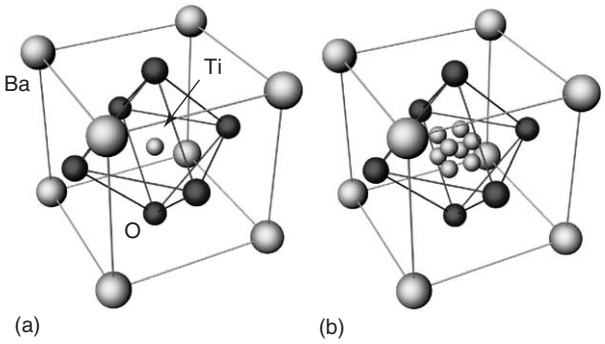


Figure 2 Possible Ti ion positions compatible with the cubic O_h symmetry of the paraelectric phase of $BaTiO_3$: (a) static Ti ion at the center of the oxygen cage of the ideal perovskite lattice, (b) Ti dynamically disordered among eight off-center sites displaced along the [111] body diagonals.

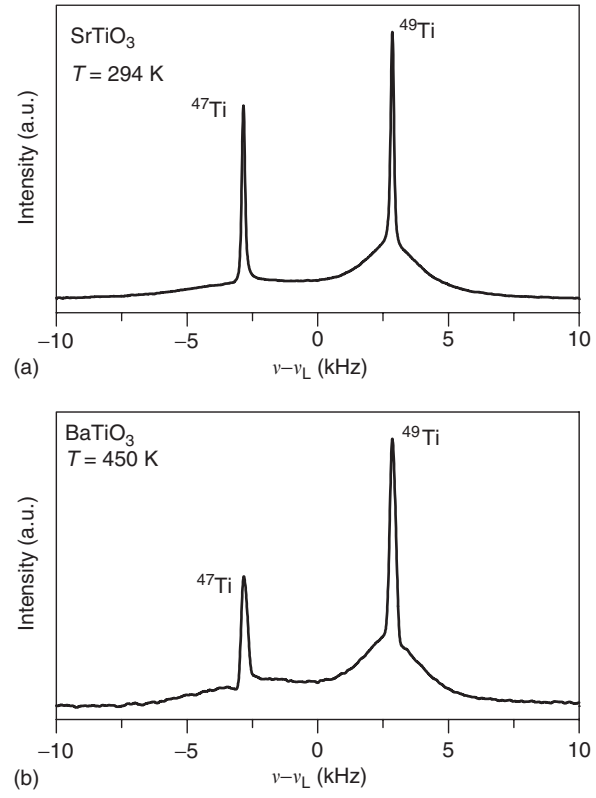


Figure 3 ^{47}Ti and ^{49}Ti NMR spectra of (a) $SrTiO_3$, and (b) $BaTiO_3$ single crystals in the cubic phase above T_c. The spectra show, in addition to the two central $1/2 \leftrightarrow -1/2$ lines, a broad background due to unresolved satellite transitions. The presence of these quadrupole-induced satellites demonstrates local tetragonal breaking of the macroscopic cubic symmetry.

$SrTiO_3$ demonstrates the local breaking of the O_h symmetry of the cubic phase. Without this breaking, there should be no satellite background around the central lines. The width of this unresolved satellite background shows that one has dynamic Ti

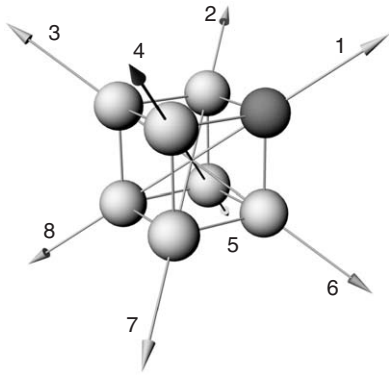


Figure 4 The eight site model of Ti disorder in $BaTiO_3$ and $SrTiO_3$.

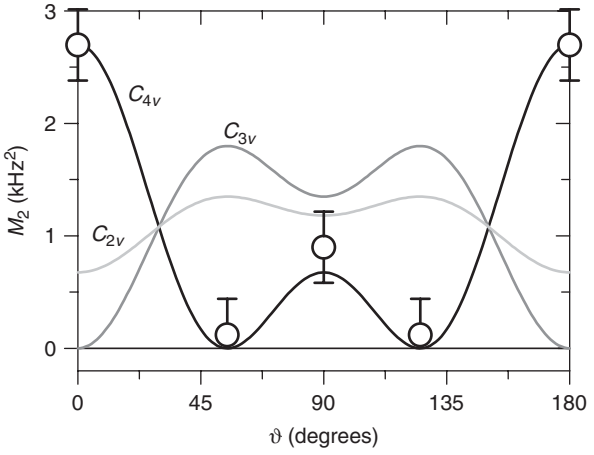


Figure 5 Angular dependence of the second moments M_2 of the ^{47}Ti quadrupolar background in $BaTiO_3$. For cubic symmetry there should be no angular dependence. The calculated values for tetragonal (C_{4v}), orthorhombic (C_{2v}), and rhombohedral (C_{3v}) local symmetry breaking in the paraelectric phase are shown as well. ϑ is the angle between the cubic direction [0 0 1] and the external magnetic field.

disorder among a subset of off-center sites (**Figures 3** and **4**).

The angular dependence of the second moments $M_2(\vartheta)$ of the ^{47}Ti spectra (**Figure 5**) shows that local tetragonal breaking of the cubic symmetry in $BaTiO_3$, $SrTiO_3$, and $SrTiO_3$[18] is dealt with.

The results can be understood within the eight-site order–disorder model (**Figure 4**). Sites 1, 2, 3, and 4 are each visited 25% of the time, whereas sites 5, 6, 7, and 8 are unvisited resulting in a local tetragonal state with a polarization along the [001] unit cell edge. The bias slowly switches so that after some time the sites 1, 2, 5, and 6 are visited 25% of the time and the polarization is along [010], etc. Thus, even in the cubic paraelectric phase, each unit cell is polarized. Only the time-averaged macroscopic polarization is zero. It should be stressed that two

distinct timescales as demonstrated by the T_2 data are dealt with here. On a short timescale, the Ti ions hop rapidly between four [111] off-center positions giving rise to a dipole moment along the cubic directions [100], etc. On a longer timescale, these tetragonal nanodomains can reorient between six equivalent directions.

Orthorhombic and rhombohedral symmetry breaking does not agree with the $M_2(\vartheta)$ data in the cubic phase. Below T_c the broad component (i.e., the unresolved satellite background) disappears and is replaced in $BaTiO_3$ by well-resolved satellites showing the presence of 90° ferroelectric domains. The tetragonal nanodomains become macroscopic and static below T_c.

One further point should be stressed. Assuming a fast motional exchange of the Ti ions among the off-center sites and site occupation probabilities n_i, one finds the motionally averaged EFG tensors for the cubic (O_h) and tetragonal (C_{4v}) phases as

$$\langle \underline{\underline{V}} \rangle = \sum_{i=1}^{8} n_i \underline{\underline{V_i}} \qquad [3]$$

leading to

$$\langle \underline{\underline{V}} \rangle_{O_h} = \frac{1}{2}(3\cos^2\theta - 1)\underline{\underline{V_0}}$$
$$\underline{\underline{V_0}} = eq \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [4]$$

and

$$\langle \underline{\underline{V}} \rangle_{C_{4v}} = 2(n_1 + n_5)(3\cos^2\theta - 1)\underline{\underline{V_0}} \qquad [5]$$

In the cubic phase $n_i = 1/8$, $i = 1,\ldots,8$, whereas in the tetragonal phase $n_1 = n_2 = n_3 = n_4 \neq n_5 = n_6 = n_7 = n_8$. In the undistorted cubic and tetragonal phases, the body diagonals [111] lie at the magic angle $\theta_M = 54.7°$ with respect to the cubic cell edges. As a result of this, $3\cos^2(\theta = \theta_M) - 1 = 0$, and the motionally averaged EFG tensor $\langle \underline{\underline{V}} \rangle$ vanishes though the EFG tensors at the eight off-center sites are different from zero. In fact, the tetragonal bias induces a slight deformation of the unit cell, $c/a \neq 1$, in the tetragonal nanodomains in the cubic phase so that $\langle \underline{\underline{V}} \rangle \neq 0$. The angle between the body diagonals and the unit cell edges deviates from the magic angle, $\theta = \theta_M + \delta$. In the tetragonal clusters inside the macroscopically cubic phase $\delta \approx 0.01°$. The situation is similar in $SrTiO_3$ and $SrTiO_3$[18].

On further cooling, $BaTiO_3$ undergoes a tetragonal–orthorhombic and at still lower temperatures, an orthorhombic–rhombohedral phase transition. These transitions are connected with a rotation of the spontaneous polarization due to changes in the site occupation probabilities n_i resulting from the biased freeze-out of dynamical Ti disorder. In the low T rhombohedral phase, the disorder is frozen out and only one of the eight sites is occupied.

Relaxors

"Weak" Relaxors

A well-known example of a prototype incipient ferroelectric or a "weak" relaxor is substitutionally diluted $KTa_{1-x}Nb_xO_3$. Pure $KTaO_3$ does not undergo any ferroelectric transition as $T \to 0$. On decreasing temperature, it shows a strong increase in the dielectric constant related to the softening of its ferroelectric phonon mode. $KNbO_3$, on the other hand, is a perovskite ferroelectric displaying transitions from the paraelectric cubic to the ferroelectric tetragonal, orthorhombic, and rhombohedral phases in complete analogy to $BaTiO_3$. Substituting Nb for Ta induces a diffuse ferroelectric phase transition in $KTa_{1-x}Nb_xO_3$ for $x > 0.008$, so that the transition temperature T_c varies with x as $T_c(x) = A(x - x_c)^{1/2}$. Within the experimental accuracy, the Ta ions occupy the center of symmetry position, whereas the Nb ions are off-center even in the cubic phase and move in a multiwell potential as in the case of $BaTiO_3$. For small x, one has a transition to a dipolar glass phase ($x \leqslant 0.06$) with no symmetry breaking. For larger x, one finds the same sequence of phase transitions as in $BaTiO_3$.

The breaking of the average cubic symmetry in $KTa_{1-x}Nb_xO_3$ far above $T_c \cong 150$ K can be easily seen

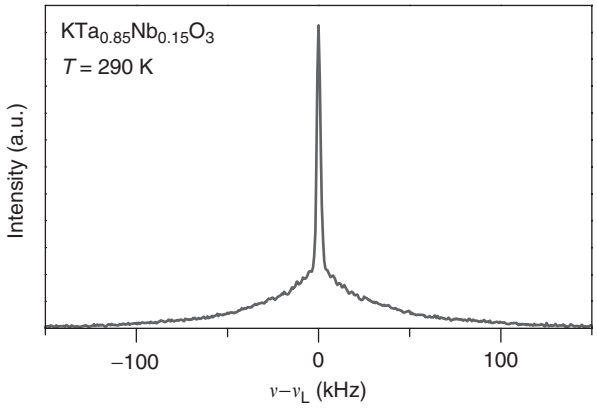


Figure 6 ^{93}Nb ($I = 9/2$) NMR spectrum of a $KTa_{0.85}Nb_{0.15}O_3$ single crystal in the cubic phase at room temperature. The presence of the quadrupole-coupling-induced unresolved satellite background demonstrates that the Nb ions are disordered between off-center sites in analogy to the Ti case in $BaTiO_3$.

from the nonzero Nb quadrupole coupling-induced unresolved satellite background for $x = 0.15$ (**Figure 6**).

"Strong" Relaxors

"Strong" perovskite relaxors such as $PbMg_{1/3}Nb_{2/3}O_3$ (PMN) are intermediate between dipolar glasses and classical ferroelectrics. They exhibit both site and charge disorder, and are therefore characterized by the presence of both random interactions and random fields. They show a broad and large, frequency-dependent peak in the dielectric constant (~ 300 K for PMN) and huge electro-mechanical coupling. In analogy to dipolar glasses, they show no macroscopic symmetry breaking and the structure remains cubic down to liquid helium temperatures. This is true even on the micrometric scale. On the nanometric scale, on the other hand, the local structure is different from the average structure. Burns and Dacol suggested that randomly oriented polar nanoclusters appear in PMN below 925 K so that one deals with a heterostructure consisting of two phases.

A characteristic difference between dipolar glasses such as $Rb_{1-x}(NH_4)_xH_2PO_4$ and relaxors is that, in relaxors a ferroelectric state appears for electric fields larger than a critical field, $E > E_c$, whereas this is not the case in dipolar glasses. This difference is obviously due to polar nanoclusters, which do not exist in dipolar glasses.

The local structure of PMN has been recently studied by field-cooled (FC) and zero field-cooled (ZFC) ^{207}Pb ($I = 1/2$) NMR. Since ^{207}Pb has no electric quadrupole moment, the dominant interaction which allows for the study of polar nanoclusters, is here the chemical shift. If the Pb nuclei in the polar clusters are displaced from their high-symmetry cubic perovskite sites, the ^{207}Pb chemical shift tensors become anisotropic. They can be written as a sum of a scalar part σ_0 and a traceless second-rank tensor part $\underline{\underline{\sigma_a}}$:

$$\underline{\underline{\sigma}} = \sigma_0 \underline{\underline{1}} + \underline{\underline{\sigma_a}} \qquad [6]$$

The eigenvalues of the traceless part are $-\sigma_a/2$, $-\sigma_a/2$, and σ_a. The largest principal axis of $\underline{\underline{\sigma_a}}$ should be parallel to the direction of the average shift of the Pb nucleus, that is, to the direction of the polarization p_i of a given polar cluster i.

For a general orientation of the crystal with respect to the direction of the magnetic field $\boldsymbol{B}_0$ specified by (θ,ϕ), the ^{207}Pb NMR transition frequency is given by

$$\nu_{Pb} = \nu_L[1 + \sigma_0(p) + \sigma_{ZZ}(p, \vartheta, \varphi, \theta, \phi)] \qquad [7]$$

where ν_L is the unperturbed Pb Larmor frequency, σ_0 is the isotropic part of the chemical shift tensor, and σ_{ZZ} is the component of $\underline{\underline{\sigma_a}}$ along the magnetic field direction. The angles (ϑ, φ) determine the orientation of the eigenframe of $\underline{\underline{\sigma_a}}$ with respect to the crystal fixed frame.

A simple nearly isotropic Pb spectrum of roughly Gaussian-line shape (**Figure 7**) is seen for both ZFC and FC spectra above 210 K. Its half-width at 410 K is ~ 30 kHz, that is, much larger than what can be accounted for by nuclear dipolar interactions. The line shape reflects a spherical shell-type distribution of chemical shift tensors due to nearly isotropic Pb displacements of various magnitudes from the cubic positions. The corresponding polarization distribution function is the one predicted by the spherical random bond–random field model. Below 210 K, an anisotropic line appears in addition to the isotropic component in both FC and the ZFC spectra. For $E > E_c$, a sudden increase in the intensity of the anisotropic component takes place (**Figure 7**) at 210 K. This signals a first-order type phase transition into an electric field-induced ferroelectric phase. The angular dependence of the anisotropic line shows that it corresponds to frozen-out Pb shifts along the [111] direction, that is, the direction of the spontaneous polarization. About 50% of the Pb nuclei still reside in the isotropic spherical glass matrix. The transition at T_c is thus an orientational percolation transition of frozen ferroelectric polar clusters with the polarization and the Pb shift parallel to the [111] direction. A similar though smaller increase in the concentration of frozen-out clusters with a Pb shift along [111] is seen also in the absence of an electric field. The concentration of ferroelectric clusters here does not seem to be large enough to induce a transition to the ferroelectric state. PMN thus seems to be an incipient ferroelectric.

The temperature dependence of the spin–spin relaxation time T_2 reflecting the nanocluster dynamics is shown in **Figure 8**. On cooling down from 400 K, T_2 decreases with decreasing temperature demonstrating the rapid slowing down of the nanocluster dynamics in the spherical glass matrix. A very pronounced T_2 minimum appears at ~ 255 K, that is, at the temperature where a maximum of the dielectric susceptibility is found. Here the inverse nanocluster correlation time is in the kHz range. The same mechanism, that is, fluctuations in the orientation of the nanocluster polarization, related to Pb jumping between several off-center sites, is responsible for the dielectric losses as well as for the T_2 behavior. The spin–lattice relaxation time T_1 is here by three orders of magnitude longer than T_2, demonstrating that one deals with a frequency distribution peaked at low frequencies. At $T > 250$ K, T_2 only weakly depends on the frequency offset, demonstrating that the

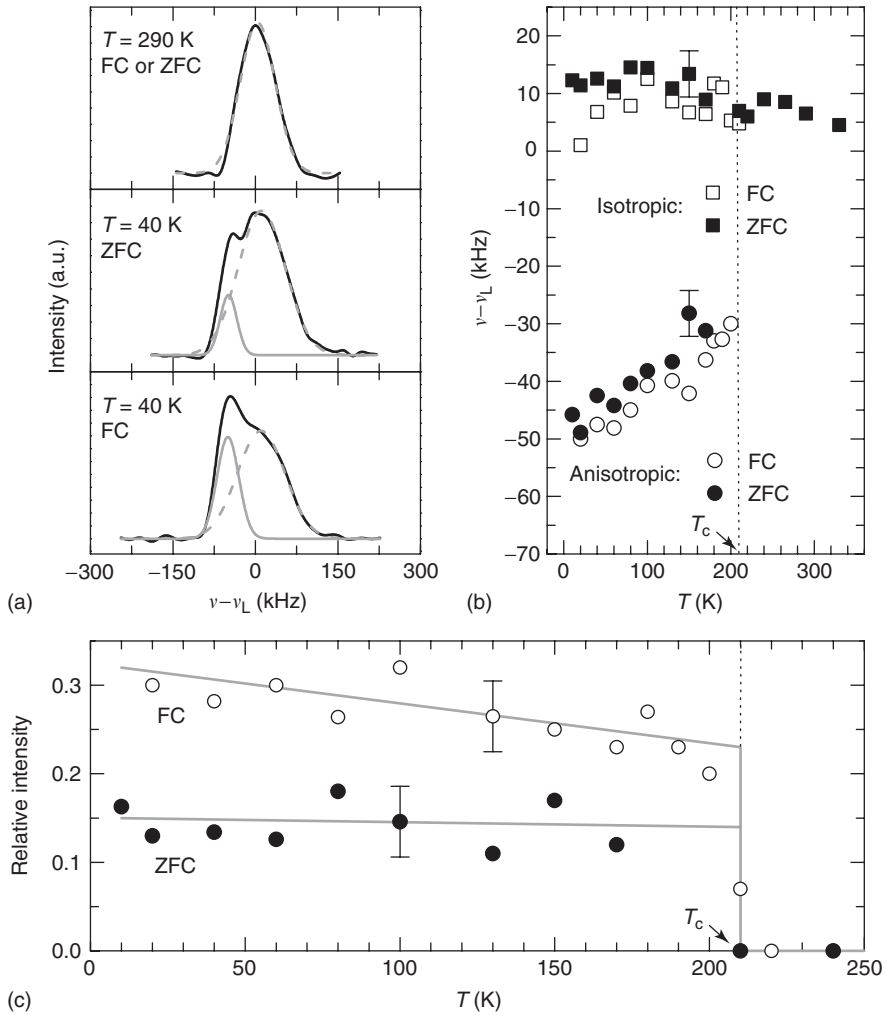


Figure 7 (a) Decomposition of the FC and ZFC ^{207}Pb NMR spectra of a PMN single crystal at 9.1 T and at $\boldsymbol{B}_0 \| \boldsymbol{E} \| [111]$ into isotropic and anisotropic components, (b) T-dependence of the positions of the isotropic and anisotropic ^{207}Pb NMR lines, (c) T-dependence of the relative intensity of the anisotropic component. Here $E = 3\,\mathrm{kV\,cm^{-1}} > E_c$.

nanocluster dynamics is similar in different parts of the crystal.

There is no difference in the FC and ZFC T_2 values above $T_c = 210\,\mathrm{K}$. Below the percolation transition at T_c, on the other hand, T_2 strongly depends on the frequency offset in the FC measurements, that is, the nanocluster dynamics is different in different parts of the crystal as expected for a two-component system. The FC and ZFC data differ significantly. A second rather flat T_2 minimum appears at $\sim$140–150 K in the FC data. It can be attributed to the ferroelectric [111] polarized clusters, whereas the first minimum is due to the spherical glass matrix.

H-Bonded Systems

The chemical, physical, and biological properties of H-bonded systems and biomolecules are determined by the local structure of the H-bonds, that is, by the position of the proton or deuteron in the hydrogen bonds and its dynamics. ^{17}O and deuteron quadrupole coupling, in particular, have been shown to be very sensitive indicators of that. The case of KH_2PO_4 is illustrated in **Figure 9**. Here, one has in the paraelectric phase, a fast exchange of the protons between the two equilibrium sites in the H-bond, ^{17}O–H…O and ^{17}O…H–O. These two sites correspond to the "close" respectively "far" positions of the proton with respect to a given ^{17}O ($I = 5/2$) nucleus. Due to fast exchange, only one averaged ^{17}O nuclear quadrupole resonance (NQR) line is seen in the paraelectric phase. In the ferroelectric phase below T_c, however, the proton motion is frozen out and one sees two separate ^{17}O NQR lines corresponding to the "close" and "far" proton positions. Still another technique, which allows for a very precise measurement of the ^{17}O–H distance, is

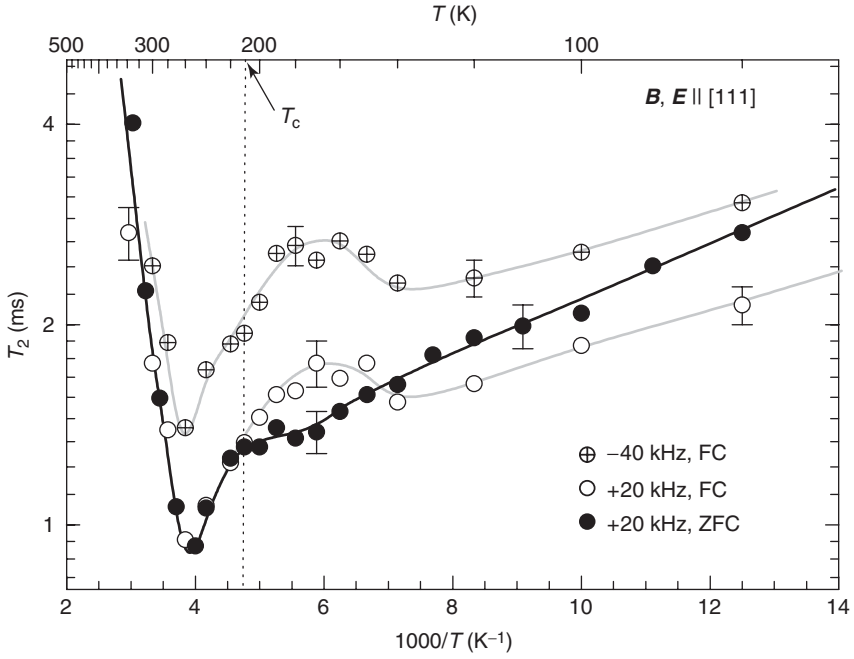


Figure 8 Temperature dependence of the ^{207}Pb spin–spin relaxation time T_2 in single crystal PMN reflecting the nanocluster dynamics in the FC and ZFC case.

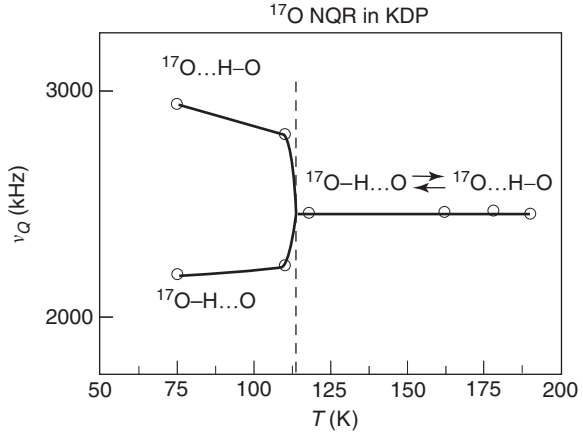


Figure 9 ^{17}O ($I = 5/2$) quadrupole resonance transitions in KH_2PO_4 obtained by the quadrupole double resonance technique. Below T_c are seen separate lines for the proton in the "far" ^{17}O…H–^{16}O and "close" ^{17}O–H…^{16}O positions whereas above T_c only one "average" line due to fast proton exchange between the two positions in the H-bond is seen. The ^{16}O nucleus has no quadrupole moment and does not contribute to the spectrum.

based on the measurement of the proton dipolar splitting of the ^{17}O NQR line by proton-^{17}O double resonance.

A significant progress in the understanding of the proton dynamics and proton transfer in solids has been recently achieved by two-dimensional (2D) deuteron and ^{31}P exchange NMR spectroscopy. Due to the so-called Pauling–Slater ice rules, only

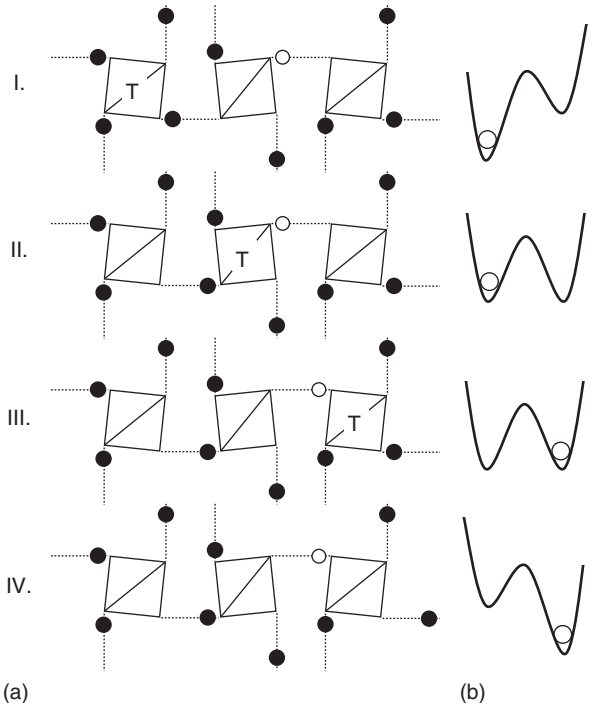


Figure 10 (a) Motion (steps I–IV) of Takagi H_3PO_4 defect (T) and (b) shape of the O–H…O H-bond potential in the proton (or deuteron) glass $Rb_{1-x}(NH_4)_xH_2PO_4$ or $Rb_{1-x}(NH_4)_xH_2AsO_4$. After several Takagi defect visits, the O–H…O potential linking different H_2PO_4 Slater groups becomes effectively of the symmetric double minimum type, thus explaining the tetragonal symmetry of these glasses and of the paraelectric phase of KH_2PO_4.

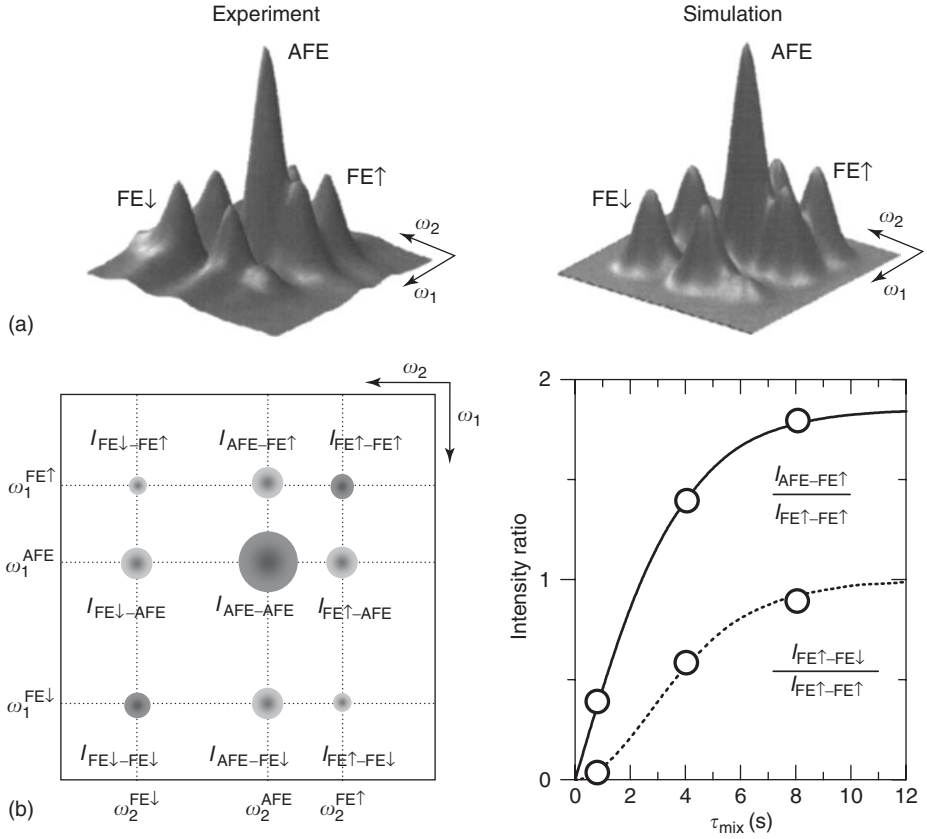


Figure 11 (a) ^{31}P 2D exchange NMR spectrum of $Rb_{1-x}(ND_4)_xD_2PO_4$ with $x = 0.5$ showing the exchange between the four antiferroelectric and the two ferroelectric Slater H_2PO_4 configurations due to the motion of Takagi defects, (b) Corresponding schematic 2D contour plot and dependence of the intensities of the 2D exchange peaks on the mixing time τ_{mix}.

two out of the four protons linking adjacent PO_4 groups via double minimum type O–H…O bonds, are close to each PO_4 ion in KH_2PO_4 type ferroelectrics and proton or deuteron glasses. The intrabond proton transfer from one well to another creates a Takagi HPO_4–H_3PO_4 defect pair which has a higher energy than the Slater pair H_2PO_4–H_2PO_4 it is originating from. In contrast to that, a H-bond linking an unpaired Takagi defect with a Slater group has a symmetric double-well potential (**Figure 10**). The visits of Takagi groups to a given Slater group invert the bias of the H-bonds (**Figure 10**), thus symmetrizing the double-well H-bond potential on a long enough timescale.

The switching among different Slater H_2PO_4 configurations induced by unpaired Takagi H_3PO_4 and HPO_4 diffusion has been indeed directly detected in a deuteron glass $Rb_{0.5}(ND_4)_{0.5}H_2P0_4$ by ^{31}P 2D exchange NMR (**Figure 11**). The resulting symmetrization of the O–H…O double minimum potential has been observed by 2D deuteron exchange NMR in $Rb_{0.68}(ND_4)_{0.32}H_2AsO_4$ (**Figure 12**). τ_{mix} is the mixing time of the 2D NMR exchange experiment pulse sequence, that is, the time over which the correlation of proton configurations in the H-bond network is measured.

Jahn–Teller Effect in C_{60}^-

Heisenberg predicted in 1928 that ferromagnetism is not expected to be found in organic systems made only of first row elements such as hydrogen, carbon, oxygen, and nitrogen. In view of that, it is rather remarkable that the C_{60}-based compound $TDAE^+$-C_{60}^- (here TDAE stands for tetrakis-dimethylaminoethylene) exhibits ferromagnetism below 16 K. This has been ascribed to the possible Jahn–Teller distortion of the spherical C_{60}^- ion into a rugby-like ball shape and the resulting redistribution of the unpaired electronic spin density into a belt-like form. The expected Jahn–Teller distortions are rather small (≈ 0.01 Å) and could not be observed by scattering techniques. Recently, however, ^{13}C NMR of the C_{60}^- ion provided for a direct evidence of this effect. The observation is made possible by the changes in the hyperfine Fermi electron–nuclear coupling shift of the ^{13}C NMR lines which accompany the Jahn–Teller distortion and act as a

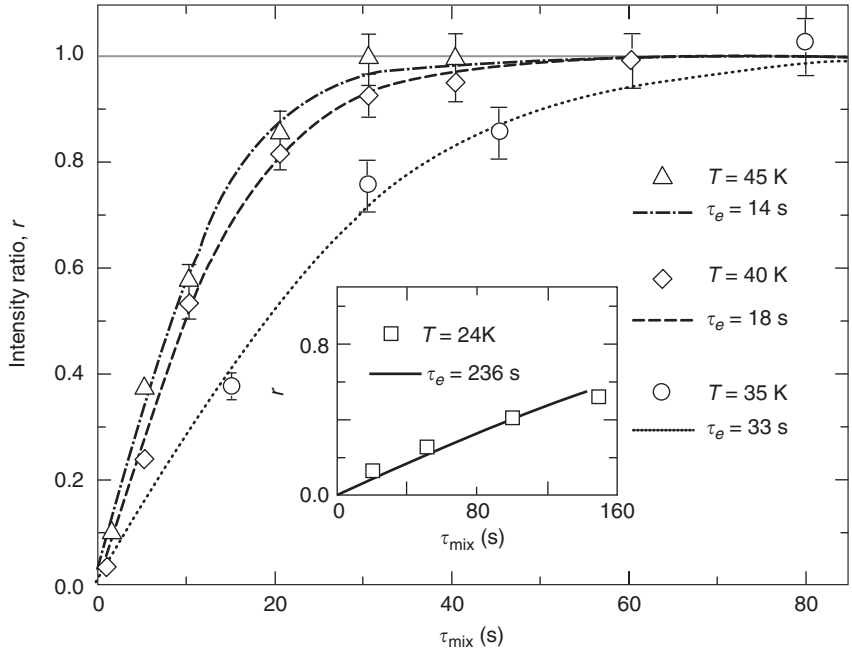


Figure 12 Dependence of the cross to diagonal peak intensity ratio r on the mixing time for 2D deuteron exchange NMR in $Rb_{0.68}(ND_4)_{0.32}D_2AsO_4$ showing the effective symmetry of the double minimum O–H…O potential as well as the exchange timescale τ_e. r should be smaller than 1 for an asymmetric potential and equal to 1 for a symmetric double minimum potential as indeed observed.

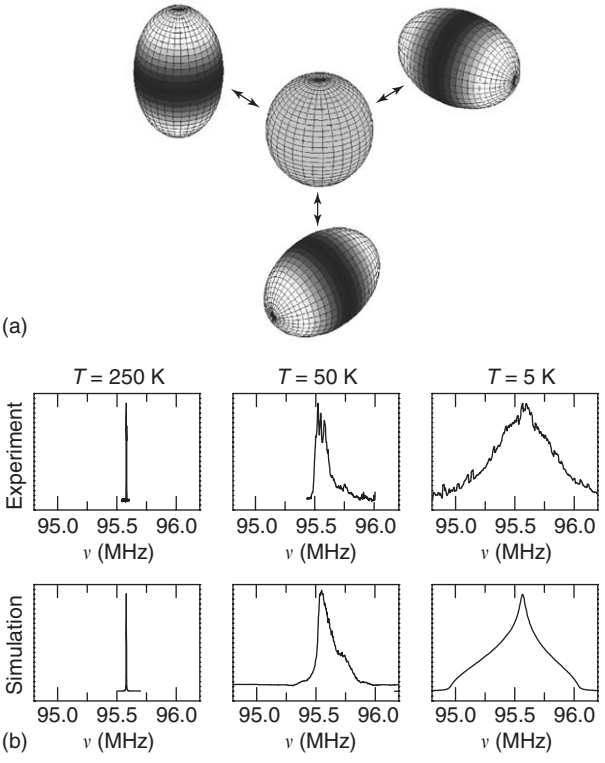


Figure 13 (a) Dynamic Jahn–Teller distortions of the C_{60}^- ion in the organic ferromagnet $TDAE^+$-C_{60}^-. The shaded areas show the belt-like concentration of the unpaired electron spin density around the equator in the distorted configurations, (b) ^{13}C NMR of the C_{60}^- ions in the temperature region of the freeze-out of the dynamic Jahn–Teller distortion, together with the theoretical simulations.

magnifying glass. In the undistorted C_{60}^- ion (or in a C_{60}^- ion exhibiting very fast exchange between different distorted configurations), all ^{13}C nuclei in the C_{60}^- ion are equivalent resulting in a narrow ^{13}C NMR line (**Figure 13**). In the frozen distorted case, on the other hand, the coupling Hamiltonian, which is a sum of the chemical shift term, the Fermi contact term and the electron nuclear dipole–dipole coupling term

$$\mathcal{H} = \mathcal{H}_{\mathrm{cs}} + \mathcal{H}_{\mathrm{hf,F}} + \mathcal{H}_{d_e,d_n} = \sum_j \boldsymbol{I}_j \underline{\underline{\mathcal{K}_t}} \boldsymbol{B}_0 \qquad [8]$$

becomes different for different nuclei.

This results in a huge increase in the width of the ^{13}C NMR spectrum when the dynamic Jahn–Teller distortions freeze out (**Figure 13**). The resulting belt-like distribution of the unpaired spin density enhances the ferromagnetic coupling between neighboring C_{60}^- ions (**Figure 14**), giving rise to an increase in T_c by three orders of magnitude. Similar Jahn–Teller distortions should take place in other fullerenes, for example, CsC_{60}.

Incommensurate Systems

In incommensurate (I) systems, the periodicity of at least one of the modulation waves cannot be expressed as a rational fraction of the periodicity of the underlying lattice. As a result of that, the translational

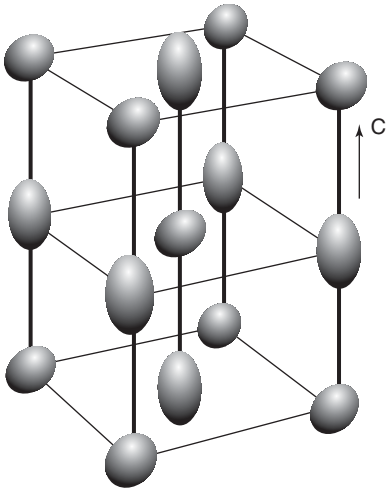


Figure 14 Lattice of Jahn–Teller distorted C_{60}^{-} ions leading to maximum ferromagnetic exchange interaction.

lattice periodicity is lost in spite of the existence of perfect long-range order. Phase transitions leading to such I phases are described by order parameters with a minimum dimensionality two ($n=2$), four ($n=4$), or six ($n=6$), resulting in systems with one (1-q), two (2-q), and three (3-q) I modulation waves. Examples of 1-q systems are, for example, Rb_2ZnCl_4, 2-q systems $AlPO_4$, and 3-q systems Ag_3AsS_3.

The modulation waves are independent if the modulation vectors $\boldsymbol{q}$ are independent and the corresponding phases are not correlated. If, however, the number of modulation waves is larger than the dimension of the space spanned by the corresponding wave vectors, and their phases are correlated, the modulation waves are not independent. The properties of such a system now strongly depend on the relative phases of the modulation waves. Such a situation occurs, for example, in quartz and berlinite ($AlPO_4$) where $n=6$, but the three modulation waves are coplanar, $D=2$, $\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3=0$, and their relative phase is fixed, $\psi_1+\psi_2+\psi_3=0$, $\pm\pi/2$, or π, as well as in several charge density wave (CDW) systems, for example, in 2H-$TaSe_2$. It should be stressed that the properties of the above 3-q, $D=2$ systems should be qualitatively different from those of a noncoplanar 3-q, $D=3$ I system with three independent modulation waves.

NMR and NQR are ideally suited for a determination of the number of independent modulation waves in I systems and their nature, as well as for the determination of their relative phases if the modulation waves are correlated (**Figure 15**).

Solitons have been observed in 1-q insulators such as Rb_2ZnCl_4 and in a number of coplanar 3-q CDW systems, for example, in 2H-$TaSe_2$ by ^{77}Se NMR, as well as in proustite by ^{75}As NQR.

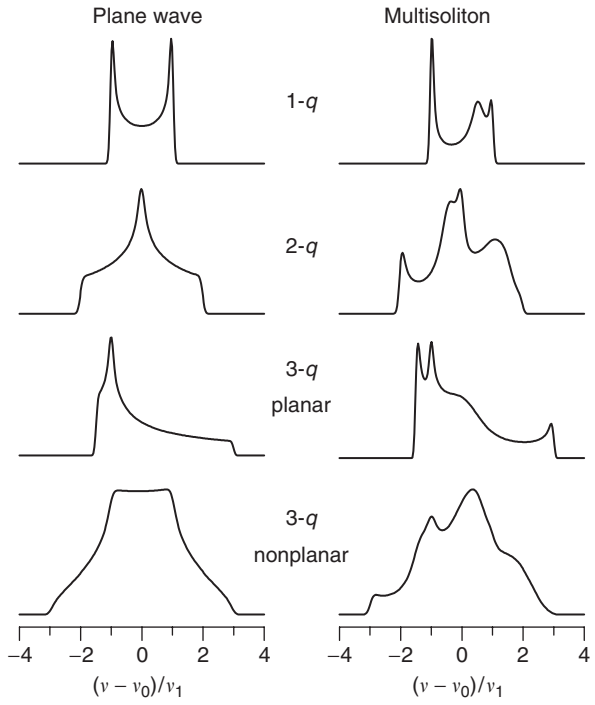


Figure 15 Theoretical NQR line shapes for the 1-q, 2-q, planar 3-q, and nonplanar 3-q incommensurately modulated systems in the plane wave and multisoliton lattice limits for a linear relation between the resonance frequency and nuclear displacement. The line shapes are obtained from eqn [13]. For the soliton limit, the phases are obtained as solutions of the sine-Gordon eqn [15]. In the planar 3-q case, the position of the central logarithmic singularity depends on the relative phases of the modulation waves even in the plane wave limit, while in this limit the spectrum is independent of the initial phases for the noncoplanar 3-q case.

The I state in proustite Ag_3AsS_3 can be described by a six-component order parameter

$$Q_{\pm l}=\rho_l \exp[\pm i\phi_l(x_l)], \quad l=1,2,3 \qquad [9]$$

where ρ_l stands for the amplitude and ϕ_l for the phase of the three independent I modulation waves. If the 1-q multidomain state is realized below T_I, then for a given domain, one of the following possibilities is realized:

$$\begin{aligned} &(1)\quad \rho_1=\rho_2=0, \quad \rho_3\neq 0 \\ &(2)\quad \rho_2=\rho_3=0, \quad \rho_1\neq 0 \\ &(3)\quad \rho_1=\rho_3=0, \quad \rho_2\neq 0 \end{aligned} \qquad [10a]$$

Each domain is here modulated in a single direction, but the direction of the modulation wave vector varies from one domain to another. For the genuine 3-q state, one has on the other hand a simultaneous I

modulation along three different directions

$$\rho_1 = \rho_2 = \rho_3 = \rho \neq 0 \qquad [10b]$$

In the general case, the dependence of the NQR resonance frequency ν_i on the nuclear displacements u_i in the I phase is nonlocal,

$$\nu_i = \nu[u_i(x_i), u_j(x_j), \ldots], \quad i \neq j \qquad [11]$$

where j runs over all ions that contribute to the resonance frequency at the ith site. Expanding the above relation in powers of the displacements, one finds for the ^{75}As NQR frequency in proustite in the 3-q case:

$$\nu_i(\phi_1, \phi_2, \phi_3) = \nu_0 + \nu_1 \sum_{l=1}^{3} \cos\,\phi_l + \nu_2 \sum_{l=1}^{3} \cos(2\phi_l + \phi_{20}) + \cdots \qquad [12]$$

where ν_1 is proportional to the amplitude of the modulation wave, that is, to the order parameter, $\nu_1 \propto \rho \propto (T_I - T)^{\beta}$, and $\nu_2 \propto (T_I - T)^{\tilde{\beta}}$ with $\tilde{\beta} \approx 2\beta$. ϕ_{20} describes the relative phase shift between the linear and quadratic terms. The I frequency distribution $f_{3\text{-}q}$ is now obtained in the constant amplitude approximation as

$$f_{3\text{-}q}(\nu) = K \iiint \delta\{\nu - \nu[\phi_1(x_1), \phi_2(x_2), \phi_3(x_3)]\}\, dx_1\, dx_2\, dx_3 \qquad [13]$$

where the x_l are variables along the directions of the modulation wave vectors. In the 1-q case, expression [13] simplifies to the well-known expression

$$f_{1\text{-}q}(\nu) = K \int \delta\{\nu - \nu[\phi(x)]\}\, dx = \frac{\text{const}}{d\nu/dx} = \frac{\text{const}}{(d\nu/d\phi)(d\phi/dx)} \qquad [14]$$

The resulting line shapes for 1-q, 2-q, planar 3-q, and noncoplanar 3-q cases are shown in **Figure 15** for the plane wave as well as for the multisoliton lattice limits. In the linear case ($\nu_2 = 0$), one finds, for a 1-q plane-wave type modulation, a frequency distribution limited by two edge singularities at $\nu - \nu_0 = \pm\nu_1$, where $d\nu/d\phi = 0$ (**Figure 15**). In the 2-q case, a line shape is found, characterized by a logarithmic singularity in the center. For the planar 3-q case, the line shape is similar to that of the 2-q case, but the position of the central singularity depends on the relative phases between the modulation waves. For the noncoplanar 3-q phase with three independent modulation waves, the NMR spectrum is bell-shaped and does not depend on the initial phases in the plane-wave limit. In contrast to the 1-q and 2-q cases and to the planar 3-q case, there are no singularities in the NMR spectrum but there are singularities in the derivative spectrum. As long as ϕ is a linear function of x, $d\phi/dx$ is nonzero and constant, and this term does not affect the spectral line shape. In the nonlinear case of a multisoliton lattice, on the other hand, $d\phi/dx \to 0$ in the commensurate regions, leading to the appearance of the characteristic "commensurate" lines in the spectrum.

From the Landau type free energy expansion model of Pokrovsky, one finds that in the simplest case the independent phases of the three modulation waves are obtained as solutions of the three

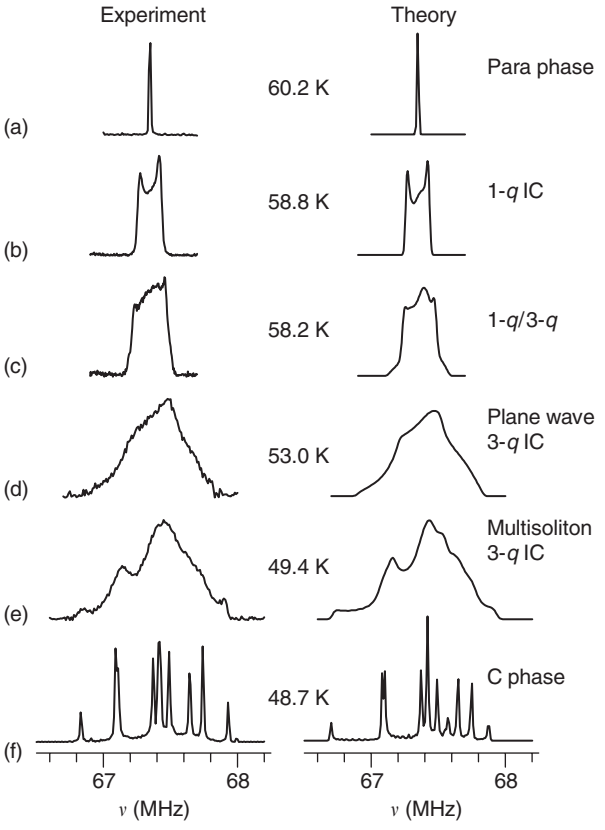


Figure 16 Comparison between the observed and theoretical ^{75}As NQR line shapes in proustite: (a) $T = 60.1$ K, para phase, (b) $T = 58.8$ K, 1-q I phase, (c) $T = 58.2$ K, coexistence of 1-q and 3-q phases, (d) $T = 53$ K, plane-wave regime of 3-q I phase, (e) $T = 49.4$ K, multisoliton regime of 3-q I phase and, (f) $T = 48.7$ K, commensurate phase.

uncoupled sine-Gordon equations:

$$\phi''(x_l) = -6Dt^2 \sin[6\phi(x_l)], \quad l = 1, 2, 3 \quad [15]$$

Here $t = 1 - T/T_I$ and D is a constant dependent on the parameters of the Landau free energy expansion first derived for proustite by Pokrovsky. The solutions for ϕ_l are elliptic amplitudes as in the 1-q case. The multisoliton line shapes in **Figures 15** and **16** are now obtained by inserting solutions of eqn [15] into eqn [13].

The temperature dependence of the soliton density n_{sl} on approaching T_I can be relatively well described by the Pokrovsky model, which in the simplest approximation predicts

$$n_{sl} = \frac{\pi}{2K(k)} \propto \frac{1}{\ln[(T - T_I)/T_I]}, \quad l = 1, 2, 3 \quad [16]$$

where $K(k)$ is a complete elliptic integral of the first kind. Here k, which depends on the amplitude of the modulation wave, is determined by a minimization of the free energy. The constant k is also directly related to the spatial derivative of the phase of the modulation wave $d\phi/dx$, and is thus experimentally accessible via NMR and NQR line shape measurements. A comparison between observed and theoretical ^{75}As NQR line shapes in proustite is presented in **Figure 16** and a good agreement is found.

See also: Ferroelectricity; Jahn–Teller Effect; Solid-State NMR Structural Studies of Proteins.

PACS: 61.43.–j; 61.43.Fs; 64.70.Pf; 64.70.Rh; 71.20.Tx; 71.70.Ej; 76.30.Fc; 76.60.Gv; 76.60.–k; 76.60.Pc; 77.84.Dy

Further Reading

Allemand PM, Khemani KC, Koch A, Wudl F, Holczer K, *et al.* (1991) Organic molecular soft ferromagnetism in a fullerene-C_{60}. *Science* 253: 301–303.

Blinc R and Apih T (2002) NMR in multidimensionally modulated incommensurate and CDW systems. *Progress in NMR Spectroscopy* 41: 49–82.

Blinc R and Žekš B (1974) *Soft Modes in Ferroelectrics and Antiferroelectrics*. Amsterdam: North-Holland.

Blinc R, Seliger J, Osredkar R, and Prelesnik T (1973) Double-resonance study of ^{17}O quadrupole coupling in paraelectric and ferroelectric KH_2PO_4. *Chemical Physics Letters* 23: 486–488.

Blinc R, Prelovšek P, Rutar V, Seliger J, and Žumer S (1986) Experimental observations of incommensurate phases. In: Blinc R and Levanyuk AP (eds.) *Incommensurate Phases in Dielectrics-Fundamentals*, vol. 1, pp. 143–276. Amsterdam: Elsevier BV.

Blinc R, Jeglič P, Apih T, Seliger J, Arčon D, *et al.* (2002) Beltlike C_{60}^{-} electron spin density distribution in the organic ferromagnet TDAE-C_{60}. *Physical Review Letters* 88: 086402.

Blinc R, Laguta VV, and Zalar B (2003) Field cooled and zero field cooled ^{207}Pb NMR and the local structure of relaxor $PbMg_{1/3}Nb_{2/3}O_3$. *Physical Review Letters* 91: 247601.

Burns G and Dacol FH (1983) Crystalline ferroelectrics with glassy polarization behavior. *Physical Review B* 28: 2527–2530.

Chaves AS, Barreto FCS, Nogueira RA, and Žekš B (1976) Thermodynamics of an eight-site order–disorder model for ferroelectrics. *Physical Review B* 13: 207–212.

Cross LE (1987) Relaxor ferroelectrics. *Ferroelectrics* 76: 241–267.

Höchli UT, Knorr K, and Loidl A (2002) Orientational glasses. *Advances in Physics* 51: 589–798.

Kawamoto T (1997) A theoretical model for ferromagnetism of TDAE-C_{60}. *Solid State Communications* 101: 231–235.

Kind R, Cereghetti PM, Jeitziner CA, Zalar B, Dolinšek J, *et al.* (2002) Slater ice rules and H-bond dynamics in KDP-type systems. *Physical Review Letters* 88: 195501.

Pokrovsky LV and Pryadko LP (1987) Thermodynamic description of the incommensurate phase in proustite. *Fizika Tverdogo Tela* 29: 1492–1496.

Slater JC (1941) Theory of the transition in KH_2PO_4. *Journal of Chemical Physics* 9: 16–33.

Nonlinear Dynamics and Nonlinear Dynamical Systems

G Casati, Università degli Studi dell'Insubria, Como, Italy
F Haake, Universität Duisburg-Essen, Essen, Germany

Remarks on Classical Chaos

Quantum chaos is a recent field of investigation of great importance for fundamental physics as well as in different research areas such as atomic physics, condensed matter, nuclear physics and, more recently, quantum computers. In order to understand the phenomena, it is necessary to start from the conception of dynamical chaos in classical mechanics. To simplify the discussion, only Hamiltonian (nondissipative) systems are considered as they are also more fundamental.

Dynamical chaos destroys the deterministic image of classical physics and shows that, typically, the trajectories of the deterministic equations of motion are, in some sense, random and unpredictable. Such surprising classical behavior is rooted in the

exponential instability of motion. The local instability is described by the linearized equations of motion:

$$\begin{aligned}\dot{\xi} &= \frac{\partial^2 H}{\partial q \partial p}\,\xi + \frac{\partial^2 H}{\partial p^2}\,\eta \\ \dot{\eta} &= -\frac{\partial^2 H}{\partial q^2}\,\xi - \frac{\partial^2 H}{\partial q \partial p}\,\eta\end{aligned} \qquad [1]$$

where $H = H(q, p; t)$ is the Hamiltonian, (q, p) are the coordinates of the $2f$-dimensional phase space and $\xi = \mathrm{d}q, \eta = \mathrm{d}p$ are f-dimensional vectors in the tangent space. The coefficients of the linear eqns. [1] are taken on the reference trajectory and, therefore, depend explicitly on time.

An important quantity characterizing the stability of the motion on the reference trajectory is the so-called Lyapounov exponent λ, which is defined as the limit

$$\lambda = \lim_{|t|\to\infty} \frac{1}{|t|} \ln \frac{d(t)}{d(0)} \qquad [2]$$

with $d^2 = \xi^2 + \eta^2$, the length of the tangent vector. The positivity of the Lyapounov exponent $\lambda > 0$ means exponential instability of motion.

The reason why exponentially unstable motion is called chaotic is due to the fact that almost all orbits, though deterministic, are unpredictable, in the framework of fixed finite precision for specifying data and doing calculations. Indeed, according to the Alekseev–Brudno theorem in the algorithmic theory of dynamical systems, the information $I(t)$ associated with a segment of trajectory of length t is equal, asymptotically, to

$$\lim_{|t|\to\infty} \frac{I(t)}{|t|} = h \qquad [3]$$

where h is the so-called KS (Kolmogorov–Sinai) entropy, which is positive when $\lambda > 0$. This means that in order to predict a new segment of a chaotic trajectory, one needs an additional information proportional to the length of this segment and independent of the previous length of the trajectory. In this situation, information cannot be extracted from the observation of the past history of motion. If, on the other hand, the instability is not exponential but power-law, the required information for unit time is inversely proportional to the previous length of the trajectory and, asymptotically, prediction becomes possible.

Quite clearly, prediction is possible even for a chaotic system inside a short time interval. More precisely, one can introduce the "randomness parameter" r,

$$r = \frac{h|t|}{|\ln \mu|} \qquad [4]$$

where μ is the accuracy at which a trajectory is recorded. Prediction is possible only inside the finite interval corresponding to $r < 1$ while, for $r > 1$, the motion is not distinguishable from a completely random motion.

Exponential instability implies a continuous frequency spectrum. Different parts of a trajectory are then statistically independent, and the latter property may be seen as the basis of statistical descriptions of chaotic many-body dynamics in terms of a few macroscopic variables, both with respect to equilibrium and effectively irreversible equilibration.

Is exponential instability necessary for a meaningful statistical description? While for integrable systems the instability is only linear in time, the converse is not true. On the contrary, examples of systems with linear instability have recently been found which exhibit normal diffusion and lead to normal transport coefficients.

Dynamical chaos represents a limiting case of classical motion. The opposite limiting case is given by completely integrable f-freedom systems, which possess f integrals of motion. Trajectories then wind around an f-torus with f discrete frequencies; the resulting quasiperiodic motion does not allow for "mixing": a cloud of points on a fixed f-torus draws out a bundle of trajectories such that close-by points separate only linearly. Nevertheless, the motion can be ergodic on the f-torus, provided the f frequencies are incommensurate.

Instead of using trajectories (in what might be called "Newton picture"), classical dynamics can be described in terms of distribution functions ("Liouville picture"). Distribution functions obey the linear Liouville equation. Therefore, the condition of nonlinearity for the dynamical chaos to occur refers to the description in terms of trajectories. Dynamical chaos implies the approach of any initially smooth distribution to a steady state: a cloud of points on the energy shell will start spreading out exponentially, and eventually cover that shell with uniform density. In particular, a chaotic evolution of the distribution function is nonrecurrent while, according to Poincaré's theorem, a single trajectory recurs to any neighborhood of some initial point infinitely many times.

Chaos in a Standard Model

The distinction of regular and chaotic motion survives quantization, even though the distinction

criteria change. In particular, the alternative of exponential or power-law divergence of trajectories disappears in quantum mechanics, Heisenberg's uncertainty principle forbidding the notion of trajectories. Conversely, genuine quantum criteria for chaos such as quantum dynamical localization or level repulsion which one is headed for, have no classical meaning. Not surprisingly, classical and quantum notions of chaos are complementary and reflect the basic particle–wave duality.

There is a common structure underlying both classical and quantum chaos, deeper than the more tangible but complementary phenomena. All autonomous single-freedom ($f = 1$) systems are integrable and show no chaos of any sort. For $f \geqslant 2$, integrability, and thus regular motion, requires that the dynamics can be split into f single-freedom ones; chaos can arise only if such reduction is impossible. The foregoing "definition" applies to classical point mechanics (where the issue is separability of, say, the Hamilton–Jacobi equation), to classical wave mechanics (where wave chaos comes with a nonseparable wave equation, see **Figures 5d–5f** below), and finally to quantum mechanics where Schrödinger's equation reigns.

For systems with a compact classical energy shell and, correspondingly, a discrete quantum energy spectrum, the complementarity mentioned brings about a conceptual difficulty worthy of brief discussion. How can the correspondence principle (which requires classical behavior to arise as a limiting case of quantum behavior) be reconciled with a discrete quantum energy spectrum when the limit is to be chaotic and thus characterized by a frequency continuum? The answer to that question must lie in the existence of $\hbar$-dependent timescales and, equivalently, energy scales. For quantum features to become manifest, one must resolve discrete energy levels (whose spacings vanish as $\hbar \to 0$), or sustain observation times which diverge in the formal limit $\hbar \to 0$. Examples are given below.

In order to illustrate quantum chaos, a simple model, the so-called kicked rotator is taken. Even though this model possesses a single pair of canonically conjugate variables, an angle θ and an angular momentum (alias action) p with the Poisson brackets $\{p, \theta\} = 1$, it allows for chaos due to periodic driving. Its classical Hamiltonian reads

$$H = \frac{p^2}{2I} + k\,\frac{I}{T}\cos\theta\;\delta_T(t) \qquad [5]$$

where k is the dimensionless strength and T the period of the perturbation, $\delta_T(t)$ a periodic δ-function of period T, and I a moment of inertia. The system [5] allows for a stroboscopic description by Chirikov's standard map

$$\begin{aligned} p_{\tau+1} &= p_\tau + k\,\frac{I}{T}\sin\theta_{\tau+1} \\ \theta_{\tau+1} &= \left(\theta_\tau + p_\tau\,\frac{T}{I}\right) \mathrm{mod}(2\pi) \end{aligned} \qquad [6]$$

where the integer τ counts the number of delta kicks. The motion [6] takes place on the cylindrical phase space $|p| < \infty,\ 0 \leqslant \theta \leqslant \pi$; alternatively, the motion can be restricted to the compact torus $-p_{\max}/2 < p \leqslant p_{\max}/2$, with $p_{\max}$ defining the "circumference" in the p direction. (Some authors employ the dimensionless action $P = (T/I)p$ to underscore the fact that the classical dynamics is controlled by the single parameter k.)

The kicked rotator allows for a convenient study of classical and quantum motion since it is simple and nevertheless displays the rich complexity and variety of time-dependent systems. Moreover, it provides a local description for a broad class of physical systems, including experimentally observed ones (see below).

For $k \gg 1$, the classical motion can be considered ergodic, mixing, and exponentially unstable with the Lyapounov exponent $\lambda \approx \ln(k/2)$, negligibly small stability islands apart. In particular, the action variable p displays a random-walk type of motion with $\langle p_\tau^2 \rangle = \langle p_0^2 \rangle + D\tau$ and the diffusion constant $D = (k^2/2)(I/T)^2$; of course, for the toroidal phase space, the foregoing diffusion law is valid only as long as the width in p remains smaller than the circumference, that is, up to a characteristic time $\tau_{\mathrm{cl}} \sim 2(p_{\max}T/I)^2/k^2$; thereafter, and indeed on that latter timescale, any nonsingular distribution function equilibrates to the ergodic state, $f(p, \theta, \tau) \to 1/2\pi p_{\max}$.

Quantum mechanically, the Hamiltonian [5] and the map [6] remain valid with p and θ operators. The one-step map for the wave function $\psi_\tau(\theta)$ reads $\psi_{\tau+1} = U\psi_\tau$ with the unitary Floquet operator

$$U = \exp\left(-\mathrm{i}\frac{kI}{T\hbar}\cos\theta\right)\exp\left(-\mathrm{i}p^2\frac{T}{2I\hbar}\right)\psi_\tau \qquad [7]$$

and $p = (\hbar/\mathrm{i})\partial/\partial\theta$. With the classical motion confined to a torus of volume $2\pi p_{\max}$, the wave function $\psi(\theta)$ lives in a Hilbert space whose dimension N is (the integer part of) $p_{\max}/\hbar$. (Authors employing $P = p(T/I)$ encounter the dimensionless variant $\hbar T/I$ of Planck's constant.)

Timescales of Quantum Chaos and Quantum Dynamical Localization

In order to understand the existence of different timescales, the classical and quantum evolution are compared starting from the same initial conditions. According to Ehrenfest's theorem, a quantum packet follows a beam of classical orbits as long as the packet remains narrow. During this time interval, the quantum motion is exponentially unstable and random as the classical one. However, the initial size of the quantum packet is bounded from below by the elementary quantum phase space cell of order $\hbar$, while the final size, for a bounded motion, is a characteristic value of the action variable, in this case of the order $p_{\max}$. Starting from an initial minimum-uncertainty wave packet of size $\Delta\theta_0 \Delta p_0 \backsim \hbar$, the time for exponential spreading is given by

$$\tau_r \sim \frac{\ln(p_{\max}/\hbar)}{\lambda} = \frac{\ln(p_{\max}/\hbar)}{\ln(k/2)} \qquad [8]$$

This is the random timescale introduced by Berman and Zaslavsky, often referred to as Ehrenfest time. Note that τ_r increases indefinitely as $\hbar \to 0$, in agreement with the correspondence principle.

Even after the random timescale, the quantum motion follows the classical diffusive energy growth

$$\left\langle \frac{p^2}{2I} \right\rangle_\tau - \left\langle \frac{p^2}{2I} \right\rangle_0 = \frac{1}{2I}\hbar^2 \sum_{n=-N/2}^{+N/2} n^2 |c_n|^2 \approx \frac{k^2}{4}\frac{I}{T^2}\tau \qquad [9]$$

up to some further timescale τ_b to be determined presently; here, c_n are the coefficients of the expansion of the wave function in the angular momentum representation, $p\varphi_n = \hbar n \varphi_n, -N/2 < n < N/2$.

The second timescale τ_b, at which the quantum evolution breaks away from the classical diffusion [9], is related to the phenomenon of quantum localization. The eigenvectors of the Floquet operator U given in [7] turn out localized in the p-representation with a width $\hbar\xi$; the dimensionless "localization length" ξ will be determined by a self-consistency argument presently, and self-consistency obviously requires $\xi \ll N$. Assuming localization, an initial minimum-uncertainty wave packet can involve only about ξ eigenstates of U with appreciable weight. Consequently, the width in p of the time-evolved packet $\psi_\tau = U^\tau \psi_0$ can never exceed the order of magnitude given by $\hbar\xi$; more quantitatively, the angular-momentum distribution reaches a steady state of approximately exponential form

$$|c_n|^2 \sim \exp(-2|n - n_0|/\xi) \qquad [10]$$

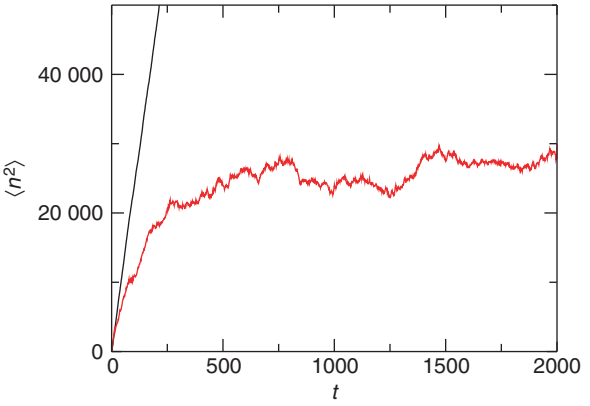


Figure 1 Classical (black) and quantum (red) mean energy as a function of time obtained by numerically iterating maps [6] and [7] for the same parameter values of **Figure 3**.

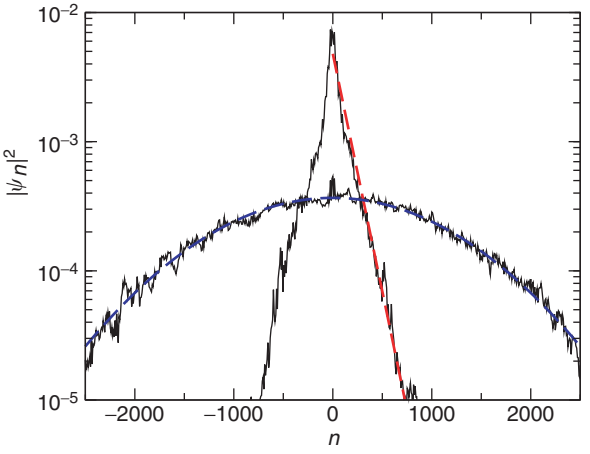


Figure 2 Quantum vs. classical probability distributions over the momentum basis at $t = 5000$. The blue curve gives the Gaussian classical distribution while the red line has slope $\xi = D$, according to [11].

In **Figures 1** and **2**, classical and quantum diffusive behavior obtained by numerical integration of the classical [6] and quantum [7] standard map are compared.

The "break time" τ_b can then be estimated as $D\tau_b \approx (\hbar\xi)^2$. An independent condition on τ_b is that the ξ eigenstates of U, appreciably excited in the wave packet in question, have a mean spacing $2\pi/\xi$ of the associated eigenphases of U. The quasiperiodic behavior of $\langle p^2 \rangle_\tau - \langle p^2 \rangle_0$ for times beyond τ_b can become manifest only for times larger than the inverse of the mean spacing mentioned, and the estimate $\tau_b \approx \xi$ thus results. The two conditions on ξ and τ_b reveal these two dimensionless quantities as roughly equal and given by the classical diffusion constant as

$$\tau_b = \xi = D/\hbar^2 \qquad [11]$$

Clearly, $\xi \ll N = p_{\max}/\hbar$ must hold for consistency (unless the classical phase space is left as the cylinder $|p| < \infty, 0 < \theta \leqslant 2\pi$). Note that both τ_b and ξ diverge in the classical limit and this is how the correspondence principle is satisfied.

Figure 3 shows the evolution of an initially narrow wave packet, demonstrating both the random (**Figures 3a–3c**) and the relaxation timescale (**Figures 3d–3f**) compared with the classical evolution.

The limitation of classical diffusion and randomness by quantum dynamical localization has been discovered in 1979 by G Casati *et al*. It can be shown that quantum dynamical localization is analogous to the famous Anderson localization of electronic motion in disordered solids. Interesting differences are that (1) in this case, localization takes place in angular-momentum space instead of coordinate space, and (2) no disorder is introduced in this model.

The above picture has been confirmed by a theoretical and experimental analysis of the microwave excitation and ionization of highly excited hydrogen atoms as well as in other atomic systems.

It is important to add that dynamical localization, even though one of the prominent manifestations of quantum chaos, is not shared by all chaotic dynamics. First of all, the rotator on the torus discussed above displays localization only under the condition $\xi \ll N$ already mentioned. Further below, the section on universality in quantum chaos will dwell on an important alternative to localization, that is, level repulsion. Here, it would be apt to point out that the kicked rotator on the "cylinder," as the prototype of dynamics with localization, can be seen as a limiting case of the so-called kicked top, which in general does not display localization. The kicked top features an angular-momentum vector $\boldsymbol{J}$ moving with conserved length, $\boldsymbol{J}^2 = \hbar^2 j(j+1) = \text{const.}$, with the quantum number j taking on integer or half-integer values; the classical limit is attained as j grows large. The classical phase space is a sphere (of radius $\sim \hbar j$). Suitable periodic kicking leads to globally chaotic classical motion on the sphere, islands of regular motion being of negligible size. Moreover, a small cloud of initial points spreads over the whole sphere within a few iterations of the classical stroboscopic map. Quantum mechanically, no localization arises: when the quasienergy eigenfunctions are visualized by their Wigner or Husimi functions, they all turn out to "live everywhere" on the sphere; the quasienergies (the eigenphases of the Floquet operator U) display level repulsion. Only very special choices of the angular-momentum dynamics entail a certain anisotropy, such that a cloud of points initially confined to an "equatorial waistband" rapidly spreads therein but diffuses only extremely slowly, by comparison, in the transversal directions. Under such conditions, the part of phase space actually explored within finite times is practically indistinguishable from the cylinder or the torus characterizing the kicked rotator. The quantum quasienergy eigenfunctions then become localized, as for the rotator, and the dimensionless localization length is much smaller than the size $2j + 1$ of the Hilbert space; conversely, no localization can be manifest when the localization length would not fit into the Hilbert space.

Dynamical Stability of Quantum Motion

In previous sections, it has been noted that analytical, numerical, and experimental results show that classical diffusive behavior can be suppressed or at least strongly inhibited by quantization. In addition to this, strong empirical evidence has been obtained that the quantum evolution is very stable, in sharp contrast to the extreme sensitivity to initial conditions and rapid loss of memory which is the very essence of classical chaos. In computer experiments, this effect leads to irreversibility. Indeed, even though the exact equations of motion are reversible, nevertheless any, however small, imprecision in solving them, such as computer round-off errors, is magnified by the exponential instability of orbits to the extent that initial conditions are effaced and reversibility is destroyed. **Figure 4** demonstrates the stability of quantum chaotic motion for the chaotic diffusive ionization of a Rydberg hydrogen atom in a microwave field. Here, the electron velocity was reversed after $\tau = 60$ periods of the microwave and the backward motion was observed. In both classical and quantum mechanics, the hydrogen atom would be expected to find its way back to the initial state. However, due to the finite computer precision, in the classical case such a return is not observed. The system retraces its history backward just for a few periods of the field and then, again, diffusive excitation occurs. Instead, in the quantum case, almost exact reversion is obtained. This result can be understood by taking into account that the random timescale is much shorter than the relaxation timescale. Yet, the accuracy of the reversal is surprising.

This property may turn out to be relevant in practical implementations of quantum computation, in which the stability of quantum computing in the presence of unavoidable imperfections in the quantum computer hardware is important for quantum information processing.

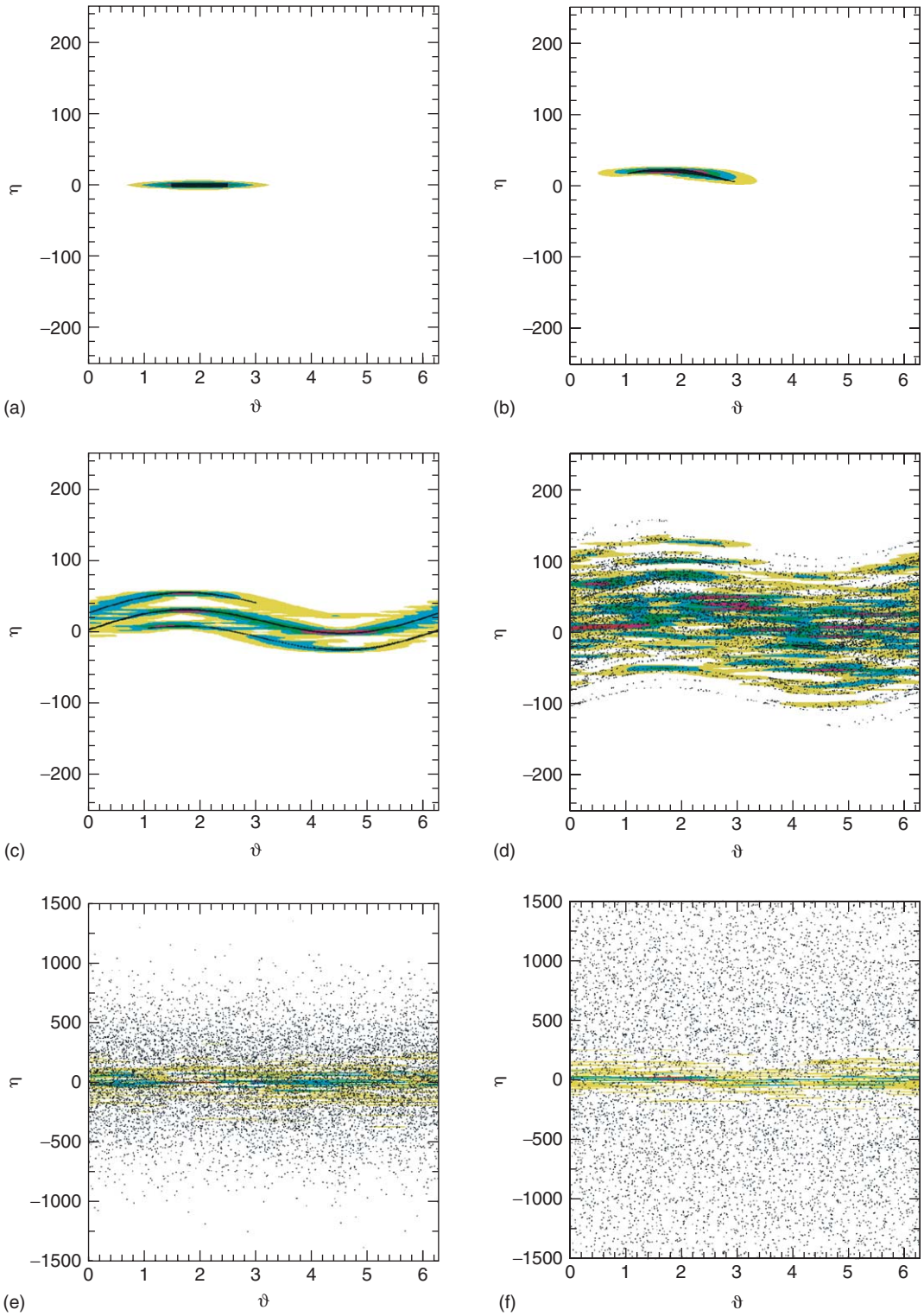


Figure 3 Classical and quantum evolution for the Chirikov standard map for parameter values $K = 5, L = 300, N = 2^{13} = 8192$. The classical evolution is computed by iterating the classical map [6] starting at time $t = 0$ with an ensemble of 10^4 orbits chosen in the interval $(\theta, n) \in [2 - 0.5, 2 + 0.5] \times [-2, 2]$. The quantum evolution is obtained by iterating the quantum map [7] starting from an initial least-spreading Gaussian wave packet of size $\Delta\theta_0 = (\Delta n_0)^{-1} \sim \sqrt{T} \approx 0.5$ centered on the initial classical density of points. The black dots represent the classical phase space density. In the quantum case, the Husimi function is plotted; different colors correspond to different intensity of the Husimi function: from red (maximum intensity) to white (zero intensity). The figure shows snapshots at different times: (a) $t = 0$, (b) $t = 1$, (c) $t = 3$, (d) $t = 10$, (e) $t = 500$, (f) $t = 5000$. Recall that the Husimi function on a given point is obtained by the projection of the evolving quantum state on the coherent state centered on that point.

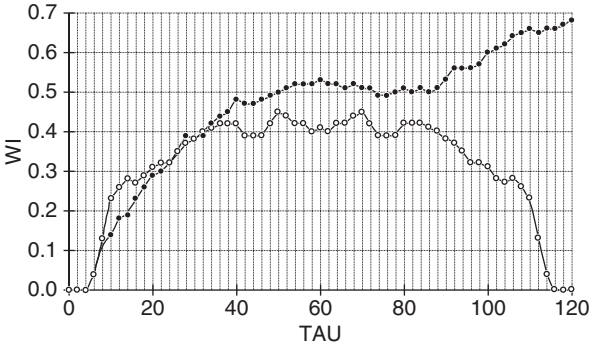


Figure 4 Classical (closed circles) and quantum (open circles) ionization probability as a function of time for a hydrogen atom under an external microwave field. Notice the perfect specular symmetry of the quantum curve about the time of reversal $\tau = 60$.

Universal Spectral Fluctuations

Interesting quantum signatures of classical chaos are encountered when quantum (quasi-) energy spectra of classically chaotic systems are scrutinized for fluctuations in the sequence of levels. One finds those levels correlated within local clusters, but statistically independent as their distance becomes much larger than the mean level spacing Δ. The correlation decay on the scale Δ is empirically found to be system-independent, within certain universality classes distinguished by symmetry. One speaks of universal spectral fluctuations.

A popular indicator of spectral universality is the level spacing distribution $P(S)$, obtained by throwing the spacings of neighboring energy levels into a histogram. **Figure 5** depicts such histograms for six different dynamics. In all cases, the histogram is a rugged variant of one and the same smooth curve, the so-called Wigner distribution which is discussed presently. One may note that this Wigner distribution rises linearly out of the origin and, with growing spacing S, passes through a maximum and eventually falls off like a Gaussian; negligible probability is thus assigned to spacings very small compared to the mean spacing, such that one can say that levels repel one another. Level repulsion is discussed below.

The dynamics referred to in **Figure 5** are all time-reversal invariant and belong to the so-called orthogonal universality class for which the antiunitary time-reversal operation squares to unity, $\mathscr{T}^2 = 1$. Two further universality classes are known as the symplectic ($\mathscr{T}$- invariance with $\mathscr{T}^2 = -1$, as typical for systems with half-integer spin) and the unitary (no $\mathscr{T}$- invariance) ones. More recently, seven more classes were discovered and shown to have realizations in disordered electronic systems, some of them superconducting. The classification is now considered complete inasmuch as the ten classes are one-to-one with the ten symmetric spaces sorted out by Cartan.

Universality goes further. As revealed by experiments and numerical work, fluctuations on the scale Δ turn out to be the same for the energy spectra of Hamiltonians as for the quasienergy spectra of Floquet operators, provided full classical chaos and the same symmetries reign. While proving that statistical similarity of autonomous and periodically driven dynamics is difficult, an intuitive understanding can be reached as follows. An energy spectrum of N levels E_n, $n = 1, \ldots, N$, of a Hamiltonian H corresponds to a quasienergy spectrum of the time evolution operator $e^{-iHt/\hbar}$, with quasienergies $E_n t/\hbar$. If the time t is chosen such that these N quasienergies just fit into the interval $[0, 2\pi]$, the quasienergy spectrum in question becomes similar to the Floquet spectra of periodically driven systems with period t. Only one difference is obvious: the lowest level E_1 of H gives rise to the lowest quasienergy $E_1 t/\hbar$, and similarly the highest energy E_N to the highest quasienergy $E_N t/\hbar$; even though close modulo 2π and being nearest neighbors as quasienergies, these two levels do not repel one another. In the quasienergy spectrum of a periodically driven system, on the other hand, there are no lowest and highest levels, and level repulsion is manifest for all neighboring levels. However, that difference can hardly be expected to play any role for the universal spectral fluctuations and correlations: as already emphasized, the universal fluctuations are of local character, that is, are supported by clusters of relatively few levels extending over a few mean level spacings. Many such local clusters exist for $N \gg 1$, such that exceptional ones at the edges of a spectrum have negligible weight. At any rate, with appeal to the statistical equivalence of energies and quasienergies, the subsequent discussion of universal spectral fluctuations is confined to energy spectra of autonomous dynamics (with finite N).

Experiments on microwave resonators and elastic vibrations of quartz blocks have brought into evidence fluctuations in the respective frequency spectra (see **Figures 5d–5f**) coinciding with the ones found for quantum chaos, the classical nature of the observed electromagnetic and elastic waves notwithstanding. This highly important finding suggests that one could see quantum chaos as a special case of a wider class of phenomena, wave chaos. Whichever wave phenomenon is looked at, as long as the pertinent ray limit (the classical limit in the case of quantum matter waves) displays chaos in the sense of positive Lyapounov exponents, it is now known to display universal fluctuations in the pertinent frequency or energy spectrum. The wave equations

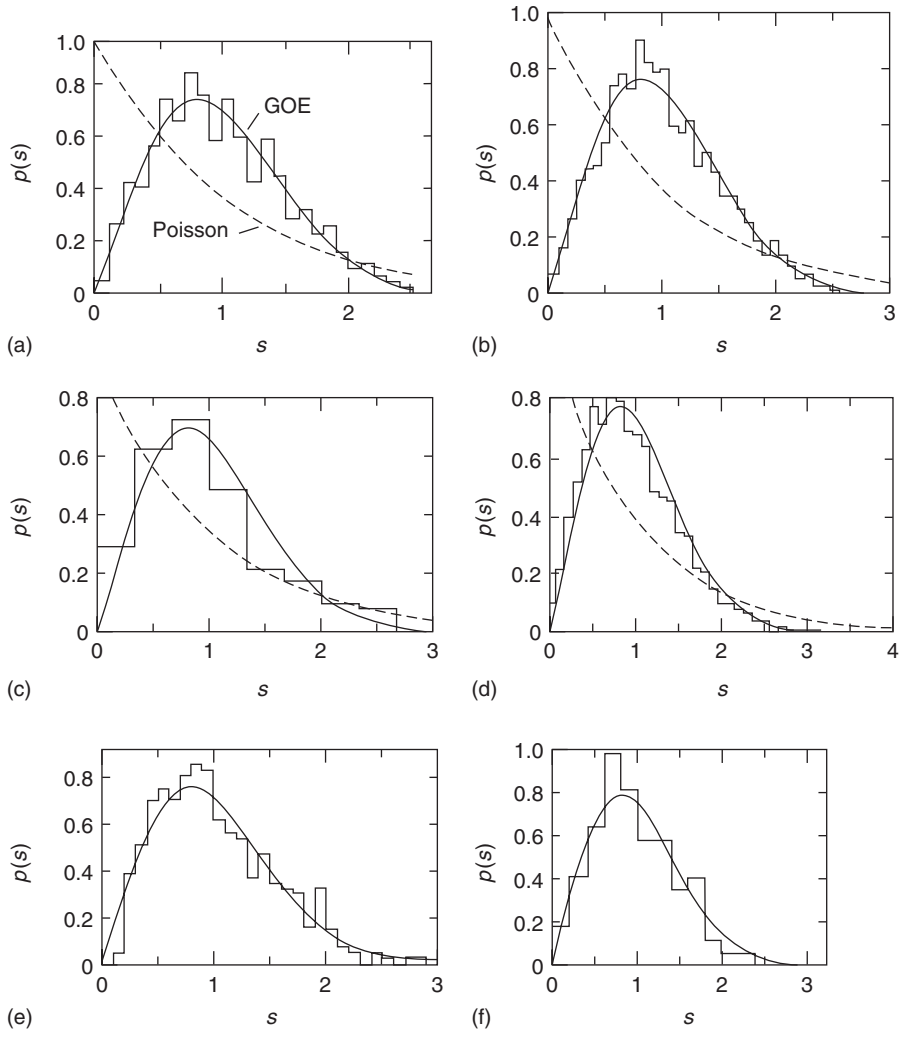


Figure 5 Level spacing distributions for (a) the Sinai billiard, (b) a hydrogen atom in a strong magnetic field, (c) an NO_2 molecule, (d) a vibrating quartz block shaped like a three-dimensional Sinai billiard, (e) the microwave spectrum of a three-dimensional chaotic cavity, (f) a vibrating elastic disk shaped like a quarter stadium. (Courtesy of Stöckmann H-J (1999) *Quantum Chaos*: *An Introduction*. Cambridge: Cambridge University Press.)

differ from case to case, but common to all chaotic waves is nonseparability: the respective boundary value problems defy solution by separation into one-dimensional problems. Such nonseparability is, of course, shared by wave and ray chaos, as well as by quantum and classical chaos.

Finally, and perhaps surprisingly, universal spectral fluctuations can also be found in a completely different context, for random matrices. Indeed, when throwing dice to specify the elements of Hermitian (or unitary) matrices, respecting nothing but symmetry (in the sense of the classes mentioned), one finds the level spacing distribution as well as other indicators of spectral fluctuations to be the same as for Hamiltonians (or Floquet operators) from the same symmetry class.

Random-Matrix Theory

Random matrices play an important role in the theory of quantum and wave chaos since certain random-matrix ensembles yield analytic expressions for many indicators of universal spectral behavior. For the three "classic" classes (orthogonal, unitary, symplectic), Wigner and Dyson have introduced the "Gaussian" ensembles (GOE, GUE, GSE) of Hermitian $N \times N$ matrices and the "circular" ensembles (COE, CUE, CSE) of unitary $N \times N$ matrices. The original motivation for investigating such random-matrix ensembles was unrelated to chaos since the latter notion had not yet seen the light of the world. Wigner and Dyson rather set out to understand the fluctuations in nuclear spectra and, for lack of good

Hamiltonians describing the innards of nuclei, resorted to random Hermitian matrices. Similarly, scattering processes involving nuclei and their constituents were modeled by means of random unitary matrices.

In respect of limits of space, this discussion is restricted to the three Gaussian ensembles GOE, GUE, and GSE which model Hamiltonians of the respective symmetries. The ensembles in question are defined by the following Gaussian density of the independent matrix elements of the Hermitian H,

$$W(H) = A\mathrm{e}^{-B\ \mathrm{Tr}\ H^2} \qquad [12]$$

where the constants A and B are fixed by normalization and the choice of a unit of energy.

The number $\mathcal{N}$ of independent elements in a matrix H differs for the three classes. In the orthogonal class, the matrices H are real symmetric and $N \times N$ such that $\mathcal{N} = N(N+1)/2$ independent real elements arise. The unitary class has complex Hermitian $N \times N$ matrices H, whose elements have $\mathcal{N} = N^2$ independent real and imaginary parts. Hamiltonians from the symplectic class have a more complicated matrix structure, often called quaternion real. Quaternion real $2N \times 2N$ matrices H consist of N^2 2×2 blocks; each block is formed with a pair of basis vectors $\{\psi_i, \mathcal{T}\psi_i, i = 1, \ldots, N\}$, which are orthogonal due to $\mathcal{T}^2 = -1$; eigenvectors forming different blocks are mutually orthogonal. All eigenvalues are doubly degenerate (Kramers' degeneracy follows from $\mathcal{T}^2 = -1$). To have N such pairs of levels, the matrix H must be $2N \times 2N$; its elements can be shown to have $\mathcal{N} = N(2N-1)$ independent real and imaginary parts. For all three ensembles, the normalization of the densities, $\int \mathrm{d}H\ W(H) = 1$, involves an $\mathcal{N}$-fold integral over the (real and imaginary parts of the) independent matrix elements with the measure $\mathrm{d}H$, the product of the respective differentials.

One may note that the aforementioned structures of H, that is, real symmetric, complex Hermitian, and quaternion real, as well as the distribution [12] are preserved under, respectively, orthogonal, unitary, and symplectic transformations of H. In fact, it is these invariances which have led to the names of the respective universality classes and their associated matrix ensembles.

The $\mathcal{N}$ independent real parameters in a random matrix H can be organized in two subsets, the N eigenvalues and $\mathcal{N} - N$ parameters, characterizing the eigenvectors (or, equivalently, the diagonalizing orthogonal, unitary or symplectic transformations). Upon integrating out the latter, one finds the eigenvalue distributions

$$\rho(E_1, \ldots, E_N) = A' \left(\prod_{i<j}^{1,\ldots,N} |E_i - E_j|^\beta \right) \times \exp\left(-B' \sum_{i=1}^{N} E_i^2 \right) \qquad [13]$$

here, the parameter β takes the values $1, 2, 4$ in the orthogonal, unitary, and symplectic class, respectively.

The simplest nontrivial ensemble average arises for the distribution of the level spacing for $N = 2$. One finds the "Wigner surmizes,"

$$P_\beta(S) = A'' S^\beta \mathrm{e}^{-B'' S^2} \qquad [14]$$

with the constants A'', B'' again fixed by normalization, $\int \mathrm{d}S\ P(S) = 1$, and the choice of a unit of energy, like $\int \mathrm{d}S\ SP(S) = 1$. The parameter β is now revealed as the degree of level repulsion, since the spacing distribution vanishes like S^β for $S \to 0$; one speaks of linear, quadratic, and quartic level repulsion in the orthogonal, unitary, and symplectic classes, respectively. The orthogonal case $P_1(S)$ is plotted as the smooth curve in **Figure 5**. Surprisingly, that curve turns out an excellent approximation to the spacing distribution of large spectra. Similarly good and surprising is the agreement of the Wigner surmizes $P_2(S)$ and $P_4(S)$ with the spacing distributions of large spectra in the unitary and symplectic class.

When the spacing distributions for $N > 2$ are calculated for the three Wigner–Dyson ensembles, no simple analytic expressions arise. In particular, the asymptotic densities for $N \to \infty$ are known, both as converging infinite products and series. In confirmation of numerical and empirical findings, the difference between the Wigner surmizes based on $N = 2$ and the asymptotic spacing distributions based on $N \to \infty$ is small: the integrated root-mean-square difference is found as

$$\left\{ \int_0^\infty \mathrm{d}S (P_\beta^{(N=2)}(S) - P_\beta^{(N\to\infty)}(S))^2 \right\}^{1/2} < 1.6 \times 10^{-4}$$

for all three Wigner–Dyson classes. Such insensitivity of the spacing distributions to the matrix size gives great practical importance to the Wigner surmizes. To verify that a single large spectrum is more faithful to the asymptotic distribution $P_\beta^{(N\to\infty)}(S)$ than to the Wigner surmize $P_\beta^{(N=2)}(S)$, several thousands of levels must be available.

In order to appreciate the faithfulness of large spectra to the predictions of the random-matrix theory (RMT) based on ensembles of two-level spectra,

the locality of correlations among levels may again be recalled. It appears that correlations between neighboring levels dominate the local correlations. It in fact turns out, that at least the universal small-S behavior, $P(S) \propto S^{\beta}$, can be understood by applying the familiar perturbation theory of nearly degenerate levels to pairs of neighboring levels.

Level repulsion of the various degrees just characterized is, in fact, a highly important quantum signature of chaos, since classically integrable dynamics with two or more freedoms usually display level clustering. The pertinent spacing distribution $P(S) = \mathrm{e}^{-S}$ is often referred to as Poissonian since a fully uncorrelated sequence of levels generated by a Poissonian random process would yield just that exponential.

Most classical dynamics have a mixed phase space, where chaotic and regular trajectories coexist. The universal spectral fluctuations prevail provided the classical islands of regular motion are overwhelmed in relative size by a single "chaotic sea."

Recently, exceptional dynamics were discovered that do display full classical chaos but no quantum level repulsion; they have quantum symmetries and conservation laws (the so-called Hecke symmetries), which disappear in the limit $\hbar \to 0$.

In contrast to the spectral fluctuations on the scale Δ, the energy dependence of the mean density of levels on larger scales is system specific, as is, of course, the precise location of the N levels.

This brief exposition of level repulsion and localization may be concluded by stating that these two quantum manifestations of chaos are, in a sense, mutually exclusive. Under conditions of localization, (quasi-)energy eigenfunctions, whose centers are much further apart than a localization length, have exponentially small overlap; their (quasi-)energies then have no reason to repel one another. When building up the spacing distribution $P(S)$ from all the nearest-neighbor spacings in the spectrum, one must expect and does indeed find the exponential distribution characteristic of statistically independent levels. However, upon building different distributions $P(S)$, each from a group of eigenfunctions with strong mutual overlap, those individual distributions do reveal universal level repulsion; by superposing the distributions from different groups, the statistical basis for comparing with RMT can be improved.

Why Universality?

Universal spectral fluctuations were conjectured early on in the development of the field of quantum chaos and later supported by a wealth of numerical and empirical data, and, as already indicated, confronted with exceptions. No mathematically hard formulation of the conjecture is available to date, leave alone proof. Nevertheless, some understanding has arisen along several lines of thought.

First progress was brought about by "parametric level dynamics," the change of energy levels $E_i(\lambda)$ and eigenfunctions $\psi_i(\lambda)$ under variation of the weight λ of a perturbation in the Hamiltonian $H = H_0 + \lambda H_1$. A closed system of $\mathcal{N}$ ordinary first-order differential equations for the N eigenvalues and $\mathcal{N} - N$ parameters representing the eigenfunctions can be established which, perhaps surprisingly, takes the form of classical Hamiltonian equations for a certain N-particle system of the Calogero–Moser–Wojciechowski type. The parameter λ plays the role of time, and the eigenvalues E_i are the coordinates of the fictitious particles. Such dynamics had been noted previously in the mathematical literature for being integrable, their nontrivial two-body interaction notwithstanding. Before their independent rediscovery in the context of parametric level dynamics, they were not paid much attention in physics; now, they may be seen as an elegant and closed form of the perturbation theory since indeed, their solution in terms of a power series in λ is nothing but the perturbation theory with respect to H_1. More importantly for the present context, the fictitious N-body system invites a statistical description. Inasmuch as the spectral fluctuations in the original quantum system do not change as λ varies, equilibrium statistical mechanics for the fictitious "gas" is called for. One may employ a generalized canonical ensemble which nails down all constants of the motion in the mean, through appropriate Lagrange parameters, and allows for equipartition otherwise. In fact, such a canonical ensemble turns out to yield the spacing distribution as coinciding with that of the RMT, up to corrections of the order $1/N$. One might argue that nothing is gained by replacing one ensemble (RMT) with another. However, the generalized canonical ensemble of level dynamics yields ensemble averages equalling λ averages, due to the ergodic behavior of the fictitious gas in the part of its phase space left accessible by the constants of the motion. For λ averages to equal ensemble averages, the λ averages must be extended over a sufficiently large parameter "window" $\delta\lambda$; that window must be taken somewhat larger than a collision "time" λ_{coll} for the fictitous gas, that is, the mean "temporal" separation between "avoided collisions" of neighboring particles in the fictitious gas; for the original quantum levels, λ_{coll} is the mean parameter distance between subsequent avoided crossings of neighboring levels. Inasmuch as that collision time is found to scale with the number N of "particles" such that $\lambda_{\mathrm{coll}} \to 0$ for $N \to \infty$, and since $N \to \infty$ in the classical

limit $\hbar \to 0$, one concludes that $\delta\lambda$ can be considered vanishingly small classically. Consequently, the canonical ensemble in question and thus RMT also have a status like that of equilibrium statistical mechanics for a stationary N-body system.

Further support for universality *à la* RMT comes from Gutzwiller's semiclassical periodic-orbit theory. The density of levels $\rho(E) = \sum_{i=1}^{N} \delta(E - E_i)$ is given there as a formal sum over classical periodic orbits, $\rho \sim \sum_\gamma A_\gamma \cos(S_\gamma/\hbar)$, where S_γ and A_γ are the action and stability coefficients of the γth orbit. The two-point correlator of the level density thus becomes a sum over pairs of orbits, each pair represented by a phase factor $e^{i(S_\gamma - S_{\gamma'})/\hbar}$. Since the action difference $S_\gamma - S_{\gamma'}$ "knows" no quantum unit $\hbar$, most pairs interfere destructively in their contribution to the correlation function. However, a paradigmatic discovery by Sieber and Richter as well as successive work brought to light families of pairs of orbits which do make finite contributions and, in fact, give the correlation function predicted by RMT.

The orbit pairs in question have periods of the order of the Heisenberg time T_H. The two orbits in a relevant pair are practically indistinguishable nearly everywhere. To understand the minute but consequential difference within a pair, it is useful to recall that an orbit cannot cross itself in phase space; however, a long orbit can, even must, have close "self-encounters," in which two orbit stretches come very close (possibly up to time reversal in $\mathscr{T}$-invariant systems). Given the fact that the smallest approaches that are discernible quantum mechanically have an order magnitude given by Planck's constant, the durations of such self-encounters have the typical order of the Ehrenfest time T_E'. On either side of a self-encounter, the orbits in a relevant pair have a long loop within which both orbits practically coincide; the two orbits differ in the order of traversals of the loops due to different connections within the short self-encounters. It is due to the relative smallness of the Ehrenfest time ($T_E/T_H \to 0$ in the limit $\hbar \to 0$) that the typical action difference can be of the order of the Planck's constant, thus enabling the orbit pairs to contribute constructively to the correlation function. The frequency of occurrence of such orbit pairs can be calculated by assuming ergodicity. It is thus clear that the latter universal property of classical chaos is intimately related to the universal spectral fluctuations characteristic of quantum chaos.

Conclusions

Chaos, both classical and quantum, requires two or more degrees of freedom and impossibility to reduce the dynamics to single-freedom ones.

From a quantum point of view, classical dynamical chaos appears as a transient, observable until quantum effects due to the discreteness of (quasi-)energy spectra take over. The correspondence principle is respected inasmuch as the breakaway from exponentially unstable classical motion to quantum quasiperiodicity takes place at a time formally diverging as $\hbar \to 0$. Quite analogous is the relation of classical waves to their ray limits.

Classical and quantum chaos have complementary manifestations reflecting the familiar particle–wave duality. Intrinsically quantum (or wave) manifestations of chaos are quantum dynamical localization and universal spectral fluctuations. Conditions under which these mutually exclusive phenomena arise have been discussed.

Universal spectral fluctuations can be modeled by suitable ensembles, of large random matrices. The status of random-matrix ensembles, relative to individual dynamics, bears analogies to the status of the ensembles, employed by equilibrium statistical mechanics, relative to individual many-body systems. Universal spectral statistics are beginning to be understood in terms of Gutzwiller's semiclassical "periodic-orbit theory."

See also: Lattice Dynamics: Anharmonic Effects.

PACS: 05.45. – a; 05.45.Mt

Further Reading

Alekseev VM and Yakobson MV (1981) *Physics Report* 75: 287.
Arndt M, Buchleitner A, Mantegna RN, and Walther H (1991) *Physical Review Letters* 67: 2435.
Bay-field JE, Casati G, Guarneri I, and Sokol DW (1989) *Physical Review Letters* 63: 364.
Benenti G, Casati G, and Strini G (2004) *Principles of Quantum Computation and Information*, vol. I: Basic Concepts (Singapore: World Scientific).
Berman GP and Zaslavsky GM (1978) *Physica A* 91: 450.
Bohigas O, Giannoni MJ, and Schmit C (1984) *Physical Review Letters* 52: 1.
Casati G and Chirikov BV (1995) *Quantum Chaos*. Cambridge: Cambridge University Press.
Casati G, Chirikov BV, Ford J, and Izrailev FM (1979) *Lectures Notes in Physics* 93: 334.
Casati G, Valz-Gris F, and Guarneri I (1980) *Letters Nuovo Cimento* 28: 279.
Casati G, Chirikov BV, and Shepelyansky DL (1984) *Physical Review Letters* 53: 2525.
Casati G, Chirikov BV, Guarneri I, and Shepelyansky DL (1986) *Physical Review Letters* 56: 2437.
Fishman S, Grempel D, and Prange R (1982) *Physical Review Letters* 49: 509.
Galvez EJ, Sauer BE, Moorman L, Koch PM, and Richards D (1988) *Physical Review Letters* 61: 2011.
Gutzwiller M (1990) *Chaos in Classical and Quantum Mechanics*. New York: Springer.

Haake F (2000) *Quantum Signatures of Chaos*. Berlin: Springer.
Haake F and Shepelyansky DL (1988) *Europhysical Letters 5*: 671.
Lichtenberg A and Lieberman M (1983) *Regular and Stochastic Motion*. Berlin: Springer.
Mehta ML (1967) *Random Matrices and the Statistical Theory of Energy Levels*, 2nd edn. 1991. New York: Academic.
Müller S, Heusler S, Braun P, Haake F, and Altland A (2005) arXiv/nlin.CD/0503052.
Sieber M and Richter K (2001) *Physica Scripta T* 90: 128.
Turek M, Spehner D, Müller S, and Richter K arXiv/nlin.CD/0409012.
Zirnbauer M (1996) *Journal of Mathematical Physics* 37: 4986; Altland A and Zirnbauer M (1997) *Physical Review B* 55: 1142.

Nonlinear Optics

V Degiorgio and I Cristiani, Università di Pavia, Pavia, Italy

Introduction

The propagation of optical waves inside a medium is described by the wave equation:

$$\nabla^2 \boldsymbol{E} = \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} + \mu_o \frac{\partial^2 \boldsymbol{P}}{\partial t^2} \qquad [1]$$

where $\boldsymbol{E}(\boldsymbol{r}, t)$ is the electric field and $\boldsymbol{P}(\boldsymbol{r}, t)$ is the macroscopic polarization. Traditional optics is based on the simplified assumption that $\boldsymbol{P}$ is linearly related to $\boldsymbol{E}$ through a differential equation describing the noninstantaneous response of the medium. In the frequency domain, $\boldsymbol{P}(\boldsymbol{r}, \omega) = \varepsilon_0 \chi(\omega) \boldsymbol{E}(\boldsymbol{r}, \omega)$, where $\boldsymbol{E}(\boldsymbol{r}, \omega)$ and $\boldsymbol{P}(\boldsymbol{r}, \omega)$ are the electric field and polarization Fourier transforms, and $\chi(\omega)$ is the frequency -dependent electric susceptibility that is related to the index of refraction of the material through the expression: $n(\omega) = \sqrt{1 + \chi(\omega)}$. Actually the relation between $\boldsymbol{P}$ and $\boldsymbol{E}$ is intrinsically nonlinear and can be considered approximately linear when the optical field is much smaller than the internal field seen by the electron. Whereas in the past the optical intensities available with conventional sources were too weak to explore optical nonlinearities, it is now very easy to generate nonlinear optical effects by exploiting laser beams. The term nonlinear optics should apply to all the physical situations in which the linear approximation is not adequate. This article is limited to a treatment of those nonlinear optical interactions that are nonresonant. Resonant interactions, such as those responsible for laser behavior, optical bistability, or electromagnetically induced transparency, are discussed elsewhere in this encyclopedia.

Nonlinear optics has enhanced the understanding of fundamental light–matter interactions and provided new powerful methods for the investigation of materials. At the same time, nonlinear optics has a strong impact on technology as it offers the way to develop coherent radiation sources, emitting at new frequencies that are not usually available with lasers, and offering tunability from ultraviolet to visible and near infrared. Important applications also concern optical devices performing the same functions provided at radio frequencies by electronic devices, such as modulators, switches, and logical circuits.

Physical Origin of Nonlinearity

A simple physical model describing the relation between $\boldsymbol{P}$ and $\boldsymbol{E}$ is represented by the Lorentz model that treats a bound electron as a harmonic oscillator. Clearly, the assumption that the electron is in a parabolic potential well is valid only for small displacements from the equilibrium position. As shown in **Figure 1**, if the exciting field is strong enough, the electron displacement will also probe a range of distances over which the potential is no longer parabolic. The potential can be written as a Taylor expansion in the displacement x: $V(x) = V_0 + \alpha x^2 + (1/2)\beta x^3 + (1/3)\gamma x^4 + \cdots$. It can be easily seen by a perturbative approach that the

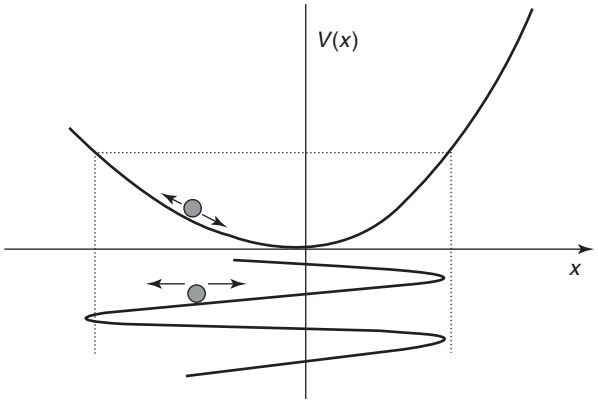


Figure 1 Particle motion in an asymmetric potential well under the action of a sinusoidal field: the oscillation is not symmetric; thus, it contains higher-harmonic contribution.

cubic term in $V(x)$ can give rise to a polarization term proportional to the square of the field, and the quartic term gives a term proportional to the cube of the field.

Note that a cubic term in $V(x)$ can exist only if the material structure is not centrosymmetric, as it happens with a crystal having a noncentrosymmetric unit cell or a system of oriented polar molecules. On the macroscopic scale the relation between $\boldsymbol{P}(\boldsymbol{r}, \omega)$ and $\boldsymbol{E}(\boldsymbol{r}, \omega)$ can be written as

$$\boldsymbol{P} = \varepsilon_0 \chi^{(1)} \boldsymbol{E} + \varepsilon_0 \chi^{(2)} \boldsymbol{E}^2 + \varepsilon_0 \chi^{(3)} \boldsymbol{E}^3 + \cdots \qquad [2]$$

Materials presenting a second-order nonlinearity, when illuminated with a beam at angular frequency ω, can generate a second-harmonic beam at frequency 2ω. If two beams at frequencies ω_1 and ω_2 are sent into the material, a third beam at the sum or difference frequency $\omega_3 = \omega_1 \pm \omega_2$ can be generated. Such a process is used in optical parametric amplifiers and oscillators. In the case of third-order materials, besides the generation of new frequencies through various frequency-mixing processes, phenomena such the optical Kerr effect, self-focusing, self-phase modulation, and optical phase-conjugation can be observed and exploited for the realization of optical devices. In addition, stimulated Raman and Brillouin scattering also belong to the family of third-order effects.

Second-Order Nonlinear Effects

Second-Harmonic Generation

In order to illustrate the basic aspects of nonlinear interactions, it is better to start with the simplest situation: a second-order material illuminated by a monochromatic plane wave propagating along the z-axis – $E_1 \exp[\mathrm{i}(\omega_1 t - k_1 z)]$. It is easy to see from eqn [2] that such a wave gives rise to a nonlinear polarization oscillating at $2\omega_1$. In turn, the nonlinear polarization, once inserted into eqn [1], behaves like a source for a second-harmonic (SH) optical wave. Limiting the consideration to stationary waves in the slowly-varying-envelope-approximation and expressing the pump and SH fields E_1 and E_2 as $E_1 = A_1(z)\exp(\mathrm{i}k_1 z)$ and $E_2 = A_2(z)\exp(\mathrm{i}k_2 z)$, respectively, the evolution of the SH generation phenomenon can be described through the following equations:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z} A_1(z) &= \mathrm{i}\frac{\omega_1}{cn_1} d_{\mathrm{eff}} A_2(z) A_1^*(z) \exp(-\mathrm{i}\Delta k z) \\ \frac{\mathrm{d}}{\mathrm{d}z} A_2(z) &= \mathrm{i}\frac{\omega_1}{cn_2} d_{\mathrm{eff}} A_1^2(z) \exp(\mathrm{i}\Delta k z) \end{aligned} \qquad [3]$$

where $d_{\mathrm{eff}} = \chi^{(2)}/2$ is called the effective nonlinear coefficient, $\Delta k = k_2 - 2k_1$ is the wave vector mismatch, c is the light velocity, and n_1 and n_2 represent the refractive indices experienced by the pump and SH fields during the propagation through the nonlinear medium. It is important to remark that, in scientific literature, the nonlinear properties of materials are usually expressed in terms of the d_{eff} coefficient values. It can be seen from eqns [3] that the condition for efficient SH generation is: $\Delta k = 0$. Such a condition is called the phase-matching condition. It expresses the requirement that the nonlinear polarization wave travels at the same phase velocity as the SH wave. It should be noted that the process consists in an exchange of energy (and momentum) between waves in which the nonlinear medium acts as a catalyst. If the process is described by a quantum-mechanical approach, two photons of the fundamental wave are converted into an SH photon; the phase-matching condition corresponds to momentum conservation in the nonlinear process.

In the case $\Delta k = 0$, the propagation equations can be solved analytically, giving the SH intensity as a function of the propagation distance z:

$$I_2(z) = I_1(0)\tanh(z/L_{\mathrm{SH}}) \qquad [4]$$

where the constant L_{SH} is inversely proportional to the nonlinear coefficient of the medium and to the square root of the input intensity. Equation [4] shows that the input intensity can be fully converted into the SH provided that the propagation distance is sufficiently long (the crystal length L should be larger than L_{SH}). As an example, by using a lithium niobate crystal with a second-order coefficient

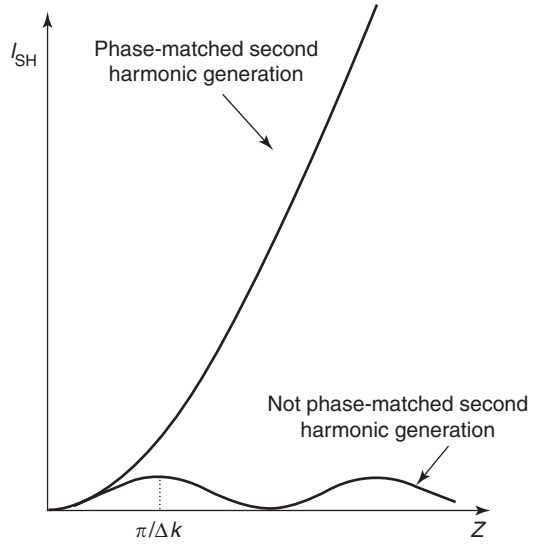


Figure 2 SH intensity as a function of propagation distance in a nonlinear crystal.

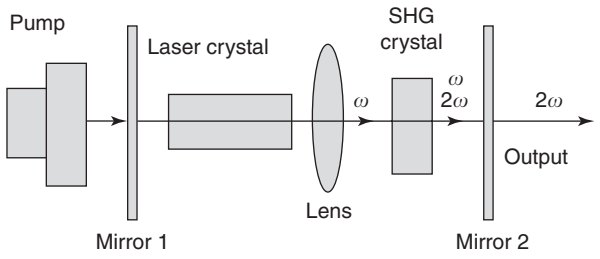


Figure 3 Scheme of an SH laser: mirrors 1 and 2 have high reflectivity for the frequency ω corresponding to the laser transition, but are transparent for 2ω.

$d_{\text{eff}} = 5\,\text{pm}\,\text{V}^{-1}$ and a neodymium laser ($\lambda = 1064$ *nm*) beam with an intensity of $10\,\text{MW}\,\text{cm}^{-2}$, $L_{\text{SH}} = 1\,\text{cm}$. When the phase-matching condition is not satisfied, $I_2(z)$ is an oscillating function of z (see **Figure 2**): the SH intensity achieves a maximum after a distance called coherence length $l_c = \pi/\Delta k$ and vanishes at $z = 2l_c$.

SH generation of a laser beam becomes particularly efficient if the nonlinear crystal is placed inside a laser cavity (see **Figure 3**). At present, such an approach represents one of the most convenient methods of generating coherent green and blue light by using near infrared solid-state lasers. As an example, a 532 nm green light is efficiently generated by intracavity frequency doubling of a neodymium laser.

Phase Matching

Recalling that the modulus of $\boldsymbol{k}$ is related to the angular frequency by the relation $k = (\omega/c)n(\omega)$, the phase-matching condition $\Delta k = 0$ requires that the index of refraction of the second harmonic be equal to that of the fundamental. In normal cases, n is an increasing function of frequency, so that phase matching is not satisfied for beams having the same polarization. A very clever method of obtaining phase matching is that of exploiting the crystal birefringence to compensate for the material dispersion (see **Figure 4**). Recall that, in a crystal possessing an optical axis, the propagation velocity of a linearly polarized plane wave depends on its polarization direction: a wave polarized perpendicularly to the optical axis sees the "ordinary" index of refraction n_o, whereas a wave polarized along the optical axis sees the "extraordinary" index n_e. A typical nonlinear crystal is lithium niobate, a uniaxial crystal presenting negative birefringence, that is, $n_o > n_e$. It can be shown that there is a frequency ω_{pm} for which $n_o(\omega_{\text{pm}}) = n_e(2\omega_{\text{pm}})$. In such a situation, an ordinary input beam at frequency ω_{pm}, propagating perpendicularly to the optical axis, can be efficiently converted into an extraordinarily polarized SH beam.

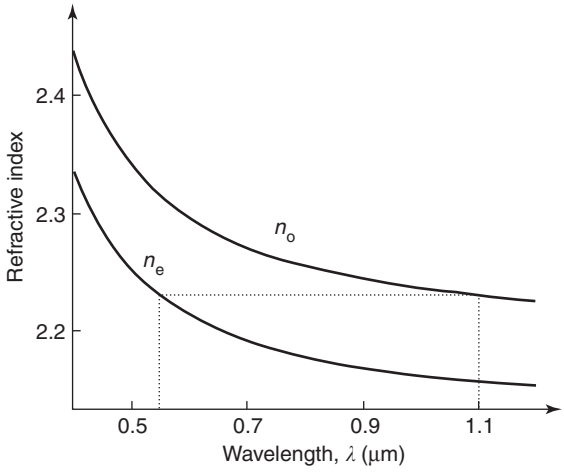


Figure 4 Birefringence phase matching.

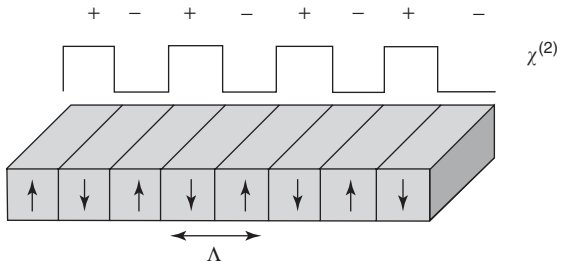


Figure 5 Periodically poled ferroelectric crystal presenting a uniform $\chi^{(1)}$ and an oscillating $\chi^{(2)}$.

Noting that the index of refraction of the extraordinary beam can be changed continuously between n_o and n_e by changing the angle θ between $\boldsymbol{k}$ and the optical axis, ω_{pm} can be tuned by appropriately rotating the crystal.

Since the indices of refraction depend on temperature, it is possible to tune ω_{pm} finely by varying T. More degrees of freedom are added to the SH experiment if two input beams having the same frequency, but different direction or different polarization, are considered.

From the point of view of applications, a very important trend that emerged in recent years is that of considering nonlinear interactions in microstructured materials and in waveguides. It goes beyond the scope of this article to approach such a subject; however, mention may be made here that a microstructured material, in which the value of the $\chi^{(2)}$ coefficient varies with a spatial periodicity Λ, can exchange momentum with the propagating waves in multiples of the unit $2\pi/\Lambda$, so that, in principle, phase matching can be achieved at any desired frequency by appropriately choosing the period Λ (see **Figure 5**). Concerning waveguided propagation, the main difference with respect to bulk propagation is that the

waveguide can support different spatial modes possessing different phase velocities even if they have the same frequency; such a "modal dispersion" also opens new possibilities of phase matching.

This section concludes with an important consideration: whereas the "ideal" phase-matching condition is $\Delta k = 0$, the "practical" condition in a crystal of finite length L is $\Delta kL < \pi$, or, equivalently, $l_c > L$. As a consequence, there is a frequency interval, centered on ω_{pm}, in which efficient SH generation can be achieved. The existence of a nonzero spectral acceptance makes it possible to use short pulses, instead of stationary beams, in nonlinear optical experiments.

Parametric Amplification and Oscillation

If two beams, possessing different frequencies ω_1 and ω_2, and different wave vectors $\boldsymbol{k_1}$ and $\boldsymbol{k_2}$, are sent into a second-order nonlinear crystal, the nonlinear polarization can give rise to new frequencies representing the SH of each beam or the sum/difference of the input frequencies. By choosing a specific phase-matching condition, only one process is selected among all possible processes. If the difference-frequency generation (DFG) is selected, and it is assumed that $\omega_1 > \omega_2$, a new beam at frequency $\omega_3 = \omega_1 - \omega_2$ and wave vector $\boldsymbol{k_3}$ is obtained, provided that the condition $\boldsymbol{k_3} = \boldsymbol{k_1} - \boldsymbol{k_2}$ is satisfied (see **Figure 6**).

All the phase-matching methods discussed in the previous section in connection with SH generation can also be applied to the sum/difference frequency generation. By using the same approach outlined for SH generation, it is easy to derive nonlinear propagation equations for the sum/difference fields having a structure similar to that of eqns [3]. In the DFG process, one photon at frequency ω_1 is annihilated and two photons at frequencies, respectively ω_3 and ω_2, are created, so that the process not only produces a new coherent beam at ω_3, but also amplifies the beam at ω_2. Such an amplification process is called parametric amplification. Whereas standard optical amplifiers work only at frequencies corresponding to transitions among energy levels, parametric amplifiers use transparent crystals and can work, in principle, over the whole transparency window of the crystal, provided that phase matching can be achieved (see **Figure 7**).

Recalling that an optical amplifier can become a laser oscillator once inserted into an optical cavity, a similar approach can transform the parametric amplifier into a parametric oscillator. Whereas in the laser case, the trigger to cavity oscillation comes from spontaneous emission from the upper level of the amplifying medium, the initial seed for parametric oscillation comes from parametric fluorescence, the process by which the "pump" photon at frequency ω_1 spontaneously breaks down into two photons at frequencies ω_2 (signal) and ω_3 (idler). Potentially, parametric fluorescence can yield any couple of frequencies satisfying energy conservation, the choice among all possible couples being determined by the phase-matching condition. The tuning of the parametric oscillator is generally achieved by rotating the crystal or by changing the operating temperature.

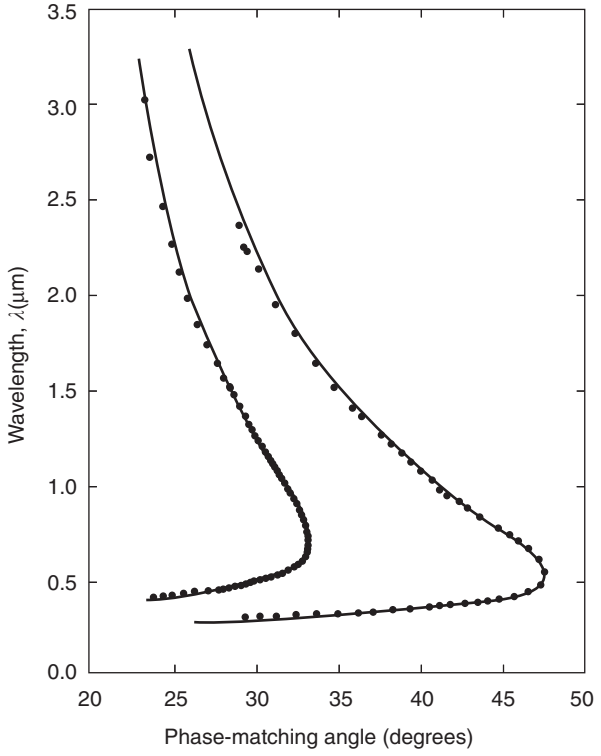


Figure 6 Tuning curves for an optical parametric oscillator pumped by the third- and fourth-harmonic of a neodymium laser. (Reproduced with permission from Cheng *et al.* (1988) *Applied Physics Letters* 53(3): 175–177.)

Nonlinear Optical Materials

The second-order processes described in the previous sections require the availability of noncentrosymmetric media. Such media are intrinsically anisotropic, so that the electric susceptibilities become tensorial quantities. In particular, $\chi^{(2)}$ is a third-order tensor, possessing 27 components χ_{ijk}. This fact makes the description of beam propagation inside the crystal and also the search for phase matching conditions rather complicated. Such descriptions are beyond the scope of this article; however, the following example is taken up: consider an input field polarized along the x-axis (this implies $j = k = 1$);

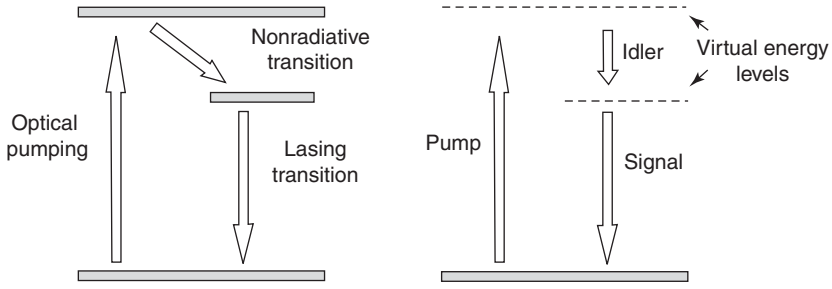


Figure 7 Comparison between the three-level laser, using real energy levels, and the optical parametric oscillator, using virtual levels.

birefringence phase matching requires that the generated nonlinear polarization lies along the z-axis (this implies $i = 3$), so that the useful component of the $\chi^{(2)}$ tensor is, in this case, χ_{311}. As a general consideration, the most powerful phase matching method, that is, birefringence phase matching, exploits nondiagonal components of $\chi^{(2)}$, whereas several interesting crystals present very large diagonal components of $\chi^{(2)}$. Such components can be exploited only if the phase-matching conditions are achieved by means of the periodical poling technique.

The commonly used second-order materials are inorganic crystals with oxygen polyedra, which can be viewed as ionic crystals where one of the ions is replaced by an oxygen polyedron in an ionized state, stabilized in a noncentrosymmetric configuration by the surrounding ions (usually of metal elements). Examples of such crystals are the ABO_3 type of ferroelectrics, such as $LiNbO_3$ and $KNbO_3$, borates, such as BaB_2O_4 and LiB_3O_5 (usually indicated, respectively, as BBO and LBO), and the isomorphic family with the generic formula $MTiOXO_4$, such as potassium titanyl phosphate $KTiOPO_4$ (usually indicated as KTP).

It is important to note that the requirements for nonlinear materials are not only suitable transparency and high nonlinearity, but also sufficient birefringence for phase-matching, high-optical-damage threshold, good thermal, mechanical, and chemical properties, and possibility of growing large crystals of good optical quality.

Families of crystals possessing a second-order nonlinearity larger than that of inorganic crystals are organic molecular crystals and heteropolar semiconductors. However, at present, they are less interesting for applications because the former has insufficiently good mechanical and chemical properties, whereas the latter is not birefringent. In addition, both families have a limited transparency range.

A possibility that is actively explored is that of starting from an amorphous material (such as a polymer or a silicate glass) and introducing an orientation of noncentrosymmetric molecular groups by heating at a temperature above the glass transition and applying an external field, usually an electric field (electrical poling).

Third-Order Nonlinear Effects

Optical Kerr Effect

If a third-order material illuminated by a monochromatic plane wave at frequency ω propagating along the z-axis is considered, it results from eqn [2] that such a wave gives rise to a nonlinear polarization oscillating at $3\omega_1$. In turn, the nonlinear polarization, once inserted into eqn [1], yields a third-harmonic (TH) optical wave. The entire treatment developed for SH generation, including the phase-matching considerations, can be easily translated to the TH case, and is not repeated here. Instead, the attention is on the polarization term having the expression

$$P(z,\omega) = \varepsilon_0[\chi^{(1)}E(z,\omega) + \chi^{(3)}\alpha|E(z,\omega)|^2E(z,\omega)] \quad [5]$$

where α is a numerical factor that can assume values ranging from 1/2 to 3/2, considering different third-order phenomena. From eqn [5], it is possible to define an intensity-dependent susceptibility: $\chi^{(1)}_{\text{eff}} = \chi^{(1)} + \chi^{(3)}\alpha|E(z,\omega)|^2$. By recalling the relation between index of refraction and electric susceptibility and assuming that the nonlinear contribution is small in comparison with the linear one, one obtains

$$n_{\text{eff}} = n + n_2|E|^2 \quad [6]$$

where $n_2 = \alpha\chi^{(3)}/(2\varepsilon_0 c\chi^{(1)})$. A more careful analysis shows that the overall effect consists in an intensity-dependent deformation of the index ellipsoid of the medium. In particular, an initially isotropic material becomes uniaxial in the presence of a strong linearly polarized optical beam.

Equation [6] expresses the optical Kerr effect. Some phenomena due to such an effect are briefly discussed below.

Optical switch The phase and the polarization of an optical signal going through a Kerr medium can be changed by illuminating the medium with a strong auxiliary (pump) beam. This opens the way to the realization of all-optical modulators and switches. If the Kerr effect is fast, as it happens when it is due to the displacement of electron clouds, the response time of the device can be as short as the duration of pump pulses (less than 1 ps).

Self-focusing Considering a pump with a Gaussian transverse intensity profile (this is the output beam from a single-mode laser), the optical Kerr effect produces a transverse index-of-refraction profile that follows the shape of the beam. If n_2 is positive, a greater index of refraction is induced on-axis than in the wings of the beam, creating in this way, a positive lens that tends to focus the beam. This effect is called self-focusing. If the input power is larger than a threshold value, self-focusing can exactly balance diffraction, so that the beam propagates without changing its profile. This phenomenon is called self-trapping. If the input power is further increased, the beam will catastrophically focus, causing in many cases optical damage inside the medium. To conclude this short discussion about self-focusing, it should be mentioned that the effect is very usefully exploited in the so-called Kerr-lens passive mode-locking method for the generation of ultrashort laser pulses.

Self-phase modulation Considering a beam with a time-dependent intensity profile (optical pulse), the optical Kerr effect gives rise to a self-induced nonlinear phase shift that is also time dependent. Generally speaking, this has the consequence of broadening the frequency spectrum of the optical pulse. It should be recalled that an optical pulse propagating in a linear dispersive medium has a time duration increasing progressively with the propagation distance. Among the various phenomena occurring during the nonlinear propagation of short pulses, it is worth mentioning the existence of optical solitons, that is, pulses of appropriate shape, intensity, and time duration that propagate without changing shape and duration in a medium presenting anomalous group-velocity dispersion. Such a phenomenon is made possible by a compensation between the effects of linear dispersion and self-phase modulation. Particularly important for applications to optical communications is the study of nonlinear propagation in single-mode optical fibers: although the nonlinearity of glass is very small, nonlinear effects can be significant because the intensity can be high and the propagation distance is very large.

Four-Wave-Mixing and Optical Conjugation

The term four-wave-mixing (FWM) refers to the interaction of four waves in a nonlinear medium via the third-order polarization. In the presence of three monochromatic beams at frequencies ω_1, ω_2, ω_3, a nonlinear polarization term oscillating at the frequency $\omega_4 = \omega_1 + \omega_2 - \omega_3$, with wave vector $\boldsymbol{k_4} = \boldsymbol{k_1} + \boldsymbol{k_2} - \boldsymbol{k_3}$, is generated:

$$P = \varepsilon_0 \chi^{(3)} A_1 A_2 A_3^* \ \exp[\mathrm{i}(\boldsymbol{k_1} + \boldsymbol{k_2} - \boldsymbol{k_3}) \cdot \boldsymbol{r}] \quad [7]$$

Among the various phenomena originated by the nonlinear polarization of eqn [7], two particularly interesting processes are selected (see **Figure 8**).

Frequency-shifting scheme In practical situations, a high-intensity beam (pump beam) at frequency $\omega_p = \omega_1$ is sent collinearly with the weak beam (signal beam) 2. The new beam 3 is also collinear, and has a frequency $\omega_3 = 2\omega_1 - \omega_2$. Note that $\omega_3 - \omega_1 = \omega_1 - \omega_2$, that is, ω_3 is the mirror image of ω_2 with respect to the pump frequency. If, for instance, the pump beam is continuous and the signal is a sequence of pulses, the frequency-shifted beam 3 is a temporal replica of the signal beam. It should be added that, in the absence of materials having a large $\chi^{(3)}$, it is usually more convenient to perform the same operation by using a cascade of two second-order processes in a $\chi^{(2)}$-material. Frequency-shifting devices can be usefully applied in optical communications systems.

Phase-conjugate mirror If a degenerate case in which all the frequencies coincide is considered, $\omega_1 = \omega_2 = \omega_3 = \omega$, the new field has frequency $\omega_4 = \omega$ and wave vector $\boldsymbol{k_4} = -\boldsymbol{k_3}$, that is, it counterpropagates with respect to beam 3. The nonlinear

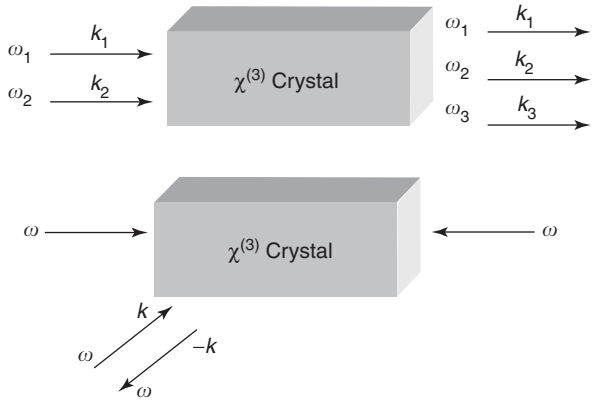


Figure 8 Schematic wave-mixing processes: frequency shifter and conjugated mirror.

medium illuminated by beams 1 and 2 acts as a mirror for beam 3: this is what is called a phase-conjugate mirror. Such a mirror has some unique properties compared to an ordinary mirror. First of all, it functions as a retroreflector for any incident angle θ, within the acceptance angle of the experiment. Any phase change $\phi(\boldsymbol{r})$ impressed on beam 3 by propagating, for example, through an inhomogeneous medium, will result in exactly the opposite phase $-\phi(\boldsymbol{r})$ impressed on the reflected wave, in this way correcting the phase distortion during the back-propagation. In terms of energy balance, the FWM process subtracts one photon from beam 1 and one photon from beam 2, and adds one photon to beam 3 and one photon to beam 4. As a consequence, the reflection coefficient of the conjugated mirror is proportional to the product of the intensities of beams 1 and 2 and can even become larger than 1.

This section concludes with a brief discussion on third-order nonlinear materials. It is useful to distinguish between slow and fast nonlinearities. Slow nonlinearities involve the rotation and alignment of molecular groups under the action of an optical field, such as it may happen in liquid crystals. The effects are large, but the response time can be of the order of milliseconds or even longer. Fast nonlinearities involve the displacement of electron clouds, response times can be in the range of picoseconds but the effects are small. In order to enhance electronic nonlinearities one can operate near a resonance, but the drawback is that response times are now controlled by the material decay time. The most interesting materials are semiconductors or conjugated polymers, but, in practice, the largest fast third-order nonlinearities are obtained by using a cascade of second-order processes in a noncentrosymmetric material.

Stimulated Raman and Brillouin Scattering

Raman scattering is an inelastic light scattering process in which a quantum of excitation is exchanged between the optical field and the scattering medium. A typical case is that of a molecule possessing a vibrational mode with angular frequency ω_v, illuminated by a light beam at frequency ω. The spectrum of scattered light, observed at any scattering angle, will contain, besides the unshifted Rayleigh line at ω, also an additional line, called Stokes line, at frequency $\omega_s = \omega - \omega_v$. The energy difference between the incident and scattered photon is taken to excite the molecular vibration. If the illuminated medium includes molecules that are already in an excited vibrational state, the spectrum of scattered light will contain a third line, called the anti-Stokes line, at frequency $\omega_{as} = \omega + \omega_v$, describing a scattering event in which the molecule has given its vibrational energy to the scattered photon. Raman scattering may exist also in association with rotational energy levels of molecules. Besides localized molecular excitations, there are also collective excitations producing Raman scattering, such as optical phonons in crystals or plasma waves in an electron gas. It should be noted that not all elementary excitations are Raman active; there are selection rules, usually different from those controlling radiative transitions.

Stimulated Raman scattering (SRS) is observed when two light beams, at frequencies ω and ω_s, are sent into the medium. The Stokes photons generated by SRS are emitted with the same wave vector as that of the incident Stokes field and added coherently to that field. The result is that the field at ω_s is amplified, and, at the same time, more molecules are excited to the vibrational level. SRS can be used not only to amplify a signal at ω_s, but also to generate a coherent wave at ω_s. In fact, if a powerful beam at ω is sent into the medium, it will create Stokes photons by spontaneous Raman scattering, and the created Stokes photons can constitute the seed for an amplification process yielding an intense coherent Stokes field. Note that the spontaneous Stokes photons are emitted isotropically in space, whereas the stimulated Stokes beam will appear in a specific direction satisfying the criterion of maximum overlap between the pump and the Stokes beam. In practice, SRS is observed only in the forward and backward directions.

The growth of the Stokes field along the direction z is described by the equation

$$P_S(z) = P_S(0)\exp\left[\int_0^z g_R P_p(z)\mathrm{d}z\right] \qquad [8]$$

where g_R (in units of $\mathrm{m\,W^{-1}}$) is called the Raman gain and $P_p(z)$ is the pump power. SRS is a third-order process; in fact, it can be shown that g_R is proportional to the imaginary part of $\chi^{(3)}$.

If the scattering medium is placed inside an optical cavity, one can build a Raman laser. The Raman laser is different from a normal laser because the pump frequency ω_p can be arbitrary, whereas in normal lasers ω_p must correspond to a transition between the ground level and an excited level of the active medium. As shown in **Figure 9**, the energy transfer between pump and Stokes can be so efficient that a sequence of Stokes lines of decreasing frequency can be coherently generated. Raman lasers utilizing

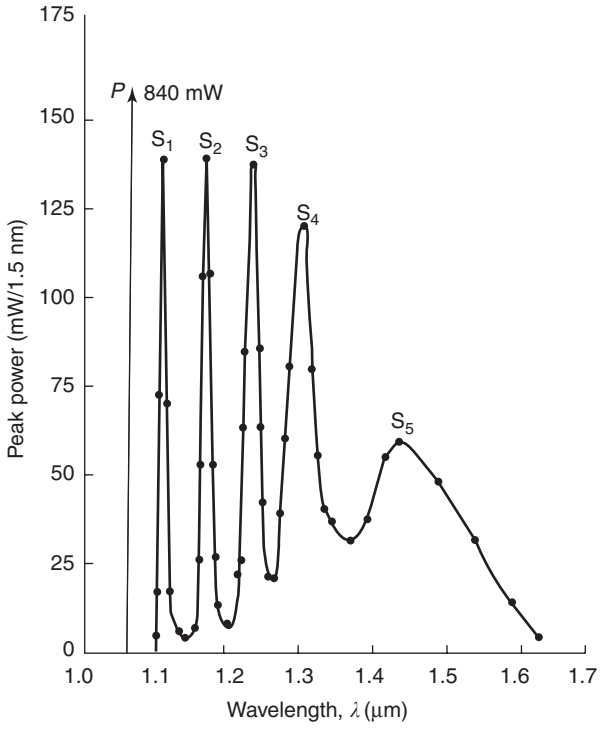


Figure 9 Generation of a sequence of Stokes lines in an optical fiber by stimulated Raman scattering. (Reproduced with permission from Cohen *et al.* (1978) *IEEE Journal of Quantum Electronics* QE-14(11): 855–859.)

optical fibers as active medium represent very efficient and powerful sources of coherent radiation: as an example, fiber Raman lasers pumped by a neodimium-laser ($\lambda = 1064\,\text{nm}$) and working on the fifth Stokes line can emit at a wavelength $\sim 1500\,\text{nm}$ with an overall efficiency of 30%.

Any medium at a nonzero temperature presents local fluctuations in the thermodynamic variables that give origin to light scattering. It is assumed that the incident light is characterized by a wave vector $\boldsymbol{k}$ and an angular frequency ω. In the case of a solid or liquid, the frequency spectrum of scattered light contains an unshifted line (due to nonpropagating density fluctuations) and two lines, called Brillouin lines, centered at frequencies $\omega \pm \omega_B$ that are due to density fluctuations propagating as spontaneous acoustic waves. If the scattering direction is fixed, hence the wave vector of the scattered field, $\boldsymbol{k}'$, only those spontaneous waves possessing a wave vector $\boldsymbol{K} = \pm(\boldsymbol{k}' - \boldsymbol{k})$ contribute to light scattering in that direction. The frequency shift ω_B of the Brillouin lines corresponds to the frequency of the acoustic phonon which is exchanged in the scattering process. Recalling that the velocity of acoustic waves is given by $u_a = \omega_B/K$, the Brillouin frequency shift can also be seen as a Doppler shift due to diffraction from a traveling phase grating. Although Raman shifts are typically in the range of several terahertz, Brillouin shifts are of the order of some gigahertz.

Stimulated Brillouin scattering (SBS) can occur in a way fully similar to that of SRS. Only the interaction $\boldsymbol{K} = -(\boldsymbol{k}' - \boldsymbol{k})$, corresponding to the situation in which the incident wave reinforces the spontaneous acoustic wave, can give rise to a stimulated effect, generating a coherent wave at the frequency $\omega - \omega_B$. In practice, SBS is observed only in the backward scattering direction. It is interesting to mention that the backward Brillouin wave behaves as the phase conjugate of the pump wave. Therefore SBS can create a phase-conjugate mirror, similar to FWM, but in a simpler way, because only one beam is required.

See also: Metals and Metallic Alloys, Optical Properties of; Optical Fibers; Optical Properties of Materials; Semiconductor Optics; Sum Rules and Kramers–Kronig Relations in Nonlinear Optics.

PACS: 42.65.An; 42.65.Dr; 42.65.Es; 42.65.Hw; 42.65.Ky; 42.65.Yj

Further Reading

Agrawal GP (1995) *Nonlinear Fiber Optics*, 2nd edn. New York: Academic Press.

Boyd RW (1992) *Nonlinear Optics*. New York: Academic Press.

Degiorgio V and Flytzanis C (eds.) (1995) *Nonlinear Optical Materials: Principles and Applications*. Amsterdam: IOS Press.

Sutherland RL (2003) *Handbook of Nonlinear Optics*. New York: Dekker.

Nomenclature

$A(z)$	electric field slowly varying envelope
c	$(3 \times 10^8\,\text{m s}^{-1})$ velocity of light in vacuum
d_{eff}	effective nonlinear coefficient
$\boldsymbol{E}(\boldsymbol{r}, t)$	electric field
g_R	Raman gain
I	intensity
k	electric field wave vector
l_c	second-harmonic generation coherence length
L_{SH}	characteristic second-harmonic generation length
n	index of refraction
n_e	extraordinary index of refraction
n_o	ordinary index of refraction
n_2	nonlinear refractive index
P_p	pump power
$\boldsymbol{P}(\boldsymbol{r}, t)$	macroscopic polarization
P_s	Stokes power
u_a	velocity of the acoustic wave
ε_0	$(8.85 \times 10^{-12}\,\text{F m}^{-1})$ permittivity vacuum

λ	wavelength	ω	angular frequency
Λ	spatial periodicity	ω_{as}	anti-Stokes frequency
μ_o	$(4\pi\times10^{-7}\,\mathrm{N\,A^{-2}})$ permeability vacuum	ω_{pm}	phase-matching frequency
$\chi^{(1)}$	linear susceptibility	ω_s	Stokes frequency
$\chi^{(2)}$	second-order susceptibility	ω_v	vibrational mode with angular frequency
$\chi^{(3)}$	third-order susceptibility		

Nuclear Fission and Fusion

D M Brink, University of Oxford, Oxford, UK

Introduction

Nuclear fission is a process in which a heavy atomic nucleus divides into two massive fragments. The phenomenon was discovered by Hahn and Strassmann in 1938 in experiments with uranium (atomic number $Z = 92$). They noticed that radioactive products formed from uranium under neutron bombardment included isotopes of barium with atomic number $Z = 56$. Hahn and Strassmann sent their results to their collaborator Lise Meitner in December 1938 and within a few weeks Meitner and Frisch (1939) arrived at an interpretation. They wrote: "the particles in a heavy nucleus would be expected to move in a collective way which has some resemblance to the movement of a liquid drop. If the movement is made sufficiently violent by adding energy, such a drop may divide itself into two smaller drops." After division the drops would repel each other and gain an energy of $\sim$200 MeV.

Fission is an important source of energy and this article begins with a very short introduction to the physics of nuclear reactors. The main focus is on some advances in the understanding of nuclear fission which have been achieved over the past 65 years. Transuranic elements, for example, uranium, are produced in nuclear reactors. They are also created using fusion reactions with beams from heavy ion accelerators. A part of this article is devoted to the study of these nuclei and the quest for superheavy nuclei. The last part of the article is about fusion reactions with light nuclei.

Nuclear Reactors

The fissioning nucleus studied by Hahn and Strassmann was the isotope ^{235}U of uranium. As argued by Meitner and Frisch, the nucleus ^{236}U produced by absorbing a neutron is unstable and splits into two unequal fragments, for example

$$^{235}\mathrm{U} + \mathrm{n} \rightarrow {}^{236}\mathrm{U}^{*} \rightarrow {}^{93}\mathrm{Rb} + {}^{141}\mathrm{Cs} + 2\mathrm{n} \qquad [1]$$

Other similar decays occur but always with fragment mass numbers $A_1 \sim 95$ and $A_2 \sim 140$. The energy released can be calculated from the measured masses of the isotopes and is 196 MeV in the above example. There is a large energy release because the fragments are more strongly bound than the initial nucleus.

Nuclear power can be produced in fission reactors with the isotope ^{235}U as fuel. The process involves a chain reaction in which neutrons produced in the fission of one uranium nucleus can cause fission in other uranium nuclei. The fission cross section for neutrons on ^{235}U is a maximum for neutrons with energies of less than 1 eV, and in a thermal reactor the neutrons are slowed down by collisions with a moderator, such as carbon or hydrogen.

Natural uranium contains only 0.7% of the isotope ^{235}U. The major component is ^{238}U. This isotope absorbs neutrons but does not fission and instead produces the plutonium isotope ^{239}Pu. The isotope ^{239}U produced by neutron capture decays by beta decay, first to ^{239}Np and then to ^{239}Pu. Several transuranic elements are produced in reactors by further capture of neutrons followed by beta decay. These transuranic elements as well as the fission products are all radioactive and constitute an important component of "nuclear waste."

The plutonium isotope ^{239}Pu can also undergo fission when it absorbs neutrons and is a potential fuel for nuclear reactors. The fission cross section for plutonium is small for slow neutrons but becomes large for neutrons with energies greater than 1 keV. Plutonium burns best in a "fast reactor" where the fission neutrons are slowed down, but not too much.

Stability against Fission

Very soon after the discovery of the fission process, Bohr and Wheeler proposed a detailed theoretical treatment based on the comparison between the nucleus and a liquid drop. All subsequent developments preserve aspects of their model.

Bohr and Wheeler assumed that the liquid drop representing the nucleus passes through a sequence of shapes on the pathway to fission. Starting from the initial metastable equilibrium shape the nucleus becomes very deformed, then it develops into a dumbbell shape with two fragments joined by a neck. Finally the neck breaks, the fragments separate and are accelerated away from each other by the strong Coulomb repulsion acting between them.

In a simple picture one can introduce a parameter q associated with each of the shapes on the pathway to fission and function $V(q)$ representing the energy of the nucleus for each value of q. It is convenient to choose $q = 0$ for the initial shape and $q = R$, where R is the distance between the centers of mass of the fragments, when the two fragments are separated. For nuclei which are metastable against fission, $V(q)$ has a local minimum at $q = 0$, increases to a maximum $V(q_B)$ at the fission barrier $q = q_B$, and then decreases monotonically for $q > q_B$. When the fragments are separated, they are accelerated away from each other by the electrostatic repulsion acting between them.

The liquid drop model predicts that the equilibrium shape of a nucleus is spherical with $q = 0$. For a small deviation from the spherical shape along the fission path, the coordinate q represents a quadrupole deformation and $V(q) \approx V(0) + (1/2)Cq^2$, where

$$C \propto \tfrac{2}{5}\, a_s A^{2/3} - \tfrac{1}{5}\, a_c Z^2 A^{-1/3} \qquad [2]$$

where a_s and a_c are the surface energy and Coulomb energy parameters in the semi-empirical mass formula. The surface energy increases and the Coulomb energy decreases as the nucleus deforms. The stability of the spherical shape is determined by the balance between the surface energy and the Coulomb energy. The spherical shape is stable if the restoring force parameter $C > 0$ or

$$x = \frac{a_c}{2a_s}\,\frac{Z^2}{A} = \frac{2E_s(0)}{E_c(0)} < 1 \qquad [3]$$

The fissility parameter x was introduced by Bohr and Wheeler as a measure of the stability of a heavy nucleus. Using standard values for the parameters $a_s \approx 16.8$ MeV and $a_c \approx 0.72$ MeV, the liquid drop model predicts that a nucleus is unstable against fission when $Z^2/A > (Z^2/A)_{\text{limiting}} \approx 47$. There are important shell corrections to the liquid drop model but these simple arguments show there is a limit to the possibility of creating very heavy elements. The nucleus ^{235}U has $Z^2/A = 36$ while one of the heaviest known transuranic elements with $Z = 114$ and $A = 298$ has $Z^2/A = 44$.

Fission Barriers

When a nucleus is metastable against fission, the difference $V_B = V(q_B) - V(0)$ is called the fission barrier height or the fission activation energy. When the excitation energy E^* is less than V_B, then the nucleus can fission only by quantum tunneling through the barrier and the fission probability is very small. When $E^* > V_B$, then fission decay rate is large. This can be illustrated by the example of ^{236}U. Alpha decay is the dominant decay mode of the ground state of ^{236}U with a fission probability of only 10^{-7}%. When ^{236}U is formed by thermal neutron capture on ^{235}U, the compound nucleus, has an excitation energy of 6.5 MeV, which is greater than, the activation energy $V_B = 6.2$ MeV and fission is the dominant decay mode. Another example is ^{239}U, which can be formed by neutron capture on ^{238}U. When ^{238}U captures slow neutrons, then the excitation energy $E^* =$ 4.8 MeV is less than the activation energy 6.6 MeV and the fission probability is very small. On the other hand, when ^{238}U is bombarded with neutrons with an energy greater than 2 MeV, then the excitation energy of the ^{239}U is larger than V_B and the fission cross section is comparable with that for neutrons with the same energy on ^{235}U.

There is a general trend for fission barrier heights to decrease with Z for transuranic elements, but there are fluctuations. Nuclei with atomic numbers less than $Z = 92$ can undergo fission, provided the excitation energy is large enough. The fission barrier for ^{208}Pb is $\sim$28 MeV and for ^{216}Rn it is $\sim$13 MeV. Their fission probabilities are extremely small.

The Dynamics of Fission

In the previous sections, it was assumed that there was a unique path for the fission of a nucleus. The situation is more complicated than this for the liquid drop model and also for real nuclei. The fission of a nucleus such as ^{236}U does not lead to a unique pair of fission fragments. At the stage where the nucleus becomes very deformed and develops into a dumbbell shape, the number of nucleons on each side of the neck is not unique. When the neck breaks, there is a distribution in the partition of mass and charge

between the fragments and there is a high probability that the fragments will be left in excited states.

The distribution of mass and charge between the fragments can be measured by identifying the fission fragments. The liquid drop model predicts that, on the average, the mass is equally distributed between the two fragments, but in many cases the mass distribution is asymmetric. For example in the fission of ^{235}U by thermal neutron capture, the mass distribution has two peaks, one centered at $A \approx 95$ and the other at $A \approx 140$. This asymmetric fission is a shell effect and is discussed in the next section. Typically the spread of each peak in the mass distribution is in the range $\delta A \approx 10$–15 mass units.

The distribution in excitation energy of the fission fragments can be investigated by measuring the kinetic energy TKE of the fragments. The total energy released is equal to the Q-value, $Q = \delta M c^2$, where δM is the difference between the mass of the initial nucleus and the sum of the masses of the fragments. The total excitation energy of the fragments is given by $\delta E_{\text{exc}} = Q - \text{TKE}$. Some nuclei may have more than one fission barrier, each with its own characteristic mass distribution and TKE distribution. The existence of more than one barrier can often be detected by making a careful analysis of the mass and TKE distributions of the fragments.

Decay Rates

Any quasistationary state of a nucleus which can decay by fission may also have a branch for decay by α-emission, γ-emission, neutron emission, or various other processes. The decaying state has a total width Γ related to the mean life τ by $\Gamma = \hbar/\tau$. The total width is a sum of partial widths $\Gamma = \Gamma_{\text{f}} + \Gamma_{\alpha} + \Gamma_{n} + \cdots$ for each decay process and the fission probability or branching ratio is Γ_{f}/Γ. The mean life τ and fission branching ratio of the ground state of a nucleus can be measured directly. At higher excitation, energy states may appear as resonances in nuclear reactions, for example, low-energy neutron scattering reactions. Then Γ is the width of a resonance. At higher excitations resonances overlap and the width of a state is not a measurable quantity, but decay rates and branching ratios still have a meaning.

The fission width of a state can be estimated in various ways depending on the excitation energy of the nucleus. A nucleus in the ground state can decay by quantum tunneling through the fission barrier. The fission width is approximately

$$\Gamma_{\text{f}} \approx \frac{\hbar\omega_{\text{vib}}}{2\pi} P \qquad [4]$$

where ω_{vib} is a characteristic frequency for the oscillation of the fission coordinate q in the ground state and P is the transmission factor for penetrating the fission barrier along the tunneling path. A simple estimate for P is

$$P = \exp\left(-\frac{2\pi\Delta E}{\hbar\omega_{\text{B}}}\right) \qquad [5]$$

where ΔE is the energy of the decaying state below the top of the barrier and ω_{B} is related to the curvature of the top of the barrier. Typical values of ω_{B} lie in the range 0.5–1 MeV.

The states in nucleus with an excitation energy of several MeV have a complicated structure and statistical arguments can be used to estimate fission widths in terms of the average spacing $D(E^*)$ of levels of the excitation energy E^*. The quantity $\hbar/D$ represents the probability per unit time for concentrating the excitation energy into the motion of the fission coordinate. For excitation energies below the fission barrier, the width can be estimated from

$$\Gamma_{\text{f}} \approx \frac{D(E^*)}{2\pi} P \qquad [6]$$

The transition state theory of Bohr and Wheeler can be used to estimate fission decay rates of states well above the fission barrier. They assumed that internal degrees of freedom could be excited at the barrier and a formula can be obtained for the fission rate

$$\frac{\mathrm{d}N_{\text{f}}}{\mathrm{d}t} \approx \frac{D(E^*)}{2\pi\hbar} N^* \qquad [7]$$

Here $N^* \geqslant 1$ is the effective number of internal states available at the barrier.

Asymmetric Fission and Shell Effects

The shell model or independent-particle model of a nuclear structure assumes that a nucleon in the nucleus moves in a single-particle potential which represents its average interaction with other nucleons. Nucleons occupy single-particle levels in the shell model potential and fill up the lowest levels allowed by the Pauli exclusion principle. Closed shells occur when there is a rather large energy gap between the last filled level and the lowest available empty level. The ordering of single-particle levels is such that closed shells occur for neutron or proton numbers equal to 2, 8, 20, 28, 50, 82, … . Heavier nuclei with closed shells of neutrons and protons are ^{40}Ca, ^{48}Ca, ^{90}Zr, ^{132}Sn, and ^{208}Pb. Closed-shell nuclei and

especially double closed-shell nuclei are particularly stable and have binding energies larger than the predictions of the liquid drop model.

Taking fluctuations into account, the liquid drop model predicts that the mass distribution of fission fragments should be symmetric about the average value $A/2$, but in most cases the measured distribution is asymmetric. For example, the mass distribution ^{235}U has one peak centered at $A_1 \approx 95$ and another at $A_2 \approx 140$. The asymmetric distribution of fragments is very common when the excitation of the fissioning nucleus is less than ~ 6 MeV. There is remarkable systematics in the fragment mass distribution for the fission of nuclei with $256 > 228$. In all these cases, the average mass of the heavy fragment lies in the range 137–142. This is not so far from the mass of the double closed-shell nucleus ^{132}Sn. This systematics extends to the superheavy nuclei ^{286}112 and ^{292}114. In both these cases, the average mass of the light fragment is near to the mass of ^{132}Sn.

In 1966, Strutinsky invented a way of adding shell corrections to the liquid drop model. The total nuclear binding energy in the shell-correction approach is written as

$$E_{\text{tot}}(N, Z, \beta) = E_{\text{LDM}}(N, Z, \beta) + \delta E_{\text{n}}(N, \beta) + \delta E_{\text{p}}(Z, \beta) \qquad [8]$$

where β stands for a set of deformation parameters, E_{LDM} is the liquid drop energy as a function of deformation while δE_{n} and δE_{p} are shell corrections calculated for neutrons and protons moving in appropriate deformed single-particle potentials. This simple model was very helpful for understanding how shell effects influenced asymmetric fission and generalizes the simple argument of the dominance of the double closed-shell nucleus ^{132}Sn. Shell effects disappear in a nucleus with a large excitation energy and the asymmetry in the fission fragment distribution disappears.

Fission Isomers

Many different processes associated with fission have shown that the fission barrier does not always have a single maximum as predicted by the liquid drop model. According to present-day understanding, many nuclei have a double-humped fission barrier. The first minimum (I) in the double-humped potential $V(q)$ contains the ground state. Then there is a barrier followed by a second minimum (II) and finally another barrier which leads to fission. Quasi-stationary states in the second minimum are called fission isomers.

The ground state in the first minimum (I) is normally not spherical, but has a prolate equilibrium shape. The first minimum also contains type (I) excited states which have deformations similar to that of the ground state. Type (II) levels are located around the second minimum and are separated from the type (I) states by the first hump of the potential barrier. The type (I) states have deformations $\delta \sim 0.35$ while the deformations of the type (II) states are more than twice as large. Fission from a class (I) state occurs when the fission coordinate tunnels through the first barrier into the second minimum (II) and finally through the outer barrier.

The cross section for scattering of low-energy neutrons by ^{240}Pu has sharp resonances with a spacing ~ 10 eV. Each resonance corresponds to a state of the compound nucleus ^{241}Pu. The wave functions of the compound states are located around the first minimum in the double-humped potential. Most of the resonance states have a relatively small fission width, but there is a group of about seven resonances corresponding to neutron energies lying between 750 and 820 eV which have much larger fission widths. There is another group of resonances with large fission widths clustered around 1400 eV and a third at 1950 eV. This behavior can be explained by the double-humped potential. If the energy of a compound state in the first minimum is very close to the energy of a type (II) state, then there is a resonance effect which enhances the penetration of the double barrier. The penetration probability of the double barrier can be much larger than that for one of the individual barriers.

Fission isomers can be studied in other ways. For example, states in the second minimum of ^{239}U can be excited in the reaction ^{238}U(d, p) and γ-rays from the isomeric states observed. Superdeformed states have been observed in nuclei with mass numbers $A \sim 130$, 150, and 190. These have very similar properties to fission isomers.

Nuclear Fusion and Transuranic Elements

Bohr and Wheeler showed that elements with large enough atomic number must be unstable against fission because of the electrostatic repulsion between protons in the nucleus. For the past half century, there has been an intensive effort to create nuclei with larger and larger atomic number to try to reach the limit of stability. There have been suggestions that shell effects might give enough extra binding to elements with $Z \approx 114$ to make them more stable. There is strong evidence that the element with $Z = 114$ has already been produced in a collaboration between scientists from The Lawrence Livermore

Laboratory in California and the Dubna Laboratory in Russia. They bombarded a plutonium ^{244}Pu target with ^{48}Ca nuclei to create the new element.

To date, more than 20 transuranic ($Z > 92$) elements and 2000 isotopes have been discovered and various properties, including masses and lifetimes, have been studied. The names with atomic numbers and symbols of the first few new elements are neptunium (93Np), plutonium (94Pu), americium (95Am), curium (96Cm), berekelium (97Bk), californium (98Cf), einsteinium (99Es), and fermium (100Fm). Isotopes of these elements can be produced by a variety of reactions. For example, curium was first made in 1944 by the reaction

$$^{4}\mathrm{He} + ^{239}\mathrm{Pu} \rightarrow ^{265}\mathrm{Cm} + \mathrm{n} \qquad [9]$$

Isotopes up to ^{254}Fm and beyond can be produced by intense neutron irradiation of lower elements, such as plutonium, using a process of successive neutron capture interspersed with β-decays until these mass numbers and atomic numbers are reached.

Heavier elements are produced by fusion reactions. For example, the element hassium with $Z = 108$ is produced in the reaction

$$^{58}\mathrm{Fe} + ^{208}\mathrm{Pb} \rightarrow ^{265}\mathrm{Hs} + \mathrm{n} \qquad [10]$$

These reactions are very delicate because if the intermediate compound nucleus has too large an excitation energy, it will not survive. For the reaction to succeed, the incident energy of the projectile must be just large enough to cross the Coulomb potential barrier.

Fusion Reactions with Light Nuclei

Fusion reactions are the source of energy production in stars and are responsible for the conversion of protons into heavier elements. They are also a potential energy source for modern industrial society. Energy production in stars comes mainly from the conversion of hydrogen into helium. The overall result of a typical sequence of reactions is

$$4\,^{1}\mathrm{H} \rightarrow ^{4}\mathrm{He} + 2\mathrm{e}^{+} + 2\nu \qquad [11]$$

with the release of 27.7 MeV for each ^{4}He atom produced. Fission reactions in nuclear reactors are initiated by neutron absorption. The cross sections are large because the neutrons have no charge. Cross sections for reactions between charged nuclei are very small at low energies because of the electrostatic repulsive forces acting between them.

The first step in the production of helium from hydrogen in stars is the conversion of protons into deuterons:

$$^{1}\mathrm{H} + ^{1}\mathrm{H} \rightarrow ^{2}\mathrm{H} + \mathrm{e}^{+} + \nu \qquad [12]$$

This is a weak interaction process and the cross section is very small, but the reaction rate is significant because there are so many protons. After deuterons have been produced, other reactions can occur, for example

$$^{2}\mathrm{H} + ^{1}\mathrm{H} \rightarrow ^{3}\mathrm{He} + \gamma \qquad [13]$$

and

$$^{3}\mathrm{He} + ^{3}\mathrm{He} \rightarrow ^{4}\mathrm{He} + 2\,^{4}\mathrm{H} \qquad [14]$$

An accurate knowledge of thermonuclear reaction rates for reactions such as these is important for understanding the generation of energy and other physical processes in stars.

All the above reactions are inhibited by the Coulomb repulsion between the nuclei. The central temperature in a star such as the sun is $\sim 1.5 \times 10^{7}$ K which is equivalent to 1.3 keV. This is much less than the height of a typical Coulomb barrier and the reaction cross sections become very small. The cross section for a reaction between two charged nuclei for incident energies below the top of the Coulomb barrier can be written as [2]

$$\sigma(E) = \frac{S(E)}{E} \exp(-2\pi\eta) \qquad [15]$$

where the Sommerfeld parameter is given by

$$2\pi\eta = 2\pi Z_1 Z_2 (e^2/\hbar v) = 31.29 Z_1 Z_2 (\mu/E)^{1/2} \qquad [16]$$

The quantities Z_1 and Z_2 are the charges of the two nuclei, μ is the reduced mass in atomic mass units, and E is the center of mass energy in keV. The astrophysical S-factor $S(E)$ depends on the reaction and is normally a rather slowly varying function of energy at sub-barrier energies except for possible resonances. On the other hand, the exponential factor containing the Sommerfeld parameter is a very rapidly varying function of the energy, and cross sections become very small for low incident energies.

The reaction ^{3}He(^{3}He, ^{4}He)2p is a useful example to illustrate these points. The energy of the Coulomb barrier is ~ 2 MeV while the S-factor $S(E)$ lies in the range 4–8 MeV b as E varies between 1 MeV and 20 keV. The cross section for this reaction is ~ 10 mb when the relative energy is $E = 500$ keV, it drops to

about μb for $E = 100$ keV and to 0.3×10^{-12} b when $E = 20$ keV. One of the challenges for nuclear physicists is to measure these very small cross sections. The cross section has been measured at 20 keV and this is about the lower limit for present-day techniques.

See also: Neutron Sources; Nuclear Models and Methods; Quantum Mechanics: Nuclei; Radioactivity.

PACS: **24.75; 28.41; 28.52**

Further Reading

Armbruster P (2000) On the production of superheavy elements. *Annual Review of Nuclear and Particle Science* 50: 441.

Bjornholm S and Lynn (1980) The double-humped fission barrier. *Review of Modern Physics* 52: 725.

Bohr N and Wheeler JA (1939) The mechanism of nuclear fission. *Physical Review* 36: 426.

Fowler WA (1984) Experimental and theoretical astrophysics: the quest for the origin of the elements. *Review of Modern Physics* 56: 149.

Oganessian YT and Lazarev YA (1985) Heavy ions and nuclear fission. In: Bromley DA (ed.) *Treatise on Heavy Ion Science*, vol. 4, p. 1. New York: Plenum.

Ragnarsson I, Nilsson SC, and Sheline (1978) Shell structure in nuclei. *Physics Reports* 45: 1.

Thirolf PG and Habs D (2002) Spectroscopy in the second and third minimum of actinide nuclei. *Progress in Nuclear and Particle Physics* 49: 325.

Vandenbosch R and Huizenga J (1973) *Nuclear Fission*. New York: Academic Press.

Wagemans C (1991) *Nuclear Fission Process*. Boca Raton, FL: CRC Press.

Nuclear Models and Methods

R A Broglia, University of Milan, Milan, Italy

Independent-Particle Model

In classical mechanics, a system of particles of mass m which interact through a two-body force (particle–particle potential) of the type shown in **Figure 1** is described by the Hamiltonian

$$H = \frac{p^2}{2m} + v(r_{12}) \qquad [1]$$

The lowest-energy state occurs at zero temperature. It is obtained by minimizing, for each particle, its potential energy with respect to all the neighbors. Such static, localized particles provide the basic design for the ground state of molecules and crystals (**Figure 2**). While the laws governing the motion of these particles are invariant with respect to translation, in particular $v(r_{12})$, the solution of such equations violates translational and/or rotational invariance. In other words, the solution of the classical equations may display a lower symmetry than the original Hamiltonian. This phenomenon, known as "spontaneous symmetry breaking," is the basis of the "emergent properties" of the system of particles described by the Hamiltonian given in eqn [1]. In fact, the systems in **Figure 2a** display localization and rigidity, properties neither contained in each individual particle nor in H.

In quantum mechanics, momentum and coordinate are conjugate variables fulfilling Heisenberg's relation

$$\Delta x \Delta p_x \geqslant \hbar \qquad [2]$$

From this relation, one can calculate the (zero-point) energy associated with the localization

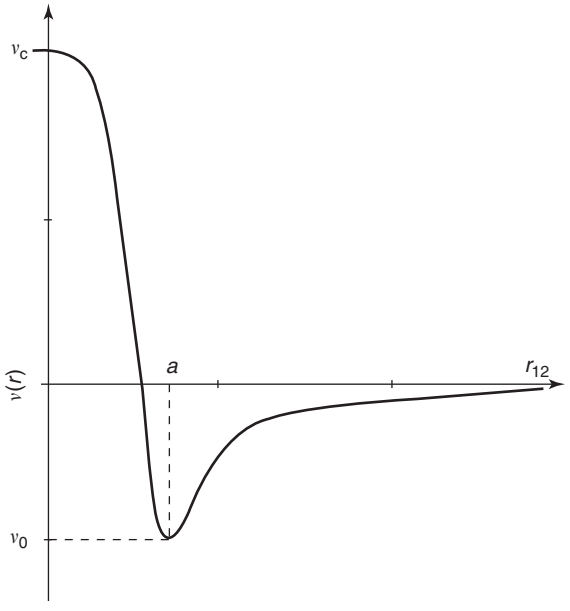


Figure 1 Schematic representation of a two-body force in a system of particles, displayed as a function of the relative particle–particle coordinate $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$. It displays a repulsive core at $r_{12} = 0$ of strength v_c and an attractive minimum of strength v_0 at $r_{12} = a$.

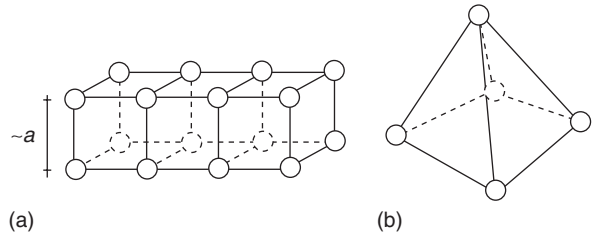


Figure 2 Examples of the state of minimum energy of a system of classical particles described by the Hamiltonian given in eqn [1].

of a particle within a volume of radius a, that is, $\sim\hbar^2/ma^2$. Thus, the particles described by the Hamiltonian given in eqn [1] may be delocalized because the potential energy gain of the classical (localized) configuration may be overwhelmed by the quantal zero-point energy. Consequently, systems for which the "quantality" parameter

$$Q = \frac{\text{zero-point energy of localization}}{\text{potential energy of localization}} \sim \frac{\hbar^2}{ma^2}\frac{1}{|v_0|} \quad [3]$$

is such that $Q \ll 1$ are expected to display particle localization, while systems associated with larger values of Q are expected to display particle delocalization. In the case of nucleons ($^1S_0$ nucleon–nucleon potential), $v_0 = -100$ MeV, $a \approx 1$ fm, $\hbar^2/M \approx 40$ MeV fm^2 ($m = M \approx 10^3$ MeV) and $Q = 0.4$. It is thus expected that nucleons, inside the nucleus, are delocalized.

Because nucleons do not occupy all the space available to them but are self-confined, it is said that the nucleus is a quantum liquid. This is not completely correct, in the sense that the nucleus behaves as a non-Newtonian solid. It deforms plastically under strain and reacts elastically to an instantaneous stress.

In the case where the particles under consideration are Ne_{20} atoms, $m = 20M$, $a \approx 3.1$ Å, and $v_0 = 3.1$ meV. Consequently, $Q \approx 0.007$ and N_{20} is, at $T = 0$, a solid, being a gas at room temperature.

Summing up, the fact that $Q \ll 1$ implies localization and thus spontaneous symmetry breaking, while $Q \sim 1$ does not, is an example of the fact that while potential energy always prefers special arrangements, fluctuations, quantum or classical, favor symmetry.

Mean Field

Because $Q \sim 1$ in the nuclear case, it is likely that the mean-field theory is applicable to the description of the motion of the nucleons. In fact, the marked variation of the binding energy per particle as a function of mass number $A = N + Z$ (cf. **Figure 3**) (where N and Z stand for the number of neutrons and protons, respectively), for specific values of N and of Z (magic numbers), testifies to the fact that nucleons in the nucleus display, in states close to the Fermi energy, a long mean free path, as compared to the size of the nucleus ($\approx$3–7 fm).

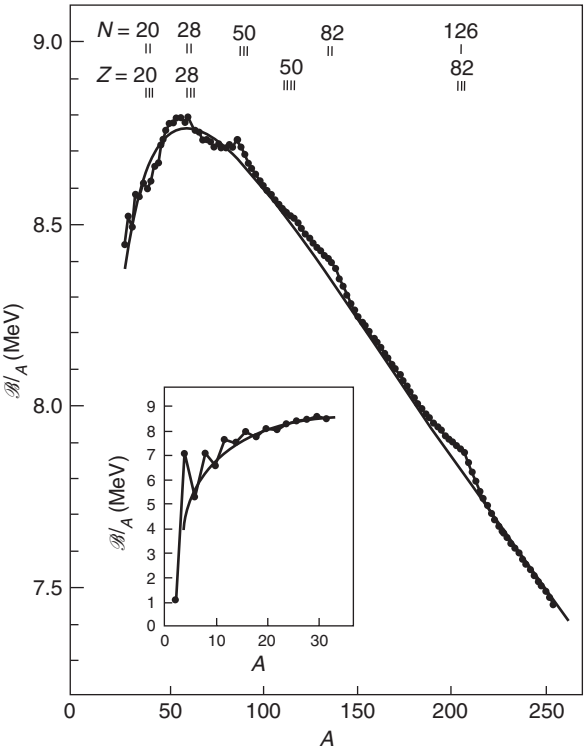


Figure 3 Binding energy per nucleon as a function of mass number.

Special stability is adscribed to nuclei displaying particular values of N and Z (magic numbers), in particular to $^{208}_{82}Pb_{126}$ containing 82 protons and 126 neutrons. In analogy with the atomic case in general, and with the noble gases in particular, this result is connected with the filling of shells, single-particle orbitals of a single-particle potential $U(r)$ generated by all the nucleons, and where each nucleon moves independently. The validity of the independent-particle model implies that the matrix elements of $U(r)$ are much larger than those of $(v(r_{12}) - U(r))$. In other words, that

$$H_{MF} = T + U(r) \quad [4]$$

is a good approximation to eqn [1].

Nucleon–nucleus elastic scattering data and one-particle stripping and pickup experiments provide the information needed to determine the parameter determining $U(r)$, and the energy and occupation probability of single-particle states.

The properties of the single-particle levels close to the Fermi energy, that is, with $E = |\varepsilon - \varepsilon_F| < 5$ MeV

(valence levels) can be quite accurately described by making use of a single-particle Saxon–Woods potential (a spin–orbit and a Coulomb term has to be added),

$$U(r) = \frac{U}{1 + \exp((r - R_0)/a)} \qquad [5]$$

where $R_0 = 1.2A^{1/3}$ fm is the nuclear radius and $a = 0.65$ fm the diffusivity of the potential, while the relation

$$U = U_0 + V_1 \frac{N - Z}{A} \tau_z \qquad [6]$$

defines the depth of the potential, with $U_0 = -45$ to -50 MeV. To describe the centroid of levels with $E > 5$–10 MeV, a term $0.4E$ has to be added to U, where $E = |\varepsilon - \varepsilon_F|$ is the absolute value of the single-particle energy measured from the Fermi energy. In other words, the depth of the empirical Saxon–Woods potential is

$$\begin{matrix} U & (E < 5\,\text{MeV}) \\ U + 0.4E & (E > 5 - 10\,\text{MeV}) \end{matrix} \qquad [7]$$

The last term in eqn [6] is closely related to the symmetry energy in the nuclear mass formula of Weizäcker $V_1 = 30$ MeV and $\tau_z = +1$ for neutrons and $\tau_F = -1$ for protons. This term expresses the fact that one has to pay a price to separate protons from neutrons, or to have a system with a neutron excess, needed in order to decrease the Coulomb repulsion among protons.

Hartree Theory

The mean-field theory, in the first version proposed by Hartree, finds that the energy of the system acquires a minimum by placing the nucleons in the potential

$$U_H(r) = \int d^3r' \rho(r') v(|\boldsymbol{r} - \boldsymbol{r}'|) \qquad [8]$$

where

$$\rho(r) = \sum_{i \in \text{occ}} |\varphi_i(\boldsymbol{r})|^2 \qquad [9]$$

is the density of the system, sum of the modulus squared of the single-particle wave functions, solutions of the Schrödinger equation

$$\begin{aligned} &\left(-\frac{\hbar^2}{2m} \nabla_r^2 + U_H(r) \right) \varphi_j(\boldsymbol{r}) \\ &\quad = -\frac{\hbar^2}{2m} \nabla_r^2 \varphi_j(r) + \sum_{i \in \text{occ}} \int d^3r' \varphi_i^*(\boldsymbol{r}') \\ &\qquad \times v(|\boldsymbol{r} - \boldsymbol{r}'|) \varphi_i(\boldsymbol{r}') \varphi_j(\boldsymbol{r}) \\ &\quad = \varepsilon_j \varphi_j(\boldsymbol{r}) \end{aligned} \qquad [10]$$

The sum in eqn [9] is over all the occupied states, while $v(|\boldsymbol{r} - \boldsymbol{r}'|)$ is the nucleon–nucleon potential (two-body interaction, cf. **Figure 1**). In other words, at equilibrium the system of particles displays a self-consistent relation between density $\rho(r)$ and potential $U(r)$, mediated by the two-body interaction $v(|\boldsymbol{r} - \boldsymbol{r}'|)$. The Hartree potential $U_H(r)$ reproduces rather well, the main features of the empirical Saxon–Woods potential for level close to ε_F.

Equations [9] and [10] are self-consistent partial differential equations, to be solved by iteration, starting from an ansatz for $U_H(r)$ (e.g., square well potential).

Constrained Hartree Theory

It is to be noted that in solving eqns [8]–[10], one has to specify the shape of the system, for example, whether the system is expected to be spherical or deformed. That is, the Hartree theory is always a constrained Hartree theory, in the sense that one looks for a minimum of energy for the particles moving in a potential whose density is either spherical symmetric or behaves as a tensor of rank L. It should be noted that not all shape constrains lead to a self-consistent solution.

For a closed-shell system, it is quite natural that the absolute minimum of the system is spherical (cf. **Figure 4a**), while a system with a number of nucleons outside the closed shell will deform under the polarization effect of these nucleons, just as the Moon and the Sun polarize the Earth (Jahn–Teller effect, cf. **Figure 4b**). In this case, and assuming a quadrupole, axial symmetric deformation,

$$R = R_0(1 + \beta_2 Y_{20}(\theta, \phi)) \qquad [11]$$

is the radius of the Saxon–Woods potential and the system defines a privileged direction in space, violating rotational invariance, in spite of the fact that all nucleons, as well as the Hamiltonian introduced in eqn [1] are invariant with respect to rotations. This result again is an example of the phenomenon of spontaneous symmetry breaking.

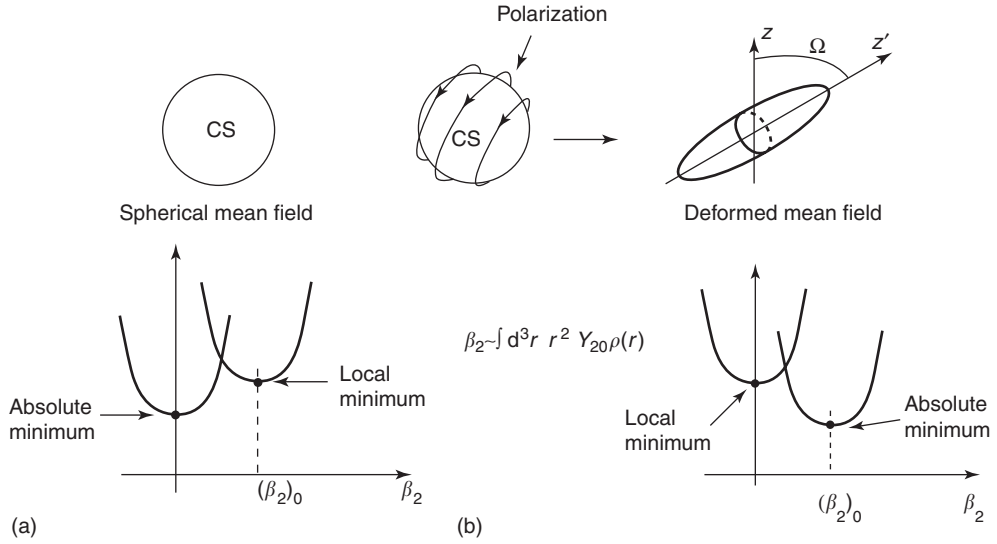


Figure 4 Schematic representation of the shapes of closed-shell (a) and open-shell (b) nuclei. The behavior of the total Hartree energy is also shown as a function of a deformation parameter, in this case the quadrupole deformation parameter.

The ground-state wave function, resulting from the filling of the lowest single-particle orbits of a deformed Saxon–Woods potential (eqns [5], [6] and [10], with $U_{\mathrm{H}}(r) = U(r)$ and [11]), can be written as

$$\Phi_0 = \prod_{i=1}^{A} \varphi_{\nu_i}(\boldsymbol{r}_i) \qquad [12]$$

It is a wave packet of different angular momenta, being associated with fixed values of the Euler angles Ω (**Figure 4b**), that is,

$$\Phi_0(\Omega) = \sum_I c_I \Psi_I \qquad [13]$$

It can be shown that there is a term in the residual interaction $(v - U)$, neglected in H_{MF}, which leads to zero-point fluctuations in the orientation of the system. Taking this term into account leads to solutions of H which transform in an irreducible way with respect to rotations, that is,

$$\Phi_{IMK}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A) \sim \int d\Omega \mathscr{D}^{I*}_{MK}(\Omega)\Phi_0(\Omega; \boldsymbol{r}_1, \ldots, \boldsymbol{r}_A) \qquad [14]$$

Here, $\mathscr{D}^{I*}_{MK}(\Omega)$ is a rotational matrix, $\Omega \sim (\theta, \phi, \psi)$ the corresponding Euler angles, K and M being the projections of the total angular momentum I on the intrinsic axis z' fixed to the body and on the laboratory axis z, respectively (cf. **Figure 5**). The states described by the wave function Φ_{IMK} are the members of a rotational band (**Figure 6**) associated with the different, quantized, frequencies of rotation. Rotational

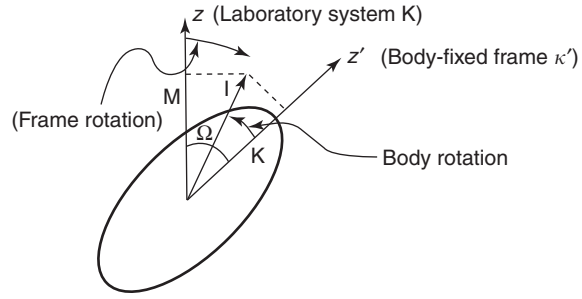


Figure 5 Schematic representation of a deformed rotating nucleus with angular momentum I, and with projection K on the body-fixed frame and M on the laboratory frame.

bands of this type have been observed systematically throughout the mass table.

The mean-field solutions Φ_{IMK} fulfill the requirement that the eigenfunctions of a Hamiltonian, which is invariant with respect to a group of symmetry (rotation group in the present case), should transform according to the irreducible representations of this group, that is, they must behave as a tensor with respect to rotations.

From a mathematical point of view, one can say that Ω and I being conjugate variables, one can choose which one to conserve in working out the solution of the problem. In the present case, it turns out that it is simpler to solve the problem in a representation in which Ω has a fixed value (and, consequently, I is completely indefinite, $\Delta\Omega\Delta I \geqslant \hbar$). Once this solution is found, one can, through a change of representation (Fourier transform), go to the representation in which I is a good quantum number. This is precisely what is accomplished through the integral (projection) given in eqn [14]; a projection which is

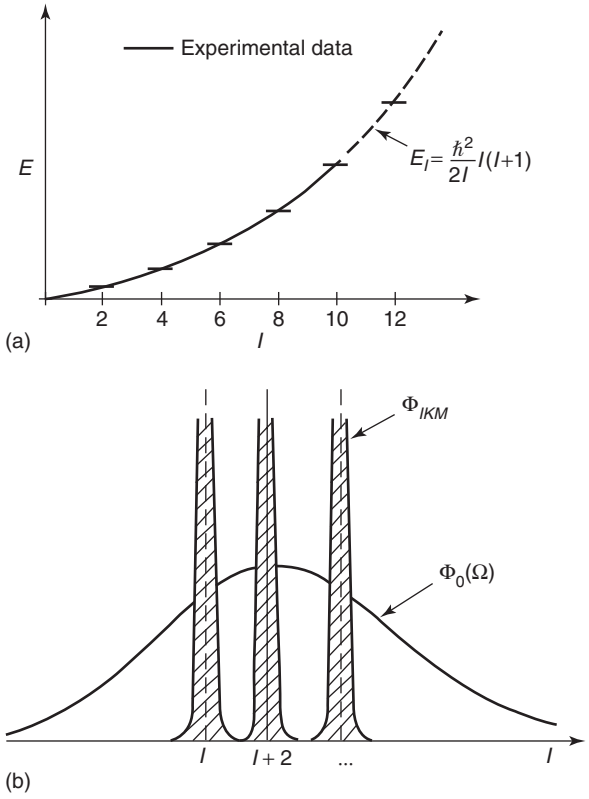


Figure 6 Schematic representation of nuclear rotational bands (a) and of the wave packet $\Phi_0(\Omega)$ of good angular momentum eigenstates Φ_{IKM}.

equivalent to

$$\Phi_{IMK} \sim \int \mathrm{d}\Omega \, \mathrm{e}^{-\mathrm{i}\boldsymbol{\Omega}\cdot\boldsymbol{I}} \Phi_0(\Omega) \qquad [15]$$

in keeping with the fact that

$$\mathscr{D}^{I*}_{MK}(\Omega) = \langle IK|\exp(-\mathrm{i}\boldsymbol{\Omega}\cdot\boldsymbol{I})|IM\rangle \qquad [16]$$

the operator

$$\mathscr{R}(\Omega) = \mathrm{e}^{-\mathrm{i}\boldsymbol{\Omega}\cdot\boldsymbol{I}}$$

generating a general rotation in space.

From a physical point of view, nuclear rotational motion is associated with a spontaneous breaking of rotational symmetry (deformation), a phenomenon which is at the basis of the variety of collective modes displayed by the nuclear many-body system.

In the description of many-body systems having a shape deviating from spherical symmetry (non-spherical nuclei, molecules, etc.), it is convenient to employ an intrinsic (or body-fixed) coordinate frame (**Figure 5**). The transformation of operators from the fixed frame (the laboratory system) to the intrinsic frame involves special features as a result of the fact that the orientation angles of the intrinsic frame are to be regarded as dynamical variables. The states of orientation can be specified by the angular variables or by the associated angular momenta.

In fact, the state of orientation of the body-fixed system is completely specified by the three angular momentum quantum numbers *IKM*, representing the conjugate variables of the three orientation angles $\Omega = (\phi, \theta, \psi)$.

Hartree–Fock Theory

Because nucleons are fermions, the product $\varphi_i(\boldsymbol{r}')\varphi_j(\boldsymbol{r})$ appearing in eqn [10] has to be replaced by the antisymmetric product $\varphi_i(\boldsymbol{r}')\varphi_j(\boldsymbol{r}) - \varphi_i(\boldsymbol{r})\varphi_j(\boldsymbol{r}')$. In this way, a particle does not interact with itself, and eqn [10] can be written as

$$\left[-\frac{\hbar^2}{2m}\nabla_r^2 + U_{\mathrm{H}}(r)\right]\varphi_j(\boldsymbol{r}) + \int \mathrm{d}^3r' U_x(\boldsymbol{r},\boldsymbol{r}')\varphi_j(\boldsymbol{r}') = \varepsilon_j\varphi_j(\boldsymbol{r}) \qquad [17]$$

where the nonlocal potential

$$U_x(\boldsymbol{r},\boldsymbol{r}') = -\sum_{i\in\mathrm{occ}} \varphi_i^*(\boldsymbol{r}')v(|\boldsymbol{r}-\boldsymbol{r}'|)\varphi_i(\boldsymbol{r}) \qquad [18]$$

is the so-called exchange or Fock potential.

It is, as a rule, not simple to deal numerically with nonlocal potentials. There are a number of ways in which one can make local U_x, for example in the LDA. Another way is through the so-called *k*-mass. In fact, for most purposes the effect of the term containing U_x in eqn [17] can be taken into account by replacing m in the kinetic energy term by

$$m_k = m\left(1 + \frac{m}{\hbar^2 k}\frac{\partial \tilde{U}_x}{\partial k}\right)^{-1} \qquad [19]$$

where $\tilde{U}_x(k)$ is the Fourier transform of the exchange potential. Making use of eqn [6] ($U = U_0 + V_1((N-Z)/A)\tau_z + 0.4E$) and of the fact that $E = (\hbar^2/2m)|k^2 - k_{\mathrm{F}}^2|$, one obtains

$$m_k \sim 0.7m \qquad [20]$$

The (local) equation

$$\left(-\frac{\hbar^2}{2m_k}\nabla_r^2 + U'(r)\right)\varphi_j(\boldsymbol{r}) = \varepsilon_j\varphi_j(\boldsymbol{r}) \qquad [21]$$

with $U'(r) = (m/m_k)U_H(r)$, leads to a set of levels which display the following properties:

1. the sequence of levels essentially coincides with those observed experimentally;
2. the density of levels coincide also with that observed experimentally for single-particle states with $E = |\varepsilon - \varepsilon_F| > 5$–10 MeV but is much lower for states with $E \leqslant 5$ MeV (valence levels); and
3. in the mean-field theory, all levels are sharp (infinite mean free path) while, experimentally, levels with $E > 5$–10 MeV display a damping width, that is, a finite lifetime.

The result (1) is quite natural, as $U_H(r)$ is quite similar to a Saxon–Woods potential with the parameters given in connection with eqns [5] and [6]. The property (2) emerges because a Schrödinger equation with $m = M$ (the bare nucleon mass) and a Saxon–Woods potential of depth $U + 0.4E$ gives essentially the same single-particle levels as a Schrödinger equation with $m \sim 0.7M$ and depth equal to $1.4U$ does. This result explains the agreement between theory and experiment concerning the density of levels with $E > 5$–10 MeV.

On the other hand, the density of levels of a Schrödinger equation with $m = M$ and depth U is much higher than the one with $m_k = 0.7M$, as the energy differences $\varepsilon - \varepsilon_F$ resulting from the m_k Schrödinger equation are $\sim(\varepsilon - \varepsilon_F)/0.7$.

Properties (2) and (3) are intimately connected with the fact that the mean-field theory defines a rigid surface, that is, where nucleons reflect elastically on it and each orbital closes on itself (cf. **Figure 7**), the resulting mean free path being infinite. In other words, when considering the motion of an individual nucleon, the self-consistency existing between density and potential (cf. eqn [8]), guarantees that the motions of the $A - 1$ remaining particles adjust at each instant of time to allow a nucleon to move independently of the others, filling their pushings and pullings to change momentum at the surface.

Now, the nuclear surface can be viewed as a quantal membrane with a surface tension $\gamma \approx 1$ MeV fm^{-2}. It is likely that it can fluctuate collectively, thus leading to an effective increase of the nuclear radius and thus to a decrease of the kinetic energy of the nucleons – in other words, to an effective increase of the single-particle level density.

Zero-point fluctuations are also expected to lead to a damping width of single-particle levels lying away from $\varepsilon_F (> 5$ MeV). In fact, assuming an orbital concentrated along the equator, the state will feel a wider and a narrower nucleus, as a function of time, thus giving rise to a distribution of the single-particle strength (cf. **Figure 8**).

Diagonalization of the Complete Hamiltonian

A possible way of obtaining the "exact" wave function of the system is to diagonalize the residual interaction $(v - U)$ in a complete basis of determinants $\Phi_n^0(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A)$ built from single-particle wave functions $\varphi_\nu(\boldsymbol{r})$, solutions of the mean-field Hamiltonian $H_{MF} = T + U$ (cf. **Figure 9**). In other words, one first calculates the absolute and the local minima and associated wave functions of the constrained mean-field Hamiltonian. In this basis, one calculates the matrix elements of $(v - U)$, that is,

$$\int \mathrm{d}^3\{r_i\}\Phi_n^0(\{\boldsymbol{r}_i\})(v - U)\Phi_{n'}^0(\{\boldsymbol{r}_i\}) = \langle v_1 v_2|(v - U)|v_{1'} v_2\rangle O^{A-2} \qquad [22]$$

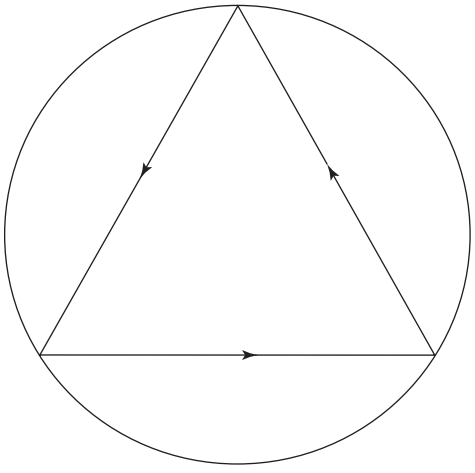


Figure 7 Schematic representation of a single-particle orbital in mean field.

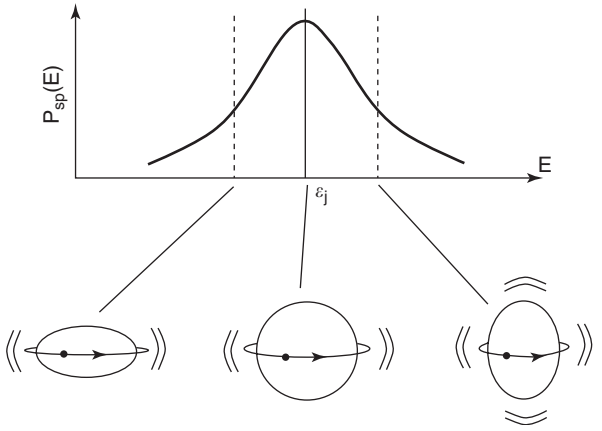


Figure 8 Schematic representation of the single-particle strength function resulting from the coupling of a single-particle orbital to a quadrupole fluctuation of the core.

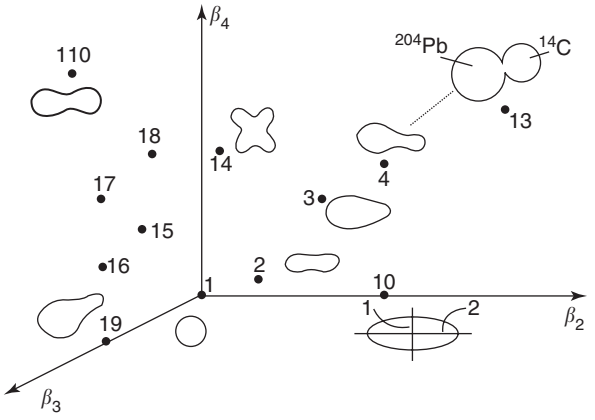


Figure 9 Schematic representation of the space of mean-field solutions of a (heavy) atomic nucleus as a function of the quadrupole (β_2), octupole (β_3), hexadecapole (β_4), etc., deformation parameters.

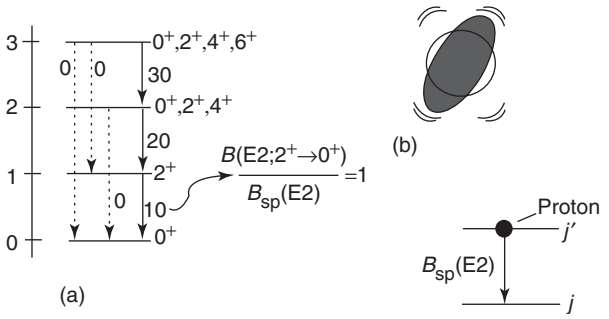


Figure 10 Schematic representation of harmonic quadrupole vibrations containing up to three phonons. The transition probabilities (e.g., $B(E2;2\rightarrow 0) = 10$) is measured in single-particle units implying that it is ten times stronger than a transition induced by the jump of a single proton from an excited state j' to a state j of lower energy.

where $O = \int \mathrm{d}^3 r(\varphi_\nu(\boldsymbol{r}))^*_n(\varphi_{\nu'}(\boldsymbol{r}))_{n'}$ is the overlap between two single-particle wave functions associated with the local minima n and n'. Diagonalizing the resulting matrix, one obtains

$$\Phi = \sum_n a^0_n \Phi^0_n \qquad [23]$$

which is the "exact" solution of the problem.

In fact, to obtain such a solution, one should, in principle, also include particle–hole (and eventually "many particle–many hole") states of the different local and the absolute minimum. In any case, the function Φ given above provides an accurate description of the exact ground state of the system.

Within this picture, the phenomenon of fission and of α- and exotic decay (e.g., of the process $^{223}\mathrm{Ra}\rightarrow{}^{209}\mathrm{Pb} + {}^{14}\mathrm{C}$) are associated with the probabilities $|a^0_{110}|^2$, $|a^0_4|^2$, and $|a^0_{13}|^2$, respectively, associated with conformations in space where a nucleus is necking in (• 1 1 0), or has a bulge of 2 protons and 2 neutrons (• 4), or corresponds to a ^{14}C nucleus on the surface of ^{209}Pb (•13). To achieve such conformations, the system has to move around a number of pairs of particles equal to the emitted particle (e.g., about half of the nuclear mass number in the case of symmetric fission). This corresponds to a crossing of the same number of empty and occupied states as a function of the deformations. The liquid-drop model (surface tension $\gamma = 1\,\mathrm{MeV\,fm^{-2}}$) can be used to calculate the restoring force associated with these deformations, while the inertia is determined by the pairing correlation (superfluidity) of the nucleons (see section on pairing).

Collective Vibration

In the nucleon–nucleus scattering experiments used to determine the (empirical) single-particle potential $U(r)$, one looks only at one reaction channel, namely, the elastic channel, where neither energy nor angular momentum is transferred between the projectile and the target, but only linear momentum is. Of course, there are, in general, other channels which are open at any given bombarding energy, and which act on the elastic (entrance) channel to drain intensity from it (absorption, imaginary part of the optical potential).

Now the focus is on the inelastic channel. Reactions of the type $A(p,p')A^*$, where not only linear momentum (elastic scattering) but also energy and angular momentum are exchanged (inelastic scattering) between a proton p and a nucleus A, which is left in an excited state A^*, have shown that the nuclear surface can vibrate as a whole in well-defined (normal) modes. In particular, quadrupole vibrations, of which states of up to three phonons have been observed (cf. **Figure 10**).

The experimental data can be parametrized at profit making use of the (empirical) harmonic oscillator Hamiltonian

$$H_{\mathrm{coll}} = \sum_{LM}\left(\frac{\pi^2_{LM}}{2D_L} + \frac{|\alpha_{LM}|^2}{2C_L}\right) \qquad [24]$$

where D_L and C_L are the inertia and the restoring forces of the mode with multipolarity L. Note that the static quadrupole deformation of the mean field (eqn [11] and **Figure 5**) can be viewed as a quadrupole vibration with a restoring force C_L that becomes zero.

The parameters C_L and D_L (which can also be calculated by making use of $(v - U)$, cf. the section

on collective vibrations), can be determined from the experimental data, that is, from the energy $\hbar\omega_L$ and the transition probability $B(EL; 0 \rightarrow L)$ of the modes, where

$$\hbar\omega_L = \hbar\sqrt{\frac{C_L}{D_L}} \qquad [25]$$

and

$$B(EL; 0 \rightarrow L) \sim |\langle n_L = 1|\alpha_{LM}|0\rangle|^2 = \frac{\hbar\omega_L}{2C_L} \sim \beta_L^2 \qquad [26]$$

in keeping with the fact that the coordinate operator of a harmonic oscillator can be written in second quantization as

$$\alpha_{LM} = \sqrt{\frac{\hbar\omega_L}{2C_L}}(\Gamma_{LM}^\dagger + \Gamma_{LM}) \qquad [27]$$

Here $|LM\rangle = |n_L = 1\rangle = \Gamma_{LM}^\dagger|0\rangle$ is the one-phonon state, $\Gamma_{LM}^\dagger$ and Γ_{LM} being the creation and annihilation boson operators fulfilling the commutation relation

$$[\Gamma_{LM}, \Gamma_{LM}^\dagger] = 1 \qquad [28]$$

In other words, α_{LM} is the coordinate of the harmonic oscillator (Dirac representation). The Hamiltonian H_{coll} (eqn [24]) can then be written as

$$H_{\text{coll}} = \sum_{LM}\left(\Gamma_{LM}^\dagger\Gamma_{LM} + \frac{1}{2}\right)\hbar\omega_L \qquad [29]$$

Consequently, with each degree of freedom there is associated a zero-point energy $\frac{1}{2}\hbar\omega_L$ and thus a zero-point fluctuation (cf. **Figure 10b**) leading to fluctuations of the nuclear radius (conserving volume so as to keep the density unchanged), which can be parametrized as (cf. also eqn [11])

$$R = R_0\left(1 + \sum_{LM}\alpha_{LM}Y_{LM}^*(\hat{r})\right) \qquad [30]$$

in keeping with the fact that the spherical harmonics provide a complete basis of eigenfunctions of angular momentum.

Beyond Mean Field: Particle–Vibration Coupling

Inserting the expression given in eqn [30] in the expression of the empirical potential given in eqn [5] and expanding to lowest order in α_{LM} (note that $\beta_L^2 \ll \beta_L$, where $\beta_L = \sqrt{2L+1}\sqrt{\hbar\omega_L/(2C_L)}$), one obtains

$$H = H_M + H_{\text{coupl}} + H_{\text{coll}} \qquad [31]$$

where

$$H_{\text{coupl}} = -\kappa\hat{\alpha}\hat{F} \qquad [32]$$

with

$$\hat{F} = \sum_{\nu_1\nu_2}\langle\nu_1|F|\nu_2\rangle a_{\nu_1}^\dagger a_{\nu_2} \qquad [33]$$

and

$$F = -\frac{1}{\kappa}R_0\frac{\partial U(r)}{\partial r}Y_{LM}^*(\hat{r}) \qquad [34]$$

The Hamiltonian H_{coupl} thus couples the motion of a single nucleon with the collective vibrations of the surface, with a matrix element (cf. **Figure 11**):

$$\langle n_\alpha = 1, \nu'|H_{\text{coupl}}|\nu\rangle = \Lambda_\alpha\langle\nu'|F|\nu\rangle = \langle n_\alpha = 1, \nu\nu'|H_{\text{coupl}}|0\rangle \qquad [35]$$

where

$$\Lambda_\alpha = -\kappa\sqrt{\frac{\hbar\omega_\alpha}{2C_\alpha}} \sim -\frac{\kappa\beta_\alpha}{\sqrt{2L_\alpha + 1}} \qquad [36]$$

is the particle–vibration coupling strength. Because $\beta_L^2 \ll \beta_L$, one can usually treat the particle–vibration coupling in the weak coupling situation. Consequently, H_{coupl} can be treated in the perturbation

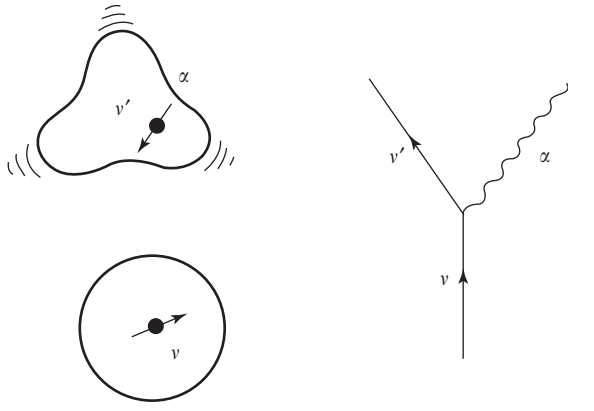


Figure 11 Schematic representation of the process by which a nucleon excites the vibrations of the surface.

theory. To second order, one finds

$$\left(-\frac{\hbar^2}{2m}\nabla_r^2 + U_{\mathrm{H}}(r)\right)\varphi_{\gamma}(r) + \int \mathrm{d}^3r' U_x(\boldsymbol{r},\boldsymbol{r}')\varphi_j(\boldsymbol{r}') + (\Delta E + \mathrm{i}W_j)\varphi_j(\boldsymbol{r}) \sim \left(-\frac{\hbar^2}{2m_k}\nabla_r^2 + U_{\mathrm{H}}''(r) + \Delta E_j + \mathrm{i}W_j\right)\varphi_j(\boldsymbol{r}) \quad [37]$$

$$= \varepsilon_j\varphi_j(\boldsymbol{r}), \quad \left(U_{\mathrm{H}}'' = \frac{m}{m_k}U\right) \quad [38]$$

where (cf. **Figure 12**) $\Delta E_j^{(\omega)}$ and $W_j^{(\omega)}$ are the real and imaginary contributions to the self-energy calculated in the second-order perturbation theory.

For most purposes, ΔE can be treated in terms of an effective mass

$$m_\omega = m(1+\lambda) \quad [39]$$

where

$$\lambda = -\frac{\partial \Delta E}{\partial \omega} \quad [40]$$

is the mass enhancement factor, while $Z_\omega = m/m_\omega$ is the quasiparticle strength (discontinuity of the Fermi energy).

Consequently, eqn [38] can be rewritten as

$$\left(-\frac{\hbar^2}{2m^*}\nabla_r^2 + U_{\mathrm{H}}' + \mathrm{i}W(\omega)\right)\varphi_j(\boldsymbol{r}) = \varepsilon_j\varphi_j(\boldsymbol{r}) \quad [41]$$

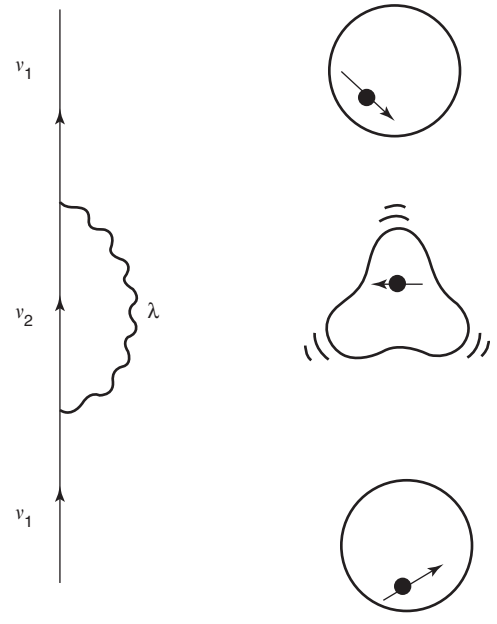


Figure 12 Self-energy graph for a single particle.

with

$$m^* = \frac{m_k m_\omega}{m} \quad [42]$$

and $U_{\mathrm{H}}' = (m/m^*)U$. Because $\lambda \sim 0.5$ (i.e., the dressed single-particle m_ω is heavier than the bare nucleon, as it has to carry a vibration along), $m^* \approx 1$, and $Z_\omega \approx 0.7$. Furthermore, due to the fact that $\hbar\omega_\alpha \approx 2$–$2.5$ MeV, the range of single-particle energy $E = |\varepsilon - \varepsilon_{\mathrm{F}}|$ over which the particle–vibration coupling processes displayed in **Figure 12** is effective and is approximately $\pm 2\hbar\omega_\alpha \sim 4$–$5$ MeV.

It should be noted that ΔE_j indicates the shift in energy of the energy centroid of the "dressed" single-particle state due to the coupling to the intermediate (more complex states) $\alpha' \equiv (\nu', \alpha)$, while $\Gamma = 2W$ measures the energy range over which the single-particle state spreads due to the coupling. While all states contribute to ΔE ("off the energy shell process," i.e., processes which do not conserve the energy), essentially only "on the energy processes," that is, processes which conserve the energy, contribute to Γ. The quantity $\hbar/\Gamma$ is the lifetime of the single-particle state, while Γ is the range of energy over which the single-particle state is distributed due to processes of the type shown in **Figure 12** (cf. also **Figure 8**).

Induced Interaction

A nucleon at the Fermi energy which creates, by bouncing inelastically off the nuclear surface, has no other choice but to reabsorb it at a later instant of time (virtual process, **Figure 12**). In the presence of another nucleon, the vibration excited by one nucleon may be absorbed by the second one (**Figure 13**),

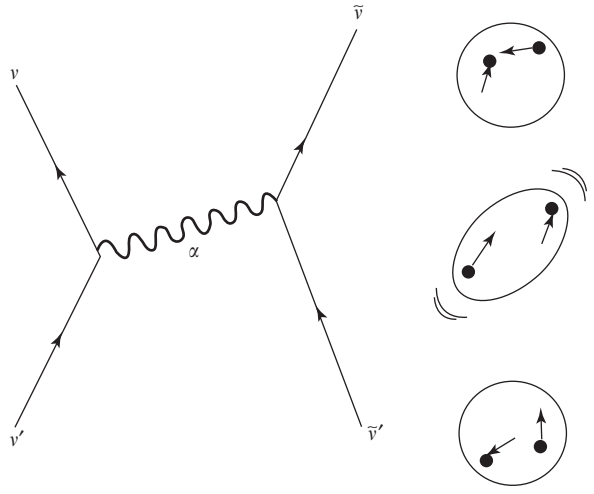


Figure 13 Schematic representation of the exchange of phonons between nucleons.

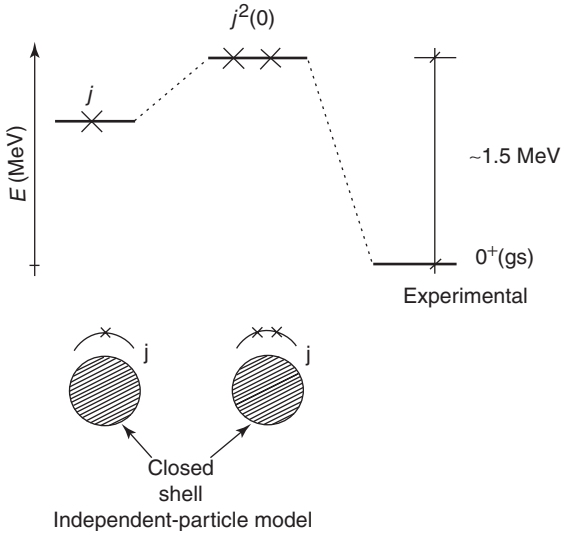


Figure 14 Schematic representation of the predictions of the independent-particle model for one and two particles outside the closed shell, in comparison with the experimental findings (e.g., for the case of ^{210}Pb, where $j = g_{9/2}$).

the exchange of a vibration leading to an (induced) interaction.

Simple estimates of this induced interaction lead to the case of, for example, ^{210}Pb, that is, of two neutrons outside the closed-shell nucleus ^{208}Pb, to values of the matrix element for pairs of particles coupled to angular momentum $J^\pi = 0^+$ of nearly -0.8 MeV, when summed over all the different multipolarities of low-lying collective surface vibrations which have been observed ($L^\pi = 2^+, 3^-, 5^-$). The fact that one considers particles coupled to angular momentum zero is because the associated orbitals have maximum overlap, which best considers the (pairing) interaction. In the case of two particles outside the closed shell, one would then expect the ground state to display, due to this mechanism, a correlation energy of 0.8 MeV larger than that predicted by the independent-particle model (cf. **Figure 14**). Experimentally, this correlation is ~ 1.5 MeV, indicating that the bare nucleon–nucleon interaction provides about half of the pairing interaction required by the experimental findings.

From this result, one can conclude that the pairing interaction induced by the process depicted in **Figure 13**, renormalizes the bare nucleon–nucleon potential significantly.

Collective Vibrations: Microscopic Description

In a self-substained vibration (normal mode), the variations of the potential must be self-consistent

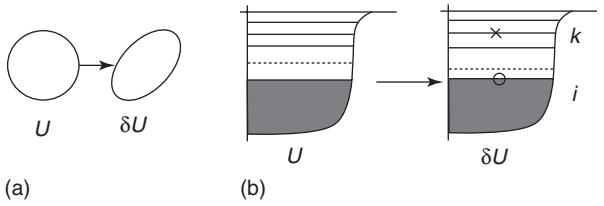


Figure 15 Complementary views of the nucleus: (a) as a system confined by an elastic surface, or (b) as a many-particle determinant.

with those of the density. In other words, the relation given in eqn [8] is also required to be valid for the dynamical situation, that is,

$$\delta U = \int \mathrm{d}^3 r' \delta\varrho v(|\boldsymbol{r} - \boldsymbol{r}'|) \qquad [43]$$

In fact, one can excite the system: (a) by acting collectively on the nucleus and deforming slightly the potential U (thus leading to δU), (b) promoting a particle from below to above the Fermi energy with a given probability (thus leading to $\delta\varrho$) (**Figure 15**), in keeping with the fact that a single-particle external field can change the state of motion of one particle at a time.

Consequently, collective vibrations in nuclei can be viewed as correlated particle–hole excitation, created by the operator

$$\Gamma^\dagger_\alpha = \sum_{ki} X^\alpha_{ki}\Gamma^+_{ki} + Y^\alpha_{ki}\Gamma_{ki} \qquad [44]$$

expressing the fact that while the collective modes look very simple in the collective representation of vibrations, they appear quite a bit more complicated (although entirely microscopic), in the particle–hole basis of the operators $\Gamma^+_{ki} = a^\dagger_k a_i$ and $\Gamma_{ki} = a^\dagger_i a_k$, where k denotes a level above the Fermi energy and i one below (cf. **Figure 15b**). Within the approximation where Γ^+_α and Γ_α fulfill boson commutation relation, Γ^+_{ki} and Γ_{ki} also are considered as (quasi-) boson operators. Within this approximation, and making use of the unitary transformation expressed by eqn [44], one can show that $\hat{\alpha} = \hat{F}$. Consequently,

$$H_{\mathrm{coupl}} = -\kappa\hat{\alpha}\hat{F} = -\kappa\hat{F}\hat{F} \qquad [45]$$

From the relations

$$[H_{MF} + H_{\mathrm{coupl}}, \Gamma^+_\alpha] = \hbar\omega_\alpha\Gamma^+_\alpha$$

and

$$[\Gamma_\alpha, \Gamma^+_{\alpha'}] = \delta(\alpha, \alpha')$$

valid for harmonic vibrations, one obtains, for the case of, for example, a dipole vibration

$$\hbar\omega_{\mathrm{D}} \sim \frac{80}{A^{1/3}}\ \mathrm{MeV} \sim \frac{100}{R}\ \mathrm{MeV} \qquad [46]$$

($R = 1.2A^{1/3}$ fm), a number which is very close to the experimental value (cf. **Figure 16**). The inverse dependence on the radius (momentum dependence) is typical of elastic vibrations.

In the case of quadrupole vibrations, one obtains

$$\hbar\omega_{\mathrm{Q}} = \frac{60}{A^{1/3}}\ \mathrm{MeV} \approx \frac{72}{R}\ \mathrm{MeV} \qquad [47]$$

which provides an overall account of the experimental findings.

In other words, the nucleus reacts elastically with a dipole mode of energy $\sim$16 MeV and carrying $\sim$100% of the energy weighted sum rule (EWSR) and with $\sim$l2 MeV and again $\sim$100% of the EWSR to an external quadrupole field.

While the mean-field theory predicts the states to be sharp, experimentally, they display a resonant behavior with a width of the order of 30% of the centroid energy. Similar to what was found in the case of single-particle motion, the coupling of the giant resonance to the zero-point fluctuations of the ground state accounts for the spreading width. This is because elastic vibrations are inversely proportional to the nuclear radius (cf. eqn [47]). In particular, in the case of the dipole resonance, quadrupole fluctuations of the nuclear surface are particularly effective (note that a 1^- can, to give again 1^-, couple only to 0^+ and to 2^+).

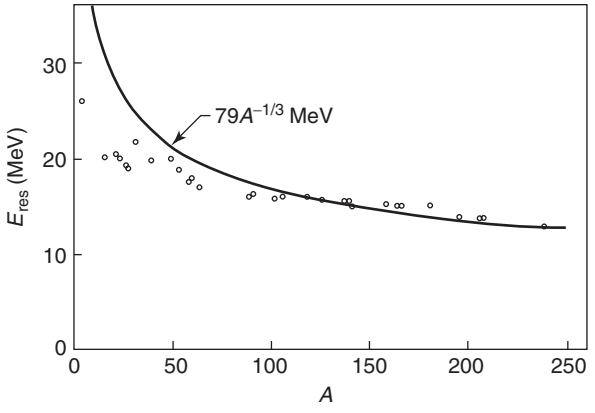


Figure 16 Schematic representation of the energy centroid of the GDR (giant dipole resonance). The open circles represent experimental data.

Pairing: Experimental Facts

The nuclear binding energy is found to exhibit a systematic variation, depending on the evenness or oddness of Z and N,

$$\delta\mathscr{B} = \begin{cases} \Delta & Z \text{ even}, N \text{ even} \\ 0 & A \text{ odd} \\ -\Delta & Z \text{ odd}, N \text{ odd} \end{cases} \qquad [48]$$

where

$$\Delta = -\frac{1}{2}\{\mathscr{B}(N-1,Z) + \mathscr{B}(N+1,Z) - 2\mathscr{B}(N,Z)\}$$

and $\mathscr{B}(N,Z)$ is the total binding energy and represents the difference between the observed mass $\mathscr{M}$ (or, equivalently, the total nuclear energy $\mathscr{E}$) in the ground state, and the masses of the separated nucleons,

$$\mathscr{M}(N,Z) = \frac{1}{c^2}\mathscr{E}(N,Z) = NM_n + ZM_p - \frac{1}{c^2}\mathscr{B}(N,Z)$$

The observed pairing energies (**Figure 17**) can be parametrized according to

$$\Delta \sim \frac{12}{\sqrt{A}}\ \mathrm{MeV} \qquad [49]$$

The large odd–even effect observed may be described in terms of pairwise correlations of identical particles coupled to angular momentum zero, which contribute an additional binding energy 2Δ per pair, for nucleons on top of the Fermi surface. Consistent with the above findings, the lowest noncollective excitations in even–even nuclei are found at an energy of $\sim 2\Delta$.

Pairing Interaction

A schematic interaction producing such an effect is the pairing force with constant matrix elements,

$$H_{\mathrm{p}} = -G \sum_{\nu_1 \nu' > 0} a^\dagger_\nu a^\dagger_{\bar{\nu}} a_{\bar{\nu}'} a_{\nu'} \qquad [50]$$

where the state $a^\dagger_{\bar{\nu}}|0\rangle = |\bar{\nu}\rangle$ is the time-reversal state to $a^\dagger_\nu|0\rangle = |\nu\rangle$.

The mean-field (BCS) solution leads to

$$(H_{\hat{p}})_{\mathrm{MF}} = U + H_{11} \qquad [51]$$

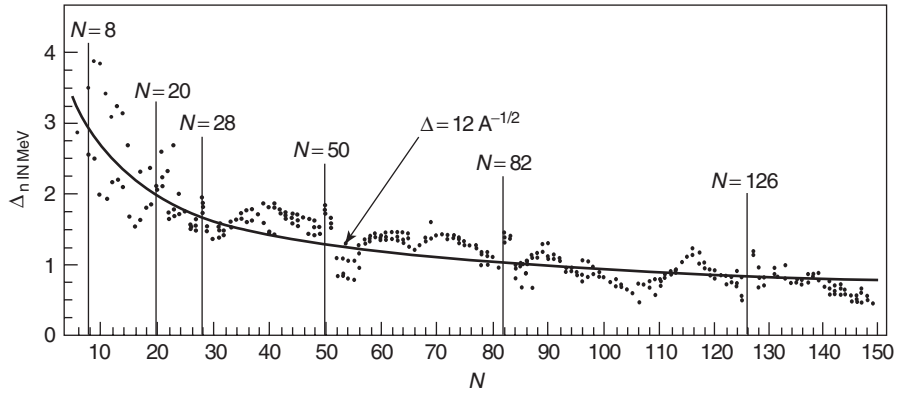


Figure 17 Odd–even mass difference for neutrons.

where

$$U = 2\sum_{v>0}(\varepsilon_v - \lambda)V_v^2 - \frac{\Delta^2}{G} \qquad [52]$$

and

$$H_{11} = \sum_{v>0} E_v \alpha_v^\dagger \alpha_v \qquad [53]$$

where

$$E_v = \sqrt{(\varepsilon_v - \lambda)^2 + \Delta^2} \qquad [54]$$

is the quasiparticle energy, while

$$\alpha_v^+ = U_v \alpha_v^+ - V_v \alpha_{\bar{v}}$$

is the quasiparticle creation operator. The occupation amplitudes U_v, V_v are

$$U_v = \frac{1}{\sqrt{2}}\left(1 + \frac{\varepsilon_v - \lambda}{E_v}\right)^{1/2}$$
$$V_v = \frac{1}{\sqrt{2}}\left(1 - \frac{\varepsilon_v - \lambda}{E_v}\right)^{1/2}$$

The parameters U_v and V_v which completely define the BCS mean-field solution of the pairing Hamiltonian depend on the Fermi energy λ and the gap parameter Δ. The equations determining these parameters are the number equation

$$N = \langle \mathrm{BCS}|\hat{N}|\mathrm{BCS}\rangle = 2\sum_{v>0} V_v^2 \qquad [55]$$

where

$$\hat{N} = \sum_{v>0} a_v^\dagger a_v \qquad [56]$$

is the particle-number operator, and the gap equation

$$\Delta = G\langle \mathrm{BCS}|P^+|\mathrm{BCS}\rangle = G\sum_{v>0} U_v V_v \qquad [57]$$

In other words, equations

$$\begin{aligned} N &= 2\sum_{v>0} V_v^2 \quad \text{(number equation)} \\ \frac{1}{G} &= \sum_{v>0}\frac{1}{2E_v} \quad \text{(gap equation)} \end{aligned} \qquad [58]$$

allow one to calculate λ and Δ from the knowledge of ε_v and G.

While the ground-state energy is equal to U, the energy of the lowest excited state, that is, that of a two-quasiparticle state, generalization of a particle–hole excitation in the case of normal ($G = 0$) systems, is

$$\begin{aligned} H_{11}|v_1 v_2\rangle &= H_{11}\alpha_{v_1}^\dagger \alpha_{v_2}^\dagger|\mathrm{BCS}\rangle \\ &= (E_{v_1} + E_{v_2})|v_1 v_2\rangle \end{aligned} \qquad [59]$$

Consequently, there are no excited states with energy less than 2Δ, the minimum value of $(E_{v_1} + E_{v_2})$.

By properly adjusting G for a given value of particles N, one can make this excitation coincide with twice the value given in eqn [49]. This value is

$$G \sim \frac{25}{A}\ \text{MeV} \qquad [60]$$

Making use of this value and of the empirical relation $\Delta = 12/\sqrt{A}$ MeV, one can write

$$\frac{\Delta^2}{G} \sim 5\ \text{MeV}$$

for the second term in eqn [52]. It should be noted that this gain in binding energy is partially compensated by an increase in single-particle energy

associated with the first term and by the fractional occupation of the single-particle levels around the Fermi surface. The summed effect amounts to only $\sim 1\,\text{MeV}$ ($\sim\Delta$) of extra binding energy acquired by the even–even system due to pairing with respect to the odd–even system. This quantity, which is very important to characterize the structure of a nucleus close to the ground state, is still very small compared to the total binding energy of the system ($\sim A \times 8\,\text{MeV} \approx 1\,\text{GeV}$).

In other words, a system of independent particles is affected only on a small region ($\sim 2\Delta$) around the Fermi energy as compared to this energy ($2\Delta/\lambda = 2\Delta/\varepsilon_F \sim 2-3\,\text{MeV}/36\,\text{MeV} \sim 0.1$). What is actually modified is the occupation of the single-particle levels around λ. For single-particle states fulfilling the condition $|\varepsilon_v - \lambda| \gg 1.5\Delta$, the system retains the single-particle properties. For single-particle states such that $|\varepsilon_v - \lambda| \lesssim 1.5\Delta$, the occupancy of the levels is strongly modified and the system is now made of nucleons coupled to angular momentum $J = 0$ (singlet states, i.e., $S = 0$ and $L = 0$, $J = L + S$). These pairs are known as Cooper pairs, and they behave like bosons. In nuclei, the number of Cooper pairs is small, typically 4–6. One would thus expect strong fluctuations of the associated pairing gap, fluctuations which may blur many of the sharp properties found in the case of infinite systems (bulk matter, thermodynamic limit).

BCS Wave Function

Because the wave function $|\text{BCS}\rangle$ is the quasiparticle vacuum, it does not have a fixed number of particles. In fact, making use of the gauge transformation

$$\mathscr{G}(\phi) = \mathrm{e}^{-(\mathrm{i}\hat{N}/2)\phi} \qquad [61]$$

which measures the particle-number properties associated with any operator or wave function, in the same way in which

$$\mathscr{R}(\Omega) = \mathrm{e}^{-\mathrm{i}\boldsymbol{I}\cdot\boldsymbol{\Omega}} \qquad [62]$$

probes the tensor properties of operators and of wave functions, one can determine the particle-number content of operators and wave functions. In particular,

$$a_v'^{\dagger} = \mathscr{G}(\phi) a_v^{\dagger} \mathscr{G}^{-1}(\phi) = \mathrm{e}^{-(\mathrm{i}/2)\phi} a_v^{\dagger} \qquad [63]$$

while

$$\begin{aligned} \mathscr{G}(\phi) H \mathscr{G}^{-1}(\phi) &= \mathscr{G}(\phi)(H_{\text{sp}} + H_{\text{p}})\mathscr{G}^{-1}(\phi) \\ &= H_{\text{sp}} + H_{\text{p}} \end{aligned} \qquad [64]$$

On the other hand,

$$\mathscr{G}(\phi) V_{\text{p}} \mathscr{G}^{-1}(\phi) = -\Delta(\mathrm{e}^{-\mathrm{i}\phi} P^{\dagger} + \mathrm{e}^{\mathrm{i}\phi} P) + \frac{\Delta^2}{G} \qquad [65]$$

and

$$a_v'^{\dagger} = \mathscr{G}(\phi) \alpha_v^{\dagger} \mathscr{G}^{-1}(\phi) = U_v \mathrm{e}^{(\mathrm{i}/2)\phi} a_v^{\dagger} - V_v \mathrm{e}^{(\mathrm{i}/2)\phi} a_{\bar{v}} \qquad [66]$$

In other words, while the total Hamiltonian $H_{\text{sp}} + H_{\text{p}}$ conserves the number of particles, neither the pairing mean field V_{p} nor the quasiparticle transformation does so. It is then not suprising that the wave function $|\text{BCS}\rangle$ is a wave packet in the number of particles and that the correct way to express it is in an intrinsic system of reference, that is,

$$\begin{aligned} &|\text{BCS}(\Phi = 0)\rangle_{\mathscr{K}'} \\ &\quad = \prod_{v>0} (U_v + V_v a_v'^{\dagger} a_{\bar{v}}'^{\dagger}) |0\rangle \\ &\quad = \prod_{v>0} (U_v + \mathrm{e}^{-\mathrm{i}\phi} a_v^{\dagger} a_{\bar{v}}^{\dagger}) |0\rangle \\ &\quad = \left(\prod_{v>0} U_v \right) \sum_n \mathrm{e}^{-\mathrm{i}n\phi} \left(\sum_{v>0} c_v a_v^{\dagger} a_{\bar{v}}^{\dagger} \right)^n |0\rangle \\ &\quad = |\text{BCS}(\Phi)\rangle_{\mathscr{K}} \end{aligned} \qquad [67]$$

The BCS state expressed in the intrinsic reference frame has gauge angle $\phi = 0$. It is equal to the BCS state referred to the laboratory system $|\text{BCS}(\phi)\rangle_{\mathscr{K}}$, which is then rotated by an angle ϕ. Consequently, the BCS wave function does not have a fixed number of particles, but a fixed orientation in gauge space (cf. **Figure 18**).

There is nothing wrong about this. In fact, one finds that it is easier to solve the pairing problem for fixed ϕ, and not N, in a similar way that another problem may be simpler to solve for fixed p instead of fixed r, or fixed Ω instead of I. Because ϕ and N are conjugate variables, in a similar way that p and x,

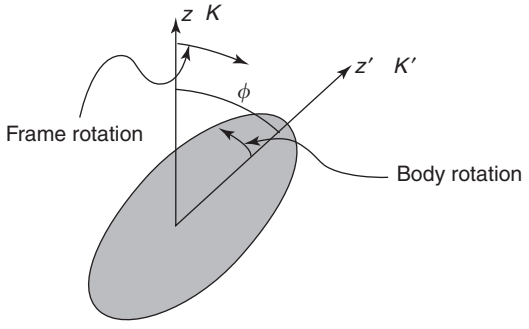


Figure 18 Schematic representation of the deformation in gauge space associated with the wave function $|\text{BCS}\rangle$ (eqn [67]).

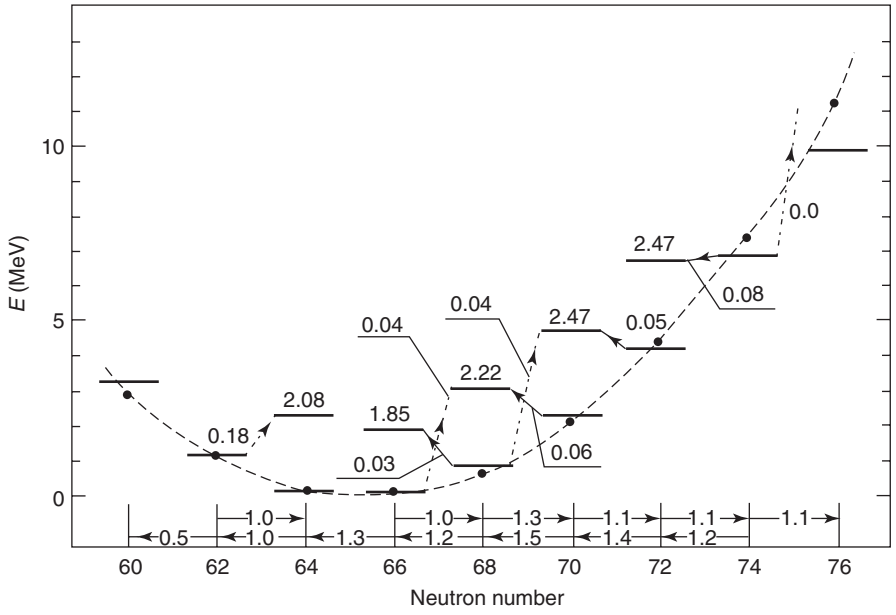


Figure 19 Experimental energies of the $J^{\pi} = 0^{+}$ states of the even Sn isotopes. The heavy drawn lines represent the values of the expression $E = -B(N) + E_{\text{exc}} + 8.58N + 45.3$ (MeV), $B(N)$ being the binding energies (in MeV) of the Sn isotopes. The dashed line represents the parabola $0.10(A - 65.4)^2$, which corresponds to a rotational energy parameter $\hbar^2/2\mathcal{J} = 0.10$ MeV. Also displayed is the excited pairing rotational band associated with the pairing vibrational mode. In all cases where more than one $J^{\pi} = 0^{+}$ state has been excited below 3 MeV in two-neutron transfer processes, the energy $\sum_i \sigma(0_i)E(0_i^+)/\sum_i \sigma(0_i^+)$ of the centroid is quoted, as well as the corresponding cross section $\sum_i \sigma(0_i^+)$. The quantity $\sigma(0_i^+)$ is the relative cross section with respect to the ground-state cross section. The numbers along the abscissa are the ground state (p, t) and (t, p) cross sections normalized to the ^{116}Sn(gs)$\leftrightarrow{}^{118}$Sn(gs) cross section.

and I and Ω are, one can transform the BCS solution so as to conserve the particle number through a Fourier transform. Thus, the solution of $H_{\text{sp}} + H_{\text{p}}$ with fixed particle number is then

$$|N_0\rangle_{\mathscr{K}} \sim \int \mathrm{d}\phi \mathrm{e}^{\mathrm{i}N_0/2}|\mathrm{BCS}(\phi)\rangle_{\mathscr{K}} \sim \left(\sum_{\nu>0} c_\nu a_\nu^\dagger a_{\bar{\nu}}^\dagger\right)^{N_0/2}|0\rangle \qquad [68]$$

These are the members of a pairing rotational band, that is, the ground states of a series of isotopes such as the Sn isotopes, which are strongly excited in two-particle transfer reaction (cf. **Figure 19**). In any case, the spontaneous symmetry breaking of particle conservation defines an emergent property in the nuclear system, not contained in the original Hamiltonian, namely, a generalized rigidity which defines a privileged orientation in gauge space, and which expresses itself in nature as a collective pairing rotational band.

See also: Molecular Clusters; Nuclear Fission and Fusion; Quantum Mechanics: Elementary Particles; Quantum Mechanics: Foundations; Quantum Mechanics: Nuclei.

PACS: **21.60.** – **n; 97.60.Jd; 21.30.Fe; 21.60.Jz; 21.10.** – **k; 23.20.Lv; 21.10.** – **k**

Further Reading

Bertsch GF and Broglia RA (1994) *Oscillations in Finite Quantum Systems*. Cambridge: Cambridge University Press.

Bohr A and Mottelson BR *Nuclear Structure*, vols. I and II, Benjamin (1969) and (1975).

Bortignon PF, Bracco A, and Broglia RA (1998) *Giant Resonances: Nuclear Structure at Finite Temperatures*. Amsterdam: Harwood Academic Publishers.

Brink D and Broglia RA (2004) *Nuclear Superfluidity. Pairing in Finite System*. Cambridge: Cambridge University Press.

Ring P and Schuck P (1980) *The Nuclear Many-Body Problem*. Heidelberg: Springer.

Numerical Approximation and Analysis

C Brezinski, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France

Introduction

When the solution of a mathematical problem cannot be obtained in an analytical form (e.g., the solution of a differential equation) or when it requires too many arithmetical operations (as the solution of a system of linear equations by Cramer's rule), then it is necessary to make use of a numerical method. Usually, such a method only provides an approximation of the solution. Numerical analysis is the branch of mathematics where constructive methods (that is methods able to construct effectively, numerically, the solution) are defined and studied.

Numerical approximation covers the approximation of functions in various senses (interpolation, best approximation, Padé approximation, etc.), the numerical solution of differential and integral equations, the computation of definite integrals, the solution of systems of linear and nonlinear equations, the computation of eigenelements, etc. It is beyond the scope of this article to enter into details and the aim of the article is to give a flavor of the domain and to describe the main ideas.

A numerical method is implemented on a computer. So, each arithmetical operation is affected by an error since computers work with numbers having a finite number of digits. So, the result which is finally obtained depends not only on the quality of the approximation used but also on the propagation of these rounding errors. Taking this point into account is fundamental when using a numerical method.

Approximation of Functions

In many situations, one needs to approximate a function f by another function g. The procedure to be used depends highly on the information on f and on the form of g. The function g can be a polynomial or a combination of given functions satisfying some conditions, or a rational function, or a spline (piecewise polynomials). If f is known at some given points x_i, then interpolation can be used. Interpolation determines the unknown parameters involved in g and provides approximations $g(x_i)$ of $f(x_i)$. If the numbers of parameters are less than the numbers of interpolation points, f can be approximated in the least squares sense. If the information on f is analytic, approximation in the sense of some norm can be used.

If the Taylor expansion of f around zero is known, a quite powerful procedure is Padé approximation. It is assumed that the formal expansion

$$f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$$

is known. A rational approximation is derived with a numerator $N_p(z) = a_0 + \cdots + a_p z^p$ and a denominator $D_q(z) = b_0 + \cdots + b_q z^q$ so that the series expansion of $N_p(z)/D_q(z)$ agrees with the expansion of f as far as possible, which means

$$f(z)D_q(z) - N_p(z) = \mathcal{O}(z^{p+q+1})$$

Such a rational function is called a Padé approximant and it is denoted by $[p/q]_f(z)$. The coefficients of N_p and D_q given by

$$\begin{aligned} a_0 &= b_0 c_0 \\ a_1 &= b_0 c_1 + b_1 c_0 \\ &\vdots \\ a_p &= b_0 c_p + \cdots + b_q c_{p-q} \\ 0 &= b_0 c_{p+1} + \cdots + b_q c_{p-q+1} \\ &\vdots \\ 0 &= b_0 c_{p+q} + \cdots + b_q c_p a_0 \end{aligned}$$

with $c_i = 0$ for $i < 0$. Setting $b_0 = 1$ and solving the system formed by the last q equations gives the b_i. Then the a_i are directly obtained by the first p relations.

Padé approximants are widely used in many problems of mathematical physics.

Ordinary Differential Equations

Consider a Cauchy problem for a system of k ordinary differential equations

$$\begin{aligned} y'(x) &= f(x, y(x)), \quad x \in [a, b] \\ y(a) &= y_0 \end{aligned}$$

where $f : \mathbb{R} \times \mathbb{R}^k \mapsto \mathbb{R}^k$, $y \in \mathbb{R}^k$, and $y_0 \in \mathbb{R}^k$ is given.

A numerical method for the integration of this system consists in constructing approximations y_n of the exact solution $y(x_n)$ at some chosen points $a = x_0 < x_1 < \cdots < x_n < \cdots < x_N = b$ in the interval of integration. There are two types of procedures for obtaining these approximations: one-step methods (where y_{n+1} is deduced only from y_n) and multistep methods (where y_{n+1} depends on $y_n, \ldots, y_{n-p}$).

First choose a fixed step size $h = (b - a)/N$ and set $x_n = a + nh$ for $n = 0, \ldots, N$. Of course, for $y_0, \ldots,$

y_N to be approximations of $y(x_0), \ldots, y(x_N)$, the formula for computing y_{n+1} has to satisfy some properties, namely the consistency, the stability, and the convergence which means that the error tends to zero when h tends to zero. Another quite important property of a method is its order r which explains the behavior of the error with h, $y_n - y(x_n) = \mathcal{O}(h^r)$.

In a one-step method, the approximate solution is computed by

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h), n = 0, \ldots, N-1$$

A Runge–Kutta method consists in

$$\begin{aligned} k_i &= f(x_n + \theta_i h, y_n + h\eta_i) \\ \eta_i &= a_{i1}k_1 + \cdots + a_{im}k_m \\ \Phi(x_n, y_n, h) &= c_1 k_1 + \cdots + c_m k_m \end{aligned}$$

If $a_{ij} = 0$ for $j \geqslant i$, the method is explicit (which means that each k_i can be computed separately) while, if not, it is implicit and the k_i are given as the solution of a system of (nonlinear) equations. Of course, an implicit method is much more difficult to implement but, in the case of a stiff differential equation (that is an equation whose solution varies rapidly), such an implicit method has to be used.

In a multistep method, $y_0, \ldots, y_{p-1}$ have first to be computed by a one-step method and then $y_p, \ldots, y_N$ are obtained by

$$a_0 y_n + \cdots + a_p y_{n+p} = h(b_0 f_n + \cdots + b_p f_{n+p})$$

for $n = 0, \ldots, N$–p, where $f_i = f(x_i, y_i)$. If $b_p = 0$ the method is explicit, otherwise it is implicit.

The main practical problem about the numerical integration of a differential equation concerns the precision of the approximate solution. If a desired precision has to be achieved, the first step is to choose the method. Then, h has to be chosen and adapted along the interval of integration in order to obtain the desired accuracy. This means, of course, that a procedure for estimating the error has to be implemented.

Integral Equations

Let g and K be two known functions. There are two kinds of integral equations depending whether the unknown function f appears only under the integral sign or also outside it. In these two kinds, two types arise depending on the interval of integration. If the interval of integration has fixed endpoints, a Fredholm equation is obtained.

$$\begin{aligned} \text{1st kind:} &\quad \int_a^b K(x,t)f(t)\,\mathrm{d}t = g(x) \\ \text{2nd kind:} &\quad f(x) - \lambda \int_a^b K(x,t)f(t)\,\mathrm{d}t = g(x) \end{aligned}$$

where $\lambda \neq 0$ is a parameter. If the right endpoint is a variable, a Volterra equation is obtained.

$$\begin{aligned} \text{1st kind:} &\quad \int_a^x K(x,t)f(t)\,\mathrm{d}t = g(x) \\ \text{2nd kind:} &\quad f(x) - \lambda \int_a^x K(x,t)f(t)\,\mathrm{d}t = g(x) \end{aligned}$$

Fredholm equations reduce to Volterra if the kernel K is such that $K(x, t) = 0$ for $x \leqslant t \leqslant b$.

To obtain approximations f_j of the function f at some points x_j, assume that the function g is known at some points x_i. So

$$g(x_i) = \int_a^b K(x_i, t)f(t)\,\mathrm{d}t, \quad i = 1, \ldots, N$$

and similarly for the other integral equations. The definite integral is evaluated by a quadrature formula

$$\int_a^b K(x_i, t)f(t)\,\mathrm{d}t \simeq \sum_{j=1}^{M} a_{ij} f(t_j)$$

Thus, the approximations $f_j \simeq f(t_j)$ is obtained by solving the system of linear equations

$$g(x_i) = \sum_{j=1}^{M} a_{ij} f_j, \quad i = 1, \ldots, N$$

If $N \neq M$, the system is solved in the least squares sense. Otherwise, a square system is obtained. However, it must be noticed that this system is usually ill-conditioned and thus preconditioning and/or regularization have to be used (see the next section). This is the case with the inversion of the Laplace transform, one of the most important integral equations.

Fredholm equations of the second kind can be solved by the method of successive approximations. If $\mathscr{K}$ denotes the integral operator which maps f into $\int_a^b K(x,t)f(t)$, then this integral equation is $(I - \lambda\mathscr{K})f = g$. So, f is given by its Neumann series

$$f = (I - \lambda\mathscr{K})^{-1} g = \left(I + \sum_{i=0}^{\infty} \lambda^i \mathscr{K}^i\right) g$$

In other terms, set $f_0 = g$ and then compute

$$f_i(x) = g(x) + \lambda \int_a^b K(x,t) f_{i-1}(t)\, \mathrm{d}t, \quad i = 1, 2, \ldots$$

This method converges if $\|\lambda \mathcal{K}\| < 1$.

Systems of Linear Equations

The choice of a method for solving a system of p linear equations in p unknowns, $Ax = b$, highly depends on the value of p. There are two classes of methods: direct methods and iterative methods.

When p is not too large (i.e., up to 10^4), an elimination procedure, such as Gaussian elimination, or Cholesky decomposition if the matrix is symmetric and positive definite, or Householder factorization, can be used. They all consist in transforming the system into an equivalent one (i.e., with the same solution) but with a matrix which is upper triangular. Then, such a system can be easily solved starting from the last equation.

When p is very large (several millions) and the matrix of the system comes out from the discretization, by finite differences or finite elements, of a partial differential equation, it is sparse (which means that most elements are zero). However, a direct method could not be used since the upper triangular matrix obtained by a direct method could be full. In that case, iterative methods are preferred. These methods split into two subclasses: relaxation methods and projection methods. Relaxation methods are nowadays mostly used as preconditioners (see below). Projection methods consist in generating a sequence of approximate solutions belonging to subspaces of increasing dimensions. Usually, these subspaces are Krylov subspaces and thus the name of such methods. Among them, the most popular are Lanczos method, the BiCGStab, and GMRES.

A fundamental topic of the solution of a system of linear equations is the condition number $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ of the matrix A, a number greater than 1. If this number is large, then a small perturbation of the matrix A and/or the right-hand side b can produce a large variation in the exact solution of the system. The system is said to be ill-conditioned. Also, the speed of convergence of some Krylov subspace methods depends on the condition number: the larger it is the slower is the convergence. Thus, it is important, when a matrix is ill-conditioned to precondition the system, that is to consider the new system $M^{-1}Ax = M^{-1}b$, where the matrix M is an approximation of A easy to invert (or, equivalently, a system with M as its matrix is easy to solve) and chosen so that $\kappa(M^{-1}A)$is smaller than $\kappa(A)$. There is no universal preconditioner and almost each system is a particular case.

Another procedure for solving an ill-conditioned system is to use regularization. It consists in solving a perturbed system. A vector x which minimizes

$$\|Ax - b\|_2 + \lambda \|Hx\|$$

is derived where λ is a parameter and H a matrix. The solution x_λ depends on $\lambda \geqslant 0$ (and also on H). When λ is close to 0, x_λ is close to $x = A^{-1}b$ but, since A is ill-conditioned, it is badly computed. If λ is large, x_λ will be correctly computed but it will not be a good approximation of x. So, there is a balance to find between large and small values of λ, and this is the main difficulty of regularization. It can be avoided by using an extrapolation technique.

See also: Semi-Empirical and Empirical Calculations.

PACS: 02.60. – x

Further Reading

Baker GA Jr. and Graves-Morris P (1996) *Padé Approximants*, 2nd edn. Cambridge: Cambridge University Press.

Björck Å (1996) *Numerical Methods for Least Squares Problems*. Philadelphia: SIAM.

Brezinski C (1997) *Projection Methods for Systems of Equations*. Amsterdam: North-Holland.

Brezinski C and van Iseghem J (1994) Padé approximations. In: Ciarlet PG and Lions JL (eds.) *Handbook of Numerical Analysis*, vol. III, pp. 47–222. Amsterdam: North Holland.

Butcher JC (1987) *The Numerical Analysis of Ordinary Differential Equations*. Chichester: Wiley.

Davis PJ (1975) *Interpolation and Approximation*. New York: Dover.

Kythe PK and Puri P (2002) *Computational Methods for Linear Integral Equations*. Boston: Birkhäuser.

Meurant G (1999) *Computer Solution of Large Linear System*. Amsterdam: North-Holland.

Saad Y (1996) *Iterative Methods for Sparse Linear Systems*. Boston: PWS Publishing Company.

O

Onsager Relations

C Cercignani, Politecnico di Milano, Milan, Italy

Introduction

The second law of thermodynamics indicates that not all the processes compatible with the first law can actually occur. Whereas one can easily perform work to heat up the system, it is not always enough to supply heat to increase the mechanical energy. At least two heat sources at different temperatures are needed as shown by Carnot's argument (sometimes one of the sources may be naturally supplied by the environment). Essentially, the second law states that heat can never pass from a colder to a warmer body without some other related change occurring at the same time. The modern statement of this principle of dissipation is based upon the notion of entropy, introduced by Clausius in 1865. Clausius showed that for every thermodynamic system there exists a function of the state of the system, its entropy denoted by S. The latter is defined by a differential relationship for the increased $\mathrm{d}S$ of the entropy of the system in an infinitesimally short time interval, during which a time "reversible" process occurs:

$$\mathrm{d}S = \frac{\mathrm{d}^{*}Q}{T} \qquad [1]$$

where $\mathrm{d}^{*}Q$ is the heat supplied to the system and T is the absolute temperature. (The notation $\mathrm{d}^{*}Q$ emphasizes that Q is not a state function and hence, in mathematical language, $\mathrm{d}^{*}Q$ is not an exact differential; "time-reversible" means, of course, that the process formed by following the states in a reverse time order is also permissible). For processes that are irreversible, one can only state that the increase of entropy in a process leading from an equilibrium state to another is larger than the increase that would occur in a reversible process between the same two states:

$$\mathrm{d}S \geqslant \frac{\mathrm{d}^{*}Q}{T} \qquad [2]$$

In order to transform this meager statement into a useful tool, it is necessary to build some structure for the evolution of the thermodynamic state of a system.

The irreversible processes are also called transport processes because they arise from reversible motions at a molecular level, which transfer a quantity (mass, momentum, energy) in a way that is not recognized as a convection at a macroscopic level. Thus, the microscopic momentum flow splits into a macroscopic momentum flow (or convection of momentum) and a viscous stress; the microscopic energy flow splits into a macroscopic energy flow (or convection of energy), the work of viscous stresses, and heat conduction.

In the following sections, the microscopic theory of transport coefficients and reciprocity relations based on the Boltzmann equation, in the simple but important case of a mixture of ideal gases (which applies, with a few modifications, to the case of electrons as well) is first discussed. This is the only accurate way of proving the symmetry properties of the transport coefficients. Then, the macroscopic theory of the transport coefficients and their symmetry relations, due to Onsager, and usually called Onsager relations, is dealt with.

The Boltzmann Equation

The next step was provided by Boltzmann, who studied perfect gases with the tools of kinetic theory. To explain the mechanical origin of irreversibility, Boltzmann considered what happens to the distribution of velocities of gas molecules when collisions occur. This led him to formulate a kinetic equation, subsequently called the Boltzmann equation, in which two-body collisions (like those between two billiard balls) play a leading role. Certain collisions (called direct collisions) cause decreases in the number of molecules with a certain velocity, while other collisions (called restoring collisions) increase that number. The occurrence of both direct and restoring collisions corresponds to the inherent reversibility of molecular events. Not only did Boltzmann show that the equation bearing his name admits Maxwell's distribution as an equilibrium solution, he also gave a heuristic proof that it is the only possible

one. To this end he introduced a quantity, which he denoted by E and was later (as here) denoted by H, defined in terms of the molecular velocity distribution. He then demonstrated that, as a consequence of his equation, this function must always decrease in an isolated system or, at most, remain constant, the latter case occurring only if a state of statistical equilibrium prevails. Thus, the Boltzmann equation is not time reversible. Boltzmann's result is usually quoted as the "H-theorem" and indicates that H must be proportional to minus the entropy. H has an interesting interpretation as a measure of order in the molecular distribution; as a consequence, S is a measure of randomness or chaos.

Despite its successes, the Boltzmann equation involves conceptual difficulties. Because it is time irreversible, it violates the recurrence theorem of mechanics. This theorem says that the particles composing a system of finite energy and size will return, at some future time, to very nearly their initial condition. Boltzmann proposed the correct way out of this paradox: his equation should be interpreted as describing what happens to most, not all, of the initial data. To illustrate this, two former students of Boltzmann, Paul and Tatyana Ehrenfest, introduced a picturesque model, colloquially referred to as the dog-flea model. Consider two dogs lying next to one another, with a total of N fleas shared between them. If the fleas jump only from one dog to the next, then after a time, the number on each dog will have changed while the total number of fleas will be the same. If the dogs are identical, after a long period of time, it is expected that each dog will have on the average $N/2$ fleas. This will be true even if all the fleas originally resided on only one dog. However, if one waits a much longer time, all the fleas will be back on the original dog. Something similar holds for molecular dynamics; the recurrence time is, however, so long (extremely longer than the estimated age of the universe) that phenomena of this kind, which would violate the second law, are never witnessed. The advantage of the Boltzmann equation is that its rigorous derivation can only be obtained by a limiting procedure, according to which these paradoxical events disappear because they occur at a time $\bar{t}$ such that $\bar{t} \to \infty$, when the Boltzmann equation holds.

Moreover, it is possible to derive the laws of fluid flow, including the linear phenomenological equations, from the Boltzmann equation and to obtain explicit expressions for the heat conductivity and the viscosity of gases that agree with experimental measurements.

The quantity H was essentially defined by Boltzmann as the integral of $f \log f$ with respect to the molecular velocity, where log denotes the natural logarithm and f is the one-particle distribution function. The latter, in the case of a monatomic gas, is a function of time t, position $\boldsymbol{x}$, and molecular velocity $\boldsymbol{v}$ and gives the expected number density of molecules in the six-dimensional phase space described by position $\boldsymbol{x}$ and velocity $\boldsymbol{v}$. The distribution function f satisfies the Boltzmann equation that can be written in the following form:

$$\begin{aligned}&\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{x}} \\ &= \int_{\Re^3} \int_{S^2} [f(\boldsymbol{x}, \boldsymbol{v}', t) f(\boldsymbol{x}, \boldsymbol{v}'_*, t) - f(\boldsymbol{x}, \boldsymbol{v}, t) \\ &\quad \times f(\boldsymbol{x}, \boldsymbol{v}_*, t)] |\boldsymbol{V}| \sigma_{\mathrm{d}}(|\boldsymbol{V}|, \boldsymbol{V} \cdot \boldsymbol{n}) \mathrm{d}\boldsymbol{v}_* \, \mathrm{d}\boldsymbol{n} \end{aligned} \quad [3]$$

where $\boldsymbol{V} = \boldsymbol{v}_* - \boldsymbol{v}$ is the relative velocity and σ_{d} the differential cross section. The velocities before a collision, $\boldsymbol{v}'$ and $\boldsymbol{v}'_*$, are related to those after the same collision, $\boldsymbol{v}$ and $\boldsymbol{v}_*$, by

$$\boldsymbol{v}' = \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_* + |\boldsymbol{V}|\boldsymbol{n}), \quad \boldsymbol{v}'_* = \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_* - |\boldsymbol{V}|\boldsymbol{n}) \quad [4]$$

where $\boldsymbol{n}$ is a unit vector ranging over the unit sphere S^2, directed as the relative velocity after the collision.

For polyatomic gases, f depends on further variables describing the internal degrees of freedom of the molecule and the Boltzmann equation becomes more complicated.

By a detailed consideration of the properties of his equation, Boltzmann showed that the time derivative of H is never positive and vanishes if and only if the velocity distribution is a Maxwellian, provided the system cannot exchange heat with the surroundings. The proof of this result requires that the space derivatives give zero contribution (such is the case for a gas in a spatially homogeneous state or inside a box with specularly reflecting walls). It is clear that H corresponds to the entropy except for a negative constant factor. The extension of this result to non-isolated systems was only provided in the second half of the twentieth century and permits one to deal with heat sources.

The molecular description is complicated since it must cover situations varying from ordinary temperature and pressures to nearly vacuum conditions. The case when these conditions occur was studied already by Maxwell and Boltzmann and was systematized by Chapman and Enskog in the past century. The final result is better stated if one refers to the entropy per unit volume s:

$$\frac{\partial s}{\partial t} = \sigma = \sigma^{(r)} + \sigma^{(i)} \quad [5]$$

where σ is the entropy source per unit volume, decomposed into a reversible part

$$\sigma^{(r)} = -\nabla \cdot \left(s\boldsymbol{u} + \frac{\boldsymbol{q}}{T} \right) \quad [6]$$

and an irreversible one $\sigma^{(i)}$. Here, $\boldsymbol{q}$ is the heat flow vector and $\boldsymbol{u}$ the (bulk) velocity vector; $\sigma^{(i)}$ in turn must be decomposed into two parts: the entropy transfer associated with the irreversible processes and a volume source that is non-negative. This decomposition is discussed below.

According to the Chapman–Enskog solution of the Boltzmann equation for a simple gas (as opposed to a mixture) and in agreement with continuum mechanics, σ can be expressed as a sum of two terms, which can be written as a linear combination of $\sum_{ij} q_{ij}\eta_{ij}$ and $\sum_i q_i g_i$; here q_{ij}, η_{ij}, q_i, and g_i denote the components of the stress deviator, the rate of strain tensor, the heat flow vector, and the temperature gradient, respectively. The same theory shows that in each of the above pairs of tensors and vectors, the first element depends linearly on the second with coefficients which may only depend on temperature and number density. This result must be postulated in continuum mechanics. The extension of the theory to mixtures (see next section for details) shows that the pattern repeats itself: there are more gradients (the concentration densities also occur) and more quantities (such as the diffusion velocity) may depend on these quantities, but the entropy source is always a linear combination of terms which are products of two factors: a gradient and another factor describing a phenomenon typically associated with that gradient. New phenomena, such as thermal diffusion, were discovered by this kinetic theory approach, that has, however, the serious limitation of being essentially restricted to rarefie(l)d gases. During World War II, thermal diffusion, an example of a so-called cross-coupling in which a temperature gradient causes a diffusion flux, was used to separate fissionable isotopes of uranium.

Electric phenomena can also be studied by means of a Boltzmann equation describing the statistical behavior of electrons. Results qualitatively similar to those found for a gas can be proved in this case as well. In the next two sections, the Chapman–Enskog theory for the general case of a mixture is briefly described.

Transport Coefficients For Gaseous Mixtures

The transport coefficients of gas mixtures are derived below by using the Chapman–Enskog method. It is assumed that there are n gaseous species.

First of all, one must write the Boltzmann equation for a mixture:

$$\frac{\partial f_a}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f_a}{\partial \boldsymbol{x}} = \sum_{b=1}^{n} \int (f'_a f'_b - f_a f_b) |V_{ba}| \sigma_{ab} \, \mathrm{d}\boldsymbol{n}_{ba} \, \mathrm{d}^3\boldsymbol{v}_b \quad [7]$$

where f_a is the one-particle distribution function for the ath species.

The one-particle distribution function of constituent a is written as

$$f_a = f_a^{(0)}(1 + h_a) \quad [8]$$

where $f_a^{(0)} h_a$ represents the deviation from the Maxwell distribution function:

$$f_a^{(0)} = n_a (2\pi RT)^{-3/2} \exp(-|\boldsymbol{v} - \boldsymbol{u}|^2 / 2RT) \quad [9]$$

where $\boldsymbol{u}$ is the bulk velocity with components u_i and R the gas constant.

Then, it is meaningful to assume that the local Maxwellian contains the entire information about the densities of all species, the bulk velocity, and the temperature. As a consequence, one obtains the following constraints for the deviation $f_a^{(0)} h_a$ from the Maxwell distribution function:

$$\int f_a^{(0)} h_a \, \mathrm{d}^3\boldsymbol{v} = 0 \quad [10]$$

$$\sum_{a=1}^{n} m_a \int \boldsymbol{v} f_a^{(0)} h_a \, \mathrm{d}^3\boldsymbol{v} = 0 \quad [11]$$

$$\sum_{a=1}^{n} m_a \int |\boldsymbol{v}|^2 f_a^{(0)} h_a \, \mathrm{d}^3\boldsymbol{v} = 0 \quad [12]$$

The Chapman–Enskog method can be presented in many ways; since most treatments restrict the computations to first-order perturbations, the simplest possible presentation is adopted. h_a is computed as a perturbation produced by the fact that the derivatives of $f_a^{(0)}$ with respect to time and space variables are nonzero, when these derivatives produce a small change on the scale of the mean free path and mean free time. Hence, if one inserts the distribution function [8] into the Boltzmann equation, one keeps the zeroth order terms on the left-hand size and the first-order terms on the right-hand side, thus obtaining

$$\partial_t f_a^{(0)} + \boldsymbol{v} \cdot \partial_{\boldsymbol{x}} f_a^{(0)} = \sum_{b=1}^{n} \int f_a^{(0)} f_b^{(0)} (h'_a + h'_b - h_a - h_b) |V_{ba}| \sigma_{ab} \times \mathrm{d}\boldsymbol{n}_{ba} \, \mathrm{d}^3\boldsymbol{v}_b \quad [13]$$

The left-hand side of [13] is transformed by inserting the Maxwell distribution function into the streaming term $\partial_t f_a + \boldsymbol{v}\cdot\partial_{\boldsymbol{x}} f_a$, which is written as

$$\frac{\partial f_a}{\partial t} + \boldsymbol{v}\cdot\frac{\partial f_a^{(0)}}{\partial \boldsymbol{x}} = f_a^{(0)}\left\{\boldsymbol{v}\cdot\partial_{\boldsymbol{x}}\left(\frac{\mu_a}{RT}\right) + \partial_t\left(\frac{\mu_a}{RT}\right) + \frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{2RT^2}\boldsymbol{v}\cdot\partial_{\boldsymbol{x}}T + \frac{(\boldsymbol{v}-\boldsymbol{u})}{RT}\cdot(\boldsymbol{v}\cdot\partial_{\boldsymbol{x}})\boldsymbol{u} + \frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{2RT^2}\partial_t T + \frac{(\boldsymbol{v}-\boldsymbol{u})}{RT}\cdot\partial_t\boldsymbol{u}\right\} \quad [14]$$

where, for convenience, the so-called chemical potential $\mu_a = RT\log n_a$ has been introduced. From [14] it follows that [13] reduces to

$$f_a^{(0)}\left\{\boldsymbol{v}\cdot\partial_{\boldsymbol{x}}\left(\frac{\mu_a}{RT}\right) + \partial_t\left(\frac{\mu_a}{RT}\right) + \frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{2RT^2}\boldsymbol{v}\cdot\partial_{\boldsymbol{x}}T + \frac{(\boldsymbol{v}-\boldsymbol{u})}{RT}\cdot[\boldsymbol{v}\cdot\partial_{\boldsymbol{x}}\boldsymbol{u}] + \frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{2RT^2}\partial_t T + \frac{(\boldsymbol{v}-\boldsymbol{u})}{RT}\cdot\partial_t\boldsymbol{u}\right\} = \sum_{b=1}^{n}\int f_a^{(0)}f_b^{(0)}(h'_a + h'_b - h_a - h_b)|V_{ba}|\sigma_{ab}\times \mathrm{d}\boldsymbol{n}_{ba}\,\mathrm{d}^3\boldsymbol{v}_b \quad [15]$$

It can be seen that the balance equations associated with eqn [13] (or [15]) are correct only to the zeroth order because the first-order terms on the left-hand side have been neglected. Thus, they do not contain the irreversible transport phenomena: they describe a mixture of Eulerian fluids (no viscosity, no heat conduction) with vanishing diffusion fluxes. In spite of that, they can be used correctly on the left-hand side, because this is correct to the zeroth order only. The procedure iterates on this to obtain the first terms which describe transport phenomena.

If one integrates eqn [15] over $\boldsymbol{v}$, one obtains the balance equation for the particle number density of constituent a that reads

$$\partial_t n_a + \partial_{\boldsymbol{x}}\cdot(n_a\boldsymbol{u}) = 0 \quad [16]$$

It has already been remarked that in a Eulerian fluid mixture, the pressure deviator and the heat flux of the mixture vanish. Hence, the balance equations for the momentum density and energy density of the mixture reduce to the balance equations for copies of a single Eulerian fluid that read

$$nc_v(\partial_t T + \boldsymbol{u}\cdot\partial_{\boldsymbol{x}}T) + p\partial_{\boldsymbol{x}}\cdot\boldsymbol{u} = 0 \quad [17]$$

$$nm(\partial_t\boldsymbol{u} + \boldsymbol{u}\cdot\partial_{\boldsymbol{x}}\boldsymbol{u}) = -\partial_{\boldsymbol{x}}p \quad [18]$$

where c_v is the heat capacity per particle at constant volume of the mixture.

Now the time derivatives $\partial_t(\mu_a/T)$, $\partial_t T$, and $\partial_t\boldsymbol{u}$ are eliminated from [15] by using eqns [16]–[18] and the following integral equation is obtained for h_a:

$$f_a^{(0)}\left\{-\boldsymbol{v}\cdot\sum_{b=1}^{n}(c_b-\delta_{ab})\partial_{\boldsymbol{x}}\left(\frac{\mu_b-\mu_a}{RT}\right) + \left(\frac{|\boldsymbol{v}|^2}{2RT} - \frac{5}{2}\right)\times\frac{\boldsymbol{v}}{RT^2}\cdot\partial_{\boldsymbol{x}}T + 2(\boldsymbol{v}-\boldsymbol{u})\cdot[(\boldsymbol{v}-\boldsymbol{u})\cdot\partial_{\boldsymbol{x}}]\boldsymbol{u} - \frac{2}{3}|\boldsymbol{v}-\boldsymbol{u}|^2\partial_{\boldsymbol{x}}\cdot\boldsymbol{u}\right\} = \sum_{b=1}^{n}\int f_a^{(0)}f_b^{(0)}(h'_a + h'_b - h_a - h_b)\times|V_{ba}|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}_{ba}\,\mathrm{d}^3\boldsymbol{v}_b \quad [19]$$

where the particle fraction $c_a = n_a/n$ of constituent a is introduced.

The integral equation [19] has a solution of the form

$$h_a = A_a\frac{\boldsymbol{v}}{RT^2}\cdot\partial_{\boldsymbol{x}}T + \frac{\boldsymbol{v}}{k}\cdot\sum_{b=1}^{n}B_{ab}(c_b-\delta_{ab})\partial_{\boldsymbol{x}}\left(\frac{\mu_b-\mu_a}{RT}\right) + 2C_a(\boldsymbol{v}-\boldsymbol{u})\cdot[(\boldsymbol{v}-\boldsymbol{u})\cdot\partial_{\boldsymbol{x}}]\boldsymbol{u} - \tfrac{2}{3}C_a|\boldsymbol{v}-\boldsymbol{u}|^2\partial_{\boldsymbol{x}}\cdot\boldsymbol{u} \quad [20]$$

where A_a, B_{ab}, and C_a are determined from the integral equations that are obtained by inserting [20] into eqn [19], viz,

$$-\frac{\boldsymbol{v}}{R}(c_d-\delta_{ad}) = \sum_{b=1}^{n}\int f_b^{(0)}[B'_{ad}\boldsymbol{v}' + B'_{bd}\boldsymbol{v}'_b - B_{ad}\boldsymbol{v} - B_{bd}\boldsymbol{v}_b]\times|V_{ba}|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}_{ba}\,\mathrm{d}^3\boldsymbol{v}_b \quad [21]$$

$$\frac{\boldsymbol{v}}{RT^2} = \sum_{b=1}^{n}\int f_b^{(0)}[A'_a\boldsymbol{v}' + A'_b\boldsymbol{v}'_b - A_a\boldsymbol{v} - A_b\boldsymbol{v}_b]\times|V_{ba}|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}_{ba}\,\mathrm{d}^3\boldsymbol{v}_b \quad [22]$$

$$\frac{\boldsymbol{v}\boldsymbol{v}}{RT} = \sum_{b=1}^{n}\int f_b^{(0)}[C'_a\boldsymbol{v}'\boldsymbol{v}' + C'_b\boldsymbol{v}'_b\boldsymbol{v}'_b - C_a\boldsymbol{v}\boldsymbol{v} - C_b\boldsymbol{v}'_b\boldsymbol{v}_b]|V_{ba}|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}_{ba}\,\mathrm{d}^3\boldsymbol{v}_b \quad [23]$$

Since one is interested only in general expressions for the transport coefficients, one need not solve the integral eqns [21]–[23]. The transport coefficients are expressed in terms of the functions A_a, B_{ab}, C_a, which can be obtained by solving the aforementioned integral equations.

It is remarked that the balance equations containing the irreversible transport phenomena, correct to first order, will occur as a compatibility condition for the equations at the next order. This is because of the iteration nature of the Chapman–Enskog procedure.

One can calculate the constitutive terms of a mixture of gases in the second Chapman–Enskog approximation. The constitutive terms are: the pressure deviator of the mixture q_{ij}, the heat flux of the mixture q_j, and the diffusion flux $\boldsymbol{J}_a$ of the constituent a.

To begin with, the constitutive equation is determined for the pressure deviator of the mixture, which can be obtained through the insertion of [8] together with [20] into its definition

$$q_{ij} = \sum_{a=1}^{n} m_a \int [(v_i - u_i)(v_j - u_j) - \tfrac{1}{3}|\boldsymbol{v} - \boldsymbol{u}|^2 \delta_{ij}] f_a \, \mathrm{d}^3\boldsymbol{v} \qquad [24]$$

yielding

$$q_{ij} = -\mu\left[\partial_{x_i} u_j + \partial_{x_j} u_i - \tfrac{2}{3}(\partial_{\boldsymbol{x}} \cdot \boldsymbol{u})\delta_{ij}\right] \qquad [25]$$

where the expression for the coefficient of shear viscosity μ is

$$\mu = \frac{1}{15}\sum_{a=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^4 f_a^{(0)} C_a \, \mathrm{d}^3\boldsymbol{v}_A \qquad [26]$$

Further, the heat flux of the mixture, defined by

$$q_j = \sum_{a=1}^{n} \frac{m_a}{2} \int (v_j - u_j)|\boldsymbol{v} - \boldsymbol{u}|^2 f_a \, \mathrm{d}^3\boldsymbol{v} \qquad [27]$$

can be shown to be

$$q_j = -\lambda \partial_{x_j} T \qquad [28]$$

The expressions of thermal conductivity λ reads

$$\lambda = \frac{1}{3}\sum_{a=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^2 f_a^{(0)} A_a \, \mathrm{d}^3\boldsymbol{v} \qquad [29]$$

The last constitutive equation to be determined is the one for the partial diffusion flux $\boldsymbol{J}_a$ with components J_{ai}, which is obtained by inserting [8] together with [20] into the definition of the partial flow:

$$J_{ai} = L_a \partial_{x_i} T + \sum_{b=1}^{n} L_{ab} \partial_{x_i}\left(\frac{\mu_b - \mu_a}{T}\right) \qquad [30]$$

In [30], the coefficients L_a and L_{ab} are related to the thermal-diffusion cross-effect and to the matrix of the diffusion coefficients, respectively. Their expressions are given by

$$L_a = \frac{1}{3}\int |\boldsymbol{v} - \boldsymbol{u}|^2 f_a^{(0)} A_a \, \mathrm{d}^3\boldsymbol{v} \qquad [31]$$

$$L_{ab} = \frac{1}{3}\int |\boldsymbol{v} - \boldsymbol{u}|^2 f_a^{(0)} B_{ab} \, \mathrm{d}^3\boldsymbol{v} \qquad [32]$$

Onsager's Reciprocity Relations from the Boltzmann Equation

In this section, one can analyze the relationships between the transport coefficients of a chemically reacting mixture of gases which are known as the Onsager reciprocity relations. The starting point is provided by the formulas derived in the previous section from the Boltzmann equation.

In order to determine the Onsager reciprocity relations, one needs the following relationship which is valid for arbitrary functions h and ψ:

$$\sum_{a,b=1}^{n} \int f_a^{(0)} f_b^{(0)} \psi_b (h'_a + h'_a - h_a - h_a) \times |V|\sigma_{ab} \, \mathrm{d}\boldsymbol{n} \, \mathrm{d}^3\boldsymbol{v}_a \, \mathrm{d}^3\boldsymbol{v}_b = \sum_{a,b=1}^{n} \int f_a^{(0)} f_b^{(0)} h_b (\psi'_a + \psi'_a - \psi_a - \psi_a) \times |V|\sigma_{ab} \, \mathrm{d}\boldsymbol{n} \, \mathrm{d}^3\boldsymbol{v}_a \, \mathrm{d}^3\boldsymbol{v}_b \qquad [33]$$

The above equation can be proved by using the symmetry properties of the elastic collision term and by interchanging the dummy indices a and b in the sums.

For the determination of the first Onsager reciprocity relation, it is first noted that due to the constraint $\sum_{a=1}^{n} \boldsymbol{J}_{ai} = 0$, one must have according to [30]–[32] that

$$\sum_{a=1}^{n} L_a = \frac{1}{3}\sum_{a=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^2 f_a^{(0)} A_a \mathrm{d}^3\boldsymbol{v} = 0 \qquad [34]$$

$$\sum_{a=1}^{n} L_{ab} = \frac{1}{3}\sum_{a=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^2 f_a^{(0)} B_{ab} \mathrm{d}^3\boldsymbol{v} = 0 \qquad [35]$$

Based on the above constraints, the two coefficients L_a and L_{ab} can be written as

$$L_a = -\frac{1}{3}\sum_{b=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^2 f^{(0)} A_a (c_b - \delta_{ab}) \, \mathrm{d}^3\boldsymbol{v} \qquad [36]$$

$$L_{ab} = -\frac{1}{3}\sum_{d=1}^{n} \int |\boldsymbol{v} - \boldsymbol{u}|^2 f_d^{(0)} B_{db} (c_d - \delta_{ad}) \, \mathrm{d}^3\boldsymbol{v} \qquad [37]$$

By multiplying [19] with $R(\boldsymbol{v}_a - \boldsymbol{u})f_a^{(0)}B_{ac}/3$, integrating the resulting equation over all values of $\mathrm{d}^3\boldsymbol{v}_a$, and after that summing over all constituents yields

$$
\begin{aligned}
&-\frac{1}{3}\sum_{a=1}^{n}\int |\boldsymbol{v}_a - \boldsymbol{u}|^2 f_a^{(0)} B_{ac}(c_d - \delta_{ad})\mathrm{d}^3\boldsymbol{v}_a = L_{dc} \\
&= \frac{R}{6}\sum_{a,b=1}^{n}\int f_a^{(0)} f_b^{(0)}(\boldsymbol{v}_a - \boldsymbol{u})B_{ac}\cdot[B'_{ad}(\boldsymbol{v}'_a - \boldsymbol{u}) \\
&\quad + B'_{bd}(\boldsymbol{v}'_b - \boldsymbol{u}) - B_{ad}(\boldsymbol{v}_a - \boldsymbol{u}) \\
&\quad - B_{bd}(\boldsymbol{v}_b - \boldsymbol{u})]|V|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}^3\boldsymbol{v}_a\,\mathrm{d}^3\boldsymbol{v}_b \\
&= \frac{R}{6}\sum_{a,b=1}^{n}\int f_a^{(0)} f_b^{(0)}(\boldsymbol{v}_a - \boldsymbol{u})B_{ad}\cdot[B'_{ac}(\boldsymbol{v}_b - \boldsymbol{u}) \\
&\quad + B'_{bc}(\boldsymbol{v}'_b - \boldsymbol{u}) - B_{ac}(\boldsymbol{v}_a - \boldsymbol{u}) - B_{bc}(\boldsymbol{v}_b - \boldsymbol{u})] \\
&\quad \times |V|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}^3\boldsymbol{v}_a\,\mathrm{d}^3\boldsymbol{v}_b = L_{cd}
\end{aligned}
\qquad [38]
$$

Hence, the first Onsager reciprocity relation which shows that the matrix of the diffusion coefficients is symmetric (i.e., $L_{cd} = L_{dc}$) has been proved.

The expression for the coefficient of the thermal-diffusion effect L_d^* is obtained through the multiplication of [20] by $R(\boldsymbol{v}_a - \boldsymbol{u})f_a^{(0)}A_a/3$, integration of the resulting equation over all values of $\boldsymbol{v}_a$, and subsequent summing over all constituents:

$$
\begin{aligned}
&-\frac{1}{3}\sum_{a=1}^{n}\int |\boldsymbol{v}_a - \boldsymbol{u}|^2 (c_d - \delta_{ad}) f_a^{(0)} A_a \mathrm{d}^3\boldsymbol{v}_a = L_d^* \\
&= \frac{R}{6}\sum_{a,b=1}^{n}\int f_a^{(0)} f_b^{(0)}(\boldsymbol{v}_a - \boldsymbol{u})A_a\cdot[B'_{ab}(\boldsymbol{v}'_a - \boldsymbol{u}) \\
&\quad + B'_{ba}(\boldsymbol{v}'_b - \boldsymbol{u}) - B_{ab}(\boldsymbol{v}_a - \boldsymbol{u})] - B_{ba}(\boldsymbol{v}_b - \boldsymbol{u})] \\
&\quad \times |V|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}^3\boldsymbol{v}_a\,\mathrm{d}^3\boldsymbol{v}_b
\end{aligned}
\qquad [39]
$$

In the same way, it follows the expression for the coefficient of the diffusion-thermal effect L_d^*, that is, multiplying [23] by $R(\boldsymbol{v}_a - \boldsymbol{u})f_a^{(0)}B_{da}/3$, integrating the resulting equation over all values of $\mathrm{d}^3\boldsymbol{v}_a$ and thereafter summing over all constituents:

$$
\begin{aligned}
&\frac{1}{3T^2}\sum_{a=1}^{n}\int |\boldsymbol{v}_a - \boldsymbol{u}|^2\left(\frac{|\boldsymbol{v}_a - \boldsymbol{u}|^2}{2RT} - \frac{5}{2}\right) f_a^{(0)} B_{da}\mathrm{d}^3\boldsymbol{v}_a = L_d^* \\
&= \frac{R}{6}\sum_{a,b=1}^{n}\int f_a^{(0)} f_b^{(0)}(\boldsymbol{v}_a - \boldsymbol{u})\cdot B_{da}[A'_a(\boldsymbol{v}'_a - \boldsymbol{u}) \\
&\quad + A'_b(\boldsymbol{v}'_b - \boldsymbol{u}) - A_a(\boldsymbol{v}_a - \boldsymbol{u}) - A_b(\boldsymbol{v}_b - \boldsymbol{u})] \\
&\quad \times |V|\sigma_{ab}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}^3\boldsymbol{v}_a\,\mathrm{d}^3\boldsymbol{v}_b
\end{aligned}
\qquad [40]
$$

The first equality above can be written, thanks to the constraint [12] and to the definition of the thermal-diffusion coefficient [30].

One can compare now the expressions [39] and [40], and by using the property [33], one can infer the last Onsager reciprocity relation that relates the coefficients of the thermal-diffusion effect and of the diffusion thermal effect $L_d^* = L_d$.

The results obtained from the Boltzmann equation in this section suggests that one looks for an explanation of the general pattern, having in mind media more general than rarefie(l)d gases.

The Macroscopic Theory of Irreversible Processes

Recall that the entropy is a function of some local parameters x_k (number density, internal energy, etc.). As a consequence,

$$
\mathrm{d}s = \sum_k F_k\,\mathrm{d}x_k \qquad [41]
$$

where $F_k = \partial s/\partial x_k$ are the "thermodynamic forces," functions of the parameters x_k. The local parameters must be chosen suitably: they must satisfy some balance equations of the following form:

$$
\frac{\partial x_k}{\partial t} = \xi_k = \xi_k^{(\mathrm{r})} + \xi_k^{(\mathrm{i})} \qquad [42]
$$

where ξ_k is the source per unit volume, decomposed into a reversible part $\xi_k^{(\mathrm{r})}$ and an irreversible one $\xi_k^{(\mathrm{i})}$. The irreversible sources are due to transport phenomena and, hence, have the general expression:

$$
\xi_k^{(\mathrm{i})} = -\nabla\cdot\boldsymbol{J}_k^{(\mathrm{i})} \qquad [43]
$$

where $\boldsymbol{J}_k^{(\mathrm{i})}$ is the irreversible flux of x_k. The time derivative of the entropy is given by

$$
\begin{aligned}
\frac{\partial s}{\partial t} &= \sum_k F_k\frac{\partial x_k}{\partial t} = \sum_k F_k\xi_k \\
&= \sum_k F_k\xi_k^{(\mathrm{r})} - \sum_k F_k\nabla\cdot\boldsymbol{J}_k^{(\mathrm{i})}
\end{aligned}
\qquad [44]
$$

The part of the entropy source associated with irreversible processes is thus given by

$$
\begin{aligned}
\sigma^{(\mathrm{i})} &= -\sum_k F_k\nabla\cdot\boldsymbol{J}_k^{(\mathrm{i})} = -\sum_k \nabla\cdot(F_k\boldsymbol{J}_k^{(\mathrm{i})}) \\
&\quad + \sum_k \boldsymbol{J}_k^{(\mathrm{i})}\cdot\nabla F_k
\end{aligned}
\qquad [45]
$$

and decomposes naturally into a transport term related to the entropy flux $\sum_k F_k\boldsymbol{J}_k^{(\mathrm{i})}$ associated with irreversible processes, and a dissipation term

$$
\sigma^{(\mathrm{d})} = \sum_k \boldsymbol{J}_k^{(\mathrm{i})}\cdot\nabla F_k \qquad [46]
$$

which provides the volume source of entropy, that is frequently referred to as the entropy source, without any further qualification. The factor multiplying a flux $\boldsymbol{J}_k^{(\mathrm{i})}$, ∇F_k, is called a (generalized) force or affinity. The source $\sigma^{(\mathrm{d})}$ is non-negative for the second law to be satisfied; in fact, the transport term, being in the form of a divergence, transforms, upon integration, into a surface integral and is thus related to processes occurring on the boundary.

The fluxes must depend on the forces; otherwise one could reverse the fluxes, keeping the forces constant, and change the sign of $\sigma^{(\mathrm{d})}$. The dependence can be very general but in most cases it is Markovian, that is, the fluxes at a time-instant depend on the forces at the same time only. Further, in many cases, the relation is linear:

$$\boldsymbol{J}_k^{(\mathrm{i})} = \sum_j L_{jk} \nabla F_j \qquad [47]$$

This circumstance holds in many cases; the processes described by these laws can still be nonlinear because the coefficients L_{jk} (called kinetic coefficients) are functions of the thermodynamic forces F_l. In addition to the kinetic and transport equations that were established experimentally in the nineteenth century, the linear laws suggested new types of transport phenomena. Indeed, the chemist Lars Onsager and later the physicist Josef Meixner were led to a connection between the phenomenological kinetic coefficients and thermodynamics. Because the entropy source is positive, the kinetic coefficients are not arbitrary. In fact, the expression of the source

$$\sigma^{(\mathrm{d})} = \sum_{jk} L_{jk} \nabla F_j \cdot \nabla F_k \qquad [48]$$

indicates that all the diagonal kinetic coefficients, L_{jj}, must be positive. Furthermore, the off-diagonal kinetic coefficients $L_{jk}(j \neq k)$ must satisfy the inequalities that arise by the fact that the kinetic coefficients form a positive definite matrix. Thus, the second law constrains the size of the off-diagonal kinetic coefficients.

The Reciprocity Relations

In the case of linear relations, the coefficients L_{jk} are assumed to satisfy a symmetry relation

$$L_{jk}(\boldsymbol{B}) = \varepsilon_j \varepsilon_k L_{kj}(-\boldsymbol{B}) \qquad [49]$$

where $\boldsymbol{B}$ is a magnetic field that can act on the system. The factor ε_j is defined as follows: it equals 1 if the parameter F_j does not change when time is reversed, and -1 if the parameter F_j changes sign when time is reversed. Temperature and concentrations belong to the first class, velocity (or momentum) to the second. The extension to intermediate cases is not difficult, but does not occur in practice.

The coupling of a flux and a force corresponding to unequal values of j and k, occurring when the corresponding off-diagonal kinetic coefficient does not vanish, suggests the kind of irreversible phenomena that might occur: thus, a temperature difference might give rise not only to a heat flux but also to an electric current, while a potential difference can give rise to an electric current and a heat flux. Experimental measurements demonstrate not only that these phenomena, known as the Seebeck and Peltier effects, exist but that the two off-diagonal coupling coefficients are equal within experimental error.

The above symmetry conditions were first established by Onsager and further discussed by Casimir. They are called the Onsager–Casimir reciprocity relations. It must be stressed that the proof of these relations does not belong to standard thermodynamics, but requires a study of the microscopic theory and a number of assumptions. A careful analysis shows that the alleged proof can, at most, be considered as a plausibility argument.

Another condition on the off-diagonal kinetic coefficients, frequently called the Curie principle, dictates restrictions related to space symmetries. For example, it implies that the thermodynamic force for a chemical reaction in a fluid, which has no directional character, doe not contribute to the heat flux, which is a vector. Experiments to test the validity of the reciprocal relations are not always easy. Some of the most accurate deal with mass diffusion at uniform temperature. For example, in a solution of table salt (NaCl) in water, there is only a single thermodynamic force for diffusion (the gradient of the chemical potential of NaCl), but in an aqueous solution of NaCl and potassium chloride (KCl), gradients in the chemical potentials of both salts act as thermodynamic forces for diffusion. Within the limits of experimental uncertainty, the off-diagonal kinetic coefficients for diffusion in this and other three-component solutions have been shown to be equal. The magnitude of the off-diagonal kinetic coefficients is frequently comparable to the diagonal ones, which has made phenomena related to off-diagonal kinetic coefficients a significant tool in electrochemistry, membrane biology, and geophysics.

It would be useful to have a better theoretical justification of the Onsager–Casimir relations and a better discussion of their experimental validity, but these do not seem to be available. One observes that the modern thermomechanics of continua dispenses with the Onsager–Casimir relations and discusses

irreversible phenomena with the help of the Clausius–Duhem inequality

$$\frac{\mathrm{d}S}{\mathrm{d}t} \geqslant \int_{\partial\Omega} \frac{q}{T}\,\mathrm{d}A \qquad [50]$$

assumed to hold for an arbitrary region Ω, with boundary $\partial\Omega$, occupied by a continuous medium. The constitutive relations that relate the various state functions and the "fluxes" to the "forces" are assumed to obey the postulate of equipresence: each quantity which is not taken as basic depends upon all the basic ones, unless this is forbidden by the second principle. The basic quantities are usually the number (or mass) density, the velocity, the thermal energy, and the concentrations. The class of materials (elastic, viscous, viscoelastic, etc.) are defined by the type of constitutive relations that are admitted.

See also: Irreversible Thermodynamics and Basic Transport Theory in Solids.

PACS: 05.20.Dd; 51.10.+y; 51.20.+d

Further Reading

Cercignani C (1988) *The Boltzmann Equation and Its Applications*. New York: Springer.
Cercignani C (1990) *Mathematical Methods in Kinetic Theory*, 2nd edn. New York: Plenum.
Cercignani C (2000) *Rarefied Gas Dynamics. From Basic Concepts to Actual Calculations*. Cambridge: Cambridge University Press.
Chapman S and Cowling TG (1940) *The Mathematical Theory of Non-Uniform Gases*. London: Cambridge University Press.
De Groot SR and Mazur P (1962) *Non-Equilibrium Thermodynamics*. Amsterdam: North-Holland.
Ferziger JH and Kaper HG (1972) *Mathematical Theory of Transport Processes in Gases*. Amsterdam: North-Holland.
Hirschfelder JO, Curtiss CF, and Bird RB (1954) *Molecular Theory of Gases and Liquids*. New York: Wiley.
Onsager L (1931) *Physical Review* 37: 405; 38: 2265.
Truesdell C (1984) *Rational Thermodynamics*, 2nd edn. Berlin: Springer.

Optical Absorption and Reflectance

Y Toyozawa, University of Tokyo, Tokyo, Japan

Introduction

The interaction of electromagnetic waves with matter causes a variety of optical processes, such as reflection, refraction, transmission, absorption, elastic and inelastic scattering, and luminescence, which can be described in terms of relevant optical constants such as reflectivity, refractivity, absorption constant, scattering cross sections, quantum yield, and so on. The optical constants are related with more fundamental material constants: dielectric constant $\varepsilon(\omega)$ or electric conductivity $\sigma(\omega)$, which are response coefficients of matter to an incident electromagnetic wave of circular frequency ω. The real and imaginary parts of $\varepsilon(\omega)$ (and of $\sigma(\omega)$) are related through the Kramers–Kronig relations originating from the causality of the response against external force.

The dispersion, ω-dependence, of the dielectric constant provides one with detailed information on elementary excitations and their interactions. Thus, the spectroscopic study of matter not only brings a fundamental knowledge on microscopic motions taking place within the matter, but also provides a useful basis for a variety of optoelectronic applications of the material.

For the sake of better comprehensibility, the definition of dielectric constant is taken up first, and then, typical microscopic models for that are presented.

Electromagnetic Fields in a Medium

Maxwell equations for the electromagnetic fields in a medium are given by

$$\nabla \times \boldsymbol{E} + \partial \boldsymbol{B}/\partial t = 0, \quad \nabla \times \boldsymbol{H} - \partial \boldsymbol{D}/\partial t = \boldsymbol{J}_{\mathrm{ext}},$$

$$\nabla \cdot \boldsymbol{D} = \rho_{\mathrm{ext}}, \quad \nabla \cdot \boldsymbol{B} = 0$$

For weak electric and magnetic fields, linear relations hold for the electric flux density $\boldsymbol{D} = \varepsilon \boldsymbol{E}$, and the magnetic flux density $\boldsymbol{B} = \mu \boldsymbol{H}$, with dielectric constant ε and magnetic permeability μ. The difference of $\boldsymbol{D}$ from that in vacuum is due to the polarization of the medium: $\boldsymbol{P} = (\varepsilon - \varepsilon_0)\boldsymbol{E} = \chi \boldsymbol{E}$, where ε_0 is the dielectric constant of vacuum and χ is the electric susceptibility (polarizability). The convention followed is that only the foreign charge ρ_{ext} and foreign current $\boldsymbol{J}_{\mathrm{ext}} = \rho_{\mathrm{ext}} \boldsymbol{v}$ appear as inhomogeneous terms on the right-hand side with the understanding that ρ and $\boldsymbol{J}$ of internal origin (due to the constituent particles of the medium) have been removed to the left-hand side to be included as $-\rho = \nabla \cdot \boldsymbol{P}$ and $-\boldsymbol{J} = -\partial \boldsymbol{P}/\partial t$ in the third and second of the Maxwell

equations, respectively. The reason for doing this is that bound and mobile electrons in the medium are indistinguishable under an alternating electric field. For a weak field, a linear relation $\boldsymbol{J} = \sigma \boldsymbol{E}$ holds, where σ is the electric conductivity. ε, σ, χ, and μ are tensors in an anisotropic medium, but unless otherwise stated, the discussion is confined to the isotropic medium in which they can be treated as scalars.

Anticipating that the homogeneous set of Maxwell equations has a solution of a plane wave form, one expresses $\boldsymbol{E} = \boldsymbol{E}_0 \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r} - \mathrm{i}\omega t)$ and obtains

$$[(\varepsilon\mu\omega^2 - k^2) + \boldsymbol{k}\boldsymbol{k}\cdot]\boldsymbol{E}_0 = \boldsymbol{0}$$

There are two types of solutions. One is the transverse wave with $\boldsymbol{E}_0 \perp \boldsymbol{k}$, $\boldsymbol{H}_0 = (\mu\omega)^{-1}\boldsymbol{k} \times \boldsymbol{E}_0$, for which $\omega = ck$, $c = (\varepsilon\mu)^{-1/2}$. This is the electromagnetic wave in the medium with velocity c, which is different from $c_0 = (\varepsilon_0\mu_0)^{-1/2}$ in vacuum. Another is the longitudinal wave ($\boldsymbol{E}_0 \| \boldsymbol{k}$), which has to satisfy $\varepsilon\mu = 0$. If one confines to the nonmagnetic medium, it is possible to express $\mu = \mu_0$ (the same as in vacuum), but in general, the ω-dependence (dispersion) and $\boldsymbol{k}$-dependence (spatial dispersion) of the dielectric constant $\varepsilon(\omega, \boldsymbol{k})$ need to be assumed. It satisfies the relation $\varepsilon(-\omega, -\boldsymbol{k}) = \varepsilon(\omega, \boldsymbol{k})^*$, since $\boldsymbol{D}$ should be real when $\boldsymbol{E}$ is. Therefore, the longitudinal wave appears only at ω satisfying $\varepsilon(\omega, \boldsymbol{k}) = 0$, which means that $\boldsymbol{E} = -\varepsilon_0^{-1}\boldsymbol{P}$. Namely, this longitudinal oscillation arises due to the self-induced depolarizing electric field, being independent of the electromagnetic wave. A few examples are mentioned later.

For a spectral region below soft X-rays, the wavelength $\lambda = 2\pi/k$ is much greater than the interatomic distance of condensed matter, and the $\boldsymbol{k}$-dependence is hereafter ignored unless otherwise stated.

From the above argument with $\partial\boldsymbol{P}/\partial t \equiv \boldsymbol{J}$, one finds that $\varepsilon(\omega)$ and $\sigma(\omega)$ are not independent but are related by $(\varepsilon(\omega) - \varepsilon_0) = \sigma(\omega)/(-\mathrm{i}\omega)$, namely, $\varepsilon_1(\omega) - \varepsilon_0 = -\sigma_2(\omega)/\omega$ and $\varepsilon_2(\omega) = \sigma_1(\omega)/\omega$ where the suffixes 1 and 2 refer to real and imaginary parts. $\sigma_1(\omega)$, which is related to Joule's heat and $\varepsilon_2(\omega)$, represent the "dissipative" parts which are always non-negative as they should be, while $\varepsilon_1(\omega)$ and $\sigma_2(\omega)$ represent the "reactive" parts. Due to the causality between the electric field and its responses, the real and imaginary parts of $\varepsilon(\omega)$ are related by Kramers–Kronig relations:

$$\varepsilon_1(\omega) - \varepsilon_0 = \pi^{-1}P\int_{-\infty}^{\infty} \varepsilon_2(\omega')\,\mathrm{d}\omega'/(\omega' - \omega)$$

$$\varepsilon_2(\omega) = \pi^{-1}P\int_{-\infty}^{\infty} (\varepsilon_1(\omega') - \varepsilon_0)\,\mathrm{d}\omega'/(\omega' - \omega)$$

Classical Models for Mobile and Bound Charges, and Their Correspondence to Quantum Mechanical Systems

Consider free electrons in metals with scattering rate τ^{-1}. Under an oscillating electric field, the equation of motion for the velocity, $\mathrm{d}v/\mathrm{d}t - \boldsymbol{v}/\tau = (e\boldsymbol{E}/m)\exp(-\mathrm{i}\omega t)$, gives a forced oscillation solution:

$$\boldsymbol{v} = (e\boldsymbol{E}/m)(-\mathrm{i}\omega + \tau^{-1})^{-1}\exp(-\mathrm{i}\omega t)$$

from which follows the complex conductivity:

$$\sigma(\omega) = (ne^2/m)(-\mathrm{i}\omega + \tau^{-1})^{-1}$$

where n is the number density of electrons with charge e and mass m. While it gives a static conductivity $\sigma = ne^2\tau/m$, it gives an imaginary conductivity and hence, negative susceptibility

$$(\varepsilon(\omega) - \varepsilon_0) = \mathrm{i}\sigma(\omega)/\omega = -ne^2/m\omega^2$$

for high frequency of spectroscopic interest such that $\omega \gg \tau^{-1}$ ($\sim 10^{14}\ \mathrm{s}^{-1}$). Rewriting $\varepsilon_1(\omega)/\varepsilon_0 = 1 - \omega_\mathrm{p}^2/\omega^2$, one finds that $\varepsilon_1(\omega)$ is negative and $\varepsilon_2(\omega) = 0$ at $\omega < \omega_\mathrm{p}$, where ω_p defined by $\omega_\mathrm{p}^2 = ne^2/\varepsilon_0 m$ is the circular frequency of the plasma oscillation (the longitudinal charge density wave) satisfying $\varepsilon(\omega) = 0$ as mentioned before, and is $\sim 10^{16}\,\mathrm{s}^{-1}$ in normal metals.

As a classical model for a bound electron in an insulator, consider a set ($j = 1, 2, \ldots$) of damped harmonic oscillators with electric charge $\sqrt{f_j}e$, circular frequency ω_j, and damping constant γ_j. The equation of motion for each oscillator under an alternating electric field,

$$\mathrm{d}^2\boldsymbol{r}_j/\mathrm{d}t^2 + \gamma_j\mathrm{d}\boldsymbol{r}_j/\mathrm{d}t + \omega_j^2\boldsymbol{r}_j = \sqrt{f_j}(e/m)\boldsymbol{E}\exp(-\mathrm{i}\omega t)$$

leads to the electric polarizability

$$\begin{aligned}\chi(\omega) &= \varepsilon(\omega) - \varepsilon_0 \\ &= N_0\sum_{j(\neq 0)} (e^2/m)f_j/(\omega_j^2 - \omega^2 - \mathrm{i}\gamma_j\omega)\end{aligned}$$

where N_0 denotes the number of such sets of oscillators per unit volume of the medium.

The dimensionless constant f_j turns out to correspond to the quantum mechanical oscillator strength. In fact, the time-dependent Schrödinger equation for an electron starting from the ground state $j = 0$ of an atom subject to perturbation, $-\boldsymbol{m}\cdot\boldsymbol{E}\exp(-\mathrm{i}\omega t)$ with dipole moment $\boldsymbol{m} \equiv e\boldsymbol{r}$, gives

$$\langle \boldsymbol{m} \rangle_t = \boldsymbol{E}\exp(-\mathrm{i}\omega t)\sum_{j(\neq 0)} (2\omega_{j0}|\boldsymbol{m}_{0j}|^2/\hbar)(\omega_{j0}^2 - \omega^2)^{-1}$$

where $\hbar\omega_j \equiv \hbar\omega_{j0} \equiv E_j - E_0$ is the jth excitation energy of the atom. Comparison with the above expression for polarizability (γ_j being neglected) indicates that $f_j = f_{j0} = 2m\omega_{j0}|\boldsymbol{r}_{j0}|^2/\hbar$ is the oscillator strength of the transition $0 \rightarrow j$, satisfying the sum rule $\sum_{j(\neq 1)} f_{ji} = 1$ per one electron of the atom as can be derived from the commutation relation $xp - px \equiv [x, p] = i\hbar$ and the Heisenberg equation of motion, $p \equiv m\,\mathrm{d}x/\mathrm{d}t = m[x, H]/i\hbar$ (H is the Hamiltonian). Note that the (circular) frequency ω_j of the classical oscillation corresponds to the transition frequency ω_{j0} of the quantum mechanical system (correspondence principle). In the case of a harmonic oscillator, only the transition $i \rightarrow j = i \pm 1$ is allowed with $f_{i+1,i} = i + 1, f_{i-1,i} = -i$ and hence, $f_1 = 1$ and $f_0 = f_{i(\geqslant 2)} = 0$, in contradistinction to the case of an atom, which is so to speak, an anharmonic oscillator. This is the etymology of "oscillator strength" introduced above – the "oscillator equivalence" of the strength of transition. In the case of a medium consisting of N_0 identical atoms per unit volume, each with n_0 electrons, the sum rule can also be written as $\int_0^\infty \varepsilon_2(\omega)\omega\,\mathrm{d}\omega = \pi N_0 n_0 (e^2/m)$.

Transverse and Longitudinal Waves

It is to be noted that $\varepsilon_1(\omega)$ consists of a series ($j = 1, 2, \ldots$) of poles at $\omega_j^{\mathrm{t}} \equiv \omega_j$ representing transverse waves, each being followed by a zero at $\omega_j + \Delta_j \equiv \omega_j^{\mathrm{l}}$ representing longitudinal wave. The longitudinal–transverse ("lt") splitting is given by $\Delta_j = N_0 |\boldsymbol{m}_{0j}|^2/\hbar\varepsilon_j'$ with screening constant ε_j' for the dipole–dipole interaction being given by the "residual" dielectric constant as defined by

$$\varepsilon_j'(\omega) \equiv \varepsilon_0 + (N_0 e^2/m) \sum_{j'(\neq j)} f_{j'}/(\omega_{j'}^2 - \omega^2)$$

at $\omega = \omega_j^{\mathrm{l}}$. The free electron model for metal corresponds to the case with vanishing ω_j of the oscillator model with Δ being ω_{p}. The lt splitting Δ in either case originates from the depolarizing force as mentioned before. Thus, the dipole–dipole interaction energy $N_0|\boldsymbol{m}_{0j}|^2$ (note that $N_0 \sim r^{-3}$, where r is the interatomic distance) as well as the plasmon energy $\hbar\omega_{\mathrm{p}}$ have been obtained without explicit consideration of the Coulomb interaction between electrons to which they are to be attributed. In fact, the Coulomb interaction has been incorporated implicitly in the form of a depolarizing field. This phenomenological method allows one to give even the appropriate screening constant to be used for the dipole–dipole interaction, which belongs to a many-body effect of the higher-order perturbation theory.

Such conjugation of transverse and longitudinal waves holds not only for electronic excitation (transverse and longitudinal excitons) but also for lattice vibrations (transverse and longitudinal optical modes).

Quantization of Radiation Field and the Spontaneous Emission of Radiation

So far, the material system has been described, first by classical mechanics of harmonic oscillators and then by quantum mechanics of atoms, whereby the electromagnetic fields have been left as classical quantities. Historically, however, the quantum theory started with the quantization of an electromagnetic field, namely, with Planck's hypothesis on discrete energy levels of a radiation field, which was later substantiated by Einstein with his model of a quantum of the radiation field, that is, a photon. Einstein's argument on thermal equilibrium between a radiation field and two level atoms not only necessitated the spontaneous emission (as the consequence of quantization) in addition to the induced emission, but also related the rates B and A of absorption and emission of a photon in a way which proved to be correct even under the quantum mechanics established later than his prediction.

The intensity of a line spectrum is represented by the oscillator strength f_j as mentioned before, or by the rate of optical absorption per spectral density $w(\omega)$ of the incident radiation energy, $B_{0j} = \pi|\boldsymbol{m}_{0j}|^2/3\varepsilon_0\hbar^2$. According to Einstein, this B coefficient is related with his A coefficient,

$$A_{j0} = (g_0/g_j)B_{0j}\hbar\omega_{j0}^3/\pi^2 c_0^3 = \tau_j^{-1}$$

which is the rate of the spontaneous emission of a photon from the excited state j, where g_j is the degeneracy of the state j. The damping constant γ_j in the oscillator model introduced above and hence, the spectral width of the absorption line is given by the reciprocal of this natural (radiative) lifetime τ_j of the final state j in the case of an isolated atom, in accordance with the uncertainty principle between energy and time.

In the case of an atom or a molecule in a gas with density not too low, γ_j is dominated by larger rate processes, such as collision with other atoms or molecules. (In the case of a molecule, the excited states j are characterized not only by electronic but also by vibrational states, so that the absorption spectrum is much more complicated).

In insulators, the discrete lines in the absorption spectrum correspond to the transverse exciton states with γ_j being dominated by the rate of exciton

scattering by phonons, and the continuous spectra correspond to electronic transitions from filled bands to empty bands. In the spectral region of much lower frequency, there are discrete lines due to the excitation of a phonon of transverse optical modes, with γ_j being governed by the rate of phonon–phonon scattering due to the anharmonicity of lattice vibrations.

Relations between Optical Constants and Dielectric Constant

In order to experimentally obtain the complex dielectric function $\varepsilon(\omega)$, the basic quantity in which the dynamics of microscopic motions in the medium are directly reflected as mentioned above, one must resort to measurements of various optical constants which are functionally related with $\varepsilon(\omega)$ as follows.

When the electromagnetic wave comes from vacuum into a medium with dielectric constant $\varepsilon_1(\omega)+\mathrm{i}\varepsilon_2(\omega)$ $(\varepsilon_2(\omega)\geqslant 0)$, and magnetic permeability $\mu=\mu_0$ (the same as in vacuum), the wave vector becomes $\sqrt{(\varepsilon_1(\omega)+\mathrm{i}\varepsilon_2(\omega))/\varepsilon_0}\equiv\tilde{n}(\omega)\equiv n(\omega)+\mathrm{i}\kappa(\omega)$ times of its value k in vacuum (the square root with non-negative $n(\omega)$ and $\kappa(\omega)$ is to be chosen). This means that the amplitude of the electromagnetic wave decays as $\exp(-\kappa kx)$ in the medium with the extinction coefficient $\kappa(\omega)$, and the wavelength becomes n^{-1} times of its value $\lambda=2\pi/k$ in vacuum. n is the refractive index with which the angle θ_r of the refracted beam (measured from the perpendicular of the surface) is determined against the given angle θ_i of incidence through the Snell's law $n=\sin\theta_\mathrm{i}/\sin\theta_\mathrm{r}$. From this law, it also follows that the light beam incident from a medium of refractivity n' upon a medium of smaller refractivity n'' is subject to total reflection (because of no solution for θ_r), when the angle θ_i of incidence is greater than the critical angle θ_c, given by $\sin\theta_\mathrm{c}=n''/n'$. $\tilde{n}$ defined above, is called the complex refractive index.

For normal incidence, the reflectivity, the ratio of the electromagnetic energies of the reflected and incident waves, is given by

$$\begin{aligned}R(\omega)&=|(\tilde{n}-1)/(\tilde{n}+1)|^2\\&=\{(n-1)^2+\kappa^2\}/\{(n+1)^2+\kappa^2\}\end{aligned}$$

To be more elaborate, the ratios $H_\mathrm{r}/H_\mathrm{i}$ and $-E_\mathrm{r}/E_\mathrm{i}$ of the magnetic and electric fields between reflected and incident waves, inclusive of their phases, are given by the complex reflectivity $\tilde{r}(\omega)=(\tilde{n}-1)/(\tilde{n}+1)$. Rewriting this as $\tilde{r}(\omega)=\sqrt{R(\omega)}\exp[\mathrm{i}\theta(\omega)]$, one obtains the relations:

$$\begin{aligned}n(\omega)&=(1-R)/(1+R-2\sqrt{R}\cos\theta)\\\kappa(\omega)&=2\sqrt{R}\sin\theta/(1+R-2\sqrt{R}\cos\theta)\end{aligned}$$

From the analyticity of the complex refractivity $\tilde{n}(\omega)$, one obtains the Kramers–Kronig relations between its real and imaginary parts,

$$n(\omega)-1=\pi^{-1}P\int_0^\infty 2\omega'\kappa(\omega')\mathrm{d}\omega'/(\omega'^2-\omega^2)$$

and

$$\kappa(\omega)=-2\omega\pi^{-1}P\int_0^\infty (n(\omega')-1)\mathrm{d}\omega'/(\omega'^2-\omega^2)$$

Now the intensity of light in the medium with a complex dielectric constant decays as $\exp(-\alpha(\omega)x)$, and the so-defined absorption coefficient $\alpha(\omega)$ is related with the extinction coefficient by $\alpha(\omega)=2\omega c_0^{-1}\kappa(\omega)$, as mentioned before. If one measures the absorption coefficient $\alpha(\omega)$ over the entire region, one can obtain $n(\omega)$, $\kappa(\omega)$ and hence, the dielectric function $\varepsilon(\omega)$.

Another method of obtaining $\varepsilon(\omega)$ is the Kramers–Kronig analysis of reflectivity. Assuming the analyticity of the logarithm of the complex refractivity $\ln\tilde{n}(\omega)$ on the upper half of the complex ω-plain, one can derive

$$\theta(\omega)=-\omega\pi^{-1}\int_0^\infty \ln[R(\omega')/R(\omega)]\mathrm{d}\omega'/(\omega'^2-\omega^2)$$

where, $0\leqslant\theta(\omega)\leqslant\pi$. From the reflectivity spectrum $R(\omega)$ measured over the entire region, one obtains $\theta(\omega)$ and hence $n(\omega)$ and $\kappa(\omega)$, and finally $\varepsilon(\omega)$.

Polariton Picture for the Electromagnetic Wave in a Dispersive Medium, the Optical Absorption and the Reflectance

To see more clearly the behaviors of the electromagnetic wave with frequency near a pole ω_j^t of the lossless dispersion $\varepsilon(\omega)$, consider a single pole $\omega_\mathrm{t}\equiv\omega_j^\mathrm{t}$, which is well isolated from other poles. First invert the abscissa and the ordinate to show ω as a function of ε as shown in **Figure 1b**. Then plot in **Figure 1a** the real and imaginary parts of the complex wave vector $\tilde{n}(\omega)k\equiv n(\omega)k+\mathrm{i}\kappa(\omega)k\equiv k_\mathrm{r}+\mathrm{i}k_\mathrm{i}$ towards the right and left, respectively, of the abscissa k against ω as ordinate. Finally, one plots in **Figure 1c** the reflectivity $R(\omega)$ against ω in the same way. **Figure 1a** indicates that the ω–k dispersion of the electromagnetic wave is distorted as a result of resonant mixing with the polarization wave of the medium. A quantum of this mixed wave is called polariton. Note that the asymptotic light velocities of the upper and lower branches of the polariton are different due to the contribution of the resonance to the dielectric constant.

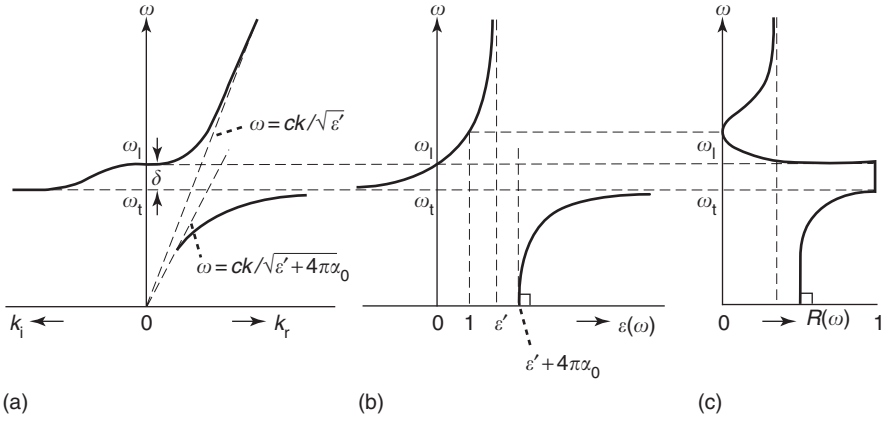


Figure 1 (a) Spatial dispersion of polariton, $\omega(k)$. (b) Dielectric dispersion, $\varepsilon(\omega)$. (c) Reflectivity, $R(\omega)$. (Reproduced with permission from Toyozawa Y (2003) *Optical Processes in Solids*. Cambridge: Cambridge University Press.)

There is no propagating wave in the medium between ω_t and ω_l, where the incident electromagnetic wave is totally reflected as shown in **Figure 1c**. Instead, there is an evanescent mode decaying as $\exp(-k_i x)$ toward the inside. The same holds in metals ($\omega_t = 0$) in much wider spectral regions, $\tau^{-1} \ll \omega < \omega_p$. Thus, the visible light is almost totally reflected on the metal surface, the reason why metals glitter.

Each pole in $\varepsilon_1(\omega)$ is blurred by the damping constants γ, and the corresponding resonance line in $\varepsilon_2(\omega)$ is broadened by γ. The spectrum of $\varepsilon_2(\omega)$ represents what is called absorption, but one must be careful in relating it with the spectrum of the absorption constant. The observed spectrum of $R(\omega)$ in some insulators shows strong resonance peaks often as high as almost reaching unity, and the dips as deep as nearly vanishing (as shown in **Figure 1c** in the idealistic situation of vanishing γ). In such a situation ($\gamma \lesssim \Delta$), the spectral shape of the absorption constant

$$\alpha(\omega) = (c_0 \varepsilon_0)^{-1} (\omega / n(\omega)) \varepsilon_2(\omega)$$

is significantly different from that of the basic quantity $\varepsilon_2(\omega)$, because of the factor $n(\omega)$ which varies significantly around the resonance. This is the reason why the Kramers–Kronig type analysis is necessary to obtain, from the observed optical constants, a more basic quantity $\varepsilon_2(\omega)$, so as to facilitate a quantitative comparison with particular theoretical models. For weak resonances (small oscillator strengths) such as weak exciton lines and impurity absorption bands, it is necessary to approximate $n(\omega)$ as a constant and to express

$$\alpha(\omega) = \text{const. } \omega \varepsilon_2(\omega) = \text{const. } \sigma(\omega)$$

If the lower branch of the polariton has a $\boldsymbol{k}$-dependence as in the case of an exciton with finite effective mass, one needs an additional boundary condition (abbreviated as ABC) besides the Maxwell boundary condition, in order to see how the incoming electromagnetic wave is converted to the lower and upper branch polaritons. More basically, this problem is to be studied in the context of a nonlocal theory with spatial dispersion $\varepsilon(\boldsymbol{k}, \omega)$ and surface structure being considered.

The Electric Quadrupole and Magnetic Dipole Transitions

So far, the discussion has been confined to the electric dipole allowed transitions. When the transition matrix elements μ_{j0} of the dipole moment between the states j and 0 vanish, one must explore the possible transitions due to higher multipoles. To see this, one starts with the canonical expression for the kinetic energy of an electron in an electromagnetic field, $(\boldsymbol{p} - \mathrm{e}\boldsymbol{A}(\boldsymbol{r}))^2/2m$, and takes the linear term in the vector potential $\boldsymbol{A}$, $H_1 = -(e/2m)[\boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{r}) + \boldsymbol{A}(\boldsymbol{r}) \cdot \boldsymbol{p}]$ as the interaction of an electron with the electromagnetic field of which the electric field is given by $-\partial \boldsymbol{A}/\partial t = \boldsymbol{E} = \boldsymbol{E}_0 \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r} - \mathrm{i}\omega t)$, and hence, the amplitude of magnetic flux density by $\boldsymbol{B}_0 = \omega^{-1} \boldsymbol{k} \times \boldsymbol{E}_0$. Considering that the wave vector of the electromagnetic wave under consideration is much smaller than the reciprocal of atomic radius a, one takes the power series expansion $\exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}) = 1 + \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}$ neglecting higher order terms, to obtain $H_1 = [-(e/\mathrm{i}m\omega)\boldsymbol{p} \cdot \boldsymbol{E}_0 - (e/\mathrm{i}m\omega)\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{Q} \cdot \boldsymbol{E}_0 - (e/2m)\boldsymbol{l} \cdot \boldsymbol{B}_0] \exp(-\mathrm{i}\omega t)$, where $Q_{\alpha\beta} \equiv (1/2)(p_\alpha r_\beta + r_\beta p_\alpha)$ is the symmetric tensor, and $\boldsymbol{l} \equiv \boldsymbol{r} \times \boldsymbol{p}$ is the orbital angular momentum. Making use of the equation of motion, $\boldsymbol{p}/m = \mathrm{d}\boldsymbol{r}/\mathrm{d}t = (\mathrm{i}/\hbar)[H, \boldsymbol{r}]$, one can find that the transition matrix elements of the first and second terms of H_1 represent those of the interactions of electric dipole and electric quadrupole moments with

the electric field, respectively. The third term represents the Zeeman term for the orbital magnetic moment in the magnetic field as is obvious from the well-known magnetogyric factor $\gamma = (e/2m)$, and is to be combined with that for the spin magnetic moment, not considered so far, with twice as large a magnetogyric factor, to give $-(e/2m)(\boldsymbol{l} + 2\boldsymbol{s}) \cdot \boldsymbol{B}_0$, where $\boldsymbol{s}$ is the spin angular momentum.

When the electric dipole transition (E_1) is forbidden between states 0 and j, one must consider the electric quadrupole transition (E_2) and/or the magnetic dipole transitions (M_1), whose oscillator strengths are of the order $(ka)^2 \sim 10^{-6}$ in contrast to the order of unity in (E_1) transitions as explained above.

See also: Dielectric Function; Electron–Phonon Interactions and the Response of Polarons; Excitons: Theory; Insulators, Optical Properties of.

PACS: 77.22.Ch; 78.20. – e; 78.40. – q

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Saunders College.

Cho K (2003) *Optical Response of Nanostructures*. Berlin: Springer.

Davydov AS (1971) *Theory of Molecular Excitons*. New York: Plenum Press.

Heilter W (1954) *The Quantum Theory of Radiation*, 3rd edn. Oxford: Clarendon Press.

Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.

Loudon R (1973) *The Quantum Theory of Light*, 2nd edn, section 1.5. Oxford: Clarendon Press.

Panofsky WKH and Phillips M (1962) *Classical Electricity and Magnetism*, 2nd edn. Reading: Addison-Wesley.

Stern P (1963) Kramers–Kronig analysis of reflectivity. In: Seitz F and Turnbull D (eds.) *Solid State Physics*, vol. 15, pp. 333–340. New York: Academic Press.

Toyozawa Y (2003) *Optical Processes in Solids*. Cambridge: Cambridge University Press.

Wooten F (1972) Kramers-Kronig analysis of reflectivity. In: *Optical Properties of Solids*, pp. 240–250. New York: Academic Press.

Optical Bistability

R Bonifacio and F Casagrande, Università degli Studi di Milano, Milan, Italy

Introduction

With the advent of the laser, in 1960, an entirely new field, nonlinear optics quickly developed, and progress in quantum optics and cavity quantum electrodynamics received a huge impulse. It was in this period that a great excitement was raised by optical bistability (OB), a phenomenon that can be described as a kind of passive counterpart of the laser.

In a laser system, an atomic medium is placed in an optical cavity where a transition is resonant with one cavity mode. Atoms are excited so that population inversion is created in the relevant levels, and the stimulated emission overcomes the absorption. If the resulting amplification can balance the cavity losses, a stationary oscillation establishes in the resonator, whose output is the laser field.

Consider a system of this kind, but with two main differences: the atomic medium is not pumped, but the optical cavity is driven by a coherent external field. This system can exhibit the phenomenon of "optical bistability": the intensity of the light transmitted from the resonator can take two different values corresponding to the same intensity of the incident field.

The first predictions of this effect were made in 1969 by Szöke, Daneu, Goldhar, and Kurnit, and independently by Seidel in an application for a US patent. The first experimental observation was achieved in 1975 by Mc Call, Gibbs, and Venkatesan, who used vapors of sodium atoms in a Fabry–Perot cavity driven by a CW dye laser. One year later a basic theoretical model of optical bistability was presented by Bonifacio and Lugiato.

The fundamental physics of this phenomenon is outlined, and the main developments in the theoretical understanding and the experimental implementations of OB and related applications are reviewed below.

Atomic Polarization and Dielectric Susceptibility

The main features of optical bistability can be understood by considering a semiclassical treatment.

A classical monochromatic radiation field, $\boldsymbol{E}(\boldsymbol{x}, t) = \boldsymbol{E}_0 \cos(\omega t - \boldsymbol{k} \cdot \boldsymbol{x})$, induces a polarization $\boldsymbol{P}(\boldsymbol{x}, t)$ in an ensemble of N two-level atoms in a volume V, where

$$\begin{aligned}\boldsymbol{P}(\boldsymbol{x}, t) &= \varepsilon_o \,\mathrm{Re}\{\chi \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)]\} \\ &= \varepsilon_o [\chi_d \cos(\omega t - \boldsymbol{k} \cdot \boldsymbol{x}) \\ &\quad + \chi_a \sin(\omega t - \boldsymbol{k} \cdot \boldsymbol{x})] \boldsymbol{E}_0\end{aligned}$$

The complex dielectric susceptibility, $\chi = \chi_d + i\chi_a$, describes the response of the medium to the field. The induced atomic polarization exhibits two components, in phase and in quadrature with the field.

For weak atom–field interaction, the susceptibility is independent of the field amplitude, and this is the realm of linear optics. The dynamics of weakly excited atoms can be simply modeled by forced, harmonically oscillating dipoles. For a low-density atomic gas, where the refractive index $n = \sqrt{1+\chi} \cong 1 + \chi/2$, the real and imaginary parts of susceptibility, χ_d and χ_a, control the dispersive and absorptive effects via two parameters, the wave number κ and the absorption coefficient α, respectively, where

$$\kappa = \frac{\omega}{2c}\,\chi_d, \quad \alpha = \frac{\omega}{2c}\,\chi_a$$

For large enough field amplitude, however, the dielectric susceptibility becomes a nonlinear function of the field intensity $I = E_0^2/2$. In fact, the optical Bloch equations provide the following expression for the susceptibility:

$$\chi = \frac{N\mu^2}{V\varepsilon_0\hbar}\,\frac{\omega_a - \omega + i\gamma}{(\omega_a - \omega)^2 + \gamma^2 + (\mu^2/\hbar^2)I}$$

where μ is the modulus of the dipole matrix element for the relevant atomic transition, ω_a the Bohr transition frequency, and γ the spontaneous emission rate, here assumed as the only source of atomic decay, phenomenologically added to the Bloch equations.

The above result shows that both the dispersive (in phase) and the absorptive (in quadrature) components of induced polarization can exhibit a nonlinear dependence on the field amplitude, that is one of the clues to optical bistability in the corresponding dispersive or absorptive regimes.

Absorptive Optical Bistability: Steady-State Equation

In the case of resonance between the atomic transition and the radiation field, $\omega_a = \omega$, the dispersive component of susceptibility vanishes, $\chi_d = 0$, from the expression of the on-resonance absorptive component, χ_a, and one can write the absorption coefficient as

$$\alpha = \frac{\alpha_0}{1 + I/I_S}$$

where $\alpha_0 = \omega N\mu^2/2cV\varepsilon_0\hbar\gamma$ is the small signal value, and $I_S = \hbar^2\gamma^2/\mu^2$ the saturation intensity. The above expression shows that, for increasing field intensity, the absorption undergoes saturation: the coefficient α can decrease well below the small signal value α_0, until the absorber eventually becomes transparent. In this limit, the populations of the two relevant levels are equal. The saturation of absorption is a basic ingredient for the occurrence of absorptive optical bistability, and is the passive equivalent of gain saturation in laser systems.

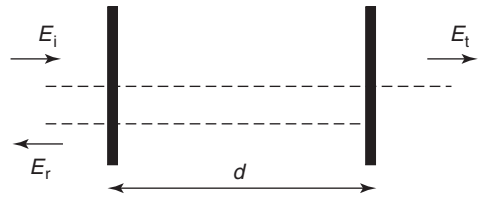


Figure 1 Fabry–Perot cavity with two partially reflecting mirrors at a distance d. The incident field E_i is partially reflected (E_r) and partially transmitted (E_t) by the optical cavity.

Absorptive OB can be observed when, for example, the atomic system fills a Fabry–Perot cavity (**Figure 1**), which is driven by a coherent incident field, E_i, oscillating at the same frequency ω of both the cavity mode and the atomic transition. The cavity mirrors provide an intrinsic feedback mechanism, by which the reflected cavity field recombines with the injected driving field. Feedback adds to nonlinearity for the occurrence of OB.

In the semiclassical treatment of system dynamics, the optical Bloch equations which describe the atomic polarization and the level of populations induced by the cavity field, are self-consistently coupled to the Maxwell equation for the cavity-field amplitude, generated by the atomic polarization as well as by the injected field, which acts as a source term. At steady state, the field propagation equation can be integrated with respect to the relevant space coordinate and with proper boundary conditions. When standing-wave effects can be averaged, the treatment for a Fabry–Perot cavity becomes equivalent to the simpler one of a unidirectional ring cavity. In this case a simple analysis shows that, at steady state, in the small absorption limit $\alpha d \ll 1$, where d is the cavity length, the ratio of transmitted to incident field amplitude can be written as

$$\frac{E_t}{E_i} = \frac{1}{1 + \alpha d/T}$$

where $T = I_t/I$ is the mirror transmittivity. By inserting the expression for the absorption coefficient α, and solving for the incident field amplitude, one derives the stationary state equation for

absorptive OB:

$$y = x + 2C\frac{x}{1+x^2}$$

It provides a nonlinear relation between the dimensionless transmitted field amplitude, x, and incident field amplitude, y, defined as $x \equiv E_t/\sqrt{I_S T}$, $y \equiv E_i/\sqrt{I_S T}$. One can notice the two contributions to the transmitted field, due to the injected field, and to radiation from the absorptive component of atomic polarization in quadrature with the field. The latter contribution is an atomic cooperative one. Actually, the steady-state equation is ruled by a single parameter, the cooperation parameter C, which can be written in terms of both macroscopic and microscopic system parameters as follows:

$$C = \frac{\alpha_0 d}{2T} = \frac{Ng^2}{2k\gamma}$$

where g is the atom–cavity mode coupling frequency, and k the cavity line width.

Absorptive Optical Bistability: Hysteresis Cycle

In **Figure 2**, the transmitted field amplitude x is plotted as a function of the incident field amplitude y, obtained by inverting the steady-state relation $y = y(x)$, for different values of the cooperation parameter C.

The dashed line for $C = 0$ (**Figure 2a**) shows for reference the empty cavity limit, $x = y$. For $0 < C < 4$ (**Figure 2b**), the curve exhibits a nonlinear monotonic behavior. It connects a cooperative, low-transmission linear regime, where $x \approx y/(1 + 2C)$ for $x^2 \ll 1$, with a one-atom, high-transmission linear regime, where $x \approx y$ for $x^2 \gg 1$. In between, $y(x)$ exhibits a region with a steep slope where the system can operate as an optical transistor, since a small amplitude incident field is transformed into a large amplitude output field.

$C = 4$ (**Figure 2c**) is a critical value. For $C > 4$ (**Figure 2d**), the curve exhibits the characteristic S-shaped behavior which is the signature of bistability. In fact, for $C > 4$ the equation $x = x(y)$ admits three real solutions on a range of input values, however, the portion of the curve with negative slope is unstable, that is, the corresponding states are unphysical. Though this instability may be easily checked mathematically, it can be readily understood in physical terms by realizing that on the negative-slope portion of $x(y)$ a decrease of the incident field would correspond to an increase of the transmitted one. Hence, there are only two accessible stationary states, corresponding to a single output value (bistability).

Which is really the state of the system depends on its history, as illustrated in **Figure 3**. For example, consider a value $y = y_1$ in the bistable range, $y_d < y_1 < y_u$. Depending on whether y_1 is approached for increasing or decreasing values of y, the system will be in the state (x_1, y_1) on the lower branch of the hysteresis cycle, or in state (x_3, y_1) on the upper branch, respectively.

In general, starting with a small incident field amplitude, the system remains on the lower branch for increasing y, up to $y = y_u$, at this value the system jumps on the upper branch. Conversely, starting with large incident amplitude, the system remains on the upper branch for decreasing y, until $y = y_d < y_u$,

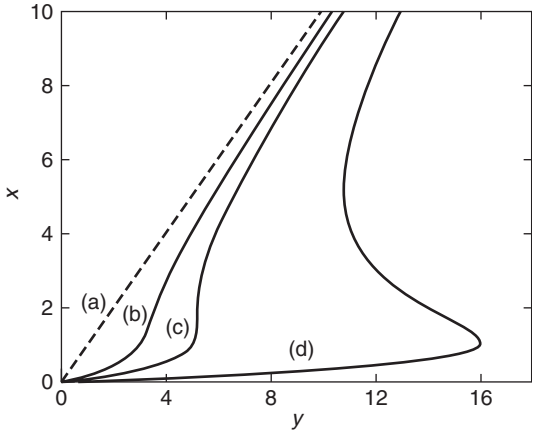


Figure 2 Dimensionless transmitted field amplitude x vs. incident field amplitude y in the stationary state of absorptive optical bistability. (a) $C = 0$, (b) $C = 2$, (c) $C = 4$, (d) $C = 15$.

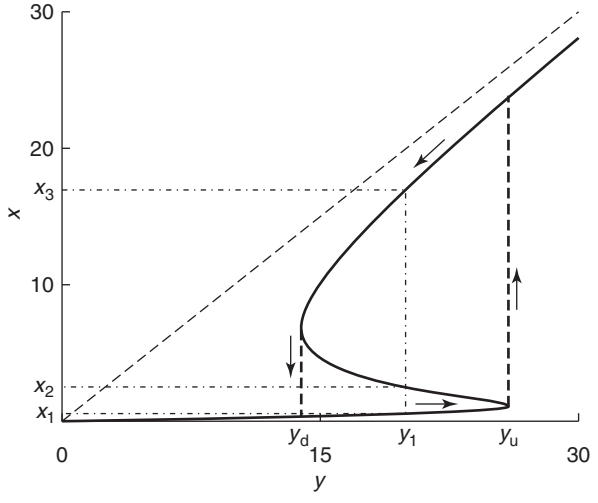


Figure 3 Hysteresis cycle in absorptive optical bistability for transmitted field amplitude x vs. incident field amplitude y, for cooperation parameter $C = 25$.

when it jumps down to the lower branch and thus closes the cycle.

The physical mechanism is as follows. For small intensity, the incident field is mostly reflected due to strong atomic absorption in the cavity. By increasing the incident field intensity, the absorption decreases, until the atomic transition is saturated, and suddenly the transmittivity approaches the empty cavity limit. If the incident intensity is decreased from this limit, the absorber remains bleached well below the upper transition value $y = y_u$, until at $y = y_d$ the intensity is so low that the transmitted field comes back to the small initial values.

The main theoretical predictions were confirmed in a series of experiments by Kimble and co-workers, culminating in a quantitative test by Orozco, Kimble, and Rosenberger in 1987.

Though hysteresis effects are quite familiar in condensed matter physics, for example, in magnetism and ferroelectricity, the relevant point here is that optical bistable systems are open systems driven out of thermodynamical equilibrium. In particular, if one considers the steady-state curves $y(x)$ for different values of C, instead of the curves $x(y)$ as in **Figure 2**, they closely resemble the van der Waals curves which describe the transition from gas to liquid in a system at thermodynamical equilibrium. More precisely, x, y, and C play the role of pressure, volume, and temperature, respectively, and the curve for $C = 4$ is equivalent to the critical isotherm. Hence, just like the laser is a prototype of a first-order phase transition in out-of-equilibrium (dissipative) systems, OB represents a prototype of second-order phase transition in the same framework.

Dispersive Optical Bistability

A similar behavior also shows up in the dispersive regime of optical bistability, though the underlying physical mechanism is different. In this regime, the system operates well off-resonance, so that the atomic response is ruled by the dispersive component χ_d of dielectric susceptibility, whereas the absorptive part χ_a is negligible. Hence, the saturation of dispersion is crucial in this case. If one evaluates the expression of χ_d in a large detuning limit, where in the denominator γ^2 is negligible and $(\omega_a - \omega)^2 \gg (\mu^2/\hbar^2)I$, then the dispersive susceptibility can be approximated as follows:

$$\chi_d \cong \frac{N\mu^2}{V\varepsilon_0\hbar}\frac{1}{\omega_a - \omega}\left[1 - \frac{(\mu^2/\hbar^2)I}{(\omega_a - \omega)^2}\right] \equiv \chi_0 + \chi_2 I$$

In this limit, the (real) refractive index of the atomic medium can be written in the characteristic form of an optical Kerr medium: $n = n_0 + n_2 I$, where $n_0 = 1 + \chi_0/2$ and $n_2 = \chi_2/2$. The atoms thus behave as a nonlinear dielectric medium with a third-order susceptibility.

Kerr nonlinearity and intrinsic feedback conspire for the occurrence of dispersive OB in a coherently driven optical cavity, where at steady state the transmitted field intensity can assume two different values for the same incident intensity.

The jump, for example, from low to high transmission branch of the hysteresis cycle is explained as follows: for increasing intensity of the incident field, the variation of the refractive index, hence the effective wave number, can shift the cavity frequency ω to coincide with the driving field frequency ω_0, that is, on resonance, where the transmission is suddenly raised. The reverse process is similarly explained by a return to off-resonance conditions.

Just dispersive effects allowed the mentioned first observation of optical bistability.

Dynamical Effects

The dynamics of optical bistable systems has been the object of careful investigation since the early days, especially after the predictions of instabilities of different kinds in the late seventies. McCall predicted regenerative pulsations in the presence of nonlinearities with some specific features, and demonstrated their occurrence with a hybrid electrooptical device.

Bonifacio and Lugiato showed by a many-mode model that, under proper conditions, a portion of the high-transmission branch of the hysteresis cycle can be unstable. In this case, the system can operate as a kind of transformer of continuous light into pulsed light, where the self-pulsing oscillation occurs with a period equal to the cavity round-trip time.

Another instability mechanism was suggested by Ikeda in 1979 by investigating dispersive bistability with time delay effects, namely, when the time spent by radiation to propagate from the output to the input end of the nonlinear medium is longer than the atomic decay time. This instability can lead to period doubling and chaos, as first observed in 1981 by Gibbs, Hopf, Kaplan, and Shoemaker. It was then shown that the Ikeda instability emerges in a suitable limit of the general treatment by Bonifacio and Lugiato.

A number of other instabilities were investigated, including those due to the transverse structure of the cavity field, or originated by counter-propagating radiation beams in bidirectional ring cavities containing a saturable absorber and driven by an external coherent field. These features stimulated further studies in nonlinear and chaotic dynamics as well as in nonequilibrium statistical mechanics.

Quantum Treatment and Experiments in Cavity Quantum Electrodynamics

Quantum statistical treatments of optical bistability allowed a proper description of system dynamics, including fluctuations, which are especially relevant near the switching conditions. Purely quantum features could be investigated, such as the photon number distribution, and nonclassical effects could show up in the statistical properties of the transmitted field. The semiclassical treatment corresponds to an approximate approach, at the level of the expectation values of the relevant system operators and where all correlations are neglected.

First of all, the quantum approach showed that, in fact, only one of the two accessible stationary states of the system is absolutely stable, the other one is a metastable state, that is physically accessible only for limited time intervals. Hence, in the bistable parameters range the system switches from one to the other state, and the corresponding switching time is one of the main parameters for practical applications of optical bistable devices.

The quantum theory also showed that the spectrum of transmitted light from an optical bistable system exhibits remarkable features. In the low-transmission branch, for instance, the spectrum is narrower than the natural line, in agreement with the cooperative nature of the emission process. Near the switch-up region, the spectrum is broadened by the onset of large fluctuations. In the high-transmission branch, for large driving field intensity, the spectrum approaches the three-peaked structure characteristic of resonance fluorescence, in agreement with the one-atom nature of system behavior in that limit. Hence, the switching represents a transition between cooperative and non-cooperative radiation processes.

Progress in cavity quantum electrodynamics in the strong coupling regime, where the coherent atom–field interaction dominates the incoherent decay processes, has allowed the experimental investigation of absorptive OB in the limit of a few atoms and photons in a high-finesse cavity, much like it has allowed achieving maser and laser operation with single atoms in microwave and optical resonators, respectively.

In 1991, Rempe, Thompson, Brecha, Lee, and Kimble could observe optical bistability with a number $N \geqslant 15$ of cesium atoms in the cavity, and cooperation parameter $C \geqslant 30$. For smaller numbers of atoms, they reported the breakdown of the semiclassical approximation and the onset of a purely quantum regime, where fluctuations play a basic role, and a single atomic emission has a relevant effect on system dynamics. In this regime, they could observe the occurrence of the effects of photon antibunching and sub-Poissonian photon statistics, as first predicted by Casagrande and Lugiato in 1980.

Semiconductors and Optical Bistable Devices

The first demonstrations of optical bistability raised great hopes in a new all-optical technology based on the transmission, control, and manipulation of light. Optical bistable devices were potentially considered as new kinds of chips for the ultimate goal to control light by light. Actually, they can operate as optical memory elements, transistors, logical gates, and switches, and can perform other basic operations for communications and computation, where photons are used instead of electrons.

After some experiments with solid (ruby) atomic samples and with Kerr liquids in Fabry–Perot resonators, relevant advances were readily achieved in experiments with semiconductors, starting around 1980 from gallium arsenide (GaAs) and indium antimonide (InSb). In bulk GaAs, for instance, the optical Fabry–Perot cavity was reduced to a microcavity, only a few micrometers long, simply consisting of a semiconductor layer with two partially reflecting faces. Under coherent excitation, the material provides strong nonlinearities due to excitonic effects at wavelengths near the bandgap. An exciton is an electron–hole pair bound system. Exciton resonances are responsible for sharp peaks in the absorption edge. The mechanism for absorption saturation is the creation of an increasing number of partially overlapping excitons, such that the Coulomb potential responsible for their bonding is eventually screened and their lifetime shortened by ionization.

Optical bistability with InSb was based on different mechanisms than excitonic absorption. The use of this material allowed the first observation of the effect of optical multistability. Optical bistable devices with semiconductors exhibited quite smaller dimensions, much shorter switching times, and much lower power levels, than with atomic systems. Switch-on times of a few picoseconds were soon reached, though switch-off times were longer due to slower recombination dynamics. Operation at room temperature with remarkably low power was obtained with GaAs super-lattice devices, based on multiple quantum well structures with alternating layers of GaAs and $Al_xGa_{1-x}As$. Many other bulk or low dimensional semiconductor structures, as well as organic materials, liquid crystals, not to mention hybrid electrooptical devices, were tested to implement optical bistable devices.

The dream of all-optical computers, however, slowly faded for the difficulties, with respect to electronic technology, in the realization of high-density arrays of fast optical gates able to operate at low enough switching energies and powers. On the other hand, increasing attention has been paid to the potentiality for applications in optical communications and in specific sectors of computing such as the interconnections.

The subsequent explosive progress in microstructured materials has however opened, again, very promising perspectives to optical bistability and related applications, mainly related to the so-called photonic crystals, composite periodic dielectric materials, where, for example, dielectric rods of high refractive index are embedded in a low refractive index material. Photonic crystals appear as the analog of semiconductors for photons, exhibiting photonic bandgaps like the electronic bandgaps in semiconductor materials. Their properties can be engineered with great flexibility. At present, optical bistable devices based on photonic bandgap microstructures are expected to exhibit quite remarkable properties, such as sizes on the order of a radiation wavelength, switching times shorter than 1 ps, and dissipation of power not larger than a few milliwatts.

The progress in optics and material sciences really seems to renew the expectations raised in the early days of optical bistability, perhaps even as regards all-optical computing.

See also: Devices for Telecommunications; Interaction of Light and Matter; Nonlinear Optics; Photon Statistics and Coherence Theory; Semiconductor Optics.

PACS: 42.65.Pc; 42.65.Sf; 42.50.Fx; 42.50.Nn; 42.70.Nq; 42.70.Ta

Further Reading

Arecchi FT and Harrison RG (eds.) (1987) *Instabilities and Chaos in Quantum Optics*. Berlin: Springer.

Bonifacio R and Lugiato LA (1976) Cooperative effects and bistability for resonance fluorescence. *Optics Communications* 19: 172.

Gibbs HM (1985) *Optical Bistability: Controlling Light with Light*. Orlando: Academic Press.

Gibbs HM, Mc Call SL, and Venkatesan TNC (1976) Differential gain and bistability using a sodium-filled Fabry–Perot interferometer. *Physical Review Letters* 36: 1135.

Grynberg G, Aspect A, and Fabre C (1997) *Introduction aux Lasers et à l'Optique Quantique*. Paris: Ellipses.

Lugiato LA (1984) Theory of optical bistability. *Progress in Optics*, vol. XXI, pp. 71–216. Amsterdam: Elsevier.

Mandel L and Wolf E (1995) *Optical Coherence and Quantum Optics*. Cambridge: Cambridge University Press.

Rempe G, Thompson RJ, Brecha RJ, Lee WD, and Kimble HJ (1991) Optical bistability and photon statistics in cavity quantum electrodynamics. *Physical Review Letters* 67: 1727.

Saleh BEA and Teich MC (1991) *Fundamentals of Photonics*. New York: Wiley.

Szöke A, Daneu V, Goldhar J, and Kurnit NA (1969) Bistable optical element and its applications. *Applied Physics Letters* 11: 376.

Optical Fibers

V Degiorgio and I Cristiani, Università degli Studi di Pavia, Pavia, Italy

Introduction

Immediately after the laser invention, it was realized that laser beams with a carrier frequency of some 10^{14} Hz have the potential to transport information with a much larger frequency band than either radio or microwaves. Atmospheric transmission is limited to line-of-sight communications and is strongly dependent on visibility conditions. It was, therefore, natural to think about a waveguide having the function of creating a protected path for the light beam, and also of compensating diffraction by optical confinement. Indeed, an optical wave having a finite transversal size broadens during propagation in free space. If one considers the output of a typical laser working at the wavelength λ, this is represented by a Gaussian wave with a beam diameter $d \sim 1$ mm. The diffraction angle θ (also called the divergence of the beam) is of the order of λ/d. Taking $\lambda = 1\,\mu$m, it is found that at the distance $L = 1$ km the beam diameter becomes 1 m. At present, guided-wave optics is an important technology with applications that are not limited to the transmission of optical signals, but also involve the fabrication of integrated optical and optoelectronic devices, such as lasers and modulators. The basic concept of optical confinement is simple. A dielectric medium of refractive index n_1, embedded in a medium of lower refractive index n_2, acts as a light trap within which optical rays remain confined by multiple internal reflections at the boundaries. The optical waveguide may have the shape of a slab, strip, or cylinder. The most widely

used is the optical fiber, which is made of two concentric cylinders of glass or polymeric material. Silica glass fibers represent, by far, the most important family of fibers, because they can transmit optical pulses over long distances with small attenuation and limited pulse broadening. The fibers made of polymeric material, called plastic optical fibers (POF), although presenting much larger losses than silica fibers, have the advantage of a much cheaper fabrication process, and can be competitive for the short links required in a variety of applications, including communication networks inside a building or a vehicle.

Structure and Modes

The structure of an optical fiber is shown in **Figure 1**. Typically, the cladding is made of pure silica, whereas the core is doped with germania during manufacturing in order to increase its refractive index and introduce a positive index step between core and cladding. In many cases, instead of having a step in the refractive index profile, the refractive index of the core is graded from a maximum value at its center to a minimum value at the core-cladding boundary. The fiber is then called a graded-index fiber, whereas the fiber shown in **Figure 1** is called a step-index fiber. For optical-communication fibers, the cladding diameter is typically 125 μm, and the core diameter can vary from 9 to 50 μm depending on the application.

An elementary description of light trapping in the core of a step-index fiber can be based on total internal reflection: since $n_1 > n_2$, when the angle of incidence β on the core-cladding interface is larger than the limit angle $\beta_c = \arcsin(n_2/n_1)$, total internal reflection occurs, so that the beam can propagate for an indefinite distance within the fiber core without radiating outside any optical power. If an optical ray whose propagation direction forms an angle α with the axis of the fiber is considered, it is easy to derive that the maximum acceptance angle, α_{max}, is given by

$$\sin \alpha_{max} = n_1 \sqrt{1 - \sin^2 \beta_c} = \sqrt{n_1^2 - n_2^2} \qquad [1]$$

The quantity $\sin \alpha_{max}$ is called numerical aperture of the fiber, and is indicated with the symbol *NA*. Considering a realistic case in which $n_1 = 1.48$, $n_2 = 1.46$, one finds that the numerical aperture is 0.24, and $\alpha_{max} = 14°$.

The geometrical picture of light guidance by multiple reflections is not adequate to explain the propagation properties of the fiber. In particular, geometrical optics suggests a complete confinement

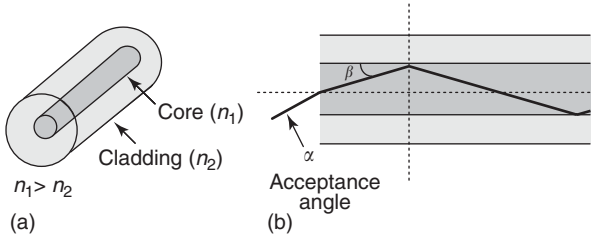


Figure 1 (a) Scheme of an optical fiber; (b) Waveguiding effect by total internal reflection.

of the light beam within the core, whereas wave optics predicts a nonzero intensity in the cladding region (evanescent wave). The general treatment requires solving the wave equation with the appropriate boundary conditions. Since it goes beyond the scope to discuss such a treatment, the article is limited to a description of the main results. Assuming that the electromagnetic wave is monochromatic (i.e., it contains a single frequency ν) and that nonlinear optical effects are negligible, it is found that there is a set of field configurations, called fiber modes, which can propagate without attenuation maintaining the same transversal distribution and the same polarization at all distances along the waveguide axis. In most practical cases, the full set of modes can be approximated by the so-called linearly polarized (LP_{mn}) modes. An LP_{mn} mode has n field maxima along a radius vector and $2m$ field maxima round a circumference. By calling E_{mn} the electric field associated to the mode mn, where the indices m and n are integer numbers, and choosing the z-axis coincident with the fiber axis, the fiber mode is characterized by a transversal field distribution $A_{mn}(x, y)$ and a propagation constant β_{mn}, both depending on m and n:

$$E_{mn}(x, y, z, t) = A_{mn}(x, y) e^{i(\beta_{mn} z - \omega t)} \qquad [2]$$

where $\omega = 2\pi\nu$ is the angular frequency. **Figure 2** shows the amplitude profile of the fundamental mode LP_{01} that is approximately described by a Gaussian function:

$$A_{01}(r) = A_o \exp\left(-\frac{r^2}{w^2}\right) \qquad [3]$$

where $r = (x^2 + y^2)^{1/2}$ is the distance from the fiber axis. It is seen that the field amplitude $A_{01}(x,y)$ is mainly concentrated inside the core, but presents exponentially decaying tails extending into the cladding region. Higher-order modes present more complicated profiles extending more and more deeply into the cladding as the indices m and n become larger and larger.

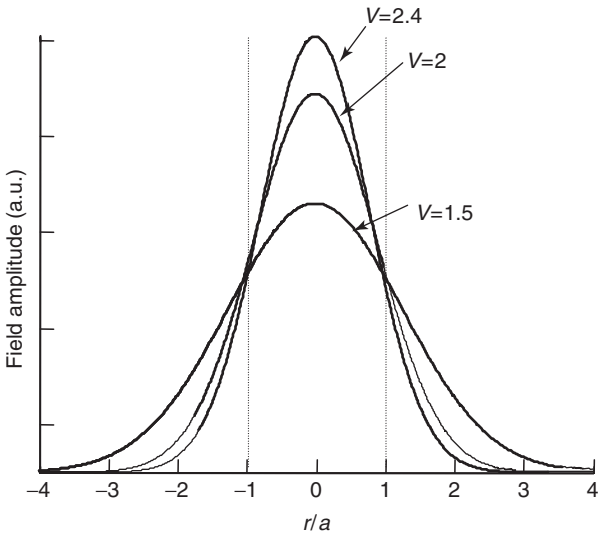


Figure 2 Field profile of the fundamental propagation mode LP_{01} for various values of the normalized frequency V.

In order to discuss fiber properties, it is helpful to introduce the dimensionless parameter V, called normalized frequency, defined as

$$V = \frac{2\pi a}{\lambda}\sqrt{n_1^2 - n_2^2} \qquad [4]$$

where $\lambda = c/v$ is the wavelength. The plots in **Figure 2** indicate that the beam size w is a decreasing function of V. This can be easily understood if one notes that the beam is more tightly confined when the core size is large in comparison with the wavelength, and when the difference in refractive index between core and cladding is large.

From the propagation constant β_{mn}, one can derive the effective index of refraction n_{mn} seen by the mode mn, by using the relation: $\beta_{mn} = 2\pi n_{mn}/\lambda$. The value of n_{mn} will be clearly intermediate between n_1 and n_2, approaching n_1 more closely when the field mode is more tightly confined.

A very important general property of waveguides is that, for a fixed waveguide configuration and a fixed λ, a fundamental mode always exists, whereas higher-order modes can be sustained only under specific conditions. In the case of optical fibers, the parameter controlling the single-mode or multimode behavior is the normalized frequency V. If V is sufficiently small, the fiber is single-mode. When V becomes larger than a critical value V_c, the second mode can also be sustained. By progressively increasing V, more and more modes can be guided. For $V \gg 1$, the number N of guided modes is given by: $N \cong V^2/2$.

The value of V_c is dependent on the fiber structure. In the case of the step-index fiber, $V_c = 2.405$. Fibers that are single-mode at $\lambda = 1500\,\text{nm}$ are usually designed with values of V close to 2, by choosing a core radius $\sim 4.5\,\mu\text{m}$ and a difference $n_1 - n_2 \sim 5 \times 10^{-3}$. It is to be noted that, since V is inversely proportional to λ, a fiber that is single-mode in the near infrared will be multimode in the visible.

Multimode fibers have typically a core diameter $\sim 50\,\mu\text{m}$. They are easier to align and less expensive than single-mode ones. However, as discussed later in the section "Dispersion," they are not fit for pulse transmission over long distances because pulse-broadening effects are large.

Attenuation

In principle, a guided mode can propagate in an optical fiber for arbitrary distances without suffering radiation losses. In practice, however, the propagating wave is always attenuated. Assuming that the losses are linear, the optical power P_o launched into the fiber decays exponentially along the fiber following the law:

$$P(z) = P_o e^{-\alpha_o z} \qquad [5]$$

where α_o, expressed in km^{-1}, is the attenuation coefficient. The quantity normally given in fiber specifications is the coefficient $\alpha = 4.34\ \alpha_o$ expressing the attenuation in dB km^{-1}.

The attenuation is due to two phenomena, light scattering and absorption. Light scattering is originated by the microscopic inhomogeneities that exist in the glass at the drawing temperature because of thermal fluctuations. Such inhomogeneities are frozen during the rapid cooling process, and give rise to spatial fluctuations of the refractive index over a submicrometric scale. Rayleigh scattering from these fluctuations has a cross section that scales according to λ^{-4}. As a consequence, scattering losses rapidly decrease as a function of wavelength.

Optical fibers are usually made of fused silica glass (SiO_2). This material has two strong absorption bands, an ultraviolet band $\sim 0.3\,\mu\text{m}$ due to electronic transitions and an infrared band $\sim 1.7\,\mu\text{m}$ due to the vibrational mode of the Si–O bond. The presence of water impurities gives other absorption peaks due to vibrations of the OH^- ion, the most important one near 1370 nm. Because of the combined effect of scattering and absorption, the attenuation curve for silica glass, as shown in **Figure 3**, has an absolute minimum at 1550 nm and a secondary minimum $\sim 1300\,\text{nm}$. The earliest optical communication schemes were working at $\sim 800\,\text{nm}$, where the attenuation has a local minimum, not apparent in **Figure 3**. The value of the attenuation at the absolute

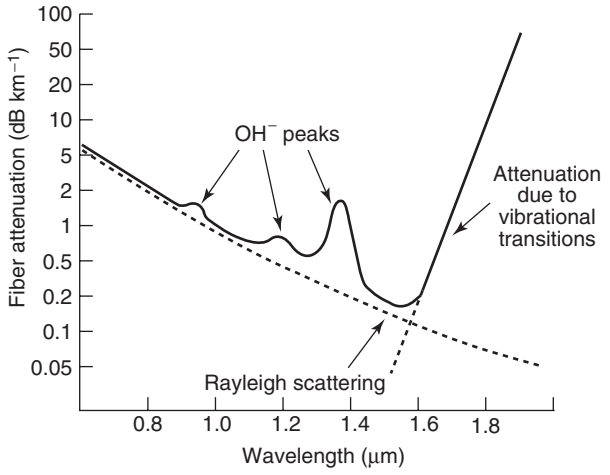


Figure 3 Dependence of the attenuation coefficient α of silica glass on the wavelength λ.

minimum, α_{min}, is $\sim 0.15\,dB\,km^{-1}$, corresponding to a reduction of the input power by a factor of 2 after a propagation length of 20 km. Optical communication systems operate at the wavelength corresponding to the minimum attenuation.

If different oxides are added to SiO_2, the position of α_{min} is only slightly changed because essentially all oxides present vibrational transitions in the range 1.5–2 μm. A significant shift of the minimum attenuation can be obtained by using glasses made with mixtures of fluorides, such as the so-called ZBLAN glasses that contain Zr, Ba, La, Al, and Na fluorides. Since vibrational levels of fluorides are localized in the range 5–8 μm, ZBLAN glasses present the attenuation minimum at $\sim 3\,\mu m$, with a value of α_{min} one order of magnitude smaller that of silica glasses. It is, however, rather difficult to draw ZBLAN fibers of good optical quality.

It should be mentioned that fiber losses increase considerably if the fiber is curved. This can be easily understood in the frame of geometrical optics: if the fiber is curved, there is a change of the incidence angle at the core-cladding interface and some rays may not be totally reflected. Such a radiation loss affects, in particular, the modes presenting a field distribution that extends considerably into the cladding, as it happens when the normalized frequency V is small.

Dispersion

In optics, the term dispersion indicates that the index of refraction of the medium changes with the frequency of the propagating wave, $n = n(\omega)$. An optical pulse with duration Δt has a spectral broadening $\Delta\omega \geqslant \Delta t^{-1}$, and cannot be treated as a monochromatic wave. Each frequency component travels in the dispersive medium with a different phase velocity $v_\phi = c/n(\omega)$. As a consequence, the pulse shape changes during propagation. If it is assumed that the input pulse is transform-limited, which means that all the frequency components have the same phase, the effect of dispersion is a progressive increase of the pulse duration as the propagation length increases.

If the wave propagates along an optical fiber, its velocity is determined by the effective index of refraction n_{mn} defined above. It is to be noted that, even in the absence of material dispersion, the effective index of refraction changes with frequency because the mode profile is frequency-dependent. Therefore, the fiber is dispersive even in the absence of dispersion of the materials constituting the fiber. This results in a waveguide contribution that must be added to the material contribution to obtain the total fiber dispersion.

Laser pulses have a spectral broadening $\Delta\omega$, normally much smaller than the central frequency ω_0. This allows one to treat the wave propagation in a dispersive medium by using a perturbative approach. The propagation constant $\beta(\omega)$ is expanded in a power series around ω_o, and the series is truncated to the quadratic term:

$$\beta(\omega) = \beta_o + \beta_1(\omega - \omega_o) + \tfrac{1}{2}\beta_2(\omega - \omega_o)^2 \qquad [6]$$

where β_j is the jth derivative of $\beta(\omega)$ evaluated at $\omega = \omega_o$. By using eqn [6] inside the propagation equation, it can be shown that the reciprocal of β_1 is the group velocity (i.e., the velocity of the pulse envelope) and that β_2 is the parameter responsible for pulse broadening, usually called group-velocity dispersion (GVD). If it is assumed that the input pulse is a transform-limited Gaussian pulse having full-width at half maximum equal to Δt_o, the pulse duration Δt after a path length z is given by

$$\Delta t = \sqrt{(\Delta t_o)^2 + \left(\frac{|\beta_2|z}{\Delta t_o}\right)^2} = \Delta t_o\sqrt{1 + \left(\frac{z}{L_D}\right)^2} \qquad [7]$$

where $L_D = (\Delta t_o)^2/|\beta_2|$ is called the dispersion length. Equation [7] indicates that dispersion effects become more and more important as the pulse duration becomes shorter and shorter. Considering, for instance, a standard single-mode optical-communication fiber at 1550 nm, $\beta_2 = -20\,ps^2\,km^{-1}$, and $\Delta t_o = 10\,ps$, one finds $L_D = 5$ km. This means that the pulse duration is more than doubled when $z = 10$ km. Clearly, pulse broadening can be very detrimental in long communication links because it produces a smearing out of individual bits, leading to higher bit error rates. In order to reduce such an effect, fibers with low value of

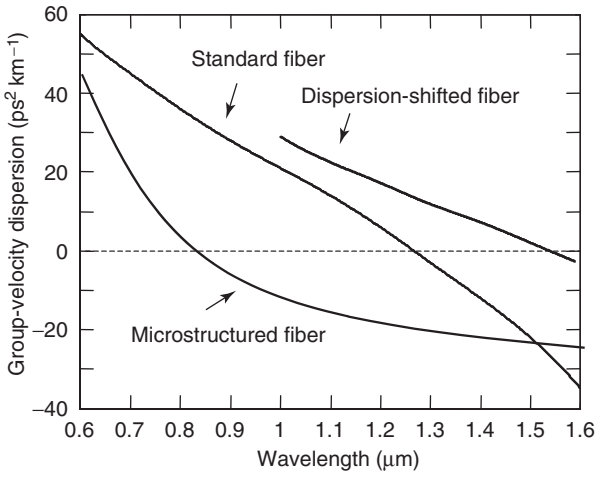


Figure 4 The dispersion parameter β_2 plotted as a function of wavelength for a standard single-mode fiber, a dispersion-shifted fiber, and a microstructured fiber.

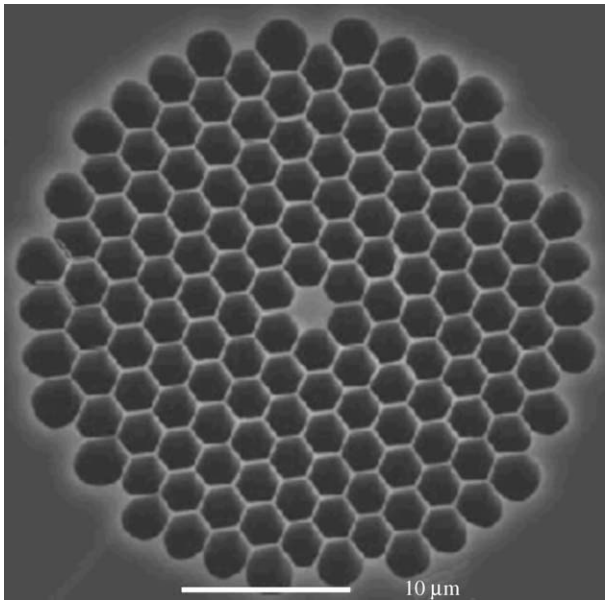


Figure 5 Cross section of a microstructured fiber made of silica with air holes running along the fiber axis. The silica core has a diameter of $\sim 2\,\mu$m.

β_2 have to be used. These special fibers, called dispersion-shifted fibers, are fabricated by an appropriate choice of refractive-index profile. **Figure 4** shows the behavior of β_2 as a function of λ for a standard single-mode fiber and for a dispersion-shifted fiber.

Multimode fibers present an additional dispersion mechanism due to the fact that different modes, even if oscillating at the same frequency, have different field profiles, hence different effective refractive indices. Therefore, if the pulse input power is coupled simultaneously to several modes, modal dispersion gives a contribution to pulse broadening that can be larger than that due to material and waveguide dispersion.

It should be noted that in this section, the refractive index has been considered as a real quantity, whereas the presence of attenuation would imply that the refractive index, hence the propagation constant, is a complex quantity.

Special Fibers

Besides step-index and graded-index fibers, other types of fibers having a more complicated index profile are used for special purposes. The previous section mentioned dispersion-shifted fibers, designed to present a reduced dispersion in a specific wavelength range. Another useful category is that of polarization-maintaining fibers, fibers presenting some anisotropy in the index profile that makes them birefringent. A very recent development consists in fibers presenting air holes running along the fiber axis. An example is shown in **Figure 5**. These are called microstructured fibers, and may have a variety of applications connected with the fact that the air–silica structure allows single-mode behavior over a very wide wavelength range and gives more freedom in choosing the area of the propagating mode or the shape of the dispersion curve.

Fiber Bragg Gratings

It is very helpful to be able to manipulate optical signals traveling inside fibers without extracting them from the fiber. This not only reduces losses, but also makes the optical system more flexible. Following the all-fiber approach, various optical components and devices have been conceived and built in order to perform a variety of functions, such as beam splitters and combiners, mirrors, filters, polarizers, amplifiers, and even lasers. This section describes frequency-selective mirrors obtained by introducing inside the fiber core, a modulation of the refractive index with a given periodicity. Such a modulation acts as a phase grating and the structure is indeed called "fiber Bragg grating" (FBG). The behavior can be understood by similarity with a multilayer dielectric mirror: a sequence of weakly reflecting surfaces can become a high-reflectivity structure if a constructive interference condition is established among all the reflected fields. Calling Λ the grating period, such a condition requires that the phase shift suffered by the field propagating over the distance Λ be π (or a multiple of π). The condition is satisfied for the wavelength λ_o given by

$$\lambda_o = 2n\Lambda \qquad [8]$$

where n is the average refractive index of the fiber core. If Δn is called the amplitude of the refractive-index modulation and N the number of grating periods, the peak reflectivity is given by

$$R_{\max} = \tanh^2\left(N\frac{\Delta n}{n}\right) \qquad [9]$$

Typically the grating length L is of the order of millimeters, that is, $N = L/\Lambda$ is in the range 10^3–10^4. Therefore, $R_{\max}$ can be close to 100%, even if Δn is small. The bandwidth of the FBG mirror is a decreasing function of Δn, so that the reflectivity peak can become very narrow for small Δn.

FBGs are fabricated by exploiting the photosensitivity of the Ge-doped core of silica fibers. The presence of Ge atoms leads to the formation of oxygen-deficient bonds that represent defects in the silica matrix. The energy required to break the bond is ~ 5 eV. The absorption of a 244 nm photon from an excimer laser can break the defect-bond, resulting in a change of absorption spectrum and in a corresponding change in the refractive index. Since index changes occur only in the regions where ultraviolet light is absorbed, by illuminating the fiber core with a periodic intensity pattern, an index-of-refraction grating is generated.

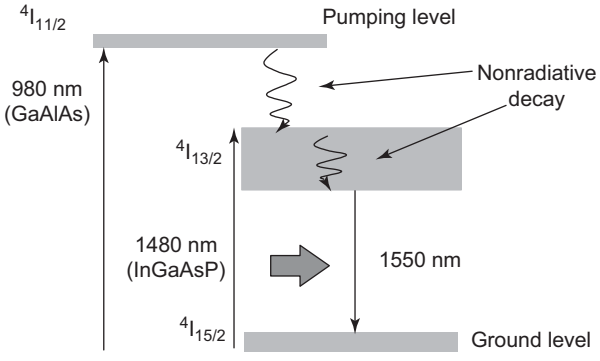


Figure 6 Energy levels of the Er^{3+} ion.

Fiber Amplifiers and Lasers

An important family of materials used for optical amplification is represented by dispersions of rare-earth ions, such as neodymium or erbium or ytterbium, in a solid matrix that can consist of a crystal or a glass. An interesting extension of the idea of the rare-earth solid-state optical amplifier is the fiber amplifier, which utilizes silica fibers doped with rare-earth ions. Optical amplification is extremely important in long-distance optical-communication links in order to compensate for fiber attenuation, so that the optical power can be maintained at sufficiently high levels along the path to ensure a satisfactory signal-to-noise ratio. Doped fibers can provide wide-band amplification at wavelengths ~ 1550 nm, corresponding to the attenuation minimum for silica fibers.

The most common fiber amplifier is based on erbium-doped silica fibers and uses a transition of the Er^{3+} ion in a wavelength range extending from 1530 to 1560 nm. The energy level diagram of the ion Er^{3+} is shown in **Figure 6**. The upper level of the 1530 nm transition is metastable (with a lifetime of 12 ms), thus permitting one to easily achieve population inversion by exciting the ion with radiation at 980 or 1480 nm. The pump field comes from a semiconductor laser, either InGaAs (980 nm) or InGaAsP (1480 nm). A scheme of the amplifier configuration is shown in **Figure 7**. The pump field and the signal field are injected into the single-mode doped fiber by using a dichroic coupler. After the doped fiber, an optical isolator is inserted in order to eliminate back-reflections. The length of the amplifying fiber is typically ~ 10 m, and the optical gain can be as large as 10 000. The efficiency of the amplifier is large

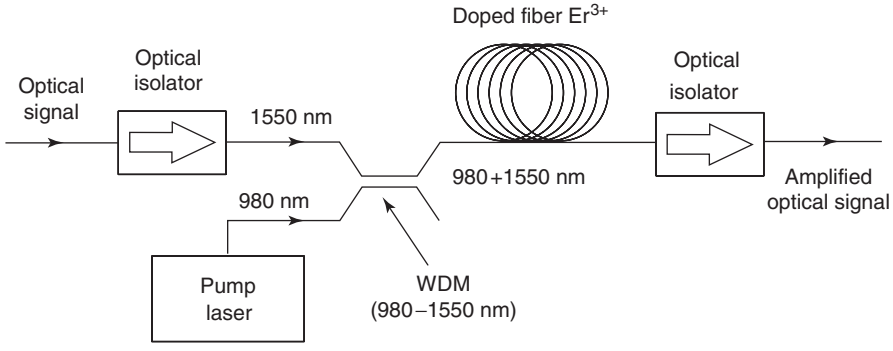


Figure 7 Scheme of an Erbium-doped fiber amplifier, pumped by a semiconductor laser. WDM (Wavelength Division Multiplexer) indicates a wavelength dependent combiner.

because the single-mode fiber ensures a very good spatial overlap between pump field and signal field. The conversion efficiency can be further improved by using a mixed erbium–ytterbium doping.

Any optical amplifier can be converted into a laser oscillator if a positive feedback scheme is provided. The fiber amplifier can become a fiber laser by placing the amplifier in between two mirrors, constituted, for instance, by two FBGs. Fiber lasers can work either continuously or pulsed, with a large power output. Since the fiber is very thin, the surface-to-volume ratio is large, thus permitting very efficient heat dissipation without using the water-circulation cooling system typical of power lasers. In order to improve the coupling between the semiconductor pump laser and the amplifier, a new fiber geometry involving a double cladding was introduced. A very successful fiber laser using such a geometry is the ytterbium laser, which is pumped at 920 nm and can emit in a wide band centered $\sim$1100 nm. Continuous-wave Yb fiber lasers with an output power larger than 100 W and efficiency of 75–80% are now available.

Nonlinear Fiber Optics

Nonlinear optical effects are particularly efficient in single-mode optical fibers despite the relatively small third-order susceptibility of silica glass. The reason for this is that the spot size of the propagating beam is small and the interaction length can be very long due to waveguiding and low losses. A point of great conceptual relevance is that most of the treatments of nonlinear propagation are developed for plane waves, whereas experiments in bulk materials are performed with laser beams that can be considered as Gaussian waves. Therefore, diffraction can play a significant role in the experiments, making the comparison with the plane-wave theory sometimes difficult. The nonlinear propagation inside the optical fiber can be treated in most cases by assuming that the transversal field distribution is invariant along the propagation distance, so that the propagation problem becomes spatially unidimensional as in the case of the plane-wave approach. This makes the comparison between theory and experiment much simpler.

The two most important nonlinear phenomena in nonlinear fiber optics are the optical Kerr effect and stimulated Raman scattering (SRS).

The optical Kerr effect, that is, the linear dependence of the index of refraction on the beam intensity, gives rise to a self-induced phase shift experienced by the optical field during its propagation, called self-phase modulation (SPM). SPM is responsible for the spectral broadening of optical pulses. In normal dispersion, the consequence is that SPM gives an additional contribution to the spreading out in time of the pulse. There is, however, an interesting situation, corresponding to positive Kerr effect and negative GVD, in which there is a compensation between the two effects, so that an optical pulse of appropriate intensity and shape (optical soliton) can propagate with constant shape and amplitude over a distance that is limited in practice only by the linear attenuation in the fiber. Indeed, several beautiful experiments showing the onset and development of solitonic propagation have been performed. With standard fibers, the wavelength range of negative GVD is above 1.3 μm, but the new fibers presenting an air–silica microstructure (see the section "Special fibers") can shift the negative GVD region to lower wavelengths, as shown in **Figure 4**, so that solitonic propagation in the visible can be observed.

Concerning SRS, it is recalled that, starting from a pump wave at frequency ν_p, the effect of SRS is that of generating a strong signal, called the Stokes wave, at a frequency $\nu_s = \nu_p - \nu_v$, where ν_v is a vibrational frequency typical of the considered medium. In the case of silica fibers, the main effect is associated with vibrations of the Si–O bond, occurring at $\nu_v =$ 12 THz. SRS can be used to amplify a weak signal at frequency ν_s in presence of a pump ν_p. Raman amplification is very attractive for applications in optical communications because, at variance with what happens in erbium-doped fiber amplifiers, it is nonresonant, so that it can be used for any ν_s, provided the appropriate pump is available.

See also: Disordered Solids and Glasses, Electronic Structure of; Laser Radiation Sources; Nanostructures, Electronic Structure of; Nonlinear Optics; Semiconductor Devices.

PACS: 42.81.–i; 42.82.–m; 42.25.Bs; 42.79.Gn; 42.55.Wd; 42.65.Dr; 42.65.Jx; 42.65.Tg

Further Reading

Agrawal GP (1995) *Nonlinear Fiber Optics*, 2nd edn. New York: Academic Press.

Concerning microstructured fibers, see, for example, several articles in the special issue of 2003. *Applied Physics B* 77(2–3).

Cregan RF, *et al.* (1999) Single mode photonic bandgap guidance of light in air. *Science* 285: 1537–1539.

Saleh BEA and Teich MC (1991) *Fundamentals of Photonics*. New York: Wiley.

Wilson J and Hawkes J (1998) *Optoelectronics: An Introduction*. Harlow: Prentice Hall.

Optical Instruments

M de Angelis, Istituto di Cibernetica CNR, Pozzuoli (NA), Italy
G M Tino, Dipartimento di Fisica dell' Università di Firenze, Sesto Fiorentino (FI), Italy

Introduction

This article describes some complex devices which are commonly used in the laboratory for analysis of condensed matter, instruments which measure the fundamental quantity of light, and use the interaction between light radiation and matter. It is assumed, that the reader is confident with simple optical instruments such as lenses, mirrors, prisms, and diffraction gratings. A complete description of the fundamental optical phenomena is given elsewhere in this encyclopedia and can be easily found in literature on fundamental optics. Before describing optical instruments in detail, a brief overview of light behavior is given. Among the instruments commonly used in the laboratory, the spectrometer or devices based on scattering, absorption, and reflection are widely described in other sections of this encyclopedia. Also interferometric methods used in microscopy are illustrated elsewhere. Here, the focus is on the following instruments: refractometer, interferometer, polarimeter, ellipsometer, and radiometer.

Optical Terminology and Fundamental Optical Phenomena

Light is an electromagnetic wave with wavelengths between 0.1 nm and 100 μm. Visible light corresponds to wavelength in the range of 400–700 nm, vacuum ultraviolet light is in the range 0.1 μm to 10 nm; at the other end of the wavelength scale between 100 and 1000 μm is the microwave region. **Figure 1** shows electromagnetic wavelength regions and **Table 1** shows important parameters that are used to characterize light and the media through which it passes. At this point, a few comments are necessary. First of all in free space, light behaves like a transverse wave which means that the variations of the electric vector $\boldsymbol{E}$, are perpendicular to the propagation direction. In electromagnetic waves, the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ are perpendicular to each other and to the propagation vector $\boldsymbol{k}$ (**Figure 2**). The simplest case of a three-dimensional wave is the plane wave in which all the surfaces upon which a disturbance has a constant phase form a set

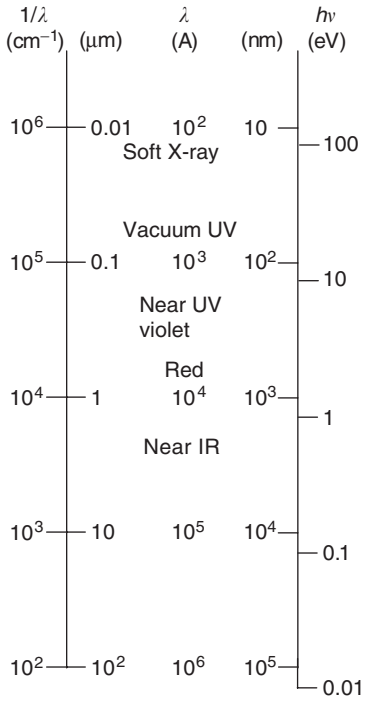


Figure 1 Diagram relating wave number, photon energy, and wavelength in light radiation.

of planes, each generally perpendicular to the propagation direction. There are quite practical reasons for studying this sort of wave, one of which is that by using optical devices, one can readily produce light resembling plane waves, and this case is taken up in detail in the following.

The mathematical expression for a plane wave that is perpendicular to a given vector $\boldsymbol{k}$ in complex representation is

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}\pm\omega t)} \qquad [1]$$

The vector $\boldsymbol{k}$ whose magnitude is the propagation number k is called the propagation vector, and it is linked to the wave quantity by the following useful relations:

$$k = \frac{2\pi}{\lambda} = \frac{2\pi n}{\lambda_0} \qquad [2]$$

where λ is the wavelength, and n is the refractive index of light. The speed of light is

$$v = \lambda\nu = n\lambda_0\nu = nc \qquad [3]$$

where c is the velocity of light in vacuum

$$c = 2.997\,924\,58 \times 10^8 \mathrm{m\,s^{-1}} \qquad [4]$$

Table 1 Fundamental parameters of electromagnetic radiation and optical media

Parameter	*Symbol*	*Value*	*Units*
Velocity of light in vacuo	$c_0 = \sqrt{\mu_0 \varepsilon_0}$	2.997 924 58 x 10^8	$m\,s^{-1}$
Permeability of free space	μ_0	4π x 10^{-7}	henry m^{-1}
Permittivity of free space	ε_0	8.854 16 x 10^{-12}	farad m^{-1}
Velocity of light in a medium	$c = \sqrt{\mu_r \mu_0 \varepsilon_r \varepsilon_0} = c_0/n$		$m\,s^{-1}$
Refractive index	$n = \sqrt{\mu_r \varepsilon_r}$		dimensionless
Relative permeability of a medium	μ_r	usually 1	dimensionless
Dielectric constant of a medium	ε_r		dimensionless
Frequency	$\nu = c/\lambda$		Hz
Wavelength in vacuo	$\lambda_0 = c_0 \nu$		m
Wavelength in a medium	$\lambda = c\nu = \lambda_0/n$		m
Wave number	$\bar{\nu} = 1/\lambda$		cm^{-1}
Wave vector	$k = 2\pi/\lambda$		m^{-1}
Photon energy	$E = h\nu$		J (1.6×10^{-19} J = 1 eV)
Electric field of wave	$\boldsymbol{E}$		$V\,m^{-1}$
Magnetic field of wave	$\boldsymbol{B}$		$A\,m^{-1}$
Poynting vector	$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{B}$		$W\,m^{-2}$
Intensity	$I = \langle \boldsymbol{S} \rangle_{av} = \lvert E \rvert^2 \varepsilon_0 c^2$		$W\,m^{-2}$

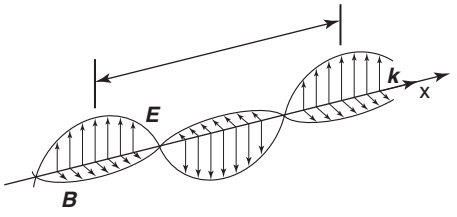


Figure 2 Picture of a plane polarized light. For plane polarized light, the ***E*** and ***B*** vector remain in perpendicular planes to the propagation vector ***k***.

and where the wave frequency

$$v = \omega/2\pi \qquad [5]$$

It is to be noticed that at any given time the surfaces joining all points of equal phase are known as wave fronts.

The velocity of light in vacuum is $c = \sqrt{\mu_0 \varepsilon_0}$; in a medium, it depends both on the relative magnetic permeability μ_r and dielectric constant ε_r of the medium, $\mu_r = 1$ for all practical optical materials so that the refractive index and the dielectric constant are related by $n = \sqrt{\varepsilon_r}$. When light propagates in a non-isotropic medium, such as a crystal, n and ε_r will generally depend on the direction of the wave propagation and its polarization state.

The Poynting vector

$$\boldsymbol{S} = c^2 \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B} \qquad [6]$$

is a vector that points in the direction of energy propagation of the wave. The local direction of the Poynting vector at a point in a medium is called the ray direction. The time-averaged value of the magnitude of the Poynting vector is the irradiance and is given by

$$I = \langle S \rangle = \varepsilon_0 c \langle E^2 \rangle \qquad [7]$$

where ε_0 is the permittivity of free space, $8.8542 \times 10^{-12}\,F\,m^{-1}$.

Since light in vacuum is a transverse wave, an appreciation of its vectorial nature is of great importance, which means that all phenomena of optical polarization should be treated in terms of a vector wave picture. Before going into details of optical devices, two things should be noted. One is that electromagnetic waves propagating in real material media are generally not transverse. Another is that for unpolarized light, in which the wave vector changes direction randomly and rapidly, or in some other cases, scalar approximation becomes useful, as in the theory of interference and hence the simpler scalar representation of the light wave is considered wherever possible.

The parameter used to describe the interaction of light with the material is the complex index of refraction $\tilde{n}$, which is a combination of a real part and an imaginary part and is given by

$$\tilde{n} = n + i\kappa \qquad [8]$$

The real part is called the index of refraction and the imaginary part the extinction coefficient. For a dielectric material such as glass, none of the light is absorbed and $\kappa = 0$; in this case the concern is only n, as previously defined. In an absorbing medium, κ describes the attenuation of the wave field in the

medium and it is related to the absorption linear coefficient. A complex index of refraction is a characteristic of an absorbing medium.

Real and imaginary parts of the index of refraction are, in fact, functions of the wavelength λ; this is why when white light enters a prism, it emerges with the various colors separated. This phenomenon is called dispersion and is described by the Kramers–Kronig equations for both the real and imaginary part of the complex index of refraction or by an approximated expression by power series of $1 = \lambda$ using Cauchy coefficient. For optical materials, the change of index with wavelength is very gradual, and often negligible, unless the wavelength approaches a region where the material is not transparent. For this reason, in this article the description of optical systems is simplified assuming that the light is monochromatic. When light is reflected by a plane boundary between two media of different refractive index, the three rays, that is, incident, reflected, and refracted lie in the same plane and the angle of incidence is always equal to the angle of reflection as shown in **Figure 3**. When a light ray crosses the boundary, the angle of refraction θ_2 is related to the angle of incidence accordingly to Snell's law

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1} \qquad [9]$$

If $n_2 < n_1$, there is a maximum angle of incidence for which there can be a refracted wave because $\sin\theta_2$ cannot be greater than unity for a real angle. This is called the critical angle θ_c and is given by $\sin\theta_c = n_2/n_1$. Total internal reflection is illustrated in **Figure 3** when θ_1 exceeds θ_c and the boundary acts as a very good mirror.

One or both the media in **Figure 3** may be anisotropic. The incident wave will then generally split into two components, and assist a double-refraction phenomenon. One of the components obeys Snell's law directly (the ordinary wave) while for the other component, Snell's law is modified (the extraordinary wave).

Looking at the wave field at a boundary between two media, it can be seen that there is still an interdependence sheared by the amplitudes which can be evaluated and which depends on polarization as in a case in which E is perpendicular as well as parallel to the plane of incidence. There are two expressions that are general statements that apply to all linear, isotropic, and homogeneous media; they are called Fresnel equations. For electric field of the wave field perpendicular to the plane of incidence (S polarization)

$$r_s = \left(\frac{E_{0r}}{E_{0i}}\right)_s = \frac{n_i\cos\theta_i - n_t\cos\theta_t}{n_i\cos\theta_i + n_t\cos\theta_t} \qquad [10]$$

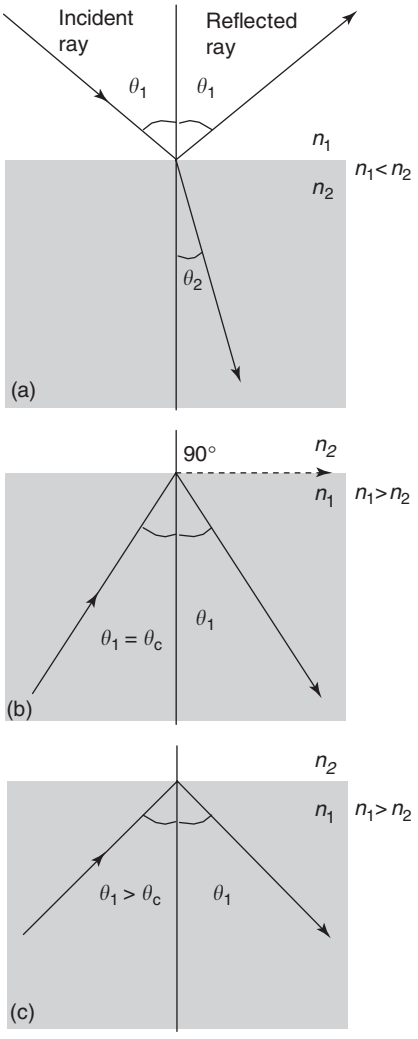


Figure 3 (a) Reflection and refraction of a light ray at the boundary between two different isotropic media of refractive index n_1 and n_2. The case shown is for $n_2 > n_1$. (b) Critical angle and (c) total internal reflection for $n_1 > n_2$.

$$t_s = \left(\frac{E_{0t}}{E_{0i}}\right)_s = \frac{2n_i\cos\theta_i}{n_i\cos\theta_i + n_t\cos\theta_t} \qquad [11]$$

Here r_s denotes the amplitude reflection coefficient and t_s is the amplitude transmission coefficient.

For electric field of the wave field parallel to the plane of incidence (P polarization)

$$r_p = \left(\frac{E_{0r}}{E_{0i}}\right)_p = \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_i\cos\theta_t + n_t\cos\theta_i} \qquad [12]$$

$$t_p = \left(\frac{E_{0t}}{E_{0i}}\right)_p = \frac{2n_i\cos\theta_i}{n_i\cos\theta_t - n_t\cos\theta_i} \qquad [13]$$

When the reflecting surface is not a dielectric and κ is nonzero, the index of refraction in the Fresnel equations is the complex index of refraction and the Fresnel equations are in general complex number.

Refractometer

The index of refraction is a topical parameter for isotropic material. Most of the refractive index measurements are based on "prism refraction." Since the angle of deviation between the incident beam and the refracted beam from a prism varies with the angle of the prism, with the angle of incidence of the ray on the prism, and with the refractive index of the prism, measuring angles involved in the process gives the value of the refractive index of the prism. The deviation of a ray of light passing through a simple prism can be described with the aid of **Figure 4**. The various angles α, i, r, and θ satisfy Snell's law irrespective of the polarization state of the input beam, provided the prism is made of an isotropic material. The deviation θ of a ray depends on the incidence angle i and increases as the angle of incidence is increased. From the symmetry of the system, it follows that the deviation is minimum for the ray that passes through the prism symmetrically. When the prism is used in the position of minimum deviation

$$i = \tfrac{1}{2}\alpha + \tfrac{1}{2}\theta \qquad [14]$$

Hence, in the minimum deviation condition, the prism index of refraction is

$$n = \frac{\sin(\alpha + \theta)/2}{\sin \alpha/2} \qquad [15]$$

When θ and α have been measured, the refractive index of the prism material can be calculated. If the angles are measured to about one second, n can be found to $\sim \pm 0.000\,01$.

The configuration can be used to determine the refractive index of a liquid by employing a hollow prism whose faces are parallel-sided plates of glass. The walls of such a vessel have no effect on the angle of deviation, and the angle of the liquid prism is the same as that of the vessel. Hence, α and θ can be measured in the usual way. It is important to point out that the refractive index of liquid varies rapidly with temperature, and there is usually no point in attempting an accuracy better than 0.0001 in the refractive index unless special care is given to temperature control.

Widely used is the "critical angle prism refractometer" which uses a critical angle configuration: the idea is to look at the light totally reflected by the internal prism wall. An alternative method is to position the incident beam so that rays fall on the prism at grazing or near-grazing incidence. The rays falling on the face BA of **Figure 4b** at grazing incidence emerge from face AC at angle ϕ, and those rays incident on BA at angles less than 90° emerge from AC at an angle greater than ϕ. No light emerges at angles less than ϕ. When a detector is placed in front of the AC plane, with the arrangement shown in the figure, the right-hand side of the field view is dark and the left-hand side of the field view is bright; the edge between the bright and dark zones is used to measure ϕ. In this configuration

$$n^2 = 1 + \left(\frac{\sin\phi + \cos\alpha}{\sin\alpha}\right)^2 \qquad [16]$$

gives the refractive index of the prism material.

The same method may be employed to determine the refractive index of a liquid of which only a small

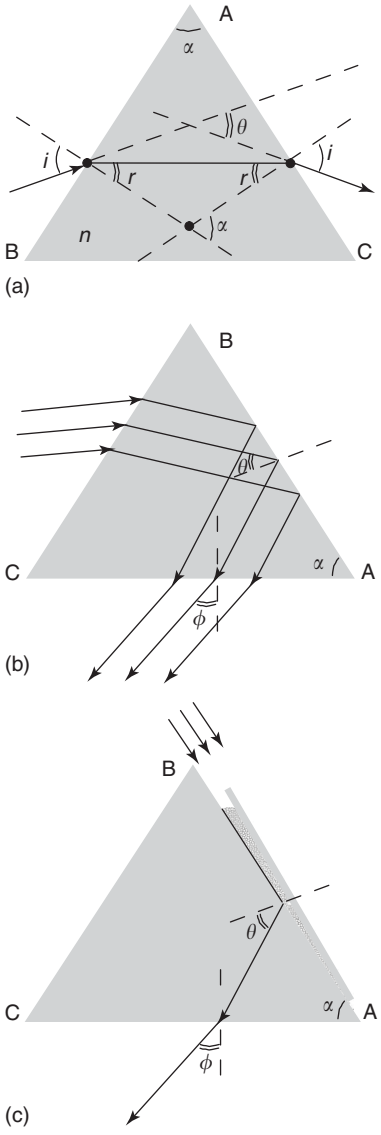


Figure 4 (a) Symmetrical path of a ray through a prism. (b) Critical angle prism refractometer with light at grazing incidence. (c) Critical angle method for liquid sample.

quantity is available, provided the index of the prism is greater than that of the liquid. A drop of the liquid is placed on the face BA and a thin glass plate is placed over it. If necessary, it is possible to have a wedge form of the liquid to be examined, and the glass is held in position by a capillary action. In this case, the refraction index of the liquid is

$$n_0 = \sin\alpha\sqrt{n^2 - \sin^2\phi} - \cos\alpha\sin\phi \qquad [17]$$

It should be noted that the maximum possible value of r is the critical angle, and that there is a grazing incidence and a grazing emergence. Hence, the angle of a prism used in the upper configuration must not be greater than twice the critical angle, otherwise it is impossible for rays to be transmitted through the prism.

A two-prism configuration with the critical-angle method is used in the Abbe refractometer, shown in **Figure 5a**. The liquid under test is placed between two prisms of high refractive index. The prism P corresponds to the prism to be used in the total reflection configuration. Light that strikes P in grazing incidence emerges at the angle ϕ to be measured. The measurement of the emerging angle would enable one to determine the refractive index of the liquid, if α and the refractive index n of P are known.

Critical-angle based refractometers have three main disadvantages: (1) they measure light rays coming from the prism on a field view where bright and dark areas have to be separated and measured, which is a basically asymmetric system, (2) the refractive index of the glass prism must be greater than that of the specimen under test, and (3) the measured index corresponds to a surface layer and may differ from that of the main bulk of the specimen. The Hilger–Chance instrument (**Figure 5b**) was developed to overcome these difficulties. The composite glass block is made by combining two prisms as indicated. Monochromatic collimated light passes through the V-block and specimen. The direction of the emergent light depends upon the refractive index of the specimen. It will be seen that two perpendicular plane surfaces must be produced on the specimen, but it is not necessary for these to be of the highest quality since a film of liquid can be placed between the faces of the specimen and the V-block. Since the method is not a critical-angle method, the refractive index of the specimen can be greater than or less than that of the V-block. The accuracy found with this instrument is 0.000 01. The instrument can be used for liquids if the V-block is provided with side plates to form a trough, but 2–3 cc of liquid are required.

Finally, accurate measurements of the index of refraction up to 0.000 001 can be done using interferometry which measures the optical path difference (geometrical difference per index of refraction), and hence provides the index of refraction of the sample under test if the geometric path (usually thickness of the sample) is well known. The interferometric technique is widely used for several samples, isotropic or not, also for polarizing active material. Ellipsometry remains at the moment the most powerful method for investigating complex systems, such as nonisotropic multilayered samples or gradient varying samples.

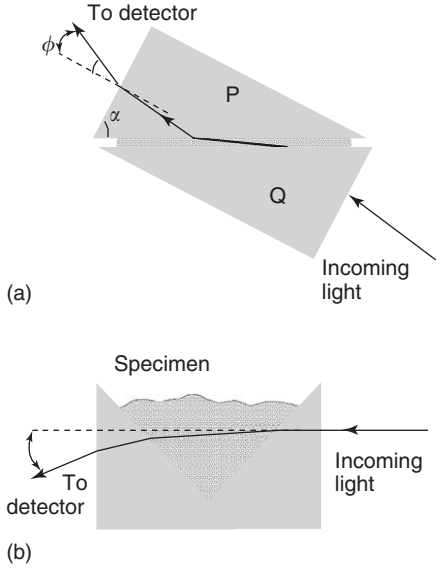


Figure 5 (a) Abbe refractometer and (b) Hilger–Chance instrument.

Interferometers

Interferometers are complex optical devices based on interference phenomenon. Optical interference occurs when two or more light waves interact yielding a resultant irradiance that deviates from the sum of the component irradiances.

A few of the more important interferometric systems will be examined here. Interferometric devices are divided into two groups: wave front splitting and amplitude splitting. In the first case, portions of the primary wave front are used as sources to emit secondary waves, or in conjunction with optical devices to produce virtual sources of secondary waves. These secondary waves are then brought together to interfere. In the case of amplitude splitting, the primary wave itself is divided into two segments, which travel different paths before recombining and interfering.

In accordance with the principle of superposition, the light field or the electric field intensity $\boldsymbol{E}$, at a

point in space, arising from the separate field $\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots$ of various contributing sources is given by

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots \qquad [18]$$

Taking the time average of both sides, one finds that the irradiance in the case of two superposed fields becomes $I = I_1 + I_2 + I_{12}$ provided that $I_1 = \langle \boldsymbol{E}_1^2 \rangle$, $I_2 = \langle \boldsymbol{E}_2^2 \rangle$, and $I_{12} = 2\langle \boldsymbol{E}_1 \cdot \boldsymbol{E}_2 \rangle$, where $\langle\ \rangle$ is the time average over a time interval T much greater than the period of the harmonic function. The latter expression is known as the interference term and is $I_{12} = \langle \boldsymbol{E}_{01} \cdot \boldsymbol{E}_{02} \cos\delta \rangle$, and $\delta = (\boldsymbol{k}_1 \cdot \boldsymbol{r} - \boldsymbol{k}_2 \cdot \boldsymbol{r} + \phi_1 - \phi_2)$ is the phase difference arising from a combined path length and initial phase-angle difference.

The most common situations correspond to $\boldsymbol{E}_{01}$ parallel to $\boldsymbol{E}_{02}$. In that case, the irradiance reduces to the value found in the scalar case: $I_{12} = E_{01}E_{02}\cos\delta$ and the total irradiance is

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta \qquad [19]$$

If a screen is placed in the region of interference, bright and dark zones would be visible, which are known as interference fringe pattern.

At various points in space, the resultant irradiance can be greater, less than, or equal to $I_1 + I_2$, depending on δ. When $\delta = 0, \pm 2\pi, \pm 4\pi, \ldots, \cos\delta = 1$ and maximum of irradiance is obtained. The waves are said to be in phase and one speaks of this as total constructive interference. When $0 < \cos\delta < 1$, the waves are out of phase. As $\delta = \pi/2,\ \cos\delta = 0$ the optical disturbance is said to be 90° out of phase, and $I = I_1 + I_2$. For $-1 < \cos\delta < 0$, one has the condition of destructive interference. The minimum in the irradiance results when the waves are 180° out of phase, $\cos\delta = -1$. This occurs when $\delta = \pm\pi, \pm 3\pi, \pm 5\pi, \ldots$, and it is referred to as total destructive interference.

A special case arises when $I_1 = I_2 = I_0$, the two sources have equal irradiance and

$$I = 2I_0(1 + \cos\delta) = 4I_0\cos^2\frac{\delta}{2} \qquad [20]$$

for which it follows that $I_{\min} = 0$ and $I_{\max} = 4I_0$.

When two beams interfere to produce a stable pattern, they must have very nearly the same frequency. A significant frequency difference would result in a rapidly varying, time-dependent phase difference, which in turn would cause I_{12} to average to zero during the detection interval. The most common means of overcoming this problem is to make one source serve to produce two coherent secondary sources.

It is to be remembered that because of the granular nature of the emission process, conventional quasi-monochromatic sources produce light that is a mix of photon wave trains. At each illuminated point in space, there is a net field that oscillates for less than 10 ns or so before it randomly changes phase. This interval over which the lightwave resembles a sinusoid is a measure of what is called its temporal coherence t_c. The average time interval during which the lightwave oscillates in a predictable way is designated as the coherence time of the radiation. The longer the coherence time, the greater the temporal coherence of the source. The corresponding spatial extent over which the lightwave oscillates in a regular, predictable way is called the coherence length $l_c = c \times t_c$, where c is the velocity of light. The most common means to produce interference, as will be seen, is to make one source serve to produce two coherent secondary sources that will interfere. An interference fringe pattern will be seen as long as the difference between the optical paths of the two secondary sources is no longer than the coherence length of the lightwave. Nowadays, laser sources are quite cheap and easier to use than lightwave sources with a temporal coherence ranging up to at least 1 ms, corresponding to Km of coherence length.

Wave Front Splitting Interferometer

A hypothetical monochromatic plane wave illuminating two parallel narrow closely spaced slits, S_1 and S_2 is to be considered. This is shown in a two-dimensional view in **Figure 6**. When symmetry exists, the segments of the primary wave front arriving at the two slits will be exactly in phase, and the slits will constitute two coherent secondary sources. One expects that the two waves coming from S_1 and S_2 overlap, interference will occur (provided the optical path difference is less than the coherence length, $c \cdot t_c$). In a realistic physical situation, the distance between the screen and the slits will be very large in comparison with the distance a between the two slits, and all the fringes would be fairly close to the center O of the screen. One can express the path difference as

$$r_1 - r_2 = a\theta \qquad [21]$$

since $\theta \backsim \sin\theta$.

Following this approximation $\theta = y/s$, $r_1 - r_2 = ay/s$. Constructive interference will occur when $r_1 - r_2 = m\lambda$; thus from the last two relations, one obtains

$$y_m = \frac{s}{a}m\lambda \qquad [22]$$

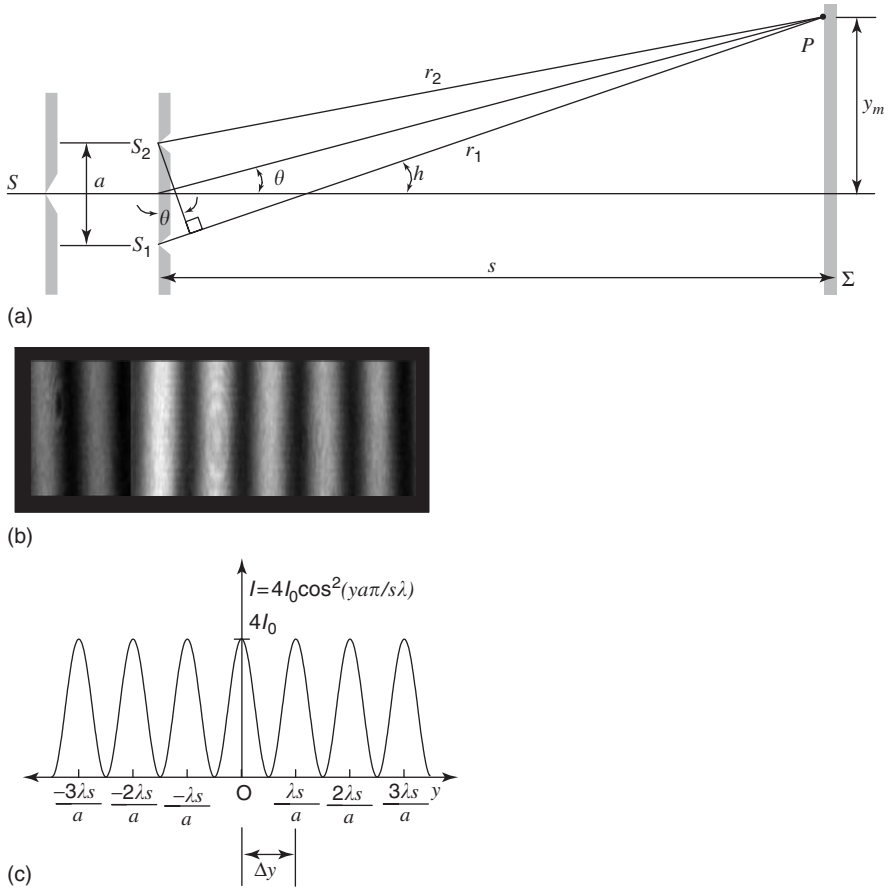


Figure 6 (a) Geometry of Young's experiment which is the most pedagogical example of wave-front splitting interferometer. A path-length difference of one wavelength corresponds to $m = \pm 1$ and to the first maximum. (b) Irradiance on a photoplate or a CCD camera is shown together with (c) the idealized irradiance vs. distance curve.

This gives the position of the mth bright fringe on the screen, if the maximum at 0 is counted as the zeroth fringe.

The interferometric configuration discussed above is known as Young's experiment. The same physical and mathematical considerations apply directly to a number of other wavefront splitting interferometer. Most common among these are Fresnel double mirror, Fresnel double prism, diffraction grating, and mirror configuration.

Amplitude Splitting Interferometer

Suppose that a light wave was incident on a semi-transparent mirror, or simply a sheet of glass. Part of the wave is transmitted and another part is reflected. Both the transmitted and the reflected waves will have lower amplitudes than the original one. When the two separate waves are brought together again at the detector, interference results as long as the original coherence between the two are not destroyed and if the path length differs by a distance lower than the coherence length.

Dielectric films The simpler amplitude splitting interference effects are observable in sheet transparent materials. Consider the simple case of a transparent parallel plate of dielectric material having a thickness d (**Figure 7**). Suppose that the film is nonabsorbing and that the amplitude–reflection coefficients at the interfaces are so low that only the first two reflected beams E_{1r} and E_{2r} need be considered. One may consider S to be a monochromatic point source. The film serves as an amplitude-splitting device, so that E_{1r} and E_{2r} are parallel on leaving the film and can be brought together at a point P on the focal plane of a lens. From **Figure 7**, the optical path length difference for the two reflected beams is given by

$$\Lambda = 2n_s d \cos\theta_t \qquad [23]$$

The corresponding phase difference associated with the optical path-length difference is then just the product $k_0\Lambda$, where $k_0 = 2\pi/\lambda_0$. If the film is immersed in a single medium $n_1 = n_2 = n$ and there

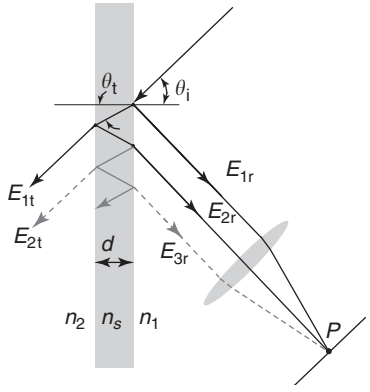


Figure 7 Thin layer interference.

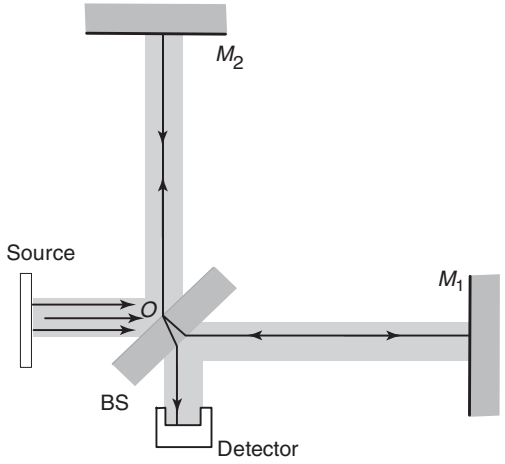


Figure 8 Michelson interferometer.

is an additional phase shift arising from the reflections between the two light beams themselves, one internally and one externally reflected, there will be a relative phase shift of π radians. Accordingly, $\delta = k_0\Lambda \pm \pi$. In the reflected light, an interference maximum, a bright spot appears at P when δ is an even multiple of π. In that case

$$d\cos\theta = (2m+1)\frac{\lambda_0}{4n_s} \qquad [24]$$

This also corresponds to a minima in the transmitted light. Interference minima in reflected light (maxima in transmitted light) result when δ is an odd multiple of π. For such cases

$$d\cos\theta = 2m\frac{\lambda_0}{4n_s} \qquad [25]$$

The thickness of sheet transparent materials commonly used in optics, varies from films less than the length of a light wave, to plates several centimeters thick. A layer of material is referred to as a thin film for a given wavelength of electromagnetic radiation when its thickness is of the order of that wavelength. The applications of thin dielectric films are manifold, as are their structures which extend from the simplest single coatings to intricate arrangements of hundreds of layers. Coatings for high reflecting mirrors, to eliminate unwanted reflections, nonabsorbing beamsplitters, dichroic mirrors, multilayer broad, and narrow bandpass filters are examples of the thin film optical devices.

Mirrored interferometer Among the amplitude-splitting interferometers that utilize mirrors and beam splitters, the best known and historically most important is the Michelson interferometer. Its configuration is illustrated in **Figure 8**: a quasimonochromatic source emits a wave, part of which travels to the right. The beam splitter O divides the wave into two, one segment traveling to the right and one up into the background. The two waves are reflected by mirrors M_1 and M_2 and return to the beam splitter. Part of the wave coming from M_2 passes through the beam splitter going downward, and part of the wave coming from M_1 is deflected by the beam splitter toward the detector. Thus the two waves are united, and interference can be expected. The optical path difference for these rays is $2d$, where d is the difference length between the two interferometer arms (OM_1 and OM_2), which represents a phase difference of $k_0 2d$. There is an additional phase term arising from the fact that the wave traversing the arm OM_2 is internally reflected in the beam splitter, whereas the OM_1 wave is internally reflected at O. If the beam splitter is simply uncoated, the relative phase shift resulting from the two reflections will be π radians. Destructive rather than constructive interference will occur when $2d = m\lambda_0$, where m is an integer. If the light source is slightly divergent, an observer will see a circular fringe system. When the mirrors of the interferometer are inclined with respect to each other, making a small angle (i.e., when M_1 and M_2 are not quite perpendicular), a pattern of straight parallel fringes is visible.

It is apparent that the Michelson interferometer can be used to make extremely accurate length measurements. As the movable mirror is displaced by $\lambda_0/2$, each fringe will move to the position previously occupied by an adjacent fringe. One need only count the number of fringes N, or a portion of a fringe, that has moved past a reference point to determine the distance traveled by the mirror $\Delta d = N(\lambda_0/2)$.

Very accurate measurements of wavelengths are also performed using the Michelson interferometer. The beam of a reference light source, usually a very stable and well-known wavelength λ_r laser, and the beam within the unknown wavelength λ_x are aligned into the interferometer with identical paths but opposite directions. This produces two different fringe systems to be revealed by two different detectors. While the movable mirror is displaced, sinusoidal signals are visible by the detectors: the rates of the interference maxima are counted by the detectors for both the known and unknown wavelengths. From the ratio of both counting rates, the unknown wavelength λ_x can be obtained by $\lambda_x/\lambda_r = N_r/N_x$.

The Twyman–Green interferometer is essentially a variation of the Michelson interferometer. It is an instrument of great importance in the domain of modern optical testing. Illustrated in **Figure 9**, the setup has a lens L_1, used to provide a source of incoming plane waves, and a lens L_2, which permits all the light, from the aperture to enter the detector so that the entire field can be seen.

Twyman–Green interferometers are used to test the quality of a surface or of a lens with respect to a reference wave front. For instance, if a reference flat high-quality surface is used for mirror M_1, the mirror M_2 can be tested and any wave front distortions produced by the surface of M_2 will appear in the final interferogram, provided that high-quality low-deformation lenses L_1 and L_2 have been used.

The Mach–Zehnder interferometer shown in **Figure 10** consists of two beamsplitters and two totally reflecting mirrors. The two waves within the apparatus travel along separated paths. A difference between the optical paths can be introduced by a slight tilt of one of the mirrors. Since the two paths are separated, the interferometer finds myriad applications. An object interposed in one beam will alter the optical path-length difference, thereby changing the fringe pattern. This results in a very accurate determination of the index of refraction of the sample and its local variations. For these reasons, a common application of the device is to observe the density variations in liquid or gas flow patterns.

Another amplitude-splitting device is the Sagnac interferometer shown in **Figure 11**. It is to be noticed that the main feature of the device is that there are two identical but oppositely directed paths taken by the beams that form closed loops before they are united to produce interference. A slight shift in the orientation of one of the mirrors will produce a path-length difference and a resulting fringe pattern changing. When the entire interferometer (mirrors, source, and detector) rotates about a perpendicular axis passing through its center, the rotation effectively shortens the path taken by one beam in comparison to that of the other. In the interferometer, the result is a fringe shift proportional to the angular

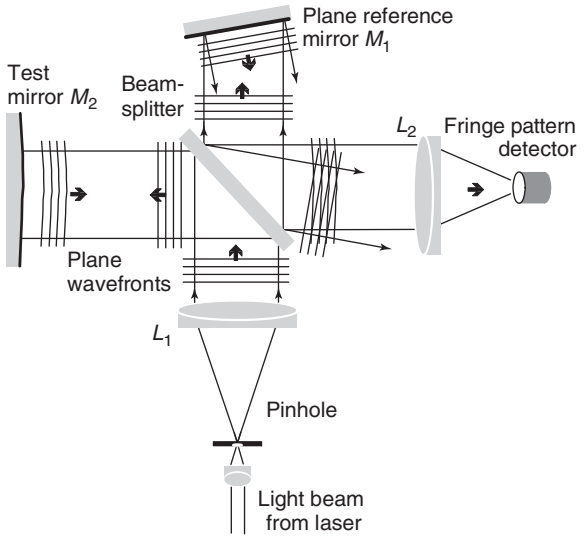


Figure 9 Twyman–Green interferometer and arrangement for a mirror test.

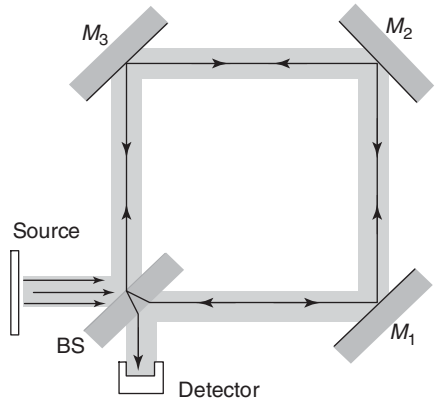


Figure 10 Mach–Zehnder interferometer.

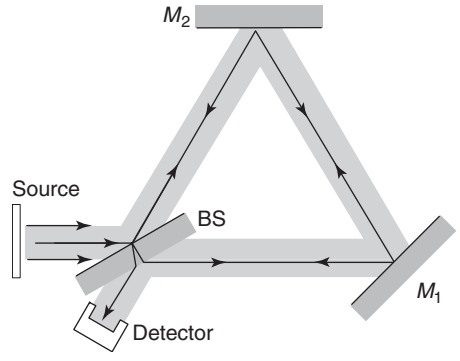


Figure 11 Sagnac interferometer.

speed of rotation $\Omega : \Delta N = 4A\Omega/c\lambda$, where A is the area formed by the beams of light.

Shearing interferometer The method of lateral shearing interferometry consists of displacing the light beam wave front laterally by a small amount and obtaining the interference pattern between the original and the displaced wave fronts. **Figure 12** schematically illustrates the principle in one of the simpler experimental configurations, based on a plane parallel plate. There are many physical arrangements to obtain lateral shear. The most popular arrangements derive from the use of beam dividers, that means, of amplitude splitting interferometers. In this way, the amplitude of the incident wave front is divided, but the shape of the wave front does not change. If the wave front is considered nearly planar, it can be expressed as $W(x, y)$, where (x, y) are the coordinates of point P. When this wave front is sheared in the x direction by an amount s, the sheared wave front is expressed by $W(x - s, y)$. The resultant path difference at P between the original and the sheared wave front is:

$$\Delta W = W(x - s, y) - W(x, y) \qquad [26]$$

The path difference ΔW may be obtained at various points on the wave front from the relation: $\Delta W = m\lambda$, where m is an integer indicating the order of the interference fringe. It is expected that for a perfect, plane wave $\Delta W = 0$ a null fringe will occur, but as soon as the wave front is distorted and ΔW is not zero, a fringe system, sometimes rather complicated, is visible. When the wave front is laterally sheared, normally it is assumed that the new wave front is not tilted with respect to the original one. It is possible to obtain a known amount of tilt between the two wave fronts by using a wedge with a small angle instead of a parallel plate in **Figure 12** setup. In such a case, the optical path difference associated with this tilt may be presented as a linear function of the coordinate along the tilt, and a system of straight fringes is generated. A deviation from the straight fringe system is a measure of the wave front deformation.

Lateral shearing interferometers are commonly used for checking the collimation of a lens, or measuring lens parameters such as focal length or the refractive index, for studying of homogeneity of transparent solid samples, and for flow studies. Another application involves the measurement of surface imperfections of concave or convex surfaces.

Here, it is important to note that Mach–Zehnder, Twyman–Green, and shearing interferometers are widely used in noncontact measurement of surface height profiles. They can be used to determine surface roughness or geometries of small features. All these interferometric profile measurements are noncontact and therefore do not harm the surface under test. They are very sensitive and can measure heights with a precision of up to 0.1 Å RMS (root mean square). Quantitative phase-measurement techniques are generally used to determine the phase difference between an object point and a reference. This phase is then converted into height information. Because of

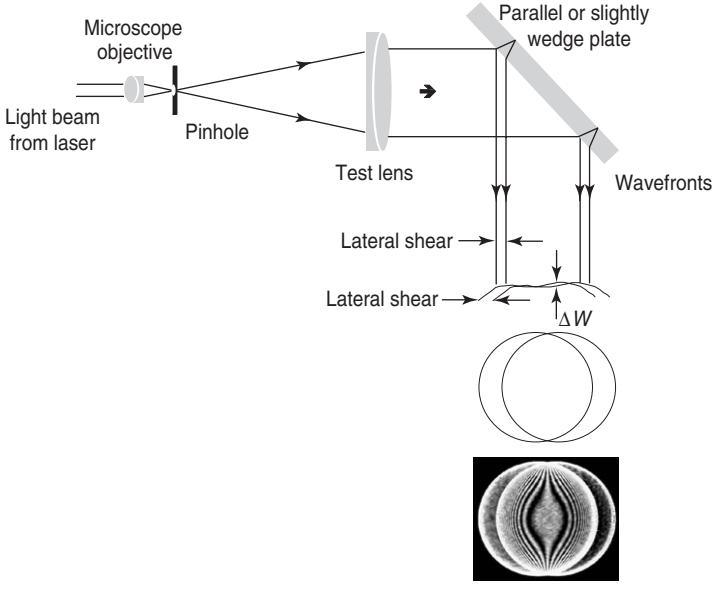


Figure 12 Schematic diagram of a lateral shearing interferometer using a plane parallel plate of glass. Diagram also illustrates the original and the sheared wave front, and lateral shearing interferogram of a wave front distorted by spherical aberration.

the precision in measurement and noncontact configuration, interferometric optical profilometers are appropriate for on-line process control. All these instruments measure heights relative to a reference surface rather than absolute distance. All profiling instruments can be calibrated relative to an existing, traceable standard; however, most of these instruments are better than existing standards, so it is hard to determine their absolute accuracy.

The main limitation of interferometric optical profilers is due to the phase shift induced by reflection of the object beam from the test surface. Every material induces a phase shift depending on its complex index of refraction. As long as a surface comprises of a single material and does not have a transparent coating on it, the phase shift does not cause a problem. When two dissimilar materials are next to each other on the surface, they will have different phase shifts upon reflection and the height difference at the boundary will be incorrect. This problem can be overcome using a coating with an opaque material, or by knowing the optical constants of the different materials.

Multiple beam interferometer The multiple beam interferometer (also known as Fabry–Perot interferometer after its inventors) consists of two plane parallel highly reflecting surfaces or mirrors separated by some distance d (**Figure 13**). Light coming in the volume between the mirrors can be partly reflected and transmitted by the surfaces. The distance d generally ranges from several millimeters to several centimeters when the apparatus is used interferometrically, and to few tens of centimeters or more when it serves as a laser resonant cavity. Consider a point source S_1 and one ray emitted, entering by way of the partially reflecting plate, it is multiply reflected within the gap. The transmitted rays are collected by a lens and brought to a screen. The multiple waves generated in the cavity, arriving at P from S_1 are coherent among themselves. All the rays incident on the parallel plates at a given angle will result in a circular fringe of uniform irradiance. The interference bands will be narrow concentric rings, corresponding to multiple-beam transmission pattern. The phase difference between two successively transmitted waves is

$$\delta = \frac{4\pi n}{\lambda_0} d \cos\theta_t + 2\phi \qquad [27]$$

where ϕ is an additional phase shift introduced by metallic films at the surfaces, if present. ϕ can be considered constant and, in general, d is so large and λ_0 so small that ϕ can be neglected. One can consider that reflectance R is the fraction of the flux density reflected at each incidence and T the transmittance. One can neglect absorption from the metal films that are used to increase the reflectance, and consider $R + T = 1$. The irradiance of the transmitted beam is

$$I_t = I_0 \frac{(1 - R^2)}{(1 - R^2) + 4R \sin^2(\delta/2)} \qquad [28]$$

and the irradiance of the reflected beam is

$$I_t = I_0 R \frac{4 \sin^2(\delta/2)}{(1 - R^2) + 4R \sin^2(\delta/2)} \qquad [29]$$

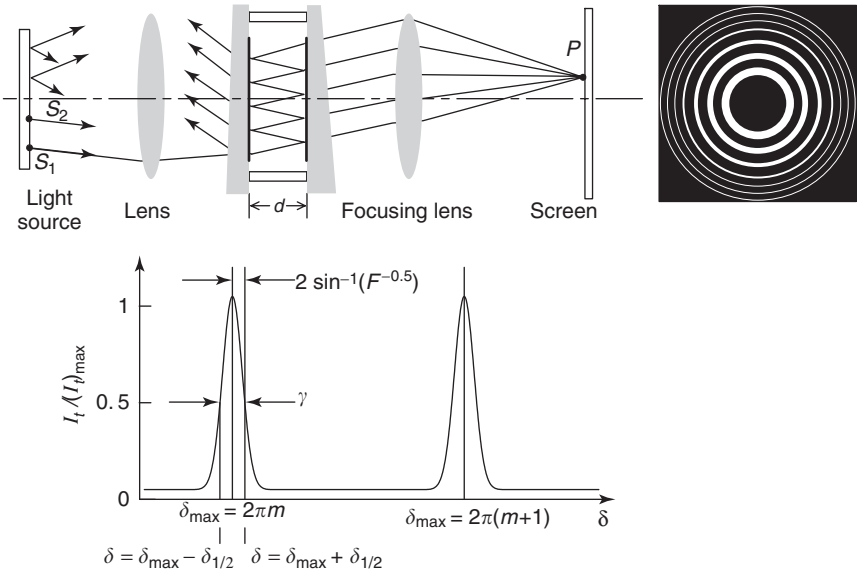


Figure 13 Multiple beam interferometer and intensity profile of Fabry–Perot fringes.

Multiple-beam interference has resulted in a redistribution of the energy density in comparison to the sinusoidal two-beam pattern. Peaks in the transmission occur at specific values of the phase difference $\delta = 2\pi m$. The frequency range δv between two maxima is the free spectral range of the interferometer. For normal incidence, the free spectral range becomes $\delta v = c/2nd$. Accordingly, the irradiance will drop to half of its maximum value whenever δ shifts from $\delta_{\max}$ of $2(1-R)/\sqrt{R}$. In frequency units, the halfwidth becomes $\Delta v = (c/2nd)((1-R)/(\pi\sqrt{R}))$. The ratio $\delta v/\Delta v$ to the free spectral range δv and the halfwidth Δv is called the finesse F of the interferometer:

$$F = \frac{\Delta v}{\delta v} = \frac{\pi\sqrt{R}}{1-R} \qquad [30]$$

The finesse is determined only by the reflectivity R of the surfaces.

The Fabry–Perot interferometer is frequently used to examine the detailed structure of spectral lines. Two incident waves with frequencies v_1 and $v_2 = v_1 + \Delta v$ produce two systems of fringes partly superimposed. This will traduce into two transmitted irradiances corresponding to the two different wavelengths. The spectral resolution $v/\Delta v$ of the interferometer is determined by the free spectral range δv and by the finesse F. A Fabry–Perot interferometer with $d = 1$ cm, $R = 0.98$ for a wavelength light in the visible range, that is, $\lambda = 500$ nm, having negligible wedge and high-quality surfaces, so that the finesse is mainly determined by the reflectivity, achieves $F = 155$, a resolving power $\Delta\lambda/\lambda = 10^{-7}$. This means that the line width of the instrument is $\delta\lambda = 5 \times 10^{-5}$ nm and in frequency unit, $\delta v = 60$ MHz.

Both the applications and configurations of the Fabry–Perot interferometer are numerous indeed. Solid etalon, or air-spaced plane parallel reflecting surfaces are used as transmission filters. Using spherical mirrors where the demand for parallelism is dropped, with high-reflectivity dielectric coating can achieve high large finesse $F > 10\,000$. Spherical mirror Fapry–Perot configurations are also used for laser cavities.

As a final comment on interferometers, it needs to be pointed out that nowadays much of the progress in the field of interferometry has come not from inventing new forms of interferometers, but from powerful fringe analysis methods, such as phase-shift interferometer analysis in optical testing, mostly helped by computer data treatment, and from making user-friendly instruments. Interferometry has the great advantage of a high-precision and noncontact capability in measurement. Commercial interferometric systems can achieve a resolution in the direction of propagation of the light beams of 2 nm and in the transversal plane ten times larger. This corresponds also to a resolution in the index of refraction measurement resolution of 10–6 and in highly controlled laboratory experiments in a 10–8 resolution. The major impediment of interferometry is its sensitivity: because of this, instruments tend to measure the whole surrounding environment, which can adversely affect accuracy.

Polarimeter and Ellipsometer

As discussed above, light may be treated as transverse electromagnetic waves. Thus far, only linear polarized light has been considered, that is, light for which the orientation of the electric field is constant, although its magnitude and sign vary in time. The electric field, therefore, resides in a fixed plane that contains both $\boldsymbol{E}$ and $\boldsymbol{k}$, the electric field vector and the propagation vector in the direction of motion, and is known as the plane of vibration.

One can resolve any plane polarized wave into two orthogonal components with the appropriate relative phase difference. For a wave propagation along the z axis, the two orthogonal electric field components can be represented in the form

$$\boldsymbol{E}_x(z,t) = \mathbf{i}\cos(kz - \omega t) \qquad [31]$$

$$\boldsymbol{E}_y(z,t) = \mathbf{j}\cos(kz - \omega t + \varepsilon) \qquad [32]$$

If ε is equal to zero or an integral of $\pm 2\pi$, the resultant electric field becomes

$$\boldsymbol{E} = (\mathbf{i}E_{0x} + \mathbf{j}E_{0y})\cos(kz - \omega t) \qquad [33]$$

and has a fixed amplitude. If ε is an odd integer multiple of $\pm\pi$, the electric field is again linearly polarized but the plane of vibration has been rotated. Circular polarization occurs when both constituent waves have equal amplitudes and their relative phase difference $\varepsilon = \pm\pi/2 + 2m\pi$, where $m = 0, \pm 1, \pm 2, \ldots$ the consequent wave is given by

$$\boldsymbol{E} = E_0[\mathbf{i}\cos(kz - \omega t) \pm \mathbf{j}\cos(kz - \omega t)] \qquad [34]$$

The amplitude is constant but the electric field $\boldsymbol{E}$ is time varying and it rotates clockwise (right-circularly polarized light) or counterclockwise (left-circularly polarized) at an angular frequency ω.

Both linear and circular polarized light may be considered to be special cases of elliptically polarized light. This means that, in general, the resultant electric field vector will rotate and change its magnitude as well. In such cases, the endpoint of

$\boldsymbol{E}$ will trace out an ellipse in a fixed space perpendicular to $\boldsymbol{k}$.

Instruments employed to measure polarization of a light beam after its interaction with a sample are commonly known as polarimeters or ellipsometers. A polarimeter is used when a light beam is transmitted through the sample whereas an ellipsometer is commonly used when a light beam is reflected by the sample. Both, the polarimeter and ellipsometer, use polarizers and retarders.

An optical device with an input of nonpolarized light and an output of some form of polarized light is known as a polarizer. For instance, an instrument that separates the two linear orthogonal components of electric field, discarding one and passing on the other, is a linear polarizer. Depending on the output it could also be classified as circular or elliptical polarizer. Polarizers come in many different configurations, but they are all based on one of the four fundamental physical mechanisms: dichroism or selective absorption, reflection, scattering, and birefringence or double refraction.

Polarizers are used in two different ways. If the polarizer is used to convert unpolarized light to polarized light, it is called the "polarizer." If it is used to determine the state of polarized light (usually by locating the null), it is called the "analyzer."

The next optical element to be considered is called the "retarder" or "quarter-wave-plate." The wave plate, in general, is an anisotropic optical element with a fast axis and a slow axis, both of which are perpendicular to each other and to the direction of propagation of light. The component of the wave which is aligned with the fast axis passes through the optical element faster than the component aligned with the slow axis. The thickness of the wave plate can be chosen such that the phase difference is exactly 90°.

Polarimeter

Polarimeters are optical instruments used for determining the polarization properties of light beams and samples. Light-measuring polarimeters determine the polarization state of a beam of light and give its polarization characteristics. Sample-measuring polarimeters are used in light–matter interaction in which an incident beam can be transmitted, reflected, diffracted, or scattered by the sample. Polarimeters are used to determine the relationship between the polarization states of incident beams and output beam. Polarimeters are composed of a polarization generator consisting of all elements needed to produce a beam of a known polarization state, and a polarization analyzer to perform a particular polarization component in the output light beam.

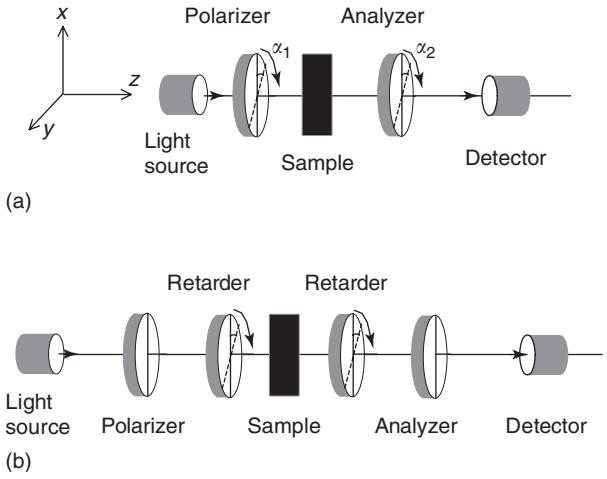


Figure 14 (a) The dual rotating polarizer polarimeter consists of a source, a linear polarizer which rotates, the sample, a second linear polarizer with a varying angular orientation, and the detector. (b) The dual rotating retarder polarimeter consists of a source, a fixed linear polarizer, a retarder which rotates, the sample, a second retarder with a varying angular orientation, a fixed linear polarizer, and the detector. The polarimeter measures, in a complete way, all the polarization characteristics.

A common polarimetric configuration is shown in **Figure 14a**. Light from the source passes through a linear polarizer whose orientation α_1 is adjustable. The generated linearly polarized light interacts with the sample and is analyzed by a second linear polarizer whose orientation α_2 is also adjustable. This polarimeter is sufficient only when linear polarization effects are dominant. Whenever circular polarization effects are present in the sample, elliptical properties of the output light must be considered and a more complete polarimeter is needed as shown in **Figure 14b**.

Typical applications of polarimeters include calibration of polarization elements for optical systems, but they are also widely used to analyze any optically active substances. There is a large number of organic and inorganic substances that are optically active in their crystalline, liquid, or dissolved state. Being optically active means that these substances are able to rotate the direction of oscillation of polarized light around a determined angle. A well-known optical substance is sugar: if it is dissolved in an optically inactive liquid like water, the degrees of angle of rotation depend – amongst other factors – on the concentration of the solution. This angle of rotation can be determined by a polarimeter with the highest possible precision. Analysis of active sugar solutions, and of any active solutions, is the most important application of polarimeters in the quality control, determination of concentration, and purity control in food, chemical, and pharmaceutical industries.

Ellipsometer

In ellipsometry, a collimated beam of a well-known polarized and monochromatic light is incident on a sample surface and the state of polarization of the reflected light is analyzed. In a typical scheme shown in **Figure 15a**, light is linearly polarized at a known but arbitrary plane of vibration identified by angle θ_i with the plane of incidence. The two components of polarization being *S* and *P*, are taken respectively parallel and perpendicular to the plane of incidence. For optically isotropic structures, ellipsometry is carried out at oblique incidence and reflected light results to be elliptically polarized. For incident electric field components E_{is} and E_{ip}, the corresponding field components of the reflected light are given by

$$E_{rs} = r_s E_{is}, \quad E_{rp} = r_p E_{ip} \qquad [35]$$

where r_p and r_s are the complex Fresnel reflection coefficients of the sample for *P* and *S* polarized light.

By taking the ratio of the respective sides of the two equations, one gets

$$\rho = r_p/r_s = \frac{E_{is}/E_{ip}}{E_{rs}/E_{rp}} \qquad [36]$$

The second term in the equation describes the incident and reflected polarization states of light. Therefore, pure polarization measurements determine the ratio of the complex reflection coefficient for the *S* and *P* polarizations. These ratios are subsequently related to the structural and optical properties of the ambient–sample interface region by invoking an appropriate model and the electromagnetic theory of reflection.

The ellipsometric measurement is normally expressed in terms of Ψ and Δ:

$$\tan\Psi\, e^{i\Delta} = \rho = \frac{r_p}{r_s} \qquad [37]$$

Referring to **Figure 15**, the phase difference between the parallel component and the perpendicular component of the incoming wave is denoted as δ_1, and the phase difference between the parallel and the perpendicular component of the outgoing wave as δ_2. It results that the parameter $\Delta = \delta_1 - \delta_2$ is the change in phase difference that occurs upon reflection and its value can be from 0° to 360°. Without regard to phase, the amplitude of both perpendicular and parallel components may change upon reflection, their ratio represented by the tan Ψ.

As ellipsometry measures a ratio quantity, the measurement is potentially more precise and accurate than a traditional intensity reflectance or transmission measurement. On the other hand, since ellipsometry is a polarization-based measurement, relatively complex optical instrumentation is required. Spectroscopic ellipsometry measures the complex ratio ρ as a function of wavelength, whereas variable angle spectroscopic ellipsometry performs the above measurement as a function of both wavelength and angle of incidence.

In case multiple interfaces have to be analyzed, it is possible to derive the ratio of the amplitude of the resultant reflected wave to the amplitude of the incident wave as a function of the Fresnel reflection coefficients for the single interfaces, wavelength of light, and thickness of layers.

For many instruments used today, the intensity measurement of the reflected light is related to

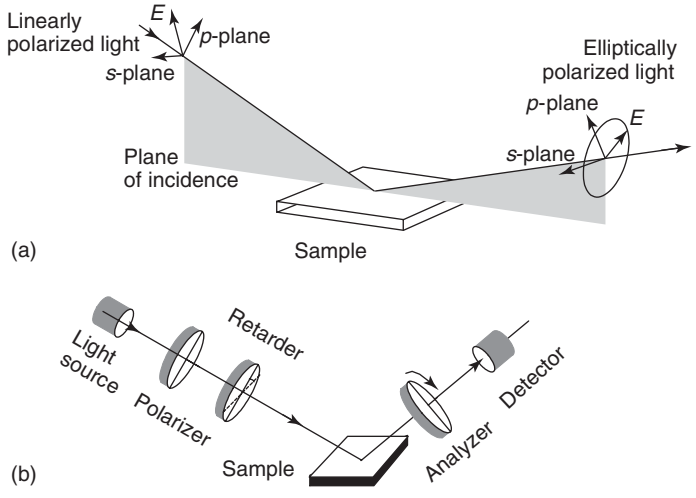


Figure 15 (a) Interaction of polarized light with a sample. (b) A rotating element ellipsometer. In this ellipsometer, only the analyzer rotates. The intensity measurement can be made with the quarter-wave-plate retarder either in the beam or outside it.

rotation of the analyzer of polarization on the reflected beam. This instrument is shown schematically in **Figure 15b**. In this case the polarizer on the incident beam is set at a fixed angle, usually 45°. The quarter wave plate (QWP) is fixed with the fast axis perpendicular to the plane of incidence and can be placed in the beam or removed from it. Only one of the polarizing elements (the analyzer) is rotated. The intensity I measured as a function of time t by the light detector can be expressed as

$$I(t) = I_0[1 + \alpha\cos 2A(t) + \beta\sin 2A(t)] \qquad [38]$$

where $A(t) = 2\pi ft + A_c$ with f being the angular frequency of the rotation and A_c being a constant phase offset. I_0 is the average intensity, and α and β are the Fourier coefficients. For a rotating analyzer instrument, the optimum precision is obtained when the light incident on the rotating analyzer is circularly polarized. If the phase shift induced by the reflection is $\sim 90°$, the optimum precision would be realized by using linearly polarized light (no QWP). The light starts out linearly polarized (components in phase) and the reflection changes it into near-circularly polarized light. If, on the other hand, the phase shift due to the reflection is either none or $\sim 180°$ without the QWP, the light incident on the analyzer would be almost linearly polarized. In this case the QWP is used. Typically, Ψ is measured without the QWP in place and Δ is measured both with and without the QWP. Depending on the value obtained, one or the other of the measurements is used.

The Fourier coefficients α and β are given by

$$\alpha = -\cos 2\Psi \quad \text{and} \quad \beta = \sin 2\Psi \cos \Delta_m \qquad [39]$$

where Δ_m is the measured phase difference between the parallel and the perpendicular components of the light at the analyzer. These equations can be inverted to give

$$\cos \Delta_m = \pm\sqrt{\frac{\beta^2}{1-\alpha^2}} \qquad [40]$$

and

$$\tan \Psi = \sqrt{\frac{1+\alpha}{1-\alpha}} \qquad [41]$$

If the QWP is not used, the phase shift induced by the reflection is given by $\Delta = \Delta_m$. If the QWP is in place and the light incident on the sample is circularly polarized, then it is necessary to add the retardation of the QWP to the measured value. If the QWP is perfect, then $\Delta = \Delta_m + 90°$.

Many desired material parameters can be extracted by this technique, including layer thickness, surface and/or interfacial roughness, and optical constants.

$$n = n_0 + \mathrm{i}n_1 \qquad [42]$$

where $n_0 = c/v$ is the index of refraction and n_1 is the extinction coefficient and takes care of absorption phenomena. The ellipsometric data measured on a bulk sample can be directly inverted into the optical constants of the material, assuming that surface oxide or roughness effects are negligible:

$$\begin{aligned}\tilde{n} &= n_0 + \mathrm{i}\kappa \\ &= \sin\phi^2\left[1 + \tan\phi^2 \times \left(\frac{1-\rho}{1+\rho}\right)^2\right]\end{aligned} \qquad [43]$$

where ϕ is the angle of incidence and ρ is measured by an ellipsometer.

This equation gives the averaged and a first appreciation of physical parameters under test. A full optical and microstructure analysis of real materials including multilayer structures, requires use of an appropriate model and regression analysis. It should be noted that ρ is the quantity measured by an ellipsometer and, assuming the instrument is operating properly, it is correct. The calculated sample parameters, such as thickness and index of refraction are correct, depend on whether the model assumed is correct. Typical examples of applications are determination of thickness and index of refraction of thin-film stacks up to five layers, material, optical constant of photoresist which is critical for industrial lithography, measurement of thickness and compositions of complex compound semiconductor structures, optical multilayer coatings, and variation in optical coating index.

Radiometer

Radiometry deals with the measurement of amounts of light. Photometry takes into account the response of the human eye which rises to a nonlinear and wavelength-dependent subjective impression of radiometric quantities.

Radiant power W, measured in watts, is the total amount of energy emitted by a light source per second. The spectral variation of radiant power can be specified in terms of the radiant power density per unit wavelength interval W_λ

$$W = \int_0^\infty W_\lambda \, \mathrm{d}\lambda \qquad [44]$$

If the light source emits radiation only for some specific duration, it is more useful to specify the source in terms of its radiant energy output Q_e, measured in joules. If the source emits radiation for a time T, one can write

$$Q_e = \int_0^T W(t)\,dt \qquad [45]$$

The amount of power emitted by a source in a particular direction per unit solid angle is called radiant intensity I_e, and is measured in watts per steradian. In general,

$$W = \oint I_e(\Omega)\,d\Omega \qquad [46]$$

where the integral is taken over a closed surface surrounding the source. If I_e is the same in all directions, the source is said to be an isotropic radiator. At distance r from such a source, if r is much greater than the dimension of the source, the radiant flux crossing a small area ΔS is

$$\phi_e = \frac{I_e \Delta S}{r^2} \qquad [47]$$

The irradiance at this point, measured in $W\,m^{-2}$, is

$$E_e = \frac{I_e}{r^2} \qquad [48]$$

which is equal to the average value of the Poynting vector measured at that point. The radiant flux emitted or reflected or scattered per unit area of a surface is called the radiant emittance M_e, and it is measured in $W\,m^{-2}$. For an extended source, the radiant flux emitted per unit solid angle per unit area of the source is called its radiance or brightness L_e:

$$L_e = \frac{\delta I_e}{\delta S_n} \qquad [49]$$

where the area δS_n is the projection of the surface element of the source in the direction being considered. When the light emitted from a source or scattered from a surface has a radiance that is independent of the viewing angle, the source or surface is called a perfectly diffuse or Lambertian radiator. For such a source, the radiant intensity at an angle θ to the normal to the surface is clearly

$$I_e(\theta) = I_e(0)\cos\theta \qquad [50]$$

The total flux emitted per unit area of such a surface is its radiant emittance, which, in this case, is $M_e = \pi I_e(0)$.

For plane waves, since all the energy in the wave is transported in the same direction, it is customary to specify the radiant flux crossing unit area normal to the direction of propagation, and call this the intensity I of the plane wave. Because lasers emit radiation into an extremely small solid angle, they have very high radiant intensity, and it is once again more usual to refer simply to the intensity of the laser beam at a point as the energy flux per second per unit area. Presently, radiometry involves the measurement of the energy content of electromagnetic radiation fields and the determination of how this energy is transferred from a source, through a medium, and to a detector, independently of its effect on a particular detector. The measurement of the effect of the transfer of radiation, such as absorption, reflection, or scatter is usually called spectrophotometry.

In common practice, radiometry is divided according to regions of the spectrum in which the same experimental technique can be used. Thus, vacuum ultraviolet, visible and near visible, intermediate-infrared, far infrared, and microwave regions are considered separate fields in radiometry.

Photometry concerns radiometry related to human vision, and measurement of the ability of electromagnetic radiation to produce a visual sensation in a physically realizable manner. The response of the human eye gives rise to a nonlinear and wavelength-dependent subjective impression of radiometric quantities. The response function of the human eye extends roughly from 400 to 700 nm with a peak at 555 nm. In photometry, radiometric quantities are used with different names and units. The luminous intensity, for instance, is the equivalent of the radiometric quantity, radiant intensity, and in photometry is measured as candela where 1 candela = 1 lumen str^{-1}.

Today, measurement of the energy content of electromagnetic radiation is performed with an electrically calibrated detector, which is a device for measuring absolute radiant power in comparison to electrical power. It consists of a thermal detector that has a radiation-heater within the surface, or the heater is in good thermal contact with the surface. When the device is irradiated, the detector senses the temperature of the radiantly heated surface. The radiation source is then blocked and the power to the electrical heater adjusted to reproduce the temperature of the radiantly heated surface. The electrical power to the heater is measured and equated to the radiant power on the surface. As the measurement has to be accurate, differences between the radiant and the electrical heating modes must be evaluated.

Detectors are distinguished by two types of radiant power absorber configurations: a flat surface coated

with a highly absorbing material or a cavity-shaped, light-trapping detector. Cavity-shaped radiometers are usually more accurate over a greater spectral range than flat-surface radiometers. However, a flat-surface receiver can be fabricated with less thermal mass than a cavity-shaped receiver, and, therefore, may have greater sensitivity and a faster response time.

Electrically calibrated detectors are in widespread use today; they are commercially available and are calibrated in the factory with a calibration lamp. They use either a thermocouple, a thermopile, or a bolometer. They operate at ambient temperature and are commonly used for solar/artificial lighting and for laser power measurements.

See also: Crystal Optics; Holography; Optical Properties of Materials.

PACS: 42.87. – d; 07.60.Hv; 07.60.Ly; 07.60.Fs; 07.60.Dq

Further Reading

Azzam RMA and Bashara NM (1987) *Ellipsometry and Polarized Light*, 2nd edn. Amsterdam: North-Holland.

Born M and Wolf E (1997) *Principles of Optics*, 6th edn. Cambridge: Cambridge University Press.

Demtroder W (1996) *Laser Spectroscopy*, 2nd edn. Berlin: Springer.

Hariharan P (1985) *Optical Interferometry*. Sydney: Academic Press.

Malacara D (1991) *Optical Shop Testing*. New York: Wiley.

Optical Microscopy

G F Vander Voort, Buehler Ltd., Lake Bluff, IL, USA

Metallurgical microscopes differ from biological microscopes primarily in the manner by which the specimen is illuminated due to the opacity of metals. Unlike biological microscopes, metallurgical microscopes must use reflected light, rather than transmitted light. **Figure 1** is a simplified ray diagram of a metallurgical microscope. The prepared specimen is placed on the stage with the surface perpendicular to the optical axis of the microscope and is illuminated through the objective lens by light from a lamp or arc source. This light is focused by the condenser lens into a beam that is made approximately parallel to the optical axis of the microscope by a half-silvered mirror. The light then passes through the objective onto the specimen. It is then reflected from the surface of the specimen, back through the objective, the half-silvered mirror, and then to the eyepiece to the observer's eye, or to a camera port or a film plane.

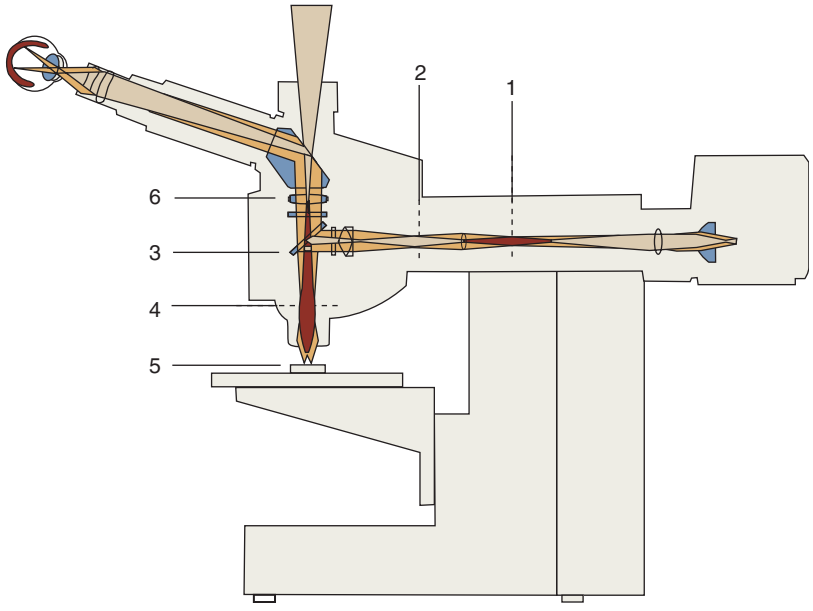


Figure 1 Schematic showing the light path in an upright reflected light microscope using bright-field illumination (1. aperture diaphragm; 2. luminous-field diaphragm; 3. reflector; 4. objective pupil; 5. sample surface; 6. tube lens). (Reproduced with permission of Buehler, Ltd.)

After the specimen has been properly sectioned and polished, it is ready for examination. However, in many cases, it is important to begin the examination with an as-polished specimen, particularly for quality control or failure analysis work. Certain microstructural constituents, such as nonmetallic inclusions, intermetallic phases, nitrides, cracks, or porosity, are best evaluated in the as-polished condition as etching reveals other microstructural features that may obscure such details. In the as-polished condition, there is a minimum of extra information for the observer to deal with which makes examination of these features most efficient. Some materials do not require etching and, in a few cases, examination is less satisfactory after etching.

For most metals, etching must be used to fully reveal the microstructure. A wide range of processes are used, the most common being etching with either acidic or basic solutions. For most metals and alloys, there are a number of general-purpose etchants that should be used first to reveal the microstructure. Subsequent examination may indicate that other more specialized techniques for revealing the microstructure may be useful. For example, the examination may reveal certain constituents that the metallographer may wish to measure. Measurements, particularly those performed with automatic image analyzers, will be simpler to perform and more reliable if the desired constituent can be revealed selectively, and numerous procedures are available to produce selective phase contrasting.

Henry Clifton Sorby

When Henry Clifton Sorby made his first examination of the microstructure of iron on 28 July 1863, he was using a transmitted light petrographic microscope made by Smith, Beck, and Beck of London that he had purchased in 1861. Although capable of magnifications up to 400×, most of his examinations were conducted at 30, 60, or 100× and his first micrographs were produced at only 9×. The objective lenses of this microscope were equipped with Lieberkühns silvered concave reflectors for focusing light on opaque specimens.

Sorby quickly realized that reflected light produced with the Lieberkühns reflectors was inadequate and he developed two alternate methods for this purpose. Subsequently, others developed vertical illuminators using prisms or plane glass reflectors and Sorby's systems were not further developed. In 1886, Sorby reported on the use of "very high power" (650×) for the study of pearlite. This was accomplished using a 45° cover glass vertical illuminator made for him by Beck.

For many years, photomicroscopy was conducted using specially built reflected microscopes known as metallographs. These devices represented the "top-of-the-line" in metallurgical microscopes and were essential for quality work. In the late 1960s and early 1970s, manufacturers developed metallographs that were easier to use from the eyepiece viewing position. The temperamental carbon-arc light sources were replaced by xenon-arc lamps. The unwieldy bellows systems for altering magnification were replaced by zoom systems. Vertical illuminators, previously using single-position objectives, were equipped with four-to-six position rotatable nosepiece turrets to minimize objective handling. The light path was deflected so that the film plane was conveniently near at hand. Universal-type vertical illuminators and objective lenses were introduced so that the illumination mode could be readily switched from bright field to dark field, polarized light, or differential interference contrast (DIC). Such systems were very attractive in that separate vertical illuminators and objectives were no longer needed for each illumination mode and handling was eliminated. Exposure meters were also added at this time, chiefly as a result of the rising popularity of instant films where such devices are needed to minimize film wastage. However, by the late 1970s, these large metallographs had become very expensive, too expensive for most laboratories.

In 1979, microscope manufacturers began to introduce very high quality, reasonably priced, compact metallographs. These microscopes can be obtained with a wide variety of objective lens types and auxiliary accessories to meet the metallographer's needs. They are available with universal vertical illuminators that permit easy switching from one illumination mode to another using the same set of objective lenses. Furthermore, the manufacturers have introduced new glass compositions and lens formulations, generally by computer-aided design, for improved image contrast and brightness. Tungsten–halogen filament lamps have largely replaced xenon-arc lamps as the preferred light source.

Microscope Components

Light Sources

The amount of light lost during passage from the source through a reflecting-type microscope is appreciable because of the intricate path the light follows. For this reason, it is generally preferable that the intensity of the source be high, especially for photomicroscopy. Several types of light sources are used including tungsten-filament lamps, tungsten–halogen lamps, quartz–halogen lamps, and xenon-arc bulbs.

Tungsten-filament lamps generally operate at low voltage and high current. They are widely used for visual examination because of their low cost and ease of operation.

Tungsten–halogen lamps are the most popular light source today due to their high light intensity. They produce good color micrographs when tungsten-corrected films are employed. Light intensity can be varied easily to suit the viewing conditions by adjusting a rheostat, a distinct advantage over arc lamps.

Xenon-arc lamps produce extremely high intensity, and their uniform spectra and daylight color temperature makes them suitable for color photomicrography. The first xenon lamps produced ozone, but modern units have overcome this problem. Light output is constant and can only be reduced using neutral-density filters.

Condenser

An adjustable lens free of spherical aberration and coma is placed in front of the light source to focus the light at the desired point in the optical path. A field diaphragm is placed in front of this lens to minimize internal glare and reflections within the microscope. The field diaphragm is stopped down to the edge of the field of view.

A second adjustable-iris diaphragm, the aperture diaphragm, is placed in the light path before the vertical illuminator. Opening or closing this diaphragm alters the amount of light and the angle of the cone of light entering the objective lens. The optimum setting for the aperture varies with each objective lens and is a compromise among image contrast, sharpness, and depth of field. As magnification increases, the aperture diaphragm is stopped down. Opening this aperture increases image sharpness, but reduces contrast; closing the aperture increases contrast, but impairs image sharpness. The aperture diaphragm should not be used for reducing light intensity. It should be adjusted only for contrast and sharpness.

Filters

Light filters are used to modify the light for ease of observation, for improved photomicroscopy, or to alter contrast. Neutral-density filters are used to reduce the light intensity uniformly across the visible spectrum. Various neutral-density filters, with a range of $\sim$85% of to 0.01% transmittance, are available. Most light microscopes offer selection of at least two such filters.

Selective filters are used to balance the color temperature of the light source to that of the film. This is often necessary for faithful reproduction of color images, depending on the light source used and the film type. A green or yellow-green filter is widely used in black-and-white photography to reduce the effect of lens defects on image quality. Most objectives, particularly the low-cost achromats, require such filtering for best results.

Polarizing filters are used to produce plane-polarized light (one filter) or crossed-polarized light (two filters rotated to produce extinction) for examinations of metals with noncubic crystallographic structures.

Objectives

The objective lens forms the primary image of the microstructure and is the most important component of the light microscope. The objective lens collects as much light as possible from the specimen and combines this light to produce the image. The numerical aperture (NA) of the objective, a measure of the light-collecting ability of the lens, is defined as

$$\mathrm{NA} = n \sin \alpha$$

where n is the minimum refraction index of the material (air or oil) between the specimen and the lens, and α is the half-angle of the most oblique light rays that enter the front lens of the objective. Light-collecting ability increases with α. The setting of the aperture diaphragm will alter the NA of the condenser and therefore the NA of the system.

The most commonly used objective is the achromat, which is corrected spherically for one color (usually yellow-green) and for longitudinal chromatic aberration for two colors (usually red and green). Therefore, achromats are not suitable for color photomicroscopy, particularly at higher magnifications. Use of a yellow-green filter yields optimum results. However, achromats do provide a relatively long working distance, that is, the distance from the front lens of the objective to the specimen surface when in focus. The working distance decreases as the objective magnification increases. Most manufacturers make long-working-distance objectives for special applications, for example, in hot-stage microscopy. Achromats are usually strain free, which is important for polarized light examinations. Because they contain fewer lenses than other more highly corrected lenses, internal reflection losses are minimized.

Semiapochromatic or fluorite objectives provide a higher degree of correction of spherical and chromatic aberration. Therefore, they produce higher quality color images than achromats. The apochromatic objectives have the highest degree of correction, produce the best results, and are more expensive. Plano objectives have extensive correction for flatness

of field, which reduces eyestrain, and are often found on modern microscopes.

With parfocal lens systems, each objective on the nosepiece turret will be nearly in focus when the turret is rotated, preventing the objective front lens from striking the specimen when lenses are switched. Many objectives also are spring loaded, which helps prevent damage to the lens. This is more of a problem with high-magnification objectives, because the working distance can be very small.

Certain objectives are designed for use with oil between the specimen and the front lens of the objective. However, oil-immersion lenses are rarely used in metallography, because the specimen and lens must be cleaned after use. However, they do provide higher resolutions that can be achieved when air is between the lens and specimen. In the latter case, the maximum possible NA is 0.95; oil-immersion lenses produce a 1.3 to 1.45 NA, depending on the lens and the oil used. Objective magnifications from $\sim$25 to 200$\times$ are available, depending upon the manufacturer. Use of oil also improves image contrast, which is valuable when examining low-reflectivity specimens, such as coal or ceramics.

Eyepieces

The eyepiece, or ocular, magnifies the primary image produced by the objective; the eye can then use the full resolution capability of the objective. The microscope produces a virtual image of the specimen at the point of most distinct vision, generally 250 mm (10 in.) from the eye. The eyepiece magnifies this image, permitting achievement of useful magnifications. The standard eyepiece has a 24-mm-diameter field of view, whereas wide-field eyepieces for plano-type objectives have a 30-mm-diameter field of view, which increases the usable area of the primary image. Today, the wide-field plano-objective is standard on nearly all metallurgical microscopes.

The simplest eyepiece is the Huygenian, which is satisfactory for use with low- and medium-power achromat objectives. Compensating eyepieces are used with high NA achromats and the more highly corrected objectives. Because some lens corrections are performed using these eyepieces, the eyepiece must be matched with the type of objective used. The newer, infinity-corrected microscopes do not perform corrections in the eyepieces, but in a tube lens. So, eyepieces are simpler in the infinity-corrected microscopes.

Eye clearance is the distance between the eye lens of the ocular and the eye. For most eyepieces, the eye clearance is 10 mm or less – inadequate if the microscopist wears glasses. Simple vision problems, such as near-sightedness, can be accommodated using the fine focus adjustment. The microscope cannot correct vision problems such as astigmatism, and glasses must be worn. High-eyepoint eyepieces are available to provide an eye clearance of $\sim$20 mm, necessary for eyeglasses.

Eyepieces are commonly equipped with various reticles or graticules for locating, measuring, counting, or comparing microstructures. The eyepiece enlarges the reticle or graticule image and the primary image. Both images must be in focus simultaneously. Special eyepieces are also produced to permit more accurate measurements than can be made with a graticule scale. Examples are the filar-micrometer ocular or screw-micrometer ocular. Such devices can be automated to produce a direct digital readout of the measurement, which is accurate to $\sim$1 μm.

A 10$\times$ magnification eyepiece is usually used, to obtain standard magnifications; some systems require other magnifications; such as 6.3$\times$. However, higher power eyepieces, such as 12, 15, 20, or 25$\times$, are also useful in certain situations. The overall magnification is found by multiplying the objective magnification, M_o, by the eyepiece magnification, M_e. If bellows or a zoom system is also used, the magnification should be altered accordingly.

Stage

A mechanical stage is provided for focusing and moving the specimen, which is placed on the stage and secured using clips. The stage of an inverted microscope has replaceable center-stage plates with different size holes. The polished surface is placed over the hole for viewing. However, the entire surface cannot be viewed, unless the specimen is smaller than the hole and it is mounted. At high magnifications, it may not be possible to focus the objective near the edge of the hole due to the restricted working distance.

Using the upright microscope, the specimen is placed on a slide on the stage. Because the polished surface must be perpendicular to the light beam, clay is placed between the specimen bottom and the slide. A piece of lens tissue is placed over the polished surface, and the specimen is pressed into the clay using a leveling press. However, pieces of tissue may adhere to the specimen surface. An alternative, particularly useful with mounted specimens, is to use a ring instead of tissue to flatten the specimen. Aluminum or stainless steel ring forms of the same size as the mounts (flattened slightly in a vise) will seat on the mount rather than on the specimen.

The upright microscope allows viewing of the entire surface with any objective, and the operator can

see which section of the specimen is being viewed – a useful feature when examining specific areas on coated specimens, welds, and other specimens where specific areas are to be examined. For mounted specimens, an auto-leveling stage holder can eliminate leveling specimens on clay.

The stage must be rigid to eliminate vibrations. Stage movement, controlled by *x*- and *y*-micrometers, must be smooth and precise; rack and pinion gearing is normally used. Many stages have scales for measuring the distances in the *x*- and *y*-directions. The focusing controls often contain rulings for estimating vertical movement. Some units have motorized stages and focus controls.

A circular rotatable stage plate may facilitate polarized light examination. Such stages, common for mineralogical or petrographic studies, are graduated to permit measuring the angle of rotation. A rectilinear stage is generally placed on top of the circular stage.

Stand

Bench microscopes require a rigid stand, particularly if photomicroscopy is performed on the unit. The various pieces of the microscope are attached to the stand when assembled. In some cases, the bench microscope is placed on a separate stand that also holds the photographic system.

Resolution

To see microstructural detail, the optical system must produce adequate resolution, or resolving power, and adequate image contrast. If resolution is acceptable but contrast is lacking, detail cannot be observed. In general, the ability to resolve two points or lines separated by distance d is a function of the wavelength, λ, of the incident light and the numerical aperture, NA, of that objective.

$$d = \frac{k_\lambda}{\text{NA}}$$

where k is 0.5 or 0.61. **Figure 2** illustrates this relationship for $k = 0.61$ and four light wavelengths. Other formulas have also been reported. This equation does not include other factors that influence resolution, such as the degree of correction of the objectives and the visual acuity of the microscopist. It was based on the work of Abbe under conditions not present in metallography, such as self-luminous points, perfect black-white contrast, transmitted-light examination, an ideal point-light source, and absence of lens defects.

Using the above equation, the limit of resolution for an objective with an NA of 1.3 is $\sim 0.2\,\mu\text{m}$. To see lines or points spaced 0.2 μm apart, the required magnification must be determined by dividing the resolving power of the human eye, which is difficult to determine under observation conditions. Abbe used a value of 0.3 mm at a distance of 250 mm – the distance from the eye for optimum vision. For light with a mean wavelength of 0.55 μm, the required magnification is 1100 times the NA of the objective. This is the origin of the $1000 \cdot \text{NA}$ rule for the maximum useful magnification. Any magnification above $1000 \cdot \text{NA}$ is termed "empty," or useless.

Strict adherence to the $1000 \cdot \text{NA}$ rule should be questioned, considering the conditions under which it was developed, which are certainly far different

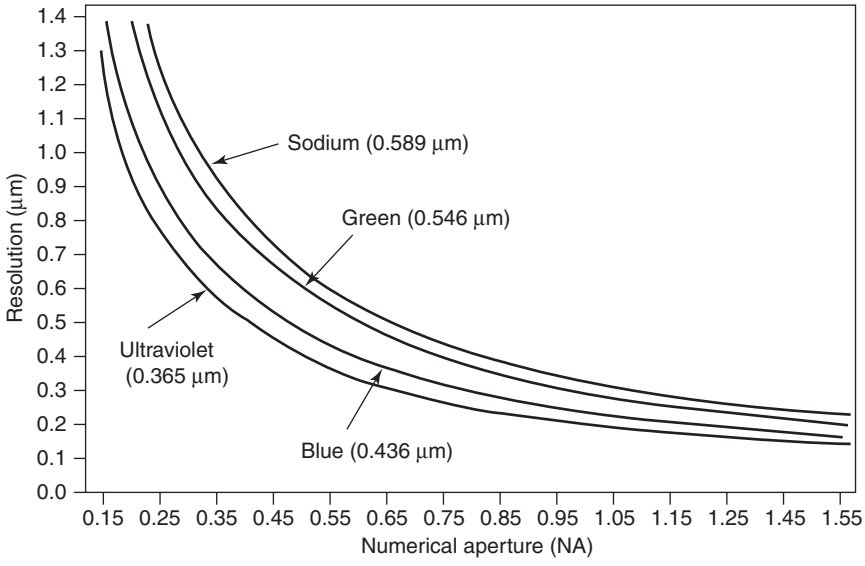


Figure 2 Influence of objective numerical aperture and wavelength on the resolution of the light microscope. (Reproduced with permission from Buehler, Ltd.)

from those encountered in metallography. According to the Abbe analysis, for a microscopist with optimum 20/20 vision and for optimum contrast conditions and a mean light wavelength of 550 nm, the lowest magnification that takes full advantage of the NA of the objective is 500 times the NA. This establishes a useful minimum magnification to be used with a given objective. It has been suggested that the upper limit of useful magnification for the average microscopist is $2200 \cdot \text{NA}$, not $1000 \cdot \text{NA}$.

Depth of Field

Depth of field is the distance along the optical axis over which image details are observed with acceptable clarity. Those factors that influence resolution also affect depth of field, but in the opposite direction. Therefore, a compromise must be reached between these two parameters, which is more difficult as magnification increases. This is one reason light etching is preferred for high-magnification examination. The depth of field, d_f, can be estimated from:

$$d_f = \lambda \frac{(n^2 - \text{NA}^2)^{1/2}}{\text{NA}^2}$$

where n is the refractive index of the medium between the specimen and the objective ($n \approx 1.0$ for air), λ is the wavelength of light, and NA is the numerical aperture. This equation shows that depth of field increases as the NA decreases and when longer wavelength light is used, as shown in **Figure 3**.

Imaging Modes

Most microscopical studies of metals are made using bright-field illumination. In addition to this type of illumination, several special techniques (oblique illumination, dark-field illumination, DIC microscopy, and polarized-light microscopy) have particular applications for metallographic studies.

Nearly all microscopes using reflected or transmitted light employ Köhler illumination because it provides the most intense, most even illumination possible with standard light sources. The reflected light microscope has two adjustable diaphragms: the aperture diaphragm and the field diaphragm, located between the lamp housing and the objective. Both are adjusted to improve illumination and the image. To obtain Köhler illumination, the image of the field diaphragm must be brought into focus on the specimen plane. This normally occurs automatically when the microstructural image is brought into focus. The filament image must also be focused on the aperture diaphragm plane. This produces uniform illumination of the specimen imaged at the intermediate image plane and magnified by the eyepiece.

Bright Field

In bright-field illumination, the surface of the specimen is normal to the optical axis of the microscope, and white light is used. A ray diagram for bright-field illumination is illustrated in **Figure 4**. Light that passes through the objective and strikes a region of the specimen surface that is perpendicular to the beam will be reflected back up the objective through the eyepieces to the eyes where it will appear to be bright or white. Light that strikes grain boundaries, phase boundaries, and other features not perpendicular to the optical axis will be scattered at an angle and will not be collected by the objective. These

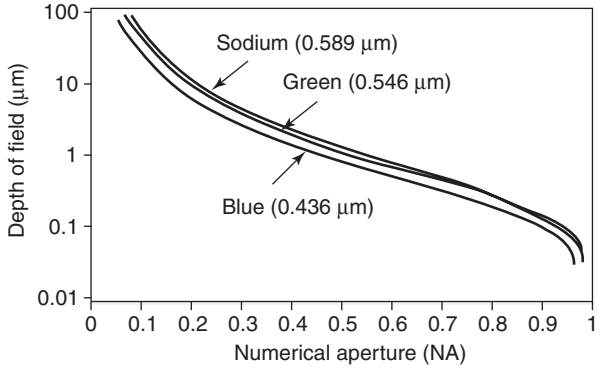


Figure 3 Influence of the objective numerical aperture and wavelength on the depth of field of the light microscope. (Reproduced with permission from Buehler, Ltd.)

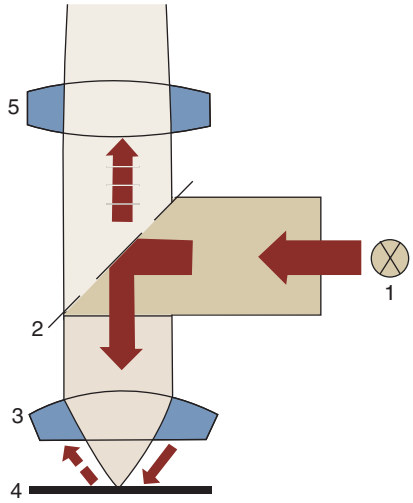


Figure 4 Schematic showing the light path in bright-field illumination (1. light source; 2. color-neutral beam splitter; 3. reflected-light objective; 4. sample surface; 5. tube lens). (Reproduced with permission from Buehler, Ltd.)

regions will appear to be dark or black in the image. Bright-field is the most common mode of illumination used by metallographers.

Oblique Illumination

The surface relief of a metallographic specimen can be revealed using oblique illumination. This involves offsetting the condenser lens system or, as is more usually done, moving the condenser aperture to a position slightly off the optical axis. Although it should be possible to continually increase the contrast achieved by oblique illumination by moving the condenser farther and farther from the light axis, the numerical aperture of a lens is reduced when this happens because only a portion of the lens is used. For this reason, there is a practical limit to the amount of contrast that can be achieved. Illumination also becomes uneven as the degree of "obliqueness" increases. Since DIC systems have been available, oblique illumination is rarely offered as an option on new microscopes.

Dark Field

Another method that can be used to distinguish features not in the plane of the polished-and-etched surface of a metallographic specimen is dark-field (also called dark ground) illumination. This type of illumination (see **Figure 5** for a ray diagram) gives contrast completely reversed from that obtained with bright-field illumination – the features that are light in bright-field will be dark in dark-field, and those that are dark in bright-field will be light, appearing to be self-luminous, in dark-field. This highlighting of angled surfaces (pits, cracks, or etched grain boundaries) allows more positive identification of their nature than can be derived from a black image with bright-field illumination. Due to the high image contrast obtained and the brightness associated with features at an angle to the optical axis, it is often possible to see details not observed with bright-field illumination.

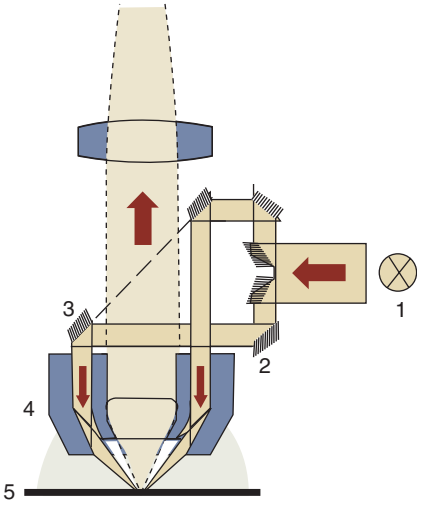


Figure 5 Schematic showing the light path in dark-field illumination (1. light source; 2. mirror step assembly; 3. mirror with an oval hole; 4. reflected-light objective; 5. sample surface). (Reproduced with permission from Buehler, Ltd.)

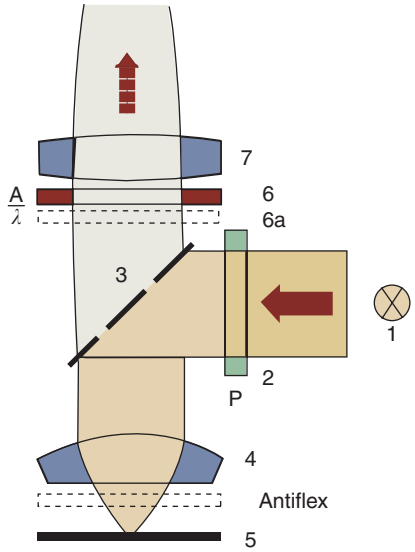


Figure 6 Schematic showing the light path in polarized light with an optional sensitive tint (lambda plate) filter (1. light source; 2. polarizer; 3. control-neutral beam splitter; 4. reflected-light objective; 5. sample surface; 6. A. analyzer; 6a. λ. optional lambda plate; 7. tube lens). (Reproduced with permission from Buehler, Ltd.)

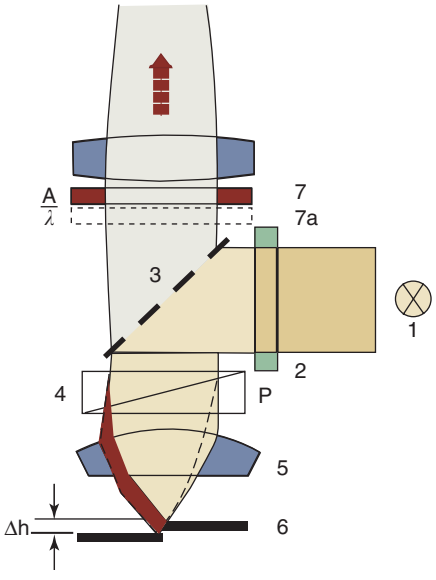


Figure 7 Schematic showing the light path in DIC illumination (1. light source; 2. polarizer; 3. control-neutral beam splitter; 4. birefringent prism, DIC prism; 5. reflected-light objective; 6. sample surface; 7. analyzer; 7a. lambda plate). (Reproduced with permission from Buehler, Ltd.)

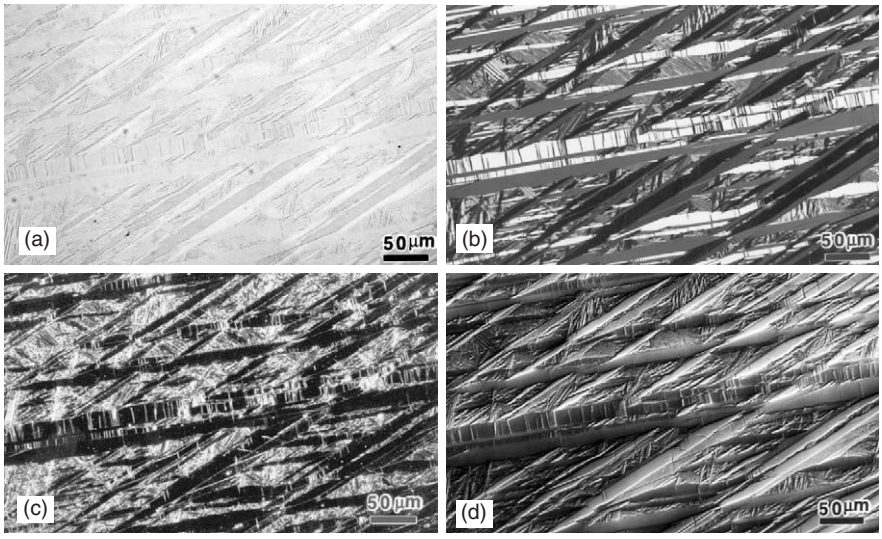


Figure 8 Microstructure of an as-polished specimen of Cu–11.8%Al quenched from the beta field to produce martensite viewed with: (a) bright-field, (b) polarized light, (c) dark-field, and (d) Nomarski DIC. Micrographs at 200×. (Reproduced with permission from Buehler, Ltd.)

Polarized-Light Microscopy

Because many metals and metallic and nonmetallic phases are optically anisotropic, polarized light is particularly useful in metallography. Polarized light is obtained by placing a polarizer (usually a Polaroid filter) in front of the condenser lens of the microscope and placing an analyzer (another Polaroid filter) before the eyepiece, as illustrated in **Figure 6**. The polarizer produces plane-polarized light that strikes the surface and is reflected through the analyzer to the eyepieces. If an anisotropic metal is examined with the analyzer set 90° to the polarizer, the grain structure will be visible. However, viewing of an isotropic metal (cubic metals) under such conditions will produce a dark, "extinguished" condition (complete darkness is not possible using Polaroid filters). Polarized light is particularly useful in metallography for revealing the grain structure and twinning in anisotropic metals and alloys and for identifying anisotropic phases and inclusions.

DIC Microscopy

When crossed polarized light is used along with a double quartz prism (Wollaston prism) placed between the objective and the vertical illuminator (**Figure 7**), two light beams are produced which exhibit coherent interference in the image plane. This leads to two slightly displaced (laterally) images differing in phase ($\lambda/2$) that produces a height contrast. The image produced reveals topographic detail somewhat similar to that produced by oblique illumination but without the loss of resolution. Images can be viewed with natural colors similar to those observed in bright-field, or artificial coloring can be introduced by adding a sensitive tint plate.

As an example of the use of these different imaging modes, **Figure 7** shows the microstructure of an aluminum bronze specimen that was water-quenched from the beta field forming martensite. The specimen was prepared but not etched. A minor amount of relief was introduced in final polishing. **Figure 8a** shows the surface in bright-field illumination. Due to the relief present, a faint image of the structure is visible. Dark-field (**Figure 8c**) and DIC (**Figure 8d**) make use of this relief to reveal much more detail than in bright-field. However, as the martensite is noncubic in crystal structure, it responds well to crossed-polarized light (**Figure 8b**). Of course, not every material can be examined using all four illumination modes, but there are many cases where two or more modes can be effectively utilized to reveal more information than given by bright-field alone.

See also: Confocal Optical Microscopy; Scanning Near-Field Optical Microscopy; Transmission Electron Microscopy.

PACS: 07.60.Pb; 42.15.Eq; 42.15.Dp; 42.79.Ci

Further Reading

Gifkins RC (1970) *Optical Microscopy of Metals*. NY: American Elsevier.

Richardson JH (1971) *Optical Microscopy for the Materials Sciences*. NY: Dekker.

Vander Voort GF (1984) *Metallography: Principles and Practice*. NY: McGraw-Hill; ASM International, Materials Park, OH, 1999.

Vander Voort GF (ed.) (2004) Metallography and microstructures. *ASM Handbook*, vol. 9. Materials Park, OH: ASM.

Optical Properties of Materials

G Guizzetti and M Patrini, Università degli Studi di Pavia, Pavia, Italy

Introduction

The optical properties of materials are of major interest from both the applicative and the fundamental points of view. They are indeed quantified by macroscopic functions, such as the refractive index, the absorption coefficient and reflectivity, a knowledge of which is necessary in designing and implementing optic or optoelectronic components and devices. They explain the nature of the material (e.g., allowing the distinction between metals, semiconductors, and insulators), and of the material homogeneity, crystal structure, and composition. Moreover, the macroscopic optical response should be related to microscopic properties, such as the electronic and vibrational structures, by means of classical models (e.g., Drude–Lorentz model) and quantum treatments (e.g., for interband and intraband optical transitions).

In this article, the optical properties of macroscopic materials based only on classical arguments are described. Starting from Maxwell's equations in a medium, the propagation of electromagnetic (em) waves in isotropic, linear, and homogeneous materials is treated. To this aim, complex macroscopic quantities as the dielectric function, the optical conductivity, and the refractive index are introduced, together with the dispersion relations which they obey. Some hints at their extensions to anisotropic and nonlinear materials are given.

The reflection and transmission properties of em waves are then analyzed at the interface between different materials, and the appropriate functions are defined for unpolarized light waves. Their extension for polarized waves follows, with the addition of the ellipsometric functions which are particularly interesting from the application point of view. Finally, links between the different optical functions and brief hints at their measurement techniques are given.

Maxwell's Equations in a Medium

The theoretical framework for the optical properties of a material is given by Maxwell's equations which describe the propagation of an em wave in a medium and which in SI units are

$$\nabla \cdot \boldsymbol{D} = \rho, \qquad \nabla \cdot \boldsymbol{B} = 0$$
$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \nabla \times \boldsymbol{H} = \boldsymbol{j} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad [1]$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic fields, $\boldsymbol{D}$ the electric displacement, $\boldsymbol{B}$ the magnetic induction, ρ the external introduced charge density, $\boldsymbol{j}$ the current density.

The relations between the electric and magnetic fields are

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \qquad \boldsymbol{H} = \frac{\boldsymbol{B}}{\mu_0} - \boldsymbol{M}\partial \qquad [2]$$

where $\boldsymbol{P}$ is the polarization (electric dipole moment per unit volume) and $\boldsymbol{M}$ the magnetization (magnetic moment per unit volume), ε_0 and μ_0 the vacuum permittivity and magnetic permeability.

The current density $\boldsymbol{j}$ in the more general case contains two terms

$$\boldsymbol{j} = \boldsymbol{j}_{\text{ext}} + \boldsymbol{j}_{\text{cond}} \qquad [3]$$

where $\boldsymbol{j}_{\text{ext}}$ is the current density introduced into the medium from an external source and $\boldsymbol{j}_{\text{cond}}$ is the conduction current which is induced by a driving em field and which arises from the motion of charges in phase with the field.

All these quantities are macroscopic, that is, the average of their microscopic counterparts over a volume ΔV with linear dimensions small compared to the wavelength λ of the em wave, but large enough to contain many atoms. Then the medium can be treated as continuous from the viewpoint of optical properties, although an ideally perfect crystal is not rigorously homogeneous on a microscopic scale. For instance, the internal charge density has crystal periodicity and varies sharply in the unit cell.

In the following, homogeneous, isotropic, nonmagnetic ($\boldsymbol{B} = \mu_0 \boldsymbol{H}$) materials are treated in the absence of external charges ($\rho = 0$) and currents ($\boldsymbol{j}_{\text{ext}} = 0$). Moreover, the linear response regime is considered, that is, the variations in the internal charge density and in the current density induced by a driving em field of small intensity are proportional to the field. In these assumptions, the appropriate constitutive equations of the medium are:

$$\boldsymbol{P} = \varepsilon_0 \chi \boldsymbol{E}, \qquad \boldsymbol{j}_{\text{cond}} = \sigma \boldsymbol{E} \qquad [4]$$

where χ is the electric susceptibility and σ is the conductivity. The susceptibility χ and the conductivity σ

contain the response of the medium to the em field E. They both appear at the right-hand side of the last of Maxwell's equations [1] and represent two different currents, the conduction current $\boldsymbol{j}_{\text{cond}}$ and the polarization current $\boldsymbol{j}_{\text{pol}} = \partial \boldsymbol{P}/\partial t = \varepsilon_0 \chi(\partial \boldsymbol{E}/\partial t)$. Since periodic solutions of Maxwell's equations are considered, $\boldsymbol{j}_{\text{cond}}$ and $\boldsymbol{j}_{\text{pol}}$ are $\pi/2$ out-of-phase: $\boldsymbol{j}_{\text{cond}}$ is in-phase with $\boldsymbol{E}$ and gives a time-average power absorption per unit volume $\boldsymbol{P}_{\text{a}} = \langle \boldsymbol{j} \cdot \boldsymbol{E} \rangle = \sigma \langle \boldsymbol{E}^2 \rangle$; $\boldsymbol{j}_{\text{pol}}$, instead, is $\pi/2$ out-of-phase with $\boldsymbol{E}$ and does not provide any absorption. It is stressed that $\boldsymbol{j}_{\text{cond}}$ and $\boldsymbol{j}_{\text{pol}}$, and thus χ and σ, are not independent quantities because they have the same origin: the motion of charges (free or bound) induced by the em field. As it is shown below, the real quantities χ and σ are better represented by the real and imaginary parts of a single complex quantity. Moreover, the conductivity σ which appears in eqn [4] is sometimes called optical conductivity, which in the limit of zero-frequency equals the usual DC electrical conductivity, that is, for a static field $\boldsymbol{E}\,\sigma \neq 0$ only in conductors, while for time-variable fields $\boldsymbol{E}(t)\,\sigma \neq 0$ for both conductors and insulators.

An alternative parameter which characterizes a material is defined through eqn [2] by

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} = \varepsilon_0 \boldsymbol{E} + \varepsilon_0 \chi \boldsymbol{E} = \varepsilon \boldsymbol{E} \qquad [5]$$

where the quantity $\varepsilon = \varepsilon_0(1 + \chi)$ is termed the dielectric constant or dielectric function (as is evident below, the second nomenclature is more appropriate when dealing with time-dependent fields). The dimensionless quantity $\varepsilon_{\text{r}} = \varepsilon/\varepsilon_0 = (1 + \chi)$ is the relative dielectric function.

ε and σ are the parameters which allow one to completely characterize the optical response of the material. They describe the velocity of the wave in the medium and its spatial attenuation, that is, the energy absorbed by the medium, respectively. ε and σ are functions of frequency (this property is called time dispersion) and the wave vector (spatial dispersion) of the em wave. Moreover, they depend also on internal/external physical parameters, such as temperature, pressure, magnetic or electric static fields.

In the aforementioned hypotheses and from Maxwell's equations, the following wave equation can be derived:

$$\nabla^2 \boldsymbol{E} = \varepsilon_{\text{r}} \varepsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} + \sigma \mu_0 \frac{\partial \boldsymbol{E}}{\partial t} \qquad [6]$$

This is the so-called d'Alembert equation for a wave propagating in a generic absorbing medium. Solutions of eqn [6] are necessarily restricted to transverse waves because $\nabla \cdot \boldsymbol{E} = 0$ in the absence of a net charge density ρ. In vacuum ($\varepsilon_{\text{r}} = 1,\ \sigma = 0$), the wave velocity coincides with the velocity of light $c = (\sqrt{\varepsilon_0 \mu_0})^{-1}$, while in a nonabsorbing medium ($\sigma = 0$) the phase velocity is $v = (\sqrt{\varepsilon_{\text{r}} \varepsilon_0 \mu_0})^{-1} = c/\sqrt{\varepsilon_{\text{r}}} = c/n$, where $n = \sqrt{\varepsilon_{\text{r}}} = c/v$ is the refractive index of the medium. The wave amplitude attenuation is described by the second term on the right-hand side of eqn [6].

The Complex Dielectric Function

It must be emphasized that, in the case of time-dependent em fields, there is generally a phase shift between the motion of the charge carriers and the electric field $\boldsymbol{E}$ (similar to what happens in AC circuits). In such cases the complex formalism is particularly convenient to describe these phase shifts, in addition to describing dispersion and absorption effects at the same time. Complex fields are usually defined, the real parts of which have physical meaning and are the ones involved in Maxwell's equations. For instance, $\boldsymbol{E}(\boldsymbol{r}, t) = \text{Re}[\tilde{\boldsymbol{E}}(\boldsymbol{r}, t)]$, where the tilde distinguishes the complex quantity from its real counterpart.

Since the experimental measurements of the optical properties of solids are usually conducted with monochromatic light, a solution of eqn [6] is considered in the form of a monochromatic plane wave:

$$\tilde{\boldsymbol{E}}(\boldsymbol{r}, t) = \tilde{\boldsymbol{E}}_0 \exp[\text{i}(\tilde{\boldsymbol{q}} \cdot \boldsymbol{r} - \omega t)] \qquad [7]$$

where $\tilde{\boldsymbol{E}}$ is the complex electric field, $\tilde{\boldsymbol{E}}_0 = \boldsymbol{E}_0 \text{e}^{\text{i}\varphi_0}$ is the polarization vector with an arbitrary phase φ_0, $\boldsymbol{r}$ is the position vector with respect to an arbitrary coordinate system, $\omega = 2\pi v$ is the angular frequency, and $\tilde{\boldsymbol{q}} = \boldsymbol{q}_1 + \text{i}\boldsymbol{q}_2$ is the complex wave vector of the em wave in the medium. The absolute value of the wave vector in vacuum is $q = 2\pi/\lambda = \omega/c$. In fact, $\tilde{\boldsymbol{q}}$ describes the phase and spatial attenuation of the em wave, its explicit form being written as

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{r}, t) &= \text{Re}\ \tilde{\boldsymbol{E}}(\boldsymbol{r}, t) \\ &= \text{Re}\left[\tilde{\boldsymbol{E}}_0 \exp[\text{i}(\tilde{\boldsymbol{q}} \cdot \boldsymbol{r} - \omega t)]\right] \\ &= \text{Re}\left[\tilde{\boldsymbol{E}}_0 \exp[\text{i}(\boldsymbol{q}_1 \cdot \boldsymbol{r} - \omega t)]\right] \exp(-\boldsymbol{q}_2 \cdot \boldsymbol{r}) \\ &= \boldsymbol{E}_0 \cos(\boldsymbol{q}_1 \cdot \boldsymbol{r} - \omega t + \varphi_0) \exp(-\boldsymbol{q}_2 \cdot \boldsymbol{r}) \end{aligned} \qquad [8]$$

In general, $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are perpendicular to the constant phase planes and to the constant amplitude planes, respectively. If $\boldsymbol{q}_1 || \boldsymbol{q}_2$, the wave is homogeneous.

By inserting the monochromatic plane wave expression (eqn [7]) in eqn [6], the wave dispersion relation is obtained:

$$\tilde{\boldsymbol{q}} \cdot \tilde{\boldsymbol{q}} = \mu_0 \left(\varepsilon + \text{i}\frac{\sigma}{\omega}\right) \omega^2 = \mu_0 \tilde{\varepsilon} \omega^2 \qquad [9]$$

where a complex dielectric function $\tilde{\varepsilon}$ has been introduced as

$$\tilde{\varepsilon} \equiv \varepsilon_1 + i\varepsilon_2 = \varepsilon + i\frac{\sigma}{\omega} \quad [10]$$

The real part $\varepsilon_1 = \varepsilon$ describes the wave dispersion while the imaginary part $\varepsilon_2 = \sigma/\omega$ accounts for the dissipation of energy by the material. Alternatively, the complex conductivity $\tilde{\sigma} \equiv \sigma_1 + i\sigma_2 = -i\omega\tilde{\varepsilon} = \sigma - i\omega\varepsilon$ is sometimes used instead of $\tilde{\varepsilon}$.

The possibility of a phase shift between $\tilde{\boldsymbol{D}}$ (or $\tilde{\boldsymbol{j}}$) and $\tilde{\boldsymbol{E}}$ is now easily included by defining $\tilde{\boldsymbol{D}} = \tilde{\varepsilon}\tilde{\boldsymbol{E}}$ (or $\tilde{\boldsymbol{j}} = \tilde{\sigma}\tilde{\boldsymbol{E}}$). Moreover, when considering an em field given by eqn [7], the displacement current $\partial\boldsymbol{D}/\partial t$ and the conduction current $\boldsymbol{j}_{\text{cond}}$ which appear in Maxwell's equations can be combined as follows:

$$\begin{aligned}\boldsymbol{j}_{\text{cond}} + \frac{\partial \boldsymbol{D}}{\partial t} &= \sigma\boldsymbol{E} + \varepsilon\frac{\partial \boldsymbol{E}}{\partial t} = \text{Re}[(\sigma - i\omega\varepsilon)\tilde{\boldsymbol{E}}] \\ &= \text{Re}[\tilde{\sigma}\tilde{\boldsymbol{E}}] = \text{Re}[-i\omega\tilde{\varepsilon}\,\tilde{\boldsymbol{E}}]\end{aligned} \quad [11]$$

Thus the response function $\tilde{\varepsilon}$ accounts not only for the displacement current $\partial\boldsymbol{D}/\partial t$, which is $\pi/2$ out-of-phase with $\boldsymbol{E}$ and does not dissipate energy, but also, in its imaginary part, for the conduction current $\boldsymbol{j}_{\text{cond}}$, which is in-phase with $\boldsymbol{E}$ and dissipates energy. The same holds for $\tilde{\sigma}$. When the optical response of a medium is calculated according to a specific model, either $\tilde{\varepsilon}$ or $\tilde{\sigma}$ must be introduced (but not both), since each one already includes both the in-phase and out-of-phase components.

In a general treatment, ε and σ are functions of space and time since the space–time representation of Maxwell's equations has been adopted. They are response functions or linear integral operators that connect the fields $\boldsymbol{D}(\boldsymbol{r}, t)$ and $\boldsymbol{j}(\boldsymbol{r}, t)$ with the electric field $\boldsymbol{E}(\boldsymbol{r}', t')$ at all other positions and all earlier times. Thus, in general,

$$\boldsymbol{D}(\boldsymbol{r}, t) = \int \mathrm{d}\boldsymbol{r}' \int_{-\infty}^{t} \varepsilon(\boldsymbol{r}, \boldsymbol{r}', t, t')\boldsymbol{E}(\boldsymbol{r}', t')\mathrm{d}t' \quad [12]$$

that is, $\boldsymbol{D}$ $(\boldsymbol{r}, t)$ and $\boldsymbol{E}$ $(\boldsymbol{r}, t)$ are not proportional to each other, while this is true for their Fourier components. In fact, fields and sources should be decomposed into a complete set of monochromatic plane waves, for example

$$\begin{aligned}\boldsymbol{E}(\boldsymbol{r}, t) = \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{q} \int_{-\infty}^{+\infty} \tilde{\boldsymbol{E}}(\boldsymbol{q}, \omega) \\ \times \exp[i(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)]\mathrm{d}\omega\end{aligned} \quad [13]$$

where $\tilde{\boldsymbol{E}}$ $(\boldsymbol{q}, \omega)$ is the Fourier transform of $\boldsymbol{E}(\boldsymbol{r}, t)$, with

$$\begin{aligned}\tilde{\boldsymbol{E}}(\boldsymbol{q}, \omega) = \left(\frac{1}{2\pi}\right)^4 \int_{-\infty}^{+\infty} \mathrm{d}\boldsymbol{r} \int_{-\infty}^{+\infty} \boldsymbol{E}(\boldsymbol{r}, t) \\ \times \exp[-i(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)]\mathrm{d}t\end{aligned} \quad [14]$$

Operating the Fourier transform of eqn [12] and utilizing the convolution theorem, one obtains

$$\tilde{\boldsymbol{D}}(\omega, \boldsymbol{q}) = \tilde{\varepsilon}(\omega, \boldsymbol{q})\tilde{\boldsymbol{E}}(\omega, \boldsymbol{q}) \quad [15]$$

where the complex dielectric function $\tilde{\varepsilon}(\omega, \boldsymbol{q})$, Fourier transform of $\varepsilon(\boldsymbol{r}, t)$, shows up as the proportionality constant between the complex field components. An analogous relation $\tilde{\boldsymbol{j}}(\omega, \boldsymbol{q}) = \tilde{\sigma}(\omega, \boldsymbol{q})\tilde{\boldsymbol{E}}(\omega, \boldsymbol{q})$ can be obtained.

The frequency and wave vector dependence of $\tilde{\varepsilon}(\omega, \boldsymbol{q})$ and $\tilde{\sigma}(\omega, \boldsymbol{q})$ describe the time and the spatial dispersion of the material, respectively. The spatial dispersion is associated with the presence of natural lengths in the medium, such as the dimension of atoms and interatomic spacings; its effect is weak for wavelengths which are long as compared to the natural lengths. In the optical spectral range, the spatial dispersion should be neglected and $\tilde{\varepsilon}(\omega, \boldsymbol{q}) = \tilde{\varepsilon}(\omega)$; this is not possible at X-ray wavelengths. Moreover, it may be noted that the wave dispersion relation (eqn [9]), which binds ω and $\boldsymbol{q}$, does not imply, in general, a connection between time dispersion and spatial dispersion. In principle, it is possible to investigate separately the temporal response of a medium (and hence the time dispersion in ω) by subjecting it to a spatially uniform field that oscillates in time, or the spatial response (and hence the spatial dispersion in $\boldsymbol{q}$) with a static field that is spatially periodic.

In an anisotropic medium, the polarization $\boldsymbol{P}$ and $\boldsymbol{j}_{\text{cond}}$ generally lie in a direction different from that of $\boldsymbol{E}$. This situation can be handled by representing the complex dielectric function as a tensor $\tilde{\varepsilon}_{ij}$. The real and imaginary parts of $\tilde{\varepsilon}_{ij}$ are symmetric also for an anisotropic medium; thus, it is possible to put each of them into diagonal form in an appropriate set of coordinate axes (principal axes). The two sets do not coincide in general, until the crystal has symmetry at least as high as an orthorhombic one. ln the particular case of cubic crystal symmetry, $\tilde{\varepsilon}_{ij}$ reduces to a scalar quantity.

In a nonlinear dielectric medium, it is possible to expand the relation between $\boldsymbol{P}$ and $\boldsymbol{E}$ in a Taylor series about $\boldsymbol{E} = 0$. Thus $\boldsymbol{P} = \varepsilon_0(\chi\boldsymbol{E} + \chi^{(2)}\boldsymbol{E}^2 + \chi^{(3)}\boldsymbol{E}^3 + \cdots)$, where the $\chi^{(n)}$ coefficient describes the nth order nonlinear effect. The eqn [6] is not applicable to em waves in nonlinear media. However, Maxwell's equations can be used to derive a

nonlinear partial differential equation that these waves obey.

Finally, the real and imaginary parts of $\tilde{\varepsilon}$ are not independent quantities, being related to one another by the integral relations, the so-called Kramers–Kronig relations. These relations follow rigorously from the requirement of causality (i.e., there can be no effect before the cause) and apply to the real and imaginary part of whatever linear response function. Thus, the complex dielectric function $\tilde{\varepsilon}(\omega)$ obeys the following relations:

$$\begin{aligned}\varepsilon_1(\omega) &= \varepsilon_0 + \frac{2}{\pi}P\int_0^\infty \frac{\omega'\varepsilon_2(\omega')}{\omega'^2-\omega^2}\,\mathrm{d}\omega' \\ \varepsilon_2(\omega) &= -\frac{2\omega}{\pi}P\int_0^\infty \frac{\varepsilon_1(\omega')-\varepsilon_0}{\omega'^2-\omega^2}\,\mathrm{d}\omega'\end{aligned} \qquad [16]$$

where P stands for the principal value of the integral.

It may be recalled that $\tilde{\varepsilon}$ and $\tilde{\sigma}$ describe the macroscopic properties of materials, but they can be related to polarizabilities on the atomic scale and give reasons for the microscopic properties of condensed matter, for example, electronic and vibrational structures.

Complex Refractive Index

In analogy with the statement $n^2(\omega) = \varepsilon_\mathrm{r}(\omega)$ valid in nonabsorbing media, where n is the refractive index, it is useful to define a complex quantity, the so-called complex refractive index $\tilde{n}(\omega)$ that is related to the relative complex dielectric function $\tilde{\varepsilon}_\mathrm{r}(\omega) = \tilde{\varepsilon}(\omega)/\varepsilon_0$ by the relation:

$$\tilde{n}^2(\omega) = \tilde{\varepsilon}_\mathrm{r}(\omega) \qquad [17]$$

The real part of $\tilde{n}(\omega)$ coincides with the usual refractive index while the imaginary part is the extinction coefficient of the medium. Thus

$$\tilde{n}(\omega) = n(\omega) + \mathrm{i}k(\omega) \qquad [18]$$

and then the real and imaginary parts of $\tilde{\varepsilon}$, including the absorption contribution, should be written as

$$\varepsilon_{\mathrm{r}1} = n^2 - k^2, \qquad \varepsilon_{\mathrm{r}2} = 2nk \qquad [19]$$

Also, the inverse relations are derived as

$$n = \left\{\frac{1}{2}\left[\varepsilon_{\mathrm{r}1} + (\varepsilon_{\mathrm{r}1}^2 + \varepsilon_{\mathrm{r}2}^2)^{1/2}\right]\right\}^{1/2}, \qquad k = \frac{\varepsilon_{\mathrm{r}2}}{2n} \qquad [20]$$

The dispersion relation of eqn [9] assumes the following expression:

$$\tilde{\boldsymbol{q}}\cdot\tilde{\boldsymbol{q}} = \mu_0\tilde{\varepsilon}\omega^2 = \mu_0\varepsilon_0\tilde{\varepsilon}_\mathrm{r}\omega^2 = \frac{\omega^2}{c^2}\tilde{n}^2 \qquad [21]$$

If the wave is homogeneous, that is, $\boldsymbol{q}_1 \| \boldsymbol{q}_2$, it holds $\tilde{\boldsymbol{q}} = (\omega/c)\tilde{n}\boldsymbol{u}_q$, where $\boldsymbol{u}_q$ is the unit vector parallel to $\tilde{\boldsymbol{q}}$, and then $\boldsymbol{q}_1 = (\omega/c)n\boldsymbol{u}_q, \boldsymbol{q}_2 = (\omega/c)k\boldsymbol{u}_q$. Thus, eqn [7] should be written as

$$\begin{aligned}\tilde{\boldsymbol{E}}(\boldsymbol{r},t) = \tilde{\boldsymbol{E}}_0 &\exp\left[-\frac{\omega}{c}k\boldsymbol{u}_q\cdot\boldsymbol{r}\right] \\ &\times\exp\left[\mathrm{i}\left(\frac{\omega}{c}n\boldsymbol{u}_q\cdot\boldsymbol{r}-\omega t\right)\right]\end{aligned} \qquad [22]$$

The extinction coefficient k measures the attenuation of the wave amplitude with the propagation distance inside the medium, while n determines the phase velocity c/n in the medium. The penetration depth δ or classical skin depth, defined as the distance at which the field amplitude drops by $1/e$, is $\delta = c/\omega k$.

In anisotropic material, $\tilde{n}(\omega)$ is no more a scalar quantity, and as it is stated for $\tilde{\varepsilon}$, it should be a tensor. For example, in a uniaxial crystal (i.e., $\tilde{\varepsilon}_{xx} = \tilde{\varepsilon}_{yy} \neq \tilde{\varepsilon}_{zz}$) ordinary ($n_\mathrm{o}$) and extraordinary ($n_\mathrm{e}$) refractive indexes are defined, for light propagation in the xy plane and along the z direction, respectively.

Just as ε_1 and ε_2, n and k also satisfy Kramers–Kronig relations.

An additional important macroscopic quantity of the medium is the absorption coefficient α, which describes the relative decrease in the wave intensity I with unit distance (in the propagation direction, i.e., $\boldsymbol{u}_q \| \mathrm{d}\boldsymbol{r}$). Since the wave intensity (i.e., the power which is incident on the unit area perpendicular to $\boldsymbol{u}_q$) is $I = nc\varepsilon_0 E^2/2$, from eqn [22] it follows that

$$\alpha(\omega) = \frac{2k\omega}{c} = \frac{4\pi k}{\lambda} = \frac{\omega\varepsilon_{\mathrm{r}2}}{nc} \qquad [23]$$

The exponential attenuation of I after a propagating distance d accounts for the phenomenological Lambert–Beer law:

$$I = I_0\mathrm{e}^{-\alpha d} \qquad [24]$$

where I_0 is the incident wave intensity. Differently from n and k, α is a dimensional quantity and is normally measured in cm^{-1}. The αd coefficient that reflects both the physical and chemical properties, and the geometry of the medium is referred to as optical density.

Reflection and Transmission Properties

Unpolarized Light

While the refractive index n is measurable with a high degree of accuracy through experimental techniques that do not require absolute intensity measurements, such as minimum deviation, Brewster angle and interferometric techniques, the complete

determination of n and k (or α) of a medium generally requires the measurement of the intensity reflected I_r, transmitted I_t, and/or absorbed I_a with respect to the intensity I_0 of an em wave incident on the material surface. From the energy conservation law, it follows that

$$I_0 = I_r + I_t + I_a \qquad [25]$$

Dividing both sides by I_0, one obtains:

$$\mathscr{R} + \mathscr{T} + \mathscr{A} = 1 \qquad [26]$$

where $\mathscr{R}$, $\mathscr{T}$, and $\mathscr{A}$ are the reflectance, the transmittance and the absorbance of the medium, respectively. The connection between these quantities and the optical functions of the medium can be obtained by Maxwell's equations with appropriate surface boundary conditions, that is, the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$ and the normal components of $\boldsymbol{D}$ and $\boldsymbol{B}$ at the surface of the material are continuous.

Hereafter, only specularly reflected light is dealt with: this condition is well matched when surface irregularities and roughness have length scales much smaller than the light wavelength λ. In this case, diffuse reflection losses are negligible.

Consider the simplest case of a plane monochromatic wave, propagating in a medium with refractive index $\tilde{n}_0$, which is normally incident on the surface of a semi-infinite medium with refractive index $\tilde{n}$ giving rise to the phenomena of reflection and refraction. Applying the boundary conditions both to the electric field $\boldsymbol{E}$ and to the magnetic field $\boldsymbol{H}$ (with $\boldsymbol{E} = \mu_0 \boldsymbol{H} \times \boldsymbol{v}$ in a transverse electromagnetic wave), the tangential components of the incident (0), reflected (r) and refracted (i.e., transmitted, t) fields must obey

$$\begin{aligned} \tilde{E}_t &= \tilde{E}_0 + \tilde{E}_r \\ \tilde{n}\tilde{E}_t &= \tilde{n}_0\tilde{E}_0 - \tilde{n}_0\tilde{E}_r \end{aligned} \qquad [27]$$

By solving for $\tilde{E}_r$ and $\tilde{E}_t$, one has

$$\begin{aligned} \tilde{r} &= \frac{\tilde{E}_r}{\tilde{E}_0} = \frac{\tilde{n}_0 - \tilde{n}}{\tilde{n}_0 + \tilde{n}} \\ \tilde{t} &= \frac{\tilde{E}_t}{\tilde{E}_0} = \frac{2\tilde{n}_0}{\tilde{n}_0 + \tilde{n}} \end{aligned} \qquad [28]$$

where $\tilde{r}$ and $\tilde{t}$ are complex quantities in the more general case, and are named reflection and transmission Fresnel coefficients, respectively. Then, the reflected and transmitted waves change in amplitude and phase with respect to the incident one, and while $\tilde{r}$ change only in sign inverting the propagation direction, $\tilde{t}$ changes the absolute value. It is noted that $\tilde{t} = 1 + \tilde{r}$ and its value should exceed unity. Moreover, in transparent media (n_0 and n real) the phase change in reflection is π when $n_0 < n$.

At the interface of the two media, the reflected and transmitted intensity relative to the incident one are called reflectivity R and transmittivity T, and are given by:

$$\begin{aligned} R = \frac{I_r}{I_0} = |\tilde{r}|^2 &= \left|\frac{\tilde{n}_0 - \tilde{n}}{\tilde{n}_0 + \tilde{n}}\right|^2 \\ &= \frac{(n - n_0)^2 + (k - k_0)^2}{(n + n_0)^2 + (k + k_0)^2} \end{aligned} \qquad [29]$$

$$\begin{aligned} T = \frac{I_t}{I_0} &= \mathrm{Re}\left(\frac{\tilde{n}}{\tilde{n}_0}\right)|\tilde{t}|^2 \\ &= \mathrm{Re}\left(\frac{\tilde{n}}{\tilde{n}_0}\right)\left|\frac{2\tilde{n}_0}{\tilde{n}_0 + \tilde{n}}\right|^2 \end{aligned}$$

It may be noted that R and T are $\leqslant 1$ with $R + T = 1$, as expected by the energy conservation at the interface between different media.

When considering the reflection, absorption, and transmission properties of optical materials, care must be taken to differentiate between reflectivity and reflectance, transmittivity and transmittance. To make the distinction between these terms clear, consider a sample of material with $\tilde{n}$ and thickness d, immersed in air ($\tilde{n}_0 = 1$), as shown in **Figure 1**. Coming from air, the near-normally incident beam is split into reflected and transmitted fractions (as determined by R and T) at each interface, giving rise to secondary beams: reflectance and transmittance are obtained by summing up all the multiple-reflected and multiple-transmitted elements, which account also for light absorption along the thickness d. The

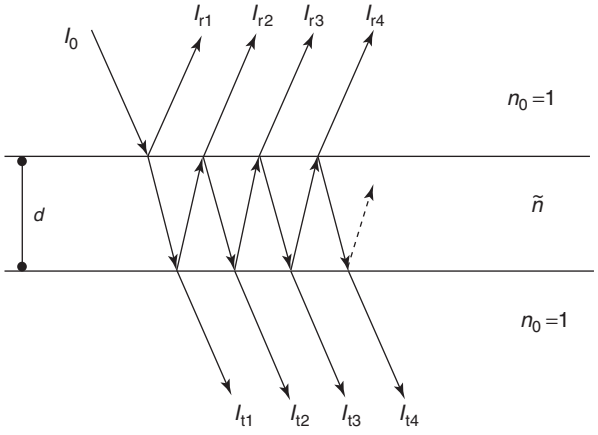


Figure 1 Reflected and transmitted beams generated by an em wave with intensity I_0 near-normally incident on a layer with complex refractive index $\tilde{n}$ immersed in air. Refracted and reflected angles are exaggerated for clearness.

total intensity reflected from the front surface is

$$I_r = I_{r1} + I_{r2} + \cdots + I_{rn} + \cdots$$
$$= (R + RT^2 e^{-2\alpha d} + R^3 T^2 e^{-4\alpha d} + \cdots) I_0 \quad [30]$$

Then, since $T = 1 - R$, the reflectance $\mathscr{R}$ is

$$\mathscr{R} = \frac{I_r}{I_0} = R + \frac{(1-R^2) R\, e^{-2\,\alpha d}}{1 - R^2\, e^{-2\,\alpha d}} \quad [31]$$

Similarly the transmittance $\mathscr{T}$ results in

$$\mathscr{T} = \frac{(1-R)^2\, e^{-\alpha d}}{1 - R^2\, e^{-2\alpha d}} \quad [32]$$

and the absorbance $\mathscr{A} = 1 - \mathscr{R} - \mathscr{T}$ is equal to

$$\mathscr{A} = \frac{(1-R)(1 - e^{-\alpha d})}{1 - R\, e^{-\alpha d}} \quad [33]$$

In the only case when the optical density $\alpha d \gg 1$, the reflectance $\mathscr{R}$ has only the first term in eqn [31] and coincides with the reflectivity R of the material; there is no more intensity transmitted. If the thickness of the sample is sufficiently low to allow transmittance, $\mathscr{T}$ depends on α, d, and R.

Polarized Light

In order to evaluate the Fresnel coefficients in the general case of a plane wave propagating in a medium with refractive index $\tilde{n}_0$ and incident at an angle ϕ_0 on a semi-infinite medium with refractive index $\tilde{n}$, it is convenient to consider separate cases with the electric vector polarized perpendicular to and parallel to the plane of incidence. There is no loss of generality in doing so, since any electric or magnetic vector may be resolved into components normal to and parallel to this plane. The subscripts s and p normally refer to the normal and parallel component, respectively.

The electric field $\tilde{\boldsymbol{E}}_0(\boldsymbol{r}, t)$ associated to an incident plane wave propagating in the z direction of a local coordinate system xyz (**Figure 2**) can be represented as

$$\tilde{\boldsymbol{E}}_0(\boldsymbol{r}, t) = \mathrm{Re}\{(\mathbf{i}\tilde{E}_{0s} + \boldsymbol{j}\tilde{E}_{0p}) \times \exp[\mathrm{i}(\tilde{\boldsymbol{q}} \cdot \boldsymbol{r} - \omega t)]\} \quad [34]$$

where $\tilde{E}_{0p}$ and $\tilde{E}_{0s}$ are the complex field components of $\tilde{\boldsymbol{E}}_0$ that describe both amplitude and phase dependence along the x (s direction) and y (p direction) axes of the system. The electric field of the reflected wave can be represented in another local coordinate system as

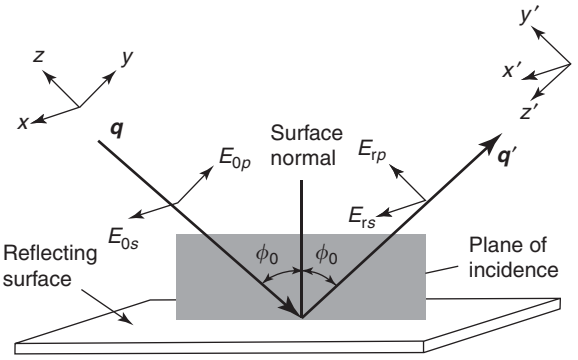


Figure 2 Reflection of an em wave from a surface. The field components E_s and E_p are shown, together with the coordinate axes for the incident and reflected waves. The plane of incidence contains the incoming wave vector $\boldsymbol{q}$ and the specularly outgoing $\boldsymbol{q}'$, as well as the surface normal.

$$\tilde{\boldsymbol{E}}_r(\boldsymbol{r}', t) = \mathrm{Re}\{(\mathbf{i}'\tilde{r}_s\tilde{E}_{rs} + \boldsymbol{j}'\tilde{r}_p\tilde{E}_{rp}) \times \exp[\mathrm{i}(\tilde{\boldsymbol{q}}' \cdot \boldsymbol{r}' - \omega t)]\} \quad [35]$$

An analogous expression holds for the transmitted electric field. As in the previous case, by imposing the surface boundary conditions for the electric and magnetic fields at the interface, the complex Fresnel coefficients in reflection ($\tilde{r}_p$ and $\tilde{r}_s$) and transmission ($\tilde{t}_p$ and $\tilde{t}_s$) are obtained:

$$\tilde{r}_p = \frac{\tilde{E}_{rp}}{\tilde{E}_{0p}} = \frac{\tilde{n}_0 \cos\tilde{\phi} - \tilde{n}\cos\tilde{\phi}_0}{\tilde{n}_0 \cos\tilde{\phi} + \tilde{n}\cos\tilde{\phi}_0}$$
$$\tilde{t}_p = \frac{\tilde{E}_{tp}}{\tilde{E}_{0p}} = 1 + \tilde{r}_p$$
$$\tilde{r}_s = \frac{\tilde{E}_{rs}}{\tilde{E}_{0s}} = \frac{\tilde{n}_0 \cos\tilde{\phi}_0 - \tilde{n}\cos\tilde{\phi}}{\tilde{n}_0 \cos\tilde{\phi}_0 + \tilde{n}\cos\tilde{\phi}} \quad [36]$$
$$\tilde{t}_s = \frac{\tilde{E}_{ts}}{\tilde{E}_{0s}} = 1 + \tilde{r}_s$$

where $\tilde{\phi}_0$ and $\tilde{\phi}$ are the incident and refraction complex angles which are no more purely geometrical quantities. The refracted em wave is generally inhomogeneous, that is, it is characterized by different directions for planes of constant phase (parallel to the interface and perpendicular to $\boldsymbol{q}_2$) and for planes of constant amplitude (perpendicular to $\boldsymbol{q}_1$). In particular, the generalized Snell law at the interface between two absorbing media is written as

$$\tilde{n}_0 \sin\tilde{\phi}_0 = \tilde{n}\sin\tilde{\phi} \quad [37]$$

If the ambient medium is transparent, such as vacuum or air, the quantity $n_0 \sin\phi_0 = \tilde{n}\sin\tilde{\phi}$ is real, and then $\tilde{\phi}$ is determined starting from $\tilde{n}$ and

ϕ_0. At normal incidence the distinction between parallel and perpendicular components disappears and $\tilde{r}_s = \tilde{r}_p = \tilde{r}$ and $\tilde{t}_s = \tilde{t}_p = \tilde{t}$, that is, eqns [36] reduce to eqns [28].

From the Fresnel coefficients in eqn [36] and the energy conservation at the interface, reflectivity R_j and transmittivity $T_j (j = s, p)$ are derived as

$$R_j = |\tilde{r}_j|^2, \qquad T_j = \mathrm{Re}\left(\frac{\tilde{n}\cos\tilde{\phi}}{\tilde{n}_0\cos\tilde{\phi}_0}\right)|\tilde{t}_j|^2 \quad [38]$$

Again it is verified that $R_s + T_s = 1$ and $R_p + T_p = 1$. For normal incidence, eqns [38] reduce to eqns [29]. In case of unpolarized waves, it holds

$$R = \tfrac{1}{2}(R_p + R_s), \qquad T = \tfrac{1}{2}(T_p + T_s) \quad [39]$$

At any angle of incidence ϕ_0 (except for $\phi_0 = 0$ or $\phi_0 = \pi/2$), it is $R_s > R_p$. Moreover, when both materials are transparent ($k_0 = k = 0$), and when $\phi_0 + \phi = \pi/2$ (i.e., the reflected and refracted waves are perpendicular), it follows that $R_p = 0$. The p component is refracted without losses, while the s component is only partially refracted. This particular value of ϕ_0 is called the Brewster angle ϕ_B, and from the Snell law it follows that $\tan\phi_B = n/n_0$. In **Figure 3**, R_p and R_s components versus the angle of incidence for bulk silicon immersed in air and in the transparent wavelength region ($n \cong 3.42$) are reported. The Brewster angle value is $\phi_B \cong 73.7^\circ$. If $k \neq 0$, R_p goes through a nonzero minimum in correspondence of a different value of $\phi > \phi_B$ which is called the pseudo-Brewster angle.

Ellipsometric Functions

An alternative way to describe the optical response of a material is based on the variation of the state of polarization of an em wave produced by reflection on the surface. The polarization state of an em wave is expressed by $\tilde{\chi}$, defined as

$$\tilde{\chi} = \frac{\tilde{E}_p}{\tilde{E}_s} \quad [40]$$

Since $\tilde{E}_p$ and $\tilde{E}_s$ are complex, $\tilde{\chi}$ is, in general, a complex quantity: if $\tilde{E}_p$ and $\tilde{E}_s$ are in-phase, $\tilde{\chi}$ is real and the em wave is linearly polarized. In the case of $\tilde{E}_p$ and $\tilde{E}_s$ $\pi/2$ out-of-phase, $\tilde{\chi}$ is purely imaginary; if, in addition, $|\tilde{E}_p| = |\tilde{E}_s|$ then $\tilde{\chi} = \pm i$ and the wave is circularly polarized. Elliptical polarization is obtained when $\tilde{E}_p$ and $\tilde{E}_s$ are out-of-phase by $\pi/2$ and have different amplitudes, or in the more general case when the phase angle has any value different from 0 and $\pi/2$.

The incident and reflected waves are characterized by the polarization states $\tilde{\chi}_0$ and $\tilde{\chi}_r$, respectively, that are connected by the following relationship:

$$\tilde{\chi}_r = \frac{\tilde{E}_{rp}}{\tilde{E}_{rs}} = \frac{\tilde{r}_p\tilde{E}_{0p}}{\tilde{r}_s\tilde{E}_{0s}} = \frac{\tilde{r}_p}{\tilde{r}_s}\tilde{\chi}_0 \quad [41]$$

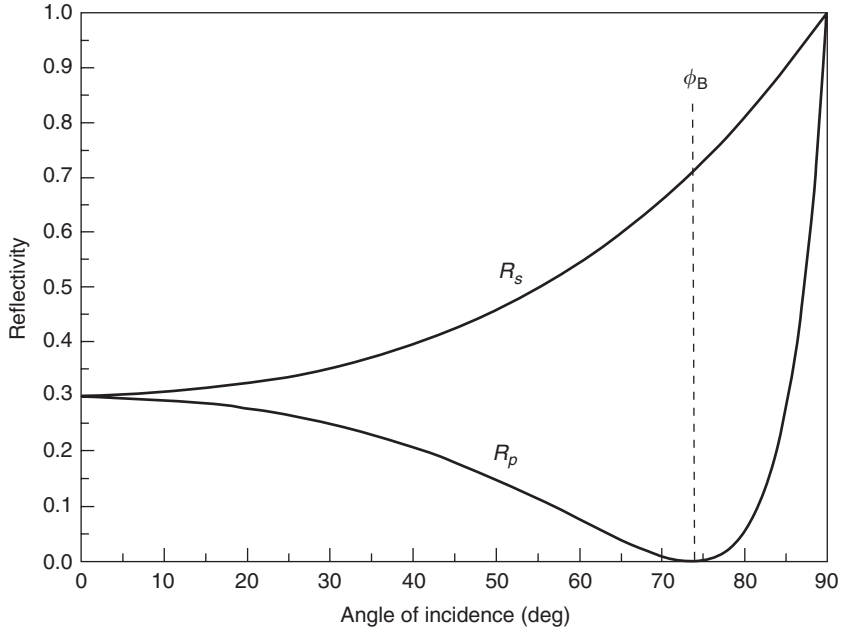


Figure 3 Polarized reflectivities R_s and R_p as a function of the angle of incidence for silicon in the subgap transparency spectral region ($n = 3.42$). R_p reaches zero value at the Brewster angle $\phi_B = 73.7^\circ$.

where $\tilde{r}_p$ and $\tilde{r}_s$ are defined in eqns [36]. Since $\tilde{r}_p$ and $\tilde{r}_s$ are complex quantities (i.e., $\tilde{r} = |\tilde{r}|e^{i\theta}$), it is evident that the reflected wave has a different polarization state with respect to the incident one.

When linearly polarized light impinges on a material surface, the reflected light is generally elliptically polarized. This ellipse of polarization is completely characterized by the ratio of the two axes and by the orientation of the major axis. From this fact, the name ellipsometry is given to the experimental technique which allows one to determine the variation in the polarization state by reflection on a sample surface.

In ellipsometry the crucial quantity is the ratio $\tilde{\rho}$, defined as the ratio between $\tilde{\chi}_r$ and $\tilde{\chi}_0$:

$$\tilde{\rho} = \frac{\tilde{\chi}_r}{\tilde{\chi}_0} = \frac{\tilde{r}_p}{\tilde{r}_s} = \left(\frac{E_{rp}}{E_{rs}}\right)\left(\frac{E_{0s}}{E_{0p}}\right)\exp\{i[(\beta_p - \beta_s)_r - (\beta_p - \beta_s)_0]\} \quad [42]$$

Here $(\beta_p - \beta_s)$ is the phase difference between p and s field components. The absolute value of $\tilde{\rho}$ gives information on the amplitude ratio of the reflected to the incident E_p and E_s, while the phase of $\tilde{\rho}$ is given by the variation of the phase difference caused by reflection. The quantity $\tilde{\rho}$ is normally written in the compact form

$$\tilde{\rho} = \tan\psi e^{i\Delta} \quad [43]$$

where

$$\tan\psi = \rho = \frac{|\tilde{r}_p|}{|\tilde{r}_s|} = \frac{r_p}{r_s} \qquad \Delta = (\beta_p - \beta_s)_r - (\beta_p - \beta_s)_0 \quad [44]$$

A single ellipsometric measurement gives $\tilde{\rho}$, that is, the ellipsometric functions $\tan\psi$ and $\cos\Delta$, which allow one to determine the optical properties of a material. As an example, in the case of a semi-infinite material with complex dielectric function $\tilde{\varepsilon}$, assuming that the ambient medium is void ($\tilde{\varepsilon}_o = 1$) and ϕ_0 is the angle of incidence, it is possible to obtain a simple relation which binds $\tilde{\rho}$ and $\tilde{\varepsilon}$:

$$\tilde{\varepsilon} = \sin^2\phi_0 + \sin^2\phi_0 \tan^2\phi_0 \left|\frac{1-\tilde{\rho}}{1+\tilde{\rho}}\right|^2 \quad [45]$$

It is then immediate to derive the real and imaginary part of $\tilde{\varepsilon} = \varepsilon_1 + i\varepsilon_2$. In addition, in transparent media, ellipsometry is widely used to determine the refractive index and the thickness of a thin film layer deposited on a substrate with known optical properties.

In conclusion, it is noted that the reflection/transmission properties of a medium are of major interest from the application point of view. Moreover, they allow a complete optical characterization of a homogeneous, isotropic medium, that is, the determination of ε_1 and ε_2 or n and k. Since there are two unknown parameters, at least two independent measurements are needed. In transparent materials, R and T should be used, or R measurements at different angles of incidence. In semi-infinite media (or with finite thickness but opaque), only reflection and ellipsometric measurements should be performed.

Moreover, the Fresnel coefficient $\tilde{r} = |\tilde{r}|e^{i\theta}$ expresses the linear relation between incident and reflected electric field amplitudes which obey causality. As stated above, it follows that a dispersion relation exists which connects its real and imaginary parts. Using the normal incidence reflectivity $R = |\tilde{r}|^2$, the dispersion relation between the absolute value R and the phase angle θ should be written as

$$\theta(\omega) = -\frac{\omega}{\pi} P \int_0^\infty \frac{\ln R(\omega')}{\omega'^2 - \omega^2}\, d\omega' \quad [46]$$

Then θ should be determined, in principle, at whatever frequency if $R(\omega)$ is known over the whole spectral interval. In practice, $R(\omega)$ is measured over a finite spectral range and some means of extrapolating experimental results over inaccessible spectral ranges are necessary. Finally, n and k are calculated on the basis of eqns [28] and [29].

See also: Optical Absorption and Reflectance; Optical Sum Rules and Kramers–Kronig Relations; Sum Rules and Kramers–Kronig Relations in Nonlinear Optics.

PACS: 42.25. – p; 42.25.Bs; 42.25.Gy; 42.25.Ja; 78.20.Ci

Further Reading

Azzam RMA and Bashara NM (1977) *Ellipsometry and Polarized Light*. Amsterdam: North-Holland.

Bassani F and Pastori Parravicini G (1974) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Born M and Wolf E (1980) *Principles of Optics*. Cambridge: Cambridge University Press.

Grosso G and Pastori Parravicini G (2000) *Solid State Physics*. San Diego: Academic Press.

Heavens OS (1962) *Optical Properties of Thin Solid Films*. London: Butterworths Sci. Publ.

Hecht E (1990) *Optics*. Boston: Addison-Wesley.

Palik ED (1985) *Handbook of Optical Constants of Solids*. Orlando: Academic Press.

Stern F (1963) Elementary theory of the optical properties of solids. *Solid State Physics* 15: 299–408.

Wooten F (1972) *Optical Properties of Solids*. New York: Academic Press.

Nomenclature

$\mathscr{A}$	absorbance (dimensionless)
$\boldsymbol{B}$	magnetic induction (T)
$\boldsymbol{D}$	electric displacement ($C\,m^{-2}$)
$\boldsymbol{E}$	electric field ($V\,m^{-1}$)
$\boldsymbol{H}$	magnetic field ($A\,m^{-1}$)
$\boldsymbol{j}$	current density ($A\,m^{-2}$)
$\boldsymbol{j}_{ext}$	external current density ($A\,m^{-2}$)
$\boldsymbol{j}_{cond}$	internal current density ($A\,m^{-2}$)
k	extinction coefficient (dimensionless)
$\boldsymbol{M}$	magnetization (magnetic moment per unit volume) ($A\,m^{-1}$)
n	refractive index (dimensionless)
$\tilde{n} = n + ik$	complex refractive index (dimensionless)
n_e	extraordinary refractive index (dimensionless)
n_o	ordinary refractive index (dimensionless)
$\boldsymbol{P}$	polarization (electric dipole moment per unit volume) ($C\,m^{-2}$)
$\boldsymbol{q}$	wave vector (m^{-1})
$\tilde{\boldsymbol{q}} = q_1 + iq_2$	complex wave vector
$\tilde{r}$	reflection Fresnel coefficient (dimensionless)
R	reflectivity (dimensionless)
$\mathscr{R}$	reflectance (dimensionless)
$\tilde{t}$	transmission Fresnel coefficient (dimensionless)
T	transmittivity (dimensionless)
$\mathscr{T}$	transmittance (dimensionless)
v	phase velocity ($m\,s^{-1}$)
ν	frequency (s^{-1})
α	absorption coefficient (cm^{-1})
αd	optical density (dimensionless)
δ	classical skin depth (nm)
ε	dielectric function ($C^2\,N^{-1}\,m^{-2}$)
$\tilde{\varepsilon} = \varepsilon_1 + \varepsilon_2$	complex dielectric function
$\varepsilon_0 = 8.859 \times 10^{-12}$	vacuum permittivity ($C^2\,N^{-1}\,m^{-2}$)
ε_r	relative dielectric function (dimensionless)
λ	wavelength (nm)
$\mu_0 = 4\pi \times 10^{-7}$	vacuum magnetic permeability ($T\,m\,A^{-1}$)
ρ	external introduced charge density ($C\,m^{-3}$)
σ	optical conductivity ($\Omega^{-1}\,cm^{-1}$)
$\tilde{\sigma} = \sigma_1 + \sigma_2$	complex conductivity
ϕ_B	Brewster angle (°)
χ	electrical susceptibility (dimensionless)
$\tilde{\chi}$	polarization state (dimensionless)
ω	angular frequency (s^{-1})
$\tan\psi$, $\cos\Delta$	ellipsometric functions (dimensionless)

Optical Properties of Surface Layers Enhanced Raman Scattering

G Mattei, CNR, Istituto dei Sistemi Complessi, Rome, Italy

Introduction

The subject of this article is the description of the properties of metal surfaces and metal particles that are sources of strong enhancement of the intensity of the scattered Raman light from molecules and thin films, which are in contact with them or in their close vicinity in the nanoscale range. First, the origin of the enhancement of the Raman intensity is discussed showing that the main source of the effect is the amplification of the electromagnetic (e.m.) field at the interface of two media (e.g., metal/dielectric), when surface e.m. waves and/or surface or localized plasmons are excited. The cases of flat and corrugated metal surfaces, single metal particles, and metal clusters are discussed. The last section is devoted to the illustration of the surface-enhanced Raman scattering (SERS), commonly characterized by enhancement of the Raman intensity of 4–8 orders of magnitude, and originated from metal (e.g., Au, Ag, Cu) surfaces which are rough on nanoscale (10–100 nm) and from metal nanoparticles and nanoclusters. Not all the subjects are treated in detail since some of them are presented exhaustively in other articles of this encyclopedia.

The Origin of the Enhancement of Raman Intensities

As it is well known, Raman scattering is a nonlinear process (e.g., two photons and one quantum of vibration are involved) which is characterized by a very low cross section σ (in m^2) defined by the relation: $P_{sc} = \sigma I_{in}$, where P_{sc} is the power (in W) of the scattered light and I_{in} is the intensity of the exciting light (in Wm^{-2}). For the normal Raman scattering, the cross section (average value $\approx 10^{-33}$ m^2) is nine orders of magnitude smaller than that of the infrared absorption process and this prevents the detection of a very small amount of material as it is necessary, for example, in surface science. To achieve the surface

sensitivity (detection limit of a fraction of the monolayer), some enhancement in the process is necessary. Two mechanisms are possible sources of this enhancement. The first one is related to the cross section itself. In some systems, molecules as well as crystals and films, at special frequencies σ present a resonant behavior due to the direct absorption of the incident photons by suitable electronic levels coupled with vibrational states. In this condition, that is, resonant Raman scattering, the intensity of the scattered light can be enhanced up to six orders of magnitude. In some cases, chemisorption on metal is accompanied by the formation of new electronic levels in the visible range for the adsorbate/metal system, allowing the resonant Raman scattering to occur. Of course, this resonant process (average enhancement factor 100), often indicated as "chemical effect," is strongly dependent on the frequency of the exciting light and is specific to the excited system. The second source of Raman intensity amplification is related to the enhancement of the local electric field seen by the excited medium (e.g., adsorbed molecules on a metal surface or on a small metal particle). Indeed I_{in} being proportional to the square of the electric field amplitude of the excitation light, an increase of this amplitude by a factor 10^2 implies an enhancement of the Raman intensity by a factor 10^4, for example. This enhancement mechanism of the Raman intensity, called "electromagnetic effect," is somehow independent of the excited system and presents some specificity that allows one to discriminate it from the "chemical effect." Of course, in some cases both the effects can contribute to the total Raman enhancement. The following sections illustrate the e.m. effect through the use of system models, namely various metal surfaces. Undoubtedly, the properties of the surface itself play an important role in the Raman enhancement. Among them, the surface shape and size, and also the frequency-dependent dielectric function of the metal, have to be considered.

Flat Metal Surface: Raman Enhancement and Surface Selection Rules

Consider the case of molecules adsorbed (or in the near proximity) on a flat metal mirror where an e.m. wave, impinging on with an angle of incidence θ, is reflected (**Figure 1**).

The molecules see a local electric field at surface, $\boldsymbol{E}^{su}$, that is due to the coherent superposition of the electric fields $\boldsymbol{E}^{i}$ and $\boldsymbol{E}^{r}$ of the incident and the reflected e.m. waves. Considering the two polarizations, s (TE) and p (TM), one has, for the x, y, and z

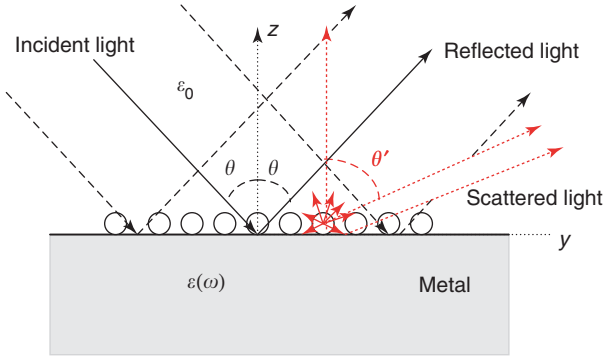


Figure 1 Scheme of the Raman scattering of molecules (circles) adsorbed on the flat interface between a dielectric and a metal.

components of the local field,

$$E_x^{su} = E_{s,0}^{i}(1 + r_s)$$
$$E_y^{su} = E_{p,0}^{i}(1 - r_p)\cos(\theta)$$
$$E_z^{su} = E_{p,0}^{i}(1 + r_p)\sin(\theta)$$

where $E_{s,0}^{i}$ and $E_{p,0}^{i}$ are the components of the incident electric field of s (TE) and p (TM) polarization, and r_s and r_p are the Fresnel coefficients, being:

$$r_s = [\cos(\theta) - (\varepsilon - \sin^2(\theta))^{1/2}]/[\cos(\theta) + (\varepsilon - \sin^2(\theta))^{1/2}]$$

$$r_p = [\varepsilon\cos(\theta) - (\varepsilon - \sin^2(\theta))^{1/2}]/[\varepsilon\cos(\theta) + (\varepsilon - \sin^2(\theta))^{1/2}]$$

It is to be noted that the dielectric function $\varepsilon(\omega)$ is a complex function of the frequency ω (here the dependence on the wave vector of the light is neglected). In the case of metals, the real part $\mathrm{Re}[\varepsilon(\omega)]$ can be strongly negative below the plasma frequency ω_p and almost zero around ω_p. It is easy to demonstrate that when $|\varepsilon|$ is very large, $r_s \to -1$ and $r_p \to +1$; it is just the opposite in the limit of very small $|\varepsilon|$, $r_s \to +1$ and $r_p \to -1$.

In the light scattering process, the local electric field $\boldsymbol{E}^{su}(\omega_{ex})$ induces on the molecules at the surface a dipole moment $\boldsymbol{\mu}$, given by: $\boldsymbol{\mu}(\omega_{sc}) = \alpha\boldsymbol{E}^{su}(\omega_{ex})$, $\hat{\alpha}$, the polarizability of the molecules being a tensor of the second rank. The frequencies ω_{ex} and ω_{sc} are those of the incident and scattered light, respectively. $|\omega_{ex} - \omega_{sc}| = \omega_{vibr}$ is the frequency of the normal vibrational mode of the molecule involved in the scattering process. The scattered light, collected along a direction θ' far away from the surface, is then the superimposition of the light emitted directly in this direction together with the scattered light directed toward the metal surface and then reflected in the same direction (**Figure 1**). Hence, the intensity of

the scattered light, that is proportional to the square of the amplitude of the scattered electric field, will be affected by two surface contributions due to the reflection of the incident and of the scattered light as well. From the above expressions, the Raman intensities for the possible polarization (s and/or p) of the incident and scattered light can be obtained:

$$I_{ss} \propto \omega_{sc}^4 |\alpha_{xx}(1+r_s)(1+r'_s)|^2$$

$$I_{ps} \propto \omega_{sc}^4 |\alpha_{xy}(1-r_p)(1+r'_s)\cos(\theta) + \alpha_{zy}(1+r_p)(1+r'_s)\sin(\theta')|^2$$

$$I_{sp} \propto \omega_{sc}^4 |\alpha_{yx}(1+r_s)(1-r'_p)\cos(\theta') + \alpha_{zx}(1+r_s)(1+r'_p)\sin(\theta')|^2$$

$$I_{pp} \propto \omega_{sc}^4 |[\alpha_{yy}(1-r_p)\cos(\theta) + \alpha_{yz}(1+r_p)\sin(\theta)](1-r'_p)\cos(\theta') + [\alpha_{zy}(1-r_p)\cos(\theta) + \alpha_{zz}(1+r_p)\sin(\theta)](1+r'_p)\sin(\theta')|^2$$

In these equations, all the quantities including the Raman polarizability components are referred to the frame of axis indicated in the figure with z along the metal surface normal. The primes indicate those quantities that are to be evaluated at the frequency and angle of the scattered light. It is clear that the observed scattered intensity depends in a quite complex way on the chosen scattering geometry (through the angles, the polarizations, and the Fresnel coefficients), the optical properties of the metal (through its frequency-dependent dielectric function), the symmetry of the molecular vibration, and the orientation of the molecule with respect to the surface (through the α_{ij} components of the polarizability). These relations determine the surface selection rules that control the Raman scattering of molecules adsorbed or posed on a flat metal surface. According to these selection rules, the Raman spectra can be very different from those of free molecules: the relative band intensities can be altered, some band can be strongly attenuated, others can be enhanced with respect to the free molecules. Even the dependence on the scattering frequency of the Raman intensity for single modes can deviate significantly from the ω^4 rule typical of the Raman scattering of free molecules. Finally, it is easy to verify that the maximum expected enhancement of the Raman intensity for any mode is 16 but for real systems, this value is seldom obtained, while enhancement values around 6–10 can be usually measured on flat metal mirrors.

Electromagnetic Enhancement due to Surface Plasmons and Surface Electromagnetic Waves: Flat and Roughed Surfaces

As it is well known, at the vacuum/metal interface, surface collective charge modes, named surface plasmons (sp), polaritons do exist. They are evanescent waves produced by the oscillations of the surface charges, and propagate along the interface with a real wave vector K, given by

$$K = (\omega/c)\sqrt{\varepsilon_0\varepsilon/(\varepsilon_0+\varepsilon)}$$

In this equation, which represents the sp dispersion curve (ω vs. K), ε_0 and ε are the dielectric function of the dielectric (e.g., the air) and of the metal, respectively. No radiation is emitted from the surface in the two semispaces. The electric field that can be very intense at the surface, decays exponentially with the distance from the interface. Indeed, the components of the wave vector perpendicular to the surface are basically purely imaginary being: $K_\perp = \mathrm{i}(\omega/c)|\varepsilon_i^2/(\varepsilon_0+\varepsilon)|^{1/2}$, where ε_i is equal to ε_0 in the dielectric medium and to ε in the metal, respectively.

The sp condition of existence and their dispersion curve are illustrated in **Figure 2**, where for simplicity the metal dielectric function, ε, has been approximated with that of a free electron gas, namely: $\varepsilon = 1-(\omega_p/\omega)^2$. In the figure, "reduced" adimensional quantities have been used for frequency, ω/ω_p, and surface wave vector, cK/ω_p.

It is clear from the comparison of the frequency dependences of the dielectric function and of the sp dispersion that sp do exist only for $\omega<\omega_p$, where the bulk modes cannot propagate in the metal, being $\varepsilon<0$. For large K values, the sp polaritons dispersion curve tends to frequency $\omega_p/(1-\varepsilon_0)^{1/2}$ for which it is $\varepsilon = -\varepsilon_0$. Since the sp are not radiative modes, they cannot be directly excited by light propagating in one of the two media. Indeed, for a given frequency ω, the component parallel to the surface, $k_{||}$ of the light wave vector will be always smaller than the corresponding component, K, of the sp wave vector. As a matter of fact, it is always $k_{||} = (\omega/c)\varepsilon_0^{1/2}\sin\theta_i < (\omega/c)\varepsilon_0^{1/2} < K$, where θ_i is the angle of incidence of the light on the metal surface.

Hence, the conditions for the energy and wave vector conservation cannot be simultaneously satisfied. Nevertheless, sp excitation by light can be obtained in special conditions, namely: by grating coupling and in attenuated total reflection (ATR) geometry.

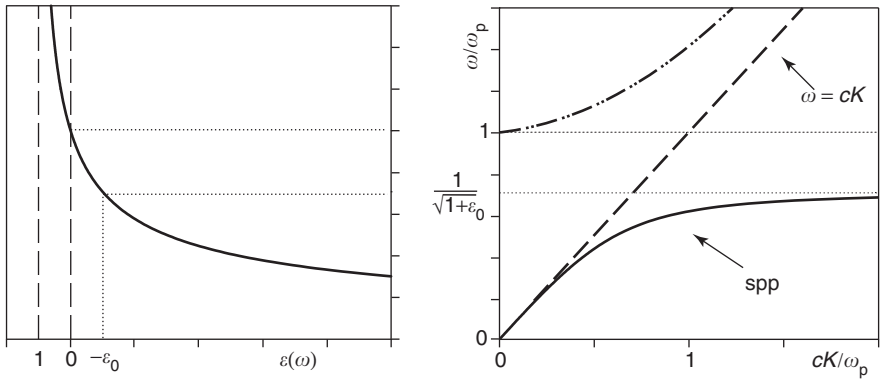


Figure 2 On the left, dielectric function vs. reduced frequency, ω/ω_p, (ω_p is the plasma frequency) for a metal in the free electron gas approximation; on the right, surface plasmon polaritons (spp) dispersion curve for the metal/dielectric interface. K is the real wave vector parallel with the surface. The dashed straight line represents the photon dispersion in vacuum.

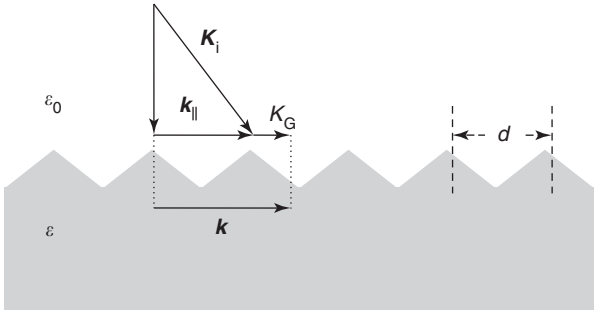


Figure 3 Scheme of the spp excitation of a metal grating of period d by incident light of wave vector K_i. $K_G = (2\pi n/d)$ is the surface wave vector due to the grating. $k_{||}$ and K are the surface wave vectors parallel to the surface of the incident light and of the spp excitations.

When a metal surface is sinusoidally modulated with a period d to form a grating, if the amplitude of the modulation is small compared with the wavelength of light, there is a modification of the light wave vector parallel to the surface, $k_{||}$, into $K_n = (k_{||} + 2\pi n/d)$ with $n = 0, \pm 1, \pm 2$, etc. It is then possible that, for suitable value of θ and n, the coupling condition, $K_n = K$, be satisfied, allowing the sp excitation by light radiation (**Figure 3**).

Wood anomalies in gratings, experimentally well known since the beginning of the twentieth century, can be attributed to this effect. Besides, sp can have their wave vector modified by a surface modulation and become radiative waves. On this basis, it is clear that a rough metal surface can allow to some extent the excitation of sp waves by light and their radiative emission.

Another way to excite sp is the use of a prism in total reflection condition to generate surface e.m. evanescent waves that can couple with the spp modes. The ATR in the so-called "Otto configuration" is illustrated in **Figure 4a**, where a prism

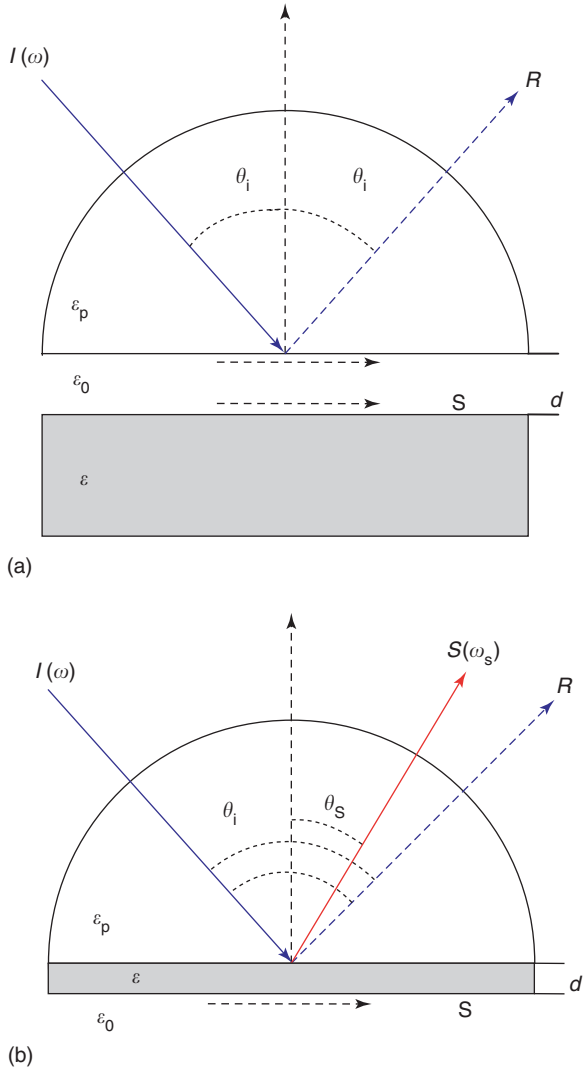


Figure 4 ATR geometries: (a) Otto configuration, (b) Kretschmann configuration: I, R and S represent the incident, the reflected, and the scattered beams, respectively; d is the thickness of the interlayer between prism and metal (a) or of the metal film (b).

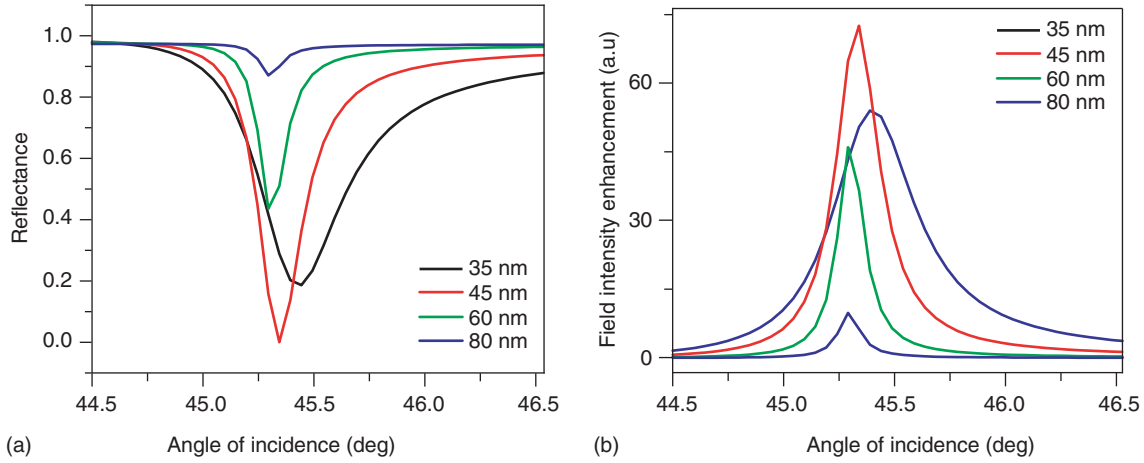


Figure 5 ATR in the Kretschmann configuration, wavelength 530 nm. (a) Angular dependence of the reflectivity in the prism for various Ag film thicknesses. (b) Angular dependence of the electric field intensity at the air/metal interface for various Ag film thicknesses.

(dielectric constant ε_p) is separated by a thin layer of thickness d and dielectric constant $\varepsilon_0 < \varepsilon_p$ from a metal surface S. If a light beam strikes the interface prism/interlayer at an angle of incidence θ_i greater than the critical angle $\theta_c = \arcsin(\varepsilon_0/\varepsilon_p)^{1/2}$, total reflection occurs: no light is transmitted in the ε_0 medium. At the interface, an evanescent e.m. wave is produced with real wave vector parallel to the surface $k_{||} = (\omega/c)\varepsilon_p^{1/2} \sin\theta_i > (\omega/c)\varepsilon_0^{1/2}$ and perpendicular wave vector component purely imaginary in the ε_0 media. The amplitude of the electric field decays rapidly in the interlayer, but, if the thickness d is of the order of the wavelength, the surface e.m. waves can interact with the sp evanescent waves generated at the interface metal/interlayer when the wave vector ($k_{||} = k$) and the frequency ($\omega_{light} = \omega_{spp}$) conservation laws are satisfied. When this resonant coupling occurs, the intensity of the reflected beam, R, is attenuated or "frustrated" allowing the detection of sp modes.

Since both the Snell's law (i.e., the conservation of the parallel component of the wave vector at the interface) and the dispersion relation of sp are symmetrical with respect to ε_0 and ε, it is possible to interchange the role of these two media with respect to the prism. This is shown in **Figure 4b** where the ATR in the so-called "Kretschmann configuration" is illustrated. The thickness d of the thin metal film must be sufficiently small to allow enough energy to reach the metal/dielectric interface. Again, attenuation of the intensity of the reflected light as a function of the frequency and of the angle of incidence allows the detection of sp waves.

Calculated angular dependence of the reflectivity at 530 nm as a function of the angle of incidence in the Kretschmann configuration (prism ($\varepsilon_p = 2.1$), metal film ($\varepsilon = \varepsilon' + i\varepsilon'' = -17.35 + i0.56$), air ($\varepsilon_0 = 1$)) are presented in **Figure 5a** for different values of d. The minimum of the reflectivity, corresponding to the excitation of the sp at the interface metal/air, is deepest for an optimal film thickness. Note that the angle and the angular width of this minimum depend on the thickness, because in the case of thin film the sp modes of the two interfaces fill each other, and influence and modify the corresponding dispersion curves depending on the distance d between the interfaces.

An important consequence of the excitation of the sp waves by ATR is that the intensity of the electric field at the metal/dielectric (metal/air, in this case) interface can be significantly larger than that of the incident e.m. wave. As an example, in **Figure 5b**, the field intensity enhancement, namely the ratio between the intensity of the electric field at the metal/air interface and the one of the incident wave at the prism/metal interface, is reported as a function of the angle of incidence, for the case of **Figure 5a**. As it is clear, in the optimal condition, an enhancement of two orders of magnitude can be achieved. As a consequence, enhanced Raman scattering from molecules adsorbed on the metal/air interface can be excited in this configuration. Considering that the scattered light can be also converted in sp excitations and, through the ATR system, into light emitted in the prism at a suitable angle, a final enhancement of about four orders of magnitude can be obtained. Indeed, in the literature several experimental evidences of enhanced Raman scattering using ATR configuration, with the enhancement factor ranging from a few tens to ten thousand, have been reported. For example, studies by enhanced Raman spectroscopy of metal/liquid interfaces and Langmuir–Blodget

layers on smooth Ag surfaces have been reported. Besides, a combination of ATR spp excitation and near-field Raman spectroscopy has been used to scan, with high sensitivity, the species present on the electrode surfaces.

Finally, for completeness, it is important to note that the conversion of traveling e.m. waves into surface evanescent e.m. waves can be used to enhance Raman scattering intensity at the interface between two dielectric media, when the exciting light strikes the interface from the denser medium in total reflection condition. As a matter of fact, the Raman scattering excited in a total reflection condition has been used for studying adsorbates, films, and even the surface phonon-polaritons dispersion at the interface between a prism and a transparent crystal.

Electromagnetic Enhancement due to Metal Particles

The following case is considered first, that is, the case of single isolated metal sphere of radius R, small with respect to the light wavelength λ: that is, $R \ll \lambda$ ("Rayleigh limit" or "electrostatic approximation"). If such a sphere, of dielectric function $\varepsilon(\omega)$ and embedded in a medium of dielectric function ε_0, is irradiated by an e.m. wave with electric field amplitude E_i polarized in the z direction (see scheme in **Figure 6**), a dipole moment $p = gR^3E_i$ is generated inside the sphere. The potential, V, outside the sphere, due to the incident electric field and the electric dipole p, expressed in spherical coordinates (r, θ, φ), is

$$V(r,\theta) = E_i(r - gR^3/r^2)\cos(\theta)$$

where $g = (\varepsilon - \varepsilon_0)/(\varepsilon + 2\varepsilon_0)$. The radial, E_n, and tangential, E_t, components of the electric field at the surface of the sphere can be obtained in a standard way from the potential. The intensity of the two electric components, averaged over all the solid angle, are proportional to

$$\bar{E}_n^2 \propto |1 + 2g|^2$$
$$\bar{E}_t^2 \propto 2|1 - g|^2$$

Because of the resonant denominator in g, when it is $\mathrm{Re}[\varepsilon(\omega)] = -2\varepsilon_0$ (condition of excitation of localized sp of the metal sphere) and the $\mathrm{Im}[\varepsilon(\omega)]$ is small (as in the case of coinage metals), the electric field intensity at the surface of the metal particle can be much larger that that of the incoming radiation.

Let the light scattering process be considered now, assuming that the molecules are adsorbed on the spherical metal particle with the z-axis of the molecule-fixed coordinate frame along the normal to the surface. The average Raman intensity can be simply considered as dependent on the average intensity of the incident and scattered fields. The metal particle enhances not only the incident electric field on the molecule, but also the Raman scattered field. It acts as an antenna, through the dipole induced in it by the molecular dipole, which amplifies the scattered light intensity. Considering the polarizability tensor of the molecule, three kinds of vibrations can be distinguished. The α_{zz}-type modes require a radial field, E_n, to be excited and produce an induced dipole with only a radial component. Hence, the average intensity due to only the radial electric field will contribute twice to the detected Raman intensity: once in excitation and once again in emission. Similar considerations can be made for the other two kinds of modes, $\alpha_{xx}, \alpha_{yy}, \alpha_{xy}$ (only tangential components), and α_{xz}, α_{yz}, (mixture of tangential and radial components).

The intensity of enhanced Raman spectra of these kinds of modes will therefore depend on the frequency through the following functions:

α_{zz}-type

$$\bar{E}_n^2 \bar{E}_n'^2 \propto |1 + 2g|^2 |1 + 2g'|^2$$

$\alpha_{xx}, \alpha_{yy}, \alpha_{xy}$, -type

$$\bar{E}_n^2 \bar{E}_n'^2 \propto 4|1 - g|^2 |1 - g'|^2$$

α_{xz}, α_{yz}-type

$$\tfrac{1}{2}(\bar{E}_n^2 \bar{E}_t'^2 + \bar{E}_t^2 \bar{E}_n'^2) \propto [|1 + 2g|^2 |1 - g'|^2 + |1 - g|^2 |1 + 2g'|^2]$$

where the prime indicates quantities to be evaluated at the frequency of the scattered light. In **Figure 7** the dependence on the frequency of these functions is

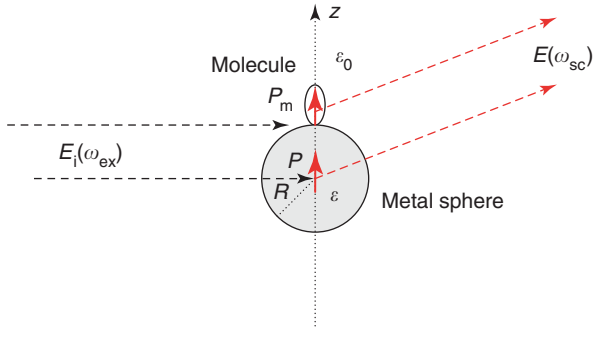


Figure 6 Scheme of the Raman scattering from a molecule adsorbed on a metal spherical particle of radius R. E_i and E indicate the electric field of the incident and of the scattered light, respectively; p_m and p are the electric dipole induced in the molecule and in the metal particle, respectively.

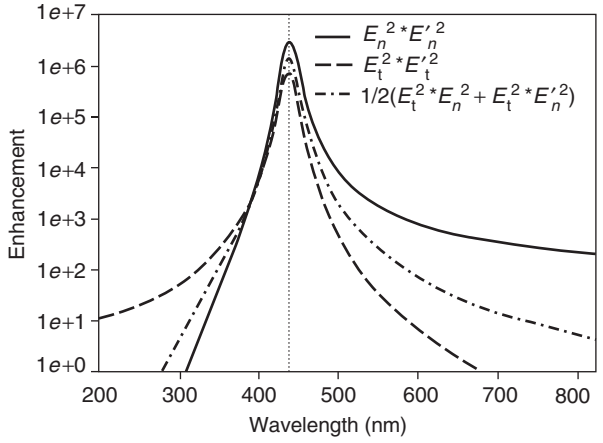


Figure 7 Wavelength dependence of the enhancement factors for the three kinds of molecular modes for a molecule adsorbed on a metal particle of radius R in the Rayleigh limit ($R \ll$ wavelengthof the exciting light).

presented for a fictitious free-electron metal particle. For simplicity, it has been considered that $\varepsilon_0 = 1$ and $g = g'$.

As expected, all three functions have a sharp maximum (corresponding enhancement factor 10^6) at a wavelength for which $\mathrm{Re}[\varepsilon(\omega)] = -2$. Besides, at longer wavelengths the α_{zz}-type modes dominate in the Raman spectrum, whereas at lower wavelengths all the modes are almost equally enhanced. It is clear that the number and the relative intensity of the bands present in the enhanced Raman spectrum can vary significantly with the frequency of the excitation light. Indeed, often very rich spectra can be measured with the distribution of the intensity that apparently violates the surface selection rules.

The model discussed above, although quite simple, provides the main features associated with the enhanced Raman scattering of molecule adsorbates on small metal particles. A more realistic model must include the effects of the particle shape and size distribution.

As the particle size increases beyond the Rayleigh regime, the used electrostatic approximation is not longer valid. The computer-intensive Lorenz–Mie or electrodynamics formalism, which accurately accounts for the effects due to the phase retardation and the quadrupole and higher order poles, must be used to obtain an exact solution. Also the conversion of the near-field electric field due to the molecule through the metal particle acting as an antenna into far-field radiation must be then taken into account. Generally, as a consequence, the intensity enhancement of the Raman scattered radiation can be significantly reduced.

Considering the shape effect, in the case of an isolated spheroid of eccentricity b/a (a and b being the length of the short and the long principal axis, respectively), in presence of an external electric field (due to an e.m. wave) oriented along the b-axis, it can be demonstrated that the Raman enhancement is always larger for the more aspherical particles. In the limiting case of $(b/a) \to \infty$, the calculated enhancement at the tip positions can be of some order of magnitude larger than for the case of spherical particles ("lightning rod effect").

Finally, to treat the case of interacting metal particles and of metal clusters that are also of great practical relevance, various model systems have been considered, for example, two interacting metal spherical particles ("bispheres"), aggregates of metal spheres, set of parallel metal cylinders of variable diameter and interdistance, etc. The calculation of the optical properties of these systems and of the effects on the Raman scattering is generally quite complex and can include retardation effects and high-order multipoles. The general result is that the interaction between two plasmon-resonant nanoparticles leads to additional resonances (at longer wavelengths) for the coupled system. Besides, the electric field in the gap between the particles or around the contact region (for two contacting particles) can be much strong in other places near the particles and stronger than in any place around an isolated particle. Hence, the Raman enhancement can be much larger than for isolated metal nanoparticles. Indeed, enhancement value as large as 10^8–10^9 has been predicted for such metal systems.

It has been recognized early that small-particle composites, or disordered clusters present in metal colloidal solution, as well as some rough metal surface have a fractal character. Studies have been developed to relate the enhancement of the local electric fields in these systems with their fractal dimension. An interesting example is illustrated in **Figure 8**, where the calculated dependence on the wavelength, λ, of the local electric field enhancement factor for three systems made by 500 metal spherical nanoparticles ($d \ll \lambda$) is presented.

In the case of a gas of spherical particles (RGP curve), the characteristic sharp feature due to the excitation of the sp in isolated spheres is clearly observed below 400 nm. The absence of other features above that wavelength indicates a very lithe interaction among the particles. For the ordered particle aggregate (CPSP curve), the single-particle excitation is attenuated and the enhancement increases in the long-wavelength region of the spectrum as a result of the mutual particle interactions. Clearly, the largest enhancement in the local field, up to 10^3 times the one of isolated particle, is predicted for the fractal system (CCA curve). On this basis, huge

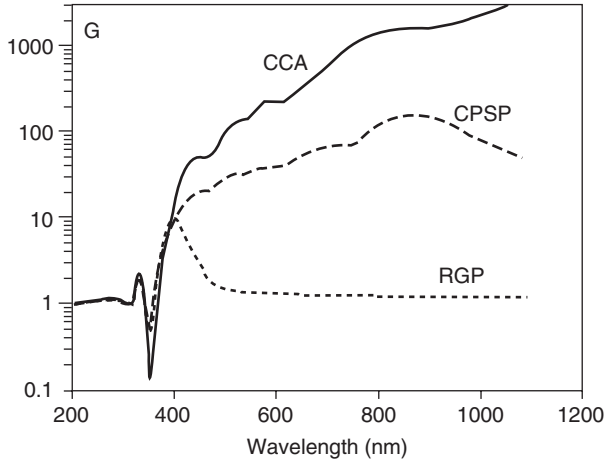


Figure 8 Wavelength dependence of the averaged local field intensity for a gas of spherical particles (RGP), a close-packed crystal of spherical particles (CPSP), and a fractal aggregate (CCA). (Reproduced from Shalaev VM (1996) *Physics Reports* 272: 61–137, with permission from Elsevier.)

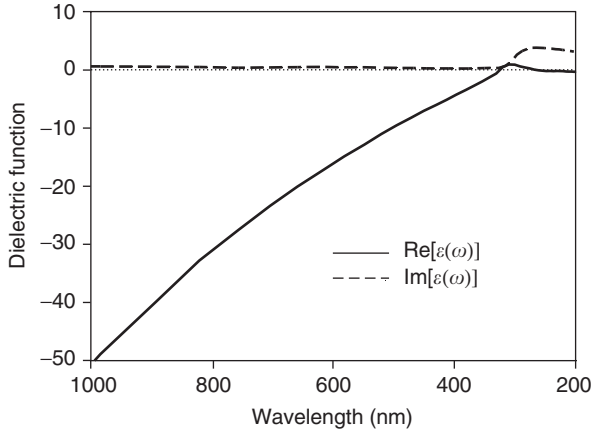


Figure 9 Experimental dielectric function, $\varepsilon(\omega)$, of a silver film. The combined contributions of the free electron and interband transitions cause $\mathrm{Re}[\varepsilon(\omega)]$ of Ag to approach 0 in the visible range even though the free electron plasma frequency is in the ultraviolet. In the same range $\mathrm{Im}[\varepsilon(\omega)]$ is quite small.

enhancement factors for the intensity of Raman scattering from molecules on fractal systems can be expected.

Surface-Enhanced Raman Scattering

SERS was discovered, although not recognized as such, by M Fleishman and co-workers in 1974. They observed an intense Raman signal from pyridine adsorbed onto a roughened silver electrode surface from an aqueous solution. They attributed, erroneously, the observed enhancement of the Raman signal to an increase of the surface area due to the roughening process. In 1977, the nature of this enhancement was attributed by D L Jeanmair and R V Van Duyne, and independently by M G Albert and J A Creighton to a new phenomenon related with the increase of the effective Raman cross section of adsorbates on some rough metal surfaces. Since then, many studies and works have being carried out to understand the fundamental aspects as well as to exploit all the applicative potentialities related with SERS. A wide variety of metal substrates have been found to exhibit SERS: electrochemically roughened electrodes, chemically etched surfaces, colloids (especially aggregated colloids), island films, particles grafted on silanized glasses, regular particle arrays, etc. The dominant metals in SERS are the coinage (Cu, Au, Ag) and alkali (Li, Na, K) metals. Moreover, although with lower enhancement, transition metals (e.g., Pt, Ru, Rh, Pd, Fe, Co, Ni and their alloys) have also been used. Generally, the observed enhancement ranges for 10^4 up to 10^8 in standard SERS measurements. The largest enhancement occurs for surfaces which are rough on the nanoscale (10–100 nm).

The main observed features of SERS spectra are the followings: (1) the intensities of the bands generally fall off with increasing vibrational frequency (e.g., C–H stretching vibrations are relatively weak, overtones and combination bands are seldom observed), (2) selection rules are relaxed, even forbidden Raman modes can be detected, (3) the spectra tend to be completely depolarized, (4) excitation profiles differ from the dependence on the fourth power of the frequency, ω_{in}^4, typical of nonresonant Raman scattering, and (5) although the enhancement is mainly concentrated on the first layer of adsorbates, often the effect can extend to a quite long range (up to tens of nanometers).

As suggested before, two mechanisms must be considered to explain SERS, namely the e.m. and the chemical effects. Previous sections describe the foundation of the e.m. effect of the SERS, since the behavior and the properties of metal substrates and particles that sustain the SERS effect can be described using the models presented there. First of all, the dominance of the coinage metals arises from the optical properties of these metals as indicated for Ag in **Figure 9**.

Indeed, the resonance condition with the sp (e.g., $\mathrm{Re}[\varepsilon(\omega)] = -2\varepsilon_0$) is satisfied at visible frequency generally used for Raman spectroscopy. Besides, the imaginary part of the dielectric function, $\mathrm{Im}[\varepsilon(\omega)]$, which is related with losses in the material, is very small at the resonance frequency, thus allowing strong enhancement of the local electric field on the rough surface and colloidal particle made with these metals.

The e.m. model explains many of the experimental features of SERS listed above, including the order of magnitude of the observed enhancement of the Raman intensities. For example, in feature (1), the need to have simultaneously the frequency of the incident and of the scattered light as close as possible to the resonance frequency implies a more effective enhancement for the vibrations with lower Raman shift. The modification of the Raman selection rules as well as the deviation from the ω_{in}^4 dependence of the Raman intensity for SERS are also suggested from the model. The dipole decay law with the distance explains the long-range effect of the SERS enhancement often observed. Finally, the observed depolarization can be easily explained considering that a SERS-active surface is a collection of roughness features of different sizes and shapes into which the molecules adsorb in a variety of orientations and the occurrence of multiple scattering.

Although the e.m. model is able to explain most of the experimental SERS features, some observations strongly support the existence of a second mechanism of Raman enhancement of adsorbates on metal, namely the "chemical effect." Synonyms found in literature for chemical effect in SERS are charge-transfer excitations, short-range effects, atomic-scale roughness, SERS-active sites, adatoms, and energy-transfer excitations. Some SERS measurements indicate a dependence of the observed signal on the special studied adsorbates that cannot be explained by an e.m. effect. For example, the SERS spectra of CO and N_2 molecules that exhibit near identical polarizability differ from a factor 200, hard to explain invoking only the e.m. enhancement. A second evidence supporting the chemical effect is the potential dependence of SERS measured in electrochemical experiments. When the potential is scanned at a fixed laser frequency, or the laser frequency is scanned at a fixed potential, broad resonances are observed.

A simple Raman resonance mechanism can provide the explanation of this chemical effect. It is not unusual that the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the adsorbate, which arise from the chemisorption, are symmetrically positioned in energy with respect to the Fermi level of the metal, as indicated in **Figure 10**.

Since for molecules normally studied by SERS the LUMO is typically in the ultraviolet, the charge transfer excitations (either from the metal to the molecule or vice versa) can occur at about half the energy of the intrinsic intramolecular excitations, namely in the visible region of the spectrum.

In recent years, the advent of new very sensitive Raman instrumentation with high spatial resolution,

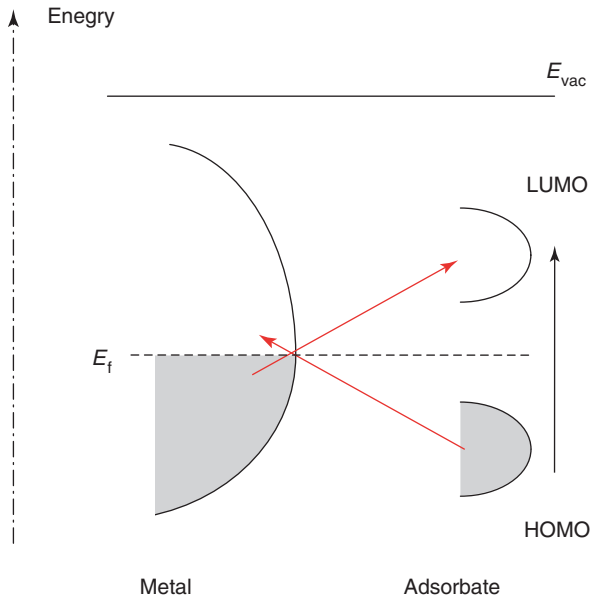


Figure 10 Typical energy level diagram for a molecule adsorbed on a metal surface. Possible charge transfer excitations are shown.

for example, micro-Raman and near-field systems, allowed the selective observation of single particles by SERS. Those experiments demonstrated that SERS from adsorbed species on metal particles show huge enhancement, around 10^{12}–10^{15}, allowing the detection of the Raman signal from a single molecule and opening new fields of applications. In the case of experiments performed on aggregate clusters of colloidal particles, their fractal nature has been invoked to explain the observed enhancement whereas in the case of single particles (in the dimension range of 100–500 nm), the multiplicative effect of resonance Raman enhancement of the molecule and surface enhancement at the metal surface has been considered the main source of the effect. Besides, not all the particles showed an efficient enhancement but only a few called "hot spots" did so. In most cases there is a time dependence of the intensity of every Raman band, which is called "blinking." Finally, this huge Raman enhancement can be reconciled with the more commonly observed value (10^6–10^8) for SERS considering the particles population averaging that occurs in conventional experiments (macro-Raman). These last findings have attracted considerable attention from both basic and practical viewpoints. Indeed, a detailed understanding of the mechanisms operating to generate such a huge local Raman enhancement is not achieved yet and, hence, represents a very interesting area for study. Potential applications are, for example, in analytical chemistry as an ultrasensitive detection technique and in the

development of micro-instrumentation, by combining scanning near-field microscopy with SERS.

See also: Nanostructures, Optical Properties of; Plasmons; Polaritons; Scattering, Inelastic: Raman; Semiconductors, Optical and Infrared Properties of; Time-Resolved Optical Spectroscopies; Treated Surfaces, Optical Properties of.

PACS: 78.; 78.20.−e; 78.20.Ci; 78.30.−j; 33.20.Fb; 68.43.Pq; 78.67.Bf; 78.68.+m; 78.40.Kc; 71.36.+c; 71.45.Gm

Further Reading

Agranovich VM and Mills DL (1982) Surface polaritons: electromagnetic waves at surfaces and interfaces. In: Agranovich VM and Maradudin AA (eds.) *Modern Problem in Condensed Matter Sciences*, vol. I. Amsterdam: North-Holland.

Campion A and Kambhampati P (1998) Surface-enhanced Raman scattering. In: *Chemical Society Reviews*, vol. 27, pp. 241–250. Royal Society of Chemistry.

Cardona M and Güntherodot G (eds.) *Light Scattering in Solids IV*, pp. 289–418. Berlin: Springer.

Chang RK and Furtak TE (eds.) (1982) *Surface Enhanced Raman Scattering*. New York: Plenum.

Kneipp K, Kneipp H, Itzkan I, Dasari RR, and Feld MS (1999) Ultrasensitive chemical analysis by Raman spectroscopy. In: *Chemical Reviews*, vol. 99, pp. 2957–2975. American Chemical Society.

Kneipp K, Kneipp H, Itzkan I, Dasari RR, and Feld MS (2002) Nonlinear Raman probe of single molecules attached to colloidal silver and gold clusters. In: Shalaev VM (ed.) *Optical Properties of Nanostructured Random Media*, Topics Appl. Phys. vol. 82, pp. 227–249. Berlin: Springer.

Kneipp K, Kneipp H, Itzkan I, Dasari RR, and Feld MS (2002) Surface-enhanced Raman scattering and biophysics. *Journal of Physics: Condensed Matter* 14: R597–R624.

Laserna J (ed.) (1996) *Modern Technique in Raman Spectroscopy*. Chichester: Wiley.

Moskovits M (1985) Surface enhanced Raman spectroscopy. In: *Reviews of Modern Physics*, vol. 57, pp. 783–826. American Physical Society.

Otto A (1984) Surface-enhanced Raman scattering: 'classical' and 'chemical' origins. In: Cardona M and Güntherodot G (eds.) *Light Scattering in Solids IV*, pp. 289–418. Berlin: Springer.

Otto A, Mrozek J, Grabhorn H, and Akemann W (1992) Surface-enhanced Raman scattering. *Journal of Physics: Condensed Matter* 4: 1143–1212.

Sánchez-Gil JA, García-Ramos JV and Méndez ER (2000) Microstructures in rough metal surfaces: electromagnetic mechanism in surface-enhanced Raman spectroscopy. In: Moreno F and Gonzáles F (eds.) (1998) *Lectures*, LNP 534, pp. 215–235. Berlin: Springer.

Shalaev VM (1996) Electromagnetic properties of small-particles composites. *Physics Reports* 272: 61–137.

Optical Sum Rules and Kramers–Kronig Relations

F Bassani, Scuola Normale Superiore, Pisa, Italy

Introduction

The optical properties of atoms, molecules, and solid materials have been an active field of research for the past two centuries and have given the most precise information on the microscopic structure of matter. Since the early nineteenth century, it was known that absorption peaks in the light intensity occur at specific frequencies and that a corresponding dispersion occurs in the frequency dependence of the dielectric function $\varepsilon(\omega)$ (or in the refractive index $n(\omega)$).

It may be recalled that in any medium all the optical functions (dispersion, absorption, reflection, etc.) depend on the frequency ω and on the momentum of the photon $\hbar\boldsymbol{k} = \hbar(2\pi/\lambda)$, $\hbar$ being the reduced Planck constant. They can all be expressed in terms of the dielectric functions $\tilde{\varepsilon}(\boldsymbol{k}, \omega)$ extended to the complex plane. From Maxwell equations, it can be seen that its imaginary part can be expressed in terms of the conductivity σ, so that

$$\tilde{\varepsilon}(k,\ \omega) = \varepsilon_1(k,\ \omega) + \mathrm{i}\varepsilon_2(k,\ \omega) = 1 + 4\pi\chi(\boldsymbol{k},\ \omega) + \frac{4\pi\sigma(\boldsymbol{k},\ \omega)}{\omega}\mathrm{i} \qquad [1]$$

where the real susceptibility χ is defined in terms of the polarization $\boldsymbol{P}$ (dipole moment per unit volume) and of the electric field $\boldsymbol{E}$ by

$$\chi(k,\ \omega) = \frac{\boldsymbol{P}(\boldsymbol{k},\ \omega)}{\boldsymbol{E}} \qquad [2]$$

If the susceptibility is also extended to the complex plane to include the conductivity, so that

$$\tilde{\chi}(\boldsymbol{k},\ \omega) = \chi(k,\ \omega) + \frac{\mathrm{i}\sigma(\boldsymbol{k},\ \omega)}{\omega} = \chi_1 + \mathrm{i}\chi_2 \qquad [3]$$

where the subscripts 1 and 2 denote the real and imaginary parts, respectively, the usual expression for the complex dielectric function obtains $\tilde{\varepsilon}(\boldsymbol{k},\ \omega) = 1 + 4\pi\tilde{\chi}(\boldsymbol{k},\ \omega)$. This also defines a complex

refractive index

$$\tilde{n}(\boldsymbol{k}, \omega) = \sqrt{\tilde{\varepsilon}}(k, \omega) = n(\boldsymbol{k}, \omega) + \mathrm{i}\kappa(\boldsymbol{k}, \omega) \qquad [4]$$

The absorption of radiation can be related to the conductivity induced by the electric field by considering the energy lost per unit time per unit volume

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \boldsymbol{J}\cdot\boldsymbol{E} = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t}\cdot\boldsymbol{E} = \sigma E^2 \qquad [5]$$

so that the absorption coefficient $\eta(\omega)$, defined as $\mathrm{d}W/\mathrm{d}x = -\eta W = (\mathrm{d}W/\mathrm{d}t)\ (n/c)$, is given by

$$\eta = \left|\frac{\mathrm{d}W}{\mathrm{d}t}\right| \frac{n}{Wc} = \frac{4\pi\sigma}{nc} = \frac{\omega\varepsilon_2}{nc} = \frac{2\omega\kappa}{c} \qquad [6]$$

Also the reflection and transmission amplitudes can be obtained from the dielectric function through Fresnel relations, which in the case of normal incidence from vacuum to a medium are

$$\frac{E_r}{E_0} = r = \frac{\tilde{n}-1}{\tilde{n}+1}, \qquad \frac{E_t}{E_0} = t = \frac{2}{\tilde{n}+1} \qquad [7]$$

$R = |r|^2$ and $T = n|t|^2$ being the reflectivity and the transmittance, respectively.

In crystalline materials, the optical constants are tensors (second-rank tensors in the linear approximation), but this aspect may be neglected and isotropic media may be considered for the purpose. The results obtained will apply to all tensorial components.

For electromagnetic waves, the magnetic susceptibility is considered unitary and one can also use the dipolar approximation ($\hbar\boldsymbol{k} \simeq 0$) because of the very small light momentum (neglect of spatial dependence).

The above-described optical functions give the electromagnetic properties of matter. Their detailed dependence on frequency is related to the microscopic structure of the material considered and to its interaction with the radiation field. Here, the general properties of the optical functions that depend only on time causality are considered.

Elementary Description of Sum Rules

The traditional phenomenological model for the optical constants is the Lorentz oscillator model, where the equation of motion from the equilibrium position of any specific electron in terms of any oscillator frequencies ω_j and of damping γ_j, is given by

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \omega_j^2 x + \frac{\gamma_j\,\mathrm{d}x}{\mathrm{d}t} = \frac{e\boldsymbol{E}}{m} \qquad [8]$$

From this, for the complex dielectric function in a volume V for a number of resonance frequencies ω_j, one obtains

$$\tilde{\varepsilon}(\omega) = 1 + \frac{4\pi e^2}{mV}\left(\sum_j \frac{f_j}{(\omega_j^2 - \omega^2) - \mathrm{i}\gamma_j\omega}\right) \qquad [9]$$

where the oscillator strength f_j gives the intensity of the resonance frequency ω_j with damping constant γ_j. The conservation of the electron number and the asymptotic behavior ($\omega \to \infty$) of [9], made to coincide with that of any free electrons ($f_i = 1,\ \omega_j = 0,\ \gamma_j = 0$), immediately gives the f-sum rule, first demonstrated in 1925 by W Thomas, F Reiche, and W Kuhn, and called the TRK sum rule:

$$\sum_j f_j = n \qquad [10]$$

This states that the sum of the oscillator strengths of all optical transitions is equal to the total number of electrons n.

Although the TRK sum rule was derived with classical mechanics, it is worth mentioning that the quantum description of the oscillator strength implies that, for the sum rule to hold, the commutation relations between momentum and position operator must also hold, as pointed out by Heisenberg. In fact one can see, by expressing the optical functions in terms of the transition probability rates, observing that electrons are excited from state i to state f, with a probability rate given by the Fermi golden rule

$$P_{if} = \frac{2\pi}{\hbar}|\langle f|eEx|i\rangle|^2\delta(E_f - E_i - \hbar\omega) \qquad [11]$$

that the quantum mechanical expression for the imaginary part of the dielectric function is

$$\varepsilon_2(\omega) = 4\pi^2 \sum_{ij} |\langle f|ex|i\rangle|^2\delta(E_f - E_i - \hbar\omega) \qquad [12]$$

Considering in expression [9] that

$$\lim_{\gamma\to 0} \frac{\gamma}{(\omega_j - \omega)^2 + (\gamma/2)^2} = 2\pi\delta(\omega_j - \omega) \qquad [13]$$

and comparing with [12] it is obtained that the quantum mechanical expression for the oscillator strength is

$$f_{fi} = \frac{2m}{\hbar^2}|\langle f|x|i\rangle|^2(E_f - E_i) \qquad [14]$$

Taking into account the general Heisenberg expression for the time dependence of the displacement x,

$$m\dot{x} = \frac{m}{\mathrm{i}\hbar}[x, H] \qquad [15]$$

where the square brackets denote commutators, and considering the expression for eigenstates,

$$H|j\rangle = E_j|j\rangle \qquad [16]$$

one obtains for the oscillator strength

$$\begin{aligned} f_{fi} &= \frac{2m}{\hbar^2}\langle i|x|f\rangle\langle f|x|i\rangle(E_f - E_i) \\ &= \frac{1}{\mathrm{i}\hbar}(\langle i|x|f\rangle\langle f|p_x|i\rangle \\ &\quad - \langle i|p_x|f\rangle\langle f|x|i\rangle) \end{aligned} \qquad [17]$$

which, summing over the final states, gives

$$\sum_f f_{fi} = \frac{1}{\mathrm{i}\hbar}\langle i|[x, p_x]|i\rangle = 1 \qquad [18]$$

because of Heisenberg commutation rules $[x, p_x] = \mathrm{i}\hbar$. Summing over all occupied initial states, one obtains immediately the TRK sum rule [10].

In general, the oscillator strength depends continuously on the frequency, so that the TRK sum rule can be expressed in terms of the imaginary part of the dielectric function as

$$\int_0^\infty \omega\varepsilon_2(\omega)\,\mathrm{d}\omega = \frac{2\pi^2 e^2}{m}\,\varrho \qquad [19]$$

which relates the total radiation absorbed to the total electron density $\varrho = n/V$.

This is a very important property, because it gives a general constraint which the absorption of the medium must obey. When the sum rule [19] is saturated at a given frequency, no further absorption can occur at higher frequencies. This explains why in light elements (low electron density) hard X-rays are not absorbed and hard γ-rays pass through all materials. The sum rule [19] is also useful to check the results of approximate calculations or experiments. It must always be obeyed by the exact optical functions.

Kramers–Kronig Dispersion Relations

A connection between sum rules and the causality principle was established on the basis of the relation between the real and the imaginary parts of the susceptibility $\tilde{\chi}(\omega)$. This can be obtained by considering the complex ω plane and observing that the susceptibility does not have poles on the upper plane because the polarization can be expressed from the time-dependent response operator $G(\tau)$ as

$$P(t) = \int_{-\infty}^{+\infty} G(\tau)E(t-\tau)\,\mathrm{d}\tau \qquad [20]$$

with the condition $G(\tau) = 0$ for $\tau < 0$ due to time causality. Then, taking the Fourier transform, one obtains

$$\tilde{P}(\omega) = \tilde{\chi}(\omega)E(\omega) \qquad [21]$$

with

$$\tilde{\chi}(\omega) = \int_0^\infty G(\tau)\exp(\mathrm{i}\omega\tau)\,\mathrm{d}\tau \qquad [22]$$

The fact that $\tilde{\chi}(\omega)$ is analytic in the upper complex plane gives, using the Cauchy theorem on a contour closed to infinity where $\tilde{\chi}(\omega)$ goes to zero (see **Figure 1**)

$$\tilde{\chi}(\omega) = \frac{1}{\pi\mathrm{i}}\fint_{-\infty}^{+\infty} \frac{\tilde{\chi}(\omega')}{\omega' - \omega}\,\mathrm{d}\omega' \qquad [23]$$

where the line on the integral denotes that the principal part must be considered. Since $\chi(-\omega) = \chi^*(\omega)$, the standard form of Kramers–Kronig (KK) relations is obtained:

$$\chi_1(\omega) = \frac{2}{\pi}\fint_0^\infty \frac{\omega'\chi_2(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega' \qquad [24a]$$

$$\chi_2(\omega) = \frac{-2\omega}{\pi}\fint_0^\infty \frac{\chi_1(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega' \qquad [24b]$$

In the case of metals, besides the pole at ω, the function $\tilde{\chi}(\omega')/(\omega' - \omega)$ has an additional pole at $\omega' = 0$ because $\sigma(0) \neq 0$. This can be taken into account by considering the function $\tilde{\chi} - (\sigma(0)/\omega)\mathrm{i}$, which does not have such a pole and is analytic

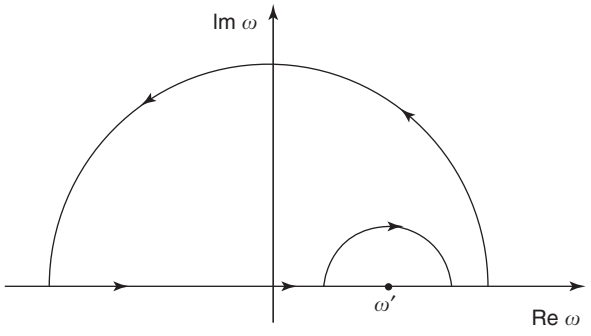


Figure 1 Indication of the contour on the complex plane where the integral of the function $\chi(\omega')/(\omega' - \omega)$ by Cauchy theorem is equal to $\mathrm{i}\pi\chi(\omega)$.

on the upper plane. Then dispersion relations [23] and [24] apply, provided the term $-\sigma(0)/\omega$ is added to the imaginary part of $\tilde{\chi}(\omega)$ and the principal part of the integrals is considered along the entire real axis, including the point at $\omega = 0$, where $\tilde{\chi}_2(\omega')$ has a pole.

The above KK relations are very useful because they relate the dispersion to the absorption. Similar relations can be obtained for the other optical constants $\tilde{\sigma}$, $\tilde{n}(\omega)$, r, and $\mathrm{Im}(1/\varepsilon)$.

They can all be connected to specific sum rules by considering the asymptotic behavior which can be obtained from the KK relations, and comparing it to that which is obtained from a specific model of the material system. Since the asymptotic behaviors must coincide, sum rules are obtained. It can be seen immediately that the TRK sum rule [10] or [19] can be obtained when one considers the asymptotic behavior from [24a] by using the so-called super-convergence theorem, which amounts to letting $\omega' \to 0$ in the denominator, and comparing it with the asymptotic behavior from expression [9] ($\omega_j \to 0$ in the denominator).

As a consequence of the above discussion, one can conclude that sum rules and KK dispersion relations derive from the causality principle only, and from a knowledge of the short-time (or high-frequency) behavior of the system. Complicated properties due to electron–electron interaction or external potentials are not relevant for sum rules and Kramers–Kronig relations. Also the quantum mechanical approach is not essential.

A revival of the field of sum rules started in the early 1970s after the discovery of the refractive index sum rule by Altarelli, Dexter, Nussezveig, and Smith (ADNS), which states for the real part of the refractive index

$$-\!\!\!\!\!\!\int_0^\infty (n(\omega) - 1)\, \mathrm{d}\omega = 0 \qquad [25]$$

After that new sum rules and more KK-type relations were derived for all optical functions and are now being used to interpret a variety of effects. New KK relations and sum rules have also been found in nonlinear optics. Sum rules can now be explored in the entire frequency range using the properties of synchrotron radiation. They are also being used as stringent tests for approximate theories.

KK Relations and Linear Sum Rules

The KK relations [24] for the complex susceptibility function have been derived by direct integration, but causality and KK relations have more direct implications; they can be summarized in what is known as a Titchmarsh theorem, which states the following: any square integrable function $f(\omega)$ which fulfills one of the three conditions below fulfills all three of them:

1. the inverse Fourier transform $F(t)$ of $f(\omega)$ vanishes for $t < 0$,
2. $f(\omega)$ is the limit for $\varepsilon \to 0^+$ of a function $\tilde{f}(\omega + \mathrm{i}\varepsilon)$ that is analytic in the upper half plane and square integrable on the real axis and on any line parallel to it, and
3. $\tilde{f}(\omega)$ verifies the KK identity

$$\tilde{f}(\omega) = \frac{1}{\pi \mathrm{i}} -\!\!\!\!\!\!\int_{-\infty}^{\infty} \frac{\tilde{f}(\omega')}{\omega' - \omega}\, \mathrm{d}\omega' \qquad [26a]$$

or equivalently for $F(t)$ real,

$$f_1(\omega) = \frac{2}{\pi} -\!\!\!\!\!\!\int_0^\infty \frac{\omega' f_2(\omega')}{\omega'^2 - \omega^2}\, \mathrm{d}\omega'$$
$$f_2(\omega) = -\frac{2\omega}{\pi} -\!\!\!\!\!\!\int_0^\infty \frac{f_1(\omega')}{\omega'^2 - \omega^2}\, \mathrm{d}\omega' \qquad [26b]$$

The above theorem proves that KK relations are a direct consequence of time causality, but it also implies that any combinations of analytic functions satisfy KK relations if they are square integrable. For establishing KK relations, it is not necessary to know the explicit form of the causal response function, which is determined by the specific operator being considered, by the Hamiltonian, and by the time-dependent perturbing potential. One only needs to know its asymptotic behavior which can be obtained in the classical approximation from the Lorentz oscillator model, or in the quantum theory from the quantum expression of the response function at time zero. The same holds for the sum rules, which are obtained by comparing the above described asymptotic behavior with that obtained from KK relations.

Consider the optical case in the dipole approximation. The interaction Hamiltonian can then be written as

$$H' = -\boldsymbol{P} \cdot \boldsymbol{E}(t) \qquad [27]$$

where $\boldsymbol{P} = \sum_i e\boldsymbol{r}_i$ is the polarization (it can be used as an operator in quantum mechanics) and $\boldsymbol{E}(t)$ is the electric field, which is taken as a sinusoidal function. While in the classical approximation one considers directly the time dependence of the $\boldsymbol{P}_+$ (Lorentz oscillator model), in the quantum mechanical approach the response function can be obtained from the trace of the density matrix operator ϱ_0 on the ground state

of the system, and to the first order is given by

$$G_p(\tau) = \mathrm{Tr}\left\{\frac{1}{-\mathrm{i}\hbar}\,[P(-\tau),\ \varrho_0]P\right\} \qquad [28]$$

where $P(\tau)$ is defined in the interaction representation. This gives, for the susceptibility, the expression

$$\chi(\omega) = \frac{1}{\hbar}\sum_n |\langle 0|P|n\rangle|^2 \times\left\{\frac{1}{\omega-\omega_{n0}+\mathrm{i}\varepsilon}-\frac{1}{\omega+\omega_{n0}+\mathrm{i}\varepsilon}\right\} \qquad [29]$$

where ε is an arbitrarily small frequency which accounts for the adiabatic switching of the perturbation.

The sum rules can now be obtained from the two expressions of the asymptotic behavior, which must coincide. The first is given by the zero-time response function and its derivatives, and can be obtained integrating by part the Fourier transform of $G(\tau)$, with $G(\infty) = 0$:

$$\underset{\omega\to\infty}{\chi}(\omega) = \int_0^\infty G(\tau)\exp(-\mathrm{i}\omega\tau)\,\mathrm{d}\tau = \frac{G(0^+)}{-\mathrm{i}\omega}+\frac{G'(0^+)}{-\omega^2}+\frac{G''(0^+)}{-\mathrm{i}\omega^3} + \frac{G'''(0^+)}{\omega^4}+\cdots \qquad [30]$$

where the time derivatives are denoted with the apex. The second is obtained from the KK relations [24] with the superconvergence theorem,

$$\underset{\omega\to\infty}{\chi_1} = -\frac{2}{\pi\omega^2}\int \omega'\chi_2(\omega')\,\mathrm{d}\omega' + O(\omega^{-2}) \qquad [31a]$$

$$\underset{\omega\to\infty}{\chi_2} = -\frac{2}{\pi\omega}\int \chi_1(\omega')\,\mathrm{d}\omega' + O(\omega^{-1}) \qquad [31b]$$

By computing the limits of expression [30], it can be seen that $G(0^+) = 0$, $G'(0^+) = \omega_\mathrm{p}^2/4\pi$, etc. One can see immediately that two sum rules are obtained: The first from the asymptotic behavior of the real part is the TRK sum rule

$$\int_0^\infty \omega'\chi_2(\omega')\,\mathrm{d}\omega' = \frac{\pi}{2}\frac{\mathrm{d}}{\mathrm{d}t}G(t)|_{t=0^+} = \frac{1}{8}\omega_\mathrm{p}^2 \qquad [32]$$

where $\omega_\mathrm{p}^2 = 4\pi e^2\varrho/m$. The second refers to the asymptotic behavior [31b] of the imaginary part of the susceptibility which, by comparing with [30] is

$$\int_0^\infty \chi_1(\omega')\,\mathrm{d}\omega' = \frac{\pi}{2}G(t)|_{t=0^+} = 0 \qquad [33]$$

where the second equality can be obtained from the quantum mechanical expression [28] of $G(t)$ or from the Lorentz oscillator model.

In the case of metals, the imaginary part of $\chi(\omega)$ has a pole $\omega = 0$ due to the fact that the conductivity $\sigma(\omega)$ is different from zero at zero frequency. As shown for the KK relations, this can be handled by adding to $\tilde{\chi}(\omega)$ the term $(-(\sigma(0)/\omega)\mathrm{i})$ which cancels the pole, preserving the properties of the function. This is the dominant term in the asymptotic behavior for $\omega\to\infty$, so that, by comparing with [31b] the sum rule for metals is obtained:

$$\int_0^\infty \chi_1(\omega')\,\mathrm{d}\omega' = \frac{\pi}{2}\sigma(0) \qquad [34]$$

Other similar sum rules can be obtained for all optical constants (conductivity σ, refractive index $\tilde{n}$, $\varepsilon^{-1}(\omega)$) in the same way as described above, because they are all related to the real and imaginary parts of the susceptibility. Since conductivity is related to the susceptibility by $\sigma(k,\omega) = -\mathrm{i}\omega\chi$, the sum rule [32] gives directly

$$\int_0^\infty \mathrm{Re}\ \sigma(k,\ \omega)\,\mathrm{d}\omega' = \frac{\omega_\mathrm{p}^2}{8} \qquad [35]$$

which has been used in superconductors. Analogous to eqn [33] is the ADNS sum rule [25] for the real part of the index of refraction, which can be obtained immediately from the observation that the dispersion relations [26b] apply to $\tilde{n}(\omega) - 1$ for both the metallic and nonmetallic media, and that the asymptotic behavior of the complex refractive index is

$$\underset{\omega\to\infty}{\tilde{n}}(\omega) \simeq 1 - \frac{\omega_\mathrm{p}^2}{2\omega^2} + O(\omega^{-2}) \qquad [36]$$

As a consequence, the ADNS sum rule [25] is obtained for both metallic and nonmetallic media.

Longitudinal Sum Rules

Sum rules can also be obtained for the longitudinal dielectric function $\varepsilon_\mathrm{l}(\boldsymbol{k}, \omega)$. This is given by the response of the electron density to an external scalar perturbation, which in $\boldsymbol{k}$ space takes the form

$$\Delta\varrho(\boldsymbol{k}, \omega) = \chi_\mathrm{l}(\boldsymbol{k}, \omega)\phi(\boldsymbol{k}, \omega) \qquad [37]$$

The relation between $\chi_\mathrm{l}(\boldsymbol{k}, \omega)$ and $\varepsilon_\mathrm{l}(\boldsymbol{k}, \omega)$ is obtained from the Maxwell equations

$$\varepsilon_\mathrm{l}^{-1}(k,\ \omega) = 1 + \frac{4\pi e}{k^2}\chi_\mathrm{l}(\boldsymbol{k}, \omega) \qquad [38]$$

A number of phenomena are related to the longitudinal dielectric function, such as the screening of charges, the scattering of electrons, with the expression for the probability for unit time and unit volume that an electron transfers momentum $\boldsymbol{k}$, and energy $\hbar\omega$ to the medium:

$$P(\boldsymbol{k}, \omega) = -\frac{8\pi e^2}{k^2 \hbar V} \operatorname{Im} \varepsilon_{\mathrm{l}}^{-1}(\boldsymbol{k}, \omega) \qquad [39]$$

The sum rules for $\varepsilon_{\mathrm{l}}^{-1}(\boldsymbol{k}, \omega)$ can be obtained with the same procedure used for the optical susceptibility and are

$$\int_0^\infty [\operatorname{Re} \varepsilon_{\mathrm{l}}^{-1}(k, \omega) - 1] \, \mathrm{d}\omega = 0 \qquad [40]$$

and

$$\int_0^\infty \omega \operatorname{Im} \varepsilon_{\mathrm{l}}^{-1}(k, \omega) \, \mathrm{d}\omega = \frac{\pi}{2} \omega_{\mathrm{p}}^2 \qquad [41]$$

where $\omega_{\mathrm{p}}^2 = (4\pi e^2/m)\varrho$ is the plasma frequency. This can be referred to the density of the valence electrons to give the contribution of plasma resonances in metals, but when ϱ includes all the electrons of the medium the sum rule is saturated for $k \to 0$ so that no other longitudinal excitations are present.

Relativistic and Spatial-Dependent Corrections

One can extend the sum rules beyond the dipole approximation, by considering spatial dispersion and the relativistic definition of the current from the Dirac equation. One must consider the fact that $\boldsymbol{k}$ is related to ω by the expression $\varepsilon(k, \omega)\omega^2 = k^2(\omega)c^2$ so that one must take this into account when performing the integration over frequency.

One can be limited to first-order contributions because of the small value of $|\boldsymbol{k}|$ in optics. The lowest-order contribution is of order k^2 because the linear term vanishes due to time reversal. Following Ginzburg and Meiman, a Taylor expansion about $k = 0$ can be obtained as

$$\varepsilon\left(\frac{\partial^2 \varepsilon}{\partial k^2} \frac{\omega\sqrt{\varepsilon}}{c}, \, \omega\right) \simeq \varepsilon(0, \omega) + \frac{\omega^2}{2c^2} \varepsilon\left(\frac{\omega\sqrt{\varepsilon}}{c}, \, \omega\right) \frac{\partial^2 \varepsilon}{\partial k^2}\bigg|_{t=0} \qquad [42]$$

which gives

$$\varepsilon\left(\frac{\omega\sqrt{\varepsilon}}{c}, \, \omega\right) \simeq \frac{\varepsilon(0, \omega)}{1 - (\omega^2/2c^2)\,(\partial^2 \varepsilon/\partial k^2)|_{k=0}} \qquad [43]$$

The calculation of the second derivative of the dielectric function with an inclusion of relativistic corrections has been carried out by Scandolo, Bassani, and Lucarini; it is rather complex and is not to be reported here. When the result is substituted in the frequency dependence [43], its asymptotic behavior is obtained, and the sum rules can be derived in the usual way with spatial dispersion and relativistic corrections. The sum rule on the real part of the susceptibility is unchanged and gives zero, and that on the imaginary part is modified by a contribution which depends on the expectation value of the kinetic energy of the electrons on the total ground state of the system $T_0 = \langle 0|\frac{1}{2}mv^2|0\rangle$. The result is

$$\int_0^\infty \omega' \varepsilon_2(\omega') \, \mathrm{d}\omega' = \frac{\pi}{2} \omega_{\mathrm{p}}^2 \left(1 - \frac{2}{3} \frac{T_0}{mc^2}\right) \qquad [44]$$

This correction is very small, of the order of $\approx 10^{-3}$ in Al, where tests of the sum rule have been made, but it may be easily detected in heavy materials using synchrotron radiation up to very high frequencies.

Applications

Applications of the above described sum rules and KK relations are innumerable and have been of great help in understanding the basic properties of optics in condensed matter. They explain immediately why above a given frequency, all materials are transparent to the electromagnetic radiation.

They show immediately the strict connection between the absorption edge in solids and the value of the static dielectric constant.

The KK dispersion relations have allowed the determination of the optical constants from the measurements of only one of them on the entire optical spectrum (for instance, the reflectivity in semiconductors, where the transmittance cannot be measured because it is too small). In that case, KK dispersion relations can be derived for the reflectivity amplitude. They connect the measured reflectivity R to the phase θ through the function

$$\ln r = \ln |r| + \mathrm{i}\theta = \tfrac{1}{2} \ln R(\omega) + \mathrm{i}\theta(\omega) \qquad [45]$$

which has been proven to have the required analytic behavior in the upper half of the complex ω plane. One can then obtain the phase, and from eqn [7] the complex refractive index and all optical constants. This has allowed the knowledge of the optical transitions in solids, and consequently of the electronic band structure over an extended energy range.

The availability of synchrotron radiation of frequency over a very extended range has furthermore

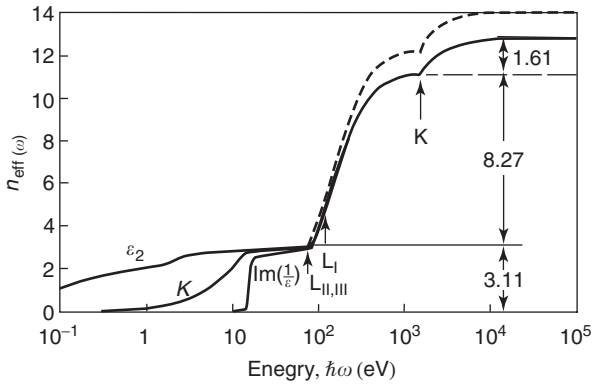


Figure 2 Sum rule for ε_2, κ, and $-\text{Im}(\varepsilon_l^{-1})$: in Al, the solid line shows the saturation at $n = 13$. (E Shiles, T Sasaki, M Inokuti, and DY Smith (1980) *Physical Review B* 22: 1612.)

allowed the direct verification of the sum rules. **Figure 2** reports the results obtained on the metal Al. Further, more detailed analysis on heavy materials would allow the verification of the relativistic and spatial dispersive correction given by expression [44].

See also: Elemental Semiconductors, Electronic States of; Interaction of Light and Matter; Optical Absorption and Reflectance; Optical Properties of Materials; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Optics.

PACS: **78.20. – e; 78.40. – q; 11.55.Hx**

Further Reading

Bassani F and Altarelli M (1983) Interaction of radiation with condensed matter. In: Kock EE (ed.), *Handbook on Synchrotron Radiation*. Amsterdam: North Holland. pp. 463–606.

Bassani F and Pastori-Parravicini G (1975) *Electronic Properties and Optical Transitions in Solids*. Oxford: Pergamon Press.

Greenway DL and Harbeke G (1968) *Optical Properties and Band Structure of Semiconductors*. Oxford: Pergamon Press.

Lucarini V, Bassani F, Peiponen KE, and Saarinen JJ (2003) Dispersion theory and sum rules in linear and nonlinear optics. *Rivista del Nuovo Cimento* 26(12): 1–120.

Scandolo S, Bassani F, and Lucarini V (2001) Spatial dispersion and relativistic effects in the optical sum rules. *European Physical Journal B* 23: 319.

Titchmarsh EC (1968) *The Theory of Functions*. Oxford: Oxford University Press.

Organic Semiconductors

R Scholz, Technische Universität Chemnitz, Chemnitz, Germany

Introduction: Materials

Organic semiconductors are molecular materials with vanishing density of states around the Fermi energy and an energetic gap of ~ 1–5 eV between the highest occupied and the lowest unoccupied states. Contrary to inorganic semiconductors, the optical and transport gaps may differ by more than 1 eV because they are related to different kinds of crystal excitations: the optical response is dominated by Frenkel excitons arising from neutral molecular excitations, and the transport gaps refer to charge transfer (CT) between different molecular sites in the limit of infinite intermolecular distance.

The highest valence band results from the highest occupied molecular orbital (HOMO), and the lowest conduction band from the lowest unoccupied molecular orbital (LUMO). Two contributions to the band dispersion have to be distinguished: strong interactions between adjacent covalently bound building blocks in long oligomers or polymers, and weak intermolecular interactions. The latter give rather small transfer matrix elements between adjacent molecules, restricting the bandwidths of the charge carriers to typically less than 0.5 eV.

Three material classes can be distinguished: polymers, crystalline or amorphous solids composed of identical polycyclic molecules, and organic salts composed of two different kinds of molecules. In the last case, the condensed phase favors electron transfer from the sites with the smaller ionization potential toward the molecules with the stronger electron affinity. The electronic states of polymers and oligomers with a large number of repeating units are covered elsewhere in the encyclopedia (refer "See also" section), and the present article concentrates on the properties of solids composed of identical molecules. Some prototypic planar molecules forming one or several semiconducting crystal phases are discussed, compare **Figure 1**. The optical properties of the monomers are related to the electronic orbitals involved in the dipole-allowed transitions, highlighting the importance of the deformation in the relaxed excited state of the molecule both for the optical response and for resonant Raman spectra.

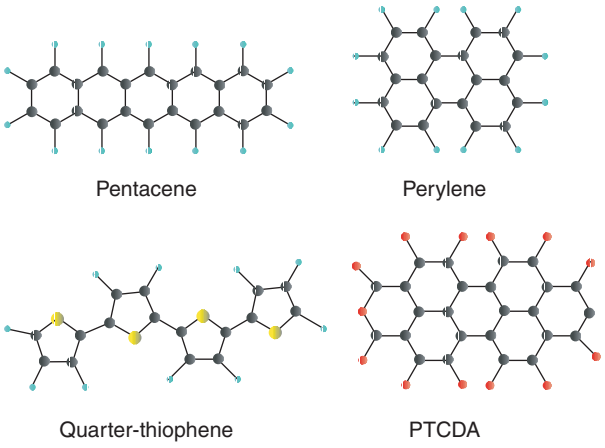


Figure 1 Molecular semiconductors discussed in the present article: pentacene as an example for the polyacene series, perylene, quarter-thiophene (T4), and PTCDA with geometries optimized at the B3LYP/6-31G(d) level.

Based on these ingredients for a single molecule, it is shown that a neutral molecular excitation can be transferred between different sites in the crystal, resulting in excitonic Bloch waves known as Frenkel excitons. For the particular case of 3,4,9,10-perylene, tetracarboxylic dianhydride (PTCDA), it is demonstrated that the anisotropy of the dielectric tensor in the visible and the dominating low-temperature photoluminescence (PL) band can be related quantitatively to the dispersion of the Frenkel excitons arising from the molecular HOMO–LUMO transition.

If CT excitations between adjacent sites in the molecular crystal are energetically below the Frenkel excitons, self-trapped excitons may form, resulting in PL bands with long radiative lifetimes corresponding to the small intermolecular transition dipoles. Two microscopic models can be distinguished: excimers composed of two neutral molecules sharing an electronic excitation, and pairs of oppositely charged molecules. In both cases, the self-trapped excitons can be assigned to pairs of internally deformed molecules involving a reduced intermolecular distance with respect to the crystal geometry.

The electronic band structure of a molecular solid is related to the overlap between the frontier orbitals on adjacent molecules. Starting from a tight-binding approach, it is shown that key parameters such as the bandwidth of the charge carriers show a pronounced temperature dependence, an important ingredient for the interpretation of the observed temperature-dependent mobilities.

Concerning devices based on organic semiconductors, the two major applications are organic field effect transistors and organic light emitting diodes. In the latter case, the disadvantage of rather low-charge carrier mobilities with respect to inorganic semiconductors is counterbalanced by favorable molecular properties such as the tunability across the visible by choosing suitable compounds, the high-internal quantum efficiency, and a relatively low power consumption.

Intermolecular Interactions and Crystal Phases

Several organic crystals with semiconducting properties are composed of planar molecules (see the examples in **Figure 1**), and different polymorphs have been observed for various species. In molecules with partly ionic bonds such as PTCDA, the net charges on the atoms surrounding these bonds result in a large electric quadrupole moment of the entire molecule, determining, in turn, preferential geometries for neighboring molecules in the solid. This quadrupole moment favors an orthogonal arrangement of the two basis molecules in the crystal unit cell, and as a consequence, the unit cell should be quadratic. In both crystal phases of this compound, a symmetry breaking toward a rectangular unit cell allows an interaction of coplanar molecules in the neighboring unit cells via hydrogen bridges, increasing the number of neighbors with similar intermolecular binding energies from 4 to 6.

The oblique orientation of the third lattice vector with respect to the molecule normal reduces the Coulomb repulsion between the stack neighbors, favoring instead short distances between oppositely charged C and O atoms in consecutive crystal planes. The precise stacking geometry results from a combination of the Coulomb interaction between the atomic net charges, a repulsive potential arising from the intermolecular overlap of the π orbitals, and an attractive van der Waals tail determined by the interactions between the molecular polarizabilities. The latter part cannot be obtained by summing over pairs of independent atoms, but the polarizability of each molecule has to be related to the molecular orbitals and the electronic excitations between them. Therefore, a realistic estimate of the interaction energy between the stack neighbors requires advanced quantum chemical approaches such as the second-order Møller–Plesset perturbation theory where the dispersion interaction involving excited Slater determinants is included.

For the case of α-PTCDA (see **Figure 2**), the interactions between neighboring molecules in the crystal are summarized in **Table 1**. The results have been obtained with B3LYP, Hartree–Fock (HF) and the second-order Møller–Plesset perturbation theory (MP2) in the 3–21G variational basis set. In all cases, the

molecular geometries rely on the atomic positions obtained with X-ray diffraction and a readjustment of the hydrogen atoms with a DFT calculation for the three-dimensional crystal. As the elimination of the basis set superposition error results in substantial changes of the interaction energies, only the counterpoise corrected energies are discussed in the following. For the coplanar neighbors, the binding energies depend only weakly on the microscopic approach used, but especially for the stack neighbors, the interaction at the HF and B3LYP level remains repulsive, whereas an MP2 calculation including dispersion interactions gives an overall negative binding energy. From an estimate of the basis set limit of the repulsive HF interaction between two stack neighbors and of their dispersion interaction, it can be concluded that the basis set limit of the MP2 calculation gives a binding energy of -0.91 eV for the stack. Without further modifications of the other intermolecular interactions reported in **Table 1**, this increases the cohesive energy in α-PTCDA at the MP2 level from -2.14 to -3.36 eV. Despite the large size of the PTCDA molecule, the estimated basis set limit of the MP2 binding energy in a stacked dimer is only about twice as large as the MP2 basis set limit obtained for a slipped naphthalene dimer in a stacked geometry, indicating that this energetic range for the van der Waals interaction is typical for various molecular crystals forming stacks. On the other hand, the large intermolecular binding energy between coplanar PTCDA neighbors is a pecularity of the large quadrupole moment and the vicinity of oxygen and hydrogen atoms in both molecules. The strength of these interactions is only weakly dependent on the quantum chemical approach used. Therefore, it can be concluded that the partly ionic carboxylic bonds with negatively charged oxygens and positively charged carbons are of crucial importance for the specific arrangement of the two basis molecules in the unit cell. For molecules without such dipolar groups in the periphery, the interaction between coplanar neighbors would be much smaller, so that noncoplanar arrangements may be more favorable, as observed, for example, for the polyacene series.

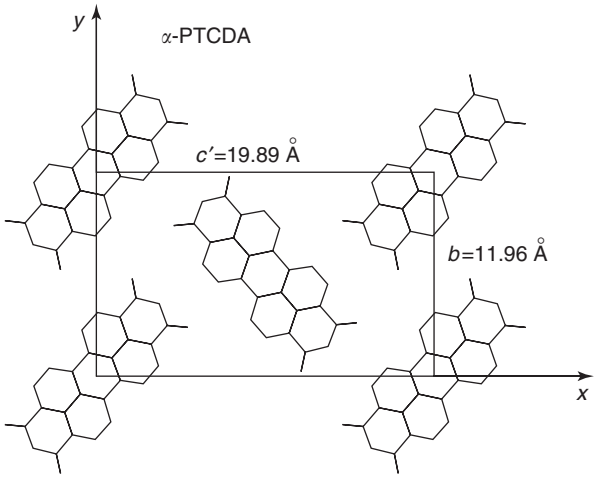


Figure 2 Unit cell of the monoclinic α phase of PTCDA, in the projection coinciding most closely with the molecular orientation. The lattice vectors are $a = 3.72$ Å, $b = 11.96$ Å, and $c = 17.34$ Å, and the direction c' visualized is a linear combination of the lattice vectors c and $2a$.

Frontier Orbitals and Relaxed Excited State of a Single Molecule

The electronic states of polycyclic molecules can be calculated with a large variety of quantum chemical techniques. In this section, the molecules shown in **Figure 1** are investigated with the B3LYP approach in the 6–31G(d) variational basis set as implemented in the GAUSSIAN98 program package. The HOMO and the LUMO are π states resulting from the atomic $2p_z$ orbitals oriented along the normal of the molecular plane. They are composed of lobes with alternating

Table 1 Total energy of different molecular dimers in α-PTCDA

Geometry	*Distance* (Å)	*Type*	*Sites*	*B3LYP* (eV)	*HF* (eV)	*MP2* (eV)
b	11.96	*AA*	2	-0.24 (-0.52)	-0.30 (-0.49)	-0.23 (-0.50)
$a+(b+c)/2$	11.60	*AB*	4	-0.18 (-0.47)	-0.21 (-0.40)	-0.20 (-0.50)
a	3.72	*AA*	2	$+0.76$ ($+0.21$)	$+0.94$ ($+0.41$)	-0.30 (-1.02)
$(b+c)/2$	10.53	*AB*	4	$+0.01$ (-0.13)	$+0.03$ (-0.08)	-0.06 (-0.19)
$a-(b+c)/2$	13.64	*AB*	4	-0.01 (-0.03)	-0.02 (-0.02)	-0.02 (-0.02)
Total			16	$+0.33$ (-3.15)	$+0.48$ (-2.16)	-2.14 (-5.88)

Both for the nearly coplanar molecules (first two lines) and for the interactions between consecutive layers of the crystal (next three lines), the type of the basis molecules involved is indicated by *AA* or *AB,* respectively. Multiplied by the number of the respective neighboring sites and summed over the five types of dimers investigated, these dimer energies result in an estimate for the total cohesive energy in the crystalline phase (last line). All results have been obtained with the GAUSSIAN98 software package using the 3–21G variational basis set, with (without) counterpoise corrections of the basis set superposition error.

sign, and as both orbitals are in the central energetic region of all the π states, each lobe is typically delocalized over two atomic sites (cf. **Figures 3–6**). For perylene and PTCDA, in each half of the molecule the higher-lying LUMO has one node plane and one lobe more than the HOMO, so that the energetic ordering is in accordance with the well-known node counting rules.

The deformation patterns shown in **Figures 3–6** correspond to the difference between a CIS/6–31G(d) calculation for the minimum of the excited state potential and the HF/6–31G(d) geometry for the electronic ground state. In all cases, the $S_0 \rightarrow S_1$ excitation arises from the dipole-allowed HOMO–LUMO transition. For quarter-thiophene, perylene, and PTCDA, the transition dipole moment is oriented along the long axis of the molecules, whereas the lowest electronic transition in pentacene has a transition dipole moment along the short axis.

The deformation pattern in the relaxed excited state of each molecule can be understood from the bonding regions and nodal planes of HOMO and LUMO: for bonds where a bonding region of the HOMO is replaced by an antibonding node of the LUMO, the bond length increases, whereas nodes of the HOMO replaced by a bonding region of the

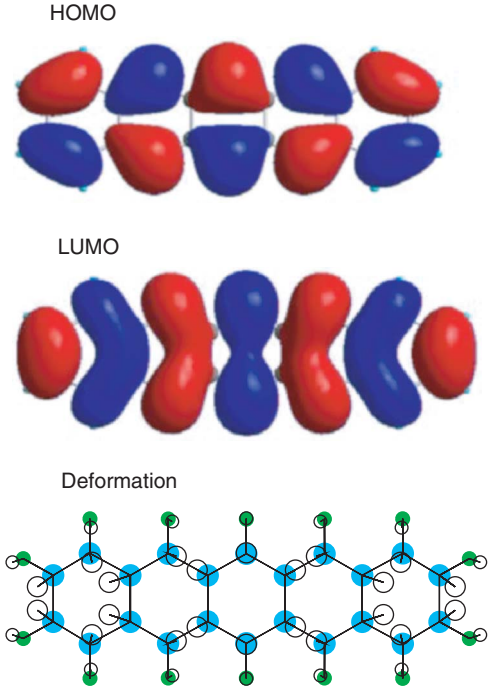


Figure 3 HOMO, LUMO, and deformation in the relaxed excited state of pentacene, multiplied by 20. The orbitals have been calculated with B3LYP/6–31G(d) in the geometry optimized with the same method, and the deformation is the difference between geometry optimizations using CIS/6–31G(d) for the excited state and HF/6–31G(d) for the ground state.

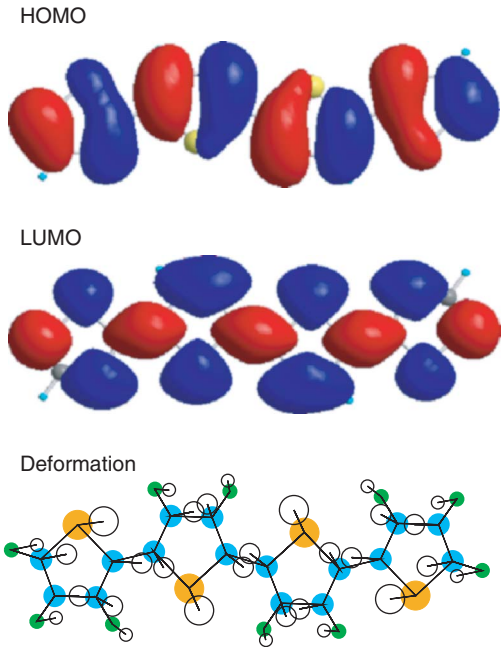


Figure 4 HOMO, LUMO, and deformation in the relaxed excited state of quarter-thiophene.

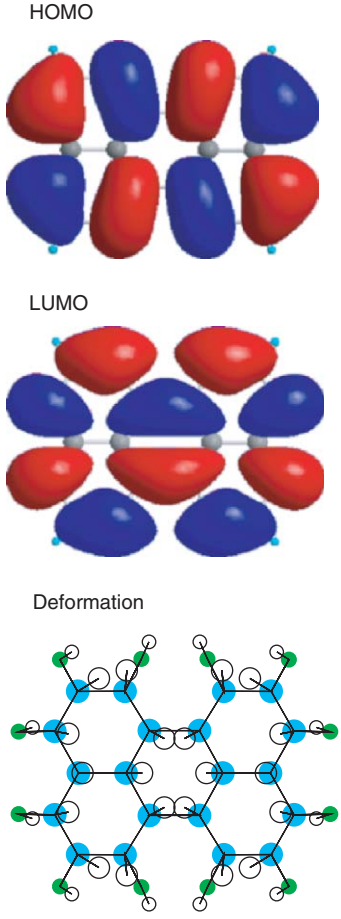


Figure 5 HOMO, LUMO, and deformation in the relaxed excited state of perylene.

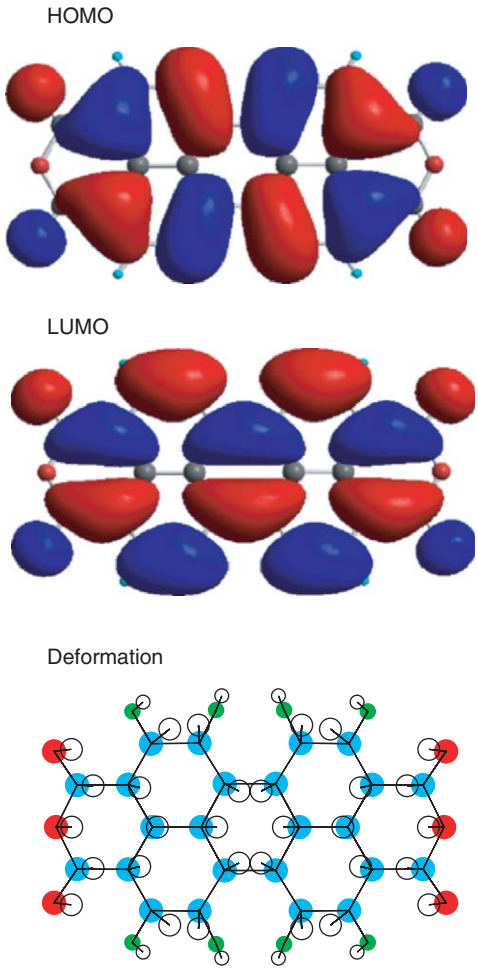


Figure 6 HOMO, LUMO, and deformation in the relaxed excited state of PTCDA.

LUMO result in a shorter bond. As the commutative point groups of the four molecules shown contain the inversion, the charge density of each orbital transforms according to the totally symmetric representation A_g, restricting the deformation in the relaxed excited state to the same symmetry.

Vibrational Spectroscopy

Infrared Spectroscopy

For each molecular species, the vibrational properties of the molecules give a unique fingerprint related to the specific bonding mechanisms involved. The strongest response of the infrared (IR)-active modes arises from bonds with a significant ionic CT, so that IR spectroscopy gives direct insight into the charge distribution over the molecule, and as discussed in an earlier section, the resulting molecular quadrupole moment has a significant influence on the intermolecular interactions in the condensed phase.

In **Figure 7**, the calculated IR spectra of the molecules visualized in **Figure 1** are reported on the same scale. This comparison reveals that the covalent bonds in pentacene and perylene result in quite small net charges on the constituting atoms. The calculated IR activity of quarter-thiophene arises mainly from two out-of-plane C–H bending modes at 696 and 802 cm^{-1}, a symmetric C–S–C in-plane stretch mode at 830 cm^{-1}, a mode at 1245 cm^{-1} where the pattern is a superposition of C–S stretching and C–C stretching of the bonds between pentagons, and a mode at 1557 cm^{-1} dominated by C–C stretching in the pentagons and C–H bending. The most interesting case is PTCDA, where the vibrational fingerprint is largely determined by the O=C–O–C=O end groups. The asymmetric C–O–C stretch contributes strongly to the modes at 1070 and 1173 cm^{-1}, the symmetric C–O–C stretch to the modes at 1151 cm^{-1}, 1171 cm^{-1}, and 1323 cm^{-1}, where in each group the mode with the highest frequency has significant contributions from C–C stretching modes involving the carbon atoms of the carobxylic groups. Among the IR modes with the highest frequencies, the one at 1645 cm^{-1} has mainly elongations of the carbon atoms, and the two strong modes at 1821 cm^{-1} and 1854 cm^{-1} are C=O stretching modes. Even though the 6–31G(d) basis set used for the calculation of the vibrational properties is reasonably well converged, the calculated IR activity of the out-of-plane modes below 1000 cm^{-1} is far below the observed spectra for crystalline material, indicating that in the free molecule the most strongly elongated H-atoms do not carry a significant charge. However, in the crystalline phase, the CT between oxygen and hydrogen atoms of neighboring molecules results in small positive net charges on the hydrogens, with a strong impact on the IR activity of the out-of-plane C–H wagging modes.

Resonant Raman Spectroscopy and Vibronic Progressions in Absorption and PL

For the Raman-active breathing modes, the elongation in the relaxed excited state of each molecule can be obtained from a projection of the deformation pattern in **Figures 3–6** on the complete set of vibrational eigenvectors of each molecule. As a result, for each of the A_g-symmetric breathing modes $\hbar\omega_j$, the corresponding reorganization energies

$$\lambda_j = g_j^2 \hbar\omega_j \qquad [1]$$

can be expressed in terms of a dimensionless vibronic coupling constant g_j^2. This quantity determines the transition probabilities from the electronic and vibrational ground state $|0_g\rangle$ to the different vibronic

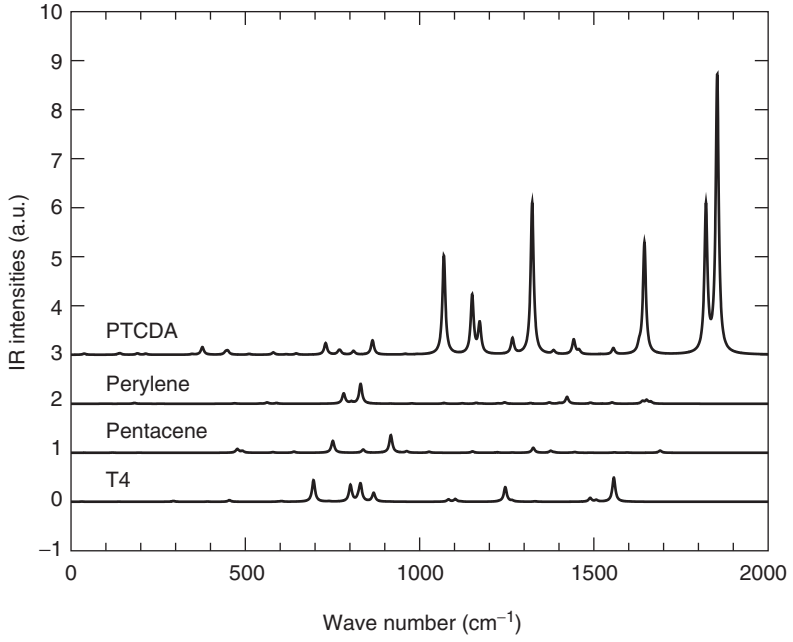


Figure 7 IR intensities calculated at the B3LYP/6–31G(d) level. For visualization purposes, the IR-active modes are broadened with Lorentzian lineshapes with a full width half maximum (FWHM) of 10 cm^{-1}, and all curves are reported on the same but arbitrary scale. From bottom to top: quarter-thiophene, pentacene, perylene, and PTCDA, as annotated. Due to inherent approximations of the B3LYP method, the calculated mode frequencies shown in this figure are too high by ~4%, and as a consequence, it is common practice to rescale them by a factor of ~0.96. The reference lines of the different molecules have been shifted for clarity.

levels $|v_e\rangle$ associated to the excited state potential:

$$P(|0_g\rangle \rightarrow |v_e\rangle) \propto \langle v_e|0_g\rangle^2 = P_v(g_j^2) = \frac{g_j^{2v}}{v!}e^{-g_j^2} \qquad [2]$$

where $P_v(g_j^2)$ is a Poisson distribution. If several internal modes are elongated in the relaxed excited state geometry, the calculation of the linear optical response requires the multiplication of the transition probabilities associated to the transitions $|0_g\rangle \rightarrow |v_e\rangle$ of different breathing modes $\hbar\omega_j$. The transition probabilites $P(|0_e\rangle \rightarrow |v_g\rangle)$ observed in PL can be described by a similar equation:

$$P(|0_e\rangle \rightarrow |v_g\rangle) \propto \langle v_g|0_e\rangle^2 = P_v(g_j^2) = \frac{g_j^{2v}}{v!}e^{-g_j^2} \qquad [3]$$

To leading order, the resonant Raman cross section of each mode is proportional to the vibronic coupling constant g_j^2,

$$\sigma_R(\hbar\omega_j) \propto g_j^2(\hbar\omega_j)^2[1 + n_{th}(\hbar\omega, k_BT)] \qquad [4]$$

where n_{th} $(\hbar\omega_j, k_BT)$ is the number of thermally excited vibrational quanta at temperature T according to Bose–Einstein statistics. Equations [2] and [4] reveal that the linear optical properties, PL and resonant Raman spectra give complementary information on the coupling between electronic excitations and vibrational modes.

Albeit the experimental Raman spectra are usually obtained from the crystalline phase, calculations for a single molecule remain a reasonable approximation because the intermolecular influence on the internal vibrational properties is rather weak.

Clear vibronic progressions can be obtained from the absorption spectra of frozen solutions, laser-induced fluorescence spectra in weakly interacting solvents such as superfluid helium, and PL at low temperature. For excitations to higher vibronic levels, nonradiative recombination mechanisms may reduce the laser-induced fluorescence signals with respect to the absorption coefficient so that the PL excitation spectra deviate considerably from the absorption coefficient. **Figure 8** shows the PL spectra of quarter-thiophene in a frozen solution of tetradecane at a temperature $T = 4.2$ K, allowing a clear assignment of the elongations of various internal breathing modes.

For more strongly interacting solvents and higher temperature, the absorption and PL spectra of dissolved monomers are considerably broadened, so that individual vibronic lines merge into broader bands. An example of this behavior is dissolved PTCDA, where the most strongly elongated breathing modes between 1100 and 1700 cm^{-1} merge into a vibronic progression with a spacing of ~1400 cm^{-1},

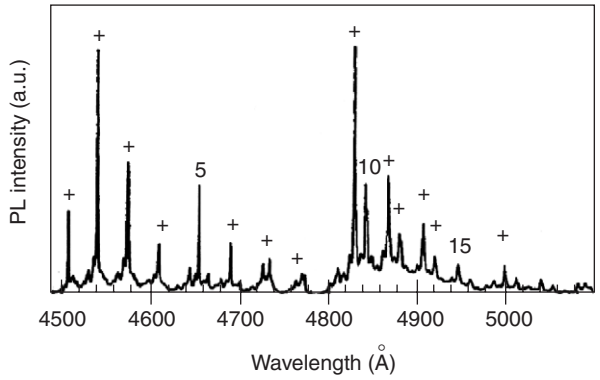


Figure 8 PL spectrum of a frozen solution of quarter-thiophene in tetradecane at $T = 4.2$ K, as a function of wavelength. The PL has been excited selectively on the lowest transition of one of the four different sites resolved in PL excitation spectra. The difference of 1478 cm^{-1} between the peaks 9 and 1 corresponds to the calculated mode at 1503 cm^{-1} in **Figure 9**. (After Birnbaum D, Fichou D, and Kohler BE (1992) The lowest energy singlet state of tetrathiophene, an oligomer of polythiophene. *Journal of Chemical Physics* 96: 165.)

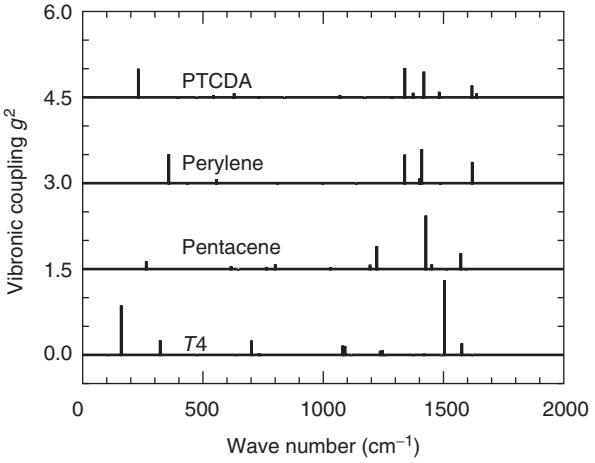


Figure 9 Vibronic coupling constants g_j^2 for the internal breathing modes $\hbar\omega_j$, calculated from a projection of the deformation between the CIS/6–31G(d) relaxed excited state geometry and the HF/6–31G(d) ground-state geometry on the librational eigenvectors obtained for the ground-state potential energy surface with B3LYP/6–31G(d). The reference lines for the different molecules have been shifted for clarity.

interpreted as the elongation of an effective internal vibration.

Figure 9 shows the vibronic coupling constants g_j^2 of the breathing modes in the ground-state potential related to the reorganization energy of each mode according to eqn [1]. As PL and resonant Raman spectroscopy probe the vibrational properties on this potential surface, both spectra can be calculated with these parameters according to eqns [2] and [4].

On the other hand, the modeling of vibronic progressions in optical absorption requires a knowledge of the vibrational properties in the excited state potential. Especially for floppy molecules such as quarter-thiophene, the breathing modes may be quite different, partly due to changes of the potential curvature, and partly due to Dushinksy rotations between the vibrational eigenvectors on the ground- and excited-state potential surfaces.

For all the four molecules visualized in **Figure 9**, a low-frequency mode below 400 cm^{-1} is strongly elongated. In each case, it corresponds to a compressive in-phase movement of all atoms, with the elongations oriented mainly along the long axis of the molecule. The modes in the region 500–800 cm^{-1} have a similar in-phase pattern oriented along the short axis. In quarter-thiophene, the dominating mode at 1503 cm^{-1} is a combination of C–C stretching with a symmetric C–S–C bending motion. For the other three molecules, the breathing modes between 1100 and 1700 cm^{-1} involve a combination of C–C stretching modes and C–H bending modes, and their elongation is driven by the modulation of the C–C bond lengths in the relaxed excited state geometry. The similarity between the modes contributing to perylene and PTCDA is a direct consequence of the similarity of the HOMO and LUMO in both compounds, and of the fact that the terminal O = C–O–C = O groups of PTCDA produce vibrational modes with very small elongations in the perylene core where the differences between the HOMO and LUMO patterns provide the driving force for the elongations.

Absorption and PL Spectra in the Crystalline Phase

For molecular crystals with weak intermolecular interactions such as quarter-thiophene, the low-temperature absorption and PL spectra show clear vibronic progressions of internal vibrations resembling dissolved monomers, (see **Figure 10**). On the other hand, in crystals with strong intermolecular interactions such as PTCDA, the optical excitation induces large elongations of external phonon modes which can be mapped with resonant Raman spectroscopy. It can be shown that these elongations of the librational phonons and the lowest internal breathing modes allow a quantitative assignment of the large broadenings of the absorption spectra observed even at low temperature. As a matter of fact, already at low temperature the fine structure of the vibronic progression is hidden under a line width with FWHM of $\sim$80 meV, so that the high-frequency breathing modes dominating the resonant Raman

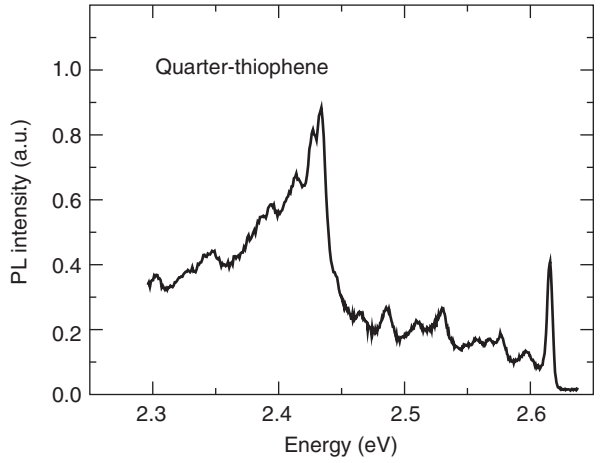


Figure 10 PL spectrum of a 25 Å thick film of quarter-thiophene on highly oriented pyrolitic graphite recorded at $T = 20$ K. The difference of 1469 cm^{-1} between the two most prominent peaks corresponds to the calculated mode at 1503 cm^{-1} in **Figure 9**. (After Gebauer W, Väterlein C, Soukopp A, Sokolowski M, and Umbach E (1996) High-resolution optical spectra from ultrathin quaterthiophene films. *Thin Solid Films* 284–285: 576.)

spectra merge again into an effective high-frequency internal vibration. In the solid phase, the higher harmonics of the internal vibrations are subject to additional broadening mechanisms induced by a combination of exciton transfer and spreading of the internal deformation over molecules surrounding the optically excited site.

The stronger intermolecular interactions in crystalline PTCDA reduce the PL efficiency with respect to the monomer. On the other hand, these interactions are also responsible for the fact that single crystals of α-PTCDA show a large variety of radiative recombination channels (see **Figure 11**). For the four strongest PL bands, the observed decay time τ can be fitted to a simple model for radiative recombination with nonradiative PL quenching at higher temperatures:

$$\frac{1}{\tau} = \gamma = \gamma_{\mathrm{rad}} + \gamma_{\mathrm{nrad}} \exp\left[\frac{-\Delta_{\mathrm{act}}}{k_{\mathrm{B}} T}\right] \qquad [5]$$

where $\gamma_{\mathrm{rad}} = \tau_{\mathrm{rad}}^{-1}$ is the inverse of the radiative lifetime, γ_{nrad} a nonradiative recombination rate, and Δ_{act} the corresponding activation barrier (see **Figure 12**).

As discussed in the next section, the fastest recombination channel with a radiative decay time of 13 ns arises from a vertical recombination starting at the dispersion minimum of the Frenkel excitons. The three PL bands with longer radiative lifetimes can be assigned to self-trapped excitons involving molecular dimers with internal deformations and a modified intermolecular distance (for further details see the section "Slow PL channels: self-trapped excitons").

Exciton Models for the Crystalline Phase

Optical Cycle of Frenkel Excitons

In the solid phase, the molecular HOMO–LUMO excitations can be transferred to different molecules, resulting in neutral crystal excitations called Frenkel excitons. The transfer matrix elements arise from the interaction between the molecular HOMO–LUMO transition dipoles, and for a quantitatively meaningful description, it is important to go beyond the point dipole approximation by including the delocalization of HOMO and LUMO over the entire molecular area, compare **Figures 3–6**. This can be achieved by discretizing the HOMO–LUMO transition dipole moment into a distribution of atomic overlap charges.

As the deformation in the relaxed excited state has a major impact on the absorption spectrum of the monomer, a simple scheme for the internal deformation has to be merged with the possibility of transferring molecular excitations. If the relevant breathing modes cluster in a narrow range as for a PTCDA molecule, it is convenient to introduce an effective internal mode $\hbar\omega_{\mathrm{eff}} = \sum_j g_j^2 \hbar\omega_j / \sum_j g_j^2$ with an effective vibronic coupling constant $g_{\mathrm{eff}}^2 = \sum_j g_j^2$, where the sum may be restricted to the most strongly elongated modes, for example, in the range 1100–1700 cm^{-1}. It has been shown that for crystalline PTCDA, an effective mode described by $\hbar\omega_{\mathrm{eff}} = 0.17\,\mathrm{eV} = 1370\,\mathrm{cm}^{-1}$ and $g_{\mathrm{eff}}^2 = 1.0$ gives a realistic description. Within this effective mode model, the elementary molecular excitations can be described by Bosonic creation and annihilation operators $b_{n\alpha\nu}^{+}$ and $b_{n\alpha\nu}$, respectively, where

$$b_{n\alpha\nu_{\mathrm{e}}}^{+} = |n\alpha\nu_{\mathrm{e}}\rangle\langle n\alpha 0_{\mathrm{g}}| \qquad [6]$$

excites the basis molecule α in the unit cell at the lattice site n from the electronic and librational ground state $|0_{\mathrm{g}}\rangle$ to the vibronic sublevel $|\nu_{\mathrm{e}}\rangle$ on the excited state potential. Based on these operators, a Hamiltonian including the possibility of transferring neutral excitations between different molecules can be written in the following form:

$$H^{\mathrm{e}} = \sum_{n} \sum_{\alpha=A,B} \sum_{\nu_{\mathrm{e}}} E_{0_{\mathrm{g}}\nu_{\mathrm{e}}} b_{n\alpha\nu_{\mathrm{e}}}^{+} b_{n\alpha\nu_{\mathrm{e}}} + \sum_{n\alpha\nu_{\mathrm{e}}} \sum_{m\beta\mu_{\mathrm{e}}} t_{n\alpha\nu_{\mathrm{e}};m\beta\mu_{\mathrm{e}}} b_{n\alpha\nu_{\mathrm{e}}}^{+} b_{m\beta\mu_{\mathrm{e}}} \qquad [7]$$

where the transfer matrix elements $t_{n\alpha\nu_{\mathrm{e}};m\beta\mu_{\mathrm{e}}}$ are the product of an electronic and a vibronic part,

$$t_{n\alpha\nu_{\mathrm{e}};m\beta\mu_{\mathrm{e}}} = T_{n\alpha;m\beta} S_{\nu_{\mathrm{e}}0_{\mathrm{g}}} S_{\mu_{\mathrm{e}}0_{\mathrm{g}}} \qquad [8]$$

$S_{\nu_{\mathrm{e}}0_{\mathrm{g}}} = \langle \nu_{\mathrm{e}}|0_{\mathrm{g}}\rangle$ and $S_{\mu_{\mathrm{e}}0_{\mathrm{g}}} = \langle \mu_{\mathrm{e}}|0_{\mathrm{g}}\rangle$ are the vibronic Franck–Condon factors of the effective internal

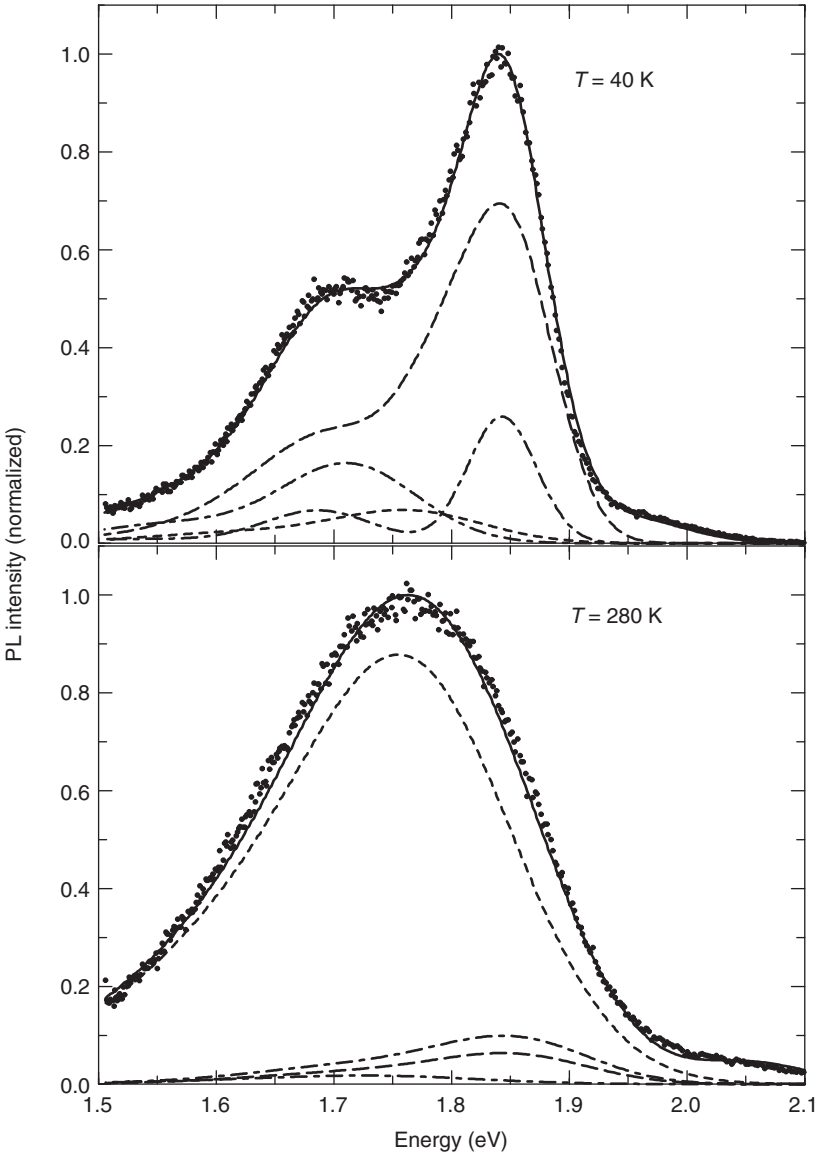


Figure 11 Decomposition of the measured time-integrated PL spectra (dots) at $T = 40$ K (upper) and $T = 280$ K (lower) into contributions from Frenkel excitons (long-dashed), excimers (short-dashed), and two CT channels assigned to recombination between an anionic and a cationic molecule (dash-dotted). The time-integrated spectra include a high-energy satellite ~ 1.95 eV at $T = 40$ K and ~ 2.05 eV at $T = 280$ K (not shown separately). (After Kobitski A Yu, Scholz R, Wagner HP, and Zahn DRT (2003) Time-resloved photoluminescence study of excitons in α-PTCDA as a function of temperature. *Physical Review B* 68: 155201.)

mode, and their squares follow a Poisson progression $S^2_{v_e 0_g} = P_{v_e}(g^2_{\text{eff}})$ according to eqn [2]. As the transfer matrix elements result from the interaction of the molecular transition dipoles, the neighboring molecules in the crystal give the largest values. For pairs of equivalent basis molecules in α-PTCDA, it can be shown that the stack neighbors exceed all other transfer matrix elements by more than one order of magnitude. Even though the transfer between both types of basis molecules remains much smaller, it induces a pronounced difference between the lineshapes of the diagonal components of the frequency-dependent dielectric tensor.

The above Hamiltonian equation [7] is based on the simplifying approximation that the electronic excitation and the deformation reside on the same molecule, so that they are transferred together. It can be block-diagonalized by a Fourier transformation to k-space, resulting in dispersion relations for the different vibronic levels of the Frenkel exciton. In the usual approximation of a vanishing photon wave vector with respect to the size of the Brillouin zone, the optical excitation results in a Frenkel exciton at the Γ point. The HOMO–LUMO transition dipole moments are oriented along the long axes of the two basis molecules, defining the plane where this transition

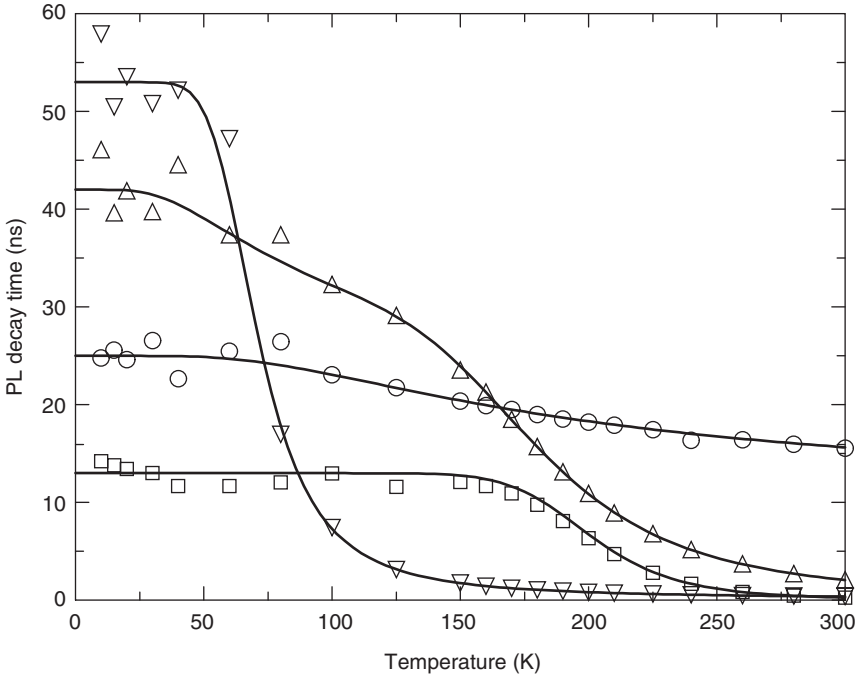


Figure 12 Lifetimes of the different PL bands, where the model curves according to eqn [5] are calculated with $\tau_{rad} = 13$ ns for the Frenkel excitons, 42 ns for the high-energy CT band around $\langle E_{PL} \rangle = 1.78$ eV, 53 ns for the low-energy CT band around $\langle E_{PL} \rangle = 1.67$ eV, and 25 ns for the PL arising from the stacked excimer. The model curve for the higher CT band involves two nonradiative decay mechanisms. (After Kobitski A Yu, Scholz R, Wagner HP, and Zahn DRT (2003) *Physical Review B* 68: 155201.)

contributes to the optical response. Due to the specific arrangement of the two molecules in the unit cell, this plane is very close to the plane visualized in **Figure 2**.

The dielectric tensor resulting from the above Hamiltonian equation [7] and realistic transfer matrix elements is visualized in **Figure 13**. In addition to the molecular ingredients discussed above, the two crystal parameters determining the optical response are the sum over all transfer matrix elements between equivalent lattice sites, $T_{AA} = 150$ meV, and the sum over all transfers between inequivalent basis molecules, $T_{AB} = 46$ meV. The redistribution of oscillator strength between the different vibronic sublevels is governed by $T_{AA} - T_{AB} = 104$ meV for ε_{xx}, and by $T_{AA} + T_{AB} = 196$ meV for ε_{yy}. These values result in different lineshapes of the two diagonal elements of the dielectric tensor, but the Davydov splitting between the vibronic sublevels remains rather small. For polycrystalline films with random azimuthal orientation of the crystallites, the average $\varepsilon_{||}(\omega)$ can be used for the definition of the refractive index $n_{||}(\omega)$, (see **Figure 14**).

The above approach to exciton transfer can easily be generalized to the exciton dispersion in the entire Brillouin zone displayed in **Figure 15**. After the optical excitation toward the dispersion maximum at Γ, the Frenkel excitons can be scattered by librational modes, resulting eventually in a thermalized exciton population around the dispersion minimum at the boundary of the Brillouin zone. These low-lying exciton states are obvious candidates for a strongly redshifted PL band.

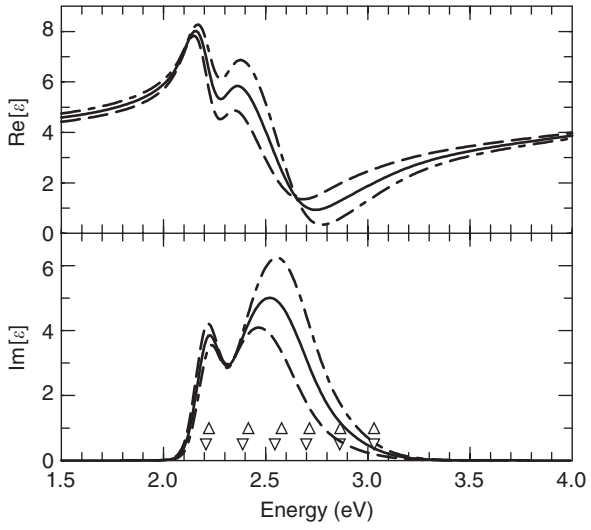


Figure 13 Calculated diagonal components of the frequency-dependent dielectric tensor in α-PTCDA. Dashed: $\varepsilon_{xx}(\omega)$, dash-dotted: $\varepsilon_{yy}(\omega)$, solid: average $\varepsilon_{||}(\omega)$. The △ and ▽ signs indicate the positions of the upper and lower Davydov components, contributing to $\varepsilon_{yy}(\omega)$ and $\varepsilon_{xx}(\omega)$, respectively. (After Vragović I and Scholz R (2003) Frenkel exciton model of optical absorption and photoluminescence in α-PTCDA. *Physical Review B* 68: 155202.)

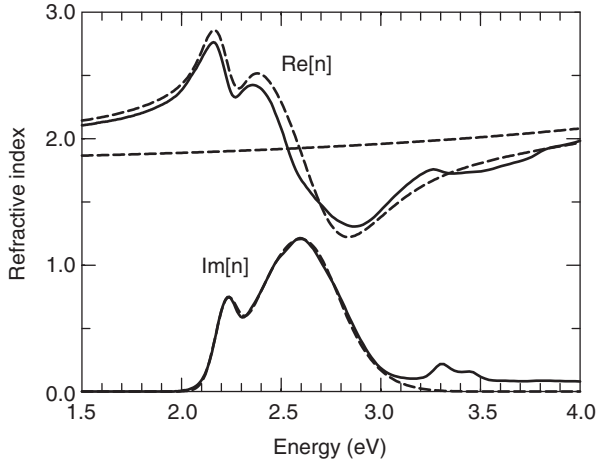


Figure 14 Real and imaginary parts of the refractive index of PTCDA films. Solid: experimental data obtained on PTCDA films deposited on quartz glass, dashed: present model calculation at the Γ point of the Brillouin zone for $n_{||}(\omega) = [\varepsilon_{||}(\omega)]^{1/2}$, and background of the real part of the refractive index. (After Vragović I and Scholz R (2003) *Physcial Review B* 68: 155202.)

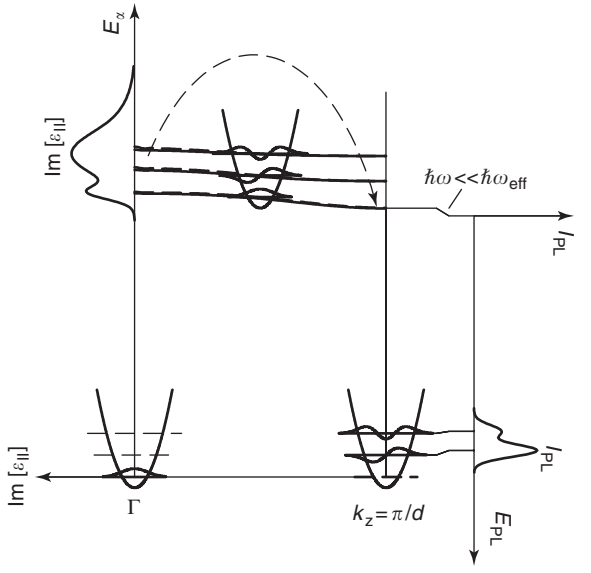


Figure 15 Schematic model of the optical excitation and recombination cycle in crystalline PTCDA at low temperature. Vibronic levels ν_g not involved in the absorption and PL transitions are visualized as dashed lines. (After Vragović I and Scholz R (2003) *Physical Review B* 68: 155202.)

Assuming a vertical PL transition, the final state after recombination has to conserve the wave vector of the thermalized Frenkel exciton, so that the electronic and vibronic ground state with a vanishing wave vector can be excluded as an acceptable final state. This lack of the lowest final state gives the largest contribution of 0.17 eV to the overall Stokes shift of 0.41 eV between the lowest absorption subband and the highest PL feature, followed by the dispersion of the Frenkel excitons extending over a range of 0.12 eV. A Stokes shift of 0.09 eV related to the low-frequency internal vibrations and to the external phonons can be estimated from resonant Raman spectra, and the correlation between different vibronic sublevels of the final states induced by exciton transfer is about 0.03 eV. The observed radiative lifetime of $\tau_{rad} = 13$ ns can be reproduced quantitatively, and the slowing down with respect to dissolved monomers with $\tau_{rad} = 4$ ns results from the lack of the dipolar coupling to the final state $|0_g\rangle$. Electron energy loss spectroscopy can be used to map the dependence of the dipolar coupling strength on the wave vector, corroborating the validity of the excitonic dispersion in **Figure 15** and the calculated distribution of the vibronic overlap factors over the different branches.

Slow PL Channels: Self-Trapped Excitons

In several organic crystals with small intermolecular distances, PL channels with decay times exceeding 20 ns have been observed (see **Figures 11** and **12** for α-PTCDA). As this range for the radiative recombination time is far above the PL decay time observed on dissolved monomers of the same species, the molecular HOMO–LUMO transition dipole can be ruled out as a possible assignment. On the other hand, intermolecular CT transitions are expected to have very small transition dipoles, and for α-PTCDA, it can be shown that the three slower PL channels arise from different types of deformed excited dimer states: The dominating high-temperature PL band results from an excimer in a stack geometry, the lower CT band from the corresponding anion–cation pair, and the higher CT band from the two basis molecules in the unit cell involving an anionic and a cationic molecule. In each case, the Stokes shift with respect to the corresponding CT transition in the undeformed crystal is composed of three contributions: an internal deformation of the two constituting molecules, an elongation of the external librational phonon modes, and self-trapping along the intermolecular distance.

For the stack excimer, the internal deformation of the two molecules can be defined in accordance with the deformation pattern in **Figure 6**, and as both molecules share the electronic excitation, the minimum on the excited state potential obtained with the time-dependent density-functional theory (TD-DFT) corresponds to about 45% of the deformation of the relaxed excited state of the monomer, giving a contribution of 0.21 eV to the Stokes shift.

From their resonant Raman intensities, the librational phonons are expected to contribute 0.04 eV to

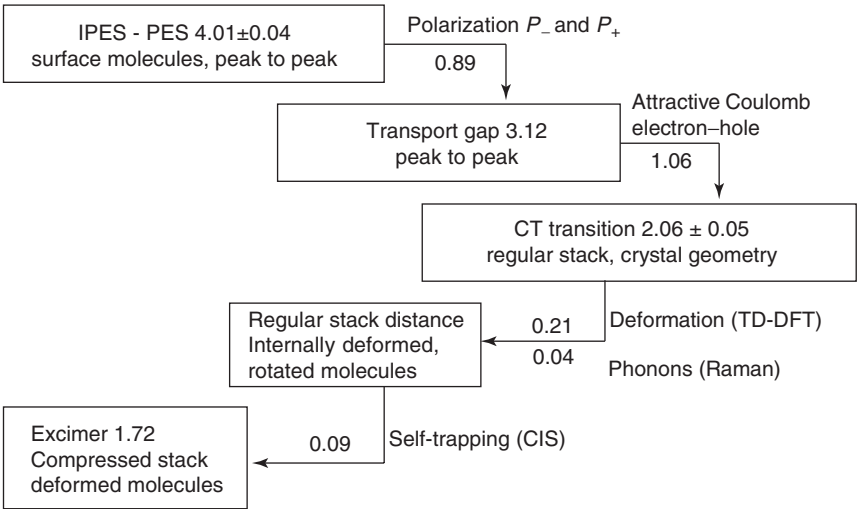


Figure 16 CT energies related to the stack geometry, in eV. The energies for the stack excimer and the undeformed stack are derived from the dimer model for a self-trapped exciton discussed in the main text, and the Coulomb attraction between the electron and hole is taken from previous microelectrostatic calculations. (Tsiper EV and Soos ZG (2001) *Physical Review B* 64: 195124.) The difference of 0.89 eV between the polarization energies for surface molecules and for molecules inside the bulk can be estimated from the difference of photoemission spectroscopy (PES) and inverse photoemission spectroscopy (IPES). (Hill IG, Kahn A, Cornil J, dos Santos DA, and Bredas JL (2000) Occupied and unoccupied electronic levels in organic π-conjugated molecules: comparison between experiment and theory. *Chemical Physics Letters* 317: 444 and Park S, Kampen TU, Braun W, and Zahn DRT (2001) *Applied Surface Science* 175–176: 249.) The experimental uncertainty of the energetic difference between PES and IPES is derived from a fit to experimental data, and the uncertainty of the CT energy in the undeformed stack is estimated from the three ingredients of the excimer red shift.

the reorganization energy in the absorption spectra of the Frenkel excitons. The internal deformation of the molecules in the stacked excimer geometry gives a Stokes shift half as large as for the monomer, so that it is reasonable to assume that the contribution of the librational phonons is reduced by the same amount. The self-trapping of the excimer along the stack distance can be investigated with a combination of configuration interaction of singles and the second-order Møller–Plesset theory, resulting in a Stokes shift of 0.09 eV. These red shifts with respect to the CT transition between stack neighbors in the geometry of the crystal ground state are summarized in **Figure 16**. The resulting estimate of 2.06 eV for the CT transition in the crystal geometry is ~ 0.25 eV below previous calculated values, and this deviation can be rationalized from the fact that the gas-to-crystal shift of the CT transition is in the same range as for the Frenkel exciton, a contribution to the CT energy not included in previous microscopic models.

The CT gap can be defined as the energy of a pair of oppositely charged molecules at infinite distance inside the crystal. It arises from polarization corrections of the energetic difference between ionization potential I and electron affinity A in vacuum:

$$E_t = I - A - P_+ - P_- \qquad [9]$$

where P_+ and P_- contain polaronic contributions due to intramolecular deformations and lattice phonons, and the energy related to the polarization induced by the charges on the ionized molecules. At finite distance, the electron–hole attraction reduces this energy,

$$E_{\text{pair}}(\boldsymbol{r}) = E_t + V(\boldsymbol{r}) < E_t \qquad [10]$$

and for stack neighbours, micro-electrostatic calculations gave $V(\boldsymbol{r}) = -1.06$ eV, resulting in an estimate of 3.12 eV for the transport gap inside the crystal. Due to smaller polarization energies for the surface molecules involved in measured photoemission and inverse photoemission spectra, the difference between the resulting HOMO and LUMO states of the surface molecules is much higher than the transport gap, (see **Figure 16**).

Electronic Band Structure and Mobility of Charge Carriers

Tight-Binding Approach

The transfer of positive or negative charges between adjacent sites in the crystal results from the overlap between the electronic π orbitals of the respective molecules. The exponential decay of the wave functions results in rather small intermolecular matrix elements, and even for the neighboring molecules, they are only of the order of $T = 0.1$ eV. The electronic band structure for a molecular crystal with

two basis molecules can be written as

$$E_{\pm}(\boldsymbol{k}) = \left(\frac{T_{\mathrm{AA}}(\boldsymbol{k}) + T_{\mathrm{BB}}(\boldsymbol{k})}{2}\right) \pm \sqrt{\left(\frac{T_{\mathrm{AA}}(\boldsymbol{k}) - T_{\mathrm{BB}}(\boldsymbol{k})}{2}\right)^2 + V_{\mathrm{AB}}^2(\boldsymbol{k})} \quad [11]$$

where the wave-vector-dependent matrix elements are given by a Fourier transform of the intermolecular transfer integrals $t_{\mathrm{AA}}(\boldsymbol{R}_n)$, $t_{\mathrm{BB}}(\boldsymbol{R}_n)$, and $t_{\mathrm{AB}}(\boldsymbol{R}_n)$:

$$T_{\mathrm{AA}}(\boldsymbol{k}) = E_{\mathrm{A}} - 2\sum_{\boldsymbol{n}} t_{\mathrm{AA}}(\boldsymbol{R}_{\boldsymbol{n}})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_n} \quad [12]$$

$$T_{\mathrm{BB}}(\boldsymbol{k}) = E_{\mathrm{B}} - 2\sum_{\boldsymbol{n}} t_{\mathrm{BB}}(\boldsymbol{R}_{\boldsymbol{n}})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_n} \quad [13]$$

$$V_{\mathrm{AB}}(\boldsymbol{k}) = -2\sum_{\boldsymbol{n}} t_{\mathrm{AB}}(\boldsymbol{R}_{\boldsymbol{n}})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_n} \quad [14]$$

For crystals with the monoclinic centrosymmetric space group $P2_1/c(C_{2h}^5)$ containing two equivalent molecules such as naphthalene, anthracene, and PTCDA, $E_{\mathrm{A}} = E_{\mathrm{B}}$ and $T_{\mathrm{AA}}(\boldsymbol{k}) = T_{\mathrm{BB}}(\boldsymbol{k})$, whereas for triclinic crystals with two inequivalent basis molecules such as tetracene and pentacene, these energies and transfer matrix elements differ. The values of the transfer matrix elements such as $t_{\mathrm{AA}}(\boldsymbol{R}_n)$ can be obtained with quantum chemical approaches or with density-functional calculations of the periodic solid. **Figure 17** shows a calculated electronic band structure for naphthalene. For the polyacene series, the total bandwidths of the HOMO and LUMO bands increase monotonously from naphthalene to pentacene.

Band Narrowing due to Polaronic Effects

Within a mixed Holstein–Peierls approach, the coupling between tightly bound electrons and phonons can be written as

$$H = \sum_{mn} \varepsilon_{mn} a_m^+ a_n + \sum_q \hbar\omega_q \left(b_q^+ b_q + \frac{1}{2}\right) + \sum_{mnq} \hbar\omega_q g_{qmn}\left(b_q^+ + b_{-q}\right) a_m^+ a_n \quad [15]$$

where $a_m^{(+)}$ annihilates (creates) an electron at the site $\boldsymbol{R}_m$ with energy ε_{mm}, and $b_q^{(+)}$ annihilates (creates) a phonon with wave vector $\boldsymbol{q}$ and energy $\hbar\omega_q$. The site indices m are a shorthand notation including both the cell index and the basis molecule. In order to keep the notation readable, a sum over different local and nonlocal (librational) phonon modes λ is suppressed.

In the noninteracting system, that is, $g_{qmn} = 0$, the electronic band structure would be determined by the transfer matrix elements ε_{mn} alone. An approximate solution of the coupled Hamiltonian [15] can be achieved by a nonlocal canonical transformation according to

$$H \rightarrow \tilde{H} = e^S A e^{S+} \quad [16]$$

$$S = \sum_{mnq} g_{qmn}(b_q^+ - b_{-q}) a_m^+ a_n \quad [17]$$

By an application of the Baker–Campbell–Hausdorff theorem and the Wick theorem, the transformed Hamiltonian decouples into independent electron and phonon contributions,

$$\tilde{H} = \sum_{mn} \tilde{E}_{mn} a_m^+ a_n + \sum_q \hbar\omega_q \left(b_q^+ b_q + \frac{1}{2}\right) \quad [18]$$

where the on-site energies $\tilde{E}_{mm}$ and transfer integrals $\tilde{E}_{mn}$ are renormalized by the presence of local and nonlocal phonon couplings:

$$\langle \tilde{E}_{mm} \rangle = \varepsilon_{mm} - \Delta_{mm} \quad [19]$$

$$\langle \tilde{E}_{mn} \rangle = \varepsilon_{mn}\mathrm{e}^{-X_{mn}} - \Delta_{mn}\frac{1 - \mathrm{e}^{-X_{mn}}}{X_{mn}} \quad [20]$$

The polaron shifts Δ_{mn} are given by

$$\Delta_{mn} = \sum_{qj} \hbar\omega_q g_{qmj} g_{-qjn} \quad [21]$$

and X_{mn} stands for

$$X_{mn} = \sum_q \left(\frac{1}{2} + N_q\right) G_{qmn} \quad [22]$$

where N_q is the thermal occupation of the phonon mode according to Bose–Einstein statistics. G_{qmn} summarizes both local and nonlocal phonon couplings:

$$G_{qmn} = |g_{qmm} - g_{qnn}|^2 + \sum_{l \neq m,n} (|g_{qml}|^2 + |g_{qnl}|^2) \quad [23]$$

For practical calculations, the phonons can be approximated by dispersionless branches λ, simplifying the above expression [22] to

$$X_{mn} = \sum_{\lambda} \left(\frac{1}{2} + N_\lambda\right)(G_{\lambda mm} + G_{\lambda nn} - g_{\lambda mn}^2) \quad [24]$$

$$G_{\lambda jj} = g_{\lambda jj}^2 + \frac{1}{2}\sum_{l \neq j} g_{\lambda jl}^2 \quad [25]$$

For the polyacene series, this formalism has been applied to the temperature dependence of the polaron bandwidth (see **Figure 18** for naphthalene and tetracene). At zero temperature, the local part of eqns [24] and [25] gives a reduction of the transfer matrix

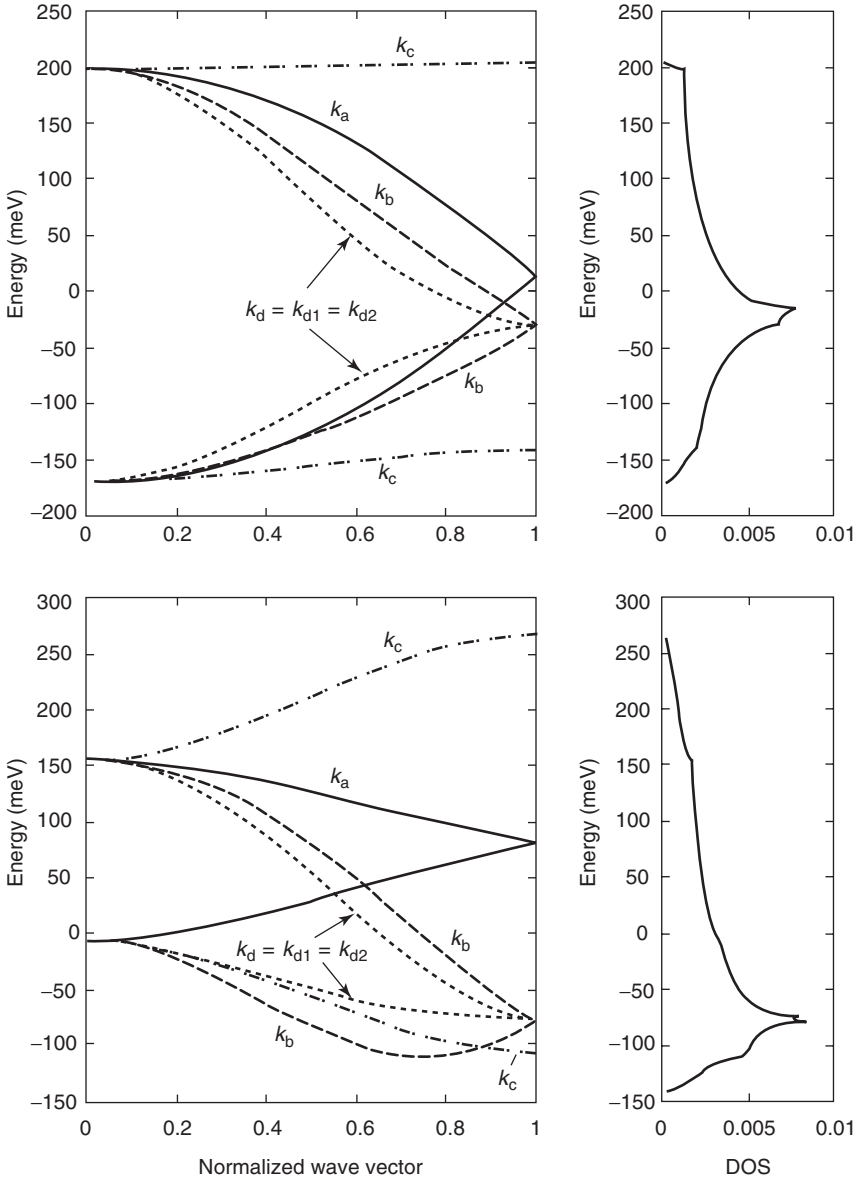


Figure 17 Shape of the LUMO band (upper panel) and the HOMO band (lower panel) of naphthalene in the major crystal directions within the first Brillouin zone, with the corresponding density of states (DOS) (right panels). (After Cheng YC, Silbey RJ, da Silva Filho DA, Calbert JP, Cornil J, and Bredas JL (2003) *Journal of Chemical Physics* 118: 3764.)

elements in eqn [20] resembling the exponential prefactor $S_{00}^2 = e^{-g^2}$ in the Poisson distribution [2]. In this regime, only this lowest vibronic level is occupied by charge carriers, reducing the bandwidth of the lowest vibronic branch by S_{00}^2. Therefore, already at low temperature, the polaron bandwidths reported in **Figure 18** are far below the purely electronic bandwidth shown in **Figure 17**. Due to the low energies of the phonon modes coupling to the electron transfer, at higher temperature the thermal occupation numbers N_λ may become rather large, resulting in a further reduction of the polaron bandwidth with respect to low temperature.

Temperature Dependence of the Charge Carrier Mobilities

Figure 19 reports the measured mobility of electrons and holes in naphthalene. In each crystal direction, the holes show higher mobilities than electrons, but the temperature dependence is quite different: for holes, the exponent n in $\mu_h(T) \propto T^{-n}$ increases with rising temperature, whereas for electrons it decreases.

In thermal equilibrium, the mobility μ_α in the direction $\boldsymbol{e}_\alpha$ can be obtained from the Kubo formula based on the current–current correlation function. The current is composed of a pure electronic part $\boldsymbol{j}^{(I)}$

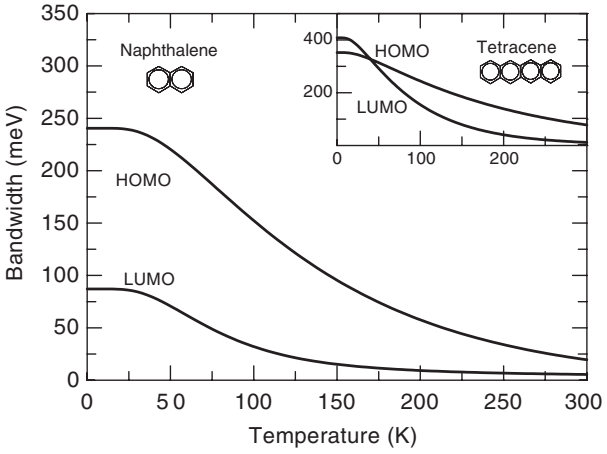


Figure 18 Effective HOMO and LUMO polaron bandwidths vs. temperature for naphthalene and tetracene. (After Hannewald K, Stojanovic VM, and Bobbert PA (2004) *Journal of Physics Condensed Matter* 16: 2023.)

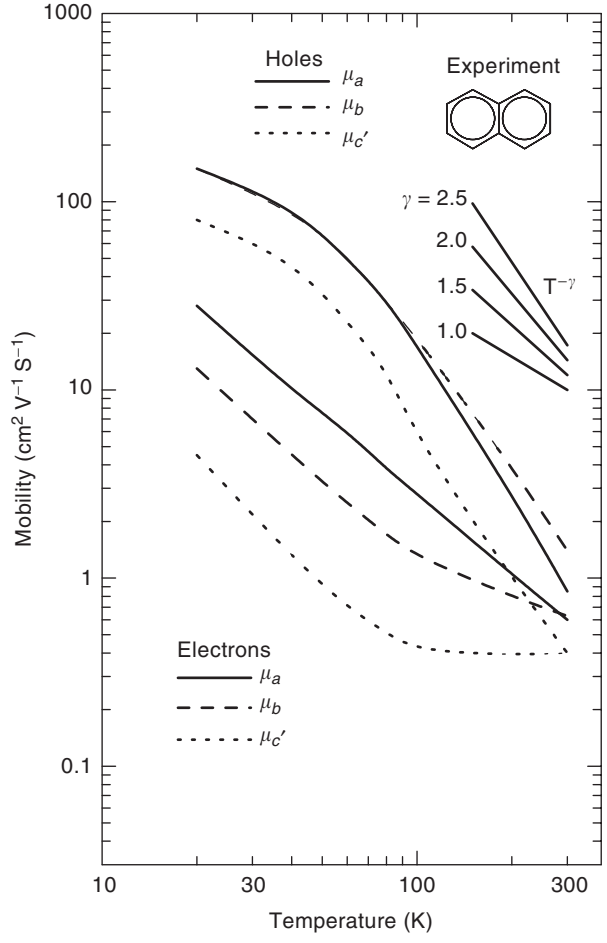


Figure 19 Measured charge carrier mobilities vs. temperature for naphthalene crystals. (After Karl (1985); Hannewald K and Bobbert PA (2004) *Applied Physics Letters* 85: 1535.)

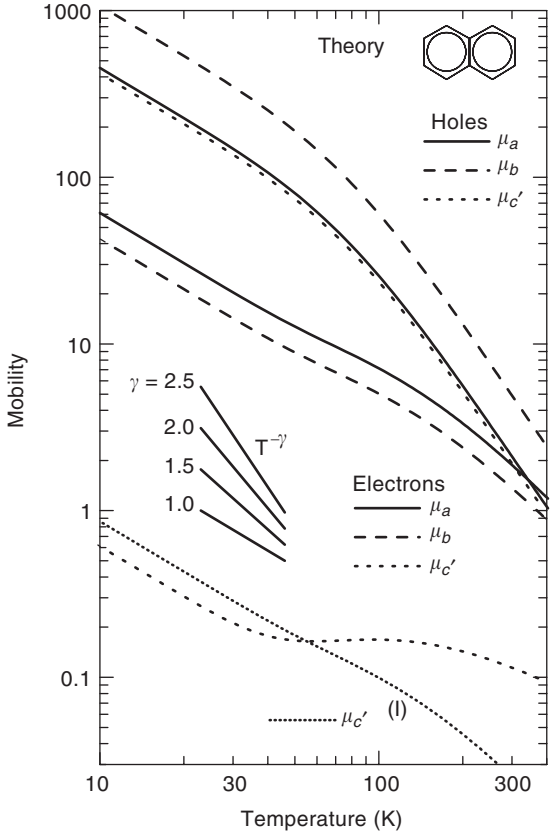


Figure 20 Calculated charge carrier mobilities vs. temperature for naphthalene crystals, as obtained from eqns [28] and [30] in arbitrary units. (After Hannewald K and Bobbert PA (2004) *Ab initio* theory of charge-carrier conduction in ultrapure organic crystals. *Applied Physics Letters* 85: 1535.)

and a phonon-assisted part $\boldsymbol{j}^{(II)}$, given by

$$\boldsymbol{j}^{(I)} = \frac{e}{i\hbar}\sum_{mn}(\boldsymbol{R}_m - \boldsymbol{R}_n)\varepsilon_{mn}a_m^+ a_n \qquad [26]$$

$$\boldsymbol{j}^{(II)} = \frac{e}{i\hbar}\sum_{mnq}(\boldsymbol{R}_m - \boldsymbol{R}_n)\hbar\omega_q g_{qmn}(b_q^+ + b_{-q})a_m^+ a_n \qquad [27]$$

The mobility $\mu_\alpha^{(I)}$ can be cast into the following form:

$$\mu_\alpha^{(I)} = \frac{e}{2k_B T\hbar^2}\sum_{m\neq n}(R_{m\alpha} - R_{n\alpha})^2\varepsilon_{mn}^2 \times \int_{-\infty}^{\infty} dt \left[e^{-\sum_\lambda G_\lambda[1+2N_\lambda-\phi_\lambda(t)]}\right]^2 e^{-\Gamma^2 t^2} \qquad [28]$$

where $R_{m\alpha}$ is the Cartesian component α of the molecular site, G_λ is defined as in eqn [25], Γ is a phenomenological broadening parameter, and

$$\phi_\lambda(t) = (1+N_\lambda)e^{-i\omega_\lambda t} + N_\lambda e^{i\omega_\lambda t} \qquad [29]$$

The bandwidth narrowing of eqn [20] is represented by the exponential of $-\sum_\lambda G_\lambda[1+2N_\lambda]$ in eqn [28]. It accounts for coherent scattering processes, whereas the part of the exponent containing $\phi_\lambda(t)$ describes

incoherent scattering processes involving changes in the phonon occupation.

The total mobility μ_α including the phonon-assisted current $\boldsymbol{j}^{(II)}$ can be obtained from an expression of the same form as eqn [28], with the replacement

$$\varepsilon_{mn}^2 \rightarrow (\varepsilon_{mn} - \Delta_{mn})^2 + \frac{1}{2}\sum_\lambda (\hbar\omega_\lambda g_{\lambda mn})^2 \phi_\lambda(t) \qquad [30]$$

In **Figure 20**, the calculated electron and hole mobilities are reported as a function of temperature. The experimental observation that the hole mobilities are larger than the electron mobilities is reproduced, together with a qualitatively reasonable temperature dependence of the exponent n in $\mu_h(T) \propto T^{-n}$. For temperatures above $T = 50\,\mathrm{K}$, the phonon-assisted electron current $\boldsymbol{j}^{(II)}$ along c' increases, resulting in a saturation of the mobility at higher temperatures, as observed experimentally.

However, for electrons the anisotropy of the mobility is severely overestimated. As the electronic transfer matrix elements and the electron–phonon couplings were determined with density-functional methods, this deficiency can be related to the lack of van der Waals corrections in the definition of the total energy of the molecular crystal.

Summary

From the experimental point of view, the field of organic semiconductors is now approaching a mature state, and the first technological applications including organic light emitting devices and organic field effect transistors are already emerging in the market. However, due to the size of the molecules involved, quantitatively reliable model calculations concerning spectroscopic properties and charge carrier mobilities are just emerging. For practical applications, it remains a major challenge to suppress unwanted impurities in the organic host which can act as traps for charge carriers, and to introduce dopants in a controlled way. Concerning optoelectronic applications, the possibility of using electrophosphorescent molecules has allowed the design of light emitting devices with highly competitive efficiencies with respect to inorganic semiconductors and gas discharge tubes.

See also: Excitons in Crystals; Polarons; Polymers and organic Materials, Electronic States of; Quantum Mechanics: Molecules.

PACS: 78.40.Me; 71.35.Aa; 78.55.Kz; 78.66.Qn; 78.30.Jw; 78.20.Bh; 71.15.Qe; 31.70.Ks; 72.80.Le; 71.15.Nc; 78.55. – m

Further Reading

Agranovich VM and Bassani F (eds.) (2003) *Electronic Excitations in Organic Based Nanostructures*. Amsterdam: Elsevier.

Agranovich VM and La Rocca GC (eds.) (2002) Proc. of Int. School of Physics "E. Fermi," course CXLIX. *Organic Nanostructures: Science and Applications*. Amsterdam: IOS Press.

Birks JB (1970) *Photophysics of Aromatic Molecules*. London: Wiley.

Broude VL, Rashba EI, and Sheka EF (1985) *Spectroscopy of Molecular Excitons*. Berlin: Springer.

Craig DP and Walmsley SH (1968) *Excitons in Molecular Crystals*. New York: Benjamin.

Davydov AS (1971) *Theory of Molecular Excitons*. New York: Plenum Press.

Farchioni R and Grosso G (eds.) (2001) *Organic Electronic Materials: Conjugated Polymers and Low Molecular Weight Organic Solids*. Berlin: Springer.

Hoffmann M and Soos ZG (2002) *Physical Review B* 66: 024305.

Karl N (1985) In: Hellwege K-H and Madelung O (eds.) *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, vol. III 17 i, p. 106. Berlin: Springer.

Matsui A (1990) *Journal of the Optical Society of America B* 7: 1615.

Park S, Kampen TU, Braun W, and Zahn DRT (2001) *Applied Surface Science* 175–176: 249.

Pope M and Svenberg CE (1980) *Electronic Processes in Organic Crystals and Polymers*. New York: Oxford University Press.

Silinsh EA and Capek V (1997) *Organic Molecular Crystals*. Berlin: Springer.

Tsuzuki S, Honda K, Uchimaru T, and Mikami M (2004) *Journal of Chemical Physics* 120: 647.

Wewer M and Stienkemeier F (2004) *Journal of Chemical Physics* 120: 1239.

Orientation Texture

S I Wright, TSL/EDAX, Draper, UT, USA

Introduction

Many single crystals exhibit properties that vary with direction (anisotropy). For example, consider the elastic modulus of a copper single crystal. The elastic modulus measured for an applied uniaxial tensile stress parallel to the $\langle 1\,1\,1 \rangle$ crystal direction (i.e., the body diagonal) is nearly three times larger than the modulus measured with a stress axis parallel to $\langle 1\,0\,0 \rangle$ (i.e., normal to the cube faces) (**Figure 1**).

Most materials are not single crystals but polycrystals. To first order, a given property that varies

with direction, such as the elastic modulus, can be approximated for a polycrystal by averaging the property for each constituent crystal. For example, if the elastic stiffness of a single crystal is represented as a fourth-order tensor – C_{ijkl}, then one approximation of the average elastic stiffness tensor ($\bar{C}_{ijkl}$) for the polycrystal is given as follows:

$$\bar{C}_{ijkl} = \frac{1}{V}\sum_{n=1}^{N} v(\boldsymbol{g}^n) g^n_{ip} g^n_{jq} g^n_{kr} g^n_{ls} C_{pqrs} \qquad [1]$$

where V is the total volume of the polycrystal, $\boldsymbol{g}^n$ is the orientation, $v(\boldsymbol{g}^n)$ is the volume of the nth grain in the polycrystal in matrix form, and N is the total number of grains.

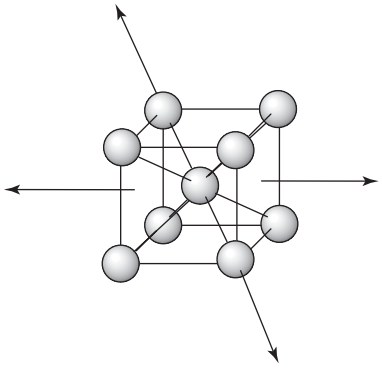


Figure 1 Schematic of elastic moduli measured for stresses applied along the ⟨1 1 1⟩ (body diagonal) and ⟨1 0 0⟩ (normal to the cube faces) crystal directions in a body-centered cubic crystal.

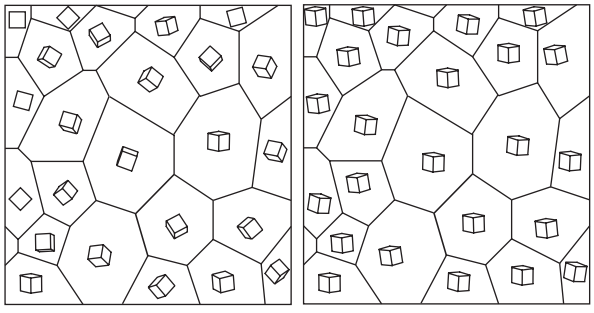


Figure 2 Schematic of a polycrystal with a random texture and a polycrystal with a strong texture.

If a polycrystal is composed of grains where the crystal lattices are all in similar orientations, then the properties of the polycrystal would mimic the inherent anisotropy of the constituent crystals. If the grains were all oriented randomly, then the properties of the polycrystal would not reflect the anisotropy of the constituent crystals, but would instead be isotropic (i.e., the properties would not vary with direction) (**Figure 2**).

Real polycrystals may vary between these two extremes depending on the processing route used to form the material. The terms crystallographic texture, orientation texture, or simply texture are used to describe the statistical distribution of grain orientations (the terms lattice preferred orientation (LPO) or fabric are also used within the geological community). In fact, the term grain (in most cases) generally describes a spatial domain where the orientation of the crystal lattice is the same or at least very nearly the same. A grain boundary is then an interface between two crystal lattices of different orientation.

As nearly all materials forming processes produce textured materials, the characterization of texture is an important element to understanding the relationship between microstructure and properties as well as the evolution of microstructure during processing. This article reviews the definition of crystallographic orientation and graphical representations of orientations, techniques for the measurement of texture, the mathematical formalism used in processing texture data, and a few examples of texture analysis.

Crystallographic Orientation

Orientation Descriptions

In order to describe a crystallographic orientation, two coordinate systems must be defined: one for the sample, referred to as $\boldsymbol{X}^{\mathrm{S}}$ and one for the crystallite, $\boldsymbol{X}^{\mathrm{C}}$ (**Figure 3**). The choice of sample axes is somewhat arbitrary; however, the choice usually follows

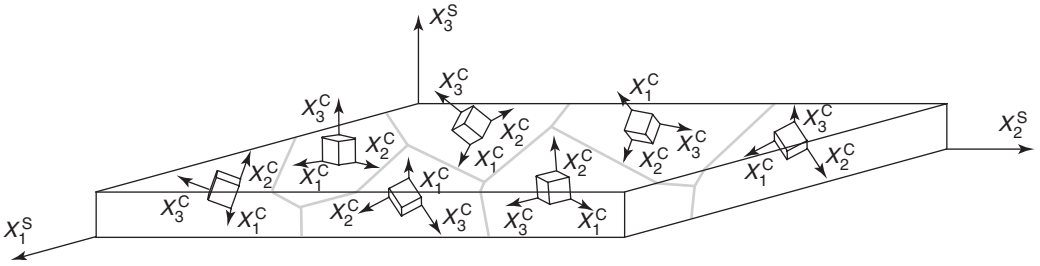


Figure 3 Schematic of the coordinate system fixed to the sample, $\boldsymbol{X}^{\mathrm{S}}$ and the coordinate system fixed to the crystals, $\boldsymbol{X}^{\mathrm{C}}$.

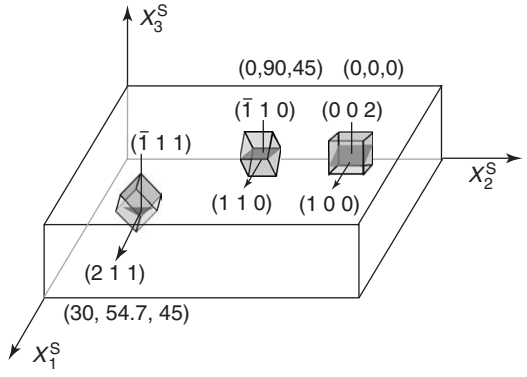


Figure 4 Two equivalent descriptions of crystallographic orientation: Euler angles and standard plane normal and crystal direction notation (i.e., (*hkl*)[*uvw*]).

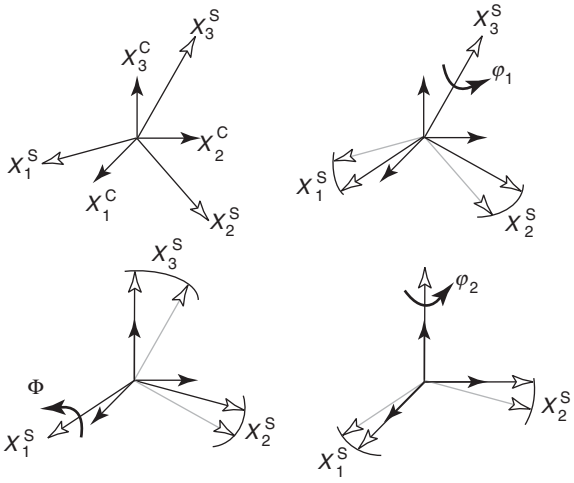


Figure 5 Schematic showing the definition of the three Euler angles, φ_1, Φ, φ_2.

the sample shape or processing used to form the sample. For example, consider a coupon obtained from a rolled sheet. It is common to assign X_1^S to the rolling direction, X_2^S to the transverse direction, and X_3^S to the sample normal direction. The coordinate system for the crystallite generally follows that traditionally used in defining the unit cell, that is, X_1^C is assigned to the [1 0 0] crystal direction, X_2^C to [0 1 0], and X_3^C to [0 0 1]. Both coordinate systems are right-handed Cartesian coordinate systems. While the assignment of the reference frame to a cubic crystal, as given here, is fairly straightforward, the assignment is less straightforward for other crystal systems such as hexagonal ones. Nonetheless, a Cartesian system can be assigned, for example, for a hexagonal crystal the convention is $X_1^C || [2\,\bar{1}\,\bar{1}\,0]$, $X_2^C || [0\,1\,\bar{1}\,0]$, and $X_3^C || [0\,0\,0\,1]$.

There are many ways to describe the orientation of a crystal with respect to the coordinate system of a sample. The most common is the (*hkl*)[*uvw*] notation where (*hkl*) are the Miller indices of the plane perpendicular to the sample normal (generally X_3^S) and [*uvw*] describes the crystal direction parallel to X_1^S (**Figure 4**). However, as integers are used, this description is not as convenient for describing any arbitrary orientation. In texture analysis, a common way to describe the orientation of the crystal lattice with respect to a sample reference frame is to use three Euler angles (**Figure 5**). In this case, the Euler angles describe three successive rotations that bring the sample reference frame into coincidence with the crystal reference frame.

The various descriptions are equivalent and can be converted from one description to another. A matrix, ***g***, can also be used to describe the orientation. It can be derived from the Euler angles according to

$$\begin{bmatrix} \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2\cos\Phi & \sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2\cos\Phi & \sin\varphi_2\sin\Phi \\ -\cos\varphi_1\sin\varphi_2 - \sin\varphi_1\cos\varphi_2\cos\Phi & -\sin\varphi_1\sin\varphi_2 + \cos\varphi_1\cos\varphi_2\cos\Phi & \cos\varphi_2\sin\Phi \\ \sin\varphi_1\sin\Phi & -\cos\varphi_1\sin\Phi & \cos\Phi \end{bmatrix} \quad [2]$$

The matrix representation is particularly useful for describing orientations in computer algorithms.

The rotation of the sample reference frame into coincidence with the crystal reference frame is termed a passive rotation. Conversely, rotating the crystal reference frame into coincidence with the sample reference frame is termed an active rotation.

Symmetry

The three Euler angles form a bounded space: $0 \leqslant \varphi_1 < 2\pi,\ 0 \leqslant \Phi < \pi,\ 0 \leqslant \varphi_2 < 2\pi$. As most crystals exhibit some degree of symmetry, for any given crystallographic orientation, there exist other symmetrically equivalent orientations. For example, consider a material with cubic crystal symmetry. There are 24 symmetry elements for a cube: the fourfold rotations about axes normal to the faces of the cube (i.e., 90° rotations about the $\langle 0\,0\,1 \rangle$ crystal directions), the twofold rotations about the face diagonals

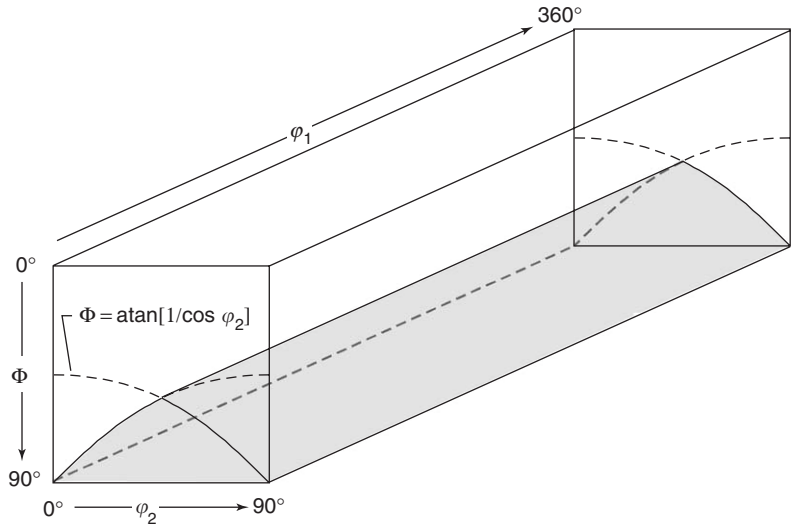


Figure 6 The asymmetric domain for cubic crystal symmetry.

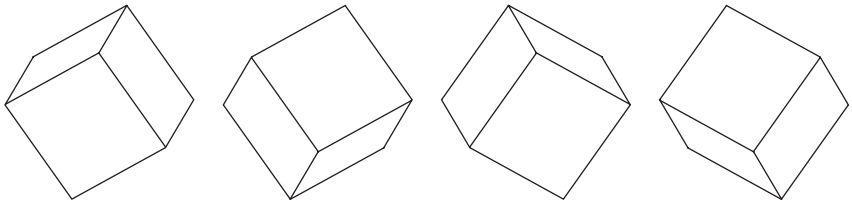


Figure 7 Four orthotropically equivalent orientations.

(180° rotations about $\langle 1\,1\,0 \rangle$), and the threefold rotations about the body diagonals (120° about $\langle 1\,1\,1 \rangle$). Thus, for any orientation of a cubic crystal, there are 24 equivalent orientations within the Euler space. A space can be defined within Euler space containing only asymmetric orientations. This subspace is termed the asymmetric domain or fundamental zone (**Figure 6**). Since texture is a statistical distribution of orientations, another type of symmetry must also be considered, namely, sample symmetry. Sample symmetry arises from the symmetry inherent to the forming process. Consider, for example, the rolling of sheet metals. This forming operation has orthotropic symmetry. For a rolled cubic material, the four orientations shown in **Figure 7** would appear with the same statistical frequency in the microstructure. Thus, the term statistical symmetry is a more accurate description. If the matrix notation is used, then the symmetry operations can be applied as follows:

$$\boldsymbol{g}^{\text{e}} = \boldsymbol{L}_i \boldsymbol{g} \boldsymbol{R}_j \qquad [3]$$

where $\boldsymbol{g}^{\text{e}}$ denotes an equivalent orientation, $\boldsymbol{L}_i$ is one of the crystal symmetry elements, and $\boldsymbol{R}_j$ is one of the statistical symmetry elements. Because of the form of this relation, the crystal symmetry is sometimes referred to as left-handed symmetry, and the statistical symmetry as right-handed.

Graphical Descriptions

Pole figures can be used to provide a graphical representation of orientation. A pole figure is a two-dimensional graphical representation of orientation, showing the orientation of a selected plane normal (a pole) with respect to the sample reference frame. The pole is a vector in three-dimensional space. If one imagines a hemisphere surrounding a crystal unit cell, then a pole would intersect the hemisphere at some location in three-dimensional space. The intersection location can be projected into a two-dimensional plot using a variety of projection schemes. An example of a pole figure is shown in **Figure 8**, this pole figure shows the $(0\,0\,0\,1)$ and $(1\,0\,\bar{1}\,0)$ poles for a hexagonal crystal, note that there is only one pole for the $(0\,0\,0\,1)$ plane and three for the $(1\,0\,\bar{1}\,0)$ plane. The additional poles in the $(1\,0\,\bar{1}\,0)$ case arise from crystal symmetry. In the pole figures, poles from the negative hemisphere are shown in gray.

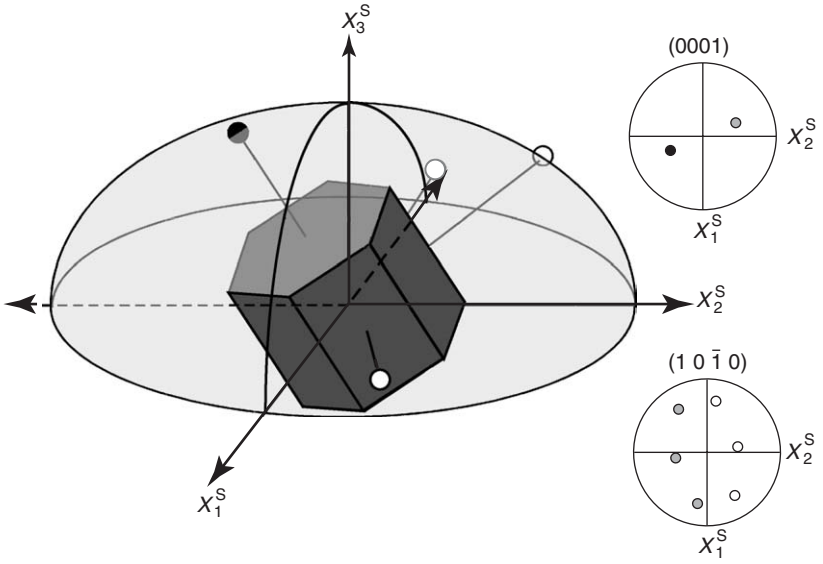


Figure 8 Schematic showing the orientation of a hexagonal crystal and the corresponding (0 0 0 1) pole figure.

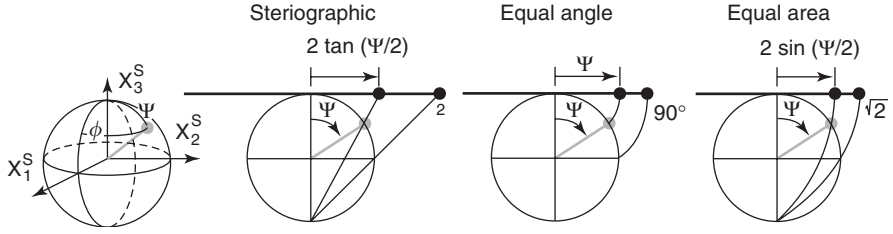


Figure 9 Stereographic, equal-angle, and equal-area pole figure projections.

Displaying the negative poles is sometimes referred to as including inversion symmetry. **Figure 9** shows three common "projection schemes" found in the literature (the equal angle is not a true projection; but, rather, a mapping of the polar angle).

A complement to the pole figure is the inverse pole figure. An inverse pole figure shows the location of a sample direction relative to the crystal reference system. For pole figures the plane of interest is selected, whereas in inverse pole figures the sample direction of interest is selected. It is a bit more intuitive to think of an inverse pole figure as showing which plane normal in the crystal is parallel to the selected sample direction. **Figure 10** shows an example. In this case, it shows the poles aligned with the sample normal for three different orientations as well as the poles aligned with the X_1^S sample direction. Because of crystal symmetry, the full circle need not be displayed; only the asymmetric portion is needed.

While pole figures and inverse pole figures are useful representations of orientations, they only provide a partial description of orientation. For example, consider the (0 0 0 1) pole figure shown in **Figure 8**. While the orientation of the (0 0 0 1) pole is known, the rotation of the crystal about this pole is unknown. A second pole figure (such as the (1 0 $\bar{1}$ 0)) is needed to completely fix the orientation. When the orientations of many crystallites are displayed, the use of multiple pole figures is quite limited. In this case, it is more productive to show the orientations in constant angle sections of the Euler space. **Figure 11** shows 30 000 discrete orientations from rolled copper in a portion of the Euler space along with the corresponding (1 1 1) pole figure.

Misorientation

To this point, the orientation of a crystal has been described relative to the sample coordinate system. However, the orientation of one crystal could also be described relative to the coordinate system associated with a neighboring crystal. Such a description is often denoted a misorientation (or alternatively, a disorientation or orientation difference). While misorientations can be described similar to orientations, that is, in terms of the Euler angles or matrices, it is

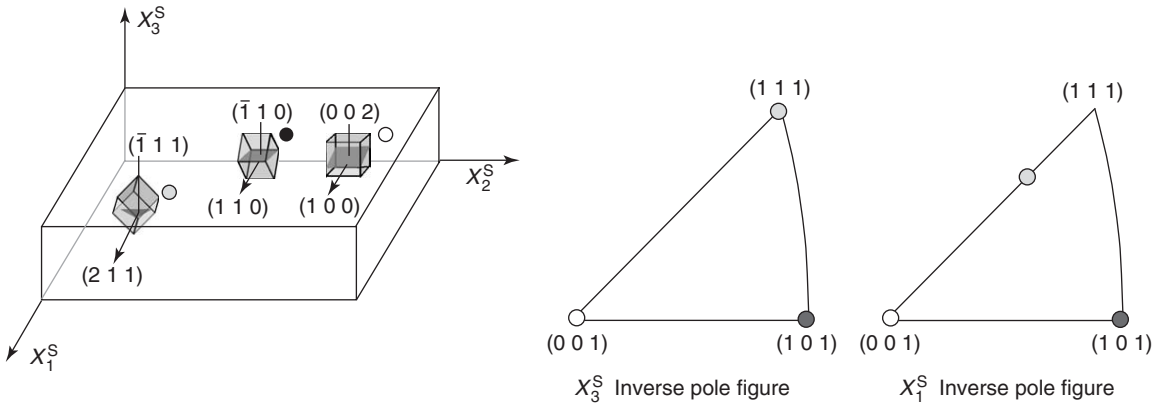


Figure 10 Schematic showing the representation of orientation in inverse pole figures.

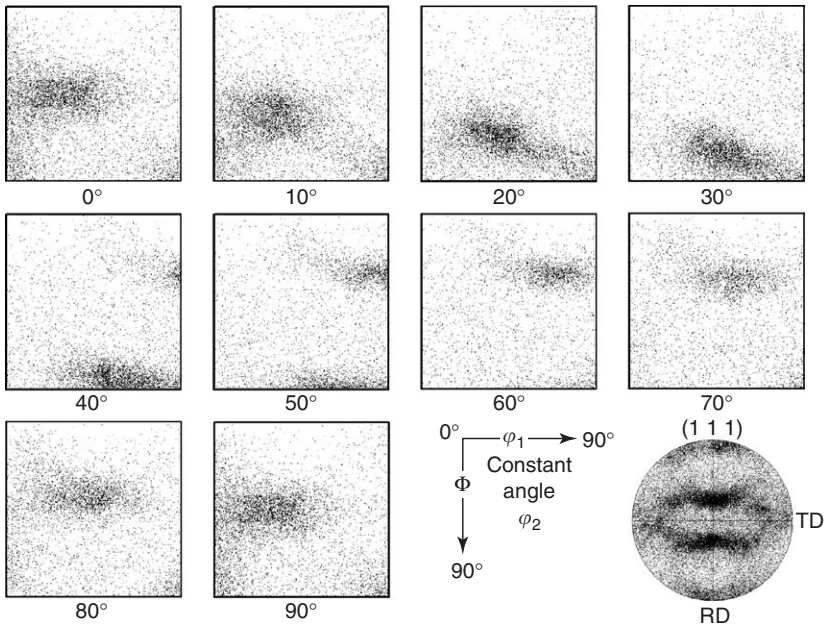


Figure 11 Plot of 30 000 discrete orientations from rolled copper in constant-angle sections through Euler space accompanied by a (1 1 1) pole figure of the same data.

usually more convenient to describe them in terms of axis–angle pairs. For two crystals in any two orientations, it is always possible to find a common axis around which one crystal can be rotated until it is in coincidence with the other. The misorientation can thus be described as a rotation about the common axis. As with orientations, crystal symmetry plays a role in the description of misorientations. For example, for a crystal with cubic symmetry there are 24 symmetrically equivalent orientations. Thus for a misorientation between two cubic crystals, there are 576 (24×24) symmetrically equivalent misorientations. This is realized in the axis–angle space as essentially 24 axes of rotation, each with a specific angle of rotation associated with it and 24 different permutations of the indices describing the axes. Generally, the angle of a permutation with minimum misorientation is used to describe the misorientation.

Texture Measurement Techniques

X-Ray Diffraction

There are essentially two ways of measuring textures, measuring pole figures or measuring the orientations of individual grains. Until recently, most textures have been measured using pole figures obtained by X-ray or neutron diffraction. The governing equation for a pole figure is Bragg's law:

$$\lambda = 2d_{hkl} \sin \theta \qquad [4]$$

where λ is the wavelength of the X-rays, d_{hkl} is the interplanar spacing of the pole figure to be measured (recall that the pole is the normal to a (hkl) plane), and θ is the Bragg angle as shown in **Figure 12**. In a pole figure measurement, a sample is placed in a goniometer so that the sample can be rotated while under the X-ray beam. For each rotation (in both Ψ and ϕ as shown in **Figure 12**), the signal from those grains satisfying Bragg's law is measured as shown in **Figure 13**. An intensity value is measured at successive tilts (Ψ) and rotations (ϕ) of the sample. This is typically done for a series of pole figures. For example, for a base-centered cubic material, the (1 1 0), (2 0 0), and (1 1 2) pole figures are typically measured. A similar approach is employed for neutron diffraction. The major difference is that X-ray pole figures are typically measured in a reflective mode, whereas neutron diffraction is done in transmission. Thus, neutron measurements have the advantage of providing a volumetric measure of texture, whereas the signal for X-rays comes from the surface of the sample. Various corrections must be applied to the pole figures to account for defocusing and absorption effects.

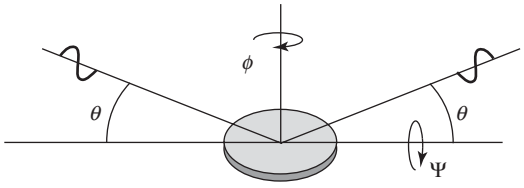


Figure 12 Diagram showing the pole figure measurement geometry.

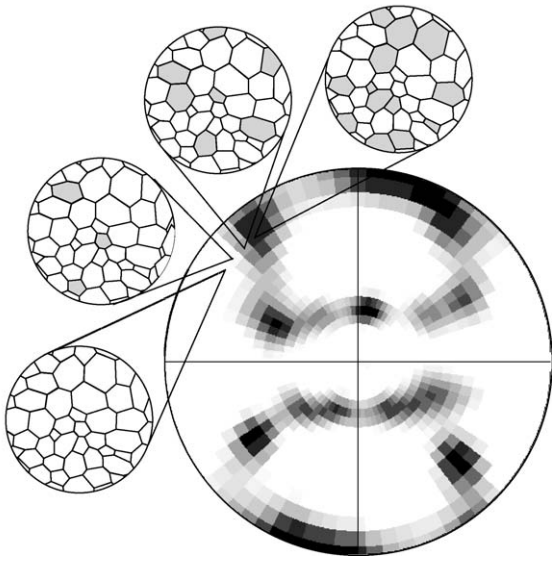


Figure 13 Schematic showing grains satisfying the Bragg condition for successive 5° rotations about the sample normal (ϕ).

Electron Backscatter Diffraction

In the last decade, the automation of electron backscatter diffraction (EBSD) in the scanning electron microscope (SEM) has made it practical to measure textures using the orientations of individual grains. Other methods for measuring individual orientations exist, including using X-rays from synchrotron radiation, optical microscopy using etch pits, or selected area or Kikuchi diffraction patterns in the transmission electron microscope. However, EBSD has become the most practical method for texture measurements via individual orientations.

In EBSD, a highly tilted sample ($\sim 70^\circ$) is placed in the SEM. When a stationary electron beam impinges on a grain in the sample, a diffraction pattern is generated as shown in **Figure 14**. The pattern is a function of the orientation of the lattice within the diffracting volume. As the beam is moved from grain to grain, the diffraction pattern changes.

The pattern is imaged on a phosphor screen and transmitted to a computer using a high-gain camera, typically a CCD camera. The orientation of the lattice in the diffracting volume can be derived from the pattern. This can be done automatically in the computer. The computer also controls the location of the beam on the sample. Typically, the beam is stepped across the sample to form a regular scan grid. An orientation measurement is made at each point in the grid (**Figure 15**). This process is fully automated and current rates exceed 50 measurement points per second. Despite the rapid measurement provided by automation, EBSD does not generally sample as

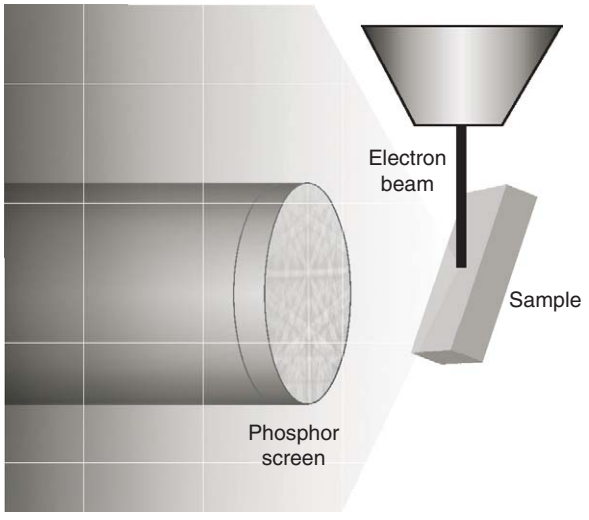


Figure 14 Schematic of EBSD.

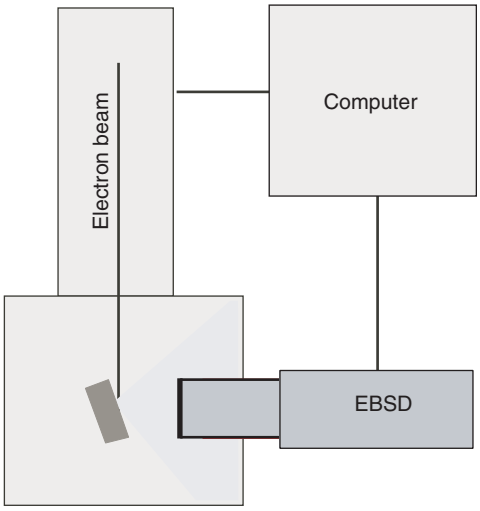


Figure 15 Schematic of automated EBSD.

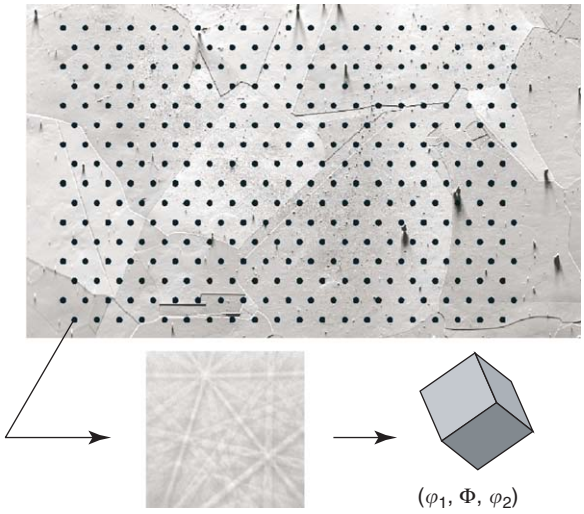


Figure 16 Schematic of the determination of the ODF from individual orientation measurements.

many grains as the conventional X-ray pole figure technique. However, EBSD has the advantage of spatially specific measurements of orientations allowing not only the distribution of orientations to be characterized, but misorientations and local texture gradients as well.

Mathematical Constructs

Basic Equation

The basis of texture analysis is the orientation distribution function or ODF. The ODF is a probability density function describing the probability of finding a grain with an orientation, g, within a given angular distance, Δg, of a specified orientation, g_o, in a polycrystal; or, alternatively, the volume fraction of material oriented within Δg of g_o. These two descriptions are related through the following equation:

$$\frac{\Delta V_{(g_o+\Delta g)}}{V} = \oint_{g\in(g_o+\Delta g)} f(g)\,\mathrm{d}g \qquad [5]$$

where the quantity to the left of the equals sign is the volume fraction of material oriented within Δg of g_o, and the function $f(g)$ is the ODF. It should be noted that the following condition must be met for a density function:

$$\oint f(g)\,\mathrm{d}g = 1 \qquad [6]$$

where the integration is carried out through all of the orientation space. The ODF gives the probability of finding a grain in a polycrystal with a specific orientation. If, for example, $f(g)$ is 3, then the chance of finding a grain with orientation, g, would be three times greater than that expected in a polycrystal with a random distribution of orientations.

The objective of texture analysis is to process the pole figures or the individual orientation measurements to determine the ODF. It is easiest to conceptualize the ODF determination using the individual orientation measurements. Consider a set of orientation measurements, where the orientations can be described as points on a number line as shown schematically in **Figure 16**. If the line is partitioned into a finite set of bins, then the number of orientations in each box can be counted and a histogram constructed as shown. A curve can then be fit to the histogram. The resulting curve is essentially the ODF.

It should be noted that the distribution of orientations shown schematically in **Figure 16** could be described using other distributions as shown in **Figure 17** depending on the size of the bins used to form the histogram. The choice is somewhat arbitrary; however, the choice can depend on the property being studied. For example, elastic anisotropy could be studied using the lower-order curve, whereas piezoelectricity requires a higher-order description of texture.

While there are multiple methods of performing the data reduction from the pole figures or discrete orientation data, two have emerged as the most common: the so-called harmonic and the WIMV methods. The WIMV (an acronym formed from the names of the originators of the technique) is essentially a binning method as outlined previously, whereas the harmonic method uses a series expansion approach, converting the orientation measurements directly to an ODF.

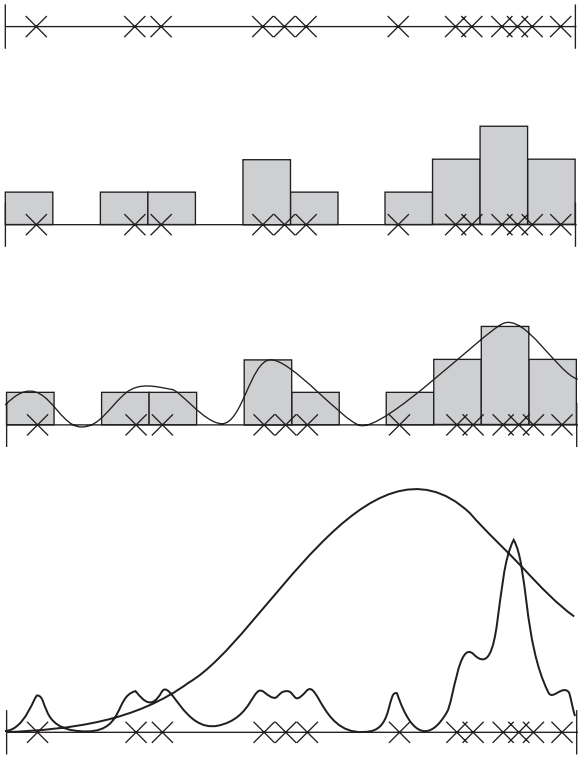


Figure 17 Schematic showing lower- and higher-order descriptions of the texture shown in **Figure 16**.

WIMV Method

A pole figure can be derived through integration of the ODF along a specific path in the orientation space according to

$$P_h(\boldsymbol{y}) = \frac{1}{2\pi}\int_{h\|y} f(g)\ \mathrm{d}\chi \qquad [7]$$

where $P_h(\boldsymbol{y})$ is the pole figure, $\boldsymbol{h}$ represents the (hkl) of the pole figure (a vector in the crystal reference frame), $\boldsymbol{y}$ is a vector in the pole figure (i.e., a vector in the sample reference frame represented by the rotation and tilt angles, ϕ and Ψ introduced previously). The integration is carried out along the path, χ, where $\boldsymbol{h}$ and $\boldsymbol{y}$ are coincident.

Pole figures are measured on discrete grids (typically 5° steps in ϕ and Ψ). If the ODF is also discretized into an array (e.g., in 5° steps in φ_1, Φ, and φ_2), then a matrix can be determined using an iterative process relating the intensities at the grid points in the measured pole figures with the ODF through eqn [6]. The mapping of the pole figure measurement grid to the ODF array is not one-to-one; thus, some interpolation techniques must be employed. One nice feature of this technique is that the resulting ODF is "smoothed" to the same degree as the measured pole figures. Thus a choice of the degree of smoothing to be applied is not required as in the other methods. The smoothness of the pole figure measurements is a function of the grid selection as well as the experimental setup (e.g., slit sizes).

Analogous to the WIMV pole figure inversion procedure, the ODF can be calculated from single orientation measurements using a binning approach. The Euler space is divided up into bins (e.g., $5^\circ \times 5^\circ \times 5^\circ$). For each orientation in the set, the corresponding bin is incremented. In addition, the symmetrically equivalent bins must also be incremented. Once all orientations have been accounted for, the values in the bins must be modified so that the normalization condition in eqn [5] is satisfied. Generally, the ODF is smoothed using a Gaussian filter. One of the challenges of this technique is the ambiguity of the Euler space at $\Phi = 0^\circ$. At $\Phi = 0$, φ_1 and φ_2 are not fully defined but rather governed by $\varphi_1 + \varphi_2 = C$. The volume of the bins vary as a function of Φ, bins centered on $\Phi = 2.5^\circ$ have a volume of $0.044\,\mathrm{rad}^3$, and bins centered on $\Phi = 87.5^\circ$ have a volume of $0.999\,\mathrm{rad}^3$ assuming they are $5^\circ \times 5^\circ \times 5^\circ$ bins. In addition, artifacts can arise simply due to the artificial boundaries introduced in the ODF by the binning technique. Lower- and higher-order descriptions of the ODF can be achieved through smoothing or by adjusting the size of the bins.

Harmonic Method

The ODF can be described by a series of generalized spherical harmonics (similar to the Fourier series expansion method used in many areas of physics). There are various ways of performing the series expansion, the most common method is that introduced by Bunge, namely,

$$f(g) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{n=-l}^{l} C_l^{mn} T_l^{mn}(g) \qquad [8]$$

where the $T_l^{mn}(g)$ are the spherical harmonic functions and are known. The C_l^{mn} coefficients are a function of the texture, and can be calculated from pole figures or individual orientation measurements. If the crystal and statistical symmetries are enforced on the calculation, then the number of C_l^{mn} coefficients is greatly reduced. For an expansion carried out to $l = 22$, a sample with cubic crystal symmetry and orthotropic sample symmetry would have 179 coefficients compared to 16 215 coefficients if the symmetry is not applied.

Pole figures can be calculated from the C coefficients using a series of spherical surface harmonics

according to

$$P_{h_i}(\boldsymbol{y}) = \sum_{l=0}^{\infty} \frac{4\pi}{2l+1} \sum_{m=-l}^{l} \sum_{n=-l}^{l} C_l^{mn} k_l^{*m}(\boldsymbol{h}_i) k_l^n(\boldsymbol{y}) \quad [9]$$

where $\boldsymbol{h}_i$ represents the (hkl) of the ith pole figure (a vector in the crystal reference frame), $\boldsymbol{y}$ is a vector in the pole figure (i.e., a vector in the sample reference frame represented by the rotation and tilt angles, ϕ and Ψ introduced previously), and the k functions are spherical surface harmonics. Since the surface harmonics and the pole figures are known quantities, the C coefficients can be obtained by finding the solution to a large number of linear equations.

The calculation of the coefficients using single orientation measurements is somewhat simpler, namely,

$$C_l^{mn} = \frac{2l+1}{N} \sum_{i=1}^{N} K T_l^{mn}(g_i) \quad [10]$$

where g_i represents one orientation from a set containing N individual orientation measurements. K is a smoothing factor. If single orientation measurements are made where only one measurement per grain is recorded, then a weighting factor proportional to the grain area must also be introduced.

One disadvantage of the harmonic method is the presence of negative values in the ODF. Negative values are not allowed for density functions but can arise from the series expansion, particularly when the textures are very sharp. As with any series expansion, truncation errors arise. However, the rank (l) of the series expansion has some correlation to physical properties. For example, the elastic constants of a crystal can be described using a fourth-order tensor. Thus, for characterizing elastic anisotropy arising from texture, the harmonic series need only be expanded out to $l = 4$.

The choice of the degree of the Gaussian smoothing is somewhat arbitrary. Generally, it is assumed that a sharper texture requires a smaller filter (say, 1–2°), whereas larger values (say 5–10°) are suggested for more random distributions. In addition, the greater the orientation measured, the smaller the Gaussian can be as well. Some comparisons between X-ray data and EBSD data suggest a value of 7.5°. However, with the increased speed in modern EBSD systems, larger data sets are being measured and therefore smaller values of 3–5° are now more appropriate.

Examples

Thin Films

Thin films are typically characterized by the alignment of a particular crystallographic axis with the sample normal. **Figure 18** shows a (1 1 1) pole figure for an aluminum thin film. The axisymmetric character of this material is sometimes called a (1 1 1) fiber texture. Some thin film properties (such as, the mean time to failure in microelectronic devices) have been correlated to the sharpness of the fiber texture. An axisymmetric texture can be more simply analyzed by a pole plot as shown in **Figure 18**. The pole plot is essentially a slice through the pole figure as shown. For more accurate results, the pole plot can be improved by averaging azimuthally.

While the pole plot is helpful for analyzing thin film textures, calculating the pole figures to verify that the texture is truly axisymmetric is always recommended. For example, in a copper thin film, a

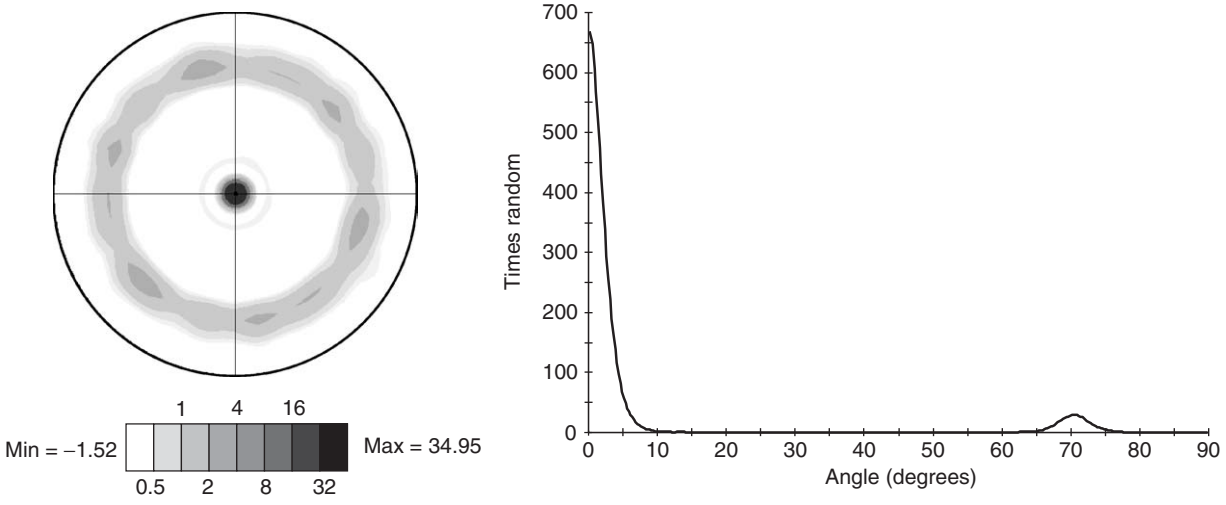


Figure 18 (1 1 1) pole figure and corresponding pole plot from an aluminum thin film.

(1 1 1) fiber texture is expected. However, other fibers may also be present. The axisymmetric nature of the texture can be confirmed using a pole figure, and the presence of other fibers discovered using an inverse pole figure as shown in **Figure 19**.

Rolled Sheet

One of the most studied materials using texture analysis is the rolled aluminum sheet for beverage cans. The texture of this material must be as well controlled as possible in order to produce stock well suited for the deep drawing of cans. Typical (1 1 1) pole figures for rolled aluminum are shown in **Figure 20**.

The bulk of the orientations typically lie within a "fiber" in the Euler space as shown in **Figure 21**. In order to simplify the analysis of rolled face-centered cubic materials such as aluminum, components associated with key locations along this fiber have been identified. These key components are shown in the Euler space and in a (1 1 1) pole figure.

A rolling texture can then be simply characterized by the intensities of these ideal components. This approach has been successfully used in developing thermomechanical processes to achieve textures that optimize the deep-drawing performance of beverage can stock.

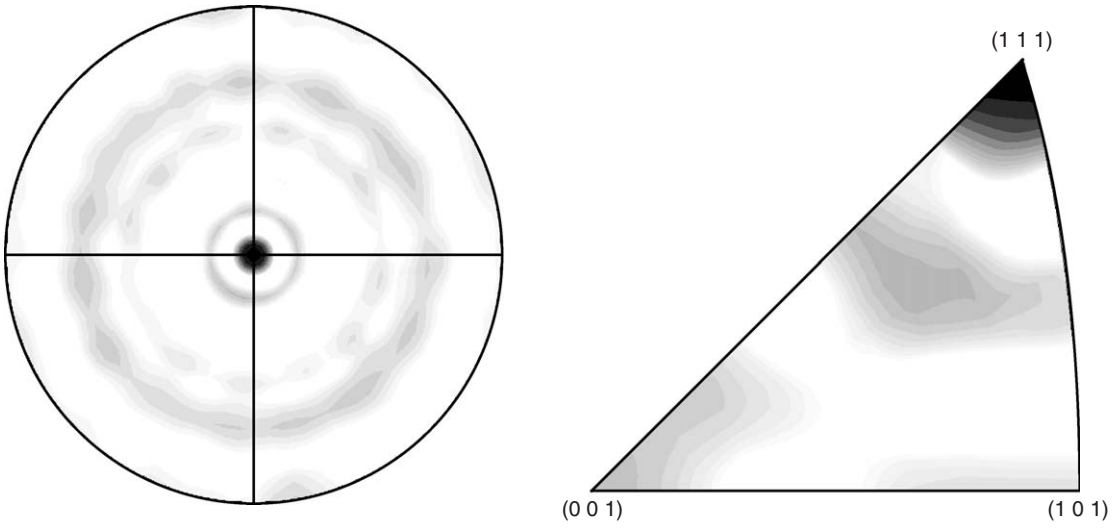


Figure 19 A (1 1 1) pole figure and an inverse pole figure for a copper thin film.

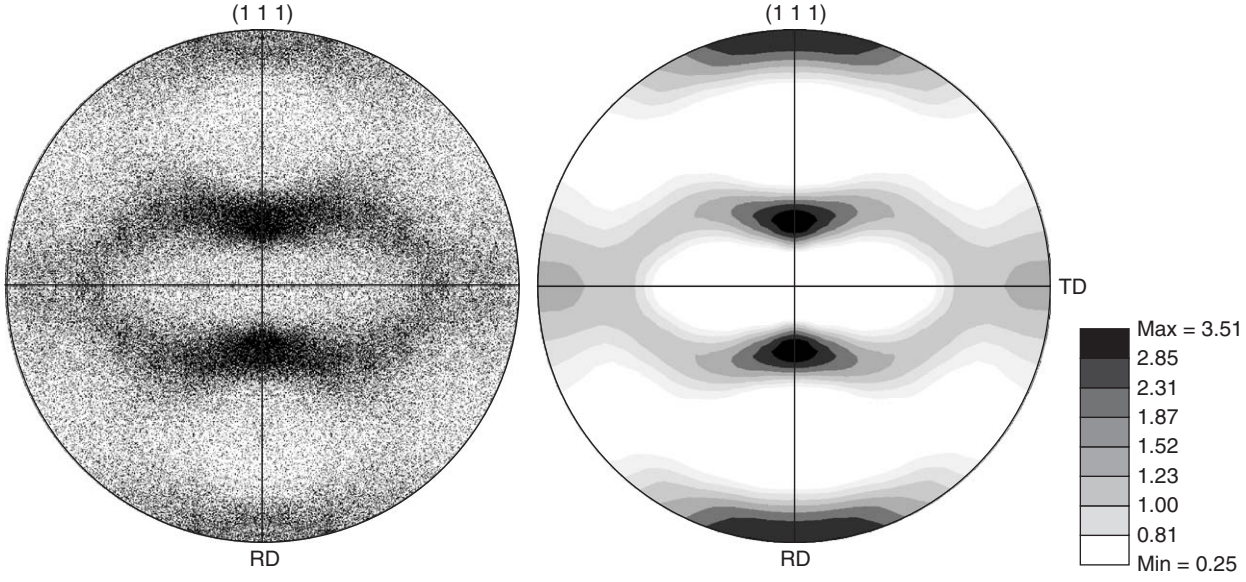
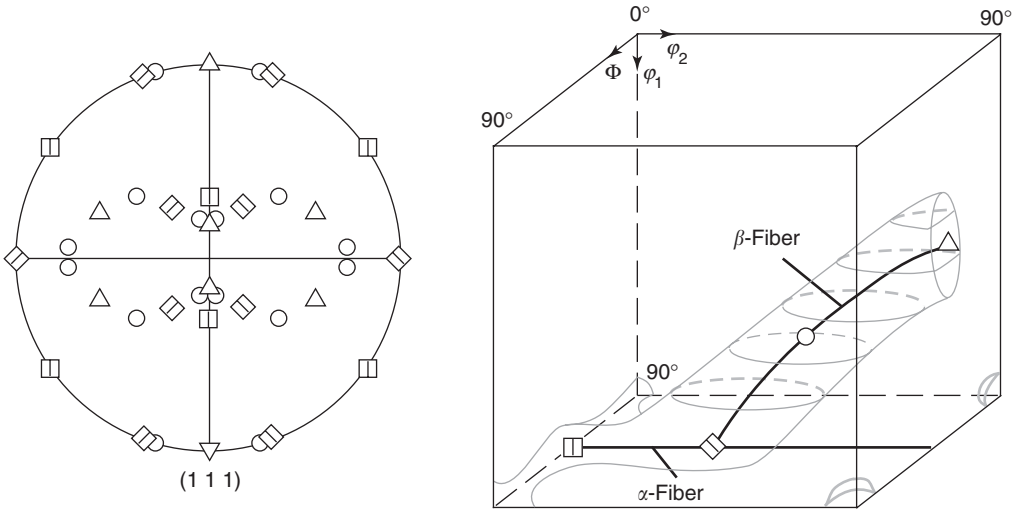


Figure 20 Typical pole figures showing discrete orientations and calculated from harmonic coefficients for rolled aluminum.



Name	Indices	$(\varphi_1, \Phi, \varphi_2)$
△ Copper	(1 1 2)<1 1 $\bar{1}$>	90°, 35°, 45°
○ S3	(1 2 3)<6 3 $\bar{4}$>	59°, 37°, 63°
◇ Brass	(1 1 0)<$\bar{1}$ 1 2>	35°, 45°, 0°
◫ Goss	(1 1 0)<0 0 1>	0°, 45°, 0°

Figure 21 ODF and pole figures showing ideal components for face-centered cubic rolling textures.

See also: Crystal Symmetry; Meso- and Nanostructures; Micro- and Macrostructures; Scanning Probe Microscopy.

PACS: 81.05.–t; 81.40.–z

Further Reading

Bunge HJ (1993) *Texture Analysis in Materials Science: Mathematical Methods*. Göttingen: Cuvillier Verlag.

Dillamore I and Roberts W (1965) Preferred orientation in wrought and annealed metals. *Metallurgical Reviews* 10: 271–380.

Field DP (2003) Textured structures, In: *ASM Handbook Volume 9 – Metallography and Microstructures*. Materials Park, Ohio: ASM International (in press).

Gambin W (2001) *Plasticity and Textures*. Dordrecht: Kluwer Academic Publishers.

Kallend JS, Kocks UF, Rollett AD, and Wenk HR (1991) Operational texture analysis. *Materials Science and Engineering* A132: 1–11.

Kocks UF, Tomé CN, and Wenk HR (1998) *Texture and Anisotropy: Preferred Orientations in Polycrystals and Their Effect on Materials Properties*, 1st edn. Cambridge: Cambridge University Press.

Randle V and Engler O (2000) *Introduction to Texture Analysis: Macrotexture, Microtexture and Orientation Mapping*. Amsterdam: Gordon Breach.

Schwartz AJ, Kumar M, and Adams BL (2000) *Electron Backscatter Diffraction in Materials Science*. New York: Kluwer Academic/Plenum Publishers.

Tenckhoff E (1988) *Deformation Mechanisms, Texture and Anisotropy in Zirconium and Zircaloy*. Philadelphia: ASTM.

Wenk HR (1985) *Preferred Orientation in Deformed Metals and Rocks: An Introduction to Modern Texture Analysis*. Orlando: Academic Press.

Wenk HR (2003) Texture and anisotropy. In: *Plastic Deformation of Minerals and Rocks, Reviews in Mineralogy and Geochemistry*, vol. 51, pp. 291–321.

Nomenclature

C_{ijkl}	elastic stiffness tensor
$\bar{C}_{ijkl}$	polycrystal average elastic stiffness tensor
C_l^{mn}	series expansion coefficients for the harmonic series expansion of the ODF
d_{hkl}	interplanar spacing
$f(g)$	orientation distribution function (ODF)

g	crystallographic orientation
$\boldsymbol{g}$	matrix description of crystallographic orientation
$\boldsymbol{g}^{e}$	symmetrical equivalent orientation in matrix form
g_o	specific crystallographic orientation
$\boldsymbol{h}$	vector describing a plane normal
(hkl)	Miller indices describing a crystallographic plane
$k_l^n(\boldsymbol{y})$	spherical surface harmonic functions
K	smoothing factor for ODF coefficient calculation
$\boldsymbol{L}_i$	element of a crystal symmetry group in matrix form
N	number of individual orientation measurements
$P_h(\boldsymbol{y})$	pole figure
$\boldsymbol{R}_i$	element of a sample symmetry group in matrix form
$T_l^{mn}(g)$	generalized spherical harmonic function
$[uvw]$	Miller indices describing a crystallographic direction
V	sample volume
$\boldsymbol{X}^C$	crystal reference frame
X_i^C	the ith axis of the crystal reference frame
$\boldsymbol{X}^S$	sample reference frame
X_i^S	the ith axis of the sample reference frame
$\boldsymbol{y}$	vector describing a sample direction
Δg	misorientation
$\Delta V_{(g_o+\Delta g)}$	sample subvolume containing grains oriented within a tolerance of specific orientation
θ	Bragg angle
λ	wavelength
ϕ	rotation angle (azimuthal angle) in pole figures
χ	specific path through orientation space
Ψ	tilt angle (polar angle) in pole figures
φ_1, Φ, φ_2	Euler angle description of crystallographic orientation

P

Paramagnetism

P F de Châtel, New Mexico State University, Las Cruces, NM, USA

Paramagnetic materials do not order magnetically and have a positive magnetic susceptibility, which means the magnetization is pointing in the same direction as the magnetic field. Inevitably, there is also a diamagnetic susceptibility. The material is classified as paramagnetic if the sum of the two contributions is positive. Nevertheless, the paramagnetic susceptibility can be determined in both kinds of materials. In addition, the paramagnetic susceptibility of magnetically ordered materials (ferro-, ferri-, and antiferromagnets) above the transition temperature can be a subject of study.

In the most common paramagnetic materials, the response to an external magnetic field is due to the alignment of atomic magnetic moments under the torque exerted by the field. The potential energy of the moment $\boldsymbol{m}_{\mathrm{at}}$ in a field $\boldsymbol{B}$ is

$$V = -\boldsymbol{m}_{\mathrm{at}} \cdot \boldsymbol{B} \qquad [1]$$

The magnitude of the atomic moment is of the order of the Bohr magneton, $\mu_{\mathrm{B}} = e\hbar/2m_{\mathrm{e}} = 9.27\times 10^{-24}\ \mathrm{J\ T^{-1}}$, where e is the electron charge, $\hbar = h/2\pi$ the Planck constant, and $\boldsymbol{m}_{\mathrm{e}}$ the electron mass. Equation [1] shows that $\boldsymbol{m}_{\mathrm{at}}$ tends to align parallel to the magnetic field and the energy gained by alignment in a field of 1 T is $\sim 10^{-23}$ J, less than 1% of the thermal energy per degree of freedom, $\frac{1}{2}k_{\mathrm{B}}T$, at room temperature ($k_{\mathrm{B}}$ is the Boltzmann constant). At ambient temperature, in a regular electromagnet, one should not expect a substantial paramagnetic magnetization. On the other hand, if the moments are free, that is, no interaction hampers their rotation, lowering the temperature can reduce the thermal fluctuations, whereupon the paramagnetic susceptibility will increase. If none of the constituent atoms of a material has a magnetic moment, there can still be a paramagnetic moment, due to moments induced by the magnetic field. The resulting Van Vleck susceptibility is independent of temperature (see below).

The value of the atomic moment depends on the number of electrons and on their angular momentum. The filled shells of the ion cores and of fully ionized species, like O^{2-} or Cu^{+}, have zero angular momentum and hence no magnetic moment. The outermost electron shells are often but partially occupied and can carry a magnetic moment. However, in crystalline solids, these shells provide the delocalized Bloch states and contribute to the cohesive energy of the crystal. This requires a considerable overlap of the atomic wave functions on neighboring sites. The corresponding states are severely perturbed and cannot be labeled by the angular momentum quantum numbers ℓ and m_ℓ. In the case of transition metals, lanthanides, and actinides, under certain circumstances, there is a distinction between the outermost s and p states and the more localized states in the d and f shells, which are not fully occupied either. The former are responsible for the cohesion and determine the interatomic distances, while the latter remain localized and can form localized moments. This subtle division of electronic states makes transition and rare-earth metals and their compounds the most important ordered magnetic materials and also the strongest paramagnets.

A given number of d or f electrons will form a total angular momentum in accordance with Hund's rules, which require that (1) the total spin angular momentum S be as large as possible without violating the Pauli principle; (2) the total orbital angular momentum L be as large as possible without violating the Pauli principle and requirement (1); and (3) S and L be aligned antiparallel and form the total angular momentum, $J = |L - S|$, if the shell is less than half filled, and parallel, $J = L + S$, if it is more than half filled.

The hierarchy of the three rules of Hund reflects a hierarchy among the interactions involved. Of these, the exchange interaction, which favors the parallel-spin arrangement, rule (1), is the strongest. The preference for the alignment of orbital moments, rule (2), is dictated by the less important correlation energy, while the spin–orbit coupling, responsible for rule (3), is ineffective in transition metals and quite important in the rare earths. **Table 1** shows the ground states

Table 1 The ground states and effective moments of transition-metal ions

Ion	*Configuration*	*Ground state*	$\mu_{\text{eff}} = g\sqrt{J(J+1)}$	$\mu_{\text{eff}} = 2\sqrt{S(S+1)}$
Ti^{3+}, V^{4+}	$3d^1$	$^2D_{3/2}$	1.55	1.73
V^{3+}	$3d^2$	$^3F_2$	1.63	2.83
Cr^{3+}, V^{2+}	$3d^3$	$^4F_{3/2}$	0.77	3.87
Mn^{3+}, Cr^{2+}	$3d^4$	$^5D_0$	0	4.90
Fe^{3+}, Mn^{2+}	$3d^5$	$^6S_{5/2}$	5.92	5.92
Co^{3+}, Fe^{2+}	$3d^6$	$^5D_4$	6.70	4.90
Co^{2+}	$3d^7$	$^4F_{9/2}$	6.63	3.87
Ni^{2+}	$3d^8$	$^3F_4$	5.59	2.83
Cu^{2+}	$3d^9$	$^2D_{5/2}$	3.55	1.73

Table 2 The ground states and effective moments of tripositive rare-earth ions

Ion	*Configuration*	*Ground state*	$\mu_{\text{eff}} = g\sqrt{J(J+1)}$	μ_{eff} *(measured)*
Ce^{3+}	$4f^1$	$^2F_{5/2}$	2.54	2.4
Pr^{3+}	$4f^2$	$^3H_4$	3.58	3.5
Nd^{3+}	$4f^3$	$^4I_{9/2}$	3.62	3.5
Pm^{3+}	$4f^4$	$^5I_4$	2.68	–
Sm^{3+}	$4f^5$	$^6H_{5/2}$	0.84	1.5
Eu^{3+}	$4f^6$	$^7F_0$	0	3.4
Gd^{3+}	$4f^7$	$^8S_{7/2}$	7.94	8.0
Tb^{3+}	$4f^8$	$^7F_6$	9.72	9.5
Dy^{3+}	$4f^9$	$^6H_{15/2}$	10.63	10.6
Ho^{3+}	$4f^{10}$	$^5I_8$	10.60	10.4
Er^{3+}	$4f^{11}$	$^4I_{15/2}$	9.59	9.5
Tm^{3+}	$4f^{12}$	$^3H_6$	7.57	7.3
Yb^{3+}	$4f^{13}$	$^2F_{7/2}$	4.54	4.5

following from the three rules for transition-metal ions in the conventional notation $(2S+1)X_J$, where $X = S$, P, D, and F stand for $L = 0$, 1, 2, and 3, respectively. Also shown are the effective moments, calculated for the $L-S$ coupled as well as spin-only states, which will be discussed below. **Table 2** gives the corresponding data for the tripositive rare-earth ions (H and I stand for $L = 5$ and 6, respectively).

If the quantum numbers of the ground state are known, the atomic magnetic moment can be calculated as $m_{\text{at}} = g_J \mu_{\text{B}} J$, where

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

is the Landé g factor.

There is no classical analog of the above procedure of determining the magnitude of the atomic moments as there should not be one, considering that the Bohr–van Leeuwen theorem precludes orbital magnetism in a classical system. What is often referred to as the classical derivation of the paramagnetic magnetization is the application of classical statistics to the quantum mechanical objects $\boldsymbol{m}_{\text{at}}$. In the context of classical statistics, the partition function involves the potential energy (eqn [1]), where $\boldsymbol{m}_{\text{at}} \cdot \boldsymbol{B} = m_{\text{at}} B \cos \vartheta$, where ϑ is a continuous variable and integration is to be performed over φ and ϑ, the angular coordinates of the vector $\boldsymbol{m}_{\text{at}}$. In the proper quantum mechanical description, a summation over the possible values of the z component of $\boldsymbol{m}_{\text{at}}$ is required and the partition function takes the form

$$Z = \left(\sum_{M_J=-J}^{J} \exp(g_J \mu_{\text{B}} B M_J / k_{\text{B}} T) \right)^{N_{\text{at}}} \qquad [2]$$

where N_{at} is the number of atoms or ions. Summation of the geometrical series leads to

$$Z = \left(\frac{\sinh\left(g_J \mu_{\text{B}} B (J + 1/2)/k_{\text{B}} T\right)}{\sinh(g_J \mu_{\text{B}} B / 2 k_{\text{B}} T)} \right)^{N_{\text{at}}}$$

Deriving the free energy as $F = -(1/k_{\text{B}}T) \ln Z$ and using $M = (-1 \partial F / V \partial B)$, the magnetization is found in the form

$$M = \frac{N_{\text{at}}}{V} g_J \mu_{\text{B}} B_J (g_J \mu_{\text{B}} B)$$

where

$$B_J(x) = \frac{2J+1}{2J}\coth\frac{2J+1}{2J}x - \frac{1}{2J}\coth\frac{1}{2J}x$$

is the Brillouin function. For large values of J, $B_J(x)$ approaches the Langevin function, $L(x) = \coth x - 1/x$ which is the result of the classical treatment outlined above. The susceptibility is found to obey Curie's law,

$$\chi_C = \left(\frac{\partial M}{\partial H}\right)_{H=0} = \frac{N_{at}}{V}\mu_0\frac{(g_J\mu_B)^2 J(J+1)}{3k_BT} = \frac{C}{T} \quad [3]$$

where $\mu_0 = 4\pi \times 10^{-7}$ T mA^{-1} is the permeability of free space and the Curie constant, C, is implicitly defined by this equation. The Curie law is well obeyed by various salts and other compounds, where the atoms or ions carrying magnetic moments are far enough from each other, so that their interaction, the exchange interaction, is negligible. In these cases the effective magnetic moment, $\mu_{eff} = g_J\mu_B\sqrt{J(J+1)}$, can be determined from the slope of χ^{-1} versus T plots. In paramagnetic rare-earth materials, the measured μ_{eff} values agree with the Hunds rule values listed in **Table 2**. In most transition-metal oxides and salts, however, the μ_{eff} values calculated from the Hunds rule spin quantum numbers alone are closer to the measured values than the ones involving J, indicating that crystal fields are strong enough to suppress ("quench") the orbital angular moment. If the exchange interaction between magnetic moments is not negligible, the Curie–Weiss law, $\chi_{CW} \propto (T-\theta)^{-1}$, often holds. To account for the interaction, whose source was unknown at that time (1907), Weiss has introduced the molecular field H_m, which acts, in addition to the external field H, on each atomic magnetic moment. As H_m is supposed to describe the tendency of atomic moments to point in the same direction, it is assumed to be proportional with the magnetization, $H_m = \lambda M$. Substituting this expression in the expression for the magnetization in the total field,

$$M = \chi_C(H + H_m) = \frac{C}{T}(H + \lambda M)$$

the magnetization can be expressed as $M = \chi_{CW}H$ with

$$\chi_{CW} = \frac{\chi_C}{1 - \lambda\chi_C} = \frac{C}{T-\theta} \quad [4]$$

where $\theta = C\lambda$ is the paramagnetic Curie temperature. It is seen that the slope of χ_{CW}^{-1} versus T plots are determined by the effective moment the same way as in the case of the Curie law, but the straight lines representing the measured data do not go through the origin. Deviations from the straight-line behavior may be due to crystal fields.

In most metals, the effect of the crystalline environment is so strong that all couplings imposed by Hund's rules are broken and all but the ion-core electrons are delocalized. In rare-earth materials, the f electrons are localized and form magnetic moments conforming to Hund's rules. In insulators, the d electrons of transition-metal cations remain localized, but the spin–orbit coupling is broken and the orbital angular momentum itself is destroyed. The latter process is called the quenching of angular momentum, not a particularly enlightening term. What it stands for is that, the spherical symmetry around a particular ion being broken by its neighbors in the crystal, the stationary states are no more eigenstates of the angular momentum. Due to the symmetry of the crystalline environment, some degeneracy of the eigenstates may remain. However, if all matrix elements of the three components of the angular momentum, $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$, between the degenerate wave functions vanish, its expectation value vanishes for any linear combination of these functions; the angular momentum is "quenched." If that is not the case, the angular momentum may be partially quenched and can be described by an effective angular momentum operator working within the multiplet of the lowest-lying crystal-field level.

Partial or complete quenching of the orbital angular momentum influences the temperature dependence of the paramagnetic susceptibility if the temperature is high enough for the thermal energy k_BT to be comparable to the splitting between the lowest and next-to-lowest crystal-field levels. At temperatures exceeding Δ/k_B, Δ being the total splitting of the ground-state multiplet, the Curie or Curie–Weiss susceptibility is recovered with the effective moment appropriate to that multiplet.

Higher-lying crystal-field states can influence the paramagnetic susceptibility also at very low temperatures, $T \ll \Delta/k_B$. Even though the thermal occupancy of such states is negligible in this limit, the magnetic field will mix them into the ground state via the perturbing Hamiltonian (eqn [1]). Taking account of this perturbation to second order, the corresponding contribution to the susceptibility, the Van Vleck paramagnetic susceptibility, is found to be

$$\chi_{VV} = \frac{N_{at}}{V}\mu_0 2\sum_i \frac{\left|\int \psi_0^*(\hat{m}_{at})_z\psi_i d\tau\right|^2}{\varepsilon_i - \varepsilon_0} \quad [5]$$

independent of temperature. Here ψ_0,ψ_i are the wave functions and ε_0,ε_i the energies of the ground state and higher-lying crystal-field states, respectively. The

operator representing the atomic magnetic moment is

$$\hat{\boldsymbol{m}}_{\mathrm{at}} = \mu_{\mathrm{B}}(\hat{\boldsymbol{L}} + 2\hat{\boldsymbol{S}})$$

evidently not proportional to $\hat{J}$. Therefore, $(\hat{m}_{\mathrm{at}})_z$ has nonvanishing matrix elements between the Hund's rule ground state and higher-lying eigenstates of different J values as well. In fact, Van Vleck derived eqn [5] for that case, with ions having small spin–orbit splitting in mind. However, the term "Van Vleck susceptibility" is used with reference to any temperature-independent contribution to the paramagnetic susceptibility due to the quantum-mechanical admixture of higher-lying states.

As mentioned above, crystal-fields in insulators constitute a subtle effect, compared to the full delocalization of conduction electrons in metals. The consequence of the latter for the paramagnetic response is indeed drastic. The conduction electrons, instead of following the Curie law (eqn [3]) with $g = 2$ and $S = 1/2$, contribute only the Pauli paramagnetism, a small, temperature-independent susceptibility,

$$\chi_P = \mu_0 \mu_{\mathrm{B}}^2 \rho(\varepsilon_{\mathrm{F}}) \qquad [6]$$

where $\rho(\varepsilon_{\mathrm{F}})$ is the density of states per unit volume at the Fermi energy. It is indeed the Pauli principle that suppresses the $1/T$ divergence of χ: most of the electrons are unable to flip their spin, because the opposite-spin state of the same energy is occupied. Only the electrons in states with energies within $k_{\mathrm{B}}T$ from the Fermi energy are given the choice of m_{s} values implied by eqn [2]. As the density of such electrons is about $k_{\mathrm{B}}T\rho(\varepsilon_{\mathrm{F}})$, the temperature in the denominator of the Curie susceptibility is canceled out and eqn [6] can be seen to follow from eqn [3].

In the alkali metals, $\rho(\varepsilon_{\mathrm{F}}) \approx 10^{47}\ \mathrm{J}^{-1}\ \mathrm{m}^{-3}$ whence from eqn [6] $\chi_P \approx 10^{-5}$, in good agreement with the paramagnetic susceptibility derived from experiment. In transition metals, the density of states is about one order of magnitude larger, but even so, eqn [6] underestimates the paramagnetic susceptibility, which is enhanced by the exchange interaction between the d electrons. Like in the derivation of the Curie–Weiss law, here too the effect of this interaction can be described by the molecular field. The result,

$$\chi = \frac{\chi_P}{1 - \lambda\chi_P} \qquad [7]$$

is formally similar to eqn [4], but it should be noted that it is independent of temperature in this case. Equation [7] gives the exchange-enhanced paramagnetic susceptibility, $S = (1 - \lambda\chi_P)^{-1}$ is the Stoner enhancement factor. Among the transition metals, the largest value of S is observed in palladium, where $S = 10$.

See also: Diamagnetism; Ferromagnetism; Localized and Itinerant Magnetism; Magnetism, History of.

PACS: 75.20.Ck; 75.20.En; 75.30.Cr

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Reinhart and Winston.
Chikazumi S (1965) *Physics of Magnetism*. New York: Wiley.
Kittel C (1986) *Introduction to Solid State Physics*, 6th edn. New York: Wiley.
O'Handley RC (2000) *Modern Magnetic Materials: Principles and Applications*. New York: Wiley.
Van Vleck JH (1932) *The Theory of Electric and Magnetic Susceptibilities*. Oxford: Clarendon Press.

Nomenclature

$\boldsymbol{B}$	magnetic field (T)
H_{m}	molecular field ($\mathrm{A\,m^{-1}}$)
$\boldsymbol{L}$	orbital angular momentum ($\mathrm{kg\,m^2\,s^{-1}}$)
$\boldsymbol{m}_{\mathrm{at}}$	atomic magnetic momentum ($\mathrm{kg\,m^2\,s^{-1}}$)
M	magnetization ($\mathrm{A\,m^{-1}}$)
S	Stoner enhancement factor (dimensionless)
$\rho(\varepsilon)$	density of states ($\mathrm{J^{-1}\,m^{-3}}$)

Periodicity and Lattices

F J Lamelas, Boise State University, Boise, ID, USA

Lattices in Two and Three Dimensions

A lattice is a periodic array of points generated by translation vectors (quasiperiodic lattices are discussed separately later). As an example, consider a two-dimensional rectangular lattice generated by the orthogonal vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ (**Figure 1**).

The lattice is assumed to be infinite in extent. Points on this lattice have position vectors

$$\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 \qquad [1]$$

where n_1 and n_2 are integers.

Lattices can be classified according to their symmetry properties. For example, the rectangular lattice in **Figure 1** has reflection lines which pass through the vertical and horizontal rows of points, and twofold rotation points coincident with the lattice points. Additional symmetry elements occur between the lattice points, as shown in **Figure 2**.

According to the lattice condition, all points of the lattice must have the same environment; therefore, not all regular arrays of points are lattices. For example, in one dimension there is only one possible lattice: an array of equally spaced points, as shown in **Figure 3**.

If one classifies all two-dimensional lattices according to their symmetry elements, it is possible to show that there are only five distinct types, as shown in as **Figure 4**.

The translation vectors define the edges of unit cells which are building blocks of the lattice. In two dimensions, the shape of the unit cell is specified by three lattice parameters: the magnitudes a and b of the translation vectors, and γ, the angle between the vectors (the magnitudes of the translation vectors are typically defined as $a = |\boldsymbol{a}_1|$, $b = |\boldsymbol{a}_2|$, and, in three dimensions, $c = |\boldsymbol{a}_3|$).

The choice of the unit cell is by no means unique. One could, for example, double the unit cells of **Figure 4** and generate the same lattices. Unit cells containing only one lattice point are primitive; in this case, the positions of all lattice points can be written as integer sums of the translation vectors, as in eqn [1]. For the centered unit cell of **Figure 4c**, the lattice points at the centers are given by half-integer sums of the translation vectors:

$$\boldsymbol{R}_{\text{center}} = (n_1 + 1/2)\boldsymbol{a}_1 + (n_2 + 1/2)\boldsymbol{a}_2 \qquad [2]$$

Even primitive cells are not unique, since one could, for example, define an oblique primitive cell for the square lattice, as shown in **Figure 5**.

In general, the conventional unit cell is chosen so as to be as small as possible, consistent with the underlying symmetry of the lattice. For example, the centered rectangular lattice (**Figure 4c**) could be generated by an oblique primitive cell, but its symmetry elements are more obviously related to a rectangular unit cell (the symmetry elements of the

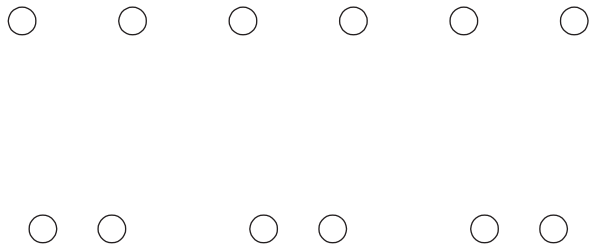


Figure 3 Two periodic one-dimensional arrays of points. The lower example does not satisfy the lattice condition since some points have their nearest neighbor on the right, and some on the left.

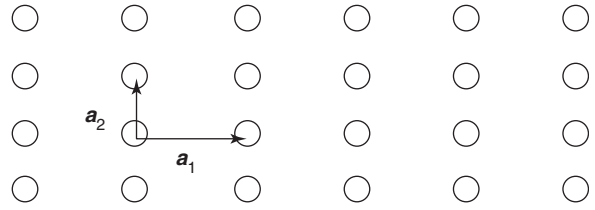


Figure 1 Rectangular lattice with translation vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

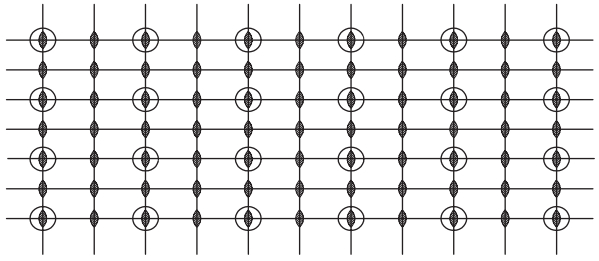


Figure 2 Symmetry elements of the rectangular lattice. The points of the lattice are indicated by open circles, reflection lines by solid lines, and twofold rotation points by lens-shaped symbols.

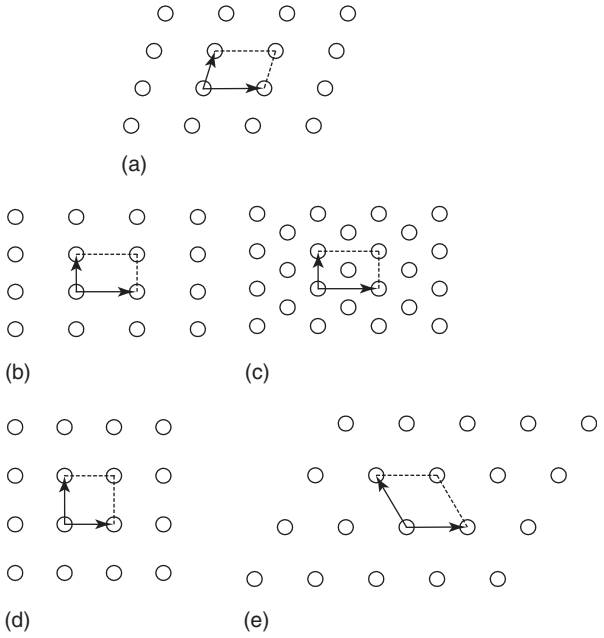


Figure 4 The five two-dimensional lattices: (a) oblique, (b) rectangular, (c) centered rectangular, (d) square, and (e) hexagonal. Outlines of unit cells are shown by the dashed lines. Except for (c), all of the unit cells shown here are primitive.

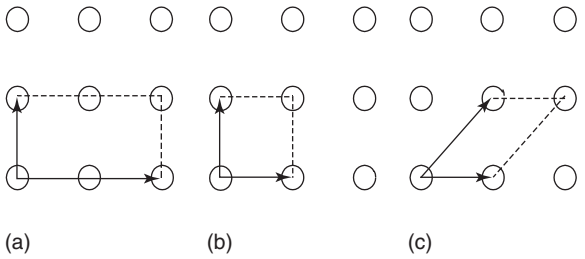


Figure 5 Several unit cells which can generate a square lattice, including (a) a nonprimitive cell containing two lattice points, (b) a conventional primitive cell, and (c) an oblique primitive cell. In counting the number of lattice points assigned to a cell, points at corners, on sides, and within the cell are given weights of 1/4, 1/2, and 1.

Table 1 The four two-dimensional systems

System	*Lengths*	*Angle* (°)
Oblique	$a \neq b$	$\gamma \neq 90$
Rectangular	$a \neq b$	$\gamma = 90$
Square	$a = b$	$\gamma = 90$
Hexagonal	$a = b$	$\gamma = 120$

centered rectangular lattice are identical to those of **Figure 2**, with additional twofold rotation points along the cell diagonal). In other words, symmetry considerations can lead one to choose a conventional cell that is larger than a primitive cell.

The two-dimensional lattices of **Figure 4** can be grouped within four systems based on the relative magnitudes of the translation vectors and the angle between them, as shown in **Table 1**.

Note that the number of lattices (five) is not equal to the number of systems (four), since there are two rectangular lattices (primitive and centered).

A periodic physical system such as a naturally occurring crystal can be described as a lattice with repeating units known as a basis attached to each lattice point. The basis may consist of a single atom, a molecule, or a larger assembly. The space group that characterizes the entire structure results from the combined symmetries of the lattice and the basis. Since bases with different symmetries can be attached to a given lattice, one expects a greater number of space groups than lattices. In two dimensions, there are 4 systems, 5 lattices, and 17 space groups. An example of two structures with the same lattice but different space groups is given in **Figure 6**.

Different space groups also arise according to the arrangement of asymmetric molecules on similar lattices. For example, if every other molecule is flipped after displacement along a lattice translation vector, then a glide-reflection symmetry element is produced. (In the glide-symmetry operation, the structure is translated along the glide line and then reflected across the line.) Examples of two space groups with and without glide symmetry are shown in **Figure 7**.

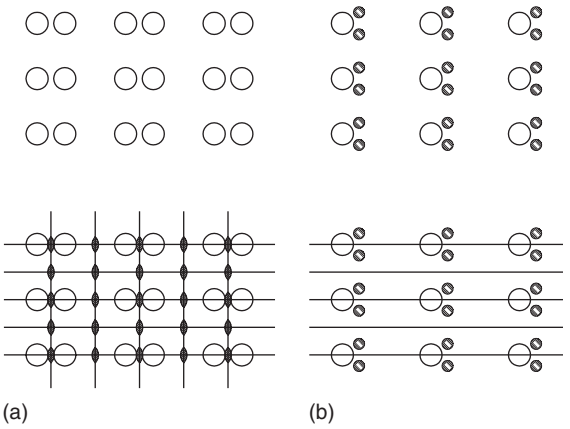


Figure 6 Upper drawings: two structures with identical rectangular lattices but different bases. Lower: the same structures repeated with symmetry elements superimposed. (a) Space group *p2mm*, indicating a primitive lattice (*p*) with twofold rotation points (2), and two sets of mirror lines (*mm*), (b) space group *pm*, indicating a primitive lattice with a single set of mirror lines.

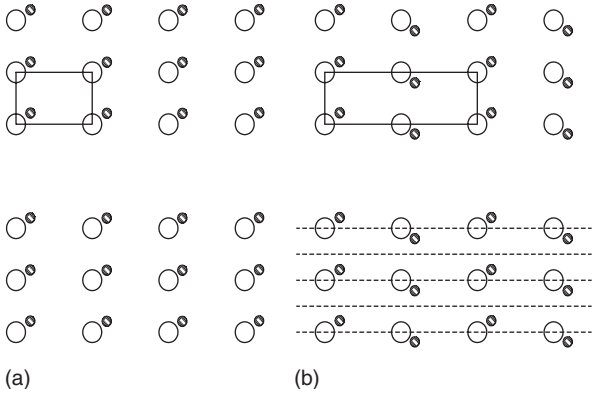


Figure 7 Upper drawings: arrangements of identical molecules with and without glide-line symmetry. Unit cells are shown by the solid rectangles. Lower: the same structures with symmetry elements superimposed. (a) Space group *p*1: primitive lattice with no rotational, mirror, or glide symmetry. (b) Space group *pg*: primitive lattice with glide-line symmetry. (Note: the unit cell contains two molecules but only one lattice point, and is therefore primitive.)

In three dimensions, lattices are generated by three translation vectors: $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. It can be shown that there are 14 symmetrically distinct lattices which satisfy the lattice condition. These are the Bravais lattices; the conventional unit cells are shown in **Figure 8**. The shapes of the unit cells are characterized by six lattice parameters: the magnitudes a, b, and c of the translation vectors, and the angles α, β, and γ, shown in **Figure 9**.

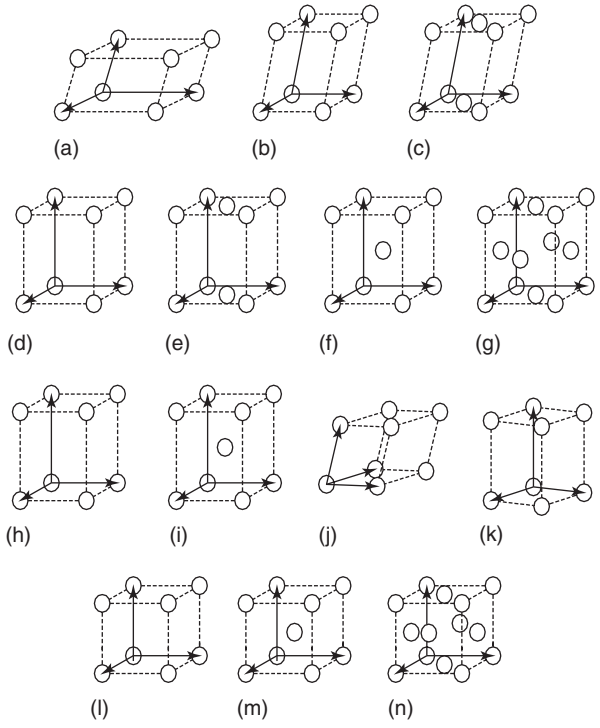


Figure 8 The 14 Bravais lattices, with lattice translation vectors shown as solid arrows. (a) triclinic, (b) monoclinic, (c) base-centered monoclinic, (d) orthorhombic, (e) base-centered orthorhombic, (f) body-centered orthorhombic, (g) face-centered orthorhombic, (h) tetragonal, (i) body-centered tetragonal, (j) trigonal, (k) hexagonal, (l) cubic, (m) body-centered cubic, and (n) face-centered cubic.

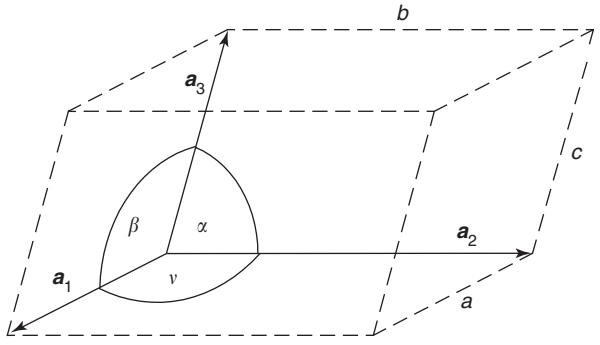


Figure 9 Three-dimensional lattice parameters. The lengths of the cell edges are the magnitudes of the translation vectors *a*, *b*, and *c*. The interaxial angles $\boldsymbol{a}_2 \wedge \boldsymbol{a}_3$, $\boldsymbol{a}_3 \wedge \boldsymbol{a}_1$, and $\boldsymbol{a}_1 \wedge \boldsymbol{a}_2$ are given by α, β, and γ.

The 14 Bravais lattices are distributed among seven crystal systems, with relative lattice parameters given in **Table 2**.

The combined symmetries of the 14 Bravais lattices and all possible bases lead to 230 space groups in three dimensions, first enumerated late in the nineteenth century and described explicitly in the International Tables for X-ray Crystallography. The classification of

Table 2 The seven three-dimensional crystal systems

System	*Lengths*	*Angles*
Triclinic	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$
Monoclinic	$a \neq b \neq c$	$\alpha = \beta = 90^\circ \neq \gamma$
Orthorhombic	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Trigonal	$a = b = c$	$\alpha = \beta = \gamma \neq 90^\circ$
Hexagonal	$a = b \neq c$	$\alpha = \beta = 90^\circ,\ \gamma = 120^\circ$
Cubic	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

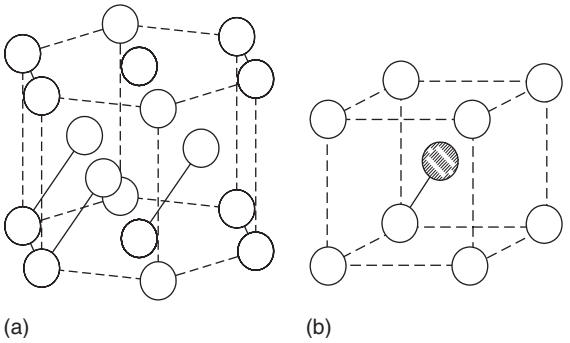


Figure 10 Examples of structures produced by adding a basis to a Bravais lattice. (a) h.c.p cobalt. Pairs of atoms in the basis are connected by solid lines. The dashed lines show the edges of a hexagonal repeating structure with three times the volume of the conventional unit cell of **Figure 8k**. (b) Cesium chloride. Open and filled circles represent oppositely-charged ions. The primitive cubic unit cell (also known as a simple cubic unit cell) is identical to that of **Figure 8l**.

periodic structures as members of specific space groups is important in predicting properties, such as allowed reflections in X-ray diffraction, and allowed vibrational modes in optical spectroscopy.

A three-dimensional crystal can be "constructed" by placing a specific basis at each of the points of one of the 14 Bravais lattices. In the simplest case, the basis consists of a single atom and the full structure is simply that of the Bravais lattice. An example is face-centered cubic (f.c.c.) aluminum, which has the unit cell shown in **Figure 8n**. As the complexity is increased, the basis can include more than one atom of the same element, as in hexagonal close-packed (h.c.p.) cobalt. In this structure, a basis consisting of two cobalt atoms is attached to the hexagonal lattice points, as shown in **Figure 10a**. In the case of compounds, the basis contains two or more elements, as in cesium chloride, where the two-ion basis is attached to a primitive cubic lattice, as in **Figure 10b**.

In all cases, it is the Bravais lattice (rather than the details of the basis) which determines the translational symmetry of the structure. That is, the crystal is invariant after a translation by a Bravais lattice vector, no matter what the basis is. This spatial

periodicity results in the great utility of Fourier methods, as shown in the next section.

The Reciprocal Lattice, Diffraction, and Electronic Band Structure

A signal which is periodic in time can be conveniently described in terms of the strength of its harmonics. The signal can be defined completely in either the time or the frequency domain, using variables which have dimensions of time or inverse time. Similarly, a property such as the electron density of a spatially periodic structure can be described directly, in a space which has dimensions of length, or by a Fourier expansion in reciprocal lattice vectors, which have dimensions of inverse length. Many of the properties of a solid, such as its electronic band structure, are formulated most easily in reciprocal space (rather than direct space). Even from a purely structural standpoint, the reciprocal-space description is essential in the diffraction techniques which are used to solve crystal structures.

If one wishes to use the reciprocal lattice in order to describe physical properties, it is useful to define the reciprocal lattice as follows (in crystallography, the definition of the reciprocal lattice may differ slightly from that given here, for example, by the omission of the factors of 2π). Consider a solid with lattice translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. The reciprocal lattice vectors are defined by

$$\boldsymbol{b}_1 = 2\pi \frac{\boldsymbol{a}_2 \times \boldsymbol{a}_3}{\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3}, \quad \boldsymbol{b}_2 = 2\pi \frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3}, \quad \boldsymbol{b}_3 = 2\pi \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3} \qquad [3]$$

The denominator of these expressions $(\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3)$ is the volume of the unit cell in direct space. The vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ are the translation vectors which define the reciprocal lattice, just as $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ define the Bravais lattice in direct space. An arbitrary translation vector, which connects two points of the Bravais lattice, has the form

$$\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \qquad [4]$$

where n_1, n_2, and n_3 are integers. Similarly, an arbitrary reciprocal lattice vector can be written as

$$\boldsymbol{G} = m_1\boldsymbol{b}_1 + m_2\boldsymbol{b}_2 + m_3\boldsymbol{b}_3 \qquad [5]$$

The form of the numerators in eqn [3] leads to a compact expression for dot products between direct-space and reciprocal-space translation vectors. Note, for example, that $\boldsymbol{a}_2 \cdot \boldsymbol{b}_1 = \boldsymbol{a}_3 \cdot \boldsymbol{b}_1 = 0$, and $\boldsymbol{a}_1 \cdot \boldsymbol{b}_1 = 2\pi$. In general,

$$\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij} \qquad [6]$$

This implies that any function of the form $e^{i\boldsymbol{G}\cdot\boldsymbol{r}}$ will have the translational symmetry of the Bravais lattice, because, for an arbitrary lattice translation vector $\boldsymbol{R}$,

$$\begin{aligned} e^{i\boldsymbol{G}\cdot(\boldsymbol{r}+\boldsymbol{R})} &= e^{i\boldsymbol{G}\cdot\boldsymbol{r}} e^{i\boldsymbol{G}\cdot\boldsymbol{R}} \\ &= e^{i\boldsymbol{G}\cdot\boldsymbol{r}} e^{i(m_1\boldsymbol{b}_1+m_2\boldsymbol{b}_2+m_3\boldsymbol{b}_3)\cdot(n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2+n_3\boldsymbol{a}_3)} \\ &= e^{i\boldsymbol{G}\cdot\boldsymbol{r}} e^{2\pi i(m_1n_1+m_2n_2+m_3n_3)} = e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \end{aligned} \qquad [7]$$

For a Bravais lattice where the translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ are all perpendicular (these are the orthorhombic, tetragonal, and cubic cases), the corresponding reciprocal lattice can be immediately determined by inspection of eqn [3]. In this case, $\boldsymbol{b}_1 \parallel \boldsymbol{a}_1$, and $b_1 = 2\pi/a_1$. Similar relations hold for $\boldsymbol{b}_2$ and $\boldsymbol{b}_3$, that is, the reciprocal-space translation vectors are parallel to those in direct space, with magnitudes which are inversely proportional.

If the direct-space lattice vectors are not all perpendicular, then the reciprocal lattice vectors are not, in general, parallel to them, because, for example, $\boldsymbol{a}_2 \times \boldsymbol{a}_3$ may not be parallel to $\boldsymbol{a}_1$. Nonetheless, the reciprocal lattice vectors still have a straightforward interpretation in terms of lattice planes.

Lattice planes are labeled using Miller indices, which are assigned by taking the reciprocal of the intersection of the plane with each of the direct-space axes, as shown in **Figure 11**.

In **Figure 11a**, the first (1 0 0) plane which does not contain the origin, intersects the x-axis at a distance a from the origin. The plane is also parallel to the y and z axes; intersections with those axes are at infinity. Taking the reciprocal of these distances in lattice-parameter units, one obtains the Miller indices $(hkl) = (1\,0\,0)$ for these planes. Similarly, in **Figure 11b** the intersections at (a, a, ∞) give $(hkl) = (1\,1\,0)$. In **Figure 11c**, the intersections are $(\infty, a, a/2)$, giving Miller indices $(hkl) = (0\,1\,2)$.

If h, k, and l are integers, it can be shown that in any of the crystal systems a reciprocal lattice vector of the form

$$\boldsymbol{G}_{hkl} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3 \qquad [8]$$

will be normal to the lattice planes with indices (hkl). The magnitude will be inversely proportional to the spacing of the planes, specifically,

$$\boldsymbol{G}_{hkl} = \frac{2\pi}{d_{hkl}} \qquad [9]$$

where d_{hkl} is the spacing of the (hkl) planes.

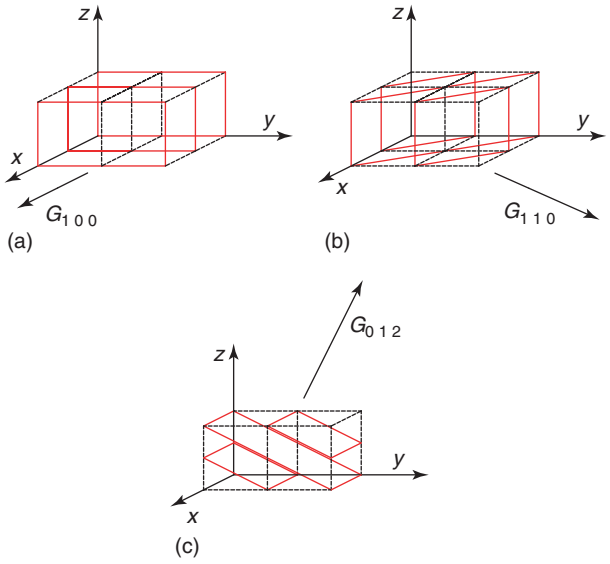


Figure 11 Several examples of sets of planes in a cubic crystal, including (a) (1 0 0) planes, (b) (1 1 0) planes, and (c) (0 1 2) planes. Unit cells are shown with dashed lines and the sets of planes are in red. The corresponding reciprocal lattice vectors are normal to the planes and have magnitudes inversely proportional to the plane spacing.

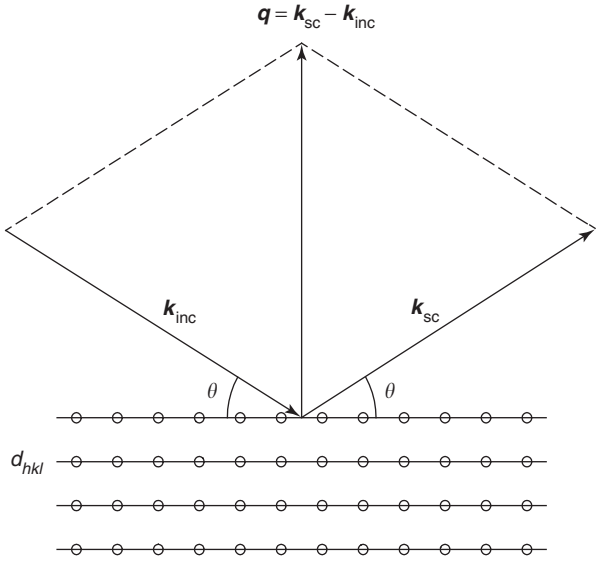


Figure 12 Bragg scattering geometry. A beam with wave vector $\boldsymbol{k}_{inc}$ is incident with angle θ upon a set of atomic planes with spacing d_{hkl}. The scattered wave vector is $\boldsymbol{k}_{sc}$ and the scattering vector is $\boldsymbol{q}$. The small circles represent atoms or lattice points in direct space.

The relationship between $\boldsymbol{G}_{hkl}$ and the (hkl) planes leads to an important formulation of the diffraction condition in the reciprocal space of the crystal. **Figure 12** shows the geometry of the Bragg diffraction. The incident beam has wavelength λ and propagates along the direction given by the wave vector $\boldsymbol{k}_{inc}$. The magnitude of the wave vector is defined by

$$k_{inc} = k_{sc} \equiv k = \frac{2\pi}{\lambda} \qquad [10]$$

where the first equality indicates that the scattering is elastic. By setting the phase shift between waves scattered from neighboring planes equal to 2π, one obtains the Bragg condition for constructive interference (or diffraction) from the hkl planes:

$$\lambda = 2\, d_{hkl} \sin\theta \qquad [11]$$

By inspection of **Figure 12**, and using eqns [10], [11], and [9], one obtains

$$q = 2k \sin\theta = 2\pi \frac{2 \sin\theta}{\lambda} = \frac{2\pi}{d_{hkl}} = G_{hkl} \qquad [12]$$

In other words, Bragg's law (eqn [11]) can be reformulated as eqn [12], which indicates that if constructive scattering from the (hkl) planes is to occur, then the magnitude of the scattering vector is equal to the magnitude of the corresponding reciprocal lattice vector. Since both $\boldsymbol{q}$ and $\boldsymbol{G}_{hkl}$ are normal to the scattering planes, eqn [12] holds for the vector quantities as well:

$$\boldsymbol{q} = \boldsymbol{G}_{hkl} \qquad [13]$$

Equation [13] is the basis of the Ewald sphere representation of the diffraction condition, shown in **Figure 13**. In an X-ray, electron, or neutron diffraction measurement, one maps out intensities associated with specific reciprocal lattice points. In essence, a diffraction measurement yields an image or a projection of the reciprocal lattice, and indirectly, information on the actual atomic ordering in the direct-space lattice.

The motion of an electron in a crystal is strongly affected by the periodic nature of the lattice. A nearly free valence electron interacts with relatively-fixed ion cores and experiences a potential, which can be written as an expansion in reciprocal-lattice vectors:

$$U(\boldsymbol{r}) = \sum_j c_j e^{i\boldsymbol{G}_j \cdot \boldsymbol{r}} \qquad [14]$$

As shown by eqn [7], a potential-energy function written in this form will have the periodicity of the Bravais lattice:

$$U(\boldsymbol{r} + \boldsymbol{R}) = U(\boldsymbol{r}) \qquad [15]$$

for any lattice vector $\boldsymbol{R}$. One might expect that the electron would be scattered erratically during its interaction with the ions; however, this is not the case. It can be shown that solutions of the Schrödinger equation in a periodic potential, such as that of

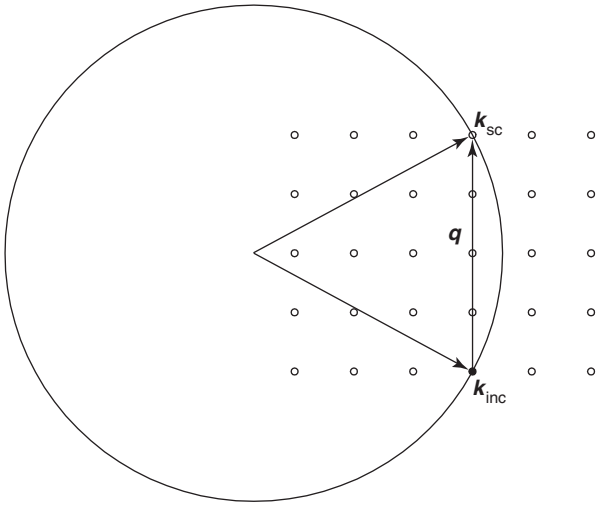


Figure 13 The Ewald sphere construction of the diffraction condition $\boldsymbol{q} = \boldsymbol{G}_{hkl}$. This figure is drawn in reciprocal space; the small circles represent reciprocal lattice points, not atoms. The tip of the incident wave vector is placed at the origin, shown as a filled circle. The surface of the sphere gives the possible locations of the tip of an elastically scattered wave vector. When the diffraction condition is met, the scattered wave vector coincides with a reciprocal lattice point. In the example shown here, if the *z*-axis is assumed to be vertical, the Bragg condition for the (0 0 4) reflection is met, that is, $\boldsymbol{q} = \boldsymbol{G}_{004}$.

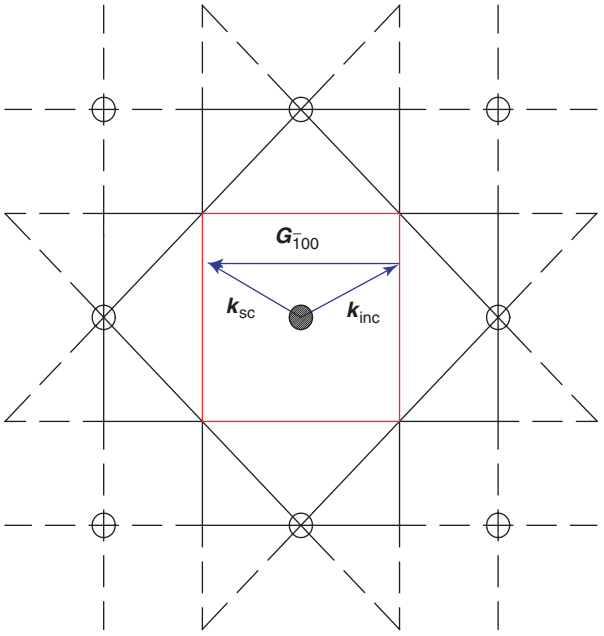


Figure 14 Construction of Wigner–Seitz cells by drawing perpendicular bisectors of reciprocal lattice vectors. The first Brillouin zone is shown in red. Electron wave vectors and a scattering vector are shown in blue. The notation $\boldsymbol{G}_{\bar{1}00}$ indicates that the scattering vector runs from the origin to the first point to the left of the origin.

eqn [14], are Bloch functions:

$$\Psi(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \qquad [16]$$

where $\hbar\boldsymbol{k}$ is the momentum of the electron, $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ is a plane wave propagating in the direction of $\boldsymbol{k}$, and $u_{\boldsymbol{k}}(\boldsymbol{r})$ is a function which has the periodicity of the Bravais lattice, that is,

$$u_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = u_{\boldsymbol{k}}(\boldsymbol{r}) \qquad [17]$$

Equation [16] indicates that electrons can propagate freely through the lattice as plane waves modulated by the lattice. The energy $\varepsilon(\boldsymbol{k})$ of an electron in a particular momentum state will have a specific form that depends on the potential energy function $U(\boldsymbol{r})$. (For a free electron (when $U = 0$), $\varepsilon \propto k^2$.)

The effect of the lattice on the electron becomes very strong when the wave vector $\boldsymbol{k}$ takes on particular values. In order to see how this comes about, consider the region of reciprocal space near the origin. Then draw a series of lines which bisect the reciprocal-lattice vectors, starting with the shortest ones, as shown in **Figure 14**. This construction divides the reciprocal space into Wigner–Seitz cells. In reciprocal space, the central cell which surrounds the origin is known as the first Brillouin zone (Wigner–Seitz cells can also be drawn in direct space. As an example, a primitive unit cell can always be constructed by generating such a cell about a Bravais lattice point).

Electrons with wave vectors within the Brillouin zone have propagating Bloch wave functions and can move through the lattice. However, if the wave vector $\boldsymbol{k}_{\mathrm{inc}}$ terminates anywhere on the zone boundary, the electron can undergo Bragg diffraction to a new state $\boldsymbol{k}_{\mathrm{sc}}$, since the difference between these wave vectors is equal to a reciprocal lattice vector. One can see the that $\boldsymbol{k}_{\mathrm{inc}}$ satisfies the condition for Bragg diffraction by comparing **Figure 14** to the Ewald sphere construction of **Figure 13**.

When the Schrödinger equation is solved in a periodic potential (one of the simplest examples being the one-dimensional Kronig-Penney model), one finds that diffraction at the Brillouin zone boundary is accompanied by a discontinuity in the energy $\varepsilon(\boldsymbol{k})$, which is otherwise continuous within the zone. That is, a fundamental property of electrons in solids - the bandgap - follows directly from the periodicity of the Bravais lattice.

Variations of Lattices: Superlattices, Surface Structures, and Quasicrystals

The discussion so far is suited to the description of three-dimensional crystalline elements and simple compounds. Also, the structures are assumed to be

infinite in extent and periodic. This section will briefly consider several departures from these assumptions.

A superlattice is a periodic structure which has a somewhat large or complex unit cell composed of smaller building blocks, something like a pedestrian walk made of a sequence of concrete and asphalt blocks, as shown in **Figure 15**.

If all of the blocks were concrete, the structure could be described by a unit cell consisting of one block. However, with the arrangement shown here, the unit cell is lengthened to four blocks.

Several examples of superlattices, which are important in condensed matter physics, are shown schematically in **Figure 16**. In an ordered metallic alloy, specific elements (shown as open and filled circles in **Figure 16a**) occupy specific sites on a lattice. The ordering produces a unit cell which is larger than that which would characterize a pure element, or a random alloy (from a diffraction point of view, a random alloy can be approximated as a single element with scattering properties intermediate between those of the individual elements). According to eqn [3], if the unit cell in direct space is enlarged, that is, if the translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ increase in length, then the corresponding reciprocal lattice vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ decrease in length. The reciprocal lattice points become denser and more peaks appear in the diffraction pattern. The new peaks are known as superlattice reflections.

Superlattices can also be produced by the intercalation of a foreign substance in a layered structure, such as graphite or one of the clay minerals, as shown in **Figure 16b**. As with the ordered alloy, the enlargement of the unit cell (relative to that of the pure host) leads to additional superlattice reflections. As in the alloy, the increase in cell size often occurs in more than one direction; therefore, the array of new reflections can be one-, two-, or three-dimensional.

An artificial superlattice, such as a semiconductor quantum-well structure or a magnetic multilayer, is prepared by the sequential deposition of thin layers of dissimilar materials onto a substrate (**Figure 16c**). In this case, the unit cell is enlarged in only one direction and the corresponding new features in the reciprocal lattice would be rows of points normal to the layers.

Two-dimensional lattices known as surface reconstructions are often formed on clean metals and semiconductors. Although the reconstruction resides on a bulk solid, its structure and reciprocal lattice can, to a certain extent, be considered as that of a very thin isolated layer. The main features of the diffraction pattern can be understood by considering the reconstruction as a limiting case: a three-dimensional structure where the interlayer spacing has become very large, as shown in **Figure 17**. If the unit cell is very large in the vertical direction (**Figure 17a**), the corresponding reciprocal lattice points become dense (**Figure 17b**). In the actual limit of a two-dimensional layer, the reciprocal lattice becomes a series of rods normal to the surface, and in diffraction one observes a set of streaks, rather than the discrete peaks which occur with a three-dimensional structure. One can also see that streaks arise by considering the exponential terms that occur in a structure-factor calculation of the diffraction intensity: $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) = \exp(\mathrm{i}(k_x x + k_y y + k_z z))$. If the surface layer lies in the $z = 0$ plane, then the scattering intensity has no k_z dependence, that is, the diffraction features consist of streaks of uniform intensity normal to the surface. One also obtains streak-like

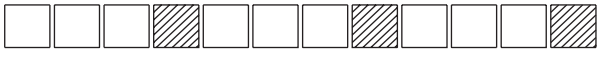


Figure 15 A one-dimensional superlattice.

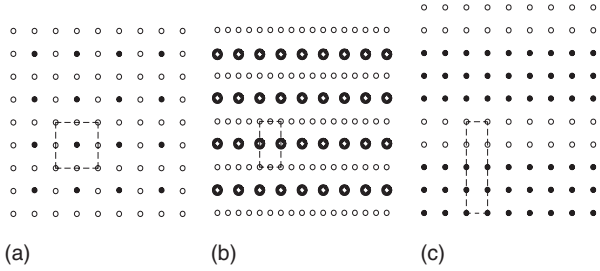


Figure 16 Examples of superlattices, including (a) an ordered alloy, (b) an intercalated layered structure, and (c) a superlattice prepared by the sequential deposition of two compounds. In all cases, the unit cell is indicated by dashed lines.

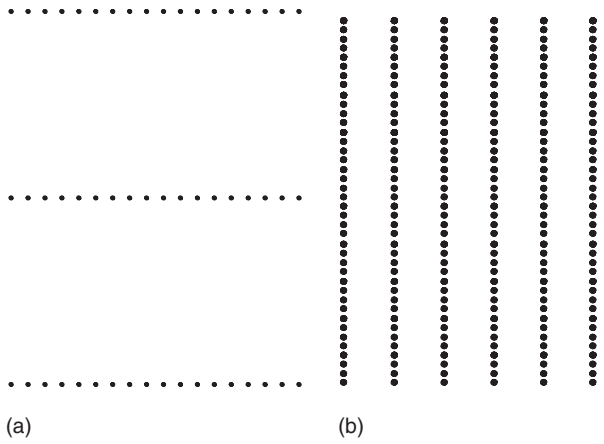


Figure 17 Isolated layers in a solid in the limit where the interlayer spacing becomes very large in direct space (a). The corresponding reciprocal lattice is shown in (b).

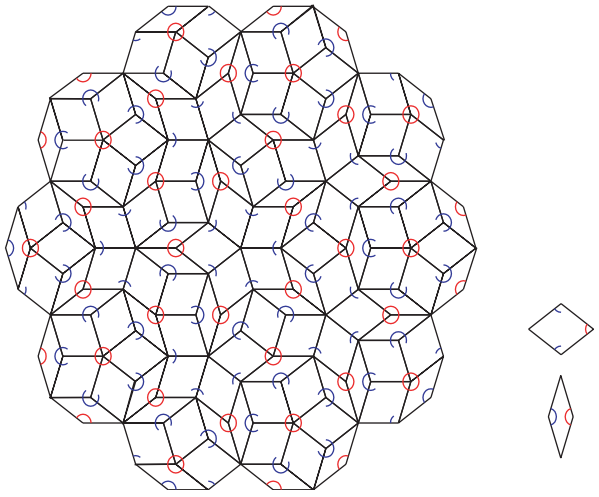


Figure 18 Portion of a Penrose tiling made with the two decorated building blocks shown to the right. The large and small building blocks are rhombuses with angles of 72° and 36°. The tiling was constructed subject to the rule that the decorations (colored arcs) on adjoining blocks must match.

diffraction features in scattering from the surface of a perfectly terminated (atomically-flat) crystal. However, in contrast to the case of the two-dimensional layer, the intensity of the streaks falls off quickly.

Finally, consider quasicrystalline ordering, where a solid is assembled according to a prescription which is ordered, but not periodic. A one-dimensional example is a Fibonacci lattice, where two building blocks A and B are repeated in the sequence

A BA ABA BAABA ABABAABA…

where the spaces are included for clarity. Each succeeding cluster of blocks is the sum of the two previous clusters. In practice, such a structure can be assembled by the sequential deposition of layers of two distinct materials. The structure is not periodic, since the pattern does not repeat, but it is ordered and it can be shown that it gives rise to a characteristic diffraction pattern with sharp peaks.

A two-dimensional example of a quasicrystal is a Penrose tiling, made up of two building blocks as shown in **Figure 18**. A fivefold symmetry does not occur among the periodic two-dimensional lattices (**Figure 4**), but the Penrose tiling of **Figure 18** has a fivefold symmetry about its central point. Like the Fibonacci lattice, the Penrose tiling is aperiodic, in that there is no repeating unit cell, yet it is ordered, as is clear from the perfect symmetry of **Figure 18**.

Quasiperiodic lattices are more than an interesting exercise in mathematical crystallography, since in the 1980s it was discovered that certain metallic alloys crystallize with icosahedral symmetry. These alloys yield diffraction patterns with a fivefold symmetry, which is forbidden for an untwinned crystal with translational symmetry. The icosahedral alloys are thought to be three-dimensional quasicrystals with an ordered aperiodic structure, formed by the optimization of near-neighbor packing during solidification. Although the actual physical structures are aperiodic, they can be represented as projections into three-dimensional space of periodic structures in a higher-dimensional space. Therefore, an understanding of the symmetry properties of periodic structures turns out to be very useful, even when the physical system is not periodic.

See also: Crystal Structure; Electron Gas (Theory); Electronic Structure (Theory): Atoms; Epitaxy; Liquid Crystals; Meso- and Nanostructures; Micro- and Macrostructures; Point Groups; Quasicrystals; Quasicrystals, Electronic Structure of; Space Groups; Structure Types and Classifications.

PACS: **61.50.Ah; 61.66. – f; 61.10.Nz**

Further Reading

Als-Nielsen J and McMorrow D (2001) *Elements of Modern X-Ray Physics*. New York: Wiley.

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Saunders College.

Hammond C (1997) *The Basics of Crystallography and Diffraction*. Oxford: Oxford University Press.

Henry NFM and Lonsdale K (eds.) (1965) *International Tables for X-Ray Crystallography, Vol. 1: Symmetry Groups*, 2nd edn. Birmingham: The Kynoch Press.

Jaric MV (ed.) (1988) *Introduction to Quasicrystals*. Boston: Academic Press.

Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.

Phase Diagram *See* Conductivity, Electrical; Conductivity, Thermal; Low-Temperature Properties of Materials; Mass Transport; Onsager Relations; Quantum Hall Effect; Radioactive Tracers; Thermal Conductivity at Low Temperatures; Thermoelectric Effect.

Phase Transformation

J H Perepezko, University of Wisconsin, Madison, WI, USA

Introduction

A central focus of phase transformation thermodynamics is the examination of relative phase stability. Under most conditions, phase stability is considered in terms of the response of a material to changes in temperature, pressure, composition, and size scale. The changes are often applied rapidly in order to establish undercoolings, overpressures, and supersaturations that place the initial stable phase in a metastable (or energized) condition. The relaxation of a metastable state to the stable state can proceed through a succession of various other metastable states, but the rate of the relaxation processes is usually limited by diffusional kinetics for thermally activated reactions. In general, it is often worthy to also identify the rate-limiting stage of a phase transformation reaction as either the nucleation or the growth stage, since the thermodynamic analysis for each stage can be different.

The essential thermodynamics that governs the relative phase stability for phase transformation reactions can be represented in analytical expressions for the Gibbs free energy as a function of temperature, pressure, composition, size scale, and other appropriate variables. These relationships can also be presented in a graphical form that facilitates the interpretation and visualization of the free energy changes, the metastable phase options, and various reaction pathways that often occur in kinetic competition.

A useful concept for the study of phase transformation processes is the distinction between closed and open system processes. In a closed system, an energized state is achieved through a rapid temperature, pressure, or composition change to create a certain level of undercooling or supersaturation (i.e., a metastable state), which then releases the excess free energy during relaxation toward the equilibrium stable state. With an open system, the energized state is often attained by a continuous incremental excess energy input to an initial state through the incorporation of excess lattice defects or solute on a localized spatial scale and time interval that is short compared to the relaxation time so that the relaxation process can be influenced by external factors. Regardless of the type of phase transformation or synthesis method, there are some common themes in the phase stability. The fundamental issues include the phase selection during the initial synthesis and the phase stability after synthesis and during subsequent treatment.

In order to treat the basic issues in phase stability within the limited coverage that is available, some of the key points are examined by considering selected thermodynamic relations for macroscopic systems and their modification for high interfacial area nanoscale systems. A central issue is the free-energy change (i.e., driving free energy) associated with a phase transformation reaction. As a means of illustrating the use of the relevant thermodynamic concepts, some examples are considered for the evaluation of the driving free energy and the ranking of different levels of metastability.

Phase Stability Thermodynamics

Equilibrium

There are a number of essentially equivalent approaches to describe thermodynamic equilibrium. For a closed system, the general conditions for equilibrium are represented by a maximum in entropy, S, or a minimum in free energy, G, with respect to a system variable z, where z is a variable such as pressure P, temperature T, or volume V as given by

$$\frac{dS}{dz} = 0 \quad \text{and} \quad \frac{d^2S}{dz^2} \leqslant 0 \qquad [1]$$

or

$$\frac{dG}{dz} = 0 \quad \text{and} \quad \frac{d^2G}{dz^2} \geqslant 0 \qquad [2]$$

These conditions prevail either for local minima/maxima that represents a metastable equilibrium or for an absolute minima/maxima that holds for a stable equilibrium as illustrated in **Figure 1**.

Pure Components

In the analysis of relative phase stability and the driving forces for different phase transformation reactions, the Gibbs free energy G is the main quantity of interest. By definition, $G = H - TS = E + PV$, where H is the enthalpy, E is the internal energy, and PV represents the system work. For reactions between two phases, the free energy change is the difference, $\Delta G = \Delta H - T\Delta S$. For example, during solidification of a liquid, $G_s - G_1 = (H_s - H_1) - T(S_s - S_1)$ or $\Delta G_f = H_f - T\Delta S_f$. At equilibrium, $\Delta G_f = 0$ so that $\Delta S_f = \Delta H_f / T_m$. If ΔS_f is taken as

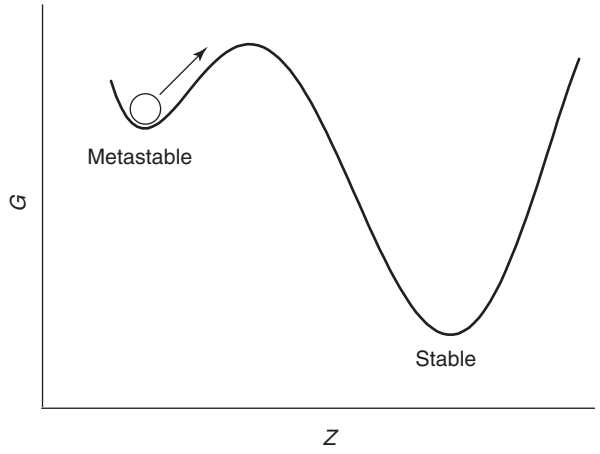


Figure 1 Schematic comparison of metastable (i.e., local) and stable (i.e., absolute) equilibrium.

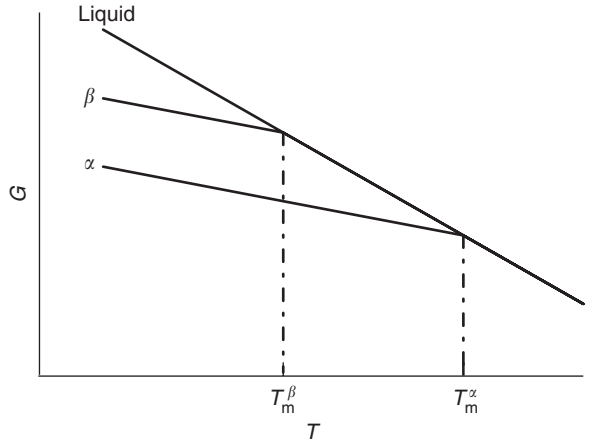


Figure 2 Molar free energy as a function of temperature for a single component. The melting point of the stable α-phase $T_m\alpha$ and the metastable β-phase $T_m\beta$ is indicated.

independent of T (i.e., neglecting the small heat capacity correction), then $\Delta G_f = \Delta H_f[1 - (T/T_m)] = \Delta H_f(T_m - T)/T_m = \Delta H_f \Delta T/T_m$, so that the driving free energy (i.e., driving force) available to accomplish the reaction is proportional to the undercooling, ΔT. The dependence of the molar free energy on temperature is illustrated in **Figure 2** for a liquid, a stable α-phase and a metastable β-phase, where it is evident that the formation of the metastable β-phase requires a minimum liquid undercooling below the stable α-phase melting point of $\Delta T = T_m^\alpha - T_m^\beta$.

From the second law of thermodynamics, several differential relations can be developed (i.e., Maxwell relations). The relation for the Gibbs free energy for a single phase is

$$dG = VdP - SdT \qquad [3]$$

For the liquid–solid case at equilibrium, $dG = 0$ and

$$V_s dP - S_s dT = V_1 dP - S_s dT \qquad [4]$$

or

$$\left(\frac{dP}{dT}\right)_{eq} = \frac{S_s - S_1}{V_s - V_1} = \frac{\Delta S_f}{\Delta V_f} = \frac{\Delta H_f}{\Delta V_f T_m} \qquad [5]$$

which is the Clapeyron equation for the pressure dependence of the melting point. When eqn [3] is applied to other two-phase coexistence such as solid–liquid, solid–vapor, and liquid–vapor equilibrium, the pressure–temperature phase diagram for a single component is developed as shown in **Figure 3**. The solid lines represent the two-phase coexistence conditions, where there is one degree of freedom according to the phase rule (i.e., $p + f = c + 2$, where

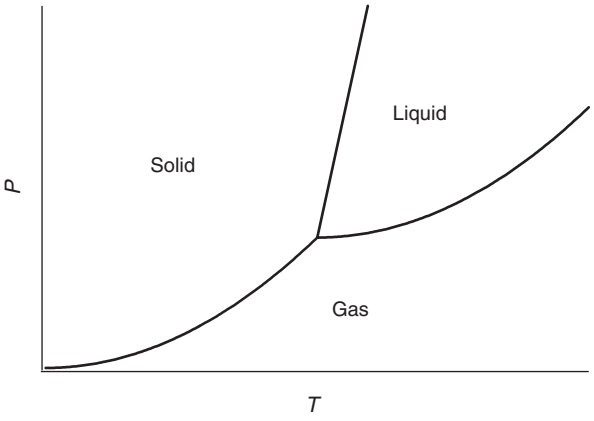


Figure 3 A schematic pressure vs. temperature diagram for a single component.

p is the number of phases, f is the number of degrees of freedom, and c is the number of components). The lines intersect at an invariant triple point where $f = 0$. It is worth noting that $(\partial G/\partial T)_P = -S$ and $(\partial G/\partial P)_T = V$ as illustrated in **Figure 2**.

Classification of Phase Transformations

The pure component reactions that are depicted in **Figure 3** are often classified as first-order phase transformations. At the equilibrium transformation temperature T_e, $G^\alpha = G^\beta$, but the slopes are different (i.e., $dG^\alpha/dT \neq dG^\beta/dT$). This gives rise to a discontinuity in S, V, and H at T_e and a singularity in heat capacity $C_p = T(dS/dT) = T(d^2G/dT^2)$ as illustrated in **Figure 4**.

During a second-order phase transformation, S, V, and H are continuous at the transformation temperature or critical temperature, T_c (i.e., there is no

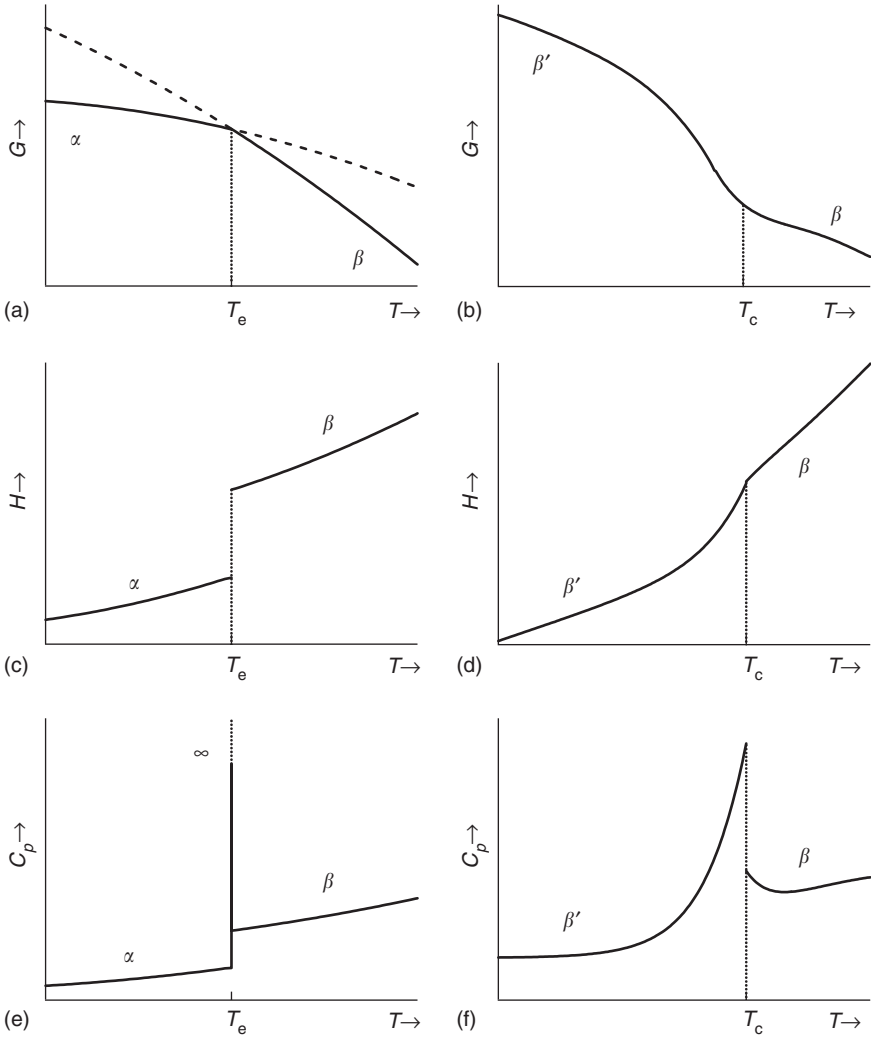


Figure 4 The thermodynamic characteristics of (a, c, and e) first-order and (b, d, and f) second-order phase transformations.

latent heat), but there is a discontinuity in heat capacity as indicated in **Figure 4**. Second-order phase transformations are often associated with electronic effects such as the paramagnetic/ferromagnetic transition, but they are also observed for certain order–disorder changes in alloys.

The distinction between first- and second-order phase transformations also has an impact on the presentation of the equilibria in a phase diagram and the progress of the phase transformation. For example, long-range ordering of unlike atoms in a crystal lattice develops due to an attractive interaction between unlike atoms. During a first-order reaction, the development of an ordered atomic arrangement initiates at separate regions as the temperature falls below T_e as shown in **Figure 5**. A sharp interface separates the transformed (i.e., ordered) regions and the untransformed (i.e., disordered) regions. Across this interface, there is a jump in the long-range order parameter, η that varies from zero for a random arrangement to unity for a perfectly ordered arrangement. The coexistence at T_e is represented on the phase diagram as a two-phase field between the ordered and the disordered phase. In contrast, during a second-order reaction, the development of an ordered arrangement initiates throughout the volume at the critical temperature T_c and progressively develops with increasing values of η as the temperature is lowered below T_c. There is no two-phase coexistence of an ordered and a disordered phase, so that the phase diagram does not display a two-phase field. Instead, the phase diagram shows the locus of T_c as a function of composition.

Alloy Solutions

For alloy solutions, the chemical terms for the free energy are included by adding the partial molar free

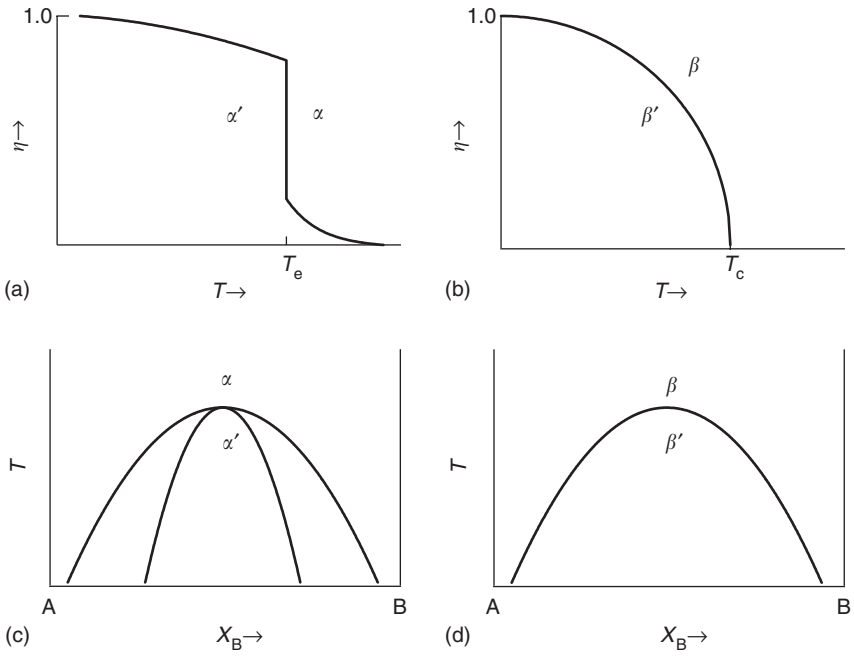


Figure 5 Schematic illustration of the temperature dependence of η (a and b) and the phase diagram characteristics (c and d) for (a and c) first-order and (b and d) second-order phase transformations.

energy for each component, i or j to eqn [3] as

$$dG = VdP - SdT + \left(\frac{\partial G}{\partial n_i}\right)_{T,P,nj} dn_i + \left(\frac{\partial G}{\partial n_j}\right)_{T,P,ni} dn_j \quad [6]$$

where n is the number of moles. The partial molar free energy represents the chemical potential, μ_i, that is defined as

$$\left(\frac{\partial G}{\partial n_i}\right)_{T,P,nj} = \mu_i = \mu_i^0 + RT \ln a_i \quad [7]$$

where μ_i^0 is the standard state chemical potential, R is the gas constant, and a_i is the activity which is the product of an activity coefficient γ_i and the mole fraction, X_i (note that for a binary solution $X_i + X_j = 1$). For alloy solution at constant T and P with a given composition X_0 in an A–B system, the molar free energy is

$$G = (1 - X_0)\mu_A + X_0\mu_B \quad [8]$$

and as indicated in **Figure 6**, eqn [8] describes the tangent to the G versus X_B curve at X_0. For reaction between two coexisting solution phases such as liquid and solid in an alloy with overall composition X_0

$$\Delta G = (1 - X_0)(\mu_A^s - \mu_A^1) + X_0(\mu_B^s - \mu_B^1) \quad [9]$$

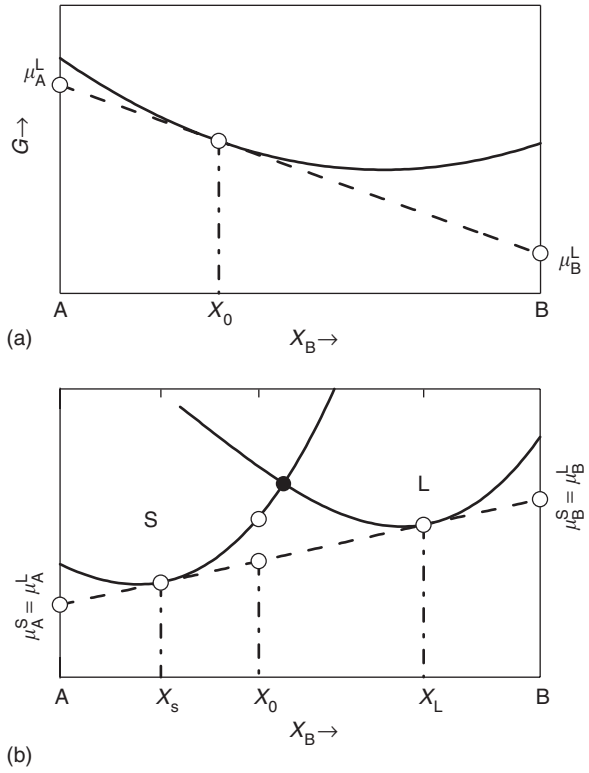


Figure 6 Schematic free energy vs. composition diagrams for a binary alloy of (a) a single-phase liquid solution with the component chemical potentials μ_A^L and μ_B^L determined by the pure component intersections with the tangent to the G vs. X_B curve at X_0 and of (b) a two-phase liquid–solid equilibrium with common tangent points at X_s and X_L.

At equilibrium, $\Delta G = 0$, so that $(\mu_A^s = \mu_A^1) = (\mu_B^s = \mu_B^1)$ and the two phases share a common tangent.

For the examination of interactions in solutions, it is often useful to represent the free energy change for mixing of pure components to form a solution ΔG_m in terms of an enthalpy of mixing ΔH_m and a mixing entropy ΔS_m. The simplest model is based upon an ideal solution where there are no nearest neighbor interactions, so that $\Delta H_m = 0$ and $\gamma_i = 1$ and the ideal entropy of mixing is given by

$$\Delta S_m = -R[(1 - X_B)\ln(1 - X_B) + X_B \ln X_B] \quad [10]$$

Most real solutions are not ideal. When the departure from ideal behavior is modest so that the molar volume change upon mixing to form a solution is negligible, the enthalpy of mixing can be represented in terms of the nearest neighbor pairwise interaction energies E_{AA}, E_{BB}, and E_{AB} as

$$\Delta H_m = ZN_a X_A X_B[E_{AB} - 0.5(E_{AA} + E_{BB})] \quad [11]$$

or

$$\Delta H_m = \Omega X_A X_B \quad [12]$$

where Z is the lattice coordination number, N_a is Avogadro's number and $\Omega = ZN_a[E_{AB} - 0.5(E_{AA} + E_{BB})]$ is a constant for a regular solution. The entropy of mixing is given by the ideal solution expression eqn [8] to yield

$$\Delta G_m = \Delta H_m - T\Delta S_m = \Omega X_A X_B + RT \times [(1 - X_B)\ln(1 - X_B) + X_B \ln X_B] \quad [13]$$

The composition dependence of ΔG_m is shown in **Figure 7** for different values of Ω and temperature. For exothermic solutions, $\Omega < 0$ and mixing results in a free energy decrease at all temperatures. In this case, the preference for unlike atom pairings results in a tendency for the development of ordering in the solid solution. When $\Omega > 0$, however, the behavior changes with temperature. At high temperatures, $T\Delta S_m$ is greater than ΔH_m for all compositions and the free energy curve has a positive curvature at all points. At low temperatures, however, $T\Delta S_m$ is smaller than ΔH_m for the central composition range and ΔG_m displays a change in curvature with a negative curvature near the pure component compositions that changes to a positive curvature (i.e., unstable solutions) in the central composition range. The positive value of Ω indicates a preference for pairings of like atoms and results in a tendency for the development of clustering in the solid solution.

The composition dependence of ΔG_m yields a concave curve when $\Delta H_m < 0$ as illustrated in **Figure 6a**. In addition, in **Figure 6b**, the equilibrium condition of the equality of the chemical potential for each component in the coexisting phases is illustrated by the common tangent construction. The various phase equilibria that are defined by a collection of free energy versus composition diagrams at different temperatures are the basis of the equilibrium phase diagram. As demonstrated in the schematic diagram in **Figure 8**, the common tangent points on the free energy versus composition diagram define the tie lines between coexisting phases at each temperature on the phase diagram. Within the two-phase regions defined by the common tangents, the free energy curves intersect. The point of intersection defines the limit to partitionless transformation. For liquid compositions to the left of the intersection (i.e., A-rich compositions), there is a free energy decrease for a composition invariant or partitionless solidification. For liquid compositions to the right of the intersection, a partitionless solidification is impossible since the free energy would increase. The locus of the intersection points as a function of temperature is often called a T_0 curve.

When $\Delta H_m > 0$, the composition dependence of ΔG_m displays a change in curvature as noted in **Figure 7d** and illustrated in more detail in **Figure 9**. As shown in **Figure 9**, the phase diagram for a solution with $\Delta H_m > 0$ exhibits a miscibility gap at low temperature where a homogeneous solution, α, decomposes into two solutions α_1 and α_2 without overcoming an activation energy barrier. Due to the change in curvature of the G versus X_B curve, there is a fundamental change in the stability of solutions within the miscibility gap and a resultant change in the mechanism of the decomposition. For example, solutions with compositions between X_1 and the turning point (i.e., where $d^2G/dX_B^2 = 0$) or between the turning point and X_2 are metastable, since $d^2G/dX_B^2 > 0$. This signifies that fluctuations in composition that are necessary to accomplish the decomposition reaction result in an increase in free energy that represents a barrier to the reaction. For these solutions, a nucleation process is required to initiate the decomposition reaction. In contrast, solutions with compositions between the turning points are unstable, since $d^2G/dX_B^2 < 0$. In this case, fluctuations in composition are spontaneous, since they lead to a reduction in free energy. The demarcation between metastable and unstable solutions within the miscibility gap is represented by the spinodal curve that refers to the locus of temperatures and compositions, where $d^2G/dX_B^2 = 0$.

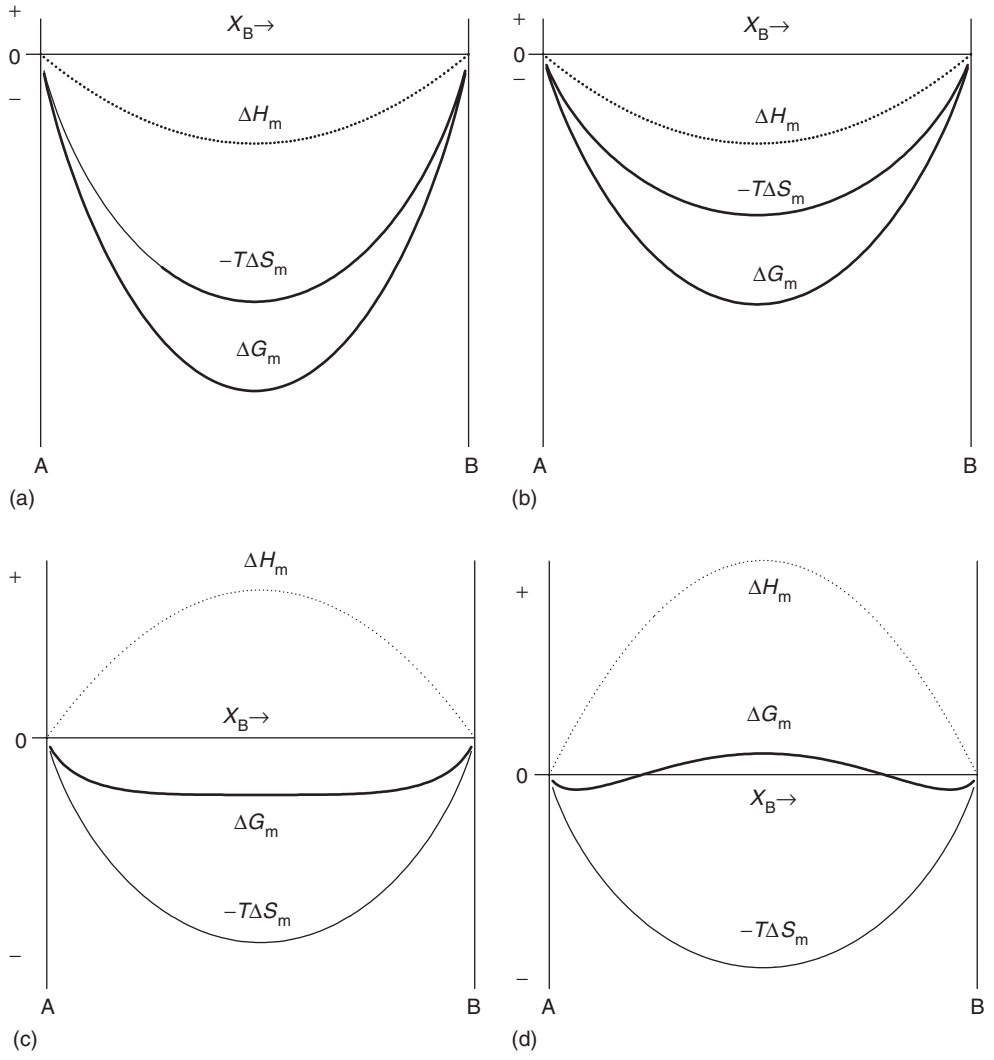


Figure 7 The influence of temperature and ΔH_m on ΔG_m when (a) $\Omega < 0$, T high; (b) $\Omega < 0$, T low; (c) $\Omega > 0$, T high; and (d) $\Omega > 0$, T low.

Phase Transformation Energetics

An important application of the thermodynamic description of phase stability is the representation of the free energy change that accompanies a phase transformation (i.e., the driving free energy). A change of phase requires a departure from equilibrium as the system develops a supersaturation or an undercooling that represents the driving free energy as indicated in **Figure 10**. For a dilute solution where the activity is represented by the mole fraction, the free energy change associated with a precipitation reaction where the α-phase composition changes from X_0 to X_α as pure β develops at temperature T_1 is given by

$$\Delta G = RT_1 \ln\left(\frac{X_0}{X_\alpha}\right) \qquad [14]$$

Since the solvus phase boundary composition, X_s, is represented as

$$X_S = \exp(\Delta S_S)\exp\left(-\frac{\Delta H_S}{RT}\right) \qquad [15]$$

where ΔS_s is the entropy of solution and ΔH_s is the enthalpy of solution, the free energy change for precipitation is expressed as

$$\Delta G = \frac{\Delta H_S \Delta T}{T_S} \qquad [16]$$

demonstrating again that a supersaturation can be expressed in terms of an equivalent undercooling. It is to be noted that the expressions in eqns [12] or [14] refer to the overall ΔG for the reaction, $\alpha(X_0) \rightarrow \alpha(X_\alpha) + \beta(X_\beta)$. During the initial nucleation

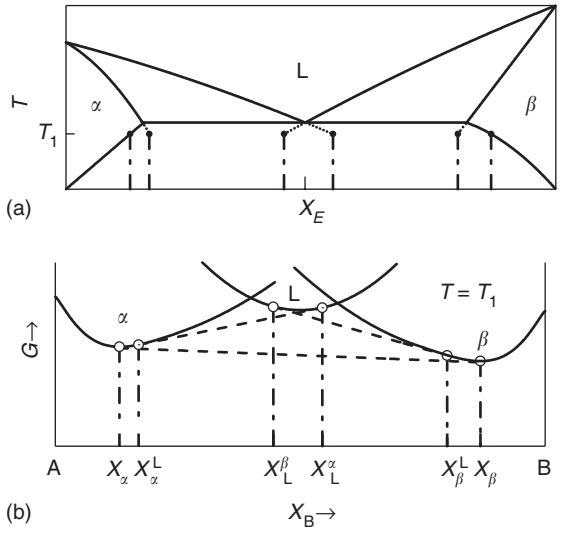


Figure 8 (a) Schematic phase diagram and (b) free energy vs. composition diagram for a binary alloy that exhibits a eutectic reaction: $L \rightarrow \alpha + \beta$. The relationship between the phase boundaries on the phase diagram and the common tangents is illustrated for a temperature T_1, where both a stable $\alpha + \beta$ and the metastable $L + \alpha$ and $L + \beta$ two-phase equilibria could occur for different kinetic conditions. The location of the T_0 points are indicated by filled circles.

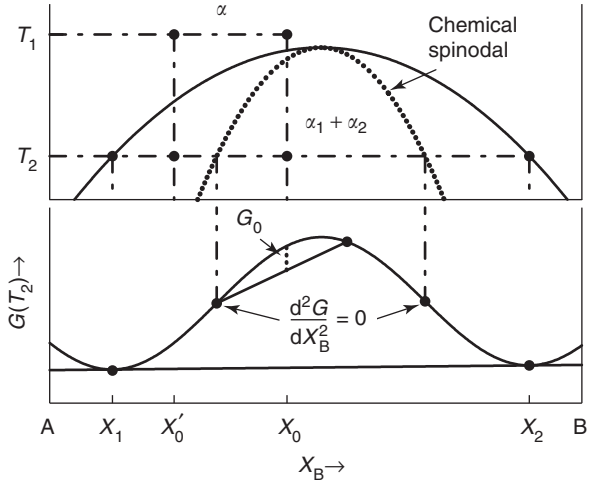


Figure 9 (a) Schematic phase diagram and (b) free energy vs. composition diagram for alloys between the spinodal points, which are unstable and can decompose into two coherent phases α_1 and α_2 without overcoming an activation energy barrier. Alloys between the miscibility gap and the spinodal are metastable and can decompose only after nucleation of the other phase.

stage when a small quantity of β forms, the α-matrix composition is not changed much from the X_0 value. In this case, the maximum driving free energy for nucleation of β, ΔG_N, is given by the separation between the tangent to the G_α curve at X_0 and a parallel tangent to the G_β curve as illustrated in **Figure 10**. The tangent to the G_β curve occurs at the

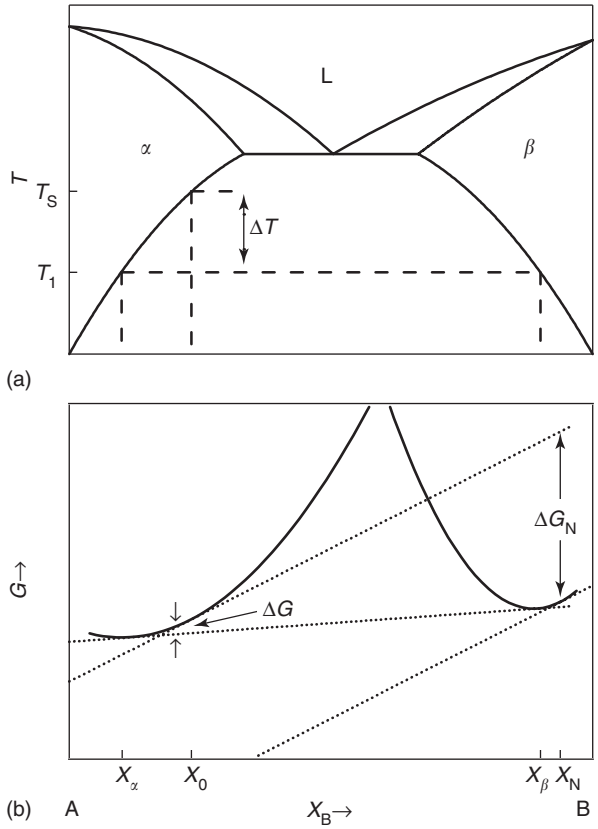


Figure 10 (a) Schematic phase diagram and (b) illustration for the development of a supersaturation $[X_0/X_\alpha]$ as an α- phase alloy of composition X_0 is undercooled by an amount ΔT below the α- phase solvus boundary. During the initial formation of a β-phase precipitate, the maximum driving free energy for nucleation, ΔG_N is given by the parallel tangent construction.

composition X_N, which is the favored composition for a β-nucleus (i.e., the composition with the largest driving free energy).

Similarly, for a phase transformation involving the decomposition of an initial phase such as the liquid with composition X_E in **Figure 8** into two phases such as α of composition X_α and β of composition X_β, the ΔG for reaction can be evaluated from eqn [8]. The free energy change is given by the separation between the tangent to the G_L curve at X_E and the common tangent to the G_α and G_β curves as

$$\Delta G = (\mu_\alpha^B - \mu_L^B)X_E + (\mu_\alpha^A - \mu_L^A)(1 - X_E) \quad [17]$$

Phase Stability Hierarchy

Nucleation-Controlled Reactions

Throughout the analysis of transformation behavior, it is commonly recognized that nucleation control is an important part of the initial stage of a reaction. Under nucleation control, a strong temperature

dependence of the product phase selection and number density is developed, which leads to a nanoscale microstructure. Nucleation limitations in diffusion reactions, especially those involved in reactive diffusion, also represent a form of nucleation control that is important in nanostructure synthesis. In terms of alloy metastability, nucleation control is important in allowing access to metastable states. In effect, the observation of nucleation control is directly related to the factors that promote the development of large undercooling or supersaturation which enables the expression of kinetic transitions through competitive phase selection resulting in the refinement of the product size to the nanoscale level.

Competitive Phase Selection and Metastability

A common theme in the development of new microstructural options by phase transformations is the occurrence of metastable structural states. Often, the reactions that occur during the freezing of undercooled liquids or during other rapid transformations are viewed as nonequilibrium processes. However, it is also evident that some of the departures from full equilibrium can be considered in terms of different levels of metastability. In fact, a hierarchy of equilibrium is identified based upon the severity of the kinetic constraints that affect the capability of a material to relax toward full equilibrium during processing. As the rate of reaction becomes faster, kinetic constraints that arise from nucleation and growth limitations associated with an equilibrium product phase formation develop and expose alloy metastability that often coincides with nanostructure synthesis. For the suppression of the equilibrium phase or the formation of a kinetically favored metastable phase, it is still possible to analyze reactions in terms of a metastable equilibrium that is used locally at interfaces. The transition from stable to metastable equilibrium is illustrated in **Figure 11** where the kinetic suppression of an equilibrium γ-phase (see **Figure 11a**) yields a metastable eutectic involving the α- and β-phases. Moreover, it is expected that the application of the appropriate local equilibrium is used when the processing involves nanoscale structures. Under extreme conditions, loss of interfacial equilibrium for either a stable or metastable phase is developed even when interfacial

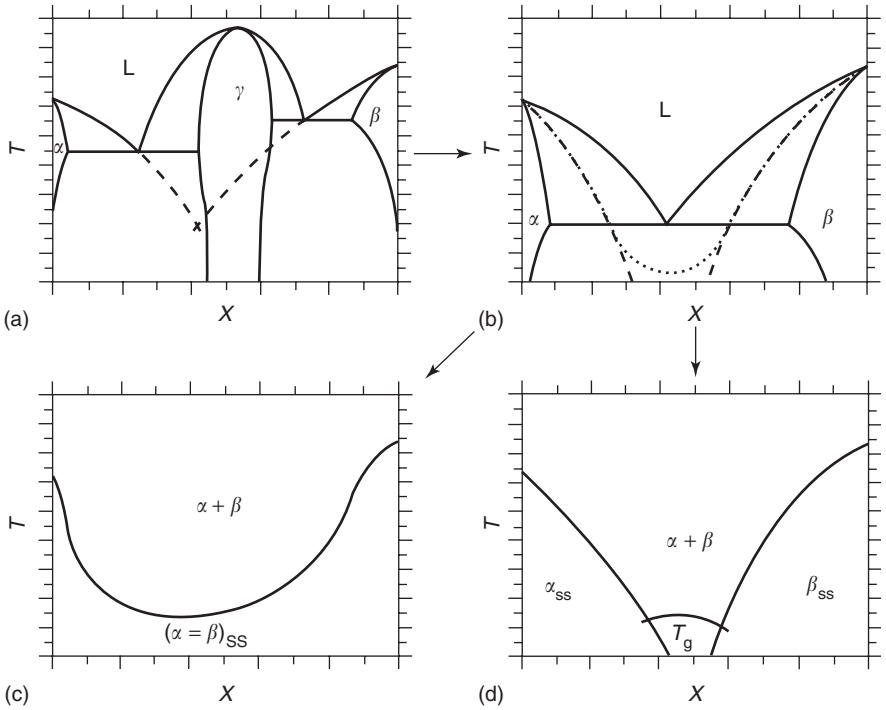


Figure 11 Schematic illustration of temperature vs. composition diagram for some of the levels in the hierarchy of equilibrium. (a) The equilibrium phase diagram of a system with an intermediate phase; dashed lines represent the metastable extensions of the liquidus and solidus curves for the primary solution phases. Solidification under metastable equilibrium conditions can result in a bypassing of the intermediate phase to yield a metastable eutectic phase diagram between α and β. (b) The T_0 curves for the primary solutions. The extensions of these curves to temperatures below the metastable eutectic are indicated as dashed curves. (c) If the primary phases have different crystal structures and low mutual solubility, then the T_0 curves might not intersect. Such a situation favors glass formation in the composition range, where the T_g curve is greater than the T_0 curves. (d) For alloys based upon components with different structures each crystal phase has a T_0 curve.

relaxation becomes too slow. With the loss of interfacial equilibrium, thermodynamics can still be used to restrict the possible range of compositions that exists at an interface at various temperatures, since the selection must yield a net reduction in the free energy of the system. One way to represent the thermodynamic restrictions is based upon the application of T_0 curves which represent the locus of temperatures and compositions, where the free energies of two phases are equal (as illustrated in **Figure 6**) for liquid and solid phases, and thus define the limiting condition for partitionless transformation. For example, as interfacial equilibrium is lost, the liquidus and solidus boundaries in **Figure 11b** collapse to the T_0 curves. With isomorphous systems that exhibit complete solubility, the T_0 curve is continuous with composition (see **Figure 11c**) while for alloys based upon components with different structures, each crystal phase has a T_0 curve (see **Figure 11d**). At temperatures and compositions above the T_0 curves, solute partitioning is required for solidification (see **Figure 11d**). Because of the diffusional constraint due to solute partitioning, a crystallization reaction can be inhibited by quenching to promote glass formation. Within the overall hierarchy of stability, the systematic examination of the different levels of kinetic constraints provides useful insight into the thermodynamic analysis of alloy metastability and phase selection.

Constrained Equilibrium

There are a number of important cases, where phase transformation reactions proceed under various constraints that involve externally imposed conditions or internally generated limitations. A common example is the development of a metastable phase because of favorable nucleation and growth kinetics compared to the stable phase. This is a kinetically constrained equilibrium. Another kinetically constrained equilibrium can be developed in alloys, where one component exhibits a significantly higher diffusivity than the other components. This is commonly observed in alloy steels where the high diffusivity component is an interstitial component such as carbon and the other slow diffusivity components are substitutional species. The high diffusivity component is able to establish an equality of the chemical potential in the coexisting phases before the slow diffusivity components. This is called "paraequilibrium." Of course, since paraequilibrium is constrained kinetically, it

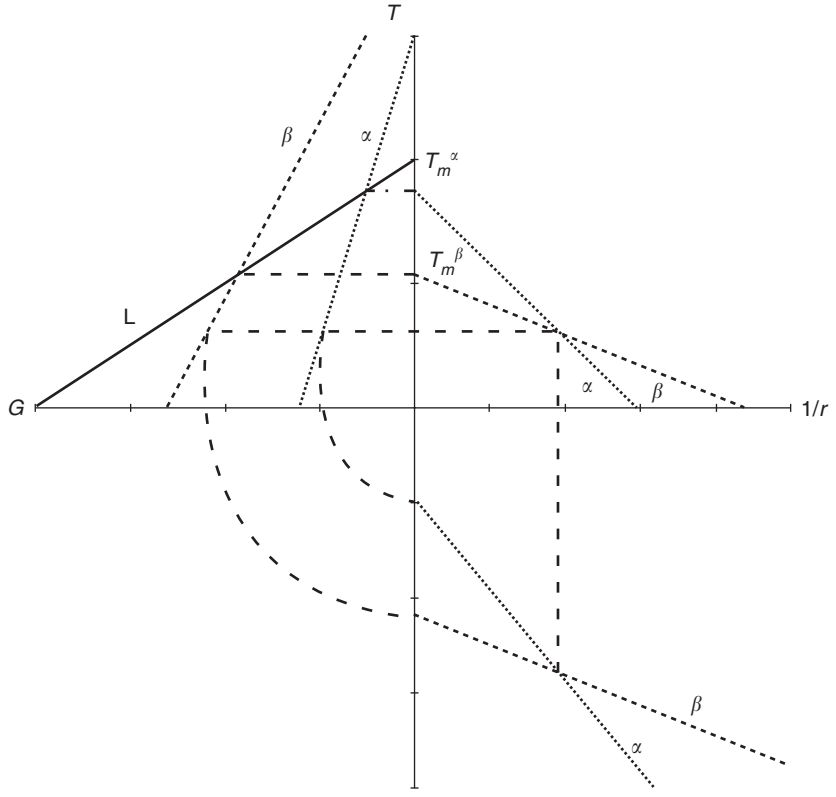


Figure 12 Schematic illustration of the influence of curvature (1/*r*) on the melting point of the α-phase and the β-phase and on the relative phase stability when $\sigma_{\alpha L} V_m / \Delta S_f^{\alpha} > \sigma_{\beta L} V_m / \Delta S_f^{\beta}$.

will dissipate with time as the alloy approaches the stable equilibrium.

During the initial stages of a phase transformation reaction in the solid state or during deposition from the vapor, it is often observed that the product phase exhibits an epitaxial crystallographic relationship with the transforming phase or the substrate involving the interfacial matching of lattice planes and directions. The appearance of a crystallographic matching results in a reduction of the interfacial energy that in turn yields favorable nucleation kinetics. However, the matching is rarely perfect so that there remains a residual interfacial strain. The resulting elastic strain energy represents another component to the free energy. With the progress of the reaction, the product phase grows and there is a loss of coherency and elastic strain energy.

The elastic strain energy is usually not partitioned equally between the transforming, and product phase so that the constraint of lattice matching across the interphase interface alters the phase equilibrium. This constrained equilibrium is designated as "coherent equilibrium." Because the level of elastic strain energy depends on the volume, the influence of coherent equilibrium is a function of composition in the two phase regions (i.e., relative amounts of phases). This yields new behavior that is beyond the scope of coverage in this discussion.

Thermodynamics of Nanoscale Systems

The nanoscale is often reported as a linear dimension but the important interfacial effects should be considered in terms of the interfacial area per unit volume (A/V). For example, for a sphere: $A/V = 3/r = 3 \times 10^7\ \mathrm{m}^{-1}$ for $r = 100$ nm. This is very significant.

Interfacial effects can be included in the free energy as

$$\mathrm{d}G = V\mathrm{d}P - S\mathrm{d}T + \sum \mu_i \mathrm{d}n_i + \sum \sigma_i \mathrm{d}A_i \qquad [18]$$

where σ_i is the interfacial energy and A_i is the interfacial area.

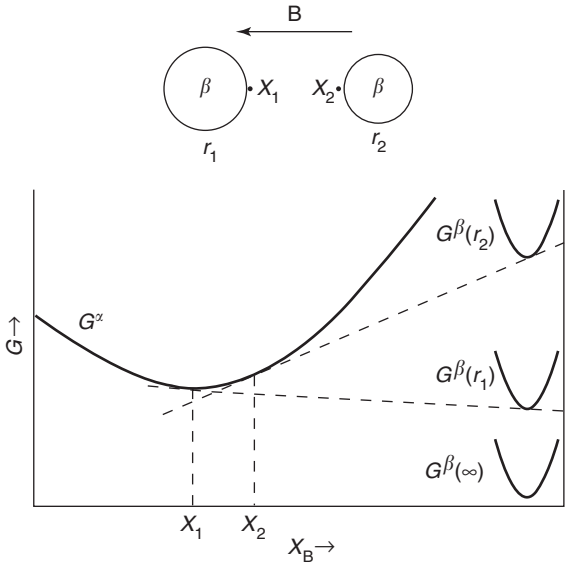


Figure 14 Schematic free energy vs. composition diagram demonstrating the influence of interface curvature of a β-phase particle on the solubility.

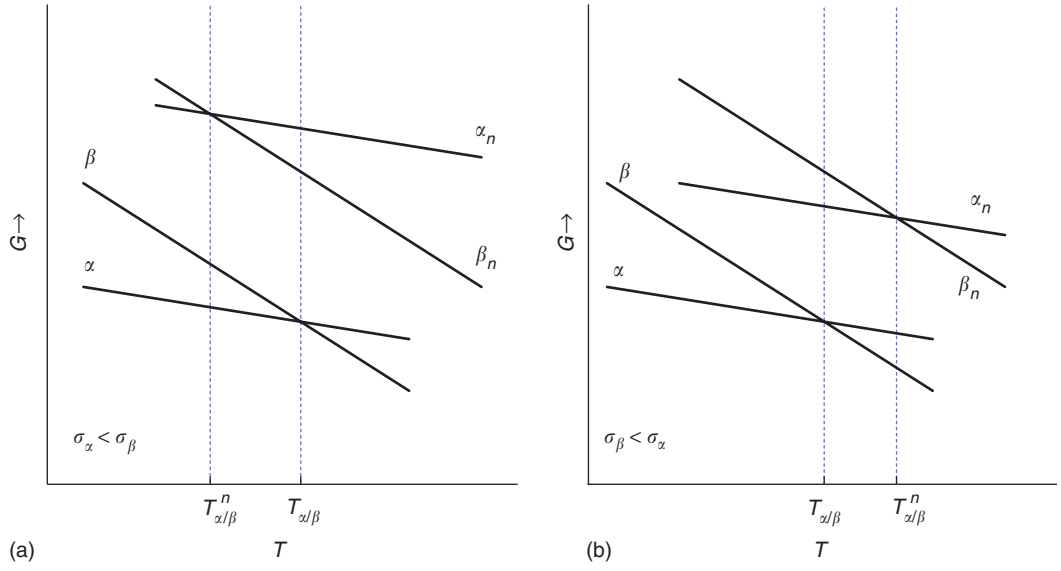


Figure 13 Schematic Illustration of the modification of the stability of the α-phase at the nanoscale. (a) when $\sigma_\alpha < \sigma_\beta$, there is an enhancement of α-phase stability and (b) when $\sigma_\beta < \sigma_\alpha$, there is a reversal of phase stability between the α-phase and the β-phase.

Table 1 Comparison of the driving free energies for some phase transformation reactions

Reaction process	*Free energy*	*Typical value*[a] ($J\,mol^{-1}$)	*Remarks*
Crystallization polymorphic transition	$\Delta H_f \Delta T/T_m$	3×10^3	
Mixing/Interdiffusion	$RT(X_A \ln X_A + X_B \ln X_b)$	5×10^3	Ideal solution behavior
Oxidation	$\Delta G^0 = -RT \ln K$	$5 \times 10^4 - 5 \times 10^6$	ΔG^0 – Formation of oxide K – Equilibrium constant
Sublimation/condensation	$\Delta H_v\ \Delta T/T_s$	10^4–10^5	ΔH_v – Sublimation enthalpy T_s – Sublimation temperature
Grain growth, coarsening	$2\sigma/r$	20 for $r = 1\,\mu m$ 2×10^3 for $r = 10\,nm$	$\sigma = 1\,J\,m^{-2}$
Precipitation	$RT\ln(X_\alpha/X_0)$	10^4	$X_\alpha/X_0 = 10$
Cold work (stored energy)	$\rho G\boldsymbol{b}^2$	10^2–10^3	G – Shear modulus b – Burgers vector ρ – Shear modulus

[a]Typical value at $T = 1000\,K$.

The increment in free energy due to interfaces is represented by the Gibbs–Thomson relation in terms of the interface curvature κ that is determined by the principal radii, r_i , as

$$\Delta G = \kappa\sigma V_m = (r_1^{-1} + r_2^{-1})\sigma V_m \qquad [19]$$

where V_m is the molar volume.

For a sphere, $\Delta G = 2\sigma V_m/r$. The effect of curvature on the melting point and the relative phase stability is indicated in **Figure 12** where the dependence of melting point on curvature can be expressed from eqns [3] and [13] as

$$T_m(r) = T_m(\infty) - \frac{2\sigma_{SL} V_m}{r\Delta S_f} \qquad [20]$$

There are two key points that are illustrated in **Figure 12**. First, the macroscopic melting points for the α- and β-phases, that is, T_m^α and T_m^β, decrease with increasing curvature. Second, the β-phase that is metastable for macroscopic sizes becomes the stable phase with respect to α-phase at high curvature values. Furthermore, the phase stability relations for a macroscopic scale system can be modified in different ways depending on the relative magnitude of σ for each phase for a nanostructured system. As illustrated in **Figure 13**, when $\sigma_\alpha < \sigma_\beta$, the stability of α is enhanced at the nanoscale. However, when $\sigma_\beta < \sigma_\alpha$, there is a reversal of phase stability at the nanoscale. There are a number of examples of the reversal in relative phase stability when the size scale changes from macroscopic to nanostructured.

The excess free energy due to curvature that acts to control the relative phase stability in nanoscale materials also acts to modify other thermodynamic behavior. For example, with the same analysis that was applied to express the curvature dependence of the melting point, the dependence of the solubility may be represented in terms of particle size. This behavior is illustrated in **Figure 14**, where it is evident that the increase in chemical potential with decreasing particle size will act to drive a diffusive flux between neighboring particles in a particle-size distribution such as that produced by a precipitation reaction. An important consequence of the interparticle transport is the dissolution of the finest particles along with a concomitant increase in the average particle size in the distribution. This coarsening behavior is central to the kinetic stability of a nanoscale microstructure.

A useful perspective on the different phase transformation reactions can be developed by considering the relative magnitudes of the driving free energies associated with each of the reactions. A partial summary of the representative magnitudes for the driving free energies for a selection of reactions is presented in **Table 1**. The listings in **Table 1** indicate that for typical processes involving chemical changes or phase transitions, the driving free energies range from a few to $100\,kJ\,mol^{-1}$. While the excess free energy associated with a nanoscale microstructure is only a few $kJ\,mol^{-1}$, this level is sufficient to alter the relative phase stability and to modify the path of kinetic reactions.

See also: Allotropy and Polymorphism; Alloys: Aluminum; Alloys: Copper; Alloys: Iron; Alloys: Magnesium; Alloys: Overview; Alloys: Titanium; Ceramic Materials; Diffusionless Transformations; Glasses; Incommensurate Phases; Irreversible Thermodynamics and Basic Transport Theory in Solids; Liquid Crystals; Phase Transformations, Mathematical Aspects of; Phases and Phase Equilibrium; Specific Heat; Thermodynamic Properties, General.

PACS: 64.60. – i; 64.70. – p; 64.75. + g; 81.30. – t; 82.60.Lf; 82.60.Qr

Further Reading

Boettinger WJ and Perepezko JH (1993) Fundamentals of solidification at high rates. In: Liebermann HH (ed.) *Rapidly Solidified Alloys: Processes, Structures, Properties, Applications*, pp. 17–78. New York: Decker.

Cahn JW (1961) Spinodal decomposition. *Acta Metallurgica* 9: 795.

Cahn JW (1968) On spinodal decomposition. *Transactions of the AIME* 242: 166.

Christian JW (1975) *The Theory of Transformations in Metals and Alloys*, 2nd edn. Oxford: Pergamon.

Garvie RC (1978) Stabilization of the tetragonal structure in zirconia microcrystals. *Journal of Physical Chemistry* 82: 218–224.

Gaskell DR (1995) *Introduction to the Thermodynamics of Materials*, 3rd edn. London: Taylor & Francis.

Hillert M (1998) *Phase Equilibria, Phase Diagrams and Phase Transformations: Their Thermodynamic Basis*. Cambridge: Cambridge University Press.

Johnson WC (1999) Influence of elastic stress on phase transformations. In: Aaronson HI (ed.) *Lectures on the Theory of Phase Transformations*, pp. 35–134. Warrendale PA: TMS.

Lupis CHP (1983) *Chemical Thermodynamics of Materials*. Oxford: Elsevier.

Porter DA and Easterling KE (1992) *Phase Transformations in Metals and Alloys*. New York: Chapman and Hall.

Soffa WA and Laughlin DE (1989) Decomposition and ordering processes involving thermodynamically first-order order→disorder transformations. *Acta Metallurgica* 37(11): 3019–3028.

Trivedi RK (1999) Theory of capillarity. In: Aaronson HI (ed.) *Lectures on the Theory of Phase Transformations*, pp. 135–165. Warrendale, PA: The Metals, Minerals and Materials Society.

Phase Transformations, Mathematical Aspects of

J-C Tolédano, Ecole Polytechnique, Palaiseau, France

Introduction

Phase transformations occur in a large variety of systems subjected to a change in temperature or pressure, for example, fluids, liquid crystals, crystalline or noncrystalline solids, magnetic compounds, superconductors, and quasicrystals. Their existence is often revealed by subtle effects, which are characteristic of the system considered: progressive or sudden onset of magnetization, small mechanical shear, an optical rotatory power, additional X-ray or neutron scattering directions of diffraction, etc.

Except within simplified microscopic models of physical systems, the occurrence of a phase transformation at a given temperature T_c is taken as an experimental fact and as a starting point of the theory. The objective is then to relate the atomic-scale configuration of the system to the laws governing the more or less abrupt variations of physical properties, which are observed experimentally. The standard framework of these so-called phenomenological theories corresponds to the situation of continuous phase transformations, that is, transformations in which, although sharp anomalies of certain physical quantities are observed, the atomic (and magnetic) configuration undergoes infinitesimal modifications at T_c (or p_c) and changes in a continuous manner on either sides of the transformation point (in the following, the temperature T will be assumed to be the only relevant intensive thermodynamic quantity).

The fundamental mathematical quantity in a physical theory of continuous phase transformations is the order parameter. The concept of order parameter was introduced by L D Landau in 1935 in the first unified theory of phase transformations, which not only addresses the interpretation of physical phenomena out of the so-called critical range of temperatures of the transformation, but also is relevant to the more recent renormalization group theory of the critical phenomena. The next section introduces this concept and specifies its mathematical properties with reference to the mathematical theory of group representations. The following sections describe the implications of these properties for the prediction of symmetry breaking at a phase transformation, or for the form of the invariant polynomials involved in the expression of the free energy, or alternatively in the effective Hamiltonian density appearing in the phenomenological theories. Attention is also given to criteria excluding the continuous character of a phase transformation. For a discussion of other important aspects, for example, methods of analysis of the minima of the free energy, extensions of the theory to the cases of discontinuous and reconstructive transitions, and the prediction of domain/variant orientations resulting from the breaking of symmetry, the reader is referred to the Further reading section.

The Order Parameter and Its Mathematical Symmetry Properties

The introduction of an order parameter relies on the following remarks.

First, it is observed that the values of thermodynamical functions defined for a system, such as the free energy F, can be formally expressed, for any state of the system (equilibrium or nonequilibrium), as a function $F(x_i|T)$ of the various degrees of freedom x_i specifying an atomic (or magnetic) configuration (e.g., a set of values of the coordinates of all the atoms and their spin orientations), the temperature being an externally controlled parameter. At a given temperature T, the equilibrium values $x_i^{\mathrm{eq}}(T)$ of these degrees of freedom are determined by locating the absolute minimum of $F(x_i|T)$. One has

$$F_{\mathrm{eq}}(T) = F(x_i^{\mathrm{eq}}|T) = \min|x_i\ F(x_i|T) \qquad [1]$$

Hence $F(x_i|T)$ serves as a variational function to determine the thermodynamic equilibrium of the system.

Second, it is emphasized that, for a continuous phase transformation, $x_i^{\mathrm{eq}}(T)$ vary continuously across the transformation temperature T_{c}, and therefore, their values in the vicinity of T_{c} are close to their equilibrium values, $x_i^* = x_i^{\mathrm{eq}}(T_{\mathrm{c}})$. Consequently, one can restrict the study of $F(x_i|T)$ and of its minima to values of the x_i close to x_i^*. Moreover, if $F(x_i|T)$ is assumed to be a function without mathematical singularities at T_{c}, it can be expressed as a Taylor expansion of the small increments of the degrees of freedom $\delta x_i = x_i - x_i^*$:

$$\begin{aligned} F(x_i|T) = {} & F(\delta x_i|x_i^*, T) = F_0(x_i^*, T) \\ & + F_1(\delta x_i|x_i^*, T) + F_2(\delta x_i|x_i^*, T) \\ & + \cdots + F_p(\delta x_i|x_i^*, T) + \cdots \end{aligned} \qquad [2]$$

where F_p is a homogeneous polynomial of degree p in the increments δx_i with coefficients depending on the parameters x_i^* and T.

The mathematical formulation of the theory is based on two additional observations. The first concerns the invariance (i.e., the unchanged expression) of F when certain geometrical operations are applied to the degrees of freedom x_i. Indeed, the free energy being a function of the relative positions of the microscopic constituents of the system (or the relative orientation of the magnetic moments), and not of their absolute position in space, its expression $F(x_i|T)$ as a function of x_i must be invariant under any orthogonal transformation or translation acting globally on the system. However, if F is considered as a function of the increments δx_i, and can therefore be written as $F = F(\delta x_i|x_i^*, T)$, it will only be invariant under the restricted set of the former transformations which also preserve the parameters x_i^* implied in this function. This restricted set consists of the operations that leave unchanged the equilibrium atomic (and spin) configuration $x_i^* = x_i^{\mathrm{eq}}(T_{\mathrm{c}})$ of the system at the temperature T_{c}. These operations represent the symmetries of the considered configuration. This set forms a mathematical group $\mathscr{G}^{\bigstar}$. It can be shown that the $\mathscr{G}^{\bigstar}$-invariance of $F(\delta x_i|x_i^*, T)$ holds necessarily for each of the homogeneous polynomials $F_p(\delta x_i|x_i^*, T)$ of degree p of its Taylor expansion (eqn [2]).

The second key aspect of the mathematical theory concerns the replacement of the increment variables δx_i by their linear combinations, which have the property of being adapted to the group $\mathscr{G}^{\bigstar}$. Specifically, if, for the sake of simplicity, a single increment variable δx is considered, it is written as

$$\delta x = \sum_{\alpha,j} \eta_j^{\alpha} \cdot \xi_j^{\alpha} \qquad [3]$$

where, for a given value of an index α, the m degrees of freedom ξ_j^{α} $(j = 1, 2, \ldots, m)$ form a basis of an m-dimensional vector space $\mathscr{E}^{\alpha}$ invariant and irreducible under the action of $\mathscr{G}^{\bigstar}$. This property means, first, that each operation g belonging to $\mathscr{G}^{\bigstar}$ acts on a given ξ_i^{α} by transforming it into a linear combinations of the m variables ξ_j^{α} (same index α), and that, therefore, the space $\mathscr{E}^{\alpha}$ is globally invariant under this action. In addition, $\mathscr{E}^{\alpha}$ does not contain any subspace possessing this property of global invariance under the set of all elements of $\mathscr{G}^{\bigstar}$. The index α labels the set of irreducible representations Γ^{α} of this group, which is established once and for all on the basis of the mathematical structure of a group. This index specifies a "simple and standard" mode of action of $\mathscr{G}^{\bigstar}$ on an invariant vector space. For any group, one of the irreducible representations, Γ^0, of dimension $m = 1$, is the trivial representation whose unique basis vector is unchanged by the action of any element $\mathscr{G}$ of the group.

The action of $\mathscr{G}^*$ on δx can, in fact, be considered on the basis of eqn [3] in two equivalent ways: either by considering, as assumed above, that the η_j^{α} are mere coefficients unchanged by this action and that $\mathscr{G}^{\bigstar}$ only acts on the ξ_j^{α}, or, alternatively, that the ξ_j^{α} constitute a given set of fixed degrees of freedom and that the η_j^{α} are transformed by the action of $\mathscr{G}^{\bigstar}$ according to the representation Γ^{α}. The latter option is retained in the following. It has the advantage that the ξ_j^{α} can be chosen as fixed quantities with finite and normalized values and that both the symmetry properties and the infinitesimal character of δx are attributed to the η_j^{α} coefficients.

The preceding statements on the irreducibility of Γ^{α} and of $\mathscr{E}^{\alpha}$ need to be completed, because the mathematical theory of group representations considers vector spaces, and representations, based on complex numbers. Since the degrees of freedom are real quantities, it must be stated that Γ^{α} is either a real irreducible representation or a direct sum of two

complex conjugate irreducible representations. The action of $\mathcal{G}^\star$ on the $\eta_{\alpha,j}$ coefficients, for given α, is then defined by a set of orthogonal matrices $M^\alpha(\mathcal{G})$ associated with the various elements $_\mathcal{G} \in \mathcal{G}^\star$.

A series of central results can be derived from the $\mathcal{G}^*$ invariance of F and by using the variables η_j^α in $F(\eta_j^\alpha | x_i^*, T)$:

- For the theory, the very definition of a continuous phase transformation at T_c is based on the occurrence of a change in the symmetry of the equilibrium atomic configuration. Namely, the symmetry group $\mathcal{G}$ of the equilibrium configuration $x_i^{eq}(T)$ is $\mathcal{G} = \mathcal{G}^*$ on one side of T_c (generally for $T \geqslant T_c$) and $\mathcal{G} \subset \mathcal{G}^*$ on the other (for $T < T_c$). In other terms, in the situation generally encountered (which will be considered as always realized in the following), there is a breaking (also referred to as lowering) of the symmetry on decreasing the temperature across T_c.
- The trivial representation Γ^0 is not present in the linear expansion of δx (eqn [3]) and, accordingly, there is no term F_1 of degree one in the Taylor expansion of F.
- The term F_2 of degree two in this expansion necessarily has the form

$$F_2(\eta_j^\alpha | x_i^*, T) = \sum_\alpha A_\alpha(x_i^*, T) \left[\sum_j (\eta_j^\alpha)^2 \right] \quad [4]$$

  where all the variables η_j^α relative to each irreducible representation α contribute to a single sum of squares with coefficient A_α.
- In the neighborhood of T_c, $A_\alpha > 0$ for all A_α except one which vanishes and changes sign at T_c. Namely, if the corresponding index is written as $\alpha = \alpha^*$, then, with $a > 0$,

$$A_{\alpha^*} = a(T - T_c) \quad [5]$$

The m-dimensional set of m degrees of freedom $(\eta_j^{\alpha^*})$ associated with the vanishing coefficient A_{α^*} is called the order parameter of the phase transformation.

- Whenever an A_α is strictly positive at a given temperature, the equilibrium values of the corresponding components $(\eta_j^\alpha)^{eq}(T)$ are necessarily equal to zero, as straightforwardly deduced by minimizing the leading term in the Taylor expansion (eqns [2] and [4]). Hence, for $T > T_c$, all the components vanish. For $T < T_c$, in the approximation where just the leading term is considered, the only degrees of freedom which can take nonzero equilibrium values are the m components of the order parameter. This result implies that the change in atomic configuration as well as the main physical anomalies induced by the phase transformation can be studied by considering the free energy as a function of the sole components of the order parameter. The other variables (η_j^α), with $\alpha \neq \alpha^*$, can be ignored, in the first place, since their equilibrium values are equal to zero at any temperature. For the sake of simplicity, the index α can then be dropped. The free energy, whose minima determine the equilibrium configuration of the system at each temperature, is given by

$$F(\eta_j | T) = F_0 + a(T - T_c) \left[\sum_j (\eta_j)^2 \right] + \cdots + F_p(\eta_j | T) + \cdots \quad [6]$$

In summary, given the symmetry group $\mathcal{G}^*$ of the equilibrium configuration of a system at the transformation temperature T_c (and above it), the order parameter of a phase transformation is a set of m degrees of freedom $(\eta_1, \eta_2, \ldots, \eta_m)$ whose symmetry properties, that is, changes under the action of the elements of $\mathcal{G}^*$, are defined by the type of irreducible representation Γ^{α^*} associated with it. Physically, η_j are linear combinations of increments δx_i of the degrees of freedom of the system (e.g., atomic position, orientations of magnetic moments). The specificity of this combination is that its second-degree contribution to the Taylor expansion (eqn [2]) has a coefficient A_{α^*} that vanishes and changes sign at T_c.

$\mathcal{G}^\star$ Groups

The groups $\mathcal{G}^\star$ of operations encountered in the study of real physical systems are of various types. A standard situation of reference is that of phase transformations taking place in crystalline solids. $\mathcal{G}^*$ is then one of the 230 crystallographic space groups, obtained by combining specific orthogonal transformations (among which rotations through a restricted set of angles), an infinite discrete set of lattice translations, and certain fractions of these translations. The irreducible representations Γ^α of these groups, denoted $\Gamma^{\boldsymbol{k},n}$, form an infinite set, where the index $\boldsymbol{k}$ is a vector of the reciprocal space, whose three components can take a continuous set of values, and n is an integer taking a finite number of values. It can be shown that if the considered phase transformation leads to a low symmetry phase which is also crystalline, not all the irreducible representations $\Gamma^{\boldsymbol{k},n}$ are acceptable to describe the symmetry properties of the order parameter (Lifshitz criterion). The vector $\boldsymbol{k}$ must end on the so-called high-symmetry points of the Brillouin zone of the crystal, its components

being a few specific fractions of the reciprocal lattice vectors. For a given group $\mathcal{G}^*$, the task of enumerating the possible symmetry properties of order parameters of phase transformations between crystalline phases is then a finite one. Order parameter dimensions range from $m = 1$ to $m = 8$.

The irreducible representations $\Gamma^{k,n}$ with k different from the above restricted set of vectors, in particular when their projections on the reciprocal lattice vectors are not all rational numbers, are not devoid of physical interest. The corresponding order parameters describe phase transformations between a crystalline phase and a so-called incommensurately modulated crystal phase. In these phases, atomic positions deviate slightly from their positions in a perfect crystal, the deviations being periodic in space; further, the ratio of atomic periods in these phases to those in the reference crystal lattice are irrational numbers.

Another standard situation is that of phase transformations in crystals involving the onset of an organized configuration of magnetic moments. The symmetry group $\mathcal{G}^*$ then includes, in addition to the translations and orthogonal transformations already considered for crystals, the time-reversal operation $\mathcal{R}$ which reverses the directions of magnetic moments. This operation is of a specific mathematical nature because, unlike the geometrical operations considered above, it is associated with an antilinear operator. This feature introduces differences in the formulation of the order parameter symmetry that cannot be simply associated with an irreducible representation of $\mathcal{G}^*$.

Liquid crystals represent another class of substances that undergo a large variety of phase transformations. The relevant $\mathcal{G}^*$ groups are, as for crystalline solids, combinations of orthogonal transformations and translational. However, in this case, the orthogonal transformations generally involve a continuous set of rotations around an axis, and the translational operations form continuous sets at least in one direction.

$\mathcal{G}^*$-Invariant Free Energy

Within the framework of Landau's phenomenological theory, it can be shown that the presence of a term F_3 of degree three in the expansion (eqn [6]) is incompatible with the continuous character for the phase transformation. Imposing the absence of such a term (Landau criterion) puts a mathematical restriction on the irreducible representations Γ^α acceptable to describe the symmetry properties of an order parameter. Its mathematical formulation states that the symmetrized third power of the representation Γ^α must not contain the trivial representation of $\mathcal{G}^\bigstar$. Certain irreducible representations of the crystallographic groups do not comply with this criterion and are, therefore, not eligible for describing the symmetry properties of order parameters.

One fourth-degree term, at least, F_4 is always present in the F-expansion. Indeed, for any irreducible representation Γ^α, the term $[\sum_j (\eta_j)^2]^2$ is $\mathcal{G}^*$-invariant, being the square of the invariant second-degree term. In general, the $\mathcal{G}^*$-invariant fourth-degree term is a linear combination:

$$F_4(\eta_j, \beta_k) = \sum_k \beta_k f_4^{(k)}(\eta_j) \qquad [7]$$

where the β_k are arbitrary coefficients and each $f_4^{(k)}$ is $\mathcal{G}^*$-invariant and can be considered as a basis vector of the vector space $\mathcal{P}_4$ of $\mathcal{G}^*$-invariant fourth-degree polynomials. An example is given in a later section.

Symmetry Breaking

The restriction of the sum in eqn [3] to only the components η_j of the order parameter shows that the equilibrium configuration of the system, defined by the $\delta x^{\mathrm{eq}}(T)$, is determined by

$$\delta x^{\mathrm{eq}}(T) = \sum \eta_j^{\mathrm{eq}}(T)\xi_j \qquad [8]$$

where the η_j^{eq} are the values of the η_j that minimize the free energy (eqn [6]). At least one of the η_j^{eq} components has nonzero value below T_c. The nonzero components are specified by the terms $F_p(\eta_j|T)$ of degrees greater than 2. In most cases, the fourth-degree term is sufficient to determine them.

Equation [8] can be interpreted in the following manner: the equilibrium configuration $\delta x^{\mathrm{eq}}(T)$ is a nonzero vector with components η_j^{eq} in the m-dimensional space $\mathcal{E}^\alpha$ of the order parameter. While this space is globally invariant under the group $\mathcal{G}^*$, a given vector of this space, with components η_j, cannot be invariant under all the elements of this group (due to the irreducibility and the nontrivial character of $\mathcal{E}^\alpha$). Only a subgroup $\mathcal{G} \subset \mathcal{G}^*$ leaves this vector unmoved. $\mathcal{G}$ is therefore, the symmetry group of the equilibrium atomic (or spin) configuration $(x^* + \delta x^{\mathrm{eq}})$ below T_c (since x_i^* is invariant under $\mathcal{G}^* \supset \mathcal{G}$). The symmetry breaking occurring at T_c is, therefore, $\mathcal{G}^* \to \mathcal{G}$.

For a given irreducible representation Γ^α of $\mathcal{G}^\bigstar$, there is a finite number of these subgroups that can be systematically enumerated. They are called the small groups or isotropy groups of the vector defined by the η_j. Enumerating the set $\{\mathcal{G}\}$ of these small groups $\mathcal{G}$, for all directions in $\mathcal{E}^\alpha$, provides a collection of the potential symmetry changes which may be observed across a continuous phase transformation, for the considered order parameter symmetry. The

actual symmetry change is associated with the vector direction (η_j^{eq}) that minimizes F (eqn [6]).

It has been conjectured that the step of minimizing F could be avoided by relying on a group-theoretical rule: the maximal small-group rule. Thus, the "equilibrium direction" (η_j^{eq}) can be assumed to define necessarily a small group $\mathcal{G}_m \in \{\mathcal{G}\}$ complying with the condition that if $\mathcal{G}$ is any other small group ($\mathcal{G} \in \{\mathcal{G}\}$) distinct from $\mathcal{G}_m$, then

$$\mathcal{G}_m \not\subset \mathcal{G} \qquad [9]$$

This conjecture, although verified in most current situations (in particular for order parameter dimension $m = 1-3$), has been invalidated by a few theoretical examples of order parameter symmetries for which the calculated minimum of the free energy corresponds to a direction (η_j^{eq}) with a nonmaximal small group $\mathcal{G}$.

Intrinsic Symmetries of the Order Parameter and of the Free Energy

The "physical" groups $\mathcal{G}^\star$ are not the ones most effectively involved in the mathematical procedure exploiting the symmetry properties of the order parameter. Indeed, the action of the group $\mathcal{G}^\star$ in the space $\mathcal{E}^\alpha$ (i.e., the manner by which $\mathcal{G}^\star$ acts on η_j) is entirely specified by the set of matrices $M^\alpha(g)$ expressing this action. These matrices themselves form a group G^* which generally has a different structure than $\mathcal{G}^\star$, the correspondence $\mathcal{G}^\star \to G^*$ being only a homomorphism. G^* is often simpler than $\mathcal{G}^\star$. While $\mathcal{G}^\star$ is, for instance, an infinite crystallographic space group involving orthogonal transformations and translations, G^* can be a finite group of orthogonal transformations. G^* is, therefore, a subgroup of the so-called full orthogonal group in m dimensions, denoted $O(m)$. In addition, it is an irreducible subgroup, that is, one that leaves no subspace of $\mathcal{E}^\alpha$ invariant. The subgroups G^* can be enumerated for each order parameter dimension m, while $\mathcal{G}^\star$ also depends on the nature of the system. The G^* groups characterize the intrinsic symmetry of the order parameter. Their enumeration is easy for small values of m. Hence, for $m = 1$, $\mathcal{E}^\alpha$ is a line containing the origin O, and G^* is one of the two possible groups denoted as C_1 and $\overline{C_1}$, where C_1 is the trivial group reduced to the sole identity and $\overline{C_1}$ is the group of order 2 generated by the symmetry about the origin O. For $m = 2$, $\mathcal{E}^\alpha$ is a plane. There are two infinite discrete sets of groups G^*: groups C_n generated by the rotations of angle $2\pi/n$ around the origin O (with n a finite or infinite integer), and groups C_{nv} products of C_n by a reflection in a line passing through the origin. For $m = 3$, $\mathcal{E}^\alpha$ is the three-dimensional space, and G^* is either identical to $O(3)$ or to one of its irreducible subgroups, namely, the full symmetry groups of a cube, a regular tetrahedron, and a regular icosahedron. The Landau criterion, which forbids the existence of a third-degree term in the free energy, has the effect of reducing the number of acceptable intrinsic symmetries. Hence, for $m = 1$, among the two groups considered, only $\overline{C_1}$ is relevant. For $m = 2$, C_3 and C_{3v} are not acceptable, and for $m = 3$, the symmetry of the tetrahedron is excluded. Once the homomorphism between $\mathcal{G}^\star$ and G^* has been specified, the determination of the small groups $\mathcal{G} \subset \mathcal{G}^\star$ of the various directions in $\mathcal{E}^\alpha$ can be replaced by the working out of the small groups $G \subset G^*$ in $\mathcal{E}^\alpha$. This task can be achieved in a unified way for all the transformations sharing the same intrinsic order parameter symmetry.

Another type of intrinsic symmetry can be defined for the $\mathcal{G}^\star$-invariant polynomials of a given degree belonging to the free energy. Hence, one can define a homomorphism $G^* \to G^*_{(4)}$ in which the latter group acts in the vector space $\mathcal{P}_4$ of $\mathcal{G}^\star$-invariant fourth-degree polynomials defined by eqn [7]. Again, the set of $G^*_{(4)}$ groups is more restricted than the set of G^* groups. Since the entire Taylor expansion F is G^*-invariant, one has necessarily $G^* \subset G^*_p \subset O(m)$. There are only three distinct forms of fourth-degree terms for $m = 2$, corresponding to $G^*_{(4)} = C_4, C_{4v}$, or $O(2)$, and two for $m = 3$ corresponding to $G^*_4 = O_H$ (cubic symmetry) and $O(3)$. To each of these groups corresponds a general expression of fourth-degree contribution to the free energy. For instance, for $m = 3$ and $G^*_{(4)} = O_H$, this is given by

$$F_4 = \beta_1(\eta_1^2 + \eta_2^2 + \eta_3^2)^2 + \beta_2(\eta_1^4 + \eta_2^4 + \eta_3^4) \qquad [10]$$

Similar results can be derived for other values of m, and for terms of degree higher than 4.

Extension to the Formulation of the Renormalization-Group Hamiltonian

The so-called renormalization-group theory in reciprocal space, which deals with the critical behavior at a continuous phase transformation, relies on an "effective Hamiltonian density" which is given by

$$H(r) = \alpha \sum_j \phi_j^2(r) + \sum_k \beta_k f_4^{(k)}(\phi_j) + \sum (\nabla \phi_j)^2 \qquad [11]$$

Its spatially homogeneous terms, with coefficients α and β_k, are identical to the fourth-degree Taylor expansion (eqn [6]). Hence, the intrinsic symmetries G^* and $G^*_{(4)}$ defined earlier can be fully utilized to establish a set of relevant Hamiltonian densities.

In this theory, one generates a flow of Hamiltonian densities all given by eqn [11] and characterized by α and β_k coefficients which may converge toward a "stable fixed-point Hamiltonian" defined by the set of coefficients α^* and β_k^* and associated with a given physical critical behavior. The existence of such stable fixed points is considered as a requirement for the occurrence of a continuous transformation in the system.

A mathematically formulated criterion forbidding such an occurrence has been stated. It relies on consideration of groups acting in the vector space $\mathscr{P}_4$ of polynomials earlier. Let G_C be the small group of the direction (β_k^*) in this space. It may also be the invariance group of other directions in $\mathscr{P}_4$. Let $\{(\beta_j)\}$ represent their set (obviously containing the (β_k^*)). One can define a group G_N acting in $\mathscr{P}_4$ and leaving globally invariant the set $\{(\beta_j)\}$ (while G_C leaves every direction unmoved). The criterion states that a necessary condition for the stability of the considered fixed-point Hamiltonian (β_k^*) is that $G_C = G_N$.

See also: Allotropy and Polymorphism; Crystal Symmetry; Diffusionless Transformations; Group Theory; Incommensurate Phases; Liquid Crystals; Magnetic Point and Space Groups; Phase Transformation; Point Groups; Quasicrystals; Space Groups.

PACS: 64.70.Kb; 64.70.Rh; 75.10.−b; 75.25.+z; 77.84.−s; 64.60.i; 64.60.Cn; 64.60.Ak; 61.30.Cz; 61.44.Fw; 61.66.Dk; 62.20.Fe

Further Reading

Izyumov Yu A and Syromiatnikov VN (1990) *Phase Transitions and Crystal Symmetry*. Dordrecht: Kluwer Academic Publishers.

Janovec V and Kopsky V (2003) In: Authier A (ed.) *International Tables of X-Ray Crystallography*, vol. D.

Landau LD and Lifschitz EM (1958) *Statistical Physics*. London: Pergamon.

Lyubarskii Yu (1960) *The Application of Group Theory in Physics*. London: Pergamon.

Michel L (1980) Symmetry defects and broken symmetry. *Reviews of Modern Physics* 52: 617–651.

Stokes TH and Hatch DM (1988) *Isotropy Subgroups of the 230 Crystallographic Space Groups*. Singapore: World Scientific.

Toledano P and Dmitriev V (1996) *Reconstructive Phase Transitions in Crystals and Quasicrystals*. Singapore: World Scientific.

Tolédano JC and Tolédano P (1987) *The Landau Theory of Phase Transitions*. Singapore: World Scientific.

Nomenclature

E^α	vector space of the order parameter
F	variational free energy
$f_4^{(k)}$	$\mathscr{G}^\star$-invariant fourth-degree independent polynomial
$F_{\text{equilibrium}}$	equilibrium free energy
$\mathscr{G}$	low-symmetry group
G	subgroup of G^* and small group of a direction in the order parameter space
$\mathscr{G}^\star$	high-symmetry group
G^*	intrinsic symmetry-group of the order parameter
$G^*_{(4)}$	intrinsic symmetry group of fourth-degree polynomials
$\mathscr{P}_4$	vector space of fourth-degree polynomials
T_c	transformation temperature
α	label of an irreducible representation
β_k	coefficient of the fourth-degree expansion
$\Gamma^{k,n}$	irreducible representation of a crystallographic space group
Γ^α	irreducible representation of a group
η_j	components of the order parameter

Phases and Phase Equilibrium

M A White, Dalhousie University, Halifax, NS, Canada

Introduction

A phase is a homogeneous region of matter, separated from other homogeneous regions by phase boundaries. The different regions (phases) will have different physical and/or chemical properties. For example, the contents of a beaker with ice cubes floating in water form a two-phase system, as the ice and water are two separate phases. When more than one phase exists, the material is heterogeneous, but the length scale of heterogeneity can vary from nanometer to macroscopic dimensions.

Although phase equilibria can involve many types of phases, this article concentrates on solid–solid and solid–liquid equilibria.

Phase Stability for Pure Materials

If more than one phase of a pure material is possible, the phase with the lowest energy will be stable with respect to all others. The stability condition can

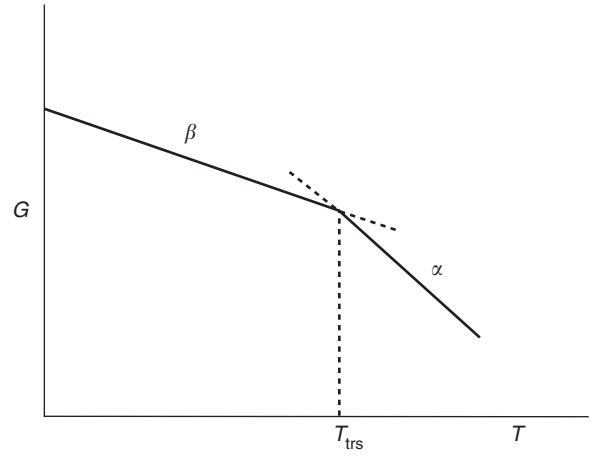


Figure 1 Gibbs energy as a function of temperature, showing a phase transformation at T_{trs}. The stable phase, shown by a solid line, is the one with the lowest G. For $T < T_{trs}$, phase α would be metastable with respect to phase β.

change as the temperature, or pressure, or volume, or other variables change.

If the temperature and pressure are the important variables, as in most laboratory conditions, then the phase with the lowest Gibbs energy (also commonly referred to as "Gibbs free energy"), G, will be the stable phase for those conditions. Since G is a state function of T and P such that

$$dG = V\,dP - S\,dT$$

it follows that

$$\left(\frac{dG}{dT}\right)_P = -S$$

so, for a given phase, G decreases as the temperature increases. If the $G(T)$ curve for two phases, α and β, cross (see **Figure 1**), the phase with the lower G will be stable. At the temperature where phases α and β are in equilibrium, $G_\alpha = G_\beta$. At a temperature at which $G_\alpha > G_\beta$, phase β is more stable than phase α. However, in these circumstances, α could exist in favor of β if the kinetic barrier for $\alpha \rightarrow \beta$ is too high. In that case, α would be metastable with respect to β because it is at a local energy minimum, but not at the global energy minimum. For example, liquid water can be supercooled below 0°C and is metastable with respect to ice for $T < 0$°C. Supercooled water owes its existence to the kinetic barrier to ice formation.

Polymorphism

For many pure materials, different phases are stable, depending on the temperature, pressure, electric field, magnetic field, etc.

Just within the temperature parameter, complex materials can show intermediate phases (known as mesophases) with intermediate amounts of disorder. In general, materials become more disordered as the temperature increases. For molecular materials with particular molecular shapes, two general categories of mesophases are possible.

If the material is composed of molecules which could easily glide past one another, because they are either rod-shaped or disk-shaped, at some intermediate temperature the material could form a liquid crystal. In the liquid crystalline mesophase, the molecules can glide past one another easily, and the material will conform to the shape of its container, making it appear liquid-like. However, if the molecules are aligned, there will be some crystal-like properties as well, for example, diffraction of light. Therefore, the term "liquid crystal" is quite apt. However, this is a mesophase: at lower temperatures the material will form a normal ordered solid, and at higher temperatures it will form an isotropic liquid.

On the other hand, if the molecules are nearly spherical, the barrier to reorientation can be so low that an orientationally disordered mesophase is possible. C_{60} is known to form such a phase: the C_{60} molecules are on the lattice points of a face-centered-cubic lattice, but they rotate rather freely so that they are orientationally disordered. At lower temperatures, the C_{60} molecules become more ordered.

Liquid crystals and orientationally disordered solids are two examples of polymorphism, which is literally the existence of more than one type of solid phase for a given composition. Polymorphism is quite common. Molecular materials are particularly prone to polymorphism because the Gibbs energy surface can have many energy minima, which can be explored as the external parameters (T, P, etc.) are varied. Even ionic materials can exhibit polymorphism due to phase changes related to electronic, magnetic, and structural properties. Examples include metal → superconductor transitions, ferromagnetic → paramagnetic transitions, and transitions with changes in crystal structure. Polymorphism plays an important role in many materials, from pharmaceuticals, where different polymorphs might have different biological activity, to geological materials, where polymorphism can be induced by extremes of temperature and pressure.

A phenomenon related to polymorphism is the existence of glass phases. A glass can be formed by supercooling a liquid. Below the glass transition temperature, T_g, the supercooled liquid becomes a rigid glass. Glass is essentially a "frozen liquid" in that its atoms are not ordered. However, the glass phase is almost always metastable with respect to a stable crystalline phase.

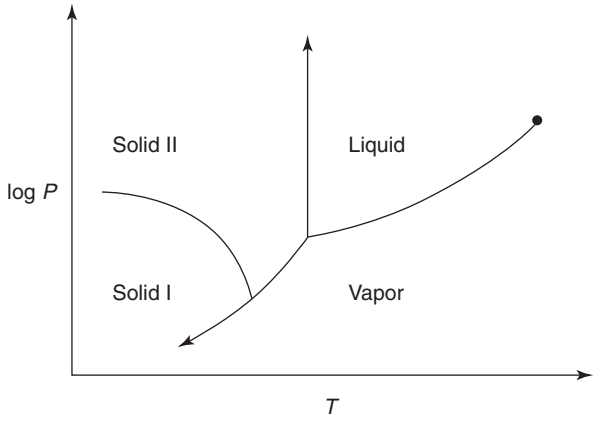


Figure 2 Schematic phase diagram for a pure (one-component) material, showing two different stable solids, labeled as I and II.

A schematic, equilibrium P–T phase diagram for a pure component with two stable solid phases is shown in **Figure 2**.

The Phase Rule

The phase rule, first devised by J Willard Gibbs, relates the number of degrees of freedom in a system, f, to the number of phases, p, and number of chemical components, c. Thus far, only single-component (i.e., pure, $c = 1$) systems have been considered, but many materials are composed of more than one component. The number of degrees of freedom is given by:

$$f = c - p + n$$

where n is the number of externally controllable variables (T, P, etc.) which can, in principle, be varied and change the state of the system.

For a pure system, $c = 1$, and if T and P can vary, $n = 2$. If a solid and liquid are in equilibrium, $p = 2$ and thus $f = 1$. This indicates that there is only one degree of freedom, and T and P cannot be varied independently while maintaining the two-phase equilibrium. This is apparent from the two-phase coexistence lines in **Figure 2**.

Phase Equilibria for Pure Materials

From the fact that two phases in equilibrium have the same value of G, and if the equilibrium is shifted by changing P and T such that equilibrium is maintained, the changes in G for the two phases are equal:

$$\mathrm{d}G^{\alpha} = \mathrm{d}G^{\beta}$$

and both can be represented by the corresponding pressure and temperature contributions:

$$\mathrm{d}G^{\alpha} = V^{\alpha}\,\mathrm{d}P - S^{\alpha}\,\mathrm{d}T = V^{\beta}\,\mathrm{d}P - S^{\beta}\,\mathrm{d}T$$

leading to the Clapeyron equation

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{\Delta_{\mathrm{trs}}S}{\Delta_{\mathrm{trs}}V} = \frac{\Delta_{\mathrm{trs}}H}{T_{\mathrm{trs}}\Delta_{\mathrm{trs}}V}$$

which quantifies the slopes of the phase equilibrium lines in **Figure 2** in terms of the changes in the thermodynamic properties due to the change in phase.

For the special case of the liquid–vapor equilibrium line, the Clapeyron equation can be simplified by taking into account that the molar volume of the gas is much greater than that of the liquid, and, since the vapor pressure is low the gas is likely to be ideal, that is, $\Delta_{\mathrm{trs}}V = V_{\mathrm{gas}} - V_{\mathrm{liquid}} \approx V_{\mathrm{gas}} = RT/P$ (where R is the gas constant); thus, for two liquid–vapor equilibrium points, (P_1, T_1) and (P_2, T_2),

$$\ln\left(\frac{P_2}{P_1}\right) = -\frac{\Delta_{\mathrm{vap}}H}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)$$

which is known as the Clausius–Clapeyron equation. For the liquid–vapor equilibrium, the P–T lines, as shown in **Figure 2**, can be known from the enthalpy change on vaporization, $\Delta_{\mathrm{vap}}H$, and vice versa.

Chemical Potential

For any multicomponent system, an important consideration for phase equilibrium is the chemical potential for component i, μ_i:

$$\mu_i = \left(\frac{\partial G}{\partial n_i}\right)_{T,P,n_{j\neq i}}$$

where n_i is the number of moles of component i. At equilibrium, for each component i:

$$\mu_i^{\alpha} = \mu_i^{\beta} = \cdots$$

where α, β,… represent the phases. In general, the chemical potentials of different components are different. The Gibbs energy is then given by

$$G = \sum_i X_i\mu_i$$

where $X_i = n_i/n_{\mathrm{tot}}$ and n_{tot} is the total number of moles. At equilibrium, the Gibbs–Duhem relation holds:

$$S\,\mathrm{d}T - V\,\mathrm{d}p + \sum_i n_i\,\mathrm{d}\mu_i = 0$$

Binary Systems

General Concepts

To describe a binary (two-component) mixture, three variables are required to describe the equilibrium phases: composition (usually expressed as mass percent or mole fraction), temperature, and pressure.

Although pressure can change the equilibrium phases, it is the most difficult variable to manipulate experimentally, and small pressure changes usually have little influence on phase equilibria, so binary phase diagrams are generally expressed as the isobaric presentation of temperature as a function of composition.

For a system composed of two components A and B, such that $n_A + n_B = 1$ mole, the Gibbs–Duhem equation becomes

$$S_{total}\,dT - V_{total}\,dP + X_A\left(\left(\frac{\partial\mu_A}{\partial T}\right)_P dT + \left(\frac{\partial\mu_A}{\partial P}\right)_T dP + \left(\frac{\partial\mu_A}{\partial X_B}\right)_{T,P} dX_B\right) + X_B\left(\left(\frac{\partial\mu_B}{\partial T}\right)_P dT + \left(\frac{\partial\mu_B}{\partial P}\right)_T dP + \left(\frac{\partial\mu_B}{\partial X_B}\right)_{T,P} dX_B\right) = 0$$

and with

$$d\mu_A = \left(\frac{\partial\mu_A}{\partial T}\right)_P dT + \left(\frac{\partial\mu_A}{\partial P}\right)_T dP + \left(\frac{\partial\mu_A}{\partial X_B}\right)_{T,P} dX_B$$

it follows that

$$(1 - X_B)\left(\frac{\partial\mu_A}{\partial X_B}\right) + X_B\left(\frac{\partial\mu_B}{\partial X_B}\right) = 0$$

which is a general binary solution result. For $X_B = 1 - X_A$, the Gibbs energy is given by

$$G = \mu_A + X_B(\mu_B - \mu_A)$$

which is the equation for a straight line for $G(X_B)$ with slope $(\mu_B - \mu_A)$ and intercept μ_A. Such a line would be a tangent in a graph of $G(X_B)$. Recalling that the phase equilibrium condition is the equivalence of μ_i, the chemical potential of each component i in each phase, it follows that the $G(X)$ curves for the phases in equilibrium must share a common tangent. This concept will prove useful in the analysis of $G(X)$ for binary phase diagrams (*vide infra*).

Solid Solutions

One of the simplest binary phase diagrams is one in which both components are completely miscible in each other in all proportions, forming a solid solution. Examples include metals which form alloys and many mineral systems. In general, solid solutions require the two components to have very similar interatomic interactions (similar types of atoms, and similar sizes). A generalized binary phase diagram showing formation of a solid solution is shown in **Figure 3**.

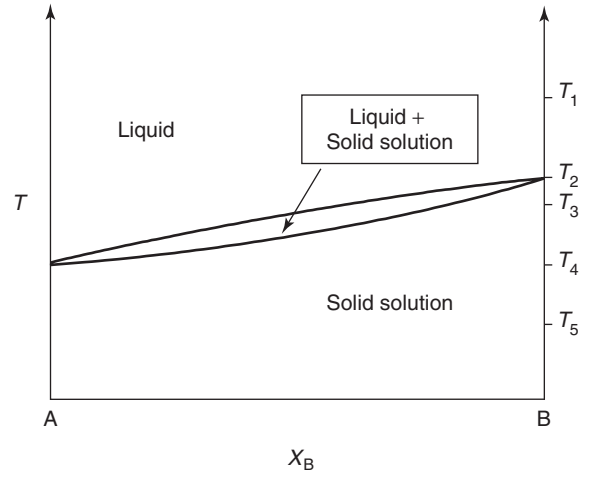


Figure 3 A schematic binary phase diagram showing complete miscibility of both components to form a solid solution.

The corresponding Gibbs energy diagrams for solid solution formed under various isobaric, isothermal conditions are shown in **Figure 4**. At temperatures below the melting point of either component, the solid solution is the most stable phase, due to favorable enthalpic effects and/or entropic considerations.

A consideration of the Phase Rule for the various regions of **Figure 3**, with $c = 2$ and $n = 1$, indicates that $f = 2$ in the single-phase regions (i.e., T and X can vary independently) and $f = 1$ in the two-phase regions (i.e., T and X are no longer independent variables).

Immiscible Solids

At the other extreme, sometimes components in a binary system are completely immiscible in the solid state. This would give rise to a phase diagram such as the one illustrated schematically in **Figure 5**. The depressed freezing point of component A on the addition of component B is given by T in

$$\ln X_A = \frac{\Delta_{fus}H_A}{R}\left(\frac{1}{T_A} - \frac{1}{T}\right)$$

where $\Delta_{fus}H_A$ is the enthalpy change on melting (fusion) of the pure component A, T_A is the melting temperature of pure A, and R is the (ideal) gas constant. Similarly, the freezing point of B is depressed on the addition of A, and where they meet is a special point, called the eutectic. At the eutectic composition, immiscible solids melt to directly give the liquid, whereas at off-eutectic compositions melting leads to a temperature range over which solid and liquid coexist.

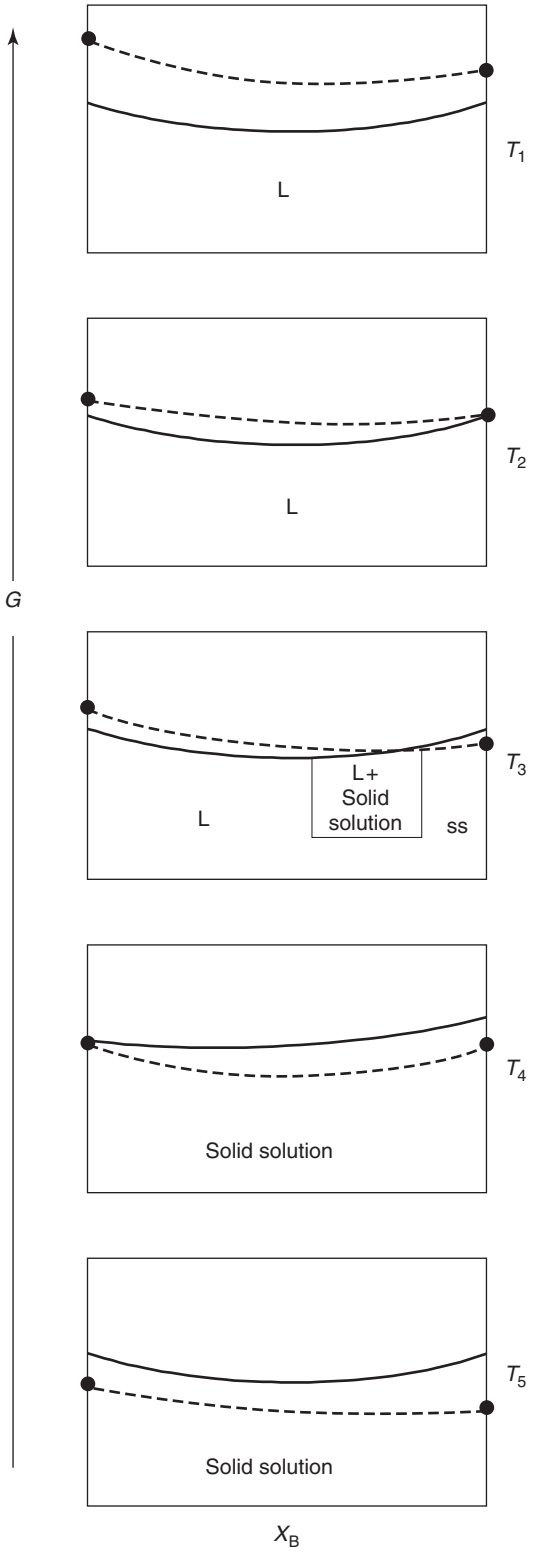


Figure 4 Gibbs energy as a function of composition for a binary system which forms a solid solution, with the temperature as indicated in **Figure 3**. The dots indicate Gibbs energies for fixed compositions, the solid curves indicate G for the liquid, and the dashed line is G for the solid solution. Note that the stable phase has the lowest value of G. When two phases are in equilibrium, their $G(X_B)$ curves share a common tangent, as at T_3.

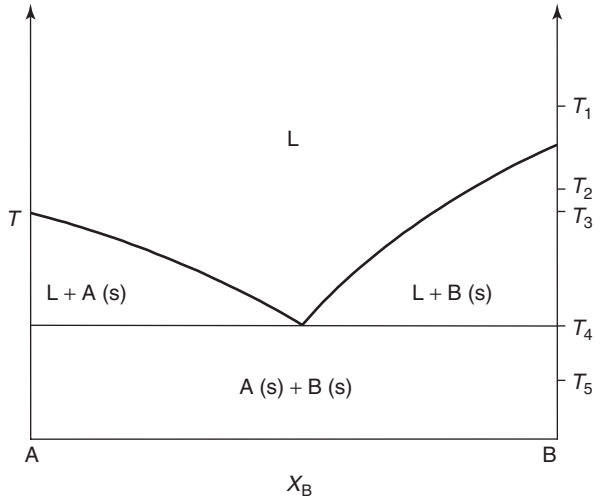


Figure 5 A schematic binary phase diagram showing complete immiscibility of the components.

In the case of immiscible solids, the components would be very dissimilar in their packing and/or interaction energies, making it more favorable for the solid to be a heterogeneous mixture of the two pure (solid) components. The corresponding Gibbs energy diagrams for immiscible solids under various isobaric, isothermal conditions are shown in **Figure 6**.

Again, the Phase Rule for the various regions of **Figure 5**, with $c = 2$ and $n = 1$, gives $f = 2$ in the single-phase regions, $f = 1$ in the two-phase regions, and now $f = 0$ at the eutectic (three phases). The invariance of the eutectic makes it useful for thermometry calibration, as are triple points in single-component systems ($f = 0$ also).

Congruently Melting Compounds

Often the two components in a binary mixture have sufficiently favorable interactions that a stoichiometric compound is formed at a certain composition. An example of a 1:1 compound of component A and component B, of empirical formula AB, is shown in **Figure 7**. This particular compound melts directly at a fixed temperature to give a liquid of the same composition, a feature known as congruent melting.

In some ways, the phase diagram of **Figure 7** is like two **Figure 5** phase diagrams, side by side. However, there is a distinction in the slope of the liquid line: at the compound composition, the slope is zero, whereas the slope is nonzero at the components (A or B in both **Figures 5** and **7**). The reason for the difference stems from the difference in composition of the liquid: pure A melts to give liquid A, and likewise for pure B, but compound AB melts to give A(liq) miscible with B(liq), and the freezing point depression equation given above no longer holds.

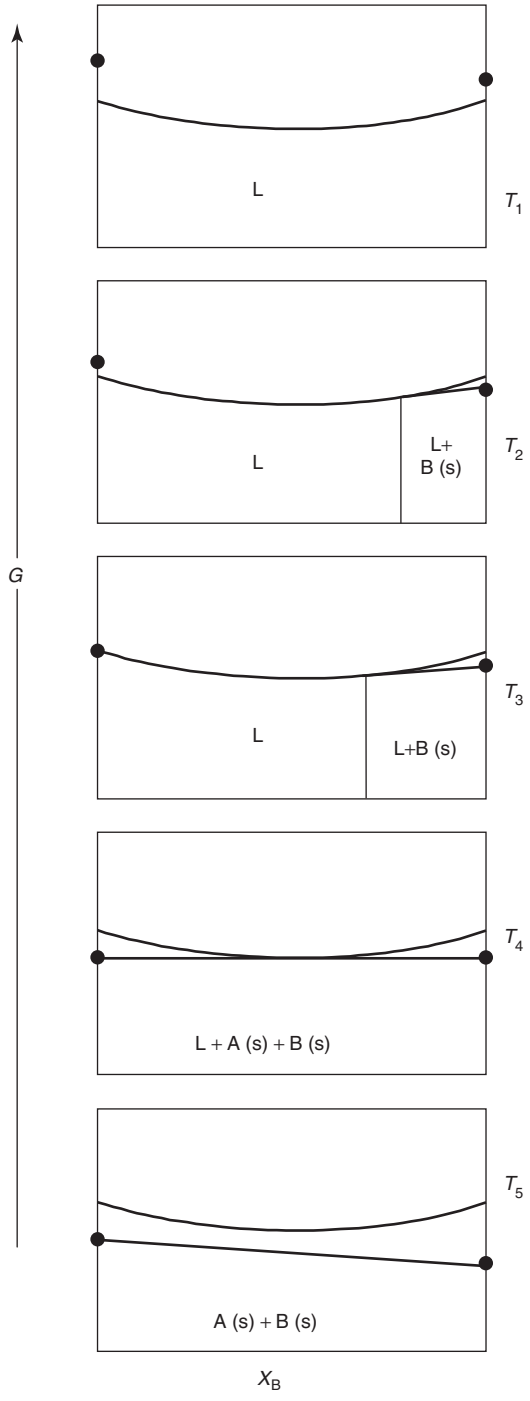


Figure 6 Gibbs energy as a function of composition for a binary system in which the solids are immiscible, with the temperatures as indicated in **Figure 5**. The dots indicate Gibbs energies for fixed compositions and the solid curves indicate *G* for the liquid.

The Gibbs energy diagrams corresponding to **Figure 7** under various isobaric, isothermal conditions are shown in **Figure 8**. Here there is a new feature compared with the previous $G(X)$ curves, the compound AB, since it is sufficiently stable to influence the phase diagram.

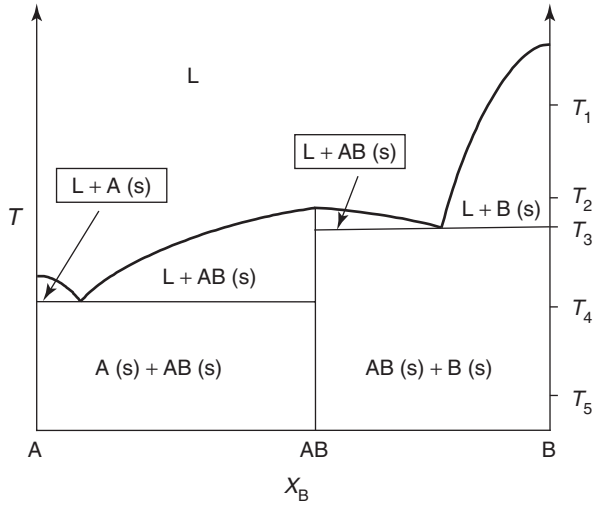


Figure 7 A schematic binary phase diagram showing the formation of a congruently melting compound of empirical formula AB.

As for the other isobaric binary phase diagrams, $f = 2$ in the one-phase regions, $f = 1$ in the two-phase regions, and $f = 0$ at the eutectic. Note that at the melting point of the congruently melting compound, $c = 1$, $n = 1$ so $f = 0$ as for a pure compound.

Incongruently Melting Compounds

Not all compounds melt congruently. Some melt with decomposition of the compound to form a liquid and one of the components, as shown in **Figure 9**. In this case, heating AB leads to

$$\mathrm{AB\ (s)} \rightarrow \mathrm{liquid} + \mathrm{B(s)}$$

and AB is said to melt incongruently.

Compounds which melt incongruently do so because the relatively low (i.e., favorable) Gibbs energy of one of the components (see **Figure 10**), and especially its rapid change with temperature.

Since

$$\left(\frac{\partial G_\mathrm{A}}{\partial T}\right)_P = -S_\mathrm{A}$$

it follows that a highly entropically stabilized component could favour incongruent melting.

Again, from the Phase Rule, $f = 2$ in the one-phase regions, $f = 1$ in the two-phase regions. Now, $f = 0$ at the incongruent melting point of the binary compound. (This is called the peritectic temperature, and the peritectic point is shown in **Figure 9**.) For a warming curve at the composition of the incongruently melting compound, the otherwise increasing temperature will halt at the peritectic temperature ($f = 0$) until the solid (AB in the case of **Figure 9**) has converted to liquid + B(s).

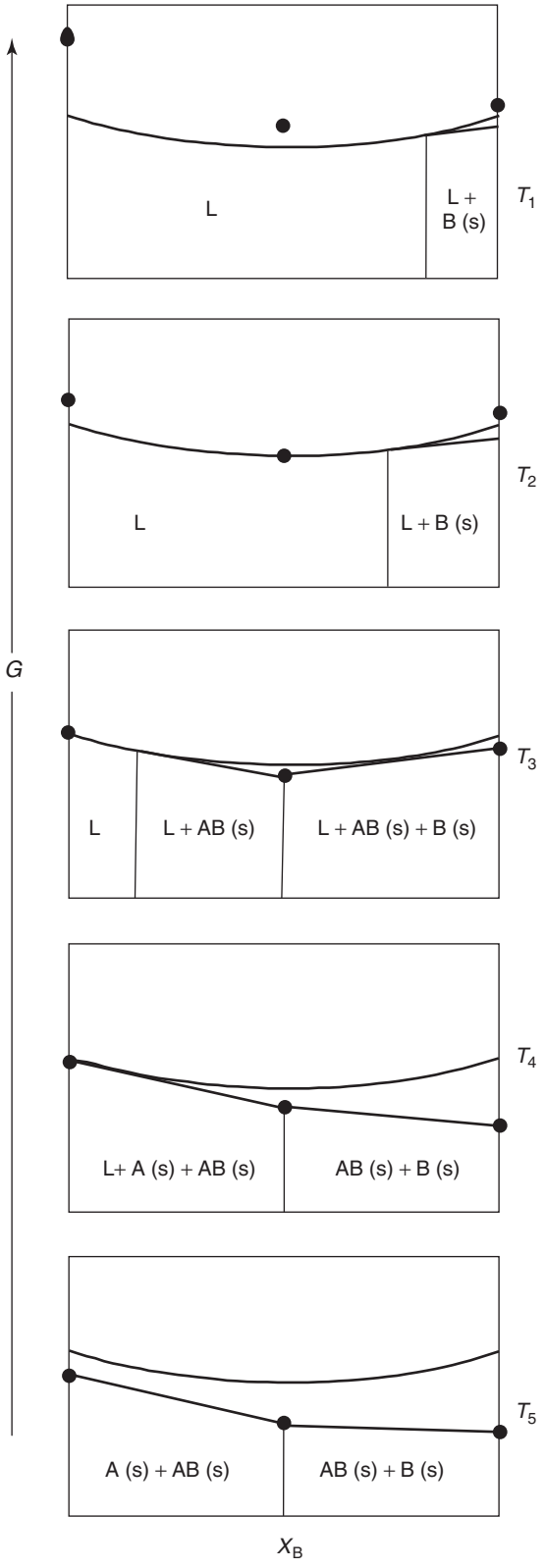


Figure 8 Gibbs energy as a function of composition for a binary system in which a congruently melting compound, AB, forms. The temperatures are as indicated in **Figure 7**. The dots indicate Gibbs energies for fixed compositions and the solid curves indicate *G* for the liquid.

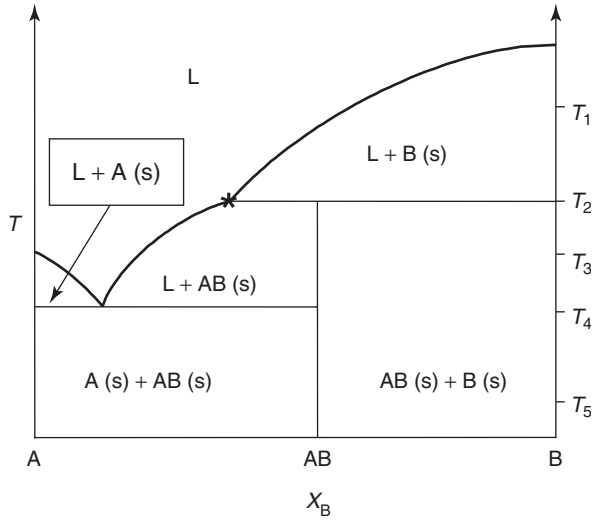


Figure 9 A schematic binary phase diagram showing the formation of an incongruently melting compound of empirical formula AB. The peritectic is shown as ✶.

Proportions of Phases in Binary Systems

In any binary system, the phases present can be surmised from the ends of an imaginary horizontal line, called a tie line. The relative proportions of the phases are inversely proportional to the distances from the ends of the tie line, given as a lever with the fulcrum at the composition under consideration. For example, at the composition shown by X_B^* in **Figure 11**, two phases coexist at temperature T*, liquid of composition X'_B, and solid B ($X_B = 1$). The relative molar proportions of liquid and B at X_B^* are given by

$$\frac{n_{liq}}{n_B} = \frac{1 - X_B^*}{X_B^* - X'_B}$$

This lever rule is a general relationship: the closer to the phase, the higher the proportion. Furthermore, the lever rule applies to both mole fraction (for a diagram expressed in *X* as for **Figure 11**), and mass fraction or mass percent (for a phase diagram expressed in mass fraction or mass percent, respectively).

Ternary Systems

When three components are mixed, making a ternary system, there are now four independently manipulable variables, *viz*. *T*, *P*, and two mole fractions (the third being given by difference). For ease of presentation, ternary phase diagrams are normally shown as isothermal isobaric slices, with the compositions given by triangular coordinates, as shown in **Figure 12**.

The sides of the triangle represent isobaric, isothermal slices of the binary system represented by the ends of the sides. For example, the side opposite the

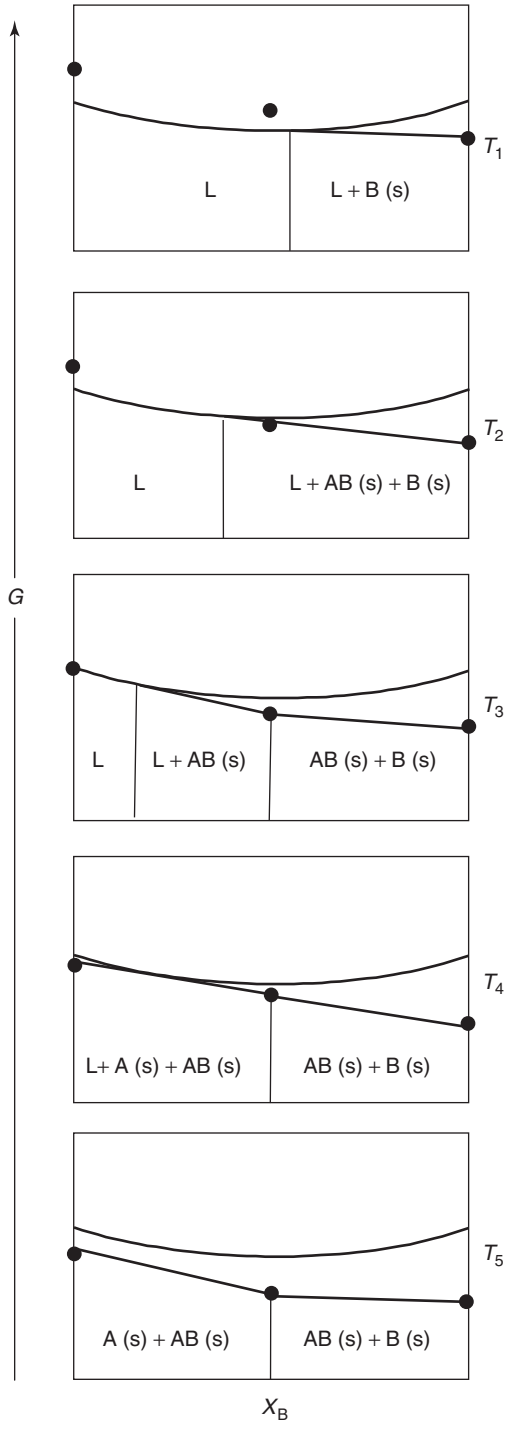


Figure 10 Gibbs energy as a function of composition for a binary system in which an incongruently melting compound, AB, forms. The temperatures are as indicated in **Figure 9**. The dots indicate Gibbs energies for fixed compositions and the solid curves indicate *G* for the liquid.

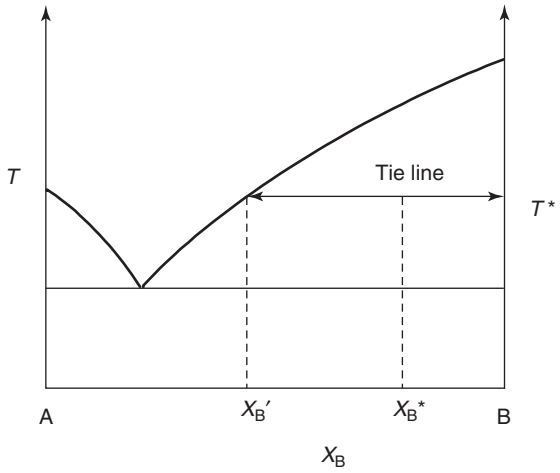


Figure 11 Illustration of the lever principle. See the text for details.

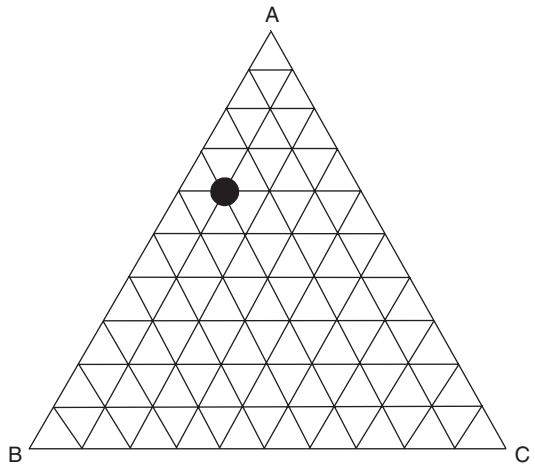


Figure 12 Graphical representation of compositions for a ternary (three-component) phase diagram. The lines correspond to constant composition in a single component. The distance from a component (A, B, or C) is inversely proportional to its concentration. For example, the B–C line at the bottom corresponds to $X_A = 0$. Each line parallel to that is at a different composition in A, at 10% steps in A. Similarly, the lines leading away from the B and C vertices correspond to constant composition in B and C, respectively. The point shown by • corresponds to 60% A, 30% B, 10% C. The ternary diagram can be expressed as mole fraction, or mass percent.

C vertex corresponds to $X_C = 0$, that is, the A–B binary system.

An example of a ternary phase diagram is shown in **Figure 13**. Here there is partial solubility of each of the components in the other. In general, ternary phase diagrams can illustrate all the features discussed so far, namely solid solution, solid immiscibility, congruently melting compound formation, and incongruently melting compound formation. Again, the number of degrees of freedom in the system is specified by the Phase Rule. In particular, note that for an isothermal, isobaric ternary system, that is, $c = 3$, $n = 0$, the maximum number of phases is three.

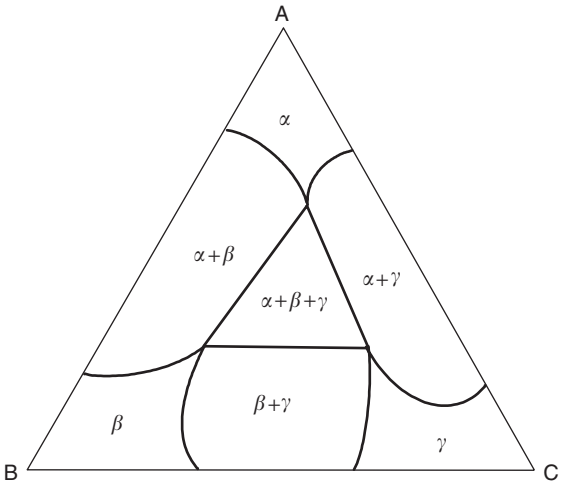


Figure 13 A schematic isobaric, isothermal ternary phase diagram, showing partial miscibility. The phases α, β, and γ are solid solutions rich in A, B, and C, respectively. Note that no more than three phases coexist.

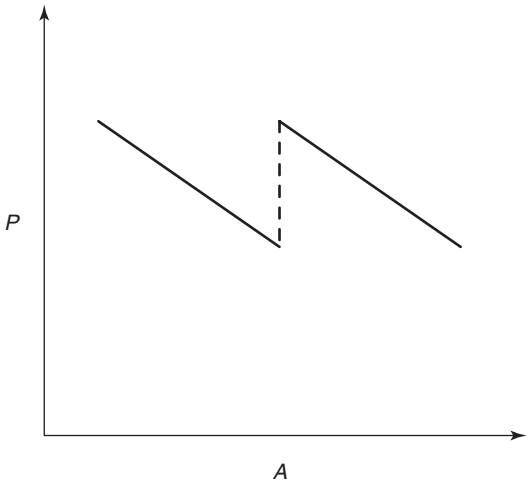


Figure 14 The pressure as a function of surface area, *A*, for a Langmuir–Blodgett film on a liquid surface. At low surface coverage (large *A*), the coverage is less than a monolayer. As the area is decreased, the pressure increases until the monolayer coverage and then at lower surface area the layer collapses, as shown by drop in *P*. At still lower surface areas, the film is greater than a monolayer thick and still shows compressibility.

Other Phase Equilibria

Thus far, only bulk phases of matter have been considered. However, if the material was two-dimensional, such as in a Langmuir–Blodgett film on a liquid surface, other phase equilibria would need to be considered. The pressure–surface area graph of **Figure 14** indicates this concept.

In the case of micelle formation by amphoteric molecules, the important phase equilibrium parameters are temperature and micelle concentration.

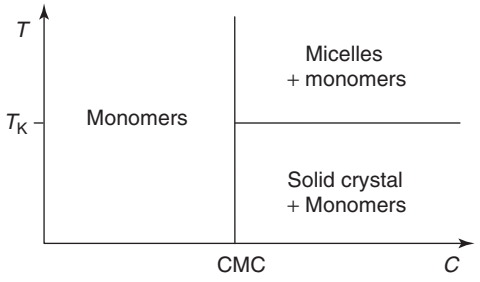


Figure 15 Phase diagram showing the temperature–concentration relationships for amphoteric molecules in solution. At low concentration, the molecules do not aggregate and appear as monomers. Above the critical micellar concentration (CMC), the molecules aggregate into micelles if the temperature is above the Krafft temperature, T_K. At high concentrations and low temperatures, a crystalline solid prevails.

This is illustrated in the phase diagram shown in **Figure 15**.

One might expect quite a different phase behavior when matter is finely divided (e.g., nanoparticle sized) due to the large surface-to-volume ratio. This can change phase stability in the presence of thermal, pressure, and other influences.

Furthermore, even within bulk materials, many other factors can influence phase behavior, including electric field, magnetic field, and strain field.

See also: Bose–Einstein Condensation; Composites: Overview; Glasses; Phase Transformation; Quasicrystals.

PACS: 61.43.–j; 61.30.–v; 64.; 64.10.+h; 64.60.My; 64.70.Kb; 64.70.Pf; 68.05.–n; 68.18.–g

Further Reading

Bergeron CG and Risbud SH (1997) *Introduction to Phase Equilibria in Ceramics*. Westerville, OH: American Ceramic Society.

Berry RS, Rice SA, Ross J (1980) *Physical Chemistry*, pp. 932–942. New York: Wiley.

Campbell AN and Smith NO (1951) *The Phase Rule and Its Applications*, 9th edn. New York: Dover.

Cedeno CL and Harne MA (2003) *Phase Equilibria Diagrams – 2003 Cumulative Index*. Westerville, OH: American Ceramic Society.

Cunningham JR and Jones DK (1991) *Experimental Results for Phase Equilibria and Pure Component Properties*. New York: American Institute of Chemical Engineers.

Hillert M (1998) *Phase Equilibria, Phase Diagrams, and Phase Transformations: Their Thermodynamic Basis*. Cambridge: Cambridge University Press.

King MB (1969) *Phase Equilibrium in Mixtures*. Oxford: Pergamon.

Oonk HAJ (1981) *Phase Theory*. Amsterdam: Elsevier.

Sandler SI (1993) *Models for Thermodynamic and Phase Equilibria Calculations*. New York: Dekker.

Saunders N and Miodownik AP (1998) *Calphad: Calculation of Phase Diagrams: A Comprehensive Guide*. Pergamon Materials Series, vol. 1. Oxford: Pergamon.

White MA (1999) *Properties of Materials*, pp. 158–193. New York: Oxford University Press.

Nomenclature

A	area
c	number of chemical components
f	number of degrees of freedom
G	Gibbs energy
n	number of moles
p	number of types of phases
P	pressure
R	ideal gas constant
S	entropy
T	temperature
T_{trs}	temperature of transition
V	volume
$\Delta_{fus}H$	enthalpy change on fusion (melting)
μ_i	chemical potential of component i

Phonons *See* Electron–Phonon Interactions and the Response of Polarons; Lattice Dynamics: Vibrational Modes; Mechanical Properties: Elastic Behavior; Polarons.

Photoelectron Spectromicroscopy

G Margaritondo, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Spectromicroscopy is the class of experimental techniques that combine spectroscopic performances with high spatial resolution. The name is most commonly used for photoelectron spectromicroscopy, which is based on the different spectroscopic techniques exploiting the photoelectric effect. Two different approaches are discussed: photon-focusing photoelectron spectromicroscopy and electron-optics spectromicroscopy. The discussion includes a few selected examples and an analysis of relevant performances. Finally, other types of spectromicroscopy are briefly reviewed.

Spectromicroscopy

The term "spectromicroscopy" broadly applies to all cases in which a spectroscopic technique is implemented with high spatial resolution. Therefore, many techniques can be technically considered under the term "spectromicroscopy." The discussion is focused on a leading subset of this class: spectromicroscopy techniques related to the photoelectric effect.

Besides photoemission spectromicroscopy, the important types of spectromicroscopy are: (1) electron energy loss spectroscopy (EELS), which can be performed with high lateral resolution providing local information on the electronic and chemical structure of surfaces; (2) auger spectroscopy, a leading chemical probe, can also achieve high lateral resolution; (3) X-ray-excited fluorescence, a powerful and relatively straightforward technique for chemical analysis. When performed with a high-brightness synchrotron source and an X-ray lens, the technique can achieve high lateral resolution becoming a powerful microchemical probe called fluorescence microprobe.

Near-field microscopy is a well-known approach that beats the diffraction limits of far-field microscopy by working with evanescent waves. In some cases, the technique can be implemented in a spectroscopic mode. For example, a near-field infrared microscope can perform vibrational spectroscopy in localized areas while still beating the diffraction limit. **Figure 1** shows an example of this approach.

The Development of Photoelectron Spectromicroscopy

Photoelectron spectroscopy, the class of experimental techniques based on the photoelectric effect, has emerged since the late 1960s as a leading probe of the electronic structure in condensed matter systems. The progress in the experimental tools progressively eliminated the historical limitations of such techniques, enhancing their impact and expanding their field of application. In this evolution, the development of photoelectron spectromicroscopy is one of the milestones.

Until the late 1980s, photoemission experiments were performed without lateral resolution – whereas the natural surface sensitivity of the photoelectron emission process automatically provided atomic-scale resolution in the direction perpendicular to the emitting surface. Technical problems, primarily related to the low signal level, forced the photoemission experiments to probe a sample area of the order of $1 \times 1\,\text{cm}^2$ or even larger. This was a severe limitation in the study of laterally inhomogeneous systems and processes.

In 1988–89, the first example of photoemission experiment with high lateral resolution was presented (see, e.g., **Figure 2**), opening up new opportunities in

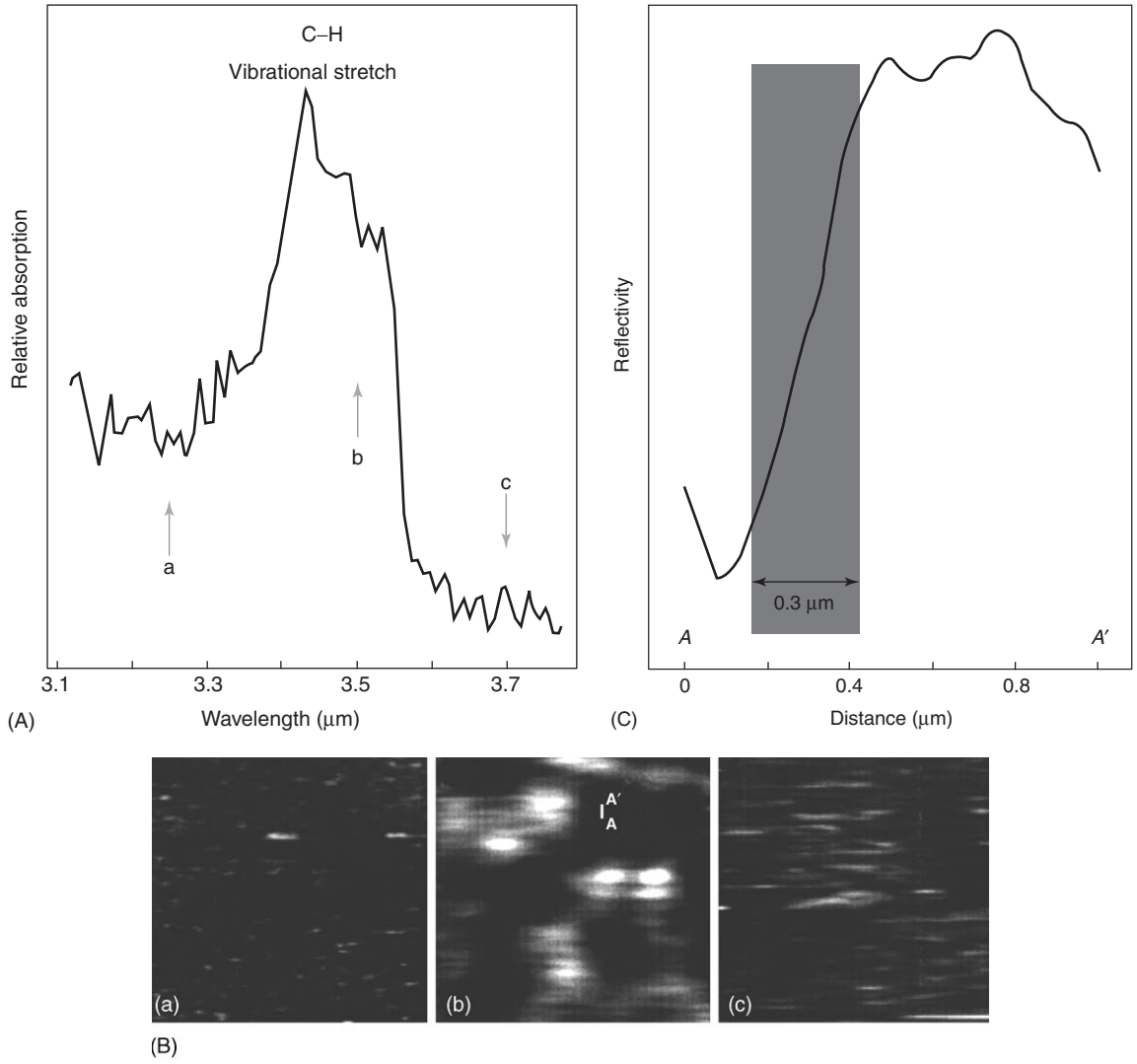


Figure 1 One of the first examples of near-field spectromicroscopy: (A) the optical absorption band corresponding to the C–H stretch vibrational mode; (B) $10 \times 10\,\mu m^2$ near-field images taken with the photon wavelengths labeled as (a), (b), and (c) in **Figure 9a**, inside and outside the C–H absorption band. The central image shows H-related features revealing an unexpected hydrogen contamination of the examined diamond film. (C) The intensity line scan along the *AA′* line of **Figure 9b** demonstrates a lateral resolution well below the diffraction limit. The data were obtained at the Vanderbilt University Free Electron Laser. (Results derived from Cricenti A, Generosi R, Luce M, Perfetti P, Margaritondo G, *et al.* (2001) Spectroscopic scanning near-field optical microscopy with a free-electron laser: CH_2 bond imaging in diamond films. *Journal of Microscopy* (Oxford) 202: 446.)

the general field of photoemission techniques. Lateral resolution could be achieved thanks to the advent of high-brightness synchrotron radiation sources such as the undulators. In the subsequent years, photoemission spectromicroscopy in different forms rapidly expanded becoming one of the most prolific subdomains of photoelectron spectroscopy.

One important aspect of this evolution was the impact on the life sciences. Without lateral resolution, photoelectron spectroscopy had been almost useless in biology, biophysics, and the medical research. Spatial dimensions such as the size of cells and cell components determine in fact the minimum spatial resolution to provide experimental information in these domains. Before the late 1980s, photoelectron spectroscopy was, as mentioned, very far from such lateral resolution levels.

After the advent of photoelectron spectromicroscopy, photoelectron-based probes became instead quite important in biology and medical research. They provide in fact very valuable microchemical information without requiring extensive sample preparation and the consequent possibility of artifacts. On the other hand, photoelectron spectromicroscopy requires ultrahigh vacuum and this constitutes a severe limitation in biological applications.

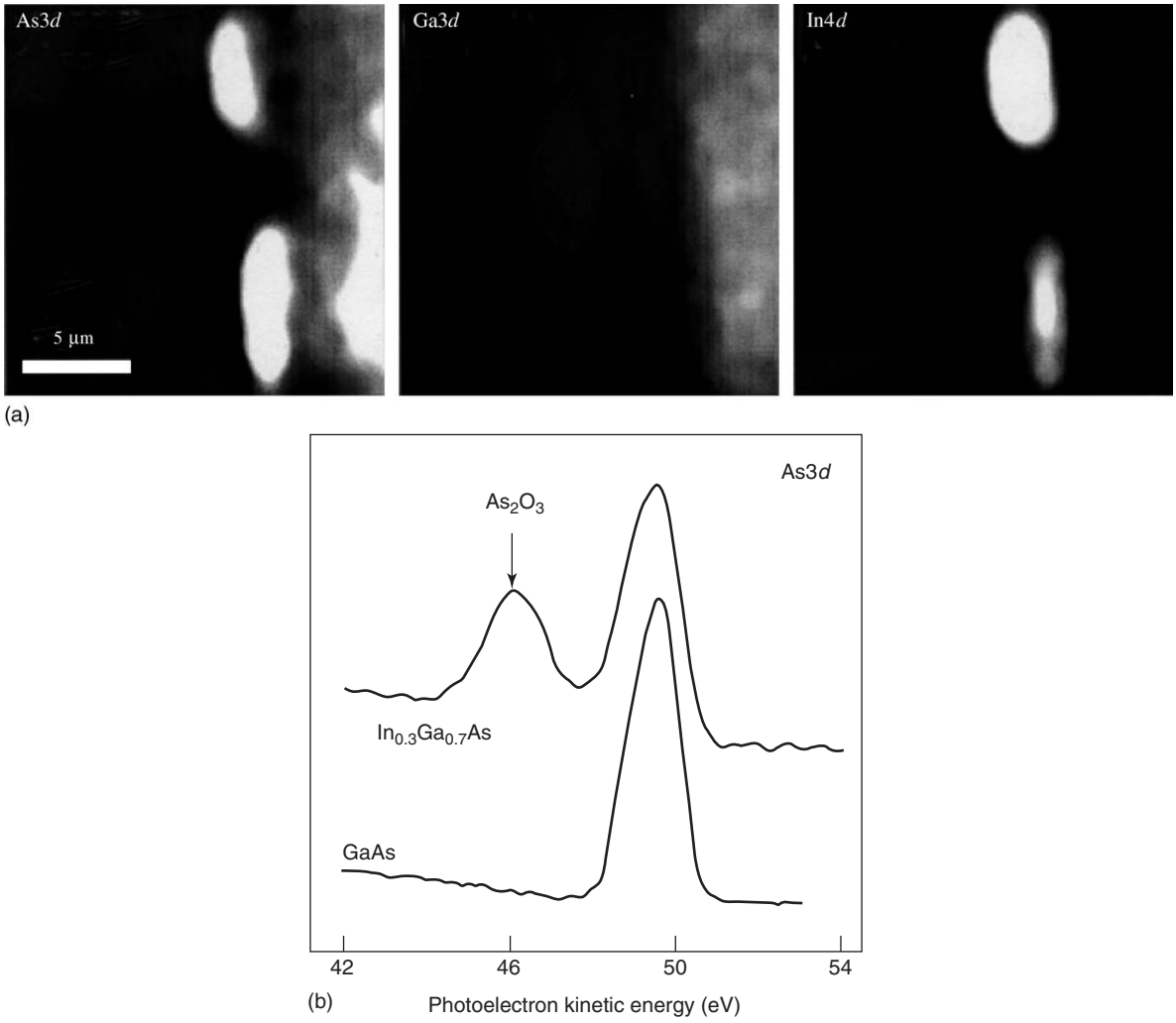


Figure 2 One of the first results of photoelectron spectromicroscopy, obtained at the University of Wisconsin Synchrotron Radiation Center with the "MAXIMUM" spectromicroscope. (Results derived from Ng W, Ray-Chaudhuri AK, Liang S, Singh S, Solak H, *et al.* (1994) Photoemission spectroscopy reaches the 1000 Å scale. *Nuclear Instruments and Methods* A347: 422.) (a) Two-dimensional photoemission intensity maps across a GaAs-$In_{0.3}Ga_{0.7}As$ interface for the As3*d*, Ga3*d*, and In4*d* core-level peaks. (b) Local As3*d* spectra taken in two spots at the two sides of the interface. Note the As_2O_3 component present on the $In_{0.3}Ga_{0.7}As$ side but not on the GaAs side.

Two Approaches to Photoelectron Spectromicroscopy

There exist basically two approaches to achieve lateral resolution in a photoemission experiment (**Figure 3**). In the first case, the primary photon beam is focused into a small sample area and photoelectrons are excited only from that area. This approach leads to two complementary data-taking procedures. In the first procedure, the photon beam is kept in a constant position on the sample surface and photoemission data are taken from that small spot. In the second case, the focused beam is scanned on the sample surface (or, rather, the samples are scanned with respect to the beam) while photoelectrons are collected for a constant set of parameters. This produces two-dimensional "images" of the photoemission intensity.

Consider (**Figure 4**) the case in which the focused spot is scanned while the collected photoelectron energy corresponds to photoelectrons emitted from a given core level to a given element. The photoemission intensity roughly corresponds to the concentration of the corresponding element. Therefore, the "images" provide two-dimensional chemical maps of the lateral distribution of different component elements.

Furthermore, a core-level energy for a given element changes slightly depending on the oxidation state of the element. Therefore, by fine-tuning the collected photoelectron energy one can obtain chemical maps of a given element in different chemical

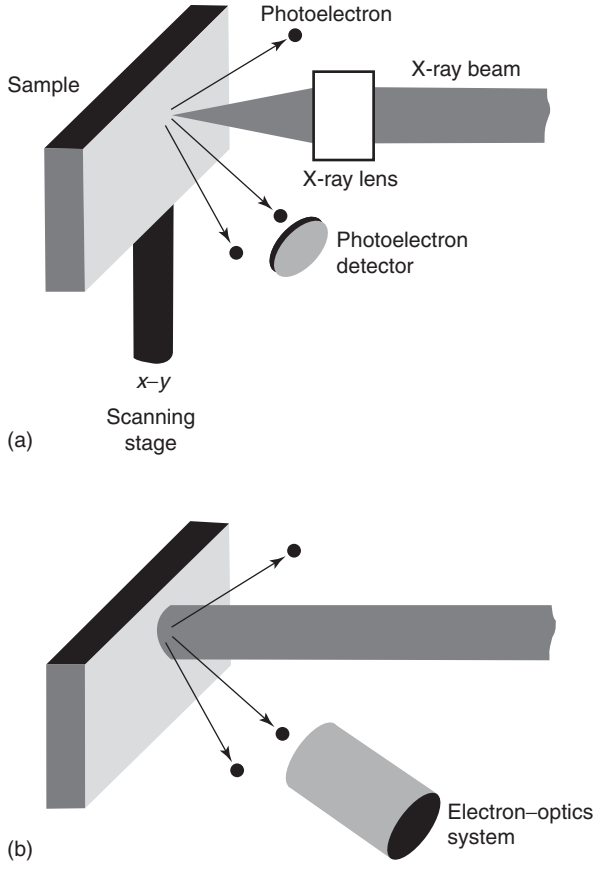


Figure 3 The two implementation modes of photoelectron spectromicroscopy: in (a) the lateral resolution is achieved by focusing the photon beam. In (b) the resolution is given by an electron lens system.

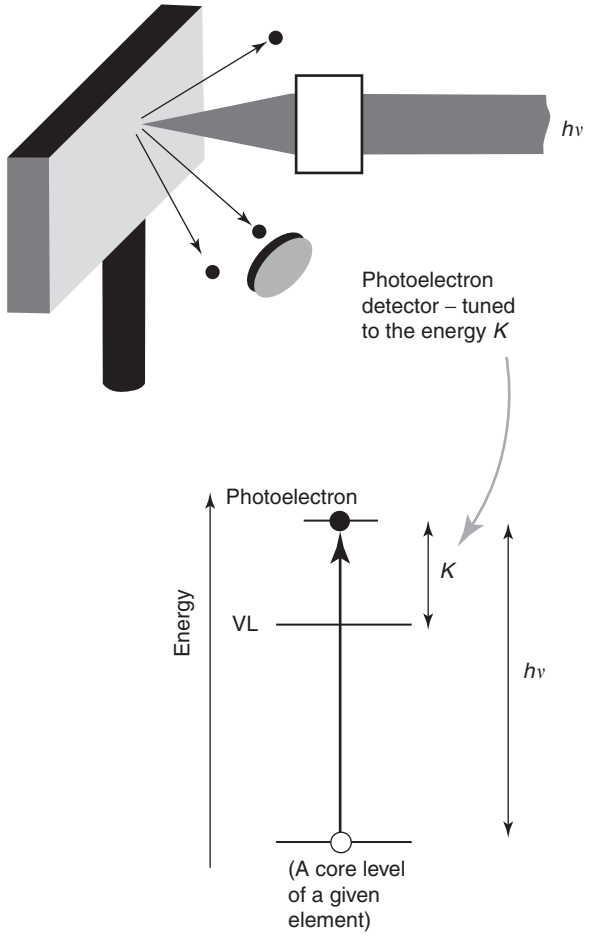


Figure 4 Chemical microanalysis with the photon-focusing mode of photoelectron spectromicroscopy. The photoelectron detector is tuned to a kinetic energy corresponding to the photoelectron emission from a given core level of a given chemical element. The corresponding photoemission intensity reveals the presence and concentration of that element.

environments. Since photoelectron spectromicroscopy has the high surface sensitivity typical of nearly all photoemission experiments, the chemical maps concern the region near the sample surface and provide extremely valuable information on surface processes such as corrosion, passivation, and catalysis.

The second approach to photoelectron spectromicroscopy (see again **Figure 3**) does not require focusing the photon beam. Lateral resolution is achieved by processing the emitted photoelectrons with an electron–optical system roughly equivalent to an electron microscope. Several years of practical experience demonstrated that the two approaches have advantages and disadvantages with respect to each other, and are largely complementary.

Practical Implementation: Photon-Focusing Spectromicroscopy

Focusing photon beams is – or at least was – a major technical problem in the implementation of the first of the above two approaches to photoelectron spectromicroscopy. The photons must have sufficient energy to enable electrons inside the solid to overcome the surface energy barrier and become photoelectrons. The corresponding smallest photon energies fall in the ultraviolet range.

Furthermore, the energy of the photons must be sufficient to extract photoelectrons from the core levels of relevant chemical elements. The corresponding photon energies range from the vacuum ultraviolet to the hard X-rays.

Focusing is a relatively easy task only in a very small portion of this spectral range, corresponding to its lowest photon energies (a few eV). For higher photon energies (X-rays), focusing is a difficult task.

The reasons are quite simple: (1) solid materials absorb photons in a large portion of the vacuum ultraviolet and soft-X-ray ranges, making it impossible to use transmission optical devices similar to optical lenses. (2) Hard X-rays can travel through solid materials, but the X-ray refractive index is very

close to that of vacuum ($n = 1$), drastically limiting the effectiveness of refractive focusing devices. (3) The reflection of X-rays is extremely limited except at near-grazing incidence, and this tendency becomes stronger when the wavelength decreases. This rules out reflection-focusing devices working at near-normal incidence.

Effective technical solutions for these problems have been conceived and implemented. **Figure 5** shows a series of X-ray-focusing devices suitable for different spectral domains.

Focusing Mirrors

Mirrors can provide X-ray focusing as long as the angle of incidence is sufficiently close to grazing incidence. This, however, complicates the fabrication of mirrors and makes them quite expensive. The surface profile must in fact meet stringent requirements over a large area.

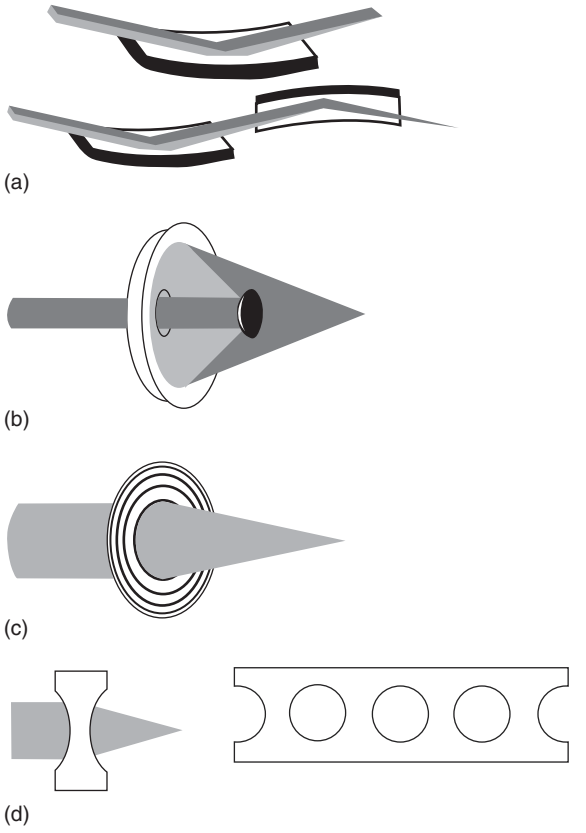


Figure 5 Different X-ray focusing devices. (a) A single grazing-incidence mirror (top) focuses only in one direction, whereas (bottom) a combination of two mirrors (Kirkpatrick–Baez lens) provides bidirectional focusing. (b) The Schwarzschild objective is formed by a convex spherical lens and a concave spherical lens, with nongrazing-incidence reflection enhanced by multilayer coatings. (c) A Fresnel zone plate (FZP). (d) (Left) a single biconcave refractive lens for hard-X-rays, and (right) a compound refractive lens (CRL).

The focusing surface can be toroidal, parabolic, or spherical (with some degree of aberration). Machining such surfaces with the required accuracy is technically difficult. It is possible to use cylindrical surfaces that are easier to fabricate (e.g., bent mirrors), but they focus only in one direction. Bidirectional focusing requires therefore a pair of mirrors combined together in a "Kirkpatrick–Baez" lens (**Figure 5a**).

Schwarzschild Objectives

These devices (**Figure 5b**) focus visible light as well as X-rays, and this facilitates the alignment of the X-ray optical system. The objective is the combination of two reflecting spherical surfaces. Note that the reflection is not at grazing incidence and this would limit the throughput.

The problem can be solved by a suitable multilayer coating that enhances reflection. The corresponding technology is, however, quite sophisticated. Furthermore, these lenses can only work in a limited spectral range that is typically located below 100 eV of photon energy. Even with these limitations, the Schwarzschild objectives did play a significant role in the development of photoelectron spectromicroscopy.

Fresnel Zone Plates

The fresnel zone plates (FZPs) are widely used as focusing devices in different spectral ranges. As illustrated in **Figure 5c**, the working principle is based on the interference of rays passing through a series of alternating transmitting and blocking circular zones. The blocking zones are supported by a very thin substrate to provide reasonable transmission in the X-ray domain.

The main difficulty in fabricating FZPs for X-rays is related to the dimension of the zones. The radial size of the zones decreases on going from the center to the periphery of the plate; furthermore, the size decreases as the wavelength decreases. For X-ray wavelengths, the outermost-zone size is in the order of hundreds of angstroms. The fabrication requires sophisticated – and expensive – electron-beam lithography techniques. These problems notwithstanding, FZPs are extensively used in X-ray microscopy and spectromicroscopy.

Refractive Lenses and Compound Refractive Lenses

For hard X-rays, the difficulties related to the limited transmission become progressively less important and it is possible to use refractive lenses. There exist, however, significant differences between the refractive lenses for visible light and those for X-rays. The refractive index for X-rays is smaller in a

material than in vacuum, whereas the contrary is true for visible light. An X-ray focusing lens, therefore, is not biconvex but biconcave (**Figure 5d**).

As already mentioned the refractive index differences between a material and vacuum are extremely limited. As a consequence, the focal length of a single biconcave refractive lens would be unreasonably large. The solution is provided by a series of lenses in the form of spherical concavities inside a block, as also shown in **Figure 5d**. Such a device is known as "compound refractive lens" or CRL.

Practical Implementation: Electron-Optics Spectromicroscopy

Figure 6 shows a typical electron-optics system for the implementation of this second approach to photoelectron spectromicroscopy. The standard name for this type of instrument is photoelectron emission microscope (PEEM).

Without photon focusing, it becomes easier to perform photoemission techniques that require changing the photon energy. The most frequent way to use a PEEM is to perform "partial yield" (PY) measurements with high lateral resolution. A PY experiment is implemented by taking the intensity of the "secondary" photoelectrons as a function of the photon energy – see **Figure 7**.

The secondary photoelectrons are the low-kinetic-energy electrons emitted from the sample after a loss of energy with respect to the sum of the initial energy in the sample plus the photon energy. Such losses are typically due to multistep phenomena. The yield of secondary photoelectrons is proportional to the number of absorbed photons. Therefore, a PY curve closely resembles the plot of the X-ray absorption coefficient as a function of the photon energy.

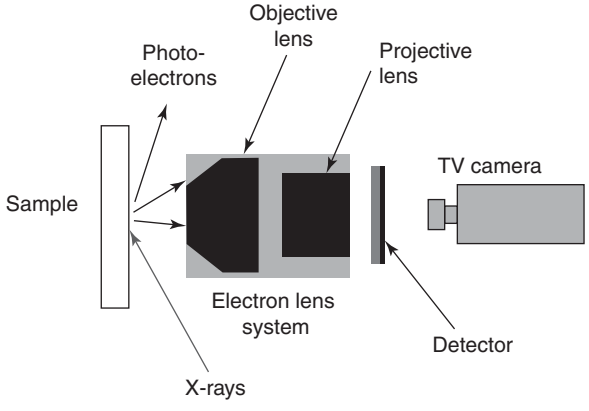


Figure 6 Schematic experimental setup of a photoelectron emission microscope (PEEM) for the second mode of photoelectron spectromicroscopy.

Note, however, that the PY technique is surface sensitive – although the surface sensitivity is not as high as for other photoemission techniques due to the low kinetic energy of the secondary photoelectrons. Therefore, a PY curve primarily reflects the absorption coefficient of the surface rather than that of the bulk.

With a PEEM, PY curves can be taken from small sample areas. This is a very powerful way to explore the local chemical composition and chemical properties. An X-ray absorption spectrum is dominated by the absorption threshold of the component elements in the specimen and reveals their presence and chemical status. This approach is particularly convenient when local chemical information must be obtained rapidly.

The technique is typically implemented by first taking overall images corresponding to the partial-yield intensity emitted from different sample areas with a fixed photon energy (or with a "white" photon

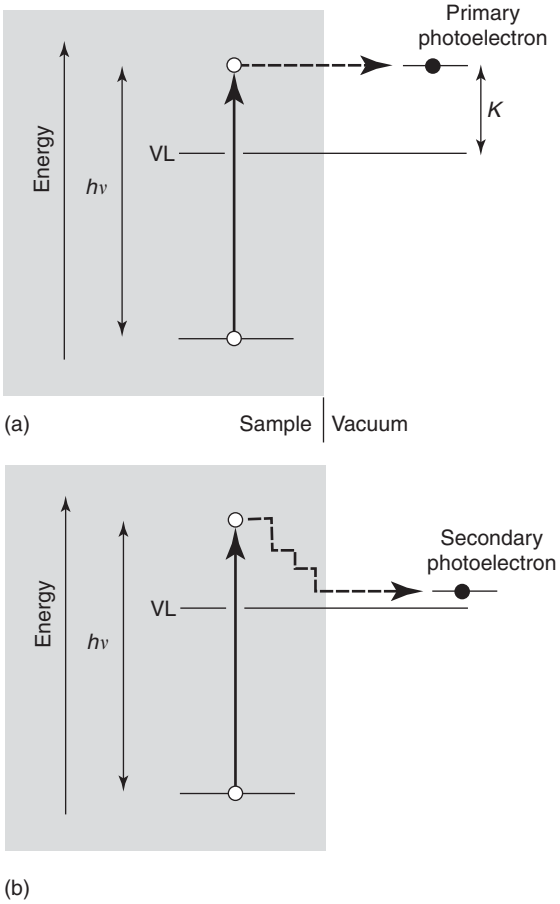


Figure 7 (a) Primary photoelectrons are emitted with an energy (above the vacuum level VL) corresponding to the initial energy in the solid plus the photon energy, without losses. (b) A secondary photoelectron loses part of its energy, typically through multiple losses, before leaving the sample. The partial yield (PY) mode of photoelectron spectroscopy and spectromicroscopy is based on the detection of secondary photoelectrons.

beam). Such images reveal the overall morphology of the specimen. Then, specific areas are selected for a more detailed analysis and PY spectra are taken from them.

Selected Examples and Performances

Figure 8 shows an example of photoelectron spectromicroscopy in the photon-focusing mode. The spectra illustrate a specific core level (Ge3*d*) of a metal–semiconductor interface (a thin germanium film on a semiconductor substrate).

This particular system has been considered as a typical example of nonreactive interface in which no chemical reaction takes place. On the contrary, spectra taken in different parts of the interface show clear differences, providing evidence for local chemical reactions that were not visible in a normal photoemission experiment.

Figure 9 shows an example of electron-optics photoelectron spectromicroscopy obtained with a PEEM operating in the PY mode. Note that the experiment concerns a biological sample: the PEEM technique is in fact particularly suited for studies in the life sciences and in medical research.

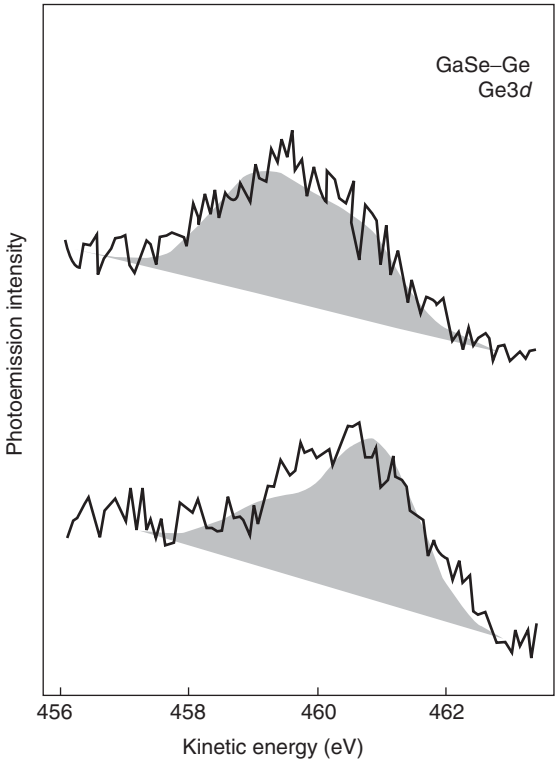


Figure 8 An example of photon-focusing photoelectron spectromicroscopy result (from Almeida J, Vobornik I, Berger H, Kiskinova M, Kolmakov A, *et al.* (1997) Spectromicroscopic evidence of Ge–GaSe chemical reactions: not a Schottky system. *Physical Review B* 55: 4899): two spectra taken at two different spots of a thin Ge overlayer on a GaSe substrate reveal different spectral line shapes for the Ge3*d* peak – and put in evidence unexpected local chemical reactions. The results were obtained at the Elettra synchrotron in Trieste.

Resolution and Sensitivity

The most relevant performances for photoelectron spectromicroscopy are the lateral resolution and the

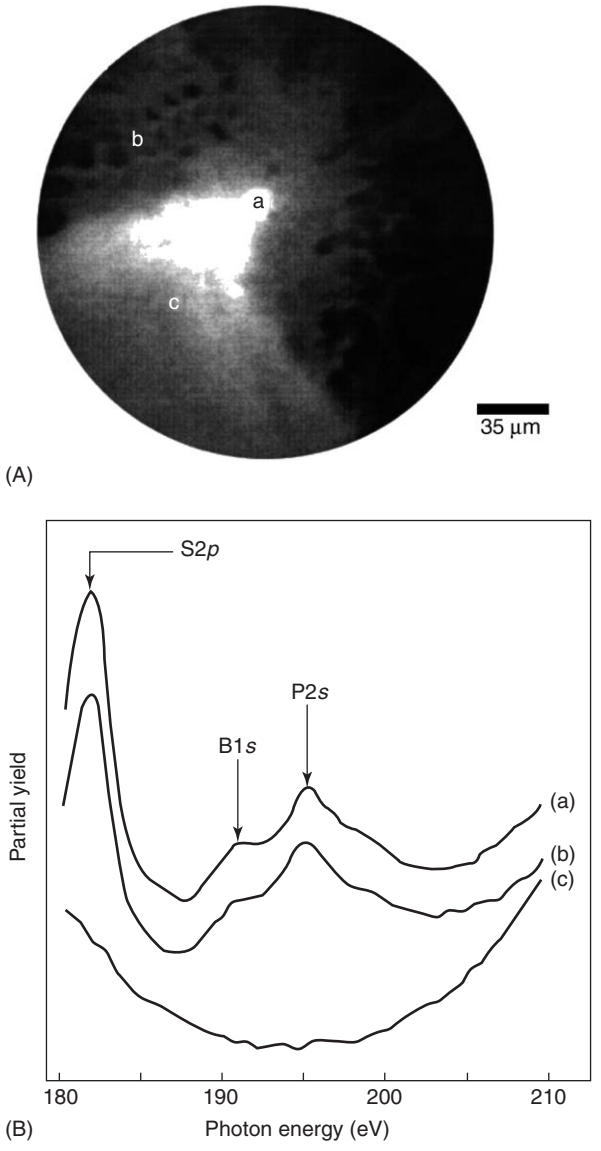


Figure 9 (a) PEEM microimage taken for a cancerous brain cell specimen with pre-injected boron, as required for the BNCT (boron neutron capture therapy). (b) PEEM PY spectra taken in specific spot reveal the localization of boron from the corresponding signal. The data were obtained with the PEEM-class spectromicroscope "MEPHISTO" at the University of Wisconsin Synchrotron Radiation Center. (Data derived from Gilbert B, Redondo J, Baudat P-A, Lorusso GF, Andres R, *et al.* (1998) Spectromicroscopy of boron for the optimization of boron neutron capture therapy (BNCT) for cancer. *Journal of Physics D* 31: 2642.)

sensitivity to trace chemicals. Concerning lateral resolution, for both approaches to photoelectron spectromicroscopy it is necessary to distinguish between the "system" (optical or electron–optical) resolution and the real resolution that can be obtained while performing spectromicroscopy.

In the spectromicroscopy mode the resolution is in fact also limited by practical factors such as the signal level. Consider for example the comparison between the resolution delivered by an electronic lens system used as an electron microscope and the corresponding resolution when the system is used in a PEEM. In the first case, the lateral resolution can reach a few nanometers. In the spectromicroscopy mode, the best practical resolution is ~ 20 nm. The standard resolution in most PEEM experiments is worse than this optimal value.

Likewise, X-ray focusing devices can deliver a lateral resolution better than the corresponding real resolution in spectromicroscopy. Consider for example a Fresnel zone plate: its optical resolution depends on the size of the outermost zone, and can be better than 10 nm. The corresponding "spectromicroscopy" resolution, although at the submicron level, does not reach comparable values – and is typically limited to 100 nm.

Note that the actual lateral resolution in a spectromicroscopy experiment is often the result of a compromise between conflicting requirements. For example, lateral resolution can be limited in order to achieve a good energy resolution while keeping the signal at a reasonable level.

Chemical sensitivity is an important point for spectromicroscopy. In a normal photoemission experiment without lateral resolution, the minimum detectable relative concentration is 0.01–0.001. This is, of course, the local concentration in the near-surface region probed by photoemission. If the minority species is concentrated in that region, the equivalent relative concentration limit for the entire specimen is much lower.

Spectromicroscopy concentrates the analysis to areas of $\sim 1\,\mu m^2$ or less. If the minority species is not homogeneously distributed, by focusing on high-concentration areas the equivalent detectable concentration limit can be lowered.

Future Developments

Spectacular improvements in the instrumentation and in particular in the photon sources open up many new avenues. On the one hand, additional spectroscopy techniques can be implemented with high lateral resolution becoming spectromicroscopies. Furthermore, it becomes increasingly possible to simultaneously achieve high lateral resolution and other extreme performances such as ultrahigh energy resolution. One of the most attractive possibilities is the coupling of time resolution and lateral resolution to study dynamic phenomena on a microscopic scale.

See also: Core Photoemission; Deep-Level Spectroscopy and Auger Spectra; Lattice Dynamics: Vibrational Modes; Low-Energy Electron Microscopy; Optical Microscopy; Surfaces and Interfaces, Electronic Structure of; Synchrotron Radiation; Valence Photoemission; X-Ray Absorption Spectroscopy.

PACS: 79.60.−i, 68.37.Xy, 68.37.Uv, 82.80.Pv, 87.64.Lg, 42.79.Ci

Further Reading

Ade H (ed.) (1997) *Spectromicroscopy with VUV Photons and X-rays*. Amsterdam: Elsevier.

Bauer E (2001) Photoelectron spectromicroscopy: present and future. *Journal of Electron Spectroscopy* 114: 975–987.

Gunther S, Kaulich B, Gregoratti L, and Kiskinova M (2002) Photoelectron microscopy and applications in surface and materials science. *Progress in Surface Science* 70: 187–260.

Gunther S, Kolmakov A, Kovac J, and Kiskinova M (1998) Artefact formation in scanning photoelectron emission microscopy. *Ultramicroscopy* 74: 35–51.

Kinoshita T (2002) Application and future of photoelectron spectromicroscopy. *Journal of Electron Spectroscopy* 124: 175–194.

Margaritondo G (1988) *Introduction to Synchrotron Radiation*. New York: Oxford University Press.

Margaritondo G (1995) The information content of photoemission spectroscopy and spectromicroscopy. *Surface Review and Letters* 2: 305–313.

Margaritondo G (2003) *Elements of Synchrotron Radiation for Chemistry, Biology and Medical Research*. New York: Oxford University Press.

Margaritondo G and Cerrina F (1990) Overview of soft-X-ray photoemission spectromicroscopy. *Nuclear Instruments and Methods* A291: 26–35.

Margaritondo G and De Stasio Gelsomina (1997) Photoelectron spectromicroscopy with synchrotron radiation: applications to neurobiology. In: Ade H (ed.) *Spectromicroscopy with VUV Photons and X-Rays*, pp. 137–147. Amsterdam: Elsevier.

Margaritondo G and De Stasio Gelsomina (1997) Synchrotron spectromicroscopy for the life sciences: general considerations and special procedures. *International Journal of Imaging Systems and Technology* 8: 188–203.

Thieme J, Schmahl G, Rudolph D, and Umbach E (eds.) (1998) X-ray Microscopy and Spectromicroscopy: Status Report from the Fifth International Conference, Wurzburg, August 19–23, 1996. Heidelberg: Springer.

Photon Statistics and Coherence Theory

P Mataloni, Università di Roma "La Sapienza," Rome, Italy

Introduction

The study of optical fluctuations represents an important tool in understanding the fundamental properties of light. One can distinguish among different types of light sources. In a common spectroscopic source, such as a gas discharge lamp, the different atoms are excited by an electrical discharge and emit their radiation independently of one another. The radiation produced by this source is called chaotic light. Several varieties of light, as those produced by a thermal cavity or a filament lamp, be-long to the same class. A second type is the coherent light, the best approximation of which is given by a single-mode laser operating well above the threshold at a single frequency and in a well-defined spatial direction.

Chaotic and coherent light have quite different statistical properties. These are measured by first- and second-order optical interference experiments that allow one to study the electric field intensity correlations between different portions of the light beam, with different coordinates in time and/or space. The fundamental component for optical interference is represented by the optical beam splitter, present in several kinds of interferometers, such as the Mach–Zehnder (MZ) and the Brown–Twiss (BT) interferometer. A beam splitter is basically a semitransparent mirror with, in general, equal probability to reflect or transmit the light beam, $R = T$. MZ is an example of first-order interferometer, while the BT represents the typical scheme for experiments of second-order interference.

Depending on its nature, the light can show either a classical or nonclassical behavior. Classical and quantum theories predict identical interference effects for coherent and chaotic light, or light with intermediate coherence properties. Exclusively quantum effects occur with other kinds of light that cannot be classically described. No difference is found between classical and quantum predictions for first-order interference experiments, while a second-order interference allows one to distinguish between classical and nonclassical light.

Chaotic Light

This article starts by examining some phenomena related to the behavior of chaotic light by looking at the time dependence of the amplitude or intensity of the light beams. It is known from spectroscopy that the line broadening processes present in a source are responsible for the fluctuations of the electric field and intensity occurring around an average value, on a timescale which is inversely proportional to the bandwidth of the spectrum. The unpredictable temporal fluctuations and the frequency spectrum of a source are complementary manifestations of the same physical properties of the radiating atoms that constitute the light source. The superposition of emission of many independent atoms that radiate with different frequencies and phases, is responsible for the generation of the chaotic light.

The temporal fluctuations of chaotic light can be investigated by considering the model of an N-excited-atom source which is broadened by collisions occurring at a rate $\gamma_{\text{coll}}[\text{s}^{-1}]$, that is, with an average time $\tau_c = 1/\gamma_{\text{coll}}$. The wave train radiated by a single atom is shown in **Figure 1**, where the electric field amplitude is schematically represented as a function of time. The total electric field radiated by the source is expressed as

$$E(t) = \sum E_i(t) = E_0 \exp(-2\pi i \nu_0 t) \times [\exp(i\phi_1(t)) + \exp(i\phi_2(t)) + \cdots + \exp(i\phi_N(t))]$$

where E_0 and ν_0 represent the characteristic amplitude and frequency of the field radiated by every atom and $\phi_i(t)$ $(i = 1, 2, \ldots, N)$ represents the random phase of each one of the uncorrelated atoms. The average spectrum is obtained by the Fourier decomposition of $E(t)$ and consists of a set of frequencies centered at ν_0.

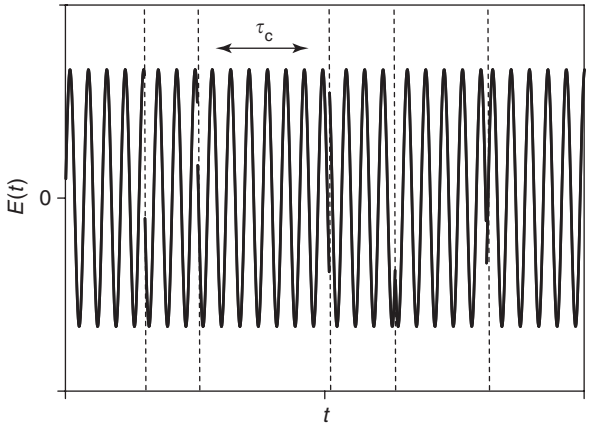


Figure 1 Sequence of wave trains for chaotic light.

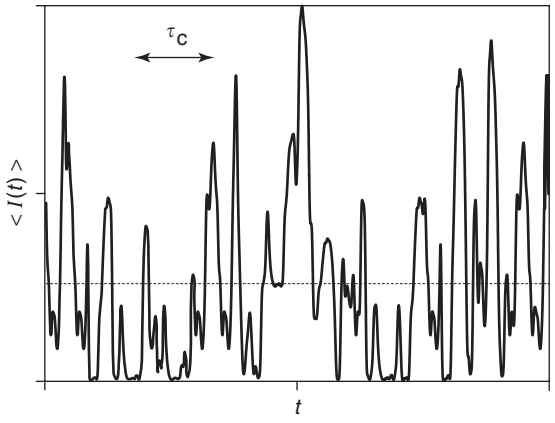


Figure 2 Temporal evolution of chaotic light.

The intensity of the beam is proportional to $|E(t)|^2$. It is a random function of time, varying on a timescale of the order of τ_c. The cycle-averaged intensity $I(t)$ is defined as the average of the intensity over a cycle of the oscillation at frequency v_0. In many cases, the corresponding fluctuations are too rapid to be detected, so that one is able to measure the average of the fluctuations over the response time of the detector. **Figure 2** shows the computer simulation of a typical temporal evolution of the cycle-averaged intensity for a chaotic light beam. The timescale is normalized to the coherence time τ_c. The average intensity $\langle I(t) \rangle$ is the result of the ensemble average over a large number of values of $I(t)$, measured over a period of time much longer than τ_c:

$$\begin{aligned}\langle I(t) \rangle &= \langle |E(t)|^2 \rangle \\ &= E_0^2 \langle |\exp(i\phi_1(t)) \\ &\quad + \exp(i\phi_2(t)) + \cdots \\ &\quad + \exp(i\phi_N(t))|^2 \rangle = NE_0^2\end{aligned}$$

This result is a consequence of the zero average contribution of the cross-terms between the phase factors of different radiating atoms. The dotted line in **Figure 2** indicates the value of $\langle I(t) \rangle$ for the fluctuating intensity.

First-Order Coherence

The first-order correlation function of the electric field at times t and $t + \tau$ is defined as

$$G^{(1)}(\tau) = \langle E^*(t)E(t+\tau) \rangle$$

It expresses how the fluctuations of the field $E(t + \tau)$ are affected by the value of field E at time t. It is found that $G^{(1)}(0) = I$ at $\tau = 0$. Furthermore, $G^{(1)}(-\tau) = G^{(1)*}(\tau)$, because of its Hermitian symmetry.

The degree of first-order temporal coherence is defined as

$$\begin{aligned}g^{(1)}(\tau) &= G^{(1)}(\tau)/G^{(1)}(0) \\ &= \langle E^*(t)E(t+\tau) \rangle / \langle E^*(t)E(t) \rangle\end{aligned}$$

a complex function, independent of the intensity, with the properties: $g^{(1)}(-\tau) = g^{(1)*}(\tau)$ and $0 \leqslant |g^{(1)}(\tau)| \leqslant 1$. The value of $|g(\tau)|$ is a measure of correlation between $E(t)$ and $E(t + \tau)$. In the ideal case of a monochromatic plane wave of amplitude $E_0, E(t) = E_0 \exp(2\pi i v_0 t)$. The degree of first-order temporal coherence becomes

$$g^{(1)}(\tau) = \exp(-2\pi i v_0 t)$$

and $|g(\tau)| = 1$, then $E(t)$ and $E(t + \tau)$ are completely correlated for all the values of τ.

The model of a collision-broadened light source can be assumed to calculate the degree of first-order coherence. The correlation function of the whole light beam is obtained from the single-atom contribution,

$$\langle E^*(t)E(t+\tau) \rangle = NE_0^2 \exp(-2\pi i v_0 \tau - \tau/\tau_c)$$

As a consequence the first-order coherence degree is

$$\begin{aligned}g^{(1)}(\tau) &= \exp(-2\pi i v_0 \tau - \tau/\tau_c) \\ &= \exp(-2\pi i v_0 \tau - |\tau|/\tau_c)\end{aligned}$$

with modulus

$$|g^{(1)}(\tau)| = \exp(-|\tau|/\tau_c)$$

Note that the exponential shape of $|g^{(1)}(\tau)|$ is consistent with the characteristic Lorentzian frequency spectrum valid for a collision-broadening mechanism, with coherence time τ_c.

Insofar as other broadening mechanisms are concerned, one can consider the case of chaotic light generated by atoms in a gas characterized by a spread in their velocities. The consequent Doppler effect leads to a Gaussian distribution of the emission frequencies. In this case the degree of first-order coherence is

$$g^{(1)}(\tau) = \exp[(-2\pi i v_0 \tau - \pi/2(\tau/\tau_c)^2]$$

with $\tau_c = \sqrt{\pi}/\Delta$ (Δ^2 corresponds to the variance of the Gaussian distribution).

The modulus of $g^{(1)}(\tau)$ for chaotic light, calculated in the two cases of Lorentzian and Gaussian frequency spectrum, is shown in **Figure 3**. The dashed line $g^{(1)}(\tau) = 1$ represents the case of a plane wave.

Light emitted from a partially coherent source is modeled as a sequence of random, uncorrelated wave packets, each one with a characteristic coherence length $l_c = c\tau_c$. The coherence times and the corresponding values of l_c are shown in **Table 1** for different

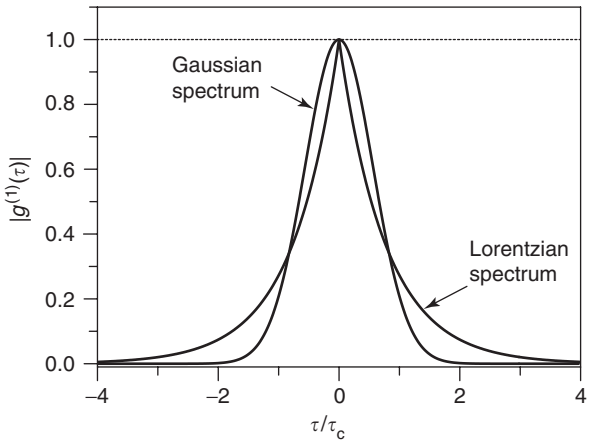


Figure 3 First-order coherence degree expressed as a function of units of τ_c for collision-broadened and Doppler-broadened chaotic light. Dashed line corresponds to the case of coherent light.

Table 1 Coherence times and coherence length of different light sources

Source	τ_c	l_c
White light (400–700 nm)	3.6 fs	1.08 μm
LED ($\lambda = 800$ nm, $\Delta\lambda = 50$ nm)	43 fs	0.01 mm
Spectral lamp (Hg)	10 ps	3 mm
Multimode CW He–Ne laser ($\lambda = 633$ nm)	0.67 ns	20 cm
Single-mode Nd : YAG laser ($\lambda = 1.06$ μm)	3 ms	900 m

light sources. The case of a single-mode laser, whose coherence length is of the order of several hundreds of meters, represents the best approximation of a monochromatic plane wave with $l_c = \infty$.

Similarly to the case of temporal correlation, one can adopt the same formalism to investigate the degree of first-order correlation of two different spatial points of the wave front, with coordinates $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$,

$$g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2)/\sqrt{[I(\boldsymbol{r}_1)I(\boldsymbol{r}_2)]}$$

with $0 \leqslant |g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2)| \leqslant 1$. For a given value of $\boldsymbol{r}_2$ and varying $\boldsymbol{r}_1$, the spatial extent of the region where $|g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2)|$ is greater than a reference value (e.g., 1/2 or 1/e) uniquely determines the radius of the coherence area. In the ideal case of a coherent plane wave $|g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2)| = 1$, whatever the distance $\boldsymbol{r}_1 - \boldsymbol{r}_2$. In this case the coherence area is infinitely large. In the general case, the first degree of coherence $g^{(1)}(\boldsymbol{r}_1, t_1, \boldsymbol{r}_2, t_2)$ can be defined for any pair of space–time points.

Depending on the value of $|g^{(1)}|$, a light can be classified as coherent ($|g^{(1)}| = 1$), incoherent ($|g^{(1)}| = 0$), or partially coherent ($0 \leqslant |g^{(1)}| \leqslant 1$). The coherence degree of a light beam is measured by first-order interferometry.

So far, a classical model of light has been adopted to describe first-order coherence processes. Also, in the case of light beams that need to be described in terms of quantum states of radiation, no difference is found between classical and quantum predictions for first-order interference experiments, as already stated. The possibility of distinguishing between classical and nonclassical light resides on second-order interference experiments.

Interference of Two Partially Coherent Waves

Consider two partially coherent waves $\boldsymbol{E}_1(\boldsymbol{r}, t)$ and $\boldsymbol{E}_2(\boldsymbol{r}, t)$, with intensities $\langle |\boldsymbol{E}_1(\boldsymbol{r}, t)|^2 \rangle$ and $\langle |\boldsymbol{E}_2(\boldsymbol{r}, t)|^2 \rangle$. The two waves correspond, for example, to the fields coming out of a two-arm interferometer, such as a Mach–Zehnder or Michelson scheme. A Mach–Zehnder interferometer, composed of two equal beam splitters (BS) and two total reflecting mirrors (M), is sketched in **Figure 4**. The light beam is incident on one of the two input ports of the first BS and is divided into two equal beams, with electric field $\boldsymbol{E}_1(t)$ and $\boldsymbol{E}_2(t)$, traveling through the two optical paths 1 and 2. The two beams meet together in the second BS, where interference between the two electric fields can occur, depending on the mutual phase conditions.

By the above assumptions, one has

$$g_{12} = \langle \boldsymbol{E}_1^* \cdot \boldsymbol{E}_2 \rangle / \sqrt{(I_1 I_2)}$$

The average intensity corresponding to the two-wave superposition is

$$\begin{aligned} I &= \langle |E_1 + E_2|^2 \rangle \\ &= I_1 + I_2 + 2\sqrt{(I_1 I_2)}\ \mathrm{Re}(g_{12}) \\ &= I_1 + I_2 + 2\sqrt{(I_1 I_2)}|g_{12}|\ \cos\phi \end{aligned}$$

where ϕ is the phase of g_{12}. It can be suitably adjusted by varying, within half of the wavelength $\lambda = c/\nu$, the delay between the two optical paths 1 and 2. Optical interference arises exactly from the third term of the above expression, that is, from the value of the mutual correlation degree, g_{12}. Two important cases are considered, depending on the limit values of g_{12}.

1. $|g_{12}| = 1$ (coherent waves) $\Rightarrow$ $I = I_1 + I_2 + 2\sqrt{(I_1 I_2)}\cos\phi$,
2. $|g_{12}| = 0$ (incoherent waves) $\Rightarrow$ $I = I_1 + I_2$.

In the general case of partially coherent waves, the intensity I is a sinusoidal function of ϕ; the larger the $|g_{12}|$, the larger the amplitude of the oscillation. In this case, the measurement of the interference

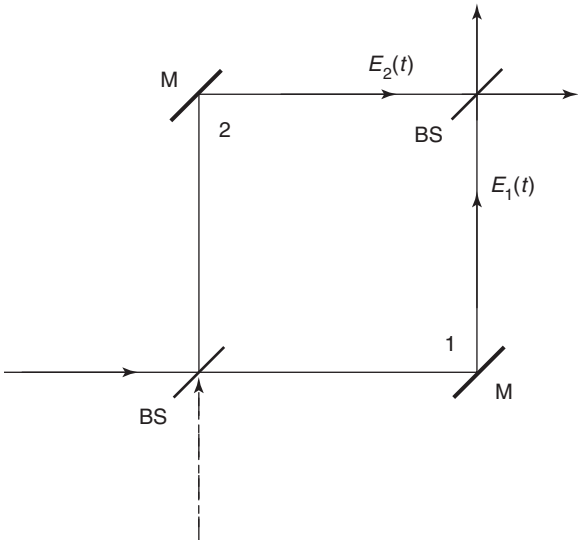


Figure 4 Mach–Zehnder interferometer.

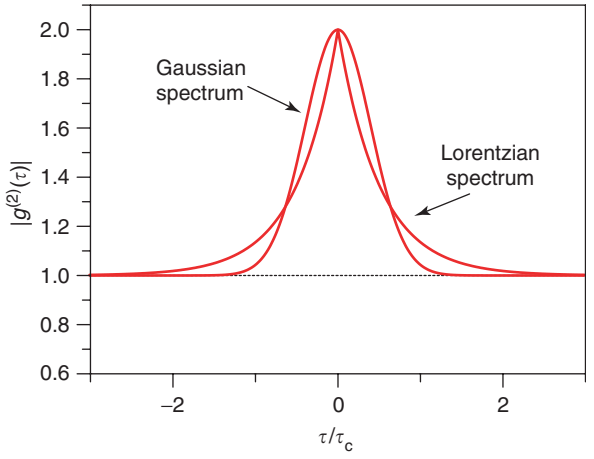


Figure 5 First-order coherence degree expressed as a function of units of τ_c for collision-broadened and Doppler-broadened chaotic light. Dashed line corresponds to the case of coherent light.

strength as a function of the optical path delay gives us all the necessary information about the temporal or spatial behavior of the correlation degree, allowing the measurement of the temporal coherence length or the coherence radius of the light beam. One can express the modulation depth by the visibility parameter,

$$V = (I_M - I_m)/(I_M + I_m)$$

with I_M and I_m corresponding to the maximum ($\phi = 0$) and minimum ($\phi = \pi$) value of the intensity, respectively.

By expressing the visibility as a function of $|g_{12}|$, one has

$$V = [2\sqrt{(I_1 I_2)}/(I_1 + I_2)]|g_{12}|$$

In a symmetrical interferometer ($I_1 = I_2$), $V \equiv |g_{12}|$. In this case, for $|g_{12}| = 1$, $I_M = 4I$ and $I_m = 0$.

Second-Order Coherence

As for the first-order coherence, consider two time measurements in which many pairs of samples of the cycle-averaged intensity are measured at a fixed delay τ. The classical definition of the intensity correlation function, valid either for chaotic or coherent light, corresponds to the average of the products of each pair of samples:

$$\begin{aligned} G^{(2)}(\tau) &= \langle I(t)I(t+\tau)\rangle \\ &= \langle \boldsymbol{E}^*(t)\boldsymbol{E}^*(t+\tau)\boldsymbol{E}(t+\tau)\boldsymbol{E}(t)\rangle \end{aligned}$$

The correlation function $G^{(2)}(\tau)$ can be normalized to the square of the long-time-averaged intensity $\langle I(t)\rangle$ in order to obtain the degree of second-order temporal coherence:

$$\begin{aligned} g^{(2)}(\tau) &= G^{(2)}(\tau)/G^{(2)}(0) \\ &= \langle I(t)I(t+\tau)\rangle / \langle \boldsymbol{E}^*(t)\boldsymbol{E}(t)\rangle^2 \end{aligned}$$

Because of symmetry properties, $g^{(2)}(\tau) = g^{(2)}(-\tau)$. It can be demonstrated that $g^{(2)}(\tau)$ satisfy the following relation at zero time delay, $\tau = 0$:

$$1 \leqslant g^{(2)}(0) \leqslant \infty$$

Since $g^{(2)}(\tau)$ is a positive function, for a finite time delay, $\tau \neq 0$, one has

$$0 \leqslant g^{(2)}(\tau) \leqslant \infty$$

Furthermore, since $\langle I(t)I(t+\tau)\rangle \leqslant \langle I(t)^2\rangle$, the degree of second-order coherence never exceeds its value for zero time delay. Hence, the following inequality holds:

$$g^{(2)}(\tau) \leqslant g^{(2)}(0)$$

For a large number N of radiating atoms, it can be demonstrated that

$$g^{(2)}(\tau) = 1 + |g^{(1)}(\tau)|^2$$

By considering the limiting values of $|g^{(1)}(\tau)|$ for $\tau = 0$ and $\tau \gg \tau_c$, for chaotic light, $g^{(2)}(0) = 2$ and $g^{(2)}(\tau \gg \tau_c) = 1$.

In the two cases of collision and Doppler-broadened light, one can write, respectively,

$$g^{(2)}(\tau) = 1 + \exp(-2|\tau|/\tau_c)$$

$$g^{(2)}(\tau) = 1 + \exp[-\pi/(\tau/\tau_c)^2]$$

Figure 5 shows the behavior of $g^{(2)}(\tau)$ of both the chaotic lights with Lorentzian and Gaussian frequency

distribution. Starting from $g^{(2)}(0) = 2$ at zero time delay, the intensities $I(t)$ and $I(t+\tau)$ become more and more uncorrelated by increasing τ and $g^{(2)}(\tau) \to 1$. In contrast to chaotic light, in the case of a classical stable wave, one has $g^{(2)}(\tau) = 1$ for any value of τ. This case is represented by the horizontal dashed line in **Figure 5**.

Similarly to the case of first-order coherence, the definition of second-order coherence degree is generalized to the spatial case,

$$g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \boldsymbol{E}^*(\boldsymbol{r}_1)\boldsymbol{E}^*(\boldsymbol{r}_2)\boldsymbol{E}(\boldsymbol{r}_1)\boldsymbol{E}(\boldsymbol{r}_2)\rangle / \langle |\boldsymbol{E}(\boldsymbol{r}_1)|^2\rangle \langle |\boldsymbol{E}(\boldsymbol{r}_2)|^2\rangle$$

and, in general, $g^{(2)}(\boldsymbol{r}_1, t_1, \boldsymbol{r}_2, t_2)$ can be defined for any pair of space–time points.

All the measurements of $g^{(2)}(\tau)$ are based on the original experimental apparatus invented by Brown and Twiss (BT) and are illustrated in **Figure 6**. Quasi-monochromatic radiation, properly selected in the single spatial mode, is incident upon a symmetric ($R = T$) optical beam splitter, so that half of the original intensity $I(t)$, $I_1(t)$, is directed to detector D_1, the other half, $I_2(t)$, to detector D_2. The electric signals of the detectors are used to determine the average of the product of $I_1(t)$ and $I_2(t+\tau)$ by means of a correlator. This operation allows to measure the intensity fluctuations. By making the assumption of a stationary field, $\langle I_1(t)\rangle = \langle I_2(t+\tau)\rangle$, and detectors with instantaneous response time, if these are placed at the same distance from the beam splitter, one has

$$g_{12}^{(2)}(\tau) = \langle I_1(t) I_2(t+\tau)\rangle / [\langle I_1(t)\rangle \langle I_2(t)\rangle] = \langle I(t) I(t+\tau)\rangle / \langle I(t)\rangle^2$$

By this scheme, any measurement affected by the temporal or spatial coherence of the light can be performed. Indeed, the original BT experiment was carried out to measure the angular diameter of stars by the observation of spatial coherence effects.

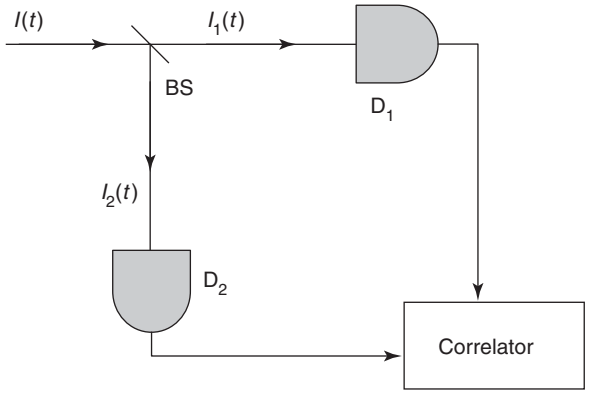


Figure 6 Brown–Twiss interferometer.

Photon Statistics of Coherent and Chaotic Light

The peculiar significance of all the measurements of degrees of second-order coherence resides on the correspondence between the classical and nonclassical radiation. So far the main aspects of optical coherence have been introduced in terms of the classical wave theory of light. The particle aspect of light is considered whenever it is possible to count photons (or rather, the photoelectrons ejected from a photoemissive surface by the absorption of photons). It is possible to demonstrate that in a typical photon-counting experiment, which allows one to measure the number of photoelectrons ejected by a light beam with average number of photons $\langle n\rangle$ during an appropriate time interval ΔT, the probability distribution $P_n(\Delta T)$ for n photoelectrons counted during the time ΔT is the Poisson distribution

$$P_n(\Delta T) = (\langle n\rangle^n / n!) \exp(-\langle n\rangle)$$

This expression applies either to a stationary coherent (laser) beam or to a chaotic radiation with coherence time $\tau_c \ll \Delta T$, in order to average the intensity fluctuations. In the opposite case, $\Delta T \ll \tau_c$, the probability of finding exactly n photons in the field, in thermal equilibrium at some temperature T, is governed by the Boltzmann factor

$$p(n) = \frac{[\exp(-nh\nu/k_B T)]}{[\sum_m \exp(-mh\nu/k_B T)]}$$

In this case, one obtains the so-called Bose–Einstein distribution,

$$P_n(\Delta T) = \langle n\rangle^n / (\langle n\rangle + 1)^{\langle n\rangle + 1}$$

It is clear from the above equations that a laser and a chaotic source may be characterized by the same average number, $\langle n\rangle$, of photons, the same frequency and bandwidth, and the same first-order coherence properties, but will have completely different photon statistics.

The mean square deviation of the photon number for the Poisson distribution is given by

$$\langle \Delta n^2\rangle = \langle n\rangle$$

while, for the Bose–Einstein distribution,

$$\langle \Delta n^2\rangle = \langle n\rangle^2 + \langle n\rangle$$

Chaotic light belongs to the category of super-Poissonian light, because it presents additional fluctuations to those of coherent light. One may ask if sub-Poissonian light ($\langle \Delta n^2\rangle < \langle n\rangle$) exists. This is the case of number state light; its genuine quantum properties will be illustrated in the next section. The

relative root mean square distributions for coherent and chaotic light are

$$\Delta n_{\text{rms}}/\langle n\rangle = 1\sqrt{\langle n\rangle}$$
$$\Delta n_{\text{rms}}/\langle n\rangle = \sqrt{(1+1/\langle n\rangle)}$$

From the above two expressions, one can conclude that the photon number fluctuations for a chaotic field can be much more pronounced than in the case of a laser field with increasing of $\langle n\rangle$.

Brown–Twiss Correlations and Nonclassical Light

The quantum definition of the degree of second-order coherence is defined analogously to the classical case by correlating two intensities at two different space–time points. Consider a light beam containing a mean number, $\langle n\rangle$, of photons, traveling in a single mode of the electromagnetic field, incident upon the BS (**Figure 6**). The degree of second-order coherence is expressed in terms of the mean and mean square photon number as

$$g^{(2)}(\tau) = \langle n(n-1)\rangle/\langle n\rangle^2$$
$$= [\langle n^2\rangle - \langle n\rangle]/\langle n\rangle^2$$

This expression can be explained as follows: the probability of counting a photon is proportional to the number of photons, n, in the field. Since the number of photons available in a second measurement is reduced by one unit, the probability of counting a second photon is proportional to $n-1$. As a consequence, the two-photon counting probability is proportional to $\langle n(n-1)\rangle = \langle n^2\rangle - \langle n\rangle$. Since the photon number variance is $(\Delta n)^2 \equiv \langle n^2\rangle - \langle n\rangle^2$, one can express the degree of second-order coherence as

$$g^{(2)}(\tau) = 1 + [\langle \Delta n\rangle^2 - \langle n\rangle]/\langle n\rangle^2$$

Since $(\Delta n)^2 \geqslant 0$, it must be that $\langle n^2\rangle \geqslant \langle n\rangle^2$. Hence, the following inequality holds:

$$g^{(2)}(\tau) \geqslant 1 - 1/\langle n\rangle \quad (\langle n\rangle \geqslant 1)$$

a result independent of the value of τ, including $\tau = 0$. It follows that one can write $g^{(2)}(0)$ instead of $g^{(2)}(\tau)$.

The signature of nonclassical light is given by the range of values, $1 - 1/\langle n\rangle \leqslant g^{(2)}(0) < 1$ accessible to the degree of second-order coherence, with $g^{(2)}(0) = 0$ for $\langle n\rangle < 1$.

For $\langle n\rangle = 1$, $g^{(2)}(0) = 0$, a result arising from the impossibility of the incident photon to be splitted by the beam splitter. As a consequence, no correlation exists between the signals of detectors D_1 and D_2 in this case. In the limit case of a beam with exactly n photons (photon number state), one has $\langle n\rangle = n$, and

$$g^{(2)}(0) = (n-1)/n \quad (n \geqslant 2)$$
$$g^{(2)}(0) = 0 \quad (n = 0, 1)$$

Figure 7 shows the degree of second-order coherence as a function of $\langle n\rangle$ for different kinds of light.

Finally, consider once more the theoretical predictions for chaotic and coherent light, **Figure 5**. The Brown–Twiss correlation at $\tau = 0$ for chaotic light, $g^{(2)}(0) > 1$ is also referred to as photon bunching, a term which indicates that photons tend to arrive simultaneously at detectors D_1 and D_2. Note that no photon bunches occur for coherent light, $g^{(2)}(0) = 1$. On the contrary, photon antibunching occurs for nonclassical light, such as photon number states, which lies in the region $0 \leqslant g^{(2)}(0) < 1$. In this case photons tend to arrive well spaced in time.

Quantum light has great importance in modern quantum optics because of its applications whenever transmission of signals through single photons is required, as in quantum cryptography and quantum computing. The single-photon regime can be attained by strongly attenuating a laser beam to ensure that the probability of having more than one photon becomes negligible. However, such attenuated beams differ from true single photons because of the intrinsic Poissonian statistics of the photon. Significant advances in the generation of single-photon states have been obtained in recent years, such as quantum dots in microcavities, parametric down-converters, falling neutral atoms and trapped ions in cavities, defects in diamond nanocrystals, and single molecules in a solid matrix.

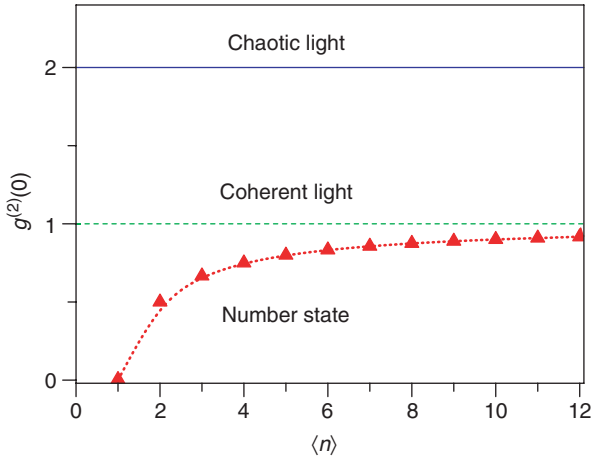


Figure 7 Degree of second-order coherence expressed as a function of $\langle n\rangle$ for chaotic light, coherent light, and quantum light.

Two-Photon States

As noted already, the optical beam splitter is the fundamental element for analyzing the statistical properties of light. Interesting interference effects occur when radiation given by two-photon states enters a beam splitter. These states, of high relevance for quantum information applications, can be generated by the nonlinear optical process of parametric down-conversion. Here two photons of frequency ν_1 and ν_2, also called as the signal and idler, are created by the annihilation of a "pump" photon of higher frequency ν_p, belonging to a laser beam which shines a crystal with no inversion symmetry, that is, with nonlinear dielectric susceptivity $\chi^{(2)} \neq 0$. Two-photon generation must obey the energy and momentum conservation laws:

$$\nu_p = \nu_1 + \nu_2; \qquad \boldsymbol{k}_p = \boldsymbol{k}_1 + \boldsymbol{k}_2$$

where $\boldsymbol{k}_p$, $\boldsymbol{k}_1$, and $\boldsymbol{k}_2$ correspond to the wave vectors of the pump, signal, and idler photons, respectively. Depending on the characteristics of the crystal, signal and idler photons can travel along the same direction or along different directions, and may have parallel or orthogonal polarizations.

The partition process on a beam splitter of two equal photons ($\nu_1 = \nu_2 = \nu_p/2$) traveling along two symmetric directions, a and b, can be studied by means of the Ou–Mandel interferometer, which is shown in **Figure 8**. Two cases are possible:

1. The two photons meet the beam splitter at different times. In this case, the probability $P(n_c, n_d)$, of finding n_c photon on mode c and n_d photons on mode d follows the classical binomial distribution,

   $$P(2,0) = R^2, \quad P(1,1) = 2RT, \quad P(0,2) = T^2$$

   For a symmetric beam splitter, $P(0,2) = P(2,0) = \frac{1}{2}P(1,1)$.
2. The two photons arrive simultaneously at the beam splitter. In this case, the output probability distribution is completely different:

   $$P(2,0) \equiv P(0,2) = 2RT, \quad P(1,1) = (R-T)^2$$

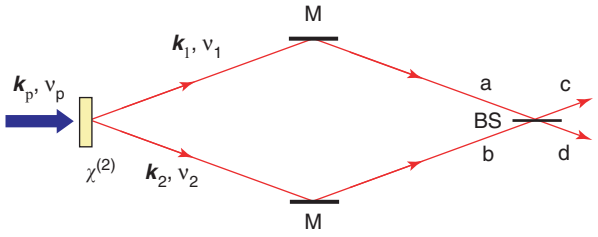


Figure 8 Ou–Mandel interferometer.

It is interesting to note that, for $R = T$, $P(1,1) = 0$, that is, in this particular case the photons emerge together in the same output arm, either c or d. The case of two-photon interference has no analog in the classical theory of light. Indeed, with two laser beams the partition probability is always binomial.

See also: Electrodynamics: Continuous Media; Electrodynamics: Quantum Electrodynamics; Laser Radiation Sources; Optical Instruments; Statistical Mechanics: Quantum.

PACS: 42.50.Ar; 42.25.Hr; 42.25.Kb

Further Reading

Brown RH and Twiss RQ (1956) Correlation between photons in two coherent beams of light. *Nature* 177: 27.

Hong CK, Ou ZY, and Mandel L (1987) Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters* 59: 2004.

Loudon R (2000) *The Quantum Theory of Light*, 3rd edn. New York: Oxford University Press.

Mandel L and Wolf E (1965) Coherence properties of optical fields. *Reviews of Modern Physics* 37: 231.

Milonni PW and Eberly JR (1988) *Lasers*. New York: Wiley.

Saleh BEA and Teich MC (1991) *Fundamentals of Photonics*. New York: Wiley.

Zou XT and Mandel L (1990) Photon bunching and sub-Poissonian photon statistics. *Physical Review A* 41: 476.

Photonic Bandgap Materials, Electronic States of

M Agio and L C Andreani, Università degli Studi di Pavia, Pavia, Italy

Introduction

Photonic crystals are materials whose dielectric function is periodic in one, two, or three spatial dimensions (1D, 2D, 3D). The propagation properties of photons in periodic media are analogous to those of electrons in crystalline solids, as covered by several sections of the encyclopedia. Translational invariance leads to the validity of the Bloch theorem (called Floquet theorem in optics) and to the formation of allowed and forbidden frequency regions. The allowed states of the electromagnetic field are organized into "photonic

bands," in analogy to electron bands in periodic solids. A frequency window for which light cannot propagate is named a "photonic gap" and periodic dielectric media are often called "photonic bandgap materials."

The concept of a 1D periodic dielectric medium has been a familiar one in optics, where it is known as a distributed Bragg reflector (or dielectric mirror) with well-defined stop bands. However, a stop band (or photonic gap) in 1D forbids propagation of light only in a limited angular cone. The field of photonic crystals is considered to have started in 1987 with seminal papers dealing with the control of spontaneous emission and with light localization in three dimensions. Indeed, the presence of a full photonic bandgap in 3D leads to the suppression of vacuum fluctuations of the electromagnetic field and, therefore, to the inhibition of spontaneous emission of a light source at frequencies within the bandgap. Moreover, the presence of disorder may lead to a weak localization of light at frequencies close to a band edge. Photonic crystals are now of great interest for basic physical properties related to the control of light propagation, radiation–matter interaction and quantum electrodynamic effects, localization and microcavities, nonlinear properties as well as for applications to optoelectronic and photonic devices such as lasers, optical fibers, filters, dispersion compensators, and integrated optical interconnects.

The article starts with the general properties of the photonic band structure and discusses the main theoretical approaches for calculating photonic bands and optical properties. Then photonic bands in three and two dimensions as well as photonic crystals embedded in planar waveguides (also known as photonic crystal slabs) are described. Defect states in photonic crystals are then considered. This section concludes with a few remarks about methods for measuring the photonic band dispersion.

Photonic Band Structure: General Properties

Maxwell equations for the electromagnetic field in matter with neither free charges nor currents can be cast into a second-order equation for the harmonic components of the electric field

$$\nabla \times \nabla \times \boldsymbol{E}(\boldsymbol{r}) = \left(\frac{\omega^2}{c^2}\right)\varepsilon(\boldsymbol{r})\boldsymbol{E}(\boldsymbol{r}) \qquad [1]$$

or of the magnetic field

$$\nabla \times \left[\frac{1}{\varepsilon(\boldsymbol{r})}\nabla \times \boldsymbol{H}(\boldsymbol{r})\right] = \frac{\omega^2}{c^2}\boldsymbol{H}(\boldsymbol{r}) \qquad [2]$$

Here, nonmagnetic media ($\mu(\boldsymbol{r}) = 1$) are assumed. The first equation should be combined with the divergence equation $\nabla \cdot \varepsilon(\boldsymbol{r})\boldsymbol{E}(\boldsymbol{r}) = 0$, while the second one requires $\nabla \cdot \boldsymbol{H}(\boldsymbol{r}) = 0$. The equation for the magnetic field is normally taken as the starting point for photonic band structure computations for two reasons. First, it has the form of an eigenvalue problem $\boldsymbol{\Theta}\boldsymbol{H}(\boldsymbol{r}) = (\omega^2/c^2)\boldsymbol{H}(\boldsymbol{r})$, with $\boldsymbol{\Theta}$ being a Hermitian operator. Second, the divergence equation for the magnetic field does not contain the dielectric function and is easier to implement. Working with a Hermitian eigenvalue problem is convenient from a computational point of view and has strong analogies with the matrix formulation of quantum mechanics.

The basic electromagnetic equations have a very useful property of "scale invariance." If $\boldsymbol{H}(\boldsymbol{r})$ is a solution at frequency ω corresponding to the dielectric constant $\varepsilon(\boldsymbol{r})$, then the scaled system with dielectric constant $\varepsilon'(\boldsymbol{r}) = \varepsilon(\boldsymbol{r}/s)$ has a solution $\boldsymbol{H}'(\boldsymbol{r}) = \boldsymbol{H}(\boldsymbol{r}/s)$ with frequency $\omega' = s\omega$. Reducing the length scale of the system by a factor s leads to eigenfrequencies that are multiplied by the same factor. Thus, there is no fundamental length scale for the photonic problem (unlike for electrons in solids, for which the Bohr radius is the natural unit of length).

A periodic dielectric medium (in 3D, say) is invariant under translations by vectors $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$ ($n_1, n_2, n_3 \in Z$), where $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ are primitive vectors and the set of all $\boldsymbol{R}$'s forms a "Bravais lattice." The dielectric constant satisfies the relation

$$\varepsilon(\boldsymbol{r} + \boldsymbol{R}) = \varepsilon(\boldsymbol{r}) \qquad [3]$$

and the magnetic field has the form implied by the Bloch–Floquet theorem:

$$\boldsymbol{H}_{n\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{H}_{n\boldsymbol{k}}(\boldsymbol{r}) \qquad [4]$$

where $\boldsymbol{k}$ is the Bloch vector (which may be restricted to the first Brillouin zone or Wigner–Seitz cell of the reciprocal lattice) and n is a discrete band index. The frequency eigenvalues have the form $\omega = \omega_n(k)$ and the set of all frequencies for $\boldsymbol{k}$ spanning the first Brillouin zone is called a photonic band. In addition to translational invariance leading to Bloch vector conservation, discrete rotational symmetries of the lattice may also be used to analyze the photonic bands, again in close similarity to the electronic problem. The density of states (DOS) is defined as

$$N(\omega) = \sum_n \int_{\mathrm{BZ}} \delta(\omega - \omega_n(\boldsymbol{k}))\mathrm{d}^3\boldsymbol{k} \qquad [5]$$

and is strongly modified as compared to the photonic DOS in vacuum or in a homogeneous medium, which increases with the square of the frequency.

Theoretical Methods

The most common method for calculating photonic band structures consists of expanding the magnetic field on a basis of plane waves that satisfy the Bloch theorem and the divergence equation:

$$\boldsymbol{H}_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G},\lambda} c_n(\boldsymbol{k}+\boldsymbol{G},\lambda)\hat{\varepsilon}(\boldsymbol{k}+\boldsymbol{G},\lambda)\mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} \quad [6]$$

where G are reciprocal lattice vectors and $\hat{\varepsilon}(\boldsymbol{k}+\boldsymbol{G},\lambda), \lambda = 1,2$ are unit and mutually orthogonal polarization vectors perpendicular to $\boldsymbol{k}+\boldsymbol{G}$. The second-order eqn [2] is transformed into the matrix equation

$$\sum_{\boldsymbol{G}'\lambda'} \boldsymbol{H}_{\boldsymbol{G}\lambda,\boldsymbol{G}'\lambda'} c_n(\boldsymbol{k}+\boldsymbol{G}',\lambda') = \frac{\omega^2}{c^2} c_n(\boldsymbol{k}+\boldsymbol{G},\lambda) \quad [7]$$

where

$$\boldsymbol{H}_{\boldsymbol{G}\lambda,\boldsymbol{G}'\lambda'} = |\boldsymbol{k}+\boldsymbol{G}||\boldsymbol{k}+\boldsymbol{G}'| \times \varepsilon^{-1}(\boldsymbol{G},\boldsymbol{G}') \begin{bmatrix} \hat{\varepsilon}_2\cdot\hat{\varepsilon}'_2 & -\hat{\varepsilon}_2\cdot\hat{\varepsilon}'_1 \\ -\hat{\varepsilon}_1\cdot\hat{\varepsilon}'_2 & \hat{\varepsilon}_1\cdot\hat{\varepsilon}'_1 \end{bmatrix} \quad [8]$$

and $\varepsilon^{-1}(\boldsymbol{G},\boldsymbol{G}') \equiv \varepsilon^{-1}(\boldsymbol{G}-\boldsymbol{G}')$ is the Fourier transform of $\varepsilon^{-1}(\boldsymbol{r})$. Numerically, the linear eigenvalue problem is solved by retaining a finite number of reciprocal lattice vectors and using standard matrix diagonalization routines or more efficient methods based on fast Fourier transforms and iterative eigensolvers. The accuracy of the method depends on the number of plane waves in the basis and on their ability to reproduce the spatial variations of the magnetic field in the periodic structure and Maxwell boundary conditions when dielectric discontinuities are present. It turns out that convergence of the method is substantially improved when $\varepsilon^{-1}(\boldsymbol{G},\boldsymbol{G}')$ is obtained by first calculating $\varepsilon(\boldsymbol{G},\boldsymbol{G}') \equiv \varepsilon(\boldsymbol{G}-\boldsymbol{G}')$ as the Fourier transform of $\varepsilon(\boldsymbol{r})$ and inverting the resulting matrix numerically. Still, convergence of the method should always be checked and it may become problematic in special situations (e.g., high dielectric discontinuities and/or complex bases in the unit cell with close-packing dielectric spheres).

Just like plane-wave expansion, other techniques for photonic band structure calculations are borrowed from corresponding methods for the electronic problem. For example, in the case of spherical (for 3D) or circular bases (for 2D), the Korringa–Kohn–Rostoker method based on expanding the Green function in the basis of symmetry-adapted functions satisfying appropriate boundary conditions can be applied. This method has the advantage that it can easily incorporate frequency dispersion of the dielectric function (thus it can be used, for example, for photonic crystals containing metallic components), but it can be applied easily only to structures with bases of special symmetry.

The calculation of photonic bands assumes ideal structures, that is, infinitely extended in space and without imperfections. On the other hand, real samples are always finite and carry a certain degree of imperfection. Furthermore, there are quantities such as the transmission and reflection coefficients that manifest the existence of photonic bandgaps and are relatively easy to obtain experimentally. Among the numerical methods that are able to compute transmission and reflection, the most widely used are the transfer matrix, the scattering matrix, and the finite-difference time-domain method.

The transfer matrix method works in real space. The Maxwell curl equations are discretized on a mesh of subcells, so that the field propagation in the structure is defined by a nearest-neighbor interaction. The matrix of the resulting linear system of equations is called a "transfer matrix," as it transfers the fields from a subcell to another. Since the interaction is limited to the first neighbors, the matrix is sparse. Propagation across the whole structure is obtained by multiplication of sparse matrices. Once the transfer matrix is obtained, the calculation of transmission and reflection is straightforward. It is worth mentioning that when the system is periodic, the eigenvalues of the transfer matrix of a single unit cell yield an alternative method for computing the photonic band structure. This is especially useful when dealing with dispersive media.

The scattering matrix method is more suitable for the study of photonic crystal slabs (see description below). It computes reflection, transmission, and diffraction coefficients for propagation along a direction perpendicular or parallel to the waveguide plane. Maxwell equations are first solved in each layer by plane wave expansion. Then, using the scattering matrix (S), instead of the transfer matrix, boundary conditions are imposed at each interface. The S-matrix connects the ingoing amplitudes to the outgoing amplitudes. The scattering matrix $S(0,N)$ of the whole structure is constructed starting from the scattering matrix of the first interface $S(0,1)$, by applying the recursion formula $S(0,l) -> S(0,l+1)$. The S-matrix yields better numerical stability than the transfer matrix.

When working with more complex systems, with imperfections or explicitly designed defects, a more flexible numerical approach is required. This is fulfilled by the finite-difference time-domain (FDTD) method. The derivatives in the Maxwell curl equations

$$\nabla\times\boldsymbol{E} = -\frac{1}{c}\frac{\partial\boldsymbol{B}}{\partial t}, \qquad \nabla\times\boldsymbol{H} = +\frac{1}{c}\varepsilon(\boldsymbol{r})\frac{\partial\boldsymbol{E}}{\partial t} \quad [9]$$

are discretized by central finite differences, both in space and time. The electromagnetic field is thus defined on two meshes in space and time, one for the $\boldsymbol{E}$ field and one for the $\boldsymbol{H}$ field. These meshes are interlaced in order to have each $\boldsymbol{E}$ field position surrounded by nearest neighbor $\boldsymbol{H}$ field positions and vice versa. The equations are recast in the so-called "Yee algorithm," which is illustrated here for a two-dimensional case:

$$\boldsymbol{E}_x|_{i,j+1/2}^{n+1/2} = \boldsymbol{E}_x|_{i,j+1/2}^{n-1/2} + \frac{c\Delta t}{\varepsilon_{i,j+1/2}\Delta y}\left(\boldsymbol{H}_z|_{i,j+1}^{n} - \boldsymbol{H}_z|_{i,j}^{n}\right) \quad [10]$$

$$\boldsymbol{E}_y|_{i+1/2,j}^{n+1/2} = \boldsymbol{E}_y|_{i+1/2,j}^{n-1/2} - \frac{c\Delta t}{\varepsilon_{i+1/2,j}\Delta x}\left(\boldsymbol{H}_z|_{i+1,j}^{n} - \boldsymbol{H}_z|_{i,j}^{n}\right) \quad [11]$$

$$\begin{aligned}\boldsymbol{H}_z|_{i,j}^{n+1} = \boldsymbol{H}_z|_{i,j}^{n} &- \frac{c\Delta t}{\Delta x}\left(\boldsymbol{E}_y|_{i+1/2,j}^{n+1/2} - \boldsymbol{E}_y|_{i-1/2,j}^{n+1/2}\right) \\ &+ \frac{c\Delta t}{\Delta y}\left(\boldsymbol{E}_x|_{i,j+1/2}^{n+1/2} - \boldsymbol{E}_x|_{i,j-1/2}^{n+1/2}\right)\end{aligned} \quad [12]$$

Looping on i, j (space), and n (time) simulates the propagation of the electromagnetic field through the structure. Once the Yee algorithm has been implemented, the discretization of the dielectric constant $\varepsilon(r)$ is the only step that needs to be adapted to the specific simulation. This is a great advantage of FDTD with respect to the transfer matrix method. A Gaussian pulse is sent along a chosen direction, while detection lines collect the incident and outgoing power. The transmission coefficient is simply the ratio of outgoing to incident power. The reflection coefficient can be obtained in a similar manner. When the system is periodic, it is also possible to obtain the band structure from an FDTD simulation. Specific implementations of the FDTD method have been developed to treat media with strong frequency dispersion, such as metallic or metallo-dielectric systems.

Three-Dimensional Photonic Crystals

The first photonic structure that has been shown to possess a complete bandgap in three dimensions is the diamond lattice of dielectric spheres in air or air spheres in a dielectric material. The photonic gap opens provided the dielectric constant is large enough and it exists in a wide range of filling factors. Another simple 3D structure with a complete bandgap is the f.c.c. lattice of air spheres in a dielectric material: this is often called the "inverse opal" structure. An example of photonic bands and DOS for this structure is shown in **Figure 1**. The frequencies are given in dimensionless units $\omega a/(2\pi c) \equiv a/\lambda$, where a is the lattice constant, and as such they are invariant under a scale transformation. A full bandgap forms between the eighth and ninth bands with a relative width $\Delta\omega_g/\omega_g \approx 4\%$. On the other hand, no gap is formed between lower-lying bands, but a pseudogap with low photonic DOS occurs around $a/\lambda = 0.52$: a degeneracy between the second and the third bands at the W point and along the XU line prevents a gap from opening.

The inverse opal is derived from the (direct) opal structure, that is, an f.c.c. lattice of dielectric spheres in air: indeed, the opal is a natural gem consisting of a close-packed lattice of silica spheres with an f.c.c. arrangement. Artificial opals can be obtained

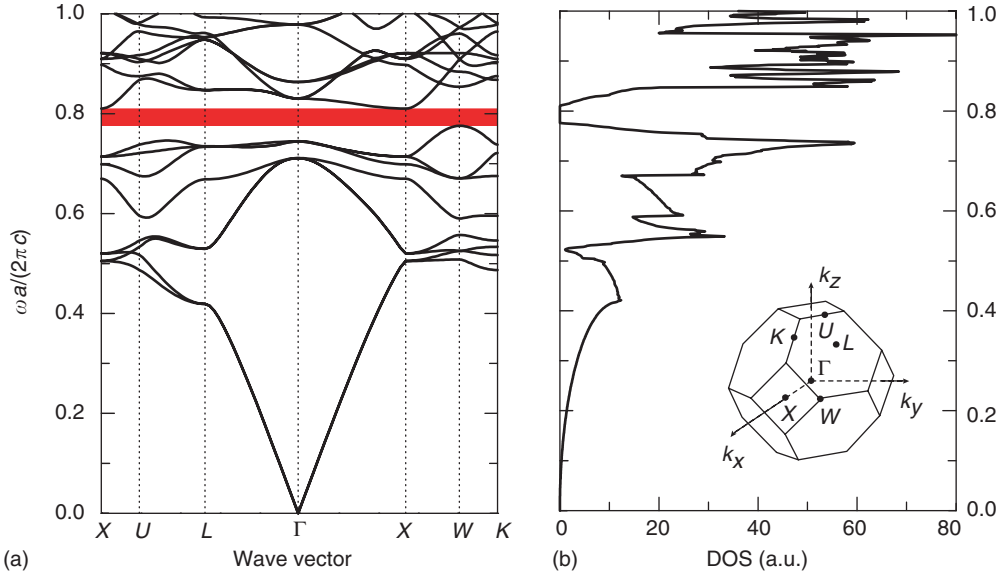


Figure 1 (a) Photonic bands and (b) DOS for an f.c.c. structure of close-packing air spheres in a dielectric material with $\varepsilon = 12$. The shaded area in (a) denotes the photonic gap. Inset: Brillouin zone and symmetry points of the f.c.c. lattice.

by self-assembling of spheres in a colloidal solution. The voids are then filled with a high-index material (Si or TiO_2) and the template is subsequently removed to obtain the inverse structure. This procedure allows one, in principle, to realize a photonic crystal with a complete bandgap in the optical region. In real systems, however, a small amount of disorder is always present: the gap shown in **Figure 1** is not very robust as it is relatively narrow and is formed between higher-lying bands, which are very sensitive to disorder.

Many other structures possessing a complete photonic gap in three dimensions have been studied. Two of them are shown in **Figure 2**. The structure in **Figure 2a**, called the "Yablonovite" (from the name of its proponent, E Yablonovitch) is obtained by drilling three sets of cylindrical holes with a triangular pattern, at 35.26° from the surface normal and at 120° with respect to each other. It represents a distortion of the diamond structure of air spheres in a dielectric material, where the surface is a (1 1 1) plane of the crystal and the holes correspond to [1 1 0] open channels of the diamond structure. The Yablonovite can be realized, in principle, at infrared and optical wavelengths by deep lithography. The structure in **Figure 2b**, often called "woodpile," is obtained by stacking four alternate layers of dielectric rods: starting from a first layer, the second layer is rotated by 90° with respect to the first, and the third layer is shifted by $a/2$ (a is the lattice constant) with respect to the first one. The woodpile structure can be realized by a series of stacking steps. The fabrication procedures of Yablonovite and of woodpile are examples of "top-down" and "bottom-up" approaches, respectively.

Two-Dimensional Photonic Crystals

2D photonic crystals have a dielectric constant that is periodic in a plane (xy) and homogeneous in the third (z) direction. The most interesting situation is when light propagates in the xy plane, that is, for a vertical wave vector $k_z = 0$. The system is invariant under specular reflection with respect to the mirror plane xy; thus the electromagnetic eigenmodes can be classified as even or odd with respect to this mirror symmetry. Even states have nonvanishing field components (E_x, E_y, H_z) and are called H- (or TE-) polarized modes, while odd states have nonzero field components (H_x, H_y, E_z) and are called E- (or TM-) polarized. Unlike in the 3D case, there is a strong polarization dependence of the photonic bands.

Most 2D structures that have been studied are based on a square or on a hexagonal Bravais lattice. In general, the most favorable situation for a complete photonic gap is to have a structure with the highest possible rotational symmetry: this helps in achieving an overlap between bandgaps along different directions of the Brillouin zone. The square Bravais lattice has a fourfold symmetry axis, while the hexagonal lattice has a sixfold rotational axis that is the maximum possible symmetry for periodic Bravais lattices. Higher rotational symmetries are possible for nonperiodic structures, such as quasicrystals.

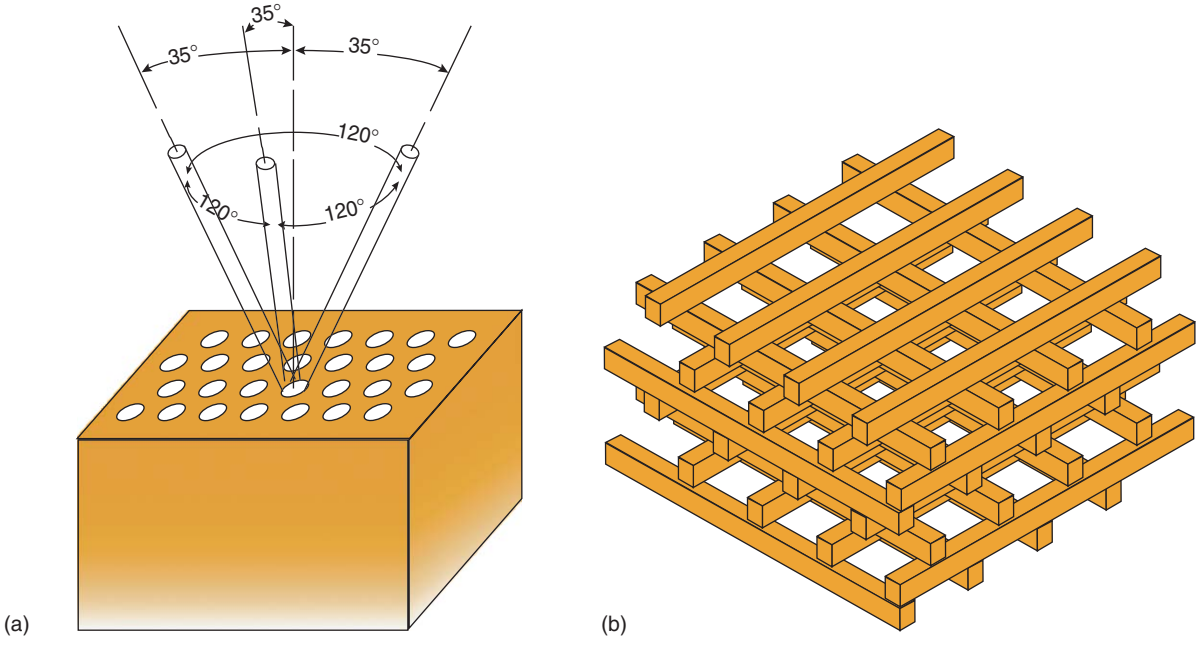


Figure 2 Schematic plots of the Yablonovite (a) and woodpile (b) photonic structures.

Figure 3 shows two examples of 2D photonic bands in structures with a full bandgap for all directions and polarizations, namely the triangular lattice of air holes in a dielectric medium (**Figure 3a**) and the graphite or honeycomb lattice of dielectric pillars in air (**Figure 3b**). Both structures are based on the 2D hexagonal Bravais lattice. The triangular lattice of holes has a large gap between the first and the second

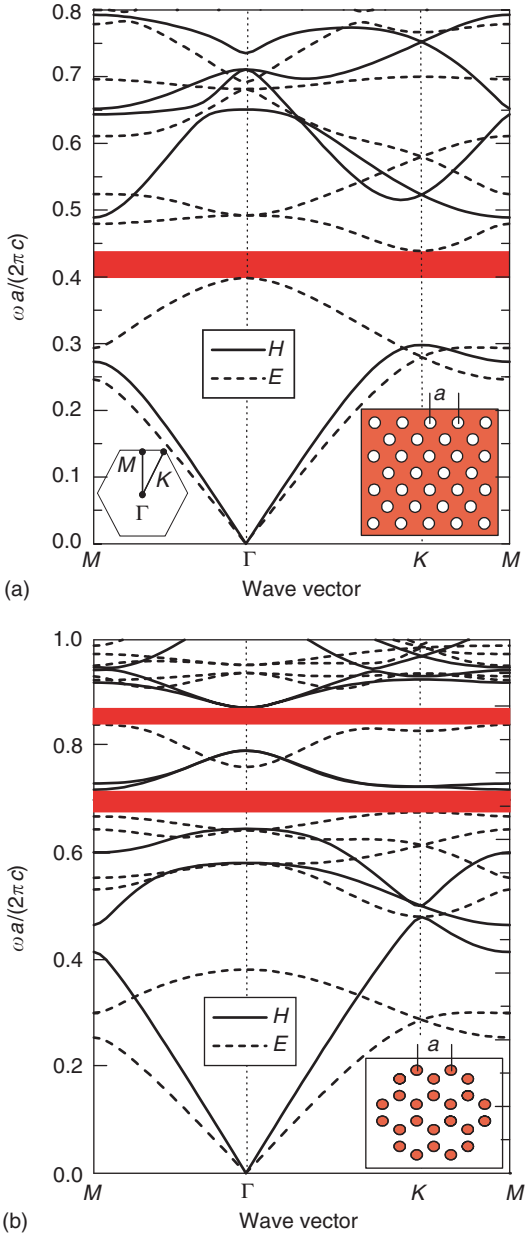


Figure 3 Photonic bands of (a) a triangular lattice of air holes with radius $r/a = 0.45$ in a dielectric material and (b) a graphite lattice of dielectric pillars with radius $r/a = 0.18$ in air. The dielectric material has $\varepsilon = 12$ in both cases. Solid (dashed) lines represent *H*- (*E*-) polarized modes: they are even (odd) with respect to a horizontal mirror plane. The shaded areas denote the photonic gaps. Insets: structures and 2D Brillouin zone.

H-band and a smaller gap between the second and the third *E*-band, which overlap in the region shown by a dashed area in **Figure 3a**. For the assumed value of the dielectric constant ($\varepsilon = 12$), the *H*-gap is found to open for hole radii $r/a > 0.16$, while the *E*-gap opens only for hole radii $r/a > 0.4$; therefore, a polarization-independent gap exists only for relatively high values of the air fraction. The graphite lattice of pillars, instead, has two complete bandgaps between higher-lying bands. Notice that the graphite lattice of pillars supports a large gap between the second and third *E*-bands, which however does not overlap any *H*-gap.

As a general rule, structures with a connected dielectric lattice (like that of **Figure 3a**) tend to have larger gaps for *H*-polarization, while structures with disconnected dielectric pillars (like that of **Figure 3b**) tend to support larger *E*-gaps. This criterion, which can be traced back to Maxwell boundary conditions for the nonvanishing field components, is generally followed for both square and hexagonal Bravais lattices. The overlap of *H*- and *E*-gaps is a stronger condition that occurs only in a few cases, those shown in **Figure 3** being the most common examples. In real systems, the complete gap shown in **Figure 3a** is easier to realize, because it is lower in frequency and less sensitive to disorder. For technological applications, a polarization-sensitive gap is often sufficient and the commonly used structure is the triangular lattice of air holes for smaller air fractions than shown in **Figure 3a**, in order to reduce radiation losses.

2D photonic structures can be defined by a top-down procedure based on lithography and etching. The prototype of a 2D photonic crystal is macroporous silicon, which is obtained by wet etching of Si in an electrochemical cell. It should be remembered that the photonic bands such as those shown in **Figure 3**, refer only to propagation in the *xy* plane: in a system that can be considered as homogeneous in the *z*-direction, the bands retain an out-of-plane dispersion and light is free to propagate in the vertical direction. This mechanism is exploited in a new class of optical fibers, which are known as holey- or photonic crystal fibers.

Photonic Crystal Slabs

In order to achieve a better control of light propagation, a 2D photonic structure can be embedded in a planar (slab) dielectric waveguide, thereby realizing a "photonic crystal slab." A few possible structures are shown in **Figure 4**. The theoretically simplest system is the "air bridge" (**Figure 4a**), which consists of a self-standing patterned dielectric membrane surrounded

by air. The air bridge has the highest refractive index contrast between the core and the cladding. **Figure 4b** exemplifies the silicon-on-insulator (SOI) structure, consisting of an air/Si/SiO_2 waveguide in which only the Si layer is patterned. A system with a weak refractive index contrast is shown in **Figure 4c**, and it can be realized with GaAs/AlGaAs or with InP/InGaAsP waveguides: both the core and the cladding need to be patterned in order to maintain the guiding properties of the slab. The fabrication of photonic crystal slabs requires epitaxial growth, lithography, and dry (reactive-ion) etching steps.

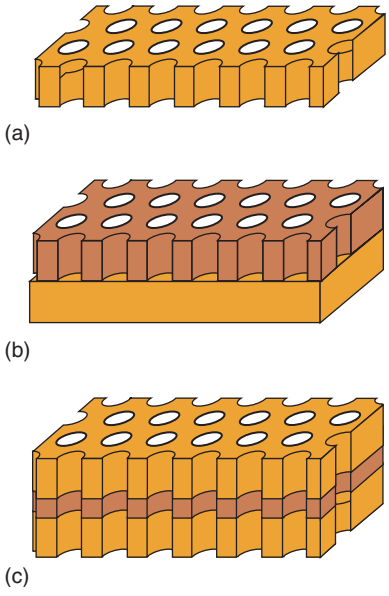


Figure 4 Schematic plots of photonic crystal slabs: (a) air bridge, or self-standing membrane, (b) SOI, where only the Si core layer is patterned, and (b) AlGaAs/GaAs/AlGaAs, where all layers are patterned.

Photonic crystal slabs are characterized by the light-line issue. Only photonic modes that lie below the light dispersion of the cladding material(s) in the $k-\omega$ plane are truly guided and stationary, while modes that lie above the cladding light line(s) are radiative and only quasi-guided. For a specified value of the in-plane wave vector, the truly guided modes lie in a frequency region of the discrete spectrum, while quasi-guided modes lie in the region of the continuous spectrum and appear as resonances: their line width can be related to an imaginary part of the mode frequency.

An example of mode dispersion in a photonic crystal slab is shown in **Figure 5** and it refers to an air bridge patterned with a triangular lattice of holes. **Figure 5a** shows the real part of the frequency dispersion in the Γ–K direction. The cladding light line (dotted line in the figure) separates the regions of guided and quasi-guided modes. The planar waveguide is monomode up to a frequency $\omega a/(2\pi c) \cong 0.61$, where a second-order mode appears. $\mathrm{Re}(\omega)$ is generally larger than for the ideal 2D case due to a vertical confinement in the slab waveguide: this effect is particularly pronounced for a strong

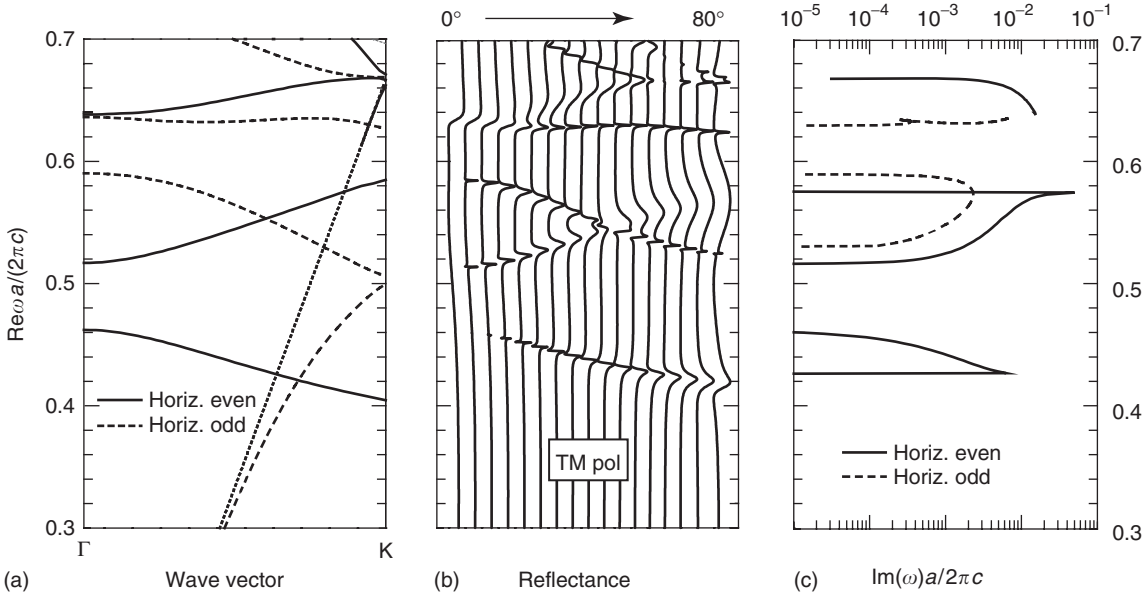


Figure 5 (a) Real part of the frequency; (b) specular reflectance from the surface at angles of incidence from 0° to 80° in steps of 5°; and (c) imaginary part of the frequency for a photonic crystal slab with a triangular lattice of air holes of radius $r/a = 0.3$ patterned in a membrane with core thickness $d/a = 0.3$ and dielectric constant $\varepsilon = 12$. All curves calculated for TM polarized (even) modes with respect to the plane of incidence along the Γ–K orientation. Solid (dashed) lines in (a) and (c) correspond to even (odd) modes with respect to a horizontal symmetry plane. The dotted line in (a) represents the dispersion of light in air.

index-contrast slab. **Figure 5b** displays the calculated reflectance of a plane wave incident on the surface of the slab, for increasing values of the angle of incidence. Spectral structures appear in reflectance spectra, which mark the excitation of photonic modes and whose evolution as a function of the incidence angle corresponds to the $\omega(k)$ dispersion of **Figure 5a**: the wave vector parallel to the surface is conserved and it equals $k = (\omega/c)\sin\theta$ (modulo a reciprocal lattice vector). **Figure 5c** shows the imaginary part of the frequency of quasi-guided modes plotted as a function of the real part. The imaginary part depends strongly on the mode index and on the wave vector, and it vanishes when the real part of the dispersion crosses the light line and goes into the guided mode region. It can be seen that an increasing imaginary part in **Figure 5c** corresponds to an increasing line width of the reflectance structure in **Figure 5b**.

The imaginary part of the mode frequency leads to an imaginary part of the wave vector through the relation $\mathrm{Im}(k) = \mathrm{Im}(\omega)/v_g$, where $v_g = \mathrm{d}\omega/\mathrm{d}k$ is the group velocity. Physically, a propagating quasi-guided mode in an ideal photonic crystal slab is subject to intrinsic radiative losses due to diffraction out of the 2D plane. In real systems, extrinsic diffraction losses related to fabrication aspects (insufficient hole depth, nonvertical profile, roughness of the sidewalls) are present for both guided and quasi-guided modes.

Defect States

It is well known in the electronic structure theory that localized defects introduce impurity states in the gap. An analogous phenomenon takes place in photonic crystals. A common example is a planar defect in a 1D photonic crystal, which is known as a planar (or Fabry–Pérot) microcavity. Linear and point defects in photonic crystals of higher dimensionality also give rise to defect modes within the photonic bandgap. The simplest kinds of defects can be treated with the methods introduced for bulk photonic crystals, provided the defect is repeated with supercell periodicity.

An example of photonic bands for a line defect, or channel waveguide in a photonic crystal slab is shown in **Figure 6**. The structure consists of a single missing row of holes in the Γ–K direction of the triangular lattice (see inset). The defect modes are found outside the projections of the 2D bands onto the Γ–K direction, which are denoted by shaded regions. The defect mode that falls below the photonic gap of the triangular lattice is called "index-guided," because it is laterally confined by the dielectric discontinuity between the channel region and the surrounding photonic crystal. The modes that exist only within the bandgap of the 2D lattice are called "gap-guided" and are laterally confined by the photonic gap, that is, by the lack of electromagnetic eigenmodes in the surrounding material. Both index- and gap-guided modes can be seen in **Figure 6**. Notice that the defect modes can be either guided or quasi-guided, according to their frequency with respect to the cladding light line. Truly guided defect modes allow, in principle, for straight propagation without losses.

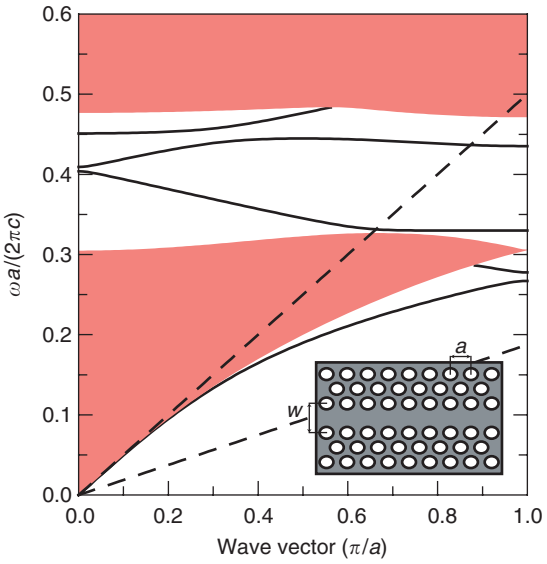


Figure 6 Dispersion of defect modes of a linear waveguide in a membrane with thickness $d/a = 0.3$ and dielectric constant $\varepsilon = 12$. The structure is shown in the inset and the holes in the triangular lattice have a radius $r/a = 0.36$. Only horizontally even and vertically odd modes are shown. The shaded regions correspond to the projected bands of the periodic lattice and the dashed lines denote the light dispersion in the dielectric material and in air.

The linear defect structure in a photonic crystal lends itself to the realization of ultra-compact waveguides with sharp bends, which cannot be obtained with conventional dielectric waveguides based on total internal reflection. **Figure 7** shows an example of wave propagation through a 120° bend in the triangular lattice of holes – the electromagnetic pulse is guided through the bend by the existence of a photonic gap in 2D. The FDTD method is particularly useful to model this kind of structures.

Measuring the Photonic Band Structure

In conclusion, a few experimental methods for measuring the photonic band structure are mentioned. The frequency window of a photonic bandgap can be obtained by transmission (or reflection) measurements.

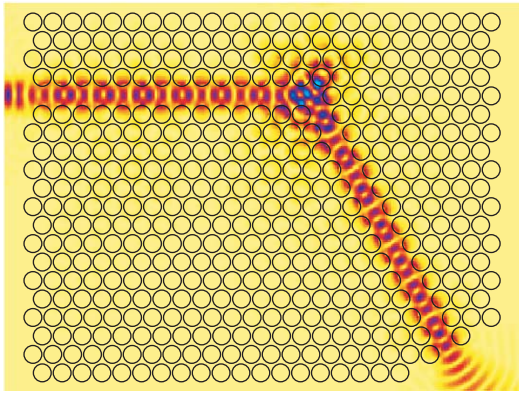


Figure 7 Electric field profile for an electromagnetic wave propagating through a 120° bend designed in a triangular lattice of air holes (FDTD calculation in two dimensions). The modulus of the field is represented in a linear scale.

The conceptually simplest method for obtaining the photonic band dispersion in the direction of light propagation is to measure the phase delay $\phi(\omega)$ of an electromagnetic wave through the photonic crystal: if d is the thickness of the sample, the wave vector at any given frequency can be obtained from the relation $\phi(\omega) = k(\omega)d$. This method has been applied by means of phase-sensitive measurements, first in the microwave spectral region and later in the near infrared and visible regions. The band dispersion in the direction of light propagation can also be obtained by analyzing Fabry–Pérot fringes in reflection or transmission through a sample of finite size. The photonic dispersion in two or three dimensions can also be derived by measuring the beam propagation (i.e., the refraction) of a light beam impinging on the photonic crystal at nonnormal incidence. The 2D band dispersion above the light line in photonic crystal slabs can be obtained by variable-angle reflectance measurements from the crystal surface. The principle is apparent from **Figure 5b**: the spectral position of a resonant structure in reflectance yields both the frequency and the parallel wave vector of a photonic mode. The same variable-angle reflectance technique can be applied to 2D and 3D photonic crystals, and in this case, it leads to the dispersion of photonic bands in a plane parallel to the crystal surface. Finally, the photonic dispersion can also be obtained by angle-resolved luminescence measurements in the far field: this is the photonic analog of angle-resolved photoemission spectroscopy for electronic systems.

See also: Crystal Symmetry; Electrodynamics: Continuous Media; Electronic Structure Calculations: Plane-Wave Methods; Nuclear Models and Methods; Tight-Binding Method in Electronic Structure.

PACS: 42.70.Qs

Further Reading

Bowden CM and Zheltikov AM (eds.) (2002) Nonlinear photonic crystals. *Feature Issue, Journal of the Optical Society of America B* 19: 1961–2296.

Busch K, Lölkes S, Wehrspohn RB, and Föll H (eds.) (2004) *Photonic crystals: advances in design, fabrication, and characterization*. Weinheim: Wiley-VCH.

Joannopoulos JD, Meade RD, and Winn JN (1995) *Photonic Crystals – Molding the Flow of Light*. Princeton NJ: Princeton University Press.

John S (1987) Strong localization of photons in certain disordered dielectric superlattices. *Physical Review Letters* 58: 2486–2489.

Johnson SG and Joannopoulos JD (2001) *Photonic Crystals: The Road from Theory to Practice*. Boston: Kluwer Academic Publishers.

Krauss TF and Baba T (eds.) (2002) Photonic crystal structures and applications. *Feature Section, IEEE Journal of Quantum Electronics* 38: 724–963.

Rarity J and Weisbuch C (eds.) (1996) *Microcavities and Photonic Band Gaps: Physics and Applications*, NATO ASI Series E vol. 324: Applied Sciences, Dordrecht: Kluwer Academic.

Sakoda K (2001) *Optical Properties of Photonic Crystals*. Springer Series in Optical Sciences, vol. 80. Berlin: Springer.

Sigalas MM, Ho KM, Soukoulis CM, Biswas R, and Tuttle G (1999) Photonic crystals. In: Webster J (ed.) *Wiley Encyclopedia of Electrical and Electronics Engineering*, vol. 16, pp. 345–359. New York: Wiley.

Slusher RE and Eggleton BJ (eds.) (2003) *Nonlinear Photonic Crystals*. Springer Series in Photonics, vol. 10. Berlin: Springer.

Soukoulis CM (2001) Photonic band gap materials. In: Fouque JP (ed.) *Diffuse Waves in Complex Media*, pp. 93–107. Dordrecht: Kluwer.

Soukoulis CM (ed.) (2001) *Photonic Crystals and Light Localization in the 21st Century*, NATO Science Series, vol. C563. Dordrecht: Kluwer.

Yablonovitch E (1987) Inhibited spontaneous emission in solid-state physics and electronics. *Physical Review Letters* 58: 2059–2062.

Photons *See* Photon Statistics and Coherence Theory; Polaritons; Statistical Mechanics: Quantum.

Photosynthesis, Physics of

J L Herek, FOM Institute for Atomic and Molecular Physics, Amsterdam, The Netherlands

Introduction

Photosynthesis is the transformation of light energy into fuel that drives essential life processes in a variety of organisms. The physics of photosynthesis is photophysics: the collection and handling of sunlight and its conversion to a more stable form of energy. The subsequent utilization of this light energy involves a series of chemical reactions, which taken together comprise the global biological function of photosynthesis. In this article, the focus is on the fundamental steps of photosynthesis. The article begins by looking at the nanoscale machines involved, and then follows the path of an absorbed photon of sunlight as it journeys through the photosystem until eventually being converted into chemical potential.

It is worth bearing in mind that the total time of this phase of photosynthesis is no more than 100 ps, yet it is packed with numerous small steps, each of which has now been resolved in exquisite detail. Since the 1980s, photosynthetic complexes have been crystallized and analyzed to the atomic level. At the same time, lasers have been replacing the Sun in laboratory studies with increasingly better time resolution, such that one can now initiate and follow the flow of energy and charge in real time. This combination of detailed structural and dynamic information has opened a new window on photosynthesis, revealing the importance of these fundamental physical processes on the overall biological function.

Photosynthetic Machinery

Photosynthesis occurs in plants, algae, and some bacteria. The photosynthetic apparatus of all these organisms contains two key components responsible for the primary photophysics: light-harvesting (LH) complexes and reaction centers (RCs). Both are sophisticated pigment–protein structures located within cell membranes, and while they are necessarily coupled, each has a unique task. The light-harvesting complexes capture photons and transfer the solar energy both within and between complexes until it reaches the reaction centers, where this energy is then used to shuttle electrons across the photosynthetic membrane.

The essential components of photosynthesis can be viewed as nanoscale machines arranged in a complex web (rather than a linear assembly line operation) for optimal performance. **Figure 1** shows the photosynthetic apparatus responsible for collection and conversion of solar energy for the case of purple nonsulfur bacteria. These primitive organisms serve as prototype species for studies of the physics of photosynthesis. Unlike plant systems, the various components of the photosynthetic apparatus in these bacteria can be easily isolated and purified. In some

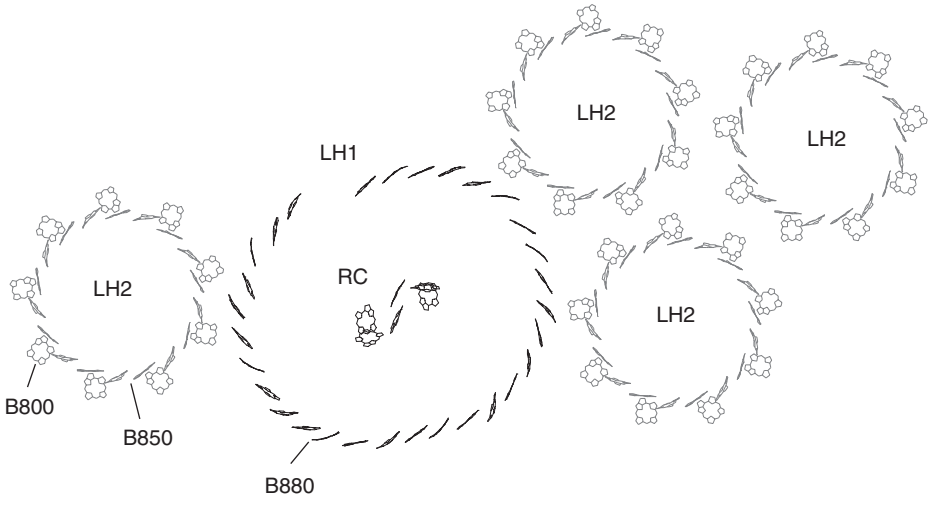


Figure 1 Schematic representation of the photosynthetic machinery of purple nonsulfur bacteria, based on crystallographic data. For clarity, the protein and carotenoid molecules are omitted. The peripheral light-harvesting complex LH2 contains two circular aggregates of bacteriochlorophyll molecules, labeled B800 and B850 according to their characteristic absorption maxima at 800 nm and 850 nm. Likewise, the larger core light-harvesting complex LH1 contains a broken ellipse of aggregated bacteriochlorophylls, the B880 ring. Nestled within the LH1 complex is the photosynthetic reaction center (RC). The supramolecular aggregate of LH2, LH1, and RC complexes is called the photosynthetic unit.

species, these components have been crystallized, allowing atomic-level structural information to be extracted. Also, the genetic makeup of these systems is known, allowing site-specific mutagenesis in which individual amino acids in the protein matrix can be altered. Such mutants are invaluable in studies aimed at detailed structure–function relations.

Figure 1 shows schematically a top-down look at the photosynthetic membrane of purple bacteria, based on real structural data. The ring-like structures are light-harvesting (LH) complexes. These complexes contain cylindrical aggregates of alpha-helical proteins (not shown) that span the lipid bilayer membrane and noncovalently bind the pigment molecules in circular or elliptical arrays. Nestled within one of these rings is another pigment–protein complex, the so-called reaction center (RC). This complex contains two branches with near mirror symmetry, also spanning the photosynthetic membrane. The overall supramolecular aggregate of peripheral light-harvesting complexes LH2 and core complexes containing the RC and surrounding LH1 complex is called the photosynthetic unit (PSU), though the ratio of LH2 to LH1-RC varies according to growth conditions such as temperature and light intensity.

In the intricate network of molecules arranged within the protein matrix, the function is determined by the location. While a chlorophyll molecule in the reaction center plays an essential role in photochemistry, a chlorophyll molecule located in an antenna complex functions in a completely different manner. These functions are described in the following sections.

Harvesting Light

The first step of photosynthesis is the collection of sunlight. How is light captured by living things? Illuminated molecules may either reflect or absorb incident light. Molecules that are especially capable of absorbing light in the visible region of the spectrum and give color to living things are called pigments. The pigments of photosynthesis are primarily chlorophylls, of which there are many different varieties. These molecules efficiently absorb light in the red and blue regions of the solar spectrum, while reflecting away most of the green (hence the predominance of this color in Nature). In addition, a variety of other less-abundant pigments enhance the light-absorbing capabilities of photosynthetic organisms by capturing other regions of the visible spectrum. Carotenoids, for example, have strong absorption in the blue-green region, but reflect away the yellow to red wavelengths. These pigments are largely responsible for the autumn "colorama," showcasing the presence of the carotenoids that becomes gloriously apparent as the chlorophyll in the leaves of the trees begins to break down.

When a molecule absorbs a photon of visible light, one of its electrons is promoted to a higher energy state. The molecule is then said to be in an electronic excited state. The wavelengths of light that are capable of creating excited states for a given molecule are given by an absorption spectrum. In **Figure 2**, the absorption spectra of chlorophyll and carotenoid molecules are shown, in comparison with the solar irradiance spectrum on Earth. The entire visible range of light (400–700 nm) as well as much of the near-infrared (up to 1000 nm) is used to drive photosynthesis in different organisms.

Understanding the excited states in photosynthesis is a key area of biophysics research, as it is from these states that all physical processes such as energy transfer and trapping as well as chemical reactions may occur. The nature of the excited state depends not only on the characteristic properties of the individual molecule, but also on its specific location within the photosynthetic complex. Interactions between neighboring pigments and, to some extent, the surrounding protein can shift the excited-state energies and thereby affect dynamics. For example, in photosynthetic purple bacteria, the primary pigment is bacteriochlorophyll. The main absorption of this pigment when isolated in solution is at $\sim$780 nm. In the photosynthetic unit, however, the absorption band shifts according to its location (and corresponding function): in the peripheral light-harvesting complex (LH2),

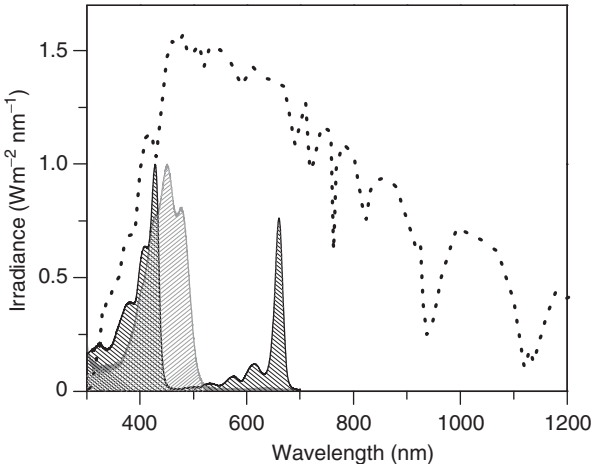


Figure 2 Solar irradiance spectrum and absorption spectra of photosynthetic pigments. Dotted curve: Intensity profile of sunlight measured on the Earth's surface. Black hatch: Absorption spectrum of chlorophyll, the most abundant photosynthetic pigment in Nature. Grey hatch: Absorption spectrum of the carotenoid beta-carotene, an accessory pigment. The spectra of the pigments are in absorption units (scale not shown).

there are two distinct binding sites for bacteriochlorophyll, giving rise to absorption bands at 800 and 850 nm, respectively. These binding sites are correspondingly labeled B800 and B850; see **Figure 1**. In the B800 ring, the individual pigments are well separated (by $\sim$20 Å), and the spectral properties are determined mainly by weak interactions with the protein environment. In the B850 ring, however, the pigments are tightly packed in a "waterwheel" arrangement; the further red-shifted absorption of the molecules here is due to the much stronger excitonic coupling between adjacent pigments. The ring structure also enhances absorption and acts as an energy storage unit. In the LH1 core light-harvesting complex, which directly couples to the reaction center, the bacteriochlorophyll absorb at 880 nm, and have an arrangement similar to the B850 ring. This tuning of absorption properties with location is crucial to the overall function of photosynthesis, creating an energy hierarchy that efficiently funnels light energy into the reaction center.

Antenna Action: Energy Transfer

While the reaction center is the hotspot for photosynthetic energy conversion, by itself it can only capture a few photons per second (given its absorption cross section, the solar flux, and its overlap with the solar irradiance spectrum). Yet, the turnover time of the reaction center function is of the order of 10 ms, that is, it is capable of a much higher throughput. Hence, an antenna system evolved to catch extra photons and feed them to the reaction center.

Antenna networks are highly regulated to maximize the efficiency of energy collection while at the same time avoiding photodamage that can occur when too much light is absorbed. These functions occur by rapid and directional energy transfer reactions. Following absorption of solar light, the pigment molecules are in electronically excited states. In principle, this excitation energy may then be used directly to drive a chemical reaction. However, in photosynthesis, the energy is first transferred between many molecules, over quite large distances, before it arrives at the site of action.

The antenna system increases the amount of sunlight that can be absorbed compared to a single pigment, thereby increasing the number of photons available for photochemistry. A photosynthetic antenna is analogous to a satellite dish, collecting energy from a large area and concentrating it at a small receiver site (the reaction center), thereby greatly increasing the signal strength. Antenna networks contain an additional 100–10 000 pigments per reaction center. Molecules located in the antenna do not perform any chemistry, but simply transfer the collected energy between each other in a well-defined and efficient manner. For optimal functioning, energy collected by the antenna pigments must reach the reaction center within $\sim$100 ps; hence, the transfer between individual pigments should be of the order of 1 ps or less. This transfer process is purely physical and is governed by the energetic interactions between molecules.

Advances in ultrafast laser spectroscopy have allowed detailed measurements of these energy transfer dynamics in real time. In the laboratory, a femtosecond laser replaces the sun and initiates the process of interest. As the absorption of the antenna pigments is determined by their location, the laser may be tuned to be resonant with a given molecule, thereby isolating a specific energy transfer step. For example, in the LH2 complex of purple bacteria, the B800 to B850 energy transfer can be targeted by tuning an excitation pulse to 800 nm, and then probing the arrival of energy by monitoring the buildup of the transient absorbance signal at 850 nm with a second delayed femtosecond pulse. By varying the timing of the probe pulse, the dynamics of the process may be followed at discrete time steps, which, when combined, yield a kinetics trace.

All photosynthetic systems have an inherent energy gradient that dictates the flow of energy among the pigments. In purple bacteria, this is given by B800 $\rightarrow$ B850 $\rightarrow$ B880 $\rightarrow$ RC (note that the valuable contributions of carotenoids, which perform accessory light harvesting in addition to other important functions, are ignored). The associated time constants for these transfer steps are shown in **Figure 3**. The lifetime of an isolated antenna complex, decoupled from the reaction center, is of the order of 1 ns. During this time, the excitation energy moves freely among the lowest-energy pigments of the antenna network, until eventually dissipated by nonradiative relaxation processes. In the full photosynthetic unit, where antennas are connected to reaction centers, the excitation is trapped within 30–50 ps. Hence under physiological conditions, nearly all the energy is utilized for photochemistry.

Excitation transfer in biological systems has usually been described as an incoherent hopping process following the theory of Förster. It is hard to overestimate the physical insight obtained by this approach. At the same time, a number of experimental observations for photosynthetic pigment–proteins do not fit into this picture and call for a more general description of the problem. The incoherent Förster description of electronic energy transfer assumes that the interaction between the involved molecules is very weak. In the opposite limit of strong interaction between the molecules, a molecular exciton description

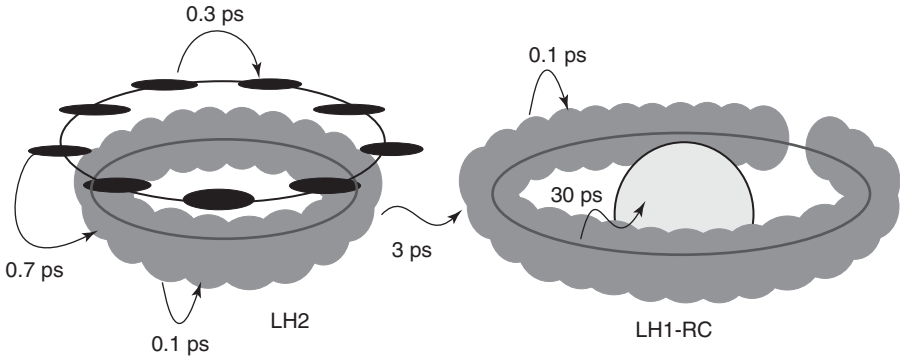


Figure 3 Energy transfer times between the various pigments in the photosynthetic units of purple bacteria.

is appropriate. In this limit, the excited states of the individual molecules combine to exciton states common for all the interacting molecules (the aggregate), the excitation is coherently delocalized over a number of pigment molecules and the absorption spectrum exhibits a prominent exciton band structure. Time evolution in the exciton picture occurs via phonon-induced relaxation between exciton levels. These two descriptions, incoherent hopping and exciton relaxation, are the two qualitatively different limiting cases of the general process of excitation dynamics. What description should be chosen for a particular observation depends on the system properties and the experimental conditions. It may happen that the real situation is somewhere in between these two limiting cases, in which case the analysis is particularly challenging.

The important factor in deciding which process(es) occur is the ratio of the coupling between electronic transitions (V) and the disorder (Δ). The disorder may be either static (spectral inhomogeneity) or dynamic (electron–phonon interaction). If V/Δ is much less than unity, then the interactions are considered very weak and energy transfer occurs in the incoherent Förster hopping regime. In the opposite case, if $V/\Delta \gg 1$, the interaction is very strong and the exciton picture is used. In both cases, the transfer rate can be calculated using the first-order perturbation theory leading to the Fermi Golden rule. In the hopping description, the electronic coupling (V) acts as a perturbation; in the exciton picture, an electron–phonon coupling (dynamic disorder) is the perturbation.

In the weak coupling incoherent limit, the rate of energy transfer (energy hopping) from a donor to an acceptor molecule is determined by three different molecular properties: the electronic coupling between donor and acceptor (V), the overlap of donor fluorescence and acceptor absorption spectra (J), and the relative orientation of the molecules (κ), and given by

$$K_{\mathrm{ET}} \sim V^2 J \kappa^2$$

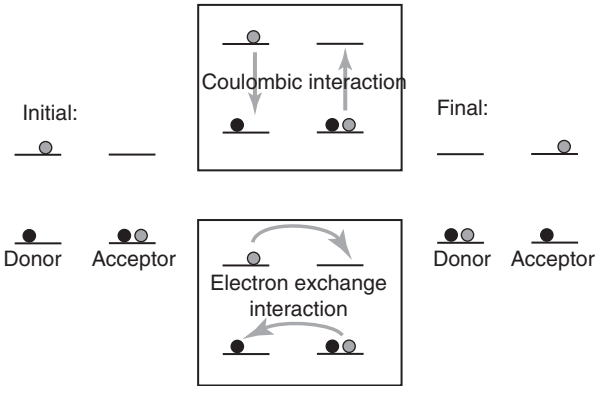


Figure 4 Interactions governing excitation energy transfer. The Coulombic interaction promotes simultaneous deactivation of the donor and excitation of the acceptor. The electron exchange interaction involves transfer of excitation energy via direct exchange of electrons between donor and acceptor molecules.

A few things should be noted. The electronic coupling (V) has two contributions, one short-range electron-exchange interaction (which decays exponentially with the intermolecular distance R) and one long-range Coulombic interaction (which scales as R^{-6}). These two cases are described in **Figure 4**. The interaction accounted for in the Förster mechanism and implicated by the transition dipoles is the generally dominating dipole–dipole contribution to the Coulombic interaction; higher multipole terms are also generally present, but of much less importance for molecules with strong transition dipole moments (e.g., strongly allowed optical transitions). The electron exchange term in the interaction may become important at very short distances, when there is overlap of the donor and acceptor molecular orbitals. This is the so-called Dexter mechanism, and in particular for molecules with forbidden optical transitions, it may become important. Energy transfer in photosynthetic antenna complexes between the lowest forbidden excited state of carotenoid molecules and the lowest excited state of chlorophyll molecules has been discussed in terms of the Dexter mechanism.

In the exciton picture, the excitation energy is delocalized over the interacting pigment molecules. However, the extent of delocalization is determined by the V/Δ ratio, and in a real system, exciton delocalization is generally smaller than the physical size of the molecular aggregate. The localization process may also have a dynamic part due to the dynamic contribution to the disorder Δ, and the exciton thereby changes its size in time. The relevant aspects of the energetics and dynamics in the exciton limit are the exciton level structure, the dynamics of exciton relaxation among the set of exciton levels, the extent and dynamics of delocalization of the exciton, and the motion of the whole exciton over the aggregate of interacting molecules. The energies of the exciton states are obtained from the Hamiltonian of the interacting molecules and the dynamics may be obtained from the density matrix of the system.

Energy Conversion and Storage

Electron transfer reactions are Nature's solution to the problem of converting the short-lived energy of electronically excited molecules into a more stable form of energy, which can, in addition, be transformed into various energy-rich substances with practically unlimited storage potential. Nature has devised highly efficient systems that optimize electron transfer processes while at the same time avoiding back-reactions and quenching processes.

When the excitation energy reaches the photosynthetic reaction center, it launches a sequence of electron transfer reactions that create a charge-separated state in which positive and negative charges are separated over the lipid bilayer membrane. Energy stored in this manner can then be used to pump protons across the membrane. The resulting proton motive force drives the synthesis of energy-rich compounds such as adenosine triphosphate (ATP). ATP is thermodynamically stable for a sufficiently long time to serve as the "fuel" for many subsequent biochemical processes. Once stored in the form of ATP, the energy can be used in the cell at any arbitrary instant for transport, synthesis, or other chemical reactions that require input energy.

Again, of all the photosynthetic organisms, the reaction center of anoxygenic purple bacteria is best understood. Though in contrast to the wide variety in the configurations of antenna complexes, the reaction centers of bacteria and plants share a rather common structural and functional motif. The reaction center consists of a number of protein subunits that form a scaffold (not shown) to hold the photochemically active molecules. While a number of these molecules are identical or similar to the (bacterio)chlorophyll pigments found in the antenna, here the molecules are called cofactors to distinguish their different function. The structure of the reaction center from *Rhodobacter sphaeroides* is shown in **Figure 5**. The cofactors are bacteriochlorophylls (P and B in **Figure 5**), bacteriopheophytins (H), quinones (Q), and an iron atom (located between the quinones; not shown). The cofactors are arranged into two nearly symmetric branches (A and B), joined at the top by the so-called "special pair" of bacteriochlorophyll molecules that are located so close together as to form a super molecule or dimer.

When the excitation energy collected by the antenna network finally arrives at the reaction center, it initiates a cascade of electron transfer reactions. These reactions have been visualized by a variety of spectroscopic techniques, in particular ultrafast transient absorption spectroscopy, often combined with genetic manipulations of the native structure to better elucidate structure–function relations. Upon light irradiation, the excited dimer (P^*) transfers an electron to the bacteriochlorophyll (B_A) and the bacteriopheophytin (H_A) on the A side of the branch within a few ps. The resulting radical pair ($P^+H_A^-$) decays in about 200 ps moving the electron onward to the quinones.

The most striking feature of the reaction center activity is the unidirectional transport down branch A (the active branch), despite the apparent twofold symmetry of the structure. The electron is 200 times

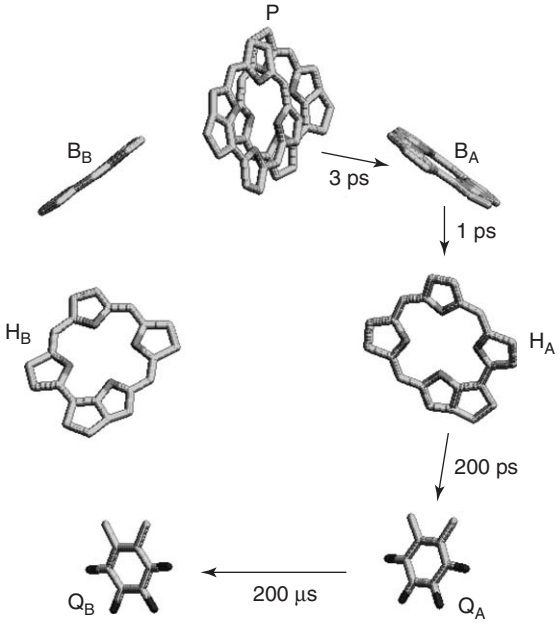


Figure 5 X-ray structure of the active molecules in the reaction center from the photosynthetic purple bacteria *Rhodobacter sphaeroides*, showing the pathway and time constants of electron transport via the active branch A.

more likely to follow branch A than branch B. While considerable experimental and theoretical efforts have tackled this discrepancy, it remains as one of the most intriguing puzzles in the primary processes of photosynthesis.

The high quantum yield of primary charge separation (as much as 0.95 in some systems) comes with a price: a lower energetic yield. The final radical pair state of the bacterial RC contains 0.55 eV, which is only 40% of the excited state energy of the primary electron donor (or around a third of the average photon energy absorbed by the antenna). Perhaps the energetic efficiency of the RC could be improved upon in synthetic RC-like structures, though model systems (consisting of electron donor and acceptor molecules connected by covalent links) have shown quantum efficiencies of 30% or less. Most likely, the rigidity of the protein surrounding the donor and acceptor pigments is essential to achieve such high quantum efficiency.

From Natural to Artificial Photosynthesis

An understanding of how Nature works will allow one to use the basic principles for other purposes. The most obvious application of artificial photosynthesis is the conversion of sunlight to useful forms of energy, such as electricity by photovoltaic devices or fuel by photochemical reactors. Modern solar cells are already being built on principles gleaned from natural photosynthesis. Sensitized semiconductor materials are environmentally friendly and becoming cost-effective alternatives to silicon. Solar cells based on this material utilize sensitizer molecules that act as antenna pigments to collect sunlight in a wide spectral region and funnel the solar energy into the semiconductor by way of an electron transfer process. Ongoing research will lead to the development of new, more efficient solar light-harvesting technologies that not only mimic Nature, but perhaps also surpass it.

See also: Elementary Excitations in Biological Systems; Excitons: Theory; Optical Absorption and Reflectance; Time-Resolved Optical Spectroscopies.

PACS: 82.39.Rt; 82.53.Ps; 87.14.Ee; 87.15.Kg; 92.20.Lw

Further Reading

Amesz J and Hoff AJ (eds.) (1996) *Biophysical Techniques in Photosynthesis*. Dordrecht: Kluwer.

Blankenship RE (2002) *Molecular Mechanisms of Photosynthesis*. Oxford: Blackwell Science.

Hall DO and Rao K (1999) *Photosynthesis*, 6th edn. London: Cambridge University Press.

Ke B (2001) *Photosynthesis: Photobiochemistry and Photobiophysics*. Dordrecht: Kluwer.

May V and Kühn O (2000) *Charge and Energy Transfer Dynamics in Molecular Systems*. Berlin: Wiley-VCH.

Ritz T, Damjanovic A, and Schulten K (2002) The quantum physics of photosynthesis (review). *ChemPhysChem* 3: 243–248.

Sundstrom V, Pullerits T, and van Grondelle R (1999) Photosynthetic light-harvesting: reconciling dynamics and structure of purple bacterial LH2 reveals function of photosynthetic unit (review). *The Journal of Physical Chemistry B* 103: 2327–2346.

van Amerongen H, van Grondelle R, and Valkunas L (2000) *Photosynthetic Excitons*. Singapore: World Scientific.

Piezoelectricity

D Damjanovic, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

The piezoelectric effect describes the linear relationship between the mechanical (stress σ and strain x) and electrical (electric field E, and electric displacement D or polarization P) fields that is observed in certain classes of materials. The "direct" piezoelectric effect, which describes an electric field or electric displacement generated by mechanical stress or strain, was discovered by Pierre and Jacques Curie in 1880. The so-called "converse" effect, which describes mechanical strain or stress generated by an electric field or charge was theoretically predicted by G Lippmann in 1881 and confirmed experimentally by the Curies in 1882. The term "piezo" has its origin in the Greek word for pressure (e.g., piezoelectricity which means electricity produced by pressure) and was proposed by W Hankel in 1881. The formulation of piezoelectricity in tensor form followed the 1894 work of Voigt.

Thermodynamic Formulation of the Piezoelectricity

A convenient way to formally derive piezoelectric effect is by using thermodynamics of polarizable, elastically deformable solids. The reversible change dU in the internal energy U of an elastic, dielectric material that is subjected to a small change of the strain dx_{ij}, electric displacement dD_i, and entropy dS is given by $dU = T\,dS + \sigma_{ij}\,dx_{ij} + E_i\,dD_i$, where T is the

temperature of the material, indices $i, j = 1, 2, 3$ indicate components of the tensors, and summation over repeated indices is assumed. The choice of the independent variables is arbitrary, and for experimental convenience one generally chooses the set T, σ_{ij}, E_i instead of S, x_{ij}, D_i used above to formulate dU. The change of independent variables from one set to another is accomplished by performing a Legendre transformation of U by adding to it, the expression $-TS - x_{ij}\sigma_{ij} - E_iD_i$. The resulting free energy function $G = U - TS - \sigma_{ij}x_{ij} - E_iD_i$ is known as the Gibbs free energy. The differential of G is $dG = -S\,dT - x_{ij}\,d\sigma_{ij} - D_i\,dE_i$ and, therefore, $x_{ij} = -(\partial G/\partial\sigma_{ij})_{T,E}$ and $D_i = -(\partial G/\partial E_i)_{T,\sigma}$. The subscripts indicate that the variables are kept constant. The derivatives $d_{ijk}^{\text{converse}} = (\partial x_{ij}/\partial E_k)_T = -(\partial^2 G/\partial E_k\partial\sigma_{ij})$ and $d_{kij}^{\text{direct}} = (\partial D_k/\partial\sigma_{ij})_T = -(\partial^2 G/\partial\sigma_{ij}E_k)$ define the converse and the direct piezoelectric effect, where d_{ijk} are elements of the tensor of the piezoelectric coefficient. Since the order in which the second derivatives are taken is irrelevant, one obtains the following Maxwell relations:

$$d_{kij}^{\text{direct}} = \left(\frac{\partial D_k}{\partial\sigma_{ij}}\right)_T = -\left(\frac{\partial^2 G}{\partial\sigma_{ij}\partial E_k}\right)$$
$$= -\left(\frac{\partial^2 G}{\partial E_{ik}\partial\sigma_{ij}}\right) = \left(\frac{\partial x_{ij}}{\partial E_k}\right)_T = d_{ijk}^{\text{converse}}$$

which demonstrate the thermodynamic equivalence of the converse and the direct piezoelectric effect. The piezoelectric effect can be formulated in integrated form as

$$D_k = d_{kij}^T\sigma_{ij}$$

and

$$x_{ij} = d_{ijk}^T E_k$$

The superscript T indicates that the piezoelectric coefficient is considered under isothermal conditions and is usually not written explicitly. In nonpolar materials, adiabatic and isothermal coefficients are identical while in polar (pyroelectric) materials their difference is usually negligible.

The piezoelectric tensor equations are generally written in reduced matrix form, according to the following convention by Voigt: $d_{ijk} = d_{im}$, where $m = 1, 2, 3$ for $jk = 11, 22, 33$, and $d_{ijk} = 1/2\,d_{im}$, where $m = 4, 5, 6$ for $jk = 23$ or 32; 13 or 31; and 12 or 21, respectively. The piezoelectric effect is thus expressed in reduced form as $x_m = d_{mi}E_i$ and $D_i = d_{im}\sigma_m$. Note that while d_{ijk} transforms as tensor, d_{im} does not.

The signs of the piezoelectrically generated charge D (or polarization P) and strain x depend on the direction of applied mechanical and electric fields, respectively. This distinguishes the piezoelectric effect from electrostriction in which the electric field of either sign produces the same sign of the strain. This sign of a piezoelectric coefficient is a matter of convention. For example, a d_{kk} is defined as positive, if tension σ_k parallel to X_k axis causes a potential difference on two parallel faces perpendicular to X_k, such that the positive terminal coincides with the face of the crystal on which $+X_k$ points from inside the crystal. A given piezoelectric coefficient d_{ij} can be either positive or negative, and can change the sign with temperature or frequency. It is common to call piezoelectric coefficients measured in the direction of the applied field, longitudinal coefficients (d_{11}, d_{22}, d_{33}), and those measured in the direction perpendicular to the field, transverse coefficients (e.g., d_{31}, d_{23}). Other piezoelectric coefficients are known as shear coefficients (e.g., d_{15}, d_{26}) (**Figure 1**).

The remaining combinations of the electric and mechanical variables (E, x;D, x;D, σ) define additional thermodynamic functions from which the following pairs of direct and converse piezoelectric relations can be derived: $\sigma_m = -e_{im}E_i$ and $D_i = e_{im}x_m$; $\sigma_m = -h_{im}^x D_i$ and $E_i = -h_{im}x_m$; $x_m = g_{im}D_i$; and $E_i = -g_{im}\sigma_m$, where e, g, and h are the piezoelectric coefficients. The relationships among the four piezoelectric

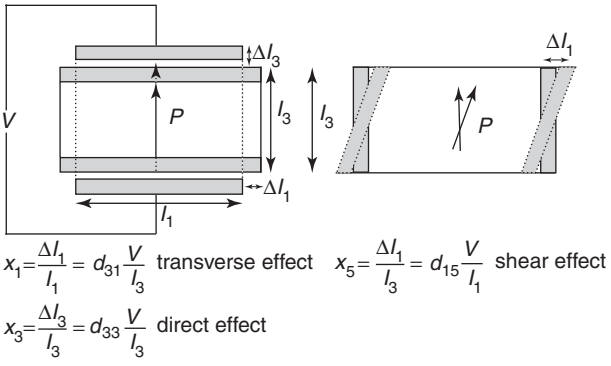


Figure 1 Illustration of the converse longitudinal, transverse, and shear piezoelectric effect in a polar material. The shaded areas indicate conductive layers (electrodes) applied on the sample surfaces to collect charge or apply electric field.

coefficients in reduced matrix notation are given by

$$d_{im} = e_{in}s^E_{nm} = \varepsilon^\sigma_{ij}g_{jm} \ \ (\mathrm{m\,V^{-1}\ or\ C\,N^{-1}})$$
$$e_{im} = d_{in}c^E_{nm} = \varepsilon^x_{ij}h_{jm} \ \ (\mathrm{C\,m^{-2}\ or\ N(V\,m)^{-1}})$$
$$g_{im} = h_{in}s^D_{nm} = \beta^\sigma_{ij}d_{jm} \ \ (\mathrm{m^2\,C^{-1}\ or\ V\,m\,N^{-1}})$$
$$h_{im} = g_{in}c^D_{nm} = \beta^x_{ij}e_{jm} \ \ (\mathrm{N\,C^{-1}\ or\ V\,m^{-1}})$$

The units of each coefficient are indicated in parentheses. The difference between the coefficients can be illustrated by taking as an example g and d coefficients. The matrix indices are omitted for simplicity. Relation $D = d\sigma$ describes the piezoelectric charge measured on electrically shorted samples, with free flow of the charge into the external electric circuit. If the sample is open circuited, this piezoelectric charge will accumulate on the sample surface and generate a voltage (or electric field E) across the sample. The field depends on the capacitance (i.e., dielectric permittivity ε) of the sample (charge = capacitance × voltage) and is related to the stress by $E = -g\sigma$, where $g = d/\varepsilon$.

Piezoelectric Constitutive Equations and Electromechanical Coupling Factors

Following the thermodynamic approach given above, the total strain in a piezoelectric material under the combined influence of electric and mechanical fields, under isothermal conditions and in the limit of the linear response, is given by

$$x_m = s^E_{mn}\sigma_n + d_{im}E_i \qquad [1]$$

where the first term denotes the Hook's law. Under the same conditions, the total charge density (or dielectric displacement) generated by a piezoelectric material is given by

$$D_i = d_{im}\sigma_m + \varepsilon^\sigma_{ij}E_j \qquad [2]$$

where the second term indicates polarization of the material by an electric field. The pairs of equations of this type are known as the piezoelectric constitutive equations. Terms describing coupling to thermal, magnetic, and other effects can be added to constitutive relations as necessary. Analog constitutive equations can be written for g, e, and h coefficients.

Analysis of any problem involving piezoelectric effect must usually involve full constitutive equations and not just its piezoelectric part. A common experimental situation is illustrated in **Figure 2**, where an external stress σ_3 is applied on a piezoelectric sample and the generated charge is collected on the top and bottom surfaces of the sample. Application of a stress deforms the sample both longitudinally and laterally. If the sample is short-circuited and free to deform laterally, the only stress present is the applied stress, and the piezoelectric charge collected from the electroded top and bottom faces is given by $D_3 = d_{33}\sigma_3$ (e.g., the case of a ferroelectric sample poled along its thickness as shown in **Figure 2**). However, if the sample is clamped laterally (e.g., a piezoelectric thin film deposited on a thick, perfectly rigid substrate), the clamping will be equivalent to having additional, effective stresses, σ_1 and σ_2, opposing the lateral deformation of the sample. The piezoelectric charge is then given by $D_3 = d_{31}\sigma_1 + d_{32}\sigma_2 + d_{33}\sigma_3$. The lateral stresses can be calculated from the clamping condition $x_1 = x_2 = 0 = s^E_{1n}\sigma_n = s^E_{2n}\sigma_n$. The effective piezoelectric coefficient is then given by $d^{\mathrm{eff}}_{33} = d_{33} + f(s_{mn}, d_{im})$. In many cases, function $f(s_{mn},d_{im})$ is negative so that the effective d^{eff}_{33} can be as much as 50% smaller than the true piezoelectric coefficient. In practice, the boundary conditions can be more complex and it is often necessary to consider constitutive equations with mixed independent variables (e.g., strain and stress). This is the case when the substrate also deforms, in which case f includes elastic properties of the substrate, and can become positive, increasing the apparent piezoelectric response. The influence of function f can be minimized by changing the aspect ratio of the sample.

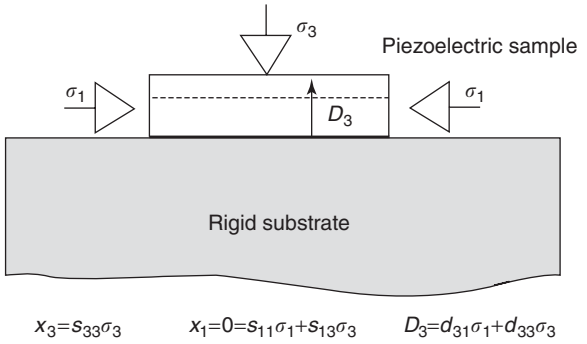


Figure 2 Illustration of the clamping effect and use of both pairs of constitutive equations for a two-dimensional piezoelectric material. The only externally applied stress is σ_3, whereas the lateral stress σ_1 is a consequence of the clamping.

Because of the piezoelectric coupling between the electrical and elastic variables, the values of the dielectric permittivity and elastic compliance (or stiffness) in piezoelectric materials will not be the same if measured under different experimental conditions. If an electric field is applied on a clamped one-dimensional sample, the resulting piezoelectric stress is $\sigma = -eE = -(d/s^E)E$. Replacing this stress into expression $D = d\sigma + \varepsilon^\sigma E$ for the dielectric displacement, one obtains $D = (\varepsilon^\sigma - d^2/s^E)E = \varepsilon^x E$ where $\varepsilon^x = \varepsilon^\sigma(1 - d^2/s^E\varepsilon^\sigma) = \varepsilon^\sigma(1 - k^2)$ is called the clamped (zero strain) dielectric constant. The parameter k, defined by $k^2 = d^2/s^E\varepsilon^\sigma$, is known as the electromechanical

coupling coefficient. In some ferroelectrics, k may be as large as 0.7–0.9 leading to as much as 50–80% difference between the free (zero stress) and clamped (zero strain) dielectric constant. It can be shown in a similar way that the elastic compliance s^D, measured under open circuit conditions (zero or constant D), and compliance s^E, measured under short circuit conditions (zero or constant E), are related by $s^D = s^E(1 - d^2/s^E\varepsilon^X) = s^E(1 - k^2)$. The coupling coefficients depend on measurement direction. In many applications, the coupling coefficients are the most important figure of merit of a piezoelectric material.

The coupling coefficients may be related to the ratio of the input mechanical (or electrical) energy and the stored electrical (or mechanical) energy. However, when defined as the energy ratio, the coupling coefficient is well defined only for simple geometries. For more complex shapes, this energy ratio depends on external variables and thus cannot represent a property of the material.

Symmetry Relations and Piezoelectricity

Because the strain and stress are symmetrical tensors, the piezoelectric-coefficient tensor is also symmetrical with respect to the same indices, $d_{ijk} = d_{ikj}$. The number of independent piezoelectric coefficients is thus reduced from 27 to 18. The number of independent elements of d_{ijk} may be further reduced by the symmetry of the material.

By using Neumann and Curie principles, it is possible to predict which materials can exhibit piezoelectric effect. The application of an electric field (which is a vector – a first rank tensor – and thus possesses noncentrosymmetric symmetry) changes dimensions of any insulating material by separating the centers of the negative and positive charges. The stress field (a tensor of fourth rank, with center of symmetry), can polarize a material (separate centers of positive and negative charges) only if there is at least noncentrosymmetric distribution of charges in the material. It is straightforward to see why piezoelectric effect cannot exist in a centrosymmetric material. If stress (which is centrosymmetric) is applied on a centrosymmetric material and the crystal gets polarized, then inversion of the crystal and stress will lead to the same crystal-stress state, but polarization would be reversed. This is possible only if the polarization is zero. Thus, the piezoelectric and other effects described by odd-rank tensors are forbidden by the symmetry in crystals that belong to 11 centrosymmetric groups and in noncentrosymmetric point group 432. Crystals that belong to the remaining 20 noncentrosymmetric groups can exhibit the piezoelectric effect. These 20 point groups are sometimes called piezoelectric point groups. Ten of the piezoelectric point groups possess a unique polar axis and may exhibit, in the absence of an external field, a spontaneous polarization vector, P_s. All ferroelectric crystals, which exhibit the largest piezoelectric coefficients (see the next section), belong to these ten polar point groups. Ferroelectric materials are different from other polar materials in the fact that their spontaneous polarization can change direction under external fields. This feature (see the next section) is essential in obtaining piezoelectric effect in polycrystalline ferroelectric materials. Thus, some piezoelectrics are pyroelectric and some pyroelectrics are ferroelectrics, while all pyroelectrics and ferroelectrics are piezoelectric. For example, SiO_2 is a piezoelectric but nonpolar material. ZnO is a polar, thus pyroelectric and piezoelectric material, but is not ferroelectric (direction of P_s cannot be switched by an external field). $BaTiO_3$ is a piezoelectric, polar (pyroelectric), and ferroelectric material. It should be emphasized that the value of a piezoelectric coefficient can still be zero even if the symmetry of the material allows for its presence.

Crystallographic point groups with discrete symmetry elements are insufficient to describe polycrystalline and noncrystalline materials. Properties of a ceramic with a random orientation of crystallites (or grains) are, for example, identical in all directions and may be described by introducing the symmetry axis of an infinite order, whose symbol is ∞. The point groups that contain infinity symmetry axes are called the limit groups of symmetry or Curie groups. Ceramic materials with a random orientation of grains possess spherical symmetry, $\infty\infty m$, which is centrosymmetric, and cannot exhibit piezoelectric effect even if the symmetry of each grain belongs to one of the piezoelectric or polar point groups. If the ceramic is, however, ferroelectric, the spontaneous polarization direction in each grain may be reoriented by an external electric field. Such poled or polarized ceramic possesses symmetry of a cone, ∞m, and exhibits piezoelectric effect.

Since the piezoelectric effect is a tensor, its properties are dependent on the orientation of the material. A triclinic material belonging to point group 1 exhibits 18 different piezoelectric coefficients, whereas poled ferroelectric ceramics, which can be described by point group *6mm*, possess only three independent piezoelectric coefficients: d_{33}, $d_{31} = d_{32}$, and $d_{24} = d_{15}$.

Microscopic Origin of the Piezoelectricity and Contributions to the Piezoelectric Effect

As the piezoelectric effect is present in a vast number of compounds, ranging from simple compositions

such as quartz to complex solid-state solutions, natural and synthetic organic and biological materials to many heterogeneous materials, it is clear that there are many different microscopic mechanisms that can lead to a piezoelectric response. The first theory of the piezoelectric effect in crystals was proposed by Born in 1933. In his theory of dynamics of crystal lattices, he formulated the piezoelectric effect in terms of displacement of positive and negative charges but the quantitative agreement of the calculated response and experimental values was unsatisfactory, especially for crystals with a mixed covalent–ionic bonding.

Theoretically, there was for some time a disagreement in the literature even about the formal proof that piezoelectricity is a well-defined bulk property. Presently, the most advanced theoretical approach in describing the piezoelectric response of both compositionally simple piezoelectrics and complex perovskite solid solutions, is based on first-principle calculations. In perovskite ferroelectrics, the strain coefficient e_{ij} can be calculated as

$$e_{im} = \frac{1}{2\pi\Omega}\sum_{\alpha} R_{\alpha,i}\frac{\mathrm{d}}{\mathrm{d}x_m}(\Omega G_{\alpha,l}P_l)$$
$$= e_{im,c} + \sum_{k}\frac{ea_i}{\Omega}Z^*_{k,ii}(k)\frac{\mathrm{d}u_i(k)}{\mathrm{d}x_m}$$

where P_l are components of the total polarization calculated as a Berry phase of the wave functions resulting from first-principle calculations, Ω is the cell volume, and $R_{\alpha,j}$ and $G_{\alpha,l}$ are components of three ($\alpha = 1, 2, 3$) real-space and reciprocal lattice vectors, respectively. $e_{im,c}$ is the so-called clamped-ion or homogeneous-strain contribution that corresponds to applied macroscopic strain and zero internal strain (internal coordinates are scaled with cell size and shape). The last term corresponds to change in polarization induced by internal distortions only; it is related to the effective charge tensor Z^* and to the response of the internal atomic coordinates to a macroscopic strain ($\mathrm{d}u_i/\mathrm{d}x_j$). k runs over all atoms of the cell, e is the electronic charge, and a_i are lattice constants along the i-direction. It is the last term in the above expression that is of most interest in identifying the contribution of each atom in the piezoelectric response of a crystal.

Piezoelectric d_{im} coefficients can be calculated from first-principle approaches by considering the free energy $F = U - P_iE_i$. The internal energy U is obtained from first-principle calculations, P_i as a Berry phase, and E_i are components of the external field. The piezoelectric d coefficient can be calculated as a slope of strain versus electric field, where the strain is calculated from the free energy. These first-principle calculations lead not only to a satisfactory agreement between calculated and experimental piezoelectric coefficients, but can, in principle, be used to predict new materials with predefined piezoelectric properties.

From the phenomenological point of view, the piezoelectric effect in ferroelectric materials may be described reasonably well by the Landau–Ginzburg–Devonshire approach. In this framework, the piezoelectricity has its origin in electrostriction biased by the spontaneous polarization. The d coefficient can, for example, be expressed as $d = 2\varepsilon QP_s$ where Q is an electrostrictive coefficient.

Recently, a high piezoelectric or piezoelectric-like response was discovered in materials in which no piezoelectric effect would be expected. Strontium titanate, $SrTiO_3$, is a simple cubic centrosymmetric perovskite at room temperature; however, high piezoelectric response ($d_{31} > 1000\,\mathrm{pm\,V^{-1}}$) was discovered at cryogenic temperatures. The mechanism of the piezoelectric response in $SrTiO_3$ is presently not understood, but its divergence as $T \to 0\,\mathrm{K}$ suggests that it could be associated with a quantum critical point and field-induced phase transition into a piezoelectric and possibly ferroelectric phase at $\mathrm{T} \approx 0\,\mathrm{K}$. A piezoelectric-like response, where strain is reversibly proportional to charge, has been reported in the platinum metal with nanoporosity when submerged in an electrolyte, and an electric potential is applied between the metal and electrolyte. The piezoelectricity is related to the charge-induced strain at the interface (metal/pores) layer, and strain levels comparable to those in the best piezoelectric ceramics have been observed.

Polyvinylidene fluoride (PVDF) is a semicrystalline polymer, consisting of basic CH_2–CF_2 units which form a head-to-tail structure CH_2–CF_2–$(CH_2$–$CF_2)_n$–CH_2–CF_2. The crystalline part consists of a nonpolar phase and a highly polar phase, in which dipoles are due to the distribution of H^- and F^+ ions on opposite sides of the polymer chain. As grown, PVDF is macroscopically nonpolar due to random orientation of polymer chains, but can be rendered polar by the application of mechanical stress and electric field.

In some materials, the piezoelectric effect can be due to several processes. In ferroelectric $AgNa(NO_2)_2$, for example, the piezoelectric effect consists of two contributions: one related to ionic and electronic polarizations of the lattice, and the other to ordering of NO_2 radicals. The contribution of radicals relaxes at relatively low frequencies (30 kHz), so that the piezoelectric response exhibits a strong Debye-like relaxation.

Besides the piezoelectric response arising from the molecular or atomistic mechanisms (sometimes called intrinsic response), there are a number of the so-called

extrinsic contributions to the piezoelectric response whose value can be comparable to that of the intrinsic effect. In ferroelectric materials, for example, the displacement of ferroelectric-ferroelastic domain walls (e.g., 90° domain walls) at weak electric and elastic fields will modify both the polarization and shape of the sample, and thus contribute to the piezoelectric response. Both theoretical estimates and experimental studies show that this contribution of moving domain walls can account for as much as 50–70% of the total piezoelectric response. The movement of domain walls is nonlinear and hysteretic even at relatively weak fields, so that the piezoelectric response in ferroelectric materials often exhibits a charge–pressure and strain–electric field hysteresis (**Figure 3**). The piezoelectric hysteresis and nonlinearity are undesirable in many high-precision applications such as piezoelectric actuators used in scanners.

Besides the nonlinearity, the piezoelectric response in many ferroelectric materials exhibits a frequency dependence and creep. In heterogeneous materials, the dielectric (or elastic) dispersion may couple with elastic (or dielectric) and piezoelectric properties through the piezoelectric effect leading to dispersion of the latter. Interestingly, the charge–pressure or strain–electric field relationship can exhibit clockwise or counterclockwise hysteresis. This is possible because the strain and electric field, and charge density and pressure, are not conjugate work variables, and the corresponding hystereses do not possess energy units.

In complex materials, such as lead zirconate titanate, $Pb(Zr,Ti)O_3$ or PZT, it is possible to design properties to enhance or reduce extrinsic contributions, thus arriving at materials that possess either higher (but less stable) or more stable (but generally lower) properties. The former materials are called soft, and latter hard. Typical soft PZT is obtained by doping the ceramic with donor cations (e.g., replacing part of $(Ti,Zr)^{+4}$ with Nb^{+5}), whereas hard characteristics are obtained by doping the ceramics with acceptor cations (e.g., replacing part of $(Ti,Zr)^{+4}$ with Fe^{+3}). The amounts of dopants necessary to significantly change the properties are small, in the order of 0.5–2%. However, commercial hard and soft compositions differ also in composition, structure, and contain large concentrations of other additives, besides containing acceptor and donor dopants.

Piezoelectric Materials

The piezoelectric effect is widely encountered in nature and many synthetic materials. The best known natural piezoelectric material is quartz, SiO_2, while

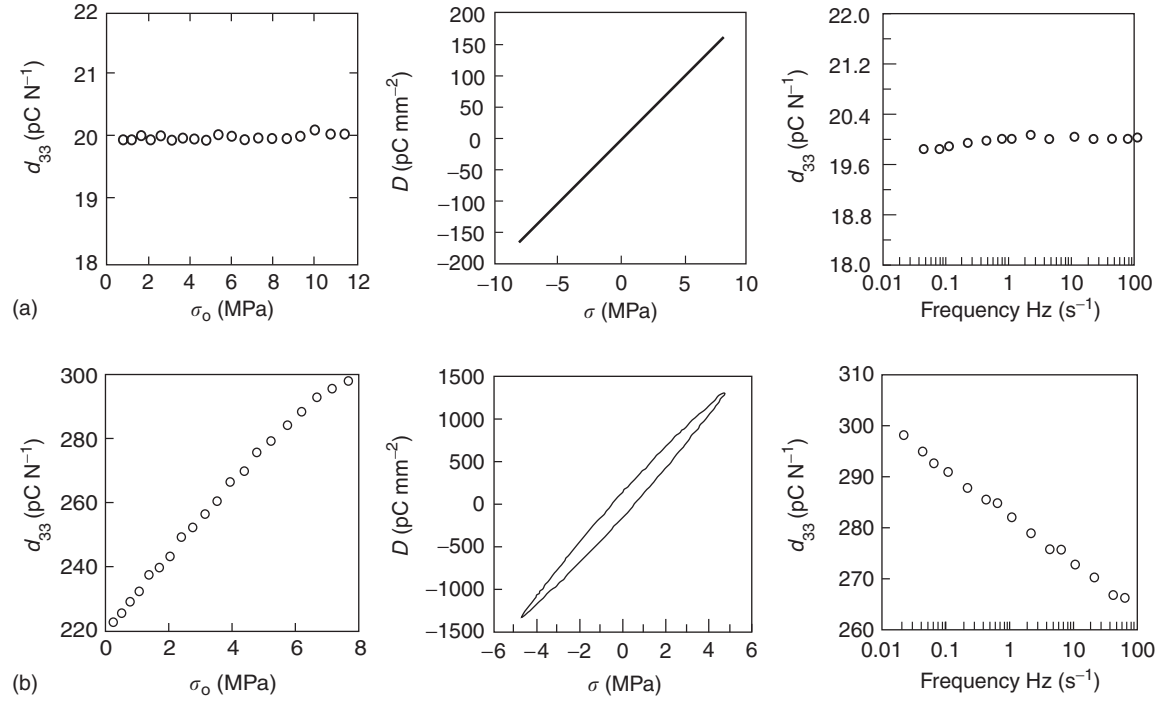


Figure 3 (a) Linear, nonhysteretic, and dispersion-free direct piezoelectric effect in $SrBi_4Ti_4O_{15}$ ceramics. (b) The direct piezoelectric effect in soft $Pb(Zr,Ti)O_3$ ceramics exhibits strong nonlinearity, charge–pressure hysteresis, and frequency-dependent response. σ_o indicates amplitude of the dynamic pressure. The charge–pressure curve is centered numerically whereas the sample is always under compressive stress during the measurements.

among thousands of synthetic materials, the most widely used and best understood are ferroelectric ceramics belonging to solid solutions of lead zirconate and lead titanate, $Pb(Zr,Ti)O_3$ (PZT). Quartz exhibits nearly perfect, unhysteretic, linear and stable, but a weak piezoelectric response. The piezoelectric coefficients of PZT are about two orders of magnitude higher than in quartz (**Table 1**). At present, PZT is not available in single crystal form; it is prepared as ceramics that need to be poled to exhibit the piezoelectric response. The best piezoelectric properties in PZT are observed near 1:1 Zr : Ti ratio. In this region (called morphotropic phase boundary region or MPB), the crystal structure of PZT changes from tetragonal ($PbTiO_3$-rich side) to rhombohedral ($PbZrO_3$-rich side) via an intermediary monoclinic structure (**Figure 4**). The following arguments are used to interpret the origin of exceptionally large piezoelectric properties in the MPB region: (1) because the free energies of the rhombohedral, tetragonal, and monoclinic phases at MPB are close to each other, there is a large number of available directions (6 tetragonal, 8 rhombohedral, 22 monoclinic) along which the polarization vector can reorient during the poling process; thus, a ceramic can be poled more efficiently, (2) at MPB, which represents a composition-induced phase transition region, some piezoelectric coefficients exhibit a maximum, just as in the vicinity of a temperature-induced phase transition, (3) those grains whose polar direction is oriented along an oblique direction with respect to the measuring field will, in general, exhibit a piezoelectric response which combines longitudinal, transverse, and shear effects; this increases the piezoelectric response because shear effects tend to be large at the MPB.

Recently, a new class of piezoelectric crystals was discovered with longitudinal and transverse piezoelectric coefficients reaching, along certain crystallographic directions, values of 1000–2500 $pC\,N^{-1}$ with shear coefficients in the range 4000–5000 $pC\,N^{-1}$, and coupling coefficients as large as 0.95. These crystals are solid solutions of a relaxor ferroelectric, such as $Pb(Zn_{1/2}Nb_{2/3})O_3$ or $Pb(Mg_{1/2}Nb_{2/3})O_3$, and ferroelectric $PbTiO_3$. As in PZT, the properties are maximized in the vicinity of the MPB (**Figure 4**), and along certain crystallographic directions, as shown in **Figure 5**. When a crystal is poled along a nonpolar direction, a special domain structure results, which is referred to as "engineered domain structure." In some

Table 1 Piezoelectric properties of selected materials

Material	*Piezoelectric coefficient* ($pC\,N^{-1}$)	*Coupling coefficient*	*Comment*
Soft PZT	$d_{33} = 600$	$k_{33} = 0.75$	Poled ceramic
Hard PZT	$d_{33} = 200$ $d_{15} = 500$	$k_{33} = 0.67$	Poled ceramic
PMN-PT	$d_{33} = 2500$ $d_{15} = 4900$	$k_{33} = 0.93$	Single crystal; d_{33} long [0 0 1] pseudocubic direction
SiO_2	$d_{11} = 2.3$	$k_{11} = 0.1$	Single crystal
PVDF	$d_{31} = 30$ $d_{33} = -15$	$k_{33} = 0.2$	Poled polymer

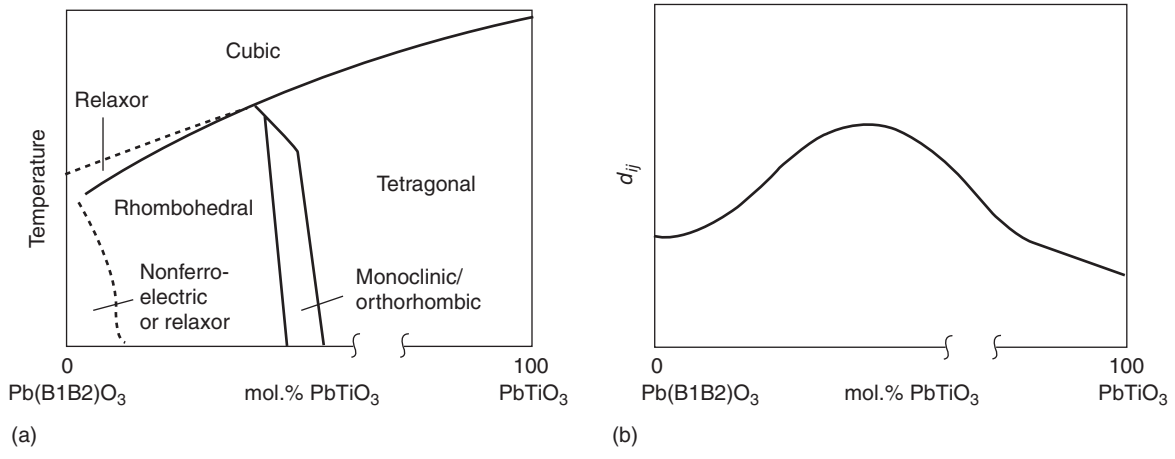


Figure 4 Schematic illustration of (a) temperature-composition phase diagram, and (b) piezoelectric coefficient–composition dependence, as observed in many solid solutions of simple and complex perovskites (e.g., $PbZrO_3$ and $Pb(Mg_{1/2}Nb_{2/3})O_3$) with $PbTiO_3$.

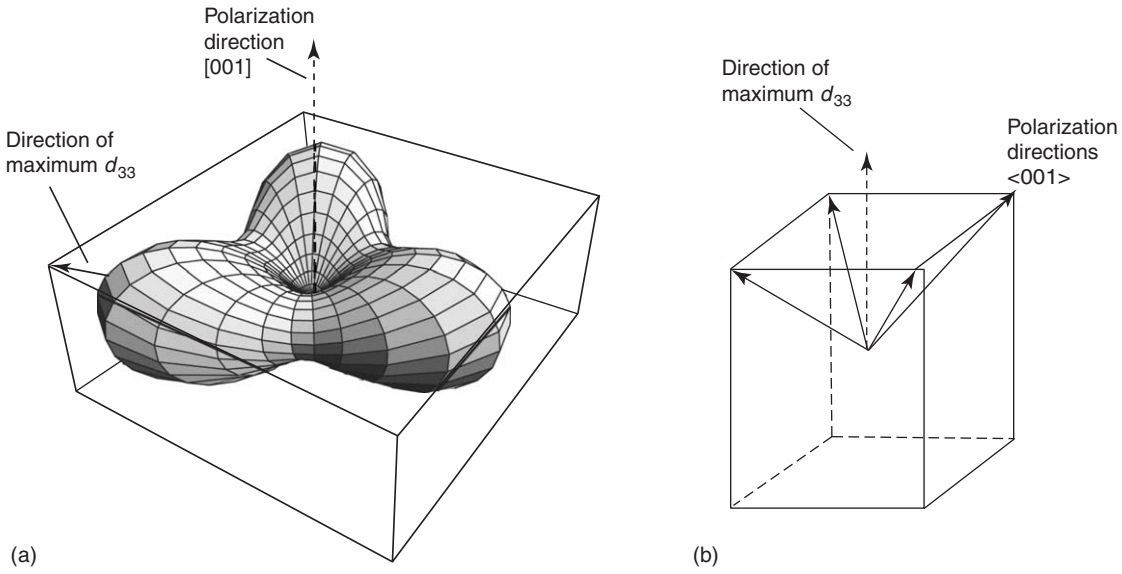


Figure 5 Dependence of the longitudinal d_{33} piezoelectric coefficient on crystal orientation for (a) crystals belonging to point group *3m*, such as rhombohedral PZT, $BaTiO_3$, or $Pb(Zn_{1/2}Nb_{2/3})O_3$-$PbTiO_3$, (b) the four domain states in such a crystal poled along pseudocubic [0 0 1] direction.

cases (e.g., rhombohedral crystals poled along $\langle 0\,0\,1 \rangle$ pseudocubic direction, **Figure 5**), the engineered domain state not only gives a large piezoelectric response but is also stable and hysteresis free.

Polymer–ceramic piezoelectric composites, which have been developed during the 1970s and 1980s, are now routinely used for medical imaging and underwater (SONAR) applications. The polymer phase is nonpiezoelectric while ceramic is a PZT composition. Different connectivities of the two phases are used. The role of the polymer is mainly to reduce the density and thus improve acoustical impedance matching to human body and water, and to reduce transverse piezoelectric effect and thus reduce cross-talk of elements in transducer arrays. The most important synthetic organic piezoelectric is PVDF.

PZT, quartz, and PVDF are the most widely used piezoelectric materials. Other than these materials, $LiNbO_3$ and $LiTaO_3$, tourmaline, modified $PbTiO_3$, lead metaniobate $PbNb_2O_6$, $BaTiO_3$, and materials belonging to the family of bismuth titanate-based aurivillius structures are used for special applications. There are some attempts to partly replace quartz by crystals belonging to the langasite $La_3Ga_5SiO_{14}$ family, and $GaPO_4$, both of which exhibit a higher piezoelectric response than quartz, temperature-stable properties and operating temperatures significantly higher than quartz. Recent developments in the area of piezoelectric thin films have raised interest in PZT and AlN films. Presently, there are attempts to develop lead-free piezoelectric materials (e.g., $KNbO_3$, $Na_{0.5}Bi_{0.5}TiO_3$ based compositions) with properties comparable to those of PZT. The *ab initio* theories appear to be a promising approach in predicting new materials with high piezoelectric properties.

Many organic, synthetic and natural, and biological materials are piezoelectric, including nylon, wood, hair, bones, and eye tissue. The piezoelectric effect in bovine eyes is several times higher than in quartz. The origin of piezoelectricity in animal and human tissue can be traced to collagen. However, presently, it is not clear whether the piezoelectric effect in living tissue has any physiological role.

Applications of Piezoelectric Materials

Piezoelectric applications can be divided into three general categories: sensors, actuators, and high-frequency transducers. The piezoelectric sensors are based on the direct effect and are used to measure and detect force, acceleration, and vibration by transforming them into electric (charge or voltage) signals. A typical PZT ceramic with d_{33} of 300 $pC\,N^{-1}$ will generate 3 nC when pressed with 10 N, an easily measured quantity. If short circuited, a cube of PZT with side 1 cm and with relative dielectric constant of 1000 will give voltage of $\sim$30 V when pressed with 10 N. The piezoelectric sensors are thus very sensitive. They are not effective at constant or very slowly varying forces, since all materials possess some conductivity and piezoelectric charge will eventually leak leading to drift and loss of the signal.

The actuators are based on the converse piezoelectric effect and are used to generate force and small displacement. The same PZT cube as above will change its size by only 0.3 nm when subjected to

100 V. This is excellent for small displacement control, for example, in atomic force and electron microscope scanners, but different schemes have to be worked out to amplify displacements of piezoelectrics for applications where μm range movement or higher is required (e.g., actuators for fuel injection).

Piezoelectric materials can be forced into mechanical resonance if driven by an electric field with the same frequency as the natural mechanical resonance frequency of the piezoelectric sample. A case which can be considered is that of a thin bar made of ferroelectric ceramics and poled along thickness direction, such that length ≫ thickness, width > thickness (**Figure 6a**). This reduces the problem to a one-dimensional case. An electric field is applied along the thickness (X_3 direction). At low frequencies of the driving field, the piezoelectric strain inside the sample can be considered to be homogeneous. However, as the wavelength of an electrically generated acoustical wave becomes comparable to the dimensions of the sample, strain becomes nonuniform. The Newton equation for a longitudinal wave propagating along X_1 direction is given by $\rho(\partial^2\delta_1/\partial t^2)\,\mathrm{d}V = F$, where δ_1 is the displacement of an elemental volume dV, and $F = (\partial\sigma_1/\partial X_1)\,\mathrm{d}V$ is the force. However, due to piezoelectric coupling, the stress depends on the electric field. The Newton equation then becomes modified by the piezoelectric constitutive eqn [1] resulting in the wave equation:

$$\begin{aligned}\partial^2\delta_1/\partial t^2 &= (1/s_{11}^E\rho)(\partial^2\delta_1/\partial X_1^2)\\ &= (v^E)^2\partial^2\delta_1/\partial X_1^2\end{aligned} \qquad [3]$$

which is obtained considering that $\partial E_3/\partial X_1 = 0$ and $x_1 = \partial\delta_1/\partial X_1$, where v^E is the velocity of the longitudinal elastic waves in the bar. For simple harmonic motion, $\delta_1 = \delta_0 e^{i\omega t}$. Taking into account the boundary

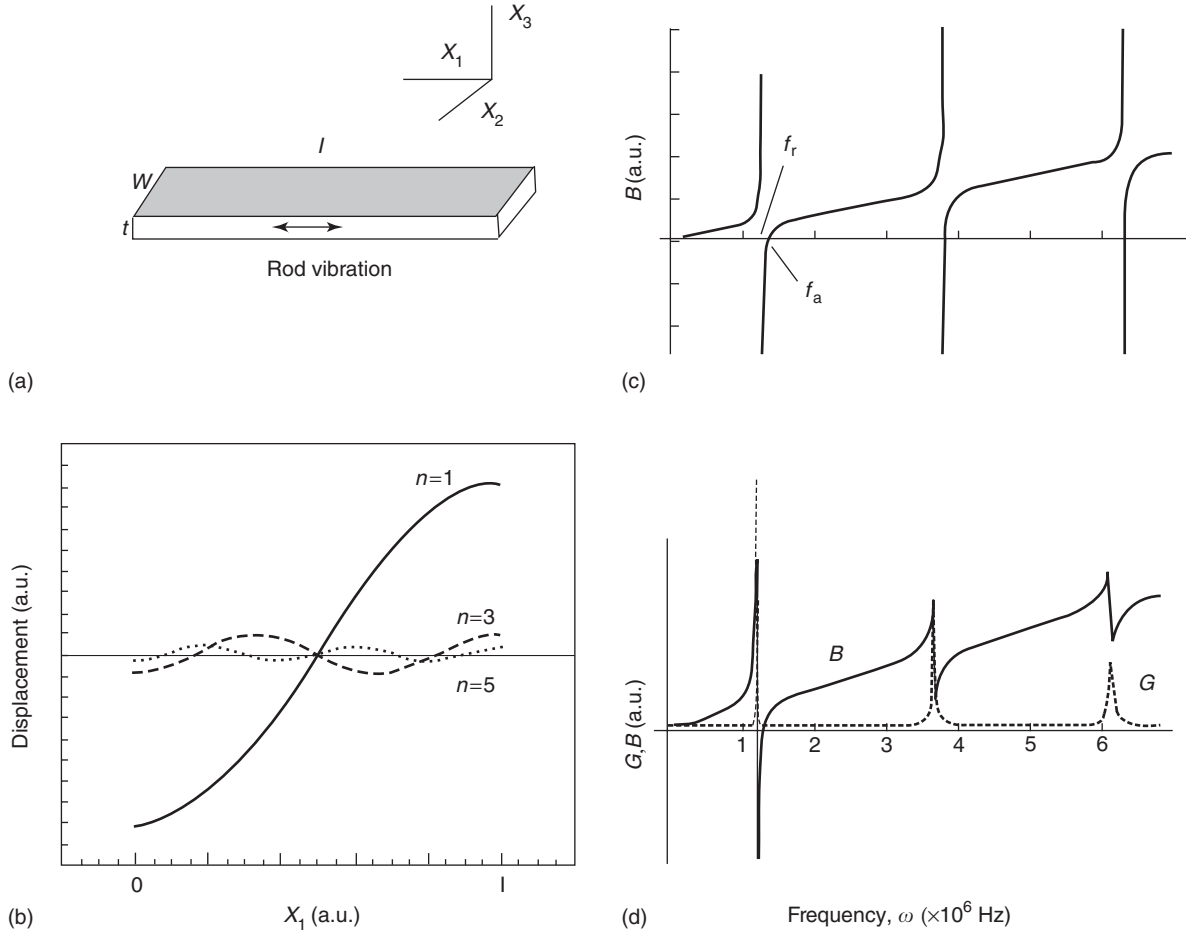


Figure 6 (a) Piezoelectric bar resonator made of a ferroelectric ceramic and poled along thickness. The shaded area indicates the electrode. Another electrode is placed on the bottom surface of the sample and the electric field is applied on these electrodes. (b) Mechanical longitudinal displacement as a function of position for the bar in transverse resonant mode (controlled by d_{31}). Admittance (Y) as a function of frequency for (c) lossless ($Y = iB$) and (d) lossy resonator ($Y = G + iB$). Resonant (f_r) and antiresonant (f_a) frequencies are marked in (c).

condition $\sigma_1 = 0$ at $X_1 = 0, l$ and eqn [1], expression [3] can be solved for δ_1 (**Figure 6b**). On the other hand, the electrical admittance Y of the vibrating bar is obtained from the ratio of the current I and voltage V across the vibrating bar $Y = I/V = w \int(\partial D_3/\partial t)\, \mathrm{d}X_1 / \int E_3\, \mathrm{d}X_3$, and can be calculated explicitly as a function of materials coefficients d, s, and ε using relations [1]–[3].

$$Y = \mathrm{i}\frac{\omega w l}{t}\left(\varepsilon_{33}^{\sigma} - \frac{d_{31}^2}{s_{11}^E}\right) + \mathrm{i}\frac{2wd_{31}^2}{(\sqrt{\rho s_{11}^E})s_{11}^E t}\tan\frac{\omega l(\sqrt{\rho s_{11}^E})}{2} \qquad [4]$$

Resonance (admittance becomes infinite) occurs when $\omega = n\pi/(l\sqrt{\rho s_{11}^E})$, where $n = 1, 3, 5, \ldots$. This is the same frequency at which mechanical resonance of an unloaded frictionless bar occurs. Resonances for $n > 1$ are called overtones.

In the presence of losses, purely imaginary admittance $Y = \mathrm{i}B$, where B is called susceptance and is shown in **Figure 6c**, will exhibit a real part, G, called conductance, $Y = G + \mathrm{i}B$, **Figure 6d**. Using characteristic resonant frequencies, the material coefficients can be calculated. For example, at resonance, f_r, admittance $Y \to \infty$, so that $\tan\lfloor \omega l(\sqrt{\rho s_{11}^E})/2 \rfloor \to \infty$ and $f_\mathrm{r} = 1/(2l\sqrt{\rho s_{11}^E})$, which can be used to calculate the elastic constant. At antiresonance, f_a, admittance $Y = 0$ and therefore

$$\tan[(\pi/2)(f_\mathrm{a}/f_\mathrm{r})]/[(\pi/2)(f_\mathrm{a}/f_\mathrm{r})] = (k_{31}^2 - 1)/k_{31}^2$$

which permits calculation of k_{31} and thus $d_{31} = k_{31}\sqrt{s_{11}^E \varepsilon_{33}^{\sigma}}$. The permittivity can be measured separately at subresonant frequencies. Values of the characteristic frequencies are modified in the case of lossy resonators.

High-frequency and resonant properties of piezoelectric materials have been widely used in many high-frequency applications, including ultrasonic medical and nondestructive imaging, frequency control (time measurements), filters, for chemical, biological and thickness monitoring, and in piezoelectric transformers.

See also: Crystal Symmetry; Ferroelectricity; Pyroelectricity.

PACS: 77.22. – d; 77.65. – j; 77.80. – e; 77.65.Bn; 77.65.Fs; 77.84.Dy

Further Reading

Ballato A (1996) Piezoelectricity: history and new thrusts. *IEEE Ultrasonics Symposium* 1: 575–583.

Bellaiche L (2002) Piezoelectricity of ferroelectric perovskites from first principles. *Current Opinion on Solid State Material Science* 6: 19–25.

Berlincourt DA, Curran DR, and Jaffe H (1964). Piezoelectric and piezomagnetic materials and their function in transducers. In: Mason WP (ed.) *Physical Acoustics-Principles and Methods*. vol. I, Part A. pp. 169–256. New York: Academic Press.

Cady WG (1964) *Piezoelectricity*. New York: Dover.

Damjanovic D, Demartin Maeder M, Voisard C, and Setter N (2001) Maxwell–Wagner piezoelectric relaxation in ferroelectric heterostructures. *Journal of Applied Physics* 90: 5708–5712.

Hamano K and Yamaguchi T (1982) Piezoelectric relaxation in ferroelectrics and polymers. *Ferroelectrics* 42: 23–33.

Herbert JM (1982) *Ferroelectric Transducers and Sensors*. New York: Gordon and Breach.

Ikeda T (1990) *Fundamentals of Piezoelectricity*. Oxford: Oxford University Press.

Jaffe B, Cook WR, and Jaffe H (1971) *Piezoelectric Ceramics*. New York: Academic Press.

Kawai H (1969) The piezoelectricity of polyvinylidene fluoride. *Japanese Journal of Applied Physics* 8: 975–976.

Lines ME and Glass AM (1979) *Principles and Applications of Ferroelectrics and Related Materials*. Oxford: Clarendon.

Muralt P (2000) Ferroelectric thin films for microsensors and actuators: a review. *Micromechanics and Microengineering* 10: 136–146.

Newnham R, Skinner DP, and Cross LE (1978) Connectivity and piezoelectric–pyroelectric composites. *Materials Research Bulletin* 13: 52.

Nye JF (1985) *Physical Properties of Crystals*. Oxford: Oxford University Press.

Park SE and Shrout TR (1997) Ultrahigh strain and piezoelectric behaviour in relaxor based ferroelectric single crystals. *Journal of Applied Physics* 82: 1804–1811.

Resta R (2003) *Ab initio* simulation of the properties ferroelectric materials. *Modelling and Simulation in Materials Science and Engineering* 11: R69–R96.

Tagantsev AK (1991) Electrical polarization in crystals and its response to thermal and elastic perturbations. *Phase Transitions* 35: 119–203.

Plasmons

W L Mochán, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

Introduction

Consider a metal whose conduction electrons move freely. In equilibrium, the negative electronic charge is balanced by the positive charge of the ionized metal atoms. If this balance is somehow disturbed by perturbing the electron density, increasing it within some region $\mathscr{R}$ which becomes thus negatively charged, the electrons would be subsequently driven away from that region by their mutual repulsion. Equilibrium is however not restored immediately, as the electrons have inertia and would keep on moving away from $\mathscr{R}$ even after the system becomes neutral. Thus, after some time interval, $\mathscr{R}$ would become positively charged. The electrons would then be attracted to this positive charge, arresting their outward motion and inducing an inward motion. By overshooting the equilibrium situation again, $\mathscr{R}$ would become negative again and the whole cycle would repeat itself until the energy associated with the perturbation is lost through Joule heating or some other dissipative process. Thus, an oscillation of the charge carrier density may be established within conductors. Due to the analogy between the electronic dynamics within a metal and the motion of charged ions within a plasma, the charge oscillations described above are known as plasma oscillations, and they have a characteristic frequency ω_p which depends on the mass, density, and charge of the carriers, known as the plasma frequency. Typically, this frequency is much higher than the thermal frequency $k_B T/\hbar$, where k_B is Boltzmann's constant, T is the temperature, and $\hbar$ is Planck's constant, so that the quantization of the plasma oscillations has to be taken into account when describing the electronic dynamics of solids. A quantized plasma oscillation of a conductor, with a typical energy $E_p = \hbar\omega_p$, is known as a "plasmon."

Simple Models

To estimate the plasma frequency, consider the most simple model of a conductor: that of a gas of electrons of number density n_0 moving freely in a homogeneous positively charged background and interacting with the self-consistent field $\boldsymbol{E}$. If somehow, a charge accumulation Q localized in a given region $\mathscr{R}$ is produced due to a compression or a rarefaction of the electron gas (**Figure 1**), it produces a Coulomb field $E(\boldsymbol{r}) = Q\boldsymbol{r}/r^3$ far away, where the vector $\boldsymbol{r}$ goes from $\mathscr{R}$ toward the observation position. Each electron becomes accelerated according to Newton's second law, $m\mathrm{d}^2\boldsymbol{r}/\mathrm{d}t^2 = -e\boldsymbol{E}(\boldsymbol{r},t)$, where m and $-e$ are the electronic mass and charge, respectively, so that $\partial\boldsymbol{j}(\boldsymbol{r},t)/\partial t = (n_0e^2/m)\boldsymbol{E}(\boldsymbol{r},t)$, where $\boldsymbol{j} = -n_0e\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ is the electric current density at $\boldsymbol{r}$. Integrating over a large closed surface Σ that surrounds the charge completely, one obtains $(\mathrm{d}/\mathrm{d}t)\int_\Sigma \mathrm{d}\boldsymbol{a}\cdot\boldsymbol{j} = (ne^2/m)\int_\Sigma \mathrm{d}\boldsymbol{a}\cdot\boldsymbol{E}$, which becomes a simple differential equation for the charge

$$\frac{\mathrm{d}^2Q}{\mathrm{d}t^2} = -\frac{4\pi ne^2}{m}Q \qquad [1]$$

by introducing the continuity equation $\int_\Sigma \mathrm{d}\boldsymbol{a}\cdot\boldsymbol{j} = -\mathrm{d}Q/\mathrm{d}t$ and Gauss' law $\int_\Sigma \mathrm{d}\boldsymbol{a}\cdot\boldsymbol{E} = 4\pi Q$. Its solution $Q(t) = Q(0)\mathrm{e}^{-\mathrm{i}\omega_p t}$ oscillates with the bulk plasma frequency

$$\omega_p = \left(\frac{4\pi ne^2}{m}\right)^{1/2} \qquad [2]$$

Typical values for the energy of a quantized plasma wave or plasmon are $E_p = \hbar\omega_p \sim 5$–$15$ eV.

If the charge were located at the surface of a semi-infinite conductor, half of the field lines would be in vacuum and would induce no electric current

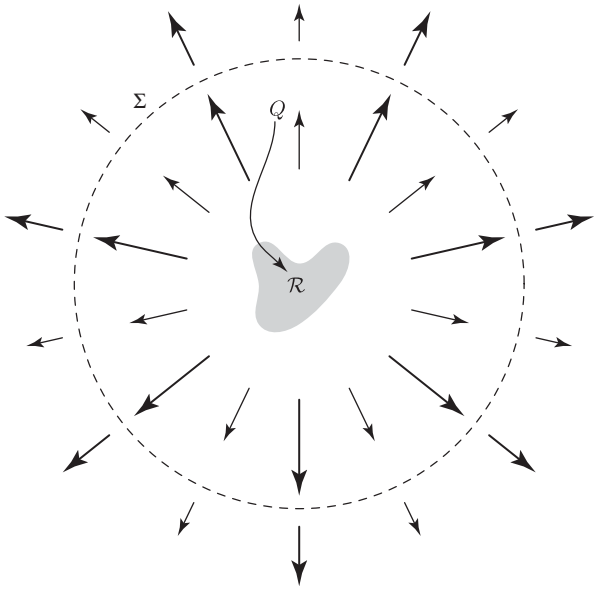


Figure 1 Charged region $\mathscr{R}$ (gray) in a homogeneous conductor. The charge Q contained within the Gaussian surface Σ (dashed circle) produces a field $\boldsymbol{E}(\boldsymbol{r},t)$ (wide arrows) which induces a current density $\boldsymbol{j}(\boldsymbol{r},t)$ (thin arrows) which flows across Σ, modifying Q over time t and yielding a plasma oscillation.

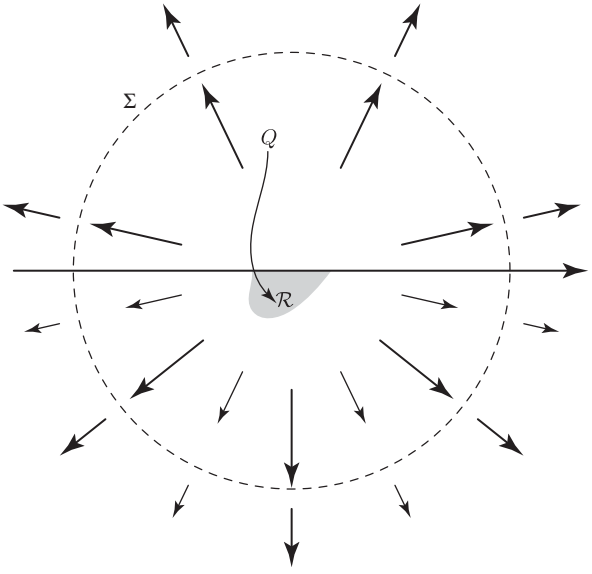


Figure 2 Charged region $\mathscr{R}$ at the surface of a semi-infinite conductor. The field $\boldsymbol{E}(\boldsymbol{r}, t)$ (wide arrows) is as in **Figure 1** but the current density $\boldsymbol{j}(\boldsymbol{r}, t)$ (narrow arrows) is only induced in the half space within the conductor. Thus, there is charge flow only across half the Gaussian surface Σ (dashed circle).

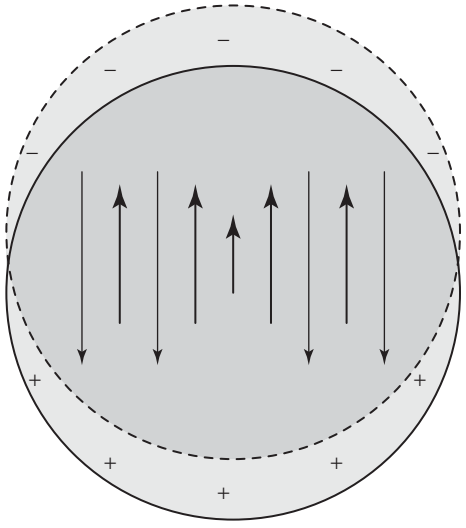


Figure 3 Metallic sphere (solid circle) all of whose electrons are displaced the same distance ζ (short arrow), occupying a displaced region (dashed circle), producing negative (top) and positive (bottom) surface charges (light gray). The interior (dark) remains neutral. The polarization (narrow arrows) and electric field (wide arrows) within the sphere are shown schematically.

(**Figure 2**), the current flow across the surface Σ would then be cut in half, and the charge would obey

$$\frac{\mathrm{d}^2 Q}{\mathrm{d}t^2} = -\frac{\omega_\mathrm{p}^2}{2} Q \qquad [3]$$

instead of eqn [1]. The solution $Q(t) = Q(0)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{s} t}$ is an oscillation with the surface plasma frequency

$$\omega_\mathrm{s} = \frac{\omega_\mathrm{p}}{\sqrt{2}} \qquad [4]$$

As a last simple example, consider a small spherical metallic particle which is initially perturbed by displacing all of its electrons by a distance $\boldsymbol{\zeta}$ (**Figure 3**). Thus, by a distance there would be an accumulation of electrons on one hemisphere and a deficiency of electrons on the other hemisphere. The ensuing surface charge $\sigma = \boldsymbol{P} \cdot \hat{\boldsymbol{n}}$ produces a homogeneous electric depolarization field $\boldsymbol{E} = -(4\pi/3)\boldsymbol{P}$, where $\boldsymbol{P} = -n_0 e\boldsymbol{\zeta}$ is the electric polarization. This field accelerates the charges, according to $m\mathrm{d}^2\boldsymbol{\zeta}/\mathrm{d}t^2 = -e\boldsymbol{E}$, which becomes

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{\zeta} = -\frac{\omega_\mathrm{p}^2}{3}\boldsymbol{\zeta} \qquad [5]$$

upon substitution. Its solution $\boldsymbol{\zeta}(t) = \boldsymbol{\zeta}(0)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{d} t}$ oscillates with the dipolar surface plasma frequency

$$\omega_\mathrm{d} = \frac{\omega_\mathrm{p}}{\sqrt{3}} \qquad [6]$$

These examples show that different systems may have plasmons, that is, collective excitations that involve fluctuations in the charge density, with different frequencies which depend not only on the properties of the material but also on geometrical properties such as size and shape.

Dispersion Relations

The previous section discussed a few simple models that predict charge oscillations of frequencies ω_p, ω_s, and ω_d at infinite, semi-infinite, and small conductors. In this section, these quantities will be calculated in a more general context, and expressions which are independent of the particular model made for the system will be obtained.

Bulk Plasmons

According to Gauss law,

$$\nabla \cdot \boldsymbol{D} = 0 \qquad [7]$$

in the absence of external charge, where $\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P}$ is the displacement field and $\boldsymbol{P}$ the polarization. Within a linear homogeneous isotropic material, the displacement may be written in terms of the electric field,

$$\boldsymbol{D} = \varepsilon\boldsymbol{E} \qquad [8]$$

where the dielectric function ε is a property of the material. Thus, Gauss law implies that $\varepsilon\nabla \cdot \boldsymbol{E} = 0$. The trivial solution $\nabla \cdot \boldsymbol{E} = 4\pi\rho = 0$ corresponds to the null charge density ρ within a homogeneous system and a transverse electric field. However, there is

a second frequently overlooked possibility, namely, that $\nabla \cdot \boldsymbol{E} = 4\pi\rho \neq 0$ but $\varepsilon = 0$. Thus, the electric field may have a longitudinal contribution, and charge fluctuations may appear within a solid, but only if the dielectric function is null.

Matter responds differently to electromagnetic fields of different frequencies or colors, as demonstrated by Newton's prism experiments. Consequently, the dielectric function depends on the frequency, that is, $\varepsilon \rightarrow \varepsilon_\omega$. According to the discussion above, charge oscillations may exist but only for those frequencies for which

$$\varepsilon_\omega = 0 \qquad [9]$$

This result may be applied to the Drude model, in which electrons are assumed to respond freely and independently to a perturbing electric field. One would expect that independent electrons would be unable to support collective oscillations. However, the electron–electron interaction is partially incorporated into the Drude model by identifying the perturbing field with the self-consistent electric field $\boldsymbol{E}$, which includes the mean value of the field produced by the electrons themselves, and not as an external field. The Drude dielectric function is

$$\varepsilon^{\mathrm{D}}_\omega = 1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2} \qquad [10]$$

and leads to plasma oscillations at its zero

$$\omega^{\mathrm{D}} = \omega_{\mathrm{p}} \qquad [11]$$

Dissipation may be incorporated by changing $\omega^2 \rightarrow \omega^2 + \mathrm{i}\omega/\tau$ in eqn [10] where τ is the mean time between electronic collisions.

Plasmon excitations within a solid are possible only due to the temporal dispersion of the dielectric function, as the dielectric constant may only vanish at some finite frequencies but never in the static limit. The meaning of dispersion may be understood if the frequency dependence is introduced explicitly into eqn [8] (i.e., $\boldsymbol{D}_\omega = \varepsilon_\omega \boldsymbol{E}_\omega$), which is then Fourier transformed using the convolution theorem, becoming

$$\boldsymbol{D}(t) = \int \mathrm{d}t' \varepsilon(t - t')\boldsymbol{E}(t') \qquad [12]$$

The Fourier transform of ε_ω yields the delay-dependent dielectric function $\varepsilon(t - t')$, which is a measure of the response $\boldsymbol{D}(t)$ of the system at time t to a perturbing field $\boldsymbol{E}(t')$ which acts at a time t'. Causality demands $\varepsilon(t - t') = 0$ whenever $t' > t$. Just as the response of the system at a given time depends on the excitation at previous times, it is also the case that the response of the system at a given position $\boldsymbol{r}$ depends on the excitation at nearby positions $\boldsymbol{r}'$, and not only at $\boldsymbol{r}$. Thus, eqn [12] should be generalized to become

$$\boldsymbol{D}(\boldsymbol{r}, t) = \int \mathrm{d}t' \int \mathrm{d}^3 r' \varepsilon(\boldsymbol{r}, \boldsymbol{r}', t - t')\boldsymbol{E}(\boldsymbol{r}', t') \qquad [13]$$

The dependence of ε on $\boldsymbol{r}$ and $\boldsymbol{r}'$ is known as spatial dispersion or nonlocality. Within a homogeneous system, ε depends only on the interval $\boldsymbol{r} - \boldsymbol{r}'$ between the excitation and the observation point, so that the RHS of eqn [13] becomes a convolution. A space and time Fourier transform then yields a simple algebraic relation

$$\boldsymbol{D}_{q\omega} = \varepsilon_{q\omega}\boldsymbol{E}_{q\omega} \qquad [14]$$

in which nonlocality is equivalent to a dependence of the dielectric response $\varepsilon_{q\omega}$ on the wave vector $\boldsymbol{q}$ besides the frequency ω. Within nonlocal media, $\boldsymbol{q}$ defines a preferred direction even within isotropic media. Therefore, the nonlocal response of the system to a transverse field $\boldsymbol{E}^{\mathrm{T}}_{q\omega} \perp \boldsymbol{q}$ and to a longitudinal field $\boldsymbol{E}^{\mathrm{L}}_{q\omega} \parallel \boldsymbol{q}$ might differ; hence,

$$\boldsymbol{D}^{\mathrm{T}}_{q\omega} = \varepsilon^{\mathrm{T}}_{q\omega}\boldsymbol{E}^{\mathrm{T}}_{q\omega} \qquad [15]$$

$$\boldsymbol{D}^{\mathrm{L}}_{q\omega} = \varepsilon^{\mathrm{L}}_{q\omega}\boldsymbol{E}^{\mathrm{L}}_{q\omega} \qquad [16]$$

instead of the single eqn [14].

Identifying $\nabla \rightarrow \mathrm{i}\boldsymbol{q}$ in Fourier space, the dispersion relation $\omega = \omega_q$ versus $\boldsymbol{q}$ of bulk plasmons is deduced from eqn [7] and is expressed implicitly by

$$\varepsilon^{\mathrm{L}}_{q\omega} = 0 \qquad [17]$$

that is, for a given wave vector $\boldsymbol{q}$, longitudinal waves may propagate provided their frequency ω_q is such that the longitudinal dielectric function vanishes.

Notice that plasmons do not propagate within local media as their group velocity $\boldsymbol{v}_{\mathrm{g}} = \nabla_q \omega_q$ vanishes, but in a spatially dispersive medium, the frequency of the plasmon does depend on the wave vector and plasmons may propagate and transport energy. A simple model that illustrates plasmon propagation is the hydrodynamic model, a generalization of the Drude model, in which the conduction electrons feel the forces due to variations of the pressure beyond their interaction with the self-consistent field. The pressure is mainly due to the Pauli principle which allows only one fermion to occupy each electronic state. To increase the density of an electronic fluid, higher energy levels have to be occupied and this requires work. The hydrodynamic longitudinal dielectric function is

$$\varepsilon^{\mathrm{H}}_{q\omega} = 1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2 - \beta^2 q^2} \qquad [18]$$

where β is a measure of the compressibility of a Fermion gas of density n_0. A calculation of β assuming local equilibrium of the electron gas yields

$\beta = \sqrt{1/3}v_F$, where $v_F = \hbar k_F/m$ is the Fermi velocity, $\hbar k_F = \sqrt{2mE_F}$ is the Fermi momentum, and E_F is the Fermi energy, that is, the kinetic energy of the highest occupied state. More elaborate calculations are consistent with a frequency-dependent β with the value $\beta = \sqrt{3/5}v_F$ at typical plasmon frequencies. The dispersion relation [17] of bulk plasmons becomes

$$\omega = (\omega_p^2 + \beta^2 q^2)^{1/2} \qquad [19]$$

which is illustrated schematically in **Figure 4**.

Notice that eqn [19] may also be solved for q, given the value of ω. For $\omega < \omega_p$ the solution is imaginary, meaning that charge fluctuations may exist for frequencies below ω_p, but they do not propagate. Rather, they decay exponentially within a distance $\lambda_s = 1/|q|$. In the static limit, one obtains $\lambda_s = \lambda_{TF}$ from eqn [19], where $\lambda_{TF} = \beta/\omega_p$ is the Thomas–Fermi screening length.

One of the main drawbacks of the hydrodynamic model is that it ignores the single-particle excitations. For a free-, independent-electron conductor, these consist of electrons which are taken from occupied states with momentum $\hbar\boldsymbol{k}$ within the Fermi sphere (i.e., $\hbar k < \hbar k_F$) and excited into empty states with momentum $\hbar(\boldsymbol{k} + \boldsymbol{q})$ outside of the Fermi sphere, leaving a hole in the initial state. The momentum of these electron–hole (e–h) pairs is $\hbar\boldsymbol{q}$ and its energy $\hbar^2[(\boldsymbol{k} + \boldsymbol{q})^2 - k^2]/2m$ may take a quasicontinuum of values bounded by $\hbar^2(q^2 \pm 2qk_F)/2m$. Plasmons with wave vector q and energy $\hbar\omega$ within these bounds, displayed in **Figure 4** by dashed lines, are short-lived as they may decay through the excitation of e–h pairs, a mechanism known as Landau damping. Thus, plasmons may exist as well-defined collective excitations only within a short range of wave vectors and frequencies.

An alternative formulation for the calculation of the plasmon dispersion may be obtained from the inhomogeneous Gauss' law, which may be written in Fourier space with the help of eqn [16] as

$$\rho_{q\omega} = \frac{\rho^{ext}_{q\omega}}{\varepsilon^{L}_{q\omega}} \qquad [20]$$

where $\rho^{ext} = i\boldsymbol{q} \cdot \boldsymbol{D}_{q\omega}/4\pi$ is the external, and $\rho = i\boldsymbol{q} \cdot \boldsymbol{E}_{q\omega}/4\pi$ the total charge density within the conductor. Thus, the inverse dielectric function $1/\varepsilon^{L}_{q\omega}$ may be interpreted as a charge–charge response function. A finite response may be obtained from an infinitesimal excitation, that is, the system may oscillate spontaneously, only when the response function has a pole. Thus, the plasmon dispersion relation is given by the pole of $1/\varepsilon^{L}_{q\omega}$, which, naturally, coincides with eqn [17].

Surface Plasmons

As suggested by the remarks above, the normal modes of an arbitrary system may be obtained by searching for the poles of its response functions. Consider an electromagnetic wave of amplitude E^i impinging on the surface of an isotropic semi-infinite conductor (**Figure 5**). The wave sets in motion the charges within the conductor which radiates a reflected wave of amplitude E^r. For s (transverse electric) and p (transverse magnetic) incident polarizations, the reflected wave has the same polarization, so one may define scalar reflection amplitudes $r_s = E^r_s/E^i_s$ and $r_p = E^r_p/E^i_p$, given by Fresnel's formulas:

$$r_s = \frac{q_\perp - k_\perp}{q_\perp + k_\perp} \qquad [21]$$

$$r_p = \frac{\varepsilon_\omega q_\perp - k_\perp}{\varepsilon_\omega q_\perp + k_\perp} \qquad [22]$$

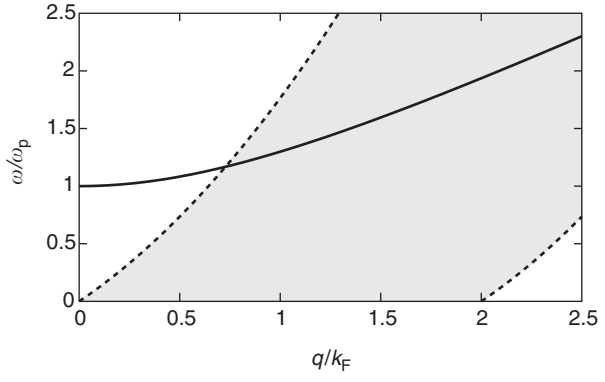


Figure 4 Dispersion relation for bulk plasmons calculated for Na within the hydrodynamic model (solid). The shadowed area between dashed lines corresponds to the region where plasmons might decay due to excitation of electron–hole pairs.

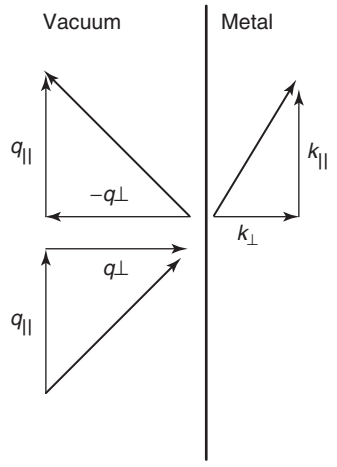


Figure 5 Electromagnetic wave incident on the surface of a semi-infinite system. The wave vectors of the incident, reflected, and transmitted waves, and their perpendicular and parallel components are shown.

where $\boldsymbol{q} = (\boldsymbol{q}_\parallel, q_\perp)$ is the wave vector of the incident wave in vacuum, and $\boldsymbol{k} = (\boldsymbol{k}_\parallel, k_\perp)$ is the wave vector of the transmitted wave into the metal. Snell's law requires that $\boldsymbol{k}_\parallel = \boldsymbol{q}_\parallel$ and the electromagnetic wave equations in vacuum and within the metal require

$$q_\perp = \left[\left(\frac{\omega}{c}\right)^2 - q_\parallel^2\right]^{1/2} \qquad [23]$$

and

$$k_\perp = \left[\varepsilon_\omega\left(\frac{\omega}{c}\right)^2 - q_\parallel^2\right]^{1/2} \qquad [24]$$

The reflection amplitudes [21] and [22] have the nature of surface response functions and therefore, their poles yield the surface normal modes of the system. One may note first that conservation of energy apparently implies that the reflectances $R_s = |r_s|^2$ and $R_p = |r_p|^2$ ought to be smaller than unity, as it is not possible to reflect more energy than is incident upon the surface. Thus, it would seem that r_s and r_p can have no pole. However, if $q_\parallel > \omega/c$, that is, outside of the light cone, eqn [23] yields an imaginary value for $q_\perp$. In this case, the incident and reflected waves do not propagate but are rather evanescent waves, each of which carries no energy along the normal to the surface. Only in this situation can the reflection amplitudes display poles. As both the reflected and transmitted waves should decay away from the surface, the signs of Im $q_\perp$ and Im $k_\perp$ should be equal to each other. Equations [21] and [22] imply that there is no *s*-polarized mode and that a *p*-polarized mode is only possible if $\varepsilon_\omega < 0$. Substituting eqns [23] and [24] into the normal mode condition $\varepsilon_\omega q_\perp + k_\perp = 0$, one obtains the surface plasmon dispersion relation

$$q_\parallel = \frac{\omega}{c}\left(\frac{\varepsilon_\omega}{\varepsilon_\omega + 1}\right)^{1/2} \qquad [25]$$

Figure 6 shows the dispersion relation of surface plasmons (SP) calculated from eqn [25] with a Drude dielectric function. The dispersion relation lies outside of the light cone. For large $q_\parallel$, ω approaches asymptotically a constant frequency given by ω_s (eqn [4]), while for small $q_\parallel$ it approaches the light cone.

For very large wave vectors $q_\parallel \gg \omega/c$, the phase velocity of the SP, $v = \omega/q_\parallel$, becomes much smaller than the speed of light c, so that the retardation of the electromagnetic field may be ignored. It is common to set $c = \infty$ in this nonretarded region and treat the problem as quasistatic, that is, solving the equations of electrostatics. In this limit, $q_\perp \approx k_\perp \approx iq_\parallel$, and therefore the reflection amplitude [22] becomes

$$r_p \approx -g = \frac{\varepsilon_\omega - 1}{\varepsilon_\omega + 1} \qquad [26]$$

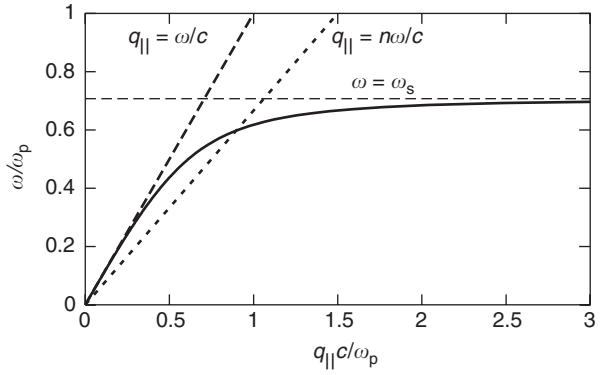


Figure 6 SP dispersion relation calculated within the Drude model (solid). The asymptotic frequency $\omega_s = \omega_p/\sqrt{2}$, the vacuum light cone $q_\parallel = \omega/c$, and the light cone of glass ($n = 1.5$) are indicated by dashed lines.

The nonretarded surface response function g is the quotient between the external and the induced scalar potentials, and its expression [26] may be recognized as the image charge induced within a conductor by the presence of a unit point charge within vacuum. The pole of g,

$$\varepsilon_\omega = -1 \qquad [27]$$

yields the nonretarded SP frequency. This equation generalizes the previous nonretarded result [4] to arbitrary materials.

More refined microscopic theories of metallic surfaces should account for the thin but finite transition region or selvedge where the properties change smoothly from those of the bulk metal to those of vacuum, and for the spatial dispersion of the metal. Their effect may be characterized within the semi-infinite jellium model by the distance from the nominal surface of the metal to the centroid of the induced charge density,

$$d_\perp \equiv \frac{\int \mathrm{d}z z\rho(z)}{\int \mathrm{d}z \rho(z)} \qquad [28]$$

The response function g depends on $\boldsymbol{q}_\parallel$ and ω and may be written as

$$g_{q_\parallel \omega} \approx \frac{1 - \varepsilon_\omega + (1 - \varepsilon_\omega) q_\parallel d_\perp}{1 + \varepsilon_\omega + (1 - \varepsilon_\omega) q_\parallel d_\perp} \qquad [29]$$

with poles at

$$\varepsilon_\omega + 1 = (\varepsilon_\omega - 1) q_\parallel d_\perp \qquad [30]$$

Using the Drude model for the local limit ε_ω of the bulk dielectric function, one obtains

$$\omega = \omega_s \sqrt{1 - q_\parallel d_\perp} \approx \omega_s - \omega_s q_\parallel d_\perp / 2 \qquad [31]$$

Thus, instead of reaching an asymptotic value ω_s, which is actually a bulk property determined by ε_ω,

the nonretarded SP has a dispersion determined by $d_\perp$, a quantity that is very sensitive to the surface condition. The sign of the dispersion is positive or negative corresponding to whether the induced surface charge lies inside ($d_\perp < 0$) or outside ($d_\perp > 0$) of the metal. The latter is the case for simple metals.

Thin Films

The reflection amplitude of a thin film may be calculated from that of each of its surfaces by summing the amplitudes of the multiply reflected waves. Thus, for the thin film of **Figure 7**

$$r = r_{12} + t_{12}r_{23}t_{21}e^{2ik_\perp d}\sum_{n=0}^{\infty}(r_{21}r_{23}e^{2ik_\perp d})^n \quad [32]$$

$$= r_{12} + \frac{t_{12}r_{23}t_{21}e^{2ik_\perp d}}{1 - r_{21}r_{23}e^{2ik_\perp d}} \quad [33]$$

where each successive term in the geometric series corresponds to an additional trip from the 1–2 interface toward the 2–3 interface and back, $k_\perp$ is the normal wave vector within the film, and r_{ab} and t_{ab} denote the reflection and transmission amplitudes of the surface ab for a wave moving from medium a toward b. The dispersion relation of the thin film SP is then given implicitly by the poles of r:

$$r_{21}r_{23}e^{2ik_\perp d} = 1 \quad [34]$$

For an unsupported film $r_{21} = r_{23} = -r_{12}$, and in the unretarded regime, using eqn [26] for r_{21},

$$\left(\frac{\varepsilon_\omega - 1}{\varepsilon_\omega + 1}\right)^2 = e^{2q_\parallel d} \quad [35]$$

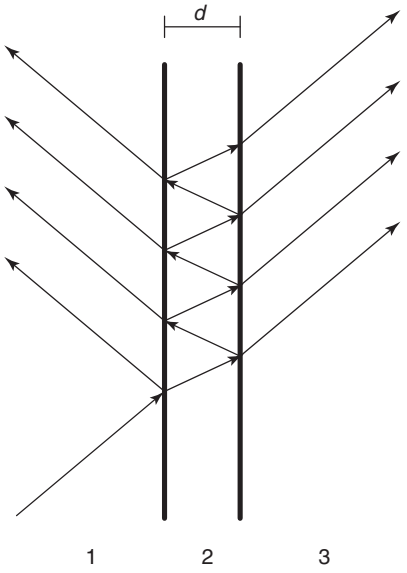


Figure 7 Reflection of an electromagnetic wave from a thin film 2 of width d bounded by media 1 and 3.

with solutions

$$\varepsilon_\omega = -\begin{cases} \coth(q_\parallel d/2) & \text{(even)} \\ \tanh(q_\parallel d/2) & \text{(odd)} \end{cases} \quad [36]$$

according to the parity of the surface charge. **Figure 8** illustrates the dispersion relation of unretarded SPs within a thin unsupported film described by the Drude model. Each surface supports by itself an SP of frequency ω_s. However, the field of each SP penetrates the film a distance $1/q_\parallel$ and thus may couple to the SP on the opposite surface, lifting the degeneracy. For small wave vectors, the odd coupled mode oscillates with the bulk plasma frequency ω_p, while the even coupled mode acquires an acoustical dispersion propagating with speed $v_s = \omega/q_\parallel = \omega_p d/2 = 2\pi n_0 de^2/m$, that is,

$$v_s = 2\pi n_s e^2/m \quad [37]$$

where $n_s = n_0 d$ is the number of electrons per unit area in the film.

Spherical Particles

The scalar potential $\psi_{lm}^{ind}(r,\theta,\phi) = C_{lm}^{ind}Y_{lm}(\theta\phi)r^{-l-1}$ produced at $r > a$ by the charges induced at the surface of a metallic sphere of radius a, subject to an external potential $\psi_{lm}^{ext}(r,\theta,\phi) = C_{lm}^{ext}Y_{lm}(\theta,\phi)r^l$ with angular momentum lm may be obtained in the nonretarded limit by solving Laplace's equation $\nabla^2\psi = 0$ within the sphere, and in the surrounding vacuum and imposing boundary conditions at $r = a$. This procedure yields the response function of the sphere,

$$S_{lm} \equiv \frac{C_{lm}^{ind}}{C_{lm}^{ext}} = -\frac{\varepsilon_\omega - 1}{\varepsilon_\omega + (l+1)/l}\,a^{2l+1} \quad [38]$$

for each value of l and m. Thus, the collective normal modes of the sphere, that is, its multipolar

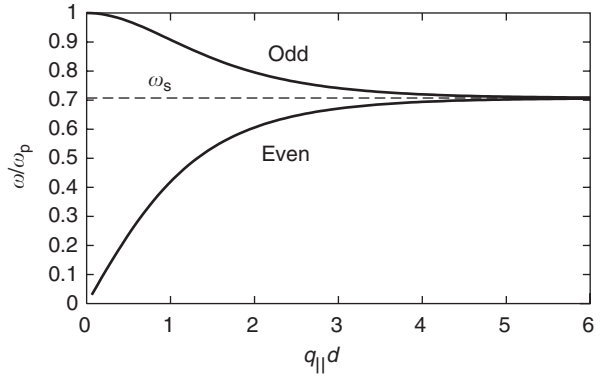


Figure 8 Dispersion relation of the odd and even SPs of a thin metallic film described by the Drude model. The SP frequency of a single surface is indicated by a dashed line.

surface plasmons, are given by the zeroes of the denominator:

$$\varepsilon_\omega = -\frac{l+1}{l} \qquad [39]$$

For the Drude model without dissipation,

$$\omega = \sqrt{\frac{l}{2l+1}}\,\omega_p \qquad [40]$$

This result yields ω_d for the case $l=1$. Thus, eqn [39] is a generalization of eqn [6] to conductors with arbitrary dielectric function, and yields modes with arbitrary angular momentum l. Notice that in the limit $l \to \infty$, eqn [39] coincides with the SP relation for a flat surface [27].

Observation

As plasmons are charge oscillations, they may couple to charged particles. A charged particle shot through the bulk of a conductor with velocity $\boldsymbol{v}$ induces screening charges of the opposite sign in its neighborhood. However, as the projectile is in motion and there is some delay before the screening is fully developed, the center of the screening charge lags behind the position of the projectile, as shown schematically in **Figure 9**. Thus, the attraction force toward the screening charge slows down the projectile, taking energy and momentum from it. Quantum mechanically, the particle loses energy $\hbar\omega$ and momentum $\hbar\boldsymbol{q}$ due to the emission of plasmons with frequency ω and wave vector $\boldsymbol{q}$. Only plasmons with velocities that match that of the particle, that is, with $\omega = \boldsymbol{q}\cdot\boldsymbol{v}$ are emitted, and with a probability per unit time which is proportional to the loss function $-\mathrm{Im}(1/\varepsilon^{L}_{q\omega})$. The loss function displays a sharp peak whenever $(\boldsymbol{q},\omega)$ approaches the plasmon dispersion relation [17]. For example, plasmon excitation may be observed with electron energy loss spectroscopy (EELS), in which high-energy electrons are shot into a metallic thin film in a transmission electron microscope, and the intensity of the electrons which are inelastically scattered into different directions is measured.

Consider now a relatively low-energy electron which collides with the surface of a conductor where it is specularly reflected. On its way toward the surface and back, it induces screening charges which produce a field that acts back on the projectile. According to the image theory, the field may be calculated in terms of a fictitious image charge within the metal. The image charge is not a point charge and it lags behind the ideal image position due to the delay in the development of the screening. Thus, the energy and momentum gained from the image force as the projectile approaches the surfaces are smaller than those lost as the projectile rebounds (**Figure 10**). Quantum mechanically, the electron loses energy $\hbar\omega$ and momentum $\hbar\boldsymbol{q}_\parallel$ along the surface by emitting SPs with a probability proportional to the surface loss function $-\mathrm{Im}\, g_{q_\parallel \omega}$. The loss function displays a sharp peak whenever $(\boldsymbol{q}_\parallel,\omega)$ approaches the SP dispersion relation [30], which may therefore be measured with low-energy reflection EELS.

Besides being excited by external charges, plasmons may also be excited by electromagnetic fields. The electric field of p-polarized light impinging on the surface of a metal may push the charge against the surface, and the resulting charge accumulation may subsequently propagate toward the bulk as a plasmon or along the surface as an SP. In the former case, the reflectance is slightly reduced below that given by the Fresnel amplitude [22] at frequencies above ω_p. In order to couple to SPs, the

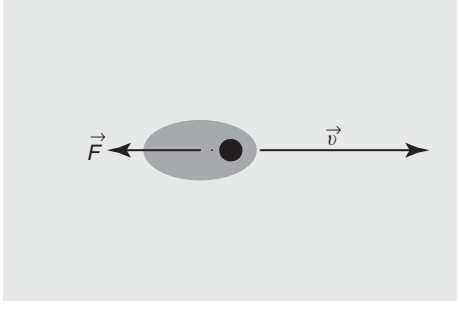


Figure 9 An external charge (black dot) moving with speed $\boldsymbol{v}$ within the bulk of a conductor (light gray), induces an excentric screening charge (dark gray) which produces a decelerating force $\boldsymbol{F}$.

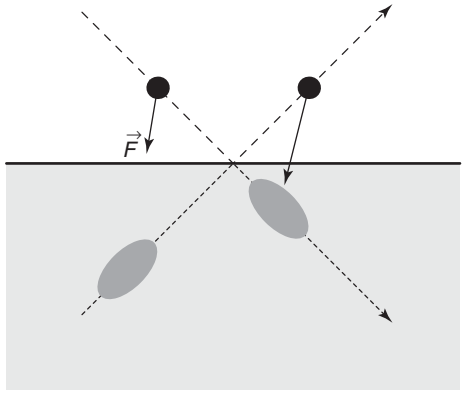


Figure 10 An external electron (black dot) collides with the surface of a metal (light gray) where it is specularly reflected. Along its trajectory (dashed), it interacts with its image (dark gray) being accelerated before the collision and decelerated afterwards by different forces (arrows), losing energy and momentum in the process.

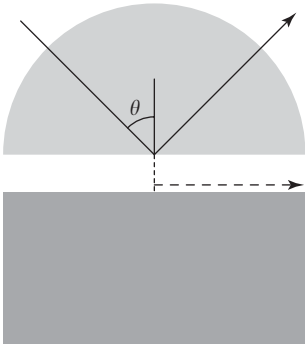


Figure 11 ATR experiment. *P* polarized light is incident at an angle θ on the surface of a glass semicylinder where it is totally reflected (solid arrow). The evanescent transmitted field (short dashes) reaches the surface of the metal, where it may excite an SP (dashed), attenuating the total reflection.

incoming wave should have a wave vector $q_{\parallel} > \omega/c$. This may be accomplished in an attenuated total reflection (ATR) experiment in which the surface is illuminated through a glass prism with an index of refraction $n > 1$ (**Figure 11**), in which the light line $n\omega/c$ extends beyond that of vacuum (**Figure 6**). By varying θ and/or ω, the region $\omega/c \leqslant q_{\parallel} = n(\omega/c)\sin\theta \leqslant n\omega/c$ may be explored (**Figure 6**). The reflectance of the system is close to 1 due to total reflection in the glass–air interface, except for a sharp dip when $(\boldsymbol{q}_{\parallel}, \omega)$ crosses the dispersion relation of the SP [25]. SPs may also be excited at a rough surface, where $\boldsymbol{q}_{\parallel}$ is no longer a conserved quantity and SPs may couple to ordinary light. The electric field close to the surface may be amplified several orders of magnitude due to SP excitations. This amplification is partly responsible for the enhanced Raman scattering by molecules adsorbed at rough metallic surfaces.

Final Remarks

Plasmons are present and have been studied in manifold systems: quasi-two-, one-, and zero-dimensional structures, in the presence of strong static electric and magnetic fields, in small particles of various shapes, etc. Their frequency may be obtained from the poles of appropriate response functions, typically, the total charge–external charge response, commonly identified with the inverse dielectric operator $\rho = \hat{\varepsilon}^{-1}\rho^{\text{ext}}$. The operator $\hat{\varepsilon}^{-1}$ may be represented by a matrix (ε^{-1}) introducing a suitable basis $|\alpha\rangle$, and each of the matrix elements $(\varepsilon^{-1})_{\alpha\beta}$ depends on frequency and may depend on other conserved quantities, such as momentum or angular momentum, if they are present. Plasmons are then given by the poles of the (ε^{-1}), or equivalently, from $\det(\varepsilon_{\alpha\beta}) = 0$. Plasmons play a fundamental role in the screening of electromagnetic fields within matter, and may be detected through inelastic scattering of charged particles or through resonant scattering of light.

Acknowledgment

This work was partially supported by DGAPA-UNAM under grant IN117402.

See also: Electrodynamics: Continuous Media; Electron Gas (Theory); Excitons in Crystals; Local Field Effects; Metals and Metallic Alloys, Optical Properties of; Molecular Clusters; Molecular Monolayers and Thin Films; Optical Absorption and Reflectance; Optical Properties of Materials; Optical Sum Rules and Kramers–Kronig Relations; Polaritons; Polarizabilities; Semiconductors, Optical and Infrared Properties of; Small Particles and Clusters, Optical Properties of; Spin Density Waves and Magnons; Surfaces and Interfaces, Electronic Structure of; Surfaces, Optical Properties of; Treated Surfaces, Optical Properties of.

PACS: 71.45.Gm; 73.20.Mf; 78.20.−e; 78.67.−n; 78.68.+m; 78.70.−g; 36.40.Gk

Further Reading

Halevi P (ed.) (1992) *Spatial Dispersion in Solids and Plasmas*. Amsterdam: North-Holland.
Halevi P (ed.) (1995) *Photonic Probes of Surfaces*. Amsterdam: North-Holland.
Ibach H and Luth H (1995) *Solid-State Physics: An Introduction to Principles of Materials Science*, 2nd edn. Berlin: Springer.
Ibach H and Mills DL (1982) *Electron Energy Loss Spectroscopy and Surface Vibrations*. New York: Academic Press.
Kittel C (1953) *Introduction to Solid State Physics*. New York: Wiley.
Kittel C (1963) *Quantum Theory of Solids*. New York: Wiley.
Liebsch A (1997) *Electronic Excitations at Metal Surfaces*. New York: Plenum.
Raether H (1980) *Excitation of Plasmons and Interband Transitions by Electrons*. Berlin: Springer.
Raether H (1988) *Surface Plasmons*. Berlin: Springer.

Plastic Behavior *See* Mechanical Properties: Plastic Behavior.

Point Defects

W B Fowler, Lehigh University, Bethlehem, PA, USA

Introduction

As its name implies, a "point defect" is an atomic-scale imperfection in an otherwise perfect crystal. In this respect, it may be contrasted with an "extended defect," which extends through a large portion of the crystal. Much of solid-state technology is associated with point defects, and indeed this topic is central to many of the articles in this encyclopedia. Hence, this article will address general properties of point defects in condensed matter, while other articles delve into more specific properties and applications.

Types of Point Defects

In a solid that contains no chemical impurities, the fundamental intrinsic point defects are vacancies, interstitials, and antisites. A "vacancy" (also called a "Schottky defect") is a missing atom from a normally occupied site, while an "interstitial" is an extra atom in a normally unoccupied site. A "Frenkel defect" results when an atom is moved from a normal site to an interstitial site. In compound solids (e.g., GaAs), an "antisite" defect exists when an atom of one of the elements is located at a site normally occupied by that of another element.

A compound solid has another degree of freedom that can lead to vacancies and interstitials, namely chemical stoichiometry. A slight abundance of one component can lead to vacancies of another component, for example. This issue is particularly relevant in complex oxides in which standard preparations often yield nonstoichiometric crystals, such as the perovskite crystal, strontium iron niobium oxide.

Chemical impurities may be present substitutionally (taking the place of a normal atom) or interstitially. Furthermore, certain chemical impurities may induce the existence of vacancies or interstitials. For example, if a second-row impurity ion such as Ca^{2+} substitutes for Na^{+} (a first-row ion) in an insulating ionic solid such as NaCl, it is likely that there will be an Na^{+} vacancy nearby, in order for overall charge conservation to occur.

Particularly in semiconductors, but also in insulators, specific point defects may exist in different charge states. The charge states will clearly influence the electrical and optical properties of the solid. Charge states may be changed by varying the chemical potential (Fermi energy) of the solid, by various photoionization processes, or by charge injection.

While the simplest notion of atoms in solids is that of spheres of different sizes, an understanding of point defects (especially in nonmetals), in fact, relies heavily on chemical bonding ideas. For example, interstitials in semiconductors almost never simply sit at empty sites, interacting weakly with their neighbors. Rather, the interstitial atom and its neighbors reorganize their bonding and their positions, often forming the so-called split interstitial whose structure is governed by chemical bonding concepts.

Control of Point Defects

The control of point defects in solids is a central role of materials physics, and indeed, it is often the ultimate source of success or failure in a particular application.

The first step in point defect control is in the crystal growth process. It is at this point that desired impurities may be introduced and undesired impurities may be avoided. The outcome depends on the purity of the starting material, the detailed growth process, and the growth environment.

Even an elemental solid such as silicon, prepared with care, will contain intrinsic defects as a consequence of thermal equilibrium. Although a perfect crystal structure represents the most stable arrangement energetically, in thermal equilibrium a number of defects will exist as determined by Boltzmann statistics:

$$N_{\mathrm{D}} = BN \exp(-AE/kT) \qquad [1]$$

Here N_{D} is the number of defects, N the number of atomic sites, and A and B are dimensionless factors of order 1. E is the energy required to create the defect, k is Boltzmann's constant, and T is the kelvin temperature. If, for example, the creation energy is 0.35 eV, the room-temperature equilibrium fraction of defects can be $\sim 10^{-6}$, large enough to have noticeable effects. Commonly, the creation energy is larger, but even then, exposure of the solid to high temperatures during growth or processing can lead to a significant fraction of intrinsic defects, which, if cooling occurs faster than thermal equilibrium can be established, and can be "frozen in" at concentrations appropriate to the higher temperature.

In some cases, additional thermal processes may reduce the defect concentration through annealing. On the other hand, thermal treatments that are necessary for the construction of complex semiconductor

devices may yield unwanted outcomes relative to the point defects in portions of the device, such as thermally stimulated diffusion of impurity atoms from one region of the device to another.

In thin films, as those that occur for example in semiconductor devices, applied electric fields may also lead to changes in the location or configuration of point defects. Electric field-induced diffusion of atoms or ions may occur. Or, trapping of electrons or holes may lead to changes in defect configurations.

Point defects may be created or modified by external ionizing irradiation, either by photons or by high-energy particles. A number of detailed processes may take place, depending on the solid and the properties of the radiation. Heavy particles (such as neutrons or α-particles) of sufficient energy can displace atoms in the crystal by momentum- and-energy-conserving "billiard-ball" collisions. For lighter particles such as electrons or γ-rays, the requirements of both energy and momentum conservation require very high energies for the lighter particle.

In some nonmetallic crystals, irradiation can introduce point defects by simply ionizing atoms or ions, in which case the energy required is the bandgap, of the order of several electronvolts. In such cases, the defect creation occurs via a chemical reaction between the ionized atom, or the atom in an excited state, and other atoms in its vicinity.

Point Defect Diffusion

Point defect diffusion is a process in which a defect moves a macroscopic distance as the result of a series of short-range, statistical motions. Many specific mechanisms of defect diffusion may exist, depending on the defect or defects, the host, and other conditions such as temperature and external fields. The interaction of point and extended defects may lead to large enhancements of point defect diffusion.

An impurity at an interstitial site is considered, for example. Thermally activated motion over a potential barrier will be required for the impurity to move to an equivalent nearest-neighbor site. If there is just one impurity and there are no additional fields, its net motion will be a "random walk"; its long-range displacement will be hampered by its tendency to jump back toward, in addition to away from, the initial site, and no particular direction of motion will be favored. Calculations using statistical methods show that the net distance traveled by the impurity will vary as the square root of time.

If, however, an external field yields a net force in one direction, then the probability of jumps in that direction will be greater than the probability of jumps in the reverse direction; so, there will be a net motion of the impurity in the direction of the net force.

Furthermore, diffusion can take place in the absence of external fields if there is a concentration gradient of impurity atoms. An impurity will tend to jump from a region of higher concentration to a region of lower concentration, since in a direction of higher concentration the neighboring interstitial sites are more likely to be already occupied. This effect is described by Fick's first law:

$$J(x) = -D\frac{\partial C(x)}{\partial x} \qquad [2]$$

Here $J(x)$ is the impurity flux in the x-direction, $C(x)$ is the impurity concentration, and D is the diffusion coefficient. The diffusion coefficient D contains all the material properties, and it is also a function of temperature. Since the diffusion process involves thermal motion over a barrier, the diffusion coefficient D often varies with temperature, T, in a way similar to eqn [1]:

$$D = D_0 \exp(-\Delta E/kT) \qquad [3]$$

Here, ΔE is the height of the potential barrier and k is Boltzmann's constant.

Point Defects in Normal Metals and Superconductors

Because metals contain nearly free electrons, charge neutrality in the region of a point defect will automatically be satisfied. Hence, issues such as charge-compensating defects or charge-state effects, which are of major importance in nonmetals, are not relevant in metals. Nevertheless, point defects are still of importance in affecting the properties of metals. Two examples are briefly mentioned here: electrical properties and magnetic properties.

Even though charge neutrality in the region of point defects in metals exists, point defects still scatter conduction electrons. This is particularly important at low temperatures, where phonon scattering in the perfect crystal is reduced while impurity scattering remains.

Magnetic impurities can be important in both magnetic and nonmagnetic metals. In the latter case, the magnetic moment of the impurity is partially, but not completely, screened by the conduction electrons. In the former case, interactions between the magnetic atoms are important and will help determine the magnetic properties of the defective material.

In most cases, point defects have negative effects on superconductivity, since superconductivity involves coherent effects between charged particles. In high-T_c superconductors, however, impurities can

increase the number of electrons or holes. Furthermore, impurities have been of some use in studying the mechanisms for superconductivity.

Point Defects in Semiconductors

Semiconductors are useful in electronics because it is possible to add impurities in a way that leads to specified concentrations of electrons or holes that, in turn, can carry current. This is called "doping." In the simplest elemental semiconductors, such as silicon (a group IVb element), each atom has four valence electrons and forms bonds to four neighbors. If a phosphorus (a group Vb element) atom replaces a silicon atom, four of the phosphorus valence electrons form bonds, and the fifth is only weakly bound to the phosphorous atom. This weakly bound electron can be thermally ionized at normal temperatures, yielding a conduction electron. The result is *n*-type conduction.

If, on the other hand, a boron (a group IIIb element) atom replaces a silicon atom, a valence electron is needed to form bonds with the four neighbors, leaving a missing electron – a hole – in the valence band, weakly bound to the boron. This hole can be thermally ionized at normal temperatures, yielding a positively charged conducting "particle" – a "hole." The result is *p*-type conduction.

This simple picture does not always hold. For example, nitrogen is also a group Vb element, but a nitrogen atom substituting for a silicon atom does not yield a weakly bound electron. Rather, the nitrogen relaxes from the center of the substitutional site and forms bonds with the neighboring silicons in which all the electrons are involved and none are left to be weakly bound. This example shows that, in general, localized ("deep") and delocalized ("shallow") defect states may coexist, and which one is more stable depends on details.

In some cases, doping in compound semiconductors (such as GaAs) follows similar logic. Silicon in GaAs is amphoteric; if Si replaces Ga the system is *n*-type, whereas if it replaces As the system is *p*-type. In principle, *n*- or *p*-type conductivity may also be associated with intrinsic defects. For sometime, it was thought that *n*-type GaN resulted from nitrogen vacancies, but eventually it was realized that the origin of the conduction electrons – at least in most cases – comes from Si replacing the Ga.

It has been found difficult to achieve *p*-type doping in a number of compound semiconductors, and success has required some "tricks." For example, Mg on an N site will yield *p*-type GaN, but it is difficult to prepare GaN:Mg. The trick here is to include hydrogen in the processing; it is possible to prepare GaN with an Mg:H complex on the N site. Subsequent irradiation or heat treatment then removes the hydrogen, leaving behind the desired structure.

Along with deliberately introduced dopants, many other kinds of defects may occur in semiconductors. Some are relatively benign; for example, substitutional C in Si does not trap electrons or holes; so it may occur in a relatively large concentration and have little effect on the electronic properties of Si. Others may themselves trap electrons or holes, or they may combine with dopants in a way that "passivates" the electrical activity of the defect. For example, substitutional B in Si may combine with an H atom, resulting in a B–H pair in which the H is bond-centered (i.e., located within a B–Si bond) in the positive or the neutral charge state. In this configuration, the H supplies the fourth bonding electron, so there is no weakly bound hole and *p*-type conduction is suppressed.

In general, hydrogen is a common and significant impurity in semiconductors. Semiconductor processing involves compounds that contain hydrogen, so it is commonly present. Hydrogen is also amphoteric; it may passivate both donors and acceptors. In fact, hydrogen is, in most cases, a "negative-U" defect. This means that it will never occur as a neutral atom in thermal equilibrium; for any pair of hydrogen atoms (in general, well-separated), the sum of the energies of H^+ and H^- is lower than the energy of two neutral H's. If in fact the H's are mobile, it is likely that interstitial H_2 molecules will form, whose energy is still lower and thus represents the most stable state.

In general, defect motion is of great importance in semiconductor technology, and it has been widely studied. As noted earlier, atomic diffusion may occur during high-temperature processing, and in semiconductors this may yield deleterious effects on resulting devices. Processes that involve impurities coupled with intrinsic defects are common. For example, a substitutional Au atom in Si can switch places with an interstitial Si, forming an interstitial Au, which in turn has a very high diffusivity.

An important phenomenon in optical and some electrical applications is recombination-enhanced diffusion. In some situations, the energy gained from the recombination of an electron and a hole is sufficient to allow atomic motion over potential barriers. For example, optical excitation in ZnSe can induce a diffusion jump of an interstitial Zn atom.

Diffusion of impurity atoms in semiconductors may be enhanced by electric fields. For example, N and B may be induced to diffuse in SiC under high-temperature conditions in the presence of a field associated with a built-in *p*–*n* junction in the material.

Another important issue is the effect of charge states on diffusion. Because the structure and energetics of

defects may be strongly dependent on charge states, so may be the diffusion coefficient. Hence, atomic diffusion phenomena observed in *n*-type and *p*-type semiconductors may be quite different.

Point Defects in Insulators

Insulating solids share with semiconductors the existence of a bandgap; a region of forbidden energy between filled and empty electronic states. They also share some physics, such as the possible existence of defect states of different charge.

A fundamental distinction between insulators and semiconductors is that defect states in insulators are almost never sufficiently shallow in energy to provide charge carriers at normal temperatures. Hence, defects in insulators are almost always "deep."

There are many distinctions between insulators and semiconductors concerning important properties and applications. For example, the bandgaps of insulators span the energy range of visible photons, so defects in insulators may have important optical properties. Indeed, in the early German literature on defects in alkali halides (e.g., NaCl), such defects were called "farbzentren" (or "color centers"). Many precious gems have significant colors because of naturally occurring point defects in an insulating host. One result of these properties was the invention of the ruby laser, and insulator hosts have proven useful in connection with other optical devices and lasers as well.

Point defects in nonmetallic crystals will, if they have energy levels within the gap between the valence and conduction bands, lead to characteristic optical absorptions that involve the excitation of electrons or holes from a ground to an excited state. In many cases, there is also luminescence as the excited defect returns to the ground state.

The general theory of optical absorption, as well as some aspects of luminescence, is similar for deep semiconductor and insulator defects. In both cases, the electron–phonon interaction is involved. The defect in its ground electronic state has neighboring atoms that vibrate about sites whose equilibrium positions are determined, in part, by the defect wave function. When the defect is electronically excited, the neighboring atoms, responding to a new defect wave function, relax to new equilibrium positions. This process may be represented by a configuration coordinate diagram (**Figure 1**) of energy, E, versus an average position of the neighboring atoms, R. This idealized case consists of two displaced parabolas representing the harmonic energy of the vibrating atoms in ground (m) and excited (k) states. η and γ denote the quantized vibrational levels associated with the parabolas. An optical transition, in general, involves a change of both electronic and vibrational quantum numbers.

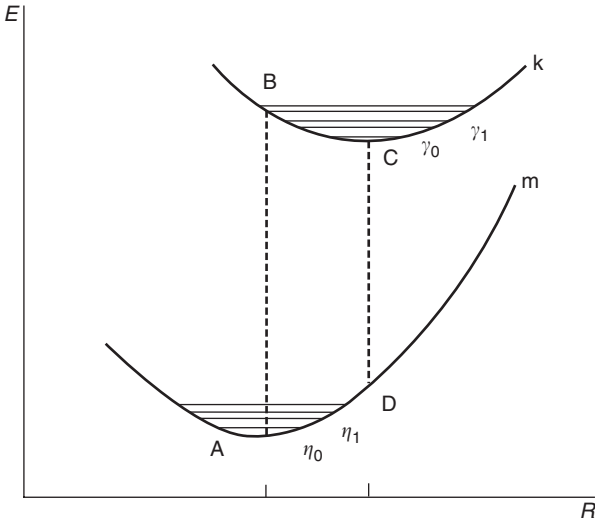


Figure 1 Configuration coordinate diagram for a model system.

In the weak-coupling limit, where the excited-state parabola is nearly above that of the ground state, the strongest transitions involve no change in the vibrational quantum number. In this case, at low temperature, if only the zero-point vibrational level of the ground state is occupied, the strongest transition will be a sharp "zero-phonon" transition from A to C to the zero-point level of the excited state.

In the strong-coupling limit shown in **Figure 1**, with both horizontal and vertical displacements of the parabolas, the zero- and few-phonon transitions are weak because the vibrational wave functions of the low-lying displaced oscillators do not overlap strongly, and the optical absorption is characterized by a broad multiphonon transition A to B which is approximately vertical in the configuration coordinate R.

In the absence of nonradiative processes, the excited defect will luminesce, emitting a photon as it returns to the ground state. In general, the radiative lifetime is at least several orders of magnitude longer than the vibrational relaxation rates, so the excited defect will relax to its lowest vibrational state, for example, at low temperature it will relax to the zero-point state C before the defect radiates. In the weak-coupling limit, the zero-phonon line C to A will be identical to that for absorption, while few-phonon lines will lie at lower photon energy (in contrast to absorption, where they lie at higher energies). In the strong-coupling limit, the broad multiphonon emission C to D will be "Stokes-shifted" to lower photon energy from that for absorption, as shown in **Figure 1**.

In the strong-coupling limit, then, the electron–phonon coupling has substantial effects. For example, for an electron trapped at a Cl^- vacancy in KCl (an F

center), the optical absorption peak at low temperature is at 2.3 eV with a width at half-maximum of 0.16 eV, while the luminescence peak is at 1.2 eV with a width of 0.26 eV.

Summary

Point defects in solids are of great importance and can have a wide variety of properties and applications. Other articles cited in the "See also" section delve more deeply into these topics. The books listed in the "Further reading" section treat the properties and effects of point defects in considerable depth and breadth. Internet searches under "crystal point defect" will provide up-to-date connections, as well.

See also: Insulators, Impurity and Defect States in; Insulators, Optical Properties of; Jahn–Teller Effect; Mechanical Properties: Strengthening Mechanisms in Metals; Metals and Alloys, Impurity and Defect States in; Semiconductors, Impurity and Defect States in.

PACS: 61.72. – y; 71.55. – i; 78.40. – q; 78.55. – m

Further Reading

Ashcroft N and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.

Borg RJ and Dienes GJ (1988) *An Introduction to Solid-State Diffusion*. San Diego: Academic Press.

Flynn CP (1972) *Point Defects and Diffusion*. Oxford: Clarendon Press.

Fowler WB (1968) *Physics of Color Centers*. New York: Academic Press.

Fowler WB and Itoh N (1990) *Atomic Processes Induced by Electronic Excitation in Non-Metallic Solids*. Singapore: World Scientific.

Hayes W and Stoneham AM (1985) *Defects and Defect Processes in Nonmetallic Solids*. New York: Wiley.

Jackson KA and Schröter W (eds.) (2000) *Handbook of Semiconductor Technology*, vol. 1. Weinheim: Wiley-VCH.

Kaxiras E (2003) *Atomic and Electronic Structure of Solids*. Cambridge: Cambridge University Press.

Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.

Levinshtein ME and Shur M (1997) *Semiconductor Technology: Processing and Novel Fabrication Techniques*. New York: Wiley.

Murch GE and Nowick AS (eds.) (1984) *Diffusion in Crystalline Solids*. New York: Academic Press.

Stoneham AM (1975) *Theory of Defects in Solids*. Oxford: Oxford University Press.

Yu PY and Cardona M (1996) *Fundamentals of Semiconductors*. Berlin: Springer.

Glossary

Antisite defect	an atom on the "wrong" site in a crystal that contains more than one type of chemical constituent.
Atomic diffusion	the motion of an atom over a macroscopic distance as the result of a series of individual microscopic motions which are statistical in nature.
Chemical impurity	an atom, ion, or (small) molecule that is chemically different from the perfect-crystal constituents.
Color centers	point defects in insulators that generate optical activity in the visible.
Configuration coordinate diagram	plot of energy versus atomic position for a molecule or defect.
Deep defects	defects in semiconductors or insulators in which the electron or hole is strongly bound and is not likely to be unbound at room temperature.
Doping	the introduction of chemical impurities into a crystal during the growth process.
Extended defect	an imperfection that extends through a large portion of a crystal.
F center	an electron trapped at a halogen-ion vacancy in an alkali halide crystal.
Fermi energy	the chemical potential of a solid.
Frenkel defect	a vacancy-interstitial pair.
Hydrogen passivation	the change of the nature of the electrical properties of defects by the incorporation of hydrogen into the defect.
Interstitial	an atom or ion at a position other than a perfect-crystal site.
n-Type	a semiconductor with an excess of donors.
Negative-U	a situation in which the "effective" Coulomb interaction between electrons or holes is negative.
p-Type	a semiconductor with an excess of acceptors.
Perfect crystal	a periodically repeated array of identical building blocks such as atoms or arrays of atoms.
Point defect	an imperfection in a perfect crystal that is of the order of atomic size.
Schottky defect	an isolated vacancy in an otherwise perfect crystal.
Shallow donors and acceptors	defects in semiconductors in which the electron (donor) or hole (acceptor) is weakly bound and is likely to be unbound at room temperature.
Stokes shift	the energy difference between absorption and emission bands.

Substitutional defect	a defect that occurs on a perfect-crystal atom site.
Vacancy	the absence of an atom or ion from its perfect-crystal site.
Zero-phonon transition	an optical transition from the zero-point vibrational level of one electronic state to the zero-point level of another state.

Point Groups

H Klapper, Mineralogisch-Petrologisches Institut, Universität Bonn, Bonn, Germany
Th Hahn, Institut für Kristallographie, RWTH Aachen, Germany

Brief Review of Groups

Definition: A group $\mathcal{G}$ is a set of N elements g_i with a composition law ("multiplication" $g_i \circ g_j$) of any two elements g_i and $g_j \in \mathcal{G}$ obeying the following four group axioms:

1. The "product" of any two group elements g_i and $g_j \in \mathcal{G}$ is again an element of the group: $(g_i \circ g_j) = g_k \in \mathcal{G}$ (axiom of closure).
2. There exists always a neutral element (identity) e, such that $(e \circ g_i) = (g_i \circ e) = g_i$ for any $g_i \in \mathcal{G}$.
3. For every $g_i \in \mathcal{G}$, there exists an inverse element $g_i^{-1} \in \mathcal{G}$ such that $(g_i \circ g_i^{-1}) = (g_i^{-1} \circ g_i) = e$.
4. For any three elements g_i, g_j, $g_k \in \mathcal{G}$, the associative law holds: $(g_i \circ g_j) \circ g_k = g_i \circ (g_j \circ g_k)$.

The finite number N of (different) elements of a group is the order of the group (finite groups). If the group contains infinitely many (different) elements, the group is said to be of infinite order (infinite group).

A finite group $\mathcal{G}$ is called cyclic, if it can be generated by one element g: $g^1 = g, g^2, g^3, \ldots, g^N = e$ and is of order N.

If the relation $(g_i \circ g_k) = (g_k \circ g_i)$ holds for all pairs g_i, $g_k \in \mathcal{G}$, the group is called commutative or Abelian. All groups containing only binary elements g_i (i.e., $g_i \circ g_i = g_i^2 = e$, or equivalently, $g_i = g_i^{-1}$) and all cyclic groups are commutative.

A subset $\mathcal{H}$ of a group $\mathcal{G}$, $\mathcal{H} \subset \mathcal{G}$, is called a subgroup of $\mathcal{G}$, if $\mathcal{H}$ obeys the group axioms. In this case, $\mathcal{G}$ is a supergroup of $\mathcal{H}$. For finite groups, the ratio of the orders of group and subgroup, $N(\mathcal{G})$ and $N(\mathcal{H})$, is the index $[i]$ of $\mathcal{H}$ in $\mathcal{G}$: $[i] = N(\mathcal{G})/N(\mathcal{H}) = |\mathcal{G}|/|\mathcal{H}|$. According to the theorem of Lagrange, for finite groups the index $[i]$ is always an integer.

An illustrative representation of a finite group $\mathcal{G}$ of order N is the multiplication table or group table of $\mathcal{G} = \{e, g_2, g_3, g_4, \ldots, g_N\}$. It is an $(N \times N)$ square array of all products according to the abstract scheme of **Table 1**. Note that in each of the N rows and N columns of the table, each group element appears exactly once.

Groups of the same order and with the same structure of the group table, that is, with strict correspondence of their ordered elements and their products (irrespective of the names and symbols of the group elements), are isomorphic. If the group table is symmetric with respect to the main diagonal, the group is commutative.

In crystallographic groups, the group elements are symmetry operations (motions, isometries). The "multiplication" is the successive application of two operations. As an example, the group tables of the two crystallographic point groups 4 (fourfold 90° rotations, tetragonal) of order 4 (**Table 2**) and $2_x/m_x\, 2_y/m_y\, 2_z/m_z$ (short *mmm*, orthorhombic) of order 8 (**Table 3**) are discussed. **Table 3** illustrates the

Table 1 Multiplication table of a group $\mathcal{G} = \{e, g_2, g_3, g_4, \ldots, g_N\}$ of order N

	e	g_2	g_3	...	g_N
e	e	g_2	g_3	...	g_N
g_2	g_2	g_2^2	$g_2 g_3$	...	$g_2 g_N$
g_3	g_3	$g_3 g_2$	g_3^2	...	$g_3 g_N$
⋮	⋮	⋮	⋮	⋮	⋮
g_N	g_N	$g_N g_2$	$g_N g_3$	...	g_N^2

Table 2 Group table (4×4) of the tetragonal crystallographic point group 4: successive 90° rotations forming a cyclic group of order 4, i.e., $4^4 = 1$ (this group is commutative)

	1	4^1	$4^2 = 2$	$4^3 = 4^{-1}$
1	1	4^1	2	4^3
4^1	4^1	2	4^3	1
$4^2 = 2$	2	4^3	1	4^1
$4^3 = 4^{-1}$	4^3	1	4^1	2

There is one subgroup of order $N = 2$ and index $[i] = 2$ with elements {1,2} (180° rotation).

Table 3 Group table (8×8) of the orthorhombic crystallographic point group $2_x/m_x\,2_y/m_y\,2_z/m_z$ (short *mmm*) of order $N = 8$

	1	2_y	m_y	$\bar{1}$	⋮	2_x	m_x	2_z	m_z
1	1	2_y	m_y	$\bar{1}$	⋮	2_x	m_x	2_z	m_z
2_y	2_y	1	$\bar{1}$	m_y	⋮	2_z	m_z	2_x	m_x
m_y	m_y	$\bar{1}$	1	2_y	⋮	m_z	2_z	m_x	2_x
$\bar{1}$	$\bar{1}$	m_y	2_y	1	⋮	m_x	2_x	m_z	2_z
...	...	...	...	...					
2_x	2_x	2_z	m_z	m_x		1	$\bar{1}$	2_y	m_y
m_x	m_x	m_z	2_z	2_x		$\bar{1}$	1	m_y	2_y
2_z	2_z	2_x	m_x	m_z		2_y	m_y	1	$\bar{1}$
m_z	m_z	m_x	2_x	2_z		m_y	2_y	$\bar{1}$	1

The top left (4×4) array (separated by dotted lines) is the group table of the subgroup $2_y/m_y$ of order $N = 4$ and index $[i] = 2$. The group *mmm* can be obtained from the subgroup $2_y/m_y$ by "extension" of its (4×4) group table by any of the additional operations 2_x, m_x, 2_z, or m_z (here by 2_x as the fifth element), each leading to the same supergroup *mmm* of index $[i] = 2$. Both, group and supergroup, are commutative, because all group elements g_i are binary.

various group–subgroup relations within this group: there are seven subgroups of order $N = 4$ and index $[i] = 2$. One of these subgroups, $2_y/m_y$, is displayed in the upper left (4×4) array in **Table 3**. The full (8×8) group table of $2_x/m_x\,2_y/m_y\,2_z/m_z$ is obtained from this $2_y/m_y$ subtable by adding any of the operations 2_x, m_x, 2_z, or m_z. From **Table 3**, it is easily derived that the seven subgroups of order 4 are $2_x/m_x$, $2_y/m_y$, $2_z/m_z$, $2_x2_y2_z$, $2_xm_ym_z$, $m_x2_ym_z$, and $m_xm_y2_z$, which are all isomorphic. There are seven further isomorphic subgroups of order 2 and index 4.

Introduction to Point Groups

Three kinds of three-dimensional groups are fundamental to the study of crystals: point groups, translation groups, and space groups. This article deals with the first type of groups, the point groups. These groups describe and govern the symmetry of finite (nonperiodic, i.e., not containing translations) objects, such as molecules, crystals, flowers, and objects of art, as well as the local symmetries ("site symmetries") of points in a periodic pattern (ornament, wallpaper, crystal structure).

In a more subtle fashion, one can distinguish two kinds of point-group symmetries:

1. Symmetry in "vector space": symmetry of a set of vectors, for example, of the face normals of a crystal ("crystal form"), of the translation vectors of a crystal lattice, and of the interatomic vectors of a crystal structure (Patterson vectors). These symmetries are visualized best as vector bushels originating from a common origin.
2. Symmetry in "point space": symmetry of any finite polyhedron, of a molecule, or of an atomic group; symmetry around an arbitrary point in an infinite crystal structure ("site symmetry") or around a lattice node in an infinite point lattice ("lattice point group").

It should be noted that in many textbooks on crystals, two terms are in general use: "point group" as the symmetry of a particular crystal and "crystal class" as the set of all crystals with the same point-group symmetry. Both receive the same symbol. It is of interest that the 32 three-dimensional crystallographic point groups were derived already around 1830.

General and Crystallographic Point Groups

Point groups can be subdivided into crystallographic and noncrystallographic point groups. Both together are called general point groups.

General Point Groups

A "general point group" is a group of rigid motions (symmetry operations, isometries) which all leave at least one point of space invariant.

This definition leads to infinitely many mono-axial point groups, because it admits infinitely many rotation groups with rotation angles $\varphi = 360°/n$ for all integer values of n. All these rotation groups are of finite order $N = n$. Furthermore, there are some point groups of order $N = \infty$, because they contain rotations (or rotoinversions) of order infinity, that is, rotations with all possible (including arbitrarily small) rotation angles. These rotation groups of infinite order are symbolized as ∞ and $\overline{\infty} = \infty/m$ in Hermann–Mauguin notation and as C_∞ and $C_{\infty h} = C_{\infty i} = S_\infty$ in Schoenflies notation. They appear as symmetry elements of cones, cylinders, circles, and spheres. For details see the "Further reading" section, especially Vol. A of the *International Tables*.

The condition "leave at least one point of space invariant" means that all symmetry elements of the group (axes, planes, and point) intersect at this one point, thus excluding any translations among the symmetry operations. Invariance of more than one point is of course possible, but always leads to infinitely many invariant points. Examples are all "mono-axial" groups n (one rotation axis only), which have an invariant line, and the reflection group m, which has an invariant plane. The intersection of several

rotation axes in one point or the presence of a center of symmetry $\bar{1}$, leads to only one invariant point which represents the center of the object.

Crystallographic Point Groups

A "crystallographic point group" is a group of rigid motions which all leave at least one point of space invariant and, in addition, map a point lattice onto itself ("crystallographic restriction").

As described elsewhere in this encyclopedia, the crystallographic restriction (lattice condition) restricts the permissible rotation angles to multiples of 180°, 120°, 90°, and 60° (two-, three-, four-, and sixfold axes). This reduces the infinitely many point groups to 10 for two dimensions and to 32 for three dimensions. The highest group orders are $N = 12$ for two and $N = 48$ for three dimensions.

Among the 32 three-dimensional crystallographic point groups, there are 11 centrosymmetric groups (so-called Laue groups), which are important in X-ray diffraction, and seven lattice point symmetries (holohedries), defining the seven crystal systems. It should be noted that the 32 three-dimensional crystallographic point groups discussed below can be attributed to 18 isomorphism classes (abstract point groups). A classification of the crystallographic point groups with respect to physical properties is given in **Table 5**.

Among the three-dimensional noncrystallographic point groups there are the two icosahedral point groups, which contain 6 fivefold, 10 threefold, and 15 twofold rotation axes. They have recently received particular attention due to their occurrence in quasicrystals and fullerenes.

In modern crystallography, point groups in four, five, and six dimensions have been derived and are used for the treatment of incommensurate and quasicrystals. There are 271 point groups in four, 955 in five, and 7104 in six dimensions, with highest orders $N = 1152$, 3840, and 51840, respectively.

Symbols of the 32 Crystallographic Point Groups

The two systems of symbols for the crystallographic point groups are briefly reviewed here:

1. Schoenflies symbols, first presented in 1891 in Schoenflies' classical book, are based on capital letters for rotation groups C, D, S, T, O with subscripts s, h, v, d for mirror planes in various orientations, and subscript i for an inversion center. The noncrystallographic icosahedral and sphere groups are denoted by letters I_h and K_h (for German "Kugelgruppe").
2. The Hermann–Mauguin or international symbols are based on the "blickrichtungen" (symmetry directions) of the various lattices (holohedries), with specific numbers 2, 3, 4, 6 for symmetry axes, m for a mirror plane, and $\bar{1}$ for an inversion center. The Hermann–Mauguin symbols for the icosahedral and the sphere groups are $2/m\bar{3}\bar{5}$ (often written as $\bar{5}\bar{3}2/m$) and $2/m\overline{\infty}$ (often written as $\infty\,\infty m$).

Symbols and groups orders of the 32 three-dimensional crystallographic point groups, both in Schoenflies and Hermann–Mauguin notation, are illustrated in **Figure 1**.

Description of Point Groups

A formal description of point groups is already given by their group tables, in the section "Brief review of groups" and in **Tables 1–3** of this article.

Algebraically, the point groups are usually represented by matrix groups, that is, by the set of (3×3) matrices $\boldsymbol{W}$ of their point group operations (motions, isometries) with determinants $+1$ and -1 or, in a more condensed form, by either the N (contravariant) coordinates x, y, z, of the symmetrically equivalent points, or the N (covariant) Miller indices $\{hkl\}$ of the symmetrically equivalent faces of the "general face form" (crystal form). N is the order of the point group.

Example: point group $2_y/m_y$ (cf. subgroup in **Table 3**), consisting *of* $N = 4$ group elements 1, 2_y, m_y, $\bar{1}$:

1. coordinates of equivalent general points: x, y, z (1), $-x, y, -z$ (2_y), $x, -y, z$ (m_y), $-x, -y, -z$ $(\bar{1})$;
2. Miller indices of equivalent general faces: (hkl) (1), $((\bar{h}, k, \bar{l})$ (2_y), $(h\bar{k}l)$ (m_y), $(\bar{h}\bar{k}\bar{l})(\bar{1})$.

A complete algebraic representation of the 10 two- and the 32 three-dimensional crystallographic point groups, as well as of the two icosahedral groups is given in Vol. A of the *International Tables* (cf. "Further reading" section).

Geometrically, a point group is normally represented by means of the stereographic projection of its "framework of symmetry," that is, the spatial assemblage of its symmetry elements (rotation axes, mirror planes, and center of symmetry).

Figure 1 shows the stereographic projections of the symmetry frameworks of the 32 three-dimensional crystallographic point groups, as well as the two icosahedral groups. The graphical symbols of rotation

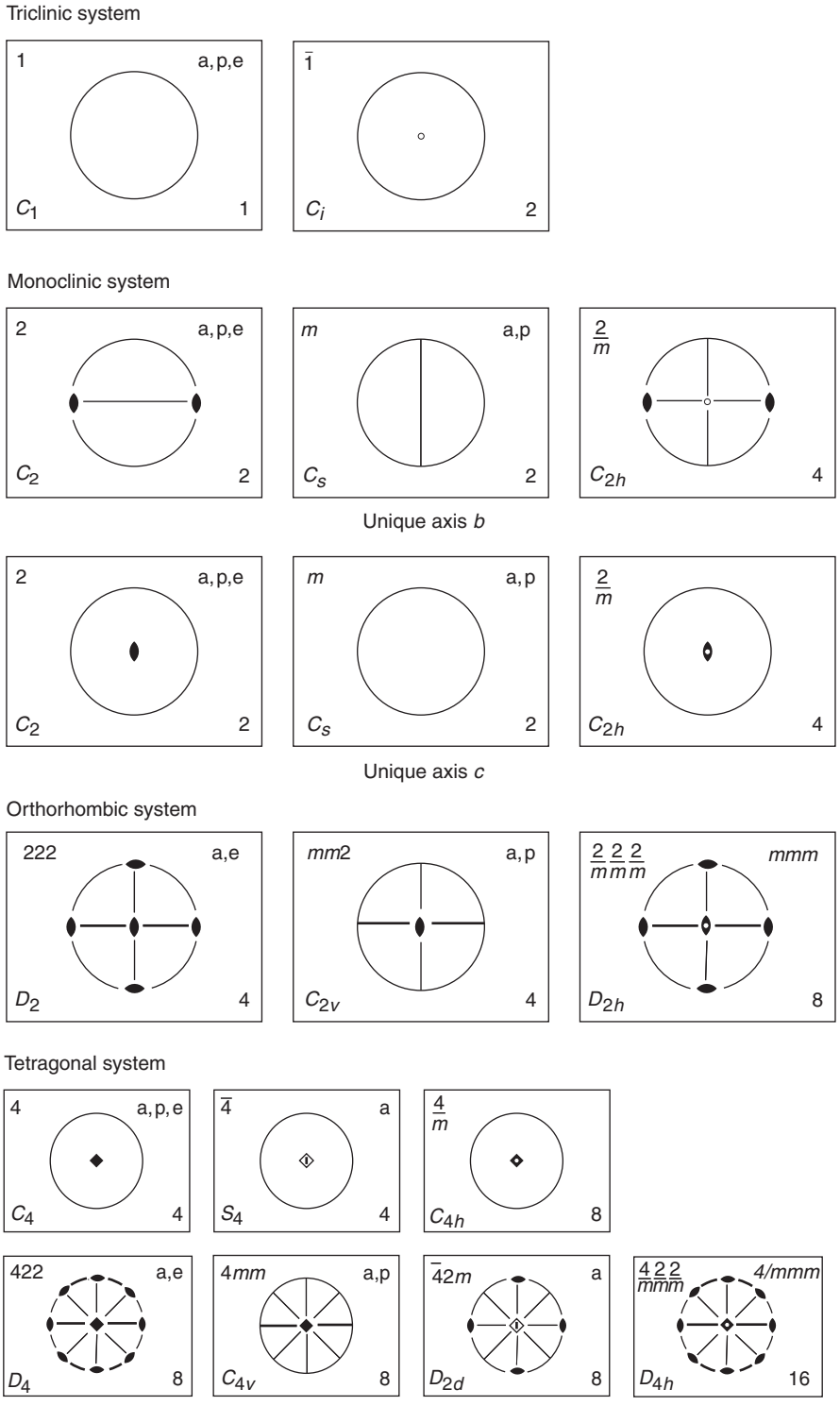


Figure 1 Stereographic projections of the symmetry elements of the 32 crystallographic point groups and the two icosahedral groups. Heavy solid lines or circles indicate mirror planes, thin lines are used for the projection circle and for auxiliary lines. Around each projection, the full (top left) and short (top right, if different) Hermann–Mauguin symbols, the Schoenflies symbol (bottom left) and the group order *N* (bottom right) are given. In the top right position the letters "a", "p", and "e" indicate that the point group is acentric (noncentrosymmetric), polar (pyroelectric) and enantiomorphic, respectively (cf. **Table 5**). No letter indicates a centrosymmetric point group. (Reproduced with permission from Hahn Th and Klapper H (2002) *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*, 5th edn. Dordrecht: Kluwer Academic; © International Union of Crystallography.)

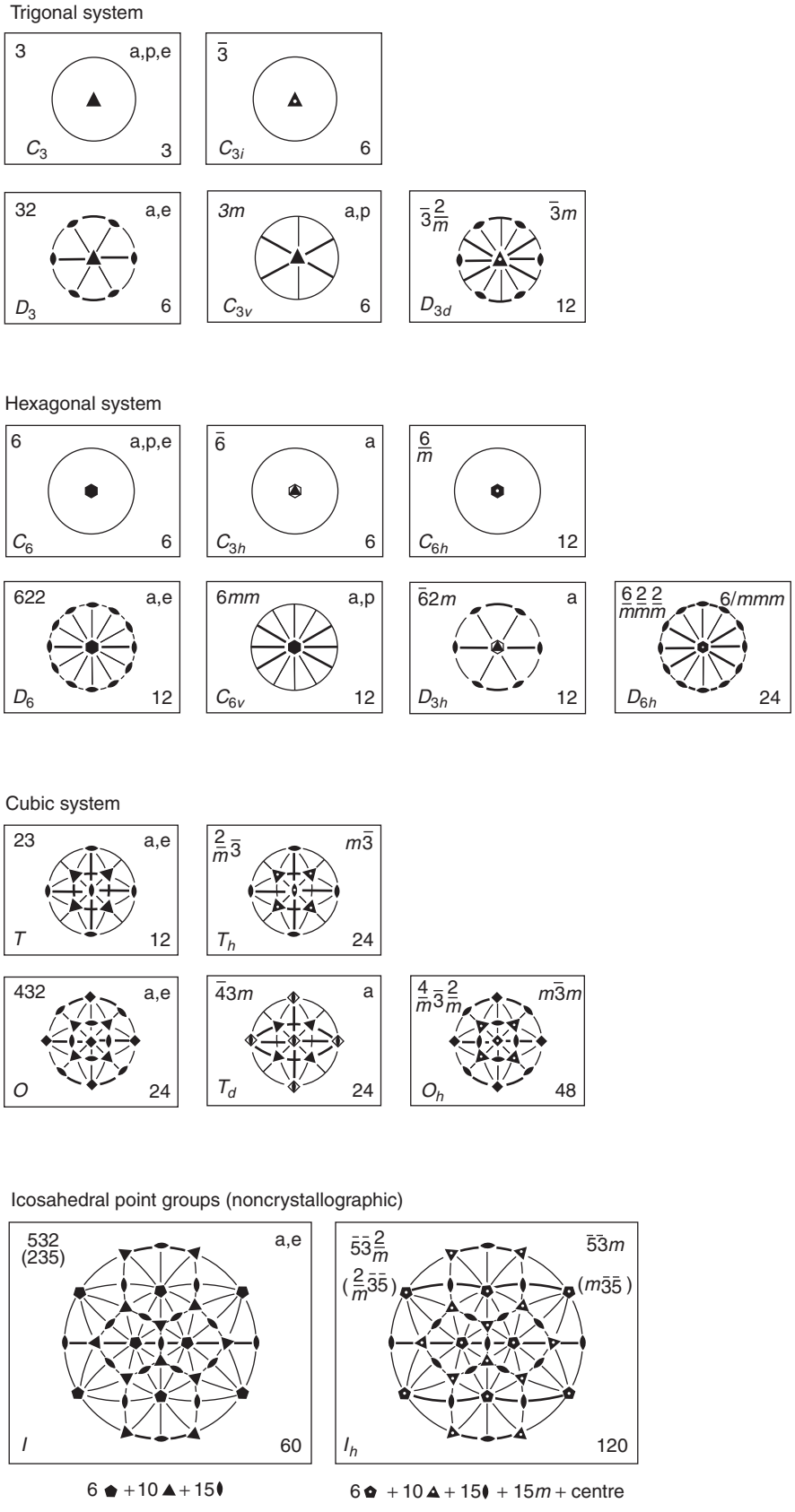


Figure 1 Continued.

and rotoinversion axes, mirror planes, and inversion centers are explained elsewhere in this encyclopedia.

The most frequent use of these stereographic projections in mineralogy and crystallography is the characterization of the crystal morphology by the normals of the crystal faces (face poles), leading to the face forms (crystal forms), that is, sets of symmetrically equivalent faces. Each point group (order N) has a "general" face form $\{hkl\}$ with N faces, each one of "face symmetry" 1, and up to six "special" face forms with N/q faces with "face symmetry" >1 of order q, for example, $\{hk0\}$, $\{1\,1\,1\}$, $\{1\,0\,0\}$. Example: the face form cube $\{1\,0\,0\}$ has six faces with "face symmetry" $4\,mm$ (order 8) in point group $4/m\,\bar{3}2/m$ (order $6 \times 8 = 48$).

Complete listings of all general and special face forms of the crystallographic point groups and the icosahedral groups are presented in Vol. A of the *International Tables* (see the "Further reading" section).

Subgroups of the Crystallographic Point Groups

Group–subgroup relations, as defined in the first section, are of particular importance for physical phenomena such as phase transitions (symmetry breaking in the Landau theory), twinning, and domain structures. In solid-state chemistry, they are widely used for the classification of structural relations in families of related compounds (structure types).

Figure 2 displays the sub- and supergroups of the two- and three-dimensional crystallographic point groups in the form of two "family trees." Each tree has two "summits": for two dimensions (**Figure 2a**) the hexagonal holohedry $6mm$ (order $N = 12$) and the tetragonal holohedry $4mm$ ($N = 8$), for three dimensions (**Figure 2b**) the cubic holohedry $4/m\bar{3}\,2/m$ ($N = 48$) and the hexagonal holohedry $6/m\ 2/m\ 2/m$ ($N = 24$). Their subgroups "mix" on the levels $N = 4$ (**Figure 2a**) and $N = 12$ (**Figure 2b**), where common subgroups of both branches appear first. Reading the diagrams downward, solid and dashed lines connect a point group $\mathcal{G}$ with all its maximal subgroups $\mathcal{H}$. The subgroup index $[i]$ is given by the ratio of the group orders of $\mathcal{G}$ and $\mathcal{H}$: $[i] = N(\mathcal{G})/N(\mathcal{H})$. If the diagrams are read upward, the lines connect a point group $\mathcal{H}$ with all its minimal supergroups $\mathcal{G}$, with supergroup index $[i] = N(\mathcal{G})/N(\mathcal{H})$. The index $[i]$ is always an integer, according to the theorem of Lagrange.

In the diagrams, solid lines refer to normal subgroups, dashed lines to sets of conjugate subgroups. Two subgroups $\mathcal{H}_1$ and $\mathcal{H}_2$ are conjugate in $\mathcal{G}$, if (at least) one symmetry operation g_j of $\mathcal{G}$ exists which maps $\mathcal{H}_1$ onto $\mathcal{H}_2$: $g_j^{-1} \circ \mathcal{H}_1 \circ g_j = \mathcal{H}_2$. Conjugate subgroups are "symmetrically equivalent" in $\mathcal{G}$. If a subgroup $\mathcal{H}$ is equivalent only to itself (self-conjugate) and to no other subgroup, $\mathcal{H}$ is a normal (invariant) subgroup of $\mathcal{G}$.

Example: the cubic point group $2/m\bar{3}$ (short : $m\bar{3}$) has six maximal subgroups. They are:

- two normal subgroups, one 23 (cubic) of index [2], and one $2/m2/m2/m$ (short: mmm, orthorhombic) of index [3];
- one set of four conjugate subgroups $\bar{3}$ (trigonal) of index [4]. They are symmetrically equivalent (conjugate) under the "lost" twofold rotations of $2/m\bar{3}$.

In three dimensions, [4] is the largest index of a maximal subgroup of a crystallographic point group. In two dimensions, it is [3]. Note that orthorhombic $2/m2/m2/m$ has three normal (nonconjugate) subgroups of the same type, $mm2$, $m2m$, and $2mm$, whereas the cubic point group 23 has four conjugate (symmetrically equivalent) subgroups of type 3.

When a crystal of symmetry $\mathcal{G}$ undergoes a (displacive) phase transition to a lower point group symmetry $\mathcal{H}$, it splits into (twin) domains which are related by the lost symmetry operations of $\mathcal{G}$. The number of domain states (orientation states) equals the subgroup index $[i] = N(\mathcal{G})/N(\mathcal{H})$.

Examples

1. Quartz (SiO_2) possesses the point group 32 of order $N = 6$ at room temperature. Upon heating above $T_c = 573°C$, it transforms into "high-quartz" with point group 622 and order 12. Upon cooling back below T_c, it loses the twofold axis 2 (contained in the hexagonal axis 6) and returns to "low-quartz" with symmetry 32, whereby it splits into domains of $[i] = 2$ orientation states ("Dauphiné twin"). This is an example of a nonferroelastic transition and merohedral twinning.

2. Barium titanate ($BaTiO_3$) (and other perovskite compounds) undergoes, by cooling below 125°C, a ferroelastic and ferroelectric cubic-to-tetragonal phase transition from point group $4/m\bar{3}\,2/m$ (order 48) to $4\,mm$ (order 8) of index $[i] = 6$. This one-step transformation can be rationalized as a succession of two maximal-subgroup steps:

(a) Reduction from $4/m\bar{3}\,2/m$ to $4/m2/m2/m$ (order 16), both centrosymmetric, with subgroup index [3]. This leads to a set of three conjugate tetragonal subgroups with their tetragonal axes oriented along the former three fourfold cube axes as well as to three ferroelastic 90° domain states.

(b) Further reduction of each of the three conjugate $4/m2/m2/m$ subgroups to its normal subgroup

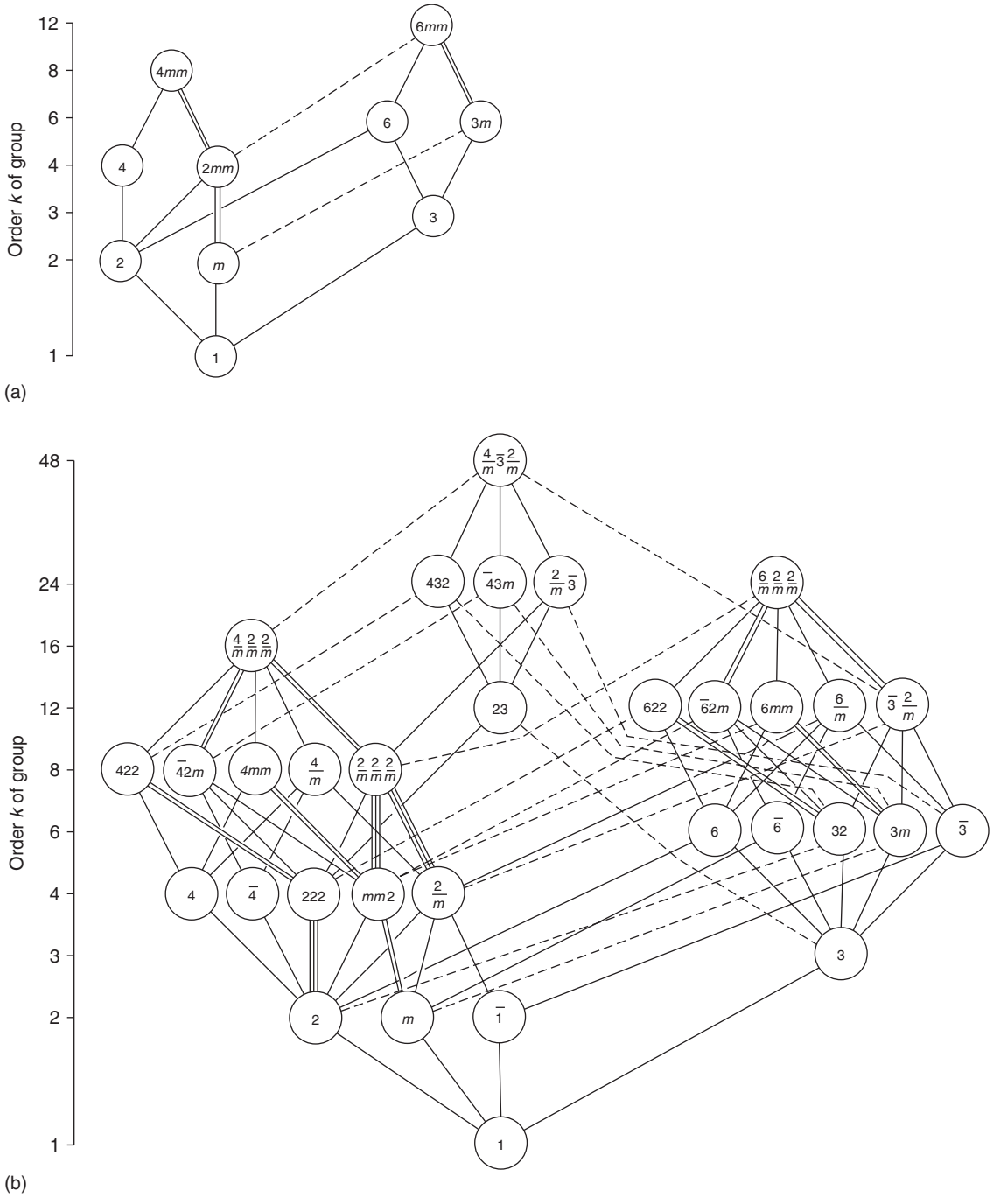


Figure 2 Sub- and supergroup relations of the crystallographic point groups for (a) two and (b) three dimensions. The group orders N are given at the left margin of the diagrams. Solid lines refer to normal subgroups, dashed lines to sets of conjugate subgroups. (Reproduced with permission from Hahn Th and Klapper H (2002) *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*, 5th edn. Dordrecht: Kluwer Academic; © International Union of Crystallography.)

4*mm* of index [2], leading to two antiparallel ferroelectric domain states (180° domains).

The total index of both transition steps is $[3] \times [2] = [6]$. As a result, six different "orientation states" (domain states) are generated: three ferroelastic 90° domain states (pseudo-merohedral twins) [due to step (1)], each one being further split into two ferroelectric 180° domain states (merohedral inversion twins) [result of step (2)].

Symmetry of Physical Properties

Definition of Tensor Symmetry

Physical properties of crystals are strongly governed by symmetry. This holds in particular for the (non-scalar) anisotropic properties described by tensors, and for nontensorial properties such as cleavage and plasticity for which a comprehensive mathematical treatment does not exist. The components of the tensors describing crystal properties are usually referred to a Cartesian coordinate system, that is, three orthogonal axes X_1, X_2, X_3 with equal unit lengths.

Whereas the anisotropic properties themselves are independent of the choice of the coordinate system, their representation by tensors is not, that is, the tensor components change if the coordinate system is transformed. As a consequence, the symmetry of a physical property can be defined in two alternative, but equivalent ways:

1. a rigid motion (rotation, reflection, and inversion), represented by a unitary matrix W ($\det W = \pm 1$), is applied to the crystal, whereby the coordinate system is kept stationary. All motions, for which the tensor components of the transformed crystal are unchanged, are symmetry operations of the tensor property $\mathscr{P}$;
2. alternatively, an arbitrary motion W is applied to the coordinate system with the crystal kept stationary. All motions, for which the tensor components are unchanged under the coordinate transformations W, are symmetry operations of the tensor property $\mathscr{P}$. In both cases the rotation matrices are inverse to each other.

As a consequence of this invariance, some of the tensor components are equal or even zero. The number of independent components is decreased when the symmetry of the crystal increases. Thus, an increase of the point-group symmetry from 1 (triclinic) to $4/m\bar{3}\ 2/m$ (cubic) or to the sphere group $\infty\,\infty m$ (i.e., isotropy) reduces the number of tensor components for symmetrical second-rank tensors from 6 to 1, and for double-symmetrical fourth-rank tensors (e.g., elasticity) from 21 to 3 (cubic) and 2 (isotropy). Even more drastic is this reduction for all tensors of odd rank (such as pyroelectricity and piezoelectricity): all components are zero if an inversion center is present, that is, properties described by odd-rank tensors do not exist in centrosymmetric crystals. This is further elucidated below.

Neumann's Principle

A useful group–subgroup relation between physical and crystallographic symmetries is provided by "Neumann's principle" (first stated in 1833). It states that the symmetry of all physical properties of a crystal is higher than (or at least equal to) its crystallographic point-group symmetry, or in the language of groups: the crystallographic symmetry $\mathscr{G}$ is a proper or improper subgroup of any physical symmetry $\mathscr{P}$, $\mathscr{G} \subseteq \mathscr{P}$. This is illustrated in **Table 4** for second-rank tensors which are, as all even-rank tensors, centrosymmetric. Their symmetries can be recognized by their representation surfaces, which are always ellipsoids or hyperboloids. The symmetry of a general ellipsoid or hyperboloid is $2/m\ 2/m\ 2/m$, but it degenerates to cylindrical and spherical symmetries with point groups $\infty/m\ 2/m$ ($\overline{\infty}2/m$) and $\infty\,\infty m$ ($2/m\,\overline{\infty}$) for uniaxial (i.e., tetragonal, trigonal, hexagonal) and cubic crystals. It is obvious that in all cases the crystallographic symmetry is a proper subgroup of the "physical symmetry," except for the orthorhombic holohedry, for which they are identical. Note that all point groups within the same crystal

Table 4 Symmetry of second-rank polar tensor properties and optical birefringence (refractive-index ellipsoids) in the seven crystal systems

Crystal system	*Holohedry*	*Type of ellipsoid or hyperboloid*	*Symmetry of tensor property*	*Optical birefringence (index ellipsoid)*
Triclinic	$\bar{1}$	General	$2/m\ 2/m\ 2/m$	Biaxial
Monoclinic	$2/m$	General	$2/m\ 2/m\ 2/m$	Biaxial
Orthorhombic	$2/m\ 2/m\ 2/m$	General	$2/m\ 2/m\ 2/m$	Biaxial
Tetragonal	$4/m\ 2/m\ 2/m$	Rotational	$\infty/m\ 2/m$	Uniaxial
Trigonal	$\bar{3}\,2/m$	Rotational	$\infty/m\ 2/m$	Uniaxial
Hexagonal	$6/m\ 2/m\ 2/m$	Rotational	$\infty/m\ 2/m$	Uniaxial
Cubic	$4/m\,\bar{3}\,2/m$	Spherical	$\infty\,\infty m(2/m\,\overline{\infty})$	Isotropic

In column 5 the terms "biaxial" and "uniaxial" refer to the number of "optical axes", i e., of circular sections through the index ellipsoid. For uniaxial crystals, the rotation axis of the tensor surface coincides with the main crystallographic symmetry axis. For orthorhombic crystals, the three principal tensor axes and the crystallographic symmetry axes are parallel. For monoclinic crystals, one principal tensor axis is fixed by symmetry; it is parallel to the monoclinic symmetry axis. For triclinic crystals, none of the principal tensor axes is fixed by symmetry, allowing their "orientation dispersion", e.g., with temperature.

Table 5 The 21 noncentrosymmetric crystallographic point groups (crystal classes) and the occurrence (+) of specific crystal properties

Crystal system	*Noncentrosymmetric point group*	*First-rank tensor (e.g., pyroelectricity)*	*Third-rank tensor (e.g., piezoelectricity)*	*Enantiomorphism*	*Optical activity (opt. gyration)*
Triclinic	1	+ (3) [*uvw*]	+ (18)	+	+ (6)
Monoclinic	2	+ (1) [0 1 0][a]	+ (8)	+	+ (4)
	m	+ (2) [*u* 0 *w*][a]	+ (10)	−	+ (2)
Orthorhombic	222	−	+ (3)	+	+ (3)
	*mm*2	+ (1) [0 0 1]	+ (5)	−	+ (1)
Tetragonal	4	+ (1) [0 0 1]	+ (4)	+	+ (2)
	$\bar{4}$	−	+ (4)	−	+ (2)
	422	−	+ (1)	+	+ (2)
	4*mm*	+ (1) [0 0 1]	+ (3)	−	−
	$\bar{4}$ 2*m*	−	+ (2)	−	+ (1)
Trigonal	3	+ (1) [0 0 1][b]	+ (6)	+	+ (2)
	32	−	+ (2)	+	+ (2)
	3*m*	+ (1) [0 0 1][b]	+ (4)	−	−
Hexagonal	6	+ (1) [0 0 1]	+ (4)	+	+ (2)
	$\bar{6}$	−	+ (2)	−	−
	622	−	+ (1)	+	+ (2)
	6*mm*	+ (1) [0 0 1]	+ (3)	−	−
	$\bar{6}$ 2*m*	−	+ (1)	−	−
Cubic	23	−	+ (1)	+	+ (1)
	432	−	−	+	+ (1)
	$\bar{4}$ 3*m*	−	+ (1)	−	−
Icosahedral	532 (235)	−	−	+	+ (1)
Spherical	$\infty\,\infty$ (2∞)	−	−	+	+ (1)

The number of nonzero independent tensor components is given in parentheses. For comparison, the noncentrosymmetric icosahedral and the sphere group (isotropy) are added. For first-rank tensors the direction [*uvw*] of the property vector is given. There are 10 pyroelectric, 20 piezoelectric, 11 enantiomorphic and 15 optically active crystal classes.
[a] Unique axis *b*
[b] Hexagonal axis.

system (e.g., orthorhombic 222, *mm*2, 2/*m*2/*m*2/*m*) exhibit the same second-rank tensor symmetry.

The more the physical symmetry approximates the crystallographic symmetry, the higher the rank of its tensor. For example, in cubic crystals tensor properties of second rank are isotropic, whereas fourth-rank tensors (e.g., linear elasticity) are not isotropic; they exhibit cubic crystallographic symmetry. For hexagonal crystals, on the other hand, the sixfold axis is still a continuous rotation axis also for fourth-rank tensors.

Tensors of Odd Rank

Physical properties represented by odd-rank tensors are of particular importance for technological applications. Pyroelectricity and piezoelectricity under hydrostatic pressure are properties of first rank, represented by maximally three independent coefficients (i.e., by a vector). Important third-rank properties are piezoelectricity (under general stress), optical second-harmonic generation, and linear electrooptics. Odd-rank tensors are subject to strong constraints by symmetry, in particular by an inversion center which renders all tensor components zero, that is, properties of odd rank do not exist in centrosymmetric crystals.

Pyroelectricity and ferroelectricity Pyroelectricity originates from a permanent electric dipole moment of the unit cell of the crystal structure or, in macroscopic terms, from an intrinsic ("spontaneous") electrical polarization. This polarization is changed by heating and cooling, thus giving rise to electric charges on certain crystal faces.

In a crystal, a spontaneous polarization can be present only along a polar direction which has no

symmetrically equivalent directions. (In the literature, the condition for the occurrence of pyroelectricity is frequently expressed as "a unique (or singular) polar axis." This term, however, is misleading for point groups *m* and 1.) Such polar directions occur in the following ten ("pyroelectric") point groups: 6, 4, 3, 2, 1, and their combination with a "parallel" mirror plane: *6mm*, *4mm*, *3m*, *mm2*, *m* (**Table 5**). In point groups with a rotation axis, the electric polarization ("pyroelectric vector" $\boldsymbol{p}$) is fixed parallel to this polar axis. For point group *m*, the pyroelectric vector $\boldsymbol{p}$ is oriented parallel to the mirror plane, and for point group 1 any direction of $\boldsymbol{p}$ is possible.

A polar crystal is ferroelectric if the direction of the spontaneous polarization can be reversed ("switched") by an electric field, leading to a hysteresis behavior. Thus, ferroelectricity can only occur in the ten pyroelectric point groups.

Piezoelectricity Among the third-rank property tensors, piezoelectrity and its converse effect, the piezoelectric strain ("linear electrostriction"), are the most important properties for technological applications (pressure sensors, resonators, etc.). Well-known examples are single crystals of quartz (SiO_2) and textured ("poled") ceramics of barium titanate ($BaTiO_3$). Of similar significance are, nowadays, the third-rank tensors of optical second-harmonic generation (for optical frequency doubling, for example, potassium dihydrogen phosphate, KDP) and of linear electrooptics (Pockels effect) (for light modulation, e.g., KDP).

Third-rank properties occur in all noncentrosymmetric crystal classes, except for the cubic class 432, where the combination of three perpendicular equivalent fourfold axes causes all tensor components to vanish. The combination of three equivalent twofold axes, present in the two cubic classes 23 and $\bar{4}3m$, however, still admits one independent nonzero component. The tensor is the same for both classes and exhibits the symmetry $\bar{4}3m$ (**Table 5**).

Piezoelectricity under hydrostatic pressure is possible only in the ten pyroelectric crystal classes (see **Table 5**). An example for use of this effect for pressure sensors (e.g., formerly for submarines) is provided by $Li_2SO_4 \cdot H_2O$ (point group 2) single crystals.

Enantiomorphism and Optical Activity (Optical Gyration)

Enantiomorphism is a qualitative property stating that an object (molecule, crystal) may have a right- and a left-handed ("mirror image") form or structure. This feature requires the absence of any "hand-changing" operation ($\det \boldsymbol{W} = -1$), such as a reflection or rotoinversion (including the inversion $\bar{1}$). Thus, only the 11 non-centrosymmetric crystal classes containing only proper rotations ($\det \boldsymbol{W} = +1$) allow enantiomorphism (**Table 5**). Enantiomorphism, of course, is also possible in noncrystallographic symmetry groups without mirror planes and rotoinversion axes.

Optical gyration is described by a second-rank tensor, but in contrast to the polar tensors of second rank shown in **Table 4**, it has an axial character, that is, it exhibits the features of a screw. Thus, the components of the axial gyration tensor change their signs if the handedness of the coordinate system is changed from right to left, and vice versa. Consequently, the invariance condition for tensor components g_{ij} under symmetry transformations, as described in the section "Definition of tensor symmetry" for polar tensors, has to be modified as follows: for all "hand-changing" symmetry operations $\boldsymbol{W}$ (i.e., $\det \boldsymbol{W} = -1$), the transformed components $\tilde{g}_{ij}$ must equal their negative counterparts: $\tilde{g}_{ij} = (-1)g_{ij}$. From this, it follows directly that optical activity is forbidden in all centrosymmetric crystal classes. For proper rotations which preserve the handedness, a change of sign does not occur. Thus, all symmetry groups allowing enantiomorphism also permit optical activity. Crystals belonging to one of the enantiomorphic crystal classes may occur in two forms, either left- or right-handed, which exhibit opposite senses of rotation of the plane of polarization. A well-known example is provided by quartz SiO_2 (enantiomorphic crystal class 32), which crystallizes in a right- and a left-hand form with opposite senses of optical activity.

In addition to the 11 enantiomorphic crystal classes, there are four nonenantiomorphic classes exhibiting optical activity: *m*, *mm2*, $\bar{4}$, $\bar{4}2m$. In these cases, optical activity does not occur for all directions of light propagation, viz. not those parallel to the $\bar{4}$ axis and parallel or normal to the mirror plane *m*. All other propagation directions exhibit activity, whereby directions symmetrical with respect to $\bar{4}$ and *m* show opposite rotation senses. Thus, the four symmetry groups above show both (right and left) rotation senses in one and the same crystal.

The optical activity is isotropic in cubic crystals, as it is in optically active liquids (e.g., in solutions of dextro- or levo-sugar). In this case, the symmetry of the optical activity is described by the noncentrosymmetric sphere group $\infty\infty(2\infty)$. Geometrically, this symmetry group can be represented by a (right- or left-hand) rotating sphere or by a sphere which is densely covered with either right- or left-hand swirls.

Curie's Principle

Until now the symmetry group $\mathcal{G}_P$ of physical properties of a "free" macroscopic crystal *C* with point

group $\mathcal{G}_C$ was discussed. Now, the symmetry $\mathcal{G}_{CF}$ of a crystal C which is subject to an external influence F, for example, to an electric field, to uniaxial stress, to temperature changes, etc., is considered. For this treatment, point-group symmetries $\mathcal{G}_F$ of the external influences (Curie groups) are defined as follows:

1. Homogeneous electric field $\boldsymbol{E}$: ∞m (polar continuous rotation axis with "parallel" mirror planes, that is, symmetry of a stationary cone). Note that a rotating cone (left- or right-hand) represents geometrically the polar enantiomorphous group ∞; a stationary cone represents the polar group ∞m with "vertical" mirror planes.
2. Homogeneous magnetic field $\boldsymbol{H}$: ∞/m (symmetry of a rotating cylinder).
3. Uniaxial stress σ_{ii}: $\infty/m\ 2/m$ (symmetry of a stationary centrosymmetric cylinder).
4. Temperature change ΔT or hydrostatic pressure $-\sigma$ (scalars): $\infty\ \infty\ m$ (symmetry of a stationary centrosymmetric sphere).
5. Shear stress σ_{ij}: $2/m\ 2/m\ 2/m$ (orthorhombic).

According to Curie's principle of 1894, the point-group symmetry $\mathcal{G}_{CF}$ of the crystal under the external field is the intersection symmetry of the symmetry $\mathcal{G}_C$ of the crystal without field and the symmetry $\mathcal{G}_F$ of the field without crystal: $\mathcal{G}_{CF} = \mathcal{G}_C \cap \mathcal{G}_F$; that is, $\mathcal{G}_{CF}$ is a (proper or improper) subgroup of both groups $\mathcal{G}_C$ and $\mathcal{G}_F$.

As examples, consider the effect of an electric field ($\mathcal{G}_F = \infty m$) and a uniaxial stress ($\mathcal{G}_F = \infty/m\ 2/m$) along one of the (four) threefold rotoinversion axes $\bar{3}$ of cubic crystals with point groups $\mathcal{G}_C = 2/m\bar{3}$ and $\mathcal{G}_C = 4/m\bar{3}\,2/m$

- Electric field parallel [1 1 1]:

$$2/m\,\bar{3} \cap \infty m:$$
$$\mathcal{G}_{CF} = 3 \quad \text{along } [1\,1\,1]$$
$$\text{(pyroelectric, optically active)}$$

$$4/m\,\bar{3}\,2/m \cap \infty m:$$
$$\mathcal{G}_{CF} = 3m \quad \text{along } [1\,1\,1]$$
$$\text{(pyroelectric)}$$

- Uniaxial stress parallel [1 1 1]:

$$2/m\bar{3} \cap \infty/m\,2/m:$$
$$\mathcal{G}_{CF} = \bar{3} \text{ along } [1\,1\,1]$$
$$4/m\bar{3}\ 2/m \cap \infty/m\,2/m:$$
$$\mathcal{G}_{CF} = \bar{3}2/m \quad \text{along } [1\,1\,1]$$

Note that, in contrast to uniaxial stress, the electric field destroys centrosymmetry, leading to pyroelectricity and even optical gyration. The polarity and the gyration sense are reversed upon reversal of the electric field. If the electric field and the uniaxial stress were applied to an arbitrary (nonsymmetry) direction of the above cubic crystals, point groups 1 and $\bar{1}$ would result in the two cases.

If a scalar influence, which is represented by the centrosymmetric sphere group $\mathcal{G}_E$, for example, a temperature change ΔT, is applied to a crystal, its symmetry is not changed, $\mathcal{G}_{CF} = \mathcal{G}_C$, provided that no phase transition is induced. For further reading on the symmetry of physical properties, refer to the "Further reading" section.

See also: Crystal Symmetry; Crystal Tensors: Applications; Crystal Tensors: Introduction; Group Theory; Magnetic Point Groups and Space Groups; Space Groups.

PACS: 61.50.Ah; 61.68.+n; 02.20.Hj

Further Reading

Authier A (ed.) (2003) *International Tables for Crystallography, Vol. D, Physical Properties of Crystals*. Dordrecht: Kluwer Academic.

Bloss ED (1971) *Crystallography and Crystal Chemistry*. New York: Holt, Rinehart and Winston.

Buerger MJ (1965) *Elementary Crystallography*. New York: Wiley.

Buerger MJ (1971) *Introduction to Crystal Geometry*. New York: McGraw-Hill.

Hahn Th and Klapper H (2002) Point groups and crystal classes. In: Hahn Th (ed.) *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*, 5. Edition, Dordrecht: Kluwer Academic.

Kleber W, Bautsch HJ, Bohm J, and Kleber I (1990) *Einführung in die Kristallographie*, 17., 18. Edition 1998. Berlin: Verlag Technik; now München: Oldenbourg.

Ledermann W (1976) *Introduction to Group Theory*. London: Longman.

McKie D and McKie Ch (1986) *Essentials of Crystallography*. Oxford: Blackwell.

Nye JF (1957) *Physical Properties of Crystals. Their Representation by Tensors and Matrices*, 1. Edition 1957, Revised Edition 1985. Oxford: Clarendon Press.

Paufler P (1986) *Physikalische Kristallographie*. Berlin: Akademie-Verlag and Weinheim: Verlag Chemie (VCH).

Philips FC (1971) *An Introduction to Crystallography*, 4. Edition. London: Longman.

Schwarzenbach D (1996) *Crystallography*. Chicester: Wiley. (original French edition 1996, German edition 2001).

Shuvalov LA (ed.) (1988). *Modern Crystallography*. Vol. 4, *Physical Properties of Crystals*. Berlin: Springer.

Vainstein BK (1994) *Modern Crystallography*. Vol. 1, *Fundamentals of Crystals. Symmetry and Methods of Structural Crystallography*, 2. Edition. Berlin: Springer.

Nomenclature

e	neutral element, identity
F	external physical influence (external field) on a crystal
g_i	element of $\mathscr{G}$
$\mathscr{G}$	point group of a crystal
$\mathscr{H}$	subgroup
$[i]$	index of a subgroup or supergroup; number of orientation states of domains
$N(\mathscr{G}) = \lvert\mathscr{G}\rvert$	order of group $\mathscr{G}$
$\boldsymbol{p}$	electric polarization vector, pyroelectric vector
P	physical (tensor) property of a crystal
$\mathscr{P}$	symmetry group of the property P
W	(3×3) matrix of a point-group operation
σ_{ii}, σ_{ij}	uniaxial stress, shear stress
$-\sigma$	hydrostatic pressure

Polaritons

M Litinskaya, Institute of Spectroscopy, Troitsk, Russia

Introduction

Definition

By definition, polaritons are quasiparticles in a medium that form as a result of interaction and mixing of light with dipole active transitions of the medium. The classical theory of light waves in ionic crystals near optical phonon frequencies was proposed by K B Tolpygo and K Huang in the early 1950s. Later, the quantum theory of light waves near excitonic resonances and the concept of polaritons were developed in works by U Fano, J J Hopfield, and V M Agranovich.

Consider an electromagnetic wave propagating in a medium, and let the frequency of the wave be close to some resonance frequency of the medium. These frequencies are generally associated with some quasiparticles. They can be of electronic nature, such as Frenkel or Wannier–Mott excitons, or related to lattice vibrations, such as optical phonons in ionic crystals. The electromagnetic field polarizes the medium. The polarization, which is high near the resonance frequency, in turn influences the electromagnetic field. One can regard it as an interaction between the "bare" photon and "bare" medium excitation ("light–matter interaction"). When the interaction is strong enough, the excitation spectrum of the medium essentially changes. Near the resonance between the light mode and the mode of the medium excitation, their dispersion curves transform into two – the "lower" and the "upper" – split polaritonic branches showing anticrossing behavior. In other words, a gap of the size 2Δ, where Δ is the coupling strength defined below, opens in the excitation spectrum. This situation is referred to as the strong coupling regime. Roughly speaking, polaritons possess the properties of both "bare" photons and "bare" medium excitations. This mixed character of polaritons leads to many interesting physical properties.

Usually, one indicates the type of crystal excitations which participates in the formation of polaritons by adding prefixes: "exciton–polaritons," "plasmon–polaritons," and "optical phonon–polaritons." The quasiparticles, which appear as a result of interaction between an electromagnetic field and the resonances of magnetic permeability, are known as "magnetic polaritons."

Theoretical Approaches

Many of the physical processes concerning polaritons, as well as their dispersion law, can be described by means of the macroscopic (semiclassical) approach. It utilizes the system of Maxwell's equations for the fields in the medium. For spatially restricted systems, the problem must be supplemented by the fields outside the crystal and by proper boundary conditions. These equations must be combined with the material equation, which relates the fields $\boldsymbol{D}$ and $\boldsymbol{E}$ through the tensor of dielectric permeability $\varepsilon_{ij}(\omega, \boldsymbol{k})$. The tensor $\varepsilon_{ij}(\omega, \boldsymbol{k})$ can be obtained from the microscopic theory or from experiment. In an isotropic medium with the account of spatial dispersion,

$$\varepsilon_{ij}(\omega, \boldsymbol{k}) = \varepsilon_{\mathrm{t}}(\omega, k)\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) + \varepsilon_{\mathrm{l}}(\omega, k)\frac{k_i k_j}{k^2} \qquad [1]$$

where $\varepsilon_{\mathrm{t}}(\omega, k)$ and $\varepsilon_{\mathrm{l}}(\omega, k)$ are, respectively, transverse and longitudinal dielectric permeabilities, and δ_{ij} is the Kronecker symbol. The condition $\varepsilon_{\mathrm{l}}(\omega, k) = 0$ determines the dispersion $\omega = \omega_{\mathrm{l}}(k)$ of a longitudinal wave, which does not interact directly with the transverse photons. In turn, the poles $\omega = \omega_{\mathrm{t}}(k)$ of the function $\varepsilon_{\mathrm{t}}(\omega, k)$ determine the transversal "bare"

excitations of the medium. As shown below, polaritons are just the transverse normal modes in the medium near these frequencies. The macroscopic description is effective as far as the wavelength of light exceeds the lattice constant and the excitonic Bohr radius strongly.

In the framework of the microscopic (quantum) approach, polaritons appear as a result of the account of retardation. It is well-known that excitons are the excitations in a medium in the approximation when only the instantaneous Coulomb interaction between the molecules (or atoms) is taken into account (i.e., in the limit $c \to \infty$). The inclusion of retardation, as known from quantum electrodynamics, is equivalent to taking into account the interaction with a transverse electromagnetic field. Hence, polaritons appear as a result of interaction between excitons and transverse photons. Such a description can also be applied to low-dimensional structures (such as, a monolayer or a chain of molecules), when a dielectric tensor cannot be introduced and the macroscopic description is not valid. This approach utilizes the second-quantization formalism, and polaritons appear as a result of diagonalization of the full Hamiltonian of the crystal, which is a sum of Hamiltonians of "bare" transverse photons, "bare" excitons, and the Hamiltonian of their interaction. The total Hamiltonian can be diagonalized by operators that are linear combinations of the operators of "bare" photons and "bare" excitons. Near the resonance, these operators are of the form

$$\begin{aligned} \xi^{+}_{\boldsymbol{k},\rho} &= C^{(\rho)}_{\text{phot}}(k) a^{+}_{\boldsymbol{k}} + C^{(\rho)}_{\text{exc}}(k) B^{+}_{\boldsymbol{k}} \\ \xi_{\boldsymbol{k},\rho} &= C^{(\rho)*}_{\text{phot}}(k) a_{\boldsymbol{k}} + C^{(\rho)*}_{\text{exc}}(k) B_{\boldsymbol{k}} \end{aligned} \qquad [2]$$

where $a^{+}_{\boldsymbol{k}}$, $a_{\boldsymbol{k}}$ and $B^{+}_{\boldsymbol{k}}$, $B_{\boldsymbol{k}}$ are, respectively, the creation and annihilation operators of photons and excitons, and $\rho = \text{U}, \text{L}$ denote the upper or the lower polaritonic branch, respectively. The coefficients $C^{(\rho)}_{\text{phot}}(k)$ and $C^{(\rho)}_{\text{exc}}(k)$ define, respectively, the amplitudes of the photon and the exciton in the polaritonic states. Their squared modules give the weights of the "bare" excitations in the polaritonic states. With the change of the wave vector k, they vary between 0 and 1. At the resonance $|C^{(\text{U,L})}_{\text{phot}}(k_{\text{res}})|^2 = |C^{(\text{U,L})}_{\text{exc}}(k_{\text{res}})|^2 = 1/2$.

Within linear optics, these two approaches yield identical results, and the choice between them is just a matter of convenience. Quite often, a semiclassical (macroscopic) approach is used to calculate the optical response of a system. Quantum (microscopic) formalism describes polaritons as quasiparticles, and can be used to discuss the collective properties of polaritons and their statistics. It is convenient for a discussion on nonlinear processes, and allows one to describe the eigenfunctions of the crystal (in particular, the polaritonic vacuum state is not equivalent to an independent-particle vacuum).

When Strong-Coupling Regime Is Destroyed

The strong-coupling regime can be destroyed by different dephasing processes. Many of them (such as, exciton–phonon interactions and scattering by sample disorder) influence polaritons through their excitonic components. In semiconductor microcavities, the lifetime of polaritons is also restricted by the finite lifetime of the cavity photon within the cavity due to a finite transmission of the cavity mirrors. Roughly, the strong-coupling regime holds when the exciton–photon coupling strength Δ is larger than all the involved dephasing rates of polaritons. If this condition is violated, the system is in a so-called weak-coupling regime, when the gap in the spectrum of elementary excitations does not form.

The role of phonons can be diminished by decreasing the temperature. One can say, for instance, that in good GaAs crystals, the strong-coupling regime holds below 20 K in bulk, and below 70 K in microcavities.

Disorder appears as impurities in bulk crystals, and interface roughnesses and/or alloy fluctuations for cavity and surface polaritons. Bulk polaritons are very sensitive to the disorder. In turn, in state-of-the-art nanostructures, the role of disorder can be brought almost to nothing. The main effect of disorder is that it introduces local excitonic levels, which trap polaritonic states and eventually destroys them.

Polaritons in Bulk

Dispersion Equation

In the framework of the macroscopic approach, the dispersion equation for polaritons in a bulk medium can be obtained by solving the wave equation for the fields $\boldsymbol{D}$ and $\boldsymbol{E}$:

$$\Delta \boldsymbol{E} + \frac{\omega^2}{c^2} \boldsymbol{D} - \text{grad div } \boldsymbol{E} = 0 \qquad [3]$$

combined with the material equation. In an isotropic medium, neglecting dissipation, the dielectric permeability $\varepsilon_{\text{t}}(\omega, k)$ is

$$\varepsilon_{\text{t}}(\omega, k) = \varepsilon_{\text{b}} + \frac{A}{\omega^2_{\text{t}}(k) - \omega^2} \qquad [4]$$

where ε_{b} is the background dielectric constant, and A is proportional to the oscillator strength f_{osc} of the transition. Omitting the resonance term yields "bare" photon approximation. Further, $\omega_{\text{t}}(k)$ is the dispersion of the crystal excitation, which, for small k, can

be written in the effective-mass approximation:

$$\omega_t(k) = \omega_t + \frac{\hbar k^2}{2m_t}, \quad \omega_t \equiv \omega_t(k=0) \qquad [5]$$

Here, m_t is the effective mass of the excitation, and it determines the curvature of the exciton dispersion. If the dependence of the dielectric permeability on the wave vector (i.e., spatial dispersion) is neglected, then $\omega_t(k) = \omega_t$ and $\varepsilon_t(\omega, k) \equiv \varepsilon(\omega)$. This usually applies to optical phonon–polaritons and some of Frenkel exciton–polaritons, while for Wannier–Mott exciton–polaritons, the dependence of $\varepsilon_t(\omega, k)$ on k is usually important.

In an infinite isotropic medium, the wave eqn [3] has three solutions. One is a longitudinal wave, and its dispersion law $\omega = \omega_l(k)$ is determined by $\varepsilon_l(\omega, k)$. In addition, there are two twice-degenerate (corresponding to two possible polarizations) transverse waves, which are polaritons. Their dispersion law can be found from the equation

$$\frac{c^2k^2}{\omega^2} = \varepsilon_t(\omega, k) = \varepsilon_b + \frac{A}{\omega_t^2(k) - \omega^2} \qquad [6]$$

Solving this equation for $\omega \sim \omega_t$, one finds

$$(\omega - \omega_{phot}(k))(\omega - \omega_t(k)) = \Delta^2 \qquad [7]$$

where $\Delta = \sqrt{A/(4\varepsilon_b)} \propto \sqrt{f_{osc}}$ is the coupling strength, and $\omega_{phot}(k) = ck/\sqrt{\varepsilon_b}$ is the photon dispersion in the bulk. Near the resonance (i.e., for $\omega \sim \omega_t(k) \approx \omega_{phot}(k)$), the polaritonic branches split. The smallest separation between them is at the resonance, and is equal to 2Δ. Far from the resonance, the polaritonic effect is usually not important, and polaritonic dispersion curves tend to that of the photon and the exciton. The solutions of [7] $\omega = \omega_{U,L}(k)$ are

$$\omega_{U,L}(k) = \frac{1}{2}\left[\omega_{phot}(k) + \omega_t(k) \pm \sqrt{(\omega_{phot}(k) - \omega_t(k))^2 + 4\Delta^2}\right] \qquad [8]$$

The dispersion curves of bulk excitations (including the longitudinal wave, which does not couple to light) are plotted in **Figure 1** ((**a**) without and (**b**) with spatial dispersion).

The dispersion of bulk exciton–polaritons has been measured for many semiconductors (such as, CuCl, GaAs, CdS) in a variety of optical experiments (for instance, two-photon absorption measured at different angles or hyper-Raman scattering measurements).

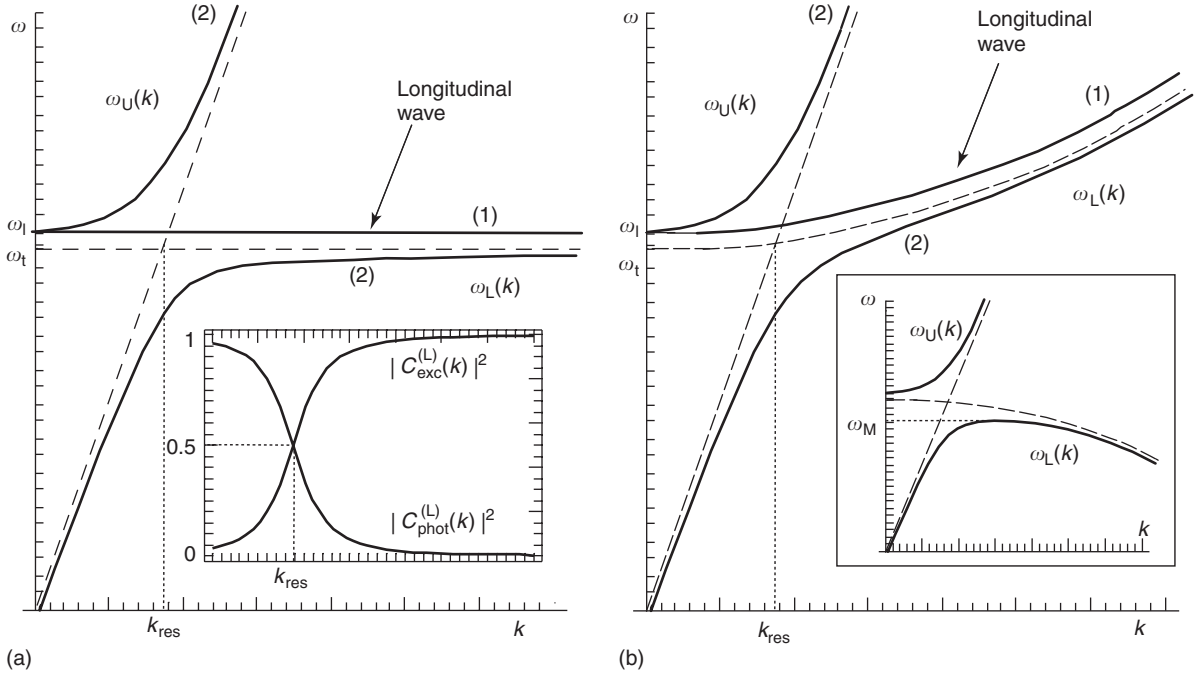


Figure 1 Dispersion of bulk polaritons and of the longitudinal wave (solid lines). The numbers in brackets show the degeneracy of each mode. Dashed lines correspond to the dispersions of "bare" photon and "bare" transverse exciton. (a) Spatial dispersion is neglected. The inset shows the dependence on the wave vector of the exciton and photon weight coefficients $|C_{exc}^{(L)}(k)|^2$ and $|C_{phot}^{(L)}(k)|^2$ defined in eqn [2] for the lower branch. (b) Spatial dispersion is taken into account, and $m_t > 0$, $m_l > 0$. The inset shows the spectrum of transverse waves for $m_t < 0$.

Anisotropic Media

The properties of polaritons in anisotropic crystals (e.g., organic crystals) are determined by the structure of the dielectric tensor $\varepsilon_{ik}(\omega, \boldsymbol{k})$ (see [1]), which relates the fields as

$$D_i(\omega, \boldsymbol{k}) = \varepsilon_{ik}(\omega, \boldsymbol{k}) E_k(\omega, \boldsymbol{k})$$

Then the equation analogous to [6], which yields the dispersion of polaritons $\omega = \omega(k)$, can be obtained from the zeros of the following determinant:

$$\left| \frac{\omega^2}{c^2} \varepsilon_{ik}(\omega, \boldsymbol{k}) - k^2 \delta_{ik} + k_i k_k \right| = 0 \qquad [9]$$

For an arbitrary direction of the wave vector $\boldsymbol{k}$, polaritons in anisotropic crystals have a mixed longitudinal–transverse character.

Spatial Dispersion and Additional Waves

Spatial dispersion is a result of the ability of excitations to propagate in the crystal due to internal interaction forces. As mentioned already, it results in the dependence of dielectric permeability on the wave vector.

If spatial dispersion is neglected, then there are no bulk waves in the interval $\omega_t < \omega < \omega_l = \omega_l(0)$, and for every other frequency, there exists only one propagating bulk wave (see **Figure 1a**).

Spatial dispersion results in the appearance of the so-called "additional waves" (waves with the same energy $\hbar\omega$, but different wave vectors $k(\omega)$), first noticed by S I Pekar. Near a resonance, it is enough to take into account the dispersion of the excitonic resonance, as it is done in [5] for an isotropic medium. Then the curve $\omega_t(k)$ bends up or down, depending on the sign of m_t. The curve $\omega_l(k)$ also bends depending on the sign of the effective mass m_l of the longitudinal exciton, since in the general case, $m_l \neq m_t$. Several examples are given below. First, in an isotropic nongyrotropic medium with $m_t > 0$, for $\omega > \omega_l$, there exist two propagating transverse waves, which are the lower and the upper polaritons (see **Figure 1b**). In the "forbidden" region ($\omega_t < \omega < \omega_l$), there appears a state of the lower polaritonic branch. In **Figure 1b**, m_l is also positive, that is typical for inorganic semiconductors. Then, for $\omega > \omega_l$, the longitudinal wave also exists. Second, if $m_t < 0$, as it is in some molecular crystals, there exist two propagating transverse waves (both are lower-branch polaritons) for $\omega < \omega_M < \omega_t$ (see the inset in **Figure 1b**). As regards the longitudinal wave, if $m_l > 0$, then it exists for $\omega > \omega_l$, together with the upper polaritonic branch, and if $m_l < 0$, then the longitudinal wave is the third wave for $\omega < \omega_M$, and it is the only one wave for $\omega_M < \omega < \omega_l$. Third, in a cubic gyrotropic crystal, the dispersion curve of the exciton splits into two branches with linear dispersion: $\omega_t(k) = \omega_t \pm \alpha k + O(k^2)$. The corresponding polaritonic spectrum with three propagating waves for $\omega < \omega_G < \omega_t$ and one propagating wave for $\omega > \omega_G$ is shown in **Figure 2**.

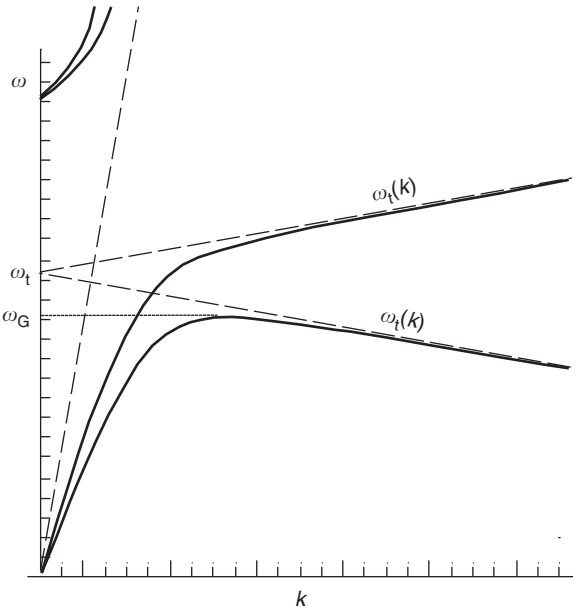


Figure 2 Dispersion of bulk polaritons (thick solid lines) in a gyrotropic cubic crystal. Dashed lines show the dispersion of "bare" excitons and "bare" photon.

For spatially restricted media (e.g., when calculating reflectance of a bulk crystal) at the frequencies where the additional waves exist, one has to introduce the so-called additional boundary conditions (ABCs), since the usual Maxwell's boundary conditions are not sufficient to find the relation between the amplitudes of the waves. From a microscopic point of view, ABCs describe the behavior of the fields close to the surface. The character of ABCs essentially depends on the properties of the medium, its surface, and on the properties of excitons. In the most general case, in the framework of linear optics, ABCs are constructed as some linear combinations of the fields $\boldsymbol{E}$, $\boldsymbol{P}$, and their derivatives. The first ABC proposed by S I Pekar is $\boldsymbol{P} = 0$ at the boundary of the medium. It is good for molecular crystals, but is often suitable for semiconductors also. A widely used approach proposed by J J Hopfield and D G Thomas introduces a "dead layer" of some thickness l near the surface, where excitons cannot penetrate. The thickness l and other phenomenological constants entering ABCs are to be determined from a microscopic theory or from comparison with experiment.

Surface Polaritons

In a medium with a surface, specific surface states, known as surface polaritons, exist in certain (ω, k) regions. Surface polaritons are normal modes that propagate along the medium surface (or along an interface between two media, or along interfaces of layered structures). The electromagnetic field in these modes decays exponentially with the distance from the surface. It will be shown below that surface polaritons are usual solutions of Maxwell's equations applied to the case when the dielectric permeability is negative (the situation realizable near a resonance frequency of the medium). Surface polaritons yield information on the surface of the crystal. They may also be important as possible final states in the different processes of interaction of electromagnetic waves in crystals.

Surface Polaritons at the Interface between Two Semi-Infinite Media

Let a semi-infinite isotropic medium with the dielectric permeability $\varepsilon(\omega)$ (spatial dispersion is neglected) be contiguous with an isotropic half-space with dielectric constant ε_{out}. Let $\boldsymbol{k}_{\parallel}$ be a two-dimensional wave vector parallel to the surface. The surface mode is TM-polarized (see the inset in **Figure 3**) and has the following dispersion relation:

$$\frac{\varepsilon(\omega)}{\sqrt{k_{\parallel}^2 - \omega^2\varepsilon(\omega)/c^2}} = -\frac{\varepsilon_{\text{out}}}{\sqrt{k_{\parallel}^2 - \omega^2\varepsilon_{\text{out}}/c^2}} \quad [10]$$

For $\varepsilon_{\text{out}} > 0$, the dispersion equation of surface polaritons has solutions only if $\varepsilon(\omega) < 0$. The dispersion equation can be rewritten in a more convenient form:

$$k_{\parallel}^2 = \frac{\omega^2}{c^2}\frac{\varepsilon(\omega)\varepsilon_{\text{out}}}{\varepsilon(\omega) + \varepsilon_{\text{out}}} \quad [11]$$

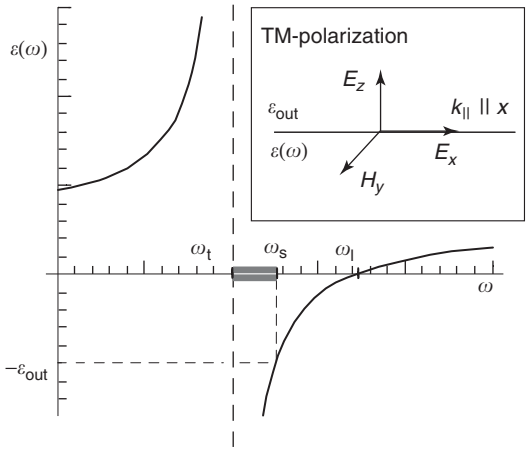


Figure 3 A graph of the function $\varepsilon(\omega)$. The frequencies marked in gray correspond to the interval of existence of surface polaritons at the interface with a material with a dielectric constant ε_{out}.

From here, it is seen that the permissible interval of the surface polariton frequencies is more narrow, since for real $k_{\parallel}$ and negative dielectric permeability, $\varepsilon(\omega)$, in fact, must be less than $-\varepsilon_{\text{out}}$. Let ω_S denote the upper boundary of the existence of surface polaritons: $\varepsilon(\omega_S) = -\varepsilon_{\text{out}}$. When ε_{out} approaches ε_b, $\omega_S \to \omega_l$. For a dielectric material, the function $\varepsilon(\omega)$ is given by [4]; it is shown in **Figure 3**, and the frequency interval where a surface polariton exists is marked in gray. For a metal, when $\varepsilon(\omega) = 1 - \omega_p^2/\omega^2$ (ω_p is the plasma frequency), surface plasmon–polariton exists for $0 < \omega < \omega_p/\sqrt{1 + \varepsilon_{\text{out}}}$.

Figure 4a shows the dispersion of surface polaritons without spatial dispersion. The thin dashed line shows the dispersion of light outside the crystal (propagating parallel to the crystal surface), and the thick dashed lines show bulk polariton dispersion curves. All these dispersion curves do not intersect with the dispersion curve of the surface polariton, hence surface polaritons cannot be excited by an incident photon (or, once excited, cannot transform into bulk polaritons or leave the crystal as photons) – the laws of conservation of energy and wave vector cannot be simultaneously fulfilled for such processes. To excite surface polaritons, one has to make use of a prism (attenuated total-reflection technique) or a grating on the crystal surface. Surface polaritons have been seen as dips in the total reflection of TM-polarized light in many such experiments.

Spatial dispersion complicates the problem, because then a lower "bulk" polariton appears for large wave vectors in the frequency interval where a surface polariton exists (see **Figure 4b**). As discussed above, it requires introduction of ABCs, which mix surface polaritons with bulk lower polaritons. The mixing leads to leakage of the energy of the surface wave into the interior of the crystal, providing a new (nonradiative and nondissipative) damping mechanism for surface polaritons.

Surface Polaritons in Presence of a Transition Layer

The presence of a surface transition layer can essentially alter the dispersion of surface polaritons. The transition layer may be either natural or artificial (a thin film on a substrate). As shown by V M Agranovich and A G Mal'shukov, if an oscillation in the layer is resonant with the surface polariton, a gap with a magnitude Δ_S opens in the surface polariton spectrum. This is similar to what happens to bulk electromagnetic waves near the resonance frequency with the difference that both, the excitation

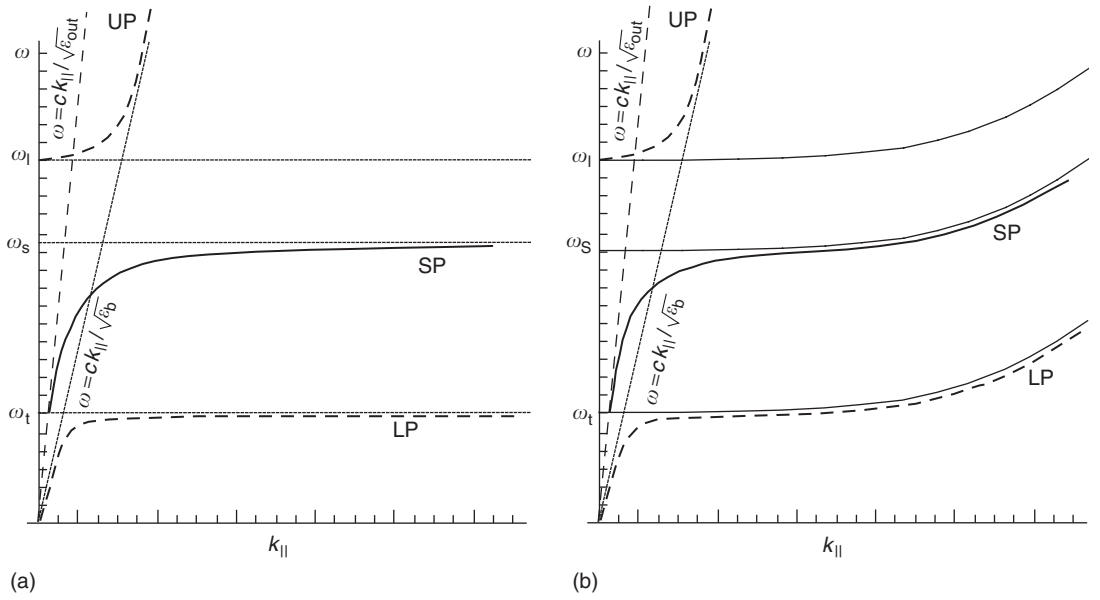


Figure 4 Dispersion of surface polariton (solid line) at an interface between a resonant isotropic crystal with dielectric function $\varepsilon(\omega)$ and an isotropic material with a dielectric constant ε_{out} (a) without and (b) with spatial dispersion. The thick dashed lines show the dispersions of bulk polaritons, and the thin dashed line shows the dispersion of the photon outside of the crystal. UP, LP, and SP denote, respectively, the upper, lower, and the surface polariton branches.

in the layer and the surface electromagnetic wave, have a two-dimensional character. While usually the effects of a thin layer are of the order of $d/\lambda \ll 1$ (d is the layer thickness, λ is the wavelength of the light), here $\Delta_S \propto \sqrt{d/\lambda} \gg d/\lambda$, and therefore the gap had been observed in a number of experiments, both in the infrared spectral region and in the spectral region of electronic transitions.

Polaritons in Nanostructures with Reduced Dimensionality

From the above discussion, it follows that in bulk crystals the interaction of excitons with an electromagnetic field does not give rise to radiative decay of the exciton. Instead, the energy oscillates back and forth between exciton and photon states, and there is no radiative decay. It happens because the total wave vector conserves (an exciton with a given wave vector interacts with only one photon with the same wave vector) and there is no density of states for radiative decay.

Another situation occurs in two- and one-dimensional (1D) crystals, where an exciton has less degrees of freedom than a photon. As noted by V M Agranovich and O A Dubovsky, in such crystals the radiative decay is restored, since in these cases only two or one of the components of the total wave vector, respectively, are conserved.

Consider polaritonic effects in quasi-2D quantum wells and quasi-1D quantum wires described elsewhere in this encyclopedia. The translation symmetry is preserved (and, consequently, the wave vector $\boldsymbol{k}_{||}$ of the exciton is a good quantum number) only in the plane of the quantum well in the first case, and only along the quantum wire in the second case. The component of the total wave vector perpendicular to $\boldsymbol{k}_{||}$ is denoted by $\boldsymbol{k}_{\perp}$. Photons are 3D, as before.

There are two kinds of states. Consider first the states with small values of the wave vector $k_{||} < k_0 \equiv \omega\sqrt{\varepsilon_b}/c$. They are degenerate with photons and can decay into them. One can regard it as a coupling of an exciton with a wave vector $\boldsymbol{k}_{||}$ with photons with the same component $\boldsymbol{k}_{||}$ parallel to the quantum well or the quantum wire, but with all possible values of $\boldsymbol{k}_{\perp}$. These photons act as a dissipative bath, and instead of forming stable states similar to polaritons in bulk, the exciton decays irreversibly.

The corresponding radiative lifetimes can be calculated treating exciton–light coupling as a perturbation. For Frenkel exciton–polaritons, the decay times in 1D and 2D are, respectively, $\tau_{1D}^{(F)} \sim (2\pi a/\lambda)\tau_0, \tau_{2D}^{(F)} \sim (2\pi a/\lambda)^2\tau_0$, where $\tau_0 \sim 10^{-9}$–10^{-8} s is the radiative lifetime of the excitation of an isolated molecule, and a is the lattice constant (i.e., the radiative decay is enhanced proportionally to the number of molecules within the wavelength of light, λ). The radiative lifetime of exciton–polaritons in 2D is then $\sim 10^{-12}$ s, and such "superradiant decay"

had been observed in anthracene. For Wannier–Mott exciton–polaritons, the lifetime of 2D exciton–polariton for $k_{\|}=0$ was found to be $\tau_{2D}^{(WM)} = m_t c\sqrt{\varepsilon_b}/(2\pi e^2 f_{osc})$. For ideal situations, such as in GaAs, it is $\sim$10 ps. The account of the effects of thermalization and disorder-induced partial localization of polaritons considerably complicates the picture both in 1D and 2D cases.

For $k_{\|} > k_0$, the modes are nonradiative. In quantum wells, to some extent, they are similar to surface polaritons, since the electromagnetic field in them decays exponentially outside of the quantum-well plane. Similarly, the propagating modes can be excited using a prism or with surface gratings. Radiative polaritons with $k_{\|} < k_0$ can be excited directly by incident light and have been observed in many semiconductor nanostructures in absorption, photoluminescence, and reflectivity spectra.

Cavity Polaritons

In multilayered nanostructures called semiconductor microcavities, both excitons and photons have a 2D character. The interaction between them leads to formation of states similar to bulk polaritons known as cavity polaritons. These states are stable in the sense that the exciton with a given 2D wave vector $\boldsymbol{k}_{\|}$ interacts with only one cavity photon with the same $\boldsymbol{k}_{\|}$, and the energy oscillates back and forth between these states. In other words, in contrast to the nanostructures where photons are 3D, there is no density of states for the radiative decay of the exciton. After several oscillations, light waves escape the cavity due to a finite transmission of the cavity mirrors. Cavity polaritons had, for the first time, been observed by C Weisbuch, M Nishioka, A Ishikawa, and Y Arakawa. The important difference with bulk polaritons is, first, that as result of a reduction of the excitonic Bohr radius due to the confinement, both the oscillator strength and the binding energy of the exciton are enhanced. As a consequence, Wannier–Mott exciton–polaritons become considerably more stable than in the bulk. Second, lower polaritons have a finite energy at vanishing $k_{\|}$, and this lowest-energy state may, in principle, work as a "trap" for relaxing polaritons.

The properties of cavity polaritons are discussed below in the context of well-investigated examples of isotropic inorganic material, such as, GaAs. However, exciton–polaritons in organic microcavities have recently attracted attention due to their large oscillator strengths (and, consequently, very large values of the splitting Δ at room temperature) in some of these materials.

The dispersion equation of cavity polaritons can be accurately obtained from Maxwell's equations with the help of the transfer-matrix method, which relates the transmission and reflectance at each interface of the layered structure. Near the resonance, one obtains the dispersion equation of the type [7] with solutions similar to [8], with the energies of "bare" exciton and photon dependent on the in-plane wave vector $k_{\|}$. A significant difference arises for the cavity photon dispersion, since the perpendicular component of its total wave vector is fixed by the boundary conditions at the microcavity mirrors, leading to parabolic (for small $k_{\|}$) dispersion of the cavity photon:

$$\omega_{phot}(k_{\|}) = \omega_c + \frac{\hbar k_{\|}^2}{2\mu} \qquad [12]$$

where $\omega_c \equiv \omega_{phot}(0) = \pi c/(L_c\sqrt{\varepsilon_b})$, L_c is the microcavity thickness, and $\mu = \pi\hbar\sqrt{\varepsilon_b}/(cL_c)$ is its effective mass, which is very small ($\sim 10^{-5} m_e$, m_e is the electron mass). The dispersion curves of cavity polaritons are plotted in **Figure 5**.

The detuning between the cavity photon and quantum-well exciton $u = \omega_c - \omega_t$ determines the ratio between exciton and photon weights in the polaritonic states for each $k_{\|}$. The value of detuning is determined by the cavity thickness and, due to a wedge-shaped microcavity profile, can be varied in experiments.

The relaxation of polaritons into the ground state $\omega_L(0)$ is impeded due to the so-called bottleneck effect observed in photoluminescence experiments. Polaritons gather near the "knee" at the lower polariton dispersion curve just below the exciton energy (near the point A in **Figure 5**), since there the acoustic

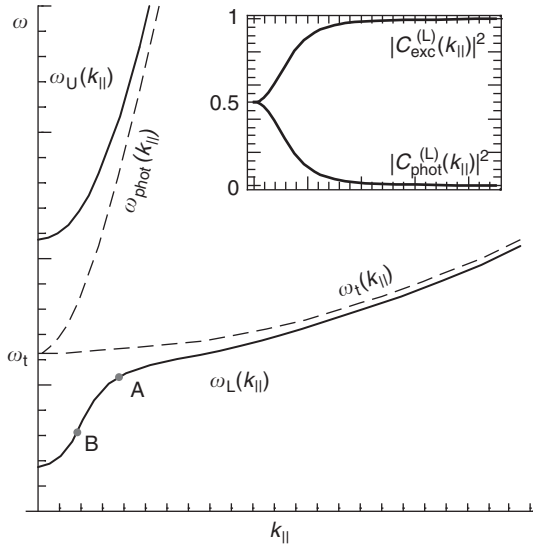


Figure 5 Dispersion of cavity polaritons (solid lines) for zero detuning. Dashed lines correspond to the dispersions of "bare" photon and "bare" exciton. The inset shows the dependence on the wave vector of the exciton and photon weight coefficients $|C_{exc}^{(L)}(k_{\|})|^2$ and $|C_{phot}^{(L)}(k_{\|})|^2$ defined in eqn [2] for the lower branch.

phonon-assisted relaxation is not effective due to an abrupt decrease of the density of final states. Qualitatively, at this point lower-branch polaritons change their character from exciton-like to photon-like, which essentially impedes the relaxation.

Much attention at the moment is devoted to the possibility of collecting a macroscopic number of polaritons in $\omega_L(0)$ state, exploiting the bosonic nature of cavity polaritons. Bosonic effects have manifested themselves in the experiments on stimulated scattering of cavity polaritons. In these experiments, lower-branch polaritons were excited resonantly near the inflection point of their dispersion curve (point B in **Figure 5**). A convex–concave character of the dispersion curve at this point allows the polariton–polariton scattering process, when one polariton is scattered into the state with $k_{\|} = 0$ ("signal"), and one is scattered to a higher-energy state ("idler"). A weak stimulation of $k_{\|} = 0$ state by a probe beam leads to a giant ($\sim$100-fold) amplification of the "signal" due to the fact that the scattering rate of bosons is proportional to $n_i(1 + n_f)$, where n_i and n_f denote, respectively, the populations of the initial and final states. It suggests a tempting possibility of Bose–Einstein condensation or Kosterlitz–Thouless transition to the superfluid phase of cavity polaritons.

See also: Excitons in Crystals; Interaction of Light and Matter; Lattice Dynamics: Vibrational Modes; Semiconductor Nanostructures.

PACS: 71.36.+c; 78.20.−e; 42.60.Da

Further Reading

Agranovich VM (1959) Dispersion of electromagnetic waves in crystals. *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* 37: 430. ((1960) *Sov. Phys. ZhETP* 10: 307).

Agranovich VM and Dubovsky OA (1966) Influence of retardation on spectrum of excitons in one- and two-dimensional crystals. *Pis'ma v Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* 3: 345. ((1966) *Journal of Experimental and Theoretical Physics (JETP) Letters* 3: 223).

Agranovich VM and Ginzburg VL (1984) *Crystal Optics with Spatial Dispersion, and Excitions*. Berlin: Springer.

Agranovich VM and Mal'shukov AG (1974) Surface polariton spectra if the resonance with the transition layer vibrations exist. *Opt. Commun.* 11: 169.

Agranovich VM and Mills DL (eds.) (1982) *Surface Polaritons*, Modern Problems in Condensed Matter Sciences. Amsterdam: North-Holland.

Bassani F (2003) Polaritons. In: Agranovich VM and Bassani GF (eds.) *Electronic Excitations in Nanostructures, Thin films and nanostructures*, vol. 31, pp. 129–184. Amsterdam: Elsevier.

Fano U (1956) Atomic theory of electromagnetic interactions in dense materials. *Physical Review* 103: 1202.

Hopfield JJ (1958) Theory of the contribution of excitons to complex dielectric constants of crystals. *Physical Review* 122: 1555.

Huang K (1951) On the interaction between the radiation field and ionic crystals. *Proceedings of the Royal Society of London Series A* 208: 352.

Kavokin A and Malpuech G (2003) Cavity polaritons. In: Agranovich VM and Taylor D (eds.) *Thin Films and Nanostructures*, vol. 32. Amsterdam: Elsevier.

Khitrova G, Gibbs HM, Jahnke F, Kira M, and Koch SW (1999) Nonlinear optics of normal-mode-coupling semicondcutor microcavities. *Reviews of Modern Physics* 71: 1591–1639.

Pekar SI (1957) The theory of electromagnetic waves in a crystal in which excitons are produced. *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* 33: 1022. ((1958) *Sov. Phys. ZhETP* 6: 785).

Tolpygo KB (1950) Physical properties of a lattice of rock salt type composed from deformable ions. *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* 20: 497 (in Russian).

Wesbuch C, Nishioka M, Ishikawa A, and Arakawa Y (1992) Observation of the coupled exciton–photon mode splitting in a semiconductor quantum microcavity. *Physical Review Letters* 69: 3314.

Yamamoto Y, Tassone F, and Cao H (2000) *Semiconductor Cavity Quantum Electrodynamics*. Berlin: Springer.

Polarizabilities

K Cho, Osaka University, Osaka, Japan

Introduction

Electromagnetic (EM) field in the presence of a matter is described by the Maxwell equations with the source terms of charge and current densities. One of the components of the current density is the time derivative of polarization, which plays a central role in the optical response of matter. The charge and current densities to be used in the Maxwell equations should be determined from the motion of the matter system in the EM field. This means that the motions of EM field and matter should, in principle, be determined self-consistently, as explicitly done usually in quantum electrodynamics.

A conventional way to avoid the tedious process of self-consistent solution was invented in the macroscopic theory of optical (or EM) response, where a polarizability constant was assigned to each material as a parameter, describing the induced polarization as a function of the incident EM field. This has allowed an easy access to the solution of the Maxwell equations in a matter system. Namely,

one looks for a solution in each part of the material with the same polarizability, and then, these solutions are connected with one another via the so-called Maxwell boundary conditions (the continuity of the normal or tangential components of EM field across a surface/interface). This determines the response of matter to an incident EM field uniquely. Another name also used for polarizability is "electric susceptibility."

Polarizability can be calculated quantum mechanically by evaluating the induced polarization from $\mathrm{Trace}\{\rho_{\mathrm{M}}\hat{\boldsymbol{P}}\}$ in the presence of an incident EM field, where ρ_{M} is the density matrix of the matter and $\hat{\boldsymbol{P}}$ the polarization (or dipole density) operator. In the semiclassical framework, an EM field is treated as a c-number, and the induced polarization defined above is usually chosen as a c-number of the matter side, to be related with the EM field. The optical processes to be covered by such a description include reflection, absorption, transmission, refraction, diffraction of macroscopic scale, and various nonlinear processes. However, in optical processes where the matter states are different before and after the process, such as Raman scattering or luminescence, one needs a more detailed treatment of induced polarization, such as the transition polarizability introduced by Born and Huang.

Macroscopic Polarizability

In macroscopic (local) theory of optical response, the polarizability relates a source electric field $\boldsymbol{E}(\boldsymbol{r},\omega)$ with the induced polarization $\boldsymbol{P}(\boldsymbol{r},\omega)$ of a matter as

$$\begin{aligned} P_\xi(\boldsymbol{r},\omega) = {} & \varepsilon_0 \sum_\eta \bar{\chi}^{(1)}_{\xi,\eta}(\omega) E_\eta(\boldsymbol{r},\omega) \\ & + \varepsilon_0 \sum_\eta \sum_\zeta \bar{\chi}^{(2)}_{\xi,\eta,\zeta}(\omega',\omega'') \\ & \times E_\eta(\boldsymbol{r},\omega') E_\zeta(\boldsymbol{r},\omega'')|_{\omega'+\omega''=\omega} \\ & + \cdots \end{aligned} \qquad [1]$$

where $\boldsymbol{r}$ and ω are position and frequency, respectively, (ξ,η) stand for the Cartesian coordinates, ε_0 the dielectric constant of vacuum, and $\bar{\chi}^{(N)}$ the polarizability of the Nth-order process. These $\{\bar{\chi}^{(N)}\}$ are tensor quantities because they give relationships between different vectors. $N=1$ corresponds to a linear process, and $N=2$, 3, etc. to the nonlinear ones of the second-, third-order, etc., respectively.

These polarizabilities are allowed to have dependence, not on position, but on frequencies, within the framework of macroscopic response. A typical way leading to the frequency dependence is the so-called Lorentz oscillator model, which regards matter as an assembly of charged oscillators. Using a model of "damped harmonic oscillator in a vibrating force" of classical mechanics, one can derive a simple expression of the frequency dependence of, for example, linear polarizability as

$$\bar{\chi}^{(1)}(\omega) = \frac{N_0 Q^2}{M} \frac{1}{\omega_0^2 - \omega^2 - \mathrm{i}\gamma\omega} \qquad [2]$$

where N_0, Q, M, ω_0, and γ are the number density, charge, mass, resonant frequency, and the damping parameter, respectively, of the oscillators. By appropriately choosing these parameters for a given material, one can rather well describe the ω-dependence of the optical spectrum. This is because the ω-dependence of $\bar{\chi}^{(1)}(\omega)$, apart from the parameter values, is essentially correct in comparison with the quantum mechanically calculated polarizability, as shown below. For a given material, there can be many different types of oscillators with different parameters.

The frequency ω_0 corresponds to the excitation energy of the matter, so that there are infinitely many of them in general. For an optical process near a resonance, it is often the case that this particular one is treated as a Lorentz oscillator explicitly and all the others as a frequency-independent background medium. Thus, the $\bar{\chi}^{(1)}$ to be used in such a treatment is the resonant term [2] plus a constant background polarizability χ_{b}. The idea of introducing background medium is a kind of "wisdom" to circumvent the tremendous complexity arising from the infinite number of freedom.

Microscopic Polarizability

Polarizabilities can be calculated quantum mechanically by evaluating the induced polarization as a functional of source electric field via a time-dependent perturbation calculation with respect to the light–matter interaction. If one does it properly, there arises a peculiarity of the microscopic response, namely, a nonlocal relationship between $\boldsymbol{P}$ and $\boldsymbol{E}$ as

$$\begin{aligned} P_\xi(\boldsymbol{r},\omega) = {} & \varepsilon_0 \sum_\eta \int \mathrm{d}\boldsymbol{r}' \chi^{(1)}_{\xi,\eta}(\boldsymbol{r},\boldsymbol{r}',\omega) E_\eta(\boldsymbol{r}',\omega) \\ & + \varepsilon_0 \sum_\eta \sum_\xi \int \mathrm{d}\boldsymbol{r}' \int \mathrm{d}\boldsymbol{r}'' \chi^{(2)}_{\xi,\eta,\zeta} \\ & \times (\boldsymbol{r},\boldsymbol{r}',\boldsymbol{r}'';\omega',\omega'') \\ & \times E_\eta(\boldsymbol{r}',\omega') E_\eta(\boldsymbol{r}'',\omega'')|_{\omega'+\omega''=\omega} \\ & + \cdots \end{aligned} \qquad [3]$$

The integral kernels $\chi^{(1)}$ and $\chi^{(n)}$ $(n=2,3,\ldots)$ are the nonlocal polarizabilities of the linear and the nth-order nonlinear processes, respectively. The

quantum mechanical expressions of, for example, the linear susceptibility at absolute zero temperature is given as

$$\chi^{(1)}_{\xi,\eta}(\boldsymbol{r},\boldsymbol{r}',\omega) = \sum_{v}\left\{\frac{\langle 0|\hat{P}_{\xi}(\boldsymbol{r})|v\rangle\langle v|\hat{P}_{\eta}(\boldsymbol{r}')|0\rangle}{E_v - E_0 - \hbar\omega - \mathrm{i}\gamma} + \frac{\langle 0|\hat{P}_{\eta}(\boldsymbol{r}')|v\rangle\langle v|\hat{P}_{\xi}(\boldsymbol{r})|0\rangle}{E_v - E_0 + \hbar\omega + \mathrm{i}\gamma}\right\} \quad [4]$$

where v runs over all the excited states of the matter from the groundstate $|0\rangle$. The parameter γ is, for idealized systems, a positive infinitesimal (0^+) representing the adiabatic switching of light–matter interaction, and for realistic systems, it is the self-energy of the excitation $E_v - E_0$, caused by nonradiative scattering, and is often approximated as a phenomenological parameter. The effect of radiative damping appears automatically by solving the coupled equations for $\boldsymbol{E}$ and $\boldsymbol{P}$, and therefore should not be included in γ. Otherwise, one would make a double counting of the interaction. The introduction of the background polarizability can be done similarly as before, by specifying the resonant frequency of interest and the rest.

Local versus Nonlocal (Macro- vs. Microscopic) Description

The nonlocal relationship between the polarization and electric field in the above expression is due to the coherence of the relevant wave functions and also to the intention to consider the spatial structure down to the subatomic scale. In the macroscopic response theory, it is approximated to be a local one as in eqn [1]. The simplified scheme of local response has been largely successful in calculating the EM or optical response of bulk matter. The validity condition for this approximation has not been discussed in detail.

A frequently encountered description for the macroscopic local theory is that it is a scheme for "the field components averaged over a region which is much larger than atomic size but much smaller than light wavelength." This means that a point in macroscopic theory has a volume satisfying the condition in the quotation marks. By this coarse graining, one loses certain information, but gains a simplified mathematical scheme. Thus, in order to judge whether or not a local description is all right for a given problem, one has to check two conditions: (1) if one can define a macroscopic "point" in the system of interest, and then (2) if the lost information by averaging is serious or not. For most nonresonant processes, local approach will be all right, because the long-wavelength approximation (LWA) is valid. For resonant processes, a nonlocal description should generally be chosen. But the case of excitations localized in a region much smaller than the macroscopic point mentioned above can be treated locally. In this case, the nonlocal polarizability [4] can be reduced to a local one [2] through the replacement

$$\hbar\omega_0 \to E_{v0}, \qquad \frac{Q^2}{M} \to \frac{2E_{v0}}{\hbar}|M_{v0}|^2 \qquad [5]$$

where M_{v0} is the electric dipole moment of the transition with the excitation energy E_{v0}.

The characteristic point of the local theory is that only electric dipole transitions can contribute to the polarization because of its essential assumption of LWA. In the nonlocal theory, there is also a possibility for other types of transitions to contribute because it takes the spatial variation of the EM field in a self-consistent way. Especially in the case of nanostructured materials, this possibility can lead to remarkable effects with respect to the size and frequency dependence of nonlinear responses. This is a reflection of the coherence of size-quantized excited states of matter, which are the counterparts of their size-quantized energy levels. In this sense, the nonlocal polarizability plays an important role for nanoscale materials.

In a microscopic description, polarizability is not a quantity to facilitate the calculation of optical response as in the macroscopic theory. The practical way of calculating optical response is rather different in the microscopic theory. Namely, it is not the polarizability itself but its components such as the matrix elements of dipole density and energy denominators that appear in the description of optical response, as shown below.

The main difference between the macroscopic and microscopic response theories lies in whether or not the boundary condition for an EM field is required. The reason why it is required in the macroscopic theory is that a medium is defined in terms of a set of position-independent polarizabilities, and that the boundaries of various optical media are introduced separately.

On the other hand, the microscopic polarizabilities include the position dependence reflecting the boundaries of matter through the boundary conditions for electrons and nuclei. Since the boundary conditions are already taken into account in the quantum mechanics of matter, it is no more necessary to consider the boundary conditions again for the EM field. The way of obtaining an optical response is the self-consistent solution of the two relations between polarization and EM field obtained from the Maxwell equations and Schrödinger equation, as shown below.

Determination of Microscopic Response

If all the matter states are included in the polarizabilities, the electric field produced by the polarization is obtained from the Maxwell equations as

$$E_\xi(\boldsymbol{r},\omega) = E_{0,\xi}(\boldsymbol{r},\omega) + \sum_\eta \int \mathrm{d}\boldsymbol{r}' G_{\xi,\eta}(\boldsymbol{r},\boldsymbol{r}',\omega) P_\eta(\boldsymbol{r}',\omega) \quad [6]$$

where $\boldsymbol{E}_0$ is an incident field, and $G_{\xi,\eta}(\boldsymbol{r},\boldsymbol{r}',\omega)$ is the (ξ,η)-tensor component of the radiation Green's function of vacuum, which satisfies the differential equation

$$\left(\nabla\times\nabla + \frac{\omega^2}{c^2}\right)\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega) = \mathbf{1}\,\delta(\boldsymbol{r}-\boldsymbol{r}') \quad [7]$$

The self-consistent solution of the EM field and the induced polarization is obtained by solving [3] and [6] simultaneously. For this purpose, one introduces the new variables

$$F_{\mu\nu}(\omega) = \int \mathrm{d}\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r},\omega)\cdot\langle\mu|\hat{\boldsymbol{P}}(\boldsymbol{r})|\nu\rangle \quad [8]$$

The induced polarization $\boldsymbol{P}$ is written as a polynomial series of these variables, which together with eqn [6] gives $\boldsymbol{E}$ also in a polynomial series of these $\{F\}$'s. Substituting the latter into eqn [8], one obtains a set of polynomial equations to determine $\{F\}$'s. For a given order of optical process, linear or Nth order nonlinear one, the solution of these simultaneous polynomial equations determines the response for a given incident field $\boldsymbol{E}_0$.

For linear process, one has a set of linear equations for $\{F\}$'s. The condition for the existence of nontrivial solution in the absence of incident field $\boldsymbol{E}_0$ is given as the vanishing of the determinant of the coefficient matrix, $\det|\boldsymbol{S}| = 0$. This is the equation for the eigenmode of the light–matter coupled system. In the rotating wave approximation, where only the first term on the right-hand side of eqn [4] is retained, one has

$$S_{\mu\nu} = (E_\mu - \hbar\omega)\delta_{\mu\nu} + A_{\mu 0,0\nu}(\omega) \quad [9]$$

The quantity $A_{\mu 0,0\nu}(\omega)$ represents the interaction energy between the induced polarization components via a transverse EM field (vacuum photons) with frequency ω as

$$A_{\mu 0,0\nu}(\omega) = -\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\langle\mu|\hat{\boldsymbol{P}}(\boldsymbol{r})|0\rangle \cdot \boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\langle 0|\hat{\boldsymbol{P}}(\boldsymbol{r})|\nu\rangle \quad [10]$$

which is a complex quantity. Since the eigenmode of the coupled system gives the resonant structure in an optical spectrum, the real and imaginary parts of $A_{\mu 0,0\nu}(\omega)$ represent the radiative shift and width of the matter excitation energy, if its frequency dependence is negligible as usually is the case. In this sense, the eigenmodes in the regime of nonlocal response are extended Lorentz oscillators, that is, a matter is an assembly of oscillating dipole densities with finite spatial extension and radiative shift and width, which can be defined precisely on a quantum mechanical basis. In atomic systems, this radiative correction is very small, but in nanoscale systems and in the bulk, this correction can be comparable to or even exceeds the level spacing of matter. For example, the difference between the dispersion curves of an exciton and polariton is just the radiative shift defined above.

Different Definitions of χ and Corresponding Matter Hamiltonians

The definition of polarizability depends on what the polarizing field is. When light is incident on an isolated matter, it polarizes the matter, and the induced oscillating polarization emits an EM field with the same frequency, and this EM field polarizes the matter further, and so on. Even if a polarizing field is purely transverse (T), the induced polarization can generally contain both longitudinal (L) and T components. Both of them act as the source term in the Maxwell equations, leading to the Maxwell field with both T and L components. Since the L component of the Maxwell field is caused by the induced charge density, the interaction of this L field with the matter polarization is a part of the Coulomb interaction among the charged particles, that is, the Coulomb interaction among the induced charge densities H_{cc}. ("Electron–hole exchange interaction" can also be rewritten in such a form.) Namely, the LT splitting of polarization waves in a crystal may be expressed as a part of matter excitation energy, but it could also be ascribed to the interaction energy of a polarization wave with an induced depolarization field.

There are two points of view about the treatment of H_{cc}, that is, as a part of (A) matter Hamiltonian, or (B) the interaction between matter and external field. For each viewpoint, there is a suitable set of "matter Hamiltonian, an external EM field, the interaction between them, and polarizability." Namely, H_{A} should contain the full Coulomb interaction, while H_{cc}, a part of the Coulomb interaction, should be excluded from H_{B}, that is, $H_{\mathrm{B}} = H_{\mathrm{A}} - H_{\mathrm{cc}}$. The fields inducing the polarization should be chosen as the full Maxwell field $\boldsymbol{E}$ in scheme (B), while it is only the T component of $\boldsymbol{E}$ in scheme (A). The linear polarizability in each case is given in the same form as eqn [4], but the energy eigenvalues and the matrix elements of polarization operator are different

because of the difference in the matter Hamiltonian. One can show the equivalence of these two points of view, at least for linear response. Depending on the problem to be considered, one can choose either of them with proper consideration of the matter Hamiltonian to be used.

As a simple example, let the electronic excitations in a crystal be considered, that is, the case where a wave vector is a good quantum number. Neglecting the effect of finite reciprocal wave vectors (Umklapp processes), one can treat $\chi^{(1)}$ as a function of $\boldsymbol{r}-\boldsymbol{r}'$ alone. Its Fourier transform $\chi^{(1)}(\boldsymbol{k},\omega)$ has a different relationship, depending on the scheme (A) and (B), with dielectric function defined as $\boldsymbol{D}(\boldsymbol{k},\omega)=\varepsilon_0\varepsilon(\boldsymbol{k},\omega)\boldsymbol{E}(\boldsymbol{k},\omega)$. From $\nabla\cdot\boldsymbol{D}=0$ and $\boldsymbol{P}=\varepsilon_0\chi_{\mathrm{A}}^{(1)}\boldsymbol{E}_{\mathrm{T}}=\varepsilon_0\chi_{\mathrm{B}}^{(1)}\boldsymbol{E}$, one can derive

$$\varepsilon=1+\chi_{\mathrm{B}}^{(1)} \qquad [11]$$

and

$$\chi_{\mathrm{B}}^{(1)}=\chi_{\mathrm{A}}^{(1)}+\chi_{\mathrm{A}}^{(1)}\cdot\frac{\boldsymbol{k}\boldsymbol{k}}{k^2}\cdot\chi_{\mathrm{B}}^{(1)} \qquad [12]$$

The excitation energies of the matter Hamiltonian appear as the poles of $\chi^{(1)}$. Therefore, all the matter excitation energies of both T and L characters in the scheme (A) appear as the poles of $\chi_{\mathrm{A}}^{(1)}$. The matter Hamiltonian in the scheme (B) does not contain H_{cc}, so that its T and L levels are degenerate. The interaction with the external L field lifts the degeneracy and gives an additional energy to the L state alone, which finally leads to the same resonant positions of the T and L levels of the light–matter coupled system. This situation in the scheme (B) is reflected in the above relationship, namely, the L mode excitations appear as the zeros of the dielectric function $\varepsilon=1+\chi_{\mathrm{B}}^{(1)}$, and the T mode excitations as the poles of $\chi_{\mathrm{B}}^{(1)}$.

The above relationship between $\chi_{\mathrm{A}}^{(1)}$ and $\chi_{\mathrm{B}}^{(1)}$ can be generalized to the case without translational symmetry. If one denotes the $(\boldsymbol{k},\boldsymbol{k}')$-Fourier components of the linear polarizabilities as $\chi_{\mathrm{A}}^{(1)}(\boldsymbol{k},\boldsymbol{k}';\omega)$ and $\chi_{\mathrm{B}}^{(1)}(\boldsymbol{k},\boldsymbol{k}';\omega)$, respectively, their relationship is generally given as

$$\chi_{\mathrm{B}}^{(1)}(\boldsymbol{k},\boldsymbol{k}')=\chi_{\mathrm{A}}^{(1)}(\boldsymbol{k},\boldsymbol{k}')+\frac{1}{2\pi^2}\int \mathrm{d}\boldsymbol{q}\,\chi_{\mathrm{A}}^{(1)}(\boldsymbol{k},\boldsymbol{q})\cdot\frac{\boldsymbol{q}\boldsymbol{q}}{q^2}\cdot\chi_{\mathrm{B}}^{(1)}(\boldsymbol{q},\boldsymbol{k}') \qquad [13]$$

with the omission of the common variable ω.

Polariton Dispersion Equation

Polarizability plays an essential role in determining the eigenmodes of light–matter coupled systems. Among various examples of such an eigenmode, the case of an exciton in a three-dimensional crystal is taken. The eigenmode of the coupled light–matter system is an exciton–polariton, whose dispersion equation is obtained from the coupled linear equations

$$0=\boldsymbol{k}\times\boldsymbol{k}\times\boldsymbol{E}(\boldsymbol{k},\omega)+\frac{\omega^2}{c^2}(1+\chi_{\mathrm{B}}^{(1)})\boldsymbol{E}(\boldsymbol{k},\omega) \qquad [14]$$

If symmetry allows pure T modes, this leads to the polariton dispersion equation $1+\chi_{\mathrm{B}}^{(1)}=(ck/\omega)^2$. A more general dispersion equation for crystal excitations is obtained from the eigenmode equation $\det|S|=0$ mentioned in connection with eqn [9]. In this generalized scheme, the Umklapp effect can easily be included through the Fourier components of the induced microscopic polarization. This effect is important for the X-ray propagation in a crystal in the dynamical scattering regime. The generalized dispersion equation has the form

$$0=\det|\{|\boldsymbol{k}+\boldsymbol{g}|^2-q^2\}\delta_{i,j}\delta_{g,g'}-q^2\hat{\boldsymbol{e}}_i(\boldsymbol{k}+\boldsymbol{g})\cdot\chi_{\mathrm{A}}^{(1)}\cdot\hat{\boldsymbol{e}}_j(\boldsymbol{k}+\boldsymbol{g}')| \qquad [15]$$

where $q=\omega/c$, $\{\boldsymbol{g}\}$ are reciprocal lattice vectors, and $\{\hat{\boldsymbol{e}}_i(\boldsymbol{k}+\boldsymbol{g});\ i=1,2\}$ are the mutually orthogonal unit vectors perpendicular to $\boldsymbol{k}+\boldsymbol{g}$. The polarizability $\chi_{\mathrm{A}}^{(1)}$ is the $\{\boldsymbol{k}+\boldsymbol{g},\boldsymbol{k}+\boldsymbol{g}'\}$ Fourier components of the polarizability of the scheme (A) mentioned before, which can contain all the resonances of the crystal.

If one neglects the matrix elements with subscripts of finite reciprocal lattice vectors in eqn [15], the reduced equation can be shown to be equivalent to the general dispersion equation obtained from eqn [14], which allows the LT mixed mode polaritons.

From the fact that the generalized dispersion equation [15] contains the effect of both resonance and Bragg scattering, it can be used in diverse situations such as polaritons, X-ray dynamical scattering, and photonic crystals in extended frameworks.

Oscillator Strength and Its Size Dependence

There is a quantity closely related with polarizability, namely, oscillator strength, which characterizes the coupling strength of a matter excitation with an EM field. It is defined as

$$f_{v0}=\frac{2mE_{v0}}{\hbar^2e^2}|M_{v0}|^2\cos^2\theta \qquad [16]$$

where θ is the angle between the transition dipole moment vector and light polarization direction. This factor is proportional to the numerator of the resonant polarizability in eqn [2], as shown in [5]. Namely, this is a useful concept within LWA. In nanoscale systems, there can be a quantum state

extending over the whole sample region, which may exhibit size-linear enhancement of its oscillator strength. As a sample size approaches the wavelength of resonant light from below, this size enhancement tends to saturate. This enhancement is not reflected on the linear polarizability, because induced (linear or nonlinear) polarization is defined as the induced dipole moment per unit volume. Following this definition, however, it has an effect on nonlinear polarizability. For example, $\chi^{(3)}$ for macroscopic response can exhibit size-linear enhancement, which is of course valid only in LWA.

Beyond the size region of LWA, oscillator strength is no more an appropriate measure of the coupling strength. As a more general measure, one could use the radiative width of each excitation, that is, the imaginary part of $A_{v0,0v}$ defined in eqn [10]. This quantity can be used irrespective of LWA. Since it is proportional to $|M_{v0}|^2$, that is, to oscillator strength, in LWA, it can be used more generally than oscillator strength.

The size dependence of the radiative width is very sensitive to the spatial structure of size-quantized excitations. Following the size evolution of a given excitation, one can observe, not only the size-linear enhancement in an LWA regime, but also size-resonant enhancement beyond the LWA. For each type of excitation, there is a characteristic size where the coupling with an EM field becomes maximal.

If the coupling with an EM field becomes very large as, for example, in a resonant Bragg reflector consisting of, for example, regular arrays of quantum wells (QWs) with electronic resonances, where the resonant light satisfies the Bragg condition, it is no more appropriate to neglect the frequency dependence of radiative correction, eqn [10]. In this situation, the coupling of the light–matter system is so strong that it is no more valid to regard its eigenmode as a matter oscillator corrected with radiative shift and width. In the above example, there is formation of photonic bands with a gap mode branch by increasing the number of QWs.

Cancellation Problem in Evaluating Nonlinear Polarizabilities

For the calculation of polarizabilities in eqn [3], the time-dependent perturbation theory with light–matter interaction is used:

$$H_{\text{int}} = -\int \mathrm{d}\boldsymbol{r}\, \hat{\boldsymbol{P}}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r}, t) \qquad [17]$$

as the perturbation Hamiltonian. The nth-order nonlinear polarizability is expressed as the n-fold time integral of the n-fold commutator of polarization operators in interaction representation. For an explicit calculation, one expands the n-fold commutator into the sum of various products of the polarization operators, and carries out the time integrations. When one sums up these integrated terms, there occurs partial or total cancellation among them for n larger than or equal to 2. A typical example is the polarization made of pure bosons such as harmonic oscillators. In this case, the innermost commutator of two polarization operators in the interaction representation at different times becomes a c-number, so that its commutator with another operator turns out to be zero. This means that all the nonlinear polarizabilities should be zero for this model. However, each of the contributions from the expanded terms of the n-fold commutator is generally nonvanishing, which means that the complete cancellation occurs in summing up these expanded terms. For general cases of light–matter interaction, matter excitations are not pure bosons, and there is interaction among them, which leads to finite nonlinear polarizability. But there always exists a partial cancellation in summing up the expanded terms.

This feature is important when one considers the size enhancement of nonlinear polarizabilities for a nanoscale matter. As mentioned above, this enhancement tends to saturate as sample size grows to be comparable to the resonant wavelength. However, each of the expanded terms of the n-fold commutator is kept enhanced also beyond LWA. The saturation occurs only through the cancellation among such terms. This means that beyond LWA the leading order terms are cancelled out, leaving the next order terms as a finite contribution. Therefore, it is essential to make a very careful treatment of cancellation in the calculation of nonlinear polarizabilities. Otherwise, one would be left with a large error in the values.

See also: Optical Sum Rules and Kramers–Kronig Relations; Polaritons; Polarons.

PACS: 78; 78.67. – n

Further Reading

Agranovich VM and Ginzburg VL (1984) *Crystal Optics and Spatial Dispersion, and Excitons*. Berlin: Springer.

Born M and Huang K (1954) *Dynamical Theory of Crystal Lattices*. Oxford: Clarendon Press.

Cho K (2003) *Optical Response of Nanostructures – Microscopic Nonlocal Theory*. Berlin: Springer.

Halevi P(ed.) (1992) *Spatial Dispersion in Solids and Plasmas*, pp. 339. Amsterdam: North-Holland.

Jackson JD (1998) *Classical Electrodynamics*, 3rd edn. New York: Wiley.

Keller O (1996) Local fields in the electrodynamics of mesoscopic media. *Physics Report* 268: 2, 3 (Elsevier).

Landau LD and Lifschitz EM (1963) *Electrodynamics of Continuous Media*. Oxford: Pergamon. (Translated by Sykes JB and Bel JS.)
Pekar SI and Kocherga OD (1983) *Crystal Optics and Additional Light Waves*. New York: Benjamin/Cummings.
Shen YR (1984) *The Principles of Nonlinear Optics*. Wiley.
Stahl A and Balslev I (1987) *Electrodynamics of Semiconductor Band Edge*. Berlin: Springer.
Toyozawa Y, (1980) In: Nakajima S, Toyozawa Y, and Abe R (eds.) *The Physics of Elementary Excitations*, ch.2. Berlin: Springer.
van Vleck JH (1932) *Theory of Electric and Magnetic Susceptibilities*. Oxford: Oxford University Press.

Nomenclature

$A_{\mu 0,0\nu}(\omega)$	radiative interaction between induced polarizations
$\boldsymbol{D}$	displacement field
$\{\hat{\boldsymbol{e}}_i(\boldsymbol{k}+\mathbf{g})$; i-1,2\}	mutually orthogonal unit vectors perpendicular to $\boldsymbol{k}+\boldsymbol{g}$
$\boldsymbol{E}$	(Maxwell) electric field
$\boldsymbol{E}_0$	incident electric field
$\boldsymbol{E}_{\mathrm{T}}$	transverse component of Maxwell field
E_v	energy eigenvalue of matter, with $v=0$ for ground state
E_{v0}	excitation energy of transition $\|v\rangle \leftarrow \|0\rangle$
f_{v0}	oscillator strength of transition $\|v\rangle \leftarrow \|0\rangle$
$F_{\mu\nu}$	expansion coefficient of induced polarization
$\boldsymbol{g}$	reciprocal lattice vector
$\boldsymbol{G}(\boldsymbol{r},\boldsymbol{r}',\omega)$	radiation Green function in vacuum
H_{A}	matter Hamiltonian for the scheme (A)
H_{B}	matter Hamiltonian for the scheme (B)
H_{cc}	Coulomb interaction among induced charge densities
H_{int}	interaction Hamiltonian between light and matter
$\boldsymbol{M}_{v0}$	transition dipole moment for $\|v\rangle \leftarrow \|0\rangle$
$N_0, Q, M, \omega_0, \gamma$	number density, charge, mass, resonant frequency, and damping parameter, respectively, of Lorentz oscillator
$\hat{\boldsymbol{P}}(\boldsymbol{r})$	polarization (dipole density) operator
$q=\omega/c$	wave number in vacuum
$\boldsymbol{r}$	position vector
$S_{\mu\nu}$	coefficient matrix to determine $F_{\mu\nu}$
ε_0	dielectric constant of vacuum
$\varepsilon(\boldsymbol{k},\omega)$	dielectric function for a crystal
θ	angle between $\boldsymbol{M}_{v0}$ and light polarization direction
ξ, η, ζ	running indices for Cartesian coordinates
ρ_{M}	density matrix of matter
$\chi_{\mathrm{A}}^{(1)}$	linear polarizability defined with respect to the transverse component of Maxwell field
$\chi_{\mathrm{B}}^{(1)}$	linear polarizability defined with respect to Maxwell field
χ_{b}	background polarizability
$\chi^{(n)}$	nth-order nonlocal nonlinear polarizability ($n=2, 3, \ldots$)
$\bar{\chi}^{(N)}$	local polarizability of the Nth order ($N=1, 2, \ldots$)
ω	frequency

Polarons

E I Rashba, Harvard University, Cambridge, MA, USA

Introduction

Charge carriers, electrons and holes, moving across a semiconductor or an insulator experience interaction with phonons. This interaction has two major effects. First, it causes electron scattering and, therefore, electrical resistance. Second, it changes the properties of electrons. Indeed, an electron deforms the lattice around it. This deformation can be described as a cloud of virtual phonons. When the electron moves, the cloud follows it. Such an electron dressed by a cloud of virtual phonons is called a polaron. The polaron concept, as applied to an electron coupled to optical phonons in a polar crystal, was introduced by Pekar in 1946. It is currently applied to electrons (holes) coupled to various types of phonons and to magnons. This concept has also been generalized as applied to electrons coupled to the fluctuations of the chemical content or the order parameter. The properties of polarons depend on the type of the phonons to which the electron is coupled, on the coupling strength, and also on the dimensionality, D, of the space. In the strong coupling regime, the deformation inside a polaron is usually large and its ability to move is suppressed. This limit corresponds to the Landau's concept of the self-trapping of electrons in a perfect lattice. Both terms, polaron and self-trapping, will be used in what follows. Similar to electrons, excitons are also coupled to phonons and dressed by them; such dressed excitons are sometimes called "excitonic polarons polaritons." While electron and hole polarons are investigated mostly through transport phenomena, excitonic polarons are observed primarily by spectroscopy. However, transport of excitonic polarons has also been investigated, especially in alkali halides.

Basic Classification of Polarons

Depending on the number N of phonons in the cloud, polarons are termed strong-coupling or weak-coupling polarons, for $N \gg 1$ and $N \ll 1$, respectively. Polarons are termed small polarons when their size r_p is close to the lattice constant a, $r_p \approx a$, and large polarons when $r_p \gg a$. Depending on the ratio of the half-width W of the electron band to the characteristic phonon frequency $\hbar\omega_{ph}$, polarons are termed adiabatic (when $W \gg \hbar\omega_{ph}$) and antiadiabatic (when $W \ll \hbar\omega_{ph}$). Free electrons and holes are always dressed by the electronic polarization of closed atomic (ionic) shells. When they move slowly, this polarization follows them nearly without retardation; therefore, it is included in the band energy of free particles. To emphasize this dressing, a term "electronic polaron" is sometimes used: "electronic polarons" are a specific type of antiadiabatic polarons.

In what follows, mostly adiabatic polarons are considered.

Polarons can also be classified by the electron–phonon coupling mechanisms. The standard Hamiltonian of this coupling is

$$H_{e-ph} = \frac{1}{N^{1/2}} \sum_{kq} \gamma_q a^+_{k+q/2} a_{k-q/2} (b_q + b^+_{-q})$$

where a_k and b_q are electron and phonon operators, respectively, $\boldsymbol{k}$ and $\boldsymbol{q}$ are their quasimomenta, and N is the number of unit cells in the normalization volume. Various mechanisms of electron–phonon coupling are discerned by the momentum dependence of the coupling constant γ_q and the phonon frequency ω_q. For the polarization coupling to optical phonons, $\omega_q = \omega_0 = \text{const.}$ and

$$\gamma_q = \frac{\sqrt{\alpha}\hbar}{q} \left(\frac{2\pi\omega_0}{v} \sqrt{\frac{2\hbar\omega_0}{m}} \right)^{1/2}$$

$$\alpha = \frac{e^2}{\hbar\kappa} \sqrt{\frac{m}{2\hbar\omega_0}},$$

$$\frac{1}{\kappa} = \frac{1}{\kappa_\infty} - \frac{1}{\kappa_0}$$

where e is the elementary charge, m is the electron effective mass, v is the unit cell volume, κ_0 and κ_∞ are the low- and high-frequency dielectric constants, respectively, and α is the Fröhlich coupling constant. In the equation for the Pekar factor $1/\kappa$, taking difference of inverse dielectric constants eliminates the fast "electronic polaron" polarization of atomic shells and retains only the polarization due to slow ionic displacements. Because of the factor q in the denominator of γ_q, the coupling to long-wavelength polar phonons, $qa \ll 1$, dominates. For nonpolar coupling to optical phonons $\gamma_q = \text{const.}$, $\omega_q = \text{const.}$, for the deformation potential coupling to acoustic phonons $\omega_q \propto q$, $\gamma_q \propto q^{1/2}$, and for piezoelectric coupling to acoustic phonons $\omega_q \propto q$, $\gamma_q \propto q^{-1/2}$.

Pekar–Fröhlich Polaron

In a Pekar–Fröhlich polaron, an electron is coupled to dispersionless optical phonons by polarization interaction. Because of the role of long-wavelength phonons, the theory is outlined in terms of an electron coupled to a dielectric continuum. This model can be solved in both limits: weak, $\alpha \ll 1$, and strong, $\alpha \gg 1$, coupling. Dressing of the electron renormalizes its spectrum; in particular, it lowers the bottom of the band and increases the effective mass. The effect is most drastic for $\alpha \gg 1$. In this limit, lattice polarization is nearly static and produces a potential well for an electron. The ground state for an electron in this well can be found from a nonlinear integral–differential Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \psi_0(r) + V[\psi_0, r]\psi_0(r) = E_0 \psi_0(r)$$

$$V[\psi_0, r] = -\frac{e^2}{\kappa} \int \frac{\psi_0^2(r')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, d\boldsymbol{r}'$$

This equation is self-consistent: an electron that is self-trapped in the ψ_0-state produces a dielectric polarization that supports, by its Coulomb field, this quantum state. It can be derived as a minimum of a functional

$$\mathcal{H}[\psi] = \min_\Psi \langle \psi | - (\hbar^2/2m)\nabla^2 + \tfrac{1}{2} V[\psi, r] | \psi \rangle$$
$$= K + V/2$$

K and V being the mean values of the kinetic energy and the energy of electron–phonon interaction. The factor 1/2 appears because the Schrödinger equation is nonlinear, and $\mathcal{H}[\psi]$ equals the total energy of the system, including the electron energy E_0 and the polarization energy $E_P = -V/2$, $\mathcal{H} = E_0 + E_P$. A scaling transformation $\psi_v(r) = v^{3/2}\psi_0(vr)$ results in

$$\mathcal{H}[\psi_0] = \min_v \{v^2 K + vV/2\}$$

Because ψ_0 is the exact ground state, the minimum of $\mathcal{H}[\psi_v]$ is achieved at $v = 1$. This equation relates K to V and leads to Pekar's virial theorem

$$K : E_P : |E_0| : |V| = 1 : 2 : 3 : 4$$

The eigenvalue $E_0 \approx -0.16 me^4/\kappa^2\hbar^2$ and the polaron radius $r_p \approx 3\kappa\hbar^2/me^2$ does not depend on ω_0. The number of phonons in the cloud $N = 2|E_0|/3\hbar\omega_0$. The strong coupling theory is applicable when $|E_0| \gg \hbar\omega_0$, and the continuum approximation when $r_p \gg a$.

The function $\psi_0(r)$ describes a strictly self-trapped electron. The translational symmetry of the problem and the finite phonon frequency ω_0 result in polaron motion: it acquires a finite effective mass m_p. Three phonon degrees of freedom are absorbed by the polaron and transformed into three zero modes of its translational motion. This results in the energy gain of $3\hbar\omega_0/2$ that, together with electron–phonon correlations, lowers the polaron energy as compared with $\mathscr{H}[\psi]$. Finally, the ground-state energy and the polaron mass are

$$E(k=0) \approx -(0.1085\alpha^2 + 2.836)\hbar\omega_0$$
$$m_p \approx 0.023\alpha^4 m$$

Landau–Pekar effective mass m_p increases steeply with α because a moving polaron drags a phonon cloud whose energy E_P grows with α, and m_p diverges when $\omega_0 \to 0$ because the number of the phonons in the cloud diverges, $N = E_P/\hbar\omega_0 \to \infty$.

In the weak coupling limit, $\alpha \ll 1$, the perturbation theory results in

$$E(k=0) \approx -\alpha\hbar\omega_0, \quad m_p \approx m/(1-\alpha/6)$$
$$N = \alpha/2$$

Because the numerical coefficients in the strong-coupling equations are small, the asymptotic expressions for $E(k=0)$ and m_p match at large α values, $\alpha \gtrsim 5$. Smooth matching can be achieved by different analytical intermediate-coupling methods, of which Feynman method has been the most widely used. More recently, numerical approaches based on the Monte Carlo method have been developed. Nevertheless, the basic physical insight into the polaron energy spectrum including its excited states is based on analytical results.

An estimate of α in terms of fundamental constants suggests that the strong coupling should be a typical one. With $m \approx m_0$, where m_0 is the electron mass in a vacuum, $\kappa \approx 1$, and $\hbar\omega_0$ estimated through the adiabatic parameter as $\hbar\omega_0/E_a \approx \sqrt{m_0/M}$, where $E_a \approx 10$ eV is an atomic scale energy and $M \approx 10^4 m_0$ is an atomic mass, one gets $\alpha \approx 10$. This estimate underscores the importance of the polaron effect in solid-state physics and chemistry of electrolytes. However, in many semiconductors α is small because of a small electron mass, $m \ll m_0$, and a large dielectric constant, $\kappa \gg 1$. For example, $\alpha \approx 0.02$ in InSb, 0.06 in GaAs, 0.6 in CdS, 1.8 in AgCl, and 4.1 in RbCl. Therefore, electron polarons are weakly coupled in most of the semiconductors and intermediately coupled in ionic insulators. In many crystals, hole polarons are coupled strongly; in these cases, they usually are small Jahn–Teller polarons.

Excited States of a Polaron

For $r \gtrsim r_p$, the potential $V[\psi_0, r]$ decreases by a Coulomb law, $V[\psi_0, r] \approx -e^2/\kappa r$. Therefore, excited states form a Rydberg-type series, their energies being scaled by $|E_0| \sim \alpha^2\omega_0$. These states have considerable widths due to multiphonon decay. According to the Franck–Condon principle, optical transition of an electron in a strong-coupling polaron can be found at a fixed lattice polarization. Such a polaron possesses an absorption spectrum that begins near the energy $\sim 3|E_0|/4$, consists of strongly overlapping bands, and is extended to high energies. While the absorption shape depends on a specific model, appearance of a peak ("polaron peak") instead of a Drude spectrum is a fingerprint of a strong-coupling polaron. A related but independent signature of a polaron is a Stokes shift caused by the polaron relaxation into a self-consistent excited state.

A different sequence of the excited states of a strong-coupling polaron has excitation energies below $\hbar\omega_0$; they cannot decay and are stable. These states develop due to the polarization of the $\psi_0(r)$-state by a fluctuating electric field of optical phonons and can be considered as polaron–phonon bound states or as local phonon modes near a polaron. A typical binding energy for these states is $\sim 0.1\hbar\omega_0$. In the $\alpha \to \infty$ limit, bound states form an infinite sequence converging to $\hbar\omega_0$. Practically, the sequence is cut by the magnitude of α and by phonon dispersion. The energy spectrum of a strong-coupling polaron is shown schematically in **Figure 1**.

For a weak-coupling polaron, the spectrum $E(k)$ has a termination point at $k_{th} \approx \sqrt{2m_p\omega_0/\hbar}$ that is the threshold for the emission of a real phonon by a fast polaron. The states with $E(k) > \hbar\omega_0$ have a finite width. For $\alpha \ll 1$, electron–phonon coupling

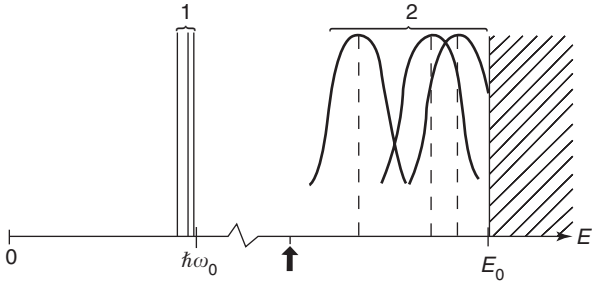


Figure 1 Schematics of the electron energy spectrum of a Pekar–Fröhlich polaron. The bottom of the polaron spectrum is chosen as an origin, polaron momentum $k = 0$. (1) polaron–phonon bound states with energies $E < \hbar\omega_0$. (2) high-energy excited states in the potential $V[\psi_0, r]$; their widths are shown schematically. The lower Stokes-shifted relaxed excited state is shown by the arrow. Two energy scales are used: $\hbar\omega_0$ on the left and E_0 on the right of the break in the E-axis. Free electron band is hatched.

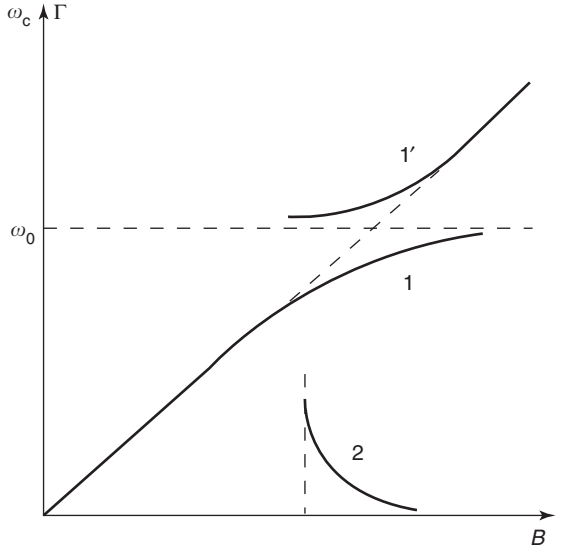


Figure 2 Dependence of the cyclotron resonance frequency $\omega_c(B)$ and the level width $\Gamma(B)$ on magnetic field B near the magneto-phonon resonance in the spectrum of a weak-coupling polaron. 1 and 1′ – the lower and upper branches of the spectrum, respectively; 2 – the width $\Gamma(B)$ of the upper branch.

manifests itself dramatically in a magneto-phonon resonance when the cyclotron frequency eB/m_pc of the polaron approaches ω_0, B being a magnetic field. The dependence $\omega_c(B)$ is shown in **Figure 2**. Near the resonance, the curve $\omega_c(B)$ splits into two branches. The lower branch is stable and is pinned to ω_0 when B increases. The upper branch is unstable because of the phonon emission, and the level width $\Gamma(B)$ reaches its maximum near the threshold. The effect is observed experimentally as the cyclotron resonance band splitting. This splitting is $\sim\alpha^{2/3}\omega_0$ near the resonance and allows direct measuring of α. The phenomenon results from the resonance enhancement of the electron–phonon coupling near the resonance that also manifests itself in the increase of N from $N\approx\alpha/2$ to $N\approx 1/2$.

Nonpolar Electron–Phonon Coupling

This problem applies to electrons and holes in nonpolar compounds, and also to excitons, because of their electrical neutrality, in any kind of crystals. In a continuum approximation, both mechanisms of nonpolar coupling (to acoustic and optical phonons) result in the same expression $V[\psi_0, r]\propto -\psi_0^2(r)$ in the strong coupling limit. The scaling transformation $\psi_v(r) = v^{D/2}\psi_0(vr)$ of a trial function $\psi_0(r)$, where D is dimensionality, leads to the functional $\mathscr{H}$:

$$\mathscr{H}[\psi_0] = \min_v \mathscr{H}(v), \quad \mathscr{H}(v) = v^2K + v^DV$$

where K and $V<0$ are kinetic and potential energies, respectively. The result depends on D critically.

$D = 1$, One-Dimensional Systems

With $D = 1$, $\mathscr{H}(v)$ reaches its minimum at a finite value of v, just as in the case of a three-dimensional (3D) large polaron. The self-trapped state can be found from a nonlinear Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_0}{dx^2} - 2c\psi_0^3 = E_0\psi_0$$

Its only normalizable solution is

$$\psi_0(\xi) = (1/\sqrt{2})\text{sech}\ \xi, \quad \xi = mcx/\hbar^2$$
$$E_0 = -mc^2/2\hbar^2$$

For coupling to nonpolar optical phonons of the frequency ω_0, the dimensionless coupling constant α is usually chosen in such a way that in the weak coupling regime the energy equals $E(k=0) = -\alpha\hbar\omega_0$. Then $E_0 \approx -\alpha^2\hbar\omega_0$, and

$$E(k=0)\approx -(\alpha^2/3 + 0.955)\hbar\omega_0$$
$$m_p \approx 32\alpha^4 m/15$$

As distinct from the large 3D polaron, there is no small numerical factors in $E(k=0)$ and m_p. Therefore, strong coupling sets in at $\alpha\gtrsim 1$. As a result, 1D geometry is favorable for large strong-coupling polarons and excitons. This is an important property of various quasi-1D-compounds, including polyacetylenes.

Similarly to a large 3D polaron, there exists an infinite sequence of stable excited states below $\hbar\omega_0$. However, as distinct from a large 3D polaron, the potential $V[\psi_0, x]$ does not support excited bound states above the $\hbar\omega_0$ threshold.

$D = 2$, Two-Dimensional Systems

Both terms in $\mathscr{H}(v)$ show the same dependence on v; hence, there exists a critical value of the coupling strength. Below it only a weak-coupling polaron exists, above it a self-trapped state is the polaron's ground state; its size is about a lattice spacing a.

$D = 3$, Three-Dimensional Systems: Holstein Polaron

There is no minimum at the curve $\mathscr{H}(v)$; hence, nonpolar coupling cannot support large strong-coupling polarons. Different options are shown in **Figure 3**.

If the coupling is strong enough, $\mathscr{H}(v)$ changes sign for a state whose size is still larger than the lattice constant (curve 1). In this case, a self-trapped state exists and is the polaron ground state. When the coupling strength decreases, the self-trapped state becomes metastable, its energy being higher than the

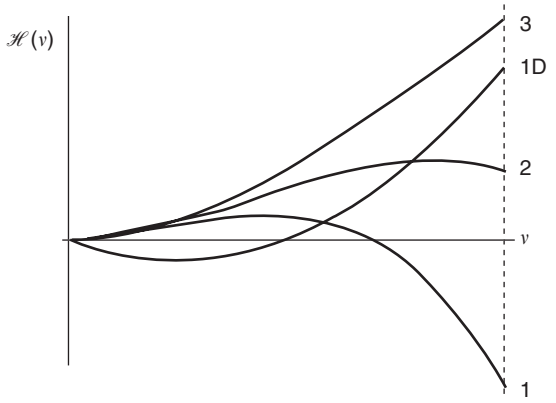


Figure 3 Different types of the dependence of the energy functional $\mathscr{H}(v)$ on v, the inverse size of a local state. 1 – strong coupling; a self-trapped state is the polaron ground state. 2 – intermediate coupling; a self-trapped state exists but is metastable. 3 – weak coupling; there is no self-trapped states. The dashed line indicates the size of the local 3D state of about a lattice constant a. For comparison, a $\mathscr{H}(v)$ curve for a strong-coupled 1D system is shown; position of its minimum indicates the size of a self-trapped state.

electron band bottom (curve 2). When coupling is weak, there is no self-trapped states (curve 3). The classification of states based on the functional $\mathscr{H}(v)$ is tantamount to the original Landau argument and is exact in the strictly adiabatic limit, $\omega_0 \to 0$. This argument says that, in a 3D quantum well, a bound state appears only when the well depth V and radius R obey the criterion $V \gtrsim \hbar^2/2mR^2$, and a self-trapped state is stable when the binding energy is large enough to ensure that the total energy of the system is lower than the bottom of the free-electron band.

Properties of small polarons depend on the crystal geometry and the details of chemical bond. Holstein model provides important insight into some of their general properties. In the lattice-site representation, the Holstein Hamiltonian reads

$$H = -t\sum_{n} a_n^+ a_{n+\delta} + \gamma \sum_{n} a_n^+ a_n (b_n + b_n^+) + \hbar\omega_0 \sum_{n} b_n^+ b_n$$

The first term is the electron kinetic energy, vectors $\boldsymbol{\delta}$ point to all nearest neighbors of the site $\boldsymbol{n}$. The second term describes intrasite electron coupling to dispersionless optical phonons through deformation potential. This Hamiltonian depends on two dimensionless constants. The ratio $\gamma/\hbar\omega_0$ measures the electron–phonon coupling strength, while the ratio $t/\hbar\omega_0$ controls the polaron evolution from the adiabatic ($t \gg \hbar\omega_0$) to antiadiabatic ($t \ll \hbar\omega_0$) limit.

For arbitrary parameter values, the problem can be solved only numerically. The results show convincingly, in a qualitative agreement with the adiabatic theory, that the switching between the weak- and strong-coupling regimes is smooth for $D = 1$ and rather sharp for $D = 3$. This is true for the dependence of both $E(k = 0)$ and m_p on $\gamma/\hbar\omega_0$. The changeover between the free and self-trapped electron regimes, which is abrupt in 3D in the adiabatic limit (and sometimes, rather improperly, termed a "phase transition"), is getting smoother when the ratio $\hbar\omega_0/t$ increases. This smoothening reflects a strong mixing of the free and self-trapped states when they are close to a resonance.

Holstein model is well suited for investigating the strong coupling regime when γ is large as compared with t and $\hbar\omega_0$, while the ratio $\hbar\omega_0/t$ can be arbitrary. Applying the Lang–Firsov transformation $\exp\{\gamma(b_n^+ - b_n)/\hbar\omega_0\}$ to all lattice sites, one eliminates the second term in H. The transformed Hamiltonian $\tilde{H}$ shows a general polaron shift of the spectrum by $-\gamma^2/\hbar\omega_0$ and a renormalization of the bandwidth, $t \to \tilde{t} = t\exp\{-(\gamma/\hbar\omega_0)^2\}$. After the transformation, the kinetic energy acquires factors dependent on phonon operators. The basic conclusions are as follows. Because the translational symmetry is not violated, the band spectrum is retained. However, the polaron band becomes exponentially narrow; therefore, band transport of polarons is easily destroyed by defects and thermal phonons. If the temperature T is not too low, the polaron transport is dominated by thermoactivated diffusion, $D \propto \exp(-E_a/k_B T)$ with $E_a \approx \gamma^2/\hbar\omega_0$, k_B being the Boltzmann constant. Activational T-dependence of the mobility is the basic transport fingerprint of small electronic and excitonic polarons.

It is a general property of polarons that their mass increases with the coupling constant by a power law in continuum models, and exponentially in lattice models.

The lattice models are especially useful for antiadiabatic polarons. Under these conditions the number of phonons in the system is nearly conserved. As a result, the lower part of the spectrum is described by the band of a dressed polaron, and the excited states by a few-particle theory as a polaron interacting with real phonons through a renormalized electron–phonon interaction. Such an approach is highly efficient in the spectroscopy of molecular excitons coupled to intra-molecular phonons where one distinguishes exciton–phonon bound states and a two-particle continuum.

Some Specific Types of Small Polarons

Strong-coupling polarons of the Holstein Hamiltonian are single-site polarons. Meantime, two-site hole polarons are typical of many solids. These are

quasimolecular ions of Xe_2^+ type in rare gas solids and of Cl_2^- type in alkali halides. While the van der Waals forces dominate in the cohesion energy of rare gas solids, such ions are bound chemically. As a result, atomic displacements are large in self-trapped states, and electron–phonon interaction is strongly nonlinear. Quasimolecular self-trapped excitons in these crystals have the same symmetry as hole polarons, with Xe_2-type self-trapped excitons mimicking excimers. Side-by-side with them, quasi-atomic self-trapped excitons usually also exist; however, they are metastable and decay into quasimolecular states. Quasimolecular hole polarons are a special case of Jahn–Teller polarons whose symmetry is broken because of the atomic level degeneracy.

Electron bubbles in liquid He and metal–ammonia solutions are different examples of nontrivial small polarons. Because of the negative work function of He, electrons injected into He produce vacuum bubbles and become self-trapped inside them. Holes in liquid He appear as positive ions surrounded by "icebergs" of solid phase.

Self-Consistency of the Adiabatic Concept of Self-Trapping

An important feature of the adiabatic concept of self-trapping is the coexistence of extended and localized states. The extended states are weakly coupled to phonons, while the localized states are strongly coupled to the same phonons. In order to prove the self-consistency of this concept, one has to check that the corrections to the energy of free states are really small, to be specific, small as compared with the self-trapping energy. The standard expression for the electron coupling constant to acoustic phonons is

$$\gamma_q = Cq\sqrt{\hbar/2\rho v\omega_q}, \quad \omega_q = sq$$

where C is the deformation potential, ρ is the density, v is the unit cell volume, and s is the sound velocity. Then the second-order correction $\delta_2 E(k=0)$ to the position of the electron band bottom and the energy of a strong-coupled single-site polaron E_p, with the accuracy to numerical factors, are

$$\delta_2 E(k=0) \approx -mC^2/\hbar\rho sa^2, \quad E_p \approx -C^2/\rho s^2 a^3$$

In both cases, the main contribution comes from short wavelength phonons. Their ratio

$$\delta_2 E(k=0)/E_p = msa/\hbar \approx \hbar\omega_D/W$$

ω_D being the Debye frequency. Therefore, this ratio equals the small parameter of the adiabatic theory of polarons. When W is of the atomic scale, $W \sim 10$ eV, this ratio is about $\sqrt{m/M}$, the basic parameter of the adiabatic theory of solids. Under the same conditions, $\delta_2 E(k=0) \sim \hbar\omega_D$ when C is of atomic scale.

A similar estimate for nonpolar coupling to optical phonons shows that $\delta_2 E(k=0) \approx -\gamma^2/W$, $E_p \approx -\gamma^2/\hbar\omega_0$, and their ratio $\delta_2 E(k=0)/E_p \approx \hbar\omega_0/W$.

Therefore, the coupling energies for nearly free and self-trapped states differ by the adiabatic parameter $\hbar\omega_{ph}/W$; hence, whenever it is small, the adiabatic theory of polarons is self-consistent.

Self-Trapping Barrier and Self-Trapping Rate

The self-consistency of the adiabatic theory suggests description of self-trapping in systems with wide electron bands, $W \gg \hbar\omega_{ph}$, in terms of the adiabatic potential surfaces $U(\ldots Q_i \ldots)$, where Q_i are phonon coordinates. With a single phonon coordinate symbolizing both the spatial extension and the magnitude of the lattice deformation, the dependence $U(Q)$ is shown in **Figure 4** for an electron (exciton) coupled to a 3D crystal lattice by a nonpolar interaction. The branching point where a local level first appears in the potential well produced by a growing lattice deformation is shown by BR; at this point the lower

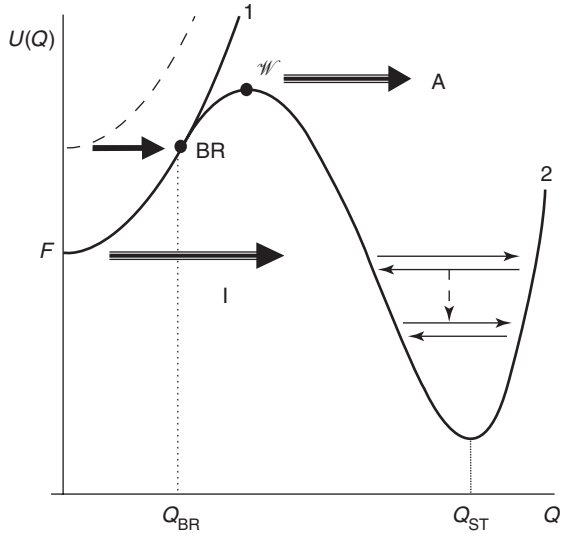


Figure 4 Adiabatic potential surfaces $U(Q)$ for a particle coupled to a 3D crystal by a nonpolar interaction. 1 – energy of the lattice deformation with a particle in a free state F at the bottom of the band. 2 – the split-off sheet of the adiabatic potential originating at the branching point BR. $\mathscr{W}$ shows the lowest saddle point of the sheet 2 that plays a role of the self-trapping barrier. Arrows I and A symbolize the instanton and activational (Arrhenius) self-trapping trajectories, respectively. A heavy arrow symbolizes self-trapping of a fast particle whose energy is shown by a dashed parabola. Oscillations around the point Q_{ST} in the self-trapped state and energy relaxation are also shown.

sheet 2 of the adiabatic potential surface splits off the free-particle surface 1. A distinctive feature of nonpolar coupling in 3D is the existence of the barrier for self-trapping shown by $\mathscr{W}$ ($\mathscr{W}$ is the lowest saddle point of the surface 2 in the multidimensional Q-space). Its energy equals

$$\mathscr{W}_{\mathrm{ac}} = 44(\hbar^3\rho s)^2/m^3C^4$$
$$\mathscr{W}_{\mathrm{op}} = 11\hbar^8/m^3(\gamma^2\omega_0 v)^2$$

for the coupling to acoustic and optical phonons, respectively. In both cases, the spatial extention of the barrier $r_{\mathrm{b}} \sim a(E_{\mathrm{p}}/W)$. Therefore, when coupling is strong, $E_{\mathrm{p}} \gg W$, the spatial extention of the barrier is large as compared with a, and the expressions for $\mathscr{W}$ in terms of macroscopic parameters are valid. The adiabatic theory is applicable when $\mathscr{W} \gg \hbar\omega_{\mathrm{ph}}$. Because $\mathscr{W}/\hbar\omega_{\mathrm{ph}} \sim (W/\hbar\omega_{\mathrm{ph}})(W/E_{\mathrm{p}})^2$, both criteria are compatible if the adiabatic parameter is large enough, $W/\hbar\omega_{\mathrm{ph}} \gg 1$. When $\mathscr{W} \lesssim \hbar\omega_{\mathrm{ph}}$, the notion of the barrier has no sense.

Free states are self-consistent solutions of the adiabatic problem, and the barrier protects free particles from fast self-trapping. Therefore, free and self-trapped states can coexist. Of course, free states are metastable, and self-trapping is sometimes considered as a decay of a "false vacuum." Self-trapping rate w is controlled by the penetration of the particle through the barrier. Thermoactivated tunneling and the activational (Arrhenius) mechanism dominate at $T < T_{\mathrm{c}}$ and $T > T_{\mathrm{c}}$, respectively. The critical temperature $T_{\mathrm{c}} \sim \hbar\omega_{\mathrm{ph}}/k_{\mathrm{B}}$, the numerical factor in it depends on specific properties of the system. In the framework of the semiclassical theory, thermoactivated tunneling can be considered as a collective tunneling of the lattice and the particle described technically as an instanton. Instanton is a notion adopted from the quantum field theory where it describes a trajectory of a classical field in imaginary time that possesses a finite Euclidean action. As applied to the self-trapping problem, an instanton begins with the tunneling of the free lattice that prepares a potential well for the particle, then the particle is trapped to a shallow energy level emerging in this well in the vicinity of the branching point, and the process is completed by joint tunneling of the coupled particle–lattice system. The probability $w(T) \propto \exp[-S(T)]$, where $S(T)$ is the semiclassical (Euclidean) action of the instanton that it acquires during its motion in imaginary time; this is a convenient way for describing the tunneling motion of the system under the barrier. The action decreases monotonically with T at $T < T_{\mathrm{c}}$ and equals $S(T) = \mathscr{W}/k_{\mathrm{B}}T$ in the Arrhenius region. Self-trapping through a $T = 0$ instanton and the Arrhenius process are shown schematically in **Figure 4**. When the particle has an initial energy E larger than $k_{\mathrm{B}}T$, which may be the case for excitons produced by light, this energy can contribute to w by reducing S. A coherent energy transfer from a particle to the tunneling lattice is suppressed by the small mass ratio; it is most efficient for $E \lesssim \mathscr{W}$.

After the instanton passes the barrier, it turns into a wave packet that oscillates between the turning points of the well with a frequency of a local mode and relaxes by emitting lattice phonons (**Figure 4**). For excitons, this dynamics can be detected by the oscillations and decay of light emission because it comes mostly from the turning points of the semiclassical lattice motion. If the energy of the packet is concentrated at a single or a few degrees of freedom, self-trapping results in lattice defect production.

Self-trapping of charged particles, electrons and holes, in ionic crystals into the Pekar–Fröhlich polaron states, and also self-trapping of electrons (holes) and excitons in 1D conductors, proceeds fast because there is no barrier for them. For example, (i) self-trapping of holes in alkali halides proceeds much faster than self-trapping of excitons, and (ii) oscillations and decay of the exciton emission from quasi-1D Pt complexes begin immediately after the excitation pulse.

Because of the atomic size of self-trapped states, calculating numerical values of their parameters, especially the self-trapping probability, is a formidable task. However, the basic qualitative predictions of the adiabatic theory are confirmed by extensive experimental data, especially as applied to excitons because they are detected optically. Coexistence of free and self-trapped states has been established in alkali halides, rare gas solids, and molecular crystals. Branching of the hot-exciton self-trapping in pyrene, including the bypassing flow of excitons that avoided the intraband energy relaxation, directly to the self-trapped states, has been observed.

Self-Trapping in Peierls Insulators

According to the Peierls theorem, a 1D metal with a single electron per lattice site is unstable and undergoes a phase transition into an insulating phase by lattice deformation. The scheme of alternating single and double bonds after the Peierls transition in a conducting polymer *trans*-polyacetylene is shown in **Figure 5**. Because double bonds are slightly shorter than single bonds, the lattice period doubles, from a to $2a$, and a gap $2\Delta = 1.8$ eV opens in the spectrum. Because a- and b-configurations of **Figure 5** have the same energy, there are two ground states, or two vacuums. A defect shown in **Figure 5** c is a kink

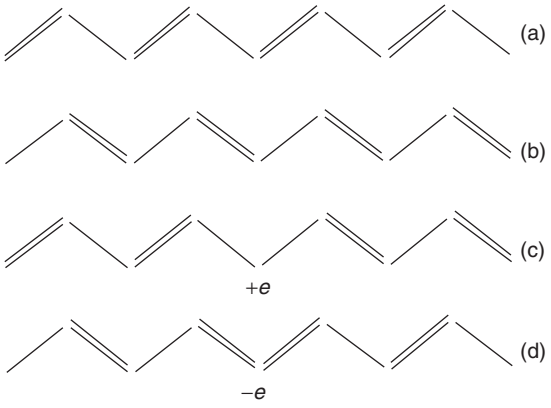


Figure 5 Schematics of valence bonds in the carbon backbone of *trans*-polyacetylene and topological solitons. (a) and (b) Two degenerate ground states with alternating bonds; (c) a soliton; (d) an antisoliton.

separating two vacuums – the *a*-vacuum on the left and the *b*-vacuum on the right; it is called a soliton, *S*. In the Heeger–Schrieffer–Su model, the energy of a deformed carbon backbone of *trans*-polyacetylene is

$$H_{\mathrm{el-ph}} = \alpha \sum_{n,s} (u_{n+1} - u_n)(a^{+}_{n+1,s} a_{n,s} + a^{+}_{n,s} a_{n+1,s})$$

where n numerates lattice sites and s spin projections. In this theory, solitons and antisolitons ($\bar{S}$) are extended defects of the size $2x_0 \approx 13a$. Their energies $\approx 0.6\Delta$ are less than Δ; hence, they are the basic low-energy excitations of the electron system. Because lattice displacements in solitons are small, only $\sim 0.03a$, soliton masses are also small, $m_S \approx 6m_0$. In the continuum approximation, the lattice displacement in a soliton, $u(x) = u_0 \tanh(x/x_0)$, creates a midgape state with the energy $E = 0$ that can be populated by one or two electrons.

There are two basic differences between solitons and polarons. First, the usual spin–charge relation is reversed in solitons. S has spin $s = 0$ when charged positively, $+e$, and spin $s = 1/2$ in the neutral state when a single electron is trapped at the midgape level. Second, a soliton is a topological excitation because a kink separates two vacuums, while a polaron only produces a local lattice distortion and does not change the vacuum. In this system, a polaron is a midgape electron bound to an $S\bar{S}$ pair; binding energy equals $\Delta/\sqrt{2}$. Because a single injected electron cannot change the vacuum, it relaxes into a polaron. However, an electron–hole pair relaxes into an $S\bar{S}$ pair.

Because the deformations are small and smooth, $u_0 \ll a$ and $x_0 \gg a$, the theory is rather accurate, and comparison with experiment is extensive and successful.

Polarons in Strongly Correlated Electron Conductors

Historically, polarons were devised for insulators and semiconductors. More recently, they were invoked in physics of strongly correlated electronic conductors, in particular, high-T_c superconductors and colossal magnetoresistance materials (cuprates, perovskite manganites, EuB_6, etc.). These materials are unstable, close to metal–insulator and magnetic transitions, show strong electron correlations, and contain $\sim 10^{-1}$–10^{-2} carriers per formula unit. Because of the presence of ions with partially filled atomic shells, they are candidates for Jahn–Teller polarons, and optical signatures of such polarons are present in the spectra. Strong correlations of Jahn–Teller polaron pairs result either in their bound states (bipolarons) or resonances inside the Fermi sea that manifest themselves in a pseudo-gap, believed to be a precursor of the superconductor gap still in the normal state. They are believed to enhance T_c over the canonical BCS mechanism.

The interplay of several competing mechanisms makes it difficult to isolate polaron effects; however, their role is well documented already.

See also: Electron–Phonon Interactions and the Response of Polarons; Electrons and Holes; Excitons: Theory; Polaritons.

PACS: 71.38. – k and all its subsections: 71.38.Cn; 71.38.Fp; 71.38.Ht, and 71.38.Mx; 71.35.Aa; 75.47.Gk; 74.72. – h

Further Reading

Alexandrov AS and Mott NF (1994) *High Temperature Superconductors and Other Superfluids*. London: Taylor and Francis.

Appel J (1968) Polarons. In: Ehrenreich H, Seitz F, and Turnbull D (eds.) *Solid State Physics*, vol. 21, p. 193. New York: Academic Press.

Brazovskii SA and Kirova NN (1984) Electron selflocalization and superstructures in quasi one-dimensional dielectrics. In: Khalatnikov IM (ed.) *Soviet Scientific Reviews*. New York: Harwood Academic.

Devreese JT (1996) Polarons. *Encyclopedia of Applied Physics*, vol. 14, p. 383. New York: VCH.

Feynman RP (1955) Slow electrons in polar crystals. *Physics Review* 97: 660.

Fomin VM and Pokatilov EP (1988) Non-equilibrium properties of charge-carriers with arbitrary coupling to the lattice in semiconductors. *Physics Reports* 158: 205.

Fröhlich H (1954) Electron in lattice fields. *Advances in Physics* 3: 325.

Fugol' IY (1978) Excitons in rare-gas solids. *Advances in Physics* 27: 1.

Heeger AJ, Kivelson S, Schrieffer JR, and Su W-P (1988) Solitons in conducting polymers. *Review of Modern Physics* 60: 781.

Holstein T (1959) Studies of polaron motion. *Annals of Physics* 8: 325 and 343.

Ioselevich AS and Rashba EI (1992) Theory of nonradiative trapping in crystals. In: Kagan Yu and Leggett AJ (eds.) *Quantum Tunneling in Condensed Matter*, p. 347. Amsterdam: North-Holland.
Krivoglaz MA (1974) Fluctuon states of electrons. *Soviet Physics–Uspekhi* 16: 856.
Levinson IB and Rashba EI (1974) Threshold phenomena and bound states in the polaron problem. *Soviet Physics – Uspekhi* 16: 892.
Nagaev EL (2001) Colossal-magnetoresistance materials: manganites and conventional ferromagnetic semiconductors. *Physics Reports* 346: 388.
Pekar SI (1946) Local quantum states of an electron in an ideal ionic crystal. *Journal of Physics USSR* 10: 341.
Ueta M, Kanzaki H, Kobayashi K, Toyozawa Y, and Hanamura E (1986) *Excitonic Processes in Solids*. Berlin: Springer.

Polymer Structures

R A Pethrick, University of Strathclyde, Glasgow, UK

Introduction

In order to understand the structure of a polymeric material, it is important to recognize the subclass of macromolecular materials to which the substance belongs. In principle, there are many features which are common to naturally occurring and synthetic polymers. The former is usually produced via a "step growth" process and hence has atoms other than carbon in their backbone, one of the exceptions to this generalization being natural rubber, which has an all-carbon backbone. Associated with the heteroatom incorporation are the amide, urea, ester, lactone, lactam, ether, etc. functions upon which so many of the structures are based. The polymer backbones or side groups may contain hydrogen-bonded entities and their interactions build the hierarchal structures that are so important in biological polymer systems. Many synthetic polymers are produced via "addition" reactions in which an active center created by a radical, cation, or anion adds to an unsaturated carbon–carbon bond and propagates the growth of a polymer chain. This type of chemistry is responsible for many of the highly used polymers, for example, poly(ethylene), poly(propylene), poly(vinylchloride), and poly(styrene). Polymers such as poly(ethyleneterephthalate) are produced via a step-growth process similar to that used by nature in much of its macromolecular synthesis. Polymers produced by addition reactions can have chain lengths that are all very similar, whereas the very nature of the step-growth process will tend to generate a broad distribution of chain lengths. The difference in the distribution of chain lengths can have a significant effect on the structures created by these polymers. A narrow distribution of chain lengths will allow a level of organization that is not observed in the broader distribution.

Influence of Molecular Structure on Morphology

Polymer materials can be divided into two types: those produced with two functions per repeat unit and form linear structures, and those which have functionality greater than 2 and form branched or cross-linked structures. The latter are often "thermoset" materials and may not be reprocessed by heating, whereas the former linear chain polymers are "thermoplastic" and can be reprocessed into new structures by heating or dispersion in a solvent. The ability for thermoplastic polymers to generate organized structures depends very much on the balance between the inherent entropy, associated with the conformational distribution between the various stable states possible for the polymer chain. For instance, in the case of poly(ethylene), there are two possible states: gauche and trans, the latter having lower energy and hence a higher probability of occurrence (**Figure 1a**).

In terms of structure building, the trans structures are ideally suited for packing into a regular crystalline ordered arrangement and creating domains which are highly ordered within the solid. The gauche structures will disrupt this packing and the crystallization processes proceed with the clustering of these higher-energy states into regions which are disordered or amorphous in nature. Certain polymers which possess bulky side chains, polystyrene being a classic example (**Figure 1b**), are unable to pack and hence the dominant structure is amorphous. If the monomers are multifunctional, then network formation leading to a thermoset structure is possible (**Figure 1c**). It is, therefore, possible to divide polymer materials into three general classes of materials: amorphous, crystalline, and network structures. In the content of networks and amorphous materials, it is important to recognize that the dynamics of the polymer chains can influence the perception of the material without requiring a major redefinition of the overall structure. In the case of an amorphous

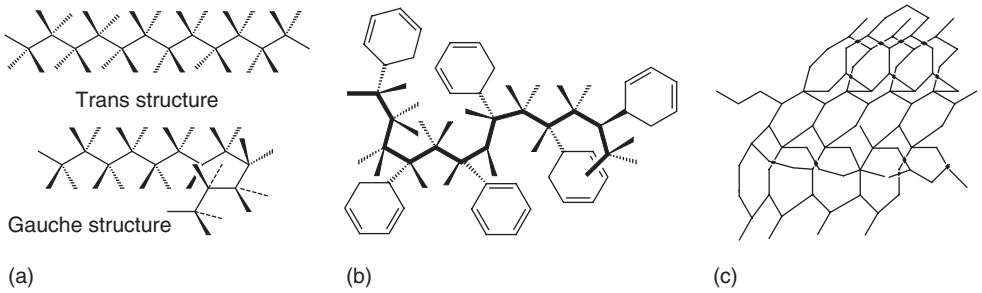


Figure 1 Possible structures for (a) poly(ethylene) chain, (b) poly(styrene), and (c) a network based on multifunctional monomers.

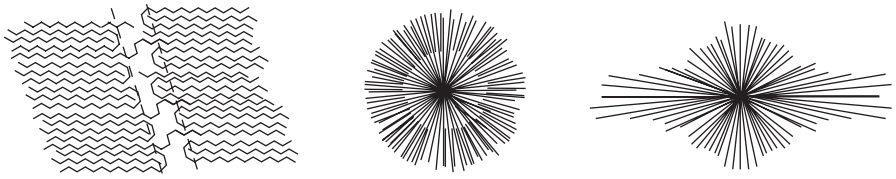


Figure 2 Chain folding in poly(ethylene) leading to lamellae formation and a pictorial representation of spherulitic growth.

polymer, if the interactions between neighboring polymer molecules are strong, then it is unlikely that there will be sufficient volume for the chains to readily change their conformations. The matrix distribution is, therefore, essentially frozen and the material is a glass. Heating the material can cause an expansion to occur, reducing the chain–chain interactions and producing sufficient volume between the chains for conformational change to occur. Creation of the "free volume" allows conformational changes, which involve cooperative motions of the polymer backbone and the material becomes flexible. This rubbery or elastomeric structure has been created without significant redistribution of the chains within the structure and can still be considered to be amorphous. It is also possible to see the same changes from glassy to rubbery behavior in a network structure.

Crystalline Polymer Structure

Only in solution-grown polymer single crystals will perfect, often rhombohedral, crystal structures be observed. In general, the structures obtained from processing crystallizable polymeric materials will reflect the thermal history associated with solidification of the material. In the melt phase, the polymer will adopt a configuration that the ideal flexible polymer would have in dilute solution and corresponds to the so-called random coil structure. The coiling behavior is a direct reflection of the gauche content in the conformational distribution. As the polymer cools down, the proportion of the trans content in the conformational distribution increases and the possibility of the alignment of these units necessary for crystal growth is encouraged. The consequence of the residual gauche content is that the elements of the chains will be disordered and can even loop back on themselves. The looping back creates a hairpin-like organization (**Figure 2**), which aids further packing of subsequent chains. The result of the folding is that the lamellae-type of structure is generated. The thickness of the lamellae, L_{lam}, is dictated by the temperature at which the crystallization process and the growth mechanism take place. The lamellae for poly(ethylene) are typically between 10 and 18 nm, for poly(4-methylpentene-1) between 11 and 15 nm, and for poly(oxymethylene) between 6 and 10 nm. If the surface energy of the fold surface is σ_f, the enthalpy for melting of the crystal phase is ΔH_M^0, and the melt temperature is T_M^0, then

$$(L_{\text{lam}})_{\text{theor}} = \frac{2\sigma_f T_M^0}{(\Delta H_M^0)(T_M^0 - T_{\text{cryst}})}$$

where T_{cryst} is the temperature to which the melt has been supercooled before crystallization occurs. The ratio of the instantaneous volume fraction ϕ of the crystalline material in the total sample to the value after infinite time ϕ_∞ indicates a randomly distributed nuclei. The probability p that a point does not lie in any one entity is proportional to the fraction:

$$p = 1 - \left(\frac{\phi}{\phi_\infty}\right)$$

If all the entities are formed simultaneously, the density of nuclei is a constant and the probability depends on the mean volume $\bar{V}_i$ and the number of entities N. The concentration of entities then becomes $v = N/V$, where V is the total volume. This leads to

$$\phi = \phi_\infty [1 - \exp(1 - v\bar{V}_i)]$$

The growth is controlled by the following factors: whether or not the nucleation is spontaneous; and the form of the growth, that is, rod, disk, or spherical. This leads to the Avrami equation which has the form

$$\phi = \phi_\infty [1 - \exp(1 - zt^n)]$$

where z and n are constants, characteristic of the particular growth mechanism. While the theoretical predictions and practical observations do not show a strong correlation, the theory has been useful for rationalizing the growth observed in similar systems.

Spherulites

Electron microscopic examination indicates that the lamellae will tend to cluster from the nucleation point and form into a floret- or spherulite-like structure (**Figure 2**), which has an interface with the growth boundary of a neighboring cluster. These boundary layers are important as they represent regions of potential weakness and can develop on cold-drawing of the polymer. Cold-drawing of predominantly isotactic poly(propylene) creates fibers with strength significantly greater than those of the originally extruded sheet of polymer (**Figure 3**).

The cold-drawing creates alignment of the lamellae and allows development of anisotropic mechanical properties. The modulus in the draw direction is significantly increased, whereas it decreases in the transverse direction due to the creation of voids. Spherulites make films opaque when their dimensions are greater than half the wavelength of light and when, in addition, inhomogeneities exist in relation to the density or the refractive index. Spherulitic poly(ethylene) is opaque, but spherulitic poly(4-methyl-pentene-1) is glass clear even when both have the same number of spherulites with the same dimensions.

Dendrites and Epitaxial Growth

Spherulites are produced because the overall crystallization rate is approximately the same in all directions. However, the growth rates of crystals in the spherulite may be directionally dependent. Conversely, if the overall crystallization rate is also direction dependent, then what are called dendrites are produced. A low-temperature solution-growth process often produces these structures. Acid etching of the polymer indicates that the interior contains a polymer chain that can be oxidized and represent an amorphous component. The remaining crystalline component has a lamellar structure of regular thickness.

Shish Kebab

High-temperature solution growth can generate a structure that resembles a "shish kebab," (from the Armenian dish) structure. When crystallization from dilute solution occurs with strong stirring, these structures are produced (**Figure 4**). The macromolecules are orientated along the direction of flow and settle out perpendicular to each other. X-ray scattering, electron diffraction, and birefringence measurements show that the chains lie parallel to the fiber axis. The fibrils produced order themselves into nucleation bundles, but the shear gradient is strongly reduced between the fibrils crystallizing out on these in the form of folded chain lamellae. The lamellae here lie vertical with respect to the fibrils. The shish kebab formation is a special case of epitaxial growth.

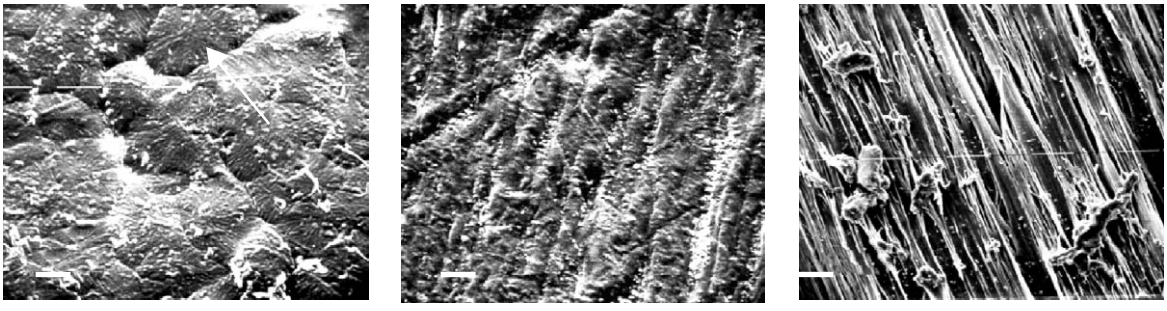


Figure 3 Electron micrographs of cold-drawn polypropylene indicating the alignment of the crystal structure, the bar is 1 μm, λ = draw ratio: $\lambda = 1$, a domain is identified by the arrow; $\lambda = 6$, the parallel texture is a consequence of the lamellae aligning in the draw direction; $\lambda = 12$, the process of drawing has generated fibers and the bridging polymer fragments are clearly seen between the fibers.

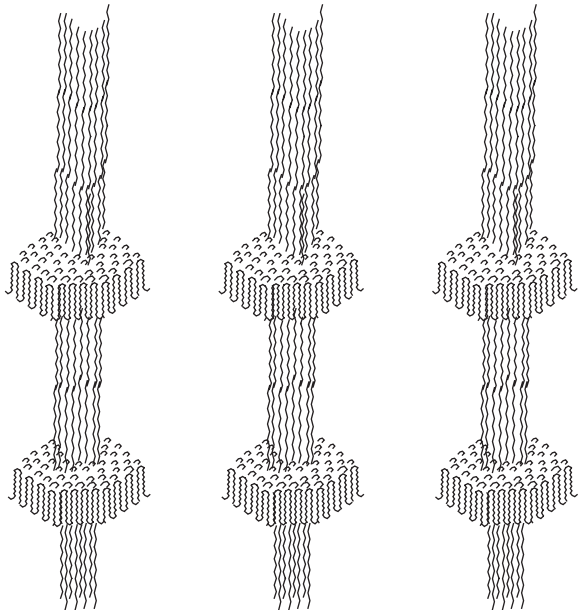


Figure 4 Schematic representation of the arrangement of chains that form a shish kebab structure.

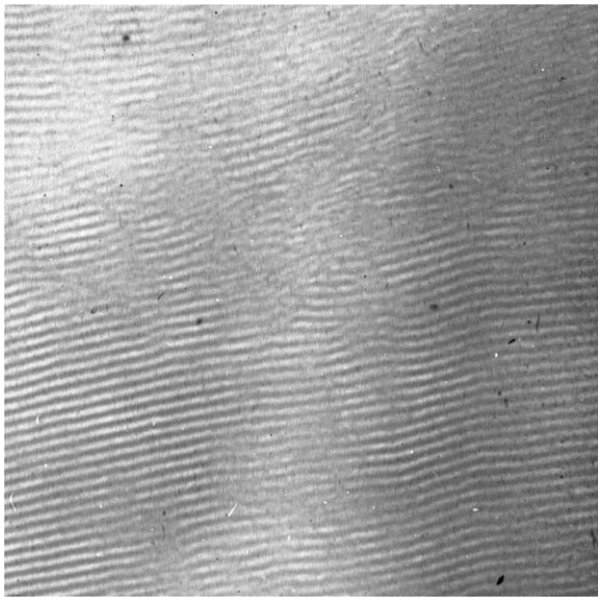


Figure 5 Transmission electron micrograph of styrene–butadiene–styrene triblock copolymer stained with osmium tetroxide. The domain sizes are 10 μm in width.

The orientated growth of one crystalline substance on another is defined as epitaxy.

Amorphous Polymer Structure

The amorphous state is characterized by a complete lack of structure or organization of the polymer chains. The chains that form this disorganized state will have trapped a certain amount of volume – "free volume." It is possible to probe this free volume using positron annihilation lifetime spectroscopy and to show that at the glass–rubber transition, T_g, there is an increase in the volume that liberates the backbone motions of the polymer chain and allows it to demonstrate a rubbery character. The polymer chains in the amorphous state will have a size which is similar to that found at the theta temperature in solutions. The theta temperature corresponds to a thermodynamic state in which the interaction of the polymer segments with one another is identical to the solvent when the polymer is in dilute solution. Most amorphous polymers and crosslinked networks do not show any structure when observed using X-ray diffraction or electron microscopy.

Block Copolymers

Due to thermodynamic incompatibility of the segments of a styrene–butadiene–styrene tri-block, copolymer phase separation can be observed in styrene (light colored) and butadiene (dark colored) phases. The different colors in the electron micrography are achieved by using osmium tetroxide that adds to the butadiene double bonds, and the osmium generates the scattering contrast (**Figure 5**).

Extrusion of the block copolymer can generate an aligned material in which the styrene and butadiene phases, orientated with the flow and perpendicular to the flow direction, form a regular ordered structure (**Figure 6**). The result of this alignment is that the modulus in the direction of the flow is dominated by the styrene phase and has a value at room temperature close to that of poly(styrene). In the transverse direction, the material exhibits a much lower modulus which is close to that of poly(butadiene), which is a rubber at room temperature. Phase separation has a dominating effect on the physical properties of a number of polymer systems, and many of the useful characteristics to be found in polyurethanes are a consequence of phase separation of the hard aromatic urethane phase from the flexible poly(ether) or poly(ester) phases.

Phase separation is very important in thermoset resins. Rubber toughening in epoxy resin can be achieved by the addition of carboxy terminated butadiene–acrylonitrile copolymer – CTBN. The CTBN is initially compatible with the epoxy amine mixture but as cure proceeds, the rubber phase separates out and forms rubbery domains (**Figure 7**). It is interesting to note that the white specs within the rubbery phases are epoxy resin, which is incorporated in the phase as it separates but subsequently is precipitated as it cures. However, since it is now in the rubbery phase, it is occluded in that phase rather than forming part of the main matrix structure.

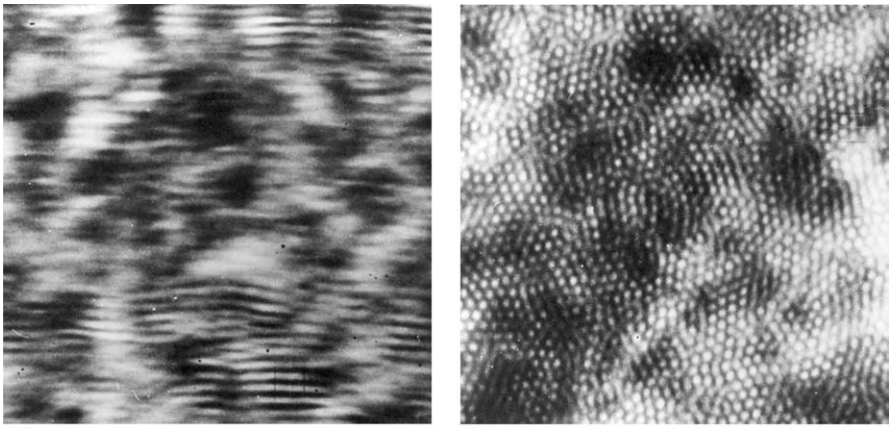


Figure 6 Electron micrographs (a) parallel and (b) perpendicular to the extrusion direction for an osmium tetroxide strained styrene–butadiene–styrene triblock copolymer. The light colored domains are ~10 nm in size.

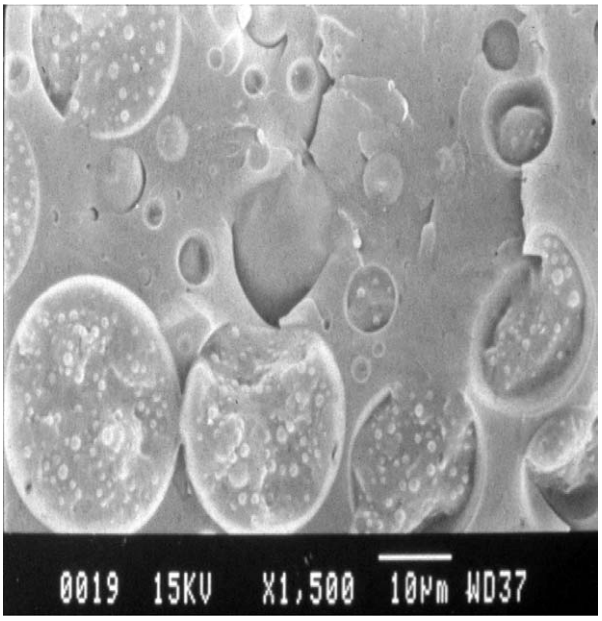


Figure 7 Electron micrograph of a carboxyl terminated butadiene–acrylonitrile copolymer phase dispersed in a thermoset amine cured epoxy resin.

These rubbery phases impart toughening to the epoxy resin by providing a mechanism for energy dissipation. Although amorphous polymers are not normally considered to exhibit structures, phase separation does create features that are important in understanding the mechanical properties of the polymers.

Liquid Crystalline and Related Structures

Polymers, like small molecules, are observed to exhibit partial ordered structures. Like their low-molar-mass analogs, polymers also exhibit the optical birefringence that is characteristic of partial degrees of alignment: nematic, smectic, and/or cholesteric organization (**Figure 8**).

These structures are distinguished on the basis of the extent to which they exhibit coherent alignment and dimensional order. In the nematic phase, the species are aligned in one direction, designated as the director. The smectic phase has alignment in at least

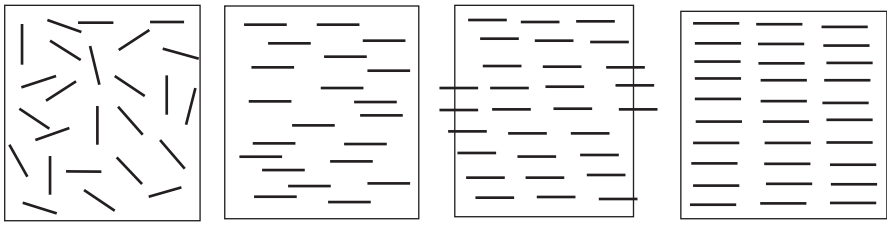


Figure 8 Schematic representation of the alignment and order in nematic and smectic phases.

two directions and may also exhibit order in a third phase depending on its designation. In the smectic A phase, the director is parallel to the plane normal, whereas in the smectic C phase the director is inclined along the plane but the perfection approaches that of the crystalline state. In the cholesteric phase, entities exhibit a regular register from plane to plane and there is a helical twist developed through the volume as the planes rotate (**Figure 9**).

In addition to these idealized structures, there are a number of characteristic features that reflect the influence of various types of characteristic defects on the material. Because the ordered domain has dimensions of the order of the wavelength of light (~ 500 nm), dramatic birefringent optical images can be observed when crossed polarization is used for the microscopic examination. These structures are rarely uniform and the director will vary slowly through space. Different phases have different characteristic textures that arise from the permitted symmetry of the defects that occur in the packing process. The textures can be used to identify the phases. However, in the case of the "fingerprint" texture seen in cholesteric liquid crystals, the texture arises simply from its helical structure as indicated in **Figure 9**. Light is rotated by the director as long as the rate of twist is not too great compared with the wavelength of light, λ.

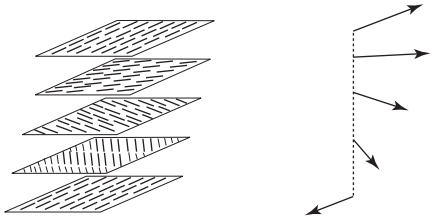


Figure 9 Schematic of the order and relative rotation of parallel planes in a cholesteric liquid crystalline phase.

Defects

Within a particular plane it is possible to envisage a twist, splay, or slow bend of the director and these structures will all have characteristic optical birefringence. Dislocations are discontinuities in the translation of atoms in a crystalline phase; however, in a liquid crystalline phase the discontinuities are in orientation, within the director field. There are various types of disclinations characterized by their "strengths." The term "strength" indicates the number of multiples of 2π that the director rotates in a complete circuit around the disclination core.

Under crossed polars, four dark "brushes" are seen where the director is parallel to one of the polarizers, and extinction occurs. The defect may be either a line or a point defect. A variety of different defects can be identified which are characteristic of a particular liquid crystalline phase.

It is not very surprising to find that polymer films can exhibit liquid crystalline order. Polymeric liquid crystallines are usually divided into main chain (MCPLC) and side chain polymer liquid crystals (SCPLC). The liquid crystalline character of the main chain polymers reflects the correlation of regular kinks in the backbone allowing organized regions to be developed. Examples of this type are poly(ethyleneterephthalate) and poly(hydroxynaphthalenic acid).

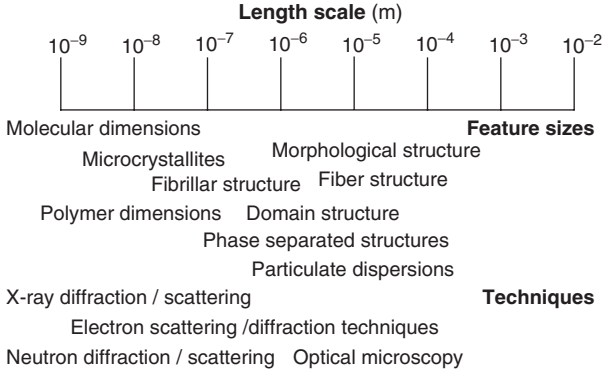


Figure 11 Length scale variations and visualization techniques.

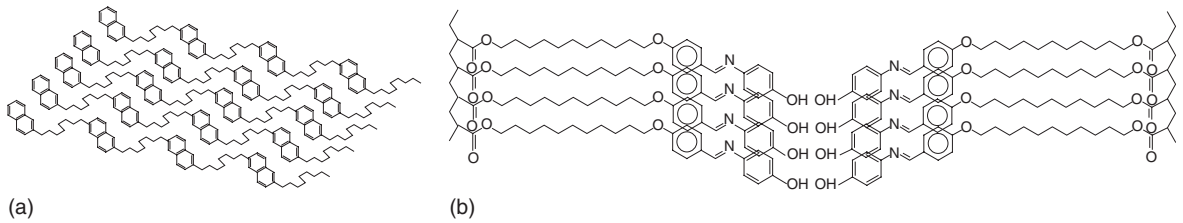


Figure 10 Schematic of backbone and side chain liquid crystalline polymers: (a) main chain liquid crystal order and (b) side chain liquid crystal order.

The flexible ethyleneoxide spacer allows the relative alignment of neighboring chains and the polymer to display liquid crystalline order. Kinks in the backbone structure of the poly(hydroxynaphthalenic acid) allows a similar alignment to be achieved and liquid crystal behavior to be observed. There has also been considerable interest in side chain liquid crystalline polymers, where the pendant entities are designed to be flexibly attached to the backbone and can themselves align to give ordered structures (**Figure 10**).

Liquid crystalline structures have been detected in a wide range of polymers including poly(ether imides), poly(4,4′-dioxy-2,2′-dimethyl-azoxybenzene dodecanedioyl)-co-polyoxyethylene 400 (PMABD-co-POE), poly(didocecyl-*p*-phenylene), poly(diacetylene) P-4BCMU, poly(*p*-phenylene terephthalamide), and aromatic polyesters. It is clear that organization ion polymer systems not only have to be considered at a molecular level, but also in terms of nano- and micro-length-scales. **Figure 11** indicates how length scales can be studied.

Biopolymeric Structures

It is well known that various cellulose derivatives can exhibit a liquid crystalline behavior; however, it is appropriate to consider how the organization can be mapped between the molecular chain and the macrostructure of a cotton fiber (**Figure 12**). Cellulose has a 1,4-β-D-linked polyanhydropyranose structure. In order to understand the physical properties, one has to be able to distinguish the principal structure at three different levels.

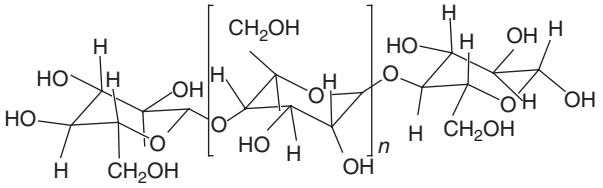


Figure 12 Structure of the cellulose molecule.

The Molecular Level

At a molecular level, the chains that consist of β-D-glucose existing in the pyranose form adopt a $^4C^1$ (chair conformation $^4C^1$ denotes that the ring which is attached through the 1- and 4-carbons is in the chair configuration), which is the lowest energy conformation. The presence of intramolecular hydrogen bonding between the hydroxyl group in the 3 position and the ring ether oxygen in the 5 position imposes a twist on the C(1)–O(1) and the C(4)–O(4) glucosidic bonds which results in the chain adopting a helical conformation.

The possibility of polymorphic forms being obtained depends on how the rigid chains pack together. The study of these forms has been the subject of active study for many years. Native cellulose exhibits the cellulose I structure. The mercerized or regenerated cellulose exhibits a different structure designated as cellulose II. Of the two forms, the latter is considered to be more thermodynamically stable. The average hydrogen bond length in the case of cellulose II is 2.72×10^{-10} m, which is remarkably shorter than the value of 2.80×10^{-10} m in cellulose I. The denser structure and greater involvement in the hydrogen bonding are reflected in the lower susceptibility of the regenerated cellulose to reaction. The polymer chains, although aligned, also exhibit

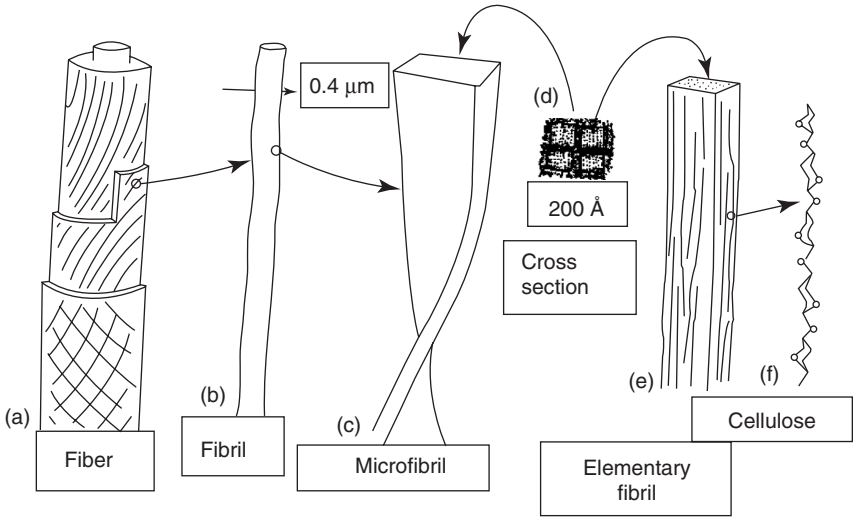


Figure 13 Morphology of a cellulose fiber showing the various levels of organization.

disorder that leads to the creation of amorphous regions and disrupts the perfection of the crystal structure.

The Supermolecular Level

The crystalline regions will organize themselves into larger structures, and the number and form that they take is dictated by a variety of factors which reflect the processes involved in the formation of the macrostructures (**Figure 13**). The crystallites are organized into the microfibrils that pack to form the fibril. This packing, however, is not random, and the microfibril having a twisted structure itself aids the formation of the fibril.

Macrostructural Level

The fibrils are themselves packed into a layered structure that reflects the twisted nature of the fibrils. The center of the cotton fiber is a hollow structure formed by the criss-crossing of the fibers to generate a tube-lumen upon which the other layers are formed. The subsequent layers adopt a common alignment of the fibers and form the so-called secondary wall structure. The outer layer of cellulose is the winding layer that is also composed of a criss-crossed layer. Outside this layer, the primary wall and the cuticle are to be found, and these often contain noncellulose materials such as lignin.

The morphology of biopolymers is dominated by hydrogen bonding and self-assembly, and there is a rich spectrum of level of structures which can be observed.

See also: Composites: Polymer–Polymer; Copolymers; Molecular Crystallography; Polymers and Organic Compounds, Optical Properties of; Polymers and Organic Materials, Electronic States of; Polymers, History of; Scattering, Elastic (General); Shubnikov–de Haas and de Haas–van Alphen Techniques; Transmission Electron Microscopy.

PACS: 61.41.+e; 61.25.Hq; 81.05.Lg; 36.20.−r

Further Reading

Balta-Calleja FJ and Vonk CG (1988) *X-Ray Scattering of Synthetic Polymers, Polymer Science Library 8*. Amsterdam: Elsevier.

Campbell D, Pethrick RA, and White JR (2000) *Polymer Characterization – Physical Techniques*. Cheltenham, UK: Stanley Thornes.

Elias HG (1977) *Macromolecules*, 185pp. London: Wiley.

Hammond C (2001) *The Basics of Crystallography and Diffraction*. Oxford: Oxford University Press.

Huang JM, Kuo JF, and Chen CY (1995) Studies on mesomorphic behaviors of segmented azoexy polyester containing polyoxyethylene. *Journal of Applied Polymer Science* 55(8): 1217–1229.

Pethrick RA (2003) Characterisation of molecular organisation in polymeric materials. In: Pethrick RA and Viney C (eds.) *Techniques for Polymer Organisation and Morphology Characterisation*. London: Wiley.

Pethrick RA, Adachi K, North AM, Harrison G, and Lamb J (1982) *Polymer* 23: 1451–1456.

Pethrick RA and Crofton D (1981) *Polymer* 22: 1048–1053.

Pethrick RA, Daly J, Fuller P, Cunliffe AV, and Datta PK (1981) *Polymer* 22: 32–36.

Pethrick RA and Datta PK (1978) *Polymer* 19: 145–148.

Pethrick RA and Delides C (1981) *European Polymer Journal* 17: 675–681.

Pethrick RA, North AM, and Wilson AD (1978) *Polymer* 19: 913–922.

Vickers ME (1995) Scattering methods. In: Richards RW and Ellis Horwood (eds.) *Polymer Science*.

Wegner G (1996) *Macromolecular Symposium*, vol. 101, pp. 257–264. Basel: Huthig and Wepf Verlag.

Woodward AE (1989) *Atlas of Polymer Morphology*. Munich: Carl Hanser.

Zhu Z (1996) *Macromolecular Symposium*, vol. 102, pp. 209–216. Basel: Huthig and Wepf Verlag.

Polymers and Organic Compounds, Optical Properties of

T Koda, University of Tokyo, Tokyo, Japan
K Kajikawa, Tokyo Institute of Technology, Yokohama, Japan

Introduction

This article deals with the optical properties of organic substances, which are by far rich in diversity and make the reductionalist view point of physics quite hard to apply. Nevertheless, during the latter half of the twentieth century, considerable progress has been made in establishing the simplest possible physical pictures, such as excitons, phonons, polaritons, and so on. This trend is still underway, aiming at the goal which may be best expressed by the following words of Mott in the preface of a book by Pope and Swenberg: "The reasons are many, but certainly the most important is that new things are to be found, which is rarely the case in the well-explored field of inorganic crystals. —and perhaps the distant dream of a new technology and one possible application to the understanding of such biological processes as do involve electronic motion."

Structures of Organic Materials

In the following, organic opto-electronic materials composed of conjugated molecules are considered. Starting from an ethylene molecule, $H_2C=CH_2$, and connecting it with another ethylene molecule one by one, either an open-end linear conjugated system or a closed-ring conjugated system is obtained, which are named in chemistry as aliphatic and aromatic molecules, respectively. In the limit of infinite chain length, the former system becomes a prototype of conjugated polymers. Typical molecular structures of the conjugated molecules are listed in **Figure 1**.

Assemblies of these conjugated molecules can be further classified into two major types, according to the nature of dominant intermolecular cohesive interaction: (1) Molecular crystals, composed of electrically neutral molecules. The van der Waals interaction between the molecules with no net charge and permanent electric dipole moment is responsible for the cohesive force. Most of the mono-molecular crystals of aromatic molecules belong to this type. (2) The charge-transfer (CT) complex crystals composed of two kinds of molecules having electric charges of opposite sign due to the intermolecular electron transfer. The effective molecular charge ρ (the degree of CT) is determined by the trade-off relation between the Coulombic energy gained by the long-range electrostatic interactions in the whole molecular assembly and the energy needed for the CT process, similar to the case of a bimolecular donor (D)-acceptor (A) CT complex $D^{+\rho}A^{-\rho}$.

There are other kinds of organic materials, though not so major in numbers. Among them, hydrogen

(a) Aromatic (ring-like) and related molecules

Family of polyacenes

Name	Structure	λ_m(nm)
Benzene		259
Naphthalene		310
Anthracene		380
Naphthacene (tetracene)		471
Pentacene		575

Some related molecules

Phenanthrene		293
Pyrene		336
Perylene		437
Porphyrine		400 (Soret band) 700 (Q-band)
Phthalocyanine		400~700 (depending on the ligand metal at the center of the ring)

(b) Aliphatic (linear) and related molecules

Family of polyenes

Name	Structure	λ_m(nm)
Ethylene		~162
Butadiene		~220
Hexatriene		~260
Octatriene		~300

Some related molecules

β-Carotene (all-trans)		~480
Trans-retinal		~370
Polyacetylene (trans-form)		~690
Poly-(*p*-phenylenevinylene) (PPV)		~430

Figure 1 Typical conjugated molecules; (a) for ring-like molecules, and (b) for chain-like molecules. Molecular structures are shown according to the IUPAC notations. λ_m represents approximate wavelength (in nm) of the lowest main absorption band.

(H)-bonded molecular crystals are important. Two atoms X, Y having electro-negativities larger than the H atom, are connected like –X–H–Y– (X, Y = N, O, P, S, …) to form a weakly coupled molecular crystal-like crystalline H_2O (ice). In organic materials and polymers, there are several important H-bonded substances including bio-polymer DNA, for instance. Often the H atom has bi-stable or multi-stable positions, showing interesting optical and dielectric phenomena due to the proton transfer. Besides, noncrystalline assembly of organic molecules is also important for technological applications. Conjugated polymer films having a spaghetti-like irregular structure of entangled fibrils may be regarded as a quasi-one-dimensional (1D) electronic system with a mixed character of amorphous and crystalline structures. Another important disordered molecular assembly is liquid crystals. These topics are described in the last part of this article.

Simple Model for π-Electron Systems

The conjugated molecules shown in **Figure 1** exhibit an obvious trend of increasing wavelengths of absorption bands as the repetition number N of the conjugation unit increases (**Figure 2**). The first theoretical model for this is known as the Kuhn model. It is assumed that the total $2N$ π-electrons including the spin freedom are confined in a 1D potential well, having width equal to the conjugation length $L(= aN)$. When the potential is flat and the wall height is infinite, it is an elementary task to calculate the quantized states for each π-electron in the Hückel approximation, neglecting the electron correlation effect. The photoexcitation of an electron takes place from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), the HOMO–LUMO gap corresponding to the lowest absorption energy due to the spin-singlet $\pi-\pi^*$ transition. To improve this simplest possible model, the flat potential may be replaced by a weak periodic potential as a perturbation. The same model is applicable to the ring-shaped molecules by using the cyclic boundary condition. This model leads to a theoretical expression for the lowest singlet $\pi-\pi^*$ transition energy ΔE as follows:

$$\Delta E = \left(\frac{\hbar^2}{2m}\right)\left(\frac{2N+1}{L^2}\right) + V_0\left(1 - \frac{1}{2N}\right)$$

Here, $\hbar$ and m are conventional fundamental constants, V_0 ($\sim 2\,\text{eV}$) is an empirical parameter for the amplitude of sinusoidal periodic potential. The first term on the right side is the increment of kinetic energy when an electron is photoexcited, while the second term is due to the perturbation of the periodic potential. In the limit of infinite length L, the first term vanishes, so that the π-electron system must exhibit a free-electron-like (metallic) behavior if $V_0 = 0$. But, when $V_0 \neq 0$, the second term remains finite in the same limit, which implies that the long conjugated molecules and polymers must be a semiconductor with a finite excitation gap.

Though the Kuhn model is useful for providing an intuitive physical picture, like the Kronig–Penny model in semiconductor physics, it has an obvious fault of completely ignoring the electron-correlation effect, which is evidently unrealistic for the π-electron systems confined within a low-dimensional molecular geometry. In the opposite limit of the strongly correlated π-electron systems, the $2N$ π-electrons may be better treated as a Fermion gas, instead of the one-electron model in the Hückel approximation. In the Fermion gas model, the finite excitation energy in the limit of $N \to \infty$ is interpreted as due to the plasma oscillation of the π-electron gas. However, the underlying physics is much more complicated than depicted by either models. In the subsequent course of studies, it has been widely accepted that the origin of finite excitation energy in long conjugated linear-chain molecules and polymers is due to the bond-alternation caused by the structural instabilities inherent in a 1D electron-lattice system and in a 1D spin-lattice system. The energy gap due to the former is called the Peierls gap, and the latter the Hubbard gap. The charge density wave (CDW) excitation, and the spin density wave (SDW) excitation are responsible for the opening of the respective gaps. The actual conjugated molecular systems have a mixed character to be treated as a 1D Hubbard–Peierls system.

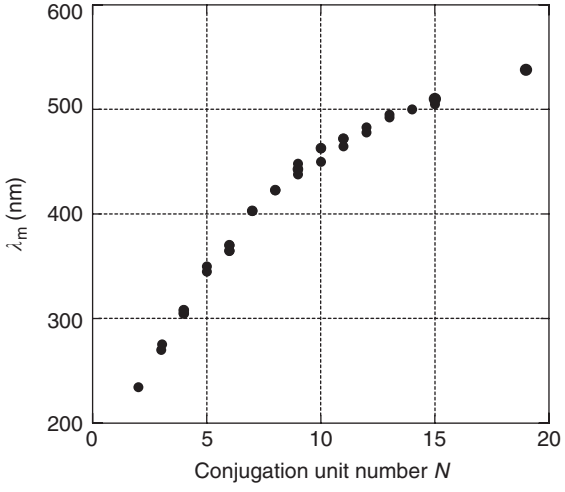


Figure 2 The absorption peak wavelength λ_m vs. conjugation unit number N for transpolyenes $[–HC=CH–]_N$. Filled circles are experimental data for specific molecules (not indicated here) in solution. (Due to H Suzuki and S Mizuhashi.)

So far, the spin-singlet excited states were considered. In the early stage of study, it had been suggested that the spin-triplet excited states were responsible for phosphorescence in organic phosphors. In subsequent studies, photoinduced transient effects in magnetic, ESR, and optical properties had confirmed that the triplet excited states are certainly photo-generated in various organic molecules and crystals. The singlet-triplet splittings are fairly large ($\leqslant 1$ eV), indicating a strong electron correlation effect in the π-conjugated organic systems.

Vibronic Spectra of Organic Molecules

The energies for the ground and excited states, E_g and E_{ex}, of a π-electron system can be expressed as a function of so-called configuration-coordinates (CC), Q_iLs, in a general form of E_i (i= g, ex)(Q_1, Q_2, … Q_i …). For simplicity, one assumes that a particular mode Q (usually of the totally symmetric breathing mode) is dominant, so that the vibronic energies for the ground and excited states (singlet and triplet) are depicted by three curves, as illustrated in **Figure 3**. Photoexcitation takes place from the vibronic levels of the ground state to those of the singlet excited state, with relative intensities depending on the Boltzmann distribution in the vibronic states of the ground states and on the Franck–Condon factor, which is defined as the overlapping integral squared of the vibrational wave functions for the ground and excited states having displaced equilibrium points. The excited electron releases the extra vibrational energy into the surroundings within a time of several vibration periods (a few picoseconds). During this relaxation process, the electron may undergo spin-flip, and makes a radiationless jump (intersystem crossing) to the triplet excited state. Under the steady photoexcitation, excited molecules find themselves near the bottom of either the singlet excited-state curve or the triplet excited-state curve. The radiative decay from the singlet excited state is observed as fluorescence with a fast decay (a few nanoseconds or less), while that from the triplet excited state as phosphorescence with a slow decay (a few milliseconds or more). The phosphorescence band is considerably red-shifted from the fluorescence band, which is usually not the case in inorganic phosphors. The excitation spectra for the phosphorescence are often useful in detecting the triplet excited state and other forbidden excited states, which are hard to observe directly by absorption.

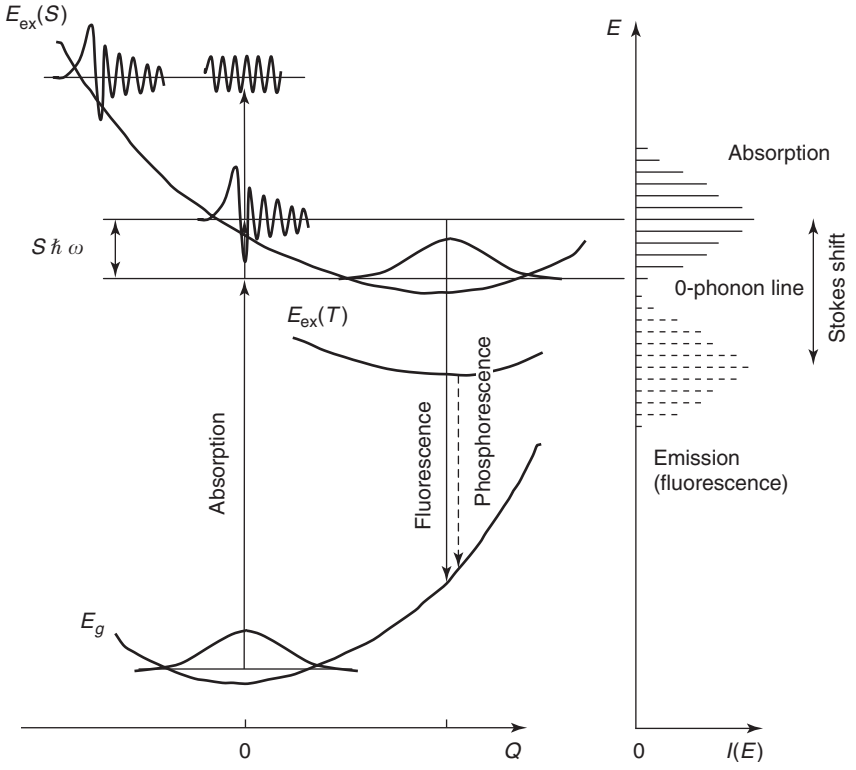


Figure 3 CC model for optical transitions in a conjugated molecule; the curves denoted E_g, $E_{ex}(S)$, and $E_{ex}(T)$ are the CC curves for the ground state, singlet, and triplet excited states, respectively. On the right, vibronic structures in the absorption and fluorescence bands are schematically shown. Those for the phosphorescence band appearing in the lower energy region are not shown.

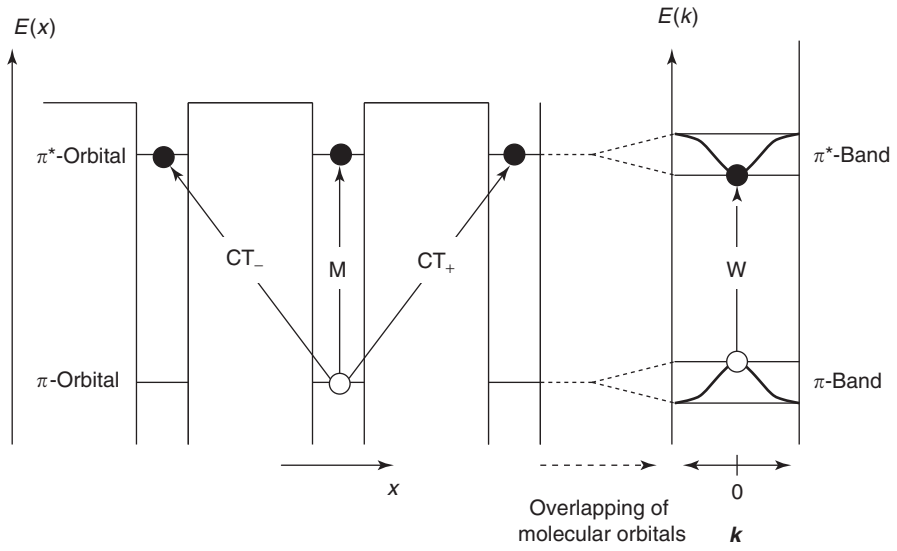


Figure 4 Schematic illustrations for elementary excitations in molecular exciton (M), charge-transfer exciton ($CT_{\pm}$), and Wannier (W) exciton. The transnational motion of respective excitons is not depicted. The filled and open circles represent electrons in the LUMO π^*-orbital (or conduction band) and holes in the HOMO π-orbital (or valence band), respectively.

According to the adiabatic (Born–Oppenheimer) approximation, the absorption spectrum at sufficiently low temperatures is composed of a series of narrow lines (o, n), corresponding to the transitions from the lowest vibronic state ($m = 0$) of the ground state to the vibronic states $n \geqslant 0$ of the singlet excited state, m and n being the vibrational quantum numbers for the ground and excited states. The relative intensity for the (o, n) components is given by the Franck–Condon factor. The result is expressed by the Poisson distribution, $I(0, n) = e^{-S}S^n/n!$, provided that both states have the same vibration frequencies. Here, S is the Huang–Rhys factor, defined as the average number of vibration quanta emitted in the thermal relaxation process after photoexcitation. In the strong vibronic coupling regime ($S \gg 1$), the absorption spectrum approaches a broad Gaussian shape, while in the weak or intermediate coupling regime ($S \leqslant 1$), the spectrum shows a sharp zero-phonon line which is followed by a series of equidistant vibronic lines on the higher energy sides. In the same way, the emission (fluorescence) spectrum at low temperatures is composed of vibronic components (m, o) approximately in a mirror-image of the absorption spectrum. The red-shift of the emission band peak from the absorption band peak is called the Stokes shift.

Exciton Concepts in Organic Crystals

Detailed optical studies of molecular crystals had started at the beginning of the twentieth century. Subsequently in the 1920s to the 1940s of the last century, further extensive studies had been made using extremely thin flakes of aromatic molecular crystals, until the conventional Kramers–Kronig analysis of reflection spectra became available in the late 1960s. From a careful comparison of absorption spectra in solutions and crystals, it had been recognized that some notable features show up when molecules are condensed into a crystalline structure. To interpret these features, a quantum mechanical picture of molecular excitons had been proposed upon the concept of the Frenkel exciton. Later on, more profound studies had been explored on the optical properties related with excitons in organic crystals. In contrast to the predominant character of 3D Wannier excitons in inorganic semiconductors, excitons in organic crystals and polymers have three different faces like Hekate, a goddess in ancient Greek mythology. They are named as molecular (Frenkel) exciton, charge-transfer (CT) exciton, and Wannier exciton, which are schematically illustrated in **Figure 4**.

Molecular (Frenkel) Excitons

The name "molecular exciton" comes from its origin in the intra-molecular $\pi - \pi^*$ excitation in the isolated molecule, while the Frenkel exciton is a theoretical concept of propagating excited states in a crystal as a linear combination of localized excitations. The character of the molecular exciton in a crystal is schematically illustrated in **Figure 5**. When a large number (N) of identical conjugated molecules are placed with sufficiently large separation from each other, the photoexcited states have the N-fold

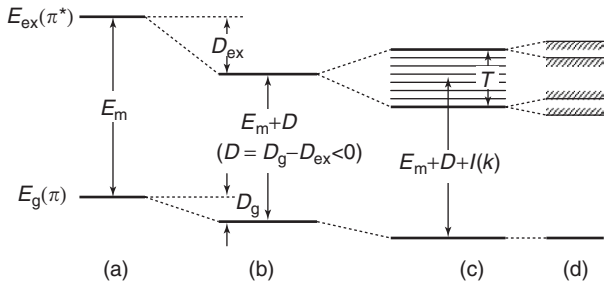


Figure 5 Schematic illustration of the energy levels for the ground and excited states in a conjugated molecule: (a) for an isolated molecule, (b) for the same molecule in solution, (c) and (d) for the excitons in the molecular crystal without and with the Davydov splitting. As for the various notations, see the text.

degeneracy with the same excitation energy E_m (**Figure 5a**). As the inter-molecular distance becomes smaller, the excitation energy is reduced to $E_m - D(>0)$, where D is the reduction in E_m due to the inter-molecular interaction (the site-shift effect). When molecules are condensed to a crystalline structure, the photo-excited state is no longer localized on a specific molecule and begins to propagate in the crystal due to the resonant energy transfer leading to formation of the Frenkel exciton band with $E(k)(= E_m - D + I(k))$.

So far, an ideal case of a rigid crystal has been considered, where all molecules are supposed to be frozen up at their equilibrium sites. Actually, the free excitons are significantly influenced by the dynamical fluctuation of potential due to the molecular vibrations. The vibronic spectra of such molecular excitons can be analyzed by the theoretical model called the dynamical coherent potential approximation (DCPA). Important parameters in this model are the exciton-bandwidth T, representing the delocalized nature of the exciton, and the Huang–Rhys factor S, representing the localized nature of the vibronic state. An example of theoretical absorption spectra by the DCPA model is shown in **Figure 6**. For simplicity, it is assumed here that the exciton is interacting with a single molecular vibration with the energy of $\hbar\omega$ by the Huang–Rhys factor S, which is taken to be unity. A series of spectra in **Figure 6** illustrate the tendency of exciton delocalization as T is increased. In the case of $T = 0$ (top of **Figure 6**), the absorption spectrum is that of an isolated molecule given by the CC model (**Figure 3**). As T increases, the intensities of phonon-side-bands are decreased, and an intense sharp line evolves at the low-energy side of the vibronic spectra. This absorption peak corresponds to the transition to the $k = 0$ edge of the exciton band under the influence of the exciton–polaron effect. It is interesting to note that such behavior of the absorption spectra

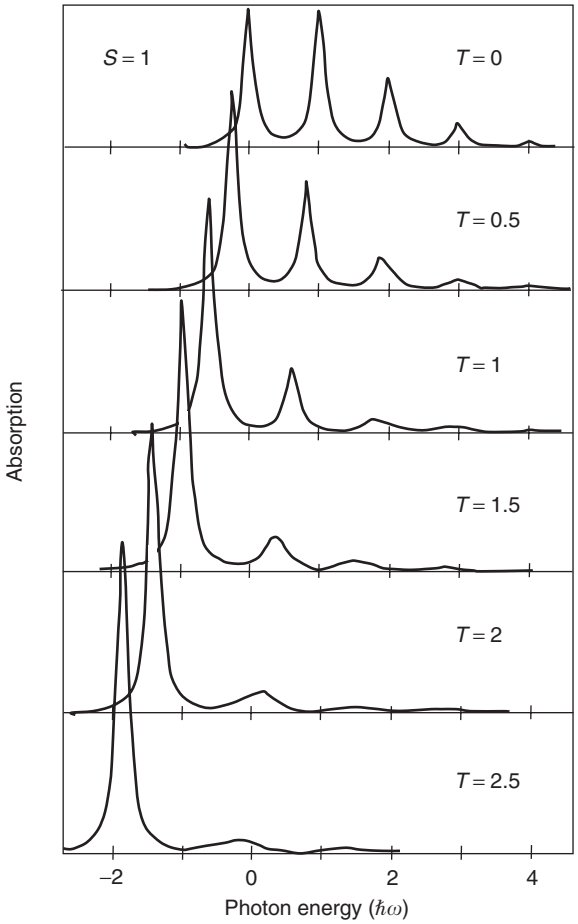


Figure 6 Theoretical spectra at a temperature of 0 K for a vibronic exciton interacting with a molecular vibration with the energy of $\hbar\omega$. The Huang–Rhys factor S is taken to be unity here, and the exciton band-width T is varied from $T = 0$ to 2.5 (in units of $\hbar\omega$). A semicircular curve with the radius of T is adopted for the density-of-state model. The $T = 0$ spectrum at the top corresponds to the localized exciton, those for $T = 0.5-2$ to the intermediate excitons, and that for $T = 2.5$ at the bottom to the nearly free exciton influenced by the polaron effect. The sharp peaks in the $T = 2-2.5$ spectra correspond to the optical transition to the $k = 0$ state of the exciton band. (Due to H Sumi.)

predicted by the DCPA is quite similar to the spectral change observed in the solutions of some organic dyes, where a single intense peak (called the J-band) appears on the low energy side of the broad molecular absorption band as the dye concentration is increased. The formation of microcrystalline aggregates and mobile molecular excitons therein are suggested to be responsible for this remarkable change.

To summarize, complicated optical properties of molecular excitons are interpreted as due to the competing trends of excitons, either being free or self-trapped. For a free exciton, the absorption intensity is concentrated on a single intense line, because very quick exciton motion tends to smooth out the local-molecular deformation (the motional narrowing).

When excitons are localized (self-trapped), the spectra become very broad and Stokes-shifted. There are a number of interesting optical phenomena related with the dynamical behavior of molecular excitons. For instance, it has been demonstrated that a fusion–fission process takes place between mobile singlet and triplet excitons in the following way: 2 triplet excitons $\rightarrow$1 singlet exciton in fusion, and vice versa in fission. In the former case, a delayed fluorescence component is observed with the same slow decay time as phosphorescence.

The prototype of molecular exciton spectra is shown in **Figure 7** for an anthracene crystal. There are two molecules at translationally nonequivalent sites, A and B, as shown in the inset. There can be two kinds of degenerate excitons propagating in the sublattices for the A and B sites. They are actually split into two polarized components due to the interaction between excited molecules on the A and B sites (the Davydov splitting). The transition dipole moment vectors for the Davydov components are suitable linear combinations of the transition moment vectors of the A and B molecules depending also on the wave vector of the exciton. The characteristic polarization dependence of the exciton spectra in **Figure 7** are due to the Davydov splitting of vibronic structures, which correspond to the intermediate coupling regime in the DCPA model (**Figure 6**). The Davydov splittings are also observed in the triplet exciton located at $\sim$1.8 eV. The Davydov splitting energies are 220 cm^{-1} (27 meV) and 21.5 cm^{-1} (3 meV), respectively, for the lowest singlet and triplet excitons. Extensive studies have been made on the various aspects of the molecular exciton in an anthracene crystal, including the fission–fusion process between the singlet and triplet excitons, the exciton polaritons, the surface excitons, and the exciton dynamics in the mixed crystals with deuterium-substituted anthracene molecules and so forth. On the higher energy side of the vibronic exciton bands, there are a series of weak features due to the intermolecular CT excitons, as mentioned in the next section. They are supposed to be responsible for the photocarrier generation process.

Aside from standard aromatic molecules shown in **Figure 1**, there are many important families of molecular crystals, such as phthalocyanines and other cyanine dyes. The molecular excitons in these crystals have a very large oscillator strength ($f \geqslant 1$), which is responsible for the remarkable metallic reflection bands due to the wide polariton-gap. Excitonic states are also supposed to play an important role in the primary stage of the photosynthesis process within the regularly arrayed chlorophyll molecules.

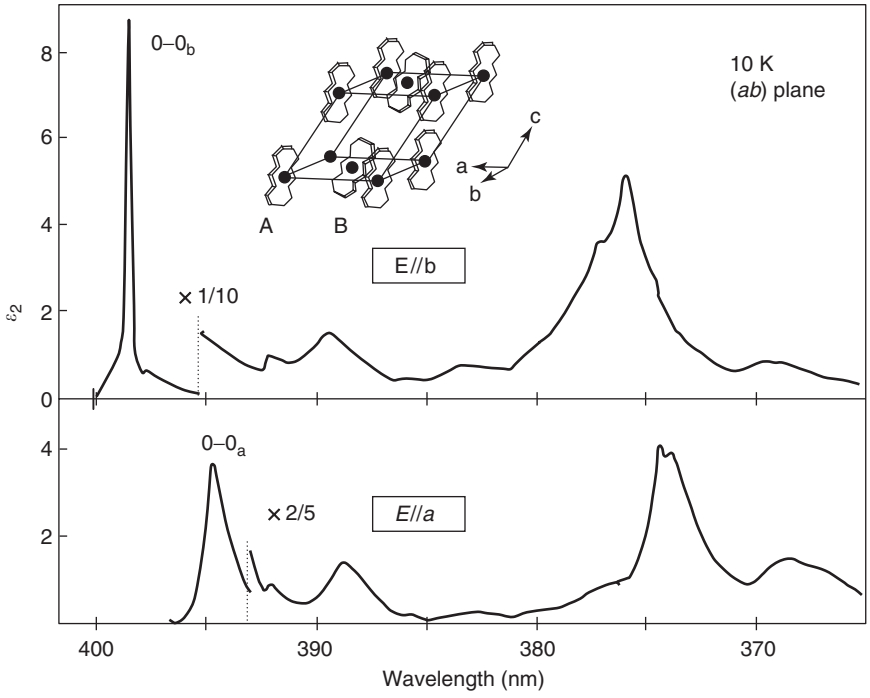


Figure 7 Polarized absorption spectra of anthracene crystal, shown by the imaginary part of complex dielectric constant $\tilde{\varepsilon}$ obtained from the Kramers–Kronig analysis of polarized reflection spectra on the (*ab*) plane at 10 K. The vibronic structures are due to the molecular excitons showing Davydov splittings due to the interaction between excited molecules in the sublattices A and B (see the inset). (Due to Y Tokura and T Koda.)

Charge-Transfer Excitons

In the CT complex crystals, the optical spectra are predominated by the CT excitations, whereas the molecular excitons play only a minor role. According to the characters of CT excitations, the CT crystals are further classified into the following three groups (**Figure** 8); (1) metal–tetracyanoquinodimethane (TCNQ) complex M^+–$TCNQ^-$ (M = K, Rb, Cu, Ag…), (2) halogen-bridged mixed-valent metal–organic complex $-M^{2+}-X-M^{4+}-X-M^{2+}-$ (M = Pt, Pd, Ni; X = Cl, Br), and (3) mixed-stack donor–acceptor (DA) complex, including anthracene–PMDA, TTF–chloranil(CA), etc. In the group (1) crystals, the electron transfer is schematically written as $TCNQ^- \cdot TCNQ^- \rightarrow TCNQ^0 \cdot TCNQ^{2-}$. This needs a finite excitation energy, ~ 1.0 eV in K–TCNQ, due to the Coulomb repulsion between the two electrons on the $TCNQ^{2-}$ molecule. For this CT process to occur, the crystal must have antiferromagnetic spin ordering on the anion sites. Such a spin-ordering occurs below T_c (~ 400 K in K^+–$TCNQ^-$), where the crystal undergoes a structural phase transition (the spin-Peierls transition) into the dimerized antiferromagnetic phase. Similar instabilities in the crystal structure are also observed in group (2) and (3) crystals due to the 1D character of the electron-lattice combined system (the Hubbard system).

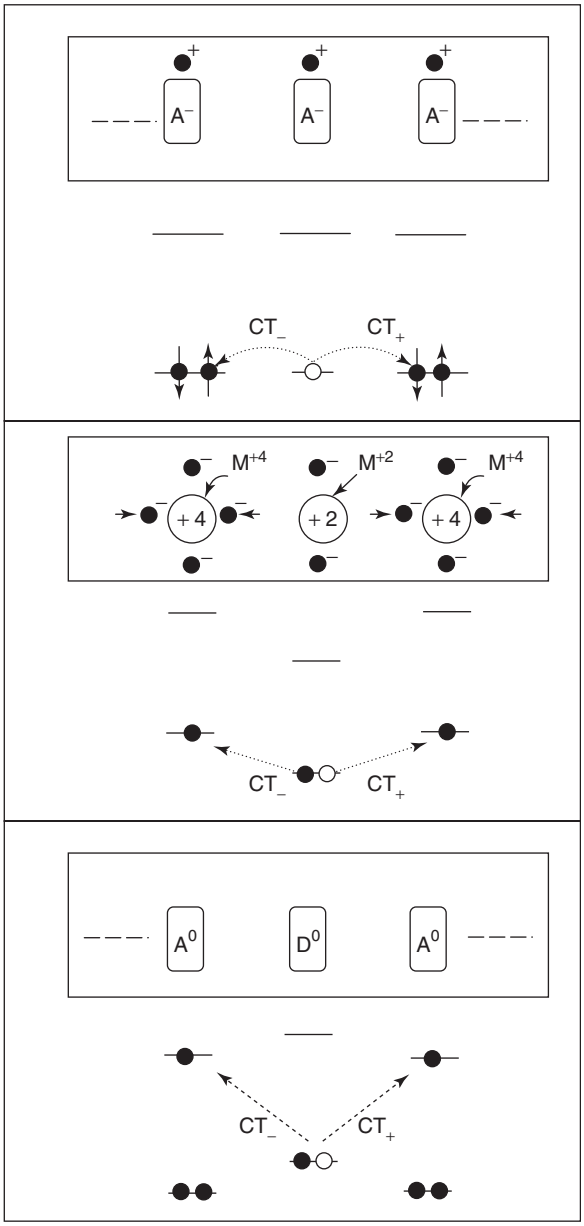


Figure 8 Schematic illustration of elementary CT excitations in three kinds of CT crystals; the top for the family of metal(monovalent)–TCNQ complex crystal, the middle for the halogen-bridged mixed-valence metal complex crystal, and the bottom for the mixed-stack donor (D)–acceptor (A) complex crystal (in the neutral phase denoted D^0A^0). The elementary CT exciton transitions are shown by arrows with dotted lines. All of these CT excitons have a quasi-1D character.

The CT exciton bands are strongly polarized along the 1D stacking axis of molecules in sharp contrast to the intra-molecular absorption bands which are polarized in the directions specified by the molecular axes. In the elementary CT process, an electron is transferred to the adjacent molecules with equal energies for the back- and force-CT excitations (CT_+, CT_-), when the crystal is not dimerized. The eigen CT states are the symmetric and antisymmetric combinations of CT_+ and CT_- states, one being optically allowed and the other forbidden. The photogenerated CT state is unstable against the local deformation and is apt to be self-trapped by the dimerization of adjacent molecules. The forbidden CT state which is not observable in ordinary absorption spectra becomes optically active, either by applying an electric field or by the lattice dimerization below T_c. It is suggested that the forbidden CT state plays an important role as an intermediate state for the large third-order nonlinear optical response of some CT crystals with $\chi^{(3)} \sim 10^{-9}$ esu ($\sim 1\,\mathrm{nm^2\,V^{-2}}$ in MKS). When the CT bands are close to the molecular exciton bands, the intensity of CT transitions is considerably enhanced. This "intensity-borrowing" effect is responsible for the high photoconductivity in some molecular crystals, such as phthalocyanine and other organic photoconductors.

As a prototype of CT crystals, a mixed-stack complex crystal composed of anthracene (An) and PMDA, which belongs to the group (3) CT crystal with a weak degree of CT of $\rho \sim 0.05$ is chosen somewhat arbitrarily. Needle-like crystals of An–PMDA show a bright red color, though the starting materials in the vapor-phase reaction are white powders of An and PMDA. This bright red color of the

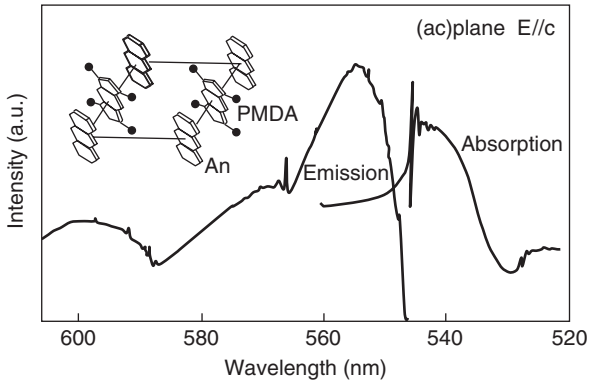


Figure 9 The absorption (obtained by Kramers–Kronig transformation) and emission spectra of An–PMDA crystal at 2 K, polarized with *E*//*c*. The *c*-axis is parallel to the mixed-stack molecular column and is the direction of preferential polarization for both spectra. The sharp absorption line at ~546 nm is due to the 0–0 component of vibronic structures. Broad and sharp features in the emission spectra are due to the free and self-trapped CT excitons coexisting at low temperatures (see the text). (Due to Y Tokura and T Koda.)

crystal is due to the strong CT band at ~540 nm, which is preferentially polarized along the stacking axis (*c*-axis) (**Figure 9**). A sharp absorption peak at ~546 nm is the zero-phonon (0–0) component of the vibronic series $(0-n)$. The emission band is not mirror-symmetric to the absorption band, indicating that the emission is due to the radiative decay of the self-trapped CT exciton. However, when the temperature is lowered near 2 K, a weak 0–0 emission line appears at the energy of 0–0 absorption line, indicating that a part of photo-generated CT excitons are moving freely in the crystal before being self-trapped. This is a rather rare case, and usually the CT excitons are promptly self-trapped after the photogeneration. In An–PMDA and other CT crystals, a weak phosphorescence band due to the radiative decay of the triplet CT excitons is also observed. They are self-trapped too by local displacement of molecules.

In other mixed-stack CT crystals with larger ρ, the CT exciton bands are shifted to the near infrared region. Hence, these crystals are not so colorful as the An–PMDA crystal. Nevertheless, they do exhibit even more interesting characteristics in the dielectric, magnetic, and structural properties, which are related with the so-called neutral–ionic (NI) instability. When the degree of CT approaches a certain critical value, the crystal ground state becomes unstable and undergoes a phase-transition from the neutral (N) phase to the ionic (N) phase. In the vicinity of the NI phase boundary, a number of novel phenomena, such as reversible photochromism, bistable electric conductivity, antiferroelectric ordering, appearance of spin- and charged-solitons, and so forth, have been observed in several mixed-stack CT crystals, in particular in TTF–chloranil crystals. Another important type of CT complex crystal is the family of segregated-stack CT crystals, such as TTF–TCNQ and BEDT–TTF–TCNQ, showing high (metallic) conductivity and even super-conductivity. Interdisciplinary studies have been extended on these "synthetic metals," forming an important field of condensed matter physics.

Wannier Excitons

The Wannier exciton is quite familiar in inorganic semiconductors, but not so much in organic molecular crystals, because of the localized nature of the π, π^* orbitals. In the case of the CT bands in anthracene crystals, spectral features detected by the electromodulation spectroscopy are assigned to the Rydberg series $E_{CT} = E_G - e^2/(4\pi\varepsilon R_{CT})$, R_{CT} being the supposed distance for the CT jump. As the limit of this series, "the bandgap energy" E_G has been estimated to be $E_G = 3.2$ eV for anthracene crystals. But this picture of CT states is considerably different from the Wannier exciton picture in the one-electron band regime, and is more properly represented by a pair of localized π^*-electron and π-hole, each located on the separated sites.

The most typical Wannier exciton in organic solids may be found in organo-silicon polymers (polysilanes) with a structure of [–RSi–SiR′-], R, R′ being organic side-group molecules. They are regarded as the σ-conjugated analog of the π-conjugated polymers, having the HOMO (σ) and LUMO (σ^*) orbitals extended along the conjugated polymer chains. Taking advantage of much easier preparation of oriented thin films by spin-coating, extensive studies have been made on the ordinary absorption, electroabsorption, two-photon excitation of luminescence, third-harmonic generation (THG) and electric-field-induced second-harmonic generation (EFISHG). Based on these experimental results, a theoretical model of quasi-1D Wannier exciton has been constructed (**Figure 10**). This model is essentially based upon a 1D-hydrogen model given by,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} - \frac{e^2}{4\pi\varepsilon_0|x|}\right)\psi_n(x) = E_n\psi_n(x)$$

where $x = x_e - x_h$, x_e, x_h being the electron and hole coordinates. Exciton states for $n = 1,2,3$ have been calculated assuming a finite potential at $x = 0$ due to the on-site Coulomb repulsion. According to this model, the exciton wave functions $\psi_n(x)$ are spread over a considerable range of the Si-chain (shown on the top of **Figure 10**). Various linear and nonlinear

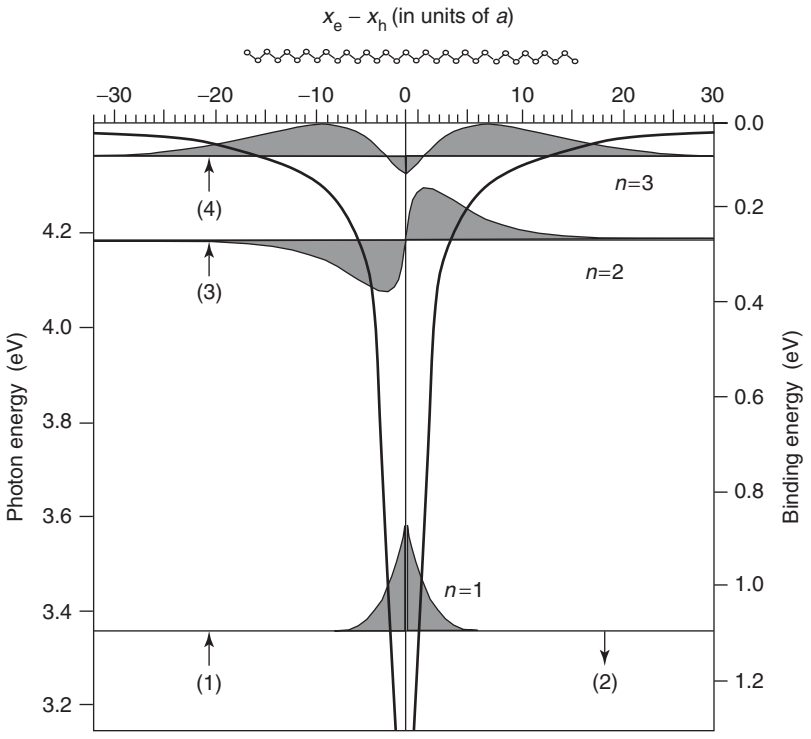


Figure 10 The quasi-1D Wannier exciton model for a polisilane (trans-polydihexylsilane: PDHS). The abscissa is the electron–hole distance along the polymer chain in units of the lattice constant (a= 2.0 Å). The energies shown on both sides of ordinates represent the photon energy from the ground state (left) and the exciton binding energy from the band edge at E_g= 4.43 eV (right). The on-site energy U at the origin ($x = 0$) of the 1D potential (solid curve) is taken as U= −5.18 eV from E_g. The envelope functions for the n = 1,2,3 excitons are indicated by curves surrounding shadowed areas. The upward arrows indicate the resonant transitions; (1) for the one- and three-photon absorptions, (2) for emission (fluorescence), (3) for two-photon absorption and electric-field-induced SHG, and (4) for three-photon absorption, respectively. The multi-photon absorptions were measured by the excitation spectra for fluorescence (2). On the top, a Si-chain structure of *trans*-PDHS is depicted. (The organic side-groups are omitted). (Due to T Hasegawa, T Koda, *et al.*)

optical transitions are indicated by vertical arrows. Observed spectra for respective transitions are in very good agreement with the theoretical results calculated by this model, indicating that trans-PDHS is an ideal molecular quantum wire having the band-gap E_g ~4.4 eV and the exciton binding energy as large as 1.1 eV.

As an example of Wannier exciton spectra in π-conjugated polymers, **Figure 11** shows the experimental results for a PDA–TS crystal obtained by the solid-state polymerization of a crystal of the diacetylene monomer R–C≡C–C≡C–R with the side group R = toluenesulfonate (TS), $-SO_3C_6H_4CH_3$. The monomer crystal is transparent in the visible region showing a pale pink color due to a small amount of short conjugated chains (origomers) generated by room light. The absorption bands (1 ~ 4) in the higher energy region are due to the $\pi-\pi^*$ transitions in the DTS monomer. After polymerization, the crystal surface turns to show a brilliant golden color in reflection. This is due to the wide polariton-gap associated with an intense absorption line at ~2 eV which is strongly polarized along the backbone chain of a conjugated polymer. This absorption line is assigned to the photogeneration of an electron–hole pair in the quasi-1D π^*-conduction and π-valence bands. Detailed studies have been made so far to elucidate the nature of the exciton states in PDA–TS and other PDAs as well by means of various linear and nonlinear spectroscopies. Similar quasi-1D Wannier exciton models are also likely to be valid in other π-conjugated polymers such as PPV (polyphenylene-vinylene). As for trans-polyacetylene, however, there is still controversy on whether the 1D Wannier exciton or the band-to-band transition is responsible for the absorption peak. It is likely that the 1D character of the π-electron system in polyacetylene is considerably modified by the 3D interchain interaction because of the very close distance between the polymer chains.

Organic Nonlinear Optical Materials

The nonlinear optical (NLO) response of organic materials is due to the fairly large anharmonicity in the

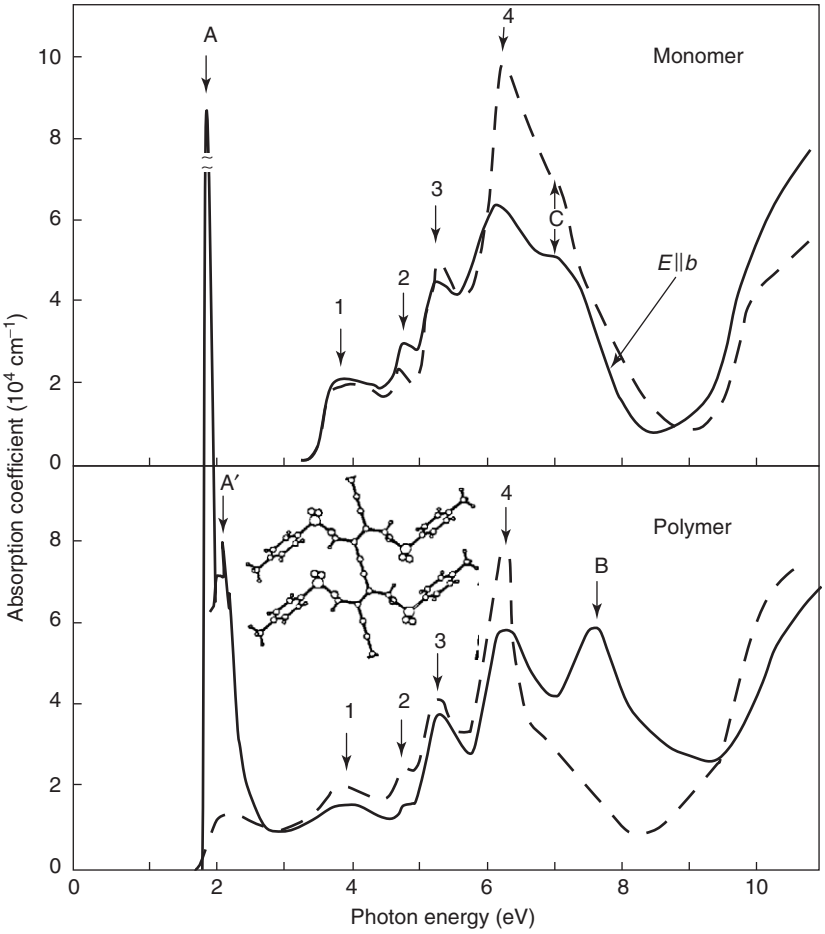


Figure 11 Polarized absorption spectra of monomeric and polymeric crystals of diacetylene (DA) with the side-group (TS); the upper part for the monomer crystal (DATS) and the lower part for the polymer crystal (PDA–TS) after thermal polymerization of the monomer crystal, both obtained by the Kramers–Kronig analysis of the reflection spectra at RT. The solid and dashed curves represent the polarized components with $E \| b$ and $E \perp b$, respectively. Absorption peak A and probably the band B are due to the Wannier excitons associated with the extended π and π^* bands on the conjugated polymer chains oriented along the *b*-axis, while the broad absorption bands (1–4) are due to the $\pi-\pi^*$ transitions within the side-group molecules. The band C in monomer crystal is assigned to the $\pi-\pi^*$ transition in the conjugated bonds of diacetylene monomer. (Due to Y Tokura, T Koda, *et al.*)

photoinduced polarization of the π-electrons, which are delocalized in the low-dimensional structures of conjugated molecules and polymers. This makes NLO spectroscopy a powerful tool for the studies of organic materials. Among others, the NLO spectroscopy for quasi-elastic and inelastic light scattering processes, such as Raman and Brillouin scattering, have been extensively employed in the studies of the structure and dynamics of organic materials. In addition, the NLO spectroscopy provides an indispensable means of detecting the forbidden electronic excited states in organic molecules and polymers. An example was shown before for the quasi-1D Wannier exciton in a σ-conjugated polymer (**Figure 10**). Another example is the two-photon excitation spectroscopy for the optically forbidden 2^1A_g state in finite linear polyenes. Also electro-absorption (reflection) spectroscopy, which is essentially due to the third-order NLO process (Kerr effect), has been widely used to detect the symmetry-forbidden excited states in conjugated polymers and CT crystals. The large third-order electric susceptibility $\chi^{(3)} \simeq 10^{-9}$ esu ($\sim 1\,\mathrm{nm^2\,V^{-2}}$ in MKS) in some low-dimensional organic materials is deemed to arise from the resonant intermediate transition between the closely separated allowed and forbidden excited states. Regarding the second-order NLO phenomena, measurement of the second-harmonic generation (SHG) from centrosymmetric organic molecules is useful for a sensitive probe to the local distortions near the surface, interface, and in other kinds of anisotropic environments which activate normally forbidden SHG of the molecule by breaking the symmetry of the local field. By scanning the laser beam, one can obtain an *in situ* image of the microscopic anisotropy in the structures.

The NLO properties of organic materials are also attractive for feasible applications in opto-electronics. The hope conceived there is multiple: the large and ultrafast NLO response, the robustness against irradiation damage, the on-demand designability of efficient organic NLO materials, and so on, which are promising for the fabrication of, for instance, optical switches or modulators in optical IC devices and the blue-emitting LEDs by efficient up-conversion of the near IR light from semiconductor lasers by the SHG process. It has been shown that, by introducing suitable bulky side-groups at both ends of conjugated molecules, the second-order electric susceptibility $\chi^{(2)}$ can be considerably enhanced owing to the push–pull effect on the distribution of the delocalized π-electrons. Many kinds of newly synthesized NLO molecules are to be found in literature. Some of these organic NLO materials may be brought into the spotlight in future technology. To bring this dream into reality, closer multidisciplinary collaborations will be essential.

Liquid Crystals

Liquid crystals (LCs) will be the most familiar functional organic materials used in various hi-tech products surrounding us. The history of LCs goes back to the discovery of a curious optical property in cholesterol. Below a certain critical temperature T_c, cholesterol shows an optical anisotropy (birefringence) like a crystal, despite the appearance of a cloudy fluid. Due to such an intermediate character, the name of "liquid crystals" was given to these novel materials. Ever since, steady and patient efforts have been devoted to the basic studies of LCs, until the boom came in 1970s to develop LC panels for watches, calculators, and so forth. The research and development of LC technology may be one of the most happy marriages between scientific curiosity and technological quest supported by active multidisciplinary collaborations.

The LC molecules are characterized by the large anisotropic shapes with either rod-like or disk-like contours. The key parameters specifying the liquid-crystalline order are the director, that is, the unit vector representing spatial orientation of the molecule, and the center position of each molecule. To elucidate the physical nature of characteristic behaviors of liquid crystals, efforts had been made early in the 1930s and the 1940s by a Russian group, and more lately by de Gennes *et al*. in the 1970s. The basic principle in most LCDs is the external control of the directors with an electric field. An LC film sandwiched by a pair of polarizers acts as an optical valve with a low working voltage, leading to yet growing applications of low-power LCDs to flat panel displays, mobile computers, LC projectors, and so on. Aside from such commercial applications, the optical properties of LCs are important and interesting subjects in soft-matter science. The NLO spectroscopy of LCs is another subject of current interest. Various NLO phenomena such as harmonic generations, optical wave-mixing, self-focusing, and the external control of these NLO properties by an electric field are being extensively investigated in LCs. Also, mixed systems of a liquid crystal and a solvent (called lyotropic liquid crystals) are interesting as a mimic biomembrane system.

Organic Electroluminescence Devices

An important application of optical properties of polymers and organic compounds, next to LCDs, will be the organic light emitting devices (organic EL or organic LED). After extensive R&D efforts in the last two decades, the efficiency, the lifetime, and color performance of these devices have been improved far beyond previous expectations even by the experts. The pioneering challenges to organic EL devices were initiated in the early 1980s to meet the increasing need of a flat display for laptop computers and other hi-tech devices. However, the way to the state-of-the-art organic EL devices has been quite hard, because of the severe working conditions such as very high electric current and harmful influence of oxidizing atmosphere. A number of breakthroughs must have been made, until an innovative idea was presented for the fabrication of a *pn*-junction type multilayer structure. The point of this idea is to split the device into essentially three separate regions, each taking part in the transport of electrons and holes and in their efficient recombination process, so as to attain better performance by means of the optimum design of the materials and device structure.

A typical structure of the EL cell is schematically illustrated in **Figure 12**. The cell consists of three layers; an Alq (an Al-chelate compound) layer, another Alq layer doped with phosphorescent $Ir(ppy)_3$ (an Ir-chelate compound of a few %), and a TPD (an aromatic diamine molecule) layer. These layers are sandwiched between the anode made of a transparent conducting film of indium–tin oxide (ITO) and the cathode made of an Mg–Ag alloy metal film. When an AC bias voltage of ~ 10 V or less is applied across the electrodes, carrier injection takes place from both electrodes into the cell: the electrons from the metal cathode with a low work-function and the holes from the ITO anode with a larger work-function. The electrons and holes are transported, respectively, across the Alq layer and the TPD layer, and

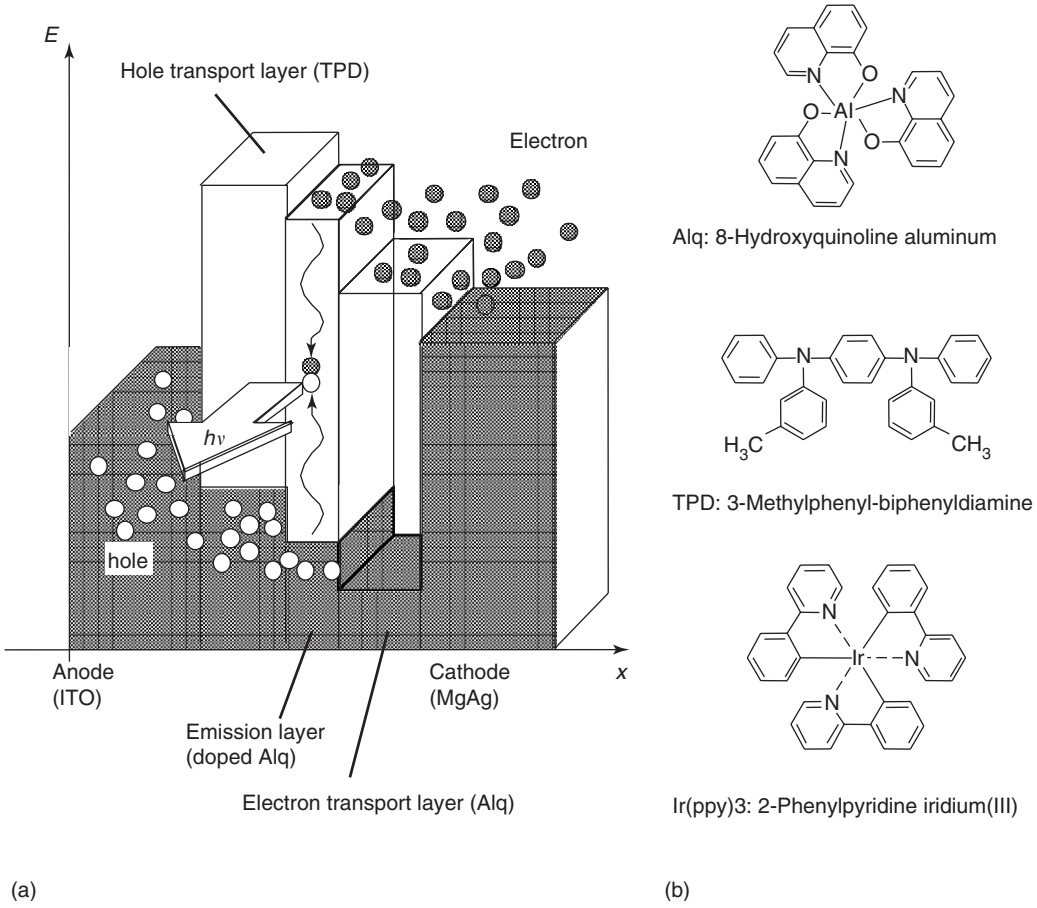


Figure 12 Schematic illustration of a typical organic electro-luminescent (EL) cell and the structures of component molecules. The EL cell shown in (a) has a *pn*-junction-type multilayer structure, composed of the transparent anode (an ITO film), the hole-injection layer (a TPD layer), the emission layer (an $Ir(ppy)_3$-doped Alq layer), the electron-transport and emission layer (an Alq film), and metal cathode (an Mg/Ag alloy film), each ~10 nm thick. The structures and the names of the component molecules (abbreviated above as Alq, TPD and $Ir(ppy)_3$) are shown in (b). The EL light is composed of two components; one being fluorescence due to the recombination of the singlet excitons in the Alq molecules and the other being phosphorescence by the recombination of the triplet excitons trapped in the $Ir(ppy)_3$ molecules. The EL light (shown by an arrow in (a) for illustration) is actually emitted out of the cell through the ITO layer vertically to the layer surface. See the text for details.

then are blocked by the layer interface forming a space charge layer near the boundary. A part of holes in the TPD layer goes through the potential barrier by tunneling into the doped Alq layer. They are bound by the electrons on the Alq molecules near the interface and form molecular excitons in either singlet or triplet state with a ratio of 1:3. These excitons diffuse into the doped Alq layer and further to the Alq layer. By the radiative decay of the singlet excitons, fluorescence is emitted from the Alq molecules, while the triplet excitons are trapped by the $Ir(ppy)_3$ molecules and are efficiently converted to the phosphorescent light due to the enhanced singlet–triplet mixing in the iridium-ligand. Other kinds of organic dye molecules can also be incorporated in the doped-Alq layer in order to give full color performance of the EL cell. As a result, the net efficiency of the EL is now approaching the level of practical light sources not only for displays but also for illumination.

Apart from these multilayer EL cells, diode-type organic EL devices (organic LEDs) have been also developed using thin films of PPV as the active layer. This type of organic EL devices has an advantage of application to the flexible and large-area EL films. In conclusion, the future scope of both types of organic EL devices is quite promising, complementing the established technology of LCDs.

Conclusion

To conclude this article, a Haiku is quoted: "Oh, how many of them are there in the fields! But each flower in its own way— In this is the highest achievement of

flower!" (Matsuo Basho, 1644–1694, translated by Maradudin A A). Simplicity and diversity may be the key words for the attractive world of polymers and organic compounds, and of their optical properties in particular. For progress in the future, one will have to go beyond reductionism in traditional condensed matter physics. Last but not the least, the authors would like to express their sincere thanks to many researchers whose names should have been quoted for their excellent pioneering works concerning the subjects dealt with in this article.

See also: Electronic Structure (Theory): Molecules; Harmonic Generation Frequency Conversion; Luminescence; Nonlinear Optics; Optical Absorption and Reflectance; Optical Properties of Materials; Polaritons; Polymer Structures; Polymers and organic Materials, Electronic States of.

PACS: 42.55.Px; 42.70.Df; 42.70.Jk; 42.70.Nq; 71.20.Rv; 71.35.Aa; 71.35.Cc; 71.36.+c; 78.20.−e; 78.40.Me; 78.55.Kz; 78.60.Fi; 78.66.Qn; 82.35.Ej; 85.60.Jb; 85.65.+h

Further Reading

Broude VL, Rashba EI, and Sheka EF (1985) *Spectroscopy of Molecular Excitons*. Berlin: Springer.

Davydov AS (1971) *Theory of Molecular Excitons*. translated by Dresner SB, New York: Plenum.

de Gennes PG (1974) *The Physics of Liquid Crystals*. New York: Oxford University Press.

Farchioni R and Grosso G (eds.) (2001) *Organic Electronic Materials: Conjugated Polymers and Low Molecular Weight Organic Solids*. Berlin: Springer.

Khoo IC (1995) *Liquid Crystals: Physical Properties and Nonlinear Optical Phenomena*. New York: Wiley.

Kippelen B (ed.) (1999) *Organic Photonic Materials and Devices*. Society of the Photo-Optical Instrumentation Engineers.

Messier J, Kajzar F, and Prasad P (eds.) (1991) *Organic Molecules for Nonlinear Optics and Photonics*, NATO ASI Series E: Applied Sciences-Vol. 194. Dordrecht: Kluwer Academic.

Miyata S and Nalwa HS (eds.) (1997) *Organic Electroluminescent Materials and Devices*. Amsterdam: Gordon and Breach.

Nasu K (ed.) (1997) *Relaxations of Excited States and Photo-induced Structural Phase Transition*. Berlin: Springer.

Pope M and Swenberg CE (1999) *Electronic Processes in Organic Crystals and Polymers*. 2nd edn. New York: Oxford University Press.

Rashba EI and Strurge MD (eds.) (1982) *Excitons*. Amsterdam: North-Holland.

Scriciftci NS (ed.) (1997) *Primary Photoexcitations in Conjugated Polymers: Molecular Exciton versus Semiconductor Band Model*. Singapore: World Scientific.

Shinar J (2003) *Organic Light-Emitting Devices*. Berlin: Springer.

Toyozawa Y (2003) *Optical Properties in Solids*. New York: Cambridge University Press.

van Amerogen H, Valkunas L, and van Grondelle R (2000) *Photosynthetic Excitons*. Singapore: World Scientific.

Wolf HC (1959) *The Electronic Spectra of Aromatic Molecular Crystals*. In: Seitz F and Tumbull D (eds.) *Solid State Physics 9*, New York: Academic Press.

Polymers and Organic Materials, Electronic States of

E Zojer* and J-L Brédas, Georgia Institute of Technology, Atlanta, GA, USA

Introduction and Overview

In this article, the results of quantum-chemical calculations will be used to provide an introductory overview on the electronic structure of organic semiconductors. The focus is on phenylene-based materials since these form a highly attractive class of π-conjugated systems. These materials can be exploited in organic photodetectors, solar cells, transistors, and electronic circuits; their main application relates to organic light emitting devices and displays, where they serve as blue emitters. Apart from their practical relevance, polyphenylenes also serve as ideal model systems to discuss the electronic properties of organic semiconductors and conjugated polymers and allow the description of general concepts that can then be applied to more complex systems.

Here, a start is made from the simple benzene molecule to show how its molecular orbitals are related to the conduction and valence bands of the corresponding polymer, poly(*p*-phenylene) (PPP) (see **Figure 1** for the chemical structure of the corresponding ten-ring oligomer 10PP).

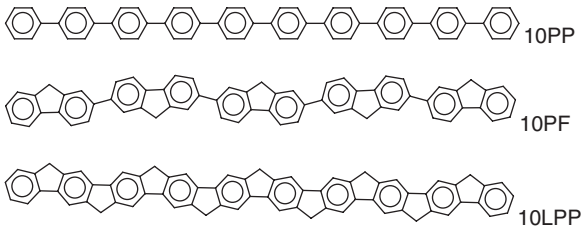


Figure 1 Chemical structures of the ten-ring oligomers of poly(*p*-phenylene) (10PP), poly(fluorene) (10PF), and ladder-type poly(*p*-phenylene) (10LPP).

*At the time of publication E Zojer was on leave from the Graz University of Technology.

In this way, a number of general properties of the excited states in phenylene-based materials can be derived. To bridge the gap between a molecular description usually found in quantum chemistry and the band structure picture prevalent in condensed matter physics, it will be shown how the molecular orbitals of phenylene chains of increasing length adopt a Bloch-type structure. A similar band-like behavior is found when considering the exciton states of extended π-conjugated molecules. To discuss the influence of molecular conformation, the electronic properties of "regular" poly(fluorene) – a partly bridged derivative of PPP – will be compared to those of its planarized conformation (the so-called β-phase) as well as of the fully bridged ladder-type PPP (see **Figure 1**). To address the role played by inter-chain interactions, the excited states of an oligophenylene dimer will be investigated. To complete the description of low-lying excited states in phenylene-based materials, excited states of different spin multiplicities will be compared. Finally, the impact of chemical defects on the electronic and optical properties of conjugated polymers will be examined by discussing the example of the ketonic defects in poly(fluorene)-type materials.

Basic Electronic Structure and Band Structure Description

To understand the electronic structure of phenylene-based materials, it is useful to start with discussing the properties of benzene rings that constitute the fundamental building blocks of these systems. (In the course of this discussion, D_{6h} symmetry is assumed for benzene.) The two highest-lying occupied and two lowest-lying unoccupied π-orbitals of benzene are shown in the left part of **Figure 2**. In a linear combination of atomic orbitals (LCAO) picture, these orbitals can be regarded as linear combinations of the p_z orbitals of the six sp^2-hybridized carbon atoms forming the ring (note that, as a result, there are two more molecular π-orbitals in benzene; their energies are, however, so different from the highest occupied molecular orbit (HOMO) and lowest unoccupied molecular orbit (LUMO) energies that they usually

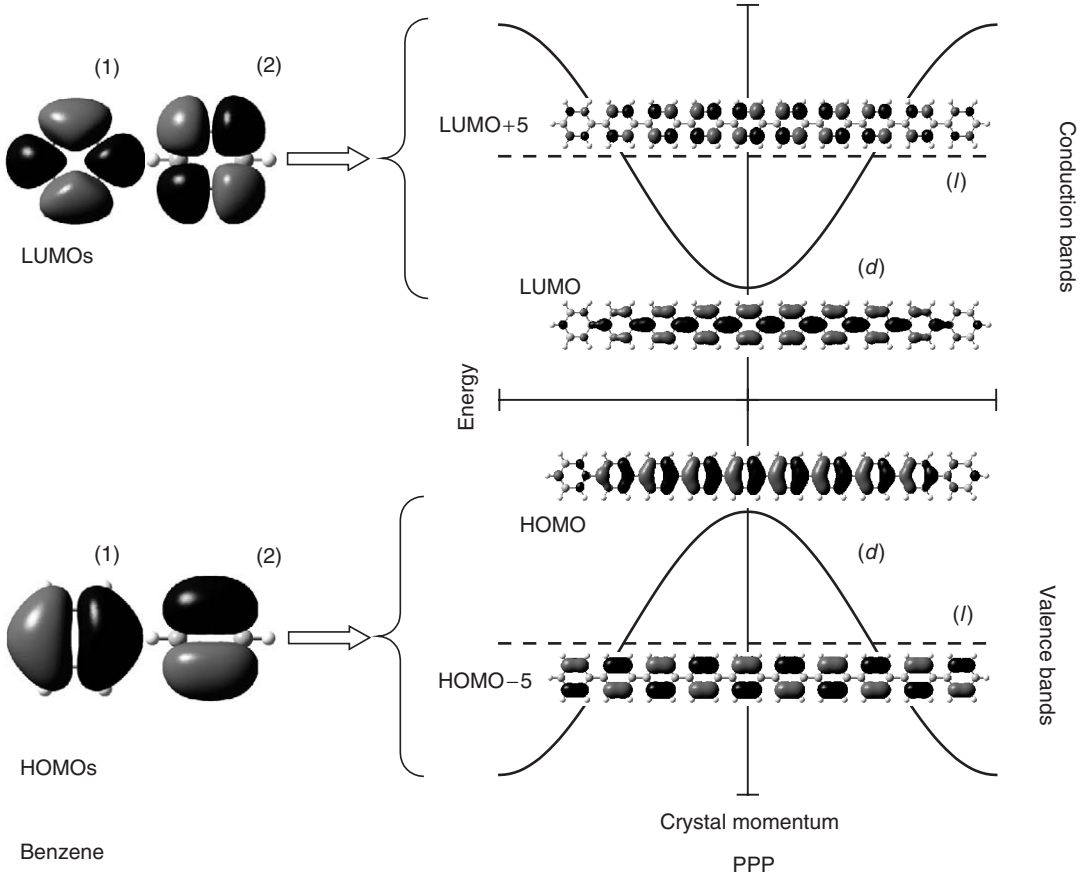


Figure 2 Left: Doubly degenerate HOMO and LUMO orbitals of benzene. Right: Hückel-type dispersed (solid lines) and nondispersed (dashed lines) conduction and valence bands of poly(*p*-phenylene). The insets show the localized and delocalized orbitals in a ten-ring oligomer, which can be associated with the four different types of bands.

do not influence the electronic and optical properties of interest).

Although the two HOMOs and LUMOs of benzene are degenerate (**Figure 2**) the molecular orbitals that derive from them display distinctively different properties when several phenylene units are linked in *para*-position to form oligo- or poly-phenylene chains. As an illustration, the HOMO and HOMO-5 of a coplanar ten-ring oligomer (10pPP) are shown on the right side of **Figure 2**. In the context of a Hückel (tight-binding) approach, the HOMO of 10pPP can be regarded as a linear combination of the benzene HOMOs of type-(1) (see **Figure 2**), while the HOMO-5 of 10pPP is a linear combination of type-(2) benzene HOMOs. The one-electron eigenstates of an infinitely long, periodic PPP chain that are similar to the HOMO of 10pPP, form the strongly dispersed valence band shown as a solid line in the Hückel-type band structure of **Figure 2**. Eigenstates like the HOMO-5 build the flat band depicted as a dashed line in the band structure. The reason for this different behavior is that for the type-(1) benzene HOMO, one finds high electron density at the carbon atoms which link the individual phenylene units to their neighbors. In contrast, the electron density on the connecting carbons vanishes for the type-(2) HOMO of benzene. Therefore, in the former case, the individual rings interact strongly in a polymer chain, while in the latter case there is only very little interaction. A strong interaction results in a large energy splitting and consequently in a strongly dispersed (wide) band; one then speaks of delocalized (d) orbitals/electronic states. The orbitals/electronic states derived from the type-(2) HOMOs are referred to as localized (l). The same reasoning also applies to the unoccupied orbitals (conduction bands) shown in the top part of **Figure 2**.

Some basic aspects regarding the optical properties of phenylene-based materials can be derived from the orbitals and the band description discussed above: The lowest energy transitions involve excitations between delocalized bands ((d)–(d) transitions) and are polarized parallel to the chain axis. Their energetic positions are strongly influenced by parameters like conjugation length and molecular conformation (*vide infra*). Experimental investigations of compounds in solution as well as quantum-mechanical calculations yield energies for the lowest optical transitions (absorption maximum) ranging from around 5.0 eV for biphenyl to about 3.7 eV for PPP (note that the latter value is extrapolated from the results obtained for finite-size systems, assuming a linear dependence of the excitation energy on the inverse number of repeat units). Transitions from delocalized into localized [(d)–(l)] and from localized into delocalized bands [(l)–(d)] are found at higher energies (around 6.2 eV according to inelastic electron scattering investigations on sexiphenyl thin films). Their dependence on the conjugation length is less pronounced than for the (d)–(d) excitations; for coplanar oligomers, they are polarized within the plane of the molecule along the short molecular axis. At higher energies (around 7.2 eV in sexiphenyl), one finds excitations involving transitions among localized levels [(l)–(l)]; these are polarized again along the long molecular axis.

The respective polarizations of these excitations can be easily understood from the symmetries of the molecular orbitals when recalling that the electric transition dipole M_{ge} for excitation from the ground state Ψ_g to the excited state Ψ_e is given by $\boldsymbol{M}_{ge} = -e\langle \Psi_g|\boldsymbol{R}|\Psi_e\rangle$. Taking (1) the center of the molecule as the origin of the coordinate system; (2) the long molecular axis as the x-axis; and (3) the molecular plane as the xy-plane, all delocalized orbitals are symmetric with respect to the xz-mirror plane. Therefore, the M_{ge} matrix element will be identically zero for a y-polarized electric field. The same applies to the (l)–(l) transitions, as here all orbitals are antisymmetric with respect to the xz-mirror plane; it also holds for z-polarized light, as all π-orbitals are antisymmetric with respect to the molecular plane. For x-polarized light and a (d)–(d) or (l)–(l) excitation, whether the M_{ge} matrix element vanishes or not, depends on the change of the orbital symmetry when going from an occupied to an unoccupied orbital. According to **Figure 2**, the HOMO is antisymmetric with respect to the yz-mirror plane while the LUMO is symmetric. Thus, the lowest-lying excited state, which is dominated by a HOMO→LUMO excitation, is optically allowed for x-polarized light. Similar arguments can be applied to show that d–l excited states can be optically allowed only for light polarized in the y-direction.

From Molecular Eigenstates to Continuous Bands

As becomes clear from the discussion, the description of π-conjugated organic materials is at the interface between condensed matter physics, molecular physics, and physical chemistry. Consequently, terms derived from each of these fields are often used simultaneously and sometimes synonymously. To bridge the gap between, on the one hand, molecular eigenstates and, on the other hand, bands, which are continuous in k-space, it is useful to consider oligomers with a relatively large, but finite, number of repeat

units. These can be described in the framework of the so-called quasiband structure model, which has been studied both theoretically and experimentally via inelastic electron scattering. The basic idea of the quasiband structure model in finite-size systems is to distribute the discrete molecular levels in k-space along the continuous bands of the parent polymers. As the crystal momentum is no longer a good quantum number for a system with finite size, a given range of $\boldsymbol{k}$-vectors has to be associated with each molecular eigenstate. In fact, the "band-like" nature of the electronic states in conjugated oligomers becomes obvious when looking at the shape of the orbitals. For 10pPP, the five highest occupied and five lowest unoccupied orbitals are shown in **Figure 3**.

Figure 3 also illustrates that the molecular orbitals adopt a Bloch-like character, that is, they can be regarded as periodically modulated plane waves. The plane-wave character is indicated by the sine-functions in **Figure 3** and the periodic modulation is given by the type-(1) occupied and unoccupied benzene orbitals from **Figure 2**. For molecular crystals based on oligomers, the quasiband structure description of the individual building block can be reconciled with the full three-dimensional band structure of the molecular crystal by going from a reduced zone scheme (relying on the crystal unit cell as a repeat unit) to an extended zone scheme, which is then consistent with assuming that the individual molecular building block (e.g., the phenylene ring) is the repeat unit.

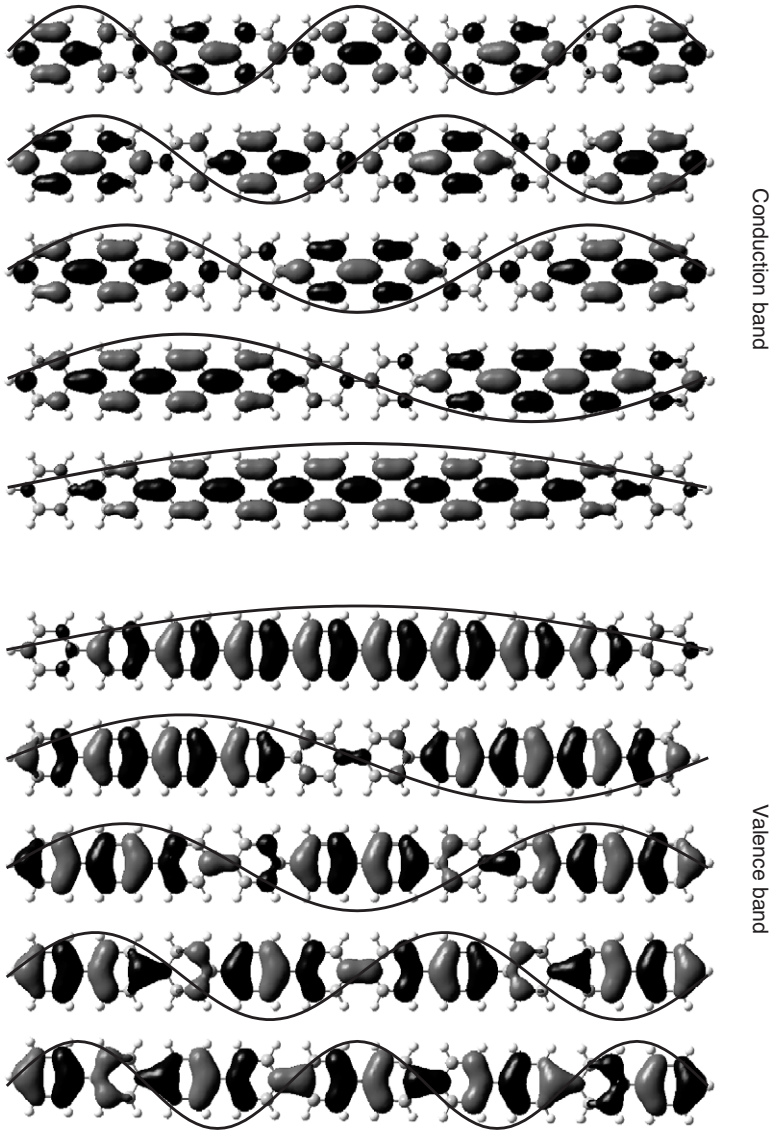


Figure 3 INDO-calculated five highest-lying occupied orbitals and five lowest-lying unoccupied orbitals of a planar ten-ring oligophenylene. The sine-wave envelope functions illustrate the Bloch-type character of the molecular orbitals.

This general evolution of the electronic states for long (periodic) chains does hold not only for one-electron eigenstates (orbitals), but also for excited states taking account of electron-correlation effects. This is shown in **Figure 4**, where the electron–hole two-particle wave functions for the lowest-lying excited states of (d)–(d) type in 10pPP are displayed.

Such plots illustrate the probability of finding the hole on atom x, while the electron is on atom y. They are calculated here for the excited states in 10pPP described by linear combinations of singly excited Slater determinants (this is referred to as a single configuration-interaction (SCI) approach, usually adequate to describe the low-lying optically allowed

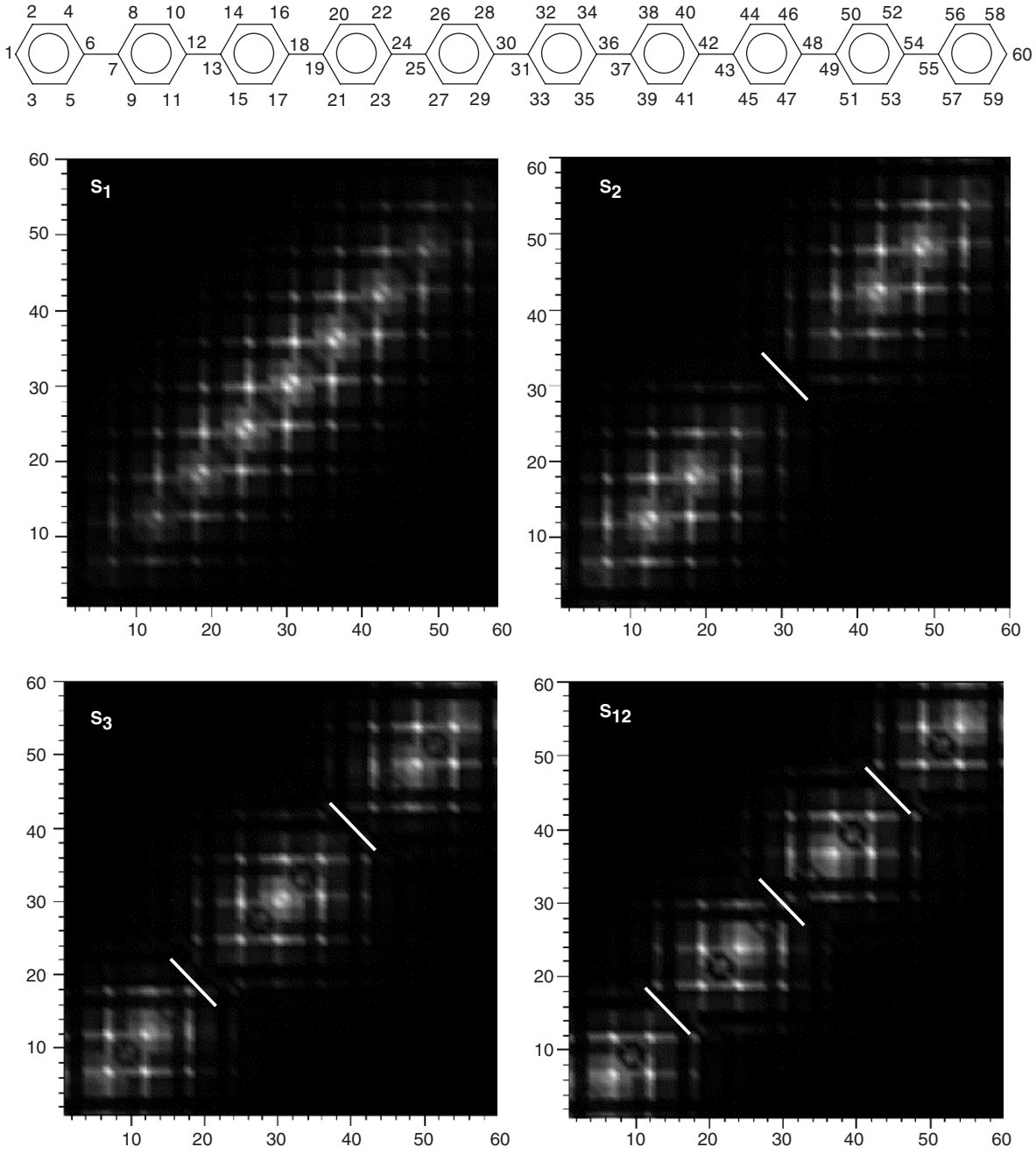


Figure 4 INDO/SCI-calculated electron–hole two-particle wave functions for the four lowest-lying (d–d) excited states in a planar ten-ring oligophenylene. The shading is proportional to the probability of finding the hole on the atom whose number is given at the horizontal axis, while the electron is at the atom whose number is given at the vertical axis. The numbers of the atoms are indicated in the molecular structure shown at the top of the graph. The diagonal white lines are a guide to the eye to show the increasing number of nodes for higher-lying states.

excited states in π-conjugated materials). In the context of the quasiband model, it is interesting to note that the S_1 state (of (d)–(d) character) presents no node. For S_2, there occurs one node (the diagonal white line in **Figure 4** is a guide to the eye); for S_3, there are two nodes, and for S_{12}, 3 nodes, etc. (the excited states from S_4 to S_{11} are of (d)–(l), (l)–(d) character). Therefore, it appears that the quasiband structure picture also holds for exciton bands, which are relevant for a comparison with spectroscopic data. This behavior is not restricted to the (d)–(d) bands; as a typical representative of a (d)–(l) excitonic state, the electron–hole wave function of S_6 is shown in **Figure 5**. The momentum-value associated with S_6 (as determined from the number of nodes) is equivalent to that of S_3 in the (d)–(d) band.

Another interesting property of the excited states that can be derived from the electron–hole wave functions is the localization of the electron–hole pair. The graphs shown in **Figure 4** indicate that the electron–hole pair is delocalized over the whole chain, as the shading goes along the full diagonal of the plots. The average electron–hole separation is, however, relatively small (calculated as 3.9 Å for the S_1 state). This is the consequence of the significant correlation of the electron and hole motions and can be seen by the relatively fast decrease of the electron–hole wave function when going away from the diagonal in the plots.

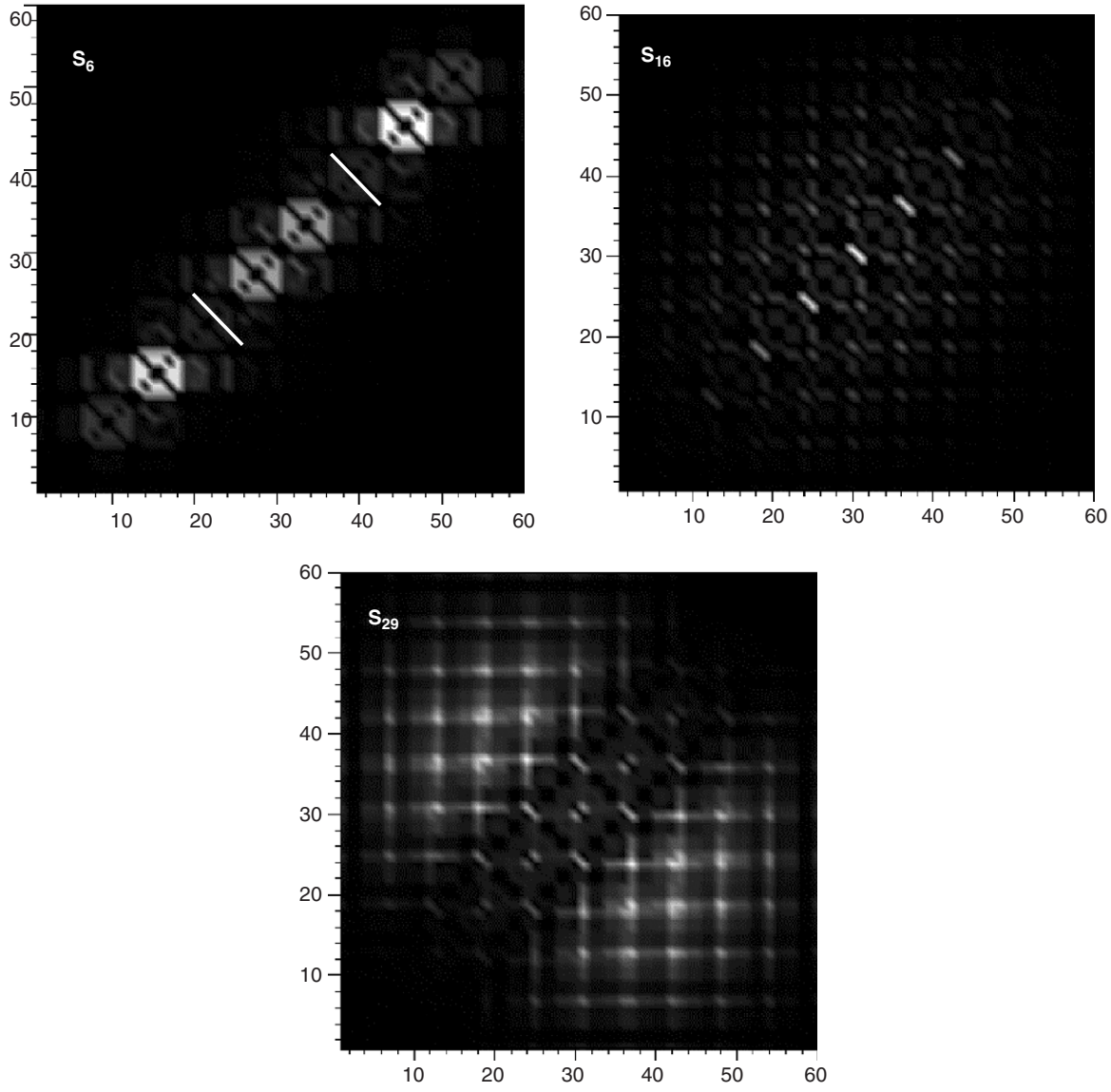


Figure 5 INDO/SCI-calculated electron–hole two-particle wave functions for some representative examples of higher-lying excited states in a planar ten-ring oligophenylene (see text). The numbering of the atoms is the same as in **Figure 4**.

As excitation energy increases, one finds exciton bands presenting a much larger electron–hole separation. An excited state representative of such bands is S_{16} at 4.8 eV (see central part of **Figure 5**). The average electron–hole separation for S_{16} is calculated to be 7.5 Å, which is more than twice as large as that for S_1. At higher energies, states with a clear charge-transfer character are found; this is illustrated by the electron–hole wave function of S_{29} in **Figure 5**: when the electron is on one half of the molecule, the hole is on the other (here, the average electron–hole separation is 14.9 Å). States like S_{16} and S_{29} are of considerable relevance for excited-state dissociation, which is found for high-energy photons in pump-probe experiments. Charge dissociation is critical, for example, in solar cells and photodetectors.

To link the properties of the exciton bands to those of the orbitals, it is useful to examine the configuration interaction description of the excited states (i.e., how the excited Slater determinants couple). For the four (d)–(d)-band states discussed above, one qualitatively obtains (writing only the dominant determinants):

$$
\begin{aligned}
|S_1\rangle &= a_1|\mathrm{HOMO}\rightarrow\mathrm{LUMO}\rangle \\
&\quad + b_1|\mathrm{HOMO}-1\rightarrow\mathrm{LUMO}+1\rangle + \cdots \\
|S_2\rangle &= a_2|\mathrm{HOMO}\rightarrow\mathrm{LUMO}+1\rangle \\
&\quad + b_2|\mathrm{HOMO}-1\rightarrow\mathrm{LUMO}\rangle + \cdots \\
|S_3\rangle &= a_3|\mathrm{HOMO}\rightarrow\mathrm{LUMO}+2\rangle \\
&\quad + b_3|\mathrm{HOMO}-2\rightarrow\mathrm{LUMO}\rangle \\
&\quad + c_3|\mathrm{HOMO}-1\rightarrow\mathrm{LUMO}+1\rangle + \cdots \qquad [1] \\
|S_4\rangle &= a_{12}|\mathrm{HOMO}\rightarrow\mathrm{LUMO}+3\rangle \\
&\quad + b_{12}|\mathrm{HOMO}-1\rightarrow\mathrm{LUMO}+2\rangle \\
&\quad + c_{12}|\mathrm{HOM}-2\rightarrow\mathrm{LUMO}+1\rangle \\
&\quad + d_{12}|\mathrm{HOMO}-3\rightarrow\mathrm{LUMO}\rangle + \cdots
\end{aligned}
$$

A most interesting aspect of these CI expansions is that the determinants that strongly mix are those for which the sum of the numbers of nodes for the occupied and unoccupied orbitals involved in the excitation are the same. (For example for S_3, a determinant in which an electron has been excited from an orbital with no node (HOMO) into an orbital with two nodes (LUMO + 2) is coupled to a determinant where both orbitals have one node (HOMO − 1 and LUMO + 1) and a determinant with two nodes for the HOMO + 2 and no nodes for the LUMO. Only the significant contribution of $|\mathrm{HOMO}-1\rightarrow\mathrm{LUMO}+1\rangle$ to the S_1 state somewhat deviates from this behavior.) In the quasiband model, this can be interpreted as momentum-conservation. This also allows one to rationalize why the states depicted in **Figure 4** apart from S_1 are either optically forbidden or have a small oscillator strength since for optical excitations the momentum transfer has to be negligibly small.

The Effect of Molecular Conformation and Intermolecular Interactions

So far, for the sake of simplicity, the oligophenylene chains have been assumed to be coplanar. However, in a nonbridged chain, this is not necessarily the case because the tendency to maximize conjugation by planarizing the (macro)molecule is counteracted by steric interactions between adjacent rings. Quantum-mechanical calculations on isolated oligophenylene chains, for example, give an inter-ring torsion angle of 40° as the optimal compromise between these two driving forces. This is confirmed by measurements on biphenyl, which displays an inter-ring twist angle of 42° in the gas phase. In solution and the molten state, the angle is found to be *c.* 32°; similar twists can be assumed for amorphous materials. These angles can also be affected when long aliphatic side chains are attached to the conjugated backbones. In the crystalline state, biphenyl has been found to be planar on average, but more recent data on terphenyl indicate an inter-ring twist of 12° in the solid state.

To address the influence of the inter-ring twist on the excited states, several PPP derivatives are discussed and their results for fully geometry-optimized and coplanar chains are compared; The systems under consideration are ten-ring oligomers of poly(*p*-phenylene) (10PP), poly(fluorene) (10PF), and ladder-type poly(*p*-phenylene) (10LPP); their chemical structures are shown in **Figure 1**. In poly(fluorene), every other connection between rings is bridged by a saturated carbon atom forming a fluorene unit that is coplanar. The torsion angle between consecutive fluorene units is similar to that in oligophenylenes. In 10LPPP, all phenylene rings are linked by a carbon bridge, which leads to a fully planar conformation of the backbone. For 10PP and 10PF, conformers forced into a coplanar conformation were also studied; in polyfluorenes a (near) coplanar conformation is in fact found for systems with nonbranched aliphatic side chains attached to the bridging carbon atoms, as a result of crystallization of these side chains (this phase is usually referred to as the *β*-phase). The transition energies and oscillator strengths for the lowest-energy absorption maxima are listed in **Table 1**.

As expected, the material with the largest bandgap is 10PP, for which the conjugation is most strongly

Table 1 INDO/SCI-calculated $S_0 \rightarrow S_1$ transition energies in ten-ring oligomers of fully geometry-optimized poly(*p*-phenylene) (10PP), poly(fluorene) (10PF), and ladder-type PPP (10LPP), as well as the coplanar conformations of poly(*p*-phenylene) (10pPP) and polyfluorene (10pPF)

Material	$E_{S_0 \rightarrow S_1}$ (eV)	*Oscillator strength*
10PP	3.84	4.52
10PF	3.58	4.34
10pPP	3.36	4.93
10pPF	3.31	4.54
10LPP	3.20	4.32

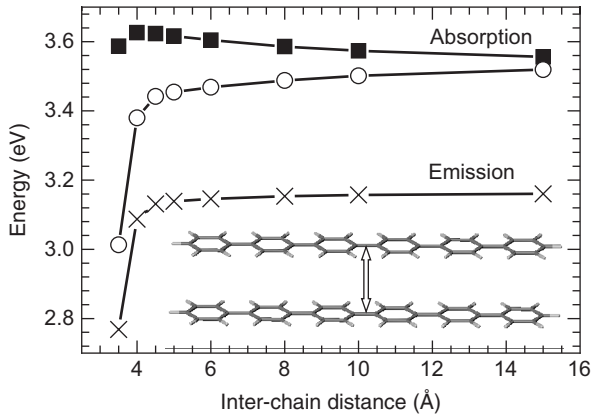


Figure 6 INDO/SCI-calculated dependence of the energies of the lowest-lying excited states in a dimer consisting of two cofacial coplanar sexiphenyl chains. The filled squares and open circles correspond to the two lowest-lying excited states for the ground-state geometry (absorption process). The crosses represent the energy of the lowest-lying excited state of the dimer in which one of the chains is in the relaxed excited-state conformation (emission process).

disturbed by the twists between all rings. The bandgap in 10PF where every other inter-ring bond is twisted, is significantly reduced by ~ 0.26 eV; for coplanar conformations (10pPP, 10pPF, and 10LPP), the energy gap is between 0.48 eV and 0.64 eV lower than in 10PP. The smallest gap is obtained for the fully bridged 10LPP system although the rings themselves are significantly distorted by the presence of the bridges. These findings are fully consistent with the experimental observations, which show that the emission and absorption maxima for the β-phase of poly(fluorene) are red-shifted by some 0.2 eV to 0.4 eV relative to regular poly(fluorene). The absorption and emission peaks in ladder-type poly(*p*-phenylene) are red-shifted by another 0.1 to 0.2 eV relative to the β-phase.

The description of excited states in phenylene-based materials given so far applies to isolated chains and also holds for molecules and polymers in solution and in amorphous films. (This is especially true for polymers with long aliphatic side chains, whose primary role is to provide solubility.) In crystals with well-ordered building blocks and short inter-chain distances, inter-molecular interactions can start to play a significant role. These have been studied by full 3D band structure calculations or by looking the properties of large molecular clusters. (Such studies show for instance that the l-bands also display significant dispersion in the direction perpendicular to the chain-axis.) When dealing with excited states, the situation quickly becomes rather complex; one then usually relies either on band structure calculations including correlation effects (e.g., by solving the Bethe–Salpeter equation) or on a correlated description of molecular clusters in a super-molecular approach. In addition, it has been recently suggested that in molecular crystals the exciton–exciton coupling is of similar magnitude as the exciton–phonon coupling; this would require a simultaneous treatment of the electronic and vibrational parts of the Schrodinger equation. Wave-propagation and polariton effects have also been found to significantly influence the optical properties of crystalline organic thin films.

In the following, the impact of intermolecular interactions on the optical properties by considering a simple model system consisting of a cofacial sexiphenyl dimer, shown in the inset of **Figure 6**, is discussed. The interaction between the two molecules results in a splitting of the molecular-orbital and exciton-state energies due to symmetric and antisymmetric linear combinations of the eigenstates of the individual molecules. The evolution of the excited-state energies related to the splitting of the lowest excited states of an individual oligomer is shown in **Figure 6**. As expected, at large inter-chain distances, the splitting is small; when the two molecules approach and start interacting significantly, the lowest excited state of the dimer gets stabilized while the next state is destabilized. Only for very short intermolecular distances (below 4 Å) would the energies of both states decrease.

For the absorption process (note that in this case the molecules in the cofacial-dimer conformation have identical geometries), the lowest-lying excited state (given by the open circles in **Figure 6**) is optically forbidden for symmetry reasons; thus a blue-shift of the absorption maximum is expected (the energy of the lowest optically allowed excited state is given by the filled squares in **Figure 6**). These selection rules relax when the symmetry is broken, for instance, when rotating one of the molecules (in such a way that the long molecular axes of the individual

oligomers are no longer parallel). Whether such symmetry-breaking occurs in molecular crystals critically depends on the space group of the system.

For the emission process, one encounters a different situation: As a result of the strong geometry relaxation that follows the excitation of a conjugated organic molecule, the symmetry of the dimer can be broken even in the cofacial conformation (i.e., when the geometry modification is centered on one of the two molecules). In this case, the lowest-lying excited state becomes optically allowed (although the oscillator strength might be relatively weak). Therefore, one has to expect a red-shift of the emission maximum upon decreasing the intermolecular distance (crosses in **Figure 6**).

Electronic Nature of the Lowest-Lying Excited States

It is important to point out that the lowest optically allowed state (usually mainly described by a HOMO→LUMO excitation) is not necessarily the lowest-lying excited state even in an isolated chain. For instance, it is well established that for polyenes (longer than butadiene) a highly correlated Ag state (with the same symmetry as the ground state) comes to lie below the optically allowed 1Bu state. This is the reason why polyacetylene does not luminesce. To describe such effects properly, it is necessary to go beyond a single configuration interaction description, since the low-lying Ag states have strong contributions from multiply excited determinants. In the case of phenylene-based materials (due to the large effective degree of bond-length alternation), the HOMO→LUMO excited state is the lowest-lying state within the singlet manifold; this is confirmed by the high luminescence quantum yields of oligo- and polyphenylenes and their bridged derivatives.

The ordering of the excited states also has to take account of triplet states. It is in fact a common feature in organic semiconductors that T_1, the lowest-lying triplet state, lies at significantly lower energies than the S_1 state. As an example, the energy diagram for a (coplanar) *p*-sexiphenyl (6PP) molecule is shown in **Figure 7** (here, the geometries and energies of the S_0 and T_1 states are obtained from density-functional theory (DFT) calculations and the transition energies to S_1 and T_n are calculated using time-dependent DFT). The T_1 state is seen to lie more than 1.3 eV below the S_1 state; there are several other excited triplet states present between T_1 and T_n (where T_n is the triplet excited state presenting the largest oscillator strength for a transition from T_1), some of which also lie below S_1. Note that

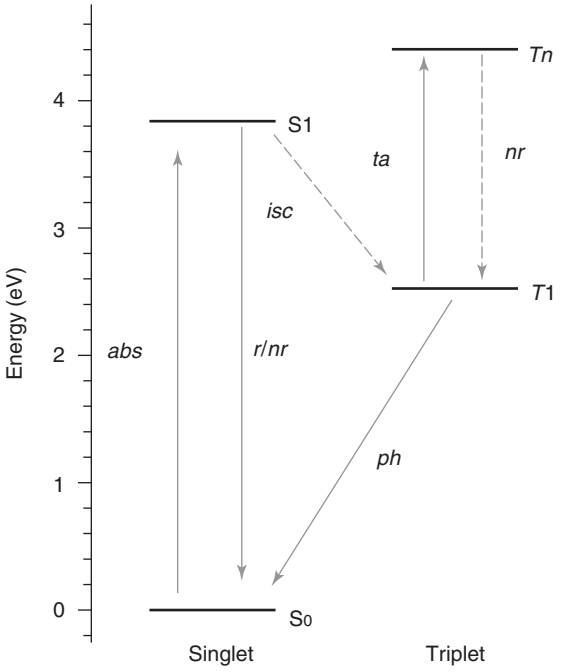


Figure 7 DFT- and TD-DFT-calculated (Jablonski) energy diagram for the low-lying optically active excited states in the singlet and triplet manifolds of coplanar sexiphenyl. The energies in the singlet manifold are given for the relaxed ground-state (S_0) geometry, while those for the triplet manifold are calculated for the relaxed T_1 geometry. The arrows represent processes coupling these excited states: *abs* – absorption; *r/nr* – radiative or nonradiative recombination; *isc* – inter-system crossing; *ta* – triplet absorption; *nr* – nonradiative recombination; *ph* – phosphorescence.

the energy of T_1 relative to S_0 and S_1 is an important factor when studying singlet/triplet formation rates in organic light emitting devices, or when adding organometallic dyes as phosphorescent triplet harvesting molecules to increase device efficiencies.

The Jablonski-diagram in **Figure 7** illustrates the typical photophysical processes linking the ground state and the various low-lying excited states. (In **Figure 7**, the relaxed S_0 geometries for the singlet and T_1 geometries for the triplet manifold is dealt with. Therefore, the energies are a proper description only for the $S_0 \rightarrow S_1$ and $T_1 \rightarrow T_n$ absorption processes. For the other processes, the energies have to be corrected by the respective relaxation energies.) Starting from the ground state S_0, the molecule can be excited via optical absorption (*abs*) into the one-photon allowed state S_1; from there, it can decay back to the ground state either radiatively or nonradiatively (*r/nr*), dissociate into positive and negative charge carriers (polarons), be excited into higher-lying singlet states S_n, or switch to the triplet manifold either via inter-system crossing (*isc*) or exciton-dissociation and subsequent nongeminate

polaron-pair recombination. Within the triplet manifold, the excited state will readily relax to the T_1 state from where the molecule can decay into the S_0 state either via phosphorescence processes (*ph*) or through the emission of phonons. Due to spin-selection rules, the T_1 state is usually long-lived. Therefore, absorption processes within the triplet manifold (*ta*) can have practical relevance for optical limiting applications and also need to be considered when designing materials for organic-laser applications.

The Role of Chemical Defects

An aspect of importance with regard to applications is the role chemical and structural defects play for the electronic and optical properties of conjugated materials. As in the inorganic semiconductor industry, great efforts have to be and are being made to eliminate chemical impurities and defects. In many instances, charge transport in organic circuits is critically influenced by charge-carrier trapping at grain boundaries and other morphological and chemical defects and the emission properties of light emitting devices can be strongly modified by sub-percent concentrations of impurities due to long-range energy migration and carrier trapping. Although "impurities" can be added intentionally as electronic dopants or emissive guest-molecules, they are usually detrimental for the performance of the device.

An example of the impact of chemical impurities is the recent understanding that the green shift of the emission in poly(fluorene) LEDs is not a consequence of aggregate or excimer formation (structural "defects"), but rather a result of chemical oxidation leading to the formation of fluorenone defects along the poly(fluorene) chains. The consequences of this type of defect for the optical properties of poly(fluorene) are now briefly discussed.

In **Table 2**, the properties of the lowest-lying excited states in a pristine ten-ring fluorene oligomer (10PF from **Table 1**) are compared with those of a model system in which the central fluorene unit has been replaced by a fluorenone (10PFK) (see chemical structure on top of **Figure 8**). Note that very similar results are obtained for ketonic defects at the end of the chain. The lowest-lying excited state in 10PF is a $\pi-\pi^*$ excited state at 3.58 eV, which is delocalized over the whole chain and is dominated by a HOMO to LUMO excitation, much like the S_1 state in coplanar poly(*p*-phenylene) discussed above. Upon geometry relaxation in the excited state, the transition energy is reduced to 3.12 eV.

In the ketone-containing molecule, 10PFK, a similar excited state, S_3, is found at an energy ~ 0.1 eV higher than that of S_1 in 10PF. There are, however, two excited states below the $\pi-\pi^*$ state in 10PFK. The lowest-lying state at 2.77 eV displays an $n-\pi^*$ character (i.e., its CI-description is dominated by an excitation from a σ-orbital localized at the C═O group to the π^* LUMO). The $S_0 \rightarrow S_1$ transition is optically forbidden due to symmetry reasons. S_2 is a weakly allowed charge-transfer $\pi-\pi^*$ (CT-$\pi-\pi^*$) state. Interestingly, in the relaxed excited-state geometry, which is relevant for emission, S_2 becomes the lowest excited state. Thus, the CT-$\pi-\pi^*$ state becomes responsible for the green emission in degraded poly(fluorene) (see right part of **Table 2**). To better understand the electronic nature of the $\pi-\pi^*$ and CT-$\pi-\pi^*$ states, it is useful to look at the orbitals involved in their CI-description and at the corresponding electron–hole two-particle wave functions. These are displayed in **Figure 8** for the relaxed excited-state geometry (the electron–hole wave function of the $n-\pi^*$ state, which is not shown here, is completely localized on the C and O atoms of the ketonic defect).

The strongest contribution to the description of the CT-$\pi-\pi^*$ state comes from a HOMO to LUMO excitation. These orbitals are shown in the inset of the upper electron–hole plot in **Figure 8**. The HOMO of 10PFK is strongly reminiscent of that of pristine poly(fluorene). The LUMO has no equivalent in pristine poly(fluorene) and is strongly localized on

Table 2 INDO/SCI-calculated excitation energies and oscillator strengths for a ten-ring poly(fluorene) chain without (10PF) and with (10PFK) a ketonic defect on the central fluorene unit. A chiral conformation is chosen and the ground-state geometry is optimized with the semi-empirical AM1 Hamiltonian, which is coupled to multi-electron CI to optimize the relaxed excited-state geometry.

	Ground-state geometry			*Excited-state geometry*		
	Type	*E* (eV)	*Oscillator strength*	*Type*	*E* (eV)	*Oscillator strength*
10PF						
S_1	$\pi-\pi^*$	3.58	4.34	$\pi-\pi^*$	3.12	3.91
10PFK						
S_1	$n-\pi^*$	2.77	0.00	CT-$\pi-\pi^*$	2.45	0.47
S_2	CT-$\pi-\pi^*$	3.05	0.42	$n-\pi^*$	2.64	0.00
S_3	$\pi-\pi^*$	3.68	3.88	$\pi-\pi^*$	3.44	3.08

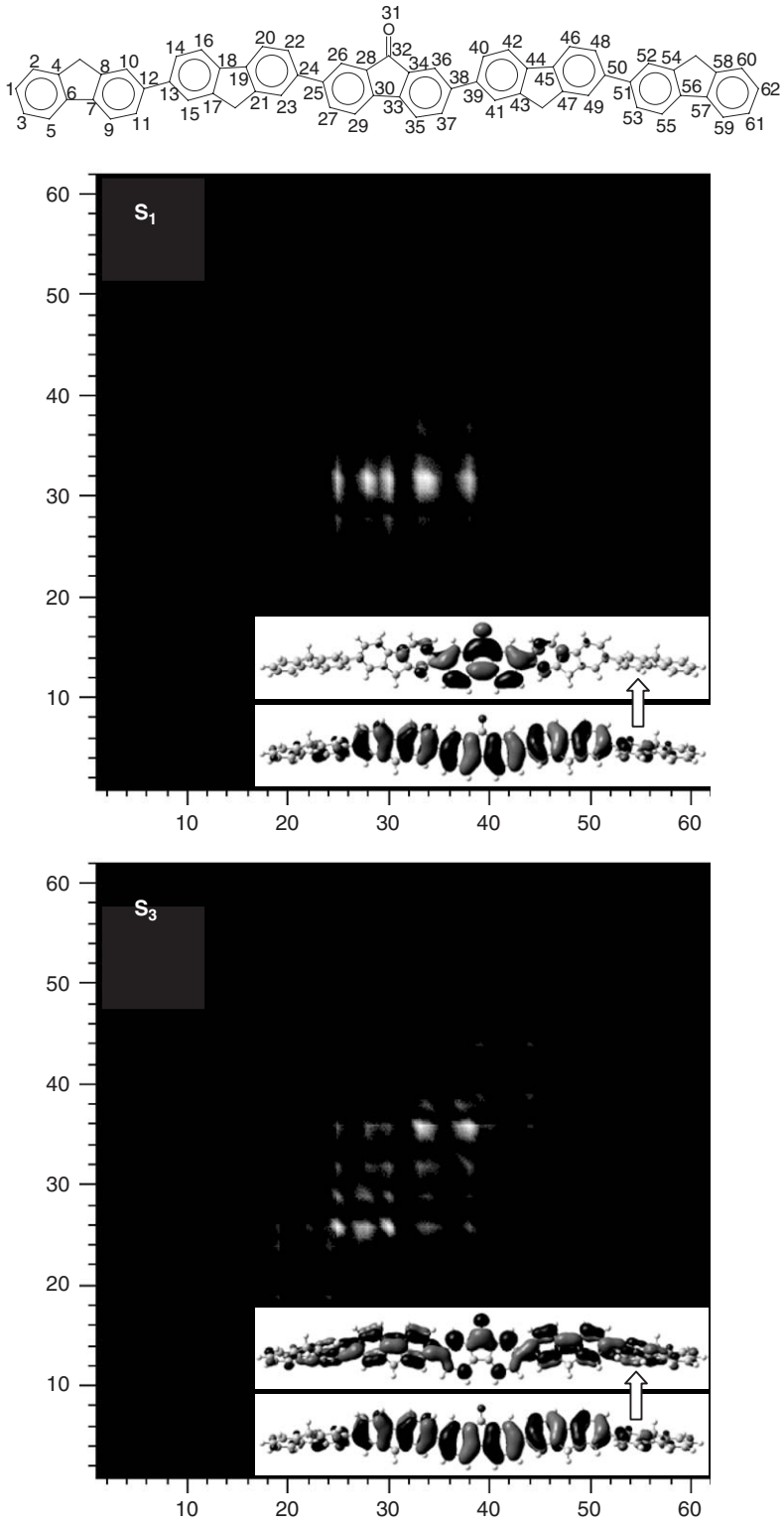


Figure 8 INDO/SCI-calculated electron–hole wave functions for the S_1 and S_3 states in a ten-ring oligofluorene in which the central fluorene unit has been replaced by a fluorenone. As the ketonic defect mainly affects the emission properties of poly(fluorene), the calculations have been performed for the relaxed S_1 geometry obtained from the AM1/CI approach. The atom site numbers used in the electron–hole plots are shown in the molecular structure at the top of the graph. The inset in the top graph shows the HOMO and LUMO orbitals which dominate the CI description of the S_1 state; the HOMO and LUMO + 1 are shown in the inset of the bottom graph.

the fluorenone unit. As a result, the electron–hole pair describing the CT-$\pi-\pi^*$ state is also localized on the fluorenone unit. The $\pi-\pi^*$ state (S_3) is delocalized over the whole chain and is dominated by a HOMO to LUMO + 1 excitation, with the latter orbital being strongly reminiscent of the LUMO in pristine polyfluorene.

This example clearly shows that the green-shifted emission in degraded poly(fluorene) stems from localized excited states resulting from the fluorenone chemical defect. Considering the fact that in poly(fluorenes) with ketonic defects, a new orbital (the LUMO from top of **Figure 8**) appears below the LUMO of pristine poly(fluorene), one can expect that this type of defect influences charge-carrier transport by acting as an electron-trap; this has been recently confirmed by experimental studies and *ab initio* calculations of ionization potentials and electron affinities in poly(fluorene) chains with and without ketonic defects.

To summarize, in this article, the example of phenylene-based materials is taken: (1) to derive the basic electronic structure of π-conjugated materials; (2) to describe the relation between molecular wave functions and band structure-type eigenstates; (3) to illustrate the impact of inter-chain interactions; (4) to examine the ordering of the excited states in the singlet and triplet manifolds; and (5) to discuss the role of chemical defects on the electronic properties.

Acknowledgments

The work at Georgia Tech was supported in part by the National Science Foundation through the STC Program under Award Number DMR-0120967 and through grants CHE-0078819 and CHE-0342321, the Office of Naval Research, and the IBM Shared University Research Program.

See also: Organic Semiconductors; Polymers, History of; Polymer Structures; Polymers and Organic Compounds, Optical Properties of.

PACS: 36.20.Kd; 31.25.Qm; 78.66.Qn; 71.20.Rv

Further Reading

André JM, Delhalle J, and Brédas JL (1991) *Quantum Chemistry Aided Design of Organic Polymers*. Singapore: World Scientific.

Brédas JL (edn.) (1999) *Conjugated Oligomers, Polymers, and Dendrimers: From Polyacetylene to DNA*, Francqui Scientific Library, vol. 4. Paris, Bruxelles: De Boeck & Larcier.

Cornil J, Beljonne D, Calbert JP, and Brédas JL (2001) Interchain interactions in organic π-conjugated materials: impact on electronic structure, optical response, and charge transport. *Advanced Materials*, vol. 13, pp. 1053–1067. Weinheim: Wiley-VCH.

Farchioni R and Grosso G (eds.) (2001) Organic electronic materials: conjugated polymers and low molecular weight organic solids. *Springer Series in Materials Science*, vol. 419. Berlin: Springer.

Hudson BS, Kohler BE, and Schulten K (1982) Linear polyene electronic structure and potential surfaces. In: Lim EC (ed.) *Excited States*, vol. 6. New York: Academic Press.

Köhler A and Beljonne D (2004) *Advanced Functional Materials*, vol. 14, pp. 11–18. Weinheim: Wiley-VCH.

Pope M and Swenberg ChE (1999) *Electronic Processes in Organic Crystals and Polymers*, 2nd edn. Oxford: Oxford University Press.

Salem L (1966) *The Molecular Orbital Theory of Conjugated Systems*. New York: W.A. Benjamin.

Sariciftci NS (ed.) (1997) *Primary Photoexcitations in Conjugated Polymers: Molecular Exciton versus Semiconductor Band Model*. Singapore: World Scientific.

Scherf U and List EJW (2002) Semiconducting polyfluorenes – towards reliable structure-property relationships. *Advanced Materials*, vol. 14, pp. 477–487. Weinheim: Wiley-VCH.

Scholes GD (2003) Long-range resonance energy transfer in molecular systems. *Annual Review of Physical Chemistry* 54: 57–87.

Skotheim TA, Elsenbaumer RL, and Reynolds JR (eds.) (1998) *Handbook of Conducting Polymers*, 2nd edn. New York: Dekker.

Soos ZG, Galvão DS, and Etemad S (1994) *Advanced Materials*, vol. 6, pp. 280–287. Weinheim: Wiley-VCH.

Tretiak S and Mukamel Sh (2002) Density matrix analysis and simulation of electronic excitations in conjugated and aggregated molecules. *Chemical Reviews* 102: 3171–3212. American Chemical Society.

Polymers, History of

M M Coleman and P C Painter, The Pennsylvania State University, University Park, PA, USA

Introduction

Although this article presents only a very brief history of polymers, it would be remiss if it did not open by acknowledging that without polymers this encyclopedia would not have been written! Mother Nature, with ingenious dexterity, used "natural" organic polymers such as proteins, polypeptides, carbohydrates, and DNA, to build the structural elements of life itself. Long before there was any fundamental understanding of the macromolecular nature of materials such as cotton, silk, wool, wood, leather, bone, and ivory, humankind found ways to use and modify these materials to produce articles for clothing, utensils, and shelter. One could write books on this subject alone, so in this short history the focus

is on how the foundations of our understanding of the nature of polymer materials were laid. In doing so, a lot of important and interesting things have been left out. Accordingly, the focus is on the history of semisynthetic and synthetic polymers, mentioning natural rubber only in passing. (Rubber actually deserves a history of its own and the interested reader is encouraged to consult the sources listed at the end of this review.)

To many people, the birth of polymer science as a discipline starts with Hermann Staudinger and his assertion that materials such as natural rubber are covalently bonded long-chain molecules, as opposed to some sort of colloidal aggregate. Jumps such as this in our understanding of nature do not occur in a vacuum, however. Staudinger was building on a wealth of work by nineteenth-century chemists and entrepreneurs, whose goal was to produce substitutes for various types of increasingly expensive or difficult to obtain natural materials such as ivory and silk.

Early History: Developments in the Nineteenth Century

Perhaps the most enduring images of the industrial revolution of the nineteenth century are those provided by the ironmasters and engineers. One thinks of a world of railways, "dark, satanic mills," and the social and political consequences of the division of labor and the creation of an urban working class. It is the world according to Charles Dickens, built on iron, steel, coal, and the power of steam. Natural polymers, such as cotton and wool, were, of course, key materials and after Charles Goodyear's discovery of vulcanization, natural rubber would join them. But, from about the middle to the end of the nineteenth century onward, various synthetic and semisynthetic polymer materials were also introduced and these were to have a profound social and economic impact.

In 1845, common household items such as combs and toothbrushes were often laboriously fashioned from bone or ivory and were thus expensive and unavailable (except in crude wooden forms) to most people. This was to change, although not immediately, with the work of the Swiss chemist, Christian Schönbein. The story (perhaps apocryphal) has it that he used his wife's apron to wipe up some spilled concentrated nitric and sulfuric acid while conducting experiments in his home. After washing the apron, he put it in front of the fire to dry, upon which it immediately flared brightly and disappeared. Schönbein had discovered guncotton, formed by nitrating the cellulose in the apron. Not only did guncotton have more explosive power than gunpowder, it gave off far less smoke. This was considered a major advance, because it lessened the confusion of the battlefield and generals could now see the people they were slaughtering. Advances in military technology were often spin-off materials and processes that are useful in everyday life, however, and this was to be no exception. Cellulose in its native form is both insoluble and infusible and cannot be cast from solution or molded to make useful objects. However, a French doctor, Louis Menard, found that nitrated cellulose would dissolve in a mixed solvent of alcohol and ether and cast to form a transparent film, called collodion. This was used to dress cuts and wounds. Crucially, it was also found that at levels of nitration of $\sim$13% or less, nitrated cellulose is no longer explosive, although still highly flammable.

Alexander Parkes, a prolific English inventor and true material scientist, had already achieved considerable fame as a metallurgist when he found that the addition of castor oil and camphor to nitrated cellulose lowered its softening temperature. The resulting material, which he modestly called Parkesine, could then be molded. Various objects made from this new material were displayed at the Great Exhibition at the Crystal Palace in 1862, where it won a medal for excellence of product. Parkes seems to have been unaware of the special role of camphor as a plasticizer, however. In some of his publications he omitted reference to it completely and the commercial exploitation of Parkesine was ultimately a failure. Nevertheless, it marked the beginning of the plastics industry.

This history now shifts across the Atlantic, where the next development came as a direct consequence of a shortage of elephants and rhinoceroses. It is a classic case of a short supply of one material, in this case ivory, providing the impetus for the development of another. The American billiard ball manufacturing company, Phelan and Callendar, offered a prize of $10 000 (a very large sum in those days) to anyone who could invent a substitute for ivory billiard balls. John Hyatt took up the challenge and chose to experiment with Parkesine. One can speculate that Hyatt might have misspent his youth in pool halls, because he was intrigued by the problem and spent several years of his life working toward a solution. He did not win the prize, but did establish the Albany Billiard Ball Company, which used a mixture of gum shellac and wood pulp to make the bulk of the ball. This was coated with nitrocellulose, which gave this new ball some interesting properties. Hyatt later reported, "We had a letter from a billiard saloon proprietor in Colorado mentioning that occasionally the violent contact of the balls would produce a mild explosion and saying that he did not care so much about it, but that instantly every man in

the room pulled a gun." It seems that controlling the degree of nitration remained a problem.

This experience with nitrocellulose prompted Hyatt to investigate its properties in more detail. He eliminated castor oil from his formulations, using only camphor, and, with good old-fashioned Yankee ingenuity, solved the manufacturing problems that had plagued Parkes. He produced a hard material that could be softened by heating to 90°C and called it celluloid. Celluloid was an enormous success. Many common objects that were previously expensive could now be mass-produced cheaply. These ranged from knife handles to false teeth, from combs to that symbol of Victorian rectitude, the stiff celluloid collar. It also provided film for the fledgling movie industry, whose stars became known as the "celluloid kings and queens."

Celluloid was not a true synthetic polymer, of course, but a chemically modified natural polymer. Its central significance to the development of the field was not so much the nature of the objects produced, although these had an important social and economic impact, but the engineering techniques and machines developed to produce these goods. An example is the artificial silk fiber produced by Count Hilaire de Chardonnet, who developed an extrusion process in which a nitrocellulose solution was forced through a small orifice producing a lustrous thread. Unfortunately, nitrocellulose lost none of its incendiary qualities in this process. The workers in Chardonnet's factory quickly dubbed the product "mother-in-law silk" (buy your mother-in-law this nice blouse and stand her in front of the fire), thus revealing an interesting, though morbid, sense of humor. Chardonnet overcame this problem by "denitrating" the fibers by passing them through a solution of aqueous ammonium hydrogen sulfide and, in effect, regenerating cellulose.

There are easier ways to make regenerated cellulose and in the early 1890s Cross, Bevan, and Beadle discovered that cellulose could be rendered soluble by sodium hydroxide and carbon disulfide. A basic solution of cellulose xanthate is formed, which regenerates cellulose upon acidification. Fibers from regenerated cellulose were called rayon, while films were known as cellophane. By the 1930s cellophane was being used to wrap an enormous range of goods, not only to protect them from spoilage, but also because they had customer appeal. A grocery chain reported an increase in sales of an astonishing 2100% in two weeks once their doughnuts were wrapped in cellophane. In a mid-twentieth century poll by the *New York Times*, cellophane was considered to be one of the key discoveries of the century to that date and was even included in a Cole Porter song (You're the Top). The impact of these materials was enormous, but cheaper synthetic polymers have now largely replaced them.

The First Synthetic Plastic

The first truly synthetic polymer materials are the "phenolic resins," originally synthesized by reacting phenol (or, as it was called at the time, carbolic acid, a compound that was, and still is, commonly used as a disinfectant) and formaldehyde (an aqueous solution of which, formalin, had been used in the embalming business for thousands of years). It was actually an Englishman, Arthur Smith, who in 1899 was awarded the first patent for the application of phenol/formaldehyde resins as electrical insulators. But it was Dr. Leo Hendrik Baekeland, a Belgium chemist born in Gent and living in the US, who made the big breakthrough.

Baekeland was an interesting man. He graduated with a Ph.D. degree, "maxima cum laude" and then taught at the University of Gent until 1889. He then left for the US, where he quickly invented Velox, a photographic paper superior to anything else then available. George Eastman, who in 1888 introduced the Kodak camera with the slogan "you press the button, we do the rest," made up his mind to buy Velox and invited Baekeland to Rochester. Baekeland had decided to ask for what was then the very large sum of $50 000, but had it in the back of his mind that he would settle for $25 000. Eastman offered him a million, allowing Baekeland to retire to his laboratory in Yonkers, New York, and look for new chemical worlds to conquer.

Baekeland decided to try and develop a less flammable replacement for shellac, a natural coating derived from the resinous secretions of tiny lac beetles. According to one source, his immediate goal was to develop a wood varnish for bowling alleys, all the rage at that time. Other sources state that Baekeland perceived the growing impact of electricity on everyday life and his goal was to produce a new synthetic insulator. It took roughly 150 000 beetles ~6 months to produce enough resin to make one pound of shellac and the demands of the fledgling electrical industry had created a shellac shortage.

Although Baekeland is given the credit for the discovery of phenolic resins, there were significant contributions from many others including Adolf Baeyer, Arthur Smith, Adolf Luft, and James Swinburne. Baekeland's key discovery (made in 1907), one that set him apart from the competition, was that phenolic resins could be made in two parts. In essence, he developed a process where a polycondensation reaction was carried out at an elevated temperature.

He then stopped the reaction at an intermediate point by cooling to ambient temperature, where the partially reacted resin solidified. In the second stage, this reasonably stable intermediate material was placed into a mold and heated under enormous pressure. This transformed the material (by continued polymerization and cross-linking reactions) into a hard, shiny, intractable, relatively brittle solid. Bakelite, as Baekeland called his new material, could be molded and machined; it was hard, relatively strong, lightweight, and resistant to heat. It was the first true synthetic polymer and there was nothing else like it on the face of the planet. It found use not only in electrical fixtures, but also in knife handles, buttons, cameras, radio, and telephone equipment, essentially anything that required a molded part. It was advertised as the "Material of a Thousand Uses." And it was.

The Dawn of Understanding

The polymeric materials discussed so far were developed in a largely "Edisonian" fashion, by trial and error, and in the absence of any fundamental understanding of the macromolecular nature of polymers. In fact, the prevailing view at that time was that materials like silk, cellulose, and natural rubber were colloidal aggregates of small molecules. Not only that, the chemistry of what were then called "high molecular compounds," which could not be purified by crystallization or distillation, was regarded with disdain by the academic chemical community and referred to as "grease chemistry."

It was therefore an act of considerable courage for Hermann Staudinger to decide to devote the bulk of his efforts to these "high molecular compounds." He asserted in lectures and papers (published in the period 1917–1920) that these materials were covalently bonded long-chain molecules. Although Staudinger, at the ripe old age of 39, was a distinguished chemist, holding academic positions at prestigious universities in Karlsruhe, Zurich, and Freiburg, convincing his colleagues was to be no easy matter. For example, Staudinger recollects in *Arbeitserinnerungen* (Working Memoirs) that he was advised in the 1920s by a famous organic chemist of the day, H Wieland, "My dear chap, give up your ideas on big molecules. There are no organic molecules with a molecular weight of more than 5000. Just clean up your products and they will crystallize and reveal themselves as low molecular weight compounds." In 1925, after giving a lecture at the Zurich Chemical Society, he was attacked by a number of leading scientists and in frustration quoted Martin Luther, "Here I stand and can do no other."

Staudinger could not accept that the physical properties of rubber, proteins, cellulose, and the like could be explained by the idea that they consisted of associations of small molecules held together by some unknown force. (It was speculated at the time that some undiscovered "vital" or life force could be involved.) Staudinger maintained that the atoms in these molecules were connected by covalent chemical bonds, as in low molecular weight materials like benzene, but these molecules just happened to be incredibly large! Despite the apparent reasonableness of this position, the dispute became almost theological in its intensity. Von Baeyer had it right when he pronounced his "Three Stages of a Triumphant Theory": First it is dismissed as untrue; second it is rejected as contrary to religion; and third it is accepted as dogma and every scientist claims to have long appreciated its truth! Staudinger was, of course, correct, and he was belatedly recognized for his achievements in 1953, when he was awarded the Nobel Prize in chemistry. Sometimes you have to outlive your enemies!

Numerous experiments performed by the late 1920s and early 1930s provided accumulating evidence in favor of the existence of macromolecules. Particularly important were Staudinger's work on polyoxymethylenes, where he systematically studied changes in the physical and chemical properties of this polymer and a homologous series of its oligomers; the X-ray diffraction studies of Meyer and Mark, establishing the chain structure of cellulose; and the molecular weight measurement made possible by Svedberg's development of the ultracentrifuge. However, it was the brilliant experimental studies of Wallace Hume Carothers, which were to remove any lingering doubt about the macromolecular, long-chain nature of polymers.

Carothers obtained his doctorate from the University of Illinois, and then took up a teaching position at Harvard. In 1928, Charles Stine persuaded Carothers to leave his teaching post and join the DuPont Company to head a newly formed fundamental research laboratory. He was given wide discretion and he clearly stated at the outset that his goal was to tackle the problem of macromolecules from a chemical approach, by synthesizing compounds of high molecular weight and known constitution. Carothers is best known for the discoveries of nylon and neoprene, but his general contributions to polymer synthesis and polymer structure/property relationships were truly seminal.

Wallace Carothers was probably the first to appreciate the enormous importance of monomer functionality in the production of condensation or step-growth polymers. Amongst many other things Carothers is also credited with categorizing the two major classes

of polymerization: condensation or step-growth (which he referred to as "C" type polymerization) and addition (called "A" type). Carothers' crowning glory, one of the major outcomes from his fundamental studies of polymerization, was the discovery of polyamides, or as they are now known universally, nylons. The first major market was ladies' stockings, introduced in the early 1940s. There was a huge demand for this product, which in these days of less formal wear is hard to imagine. Nylon production was co-opted for the war effort after Pearl Harbor was bombed (replacing silk in parachutes and so on), but when nylon stocking production was resumed in 1945 there were actually so-called nylon riots among people lining up to purchase the limited supply. This is a great story, one that would take more space than available here, but it is also a tragic one. Carothers never lived to see the extraordinary success of nylon. The bouts of depression, to which he had long been subject, became increasingly frequent after 1935 and he began to believe he was a failure as a scientist. He left his home in Wilmington, Delaware, on April 28, 1937 and checked into a Philadelphia hotel room, where he swallowed lemon juice laced with potassium cyanide.

The Post 1930s

Following Staudinger and Carothers, the synthesis of polymer materials would become based on chemical principles, rather than empirical trial and error. It was now understood what types of functional groups could react to form polymers and chemists set about making them. Nevertheless, it took years of effort by many scientists to achieve a firm understanding of the details of reaction mechanisms and their kinetics. For example, industrial scale polymerization of vinyl monomers in Germany by what became recognized as a free radical mechanism was in early work (late 1920s) initiated by exposure to sunlight. Initiation also played a crucial role in the development of a high-pressure process for the polymerization of ethylene by Fawcett and Gibson at ICI (Imperial Chemical Industries) in England in the 1930s. While trying to react benzaldehyde with ethylene at high pressure, a waxy solid was formed which they recognized to be a polymer of ethylene. A subsequent experiment blew up the apparatus and ICI banned high-pressure studies until stronger reaction vessels could be constructed. Fawcett left to join Carothers' group in America, but Perrin continued work in secret after normal working hours, risking both his life and his job. By 1936, it was recognized that small amounts of oxygen could initiate the reaction and by 1939, production of significant quantities was achieved. This turned out to be crucial for the war effort. Because polyethylene turned out to have an extremely low dielectric loss, it was ideal for radar cable insulation. As Robert Wattson Watts, the British inventor of radar, stated, "The availability of polyethylene transformed the design, production, installation and maintenance of airborne radar from the most insoluble to the comfortably manageable. And so, polyethylene played an indispensable part in the long series of victories in the air, on the sea and the land, which were made possible by radar."

By 1950, significant advances in understanding free radical polymerization, the nature of copolymerization, emulsion polymerization, etc., had been achieved. At the same time, the period between the first and second World Wars had seen the introduction of a number of important condensation polymers, which included not only nylon, but also the polyurethanes. Otto Bayer, who just happened to have the same name as the large German company for whom he worked, became head of the Central Scientific Laboratory at the ripe old age of 31. He was aware of the work of Staudinger and Carothers and reasoned that a superior "condensation" polymerization would be one where a small molecule (such as water) was not split off. He had worked on isocyanate chemistry and perceived that diisocyanates might form the basis for making such polymers. Synthesizing diisocyanates is hard and dangerous chemistry and one of Bayer's closest colleagues remarked, "If you had ever tried to make a monoisocyanate, you would not have come up with the mad idea of trying to produce diisocyanates. The reaction of phosgene on diamines would result in almost anything except diisocyanate, and would probably be thoroughly unstable, anyway." Others added, "It will explode!" This is another great story, which too cannot be told properly for want of space, but needless to say, Bayer succeeded.

Although some addition polymerizations involving catalysts and ionic initiators were performed between the two World Wars, it was not until the early 1950s that things really took off, when catalysts for the low pressure polymerization of ethylene were discovered, independently and almost simultaneously in the early 1950s, by Karl Ziegler in Germany and Robert Banks and Paul Hogan at the Phillips Petroleum company in the USA. These were not the only scientists involved, as workers at Standard Oil of Indiana were pursuing similar goals, while Giulio Natta, a Professor at the Politecnico di Milano, had become interested in the field after he heard a lecture given by Ziegler. Natta had sent some of his collaborators to work in Ziegler's laboratory to learn the art and craft of the technique and upon their return, Natta set out to synthesize polypropylene,

which could not be made free radically. They succeeded and filed for a patent just days before Ziegler's group, who by then had also succeeded in polymerizing propylene. Although Ziegler and Natta were jointly awarded the 1963 Nobel Prize in chemistry, they were no longer speaking to one another. Polypropylene also became the subject of one of the longest patent cases in history.

There is much more to this story, and the history of the development of many other major polymers and classes of polymers is equally fascinating. But although a history of polymer materials naturally focuses on how these materials were discovered and synthesized, an understanding of the physics and physical chemistry of these materials was equally important. It was Carothers, looking for a physical chemist to join his group, who hired the newly graduated Paul Flory in 1934. Carothers encouraged Flory to carry out mathematical investigations of polymerization and the subsequent impact that Flory was to have on the field of polymer science is difficult to overstate. Flory's book, *Principles of Polymer Chemistry*, is still a classic 50 years after publication and essential reading for anyone starting out in this field. In 1974, he was awarded the Nobel Prize in chemistry for his fundamental achievements, both theoretical and experimental, in the physical chemistry of macromolecules. While many other notable scientists have contributed to the theory of polymers, Pierre-Gilles de Gennes has a knack for seeing commonalities in wildly different physical systems and then reducing these observations to simple scaling arguments. His book, *Scaling Concepts in Polymer Physics*, is also a required reading for aspiring polymer physicists. In 1991, his peers recognized de Gennes' accomplishments and he was awarded the Nobel Prize in physics.

Final Words

There is much that has been omitted and many great scientists have been slighted by these omissions. The authors can only suggest that if they have sparked any interest in the reader that he or she refers to the references cited below.

See also: Polymer Structures; Polymers and Organic Compounds, Optical Properties of; Polymers and Organic Materials, Electronic States of.

PACS: 01.65.+g; 71.20.Rv; 72.80.Le; 73.61.Ph; 77.84.Jd; 78.30.Jw; 42.70.Jk; 61.41.+e; 61.82.Pv

Further Reading

Furukawa Y (1998) *Inventing Polymer Science: Staudinger, Carothers, and the Emergence of Macromolecular Science*. Philadelphia, PA: University of Pennsylvania Press.

Kaufman M (1963) *The First Century of Plastics*. London: The Plastics Institute.

Mark HF (1993) *From Small Organic Molecules to Large*. Washington, DC: American Chemical Society.

Morawetz H (1985) *Polymers: The Origin and Growth of a Science*. New York: Wiley.

Mossman STI and Morris PJT (eds.) (1994) *The Development of Plastics*. Cambridge: The Royal Institute of Chemistry.

Painter PC and Coleman MM (2003) *Painter and Coleman on Polymers*, 2-CD-ROM Set. Disk One: *Polymer Science and Engineering*. Disk Two: *The Incredible World of Polymers: Tales of Innovation, Serendipity and Perseverance*. Lancaster, PA DesTech Publications.

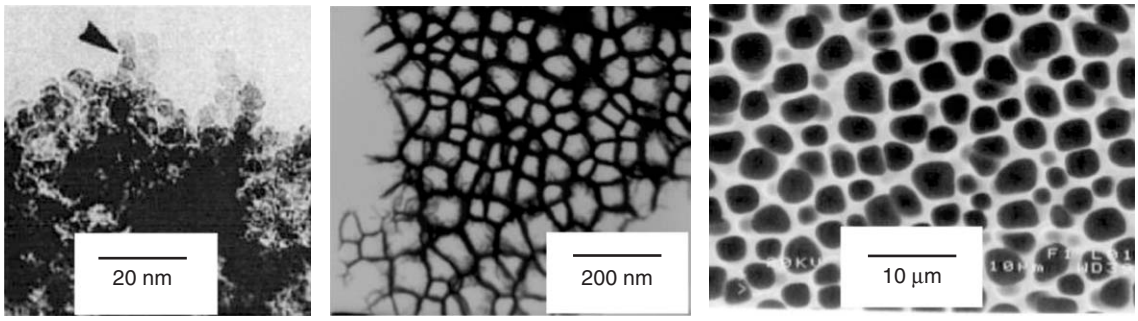
Porous Silicon

Z Gaburro, N Daldossoh, and L Pavesi, Università di Trento, Povo (Trento), Italy

This article deals with generalities and definitions of porous silicon (PSi): fabrication techniques, structural properties, chemical properties, electronic properties, electrical properties, optical properties, and actual or potential applications of PSi. Optical properties include light transport, photoluminescence, and electroluminescence.

Generalities and Definitions

PSi was discovered by Uhlir in 1956 while performing electrochemical etching of silicon. In 1990, Canham showed that certain PSi materials can have large photoluminescence (PL) efficiency at room temperature in the visible: a surprising result, since the PL efficiency of bulk silicon (Si) is very low, due to its indirect energy bandgap and short nonradiative lifetime. The reason for this was the partial dissolution of silicon, which causes (1) the formation of small silicon nanocrystals in the PSi material; (2) the reduction of the effective refractive index of PSi with respect to silicon, and hence an increased light extraction efficiency from PSi; and (3) the spatial confinement of the excited carriers in small silicon regions where nonradiative recombination centers are mostly absent. In general, PSi is an interconnec ted network of air holes (pores) in Si. PSi is classified according to the pore diameter, which can vary from



(a) (b) (c)

Figure 1 Examples of PSi structures: (a) microporous, (b) mesoporous, and (c) macroporous. ((a) Reproduced with permission from Cullis AG and Canham LT (1991) Visible light emission due to quantum size effects in highly porous crystalline silicon. *Nature* 353: 335–337; © Nature Publishing Group.)

Table 1 IUPAC classification of porous materials

Dominant pore width (nm)	*Type of material*
$\leqslant 2$	Microporous
2–50	Mesoporous
>50	Macroporous

a few nanometers to a few micrometers depending on the formation parameters (**Figure 1**). According to the general classification of porous materials, three size regimes are defined as in **Table 1**.

The word "nanoporous" is sometimes used for the smallest-pore regime to emphasize the nanometric dimension. The volumetric fraction of air of the material is called porosity (P). The internal surface of PSi per unit volume can be very large, of the order of $500\,m^2\,cm^{-3}$. The enhanced photoluminescence efficiency of PSi – compared to Si – has motivated research toward other porous semiconductor materials. For example, highly porous SiC, GaP, $Si_{1-x}Ge_x$, and Ge structures have been investigated.

Fabrication Procedures

Anodization

PSi is mostly fabricated by electrochemical anodization – often referred to as electrochemical etching – of bulk Si wafers in diluted aqueous or ethanoic hydrofluoric acid (HF). Ethanol is often added to facilitate evacuation of H_2 bubbles, which develop during the process. Typical anodization arrangements are schematically shown in **Figure 2**. Another much less common technique is stain etching, or chemical etching (with no current flow), performed with HF–HNO_3 solution.

The anodization can be performed either in potentiostatic (voltage-controlled) or in galvanostatic (current-controlled) mode. The latter is normally preferred, because it supplies the required charge for the reaction at constant rate, regardless of any evolution – during anodization – of the cell electrical impedance, ultimately leading to more homogeneous and reproducible material. The typical values of the most important anodization parameters, along with their impact on the anodization process, are reported in **Tables 2** and **3**. Three sample setups for fabrication of different PSi structures are shown in **Table 4**. The anodization can be modulated. The modulation is most easily achieved by varying the applied current density. Modulation results in controlled changes of the microstructure and the porosity of PSi along the growth direction. For example, PSi multilayer films can be simply created by varying the current density in time.

The position of pore nucleation can be controlled by etch pits, defined by a lithographic step before the anodization (**Figure 3**). This option is in practice limited to macroporous PSi, since present lowest lithographic resolution limits are in the range of a hundred nanometers. For example, ordered two-dimensional periodical matrices of pores can be formed, leading to materials with remarkable optical transport properties (2D photonic crystals, see **Figure 4**). PSi singlelayers or multilayers can be fabricated as free-standing samples. The detachment from the substrate can be achieved by applying a short pulse of high current density, that is, exceeding the electropolishing threshold, typically some hundreds of $mA\,cm^{-2}$.

Anodization Mechanisms

Typical current–voltage (I–V) characteristics of the electrochemical cell are shown in **Figure 5**. It should be stressed that the quantity having physical meaning is the current density J (at the Si/electrolyte interface),

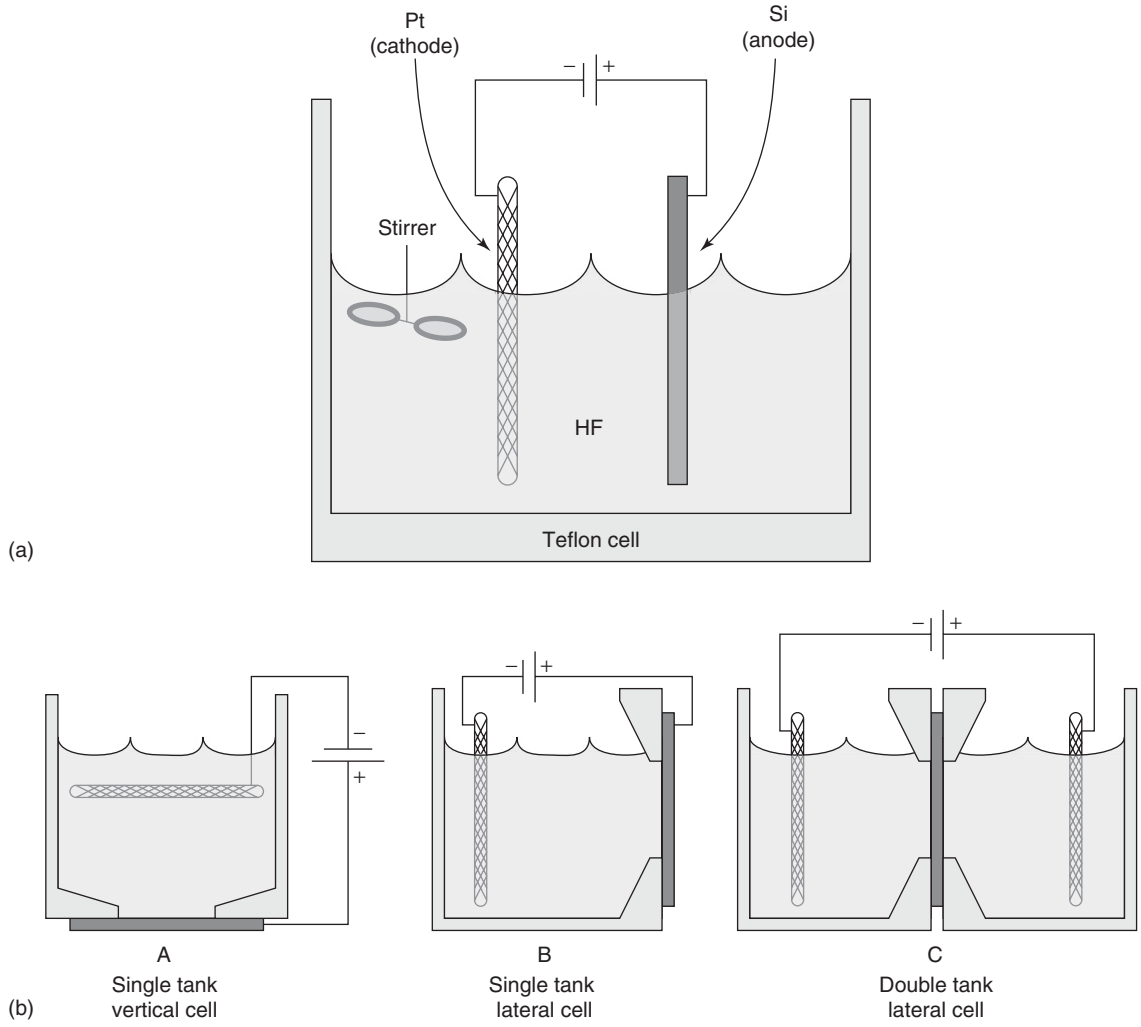


Figure 2 Schematic anodization arrangements: (a) principle and (b) practical examples (A, B, and C). Platinum and Teflon are mostly used because of their chemical resistance to HF. In the practical arrangements A, B, and C, the Si area exposed to the solution – typically $\sim 1\,cm^2$, circular – is well defined by the Teflon cell. Rubber gaskets between Si and Teflon can easily avoid spills. Type A is easier to design, type B is convenient when Si illumination is desirable (e.g., for *n*-type Si, see section "Fabrication procedures"), and type C does not require electrical contacts on Si. An optional stirrer is sometimes used to improve diffusion of HF during the anodization and H bubbles removal.

Table 2 Critical parameters of the anodization procedure

Parameter	*Typical range*	*Unit*
HF concentration	2–40	% (in weight)
Current density	0.5–150	$mA\,cm^{-2}$
Anodization time	5–1800	s
Temperature	250–300	K
Wafer resistivity (*p*-type)	0.001–100	Ω cm
Wafer resistivity (*n*-type)	0.001–100	Ω cm

A typical electrochemical solution can include, e.g., water (H_2O) in an approximately equal amount of HF, the rest being ethanol.

rather than the absolute current I. J and I scale with a fixed constant for a given cell, provided that the area of the Si sample exposed to the electrolyte is well defined and fixed.

In order to form PSi, the current at the Si side of the Si/electrolyte interface must be carried by holes, injected from the bulk toward the interface. The current must be kept between zero and the electropolishing threshold, which can be identified as the value of the first maximum of the anodic regime in the I–V curve. Useful regimes are included in the hashed region of **Figure 5**, where the electropolishing threshold voltage (for the curve marked with open circles) is $V \sim 1.3\,V$. In order to achieve significant hole current in *n*-type Si, external illumination of the sample is required, depending on the doping level. If the current exceeds the electropolishing threshold, the anodization results in a progressive, complete removal of Si. The wafer then has a mirror-like appearance.

Table 3 Effect of anodization parameters on PSi formation

An increase of	*Porosity*	*Etching rate*	*Electropolishing threshold*
HF concentration	Decrease	Decrease	Increase
Current density	Increase	Increase	
Anodization time	Increase	Almost constant	
Temperature			Increase
Wafer doping (*p*-type)	Decrease	Increase	Increase
Wafer doping (*n*-type)	Increase	Increase	

As an approximate rule of thumb, the etch rate is of the order of 1 nm s^{-1} for each mA cm^{-2} of the anodization current density.

Table 4 Three sample sets of anodization parameters, for *p*-type Si substrates, resulting in the three different structural regimes described in **Table 1**

HF concentration (%)	15	15	2
Current density (mA cm^{-2})	50	50	5
Wafer resistivity (Ω cm)	5	0.01	10
Resulting structure	Microporous	Mesoporous	Macroporous

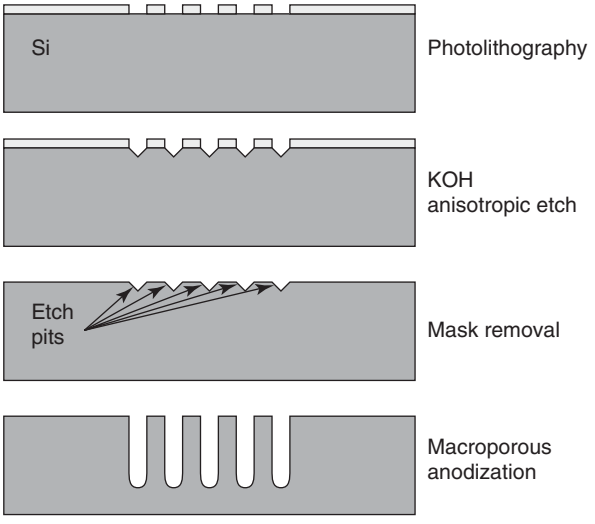


Figure 3 Schematic lithographic procedure resulting in the definition of pore-starting etch pits.

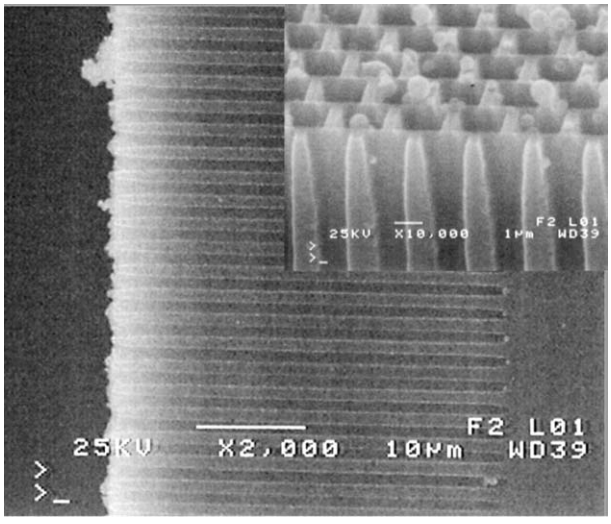


Figure 4 Example of a photonic crystal (cross section and top-view images of the same sample) realized with ordered macroporous PSi in a *p*-type doped silicon substrate. This photonic crystal has a photonic bandgap at 3.5 μm.

A hypothesized chemical reaction which could describe the anodization is

$$Si + 6HF \rightarrow H_2SiF_6 + H_2 + 2H^+ + 2e^- \qquad [1]$$

The last term – a negative charge at the interface, to be neutralized by the current flow – would explain the need of hole injection from the substrate toward the Si/electrolyte interface. Similar to most semiconductor junctions, at the Si/electrolyte interface a depletion zone is formed. The width of this depletion zone depends on the doping and may explain the different pore sizes found in p^--and p^+-type doped silicon. In addition, the depletion layer width depends on the surface curvature: the anodization preferentially occurs at the pore tips where the curvature is largest. Moreover, when the depletion zones of adjacent pores meet each other, the current flow is suddenly pinched off. Further Si etching is blocked, and pore collapsing is prevented. For this reason, the reaction is self-limited in the hashed anodization regimes of **Figure 5**, and leads to a porous structure rather than to electropolishing. As further practical consequence, in stationary conditions, the porosity remains approximately constant, whereas the overall thickness of the PSi layer grows essentially linearly in time.

Back-side illumination (most conveniently achieved using cells of type B, **Figure 2**) in potentiostatic

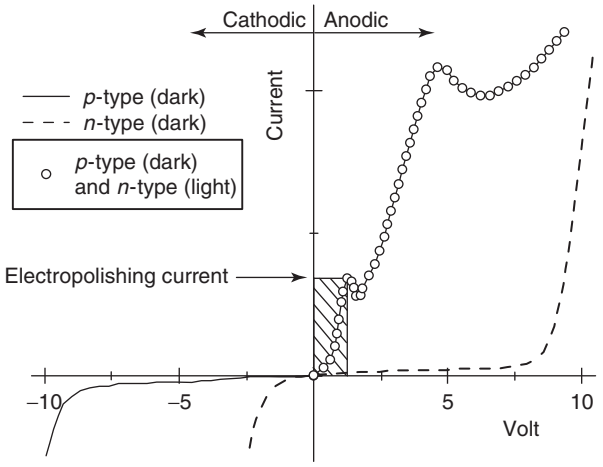


Figure 5 Typical *I–V* characteristics of an electrochemical cell for PSi fabrication (**Figure 2**). The hashed region corresponds to the useful regime where PSi can be achieved, assuming the *I–V* characteristic marked with hollow circles. In the anodic regime, the characteristics of a cell with *n*-type Si will lie in the region bounded by the characteristic in dark (dashed line) and in full light (hollow circles).

regime, with *n*-type Si, allows a useful, further control on the hole injection, both in terms of carrier flux (which is proportional to illumination intensity) and of localization of the injection, which is concentrated in the region of the pore tips, acting as hole collectors. The etching process leads to a very regular pore growth, which is most effectively exploited to fabricate macroporous photonic crystal devices (**Figure 4**).

Drying Techniques

Due to large capillary stress, drying of samples is a critical step and can result in extended cracking if special procedures are not followed. Methods to reduce or eliminate the capillary stress include pentane drying, supercritical drying, freeze drying, and slow evaporation rates. Pentane drying is the easiest to implement. Pentane has a very low surface tension, and shows no chemical interaction with PSi (unlike ethanol). Using pentane as drying liquid enables to strongly reduce the capillary tension, but since water and pentane are nonmiscible liquids, ethanol or methanol have to be used as intermediate liquids. Using this drying technique, PSi layers with porosity values up to 90% and thickness up to 5 μm exhibit no cracking pattern after drying. Supercritical drying requires a specific apparatus but is more effective in preventing cracking of the film so that PSi layers of porosity values up to 95% were demonstrated.

Oxidation Techniques

As soon as PSi samples are dried, hydrogen surface passivation is gradually replaced by native oxide. This chemical change affects most PSi properties, for example, photoluminescence and electrical conductivity. Since the characteristics of the native oxide strongly depend on many factors such as the storage conditions, evolution of sample properties is often poorly characterizable. To stabilize PSi properties, oxidation is often used. Oxidation implies the formation of a layer containing the original Si atoms. Therefore, it also reduces the nanocrystallite size, with remarkable impact on photoluminescence emission energy (blue-shift). The employment of controlled oxidation procedures is therefore a valuable post-anodization option toward a better control of PSi properties and stabilization.

Tested oxidation procedures on PSi include anodic oxidation, chemical oxidation, thermal oxidation, plasma-assisted oxidation, and irradiation-enhanced oxidation. Anodic oxidation is an electrochemical process, in which oxide layers are formed with current injection. As such, the technique is structurally selective: more conductive paths get more oxidized. This selectivity is particularly interesting for light-emitting diode (LED) applications, where more conductive paths are usually associated to large injection channels, which are poorly luminescent, and their oxidation results in enhanced electroluminescence quantum efficiency. Another application of controlled oxidation – mostly thermal – is the fabrication of small, blue-luminescent PSi nanocrystallites.

Structural Properties

Structural properties of PSi have been investigated by electron microscopy, scanning probe microscopy, X-ray scattering techniques, X-ray absorption techniques, and Raman spectroscopy. In particular, X-ray absorption spectroscopy can be performed not only by extended X-ray absorption fine structure (EXAFS) analysis, but also by X-ray excitation of optical luminescence (XEOL), a convenient tool for luminescent samples, in which absorption is probed via the measurement of the corresponding optical emission. A general feature resulting from the comparison of these techniques is that highly luminescent PSi is a composition of extensively connected Si crystals of nanometric dimensions (typically, 1–5 nm in radius). Sizes smaller than Si Bohr radius ($\sim$4.9 nm), imply significant quantum confinement effects. Raman measurements are also often used to demonstrate crystallinity in PSi skeleton and infer information on the size of Si nanocrystals.

Accurate determination of pore size distribution in mesoporous Si is usually given by the analysis of adsorption isotherms of gases at low temperature (Brunauer–Emmett–Teller, or BET, method). The

physical adsorption by a porous surface is increased relative to a nonporous one because of capillary condensation in pores. This increase in adsorption starts when the gas pressure is high enough to fill the smallest pores.

Chemical Properties

Surface chemical composition of PSi is best probed with Fourier transform infrared (FTIR) spectroscopy. FTIR signal in PSi is larger and easier to measure than in bulk Si due to much larger specific area. Chemical bonds and their IR resonance positions detected in PSi are shown in **Table 5**. In fresh, as-prepared samples, oxygen is normally absent, the dominant bonds being $Si–H_x$ groups ($x = 1$, 2, or 3). Aging is observed as a slow replacement of H by O bonds in PSi. Also, luminescence fatigue is explained by photochemical reactions occurring on the surface of the Si nanocrystals.

Electronic Properties

The first and most favorable explanation for the visible emission in PSi is the quantum confinement of excitons in nanometer-sized silicon. An empirical law, based on the effective mass approximation, links the energy gap E_g of silicon nanocrystals with their sizes $\ell : E_g(\ell) = E_g^{Si} + 88.34/(\ell)^{1.37}(\text{eV})$, when ℓ is given in Å. The energy gap opening is given by an equal energy shift of the bottom of the conduction band to high energy and of the top of the valence band to low energy. The nature of the energygap is still indirect even though a quasidirect bandgap can be formed in ultrasmall Si nanocrystals.

Electrical Properties

Electrical resistivity in PSi is five orders of magnitude higher than in intrinsic Si, because PSi is depleted of free carriers. Depletion can occur either because of the energy gap widening from quantum confinement which reduces the thermal generation of free carriers, or because of trapping of free carriers. Trapping can occur during the preparation of PSi either because the binding energy of dopant impurities are increased or because of the formation of surface states. It has been demonstrated that the dopants are still present in essentially unchanged concentration after the etching, but are in a neutral state.

The electrical transport is mainly affected by the disordered structure of the Si skeleton which restricts the conductive paths to a highly constrained geometry, which for certain porosities forms a percolated or fractal geometry. As a consequence, conductivity is thermally activated, strongly frequency dependent, and highly dispersive. Several models have been proposed to explain the electrical transport properties. They differ on the transport paths and mechanisms. The proposed transport paths range from transport in the Si nanocrystals (with diffusion or tunneling between the Si nanocrystals or at their surface) to transport in the amorphous and disordered matrix surrounding the nanocrystals, or through both. The suggested mechanisms are band transport, activated hopping in band tail, trap-controlled hopping through nanocrystals, activated deep states hopping, Pool–Frenkle processes, and activated hopping in fractal networks.

Given the large specific area per unit volume, the electrical transport is strongly influenced by external factors such as residual electrolyte and ambient atmosphere. The latter property is very interesting for sensor applications. Certain gases – for example, NO_2 – have the capability of modifying the free carrier population. Changes in the electrical conductivity in the presence of sub-ppm concentrations of such gases can be detected at room temperature operation. Other gaseous species (e.g., polar liquid vapors) also affect electrical transport via electric field interactions with charge carriers.

Table 5 Wave number positions and attributions of the absorption peaks observed in several PSi samples by Fourier transform infrared absorption FTIR measurements

Peak position (cm^{-1})	*Attribution*	*Peak position* (cm^{-1})	*Attribution*
3610	OH stretch (SiOH)	1463	CH_3 asymmetric deformed
3452	OH stretch (H_2O)	1230	$SiCH_3$ bending
2958	CH stretch (CH_3)	1056–1160	SiO stretching in O–SiO and C–SiO
2927	CH stretch (CH_2)		
2856	CH stretch (CH)	979	SiH bending in SI_2–H–SiH
2248	SiH stretch (O_3–SiH)	948	SiH bending in Si_2–H–SiH
2197	SiH stretch (SiO_2–SiH)	906	SiH_2 scissor
2136	SiH stretch (Si_2O–SiH)	856	SiH_2 wagging
2116	SiH stretch (Si_2H–SiH)	827	SiO bending in O–Si–O
2087	SiH stretch (Si_3–SiH)	661	SiH wagging
1720	CO	624	SiH bending in (Si_3SiH)

Optical Properties

Light Diffusion and Scattering

The dimension l of the PSi structures (i.e., the pore size and the porous layer thickness), compared to the optical wavelengths λ, can range from $l \ll \lambda$ all the way to $l \approx \lambda$. There are accordingly two different regimes of light propagation and interaction with PSi.

The first regime is $l \ll \lambda$. The radiation fields cannot resolve the PSi structures, and the interaction can be conveniently described by means of an "effective medium approximation," where a macroscopic dielectric constant (or a refractive index) is evaluated as a suitable combination of the dielectric constants of Si and air. The result is, in general, a complex dielectric function of space. Effective medium theories, which are routinely used for PSi, are enumerated in **Table 6**. Refractive index values are reported in **Figure 6**.

The second regime is $l \sim \lambda$. These kinds of structures are called mesoscopic: the structural size is small compared to the wave coherence length, that is, the maximum length along which phase memory and coherent phenomena are observable. The porous geometrical structure strongly influences light transport and interaction. Useful mathematical models take advantage either of the Bloch theorem (in the case of periodical structures (photonic crystals)), or of statistical and Green's function methods (in the case of aperiodic or random structures).

Mixed regimes are also possible in PSi. The most important case is the one-dimensional (1D) mesoscopic structure: the anodization parameters are set to achieve microporous PSi (size of pores $\ll \lambda$, first regime), but the anodizing current density is modulated to obtain periodical or aperiodical thin layer structures (thickness $\sim \lambda$, second regime along the PSi growth direction). Each layer is different from adjacent layers in terms of porosity and thus of effective dielectric constant (and effective refractive index). Compared to other thin-film growth techniques, PSi has the advantage of allowing a continuous tuning of the refractive index over a wide range ($1.4 \leqslant n_{\text{eff}} \leqslant 2.4$), and of being a fast and cheap technique that can lead to structures of several hundreds of layers, with typical fabrication times ranging from few minutes to few hours.

In the presence of high-index liquids (and/or their vapors), the refractive index of the porous layer changes, affecting light transport properties. The effect can be exploited for gas sensors and biosensors applications.

Table 6 Useful effective medium approximation for the dielectric function of PSi

Theory	*Formula*
Bruggeman	$P\frac{\varepsilon_M - \varepsilon_{\text{eff}}}{\varepsilon_M + 2\varepsilon_{\text{eff}}} + (1-P)\frac{\varepsilon - \varepsilon_{\text{eff}}}{\varepsilon + 2\varepsilon_{\text{eff}}} = 0$
Maxwell Garnett	$\frac{\varepsilon_{\text{eff}} - \varepsilon_M}{\varepsilon_{\text{eff}} + 2\varepsilon_M} = (1-P)\frac{\varepsilon - \varepsilon_M}{\varepsilon + 2\varepsilon_M}$
Looyenga	$\varepsilon_{\text{eff}}^{1/3} = (1-P)\varepsilon^{1/3} + P\varepsilon_M^{1/3}$

ε, dielectric function of Si; ε_M, dielectric function of host material (normally: air); ε_{eff}, calculated effective dielectric function; P, porosity.

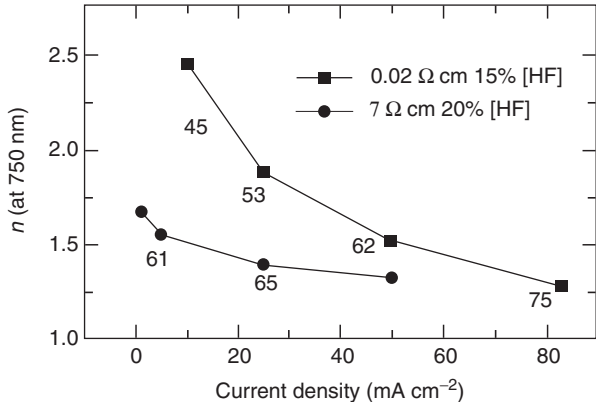


Figure 6 Refractive index as a function of current density and porosity for two different substrate doping levels.

Photoluminescence

Microporous PSi structures have been reported to luminesce efficiently in the near infrared (IR) (0.8 eV), in the whole visible range, and in the near ultraviolet (UV) (**Figure 7**).

Such a broad range of emission energies arises from a number of clearly distinct luminescent bands, which are listed in **Tables 7–9**. In addition, PSi has been used as an active host for rare-earth impurities, for example, Nd or Er, or dye solutions. Direct energy transfer between PSi and the impurity or dye is demonstrated. The most efficient and studied emission is the slow (S) band because it can be excited by electrical injection. The most popular model for the photoluminescence – especially for the S-band – is excitonic radiative recombination subject to quantum confinement. Other models are surface states models, the siloxene model, defect models, the surface hydrides model, and the hydrogenated amorphous silicon model. Coexistence of more than one mechanism is likely.

S-Band It can be tuned from close to the bulk silicon bandgap through the whole visible range. The large spectral width comes from inhomogeneous broadening and vibronic coupling of the radiative transitions. The S-band efficiency is not proportional to the inner surface area: a "threshold" porosity has

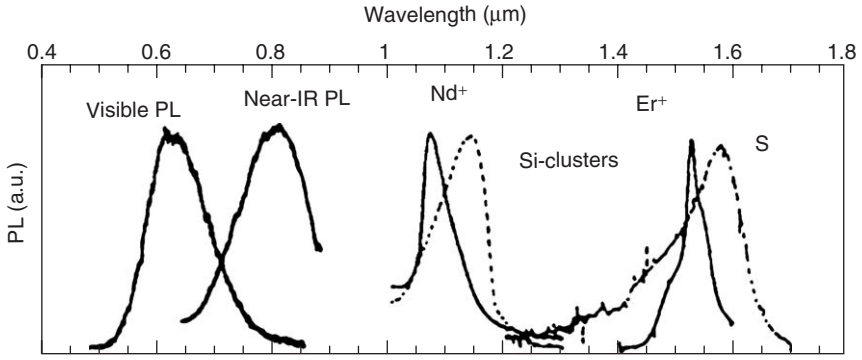


Figure 7 Room temperature photoluminescence spectra for various PSi structures which have been oxidized or implanted with some selected impurities. (Reproduced from Fauchet PM, Tsybeskov L, Duttagupta SP, and Hirschman KD (1997) Stable photoluminescence and electroluminescence from porous silicon. *Thin Solid Films* 297(1–2): 254–260, with permission from Elsevier.)

Table 7 PSi luminescence bands

Spectral range	*Peak wavelength* (nm)	*Label*
UV	~350	UV band
Blue-green	~470	Fast (F) band
Blue-red	400–800	Slow (S) band
Near IR	1100–1500	IR band

to be exceeded to achieve an efficient luminescence. Postanodization chemical etching in HF, corresponding to a porosity increase, results in a strong rise in PL efficiency and a blue shift of the visible band. External quantum efficiencies higher than 0.1% are obtainable from high-porosity PSi layers of all types, but efficiency normally decreases in the order n^--type, p^--type, n^+-type, and p^+-type doped PSi. High porosity is essential for high visible PL efficiency. The inefficient luminescence observed from inhomogeneous material of low porosity originates from microscopic areas of high porosity. Isolated nanocrystals, produced by dispersing a colloidal suspension of PSi fragments on a glass coverslip, show external quantum efficiencies $\geqslant 88\%$, with a number of bright-to-dark nanocrystals in the suspension of only 2.8%. Thus, the average quantum efficiency of a PSi layer of 1% results from a statistical distribution of high and low quantum efficiency nanocrystals.

Phonon-assisted luminescence (vibronic-coupled transitions) is demonstrated by fine structure in resonantly excited photoluminescence at low temperature. Thus, the electronic and the vibronic structure of the luminescence of PSi are similar to crystalline Si, with an energy shift consistent with quantum confinement. At large confinement energies (>0.7 eV), no-phonon quasidirect transitions dominate, because of the breakdown of k-conservation under strong confinement.

The luminescence is polarized parallel to the exciting light. This is measured by recording the polarization ratio $\rho = (I_\parallel - I_\perp)/(I_\parallel + I_\perp)$. ρ is zero on the IR band and nonzero on the S-band (tending to zero as the energy approaches that of the Si bandgap). The interpretation is that PSi is an ensemble of randomly oriented aspherical nanocrystals, preferentially aligned along the $\langle 1\,0\,0 \rangle$ direction. Recombining carriers have bulk-like wave functions which are sensitive to the shape of the nanocrystals.

Aging effects include a blue shift in the S-band and changes in the photoluminescence efficiency (both increasing and decreasing have been observed), which can be explained in terms of surface oxidation. Anticorrelation exists between the dangling bond density and the luminescence efficiency; thus, the effect of the oxidation depends on its passivation properties.

Other Bands The F-band is observed only in oxidized PSi, and it is probably originated from contaminated or defective Si oxide. Annealing in water vapor activated the blue emission indicating a possible major role of adsorbed hydroxyls in the emission process. The IR band is weak at room temperature, and becomes much stronger at cryogenic temperatures. Its origin seems to be related to dangling bonds, although no direct correlation has been demonstrated.

Absorption

The absorption coefficient has been measured in PSi by optical transmission, photoluminescence excitation (PLE), and photothermal deflection spectroscopy (PDL). Transmission spectra are shifted toward higher energy compared to bulk Si with a shift, which increases with increasing porosity. This observation is consistent with quantum confinement model. The line shape analysis of the absorption coefficient shows that its energy dependence follows a trend like that of an indirect gap semiconductor similar to Si, but displaced to higher energy.

Table 8 Some spectral characteristics of the S-band

Property	*Typical values*	*Comments*
Peak wavelength	1100–400 nm	At 300 K, depends on porosity
PL external quantum efficiency	$\geqslant$5%	At 300 K, depends on porosity
FWHM	0.3 eV	At 300 K (8 meV in PSi microcavities)
PL decay times	$\simeq 10\,\mu s$	Strongly dependent on wavelengths, temperature and aging condition
Polarizability ratio	$P \leqslant 0.2$	
Fine structure under resonant excitation	Phonon replica at 56 and 19 meV	Heavily aged PSi, energies consistent with Si phonons

Table 9 Some spectral characteristics of the F-band

Property	*Typical values*	*Comments*
Peak wavelength	480 nm	UV excitation at 300 K
PL efficiency	$\geqslant$0.1%	External quantum efficiency UV excitation at 300 K
FWHM	0.4 eV	UV excitation at 300 K
PL decay times	1 ns	Independent on wavelength and excitation conditions

A characteristic feature of PSi is the large displacement (Stokes shift) between the absorption edge and the emission peak energy (see **Figure 8**). The existence of localized states that can account for the observed shift has been demonstrated. Such states can lead to self-trapped excitons at some surface bonds of Si crystallites, such as Si–Si dimers and Si–O double bonds, giving reduced dependence of the luminescence energy on size, and a very large Stokes shift. Other models for the Stokes shift are based on Si quantum dot relaxation from distorted configurations, giving rise to new transitions involving localized states that lower the emission threshold with respect to the absorption.

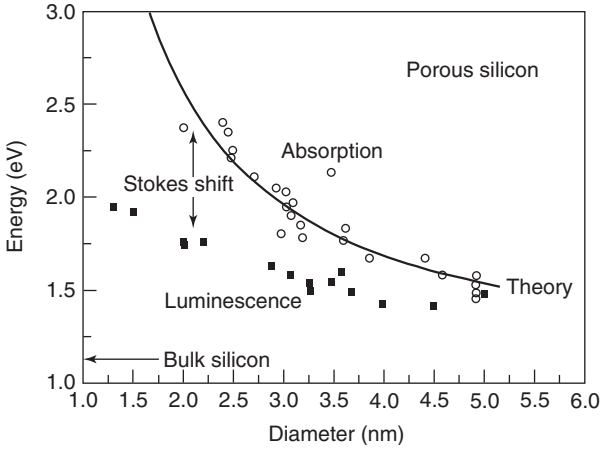


Figure 8 Compilation of optical bandgaps of PSi samples obtained from optical absorption (unfilled symbols) and luminescence (filled symbols). The lines represent calculated values with the empirical law reported in the text.

Applications

Important actual or perspective applications of PSi are shown in **Table 10**.

Light Emitting Diodes (LEDs)

Si *p*/*n* junction emits light both in forward and in reverse bias conditions. Power efficiency is $\sim 10^{-4}$ (10^{-8}) at 1.1 μm wavelength in forward (reverse) bias. Also in PSi based devices, electroluminescence (EL) is observed both in the IR and in the visible. Efficiency is higher than in bulk Si diodes, depending on the diode structure and on the PSi contact method. Initially, EL was observed with wet contacts and later with solid contacts. Major achievements in terms of quantum efficiency are summarized in **Figure 9**.

Sensors, Biosensors, Medical Applications

PSi is a twofold promising material for sensor applications. On one side, its electrical and optical properties strongly depend on the environment, because of its large specific area. Useful sensing parameters include, for example, the electrical conductivity and photoluminescence. Most sensing parameters can effectively work at room temperature. On the other side, it supplies a handy template to hold chemical and biological species. Target species can be immobilized in the PSi matrix, a practice which is especially useful for biological samples – such as oligonucleotides, biotin, or antibodies. Techniques for immobilization range from physical adsorption to the replacement of hydride bonds with Si alkyls, and to antibody bonding at functionalized PSi surface with subsequent antibody–antigen interactions. Species identification can be performed by probing optical, electrical, or chemical properties. Examples include, respectively, reflectivity or wave-guiding measurements (which depend on the dielectric constant of the species, as already discussed), conductance measurements, and

Table 10 Potential application areas of PSi

Application area	*Role of PS*	*Key property*
Optoelectronics	LED	Efficient electroluminescence
	Waveguide	Tunability of refractive index
	Field emitter	Hot carrier emission
	Optical memory	Nonlinear properties
Micro-optics	Fabry-Perot filters	Refractive index modulation
	Photonic bandgap structures	Regular macropore array
	All optical switching	Highly nonlinear properties
Energy conversion	Antireflection coatings	Low refractive index
	Photo-electrochemical cells	Photocorrosion cells
	Solar cells	
Environmental monitoring	Gas sensing	Ambient sensitive properties
Microelectronics	Micro-capacitor	High specific surface area
	Insulator layer	High resistance
	Low-k material	Electrical properties
Wafer technology	Buffer layer in heteroepitaxy	Variable lattice parameter
	SOI wafers	High etch selectivity
Micromachining	Thick sacrificial layer	Highly controllable Etching parameters
Biotechnology	Tissue bonding	Tunable chemical reactivity
	Biosensors	Enzyme immobilization

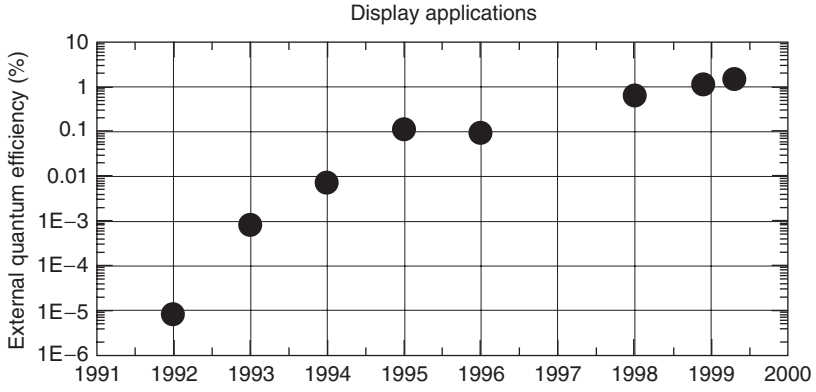


Figure 9 Time evolution of the external quantum efficiency of porous silicon-based LEDs for display applications.

pH measurements. Biocompatibility of PSi has suggested medical applications, such as *in vivo* slow release of drugs, *in vivo* diagnostic tests, and template for bone growth.

See also: Dielectric Function; Disorder and Localization Theory; Effective Masses; Electrons and Holes; Elemental Semiconductors, Electronic States of; Excitons: Theory; Interaction of Light and Matter; Light Emitting Diodes; Luminescence; Meso- and Nanostructures; Nanostructures, Electronic Structure of; Nanostructures, Optical Properties of; Optical Absorption and Reflectance; Optical Properties of Materials; Periodicity and Lattices; Photonic Bandgap Materials, Electronic States of; Quantum Devices of Reduced Dimensionality; Quasicrystals; Quasicrystals, Electronic Structure of; Semiconductor and Metallic Clusters, Electronic Properties of; Semiconductor Nanostructures; Semiconductor Optics; Semiconductors, General Properties; Semiconductors, Optical and Infrared Properties of; Sensors; Silicon, History of; Small Particles and Clusters, Optical Properties of; Surfaces, Optical Properties of; Thin Films, Mechanical Behavior of; Waves in Random Media.

PACS: 61.46.+w; 63.22.+m; 68.65.Hb; 73.21.La; 73.22.−f; 73.63.Bd; 78.67.Bf; 81.05.Rm; 81.07.Bc; 81.16.−c; 81.65.Cf; 82.45.Cc; 82.45.Vp; 82.45.Yz; 85.35.−p; 87.83.+a

Further Reading

Amato G, Delerue C, and von Bardeleben HJ (eds.) (1997) Optical and structural properties of porous silicon nanostructures. In: Manasreh MO (series edn.) *Opto-Electronic Properties of Semiconductors and Superlattices*, vol. 5, pp. 1–644. Newark: Gordon and Breach.

Bisi O, Ossicini S, and Pavesi L (2000) Porous silicon: a quantum sponge structure for silicon based optoelectronics. *Surface Science Report* 38(1–3): 5–126.

Brus L (1996) Semiconductor colloids: individual nanocrystals, opals and porous silicon. *Current Opinion Colloid and Interface Science* 1(2): 197–201.

Buriak JM (2002) Organometallic chemistry on silicon and germanium surfaces. *Chemical Review* 102(5): 1272–1308.

Canham LT (1990) Silicon quantum wire fabrication by electrochemical and chemical dissolution of wafers. *Applied Physics Letters* 57(10): 1046–1048.

Cullis AG, Canham LT, and Calcott PDJ (1997) The structural and luminescence properties of porous silicon. *Journal of Applied Physics* 82(3): 909–965.

Filler MA and Bent SF (2003) The surface as molecular reagent: organic chemistry at the semiconductor interface. *Progress in Surface Science* 73: 1–56.

Föll H, Christophersen M, Carstensen J, and Hasse G (2002) Formation and application of porous silicon. *Material Science and Engineering R* 39(4): 93–141.

Gelloz B and Koshida N (2003) Electroluminescence of nanocrystalline porous silicon devices. In: Nalwa HS and Rohwer LS (eds.) *Handbook of Luminescence, Display Materials, and Devices*, pp. 127–156. Stevenson Ranch, CA: American Scientific Publisher.

Gösele U and Lehmann V (1995) Light-emitting porous silicon, matter. *Chemical Physics* 40(4): 253–259.

Ossicini S, Pavesi L, and Priolo F (2003) *Light Emitting Silicon for Microphotonics, Springer Tracts in Modern Physics*, vol. 194. Berlin: Springer.

Sailor MJ, Heinrich JL, and Lauerhaas JM (1997) Luminescent porous silicon: synthesis, chemistry, and applications. *Studies in Surface Science and Catalysis*. 103: 209–235.

Smith RL and Collins SD (1992) Porous silicon formation mechanisms. *Journal of Applied Physics* 71(8): R1–R22.

Stewart MP and Buriak JM (2000) Chemical and biological applications of porous silicon technology. *Advanced Materials* 12(12): 859–869.

Yerokhov VY and Melnyk II (1999) Porous silicon in solar cell structures: a review of achievements and modern directions of further use. *Renewable and Sustainable Energy Reviews* 3(4): 291–322.

Powder Processing: Models and Simulations

A B Yu, The University of New South Wales, Sydney, NSW, Australia

Introduction

Granular materials, which can be either wet or dry and range in size from nanometers to centimeters, are widely encountered as particles or powder in industries and in nature. As with solids, they can withstand deformation and form heaps; as with liquids, they can flow; as with gases, they exhibit compressibility. Corresponding to the fluid- and solid-like modes, they show different flow regimes: quasi-static regime, rapid flow regime, and a transitional regime that lies in between. These features give rise to another state of matter that is poorly understood. The development of a general theory to describe the packing (statics) and flow (dynamics) of granular materials has been a problem challenging the scientific community for years.

Essentially, the existing approaches to modeling granular materials can be classified into two categories: the continuum approach at a macroscopic level and the discrete approach at a microscopic level. In the continuum approach, the macroscopic behavior is described by balance equations, for example, mass and momentum as used in the two fluid model (TFM), closed with constitutive relations together with initial and boundary conditions. This approach is preferred in process modeling and applied research because of its computational convenience. However, its effective use depends heavily on constitutive or closure relations and the momentum exchange between particles of different type. In the past, different theories have been devised for different materials and for different flow regimes. For example, models have been proposed to derive the constitutive equations for the rate-independent deformation of granular materials based on either the plasticity theory or the double shearing theory; the rapid flow of granular materials has been described by extending the kinetic theory of dense gases; the transitional regime that involves both collisional and frictional mechanisms is studied by use of the kinetic theory combined with the Mohr–Coulomb quasi-static theory. However, to date, there is no accepted continuum theory applicable to all flow conditions. As a result, phenomenological assumptions which have very limited application are necessary to obtain the constitutive relations and boundary conditions.

The discrete approach is based on the analysis of the motion of individual particles and has the advantage that there is no need for global assumptions on solids such as steady-state behavior, uniform constituency, and/or constitutive relations. Various methods have been developed in the past. A major type of discrete approach is based on the so-called discrete element method (DEM), originally developed for rock mechanics. The method considers a finite number of discrete particles interacting by means of contact and noncontact forces, and every particle in a considered system is described by Newton's equations of motion. In principle, it is similar to molecular dynamic simulation (MDS) but the forces involved differ because of the difference in

time and length scales. DEM-based simulation has been recognized as an effective method to study the fundamentals of granular materials and assess the performance of industrial processes. By means of a proper average method, it can also generate information to support continuum modeling.

This article is focused on the discrete approach. It first describes the principles and key features of the simulation method, and then discusses its application to the study of a few typical static and dynamic particle systems. Emphasis is given to the validity analysis and the usefulness of the resulting particle scale information in elucidating the fundamentals of granular materials, achieved through illustrative results.

Simulation Method

Governing Equations

A particle in a considered system can have two types of motion: translational and rotational, determined by Newton's second law of motion. During its movement, the particle may collide with its neighbor particles or wall at the contact points and interact with the surrounding fluid, through which the momentum and energy are exchanged. For the systems considered in this article, at any time t, the equations governing the translational and rotational motions of particle i are

$$m_i \frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = \boldsymbol{f}_{\mathrm{pf},i} + m_i\boldsymbol{g} + \sum_{j=1}^{k_i}\left(\boldsymbol{f}_{\mathrm{c},ij} + \boldsymbol{f}_{\mathrm{d},ij} + \boldsymbol{f}_{\mathrm{v},ij}\right) \quad [1]$$

and

$$I_i \frac{\mathrm{d}\omega_i}{\mathrm{d}t} = \sum_{j=1}^{k_i}\left(\boldsymbol{T}_{\mathrm{c},ij} + \boldsymbol{T}_{\mathrm{r},ij}\right) \quad [2]$$

where m_i, I_i, $\boldsymbol{v}_i$, and $\boldsymbol{\omega}_i$ are, respectively, the mass, moment of inertia, translational, and rotational velocities of the particle. The forces involved are: the particle–fluid interaction force, $\boldsymbol{f}_{\mathrm{pf},i}$, gravitational force, $m_i\boldsymbol{g}$, and interparticle forces between particles i and j. The torques include the interparticle torque $\boldsymbol{T}_{\mathrm{c},ij}$ and rolling friction torque $\boldsymbol{T}_{\mathrm{r},ij}$. For multiple interactions, the interparticle forces and torques are summed for k_i particles interacting with particle i.

Particle flow is often coupled with fluid (gas and/or liquid) flow. Various approaches have been proposed to simulate such a coupled particle–fluid flow. The time and length scales for fluid flow can range from discrete (MDS, lattice Boltzmann (LB), pseudo-particle (PP)) to continuum (direct numerical simulation (DNS), large eddy simulation (LES)), and other conventional techniques such as the TFM described above. Relative to particle phase, they can also be categorized into three groups: sub-particle, pseudo-particle, and computational cell. **Table 1** lists a few representative combinations of different length scales for fluid and particle phases and their relative merits in different aspects.

At this stage of development, the difficulty in particle–fluid flow modeling is mainly related to the solid phase rather than the fluid phase. Therefore, the

Table 1 Typical models for particle–fluid flow and their relative merits

Model type				
Length scale for fluid phase	Sub-particle (discrete or continuum)	Pseudo-particle (discrete)	Computational cell (continuum)	Computational cell (continuum)
Length scale for particle phase	Particle (discrete)	Particle (discrete)	Particle (discrete)	Computational cell (continuum)
Nature of coupling	Discrete + discrete or continuum + discrete	Discrete + discrete	Continuum + discrete	Continuum + continuum
Example	LB-DEM or DNS	PP-DEM	CCDM or CFD-DEM	TFM
Closure of equations	Yes (but may experience numerical difficulty for systems with strong particle–particle interactions)	No (difficulty to determine physical properties of a pseudo-particle)	Yes	No (constitutive relation for solid phase and phase interactions not generally available)
Incorporation of distribution effects of dispersed, solid phase	Yes	Yes	Yes	No
Computational effort	Extremely demanding	Very demanding	Demanding	Acceptable
Suitability for engineering application in relation to process modeling and control	Extremely difficult	Very difficult	Difficult	Easy
Suitability for fundamental research in relation to fluid physics	Most acceptable ($\boldsymbol{f}_{\mathrm{pf}}$ can be used for CCDM)	Acceptable (but only valid for well-defined PP system)	Acceptable	No

so-called combined continuum and discrete model (CCDM) or CFD–DEM is attractive because of its computational convenience as compared to DNS or LB–DEM, and capability to capture particle physics as compared to TFM. In CCDM, the continuum fluid field is calculated from the continuity and the Navier–Stokes equations based on the local mean variables over a computational cell, which are given by

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot (\varepsilon \boldsymbol{u}) = 0 \qquad [3]$$

and

$$\frac{\partial (\rho_f \varepsilon \boldsymbol{u})}{\partial t} + \nabla \cdot (\rho_f \varepsilon \boldsymbol{u}\boldsymbol{u}) = -\nabla p - \boldsymbol{F}_{pf} + \nabla \cdot (\varepsilon \boldsymbol{\tau}) + \rho_f \varepsilon \boldsymbol{g} \qquad [4]$$

where $\boldsymbol{u}$, p, and $\boldsymbol{\tau}$ are, respectively, the fluid velocity, pressure, and viscous stress tensor; $\boldsymbol{F}_{pf} = \left(\sum_{i=1}^{k_c} \boldsymbol{f}_{pf,i}/\Delta V\right)$ is the particle–fluid interaction force, and ε, ΔV, and k_c are respectively the porosity, volume, and number of particles in a computational cell.

Models for Particle–Particle and Particle–Fluid Interactions

The forces acting on a particle result from particle–particle and particle–fluid interactions. For particles, these interactions include the forces due to direct or nondirect contacts between particles, as shown in **Figure 1**. The direct contact forces include the contact force and viscous contact damping force. The nondirect contact forces include the forces associated with fine particles such as the van der Waals force, electrostatic force, capillary and viscous forces, and solid-bonding forces. Particle–fluid interactions include buoyancy, drag force, and so on. These forces have been implemented by various investigators in various ways. To date, the equations to determine some of the forces are not fully developed. Therefore, development of a more comprehensive theory and experimental techniques to quantify the interaction forces between particles and between particle and fluid under various conditions will continue to be an active research area for years to come. **Table 2** lists the equations employed to calculate the forces and torques for the particle systems considered in this article; they are also commonly used in the literature.

Numerical Solution and Boundary Treatments

Equations [1] and [2] for solid flow can be solved by an explicit time integration method, and eqns [3] and [4] for fluid flow by use of a standard computational fluid dynamics (CFD) algorithm, facilitated with suitable boundary conditions. For fluid phase, the no-slip boundary condition often applies to bed walls, together with other (inlet or outlet) conditions specified. For particle phase, the interparticle force models are also applied to the interaction between a particle and a wall, with the corresponding wall properties used; however, the wall is assumed to be so rigid that no displacement and movement of the wall result from this interaction. Depending on the geometry and flow conditions, periodic boundary conditions can be implemented to reduce the computational effort. If fluid flow is negligible, simulations can be conducted without the need for solving eqns [3] and [4].

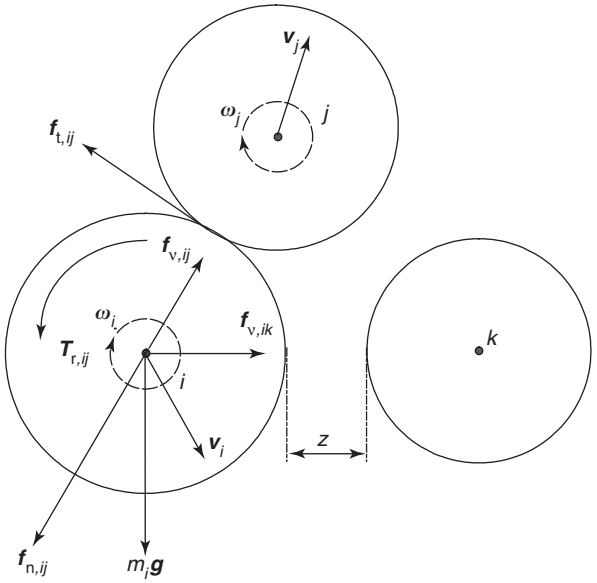


Figure 1 Schematic illustration of the forces acting on particle *i* from contacting particle *j* and noncontacting particle *k*. (Modified with permission from Yang RY, Zou RP, and Yu AB (2000) Computer simulation of the packing of fine particles. *Physical Review E* 62: 3900–3908; © American Physical Society and Yang RY, Zou RP, and Yu AB (2003) Effect of material properties on the packing of particles. *Journal of Applied Physics* 94: 3025–3034; © American Institute of Physics.)

The modeling of the solid flow by DEM is at the individual particle level, whilst the gas flow by CFD is at the computational cell level. As shown in **Figure 2**, their coupling is numerically achieved as follows. At each time step, DEM will give information, such as the positions and velocities of individual particles, for the evaluation of porosity and volumetric fluid drag force in a computational cell. CFD will then use these data to determine the gas flow field which then yields the fluid drag forces acting on individual particles. Incorporating the resulting forces into DEM will produce information about the motion of individual particles for the next time step. The fluid drag force acting on individual particles will react on the fluid phase from the particles, so that Newton's third law of motion is satisfied.

Table 2 Components of forces and torque acting on particle *i*

Force or torque	*Symbol*	*Equation*
Normal forces		
Contact	$\boldsymbol{f}_{\mathrm{cn},ij}$	$-\kappa_i \boldsymbol{\delta}_{\mathrm{c},ij}$
Damping	$\boldsymbol{f}_{\mathrm{dn},ij}$	$-\eta_i \boldsymbol{v}_{\mathrm{n},ij}$
Tangential forces		
Contact	$\boldsymbol{f}_{\mathrm{ct},ij}$	$\min\left(\gamma_{\mathrm{s}} \boldsymbol{f}_{\mathrm{cn},ij}, \kappa_i \boldsymbol{\delta}_{\mathrm{t},ij}\right)\dfrac{\boldsymbol{\delta}_{\mathrm{t},ij}}{\delta_{\mathrm{t},ij}}$
Damping	$\boldsymbol{f}_{\mathrm{dt},ij}$	$-\eta_i \boldsymbol{v}_{\mathrm{t},ij}$
Torques		
Interparticle	$\boldsymbol{T}_{\mathrm{c},ij}$	$\boldsymbol{R}_i \times (\boldsymbol{f}_{\mathrm{ct},ij} + \boldsymbol{f}_{\mathrm{dt},ij})$
Rolling friction	$\boldsymbol{T}_{\mathrm{r},ij}$	$-\gamma_{\mathrm{r}} R_i \boldsymbol{f}_{\mathrm{cn},ij} \hat{\boldsymbol{\omega}}_i$
Gravity force	$\boldsymbol{G}_i$	$m_i\, \boldsymbol{g}$
van der Waals force	$\boldsymbol{f}_{\mathrm{v},ij}$	$\dfrac{\mathrm{H}d}{6[\max(z, z_0)]^2}\dfrac{\boldsymbol{R}_i}{R_i}$
Fluid drag force	$\boldsymbol{f}_{\mathrm{pf},i}$	$0.5 c_{d0,i} \rho_f \pi \boldsymbol{R}_i^2 \varepsilon_i^2 \lvert\boldsymbol{u}_i - \boldsymbol{v}_i\rvert (\boldsymbol{u}_i - \boldsymbol{v}_i) \varepsilon_i^{-(\chi+1)}$

where: $\boldsymbol{V}_{ij} = \boldsymbol{V}_j - \boldsymbol{V}_i + \omega_j \times \boldsymbol{R}_j - \omega_i \times \boldsymbol{R}_i$, $\boldsymbol{V}_{n,ij} = (\boldsymbol{V}_{ij} \cdot n) \cdot n$, $\boldsymbol{V}_{t,ij} = (\boldsymbol{V}_{ij} \times \boldsymbol{n}) \times \boldsymbol{n}$, $\hat{\omega}_i = \dfrac{\boldsymbol{\omega}_i}{\omega_i}$,

$n = \dfrac{\boldsymbol{R}_i}{R_i}$, $\chi = 3.7 - 0.65 \exp\left[-\dfrac{(1.5 - \log_{10}\mathrm{Re}_{p,i})^2}{2}\right]$, $c_{d0,i} = \left(0.63 + \dfrac{4.8}{\mathrm{Re}_{p,i}^{0.5}}\right)^2$, $\mathrm{Re}_{p,i} = \dfrac{2\rho_f R_i \varepsilon_i \lvert\boldsymbol{u}_i - \boldsymbol{v}_i\rvert}{\mu_f}$

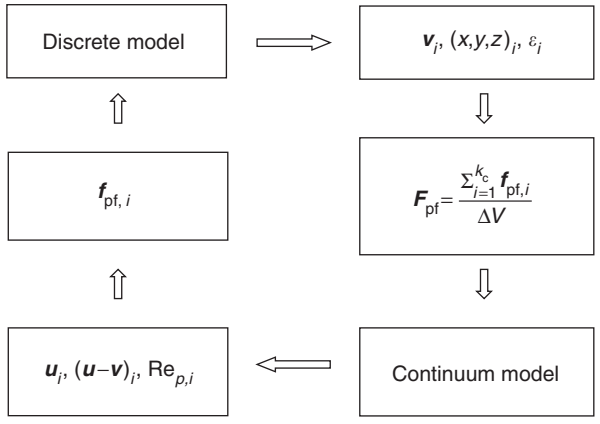


Figure 2 Coupling and information exchange between continuum and discrete models.

Model Validity and Application

Particle Packing

Random packing of particles, as a core research area in particle science and technology, is fundamental to many operations in powder processing and is useful in modeling some physical systems. Its formation is a dynamic process involving various interparticle forces as discussed above. Therefore, a successful simulation method must take into account all dynamic factors related to both geometry and force. Most of the previous methods only consider the former and ignore the latter, and hence fail to generate realistic packing structure, particularly when forces rather than the gravity are dominant. DEM-based simulation offers a promising way to overcome this difficulty.

Table 3 Parameters for glass beads

Parameter	*Symbol*	*Base value*
Particle density	ρ_{p}	$2.5 \times 10^3\,\mathrm{kg\,m^{-3}}$
Young's modulus	Y	$1.0 \times 10^7\,\mathrm{N\,m^{-2}}$
Poisson's ratio	$\tilde{\sigma}$	0.29
Sliding friction coefficient	γ_{s}	0.3
Rolling friction coefficient	γ_{r}	0.005
Normal damping coefficient	η_{i}	$2 \times 10^{-5}\,\mathrm{s^{-1}}$
Hamaker constant	H	$6.5 \times 10^{-20}\,\mathrm{N\,m}$
Minimum separation	z_0	$1 \times 10^{-20}\,\mathrm{m}$

Simulation conditions can vary to represent different packing processes, for example, the spherical growth of a packing under a centripetal force such as that used in isotropic compression or the vertical growth under gravity and other forces. Treatment such as vibration or compaction can also be implemented in a simulation. Here, only packing with vertical growth, which is typically conducted as follows, is considered. First, particles are randomly generated without overlap in a cylindrical or rectangular container. Then, the particles are allowed to settle down under gravity and other forces. This dynamic process proceeds until all particles reach stable positions with an essentially zero velocity as a result of the damping effect for energy dissipation. This packing process is equivalent to a physical operation to transform a gas-fluidized bed to a fixed bed by stopping the gas supply. Periodical boundary conditions can be applied to avoid the effect of sidewalls. **Table 3** lists the parameters normally used to simulate the packing of glass beads. Note that Young's modulus is smaller than the

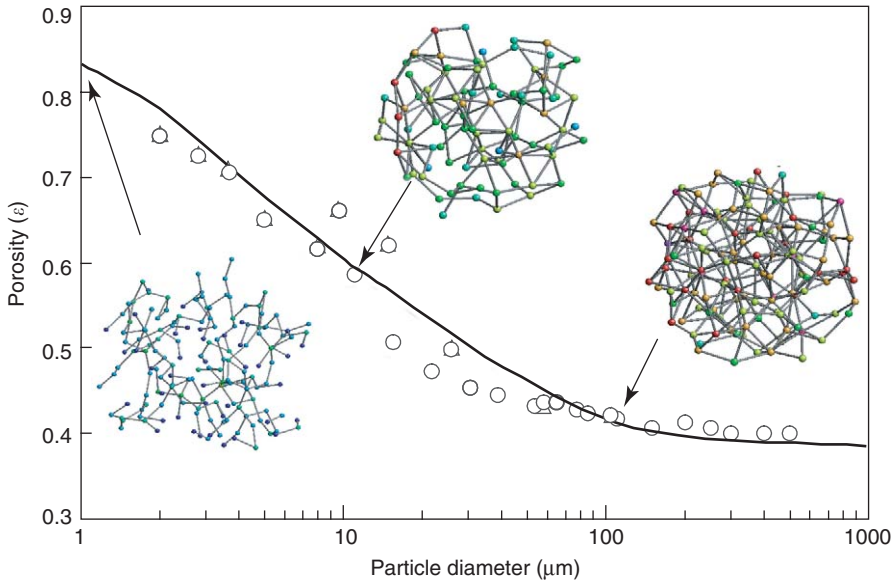


Figure 3 Porosity vs. particle size for glass beads: points, measured; line, simulated; and insets, packing structures taken from spherical samples for different sized spheres. (Reprinted with permission from Yang RY, Zou RP, and Yu AB (2000) Computer simulation of the packing of fine particles. *Physical Review E* 62: 3900–3908; © American Physical Society.)

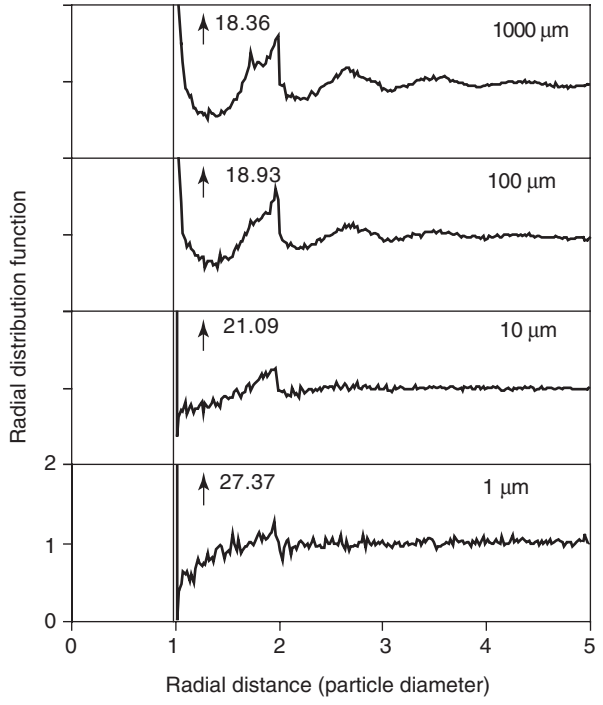


Figure 4 The radial distribution function for packings of different sized particles. (Reprinted with permission from Yang RY, Zou RP, and Yu AB (2000) Computer simulation of the packing of fine particles. *Physical Review E* 62: 3900–3908; © American Physical Society.)

actual value in order to reduce the stiffness of particles and hence the computational effort.

The validity of the simulation technique can be assessed at both macroscopic and microscopic levels. **Figure 3** shows the relationship between porosity and particle size for mono-sized spheres. The good agreement between the measured and simulated results confirms macroscopically the validity of the DEM and the important role of the van der Waals force in governing the behavior of fine spheres ranging from 1 to 100 μm. The split-second peak in the radial distribution function has been recognized as a key feature for coarse spheres with negligible effect of the van der Waals force. For years, many investigators have struggled to reproduce this feature by computer simulation. As shown in **Figure 4**, it can be readily generated with the present simulation technique. The results also show that the radial distribution function varies with particle size or porosity. The coordination number is a more sensitive parameter in characterizing the packing structure. **Figure 5** shows that the simulated coordination number distributions for a particle mixture are quantitatively comparable to those measured, confirming the validity of the DEM approach from the viewpoint of microstructure.

Further comparison can be made in terms of interparticle forces. There is a good agreement between the simulated and measured distributions of normal contact forces between particles for a packing formed under gravity (**Figure 6**). This demonstrates the capability of the DEM in generating quantitative force information which is difficult, if not impossible, to obtain with the current experimental technique.

The validated DEM model offers a convenient way to carry out controlled numerical experiments to

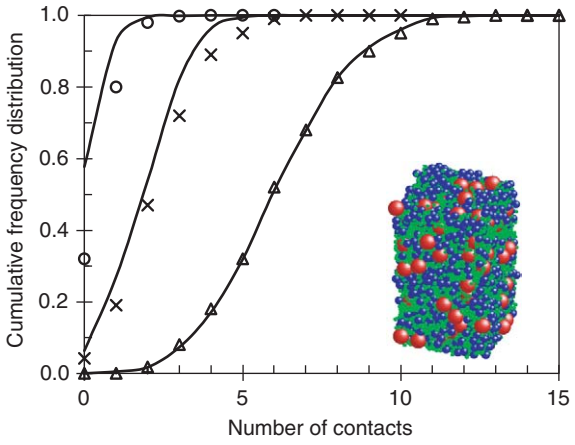


Figure 5 Distribution of coordination numbers for the packing of a ternary mixtures of spheres (inset) when the size ratios are 4:2:1 and the volume fractions are 0.25, 0.50, and 0.25 for large (red), medium (purple), and small (green) components: o, measured contacts between medium and large components; x, measured contacts between medium and medium components; Δ, measured contacts between medium and small components; lines, simulated.

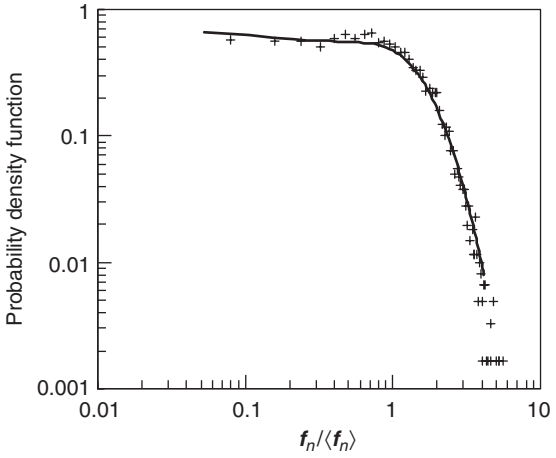


Figure 6 Frequency probability distribution of normalized normal contact forces between particles for the packing of 1 mm spheres: +, measured; −, simulated.

study the packing of particles under various conditions. For example, it has been used to quantify the effects of material properties such as the sliding and rolling friction coefficients related to the surface forces, the Hamaker constant, and the particle density related to the body forces acting on a particle, as shown in **Figure 7**. Correlations between macroscopic and microscopic or structural variables such as porosity, radial distribution function, and coordination number can be established based on DEM simulations. An important finding from this type of study is that macroscopic properties, for example, porosity, can be described as a function of interparticle forces, that is, the force ratio of the average van der Waals force and gravity force on a particle, as shown in **Figure 8**. This relationship, as an equation of state, can be generalized in terms of cohesive force relative to effective gravity force, although further work is necessary to understand its dependence on variables related to surface forces.

Structural analysis is the key to elucidating the packing mechanisms and the underlying physics. Realistic simulation of particle packing is crucial to such an analysis. In fact, the resulting structural information can be directly used to study the structural properties of porous media, including, for example, permeability related to pore-to-pore connectivity and effective thermal conductivity related to particle-to-particle connectivity. DEM-based microstructural study will play a key role in quantifying the governing mechanisms and generate reliable predictive methods for engineering application.

Particle Flow in a Horizontal Rotating Drum

Particle flow, coupled with fluid (gas and/or liquid) flow, is a very common feature in almost all operations or processes in powder processing. Understanding the fundamentals governing the flow and formulating suitable governing equations and constitutive relationships are of paramount importance to the development of better strategies for process and/or quality control. Previous studies are largely at a macroscopic or global scale, the resulting information being helpful in developing a broad understanding and design of a process of particular interest. However, the lack of quantitative fundamental understanding makes it difficult to generate a general method for reliable scaleup, design, and control/optimization of processes of different types. The bulk behavior of a particle system depends on the collective interactions of individual particles, and hence research on a particle scale has been recognized as a promising approach to overcome this difficulty. The CCDM method described above plays a very important role in this connection. This is demonstrated by the following two examples: particle flow in a rotating drum and gas fluidization of cohesive powders.

Rotating drums are widely used in various industries for various purposes including, for example, granulation/agglomeration, mixing, drying, milling, and coating. A comprehensive understanding of the flow of particles in a drum under different conditions is essential for process control and optimization. It is also very much related to the physics of granular matter, for example, particle avalanche in relation to heap formation or transition between fluid-like and

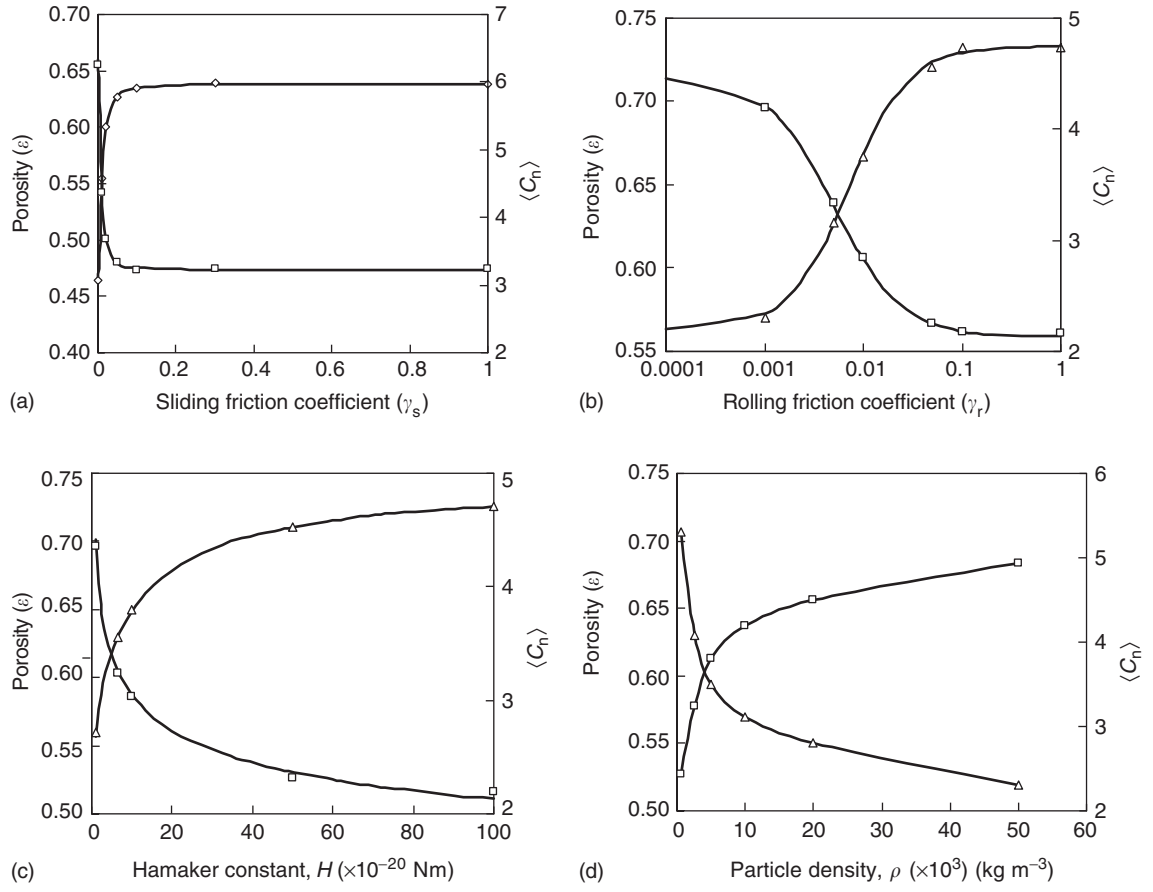


Figure 7 The effect of material properties on porosity (Δ) and mean coordination number (□): (a) sliding friction coefficient; (b) rolling friction coefficient; (c) Hamaker constant; and (d) particle density. (Modified with permission from Yang RY, Zou RP, and Yu AB (2003) Effect of material properties on the packing of particles. *Journal of Applied Physics* 94: 3025–3034; © American Institute of Physics.)

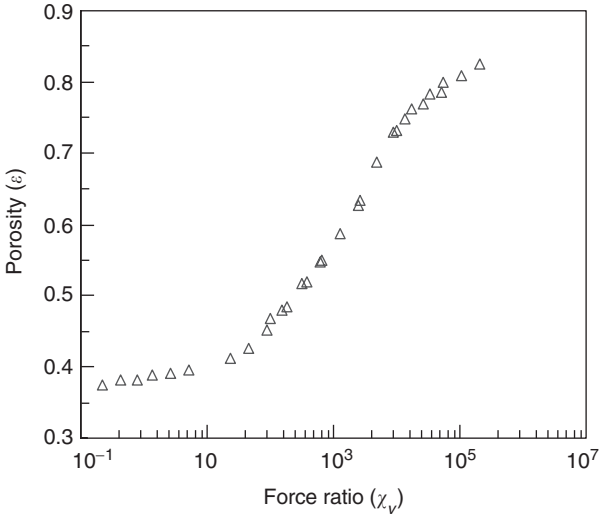


Figure 8 The relationship between porosity and force ratio for glass beads, where $\chi_v = \frac{1}{N}\left|\sum_j \boldsymbol{F}_{v,ij}/m_i g\right|$. (Modified with permission from Yang RY, Zou RP, and Yu AB (2003) Effect of material properties on the packing of particles. *Journal of Applied Physics* 94: 3025–3034; © American Institute of Physics.)

solid-like modes. This example illustrates how such a process can be simulated and analyzed.

Simulated here is the flow of glass beads in a horizontal drum of diameter 100 mm and thickness 16 mm. The drum is 35% filled with spheres of diameter 3 mm and rotates at various pre-set speeds. Periodic conditions apply axially to reduce the number of particles used; for simplicity the analysis is two-dimensional, achieved by averaging the data along the axial direction. Depending on the rotation speed, different flow regimes have been identified; but the cases considered here are in the so-called rolling regime that is commonly used in practice. The flow of particles in this regime has been well documented in the literature. Driven by the rotating drum, particles participate in collective solid rotation, reaching the filling surface inclined at an angle of repose from the rotating bulk at the bottom, pouring down this surface and getting mixed at the same time and then being re-trapped into the rotating bulk. This is indeed what is observed in the simulation, as shown in **Figure 9**.

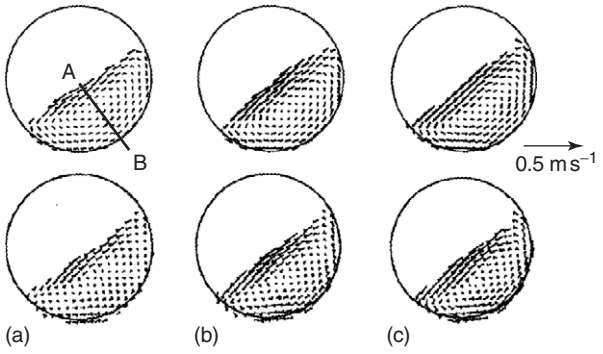


Figure 9 Velocity fields in a plane section taken from DEM simulation (top) and PEPT experiment (bottom) at rotation speeds of: (a) 20 rpm; (b) 42 rpm; and (c) 65 rpm. (Reproduced from Yang RY, Zou RP, and Yu AB (2003) Microdynamic analysis of the flow of particles in horizontal rotating drum. *Powder Technology* 130: 138–146, with permission from Elsevier.)

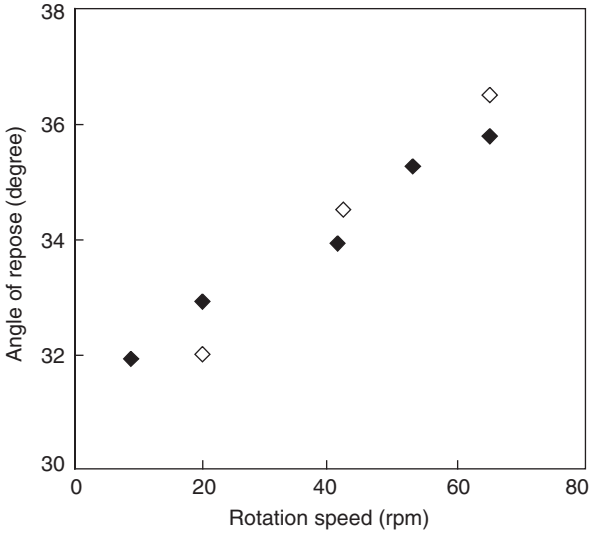


Figure 10 Dynamic angle of repose as a function of rotation speed: ◇, PEPT measurement; ◆, DEM simulation. (Reproduced from Yang RY, Zou RP, and Yu AB (2003) Microdynamic analysis of the flow of particles in horizontal rotating drum. *Powder Technology* 130: 138–146, with permission from Elsevier.)

The validity of the simulation technique can be assessed by directly comparing the simulated and measured results. It is very useful to do so at a particle scale. One of the experimental techniques to produce such data is the so-called positron emission particle tracking (PEPT). **Figure 10** shows the dependence of the dynamic angle of repose on rotation speed. The agreement between the DEM simulation and PEPT measurement is good, both indicating that the angle increases linearly with the speed of rotation for the range considered.

Figure 9 also shows that the flow patterns produced by the DEM are comparable to those generated

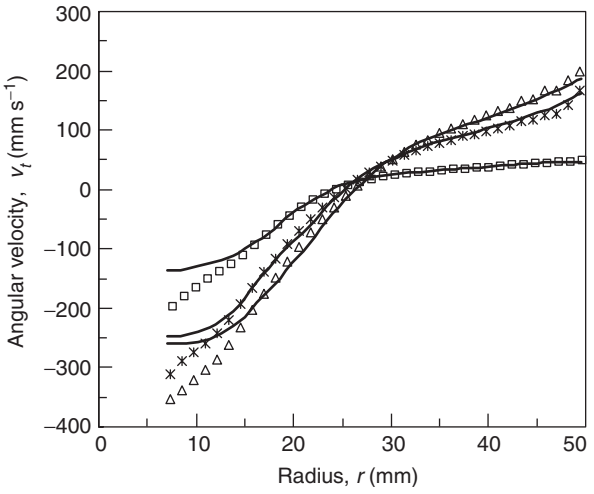


Figure 11 Angular velocity along the radius perpendicular to the flowing surface at rotation speeds of: □, 20 rpm; ✳, 42 rpm; and Δ, 65 rpm. Points are the PEPT measurements and solid lines are the simulated results. (Reproduced from Yang RY, Zou RP, and Yu AB (2003) Microdynamic analysis of the flow of particles in horizontal rotating drum. *Powder Technology* 130: 138–146, with permission from Elsevier.)

by PEPT for different rotation speeds. To be more quantitative, **Figure 11** plots the angular velocity along the radius perpendicular to the flowing surface, that is, line $\overline{AB}$ in **Figure 9**. Obviously, the agreement between the simulation and experiment is good. Satisfactory agreement between DEM and PEPT can also be observed for other processes such as bladed mixer and gas fluidization. It is noticed that the simulated velocity near the flow surface is higher than the PEPT velocity. This difference is mainly because the DEM simulation may not fully represent the PEPT experimental conditions, for example, the wall properties are simply assumed to be the same as those of particles in the DEM simulation. The selection of parameters for simulation is important to generate results matching physical experiment. The different sampling methods used in the simulation and experiment may also contribute to this observed discrepancy. Unlike PEPT, which determines the velocity field using the trajectory data of a single particle, the velocity field obtained from the DEM simulation is the average of the instantaneous velocities of all particles in the system.

The dynamics of particle flow includes at least three aspects: velocity, structure, and force. Previous studies had to be largely limited to velocity field because of the difficulty in obtaining information for the other two. Consequently, there have been problems in probing the underlying mechanics and solving practical problems reliably. DEM is an effective technique that can overcome this problem. Its results

can be readily used to establish the spatial and statistical distributions of important variables related to flow and force structures such as porosity, coordination number, and various forces acting on particles. These variables can be treated as local average variables. As an example, **Figure 12** shows some typical results when the rotation speed is 42 rpm. Nonuniformity is obvious. For example, the bottom part, that is, the rotating solid bed, has a lower porosity while the top part, that is, the flow layer, a higher porosity (**Figure 12a**). The normal contact force is important as other contact forces (e.g., tangential contact force) are directly related to it. **Figure 12b** shows the force network in terms of this force, where each stick represents one connection between two particles with its thickness representing the magnitude of the force. As expected, large forces occur to particles at the bottom wall region to provide support to the top particles. However, relatively large forces can also be found at the lower joint region between the flow layer and the rotating solid bed where particles of high velocity pouring down the bed surface impact the particles of low velocity waiting to start the bed rotation movement. Particles at the top have minimum contact forces. Notably, the force network actually also represents the flow structure of particles.

Understanding the interactions among particles is important for such a dynamic system. For example, the size enlargement or reduction process depends on two important factors: the effectiveness of a collision and the frequency of collisions between particles. The effectiveness is mainly related to the so-called initial kinetic energy or impact energy, which is proportional to the square of the relative collision velocity between two particles v_{ij} ($=|v_i - v_j|$). **Figure 12c** shows the spatial distribution of relative collision velocity. It demonstrates that the most vigorous interaction between particles is at the lower joint point between the flow layer and the rotation solid bed. It then propagates into the bed mainly along the shear band between the flow layer and rotation solid bed. As shown in **Figure 12d**, the spatial distribution of collision frequency has a similar pattern. A region of high relative collision velocity corresponds to a region of high-collision frequency, or vice versa. Note that the concept of collision differs from that of contact between particles. The contact or coordination number is determined by "freezing" the bed. In contrast, particle collision is a dynamic process and a collision must be from two noncontacting particles. **Figure 12d** is obtained when two particles collide from a gap of at least 1% of particle diameter.

The resulting information can be used to understand the performance of industrial operations. Here, particle size enlargement by agglomeration is taken as an example. Increasing rotation speed can increase the impact energy per collision and the collision frequency, giving two opposite effects on agglomeration

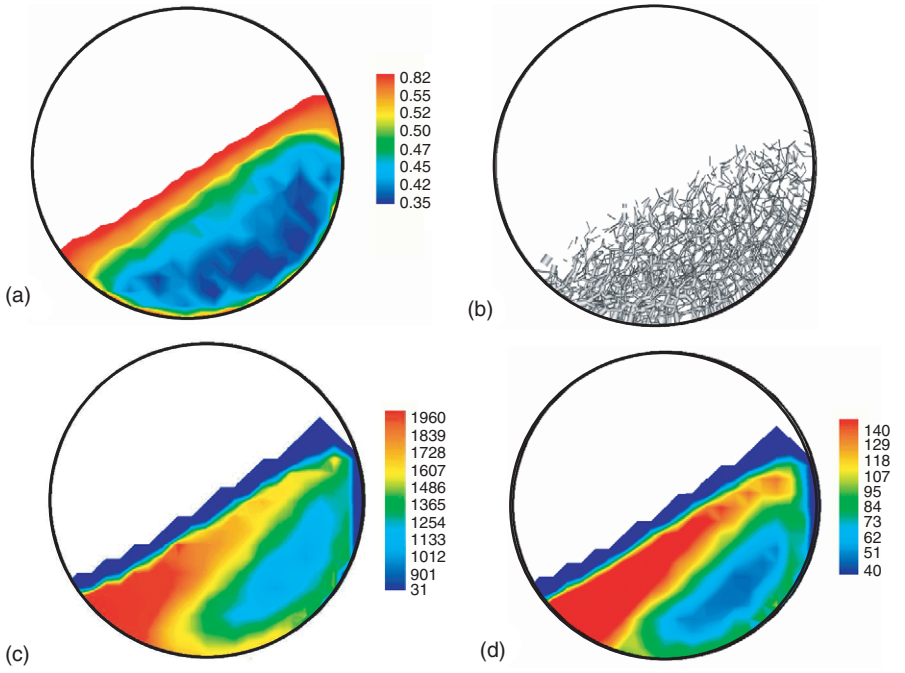


Figure 12 Spatial distribution of: (a) porosity; (b) (snapshot) force network, total force (mg); (c) collision velocity ($m\,s^{-1}$); and (d) collision frequency (s^{-1}) when the rotation speed is 42 rpm. (Reproduced from Yang RY, Zou RP, and Yu AB (2003) Microdynamic analysis of the flow of particles in horizontal rotating drum. *Powder Technology* 130: 138–146, with permission from Elsevier.)

in a drum. The increased kinetic energy will promote the breakage, crushing, and abrasion of agglomerates, decreasing the growth rate. On the other hand, the increased collision frequency means more contacts or coalescences among particles within a given time period, leading to an increased growth rate. Consequently, as schematically shown in **Figure 13**, depending on the rotation speed and operational conditions, various growth kinetics can result from the two opposite effects.

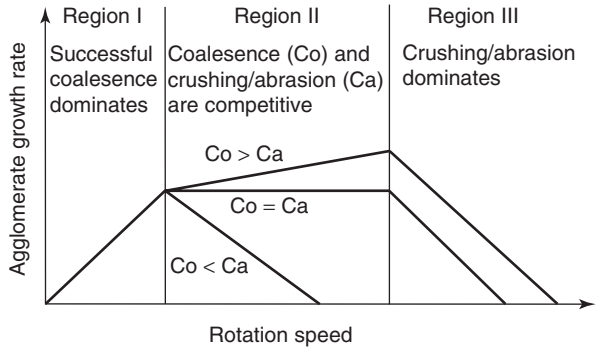


Figure 13 Schematic illustration of the effect of rotation speed on the growth kinetics of agglomerate size. (Reproduced from Yang RY, Zou RP, and Yu AB (2003) Microdynamic analysis of the flow of particles in horizontal rotating drum. *Powder Technology* 130, 138–146, with permission from Elsevier.)

Gas Fluidization of Cohesive Powders

It has long been recognized in practice that powders show four distinct behaviors, that is, cohesive, aeratable, sand-like, and spoutable, when fluidized. Theoretically, these behaviors must be related to particle–particle and particle–fluid interactions. Through controlled numerical experiments, the role of various interparticle forces can be examined. The simulations are carried out in a rectangular packed bed. Gas is then introduced uniformly at the bottom of the bed at a specified velocity to start a coupled gas–solid flow. **Table 4** lists the physical parameters used in this work. Different from the above two examples, it is assumed that

$$\gamma_r = \sqrt{R_i^2 - (R_i - 0.5\delta_{ij})^2}/R_i$$

Figure 14 summarizes the gas and solid flow patterns in the sequence of increasing gas velocity, with particles colored in a layered pattern according to their initial positions in the bed to demonstrate the particle rearrangement and solid mixing. Three states can be seen in this figure: fixed bed, expanded bed, and fluidized bed. In the fixed bed, there is no observable structural change in the bed. Gas simply flows through interstices among stationary particles.

Table 4 Parameters used for the present simulation

Solid phase		*Gas phase*	
Particle shape	Spherical	Type of gas	Air
Number of particles, N	45 000	Viscosity, μ_f	$1.8 \times 10^{-5}\,\mathrm{kg\,m^{-1}\,s^{-1}}$
Particle diameter, d	$1.0 \times 10^{-4}\,\mathrm{m}$	Density, ρ_f	$1.205\,\mathrm{kg\,m^{-3}}$
Particle density, ρ_p	$1440\,\mathrm{kg\,m^{-3}}$	Bed width	$2.5 \times 10^{-2}\,\mathrm{m}$
Spring constant, κ	$50\,\mathrm{N\,m^{-1}}$	Bed height	$4.5 \times 10^{-2}\,\mathrm{m}$
Sliding friction coefficient, γ_s	0.3	Bed thickness	$1.0 \times 10^{-4}\,\mathrm{m}$
Contact damping coefficient, η	$1.65 \times 10^{-5}\,\mathrm{kg\,s^{-1}}$	Cell width	$1.0 \times 10^{-3}\,\mathrm{m}$
Hamaker constant, H	$2.10 \times 10^{-20}\,\mathrm{N\,m}$	Cell height	$1.0 \times 10^{-3}\,\mathrm{m}$

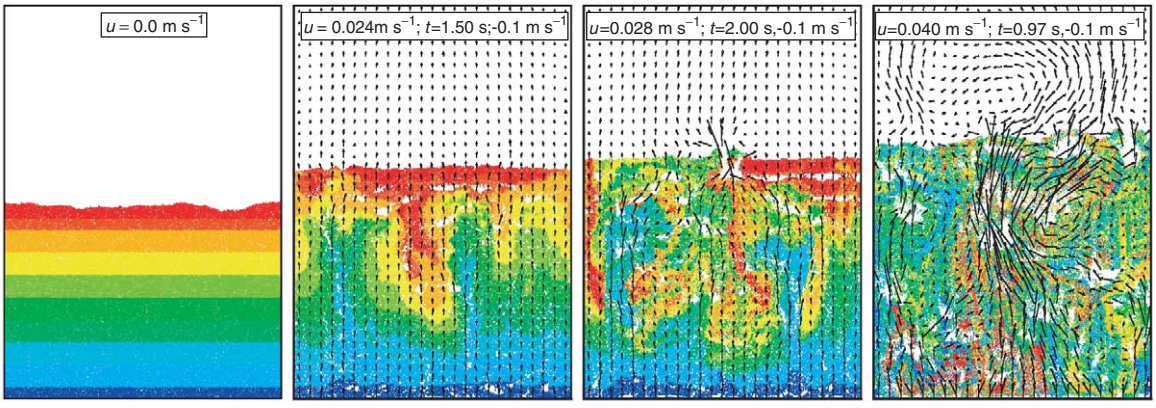


Figure 14 Particle configurations and gas flow fields for different superficial gas velocities, showing the existence of fixed ($u = 0.0\,\mathrm{m\,s^{-1}}$), expanded ($u = 0.024\,\mathrm{m\,s^{-1}}$), and fluidized beds ($u \geqslant 0.028\,\mathrm{m\,s^{-1}}$).

In the expanded bed, cavities or large pores of different size and orientation are formed. Generally speaking, the cavities show an elongated shape and are predominantly horizontal. The expanded bed terminates at the so-called minimum bubbling velocity, which is $\sim 0.028\,\mathrm{m\,s^{-1}}$ from the simulation. In the fluidized bed, sustainable bubbles appear in the bed. The predicted trends are qualitatively in good agreement with those observed.

The analysis of the forces acting on individual particles is important to understand the phenomena observed. This can be demonstrated in the analysis of the origin of bed expansion. **Figure 15** shows the variation of the overall averaged contact force, van der Waals force, and fluid drag force relative to particle gravity during the transition from a fixed bed to an expanded bed when $u = 0.024\,\mathrm{m\,s^{-1}}$. Three stages, that is, startup, transient, and stable, can be

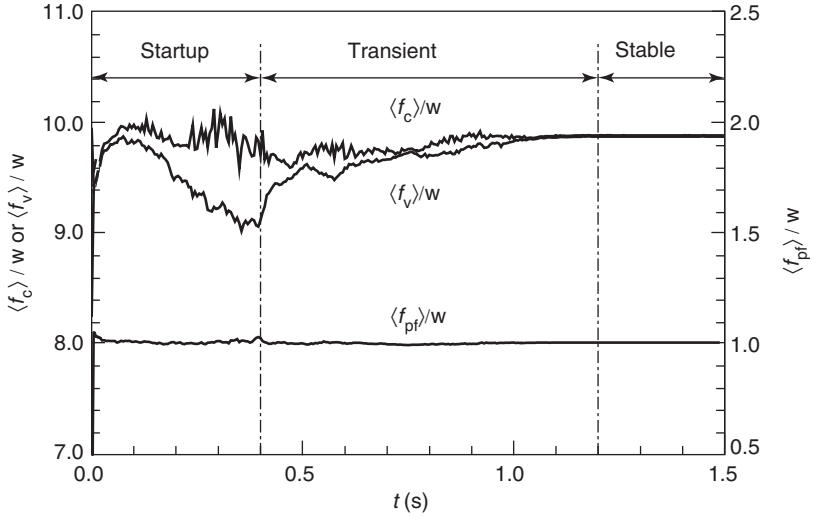
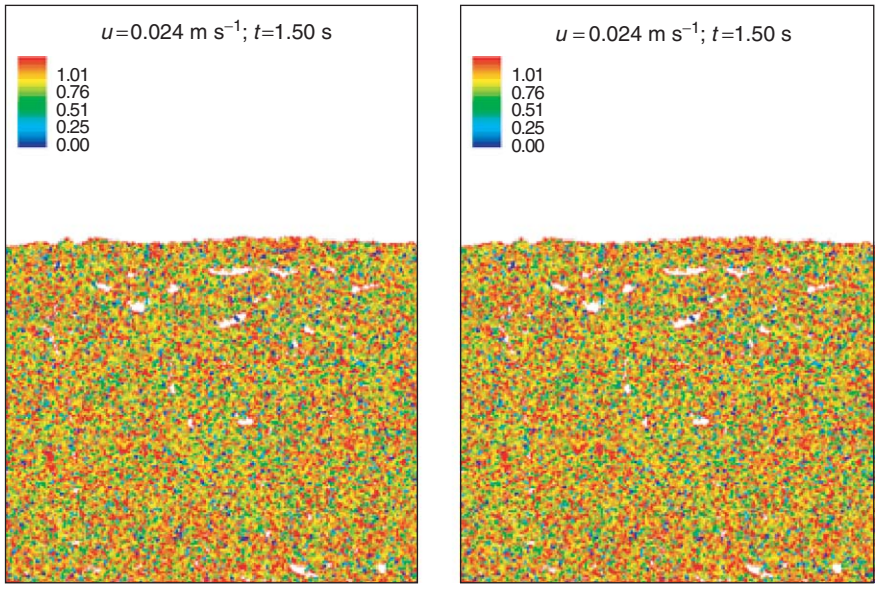


Figure 15 Variation of dimensionless ensemble-averaged contact force, van der Waals force, and fluid drag force with time when $u = 0.024\,\mathrm{m\,s^{-1}}$.



(a) (b)

Figure 16 Spatial distribution of the magnitudes of: (a) contact force; (b) van der Waals force acting on individual particles in an expanded bed when $u = 0.024\,\mathrm{m\,s^{-1}}$. The forces relative to the gravity of a particle are shown in a logarithm scale.

identified in this figure. In the startup stage, both the contact and van der Waals forces increase sharply within a short time. The two forces reach their maximum and then the contact force fluctuates at this level while the van der Waals force decreases. In the transient stage, the fluctuation in contact force is attenuated and the van der Waals force starts increasing. In the stable stage, the contact force and van der Waals force balance each other. The fluid drag force is virtually maintained at a value just balancing the weight of the bed during all three stages, a fact that highlights the rapid response from the solid phase, by changing its porosity, to the fluid drag force induced by gas flow. Clearly, the contact force must be balanced by the van der Waals force to form a stable expanded bed. Further analysis of the spatial distributions of the two forces indicates that the balance between them is observed not only at a bulk average scale but also at an individual particle scale, as shown in **Figure 16**. If the van der Waals force is unable to balance the contact force after initial disruption of the fixed bed, instability and then fluidization result.

Figure 17 shows the sequences of gas–solid flow patterns with increasing magnitude of the van der

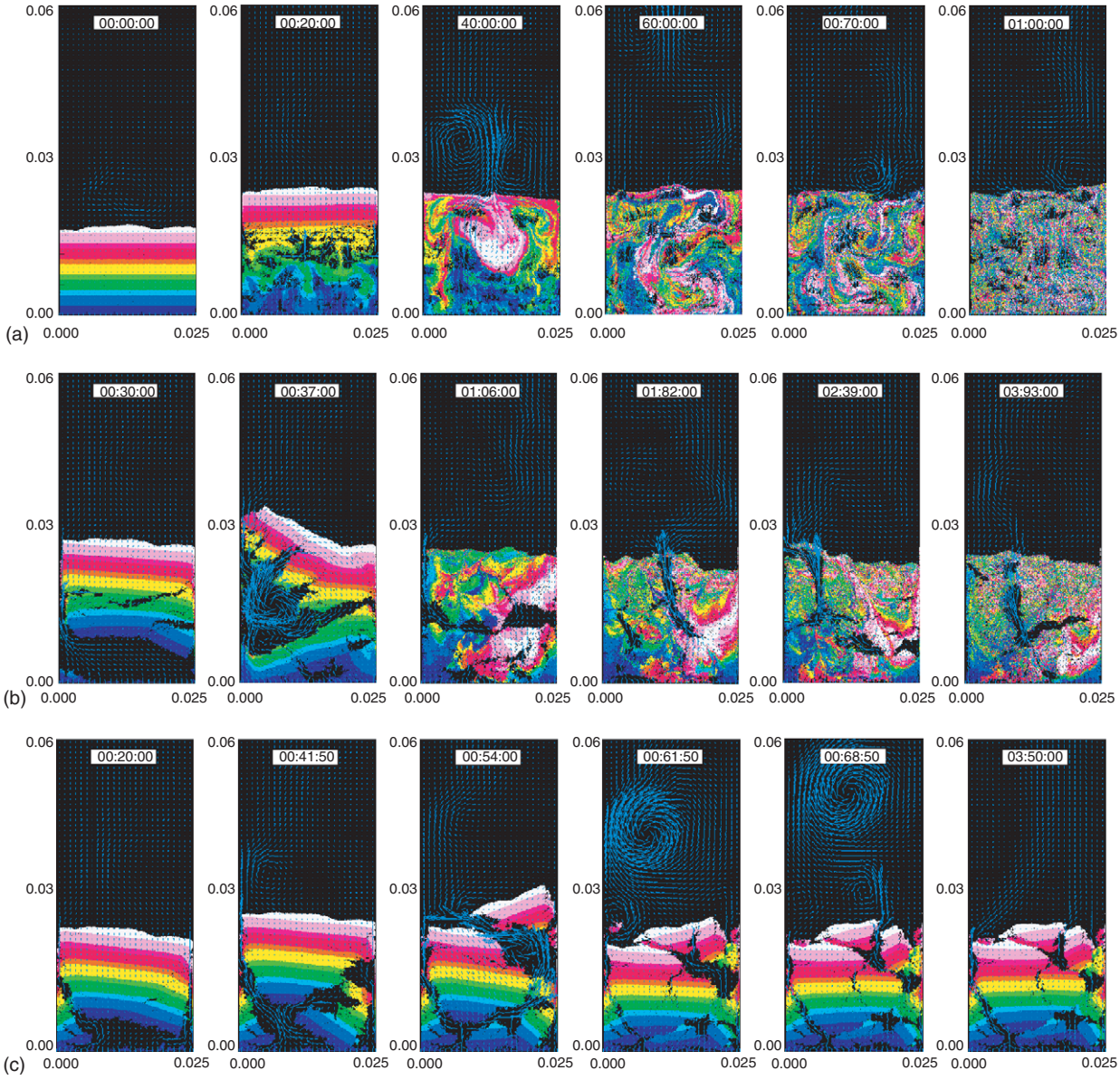


Figure 17 Gas–solid flow patterns when the ratio of superficial gas velocity to minimum fluidization velocity is 4.8 and the ratio of van der Waals force to particle weight is (a) 10, (b) 70, and (c) 300. The particles are colored according to their initial heights in the bed to visualize the solid mixing.

Waals force, where the change of the van der Waals force is achieved by adjusting the Hamaker constant. Three regimes of fluidization can be identified from this figure: good fluidization, quasi-fluidization, and defluidization. When the ratio of the maximum van der Waals force (determined when $z = z_0$) to particle gravity is less than ~ 30, good fluidization is obtained with continuous bubbles of relatively small sizes rising through the bed. When the force ratio is in the range of ~ 40–100, the fluidization quality deteriorates significantly and severe channeling occurs although the bed is still fluidizable at higher gas velocities. When the force ratio is greater than ~ 300, the fluidization is impossible by normal means. The simulation reasonably reproduces the experimentally observed phenomena, leading to the quantification of the roles of various particle–particle and particle–fluid interaction forces for this gas–solid flow system.

Concluding Remarks

The bulk behavior of granular materials depends on the collective interactions among individual particles. Particle scale modeling and analysis is therefore a key to elucidating the underlying physics and linking fundamental to applied research, particularly in developing from know-how to know-why knowledge. The examples discussed in this article clearly demonstrate that discrete particle simulation, although not perfect at this stage of development, does capture the main features and is an effective way to achieve this goal. With the rapid development of the computer technology, it offers a cost-effective way to meet the challenge of understanding the ever-changing particle science and technology resulting from the continuing development of new processes and products.

Acknowledgments

The authors are grateful to the Australian Research Council for the financial support, Drs. RY Yang, RP Zou, and BH Xu for their contribution to the case studies reported.

See also: Ceramic Materials; Computer Simulation Techniques in Condensed Matter Physics.

PACS: 45.70.−n; 45.70.Cc; 45.70.Mg; 61.43.−j; 61.43.Bn; 61.43.Gt; 46.05.+b; 83.10.Pp; 83.10.Rs

Further Reading

Allen MP and Tildesley DJ (1987) *Computer Simulation of Liquids*. New York: Clarendon Press.

Bideau D and Hansen A (eds.) (1993) *Disorder and Granular Media*. North-Holland: Elsevier.

Crowe CT, Sommerfield M, and Tsuji Y (1998) *Multiphase Flows with Droplets and Particles*. Boca Raton: CRC Press.

Cundall PA and Strack ODL (1979) A discrete numerical model for granular assemblies. *Geotechnique* 29: 47–65.

Gidaspow D (1994) *Multiphase Flow and Fluidization: Continuum and Kinetic Theory Descriptions*. San Diego: Academic Press.

Herrmann HJ, Hovi JP, and Luding S (1997) *Physics of Dry Granular Media*. Dordrecht: Kluwer Academic.

Israelachvili JN (1991) *Intermolecular and Surface Forces*, 2nd edn. London: Academic Press.

Johnson KL (1985) *Contact Mechanics*. Cambridge: Cambridge University Press.

Mehta A (ed.) (1993) *Granular Matter: An Interdiscilinary Approach*. New York: Springer.

Oda M and Iwashita K (eds.) (1999) *Mechanics of Granular Materials*. Rotterdam, The Netherlands: A.A. Balkema Publishers.

Roco MC (ed.) (1993) *Particulate Two-Phase Flow*. Stoneham, MA: Butterworth-Hainemann.

Seville JPK, Tüzün U, and Clift R (1997) *Processing of Particulate Solids*. London: Blackie Academic & Professional.

Stewart R, Bridgwater J, Zhou YC, and Yu AB (2001) Simulated and measured flow of granules in a bladed mixer – a detailed comparison. *Chemical Engineering Science* 56: 5457–5471.

Xu BH and Yu AB (1997) Numerical simulation of the gas-solid flow in a fluidised bed by combining discrete particle method with computational fluid dynamics. *Chemical Engineering Science* 52: 2786–2809.

Yu AB and Xu BH (2003) Particle scale modelling of particle-fluid flow in fluidization. *Journal of Chemical Technology and Biotechnology*. 78: 111–121.

Zhou YC, Yu AB, Stewart R, and Bridgwater J (2004) Microdynamic analysis of particle flow in a bladed mixer. *Chemical Engineering Science* 59: 1343–1364.

Nomenclature

C_n	coordination number (dimensionless)
d	diameter of particles (m)
$\boldsymbol{f}$	force acting on a particle (N)
g	gravitational acceleration ($m\,s^{-2}$)
H	Hamaker constant (Nm)
I	moment of inertia of a particle (defined as $I = 2/5 m R^2$, $kg\,m^2$)
k_c	number of particles in a computational cell (dimensionless)
k_i	number of particles interacting with particle i (dimensionless)
m	mass of a particle (kg)
N	total number of particles in a considered system (dimensionless)
p	pressure (Pa)
$\boldsymbol{R}$	radius vector of a particle (m)
Re	Reynolds number (dimensionless)
t	time (s)
$\boldsymbol{T}$	torque acting on a particle (Nm)
$\boldsymbol{u}$	fluid velocity ($m\,s^{-1}$)
$\boldsymbol{v}$	translational velocity of a particle ($m\,s^{-1}$)
$\boldsymbol{w}$	weight of a particle (N)

W	weight of a particle bed (N)
z	separation distance between two interacting surfaces (m)
z_0	smallest separation distance between two interacting surfaces (m)
ΔV	volume of a computational cell (m^3)
δ	displacement between two contacting particles (m)
ε	porosity (dimensionless)
γ	friction coefficient (dimensionless)
ρ	density ($kg\,m^{-3}$)
τ	fluid viscous stress tensor ($kg\,m^{-1}\,s^{-2}$)
μ	viscosity ($kg\,m^{-1}\,s^{-1}$)
η	viscous contact damping coefficient of a particle ($kg\,s^{-1}$)
κ	spring constant of a particle (Nm^{-1})
ω	rotational velocity of a particle (s^{-1})
ζ	arbitrary variable used in averaging operators

Subscripts

c	contact
d	damping
f	fluid phase
i	particle *i*
ij	between particles *i* and *j*
n	normal component
p	particle phase
pf	particle–fluid
pp	particle–particle
r	rolling friction
s	sliding friction
t	tangential component
v	van der Waals

Protein Folding and Aggregation

R A Broglia and G Tiana, Università degli Studi di Milano, Milan, Italy

Introduction

The protein folding problem concerns understanding how a protein of specified amino acid sequence ends up in a unique configuration which, among other things, determines its biological function. In physical terms: how does the one-dimensional information provided by the sequence of 20 different types of amino acids encodes for a unique and stable three-dimensional equilibrium conformation.

This problem has a clear biological and medical importance. The sequencing of the human genome provides information on the sequence of amino acids forming each of the hundreds of thousands of proteins which build our cells. The acquisition of sequence data by DNA sequencing is relatively quick, and vast quantities of data have become available through international efforts. On the other hand, the acquisition of three-dimensional data is still slow and is limited to proteins that either crystallize in a suitable form or are sufficiently small and soluble so that their structure can be determined by NMR in solution. Algorithms are thus required to translate the linear information into spatial information.

But the protein folding problem is quite intriguing also from a physical point of view. A protein is a system which is in a nearly-zero-entropy equilibrium state (usually referred to as "native" state; cf., e.g., **Figure 1**) in a wide range of temperatures (typically from 0°C to 60°C). Such equilibrium state has essentially no symmetries. The interactions within the protein are noticeably complicated and heterogeneous. Nonetheless, the protein does not display, as a rule, any competing low-energy states or kinetic

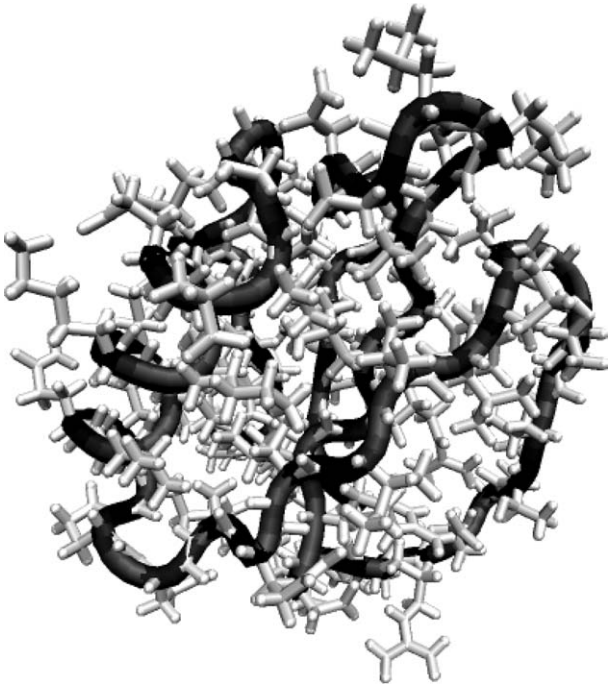


Figure 1 The native conformation of chymotrypsin inhibitor 2. The light gray parts highlight the arrangement of the atoms, which build a complicated network of interactions. In dark gray is highlighted the backbone of the protein.

traps (metastable states) typical of "frustrated" systems, but folds on a short call. The only feature typical of frustrated systems that survives in the case of proteins is the difficulty of predicting the ground state conformation from the knowledge of its elements (amino acids) and their interaction.

The goal of the physical approach to the protein folding problem is to understand the general principles of the folding mechanism of any protein. This implies that such a mechanism displays some amount of universality. There is indeed evidence that supports this scenario. In fact, aside from displaying a unique state of equilibrium (unique at the length scale of amino acids, i.e., 10^{-10} m) at biological temperatures, most proteins undergo a highly cooperative denaturation when the temperature is increased, similar to first-order phase transition (see **Figure 2**). The average folding time of proteins usually ranges from microseconds to seconds, and the distribution is usually a single- or multi-exponential, typical of a (or of few) Poissonian process. As a rule, proteins are very tolerant to point mutations. Mutations in a large number of sites have little or no effect on the folding properties of the protein, while each protein displays some key sites which, if mutated, lead to a large destabilization of the native state.

Typical cells contain around 20% proteins by weight. Consequently, evolution has provided proteins not only with the ability to fold on their own, but also to avoid aggregation, or to recognize other amino acid chains to perform their biological task. As in the case of oligomers, ordered assemblies of few single chains, which display the same features of unicity, stability, and fast folding of monomeric proteins.

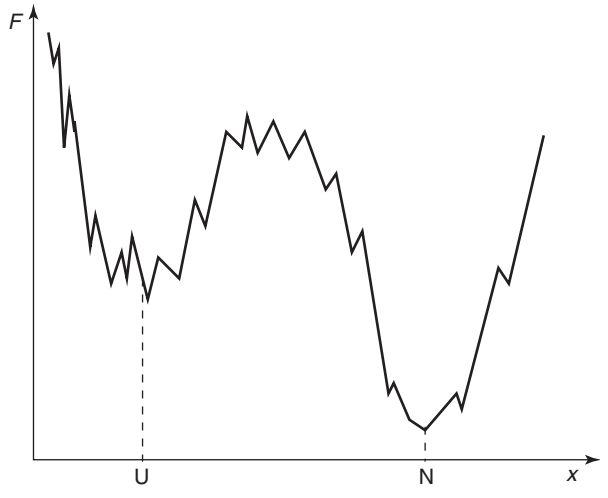


Figure 2 The typical free-energy landscape of a protein, plotted with respect to an arbitrary reaction coordinate *x*. Although the complexity of the interactions leaves some roughness in the landscape, evolution has shaped it in order to display two major wells separated by a barrier. The minimum associated with the native state (N) has essentially zero entropy and is stabilized by its attractive interactions, while the unfolded state (U) is mainly stabilized by its large entropy. As a consequence, when the temperature is increased, the free energy of the unfolded state decreases stabilizing it.

Models for Folding

The simplest (and oldest) interpretation of the folding process consists of summarizing all possible conformations in a few macroscopic states, in the same way as is done to describe chemical reactions. Usually, the thermodynamically relevant states are the unfolded states (U), which contain all the conformations where the protein chain is unstructured; the native state (N), which corresponds by definition to a single conformation; and possibly some intermediate states ($I_1, I_2, \ldots$). The folding reaction is thus viewed as a chain of events $U \rightarrow I_1 \rightarrow I_2 \rightarrow \cdots \rightarrow N$, and the dynamics of the associated probabilities is described by a mean of master equations. The distribution of folding times results in a sum of a number of exponential terms of the type $e^{-t/\tau}$ equal to the number of jumps between states that the system has to do to reach the native state. The typical time τ needed to perform each of such jumps is determined, according to Kramer's theory, by the height of the free-energy barrier ΔF, which separates the two states, and has the form $\tau \propto \exp[\Delta F/kT]$. In the case of small globular proteins, experiments often detect a single timescale for folding, indicating a two-state process between states U and N.

This simple model highlights the cooperativity of the folding process, determined by the presence of a major free-energy barrier, and provides good numerical estimates for the dependence of folding times on temperature and on mutations in protein sequence, but provides little insight into the molecular aspects of folding.

More recently, the focus has moved towards understanding the protein folding mechanism through a detailed description of the associated free energy landscape. A model which pursues this goal (the Go model) describes realistically the geometry of the protein but makes use of a simplified two-body potential function. Starting from the knowledge of the native conformation of the protein, two atoms are defined to be in (native) contact if their distance is smaller than a threshold. To such a conformation is ascribed a (negative) energy, and zero otherwise. This model, used in connection with computer simulations, is aimed at approximating the free energy with a careful entropic part and a simplified energetic part. Studying the folding of a number of small globular proteins with the help of the Go model, it is observed that the

conformations associated with the top of the free-energy barrier separating the unfolded from the native conformation have well-defined structural properties, characterized by some important native contacts.

A different approach focuses attention on the energetic content of the free-energy function and, in particular, on the heterogeneity of the interaction arising from the presence of 20 kinds of different amino acids. It is known that physical systems displaying such a heterogeneity are associated, as a rule, with a rough energy landscape with many competing low-energy states. This picture is incompatible with that of proteins, which must display a unique ground state, well separated from the others, and as few metastable states as possible. The purpose of these models is to understand what makes a protein, characterized by a well-defined amino acid sequence, different from a generic heterogeneous system, whose paradigm is a random sequence of amino acids.

The starting point of energy-based models is the study of the thermodynamics of a random heteropolymer, which is an idealization of the peptide chain only characterized by stochastically independent contact energies. Being the sum of independent stochastic contributions, the total energy of the heteropolymer follows a Gaussian distribution, according to the Central Limit Theorem. Identifying this distribution with the distribution of conformational energies of the heteropolymer, the number of conformational states as a function of energy is

$$n(E) \propto \exp\left[-N_c\left(\frac{(E - N_c\varepsilon_0)^2}{2N_c^2\sigma_\varepsilon^2} - \ln\gamma\right)\right] \qquad [1]$$

where N_c is the number of contacts in the chain (assumed to be constant), ε_0 and σ_ε are the mean and the standard deviation over the 20 types of amino acids of the contact energies, and γ is the number of conformations that the chain can build per contact. Similar to the random energy model used in the field of spin glasses, the random heteropolymer model displays an entropy given by the relation $S(E) = N_c \ln\gamma - (E - N_c\varepsilon_0)^2/(2N_c\sigma^2)$ at energies above the threshold energy $E_c \equiv N_c\varepsilon_0 - N_c\sigma(2\ln\gamma)^{1/2}$. On the other hand, below E_c the probability of finding any state of the heteropolymer is negligible, decreasing exponentially with the number of contacts N_c. Using the replica solution originally set up for spin glasses, it is possible to show that low-energy states are clustered in such a way that different clusters are separated by an energy barrier. This is a picture which can be hardly conciliated with that of proteins displaying a unique equilibrium state.

This simple model emphasizes a property of peptidic chains, that is the "frustration" due to the complicated interaction between 20 amino acids with different physical properties. "Frustration" implies the inability the system has to optimize all interactions at the same time, even in the lowest-energy state. A sequence of amino acids chosen at random is likely to be highly frustrated, and consequently not to display protein-like features. Within such a scenario, it is assumed that evolution selects protein sequences by minimizing the associated frustration. Setting this scenario within the framework of statistical mechanics, one can assume that protein-like sequences are those which display in some conformation an energy much lower than the sequence-independent threshold E_c, this conformation being the native state. In fact, if the native state is well below E_c, it does not have to compete with the other conformations whose energies are controlled by eqn [1]. Consequently, the native state will be unique and stable, more so, the larger the energy gap between the native state and E_c.

From the study of the models described above it emerges that the key ingredients of proteins to fold are their polymeric character and the heterogeneity of interactions. A powerful tool which allows to retain these two ingredients, although relaxing that of providing a detailed description of the protein geometry, is the lattice model. This model is based on two approximations: (1) the internal atomic structure of the amino acids is neglected and each of them is described as pointlike; (2) the amino acids move on the vertices of a cubic lattice of unitary side length. Accordingly, the conformational degrees of freedom are discrete. This is very convenient from a computational point of view and makes conformational entropy easy to handle, provided a contact potential is given. The simplest choice of such a potential has the form

$$U(\{r_i\},\{\sigma(i)\}) = \sum_{ij} \varepsilon_{\sigma(i)\sigma(j)}\Delta(r_i - r_j) \qquad [2]$$

where r_i and $\sigma(i)$ are the position and type of the ith amino acid, $\Delta(r_i - r_j)$ being a contact function assuming the value 1 if the ith and the jth amino acids are in contact and zero otherwise, while $\varepsilon_{\sigma\tau}$ is the element of a 20×20 interaction matrix which measures the interaction energy between amino acids of type σ and τ. A widely used interaction matrix has been determined by Miyazawa and Jernigan from the statistical analysis of the contacts of a large database of proteins, assuming that the frequency with which a given contact appears in the database measures the strength of the contact energy between the corresponding amino acids.

Within the framework of the lattice model, it is possible to obtain a data set of sequences which fold

into a selected native conformation starting from a random sequence displaying the observed composition (experimental ratio of the different types of amino acids), by swapping amino acids in such a way as to minimize the energy of the protein in the native conformation. This is a consequence of the fact that E_c does not depend on the details of the sequence, but only on its length. In other words, swapping amino acids does not change E_c, while the native energy can be decreased until it displays a large gap with respect to E_c.

A key result which emerges from such studies is that the stabilization energy of the protein is not distributed evenly throughout the native conformation, but is concentrated in a few, "hot" sites. This fact has remarkable consequences on the properties of the protein. First, it sheds light on the experimental fact that a protein is essentially insensitive to mutations in most of its sites. In fact, if a site different from the few "hot" sites is mutated, the change in native energy is not enough to fill the gap to the threshold energy E_c (which is not affected by the mutation), and consequently the protein retains its protein-like features. On the other hand, if a "hot" site is mutated, the energy of the resulting sequence in the native conformation can become larger than E_c. Consequently, the native state gets surrounded by a sea of unfolded conformation, losing its uniqueness.

Moreover, lattice model calculations indicate that "hot" sites build, as a rule, local elementary structures (LESs), that is fragments of the chain characterized by strong interactions between the "hot" sites which are close along the chain. Such LESs, which can be viewed as hidden, incipient secondary structures (like beta-sheets and alpha-helices) are very important for the folding dynamics because they bias the chain to the native state, according to a hierarchical scheme. Starting from a random conformation, LESs are formed very fast, because their stabilizing native contacts are close along the chain and, consequently, they do not need to explore a wide conformational space. Due to their large stabilization energy, LESs remain remarkably stable for the whole folding process. The rate-limiting step is then that of the assembly of LESs among themselves to build out their relative native contacts (the so-called "folding nucleus"). Being (quasi-) rigid entities, LESs make the chain effectively shorter. Furthermore, LESs interact among themselves with energies which are much larger than those with which single amino acids do. Consequently, it is unlikely that LESs form metastable structures, different from the native contacts designed by evolution. Once LESs have reached their mutual native positions, the conformational space of the protein is so restricted that the remaining amino acids fold immediately to their native position. Summing up, LESs provide an efficient way to squeeze out entropy from the initial random conformation in its way to the native state, avoiding metastable traps.

The uneven distribution of the stabilization energy and some hints about the events which take place on the folding trajectories of small real proteins can be studied by means of all-atom molecular dynamics simulations, combined with semi-empirical force fields. Results of such simulations, performed by considering specifically water, reveal the existence and allows the identification of "hot" sites. However, the simulation of the complete folding trajectory and the repetition of such simulations to obtain a statistically significant sample of the event is still computationally out of reach.

Aggregation

The term "aggregation" can mean a number of different processes, when applied to proteins. The simplest case is that of dimerization (or oligomerization), where two (or few) amino acid polypeptide chains assemble together into a conformation which displays all properties typical of the native state of monomeric proteins (e.g., uniqueness, stability, etc.). There are essentially three paradigms for this process. The best-understood is that of "exchange-domain" dimers, wherein two identical proteins have been induced by evolution to exchange part of their peptidic chain, giving rise to an interwined, symmetric conformation. The "lock-and-key" scenario is followed by those chains which first fold into monomeric native states and then assemble together, resulting at equilibrium in three populated states: denaturated, intermediate (single chain folded), and oligomeric (dock of single folded chains). On the other hand, "induced fit" oligomers first bind together in an unstructured conformation and subsequently find the native structure. While dimers (or oligomers) following the "exchange-domain" and the lock-and-key mechanism can still be described in terms of a few key sites (hot sites) controlling the whole process as in the case of single domain proteins, the "induced-fit" mechanism seems to be more subtle, implying a delicate balance between the contact energies at the interface and in the volume.

A mechanism similar to the lock-and-key mechanism can be used at larger scales, giving rise to micrometric, ordered structures. It is the case of tubulin, which builds microtubules in the cell, and actin, which builds muscle fibers. Aggregation can involve, under particular conditions, also proteins which usually fold to monomeric native conformations.

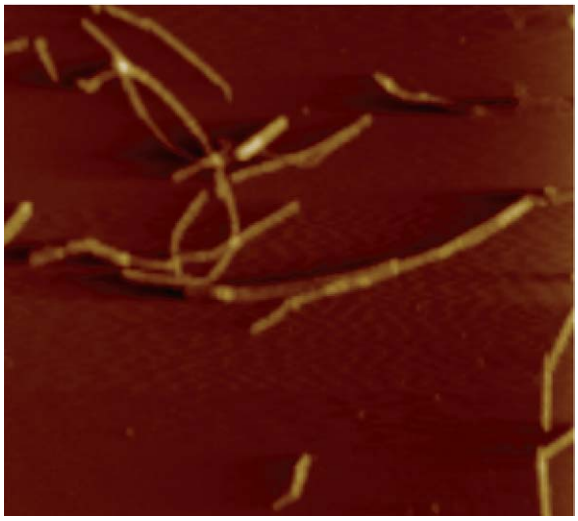


Figure 3 Fibrils made of insulin detected by atomic force microscopy. The length scale of the figure is μm.

A cell contains 17–26% of proteins by weight, and in the cytoplasm there are $\sim 10^6$ peptide chains, separated by an average distance of 3 nm. It is thus likely that evolution has selected sequences not only to fold on their own, but also to avoid the interference of other proteins. Nonetheless, mutations, changes in the chemical environment, or other factors can make proteins aggregate into (usually) biologically inactive, eventually harmful, clumps.

Aggregates can be disordered clumps, kept together by unspecific interactions, or an ordered, elongated structure usually called a fibril (see **Figure 3**). There is strong evidence that a number of diseases are related to the deposit of fibrils in particular areas of the human body (e.g., fibrils of beta-amyloid protein in the brain, in the case of Alzheimer's disease). Solid state NMR experiments have shown that, as a rule, fibrils contain parallel arrangements of specific U-like conformations of parts of the protein.

Aggregation is usually triggered by a nucleation event, and consequently the whole process can be separated into formation of the nucleus, that is, the minimum stable aggregate, and growth. Nucleation is a mechanism which reflects the entire thermodynamics of the solution, affected by the thermodynamic fluctuations of the solvent. For example, it has been shown that nucleation rate of sickle cell hemoglobin is strongly dependent on fluctuations associated with liquid–liquid demixing phase transition. Nucleation can be slowed down by stabilization of nonaggregating metastable states: light-scattering experiments and models based on molecular dynamics simulations indicate that micelle-like structures composed of a few beta-amyloid monomers compete with nuclei and discourage fibril formation. Also, the growth process often involves a nontrivial process which goes beyond mere diffusion. For example, the rate of elongation of beta-amyloid fibrils depends on temperature following an Arrhenius law, indicating that the binding of monomers to the end part of a fibril requires conformational changes.

See also: Folding and Sequence Analysis; Protein Folding, Engineering of.

PACS: 87.14.Ee; 64.60.Cn

Further Reading

Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181: 223.

Broglia RA, Shakhnovich EI, and Tiana G (2001) *Protein Folding, Evolution and Design*, Proceedings of the International School of Physics (E. Fermi). Amsterdam: IOS Press.

Broglia RA and Tiana G (2004) Lattice model of protein folding and of nonconventional drug design. *Journal of Physics: Condensed Matter* 16: R111.

Creighton TE (1992) *Protein Folding*. New York: W.H. Freeman.

Frauenfelder H and Wolynes PG (1994) Biomolecules: where the physics of complexity and simplicity meet. *Physics Today*, February 58.

Kusumoto Y, Lomakin A, Teplow DB, and Benedek G (1998) Temperature dependence of amyloid beta-protein fibrillazion. *Proceedings of National Academy of Sciences of the USA* 95: 12277.

Shakhnovich EI (1994) Proteins with selected sequences fold into unique native conformation. *Physical Review Letters* 72: 3907.

Vaiana SM, Palma-Vittorelli MB, and Palma MU (2003) Time scale of protein aggregation dictated by liquid–liquid demixing. *Proteins* 51: 147.

Protein Folding, Engineering of

S E Jackson, University of Cambridge, Cambridge, UK

Introduction

Proteins are the most diverse of biological macromolecules, both in the structures they adopt and the functions they perform. Proteins are essential in all biological processes, acting as catalysts in the form of enzymes, in structural roles such as collagen and keratin in skin and hair, in the transport of small molecules such as oxygen by the protein hemoglobin, in the immune response in the form of antibodies, and in the regulation of all cellular processes acting

as receptors and mediators of cellular signaling pathways.

Proteins are linear-chain polymers of amino acids, of which there are twenty naturally occurring building blocks. The condensation of the monomeric amino acid units into a polymer produces what is known as the polypeptide chain. Each protein has a specific sequence of amino acids which is encoded by the DNA of the gene for that protein. Proteins vary in length, the shortest are ~40–50 amino acids or residues long, whilst the longest, such as the giant muscle protein titin, can be over 10 000 residues in length. Many proteins, however, are much smaller than this, typically between 200 and 300 amino acid units. This article focuses on how, relatively small proteins, that is to say those that are 100 residues or less in length, fold.

Proteins are produced in cells on ribosomes, the cellular machines for the synthesis of proteins. When it is first made, the polypeptide chain is in a highly flexible and conformationally dynamic state. In this state, the protein is referred to as being "unfolded" or "denatured" and it is not active. The polypeptide chain has to collapse into a specific, compact, and tightly packed folded state, known as the native state, before it can perform its physiological function.

The protein-folding field emerged in the early 1960s with the seminal work of Christian Anfinsen who showed that all the information necessary to specify the final three-dimensional structure of a protein was encoded in its primary amino acid sequence. How, after synthesis, an unfolded polypeptide chain collapses into a specific, compact, and tightly packed folded state, is the protein-folding problem. Up until the late 1980s, most of the work in the field focused on proteins with disulfide bonds or *cis*-prolines, as in these systems stable intermediates were observed which could be characterized experimentally. In these cases, folding is limited not by folding *per se*, but by the formation or rearrangement of disulfide bonds or isomerization of the prolyl peptide bond. Whilst these processes are interesting in themselves, they provide little information on the conformational changes occurring in the protein that result in the formation of a stable native state. In the late 1980s, protein engineering techniques were first employed along with rapid-reaction kinetics to characterize the folding pathways of a number of small proteins. In contrast to earlier studies, in these studies, the rate-limiting steps probed corresponded to folding and not some slower rearrangement or isomerization event. This approach to studying protein folding has now been widely adopted by the experimental protein-folding community and has provided data critical for the development of new protein-folding mechanisms and in benchmarking numerous computational studies. Here, the main features of the protein engineering methods are outlined and the work in this field over the last ten years, which has led to the determination of the energy landscapes for folding at atomic resolution for several small proteins, is summarized.

Protein Engineering

With the advent of recombinant DNA technology in the 1970s and 1980s came the ability to rationally and specifically engineer proteins. Over the past twenty years, these techniques have been considerably further developed and extended to include the use of the polymerase chain reaction (PCR), which now plays a major role in the manipulation of DNA. These methods enable many proteins from a variety of sources (e.g., human, yeast, bacterial, and archaebacterial) to be produced in large quantities in bacteria, generally *Escherichia coli*. This greatly facilitates the process of obtaining sufficient quantities of pure proteins for biophysical analysis. Further, these technologies allow one to engineer proteins with ease by modifying the genes (DNA) that encode for a specific protein. A number of different methods of engineering proteins can be adopted: site-directed mutagenesis can be used to replace one amino acid in a protein with any other at a specific position in the protein thereby producing a mutant protein with a single mutation (the protein with the naturally occurring sequence is known as the "wild type"). Alternatively, multiple mutations can be made either rationally or randomly; the latter is a powerful technique when combined with a selection screen for proteins with a particular property. Proteins can also be changed more radically in a variety of ways, for example, by changing the order in which structural units or motifs are connected, resulting in a mutant protein which has the same overall structure as the wild-type protein but with different connectivities between structural subunits. These mutant proteins are known as circular permutants. These methodologies are now routine and used widely by the protein-folding community. A review of the entire field is beyond the scope of this article. Here, the focus is primarily on the results that have come from single-site-directed mutagenic studies, which have had the biggest impact on the field of protein folding. In particular, the focus is on the methods of analyzing protein-folding pathways using protein-engineering approaches and the models that have resulted from such work.

Characterizing Folding Pathways: ϕ-Value Analysis

In order to understand how proteins fold, the structure and energetics of all states on the folding pathway

need to be characterized. This includes the folded or native state, the unfolded or denatured state, as well as any intermediate states. Whereas, in the early years of protein folding it was thought that stable intermediate states were a prerequisite enabling the protein to fold on physiological timescales by reducing the amount of conformational space that needed to be sampled, in the early 1990s it was shown that a small protein, CI2, could fold very rapidly indeed without populating any such partially structured states. It is now well established that many small proteins fold in this so-called "two-state" fashion. In these cases, and others, one of the most important states to characterize on the folding pathway is the rate-limiting transition state which represents the energetic barrier to folding. Whilst the unfolded and folded states can be characterized directly using X-ray crystallography and multidimensional NMR techniques, the structure and energetics of the transition state can only be inferred from kinetic measurements. Fersht and co-workers have developed a method for analyzing folding transition states which is known as ϕ-value analysis and which uses protein-engineering approaches. The method is outlined below.

A point mutation is made to a protein and the kinetics of folding and unfolding are measured for both the wild-type protein and the mutant. In general, "nondisruptive deletions" are engineered such that the mutation removes a single interaction in the folded state of the protein and does not introduce any new interactions. In these cases, interpretation of ϕ-values is straightforward. The mutations usually destabilize the native state of the protein with respect to the denatured state (**Figure 1**). $\Delta\Delta G_{D-N}$ is the difference in the free energy of unfolding (the energy difference between D, the denatured state, and N, the native state) between the wild type and the mutant, and a positive sign indicates that the mutation is destabilizing. It is straightforward to measure this difference under equilibrium conditions using a suitable probe of the folded state, frequently intrinsic fluorescence of aromatic amino acids or far-UV circular dichroism. In order to quantitate the effect of a mutation on the transition state for folding, then $\Delta\Delta G_{\ddagger-D}$, the energy difference between the transition state ($\ddagger$) and the denatured state, between the wild type and the mutant, also needs to be measured (**Figure 1**). For simple systems, such as CI2 which fold in a two-state manner, this value can be calculated from the unfolding data (by measuring the rate constant of unfolding, k_U, which is determined by the energy barrier for unfolding), or from the folding data using the rate constant for folding, k_F. A ϕ-value is defined as the ratio of $\Delta\Delta G_{\ddagger-D}$ to $\Delta\Delta G_{D-N}$, which is a measure of the extent to which favorable interactions that the side chain of the mutated amino acid makes in the native state are formed in the transition state. Formal definitions of the different parameters are given below. Unfolding data can be used to calculate ϕ-values whether folding is two-state or not, as for small proteins no intermediate states are stable under unfolding conditions

$$\Delta\Delta G_{\ddagger-N} = -RT \ln(k_U/k'_U) \quad [1]$$

where k_U and k'_U are the rate constants of unfolding for the wild type and mutant, respectively. In this case, a ϕ-value is defined as follows:

$$\Phi = 1 - \Delta\Delta G_{\ddagger-N}/\Delta\Delta G_{N-D} \quad [2]$$

For those small proteins that fold with two-state kinetics, ϕ-values can also be calculated from folding rate constants

$$\Delta\Delta G_{\ddagger-D} = -RT \ln(k_F/k'_F) \quad [3]$$

where k_F and k'_F are the rate constants of folding for the wild type and mutant, respectively.

In this case, a ϕ-value is defined as follows:

$$\Phi = \Delta\Delta G_{\ddagger-D}/\Delta\Delta G_{N-D} \quad [4]$$

For two-state systems, ϕ-values calculated from the unfolding and folding data are the same.

Interpretation of ϕ-Values

Briefly, $\phi = 1$ when $\Delta\Delta G_{\ddagger-D} = \Delta\Delta G_{N-D}$, that is, the lost interaction energy upon mutation is the same in the native and transition states (**Figure 1**, case 1). In this case, the mutation affects only the rate of folding and not the rate of unfolding. ϕ-Values of 1 occur when the transition state is highly structured in the region of the mutation. Conversely, $\phi = 0$ when $\Delta\Delta G_{\ddagger-D} = 0$, that is, when the mutation has no effect upon the energy of the transition state relative to the unfolded state (**Figure 1**, case 2). In this case, the mutation affects the rate of unfolding but has no effect on the rate of folding. ϕ-Values of 0 occur when the transition state is completely unstructured in the region of the mutation. Besides the values of 0 and 1, fractional ϕ-values are obtained when $0 < \Delta\Delta G_{\ddagger-D} < \Delta\Delta G_{N-D}$ (**Figure 1**, case 3). In this case, the mutation has an effect on both unfolding and folding rate constants. In general, fractional ϕ-Values are more difficult to interpret; however, in some cases, there is an approximately linear relationship between the ϕ-Value and the extent of formation of nonpolar contacts in the transition state relative to the native state. In these cases, ϕ-values are a good measure of the degree to which interactions of the mutated side

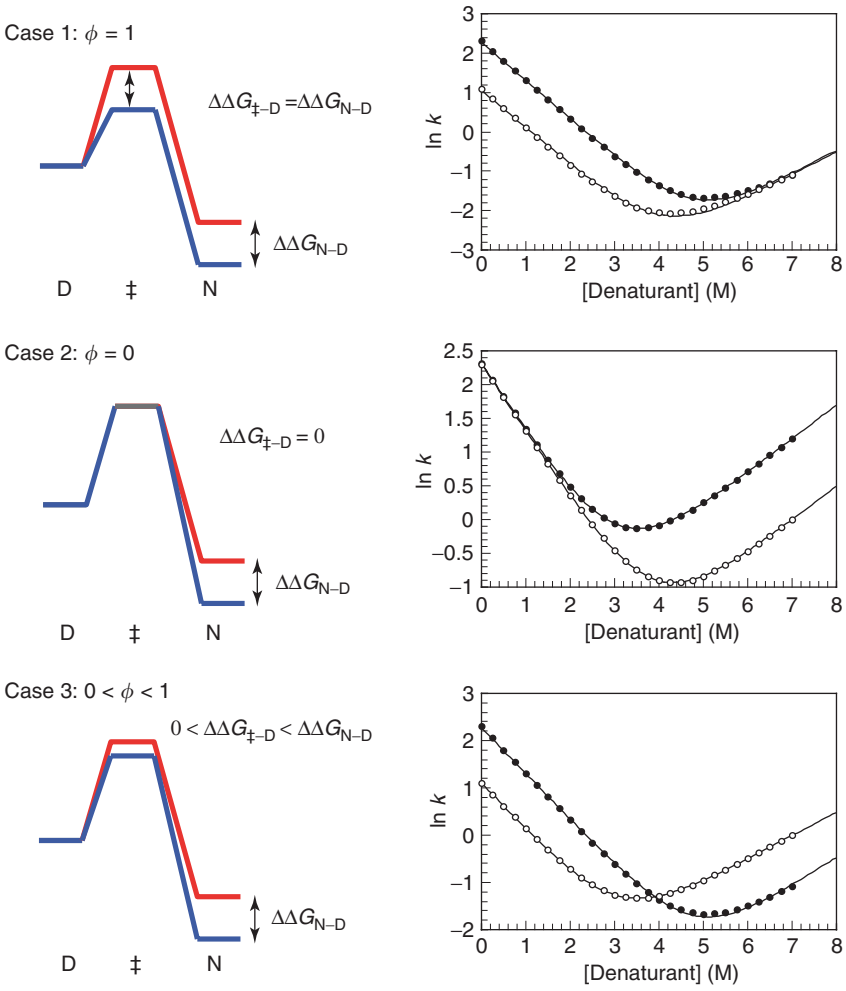


Figure 1 Free-energy diagrams showing the effect of a destabilizing mutation on the energy of the denatured (D), transition (‡), and native (N) state of a protein which folds with simple two-state kinetics. Case 1, when $\Delta\Delta G_{‡-D} = \Delta\Delta G_{N-D}$ then $\phi = 1$. In this case, the mutation affects the folding rate but not the unfolding rate – see the "chevron" plot of the rate constants for folding/unfolding as a function of denaturant concentration (right-hand side of figure). At low concentrations of the denaturant, the folding rate constant is measured and this decreases exponentially with increasing denaturant concentration, whilst at high concentrations of the denaturant, the unfolding rate constant is measured and this increases exponentially with increasing denaturant concentration. Case 2, $\Delta\Delta G_{‡-D} = 0$ then $\phi = 0$. In this case, the mutation affects the unfolding rate but not the folding rate. Case 3, $0 < \Delta\Delta G_{‡-D} < \Delta\Delta G_{N-D}$ then $0 < \phi < 1$. In this case, the mutation affects both unfolding and folding rate constants. (Reprinted with permission from Daggett V and Fersht AR (2003) The present view of the mechanism of protein folding. *Nature Reviews. Molecular Cell Biology* 4: 497–502; © Macmillan Magazines Ltd.)

chain are made in the transition state. An extensive description of the interpretation of ϕ-values and the assumptions and limitations of this method are not within the scope of this article and are discussed in detail elsewhere; see "Further reading."

In order to illustrate this method, the results of protein engineering and ϕ-value analysis on the small protein FKBP12 from the author's laboratory are discussed briefly below. FKBP12 is a small protein, the structure of which is shown in **Figure 2a**. Extensive protein engineering was performed on FKBP12 and the effect of mutations on the stability and kinetics of folding and unfolding of the protein was measured. Data were analyzed by the method described above and ϕ-values for many residues in the protein obtained. The results of these studies are summarized in **Figures 2b** and **2c**, which show the range of ϕ-values obtained and regions of the protein that have high values (are partially structured in the transition state) and those that have very low values (regions which are largely unstructured in the transition state). As expected, the ϕ-values vary between 0 and 1. In the case of FKBP12, no values of 1 were found, the highest value being 0.6. This is quite typical for small proteins. From **Figure 2**, it can be seen that residues with high values (> 0.5) cluster together and form a folding nucleus (see discussion on nucleation–condensation mechanism). In FKBP12, the

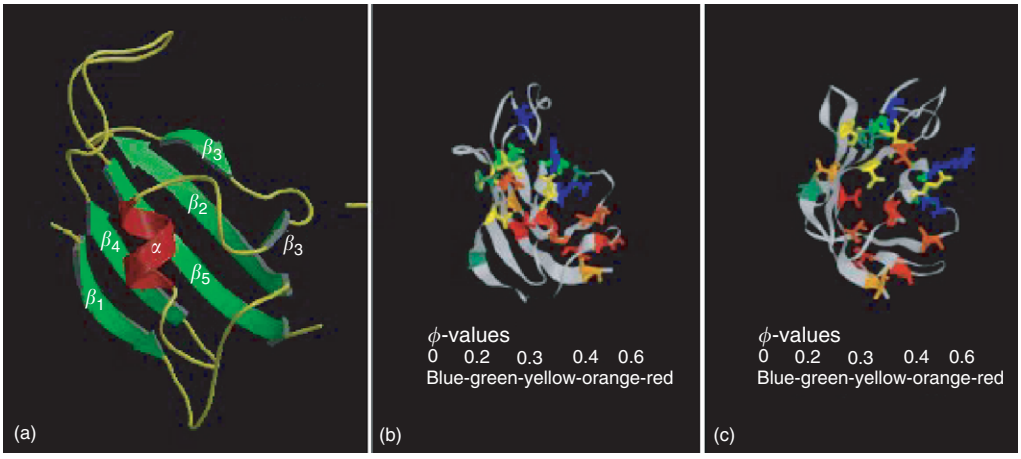


Figure 2 (a) Structure of the small, 106-residue protein FKBP12. Shown in red is the single α-helix, shown in green are the five β-strands which wrap around the helix to form the hydrophobic core. (b) Results of protein engineering studies and ϕ-value analysis. ϕ-values range from 0 (blue) to 0.6 (red). The side chains of the mutated residues are shown. (c) View of the ϕ-values and structure of FKBP12 rotated through 90°. (Reprinted with permission from Daggett V and Fersht AR (2003) The present view of the mechanism of protein folding. *Nature Reviews. Molecular Cell Biology* 4: 497–502; © Macmillan Magazines Ltd.)

nucleus is formed by residues at the C-terminus of the α-helix and residues in the middle of the central β-strands which form part of the hydrophobic core. There is a gradation of ϕ-values moving away from this nucleus, and edge strands and end of the β-strands and loop regions frequently have values close to zero indicating that these regions are highly flexible and unstructured in the transition state. FKBP12 behaves in many ways like other small proteins including CI2 (see next section).

The resolution with which folding pathways can be determined using this approach has meant that it is now the most popular way of characterizing folding transition states experimentally and ϕ-values have now been measured for a number of proteins. Results from these studies from numerous folding groups fall broadly into one of three categories illustrated by the three examples given below. Together, these studies give a detailed picture of the ways in which proteins fold.

CI2 Protein-Engineering Results: The Nucleation–Condensation Mechanism

As discussed above, the small protein CI2 folds in a simple two-state manner without populating a stable intermediate state. Thus, mechanisms which require a stable intermediate to be present cannot be applied to this protein. Protein-engineering experiments have mapped in detail the structure of, and interactions present in, the transition state. As a result of these studies, the nucleation–condensation mechanism of protein folding was proposed by Fersht and co-workers in the mid-1990s. In this mechanism, the rate-limiting step of folding is the formation of a folding nucleus which can either involve a large part of the protein structure (in this case, the protein is said to have a diffuse nucleus) or comprise just a few elements of secondary structure (in this case, the nucleus is said to be localized). After the nucleus has formed, folding involves the rapid condensation of the rest of the native structure around the nucleus in an energetically downhill process. For CI2, the nucleus is quite diffuse, involving, to some degree, most regions of the protein. However, several residues in the α-helix and in the β-sheet are the most critical in stabilizing the transition state structure. In contrast to earlier proposed mechanisms, in the nucleation–condensation model, secondary and tertiary structures are formed concomitantly, the secondary structure being inherently unstable without the presence of some stabilizing tertiary interactions. NMR spectroscopy and molecular dynamic simulations have both shown that the denatured state of CI2 has little fixed structure and approximates well to a random coil. The complete folding pathway of CI2 is shown in **Figure 3a**.

Barnase: Hierarchical Folding Mechanisms

Barnase was the first protein to be subject to an extensive protein engineering and ϕ-value analysis and its folding pathway illustrates some key differences with CI2. Barnase is a larger protein, 110 residues in length, which, in contrast to CI2, folds with three-state kinetics demonstrating the presence of a kinetically significant, populated, intermediate state on the folding pathway. A close analysis of the native structure of barnase shows that it has two structural

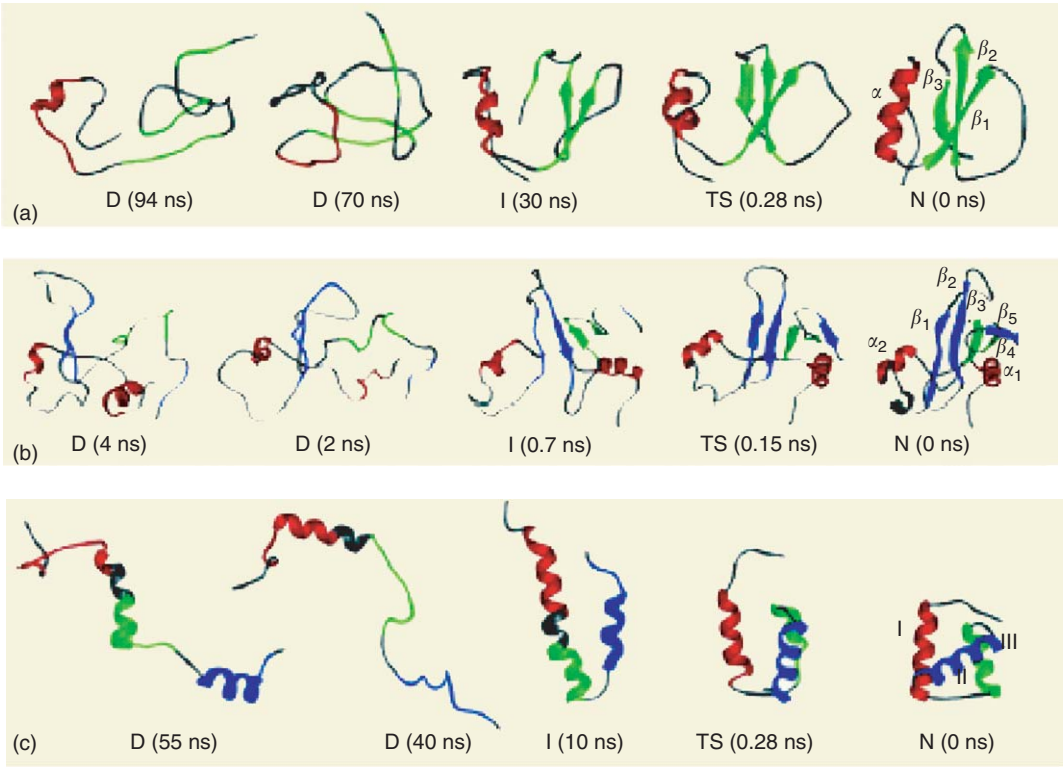


Figure 3 Schematic representations of the folding pathways of (a) CI2, (b) barnase, and (c) engrailed homeodomain. Results from protein-engineering studies and molecular dynamic simulations. (Reprinted with permission from Daggett V and Fersht AR (2003) The present view of the mechanism of protein folding. *Nature Reviews. Molecular Cell Biology* 4: 497–502; © Macmillan Magazines Ltd.)

subdomains. In contrast with CI2, barnase has a significant residual structure in the denatured state as probed by molecular dynamics simulations and NMR spectroscopy, that is, the denatured state is not a random coil and elements of both secondary and tertiary structure are transiently formed in this state. In barnase, the residual structure establishes a loose native-like topology in the denatured state which facilitates folding. From this state, the main hydrophobic core of the protein folds via a nucleation–condensation mechanism similar to CI2. The folding is nucleated by a residual fluctuating native-like helical structure in $\alpha 1$ with a $\beta 3$–$\beta 4$ hairpin. The $\alpha 1$, $\beta 3$, and $\beta 4$ elements of secondary structure form a scaffold or nucleus on which the remaining β-strands pack. This creates the metastable intermediate state observed in the kinetics. Further collapse and consolidation then occurs which brings residues involved in stabilizing the rate-limiting transition state together. During the formation of this nucleus, there is a refinement of packing interactions and consolidation of the interface between the two structural subdomains. Thus, barnase shows more hierarchical folding than CI2, with a gradual build-up of structure along the folding pathway. A schematic representation of the folding pathway of barnase is shown in **Figure 3b**.

Ultrafast Folders: The Engrailed Homeodomain

CI2 and barnase fold on the order of milliseconds to seconds at room temperature. Recently, proteins which fold with much faster rates have been identified and their folding pathways characterized using the techniques described above. The engrailed homeodomain from *Drosophila melanogaster* (En-HD) is a 61-residue, highly helical protein which folds on the microsecond timescale under physiological conditions. When unfolded in high concentrations of chemical denaturant or at very high temperatures, the unfolded state contains rather little residual secondary structure which is highly dynamic and fluctuating. Under more physiological conditions, however, at lower temperatures in the absence of denaturants, the denatured state has considerable residual structure with a high helical content and some long-range tertiary contacts. Combining experimental protein-engineering results with molecular dynamic simulations has provided a detailed picture of how this protein folds, **Figure 3c**. From the physiological denatured state (D), the transition state (‡) is formed by a reorientation of the helices and docking of them to form a partially packed hydrophobic core. In

the final steps, the helices fully dock against each other and water is expelled from the hydrophobic core. En-HD folds by first forming relatively stable elements of secondary structure (in this case α-helices) which then diffuse and collide to form the tertiary structure and the native state. This is much like the framework model proposed by Ptitsyn and coworkers in the early days of protein folding.

A Unifying Mechanism?

Outlined above are the results of detailed experimental and computational studies on the folding pathways of three proteins. These three proteins are excellent examples illustrating the different mechanisms by which proteins can fold. It is useful to include a short history of the field of protein folding here, to show how recent work has established new folding mechanisms and shed light on mechanisms which were proposed in the early days of the field when high-resolution structural information on folding pathways was not yet available. The Holy Grail of folding is, of course, to unite these different mechanisms into a simple single mechanism applicable to all proteins. This unified mechanism is presented later after a brief discussion of individual mechanisms.

Wetlaufer was one of the first to propose a nucleation type of mechanism, that of nucleation growth, back in 1973. Here, an initial nucleus of local secondary structure forms and then the tertiary structure propagates rapidly from this. As this mechanism does not require the formation of stable intermediate states – and at that time all experimental studies were focused on proteins which folded slowly and had discrete intermediates (later some of these intermediates were shown to be misfolded species off the main folding pathway) – the nucleation-growth mechanism was not widely adopted. At that time, experimental studies were limited by suitable experimental techniques and rapid-reaction kinetics had not yet been applied to complex biological problems such as protein folding. In addition, protein engineering was still in its infancy. Other mechanisms were proposed and dominated in the 1970s and 1980s: the framework model in which secondary structure folds first, then diffuses and collides to form the three-dimensional native structure; the hydrophobic collapse model, in which the hydrophobic effect drives the formation of a compact state in which many hydrophobic side chains are excluded from the solvent water, rearrangement then takes place to form the fully folded structure. This model was extended and it was proposed that a secondary structure was formed during the collapse stage to form a "molten-globule"-like intermediate state. These molten-globule states have a native-like secondary structure but a liquid-like core in which side chains retain a high degree of mobility and in which fixed tertiary interactions are not yet formed. In the early 1990s, and as a direct result of protein-engineering studies and ϕ-value analysis by Fersht and coworkers on CI2, a new mechanism – the nucleation–condensation model – was proposed.

The results and models of the folding of CI2, barnase, and En-HD are typical of the range of results obtained for the folding of a number of small proteins. Despite apparent differences in the way in which these three proteins fold, Daggett and Fersht have recently proposed a unified mechanism of folding based on the nucleation–condensation model. In their view of folding, there is a spectrum of pathways available to a protein with the framework mechanism at one end of the spectrum and the hydrophobic collapse model at the other. The nucleation–condensation mechanism lies in the center of this spectrum and the two other models can be viewed as extreme cases. Proteins which have sequences with high intrinsic propensities for forming stable elements of secondary structure, such as the engrailed homeodomain, tend to fold with mechanisms which lie at one end of the spectrum and which approximate to a framework model. Proteins for which the conformational preferences for secondary structure are weak, such as CI2, tend to fold according to a nucleation–condensation mechanism in which secondary structural elements are stabilized by tertiary interactions in the transition state. Larger proteins with more than one structural subdomain, such as barnase, fold in a hierarchical fashion but each subdomain, however, folds with a mechanism which lies somewhere on the spectrum of mechanisms found for smaller proteins. It is interesting that, although the hydrophobic collapse model can be viewed as one extreme of the folding spectrum, there are no examples so far of small proteins which fold in this way. Theoretically, proteins may fold through a molten-globule-like intermediate if hydrophobic interactions are very strong.

Computational Approaches

It is clear from the discussions on CI2, barnase, and En-HD given above that results from computational studies in conjunction with experimental results of ϕ-values from protein-engineering studies have played a critical role in the development of new folding mechanisms and in increasing our knowledge of the complexities of the energy landscape for

folding of small proteins. Although it is beyond the scope of this article, a few words on the main computational techniques used are included here. Many different computational approaches have been employed in the 1990s and applied to the protein-folding problem. These approaches range from rather coarse-grained models in which amino acids in the polypeptide chain are treated as beads on a string, a lattice-based structure is used, and simple pairwise potentials are employed. At the other end of the scale, all-atom molecular dynamic approaches have also been widely applied to the protein-folding problem. The different approaches, limited in different ways, have provided very complementary information on folding mechanisms and the fundamental questions regarding folding. Simple models and lattice-based approaches have provided support for nucleation-based mechanisms of folding, as well as being used to investigate cooperativity in folding. All-atom MD simulations are the most realistic of computational approaches but are restricted by the computational power and time necessary to simulate complex biological systems such as proteins (10 000 plus atoms). With the exception of small structural motifs, such as β-hairpins, for which folding has been simulated using distributed computing techniques, simulations of folding itself are not possible. This problem has been circumvented by simulating unfolding processes using high temperatures to accelerate the unfolding event. In these cases, the potentials used correspond to lower temperatures, the effect of the high temperature merely being to speed up the process but not change the mechanism. Using the principle of microscopic reversibility, the unfolding and folding pathways must be identical under the same conditions. Thus, it has been possible to gain information on folding pathways using these techniques. Rigorous testing of this approach has been performed with multiple simulations of the same protein under different temperatures and conditions. Results of this type of testing have shown the technique to be robust. Further, this type of approach has now been benchmarked extensively against experimental data and shown to be reliable and accurate in predicting experimental results. From these simulations, ensembles of structures corresponding to the denatured, intermediate, and transition states have been identified and characterized for a significant number of small proteins. The data obtained from these simulations is very much complementary to that obtained from ϕ-value analysis, the one generating structures whilst the other provides information on the energetics of the system.

Recently, experimental and computational approaches have been combined and new methods developed by Vendruscolo, Paci and co-workers to probe in further detail the structures of, and interactions in, transition state ensembles. Experimental ϕ-values are used as restraints in simulations, in a manner similar to the use of NOE data in NMR structure determination, and computational approaches used to generate an ensemble of conformations corresponding to the transition state. The advantage over the MD unfolding simulations described above is that a large number of conformations can be generated rapidly. This approach has been successfully applied to the small 98-residue protein acylphophatase to characterize the network of interactions that stabilize the transition state and which are formed when a few key residues (the folding nucleus) form a native-like arrangement.

See also: Folding and Sequence Analysis; Protein Folding and Aggregation; Protein Folding, Evolution and.

PACS: 87.80. – y; 82.37.Rs; 87.14.Ee; 87.15.Cc; 87.15.Aa

Further Reading

Daggett V and Fersht AR (2003) Is there a unifying mechanism for protein folding? *Trends in Biochemical Sciences* 28: 18–25.

Ferguson N and Fersht AR (2003) Early events in protein folding. *Current Opinion in Structural Biology* 13: 75–81.

Fersht AR (1999) *Structure and Mechanism in Protein Science*. London: Freeman.

Fersht AR and Daggett V (2002) *Cell* 108: 573–582.

Fersht AR, Matouschek A, and Serrano L (1992) The folding of an enzyme. 1. Theory of protein engineering analysis of stability and pathway of protein folding. *Journal of Molecular Biology* 224: 771–782.

Jackson SE (1998) How do small single-domain proteins fold? *Folding & Design* 3: R81–R90.

Matthews CR (1987) Effect of point mutations on the folding of globular proteins. *Methods in Enzymology* 154: 489–511.

Murphy KP (2001) *Protein Structure, Stability and Folding*. Totowa, NJ: Humana Press.

Pain RH (2000) *Mechanisms in Protein Folding*. New York: Oxford University Press.

Tanford C (1970) Protein folding, Part C. *Advances in Protein Chemistry* 24: 1–95.

Vendruscolo M and Paci E (2003) Protein folding: bringing theory and experiment closer together. *Current Opinion in Structural Biology* 13: 82–87.

Protein Folding, Evolution and

R A Goldstein, National Institute for Medical Research, London, UK

Introduction

Many important problems in physics involve the movement of a descriptor over some complicated, often high-dimensional, and sometimes rugged search space. The relationship between evolution and protein folding involves the search over two extremely different interacting spaces; the space of all possible protein sequences and the space of all possible protein conformations. For this and other reasons, physicists have found this area fruitful and challenging.

DNA provides the basic object of evolution for essentially all living systems, from the 580 000 base-pair genome of *Mycoplasma genitalium* to the 3.3×10^9 base-pair genome of humans to the greater-than 10^{11} base-pair genome of some amphibians and plants. Various sections of these genomes, called genes, contain instructions for constructing some biological entity, such as an RNA or a protein. It is the proteins, often combined with other elements (metallic atoms and clusters, RNA, other simple molecular "cofactors," etc.) that fulfill most of the active roles in the organism, from locomotion to metabolism to signaling. Proteins are initially synthesized as a long unbranched heteropolymer consisting of a linear arrangement of various numbers of the 20 amino acids or residues. Before they can fulfill their function, most proteins have to fold into a compact, regular three-dimensional structure.

Folding into a native conformation is a nontrivial task. Consider a moderately sized protein of 200 amino acids, each of which can be in one of three conformations. That means that, ignoring constraints such as excluded volume, the entire protein can have $3^{200} \sim 2.7 \times 10^{95}$ possible conformations. (A concentrated sample of such a protein, with one example of each conformation, would form a ball 10^8 light-years across!) Of all these conformations, some small proportion would be considered correctly folded. While factors such as excluded volume reduce the number of possible conformations, it is clear that finding a protein's folded conformation involves searching for a molecular needle in a cosmological haystack. A number of investigators have demonstrated that such folding is NP-hard.

An additional constraint is that protein folding is highly cooperative, often all-or-nothing. This cooperativity is essential for thermodynamic stability in that there is a large energy penalty for any significant unfolding. If this were not the case, given the number of different ways that some unfolding can occur, the proportion of correctly folded proteins at equilibrium would be negligible. While a few simple schemes have been suggested that allow folding to occur by a simple and gradual reduction in the free energy, these schemes are not consistent with the required and observed cooperativity.

Nevertheless, proteins solve this problem regularly. The fastest proteins fold in microseconds, with folding times of seconds to minutes more common. The difficulty is our problem, not theirs. So how are proteins able to fold so quickly?

There have been two categories of approaches to this problem. The first, more common among biochemists, has been to look for an encoding of the folding pathway. Just as a sheet of origami paper (for beginners) has dotted lines saying "fold here," may be there are patterns in the amino acid sequence that say when certain structural elements should be initiated or terminated. This view of the folding process is supported by the modular nature of proteins. Proteins exist as assemblies of simpler elements: shorter regions of the protein sequence fold into characteristic regular (α-helical and β-sheet) or nonperiodic (turn and loop) "secondary" structures, which are built into three-dimensional "supersecondary" structures containing a small number of these elements, which are in turn assembled into "tertiary" structures. Larger protein structures can be divided into "domains," regions of sequence of 100–300 residues that fold and function as separate entities. Different protein chains sometimes assemble to form "quaternary" complexes. It is argued that the modularity makes it possible for the folding code to be local.

More recently, others (such as physicists) have addressed this issue by considering the free-energy landscape of the proteins, that is, the free energy of the protein sequence as a function of its conformation, and considered how this landscape could be negotiated by a folding protein. Often this has been combined with notions that there is not a single folding pathway, but rather a multiplicity of ways that a protein can go from the large ensemble of unfolded states to the relatively small region corresponding to the folded state. Also, common in this approach is that the "folding code" is distributed in a holographic way throughout the sequence. Theoretical models based on the spin-glass theory by Wolynes and co-workers, in addition to lattice-model simulations by Shakhnovich and co-workers, have suggested that a sufficient criterion for ability to fold is for the protein

to have a native state sufficiently stable, relative to the stability of random conformations. In particular, computational models of lattice proteins that are optimized for rapid folding tend to develop such an "energy gap." More lately, there has been interest in how proteins might develop "funnels" in the free-energy landscape that direct the folding process, and how protein folding might involve the formation of a specific folding nucleus as the rate-limiting step. Again, the challenge is to reduce the search problem while not losing the essential cooperativity.

While protein folding may be NP-hard, this only means that it is impossible to find a polynomial-time general solution. One can ask instead how nature, through the process of evolution, can discover those proteins whose folding is sufficiently easy. This then brings one to the dual search spaces. Consider first the "fitness landscape" of all possible protein sequences. This space is high dimensional, with as many dimensions as the length of the protein, but is also sparse – only 20 points along each dimension. Each sequence in this space has a fitness value associated with it, corresponding to the fitness of the organism with that particular protein sequence. (In reality, the fitness landscape would have as many dimensions as the entire genome, but it is assumed that all of the other dimensions are frozen.) Each point in sequence space is mapped to a corresponding free-energy landscape, encoding the free energy as a function of the conformation of that particular sequence. Some of these free-energy landscapes correspond to fast-folding proteins, not the vast majority. The "foldability" of any particular sequence, that is, the ability of that sequence with its corresponding free-energy landscape to fold, is then a component in the fitness of that sequence in the fitness landscape. As the properties of the free-energy landscape form a contribution to fitness, evolution will result in proteins with free-energy landscapes compatible with acceptable folding rates, a process called "free-energy landscaping."

There is rather universal agreement on certain aspects of evolution. Mutations occur among somatic cells, which affect the fitness of the offspring. Most of the time, this particular mutation will be eliminated from the population. There is a finite chance that, alternatively, the mutated sequence will expand among the population and become the dominant genotype, a process referred to as "fixation." There is also a chance that this mutation will remain in a fraction of the population without eliminating other forms, and this part of the genome will be "polymorphic." Some mutations will be beneficial, others neutral, and others detrimental. While a beneficial mutation has a higher probability of fixation compared with neutral or detrimental mutations, there is a probability of fixation of a detrimental mutation just as there is a probability that a beneficial mutation will be eliminated. These probabilities will depend on the population size, with randomness increasing as the population shrinks. There is a ranging debate about the relative number of beneficial, neutral, and detrimental mutations. While all agree that the vast majority of mutations are detrimental, there are some evolutionists who believe that the mutations that are accepted are generally accepted because they are beneficial, that is, increase the fitness of the organism. Accepted mutations are, then, likely to be "adaptive." Others argue that effectively neutral mutations exceed beneficial mutations to such an extent that most evolution proceeds through "random drift," the accumulation of such neutral mutations.

These neutral mutations can have many important consequences. They allow more freedom in the search over the sequence space, enabling the relatively few adaptive changes to occur. In addition, greater random neutral drift in a population acts in a manner similar to a higher temperature, making "sequence entropy" (i.e., the number of sequences corresponding to a given trait) important in the evolutionary dynamics. Higher mutation rates and smaller population sizes correspond to larger evolutionary temperatures.

One aspect emphasized by neutralists is the dangers of the so-called "Panglossian paradigm," that is, explaining why a characteristic evolved through reference to how it is currently used, in the same way as Candide's Pangloss described how noses were obviously created for supporting spectacles. Evolution is not an optimization procedure, and the caricature embodied by the notion of "survival of the fittest" is not a perspective Darwin would appreciate. In particular, features created for one reason can be used for an alternative task, and characteristics that arise at random may fulfill no role. If random walks through viable regions of sequence space represent a "null model," claiming that an acquired characteristic is adaptive requires demonstrating that it cannot be explained by neutral drift.

There is an interesting timescale separation concerning evolution and folding. Evolution does not occur on the folding timescale, so there is an unchanging sequence that folds. Although there are accessory proteins (such as, the so-called chaperones) that may assist a protein in folding, there is little evidence that these accessory proteins provide any information about the nature of the final folded state. The protein sequence, then, determines the final folded state. In an evolutionary timescale, sequences

change much more quickly than structures, so that two sequences with near-random sequence similarities may fold into near-identical structures. Changes in the sequence inconsistent with this conserved fold will be detrimental and eliminated from the population. Thus, on the evolutionary timescale, it is the folded structure that determines (or at least constrains) the protein sequence.

Certain questions immediately arise. How do proteins evolve to solve the protein-folding problem? What are the properties of the solutions that they find? What are the other consequences of these solutions? The answers to these and similar questions are only now starting to emerge. Many perspectives will likely change quickly as one learns more about proteins, protein folding, and protein evolution.

The Formation of the Universe of Protein Folds

When examining proteins of known structure, it is obvious that not all folds are equally represented. Rather, there are very few extremely likely folds, while the vast majority of folds are rather rare. The observed distribution has been modeled with a few different functional forms, including a power law and a stretched exponential. Four basic explanations have been offered for this uneven distribution. The first is that there are certain structures that are more amenable to structural characterization, and that the distribution is an experimental artifact. The second is that there are certain folds that are compatible with a wider range of functions. The third is that this represents the results of a process of speciation and gene duplication. The fourth explanation is based on ideas of protein "designability," that is, there are certain folds that can be formed by a larger number of sequences compared with other folds. Designability could cause an uneven distribution of folds for one of two reasons. First, initial instances of a given protein would be more likely to result in a highly designable structure. Second, highly designable structures correspond to larger "neutral networks," that is, larger regions of sequence space corresponding to a viable protein. This means that these proteins could have greater freedom to evolve to acquire new characteristics (new specificities, new functionalities) while preserving the folded structure, and would also have an additional evolutionary advantage due to the higher degree of robustness to mutations. It has also been noted that highly designable proteins are more likely to have a large energy gap between the folded and unfolded states, considered in some models to be a sufficient condition for rapid folding. Detangling these different explanations will require a broader analysis of the current distribution, which will be assisted by the various sequencing projects and by current efforts to characterize the broad range of protein structures in a comprehensive manner.

What characterizes highly designable structures? Finkelstein first considered the stability penalty certain substructures would cause, arguing that this would reduce the designability. Calculations for lattice proteins using a simplified amino acid alphabet led Tang and co-workers to note the greater representation of highly symmetric structures. Shakhnovich and co-workers developed analytical models suggesting that designability was strongly influenced by the contacts present in the folded state, especially the total number of contacts as well as the different number and kind of closed loops formed by these contacts. Goldstein and co-workers emphasized the number of contacts unlikely in random structures. Strengthening such contacts, generally between amino acids well separated in the sequence, would stabilize the native state while stabilizing the fewest number of alternative conformations. Micheletti and co-workers provided a different twist with a study of the normal vibrational modes in a protein, demonstrating that protein structures have slow relaxation times relative to random graphs, and that slower relaxation times are correlated with faster folding rates. This is because these slower relaxation times are correlated with larger, low-energy motions that may indicate a smoother free-energy landscape. In this model, the particular protein structures commonly found by evolution may represent better folders in as much as they are compatible with these slower relaxation times.

Have Proteins Evolved for Fast Folding?

There is a current debate over what nature considers important during protein evolution. Have proteins evolved for rapid folding? Is this one of the major sources of selection pressure acting on protein evolution? Or is fast folding a consequence of selective pressure acting on some other characteristic, such as stability? Are other properties of folding under selective pressure?

One approach to gain insight into these questions is to consider how changes in the protein sequence affect folding. In general, the folding rate can be rather insensitive to such changes. In fact, it is often not difficult to find mutations that speed up folding, sometimes significantly. This would seem to indicate that proteins have not evolved under strong selection to fold quickly, as strong selection would likely have yielded faster folding sequences.

Mutational work can also provide information about the folding pathway, an approach pioneered by Alan Fersht. Depending upon the results, Alan Fersht and colleagues were able to identify which amino acids are involved in (presumably native) interactions in the transition state, what has been characterized as the "folding nucleus." This is quantified with a so-called "Φ value"; Φ values close to 1 indicate strong involvement of the amino acid in the transition state, while Φ values closer to 0 indicate lack of involvement. Intermediate Φ values, observed in a large number of proteins, indicate either the partial involvement of that amino acid in the transition state or heterogeneity in the folding process with a distribution of transition states. The presence of such a large number of intermediate values suggests that the folding nucleus is often somewhat diffuse. It is often observed that homologous proteins share the same folding nucleus, as indicated by similar patterns of Φ values. This could mean that there is a tendency for conservation of the folding nuclei. An alternative explanation, indicated by a number of different dynamics simulations, is that the dynamics of the folding of the protein is often relatively insensitive to the actual nature of the interactions that determine the folded state, but depends, instead, mostly on the resultant state. The similarity of the folding nuclei in related proteins may then just represent similarity of the final structure. It is unclear, if the folding pathway is conserved, what the evolutionary advantage of a conserved pathway would be. Baker and co-workers, for instance, demonstrated that it is possible to change the folding pathway through protein engineering, resulting in a more-stable, faster-folding variant.

Φ-value analysis provides an alternative method to investigate whether the folding rate is under active selective pressure. Are residues with larger Φ values, those that are important in the folding nucleus, more conserved? Such a study was performed by Shakhnovich, who observed that residues known to have high-Φ values are, indeed, more highly conserved than the average residue in a protein. This result was subsequently disputed by Baker and colleagues who pointed out a strong experimental bias in the data: namely, not all residues in any protein are characterized experimentally, and the experimentally characterized residues are, on an average, more highly conserved than those that have not been characterized. If one only considers locations with measured Φ values, there is little or no correlation between Φ value and amino acid conservation.

If it is unclear whether folding path or rate is under active selective pressure or not, what is? Artificial mutations are more likely to reduce protein stability, although even this quantity can be rather robust. This suggests that protein stability is under more selective pressure than folding rate. While theoretical calculations link these two quantities, suggesting that a certain stability is sufficient for rapid folding, one might imagine that stability might be important in its own right. All proteins spend some fraction of their time at equilibrium in an unfolded state. Reducing this time might help prevent degradation and aggregation. It is interesting to note, however, that proteins, in general, are marginally stable (stabilities around $-10\,\mathrm{kcal\,mol^{-1}}$, equivalent to only a few hydrogen bonds), that many mutations can be found that enhance protein stability, and that many proteins may actually be largely unfolded under physiological conditions. It has been widely suggested that the observed marginal stability of proteins may represent some adaptation for functionality, although it would also naturally result from neutral evolution.

One interesting approach toward analyzing folding involves the use of circular permutations. The termini of many proteins are near one another. It is possible to create a new protein starting from a random point in the original protein, proceeding to the far termini, continuing across to the beginning of the protein, and then synthesizing the original sequence up until the starting point. The result is a protein with almost the identical final structure (except for the relocation of the termini), but with an extremely different conformational landscape. Interestingly, the reordered proteins do not have appreciably slower folding rates, nor are they less stable, but rather the folding is much less cooperative, with more intermediate Φ values. This suggests that it may be the all-or-nothing nature of the folding that may be under the most selective pressure.

Folding is a difficult problem, but one aspect that may be more directly important is avoiding misfolding. For instance, there may be strong selective pressure on certain amino acids that destabilize alternatively folded states. Often the result of misfolding is the formation of protein aggregates, often connected with pathological conditions such as Alzheimer's, Creutzfeldt–Jakob disease, and type II diabetes. A wide range of proteins can form destructive aggregates under appropriate conditions. The fact that only a relatively small fraction does under physiological conditions suggest that evolution has been quite good at preventing protein misfolding. Evidence of this can be seen in the accepted mutations during evolution. Although amino acids on the outside of the protein generally change faster than internal amino acids, hydrophilicity on the exterior of the protein is sometimes more conserved than hydrophobicity on the protein interior. This would

prevent aggregation as well as discourage folding to alternative states where these amino acids are buried.

Do Proteins Fold into Their Lowest Energy State?

An example of the neutralist approach to protein folding is one of the basic issues in the field; do proteins fold into the state of lowest free energy? This hypothesis, called the "thermodynamic" hypothesis, was proposed by Anfinsen based on the ability of unfolded proteins to refold spontaneously in solution. An alternative perspective was developed by Levinthal. Noting the extremely large number of possible conformations, and the inability of a protein to achieve something resembling thermodynamic equilibrium in any reasonable timescale, led him to doubt the thermodynamic hypothesis. In the absence of any selective pressure, could neutral evolution result in proteins fulfilling the thermodynamic hypothesis? In 1998, Goldstein and co-workers proposed a model for how this could happen. Mutations that destabilize the free-energy minimum corresponding to the folded state might be highly detrimental, and thus eliminated. Conversely, mutations that destabilized other, possibly deeper minima would not have a negative effect on the protein, would be more likely to be advantageous (or at least neutral), and thus have a greater chance for fixation. The result might be the evaporation of these alternative minima over an evolutionary time, resulting in proteins that did fulfill the thermodynamic hypothesis, even in the absence of any selective advantage. The quantitative question concerns the number of expected alternative minima, the probability that they would be explored, and the probability that any minima would be lower in energy than the true native state. A simple calculation indicated that, for reasonable energy parameters, sequences with metastable folded states would be much less likely than proteins fulfilling the thermodynamic hypothesis, and thus neutral evolution would favor this hypothesis. Of course, this argument would fail when there was a specific reason why a metastable folded state would be advantageous, as seems to be the case with serpins, a group of inhibitors that prevent enzymes known as proteases from digesting other proteins, which inactivate themselves by converting from an active metastable structure to a more stable, alternative form.

Evolution and Optimization

One overall problem with a number of these analyses is that evolution is not an optimization process, any more than thermodynamics is an enthalpy-reduction process. (While Baker and co-workers have found that the sequences of proteins can be appreciably constructed by finding the sequence lowest in energy for a given backbone structure, this may be due to information about the protein sequence being embedded in this backbone structure. The reconstructed information sharply degrades as the backbone is allowed to move. These results are also incompatible with the wide range of sequences observed in homologous proteins.) There are stochastic processes underlying mutations and fixation events. Chance can operate either through the memorization of accidental events, as well as through the role of entropic considerations.

In particular, there is selection not for a single point in sequence space, but rather for a distributed somewhat heterogeneous population, what Eigen called a "pseudo-species." As this population samples a region of the fitness landscape, there is a preference for evolution to select flatter regions of the landscape, compared with narrower, higher peaks. This has been called, by Crutchfield and co-workers, the "Survival of the Flattest." Goldstein and co-workers have performed simulations indicating that quantities selected by evolution will be "buffered," that is, these quantities will be surprisingly robust to sequence change. According to this argument, quantities that are seemingly not optimized may be those under strong selective pressure.

See also: Folding and Sequence Analysis; Protein Folding and Aggregation; Protein Folding, Engineering of.

PACS: 87.14.Ee; 87.15.; 87.15.Cc; 87.23.Kg

Further Reading

Buchler NEG and Goldstein RA (1999) The effects of alphabet size and foldability requirements on protein structure designability. *Proteins* 34: 113–124.

Govindarajan S and Goldstein RA (1996) Why are some protein structures so common? *Proceedings of the National Academy of Sciences USA* 93: 3341–3345.

Govindarajan S and Goldstein RA (1997) The foldability landscape of model proteins. *Biopolymers* 42: 427–438.

Govindarajan S and Goldstein RA (1998) On the thermodynamic hypothesis of protein folding. *Proceedings of the National Academy of Sciences USA* 95: 5545–5549.

Kim DE, Gu H, and Baker D (1998) The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proceedings of the National Academy of Sciences USA* 95: 4982–4986.

Larson SM, Ruczinski I, Davidson AR, Baker D, and Plaxco KW (2002) Residues participating in the protein folding nucleus do not exhibit preferential evolutionary conservation. *Journal of Molecular Biology* 316: 225–233.

Li H, Helling R, Tang C, and Wingreen N (1996) Emergence of preferred structures in a simple model of protein folding. *Science* 273: 666–669.

Lindberg M, Tangrot J, and Oliveberg M (2002) Complete change of the protein folding transition state upon circular permutation. *Nature Structural Biology* 9: 818–822.

Micheletti C, Lattanzi G, and Maritan A (2002) Elastic properties of proteins: insight on the folding process and evolutionary selection of native structures. *Journal of Molecular Biology* 321: 909–921.

Mirny LA, Abkevich VI, and Shakhnovich EI (1998) How evolution makes proteins fold quickly. *Proceedings of the National Academy of Sciences USA* 95: 4976–4981.

Mirny L and Shakhnovich E (2001) Evolutionary conservation of the folding nucleus. *Journal of Molecular Biology* 308: 123–129.

Nauli S, Kuhlman B, and Baker D (2001) Computer-based redesign of a protein folding pathway. *Nature Structural Biology* 8: 602–605.

Onuchic JN, Luthey-Schulten Z, and Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry* 48: 545–600.

Orengo CA, Sillitoe I, Reeves G, and Pearl FM (2001) What can structural classifications reveal about protein evolution? *Journal of Structural Biology* 134: 145–165.

Ortiz AR and Skolnick J (2000) Sequence evolution and the mechanism of protein folding. *Biophysical Journal* 79: 1787–1799.

Plaxco KW, Larson S, Ruczinski I, Riddle DS, Thayer EC, *et al.* (2000) Evolutionary conservation in protein folding kinetics. *Journal of Molecular Biology* 298: 303–312.

Taverna D and Goldstein RA (2000) The distribution of structures in evolving protein populations. *Biopolymers* 53: 1–8.

Taverna D and Goldstein RA (2002) Why are proteins so robust to site mutations? *Journal of Molecular Biology* 315: 479–484.

Tiana G, Broglia RA, and Shakhnovich EI (2000) Hiking in the energy landscape in sequence space: a bumpy road to good folders. *Proteins* 39: 244–251.

van Nimwegen E, Crutchfield JP, and Huynen M (1999) Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences USA* 96: 9716–9720.

Pseudopotential Method

A Filippetti, University of Cagliari, Monserrato, Italy
G B Bachelet, Università di Roma "La Sapienza", Roma, Italy

Introduction

Pseudopotentials were introduced to model the interaction between ions and valence electrons. They effectively eliminate, from the very start, the true electron–nuclear potential and the "inner core" electrons, that is, those electronic states which are tightly bound to the nucleus, do not participate in the formation of chemical bonds, and remain approximately unchanged in atoms, molecules, and solids. This tremendous conceptual simplification also amounts to a very practical and efficient computational scheme, especially when a plane-wave basis set is adopted to expand the electronic wave functions. After 1980, this method, in connection with a density-functional description of the electron–electron interaction, evolved into a reliable prescription for the first-principles computation of electronic, structural, and dynamical properties of molecules and solids. In this context, its present accuracy and predictive power for real materials, as well as relatively complex molecular systems, leads many experimental groups to regard its results as a valuable, independent source of microscopic information.

The physical and chemical basis for this fortunate approximation is twofold. First, in an atom only a few electrons determine the formation of chemical bonds to other atoms, while most of the electrons, tightly bound to their nucleus, are practically frozen; pseudopotentials simultaneously eliminate the unimportant inner electrons and the nuclear potential responsible for their binding. Second, the quantum nature and the Fermi statistics of the electrons are such that the combination of repulsion and attraction due to the inner electrons and to the nucleus results in substantial cancellations, and can be effectively replaced by a relatively smooth ionic pseudopotential, which only acts on the few important electrons.

To fully appreciate contemporary first-principles pseudopotentials, a few key concepts are needed.

Valence and the Periodic Table

Between the end of the eighteenth century and the beginning of the nineteenth century, the pioneering work of Lavoisier, Dalton, Gay–Lussac, Avogadro, and others led to the distinction between compounds and elements, to the discovery that elements combine in fixed and multiple proportions of whole numbers to yield compounds, and, ultimately, to the modern concept of the atom. The formulas of compounds were soon systematized by assigning certain combining powers, the valences, to the elements, as if each atom had a particular number of little hooks on it to attach itself to similar hooks on other atoms. Half a century later, Mendeleev drew up his periodic table of the elements, with eight columns reflecting the remarkable correlations between the atomic valence and weight, first pointed out by Newlands.

What in his table was just a progressive number was in fact the atomic number Z (the number of protons in the nucleus or, equivalently, the number of electrons of the neutral atom) – as shown by Moseley's X-ray experiments, at the beginning of the twentieth century.

The Atomic Shell Structure

In the same years, the advent of quantum mechanics (first the Bohr model and then the Schrödinger equation) allowed a simple and consistent interpretation of the remarkable regularities of the periodic table of the elements.

Atoms are built up of successive orbital shells of electrons, labeled by the principal quantum number n and the angular quantum numbers lm, which are less and less tightly bound to the nucleus. For given nl, each of the $(2l+1)$ equivalent orbitals nlm can host, because of the Pauli principle, a maximum of two electrons. A complete s shell ($l=0$) thus contains $2(2l+1)=2$ electrons, a complete p shell ($l=1$) 6 electrons, and so on. The Pauli principle (i.e., the Fermi statistics obeyed by the electrons) is thus the reason for the progressive filling of orbital shells, as the atomic number grows. Elements of the same group of the periodic table, with different Z and different number of inner, complete core shells, have a common feature: the number of electrons and the electronic configuration of their outermost shell – the valence shell. That is why their chemical properties are so similar; and also why (disregarding transition d- and f-metals) chemical properties go by "octaves" (Newlands) and the groups are eight: it takes eight electrons to completely fill an sp valence shell.

The persistence throughout the periodic table of a recognizable hydrogen-like orbital shell structure explained a lot of experiments in the early days of quantum mechanics, but is not so obvious. Somehow, one-electron quantum numbers and orbitals $\psi_{nlm}(\boldsymbol{r}) = R_{nl}(r)Y_{lm}(\hat{\boldsymbol{r}})$ (where the vector $\boldsymbol{r}$ represents the position of the electron with respect to the nucleus, R are radial functions, and Y spherical harmonics) remain meaningful even for atoms which, unlike hydrogen, have many electrons and not just one. Why? Because, in atoms, to a good approximation, (1) the electron–electron repulsion may be replaced by its average: a mean-field, one-electron potential which screens the strong, attractive $-Z/r$ Coulombic potential of the nucleus, (2) such a mean-field potential is, in turn, well-approximated by its spherical average, so that the atomic states can still be factorized into a radial and an angular part, and thus labeled by the same quantum numbers nlm as the hydrogen atom, (3) since the nucleus is a positive point charge while the negative electronic clouds are more or less delocalized, their spherically-averaged mean-field repulsive potential provides, at finite distance from the nucleus, only a partial screening of the nucleus, and does not completely destroy the order of the energy levels of the bare $-Z/r$ nuclear potential. The validity of the mean-field approximation (1) in atoms is the cornerstone of the remarkable simplicity of their interpretation in terms of one-electron orbitals and quantum numbers, and amounts to the success of the Hartree–Fock variational solution of the true, many-electron quantum problem. Combined with the spherical average (2), it results in a set of hydrogen-like equations (atomic units $\hbar = m_e = e = 1$ (where m_e and e are electron mass and charge, respectively) are used here and in what follows, unless otherwise stated):

$$\hat{h}_{SCF}\psi_{nlm} = \left[-\frac{1}{2}\nabla_{\boldsymbol{r}}^2 - \frac{Z}{r} + \hat{v}_{screen}\right]\psi_{nlm} = \varepsilon_{nl}\psi_{nlm} \quad [1]$$

where $\hat{h}_{SCF}$ is one-electron self-consistent field Hamiltonian and ε_{nl} are its eigenvalues. Equation [1] differs from the true (nonrelativistic) Schrödinger equation appropriate to noninteracting electrons by the presence of a screening potential $\hat{v}_{screen}$, which is a functional of the occupied orbitals $\psi_{nlm}(\boldsymbol{r})$. (The concept of functional generalizes that of function. A function associates a number to each number defined in a certain interval, its domain. A functional associates a number or a function to each function defined in a certain functional space. For example, it can be seen below that for any charge distribution $n(\boldsymbol{r})$, there is a corresponding electrostatic potential $v_H(\boldsymbol{r})$; v_H is then a functional of n: $v_H = v_H[n]$.)

The presence of $\hat{v}_{screen}$ makes the entire effective Hamiltonian $\hat{h}_{SCF}$ depend on its own eigensolutions; this is a nonlinear, but numerically feasible self-consistent-field (SCF) problem. Similar self-consistent equations emerge from the density-functional theory; the reason is quite subtle, but superficially speaking, in comparison to the Hartree–Fock equations, the difference is that the screening potential $\hat{v}_{screen}$ depends on, and acts upon, the orbitals $\psi_{nlm}(\boldsymbol{r})$, in a different way. This difference, however, concerns only part of the screening potential, the so-called exchange-correlation potential, which is not crucial for a qualitative understanding of the atomic shell structure, since the main contribution of the occupied electronic states to $\hat{v}_{screen}$ always comes from the other part, their classical electrostatic (or Hartree) potential

$$v_H(\boldsymbol{r}) = \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r}' - \boldsymbol{r}'|}\mathrm{d}\,\boldsymbol{r}' \quad [2]$$

where $n(\boldsymbol{r})$ is the electronic charge density, sum of the squares of the occupied orbitals $\psi_{nlm}(\boldsymbol{r})$. In the

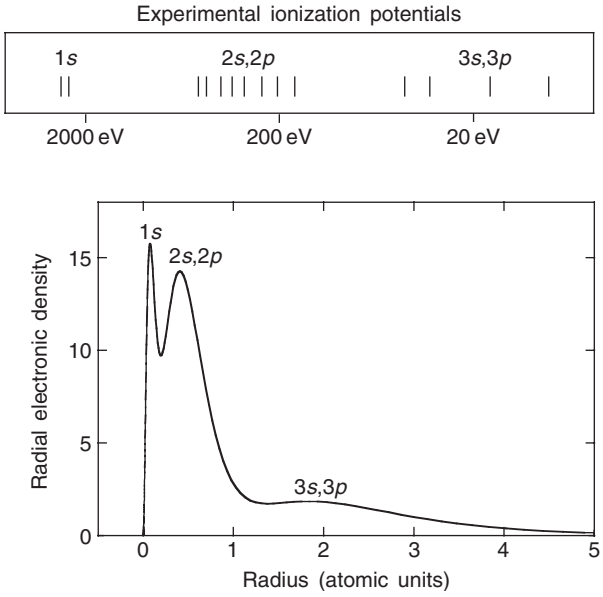


Figure 1 Atomic shell structure: the example of silicon ($Z = 14$). Top: the 14 ionization potentials of silicon, in electron volts (note the logarithmic scale). Bottom: radial electronic density as a function of the distance from the nucleus; both in energy (top) and in space (bottom), the core shells (1*s*, $n = 1$, and 2*s*, 2*p*, $n = 2$) are well separated from each other and from the valence 3*s*, 3*p* shell ($n = 3$).

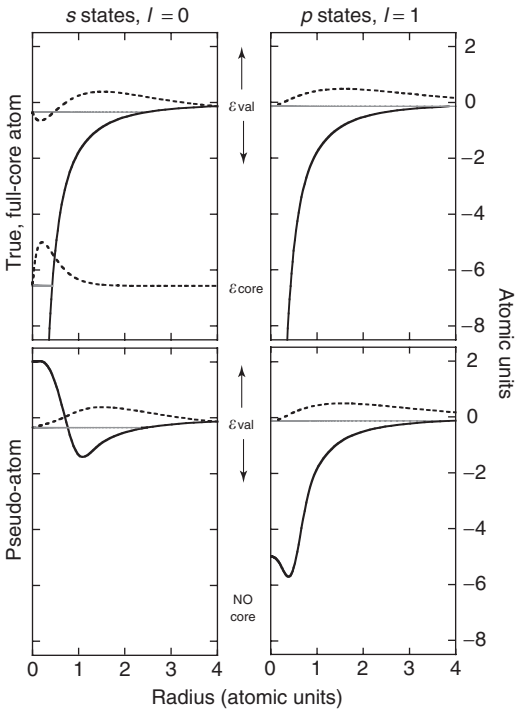


Figure 2 True vs. pseudo-atom, the example of boron ($Z = 5$). In the true neutral atom (top panels) there are five electrons. Two occupy the tightly bound 1*s* core orbital $\psi_{100} = R_{10}Y_{00}$, another two the shallow 2*s* valence orbital $\psi_{200} = R_{20}Y_{00}$ (top left), and the last one the 2*p* valence orbital (e.g., p_z, or $\psi_{210} = R_{21}Y_{10}$: top right). Here, the radial orbitals $\chi_{nl} = r\,R_{nl}$ are dashed lines and their zero is vertically shifted and lined up to the corresponding energy eigenvalue (gray line). The screened nuclear potential (top panels, black solid line) retains much of its original Coulombic shape, and a hydrogen-like shell structure results: the 1*s* orbital corresponds to a deep eigenvalue (ε_{core}, gray thick line) and is localized near the nucleus, while the 2*s* and 2*p* valence orbitals are "fatter" and correspond to a shallower energy (ε_{val}, gray thick line). In a pseudo-atom (bottom panels), the core electrons and the nucleus are simultaneously eliminated and replaced by a pseudopotential (bottom, black solid line); in boron, one is thus left with just three electrons, two in the *s* (left) and one in the *p* (right) pseudo-valence-orbital. These orbitals should ideally have the same energy eigenvalue and wave function amplitude as the true valence orbitals (top, dashes). The amplitude, though, can be the same only outside the core: inside it, the true valence 2*s* orbital (top left) is radially orthogonal to the core 1*s* orbital and changes sign (has one radial node), while, by definition, the pseudo-valence-orbital (bottom left) has no underlying core, and is thus nodeless. Similar considerations apply to energy: the pseudopotential (black solid, bottom panels) may effectively replace the true potential (black solid, top panels) only within some energy window around the valence eigenvalue ε_{val}. But this may be sufficient for an excellent approximation of the atom's behavior in a molecule or a solid, since bands and bonds spread the valence energies only a few atomic units away from the eigenvalue of the isolated atom (see text).

ground state of the neutral, atom they are occupied in order of increasing energy ε_{nl} (decreasing binding energy and localization), each by two electrons (spin, Pauli principle), except for the outermost valence states, some of which may be partially occupied or empty. The repulsive electrostatic potential v_H (eqns [1] and [2]) due to these concentric spherical clouds only partially screens the pointlike $-Z/r$ attractive singularity of the nuclear potential; a hydrogen-like shell structure survives, with the outermost (valence) orbitals usually well separated from the inner-core orbitals (see **Figure 1** and top panels of **Figure 2**).

From Free Electrons to Pseudopotentials

The atomic shell structure, the distinction between core and valence shells, and the fact that the core shells are tightly bound and closer to the nucleus, explain not only the periodic properties of the elements, but also a wealth of other experimental facts, from the energy gaps which separate different groups of atomic ionization potentials in a given atom (see **Figure 1**, upper panel) to the minor role played by the cores in the formation of chemical bonds and in the electrical and thermal properties of solids. Sommerfeld's free-electron model, for example, gave quite a reasonable picture of the conductivity and low-temperature specific heat of simple metals by completely disregarding both the core electrons and the nuclei; the valence electrons were its only ingredient, and, apart from being confined in the crystal, were subject to a completely flat potential. The success of such a crude model rests on more than just

a distinction between core and valence electrons, and is a good starting point for introducing the concept of pseudopotential. It tells that, at least in simple metals such as, say, sodium or potassium, valence electrons roughly behave as if core electrons and strongly attractive nuclei simply did not exist. The successful description of simple metals in terms of quasi-free electrons (an improvement over Sommerfeld's model based on perturbation theory) confirmed that the combined presence of strong and localized objects – core shells plus nuclear potentials – effectively results in a weak periodic perturbation, a lattice of smooth ionic potentials which slightly repel the valence electrons from the immediate vicinity of the nuclei.

If the basic idea of the pseudopotential method goes back to Fermi and in some sense also to Hellmann and Antončik, Phillips and Kleinman were the first, in a seminal paper which appeared on the *Physical Review* in 1959, to show that, inside the core region, the requirement of orthogonality of the valence states to the inner core states, which pushes the outermost maximum of the valence wave functions away from the core region (**Figure 2**, top left), acts like a repulsive potential which tends to cancel the attractive nuclear potential felt by the valence electrons, so that, more often than not, their combined effect may be well approximated by a relatively weak repulsive potential, or pseudopotential (**Figure 2**, bottom left). As soon as their "cancellation theorem" was supplemented by the key observation about the amount of arbitrariness in the choice of pseudopotentials and their general form, independently proposed by Bassani and Celli, and by Cohen and Heine in 1961, a practical prescription for their use became available. These results raised enthusiasm in the scientific community of solid-state theoretical physicists, who started a very fruitful exploitation of the pseudopotential idea, in combination with a basis set of plane waves to represent the valence wave functions. The reason is easily understood. A crystal with lattice vectors $\boldsymbol{R}$ is described by SCF equations analogous to the atomic eqn [1], where a single nucleus is replaced by a periodic array of nuclei, and the eigenstates labeled by the band index n and Brillouin-zone wave vector $\boldsymbol{k}$, satisfy the Bloch condition $\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{R}) = e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\psi_{n\boldsymbol{k}}(\boldsymbol{r})$. A plane wave $e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}$, where $\boldsymbol{G}$ is a reciprocal-lattice vector defined by $e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1$, individually satisfies the Bloch condition, and so does any linear combination of plane waves

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{n\boldsymbol{k}}(\boldsymbol{G}) e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} \qquad [3]$$

where $c_{n\boldsymbol{k}}(\boldsymbol{G})$ are one-electron wave functions in the reciprocal space. A basis set of plane waves thus represents a natural choice for solving the crystalline SCF equations. If the sum in eqn [3] includes all the reciprocal-lattice vectors $\boldsymbol{G}$ (a complete set of plane waves), the above expansion becomes an exact transformation, and turns the real-space SCF equations for $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$ into their $\boldsymbol{G}$-space counterpart

$$\sum_{\boldsymbol{G}'} \left[\frac{1}{2}|\boldsymbol{k}+\boldsymbol{G}|^2\delta_{\boldsymbol{G},\boldsymbol{G}'} + \tilde{v}(\boldsymbol{G}-\boldsymbol{G}')\right] c_{n\boldsymbol{k}}(\boldsymbol{G}') = \varepsilon_{n\boldsymbol{k}} c_{n\boldsymbol{k}}(\boldsymbol{G}) \qquad [4]$$

where $\tilde{v}$ are the Fourier components (at wave vectors $\boldsymbol{G}'' = \boldsymbol{G} - \boldsymbol{G}'$) of the spatially periodic potential due to the nuclei plus the screening potential $\hat{v}_{\text{screen}}$ due both to the core and to the valence electrons. In practice, when a numerical solution is being sought, the plane-wave expansion must be truncated so that only a finite number N_{pw} of $\boldsymbol{G}$-vectors, those satisfying $|\boldsymbol{k}+\boldsymbol{G}| < G_{\text{cut}}$, are included in the sum of eqn [3]. Here is where the elimination of cores and the introduction of pseudopotentials become crucial. If the smallest length scale in the electronic wave functions is λ_{min}, one needs at least $G_{\text{cut}} \sim 2\pi/\lambda_{\text{min}}$, and, in the expansion of eqn [3], the number of plane waves $\boldsymbol{k}+\boldsymbol{G}$ included in a sphere of radius G_{cut} is $N_{\text{pw}} \propto G_{\text{cut}}^3$. In the true, full-core atom, λ_{min} is dictated by the 1*s* core state: because of orthogonality, in the core region not only the 1*s*, but also all the higher *s* states, up to the valence shell, will have wiggles of length scale $\lambda_{\text{min}} = \lambda_{1s}$, which is much smaller than typical nearest-neighbor (NN) distances and lattice constants $\lambda_{1s} \ll d_{\text{NN}} < a$. This makes N_{pw} so large, that the $N_{\text{pw}} \times N_{\text{pw}}$ matrix corresponding to eqn [4] becomes practically intractable. (Indeed the computational effort to diagonalize an $N_{\text{pw}} \times N_{\text{pw}}$ matrix scales like N_{pw}^3 for a full diagonalization and approximately like N_{pw}^2 if only the lowest eigenvectors and eigenvalues of the eqn [4] are being sought.) If, instead, core electrons can be effectively eliminated, one is left with a crystal of valence-only, pseudo-atoms where effective ions, screened by their valence electrons alone, replace the true ions made of nuclei and core electrons. Unlike the true valence wave functions, the psuedo-wave-functions are smooth inside the core regions, since, in a pseudo-atom, the spatial wiggles due to the orthogonality to inner-core states disappear (compare the top and bottom panels of **Figure 2**); the minimum electronic length scale λ_{min} is now comparable to typical nearest-neighbor distances. So, with respect to the true crystal, a much smaller number N_{pw} of $\boldsymbol{G}$-vectors is required for an accurate expansion of the crystal wave functions (eqn [3]); the reciprocal-lattice vectors $\boldsymbol{G}'' = \boldsymbol{G} - \boldsymbol{G}'$ to be included in the crystal pseudopotential $\tilde{v} = \tilde{v}^{\text{ps}}(\boldsymbol{G}'')$ of eqn [4] are also few; a numerically affordable $N_{\text{pw}} \times N_{\text{pw}}$ saecular problem results.

Besides rationalizing the perturbative theory of the quasi-free-electron model, pseudopotentials thus

ignited some sort of gold rush in the field of the electronic states of solids: appropriate pseudopotentials for the electron-ion interaction and a basis set of a few plane waves for the expansion of the crystal wave functions were now sufficient, beyond the limits of the perturbation theory, to successfully describe the valence bands (and thus the electrical and optical properties) not only of simple metals, but also of semiconductors and insulators.

Empirical Pseudopotentials

The very argument on orthogonality between valence and core orbitals suggests that pseudopotentials should, in principle, be derived from atomic wave functions.

Luckily, it was clear from the very beginning that, for quite a number of elements, a local (i.e., angular-momentum independent, see later) pseudopotential, given in terms of a few adjustable parameters, could provide an acceptable approximation of the true electron-ion interaction. This was true for many solids, some of which were of paramount technological importance: alkali metals, simple metals such as aluminum, and also elemental and compound semiconductors. Because of this simplification (and also of the difficulties of fully self-consistent calculations in those years), much of the early, successful pseudopotentials were obtained from fits to experimental data and not from atomic wave functions. A very popular empirical procedure was, for example, to (1) directly address a valence-only problem for the crystal: in eqn [4], the core bands are thrown away from the very start; (2) forget about self-consistency, and rather (3) treat as disposable parameters, the (by now few) Fourier components $\tilde{v}^{\mathrm{ps}}(\boldsymbol{G}-\boldsymbol{G}')$ of the total crystal pseudopotential, which replace the true potential $\tilde{v}(\boldsymbol{G}-\boldsymbol{G}')$ in eqn [4], and vary them until the experimental optical absorption spectrum, the effective mass, and/or other measured quantities were optimally reproduced. In less than ten years, these (or similar) empirical pseudopotentials gave a substantial contribution to the basic understanding of many crystalline solids. In particular, they proved the ideal theoretical tool for bulk semiconductors, whose revolutionary impact on technology happened to boom precisely in those years; alternative band-structure approaches, excellent for close-packed solids and transition metals, failed to describe the highly directional covalent bonds of semiconductors.

Model Pseudopotentials

The success on ideal, periodic bulk solids fueled more ambitious tasks of great fundamental and technological interest: the description of electrons at lattice imperfections of solids, such as surfaces, vacancies, interfaces.

These imperfections imply a much lower spatial symmetry than perfect crystals, not only because they break the translational symmetry, but also because, as a consequence of their presence, a number of surrounding atoms may be displaced with respect to their position in the corresponding ideal solid. Periodicity, and thus Bloch's theorem and a plane-wave basis, may still be exploited: one may always treat an arbitrarily complicated, low-symmetry collection of atoms as the (large) unit cell, or supercell, of an artificial crystal. But the early applications to imperfect solids showed that empirical pseudopotentials, tailored to describe perfect crystals, were not transferable to low-symmetry systems: the self-consistent rearrangement of valence electrons around imperfections, with a reliable dependence of their ground-state energy on atomic positions (needed to determine equilibrium atomic displacements), were simply beyond their scope.

New, transferable pseudopotentials were needed for this purpose. In other words, potentials which adequately represent the interaction between ions and valence electrons not just in a given perfect crystal but in many different chemical or solid-state surroundings (isolated atom, molecule, ideal, or imperfect crystal). This brought many theorists back to pseudopotential definitions based on the isolated atom, and to the distinction between a bare ionic pseudopotential $v^{\mathrm{ps,\ bare}}$ which replaces the true ion (nucleus plus core electrons) and never changes, and valence electrons, which are allowed to self-consistently rearrange (i.e., form different bonds in response to different environments). Since by electrostatic arguments, outside the atomic core, such a bare pseudopotential must have a long-range, attractive Coulombic tail $-Z_{\mathrm{v}}/r$ (where r is the distance from the nucleus and Z_{v} is the valence charge, equal to the number of valence electrons in the neutral atom), the simplest bare pseudopotentials were of the form:

$$v^{\mathrm{ps,bare}} = \begin{cases} V_0, & r < r_{\mathrm{c}} \\ -Z_{\mathrm{v}}/r, & r > r_{\mathrm{c}} \end{cases} \qquad [5]$$

The only two parameters appearing in the above model (sometimes reduced to just one, the core radius r_{c}: see **Figure 3** and caption) were determined either from the isolated self-consistent atom or in other semi-empirical ways. Of course fancier models, with more parameters, were also conceived.

To use these bare pseudopotentials in, say, a molecule or a solid, one had then to screen them self-consistently with their valence electrons. For this purpose, the most popular choice within the

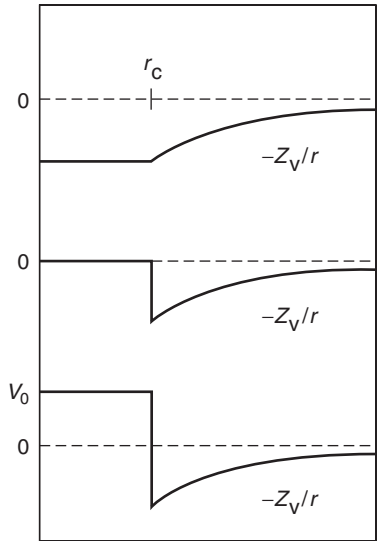


Figure 3 Model bare pseudopotentials. Lower panel: two-parameter pseudopotential. Upper and middle panels: one-parameter pseudopotentials, where the only parameter is the core radius r_c, roughly corresponding to the spatial extent of the outermost core orbitals, or to the outermost radial node of the true valence orbitals (see **Figure 2** and text).

solid-state community was the density-functional theory by Hohenberg, Kohn, and Sham, which, in 1964–65, had opened a historical avenue toward the calculation of structural and electronic properties of atoms, molecules, and extended systems. Here the screening potential (which appears, e.g., in eqn [1]) is $\tilde{v}_{\text{screen}} = v_{\text{H}} + v_{\text{xc}} = v_{\text{screen}}(\boldsymbol{r})$, where v_{H} is the Hartree potential (eqn [2]) and v_{xc} is another functional of the electronic charge density, which includes, in principle, all the exchange and correlation contributions due to the quantum many-body nature of the electronic interacting system. In practice, the functional v_{xc} was given by the (very fortunate) local density approximation, which adopts for v_{xc} an analytic expression derived from the homogeneous electron gas. This approximately yields a cube root law $v_{\text{xc}}(\boldsymbol{r}) \propto n^{1/3}(\boldsymbol{r})$ and thus, unlike the Hartree potential (eqn [2]), v_{xc} has a nonlinear dependence on the charge density: $v_{\text{xc}}[n_1(\boldsymbol{r}) + n_2(\boldsymbol{r})] \neq v_{\text{xc}}[n_1(\boldsymbol{r})] + v_{\text{xc}}[n_2(\boldsymbol{r})]$, a point which will be recalled later on. More accurate approximations of this functional are currently used. The density-functional approach has the additional bonus of a straightforward evaluation of the ground-state energy (yet another functional of the electronic density).

Molecular or solid-state systems were thus described in terms of model ions, sitting at the atomic positions, each represented by a model bare-ion pseudopotential like eqn [5], and valence electrons, self-consistently described with the density-functional theory. This approach gave important insight into the charge rearrangement around, for example, surfaces or defects, but definitely failed to provide reliable ground-state energies, and thus useful predictions on equilibrium atomic geometries. In the case of silicon, for example, it completely missed the experimentally stable diamond structure. Since, in the same decade, parallel applications of the density-functional theory to true, full-core systems (like, e.g., atoms and molecules) gave, instead, very reasonable energies, it appeared very likely that the bug was in the definition of the model pseudopotential. It was shown that model pseudopotentials not only gave poor ground-state energies for crystals, but already failed to faithfully reproduce the valence charge density of an isolated atom outside the core region; an increasing attention was thus paid to the tails of the pseudo-wave-functions in the isolated atom, with the goal of a completely coherent, first-principles definition of the pseudopotential, exclusively based on atomic wave-functions, within a given self-consistent scheme (usually the density-functional theory). The goal appeared very difficult and all the key concepts and approximations involved in the pseudopotential theory were critically revisited: is it at all possible that a smooth, local potential faithfully reproduces the effect of the true, full-core ion in a variety of different physical and chemical situations?

Nonlocal Pseudopotentials

From the original Phillips–Kleinman paper, it was actually clear that the radial orthogonality condition will, in general, produce angular-momentum dependent pseudopotentials. In boron, for example, the pseudo-valence-orbital of *s* symmetry feels an effective repulsive potential (**Figure 2**, bottom left) which is reminiscent of the orthogonality between the true 2*s* valence orbital and the 1*s* core orbital (top left). This is not the case, however, for the 2*p* valence orbitals (top right): for them, no effective repulsion is required in the pseudo-atom, because 2*p* is the lowest possible *p* state and, already in the true boron atom, its orbital has no radial nodes: at variance with the 2*s* states (top left), the 2*p* states (top right) have no underlying *p*-like core to be orthogonal to. In other words, if a pseudo-boron ion must accurately replace a true boron ion (nucleus plus core electrons), then, in the core region, the corresponding pseudopotential must be repulsive for *s* valence electrons but attractive for *p* valence electrons. The same situation holds not only for boron, but also for the entire first row of the periodic table (*sp* elements such as carbon, oxygen, or nitrogen), and, for similar reasons, for the first period of *d*- and *f*-transition elements. But the *sp* angular momentum dependence is not exclusive of

first-row atoms: second-row and heavier elements, which have both *s* and *p* electrons in the core, will have a similar, though less dramatic effect. For example, the silicon ion (not shown) has two *s*-shells but only one *p*-shell in the core, so its *s* valence electrons effectively feel a more repulsive potential than its *p* electrons. Some angular-momentum dependence may, in fact, be required for an accurate description of any element of the periodic table.

Mathematically, such angular-momentum dependent potentials are nonlocal potentials $\hat{v}^{\mathrm{NL}}$, whose action on a wave function ψ may be written, in real space, as

$$(\hat{v}^{\mathrm{NL}}\psi)_{\boldsymbol{r}} = \int \mathrm{d}\boldsymbol{r}' K(\boldsymbol{r}, \boldsymbol{r}')\psi(\boldsymbol{r}') \qquad [6]$$

where the kernel, for an atom sitting in the origin, is

$$K(\boldsymbol{r}, \boldsymbol{r}') = \sum_{lm} Y^*_{lm}(\hat{\boldsymbol{r}}) v_l(\boldsymbol{r}) \frac{\delta(r - r')}{rr'} Y_{lm}(\hat{\boldsymbol{r}}') \qquad [7]$$

and the angular-momentm projector selects the component of $\psi(\boldsymbol{r})$ onto which the corresponding $v_l(r)$ applies. An *s* electron will, for example, feel a radial potential $v_{l=0}$ (**Figure 2**, bottom left, black solid line) and a *p* electron a different potential $v_{l=1}$ (bottom right). This is sometimes called a "semi-local" potential, since it is nonlocal with respect to the direction $\hat{\boldsymbol{r}}$ (the angular coordinates θ, ϕ), but local with respect to the radial coordinate r (the electron's absolute distance from the nucleus). Nonlocal pseudopotentials of this kind attracted interest because, as mentioned, neither the local, empirical pseudopotentials (which had been successful in the basic understanding of bulk metals and semiconductors), nor the local, semi-empirical model potentials described in the previous section, could predict reliable structural properties and ground-state energies. Many groups reconsidered the possible relevance of nonlocal pseudopotentials and went back to the Phillips–Kleinman argument, but even an exact orthogonalization to the core atomic orbitals failed to produce a satisfactory solution of the transferability problem.

Norm-Conserving Pseudopotentials

The breakthrough came from a (by now classical) paper of Hamann, Schlüter, and Chiang, which appeared on *Physical Review Letters* in 1979, closely followed by a very similar, but independent paper of Kerker on the *Journal of Physics C*. They showed how to incorporate, from first-principles, the pseudopotential scheme within the density-functional theory. The key property of a transferable ionic pseudopotential, they said, is to exactly reproduce for each angular momentum, the corresponding valence wave function outside the atomic core. To obtain those exact tails, each valence state of different angular momentum must be subject to a different radial potential, as in eqns [6] and [7].

Why are wave function tails so important? Qualitatively, it is clear that they are the key ingredient for the formation of chemical bonds with neighboring atoms, and thus for accurate charge densities and ground-state energies. By the Gauss theorem, if the tails of the (normalized) pseudo-wave-functions depart from the true tails, then also the electrostatics of the pseudo-atom will inevitably fail to reproduce the true one. But the key observation was the finding that an exact reproduction of the tails ensures transferability: if the tails are right for the isolated atom, then even in a different chemical environment the pseudo-atom will behave like the true one.

Let the original Hamann–Schlüter–Chiang recipe be briefly recalled. First, one chooses a reference electronic configuration (say, the isolated neutral atom) for which the pseudopotential is constructed. For the true atom, the self-consistent radial equations, corresponding to eqn [1] with $\chi_{nl}(r) = r\,R_{nl}(r)$, are

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)}{2r^2} - \frac{Z}{r} + v_{\mathrm{screen}}\right]\chi_{nl} = \varepsilon_{nl}\chi_{nl} \qquad [8]$$

where $v_{\mathrm{screen}}(r) = v_{\mathrm{H}}[n^{\mathrm{v}} + n^{\mathrm{c}}] + v_{\mathrm{xc}}[n^{\mathrm{v}} + n^{\mathrm{c}}]$, and n^{v}, n^{c} are the charge densities associated to the occupied valence and core orbitals, respectively. One would now want to construct a pseudo-atom which, instead, satisfies

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)}{2r^2} + v_l^{\mathrm{ps,screened}}\right]\chi_l^{\mathrm{ps}} = \varepsilon_l\chi_l^{\mathrm{ps}} \qquad [9]$$

with ε_l identical to the valence ε_{nl} of eqn [8]. In eqn [9], the principal quantum number n can be dropped since the pseudo-atom has only valence states, and, for each l, there is just one valence radial pseudo-wave-function. The presence of a different radial potential $v_l^{\mathrm{ps,screened}}$ for each l means that this construction allows for a nonlocal potential of the type just discussed. Such a construction starts from the end: one first defines the desired pseudo-wave-functions and then, by inverting the eqn [9], the pseudopotential.

In the first step, for each l, an optimal pseudo-wave-function $\chi_l^{\mathrm{ps}}(r)$ is constructed by exactly matching, at an appropriate core radius r_{c}, the (numerically available) tail of the corresponding true wave function with some smooth, analytical, nodeless radial function defined between $r = 0$ and $r = r_{\mathrm{c}}$, such that the resulting pseudo-wave-function is still correctly normalized to one. For each l, this implies by construction

that the partial norm up to r_c of the true and pseudo-wave-functions is identical,

$$\int_0^{r_c} \mathrm{d}r[\chi_l^{\mathrm{ps}}(r)]^2 = \int_0^{r_c} \mathrm{d}r[\chi_{nl}(r)]^2 \qquad [10]$$

since both radial orbitals are normalized to one. Because of this property, the new generation of first-principles pseudopotentials was called norm-conserving. The norm conservation, eqn [10], not only guarantees the correct electrostatics for the chosen reference atom; its mathematics also ensures, as shown in the original paper, that the scattering properties (hence the bonding properties) of the true and pseudo-atoms closely follow each other over an energy range of $\sim 20\,\mathrm{eV}$ around ε_l; this, as confirmed by a wealth of subsequent applications, amounts to an optimum transferability. Incidentally, the remarkable consequences of eqn [10] shed new light on the links between the pseudopotential method and other electronic-structure methods.

In the second step, once such a good $\chi_l^{\mathrm{ps}}(r)$ has been built, eqn [9] may be inverted to yield the corresponding $v_l^{\mathrm{ps,screened}}$, the screened ionic pseudopotential appropriate to the angular momentum l. This radial function is shown for $l = 0$ (left) and $l = 1$ (right) as a solid black line in the bottom panels of **Figure 2**.

From $v_l^{\mathrm{ps,screened}}$, in the third and final step, the bare ionic potential $v_l^{\mathrm{ps,bare}}$ is extracted with a simple unscreening prescription.

$$v_l^{\mathrm{ps,bare}}(r) = v_l^{\mathrm{ps,screened}}(r) - v_{\mathrm{H}}[n^{\mathrm{ps},v}] - v_{\mathrm{xc}}[n^{\mathrm{ps},v}] \qquad [11]$$

where $n^{\mathrm{ps},v}$ is the pseudo-charge-density of the (just constructed) valence orbitals appearing in eqn [9]. This bare ionic pseudopotential is the actual ingredient of all subsequent calculations; its pseudo-valence-electrons will be allowed to self-consistently rearrange and produce different screening effects in response to different molecular and solid-state environments.

Pseudopotentials That Work

The Hamann–Schlüter–Chiang prescription proved extremely successful for many elements of the periodic table: when these nonlocal pseudo-ions were fed into the scheme based on plane-wave basis functions proposed in the same years by Ihm, Zunger, and Cohen, the structural energies predicted by their self-consistent screening suddenly came out in excellent agreement with experiments. Not only was the diamond structure finally obtained as the lowest-energy crystal structure for silicon, but also its bulk modulus and phonon frequencies, which involve higher derivatives of the energy with respect to the atomic positions, were very accurately reproduced from first-principles, without the need of any adjustable parameter. This suggested a fully reliable theoretical framework, at least for materials where electron correlations are not too strong and the current approximations of the exchange-correlation density functional are satisfactory for the corresponding full-core systems. A second "pseudopotential gold rush" took place, this time aimed at the prediction of unknown equilibrium geometries of surfaces, defects, dislocations, and also at phonon spectra and at the stability of new compounds.

An additional reason for this unprecedented success was a remarkable advantage of plane waves over any set of localized basis functions which – like atomic orbitals, gaussians, muffin-tin orbitals – are attached to atoms and follow their positions. With plane waves, which are independent of the atomic positions in the crystalline unit cell, interatomic forces have a particularly simple expression, related to the Hellmann–Feynman theorem. Forces are thus another almost automatic, accurate byproduct of a density-functional self-consistent calculation based on plane waves and pseudopotentials.

This advantage proved precious in the determination of equilibrium geometries of unknown low-symmetry structures (such as the 2×1 reconstruction of the silicon surface established in those years by Kosal Pandey) or phonon force constants, as shown by Martin and Kunc; but its vital importance became apparent when Car and Parrinello, in 1985, invented a method which combines molecular-dynamics simulations and the density-functional theory. This idea opened to pseudopotentials a completely new field: the real-time microscopic dynamics of the constituent atoms in covalent and metallic systems, or even in the context of chemical reactions.

Other outstanding examples of the combined power of plane waves and pseudopotentials are the linear-response perturbation theory elaborated in 1987 by Baroni, Gianozzi, and Testa, which allows the direct calculation of the full phonon dispersion curves as well as the dielectric and piezoelectric tensors, and the Berry-phase approach to the spontaneous polarization of ferroelectric materials, developed by Vanderbilt and Resta in 1993.

The improved predictive power of contemporary exchange-correlation functionals, which are superior to the original local density approximation, and also some key developments of the original Hamann–Schlüter–Chiang formulation, presented in the appendix, have made the first-principles pseudopotential method a standard working tool for the investigation of real materials and relatively complex molecular systems not only for theorists, but also for

experimentalists, both in the field of solid-state physics and chemistry. The corresponding computer programs, which in 1980 used to be the privilege of a few pioneering groups and would require large-scale computational facilities, are today available in the public domain and may be executed by an undergraduate student on a personal computer.

Appendix: Key Ingredients of Contemporary Pseudopotentials

In the previous section, for the sake of clarity and simplicity, norm-conserving pseudopotentials were presented in their plain, original version. However, the predictive power and computational feasibility reached by contemporary first-principles pseudopotential methods, mentioned at the end of the previous section, would never have been possible without a number of significant improvements and innovations over the original Hamann–Schlüter–Chiang scheme. They are briefly reviewed in this last section.

Nonlinear Core Correction

Shortly after the discovery of norm-conserving pseudopotentials, Louie, Froyen, and Cohen observed that the original unscreening procedure is the simplest, but not always the best one. As a matter of fact, it implicitly assumes that, in the true atom (eqn [8]), the screening potentials depend linearly on the electronic charge density. If it were so, then there would be no error when the screening potential due to the valence charge n^{v} alone, subtracted out at the atomic stage (eqn [11]), is self-consistently added back to the bare pseudopotential in the context of a different molecular or solid-state situation. Unfortunately, such a linear behavior exactly holds for the Hartree potential $v_{\mathrm{H}}[n]$, which, as evident from eqn [2], is a linear functional of the charge density $n(r)$ (i.e., $v_{\mathrm{H}}[n^{\mathrm{v}}+n^{\mathrm{c}}]=v_{\mathrm{H}}[n^{\mathrm{v}}]+v_{\mathrm{H}}[n^{\mathrm{c}}]$), but, as already mentioned, does not hold for the exchange-correlation functional $v_{\mathrm{xc}}[n]$, which is a strongly nonlinear functional of its argument.

Fortunately, core and valence charge densities are often well separated, and thus scarcely overlapping, and this nonlinearity may be safely disregarded: whenever n^{v} rapidly vanishes inside the core and n^{c} rapidly vanishes outside the core, then, apart from a small overlap region, the approximation $v_{\mathrm{xc}}[n^{\mathrm{v}}+n^{\mathrm{c}}]\simeq v_{\mathrm{xc}}[n^{\mathrm{c}}]+v_{\mathrm{xc}}[n^{\mathrm{v}}]$, implicit in the unscreening procedure (eqn [11]), is a good one. This is why the original, simple unscreening procedure works well for many elements of the periodic table – those with well-separated core and valence shells.

But quite a few elements of the periodic table do not share the clearcut core-valence separation (both in real space and in energy) of the examples shown in **Figures 1** and **2**. When the underlying core is shallow, as it happens in most alkali elements, or when two different principal quantum numbers are present in the valence shell (like *d*- or *f*-transition elements), and also in a few other cases, the core-valence overlap may be remarkable, and the nonlinearity of $v_{\mathrm{xc}}[n]$ can seriously hinder the main virtue of norm-conserving pseudopotentials: their transferability. For these atoms, such a desirable virtue is recovered by an alternative unscreening prescription proposed by Louie and co-workers:

$$v_l^{\mathrm{ps,bare}}(r)=v_l^{\mathrm{ps,screened}}(r)-v_{\mathrm{H}}[n^{\mathrm{ps,v}}]-v_{\mathrm{xc}}[n^{\mathrm{ps,v}}+n^{\mathrm{c}}]$$

In this way, the core charge density $n^{\mathrm{c}}(r)$ of the true atom (or an appropriate model thereof) becomes an ingredient of the pseudo-ion, and, although frozen in the reference atomic state, must be carried along with the bare-ion pseudopotential $v_l^{\mathrm{ps,bare}}$ and used in all subsequent molecular or solid-state calculations; this is done by adding it back to the pseudo-valence-charge, of course, within the exchange-correlation potential only. Such a nonlinear core correction restores the nonlinear character of the interaction between a valence electron and the ionic core, and proved to be successful in all the problematic cases mentioned above, for which it immediately became the standard solution.

Relativistic Effects and Separable Pseudopotentials

Two other important developments of modern first-principles pseudopotentials involve one of the fathers of the pseudopotential theory, Leonard Kleinman.

One concerns the possibility of including in the pseudopotential relativistic effects to order α^2, where $\alpha\simeq 1/137$ is the fine-structure constant. Kleinman showed that the Hamann–Schlüter–Chiang construction can be generalized to the relativistic case, and that the resulting pseudopotential can be conveniently used to treat heavy atoms in a nonrelativistic formalism. The reason for this is that relativistic effects on the valence electrons are limited to the core region (only there the valence radial wave functions of the true atom have wiggles, and thus a very high kinetic energy), and can be lumped together with other properties of the core to yield a Schrödinger-like pseudo-atom. This development opened the way to the construction of norm-conserving pseudopotentials for the whole periodic table, from hydrogen to plutonium.

The second development, due to Kleinman and Bylander, amounts to an approximate transformation of the kernel K from its original semilocal form,

given by eqn [7], into a fully nonlocal, but separable kernel $\tilde{K}$ of the form

$$\tilde{K}(\boldsymbol{r},\boldsymbol{r}') = \sum_{lm}\sum_{\beta} Y^*_{lm}(\hat{\boldsymbol{r}})w^*_{\beta l}(r)w_{\beta l}(r')Y_{lm}(\hat{\boldsymbol{r}}') \quad [12]$$

This brings no advantage to the transferability of the pseudopotential, but there is a consistent computational advantage over the original semilocal form. Indeed, the semilocal form requires, to be calculated and stored in $\boldsymbol{G}$-space, an $N_{\mathrm{pw}} \times N_{\mathrm{pw}}$ matrix (if N_{pw} is the number of plane-wave basis functions), whereas a separable pseudopotential can be factorized as a product of two N_{pw} dimensional vectors. For this reason, the Kleinman–Bylander form is now included in all the computer programs dealing with pseudopotentials.

Ultrasoft Pseudopotentials for Hard Atoms

The success of pseudopotentials is largely related to the fact that pseudo-wave-functions can be expanded within a tractable number of plane waves. Not surprisingly, pseudo-atoms with core radii r_{c} imply a plane-wave cutoff $G_{\mathrm{cut}} \sim 2\pi/r_{\mathrm{c}}$ and thus, if the core radius adopted in the pseudopotential construction is too small (compared to typical interatomic distances), then the corresponding pseudopotential is too hard, in the sense that the number of plane waves required to describe its norm-conserving wave functions (eqn [10]) becomes unacceptably large. This happens for the elements with one radially nodeless orbital ($2p$, $3d$, $4f$) in the valence shell. For these valence states, a norm-conserving pseudo-valence-orbital cannot be too different from the original true orbital: near $r = 0$, they both have to go like r^{l+1}; beyond r_{c}, they are identical by construction; between 0 and r_{c}, they are both nodeless and, by eqn [10], they must have the same partial norm. Under these circumstances, it is easy to see that the pseudopotential resulting from the inversion of the Schrödinger eqn [9] will inevitably have strong spatial variations in the small interval between 0 and r_{c}, and thus will be hard, requiring too many plane waves in the corresponding wave functions, as it happens, for example, in the case of the boron p pseudopotential, shown in the bottom right panel of **Figure 2**.

In order to obtain manageable pseudopotentials for these atoms, Vanderbilt proposed a radical departure from the Hamann–Schlüter–Chiang procedure: a more complicated pseudopotential scheme (based on generalized orthonormality conditions) which, by releasing the condition of norm conservation, allowed the generation of a new family of very smooth (ultrasoft), yet reliable and transferable pseudopotentials. In this scheme, the pseudo-wave-functions are allowed to be as soft as possible within the core region, so that the plane-wave cutoff G_{cut} can be reduced dramatically. The scheme works very well even for the worst cases, the hard atoms mentioned above: when applied to the study of compounds involving transition metals and rare-earth atoms as well as nitrides and oxides, the scheme has proved accurate and reliable, and has been adopted and implemented by many groups, in spite of its greater conceptual and numerical complexity.

See also: Crystal Structure; Density-Functional Theory; Electronic Structure (Theory): Atoms; Electronic Structure Calculations: Plane-Wave Methods; Electronic Structure Calculations: Scattering Methods; Hartree and Hartree–Fock Methods in Electronic Structure; Molecular Dynamics Calculations; Periodicity and Lattices.

PACS: 71.15.–m; 71.15.Ap; 71.15.Dx; 71.15.Mb; 71.15.Nc; 71.15.Pd; 71.15.Rf

Further Reading

Ashcroft NW and Mermin D (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.

Bachelet GB, Hamann DR, and Schlüter M (1982) Pseudopotentials that work: from hydrogen to plutonium. *Physical Review B* 26: 4199.

Baroni S, Dal Corso A, de Gironcoli S, and Giannozzi P (2001) Phonons and related crystal properties from density-functional perturbation theory. *Review Modern Physics* 73: 515.

Bassani F and Pastori Parravicini G (1974) *Electronic State and Optical Transitions in Solids*. New York: Pergamon.

Cohen ML, Heine V, and Weaire D (1970). In: Ehrenreich H, Seitz F, and Turnbull D (eds.), *Solid State Physics*, vol. 24. New York: Academic Press.

Payne MC, Teter MP, Allan DC, Arias TA, and Joannopoulos JD (1992) Iterative minimization techniques for *ab-initio* total energy calculations: molecular dynamics and conjugate gradients. *Review Modern Physics* 64: 1045.

Singh D (1994) *Plane Waves, Pseudopotentials, and the LAPW Method*. Boston: Kluwer.

Pyroelectricity

P Muralt, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

Introduction

Pyroelectricity provides one of the best performing principles for the detection of temperature changes. This occurs in polar matter and results in the appearance of surface charges proportional to a temperature change ΔT. The basic structure is a planar capacitor whose charge Q varies according to $\Delta Q = Ap\Delta T$, where A is the area of the capacitor and p the pyroelectric coefficient. A pyroelectric sensor thus directly generates an electric signal, facilitating the integration into electronic circuits. Pyroelectric crystals, ceramics of ferroelectric materials, as well as polymers have therefore been used since the 1960s in thermal infrared (IR) detectors, joining the earlier thermal IR detection techniques of bolometers and thermopiles. Since the late 1980s, pyroelectrics have also been investigated in the form of thin films and integrated onto micromachined structures. Pyroelectricity thus made its way from a scientific curiosity known since 2500 years to a very useful physical effect that, for instance, enables firemen to see through smoke.

Fundamentals

Pyroelectricity is the electrical response of a polar, dielectric material to a change in temperature. The origin of this effect lies in polar features that are lined up with the same orientation along at least one direction in the material. In polymers, the polar features are polar molecules or groups. One can thus identify microscopic dipoles. In inorganic materials it is often not possible to identify dipoles. It is rather the anisotropy and anharmonicity of chemical bonding that results in the center of gravity of the positive charges moving differently than the center of gravity of the negative charges upon changing the temperature. The case is illustrated schematically in **Figure 1** for a periodic one-dimensional structure. Polarization of periodic structures is a delicate issue and needs careful discussion. The essential point is that the spring constants f_1 and f_2 are different and vary differently with temperature. The macroscopic dipole originating from the way the chain is cut is of no importance because (1) this cut does not change with temperature, (2) each chain is cut in a different way and the net charge is averaged out, and (3) there is anyhow charge compensation with time (leakage). Changing the temperature leads to a relative displacement of the positive ions with respect to the negative ones. In **Figure 1**, the spring constant f_2 hardens more than f_1 upon cooling and the positive ions are shifted toward the negative ones to the right. As a result, the net charges on the electrodes are increased positively on the right, and negatively on the left. These are the pyroelectric charges. An equivalent effect results if negative charges (electrons) from the negative ions move toward the positive ions by a charge transfer process. More insight into such mechanisms of covalent nature in ionic crystals have recently been gained by first-principle calculations.

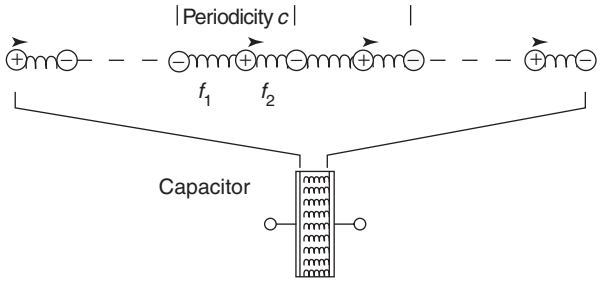


Figure 1 One-dimensional model of a periodic polar material in a capacitor structure.

In crystalline matter, pyroelectricity occurs in all materials with symmetries that allow the existence of a polar direction. No symmetry element may exist in the point group that inverts or rotates this direction by 180°. Such materials are also piezoelectric, i.e., they respond with an electric field to a change of the shape. From the 20 piezoelectric point groups, 10 are polar and allow for pyroelectricity. The pyroelectric effect is anisotropic and depends on the directions of the electroded faces with respect to the crystal axes: $p = (p_x, p_y, p_z)$. The highest pyroelectric coefficients are measured in ferroelectrics. In such substances there is a spontaneous electric polarization P_s that can be inverted by the application of an external electric field, and consequently is a measurable physical quantity. Upon heating up, P_s diminishes and becomes zero at the critical temperature of the ferroelectric–paraelectric phase transition as depicted in **Figure 2** for a second-order phase transition. The symmetry group of the ferroelectric phase is a subgroup of the symmetry group of the paraelectric phase, that is, a number of symmetry elements are lost (certainly the inversion element, if present in the paraelectric phase). The direction of P_s is not uniquely defined. Turning an existing solution by

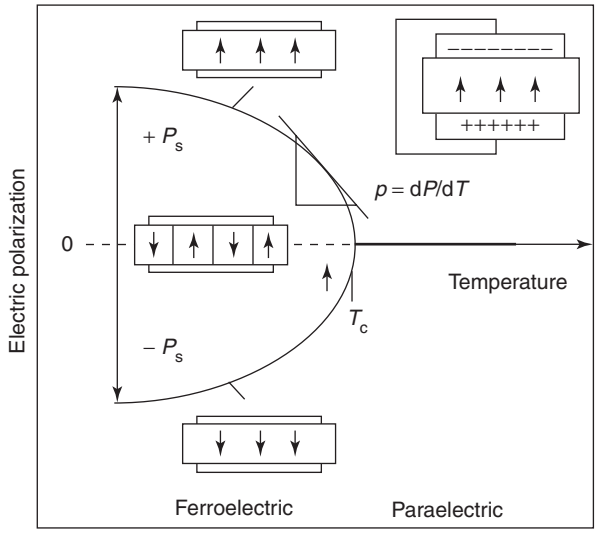


Figure 2 Schematic graph showing the bifurcation (second-order phase transition) of the electrical polarization at the critical temperature from $P = 0$ (paraelectric phase) to the two equally probable solutions $+/-$ P_s for the spontaneous polarization in the ferroelectric phase. The inserted capacitors show extreme cases of 180° domain configurations.

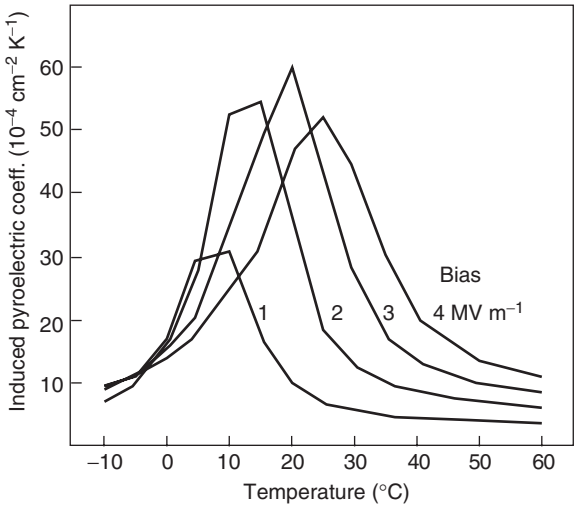


Figure 3 Experimental curves for the induced pyroelectric coefficient in lead scandium tantalate. (Courtesy of R Watton and MA Todd.)

180° (applying the lost inversion, mirror or rotation symmetry) yields always another, equally probable solution. Regions with homogeneous polarization are called domains. For achieving a macroscopic polarization, a poling process in an external electric field – often combined with heating above room temperature – must be applied to eliminate the 180° domains and to also possibly switch ferroelastic domains. The macroscopic orientation obtained after removal of the external field is called remanent polarization P_r.

The charges on the electrodes are given by the electric displacement field D perpendicular to the electrode faces. Giving the index 3 to this direction, the pyroelectric coefficient is equal to $\partial D_3/\partial T$. The electric displacement field in ferroelectrics can be written as $\boldsymbol{D}(\boldsymbol{E}) = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}(\boldsymbol{E})$, where ε_0 means the permittivity of the vacuum. The polarization part is much larger and the only temperature-dependent term. For the present discussion, one can set D equal to P

$$\Delta D_3 = \frac{\partial P_3(E)}{\partial T}\,\Delta T \cong \frac{\partial P_{r,3}}{\partial T}\,\Delta T + \frac{\partial \varepsilon_{33}}{\partial T}\,E_3 \Delta T \quad [1]$$

The first term is called the true pyroelectric effect and $p = \partial P_{r,3}/\partial T$ is the true pyroelectric coefficient. The second term arises from the first-order correction for $P_r(E)$. Pyroelectricity, arising due to this second term, is called induced pyroelectricity. Accordingly, there are two modes of operations for pyroelectric IR detectors. The first one works without an external electric field at a temperature much below the para-to-ferroelectric phase transition. The polarization P_r has previously been maximized by hot poling at an elevated electric field, assuring a good time stability of p at the application temperature around room temperature. The second mode exploits the peaking of p near the phase transition. A DC electric field needs to be applied to obtain a nonzero average polarization. One speaks of induced pyroelectric currents or sometimes of dielectric bolometer mode, since the field dependence of P can be considered as a dielectric property. The applied field has not only the effect of eliminating 180° domains, it broadens and shifts (first-order transition only) the phase transition to higher temperatures. This leads to a reduction of the signal. There is thus an optimal DC field (a few MV m^{-1}) and temperature (see **Figure 3**) at which the response is peaking. The temperature of such a device needs to be stabilized. The advantage of the first method is the simplicity of operation conditions: no DC field and no temperature stabilization are needed. The second method allows higher sensitivities.

In the following the phenomenological theory of Landau, Ginzburg, and Devonshire (LDG) is used to gain some better insight into the difference of true and induced pyroelectricity and to prepare the ground for discussing materials figure of merits. The free energy density G of the ferroelectric phase is written as

$$\begin{aligned} G(T,P) &= \tfrac{1}{2}\,\beta(T - T_0)P^2 + \tfrac{1}{4}\,\gamma P^4 + \tfrac{1}{6}\,\delta P^6 \\ E &= \beta(T - T_0)P + \gamma P^3 + \delta P^5 \end{aligned} \quad [2]$$

The derivation with respect to P yields the electric field E. The lowest energy at $E = 0$ yields the spontaneous polarization P_s. The inverse permittivity due

to the ferroelectric effect is equal to the second derivative of G with respect to P. A general relation is found for the true pyroelectric coefficient as follows: $p = -\beta\varepsilon P_s$. High pyroelectric coefficients are thus found in materials with high polarization and high coefficient β. The true pyroelectric voltage response, which is proportional to p/ε, follows the spontaneous polarization and is thus optimal at low temperatures, far below the phase transition. In case of a second-order phase transition ($\gamma > 0$), and neglecting the sixth-order term ($\delta = 0$), the spontaneous polarization is readily obtained as $P_s = ((\beta/\gamma)(T_0 - T))^{1/2}$. The true pyroelectric coefficient is derived as $p = (-1/2)(\beta/(\gamma(T_0 - T)))^{1/2}$. The pyroelectric coefficient diverges at the phase-transition temperature T_0. So does the permittivity, for which below the critical temperature T_0, the relation $\varepsilon^{-1} = \partial^2 G/\partial P^2 = 2\beta(T_0 - T)$ is obtained. The pyroelectric coefficient obviously behaves as the square root of the permittivity, that is, $p/\sqrt{\varepsilon} = -\beta\sqrt{2/\gamma} = \text{const.}$ In ideal first-order transitions ($\gamma < 0$, $\delta > 0$), the critical temperature $T_c = T_0 + 3\gamma^2/(16\beta\delta)$ is above T_0 and the polarization jumps form 0 to $\sqrt{-3\gamma/3\delta}$. The permittivity remains finite but jumps as well at T_c. The jump of P_s at T_c gives rise to large pyroelectric currents. Induced pyroelectricity is thus very large in materials with first-order phase transitions. In addition, the field needed for induced pyroelectricity does not simply broaden the transition as in the case of second-order phase transitions, but the discontinuity of P is preserved up to a certain critical field. The critical temperature is shifted to higher temperatures up to a maximal possible value of $T_{max} = T_0 + (9\gamma^2)/(20\beta\delta)$. For illustration, calculated curves are depicted in **Figure 4** for a model substance.

A further type of phase transition that attracts attention for being used in the induced pyroelectric mode is the relaxor-ferroelectric transition. It occurs typically in specifically disordered ferroelectrics, where, for instance, the B-site of the perovskite ABO_3 is a mixture of randomly distributed ions of different valency. In the above-mentioned $PbSc_{0.5}Ta_{0.5}O_3$, Sc^{3+} and Ta^{5+}, ions occupy, in a 1:1 mixture, a site that should have a $4+$ ion. If the two different ions are ordered in a periodic manner, the above discussed ideal first-order phase transition takes place; if they are disordered, the ferroelectric phase is shifted to lower temperatures. Usually, very high low-frequency permittivities are obtained around the phase transition. It is thought that random internal fields prevent a homogeneous, ferroelectric polarization. Instead, local polar regions appear that easily grow in an external field thus providing for the huge dielectric response. There are chances that the absence of domain or phase

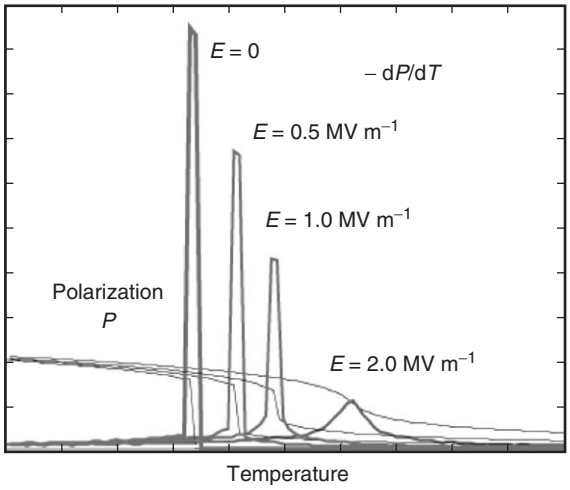


Figure 4 Theoretical behavior of the pyroelectric coefficient as a function of temperature and for various electric fields, as derived from the Landau–Devonshire model. The behavior near the phase transition in the absence of an external field is purely theoretical, since the average polarization decays to zero by thermal fluctuations.

fluctuations in the relaxor phase allow for working points with a larger dP/dT and a lower $\tan\delta$. A modeling of this issue is missing to date.

The pyroelectric charge is usually not directly measured as a charge but as a current I through a capacitor of surface A. Taking the time derivative of eqn [1]

$$I = A\frac{dD_3}{dt} = Ap\frac{dT}{dt} \qquad [3]$$

The current measurement is much less prone to drifts and also faster than charge measurements. In addition, charges are compensated with time either because of leakage currents through the pyroelectric capacitor, or by conduction through air, or along the device surface. For imaging static pictures or temperatures, a chopper is applied that modulates the infrared radiation falling onto the detector with a constant angular frequency $\omega = 2\pi f$. The frequency f amounts to 30–60 Hz for imaging applications.

Several methods have been proposed for the measurement of the pyroelectric effect of bulk or thin-film samples. The most frequently applied techniques are based on the pyroelectric current generated by a temperature ramp of constant rate (i.e., $dT/dt = \text{const.}$), yielding a constant pyroelectric current according to eqn [3]. One speaks of the Byer–Roundy technique when a single ramp is applied. It is important to check heating and cooling curves, because the release of trapped charges (thermally induced currents) might falsify the heating curve. Such trapped charges are injected into the sample during the poling procedure. An improvement of the method is

obtained by cycling the temperature periodically. The temperature is swept up and down with small amplitudes (e.g., 1 K) and with constant rates (e.g., $0.25\,\mathrm{K\,s^{-1}}$) around a given average temperature. The pyroelectric current, being proportional to the slope of the temperature, will switch between two constant values after some stabilization time. The temperature is best controlled on a small chuck equipped with Joule heating and Peltier cooling elements, placed in a chamber protecting from perturbations. For higher modulation frequencies, temperature changes are induced by chopped light absorption (Chynoweth). However, this technique is more qualitative than quantitative as it requires the determination of thermal properties until the temperature change is really known. Sometimes the polarization is measured as a function of temperature. This is rather dangerous, since the polarization must be derived from a hysteresis loop obtained by switching the polarization. First, it is not sure whether the complete polarization is switching, second fatigue could reduce polarization more efficiently than the temperature increase.

The quality of a sensor does not only depend on sensitivity, but also on the signal-to-noise ratio (S/N ratio), or in other words, the noise level. The intrinsic noise type of pyroelectrics is the Johnson noise of the capacitor element. For a capacitor of area A and thickness t, the noise current is obtained as

$$I_\mathrm{n} = \sqrt{\frac{A}{t}\,4kT(\sigma + \omega\varepsilon\,\tan\,\delta)\Delta f} \qquad [4]$$

where σ is the leakage conductivity of the pyroelectric thin film (in bulk detectors, the parallel resistor), ω the angular frequency, $\tan\,\delta$ the loss tangent, Δf the frequency bandwidth, A the surface, and t the thickness of the element. The cross over frequency $\omega_\mathrm{cn} = \sigma/(\varepsilon\,\tan\,\delta)$ between resistor type noise and dielectric noise amounts to typically 0.1–10 Hz. At usual modulation frequencies of 10–60 Hz, the dielectric noise is thus dominating. The intrinsic figure of merit of the dielectric and pyroelectric materials parameters for an optimal S/N ratio can thus be formulated as

$$\mathrm{FM} = \frac{p}{\sqrt{\varepsilon\,\tan\,\delta}} \qquad [5]$$

Various phenomena can contribute to the dielectric loss. At high frequencies (GHz and higher), the dielectric loss increases due to interactions of phonons and AC electric fields. In noncentrosymmetric materials, the dominating phonon mechanism is the quasi-Debye loss mechanism. Thermal fluctuations and relaxation of phonons lead to charge fluctuations in a similar way as dipoles give rise to loss in the Debye theory. In ferroelectrics, domain wall relaxations play a role as well. In the 1–10 GHz range, there is usually a relaxation of 180° domain wall motions. All these mechanisms yield a $\tan\,\delta$ that increases at least linearly with the dielectric constant, that is, $\tan\,\delta \propto \varepsilon^\alpha$, where α is between 1 and 4. It is difficult to make a general theory on loss mechanisms at very low frequencies of a few tens of hertz. In this region, the dielectric loss is governed by defects and ferroelectric domains. Charged or dipolar point defects relax with respect to position or orientation, structural defects may interfere with electric fields by means of the piezoelectric coupling, and defects may delay domain wall motions or give rise to a creeping of domain walls. Many of the low-frequency relaxation mechanisms show again a $\tan\,\delta$ proportional to ε. Near the critical temperature, polarization fluctuation and phonon losses (soft mode) were found to follow a Curie–Weiss law like the dielectric constant, meaning that $\tan\,\delta$ diverges as well at the critical temperature. This has, for instance, been found in 1991 by R Watton and M A Todd where the loss tangent and the dielectric constant increase by about the same factors upon approaching the critical temperature ($PbSc_{0.5}Ta_{0.5}O_3$). The materials figure of merit FM thus behaves rather like $p\varepsilon^{-1}$ than $p\varepsilon^{-1/2}$ especially near critical temperatures.

Evaluating the FM by means of the LDG theory leads to the conclusion that, in case of second-order phase transitions, there is no improvement of the intrinsic S/N ratio when going closer to the phase transition. $p\varepsilon^{-1/2}$ is, in fact, independent of temperature, and taking into account the increasing dielectric losses, the intrinsic signal to noise is expected to decrease with temperature. The case of first-order transitions is more difficult to judge. The permittivity does not diverge at the phase transition, there is only a discontinuity. Lead scandium tantalate showed an increase of FM of nearly a factor 3, hence a clear advantage for induced pyroelectricity in this case. The optimal DC field is, however, not the one at which the pyroelectric current is peaking, that is, inside the hysteretic first-order transition with mixed phase noise, but just somewhat above the critical field at which the jump in polarization is lost.

Operational Principles of Pyroelectric IR Detectors

The main application of pyroelectric materials is detection of IR radiation. The latter is measured indirectly by means of a temperature change ΔT of an absorbing structure as a result of the absorbed radiation power. The thermal properties of the absorbing structure have to be suitably designed for obtaining a

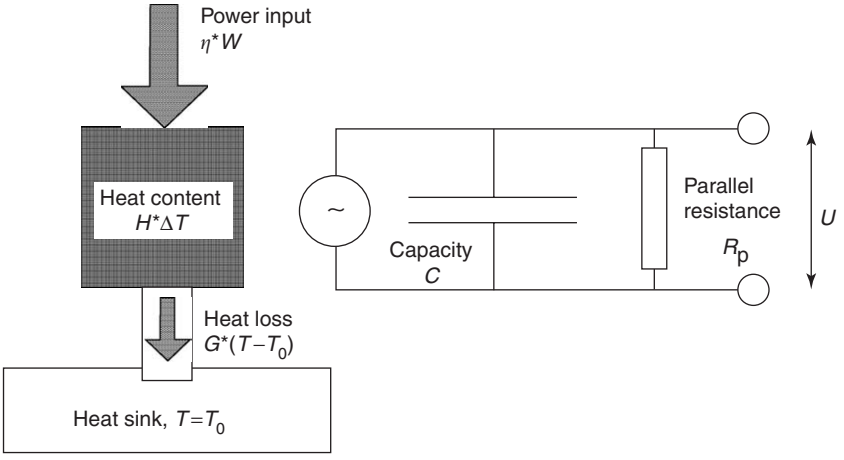


Figure 5 Schematic of thermal and electrical circuit of a pyroelectric element.

maximal temperature increase ΔT at the required response frequency. ΔT is obtained from the balance of the heat flows, as sketched in **Figure 5**. First of all, some fraction η of the IR radiation (power W) falling onto the element is absorbed, that is, transformed into heat. The temperature change at a given heat input $\eta \cdot W \cdot \Delta t$ depends on the heat conductivity G to the surrounding heat reservoir (heat sink) at temperature T_0 and the heat capacity H of the element. It is supposed that the element exhibits a uniform temperature. Thermal wavelength effects are thus neglected. This is justified in thin structures with uniform illumination by the IR radiation. (In $PbTiO_3$ with $c_p = 3.2\,\mathrm{MJ\,m^{-3}\,K^{-1}}$ and $\kappa = 3.8\,\mathrm{W\,m^{-1}\,K^{-1}}$, the thermal wavelength amounts to 600 µm at 1 Hz and 20 µm at 1 kHz.) Considering stationary solutions in the frequency domain for an IR power of $W = W_\omega \cdot e^{i\omega t}$, the temperature modulation is readily obtained as

$$\Delta T_\omega = \frac{\eta W_\omega}{G + i\omega H}, \qquad |\Delta T_\omega| = \frac{\eta W_\omega}{G\sqrt{1+\omega^2\tau_{th}^2}} \quad [6]$$

The thermal time constant $\tau_{th} = H/G$ has been introduced. There are thus two frequency domains separated by the inverse thermal time constant. The current responsivity, that is, the current response per watt of radiation power falling onto the detector element is obtained as

$$\Re_J(\omega) = \frac{J_\omega}{W_\omega} = \frac{p\eta\omega A}{G\sqrt{1+\omega^2\tau_{th}^2}} \quad [7]$$

Alternatively, the voltage across a parallel resistor R_p can be measured. In bulk devices, a value of 10 GΩ is typically applied (R_p should not exceed the gate impedance of the amplifier). In thin-film devices, leakage through the film may yield an R_p in a good value range. At higher frequencies, the dielectric loss tangent $\tan\delta$ also comes into play. The conductance Y of the complete element (thus including the parallel resistance) is obtained as

$$|Y(\omega)| = \frac{1}{R_p}\sqrt{(1+\omega\tau_{el}\tan\delta)^2 + \omega^2\tau_{el}^2} \quad [8]$$

The electrical time constant $\tau_{el} = R_p \cdot C$ has been introduced. The voltage response is proportional to the impedance $Z = Y^{-1}$. Above the angular frequency τ_{el}^{-1}, the impedance decreases as ω^{-1}. The pyroelectric element thus works like a low pass filter, cutting off the high-frequency voltage response. Omitting the small term in $\tan\delta$, the complete voltage responsivity is written as

$$\Re_V(\omega) = |Z(\omega)|\Re_J(\omega) = \frac{p\eta\omega A R_p}{G\sqrt{1+\omega^2\tau_{th}^2}\sqrt{1+\omega^2\tau_{el}^2}} \quad [9]$$

The general trends of the two responsivity functions depends on the relative size of the two time constants. In bulk devices, one generally encounters $\tau_{el} < \tau_{th}$. The inverse situation, that is, $\tau_{el} > \tau_{th}$, is typically thin-film structures. The resultant frequency behavior for thin films is schematically shown in **Figure 6**. In thin-film devices, the parallel resistance is not directly involved in the voltage response. This is very advantageous, as film leakage is not easy to control precisely. The mounting of parallel resistors can be avoided if the leakage is sufficiently large to avoid discharging through the input gate of the voltage amplifier. Observing that $C = \varepsilon A/t_p$, the response can further be developed in the intermediate region as $\Re_V \cong (\eta p t_p)/(\varepsilon G)$ in case of thin films, and $\Re_V \cong (\eta p A R_p)/H$ in the bulk case. The figure of merit given in general, $F_V = p/\varepsilon c_p$, corresponds to a

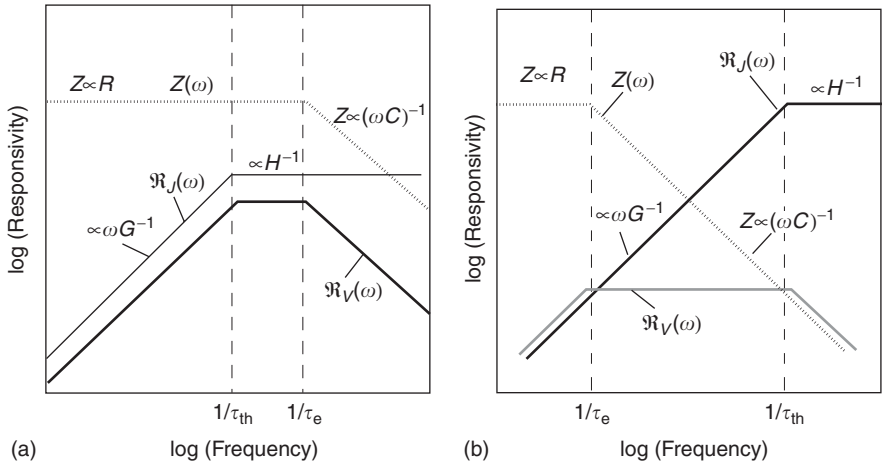


Figure 6 Log–log scheme of current and voltage response, together with the impedance for the typical bulk (a) and thin-film situation (b).

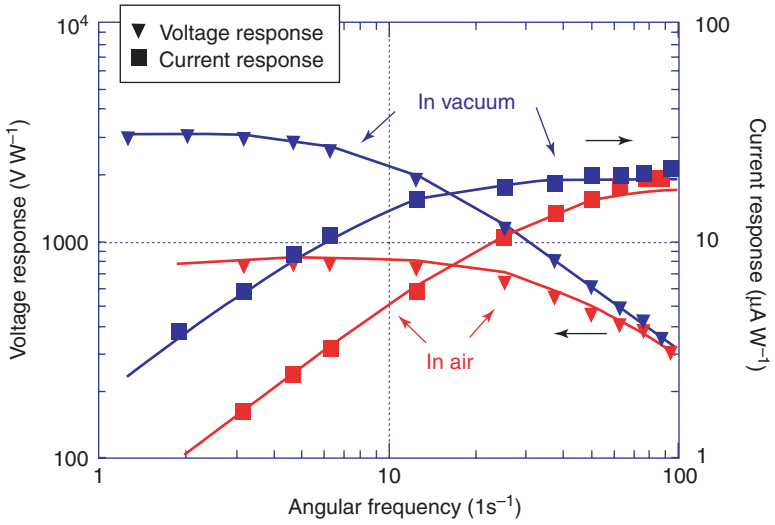


Figure 7 Voltage and current responsivity of a linear array element (1 × 12) measured in air and in vacuum. One element was 0.4 × 0.9 mm large. The curves are calculated according to eqns [7] and [9], the points are measured. (From Kohli M, Wüthrich C, Brooks KG, Willing B, Forster M, *et al.* (1997) Pyroelectric thin film sensor array. *Sensors and Actuators A* 60: 147–153.)

mix of those two situations, and is relevant in any case for the high-frequency case. It includes both the dielectric constant (note that $\varepsilon = \varepsilon_0 \varepsilon_r$) and the specific heat capacity c_p.

A small thickness is favorable to drive the current response up. The impact of the heat conductivity G at intermediate and low frequencies, and its unimportance at high frequencies, is demonstrated in **Figure 7**, which depicts the response of a pyroelectric detector in air and in vacuum, that is, once with air convection (large G), and once without air convection (small G).

The ultimate detection limit of a sensor is given by the intrinsic noise of the detector element. In order to reach this limit with the complete sensor setup, the noise contribution from the amplifier and from external sources of the environment have to be smaller than the intrinsic noise. Using low-noise amplifiers such as J-FET transistors, the theoretical limit of the intrinsic noise can indeed be reached. The thermal noise power $P_{\text{thn}} = \sqrt{4kT^2GB}$ is proportional to the square root of the heat conduction G, and is very small for optimized thin-film devices. The Johnson noise is normally dominating. The contributing noise sources are given by the real part g of the admittance Y. The noise current J_n for a bandwidth B is obtained as

$$J_n = \sqrt{4kTgB} \quad \text{with } g = R_p^{-1} + \omega C \tan\delta \qquad [10]$$

The minimal detectable power is the noise equivalent power defined by $\text{NEP} = J_n/\Re_J$. Very often, the

detectivity D^* is given instead of the NEP:

$$D^* = \frac{A^{1/2}B^{1/2}}{\text{NEP}} = \frac{\Re_I(\omega)A^{1/2}B^{1/2}}{J_n} \qquad [11]$$

The multiplication with $A^{1/2}B^{1/2}$ cancels the surface and bandwidth contribution from the noise. At high frequencies, D^* is independent of the detection area and bandwidth. For this reason, D^* is frequently used as a measure of performance to compare different detectors. The best operating conditions are found near the inverse thermal time constant, where the S/N ratio is peaking. Ideally, τ_{th} should be matched to τ_{el}^* tan δ. The materials figure of merit for the detectivity is generally given as

$$F_D = \frac{p}{c_p\sqrt{\varepsilon \tan \delta}} \qquad [12]$$

Table 1 contains a list of representative materials and their characteristics for pyroelectric applications. It can be noted that thin-film materials show as good characteristics as bulk materials, especially when they are grown epitaxially.

The detectivity D^* is a useful quantity to characterize an IR detector independent of size and detection bandwidth. However, for thermal imaging applications, D^* alone does not sufficiently characterize the complete device, which also contains an optical part. The quantity of interest is the minimal temperature difference of a blackbody target object with respect to the background temperature one still can detect. It is called the "noise equivalent temperature difference" (NETD). The procedure to evaluate this quantity is as follows: the noise voltage V_n of the detector is measured when the target temperature is in equilibrium with the background temperature. The target temperature is increased by ΔT and the signal voltage $V_s(\Delta T)$ is measured. The NETD is derived as

$$\text{NETD} = \frac{\Delta T}{V_s/V_n} \qquad [13]$$

ΔT is equal to the NETD if the signal is just equal to the noise voltage. Of course, the NETD can be related to the NEP and D^* of the detector, knowing the optical properties of the imaging system. The temperature increase ΔT has to be such that the increase of the power falling onto one pixel of the focal plane array (FPA) equals the NEP. The power increase is proportional to $(dW/dT)\Delta T$ in the wavelength interval $\Delta\lambda$ detected, where W is the power density emitted by the target surface. It is further proportional to the FPA pixel area A_D, and the transmissivity τ of the optics. The reader is referred to special literature on this topic for the derivation of general formulas. The formula given below finds

Table 1 Bulk and polymer pyroelectric materials. The figure of merits F_V, and F_D are given in the text

Material	p (μCm^{-2} K^{-1})	ε_r	tan δ	c_p (10^6 Jm^{-3} K^{-1})	F_V (m^2 C^{-1})	F_D (10^{-5} Pa$^{-1/2}$)	T_c (°C)
$LiTaO_3$, single crystal	230	47	<0.01	3.2	0.17	5–35	620
$PbZrTiO_3$ (PZFNTU), bulk ceramic	380	290	0.003	2.5	0.06	5.5	230
$PbTiO_3$, bulk ceramics	180	190	0.01	3.0	0.04	1.5	490
$PbTiO_3$ (PCWT4-24), bulk ceramic	380	220	0.01	2.5	0.08	3.4	255
P(VDF/TrFE), 80/20 copolymer film	31	7	0.015	2.3	0.22	1.4	135
PZT 15/85, (111) polycrystalline thin film on silicon (MEMS)	160–220	200–230	0.01–0.015	2.7	0.03–0.04	1.3–1.5	~420
$PbLaTiO_3$, 10–20% La polycrystalline thin film on silicon	200–576	153–550	0.01–0.024	2.7	0.02–0.15	0.7–4.1	
$PbTiO_3$, (001) thin-film epitaxy on Pt/MgO	250	97	0.006	3.2	0.09	3.4	~490

Sources: Porter SG (1981) A brief guide to pyroelectric detectors. *Ferroelectrics* 33: 193–206; Whatmore RW (1986) Pyroelectric devices and materials. *Reports on Progress in Physics* 49: 1335–1386; Bauer S and Lang SB (1996) Pyroelectric polymer electrets. *IEEE Transaction on Dielectrics and Electrical Insulation* 3: 647–676; Moulson AJ and Herbert JM (1990) *Electroceramics*. London: Chapman and Hall; Watton R, Manning PA, Perkins MJ, and Gillham JP (1996) Uncooled IR imaging: hybrid and integrated bolometer arrays. In: *Infrared Technology and Applications XXII*. Orlando: SPIE; Shorrocks NM, Whatmore RW, and Osbond PC (1990) Lead scandium tantalate for thermal detector applications. *Ferroelectrics* 106: 387–392.

frequent use:

$$\mathrm{NETD} = \frac{(4F^2+1)}{A_D \cdot \tau \cdot (\mathrm{d}W/\mathrm{d}T)_{\Delta\lambda}} \cdot \mathrm{NEP} = \frac{(4F^2+1)}{A_D^{1/2} \cdot \tau \cdot (\mathrm{d}W/\mathrm{d}T)_{\Delta\lambda}} \cdot \frac{B^{1/2}}{D^*} \qquad [14]$$

F is the f-number of the optics. Typically, an NETD of less than 100 mK is required for uncooled IR-imaging systems.

State of the Art of Pyroelectric Focal Plane Arrays

Development of pyroelectric focal plane arrays started in the 1980s. A device consisting of a reticulated, 60 μm thick pyroelectric ceramic wafer that was bonded to a CCD readout wafer by means of contact solder bumps, exhibiting a pixel diameter of 70 μm, showed an NETD of 400 mK at 25 Hz. A presently sold commercial product consists of a 14 μm thick reticulated ceramic wafer with a 50 μm pitch. The NETD could be reduced to 75 mK, and the spatial resolution could be increased. The fundamental limits are not yet reached with this camera. However, in the frame of this ceramic wafer technology, a further shrinkage of wafer thickness and pitch will become increasingly difficult. Much more promising are silicon micromachining techniques enabling still smaller and thinner pixels. Thermal cross talk between pixels is reduced very much as well. Unfortunately, there are still processing obstacles preventing integration of high quality pyroelectric thin-film structures directly onto silicon based read-out circuits. The present state of the art arrives at an NTD of 700 mK, allowing, nevertheless, for high quality IR images. In principle, micromachining technology would allow pushing down the NETD to 10–20 mK.

See also: Ferroelectricity; Piezoelectricity.

PACS: 85.50. − n; 77.70. + a; 77.55. + f; 85.85. + j

Further Reading

Buser RG and Tompsett MF (1997) Historical overview. In: Kruse PW and Skatrud DD (eds.) *Uncooled Infrared Imaging Arrays and Systems*, pp. 1–16. San Diego: Academic Press.

Chu F, Setter N, and Tagantsev AK (1993) The spontaneous relaxor-ferroelectric transition of PST. *Journal of Applied Physics* 74: 5129–5134.

Gurevich VL and Tagantsev AK (1991) Intrinsic dielectric loss in crystals. *Advanced Physics* 40: 719–767.

Hanson CM and Beratan HR (2002) Thin film ferroelectrics: breakthrough. In: Dereniak EL and Sampson RE (eds.) *Infrared Detectors and Focal Plane Arrays* VII, vol. 4721. San Jose: SPIE.

Kruse PW (1995) A comparison of the limits to the performance of thermal and photon detector imaging arrays. *Infrared Physics & Technology* 36: 869–882.

Lang SB (1974) *Sourcebook of Pyroelectricity*. London: Gordon and Breach.

Lines ME and Glass AM (1977) *Principles and Applications of Ferroelectrics and Related Materials*. Oxford: Oxford University Press.

Lloyd JM (1975) Thermal imaging systems. In: Wolfe WL (ed.), *Optical Physics and Engineering*. New York: Plenum.

Muralt P (2001) Micromachined infrared detectors based on pyroelectric thin films. *Reports on Progress in Physics* 64: 1339–1388.

Tinte S, Iniguez J, Rabe KM, and Vanderbilt D (2003) Quantitative analysis of the first principles effective Hamiltonian approach to ferroelectric perovskites. *Physical Review B* 67: 064106.

Watton R and Todd MA (1991) Induced pyroelectricity in sputtered lead scandium tantalate films and their merit for IR detector arrays. *Ferroelectrics* 118: 279–295.

Q

Quantum Cascade Lasers

A Tredicucci, NEST-INFM & Scuola Normale Superiore, Pisa, Italy

Introduction

A semiconductor heterostructure consists of a layered sequence of two or more semiconductor materials with different band gaps. The resulting band-edge discontinuity provides an effective potential profile (with wells and barriers) for the electronic motion. If the layer thicknesses are of the order of the electron wavelength, quantum confinement splits the bulk bands into a number of individual subbands. Inter-subband optical transitions then become possible, for instance involving only electrons belonging to the first conduction band. The transition energy is mainly determined by the width of the quantum well where the electron wave functions are contained.

The idea of constructing a unipolar semiconductor laser operating on inter-subband transitions is due to Kazarinov and Suris and dates back to the 1970s. The concept was immediately appealing because of the artificial nature of the transition, which allows a full control on its properties by design (energy, dipole matrix element, nonradiative scattering rate, etc.). Furthermore, two fundamental aspects are particularly relevant for the implementation in a laser. Both subbands involved originate from the same bulk conduction band and therefore are ideally characterized by the same effective mass. This produces a nearly delta-like joint density of states, maximizing the peak gain dependence on current and temperature. Furthermore, at the end of the photon emission process, the electron is still in the conduction band, and can then be recycled in a subsequent identical active region to emit another photon, for a number of times equal to the number of periods.

Despite many practical proposals and attempts, the first inter-subband laser, emitting in the mid-infrared at 4.3 μm, was realized only in 1994, at Bell Laboratories, thanks also to the development of modern semiconductor growing techniques, like molecular beam epitaxy (MBE). Population inversion was achieved through current injection via resonant tunneling, and the device was named "quantum cascade laser," reflecting the periodic potential profile with the above-mentioned "cascade" of transitions in repeated identical active regions.

In ten years, research on quantum cascade lasers (QCLs) has rapidly developed, and the devices have reached considerable performance levels. They cover the whole mid-infrared range (from 3.4 μm down to 24 μm); single-mode and tuneable emission is routinely achieved using distributed feedback (DFB) resonators; continuous wave output powers as large as 500 mW have recently been demonstrated at room temperature. Lately, far-infrared QCLs with emission frequencies in the THz range (down to 2.1 THz) have also been realized, thereby effectively bridging the gap between quantum photonic sources and classical electronic oscillators.

Physics and Design

A schematic representation of the band profile of the inter-subband laser structures originally proposed by Kazarinov and Suris is shown in **Figure 1**. In an applied external electric field, electrons are injected and traverse a series of identical quantum wells, in the process emitting a photon at every stage. In panel (a) the optical transition is within the same quantum well, between the first excited state and the ground state. Carriers then tunnel into the second excited state of the following quantum well, from where they again relax nonradiatively into the first excited state and undergo the next optical transition. In the concept of panel (b), amplification of light takes place via the tunneling of electrons from the ground state of one quantum well to an excited state of the following quantum well accompanied by the simultaneous emission of a photon.

While these first concepts illustrate well the quantum cascade principle, they turned out to be too simple for practical implementation, as two main issues had still to be properly considered. Carrier dyna mics inside the active region is mainly governed by scattering between electrons and LO phonons. In the

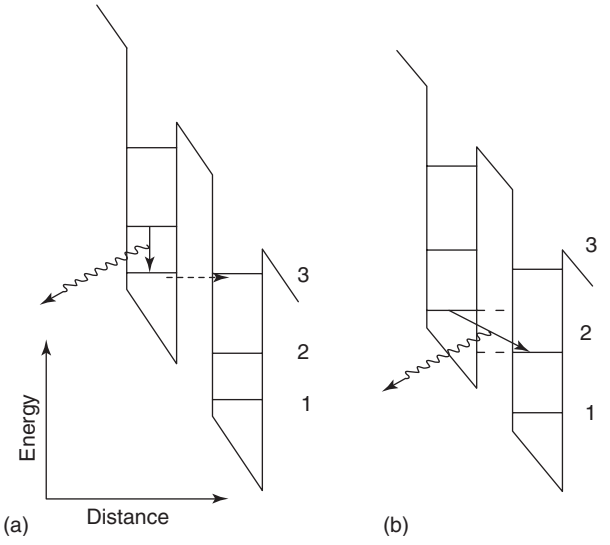


Figure 1 Schematic band-diagram representation of the original proposals by Kazarinov and Suris. (a) Electrons undergo an optical transition between levels 2 and 1 within the same quantum well and then tunnel into level 3 of the next one, from where they nonradiatively relax again into level 2. (b) Electrons undergo a photon-assisted tunnel process from one quantum well to the next, emitting their energy as photons.

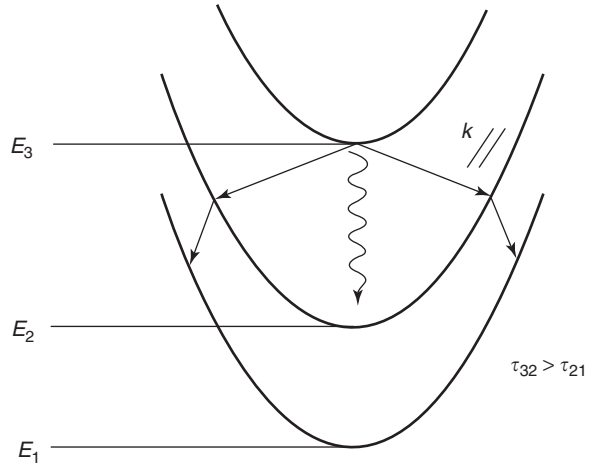


Figure 2 Schematic view of the three subbands labeled 1,2,3 in the active region of a QCL. The optical transition, represented by a wavy arrow, is between bands 3 and 2. The straight arrows indicate possible paths for scattering between electrons and LO phonons. Owing to the larger energy separation between subbands 3 and 2, the corresponding phonon wave vector is larger than in a scattering process between subbands 2 and 1. This resulting difference in the scattering rate is utilized to design population inversion.

polar III–V materials, the Fröhlich interaction is very strong and can lead to scattering rates faster than $1\,\mathrm{ps}^{-1}$. This must be compared with radiative lifetimes that are of the order of microseconds. Initially, it was thought that this would hinder population inversion, or at least make it more difficult, at transition energies above the LO phonon (36 meV in GaAs), rather than below the reststrahlen band, where LO phonon scattering is suppressed. Instead, it turned out that this strong electron–phonon scattering can be exploited to control the level occupation. The basic concept is sketched in **Figure 2**. The scattering rate for phonon emission decays as $1/Q^2$ with the phonon wave vector Q. At mid-infrared wavelengths, where the photon energy exceeds that of the phonon, the wave functions and energies of the states in the active region are engineered so that lasing takes place between subbands 3 and 2 while subbands 2 and 1 are separated by an energy close to that of the LO phonon, that is, the scattering of electrons by LO phonons becomes a quasiresonant process. Owing to the different phonon wave vectors involved in the inter-subband scattering processes between levels 3 and 2 (characteristic time τ_{32}) and 2 and 1 (τ_{21}), the lifetime of level 3, τ_3, computed as $\tau_3 \sim (1/\tau_{32} + 1/\tau_{31})^{-1}$, considerably exceeds that of level 2 $\tau_2 \sim \tau_{21}$. In essence, population inversion is achieved via resonant tunnel injection into subband 3 and using resonant phonon emission to deplete subband 2.

Solving a simple rate equation system,

$$\frac{\partial N_3}{\partial t} = \frac{J}{q_0} - \frac{N_3}{\tau_3} = 0$$

$$\frac{\partial N_2}{\partial t} = \frac{N_3}{\tau_{32}} - \frac{N_2}{\tau_2} = 0 \qquad [1]$$

the population inversion $N_3 - N_2$ can be written as

$$N_3 - N_2 = \left(1 - \frac{\tau_2}{\tau_{32}}\right) N_3 = \left(1 - \frac{\tau_2}{\tau_{32}}\right) \frac{\tau_3 J}{q_0} \qquad [2]$$

It is evident from eqn [2] that a gain in a QCL requires $\tau_{32} > \tau_2$, that is, extraction of carriers from the lower laser level must be more rapid than repopulation by carriers which have been nonradiatively scattered from the upper laser level. Using the last expression, the gain constant (in $\mathrm{m\,A}^{-1}$) of a QCL can then be written as:

$$g = \tau_3 \left(1 - \frac{\tau_2}{\tau_{32}}\right) \frac{4\pi q_0 Z_{32}^2}{\lambda_0 n \varepsilon_0 L_p} \frac{1}{\gamma_{32}} \qquad [3]$$

where Z_{32} is the dipole matrix element between states 3 and 2, λ_0 is the wavelength of the emitted light in vacuum, n the refractive index of the guided optical mode, ε_0 the vacuum permittivity, and L_p the cumulative length of the active region and the injector. γ_{32} (in eV) is the full-width at half maximum

(FWHM) of the inter-subband electroluminescence and phenomenologically replaces the expression for the joint density of states (ideally delta-like).

A second important aspect stems from the fact that a simple periodic sequence of identical quantum wells suffers from an intrinsic instability. Under current injection, the electric field across the entire superlattice will not remain constant, altering the tunneling process and leading to the build-up of charges, in turn affecting the electric field. Carrier transport will therefore differ from one superlattice period to the next, rendering the achievement of gain in such a structure very problematic. In QCLs, as demonstrated by Faist, Capasso, and co-workers in 1994, the problem of an inhomogeneous electric field along the structure can be solved by the introduction of collector/injector regions between the "active" quantum wells that host the optical transition. These regions, called injectors for short, are doped superlattice structures whose states form a miniband that collects electrons from the preceding active region, cools down the carrier distribution, and injects carriers again into the upper laser level of the following active region. The injector regions are doped in order to provide a reservoir of carriers at each stage of the cascade, ensuring electrical neutrality and thereby stabilizing the operation of the laser over a wide range of currents and voltages. A pair of one injector/collector region and one active region forms the building block (often called period or stage) that is repeated many times to give rise, under the appropriate bias, to the cascading potential. An example of a QCL design is shown in **Figure 3**. The moduli squared of the relevant wave functions are shown and the optical transition is indicated by an arrow. The wave functions belonging to the active region are drawn as thick lines. The band structure in the injector is engineered so that the majority of electrons thermalize into the ground state, which is coupled to the active region via the injection barrier. The thickness of this barrier controls the tunneling time from the injector ground state into the upper laser level and is crucial for the operation of QCLs, as its magnitude affects the injection efficiency. The injector not only collects carriers from the two lower states of the active region, but also simultaneously prevents the loss of carriers into the continuum. To this end, the minigap of the injector region is made resonant with the upper laser level. In essence, the injector can also be viewed as a Bragg reflector that possesses a high transparency for electrons with an energy equal to that of levels 2,1 and a stop-band at energies around that of level 3.

The design concept shown in **Figure 3** features an optical transition vertical in real space and a three-well active region. As explained above diagonal transitions (inter-well) are also a possibility. In general,

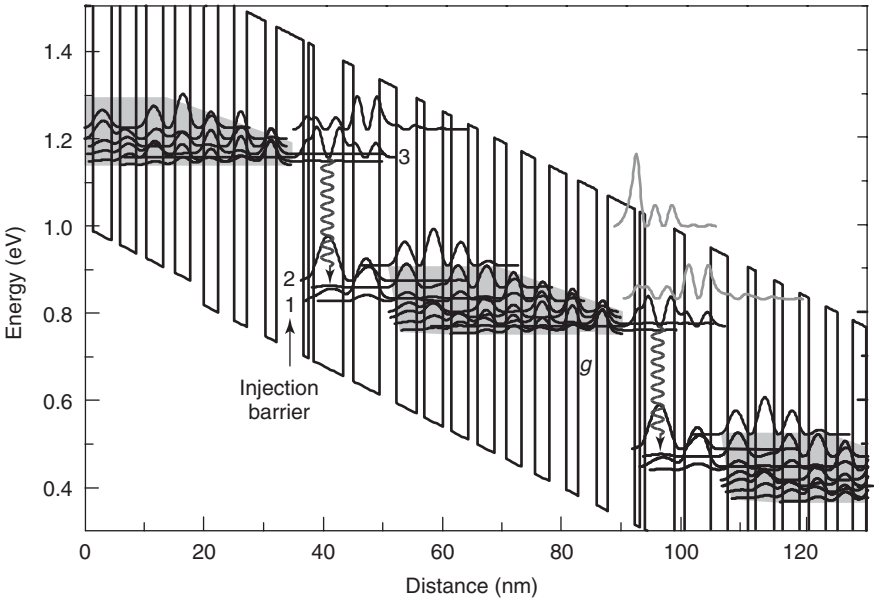


Figure 3 Schematic view of a portion of the conduction band structure of a representative $\lambda \approx 4.6\,\mu m$ QCL. Moduli squared of the most important states in the active region are drawn as thick solid lines labeled 1,2,3, while the injector ground state is labeled *g*. Carrier injection takes place from the ground state *g* into the upper laser state 3 via resonant tunneling through the injection barrier. Electrons subsequently undergo a radiative transition to level 2 (indicated by a wavy arrow), from which they are scattered mainly by LO phonon emission to state 1. Finally, they are extracted from state 1 into the states of the injector and thermalize again into the ground state of the following injector. The miniband comprising the subbands in the injector is shaded in gray. The upper laser state 3 faces a minigap downstream to prevent electrons from escaping into the continuum.

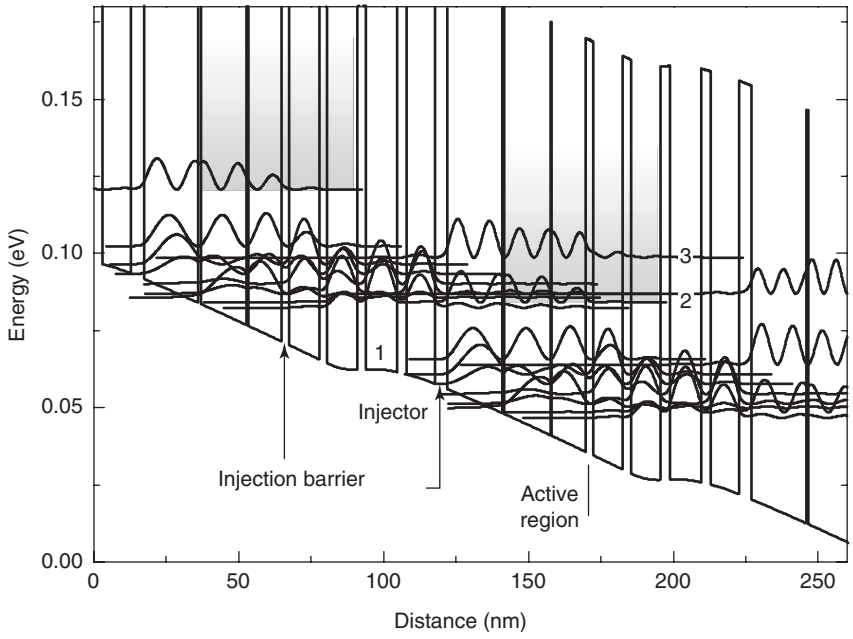


Figure 4 Conduction band energy diagram of a 4.4 THz "chirped" SL QCL under an average electric field of 3.5 kV cm^{-1}. The material is GaAs-based, and the layer thickness (in nm) are, from left to right, starting from the injection barrier 4.3/18.8/0.8/15.8/0.6/11.7/2.5/10.3/2.9/10.2/3.0/10.8/3.3/9.9, where $Al_{0.15}Ga_{0.85}As$ layers are in bold face and the 10.2 nm wide GaAs well is doped 4×10^{16} cm^{-3}. Also shown are the moduli squared of the wave functions; the optical transition occurs between states 2 and 1.

vertical transitions are preferred as they possess larger dipole matrix elements and are characterized by smaller inhomogeneous broadening. On the other hand, diagonal ones can be useful to suppress nonradiative scattering and achieve longer transition lifetimes. While the "three-well-vertical" design is probably the most classic, at least for mid-infrared QCLs, many other ideas have been proposed and successfully implemented, either to enhance performances or to obtain novel functionalities (for instance multi-wavelength or bi-directional emission). A particularly successful concept is the use of superlattice (SL) active regions, which are especially effective in simplifying the achievement of a sizeable population inversion, even in problematic situations like very small transition energies. SLs differ fundamentally from the previously described designs as lasing no longer takes place between individual subbands, but occurs between the first and second SL miniband, across the minigap. In k-space this is at the boundary of the Brillouin zone ($k_z = \pi/d$), leading to very large dipole matrix elements. Furthermore, superlattice active regions readily achieve a population inversion at $k_z = \pi/d$, owing to a favorable lifetime ratio τ_2/τ_{32}. While τ_2 is very short due to the intra-miniband relaxation, nonradiative scattering processes from the bottom of the upper miniband across the minigap now distribute among the many states of the lower miniband. This leads to a long time τ_{32}, albeit the total lifetime τ_3 remains more or less unchanged compared to that in a quantum well structure. In addition, superlattice QCLs are capable of carrying very high current densities without reaching negative differential conductance, leading to high output-power devices.

In a true superlattice, the application of bias along the SL period lifts the degeneracy of the constituent subbands; the states then become more and more localized in their parent quantum well (Wannier–Stark ladders). In order to prevent such break-up, the electric field across the active region must be balanced. In the first SL designs, this was achieved by the re-distribution of the electrons with respect to the donors, which helped in achieving a flat-band condition in almost the entire active region. This technique, however, requires the use of large dopant concentrations. Chirped superlattices instead are now widely used, in which cancellation of the electric field at the design value is obtained by varying the thickness of both wells and barriers leading to a built-in quasifield of equal magnitude and opposite sign. By careful design, truly delocalized miniband states are then recovered. A typical chirped SL design is shown in **Figure 4**, which reports the band structure of the first QCL operating at THz frequencies.

Waveguides and Devices

The resonator is an essential part of every laser. Its general requirements are low propagation losses and

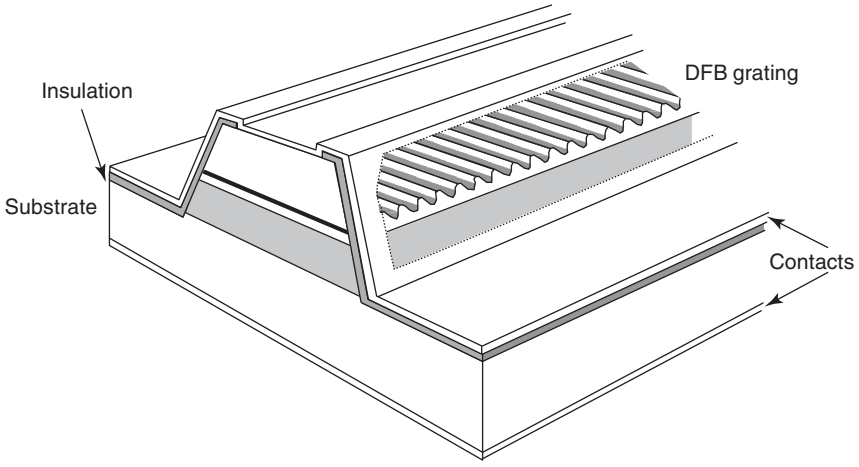


Figure 5 Schematic view of a QCL ridge. It is typically realized through optical contact lithography and wet chemical etching, with a lateral size of the order of a few wavelengths. The core of cascading stages is shaded in gray. An insulation layer (normally SiO_2 or Si_3N_4) is deposited to prevent electrical shorting when the top metallic contact has to extend also away from the ridge (for instance if the ridges are too narrow to allow wire-bonding on their top). Depicted is the case of a mid-IR laser grown on a conducting substrate. For THz QCLs, undoped substrates are preferred for their low absorption; in this case, the "bottom" contact has to be lateral. Front and end facets are obtained by cleaving along a crystal plane but coatings can be applied to alter the reflectivity. Also shown is a buried (in between core and cladding) grating, lithographically etched to provide DFB operation.

a large overlap of the optical mode with the gain medium. In QCLs, the most common resonator is the ridge-waveguide geometry as shown schematically in **Figure 5**. Light propagates along the ridge (taken to be along the x-direction) and is reflected at the back and front facets due to the step in the refractive index. Confinement in the y-direction results from the etching of the ridge. In common interband semiconductor lasers that emit in the visible to mid-infrared spectral range, dielectric waveguides are employed to achieve confinement in the z-direction. In such a waveguide the active core of the laser, in which light is generated, is sandwiched between claddings of lower refractive index leading to a bound optical state peaking in the active core and decaying exponentially into the claddings. Dielectric waveguides are indeed used also for QCLs emitting in the mid-infrared region ranging from $\lambda \sim 3.5\,\mu m$ to about $\lambda \sim 15\,\mu m$. The mode profile along z of a sample waveguide of a QCL emitting at $\lambda \sim 10.6\,\mu m$ is shown in **Figure 6**. Only the fundamental TM mode is plotted as this polarization is implied by the optical selection rules of inter-subband transitions. It is interesting to notice that, at mid-infrared wavelengths, the refractive index of a semiconductor material can be lowered considerably just by increasing the doping concentration, as the plasma frequency of the extrinsic electron gas is raised. This allows a great flexibility in the waveguide design and in the material choice. In particular effective waveguides can be realized using the same material all over the structure, with low doping in the core and large doping in the claddings (typically this is the case for GaAs QCLs).

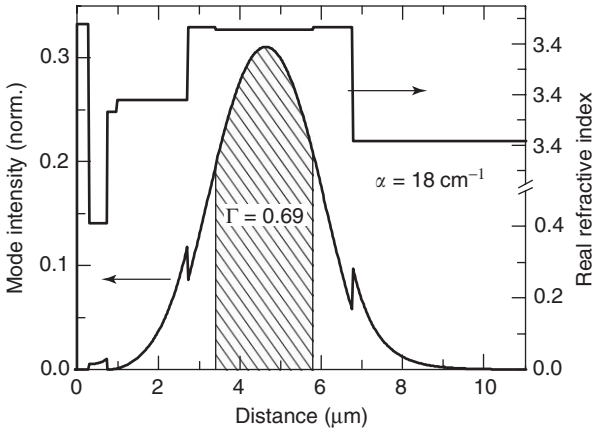


Figure 6 Mode profile of the fundamental mode of a 10.6 μm QCL. The index profile of the dielectric waveguide is shown as well. The confinement factor is calculated to be 0.64. The lower cladding layer is the InP substrate, while the top cladding is made of several AlInAs layers with various doping concentrations, of which the topmost one is highly doped (plasmon-enhanced waveguide) to decouple the mode from the metal–semiconductor interface. Above and below the stack of active regions and injectors, low-doped InGaAs layers are grown to increase the overall refractive index of the core. This results in a greater confinement factor. The discontinuities occur because the mode is TM (transverse magnetic), thus the electric field is discontinuous at the interfaces.

In determining the modal laser gain, the radiation confinement factor Γ on the active core has to be calculated by integrating the intensity over the whole

stack of cascading periods and normalizing it to the total intensity integral:

$$\Gamma = \frac{\int_{\text{core}} |E(z)|^2 \mathrm{d}z}{\int_{-\infty}^{+\infty} |E(z)|^2} \qquad [4]$$

Using eqn [3], the expression for the QCL threshold current density assumes the conventional form:

$$J_{\text{th}} = \frac{\alpha_{\text{w}} + \alpha_{\text{m}}}{g\Gamma} \qquad [5]$$

α_{w} and α_{m} being the waveguide and mirror losses respectively. The former are mainly determined by free carrier absorption in the cladding layers, which have to be doped to allow current injection; the latter are given by the impedance mismatch between the waveguide and vacuum (air) at the cleaved facets acting as front-end and back-end mirrors. In a first approximation, α_{m} can be computed considering a plane wave being reflected at the interface between the waveguide and air:

$$\alpha_{\text{m}} = -\frac{1}{L_{\text{cav}}} \ln\left(\frac{n_{\text{mod}} - 1}{n_{\text{mod}} + 1}\right) \qquad [6]$$

where n_{mod} is the modal refractive index and L_{cav} is the length of the laser cavity (typically few mm).

Although in principle there is no spectral limit for dielectric waveguides, their use at even longer wavelength is hindered by two facts: with increasing emission wavelength free carrier absorption in the cladding layers rises significantly, and, simultaneously, the necessary waveguide thickness, which scales approximately linearly with the emission wavelength, becomes incompatible with MBE growth. Yet, MBE growth is required to achieve the necessary precision and reproducibility of the layer thickness in the active core, although recently MOCVD-grown QCLs have also been reported. These problems were solved in long-wavelength QCLs by the introduction of surface-plasmon waveguides. Surface plasmons are a peculiar solution of Maxwell's equations existing at the interface between materials possessing dielectric constants ε of opposite sign. For an electric field vector perpendicular to the interface (i.e., TM-polarization), a bound state is found that propagates along the interface decaying exponentially in the perpendicular direction. Such a system is realized, for example, at the interface between a metal (negative ε) and a low-doped semiconductor. If a metal with a very large (in absolute value) dielectric constant is chosen, the penetration of the mode into the metal can be made very small and the absorption of light in the metal becomes negligible, reducing the optical losses experienced by the mode. Surface plasmon waveguides become advantageous over dielectric ones above 15 μm wavelength or so and guarantee large confinement factors even with an epitaxial thickness considerably smaller than the wavelength. **Figure 7** shows the computed mode profile in a waveguide using this concept for a $\lambda \sim 24\,\mu\text{m}$ QCL based on the InP material system. The optical confinement reaches a very high value of 88%. Propagation losses are $42\,\text{cm}^{-1}$, and are caused mainly by the doped stack of active regions and the penetration into the doped substrate. A sophisticated version of the surface plasmon waveguide featuring a second layer with negative dielectric constant (realized in a highly doped semiconductor) is customarily employed for THz QCLs.

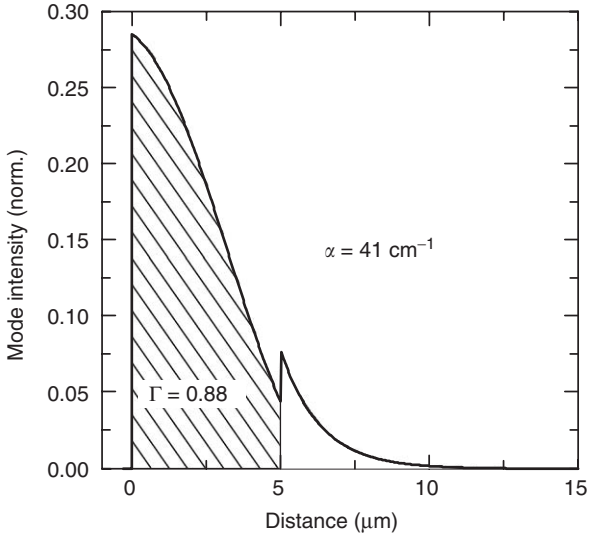


Figure 7 Waveguide of a $\lambda \approx 24\,\mu\text{m}$ QCL employing the surface plasmon to achieve a high confinement factor with a small epilayer thickness. In the present case, this is helped by the additional confinement coming from the low refractive index of the InP substrate, a feature not available in the GaAs material system.

Beyond Fabry–Perot ridge resonators, QCLs have been made with a variety of optical cavities. Suspended microdisks and cylinder lasers based on whispering gallery modes have shown the lowest threshold currents, while deformed stadium-shaped lasers have been realized to study chaotic light propagation. The interest for spectroscopic applications has stimulated the development of single-mode and tunable devices. As for interband laser diodes, this is achieved integrating a diffraction grating in the laser structure (see **Figure 5**), giving rise to distributed-feedback (DFB) resonators. Many different techniques have been developed, relying in turn on gratings patterned on top of the ridge, buried between the ridge and the cladding, two-metal gratings in the top contact, etc. In any case, the presence of

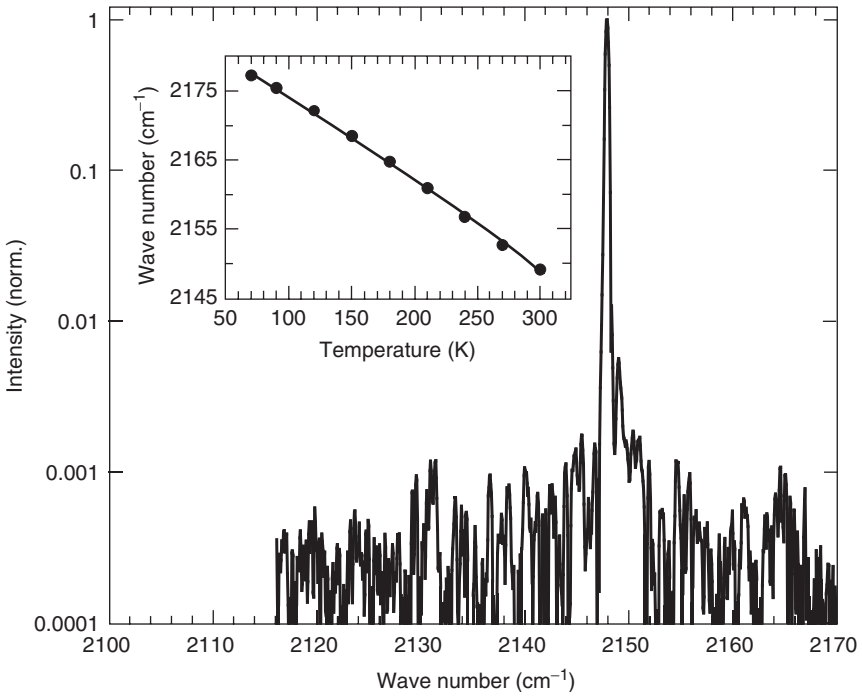


Figure 8 Single-mode laser emission spectrum of a 4.8 μm DFB QCL plotted in logarithmic scale. A side mode suppression ratio of about 30 dB is displayed. In the inset the emission frequency dependence on the laser temperature is displayed; the tuning range is $\sim 30\,\mathrm{cm}^{-1}$ from cryogenic to room temperature. (Courtesy of C Gmachl, Princeton University.)

the grating provides a strongly frequency-selective reflection, which ensures laser oscillation only on the longitudinal mode whose wavelength matches the grating period. Tunability is provided by the change of the material refractive index with temperature (which can be induced also by varying the drive current). High side-mode suppression ratios of more than 30 dB as well as very narrow line widths can be obtained. **Figure** 8 displays an exemplary emission spectrum and tuning characteristics of a 4.8 μm wavelength DFB QCL. Finally, external cavity QCLs have also been recently implemented using an external diffraction grating in Littröw configuration, to allow for a much wider tuning range.

Performance and State of the Art

The first QCLs were realized in the InGaAs/AlInAs material system-grown lattice matched to InP substrate. Later on strain-compensated structures were also introduced to increase the available conduction band offset. Indeed InP-based structures currently still provide the best QC devices for the mid-infrared spectral range. The reasons are to be found in the small electron effective mass, which improves the radiative efficiency, the rather large band offset, the good transport characteristics, the low refractive index substrates, and obviously the very mature fabrication and processing technology. Typical threshold current densities are of the order of a $\mathrm{kA\,cm}^{-2}$, with an exponential dependence on the temperature $\exp(T/T_0)$ characterized by a large T_0 in between 100 K and 200 K. Devices have been realized with a varying number of cascade stages, ranging from a single one to more than one hundred. Obviously the more stages are used in the cascade, the higher output power can be obtained; however, more periods also mean larger applied voltages (the voltage drop per period has to be constant), thereby resulting in higher electric power to be dissipated. In the best lasers, the QC principle, that above the threshold one photon per period is generated per each single electron traversing the structure, has been experimentally verified from the output power slope efficiency:

$$\frac{\partial P}{\partial I} = \frac{1}{2} N_p \frac{\alpha_\mathrm{m}}{\alpha_\mathrm{m} + \alpha_\mathrm{w}} \frac{h\nu}{q_0} \left(1 - \frac{\tau_2}{\tau_{32}}\right) \qquad [7]$$

in which N_P is the number of stages and ν the photon frequency. Peak output powers of about 2 W have been obtained in pulsed operation; at $\lambda \approx 6\,\mu\mathrm{m}$ wavelength, room-temperature continuous wave (CW) operation with about 0.5 W of optical power has been recently demonstrated in a QCL employing a buried heterostructure configuration of the ridge (see data

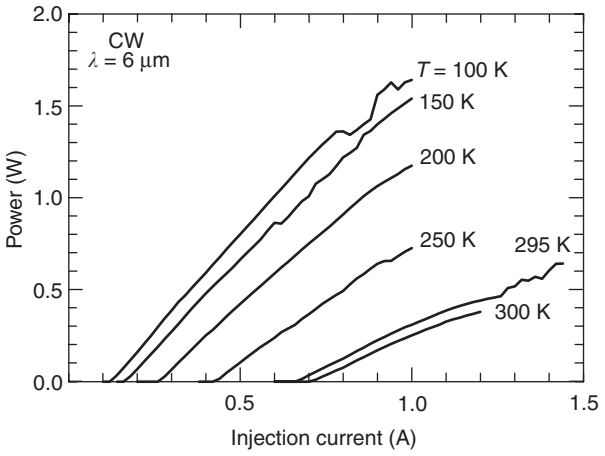


Figure 9 Light–current characteristics as recorded in continuous wave from the front facet of a $\lambda \approx 6\,\mu m$ InGaAs/AlInAs buried heterostructure QCL. A metallic high reflection coating was used on the back-facet. The power represents what collected on a thermopile detector. Data are shown for different substrate temperatures. (Courtesy of M Razeghi, Nortwestern University.)

in **Figure 9**). It actually took eight years before QCLs could be driven in CW at room temperature, owing mainly to the difficult thermal management of a device needing a few watts of electrical power.

In 1998, the first QCLs in the GaAs/AlGaAs material system were developed. While these devices cannot compete with InGaAs/AlInAs ones in the mid-infrared, GaAs QCLs are the only ones presently capable of operating in the THz range, thanks to the highest quality of epitaxial growth and the far-IR transparency of commercially available GaAs substrates. THz QCLs were first fabricated in 2002, after solving the problem of population inversion at transition energies smaller than the LO phonon resonance using SL active regions, and implementing surface-plasmon-based special waveguides. Currently, they cover the region from 2.1 THz to 4.7 THz, with output powers of several tens of mWs, up to a record temperature of 140 K (90 K in CW). Interestingly, THz QCLs display the lowest threshold current densities, down to a couple of $A\,cm^{-2}$ in an external magnetic field.

In an opposite direction, material systems with very large band offsets are being explored to extend the QC concept to the near-IR, possibly toward communication wavelengths. In this respect, an important step has been the very recent development of QCLs in InAs/AlSb heterostructures. The interest stems from the possibility of directly modulating QCLs at very high frequencies, owing to their unipolar nature, and to the ultrafast electron dynamics; such a feature is clearly highly desirable for high-throughput data transmission. Modulation at several GHz has already been proved, limited probably by parasitic capacitance, as well as self-mode-locking on ultrashort pulses.

As far as the spectral emission properties are concerned, DFB QCLs have shown excellent single-mode line widths, and are being implemented in an ever-increasing number of spectroscopic applications, mainly related to trace-gas detection. Direct and modulated absorption, photoacoustic spectroscopy, cavity ring-down, are just a few of the techniques successfully employed. Furthermore, they can be stabilized on molecular absorption lines using servo-feedback systems to bring the frequency noise down to a few Hz, thereby making them appealing also as frequency standards.

The Future

With the commercialization of mid-IR QCLs already a developing reality, open research challenges of immediate technological relevance invest the operation of QCLs in other frequency ranges. Apart from the already mentioned quest for near-IR devices, many elements of the newly developed THz QCLs are presently the object of intense investigation. The improvement of temperature performances, single-mode tunability, and the extension toward 1 THz are the most important goals. In parallel, the application of THz QCLs to imaging systems and security scanners is attracting a lot of attention from the industrial world.

In a similar spirit, nonlinear devices integrated within a QCL cavity are being explored for harmonic and difference frequency generation, with promising perspectives. This research is also stimulating the study of fundamental quantum optics effects using inter-subband transitions.

Basic aspects of the QC concept are also being revisited. The interplay between population inversion and Bloch oscillations is particularly interesting, but the possibility of realizing QCLs in lower dimensionality nanostructures (such as quantum dots or nanowhiskers) is probably the most stimulating idea.

Finally, QCLs offer a perfect laboratory system for developing studies and new devices in various fields of modern optics. Photonic crystals and surface plasmon sub-wavelength components are two examples where QCLs (thanks to their planar TM emission, long wavelength, and insensitivity to surface recombination) could have a deep immediate impact.

See also: Electron–Phonon Interactions and the Response of Polarons; Semiconductor Lasers; Semiconductor Nanostructures; Semiconductor Optics; Semiconductors, Electronic Transitions in; Semiconductors, Optical and Infrared Properties of.

PACS: 42.55.Px; 42.55.Sa; 42.60.Da; 42.72.Ai; 73.21. − b; 73.63.Hs; 78.45. + h; 78.60.Fi; 78.67. − n

Further Reading

Beck M, Hofstetter D, Aellen T, Faist J, Oesterle V, *et al.* (2002) Continuous wave operation of a mid-infrared semiconductor laser at room temperature. *Science* 295: 301–305.

Capasso F, Tredicucci A, Gmachl C, Sivco DL, Hutchinson AL, *et al.* (1999) High-performance superlattice quantum cascade lasers. *IEEE J. Select. Topics Quantum Electron* 5: 792–807.

Capasso F, Gmachl C, Paiella R, Tredicucci A, Hutchinson AL, *et al.* (2000) New frontiers in quantum cascade lasers and applications. *IEEE J. Select. Topics Quantum Electron* 6: 931–947.

Capasso F, Gmachl C, Sivco DL, and Cho AY (2002) Quantum cascade lasers. *Physics Today* 55(5): 34–40.

Choi HK (ed.) (2004) *Long-Wavelength Infrared Semiconductor Lasers*. Chichester: Wiley.

Faist J, Capasso F, Sivco DL, Sirtori C, Hutchinson AL, *et al.* (1994) Quantum cascade lasers. *Science* 264: 553–556.

Gmachl C, Capasso F, Köhler R, Tredicucci A, Hutchinson AL, *et al.* (2000) The sense-ability of semiconductor lasers. *IEEE Circuits and Devices* 16(3): 10–18.

Gmachl C, Capasso F, Sivco DL, and Cho AY (2001) Recent progress in quantum cascade lasers and applications. *Reports of Progress in Physics* 64: 1533–1601.

Köhler R, Tredicucci A, Beltram F, Beere HE, Linfield EH, *et al.* (2002) Thz semiconductor-heterostructure laser. *Nature* 417: 156–159.

Liu HC and Capasso F (ed.) (2000) Intersubband Transitions in Quantum Wells: Physics and Device Applications, vols. 1 and 2. London: Academic Press.

Quantum Computation and Chaos

G Casati and G Benenti, Università degli Studi dell'Insubria, Como, Italy

Introduction

Miniaturization provides us with an intuitive way of understanding why, in the near future, quantum mechanics will become important for computation. The electronics industry for computers grows hand-in-hand with the decrease in size of integrated circuits. This miniaturization is necessary to increase computational power, that is, the number of floating-point operations per second (flops) a computer can perform. In the 1950s, electronic computers based on vacuum-tube technology were capable of performing $\sim 10^3$ floating-point operations per second, while nowadays (2004) there exist supercomputers whose power is greater than 10 teraflops (10^{13} flops). As remarked, this enormous growth of computational power has been made possible owing to progress in miniaturization, which may be quantified empirically in Moore's law. This law is the result of a remarkable observation made by Gordon Moore in 1965: the number of transistors on a single integrated-circuit chip doubles approximately every 18–24 months. This exponential growth has not yet saturated and Moore's law is still valid. At the present time the limit is $\sim 10^8$ transistors per chip and the typical size of circuit components is ~ 100 nm. Extrapolating Moore's law, it is estimated that around the year 2020, one would reach the atomic size for storing a single bit of information. At that point, quantum effects will become unavoidably dominant.

Quantum physics sets fundamental limitations on the size of the circuit components. The first question under debate is whether it would be more convenient to push the silicon-based transistor to its physical limits or instead to develop alternative devices, such as quantum dots, single-electron transistors, or molecular switches. A common feature of all these devices is that they are at the nanometer length scale, and therefore quantum effects play a crucial role.

So far, the quantum switches that could substitute silicon-based transistors and possibly be connected together to execute classical algorithms based on Boolean logic were discussed. In this perspective, quantum effects are simply unavoidable corrections that must be taken into account, owing to the nanometer size of the switches. A quantum computer represents a radically different challenge: the aim is to build a machine based on quantum logic, that is, it processes the information and performs logic operations in agreement with the laws of quantum mechanics.

Quantum Logic

The elementary unit of quantum information is called a qubit (the quantum counterpart of the classical bit) and a quantum computer may be viewed as a many-qubit system. Physically, a qubit is a two-level system, like the two spin states of a spin-1/2 particle, the vertical and horizontal polarization states of a single photon or the ground and excited states of an atom.

A classical bit is a system that can exist in two distinct states, which are used to represent 0 and 1, that is, a single binary digit. The only possible operations (gates) in such a system are the identity

$(0\to 0,\ 1\to 1)$ and NOT $(0\to 1,\ 1\to 0)$. In contrast, a quantum bit (qubit) is a two-level quantum system, described by a two-dimensional complex Hilbert space. In this space, one may choose a pair of normalized and mutually orthogonal quantum states, called $|0\rangle$ and $|1\rangle$ (say, the eigenstates of the Pauli operator σ_z), to represent the values 0 and 1 of a classical bit. These two states form a computational basis. From the superposition principle, any state of the qubit may be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad [1]$$

where the amplitudes α and β are complex numbers, constrained by the normalization condition $|\alpha|^2 + |\beta|^2 = 1$.

A quantum computer can be seen as a collection of n qubits and therefore its wave function resides in a 2^n-dimensional complex Hilbert space. While the state of an n-bit classical computer is described in binary notation by an integer $k \in [0, 2^n - 1]$,

$$k = k_{n-1}2^{n-1} + \cdots + k_1 2 + k_0 \qquad [2]$$

with $k_0, k_1, \ldots, k_{n-1} \in [0, 1]$ binary digits, the state of an n-qubit quantum computer is

$$\begin{aligned} |\psi\rangle &= \sum_{k=0}^{2^n-1} c_k|k\rangle \\ &= \sum_{k_{n-1},\ldots,k_1,k_0=0}^{1} c_{k_{n-1},\ldots,k_1,k_0}|k_{n-1}\cdots k_1 k_0\rangle \end{aligned} \qquad [3]$$

where $|k_{n-1}\cdots k_1 k_0\rangle \equiv |k_{n-1}\rangle \otimes \cdots \otimes |k_1\rangle \otimes |k_0\rangle$. Notice that the complex numbers c_k are constrained by the normalization condition $\sum_{k=0}^{2^n-1} |c_k|^2 = 1$.

The superposition principle is clearly visible in eqn [3]: while n classical bits can store only a single integer k, the n-qubit quantum register can not only be prepared in the corresponding state $|k\rangle$ of the computational basis, but also in a superposition. It is stressed that the number of states of the computational basis in this superposition can be as large as 2^n, which grows exponentially with the number of qubits. The superposition principle opens up new possibilities for computation. When one performs a computation on a classical computer, different inputs require separate runs. In contrast, a quantum computer can perform a computation for exponentially many inputs on a single run. This huge parallelism is the basis of the power of quantum computation.

The superposition principle is not a uniquely quantum feature. Indeed, classical waves satisfying the superposition principle do exist. For instance, consider the wave equation for a vibrating string with fixed endpoints. Its solutions $|\varphi_k\rangle$ satisfy the superposition principle and one can write the most general state $|\varphi\rangle$ of a vibrating string as a linear superposition of these solutions, which are analogous to eqn [3]: $|\varphi\rangle = \sum_{k=0}^{2^n-1} c_k|\varphi_k\rangle$. It is therefore also important to point out the importance of entanglement for the power of quantum computation, as compared to any classical computation. Entanglement is the most spectacular and counter-intuitive manifestation of quantum mechanics, observed in composite quantum systems: it signifies the existence of nonlocal correlations between measurements performed on well-separated particles. After two classical systems have interacted, they are in well-defined individual states. In contrast, after two quantum particles have interacted, in general, they can no longer be described independently of each other. There will be purely quantum correlations between two such particles, independently of their spatial separation. Examples of two-qubit entangled state are the four states of the so-called Bell basis, $|\phi^{\pm}\rangle = \frac{1}{2}(|00\rangle \pm |11\rangle)$ and $|\psi^{\pm}\rangle = \frac{1}{2}(|01\rangle \pm |10\rangle)$. The measure of the polarization state of one qubit will instantaneously affect the state of the other qubit, whatever their distance is. There is no entanglement in classical physics. Therefore, in order to represent the superposition of 2^n levels by means of classical waves, these levels must belong to the same system. Indeed, classical states of separate systems can never be superposed. Thus, any computation based on classical waves requires a number of levels that grow exponentially with n. If Δ is the typical energy separation between two consecutive levels, the amount of energy required for this computation is given by $\Delta 2^n$. Hence, the amount of physical resources needed for the computation grows exponentially with n. In contrast, due to entanglement, in quantum physics a general superposition of 2^n levels may be represented by means of n qubits. Thus, the amount of physical resources (energy) grows only linearly with n.

To implement a quantum computation, one must be able to control the evolution in time of the many-qubit state describing the quantum computer. As far as the coupling to the environment is neglected, this evolution is unitary and governed by the Schrödinger equation. It is well known that a small set of elementary logic gates allows the implementation of any complex computation on a classical computer. This is very important: it means that, when one changes the problem, one does not need to modify one's computer hardware. Fortunately, the same property remains valid for a quantum computer. It turns out that, in the quantum circuit model, each unitary transformation acting on a many-qubit system can be decomposed into quantum gates acting on a single qubit

and a suitable quantum gate acting on two qubits. Any unitary operation on a single qubit can be constructed using only Hadamard- and phase-shift gates. The Hadamard gate is defined as follows: it turns $|0\rangle$ into $(|0\rangle + |1\rangle)/\sqrt{2}$ and $|1\rangle$ into $(|0\rangle - |1\rangle)/\sqrt{2}$. The phase-shift gate (of phase δ) turns $|0\rangle$ into $|0\rangle$ and $|1\rangle$ into $e^{i\delta}|1\rangle$. A generic unitary transformation acting on a many-qubit state can be transformed into a sequence of Hadamard, phase-shift, and CNOT gates, where CNOT is a two-qubit gate, defined as follows: it turns $|00\rangle$ into $|00\rangle$, $|01\rangle$ into $|01\rangle$, $|10\rangle$ into $|11\rangle$, and $|11\rangle$ into $|10\rangle$. As in the classical XOR gate, the CNOT gate flips the state of the second (target) qubit if the first (control) qubit is in the state $|1\rangle$, and does nothing if the first qubit is in the state $|0\rangle$. Of course, the CNOT gate, in contrast to the classical XOR gate, can also be applied to any superposition of the computational basis states.

The decomposition of generic unitary transformation of an n-qubit system into elementary quantum gates is in general inefficient, that is, it requires a number of gates exponentially large in n (more precisely, $O(n^2 4^n)$ quantum gates). However, there are special unitary transformations that can be computed efficiently in the quantum circuit model, by means of a number of elementary gates polynomial in n. A very important example is given by the quantum Fourier transform, mapping a generic n-qubit state $\sum_{k=0}^{2^n-1} a_k|k\rangle$ into $\sum_{l=0}^{2^n-1} b_l|l\rangle$, where the vector $\{b_0, \ldots, b_{N-1}\}$ is the discrete Fourier transform of the vector $\{a_0, \ldots, a_{N-1}\}$, that is, $b_l = \sum_{k=0}^{N-1} e^{2\pi i k l/2^n} a_k$. It can be shown that this transformation can be efficiently implemented in $O(n^2)$ elementary quantum gates, whereas the best-known classical algorithm to simulate the Fourier transform, the fast Fourier transform, requires $O(n2^n)$ elementary operations. The quantum Fourier transform is an essential subroutine in many quantum algorithms.

Quantum Algorithms

As shown above, the power of quantum computation is due to the inherent quantum parallelism associated with the superposition principle. In simple terms, a quantum computer can process a large number of classical inputs in a single run. For instance, starting from the input state $\sum_{k=0}^{2^n-1} c_k|k\rangle \otimes |0 \ldots 0\rangle$, one may obtain the output state

$$\sum_{k=0}^{2^n-1} c_k|k\rangle \otimes |f(k)\rangle \qquad [4]$$

Therefore, the function $f(k)$ is computed for all k in a single run (note that one needs two quantum registers to compute by means of a reversible unitary transformation $f(x)$). However, it is not an easy task to extract useful information from the output state. The problem is that this information is, in a sense, hidden. Any quantum computation ends up with a projective measurement in the computational basis: the qubit polarization is measured along the z-axis for all the qubits. The output of the measurement process is inherently probabilistic and the probabilities of the different possible outputs are set by the basic postulates of quantum mechanics. Given the state [4], one obtains $|\bar{k}\rangle|f(\bar{k})\rangle$ with probability $|c_{\bar{k}}|^2$, hence, the evaluation of the function $f(k)$ for a single $k = \bar{k}$, exactly as with a classical computer. However, there exist quantum algorithms that exploit quantum interference to efficiently extract useful information.

In 1994, Peter Shor proposed a quantum algorithm that efficiently solves the prime-factorization problem: given a composite odd positive integer N, find its prime factors. This is a central problem in computer science and it is conjectured, though not proven, that for a classical computer it is computationally difficult to find the prime factors. Shor's algorithm instead efficiently solves the integer factorization problem in $O((n^2 \log n \log\log n))$ elementary quantum gates, where $n = \log N$ is the number of bits necessary to code the input N. Therefore, it provides an exponential improvement in speed with respect to any known classical algorithm. Indeed, the best classical algorithm, the number field sieve, requires $\exp(O(n^{1/3}(\log n)^{2/3}))$ operations. The integer factoring problem can be reduced to the problem of finding the period of the function $f(k) = a^k \bmod N$, where N is the number to be factorized and $a < N$ is chosen randomly. The modular exponentiation can be computed efficiently on a quantum computer and one obtains the state $(1/\sqrt{N}) \sum_{k=0}^{2^n-1} |k\rangle|f(k)\rangle$. Notice that there are two quantum registers: the first one stores k, the second $f(k)$. By measuring the second register, one obtains the outcome $f(\bar{k})$. Thus, the quantum computer wave function collapses onto $(1/\sqrt{m}) \sum_{j=0}^{m-1} |\bar{k} + jr\rangle|f(\bar{k})\rangle$, where m is the number of k values such that $f(k) = f(\bar{k})$, and r is the period of $f(k)$, that is, $f(k) = f(k + r)$. To determine the period r, one has to perform the quantum Fourier transform of the first register. It can be seen that the resulting wave function is peaked around integer multiples of N/r. From the measurement of this state, one can extract the period r. It is worth mentioning that there are cryptographic systems, such as RSA, that are used extensively today and that are based on the conjecture that no efficient algorithms exist for solving the prime factorization problem. Hence Shor's algorithm, if implemented on a large-scale quantum computer, would break the RSA cryptosystem.

Other quantum algorithms have been developed. In particular, Grover has shown that quantum computers can also be useful for solving the problem of searching for a marked item in an unstructured database of $N = 2^n$ items. The best one can do with a classical computer is to go through the database until one finds the solution. This requires $O(N)$ operations. In contrast, the same problem can be solved by a quantum computer in $O(\sqrt{N})$ operations. In this case, the gain with respect to classical computation is quadratic.

Quantum Simulation of Physical Systems

The simulation of quantum many-body problems on a classical computer is a difficult task as the size of the Hilbert space grows exponentially with the number of particles. For instance, if one wishes to simulate a chain of n spin-1/2 particles, the size of the Hilbert space is 2^n. Namely, the state of this system is determined by 2^n complex numbers. As observed by Feynman in the 1980s, the growth in memory requirement is only linear on a quantum computer, which is itself a many-body quantum system. For example, to simulate n spin-1/2 particles one only needs n qubits. Therefore, a quantum computer operating with only a few tens of qubits can outperform a classical computer. Of course, this is only true if one can find an efficient quantum algorithm and if one can efficiently extract useful information from the quantum computer. Quite interestingly, a quantum computer can be useful not only for the investigation of the properties of many-body quantum systems, but also for the study of the quantum and classical dynamics of complex single-particle systems.

For a concrete example, consider the quantum mechanical motion of a particle in one dimension (the extension to higher dimensions is straightforward). It is governed by the Schrödinger equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\psi(x,t) = H\psi(x,t) \qquad [5]$$

where the Hamiltonian H is given by

$$H = H_0 + V(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \qquad [6]$$

The Hamiltonian $H_0 = -(\hbar^2/2m)\,\mathrm{d}^2/\mathrm{d}x^2$ governs the free motion of the particle, while $V(x)$ is a one-dimensional potential. To solve eqn [5] on a quantum computer with finite resources (a finite number of qubits and a finite sequence of quantum gates), one must first of all discretize the continuous variables x and t. If the motion essentially takes place inside a finite region, say $-d \leqslant x \leqslant d$, decompose this region into 2^n intervals of length $\Delta = 2d/2^n$ and represent these intervals by means of the Hilbert space of an n-qubit quantum register (this means that the discretization step drops exponentially with the number of qubits). Hence, the wave function $|\psi(t)\rangle$ is approximated as follows:

$$|\tilde{\psi}(t)\rangle = \frac{1}{\mathcal{N}}\sum_{i=0}^{2^n-1}\psi(x_i,t)|i\rangle \qquad [7]$$

where $x_i \equiv -d + (i+1/2)\Delta$, $|i\rangle = |i_{n-1}\rangle \otimes \cdots \otimes |i_0\rangle$ is a state of the computational basis of the n-qubit quantum register and $\mathcal{N} \equiv \sqrt{\sum_{i=0}^{2^n-1}|\psi(x_i,t)|^2}$ is a factor that ensures correct normalization of the wave function. It is intuitive that $|\tilde{\psi}\rangle$ provides a good approximation to $|\psi\rangle$ when the discretization step Δ is much smaller than the shortest length scale relevant for the motion of the system. The Schrödinger equation [5] may be integrated by propagating the initial wave function $\psi(x,0)$ for each time-step ε as follows:

$$\psi(x,t+\varepsilon) = \mathrm{e}^{-(\mathrm{i}/\hbar)[H_0+V(x)]\varepsilon}\psi(x,t) \qquad [8]$$

If the time-step ε is small enough, it is possible to write the Trotter decomposition

$$\mathrm{e}^{-(\mathrm{i}/\hbar)[H_0+V(x)]\varepsilon} \approx \mathrm{e}^{-(\mathrm{i}/\hbar)H_0\varepsilon}\mathrm{e}^{-(\mathrm{i}/\hbar)\ V(x)\varepsilon} \qquad [9]$$

which is exact up to terms of order ε^2. The operator on the right-hand side of eqn [9] is still unitary, simpler than that on the left-hand side, and, in many interesting physical problems, can be efficiently implemented on a quantum computer. Advantage is taken of the fact that the Fourier transform can be efficiently performed by a quantum computer. One can then write the first operator in the right-hand side of eqn [9] as

$$\mathrm{e}^{-(\mathrm{i}/\hbar)\ H_0\varepsilon} = F^{-1}\mathrm{e}^{+(\mathrm{i}/\hbar)\ \left(\hbar^2k^2/2m\right)\varepsilon}F \qquad [10]$$

where k is the variable conjugated to x and F the discrete Fourier transform. This represents a transformation from the x-representation to the k-representation, in which this operator is diagonal. Then, using the inverse Fourier transform F^{-1}, one returns to the x-representation, in which the operator $\exp(-\mathrm{i}V(x)\varepsilon/\hbar)$ is diagonal. The wave function $\psi(x,t)$ at time $t = l\varepsilon$ is obtained from the initial wave function $\psi(x,0)$ by applying l times the unitary operator

$$F^{-1}\mathrm{e}^{+(\mathrm{i}/\hbar)\ \left(\hbar^2k^2/2m\right)\varepsilon}F\mathrm{e}^{-(\mathrm{i}/\hbar)\ V(x)\varepsilon} \qquad [11]$$

Therefore, simulation of the Schrödinger equation is now reduced to the implementation of the Fourier transform plus diagonal operators of the form

$$|x\rangle \to \mathrm{e}^{\mathrm{i}cf(x)}|x\rangle \qquad [12]$$

where c is some real constant. Note that an operator of the form [12] appears both in the computation of $\exp(-\mathrm{i}V(x)\varepsilon/\hbar)$ and of $\exp(-\mathrm{i}H_0\varepsilon/\hbar)$, when this latter operator is written in the k-representation. The quantum computation of [12] is possible, using an ancillary quantum register $|y\rangle_a$, by means of the following steps:

$$\begin{aligned} |0\rangle_a \otimes |x\rangle &\to |f(x)\rangle_a \otimes |x\rangle \\ &\to \mathrm{e}^{\mathrm{i}cf(x)}|f(x)\rangle_a \otimes |x\rangle \\ &\to \mathrm{e}^{\mathrm{i}cf(x)}|0\rangle_a \otimes |x\rangle \\ &= |0\rangle_a \otimes \mathrm{e}^{\mathrm{i}cf(x)}|x\rangle \end{aligned} \qquad [13]$$

The first step is a standard function evaluation and may be implemented by means of $O(n2^n)$ elementary quantum gates. Of course, more efficient implementations (polynomial in n) are possible when the function $f(x)$ has some structure, as is the case for the potentials $V(x)$ usually considered in quantum mechanical problems. The second step in eqn [13] is the transformation $|y\rangle_a \to \mathrm{e}^{\mathrm{i}cy}|y\rangle_a$ and can be performed in m single-qubit phase-shift gates, m being the number of qubits in the ancillary register. Indeed, one may write the binary decomposition of an integer $y \in [0, 2^m - 1]$ as $y = \sum_{j=0}^{m-1} y_j 2^j$, with $y_j \in \{0, 1\}$. Therefore,

$$\exp(\mathrm{i}y) = \exp\left(\sum_{j=0}^{m-1} \mathrm{i}cy_j 2^j\right) = \prod_{j=0}^{m-1} \exp(\mathrm{i}cy_j 2^j) \qquad [14]$$

which is the product of m single-qubit gates, each acting nontrivially (differently from identity) only on a single qubit. The jth gate operates the transformation $|y_j\rangle_a \to \exp(\mathrm{i}cy_j 2^j)|y_j\rangle_a$, with $|y_j\rangle_a \in \{|0\rangle, |1\rangle\}$ vectors of the computational basis for the jth ancillary qubit. The third step in eqn [13] is just the reverse of the first and may be implemented by the same array of gates as the first but applied in the reverse order. After this step the ancillary qubits are returned to their standard configuration $|0\rangle_a$ and it is therefore possible to use the same ancillary qubits for every time-step. Note that the number of ancillary qubits m determines the resolution in the computation of the diagonal operator [12]. Indeed, the function $f(x)$ appearing in eqn [12] is discretized and can take 2^m different values.

An interesting physical model, the so-called quantum sawtooth map, that can be simulated efficiently on a quantum computer without ancillary qubits is discussed here. The sawtooth map is a prototype model in the studies of classical and quantum dynamical systems and exhibits a rich variety of physical phenomena, from complete chaos to complete integrability, normal and anomalous diffusion, and quantum dynamical localization. The sawtooth map belongs to the class of periodically driven dynamical systems, governed by the Hamiltonian

$$H(\theta, I; \tau) = \frac{I^2}{2} + V(\theta) \sum_{j=-\infty}^{+\infty} \delta(\tau - jT) \qquad [15]$$

where (I, θ) are conjugate action-angle variables $(0 \leqslant \theta < 2\pi)$. This Hamiltonian is the sum of two terms, $H(\theta, I; \tau) = H_0(I) + U(\theta; \tau)$, where $H_0(I) = I^2/2$ is just the kinetic energy of a free rotator (a particle moving on a circle parameterized by the coordinate θ), while $U(\theta; \tau) = V(\theta) \sum_j \delta(\tau - jT)$ represents a force acting on the particle that is switched on and off instantaneously at time intervals T. Therefore, it is said that the dynamics described by Hamiltonian [15] is kicked. It is easy to integrate the Hamiltonian equations of motion and find that the evolution from time tT^- (prior to tth kick) to time $(t+1)T^-$ (prior to the $(t+1)$th kick) is described by the map

$$I_{t+1} = I_t + F(\theta_t), \quad \theta_{t+1} = \theta_t + TI_{t+1} \qquad [16]$$

where $F(\theta) = -\mathrm{d}V(\theta)/\mathrm{d}\theta$ is the force acting on the particle. In the following, the special case $V(\theta) = -k(\theta - \pi)^2/2$ is considered. This map is called the sawtooth map, since the force $F(\theta) = -\mathrm{d}V(\theta)/\mathrm{d}\theta = k(\theta - \pi)$ has a sawtooth shape, with a discontinuity at $\theta = 0$. By rescaling $I \to J = TI$, the classical dynamics is seen to depend only on the parameter $K = kT$. The classical motion is stable for $-4 \leqslant K \leqslant 0$ and completely chaotic for $K < -4$ and $K > 0$. For any $K > 0$, one has normal diffusion in the action (momentum) variable: $\langle (\Delta J_t)^2 \rangle \approx D(K)t$, where $\Delta J \equiv J - \langle J \rangle$ and $\langle \ldots \rangle$ denotes the average over an ensemble of trajectories. For $K > 1$, $D(K) \approx (\pi^2/3)\ (K^2)$.

The quantum version of the sawtooth map is obtained by means of the usual quantization rules, $\theta \to \theta$ and $I \to I = -i\partial/\partial\theta$ (set $\hbar = 1$). The quantum evolution in one map iteration is described by a unitary operator U acting on the wave function ψ:

$$\psi_{t+1} = U\psi_t = U_T U_k \psi_t,$$

$$U_T = \mathrm{e}^{-\mathrm{i}TI^2/2}, \quad U_k = \mathrm{e}^{\mathrm{i}k(\theta-\pi)^2/2} \qquad [17]$$

The effective Planck constant is given by $\hbar_{\mathrm{eff}} = T$ and the classical limit corresponds to $k \to \infty$ and $T \to 0$ while keeping $K = kT$ constant.

In the following, an exponentially efficient quantum algorithm for simulation of the map [17] is

described. It is based on the forward/backward quantum Fourier transform between momentum and angle bases. Such an approach is convenient since the operator U is the product of the two operators U_k and U_T, which are diagonal in the θ and I representations, respectively. This quantum algorithm requires the following steps for one map iteration:

1. Apply U_k to the wave function $\psi(\theta)$. In order to decompose the operator U_k into one- and two-qubit gates, first of all, θ is written in binary notation:

$$\theta = 2\pi \sum_{j=1}^{n} \alpha_j 2^{-j} \qquad [18]$$

with $\alpha_i \in \{0, 1\}$. Here n is the number of qubits, so that the total number of levels in the quantum sawtooth map is $N = 2^n$. One can insert eqn [18] into the unitary operator U_k, obtaining the decomposition

$$e^{ik(\theta-\pi)^2/2} = \prod_{i,\, j=1}^{n} e^{i2\pi^2 k\left(\alpha_i 2^{-i} - (1/2n)\right)\left(\alpha_j 2^{-j} - (1/2n)\right)} \qquad [19]$$

which is the product of n^2 two-qubit gates, each acting nontrivially only on the four-dimensional sub-space spanned by the qubits i and j.
2. The change from the θ to the I representation is obtained by means of the quantum Fourier transform, which requires and $n(n+1)/2$ elementary quantum gates.
3. In the I representation, the operator U_T has essentially the same form as the operator U_k in the θ representation, and therefore it can be decomposed into n^2 two-qubit gates, similar to eqn [19].
4. Return to the initial θ representation by application of the inverse quantum Fourier transform.

Thus, overall, this quantum algorithm requires $3n^2 + n$ gates per map iteration. This number is to be compared with the $O(n2^n)$ operations required by a classical computer to simulate one map iteration by means of a fast Fourier transform. Thus, the quantum simulation of the quantum sawtooth map dynamics is exponentially faster than any known classical algorithm. Note that the resources required for the quantum computer to simulate the evolution of the sawtooth map are only logarithmic in the system size N.

As an example of the efficiency of this quantum algorithm, **Figure 1** shows the Husimi functions, taken after 1000 map iterations, for $K = -0.1$, $-\pi \leqslant J < \pi$ (torus geometry). The classical limit is obtained exponentially fast by increasing the number of qubits n, with $\hbar_{\text{eff}} = T = 2\pi/2^n$ ($k = K/T, -N/2 \leqslant I < N/2$). For this value of K, the classical limit is characterized by a complex hierarchical structure of the phase space down to smaller and smaller scales and by anomalous diffusion in the variable J. It is noted that $n = 6$ qubits are sufficient to observe the quantum localization of the anomalous diffusive propagation through hierarchical integrable islands. At $n = 9$, one can see the appearance of integrable islands, and at $n = 16$ the quantum Husimi function explores the complex hierarchical structure of the classical phase space.

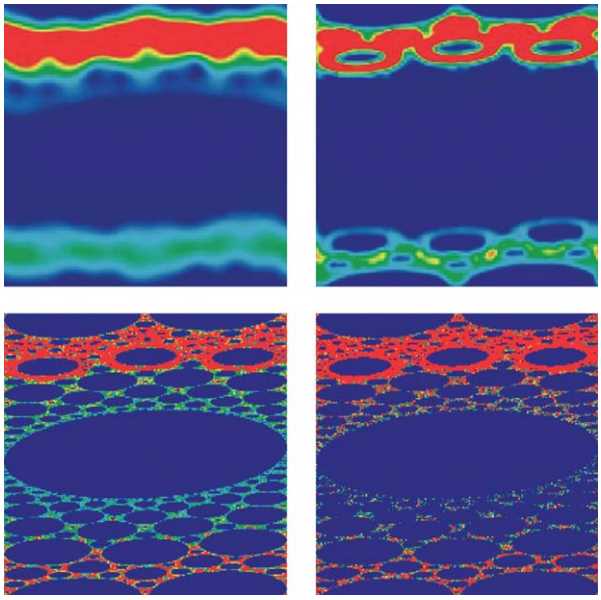


Figure 1 Husimi function for the sawtooth map in action angle variables (J, θ), with $-\pi \leqslant J < \pi$ (vertical axis) and $0 \leqslant \theta < 2\pi$ (horizontal axis), averaged in the interval $950 \leqslant t \leqslant 1000$, for $K = -0.1$, $T = 2\pi/2^n$, $n = 6$ (top left), 9 (top right), 16 (bottom left), and classical density distribution (bottom right). A momentum eigenstate, $|\psi(0)\rangle = |m_0\rangle$, with $m_0 = [0.38 \times 2^n]$ is considered as initial state at time $t = 0$; the classical plot is obtained from an ensemble of 10^8 trajectories, with initial momentum $J_0 = 0.38 \times 2\pi$ and random angles. The color is proportional to the density: blue for zero and red for maximal density.

However, there is an additional aspect to be taken into account. Any quantum algorithm has to address the problem of efficiently extracting useful information from the quantum computer wave function. Indeed, the result of the simulation of a quantum system is the wave function of this system, encoded in the n qubits of the quantum computer. The problem is that, in order to measure all $N = 2^n$ wave function coefficients by means of standard polarization measurements of the n qubits, one has to repeat the quantum simulation a number of times exponential in the number of qubits. This procedure would spoil any quantum algorithm, even in the case, like the present one, in which such algorithm could

compute the wave function with an exponential gain with respect to any classical computation. Nevertheless, there are some important physical questions that can be answered in an efficient way. Two significant examples are discussed.

Dynamical localization is one of the most interesting phenomena that characterize the quantum behavior of classically chaotic systems: quantum interference effects suppress chaotic diffusion in momentum, leading to exponentially localized wave functions. Dynamical localization can be studied in the sawtooth map model. In this case, map [17] is studied on the cylinder $[I \in (-\infty, +\infty)]$, which is cut-off at a finite number N of levels due to the finite quantum (or classical) computer memory. Similar to other models of quantum chaos, quantum interference in the sawtooth map leads to suppression of classical chaotic diffusion after a break time $t^\star \sim D_I$ where $D_I \approx \langle (\Delta I_t)^2 \rangle / t$ is the classical diffusion coefficient. For $t > t^\star$, while the classical distribution goes on diffusing, the quantum distribution reaches a steady state which decays exponentially over the momentum eigenbasis:

$$W_m \equiv |\langle m|\psi\rangle|^2 \approx \frac{1}{\ell} \exp\left[-\frac{2|m - m_0|}{\ell}\right] \quad [20]$$

where $\ell \sim D_I$ is known as the localization length and m_0 is the initial value of the momentum (the index m singles out the eigenstates of I, that is, $I|m\rangle = m|m\rangle$). Therefore, for $t > t^\star$ only $\sqrt{\langle (\Delta I)^2 \rangle} \sim \ell$ levels are populated. For the sawtooth map, $\ell \approx D_I \approx (\pi^2/3)k^2$. **Figure 2** demonstrates that exponential localization, obtained using the above-described quantum algorithm, can already be clearly seen with $n = 6$ qubits. The exponentially localized distribution, appearing at $t \approx t^\star \approx 10$, is frozen in time, apart from quantum fluctuations. The freezing of the localized distribution can be seen from comparison of the probability distributions taken immediately after $t^\star$ and at a much larger time $t = 300 \approx 25t^\star$.

The localization length can be measured by running the quantum algorithm described above, repeatedly up to time $t > t^\star$. Each run is followed by a standard projective measurement on the computational (momentum) basis. Since the wave function at time t can be written as $|\psi(t)\rangle = \sum_m \hat{\psi}(m,t)|m\rangle$, such a measurement gives outcome $\bar{m}$ with probability $W_{\bar{m}} = |\hat{\psi}(\bar{m},t)|^2$. The results of the measurements are stored in a histogram and then the localization can be extracted from a fit of the exponential decay of this histogram over the momentum basis. The number of runs and measurements required to estimate ℓ depends on the desired accuracy but not on ℓ itself. Let the gain of quantum computation of the

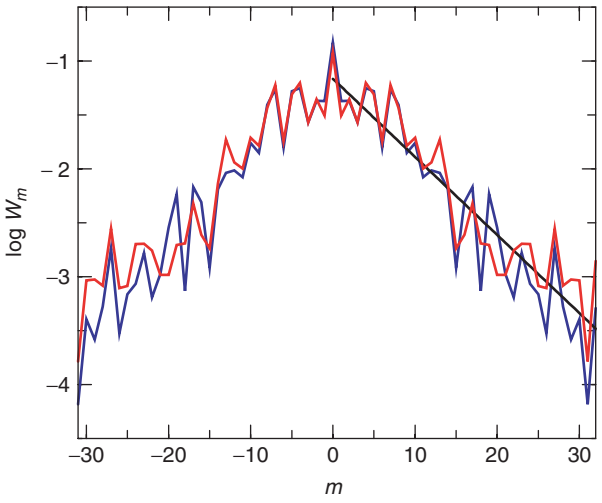


Figure 2 The probability distribution over the momentum basis for the sawtooth map with $n = 6$ qubits, $k = \sqrt{3}$, $K = \sqrt{2}$, and initial momentum $m_0 = 0$; the time average is taken in the intervals $10 \leqslant t \leqslant 20$ (red curve) and $290 \leqslant t \leqslant 300$ (blue curve). The straight line fit, $W_l \propto \exp(-2|m|/\ell)$, gives a localization length $\ell \approx 12$. Note that the logarithm is base ten.

localization length be estimated with respect to classical computation. First of all, recall that it is necessary to make about $t^\star = O(\ell)$ map iterations to obtain the localized distribution. This is true, both for the present quantum algorithm and for classical computation. It is reasonable to use a basis size $N = O(\ell)$ to detect localization (say, N equal to a few times the localization length). In such a situation, a classical computer requires $O(\ell^2 \log \ell)$ operations to extract the localization length, while a quantum computer would require $O(\ell(\log \ell)^2)$ elementary gates. Indeed, both classical and quantum computers need to perform $t \approx t^\star = O(\ell) = O(N)$ map iterations. Therefore, the quantum computer provides a quadratic speed-up in computing the localization length. It is also interesting to notice that, even though the speed-up is only quadratic, the advantage of the quantum computer in memory resources remains exponential. This point becomes crucial in many-dimensional or many-body physical models, in which, using a classical computer, it is very hard, if not impossible, to merely establish if the system is localized or not.

One may ask the following question: why is the speed-up in computing the localization length not exponential, even though a single step of the dynamical evolution can be simulated with exponential speed-up? The ultimate reason is that in the diffusive process that precedes localization, the spread of information is "slow." More precisely, the number N of quantum levels which are significant to describe the system's dynamics grow only as a square root of time. Therefore, in order to use N levels, the

dynamics must be iterated up to a timescale which is $\propto \sqrt{N}$, that is, exponential in the number of qubits.

The quantum computation can provide an exponential gain (with respect to any known classical computation) in problems that require the simulation of dynamics up to a time t which is independent of the number of qubits. In this case, provided that one can extract the relevant information in a number of measurements polynomial in the number of qubits, one should compare $O(t(\log N)^2)$ elementary gates (quantum computation) with $O(tN \log N)$ elementary gates (classical computation). This is the case of dynamical correlation functions of the form

$$C(t) \equiv \langle \psi | A^{\dagger}(t) B(0) | \psi \rangle = \langle \psi | (U^{\dagger})^t A^{\dagger}(0) U^t B(0) | \psi \rangle \qquad [21]$$

where U is the time-evolution operator [17] for the sawtooth map. Similarly, one can efficiently compute the fidelity of quantum motion, which is a quantity of central interest in the study of the stability of quantum motion under perturbations. The fidelity $f(t)$ (also called the Loschmidt echo), measures the accuracy with which a quantum state can be recovered by inverting, at time t, the dynamics with a perturbed Hamiltonian. It is defined as

$$f(t) = \langle \psi | (U_{\varepsilon}^{\dagger})^t U^t | \psi \rangle = \langle \psi | e^{iH_{\varepsilon}t} e^{-iHt} | \psi \rangle \qquad [22]$$

Here the wave vector $|\psi\rangle$ evolves forward in time with Hamiltonian H up to time t and then evolves backward in time with a perturbed Hamiltonian H_{ε}. If the evolution operators U and U_{ε} can be simulated efficiently on a quantum computer, as is the case in many physically interesting situations, then the fidelity of quantum motion can be evaluated with exponential speed-up with respect to known classical computations. The same conclusion is valid for the correlation functions [21].

Quantum Noise

Any practical implementation of a quantum computer will have to face errors, due to the inevitable coupling of quantum processors to the surrounding environment or to device imperfections. The first kind of error is known as decoherence and is a threat to the actual implementation of any quantum computation. More generally, decoherence theory has a fundamental interest beyond quantum information science, since it provides explanations for the emergence of classicality in a world governed by the laws of quantum mechanics. The core of the problem is the superposition principle, according to which any superposition of quantum states is an acceptable quantum state. This entails consequences that are absurd according to classical intuition, like the superposition of "cat alive" and "cat dead" that is considered in the Schrödinger's cat paradox. The interaction with the environment can destroy the coherence between the states appearing in a superposition (for instance, the "cat alive" and "cat dead" states). Therefore, decoherence invalidates the quantum superposition principle, which is at the heart of the power of quantum algorithms.

The presence of device imperfections, although not leading to any decoherence, also hinders the implementation of any quantum computational task. Imperfection effects can be modeled as follows: the quantum computer is seen as a lattice of interacting spins (qubits) where, due to the unavoidable presence of imperfections, the spacing between the up and down states and the couplings between the qubits are both random. The Hamiltonian of this model reads, for a linear array of n qubits, as follows:

$$H_S = \sum_i (\Delta_0 + \delta_i)\sigma_i^z + \sum_{i<j} J_{ij}\sigma_i^x \sigma_j^x \qquad [23]$$

where the σ_i are the Pauli matrices for the qubit i, and Δ_0 is the average level spacing for one qubit. The second sum in eqn [23] runs over nearest-neighbor qubit pairs and δ_i, J_{ij} are randomly and uniformly distributed in the intervals $[-\delta/2, \delta/2]$ and $[-J, J]$, respectively. It is assumed that the phase accumulation given by Δ_0 can be eliminated by standard spin echo techniques, while the other terms give unwanted phase rotations and qubit couplings. In **Figure 3**, the limits to quantum computation due to hardware

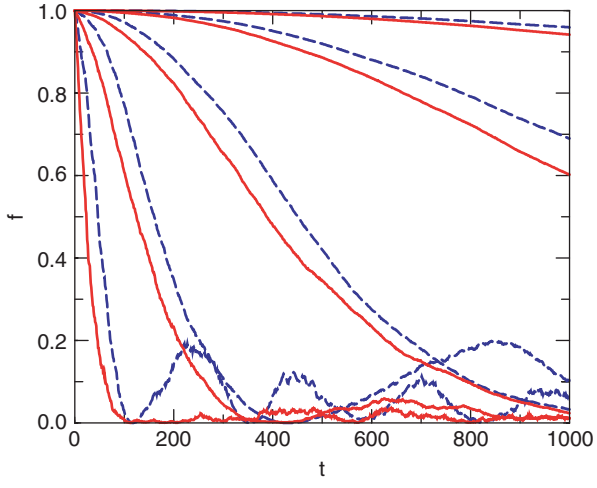


Figure 3 Fidelity as a function of t, for the sawtooth map with parameter values as in **Figure 1**, $n = 9$ qubits, $J = \delta$ (red lines) and $J = 0$ (blue lines). From top to bottom: imperfection strength $\varepsilon = \delta\tau_g = 10^{-5}$, 3×10^{-5}, 10^{-4}, 3×10^{-4}, 10^{-3}. Here τ_g denotes the time interval between consecutive quantum gates.

imperfections in a concrete example are illustrated, the quantum algorithm simulating the sawtooth map. The fidelity of quantum computation, defined by $f(t) = |\langle \psi_\varepsilon(t)|\psi_0(t)\rangle|^2$ is plotted, where $|\psi_\varepsilon(t)\rangle$ is the actual quantum wave function in the presence of imperfections and $|\psi_0(t)\rangle$ is the quantum state for a perfect computation. One can show that the fidelity drops with time according to the law $f(t) \approx \exp(-At^2)$, as expected, for small ε, from perturbation theory. The constant $A \propto n_q(\varepsilon n_g)^2$, with $n_g = 3n^2 + n$ number of quantum gates per map iteration. Therefore, a reliable quantum computation is possible up to a timescale $t_f \propto \varepsilon^{-1} n_g^{-1} n_q^{-1/2}$. A longer quantum computation would necessarily require the implementation of quantum error correcting codes.

Final Remarks

A few significant examples have been discussed showing the capabilities of a quantum computer in the simulation of complex physical systems. A quantum computer with a few tens of qubits, if constructed, would outperform a classical computer in these kinds of problems. Many proposals have been put forward to build a real quantum computer, ranging from atomic physics to solid-state devices. At present, it has been possible to demonstrate several quantum algorithms using NMR quantum processors (with up to seven qubits) and cold ions in a trap, while, in the solid-state arena, two-qubit conditional gate operation has been demonstrated using a pair of coupled superconducting qubits. In these implementations, quantum noise effects allow a reliable quantum computation only up to a few tens of quantum gates, in the best cases. Therefore, decoherence and imperfection effects appear to be the ultimate obstacle to the realization of a large-scale quantum computer.

See also: Nonlinear Dynamics and Nonlinear Dynamical Systems; Quantum Mechanics: Foundations.

PACS: 03.67.Lx; 05.45.Mt; 05.45.Pq

Further Reading

Alber G, Beth T, Horodecki M, Horodecki P, Horodecki R, *et al.* (2001) *Quantum Information – An Introduction to Theoretical Concepts and Experiments*. Heidelberg: Springer.

Benenti G, Casati G, and Strini G (2004) *Principles of Quantum Computation and Information*, *vol. I*: *Basic Concepts*. Singapore: World Scientific.

Bennett CH and DiVincenzo DP (2000) Quantum information and computation. *Nature* 404: 247.

Bouwmeester D, Ekert A, and Zeilinger A (eds.) (2000) *The Physics of Quantum Information*. Heidelberg: Springer.

Casati G and Chirikov BV (1995) *Quantum Chaos: Between Order and Disorder*. Cambridge: Cambridge University Press.

Ekert A, Hayden PM, and Inamori H (2001) Basic Concepts in quantum computation, quant-ph/0011013. In: Kaiser R, Westbrook C, and David F (eds.) *Coherent Atomic Matter Waves*, Les Houches Summer Schools, Session LXXII. Heidelberg: Springer.

Galindo A and Martin-Delgado MA (2002) Information and computation: classical and quantum aspects. *Review of Morden Physics* 74: 347.

Haake F (2000) *Quantum Signatures of Chaos*, 2nd edn. Heidelberg: Springer.

Nielsen MA and Chuang IL (2000) *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

Preskill J (1998) *Lecture Notes on Quantum Information and Computation*. Available at http://theory.caltech.edu/people/preskill.

Shepelyansky DL (2001) Quantum chaos and quantum computers. *Physica Scripta* T90: 112.

Steane A (1998) Quantum computing. *Reports of Progress in Physics* 61: 117.

Zurek WH (2003) Decoherence, einselection, and the quantum origins of the classical. *Review of Modern Physics* 75: 715.

Quantum Devices of Reduced Dimensionality

M Grundmann, Universität Leipzig, Leipzig, Germany

Introduction

A "quantum device" could be defined as a device whose functionality or principle of operation depends essentially on quantum mechanical effects. However, this definition somehow seems too vague. For example, a chair obtains its holding power from the quantum mechanical overlap of the wave functions of its constituents or a toaster relies on quantum mechanical scattering processes for converting electrical current into heat. Therefore a "quantum device" shall be defined here as a "man-made device that intentionally employs or harnesses quantum mechanical effects for its operation." This article is limited to condensed matter devices, thus completely (but intentionally) bypassing the "natural" use of quantum mechanics in living organisms.

A large number of devices already function based on the control of quantum mechanical interactions or effects. Prominent examples are the LASER (light

amplification by stimulated emission of radiation) and the SQUID (superconducting quantum interference device). A possibly larger number awaits its commercialization or even their invention. This section discusses, in particular, quantum devices with "reduced dimensionality." These shall be understood as devices that rely on structures that confine charge carriers in one, two, or all three dimensions. Such systems are called quantum films or wells, quantum wires, and quantum boxes, respectively. Given the typical de Broglie wavelength of electrons in semiconductors, typical sizes of such structures are in the 1–10 nm range. It is noted that currently (2004) the smallest feature size in mass-market consumer electronics is below 100 nm. The predicted miniaturization will hit the 10 nm mark in the years beyond 2010. Therefore quantum effects seem unavoidable in mainstream microelectronics.

The natural enemies of quantum devices are dephasing events such as scattering of an otherwise coherently evolving state, increasing temperature, either due to higher scattering rate or the modification of thermodynamic distribution functions, and increasing structural size, mostly leading to a transition to classical behavior.

Quantum Mechanical Effects

The quantum effects that are used in quantum devices are sorted by increasing order of complexity and decreasing order of fundamental understanding and employment in devices:

1. single particle effects,
2. Pauli's exclusion principle,
3. quantization of electronic states in confining potentials,
4. tunneling effect,
5. spatial and temporal coherence of the wave function,
6. spin,
7. quantum liquids, and
8. entanglement.

Examples for the use of these effects in condensed matter devices is discussed below. The list, however, may not be exhaustive as eventually any quantum mechanical effect, for example, the quantization of other degrees of freedom or still unexplored effects, may be used in devices.

Quantization in Potential Wells

Quantum devices of reduced dimensionality depend some way or another on the presence of potential wells that limit the motion of charge carriers in c dimensions compared to bulk material. The motion can be limited in one dimension within a quantum film or well (QW), in two dimensions within a quantum wire (QWR), or in all three dimensions within a quantum dot (QD) or box. These structures can be combined to double or multiple structures or even periodic arrays (**Table 1**).

Table 1 Structures with two-, one-, and zero-dimensional confinement

Confinement of motion to c dimensions	*Name*	*Combination of two*	*Periodic structure*
2	Quantum well	Double quantum well	Superlattice
1	Quantum wire	Quantum wire pair	Quantum wire array
0	Quantum dot	Quantum dot molecule	Quantum dot chain

The confining potential can arise from applied spatially modulated electric potentials or the effect of band offsets in heterostructures or a combination of both. The quantization depends on the size, shape, and barrier height of the confining potential. By controlling the barrier height between confinement structures the degree of quantum mechanical coupling can be varied.

The simplest picture arises from the particle in a rectangular box with side lengths L_x, L_y, and L_z and infinite barriers. The quantization energy is given by

$$E = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right)$$

with m being the effective mass of the particle (here taken as isotropic) and the n_i being the quantum numbers ($n_i \geqslant 1$).

If the confining potential is harmonic, the energy levels are equidistant.

$$E = (d/2 + n)\hbar\omega$$

$d = 3 - c$ shall be the dimension of the confinement potential with c being the number of confined dimensions. For a parabolic confinement potential the generalized Kohn's theorem applies, which states that the resonance energies of many electrons in the dot are the same as for a single electron and that dipole radiation interacts only with the center of mass motion of the electrons.

Since self-assembled nanostructures often involve strained systems, the confining potential depends on

the local strain and the resulting piezoelectric field which may be spatially strongly inhomogeneous.

The finite barrier height limits the number and confinement energy of confined states. In quantum dots, there is a critical minimum size for which at least one state is confined. Quantum dots smaller than the critical size do not exhibit any confined states, as opposed to quantum wells that always have at least one localized and quantized state.

Two-Dimensional Carrier Gases

A particularly important confined electron system is the two-dimensional electron gas (2DEG). Such a gas can develop in a quantum well. However, a single interface between two semiconductors of different bandgaps is sufficient to cause the creation of a 2DEG. If modulation doping, that is spatially inhomogeneous doping, is used, a high carrier concentration can be achieved in the 2DEG without the dopant atoms being located in or close to the 2DEG. In this way, impurity scattering is avoided and a high mobility is achieved. For GaAs/AlGaAs structures the mobility record (2004) stands at 3.1×10^7 Vs cm^{-2} at low temperatures. Thus, 2DEG and its manipulation are the basis for many quantum devices.

In a high electron mobility transistor (HEMT), the transport in a 2DEG is used as channel conductivity in a field effect transistor. HEMTs display high transconductance and low noise figure and are used for many high-frequency applications such as satellite receivers. In order to maximize the carrier confinement, pseudomorphic quantum wells, for example, InGaAs/AlGaAs on GaAs, are used. The highest performance is achieved with InGaAs/InAlAs quantum wells on InP substrate. In order to transfer this technology to (larger and cheaper) GaAs substrates, a metamorphic, dislocated buffer is grown, and such a device is called MHEMT. Similar strategies apply to SiGe/Si HEMTs.

Magnetotransport effects in 2DEGs at low temperature have been studied in detail for a while. In the integral quantum Hall effect (IQHE) the conductivity of the 2DEG exhibits extended plateaus for a range of magnetic fields with a well-defined conductivity value σ_{xy} that takes integer multiples of e^2h^{-1}. This quantum mechanical value is found to be independent from details of the sample (within reasonable spread) such as doping, layer thickness, and growth method. The IQHE can be used for fabrication of a standard for the resistance and the unit "ohm." The precision of the QHE normal is two orders of magnitude better than the realization of the ohm in the SI system. For this purpose, parallel and serial circuits of QHE resistors can also be used and an AC measurement technique can be devised.

It is to be noted that more complicated phenomena, the fractional quantum Hall effect, are observed and discovered in such quantum liquids at even more extreme conditions (very low temperatures and high magnetic fields).

Quantum wells are also very successfully used as active mediums in semiconductor laser diodes. Charge carriers are captured efficiently from the barrier material (high bandgap) into a quantum well or multiple quantum wells (low bandgap). Due to their small thickness a high-density carrier gas is easily achieved and inversion is reached quickly. The finite density of states at the sub-band edges provides a favorable gain spectrum. This leads ultimately to low lasing threshold current of $\sim 40\,\mathrm{A\,cm^{-2}}$. Their combination with high-finesse optical cavities using high-reflectivity dielectric layer packages (Bragg mirrors) leads to vertical surface emitting lasers (VCSEL) and tests the use of quantum electrodynamic (QED) effects (e.g., the modification of spontaneous emission lifetime – Purcell effect).

Quantum Wires

Many device applications have been envisioned for quantum wires. Since they are positioned in terms of reduced dimensionality between quantum wells and quantum dots, for some properties the ultimate solution is quantum dots. Since quantum dots can be fabricated in a self-assembled manner, they have made some of the quantum wire devices obsolete. However, quasi-one-dimensional conduction poses interesting physics (e.g., reduced scattering and Luttinger liquids). Quantum wires can be used as channels in transistors. In order to draw a sufficient current, many wires may have to be used in parallel. Due to the one-dimensional k-space, carrier scattering is reduced and thus a higher mobility should be possible. Recent progress in the self-assembled fabrication of quantum wire nanowhiskers and heterostructures incorporated in them makes them seem attractive for a variety of applications including LEDs, lasers, and sensors.

Quantum Dots

Quantum dots confine carriers in all three spatial dimensions. These are therefore sometimes termed "artificial atoms." The embedding into a solid-state matrix makes them more readily accessible to manipulation with electrodes, currents, fields, etc. than atoms in a trap. However, this comes at the price of stronger (compared to ions in a trap) coupling to external degrees of freedom, such as charge fluctuations and thermal bath.

The primary effect of quantum confinement is to create singular density of states at the quantized levels. The energetic broadening for a single quantum dot is typically small at low temperature (μeV range). With increasing temperature, additional dephasing mechanisms can increase the homogeneous broadening to the 10 meV range. An ensemble of quantum dots exhibits inhomogeneous broadening because the individual nanostructures are not exactly identical with respect to size, shape, and possibly chemical composition.

Devices based on single dots have been envisioned, realized, and investigated for storage of single electrons, emission of single or entangled photons, and quantum information processing.

The storage of single electrons is comparably straightforward. The capture of a single electron causes a second electron to need a larger energy to be on the dot due to Coulomb repulsion. While for typical shallow donors in semiconductors this effect is so strong that only the neutral state (with one electron) is localized, the Coulomb interaction energy in self-assembled quantum dots is typically in the range of 20 meV. At sufficiently low temperature for a given potential landscape, a second electron cannot enter the dot (Coulomb blockade) and the dot charge is fixed to an integer number of electrons, possibly one or zero. The storage of single electrons has been successfully demonstrated at low temperatures. Efforts are underway to maintain the effects up to liquid nitrogen temperature. The dots are typically built into a transistor structure to make a single electron transistor (SET). Here the gate electrode contains the quantum dot close to the channel. Due to electrostatic interaction the source–drain channel only becomes conductive after a sufficient source–drain voltage larger than the voltage related to the Coulomb energy has been applied. Conversely, in a single-electron memory, the conductance of the SET can be used to test whether a single electron is located on the quantum dot or not. By periodically modulating the applied bias voltages in a suitable manner with a frequency f, a single electron can be transferred by tunneling onto and off the dot from source to drain for each cycle. Such a device is called turnstile device and can be used to define a current normal via $I = fe$.

Quantum dots can be used in what is called a "photon gun." In such a device a trigger impulse causes the reliable emission of a single photon. This is a nonclassical form of light which cannot be produced from dimming a conventional source. In classical light the time difference between photons follows Poisson statistics while a photon gun is supposed to create a periodic stream of single photons. With such photons, higher data rates for quantum cryptographic transmissions can be realized. Upon trigger the pulse carriers are injected into the quantum dot. Emission on the single exciton line occurs only once for each trigger with only a small time delay between trigger and emission.

The emission of entangled photons is possible if the cascade relaxation from biexciton to exciton to empty quantum dot is used. If the single exciton state is degenerate with respect to the two in-plane polarization directions, the photons are entangled and can be used for quantum cryptographic information transmission schemes. The polarization degeneracy calls for quantum dots with certain geometrical symmetry as strain, piezoelectric fields, and geometric asymmetries contribute to the lifting of the degeneracy. The two emitted photons have different energies unless the biexciton binding energy is zero. This is generally possible and depends on the size and shape of the quantum dot.

Several schemes have been devised to use quantum dots in quantum computing schemes. The spin degree of freedom seems a good candidate for quantum computational operations due to its long dephasing time of up to several milliseconds. Many of the current experiments involve optical sampling of the quantum information with polarized light and suitable pulses. Eventually electronic means will be more useful to access the quantum dots. This has been accomplished for quantum dots defined with in-plane gates, but only at very low temperatures. Using clever combinations of the Coulomb blockade mechanisms and external bias, the occupation and the spin orientation of a single electron can be tested electronically at low temperatures. Quantum computing schemes based on cellular automata (QCA) have been proposed extensively to realize quantum computational algorithms. QCA rely on arrays of quantum dots that interact to allow for switching of their polarization state.

Quantum dots can be used as active-gain medium in semiconductor diode lasers, which record small transparency current densities due to the fairly small volume for carrier inversion. Till date, reasonable laser performance has been obtained by using only self-assembled quantum dots.

The discrete density of states leads, in an ideal situation, to very large T_0 values (small temperature-dependence of the threshold current) and zero alpha-factor (also termed line width enhancement factor). Realistically, the T_0 values have been found to be large (close to infinity) only up to 150–200 K; for higher temperatures the T_0 values are not particularly advantageous compared to conventional quantum well lasers because of loss of carriers from the

quantum dots due to thermal excitations or the Auger effect. Small α-factors have been realized, also leading to a reduction of current filamentation and a homogeneous optical near field. These properties allow for improved high power lasers.

Typically, lasing is designed to occur on the ground state of the quantum dots since this results in the lowest threshold (below $10\,A\,cm^{-2}$) for laser diodes. Excited states provide higher gain and can be made to contribute to the lasing process simultaneously with the ground state. By intentionally increasing the width of the inhomogeneous broadening, for example, by growing vertically stacked dots with slightly different geometrical properties, the gain spectrum can be made very broad. Such gain medium would allow for wide tunability.

The management of strain and chemical composition allows to create 1300 and 1550 nm emission, important for datacom and telecom applications, respectively, on GaAs substrate. In particular, the use of metamorphic buffers has been successful recently. Quantum dot lasers are found to have a strongly improved radiation hardness compared to quantum well lasers.

Table 2 Quantum devices, the effects they are based on, and applications

Device	*Quantum effect*	*Performance/ Use*
High electron mobility transistor	2D confinement	High-speed, low-noise amplifier, low-threshold laser
Quantum well laser	2D confinement	Low threshold
Quantum dot laser	0D confinement	Low threshold, small α
Quantum Hall effect	Quantum liquid	Resistance standard
Electron turnstile	Tunneling, single electrons	Current standard
Photon gun	Single-particle effect	Quantum cryptography
Quantum cellular automata	0D confinement, coherence	Quantum computing
Quantum dot	Entanglement, coherence	Quantum computing, quantum cryptography
Josephson junction	Coherence of wave function	Voltage standard
SQUID	Coherence of wave function	Sensitive magnetometer

Josephson Junctions

A Josephson junction is made up from two superconductors sandwiching a thin nonsuperconducting layer such that electrons can tunnel through the barrier. The coherence of the wave function in the superconductor leads to DC or AC currents. The DC Josephson current is proportional to the phase difference between the two superconductors. The frequency f of the AC Josephson current is proportional to the voltage V applied across the junction, $f = 2\,eV\,h^{-1}$. This allows for a voltage normal based on a frequency. A Josephson junction standard can realize a voltage with an accuracy of 10^{-10}. For a standard in the volt regime several thousand junctions are put in series. A superconducting loop with two Josephson junctions in either arm is very sensitive (10^{-14} T) to the magnetic flux enclosed (SQUID, superconducting quantum interference device).

Summary

Devices relying on quantum effects have a broad spectrum of applications (**Table 2**). Generally it is difficult to achieve operation at ambient temperatures. Great progress has been made in the field of metrology where quantum devices will become standards that can be easily reproduced, at least in a laboratory environment. Room temperature operation has been achieved, for example, for quantum dot lasers with the preservation of several advantages from the quantum nature of the gain medium.

See also: Micromechanical Devices and Systems; Quantum Cascade Lasers; Semiconductor Devices; Semiconductor Nanostructures; Tunneling Devices.

PACS: 85.35.Be; 68.65.La; 73.21.Hb; 73.21.La; 73.63.Kv; 78.67.Hc

Further Reading

Bachmair H, Göbel EO, Hein G, Melcher J, Schumacher B, *et al.* (2003) The von Klitzing resistance standard. *Physica E* 20: 14.

Borri P, Langbein W, Mørk J, Hvam JM, Heinrichsdorff F, *et al.* (1999) *Physical Review B* 60: 7784.

Brey L, Johnson NF, and Halperin BI (1989) *Physical Review B* 40: 10647.

Elzerman JM, Hanson R, Willems van Beveren LH, Witkamp B, Vandersypen LMK, *et al.* (2004) *Nature* 430: 431.

Gehrig E, Hess O, Ribbat C, Sellin RL, and Bimberg D (2004) *Applied Physics Letters* 84: 1650.

Gisin N, Ribordy G, Tittel W, and Zbinden H (2002) *Review of Modern Physics* 74: 145.

Grundmann M (2000) *Physica E* 5: 167.

Kane CL and Fisher MPA (1992) *Physical Review Letters* 68: 1220.

Klitzing Kv, Dorda G, and Pepper M (1980) *Physical Review Letters* 45: 494.

Kouwenhoven LP, van der Vaart NC, Johnson AT, Kool W, Harmans CJPM, *et al.* (1991) *Zeitschrift fürPhysik B* 85: 367–373.

Landwehr G (2003) *Physica E* 20: 1.

Lent CS, Tougaw PD, Porod W, and Bernstein GH (1993) *Nanotechnology* 4: 49.

Luttinger JM (1963) *Journal of Mathematical Physics* 4: 1154.
Santoprete R, Koiller B, Capaz RB, Kratzer P, Liu QKK, *et al.* (2003) *Physical Review B* 68: 235311.
Stier O, Grundmann M, and Bimberg D (1999) *Physical Review B* 59: 5688.
Störmer HL, Tsui D, and Gossard AC (1999) *Review of Modern Physics* 71: S298.
Tomonaga S (1950) *Progress of Theoretical Physics* 5: 544.
Yuan Z, Kardynal BE, Stevenson RM, Shields AJ, Lobo CJ, *et al.* (2002) *Science* 295: 102.

Quantum Hall Effect

J Weis, Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany

Introduction

The quantum Hall effect (QHE) and its relation to fundamental physical constants was discovered in 1980 by Klaus von Klitzing for which he received a Nobel prize in 1985. The Hall resistance R_H (Hall voltage divided by applied current) measured on a two-dimensional charge carrier system at low temperatures (typically at liquid helium temperature $T = 4.2\,\mathrm{K}$) and high magnetic fields (typically several tesla), which is applied perpendicularly to the plane of the charge carrier system, shows well-defined constant values for wide magnetic field or charge carrier density variations. These plateau values are described by $|R_H| = h/(ie^2)$, where h is the Planck constant, e is the elementary charge, and i an integer value with $i = (1, 2, 3, \ldots)$. Nowadays, this effect is denoted as integer quantum Hall effect (IQHE) since, beginning with the year 1982, plateau values have been found in the Hall resistance of two-dimensional electron systems of higher quality and at lower temperature which are described by $|R_H| = h/(fe^2)$, where f is a fractional number. The fractions $f = \{1/3, 2/3\}$ are the most prominent ones. For the discovery of these unexpected new quantum states in 1982, manifesting themselves in the fractional quantum Hall effect (FQHE), Dan C Tsui, Horst L Störmer, and Robert B Laughlin were honored with the Nobel prize in 1998.

The most important implication of the IQHE is its application in metrology where the effect is used to represent a resistance standard. Due to a small standard uncertainty in reproducing the value of the quantized Hall resistance (few parts of 10^{-9} in the year 2003), its value was fixed in 1990, for the purpose of resistance calibration, to $25812.807\,\Omega$ and is nowadays denoted as the conventional von Klitzing constant R_{K-90}.

It is generally accepted that the von Klitzing constant R_K agrees with h/e^2, and is therefore directly related to the Sommerfeld fine-structure constant $\alpha = (\mu_0\, c/2)(e^2/h) = (\mu_0\, c/2)(R_K)^{-1}$, which is a measure for the strength of the interaction between electromagnetic fields and elementary particles (please note, in the International System of Units (SI), the speed of light c in vacuum and the permeability of vacuum μ_0 are defined as fixed physical constants). The IQHE allows one to determine the fine-structure constant α with high precision, simply based on magnetoresistance measurements on a solid-state device.

The FQHE is a manifestation of correlation effects among the charge carriers interacting in the two-dimensional system, which lead to the formation of new quantum states. With an improvement in the quality and reaching lower temperatures for the charge carrier system, more and more quantum Hall states have been found. Quantum Hall systems are, therefore, used as model systems for studying the formation of correlated many-particle states, developing theory for their description, and identifying, probably, their simpler description in terms of the formation of new quasiparticles, for instance, the so-called "composite fermions."

Hall Effect

An electrical charge moving with constant velocity in a homogeneous magnetic field feels the Lorentz force acting in a direction perpendicular to both the magnetic field and the direction of motion. A current through an electrical conductor (an experimental setup is shown in **Figure 1**) leads to the Hall effect discovered by Edwin Hall in 1879. As can be derived within the Drude model, the free charge carriers in the conductor accumulate due to the Lorentz force on one side and deplete on the opposite side leading to an electrical field (Hall field) in the conductor which in back-action compensates for the Lorentz force so that, under stationary conditions, the charge carriers drift straight through the sample. If only one type of charge carrier is present, the related Hall voltage V_H, that is, the difference in the electrochemical potentials between two contacts at opposite sample edges, where the connection line between these two contacts lies perpendicular to the current direction, is simply described by $V_H = V_y = R_H I_x$

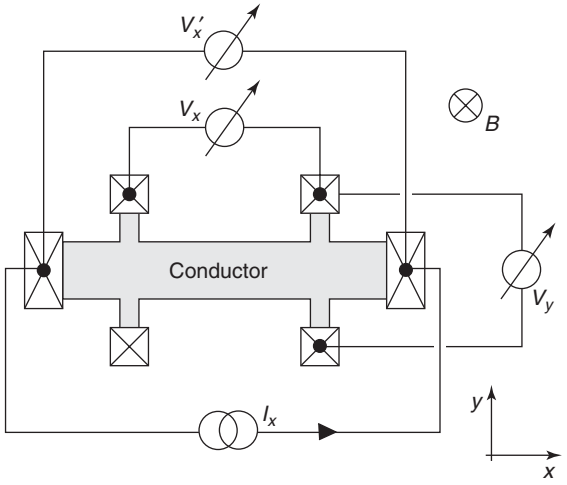


Figure 1 Setup for measuring the Hall resistance $R_{xy} = V_y/I_x$, the (four-terminal) longitudinal resistance $R_{xx} = V_x/I_x$, and the (two-terminal) longitudinal resistance $R'_{xx} = V'_x/I_x$.

with the Hall resistance R_H given by

$$R_H \equiv R_{xy} = B_z/qnd$$

for the orientation of the arrangement shown in **Figure 1**. The Hall resistance is positive or negative depending on the type of charge carriers which are either electrons ($q = -e$) or missing electrons/holes ($q = +e$). Its magnitude increases linearly with the homogeneously applied magnetic flux density B_z and the slope of the increase is given by the concentration n of these carriers and the thickness d of the conducting layer. That is why the Hall effect is used as an important diagnostic tool to determine essential properties of electrical conductors – free charge carrier concentration n and sign of q. Hall devices find various applications as magnetic field sensors.

Two-Dimensional Charge Carrier System

The QHE is observed in two-dimensional charge carrier systems, belonging to the class of so-called low-dimensional charge carrier systems. Generally, the charge carriers in a bulk conductor can be considered as independent charged quasiparticles moving, with an effective mass m^*, freely in the three-dimensional space of the crystal. In the simplest case, such a charge carrier system is treated as the Sommerfeld fermion gas, that is, the wave functions of these quasiparticles are described by plane waves and their eigenenergies by $\varepsilon = |\hbar k|^2/(2m^*)$, where k denotes the wave vector and $\hbar = h/2\pi$. The de Broglie wavelength $\lambda = 2\pi/|k|$ of the charge carriers at their Fermi level in the bulky conductor (= Fermi wavelength λ_F) is given by $\lambda_F = h/\sqrt{(2m^*\varepsilon_F)}$. Confining the charge carriers to spatial dimensions of the length of the Fermi wavelength λ_F, the charge carriers are restricted in their motion. Only certain eigenenergies are possible in the confined direction, and the distance between these eigenenergies increases with narrowing the spatial width of the confinement, until only the lowest eigenenergy is still lying below the Fermi level. Two-dimensional charge systems allow motion only in a plane, one-dimensional charge systems (denoted as quantum wires) only along a line, and in a zero-dimensional system (denoted as quantum dots or "artificial atoms"), the charge carriers are confined in a cage allowing them overall only a discrete set of eigenenergies.

The Fermi wavelength $\lambda_F = h/\sqrt{(2m^*\varepsilon_F)}$ of charge carriers in a bulky conductor is large in the case of a low Fermi energy ε_F, achieved by a low charge carrier concentration n, and in the case of a low effective mass m^*. Therefore, semiconductor and especially III–V compound semiconductor materials (for gallium arsenide (GaAs), $m^* = 0.067\,m_0$, where m_0 is the free electron mass) are preferentially used to define low-dimensional charge carrier systems by confining conduction band electron or valence band holes in one or more directions to a few tens of nanometers or less. The first two-dimensional charge carrier systems were realized in metal-oxide-semiconductor field effect transistors (MOSFETs) in the thin conducting channel at the insulator–semiconductor interface. Silicon-MOSFETs are used as electronic switches of which, nowadays, 100 millions are integrated on a single microprocessor die. The QHE was originally discovered on such an Si-MOSFET device. Two-dimensional charge carrier systems of higher quality are obtained in layers or at interfaces of III–V compound semiconductor heterostructures, which are fabricated by molecular beam epitaxy (MBE) or metal organic chemical vapor deposition (MOCVD). III–V semiconductor heterostructures find various applications as base material for low-noise and high-frequency transistors or as optoelectronic devices such as laser-diodes which belong, nowadays, to key devices in modern communication technology.

The highest quality for two-dimensional electron systems is obtained in GaAs/$Al_xGa_{1-x}As$ heterostructures based on modulation-doping. (The quality of conductors is characterized by the charge carrier mobility μ, which relates, without an applied magnetic field, the drift velocity v_D of the free charge carriers to the applied electric field E by $v_D = \mu E$. As derived from the Drude model, the mobility is obtained from measuring the electrical conductivity σ via $\sigma = q\mu n$. The larger the absolute value of the mobility μ, the larger the distance a charge carrier

can move without being scattered. Due to the spatial separation of dopants and free charge carriers in modulation-doped heterostructures, the scattering of the charge carriers on the ionized dopants is strongly reduced. At temperatures below 1 K, recently, electron mobilities up to $3 \times 10^3\ \mathrm{m^2/(Vs)} = 3000\ \mathrm{1/T}$ have been reported in electron systems at such a heterojunction, allowing free mean paths of several 100 µm. At room temperature, such devices are used as high-electron mobility transistors (HEMT) in communication technology. As shown in **Figure 2**, the donors are located in an $Al_xGa_{1-x}As$ layer, which has a conduction band bottom lying higher than that of bulk GaAs. The band offset is tunable by the ratio x by which gallium atoms are replaced by aluminum atoms on their crystal lattice sites. The electrons separate from their donors and accumulate in the GaAs at the heterojunction, because they are still attracted by the positively charged donor ions. Each electron feels a triangular-shaped confining potential which restricts the electron motion to a free motion parallel to the heterojunction plane. If the electron density and the temperature are low enough, then only the lowest subband with energy $\varepsilon_{z,0}$ is occupied and the possible eigenenergies are described by

$$\varepsilon = |\hbar k_{||}|^2/(2m^*) + \varepsilon_{z,0}$$

where $k_{||}$ denotes the wave vector of the charge carrier in the plane of the 2D system. For the two-dimensional charge system with such a parabolic dispersion relation, the density of states $D_{2D}(\varepsilon)$ is constant

$$D_{2D}(\varepsilon) = g_s m^*/2\pi\hbar^2$$

where the factor $g_s = 2$ takes into account the two possible spin orientations for each plane wave state characterized by the wave vector $k_{||}$.

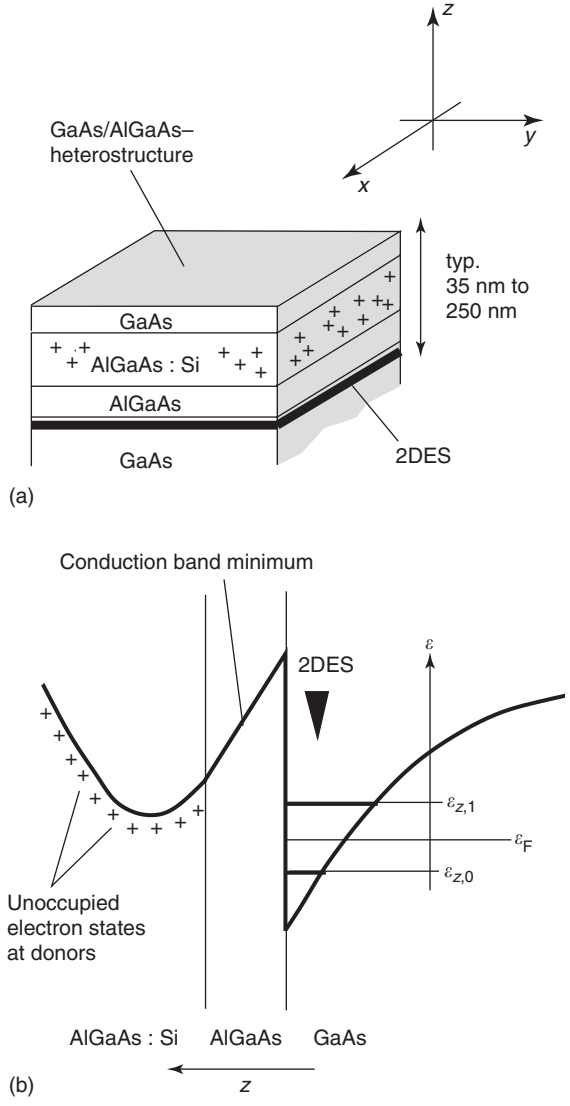


Figure 2 (a) AlGaAs/GaAs heterostructure with a two-dimensional electron system (2DES) at the heterojunction in GaAs. (b) Sketch of the respective profile of the conduction band edge in z direction, and the two lowest subband energies $\varepsilon_{z,0}$ and $\varepsilon_{z,1}$ due to the triangular-shaped confining potential in z direction.

Hall Effect on Two-Dimensional Charge Carrier System – Quantum Hall Effect

In **Figure 3**, Hall resistance curves versus magnetic flux density are shown as they are measured on two-dimensional electron systems of different quality. At low magnetic field, the Hall resistance linearly increases with the magnetic flux density B_z,

$$|R_H| = |R_{xy}| = |V_y/I_x| = |B_z/(-en_s)| \qquad [1]$$

and allows one to determine the respective sheet charge carrier density n_s of the two-dimensional electron system. At higher magnetic field, the Hall effect on the two-dimensional charge system shows, in certain magnetic field ranges, values which are independent of the magnetic field. The values of these plateaus are perfectly described by

$$|R_H| = h/ve^2 \qquad [2]$$

where v is an integer number, $i = \{1, 2, 3, \ldots\}$ or a certain fractional number f which seems to follow $f = p/q$ with $p = \{1, 2, 3, \ldots\}$ and $q = \{3, 5, 7, \ldots\}$, with exceptions like $f = 5/2$. The same plateau values are found by keeping the magnetic flux density B_z constant and varying the sheet electron concentration n_s. Therefore, varying the ratio n_s/B_z in certain ranges allows one to observe Hall resistance plateaus described by eqn [2].

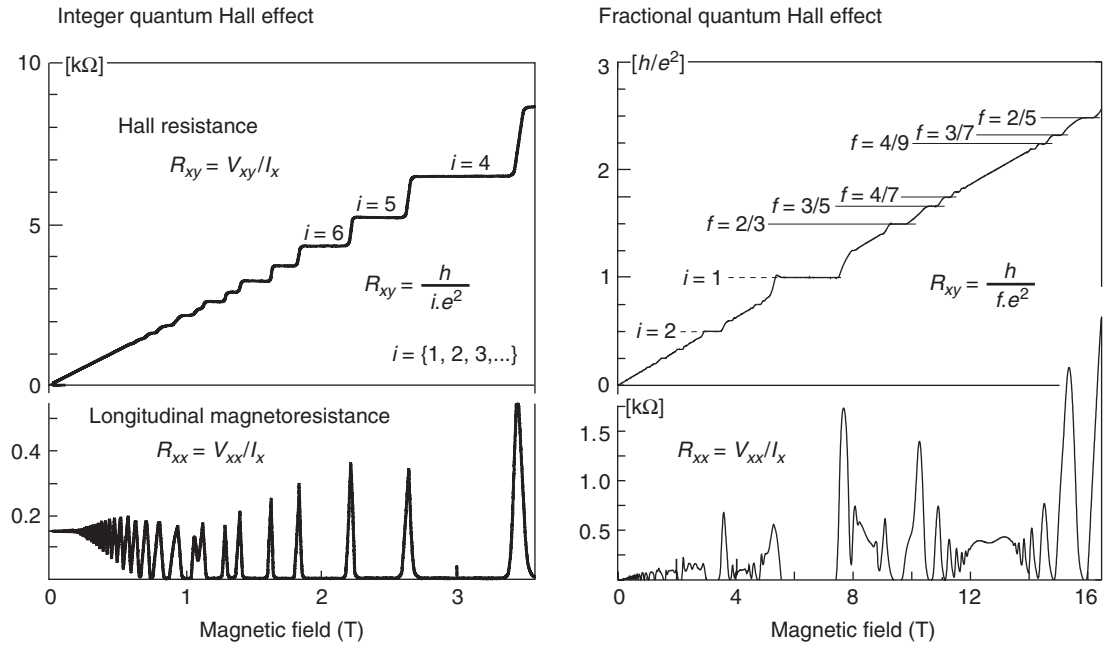


Figure 3 Quantum Hall effect: Hall resistance R_{xy} and longitudinal resistance R_{xx} measured on a 2DES as a function of the applied magnetic field. In certain magnetic field regions, plateaus are observed in R_{xy}, and at the same time R_{xx} vanishes.

Such Hall resistance values are expected due to the comparison between eqns [1] and [2] only for certain well-defined ratios $n_s/B_z = \nu\, e/h$. The meaning of this ratio can be expressed slightly differently: taking a certain area A out of a homogeneous two-dimensional electron system enclosing the integer number $N = n_s A$ of electrons, the magnetic flux penetrating this area is given by $\Phi = AB_z$. The ratio of electron number N and magnetic flux Φ in units of the magnetic flux quantum $\Phi_0 = e/h$ is then

$$\nu = N/(\Phi/\Phi_0) \qquad [3]$$

that is, for each electron, ν^{-1} magnetic flux quanta are present.

Whenever a plateau is found in the Hall resistance, the (four-terminal) longitudinal resistance $R_{xx} = V_x/I_x$ vanishes, that is, $R_{xx} = 0$. The occurrence of the Hall plateau is accompanied by a dissipationless current flowing along the sample. However, this does not mean that there is no dissipation at all. In the plateau regimes, the (two-terminal) resistance $R'_{xx} = V'_x/I_x$ measured between the current biased contacts is about the Hall resistance $|R_H|$, that is, the electrical power $|R_H|\, I_x^2$ is dissipated. The heat is created in two spot-like regions, one located close to each current-biased contact. Along the sample away from the hot spot regions, dissipation is absent, indicated by the absence of a voltage drop along the sample. In the transition regimes between Hall plateaus, dissipation occurs along the whole sample measurable by $R_{xx} \neq 0$.

The Hall plateau values with $i = 2$ and $i = 4$ are reproducible to the standard uncertainty of 10^{-9}, independent of the sheet charge carrier density, the sample geometry, and further properties of the material in which the two-dimensional charge carrier system is embedded. That is why the quantum Hall resistance has been used since 1990 as a resistance standard. Si-MOSFET or GaAs-AlGaAs-HEMT devices are used for this purpose. To trust in the plateau values obtained on a certain sample for metrological application, the flatness of the Hall plateau, the vanishing of the longitudinal resistance $R_{xx} = 0$, and the invariance of the Hall resistance value with changing the temperature are checked. Increasing the quality of the samples, the Hall plateaus get smaller. In consequence, a certain amount of disorder is required to obtain well-defined Hall resistance plateaus and, therefore, accuracy in the quantized Hall resistance value. For metrological application, the sheet density n_s of the two-dimensional electron system in a $GaAs/Al_xGa_{1-x}As$ heterostructure is typically in the range of $3\text{–}6 \times 10^{15}\,m^{-2}$ with a mobility μ of $40\text{–}80\,T^{-1}$, and the measurements are done at the temperature $T = 1.5$ K. For high-precision measurements, it is desirable to have the current level as high as possible. However, the current level is limited for the respective sample, since the QHE breaks down suddenly with increasing the current beyond a certain critical current level. The typical critical sheet

current density I_x/w is ~ 0.5–$1.5\,\mathrm{A\,m^{-1}}$. Also, a significant increase of the working temperature degrades the Hall plateau, and the plateaus finally disappear.

Theoretical Models for the Quantum Hall Effect

A complete microscopic theory, allowing one to derive the Hall plateau values and widths, still does not exist. However, it is commonly accepted that the IQHE can be described within a model of independent electrons whereas for the FQHE, many-particle correlations due to the electron–electron interaction are of importance.

Integer Quantum Hall Effect

In recent years, spatially resolved measurements of the Hall potential profile with low-temperature scanning probe microscopes on two-dimensional electron systems defined in GaAs-AlGaAs-HEMT devices have revealed the following microscopic picture of the IQHE.

The Single-Particle Energy Spectrum – Landau Levels

The magnetic field forces the electrons on cyclotron orbits. Quantum mechanically, the possible eigenenergies cluster in discrete Landau levels described by

$$\varepsilon = \varepsilon_{z,0} + (n + 1/2)\hbar\omega_c + s_z g^* \mu_B B \qquad [4]$$

with the Landau level index $n = \{0, 1, 2, \ldots\}$, and the spin quantum number $s_z = \{-1/2, 1/2\}$. The gaps between the Landau levels are either given by the cyclotron energy $\hbar\omega_c = \hbar eB/m^*$ minus the Zeeman energy $g^*\mu_B B$, or the Zeeman energy (μ_B denotes the Bohr magneton, $\mu_B = e\hbar/(2m_0)$ and g^* the effective Landé factor). The degeneracy $n_L = eB/h$ of each (spin-resolved) Landau level and the cyclotron energy ω_c increases with B. Therefore, with an increasing magnetic field, the electrons are redistributed to lower Landau levels leading to a sawtooth-like variation of the chemical potential $\mu_{ch}(n_s, B)$ for fixed sheet electron density n_s with increasing magnetic field (see **Figure 4**). The ratio of the sheet electron density n_s and Landau level degeneracy n_L is denoted as Landau level filling factor ν_L (or shortly filling factor),

$$\nu_L = n_s/n_L \qquad [5]$$

Whenever the filling factor ν_L takes an integer value, the chemical potential lies energetically between two Landau levels, and in the homogeneous electron system the respective fraction $\nu^{-1} = \{1, 1/2, 1/3, 1/4, \ldots\}$ of a magnetic flux quantum is present for each electron, consistent with eqn [3]. For integer filling factors, the electron system behaves incompressible, that is, the compressibility defined by $\kappa = (\partial\mu_{ch}/\partial n_s)^{-1}$ becomes zero due to the discontinuity in $\mu_{ch}(n_s, B)$ versus n_s.

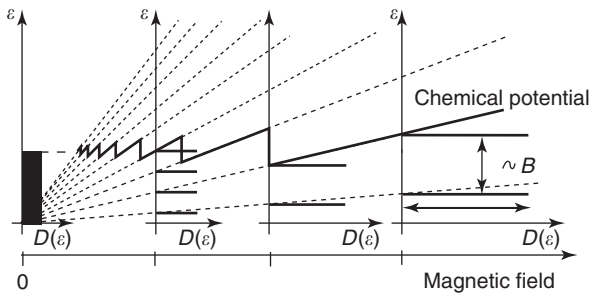


Figure 4 For an infinite homogeneous electron gas, the single-particle eigenenergy spectrum splits up into Landau levels on applying a strong magnetic field (in this sketch, the spin is neglected). The chemical energy varies in a sawtooth-like manner with an increasing magnetic field.

Dissipationless Current Flow

For an infinite two-dimensional electron system with an applied magnetic field in z direction and a homogeneous electric field in y direction, all Landau level states undergo a drift in x direction, so that the local sheet current density j_x in x direction is given by

$$j_x = \nu_L e^2/h E_y \qquad [6]$$

This Hall current is flowing without dissipation since this drift is a property of the eigenfunctions solving the respective Hamiltonian. Actually, for any smooth varying electrical field $E_y(y)$, the eqn [6] remains locally valid. Integrating the sheet current density $j_x(y)$ over a certain width w in y direction, the integral current $I_x = \int_0^w j_x(y)\,dy$ is given by the voltage drop $V_y = \int_0^w E_y(y)\,dy$ over this width,

$$I_x = \nu_L e^2/h V_y \qquad [7]$$

and is therefore independent of the details of the Hall voltage drop along the y direction as long as ν_L is constant on the width w. For integer values of ν_L, that is, whenever the electron system is incompressible, the Hall resistance values (eqn [2]) are obtained. However, this does not allow one to obtain Hall resistance plateaus. Here, theoretical models refer to disorder and inhomogeneities being present in real

two-dimensional electron systems. Especially, the edges of the sample and static potential fluctuations all over the sample lead to variations of the local filling factor within the two-dimensional electron system.

Compressible and Incompressible Regions within a Two-Dimensional Electron System

At the edges of the two-dimensional electron system, within the depletion region, the sheet electron density changes from zero to its bulk value over a typical distance of 1 μm. In thermodynamical equilibrium, the electron density profile can be calculated within a Thomas–Fermi approximation by (1) the constraint that the electrochemical potential μ_{elch} that is, the sum of local chemical potential $\mu_{\mathrm{ch}}(n_{\mathrm{s}}(r))$ plus the electrostatic energy $-e\phi(r)$, is constant within the whole two-dimensional electron system, and (2) by the Poisson equation which relates the local electrostatic potential $\phi(r)$ to the local sheet electron density $n_{\mathrm{s}}(r)$, taking into account the adequate boundary condition for $\phi(r)$, which are due to the respective sample. The problem can self-consistently be solved if (3) the chemical potential $\mu_{\mathrm{ch}}(n_{\mathrm{s}},B)$ of the two-dimensional electron system is known as a function of the sheet electron density n_{s} (see **Figure 5**). Without the applied magnetic field, the density of states is constant and, therefore, the chemical potential increases linearly with n_{s}, and the electron concentration profile increases smoothly at the edge toward the bulk. In contrast, with an applied high magnetic field, the density of states is discrete, the chemical potential increases in a step-like manner with increasing n_{s}, and therefore it is energetically favorable that the electron density profile shows regions of varying and of constant electron density. A strip-like structure is found along the edges, where the strips of constant electron density behave as an incompressible electron system whereas the regions of varying electron density behave compressible. In compressible regions, occupied and unoccupied electronic states exist at the Fermi level, whereas in incompressible regions, the occupied states lie below and the unoccupied states above the Fermi level.

The widths and positions of these strips depend on the imprinted potential profile at the edges and on the quantization energies (cyclotron energy and Zeeman energy) which increase with magnetic field. Therefore, the strips shift with increasing magnetic field away from the edges into the bulk, and the respective innermost incompressible strip covers the whole bulk at integer values of $n_{\mathrm{s}}/n_{\mathrm{L}}$.

Due to potential fluctuations within the bulk region of the real two-dimensional electron system, the local electron density $n_{\mathrm{s}}(r)$ varies – a landscape of compressible and incompressible regions is also formed there which reshapes with varying the mean sheet electron density n_{s} or the magnetic field. Close to integer values of $n_{\mathrm{s}}/n_{\mathrm{L}}$, the bulk is mainly incompressible with compressible droplets embedded.

Current Distribution in Real Samples

With these ingredients (see **Figure 6**), the IQHE is roughly described by the following qualitative picture. With the application of an electrochemical potential difference between the source and drain

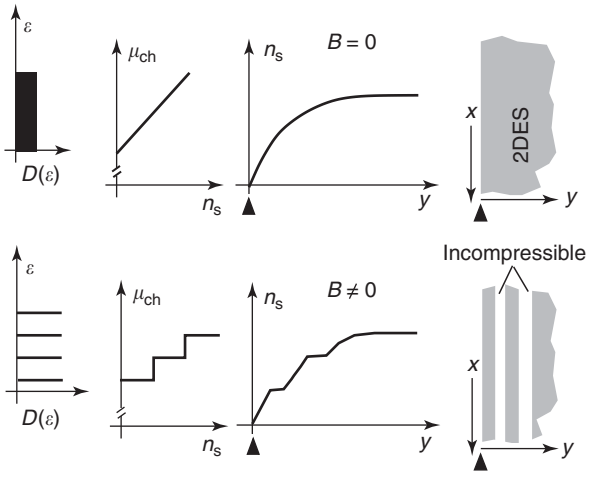


Figure 5 Formation of compressible and incompressible strips within the depletion region along the edges of the 2DES for a strong applied magnetic field.

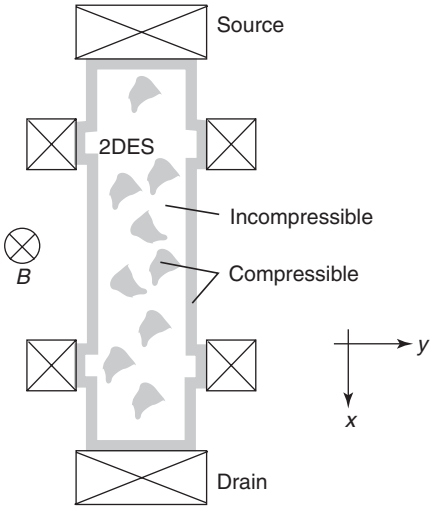


Figure 6 In the (integer) quantum Hall regime, the bulk of the 2DES is incompressible with compressible droplet regions embedded. Along the edges, the electron system has one or more compressible strips.

contact, the electrochemical potential of the source is carried by the compressible strips along one edge, the potential of the drain by the compressible strips along the opposite edge, into the sample. The difference is then present as a Hall voltage driving the Hall current without dissipation along the incompressible regions through the bulk of the sample. Since the Hall voltage drops only over incompressible regions with the same local filling factor, eqn [7] is valid. The inhomogeneities at the edges and within the bulk assure that incompressible regions extend along the whole sample for moderate variations of the magnetic field or for changes of the mean electron density. The Hall plateau values are measured as long as electron scattering from one edge to the opposite edge is suppressed by wide incompressible regions in between.

Localization

Since a certain amount of disorder is required to explain the occurrence of plateaus in the Hall resistance, the QHE has been addressed from the beginning as a localization–delocalization phenomena of electrons in a random disorder potential. The model of an incompressible and compressible landscape within the bulk region, as depicted above, is applicable in the case of a smooth varying disorder potential, and takes the electron–electron interaction self-consistently in an electrostatic manner into account. Other common models investigate the universality of the localization–delocalization phenomenon and the possibility of quantum phase transitions in such a quantum Hall system. The system is then discussed in terms of the density of states in energy where each Landau level is broadened in energy due to the static potential fluctuations. It turns out that states in the tails of the broadened Landau levels correspond to localized states, that is, they are not extended beyond a certain typical length under the respective conditions. Extended states exist only in the energetical center of a broadened Landau level. Localized states cannot contribute to the conductivity, that is, varying the Fermi level in the range of localized states, these states can be filled or emptied without affecting the transport properties – ingredients for the observation of plateaus in the Hall resistance.

Fractional Quantum Hall Effect

With increasing the magnetic field, electrons finally end in the lowest Landau level. Here the electron–electron interaction becomes dominant leading to many-electron correlations, that is, their motions are not independent of each other. For certain fractional filling factors ν, it has been found that the many-electron quantum state behaves incompressible and the respective charge excitations in the electron system are quasiparticles of fractional charge. In a later theoretical description, the electrons and flux quanta present in the system have been combined with new quasiparticles – the so-called composite particles which have either fermionic or bosonic character depending on whether the number of flux quanta attached to an electron is even or odd. This has simplified the picture of the FQHE. Around fractional ν of even denominators, such as $\nu = \{1/2, 3/2, 1/4, 3/4, 5/4, \ldots\}$, composite fermions are formed which do not see any effective magnetic field at the respective filling factor ν. The larger the denominator, the more fragile are these composite fermions. With varying magnetic field, these composite fermions survive and they now feel an effective magnetic field which enforces them to a cyclotron motion. Therefore, within the picture of composite fermions, the series of fractional quantum Hall states which lie symmetrically around $\nu = 1/2$ are interpreted as the IQHE of composite fermions consisting of an electron with two flux quanta attached. (This symmetric structure around $\nu = 1/2$ can be seen in the data of **Figure 3** for FQHE by comparing the low magnetic field region of the IQHE with the regions $\sim$12.6 T, which corresponds to $\nu = 1/2$ in this sample.) At $\nu = 1/2$, the composite fermion does not see any magnetic flux, that is, $(\nu_{\mathrm{CF}})^{-1} = 0$, whereas at $\nu \neq 1/2$, $(\nu_{\mathrm{CF}})^{-1} = |\nu^{-1} - 2|$ flux quanta are present for the composite fermion. An integer filling factor $\nu_{\mathrm{CF}} = \nu/|1 - 2\nu|$ is reached for the fractional filling factors $\nu = \{1/3, 2/5, 3/7, 4/9, 5/11, \ldots\}$ and $\nu = \{1, 2/3, 3/5, 4/7, 5/9, \ldots\}$. The fractional quantum Hall states $\nu = 2/3$ and $\nu = 2/5$ are, therefore, the integer quantum Hall states $i_{\mathrm{CF}} = 2$ of this composite fermion.

The observed exotic fractional quantum Hall state $\nu = 5/2$ is interpreted as a pairing of composite fermions into a novel many-particle ground state. Although the experimental findings support the composite fermion picture, the theoretical foundation for this description is still under debate.

See also: Ballistic Transport; Conductivity, Electrical; Effective Masses; Electron Gas (Theory); Electrons and Holes; Elemental Semiconductors, Electronic States of; Epitaxy; Nanostructures, Electronic Structure of; Quantum Devices of Reduced Dimensionality; Quantum Mechanics: Foundations; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Devices; Semiconductor Nanostructures; Transistors; Transport in Two-Dimensional Semiconductors.

PACS: 73.43. – f

Further Reading

Chakraborty T and Pietiläinen P (1995) *The Quantum Hall Effects*, 2nd edn. Berlin: Springer.
Das Sarma S and Pinczuk A (1997) *Perspectives in Quantum Hall Effects*. New York: Wiley.
Delahaye F and Jeckelmann B (2003) Revised technical guideline for reliable dc measurements of the quantized Hall resistance. *Metrologia* 40: 217.
Ezawa ZF (2000) *Quantum Hall Effects – Field Theoretical Approach and Related Topics*. Singapore: World Scientific.
Heinonen O (ed.) (1998) *Composite Fermions: A Unified View of the Quantum Hall Regime*. Singapore: World Scientific.
Janßen M, Viehweger O, Fastenrath U, and Hajdu J (1994) *Introduction to the Theory of the Integer Quantum Hall Effect*. Weinheim: VCH.
Jeckelmann B and Jeanneret B (2001) The quantum Hall effect as an electrical resistance standard. *Reports of Progress in Physics* 64: 1603.
Klitzing Kv (1986) The quantized Hall effect. *Review of Modern Physics* 58: 519.
Laughlin RB (1999) Nobel lecture: fractional quantization. *Review of Modern Physics* 71: 863.
Prange RE and Girvin S (eds.) (1987) *The Quantum Hall Effect*. New York: Springer.
Stone M (ed.) (1992) *Quantum Hall Effect*. Singapore: World Scientific.
Stormer HL (1999) Nobel lecture: the fractional quantum Hall effect. *Review of Modern Physics* 71: 875.
Yoshioka D (2002) *Quantum Hall Effect*. Berlin: Springer.

Nomenclature

B	magnetic flux density
c	speed of light in vacuum
e	elementary charge
$-e$	electron charge
E	electrical field
g^*	effective Landé factor
g_s	spin degeneracy
j	current density
$\hbar$	Planck constant
I	current
k	wave vector
m^*	effective mass
m_0	free electron mass
n	charge carrier concentration
n	Landau level index
n_L	Landau level degeneracy
n_s	sheet carrier density
N	electron number
q	charge
r	location vector
R_H	Hall resistance
R_K	von Klitzing constant
R_{K-90}	conventional von Klitzing constant
R_{xx}, R_{xy}	resistance
s_z	spin quantum number
V	voltage
α	Sommerfeld fine-structure constant
$D(\varepsilon)$	energy density of states
ε	single-particle energy
ε_F	Fermi energy
μ	charge carrier mobility
μ_0	permeability of vacuum
μ_B	Bohr magneton
μ_{ch}	chemical potential
μ_{elch}	electrochemical potential
λ	DeBroglie wavelength
λ_F	Fermi wavelength
$\hbar\omega_c$	cyclotron energy
Φ	magnetic flux density
Φ_0	magnetic flux quantum
ν	Landau level filling factor

Quantum Mechanics: Atoms

A Sasso, Università di Napoli "Federico II," Napoli, Italy

Introduction

A quantum mechanical description of the atomic structure began to emerge at the end of the nineteenth century. The period up to mid-1920s produced many fundamental discoveries that laid the foundation of modern physics.

In 1896, the Dutch physicist Zeeman observed that the spectral lines emitted by an excited atomic sample could be separated by a strong magnetic field and he ascribed this phenomenon to the electrons present in atoms. In 1906, J J Thomson proposed an atomic model where the number of electrons Z was equal to the number of positive charges (neutral atoms). According to this model, the electrons were embedded in a positive cloud and their mass was approximately half of the atomic mass.

The scattering experiments of α-rays (particles formed by two protons and two neutrons) through thin layers of matter performed by Geigerr and Marsden represented the arguments used by Rutherford for a new description of atomic structure. Indeed, these experiments suggested that the solid

matter resulted from much more transparent structures than expected from the Thomson model. Since only a small part of the α-particles deviate from their trajectories, the repulsive force should arise from scattering centers occupying a small volume localized within the atom. These scattering centers are called Rutherford atomic nuclei and include almost all the atomic mass. The size is estimated to be 10^{-15} m, hence several orders of magnitude smaller than the atomic size (10^{-10} m). By treating the nucleus as a point of charge Ze and the repulsive interaction of the α-particles (charge $2e$) in terms of the Coulomb force, Rutherford obtained the differential cross section, that is, the probability of finding a particle scattered at an angle θ:

$$\frac{d\sigma}{d\Omega} = \left(\frac{2Ze^2}{4\pi\varepsilon_0}\right)\frac{1}{4\mu^2 v^2 \sin^4 \theta/2} \qquad [1]$$

where v is the relative velocity of the α-particle (mass M_α) and the nucleus (mass M), and m is the reduced mass $\mu = M_\alpha M/(M_\alpha + M)$.

In the Rutherford model, the electrons move in the Coulomb field of the nucleus in orbits similar to the planetary system. Nevertheless, according to the laws of classical physics, an accelerating charged particle, like an electron rotating about its nucleus, radiates electromagnetic waves and hence, loses energy. This would imply that in a time 10^{-10} s, all the energy of the atom is radiated away and the electrons would collapse into the nucleus. This prediction is clearly contrary to the experiment and is one of the most important proofs that the classical laws of motion fail when applied to phenomena on the atomic scale. In addition, the atomic line structure observed in all atomic spectra constitute another important experimental evidence not explainable in terms of the Rutherford model.

A crucial effort to remove these difficulties was done by Bohr in 1913. Bohr was able to put together the concepts of Rutherford's nuclear atoms, Planck's quanta, and Einstein's photons to explain, with high precision, the line positions of the radiation emitted by atomic hydrogen.

Bohr introduced the idea that among all the electrons around the nucleus, only one, the so-called optical electron is responsible for the emission of spectral lines. He still assumed that this electron moves in circular orbits due to the effect of the electrostatic attraction of the nucleus but, instead of the infinite number of orbits allowed by classical mechanics, he postulated that only a certain set of stable orbits, which he called stationary states, are possible and that in these orbits the electron does not radiate electromagnetic energy. Each of these stationary orbits corresponds to definite energies E_a, E_b, E_c, … . The optical electron can emit radiation only when it jumps from an outer orbit toward an inner one. To obtain the frequency of the emitted radiation, Bohr used Einstein's concept of quanta of electromagnetic energy, the photons, which carry an energy $h\nu$ ($h = 6.626 \times 10^{-34}$ J s^{-1} is the Planck constant). The Bohr frequency relation is

$$h\nu = E_{fin} - E_{ini} \qquad [2]$$

where E_{ini} and E_{fin} are the energies corresponding to the initial and final electron energies, respectively. Bohr also postulated the quantization of the angular momentum, L, of the electron moving in its circular orbit, that is,

$$L = m_0 v r = n\frac{h}{2\pi} = n\hbar \qquad [3]$$

where n is a positive integer, $n = 1, 2, 3, \ldots$. Thanks to this idea, Bohr could determine the allowed energies of the electron. If the nuclear mass M is assumed to be much heavier than the electron mass m, then the nucleus can be considered to be at rest and the electrostatic force acting on the electron can be so equated to with the centripetal force:

$$\frac{Ze^2}{4\pi\varepsilon_0 r} = m_0\frac{v^2}{r} \qquad [4]$$

By combining eqns [3] and [4], the possible values of the velocity v and the orbit radius r can be easily found, as well as the total energy of the electron given by the potential energy V and the kinetic energy T:

$$E_n = V + T = -\frac{m_0}{2\hbar^2}\left(\frac{Ze^2}{4\pi\varepsilon_0}\right)^2\frac{1}{n^2}$$
$$n = 1, 2, 3, \ldots \qquad [5]$$

The quantum number n is called the principal quantum number: $n = 1$ corresponds to the ground state, while for $n \to \infty$ the energy is null, $E_\infty = 0$, and this state denotes an electron free from its nucleus (positively ionized atom). Therefore, if one indicates the ionization potential as

$$I_p = \frac{m_0}{2\hbar^2}\left(\frac{Ze^2}{4\pi\varepsilon_0}\right)^2$$

the energy required to bring the electron away from the attraction of its nucleus, then eqn [5] can be written as

$$E_n = I_p\frac{1}{n^2} \qquad [6]$$

For the case of the hydrogen atom ($Z = 1$), the frequency line ν_{ab} corresponding to the transition between two levels labeled with a and b is given by

$$\nu_{ab} = \frac{m_0}{4\pi h^3}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\left(\frac{1}{n_a^2} - \frac{1}{n_b^2}\right)$$
$$= R\left(\frac{1}{n_a^2} - \frac{1}{n_b^2}\right) \quad n_b > n_a \qquad [7]$$

where R is called the Rydberg constant ($R = 109\,677.58(8)\ \text{cm}^{-1}$). **Figure 1** shows the main orbits for hydrogen, showing the formation of the different series: transitions involving the states with $n_a = 1$, $n_a = 2$, $n_a = 3$, and $n_a = 4$ give rise to the Lyman, Balmer, Paschen, and Brackett series, respectively.

At the basis of the Bohr's model, there is the correspondence principle according to which for high quantum numbers, the quantum relations asymptotically tend to the classical ones. For instance, the frequency between the levels n and $n+k$ with $n \gg k$ becomes

$$\nu_{n_a, n_a+k} = k\,\frac{2R}{n^3} \qquad [8]$$

that is, the line frequency emitted by the atom corresponds approximately to integer multiples of a fundamental frequency $2R/n^3$, which represents the frequency of a classical oscillator characterized oscillating at the same frequency $2R/n^3$. In terms of classical physics, the orbiting frequency of an electron in an orbit of radius r:

$$\nu = \frac{1}{2\pi}\left(\frac{Ze^2}{m_0 r_n^3}\right)^{1/2} \qquad [9]$$

By combining eqns [8] and [9], the Rydberg constant for an atom with $Z = 1$ can be written in terms of fundamental constants e, m_0, c, and h, that is, $R = 2\pi^2 e^4 m_0 / ch^3$. Taking into account the finite mass M of the nucleus, the Rydberg constants

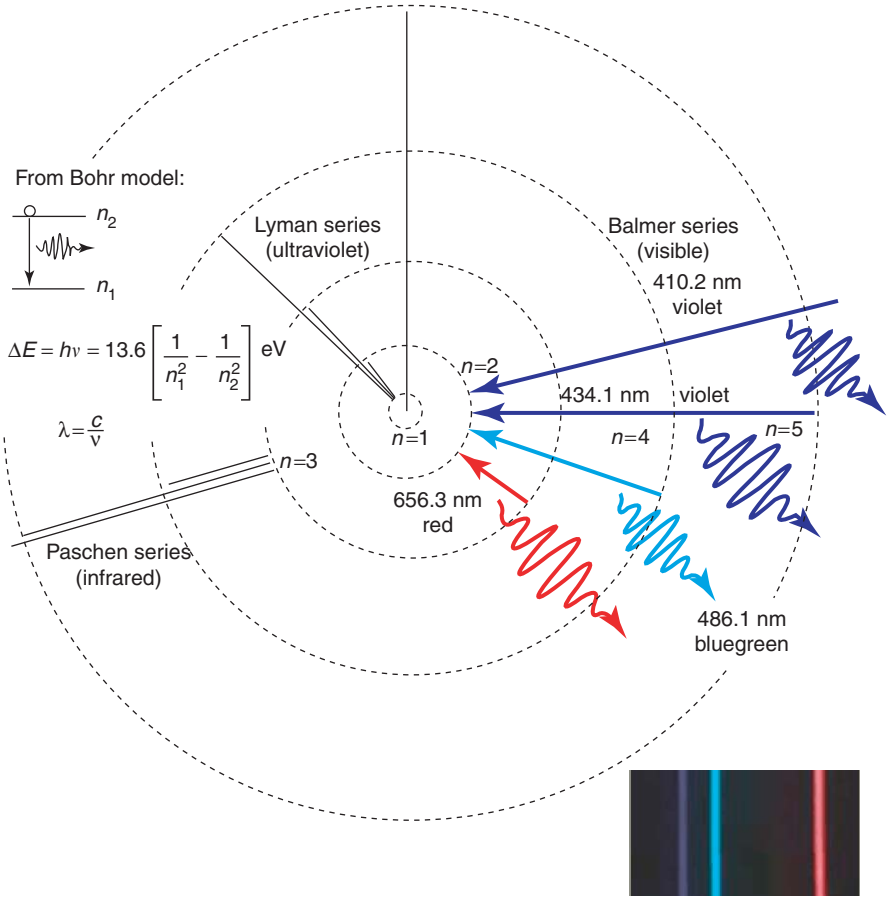


Figure 1 Schematic representation of Bohr model of hydrogen atom. The electron can occupy only quantized orbits corresponding to an energy depending on the principal quantum number n. Emitted radiation at different wavelengths derives from jumps between orbits with different principal quantum numbers n. The first three series (Lyman, Balmer, and Paschen) are shown. In the inset shown in the lower-right corner the first three lines of the Balmer series are also shown (red, visible violet, violet).

become

$$R_M = R_\infty \left(1 - \frac{m}{M}\right) \qquad [10]$$

where R_∞ denotes the Rydberg constant for an infinitely massive atom.

Although the Bohr model is rather successful, it has several unsatisfactory features: it is not able to treat atoms involving two or more electrons and it assumes circular orbits, which is quite arbitrary. Moreover, the Bohr model cannot be used to calculate the intensities of different spectral lines, that is, the transition probability that an electron jumps (absorbed or emitted photon) from a given level to another one. Some of these objections, such as a more general description in terms of elliptical orbits, were removed by the work of Wilson and Sommerfield (old quantum mechanics theory). Nevertheless, the serious question of the transition rate and other still more intriguing questions could be superseded by the quantum mechanics developed by de Broglie, Schrödinger, Heisenberg, Dirac, and others.

Stern and Gerlach Experiment

A significant improvement of the Bohr model was made by an important experiment performed by Stern and Gerlach in 1922 to measure the magnetic dipole moments of atoms. This moment arises from the fact that, according to the Bohr model of a hydrogenic atom, an electron in a circular orbit with an angular momentum $\boldsymbol{L}$ forms a current loop and produces a magnetic dipole $\boldsymbol{M}$,

$$\boldsymbol{M} = -\frac{e}{2m_0}\boldsymbol{L} = -\mu_B \frac{\boldsymbol{L}}{\hbar} \qquad [11]$$

where $\mu_B = e\hbar/2m_0$ is the Bohr magneton ($\boldsymbol{M} = -(e/2m_0)\boldsymbol{L} = -\mu_B = 9.27 \times 10^{-24}\ \mathrm{J\ T^{-1}}$).

In the Stern experiment, a beam of silver atoms was sent through a spatial region where a magnetic field was created (see **Figure 2**).

The force exerted on the dipole is given by $\boldsymbol{F} = -\nabla(-\boldsymbol{M} \times \boldsymbol{B})$, and for the case where only the magnetic field component B_z presents a gradient along the z-axis, the only force on the atoms in the beam is in the z-direction:

$$F_z = M_z \frac{\partial B_z}{\partial z} \qquad [12]$$

In a classical picture, the direction of $\boldsymbol{M}$ in the atomic beam is random and the possible values of M_z are expected to occur in the interval $-M < M_z < M$. As a consequence, atoms impinging on a collecting screen would be spread over a region symmetrically disposed about the point of no deflection. In their experiment, Stern and Gerlach found two distinct lines, symmetrically arranged around the point of no deflection. This result cannot be explained in terms of the Bohr model, although this model postulated that the orbital angular momentum occurs only in integral units of $\hbar L = l\hbar$, and $L_z = m\hbar$, with $-l < m < +l$. Indeed, the results of Stern and Gerlach for silver do not fit with this model, since the multiplicity $(2l+1)$ is 2, giving $l = 1/2$, a noninteger

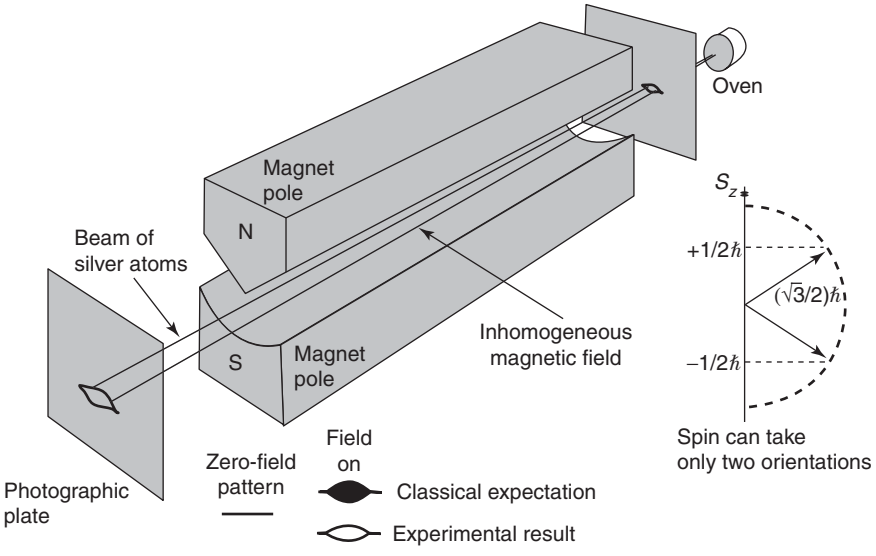


Figure 2 Stern–Gerlach experimental scheme. The silver atomic beam passes through an inhomogeneous magnetic field. One observes the splitting of the beam into two components due to the intrinsic magnetic dipole of the electron (spin).

value. In order to explain the Stern and Gerlach experiment for silver, Gouldsmit and Uhlenbeck suggested, in 1925, that the electron possesses an intrinsic magnetic moment M_s due to an intrinsic angular momentum (spin) in addition to the orbital magnetic dipole moment. Similarly to the orbital magnetic dipole moment, $\boldsymbol{M}_s$ can be written as

$$\boldsymbol{M}_s = -g_s \mu_B \frac{\boldsymbol{S}}{\hbar} \qquad [13]$$

where g_s is the spin gyromagnetic ratio. If a spin quantum number is introduced, similarly to l, the multiplicity of the spin component in a given direction is $(2s+1)$. Therefore, the Stern and Gerlach experiment can be explained if one assumes that the orbital angular momentum of silver is zero, but that its spin angular momentum is $s = 1/2$, so that the possible values of the component of the spin $\boldsymbol{S}$ in the z-direction are $\pm\hbar$.

de Broglie's Hypothesis and the Formulation of Wave Mechanics

As is well known, according to the classical physics, light, described in terms of Maxwell's electromagnetic equations, gives rise to a wide variety of typical wave phenomena, such as interference and diffraction. On the other hand, particles are well described by Newton's law.

In a quantum mechanical picture, the electromagnetic radiation emitted by atoms is quantized, hν being the energy carried by a photon; the energy of electrons in the orbits around the nucleus are also quantized. However, this sharp distinction between waves and matter appears inadequate if compared to experiments where beams of electrons, when impinging on a small aperture, similarly to photons, give rise to diffraction phenomena.

In 1924, de Broglie suggested that just as radiation has particle-like properties, electrons, or other species of particles, show wave-like behavior. With a photon is associated an energy $E = h\nu$ and momentum $p = h\nu/c$, and de Broglie suggested that the frequency and wavelength associated with a particle of energy E moving at a nonrelativistic speed v is

$$\nu = \frac{E}{h} \qquad [14]$$

$$\lambda = \frac{h}{p} = \frac{h}{m_0 v} \qquad [15]$$

This idea directly provides an explanation of the postulate of eqn [3] used in the Bohr model of the hydrogen atom. Indeed, the wavelength of an electron in a stable stationary orbit of radius r is

$$n\lambda = 2\pi r, \quad n = 1, 2, 3, \ldots \qquad [16]$$

and, for eqn [15],

$$L = rp = r\frac{h}{\lambda} = n\hbar \qquad [17]$$

which is identical to eqn [3].

The wave nature of particles is evidenced when their wavelength is comparable with the dimension of the obstacles or slits used in the apparatus. For instance, the de Broglie wavelengths associated with electrons of energy 1, 10, and 100 eV are 12, 3.9, and 1.2 Å, respectively. Davisson and Germer showed that electrons of such energy could be diffracted by crystal lattices whose spacing of atoms is of the order of few angstroms.

Wave Packet and Schrödinger Equation

A next step toward a full quantum mechanics description of an atom requires the introduction of the concept of wave packets and its physical interpretation.

Consider a wave plane of amplitude A_0, frequency $\nu = \omega/2\pi$, and wave number $k = 2\pi/\lambda$ traveling in the x-direction:

$$A(x,t) = A_0 e^{i(kx-\omega t)} \qquad [18]$$

Applying relations [14] and [15],

$$A(x,t) = A_0 e^{(i/\hbar)(px-Et)} \qquad [19]$$

The wave of eqn [19] is an infinitely long wave train. Since this is in contrast with localized "point-mass" particles, one can consider an overlapping of wave planes in the integral form

$$\psi(x,t) = \int_{k_0-\Delta k}^{k_0+\Delta k} a e^{i(kx-\omega t)} dk \qquad [20]$$

where the amplitude of each wave is taken equal to the constant a. To evaluate the integral [20], one can expand ω around the value k_0 using a Taylor series in $(k–k_0)$, that is,

$$\omega = \omega_0 + \frac{\partial \omega}{\partial k}(k - k_0) + \cdots \qquad [21]$$

When expression [21] is inserted in eqn [20], one obtains

$$\psi(x,t) = a\mathrm{e}^{\mathrm{i}(k_0x-\omega_0t)}\int_{-\Delta k}^{\Delta k} a\mathrm{e}^{\mathrm{i}((\partial\omega/\partial k)t-x)\xi}\mathrm{d}\xi$$

$$= a\mathrm{e}^{\mathrm{i}(k_0x-\omega_0t)}2\frac{\sin[((\partial\omega/\partial k)t-x)\Delta k]}{(\partial\omega/\partial k)t-x} \quad [22]$$

Equation [22] represents a wave packet localized in the region around $x = (\partial\omega/\partial k)t$ moving with a group velocity $v_{\text{group}} = \partial\omega/\partial k = \partial E/\partial p$, which is equal to the particle velocity. **Figure 3** shows the real part of eqn [22] at a given instant t: the width of the packet given by the position of the first zero point of eqn [22] (see **Figure 3**) can be easily calculated and it results in $\Delta x = 2\pi/\Delta k$; therefore, the larger the wave vector region, Δk, the narrower the wave packet extension, Δx.

This result represents the Heisenberg uncertainty principle, the uncertainty in the position is connected with the uncertainty in the momentum. Since the momentum and wave vector are connected by eqn [15], $p = \hbar k$, the Heisenberg principle can also be written as

$$\Delta x\Delta p \geqslant \hbar \quad [23]$$

This relation states that it is impossible to measure the position and the velocity of a particle simultaneously and exactly.

A similar discussion applies to another pair of observables, energy and time. In this case, the uncertainties of these quantities are related by

$$\Delta E\Delta t \geqslant \hbar \quad [24]$$

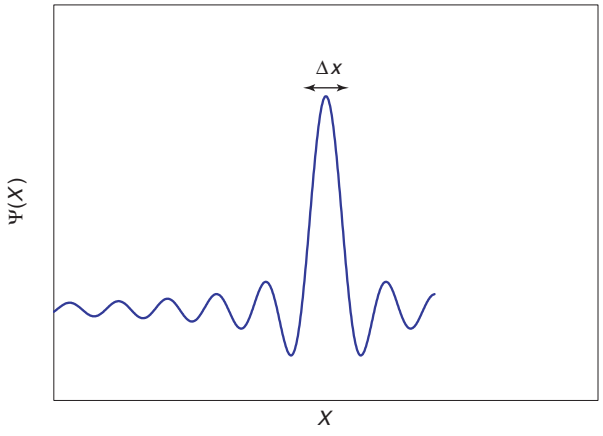


Figure 3 The real part of the wave packet described by eqn [22]. The first node is at $x_0 = \pi/\Delta k$. The rapid oscillations are due to the behavior of the $\sin x/x$ term appearing in eqn [22].

which means that in order to measure the energy with high accuracy, the measurement time has to be large.

The particle picture based on the wave packet is not perfectly equivalent to the wave picture. Indeed, two questions must be solved: the interpretation of $\psi(x,t)$ and the equation to be used for its time evolution. The interpretation given to $\psi(x,t)$ is probabilistic, in the sense that

$$|\psi(x)|^2\mathrm{d}x \quad [25]$$

represents the probability of finding the particle in an interval dx around the position x. In other words, the particle conserves its point-mass nature but it is delocalized within the extension of the wave packet. That implies that $|\psi(x,t)|^2$ is integrated over all space, the particle must be found somewhere. This normalization condition, extended in three dimensions, is written as

$$\int |\psi(x,y,z)|\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z = 1 \quad [26]$$

The equation describing the space–time evolution of the wave function ψ for a free particle is

$$-\frac{\hbar^2}{2m_0}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi \quad [27]$$

This is the basic Schrödinger equation for one-dimensional nonrelativistic force-free particle (particle free from interaction and moving with velocity very small compared to the speed of the light) and its solution automatically fulfils the relation

$$\hbar\omega = \frac{p^2}{2m_0} = \frac{\hbar^2k^2}{2m_0} \quad [28]$$

Equation [27] can be generalized in three dimensions:

$$-\frac{\hbar^2}{2m}\nabla^2\psi = \mathrm{i}\hbar\frac{\partial}{\partial t}\psi \quad [29]$$

and to the case of a particle moving in presence of a potential $V(\boldsymbol{r})$:

$$\left(-\frac{\hbar^2}{2m_0}\nabla^2 + V(\boldsymbol{r})\right)\psi(\boldsymbol{r},t) = \mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) \quad [30]$$

The expression $-(\hbar^2/2m_0)\nabla^2$ is called the kinetic energy operator and

$$H = -\frac{\hbar^2}{2m_0}\nabla^2 + V(\boldsymbol{r}) \quad [31]$$

is another operator that represents the total energy (kinetic energy plus potential energy) of the system and it is called the Hamiltonian (in quantum mechanics all physical quantities which are a result of measurement processes (e.g., the momentum p, the angular momentum l, etc.) are associated with some operators).

If the potential V does not depend on time, the time dependence can be separated in the wave function ψ as

$$\psi(\boldsymbol{r}, t) = \mathrm{e}^{-\mathrm{i}(E/\hbar)t}\Phi(\boldsymbol{r}) \qquad [32]$$

and the time-dependent Schrödinger equation [30] becomes

$$\left(-\frac{\hbar^2}{2m_0}\nabla^2 + V(\boldsymbol{r})\right)\Phi(\boldsymbol{r}) = H\Phi(\boldsymbol{r}) = E\Phi(\boldsymbol{r}) \qquad [33]$$

which is normally called the time-independent Schrödinger equation. The functions Φ and the energies E solution of eqn [33] are called eigenfunctions and eigenvalues, respectively.

Quantum Mechanic Description of the Hydrogen Atom

Consider a single electron of mass m_0 in the field of the nucleus of charge Z. In this case, the potential to be used in the Schrödinger equation is given by the Coulomb potential $V(\boldsymbol{r}) = -Ze^2/4\pi\varepsilon_0\boldsymbol{r}$. In a central potential such as this, as in classical mechanics, the angular momentum is conserved. In quantum mechanics, this is equivalent to stating that the energy and angular momentum can be simultaneously measurable or that the corresponding operators, Hamiltonian and angular momentum, commute. Although the individual components of the angular momentum operator $(\hat{l}_x, \hat{l}_y, \hat{l}_z)$ do not commute with each other, one of these components, for example, $\hat{l}_z$, commutes with the square of the total angular momentum operator; thus, $\hat{l}_z$ and $\hat{l}^2$ are simultaneously measurable.

The Schrödinger equation for an electron in a central potential can be more conveniently solved by choosing spherical polar coordinates. The Cartesian coordinates x, y, z are replaced by r, θ, ϕ, where r is the radius, θ the azimuthal angle between the axis z and the vector $\boldsymbol{r}$, and ϕ the angle which the projection of r in the x–y plane forms with the x-axis. If the spatial part of the wave function ψ is separated as

$$\Phi(r, \theta, \phi) = R(r)Y(\theta, \phi) \qquad [34]$$

then the Schrödinger equation [33] becomes

$$Y(\theta,\phi)\left[-\frac{\hbar^2}{2m_0}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r)\right]R(r) + \frac{R(r)}{2m_0r^2}\hat{l}^2Y(\theta,\phi) = ER(r)Y(\theta,\phi) \qquad [35]$$

For simplicity, the expressions for $\hat{l}_z$ and $\hat{l}^2$ are not explicitly given in spherical coordinates. If the fact that $F(\theta, \phi)$ are also eigenfuntions of the operators $\hat{l}_z$ and $\hat{l}^2$ is used, that is,

$$\hat{l}^2Y(\theta,\phi) = \hbar^2\omega Y(\theta,\phi) \qquad [36]$$

$$\hat{l}_zY(\theta,\phi) = \hbar m Y(\theta,\phi) \qquad [37]$$

eqn [35] reduces to the radial equation

$$\left[-\frac{\hbar^2}{2m_0}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) + \frac{\hbar^2\omega}{2m_0r^2}\right]R(r) = ER(r) \qquad [38]$$

where m appearing in eqn [37] is the magnetic quantum number.

It can be shown that the eigenvalues of $\hat{l}^2$ are $\hbar l(l+1)$ (with $l = 0, 1, 2, \ldots$) instead of the classical expected value of $\hbar l^2$; the eigenvalues of $\hat{l}_z$ are $\hbar m$ (with $-l < m < +l$). The eigenfunctions $Y(\theta, \phi)$ are called the spherical harmonic function and they are labeled by the integer numbers l and m as

$$Y_{l,m}(\theta,\phi) = \mathrm{e}^{\mathrm{i}m\phi}P_l^m(\cos\theta) \qquad [39]$$

$P_l^m(\cos\theta)$ being the Legendre polynomials. The spherical harmonics for $l = 1$ and $l = 2$, both in the spherical and Cartesian coordinates, have the form

$$\begin{aligned}
Y_{1,0} &= \sqrt{\frac{3}{4\pi}}\cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r},\\
Y_{1,\pm1} &= \mp\mathrm{e}^{\pm\mathrm{i}\phi}\sqrt{\frac{3}{8\pi}}\sin\theta = \mp\sqrt{\frac{3}{8\pi}}\frac{x \pm \mathrm{i}y}{r}\\
Y_{2,0} &= \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}\cos\theta - \frac{1}{2}\right) = \frac{1}{2}\sqrt{\frac{5}{4\pi}}\frac{2z^2 - x^2 - y^2}{r^2}\\
Y_{2,\pm1} &= \mp\frac{1}{2}\mathrm{e}^{\pm\mathrm{i}\phi}\sqrt{\frac{15}{2\pi}}\cos\theta\sin\theta = \mp\frac{1}{2}\sqrt{\frac{15}{2\pi}}\frac{(x \pm \mathrm{i}y)z}{r^2}\\
Y_{2,\pm2} &= \frac{1}{4}\mathrm{e}^{\pm2\mathrm{i}\phi}\sqrt{\frac{15}{2\pi}}\sin^2\theta = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\frac{(x \pm \mathrm{i}y)^2}{r^2}
\end{aligned} \qquad [40]$$

The expression for the radial part $R(r)$ of the eigenfunction $\Phi(r, \theta, \phi)$ is

$$R_{n,l}(r) = N_{n,l} e^{-k_n r} r^l L_{n+1}^{2l+1}(2k_n r) \qquad [41]$$

where n is the principal quantum number, $N_{n,l}$ is a normalization factor, and $L_{n+1}^{2l+1}(2k_n r)$ are the Laguerre polynomials. It can be shown that the orbital angular momentum quantum number l cannot be larger than n–1.

Therefore, one can conclude that the wave functions of the hydrogen atom are labeled by three quantum numbers, n, l, and m, and they may be written in the form

$$\psi_{n,l,m}(r, \theta, \phi, t) = e^{i(E_n/\hbar)t} e^{im\phi} P_l^m(\cos\theta) R_{n,l}(r) \qquad [42]$$

In the Bohr model, the electron moves in circular orbits, similarly as a planetary system; now the electron is described by a charge cloud which is mathematically represented by the wave functions [42] whose square modulus $|\psi_{n,l,m}(r, \theta, \phi)|^2$ represents the probability of finding the electron about a given position. Graphical pictures of some wave functions of the hydrogen atom are depicted in **Figure 4**.

For historical reasons, the following convention is normally assumed to indicate the orbital angular momentum l:

l	0	1	2	3	4	5
orbital	s	p	d	f	g	h

For a single-electron atom, the orbital angular momentum $\boldsymbol{l}$ and the spin angular momentum $\boldsymbol{s}$ couple and precess about their mutual resultant $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$.

In terms of energies, different wave functions having the same principal quantum number n and different l, m, and s have the same energy. These states are said to be degenerate and, for a given principal quantum number n, the degree of degeneracy is $2n^2$. The physical reason for such a degeneracy lies in the peculiar $1/r^2$ behavior of the Coulomb electric field.

When the spectral lines of the hydrogen spectrum are examined at very high resolution, they are found to be closely spaced doublets. This splitting is called fine structure and was one of the first experimental evidences for electron spin.

The small splitting of the spectral lines is attributed to an interaction between the electron spin s and the orbital angular momentum l, the so-called spin–orbit interaction.

The familiar red H_α line of hydrogen is a single line according to the Bohr theory. The straightforward application of the Schrödinger equation to the hydrogen atom gives the same result. If the wavelength of this line is calculated using the energy expression from the Bohr theory (eqn [7]), the result is 656.11 nm for hydrogen, treating the nucleus as a fixed center. If the reduced mass is used, the wavelengths become 656.47 nm for hydrogen and 656.29 nm for deuterium (see **Figure 5**). The difference between the hydrogen and deuterium lines is $\sim$0.2 nm and the splitting

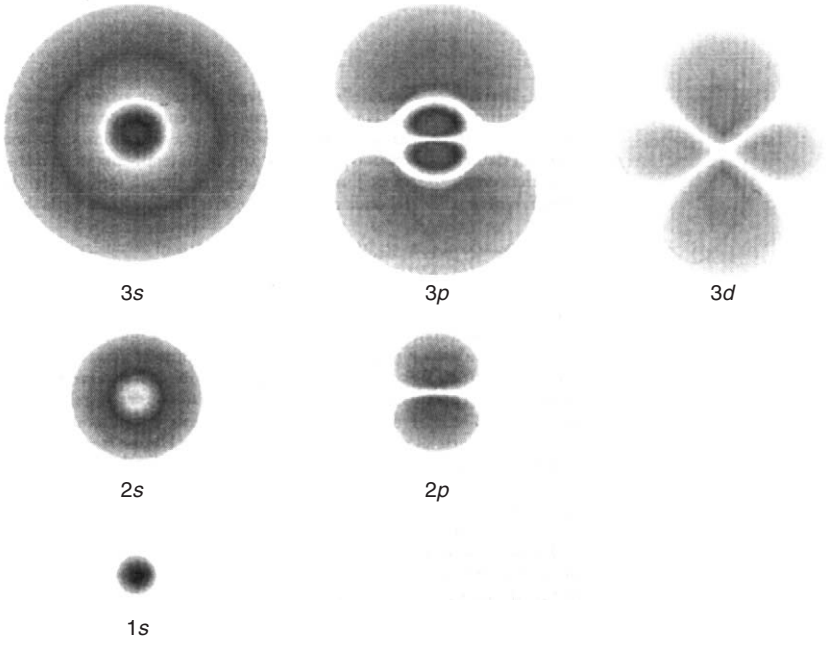


Figure 4 The density electron distribution of different hydrogen states representing the localization probability of $|\psi|^2$.

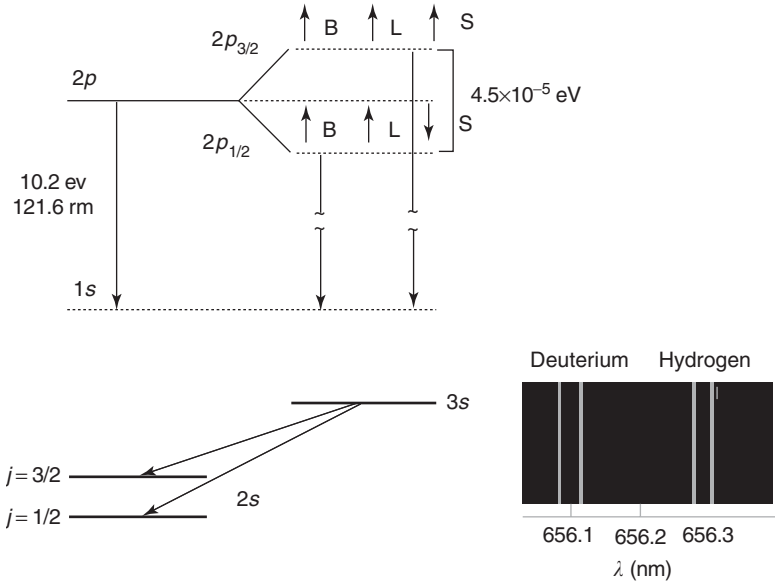


Figure 5 Spin–orbit effect in the 2*p* state of hydrogen atom (upper part). The splitting caused in the 2*p* state is responsible for the doublet observed in the $3s \rightarrow 2p$ transition, in hydrogen or in deuterium (lower part).

of each of them is ~ 0.016 nm, corresponding to an energy difference of $\sim 0.000\,045$ eV. This corresponds to an internal magnetic field on the electron of ~ 0.4 T.

More Complex Atoms: Alkali Atoms and Multi-Electron Atoms

When two or more electrons are bound to a nucleus, their individual electrons can be characterized by the three quantum numbers n, l, and m. This approach is at the basis of the construction of the periodic table of the elements. To do that, one has to take into account the Pauli principle which states that a state labeled by specific values of n, l, and m can be occupied by at most two electrons: one having spin up ($s = +\hbar/2$) and the other one with spin down ($s = -\hbar/2$). As a general rule, the electrons occupy first the levels with lowest energies starting from the ground state which correspond to zero energy. A particular state of occupation of the energy levels is called the electron configuration (or terms) of the atom: for instance for the lowest state of sodium (Na, $Z = 11$), the electronic configuration of the ground state is $1s^2\ 2s^2\ 2p^6\ 3s^1$.

A closed shell occurs when the next electron to be added would occupy the *s* state of the next higher principal quantum number n. Due to symmetry reasons, the angular momentum of closed shells is null and the atoms which reach this condition (noble gases) are particularly stable from a chemical point of view. As a general rule, the electrons in the closed shells are closer to the nucleus than the outer electrons (valence electrons).

Table 1 Ionization energy of alkali atoms

Atomic number Z	*Element*	*Ionization energy* (eV)
1	Hydrogen	13.59
3	Lithium	5.4
11	Sodium	5.14
19	Potassium	4.3
37	Rubidium	4.18
55	Cesium	3.89

Alkali atoms represent atomic systems that are very similar to the hydrogen atom and, for this reason, are also named hydrogenic atoms. Indeed, their electronic configuration corresponds to that of a noble gas plus one electron in the *s* shell. Of course, the similarity with hydrogen atom is only partial, because the inner electrons play some role. For instance, the ionization energies (the energy that has to be furnished to the outer electron to free it from the attractive potential of the nucleus) decrease with the atomic number Z (see **Table 1**).

This behavior can be understood qualitatively by recalling that the electrons in the same subshell have equivalent spatial distributions, so that their screening of one another is rather small.

The valence electron "sees" only partially the nuclear charge $+Ze$ because part of this is screened by the Z–1 inner electrons. This effect can be taken into account by introducing an effective potential $V_{\text{eff}}(r)$ for the valence electron. In this way, the many-body

Table 2 Quantum defects of $\Delta(n,l)$ for the spectra of the sodium (Na) atom

Term	$n=3$	$n=4$	$n=5$	$n=6$
$l=0$	1.373	1.357	1.352	1.349
$l=1$	0.883	0.867	0.862	0.859
$l=2$	0.010	0.011	0.013	0.011
$l=3$		0.000	–0.001	–0.008

problem involving Z interacting electrons in the nuclear potential is strongly simplified and reduced, as for the hydrogen atom, to a two-body problem. The nuclear potential does not vary as $1/r$ and this is the reason why the l degeneracy is removed in alkali atoms. The energy terms $E_{n,l}$ for a given alkali atom X can be written similarly to the hydrogen series but including an effective principal quantum number n_{eff}:

$$E_{n,l} = -R_{\text{X}}hc\frac{1}{n_{\text{eff}}} = -R_{\text{X}}hc\frac{1}{[n-\Delta n(n,l)]^2} \quad [43]$$

where R_{X} is the Rydberg constant for the alkali atom X and $\Delta n(n,l)$ is the so-called quantum defect associated to the quantum numbers n and l. **Table 2** reports the quantum defects (in general, not an integer) of the sodium atom.

Consider now the case when two electrons occupy the outer shell. This is the case of helium atom, or helium-like atoms (Be, Mg, Ca, etc.). Compared with the hydrogenic atoms, the total binding energy contains a term due to the repulsive interaction of the two electrons, that is,

$$V = -\frac{Ze^2}{4\pi\varepsilon_0 r_1} - \frac{Ze^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 r_{12}} \quad [44]$$

The repulsive term depends on the n and l quantum numbers since the electronic spatial distribution depends on these quantum numbers. In contrast with the hydrogen atom, the Schrödinger equation with the potential [44] does not have an exact solution. Approximate solutions can be obtained by using perturbative approaches.

Consider the coupling between the individual orbital and spin angular momenta of the two electrons: (l_1, s_1) and (l_2, s_2). These angular momenta are coupled to each other by means of electromagnetic interaction (e.g., the spin–orbit interaction) between the electrons in the atom. They combine according to specific quantum mechanical rules to produce the total angular momentum, J, of the atom. There are two limiting cases in angular momentum coupling: the *LS* (or Russell–Saunders) coupling, and the *j–j* coupling (or the *jj* coupling).

LS Coupling

If the spin–orbit interaction $\boldsymbol{l}_i \cdot \boldsymbol{s}_i$ is smaller than the mutual electrostatic interaction (third term in eqn [44]), the orbital momenta $\boldsymbol{l}_i$ combine vectorially to form the total orbital momentum $\boldsymbol{L}$ while the spin momenta $\boldsymbol{s}_i$ combine to form the total spin $\boldsymbol{S}$. $\boldsymbol{L}$ and $\boldsymbol{S}$ couple to form the total angular momentum $\boldsymbol{J}$. According to the rules of composition of angular momentum, the possible values of L are

$$L = |l_1 + l_2|,\ |l_1 + l_2| - 1, \ldots, |l_1 - l_2|; \quad \text{multiplicity } 2L+1 \quad [45]$$

and each value determines the term characteristic

$$L = 0,\ 1,\ 2,\ 3, \ldots \quad S,\ P,\ D,\ F, \ldots$$

Similarly, the total spin of the two electrons can be arranged such that $S=0$ (antiparallel spins) with multiplicity $2S+1=1$ (the singlet state) or $S=1$ (parallel spins) with multiplicity $2S+1=3$ (the triplet state). This can also be extended to many-electron light atoms:

$$\boldsymbol{L} = \sum_i \boldsymbol{l}_i, \quad \boldsymbol{S} = \sum_i \boldsymbol{s}_i, \quad \boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S} \quad [46]$$

For helium and the alkaline-earth atoms (Be, Mg, Ca, Sr, and Ba) in their ground configuration, the triplet state is not allowed because of the Pauli principle (two equivalent electrons cannot occupy the same state with the same quantum numbers n, l, m, s).

Depending on which values are possible for the corresponding numbers L and S for a certain configuration, a number of electrostatically split terms are obtained. Such terms are designed as ${}^{2S+1}X_J$, where the $2S+1$ denotes the spin multiplicity, X is the total orbital angular momentum (S, P, D, …), and J is the total angular momentum.

Transitions from one state to another are governed by the following selection rules:

$$\begin{aligned} \Delta l &= \pm 1 \quad \text{for the single electron} \\ \Delta L &= 0,\ \pm 1 \quad \text{for the total system} \\ \Delta S &= 0 \quad \text{for the total spin} \end{aligned} \quad [47]$$

jj Coupling

The second limiting case for angular momentum coupling occurs for heavy atoms, because in this case the spin–orbit coupling $\boldsymbol{l}_i \cdot \boldsymbol{s}_i$ for each individual electron is larger than the interactions $\boldsymbol{l}_i \cdot \boldsymbol{l}_j$ and $\boldsymbol{s}_i \cdot \boldsymbol{s}_j$. In the *jj* coupling, the angular momenta couple

according to the vectorial sum:

$$l_1 + s_1 \rightarrow j_1; \qquad l_2 + s_2 \rightarrow j_2 \qquad [48]$$

The individual angular momenta j_i couple to give the total angular momentum $J = \sum_i j_i$. In this type of coupling, the orbital angular momenta L are not defined and there are, therefore, no terms S, P, D, etc.

In real cases, the *jj* coupling is observed only in very heavy elements. In most cases, intermediate forms of coupling exist. One consequence is, for instance, that the selection rules [47] do not apply for these atoms where the intercombination between terms of different multiplicities is not strictly forbidden.

See also: Electronic Structure (Theory): Atoms; Electronic Structure (Theory): Molecules; Hartree and Hartree–Fock Methods in Electronic Structure; Quantum Mechanics: Foundations; Quantum Mechanics: Methods; Quantum Mechanics: Molecules.

PACS: 31.10; 32.10

Further Reading

Corney A (1977) *Atomic and Laser Spectroscopy*. Oxford: Clarendon.

Haken H and Wolf HC (1996) *The Physics of Atoms and Quanta: introduction to experiments and theory*, 5th edn. Berlin: Springer.

Kuhn HG (1969) *Atomic Spectra*. London: Longmans.

Sobelman II (1992) *Atomic Spectra and Radiative Transitions*, vol. 12, Springer Series Atoms Plasmas. Berlin: Springer.

Nomenclature

c	speed of light ($m\,s^{-1}$)
e	charge of electron (C)
E	total energy (J)
g_s	spin gyromagnetic ratio
h	Planck constant (J s)
I_p	ionization energy (J)
J	total angular momentum ($kg\,m^2\,s^{-1}$)
k	wave vector (cm^{-1})
L	angular momentum ($kg\,m^2\,s^{-1}$)
m_0	electron mass (kg)
p	linear momentum ($kg\,m\,s^{-1}$)
r	radius of electron orbit (m)
R	Rydberg constant (m^{-1})
S	spin vector (J s)
T	kinetic energy (J)
v	velocity ($m\,s^{-1}$)
V	potential energy (J)
Z	atomic number
ε_0	permittivity of vacuum ($F\,m^{-1}$)
λ	wavelength (m)
μ_B	Bohr magneton ($J\,T^{-1}$)
ν	frequency (Hz)
σ	cross section (m^2)

Quantum Mechanics: Critical Phenomena

S Caracciolo, Università degli Studi di Milano, Milan, Italy

Critical Phenomena

Phase transitions are present in a variety of relevant physical systems: from the familiar phenomena of solid–liquid–gas transition to the more intriguing transition to a superconducting state at lower temperature. It is also believed that phase transitions in the early universe from a hotter state with higher degree of symmetry are responsible for the universe as it is known now. The concept of phase transition, and the methods developed in physics to study it, have found fruitful applications very far from the context in which it had been introduced. For example, in computational complexity it is used to characterize the transition for an ensemble of Boolean variables under an increasing number of constraints, which may or may not be satisfied at the same time. It can be used to understand under what conditions a neural network is or is not able to recognize suitable patterns, or to distinguish if the information in a noisy communication channel can or cannot be safely transmitted.

The understanding of phase transitions is one of the most spectacular achievements of statistical mechanics. In the thermodynamical limit, that is, in the presence, formally, of an infinite number of degrees of freedom, the knowledge of the Hamiltonian of the system does not allow one to predict its state: a multiplicity of states are possible! Moreover, the density of a thermodynamic potential, say the free energy, is not necessarily an analytical function of its control parameter, say the temperature, even though it is always a continuous function.

According to the Ehrenfest classification, a phase transition is of order n when, at the transition point,

the thermodynamic potential has a discontinuity at the n-derivative and is continuous in the derivatives of lower order. Nowadays, thanks to the success of ideas related to the application of the renormalization group, it is more customary to make a distinction between critical and noncritical phase transitions, where, in the former case, the correlation length ξ, which is the typical length associated to the fluctuations in the system, diverges at the critical point. Generically, at a critical point, the system, with the disappearance of its typical scale, displays scale invariance. Suppose that the critical point is characterized by a critical value of the temperature T_c, then, the correlation length is expected to diverge with a power law

$$\xi(T) \sim \left| \frac{T - T_c}{T_c} \right|^{-\nu} = |\tau|^{-\nu} \qquad [1]$$

where the power ν is the associated critical exponent and τ is the reduced temperature, a convenient dimensionless parameter. Critical exponents are an example of a universal quantity, that is, something that does not depend on the microscopic details of the system involved, but on more generic features such as dimensionality and the internal symmetries involved. T_c and the size of the critical region, where the universal behavior is observed are, instead, nonuniversal.

Landau introduced the concept of the order parameter. He imagined a local field $\phi(x)$ whose expectation $\langle \phi(x) \rangle = \phi$ (independently from the position because of space homogeneity) vanishes in the disordered phase, but takes a finite value in the ordered phase. The idea behind this concept is that, while the Hamiltonian of the system is invariant under the transformation of a group of symmetry G, the equilibrium states of the system could display a smaller symmetry group H, where H is a subgroup of G. Then, a broken symmetry transformation will send one state into another, and if the order parameter is not left invariant by this broken transformation, it can take a nonvanishing value in the phase with lower symmetry. For example, in a ferromagnet, the Hamiltonian is invariant under the Z_2 inversion symmetry of all the spins, but below the Curie temperature it allows two distinct pure thermodynamical phases with nonvanishing value of the residual magnetization which are sent one into the other by inversion of the spins, as this symmetry is broken.

In classical equilibrium statistical mechanics in the canonical ensemble, relations among various critical exponents can be derived by assuming that, near the critical point, the density of free energy f (more precisely its singular part) expressed as a function of the reduced temperature and the field conjugate to the order parameter h (or to the magnetic field in the case of the ferromagnet) is a homogenous function

$$f(\tau, h) = b^{-d} f(\tau b^{y_\tau}, h b^{y_h}) \qquad [2]$$

where d is the dimensionality of the system, $y_\tau = 1/\nu$ and y_h is another critical exponent. By choosing $b = \xi$, at $h = 0$, because of [1], one gets $f \sim \xi^{-d}$. **Table 1** collects the most common critical exponents. Because of [2], they are not all independent. By applying the definitions, it is easily derived that

$$d\nu = 2 - \alpha, \qquad y_h = \beta\delta/\nu \qquad [3]$$

$$\alpha + 2\beta + \gamma = 2, \qquad 2 - \alpha = \beta(\delta + 1) \qquad [4]$$

called scaling relations. The first one, the Josephson relation, is an example of a hyperscaling law: a scaling relation, which involves directly the dimension d of the system. It has a less secure footing than the other relations. As an example of derivation, consider the specific heat

$$C = \frac{\partial^2}{\partial \tau^2} f(\tau, 0) = b^{-d+2/\nu} f(\tau b^{1/\nu}, 0) \qquad [5]$$

which must behave, for $\tau \to 0$, as $\tau^{-\alpha}$ independently from b, and the first relation follows. Similar

Table 1 Main critical exponents and their mean-field values

	Exponent	*Mean field*	*Definition*	*Conditions*
Specific heat	α	0	$C \propto \lvert\tau\rvert^{-\alpha}$	$\tau \to 0$, $h = 0$
Order parameter	β	1/2	$\phi \propto (-\tau)^{\beta}$	$\tau \to 0^-$, $h = 0$
Susceptibility	γ	1	$\chi \propto \lvert\tau\rvert^{-\gamma}$	$\tau \to 0$, $h = 0$
Eq. state	δ	3	$h \propto \phi \lvert\phi\rvert^{\delta - 1}$	$h \to 0$, $\tau = 0$
Correlation length	ν	1/2	$\xi \propto \lvert\tau\rvert^{-\nu}$	$\tau \to 0$, $h = 0$
Algebraic decay	η	0	$G(r) \propto \lvert r\rvert^{-d+2-\eta}$	$\tau = 0$, $h = 0$

The last two values refer to the properties of the two-point correlation function.

considerations allow one to also derive the relation

$$\gamma = (2 - \eta)\nu \qquad [6]$$

When one also considers dynamics, near the critical point, the typical scale of autocorrelation in time of fluctuations ξ_t also diverges. This phenomenon, known as critical slowing down, is well known to researchers who simulate statistical ensembles by the dynamic Monte Carlo methods, as it makes computer time more and more inefficient in producing statistically independent configurations. The dynamic critical exponent z is defined by

$$\xi_t \sim \xi^z \qquad [7]$$

and, for example, for Brownian diffusion $z = 2$.

Landau Theory

An estimate of the critical exponents can be recovered using the mean-field theory, as shown by Landau. It amounts to imposing analyticity of thermodynamic potentials also at criticality. Consider the thermodynamic potential $g(\phi,\tau)$ at a fixed temperature and an order parameter. In order to fix the ideas, consider a ferromagnet where the order parameter ϕ is the magnetization. Because of analyticity, g can be expanded in ϕ:

$$g(\phi,\tau) \approx g_c(\tau) + r(\tau)\frac{\phi^2}{2} + s(\tau)\frac{\phi^3}{3} + u(\tau)\frac{\phi^4}{4} + \cdots \qquad [8]$$

A linear term in ϕ is not present as, by definition,

$$h = \frac{\partial}{\partial\phi} g(\phi, T) \approx r(\tau)\phi + s(\tau)\phi^2 + u(\tau)\phi^3 + \cdots \qquad [9]$$

that is, one expects a linear response for a small external field. An inversion symmetry would require all odd powers of ϕ in [8] to vanish, so $s(\tau) = 0$. For $u(\tau)$ positive, the series can be truncated to study small effects. For positive $r(\tau)$, there is only a minimum of G, at $\phi = 0$, but for negative $r(\tau)$ there are two of them, among the three values of ϕ corresponding to $h = 0$. Therefore, $r(0) = 0$; thus, at criticality, eqn [9] states that $h \sim \phi^3$, which means that $\delta = 3$. By expanding r in τ, $r(\tau) \sim \tau$. The minima at $h = 0$ for negative τ must have

$$\phi = \pm\sqrt{-\frac{r(\tau)}{u(\tau)}} \sim (-\tau)^{1/2} \qquad [10]$$

thus, $\beta = 2$. For positive τ, $\phi \sim h/r(\tau)$; thus

$$\chi = \frac{\partial\phi}{\partial h} = \frac{1}{r(\tau)} \sim \tau^{-1} \qquad [11]$$

which means $\gamma = 1$. From the scaling relations, $\alpha = 0$ can be obtained. In order to study the correlation functions, fields dependent on position should be introduced to get Gaussian fluctuations which display $\nu = 1/2$ and $\eta = 0$. Josephson hyperscaling relation thus holds for these values only in $d = 4$ dimensions. By taking into account fluctuations of the order parameter by means of a local effective field theory, it can be shown that the Landau theory can hold only in dimension d greater than $d_c = 4$. In lower dimensions, fluctuations are so strong that the effective field theory needs renormalization and the scaling operators acquire anomalous dimensions. The critical behavior is dominated by nontrivial fixed points of the renormalization group flow.

Quantum Critical Phenomena

Relevant fluctuations in finite-temperature transitions are always described according to classical mechanics. The order of magnitude of the classical thermal energy per degree of freedom $k_B T$ can be compared with the zero-point quantum energy per degree of freedom $\hbar\omega_c$, where ω_c is the frequency of the fluctuation. In the critical region, there are modes with enormous autocorrelation time ξ_t, so that the associated frequency is almost vanishing and

$$\hbar\omega_c \ll k_B T_c, \qquad \tau^{z\nu} \ll T_c \qquad [12]$$

which means that at finite T_c transitions, quantum mechanical effects are completely hidden by the thermal fluctuations and the universal properties at criticality are determined by a classical critical point.

However, in some cases, it is possible by changing a parameter at our disposal, say r, to decrease $T_c(r)$ eventually up to zero at some value r_c, that is, $T_c(r_c) = 0$. As an example, at a low enough temperature, $LiHoF_4$ forms an ionic crystal where the only magnetic degrees of freedom are spins of the holmic atoms which choose a preferred axis of orientation in the crystal. In a magnetic field, perpendicular to the preferred axis, this material can be described as a classical Ising model with a ferromagnetic interaction in a transverse field. It has been experimentally found that at a fixed magnetic field, there is a critical phase transition from a paramagnetic phase disordered by thermal fluctuations to a ferromagnetic ordered phase. By increasing the magnetic field, the line of

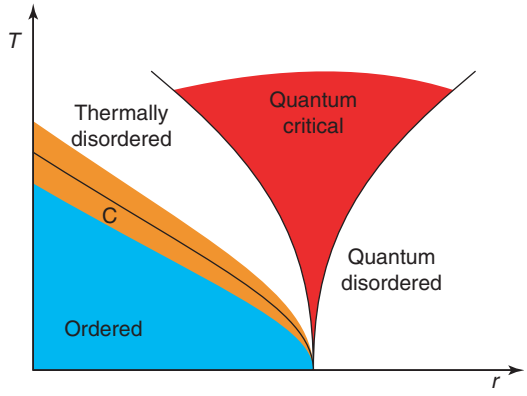


Figure 1 Phase diagram in the plane of the control parameter r and temperature T. Only the curve C is a phase boundary. It ends at the quantum critical point in $(r_c, 0)$. The quantum critical region is bounded by nonuniversal crossover lines.

phase transition reaches $T_c = 0$, at $h_c \sim 50\,\text{kOe}$. At $T = 0$, there are only quantum fluctuations and, therefore, they are only responsible for the (quantum) transition to the disordered phase. For this reason, the endpoint of the line of phase transition at $T = 0$ is called a quantum critical point. In **Figure 1**, the phase boundary curve is denoted by C. At $T = 0$, the line C terminates in the quantum critical point at r_c. In certain cases, such as a one-dimensional magnet or a two-dimensional magnet with a non-Abelian continuous symmetry to be broken, there is no ordered state at nonzero temperature, so that the ordered region below the curve C in **Figure 1** shrinks to the segment to the left of r_c at $T = 0$. Close to the boundary, the region is drawn where classical thermal fluctuations dominate and can be described in terms of a universal classical critical behavior. The disordered region at the right of r_c is a quantum paramagnet where the dominant fluctuations are quantum. The region denoted as quantum critical is dominated by the critical ground state at r_c (where long-range order is the effect of ground-state entanglement) and its thermal fluctuations. Its boundaries occur at

$$T \sim |\tau|^{z\nu} \sim \xi_t^{-1} \qquad [13]$$

(here $\tau = (r - r_c)/r_c$) and do not denote a phase boundary, but only a crossover which discriminates where thermal fluctuations become more important with respect to the quantum ones.

It is important to stress that it is not essential for the experiment that the quantum critical point is not in the physically accessible region, as it is always at $T = 0$. As it always happens around a critical point, fluctuations are well described by an effective field theory.

Phenomenologically, the previous remarks are encoded in a generalization of the scaling ansatz [2], which keeps into account one more relevant parameter

$$f(\tau, T, h) = b^{-(d+z)} f(\tau b^{y_\tau}, T b^z, h b^{y_h}) \qquad [14]$$

In this case, by choosing $b = \xi$, at $h = 0$, one gets the free energy

$$f(\tau, T, 0) = \xi^{-d} \xi_t^{-1} F(T \xi_t) \qquad [15]$$

expressed through a crossover function. It is interesting that, at variance with the classical case, in the effective volume factor, together with ξ^d, there is an explicit contribution from the time direction ξ_t. Indeed, in contrast with what occurs in classical statistical mechanics, where, generally, the contribution from the dynamics (i.e., the kinetic part of the Hamiltonian) is trivially simplified out, in quantum mechanics, because of the Heisemberg uncertainty relations (i.e., noncommutativity of coordinates and momenta), static and dynamic fluctuations cannot be disentangled. This difficulty demands more powerful tools for including quantum effects in statistical descriptions.

Path Integral

There is a formal analogy between the Boltzmann operator in quantum statistical mechanics at temperature $T = 1/(k_B \beta)$ (i.e., $\exp(-\beta \mathscr{H})$) and the time evolution operator in quantum mechanics (i.e., $\exp(-it\mathscr{H}/\hbar)$), after the identification $t = i\beta\hbar$. An analytic continuation to imaginary time, called the Wick rotation, is the starting point for the construction of the Schwinger functions, that is, correlation functions in the Euclidean field theory. A path integral representation is obtained by using the Trotter formula

$$\exp(-\beta \mathscr{H}) = \lim_{N \to \infty} \left[\exp\left(-\frac{\beta}{N} \mathscr{H} \right) \right]^N \qquad [16]$$

that is, one can reduce to consider the time evolution on a small lapse β/N where it is essentially classical, as the quantum effect of noncommutativity of operators is of the order $1/N^2$ and can be neglected in the large-N limit. For example, consider the partition function (which, in the presence of an external source, is the generator of all the correlation functions)

$$Z = \text{Tr} \exp(-\beta \mathscr{H}) \qquad [17]$$

where the trace is over the physical Hilbert space. By introducing a complete basis set for this space at each intermediate time, one has

$$Z = \lim_{N\to\infty} \sum_{\{\phi\}} \prod_{i=0}^{N} \langle \phi_i | \exp(-\Delta t \mathscr{H}) | \phi_{i+1} \rangle$$
$$= \int \mathscr{D}\varphi(x,t) \exp\left(- \int_0^\beta \mathrm{d}t \int \mathrm{d}^d x \mathscr{L}[\varphi(x,t)] \right) \quad [18]$$

where one has to set $\phi_N = \phi_0$ because a trace is being evaluated, and $\Delta t = \beta/N$. When, for example, by means of coherent states, the expectation values are parametrized by auxiliary variables φ, this formula shows that the quantum statistical mechanics in d dimensions can be seen (before the limit $N\to\infty$ is taken) as a classical statistical mechanics of a model in $d+1$ dimensions, which is put in the discrete "time" direction where the index i runs in a box of finite size β. Thus, once more, for finite β, when fluctuations have a correlation length long enough, the effective system dimension is d, but at increasing β there is a crossover to a fully $(d+1)$-dimensional system. Moreover, the path integral representation in [18] shows that, if nontrivial singularities appear in the limit of large N, when time is continuous, it behaves as a spatial dimension, so that the dynamical critical exponent $z=1$. As a consequence, three-dimensional quantum systems, the most interesting quantum systems in solid-state physics, behave at the quantum critical point as four-dimensional classical critical systems. But for $d=4$, the mean-field theory is exact (apart from logarithmic violations) and the critical exponents are those provided in **Table 1**. This conclusion, though often invalid, should also hold when, because of renormalization, $z \geqslant 1$. Indeed, the effective classical description of the soft modes which are massless in the classical theory and are associated with the fluctuation of the order parameter, does not take into account the presence of other modes which can be singular at the quantum critical point because of the vanishing of the temperature. An important example are the particle–hole excitations in an electron fluid that have a mass proportional to the temperature. Indeed, the theory of quantum phase transitions is very rich and shows a complex structure.

Examples

An example of a quantum critical point obtained by changing the magnetic field has already been noted. In other important examples, the control parameter can be the amount of disorder or the strength of the interaction.

Quantum Spin-Glass

$LiHo_xY_{1-x}F_4$ is obtained from $LiHoF_4$ with the replacement of the magnetic ion Ho^{3+} by the nonmagnetic Y^{3+}. By varying the dilution x, and the transverse magnetic field, a divergence in the nonlinear susceptibility, which shows a spin-glass transition, can be moved at zero temperature toward the quantum critical point, even though experimentally the divergence seems to become so weak that it could also be nonexistent, so that the transition could instead be of the first order.

Metal–Insulator Transition

Electrons in a metal form Bloch states which extend throughout the material, and conduction of electricity is allowed from one boundary to the other. But impurities can localize the electron states, that is, wave functions of the electrons have a significant value only over bounded regions, so that electricity can no longer be conducted to the whole sample. The quantum transition from a metal to an insulator, driven by disorder, is called Anderson localization when the interaction among the electrons is neglected. It is known that for $d=2$, in the absence of an external magnetic field, any amount of disorder is enough to produce localization. For $d=3$, when disorder is not too strong, there are both localized and extended states: they are separated in energy by the mobility edge. The difference between the Fermi energy and the mobility edge is the control parameter to the phase transition. By varying the density of electron or the amount of disorder, the transition can be obtained at zero temperature. In the presence of electron interactions, the transition, called the Anderson–Mott transition, is very poorly understood.

Superfluid–Insulator Transition

Disorder can also be responsible for a transition from a superfluid to an insulating state. If there is an attractive interaction between the electrons, superconductivity can arise. In the case of a charged superfluid or superconductor, the transition can be induced by an external magnetic field. Mott transitions can be obtained in Josephson junction arrays where the ratio of the Josephson and Coulomb interaction can be varied.

Quantum Hall Transition

Consider an electron gas confined in an interface between two semiconductors. For $d=2$, free electrons in a magnetic field occupy discrete energy (Landau) levels. By changing the ratio of the density of the

electron to the magnetic field perpendicular to the interface, different phases are obtained. While the resistivity longitudinal to the current flow vanishes, the transverse one, called the Hall resistivity, is quantized. Different states are characterized by the different values of transverse resistivity. Transitions between Hall states are quantum transitions associated with the divergence of the localization length.

Kosterlitz–Thouless Transition

This is a well-known example of a model in classical statistical mechanics which shows a phase transition for $d = 2$, even in the absence of a local order parameter. It is the $O(2)$-vector model, or the XY model, because the spin variables are constrained in a plane. At a high temperature, it is disordered by topological excitations: vortices. However, at temperatures low enough, vortex–antivortex pairs condense and become irrelevant for the Gaussian fixed points. This means that by using mapping between the classical and the quantum models, the quantum $d = 1$ model has a quantum critical point. This transition is relevant for the discussion of the low-dimensional superconductors, and has been used to describe underdoped high-temperature superconducting cuprates.

Deconfinement Transition

Particles or quasiparticles carrying fractionalized quantum numbers interacting through a gauge field, as, for example, quarks in quantum chromodynamics, can undergo a phase transition from a confined phase, where the only excitations at low energy are composite excitations with integer quantum numbers, to a higher-temperature phase, where the effect of the gauge interaction is weak and the fractionalized particles are unbound. The nature of this phase transition is not, in general, fully understood. Different phenomenologies can arise in cases in which the gauge group is discrete, Abelian, or non-Abelian. It could be driven by different topological defects of the gauge fields. In particular, monopoles, instantons, or vortices associated with the center of the gauge group can be effective.

SAT/UNSAT Quantum Transition

There is a quantum version of the classical problems of satisfiability/unsatisfiability in the complexity theory of computation. It refers to the qbits of quantum computation and is governed by a quantum critical point.

See also: Phase Transformations, Mathematical Aspects of; Quantum Computation and Chaos; Statistical Mechanics: Classical; Statistical Mechanics: Quantum.

PACS: 68.35.Rh; 64.60. – i; 05.70.Jk; 73.43.Nq

Further Reading

Cardy J (1966) *Scaling and Renormalization Group in Statistical Physics*. Cambridge: Cambridge University Press.

Domb C and Green M (and afterwards Lebowitz JL) (1971–2003) *Phase Transitions and Critical Phenomena*. London: Academic Press.

Parisi G (1988) *Statistical Field Theory*. Redwood City CA: Addison-Wesley.

Sachdev S (1999) *Quantum Phase Transitions*. Cambridge University Press: Cambridge.

Zinn-Justin J (2002) *Quantum Field Theory and Critical Phenomena*, 4th edn. Oxford: Oxford University Press.

Quantum Mechanics: Elementary Particles

G Marchesini, Università di Milano–Bicocca, Milan, Italy

Historical Introduction

By the second decade of the twentieth century quantum mechanics was a "theory": its systematic application to the microscopic world was in general successful both in describing phenomena and in solving fundamental problems.

At that time, the search for a theory to unify quantum mechanics with relativity became an important issue. It was clear that the phenomena in which the two theories were involved were that of fast (relativistic) and small (quantum) bodies, which is the world of "elementary particles." Among the known elementary particles at that time were the photon γ, electron e, proton p, and neutron n. But soon their number increased. The search for the "unified theory" proceeded in two main directions.

On the one hand, a systematic search for elementary particles, conserved (or almost conserved) quantum numbers, and symmetries was initiated. It required a systematic interplay of theoretical, experimental, and technological developments. Soon it became clear that one could classify the interactions in three groups, with increasing strength: weak, electromagnetic, and strong interactions. The well-known electromagnetic interactions are mediated by

photons. Weak interactions are responsible, for instance, for the neutron decay and the weak radiative decay of nuclei. Strong interactions are responsible for nuclear structures.

On the other hand, the search for dynamical formulations also gained momentum around the same time. The first attempt was a relativistic generalization of the Schrödinger equation. Most important among these was the Dirac equation, which successfully predicted antiparticles, fine-structure splitting of the hydrogen spectrum and the electron magnetic moment. This generalization, however, does not account for particle creation and annihilation processes that take place in elementary particle interactions. Instead, it required the formulation of the quantized (relativistic) field theory (QFT). The meaning of "elementary particles" became clear, together with their connection to quantum fields. The first QFT was quantum electrodynamics (QED), the quantization of Maxwell fields with charged currents of quantized (Dirac) electrons. As it will be clear in the following, the formulation of QFT is very constrained, in that it only requires the quantum numbers and symmetries of the fields (particles).

The two directions are thus intrinsically unified. In the following, after the description of quantum numbers and symmetries, the construction of QFT for elementary particles is illustrated.

Quantum Numbers and Symmetries

The characteristic features of elementary particles are listed in "Review of particle physics." The main structures are summarized here. The commonly used units are $\hbar$, c (light velocity) and MeV for masses (or $\mathrm{GeV} = 10^3\,\mathrm{MeV}$). Electric charges of all observed particles are multiples of the electronic charge. Each particle has an antiparticle with the same mass and spin but opposite charge (and quantum numbers in general). They are classified as leptons and hadrons, which undergo weak, electromagnetic, and strong interactions mediated by "gauge bosons" (gravitational interactions are discussed later).

Gauge Bosons

The gauge bosons that mediate the various interactions are: photon γ, weak bosons (charged $\mathrm{W}^{\pm}$ and neutral Z^0), and gluons g. They are responsible for electromagnetic, weak, and strong interactions, respectively. The photon is massless (with two transverse polarizations) and couples to charged particles. The bosons $\mathrm{W}^{\pm}$, Z^0 are massive (with three polarizations). Due to their large masses (almost hundred proton masses), their discovery (1983) was very difficult. Gluons g carry eight "color charges," are neutral and massless; gluons are not produced but are "confined" within hadronic matter.

Leptons

There are three "families" (or doublets) which undergo electromagnetic and weak interactions (they couple to γ and $\mathrm{W}^{\pm}, \mathrm{Z}^0$). The first family is $(\mathrm{e}, \nu_{\mathrm{e}})$: electron (e) is charged and massive, and neutrino (ν_{e}) is neutral and massless. They both have spin 1/2. While the electron has both polarizations, the neutrino is only left-handed, so that parity (reflection into a mirror) is violated in weak interactions. The other two families, (μ, ν_{μ}) and (τ, ν_{τ}), differ only in the masses of charge particles ($m_{\mu} \sim 200 m_{\mathrm{e}}$ and $m_{\tau} \sim 3600 m_{\mathrm{e}}$). There are recent indications that neutrinos are massive (and then oscillate one into the other with characteristic period).

Hadrons

There are few hundred hadrons divided into two groups, according to whether their spin is integer (mesons, such as π^0, $\pi^{\pm}$) or half-integer (baryons, such as n, p). They undergo all three types of interactions (couple to all gauge bosons).

It was a major discovery that all hadrons are described as "made up" of few quanta, quarks, with spin 1/2 and six quantum numbers called flavors ($f = \mathrm{u}, \mathrm{d}, \mathrm{c}, \mathrm{s}, \mathrm{t}, \mathrm{b}$). Their flavors, charges q_f, and masses m_f are

$$f = \begin{pmatrix} \mathrm{u} & \mathrm{c} & \mathrm{t} \\ \mathrm{d} & \mathrm{s} & \mathrm{b} \end{pmatrix}, \qquad q_f = \begin{pmatrix} +2 \\ -1 \end{pmatrix} \frac{e}{3}$$
$$m_f \simeq \begin{pmatrix} 0.003 & 1.2 & 4.3 \\ 0.006 & 0.1 & 178 \end{pmatrix} \mathrm{GeV} \qquad [1]$$

Mesons are made up of quark–antiquark pairs. For instance, the spinless meson π^+ is made up of quark u and antiquark d, so its charge is

$$q_{\pi^+} = \left(\frac{2}{3} + \frac{1}{3}\right) e = e$$

The negatively charged π^-, antiparticle of π^+, is made up of antiquark u and quark d, so its charge is $-e$.

Baryons are made up of three quarks. The proton, for instance, is made up of three quarks u, u, and d, so its charge is

$$q_p = \left(\frac{2}{3} + \frac{2}{3} - \frac{1}{3}\right) e = e$$

The antiproton is made up of three antiquarks u, u, and d, so its charge is $-e$. Flavor quantum numbers are conserved only in electromagnetic and strong interaction but are weakly violated.

Quarks make up hadrons but are not seen; they are "confined" within the hadrons. This fact should be explained in quantum chromodynamics (QCD), the theory for quark and gluon interaction. To construct QCD, one needs to introduce an additional quantum number for the quark, the "color" with three values. The fact that quarks (and gluons) are confined can be formulated by requiring that physical states (hadrons) are "colorless." In mathematical terms, the three color "charges" correspond to the dimension of the smallest matrices for the eight generators of the SU(3) group. Under SU(3) transformations, quarks (and gluons) transform, whereas hadrons are invariant.

The color quantum number is needed to explain the size of total cross sections and decay rates (e.g., $\pi^0 \to 2\gamma$). Moreover, it is needed for the "anomaly cancellation," a requirement for consistency of QFT for elementary particles. Anomaly cancellation should be mentioned here since it provides a strong constraint on the spectroscopy of all elementary particles. The three quark doublets in [1] parallel the three-lepton families:

$$\left\{\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}\begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}\begin{pmatrix}\mathrm{t}\\ \mathrm{b}\end{pmatrix}\right\} \quad \text{and} \quad \left\{\begin{pmatrix}\mathrm{e}\\ \nu_{\mathrm{e}}\end{pmatrix}\begin{pmatrix}\mu\\ \nu_{\mu}\end{pmatrix}\begin{pmatrix}\tau\\ \nu_{\tau}\end{pmatrix}\right\} \qquad [2]$$

Within such a scheme, the charges of leptons and quarks (including color) sum up to zero, so that fundamental conservation laws of weak currents are satisfied. A missing piece in [2] would destroy such conservation laws in the presence of an anomaly. It is interesting to observe that, because of the t-quark large mass, see [1], the t flavor was discovered much later (1994). However, long before its discovery, the theoretical framework was so well established that the general question was not if but when the t flavor would be found.

Quantum Field Theory: Electrodynamics

The construction of QFT started with electromagnetic interactions, QED, which is presented here in a way suited for generalization to other interactions. Two quantum fields are involved, the electron and the photon. The quantum field $\psi(x)$ of the electron (or any other charged fermion) can be decomposed, in perturbation theory, into fermionic operators that create electrons and positrons with momentum $\boldsymbol{k}$ and helicity $(1/2)\lambda$ (here denoted by $b^\dagger_\lambda(\boldsymbol{k})$ and $c^\dagger_\lambda(\boldsymbol{k})$, respectively):

$$\psi(x) = \int \frac{\mathrm{d}^3k}{(2\pi)^3 2E_k} \times \sum_{\lambda=\pm 1} \{\mathrm{e}^{-\mathrm{i}k\cdot x} u_\lambda(\boldsymbol{k}) b_\lambda(\boldsymbol{k}) + \mathrm{e}^{\mathrm{i}k\cdot x} v_\lambda(\boldsymbol{k}) c^\dagger_\lambda(\boldsymbol{k})\} \qquad [3]$$

with $E_k = \sqrt{m^2 + \boldsymbol{k}^2}$ and $x \cdot k = x_\mu k^\mu = tE_k - \boldsymbol{xk}$. The wave functions u_λ and v_λ are solutions of the Dirac equation with positive and negative frequency and, as $\psi(x)$, have four components (spinors) in correspondence with four degrees of freedom: two helicities for the particle and two for the antiparticle. The (free) Dirac action is

$$S_{\mathrm{Dirac}}[\psi] = \int \mathrm{d}^4x \bar{\psi}(x)[\mathrm{i}\partial\!\!\!/ + m]\psi(x) \qquad \partial\!\!\!/ = \partial_\mu \gamma^\mu, \quad \partial_\mu = \frac{\partial}{\partial x^\mu} \qquad [4]$$

with $\mu = 0, 1, 2, 3$, γ^μ being the Dirac matrices and $\bar{\psi} = \psi^\dagger \gamma_0$.

The (free) Maxwell action is given by the well-known expression

$$S_{\mathrm{Maxwell}}[A] = -\frac{1}{4}\int \mathrm{d}^4x F^{\mu\nu}(x) F_{\mu\nu}(x) \qquad F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \qquad [5]$$

where the (real) vector quantum field has the decomposition

$$A_\mu(x) = \int \frac{\mathrm{d}^3k}{(2\pi)^3 2E_k} \{\mathrm{e}^{-\mathrm{i}kx} a_\mu(\boldsymbol{k}) + \mathrm{e}^{\mathrm{i}kx} a^\dagger_\mu(\boldsymbol{k})\} \qquad [6]$$

with $a^\dagger_\mu(\boldsymbol{k})$ and $a_\mu(\boldsymbol{k})$ the creation and the annihilation operators, respectively. While there are four fields $A_\mu(x)$, the photon has only two polarizations. Related to such a redundancy is the fact that the action [5] is invariant under the "gauge" transformation

$$A_\mu(x) \Rightarrow A_\mu(x) - \frac{1}{e}\,\partial_\mu \Lambda(x) \qquad [7]$$

with $\Lambda(x)$ arbitrary. Therefore, not all of the four fields are independent.

The action of QED, including the photon–electron interaction, can be derived by starting from the free Dirac action plus the gauge invariance. This will be illustrated here since the same procedure will be followed to construct QFT for weak and strong interactions. First, [4] is invariant under the phase

transformation

$$\psi(x) \Rightarrow e^{i\Lambda}\psi(x), \qquad \bar{\psi}(x) \Rightarrow \bar{\psi}(x)e^{-i\Lambda} \qquad [8]$$

where Λ is an arbitrary constant. Then this symmetry is "promoted" by requiring that the action be invariant under [8] with $\Lambda \rightarrow \Lambda(x)$, an arbitrary function of x. The action [4] is not invariant anymore since

$$\bar{\psi}(x)\partial_\mu\psi(x) \Rightarrow \bar{\psi}(x)[\partial_\mu + i\partial_\mu\Lambda(x)]\psi(x) \qquad [9]$$

To enforce the symmetry, one needs to compensate for the $\partial_\mu\Lambda(x)$ term. This is obtained by adding a vector field $A_\mu(x)$ transforming according to [7] and by making the replacement:

$$\begin{gathered}\bar{\psi}(x)\partial_\mu\psi(x) \quad \text{replaced with} \\ \bar{\psi}(x)[\partial_\mu + ieA_\mu(x)]\psi(x)\end{gathered} \qquad [10]$$

in the Dirac action [4]. The result is invariant under the gauge transformation. This shows how powerful requiring gauge symmetry is. From the free Dirac action, one theoretically "discovers" the photon field. Such a strategy will be generalized to "discover" other gauge bosons.

In conclusion, the QED action is

$$\begin{aligned} S_{\text{QED}}[A,\psi] &= S_{\text{Maxwell}}[A] + S_{\text{Dirac}}[\psi] \\ &\quad + ie\int d^4x A_\mu(x) j^\mu(x) \\ j^\mu(x) &= \bar{\psi}(x)\gamma^\mu\psi(x) \end{aligned} \qquad [11]$$

which is a Lorentz invariant, so that unification of relativity and quantum physics is achieved. The interaction does not conserve the particle number. The last term contains three quantum fields (with three creation or annihilation operators), and then a photon could annihilate into an electron–positron pair, an electron can absorb or emit a photon, etc. This feature is unavoidable and is based on the fact that the relativistic action must be constructed only in terms of x-dependent quantum fields, a requirement which is not necessary in nonrelativistic many-body theories. A further important consequence of relativity is that quanta with integer spin must satisfy Bose statistics while quanta with half-integer spin must satisfy Fermi statistics, a consequence of causality of relativistic particle propagators.

The connection between QED action and cross-sections is based on the S-matrix formalism, similar to the nonrelativistic case, apart from nonconservation of particle number. Cross-sections calculation are done (perturbatively) in terms of Feynman rules involving point-like interaction and integration over virtual particle momenta.

The most important surprise of QFT is the presence of infinities arising from integrations over virtual momenta in the ultraviolet (UV) region. The comprehension of this problem (renormalization) has been very laborious from the conceptual and calculational points of view. In QED, all infinities can be completely absorbed into the two parameters (mass and charge) entering the action [11] (and into the normalization of the two fields). Therefore, (m, e) in S_{QED} need to be reinterpreted as UV parameters (m_0, e_0) that absorb all infinities. On the other hand, (m_0, e_0) are determined by two conditions, for instance, giving the physical mass and charge (m, e). In conclusion, all cross sections and distributions involving photons and charged particles can be computed in terms of (m, e), which, however, are not the parameters entering the action but the experimentally measured ones.

The consequences of renormalization is extraordinary from the conceptual point of view: one aims to formulate a fundamental theory, but discovers that QED has to be considered only as an effective theory. Infinities arise from the fact that local fields (defined at point x with infinite resolution) are not adequate for a fundamental description at all distances. At very short distances (much shorter than the present resolution of 10^{-16} cm), nature should be described by different fundamental objects to be discovered (strings?). However, QED remains a predictive theory since the (unknown) fundamental objects can be integrated over the very short distances and the result can be embodied only into quantum fields with UV charge and mass (m_0, e_0).

The features described for infinities in QED are general. A QFT in which all infinities can be absorbed into few UV parameters is said to be renormalizable. It turns out that the criterion is very simple: a QFT is renormalizable if the UV parameters in the action are masses and dimensionless couplings as in QED. Such a criterion allows one to identify the few interaction terms in the action.

In conclusion, quantization and relativity provides remarkably strong constraints: the quantum numbers of fields, their symmetries, and the criterion for renormalizability essentially allow one to uniquely determine the QFT, up to a few UV parameters.

Quantum Chromodynamics: Strong Interactions

The fundamental fields in QCD are the quarks (which make up the hadrons). To construct the action one proceeds as in QED and starts from the free Dirac action. The quarks are described by the fermion field

$\psi_{af}(x)$, given in [3] with flavor $f = 1, \ldots, 6$ and color $a = 1, 2, 3$ quantum numbers. The Dirac action is

$$S_{\text{Dirac}}[\Psi] = \int \mathrm{d}^4x \sum_f \bar{\Psi}_f(x) \cdot [\mathrm{i}\partial\!\!\!/ + m_f]\Psi_f(x)$$
$$\bar{\Psi}_f \cdot \Psi_f = \sum_a \bar{\psi}_{af}\psi_{af} \qquad [12]$$

with $\Psi_f(x)$ the triplet field

$$\Psi_f(x) = \begin{pmatrix} \psi_{1f}(x) \\ \psi_{2f}(x) \\ \psi_{3f}(x) \end{pmatrix}$$
$$\bar{\Psi}_f(x) = (\bar{\psi}_{1f}(x)\bar{\psi}_{2f}(x)\bar{\psi}_{3f}(x)) \qquad [13]$$

To search for the interaction and the gluon, one focuses on the field transformations. One observes that [12] is invariant under the SU(3) transformation (cf. (8))

$$\Psi_f(x) \Rightarrow U\Psi_f(x), \qquad \bar{\Psi}_f(x) \Rightarrow \bar{\Psi}_f(x)U^{-1}$$
$$U = \mathrm{e}^{\mathrm{i}u}, \qquad u = \sum_{\alpha=1}^{8} \lambda^\alpha \theta^\alpha \qquad [14]$$

with λ^α the eight (Hermitean and traceless) 3×3 matrices of SU(3) and θ^α eight real arbitrary parameters.

To find the gluon, the gauge vector boson analogous to the photon field A_μ, one promotes [14] into a local symmetry with $U(x)$ an arbitrary x-dependent SU(3) matrix. In this case, [12] is not invariant since (cf. (9))

$$\bar{\Psi}(x) \cdot \partial_\mu \Psi_f(x)$$
$$\Rightarrow \bar{\Psi}(x) \cdot [\partial_\mu + U^{-1}(x)\partial_\mu U(x)]\Psi_f(x) \qquad [15]$$

To obtain an invariant action, one needs to compensate for the $U^{-1}(x)\partial_\mu U(x)$ terms by introducing eight vector gluon fields $G^\alpha_\mu(x)$ and make in [12] the replacement (cf. (10)):

$$\bar{\Psi}(x) \cdot \partial_\mu \Psi_f(x)$$

replaced with

$$\bar{\Psi}(x) \cdot [\partial_\mu + \mathrm{i}g_{\mathrm{s}} G_\mu(x)]\Psi_f(x)$$
$$G_\mu = \sum_{\alpha=1}^{8} \lambda^\alpha G^\alpha_\mu(x) \qquad [16]$$

with g_{s} the color coupling (analogous to e) and G_μ a 3×3 traceless matrix quantum field which undergoes the gauge transformation (cf. [7])

$$G_\mu(x) \Rightarrow U(x)\left[G_\mu + \frac{\mathrm{i}}{g_{\mathrm{s}}} U^{-1}(x)\partial_\mu U(x)\right] \qquad [17]$$

The replacement [16] in [12] leads to the action

$$S_{\text{QCD}}[\psi, G] = S_{\text{Dirac}}[\psi] + S_{\text{gluons}}[G]$$
$$+ \mathrm{i}g_{\mathrm{s}} \int \mathrm{d}^4x \sum_{\alpha=1}^{8} G^\alpha_\mu(x) j^{\alpha\mu}(x) \qquad [18]$$

with the eight color currents

$$j^{\alpha\mu} = \sum_f \bar{\Psi}_f(x) \cdot \lambda^\alpha \gamma^\mu \Psi_f(x) \qquad [19]$$

The pure gluon term $S_{\text{gluons}}[G]$ is constructed as in QED by requiring invariance under [17]. One finds that it contains gluon self-interaction terms (three- and four-gluon interaction vertices) corresponding to the fact that the gluon field is a matrix in color indices, that is, the gluon has "color charge" and can interact with itself. S_{QCD} contains only the quark masses and the dimensionless gauge coupling g_{s} and is, therefore, renormalizable.

The most challenging question is that while fields are quarks and gluons, only hadrons are seen at large distances. Therefore, the main question is how the QCD action explains that colorless hadrons are made up of color quanta (color confinement). There is no answer yet to this question. Numerical studies of QCD, formulated in terms of Feynman path integrals, seem to reproduce hadron spectra and indicate that color is confined.

Perturbative methods can be used at short distances. Here the effective strength of QCD interaction (the color coupling g_{s}) is small, due to radiative correction (asymptotic freedom). This unexpected phenomenon (in QED, the effective charge increases by going to short distance) is due to gluon self-interaction and radiative corrections, from nonphysical polarization of virtual gluons. Predictions obtained for short-distance distributions are in general well confirmed by experimental data.

Electroweak Theory and Standard Model

One of the most beautiful discoveries in particle physics is the unified theory of electromagnetic and weak interactions, a basic element of the standard model.

Weak phenomena (the prototype is neutron decay) are characterized by very short distances, in contrast to electromagnetic phenomena with an infinite range

of action. This is because the photon is massless (infinite range), while the weak bosons $W^{\pm}$, Z^0 are very massive, so that, at low energy one can assume point-like weak interactions (Fermi theory). It is then surprising that these two completely different interactions could be unified into the electroweak (EW) theory.

To illustrate the EW theory, consider the first lepton family (e, ν_e) and neglect neutrino mass. The important phenomenological observation is that the neutrino is only left-handed and is described by

$$\psi_{\nu L}(x) = P_L \cdot \psi_\nu(x) = \nu_L(x) \qquad [20]$$

with ψ_ν the fermion field given by [3] and $P_L = (1-\gamma_5)/2$ the chiral operator which, for massless fermions, projects on left-handed helicity ($\lambda = -1$). The electron, instead, is present in nature with both helicities. Using the left (P_L) and right (P_R) chiral projectors, one can write the electron quantum field ($P_L + P_R = 1$)

$$\psi_e(x) = P_L \cdot \psi_e(x) + P_R \cdot \psi_e(x) = e_L(x) + e_R(x) \quad [21]$$

(ν_L, e_L, e_R are projected neutrino and electron quantum fields). The (free) Dirac action written in terms of left and right field is

$$S_{\text{Dirac}}[\psi] = i \int d^4x \{\bar{\nu}_L(x) \partial\!\!\!/ \nu_L(x) + \bar{e}_L(x) \partial\!\!\!/ e_L(x) + \bar{e}_R(x) \partial\!\!\!/ e_R(x)\} \qquad [22]$$

The electron mass term $m_e(\bar{e}_R e_L + \bar{e}_L e_R)$, which mixes the left and right electron fields, will be taken into account later.

To construct the theory, one proceeds as in QED and QCD and identifies the symmetry transformations. The Dirac action [22] is invariant under the $SU(2) \times U(1)$ transformation (the transformation of a doublet and a phase rotation)

$$\begin{pmatrix} \nu_L(x) \\ e_L(x) \end{pmatrix} \Rightarrow V \begin{pmatrix} \nu_L(x) \\ e_L(x) \end{pmatrix}$$
$$V = e^{i\nu}, \quad \nu = \sum_{i=1}^{3} \sigma_i \theta_i$$
$$\begin{pmatrix} \nu_L(x) \\ e_L(x) \end{pmatrix} \Rightarrow e^{i\Lambda Y_L} \begin{pmatrix} \nu_L(x) \\ e_L(x) \end{pmatrix} \qquad [23]$$
$$e_R(x) \Rightarrow e^{i\Lambda Y_R} e_R(x)$$

with V an SU(2) matrix, σ_i the Pauli matrices, and (θ_i, Λ) arbitrary real constants. The proportionality constants turn out to be the hypercharges ($Y_L = -1$ and $Y_R = -2$), a generalization of electric charges.

The next step is the search for the gauge vector bosons, $\gamma, W^{\pm}, Z^0$, mediating the interactions. This is done by promoting [23] into local symmetries requiring $\theta_i(x)$ and $\Lambda(x)$ to be four arbitrary x-dependent functions. If one proceeds as in the case of QED and QCD, one obtains four massless gauge vector boson fields.

To set up EW action, one has to take into account that $W^{\pm}, Z^0$ are massive, while γ remains massless, that the electron has a mass and charge, while the neutrino is massless and neutral. The simplest way to obtain all these features is to introduce two complex scalar fields $\phi_1(x)$ and $\phi_2(x)$ and write the (unique) action which is invariant under the $SU(2) \times U(1)$ transformation [23], together with

$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \Rightarrow V \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}$$
$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \Rightarrow e^{i\Lambda(Y_L - Y_R)} \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \qquad [24]$$

Moreover, one has to require that the scalar fields undergo a symmetry breaking and acquire a vacuum expectation value v in one of the four real components, in particular, in physical gauge,

$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \to \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \chi(x) \end{pmatrix} \qquad [25]$$

with $\chi(x)$ a real scalar field. The other three real scalar fields disappear from the EW action and become the longitudinal polarization of the three massive gauge bosons. The particle of $\chi(x)$ field, called Higgs meson, is spinless and neutral, and then couples only to weak bosons. Its mass is not constrained. This particle has not been observed yet and its discovery is very important, since it gives the most clear signal of spontaneous symmetry breaking at the basis of the EW theory.

Spontaneous symmetry breaking in the EW theory is similar (non-Abelian version) to the one used in the standard Bardeen–Cooper–Schrieffer theory of superconductivity, in which one requires that the single complex scalar field (Cooper pair) acquires a ground-state expectation value. This gives a finite range (inverse effective mass) for the electromagnetic field in the superconducting material.

The EW theory for the electron family contains few UV parameters: four dimensionless couplings (two associated with SU(2) and $U(1)$, the coupling of lepton family with the scalar doublet and the scalar doublet self-interaction) and the vacuum expectation value v which has dimension of mass. EW action is then renormalizable.

To complete the EW theory, one has to add the other two lepton families and the quark doublets in (2). For the leptons, one simply replicates the previous construction (with a new parameter for each lepton family, the coupling with the scalar doublet). Taking into account quarks is more subtle, since all fermions in the doublets are massive (quark family mixing). As mentioned before, only by adding together lepton and quarks families, as in [2], and the color charge for quarks, one finds a consistent theory (anomaly cancellation).

Finally, to construct the standard model, the unified theory of all interactions among the elementary particles, one has to merge EW and QCD action. The resulting theory has the symmetry $SU(3) \times SU(2) \times U(1)$, where SU(3) is the color symmetry of QCD. The standard model has been successfully tested, especially by precise measurements at LEP, the electron–positron collider at CERN. At the moment, the fact that neutrino has a mass is the only experimental indication that the theory needs to be modified.

Beyond the Standard Model

Although the standard model is theoretically very constrained and has been successfully tested, there are various reasons to expect that it is incomplete, besides the neutrino mass problem. Schematically, the reasons are: (1) spontaneous symmetry breaking, needed to provide mass to weak bosons and fermions, could be more complex than the one used here. The point will be clarified by the discovery of the Higgs meson; (2) the number of UV parameters in the action is finite but large; (3) the mass scales for various particles is not naturally explained; (4) a simple merging of EW and QCD into the standard model without constraints is not satisfactory, one would expect that they are unified at a deeper level, for instance, by a symmetry unifying SU(3) color with EW symmetry $SU(2) \times U(1)$; (5) to obtain a complete theory, one should take into account gravitational forces as described by general relativity but its quantization leads unavoidably to QFT which is nonrenormalizable (with infinite number of parameters in the action). Finally, as discussed before, the theory is formulated as QFT, which should be considered as an effective theory involving UV parameters to account for infinities. Intensive experimental and theoretical activity is focused on the study of the above questions.

From the experimental point of view, one faces challenging technological problems due to the extreme conditions: one needs to deal with very high energy (e.g., for the Higgs meson search) or with the very small cross sections involved in neutrino physics. In the last few decades, large developments in the astrophysical and cosmological measurements are producing important information for elementary particle physics.

There have been many attempts to search for a theory which solves the above problems. Super-symmetry, in which bosons and fermions are unified, is by now an important element in this search. It is so elegant, and unifies so many aspects and seems to partially explain particle mass scales that one cannot resist but think it represents natural facts. However, up to now, there are no experimental indications of boson–fermion unification.

Infinities in QFT suggest that, at very short distances, fields are not adequate for a physical description. The string theory is a candidate for a fundamental theory. It generalizes QFT: it unifies quantum and relativistic aspects for extended objects (strings) instead of point-like objects as in QFT. The surprise is that the resulting theory is not only highly constrained and mathematically beautiful, but it also requires unification with gravitation, the only force not yet considered at quantum level. Many attempts are ongoing in this direction. The important difficulty, coming as a prize for unification with general relativity, is that the scale at which one starts to resolve strings is $\sim 10^{-33}$ cm, the Planck scale. Essentially no experimental data will be available at such a short distance. Nevertheless, string theories continuously produce theoretical ideas on the construction of the fundamental theory which will supersede QFT. Recent examples are QFT in spaces with extra dimensions (more than three space and one time coordinates) which, in physical dimensions, involve extended objects. They partially solve some of the above problems and have a unifying aspect; however, they can be considered only as first attempts (their actions are not renormalizable).

See also: Group Theory; Nuclear Fission and Fusion; Nuclear Models and Methods; Quantum Mechanics: Methods; Quantum Mechanics: Nuclei; Radioactivity; Relativity; Scattering, Nuclear Resonant; Statistical Mechanics: Classical; Superconductivity: BCS Theory; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice.

PACS: 03.; 03.65. – w; 03.65.Ta; 10

Further Reading

Hagiwara K, *et al.* (2002) Review of particle physics. *Physical Review D* 66: 010001.

Okun LB (1982) *Leptons and Quarks*. Amsterdam, Netherlands: North-Holland.

Pais A (1986) *Inward Bound of Matter and Forces in the Physics World*. New York: Oxford University Press.
Schweber SS (1994) *QED and the Men Who Made It: Dyson, Feynman, Schwinger and Tomonaga*. Princeton University Press.
Weinberg S (1995) *The Quantum Theory of Fields I, II and III*. Cambridge University Press.
Wilson K (1979) Problems in physics with many scales of length. *Scientific American* 241: 158.

Quantum Mechanics: Foundations

A J Leggett, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Introduction

When one thinks about fundamental tests of the principles of quantum mechanics, the area of condensed matter physics is not the first that springs to mind. At least at first sight, such tests would seem to require the use of physical systems which are very well characterized, well isolated from their environments, and (for certain kinds of tests) propagate at or close to the speed of light. None of these conditions is at all well satisfied by a typical condensed matter system. Nevertheless, there are certain kinds of fundamental tests for which such systems are indispensable and others for which they permit substantial extensions of the results obtained on more "microscopic" systems, and this is the subject of this article. The issues of "quantum statistics" discussed are: one aspect of nonlocality in quantum mechanics, and the universal validity of the quantum description.

Quantum Statistics

It is a well-known consequence of the basic principles of the quantum field theory that "elementary" particles can be partitioned into two and only two classes, based on their possessing integral or half-odd-integral spin in units of $\hbar$; these are called bosons and fermions respectively (cf. below for the names). A very fundamental consequence of the indistinguishability in principle of elementary particles, coupled with considerations peculiar to the relativistic quantum field theory, is the spin-statistics theorem: the wave function of any quantum-mechanical system containing more than one particle must be unchanged under the exchange of any two identical bosons, and must change sign under the exchange of any two identical fermions. (A conventional, if somewhat imprecise, phrasing of this principle is that particles of integral (half-odd-integral) spin obey Bose (Fermi) "statistics.") The spin-statistics theorem has been well confirmed at the level of particle physics, for example, in the selection rules on decay processes, and in the case of fermions one of its consequences, the Pauli exclusion principle which forbids more than one fermion from occupying any single-particle state, has well-known effects for the structure of both atoms and nuclei. The spin-statistics theorem applies not only to "elementary" particles but also to complexes built up from the latter, provided that (1) the "internal" structure of the composites is identical, and (2) they remain undissociated in the regime of interest. This follows because if S is the total spin of the complex, the parity of $2S$ is just that of the total number of fermions contained in it, and this is also the parity of the wave function under the simultaneous exchange of all the fermions between two such complexes. Thus, for example, the ^{4}He atom (2 electrons, 2 protons, 2 neutrons, $S = 0$) behaves as a boson, while the ^{3}He atom (2 electrons, 2 protons, 1 neutron, $S = 1/2$) behaves as a fermion. The same conclusion obviously applies to the even and odd isotopic species of any chemical element. Does this difference in "statistics" have any effect in the microscopic limit?

Actually, it is an experimental fact that the properties of most substances do not depend qualitatively on their isotopic composition. The reason is that for the effects of the statistics to be appreciable, two conditions must hold simultaneously. The first is that the behavior predicted by a quantum-mechanical calculation should differ appreciably from that obtained in the classical approximation (if this is not true, then the whole concept of a wave function becomes unnecessary, so its symmetry clearly cannot affect any physical property); this requires that the thermal de Broglie wavelength of the particles in question should be greater than, or at least comparable to, the average interparticle spacing. The second condition is that the particles in question can exchange places fairly readily; if this is not so, it is easy to show that the symmetry of the wave function under interchange can have no substantial consequences. Now for almost all elements and compounds, at temperatures low enough for the first condition to be satisfied the system is already solid, so that the second condition is violated (in a solid the atoms change places, if at all, only very rarely). In fact, by the traditional definition the only

"condensed matter" systems where the two conditions are simultaneously satisfied are the above two stable isotopes of helium, both of which are unique in the remaining liquid under their own vapor pressure down to the lowest temperatures investigated.

The effect of the "statistics" on the behavior of liquid helium is indeed spectacular. The two isotopes, ^{4}He and ^{3}He, are essentially identical in their electronic structure and interatomic interactions, and the difference in nuclear mass does not seem to have any major qualitative effect; indeed, in the solid phase the properties of the two species are very similar (with the exception, of course, of those associated explicitly with the nuclear spin degree of freedom, which is peculiar to ^{3}He). Despite this, the properties in the liquid phase are completely different: in the case of liquid ^{4}He, the liquid enters, below the so-called λ-temperature of $\sim$2.17 K, a "superfluid" phase which displays a variety of exotic properties (persistent currents, nonclassical rotational inertia, film creep, fountain effect, etc.). By contrast, liquid ^{3}He behaves, right down to temperatures below 3 mK, in a way qualitatively similar to that of any other liquid, and in particular, shows none of the above exotic effects. (Below 2.6 mK, liquid ^{3}He does indeed become superfluid, but as shall be seen below, the character of the superfluid phase(s) is/are rather different from that occurring in ^{4}He.)

The origin of this spectacular difference in the behavior of the two isotopes (or at least of much of it) may be understood, at least qualitatively, by asking how they would behave in the complete absence of interactions between the atoms. If N identical noninteracting atoms (of either species) are considered to be moving in a volume V, the relevant single-particle energy eigenstates are plane waves of the form

$$\psi_k(\boldsymbol{r}) = V^{-1/2} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) \qquad [1]$$

If the system is considered to be in thermal equilibrium at temperature T and the standard principles of statistical mechanics are applied subject to the constraints imposed on the many-particle wave function by the spin-statistics theorem, one finds that the distribution $n_k(T)$ of atoms between the single-particle plane-wave states [1] is given by the formula

$$n_k(T) = (\exp((\varepsilon_k - \mu)/k_\mathrm{B}T) \pm 1)^{-1},\ \varepsilon_k \equiv \hbar^2 k^2/2m \qquad [2]$$

where the plus sign applies for fermions and the minus sign for bosons. In these formulas, the quantity $\mu(T)$ is the chemical potential, and its value is fixed, for either sign, by the self-consistency condition

$$\sum_k n_k(T : \mu(T)) = N \qquad [3]$$

The consequences of the sign difference are profound: for fermions, in the limit $T \to 0$, the N lowest-energy single-particle states are each occupied by a single atom, in accordance with the Pauli exclusion principle; one can say that they fill up a "Fermi sea" or "Fermi sphere" in momentum space, with a radius k_F given (for spin 1/2) by the formula $k_\mathrm{F} = (3\pi^2 n)^{1/3}$, where $n \equiv N/V$. At nonzero but low temperatures, particles can be excited to states which lie $\lesssim k_\mathrm{B}T$ in energy above the top of the Fermi sea, leaving "holes" in a similar range in the sea. Although such a "free Fermi gas" model is of course not a realistic picture of real liquid ^{3}He, it turns out that the modifications necessary to take account of the interatomic interactions do not change the situation qualitatively (this is the essential content of Landau's "Fermi-liquid" theory), and that the behavior of the resulting system is overall not very different from that of a classical liquid; in particular, it is predicted to show none of the exotic effects seen in liquid ^{4}He.

In the case of a system of noninteracting bosons, which is described by eqn [2] with the minus sign, the behavior is again not qualitatively very different from that of a normal liquid so long as the chemical potential μ has a finite negative value (which is the case at sufficiently high temperatures). As the temperature falls, however, the value of $\mu(T)$ increases, and for a system moving in free space there exists a nonzero temperature T_c (whose value depends on the density n) below which such a nonzero negative value of μ is incompatible with the self-consistency relation [3]. For $T < T_\mathrm{c}$, the chemical potential tends to zero, and a finite fraction of all the atoms (i.e., a macroscopically large number N_0) occupies the lowest-energy single-particle state, namely that with $\boldsymbol{k} = 0$. This is the phenomenon known as Bose–Einstein condensation (BEC). For a gas of atoms with the mass and density of ^{4}He, the onset temperature T_c of BEC is $\sim$3.3 K, not far from the temperature (2.17 K) of the onset of superfluidity in liquid ^{4}He. It is the almost universal belief that below this latter temperature, BEC indeed occurs in liquid ^{4}He, and that this is the origin of the complex of exotic properties observed in that system.

To illustrate how BEC can give rise to this kind of exotic behavior, the focus is, for definiteness, on the property of "nonclassical rotational inertia." Consider a free gas of N identical atoms (of any species) in a toroidal (annular) container of radius R and thickness $d \ll R$, which rotates at an angular velocity $\omega < \omega_\mathrm{c}/2$, where $\omega_\mathrm{c} \equiv \hbar/mR^2$. What is the angular

momentum of the system in thermal equilibrium? Classical intuition suggests that the system would like to rotate so that its mean velocity exactly matches that of the container, and that its angular momentum will then be simply $NmR^2\omega$. It is easy to show that a system which is well described by classical (Newtonian) mechanics will indeed behave in just this way. For a quantum-mechanical system, it turns out that the effect of the rotation is to change the "effective"energy of a single-particle state with angular momentum l from its original energy $l^2\hbar^2/2mR^2$ to $l^2\hbar^2/2mR^2 - \hbar\omega l$. For a system of any species which is not Bose-condensed, the distribution given by formula [2] is a smooth function of energy, and provided only that the thermal energy k_BT is $\gg\hbar^2/mR^2$ (a condition very well satisfied in any currently realistic experiment), the classical result for the total angular momentum is obtained. However, for a boson system below T_c, a macroscopic fraction of all the atoms must occupy the single lowest-(effective)-energy single-particle state; for $\omega < (1/2)\omega_c$, this is a state with $l = 0$, so a fraction of the atoms stays at rest in the laboratory frame and the observed angular momentum is less than the classical value. Again, this simple model is not a realistic description of liquid ^{4}He (and in fact, cannot account for all the exotic behavior of the latter), but a model which takes adequate account of the interatomic interactions while still assuming the existence of BEC, and has been very successful in explaining all the main phenomena of superfluidity in that system.

Finally, it is generally believed that below 2.6 mK the fermions in liquid ^{3}He form "Cooper pairs" – a sort of giant "diatomic molecules." (A similar phenomenon occurs for the electrons in metals below the transition temperature for superconductivity, but the Cooper pairs in ^{3}He have a more sophisticated structure). Since such pairs of fermions have integral spin, they are in effect bosons and can (and do) undergo Bose condensation. One, therefore, expects the system to show exotic properties such as those displayed by ^{4}He below the λ-temperature, and indeed such are found.

Thus, the two isotopes of liquid helium manifest, in a dramatic way, the consequences of a fundamental principle, the indistinguishability, in principle, of elementary particles.

Aharonov–Bohm Effect

In the classical electromagnetic theory, all electromagnetic effects on charged particles can be obtained from a knowledge of the electric and magnetic fields "seen" by the particle, that is, the fields at the position of the latter; the electromagnetic vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$ is introduced purely as an auxiliary function to simplify some of the calculations. However, in quantum mechanics the situation is different. Consider an experimental arrangement in which a solenoid, regarded here as of infinite length, carries a finite steady current (see **Figure 1**). The magnetic field produced by the current is entirely confined to the interior of the solenoid; outside it is zero. Now consider a beam of electrons which is split so as to pass on either side of the solenoid, and detected on a screen behind it (**Figure 1**). According to a quantum-mechanical calculation, the position of the interference fringes observed on the screen should depend on the current in the solenoid, being periodic in the flux through the latter with period h/e. This is despite the fact that there are no electric or magnetic fields at any point on the trajectories of the electrons! This strange behavior, usually known as the Aharonov–Bohm effect, can be interpreted as showing either that the electrons can "feel" fields even at points they never travel to, or that the vector potential has a "physical reality" not envisaged in the classical electromagnetic theory.

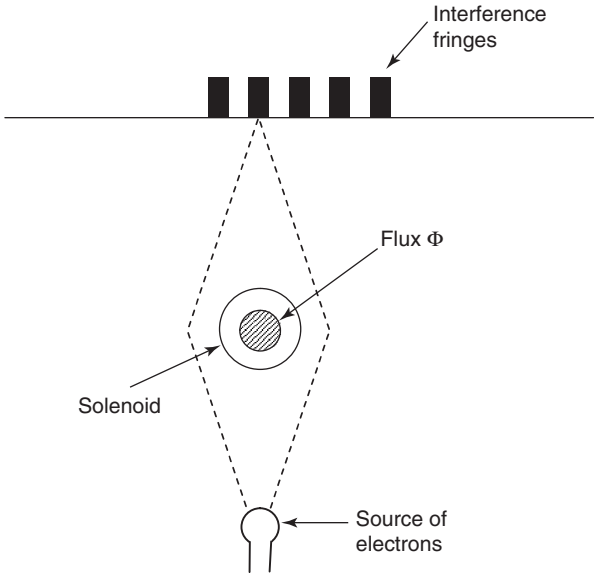


Figure 1 The Aharonov–Bohm effect. The magnetic field of the solenoid is confined to its interior, so there are no fields at any point on the paths taken by the electrons.

While the most loophole-free tests of the Aharonov–Bohm effect have been done with electron beams in free space, the use of condensed-matter systems permits one to display the effects of this "nonlocality" on a macroscopic scale. This is possible in a superconductor, since, to produce the superconducting state, the electrons must form Cooper pairs (as in the superfluid phase of liquid ^{3}He, see above), and these pairs then effectively form bosons

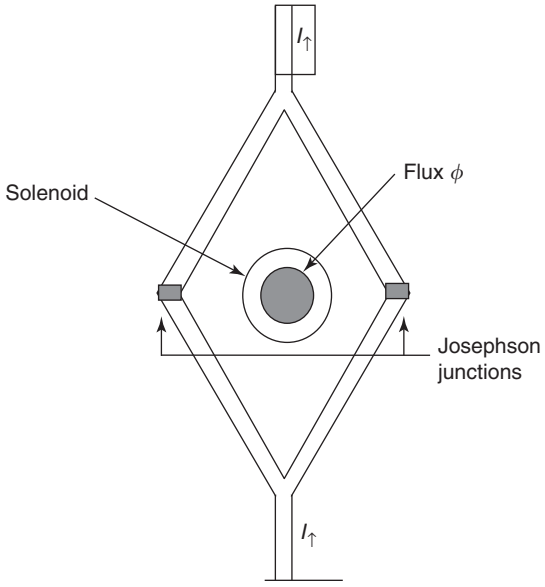


Figure 2 The Aharonov–Bohm effect in a superconducting circuit. The solenoid produces no field anywhere in the metal.

and undergo a sort of BEC (again just as in ^{3}He). Suppose now one forms the completely superconducting device shown in **Figure 2**; in the present context it is enough to know that each junction by itself can carry, without dissipation, a certain maximum ("critical") current I_c, which is assumed for simplicity to be the same for the two junctions. Thus, if one of the arms were blocked the critical current of the device would simply be I_c, while if both are open and there is no current in the solenoid S, the currents in the two arms just add and the critical current of the device is $2I_c$. If however, one imposes a current through the solenoid and thus generates a total flux ϕ through the device, it turns out that the critical current of the latter should be modulated by a factor $|\cos(\pi\phi/\phi_0)|$; thus it is periodic in the (superconducting) "flux quantum" $\phi_0 \equiv h/2e$. (The reason the period is $h/2e$ rather than h/e as in the electron beam case is that the Cooper pairs have charge $2e$, not e). As in the electron-beam case, this is despite the fact that the electrons which carry this current never experience any electric or magnetic fields. This prediction is well verified experimentally, showing that the Aharonov–Bohm effect has not merely atomic-level but also macroscopic manifestations.

Limits on the Validity of Quantum Mechanics

The quantum measurement paradox (or, as it is sometimes known after its most famous instantiation, the "Schrödinger's Cat" paradox) may be stated succinctly as follows: at the microscopic level, if one considers a standard example such as the Young's slits experiment with attenuated beams of single electrons, the fact that the probability amplitude (wave function) which in quantum mechanics describes the ensemble in question, has a finite value at each of the two slits of the intermediate screen cannot, at least without extreme contortions, be interpreted as implying that each individual electron of the ensemble passed through a definite slit (so that the only significance of the amplitude is that its square is a measure of the probability of a given electron passing through the slit in question). The evidence for the above statement is the phenomenon of interference at the final screen. Now, if one believes (as most physicists do) that quantum mechanics is, in principle, the complete truth about the physical universe, it can be applied not just to single electrons and atoms, but to arbitrarily complex objects composed of them, and in this process the formalism itself is in no way modified. This means that it is perfectly possible to devise thought-experiments (such as Schrödinger's original example involving the cat) such that the unambiguous description given by quantum mechanics of the final state of the universe (or more precisely the relevant "ensemble of universes") is a linear superposition of amplitudes corresponding to macroscopically distinct outcomes (e.g., living or dead cat). Thus, unless the interpretation of the quantum formalism is changed significantly between the microscopic and macroscopic levels, one is forced to conclude that it is not true that, for example, each individual cat was either alive or dead before inspection. Yet it is (or at least appears to be!) a fact of our direct experience that inspection reveals each cat to be in one state or the other!

In literature it is often claimed that the apparent paradox can be resolved by appealing to the phenomenon known as "decoherence": when an object is sufficiently macroscopic, and/or interacts with a sufficiently complex environment, then one effect of this interaction is, in effect, to "scramble" the relative phases of the amplitudes representing the two macroscopically distinct states, so that the reduced density matrix of the system becomes equivalent to that of a classical "mixture" of the two states (which *inter alia* gives no possibility of interference), and can thus (it is claimed) be interpreted as simply saying that one or other of the two states is definitely realized for each cat, but one does not know which, and thus have to give a probabilistic description. While the occurrence and ubiquity of the decoherence phenomenon at the macroscopic level is not in serious dispute, the validity of the last, crucial step in the argument (which is equivalent to a major change in the interpretation of the quantum formalism between

the micro- and macro-levels) is controversial in the extreme, so that it is by no means universally agreed in the physics community that the measurement paradox is "solved."

The premise underlying the above discussion is that since the pristine quantum formalism describes, apparently perfectly, the behavior of individual atoms and molecules, it must likewise describe perfectly, at least in principle, the behavior of arbitrarily complex objects composed of them. Given the situation just described concerning the measurement paradox, it would appear to make sense to ask if one is sure that this premise is actually true, or could it perhaps be that by the time one gets from the level of atoms to that of cats, some new physics comes into play which, *inter alia*, has the effect of reducing any superposition of macroscopically distinct states, not to a mixture as the "decoherence" argument would have it, but to a definite one of its branches for any given system of the ensemble. The class of theories which have these properties are termed "macrorealistic"; a number of such theories have appeared in literature, the most fully developed probably being that associated with the names of Ghirardi, Rimini, Weber, and Pearle. The question is whether one could actually devise an experimental test of the class of macrorealistic theories against quantum mechanics.

It is fairly obvious that to the extent that a superposition of macroscopically distinct states is *ipso facto* automatically subject to a degree of decoherence sufficient to reduce the density matrix to that of a classical mixture, it will be impossible to implement such a test, because the experimental predictions of the "mixture" description are, by definition, equivalent to those of a theory in which a definite outcome is realized for each system, that is, a macrorealistic theory. The experimental challenge is thus to find a system, where the superposition is both of "macroscopically distinct" states and at the same time sufficiently resistant to decoherence. Of course, the possibility of doing so rests on the definition of "macroscopically distinct," and this point has had considerable discussion in literature; for present purposes, it is adequate to define two states as having this property relative to one another if there is at least one extensive physical quantity whose expectation value in the two states is different by some large number (say 10^4–10^9, see below) in the "natural" (atomic) units. Below, for convenience, this number is denoted by the symbol Λ.

Suppose, for the sake of argument, that one has found such a system; then one can try to engineer a situation where, if quantum mechanics is indeed the whole story, at some stage the correct description of the relevant ensemble is indeed a linear superposition of two (or more) macroscopically distinct states. At that point one can, in principle, (1) test whether the predictions of quantum mechanics in fact correspond to experiment, and (2) test whether predictions made on the basis of a macrorealistic theory correspond to experiment. While it can be shown that there exist experimental situations in which a positive answer to (1) automatically implies a negative answer to (2) (i.e., confirmation of the quantum-mechanical predictions automatically refutes the whole class of macrorealistic theories), such experiments have not yet been realized, and the experiments to be described have been of type (1) only. In such tests, the interpretation of the raw data is always done in explicitly quantum-mechanical terms, and the question being asked is whether the experimentally realized state is, as predicted by theory, a superposition or rather a mixture of macroscopically distinct states.

Condensed matter experiments, which were consciously designed to test the validity of quantum mechanics at the macroscopic level were started as early as 1980. The first generation of these experiments looked for, and saw, the essentially quantum-mechanical phenomenon of tunneling out of a metastable well (often referred to, in this context, as "macroscopic quantum tunneling" or MQT). A second generation of experiments has looked more directly for the effects of superposition, by testing for NH_3-type oscillations or a related phenomenon. The system of choice is usually a bulk superconducting ring interrupted by a single Josephson junction ("RF SQUID"), or some variant thereof, and the two states in question are states in which the current circulates clockwise or counterclockwise respectively; depending on the details of the geometry, the value of Λ can range from 10^4 to 10^9. If such a system is subjected to an external flux ϕ_{ext}, the energy eigenstates are (in the absence of decoherence) linear superpositions of the above two states; for the special case $\phi_{\text{ext}} = \phi_0/2 = h/4e$, the ground state has even parity and the excited state odd parity, and the level splitting is twice the matrix element t for tunneling between the clockwise- and anticlockwise-circulating states. More generally, as a function of flux the level splitting has the form

$$\Delta E = 2\sqrt{t^2 + \text{const.}\ (\phi - \phi_0/2)^2} \qquad [4]$$

as shown by the solid line in **Figure 3**. By contrast, if the true density matrix of the system is a mixture then the splitting should be linear in the flux, as indicated by the broken line. In the experiments (which used a spectroscopic technique to measure the splitting), it was unambiguously the solid line which was found.

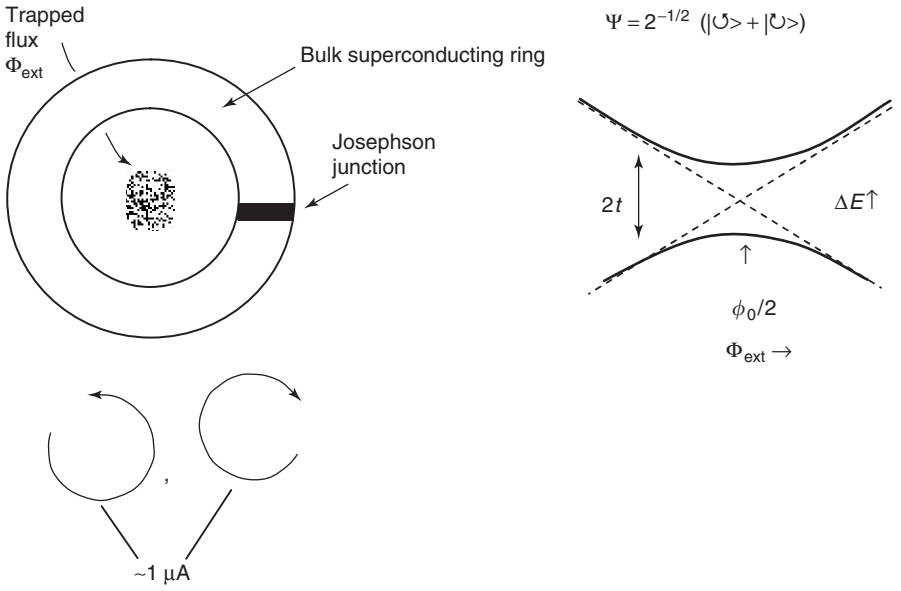


Figure 3 The predicted dependence of the energies of the two lowest states of an RF SQUID ring on the externally applied flux, under the assumption of a coherent superposition (solid lines) or a mixture (broken lines).

In one set of experiments, the value of Λ was $\sim 10^9$, which is probably a record at the time of writing. Note however that these experiments, while verifying that the behavior at this level is consistent with the predictions of quantum mechanics, have not explicitly excluded theories of the macrorealistic class.

In the last two years, experiments on similar systems have detected the NH_3-type oscillations between the clockwise and counter-clockwise-circulating states directly (although the Λ-values in these experiments have been somewhat less spectacular, $\sim 10^5$–10^6). This observation opens the way to eventual tests of type (2), in which (assuming the result is consistent with the predictions of quantum mechanics) one will be able to definitively exclude the viability of any theory of the macrorealistic class at the level of SQUID rings. Whether such a theory may become viable at a level closer to direct human experience is, of course, a matter for experiments, most likely of a different and perhaps currently unforeseeable type, in the (probably distant) future.

See also: Quantum Mechanics: Atoms; Quantum Mechanics: Elementary Particles; Quantum Mechanics: Methods; Quantum Mechanics: Molecules.

PACS: **03.65 Ta; 03.65Vf; 03.65Yz; 05.30Fk; 05.30Jp**

Further Reading

Leggett AJ (1989) In: Davies P (ed.) *The New Physics*, pp. 268–288. Cambridge: Cambridge University Press.

Leggett AJ (2002R) *Journal of Physics: Condensed Matter* 14: 415–451.

Wilks J (1970) *An Introduction to Liquid Helium*. Oxford: Clarendon Press.

Quantum Mechanics: Methods

G M Prosperi, University of Milan, Milan, Italy

Introduction

In principle, any problem in quantum mechanics can be reduced to the solution of the corresponding Schrödinger equation

$$H\psi(t) = i\hbar\frac{\partial\psi(t)}{\partial t} \qquad [1]$$

where obviously H is the Hamiltonian operator and $\psi(t)$ the state vector. As is well known, if H is time independent, this is equivalent to the solution of the

stationary equation

$$Hu_n = W_n u_n \qquad [2]$$

W_n and u_n being the eigenvalues and the eigenstates of H, respectively.

In practice, only in very special cases one can solve eqns [1] or [2] analytically. This essentially happens for a system of free particles, for the single particle or two particles with very few types of potential, for free fields and for certain simplified many-body or field theory models. In the majority of the realistic cases, it is necessary to revert to approximate methods or to make use of numerical simulations.

Leaving apart numerical procedures, methods of very general application are those based on expansions on small parameters, variational methods, semiclassical methods such as the Wentzel–Kramer–Brilouin (WKB) method or those based on the Feynman path integral formalism. Many other methods can be reduced to these.

Various types of small-parameter expansions are used: weak and strong coupling expansions in quantum mechanics and in field theories, $1/c$ expansions in evaluating relativistic corrections to nonrelativistic results, small density, low- and high-temperature expansions in quantum statistical mechanics, and so on.

The many variants of the perturbation theory are typical, in that they are based on the idea of decreasing order of terms in the perturbation expansion. Consider the Hamiltonian of a system split into two parts,

$$H = H_0 + \lambda H_1 \qquad [3]$$

λ being a small parameter, and assume that one is able to solve the eigenvalue equation for H_0,

$$H_0 u_n^{(0)} = W_n^{(0)} u_n^{(0)} \qquad [4]$$

The general idea is to solve eqn [1] or [2] by an expansion in λ. However, in the field theory and in certain versions of the many-body theory, it is often more convenient to work in the Heisenberg picture and use similar techniques for studying the ground-state expectation values of products of appropriate observables.

In bound-state problems, such as those typically considered in atomic and molecular physics (or in nuclear physics or in quark models), the Hamiltonian H_0 already includes a part of the interaction which is more tractable and λH_1 denotes an additional "perturbation." In the scattering theory or in theories involving second-quantization techniques, H_0 often includes only the free part of the Hamiltonian and λH_1, the interaction. In the latter case, λ is usually a coupling constant.

In any case, the perturbation idea is based on the convergence of some expansion (or at least on a good asymptotic approach to a quantity of interest), and sometimes it can completely break down. Then it becomes necessary to turn to other methods.

This article is restricted to the most elementary types of time-dependent and time-independent perturbation theories, to the Green function technique, to variational methods, and to the main ideas on WKB and path integral formalism. Some simple illustrative examples of application are also given.

Time-Independent Perturbation Theory

This is a recursive method for solving the eigenvalue equation [2] for a Hamiltonian of the form [3]. It consists in setting

$$\begin{aligned} u_n &= u_n^{(0)} + \lambda u_n^{(1)} + \lambda^2 u_n^{(2)} + \cdots \\ W_n &= W_n^{(0)} + \lambda W_n^{(1)} + \lambda^2 W_n^{(2)} + \cdots \end{aligned} \qquad [5]$$

in the eigenvalue equation for [3], and in equating equal-power terms in λ on the left- and right-hand sides of the equation. In this way, one can obtain an infinite set of equations; a zero-, a first-, a second-order equation, and so on. These are equivalent in principle to the original equation and should be solved recursively. The zero-order equation coincides with eqn [4], and the successive equations can be written as

$$\begin{aligned} &H_0 u_n^{(1)} + H_1 u_n^{(0)} = W_n^{(0)} u_n^{(1)} + W_n^{(1)} u_n^{(0)} \\ &H_0 u_n^{(2)} + H_1 u_n^{(1)} = W_n^{(0)} u_n^{(2)} + W_n^{(1)} u_n^{(1)} + W_n^{(2)} u_n^{(0)} \\ &\vdots \end{aligned} \qquad [6]$$

where each equation depends on the quantities determined by the solutions of the preceding ones.

The simplest situation occurs when the spectrum of H_0 is not degenerate, that is, when a single eigenfunction $u_n^{(0)}$ (unique up to a factor) corresponds to each unperturbed eigenvalue $W_n^{(0)}$. Then, since the $u_n^{(0)}$ must make an orthonormal basis (up to normalization) set, one can put in the first-order equation

$$u_n^{(1)} = \sum_s a_{ns}^{(1)} u_s^{(0)} \qquad [7]$$

and, after projecting over the vector $u_k^{(0)}$, one obtains

$$\begin{aligned} &a_{nk}^{(1)}((W_k^{(0)} - W_n^{(0)}) + \langle u_k^{(0)} | H_1 | u_n^{(0)} \rangle) \\ &\quad = W_n^{(1)} \delta_{kn} \end{aligned} \qquad [8]$$

where the Dirac notation $\langle \phi|\psi \rangle$ has been used for the scalar product and the relation $\langle u_r^{(0)}|u_s^{(0)} \rangle = \delta_{rs}$.

From [8], one has, for $k = n$,

$$W_n^{(1)} = \langle u_n^{(0)}|H_1|u_n^{(0)} \rangle \qquad [9]$$

that is, the first-order correction to $W_n^{(0)}$ is simply given by the expectation value of H_1 over $u_n^{(0)}$.

For $k \neq n$, one has instead

$$a_{nk}^{(1)} = \frac{\langle u_k^{(0)}|H_1|u_n^{(0)} \rangle}{W_n^{(0)} - W_k^{(0)}} \qquad [10]$$

Equation [8] does not determine $a_{nn}^{(1)}$; however, up to the first order, the normalization condition $\langle u_n|u_n \rangle = 1$ gives $a_{nn}^{(1)} + a_{nn}^{(1)*} = 0$. So, $a_{nn}^{(1)}$ must be purely imaginary and can be taken equal to 0; indeed, to set $a_{nn}^{(1)} = i\eta_n \neq 0$ would be equivalent to multiplying $u_n^{(0)}$ in the preceding equations by the inessential phase factor $1 + i\eta_n \simeq e^{i\eta_n}$.

After introducing into the second-order equation the expression for $W_n^{(1)}$ and $u_n^{(1)}$ obtained above, one can proceed in a similar way for the second-order equation and obtain

$$W_n^{(2)} = \sum_{s \neq n} \frac{|\langle u_n^{(0)}|H_1|u_s^{(0)} \rangle|^2}{W_n^{(0)} - W_s^{(0)}} \qquad [11]$$

and a complicated expression for $a_{nk}^{(2)}$.

In principle, the procedure could be repeated indefinitely until sufficiently small corrections are obtained and, if the resulting expansions [5] converge, eqn [2] would be solved to the required approximation. In fact, eqn [11] itself would require the evaluation of an infinite number of terms, and only in very special cases, the method is of practical application beyond the first order.

If the spectrum of H_0 is degenerate, the unperturbed eigenvectors $\{u_n^{(0)}\}$ are no longer uniquely determined. For each unperturbed eigenvalue $W_n^{(0)}$, one has an eigenspace which may have more than one dimension. Then, a basis $u_{n1}^{(0)}$, $u_{n2}^{(0)}$, ... in each of these subspaces can be chosen and the submatrices $\langle u_{ni}^{(0)}|H_0 + \lambda H_1|u_{ni'}^{(0)} \rangle$ can be considered. The eigenvalues and the eigenstates obtained by diagonalizing this type of matrices must be considered as the new first-order eigenvalues and zero-order eigenstates, and from these one can proceed as above. Actually, the treatment can be simplified by restricting to appropriate invariant subspaces if common constants of motion with respect to H_0 and H_1 exist.

The calculation of the first relativistic correction to the spectrum of the hydrogen atom may be considered as a typical example of application of the above theory. In this case, one can set

$$H_0 = \frac{1}{2m_e} \boldsymbol{p}^2 - \frac{e^2}{4\pi r} \qquad [12]$$

where in the usual Schrödinger representation, $\boldsymbol{p}^2 = -\hbar^2\nabla^2 \equiv -\hbar^2(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$. The Hamiltonian first relativistic correction, $O(1/c^2)$, can be obtained from the Dirac equation and has the form

$$H_1 = -\frac{1}{6m_e^3c^2} \boldsymbol{p}^4 + \frac{e^2}{4\pi m_e^2 c^2 r^3} \boldsymbol{S}\cdot\boldsymbol{L} + \frac{\hbar^2}{8m_e^2c^2} e^2\delta^3(\boldsymbol{x}) \qquad [13]$$

where $\boldsymbol{S}$ and $\boldsymbol{L}$ denote the spin and the orbital momentum of the electron.

The nonrelativistic energy levels and eigenstates of H_0 can be written in the form

$$W_n^{(0)} = -w_0\frac{1}{n^2} \quad \text{and} \qquad u_{nlmm_s} = R_{nl}(r)Y_{lm}(\theta,\phi)v_{m_s} \qquad [14]$$

n, l, m, m_s being the ordinary total, azimuthal, magnetic, and spin quantum numbers respectively, $w_0 = m_e e^4/8\pi^2\hbar^2 = 13.6$ eV the hydrogen ground state energy, $Y_{lm}(\theta,\phi)$ the spherical harmonics, $R_{nl}(r)$ the radial parts of the eigenfunctions, usually expressed in terms of Laguerre functions, v_{m_s} ($m_s = \pm 1/2$) ordinary two-component spinors. The spectrum of H_0 is obviously degenerate in this case (the eigenvalue depending on the quantum number n alone). However, both H_0 and H_1 commute with $\boldsymbol{L}^2$, $\boldsymbol{J}^2$, and J_z ($\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ being the total angular momentum). One can replace then the original u_{nlmm_s}, as given in [14], with the system of the eigenstates $|n; l, j, m_j\rangle$ common to H_0, $\boldsymbol{L}^2$, $\boldsymbol{J}^2$, and J_z, that can be obtained from the preceding ones by linear combination with the appropriate Clebsh–Gordan coefficients. The subspaces spanned by these new unperturbed states for fixed l, j, and m_j are invariant under the action of H_1, and H_0 is no longer degenerate in them. In conclusion, one can use the ordinary formula

$$W_{n,l,j}^{(1)} = \langle n; l; j, m_j|H_1|n; l; j, m_j \rangle \qquad [15]$$

The quantity [15] does not depend on m_j for symmetry reasons and can be calculated with some manipulations. Adding [15] to [14] one obtains, finally, the well-known expression

$$W_{n,j} = -w_0\left[\frac{1}{n^2} + \frac{\alpha^2}{n^3}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right)\right] \qquad [16]$$

which in fact depends only on n and j. In this case, in the end the small parameter is the so-called fine structure constant $\alpha = e^2/4\pi\hbar c = 1/137.036$.

In a similar way, one can evaluate spin–orbit, spin–spin and tensorial splittings, relativistic corrections, two-particle exchange effects, etc., in various bound-state problems in atomic physics, molecular physics, nuclear physics, quark model, etc.

With appropriate modifications, the method can be applied even to continuous spectrum and to scattering problems. However, in the latter connection the following time-dependent theory is more important.

Time-Dependent Perturbation Theory

In this case, it is convenient to work in the so-called Dirac or interaction picture. This consists in giving observables a time dependence of the form

$$O_{\mathrm{I}}(t) = \mathrm{e}^{(\mathrm{i}/\hbar)H_0 t} O \mathrm{e}^{-(\mathrm{i}/\hbar)H_0 t} \qquad [17]$$

and, correspondingly, in redefining the state vector as

$$\psi_{\mathrm{I}}(t) = \mathrm{e}^{(\mathrm{i}/\hbar)H_0 t}\psi(t) \qquad [18]$$

In eqns [17] and [18], O denotes the ordinary time-independent operator in Schrödinger picture and $\psi(t)$ a solution of [1].

As can easily be checked, $\psi_{\mathrm{I}}(t)$ satisfies the equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi_{\mathrm{I}}(t) = \mathrm{i}\lambda H_{1\mathrm{I}}(t) \qquad [19]$$

Introducing the state evolution operator in interaction picture by $\psi_{\mathrm{I}}(t) = U(t, t_0)\psi_{\mathrm{I}}(t_0)$, such an equation can be rewritten in an integral form as

$$\begin{aligned} U(t, t_0) &= 1 - \frac{\mathrm{i}\lambda}{\hbar}\int_{t_0}^{t} \mathrm{d}t' H_{1\mathrm{I}}(t') U(t', t_0) \\ &= 1 + \sum_{n=1}^{\infty}\left(-\frac{\mathrm{i}\lambda}{\hbar}\right)^n \int_{t_0}^{t}\mathrm{d}t_1 \int_{t_0}^{t_1}\mathrm{d}t_2 \cdots \\ &\quad \times \int_{t_0}^{t_{n-1}} \mathrm{d}t_n H_{1\mathrm{I}}(t_1) H_{1\mathrm{I}}(t_2)\cdots H_{1\mathrm{I}}(t_n) \end{aligned} \qquad [20]$$

where the obvious relation $U(t_0, t_0) = 1$ has been used, and the second line is obtained by repeated iteration of the first one.

The quantity

$$P_{r\to s}(t - t_0) = |\langle u_s^{(0)}|U(t, t_0)|u_r^{(0)}\rangle| \qquad [21]$$

gives the transition probability of the system from an initial unperturbed state $u_r^{(0)}$ to the final state $u_s^{(0)}$ during the time $t - t_0$. Approximate expressions for $P_{r\to s}(t - t_0)$ can be obtained by truncating the expansion in [20] to some definite order.

In principle, in [19] and [20] the quantity H_1 can be supposed to be either explicitly time dependent or time independent (i.e., time dependent or time independent in the Schrödinger picture). If H_1 does not depend explicitly on time, one has, up to the first-order,

$$\begin{aligned} &\langle u_s^{(0)}|U(t, t_0)|u_r^{(0)}\rangle \\ &\quad = \delta_{rs} - \frac{\mathrm{i}\lambda}{\hbar}\int_{t_0}^{\mathrm{t}} \mathrm{d}t' \langle u_s^{(0)}|H_{1\mathrm{I}}(t')|u_r^{(0)}\rangle \\ &\quad = \delta_{rs} - \mathrm{i}\frac{2\lambda}{W_s^{(0)} - W_r^{(0)}} \\ &\qquad \times \sin\left[\frac{1}{2\hbar}(W_s^{(0)} - W_r^{(0)})(t - t_0)\right] \\ &\qquad \times \mathrm{e}^{(\mathrm{i}/2\hbar)(W_s^{(0)} - W_r^{(0)})(t - t_0)} \langle u_s^{(0)}|H_1|u_r^{(0)}\rangle \end{aligned} \qquad [22]$$

and, correspondingly, for $s \neq r$

$$\begin{aligned} P_{r\to s}(t - t_0) = &\frac{4\lambda^2}{(W_s^{(0)} - W_r^{(0)})^2} \\ &\times \sin^2\left[\frac{1}{2\hbar}(W_s^{(0)} - W_r^{(0)})(t - t_0)\right] \\ &\times |\langle u_s^{(0)}|H_1|u_r^{(0)}\rangle|^2 \end{aligned} \qquad [23]$$

If the unperturbed energy levels are sufficiently dense to be treated as a continuous distribution, one can use in [23] the asymptotic formula $(\sin^2\omega\tau)/\omega^2 \to \tau\pi\delta(\omega)$ valid for large τ. The replacement is legitimate as soon as $t - t_0$ is large with respect to an appropriate relaxation time δt which will be defined, and so small that $P_{r\to s}(t - t_0) \ll 1$ (so that first-order approximation can be applied). The time δt is given by $\delta t = \hbar/\delta W$, δW specifying the order of magnitude of $W_s^{(0)} - W_r^{(0)}$ for which $\langle u_s^{(0)}|H_1|u_r^{(0)}\rangle$ remains significantly different from 0; it can be interpreted as the duration of the transition. During the considered interval of time (if it exists), $P_{r\to s}$ remains proportional to $t - t_0$ and it is possible to define the transition probability per unit time,

$$\Gamma_{r\to s} = \frac{2\pi}{\hbar}|\langle u_s^{(0)}|H_1|u_r^{(0)}\rangle|^2 \delta(W_s^{(0)} - W_r^{(0)}) \qquad [24]$$

Equation [24] is called Fermi's golden rule and has a large number of applications that range from emission and absorption of radiation by matter, to β decay, weak decay of particles, and elastic and inelastic scattering.

The scattering of a particle by a fixed potential is discussed now. The Hamiltonian may be written as

$$H = \frac{p^2}{2m} + V(r) \qquad [25]$$

Set $H_0 = p^2/2m$, $\lambda H_1 = V(r)$ and denote by $|\boldsymbol{p}\rangle$ the eigenstates of the linear momentum of the particle ($\langle \boldsymbol{p}'|\boldsymbol{p}\rangle = \delta^3(\boldsymbol{p}-\boldsymbol{p}')$, $H_0|\boldsymbol{p}\rangle = W_p|\boldsymbol{p}\rangle$, $W_p = p^2/2m$).

It can be shown that for a short-range potential, the quantity $U(t, t_0)$ has a limit for $t \to +\infty$ and $t_0 \to -\infty$ in the strong sense of the operator theory. Taking into account the second asymptotic relation $(\sin \omega\tau)/\omega \to \pi\delta(\omega)$ (for large τ) and generalizing eqn [22], one can write

$$\langle \boldsymbol{p}'|U(\infty, -\infty)|\boldsymbol{p}\rangle = \delta^3(\boldsymbol{p}'-\boldsymbol{p}) - 2\pi\hbar i\delta(W_{p'} - W_p)T(\boldsymbol{p}', \boldsymbol{p}) \qquad [26]$$

which is the S-matrix for the problem. It is to be noted now that in [23] and [24], the final state $u_s^{(0)}$ can be indifferently supposed to be normalized according to the continuum or the discrete spectrum rules, but the initial state $u_r^{(0)}$ must necessarily be a proper vector normalized to 1. It is then convenient to consider the momentum eigenstates $|\boldsymbol{p}\rangle$ referring to the infinite space (represented by $(1/(2\pi\hbar)^{3/2})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}$ in the usual Schrödinger representation) as the limit case of states $|\boldsymbol{p}\rangle_L$ referring to a finite box of side L (represented by $(1/L^{3/2})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}$, with $\boldsymbol{p}$ assuming only approate discrete values). Then, since $|\boldsymbol{p}\rangle_L = (2\pi\hbar/L)^{3/2}|\boldsymbol{p}\rangle$, one has

$$d\Gamma(\boldsymbol{p}') = d^3\boldsymbol{p}'\frac{(2\pi\hbar)^4}{L^3}\delta(W_{p'} - W_p)|T(\boldsymbol{p}', \boldsymbol{p})|^2 \qquad [27]$$

where $[\delta(W_{p'} - W_p)]^2 = ((t-t_0)/2\pi\hbar)\delta(W_{p'} - W_p)$ has also been used, which follows from a comparison between the two mentioned asymptotic relations for the δ function. Dividing [27] by the flux $\Phi = (1/L^3)(p/m)$, the differential cross section finally obtained is

$$d\sigma(\boldsymbol{p}') = d^3\boldsymbol{p}'(2\pi\hbar)^4\frac{m}{p}\delta(W_{p'} - W_p)|T(\boldsymbol{p}', \boldsymbol{p})|^2 \qquad [28]$$

This is supposed to be an exact equation. However, up to the first order, $T(\boldsymbol{p}', \boldsymbol{p}) = (1/\hbar)\langle \boldsymbol{p}'|V(r)|\boldsymbol{p}\rangle$, and one has the famous Born approximation.

The method can be easily extended to inelastic scattering, such as the collision of an electron with an atom, a proton with a nucleus, and so on.

Green Functions

If $V(r)$ in [25] is time independent, one can write

$$U(t', t) = e^{(i/\hbar)H_0t'}e^{-(i/\hbar)H(t'-t)}e^{-(i/\hbar)H_0t} \qquad [29]$$

From this equation, one immediately has

$$e^{-(i/\hbar)H\tau}U(0, t) = e^{-(i/\hbar)H\tau}e^{(i/\hbar)Ht}e^{-(i/\hbar)H_0t} = U(0, t-\tau)e^{-(i/\hbar)H_0\tau} \qquad [30]$$

and taking the limits $t \to -\infty$ or $t \to +\infty$

$$e^{-(i/\hbar)H\tau}U(0, \mp\infty) = U(0, \mp\infty)e^{-(i/\hbar)H_0\tau} \qquad [31]$$

Then, setting $|\boldsymbol{p}, \pm\rangle = U(0, \mp\infty)|\boldsymbol{p}\rangle$, one finds

$$H|\boldsymbol{p}, \pm\rangle = U(0, \mp\infty)H_0|\boldsymbol{p}\rangle = W_p|\boldsymbol{p}, \pm\rangle \qquad [32]$$

that is, the states $|\boldsymbol{p}, \pm\rangle$ are eigenstates of H corresponding to the same eigenvalue $W_p = p^2/2m$ as $|\boldsymbol{p}\rangle$ of H_0. Furthermore, setting $\tau = t$ in [30], considering again the limits $t \to \mp\infty$ and applying the expression to $|\boldsymbol{p}\rangle$, one obtains

$$\begin{aligned} e^{-(i/\hbar)Ht}|\boldsymbol{p}, +\rangle &\to e^{-(i/\hbar)H_0t}|\boldsymbol{p}\rangle \quad \text{for} \to -\infty \\ e^{-(i/\hbar)Ht}|\boldsymbol{p}, -\rangle &\to e^{-(i/\hbar)H_0t}|\boldsymbol{p}\rangle \quad \text{for} \to +\infty \end{aligned} \qquad [33]$$

which shows that $|\boldsymbol{p}, +\rangle$ and $|\boldsymbol{p}, -\rangle$ correspond to an ingoing and outgoing particle, respectively, both with momentum $\boldsymbol{p}$. Notice that, keeping in mind the relations $U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0)$ and $U^\dagger(t, t_0) = U(t_0, t)$, the S-matrix can be written in terms of such states in the alternative form,

$$\langle \boldsymbol{p}', -|\boldsymbol{p}, +\rangle = \langle \boldsymbol{p}'|U(\infty, 0)U(0, -\infty)|\boldsymbol{p}\rangle = \langle \boldsymbol{p}'|U(\infty, -\infty)|\boldsymbol{p}\rangle \qquad [34]$$

which is more frequently used in the relativistic theory. It can be shown that the set of all $|\boldsymbol{p}, +\rangle$ (the set of all $|\boldsymbol{p}, -\rangle$) together with the possible bound states makes a complete basis.

An observable in the Heisenberg picture is denoted as $O(t) = e^{(i/\hbar)Ht}Oe^{-(i/\hbar)Ht}$ and the same observable is denoted in the interaction picture with $O_0(t) = e^{(i/\hbar)H_0t}Oe^{-(i/\hbar)H_0t}$, which is obviously the Heisenberg picture for the free case. Taking into account

eqn [20], one has

$$\begin{aligned}
&\langle \boldsymbol{p}', -|O(t)|\boldsymbol{p}, +\rangle \\
&\quad = \langle \boldsymbol{p}'|U(\infty, t)O_0(t)U(t, -\infty)|\boldsymbol{p}\rangle \\
&\quad = \sum_{n=0}^{\infty}\left(-\frac{\mathrm{i}\lambda}{\hbar}\right)^n \sum_{s=0}^{n}\int_t^{\infty}\mathrm{d}t_1\cdots\int_t^{t_{s-1}}\mathrm{d}t_s\int_{-\infty}^{t}\mathrm{d}t_{s+1}\cdots \\
&\qquad \times\int_{-\infty}^{t}\mathrm{d}t_n\langle \boldsymbol{p}'|H_{1\mathrm{I}}(t_1)\cdots H_{1\mathrm{I}}(t_s)O_0(t)H_{1\mathrm{I}}(t_{s+1})\cdots \\
&\qquad \times H_{1\mathrm{I}}(t_n)|\boldsymbol{p}\rangle
\end{aligned} \tag{35}$$

The above equation can be easily generalized to the products of r operators in the following way:

$$\begin{aligned}
&\langle \boldsymbol{p}', -|T[O(t_1)O(t_2)\cdots O(t_r)]|\boldsymbol{p}, +\rangle \\
&\quad = \sum_{n=0}^{\infty}\left(-\frac{\mathrm{i}\lambda}{\hbar}\right)^n \int_{-\infty}^{\infty}\mathrm{d}\tau_1\int_{-\infty}^{\infty}\mathrm{d}\tau_2\cdots \\
&\qquad \times\int_{-\infty}^{\infty}\mathrm{d}\tau_n\langle \boldsymbol{p}'|T[O(t_1)O(t_2)\cdots O(t_r) \\
&\qquad \times H_{1\mathrm{I}}(\tau_1)H_{1\mathrm{I}}(\tau_2)\cdots H_{1\mathrm{I}}(\tau_r)]|\boldsymbol{p}\rangle
\end{aligned} \tag{36}$$

where $T[A(t_1)B(t_2)\cdots]$ denotes the time-ordered product (i.e., the product of the operators ordered according to decreasing times).

The above developments can be repeated with minor modifications for a system of N particles or in the second-quantization formalism. The single-particle states have only to be replaced with N-particle states $|\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots \boldsymbol{p}_N\rangle$ and $|\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots \boldsymbol{p}_N; \pm\rangle$. Some subtleties occur in the relativistic field theory, where the operator $U(t', t)$ converges for $t' \to +\infty$ and $t \to -\infty$ only in the weak sense.

In the second-quantization formalism and relativistic field theory, the ground-state or vacuum-correlation functions have a special interest. One can write (up to a phase factor)

$$\begin{aligned}
&\langle \Psi_0|T[O(t_1)O(t_2)\cdots O(t_r)]|\Psi_0\rangle \\
&\quad = \sum_{n=0}^{\infty}\left(-\frac{\mathrm{i}\lambda}{\hbar}\right)^n \int_{-\infty}^{\infty}\mathrm{d}\tau_1\int_{-\infty}^{\infty}\mathrm{d}\tau_2\cdots\int_{-\infty}^{\infty}\mathrm{d}\tau_n \\
&\qquad \times\langle \Phi_0|T[O(t_1)O(t_2)\cdots O(t_r) \\
&\qquad \times H_{1\mathrm{I}}(\tau_1)H_{1\mathrm{I}}(\tau_2)\cdots H_{1\mathrm{I}}(\tau_n)]|\Phi_0\rangle
\end{aligned} \tag{37}$$

where Φ_0 denotes the unperturbed ground state, and $\Psi_0 = U(0, \mp\infty)\Phi_0$ (the $\mp$ expressions are identical up to a phase factor) the actual ground state (vacuum state), the energy of which can be conventionally assumed to be 0.

The vacuum correlation functions contain, in principle, all the information about the system. They are also called Green functions and are used in many contexts. In particular, they can be related to the S-matrix and therefore to the scattering theory; they are the basis for a correct introduction of the idea of a quasiparticle and for establishing gap equations in the many-body system; they are a useful tool in the renormalization theory and for the treatment of the bound-state problem in field theory.

It is to be noted in the last connection that, using the completeness relation for the eigenstates of H $\left(\sum_n |u_n\rangle\langle u_n| = 1\right)$, one can write

$$\begin{aligned}
G(t) &= \langle \Psi_0|T[O(t)O(0)]|\Psi_0\rangle \\
&= \theta(t)\sum_n \langle \Psi_0|\mathrm{e}^{(\mathrm{i}/\hbar)Ht}O\mathrm{e}^{-(\mathrm{i}/\hbar)Ht}|u_n\rangle\langle u_n|O|\Psi_0\rangle \\
&\quad + \theta(-t)\sum_n \langle \Psi_0|O|u_n\rangle \\
&\quad \times \langle u_n|\mathrm{e}^{(\mathrm{i}/\hbar)Ht}O\mathrm{e}^{-(\mathrm{i}/\hbar)Ht}|\Psi_0\rangle \\
&= \sum_n \langle \Psi_0|O|u_n\rangle \\
&\quad \times \langle u_n|O|\Psi_0\rangle(\theta(t)\mathrm{e}^{-(\mathrm{i}/\hbar)W_n t} \\
&\quad + \theta(-t)\mathrm{e}^{-(\mathrm{i}/\hbar)W_n t})
\end{aligned} \tag{38}$$

where $\theta(t)$ is the step function: ($\theta(t) = 1$ for $t > 0$ and $\theta(t) = 0$ for $t < 0$). Then, for the Fourier transform, one obtains

$$\begin{aligned}
\hat{G}(\omega) &= \int_{-\infty}^{\infty}\mathrm{d}t\mathrm{e}^{(\mathrm{i}/\hbar)\omega t}G(t) \\
&= \mathrm{i}\hbar\sum_n \langle \Psi_0|O|u_n\rangle\frac{2W_n}{\omega^2 - W_n^2}\langle u_n|O|\Psi_0\rangle
\end{aligned} \tag{39}$$

Therefore, $\hat{G}(\omega)$ has poles at the discrete energy eigenvalues (bound states) in the complex ω plane and cuts on the continuum part of the spectrum. A similar argument can be repeated in the relativistic field theory; in this case, however, the relevant Green function is of the form $G(x) = \langle \Psi_0|T[\phi(x)\phi(0)]|\Psi_0\rangle$, $\phi(x)$ being the appropriate field. The Fourier transform, $\hat{G}(p) = \int d^4x\, \mathrm{e}^{(\mathrm{i}/\hbar)px}G(x)$ is on the space–time coordinates and the single-particle masses correspond to poles on the squared four-momentum $p^2 = p_0^2 - \boldsymbol{p}^2$.

Obviously, poles of this type are not present in the single terms of the perturbative expansion [37], if H_1 is identified with the entire interaction and, in particular, in the one-particle case, with the entire potential $V(r)$. This indicates a lack of convergence in the expansion for $\hat{G}(\omega)$ for ω close to the poles and a breakdown of the perturbative approach.

Actually, in some cases in field theories and in many-body theories, it is possible to isolate leading contributions in each term in [37] and to sum them up again, circumventing the difficulty. For instance, in a self-interacting scalar field theory, the various contributions to a two-point Green function can be

represented by Feynman graphs with vertices, internal lines, one incoming and one outgoing line (the single lines being associated with unperturbed two-point functions). Each graph can be divided into various irreducible segments (again beginning and ending with single lines but such that they cannot further be split in parts by cutting single lines). If $\mathrm{i}\Pi(p)$ is the sum of the Fourier transforms of the irreducible contributions of any order, the expansion can be reorganized in the following way:

$$\begin{aligned}\hat{G}(p) &= \frac{\mathrm{i}}{p^2-m^2} + \frac{\mathrm{i}}{p^2-m^2}\,\mathrm{i}\Pi(p)\frac{\mathrm{i}}{p^2-m^2} \\ &\quad + \frac{\mathrm{i}}{p^2-m^2}\,\mathrm{i}\Pi(p)\frac{\mathrm{i}}{p^2-m^2} \\ &\quad \times \mathrm{i}\Pi(p)\frac{\mathrm{i}}{p^2-m^2} + \cdots \\ &= \frac{\mathrm{i}}{p^2-m^2} + \frac{\mathrm{i}}{p^2-m^2}\,\mathrm{i}\Pi(p)\hat{G}(p) \end{aligned} \qquad [40]$$

The resulting equation can be solved explicitly

$$\hat{G}(p) = \frac{\mathrm{i}}{p^2-m^2+\Pi(p)} \qquad [41]$$

and the physical mass of the particle is given by the root of the denominator in p^2, m being simply a parameter with the dimension of a mass occurring in the Lagrangian or the Hamiltonian with no immediate meaning. By similar procedures one can obtain gap equations in many-body theories, Bethe–Salpeter equations for bound states in field theory, etc. Naturally, there exist even more essentially nonperturbative effects that cannot be obtained simply by resummation techniques.

Variational Method

This method uses primarily the property of the ground state to be the minimum of the so-called numerical range of H, that is, the set described by $\langle\phi|H|\phi\rangle$ for any vector ϕ such that $\langle\phi|\phi\rangle = 1$. In fact, setting

$$\phi = \sum_n c_n u_n, \quad \text{with } \sum_n |c_n|^2 = 1 \qquad [42]$$

it is clear that one has

$$\langle\phi|H|\phi\rangle = \sum_n |c_n|^2 W_n \geqslant W_0 \qquad [43]$$

W_0 being the ground-state energy.

Then, if a normalized state vector is considered, depending on a certain number of parameters $\phi(\alpha_1, \alpha_2, \ldots)$, it is clear that the minimum of the quantity

$$\langle\phi(\alpha_1,\alpha_2,\ \ldots)|H|\phi(\alpha_1,\alpha_2,\ \ldots)\rangle \qquad [44]$$

provides an upper bound to W_0, which can, however, be very close to the actual value. The vector $\phi(\alpha_1, \alpha_2, \cdots)$ is called a test function. The more appropriate the choice of such a test function, the better the estimate. In the case of a hydrogen atom (eqn [12]) setting $\phi(\alpha) = \sqrt{\alpha^3/\pi}\mathrm{e}^{-\alpha r}$, one obtains the exact expression for W_0 and the corresponding eigenfunction. This obviously is due to the specific form of the test function that includes the exact eigenfunction for a $\alpha = m_\mathrm{e}e^2/(4\pi\hbar^2)$ (the reciprocal of the Bohr radius). In general, using the variational method, the accuracy in the determination of the eigenvalue is much better than in the determination of the eigenfunction.

As a less trivial example, the determination of the ground state of the helium atom may be considered. Neglecting spin-dependent forces, the Hamiltonian is

$$\begin{aligned} H &= H_0 + \frac{e^2}{4\pi r_{12}} \\ &= \frac{\boldsymbol{p}_1^2}{2m_\mathrm{e}} + \frac{\boldsymbol{p}_2^2}{2m_\mathrm{e}} - \frac{2e^2}{4\pi r_1} - \frac{2e^2}{4\pi r_2} + \frac{e^2}{4\pi r_{12}} \end{aligned} \qquad [45]$$

where r_1, r_2, and r_{12} are the distances of the two electrons from the nucleus and the distance between the two electrons, respectively. The ground state of H_0 (i.e., of H when the interaction between the two electrons is neglected), satisfying the Pauli principle requirements, can be written in the form

$$\begin{aligned}\phi_0 &= u_{1S}(\boldsymbol{r}_1)u_{1S}(\boldsymbol{r}_2)\frac{1}{\sqrt{2}} \\ &\quad \times (v_{1,\frac{1}{2}}v_{2,-\frac{1}{2}} - v_{1,-\frac{1}{2}}v_{2,\frac{1}{2}}) \end{aligned} \qquad [46]$$

where $u_{1S}(\boldsymbol{r}) = (1/\sqrt{\pi})(2/a_0)^{3/2}\mathrm{e}^{-(2r/a_0)}$ and $a_0 = 4\pi\hbar^2/(m_\mathrm{e}e^2)$. The experimental value of W_0 is $-78.99\,\mathrm{eV}$ (i.e., 78.99 is the energy necessary to strip both electrons from the atom). If the interaction between the electrons is treated as a perturbation, one finds, applying eqn [9],

$$\left\langle \phi_0 \middle| \frac{e^2}{4\pi r_{12}} \middle| \phi_0 \right\rangle = \frac{5}{2}\,w_0 \qquad [47]$$

($w_0 = 13.6$, defined as in [14] and so $W_0 = -2\times 4w_0 + \frac{5}{2}w_0 = -\frac{11}{16}w_0 = -74.8\mathrm{eV}$. However, if one replaces in [46] the function $u_{1S}(\boldsymbol{r})$ with $u_{1S}^{\mathrm{test}}(\boldsymbol{r}) = \pi)(Z/a_0)^{3/2}\mathrm{e}^{-Zr/a_0}$, one obtains by simply rescaling

$$\langle\phi^{\mathrm{test}}|H|\phi^{\mathrm{test}}\rangle = (2Z^2 - 8Z + \tfrac{5}{4}\,Z)w_0 \qquad [48]$$

This expression is reduced to the preceding one for $Z = 2$; however, it attains its minimum for $Z = 27/16$ and using such a value, it gives $W_0 = -5.70\,w_0 = -77.5\,\text{eV}$ in much better agreement with the data. It is to be noted that Z can be interpreted as a kind of effective nuclear charge acting on each electron, due to the partial screen produced by the other electron.

In case there exist other constants of motion, such as the angular momentum or components of the angular momentum, the test function should be chosen as a common eigenfunction. Then, the same method can be used to determine the ground state of H inside every invariant subspace specified by the values of such quantities. In this way, the lowest triplet S state, singlet and triplet P states can be determined for the same helium atom by variants of the test function.

The method can be successfully applied in a similar way to a large variety of problems. In particular, it can be applied to the evaluation of the binding potential among nuclei in molecules in the Born–Openheimer approximation. This consists in treating the nuclei as fixed, and in studying the electron ground-state energy as a function of the nuclei positions. This problem can be exactly solved (in a sense) only in the very special case of the ionized hydrogen molecule, but in many cases the variational method offers a simple and at least a qualitative understanding of the properties of the molecule.

Up to now, the attention has been restricted to the case of the ground state of a system, or the ground state inside an invariant subspace of states, that is, to the actual minima of the functional $W[\phi] = \langle \phi|H|\phi \rangle$ under the constraint $\langle \phi|\phi \rangle = 1$. However, it is worth mentioning that the same functional is stationary at every eigenstate of H and vice versa, under the same constraint. Using the Lagrange multiplier theorem, the condition for $W[\phi]$ to be stationary can in fact be written as

$$\begin{aligned}\delta(W[\phi] - \lambda\langle \phi|\phi \rangle) &= \langle \delta\phi|H|\phi \rangle + \langle \phi|H|\delta\phi \rangle \\ &\quad - \lambda(\langle \delta\phi|\phi \rangle + \langle \phi|\delta\phi \rangle) \\ &= 0\end{aligned} \qquad [49]$$

Combining this equation with the one obtained by replacing $\delta\phi$ with $-\mathrm{i}\delta\phi$, one immediately obtains $\langle \delta\phi|(H - \lambda)|\phi \rangle = 0$ and then, due to the arbitrariness of $\delta\phi$,

$$H\phi = \lambda\phi \qquad [50]$$

which proves the statement.

In applying the above result to the determination of the excited states, one should choose test functions orthogonal to those corresponding to lower eigenvalues and, since approximations are involved, it is clear that this cannot be done with the same confidence as for the ground state. However, such generalization is important for theoretical developments and to justify more elaborate methods, such as, for instance, the famous Hartree–Fock method.

The Hamiltonian for a many-electron atom is

$$H = \sum_{i=1}^{Z}\left(\frac{p_i^2}{2m_\mathrm{e}} - \frac{Ze^2}{r_i}\right) + \sum_{i<j}\frac{e^2}{r_{ij}} \qquad [51]$$

which generalizes eqn [45]. One can expect that for large Z, a qualitative understanding of the spectrum can be achieved by replacing the complicated interaction acting on each electron according to [51] with an average central potential $V(r)$ and by simply writing $H = \sum_{i=1}^{Z}[(p_i^2/2m_\mathrm{e}) - V(r_i)]$. This suggests that the test function to be chosen is (once again consistent with the exclusion principle)

$$u^{\text{test}}_{a_1,a_2,\ldots a_z}(1, 2, \ldots, Z) = \det[u_{a_i}(j)] \qquad [52]$$

in which 1, 2, …,Z denote briefly the coordinates and the spin indices for the various electrons, and $u_{a_i} = R_{n_i,l_i}(r_i)Y_{l_i,m_i}(\theta_i, \phi_i)v_{m_{si}}$, the one-particle eigenfunctions corresponding to the quantum numbers $a_i = (n_i, l_i, m_i, m_{si})$ for the hypothetical potential $V(r)$. Leaving actually the radial functions $R_{nl}(r)$ undetermined and treating them as variational parameters, one obtains the Hartree–Fock equations (strictly speaking the simple form [52] applies only to complete shells, to one electron plus a complete shell, to a complete shell minus one electron).

Semiclassical Methods

Among the semiclassical methods, the WKB method is a classic one. However, in its simplest form this method applies only to problems that are or can be reduced to one-dimensional (as in the case of a particle in a central potential $V(r)$) and has today little more than historical interest.

It consists in setting $u(x) = \mathrm{e}^{(\mathrm{i}/\hbar)S(x)}$ in the eigenvalue equation

$$\frac{\hbar^2}{2m}\frac{d^2u(x)}{\mathrm{d}x^2} + V(x)u(x) = Wu(x) \qquad [53]$$

and in solving the resulting equation

$$\left(\frac{\mathrm{d}S}{\mathrm{d}x}\right)^2 - \mathrm{i}\hbar\frac{\mathrm{d}^2S}{\mathrm{d}x^2} = 2m(W - V(x)) \qquad [54]$$

by an expansion in $\hbar$,

$$S = S_0 - i\hbar S_1 + (-i\hbar)^2 S_2 + \cdots \qquad [55]$$

Keeping only the first two terms in the expansion, one obtains the two independent solutions of the original equation

$$u_{\pm}(x) = \frac{C_{\pm}}{[p(x)]^{1/2}} \exp\left(\pm\frac{i}{\hbar}\right) \int_0^x dx' p(x') \qquad [56]$$

where $p(x) = \sqrt{2m[W - V(x)]}$ and $C_{\pm}$ are arbitrary constants. In a region in which $V(x) < W$ (classically accessible region), $p(x)$ is the classical momentum of the particle in point x, and $u_{\pm}(x)$ are oscillating functions. When asking for a real solution, the latter can be also summarized as

$$u(x) = \frac{A}{[p(x)]^{1/2}} \sin\left[\frac{1}{\hbar}\int_0^x dx' p(x') + \alpha\right] \qquad [57]$$

A and α being arbitrary constants. On the contrary, in a region in which $V(x) > W$ (classically inaccessible region), [56] becomes

$$\bar{u}_{\pm}(x) = \frac{C'_{\pm}}{[\bar{p}(x)]^{1/2}} \exp\left(\pm\frac{1}{\hbar}\int_0^x dx' \bar{p}(x')\right) \qquad [58]$$

with $(\bar{p}(x) = \sqrt{2m[V(x) - W]})$. These have a stable behavior and vanish or diverge for $x \to \pm\infty$, if the region extends to such limits.

By considering the order of magnitude of S_2, one can see that the above approximations are good as long as $V(x)$ does not change appreciably for a variation of x of the order of $\hbar/p(x)$ or $\hbar/\bar{p}(x)$. Such conditions cannot hold, obviously, in the vicinity of an inversion point x_0, that is, a point for which $V(x_0) = W$ and $p(x)$ and $\bar{p}(x)$ vanish. However, using a linear approximation for $V(x)$ in the neighborhood of such a point, an appropriate rule can be given to connect solutions of the types [57] and [58] through it.

The above rule is not explicitly reported here. However, the two most interesting situations that can occur are only mentioned. Assuming that $V(x)$ vanishes for $x \to \pm\infty$, let the potential have a negative minimum $-V_0$ from where it rises regularly to 0 in both directions (potential well). The other case considered here involves a positive maximum from where similarly it decreases uniformly (potential barrier). In both cases, if W has an intermediate value between $-V_0$ and 0 or between 0 and V_0, there exist two inversion points x_1 and x_2. In the first situation, the classically nonaccessible region is the external part of the interval (x_1, x_2), and the condition for which an internal solution of the form [57] is connected simultaneously with a $\bar{u}_+(x)$ (vanishing for $x \to -\infty$) through x_1 and with a $\bar{u}_-(x)$ (vanishing for $x \to +\infty$) through x_2, respectively, is

$$2\int_{x_1}^{x_2} dx\, p(x) = \left(n + \frac{1}{2}\right)h \quad (n = 0, 1, \ldots) \qquad [59]$$

which allows only discrete values for W, for $W < 0$, and provides a justification of the famous Sommerfeld conditions of the old quantum mechanics.

In the second situation, the classically inaccessible region is the interval (x_1, x_2) itself. The energy spectrum is purely continuous and the wave functions are different from 0 in the entire x-axis. Then, there is a finite probability that the particle, initially on the left of x_1, goes through the barrier and is found on the right of x_2 after some time (tunnel effect). Such a probability is expressed by the transmission coefficient

$$\tau = \exp\left(-\frac{2}{\hbar}\int_{x_1}^{x_2} dx\, \bar{p}(x)\right) \qquad [60]$$

Equation [60] has been largely used in discussing nuclear reactions and particularly in the theory of α radioactivity.

Finally, methods related in some way to the preceding one, but of much more general application, are those based on the path integral formalism.

One may consider, for simplicity, a particle in one dimension with its position and momentum operators denoted by Q and P, and with $|q\rangle$ and $|p\rangle$ their respective eigenstates. Let $Q(t) = e^{(i/\hbar)Ht} Q e^{(-i/\hbar)Ht}$, and $|q, t\rangle = e^{(i/\hbar)Ht}|q\rangle$ be the corresponding operators and eigenvectors in the Heisenberg picture, with $H = P^2/2m + V(Q)$. For an infinitesimal ε, one has

$$\begin{aligned}
&\langle q', t+\varepsilon | q, t\rangle \\
&\quad = \left\langle q' \middle| \left(1 + \frac{i}{\hbar} H\varepsilon\right) \middle| q \right\rangle \\
&\quad = \delta(q' - q) + \frac{i}{\hbar}\varepsilon\left[\int dp \langle q'|p\rangle \frac{p^2}{2m} \langle p|q\rangle \right. \\
&\qquad \left. + \delta(q' - q) V\left(\frac{q+q'}{2}\right)\right] \\
&\quad = \int \frac{dp}{2\pi\hbar} \exp\left\{\left(\frac{i}{\hbar}\right)\varepsilon\left[p(q'-q) + \frac{p^2}{2m} + V\left(\frac{q+q'}{2}\right)\right]\right\}
\end{aligned} \qquad [61]$$

where $\langle q|p\rangle = (1/2\pi\hbar)e^{(i/\hbar)pq}$ and $\delta(q - q') = (1/2\pi\hbar)\int dp e^{(i/\hbar)p(q'-q)}$ have been used. Then, for a finite time interval (t, t'), by dividing this into a small subinterval (t_{n-1}, t_n) with $t_n = t_{n-1} + \varepsilon$, one

can write

$$\begin{aligned}
&\langle q', t'|q, t\rangle \\
&\quad = \int \mathrm{d}q_1 \mathrm{d}q_2 \ldots \mathrm{d}q_{N-1} \langle q', t'|q_{N-1}, t_{N-1}\rangle \cdots \\
&\qquad \times \langle q_2, t_2|q_1, t_1\rangle \langle q_1, t_1|q, t\rangle \\
&\quad = \int \mathrm{d}p_1 \mathrm{d}q_1 \ldots \mathrm{d}p_{N-1} \mathrm{d}q_{N-1} dp_N \\
&\qquad \times \exp\left\langle \frac{\mathrm{i}}{\hbar} \sum_{n=1}^{N} \left\{ p_n(q_n - q_{n-1}) \right.\right. \\
&\qquad \left.\left. -\varepsilon\left[\frac{p_n^2}{2m} + V\left(\frac{q_n + q_{n-1}}{2}\right)\right]\right\}\right\rangle
\end{aligned} \tag{62}$$

where $q_N = q'$. In the limit $\varepsilon \to 0$, eqn [62] can be interpreted as an integral over the infinite paths connecting q with q' and written in short

$$\begin{aligned}
&\langle q', t'|q, t\rangle \\
&\quad = \int_{q(t)=q}^{q(t')=q'} \mathscr{D}p \mathscr{D}q \exp\left\{\frac{\mathrm{i}}{\hbar} \int_t^{t'} \mathrm{d}\tau \left[p(\tau)\dot{q}(\tau) \right.\right. \\
&\qquad \left.\left. -\frac{p^2(\tau)}{2m} + V(q(\tau))\right]\right\}
\end{aligned} \tag{63}$$

If one sets formally

$$\begin{aligned}
\mathscr{D}q &= \lim_{\varepsilon \to 0}\left(\frac{m}{2\pi\hbar \mathrm{i}\varepsilon}\right)^{N/2} \mathrm{d}q_1 \mathrm{d}q_2 \cdots \mathrm{d}q_{N-1} \\
\mathscr{D}p &= \lim_{\varepsilon \to 0}\left(\frac{\mathrm{i}\varepsilon}{2\pi\hbar m}\right)^{N/2} \mathrm{d}p_1 \mathrm{d}p_2 \cdots \mathrm{d}p_N
\end{aligned} \tag{64}$$

one can also explicitly integrate the momenta and write

$$\begin{aligned}
&\langle q', t'|q, t\rangle \\
&\quad = \int_{q(t)=q}^{q(t')=q'} \mathscr{D}q \exp\left\{\frac{\mathrm{i}}{\hbar} \int_t^{t'} \mathrm{d}\tau \left[\frac{1}{2}m\dot{q}^2(\tau) - V(q(\tau))\right]\right\}
\end{aligned} \tag{65}$$

It is to be noted that eqn [63] is in terms of paths in the phase space, eqn [65] in terms of paths in the configuration space. It is also to be noted that the quantity $S(q', t'; q, t) = \int_t^{t'} \mathrm{d}\tau [\frac{1}{2}m\dot{q}^2(\tau) - V(q(\tau))] = \int_t^{t'} \mathrm{d}\tau\, L[q(\tau), \dot{q}(\tau)]$ occurring in [65] and the corresponding one occurring in [63] are the classical actions for the particular paths in the configuration and the phase space, respectively.

To evaluate $\langle q', t'|q, t\rangle$ is obviously equivalent to solving the Schrödinger equation [1]. Since the classical action is stationary over the classical solution of the equation of motion joining q and q', the quantum theory can be viewed as the result of quantum fluctuations around the classical solution.

One could now consider the Green functions in this formalism. In analogy with what has been done in the section "Green functions," one can start from the expression

$$\begin{aligned}
&\langle q', t'|Q(t_1)Q(t_2)\cdots Q(t_r)|q, t\rangle \\
&\quad = \int \mathrm{d}q_1 \mathrm{d}q_2 \cdots \mathrm{d}q_r \langle q, t|q_1, t_1\rangle q_1 \\
&\qquad \times \langle q_1, t_1|q_2, t_2\rangle q_2 \cdots \langle q_{r-1}, t_{r-1}|q_r, t_r\rangle \\
&\qquad \times q_r \langle q_r, t_r|q, t\rangle
\end{aligned} \tag{66}$$

Assuming $t_1 > t_2 > \cdots t_r$, this can be rewritten as

$$\begin{aligned}
&\langle q', t'|T[Q(t_1)Q(t_2)\cdots Q(t_r)]|q, t\rangle \\
&\quad = \int_{q(t)=q}^{q(t')=q'} \mathscr{D}q \exp\left[\frac{\mathrm{i}}{\hbar}\int_t^{t'} \mathrm{d}\tau L(q, \dot{q})\right] \\
&\qquad \times q(t_1)q(t_2)\cdots q(t_r)
\end{aligned} \tag{67}$$

or in the corresponding form in terms of phase space paths (see [63]). Obviously, written with the explicit insertion of the time-ordering prescription T, eqn [67] remains valid for any possible ordering of t_1, t_2,…, t_r.

If the times t and t' have a small negative phase $\mathrm{e}^{-\mathrm{i}\delta}$, one can perform the limits $t \to -\infty$ and $t' + \infty$ and obtain

$$\begin{aligned}
&\langle q', t'|q, t\rangle \\
&\quad = \sum_n \langle q'|u_n\rangle \mathrm{e}^{-(\mathrm{i}/\hbar+\delta)W_n(t'-t)} \langle u_n|q\rangle \\
&\qquad \to \langle q'|\Psi_0\rangle \mathrm{e}^{-(\mathrm{i}/\hbar+\delta)W_0(t'-t)} \langle \Psi_0|q\rangle
\end{aligned} \tag{68}$$

Ψ_0 being again the ground state of the system. Similarly, one also has

$$\begin{aligned}
&\left\langle q', t'|T\left[Q(t_1\mathrm{e}^{-\mathrm{i}\delta})Q(t_2\mathrm{e}^{-\mathrm{i}\delta})\cdots Q(t_r\mathrm{e}^{-\mathrm{i}\delta})\right]|q, t\right\rangle \\
&\quad \to \langle q'|\Psi_0\rangle \mathrm{e}^{-(\mathrm{i}/\hbar+\delta)W_0(t'-t)} \langle \Psi_0|q\rangle \langle \Psi_0| \\
&\qquad \times T\left[Q(t_1\mathrm{e}^{-\mathrm{i}\delta})Q(t_2\mathrm{e}^{-\mathrm{i}\delta})\cdots Q(t_r\mathrm{e}^{-\mathrm{i}\delta})\right]|\Psi_0\rangle
\end{aligned} \tag{69}$$

Combining the last two equations with [65] and [67], one obtains finally (suppressing phase δ)

$$\begin{aligned}
&\langle \Psi_0|T[Q(t_1)Q(t_2)\cdots Q(t_r)]|\Psi_0\rangle \\
&\quad = \frac{\int \mathscr{D}q \exp\left[\frac{\mathrm{i}}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}\tau L(q, \dot{q})\right] q(t_1)q(t_2)\cdots q(t_r)}{\int \mathscr{D}q \exp\left[\frac{i}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}\tau L(q, \dot{q})\right]}
\end{aligned} \tag{70}$$

where any specification of boundary conditions has been suppressed, since in the limit the right-hand expression is independent of it.

The preceding equation can be immediately extended with obvious modifications to more complicated systems, and in particular to bosonic fields and even to fermionic ones, if in the latter case one makes use of the Grassmann variable formalism. In field theory, obviously, the Lagrangian $L(q(\tau), \dot{q}(\tau))$ must be replaced by a Lagrangian density $\mathscr{L}[\phi(x)\partial^{\mu}\phi(x)]$, variables $t_1, t_2, \ldots \tau, \tau_1, \ldots$ with space–time point specifications $x_1, x_2, \ldots, \xi, \xi_1, \ldots,$ etc. The method turns out to be particularly convenient in treating gauge field theories, such as quantum electrodynamics (QED) or non-Abelian theories, as the electroweak theory and quantum chromodynamics (QCD).

Equation [70] can be used in various ways. Writing $L(q, \dot{q}) = L_0 + \lambda L_1$ enables one to establish the perturbative expansion

$$\begin{aligned}
&\langle \Psi_0 | T[Q(t_1)Q(t_2)\cdots Q(t_r)] | \Psi_0 \rangle \\
&= \left\{ \sum_{s=0}^{\infty} \left(\frac{\mathrm{i}}{\hbar}\lambda\right)^s \frac{1}{s!} \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \cdots \right. \\
&\quad \times \int_{-\infty}^{\infty} \mathrm{d}\tau_s \int \mathscr{D}q \exp\left[\frac{\mathrm{i}}{\hbar} \int_{-\infty}^{\infty} \mathrm{d}\tau\, L_0(q(\tau), \dot{q}(\tau))\right] \\
&\quad \times q(t_1)q(t_2)\cdots q(t_r) L_1(q(\tau_1), \dot{q}(\tau_1)) \\
&\quad \left. \times L_1(q(\tau_2), \dot{q}(\tau_2)) \cdots L_1(q(\tau_s), \dot{q}(\tau_s)) \right\} \\
&\quad \times \left\{ \sum_{s=0}^{\infty} \left(\frac{\mathrm{i}}{\hbar}\lambda\right)^s \frac{1}{s!} \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \cdots \int_{-\infty}^{\infty} \mathrm{d}\tau_s \int \mathscr{D}q \right. \\
&\quad \times \exp\left[\frac{\mathrm{i}}{\hbar} \int_{-\infty}^{\infty} \mathrm{d}\tau\, L_0(q(\tau), \dot{q}(\tau))\right] \\
&\quad \left. \times L_1(q(\tau_1), \dot{q}(\tau_1)) L_1(q(\tau_2), \dot{q}(\tau_2)) \cdots L_1(q(\tau_s), \dot{q}(\tau_s)) \right\}^{-1}
\end{aligned} \tag{71}$$

In the one-dimensional example considered above, L_0 can be the kinetic energy and $L_1 = -V(q)$, or better L_0 a harmonic oscillator Lagrangian and L_1 some anharmonic term. In field theory, $\mathscr{L}_0(x)$ is typically assumed to be the kinetic part of $\mathscr{L}(x)$, and $\mathscr{L}_1(x)$ the interaction. In any case, if L_1 does not depend on time derivatives of the basic variables, one has $L_1 = -H_1$ and eqn [71] has the same content as eqn [36]. If L_1 depends on time derivatives, $L_1 \neq H_1$ and [71] and [36] are formally different. However, it can be shown that the modifications, implied in the last case, in the definition of the conjugate momenta cancel the modifications in the definition of H_1 (normal dependent terms). Then, even in this case, eqn [36] can be recast in the simpler form corresponding to [71].

In fact, eqn [70] can be used for nonperturbative developments. In particular, one can fix a small but finite value for the lattice spacing ε introduced above, specify the path $q(\tau)$ in terms of a finite number of variables as in [62], and interpret the measure $\mathscr{D}q$ as in [64], without performing the actual limit. In this way, a theory with a discrete time or a discrete space–time is obtained (lattice theory), which has to be understood as an approximation to the continuous theory. To such a procedure, often there are not many alternatives in field theories for an understanding of essentially nonperturbative effects, such as color confinement in QCD. The resulting expression can be treated analytically in some cases. More frequently, it can be used for direct numerical evaluations.

Actually, to make integrals well defined, it is convenient to work in the corresponding Euclidean theory that is obtained from [62] by performing the analytic continuation to an imaginary time, $\tau \to -\mathrm{i}\tau$, $t_1 \to -\mathrm{i}t_1, \ldots$ (corresponding to $\delta = \pi/2$ in [69]). This amounts to writing

$$\begin{aligned}
&\langle \Psi_0 | T[Q_{\mathrm{E}}(t_1)Q_{\mathrm{E}}(t_2)\cdots Q_{\mathrm{E}}(t_r)] | \Psi_0 \rangle \\
&= \frac{\int \mathscr{D}q \exp\left[-\frac{1}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}\tau L_{\mathrm{E}}(q, \dot{q})\right] q(t_1)q(t_2)\cdots q(t_r)}{\int \mathscr{D}q \exp\left[-\frac{1}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}\tau L(q, \dot{q})\right]}
\end{aligned} \tag{72}$$

where one has set $Q_{\mathrm{E}}(t) = Q(-\mathrm{i}t) = \mathrm{e}^{(1/\hbar)Ht} Q \mathrm{e}^{-(1/\hbar)Ht}$ and $L_{\mathrm{E}}(q, \dot{q}) = -L(q, \mathrm{i}\dot{q})$. For example, in the one-dimensional example, one has $L_{\mathrm{E}} = \frac{1}{2}m\dot{q}^2 + V(q)$, which turns out to be formally identical to the Hamiltonian.

An alternative approach to consider the infinite time limit in [65] is to set $q' = q$ and integrate with respect to such a variable. One has then in the Euclidean formalism

$$\begin{aligned}
&\mathrm{Tr}\{T[Q_{\mathrm{E}}(t_1)Q_{\mathrm{E}}(t_2)\cdots Q_{\mathrm{E}}(t_r)]\} \\
&= \int_{\text{periodic}} \mathscr{D}q \exp\left[-\frac{1}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}\tau L_{\mathrm{E}}(q, \dot{q})\right] \\
&\quad \times q(t_1)q(t_2)\cdots q(t_r)
\end{aligned} \tag{73}$$

where the integration is over paths satisfying periodic boundary conditions, $q(t') = q(t)$. In the limit of large time, this equation has essentially the same content of [72] and is often more convenient for numerical applications. On the contrary, for finite $t' - t$ one can set $t' - t = \hbar\beta$ thinking of β as $1/kT$, T being the absolute temperature and k the Boltzmann constant. In this form, for $r = 0$, the right-hand side of [73] becomes identical to the partition function $Z(\beta) = \mathrm{Tr}\,\mathrm{e}^{-\beta H}$, and for $r \neq 0$ to the corresponding correlation functions.

This establishes a connection with statistical mechanics, and [73] becomes the basis for important developments, both analytical and numerical, in this field.

See also: Electrodynamics: Quantum Electrodynamics; Electronic Structure (Theory): Atoms; Hartree and Hartree–Fock Methods in Electronic Structure; Quantum Mechanics: Foundations; Statistical Mechanics: Quantum.

PACS: 03.65.−w; 03.65.Nk; 09.65.Sq; 05.30.−d; 31.15.Pf; 31.15.Hd

Further Reading

Caldirola P, Cirelli R, and Prosperi GM (1982) *Introduzione alla Fisica Teorica* (in Italian). Torino: UTET.

Das A (1993) *Field Theory, a Path Integral Approach*. Singapore: World Scientific.

Fetter AL and Walecka JD (1971) *Quantum Theory of Many-Particle Systems*. New York: Dover.

Galindo A and Pasqual P (1990, 1991) *Quantum Mechanics I, II*. New York: Springer.

Gasiorowicz S (1996) *Quantum Physics*, 2nd edn. New York: Wiley.

Landau LD and Lifschitz EM (1974) *Quantum Mechanics (Non-relativistic Theory)*. Oxford: Pergamon.

Le Bellac M (1988) *Des phénomènes critiques aux champs de jauge* (in French). Paris: InterEditions.

Martin PA and Rothen F (2002) *Many-Body Problems and Field Theory*. New York: Springer.

Negele JW and Orland H (1998) *Quantum Many-Particle Systems*. London: Wesley.

Siegel W. *Fields*. http://insti.physics.sunysb.edu/siegel/plan.htm/.

Weinberg S (1995, 1996, 2000) *Theory of Fields, I, II, III*. Cambridge: Cambridge University Press.

Zinn-Justin J (2002) *Quantum Field Theory and Critical Phenomena*. 4th edn. Oxford: Clarendon.

Quantum Mechanics: Molecules

V Schettino and M Mugnai, Università di Firenze, Florence, Italy

Introduction

A molecule is an ensemble of atoms held together by electrostatic forces or, in the language of chemistry, by chemical bonds. Stable molecules are considered as examples in the following; but from the quantum mechanical point of view unstable and short-lived molecules, which are amenable to direct experimental observation with available time-resolved techniques, can also be treated on the same footing. The quantum mechanics of molecules is a complicated many-body problem whose complete treatment should include the degrees of freedom of the electrons and the nuclei. In favorable cases, the problem can be simplified by exploiting the molecular symmetry that can allow some qualitative or semiquantitative understanding as well. Actually, an exact numerical solution of the Schrödinger equation can be found only for the hydrogen molecular ion H_2^+, which, at least for linear molecules, can serve as a reference, very much like the hydrogen atom for many-electron atoms. In the most general case, the quantum mechanical treatment of polyatomic molecules can only be carried to obtain approximate solutions for the molecular energy and eigenfunctions. However, the dissociation energy of a chemical bond is in the range of a few electronvolts. Therefore, to reach an accuracy of significance, in many chemical and molecular problems the solution of the Schrödinger equation of the molecule must be carried to a reasonably high level of approximation. The quantum mechanical treatment of polyatomic molecules is generally carried out in the so-called molecular orbital method or, alternatively, in the valence band approach. These methods are presented in the following and compared with the alternative approach of the density-functional theory (DFT). As an example, the molecular orbital results of biatomic molecules are described qualitatively. The application of group theory and symmetry considerations to polyatomics are also discussed. The nuclear Schrödinger equation and its separation into the translational, rotational, and vibrational terms are also presented and the connection with various spectroscopic techniques are discussed.

The Born–Oppenheimer Approximation

One is generally compelled to consider the quantum mechanics of a molecule in the framework of the Born–Oppenheimer approximation. Taking advantage of the fact that the electron mass is several orders of magnitude smaller than that of the nuclei, the total wave function is written as a product of an electronic wave function Ψ_e, which depends on the electron coordinates q_i and parametrically on the nuclear coordinates q_α, and a

nuclear part Ψ_n:

$$\Psi(q_i q_\alpha) = \Psi_e \Psi_n = \Psi_e(q_i \bar{q}_\alpha)\Psi_n(q_\alpha) \qquad [1]$$

The Schrödinger equation is then separated into an electronic equation with the Hamiltonian

$$H_e = T_e + T_{en} + T_{ee} + T_{nn} \qquad [2]$$

including, in order, the electronic kinetic energy, the electron–nucleus interaction, the interelectron repulsion and the nucleus–nucleus repulsion (the last term can be considered as a constant at any given nuclear configuration). An equation for the nuclear motion with a Hamiltonian, where the total electron energy U_t contributes to the potential, is given by

$$H_n \psi_n = (T_n + U_t)\psi_n = E_n \psi_n \qquad [3]$$

where T_n is the kinetic energy of the nuclei and U_t is the total electron energy corresponding to the Hamiltonian [2]. In essence, the problem is similar to one of solving the electronic problem at fixed nuclear configurations and searching for the minimum corresponding to the equilibrium structure (see **Figure 1**). An approximate solution of the electronic wave equation is generally attempted by the variational method.

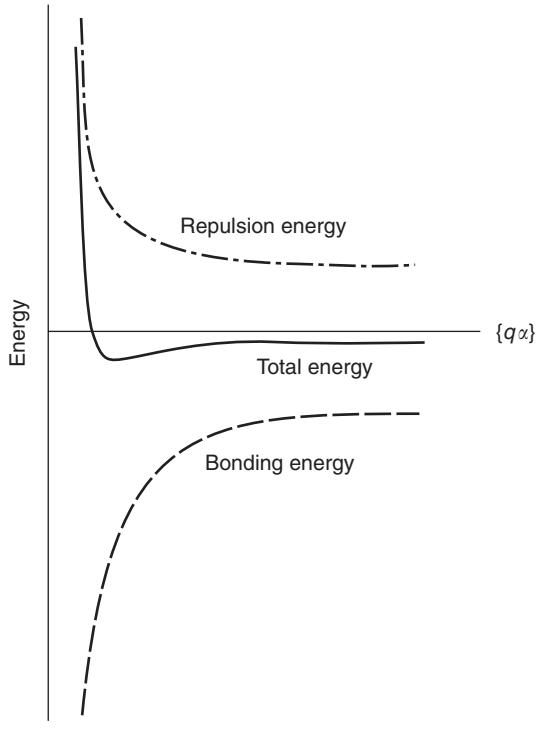


Figure 1 Variation of the bonding and repulsion energy as a function of the intermolecular coordinates q_α.

The Molecular Orbital Method

In the molecular orbital method, the multi-electron wave function $\Psi_e(q_i \bar{q}_\alpha)$ is expressed as an antisymmetrized product of single electron spin orbitals, written as a Slater determinant. The electronic wave function Ψ_e corresponds to a given electron configuration where the electrons are allocated in molecular orbitals according to the Pauli principle. For a $2n$-electron molecule, the ground-state configuration is shown in **Figure 2a** and the corresponding Slater determinant is given by

$$\Psi_e(q_i \bar{q}_\alpha) = N|\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2) \cdots \varphi_n(2n-1)\alpha(2n-1)\varphi_n(2n)\beta(2n)| \qquad [4]$$

where α and β denote the spin eigenfunction.

The general purpose of the method is to reduce the multi-electron wave equation into an effective one-electron equation for the molecular orbitals. Separating the one- and two-particle moieties of the electronic Hamiltonian, one obtains the variational integral in the form

$$E = \langle \Psi_e | H | \Psi_e \rangle = E_1 + E_2 = 2\sum_i H_{ii} + \sum_{ij}(2J_{ij} - K_{ij}) \qquad [5]$$

which exposes the contributions to the energy from the single electron in the molecular orbital φ_i in the nuclear field H_{ii} and the electron–electron repulsion including the Coulomb contribution J_{ij} that represents the electronic repulsion of electron 1 in orbital i, electron 2 in orbital j, and the exchange contribution K_{ij} arising from the antisymmetric character of the multi-electronic wave function. Applying the variational principle to the above integral and defining appropriate Coulomb $\hat{J}$ and exchange $\hat{K}$ operators, the minimization procedure (preserving

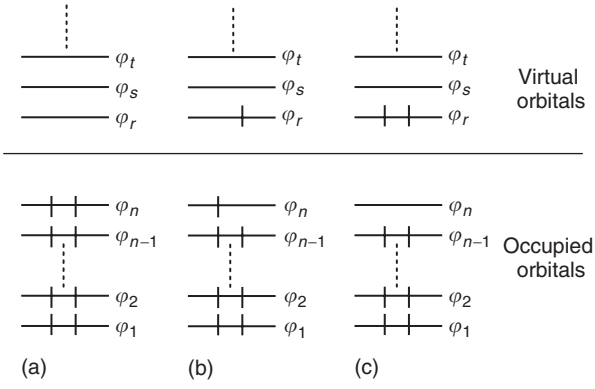


Figure 2 Electronic state configuration of a $2n$-electron molecule. (a) Ground-state configuration; (b) a singly excited configuration; and (c) a doubly excited configuration.

the orthonormalization of the basis functions) leads to the pseudo-eigenvalue-equation

$$\hat{F}\varphi_i(1) = \left(H_1^c + \sum_j (2\hat{J}_j(1) - \hat{K}_j(1))\right)\varphi_i(1) = \sum_j \varepsilon_{ij}\varphi_j(1) \quad [6]$$

called the Fock equation. The Fock operator $\hat{F}$ is an effective one-electron Hamiltonian including an averaged electron–electron repulsion. Since the operator $\hat{F}$ depends on the solutions, the molecular orbitals can be determined by an iterative procedure starting from an initially guessed form. From a practical point of view, the starting form of the molecular orbitals is chosen as a linear combination of the atomic orbitals φ_μ of the atoms that constitute the molecule

$$\phi_i = \sum_\mu c_{\mu i}\varphi_\mu \quad [7]$$

This reconduces the approach into the framework of the linear variational method. The integrodifferential Fock equation [6] is transformed into a set of linear algebraic equations in the c coefficients (Roothan equations). The problem is conveniently expressed in matrix notation. Defining the overlap matrix

$$S_{\mu v} = \langle \varphi_\mu | \varphi_v \rangle \quad [8]$$

the electron density matrix

$$P_{\mu v} = \left(\sum_i 2c_{\mu i}^* c_{vi}\right) \quad [9]$$

and the Fock matrix F (the matrix of the operator [6]), the Roothan equations can be written in the form

$$FC = CSE \quad [10]$$

which is evidently amenable to an iterative solution. The selected basis set of the atomic orbitals can be transformed into an orthonormal basis set to have the Roothan equations in the form of an eigenvalue problem.

The atomic orbitals of the basis set contain an angular part expressed as spherical harmonics and a radial part in the form of Slater orbitals expressing an exponential decay with distance. The latter, however, are not computationally convenient for the calculation of the two-electron integrals. The radial part is normally expressed as a combination of Gaussian functions. In general, the results of the Hartree–Fock approach to the electronic molecular wave function and properties will depend on the ex-tension of the basis set used for the expansion of the molecular orbitals. The fit of the predicted properties to experiments rapidly improves enlarging the basis set and the calculated values converge to the so-called Hartree–Fock limit. This, however, does not match the experiment within the limit of chemical accuracy. Expressing the multi-electron wave function as a single Slater determinant of the occupied molecular orbitals is not a correct approximation. The multi-electron wave function should be expressed as a combination of the Slater determinants corresponding to configurations of the ground state and of the singly, doubly and multiply excited states, as shown in **Figure 2b** and **2c** (configuration interaction). The difference between the Hartree–Fock and the experimental energy is called correlation energy

$$E_{corr} = E_{HF} - E_{exp} \quad [11]$$

A full configuration interaction for the obtainment of the correlation energy of a molecule is in most cases an unfeasible computational task, and an estimate of the major contribution to the correlation energy can be afforded by perturbation methods where the Hartree–Fock Hamiltonian is assumed as the unperturbed part of the Hamiltonian. However, even at a perturbation or multiperturbation level the calculation of the correlation energy is computationally too demanding for medium- or large-size molecules.

Density-Functional Theory and the Kohn–Sham Equations

The computational demand for post Hartree–Fock (correlated) calculations does not allow the application of accurate quantum mechanical methods to systems of practical interest, such as biological molecules, macromolecules, solutions, and condensed phases. A way around this problem has been furnished by the DFT based on the Hohemberg–Kohn theorems. The first theorem states that, for a multi-electron system in the ground state, the expectation value of any physical observable O is a unique functional of the electron density ρ, unique in the sense that it will not depend on the external potential (which, in the case of interest, is the nuclear potential)

$$\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = F_{HK}[\rho] \quad [12]$$

The electron density thus determines the kinetic and potential energy of the system and the total energy. One can write the total electronic energy E_v as the sum of the kinetic T, electron–electron E_{ee} and

the electron–nuclei E_{en} energy

$$E_v[\rho] = T[\rho] + E_{ee}[\rho] + E_{en}[\rho] = \int \rho(r)V(r)\,dr + F_{HK}[\rho] \quad [13]$$

where

$$F_{HK}[\rho] = T[\rho] + E_{ee}[\rho] \quad [14]$$

and $V(r)$ is the nuclear potential in the present case. The electron–electron repulsion term E_{ee} now includes the classical term and all the exchange and correlation terms. If the functional F were known, by application of the variational principle, guaranteed by the second Hohemberg–Kohn theorem, one could determine the electron density of the system. Unfortunately, this knowledge is not available. However, the Hohemberg–Kohn theorems are the foundation of a theory allowing the search of solutions of increasing accuracy.

From a practical point of view, one can resort to the Kohn–Sham scheme where it is assumed that there is an auxiliary system of noninteracting electrons subject to a local potential V_s with Hamiltonian

$$H_s = T_s + V_s \quad [15]$$

and energy

$$E_s = T_s + \int dr\, V_s\rho_s(r) \quad [16]$$

with the property of having the same electron density ρ_0 as the real system. In the Kohn–Sham approach, one considers a one-particle Schrödinger equation

$$H_{KS}\psi_i = \left(-\tfrac{1}{2}\nabla^2 + V_s\right)\psi_i = \varepsilon_i\psi_i \quad [17]$$

and the Kohn–Sham occupied orbitals give the exact electron density. The total energy of the system of interacting particles subjected to an external potential V is given by

$$E_v[\rho_0] = T_s[\rho_0] + \int dr\, V(r)\rho_0(r) - \frac{1}{2}\int dr\, dr' \frac{\rho_0(r)\rho_0(r')}{|r-r'|} + E_{xc}[\rho_0] \quad [18]$$

and the crucial quantity is the exchange correlation functional E_{xc}. As a useful approximation (LDA, the local-density approximation), the exchange correlation functional can be calculated using the homogeneous electron gas expression $\varepsilon_{xc}(n)$ substituting the constant density n with the local density $\rho(r)$ of the inhomogeneous system

$$E_{xc} = \int dr\, \rho(r)\varepsilon_{xc}[\rho(r)] \quad [19]$$

Several improvements are possible to the LDA approximation. Since the exchange correlation functional depends on the electron density, an iterative procedure can be used. A starting electron density is obtained as a superposition of the atomic electron densities for a given nuclear configuration. The main advantage of the DFT approach in quantum mechanical calculations of molecules is the drastic reduction of computational time while all exchange and correlation terms are included. This allows one to perform calculations on more complicated systems such as large molecules or macromolecules or biological systems, liquids, solids, and solutions. One of the major successes of the DFT in molecular systems has been achieved by extending the approach to molecular dynamics simulation with an on-the-fly calculation of intermolecular forces. This has allowed a reasonably realistic representation of complex systems including chemical reactions. Attempts to extend the application of the DFT to excited states are ongoing.

The Valence Bond Theory

The molecular orbitals method discussed above describes the molecule as an arrangement of nuclei in some equilibrium configuration and generating a field, in which the interacting electrons move. Therefore, in principle the structure of the atoms forming the molecule is not relevant, even though for practical purposes the molecular orbitals are expressed as a linear combination of atomic orbitals. The approach of the valence bond theory is quite different and closer to the chemical point of view that the molecule is made by interacting atoms. The valence bond theory accounts for the formation of a chemical bond as a result of the interaction of two atoms when the interatomic distance becomes reasonably short. Actually, the first quantum mechanical description of the chemical bond in the hydrogen molecule was reported by Heitler and London in 1927 according to the valence bond approach. Considering two hydrogen atoms H_A and H_B in their ground state with the unpaired electron in the respective 1*s* orbitals (φ_A and φ_B), at large separation the total wave function of the couple will be

$$\Psi_I = \varphi_A(1)\varphi_B(2) \quad [20]$$

When the two atoms are close enough such that the atomic electron clouds overlap, the above wave

function is not appropriate since the two electrons are identical particles and Ψ_I distinguishes between the two by assigning electron 1 to atom A and electron 2 to atom B. The wave function

$$\Psi_{II} = \varphi_A(2)\varphi_B(1) \qquad [21]$$

gives an equivalent description of the system and the idea of the valence bond theory is that the wave function of the system (the hydrogen molecule in the present case) is a linear combination of the two,

$$\Psi = a\Psi_I + b\Psi_{II} \qquad [22]$$

The coefficients can in general be determined variationally but, in the case of the hydrogen molecule, the symmetry requires that $a = \pm b$ and their values can be found by the normalization condition. The Ψ_+ state has energy lower than the separated atoms and corresponds to the ground state. The correct antisymmetric character of the wave function can be ensured by multiplying the orbital part by an antisymmetric spin term. Alternatively, introducing the spin directly, the Ψ_I and Ψ_{II} components can be expressed in the form of Slater determinants.

Along the same lines, the valence bond theory can be applied to more complicated molecules by considering the formation of the individual bonds by spin-coupling unpaired electrons of the atoms that form the molecule. Each possible coupling scheme of the electrons corresponds to a Slater determinant and the total wave function is a linear combination of the determinants corresponding to independent coupling schemes.

At the lowest level of approximation, the valence bond treatment of the hydrogen molecule gives better results than the molecular orbital method with the molecular orbitals expanded in the minimal basis set (the 1*s* orbitals of the two atoms in the case of the H_2 molecule). However, it can be seen that when the configuration interaction is applied to the molecular orbital method and the ionic formulas with the two electrons in turn on the same atom are included in the valence bond theory, the two methods converge to the same result.

The Molecular Orbital Description of Biatomic Molecules

It is useful to consider the results of the molecular orbital methods for the homonuclear diatomic molecules of the second row of the periodic table (Li_2, Be_2, B_2, C_2, N_2, O_2, F_2, and Ne_2), since the same concepts can be easily extended to more general cases. In the simplest approximation, the molecular orbitals are expanded in the minimal basis set, that is, in the 1*s*, 2*s*, $2p_x$, $2p_y$, and $2p_z$ orbitals of the two atoms. In principle, the molecular orbitals will be linear combinations of all the atomic orbitals of the basis set. In practice, the different overlap of the atomic orbitals and symmetry restrictions make it possible to approximately represent the molecular orbitals as linear combinations of the corresponding atomic orbitals two-by-two, as shown in **Figure 3**.

This gives an approximate ordering of the energy of the molecular orbitals. The structure of the molecule is then obtained by allocating the electrons in the molecular orbitals according to the Pauli exclusion principle.

Biatomic (and linear) molecules have axial symmetry and, therefore, the Hamiltonian commutes with the *z*-component L_z of the angular momentum (*z* is intended as the molecular axis). The molecular orbitals will then be eigenfunctions of the L_z operator as well and they can be classified according to the M_z quantum number as σ ($M_z = 0$), π ($M_z = 1$), δ ($M_z = 2$) orbitals, etc. In addition, the molecular orbitals can be also classified in terms of the atomic orbitals that are formed when the atoms separate at infinite interatomic distance (separated-atoms approximation). The molecular orbitals in **Figure 3** are labeled according to these rules. In addition, the *g* and *u* labels indicate the symmetric and antisymmetric character of the orbital with respect to the inversion center, respectively. Finally, the asterisk labels the molecular orbitals that are antisymmetric for a reflection in a plane perpendicular to the molecular axis. As can be seen from **Figure 3**, the latter are antibonding orbitals, that is, their energy is higher than the energy of the atomic orbitals from which they are formed as linear combinations.

As an example, the ground-state configurations of the nitrogen and oxygen molecules are

$$N_2: \quad (\sigma_g 1s)^2(\sigma_u^* 1s)^2(\sigma_g 2s)^2(\sigma_u^* 2s)^2(\sigma_g 2p_z)^2(\pi_u 2p_x)^2 (\pi_u 2p_y)^2$$

$$O_2: \quad (\sigma_g 1s)^2(\sigma_u^* 1s)^2(\sigma_g 2s)^2(\sigma_u^* 2s)^2(\sigma_g 2p_z)^2(\pi_u 2p_x)^2 (\pi_u 2p_y)^2(\pi_g^* 2p_y)(\pi_g^* 2p_y)$$

and the ground-state wave function can be consequently written as a Slater determinant. In particular, it can be seen that the molecular orbital approach simply explains the paramagnetic character of the oxygen molecule arising from the presence of two unpaired electrons in the degenerate π_g orbitals.

The ordering of the molecular orbitals can change significantly along the second row of the periodic table. A qualitative idea of these changes can be obtained by considering not only the above separated-atoms

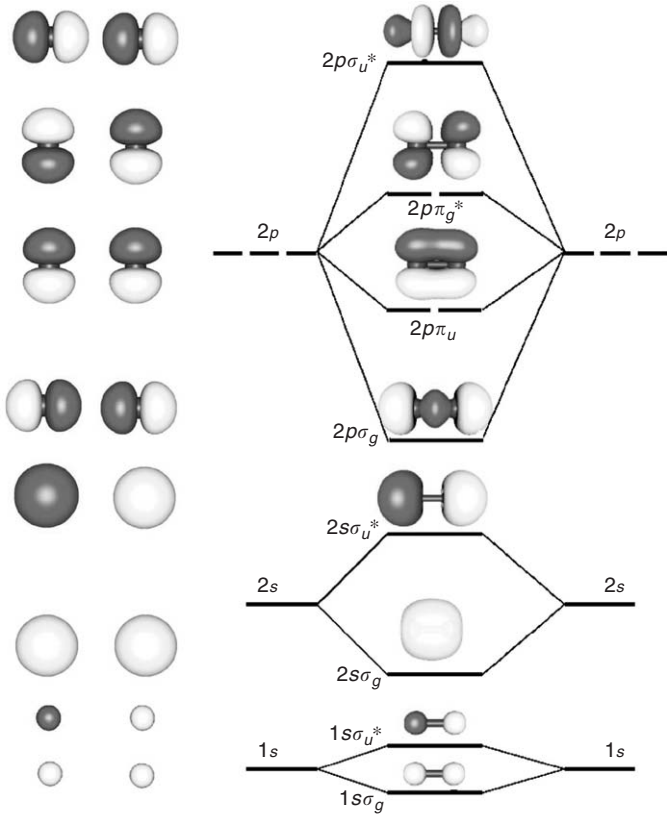


Figure 3 Schematic representation of the molecular orbitals of second row diatomic states obtained as a linear combination of the relevant atomic orbitals.

limit but also the united-atom approximation correlating the molecular orbitals with the atomic orbitals resulting at zero interatomic separation. A correlation diagram can then be obtained (shown in **Figure 4**) between the united- and separated-atoms limits that gives an idea of the changes in the molecular orbitals energy as a function of the interatomic separation. The diagram is constructed by considering that the orbital symmetry is conserved at all interatomic distances.

As discussed earlier, an accurate fit of the molecular energy and structure cannot be obtained by representing the molecular wave function as a single Slater determinant and with the expansion in the minimal basis set. For accurate prediction of molecular properties, it is necessary to use an extended basis set and to take account of the correlation energy as discussed earlier. The same qualitative description can be given for a heteronuclear diatomic molecule. In this case, however, the distinction into *g* and *u* type molecular orbitals is obviously lost and, in addition, the coefficients of the linear combinations of atomic orbitals are different for the two atoms. This will result in the presence of a nonzero electric dipole moment in heteronuclear diatomics.

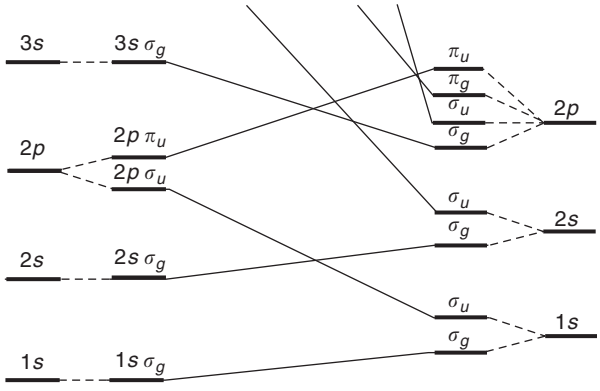


Figure 4 Molecular orbital theory. Correlation of molecular orbital energy of first and second row diatomic states between the united atom (left) and separated atom limits (right).

Quantum Mechanics and Symmetry of Molecules

The symmetry group of a molecule (point group) can be defined as the collection of symmetry operations that commute with the Hamiltonian. These are unitary transformations of two types: proper rotations (rotations of an angle $(2\pi/n)k$ around a C_n symmetry

Table 1 The irreducible representation of the symmetry group of the H_2O molecule (*yz* is the plane of the molecule)

C_{2v}	E	C_2^z	σ_{yz}	σ_{xz}
A_1	1	1	1	1
A_2	1	1	−1	−1
B_1	1	−1	1	−1
B_2	1	−1	−1	1

axis, with $k = 0, 1, \ldots, n-1$) and improper rotations (rotations around an axis followed by a reflection through a plane perpendicular to the axis). The latter include, as particular cases, the inversion through a center and the reflection through a plane. There are compatibility relations among the symmetry operations of a point group and it turns out that most molecules belong to one of the 32 crystallographic point groups. Since all the symmetry operations of a molecule commute with the Hamiltonian operator, the molecular wave functions must also be eigenfunctions of the symmetry operators. As a consequence, the molecular orbitals and the total electronic wave function form the basis for the irreducible representations of the molecular symmetry group. As an example, the water molecule belongs to the C_{2v} symmetry group, a group of order 4, including as operations the identity E, a rotation C_2, and two reflection operations σ_1 and σ_2. The molecular orbitals will necessarily behave as one of the irreducible representations of the C_{2v} group that are labeled as A_1, B_1, A_2, and B_2 (see **Table 1** and **Figure 5**).

Expanding the molecular orbitals as a linear combination of the minimal basis set (including the 1*s* orbitals of the hydrogens and the 1*s*, 2*s*, $2p_x$, $2p_y$, and $2p_z$ orbitals of the oxygen atom), one obtains seven molecular orbitals. The characters of the completely reduced representation of the molecular orbitals are the same as the characters of the reducible representation of the atomic orbitals. By analyzing the reducible representation of the atomic orbitals, one obtains the symmetry of the molecular orbitals; in the present example,

$$\Gamma = 4A_1 + 2B_1 + B_2$$

Using the appropriate projection operators, the atomic orbitals can be linearly combined into symmetry-adapted functions according to the above reduced representation. With this choice, the Fock matrix entering in the Roothan equations is block-diagonalized into three blocks of dimensions 4, 2, and 1. The resulting approximate form of the molecular orbitals of water is shown in **Figure 5**. The electronic ground state of the water molecule

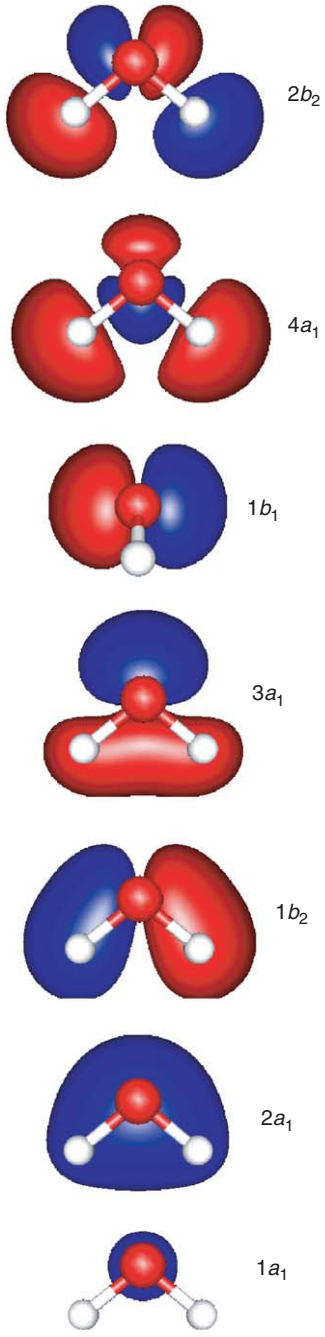


Figure 5 Schematic representation of the molecular orbitals of the H_2O molecules classified by their symmetry character and in order of increasing energy.

corresponds to the configuration

$$(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2(1b_1)^2$$

and its symmetry is given by the direct product of the representations of the occupied molecular orbitals. In the present case of a closed-shell molecule (with

double occupancy of the orbitals), the total wave function turns out to be totally symmetric with spin multiplicity equal to 1 (a singlet state $^1A_1$).

The Nuclear Schrödinger Equation

The solution of the Schrödinger equation for the nuclear motion (with the Hamiltonian defined in [3]) can be obtained by first decoupling the translational, rotational, and vibrational degrees of freedom. For this purpose, one can define a laboratory X, Y, Z and a molecule-fixed x, y, z Cartesian reference system. The position of each nucleus α, with mass m_α, is then specified by giving the three Cartesian coordinates of the origin of the molecule-fixed system, three Eulerian angles describing the orientation of the molecular axes relative to the laboratory axes and, for an n-atomic molecule, the $3n$ coordinates in the molecular reference frame x_α, y_α, z_α. Among the latter, there are two vectorial relations specifying the position of the origin of the molecule-fixed system, placed at the center of mass,

$$\sum_\alpha m_\alpha \boldsymbol{r}_\alpha = 0 \qquad [23]$$

and the orientation of the molecule-fixed system

$$\sum_\alpha m_\alpha \boldsymbol{a}_\alpha x^* \boldsymbol{r}_\alpha = 0 \qquad [24]$$

where $\boldsymbol{a}_\alpha$ denotes the equilibrium position of atom α. With this choice of the reference system, it can easily be shown that the translational degrees of freedom are completely decoupled from the rotational and vibrational degrees of freedom. Therefore, the translational motion of the molecule can be conveniently treated with the particle-in-a-box model. However, there is still a Coriolis coupling between the rotational and vibrational motions. Considering that the vibrational velocities are much higher than the rotational velocities, the Coriolis coupling can be considered to be small for vibrational motion but can be substantial for the rotational motion. In condensed phases, where the molecules cannot rotate freely, the rotational–vibrational coupling is eliminated.

The Vibrations of Molecules

The potential governing the vibrational motion of a molecule can in principle be obtained from an *ab initio* calculation of the electronic energy. In practice, the potential is expanded around the equilibrium configuration and in the harmonic approximation it is in the quadratic form

$$V = \frac{1}{2}\sum_{ij} f_{ij}\rho_i\rho_j \qquad [25]$$

where ρ is a $3n$-dimensional vector of the Cartesian displacements, from the equilibrium. To fully decouple the problem, it is appropriate to transform to a set of $3n$ normal coordinates Q_k, linear combinations of the Cartesian displacements that simultaneously diagonalize the kinetic and potential energy,

$$H = T + V = \frac{1}{2}\sum_k P_k^2 + \frac{1}{2}\sum_k \lambda_k Q_k^2 \qquad [26]$$

where P_k is the conjugate momentum and $\omega_k = \sqrt{\lambda_k}$ is the classical normal frequency of vibration. Because of the relations [23] and [24], six (five in the case of linear molecules) of the normal coordinates describe overall translations and rotations of the molecule and the corresponding frequencies are zero. Neglecting the zero-frequency normal coordinates and writing the vibrational wave function as a product of functions of the individual remaining normal coordinates,

$$\psi_v(Q_1Q_2\ldots) = \psi(Q_1)\psi(Q_2)\cdots\psi(Q_{3n-6}) \qquad [27]$$

The vibrational Schrödinger equation

$$\left(\sum_k -\frac{\hbar^2}{2}\nabla_k^2 + \sum_k \lambda_k Q_k^2\right)\psi_v(Q_1Q_2\ldots) = E_v\psi_v(Q_1Q_2\ldots) \qquad [28]$$

can be separated into $3n-6$ (or $3n-5$) harmonic oscillator equations,

$$-\frac{\hbar^2}{2}\nabla_k^2\psi(Q_k) + \lambda_k Q_k^2\psi(Q_k) = E_k\psi(Q_k) \qquad [29]$$

Therefore, the vibrational eigenfunction is the product of $3n-6$ harmonic oscillators functions

$$\psi_v = \prod_k \psi_{n_k}(Q_k) = C\exp\left(-\frac{1}{2}\sum_k \gamma_k Q_k^2\right)\prod_k H_{n_k}(\gamma_k^{1/2}Q_k) \qquad [30]$$

where C is a normalizing factor, $\gamma_k = \omega_k/\hbar$, and H_{n_k} is the Hermite polynomial of degree n_k. A vibrational state is specified by $3n-6$ quantum numbers $(n_1n_2\ldots n_k\ldots n_{3n-6})$. The vibrational energy

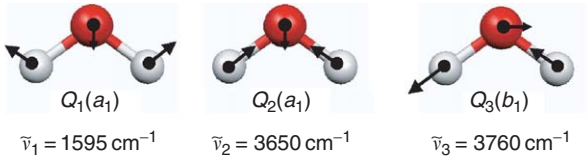


Figure 6 Normal modes of vibration of the water molecule.

is given by

$$E_v = \sum_k \hbar\omega_k (n_k + \tfrac{1}{2}) \qquad [31]$$

In **Figure 6** the vibrational normal modes (or the normal coordinates) of the water molecule are represented.

The normal coordinates form a basis for a completely reduced representation of the molecular point group. The characters of this reduced representation can be obtained from the equivalent representation of the Cartesian displacements since the normal and Cartesian coordinates are related by a linear transformation. By eliminating the zero-frequency normal coordinates, one obtains for the water molecule that the normal coordinates (or the wave functions with quantum numbers (1 0 0), (0 1 0), and (0 0 1)) classify in the irreducible representations of the symmetry group

$$\Gamma = 2A_1 + B_1$$

as seen from **Figure 6**. Qualitative information on the form of the normal coordinates can be obtained by using symmetry arguments and internal coordinates (stretching and bending of bonds and torsional coordinates) instead of Cartesian coordinates.

Symmetry arguments can also be used to obtain the selection rules for transitions between two vibrational eigenstates i and f. In the infrared absorption spectroscopy, the relevant operator is the electric dipole moment P and the transition probability is proportional to

$$|\langle \psi_f | P | \psi_i \rangle|^2 \qquad [32]$$

Symmetry considerations require that the transition probability is nonzero when the direct product of the irreducible representations of the terms appearing in the matrix element $\Gamma_f \otimes \Gamma_P \otimes \Gamma_i$ contains the totally symmetric representation. The spectrum is dominated by the fundamental transitions from the ground state (totally symmetric) to states where a single quantum number is equal to 1. These have nonzero transition probability if the corresponding normal coordinate belongs to the same irreducible representation as one of the components of the electric moment. Nonlinearities in the expansion of the electric moment and anharmonic terms of the potential make multiple transitions weakly active in the infrared spectrum. Vibrational transitions can also be observed directly in the Raman spectrum. In this case, the relevant operator is the molecular polarizability and the selection rules are consequently different.

Rotation of Molecules

The rotation of a diatomic molecule can, to a first approximation, be described by the rigid rotator model and the rotational Hamiltonian can be written as

$$H_r = \frac{1}{2I} L^2 \qquad [33]$$

where I is the inertia moment and L is the angular momentum operator. The rotational eigenfunctions are, therefore, the spherical harmonics and the rotational energy levels are given by

$$E_J = B_0 J(J+1) \qquad [34]$$

where $B_0 = \hbar^2/2I$ is the rotational constant and J is the rotational quantum number ($J = 0, 1, 2\ldots$). The rotational levels and eigenfunctions have $2J+1$ degeneracy in consideration of the possible orientation of the angular momentum. The rotational constant refers to the vibrational (and electronic) ground state. The inertia moment and the rotational constant change in vibrational excited states. The rotational constant also changes, which in turn changes the rotational level because of centrifugal distortion effects, and in each vibrational state n the energy levels of the rotating molecule should be written as

$$E_r = B_n J(J+1) - D_n J^2 (J+1)^2 \qquad [35]$$

where the centrifugal distortion constant D_n is much smaller than the rotational constant and can be related to the vibrational force constants. As a consequence of the second term in [35] in the pure rotational spectrum (selection rule $\Delta J = \pm 1$), the lines are not exactly equidistant.

The rotational problem of polyatomic molecules is considerably more complicated. The structure of the spectrum is determined by the relative values of the principal inertia moments. In the case of symmetric-top molecules, for which

$$I_{xx} = I_{yy} = I_\perp, \qquad I_{zz} = I_{||} \qquad [36]$$

the rotational Hamiltonian has the form

$$H_r = \frac{1}{2I_\perp} L^2 + \left(\frac{1}{2I_{||}} - \frac{1}{2I_\perp} \right) L_z^2 \qquad [37]$$

and the eigenvalues are

$$E_{JK} = \frac{\hbar^2}{2I_\perp} J(J+1) + \hbar^2 \left(\frac{1}{2I_{||}} - \frac{1}{2I_\perp} \right) K^2, \quad K = J, J-1, \ldots \qquad [38]$$

and the selection rule is

$$\Delta J = \pm 1, \qquad \Delta K = 0 \qquad [39]$$

See also: Density-Functional Theory ; Electronic Structure (Theory): Molecules; Hartree and Hartree–Fock Methods in Electronic Structure; Quantum Mechanics: Atoms; Quantum Mechanics: Foundations; Quantum Mechanics: Methods.

PACS: 31.15.Ar; 31.15.Ew; 31.15.Ne; 33.15.−e; 33.20.Ea; 33.20.Fb

Further Reading

Califano S (1976) *Vibrational States*. New York: Wiley Interscience.

Herzberg G (1945) *Molecular Spectra and Molecular Structure Volume 2: Infrared and Raman Spectra of Polyatomic Molecules*. New York: Van Nostrand-Reinhold.

Herzberg G (1957) *Molecular Spectra and Molecular Structure Volume 1: Spectra of Diatomic Molecules*. New York: Van Nostrand-Reinhold.

Herzberg G (1966) *Molecular Spectra and Molecular Structure Volume 3: Electronic Spectra and Electronic Structure of Polyatomic Molecules*. New York: Van Nostrand-Reinhold.

Levine IS (1991) *Quantum Chemistry*. NJ: Prentice-Hall.

McWeeny R (1973) *Quantum Chemistry*. Oxford: Pergamon.

Parr RG and Yang W (1989) *Density-Functional Theory of Atoms and Molecules*. New York: Oxford Science Publications.

Parrinello M (1997) From silicon to RNA: the coming age of *ab initio* molecular dynamics. *Solid State Communications* 102: 107–120.

Parrinello M (2000) Simulating complex systems without adjustable parameters. *Computing in Science and Engineering* 2: 22–27.

Pople JH and Beveridge DL (1970) *Approximate Molecular Orbital Theory*. New York: McGraw-Hill.

Sironi M, Raimondi M, Martinazzo R, Gianturco FA, and Cooper DL (2002) Theoretical and computational chemistry. In: Cooper DL (ed.) *Valence Bond Theory*, vol. 10, pp. 261–277. Amsterdam: Elsevier.

Szabo A and Ostlund NS (1982) *Modern Quantum Chemistry*, 1st edn. New York: McGraw-Hill.

Quantum Mechanics: Nuclei

R A Broglia, University of Milan, Milan, Italy

Introduction

The atom is the smallest unit into which matter that has the characteristic properties of a chemical element can be divided. Most of the atom is empty space. The rest consists of a cloud of negatively charged electrons whirling around a small, very dense, positively charged nucleus made of protons and neutrons which contains essentially all the mass of the atom. The electric forces bind the electrons to the nucleus giving rise to a miniature planetary-like system. The strong and Coulomb forces acting among nucleons lead to a self-bound system. Inside it, nucleons move essentially independent of each other. Bouncing in their trajectories elastically off the nuclear surface, they describe closed orbits which resemble, at a length scale five orders of magnitude smaller, those of the electrons in the atom.

The Atomic Nucleus

Nucleons, like electrons are fermions, that is, they have a half-integer spin $(1/2, 3/2, \ldots)$, and, according to quantum mechanics, obey the Pauli exclusion principle which allows only one particle to occupy each quantal state. Consequently, for the ground state of the system, the available orbitals of the mean field are filled from the bottom upwards to the Fermi energy. Under normal conditions, atomic nuclei are in their ground state, that is, at zero temperature. This is because nuclei on earth, leaving aside those that arrive in the form of cosmic rays, are isolated. In fact, for two nuclei to interact, they need to have large kinetic energies, of the order of tens of MeV ($1\,\text{MeV} = 10^6\,\text{eV}$), so as to be able to overcome the Coulomb repulsion in a collision and reach within the range of the nuclear attraction. Energies of such magnitude contrast with the energies available at room temperature ($\approx 25 \times 10^{-3}$ eV). This is the reason why one needs large machines, so called atomic smashers, to make two atomic nuclei interact. In these machines, an atom containing a heavy atomic nucleus is stripped of most of its electrons becoming a heavy ion, which is subsequently accelerated and collimated. The resulting beam, aimed at a target of other heavy atomic nuclei, eventually leads to a heavy ion reaction (**Figure 1**).

In the event where the two nuclei fuse, the energy and angular momentum of relative motion become mostly excitation energy and angular momentum of the composite system. Typical values of these quantities are set (**Figure 2**), in relation to temperatures and rotational frequencies observed in other

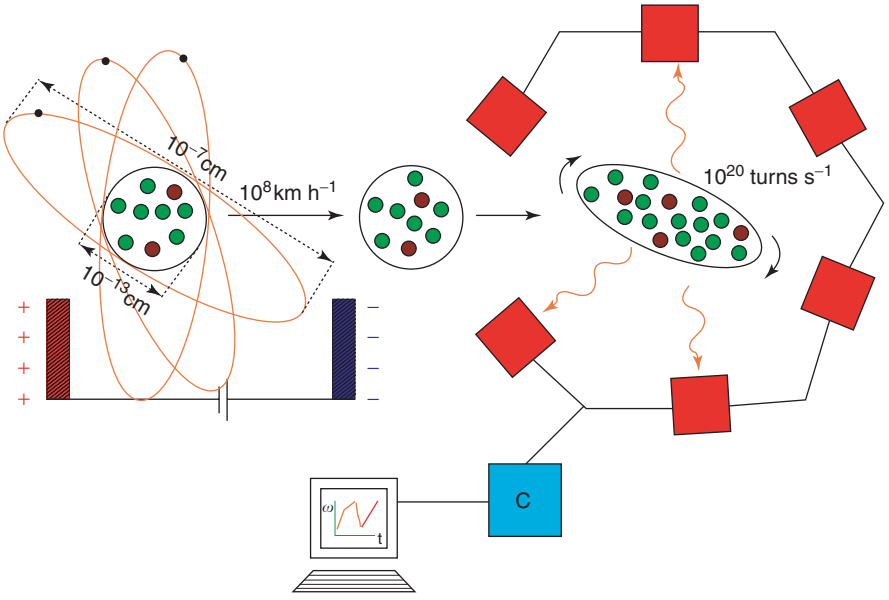


Figure 1 Schematic representation of a heavy ion accelerator used to induce nuclear reactions. A neutral atom (left part of the figure), formed of a nucleus made of protons (red dots) and neutrons (green dots), is stripped off the electrons (small black dots) moving around it. The system becomes positively charged and the resulting heavy ion is accelerated to velocities $\sim$100 millions of kilometers per hour in an electric field and smashed against the atomic nuclei of a target (middle). Out of the violent encounter between the two atomic nuclei, a number of processes can take place: Coulomb excitation, particle transfer, etc. In some cases the two nuclei fuse, leading to a highly excited, rapidly rotating compound nucleus (right) which, in the process of cooling down, emits nucleons, alpha particles and γ-rays (wavy arrows). These γ-rays (photons) can be detected by use of a so-called 4π-array, that is a set of detectors (red squares) covering a consistent fraction of the solid angle. Once the signal is collected, it is analyzed making use of specifically developed computer programs (software labeled C) and the results are eventually displayed on a screen.

physical systems. The thermalization of such a system depends, naturally, on the specific heat of the system, a quantity which is directly related to the density of levels around the Fermi energy.

Single-Particle Motion

To study the single-particle motion in nuclei, the best probe one can use are transfer reactions. For example, aiming a beam of deuterons (the isotope $^{2}_{1}\mathrm{H}_{1}(\equiv d)$ of hydrogen containing one proton and one neutron) on a target nucleus, one can learn about the properties of the single-particle neutron levels lying above the Fermi energy, by studying the spectrum and angular distribution of the emerging protons, resulting from the stripping of a neutron from the projectile (**Figure 3**). The main outcome of these studies, in particular of those carried out on the closed-shell nucleus ^{208}Pb, shows that the mean-field theory is able, in most cases, to correctly predict the sequence of single-particle levels.

Collective Motion

Nuclear vibrations are excited by bombarding nuclei with high-energy photons, nucleons, electrons, etc. The vibrations are detected by observing how photons are absorbed by the nucleus, or how protons are scattered inelastically from the nuclear surface, etc. These experiments reveal that the nucleus displays both elastic and plastic behavior. In fact, the so-called giant dipole resonance corresponding to a back- and -forth sloshing of protons against neutrons, excited in photoabsorption experiments has an energy centroid which scales with the inverse of the radius of the nucleus (a dependence observed for all nuclear resonances), that is, with the nuclear momentum, a behavior typical of elastic vibrations. Once excited into a giant dipole resonance, the nucleus vibrates at an extremely high frequency, of the order of 5×10^{21} Hz (which corresponds to a vibrational energy of $\sim$20 MeV). To make more vivid the qualification of "extremely" used above, it is useful to remember that the nominal range of human hearing extends from about 20×10^{4} to 2×10^{4} Hz, while that of human vision ranges from 4×10^{14} to 8×10^{14} Hz. In order to excite giant resonances with suitable cross sections, use has to be made of fields that change quickly with time, with frequencies of the order of those associated with single-particle motion ($1/\omega \equiv 10^{23}$ Hz).

The lowest excited states of even–even nuclei display, with very few exceptions, quadrupole and

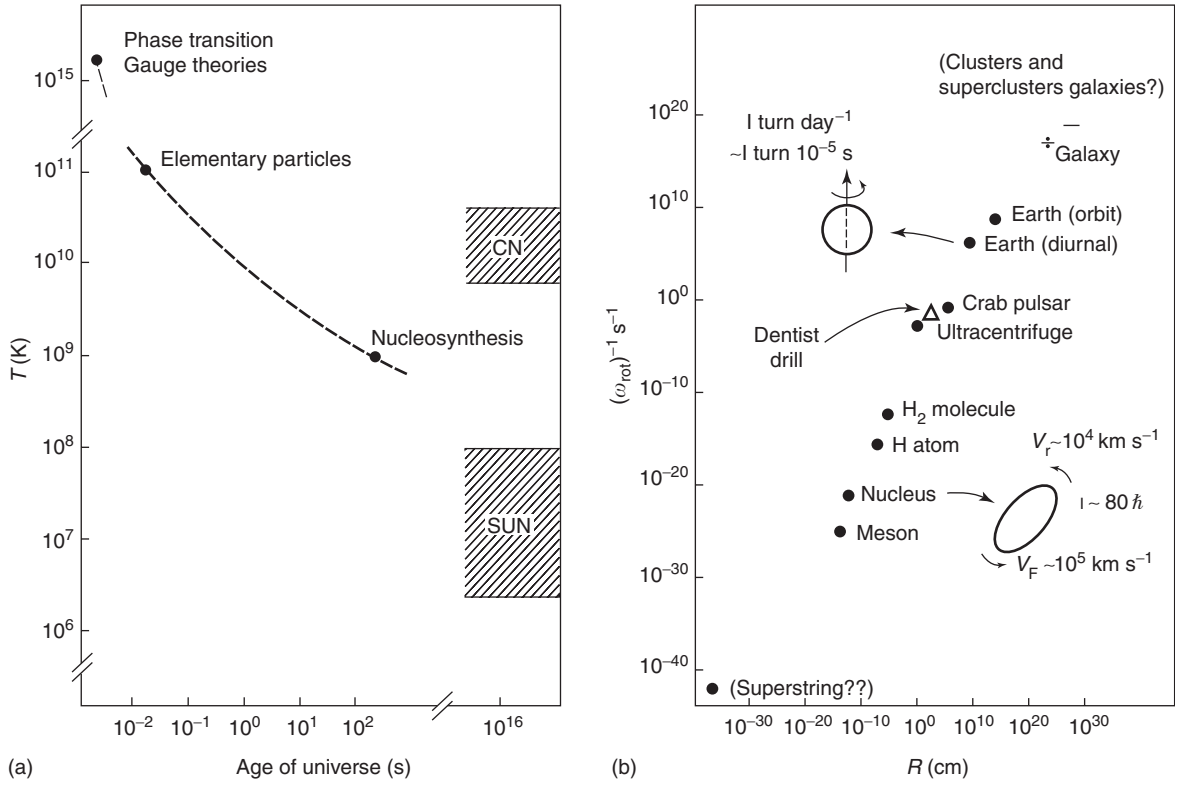


Figure 2 Temperatures and rotational frequencies typical of reactions where two heavy ions fuse and eventually, after a time of the order of a few times 10^{-21}s equilibrate all their degrees of freedom, forming a single compound nucleus. (a) The scale of temperature has been set in connection with the temperature ascribed to the Universe at different times after the Big Bang through the first 3 min to our times. (b) The inverse of the rotational period (in seconds) of a variety of rotating objects observed (or thought to exist, like superstrings which are the mathematical connotation of an elementary particle) in the Universe from the largest to the smallest, are quoted as a function of their radius. Both the inverse frequencies and the radii span 60 orders of magnitude, the relation between the two quantities being almost linear.

octupole character, that is, they correspond to surface collective vibrations of multipolarity 2 and 3. The associated energies, 1–3 MeV, show a dependence with mass number inversely proportional to the area of the nuclear surface, typical of the vibrations of a liquid drop. They reflect the plastic properties of atomic nuclei. In fact, away from closed shells, the energy of the lowest $2+$ state can become particularly low in energy (**Figure 4**). In these nuclei (^{152}Sm,^{154}Sm, or e.g.,^{238}U), the $2+$ state is the lowest member of the so-called ground-state rotational band of a spheroidal nucleus rotating around an axis perpendicular to the symmetry axis (**Figure 5**). These states are excited with large cross sections by Coulomb and by nuclear fields which change slowly in time ($1/\omega \equiv 10^{20}$ Hz), as compared with typical frequencies associated with single-particle motion.

Further evidence of the plastic behavior of the nucleus is provided by nucleon fission. This process can be viewed as the division, similar to that of a cell, of a liquid drop into two smaller droplets as a result of deformation (**Figure 6**). All these phenomena are essentially controlled by the nuclear surface.

Pairing

The theory describing the coupling of the nucleons to vibrations and to static deformations of the nuclear surface provides a unified description of liquid drop and of independent-particle behavior.

The rotational motion is one of the simplest and, at the same time, most profound examples of periodic motion (think only of the earth's rotation around itself and around the Sun, and of the succession between day and night and between summer and winter). A detailed study of the rotational frequency and of the variations of the rotational period as a function of time can provide important information on the physical properties of the system under study and on the forces acting inside it.

Strongly rotating nuclei (**Figure 2**) can be produced in the laboratory either via a Coulomb excitation (**Figure 5**) or in fusion processes (**Figure 1**). Because

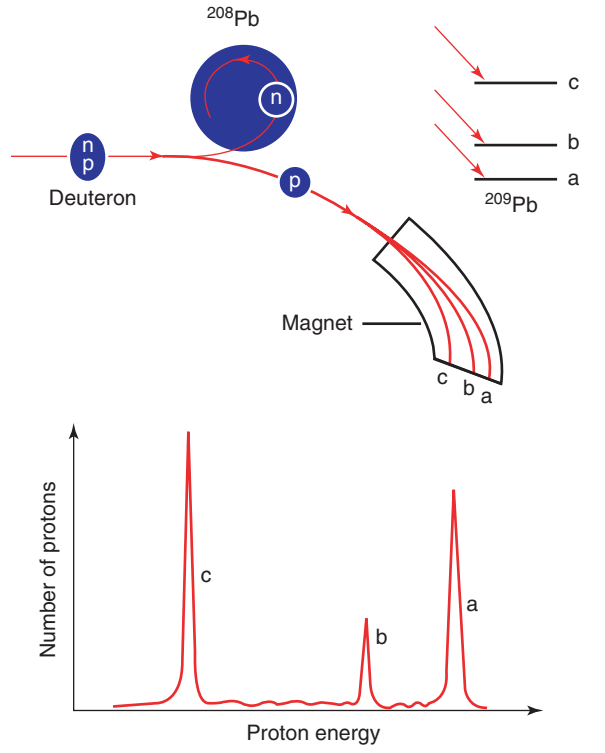


Figure 3 Nuclear reaction in which a neutron is transferred from the projectile (deuteron) to the nucleus ^{208}Pb. The energy of the outgoing proton reflects both the Q-value of the reaction and the energy of the final state of the system.

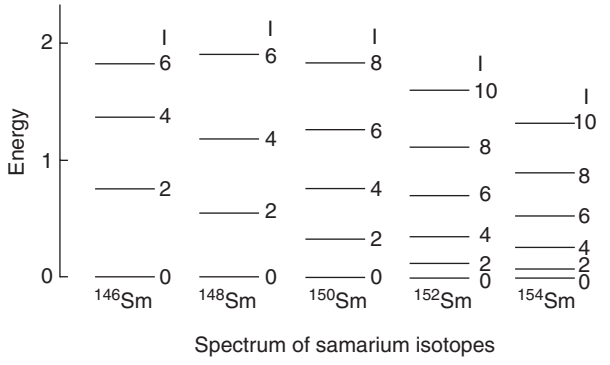


Figure 4 Sequence of positive parity states of the isotopes of ^{62}Sm. The energy scale is in MeV.

the system is charged, the rotating nucleus emits "light," that is gamma-rays which provide detailed information on the properties of the system. A surprising result emerging from such measurements is that at low rotational frequencies, the moment of inertia of the system is considerably smaller, about half of the expected (rigid) moment of inertia, while at higher rotational frequencies the moment of inertia coincides more or less with the rigid moment of inertia of the system. The explanation of this behavior is connected with the fact that the nucleus, which in connection with the present discussion can be viewed as a spheroidal container, is filled at low rotational frequencies with a nonviscous (superfluid) liquid. Spinning the container, only the matter at the poles, that is, the matter directly pushed by the walls (surface) of the container, is set up into motion, the central core remaining at rest (**Figure 7**). On the other hand, at high rotational frequencies the nucleus displays a phase transition, into a system of nucleons moving independently of each other in well-defined orbits (normal system). Each of these orbits are solidly anchored to the mean field (surface) of the nucleus. Consequently, each nucleon inside the nucleus reacts to a change in the rotational frequency with its full mass. The collective motion of the system thus coincides with that of a rigid body of the same mass, dimension, and deformation of the nucleus under study.

These phenomena are closely connected with the variety of effects belonging to the field of low-temperature physics, which go under the name of "superconductivity" when they occur in an electric charged system such as electrons in a metal, and "superfluidity" when they occur in a neutral system such as an insulating liquid. The behavior of a superconducting metal is qualitatively different from that of a normal metal, in that it conducts electricity with zero resistance (hence the name). Similarly, a liquid that is superfluid can flow through tiny capillaries without apparent friction, and can even climb up over the rim of a vessel containing it and thereby gradually empty the vessel.

Bose–Einstein Condensation

To sketch the general picture, let the focus be on the common isotope of helium (^{4}He). Such an atom has zero total spin, and therefore, according to quantum mechanics should obey Bose–Einstein statistics. Consequently, when a large number of such particles are placed in a restricted volume and are cooled down below a certain temperature (which for helium turns out to be ~ 3 K), a phenomenon called "Bose–Einstein condensation" (BEC) takes place: namely, a finite fraction of all the atoms begins to occupy a single quantum-mechanical state, the fraction increasing as the temperature decreases. Thus, in the same way that fermions are extremely individualistic, each of them occupying a different (quantal) state, bosons are extremely gregarious and love to be in the same (quantal) state of the system. The atoms in this special situation become locked together in their motion, like soldiers in a well-drilled army, and can no longer behave independently. Thus, for example, if the liquid flows through a narrow capillary, the processes of

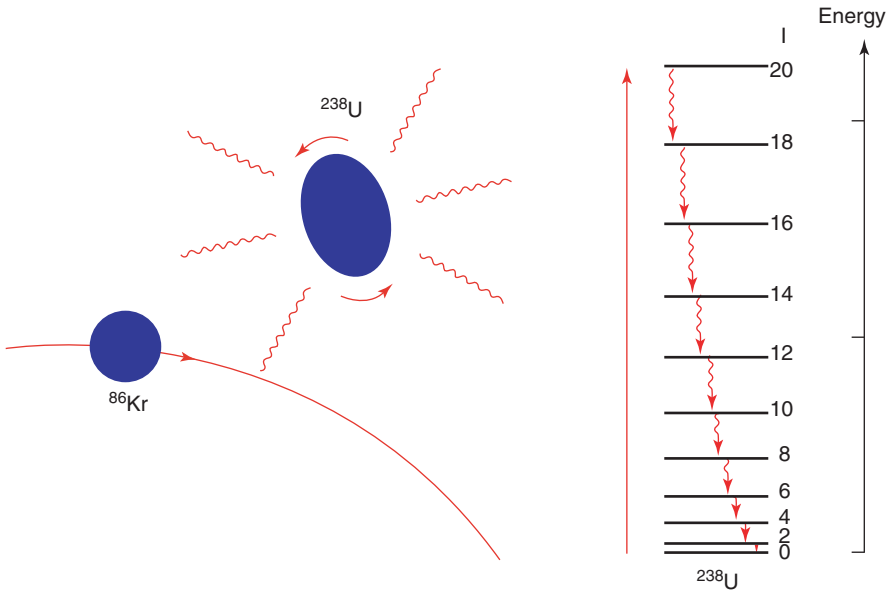


Figure 5 Excitation of the rotational motion of a (target) nucleus of ^{238}U by a ^{86}Kr projectile through the Coulomb field (Coulomb excitation). The system cools down by emitting a number of γ-rays (red wavy arrowed lines) each carrying angular momentum 2 (in keeping with the quadrupole deformation of the nucleus ^{238}U) and different energy, connecting the members of the ground-state rotational band of the system.

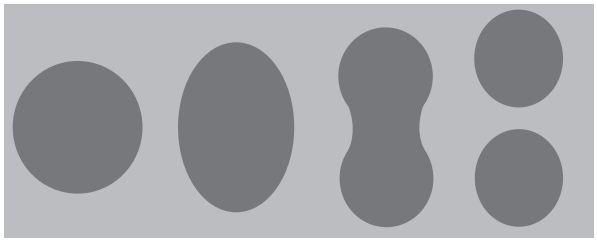


Figure 6 A schematic sequence of events in the process of spontaneous nuclear fission.

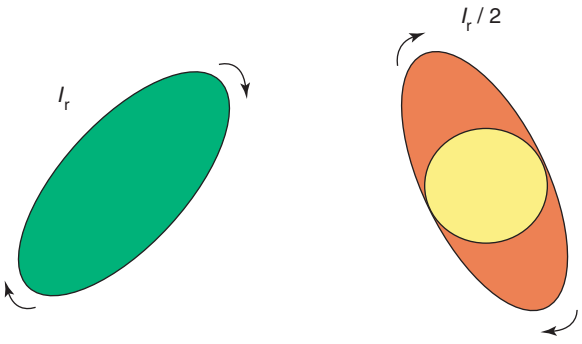


Figure 7 Schematic representation of the reaction to rotation of a deformed nucleus. In the left part of the figure, it is assumed that the system can be described in terms of the motion of independent nucleons (fermions). The associated moment of inertia is equal to the rigid moment of inertia I_r. In the right part of the figure, it is assumed that the system is filled with a superfluid fluid (made out of pairs of fermions) displaying no friction. The associated moment of inertia of this system is $\approx I_r/2$. Consequently, under rotation the central core (yellow) remains essentially at rest, the matter associated with poles (orange) being involved in the rotation.

scattering of individual atoms by roughness in the walls, which would produce so strong a viscous drag on any normal liquid as to effectively prevent it from flowing at all, is now quite ineffective, since all atoms (bosons) or none must scatter.

A similar picture applies to superconductivity in metals. However, in this case the "particles" that undergo BEC (and are therefore required to be bosons) are not individual electrons, which have spin 1/2 and are fermions. The "particles" are pairs of electrons (Cooper pairs) which form in the metal, and carry mass and charge which are double that of a single electron aside from total spin zero. Cooper pairs condense when the metal is cooled down (e.g., below 7 K in the case of lead) and occupy a single quantal state. A similar phenomenon takes place in the nuclear case, where pairs of nucleons moving in time-reversal states form Cooper pairs with total angular momentum zero (**Figure 8**). Because the nucleus in its ground state or in the states belonging to the so-called yrast band (that is the states of lowest energy for a given angular momentum, where all the energy of excitation is tied up in the collective, ordered rotational motion of the system as a whole) is at zero temperature, these Cooper pairs condense, giving rise to a superfluid nonviscous liquid.

The variation of the rotational frequency of deformed nuclei displays sharp discontinuities (**Figure 9**).

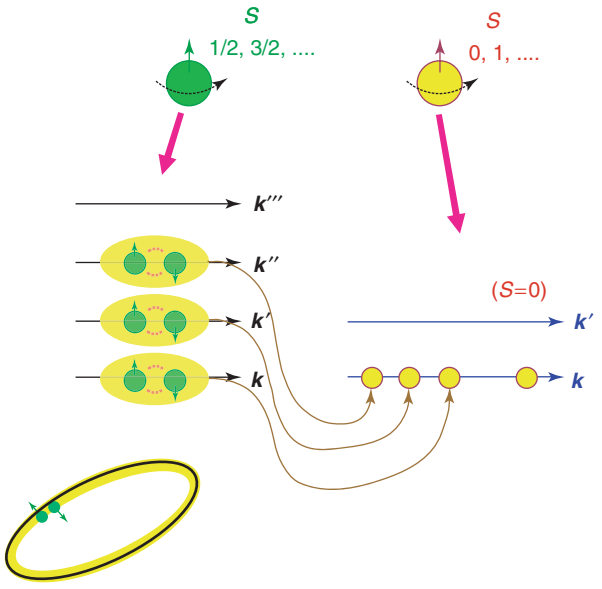


Figure 8 Particles displaying half-integer values of the spin are called fermions, fulfill Fermi–Dirac statistics, and according to the exclusion principles cannot occupy the same quantal state labeled in the figure by the quantum number ***k*** and by the intrinsic spin which can be in one of two states, namely spin up and spin down, associated with the two projections ($\pm 1/2$) of a spin $s = 1/2$ as that of the nucleon or of an electron, and explicitly shown in the figure in terms of an arrow pointing up or down respectively. Particles of integer spin are called bosons and fulfill Bose–Einstein statistics. They can all occupy the same quantal state. Coupling two fermions to integer spin, e.g., $S = 0$, one can form a boson (Cooper pair). If the temperature of the system is lower than the binding energy of the Cooper pair, a condensation phenomenon can also take place in this case. In nuclei, the interaction responsible for the formation of Cooper pairs can be viewed as a surface phenomenon effect (cf. also **Figure 11**).

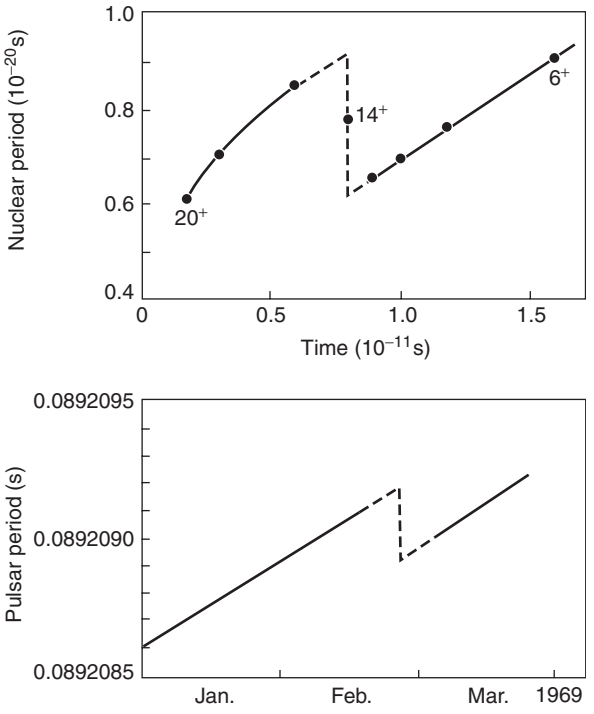


Figure 9 Plot of the rotational period vs. time for the nucleus ^{158}Er (top) and the pulsar Vela (bottom).

This phenomenon reflects the fact that rotation has an opposite effect on the motion of the two members of each Cooper pair (**Figure 10**). After a compound nucleus with high angular momentum and excitation energy has been formed in a fusion process, it cools down by evaporating particles. When the excited nucleus comes close to or eventually reaches the yrast line, it continues to slow down by emitting gamma rays each carrying two units of angular momentum. At the beginning of this process, the rotational frequency is so high that the associated energy is larger than the binding energy of each pair. Consequently, no Cooper pair is present in the system, and the nucleus behaves as a normal system displaying a rigid moment of inertia. As the nucleus continues to decrease its rotational frequency, it will eventually arrive at a (critical) frequency below which pair formation can take place, and the system can make a transition to the superfluid phase. Consequently, the moment of inertia of the system decreases by a factor of ~ 2. The rotational frequency thus increases sharply. This is because the product of the moment of inertia and the rotational frequency is equal to the total angular momentum of the system, a quantity which is conserved. Thus, the observed decrease of the rotational period is intimately related with the normal superfluid-phase transition taking place as a function of the rotational frequency. This transition is associated with a violent variation of the moment of inertia of the system, that is, to a nuclear-quake.

In the case of low-temperature superconductivity, the attraction among electrons is generated by the exchange of lattice phonons between them (and the energy needed to break the pair is of the order of a few milli (10^{-3}) electronvolts). In the nuclear case, roughly 50% of the pairing effects is due to a phenomenon similar to this one, namely the exchange of collective surface vibrations between nucleons. The other 50% is due to the strong force acting between nucleons and arising from the exchange of mesons (the carriers of the strong force) between nucleons (cf. **Figure 11**). The effect which this force has on nucleons moving in states connected by the operation of time reversal (**Figure 10**), can be measured in the scattering process between nucleons and expressed in terms of the $^1S_0$ phase shift (cf. **Figure 11** and corresponding caption). It is found that the exchange of mesons leads to a pairing force which is attractive at low relative momenta, that is, in situations similar to

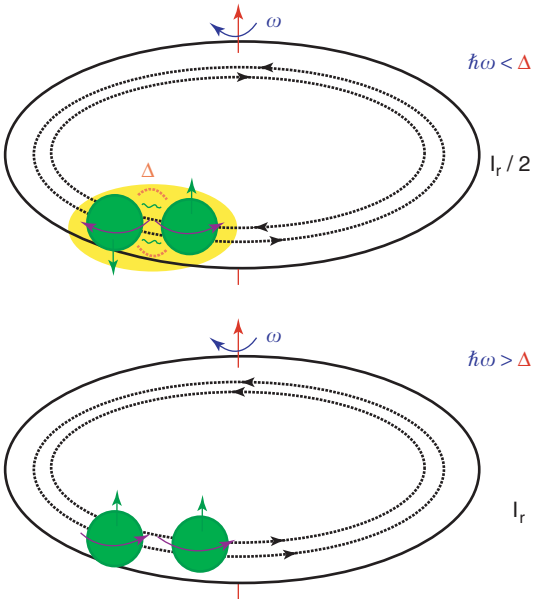


Figure 10 Pair of nucleons moving in states of time-reversal (clockwise and anticlockwise, spin up and spin down) and forming a Cooper pair in a deformed nucleus which rotates as a whole with frequency ω. The binding energy of the pair is measured by the quantity Δ, known as the pairing gap, the interaction among the nucleons being represented by red dotted curves and green wavy lines (see also **Figure 11**). For rotational frequencies such that the rotational energy $\hbar\omega$ is smaller than the pairing gap (top), the system behaves as a condensate of Cooper pairs (bosons), with the moment of inertia being one-half the rigid moment of inertia (cf. **Figure 9**). For rotational frequencies such that $\hbar\omega$ is larger than Δ, the nucleon moving clockwise in its trajectory is so much retarded in its revolution period with respect to the partner nucleon, that it cannot correlate efficiently any more with it and "align" its motion (and spin) with the rotational motion, becoming again a pair of fermions and not participating any more in the condensate (cf. **Figure 9**).

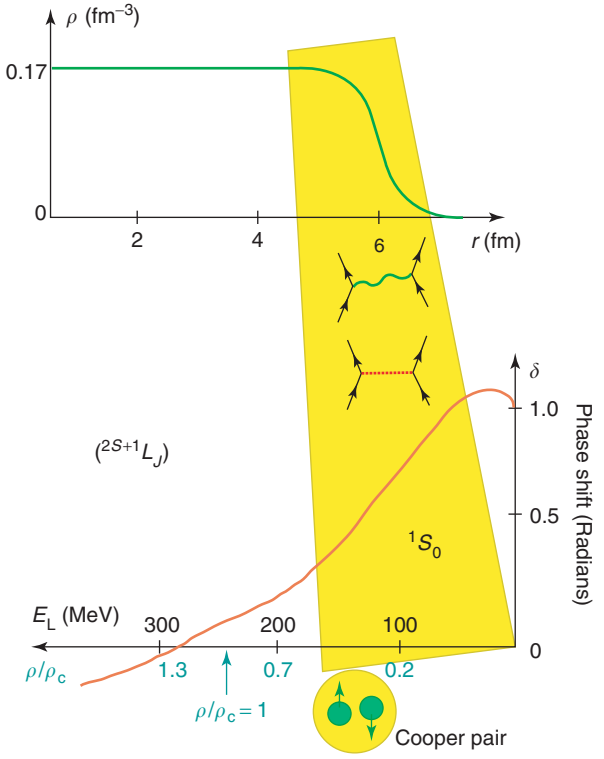


Figure 11 (top) Nuclear density ρ in units of fm^{-3} (where $\text{fm} \equiv 10^{-13}$ cm), plotted as a function of the distance r (in units of fm) from the center of the nucleus. Saturation density corresponds to $\approx 0.17\ \text{fm}^{-3}$, equivalent to $2.8 \times 10^{14}\ \text{gm cm}^{-3}$. Because of the short range of the nuclear force, the so-called strong force, the nuclear density changes from 90% of saturation density to 10% within 0.65 fm, the so-called nuclear diffusivity. (bottom) Phase parameter associated with the elastic scattering of two nucleons moving in states of time reversal, the so-called $^1S_0$ phase shift, in keeping with the fact that the system is in a singlet state of spin zero. The solution of the Schrödinger equation describing the elastic scattering of a nucleon from a scattering center (in this case another nucleon) is, at large distances from the scattering center, a superposition of the incoming wave and of the outgoing, scattering wave. The interaction of the incoming particle with the target particle changes only the amplitude of the outgoing wave. This amplitude can be written in terms of a real phase shift–or scattering phase–δ. Positive quantities of δ implies an attractive interaction, negative a repulsive one. For low relative velocities (kinetic energies E_L), i.e., around the nuclear surface where the density is low, the $^1S_0$ phase shift arising from the exchange of mesons (like, e.g., pions, represented by a horizontal dotted line) between nucleons (represented by upward pointing arrowed lines) is attractive. This mechanism provides about half of the clue to nucleons moving in time-reversal states to form Cooper pairs. These pairs behave like bosons and eventually condense in a single quantal state leading to nuclear superfluidity. Cooper pair formation is further assisted by the exchange of collective surface vibrations (wavy curve in the scattering process) between the members of the pair.

that experienced by pairs of nucleons moving on the surface of the nucleus.

The role that the exchange of nuclear surface vibrations has on Cooper pair formation is larger for small rather than for large nuclei. This is because the surface to volume ratio, and thus the curvature of the system as well as the spill-out of nucleons is larger for the light nuclei. Consequently, in these systems the collectivity of surface vibrations is larger than in the heavier systems. These effects become exacerbated in the case of very light exotic nuclei. Effects similar to these open the possibility of creating real high T_c superconductors when applied to the case of fullerene-based materials.

The central role played by the surface in the condensation process leading to nuclear superfluidity, where nucleons move without friction, is demonstrated by the fact that BEC essentially happens only at the surface of the nucleus. In other words, nucleons behave in the nuclear surface like a superfluid, while inside the nucleus they display normal properties, including friction. The pairing gap, the energy which provides a quantitative measure of the binding energy of the Cooper pairs and thus of the superfluidity of the

system, is large ($\equiv 1$ MeV) at the nuclear surface, becoming small (zero) inside the nucleus.

Exotic Forms of Nuclear Matter

Surface-controlled pairing is also believed to be responsible for the existence of exotic forms of nuclear matter at very low densities. These so-called halo nuclei, which lie at the limits of stability of the chart of nuclides (cf. **Figure 12**), are composed of essentially all surface and display a very large spill-out of nucleons.

Halo Nuclei

These nuclei exist at the neutron drip line in the chart of nuclides (see below and **Figure 13b**), and their study constitutes one of the most important subjects lying at the forefront of nuclear research.

When neutrons are progressively added to a normal nucleus, the Pauli principle forces them into

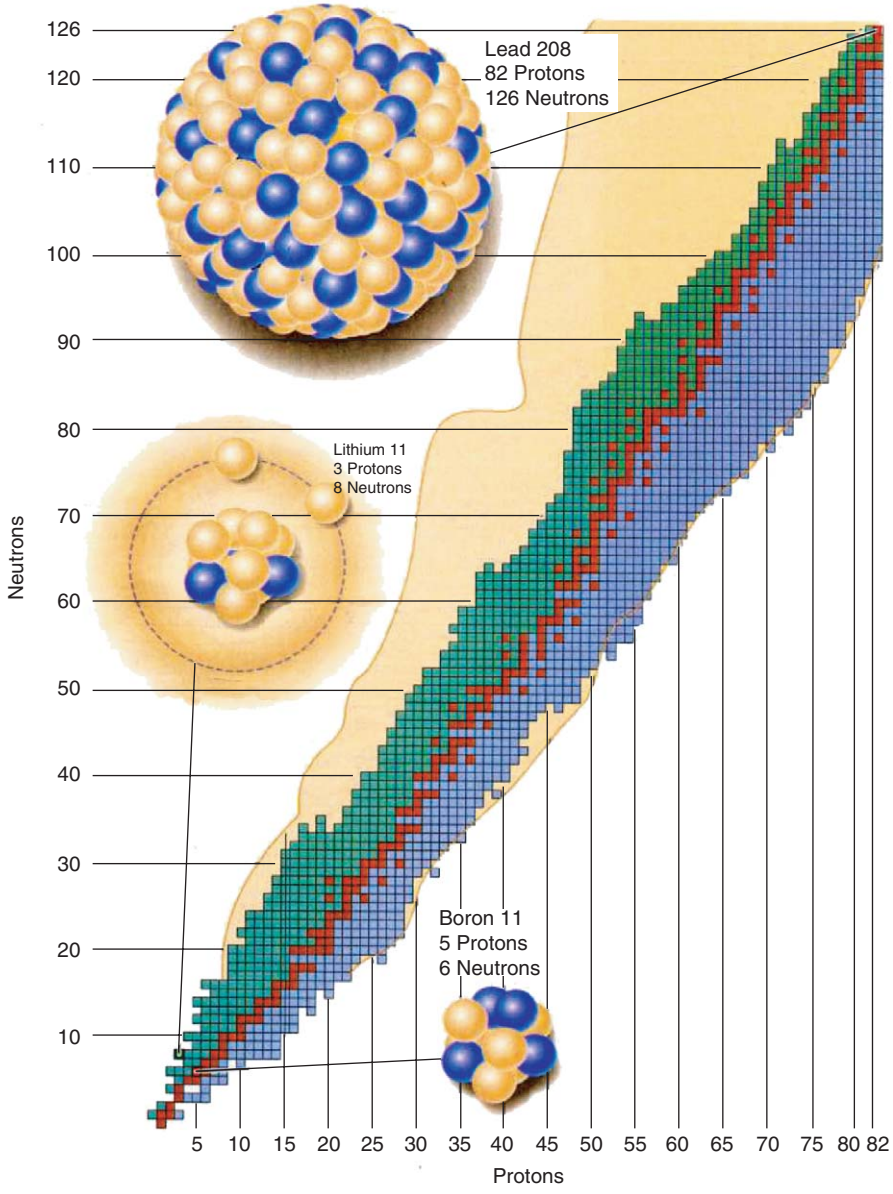


Figure 12 Nuclei are represented here by squares, positioned horizontally according to the number of protons and vertically according to the number of neutrons they contain. Stable nuclei are shown as red squares. Green squares indicate bound but unstable nuclei that have excess neutrons. Blue squares mark bound, proton-rich nuclei. The outer borders of these regions are called drip lines, along which large halo nuclei may be found. The extended halo of lithium 11, comprising two neutrons, makes this nucleus nearly twice as large as boron 11, a stable nucleus containing the same number of nucleons. ^{11}Li is half the size of a lead nucleus, ^{208}Pb, that holds 197 more particles.

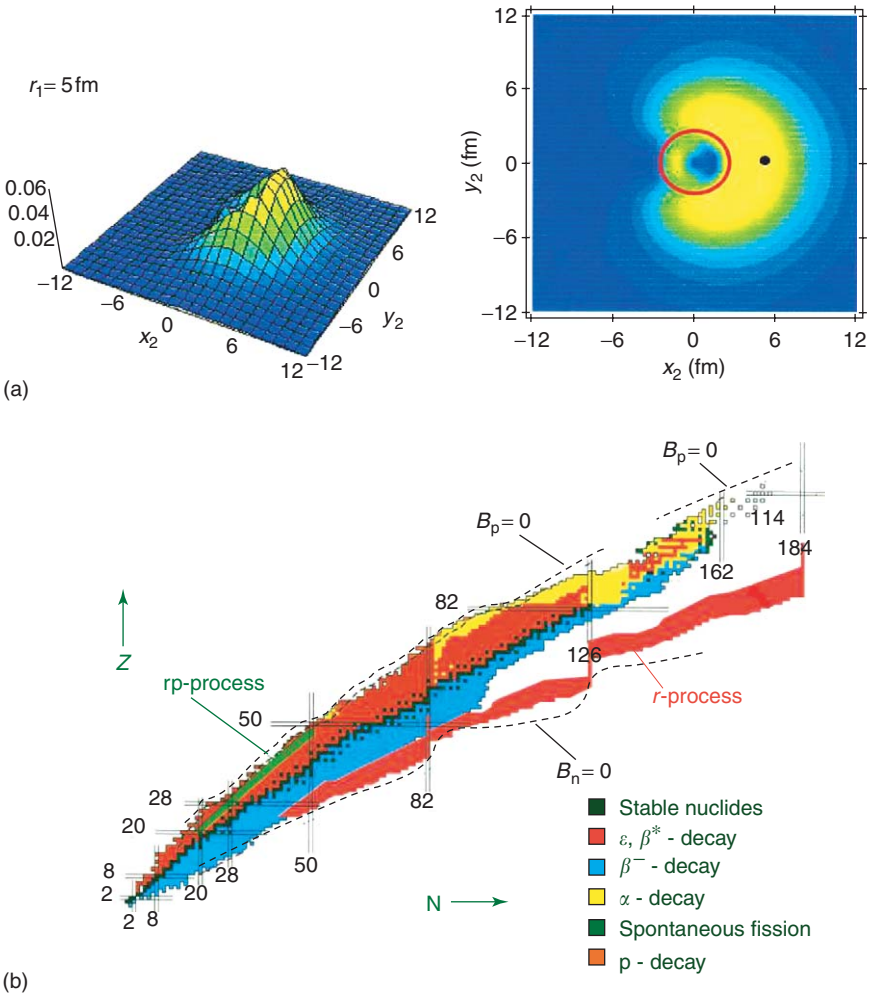


Figure 13 (a) Spatial structure of the Cooper pair which describes the two halo neutrons of ^{11}Li. Fixing one of the two neutrons at a distance of 5 fm from the origin of the nucleus (black dot) on the *x*-axis, the probability of finding the other neutron is displayed. The numbers appearing on the *z*-axis of the 3D plot displayed to the left side are in units of fm^{-2}. The projection on the *x*–*y* plane is shown on the right panel, where the core ^{9}Li is shown as a red circle. (b) Map of the nuclear landscape, as in **Figure 12**, where particular characteristics which makes it easier to appreciate the new area of research associated with exotic nuclei are emphasized. The nuclei existing naturally on Earth are marked by black squares. The magic numbers associated with increased stability (closed shells), are the convenient "landmarks" in the map and are indicated by double lines. Also shown are estimates of the borderline of stability, the so-called drip lines, where the proton and neutron binding energies B_p and B_n become zero. Two paths which astrophysicists identify as important in the creation of elements in the Universe are also displayed: the r-process and the rp-process which traverse regions yet unexplored. A third path, the s-process path, runs along the stable nuclei.

states of higher momentum. When the core becomes neutron-saturated, the nucleus expels most of the wave function of the last neutrons outside to form a halo, which because of its larger size can have lower momentum. In most cases neutrons essentially drip off from the nucleus, defining the lines of stability for neutron number in the chart of nuclides (neutron drip line, cf. **Figure 13b**, dashed lines marked $B_n = 0$). In some few, selected cases, the system becomes a halo nucleus. In halo nuclei, some of the constituent neutrons venture beyond the drop's surface and form a misty cloud or halo (similar effects are found for nuclei close to the proton drip line, cf. **Figure 13b** dashed lines marked $B_p = 0$). Not surprisingly, these extended nuclei behave very differently from ordinary ("normal") nuclei lying along the stability valley in the chart of nuclides. In particular, they are larger than normal nuclei of the same mass number. For example, ^{11}Li is twice as large as ^{11}B, a stable nucleus containing the same number of nucleons and half the size of the lead nucleus ^{208}Pb, which holds 197 more particles (**Figure 12**). In the case of ^{11}Li, the last two neutrons are very weakly bound. Consequently, these neutrons need very little energy to move away from the nucleus. There, they can remain in their "stratospheric" orbits, spreading out and forming a tenuous

halo. If one neutron is taken away from ^{11}Li, a second neutron will come out immediately, leaving behind the core of the system, the ordinary nucleus ^{9}Li. This result indicates that pairing may play an important role in the stability of ^{11}Li. In fact, the properties of the exotic nucleus ^{11}Li can be understood, even quantitatively, in terms of the simplest scenario imaginable: the formation of a single di-neutron Cooper pair which is held together to the ^{9}Li core essentially by the exchange of surface vibrations among the halo nucleons (**Figure 13a**). The importance of these vibrations is due to the high polarizability displayed by ^{11}Li.

It is to be noted that of all the atomic nuclei which are thought to be possible, only about half of them have been observed in the laboratory, and a small

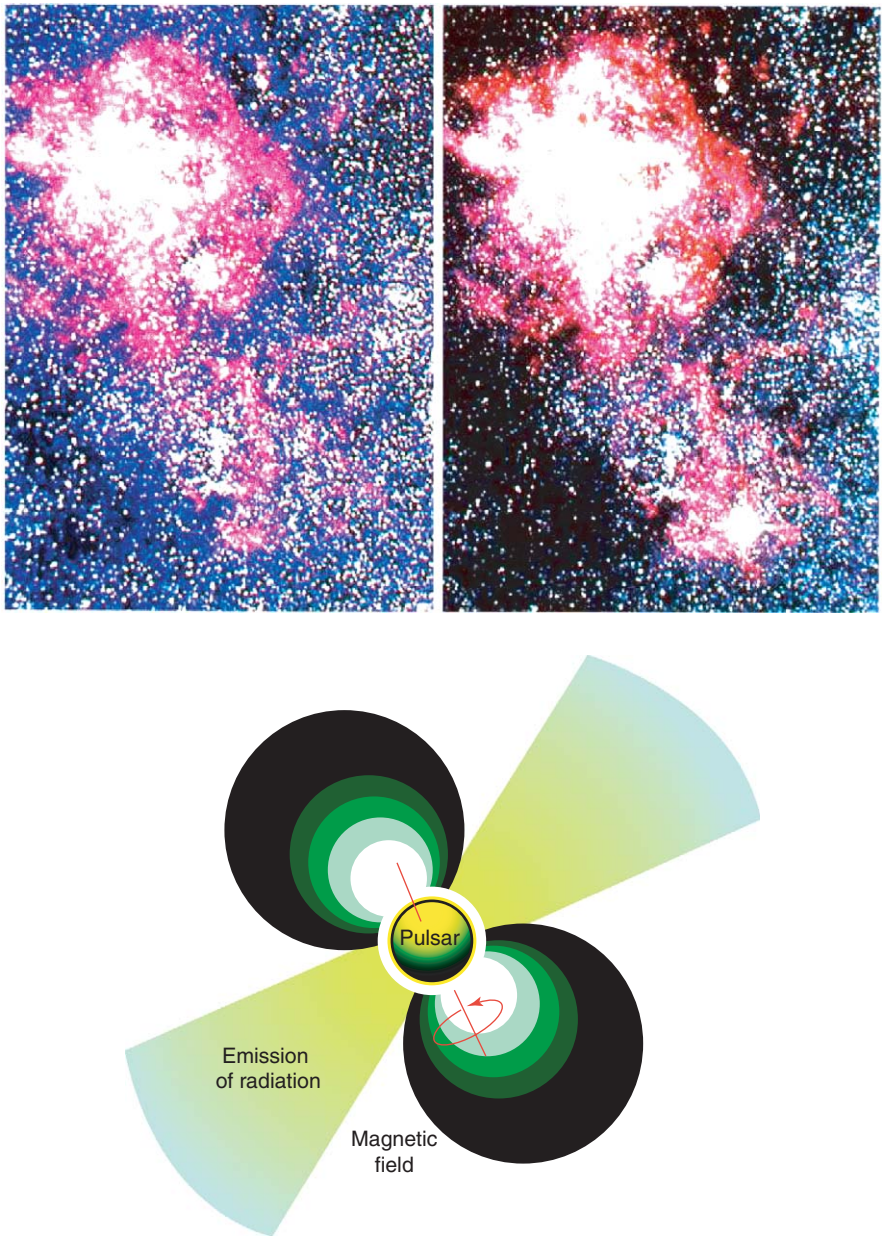


Figure 14 On 23 February 1987, the great Magellean cloud was the scene of a Supernova explosion. Supernova 1987A, as it was christened, was so bright that one could see it from the southern hemisphere and during a period of about six months with the naked eye. (top) A photo of the Magellean cloud before (left) and after (right) the date indicated above. A bright spot on the lower part of the figure is clearly visible. (bottom) Neutron stars have an intense magnetic field which accelerates electrons and protons, inducing those particles to emit beams of radiation. These beams arrive to the earth with a frequency determined by the rotation of the star giving rise to a pulsating signal like light emitted by a lighthouse. Hence, the name pulsar.

fraction exists naturally on earth (cf. **Figure 13b**). Of particular importance in this research is the mechanism for the creation of elements in the Universe. Unstable nuclei are involved in explosive burning in astrophysical environments. Prominent among these processes are explosive hydrogen burning (the rp-process) on the surface of accreting white dwarfs (novae) or neutron stars (X-ray bursts), as well as rapid neutron capture (the r-process) in very neutron-rich conditions of Supernova explosions (cf. **Figure 14**). The r- and rp-process paths shown in **Figure 13b**, traverse regions of unstable nuclei of the chart of nuclides which are as yet unexplored. A third path, the s-process path, runs along the stable nuclei.

Neutron Stars

The nuclear surface controls not only the properties of individual nuclei, but also the properties of "materials" whose building blocks are atomic nuclei. These femtometer materials display properties reflecting not only the ubiquitous role played by the interweaving of nucleons and surface vibrations, but also the marked dependence of the strong force with density. A textbook example is provided by neutron stars (pulsars). These remnants of fierce Supernova explosions are gigantic, rapidly rotating nuclei thirty kilometers across, held together by the gravitational force (**Figures 14** and **15**). Neutron stars display, in the process of cooling down through radiation and particle emission, marked glitches (starquakes) (**Figure 9**), which are likely to be connected with nuclear superfluidity. In fact, neutron stars usually rotate with such precision that they are known as the best time keepers in the universe. But ever so often their rotation rate increases. It is thought that these glitches are related to superfluidity inside the star, in particular superfluidity in the inner crust of the star, where nuclei forming a crystalline lattice are immersed in a sea of free, superfluid neutrons.

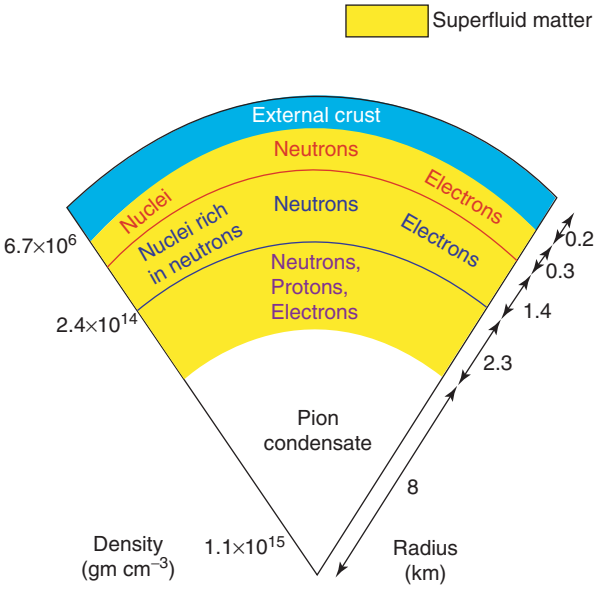


Figure 15 Cross section of a neutron star calculated making use of a so-called stiff equation of state for neutron matter densities larger than or equal to nuclear saturation density.

See also: Nuclear Fission and Fusion; Nuclear Models and Methods; Radioactivity.

PACS: 21.30.Fe; 21.10.Ky; 21.60.−n; 26.60.Fc; 97.60.Jd; 21.60.Cs; 21.60.Jz; 21.10.Re

Further Reading

Bohr A and Mottelson BR (1975) *Nuclear Structure*, Vol II. New York: Benjamin.

Brink D and Broglia RA (2004) *Nuclear Superfluidity: Pairing in Finite Systems*. Cambridge: Cambridge University Press.

Heyde K (1998) *From Nucleons to the Atomic Nucleus*. Berlin: Springer.

Nilsson SG and Ragnarsson I (1995) *Shapes and Shell in Nuclear Structure*. Cambridge: Cambridge University Press.

Quasicrystals

M de Boissieu, Laboratoire de Thermodynamique et Physico Chimie Métallurgique, St. Martin d'Hères, France

Introduction

The atomic structure of condensed matter has for a long time been described by the ideal crystal model: a unit cell, which repeats periodically toward infinity, and is decorated by a motif (atomic positions). The atomic structure is completely determined by the decoration of the unit cell. Only particular symmetries are compatible with periodicity, leading to the 14 Bravais lattices and 230 different space groups. For instance only two-, three-, four-, and sixfold rotation axes are allowed, the fivefold rotational symmetry being incompatible with periodicity.

Extremely powerful tools have been developed for solving periodic atomic structures by means of the

analysis of the intensity distribution in the X-ray or neutron diffraction pattern and complex structures such as proteins, which contain several thousands of atoms in the unit cell, are now almost routinely determined. Indeed, if an X-ray (or neutron or electron) beam is sent on a periodic crystal, it is diffracted in specific directions and forms a diffraction pattern. The diffraction pattern of a crystal is best understood in the so-called reciprocal space in which it forms a periodic lattice. The position of reciprocal lattice vectors or Bragg peaks is given by the Bragg law which relates the lattice plane d-spacing, the reciprocal lattice vectors, and the incoming radiation wavelength. The reciprocal lattice has the same point group symmetry as the crystal under study, that is, if a crystal has a threefold rotation, the reciprocal space also has threefold symmetry. The intensity distribution of the Bragg peaks depends on the atomic decoration of the unit cell, being the square of the Fourier transform of the structure. All structural information is contained in this intensity distribution, but cannot be used directly because the phase of the Fourier transform is unknown. Direct methods allow one to overcome this "phase problem" and give a starting model which is further refined against the observed Bragg peak intensity distribution.

The power of the periodic description of condensed matter is also that it allows describing their physical properties by means of the Bloch theorem, for instance. Real materials are in most cases considered as a perturbation of the ideal crystal in which defects are introduced (vacancies, dislocations, chemical disorder, etc.).

At the other extreme, if too much disorder is present, the material is described by the glass model in which only short-range order exists: it is characterized by a diffraction pattern in which Bragg reflections no longer exist, and presents a smooth variation as a function of the scattering angle.

In 1984, Shechtman *et al.* published the electron diffraction pattern of a rapidly cooled AlMn alloy which displayed Bragg reflections and fivefold symmetry axes. This was a break in the crystal description of matter, since two properties, believed to be incompatible, were present: long-range order, as evidenced by the Bragg peaks, and fivefold symmetry. An example of the Laue diffraction pattern of a high-quality Al–Pd–Mn icosahedral quasicrystal is shown in **Figure 1**: Bragg peaks are sharp and the fivefold symmetry, characteristic of the icosahedron, is clearly visible. The discovery of Shechtman was named quasicrystal and is a new kind of long-range ordered matter. Although it gave rise to a controversy in the crystallographic community, it appears now as the new paradigm for long-range ordered structure.

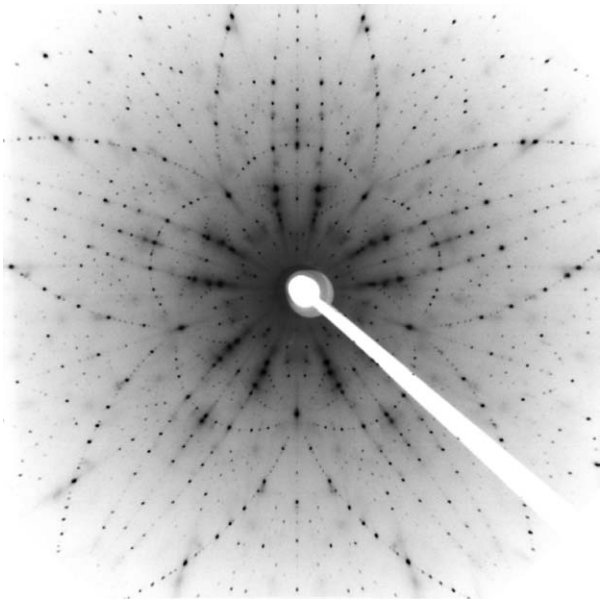


Figure 1 Fivefold axis Laue X-ray diffraction pattern of the i-AlPdMn phase. (Courtesy of W Steurer, ETH Zürich, Switzerland.)

Nonperiodic long-range order had already been observed before the discovery of the quasicrystal, in incommensurately modulated structures but could, in general, be described as a perturbation on a periodic underlying structure. Aperiodic crystals encompass long-range ordered, nonperiodic structure such as incommensurately modulated structures, composite structures (where two incommensurate lattices are in interaction), and quasicrystals. Tools for structure determination are common to all aperiodic crystals and make use of the so-called high-dimensional crystallography.

Quasicrystal Synthesis

The first quasicrystalline samples were obtained by rapid quench from the melt and were metastable. As a consequence, only small single grains ($< 1\,\mu m$) could be synthesized and the only tool to pinpoint the quasicrystalline character was electron diffraction. Since then a large number of quasicrystalline phases have been discovered, all made of metallic alloys (AlMn, AlPdMn, ZnMgY, AlNiCo, etc.). Some of them are stable at room temperature and can be obtained by slow cooling from the melt, or usual growth process such as Bridgman or Czochralski, leading to mm- or even cm-sized single-grain quasicrystals. Their composition range is, in general, very limited, almost as a line compound. The stabilizing mechanism of the quasicrystal is far from being understood, but size effects and Hume–Rothery electronic stabilization are believed to play a crucial role.

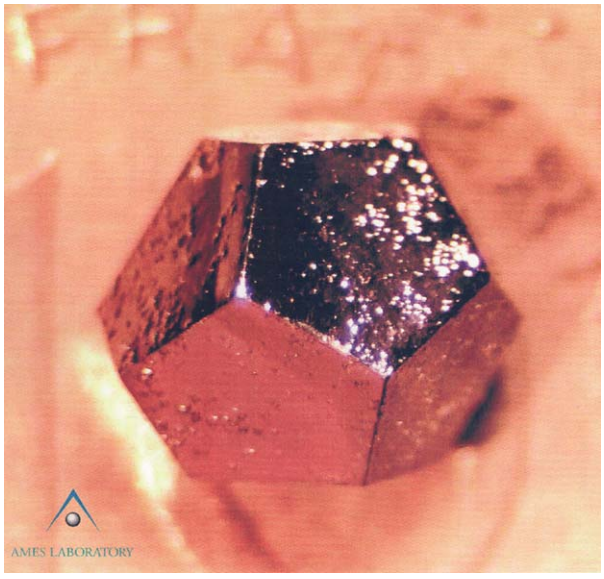


Figure 2 Single grain of the ZnMgY phase obtained by slow cooling from the melt. (Courtesy of P Canfield and I Fisner, Ames Laboratory, USA.)

Figure 2 displays an example of a single grain of the ZnMgY icosahedral phase obtained by slow cooling from the melt which shows a macroscopic icosahedral symmetry. Besides the icosahedral symmetry, quasicrystals have been obtained in systems which are a periodic stacking of quasiperiodic planes with 8-fold, 12-fold, and 10-fold symmetry. However, most of the structural studies have been carried out so far in a few icosahedral and decagonal phases, where the quasicrystalline long-range order is well defined.

High-Dimensional Crystallography

The structural analysis of quasicrystals is now best achieved using the high-dimensional description. As an illustration of this procedure, a simple 1D example is considered: the Fibonacci chain. It is made of two tiles whose length is in the ratio $\tau/1$ τ being the golden mean (1.618…). The Fibonacci chain can be built up by a substitution rule L→LS and S→L, where L and S are for the long and short segment. The resulting 1D structure is quasiperiodic and has the form LSLLSLSLLSLLS… . A compact description of this structure is achieved in a 2D periodic space. A square lattice in a 2D space is taken up to start with. The latter is decomposed in two orthogonal subspaces: the physical space, also called parallel space (E_{par}), and a complementary space called perpendicular space (E_{per}, **Figure 3**). The slope of the physical space with respect to the lattice is irrational and is equal to $1/\tau$. The square lattice is

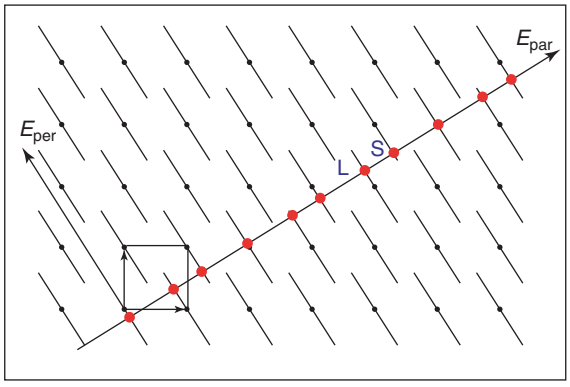


Figure 3 Illustration of the 2D description of a 1D quasicrystal, here the Fibonacci chain (see text).

decorated with segment lines, elongated along the perpendicular direction, and whose length is equal to the projection of the square onto the perpendicular space. The 1D quasicrystal is obtained as a section of the decorated periodic lattice by E_{par}: each time the E_{par} line intercepts a segment line, an atomic position is generated. The final result is the Fibonacci chain. The advantage of the high-dimensional description is that, since periodicity is recovered, usual crystallographic tools such as Patterson analysis or structure factor calculations are easily carried out. Moreover, all the information on the structure (local environment, interatomic distances, etc.) is now contained in a very compact form as the decorated square lattice. More complex structures can be generated: for instance, the segment line can be given a longer length; this will generate additional positions in the 1D quasicrystal. It is also possible to add other segment lines in the square unit cell, which could mimic another atomic species.

This procedure generalizes to the case of a 3D quasicrystal such as icosahedral phases. In this case, the higher-dimensional space is 6D and the lattice hyper-cubic. Segment lines are now 3D objects called occupation domains or atomic surfaces. The problem of structure determination is that of finding the location and the detailed shape of the occupation domains. Although it might seem a simple task, it is a complex problem, for which satisfactory solutions have been proposed only recently. Indeed, specifying the shape of a 3D occupation domain requires an infinite number of parameters, and guidelines have to be found for the modeling.

Return to the 1D case and look for the diffraction pattern of the Fibonacci chain. The diffraction pattern can be computed from the 2D decorated lattice, remembering that the section operation becomes a projection by Fourier transform. For a square lattice, the diffraction pattern can be indexed by 2-integer

indices, and each reciprocal lattice vector has a component in the parallel space and one in the "perpendicular" space. Since the structure is nonperiodic, there is a dense set of Bragg reflections, but only a finite number of them are larger than a given threshold. In the case of the icosahedral phase, 6-integer indices are necessary for indexing the diffraction pattern: they correspond to the six directions pointing toward the vertices of an icosahedron.

The route from diffraction data to the atomic structure is, in some sense, similar to what is achieved in standard crystallography: the point group and space group of the quasicrystal is determined from the diffraction spectrum (see **Figure 1**). Then, the Patterson analysis allows one, in general, to specify approximately the position and shape of the atomic surfaces; from this, a model is proposed which is then compared to available diffraction data. Before presenting experimental results along these lines, the study of the structural quality of quasicrystals is briefly outlined.

Structural Quality of Quasicrystal

As for any materials, quasicrystals contain defects (vacancies, dislocations, strain, etc.), which results in either a Bragg peak broadening or diffuse scattering. Because of the quasiperiodic arrangement, the effect of defects on the diffraction depends on both the parallel and the perpendicular component of the reciprocal lattice vector. In the theory of QC, a distortion or a fluctuation associated with the perpendicular component has been named "phason" (this notion will be referred to later), and leads to changes in the diffraction pattern associated with the perpendicular component of the scattering wave vector (Q_{per}). For instance, one can imagine that the slope of the parallel space in **Figure 3** is changing: if the slope becomes rational, this will lead to a periodic structure. Such a change is equivalent to a shear strain in the perpendicular direction and has been named "phason strain." A distribution of phason strain will result in a broadening of the Bragg reflection (which are no longer delta peaks), in very much the same way that a distribution of elastic strain in a periodic crystal leads to a Bragg peak broadening. In the early days of quasicrystal studies, most of the obtained quasicrystals displayed distorted diffraction patterns, or Bragg peak broadening, the notion of long-range quasiperiodic order even being questioned. Quasicrystals, obtained most recently, display a high quasiperiodic long-range order, with Bragg peaks which are almost resolution limited. In some cases, the structural quality is such that dynamical X-ray diffraction has been evidenced. **Figure 4** shows 2D images of Bragg peaks

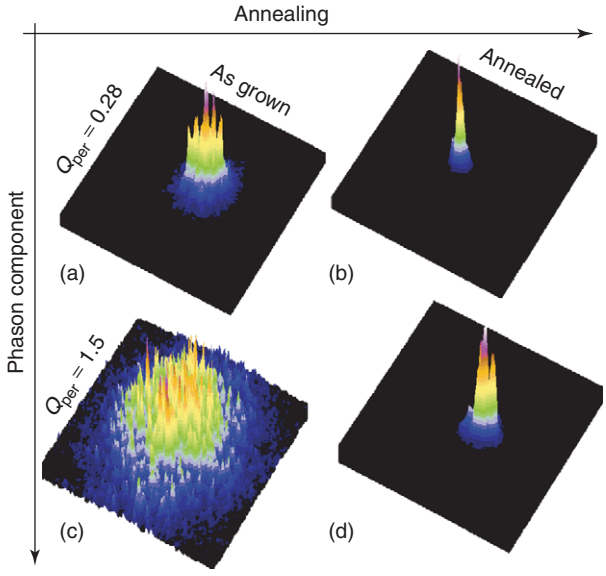


Figure 4 2D images of the Bragg reflections obtained in the i-AlPdMn phase for an as-grown single-grain quasicrystal (left) and the same sample after annealing (right). Whereas Bragg reflections are broadened (with a broadening proportional to Q_{per}) in the as-cast sample, they are almost resolution limited in the annealed sample.

obtained in a high-resolution setting. The left part shows Bragg peaks in an i-AlPdMn sample obtained by the Czokralski method, whereas the right part corresponds to the same Bragg reflections, once the single quasicrystal has been annealed. As it is evident, Bragg reflections are almost resolution limited with a coherent length which is of the order 10 μm. Of course, quasicrystals still contain some defects and are not as "perfect" as silicon for instance, but their quality is comparable to the best metallic alloys, and the notion of long-range quasiperiodic order is strongly supported by experimental results.

From Diffraction Data to a First-Order Model

As an example of structure determination, results obtained in the i-AlPdMn phase, which is stable at room temperature and can be obtained as cm-sized single grains are presented. This is a ternary alloy, with a composition of $Al_{68.9}Pd_{22.1}Mn_9$ with a very narrow composition range. The icosahedral symmetry of this phase is illustrated in **Figure 1** which displays a fivefold symmetry Laue pattern. In order to facilitate the analysis, both X-ray and neutrons diffraction data have been collected. Indeed, scattering power is very different for X-ray and neutrons, Pd atoms being mostly "visible" with X-rays, whereas Mn atoms are more "visible" with neutrons: the use

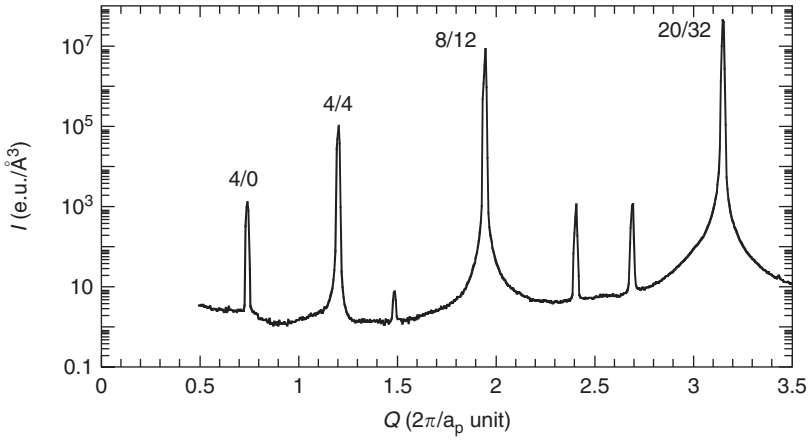


Figure 5 X-ray diffraction pattern taken along a twofold axis in the i-AlPdMn phase. The 2-number indices is a short-hand notation for the 6-indices scheme. Note that the position of the indexed Bragg reflections scales with τ, the golden mean.

of the combined set of data leads to a precise structural information. Indexing of the diffraction pattern allowed determining that the reciprocal lattice is body-centered icosahedral, with systematic absences which are similar to the cubic b.c.c. lattice, that is, 6D indices are all even or all odd. A systematic scan along a twofold direction is shown in **Figure 5**: note that the intensity scale is logarithmic, and that very weak reflections can be measured. The position of all reflections is in perfect agreement with the 6D indexing scheme, and data have been collected using a standard neutron and X-ray diffractometer. Once integrated intensities are collected, the first step is to compute the so-called 6D Patterson function. It is obtained by Fourier transform of the intensities, and represents the autocorrelation function of the 6D structure. In principle, a direct interpretation of this function is difficult, but in several cases, it has been possible to derive the position and the rough size of the atomic surfaces. For instance, in the i-AlPdMn phase, it could be demonstrated that there are only three atomic surfaces, located on the nodes and "body center" sites of the 6D cube. A section of this 6D cube is shown in **Figure 6**, with the position of the three different atomic surfaces. In **Figure 6a**, the section of the 6D periodic structure is such that it contains a fivefold axis both in the physical space and in the complementary perpendicular space. The section also shows the 6D periodicity, and is built up by a linear combination of two 6D lattice vectors with indices [100000] and [011111]. The segment lines, elongated in the perpendicular direction, are just traces of the 3D atomic surfaces. As can be seen in this figure, only three different positions in the 6D cube are occupied, labeled n1, n2, and bc1, the other positions being imposed by the face-centered icosahedral symmetry of the 6D structure (because a body-centered icosahedral reciprocal lattice corresponds to a face-centered icosahedral direct lattice). Moreover, some information on the chemical content of each atomic surface is also obtained by the

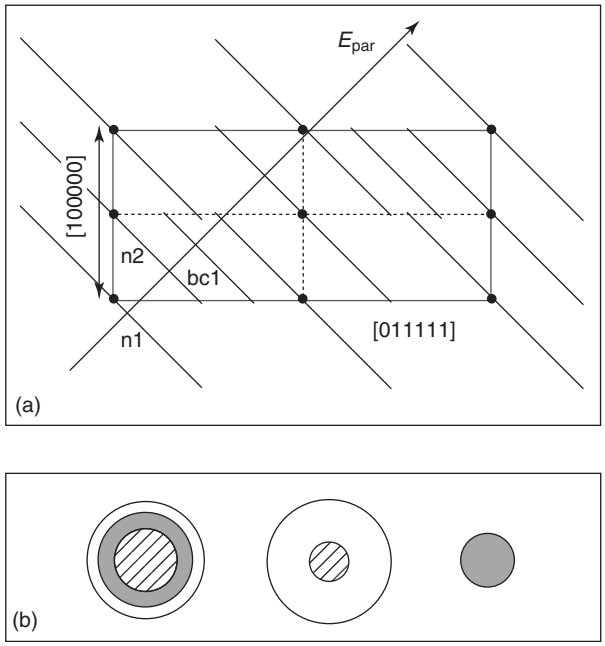


Figure 6 Sketch of the spherical model of the i-AlPdMn icosahedral phase. (a) Section of the 6D space, with a fivefold axis both in the parallel and perpendicular space. The segment lines are traces of the atomic surface. The unit cell is shown as a black rectangle. It is a doubling of a smaller unit cell shown as dashed line, due to a superstructure chemical ordering. There are two different atomic surfaces on the nodes (shown as black and gray lines, and labeled n1 and n2) and a third one on the body-center site (bc1). (b) Decomposition of the different atomic surfaces in the perpendicular space n1, n2, and bc1 (from left to right). There are spherical shells with different atomic composition; the shading stands for the different chemical species: hatched Mn, grey Pd, and white Al.

analysis of the Patterson function. In particular, it can be shown that the n1 atomic surface contains Mn, Pd, and Al atoms, whereas the n2 atomic surface only contains Mn and Al atoms. In a first crude approximation, atomic surfaces are modeled by spheres and spherical shells as illustrated in **Figure 6b**. The chemical content of each atomic surface is modeled by successive shells, with, for instance, a first Mn sphere, enclosed in a Pd shell, and finally an Al external shell for the n1 atomic surface. Although the model presented in the **Figure 6** is simplistic, it already allows reproducing the strongest Bragg peak intensities, and is thus a good first-order model. This model is not only a theoretical view of the 6D structure but it allows one to predict some of the main ingredients of the resulting 3D structure. In particular, it is possible to determine which kind of local environments are to be found in the 3D structure. For example, there are groups of atoms packed along successive shells of icosahedral symmetry, referred to as clusters in the following. Four different types of clusters are found and are packed quasiperiodically. It can also be shown that the structure can be viewed as a quasiperiodic packing of dense atomic planes. This is described again in a later section.

Modeling the Atomic Structure of Quasicrystals

Although the previous first-order model gave some information on the resulting atomic structure, it is obviously necessary to go beyond this model if any detailed information has to be predicted. This is particularly important for the theoretical study of physical properties of quasicrystals. It has to be pointed out that the comparison with diffraction data remains a powerful test, especially for weak Bragg reflections which are very sensitive to the details of the atomic structure (details of the atomic surface shape, for instance).

Modeling of the atomic structure of quasicrystals requires, in principle, an infinite number of parameters since it is an infinite structure. Although the high-dimensional description simplifies the quasicrystal picture, the number of parameters remains much larger than the one which can be refined using the diffraction data. This is why a certain number of constraints or hypotheses have been proposed to restrict models within a reasonable number of parameters.

The first one corresponds to the usual density and chemical composition constraint. The model should also be designed to avoid short distances. Although these two constraints are simple, they are already difficult to fulfill. However, they are not sufficient to restrict the modeling fully and two main routes have been followed to go further: a decoration of tiling, such as a Penrose tiling, or a modeling directly in the high-dimensional space. In both cases, a very fruitful approach has been the comparison with crystalline "approximant": these are periodic crystalline phases which share the same local environment as the crystalline phase. In the high-dimensional picture, a crystalline approximant can be obtained by changing the slope of the cut space which becomes rational. Depending on the symmetry of the periodic approximant, there is a relation between the unit cell of the 6D periodic structure and the cell parameter of the crystalline approximant. Thus, it has been very surprising that such a relationship was observed experimentally. The crystalline approximant serves then as a template for the quasicrystal: for instance, the atomic clusters which are identified in the crystal are used as "building" blocks in the quasicrystal. In the Penrose tiling decoration scheme, the difficulty is then to find the appropriate tile dimension and the cluster connectivity. Moreover, clusters only define a backbone of the structure and it is necessary to fill the "gaps" between clusters. In the high-dimensional scheme, the cluster density can, for instance, be maximized. Other constraints have been proposed on the shape of the atomic surfaces, for instance. The most fruitful approach so far has been one which combines both approaches: using both the knowledge of the local structure obtained from the analysis of a crystalline approximant and a sophisticated 6D modeling, it is possible to achieve a refinement which is similar to what is done in 3D crystallography.

Examples

Icosahedral i-AlPdMn Phase

Because the i-AlPdMn phase was the first one to be obtained with a very good structural quality and as a large single-grain quasicrystal, this is the icosahedral phase for which the most detailed analysis has been achieved. Starting from the crude spherical model and having identified some basic atomic clusters in a periodic approximant, A Yamamoto proposed a model based on a complex atomic surface construction. Starting from a decorated 3D Penrose tiling, the 6D model is built up: this leads to atomic surfaces with a complex shape. Each atomic surface is, in fact, made of small subdomains, each of the subdomains corresponding to a specific local environment. For each subdomain, it is then possible to refine not only a chemical composition, but also a small shift in the real space, which is equivalent to "relaxing" the local configuration. Both parameters are important in the success of the refinement. The quality of the

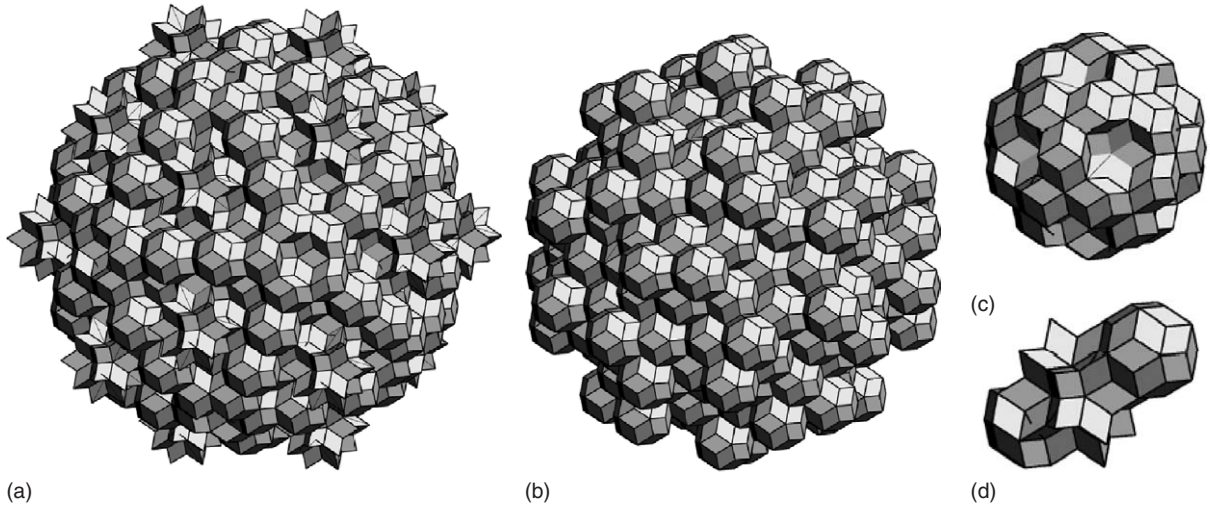


Figure 7 External shape of the atomic surfaces in the i-AlPdMn phase as proposed by A Yamamoto in his model. The average size is similar to the one of the spherical model. (a), (b), and (c) correspond to the n1, n2, and bc1 atomic surfaces of **Figure 6**, respectively. (d) The atomic surface is a supplementary one located on the mid-edge. Each atomic surface is composed of small volumes. (Courtesy of A Yamamoto, NIMS, Tsukuba, Japan.)

refinement is given by the R-factor, which is a measure of the difference between observed and calculated intensity and is found to be of the order 10% or less.

Figure 7 shows the external shape of the four atomic surfaces, which are much more complex than the spheres used in the simpler model! Note that there is a supplementary atomic surface located on a mid-edge position in the 6D cube. The smaller subdivisions are visible. This model compares well with neutron and X-ray data, in particular, for weak reflections. The agreement is, in this case, much better than for the spherical model, and demonstrates that the resulting structure is a good representation of the quasicrystal.

Once the 6D model is constructed, it is possible to deduce some fundamental properties of the resulting 3D structure. In particular, from the 6D model it is possible to determine which atomic clusters are present, and what their distribution is. The icosahedral AlPdMn phase contains two different atomic clusters, each one at least with two different chemical "decoration," which lead to at least four different clusters altogether. Two of these clusters are displayed in **Figure 8**: they correspond to the so-called pseudo-Mackay icosahedron, whose external shell is shown here. One cluster contains mainly Al and Mn atoms, whereas the other one contains mainly Al and Pd atoms. This shows that there is a strong chemical ordering in the i-AlPdMn phase. A more detailed examination shows, however, that there are some sites which are chemically disordered (at least within the experimental error), and that the structure can probably not be described uniquely with only a few

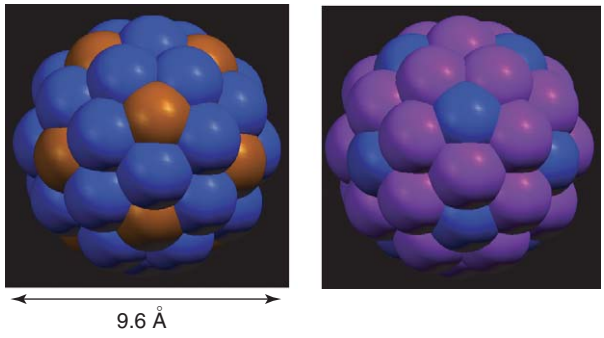


Figure 8 External shell of one of the two icosahedral clusters found in the i-AlPdMn phase. The two average chemical decorations are shown. Left, the cluster contains mainly Al and Mn atoms; right, the cluster contains mainly Pd and Al atoms.

clusters. Nevertheless, this description is useful in understanding the physical properties of quasicrystals. The arrangement of clusters in the 3D structure is perfectly deterministic, though not periodic. In particular, there is a hierarchy of cluster arrangements, with "clusters" of clusters. Another interesting property is the general one which states that if a cluster of radius R exists, then there will be at least one identical cluster in a sphere of radius $2R$. Both of these properties are responsible for the peculiar nature of wavefunctions in quasicrystals.

Another interesting approach of the description of the 3D structure is in terms of dense planes. Of course, the quasicrystal is not periodic and there are, in principle, no periodic planes. However, if we compute the projection of the structure down a high-symmetry axis, there are well visible planes

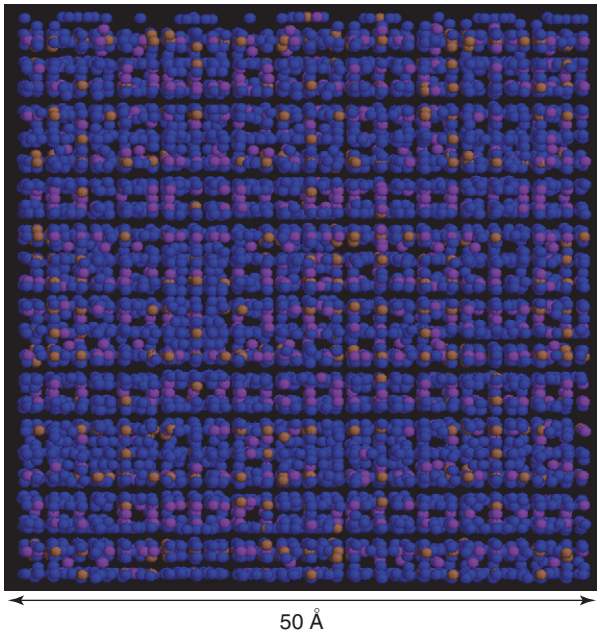


Figure 9 Side view of 50×50 Å cube of the 3D structure generated with the model of **Figure 7**. The vertical direction is parallel to a fivefold axis. Planes are visible and are grouped in small and thick layers. Blue, brown, and pink are for Al, Mn, and Pd atoms, respectively. (Courtesy of A Yamamoto, NIMS, Tsukuba, Japan.)

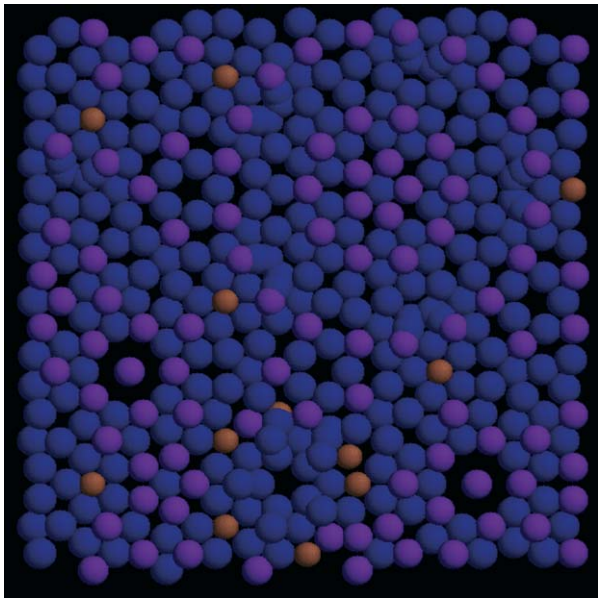


Figure 10 Examples of a fivefold atomic plane in the i-AlPdMn phase as generated by the 6D model of **Figure 7**. Same colors as **Figure 9**. (Courtesy of A Yamamoto, NIMS, Tsukuba, Japan.)

which are quasiperiodically stacked. This is exemplified in **Figure 9** which shows a lateral view of a 50×50 Å of a piece of quasicrystal generated from the 6D model. The different colors are for the different atoms, and the vertical axis is parallel to a fivefold axis. The structure in terms of dense atomic planes is clearly visible: there is a Fibonacci sequence of layers with two different thicknesses. There are some "gaps" in between the layers, which correspond to a physical separation of the planes of the order 1.2 Å. In this description, a strong chemical ordering is also visible, with the Pd atoms being concentrated in specific atomic planes. A characteristic plane is displayed in **Figure 10**. Tenfold and fivefold rings are visible, and correspond to the trace of the clusters.

Such a description in terms of dense atomic planes allows the understanding of some of the properties of quasicrystals. For instance, dense planes are certainly the stable ones and should show up as large facets in macroscopical grains: indeed this is what is observed, since the shape of i-AlPdMn single grains is frequently a dodecahedron, similar to **Figure 1**, with large facets perpendicular to a fivefold axis. Along the same lines of thought, it is possible to have some predictions on the favorable dislocation slip planes, which should be again the fivefold planes. Finally, recent studies on the surface of quasicrystals have evidenced terraces and steps, with a distribution of step heights in agreement with the result of the structure determination.

Decagonal Phase

The decagonal phase is a periodic stacking of quasiperiodic planes. If viewed along the tenfold symmetry axis, interpretation of high-resolution transmission electron microscopy images is possible. Indeed, a large number of models have been proposed from analysis of images. One should keep in mind, however, that the observed images are projections of the structure. In the "simple" decagonal phases, where the periodicity is of the order 4 Å, there are only two different atomic layers and the interpretation is relatively easy. This is not the case, however, for other decagonal phases.

The direct interpretation of images is appealing, since one can directly see the structure, although atomic resolution is difficult to achieve. The first step is to find the underlying quasiperiodic lattice, in the form of a tiling (Penrose tiling or equivalent). Note that it is not obvious, *a priori*, that a quasicrystal can be described by a decorated tiling. Various types of tilings have been used in the d-AlNiCo, most of them being a variant of the Penrose tiling which is made of two tiles. The second step is to find a decoration for the proposed tiling. This is achieved most of the time by super-imposing the tiling on the images. Since atoms are not visible, atom positions are inferred from the known approximant phases. An image is then simulated, and compared to the experimental image.

One interesting modeling has been proposed with the so-called quasi-unit cell approach. Instead of using two tiles, a single atomic cluster configuration is used to map the structure: specific overlap rules are used to generate a complete covering. **Figure 11** shows the electron diffraction pattern together with the corresponding high-resolution images of the d-AlNiCo phase. The single cluster covering is superimposed on the figure, and the decoration is shown in the enlarged part. Agreement between the simulated image and the experimental one is satisfactory. The same procedure, as the one described for the i-AlPdMn phase, has also been used. In this case, the atomic surfaces are 2D objects. Refinement of the occupation parameters and of the parallel shift is carried out in a similar way. The resulting structure is close to the one obtained by HREM images analysis, although a description in terms of a single cluster is probably too crude an approximation.

In conclusion, the atomic structure of the quasicrystal is now rather well understood. The high-dimensional approach is a very powerful tool for analyzing diffraction data and a detailed analysis is now possible. The resulting structure can be analyzed in terms of clusters or dense atomic planes. Whether the structure can be described as a decorated Penrose tiling is still an open question. However, the understanding of the atomic structure is sufficient now for the purpose of simulation or for physical properties calculations.

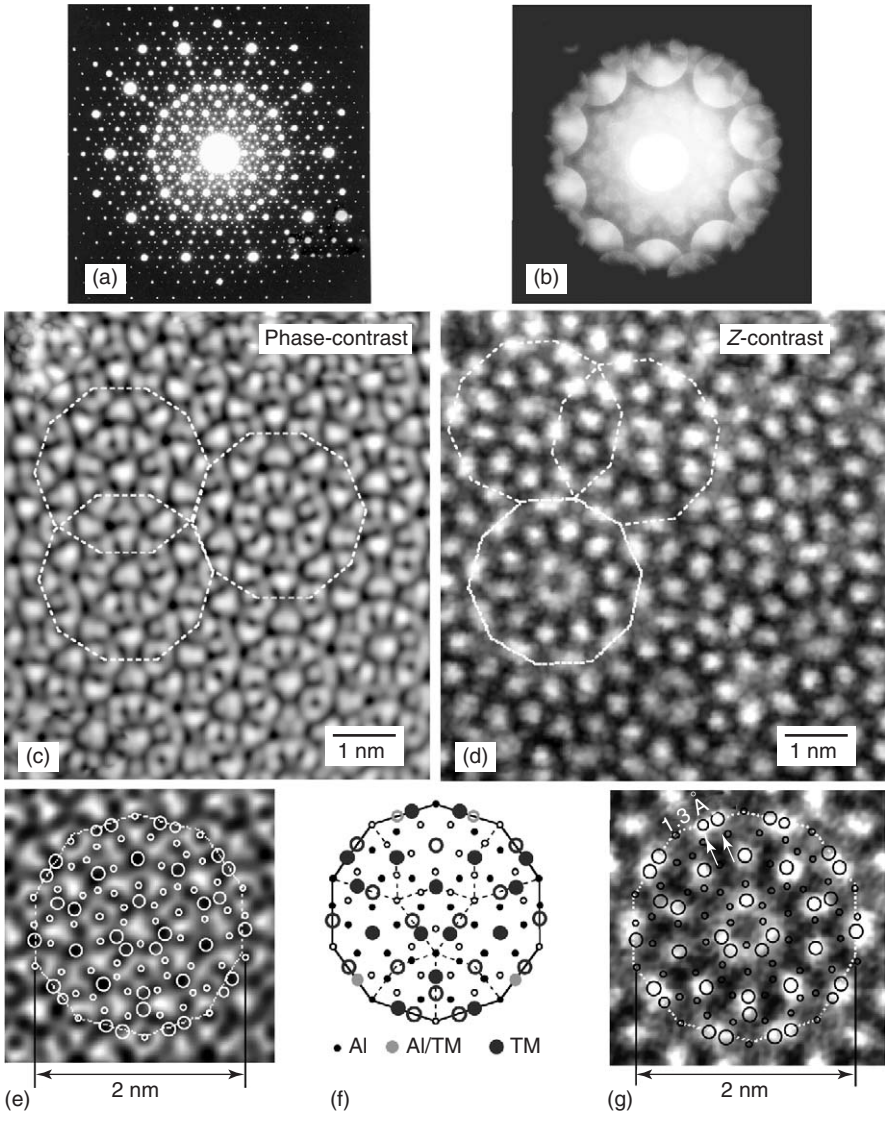


Figure 11 (a, b) Electron diffraction pattern of the AlNiCo decagonal phase, (c) corresponding phase contrast image. (d) The quasi-unit cell is outlined, corresponding Z-contrast images, and (e–g) quasi-unit cell decoration compared to the experimental images. (Courtesy of E Abe, NIMS, Tsukuba, Japan.)

See also: Crystal Growth, Bulk: Methods; Crystal Growth, Bulk: Theory and Models; Crystal Structure Determination, History of; Crystal Symmetry; Film Growth and Epitaxy: Methods; Periodicity and Lattices; Quasicrystals, Electronic Structure of; Space Groups.

PACS: 61.44.Br; 61.10.Nz; 61.12.Ld

Further Reading

DiVicenzo DP and Steinhardt PJ (eds.) (1991) *Quasicrystals: The State of the Art*. Singapore: World Scientific.
Hippert F and Gratias D (1994) *Lectures on Quasicrystals*. Paris: Les Editions de Physique.
Janot C (1994) *Quasicrystals, A Primer*. Oxford: Oxford Science Publication.
Stadnik Z (1999) *Physical Properties of Quasicrystals*. Berlin: Springer.

Quasicrystals, Electronic Structure of

U Grimm, The Open University, Milton Keynes, UK
M Schreiber, Technische Universität Chemnitz, Chemnitz, Germany

Introduction

Quasicrystals are aperiodically ordered solids. Their unusual transport properties, especially the low electric conductivity, are indications of the particular electronic structure of these materials. There are several approaches toward understanding the electronic structure of these systems. Whereas a complete understanding is still lacking, quasicrystals can heuristically be described as Hume–Rothery alloys, which qualitatively account for the experimentally observed properties. Other approaches that have been followed are based on the investigation of electrons in aperiodic potentials and on a hierarchical cluster picture of quasicrystals.

Aperiodically Ordered Solids

Quasicrystals were first discovered in a rapidly quenched AlMn alloy in the early 1980s. Today, many quasicrystalline compounds have been identified, mostly in aluminum or zinc–magnesium-based ternary alloys, some of which can be grown to form nicely faceted single crystals of macroscopic size.

The fingerprint of a quasicrystal is a diffraction which consists of Bragg peaks, like in a periodic crystal, but which shows a symmetry which is not compatible with lattice periodicity in three-dimensional space. So far, the noncrystallographic symmetries that have been observed in experiments are icosahedral symmetry, featuring axes of two-, three-, and fivefold rotational symmetry, and three classes of dodecagonal, decagonal, and octagonal quasicrystals with a single axis of twelve-, ten-, or eightfold rotational symmetry, respectively, which are periodic along the distinguished axis. In contrast to the icosahedral quasicrystals, such phases show anisotropic physical properties.

The apparent paradox of a Bragg-like diffraction with crystallographically forbidden symmetry can be resolved by interpreting quasicrystals as aperiodically ordered solids. The definition of a crystal has been adapted to include quasicrystals, alongside incommensurate structures, as aperiodic crystals. Their structure is usually described in terms of quasiperiodic tilings (or coverings) of space, which play the role of the lattice in conventional crystals. Quasiperiodic tilings can, for instance, be derived as projections of sections through higher-dimensional periodic lattices, which can accommodate the symmetries that are observed in quasicrystals. Structure models consist of decorating the basic building blocks of these tilings with atoms, very much like the unit-cell picture of a conventional periodic crystal. For icosahedral quasicrystals, this yields a six-dimensional description of the structure, for the quasicrystalline phases with one periodic axis the analogous description of a quasiperiodic plane requires at least four space dimensions, and the positions of the Bragg peaks can be indexed accordingly. Direct evidence for the structure is furthermore obtained by high-resolution electron microscopy and atomic force microscopy. While there exist sophisticated structure models for various quasicrystals, many important quetions remain unanswered, for instance concerning inherent disorder in the structure such as stochastic occupancy of atomic sites, or how the quasicrystalline structure actually grows.

Due to the aperiodic structure of these materials, important concepts used in conventional crystals, such as the notion of a Brillouin zone, do not strictly apply for quasicrystals. However, in many cases there are, albeit often less precisely defined, analogs of these concepts, such as the pseudo-Brillouin or Jones zone. This zone is constructed from the main structural peaks, it is nearly isotropic in quasicrystals and, in contrast to the case for amorphous alloys, is a sharply defined polyhedron.

Quasicrystal Surfaces

Many of the experimental techniques, such as photoemission experiments, used to probe electronic properties of quasicrystals are surface sensitive, because the escape depth for elastic electrons is of the order of a few nanometers. Therefore, it is important to know something about the surface structure in order to interpret the experimental results of such measurements.

There has been a long discussion about quasicrystal surfaces, with results depending on the preparation of the surface. However, due to recent progress in experimental techniques, both concerning the preparation of quasicrystalline samples and the image resolution of electron microscopy, it now appears to have been established that quasicrystal surfaces, prepared by sputtering in ultrahigh vacuum and suitable annealing, are themselves quasicrystalline and can be understood as a termination of the bulk structure.

Recent evidence for this is presented in **Figure 1**, which shows a high-resolution scanning tunneling microscopy image of a flat terrace of the fivefold surface of an icosahedral quasicrystal. This technique measures corrugations in the electronic density at the surface arising from the positions of the surface atoms. As shown in **Figure 1**, the pattern obtained for the particular terrace is in good correspondence with a patch of a perfect quasiperiodic Penrose tiling. The matching of the surface structure for this terrace and other terraces with a bulk model of the icosahedral quasicrystal corroborates that the surfaces in this case are consistent terminations of the bulk structure, so there should be no artifacts in experimental measurements due to surface reconstructions.

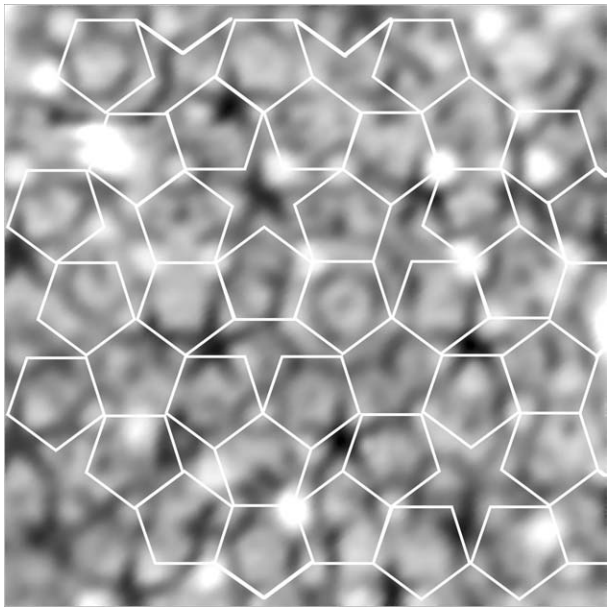


Figure 1 75 Å × 75 Å segment of a scanning tunneling microscopy image of a terrace of an icosahedral Al–Pd–Mn quasicrystal surface with a superimposed patch of a perfect Penrose tiling of edge length 7.8 Å. (Reprinted figure with permission from Papadopolos Z, Kasner G, Diehl RD, Ledieu J, Cox EJ, *et al.* (2002) Bulk termination of the quasicrystalline fivefold surface of $Al_{70}Pd_{21}Mn_9$. *Physical Review* B 66(184207): 1–13; © 2002, American Physical Society.)

Transport Properties

The particular, nearly isotropic, order in quasicrystals, which in a sense is intermediate between that of conventional periodic crystals and amorphous systems, is reflected in unusual electronic properties. Quasicrystals, despite being composed of elements which in their pure form are good conductors, show very low electric conductivity which decreases with temperature and also with the structural perfection of the quasicrystal. So their behavior is not at all metallic; it is rather reminiscent of a system at a metal–insulator transition, with no energy gap but zero conductivity at zero temperature. An example is the icosahedral phase in the Al–Pd–Re system, which shows extremely high resistivity at low temperatures, up to $10^6\,\mu\Omega\,\mathrm{cm}$, with a very strong temperature dependence – the resistivity at a temperature of 0.3 K being about 100 times as large as that at 300 K.

Also other transport properties show interesting phenomena, such as large thermoelectric powers and Hall coefficients that depend crucially on the composition and may even undergo sign changes as the temperature is changed. Though these transport properties partly resemble those of amorphous phases and crystalline complex metal alloy phases of similar composition, the anomalies appear to be particularly pronounced in quasicrystals, in the sense that the residual conductivity is lowest and the thermopower largest for the "optimal" quasicrystal composition, to an extent that such properties, alongside diffraction, are often used to measure the structural perfection of quasicrystals.

An example of an *in situ* measurement on a thin film is presented in **Figure 2**. It shows that, while the temperature dependence of the conductivity is similar in the amorphous and icosahedral phase, the transition to the quasicrystal is accompanied by a strong decrease in conductivity and increase in thermopower, and the low-temperature conductivity in the quasicrystal is nearly zero after several hours of heat treatment.

Experimental values of the conductivity $\sigma_{4\,\mathrm{K}}$ at a temperature of 4 K are in the range of a few tens to a few hundreds $(\Omega\,\mathrm{cm})^{-1}$. The increase in conductivity with temperature is well described by a relationship of the form $\sigma_{4\,\mathrm{K}} = A(\sigma_{300\,\mathrm{K}} - \sigma_{\mathrm{M}})$, which shows that

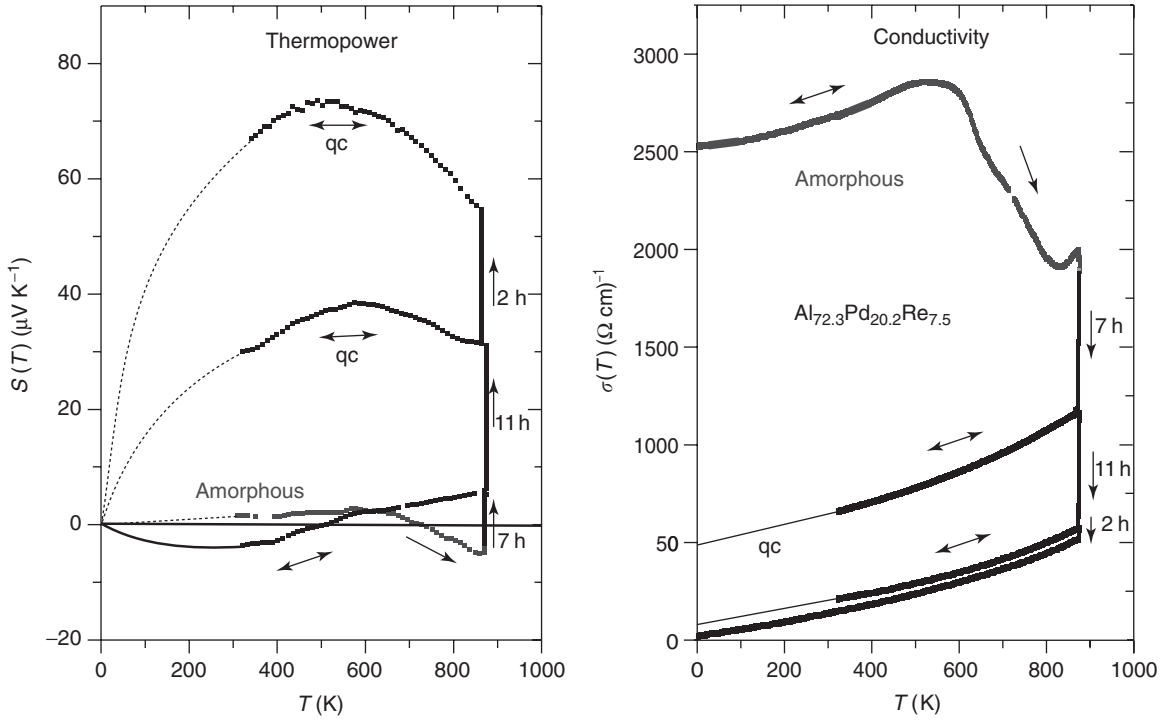


Figure 2 *In situ* measurements of the thermopower (left) and conductivity (right) of an initially amorphous AlPdRe thin film during heat treatment and formation of the quasicrystalline (qc) phase. Arrows indicate reversible and irreversible branches. (Reprinted with permission from Haberkern R (2002) Electronic transport properties of quasicrystalline thin films. In: Suck J-B, Schreiber M, and Häussler P (eds.) *Quasicrystals – An Introduction to Structure, Physical Properties, and Applications,* pp. 364–378. Berlin: Springer; © Springer.)

the low-temperature conductivity could vanish even for a finite conductivity at temperature 300 K. Experimentally, σ_M turns out to be $\sim 150(\Omega\,\mathrm{cm})^{-1}$ for high-quality icosahedral quasicrystals. Given that the actual values for $\sigma_{300\,K}$ are in the range of 200–500 $(\Omega\,\mathrm{cm})^{-1}$, this indicates that these materials are not far from a metal–insulator transition.

For the Hall effect, the temperature dependence can be reasonably well reproduced by a description of the Fermi sphere with electron and hole pockets. In a simple two-band model of free electrons, a change in composition can affect the balance between electrons and holes, and thus account for changes in the sign of the Hall coefficient. Within this picture, the Hall coefficient provides information about the matching between the Fermi surface and the pseudo-Brillouin zone. Similar comments apply to the strong variations and sign changes observed in the temperature dependence of the thermopower, which appear to be correlated to changes in the Hall coefficients and hence might also reflect the sign of majority carriers in the alloy.

Magnetic Properties

One might expect quasicrystals to show interesting magnetic properties. The reason is that most aluminum-based quasicrystals contain a fair amount of transition metals, and there is a class of magnesium–zinc-based quasicrystals containing rare-earth atoms, which provide strong localized magnetic moments. However, no long-range magnetic order has been observed in quasicrystals to date.

In aluminum-based quasicrystals, it appears that only a low fraction of the transition metal atoms in the quasicrystal actually have magnetic moments. One often finds rather small paramagnetic susceptibilities; structurally perfect quasicrystals tend to show diamagnetic behavior, and there appears to be a correlation between diamagnetism and high resistance in the samples. The situation is somewhat different in quasicrystals containing rare-earth elements. Though there have been early reports on an antiferromagnetic ordering in these quasicrystals, these results have not been confirmed and have probably been caused by the presence of conventional crystalline phases in the samples. However, spin-glass freezing transitions have been observed at relatively low temperatures.

Electronic Density of States

To date, the transport properties and the electronic structure of quasicrystals have not been completely

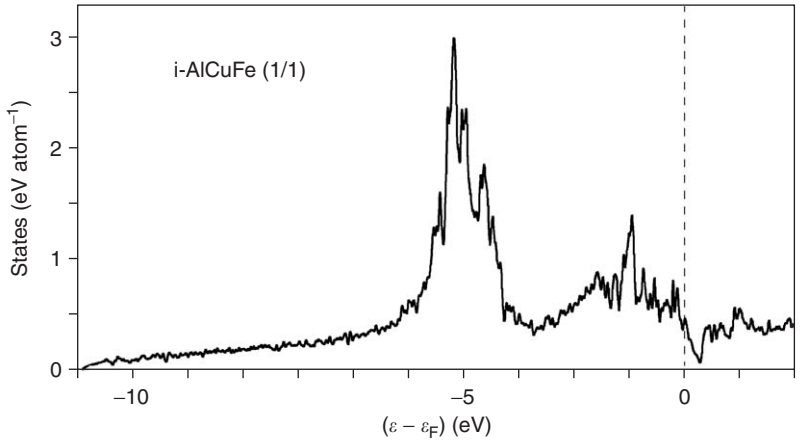


Figure 3 *Ab initio* electronic density of states for the 1/1 cubic approximant of the icosahedral AlCuFe quasicrystal, as obtained by the atomic-sphere approximation to the linear muffin-tin orbital method (ASA-LMTO). The Fermi energy ε_F is indicated by a dashed line. (Courtesy of H Solbrig (Chemnitz).)

understood. Experimental results for aluminum-based quasicrystals such as AlPdMn, which is unique in that it has both icosahedral and decagonal phases, clearly indicate that hybridization between the aluminum *sp* states and the transition-metal *d* electrons plays an important role, which is sensitive to details of the local atomic environments.

There is a rather general heuristic picture that interprets quasicrystals as Hume–Rothery solids. The idea behind this is that there exists a strong interaction between the structure and the electrons, and that quasicrystals occur precisely when the number of electrons available to the system is such that the Fermi surface in momentum space closely matches the pseudo-Brillouin zone. The matching gives rise to a resonant scattering between the conduction electrons and the static structure, leading to a deep pseudogap in the electronic density of states and a stronger tendency toward localization at the Fermi level. Thus, it offers an explanation why the effects are more pronounced in quasicrystals.

This heuristic argument, in combination with the hybridization effects mentioned above, can account for the qualitative behavior of transport properties and their dependency on the composition of the alloy, and has also been supported by experimental and theoretical investigations. It also provides an explanation for the stability of quasicrystals and their existence at certain compositions, because the pseudogap formation reduces the energy of the occupied states close to the Fermi energy, and thus lowers the total energy.

An example of an *ab initio* calculation of the electronic density of states for a periodic approximant of an icosahedral quasicrystal, that is, a cubic system with very similar short-range ordering as the quasicrystal, is presented in **Figure 3**. It clearly shows a pronounced pseudogap feature close to the Fermi energy.

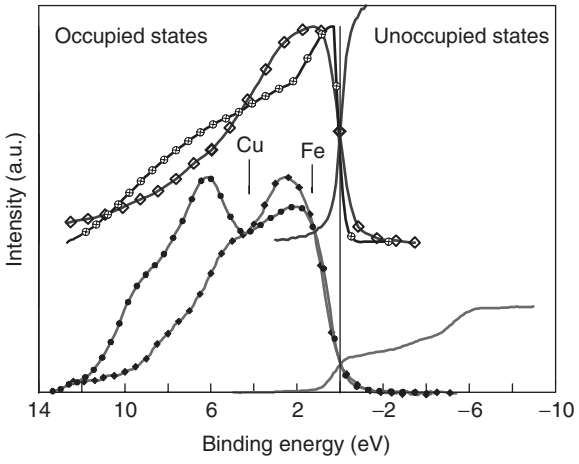


Figure 4 Occupied Al 3*p* and Al 3*s*, *d* (left side of the figure) and unoccupied Al *p* (right side of the figure) spectral distributions in f.c.c. Al (upper set of curves) and icosahedral AlCuFe (lower set of curves). Changes in the shapes of the spectral curves from f.c.c. Al to the icosahedral compound emphasize that the quasicrystal is no longer a free-electron system. The vertical bars show the maxima of the Cu 3*d* and Fe 3*d* distributions in the quasicrystal, respectively. (Courtesy of E Belin-Ferré (Paris).)

Experimentally, the electronic density of states has been investigated by X-ray emission (XES) and absorption (XAS) spectroscopy. **Figure 4** shows an example of experimentally obtained results for the aluminum partial density of states in pure faced-centered cubic (f.c.c.) aluminum and in icosahedral AlCuFe quasicrystals. Clearly, the aluminum valence band is shifted toward higher binding energies, an effect caused by the strong *sp–d* hybridization in the alloy. In addition, the observed conduction band is

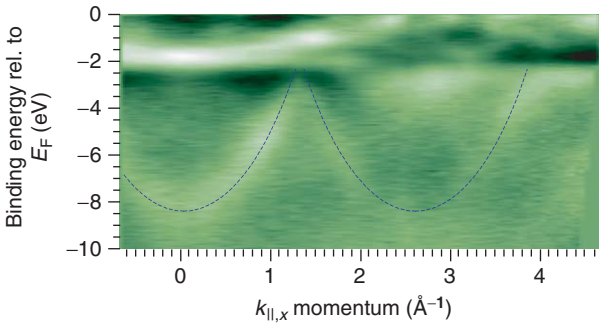


Figure 5 Intensity plot of angle-resolved photoemission spectra for AlNiCo valence electrons for photon energy 95 eV, obtained after mapping the angle into momentum $k_{\|,x}$. The dashed line indicates the parabolic band-like character of the *sp*-derived states. (Reproduced with permission from Rotenberg E, Theis W, Horn K, and Gille P (2000) Quasicrystalline valence bands in decagonal AlNiCo. *Nature* 406: 602–605; © Nature Publishing Group.)

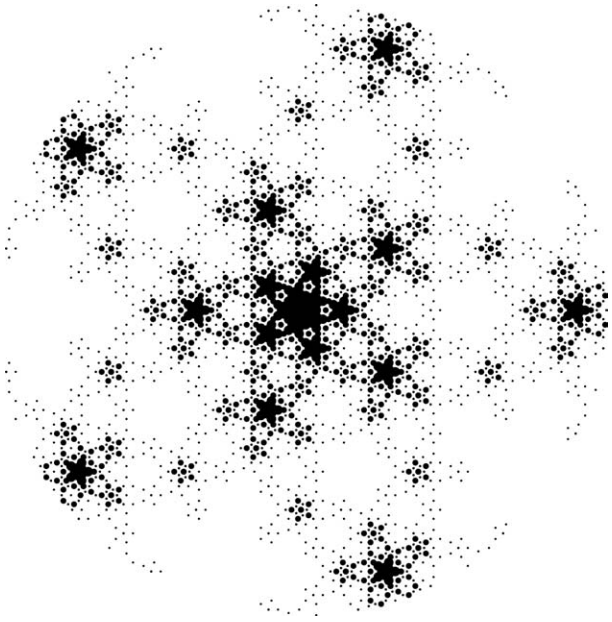


Figure 6 Sketch of the probability distribution for a multifractal wave function of a simple tight-binding model Hamiltonian on the Penrose tiling. (Reprinted with permission from Repetowicz P, Grimm U, and Schreiber M (1998) Exact eigenstates of tight-binding Hamiltonians on the Penrose tiling. *Physical Review* B 58: 13482–13490; © American Physical Society.)

much flatter in the quasicrystalline alloy than in pure aluminum. Together, this leads to a depletion of available states at the Fermi energy, as seen in **Figure 4**, and thus an experimental verification of a pronounced pseudogap at the Fermi energy, in accordance with the majority of experimental results.

However, as shown in **Figure 5**, recent angle-resolved photoemission experiments on decagonal AlCoNi revealed free-electron-like behavior of the *sp* and *d* bands, with *d* bands crossing the Fermi energy. This indicates that, even though the density of electronic states may still be very low at the Fermi energy, there might be quasi-free electronic states that can contribute to electronic transport.

While these results point to a well-defined sharp Fermi surface, previous experimental attempts to investigate the Fermi surface using the de Haas–van Alphen effect have not been very successful. There has been one published report on de Haas–van Alphen oscillations in quasicrystals, but these results have so far remained unreproduced. In other investigations, quantum oscillations have not been found. However, this does not rule out their existence, but the effects may be small due to a rather short mean free path of the electrons.

The Nature of the Electronic States

From a theoretical point of view, the particular aperiodic order present in quasicrystals must affect the nature of electronic states in such systems. On the one hand, in a conventional, periodic crystal, there are extended Bloch states, and conduction electrons can move freely through an idealized perfect crystal. This allows for electronic transport, so the system is a conductor, as long as the Fermi energy lies within a band of electronic states and not in a gap. In a three-dimensional system, this property is not destroyed by weak disorder. In a strongly disordered system, on the other hand, one expects the electrons to be localized at energetically favorable positions, and the system behaves as an insulator.

In quasicrystals, apart from the periodic direction in dodecagonal, decagonal, and octagonal phases, there are no extended Bloch waves due to the aperiodicity. There exists a competition between the aperiodicity, which tends to localize the electron, and the repetitivity of the structure (the same structural motifs appearing over and over again), which tends to have a delocalizing effect because electrons can tunnel between identical local environments. The result is that in simple model Hamiltonians one usually observes electronic states that are neither extended over the entire system nor exponentially localized, but so-called "critical" states that show a multifractal probability distribution. **Figure 6** shows an example of such a probability distribution which was obtained for a model of an electron moving on the Penrose tiling, which is a planar aperiodic tiling of decagonal symmetry.

For large classes of one-dimensional quasiperiodic systems, it can be shown in a mathematically rigorous way that the corresponding energy spectra have

the unusual property that they contain neither a continuous part nor a point part, but are purely singular continuous. Roughly, this means that the density of states is a fractal structure, such as a Cantor set, and the integrated density of states is a "devil's staircase." These rather strange spectral properties of discrete aperiodic Schrödinger operators provide another hint at the origin behind the particular behavior of quasicrystals.

However, the recent angle-resolved photoemission results mentioned above indicate the presence of quasi-free electronic states. The data, however, cannot rule out that these eigenstates are critical which might explain the weakness of the band-like emission observed in the quasicrystal when compared to that for periodic crystalline materials.

The Role of Clusters

Finally, there has been an attempt to explain the electronic properties by using a cluster-based picture of quasicrystals. This acknowledges the doubtlessly very important role of local clusters of atoms in quasicrystals. According to this approach, the quasicrystal structure is described in terms of a hierarchical sequence of clusters of atoms making up larger superclusters of clusters and so on. This hierarchical structure leads to a density of states that has spiky, self-similar features.

Whereas clusters are definitely important features of the local structure of quasicrystals, the assumption of a hierarchical cluster structure appears somewhat artificial, however, and leads to strange consequences, such as the presence of holes in the structure. Therefore, the validity of this approach appears limited, though it does predict an asymmetry of the density of states at the Fermi energy, which is in accordance with the experimental results.

Summary

The electronic structure of quasicrystals continues to be an active and controversially discussed topic of current research. While we now have a reasonable qualitative picture of some important aspects, the electronic structure is still not completely understood. Any progress is intimately linked to a better understanding of the structure formation and growth of quasicrystals, and of the theory of electronic transport in aperiodically ordered solids.

However, current experiments lead to an improved knowledge of the actual structure and the amount and type of inherent disorder, and it is conceivable that a more coherent picture of the electronic structure of these fascinating materials will emerge in the near future.

See also: Magnetic Point Groups and Space Groups; Quasicrystals.

PACS: 61.44.Br; 71.23.Ft; 75.50.Kj; 71.23.An; 72.15. – **v**

Further Reading

Baake M and Grimm U (2002) Further reading: literature on quasicrystals. In: Suck J-B, Schreiber M, and Häussler P (eds.) *Quasicrystals – An Introduction to Structure, Physical Properties, and Applications*, pp. 539–544. Berlin: Springer.

Belin-Ferré E (2002) Electron densities of states in quasicrystals and approximants. In: Suck J-B, Schreiber M, and Häussler P (eds.) *Quasicrystals – An Introduction to Structure, Physical Properties, and Applications*, pp. 338–363. Berlin: Springer.

Berger C (1994) Electronic properties of quasicrystals: experimental. In: Hippert F and Gratias D (eds.) *Lectures on Quasicrystals*, pp. 463–504. Les Ulis: Les Editions de Physique.

Damanik D (2000) Gordon-Type arguments in the spectral theory of one-dimensional quasicrystals. In: Baake M and Moody RV (eds.) *Directions in Mathematical Quasicrystals*, pp. 277–305. Providence, RI: American Mathematical Society.

Fournee V, Belin-Ferré E, Pecheur P, Tobola J, Dankhazi Z, *et al.* (2002) Electronic structure of Al-Pd-Mn crystalline and quasicrystalline alloys. *Journal of Physics: Condensed Matter* 14: 87–102.

Grimm U and Scheffer M (2001) Incommensurate crystals and quasicrystals. In: Meyers RA (ed.) *Encyclopedia of Physical Science and Technology*, 3rd edn., vol. 7, pp. 731–750. San Diego: Academic Press.

Grimm U and Schreiber M (2003) Energy spectra and eigenstates of quasiperiodic tight-binding Hamiltonians. In: Trebin H-R (ed.) *Quasicrystals – Structure and Physical Properties*, pp. 210–235. Weinheim: Wiley-VCH.

Häussler P, Barzola-Quiquia J, Haberkern R, Madel C, Lang M, *et al.* (2003) On fundamental structure forming processes – Disordered and quasicrystalline systems. In: Trebin H-R (ed.) *Quasicrystals – Structure and Physical Properties*, pp. 289–310. Weinheim: Wiley-VCH.

Krajčí M and Hafner J (2002) Phonons and electrons in quasicrystals. In: Suck J-B, Schreiber M, and Häussler P (eds.) *Quasicrystals – An Introduction to Structure, Physical Properties, and Applications*, pp. 393–420. Berlin: Springer.

Mayou D (1994) Introduction to the theory of electronic properties of quasicrystals. In: Hippert F and Gratias D (eds.) *Lectures on Quasicrystals*, pp. 417–462. Les Ulis: Les Editions de Physique.

Sire C (1994) Properties of quasiperiodic Hamiltonians: spectra, wavefunctions and transport. In: Hippert F and Gratias D (eds.) *Lectures on Quasicrystals*, pp. 505–533. Les Ulis: Les Editions de Physique.

Solbrig H and Landauro CV (2003) Electronic transport parameters and spectral fine structure: from approximants to quasicrystals. In: Trebin H-R (ed.) *Quasicrystals – Structure and Physical Properties*, pp. 254–271. Weinheim: Wiley-VCH.

Stadnik ZM (ed.) (1999) *Physical Properties of Quasicrystals*. Berlin: Springer.

R

Radioactive Tracers

P Brisset, CEA-DRT-LIST-DETECS-SSTM, Gif sur Yvette, France
J Thereska, International Atomic Energy Agency, Vienna, Austria

Introduction

Radioactive tracers were first used in industry and environment in the 1950s. Since then, there has been a continuous expansion in their usage so that, today teams around the world are actively promoting and developing both theory and practice of radiotracer methods as a unique tool to study various types of flows generally inside opaque "chemical" reactors measuring directly *in situ* mass transfer parameters inside the reactor. As industrial application fields, chemical engineering, petrol engineering, steel industry, cement industry, water treatment plants, and all types of reactors, processes, and flows existing in such industries can be quoted.

The success of radioisotope applications is due primarily to the ability, conferred by the unique properties of radioactive materials, to collect data which cannot be obtained by other investigative techniques. Though radiotracers are continually under pressure from alternative techniques, nevertheless, they continue to make an increasingly important contribution to the better management of natural resources, industrial efficiency, and environmental conservation.

The economic and technical benefits are considerably high and already proofed and recognized by end-users; benefit-to-cost ratios between 10 : 1 and 4000 : 1 may be achieved. The cost effectiveness of radioisotope applications should be widely promulgated to encourage industrialists to take full advantage of the technology.

Radiotracer methodology generally comprises four interrelated aspects, namely, experimental design, data acquisition, data treatment, and processing and data interpretation. Radiotracer technology deals with tools to prepare adequate radiotracers, to inject them in a proper way, to measure the radioactivity in the field or through sampling in the laboratory, as well as to obtain reliable experimental data, to treat them with specific software and to extract maximum of results about the process in investigation.

This article presents the industrial applications of radioactive tracers. The medical and environmental applications are considered as out of the scope but the general principles here described are applicable.

Residence Time Distribution

The method advocated by Danckwerts in 1953 consists of writing the material balances at the entrance and exit of the studied system in order to find a characteristic relationship between the corresponding concentrations.

The principle of the tracer experiment consists in a common impulse–response method: injection of a tracer at the inlet of a system and recording the concentration–time curve at the outlet. The function obtained $C(t)$ is presented in **Figure 1**.

In case the tracer inlet pulse is not a Dirac pulse, the concentration–time curve should be recorded at the inlet as well (**Figure 2**).

The residence time distribution (RTD) function, also called exit age distribution function $E(t)$, is represented mathematically by the equation

$$E(t) = \frac{C(t)}{\int_0^\infty C(t)\,\mathrm{d}t} \quad \text{or} \quad \int_0^\infty E(t)\,\mathrm{d}t = 1$$

where $C(t)$ is the tracer concentration versus time at the outlet of the system.

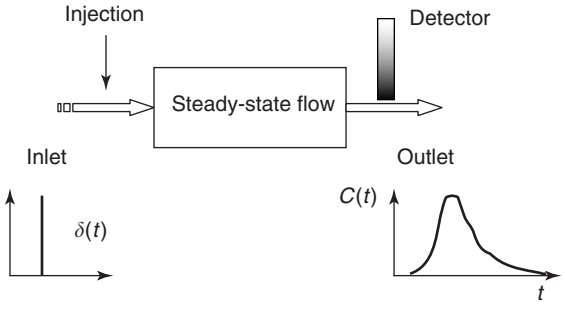


Figure 1 System impulse–response: one detector at the outlet is enough.

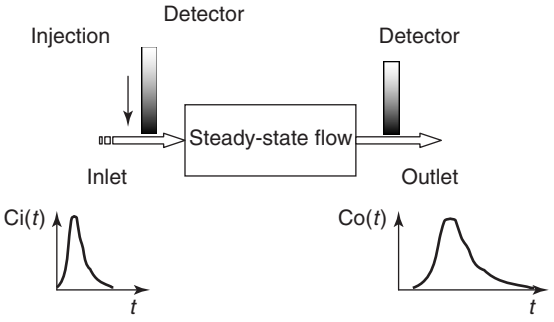


Figure 2 System response to a non-Dirac injection: two detectors at the inlet and at the outlet.

The experimental RTD, $E(t)$, is calculated from the count rate distribution at the outlet of the system $l(t)$.

There are some physical parameters which can be obtained directly by experimental RTD: $E_{\exp}(t)$ is measured as a series of numerical values. This experimental RTD is used for diagnosis of some reactor troubleshooting, for example, parallel flows, dead space, bypass, or holding-up. The direct calculation of the experimental mean residence time, the mean velocity, or tracer balance is quite useful for determination of some process parameters. More data about the hydrodynamics (flow pattern, mixing) of the system can be extracted through the modeling of the RTD curve.

Moment's Method for Simple Parameter Calculation

The moment's method is the simplest method of estimating the distribution parameters. The probability distribution with m parameters $pf(x, p_1, p_2, \ldots, p_m)$ is considered where x is a random variable. Having the empirical probability distribution, one can evaluate the estimator of the parameters by estimating the first m moments:

$$\mu_n = \int_0^\infty x^n f(x)\,\mathrm{d}x$$

The mean residence time (moment of first order (MRT)) is directly related to the flow rate Q and the effective flow volume of the system V by the relation

$$\bar{t} = \frac{V}{Q}$$

The standard deviation (SD) characterizes the mixing rate in the system (full width at half maximum (FWHM = 2SD). The higher the mixing rate, the greater the value of SD. SD equals zero in the case of lack of mixing (plug flow). For a perfect mixing system, $E(t)$ is an exponential function.

RTD Modeling

Experimental RTD is the basic information for further treatment. Throughout its modeling, the optimal parameters for process simulation and control could be determined. Modeling is realized generally by mathematical equations involving empirical or fundamental parameters, such as axial dispersion coefficients or arrangement of ideal mixers.

Evaluation of the dynamic parameters of continuous flows in vessels by optimizing the experimental RTD with the theoretical model (or RTD) is almost a common approach in field experiments (parametric approach or gray box principle). The fitting coefficient is found by using the least-squares method. Always knowing some features of the reactor performance, parametric modeling can be used to find the dynamic parameters.

Two classes of classical well-known models are mostly used: N (of equal size) fully mixed tanks in series and axial dispersion model with Peclet number Pe as a parameter of axial dispersion. In practice, however, the above ideal conditions are rarely achieved and the situation is usually somewhere between the two.

The axial dispersion model is used when the material that passes through a vessel moves along the longitudinal direction by advection as it tends to mix in the transverse section. The differential equation representing the unidirectional dispersion model is

$$\frac{\partial C(x,t)}{\partial t} = D\frac{\partial^2 C(x,t)}{\partial x^2} - u\frac{\partial C(x,t)}{\partial x}$$

where $C(x,t)$ is the concentration at a distance x at time t, D is the axial dispersion coefficient, and u is the mean velocity of advective transport.

For an instantaneous and planar injection at $t = 0$ and $x = 0$, the solution is

$$C(x,t) = \frac{M}{A\sqrt{4\pi Dt}}\,\mathrm{e}^{-(x-ut)^2/4Dt}$$

where M is the mass of a tracer injected into the cross section at the inlet.

The model parameter normally used as an index of mixing is the nondimensional Peclet number, $Pe = ux/D$ (Pe = infinite for plug flow whereas $Pe = 0$ for completely mixed flow).

Ideal stirred tanks connected in series model are frequently used to describe the systems where it is assumed that an injected tracer is immediately (in comparison to the flow rate) mixed with the entire volume of the system as a result of either mechanical mixing or some circulation (**Figure 3**).

In such a case, the concentration of the tracer at the inlet and the outlet is equal. Then the

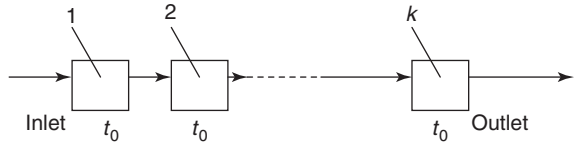


Figure 3 Scheme of ideal stirred mixing tanks in series model.

time–concentration function for the outlet is

$$\frac{dC_0(t)}{dt} = \frac{1}{\bar{t}}[C_0(t) - C_i(t)]$$

As the $C_i(t)$ function is usually a Dirac pulse $\delta(t)$, normalized $C_0(t)$ represents the RTD which in the time domain is equivalent to

$$E(t) = \frac{1}{\bar{t}} \exp\left(-\frac{1}{\bar{t}}\right)$$

It is common, in practice, to present the system as an arrangement of perfect mixers connected in series. For such a model, $E(t)$ is

$$E(t) = \frac{1}{(k-1)!}\frac{1}{t_0}\left(\frac{t}{t_0}\right)^{k-1} \exp\left(-\frac{t}{\bar{t}}\right)$$

where t_0 is the MRT for a single mixer, k is the number of mixers. The total MRT is then $\bar{t} = kt_0$.

In order to compare $E(t)$ curves for different flow conditions and mixing efficiency, normalization to dimensionless time θ is performed: $\theta = t/\bar{t}$. Then the equation takes the form

$$E(t) = \frac{k^k}{(k-1)!}\theta^{k-1} \exp(-k\theta)$$

where k is infinite for plug flow and is equal to 1 for completely mixed flow.

For continuous process vessels with high dispersion, the best model is the cascade of mixers-in-series. When dispersion is low, either the axial dispersion or the cascade of mixers follows the material transfer well. In the latter case, both models are equivalent (Villermaux): $Pe = 2^*(N-1)$.

The cascade of mixers in series model describes quite well all the simple flows with partial dispersion. Moreover, the cascade of mixers-in-series model offers the possibility to build up more complicated models, combining the mixer units in various arrangements as well as adding into them several cells or zones with different flow regimes, that is, plug flow, stagnant zone, dead volume, bypass, recirculation, etc.

Radioisotopes Used as Radiotracers

Theoretically, it can be said that all the radio nuclides existing in the Mendeleyev classification can be used as tracers. Practically, one is limited to γ emitters able to be industrially produced in nuclear reactors.

Radioisotope generators are also very important in tracer work. Such generators use the filiation chain phenomenon. That means the radioactive father decays through particle emission (e.g., β) thus producing a son which is also radioactive of shorter half-life. The father is chemically fixed in a chromatographic column. An appropriate fluid percolating in the column recovers the son (elution). The result of this operation is a radioactive solution containing the son only. There are two easily available generators for remote tracer experiments: ^{99}Mo–^{99m}Tc and ^{113}Sn–^{113m}In. The tracer is obtained in an aqueous form by the elution of the generator. An appropriate chemical treatment will give the possibility of using the radio nuclide to label solid particles (by adsorption) or liquid phase (aqueous or organic).

Some other generators also exist, such as ^{137}Cs–^{137m}Ba for an aqueous phase or ^{81}Rb–^{81m}Kr for a gas phase (**Table 1**).

Factors Affecting RTD Measurement

The formulation of RTD and its utilization for system analysis are well established. Accurate RTD formulation can be affected by tracer mixing length and detector response.

Tracer Mixing Length

Good mixing of tracer into the flowing system is a precondition for formulating the RTD.

The mixing length in-duct flows is defined as the distance beyond which the tracer concentration in the cross section is almost constant. An accurate way to determine the mixing length is the use of a computational fluid dynamics (CFD) code to model both the bulk flow and tracer injection.

There are a few theoretical formulas to calculate mixing length for various velocity and turbulent diffusion coefficient profiles as a function of injection configuration (central injection and annular injection). These formulas are known to underestimate, by a wide margin, the experimentally observed mixing lengths, probably because the simulation of the injection is too simplistic. Values for good mixing length determined experimentally in a straight pipe of circular cross section with central injection are about twice as large as theoretical values. The main reason for this discrepancy is the difference between actual and postulated flow conditions. These formulas should therefore be used with some caution.

In general, good tracer mixing requires as many as 200 pipe diameters to achieve. It is often not possible

Table 1 The most commonly used radiotracers in industry

Isotope	*Half-life*	*Gamma energy* MeV (%)	*Chemical form*	*Tracing of phase*
Bromine 82	36 h	0.55 (70) 1.32 (27)	Ammonium bromide *p*-dibromo-benzene	Aqueous Organic
Lanthanum 140	40 h	1.16 (95) 0.92 (10) 0.82 (27) 2.54 (27)	Chloride or oxide	Solid adsorbed or particles
Gold 198	2.7 d	0.41 (99)	Chloride or glass	Solid adsorbed or particles
Mercury 203	46.6 d	0.28 (86)	Mercury metal	Mercury
Iodine 131	8 d	0.36 (80) 0.64 (9)	KCl or NaCl Iodobenzen	Aqueous organic
Technetium 99m	6 h	0.14 (90)	Sodium technetate	Aqueous organic or solid
Krypton 79	35 h	0.51 (15)	Krypton	Gas
Argon 41	1.8 h	1.29 (99)	Argon	Gas
Indium 113m	1.6 h	0.39 (65)	Chloride complexed	Solid adsorbed aqueous organic

to inject the tracer at such a distance upstream to the measurement section. Every singularity in the pipe promotes turbulence that tends to decrease good mixing length. It is therefore useful to be able to reduce the length by using appropriate devices.

A substantial reduction of mixing length can be obtained by injecting the tracer through multiple orifices uniformly distributed on the pipe wall or (if possible) inside the pipe.

Injecting the tracer currently at a velocity much larger than the bulk flow velocity induces high mixing at the end of the jet. Reduction in good mixing length depends on the number and momentum of the jet, and of its angle with respect to the main flow direction. Little quantitative data are available on these effects. However, a simple jet arrangement can bring a 30% reduction with respect to a single central injection point.

Incorporating obstacles within the pipe, in the vicinity of tracer injection, produces turbulence that enhances mixing and reduces good mixing length. As an example, injecting the tracer through three triangular plates at an angle of 40° with respect to the main flow direction reduces the mixing length by one-third with respect to a central single injection point.

If a tracer is injected upstream a pump or a turbine, the mixing length is considerably reduced. Available information indicates that centrifugal pumps reduce the mixing length by $\sim$100 pipe diameters.

Detector Response

An important point dealing with RTD formulation is "what is really seen by a detector," since a wall radiation detector does not perform local measurement but collects a certain amount of information within a solid angle called the volume of sensitivity. Since the radiotracers commonly used are γ emitters, the photons they emit undergo multiple random interactions, that is, with the fluid itself, walls, screens, and collimator, until they reach the detection probe crystal scintillator and photomultiplier.

The importance of these interactions is a function of the energy of the emitted photons and the nature (density and chemical composition) of the fluid and materials. The link between the tracer concentration and detector signals is therefore not direct, the problem lying mainly in the correct representation of interactions.

A relatively simple method to calculate the detection volume is the Monte Carlo method. The basic principle is to choose randomly the initial position, energy, direction, and free path of a photon. Its new position is then calculated. If the photon has not left the system, it is going to interact with the surrounding matter. The major types of interactions to be considered are photoelectric effect, Raleigh effect, Compton effect, and pair creation. The probability of these processes is a function of the energy of the incident photon. The type of interaction is randomly chosen and the procedure is repeated until the photon

either leaves the system or is absorbed by any material of the simulated system. When a sufficiently large number of photons have been treated in this way, it is possible to build statistics in terms of energies of detected photons and numbers, which in turn can be translated into count numbers. Thus, the Monte Carlo simulation provides the detector response.

Radiation Safety

Radiotracer technology is one of the many beneficial applications of ionizing radiation that is used around the world. To ensure that people are protected from the harmful effects of radiation, such applications must comply with the international basic safety standards (BSS) or equivalent national regulations.

Any work with radioactive materials will normally require authorization from the relevant national regulatory authority. The authorized person or organization will have the prime responsibility to ensure that the radioactive materials are used safely and in compliance with the relevant regulations and standards.

Integration of RTD Tracing with CFD Simulation for Industrial Process Visualization and Optimization

The RTD method has been continuously developed and used. The treatment of RTD curves for extracting important parameters of industrial processes has achieved a good standard. Efficient RTD software was validated for modeling of various chemical engineering reactors, but the RTD method still remains a global approach. An RTD systemic analysis requires the choice of a model, which is often semi-empirical and rather idealized (combination of perfect mixers, dead volumes, etc.). There are some situations in which the RTD approach cannot be applied, that is, nonlinear systems.

The industry is looking for more predictive techniques. The CFD method is a fine and predictive analysis, which provides nice spatial pictures of the insight of a process, such as flow patterns and velocity map. A CFD simulation could be easily extrapolated to other flow conditions once validated. Owing to lack of physical experimental data, the CFD calculation provides qualitative results only, especially in systems with strong interaction of hydrodynamics with physicochemical reactions. This is the reason why CFD models have to be verified and validated by experimental tracer RTD results.

The RTD–CFD interaction is on both sides. The CFD can also be used to complement the information obtained from the RTD systemic approach. CFD provides data that can quantify the RTD systemic model, which means the CFD model can "degenerate" into a more quantitative RTD systemic analysis, providing more comprehensive results for chemical engineers. In fact, these two approaches, experimental and numerical, are complementary to each other. The RTD systemic approach detects and characterizes the main features of the flow (mixing and recirculation) while CFD enables one to locate them. The trend is to combine the RTD experimental and CFD numerical approaches in an integrated CFD–RTD method, which provides reliable quantitative results for industrial complex processes.

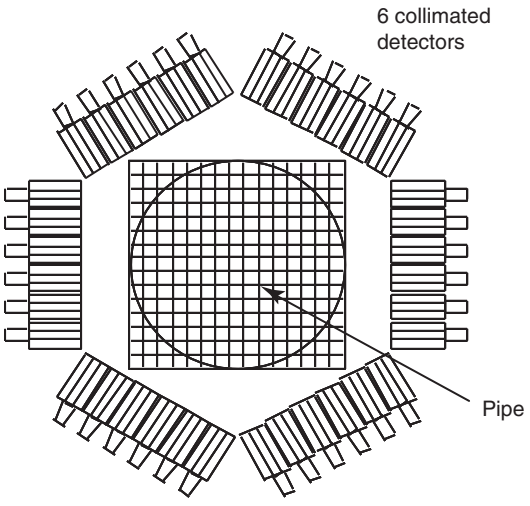


Figure 4 SPECT principle (Legoupil).

Figure 5 Example of a SPECT reconstruction of a tracer in a pipe.

Radiotracer Imaging Techniques for Industrial Process Visualization

New development is expected in introducing new radiotracers (two-dimensional (2D) and 3D imaging techniques) for localization and visualization of flow patterns in multiphase systems.

Emission tomography is the last step in research and development in radiotracer methodology. The real time imaging techniques for flow pattern visualization inside vessels are important for investigating multiphase flow systems. Industrial process gamma tomography is quite similar to nuclear medicine imaging.

Among the various techniques, single photon emission computed tomography (SPECT) shows good results on an industrial scale (**Figure 4**).

Emission tomography provides 2D maps of the count rates which, when properly interpreted, yield the radiotracer instantaneous concentration field as shown in **Figure 5**.

See also: Radioactivity.

PACS: 47.85. – g; 92.20.Td; 87.58.Fg; 89.20.Kk

Further Reading

IAEA (1990) *Guide book on Radio-Isotope Tracers in Industry*, Technical reports series no. 316. Vienna: IAEA.

IAEA (2001) Radiotracer technology as applied to industry. Final report of a co-ordinated research project, 1997–2000. IAEA-TECDOC-1262, Vienna, Austria.

Legoupil S, Berne Ph, Blet V, and Brisset P (2001) *Industrial Process SPECT Tomography*, 3rd International Conference on Chemical Engineering, ENSIC Nancy, France, Invited lecture.

Legoupil S, Pascal G, Bloyet D, and Chambellan D (1996) An experimental single photon emission computed tomograph method for 2D fluid flow analysis. *Journal of Applied Radiation and Isotopes* 48: 1507–1514.

Thyn J, Zitny R, Kluson J, and Cechak T (2000) *Analysis and Diagnositcs of Industrial Processes by Radiotracers and Radioisotope Sealed Sources*. Vydavatelstvi CVUT, Praha, ISBN 80-01-02241-2.

Villermaux J (1995) *Genie de la reaction chimique – conception et fonctionnement des reacteurs*, 2nd edn. TEC & DOC Lavoisier, ISBN 2-85206-759-5.

Radioactivity

W Greiner and D N Poenaru, J. W. Goethe Universität, Frankfurt, Germany

Introduction

All substances are made of atoms with electrons orbiting around the nucleus consisting of two kinds of nucleons: protons (p's) and neutrons (n's). Inside the nucleus, p's and n's are held together by a strong interaction overcoming the electrostatic repulsion between the positively charged p's. The first information about the atomic nucleus was obtained at the beginning of the twentieth century by studying radioactivity, and the investigations at the origin of all developments led to nuclear physics, particle physics, radiochemistry as well as many applications in medicine, biology, and industry. Radioactivity, a spontaneous disintegration of nuclei, is a natural phenomenon taking place since the very beginning of the universe. Over 60 radionuclides can be found in nature. Henri Becquerel discovered, in 1896, an invisible radiation emitted by uranium. Two other emitters, polonium and radium, were added after two years by Marie and Pierre Curie, who coined the name radioactivity. By using radium, it was possible to devise the first medical application of radioactivity. Artificial radioactivity was discovered by Irène and F Joliot-Curie in the 1930s. They obtained a new radioactive element ^{30}P by bombarding a stable nucleus ^{27}Al with α-particles. Many radionuclides for biomedical studies were produced after the invention of the cyclotron by E Lawrence in 1930.

A chemical transformation of the emitter into a different element (the daughter) takes place during this process, as shown for the first time by E Rutherford and F Soddy in 1902. An exception is the γ-decay, in which an upper excited energy level is de-excited toward a lower one of the same nucleus. The detection of radiations is based on their interactions with matter, leading mainly to ionizations.

For a long period of time, only three kinds of nuclear decay modes (α, β, and γ) were known. They illustrate three of the fundamental interactions in nature: strong, weak, and electromagnetic. Spontaneous fission became popular soon after its discovery in 1940 by K Petrzhak and G Flerov owing to both military and peaceful applications of the neutron-induced fission discovered in 1939 by O Hahn, Lise Meitner, and F Strassmann, employing the large amount of energy released during the process. Seven Nobel prizes for physics and chemistry were received (in 1903, 1908, 1911, 1921, 1935, 1939, and 1944) by almost all scientists mentioned above. At present, many other kinds of radioactivities are known, for example,

heavy ion radioactivities, proton decays, various beta-delayed decay modes (p, 2p, 3p, n, 2n, 3n, 4n, d, t, α, f), fission isomers, particle-accompanied fission. The height of the potential barrier, mainly of Coulomb nature, for emission of the charged particles is much higher than the observed kinetic energy of the emitted particles. The spontaneous decay can only be explained by quantum mechanical tunneling.

There are many applications of radioactivity. Nuclei are identified by measuring their characteristic radioactive transitions, for example, the energy spectrum of α-particles or the rotational γ-spectra, etc. Radionuclide tracers are used to study reaction mechanisms and transport processes in chemistry, biochemistry, and agriculture. Dating with ^{14}C and other nuclides became a routine in archeology and geology. Since 1921, radioelements have been employed to study the metabolism of plants. The radionuclides contributed essentially to the spectacular progress of biomedicine and genetics. The positron emission tomography (PET) is presently one of the best imaging devices in life sciences. Therapeutic radionuclides are routinely used. Other applications are: sterilization of medical instruments, industrial γ-ray radiography, smoke detectors, electric batteries for satellites (based on radioactive ^{239}Pu, ^{60}Co, or ^{90}Sr sources), improving the properties of materials or producing new materials, etc.

Binding Energy

The nuclear masses in atomic units are very close to the mass number $A = Z + N$, where Z is the number of protons and N the number of neutrons. This is the reason why usually the mass excess (or the mass defect) Δ is used:

$$\Delta(A, Z) = [M(A, Z) - A]c^2$$

where c is the speed of light. In the following, it is assumed that masses are expressed in units of energy, hence the factor c^2 is dropped. The binding energy accounts for the stability of a nucleus:

$$B(A, Z) = ZM_{\mathrm{p}} + NM_{\mathrm{n}} - M(A, Z)$$

where M_{p} and M_{n} are, respectively, the proton and neutron masses in units of energy.

The energy needed to remove a nucleon,

$$S_{\mathrm{n}}(A, Z) = M(A-1, Z) + M_{\mathrm{n}} - M(A, Z)$$

or

$$S_{\mathrm{p}}(A, Z) = M(A-1, Z-1) + M_{\mathrm{p}} - M(A, Z)$$

is called separation energy. The neutron drip line is defined by $S_{\mathrm{n}} = 0$. A negative value, $S_{\mathrm{p}} < 0$, characterizes a proton radioactive nucleus.

Decay Law

In many decay modes, a parent nucleus ^{A}Z, with mass and atomic numbers A and Z, in its ground state, is split into a daughter (heavy fragment) $^{A_1}Z_1$ and an emitted (light fragment) $^{A_2}Z_2$:

$$^{A}Z \rightarrow {}^{A_1}Z_1 + {}^{A_2}Z_2$$

with conservation of the hadronic numbers, $A = A_1 + A_2, \quad Z = Z_1 + Z_2$. This binary decay process is energetically allowed if the released energy (or the Q-value)

$$Q = M - M_1 - M_2$$

is a positive quantity. The mass excess is transformed into energy according to the Einstein formula $E = mc^2$. In the absence of any energy loss for fragment deformation and excitation, as in cold-fission phenomena or in α-decay, the total kinetic energy (TKE) is equal to the Q-value and is divided between the particles in inverse proportion with their masses, as required by the conservation of linear momentum, hence for the small fragment,

$$E_{k2} = QA_1/A$$

which can be experimentally determined.

Another important measurable quantity is the partial half-life, T, of the parent nucleus against this decay mode. The nucleus is metastable or, in other words, it exhibits a radioactivity by spontaneous emission of the particle $^{A_2}Z_2$, if both the Q-value and the potential barrier height are positive quantities: $Q > 0, E_{\mathrm{b}} > 0$. Only temporarily, the two fragments are held together by the nuclear forces. There is a finite probability, P, per unit time for penetration through the barrier by quantum-mechanical tunneling effect, as it was shown in 1928 by Gamow and independently by Condon and Gurney in the case of α-decay. The activity $\mathscr{A}$ (number of disintegrations per second) of a radioactive source is measured in the SI system in terms of becquerel; 1 Bq corresponds to the decay of one atom per second. The traditional unit is 1 curie (Ci) $= 3.7 \times 10^{10}$ decays per second. The probability of decay may be expressed in terms of its half-life, T, the time in which only half of the initial quantity remains. By assuming a disintegration rate independent of the past history of the individual decaying nuclei, the variation of the number $\mathscr{N}(t)$ of radioactive nuclei at a time t during the time dt must be proportional to $\mathscr{N}(t)$ and to dt : d$\mathscr{N} = -\lambda \mathscr{N}dt$. After integration, one obtains the exponential decay law:

$$\mathscr{N}(t) = \mathscr{N}(0)\, \exp(-t/\tau)$$

where the time constant $\tau = 1/\lambda$ defines the lifetime of the parent nucleus and λ is the disintegration constant.

The half-life, T, is defined by $\mathcal{N}(t) = \mathcal{N}(0)/2$, hence $T = \tau \ln 2 = 0.683147/\lambda$.

The quasiclassical Wentzel–Kramers–Brillouin (WKB) approximation allows one to calculate the penetrability in a one-dimensional case as

$$P = [1 + \exp(K)]^{-1}$$
$$K = \frac{2}{\hbar}\int_{q1}^{q2} \{2B_q(q)[E(q) - Q]\}^{1/2}\, \mathrm{d}q$$

where K is the action integral, $E(q)$ is the potential barrier, q_1, q_2 defined by $E(q_1) = E(q_2) = Q$ are the classical turning points, and B_q is the nuclear inertia.

In the presence of two decay modes of the same nucleus with half-lives T_1 and T_2, the total half-life is $T = T_1 \cdot T_2/(T_1 + T_2)$. The experimentally accessible range of nuclear lifetimes covers more than 45 orders of magnitude. Specific methods of measurements have been designed for different subdomains.

Decay Modes

Many nuclei are unstable against emission of: (1) protons or heavier charged particles, (2) electrons or positrons, or/and (3) photons. The first type (p- and α-decay, heavy ion radioactivities, spontaneous fission) involves the breakup of a nucleus into two fragments. When protons pass through matter, they have a low penetrating power or a short range because they ionize atoms strongly. A sheet of paper is enough to stop them. In the second type, one kind of nucleon transforms into the other via the weak interaction. With a mass smaller by three orders of magnitude than that of the nucleons and the charge of one unit, β-particles have a medium penetrating power; they are stopped by a sheet of aluminum. A photon is emitted when a state of a given nucleus passes from a higher to a lower excitation or to the ground state. The γ-rays do not directly ionize the atoms; they have a high penetrating power. A thick sheet of metal such as lead is necessary to attenuate them significantly.

Alpha Decay

An α-particle is the doubly magic ^{4}He nucleus. Nuclei with closed shells (or magic numbers $N, Z = 2, 8, 20, 28, 50, 82, 126$) are relatively stable. One of the strongest α-decay is that of ^{212}Po, leading to two double-magic nuclei ^{208}Pb and ^{4}He:

$$^{212}_{84}\mathrm{Po}_{128} \rightarrow {}^{4}_{2}\mathrm{He}_2 + {}^{208}_{82}\mathrm{Pb}_{126}$$

During this process, an energy $Q_\alpha = 8.954$ MeV is released in a form of kinetic energy of the α-particle

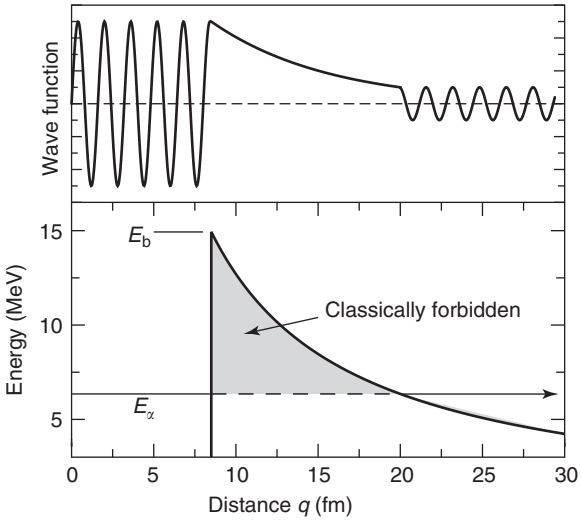


Figure 1 Quantum tunneling penetration of the potential barrier by an α-particle.

$E_\alpha = (208/212)Q_\alpha = 8.785$ MeV and a small (169 keV) recoil energy of the daughter. Both this kinetic energy and the half-life $T_{1/2} = 0.298$ μs are measurable quantities. In order to display not only the mass number, $A = 208$, and the chemical symbol, Pu, but also the number of protons, $Z = 82$, and neutrons, $N = 126$, it is written as $^{208}_{82}\mathrm{Pb}_{126}$.

According to classical physics, a particle of energy E_α less than the height E_b of a barrier could not penetrate; the region inside the barrier is classically forbidden (see **Figure 1** at the bottom). The particle is supposed to be preformed at the nuclear surface. The wave function associated with a free particle must be continuous at the barrier and will show an exponential decay inside the barrier. It must also be continuous on the far side of the barrier, so there is a finite probability that the particle will tunnel through the barrier.

There are more than 400 known α-emitters, mostly proton-rich nuclei, as shown in **Figure 2**. The neutron-rich fission fragments of ^{252}Cf shown in the same figure do not decay by α-emission because they have a negative Q_α.

Beta Decay

A beta-particle can be either an electron, e^-, emitted when a neutron is transformed into a proton, or a positron, e^+, when a proton is transformed into a neutron. A neutrino (ν) or antineutrino ($\overline{\nu}$) is simultaneously emitted. In the electron-capture (EC) process, an orbital electron is captured by the nucleus and in the atom, this de-excitation produces an X-ray. As for β^+-decay, the daughter has a lower atomic number, $Z - 1$. EC always coexists in competition

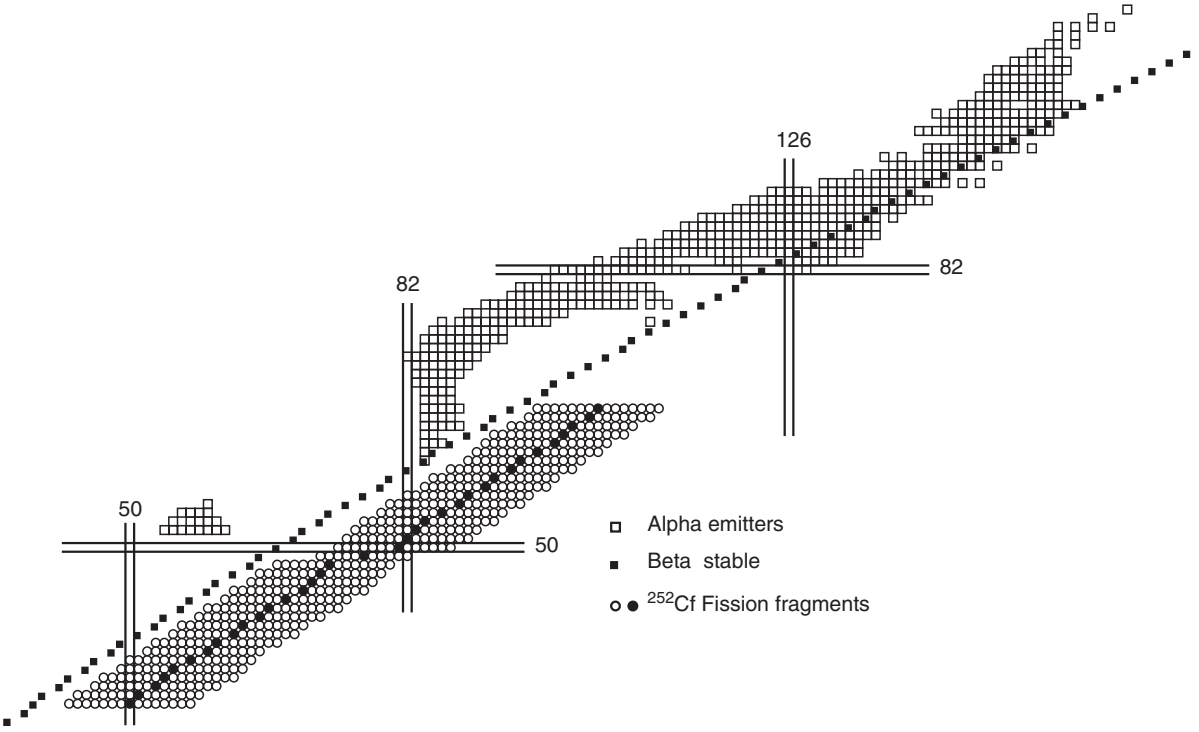


Figure 2 Nuclear chart of α-emitters and the fission fragments of ^{252}Cf. Fission fragments are neutron-rich nuclei; because of $Q_\alpha < 0$, they are stable against α-decay. (Reproduced with permission from Poenaru DN, Greiner W, Hamilton JH, and Ramayya AV (2001) *Journal of Physics G: Nuclear and Particle Physics* 27: L19–L28; © IOP Publishing Limited.)

with positron emission; that is why the resulting half-life β^+/EC is given.

For the three kinds of β-decay – β^-, β^+, and EC – one has

$$^AZ \rightarrow {}^A(Z+1) + e^- + \overline{\nu}_e$$
$$^AZ \rightarrow {}^A(Z-1) + e^+ + \nu_e$$
$$^AZ + e^- \rightarrow {}^A(Z-1) + \nu_e$$

As an example, the neutron-rich nucleus ^{212}Bi is a β^- and α-emitter with $Q_{\beta^-} = 2.254$ MeV, $T_{\beta^-} = 94.52$ min (branching ratio 64%) and $Q_\alpha = 6.207$ MeV($E_\alpha = 6.090$ MeV) $T_\alpha = 168.5$ min (branching ratio 36%), leading to a total half-life of 60.55 min.

The nucleus ^{151}Eu decays by electron conversion with $Q_{EC} = 2.261$ MeV and a half-life of $T_{EC} = 35.8$ y. Other examples are given in **Figure 3**, where a chain of α and β^- decays of one of the three natural radioactive families, ^{235}U, are shown together with two rare decay modes of ^{14}C and ^{24}Ne. The other two natural families start with α-decay of ^{232}Th and ^{238}U.

The very broad energy spectrum of electrons emitted in the β-decay, explained by E Fermi in 1934, is typical for the kinematics of three bodies in the final state. In this way, a need for existence of another particle was seen by W Pauli who introduced the neutrino. It is not easy to detect neutrinos because they interact very weakly with matter. The main sources of low-energy neutrinos are: β^- activity of fission fragments in high-power nuclear reactors, solar thermonuclear fusion processes (p–p chain and CNO cycle), stellar collapse leading to neutron stars in supernova explosions, and the capture of atomic electrons in strong sources of ^{51}Cr or other isotopes.

There are two-step processes, beta-delayed particle emissions discovered by V Karnaukhov in 1962, in which, after the population of an excited state by the β-decay of a precursor, a neutral (n, 2n, 3n, 4n) or a charged particle (p, 2p, 3p, d, t, α, light fission fragment, etc.) is rapidly emitted. In the vicinity of the drip line, the Q_β-values are rather high; when the corresponding daughters possess low separation energies for hadrons or clusters of nucleons, a beta-delayed particle emission takes place, usually from an excited state. In this way, the particle emission is much faster; the half-life is determined by the preceding β-decay of the precursor.

Gamma Decay

An excited nuclear state decays to a lower one or to the ground state of the same nucleus by emission of electromagnetic radiations, such as X-rays and light,

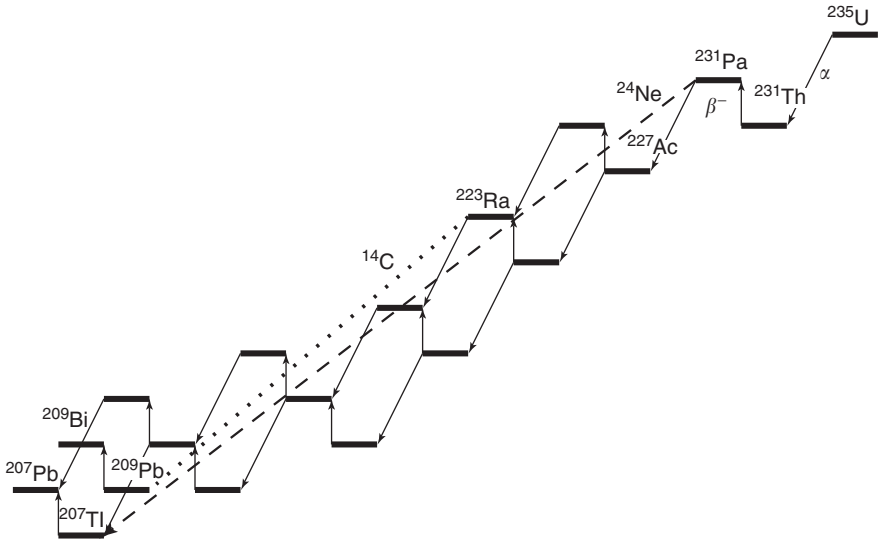


Figure 3 One of the three natural radioactive families. Starting with ^{235}U, there are several sequences of α and β^- disintegrations ending up with a stable nucleus ^{207}Pb. From ^{231}Pa to the same ^{207}Pb, one needs only one ^{24}Ne radioactivity (dashed line) followed by one β^- decay. Alternatively, from ^{223}Ra one can reach the stable nucleus ^{209}Bi with one ^{14}C decay (dotted line) followed by one β^- decay.

but with shorter wavelengths. The energy of the emitted photon is equal to the difference of the two levels. The half-lives are usually very short, under 1 ns, except for the decay of isomeric states which have a large spin and low excitation energy, hence the emission of γ-rays is strongly hindered. They decay mainly by electron conversion with half-lives in the range 10^{-10} s to 5000 yr. In an electron conversion process, the excited state decays by transferring its energy to an orbital electron which becomes free. Unlike in beta-decay, there is no change of Z or N. The isomeric or metastable states are indicated with a superscript m, as in ^{166m}Ho which decays by β^- emission ($T = 1200$ yr).

Important characteristics of a gamma transition are: the energy, the character (electric, E, or magnetic, M), and the multipolarity (dipole, $L = 1$, quadrupole, $L = 2$, etc). The order of magnitude of transition probabilities can be estimated by using the Weisskopf formula obtained in the assumption of a single nucleonic motion. For the decay constant of electric quadrupole and magnetic dipole transitions, one has: $\lambda(E2) = 7.3 \times 10^7\ A^{4/3}E^5$, $\lambda(M1) = 5.6 \times 10^{13}E^3$, where λ is obtained in s^{-1}, if E is given in MeV. When the measured $E2$ transition is much stronger than that given by the Weisskopf estimate, one may conclude that many nucleons are involved; in other words, the decaying nuclear level has a collective character.

In a level scheme, as that of $^{10}_5B_5$ in **Figure 4**, the energy levels are represented as horizontal lines, the heavy one for the ground state (gs), and transitions

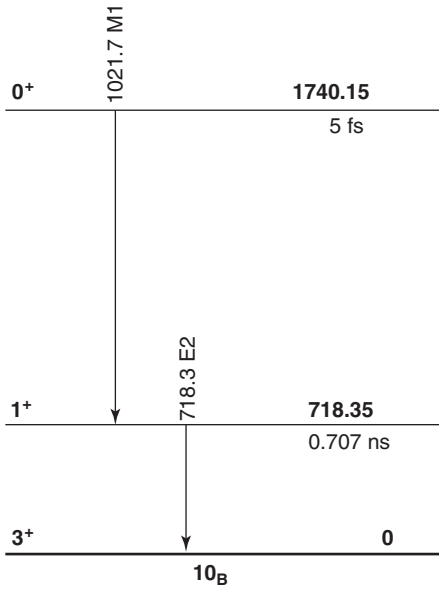


Figure 4 The energy level scheme and γ-transitions of ^{10}B.

by vertical arrows. Level energies expressed in keV are written in bold type near the right end of a level above the line. The half-life of each level is given under the line. The spin and parity are located near the left end. The energy and the multipolarity of each transition is written vertically on top of the corresponding arrow. The upper excited level at 1740.15 keV with a half-life of 5 fs and a spin and parity 0^+ decays 100% to the lower excited level at 718.35 keV (half-life 0.707 ns) with a spin and parity 1^+ by emitting M1 γ-rays with the energy

1740.15 – 718.35 = 1021.7 keV. In its turn, this level is 100% de-excited to the ground state 3^+ by an $E2$ transition of 718.3 keV.

The β^-decay of the 0^+ nucleus ^{10}Be with $Q_{\beta^-} = 555.9$ keV and a half-life of 1.51×10^6 y feeds the gs of ^{10}B. Its excited states can be fed by EC of 0^+ ^{10}C with $Q_{EC} = 3647.81$ keV and a half-life of 19.25 s in a proportion of 98.53% for the lower level and 1.47% for the upper one. The level schemes of heavy nuclei are very complex, with a lot of levels and γ-transitions.

Spontaneous Fission

Intermediate mass nuclei (fission fragments) are emitted during spontaneous fission. Fission fragments are usually deformed and excited, hence the process continues with neutron evaporation as well as β and γ decays. A neutron multiplicity of the order of 2 to 3 neutrons per fission act in fissile materials is very important for making possible the self-sustained chain reaction to produce energy in nuclear power plants. The nuclear fission is the most complex decay process, in which more than 500 different nuclides are produced. Fission characteristics not only provide basic information on this large amplitude collective motion, but they are also important for certain applications, such as nuclear power plants, safeguards, and nuclear incineration. The TKE of fission fragments increases with the fissility parameter from ~160 MeV for ^{232}Th to 270 MeV for the heaviest elements. The half-lives against spontaneous fission of even–even transuranium nuclei decrease with increasing Z from $\sim 10^{15}$ yr for U($Z = 92$) to ~1 ms for Hs ($Z = 108$). The frequently used source of ^{252}Cf decays by spontaneous fission with a partial half-life of $T_f = 85.5$ yr (branching ratio 3%) and α-decay $T_\alpha = 2.73$ yr (branching ratio 97%), leading to a total half-life of 2.64 yr.

In the usual mechanism of fission, a significant part (~25–35 MeV) of the released energy Q is used to deform and excite the fragments, which subsequently cool down by neutron and γ-ray emissions; hence the total kinetic energy of the fragments, TKE, is always smaller than Q. Since 1981, a new mechanism has been experimentally observed by C Signarbieux *et al*. – cold fission characterized by a very high TKE, practically exhausting the Q-value, and a compact scission configuration. Experimental data have been collected in two regions of nuclei: (1) thermal neutron-induced fission on some targets, such as ^{233,235}U, ^{238}Np, ^{239,241}Pu, ^{245}Cm, and the spontaneous fission of ^{252}Cf, and (2) the bimodal spontaneous mass-symmetrical fission of ^{258}Fm, ^{259,260}Md, ^{258,262}No, and ^{260}104 discovered by K Hulet *et al*. in 1986. The yield of the cold fission mechanism is comparable to that of the usual fission events in the latter region, but it is much lower (~5–6 orders of magnitude) in the former. An example of a cold fission process of ^{234}U is given in **Figure 5** in comparison with α-decay (the strongest decay mode of this nucleus) and two heavy-particle radioactivities (^{24}Ne and ^{28}Mg).

The particle-accompanied fission (or ternary fission) was observed both in neutron-induced and spontaneous fission. It was discovered in 1946 by L Alvarez and independently by T San-Tsiang *et al*. Several such processes, in which the charged particle is a proton, deuteron, triton, $^{3-6,8}$He, $^{6-11}$Li, $^{7-14}$Be, $^{10-17}$B, $^{13-18}$C, $^{15-20}$N, $^{15-22}$O, and other heavier isotopes of F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, and even Ca, have been detected. Some of them are shown in **Figure 6**.

Fissioning isomers are the first superdeformed states experimentally determined. A spontaneous fission activity of ^{242}Am with ~14 ms half-life was discovered in 1961 by S Polikanov *et al*. during the irradiation of U and Pu targets with ^{16}O and ^{22}Ne ions. Many other isomers, with half-lives in the range 5 ps to 14 ms, have been found in the region of heavy elements with $92 \leqslant Z \leqslant 97, 141 \leqslant N \leqslant 151$. The phenomenon was explained on the basis of the Strutinsky's double-humped fission barrier according to which these states are shape-isomers, characterized by a very large prolate distortion (major-to-minor semi-axis ratio ~2:1).

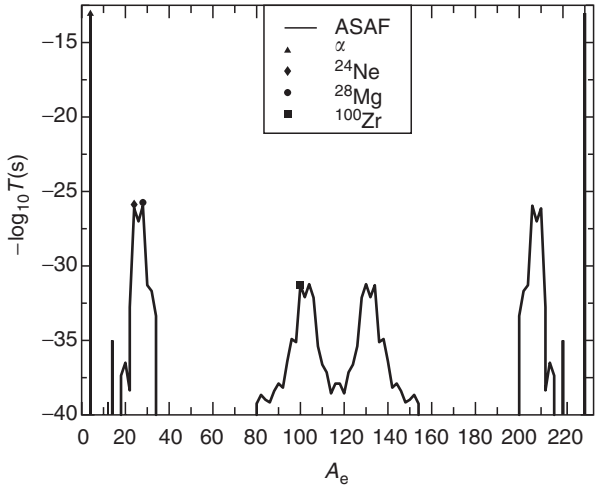


Figure 5 Lifetime spectra calculated with ASAF model for three groups of decay modes of ^{234}U: α-decay; heavy ion radioactivities (^{24}Ne and ^{28}Mg) and cold fission with light fragment ^{100}Zr. Except for CF for which T was not determined, the experimental points are in good agreement with calculations. (Reproduced with permission from Poenaru DN and Greiner W (1996) *Nuclear Decay Modes*. ch. 6, pp. 275–336. Bristol: IOP; © IOP Publishing Limited.)

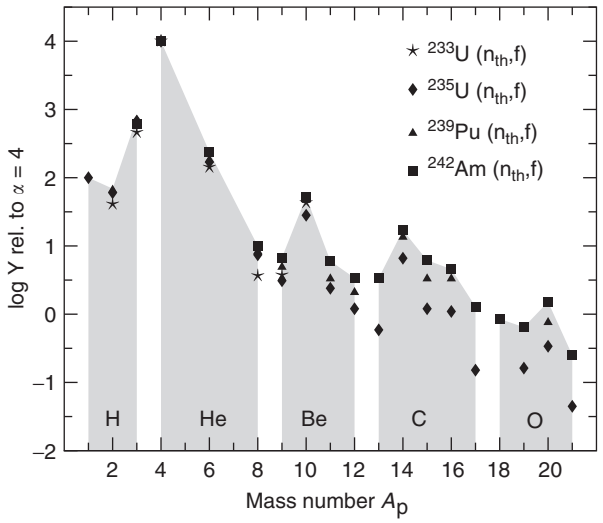


Figure 6 Systematics of experimental yield of various light-particle-accompanied fissions relative to ternary processes in which the α-particle is emitted. (Reproduced with permission from Poenaru DN, Greiner W, Hamilton JH, and Ramayya AV (2001) *Journal of Physics G* 27: L19–L28; © IOP Publishing Limited.)

Proton Radioactivity

The radioactive emission of one and two protons has been theoretically predicted by V Goldansky since 1960. S Hofmann *et al.* discovered, in 1980, the spontaneous proton emission from the ground state of ^{151}Lu with a half-life of 85 ms and a kinetic energy of 1233 keV. The relatively large angular momentum involved by making a transition from ^{151}Lu with a spin and parity $11/2^-$ to the ground state 0^+ of ^{150}Yb contributes substantially to the potential barrier by this centrifugal component. There are almost 30 proton emitters beyond the proton drip line with odd atomic numbers between 51 and 83. Examples are: ^{105}Sb, ^{109}I, ^{112,113}Cs, ^{117}La, ^{131}Eu, ^{141}Ho, ^{145,146,147}Tm, ^{150,151}Lu, ^{156}Ta, ^{160,161}Re, ^{164,165,166,167}Ir, ^{170,171}Au, ^{177}Tl, and ^{185}Bi.

The odd–even staggering of the proton separation energy S_p, responsible for the nonregular structure of the drip line, leads to situations where $S_p > 0$ for a given nucleus, while $S_{2p} < 0$. Possible candidates for two-proton radioactivity are: ^{17}Ne, ^{19}Mg, ^{39}Ti, ^{42}Cr, ^{49}Ni, ^{59}Ge, ^{152}Hf, ^{153}Ta, ^{154}W, and ^{155}Re. In 2002, there were two successful experiments at GANIL and GSI reporting the 2p radioactivity of ^{45}Fe with an energy of 1.14 MeV and a half-life of 4.7 ms.

Heavy Ion Radioactivity

Heavy ion radioactivities were predicted in 1980 by A Sandulescu, D Poenaru, and W Greiner, four years before the first experimental evidence reported by H J Rose and G A Jones. Some of the cluster emitters are members of the natural radioactive families (see **Figure 3**, where part of the disintegration chain of ^{235}U is shown). The main region of emitters experimentally observed is above $Z = 86$ with daughters around ^{208}Pb: ^{221}Fr, $^{221–224,226}$Ra, ^{225}Ac, ^{228,230}Th, ^{231}Pa, $^{230,232–236}$U, ^{236,238}Pu, and ^{242}Cm. There is a strong competition with α-decay. The largest branching ratio with respect to α-decay, $b = T_\alpha/T$, is $\sim 10^{-9.2}$.

From many decay modes with half-lives and branching ratios relative to α-decay predicted with the analytical superasymmetric fission (ASAF) model, the following have been experimentally confirmed: ^{14}C, ^{20}O, ^{23}F, $^{24–26}$Ne, ^{28,30}Mg, ^{32,34}Si. The experimental data are in good agreement with predicted values (see the examples given in **Figure 7**). A strong shell effect can be seen: as a rule, the shortest value of the half-life (maximum of $1/T$) is obtained when the daughter nucleus has a magic number of neutrons ($N_d = 126$) and/or protons ($Z_d = 82$).

The fine structure in ^{14}C radioactivity of ^{223}Ra, for the first time discussed by M Greiner and W Scheid in 1986, was discovered in 1989 by E Hourani, M Hussonnois *et al.* at IPN Orsay. Surprisingly, they observed that a transition of the daughter to the first excited state was stronger than that to the ground state.

The measurable decay constant $\lambda = vSP$ is a product of three model-dependent terms of which the frequency of assaults on the barrier per second, v, may be considered constant. The preformation probability, S, can be calculated quasiclassically as the internal barrier penetrability in a fission model. With good approximation, one can obtain one universal curve for any kind of cluster decay mode with a mass number A_e, including α-decay:

$$\log T = -\log P - 22.169 + 0.598(A_e - 1)$$

where the penetrability, P, is calculated easily for the given split. In a logarithmic scale, the equation $\log T = f(\log P)$ represents a single straight line which can be conveniently used to estimate the half-life.

Nuclear Stability

The area of experimentally identified nuclei until 2003, covering 2228 species, extends well beyond the central region of about 300 beta-stable nuclides (shown as black squares in **Figure 8**), both on the proton and neutron-rich sides. A semi-empirical approximation for the line of β-stability, $N - Z = 0.4A^2/(200 + A)$, allows one to estimate which mass number, A, or neutron number, N, corresponds to a given proton number, Z, for a β-stable nucleus. It is

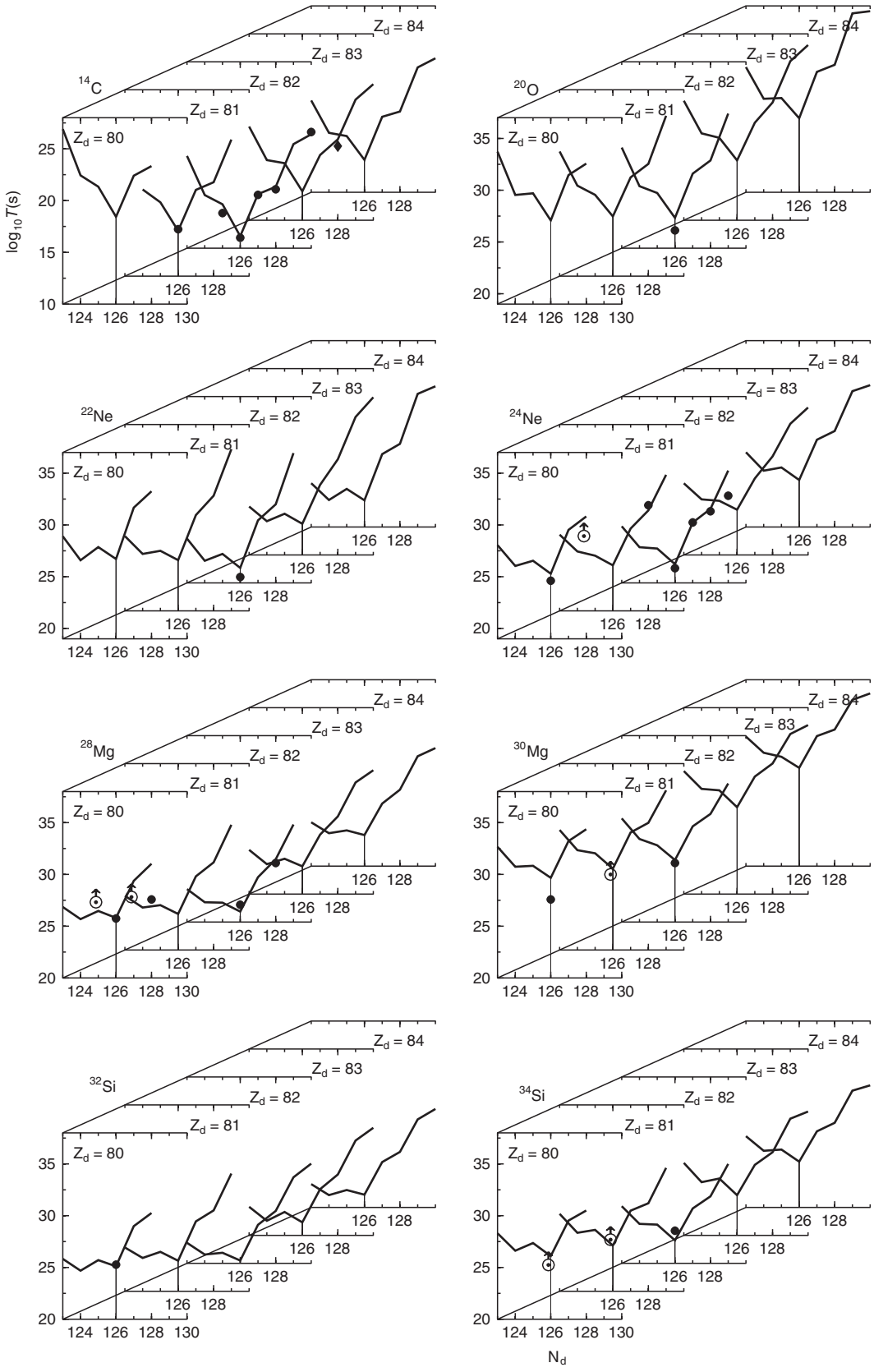


Figure 7 Predicted half-lives within the ASAF model (lines) and measurements (points) for eight kinds of cluster decay modes vs. proton (Z_d) and neutron (N_d) number of the daughter nucleus. The shell effects are clearly seen as minima of lifetime for magic numbers of the daughter, $Z_d = 82$, $N_d = 126$.

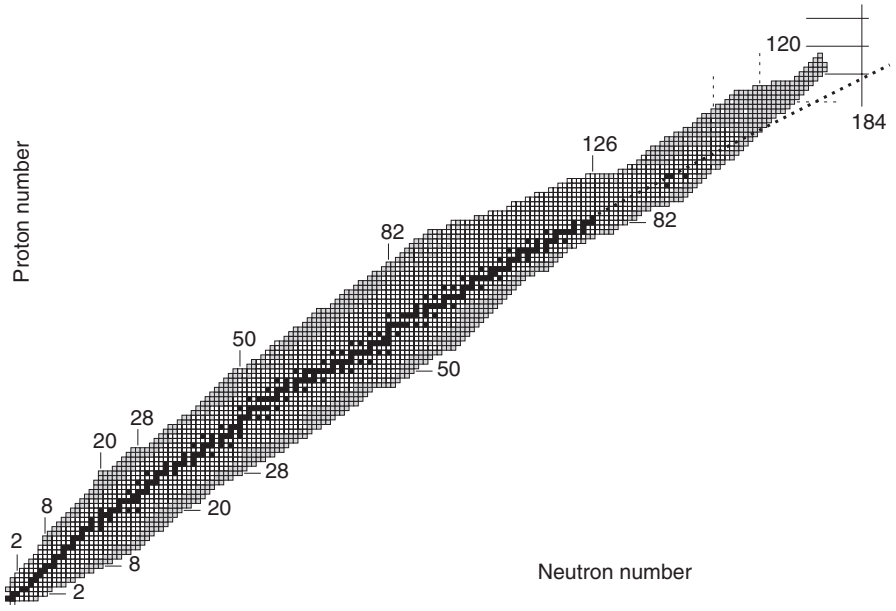


Figure 8 Nuclear chart. The nuclides found in nature are represented by black squares. The closed shell ("magic") numbers of neutrons and protons are displayed. There are ~2228 nuclides with measured masses, and for some 951 others (dashed squares on the figure), the masses were obtained by extrapolation. (From Audi G, Wapstra AH, and Bersillon O (2003) *Nuclear Physics A* 729: 337–676.)

displayed as a dotted line for $Z > 82$ in **Figure 8**. Like the noble gases He, Ne, Ar, Kr, Xe, and Rn, which have an increased stability due to their closed-shell electron configuration (for 2, 10, 18, 36, 54, and 86 electrons) nuclei with the magic number of protons or neutrons, 2, 8, 20, 28, 50, 82, shown in **Figure 8**, are particularly stable as a consequence of quantum-mechanical laws.

Many nuclides found in nature or produced artificially by nuclear reactions decay by spontaneous emission of certain particles or electromagnetic radiation. The daughter nucleus can in its turn be unstable, and the phenomenon continues until a stable configuration is reached.

The limits of nuclear stability are determined either by hadron-decay modes, or (for heavy nuclei) by α-decay or spontaneous fission. The most important limits are the proton and neutron drip lines, where the proton or the neutron has roughly zero binding energy. A large proton excess is not possible owing to the Coulomb repulsion; for light nuclei, the proton drip line is very close to $N = Z$. By combining the Coulomb and centrifugal (angular momentum) barriers, the lifetime of proton emission from an unbound state can be long enough to be experimentally detected in a proton-decay experiment. In this way, the proton drip line can be overstepped: one accedes to a nucleus situated outside the drip line. A series of proton-rich In and Sn isotopes (including the doubly magic ^{100}Sn) were produced. The heaviest known proton emitter is $^{185}_{83}\mathrm{Bi}_{102}$. The neutron drip line has been reached only for the lightest nuclides up to $^{49}_{16}\mathrm{S}_{33}$.

The existence of extremely heavy nuclei is prevented by the spontaneous fission process. When the number of protons becomes very high, the Coulomb repulsion is no longer counter-balanced by the strong interaction, and the phenomenological liquid drop model (LDM) fission barrier height may approach zero. By adding a shell correction term to the LDM energy of a superheavy nucleus, a potential barrier shows up. This shell-stabilizing property, characteristic of the superheavy nuclei, is already present in the heaviest elements known to date with $Z = 107 - 118$ (see **Figure 9**). An example of a potential energy surface of the superheavy nucleus with $Z = 120$ and $A = 304$ is given in **Figure 10**.

Owing to shell effects, the fission barriers of these nuclei are so high that they decay mainly by α-emission instead of spontaneous fission. The magic numbers in the superheavy region are still not known. The candidates are $Z = 114, 120, 126$ and $N = 184$.

Radiation Exposure

Human beings and different materials are permanently exposed to radiations coming from primordial (since the origin of the Earth), cosmogenic (formed as a result of cosmic-ray interactions), or human-produced elements. More than 2/3 of the radiation dose corresponds to natural radioactivity and about 1/4 to

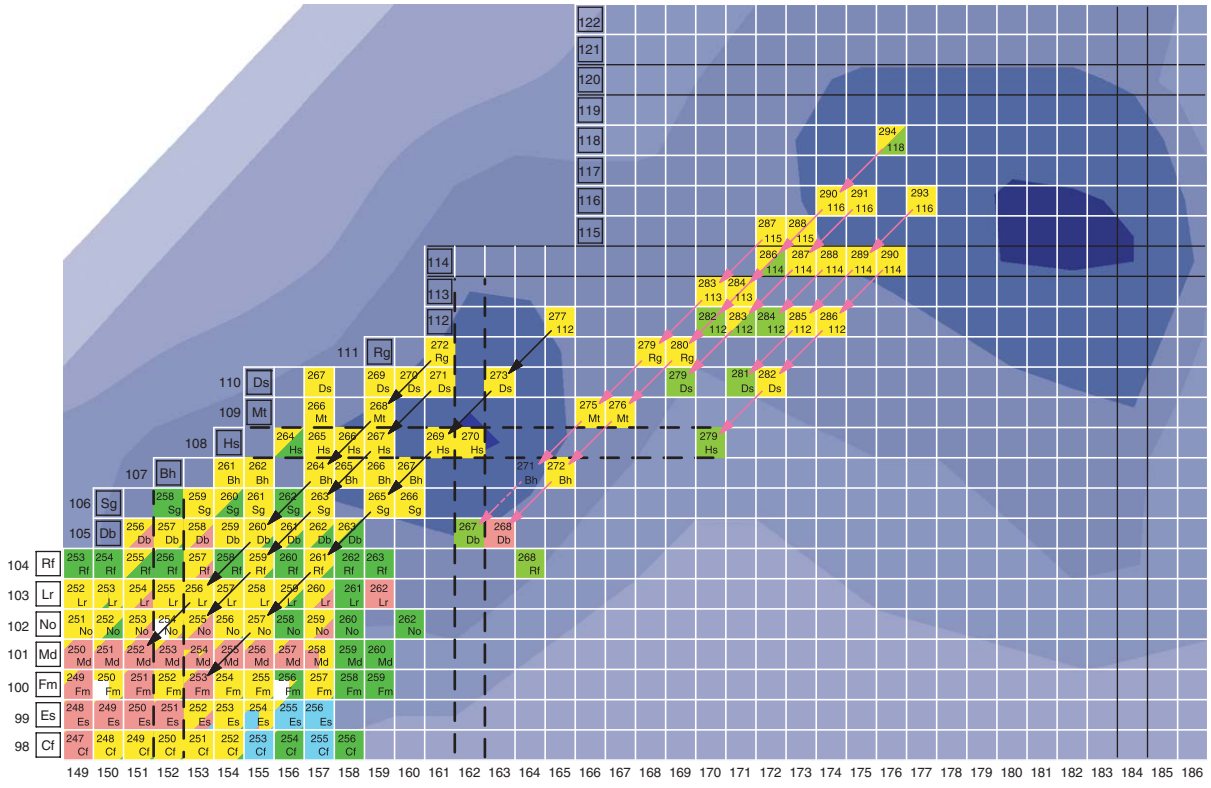


Figure 9 Part of the nuclear chart displaying the superheavy nuclei produced at GSI Darmstadt and JINR Dubna. The background shows two minima (dark blue) of shell correction energies where the most stable nuclides are expected to be found, around $Z = 114$, $N = 184$ for spherical nuclei and $Z = 108, N = 162$ for the deformed ones. (From Hofmann S and Darmstadt GSI (2004).)

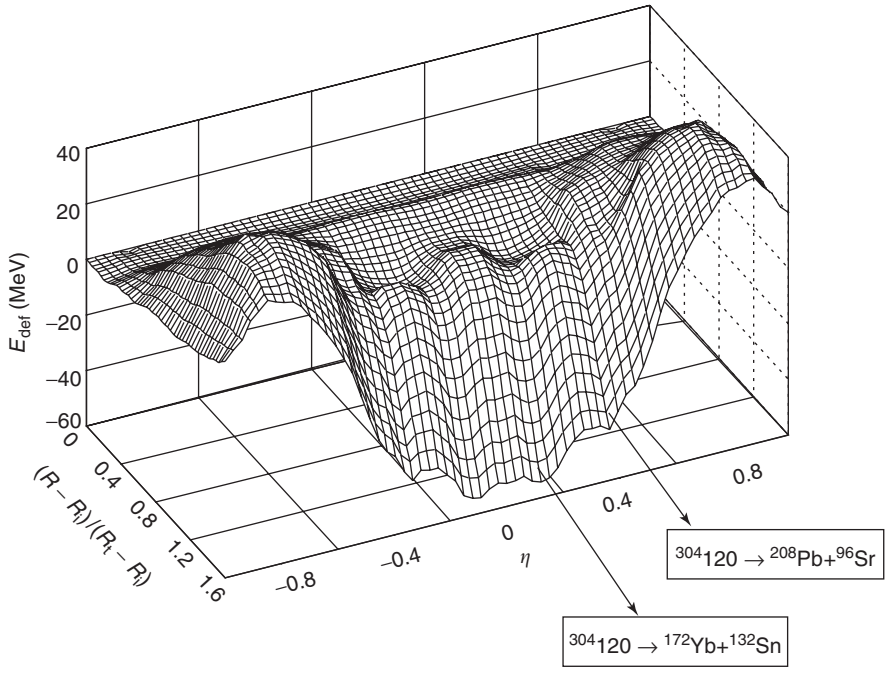


Figure 10 Potential energy surface of the superheavy nucleus $^{304}120$ vs. fragment separation and mass asymmetry $\eta = (A_1 - A_2)/(A_1 + A_2)$. The survival of this nucleus is based on the existence of a finite value of potential barrier. Two main valleys corresponding to the split with one doubly magic fragment ^{132}Sn and ^{208}Pb can be seen. (Reproduced with permission from Gherghescu RA, Poenaru DN, and Greiner W (1998) *Journal of Physics G* 24: 1149; © IOP Publishing Limited.)

medical irradiations (mainly X-rays). When passing through air, ionizations are produced. The exposure is the total electric charge on the ions produced in a given mass of air and is measured traditionally in roentgen (R). The SI unit is coulomb per kilogram: $1\,\mathrm{R} = 2.58 \times 10^{-4}\,\mathrm{C\,kg^{-1}}$.

The quantity expressing the energy absorbed in a mass unit by ionization in different materials is called the absorbed dose, D. The traditional unit 1 rad corresponds to the absorption of 100 ergs per gram of material. The SI unit is the gray (Gy) equal to the absorption of $1\,\mathrm{J\,kg^{-1}}$, hence $1\,\mathrm{Gy} = 100\,\mathrm{rad}$. Under 250 mGy, no sizeable biologic or medical effect has been observed. On the other extreme, the lethal dose is $\sim$5000 mGy.

For human beings, it is important to take into account the biological effects of different radiations. Ionizing radiation causes physical damage to cells and DNA. A quality factor (QF) is calculated for a given type and energy of radiation according to the energy deposited per unit path length. For X-rays, β and γ, which deposit little energy per unit length, $\mathrm{QF} = 1$, and for α-particles QF may reach a value as high as 20. The dose equivalent $\mathrm{DE} = D \cdot \mathrm{QF}$ is measured traditionally in rem and the SI unit is called sievert (Sv), $1\,\mathrm{Sv} = 100\,\mathrm{rem}$.

The International Commission on Radiation Protection recommended a limiting annual whole-body absorbed dose of 5 mSv per year for the general public and 50 mSv per year for professionals working with radiations. The mean value of the natural radiation dose received in France is $\sim$3.5 mSv per person in a year: $\sim$0.4 mSv from cosmic radiations coming from space, 0.4 mSv from the earth (uranium, thorium, potassium), 1.6 mSv from the elements absorbed by breathing (particularly the gas radon) and eating (potassium), and 1.1 mSv from artificial exposure of medical origin (mainly X-rays). The industrial activity represents $\sim$0.1 mSv, of which only 0.02 mSv comes from nuclear energy.

See also: Nuclear Fission and Fusion; Quantum Mechanics: Atoms; Quantum Mechanics: Elementary Particles; Quantum Mechanics: Nuclei; Radioactive Tracers.

PACS: 21.10.−k; 21.10.Dr; 21.10.Tg; 23.40.−s; 23.50.+z; 23.60.+e; 23.70.+j; 23.90.+w; 25.85.Ca; 29.25.Rm

Further Reading

Audi G (2003) *Nucleus* (Chart of nuclides and their properties). Available at http://csnwww.in2p3.fr/AMDC/web/nubdisp_en.html

Booth CN and Combley FH (2000) *Nuclear Physics* 303. Available at http://www.shef.ac.uk/physics/teaching/phy303/phy303.html

Eisenbud M and Gesell T (1997) *Environmental Radioactivity from Natural, Industrial and Military Sources*. San Diego: Academic Press.

Hofmann S (2003) *On Beyond Uranium*. London: Taylor & Francis.

Krane KS (1988) *Introductory Nuclear Physics*. New York: Wiley.

Poenaru DN and GreinerW (eds.) (1996) *Nuclear Decay Modes*. Bristol: IOP Publishing.

The special volume of *Radiochimica Acta* 70/71 (1995) devoted to One hundred years after the discovery of radioactivity. München: Oldenburg Verlag.

Nomenclature

A	mass number
$\mathscr{A}$	activity (Bq)
b	branching ratio
B	binding energy (MeV)
D	absorbed dose (Gy)
DE	dose equivalent (Sv)
E	kinetic energy (MeV)
N	neutron number
P	quantum penetrability
Q	released energy (MeV)
S	preformation probability
T	half-life (s)
TKE	total kinetic energy (MeV)
v	frequency of assaults ($\mathrm{s^{-1}}$)
Z	atomic number
Δ	mass defect (MeV)
λ	decay constant ($\mathrm{s^{-1}}$)
τ	lifetime (s)

Recovery, Recrystallization, and Grain Growth

N Hansen, Risø National Laboratory, Roskilde, Denmark
R A Vandermeer, Columbus, OH, USA

Introduction

Metallic materials harden when deformed and soften, if afterward annealed. This behavior has been known to man since the beginning of our civilization, and it is today exploited in the industrial fabrication of metals and alloys by thermomechanical processing. Hundreds of millions of tons of metallic materials are produced annually.

Examples of such processing are aluminum beverage cans, steel sheets and plates for cars and ships, and light metals such as aluminum and titanium for aeroplanes and satellites.

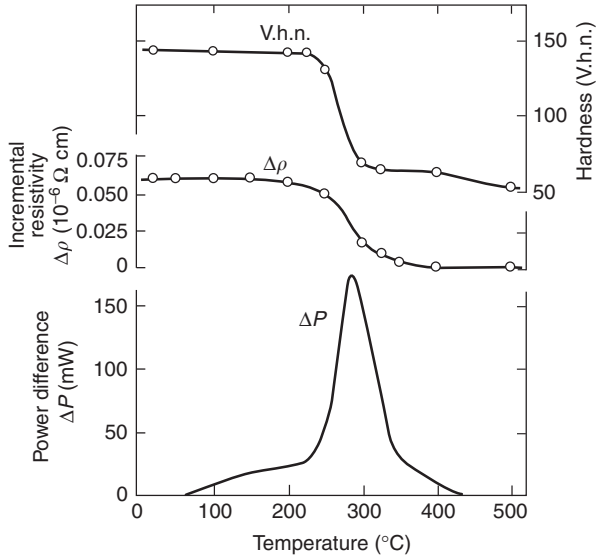


Figure 1 Change in hardness, electrical resistivity, and stored energy (represented by a power differential) during heating of deformed pure copper. (Reproduced from Clareborough HM, Hargreaves ME, and West GW (1955) *Proceeding of the Royal Society London* 232 A, 252.)

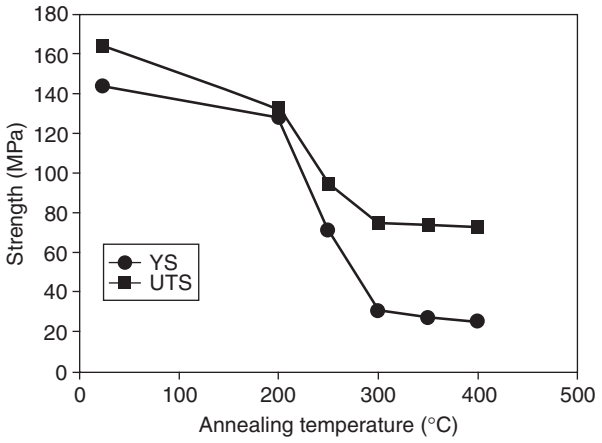


Figure 2 Yield stress (YS) and ultimate tensile strength (UTS) at room temperature after heating of a sheet of cold rolled aluminum.

During a thermomechanical treatment, the metal is exposed to a plastic deforming force, which hardens it and to an elevated temperature, which leads to softening either during deformation or after. A thermomechanical process can therefore, for example, be hot rolling or cold rolling followed by annealing, that is, a heat treatment at elevated temperatures. A simultaneous exposure to force and temperature typically gives rise to the processes of dynamic recovery and dynamic recrystallization. In the case where the heat treatment is applied after deformation, these processes are recovery, recrystallization, and grain growth, which are considered in what follows.

Recovery, recrystallization, and grain growth refer to thermally activated microstructural or even submicrostructural processes, which directly affect the properties of the metal (**Figure 1**). This is because substantial and important relationships exist between the structure of metallic materials and their properties. Therefore, to understand and model such properties, the microstructural changes taking place during recovery, recrystallization, and grain growth must first be described and quantified. For this purpose, microstructural parameters are determined linking processing and properties. This article focuses on mechanical properties, but physical and chemical properties also are influenced by the microstructure.

The mechanical properties of a metal describe its resistance to flow, when subjected to a load. Mechanical properties of a metal are therefore legion. They can be general parameters such as the yield and flow stress, or they can be more specific, for example, related to the metal's formability where the work hardening rate, elongation, etc., may be important. Furthermore, the properties depend on the deformation mode, which may be mono tonic or cyclic. Therefore, with an application in view, many different structure–property relationships must be evaluated. This article will concentrate on strength properties such as the yield stress, the flow stress, and the hardness (**Figure 2**). This is because these properties have been most frequently used in previous studies, they are easy to measure, and physically based models have been formulated for their dependency on a variety of microstructural parameters. Further properties, which directly reflect the strength of the metal, are of general interest from the point of view of applications as there always has been an incentive to develop strong metals, which can carry larger loads by smaller cross sections.

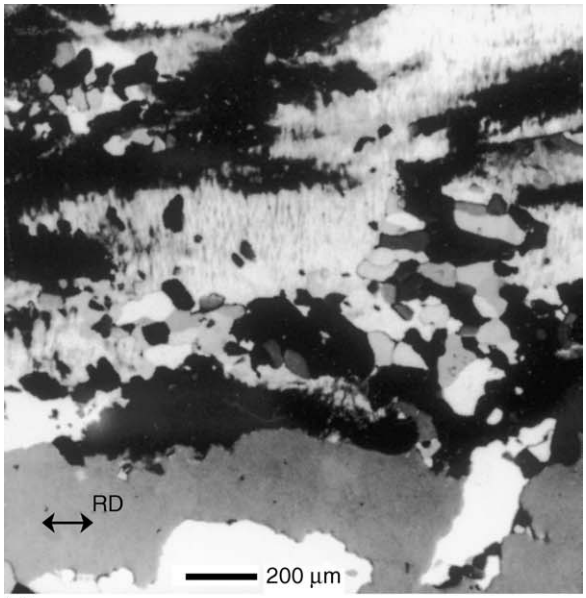
Recovery, Recrystallization, and Grain Growth (Phenomena)

Recovery refers to the structural changes associated with the annihilation of deformation-induced defects, and/or their rearrangement into lower-energy configurations giving rise to substantive changes in properties without changing the basic orientation or crystallographic texture of the material. Recovery is often characterized and described in terms of the recovery of specific properties, chief among them the hardness, flow stress, stored energy, and electrical resistivity. But, since the advent of the electron microscope, changes in microstructural features such as the dislocation density and subgrain and dislocation

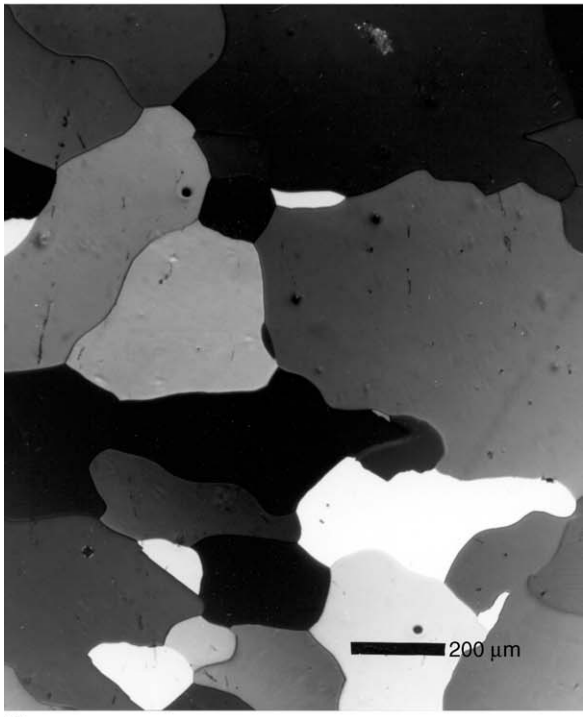
cell sizes have also been detailed along with the property changes. Usually recovery mechanisms alone, which occur primarily at the lower annealing temperatures, cannot restore the material to its pre-deformation state. The term recovery is sometimes also used to describe the phenomena associated with annealing out of quenched-in point defects such as vacancies, and the annealing out of irradiation-produced defects such as interstitials and voids.

In competition with the various recovery mechanisms acting to anneal out the deformation-induced defects is the microstructural phenomenon called primary recrystallization or simply recrystallization, which manifests itself at intermediate annealing temperatures. Recrystallization is understood to entail the evolutionary change whereby a population of new, strain-free grains emerge from the deformed volumes in a step referred to as nucleation, that is, they become identifiable at a length scale of a few micrometers, followed by the increase in size or growth of these nuclei/grains until the deformed volumes are consumed and the new grains all impinge upon one another (**Figure 3**). The hallmark of a growing grain is that its grain boundary, a portion of which is required to be of the high-angle type sweeps through the deformed volumes by a boundary migration mechanism. In so doing, the deformation-induced defects in the swept-out path are eliminated. When one new grain meets another, the segments of their grain boundaries that impinge, that is, come together to form a new grain boundary, cease to migrate. Recrystallization can create new crystallographic orientations in local regions and often a different texture is developed in the material as a result of recrystallization.

When recrystallization is complete, the material is polycrystalline but with the deformation-induced defects eliminated. In this state the substance is composed of many contiguous crystals, that is, grains, defined by an interconnected grain boundary network consisting of nodes (grain corners), branches (grain edges), and faces (grain boundaries), and having a thickness of atomic dimensions. In a thermodynamic sense, this interface array imparts an excess energy to the material, which would not be present in a single crystal. The material perceives the excess energy through the topological grain boundary network as an interfacial tension in the Gibbsian sense. Equilibration of the interfacial tensions throughout and the requirement to maintain perfect space filling causes the grain boundary faces to be curved, and so gives rise to pressure differences between grains. Given sufficient thermal activation and atomic mobility, these pressure differences cause grain boundaries to migrate toward their centers of curvature, thereby reducing the amount of grain boundary area in the material and systematically eliminating some of the grains. Concomitantly the remaining grains, on average, enlarge. This



(a)



(b)

Figure 3 (a) Partly recrystallized structure in deformed and annealed aluminum. Regions with remaining deformed (recovered) matrix material have a stripy appearance and contain new grains of different size. (b) Fully recrystallized state showing a polycrystalline structure. (Reproduced from Juul Jensen D (1997) *Orientation Aspects of Growth during Recrystallization*, Risø-R-978 (EN). Roskilde, Denmark: Risø National Laboratory.)

evolutionary process is known as grain growth and usually occurs at higher annealing temperatures. When the growth process is characterized microstructurally by a progressive increase in average size of the grains and the grain size distribution behaves in a statistically self-similar manner, that is, consecutive configurations of the system are geometrically similar in a statistical sense, the process is referred to as normal or steady state grain growth. A few special grains may also grow rapidly to sizes much larger than the average, and consume the smaller stagnant grains in a process termed abnormal grain growth. Thus, the driving force for grain growth is derived from the grain boundary energy unlike that for recovery and recrystallization, which arises from the deformation-induced defects.

Microstructural Parameters

The microstructural parameters determining the mechanical property change during annealing from fine scale dislocation structures to coarse scale grain structures, that is, microstructures must be characterized over length scales from the nanometer/submicrometer range to the micrometer/millimeter range. A thorough microstructural characterization must therefore be based on the application of a number of microscopical techniques such as optical microscopy (OM), scanning electron microscopy (SEM), and transmission electron microscopy (TEM). By these techniques, a qualitative and quantitative description of the morphological changes taking place during annealing is obtained. Combined with techniques for analyzing the local and macroscopic crystallography, that is, orientations (texture), structural parameters can be quantified.

The changes in microstructural and crystallographic parameters taking place during annealing of a deformed metallic material relate directly to the structure in the deformed state. This structure depends on a number of materials and process parameters. However, microstructural characterization and modeling of a variety of metals and alloys deformed by different thermomechanical processes have shown a universal structural evolution. Dislocations are generated during deformation and they accumulate in dislocation boundaries, which subdivide the original grains on a finer and finer scale as the degree of deformation (the plastic strain) is increased. This universal structural evolution has its origin in common principles and processes that apply to the formation and storage of dislocations during plastic deformation (**Figure 4**).

The universal structural evolution during plastic deformation is defined by a number of microstructural

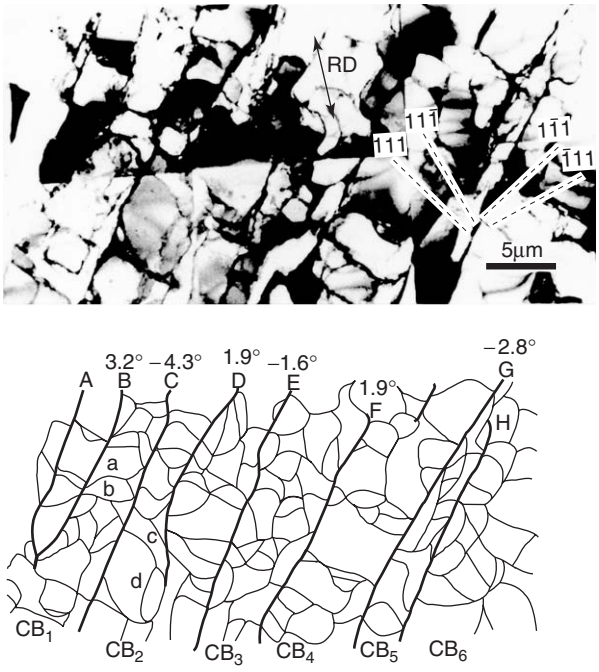


Figure 4 Transmission electron microscopy image and sketch of microstructure in aluminum, cold rolled to 10% reduction in thickness. The structure is subdivided by extended dislocation boundaries (marked A, B, C, etc.) and short cell boundaries (marked a, b, c, etc.). (Liu Q and Hansen N (1995) Geometrically necessary boundaries and incidental dislocation boundaries formed during cold deformation. *Scripta Metallurgica et Materialia* 32: 1289–1295.)

parameters, which in various ways have been related to the mechanical properties of materials in the deformed state and after recovery. Recovery leads to a coarsening of the deformed structure, that is, the recovered structure is to an extent similar to the deformed structure. However, after nucleation and growth have initiated recrystallization, the microstructure changes into a composite structure, where large nuclei are embedded in a deformed matrix. After completion of recrystallization, this composite structure is replaced by a grain structure (**Figure 3b**) characterized by the average grain size and the grain size distribution. During grain growth the structural evolution is self-similar, that is, the probability distributions of grain size collapse into one function, when scaled with the average grain size for each distribution.

Mechanical Properties and Microstructural Parameters

A large research effort has been devoted to the formulation and validation of relationships between the mechanical properties and a number of different microstructural parameters.

In the deformed state, the most general relationship expresses the dependency of the flow stress (σ_f) on the total dislocation density ρ_t:

$$\sigma_f - \sigma_0 = M_1 \alpha b G \sqrt{\rho_t} \quad [1]$$

where σ_0 is a friction stress, M_1 is the so-called Taylor factor, α is a number of the order of 0.2–0.3, b is the Burgers vector, and G is the shear modulus. Apart from a variation in α, eqn [1] is in general fairly insensitive to changes in the dislocation structure, purity, strain rate, and temperature.

In the recrystallized state, the yield stress σ_y can be related to the grain size D by the Hall–Petch equation:

$$\sigma_y - \sigma_0 = kD^{-0.5} \quad [2]$$

where k is a constant, and σ_0 is the yield stress of an undeformed single crystal orientated for multiple slip or approximately the yield stress of a very coarse untextured polycrystal. An equation similar to eqn [2], however not as universal, relates the flow stress at a particular strain $\sigma(\varepsilon)$ to the grain size:

$$\sigma(\varepsilon) - \sigma_0(\varepsilon) = k(\varepsilon)D^{-0.5} \quad [3]$$

where $\sigma_0(\varepsilon)$ and $k(\varepsilon)$ are constants at a given strain (**Figure 5**).

A characteristic structural feature in many deformed, recovered, and partly recrystallized metals is dislocation cells or subgrains. These features can be defined as follows: (1) A cell is a roughly equiaxed volume element within which the dislocation density is well below the average density and which is rotated relative to similar adjacent volume elements, from which it is separated by more or less well-formed rotational dislocation boundaries. (2) A subgrain is a dislocation cell, where the relative rotations between the lattice on either side of the boundary has become so large that the neighboring cell can no longer operate with the same combination of simultaneously operating glide systems. For cell- and subgrain-forming metals, numerous experiments have shown that an empirical relationship exists between the flow stress σ_f and the average cell or subgrain diameter d:

$$\sigma_f - \sigma_0 = k_2 G b d^{-n} \quad [4]$$

where k_2 is a constant. In contrast to the Hall–Petch equation the exponent is not a constant, but is normally found to be in the range 0.5–1.0 with values in the upper end of this interval characteristic of dislocation cell structures, and a value in the lower end characteristic of subgrain structures. The exponent is, therefore, an indirect measure of the boundary strength. For a cell structure, where the cell boundaries are of low angle and $n \simeq 1$, it follows from eqns [1] and [4] that the dislocation density must be related to the cell size. This relationship has been expressed in an equation for a mixed tilt/twist boundary:

$$\rho_t = \rho_0 + \frac{1.5\theta}{b} S_v \quad [5]$$

where ρ_0 is the dislocation density in the material between the dislocation boundaries, θ is the angle of misorientation across the boundaries, and S_v is the area of dislocation boundaries per unit volume of material being inversely proportional to the boundary spacing. For equiaxed cells, S_v is typically equal to $3/d$.

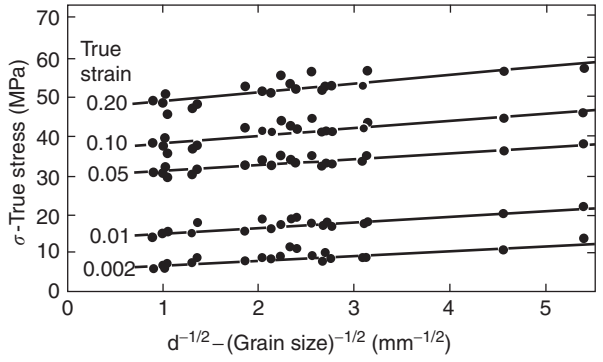


Figure 5 The dependence of yield stress and flow stress on the grain size of polycrystalline copper. (Hansen N and Ralph B (1982) The strain and grain-size dependence of the flow-stress of copper. *Acta Metallurgica* 30: 411–417.)

Structure and Mechanical Properties

The structural parameters of the deformed state relate to the presence of the original grain boundaries and the density of dislocations introduced by the deformation. On the assumption that the strength contributions of grain boundaries and dislocations can be added linearly, the flow stress can be expressed by eqn [2] when introducing the dislocation strengthening expressed by combining eqns [1] and [5]. The flow stress can then be written as

$$\sigma(\varepsilon) = \sigma_0(\varepsilon) + M_1 \alpha G \sqrt{\rho_0 + 1.5 b S_v \theta} + k(\varepsilon) D^{-0.5} \quad [6]$$

In studies of the flow stress behavior of cold rolled aluminum, nickel, and iron, good agreement between theory and experiment has been found (**Figure 6**).

In the recovered state, the microstructure may be quite similar to that of the deformed state but the structure has coarsened leading to reduction in S_v. The value of θ may also change, however, to a much lesser extent. Finally, the dislocation density between the boundaries (ρ_0) decreases as dislocations are annihilated or are integrated into the cell/subgrain

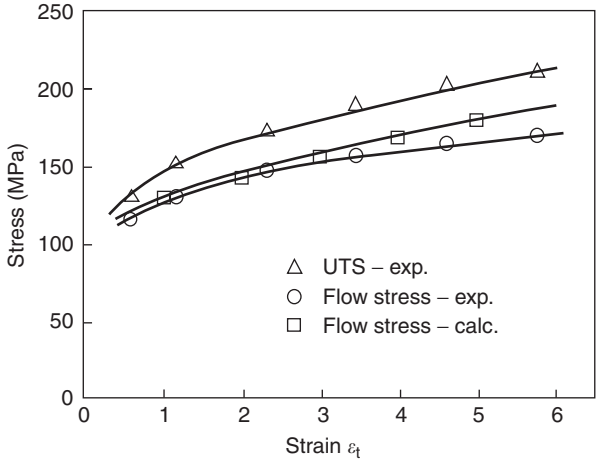


Figure 6 Comparison of experimental and calculated strength properties (ultimate tensile strength and flow stress at 0.2% offset) of aluminum cold rolled up to a strain of 5.5, which corresponds to a 99.6% reduction in thickness. (Liu Q, Huang X, Lloyd DJ, and Hansen N (2002) Microstructure and strength of commercial purity aluminum (AA 1200) cold-rolled to large strains. *Acta Materialia* 50: 3789–3802.)

boundaries. As a result the flow stress decreases with a rate, which depends on the rate of annealing processes, for example, vacancy and dislocation annihilation and migration of subboundaries, and high-angle boundaries.

In the recrystallization stage, the structure is a composite of soft recrystallized grains in a much harder matrix (**Figure 7**); the flow stress can be expressed by the equation

$$\sigma_f - \sigma_0 = V_r\sigma_r + (1 - V_r)\sigma_m \qquad [7]$$

where V_r is the volume fraction of recrystallized material, and σ_r and σ_m are the flow stresses of undeformed (recrystallized) and recovered structures, respectively. The recovered structure is strengthened mainly by cells/subgrains, that is, σ_m can be expressed by eqn [4]. During recrystallization, σ_m decreases due to coarsening of the cell/subgrain structure, and as V_r increases, the flow stress decreases rapidly.

In the recrystallized state and after normal grain growth (steady state) the mechanical properties (e.g., yield stress, flow stress, and hardness) can be expressed by a Hall–Petch type relationship, expressing proportionality between the property and the reciprocal square root of the grain size (**Figure 5**).

Kinetics

The change in mechanical properties and in the strength-determining structural parameters may be related directly to the time and temperature of annealing and expressed in rate equations.

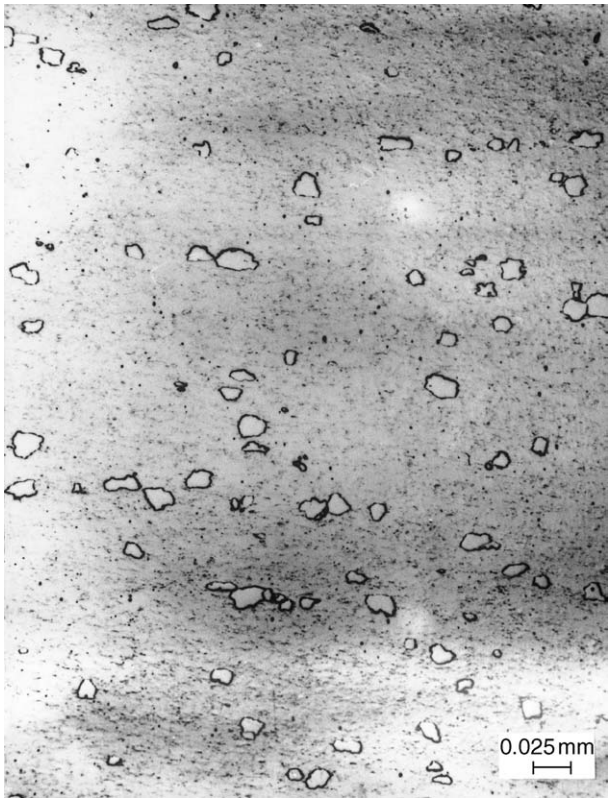


Figure 7 Early stage of recrystallization in deformed crystals of iron. (Vandermeer RA and Rath BB (1989) Modeling recrystallization kinetics in a deformed iron single crystal. *Metallurgical Transaction A* 20A: 391–401.)

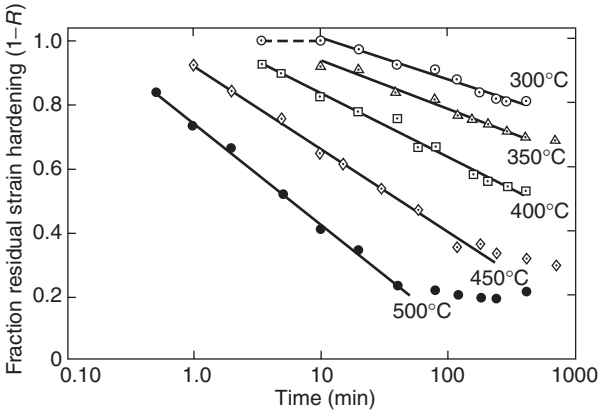


Figure 8 Recovery kinetics of flow stress of polycrystalline iron following 5% prestrain (1-*R*) represents the fraction of residual hardening, i.e., (1-*R*) is one in the deformed state and zero after recrystallization. (Michalak JT and Paxton HW (1961) Some recovery characteristics of zone melted iron. *Transactions of the Metallurgical Society of AIME* 221: 850–857.)

During recovery, the change in the mechanical properties during isothermal recovery is highest initially and decreases with time as recovery proceeds (**Figure 8**). There is no universal rate equation, which can describe all the changes during annealing;

however, one rate equation has been found to give a good empirical description:

$$\frac{dP}{dt} = -(P - P_r)K_0 \exp - \left(\frac{Q_0 - \beta(P - P_r)}{RT}\right) \quad [8]$$

where P is the instantaneous value of the property, for example, the flow stress or the hardness, P_r is the remnant value of the property which is unrecoverable after very long annealing times, T is the absolute temperature, t is the annealing time, R is the gas constant, and K_0, Q_0, and β are constants characterizing the recovery behavior. Except at large or small times, this rate equation for recovery may be integrated to yield a property change that depends approximately on the logarithm of the annealing time. Because recovery is something of a generic term used to represent various and sundry property changes, usually no single atomic mechanism can be invoked to characterize it. Thus, other rate equations may be required to adequately describe some experiments.

The change in mechanical properties with time and temperature may also be analyzed, based on rate equations describing the changes in strength-determining structural parameters, for example, the dislocation density, the cell/subgrain size, and the grain size. Common to these changes is that, they are controlled by a driving pressure (the stored energy). For a deformed structure, this driving pressure P ($N\,m^{-2}$) is proportional to the dislocation density multiplied by the energy per unit length of dislocation line or approximately

$$P = 0.5\rho_t G b^2 \quad [9]$$

where ρ_t is the total dislocation density stored in boundaries and between the boundaries. For a structure subdivided into cells/subgrains of grains, the driving pressure can be related to the total boundary energy per unit volume when

$$P \cong S_v \gamma_b \quad [10]$$

where S_v is the boundary area per unit volume and γ_b is the boundary energy per unit area. For cell/subgrain boundaries γ_b increases with the angle of misorientation across the boundary, whereas for high-angle grain boundaries ($>15^\circ$), γ_b is the grain boundary energy ($J\,m^{-2}$) taken to be a constant for all boundaries.

The rate equations for the structural parameters, describing the change in dislocation density as a function of temperature and time during recovery, involve a number of processes, which have been modeled separately. However, no universal relations have been proposed. In contrast, it is generally accepted that for boundary migration there is a simple linear relationship between velocity V and the driving pressure P:

$$V = MP \quad [11]$$

where M is the boundary mobility, which is a parameter characterizing the boundary migration mechanism. The temperature dependence of M can be expressed by an Arrhenius equation:

$$M = M_0 \exp[-Q/RT] \quad [12]$$

where M_0 is a pre-exponential "constant," and Q is the activation energy for the rate determining process (e.g., an atom transfer process). Equation [12] can be derived from either an absolute reaction rate mechanistic premise or on an irreversible thermodynamic principle.

The change in mechanical properties during recrystallization is expressed in eqn [7], where a key parameter is V_r, the volume fraction of recrystallized material. Phenomenologically, recrystallization may be envisioned as a nucleation and growth "transformation" and treated so, kinetically. Unlike recovery, the isothermal rate of recrystallization exhibits "sigmoidal" kinetics starting out low, increasing to a maximum value, and then decreasing again with time. Isothermal recrystallization kinetics are usually described experimentally in terms of the Avrami equation:

$$V_r = 1 - \exp(-Bt^k) \quad [13]$$

where V_r is the volume fraction recrystallized, t is the annealing time, and B and k are parameters that depend on the characteristics of recrystallization such as the nucleation rate, growth rate, the shape of the recrystallized grains, and the nature of the new grain impingement patterns. Ideally B and k may be constants, but often one or the other (usually the k) is not. Deviations from the ideal may be caused by competition between the recovery and recrystallization mechanisms for the stored energy of cold deformation, by the heterogeneous nature of the deformed microstructure, by changes in the growth rate mechanism, by complex impingement patterns, or combinations of these phenomena.

Finally, considering normal grain growth as an example, the driving pressure is $\gamma_b S_v$ and S_v is $2/D$, where D is the grain size measured along random lines. The boundary migration rate is therefore

$$\frac{dD}{dt} = \frac{2M\gamma}{D} \quad \text{or} \quad (D^2 - D_0^2) \cong Kt \quad [14]$$

where D_0 is the length scale of the grain size at time equal to zero (fully recrystallized grain size), and K is a temperature-dependent rate parameter. However, failure of experiment to conform to parabolic kinetics

is the rule rather than the exception and may be attributed to a number of material related factors: (1) the scaled grain size distribution has not achieved statistical self similarity, and grain growth is in a transient period, (2) pinning forces causing growth stagnation are present, for example, second phase precipitates, surface grooving of grain boundaries, and intrinsic drag forces due to soluble impurities, and (3) the onset of abnormal grain growth.

Based on eqns [2] and [14], the changes in yield stress with annealing time can be expressed as $\sigma \simeq t^{-0.25}$, or $\sigma^4 \approx 1/t$. In a similar way, the flow stress of a subgrain structure may be related to the annealing time, however, with the further complication that both the mobility of the subgrain boundaries and the driving pressure depend on the angle across the subgrain boundaries.

Ultrafine Structures

In most annealing studies, the initial deformed microstructure has been formed by deformation to low and medium strains. However, recently there has been a growing interest to explore the behavior of metals deformed to high and ultrahigh strains, for example, more than 90% reduction in thickness by rolling. The reason is that the microstructure refines with increasing strain and as a microstructural refinement induces an increase in strength, it is the intention to produce very strong metals by only applying plastic deformation, for example, to replace traditional alloying. A large strain leads to an increase in the dislocation density but also to the formation of deformation induced high-angle boundaries, which are indistinguishable from the original grain boundaries in the deformed structure (**Figure 9**). For these structures, subdivided by a mixture of dislocation boundaries and high-angle boundaries, the flow stress can be expressed by an equation of the same form as the Hall–Petch relation:

$$\sigma_f - \sigma_0 = k_2 D_b^{-0.5} \qquad [15]$$

where k_2 is a constant and D_b is the average boundary spacing. However, the constancy of k_2 assumes that the boundaries are identical as in a polycrystalline metal, whereas for structures, which are a mixture of dislocation boundaries and high-angle boundaries, k_2 is not a constant. Instead of eqn [15], eqn [6] may be used, where S_v and θ are the surface area and angle, respectively of low-angle boundaries, and D is the random spacing between high-angle boundaries.

The presence of a very fine microstructure and a large concentration of high-angle boundaries in a deformed structure significantly affects its annealing

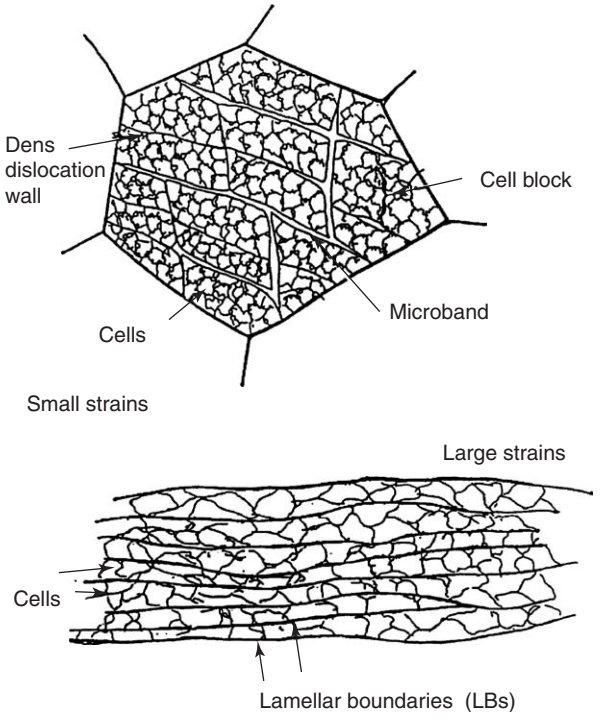


Figure 9 Schematic drawing of deformation microstructures and grain subdivision illustrating the structural refinement with increasing plastic strain. The structure is in both cases subdivided by extended boundaries and short interconnecting boundaries. The average spacing between the boundaries can at very large strain decrease below 100 nm. (Reproduced from Hughes DA and Hansen N (1997) High-angle boundaries formed by grain subdivision mechanisms. *Acta Materialia* 45: 3871–3886, with permission from Elsevier.)

behavior. Both the driving pressure and the boundary mobility are increased. Therefore, the migration rate of boundaries is increased in accordance with eqn [11]. As a result, both the rate of recovery and of recrystallization will be enhanced. The concentration of high-angle boundaries may also reach a level, where small subgrains or nuclei will be so abundant that their growth will stop early due to impingement, that is, a very fine grain structure may result. A route has thereby been established to produce metallic materials with a grain size of the order of a few micrometers or less. Such materials will have a very high strength as experiments show that strength–structure relationships as eqns [2] and [3] also are applicable at such small grain sizes.

Conclusion and Outlook

The change in mechanical properties during recovery, recrystallization, and grain growth has been and still is an important research area driven by scientific curiosity and industrial demand. The research and development at present concentrates on exploration

of mechanisms and formulation of models for the structural evolution. In parallel, advanced and new experimental methods are under development for fast and precise structural characterization in 2D and 3D over length scales from the macroscopic range to the atomic dimension. This research and development also expands the experimental and theoretical basis for the development of property structure models for the evolution of mechanical properties during annealing of deformed metals.

See also: Alloys: Aluminum; Alloys: Copper; Alloys: Iron; Alloys: Magnesium; Alloys: Titanium; Deformation Processing; Mechanical Properties: Plastic Behavior.

PACS: 61.72.Bb; 61.72.Cc; 61.72.Dd; 61.72.Ff; 61.72.Lk; 61.72.Mm; 61.72.Hm; 62.20.Fe; 81.10.Aj; 81.10.Jt; 81.40.Ef; 82.60.Nh

Further Reading

Bever MB (1957) *Thermodynamics of Kinetics and Recovery*, Proceedings Seminar on Creep and Recovery. Cleveland: American Society for Metals.

Cotterill P and Mould PR (1976) *Recrystallization and Grain Growth in Metals*. Surrey University Press.

Haessner F (ed.) (1978) *Recrystallization of Metallic Materials*, 2nd edn. Stuttgart: Dr. Riederer Verlag.

Hansen N (2001) New discoveries in deformed metals. *Metallurgica and Materials Transactions* 32A: 2917–2925.

Hansen N (2004) *Hall–Petch Relation and Grain Boundary Strengthening*. Viewpoint set no. 35. Metals and alloys with a structural scale from the micrometer to the atomic dimension. *Scripta Materialia* 51: 751–841.

Hansen N, Huang X, Juul Jensen D, Lauridsen EM, Leffers T, Pantleon W, Sabin JJ, and Wert JA (eds.) (2000) *Recrystallization-fundamental Aspects and Relations to Deformation Microstructures*, Proceedings of the 21st Risø International Symposium on Materials Science. Denmark: Risø National Laboratory, Roskilde.

Hansen N, Jones AR, and Leffers T (eds.) (1980) *Recrystallization and Grain Growth of Particle Containing Materials,* 1st Risø International Symposium on Metallurgy and Materials Science. Denmark: Risø National Laboratory, Roskilde.

Himmel L (ed.) (1963) *Recovery and Recrystallization of Metals*. London: Interscience.

Hu H, Rath BB, and Vandermeer R (1990) In: Chandra T (ed.) *A Historical Perspective and Overview of the Annealing Studies of Cold-Worked Metals in Recrystallization*, 3–16. The Minerals, Metals and Materials Society.

Hughes DA and Hansen N (2001) Graded nanostructures produced by sliding and exhibiting universal behavior. *Physical Review Letters* 87: 135503-1–135503-4.

Humpreys FJ and Hatherly M (2004) *Recrystallization and Related Annealing Phenomena*, 2nd edn. Oxford: Elsevier.

Nes E (1995) Recovery revisited. *Acta Metallurgica Mater* 43: 2189–2207.

Perryman ECW (1957) *Proceedings Seminar on Creep and Recovery*. Cleveland: American Society for Metals.

(1966) *Recrystallization, Grain Growth and Textures*. Metals Park, OH: American Society for Metals.

Recursive Methods for Electronic States

R Haydock, University of Oregon, Eugene, OR, USA

Recursive Approaches

A wide variety of recursive methods are used to calculate electronic properties of systems to which band-structure methods do not apply due to lack of periodicity, for example surfaces or other defects and disordered systems. These methods range from Gram–Schmidt orthogonalization, to the method of moments, modified moments, orthogonal polynomial expansions such as Chebyshev expansions, the Lanczos method, including the conjugate gradient method, and the recursion method (also related to the Mori projection method). In their simplest applications, recursive methods determine densities of states from path counting, for example, a tight-binding *s*-band on a square lattice. The common ingredients of recursive methods are: they start with some initial state of physical significance and apply the Hamiltonian repeatedly to it generating a sequence of states, which span the time evolution of the initial state; and they use the normalizations and overlaps between states in the sequence to construct expansions, such as continued fraction expansions of electronic Green functions, from which physical properties are calculated. Some of the advantages of recursive approaches are efficient computing because they are usually formulated in position space making Hamiltonians sparse, and the well-developed mathematical theory of moments. One of the main challenges for these methods is obtaining reliable information about singularities in the densities of states for macroscopic systems.

Path Counting

Calculation of the density of states for a tight-binding *s*-band on a square lattice, shown in **Figure 1**, illustrates the basic ideas of recursive methods. The density of states $n(E)$ is defined as the volume-normalized trace, $\mathrm{tr}\{\delta(E-\boldsymbol{H})\} = \sum \delta(E-E_n)/V$, where $\boldsymbol{H}$ is the Hamiltonian operator, E is the

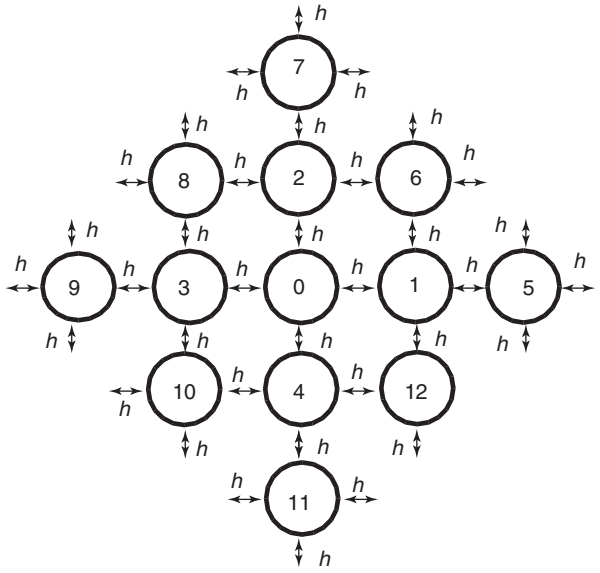


Figure 1 A tight-binding *s*-band on a square lattice with hopping matrix elements *h* between nearest-neighbor orbitals.

energy, δ is the Dirac delta distribution, and the sum is over the eigenvalues $\{E_n\}$ of $\boldsymbol{H}$ normalized by the volume V. The nth power moment μ_n of $n(E)$ is $\mathrm{tr}\{E^n\delta(E-\boldsymbol{H})\}$, or $\mathrm{tr}\{\boldsymbol{H}^n\}$, and since the orbitals on all sites are equivalent, this is $\langle 0|\boldsymbol{H}^n|0\rangle$, where the unit volume is taken to be the area of the unit cell and $|0\rangle$ is the orbital at the origin.

Evaluation of the moments can be reduced to path counting by inserting a resolution of the identity, the sum $\Sigma|\alpha\rangle\langle\alpha|$ over a complete, orthonormal set of orbitals $\{|\alpha\rangle\}$, between each factor of $\boldsymbol{H}$ in $\langle 0|\boldsymbol{H}^n|0\rangle$ to give

$$\mu_n = \sum \langle 0|\boldsymbol{H}|\alpha_1\rangle\langle\alpha_1|\boldsymbol{H}|\alpha_2\rangle\langle\alpha_2|\boldsymbol{H} \cdots \boldsymbol{H}|\alpha_{n-1}\rangle\langle\alpha_{n-1}|\boldsymbol{H}|0\rangle \quad [1]$$

where the sum is over all the orbitals $|\alpha_1\rangle$, $|\alpha_2\rangle, \ldots, |\alpha_{n-1}\rangle$. Each term in this sum is the product of matrix elements along a path of n hops starting and finishing at $|0\rangle$. In the method of moments, paths are enumerated giving 1, 0, $4h^2$, 0, and $36h^4$, for moments μ_0 through μ_4, respectively, where the nearest-neighbor hopping elements of $\boldsymbol{H}$ are h and the diagonal elements are zero. The problem of moments is to construct a positive density of states from these moments.

The classical solution to the moment problem is a continued fraction $b_0^2/E - a_0 - b_1^2/E - a_1 - b_2^2/E - a_2 - \cdots - b_n^2/E - a_n - \cdots$, where the slash "/" means division by everything to the right, and the parameters $\{a_n, b_n\}$ *are* constructed so that as E approaches the real axis from below, the imaginary part of the continued fraction is an approximate density of states, which reproduces a number of moments equal to the number of parameters in the continued fraction. While moments are useful for analysis, they are difficult numerically because they must be computed with exponentially increasing accuracy to maintain a given accuracy in the resulting density of states.

The solution to this numerical instability is to calculate the continued fraction parameters directly by constructing an orthonormal basis u_0, $u_1, u_2, \ldots u_n, \ldots$, in which the Hamiltonian is a tridiagonal matrix, equivalent to the three-term recurrence

$$\boldsymbol{H}u_n = a_n u_n + b_{n+1}u_{n+1} + b_n u_{n-1} \quad [2]$$

For the above example and referring to **Figure 1**: u_0 is $|0\rangle$, u_1 is $(|1\rangle + |2\rangle + |3\rangle + |4\rangle)/2$, and u_2 is $(|5\rangle + 2|6\rangle + |7\rangle + 2|8\rangle + |9\rangle + 2|10\rangle + |11\rangle + 2|12\rangle)/\sqrt{20}$. This gives the parameter values $a_0 = a_1 = 0$, $b_0 = 1$, $b_1 = 2h$, and $b_2 = \sqrt{5}h$.

The Projected Density of States

The electronic structure of disordered and defective electronic systems is difficult to calculate because such systems have a macroscopic number of inequivalent orbitals, and the interactions between these orbitals exceed their energy differences, and macroscopic quasi-degeneracy. Band-structure methods cannot be used because the unit cells are macroscopic, and perturbation methods cannot be used because of the small differences in the energies of the orbitals. The main effect of quasi-degeneracy is that individual states are unstable in the sense that the interactions between two such states, which are close in energy, can mix them arbitrarily.

Densities of states are more stable than those of individual states, because the only states which get strongly mixed are close in energy. However, approximate total densities of states for finite systems only converge slowly to their macroscopic limits as the ratio of surface area to volume. The density of states, which converges fastest with increasing size, is the projected density of states

$$n_0(E) = \langle 0|\delta(E-\boldsymbol{H})|0\rangle \quad [3]$$

where $\boldsymbol{H}$, is the Hamiltonian and $|0\rangle$ is the spatially localized orbital on which the states are being projected. Physically $n_0(E)$ is the probability distribution for finding an electron with energy E in the orbital $|0\rangle$. In analogy with the blackbody properties of electromagnetic spectra, $n_0(E)$ converges in the mean (integrals over smooth functions) with increasing system size, and the error at E decreases exponentially with the number of wavelengths of the state at E between $|0\rangle$ and the boundary.

Recursive Tridiagonalization

The projected density of states $n_0(E)$ is simply the Fourier transform of the time evolution (taking time in units of inverse energy) of the orbital $|0\rangle$ projected back onto $|0\rangle$, $\langle 0|\exp\{-\mathrm{i}t\boldsymbol{H}\}|0\rangle$, which can be expanded in the power moments to give $\sum(-\mathrm{i}t)^n\mu_n/n!$. This series shows that if the moments of a physical system grow faster with n than $n!$, the evolution of $|0\rangle$ is singular at $t=0$. This also shows that truncation of the series at some power of t makes the high moments zero producing an $n_0(E)$, which is nonzero only at $E=0$.

Recursive approaches are based on truncating the moment expansion in state space, rather than powers of t. The nth truncated state space or Krylov subspace is spanned by $|0\rangle$ and the first n powers of $\boldsymbol{H}$ acting on $|0\rangle$: the power states, $|0\rangle, \boldsymbol{H}|0\rangle, \boldsymbol{H}^2|0\rangle, \ldots, \boldsymbol{H}^n|0\rangle$. Recursive methods differ in the choice of vectors used to span the Krylov subspaces and the parameters generated. In the method of moments, the spanning states are just the above power states, and the parameters are the overlaps between them; these overlaps are the power moments, and there are only $2n+1$ of them because $\boldsymbol{H}$ is Hermitian making overlaps equal between pairs of states whose powers of $\boldsymbol{H}$ add up to the same total. In modified moment methods, the above power states are replaced by polynomial states, the results of polynomials in $\boldsymbol{H}$ acting on $|0\rangle$.

The Lanczos and recursion methods are special cases of modified moments in which Gram–Schmidt orthogonalizatio n is applied to the power states, tridiagonalizing $\boldsymbol{H}$ recursively, in the basis of orthonormal polynomial states $\{p_n(\boldsymbol{H})0\rangle\}$. These polynomials are also orthonormal with respect to integration over the projected density of states $\delta_{n,m} = \int p_n(E)p_m\ (E)n_0(E)\ \mathrm{d}E$, as can be seen by substituting $\langle 0|\delta(E-\boldsymbol{H})|0\rangle$ for $n_0(E)$ and replacing E by $\boldsymbol{H}$ in the polynomials. Since $\boldsymbol{H}$ is Hermitian in the inner product between states, the only nonzero terms produced by Gram–Schmidt orthogonalization are the three terms in eqn [2], corresponding to a tridiagonal representation of $\boldsymbol{H}$ in which the diagonal matrix elements are the $\{a_n\}$ and the off-diagonal matrix elements are the $\{b_n\}$.

Analytic systems of orthogonal polynomials are known for many model projected densities of states $n_0(E)$. As a result, combinations of $|0\rangle$ and $\boldsymbol{H}$ which have such $n_0(E)$ can be tridiagonalized analytically by substituting $\boldsymbol{H}$ for E in the polynomials and applying them to $|0\rangle$; or they can be tridiagonalized perturbatively if $n_0(E)$ differs only slightly from one of the analytic models. When $n_0(E)$ does not have analytic polynomials, $\boldsymbol{H}$ can be tridiagonalized numerically using eqn [2] to construct u_{n+1} from u_n and u_{n-1}, with a_n determined by orthogonality and b_{n+1} determined by the normalization. The first step of most diagonalization methods is tridiagonalization, and recursive tridiagonalization is very efficient when $\boldsymbol{H}$ is short-ranged, because the zero-valued matrix elements of $\boldsymbol{H}$ are not changed by the transformation. For systems in D spatial dimensions, the number of arithmetic operations required to construct $u_0, \ldots, u_n$ and the associated parameters only increases as n^D. This scaling is independent of the size of the system because the projected density of states is a local rather than global quantity.

It is important to note that the recursively generated states, whether power states, polynomial states, or the tridiagonal basis of orthonormal states, are generally not a complete set of states. The recursive states span only the evolution of the starting state $|0\rangle$; so if there is a symmetry or any other property of the system such as disorder, which prevents the evolution of $|0\rangle$ from exploring all the states of the system, then the recursive basis is not complete. Because the recursive states span only the evolution of a single state, they may be interpreted as an abstract, one-dimensional subsystem of the original physical system, a semibounded chain model illustrated in **Figure 2**. This abstract system evolves along the sequence $\{u_n\}$ with matrix elements $\{a_n\}$ and $\{b_n\}$. It is one-dimensional, semibounded, and hence nondegenerate, because the starting state $|0\rangle$ contains at most one state from each degenerate subspace of the states of the original system.

Another important property of recursively generated states is that they can be linearly dependent. A simple example is the sequence of power states when $|0\rangle$ is an eigenstate or a mixture of just a few eigenstates. The basis $\{u_n\}$ is orthonormal if $\boldsymbol{H}$ is exactly Hermitian and if the orthonormalization in eqn [2] is carried out exactly, but neither is the case when the recursion is numerical. As in all recursive processes, errors grow exponentially with n in u_n, a_n, and b_n; however, it has been shown that the errors introduced into the projected densities of states or related projected quantities calculated from the $\{u_n\}$, $\{a_n\}$, and $\{b_n\}$ do not accumulate and are comparable to the error in a single recursion eqn [2].

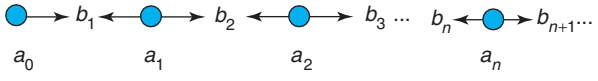


Figure 2 A chain model with site energies $\{a_n\}$ and nearest-neighbor hopping matrix element $\{b_n\}$.

Macroscopic Systems

The main applications of recursive methods are to macroscopic systems, which can only be partly tridiagonalized numerically. The challenge in these cases is to extract the maximum information about a macroscopic system from a small tridiagonal matrix. There are many parts to this problem ranging from whether it is better to approximate the original system by a finite cluster, and including methods for approximating the projected density of states and quantities calculated from it.

Cluster Size and Shape

The electronic Hamiltonian $\boldsymbol{H}$ is usually represented in terms of localized orbitals so that each orbital has nonzero matrix elements with only a few neighboring orbitals. A numerical calculation is generally limited by the total number of these orbitals which can be included, and there is often a trade-off between the number of orbitals included and the size of the tridiagonal matrix, which can be produced.

Viewing the tridiagonal matrix as a chain model (**Figure 2**), the blackbody property of $n_0(E)$ explains that the deeper errors occur in the chain, the exponentially smaller their effect on $n_0(E)$. From this it seems best to use the largest possible power cluster, the cluster of orbitals which completely contains as many power states as possible, because this gives the largest tridiagonal matrix containing only rounding errors. This simple conclusion has to be modified by consideration of subsequent approximations in the calculation. If the resulting tridiagonal matrix is going to be extended with approximate matrix elements, then an abrupt change from the numerical matrix elements to approximate matrix elements can produce unstable results. Consequently, in many calculations the changeover from numerical to approximate matrix elements is smoothed by using clusters with highly disordered boundaries, very different from the typically flat boundaries of power clusters.

Projected Densities of States

The central problem in applications to macroscopic systems is that of approximating the projected density of states, particularly its singularities. This is difficult because at singularities the blackbody property of $n_0(E)$ breaks down, the wavelengths of eigenstates go to infinity, and the spectrum becomes very sensitive to distant parts of the system. The simplest approach to this problem is through the continued fraction

$$R(z) = b_0^2/z - a_0 - b_1^2/z - a_1 - b_2^2/z - a_2 - \cdots - b_n^2/z - a_n - \cdots \quad [4]$$

which is of Jacobi form because the $\{a_n\}$ and $\{b_n\}$ are all real, and hence has singularities which are nonnegative imaginary and only occur on the real z-axis. The projected density of states $n_0(E)$ is then just $\mathrm{Im}\{R(E - \mathrm{i}\varepsilon)\}/\pi$ in the limit as ε goes to zero from the positive side.

The three main approaches to approximate the projected density of states for macroscopic systems are bounding, averaging, and continuing. While $n_0(E)$ itself cannot be bounded because it only converges in the mean, there are rigorous bounds on integrals $F = \int n_0(E) f(E)\, \mathrm{d}E$ for smooth $f(E)$ as well as for some singular $f(E)$. This is important because physical quantities can be expressed as integrals over $n_0(E)$: for example, the occupation of $|0\rangle$ is the integral of $n_0(E)$ over the Fermi function, defined by $f(E)$ is 1 for $E < E_\mathrm{F}$ and 0 for $E > E_\mathrm{F}$. The bounds on integrated quantities such as the occupation can then be used to estimate $n_0(E)$, in this example by the derivative of the occupation with respect to E_F for $E_\mathrm{F} = E$. This method has the advantage that it does not introduce spurious structure into $n_0(E)$, but the disadvantage is that it does not reproduce the exact moments when $n_0(E)$ is integrated.

The simplest form of averaging is to approximate $n_0(E)$ by $\mathrm{Im}\{R(E - \mathrm{i}\varepsilon)\}/\pi$, where ε is positive so that singularities within about ε of E are averaged into the approximation for $n_0(E)$. This is equivalent to replacing δ-distributions on the real E-axis with functions broadened by 2ε. A more sophisticated form of averaging is to maximize the entropy functional $-\int n_0(E) \ln[n_0(E)]\, \mathrm{d}E$, where ln is the natural logarithm and subject to the constraint that $n_0(E)$ has the correct moments. This produces an approximate density of states of the form $\exp\{-\mathrm{P}(E)\}$, where $\mathrm{P}(E)$ is a polynomial in E of the same degree as that of the highest moment used as a constraint.

Another way of looking at the problem is that the part of $R(E)$ in eqn [4] is known and the rest is estimated. In certain cases the singularities of $n_0(E)$ are known, e.g., band edges. It is the singularity of $n_0(E)$, which determine the behavior of $\{a_n\}$ and $\{b_n\}$ for large n; so, this known asymptotic behavior can be used to continue the numerical part of $R(E)$. Results are greatly improved by smoothing the transition from the numerical part to the asymptotic part of the continued fraction.

In many cases only a small part of the tridiagonalization of $\boldsymbol{H}$ is known numerically and it would be best to locate singularities and other structure in $n_0(E)$. This problem could be considered in terms of a chain model **Figure 2** for which the parameters only up to b_N are known. For this model, $n_0(E)$ is simply the squared amplitude of the state on site zero at energy E, for which all the parameters of the model

need not be known; just the boundary condition at site N on the chain is required. Experience in trying different boundary conditions shows that wrong boundary conditions introduce spurious reflections, which in turn produce spurious structure in $n_0(E)$. Consistent with this, a good choice of boundary condition seems to be the one which produces the least reflection, the perfectly absorbing boundary at site N.

The perfectly absorbing boundary can be defined as the one which does not reflect a wave traveling to the right along the chain. In order to implement this definition, the current $J_n = -b_n \, \mathrm{Im}\{\psi_n^* \psi_{n-1}\}$ carried by a complex wave from site $n-1$ to site n needs to be calculated. For waves ψ, which are eigenstates of the Hamiltonian for the chain, this current is conserved and $J_n = J$, independent of n, a property also known as the Christoffel–Darboux identity. The probability p_n of finding the system on the nth site is $\psi_n^* \psi_n$, and the probability P_N of finding it on one of the first N sites is the sum of p_n from $n=0$ to N, which is related to the Christoffel function. The wave which travels to the right with the least reflection is the one which minimizes P_N for a fixed, positive J. **Figure 3** shows approximate $n_0(E)$ for the tight-binding s-band on a square lattice that **Figure 1** calculated this way for chains of different lengths.

Binding Energies

The definition of the binding energy of an atom to a solid is the difference between the total energy of the solid without the atom and with it, the difference between two large numbers, each of which depends on the electronic structure of the whole system. However, binding the atom only changes the electronic structure of the solid locally, so there is a local way to calculate it. The simplest example is the binding energy of a single orbital $|0\rangle$ to the other orbitals in a system, the change in the total energy of the system due to removing the matrix elements of the Hamiltonian between $|0\rangle$ and the other orbitals. Binding energies of more complicated systems can be expressed as the sum of the binding energies as individual orbitals are decoupled from the system, one by one.

The key to the local calculation of binding energies is the observation that in the simplest case, decoupling $|0\rangle$ from the rest of the system only changes the energies of states which evolve from $|0\rangle$, the states spanned by the chain model in **Figure 2**. At zero temperature the states below the Fermi level E_F are doubly occupied by electrons of opposite spin, and those above E_F are unoccupied, so the binding energy is twice the sum of the shifts in the energies of the occupied states when $|0\rangle$ is decoupled from them. Since $\langle 0|\delta(z-\boldsymbol{H})|0\rangle$ can be represented as the limit of $\mathrm{Im}\{\langle 0|(z-\boldsymbol{H})^{-1}|0\rangle\}/\pi$ as z becomes real, $R(z)$ is $\langle 0|(z-\boldsymbol{H})^{-1}|0\rangle$, and in terms of determinants this inverse element is $\det|z-\boldsymbol{H}'|/\det|z-\boldsymbol{H}|$, where $\boldsymbol{H}'$ is $\boldsymbol{H}$ with the row and column for $|0\rangle$ each deleted. From this it can be shown that the binding energy U is

$$U = \mathrm{Im}\left\{\int_F \ln R(z)\, dz\right\}/\pi \qquad [5]$$

where the integral is around the Fermi contour which encloses the energies of the occupied states, and ln $R(z)$ is the natural logarithm of the continued fraction, single-valued because the Fermi contour avoids zeros and infinities of $R(z)$.

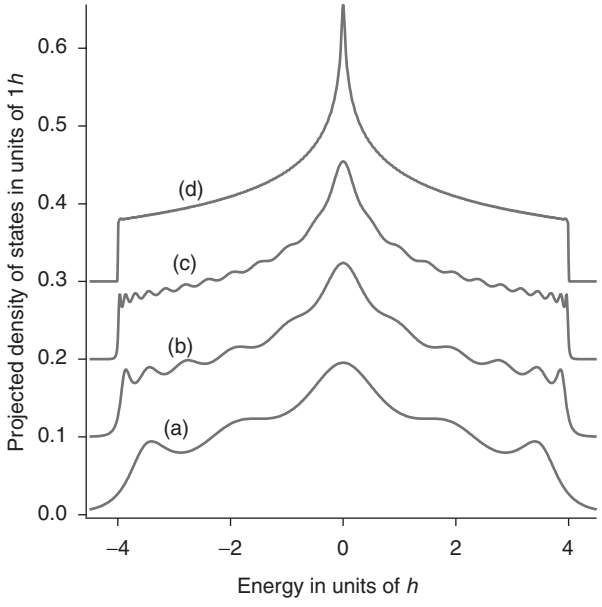


Figure 3 Projected density of states for the tight-binding s-band on a square lattice calculated using (a) 5, (b) 11, (c) 23, and (d) 199 pairs of tridiagonal matrix elements.

See also: Disordered Solids and Glasses, Electronic Structure of; Electronic Structure Calculations: Plane-Wave Methods; Electronic Structure Calculations: Scattering Methods; Insulators, Impurity and Defect States in; Metals and Alloys, Impurity and Defect States in; Quasicrystals, Electronic Structure of; Semiconductors, Impurity and Defect States in; Tight-Binding Method in Electronic Structure.

PACS: 71.15.Dx; 02.70.Hm

Further Reading

Akhiezer NI (1965) *The Classical Moment Problem and Some Related Questions in Analysis*. New York: Kemmer, Hafner.

Baker GA (1996) *Padé Approximants*. Cambridge: Cambridge University Press.

Bullett DW, Haydock R, Heine V, and Kelly MJ (1980) In: Ehrenreich, Seitz, and Turnbull (eds.) *Solid State Physics*, vol. 35. New York: Academic Press.
Chihara TS (1978) *An Introduction to Orthogonal Polynomials*. New York: Gordon and Breach.
Cullum J and Willoughby RA (1985) *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*. Boston: Birkhäuser.
Cyrot-Lackmann F (1970) Application de la Méthode des Moments aux Surfaces. *J. Phys. (Paris)*, suppl. C1, pp. 67–74.
Kittel C (1987). *Quantum Theory of Solids*, 2nd edn., pp. 338–367. New York: Wiley.
Lorentzen L and Waadeland H (1992) *Continued Fractions with Applications*. Amsterdam: North-Holland.
Pettifor DG and Weaire DL (1985) *The Recursion Method and its Applications*. Berlin: Springer.
Press WH, Flannery BP, Teukolsky SA, and Vetterling WT (2002) *Numerical Recipes in FORTRAN (C++): The Art of Scientific Computing*. Cambridge: Cambridge University Press.
Shohat JA and Tamarkin JD (1943) *The Problem of Moments*. New York: American Mathematical Society.
Silver RN and Roder H (1997) Calculation of densities of states and spectral functions by Chebyshev recursion and maximum entropy. *Physical Review E* 56: 4822–4829.
Viswanath VS and Müller G (1994) *The Recursion Method: Application to Many-Body Dynamics*. Berlin: Springer.

Red Blood Cells, Physical Properties of

A Bifone, GSK Medicines Research Centre, Verona, Italy

Introduction

Erythrocytes (also known as red blood cells) are highly specialized cells whose primary function is to transport oxygen from the lungs to the tissues, an essential process to sustain human life. The physical properties of red blood cells (RBCs) appear to have been optimized via evolutionary processes. Their biconcave shape increases the cell surface area, facilitating gas exchange with the extracellular environment. The high physical flexibility of RBCs allows them to squeeze through tiny capillaries. Most importantly, red blood cells are packed with hemoglobin, an iron-containing protein that reversibly binds oxygen. Under physiological conditions, the vast majority of oxygen molecules in the RBCs are bound to hemoglobin. The large concentration of hemoglobin in the RBC cytoplasm, close to crystallization, increases the apparent solubility of oxygen in the intracellular medium, thus improving the efficiency of oxygen delivery to the target tissue. Moreover, hemoglobin affinity for oxygen is modulated by microenvironmental factors such as oxygen and carbon dioxide concentrations, pH, that regulate oxygen delivery to meet the local oxygen demand.

Hemoglobin consists of four polypeptidic units, each containing a heme group with an iron ion in the center (**Figure 1**). When oxygen binds to the heme group in one of the four subunits, the protein undergoes conformational changes that affect the ability of the other subunits to bind oxygen. These changes in quaternary structure are responsible for the characteristic sigmoidal oxygen saturation curve of hemoglobin (i.e., the curve that describes the saturation of the oxygen-binding sites as a function of the partial pressure of oxygen in solution), a behavior referred to in terms of positive cooperativity.

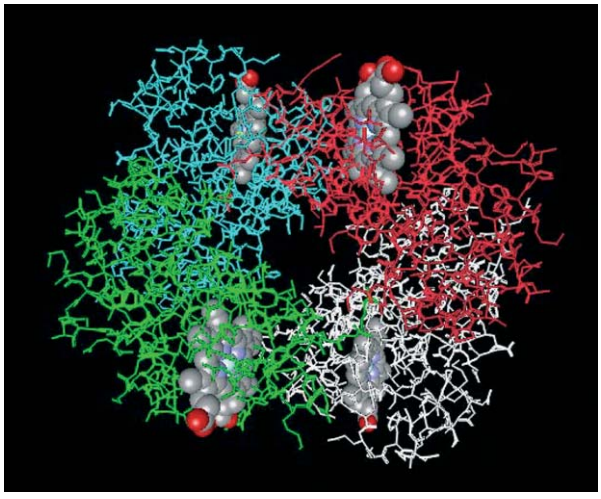


Figure 1 Stick-model of hemoglobin in the fully deoxygenated form. The four subunits are shown in different colors. The atoms of the heme groups are represented as space-filling spheres.

In the deoxygenated form, the iron atom is in a Fe(II) oxidation state, with a total electron spin $S = 2$, and the Fe–heme complex is paramagnetic. Upon oxygen binding, the electronic spin state goes from $S = 2$ to $S = 0$, and the paramagnetic contribution to the protein susceptibility vanishes. As a consequence of the very high concentration of hemoglobin in the cytoplasm, the magnetic susceptibility of the entire RBC is strongly dependent on blood oxygenation.

In this article, the magnetic properties of erythrocytes are reviewed. Besides an interest in the fundamental physical properties of biological materials, several motivations have stimulated experimental and theoretical studies on this subject. A better

understanding of magnetism in living cells is important to assess safety issues related to the exposure of the human body to strong static magnetic fields, like those used for medical diagnostics, or in passenger transport systems based on magnetic levitation. Moreover, the magnetic properties of RBCs have recently found an important application in the study of human brain function. The effects of the blood oxygenation level on RBCs' magnetic susceptibility can be detected noninvasively in living organisms with nuclear magnetic resonance (NMR) methods. This endogenous contrast mechanism is exploited in "functional magnetic resonance imaging" (fMRI), the most powerful technique to date, to map the temporal and spatial patterns of brain activity in conscious human subjects.

The next section describes the magnetic properties of RBCs, their dependence on blood oxygenation levels, and the effects of oxygenation on the field distribution in a suspension of RBCs immersed in a static magnetic field. In the subsequent section, the mechanisms whereby these magnetic field gradients affect NMR signals are discussed. In the last section, the application of this effect to brain function imaging is presented briefly.

RBCs in Static Magnetic Fields

RBCs in a strong static magnetic field tend to orient themselves with the disk plane along the field, a result of the anisotropy of the cell's diamagnetic response. Almost 100% orientation is observed in blood samples exposed to a static field of 4 T. Interestingly, neither the direction nor the degree of orientation depends on the oxygenation state of hemoglobin. This is not surprising though, as the torque that orients the cell is solely determined by the anisotropic component of the susceptibility tensor. Hemoglobin has a small anisotropic susceptibility even in the deoxygenated form. Consequently, the magnetic field-related energy is lower than thermal energy, and monomolecular hemoglobin is not oriented at field strengths up to several tesla. Moreover, the vast majority of hemoglobin molecules are in solution, and free to reorient without transferring torque to the cell body. Thus, hemoglobin does not influence the cell orientation, which can be attributed primarily to the anisotropic susceptibility of membrane components (phospholipids and membrane proteins). Interestingly, sickle RBCs behave quite differently, and orient themselves perpendicular to the magnetic field. In sickle cells most of the hemoglobin is polymerized, and forms bundles of filaments that are integrated with the cell membrane, causing the characteristic cell deformation. In this pathological situation, the anisotropic susceptibility of polymerized hemoglobin is much larger, and determines the orientation of the cell.

The isotropic susceptibility of RBCs is dominated by the contribution of hemoglobin, and is affected by changes in the oxidation state of the iron ion upon oxygen binding. For deoxygenated and oxygenated RBCs, susceptibilities $\chi_{\mathrm{rbc}}^{\mathrm{deoxy}} = -6.7 \times 10^{-6}$ and $\chi_{\mathrm{rbc}}^{\mathrm{oxy}} = -9.2 \times 10^{-6}$ have been measured, respectively. The negative sign of the magnetic susceptibility implies that the cell is diamagnetic even when hemoglobin is in its deoxygenated state. Indeed, the total magnetic susceptibility is the weighted sum of the susceptibilities of all the constituents of the cell. While the iron atom is paramagnetic in its deoxygenated form, the diamagnetic contributions from globins, water, membrane, etc. are still dominant. It is predicted that RBCs would have a paramagnetic susceptibility if all hemoglobin were in the met-state, where the iron ion is oxidized to its Fe(III) form, with a total electron spin $S = 5/2$. However, this situation is not physiologically relevant, since the concentration of met-hemoglobin in blood is normally very low. Met-hemoglobin is potentially harmful for the organism, since it has lost its ability to bind oxygen, and active biological mechanisms are in place to convert the small amount of met-hemoglobin that spontaneously forms in blood to its normal Fe(II) state.

The magnetic susceptibility of the plasma, the fluid in which RBCs are suspended, is $\chi_{\mathrm{plasma}} = -9.06 \times 10^{-6}$. Oxygen molecules are paramagnetic, and changes in oxygen tension affect plasma susceptibility slightly. However, this variation is small compared with the large changes of magnetic susceptibility of the cells, and can be neglected to a first approximation.

Even though both RBCs and plasma are diamagnetic, there is an oxygenation-dependent mismatch between their susceptibilities that causes an inhomogeneous field distribution in the cell suspension. If RBCs had a spheroidal or an ellipsoidal shape, and were sufficiently far apart from each other, the magnetic field would be uniform inside the cell, and nonuniform outside. However, RBCs have a complex, bi-concave shape (**Figure 2**), which results in an inhomogeneous distribution of magnetic field in both regions. The field distribution can be calculated by integrating the Maxwell equations over the surface of discontinuity of magnetic susceptibility. As a particular case, the solution for a spherical surface can be calculated analytically, and corresponds to a uniform field distribution. For more complex shapes, like that of an RBC, the Laplace equation has to be solved numerically. The model shown in **Figure 2** captures the essential features of the geometry of an

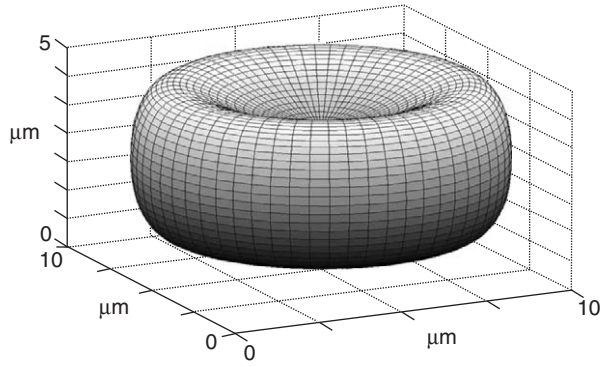


Figure 2 A physicist's model of an RBC.

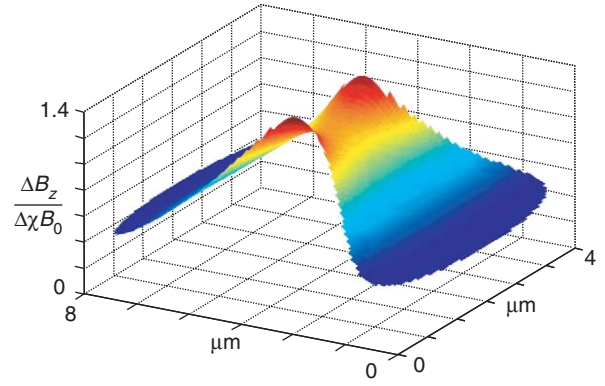


Figure 3 Distribution of the z component of the magnetic field in the plane perpendicular to the disk of an RBC. The applied magnetic field is along the plane of the cell.

RBC, and is described by the simple equations:

$$x = \sin\vartheta\cos\varphi$$
$$y = \sin\vartheta\sin\varphi$$
$$x = \cos\vartheta(0.2 + \sin^2\vartheta)$$

The field distribution inside this model-cell with its plane of symmetry parallel to the external magnetic field is shown in **Figure 3**. In a deoxygenated RBC, the internal field gradients in a static field of 1.5 T are $\sim 250\,\mathrm{mT\,m^{-1}}$. Similarly, large magnetic field gradients are present around the cell. The strength of the gradients increases with the mismatch in susceptibility between the cell and the surrounding medium, that is, with lower levels of blood oxygenation.

Effects of Oxygen Saturation Levels on Nuclear Magnetic Resonance Signals

The NMR frequency of nuclei with a nonzero spin is proportional to the nuclear gyromagnetic ratio, and to the strength of the local field experienced by the nucleus. Thus, changes in field distribution in and around RBCs affect several NMR parameters, including resonance frequency, and NMR relaxation times. These two parameters will be discussed separately in the next few paragraphs. The discussion will be focused on hydrogen nuclei, the most commonly studied nuclear species in biomedical NMR and MRI. The largest part of the hydrogen NMR signals from living organisms comes from unbound water that is present in high molar concentration in many tissues, including blood. High concentration, large gyromagnetic ratio, and high isotopic abundance greatly increase the sensitivity of NMR detection of hydrogen, thus making it possible to obtain high-resolution MR images of living organisms.

The local microscopic field experienced by a nucleus in a homogeneous medium of given susceptibility χ immersed in an externally applied field $\boldsymbol{B}_0$ is

$$\boldsymbol{B}_\mathrm{L} - \boldsymbol{B}_0 = (D_\mathrm{s} - \tfrac{2}{3})(\chi - \chi_0)\boldsymbol{B}_0$$

where D is a demagnetizing factor dependent on sample geometry, and 2/3 is the geometric factor of a sphere drawn around the nucleus. This construct is normally referred to as the "sphere of Lorentz," a notional spherical vacuum cavity surrounding the nucleus that accounts for the fact that, at an atomic scale, the medium cannot be represented as a continuum. It is interesting to note that for a spherical sample $D_\mathrm{s} = 2/3$ and the resonance frequency is independent of the changes of χ. For an RBC, the geometry is far from spherical, and a shift in resonance frequency is to be expected for water molecules in a suspension of RBCs when the cell susceptibility changes. Indeed, a difference of 0.33 ppm in the resonance frequency of the hydrogen nuclei in oxygenated, and deoxygenated human blood has been measured by Thulburn and co-workers in experiments conducted *in vitro* at high magnetic fields. In order to calculate the susceptibility effects on the NMR signals in a complex medium like blood, one needs to refine the theoretical framework.

In a homogeneous material, the actual size of the sphere of Lorentz does not explicitly appear in the calculation. However, blood is highly heterogeneous, with hemoglobin compartmentalized in discrete units (the RBCs) embedded in plasma. A quantitative calculation of the shift of the NMR line of protons in blood requires a slightly more complex construct, to account for the different characteristic length scales of the contributions to the local field arising from distant RBCs and from hemoglobin. A sphere of Lorentz around the nucleus of sufficient size to comprise a large number of erythrocytes is first drawn, so that the contributions to the local field from distant sources around the sphere vary smoothly. The field experienced by the nucleus in the center

of the sphere is

$$\boldsymbol{B}_{\mathrm{L}} = \boldsymbol{B}_0 + (D_{\mathrm{s}} - \tfrac{2}{3})\chi_{\mathrm{blood}}\boldsymbol{B}_0 + \boldsymbol{B}_{\mathrm{i}}$$

where $\boldsymbol{B}_{\mathrm{i}}$ denotes the contributions to the local field from sources inside the sphere, and χ_{blood} is the susceptibility of blood (the sphere is large on a cellular scale, and the external medium appears as a continuum). χ_{blood} depends on the relative volume occupied by RBC V_{rbc} (i.e., the hematocrit):

$$\chi_{\mathrm{blood}} = V_{\mathrm{rbc}} \times \chi_{\mathrm{rbc}} + (V_{\mathrm{rbc}} - 1)\chi_{\mathrm{pl}}$$

In order to evaluate $\boldsymbol{B}_{\mathrm{i}}$, a second sphere of Lorentz, smaller than a cell, but large on a molecular scale is drawn. If this sphere is entirely comprised within the intracellular space:

$$\boldsymbol{B}_{\mathrm{i}} = (D_{\mathrm{rbc}} - \tfrac{2}{3})(\chi_{\mathrm{rbc}} - \chi_{\mathrm{pl}})\boldsymbol{B}_0$$

For the complex shape of an RBC, the geometric factor D_{rbc} has to be evaluated numerically with procedures similar to those used to calculate the field distribution of **Figure 3**, and depends on the orientation of the cell with respect to the external magnetic field. At the high fields used by Thulburn and co-workers, the RBCs are fully oriented in the configuration of minimum energy, and $D_{\mathrm{rbc}} = 0.75$.

Using this construct, the fractional changes in local field experienced by nuclei can be calculated for the plasma and RBC compartments:

$$\Delta B_{\mathrm{pl}}/B_0 = (D_{\mathrm{s}} - \tfrac{2}{3})V_{\mathrm{rbc}}\Delta\chi_{\mathrm{rbc}}$$

$$\Delta B_{\mathrm{rbc}}/B_0 = (D_{\mathrm{s}} - \tfrac{2}{3})V_{\mathrm{rbc}}\Delta\chi_{\mathrm{rbc}} + (D_{\mathrm{rbc}} - \tfrac{2}{3})\Delta\chi_{\mathrm{rbc}}$$

Water molecules diffuse rapidly within and outside the cell, and dynamically average the entire field distribution on the typical NMR timescale. In this fast exchange regime, the NMR signal consists of a narrow resonance, whose center frequency reflects the average field experienced by the nucleus. The resonance shift for the hydrogen nuclei of a water molecule in blood is therefore proportional to the average of the local-field changes in the intra- and extracellular environments, weighted by the relative sizes of the water compartments (0.3 and 0.7, for RBC and plasma, respectively).

For a change in intracellular magnetic susceptibility of $\Delta\chi_{\mathrm{rbc}} = 2.5 \times 10^{-6}$, corresponding to the difference between oxygenated and deoxygenated RBCs, and for a cylindrical blood sample of 40% hematocrit, these calculations predict a shift of the hydrogen resonance of 0.3 ppm, in close agreement with the shift measured by Thulburn and co-workers under similar conditions. The theoretical estimate reported above shows that this shift can be attributed almost entirely to changes in blood magnetic susceptibility, rather than to direct interactions of water molecules with the paramagnetic centers.

A shift in resonance frequency is not the only, nor the strongest effect of the oxygenation-dependent magnetic properties of blood on NMR parameters. The rapid diffusion of water molecules in the field gradients inside and outside the cells, and in the vicinity of small capillaries, contributes to the irreversible loss of phase-coherence of the nuclear spin-system. The characteristic time of this process is termed spin–spin relaxation time, or T_2. For larger blood vessels, the mismatch in susceptibility between blood and the surrounding tissue results in magnetic field gradients on a larger length scale than that sampled by diffusing water molecules. The partially reversible dephasing of spins induced by these more spatially extended gradients is governed by a time constant referred to as T_2^*.

Both relaxation mechanisms are more efficient in the presence of stronger gradients, and T_2 and T_2^* are shorter in deoxygenated blood. It is important to note that changes in blood oxygenation do not only affect the NMR relaxation times of nuclear spins in blood, but also in the surrounding tissue. The vascular network is highly structured, with the smallest capillaries measuring only a few microns in diameter, and separated by a few tens of μm. Therefore, a change in blood oxygenation affects the relaxation times of a large number of spins in both intra- and extravascular compartments, and NMR relaxation times are more sensitive than the relatively small shift in resonance frequency to changes in blood oxygenation levels. MR images can be sensitized to T_2 and T_2^* by using appropriate sequences of radiofrequency pulses, thus using RBC as an endogenous source of contrast. The application of these methods to map the patterns of brain activity is discussed in the next section.

The Blood Oxygenation Level Dependent Effect, and its Application to Mapping Brain Activity

It has been known for a long time that neural activation is accompanied by changes in brain microcirculation. In 1896, Roy and Sherrington reported local increases in blood flow associated with changes in brain activity. Following this pioneering work, many studies have shown that an increase in neuronal electrical activity elicits a complex response, including changes in regional cerebral blood flow, blood volume, and metabolic rate. The net result of this complex response is a time-dependent change in blood oxygenation.

In correspondence to a local increase in brain activity, for instance in response to somatosensory

stimuli or motor tasks, the T_2 and T_2^* in the image voxel comprising the activated area become longer, causing an increase in NMR signal intensity. This phenomenon is known as the "positive" blood oxygenation level dependent (BOLD) effect. In a typical experiment, a time-series of brain MR-images is acquired, and the time-course of signal intensity is analyzed for each pixel to identify the activated brain areas. By way of example, **Figure 4** shows the pattern of brain activity in a healthy subject exposed to an oscillating check-board visual stimulus. More complex paradigms, including cognitive or emotional tasks, can be designed to investigate higher brain functions.

It has been seen in the previous section that longer T_2 and T_2^* correspond to higher levels of blood oxygenation (i.e., to weaker magnetic field gradients in the surroundings of RBCs and blood vessels). Therefore, the positive BOLD effect indicates that the increase in oxygen supply exceeds the increase in oxygen demand by the activated tissue, and more oxygen is left in the bloodstream. Consequently, the level of oxygenation in the venous part of the capillary bed is higher (the arterial blood is already close to saturation, and is not affected). Whether this implies an uncoupling of the hemodynamic and the metabolic responses to changes in brain electrical activity, or simply reflects less efficient oxygen extraction at higher flow rates is still the subject of investigation.

Several research groups have reported a transient decrease in signal intensity immediately following the onset of activation, and before the larger and longer-lasting signal overshoot. This "dip" in the signal intensity time-course is sometimes called the "negative" BOLD effect, and indicates a transient decrease in blood oxygenation. The negative BOLD signal is normally attributed to a transient mismatch between the rapid increase in oxygen consumption following the onset of activity, and the slower hemodynamic response, resulting in a temporary decrease in blood and tissue oxygenation. While the negative BOLD effect is in principle more specific to changes in neuronal activity, and is more spatially localized, its amplitude is smaller than the positive BOLD effect, and the majority of fMRI studies make use of the positive contrast.

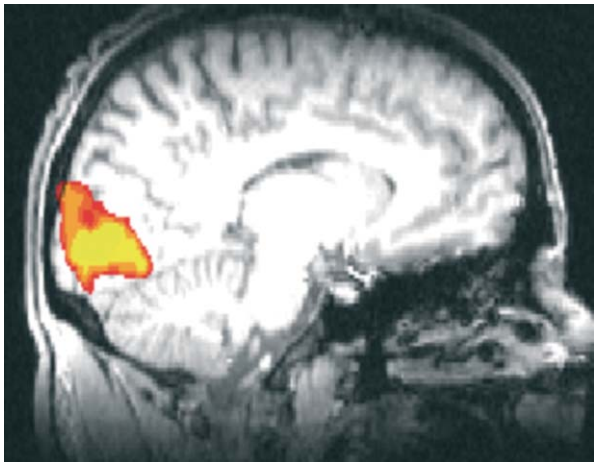


Figure 4 Functional MR image showing the pattern of brain activity elicited by a visual stimulus. (Courtesy of Dr. Peter Chiarelli of the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain.)

Despite more than a century of investigation, the physiological mechanisms underlying the hemodynamic response to changes in brain activity are not completely understood. In particular, the apparent lack of coupling between changes in oxygen consumption and increase in blood flow, ultimately responsible for the BOLD effect, is the subject of much debate in the scientific community. Nevertheless, a decade of fMRI studies has demonstrated the power of this method, which has contributed enormously to our understanding of brain function in healthy subjects and in patients.

Conclusion

RBCs are highly unusual compared to other cells, in so far as their magnetic susceptibility can change over a wide range of values, depending on oxygenation levels. The molecular mechanism of this phenomenon can be attributed to changes in electron-spin state of the heme groups in hemoglobin, the main constituent of RBCs. The unusual magnetic properties of RBC result in a nonuniform field distribution in blood, and in the vicinity of blood vessels, that can be calculated within the framework of classical magnetostatics. These relatively strong magnetic field gradients affect the NMR signals in living tissues and *in vitro* preparations containing RBCs. The calculation of the effects of the field distribution on nuclear spins must take into account the different length scales of the magnetic field gradients, ranging from the cellular scale to the macroscopic scale. Experiments and theoretical calculations show that changes in the oxygenation level for RBCs affect several NMR parameters, including the transverse relaxation times T_2 and T_2^*. Thus, NMR can be QJ;applied to measure local variations in blood oxygenation arising from changes in brain activity. This effect, known as the BOLD effect, is exploited in functional MRI, a powerful method to investigate human brain function. In principle, exogenous paramagnetic or super-paramagnetic contrast agents could also be used to manipulate blood susceptibility, and to measure hemodynamic changes in the brain.

However, the peculiar magnetic properties of red blood cells provide an endogenous source of contrast, thus allowing for totally noninvasive functional studies in human subjects.

Acknowledgments

The author wishes to thank Andrea Cherubini, for his assistance in preparing the figures and for insightful discussions, and Dr. Adam J Schwarz and Prof. Silvio Aime for critically reviewing the manuscript. **Figure 4** was kindly provided by Dr. Peter Chiarelli of the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain.

See also: Electric and Magnetic Fields in Cells and Tissues; Electromagnetic Biological Effects; Magnetic Interactions; Magnetoelasticity; Metalloproteins, Structural Determination of; Neuroscience; Paramagnetism.

PACS: 75.90.+w; 76.60.−k; 87.57.−s; 87.61.−c; 87.90.+y

Further Reading

Buxton RB and Frank LR (1997) A model for the coupling between cerebral blood flow and oxygen metabolism during neutral activation. *Journal of Cerebral Blood Flow Metabolism* 17: 64–72.

Durrant CJ, Hertzberg MP, and Kuchel PW (2003) Magnetic susceptibility: further insights into macroscopic and microscopic fields and the sphere of Lorenz. *Concepts in Magnetic Resonance* 18A(1): 72–95.

Higashi T, *et al.* (1993) Orientation of erythrocytes in a strong static magnetic field. *Blood* 82: 1328–1334.

Jackson JD (1966) *Classical Electrodynamics*. New York: Wiley.

Ogawa S, Lee TM, Kay AR, and Tank DW (1990) Brain Magnetic Resonance Imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the USA* 87: 9868–9872.

Ogawa S, *et al.* (1992) Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with Magnetic Resonance Imaging. *Proceedings of the National Academy of Sciences of the USA* 89: 5951–5955.

Thulburn KR, Watertone JC, Matthews PM, and Radda GK (1982) Oxygenation dependence of transverse relaxation time of water protons in whole blood at high field. *Biochimica et Biophisica Acta* 714: 265–270.

Voet D and Voet JG (1990) *Biochemistry*. New York: Wiley.

Relativity

V Barone, Università del Piemonte Orientale, Alessandria, Italy
P Sodano, Università di Perugia, Perugia, Italy

Introduction

In 1905, a young German physicist, Albert Einstein, published in "Annalen der Physik" an article titled "Zur Elektrodynamik Bewegter Körper" (On the electrodynamics of moving bodies), which marked the birth of "relativity," a revolutionary view of space, time, and motion, and one of the highest achievements of the human mind. The special theory of relativity (the subject of the 1905 paper) was followed ten years later by the general theory of relativity.

Broadly speaking, relativity explores the relationship between the descriptions of physical phenomena carried out by different observers. Special relativity considers only inertial observers; general relativity removes this limitation and accounts also for accelerated and freely falling observers. The crucial difference between the two theories resides in the geometry of space–time: while the space–time of special relativity (the Minkowski space) is flat, the space–time of general relativity is a curved manifold (curvature is indeed the manifestation of gravity).

Special relativity introduces new kinematic and dynamical laws. The Newtonian theory survives as an approximation, valid in the low-velocity limit, that is when the velocities of particles are much smaller than the velocity of light. Maxwell's electromagnetic theory, on the other hand, does not need to be modified, as it is inherently consistent with relativity.

This article is mainly devoted to the special theory of relativity. The general theory, which deals with gravitational phenomena and is less relevant for microphysics and mesophysics, is briefly discussed in the last section.

The Origins of Relativity

In Einstein's own words, the main motivation of special relativity was to eliminate some "asymmetries" in the electrodynamics of moving bodies. These asymmetries emerge when one combines Maxwell's electromagnetism with Newtonian mechanics. The latter is consistent with Galilean relativity, that is with the invariance of mechanical laws under the transformation

$$x' = x - vt \qquad [1]$$

between two inertial frames with constant relative velocity v (an inertial frame is, by definition, a frame in which free particles have constant speed). A fundamental assumption of Newtonian mechanics is

that time is absolute: quoting the "Principia," it "flows equably without relation to anything external." Thus, the Galilean transformation [1] is accompanied by

$$t' = t \qquad [2]$$

From [1] and [2], one obtains the velocity addition formula

$$\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{v} \qquad [3]$$

Now, consider the Maxwell equations. They contain a constant parameter c, which represents the speed of the electromagnetic waves in vacuum (i.e., the speed of light). If velocities (including c) change according to [3], one concludes that Maxwell equations are valid only in a privileged reference frame, traditionally identified with the ether. To Einstein, the existence of the ether and the frame-dependence of electromagnetic equations were unacceptable features of classical physics. He realized that, in order to avoid these problems, it was necessary to give up eqn [2] and introduce a new operational definition of time, based on a fundamental assumption about the velocity of signals: the absolute constancy of c.

The Postulates of Special Relativity

Galileo Galilei (in his "Dialogue concerning the two chief systems of the world") was the first who noticed that the two frames in uniform relative motion are physically indistinguishable. Einstein raised this empirical observation to the status of a universally valid principle, applicable to all physical phenomena (not only to mechanics). According to the "principle of relativity," the laws of physics have the same form in any inertial frame. In other terms, the inertial frames are physically equivalent and no experiment can discriminate them. The second pillar of special relativity, as already mentioned, is the "postulate of the constancy of the velocity of light": the velocity of light in vacuum, c, has the same value in any inertial frame, independent of the motion of the source. These two postulates lead to transformation laws between inertial frames, which are quite different from the Galilean transformation [1] and mix time and space variables.

Lorentz Transformations

Consider two inertial frames $\mathscr{K}(x, y, z, t)$ and $\mathscr{K}'(x', y', z', t')$, and it is supposed that their origins O and O' coincide at $t' = t = 0$. The transformation laws between $\mathscr{K}$ and $\mathscr{K}'$ must be linear, so that a uniform motion in $\mathscr{K}$ transforms into a uniform motion in $\mathscr{K}'$. If a spherical light wave is emitted at $t = t' = 0$ from $O \equiv O'$, the equation of its front is

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \qquad [4]$$

in $\mathscr{K}$, and

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \qquad [5]$$

in $\mathscr{K}'$ (recall that the velocity of light is the same). Thus,

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = \lambda(x^2 + y^2 + z^2 - c^2t^2) \qquad [6]$$

and, due to the symmetry of $\mathscr{K}$ and $\mathscr{K}'$ as postulated by the principle of relativity, λ must be equal to unity:

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2 \qquad [7]$$

From this invariance condition, one can derive the coefficients of the transformation, relating $\mathscr{K}$ to $\mathscr{K}'$ (the "Lorentz transformation" (LT)). For the configuration of **Figure 1** ($\mathscr{K}'$ and $\mathscr{K}$ are assumed to have parallel axes, and $\mathscr{K}'$ is supposed to move relative to $\mathscr{K}$ with velocity v along the positive x direction), the transformation laws are

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(t - vx/c^2) \end{aligned} \qquad [8]$$

where γ is the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - (v^2/c^2)}} \qquad [9]$$

It is to be noted that, for $v \to c$, the γ factor goes to ∞ and hence the LTs become meaningless. Therefore, c is a limiting speed. In the formal limit $c \to \infty$, that is for velocities much smaller than c (the Newtonian limit), eqns [8] reduces to the Galilean transformations [1] and [2].

Let $\boldsymbol{u} = \mathrm{d}\boldsymbol{x}/\mathrm{d}t$ be the velocity of a particle in $\mathscr{K}$ and $\boldsymbol{u}' = \mathrm{d}\boldsymbol{x}'/\mathrm{d}t$ its velocity in $\mathscr{K}'$. Differentiating

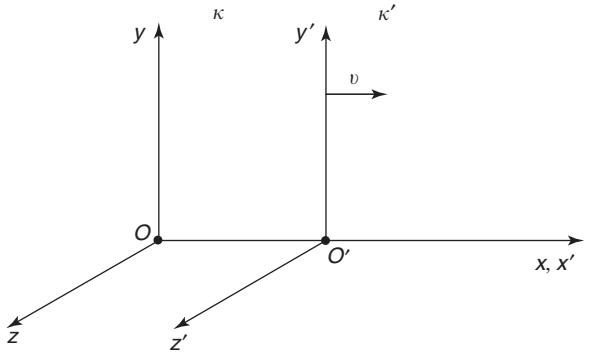


Figure 1 Two inertial frames in uniform relative motion along the x-axis.

eqn [3], one obtains the relativistic transformation of velocities:

$$u'_x = \frac{u_x - v}{1 - (u_x v/c^2)}, \qquad u'_y = \frac{u_y}{\gamma(v)[1 - (u_x v/c^2)]}$$
$$u'_z = \frac{u_z}{\gamma(v)[1 - (u_x v/c^2)]} \quad [10]$$

In particular, if $u_x = c$, the first of these relations gives $u'_x = c$, consistently with the postulate of the constancy of the velocity of light.

An LT between frames with parallel axes is called a "boost." Thus, [8] is a boost along x. A boost along an arbitrary direction is given by

$$\boldsymbol{x}' = \boldsymbol{x} + [(\gamma - 1)\boldsymbol{v} \cdot \boldsymbol{x}/v^2 - \gamma t]\boldsymbol{v}$$
$$t' = \gamma(t - \boldsymbol{v} \cdot \boldsymbol{x}/c^2) \quad [11]$$

The most general (homogeneous) LT combines a boost with a rotation of the axes.

Measurements of Time Intervals and Lengths

An immediate consequence of LTs is the relativity of simultaneity: two events A and B which are simultaneous in an inertial frame $\mathscr{K}$, $t_A = t_B = t$, are not simultaneous in another frame $\mathscr{K}'$. According to [8], their time interval in $\mathscr{K}'$ is, in fact,

$$\Delta t' \equiv t'_B - t'_A = \gamma \frac{v}{c^2}(x_A - x_B)$$

It is supposed that two events occur at the same point in a given inertial frame $\mathscr{K}'$. Let $\Delta\tau$ be their time interval measured by a clock at rest in $\mathscr{K}'$ (τ is called the "proper time"). The time interval Δt in another inertial frame $\mathscr{K}$ moving with velocity v relative to $\mathscr{K}'$ is given by

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - (v^2/c^2)}} \quad [12]$$

and is larger than $\Delta\tau$. This phenomenon is known as "time dilatation." It is worth noting that, whereas in $\mathscr{K}'$ the two events occur at the same point, in $\mathscr{K}$ they occur in two different points, and one needs two clocks to measure their time interval Δt.

Another important relativistic effect is the "length (or Lorentz–Fitzgerald) contraction". Consider a rod at rest in a certain frame $\mathscr{K}'$, and let L_0 be its length in $\mathscr{K}'$ (i.e., its proper length). An observer in another inertial frame $\mathscr{K}$ moving with velocity v relative to $\mathscr{K}'$ measures a length L which is smaller than L_0:

$$L = \sqrt{1 - (v^2/c^2)} \quad L_0 < L_0 \quad [13]$$

This phenomenon can be understood by observing that the measurements of the ends of the rod, which are simultaneous in one of the two frames, are not simultaneous in the other. Since the transverse coordinates do not change under a boost, the transformation law of volumes is the same as [13]: $V = \sqrt{1 - (v^2/c^2)}V_0$.

Doppler Effect

Consider a light source $\mathscr{S}$ and an observer $\mathscr{O}$ in relative motion. The frequency of light measured by the observer is different from the proper frequency of light (i.e., the frequency measured in the source rest–frame). This phenomenon is known as the "Doppler effect." Special relativity introduces some important corrections to the nonrelativistic predictions. Given a plane wave $e^{i(2\pi\nu t - \boldsymbol{k}\cdot\boldsymbol{x})}$, its phase $\phi \equiv 2\pi\nu t - \boldsymbol{k} \cdot \boldsymbol{x}$ counts the number of wave crests passing by a fixed point in a time interval t, and therefore is an invariant quantity:

$$2\pi\nu' t' - \boldsymbol{k}' \cdot \boldsymbol{x}' = 2\pi\nu t - \boldsymbol{k} \cdot \boldsymbol{x} \quad [14]$$

Using the LT of coordinates and the dispersion relations $\nu = ck/2\pi$ and $\nu' = ck'/2\pi$, one finds

$$\nu' = \nu \frac{\sqrt{1 - (v^2/c^2)}}{1 + (v/c)\cos\vartheta'} \quad [15]$$

where ν' is the frequency measured by the observer (supposed to be at rest in $\mathscr{K}'$), ν is the proper frequency (the frequency in the source rest-frame $\mathscr{K}$), and ϑ' is the angle between the direction of light and the x-axis (which is assumed to be the direction of the relative motion of $\mathscr{S}$ and $\mathscr{O}$). For $\vartheta' = 0$, one has the "longitudinal Doppler effect":

$$\nu' = \nu \sqrt{\frac{1 - (v/c)}{1 + (v/c)}} \quad [16]$$

If the observer $\mathscr{O}$ recedes from the source ($v > 0$), then ν' is smaller than ν; if $\mathscr{O}$ approaches the source ($v < 0$), then ν' is larger then ν. Equation [16] differs from the nonrelativistic result

$$\nu' = \frac{\nu}{1 + (v/c)} \quad [17]$$

by terms of second order in v/c. When the light propagates perpendicularly to the direction of the relative motion of $\mathscr{S}$ and $\mathscr{O}$ ($\vartheta' = \pi/2$), one obtains from [15],

$$\nu' = \nu \sqrt{1 - (v^2/c^2)} \quad [18]$$

This is the formula for the "transverse Doppler effect", which is genuinely a relativistic effect, with no

Newtonian counterpart. The first measurement of the Doppler effect to order v^2/c^2 was performed by Ives and Stilwell in 1938.

Relativistic Dynamics

The relativistic dynamics of a particle of mass m under the action of a force $\boldsymbol{F}$ is governed by the "Minkowski equation"

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\gamma\boldsymbol{v}) = \boldsymbol{F} \qquad [19]$$

which has been experimentally verified for the first time by Bucherer in 1909. Equation [19] reduces to Newton's second law of motion in the low-velocity limit ($v \ll c,\ \gamma \to 1$) whose left-hand side denotes the derivative of the relativistic momentum $\boldsymbol{p}$,

$$\boldsymbol{p} = m\gamma\boldsymbol{v} = \frac{m\boldsymbol{v}}{\sqrt{1-(v^2/c^2)}} \qquad [20]$$

This expression of $\boldsymbol{p}$ guarantees that the conservation of momentum, $\sum_i \boldsymbol{p}_i = \text{constant}$, is preserved by LTs. If the kinetic energy T is defined as the work needed to accelerate a particle, initially at rest, to a velocity v, that is,

$$\mathrm{d}T = \boldsymbol{F}\cdot\boldsymbol{v}\mathrm{d}t \qquad [21]$$

using [19], one obtains

$$T = \frac{mc^2}{\sqrt{1-(v^2/c^2)}} - mc^2 \qquad [22]$$

where the first term,

$$E = \frac{mc^2}{\sqrt{1-(v^2/c^2)}} \qquad [23]$$

is the total energy, and the second term,

$$E_0 = mc^2 \qquad [24]$$

is the energy possessed by the particle at rest. By combining [20] and [23], one gets

$$\boldsymbol{v} = \frac{c^2\boldsymbol{p}}{E} \qquad [25]$$

and the momentum–energy relation

$$E^2 = p^2c^2 + m^2c^4 \qquad [26]$$

which is relativistically invariant. In a classical framework, only the positive-energy solution of [26] makes sense,

$$E = \sqrt{p^2c^2 + m^2c^4} \qquad [27]$$

since no continuous process can generate a transition from positive-energy states (with $E \geqslant mc^2$) to negative-energy states (with $E \leqslant -mc^2$). In quantum physics, the situation is different because discontinuous changes are possible. A problem of stability then arises. In order to circumvent this difficulty, in 1930, Dirac suggested that the negative-energy levels (the so-called "Dirac sea") are completely occupied. A hole in the Dirac sea would correspond to an oppositely charged particle with positive energy (an antiparticle). In the modern quantum field theory, one does not need to introduce negative-energy states: particles and antiparticles are described by the same physical object, a quantum field.

An important peculiarity of relativistic mechanics is that it allows the existence of massless particles. If $m = 0$ in [26], one obtains the energy–momentum relation

$$E = pc \qquad [28]$$

which shows that the energy of a massless particle is proportional to its momentum. Inserting [28] into [24], one finds $v = c$: in any frame, massless particles travel at the speed of light (it is to be noticed that for $v \to c$, eqns [20] and [23] diverge, unless $m \to 0$). An important example of massless particles is given by the photons, which are the quanta of electromagnetic radiation. Combining [28] with the Planck–Einstein formula relating the photon energy E to its frequency ν,

$$E = h\nu \quad (h = \text{Planck constant}) \qquad [29]$$

one obtains for the photon momentum, the expression $p = h\nu/c$.

Mass–Energy Equivalence

The rest-energy formula $E_0 = mc^2$ is of enormous importance. It implies that mass and energy are, so to speak, "equivalent": mass can be converted into energy, and vice versa. Since E_0 is a constant, it is thought to be simply readsorbed in a redefinition of the zero-level of energy. This is not true, since mass is not conserved; therefore, the rest-energy has observable consequences. The law of energy conservation is

$$\sum_i E_i = \sum_f E_f \qquad [30]$$

where $i(f)$ labels the initial (final) state of a physical process. Splitting E into the kinetic energy T and the rest-energy mc^2, one has

$$\sum_i T_i + \sum_i m_ic^2 = \sum_f T_f + \sum_f m_fc^2 \qquad [31]$$

Except for elastic scattering, the kinetic and mass contributions are not separately conserved: one converts into the other. The effect is too small to be

appreciated in the ordinary experience (and in chemistry), but is quite relevant in nuclear and subnuclear physics. For instance, in the fission of a nucleus of ^{238}U induced by a neutron (n),

$$n + {}^{238}\mathrm{U} \rightarrow {}^{143}\mathrm{Ba} + {}^{93}\mathrm{Kr} + 3n$$

the fractional mass loss is $\Delta m/m \sim 0.8 \times 10^{-3}$ and the energy produced is $\Delta E = (\Delta M)c^2 = 2.7 \times 10^{-11}\,\mathrm{J} = 170\,\mathrm{MeV}$. An example of a process characterized by a total conversion of mass into energy is electron–positron annihilation in two photons: $\mathrm{e}^- + \mathrm{e}^+ \rightarrow 2\gamma$.

Relativistic Motion of Charged Particles

The equation of motion of a particle of charge e in an electromagnetic field is (in the Gauss system of units)

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\gamma \boldsymbol{v}) = e\boldsymbol{E} + \frac{e}{c}\,\boldsymbol{v} \times \boldsymbol{B} \qquad [32]$$

where $\boldsymbol{E}$ is the electric field and $\boldsymbol{B}$ the magnetic field. Two noteworthy cases are considered here.

Constant Electric Field

The electric field $\boldsymbol{E}$ is considered here to be directed along the x axis. With $v_x = v_y = v_z = 0$ at $t = 0$, the motion is rectilinear, and assuming $x(0) = 0$, one obtains

$$x(t) = \frac{1}{\alpha}\left(\sqrt{1 + \alpha^2(ct)^2} - 1\right), \quad (\alpha \equiv e|\boldsymbol{E}|/mc^2) \qquad [33]$$

Graphically, $x(t)$ is a hyperbola. For this reason, the one-dimensional motion of a particle acted on by a constant force is called "hyperbolic motion." If $\alpha \ll 1$, eqn [33] reduces to the nonrelativistic result $x(t) = 1/2\alpha(ct)^2$ (a parabola).

Constant Magnetic Field

If the magnetic field is directed along z-axis, the motion of the particle along this axis is uniform, whereas in the xy plane, one has

$$x(t) = \frac{v_\perp}{\Omega}\sin(\Omega t - \alpha) + x_0 \qquad [34]$$

$$y(t) = \frac{v_\perp}{\Omega}\cos(\Omega t - \alpha) + y_0 \qquad [35]$$

where $v_\perp$ is the initial transverse velocity and

$$\Omega = \frac{e|\boldsymbol{B}|}{m\gamma c} \qquad [36]$$

is the "cyclotron frequency." The particle follows a helical path centered at (x_0, y_0) and of radius $R = m\gamma v_\perp c/e|\boldsymbol{B}|$. The relativistic expressions of Ω and R differ from the nonrelativistic ones by the presence of the γ factor.

Minkowski Space

"Henceforth, space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality." These words, pronounced by Hermann Minkowski in 1908 at the Assembly of German Natural Scientists, are a perfect synthesis of the implications of relativity on the physical concepts of space and time. In the Newtonian view of physical processes, the evolution of a system is pictured as a succession of "snapshots" of $\boldsymbol{R}^3$ (the three-dimensional Euclidean space), labeled by time, which is an absolute parameter. In relativity, on the contrary, time is a frame-dependent coordinate, on the same ground as the three spatial coordinates. The primary elements of the relativistic description of physical world are the events, which are specified by four space–time coordinates, $x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$ (hereafter Greek indices take values 0,1,2, and 3, whereas Latin indices take values 1,2, and 3). From a mathematical point of view, events are points of a four-dimensional Riemannian manifold, the Minkowski space $\mathscr{M}$. Given two infinitesimally close events x^μ and $x^\mu + \mathrm{d}x^\mu$, the interval (or separation) $\mathrm{d}s$ between them is defined as

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \qquad [37]$$

where Einstein's summation convention is adopted (repeated indices are summed), and $g_{\mu\nu}$, the "metric tensor," has the form

$$g_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad [38]$$

The quadratic differential form [37] is called the metric of $\mathscr{M}$. It is related to the proper-time interval $\mathrm{d}\tau$ by $\mathrm{d}s = c\,\mathrm{d}\tau$.

A general definition of LTs is given here. A homogenous LT is a linear transformation

$$x'^\mu = \Lambda^\mu_\nu x^\nu \qquad [39]$$

which preserves the metric $\mathrm{d}s^2$. The condition $\mathrm{d}s'^2 = \mathrm{d}s^2$ implies that the 4×4 real matrix Λ^μ_ν must satisfy

$$g_{\mu\nu}\Lambda^\mu_\rho\Lambda^\nu_\sigma = g_{\rho\sigma} \qquad [40]$$

or, in matrix form,

$$\Lambda^T g \Lambda = g \qquad [41]$$

The most general transformation preserving $\mathrm{d}s^2$ is the inhomogeneous LT (or "Poincaré transformation")

$$x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \qquad [42]$$

which includes a space–time translation (a^μ is a set of four real quantities).

The Lorentz matrices Λ^μ_ν have two important properties: (1) their determinant is ± 1 and (2) their Λ^0_0 element is either $\geqslant +1$ or $\leqslant -1$. LTs with $\det \Lambda = +1$ ($\det \Lambda = -1$) are called proper (improper). LTs with $\Lambda^0_0 \geqslant +1$ ($\Lambda^0_0 \leqslant -1$) are called "orthochronous" ("antichronous"). Proper orthochronous LTs are called special, or restricted LTs. Improper and/or antichronous LTs incorporate space and time inversions. Boosts (i.e., LTs relating two inertial frames with parallel axes) and rotations are examples of restricted LTs. For instance, the Lorentz matrix B_x of a boost along the x-direction is

$$B_x(\beta) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad [43]$$

where $\beta \equiv v/c$ is the so-called "rapidity" of the boost. Any restricted LT can be decomposed in the product of a boost B and a rotation R:

$$\Lambda = B(\boldsymbol{\beta})R(\boldsymbol{\theta}) \quad \text{or} \quad \Lambda = R(\boldsymbol{\theta}')B(\boldsymbol{\beta}') \qquad [44]$$

where $\boldsymbol{\beta}$, $\boldsymbol{\beta}'$ and $\boldsymbol{\theta}$, $\boldsymbol{\theta}'$ are uniquely fixed, but in general $\boldsymbol{\beta} \neq \boldsymbol{\beta}'$ and $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$.

Four-Vectors and Four-Tensors

The most elegant way to implement the principle of relativity is to write the laws of physics in a manifestly covariant form. To understand this, some elementary notions of tensor calculus are to be introduced. A "contravariant four-vector" is a set of four quantities $A^\mu = (A^0, A^1, A^2, A^3)$, which transform, under an LT $x'^\mu = \Lambda^\mu_\nu$, as

$$A'^\mu = \Lambda^\mu_\nu A^\nu \qquad [45]$$

Multiplying A^μ by the metric tensor $g_{\mu\nu}$, one obtains the covariant four-vector A_ν,

$$A_\nu = g_{\mu\nu}A^\mu \qquad [46]$$

which transforms as

$$A_\mu = \Lambda^\nu_\mu A_\nu \qquad [47]$$

where $\Lambda^\nu_\mu \equiv (\Lambda^{-1})^\nu_\mu$. The scalar product of two four-vectors A^μ and B^μ is $A \cdot B \equiv A_\mu B^\mu$. A^μ and B^μ are orthogonal, if $A_\mu B^\mu = 0$. The (squared) norm of a four-vector A^μ is $A^2 \equiv A \cdot A \equiv A_\mu A^\mu$. If $A^2 > 0$, the four-vector A^μ is said to be "timelike"; if $A^2 < 0$, it is said to be "spacelike"; if $A^2 = 0$, it is said to be "lightlike."

A "four-tensor" of rank n is a set of 4^n quantities $T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q}$ (with $p + q = n$) which transform under an LT as

$$T'^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q} = \Lambda^{\mu_1}_{\alpha_1} \ldots \Lambda^{\mu_p}_{\alpha_p} \Lambda^{\beta_1}_{\nu_1} \ldots \Lambda^{\beta_q}_{\nu_q} T^{\alpha_1 \ldots \alpha_p}_{\beta_1 \ldots \beta_q} \qquad [48]$$

where $T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q}$ is contravariant in the indices $\mu_1 \ldots \mu_p$ and covariant in $\nu_1 \ldots \nu_q$. Note that four-vectors are four-tensors of rank 1. A four-tensor of rank 0 is a (Lorentz) scalar, that is a quantity which does not change under LTs. Contraction of all indices of a tensor quantity results in a scalar. For instance, $T^{\mu\nu}A_\mu B_\nu$ is a scalar.

The metric tensor $g_{\mu\nu} = g^{\mu\nu}$ raises or lowers the indices. For example,

$$A^\mu = g^{\mu\nu}A_\nu, \qquad A_\mu = g_{\mu\nu}A^\nu, \qquad T^\mu_\nu = g_{\nu\rho}T^{\mu\rho} \qquad [49]$$

The raising or lowering of the 1,2,3 indices yields a change of sign; the raising or lowering of the 0 index leaves the component unchanged:

$$A^0 = A_0, \quad A^1 = -A_1, \quad A^2 = -A_2, \quad A^3 = -A_3 \qquad [50]$$

According to the principle of relativity, physical laws must be invariant in form under LTs. This means that they must be of the type

$$R^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q} = T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q} \qquad [51]$$

By definition, the tensor quantities appearing on the two sides transform in the same way under an LT: thus, if [51] holds in a given inertial frame, it holds in any inertial frame. A tensor law such as [51] is said to be "manifestly covariant."

In the following, the Levi–Civita completely antisymmetric symbol $\varepsilon^{\mu\nu\rho\sigma}$ is used, defined as

$$\varepsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{for even permutations of} \\ & \quad \mu = 0,\ \nu = 1,\ \rho = 2,\ \sigma = 3 \\ -1 & \text{for odd permutations of} \\ & \quad \mu = 0,\ \nu = 1,\ \rho = 2,\ \sigma = 3 \\ 0 & \text{otherwise} \end{cases} \qquad [52]$$

Lorentz Group and Poincaré Group

The homogeneous LTs form a group, the "Lorentz group" $\mathscr{L}$. The set of restricted LTs is a subgroup of $\mathscr{L}$, called the "restricted Lorentz group" (denoted by $\mathscr{L}^\uparrow_+$), and isomorphic to SO(3,1), the group of four-dimensional pseudo-orthogonal matrices with unit determinant. The main mathematical properties of $\mathscr{L}^\uparrow_+$ are: (1) $\mathscr{L}^\uparrow_+$ is a six-parameter semisimple Lie group; (2) it is noncompact; and (3) it is doubly connected (its universal covering being SL(2, C), the

group of two-dimensional complex matrices with unit determinant). A generic representation of $\mathscr{L}_+^\uparrow$ is an invertible linear operator of the form

$$\mathscr{D}(\Lambda) = \exp\left(\frac{\mathrm{i}}{2}\,\omega_{\mu\nu}J^{\mu\nu}\right) \qquad [53]$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ are six real parameters, and $J^{\mu\nu}$ are the generators of the restricted Lorentz group, satisfying the commutation rule

$$[J^{\mu\nu}, J^{\rho\sigma}] = -\mathrm{i}(g^{\mu\sigma}J^{\nu\rho} + g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} - g^{\nu\sigma}J^{\mu\rho}) \qquad [54]$$

which defines the Lie algebra of $\mathscr{L}_+^\uparrow$. The meaning of the generators of $\mathscr{L}_+^\uparrow$ is better understood if one introduces the vector operators

$$\boldsymbol{J} = (J^{23}, J^{31}, J^{12}), \quad \boldsymbol{G} = (J^{01}, J^{02}, J^{03}) \qquad [55]$$

in terms of which the commutator [54] becomes

$$[J^i, J^j] = \mathrm{i}\,\varepsilon^{ijk}J^k \qquad [56]$$

$$[G^i, G^j] = -\mathrm{i}\,\varepsilon^{ijk}J^k \qquad [57]$$

$$[J^i, G^j] = \mathrm{i}\,\varepsilon^{ijk}G^k \qquad [58]$$

An LT is then represented by

$$\mathscr{D}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \exp\{\mathrm{i}(\boldsymbol{\theta}\cdot\boldsymbol{J} - \boldsymbol{\eta}\cdot\boldsymbol{G})\} \qquad [59]$$

where $\boldsymbol{J}$ is the generator of rotations (i.e., the "angular momentum" operator), and $\boldsymbol{G}$ is the generator of boosts (the so-called "center-of-mass motion" operator). Considering the linear combinations

$$N^i \equiv \tfrac{1}{2}(J^i + \mathrm{i}G^i), \quad M^i \equiv \tfrac{1}{2}(J^i - \mathrm{i}G^i) \qquad [60]$$

The Lie algebra ([56]–[58]) translates into:

$$[N^i, N^j] = \mathrm{i}\,\varepsilon^{ijk}N_k \qquad [61]$$

$$[M^i, M^j] = \mathrm{i}\,\varepsilon^{ijk}M_k \qquad [62]$$

$$[N^i, M^j] = 0 \qquad [63]$$

These are two independent SU(2) Lie algebras (in mathematical terms, this means that the algebra of $\mathscr{L}_+^\uparrow$ is the same as that of $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$).

Physical variables (position, momentum, angular momentum, etc.) and fields (electromagnetic field, gravitational field, Dirac field, etc.) transform as the finite-dimensional representations of $\mathscr{L}_+^\uparrow$ (which are not unitary, since $\mathscr{L}_+^\uparrow$ is a noncompact group). These representations are labeled by the eigenvalues of the Casimir operators N^2 and M^2 of the $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$ algebra, which are $n(n+1)$ and $m(m+1)$, respectively, with n and m integers or half-integers. The representation associated with the eigenvalues n, m has dimension $(2n+1)(2m+1)$ and is denoted by (n,m). The most important finite-dimensional representations of $\mathscr{L}_+^\uparrow$ are:

1. The scalar representation (0,0), in which all LTs reduce to the identity.
2. The vector representation (1/2,1/2), in which the LTs are represented by the matrices Λ_ν^μ, and the generators are $(J^{\rho\sigma})_\nu^\mu = -\mathrm{i}(g^{\rho\mu}g_\nu^\sigma - g^{\sigma\mu}g_\nu^\rho)$.
3. The two (inequivalent) spinor representations (0,1/2) and (1/2,0), in which LTs are represented by two-dimensional SL(2, C) matrices and the generators are given by:

$$(0,\tfrac{1}{2}):\ \boldsymbol{J} = \frac{\boldsymbol{\sigma}}{2}, \quad \boldsymbol{G} = \mathrm{i}\,\frac{\boldsymbol{\sigma}}{2}; \quad (\tfrac{1}{2},0):\ \boldsymbol{J} = \frac{\boldsymbol{\sigma}}{2}, \quad \boldsymbol{G} = -\mathrm{i}\,\frac{\boldsymbol{\sigma}}{2} \qquad [64]$$

   where σ^i are the Pauli matrices. The fields transforming in this representation are the two-component "Weyl spinors."
4. The four-dimensional reducible $(1/2,0)\oplus(0,1/2)$ representation, whose generators are

$$\boldsymbol{J} = \begin{pmatrix} \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix}, \quad \boldsymbol{G} = \begin{pmatrix} -\frac{\mathrm{i}}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{\mathrm{i}}{2}\boldsymbol{\sigma} \end{pmatrix} \qquad [65]$$

The fields transforming in this representation are the four-component "Dirac spinors."

The Poincaré transformations (i.e., [42] form the Poincaré group $\mathscr{P}$, which is a 10-parameter noncompact Lie group. The generators of $\mathscr{P}$ are $J^{\mu\nu}$ and the momentum operators P^μ, which generate space–time translations. The Lie algebra of $\mathscr{P}$ is defined by [54] and

$$[P^\mu, J^{\rho\sigma}] = \mathrm{i}(g^{\mu\rho}P^\sigma - g^{\mu\sigma}P^\rho) \qquad [66]$$

$$[P^\mu, P^\nu] = 0 \qquad [67]$$

The Pauli–Lubanski operator is introduced as

$$W^\mu = \tfrac{1}{2}\,\varepsilon^{\mu\nu\rho\sigma}J_{\nu\rho}P_\sigma \qquad [68]$$

The unitary irreducible representations of $\mathscr{P}$ (which are infinite-dimensional, since $\mathscr{P}$ is noncompact) are labeled by the eigenvalues of $P^2 = P^\mu P_\mu$ and $W^2 = W^\mu W_\mu$, which are the two Casimir operators of $\mathscr{P}$. The physical states of particles transform according to these representations. In particular, for massive particles one has $P^2 = M^2$ and $W^2 = -M^2 s(s+1)$, where M is the mass and $s = 0, 1/2, 1, 3/2, \ldots$, is the spin of the particle; for "massless particles," one has $P^2 = W^2 = 0$ and $W^\mu = \lambda P^\mu$, where $\lambda = \pm s$ is the helicity of the particle.

Covariant Formulation of Dynamics

The trajectory of a particle in the Minkowski space is represented by a world-line $x^\mu(\tau)$, where τ is the proper time (recall that $d\tau = ds/c$). Since τ is an invariant parameter, the derivative of x^μ with respect to τ is a four-vector, the four-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} = (\gamma(v)c, \quad \gamma(v)\boldsymbol{v}) \qquad [69]$$

The norm of u^μ is

$$u^\mu u_\mu = c^2 \qquad [70]$$

Multiplying u^μ by m, one obtains the four-momentum p^μ:

$$p^\mu = mu^\mu = (m\gamma(v)c, \ m\gamma(v)\boldsymbol{v}) = \left(\frac{E}{c}, \boldsymbol{p}\right) \qquad [71]$$

It is seen that the time component of p^μ is the energy of the particle (apart from a factor $1/c$), whereas the spatial components form the three-momentum $\boldsymbol{p}$. From [70], it follows that the norm of the four-momentum is

$$p^\mu p_\mu = m^2c^2 \qquad [72]$$

This relation (called "mass-shell relation") is equivalent to [26].

The covariant equation of motion of a free particle,

$$m\frac{du^\mu}{d\tau} = 0 \qquad [73]$$

is obtained by applying Hamilton's variational principle to the action

$$S_0 = -mc^2 \int d\tau \qquad [74]$$

It is to be noted that S_0 is proportional to the proper time of the particle. Expressing (noncovariantly) S_0 as $\int dt\, L_0$ and using $d\tau = dt\sqrt{1-(v^2/c^2)}$, the free–particle Lagrangian L_0 is

$$L_0 = -mc^2\sqrt{1-(v^2/c^2)} \qquad [75]$$

The canonical momentum $\boldsymbol{p} = \partial L_0/\partial \boldsymbol{v}$ and the Hamiltonian $H = \boldsymbol{p}\cdot\boldsymbol{v} - L_0$ coincide with [20] and [27] respectively.

For an interacting particle, the covariant form of the Minkowski equation is

$$m\frac{du^\mu}{d\tau} = \mathscr{F}^\mu \qquad [76]$$

The four-vector on the right-hand side is the four-force, which is related to the ordinary three-force $\boldsymbol{F}$ by

$$\mathscr{F}^\mu = \left(\frac{\gamma}{c}\boldsymbol{F}\cdot\boldsymbol{v}, \gamma\boldsymbol{F}\right) \qquad [77]$$

Since $\mathscr{F}^\mu$ is orthogonal to u^μ,

$$\mathscr{F}^\mu u_\mu = 0 \qquad [78]$$

the four-force must necessarily depend on the velocity of the particle (it is seen that the electromagnetic force acting on a particle is linear in u^μ).

The distributions of energy and momentum of a continuous system (field, fluid, etc.) are contained in the "energy–momentum tensor" $T^{\mu\nu}$, whose components have the following meaning: T^{00} is the energy density; T^{0j}/c is the momentum density; cT^{i0} is the energy flux density in the i direction; T^{ij} (called stress tensor) is the momentum flux density in the i direction. Consider, for instance, a perfect (i.e., completely nonviscous) fluid. In a frame $\mathscr{K}'$ where a given point P of the fluid is instantaneously at rest, the energy–momentum tensor of the fluid has the diagonal form

$$T'^{\mu\nu} = \begin{pmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & c^2\rho_m \end{pmatrix} \qquad [79]$$

where p is the pressure and ρ_m is the proper mass density at P. By a Lorentz transformation, one obtains the expression of $T^{\mu\nu}$ in an arbitrary frame:

$$T^{\mu\nu} = \left(\rho_m + \frac{p}{c^2}\right)u^\mu u^\nu - g^{\mu\nu}p \qquad [80]$$

where u^μ is the four-velocity of the fluid at P. In the case of an incoherent fluid, one has $p = 0$ (no stress), and hence $T^{\mu\nu} = \rho_m u^\mu \mu^\nu$.

Relativistic Scattering

A scattering process is schematically represented by

$$1 + 2 \to 3 + 4 + \cdots + N \qquad [81]$$

where 1 and 2 are the incoming particles and $3, 4, \ldots, N$ are the outgoing particles. One normally assumes that the initial and final states are asymptotic states, which means that, before and after the scattering, the particles are free and satisfy the mass-shell relations

$$p_a^2 = m_a^2 \quad (a = 1, 2, \ldots, N) \qquad [82]$$

The energy–momentum (i.e., four-momentum) conservation law reads

$$p_1 + p_2 = p_3 + p_4 + \cdots + p_N \qquad [83]$$

Due to [82] and [83] and to the arbitrariness in the choice of the reference frame, the number of independent kinematic variables describing the N-body process [81] is $3N - 10$.

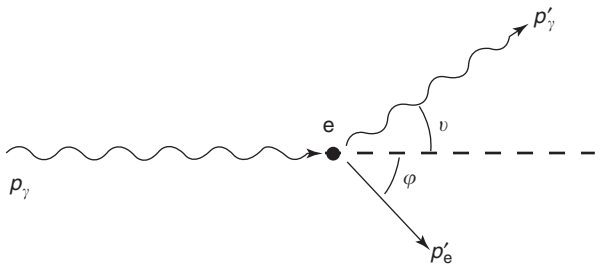


Figure 2 Compton scattering.

Elastic scattering is a two-body process of the type $1+2\rightarrow 1'+2'$, with the same initial and final particles. In this case, the total mass and the total kinetic energy of the particles are separately conserved. An important example of a relativistic elastic process is "Compton scattering," wherein a photon collides elastically with an electron (**Figure 2**):

$$\gamma + e^- \rightarrow \gamma + e^- \qquad [84]$$

In the electron rest-frame, the four-momenta are (apices denote final-state variables)

$$p_\gamma = \left(\frac{E_\gamma}{c}, \boldsymbol{p}_\gamma\right), \qquad p_e = (m_e c, \boldsymbol{0});$$
$$p'_\gamma = \left(\frac{E'_\gamma}{c}, \boldsymbol{p}'_\gamma\right), \qquad p'_e = \left(\frac{E'_e}{c}, \boldsymbol{p}'_e\right) \qquad [85]$$

with

$$E_\gamma = h\nu, \quad |\boldsymbol{p}_\gamma| = \frac{h\nu}{c}; \quad E'_\gamma = h\nu', \quad |\boldsymbol{p}'_\gamma| = \frac{h\nu'}{c} \qquad [86]$$

If four-momentum conservation is applied, one obtains

$$\nu' = \frac{\nu}{1+\varepsilon(1-\cos\vartheta)} \quad (\varepsilon \equiv h\nu/m_e c^2) \qquad [87]$$

where ϑ is the photon scattering angle. In terms of wavelengths, eqn [87] becomes

$$\lambda' - \lambda = \lambda_C(1-\cos\vartheta) \qquad [88]$$

where $\lambda_C = h/m_e c^2$ is the Compton wavelength of the electron (its numerical value is 0.024×10^{-10} m).

Consider now an inelastic production process, such as [81]. This reaction is kinematically allowed only if the initial energy is greater than a threshold value, determined by the condition

$$W^2 \equiv (p_1+p_2)^2 \geqslant (m_3+m_4+\cdots+m_N)^2 \qquad [89]$$

where W is the so-called "invariant mass" of the system. For instance, in the rest-frame of particle 2 (the laboratory frame), the energy of particle 1 must be

$$E_1 \geqslant \frac{(m_3+m_4+\cdots+m_N)^2 - (m_1^2+m_2^2)}{2m_2} \qquad [90]$$

Covariant Formulation of Electrodynamics

The most elegant way to show that electromagnetism is relativistically invariant is to put Maxwell's equations in a covariant form. The four-potential

$$A^\mu = (\varphi, \boldsymbol{A}) \qquad [91]$$

is introduced, which contains the scalar potential φ and the vector potential $\boldsymbol{A}$. The electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ are incorporated into the "field strength tensor" $F^{\mu\nu}$, an antisymmetric tensor of rank 2 defined as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \qquad [92]$$

This expression must be equivalent to

$$\boldsymbol{E} = -\boldsymbol{\nabla}\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} \qquad [93]$$

$$\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} \qquad [94]$$

so that the components of $F^{\mu\nu}$ are

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \qquad [95]$$

In terms of $F^{\mu\nu}$, the inhomogeneous Maxwell equations

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho, \quad \boldsymbol{\nabla}\times\boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi}{c}\boldsymbol{j} \qquad [96]$$

read

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} j^\nu \qquad [97]$$

where $\partial_\mu \equiv ((1/c)\partial/\partial t, \boldsymbol{\nabla})$ and the four-density j^μ is

$$j^\mu = (c\rho, \boldsymbol{j}) \qquad [98]$$

The operator ∂_ν is now applied to eqn [97]. Due to the antisymmetry of $F^{\mu\nu}$, one has $\partial_\nu\partial_\mu F^{\mu\nu} = 0$, and, therefore,

$$\partial_\nu j^\nu = 0 \qquad [99]$$

which is equivalent to the equation of continuity of electric charge,

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0 \qquad [100]$$

Inserting [92] into [97], and adopting the Lorentz gauge, $\partial_\mu A^\mu$, one obtains

$$\Box A^\nu = \frac{4\pi}{c} j^\nu \qquad [101]$$

where $\Box \equiv \partial_\mu\partial^\mu$ is the d'Alembertian. The simplest covariant expression of the homogeneous Maxwell

equations,

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \nabla \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0 \quad [102]$$

is obtained by introducing the "dual field strength" $\tilde{F}^{\mu\nu}$,

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\,\varepsilon^{\mu\nu\rho\sigma}\,F_{\rho\sigma} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix} \quad [103]$$

in terms of which eqn [102] becomes

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad [104]$$

The (symmetrized) energy–momentum tensor of the electromagnetic field is

$$T^{\mu\nu} = -\frac{1}{4\pi}\,F^{\mu\rho}F^{\nu}_{\rho} + \frac{1}{16\pi}\,F^{\rho\sigma}F_{\rho\sigma}g^{\mu\nu} \quad [105]$$

It contains the energy density $\omega \equiv T^{00}$ and the momentum density $\pi^i \equiv T^{0i}/c$, which are given explicitly by

$$\omega = \frac{1}{8\pi}(\boldsymbol{E}^2 + \boldsymbol{B}^2), \quad \boldsymbol{\pi} = \frac{1}{4\pi c}\,\boldsymbol{E} \times \boldsymbol{B} \quad [106]$$

With $F^{\mu\nu}$ and $\tilde{F}^{\mu\nu}$, the two invariants are constructed as

$$F^{\mu\nu}F_{\mu\nu} = -2(\boldsymbol{E}^2 - \boldsymbol{B}^2), \qquad \tilde{F}^{\mu\nu}F_{\mu\nu} = -4\boldsymbol{E}\cdot\boldsymbol{B} \quad [107]$$

and it is seen that, if the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are equal in magnitude and/or orthogonal in a given inertial frame, they are equal in magnitude and/or orthogonal in any inertial frame. The LTs of $\boldsymbol{E}$ and $\boldsymbol{B}$ for a boost along a generic direction are ($\boldsymbol{\beta} = \boldsymbol{v}/c$)

$$\boldsymbol{E}' = \gamma(\boldsymbol{E} + \boldsymbol{\beta} \times \boldsymbol{B}) - \frac{\gamma - 1}{\beta^2}(\boldsymbol{E}\cdot\boldsymbol{\beta})\boldsymbol{\beta} \quad [108]$$

$$\boldsymbol{B}' = \gamma(\boldsymbol{B} - \beta \times \boldsymbol{E}) - \frac{\gamma - 1}{\beta^2}(\boldsymbol{B}\cdot\boldsymbol{\beta})\boldsymbol{\beta} \quad [109]$$

A charged particle interacting with the electromagnetic field is now considered. The equation of motion of the particle is

$$m\,\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{e}{c}\,F^{\mu\nu}u_\nu \quad [110]$$

and corresponds to [32]. The action leading to [110] is

$$S = -mc^2\int \mathrm{d}\tau - \frac{e}{c}\int A_\mu(z)\,\mathrm{d}z^\mu \quad [111]$$

From this, the noncovariant Lagrangian is extracted as

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} - e\varphi + \frac{e}{c}\,\boldsymbol{v}\cdot\boldsymbol{A} \quad [112]$$

The canonical momentum $\boldsymbol{P}$ differs from the kinetic momentum $\boldsymbol{p} = m\gamma\boldsymbol{v}$ by a term proportional to $\boldsymbol{A}$:

$$\boldsymbol{P} = \frac{\partial L}{\partial \boldsymbol{v}} = m\gamma\boldsymbol{v} + \frac{e}{c}\,\boldsymbol{A} = \boldsymbol{p} + \frac{e}{c}\,\boldsymbol{A} \quad [113]$$

The Hamiltonian associated to [112] is

$$H = \left[\left(\boldsymbol{P} - \frac{e}{c}\,\boldsymbol{A}\right)^2 c^2 + m^2c^4\right]^{1/2} + e\varphi \quad [114]$$

The relativistic treatment of a system of interacting charged particles is complicated by the fact that each particle has its own proper time. It is possible, however, to write an approximate Lagrangian, valid to second order in v/c (a, b are particle indices):

$$L = \sum_a\left(\frac{1}{2}\,m_a\boldsymbol{v}_a^2 + \frac{1}{8c^2}\,m_a\boldsymbol{v}_a^4\right) - \frac{1}{2}\sum_a\sum_{b\neq a}\frac{e_ae_b}{r_{ab}} + \frac{1}{4c^2}\sum_a\sum_{b\neq a}\frac{e_ae_b}{r_{ab}}\left[\boldsymbol{v}_a\cdot\boldsymbol{v}_b + \frac{(\boldsymbol{v}_a\cdot\boldsymbol{r}_{ab})(\boldsymbol{v}_b\cdot\boldsymbol{r}_{ab})}{r_{ab}^2}\right] \quad [115]$$

This Lagrangian was derived in 1920 by C G Darwin and leads, in quantum mechanics, to the Breit equation.

The electromagnetic potentials generated by a point charge in motion are obtained by solving eqn [101] with the four-density j^μ given by

$$j^\mu(x) = ec\int_{-\infty}^{+\infty}\mathrm{d}\tau\,u^\mu(\tau)\delta^4(x - z(\tau)) \quad [116]$$

where $z^\mu(\tau)$ is the world-line of the particle. Assuming there are no incident fields, one obtains the so-called "Liénard–Wiechert potentials":

$$\varphi(\boldsymbol{x}, t) = \left[\frac{e}{R - \boldsymbol{R}\cdot\boldsymbol{v}/c}\right]_{\mathrm{ret}}$$
$$\boldsymbol{A}(\boldsymbol{x}, t) = \left[\frac{e(\boldsymbol{v}/c)}{R - \boldsymbol{R}\cdot\boldsymbol{v}/c}\right]_{\mathrm{ret}} \quad [117]$$

where

$$\boldsymbol{R}(t) = \boldsymbol{x} - \boldsymbol{z}(t) \quad [118]$$

is the position of the observation point relative to the charge (located in $\boldsymbol{z}$), and the subscript "ret" means that all quantities in the brackets are evaluated at the retarded time, defined by

$$t_{\mathrm{ret}} = t - \frac{R(t_{\mathrm{ret}})}{c} \quad [119]$$

The fields corresponding to the potentials [117] are ($\boldsymbol{\beta} \equiv \boldsymbol{v}/c, \boldsymbol{n} \equiv \boldsymbol{R}/R$)

$$\boldsymbol{E}(\boldsymbol{x},t) = e\left\{\frac{(1-\beta^2)(\boldsymbol{n}-\boldsymbol{\beta})}{R^2(1-\boldsymbol{\beta}\cdot\boldsymbol{n})^3}\right\}_{\text{ret}} + \frac{e}{c}\left\{\frac{\boldsymbol{n}\times[(\boldsymbol{n}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}]}{R(1-\boldsymbol{\beta}\cdot\boldsymbol{n})^3}\right\}_{\text{ret}} \quad [120]$$

$$\boldsymbol{B}(\boldsymbol{x},t) = [\boldsymbol{n}\times\boldsymbol{E}(\boldsymbol{x},t)]_{\text{ret}} \quad [121]$$

At large distances from the source, the dominant contributions come from the first term of [120] and the corresponding term in [121], which behave as $1/R$ and are the acceleration-dependent radiation fields.

Spin Dynamics

The spin $\boldsymbol{s}$ of a massive particle is incorporated into a four-vector S^μ orthogonal to the four-velocity, given by

$$S^\mu = \left(\gamma\boldsymbol{\beta}\cdot\boldsymbol{s}, \boldsymbol{s} + \frac{\gamma^2}{\gamma+1}(\boldsymbol{\beta}\cdot\boldsymbol{s})\boldsymbol{\beta}\right) \quad [122]$$

In the particle rest-frame $\mathscr{K}'$, the spin four-vector [122] reduces to $(0,\boldsymbol{s})$, and the equation of motion of $\boldsymbol{s}$ in presence of a magnetic field is

$$\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t'} = \frac{eg}{2mc}\,\boldsymbol{s}\times\boldsymbol{B}' \quad [123]$$

where g is the "gyromagnetic ratio." The covariant generalization of [123] in an arbitrary inertial frame $\mathscr{K}$ is the so-called "Bargmann–Michel–Telegdi equation":

$$\frac{\mathrm{d}S^\mu}{\mathrm{d}\tau} = \frac{e}{mc}\left[\frac{g}{2}F^{\mu\nu}S_\nu + \left(\frac{g}{2}-1\right)S_\rho F^{\rho\sigma}u_\sigma u^\mu\right] \quad [124]$$

From this, one can extract the corresponding equation for the spin vector $\boldsymbol{s}$, which reads

$$\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t} = \frac{eg}{2mc}\,\boldsymbol{s}\times\left[(\boldsymbol{B}-\boldsymbol{\beta}\times\boldsymbol{E}) - \frac{\gamma-1}{\gamma\beta^2}\,\boldsymbol{\beta}(\boldsymbol{\beta}\cdot\boldsymbol{B})\right] + \boldsymbol{\omega}_{\text{T}}\times\boldsymbol{s} \quad [125]$$

with

$$\boldsymbol{\omega}_{\text{T}} = -\frac{\gamma-1}{\beta^2}\,\boldsymbol{\beta}\times\frac{\mathrm{d}\boldsymbol{\beta}}{\mathrm{d}t} \quad [126]$$

The term containing $\boldsymbol{\omega}_{\text{T}}$ represents the "Thomas precession" of spin. Recalling the transformation law [109] of $\boldsymbol{B}$, it is seen that eqn [125] is related to the equation of motion [123] of $\boldsymbol{s}$ in the particle rest-frame $\mathscr{K}'$ (which is a noninertial frame) by

$$\left(\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t}\right)_{\mathscr{K}} = \left(\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t}\right)_{\mathscr{K}'} + \boldsymbol{\omega}_{\text{T}}\times\boldsymbol{s} \quad [127]$$

General Relativity

Special relativity does not describe gravitational phenomena. Newton's law of gravity is, in fact, incompatible with the principle of relativity and all the attempted covariant generalizations of this law have led to various inconsistencies and/or to wrong predictions. A coherent and empirically successful theory of gravity is provided by general relativity (Einstein, 1916), which extends the co-variance of physical laws to arbitrary coordinate transformations ("principle of general covariance"). One of the starting points of general relativity is the "equivalence principle," namely the fact that the inertial mass is equal to the gravitational mass. As noticed by Einstein, this equality (which has no explanation in Newtonian physics) allows one to cancel, or simulate, gravity by an inertial force (i.e., by a transformation to an accelerated frame), provided the space–time region where this is done is so small that nonhomogeneities of the gravitational field can be neglected. In Einstein's theory, the equivalence principle takes the form of a postulate which states that for any space–time event, it is always possible to choose a coordinate system such that in a sufficiently small neighborhood of the event at hand, physical laws are written in the same way as in inertial systems. Unlike the Minkowski space $\mathscr{M}$, the space–time of general relativity is a curved Riemannian manifold. Its metric tensor $g_{\mu\nu}(x)$ depends on the coordinates and cannot be reduced to the Minkowskian metric tensor [38]. From a physical point of view, $g_{\mu\nu}(x)$ represents the gravitational field. To be precise, in the weak-field limit, one has

$$g_{00} = 1 + \frac{2\Phi}{c^2} \quad [128]$$

where Φ is the Newtonian gravitational potential ($\Phi = -GM/r$ for a spherical mass M). Thus, general relativity establishes a fundamental link between geometry and physics: gravity manifests itself as the curvature of space–time. This curvature is given by a tensor of rank 4, the "Riemann tensor" $R_{\mu\nu\rho\sigma}$, which contains first and second derivatives of $g_{\mu\nu}(x)$. The dynamics of the gravitational field is governed by the Einstein equation

$$R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}R = \frac{8\pi G}{c^4}\,T_{\mu\nu} \quad [129]$$

where $R_{\nu\sigma} = g^{\mu\rho}R_{\mu\nu\rho\sigma}$ is the "Ricci tensor," $R = g^{\nu\sigma}R_{\nu\sigma}$ is the "curvature scalar," and $T^{\mu\nu}$ is the symmetrized energy–momentum tensor. An important point to be noticed is that gravity is generated not only by matter, but by any distribution of energy and momentum. In particular, since the gravitational field carries energy and momentum, it acts as a source of itself (this is signaled by the fact that the Einstein

equation is a nonlinear equation). Qualitatively, eqn [129] shows that the curvature is proportional to the mass density $\rho_m \sim T_{00}/c^2$,

$$\text{curvature} \sim \frac{G\rho_m}{c^2} \qquad [130]$$

Here, the constant coefficient is $G/c^2 = 7.4 \times 10^{-30}$ kg m^{-1}. The low value of this quantity indicates that only very high densities of energy and mass produce an appreciable curvature of space–time. When $G\rho/c^2 \ll 1$, the space–time is approximately flat and special relativity applies.

See also: Electrodynamics: Continuous Media; Electrodynamics: Quantum Electrodynamics; Group Theory; Nuclear Fission and Fusion.

PACS: 03.30.+p; 03.50.De; 04.20.Cv

Further Reading

Anderson JL (1967) *Principles of Relativity Physics*. New York: Academic Press.
Barut AO (1980) *Electrodynamics and Classical Theory of Fields and Particles*. New York: Dover.
Jackson JD (1975) *Classical Electrodynamics*, 2nd edn. New York: Wiley.
Landau LD and Lifshitz EM (1980) *The Classical Theory of Fields*, 4th edn. Oxford: Pergamon Press.
Lorentz HA, Einstein A, Minkowski H, and Weyl H (1952) In: Sommerfeld A (ed.) *The Principle of Relativity*. New York: Dover.
Møller C (1952) *The Theory of Relativity*. Oxford: Clarendon.
Resnick R (1968) *Introduction to Special Relativity*. New York: Wiley.
Rindler W (1960) *Special Relativity*. Edinburgh: Oliver and Boyd.
Rosser WGV (1967) *Introductory Relativity*. London: Butterworths.
Schutz BF (1990) *A First Course in General Relativity*. Cambridge: Cambridge University Press.
Weinberg S (1972) *Gravitation and Cosmology*. New York: Wiley.

Rhodopsin and the First Step in Vision

R R Birge, University of Connecticut, Storrs, CT, USA

Introduction

Rhodopsin is a membrane-bound photoreceptor protein that is responsible for dim light vision in humans and animals with image resolving eyes. This protein is a member of a special class of membrane bound proteins known as G-protein coupled receptors (GPCRs), but it is the only GPCR for which an atomic level structure has been obtained. The primary event of rhodopsin involves the photochemical conversion of a bound chromophore, 11-*cis* retinal, to the all-*trans* form. Evolution has optimized both the thermodynamics and kinetics of this process to be highly efficient, and the photochemistry is complete in roughly 200 fs while storing 134 kJ mol^{-1} of energy. The process is facilitated by the interaction of a lowest energy, strongly allowed excited state, which absorbs the photon, and a nearby forbidden state, which interacts with the lower state to produce a barrierless excited state potential energy surface. In this article, one can examine the photophysical properties of rhodopsin and the cone pigments that are responsible for color vision. The discussion not only emphasizes the photophysical aspects of rhodopsin photochemistry, but also includes an overview of the biochemistry of vision and the potential relevance of rhodopsin structure and function to other GPCR proteins. Where possible, one can refer to review articles rather than the primary literature in order to obtain relevant background information and perspective.

G-Protein Coupled Receptors

Life would not be possible without communication between the cells and the external environment. This critical capability allows cells to monitor the status of the organs and tissues and to respond to environmental threats and opportunities. Intercellular and interspecies communication is mediated by a group of membrane-bound proteins known as G-protein coupled receptors (GPCRs). Each GPCR monitors a specific compound or group of compounds and becomes activated when the target ligand (agonist) occupies the active site. The activated GPCR signals the presence of the ligand through interaction with a separate protein known as a G-protein. G-proteins are so named because they bind guanine nucleotides GDP and GTP, and the G-protein that is utilized by rhodopsin is called transducin. When rhodopsin becomes activated by light, the GDP on transducin is replaced with GTP which activates the signaling cascade responsible for the nerve impulse. The G-protein based activation process provides reliable signaling and, in the case of the vision, significant amplification. There are at least 700 GPCRs encoded in the human genome. Despite significant biological importance, the structure and function of most GPCRs are poorly understood and a high-resolution crystal structure of only one GPCR, that is, the visual

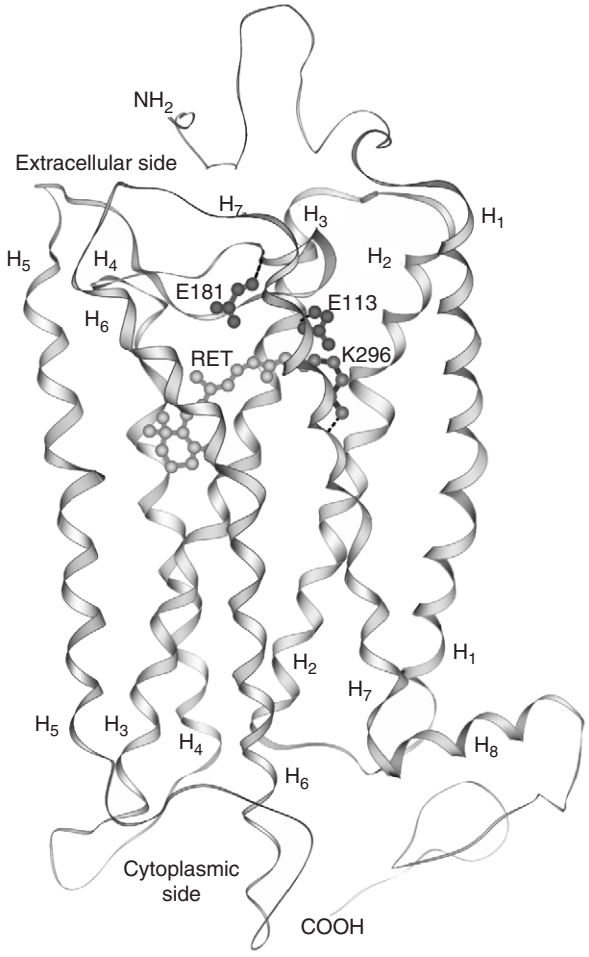


Figure 1 A ribbon picture of rhodopsin based on the crystal structure. The transmembrane helices are labeled H1–H7. A short helical segment that lies on the cytoplasmic surface of the membrane is marked H8. The 11-*cis* retinal chromophore (RET) is attached to the protein via a protonated Schiff base linkage to lysine 296 (K296). The two nearby glutamic acid residues E113 and E181 are also shown. (Adapted from Palczewski K, Kumasaka T, Hori T, Behnke CA, *et al.* (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* 289 (5480): 739–745; and Okada, Ernst O, Palczewski K, and Hofmann K (2001) Activation of rhodopsin: new insights from structural and biochemical studies. *Trends in Biochemical Science* 26: 318–324.)

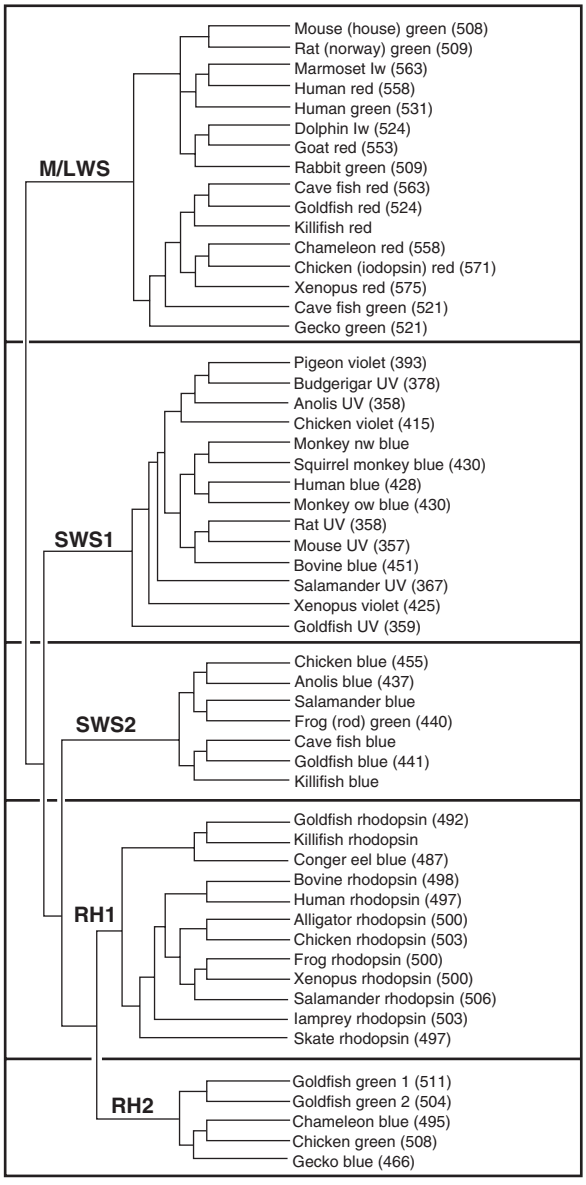


Figure 2 A dendogram showing selected vertebrate visual opsins grouped as suggested by Shozo Yokoyama. The rhodopsin pigments are responsible for scotopic (low-light) vision and are in group RH1. The cone pigments responsible for photopic (high flux, color) vision are found in the other groups. The absorption maxima, known, are listed in nanometers in parentheses. (Adapted from Ebrey T and Koutalos Y (2001) Vertebrate photoreceptors. *Progress in Retinal and Eye Research* 20: 49–94.)

pigment rhodopsin, has been reported. The three-dimensional structure of rhodopsin protein is shown in **Figure 1**.

The Visual Pigments

The rhodopsins ($\lambda_{max} \approx 500$ nm) are members of the visual pigment family, which are divided into groups based on similarities in primary structure (the sequence of amino acids in the protein) and wavelength of absorption (see **Figure 2**). The rhodopsins are found in group RH1 and are responsible for scotopic (low light level) vision. The cone pigments are found in RH2, SWS1, SWS2 or M/LWS groups and are responsible for photopic (high light level) color vision. All of these visual pigments are seven transmembrane alpha helical proteins that bind 11-*cis* retinal and initiate the light transduction signaling pathway in retinal photoreceptors. While other GPCRs interact with their ligands noncovalently, the visual pigments consist of 11-*cis* retinal covalently attached to the

apoprotein through a Schiff base linkage to a conserved lysine residue in the seventh transmembrane helix (H7, see **Figure 1**). Following absorption of light, the retinal chromophore isomerizes to the all-*trans* conformation and triggers a series of conformational changes that lead to the formation of the active state, R$*$ or meta II. All-*trans* retinal is eventually released from the vertebrate protein and visual pigments, can be regenerated from 11-*cis* retinal spontaneously. Although some consider a photon of light to represent the agonist, a more logical choice is all-*trans* retinal, which is photochemically generated from the 11-*cis* retinal cofactor. Thus, the visual pigments have a covalently attached agonist but one that starts out in an inactive form. The fact that the agonist is covalently attached allows studies of the activation process at a level not possible with the vast majority of the other GPCRs, where the mode and location of interaction of the agonist remains obscure.

The Primary Event of Vision

The primary photochemical event that follows the absorption of light by rhodopsin is localized within the light-absorbing chromophore, the 11-*cis* retinal. This chromophore is attached to the protein via a protonated Schiff base linkage as shown in **Figure 3**. Upon the absorption of light, the 11-*cis* chromophore isomerizes rapidly to form an 11-*trans* photoproduct that is schematically shown in **Figure 3**. Although experimental studies show that this chromophore is 11-*trans*, it is possible that rotations about single bonds have taken place to accommodate the change in geometry. The remarkable feature of the primary event is that it is complete in 200 fs. Detailed vibronic analyses indicate that the process is vibrationally coherent, which indicates that there is a direct path that the 11-*cis* excited state species will follow in carrying out the isomerization to the 11-*trans* ground-state photoproduct. This observation, coupled with the observation that photoisomerization also occurs with high efficiency at 4 K, indicates that the photochemistry takes place along a barrierless excited state surface. A simplified diagram presenting some of the important aspects of the photoisomerization surface is shown in **Figure 4**. In addition to having a fast photochemical conversion, rhodopsin must store enough energy in the primary event to drive the subsequent thermal reactions that activate the protein. The primary event stores $\sim 134\,\mathrm{kJ\,mol^{-1}}$, which is $\sim 57\%$ of the energy available from light absorption. The quantum efficiency of the primary event is 0.65 which means that 65% of the photons absorbed yield activated protein. Thus, the net efficiency of the protein in carrying out its photoreceptor function is 36%. The efficiency is all the more remarkable when one considers that the probability that a rhodopsin molecule will spontaneously activate to produce a false signal is extremely small, $\sim 10^{-11}$ false (thermal) events per rhodopsin molecule per second. There is no human-made device that approaches this level of speed, efficiency, and reliability. The question arises, however, why nature would optimize the visual pigment to have a primary event that is so fast. The neural signal is separated from the original photon absorption by tens of milliseconds, and one might conclude from this observation that there would be no evolutionary rationale for selection of a coherent, femtosecond isomerization process. The answer to this conundrum is simple. Natural selection optimized the kinetics of the primary event so that the significant energy storage required could be accommodated while simultaneously eliminating possible side reactions. It is noted that the primary photochemical event of bacteriorhodopsin, which stores roughly one-third as much energy in the primary event, is approximately three times slower despite having a nearly identical quantum efficiency. The interesting question is why both light-transducing proteins have converged on nearly identical quantum efficiencies, an interesting and important mechanistic observation that is not yet understood.

What remains to be examined here is the mechanistic origin of the barrierless excited state potential energy surface. As background, it is noted that the protonated Schiff base chromophore of 11-*cis* retinal does not undergo comparable photochemistry in solution. Although the quantum efficiency is only slightly lower in solution, the photoisomerization takes nanoseconds, not femtoseconds, which is many orders of magnitude slower. Clearly the protein binding site has a significant impact on the excited state potential surface.

The mechanism relies on the electronic properties of polyenes. In the late 1970s, a new excited singlet state was experimentally identified by Bryan Kohler and his graduate student, Bruce Hudson, who were both at Harvard at that time. The state is the lowest-lying excited singlet state in long-chain linear polyenes and has the same symmetry as the ground state, which under the C_{2h} point group is ^1A_g. One-photon selection rules prohibit $^1A_g^* \leftarrow {}^1A_g$ transitions (the asterisk indicates an excited state), and hence the $^1A_g^*$ state is spectroscopically forbidden which explains why nearly 50 years of polyene spectroscopy had failed to observe it. Of note is the fact that this forbidden state is responsible for the fluorescence from long-chain polyenes, and that in all-*trans* retinal, this

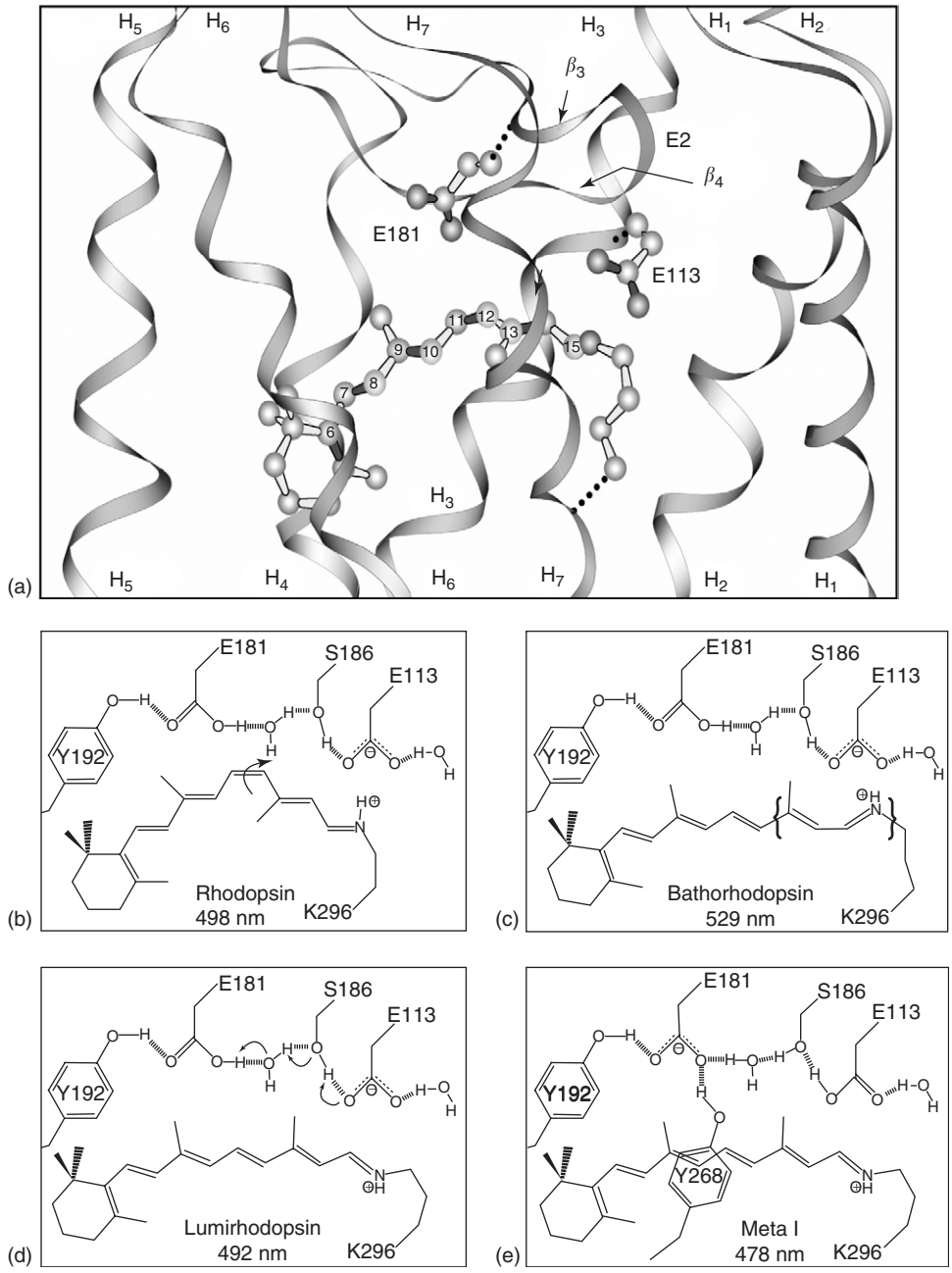


Figure 3 The second extracellular loop, E2, extends into the chromophore binding site in the form of two beta sheets (β3 and β4) as shown in (a) (coordinates from Palczewoski K, Kumasaka T, Hori T, Behnke CA, *et al.* (2000) Crystal structure of rhodospin: a protein-coupled receptor. *Science* 289(5480): 739–745). Glutamic acid residues 113 and 181 are also shown in (a), and the protonation states of these two residues as a function of photobleaching intermediate is schematically shown in panels b–e (assignments from Yan ECY, Kasmi MA, Ganim Z, Hou JM, Pan D, *et al.* (2003) Retinal counterion switch in the photoactivation of the G protein-coupled receptor rhodopsin. *Proceedings of National Academy of Sciences* USA 100: 9262–9267). The arrows in panels b and d indicate the key conformational or chemical changes associated with transformation of rhodopsin (to bathorhodopsin) and lumirhodopsin (to meta I). The chromophore geometry in bathorhodopsin remains controversial, and the brackets are intended to reflect both distortion and conformational uncertainty. The relative location of the residues is approximate and does not reflect possible movement during the thermal relaxations.

state lies below the strongly allowed $^1B_u^*$ state responsible for the strong absorption band characteristic of the visual chromophores. Experimental and theoretical studies indicate that the $^1A_g^*$ state is characterized by significant bond order reversal, which lowers the barrier to dihedral distortion around the (ground state) double bonds in the excited state. Two-photon studies indicate that in protonated Schiff

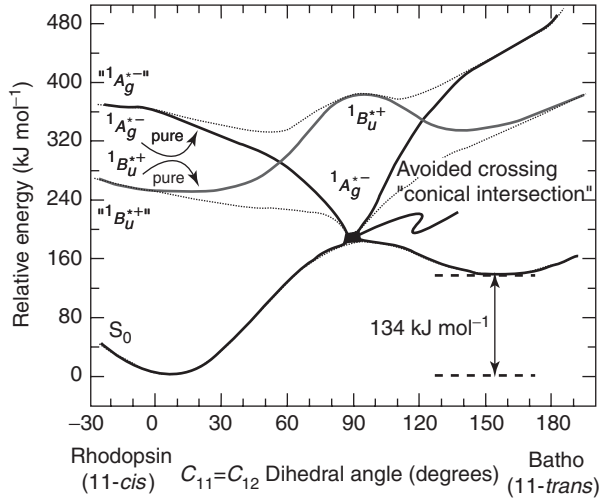


Figure 4 A schematic drawing of the adiabatic ground and excited singlet-state potential surfaces for isomerization of the chromophore in rhodopsin. Interaction between the lowest-excited $^1B_u^*$-like state and the second excited $^1A_g^*$-like state is responsible for generating a barrierless excited state potential surface for isomerization. At the orthogonal point near 90°, the excited singlet-state surface crosses the ground state surface and produces what is called a conical intersection which facilitates rapid and coherent transfer into the ground state.

bases and in the protein bound chromophore, the $^1A_g^*$ state lies above the allowed $^1B_u^*$ state. This level inversion is responsible for the barrierless excited state potential surface because as the amount of torsional distortion increases, the amount of repulsion between the two states increases. This repulsion causes the lowest-lying excited state to have lower energy as torsion increases and is responsible, at least in part, for the barrierless surface. The situation is shown graphically in **Figure 4**. It is parenthetically noted that, the level ordering shown in **Figure 4** also applies to the unprotonated Schiff base chromophores in the UV cone pigments. Protonation, conformation, and external charges are each capable of inducing level order inversion in the visual chromophores.

The protein matrix enhances the photoisomerization process by placing the primary negatively charged glutamic acid counterion, E113, near the Schiff base group. In the ground state, this group stabilizes the positive charge which is primarily localized in the imine region. The absorption of light creates a hyperpolarized charge shift that transfers electron density down the polyene chain toward the imine linkage as shown in **Figure 5**. The primary counterion, E113, is now a source of destabilization due to electrostatic repulsion between the chromophore excited state charge distribution and the nearby glutamic acid residue. This repulsion alters the excited state potential surface to have a larger initial slope toward product formation and enhances the kinetics of photoisomerization.

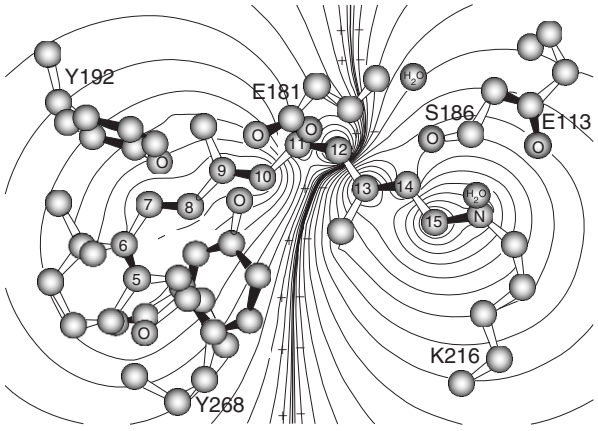


Figure 5 Electrostatic contours associated with excitation of the protein-bound 11-*cis* retinal protonated Schiff base chromophore in rhodopsin. The chromophore undergoes a significant charge shift upon excitation that transfers electron density toward the imine linkage (C14–C15 = N), and produces electrostatic repulsion between the chromophore and the primary counterion E113. The polyene atoms of the chromophore are numbered, and the nitrogen and oxygen atoms are labeled (oxygen atoms on water are labeled H_2O).

The Photobleaching Process

The primary event initiates a series of thermal reactions that ultimately produce meta II, the active form of rhodopsin. The early intermediates of this process are called bathorhodopsin (the primary photoproduct), lumirhodopsin, and metarhodopsin I (meta I). Recent site-directed mutagenesis and spectroscopic studies indicate that there is a counterion shift during the lumi to meta I transition, as shown in **Figure 3**. This counterion shift is also found in the UV cone pigment of the mouse (MUV), which is important because the chromophore is unprotonated in the MUV-rho and MUV-batho states. The MUV chromophore is ultimately protonated in the lumi intermediate via donation of a proton from the nearby (neutral) E113 residue. The remaining portion of the photobleaching sequence essentially duplicates that found in rhodopsin. This observation provides support for the importance of the counterion switch to the activation of visual pigments in general. Otherwise, why would such a convoluted sequence of events evolve if it were not a key contributor to function? From another perspective, the counterion switch mechanism makes it possible for ultraviolet vision, which uses an unprotonated chromophore to achieve short wavelength absorption. The counterion switch mechanism provides a comprehensive activation strategy that will work for both protonated and unprotonated chromophores.

See also: Elementary Excitations in Biological Systems.

PACS: 42.66.Si; 82.50.Hp; 87.15. – v

Further Reading

Berg J, Tymoczko J, and Stryer L (2002) *Biochemistry*, 5th edn. New York, NY: W. H. Freeman and Company.

Birge RR and Vought BW (2000) In: Palczewski K (ed.) *Methods in Enzymology*, vol. 315, pp. 143–163. San Diego, CA: Academic Press.

Hudson B and Kohler R (1974) Linear polyene electronic structure and spectroscopy. *Annual Review of Physical Chemistry* 25: 437–460.

Mathies R, Lin S, Ames J, and Pollard T (1991) From femtoseconds to biology: mechanism of bacteriorhodopsin's light-driven proton pump. *Annual Reviews of Biophysics and Biophysical Chemistry* 20: 491–518.

Oesterhelt D (1998) The structure and mechanism of the family of retinal proteins from halophilic Archaea. *Current Opinion in Structural Biology* 8: 489–500.

Stuart JA and Birge RR (1996) In: Lee AG (ed.) *Biomembranes*, vol. 2A, pp. 33–140. London: JAI Press.

Ruthenates

B Raveau, Laboratoire de Cristallographie, ISMRA, Caen, France

Introduction

The properties of ruthenium oxides are mainly governed by the ability of Ru to adopt, preferentially, the tetravalent state and the octahedral coordination. For this reason, in the system Ru–O, RuO_2 is the only stable oxide. The latter exhibits the rutile structure with a large axial ratio $c/a > 2/3$, so that the π^* orbitals in this system are half-filled. As a consequence, RuO_2 exhibits a metallic conductivity ($\rho \approx 10^{-5}$ Ω cm at room temperature) and its magnetic susceptibility is small, nearly temperature independent, characteristic of Pauli paramagnetism.

The rather low reactivity of RuO_2 with other oxides suggests that the number of ruthenates that have been synthesized to date is rather limited. They are of two types, Ru(IV) and Ru(V) ruthenates. The first ones tend to exhibit a metal-like behavior, whereas the second ones are generally insulators or semiconductors. In a general way, the magnetic properties of these materials are governed by the low-spin configuration of the d^4 Ru(IV) and d^3 Ru(V) species. The ruthenates can be classified, according to their structure, in five categories: perovskites, Ruddlesden and Popper (RP), hexagonal perovskites, pyrochlores, and 1D Ln_3RuO_7. It must also be emphasized that ruthenium can substitute for various transition metal species of oxides modifying their physical properties in a spectacular way.

Ruthenate Perovskites

The ruthenates $CaRuO_3$ and $SrRuO_3$ and their solid solution constitute the basis of our knowledge for the understanding of magnetism in ruthenium oxides. Both of them have an orthorhombically distorted perovskite structure (Pnma with $a \sim c \sim a_c$ and $c \sim 2a_c$), but $SrRuO_3$ is much less distorted and often described as pseudocubic. Their transport properties also exhibit a similarity: they both exhibit a metal-like behavior, so that they are often described as bad metals (**Figure 1**), with comparable resistivities at room temperature. The latter properties have been explained very early by Goodenough by the strong covalent coupling between the Ru-4*d* and O-2*p* orbitals, leading to a narrow band model. Nevertheless, the magnetic properties of these oxides are very different: $SrRuO_3$ is ferromagnetic with a T_c of 165 K, whereas $CaRuO_3$ has been described for a long time as a paramagnetic metal involving antiferromagnetic interactions and is now believed by several authors to be nearly ferromagnetic, as supported by NMR observations. The different behavior of $SrRuO_3$ compared to $CaRuO_3$ is also observed on its $\rho(T)$ curve (**Figure 1**), which shows a Fischer–Langer type anomaly at T_c. The highly correlated nature of the 4*d*-electron band in these perovskites is

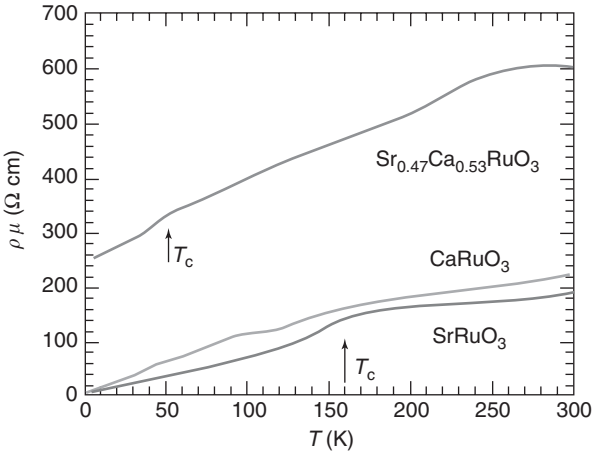


Figure 1 Resistivity vs. temperature for $Sr_{1-x}Ca_xMnO_3$ perovskites ($x = 0$, 0.53, 1). (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ ($0 \lesssim x \lesssim 1.0$). *Physical Review B* 56: 321–329; © American Physical Society.)

also confirmed by the recent studies of their solid solution $Sr_{1-x}Ca_xRuO_3$. The peculiar magnetic behavior of $CaRuO_3$, and the fact that it sits on the verge of a ferromagnetic instability is also supported by recent studies, which show that the substitution of Ti(IV) for Ru(IV) in this phase induces ferromagnetism for low Ti level (2%).

Besides, the Ru(IV) perovskites, a series of Ru-based perovskites with the generic formulation A_2LnRuO_6, have been reported for A = Ca, Sr, or Ba and Ln = La, Nd, Y, Ho, Lu, or Er. In all these oxides, ruthenium is pentavalent in contrast to $CaRuO_3$ and $SrRuO_3$. The crystal symmetry of these phases is often monoclinic ($P2_1/n$ $a \sim b \sim a_c\sqrt{2}$, $c \sim 2a_c$ $\beta \sim 90.2$) and more rarely cubic (Fm3m $a \sim 2a_c$).

All these oxides exhibit a 1 : 1 ordering between Ru(V) and Ln(III) or Ca(II) on the octahedral sites. Thus, in this structural type (**Figure 2**), one RuO_6 octahedron alternates with one LnO_6 or CaO_6 octahedron along the three crystallographic directions. In all these compounds, the Ru(V) lattice is antiferromagnetic at low temperature with T_N values ranging from 26 to 40 K. Most of them exhibit an A-type antiferromagnetic structure at low temperature (**Figure 3a**), except Ba_2LaRuO_6 which belongs to the third type (**Figure 3b**). Moreover, the lanthanide cations may be involved in the long-range antiferromagnetic order depending on their nature, as illustrated, for example, for Sr_2ErRuO_6 (**Figure 3c**) which consists of two interpenetrating A-type antiferromagnetic sublattices on Ru^{5+} and Er^{3+} sites, respectively. Though no transport measurements have been carried out, no metallic conduction should be expected for these materials at least at low temperature, due to the isolated character of the RuO_6 octahedra. Nevertheless, high-temperature magnetic susceptibility measurements, carried out on Ba_2YRuO_6, Ba_2LuRuO_6, and Sr_2LuRuO_6 suggest that Ru^{5+} *d*-electrons should be itinerant rather than localized. Such a feature may result from a direct overlap of the neighboring Ru-*t2g* orbitals. Double perovskites involving a disorder of ruthenium and of a transition element M over the same octahedral sites have been synthesized for M = Ni, Co, Fe, Cu. This is indeed the case for the compounds Sr_2MRuO_6 (M = Fe, Co), $BaLaMRuO_6$ (M = Ni, Co), and $SrLaCoRuO_6$, all of which exhibit a monoclinic symmetry, either I2/*c* or I2/*m* or $P2_1/n$, with $a \sim b \sim a_c\sqrt{2}$, $c \sim 2a_c$, and $\beta \sim 90.1^\circ$, except $BaLaNiRuO_6$ which is triclinic. These oxides exhibit a complex magnetic behavior, due to the fact that two alioelectronic cations are distributed at random over the sites of a nonfrustrated lattice. It often results in a spin-glass behavior as shown from the magnetic susceptibility curves of $BaLaCoRuO_6$ (**Figure 4a**) and of Sr_2FeRuO_6 (**Figure 4b**) which exhibit a cusp around 50 and 40 K, respectively. This behavior is significantly different from that of the double perovskites, $BaLaZnRuO_6$ and $BaLaMgRuO_6$, which exhibit an

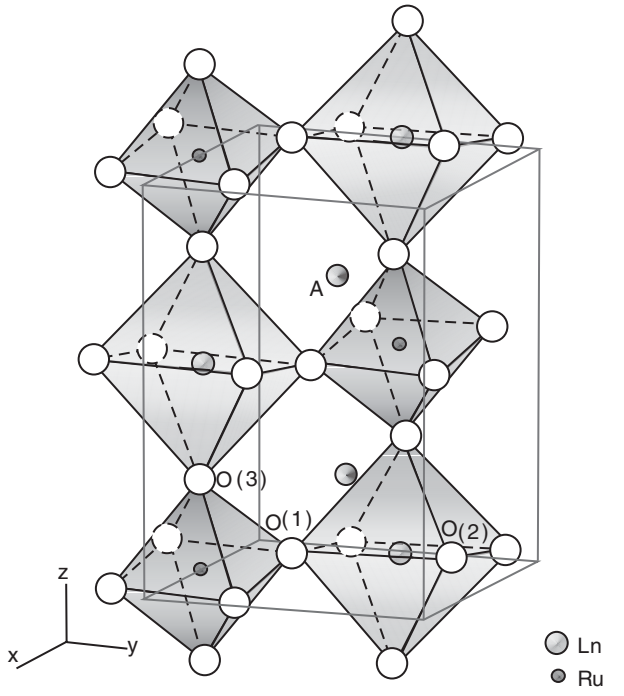


Figure 2 Crystal structure of the ordered perovskites A_2LnRuO_6. (After Battle PD, Jones CW, and Studer F (1991) *Journal of Solid State Chemistry* 90: 302–312.)

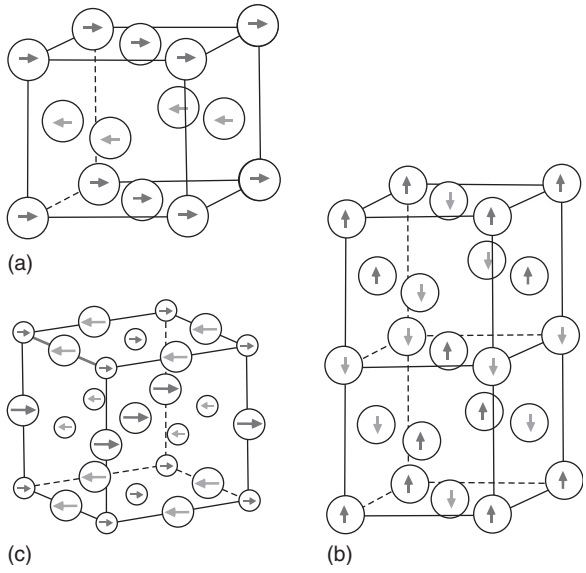


Figure 3 Antiferromagnetic structures of: (a) A-type A_2LnRuO_6 (Ln = Pr−Ln), (b) third-type Ba_2LaRuO_6, and (c) of Sr_2ErRuO_6 with two sublattices, Ru^{5+} (small circles) and Er^{3+} (large circles). (After Battle PD, Jones CW, and Studer F (1991) *Journal of Solid State Chemistry* 90: 302–312.)

A-type antiferromagnetic ordering at low temperature. This difference may be explained by the fact that the Ru^{5+} and Zn^{2+} (or Mg^{2+}) species, which are distorted according to a 1 : 1 ordering, do not interact magnetically, due to the d^{10} (Zn^{2+}) or d^0 (Mg^{2+}) configurations.

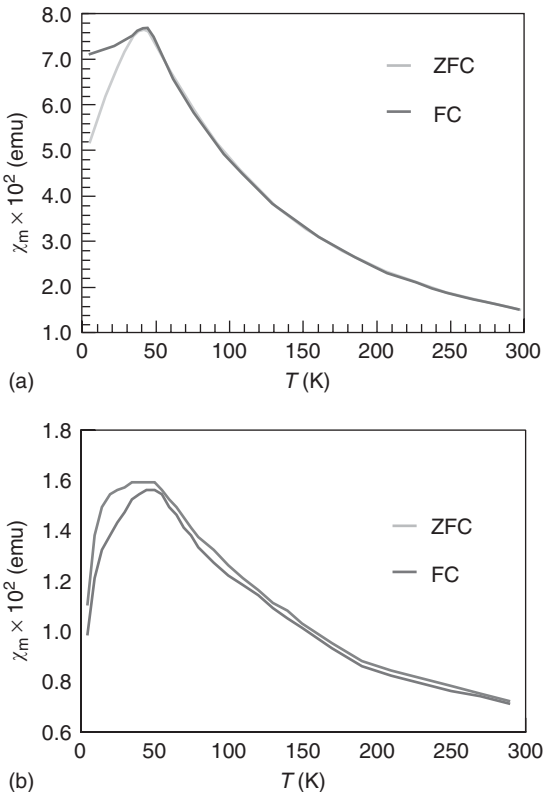


Figure 4 Magnetic susceptibility vs. temperature of (a) BaLaCoRuO$_6$ and (b) Sr_2FeRuO_6. (b) (After Battle PD, Jones CW, and Studer F (1991) *Journal of Solid State Chemistry* 90: 302–312.)

Ruddlesden and Popper Ruthenates

Ca_2RuO_4 and Sr_2RuO_4 represent the first member of the RP series $A_{n+1}Ru_nO_{3n+1}$ (A = Ca, Sr), that is, their structure (**Figure 5a**) consists of single perovskite layers "$SrRuO_3$" or "$CaRuO_3$" intergrown with single rock-salt layers "SrO" or "CaO." These quasi-two-dimensional ruthenates exhibit very different physical properties due to the different distortions of their structure. Ca_2MnO_4 exhibits a metal–insulator (MI) transition at $T_{MI} = 357$ K (**Figure 6a**). This Mott insulator is antiferromagnetic below $T_N = 110$ K. In contrast, Sr_2RuO_4 is a superconductor below $T_c = 1.5$ K and is paramagnetic above this temperature with a tendency toward ferromagnetism. The study of the system $Ca_{2-x}Sr_xMnO_3$ shows that the system changes successively with decreasing x, from $x = 2$ where it is a nonmagnetic Fermi liquid, to a ferromagnetic metal for $x = 0.5$, to an antiferromagnetically correlated metal for $0.2 < x < 0.5$, and to an antiferromagnetic insulator for $x < 0.2$. Neutron powder diffraction studies clearly demonstrate that the magnetic state changes are associated to structural transitions as shown from the phase diagram in **Figure 7**. Based on structural distortions, a magnetic phase diagram can be calculated which demonstrates that the rotation and the tilting of the RuO_6 octahedra are responsible for the ferro- and antiferromagnetism respectively, whereas the flattening of the RuO_6 octahedra controls the stability of these magnetic states. The substitution of La^{3+} for

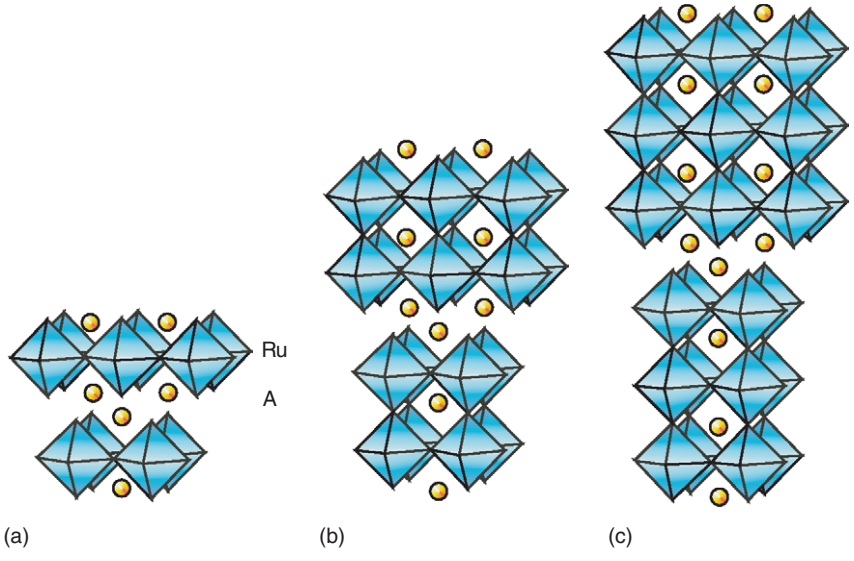


Figure 5 Structure of the RP ruthenates $A_{n+1}Ru_nO_{3n+1}$: (a) $n = 1A_2RuO_4$, (b) $n = 2A_3Ru_2O_7$, and (c) $n = 3A_3Ru_4O_{10}$.

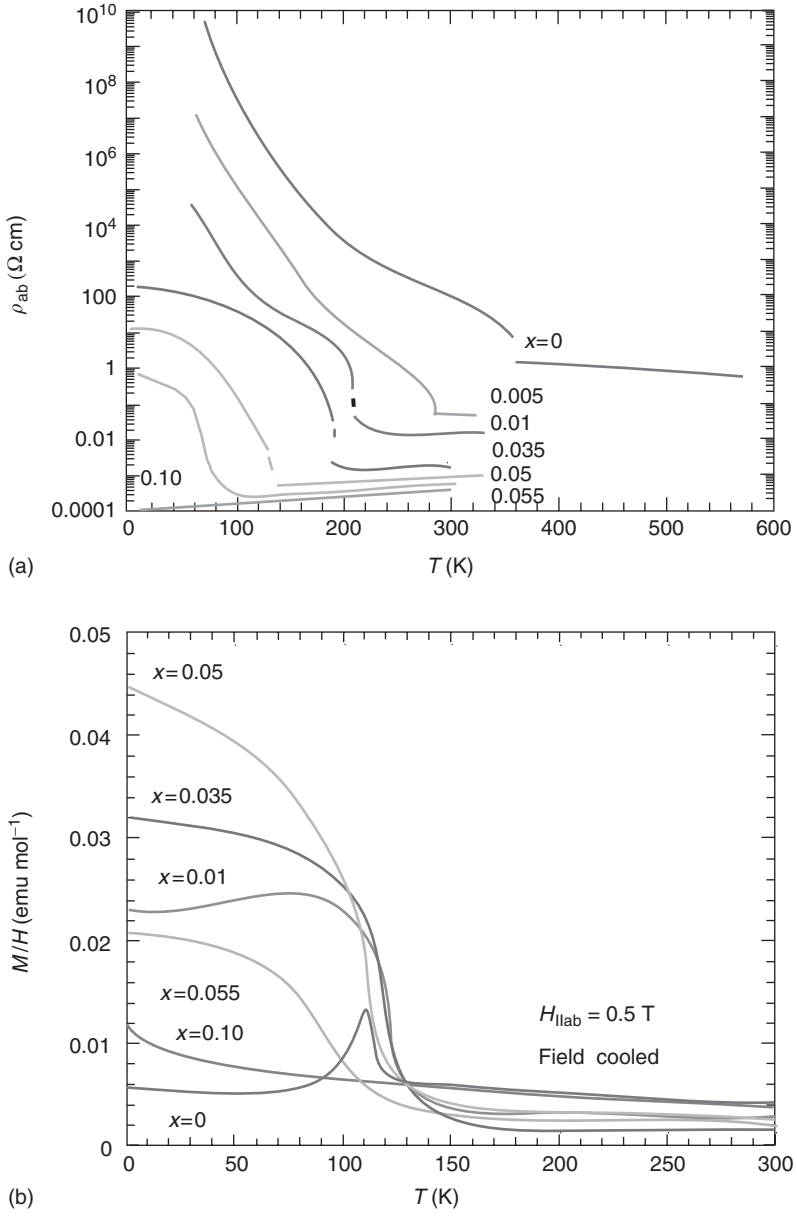


Figure 6 (a) Evolution of the resistivity and (b) of the magnetization vs. temperature for the oxides $Ca_{2-x}La_xMnO_4$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ ($0 \lesssim x \lesssim 1.0$). *Physical Review B* 56: 321–329; © American Physical Society.)

Ca^{2+} in Ca_2RuO_4 shows that the doping with very small amounts of lanthanum lowers the MI transition dramatically down to 100 K and simultaneously the resistivity at low temperature can be decreased by several orders of magnitude (**Figure 6a**). Concomitantly, ferromagnetism is induced in the material (**Figure 6b**). The Curie temperature can reach up to 100 K, and remains lower than T_{MI}, suggesting that the metal–insulator transition is not driven by the magnetic instability. These results emphasize the competition of ferromagnetism and antiferromagnetism in these ruthenates.

The $n = 2$ members of the RP series, $Ca_3Ru_2O_7$ and $Sr_3Ru_2O_7$ consist of double perovskite layers "$A_2Ru_2O_6$" intergrown with single rock-salt "SrO" layers (**Figure 5b**). $Ca_3Ru_2O_7$ exhibits, like Ca_2RuO_4, a metal-to-insulator transition as T decreases, with $T_{MI} = 48$ K. Its magnetic phase diagram (**Figure 8**) shows that it is an antiferromagnetic Mott insulator below 48 K, and becomes a metallic ferromagnet between $T_{MI} = 48$ K and $T_N = 56$ K. Above T_N, this phase becomes paramagnetic and remains metallic. Remarkably, the phase becomes metallic as the magnetic field is increased beyond 6 T, suggesting that the

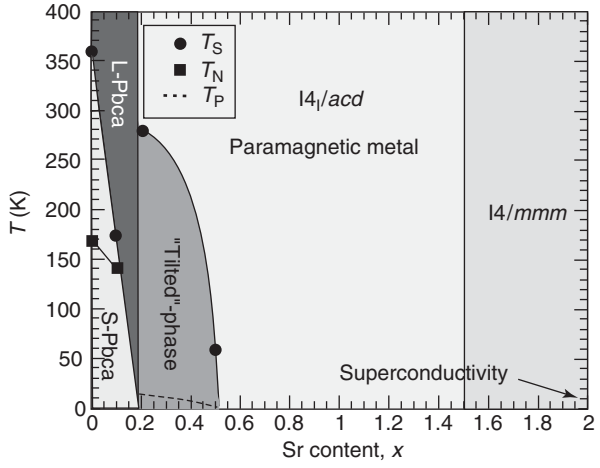


Figure 7 Phase diagram of $Ca_{2-x}Sr_xRuO_4$. (Reprinted with permission from Friedt O, Braden M, André G, Adelmann P, Nakatsuji S, *et al.* (2001) *Physical Review B* 63: 174432; © American Physical Society.)

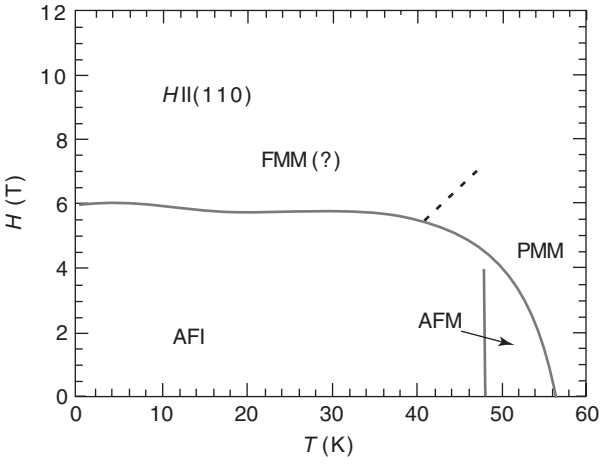


Figure 8 Magnetic phase diagram of $Ca_3Ru_2O_7$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ ($0 \lesssim x \lesssim 1.0$). *Physical Review B* 56: 321–329; © American Physical Society.)

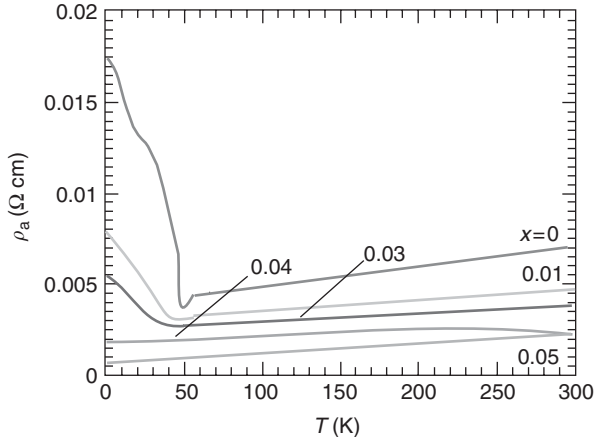


Figure 9 Resistivity vs. temperature for $(Ca_{1-x}La_x)_3Ru_2O_7$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ ($0 \lesssim x \lesssim 1.0$). *Physical Review B* 56: 321–329; © American Physical Society.)

system becomes ferromagnetic (FMM). Note that the confluence of the limits at $T = 48$ K and $H = 4.1$ T lead to a multicritical point. Similar to $Ca_{2-x}La_{x}RuO_4$, the doping of $Ca_3Ru_2O_7$ at Ca sites changes the magnetism and transport of this oxide rapidly. The resistivity is decreased as x increases up to 0.04 in $Ca_{3-x}La_xRu_2O_7$ (**Figure 9**): finally, a fully metallic state is obtained for $x = 0.05$. In contrast to $Ca_{2-x}La_xRuO_4$, no ferromagnetism is induced by doping, that is, the antiferromagnetic metallic state characteristic of $Ca_3Ru_2O_7$ persists. However, remarkably, the application of a magnetic field, even modest, induces substantial ferromagnetism. The $M(H)$ curves (**Figure 10**) show that M_{sat} decreases rapidly as x increases and disappears for $x = 0.04$. They demonstrate that the metamagnetic transition is associated with the insulator-to-metal transition. Thus, La-doped or not, $Ca_3Ru_2O_7$ exhibits a negative magnetoresistance which can reach as high as 88%, depending on the temperature and magnetic field.

In contrast to $Ca_3Ru_2O_7$, $Sr_3Ru_2O_7$ is an itinerant ferromagnet at low temperature. Single crystals of this compound are grown. The zero field cooled (ZFC) and field cooled (FC) $M(T)$ curves (**Figure 11**) of this oxide registered at low field show a large magnetic anisotropy below $T_c = 104$ K, the easy axis being near [0 0 1]. It also evidences a broad maximum below T_c at $H \perp (0\,0\,1)$ (inset **Figure 11**) suggesting the presence of noncollinear spin structure in the ferromagnetic matrix. The $M(H)$ curves registered at $H \perp (0\,0\,1)$ for low temperature ranges show that the spin system undergoes a transition to a parallel spin configuration for a critical value of the magnetic field smaller than 3 T. This reorientation transition disappears at about 66 K in agreement with additional transition observed at $T^* = 66$ K on the $M(T)$ curve (**Figure 11**). All these observations are explained by a possible canting of the spins away from the [0 0 1] directions for $T < T^*$. The resistivity also exhibits a very large anisotropy as shown from the $\rho(T)$ curves (**Figure 12**). ρ_c/ρ_{ab} is indeed close to 3 at room temperature in accordance with the bidimensional character of the structure which favors the delocalization of the carriers within the basal plane of the RuO_6 octahedra. Note that besides the anomalies at T^* and T_c on the $\rho(T)$ curve, there exists a third anomaly $\sim$250 K whose origin is unknown. Like $SrRuO_3$, this oxide exhibits negative magnetoresistance, as shown from the $MR(T)$ curve (**Figure 13**)

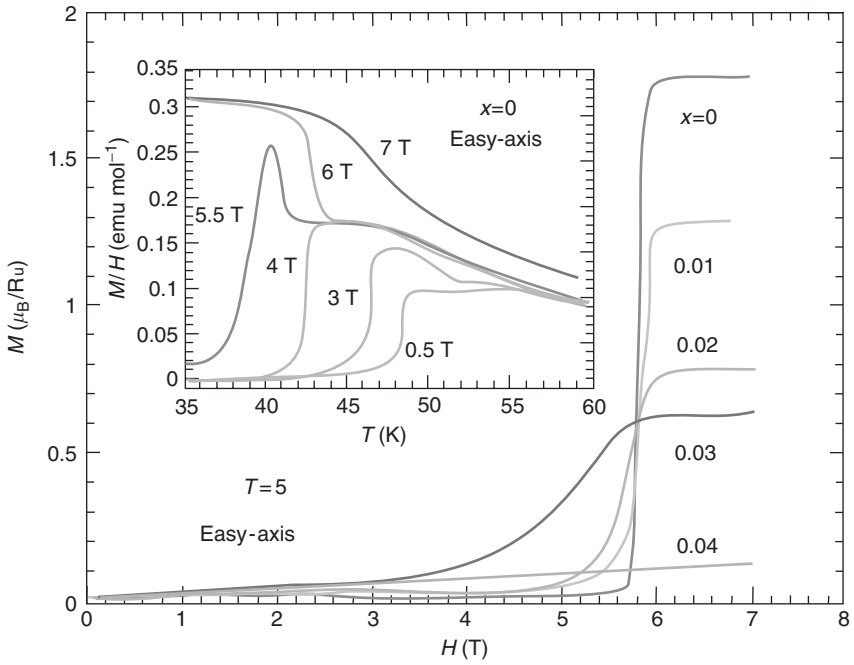


Figure 10 Magnetization vs. field at 5 K for $(Ca_{1-x}La_x)_3Ru_2O_7$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ $(0 \lesssim x \lesssim 1.0)$. *Physical Review B* 56: 321–329; © American Physical Society.)

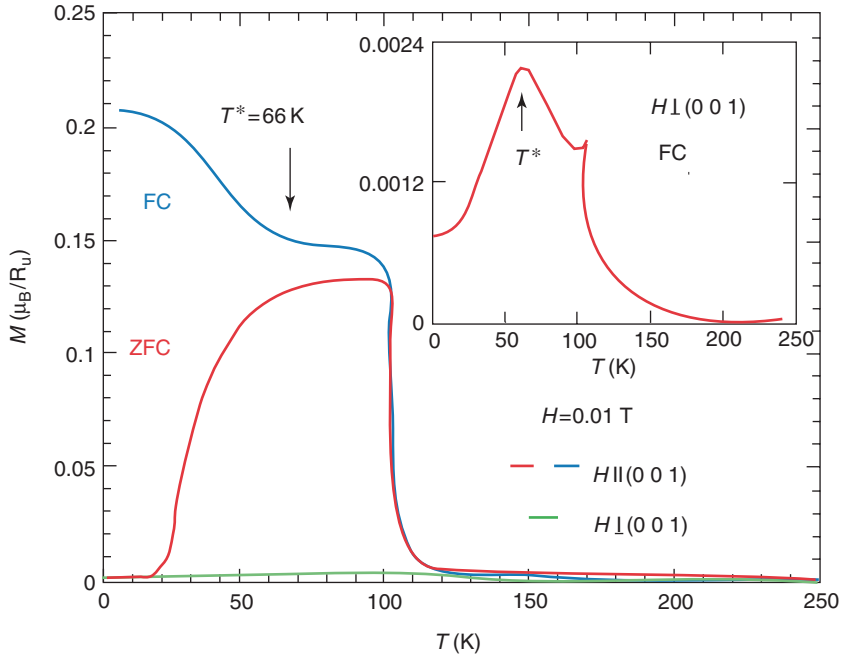


Figure 11 Magnetization vs. temperature at $H = 0.01$ T for $Sr_3Ru_2O_7$; red and blue lines: M vs. T with H II(0 0 1) measured in a zero field cooled (ZFC) and field cooled (FC) sequence; green line: M vs. T with $H\perp(0\,0\,1)$ after FC. Inset: M vs. T for FC at $H\perp(0\,0\,1)$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ $(0 \lesssim x \lesssim 1.0)$. *Physical Review B* 56: 321–329; © American Physical Society.)

with $MR = (\rho(0\mathrm{T}) - \rho(10\mathrm{T}))/\rho(0\mathrm{T})$. One indeed observes a peak of magnetoresistance of $\sim$15% below T^*, suggesting that the effect originates mainly from the spin reorientation under the magnetic field. In contrast, the effect of T_c is small compared to $SrRuO_3$ which exhibits a maximum MR of $\sim$10% at T_c.

The third member of the RP series, $Sr_4Ru_3O_{10}$ consists of triple octahedral layers "Ru_3O_9" intergrown

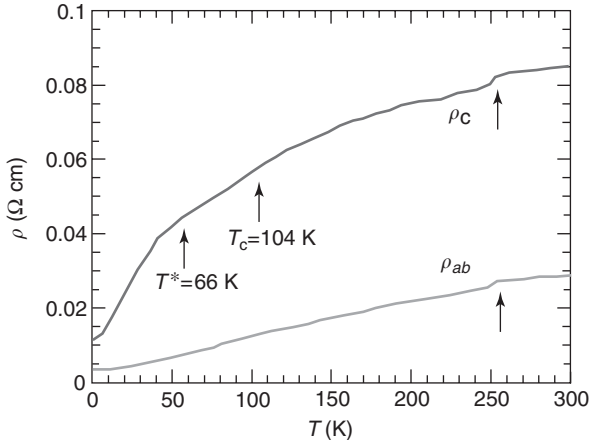


Figure 12 Resistivity vs. temperature T for $Sr_3Ru_2O_7$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ $(0 \lesssim x \lesssim 1.0)$. *Physical Review B* 56: 321–329; © American Physical Society.)

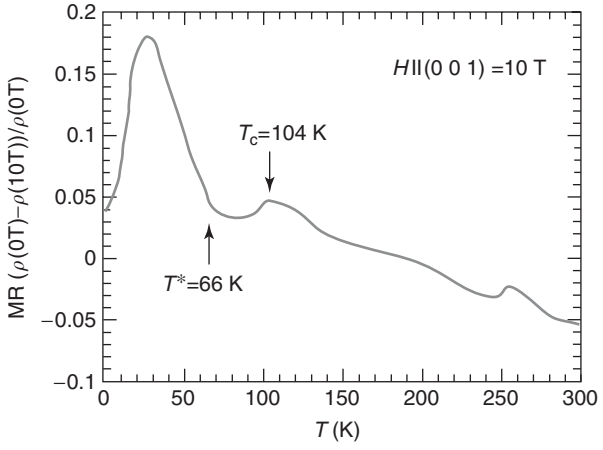


Figure 13 Magnetoresistance vs. temperature for $Sr_3Ru_2O_7$. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ $(0 \lesssim x \lesssim 1.0)$. *Physical Review B* 56: 321–329; © American Physical Society.)

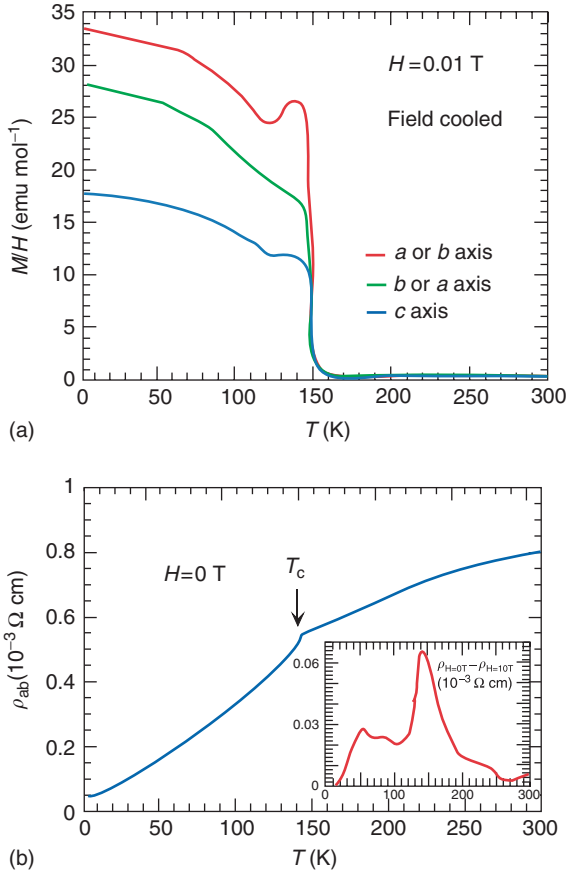


Figure 14 Magnetization (a) and resistivity (b) of $Sr_4Ru_3O_{10}$ vs. temperature. (Reprinted with permission from Cao G, McCall S, Shepard M, Crow JE, and Guertin RP (1997) Thermal, magnetic, and transport properties of single-crystal $Sr_{1-x}Ca_xRuO_3$ $(0 \lesssim x \lesssim 1.0)$. *Physical Review B* 56: 321–329; © American Physical Society.)

with single rock salt "SrO" layers. The single-crystal study of this compound shows that like $Sr_3Ru_2O_7$, this oxide is ferromagnetic, with a T_c of 148 K, intermediate between that of $SrRuO_3$ and $Sr_3Ru_2O_7$. The $M(T)$ curves (**Figure 14a**) show that, in contrast to $Sr_3Ru_2O_7$, the easy axis lies in the *ab*-plane. Also differently from the latter phase, there is no clear evidence for spin reorientation at low temperature. The saturation value of the magnetization, of 1.75 μ_β, is in agreement with the low spin configuration generally observed for Ru(IV) d^4. The $\rho(T)$ curve (**Figure 14b**) of this phase is similar to that observed for $SrRuO_3$ with a Fisher–Langer anomaly at $T_c =$ 148 K. Like for $SrRuO_3$, one observes a small magnetoresistance, maximum at T_c ($\sim$6% under 10 T) due to the disappearance of the Fisher–Langer anomaly under a magnetic field.

Comparing the magnetic properties of the different members of the RP series, and taking into consideration the properties of the perovskite $(n = \infty)$, it clearly appears that in the Sr-RP series, ferromagnetic strength and consequently T_c increases as n increases, that is, as the dimensionality of the structure increases. In the same way, in the Ca-RP series, it is remarkable that the tendency to ferromagnetism, that is, T_N decreases as n increases, that is, as the dimensionality of the structure increases.

Hexagonal Ruthenates and Relatives

For large A-site cations, as, for instance, in $BaRuO_3$, the perovskite structure is no more stable and transforms into a "hexagonal perovskite" by sharing partly the faces of the RuO_6 octahedra. In this way, three forms of $BaRuO_3$ are actually known, called 9L, 6L,

and 4L due to the number of layers L characterizing the periodicity.

Two forms 9L and 4L, have been studied as single crystals. Their structures (**Figure 15**) consist of two-face sharing octahedra and three-face sharing octahedra for the 4L and 9L, respectively. These units are then connected along c through the apices of their octahedra like in the perovskite structure. It results in two sorts of overlapping of the d orbitals of Ru(IV): direct Ru–Ru *d*-interactions for the face-sharing octahedra and superexchange Ru–O–Ru–*p* interactions forming 180° angles for the corner-shared octahedra. Both oxides exhibit an almost temperature-independent magnetic susceptibility so that no local moment is detected on ruthenium, and can be interpreted as Pauli paramagnetism of conduction electrons. $BaRuO_3$ 4L exhibits indeed a metal-like behavior in the range 300–4 K, in both in-plane and out-of-plane directions (**Figure 16**). In contrast, $BaRuO_3$ 9L exhibits a very unusual behavior: the conductivity is first metal-like from 300 to ~100 K, but then an upturn of the resistivity is observed below 100 K (**Figure 16**). Moreover, the resistivity of these materials is sensitive to annealing under different temperatures.

The ruthenates $Ba_3MRu_2O_9$, with M = Fe, Co, Ni, Cu, In, all exhibit a 6H structure built up of units of two-face shared octahedra occupied by Ru, interconnected by single MO_6 octahedra (**Figure 15c**). In these phases, ruthenium is either pentavalent or mixed valent Ru(IV)–Ru(V). In contrast to $BaRuO_3$, all these oxides are semiconductors, due to the isolation of the Ru dimers by the M atoms. Magnetic susceptibility measurements show that the M-3*d* species have local moments below 400 K and that compounds with M = Co, Ni, Cu exhibit a magnetic transition near 100 K. The magnetic structure of these oxides varies with the nature of the M element. The oxides $Ba_3MRu_3O_9$ with M = Mg, Ca, Cd, Sr also exhibit the same 6H structure with a similar ordering of the ruthenium ions in the face-shared bioctahedra and of the M cations in the isolated corner-shared octahedra. These oxides are also semiconductors as expected for pure Ru(V) phases. The magnetic susceptibility behavior suggests that the *d*-electrons of Ru(V) are delocalized between the two sites of the bioctahedra in agreement with the Heisenberg–Dirac–Van Vleck model.

Pyrochlore Ruthenates

Ruthenates with the pyrochlore structure (**Figure 17**), built up of corner-sharing octahedra were extensively studied, mainly for their transport properties, which are closely related to their oxygen nonstoichiometry. For instance, $Bi_2Ru_2O_7$ and $Pb_2Ru_2O_{6.5}$ were found to exhibit a metal-like behavior, with resistivity of $\sim 10^{-3}\,\Omega$ cm at room temperature and a Pauli paramagnetism. Based on XPS, UPS, and EELS

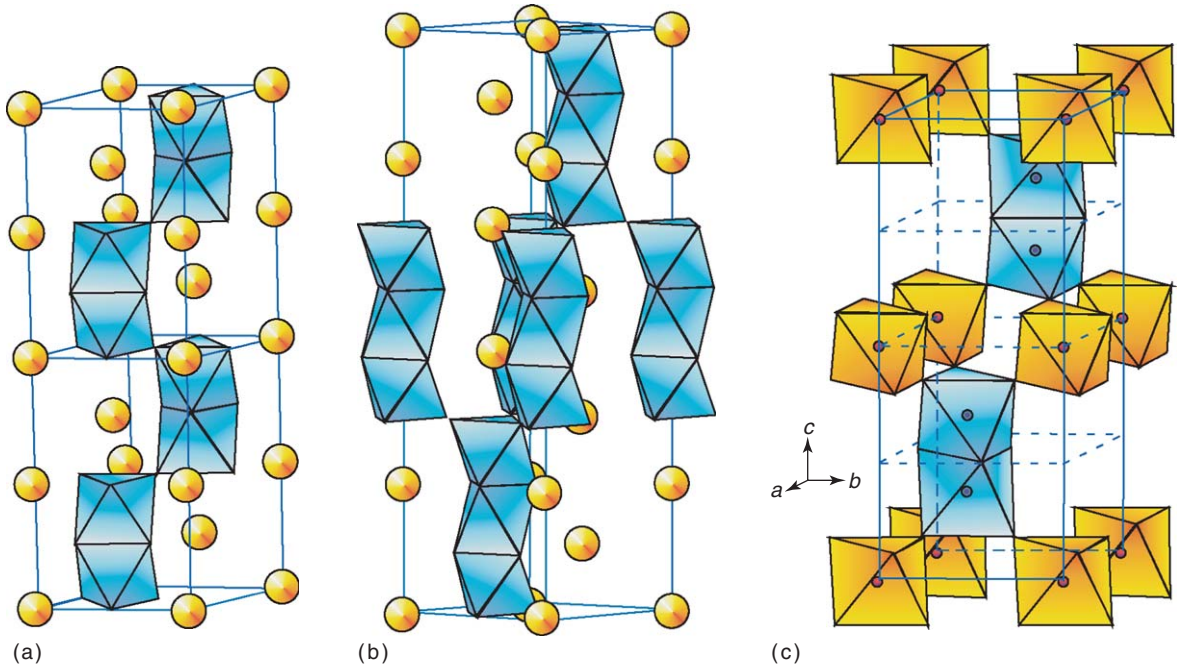


Figure 15 Structures of (a) 9L $BaRuO_3$, (b) 9L $BaRuO_3$, and (c) 6H $Ba_3MRu_2O_9$ where blue circles represent Ru and red circles M elements.

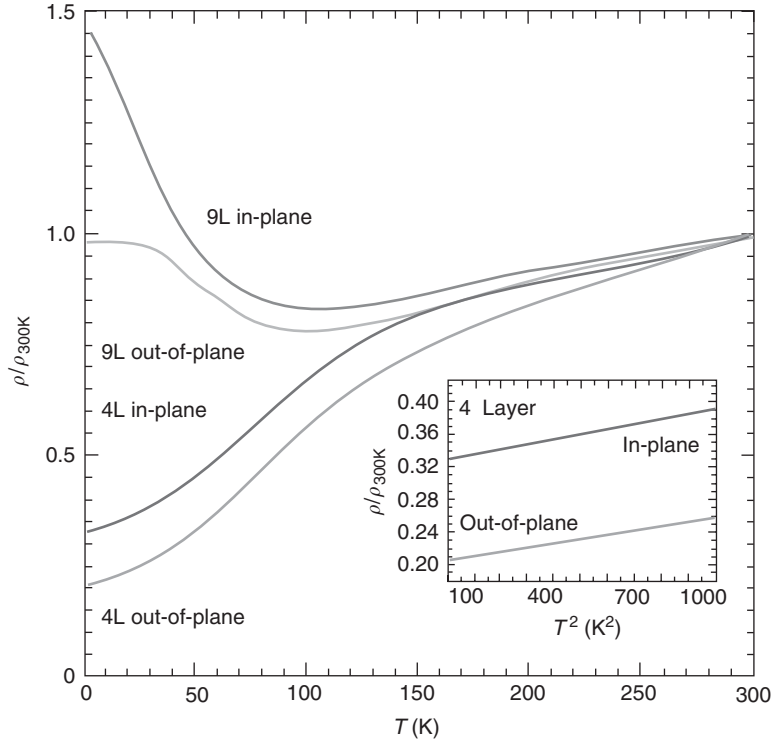


Figure 16 Resistivity vs. temperature for 4L and 9L $BaRuO_3$ forms. (Reprinted with permission from Rijssenbeek JT, Jin R, Zadorozhny Yu, Liu Y, Batlogg B, *et al.* (1999) Electrical and magnetic properties of the two crystallographic forms of $BaRuO_3$. *Physical Review B* 59: 4561–4564; © American Physical Society.)

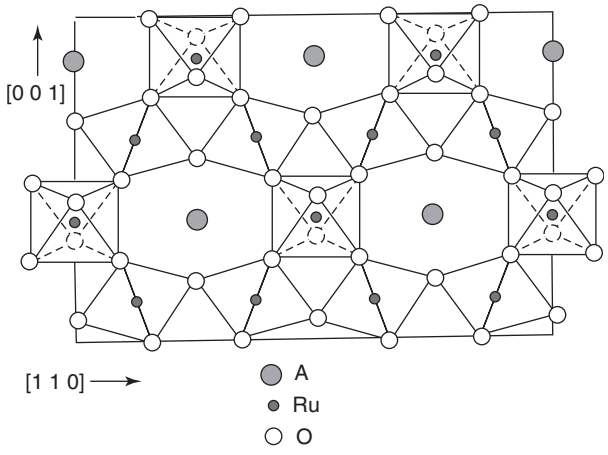


Figure 17 Pyrochlore structure of $A_2RU_2O_{7-\delta}$ ruthenates.

studies, it was shown that the *6p* states of Bi and Pb which are close to the Fermi level can participate in the conductivity by mixing with Ru-4*d*-orbitals via the 2*p*-oxygen-orbitals. In contrast, rare-earth ruthenates $Ln_2Ru_2O_7$ Ln = Pr−Lu, Y were found to be semiconductors with low activation energies. The difference in conductivity between $Ln_2Ru_2O_7$ and $Bi_2Ru_2O_7$ has been attributed to the bending of Ru–O–Ru band which is larger for $Ln_2Ru_2O_7$ (Ru–O–Ru angle $\approx 128^\circ$) than for $Bi_2Ru_2O_7$ (Ru–O–Ru angle $\approx 138^\circ$). This viewpoint is supported by the study of the solid solution $La_{2-x}Cd_xRu_2O_7$, which also exhibits metallic conductivity, for Ru–O–Ru band angles similar to $Bi_2Ru_2O_7$, but is characterized by large paramagnetic susceptibilities. Similarly, the studies of the solid solutions $Bi_{2-x}Ln_xRu_2O_7$ and $Pb_{2-x}Ln_xRu_2O_7$ with Ln = Y, Pr–Lu, clearly show that the metal–insulator crossover observed for these oxides at temperatures ranging from 40 to 80 K, is correlated to the bending of the Ru–O–Ru band and to the distortion of the RuO_6 octahedra. Finally, the pyrochlore $Tl_{2-x}Ru_2O_{7-\delta}$, which was synthesized under high pressure, is a good example of the great influence of cationic and oxygen non-stoichiometry upon the transport properties of this structural family, prepared under high pressure. In reducing conditions, these phases exhibit a metallic behavior (**Figure 18a**) whereas in neutral atmosphere (**Figure 18b**) or under high oxygen pressure (**Figure 18c**), they show a metal-to-insulator transition at about 40 and 120 K respectively. The latter transitions are correlated to structural changes versus temperature. The magnetic properties of these oxides are so far not completely elucidated showing spin-glass-like behavior and magnetic anomalies at different temperatures (40 and 120 K), depending on

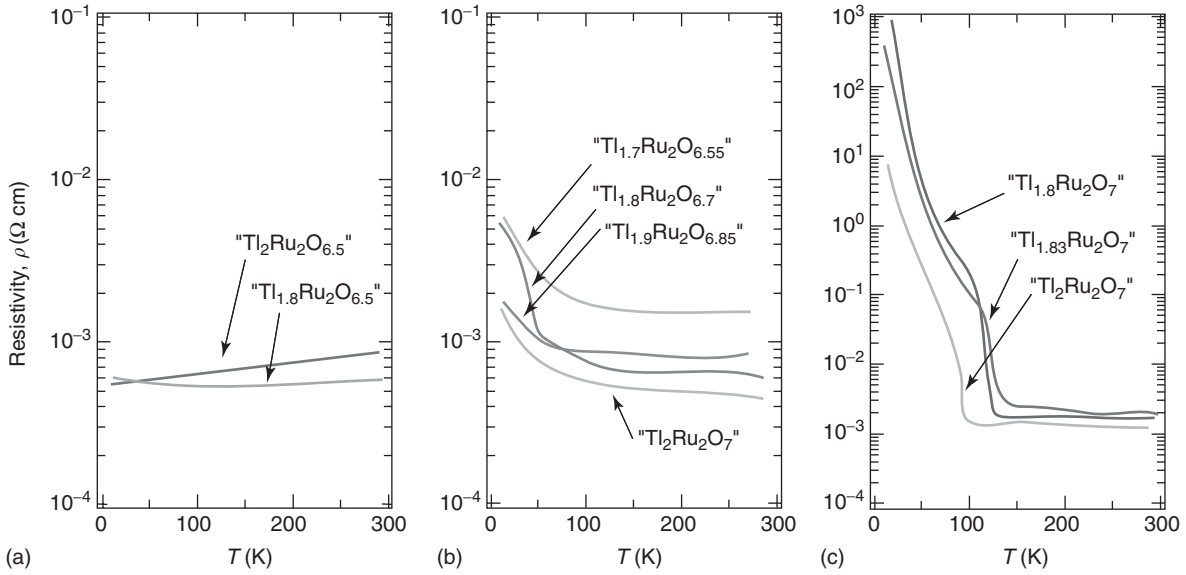


Figure 18 Temperature dependence of resistivity for thallium pyrochlores with various nominal compositions prepared under pressure (a) in reducing conditions (b) in neutral atmosphere, and (c) under high oxygen pressure. (After Takeda T, Nagata M, Kobayashi H, Kanno R, Kawamoto Y, *et al.* (1998) *Journal of Solid State Chemistry* 140: 182–193.)

the conditions of synthesis, that is, of the chemical composition of the so-studied pyrochlore.

The Quasi-One-Dimensional Ruthenates Ln_3RuO_7

The lanthanide ruthenates Ln_3RuO_7 (Ln = La−Sm) exhibit, from the viewpoint of Ru arrangement, a 1D-type structure built up of chains of corner-sharing RuO_6 octahedra. The magnetic and transport properties of La_3RuO_7 reflect its 1D geometry. One indeed observes that this phase is a semiconductor even along the chain, though the Ru–O–Ru angles of 145° are very similar to those observed for 3D pyrochlore ruthenates. The magnetic data show that this Ru(V) oxide exhibits a Curie–Weiss behavior above 50 K and an antiferromagnetic order below this temperature. In fact, a short-range antiferromagnetic order is observed in the range 20–50 K which becomes long range below 20 K.

Ruthenium as a Substituting Element in Oxides

The ability of ruthenium to adopt an octahedral coordination makes it possible to be substituted for transition elements in many oxides. As a consequence, its electronic configuration either as d^4 Ru(IV) or as d^3 Ru(V), and the broad extension of its d orbitals, suggests that such an element is susceptible to induce new physical properties in already known materials.

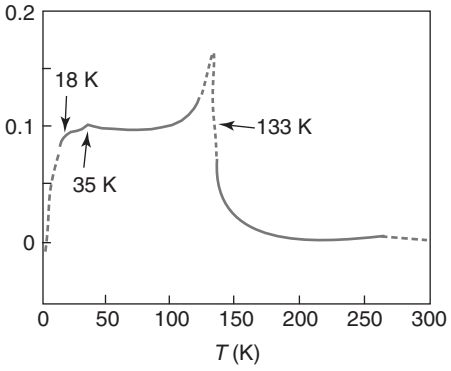


Figure 19 AC susceptibility measurements of $RuSr_2Cu_2O_8$, showing that superconductivity exists up to 35 K, with $T_c = 18$ K and $T_c = 133$ K. (Reprinted figure with permission from Chmaissem O, Jorgensen JD, Shaked H, Dollar P, and Tallon JL (2000) *Physical Review B* 61: 6401–6407; © American Physical Society.)

Two examples of this promising role of Ru to induce exciting properties are described.

The first example deals with the introduction of Ru in the 1212 structure of cuprates. It has been discovered that the 1212 ruthenocuprate $RuSr_2GdCu_2O_8$ is both a superconductor with a T_c = 15–40 K, and a ferromagnet with a $T_c = 133-136$ K (**Figure 19**). The coexistence of superconductivity and ferromagnetism below 15–40 K appears most surprising, since, according to Ginzburg investigations, such a feature should not be possible at a microscopic scale. The particular layered structure of this compound (**Figure 20**), which consists of superconducting pyramidal copper layers stacked with octahedral ruthenium layers, is certainly the key to

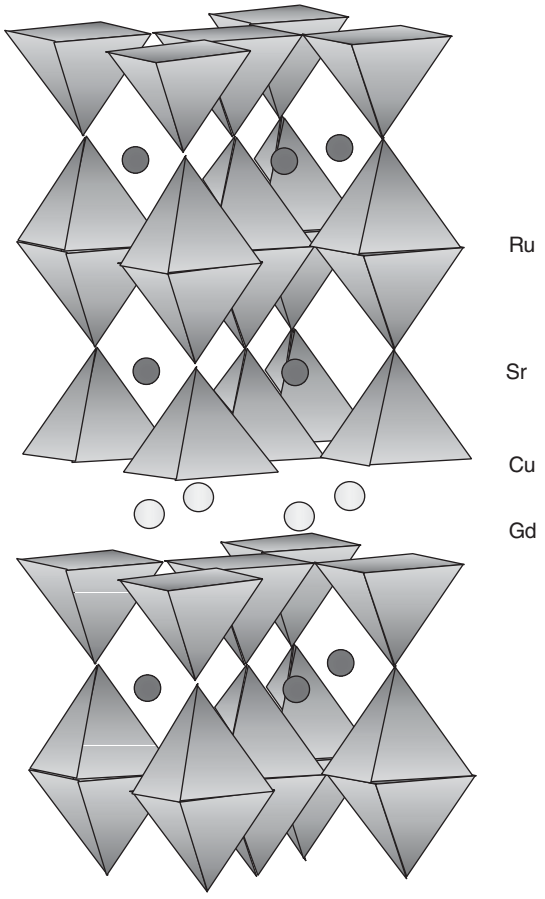


Figure 20 The 1212 structure of $RuSr_2GdCu_2O_8$ built up of layers of RuO_6 octahedra stacked with layers of CuO_5 pyramids.

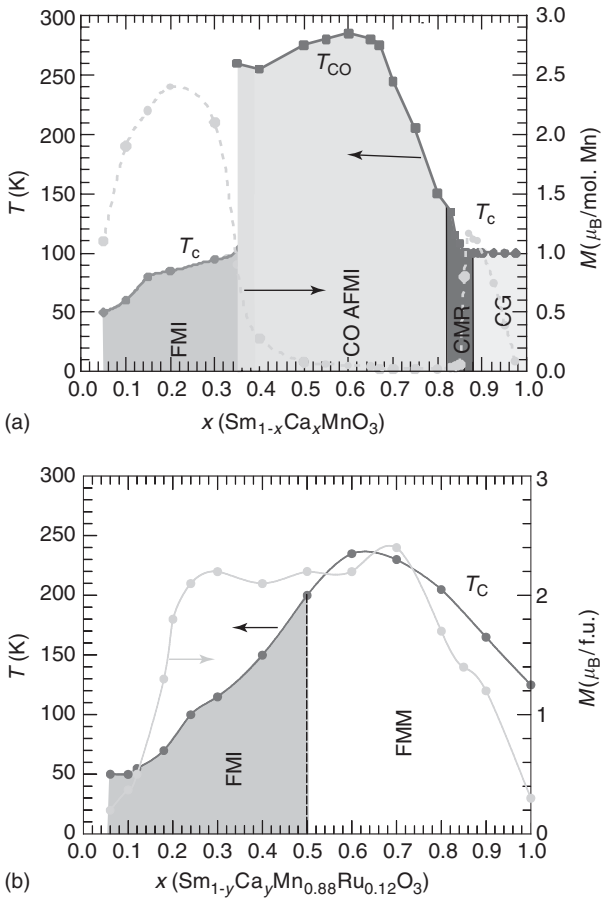


Figure 21 Magnetic phase diagrams of the manganites $Sm_{1-x}Ca_xMnO_3$ (a) and $Sm_{1-x}Ca_xMn_{0.88}Ru_{0.12}MnO_3$ (b).

the understanding of this behavior, since this structural separation should decrease the interactions between ferromagnetism and superconductivity dramatically. In any case, this extraordinary behavior suggests that numerous investigations are still in progress to help understand the physics of these materials.

The second example deals with the doping of perovskite manganites with ruthenium at the Mn site. It was shown that the partial substitution of ruthenium for manganese in charge-ordered manganites such as $Sm_{0.5}Ca_{0.5}MnO_3$, which are insulator ferromagnets, induces ferromagnetism and a metal–insulator transition, and consequently, colossal magnetoresistance (CMR) effects. Thus, very high T_c up to 220 K can be reached by Mn-site doping, changing completely the magnetic-phase diagram of the manganites. This is illustrated for $Sm_{1-x}Ca_xMnO_3$ (**Figure 21a**) doped with 12% Ru (**Figure 21b**). It can be seen that the T_c of the ferromagnetic insulating region is increased by Ru substitution, but, more importantly, that the broad charge-ordered antiferromagnetic domain (**Figure 21a**) is suppressed and replaced by a ferromagnetic metallic region with a T_c up to 200 K. This great ability of Ru to induce ferromagnetism and metal-like behavior can be explained by both, its ability to couple ferromagnetically with manganese neighbors, and the large extension of its *d*-orbitals allowing conduction paths to be enhanced.

See also: Ferromagnetism; Magnetic Materials and Applications; Magnetism, History of; Superconductivity: General Aspects.

PACS: 71.30.+h; 72.80.−r; 72.80.Ga; 74.70.Pq; 74.72.Jt; 75.47.Pq; 75.50.Cc; 75.50.Dd; 75.50.Ee; 75.50.Lk

Further Reading

André G, Adelmann P, Nakatsuji S, *et al.* (2001) *Physical Review B* 63: 174432.

Battle PD, Gibb TC, Jones CW, and Studer F (1989) *Journal of Solid State Chemistry* 78: 281–293.

Battle PD and Jones CW (1989) *Journal of Solid State Chemistry* 78: 108.

Cao G, Abboud K, McCall S, Crow JE, and Guertin RP (2000) *Physical Review B* 62: 998–1003.

Cao G, McCall S, and Crow JE (1997) *Physical Review B* 55: R672–R675.

Cao G, McCall SK, Crow JE, and Guertin RP (1997) *Physical Review B* 56: R5740–R5743.

Cao G, McCall S, Crow JE, and Guertin RP (1997) *Physical Review Letters* 78: 1751–1754.

Cao G, McCall S, Dobrosavljevic V, Alexander CS, Crow JE, *et al.* (2000) *Physical Review B* 61: R5053–R5057.

Goodenough JB (1971) *Metallic Oxides*. Oxford: Pergamon.

Kanno R, Kawamoto Y, Izumi F, and Sleight AW (1998) *Journal of Solid State Chemistry* 140: 182–193.

Khalifah P, Erwin RW, Lynn JW, Huang Q, Batlogg B, *et al.* (1999) *Physical Review B* 60: 9573–9578.

Kim SH and Battle PD (1995) *Journal of Solid State Chemistry* 114: 174–183.

Liu Y, Batlogg B, *et al.* (1999) *Physical Review B* 59: 4561–-4564.

Martin C, Maignan A, Hervieu M, Autret C, Raveau B, *et al.* (2001) *Physical Review B* 63: 174402.

Martin C, Maignan A, Hervieu M, and Raveau B (1999) *Physical Review B* 60: 12191–12199.

Rao CNR and Raveau B (eds.) (1998). *Colossal and Magnetoresistance, Charge Ordering and Related Properties of Manganese Oxides*. Singapore: World Scientific.

Tallon JL, Loram JW, Williams GVM, and Bernhard C (2000) *Physical Review B* 61: R6471–R6474.

Tokura Y (ed.) (1999) *Colossal Magnetoresistive Oxides*. New York: Gordon and Breach.

S

Scanning Near-Field Optical Microscopy

A Cricenti, ISM-CNR, Rome, Italy

Introduction

Historical Background

At the beginning of the twentieth century, a method was proposed to overcome the resolution limits of classical optical microscopy. The idea was based on raster-scanning an aperture much smaller than the wavelength in close proximity to the sample. Furthermore, in order to obtain a subwavelength resolution, such a scan should be performed at very small distances from the sample. This proposal was demonstrated to be very far ahead of its time, but due to the necessity of performing the scan at less than 10 nm from the sample, it could not be verified experimentally. The first proof of this idea came in the microwave region half a century later. Subsequently, after the invention of the scanning tunneling microscope (STM) in 1982, it became possible to rely on the introduction of experimental techniques able to maintain a strongly localized tip at distances <10 nm from the surface, whereas the tip–sample distance control was realized by monitoring a well-determined physical parameter in the interaction region (in the case of the STM, this parameter is the tunneling current between tip and surface). In 1984 scanning near-field optical microscopy (SNOM) was reinvented on the basis of the proposed method and opened the doors to a new kind of spectroscopy.

Basic Principle of SNOM

SNOM is based on the detection of the evanescent waves that are directly connected to the subwavelength features of the object: such waves do not propagate and are confined in the near-field region, decaying on distances of the order of fractions of the wavelength λ. They are revealed and transformed into propagating waves, in air or inside an optical fiber, by a nanodetector (tapered optical fiber) located few nanometers from the surface. In this way, it is possible to overcome the $\lambda/2$ limit imposed by diffraction and to obtain information on sample details absolutely not accessible to conventional microscopy. In order to try to understand this phenomenon, some very simple arguments in the theory of the electromagnetic field are first summarized.

Far Field and Near Field

What is the electromagnetic field generated by an electrical dipole placed at a given point O with the moment-vector parallel to the z-axis?

The electromagnetic field generated by an electrical dipole in a point $P(r, \theta, \phi)$ (see **Figure 1**), placed at a distance r from the origin O, and with the condition that the dipole axis $d \ll r$, can be calculated by using the Maxwell equations, starting from the potential vector A_z and the scalar potential $V(\boldsymbol{r}, t)$, assuming an oscillating retarded function of time $p = p_0 \sin \omega(t - r/c)$ for the emitted light with p' and p'' as its first and second derivative, respectively, and $p_0 =$ constant:

$$A_z = \frac{\mu_0}{4\pi}\frac{p'}{r}$$

$$V(\boldsymbol{r}, t) = \frac{1}{4\pi\varepsilon_0}\left(\frac{p'}{rc} + \frac{p}{r^2}\right)\frac{z}{r}$$

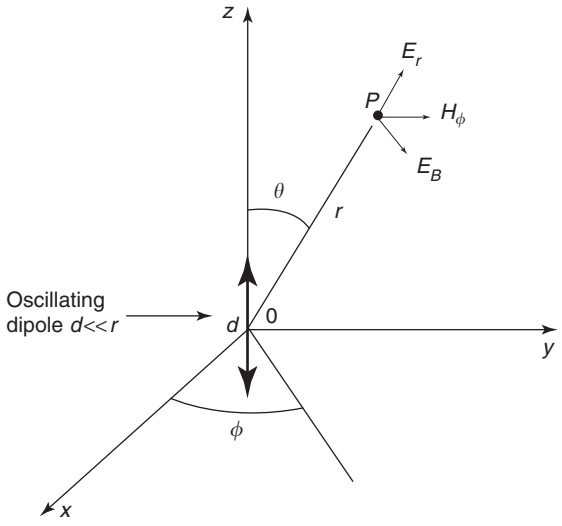


Figure 1 Diagram of the electromagnetic field generated by a dipole oscillating along z-axis.

Using spherical coordinates,

$$A_r = \frac{\mu_0}{4\pi}\frac{p'}{r}\cos\theta, \quad A_\theta = \frac{-\mu_0}{4\pi}\frac{p'}{r}\sin\theta, \quad A_\phi = 0$$

$$V = \frac{1}{4\pi\varepsilon_0}\left(\frac{p'}{rc} + \frac{p}{r^2}\right)\cos\theta$$

With

$$\boldsymbol{B} = \text{rot}\,\boldsymbol{A}$$

$$B_r = 0, \quad B_\theta = 0, \quad B_\phi = \frac{\mu_0}{4\pi}\left(\frac{\sin\theta}{r}\right)\left(\frac{p''}{c} + \frac{p'}{r}\right)$$

$$\boldsymbol{E} = -(\text{grad}\,V + \boldsymbol{A})$$

$$E_r = \frac{1}{4\pi\varepsilon_0}\left(\frac{2\cos\theta}{r}\right)\left(\frac{p'}{cr} + \frac{p}{r^2}\right)$$

$$E_\theta = \frac{1}{4\pi\varepsilon_0}\left(\frac{\sin\theta}{r}\right)\left(\frac{p''}{c^2} + \frac{p'}{cr} + \frac{p}{r^2}\right), \quad E_\phi = 0$$

Note that the electric field is perpendicular to the magnetic field and that there are terms depending on $1/r$, $1/r^2$, and $1/r^3$; different components have a different meaning. The energy associated with the electromagnetic field traveling in a certain direction with velocity c is due only to the terms containing $1/r$. In fact, for $r \gg \lambda$, i.e., for large distances,

$$B_r = 0, \quad B_\theta = 0, \quad B_\phi = \frac{\mu_0}{4\pi}\left(\frac{p''}{rc}\right)\sin\theta$$

$$E_r = 0, \quad E_\theta = \frac{1}{4\pi\varepsilon_0}\left(\frac{p''}{rc^2}\right)\sin\theta, \quad E_\phi = 0$$

The Poynting vector and its flux are given by

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} = \frac{\mu_0}{16\pi\varepsilon_0 c^3}\left(\frac{p''}{r}\right)^2 \sin^2\theta$$

$$F(\boldsymbol{S}) = \int Sr^2 \sin\theta\, \mathrm{d}\theta = \text{constant}$$

The flux does not depend on the radius of the sphere. There is no decrease or accumulation of energy when the wave is passing and the wave has an independent identity. If the condition of near-field, $r \ll \lambda$, is satisfied, the terms $1/r^2$ and $1/r^3$ present in the electromagnetic field must also be taken into account. In the calculation of the Poynting vector, new terms with higher-order dependence in r (r^{-3}, r^{-4}, and r^{-5}) are present and correspond to nonradiating field components. By time average, one obtains

$$S' = S \quad \text{and} \quad F(S) = F(S')$$

The terms $1/r^2$ and $1/r^3$ do not contribute to the traveling wave: they describe the near field that is important only for $r \ll \lambda$.

Experimental Setup

SNOM Configurations

There are many configurations in which SNOM can work, both in reflection and in transmission. Some of the most used ones are:

1. *Transmission collection mode*. The tip collects the light transmitted by the sample illuminated by a conventional source;
2. *Transmission illumination mode*. The tip plays the role of a nanosource and the light transmitted is collected by a conventional detector;
3. *External reflection collection mode*. The sample is illuminated by a laser beam and the tip collects the light scattered locally and drives it into the detector;
4. *Internal reflection collection mode*. The illumination beam is totally reflected inside a piece of glass and the tip collects the resulting evanescent waves lying on the sample surface;
5. *External illumination collection mode*. It is a combination of (1) and (2);
6. *Internal reflection illumination mode*. It is the inversion of (4) so that the light is detected beyond the critical angle (where the conditions of total reflection are realized); and
7. *Internal reflection perturbation mode*. A completely opaque tip perturbs the evanescent field on the object surface generated as in (4).

Photocurrent measurements can also be performed by directly illuminating the sample by the optical fiber. In this case the back contact of the specimen is directly fed into an *I–V* converter preamplifier with a fixed gain to reveal the electron current induced by the incoming photons.

Besides these configurations, the SNOM can work in two different modes of operation: *constant-height mode* (the tip raster-scans the surface at fixed height), and *constant-distance mode* (the tip raster-scans the surface maintaining the same gap from the sample).

For any of the SNOM arrangements, in order to monitor the tip position, a shear-force control mechanism is operating: in this configuration the tapered tip is forced to oscillate laterally (see **Figure 2**) by a piezoceramic at its resonance frequency. When the fiber is approaching the surface, the vibration amplitude begins to decrease progressively with distance due to the tip–sample interaction. By monitoring the

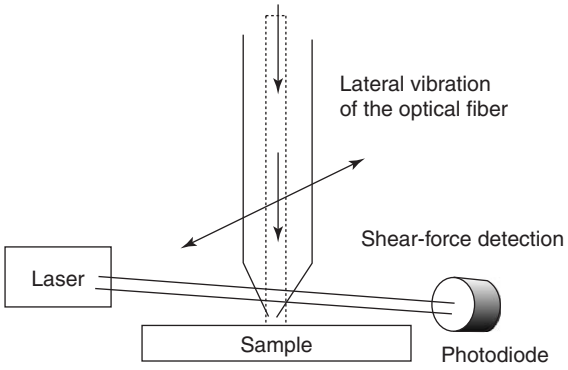


Figure 2 External reflection collection mode and shear-force control.

shear-force signal it is possible to maintain the tip at nanometric distance from the sample and to obtain a topographic image of the sample area over which the scan is realized. The electronic apparatus allows one to collect one or more optical images simultaneously, so that information concerning the local sample optical properties, the dielectric constant, etc., are obtained with a resolution well below $\lambda/2$.

The possibility to change the wavelength λ of the external laser beam allows the observation of the presence of topographic, chemical, or electronic defects in the optical image with a lateral resolution limited by the size of the fiber-tip aperture (~ 50 nm), and a vertical resolution that depends both on the light penetration depth and on the matrix element that governs the interaction between the light and the sample. Considering a penetration depth proportional to λ, we can suppose that the observed defect in the optical image is at a distance less than λ from the sample surface. This kind of investigation turns out to be very important in characterization of defects in layers below transparent films not accessible to a topographic analysis and in the observation of toxic compounds inside biological samples.

Instrument Description

As already stated, in order to bypass the diffraction limit, it is necessary to have a subwavelength aperture of the tapered fiber and small structures in the sample so that features smaller than $\lambda/2$ can be observable in the optical image. How do we create a subwavelength aperture at an optical fiber? SNOM probes for visible and near-infrared light can be fabricated starting from single-mode silica fibers (range of transmission, 450–2000 nm), whereas infrared SNOM (range of transmission, 2000–10 000 nm) probes can be made from single-mode chalcogenide (arsenic sulphide) fibers (**Figure 3**).

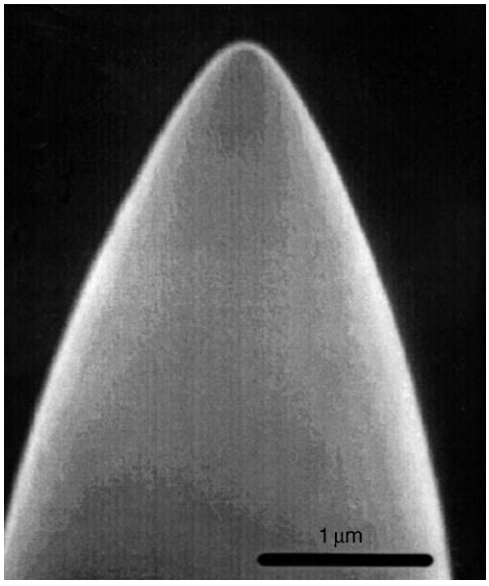


Figure 3 Scanning electron micrograph of a chalcogenide fiber SNOM tip.

One end of the fiber is interfaced to the detector while the other end is reduced to a small aperture using a standard micropipette puller or chemically etched using a protective layer etching system. The tips so obtained are then coated with a metal, aluminum, or gold, using a thermal evaporation deposition system at a pressure of 10^{-6} mbar. In order to leave an aperture, the tips are put at an angle of 25–30° above the evaporation point source and rotated to achieve a fairly uniform coating with an approximate coating thickness of 100–125 nm. With this procedure, tips with an aperture ranging between 20 and 100 nm are easily obtained.

The main components of an SNOM unit are a sample movement and a detection system combined in a way that, apart from the requirements of reliability and mechanical stability, would allow analyzing a sample with all available geometries for input/output of photons, in order to get as much information as possible on the sample. This instrument is shown in **Figure 4** and in the picture in **Figure 5** and is essentially shaped like a two-piece cylinder, the lower part containing the sample holder with its movements and the upper part hosting the fiber and the detection system. The two halves are held together by means of a screwlock that is used both to fit samples of several millimeters height and for a rough approach between fiber and sample.

The fiber tip fits tightly into a small tube, attached to a piezoceramic, used to vibrate the fiber for a few nanometers, thus allowing a shear-force signal, and mechanically attached to the top half of the cylinder. The lower half of the cylinder contains the sample

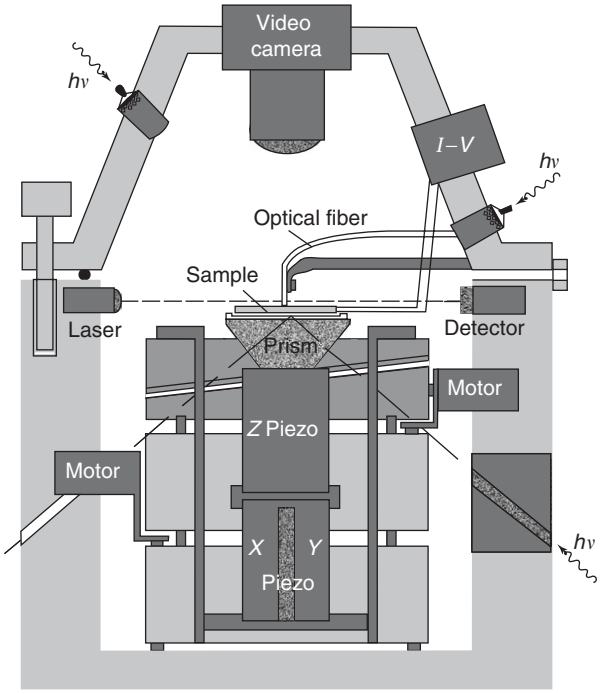


Figure 4 Schematic diagram of the mechanical part of the SNOM head (upper cylinder) and the sample holder (lower cylinder).

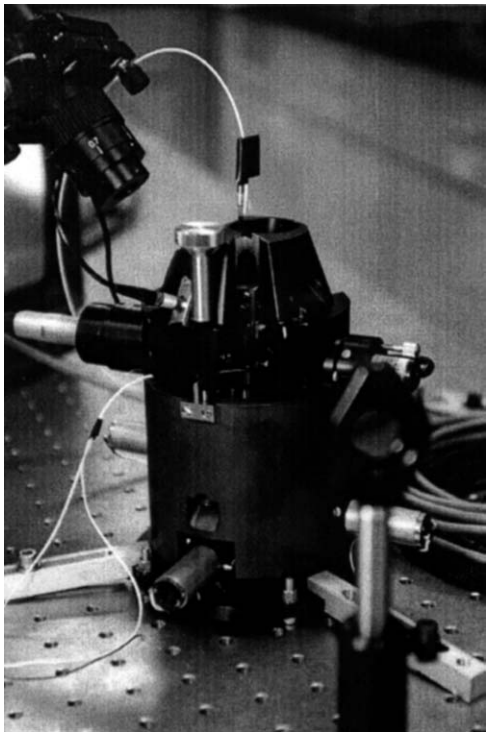


Figure 5 Diagram of the SNOM described in **Figure 4**.

holder, a piezoelectric scanner and *x*–*y*–*z* motor-controlled translators. The piezoscanner consists of two piezoceramic hollow cylinders glued together, with the upper one used for *z*-motion and a magnetic piece for holding samples attached to a magnetic support. The front and back contacts of the sample holder are used to feed a signal into an *I*–*V* converter in order to measure the electron current induced by the incident photons on the sample. The lower piezo is divided into four perpendicular sectors for *x*–*y* motion with an available *z*-motion of $\sim 8.0\,\mu m$ and an *x*–*y* scanning motion of $\sim 40.0\,\mu m$ for each axis. The piezoscanner is fixed onto a motorized *x*–*y*–*z* translator that can be moved several millimeters along the three axes. A video camera is usually mounted onto the top cylinder in order to control the rough approach between tip and sample and to position different regions of the sample below the probe.

The instrument allows measuring the local reflectivity when the sample is illuminated by an external source either directly or through the fiber: in both cases the optical fiber picks up the light reflected by the surface. Several detectors, positioned at the other end of the optical fiber, can be used to detect light depending on the photon wavelength. The detected signal is revealed by a lock-in amplifier whose analog output is converted by an analog-to-digital converter (ADC) module and read by the computer through a digital I/O card. The fiber can also detect the evanescent wave when the sample is illuminated in the back.

Figure 6 shows the layout of an SNOM experiment. The shear force is detected through a 670 nm single-mode diode laser that is focused onto the external body of the oscillating fiber and the resulting shadow is revealed by a two-sector position-sensitive detector and fed into the electronic feedback loop to keep a constant shear-force signal between fiber and sample while scanning the sample. Shear-force images are taken with a signal that is 70% of the initial amplitude value of free oscillation. A feedback circuit operates in such a way that as the sample approaches the fiber and a change in the shear-force amplitude is detected, the feedback circuit becomes active and the stepping is stopped: this method avoids fiber crashing onto the specimen.

A major requirement of any SNOM apparatus is a suitable photon source which must be both intense and tunable. Tunability is required to cover the relevant absorption bands that are present on the sample. Intensity is a critical point because of the inefficient light transmission of the narrow fiber tip. Laser modules and free electron laser (FEL) sources are ideal for this problem because of their unique combination of extreme intensity and broad tunability.

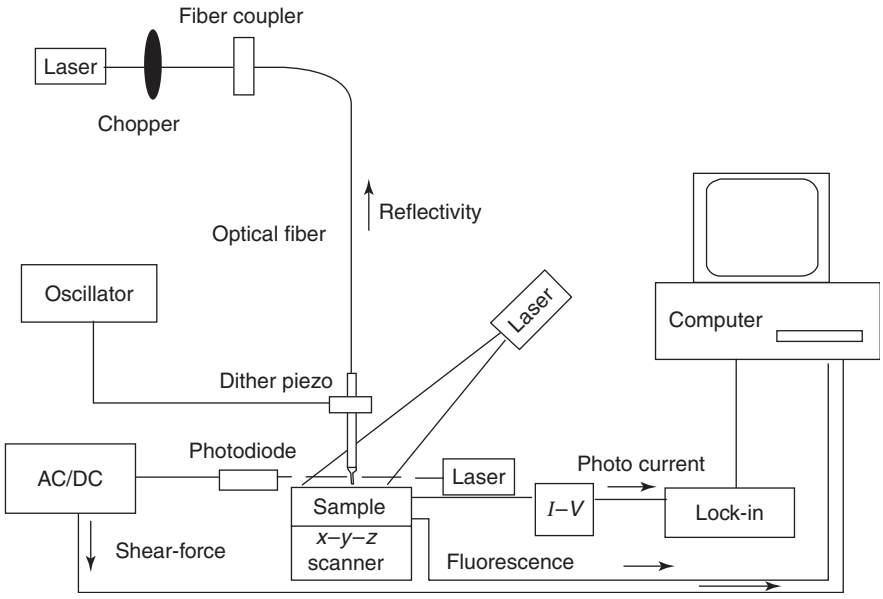


Figure 6 Electronic layout and data acquisition system for SNOM.

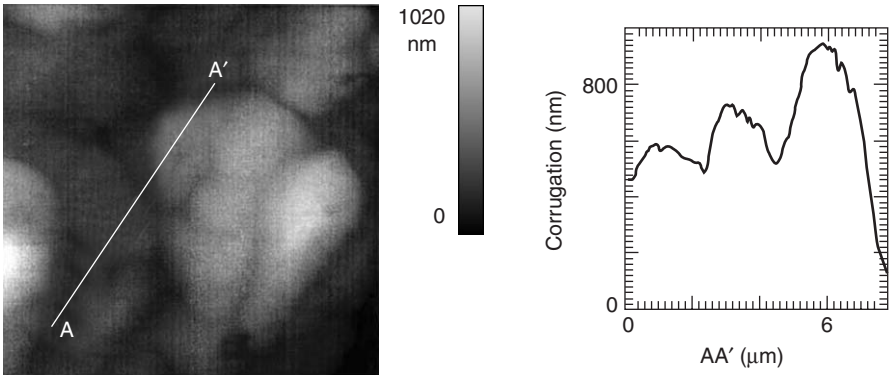


Figure 7 Shear-force topography image ($10 \times 10\,\mu m$) of diamond grains on a silicon substrate together with the corrugation along the AA′ line.

Experimental Results

Some results on representative samples will now be presented to give an overview of possible applications in materials science and biology and the great potential of the SNOM technique when used in a spectroscopy mode with an appropriate tunable source.

Diamond on Silicon

The first results concern a polycrystalline diamond film prepared by vapor-phase deposition on a silicon substrate. In principle, these should be hydrogen-free specimens after annealing to remove residual hydrogen and water. This point was tested by looking with Fourier transform infrared spectroscopy (FTIR) for absorption near 3.5 μm, due to the C–H vibrational stretch mode. In fact, the observed presence of such a band in the FTIR spectrum indicates that some hydrogen is left in the sample after annealing.

Figure 7 shows a $10 \times 10\,\mu m$ shear-force image together with the corrugation along the AA′ line. Small features on the individual diamond grains are quite well resolved: from reproducible structures in several consecutively acquired images, the best lateral resolution of the probe is estimated to be between 50 and 80 nm. Since irregularities in the probe are expected to be a large detriment to lateral resolution on rough specimens, these topographical images indicate that the probe tip itself is smoothly rounded.

Figure 8 shows a $10 \times 10\,\mu m$ reflectivity image of the same zone as in **Figure 7**, with the FEL operating at 3.5 μm, together with the corrugation along the

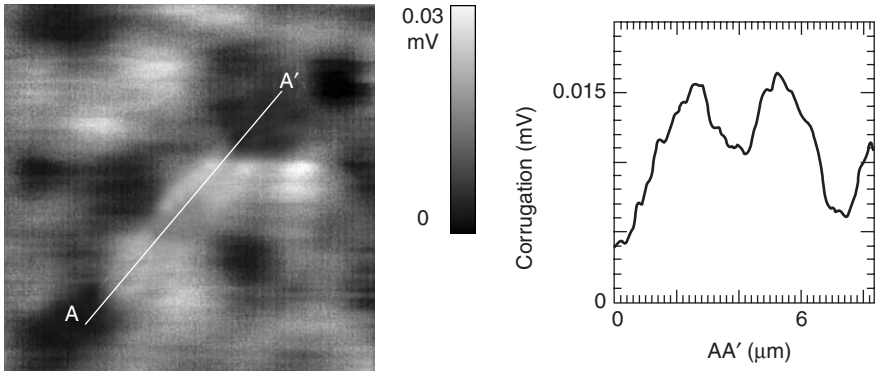


Figure 8 Reflectivity image, taken simultaneously with the topography of **Figure 7**, of diamond grains on a silicon substrate taken with a photon wavelength of 3.5 μm, together with the corrugation along the AA′ line.

AA′ line. The correlation between optical data and certain features in the topographical image can be seen clearly. When the FEL is tuned outside the C–H absorption band (to 3.2 or 3.7 μm), the topographical image is reproduced, whereas the optical image becomes featureless, indicating that the dark regions in the optical image of **Figure 8** correspond to regions where the absorption was stronger, or where there was residual hydrogen in the film. Although the relative intensity noise in the FEL (which could not be normalized out in the apparatus used for this experiment) is evident in **Figure 8**, several small features and edges are still resolved. From these, the optical lateral resolution was estimated to be 100 nm, well below the classical limit of $\lambda/2$.

Boron-Doped Silicon

B^+ ion implantation in silicon is the most useful *p*-type doping process and it is crucial to determine as-implanted and postanneal impurity distributions with the highest possible accuracy. Since implantation of energetic ions into silicon induces a large amount of crystal damage in the semiconductor consisting of high concentrations of interstitials and vacancies much higher than the implant dose, rapid thermal annealing is used to restore lattice order and to achieve dopant activation. Any annealing step is a highly nonequilibrium process and involves the formation/dissolution of extended defects, metastable phases and precipitates. As a result, both point defects and dopant atoms experience a high mobility characterized by anomalous diffusion, namely, transient enhanced diffusion. Dopant atoms may also form inactive immobile clusters and may deactivate much below their solid solubility. Since the solid solubility of B in Si is limited to 2×10^{20} atoms per cm^3, for increasingly high implant doses there is an increasing tendency for impurity atoms to cluster within the Si lattice after annealing.

In this case, a combination of internal photoemission (IPE) and SNOM has been used to investigate the electronic properties of silicon after boron implantation and annealing. Three different wavelengths have been used (633 nm (not shown), 1330 nm, and 1550 nm (not shown)) to locally illuminate the sample. Shear-force, reflectivity, and photocurrent images were collected simultaneously and compared in order to discriminate between true structures and artifacts. For the acquisition of R images the tapered optical fiber locally illuminates the sample and collects the light normally reflected at the surface by means of a fiber coupler while the photocurrent signal is measured by an *I–V* converter preamplifier with a gain of $10^8\,V\,A^{-1}$.

Depending on the level of boron concentration, the photocurrent shows intensity variations within two orders of magnitude. Boron clusters behave as metal clusters embedded into the silicon matrix and introduce gap states, which give rise to the observed photocurrent. The images reported here are for an Si(1 0 0) sample with the mean boron concentration values, in the first micron of depth, of $\sim 5 \times 10^{19}\,cm^{-3}$.

Figure 9 shows a 10×10 μm image of the shear-force (a), photocurrent (b), and reflectivity (c) taken with a photon wavelength of 1330 nm, that is below the indirect bandgap of silicon: as visible, there is no correspondence between the shear-force and photocurrent images acquired on the same silicon sample scanned area. As evidenced by the white traces (AA′) in both the images, there is a reflectivity enhancement with no correspondence in the shear-force image and that might be due to a subsurface defect. The shear-force and reflectivity images show surface-related structures not related to boron and give rise to local minima in the photocurrent image due, probably, to scattering of the impinging light. The photocurrent image shows the presence of a buried boron cluster, which has no correspondence in the

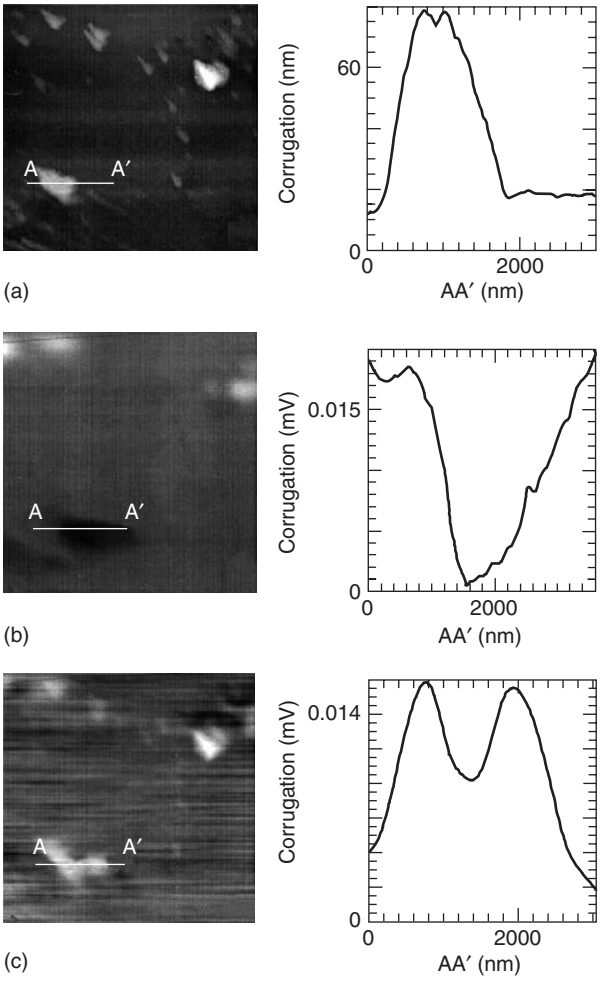


Figure 9 A $10 \times 10\,\mu m$ image of the shear-force (a), photocurrent (b), and reflectivity (c) taken with a photon wavelength of 1330 nm, for a Si(1 0 0) sample with the mean boron concentration values, in the first micron of depth, of $\sim 5 \times 10^{19}\,cm^{-3}$.

shear-force and reflectivity images and is characterized by lateral dimensions of several hundred nanometers. A resolution less than 100 nm (i.e., $\sim\lambda/13$) is observable in both the reflectivity and the photocurrent images that have been collected in many different areas of the sample. However, it has to be mentioned that these results are certainly affected by the convolution of the tip aperture with the penetration depth, $\alpha = \lambda/4\pi k$, of the light and by the subsequent diffusion of the excited carriers from the excitation volume, which should deteriorate the ultimate resolution.

Biological Growth Medium

An important first step in the SNOM evaluation of biological structures is the IR characterization of the biological growth medium. The growth medium is composed of several products, primarily sulfur and nitrogen oxide compounds, whose vibrational stretch mode absorption bands occur in the infrared $\sim 7\,\mu m$. The SNOM images presented here were taken by tuning the FEL to $\lambda = 6.95\,\mu m$, that is, at such absorption band, and at $6.6\,\mu m$, which is outside the absorption band, in order to estimate the noise-background contributions. Also in this case, the tunability is required to cover the relevant absorption bands, and intensity is a critical point because of the inefficient light transmission of the narrow fiber tip. FEL sources are ideal for this problem because of their unique combination of extreme intensity and broad tunability. The FEL photons are sent to the sample surface and detected after reflection by the narrow-point optical fiber tip mounted on the SNOM module.

Figure 10 shows two $20 \times 20\,\mu m$ SNOM reflection images obtained with $\lambda = 6.95\,\mu m$ photons (image (c)) and with the adjacent wavelength, $\lambda = 6.6\,\mu m$ (image (a)). In the reflection images, darker areas correspond to stronger absorption. The contrast between the featureless image off-absorption and the microstructures of the on-absorption image is striking. Are these microstructures really related to the growth medium constituents or are they just artifacts? To answer this question, it is important to compare the corresponding shear-force (topological) image of (b). First, the same topological image is observed at the two different wavelengths (image (b)). The topological images show the presence of several biological growth medium grains with width and height of a few micrometers. Furthermore, for artifacts, one would expect a similar optical image when collecting data at $6.6\,\mu m$ (image (c)) and $6.95\,\mu m$ (image (a)): absolutely no one-to-one correlation is observed between the two SNOM images, ruling out artifacts. This proves, in fact, that the dark regions in the on-absorption SNOM image of (c) correspond to regions with strong growth medium component concentrations related absorption.

It is quite instructive to analyze the lateral resolution of the topographic and spectroscopic-SNOM images obtained in this work. For the topographic images, line scans yield a lateral resolution of 50–80 nm, demonstrating the high quality of the fiber tips. On the other hand, similar line scans for the spectroscopic SNOM images (see the slope in the profile AA′) demonstrate a lateral resolution of 200 nm, which is well below the classical-microscopy limit ($\lambda/2$).

Fluorescence SNOM on Biological Cells

Another particularly exciting application of SNOM is in the observation of fluorophore agents inside a

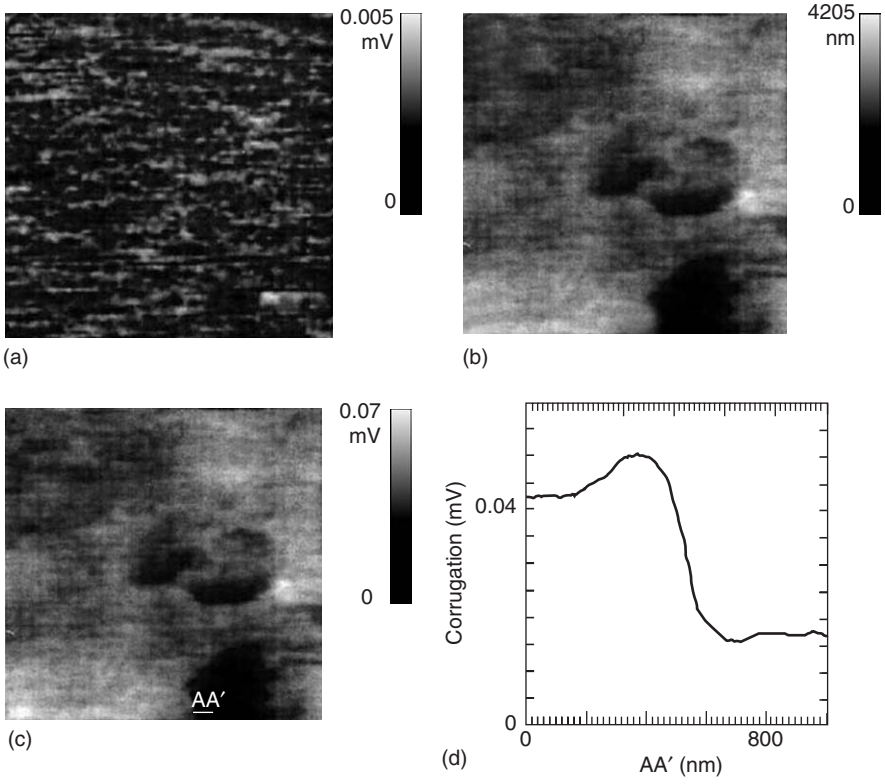


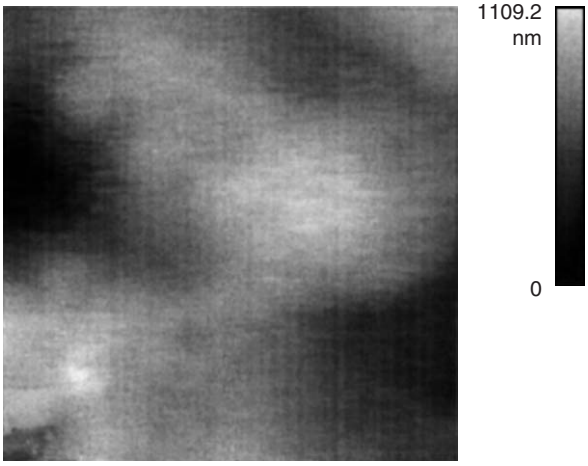
Figure 10 Left: $20 \times 20\,\mu m$ SNOM reflection images obtained with $\lambda = 6.95\,\mu m$ photons (c), corresponding to the vibrational stretch mode absorption bands of sulfur and nitrogen oxide compounds, and with an adjacent wavelength outside the absorption band, $\lambda = 6.6\,\mu m$ (a). Darker areas correspond to stronger absorption. Right: corresponding shear-force (topographic) image (b), brighter areas correspond to higher topography values, and a corrugation (d) taken along the AA′ line on the reflectivity image taken with $\lambda = 6.95\,\mu m$ in (c).

biological cell. The effect of a 50 Hz electromagnetic field (EMF) on morphological and biochemical properties of human keratinocytes (HaCaT) is being shown here. Control and exposed cells were immunolabeled with a fluorescent antibody to test the cells for the presence of $\beta 4$ integrin, a molecule marker for cell adhesion and differentiation. In this case the tapered metal-coated optical fiber is illuminating the sample through its small aperture, acting as a subwavelength light source, exciting the fluorophore locally, at a wavelength of 488 nm. Thus, only a small volume of the sample is optically illuminated, minimizing the fluorophore bleaching and the biological damage normally occurring in the optical fluorescence microscopy. The fluorescence signal is detected at 514 nm after blocking the excitation wavelength.

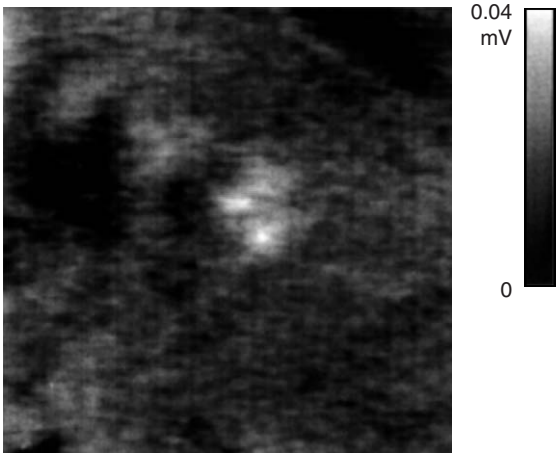
Figure 11a is a $20 \times 20\,\mu m$ topographic image of exposed cells showing that, after exposure to an electromagnetic field, cells become much closer to each other. The corresponding SNOM fluorescence image is reported in **Figure 11b**, showing strong fluorescence signal at all cell structures, demonstrating the presence in the cell membrane of the adhesion molecule $\beta 4$ integrin (a protein implicated in cell–cell adhesion, binding cytoplasmic domain of cadherin, and in signal transduction). It is worth observing that, at the border of the cells, fluorescence is accumulating in filamentous structures that can be defined as gap junctions between cells. On the contrary, no fluorescence was observed in control cells.

Perspectives

SNOM is becoming more and more important in current research in materials science and biology. The examples shown here indicate that one is far away from the limits in resolution of the technique, and the use of tunable infrared FEL radiation together with near-field optical collection opens the way to realize a nondestructive tool able to measure unknown properties of solid samples as well as of biological samples. The tunability of the source as well as the variation in the angle of incidence will allow a three-dimensional reconstruction of the shape and position of features and local inhomogeneities deeply located in the sample with the "finger-print region" between 2 and 10 μm of particular interest in view of its



(a)



(b)

Figure 11 Shear-force (a) and fluorescence (b) SNOM image ($28 \times 28\,\mu m$) of HaCaT cells exposed for 24 h to a 50 Hz magnetic field.

spectroscopic selectivity. Particularly interesting are the efforts to reliably develop the technique to realize photochemical processing on the nanometer scale, the characterization of cellular and subcellular structures, and also the *in vivo* monitoring of external agent uptake. New setups are currently very promising in the so-called SNOM "apertureless" approach that combines the high resolution obtained in STM with optical spectroscopy, and once the method is reliably established, it will overcome any limitations imposed by the limited transmission of the optical fiber.

See also: Confocal Optical Microscopy; Optical Microscopy; Scanning Probe Microscopy.

PACS: 78.20. – e; 78.40. – q; 07.79.Fc

Further Reading

Betzig E and Trautman JK (1992) Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit. *Science* 257: 189–195.

Fillard JP (1996) *Near Field Optics and Nanoscopy*, pp. 153–217. Singapore: World Scientific.

Grafstrom S (2002) Photoassisted scanning tunneling microscopy. *Journal of Applied Physics* 91(4): 1717–1753.

Hecht B, Pohl DW, Heinzelmann H, and Novotny L (1995) Tunnel near-field optical microscopy: TNOM. In: Marti O and Möller R (eds.) *Photons and Local Probes*, pp. 93–107. The Netherlands: Kluwer Academic Press.

Heinzelmann H and Pohl DW (1994) Scanning near-field optical microscopy. *Applied Physics A* 59: 89–101.

Isaacson M, Cline J, and Barshatzky H (1992) Near-field optical microscopy. In: Wickramasinghe HK (ed.) *Scanned Probe Microscopy*, pp. 23–36. New York: American Institute of Physics.

Paeasler M and Moyer P (1996) *Near-Field Optics: Theory, Instrumentation, and Applications*. New York: Wiley.

Pohl DW (1992) Nano-optics and scanning near-field optical microscopy. In: Wiesendanger R and Guntherodt H-J (eds.) *Scanning Tunneling Microscopy II*, Springer Series in Surface Science, vol. 28, pp. 233–272. Berlin: Springer.

Sandoghdar V (2001) Trends and developments in scanning near-field optical microscopy. In: Allegrini M, Garcia N, and Marti O (eds.) *Proceedings of the International School of Physics "Enrico Fermi,"* pp. 65–119. Amsterdam: IOS Press.

Van Hulst NF, Segerink FB, and Bolger B (1992) An evanescent-field optical microscope. In: Wickramasinghe HK (ed.) *Scanned Probe Microscopy*, pp. 79–94. New York: American Institute of Physics.

Van Labeke D and Barchiesi D (1993) Theoretical problems in scanning near-field optical microscopy. In: Pohl DW and Courjon D (eds.) *Near field Optics*, pp. 157–178. The Netherlands: Kluwer Academic Press.

Nomenclature

B	magnetic intensity ($A\,m^{-1}$)
c	velocity of light in free space ($m\,s^{-1}$)
d	length (m)
E	electric intensity ($V\,m^{-1}$)
$F(S)$	flux of Poynting vector
I/V	current to voltage (AV^{-1})
k	wave vector (μm^{-1})
p'	first derivative of p
p''	second derivative of p
r	distance (m)
S	Poynting vector
t	time (s)
ε_0	permittivity of free space ($F\,m^{-1}$)
λ	wavelength (μm)
μ_0	permeability of free space ($H\,m^{-1}$)
ω	angular frequency (Hz)

Scanning Probe Microscopy

J Gomez-Herrero, Universidad Autonoma de Madrid, Madrid, Spain
R Reifenberger, Purdue University, W. Lafayette, IN, USA

Introduction

The scanning probe microscope (SPM) is an extremely versatile instrument that has steadily evolved from its invention in the early 1980s. SPMs are now routinely available in many research labs throughout the world and are widely acknowledged for ushering in the study of matter at the nanoscale range.

The underlying principles of an SPM are quite simple but yet completely different in many significant ways from traditional microscopes. Essentially, the SPM works by positioning a sharp tip (often called a proximal probe) $\sim$1 nm above a substrate. The highly local information provided by the microscope is achieved by a combination of the sharpness of the tip, as well as the small separation between the tip and substrate. The critical feature of any SPM is the ability to maintain a constant tip–substrate distance (with a precision approaching a few picometers), while the tip is rastered across the substrate in a highly controlled way. To achieve this precision, a signal must be acquired that is very sensitive to the tip–substrate separation. The exact physical origin of this signal then determines the property of the substrate that is mapped by the SPM. A key discovery during the development of the SPM was the realization that with a sufficiently sharp tip, a quantitative three-dimensional image of surfaces can be obtained, often with atomic resolution.

The worldwide interest in scanning probe instruments was ignited by the research accomplishments of G Binnig and H Rohrer, at the IBM Zurich Research labs in Switzerland. These two individuals shared the Nobel prize in physics in 1986 for their seminal work in scanning probe microscopy. A reading of the published literature reveals relevant prior art that resembles the implementation of SPMs in the early 1980s. As examples, work on surface profilers (using optical deflection techniques similar to those used in current scanning force microscopes) can be found in the published work of G Shmalz in 1929. In 1972, R Young, J Ward, and F Scire developed an instrument (called a topografiner) designed to measure the surface microtopography of a substrate. This latter work used a controllable metal–vacuum–metal separation to maintain a fixed tip–substrate distance, in some sense foreshadowing by some 10 years the tunnel gap approach developed independently by Binnig and Rohrer.

In what follows, the general principles underlying all SPMs are discussed first. Then the two widely used families of SPMs are discussed – the scanning tunneling microscope (STM) for studying the surface topography of "electronically conducting" substrates and the scanning force microscope (SFM), also known as the atomic force microscope (AFM), developed to investigate the surface topography of "electronically insulating" substrates. The proliferation and development of seminal SPM technology has produced a wide variety of dual-probe implementations of SPMs (often called SxMs; where x stands for some physical variable of interest), which have led to simultaneous measurements with high lateral and vertical resolution, not only of surface topography but also of other local properties of substrates.

Generic SPM Components

Virtually all SPMs share a number of common features that are schematically illustrated in **Figure 1**. This figure shows a sharp tip placed in proximity to a substrate. A vibrationless coarse approach mechanism (not illustrated) is required to initially position the tip close to the substrate without damage. After this approach, the substrate is scanned (rastered) in a controlled way under the tip by a computer program which oversees the operation of the instrument. A specially designed signal transducer forms a central part of the microscope as it measures (with high gain) a particular physical property related to the tip–substrate interaction, producing a measurable voltage that can be monitored by the computer.

The key feature of the microscope is a feedback loop, which acts to keep the output of the signal transducer constant during the scanning procedure. This is usually accomplished with a controller, based on the general principle of a proportional–integral–differential (PID) controller. By recording the z-motion of the substrate necessary to keep the signal transducer output constant, a point by point "image" of the substrate's surface can be generated. The image consists of $N \times N$ ordered values of (x, y, z) coordinates where N is usually of the order of a few hundred points. A one-dimensional ordered array of N points is usually referred to as one scan. An SPM image is therefore made up of N individual scans. Typically, the microscope can complete from $\sim$1 to $\sim$10 scans in one second; the speed is largely

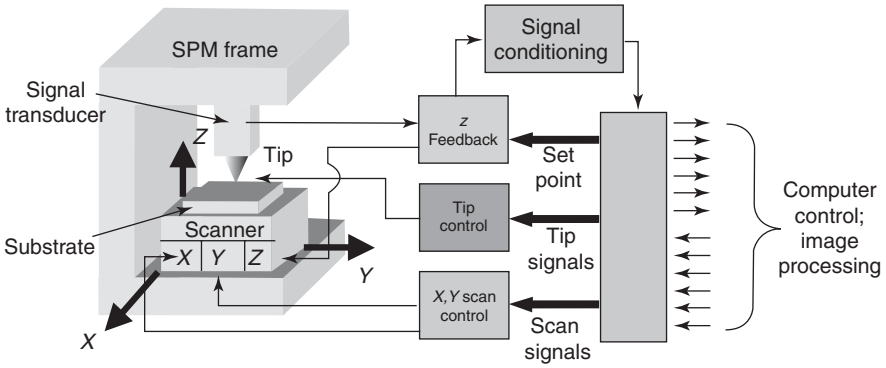


Figure 1 A schematic of a generic SPM in which a sharp tip interacts with a substrate. The design of the signal transducer allows different local properties of the substrate to be probed. The relative tip–substrate separation is controlled via a *Z*-feedback mechanism, producing an image of the substrate as it is rastered in a controlled way beneath the tip.

determined by the substrate roughness and the time constant of the feedback loop, which in turn, is related to the internal mechanical resonances of the SPM frame. Once the digital image is acquired, it can be displayed in a wide variety of two- or three-dimensional formats and further analyzed using a wide variety of sophisticated processing software.

Because of the close proximity between the tip and substrate (typically ~ 1 nm), the SPM head must be carefully isolated from uncontrolled wall and floor vibrations. Adequate vibration isolation is an absolute must and is often achieved by the use of a spring system designed to isolate the SPM from surrounding environmental vibrations. Immunity to vibration is also achieved by designing the SPM frame to be small, compact, and rigid thereby ensuring that the instrument's resonant vibrational frequencies lie in the kHz range, far removed from building vibrations which tend to occur in the 1–20 Hz regime. Very often, acoustic shielding surrounds the microscope to prevent the interaction of unwanted sound waves (audible noise) with the microscope. Care must be taken to minimize temperature gradients, since a small gradient between different pieces of the instrument can cause uncontrolled expansion or contraction of ~ 50–100 nm. If the substrate reacts spontaneously and uncontrollably with ambient environmental conditions, then further precautions must be taken to slow down or eliminate these uncontrolled surface modifications before a faithful image of the substrate's surface can be obtained. Very often, this requirement dictates either an ultrahigh vacuum environment or the submersion of the substrate under a protective liquid.

The controlled and vibrationless rastering of the substrate under the tip is critical to the operation of any SPM. This is often achieved through the use of piezoelectric ceramics which bend, contract, and expand in an appropriate manner by applying voltages

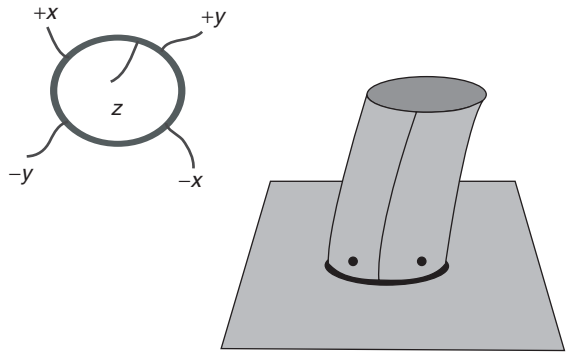


Figure 2 A schematic diagram of a segmented piezotube often used in SPMs to perform a controlled rastering in the *x*, *y*, and *z* directions.

in the 100–500 V range. Piezoelectric ceramics, that are fashioned in the shape of tubes, seem to be the current choice because of their high performance-to-cost ratio. Piezotubes (see **Figure 2**) are hollow piezoceramic cylinders (typically having a length $L \cong 2-4$ cm, an outside diameter $D \cong 0.5$ cm, and a wall thickness $w \cong 0.1$ cm), electroplated with thin metal films to form continuous inner and outer electrodes spanning the entire length of the tube. The inner electrode is continuous while the outer electrode is often segmented into four orthogonal quadrants or sectors to enable motion in an *xy* plane.

When a voltage is applied between the inner electrode and all four outer electrodes, the piezotube expands or contracts, depending on the polarity of the applied voltage. A bending motion closely resembling a pure lateral translation is achieved by applying a voltage between the inner electrode and one quadrant of the outer electrode. In order to achieve maximum translations, the applied voltage is often inverted and simultaneously applied to the opposing quadrant of the piezotube. In this way, orthogonal translations in an *x*–*y* plane can be achieved

by utilizing all four segmented sectors along the outer wall of the piezotube.

The displacements achieved by this action can be calculated from a knowledge of the d_{31} piezoelectric constant of the piezoceramic material comprising the piezotubes. Typical values of d_{31} (at room temperature) lie between -100 and $-300\times10^{-12}\,\mathrm{mV}^{-1}$. For an applied voltage V, the elongation (or contraction) of the piezotube along its axis is given by $\Delta z = L(V/w)d_{31}$. The displacement of the scanner in the x–y plane can be estimated from $\Delta x = \Delta y = (2\sqrt{2}/\pi D)\,(V/w)L^2 d_{31}$. In practice, intrinsic nonlinearities, hysteresis, aging, and creep all conspire to limit the performance of piezoceramics, which in practice often require constant calibration against established standards if high-fidelity metrology is a requirement.

The Scanning Tunneling Microscope

The STM was historically the first scanning probe microscope, and was introduced in 1982 by G Binnig and H Rohrer with the demonstration that a controllable vacuum tunneling gap could be achieved between a sharp metallic tip and a conducting substrate. The vertical resolution of the microscope is a few picometers while the lateral resolution is $\sim$0.1 nm. STM images typically span an area ranging from a few nanometers to few hundred nanometers.

To understand tunneling through a vacuum gap, it is useful to consider what happens when an electron wave with incident energy E encounters a barrier with a characteristic width d and a characteristic height V_o (**Figure 3a**). Such a barrier is present in the region of space formed by the physical gap between the tip and substrate, in which case $V_o = \varphi + \mu_t$ where φ is the work function of the tip and μ_t is the characteristic Fermi energy of the tip. The presence of this gap prevents the transit of a classical electron, but within the context of quantum mechanics, the electron has a finite probability of penetrating the barrier.

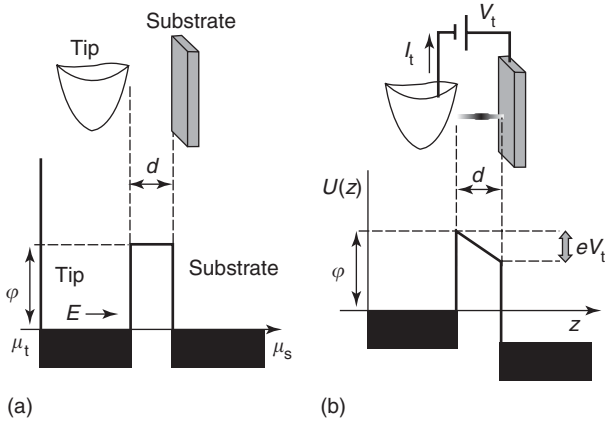


Figure 3 (a) A schematic of a potential barrier of width d between a metallic tip and metallic substrate. In equilibrium, the electrochemical potential of the tip (μ_t) and that of the substrate (μ_s) are aligned. The height of the potential barrier is $V_o = \varphi + \mu_t$, (b) The situation that develops when a bias (V_t) is applied between tip and substrate.

The basic physics required to understand how an STM operates, begins by considering electrons incident upon a barrier at an energy $E < V_o$. Such electrons can quantum mechanically tunnel through the barrier with a transmission probability T that can be obtained from a time-independent solution to the Schrödinger equation. For a square barrier, the transmission probability is given by

$$T = \frac{4E(V_o - E)}{4E(V_o - E) + V_o^2 \sinh^2(\kappa d)} \qquad [1]$$

valid when $E < V_o$, where

$$\kappa \equiv \frac{2\pi}{h}\sqrt{2m(V_o - E)}$$

m is the electron mass, and h is Planck's constant. When $\kappa d \gg 1$, an approximation appropriate for STM experiments, eqn [1] reduces to the well-known result that

$$T \approx \frac{16E(V_o - E)}{V_o^2} e^{-2\kappa d} \qquad [2]$$

If $E \cong \mu_t$, as is the case for low applied bias, then $V_o - E \cong \varphi$. Since φ is $\sim$5 eV (typical values of the work function in metals and semiconductors), the coefficient 2κ in eqn [2] is $\sim 23\,\mathrm{nm}^{-1}$.

To pass an electrical current I_t between the tip and substrate, a bias voltage V_t must be applied between the tip and the substrate (**Figure 3b**). This bias voltage distorts the shape of the square barrier, which is also rounded and lowered in height by many-body electron correlation effects not illustrated here. The electric current between tip and substrate is proportional to the transmission probability displayed in eqn [2]. After integrating over the appropriate range of energies, the tunnel current I_t is given by an expression of the general form

$$I_t \approx f(V_t, \varphi) e^{-2\kappa d} \qquad [3]$$

where $f(V_t, \varphi)$ is a function that depends on the applied voltage and the exact details of the barrier under consideration. For applied voltage differences of 1 V, typical tunnel currents encountered in STM experiments lie between 0.01 and 1 nA.

The strong exponential dependence of I_t with distance d is the important conclusion drawn from this calculation. Rough estimates using eqn [2] indicate that a change in the barrier width d by 0.1 nm causes a change in I_t by roughly a factor of 10. This large amplification implies that small tip motions can be easily detected, measured, and hence controlled.

The exquisite sensitivity of I_t to tip–substrate separation can be used as the signal transducer referred to in **Figure 1** to control the vertical tip position above a substrate. Two modes of imaging can be contemplated: (1) constant-height imaging in which the tip is moved at a fixed height above the substrate and where variations in the tunnel current due to height variations are recorded and (2) constant-current imaging in which the tip position is continually adjusted by the z-feedback loop to produce a constant tunnel current. These two modes of imaging in STM are illustrated in **Figures 4a** and **4b**, respectively.

The constant-height mode (**Figure 4a**) is of limited use, since it is only appropriate for substrates that are essentially flat at the atomic length scale. Furthermore, since variations in the tunnel current measured in the constant-height mode depend exponentially on the distance, they cannot be directly interpreted as height profiles. The three-dimensional imaging process in STM is most easily understood by considering the constant-current imaging mode (**Figure 4b**). The STM image is formed by recording the relative motion of the substrate to maintain a constant tunnel current as the tip is swept across a preselected area of the substrate. To achieve this, a high-gain current amplifier (typical gain is $\sim 10^8$ to $\sim 10^9\,\mathrm{VA}^{-1}$) is required.

Sharp tips are necessary to produce images with high lateral resolution. Common ways of producing STM tips from metal wires (such as W or Pt) with diameters of $\sim 1 \times 10^{-4}$ m rely on electrochemical etching or physical cutting. The reliable formation of sharp tips may seem like a daunting venture, but ultimately every tip formed must end with one or possibly a few atoms which ever so slightly protrude from the apex, forming a small mini-tip at the tip's apex. The presence of such mini-tips, along with the strong exponential drop-off of current with distance, provides a reasonable way to understand why the total tunnel current between tip and substrate may be dominated by an atomically small protrusion from an otherwise large tip.

More complete theories of STM clearly have shown that the tunnel current can be related to the wave function overlap between electron states in the tip with electron states in the substrate. This implies that the images obtained from an STM not only contain surface topographic information, but also information about the variation of the local density of electronic states. This complication provides a caveat against the direct translation of relative tip–substrate separation into surface topographic features. With this caution in mind, a few representative STM images from a variety of different surfaces are given in **Figure 5**. STM images are notable for the amazing detail they reveal about the atomic periodicity and surface morphology of clean, electronically conducting substrates.

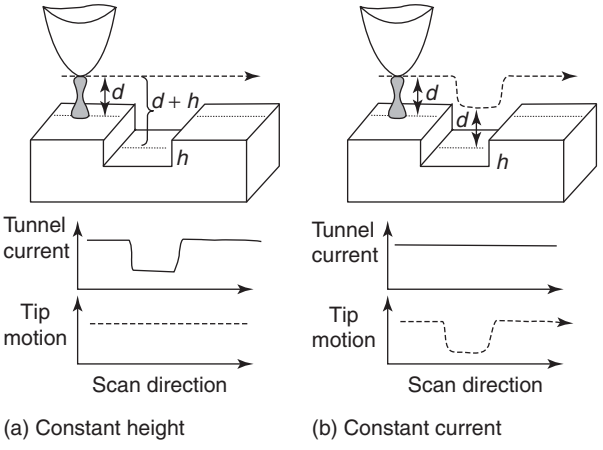
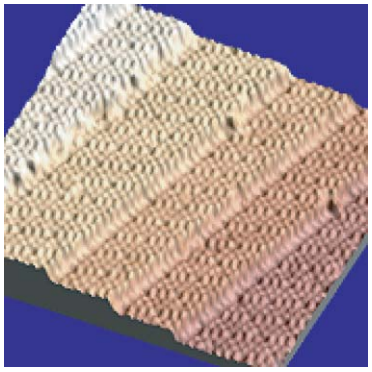


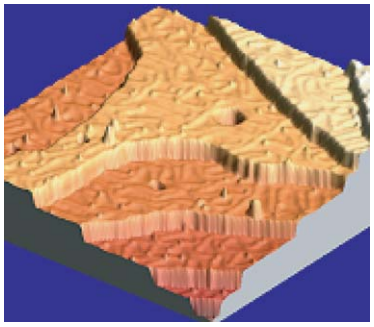
Figure 4 A schematic illustrating two modes of imaging employed in STMs. (a) The tip–substrate separation is held constant and the variations in tunnel current are measured; (b) the tunnel current is held constant by a feedback loop and the relative tip–substrate separation is varied to maintain a constant tunnel current.

The Scanning Force Microscope

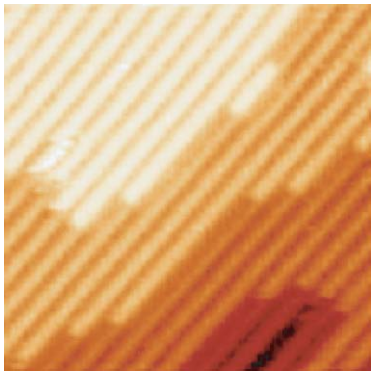
While STMs provide a quantitative map of surface topography with atomic resolution, they suffer from a fundamental limitation that the substrate studied must be sufficiently conducting to support a tunnel current. In order to overcome this difficulty, an AFM was first demonstrated in 1986 by Binnig, Quate, and Gerber. The operation of this microscope relied on the surface forces acting on a sharp tip in close proximity to a nonconducting surface. For sufficiently small tip–substrate separations, these interaction forces can range from tens of pN (10^{-12} N) to tens of μN (10^{-6} N), with typical values of a few nN (10^{-9} N). An understanding of these interaction forces is central to understanding how an AFM (also called an SFM) functions. Most importantly, these forces are not predicated on the fact that either the tip or substrate is electrically conducting. Because of the long-range nature of the interaction forces, the vertical resolution of an SFM is typically a few



(a)



(b)



(c)

Figure 5 Typical STM images. Different color schemes are employed in each image to better render the *z*-height information. (a) An image of step edges in a Si(1 1 1) substrate showing atomic periodicity. The atoms on the surface undergo a 7×7 reconstruction, creating a new unit cell different from that observed in the bulk. The field of view is 31×31 nm. (b) An image of a Au(1 1 1) surface showing atomically flat plateaus terminated by abrupt steps that are one atom in height. The small corrugation ($\sim$0.01 nm in height) observed on each flat terrace is a manifestation of the $22 \times \sqrt{3}$ surface reconstruction on this facet of Au. The field of view is 100×100 nm. (c) An atomically resolved image of Au(1 1 0) surface viewed from the top in false color format (bright is high, dark is deep). Rows of individual Au atoms forming a (2×2) reconstruction on different terraces of the substrate are clearly visible. The field of view is 9×9 nm. (Courtesy Gómez-Rodríguez JM and Mendez J, Departmento Física de la Materia Condensada, Universidad Autonoma de Madrid.)

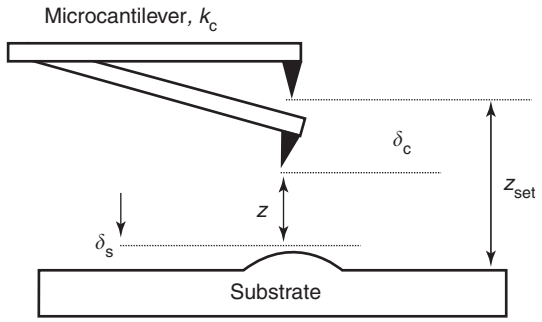


Figure 6 A schematic illustrating the sequence of events when a tip on a microcantilever is brought into close proximity to a substrate. Initially, the tip is located a distance z_{set} from the substrate. Attractive interaction forces between the tip and substrate bend the tip toward the substrate until a deflection δ_c of the cantilever brings the system into equilibrium. If the interaction forces are sufficiently strong, the substrate may also experience a distortion δ_s, which may be appreciable if the substrate is soft. The final tip–substrate separation is indicated by the parameter *z*.

picometers (comparable to an STM) while the lateral resolution is $\sim$10–20 nm, somewhat larger than for STMs. SFM images typically span an area ranging from $\sim$100 nm to around tens of micrometers.

In practice, the operation of an SFM relies on a sharp tip which is usually supported on the end of a microcantilever whose minute deflections can be carefully monitored. As shown in **Figure 6**, a microcantilever with spring constant k_c, when positioned at a distance z_{set} from a substrate, will deflect toward the substrate by an amount δ_c due to interaction forces that exist between the tip and substrate. In addition, it is possible that the surface will distort by an amount δ_s due to the action of the same forces. In general, it is difficult to determine the exact tip–substrate separation z due to a lack of knowledge about the distortion δ_s as well as the inability to accurately determine the initial tip–substrate separation z_{set}. For sufficiently small deflections, the cantilever motion can be well approximated in terms of Hook's law, which predicts a restoring force F given by $F = -k_c\delta_c$.

Table 1 provides some dimensions and relevant properties of typical microcantilevers that are commercially available. An uncertainty in cantilever thickness causes a considerable spread in the resulting spring constants. Sharp tips, with effective radius R (typically, R is between 5 and 30 nm), are routinely formed onto these cantilevers using lithographic techniques developed by the semiconductor industry. The widespread availability of microcantilevers means that interaction forces $\sim$1 nN between the tip and the substrate can be monitored and that cantilever deflections $\sim$1 nm or less can be readily detected.

Table 1 A few representative silicon cantilevers commercially available with their characteristics

Length (μm)	*Width* (μm)	*Thickness* (μm)	k_c (N m^{-1}) *(min, typical, max)*	f_o (kHz) *(min, typical, max)*
125±5	35±3	4.0±0.5	20, 40, 75	265, 325, 400
230±5	40±3	7.0±0.5	25, 40, 60	150, 170, 190
90±5	35±3	2.0±0.3	6.5, 14, 28	240, 315, 405
125±5	35±3	2.0±0.5	1.8, 5.0, 12.5	110, 160, 220
90±5	35±3	1.0±0.3	0.45, 1.75, 5.0	95, 155, 230
300±5	35±3	1.0±0.3	0.01, 0.05, 0.1	9.5, 14, 19

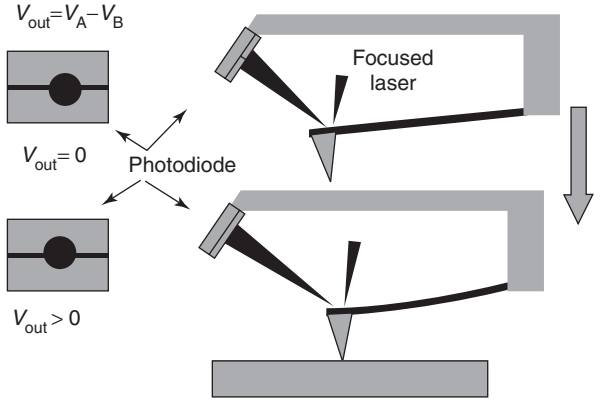


Figure 7 A common method employed to measure the deflection of a cantilever is a beam bounce technique in which a diode laser beam is reflected from a microcantilever onto a segmented photodiode. By monitoring the voltage from each segment of the photodiode, the relative motion of the reflected laser spot can be monitored and information about sub-nanometer motion of the cantilever can be inferred.

To measure cantilever motion while scanning, a high-gain transducer of cantilever deflection (the signal transducer referred to in **Figure 1**) plus a feedback mechanism is required. A variety of techniques – capacitance, optical interferometry, piezoelectric microcantilevers, and optical beam deflection – have been successfully implemented to accurately detect cantilever deflection. Each technique seems to have its own advantages. Currently, the technique most often implemented is an optical deflection scheme shown schematically in **Figure 7**.

Using this approach, a focused laser beam is deflected from a microcantilever and the reflected light is directed onto a segmented photodiode. Fine positioning of the reflected spot allows for a null condition characterized when the voltages from the appropriate photodiode segments are made to sum to zero by an external operational amplifier (not shown). A small cantilever deflection disrupts this null condition, giving rise to a voltage proportional to beam deflection. The origin of the high amplification for this particular system follows from simple geometrical considerations. For a cantilever displacement Δz, the reflected laser spot moves a distance $\Delta s = \Delta z(d/l)$, where d is the distance of the cantilever from the photodiode and l is the cantilever's length. Typically, the ratio of d/l for a microcantilever can easily be a factor of 100–500.

When discussing the nature of the interaction force between tip and substrate, it is often convenient to approximate the tip as a sphere with radius R. This sphere then interacts with the substrate via a number of possible forces, which can cause the cantilever to deflect as shown in **Figure 6**.

The exact details of the relevant interaction forces, as well as their variations on z, depend to a large extent on the composition of the tip and substrate. For the ideal case of a clean, electrically neutral tip positioned above an electrically neutral, clean substrate in ultrahigh vacuum, the interaction forces might be well-approximated by a superposition of a short-range, hard-wall repulsion (effective when the tip–substrate separation is less than ~ 0.3 nm) plus a longer range surface interaction due to the van der Waals (vdW) force acting between dipoles induced on the individual atoms comprising the tip and substrate. The z-dependence of this vdW force is related to the detailed shape of the substrate and tip. If the substrate/tip are studied under ambient conditions, hydration forces due to adsorbed water, as well as long-range electrostatic forces due to uncontrollable charging of the tip or substrate, may well dominate.

Without a detailed knowledge of the system under study, it is difficult to accurately specify a force versus distance relationship. In general, such a curve would have the approximate shape shown in **Figure 8**. This figure qualitatively illustrates (1) the attractive regime ($F<0$) in which the interaction forces cause the tip (and microcantilever) to bend toward the substrate and (2) the repulsive regime ($F>0$) that causes the microcantilever to bend away from the substrate when the tip comes into contact with it.

A qualitative appreciation of the important features of this force curve is critical to understanding how an SFM obtains an image of a substrate. By rastering the substrate beneath the tip, all the while maintaining a constant force between the tip and the

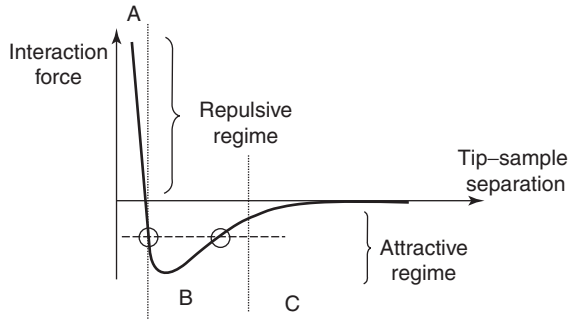


Figure 8 A schematic illustrating how the interaction force between tip and substrate varies as a function of separation. Three regions (A, B, C) are indicated. Different modes of imaging are achieved when the tip is positioned in each region.

substrate, an image closely resembling the surface topography of the substrate's surface can be obtained. There are a variety of methods that have been developed to achieve this task. The exact method employed depends on the distance between the tip and substrate.

If the tip is in region A in **Figure 8**, then imaging is performed in a "contact mode"; the tip exerts a force directly on the sample as it is scanned across it. In contact mode imaging, the direct up and down motion of the cantilever is measured while scanning. This motion can be used to produce a three-dimensional image of the substrate in much the same way as a conventional profilometer, except that now the applied force lies in the nN range and the radius of the stylus is in the 5–30 nm range. This mode of operation can be damaging, especially for soft substrates and stiff microcantilevers since significant lateral forces develop during the scanning process.

If the tip is in region C of **Figure 8**, the interaction forces are sufficiently weak so that very small deflections of the cantilever result. Since the substrate–tip separation is large, imaging in this region is often referred to as the "noncontact mode." Under these circumstances, indirect detection schemes are usually employed. As an example in noncontact mode imaging, the tip is often driven sinusoidally at a frequency near its mechanical resonance. Small position-dependent shifts in the resonance frequency occur when the substrate is rastered beneath the tip. These frequency shifts can then be used as a sensitive measure of tip–substrate separation, thereby providing the transducer signal for the feedback controller. Because of the noncontact feature of this mode, it is preferred when studying soft substrates.

If the tip is placed in region B of **Figure 8**, the interaction forces become comparable to the restoring force of the microcantilever, implying that static tip displacements, although measurable, cannot be reliably measured because of resulting instabilities. In simple terms, the instabilities arise because of the double-valued nature of the interaction force as illustrated by the horizontal dotted line in **Figure 8** which indicates that for the same value of the interaction force, there are two possible tip–sample separations. These instabilities are often referred to as jump-to-contact because the tip spontaneously snaps into contact with the substrate no matter how carefully the procedure is employed. To scan in region B, the tip must assume a time-dependent behavior which is carefully controlled by the SFM computer. During the tip's motion, it is possible that the tip might periodically come into contact with the substrate, giving rise to what is known as "intermittent-contact" or "tapping mode" imaging.

The boundaries between the different regions in **Figure 8** are not necessarily well defined, so a precise distinction between the different imaging regimes is difficult to provide. When operating an SFM in either region B or C, the SFM is often referred to as a dynamic force microscope (DFM) and the imaging process is often referred to as dynamic mode imaging.

To better appreciate the information contained in an SFM image obtained in regions B and C of **Figure 8**, it is useful to discuss in more detail the appropriate imaging modes. If the tip hovers at some distance above the substrate, then a sinusoidal modulation of the cantilever's position is usually required to monitor the interaction forces. This is often accomplished by mounting the cantilever assembly directly onto a small piezoelectric slab, which is driven at a preset angular frequency ω. During the oscillatory motion of the cantilever, the tip feels two forces: a restoring force due to the cantilever and an interaction force due to the forces acting between the tip and substrate.

Under these circumstances, the tip motion can be visualized by considering the two spring systems shown in **Figure 9**. One spring, with spring constant k_c, accounts for the restoring force of the cantilever, while the other spring, with effective spring constant k_{inter}, accounts for the interaction forces. Effectively, the tip is acted on by a single spring with effective spring constant $k_{eff} = k_c - k_{inter}$. If the interaction force curve is accurately known (rarely the case), k_{inter} could be estimated according to $k_{inter} = dF_{inter}/dz$, where it is understood that the derivative must be evaluated at the equilibrium separation between the cantilever and substrate.

When the tip is far from the sample, $k_{inter} \cong 0$ and the equation of motion of a tip with mass m can be analyzed using the equations for a damped oscillator (with damping coefficient β), driven by a sinusoidal

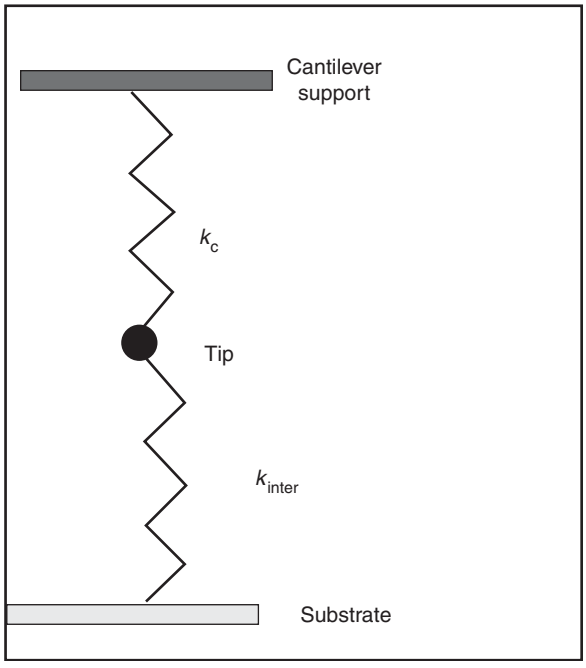


Figure 9 A simple way to understand the different forces acting on a tip when it is positioned in close proximity to a substrate. The tip is acted upon by two springs, one due to the restoring force of the cantilever, the other due to the interaction force between tip and substrate.

force $F_o \cos(\omega t)$:

$$m\ddot{z} + \beta\dot{z} + k_c z = F_o \cos(\omega t) \qquad [4]$$

The steady-state tip motion $z(t) = \mathrm{Re}(z_o e^{-i\alpha} e^{i\omega t})$ is specified by

$$z_o = \frac{F_o}{\sqrt{m^2(\omega_o^2 - \omega^2)^2 + \beta^2\omega^2}} \qquad \tan\alpha = \frac{\omega\beta}{m(\omega_o^2 - \omega^2)} \qquad [5]$$

where the resonance frequency $\omega_o = 2\pi f_o \equiv \sqrt{k_c/m}$. Clearly, the amplitude of the oscillating cantilever is maximu'm when the driving frequency equals the resonance frequency (i.e., when $\omega = \omega_o$).

As the tip moves closer to the substrate, the interaction force F_{inter} shown schematically in **Figure 8** comes into play and the equation of motion becomes

$$m\ddot{z} + \beta\dot{z} + k_c z = F_o \cos(\omega t) + F_{inter} \qquad [6]$$

The details of the tip motion now depend on F_{inter}.

A general analysis of this problem usually starts by expanding F_{inter} in a Taylor series about the equilibrium tip–substrate separation. This analysis shows that the resonance frequency (which can be accurately measured) becomes a function of the tip–substrate separation. The effective resonant frequency ω_e is given approximately by

$$\omega_e \cong \omega_o \left(1 - \left.\frac{dF_{inter}}{dz}\right|_{z=z_{eq}}\right)^{1/2} \qquad [7]$$

As a consequence, the amplitude of the cantilever's oscillation, driven at a frequency ω slightly off-resonance, will vary with position as the substrate is rastered beneath the oscillating microcantilever. The amplitude change resulting from variations in ω_e can be measured using a conventional phase-sensitive detection of the signal from the microcantilever/photodiode assembly. The substrate–tip separation can be continuously adjusted by a feedback loop to maintain either a constant resonance frequency or a constant amplitude of oscillation. By measuring the required motion of the substrate in the z-direction required to achieve this condition, it is possible to render an "image" of the surface of the substrate. Such an SFM image reflects relative changes in the tip–substrate separation required to maintain a constant resonant frequency, or equivalently a constant amplitude of the tip's oscillation.

As the equilibrium separation between the cantilever and substrate decreases and the tip passes from region C to region B, the intermittent mode of operation comes into play. The situation becomes considerably more complicated. The interaction forces are now comparable to the restoring forces of the cantilever. If $k_c = k_{inter}$ at some z_{set}, instabilities will result. If the amplitude of the tip oscillation drives the tip from region B to region A, the tip will intermittently "tap" the surface. This could produce a substrate damage, especially if the cantilever has a high spring constant. Also, nonlinearities and instabilities in the tip motion can produce a chaotic rather than a periodic motion of the cantilever. Considerable operator skill and insight is often required to produce artifact-free images under these conditions.

Another way of implementing intermittent contact mode imaging is to use an alternative approach often referred to as the "jumping mode." In this procedure, the cantilever does not undergo sinusoidal motion but instead follows a motion determined by the software controlling the SFM. In practice, the software is programmed to retract the tip, position it at the beginning of a selected scan range, and then drive the substrate toward the tip under feedback control until the cantilever bending reaches a preset loading force. At this point, the z-displacement of the substrate required to meet this condition is measured and the tip is withdrawn, moved to a nearby adjacent location above the substrate where the process is again repeated. After completing a scan, the relative

z-motion of the substrate at each point in the scan is plotted for further analysis. The advantage of this technique is that the force applied to the substrate can be carefully monitored during the imaging process. Furthermore, the lateral force imparted to the substrate while scanning is eliminated. The disadvantage is that the scan proceeds at a somewhat slower rate than when the cantilever is sinusoidally driven. **Figure 10** provides a schematic diagram of these different imaging modes.

Table 2 summarizes the above discussion by listing a few very general guidelines for cantilever selection during each mode of SFM operation.

By way of summary, in SFM, no matter which mode of imaging is employed, the motion of the z-piezo is used to form an image to keep a required signal constant. In contact mode imaging, the static deflection of the cantilever is used. In noncontact mode imaging, the amplitude of the cantilever oscillation is employed. The SFM is characterized by a lateral resolution of $\sim$10–20 nm, somewhat higher than that for STM since the tip radius R plays a central role. Atomically resolved images in SFM have been demonstrated, but usually this requires an SFM operating under ultrahigh vacuum conditions. The vertical resolution (typically better than 0.1 nm) rivals that of STM, while a force sensitivity of the order of 1 pN can be achieved.

A few typical SFM images taken in noncontact mode are given in **Figure 11**.

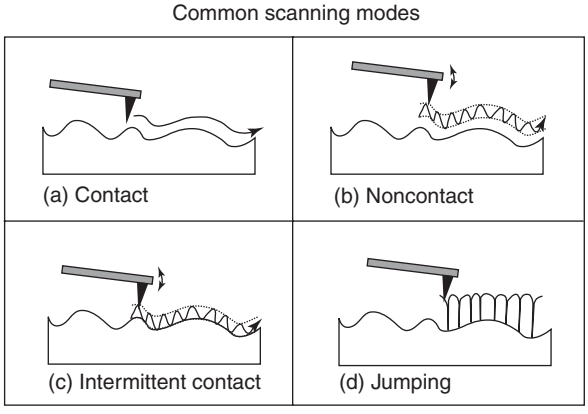


Figure 10 A schematic illustrating the different scanning modes commonly employed in SFM. (a) The contact mode of imaging where the tip is in constant contact with the substrate; (b) the noncontact mode of imaging, where the tip oscillates sinusoidally while maintaining a fixed difference between tip and substrate; (c) the intermittent contact mode where the tip "taps" the substrate during the scanning process (Note that the frequency of tip oscillation in (b) and (c) is much greater than the scanning frequency of the microscope.); (d) the jumping mode where the tip is moved into contact with the substrate, then lifted and moved to another location before contact with the substrate is re-established.

The SxM Family of Microscopes

The basic techniques described above have been extended in a number of very clever ways, producing a large family of SPMs, often referred to as SxMs, each designed to detect the local variation in some quantity of interest. This extension of SPM is often referred to as dual-probe microscopy because the tip not only measures topography but also some other physical parameter of interest with high lateral resolution. A few examples include an electrostatic force microscope (EFM or scanning Kelvin probe), a magnetic force microscope (MFM), a photon scanning tunneling microscope (PSTM), a scanning electrochemical microscope (SECM), a scanning near-field optical microscope (SNOM), a scanning capacitance microscope (SCM), scanning tunneling spectroscopy (STS), and a frictional force microscope (FFM).

Summary

The rapid evolution of SPMs since their first demonstration in the early 1980s has truly been remarkable. Largely because they are versatile and relatively inexpensive, SPMs have ushered in a worldwide interest in nanotechnology. With precise engineering, SPMs are capable of very high resolution (subatomic scale) metrology. Also, it is now clear that SPM tips can be used as tools capable of nanometer manipulation and fabrication. Active research in nanolithography is underway to controllably use the

Table 2 A qualitative comparison between the three commonly used modes in scanning force microscopy

Scanning mode	k_c (N m^{-1})	f_o, *Resonance frequency* (kHz)	*Approx. tip–substrate separation* (nm)	*Comments*
Contact	<1	–	0	Tip wear, contact mechanics
Intermittent contact	10–100	100–300	$\sim$3	Stable
	1–5	10–70	$\sim$3	Unstable
Noncontact	10–100	100–300	>5	Stable
	1–5	10–70	>5	Stable

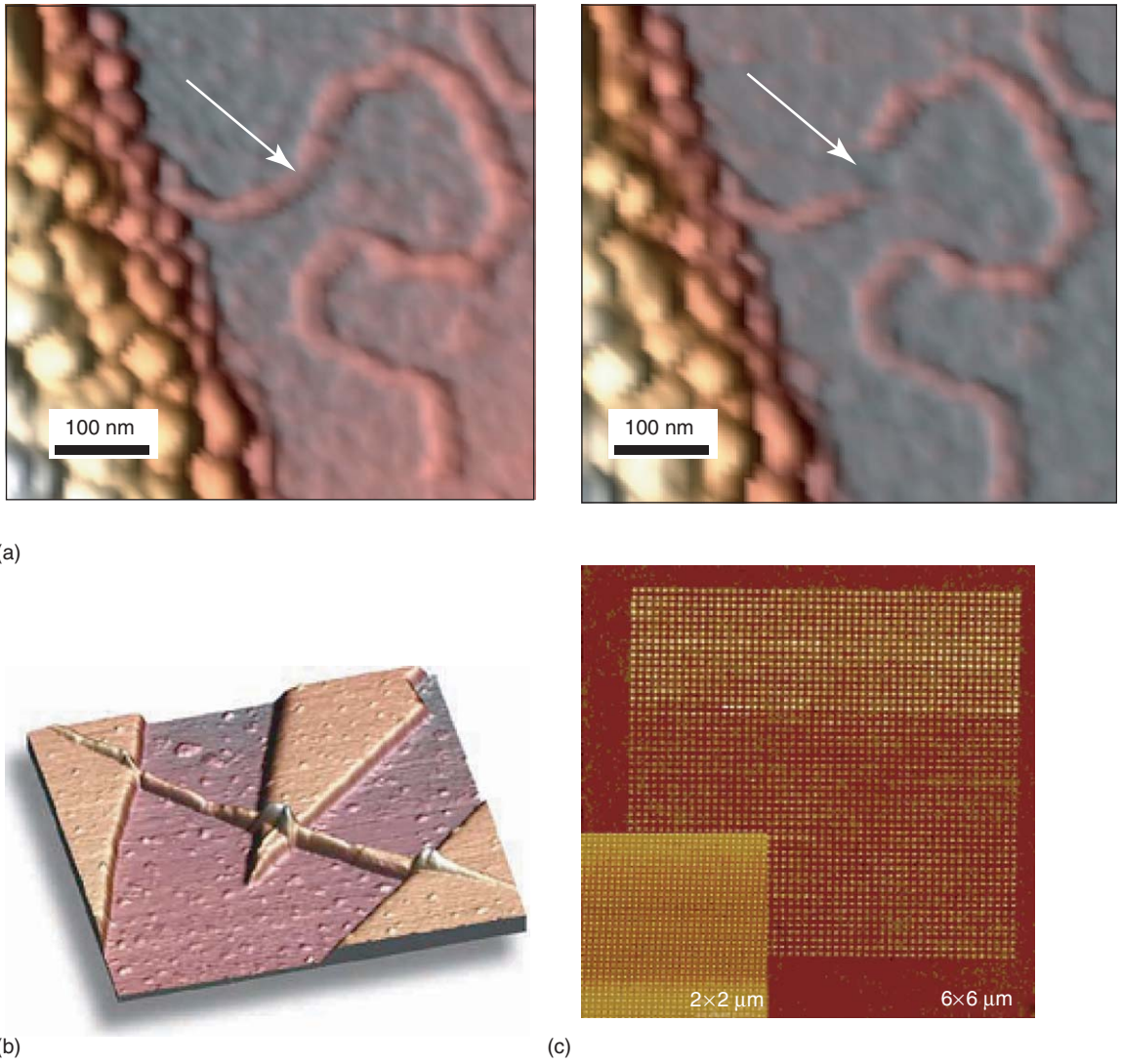


Figure 11 Representative examples of SFM images. Different color schemes are employed in each image to better render the *z*-height information. (a) An image of λ-DNA deposited on a mica substrate. The left edge of this image shows a gold contact pad deposited on top of the DNA molecule. After manipulation, the DNA molecule has been cut by the tip, at the position shown by the arrows. The size of the image is indicated by the scale bar. (b) An image of three gold contact pads deposited onto a multiwalled carbon nanotube and then imaged with an SFM in the noncontact mode. The field of view is $15 \times 15\,\mu m$. (c) An example of nanolithography performed by a scanning force tip viewed in false color (bright is high, dark is low). A periodic array of silicon oxide nanostructures has been created by applying a voltage between the tip and a Si substrate under ambient conditions. Local electrochemistry between the tip and substrate produces the protrusions which are then imaged using the same tip that created them. (Courtesy DNA image – J Gomez-Herrero, Departmento Física de la Materia Condensada, Universidad Autonoma de Madrid; Carbon nanotube – Reifenberger R, Department of Physics, Purdue University; Nanolithography – C Martin and F Perez-Murano, Instituto de Microelectronica de Barcelona; Centro Nacional de Microelectronica.)

SFM tip to locally modify a substrate in a very precise way at the nanometer length scale. SPM operation has been extended to scanning under liquids, allowing a window into the biological world. Linear parallel arrays of cantilevers have been designed and fabricated to work in a massively parallel fashion, and efforts to independently control individual cantilevers in the array have also been reported. Current indicators are that technology underlying these proximal probe microscopes will continue to improve, and the SPM class of instruments will continue to become ever more commonplace as a tool of choice to probe the properties of nanoscale objects.

See also: Biomolecules, Scanning Probe Microscopy of; Confocal Optical Microscopy; Fermi Surface Measurements; Low-Energy Electron Microscopy; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Optical Microscopy; Photoelectron Spectromicroscopy; Scanning Near-Field Optical Microscopy; Superconductivity: Tunneling; Surfaces and Interfaces, Electronic Structure of; Surfaces, Optical Properties of;

Transmission Electron Microscopy; Treated Surfaces, Optical Properties of; van der Waals Bonding and Inert Gases.

PACS: 07.79. – v; 68.37.Ef

Further Reading

Binnig G, Quate CF, and Gerber C (1986) Atomic force microscope. *Physics Review Letters* 56: 930–933.

Binnig G and Rohrer H (1991) In touch with atoms. *Review of Modern Physics* 71: S324–S330.

Binnig G, Rohrer H, Gerber Ch, and Weibel E (1982) Tunnelling through a controllable vacuum gap. *Applied Physics Letters* 40: 178–180.

Binnig G, Rohrer H, Gerber C, and Weibel E (1983) 7×7 Reconstruction of Si(1 1 1) resolved in real space. *Physics Review Letters* 50: 120–123.

Cappella B and Dietler G (1999) Force distance curves by atomic force microscopy. *Surface Science Report* 34: 1–104.

Giessibl FJ (2003) Advances in atomic force microscopy. *Review of Modern Physics* 75: 949–983.

Gracia R and Perez R (2002) Dynamic atomic force microscopy methods. *Surface Science Report* 47: 197–301.

Guntherodt HJ and Wiesendanger R (eds.) (1994) *Scanning Tunneling Microscopy I: General Principles & Applications to Clean & Adsorbate-Covered Surfaces*. New York: Springer.

Guntherodt HJ and Wiesendanger R (eds.) (1997) *Scanning Tunneling Microscopy II: Theory of STM & Related Scanning Probe Methods*. New York: Springer.

Israelachvili J (1991) *Intermolecular and Surface Forces*, 2nd edn. London, New York: Academic Press.

Julian Chen C (1993) Introduction to scanning tunneling microscopy, *Oxford Series in Optical and Imaging Sciences*. New York: Oxford University Press.

Marti O and Amrein M (eds.) (1993) *STM and SFM in Biology*. New York: Academic Press.

Martin Y, Williams CC, and Wickramasinghem HK (1987) Atomic force microscope – force mapping and profiling on a sub 100 – Å scale. *Journal of Applied Physics* 61: 4723–4729.

Meyer G and Amer NM (1988) Novel optical approach to atomic force microscopy. *Applied Physics Letters* 53: 1045–1047.

Meyer E, Hug HJ, and Bennewitz R (2004) *Scanning Probe Microscopy*. Berlin: Springer.

Sarid D (1991) Scanning force microscopy with applications to electric, magnetic, and atomic forces, *Oxford Series in Optical and Imaging Sciences*. New York: Oxford University Press.

Scattering Techniques, Compton

P E Mijnarends, Northeastern University, Boston, MA, USA and Delft University of Technology, Delft, The Netherlands
A Bansil, Northeastern University, Boston, MA, USA

Introduction

"Compton scattering" refers to a collision between a photon and a charged particle, often an electron, in which the photon loses a substantial fraction of its energy. The use of words "photon" and "collision" suggests a corpuscular nature of light, that is, that light may be viewed as a stream of particles with well-defined energies and momenta. The first indications that light can lose a part of its energy when scattered in this manner appeared at the beginning of the twentieth century. Convincing evidence was provided by the pioneering experiments in the 1920s by Arthur Compton after whom the effect has been named.

This was not the first time the suggestion had been made that light may behave as a particle. The photoelectric effect, in which a material irradiated by light can be seen to emit electrons, was earlier explained by Einstein by postulating that light of frequency v consists of particles called light quanta (and later photons), each with energy hv, where h is Planck's constant. For a given material, the photoemitted electrons possess energies, which depend on the frequency (or equivalently the wavelength, $\lambda = c/v$, where c is the speed of light), but not on the intensity of the incident light. Above a certain wavelength no electron emission is possible, as is observed to be the case, because a photon does not carry enough energy to overcome the binding energy of the electron to the material. Compton scattering, which views the photon–electron scattering to behave rather like billiard balls colliding with other billiard balls, thus reinforces Einstein's explanation of the photoelectric effect.

An outline of this chapter is as follows. The following section gives an overview of the theory of Compton scattering. The equations for energy and momentum transfer and the cross section for the Compton scattering process are given. Various forms of Compton scattering are discussed: (1) Compton scattering from the charge distribution of electrons in a material; (2) magnetic Compton scattering; and (3) $(\gamma, e\gamma)$ scattering, in which additional information is obtained by measuring the properties of the recoil electron in addition to those of the scattered photon. The third section addresses issues of instrumentation. In keeping with current trends, the emphasis is on experiments using synchrotron radiation facilities around the world, which provide bright sources of X-rays. The forth section touches upon the effect of multiple scattering of a Compton scattered photon

within the sample, which complicates the interpretation of data. The fifth section turns to methods of analyzing Compton scattering data. In those cases where an adequate data set is available, one can in principle obtain the three-dimensional momentum density of the many-body electronic ground state of the material. Much insight has been gained into the electronic structure and bonding properties in wide classes of materials via comparisons of measured Compton scattering data with corresponding *ab initio* computations. The last section concludes with a presentation of a few illustrative examples of such comparisons.

Theory

Consider first a collision between a photon and a stationary electron in free space. The incident photon has momentum $\boldsymbol{k}_0$ and energy $E_0 = |\boldsymbol{k}_0|c$. As a result of the collision, the electron recoils with momentum $\boldsymbol{p}_1$ and relativistic energy $\varepsilon_1 = [(p_1c)^2 + m^2c^4]^{1/2}$, while the momentum and energy of the scattered photon are $\boldsymbol{k}_1$ and $E_1 = |\boldsymbol{k}_1|c$, respectively. Energy and momentum conservation laws then yield

$$\varepsilon_1 + E_1 = E_0 + mc^2 \qquad [1]$$

$$\boldsymbol{p}_1 + \boldsymbol{k}_1 = \boldsymbol{k}_0 \qquad [2]$$

Equation [2] shows that the momentum gained by the recoiling electron, $\boldsymbol{p}_1 = \boldsymbol{k}_0 - \boldsymbol{k}_1 \equiv \hbar\boldsymbol{q}$, is the same as the momentum lost by the photon, which, after division by $\hbar$, is called the scattering vector $\boldsymbol{q}$ (**Figure 1**). Eliminating the electron momentum $\boldsymbol{p}_1$ from eqns [1] and [2] gives

$$E_0 - E_1 = \frac{E_0E_1}{mc^2}(1 - \cos\phi) \qquad [3]$$

Recalling the relation $\lambda = hc/E$ for photons, this can be written as

$$\lambda_1 - \lambda_0 = \frac{h}{mc}(1 - \cos\phi) \qquad [4]$$

The quantity $h/mc \sim 0.00\,243$ nm, called the Compton wavelength of the electron, is the natural unit of length in quantum electrodynamics. The shift in wavelength is seen from eqn [4] to be independent of the energy of the incident photon. Therefore, this shift can be more easily detected by using high-energy radiation since the percentage change in wavelength is then larger. Equation [3] can also be rewritten as

$$E_1 = \frac{E_0}{1 + (E_0/mc^2)(1 - \cos\phi)} \qquad [5]$$

This equation shows that for X- or γ-rays of energy E_0 comparable to $mc^2 \sim 511$ keV, the energy loss at large scattering angles ϕ is a substantial fraction of the initial energy E_0.

It is important to note that the energy conservation (eqn [1]) neglects changes in the potential energy of the system, that is, the potential energy of the electron is assumed to be the same before and after the collision. In scattering from atoms, molecules, and solids, this approximation is valid only if the energy transferred in the scattering process is much larger than relevant binding energies involved. Moreover, for high energy transfers, the scattering process is a fast one in that the duration of the interaction is too short for the system to rearrange itself. This is called the "impulse approximation," and it is at the heart of much of the interpretation of the Compton scattering spectra from materials. In the deeply inelastic regime, where the "impulse approximation" is valid, Compton scattering provides a direct probe of the many-body ground state of the electronic system.

In the more realistic case of a moving electron with initial momentum $\boldsymbol{p}_0$ and energy $\varepsilon_0 = [(p_0c)^2 + m^2c^4]^{1/2}$, the energy and momentum conservation conditions are

$$\varepsilon_1 + E_1 = \varepsilon_0 + E_0 \qquad [6]$$

$$\boldsymbol{p}_1 + \boldsymbol{k}_1 = \boldsymbol{p}_0 + \boldsymbol{k}_0 \qquad [7]$$

The scattering vector (**Figure 2**) once again is $\hbar\boldsymbol{q} \equiv \boldsymbol{k}_0 - \boldsymbol{k}_1 = \boldsymbol{p}_1 - \boldsymbol{p}_0$. Here, ε_0 and ε_1 cannot be eliminated

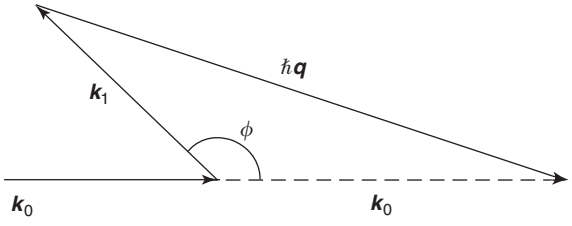


Figure 1 Momentum conservation diagram for Compton scattering of an incident photon of momentum $\boldsymbol{k}_0$ from a free electron at rest. $\boldsymbol{k}_1$ is the momentum of the scattered photon, q the scattering vector, and ϕ the scattering angle.

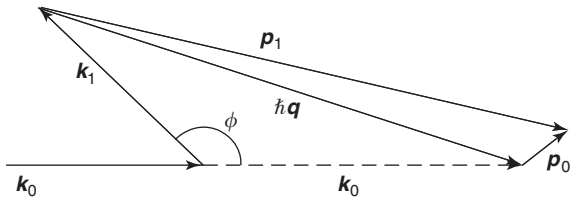


Figure 2 Momentum conservation diagram for Compton scattering of an incident photon of momentum $\boldsymbol{k}_0$ from an electron with momentum $\boldsymbol{p}_0$. Upon recoiling, the electron has a momentum $\boldsymbol{p}_1$. The initial momentum $\boldsymbol{p}_0$ is typically much smaller than $\boldsymbol{p}_1$.

simultaneously from eqns [6] and [7] to obtain an equation involving just the photon kinematics. However, elimination of $\boldsymbol{p}_1$ leads to the expression

$$(E_0 - E_1)(\varepsilon_0 + \varepsilon_1) = (\hbar^2 q^2 + 2\hbar \boldsymbol{p}_0 \cdot \boldsymbol{q})c^2 \qquad [8]$$

Substituting $\varepsilon_0 = \varepsilon_1 = mc^2$ in eqn [8], one obtains the useful nonrelativistic relation

$$E_0 - E_1 = \frac{\hbar^2 q^2}{2m} + \frac{\hbar \boldsymbol{p}_0 \cdot \boldsymbol{q}}{m} \qquad [9]$$

The term $\hbar^2 q^2/2m$ is called the Compton shift, while the term $\hbar \boldsymbol{p}_0 \cdot \boldsymbol{q}/m$ has the character of a Doppler shift in that this term is proportional to the velocity component of the electron along the direction of $\boldsymbol{q}$. It is customary to choose p_z as the component of the initial electron momentum along the scattering vector so that $p_z = \boldsymbol{p}_0 \cdot \boldsymbol{q}/q$. If $p_z = 0$, then eqn [9] becomes

$$E_0 - E_C = \frac{\hbar^2 q^2}{2m} \qquad [10]$$

where E_C is the energy corresponding to the top of the Compton profile.

In the relativistic case eqn [8] immediately yields

$$p_z = \frac{(E_0 - E_1)(\varepsilon_0 + \varepsilon_1) - \hbar^2 q^2 c^2}{2\hbar q c^2} \qquad [11]$$

The scattering vector squared is

$$\begin{aligned} \hbar^2 q^2 &= k_0^2 + k_1^2 - 2k_0 k_1 \cos\phi \\ &= (E_0^2 + E_1^2 - 2E_0 E_1 \cos\phi)/c^2 \end{aligned} \qquad [12]$$

Substituting [12] into [11], one obtains

$$p_z = \frac{(E_0 - E_1)\varepsilon_0 - E_0 E_1 (1 - \cos\phi)}{c\sqrt{E_0^2 + E_1^2 - 2E_0 E_1 \cos\phi}} \qquad [13]$$

The width of the Compton profile is very narrow compared with E_C so that $E_C \approx E_1$, and one can use the leading order nonrelativistic approximation $\varepsilon_0 = mc^2$ to obtain the following equation:

$$\frac{p_z}{mc} \approx \frac{E_0 - E_1 - (E_0 E_1/mc^2)(1 - \cos\phi)}{\sqrt{E_0^2 + E_1^2 - 2E_0 E_1 \cos\phi}} \qquad [14]$$

This expression allows one to convert the energy scale of the Compton profile to a momentum scale. Conventionally, a minus sign is added to the left-hand side of [13] and [14], so that positive values of p_z correspond to the high-energy side of the profile. Notice also that the energy and momentum of the electron do not occur on the right-hand side of eqn [14]. In a typical Compton scattering experiment the kinematics of the recoil electron is not measured, so that the data are in effect integrated over the p_x- and p_y-components and p_z is the only electron momentum component measured.

The double differential scattering cross section for spin-dependent Compton scattering from a material containing a collection of moving electrons is

$$\begin{aligned} \frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\ \mathrm{d}E_1} = (e^2/mc^2)^2 (m/2\hbar q)(E_1/E_0) \\ \times \{f J(p_z) + (\cos\phi - 1) P_c \hat{\boldsymbol{\sigma}} \\ \cdot [(\boldsymbol{k}_0 \cos\phi + \boldsymbol{k}_1)/mc] J_{mag}(p_z)\} \end{aligned} \qquad [15]$$

where

$$\begin{aligned} f = 1 + \cos^2\phi + [(E_0 - E_1)/mc^2](1 - \cos\phi) \\ + P_l \sin^2\phi \end{aligned} \qquad [16]$$

and P_c and P_l denote the Stokes parameters which give the degree of circular and linear polarization, respectively, of the incident beam, $\hat{\boldsymbol{\sigma}}$ is a unit vector along the direction of magnetization, and e^2/mc^2 is the classical electron radius r_0. Equation [15] is valid at low energies, but at high energies the factor $m/2\hbar q$ should be replaced by the more accurate expression

$$m/2\hbar q[1 + E_0/mc^2(1 - \cos\phi)]$$

Due to the presence of the factor $(\boldsymbol{k}_0 \cos\phi + \boldsymbol{k}_1)/mc$, the magnetic part of the cross section increases with the photon energy.

The first term in the cross section of eqn [15] is proportional to the so-called Compton scattering profile $J(p_z)$ due to charge scattering, which is defined as the projection of the total ground-state momentum density $\rho(\boldsymbol{p}) = \rho^{\uparrow}(\boldsymbol{p}) + \rho^{\downarrow}(\boldsymbol{p})$ onto the scattering vector $\boldsymbol{q}$ which was chosen as the z-axis, or, in other words,

$$J(p_z) = \int\int [\rho^{\uparrow}(\boldsymbol{p}) + \rho^{\downarrow}(\boldsymbol{p})]\ \mathrm{d}p_x\ \mathrm{d}p_y \qquad [17]$$

Here, $\rho^{\uparrow(\downarrow)}(\boldsymbol{p})$ denotes the momentum density of electrons with spin up (down) in a magnetic material. In a nonmagnetic material, of course $\rho^{\uparrow}(\boldsymbol{p}) = \rho^{\downarrow}(\boldsymbol{p})$. The momentum density, which is the probability distribution of electrons in momentum space, is given as

$$\rho(\boldsymbol{p}) = \sum_i n_i |\chi_i(\boldsymbol{p})|^2 \qquad [18]$$

where

$$\chi_i(\boldsymbol{p}) = (2\pi\hbar)^{-3/2} \int \psi_i(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\, \mathrm{d}\boldsymbol{r} \qquad [19]$$

is the Fourier transform of the wave function $\psi_i(\boldsymbol{r})$ and the summation runs over all (spin-dependent) states i. A property of the Fourier transform is that $\psi_i(\boldsymbol{r})$ and $\chi_i(\boldsymbol{p})$ weight inverse regions of configuration space: the valence electrons, which tend to lie farther away from the nuclei contribute to $\chi_i(\boldsymbol{p})$ at low momenta, whereas the tightly bound core electrons contribute at higher momenta. Also, since $\chi_i(\boldsymbol{p})$ is the Fourier transform of $\psi_i(\boldsymbol{r})$, it contains the same fundamental information as $\psi_i(\boldsymbol{r})$. Thus, the momentum density describes the electronic system as well as the more familiar charge density. The occupation number n_i gives the number of electrons populating the ith state. In the absence of correlations between the electrons, $n_i = 1$ for an occupied state and $n_i = 0$ when the state is empty. Integrating the Compton profile $J(p_z)$ over p_z is equivalent to integrating the momentum density $\rho(\boldsymbol{p})$ over the entire $\boldsymbol{p}$-space, which from the properties of the Fourier transform is the same as the integral of the real-space charge density. But the integral of the charge density (appropriately normalized) obviously is the number of electrons Z per formula unit. This observation makes it possible to normalize the Compton profile on an absolute scale. However, it is emphasized that the simple relationship [17] between the Compton profile and the electron momentum density (EMD) only exists by the grace of the impulse approximation.

The second term in formula [15], which can be measured due to the circular polarization P_c of the incident radiation, contains the magnetic Compton scattering profile:

$$J_{\mathrm{mag}}(p_z) = \int\int [\rho^{\uparrow}(\boldsymbol{p}) - \rho^{\downarrow}(\boldsymbol{p})]\, \mathrm{d}p_x\, \mathrm{d}p_y \qquad [20]$$

which, like the $J(p_z)$ of eqn [17], is a projection along p_z, except that $J_{\mathrm{mag}}(p_z)$ involves the difference (rather than the sum) of the up- and down-spin momentum densities. $J_{\mathrm{mag}}(p_z)$ is thus related to the ground-state momentum distribution of the electrons with unpaired spin. The magnetic term in the cross section [15] is typically smaller than the charge scattering term proportional to $J(p_z)$ by a few orders of magnitude. This is due firstly to the factor of $(\boldsymbol{k}_0 \cos\phi + \boldsymbol{k}_1)/mc \approx E/mc^2$, since the photon energy is usually only a fraction of mc^2, and secondly because the integral of the magnetic Compton profile $\int J_{\mathrm{mag}}(p_z)\, \mathrm{d}p_z$ equals the number of electrons with unpaired spin, which is much less than the total number of electrons Z. The magnetic effect, however, increases with photon energy. It is to be noted that, in contrast to the scattering of polarized neutrons, magnetic Compton scattering only involves the electron spin, and does not depend on the orbital moment.

It is helpful to consider the momentum density of the noninteracting free-electron gas at 0 K. In this case, electrons occupy all available energy levels up to the Fermi energy E_F. Since $E = p^2/2m$, the occupied states fill a sphere in momentum space, the Fermi sphere, of radius $p_F = \sqrt{2mE_F}$, which is called the Fermi momentum. Thus, $n_i = 1$ and $n_i = 0$ inside and outside the Fermi sphere, respectively. It is then easily shown from eqn [17] that the Compton profile is given by the area of the slice of the Fermi sphere at $p = p_z$ (**Figure 3**):

$$\begin{aligned} J(p_z) &= 2\pi(p_F^2 - p_z^2) \quad p_z \leqslant p_F \\ J(p_z) &= 0 \qquad\qquad\quad p_z > p_F \end{aligned} \qquad [21]$$

which has the shape of an inverted parabola. The factor of 2 in $J(p_z)$ here accounts for the spin degeneracy of free-electron states. Equation [21] describes the basic shape of $J(p_z)$ associated with the loosely bound conduction electrons in metals. The contribution of the more tightly bound core electrons varies smoothly and extends to much higher momenta, so that the total Compton profile possesses the shape of an inverted parabola sitting on top of the broader core contribution (**Figure 3**). The positions of the two kinks in $J(p_z)$ at $p_z = \pm p_F$ mark the Fermi momentum p_F. In a measured Compton profile these kinks

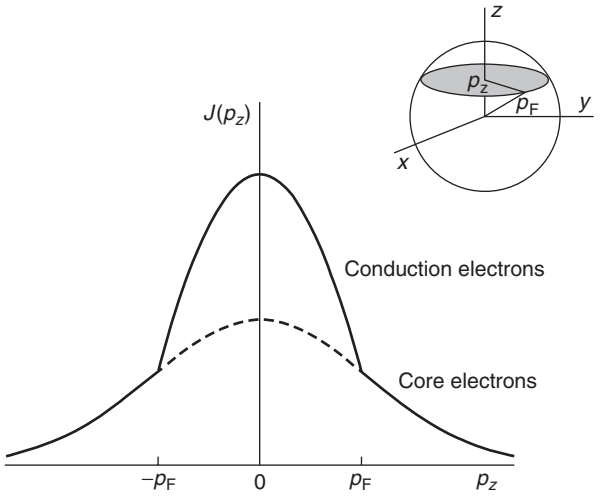


Figure 3 Schematic Compton profile $J(p_z)$ for a hypothetical metal with two core electrons and one free conduction electron per atom. The parabolic contribution of the free electron is given by the area of the slice (for different p_z values) through the Fermi sphere of radius p_F as shown in the inset.

will only appear as rapid changes in slope around the Fermi momentum due to the effect of finite experimental resolution. The width of the rapidly varying portion of the Compton profile is a direct result of the spreading of the electrons throughout the Fermi sphere due to Pauli's exclusion principle and the associated Fermi–Dirac statistics of electrons, which forbids more than one electron (of a given spin) to occupy any energy level.

In a real electron gas, electrons interact through Coulomb forces, so that electronic motions become correlated in the sense that an individual electron no longer moves independently of the positions of other electrons. The Coulomb interaction kicks some electrons out of the Fermi sphere, even at zero temperature, and in the process free-electron (noninteracting) states above the Fermi momentum become partially occupied at the expense of states inside the Fermi sphere. The occupation numbers thus become <1 below E_F and nonzero above E_F, and the momentum distribution becomes more spread out. This effect of electron correlations can be identified in experimental Compton scattering profiles.

In a crystal, electrons experience a periodic potential due to the presence of the underlying lattice. The energy spectrum of the crystal can be shown to consist of groups of levels or bands, where eigenstates within each band are characterized by the wave vector $\boldsymbol{k}$, which is restricted to lie within the first Brillouin zone. The crystal wave function $\phi_i(\boldsymbol{k})$ for the ith band for wave vector $\boldsymbol{k}$ contains momentum contributions not only at $\boldsymbol{p} = \hbar\boldsymbol{k}$, but also at momenta $\boldsymbol{p} = \hbar(\boldsymbol{k} + \boldsymbol{K})$, where $\boldsymbol{K}$ is any one of the group of reciprocal lattice vectors associated with the real-space crystal lattice (there should be no confusion with the photon momenta $\boldsymbol{k}_0$ and $\boldsymbol{k}_1$ introduced earlier). The periodic potential thus has the effect of broadening the valence electron momentum distribution somewhat like the effect of electron correlations. Moreover, the Fermi surface breaks in the momentum density of the crystal will generally occur not only at the momenta $\boldsymbol{p}_F$, but also at all the Umklapp images of the Fermi surface at momenta $\boldsymbol{p}_F + \hbar\boldsymbol{K}$ with appropriate weights.

The twofold integration of $\rho(\boldsymbol{p})$ in eqn [17] reflects the fact that in a standard Compton experiment, only the characteristics of the incoming and outgoing photons are measured and those of the recoiling electrons are ignored. It is obviously attractive to devise a way of measuring the three-dimensional momentum density $\rho(\boldsymbol{p})$. This is possible if the kinematics of the recoil electron is measured in coincidence with that of the Compton scattered photon. This is referred to as a $(\gamma, e\gamma)$ experiment since here one measures the cross section for the process in which a γ-ray photon scatters into a photon and an electron. **Figure 4** clarifies the various momenta involved. The recoil electron is emitted at an angle α (this angle does not necessarily lie in the plane defined by $\boldsymbol{k}_0$, $\boldsymbol{k}_1$, and $\hbar\boldsymbol{q}$) with respect to the primary beam direction. The angle α_0 for an electron initially at rest is given by

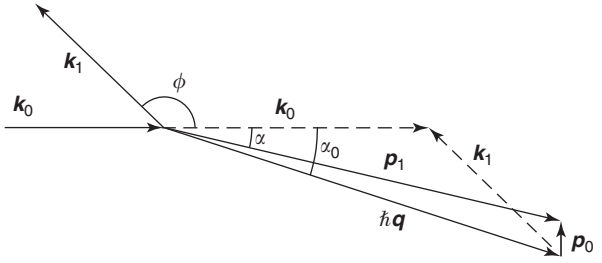


Figure 4 Momentum conservation diagram for $(\gamma, e\gamma)$ scattering, where various momenta have been defined previously in **Figures 1** and **2**. α_0 refers to the limiting case where the electron is initially at rest.

$$\cot \alpha_0 = (1 + E_0/mc^2) \tan(\phi/2) \qquad [22]$$

The components p_x and p_y of the initial electron momentum $\boldsymbol{p}_0$ cause the recoil electron to end up in solid angles closely around the direction α_0, where it can be detected with the aid of a position-sensitive detector. As before, p_z can be determined from the Doppler broadening formula [13]. Alternatively, one may measure the energy of the recoil electron by a time-of-flight method in coincidence with that of the photon. It should be emphasized that if the coincidence condition is not imposed, the energy distribution of the scattered photons as well as that of the recoil electrons is each proportional to the Compton profile. Only by observing the photon and the electron in coincidence does one obtain $\rho(\boldsymbol{p})$. Since electrons interact strongly with matter, $(\gamma, e\gamma)$ experiments require very thin samples in order to minimize the scattering of the recoil electrons.

A technique closely related to $(\gamma, e\gamma)$ scattering is $(e, 2e)$ scattering in which an incoming electron generates two outgoing electrons, the scattered and the recoiling one, and the kinematics of the two outgoing electrons is measured in coincidence. Electron scattering is even more of a problem in such an experiment. Finally, inverse Compton scattering in which a photon takes up energy by collision with an energetic electron plays a role in astrophysics.

Instrumentation

In its simplest form, a Compton scattering experiment involves a source of X- or γ-rays of a

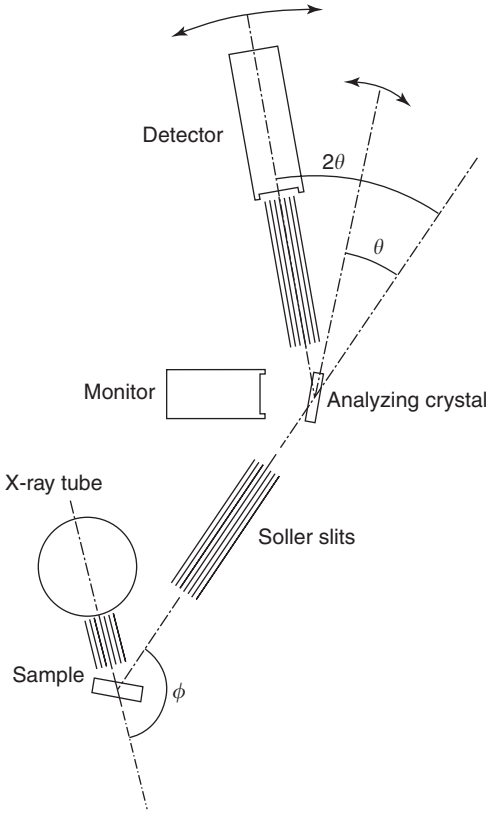


Figure 5 Schematic setup of an X-ray Compton scattering experiment.

well-defined energy (or at least with a strong monochromatic component in its energy spectrum), a sample, an energy analyzer, and a detector to register the scattered radiation (**Figure 5**). Along its path the radiation is narrowly collimated by systems of parallel Soller slits in order to achieve an accurate definition of the various angles and thus of the energy of the scattered radiation. The scattered photons are energy-analyzed by Bragg reflection from an analyzing crystal (usually LiF or Si) and detected with the aid of a scintillation detector. In such a setup, the Compton profile is obtained via a step-by-step energy scan, and the experiment therefore suffers from long data collection times and possible instabilities due to temperature variations, electronic drift, and other effects. Improvement is achieved by replacing the analyzer–detector combination with a high-resolution high-purity or lithium-drifted germanium solid-state detector, which registers the entire spectrum simultaneously, at the expense, however, of resolution.

When an X-ray tube is used, the energy of the incoming radiation is limited to several tens of keV (Mo or Ag anode) or at most to ~ 60 keV for the Kα1 line of a tungsten anode. The flux from X-ray tubes is often too low to allow the use of a primary monochromator, and the scattered spectrum therefore consists of contributions from the Kα1 and Kα2 lines, in addition to other parts of the spectrum, which complicates the interpretation of the measured profiles. Over the years, the count rates have improved with the development of more powerful X-ray tubes, sometimes with rotating anodes, and the use of large samples in combination with bent analyzing crystals, which focus the broad scattered beam onto the detector, thereby allowing radiation with a larger divergence angle to be used. Momentum resolutions for this type of setup vary between 0.15 and 0.5 a.u. full-width-at-half-maximum (1 a.u. of momentum equals 1.9929×10^{-24} kg m s^{-1}), depending on the details of the instrumental design.

The relatively low energy of the radiation provided by X-ray tubes limits Compton experiments to materials with low Z values, since at these energies Compton scattering increases as Z while the competing photoelectric absorption increases as Z^5. Also, at low photon energies, the impulse approximation, which requires the energy transferred in the scattering process to be large compared to the binding energies involved, ceases to be valid for high Z atoms. Higher-energy radiation is needed to circumvent these limitations. γ-ray sources using radioactive isotopes offer the advantage of providing radiation of a high, single energy, and often possess a long half-life $t_{1/2}$. Examples include ^{241}Am (60 keV, $t_{1/2} =$ 425 years), ^{198}Au (412 keV, 2.7 days), and ^{137}Cs (667 keV, 30 years). In combination with a solid-state detector, these sources allow the simultaneous detection of the entire Compton scattered spectrum in an apparatus without moving parts. Shielding is a serious problem at these high energies and limits the maximum scattering angle that can be used. An annular shape of the source with the scattered beam passing through the center of the ring can offer a solution. Scattering in the source itself should be kept low since it produces a tail at the low-energy side of the Compton line, which must be corrected for in data handling procedures. For this purpose, a source material with a high specific activity should be used. The best momentum resolution of ~ 0.40 a.u. that can be achieved with γ-ray sources is substantially worse than that of a high-resolution X-ray setup using an analyzing crystal. This is an important consideration in some investigations since a high momentum resolution of ~ 0.1 a.u. or better is generally needed for observing Fermi surface signatures in the Compton spectra.

The most versatile source of X-rays is synchrotron radiation, which is produced by the acceleration of charged particles, most commonly electrons, going around a circular orbit. The synchrotron radiation beam has a small angular divergence, and is emitted

in the forward direction by the orbiting particles, rather like the searchlight beam in a lighthouse. Photons emitted in the orbital plane are linearly polarized, while those emitted above and below this plane are circularly polarized. The energy range over which an intense flux of linearly or circularly polarized photons is produced can be extended by increasing the acceleration of the orbiting particles over a small portion of their trajectory. This can be achieved with the aid of a suitably designed insertion device, an "undulator" or "wiggler," which introduces a small wiggle in the particle orbit. In a symmetric wiggler the two directions of circular polarization cancel, so the insertion device should contain a certain degree of asymmetry. Since the energy spectrum of synchrotron radiation is "white," a monochromator must be used to "cut" a monochromatic slice from this spectrum. Perfect Si or Ge single crystals with a rocking curve of a width comparable to the beam divergence are often used as monochromators. These have to be cooled to maintain stability in the presence of an intense heat load of several 100 W to a few kW to which they are exposed. The scattered beam is analyzed using a curved analyzing crystal and focused onto a detector. The high intensity of radiation may in some cases cause overloading of the detector electronics. **Figure 6** shows the Compton setup in use at ESRF in Grenoble, France.

Typical synchrotron fluxes at the sources used for Compton scattering vary between 10^{10} and a few times 10^{13} photons s^{-1}, while initial photon energies range from 10 keV to >250 keV. Such fluxes are sufficiently high for carrying out a number of different types of experiments. For example, using unpolarized or linearly polarized light, conventional Compton profiles for electronic structure studies

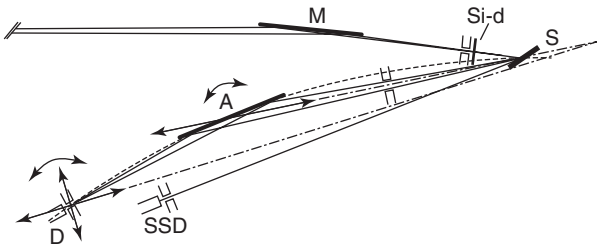


Figure 6 Drawing of the scanning spectrometer at ESRF, Grenoble, France. The beam reflected by the monochromator (M) is collimated by the entrance slit, and the beam scattered by the sample (S) is collimated by the aperture slit. The sample, analyzer crystal (A), and the detector slit lie on the dashed circle. The detector (D) is an NaI scintillation counter. A Si-diode (Si-d) and a Ge detector (SSD) are used for monitoring. The arrows indicate translations and rotations of the analyzer and detector. The monochromator part of the figure is scaled down by a factor of 10. (Courtesy of V Honkimaki.)

may be measured with a resolution of 0.10–0.15 a.u. or better. If the initial radiation is circularly polarized, the presence of the spin-dependent term in the scattering cross section [15] enables one to measure the magnetic momentum density of electrons with unpaired spin in ferromagnetic or ferrimagnetic media (since Compton scattering is an incoherent process, no magnetic effect is observed in antiferromagnetics). Since the synchrotron beam decays exponentially in time, the magnetization of the sample or the circular polarization (helicity) of the photons should be reversed every 10 s or so in the pattern $(+--+)$. A high circular polarization is achieved at the expense of a high flux. In view of the smallness of the magnetic effect and the importance of good statistics, resolution has therefore to be traded against intensity and much of the currently existing magnetic Compton scattering work is limited to a resolution of ~ 0.4 a.u. or worse.

Finally, the way is open to $(\gamma, e\gamma)$ spectroscopy. As discussed in the second section, by measuring the recoiling electron in coincidence with the scattered photon it is possible to determine the three-dimensional momentum density $\rho(\boldsymbol{p})$ of the electrons in a material. $(\gamma, e\gamma)$ measurements are only feasible with an intense flux of high-energy (100–150 keV) photons. Measurements of $\rho(\boldsymbol{p})$ on very thin (17–25 nm thick) graphite samples have been performed in this way with a momentum resolution of 0.60 (0.85) a.u. in the $p_z(p_x)$-direction. A better resolution has been obtained with a time-of-flight spectrometer. In cases where the extremely low cross section of the process makes it necessary to integrate over the energy distributions of the recoil electron as well as the scattered photon, one is left with information only on the perpendicular components p_x and p_y of the initial electron momentum. The quantity measured is then a once-integrated, two-dimensional distribution of the EMD similar to a two-dimensional angular correlation distribution of photons from positrons annihilating in a solid.

Multiple Scattering

Multiple scattering refers to events in which the Compton scattered photon is further scattered one or more times within the sample before it escapes to the detector. The presence of multiple scattering events in the measured Compton profile complicates the analysis of the data. Generally, the multiple scattering contribution tends to broaden the profile somewhat, but its detailed energy dependence on sample geometry and thickness is complex. The effect can be reduced by choosing a thin sample, but that also results in a reduction in the intensity of the primary

Compton scattering events of interest. Attempts to measure a Compton profile for various thicknesses followed by extrapolation to zero thickness suffer from the nonlinear dependence of multiple scattering on sample thickness. The most reliable results are obtained using an iterative Monte Carlo approach to simulate the effect of multiple scattering under the specific experimental conditions.

Data Analysis

Although all electrons contribute equally to the Compton profile, one is usually interested in the properties of the smaller group of valence electrons, which are mainly responsible for bonding as they undergo substantial changes when the material is formed by bringing the atoms together. The tightly bound core electrons are relatively unaffected and their contribution to the Compton profile is isotropic. One way to highlight the contribution of valence electrons in analyzing Compton data is to consider differences $\Delta J(p_z)$ between the Compton profiles measured along various pairs of directions of the scattering vector from an oriented monocrystalline specimen, for example, $\Delta J(p_z) \equiv [J_{100}(p_z) - J_{110}(p_z)]$, for the directional difference between the [1 0 0] and [1 1 0] profiles. The advantage is that the contributions from core electrons, background, and correlation effects (to the extent these are isotropic) cancel in $\Delta J(p_z)$ and the anisotropies related to the valence electrons become clearer.

As seen in the second section, the Compton profile in a crystal generally contains kinks (with appropriate weights) at the Fermi momentum $\boldsymbol{p}_F$ as well as at the images of the Fermi surface at higher momenta induced by the periodic crystalline potential. These kinks originate from breaks in the underlying three-dimensional momentum density $\rho(\boldsymbol{p})$. They can be made more clearly visible by considering the first or even the second derivative of the Compton spectrum.

Relation [17], which defines the Compton profile $J(p_z)$ as a two-dimensional integral of the momentum density $\rho(\boldsymbol{p})$, may alternatively be viewed as an integral equation for determining the three-dimensional function $\rho(\boldsymbol{p})$. The problem can then be formulated as one of "inverting" relation [17] or "reconstructing" $\rho(\boldsymbol{p})$ from a given set $J_\beta(p_z)$ of profiles corresponding to different orientations (scattering vectors) β of a monocrystalline sample. For this purpose, two approaches are commonly used. One is to expand $\rho(\boldsymbol{p})$ into a series of appropriately chosen directional harmonics. These can be the spherical harmonics, but often the crystal lattice possesses an inherent point symmetry, in which case harmonics of a higher symmetry (called lattice harmonics) are advantageous to use since one needs a smaller number of such harmonics. The expansion coefficients in this series can be expressed as Hankel or Bessel transforms of linear combinations of the set of profiles $J_\beta(p_z)$. This method is particularly useful if the momentum density $\rho(\boldsymbol{p})$ and the Fermi surface are not too far from being spherical. Methods of the second category disregard the point symmetry and solve the problem by Fourier transform techniques. The reconstruction itself can be performed in either the $\boldsymbol{p}$- or the $\boldsymbol{r}$-space. In the latter case one considers the reciprocal form factor $B(\boldsymbol{r})$ by taking the Fourier transform of $\rho(\boldsymbol{p})$:

$$B(\boldsymbol{r}) = \int \rho(\boldsymbol{p}) \exp(\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar)\, \mathrm{d}\boldsymbol{p} \qquad [23]$$

from which it is easily shown that

$$B(z) = \int_{-\infty}^{\infty} J(p_z) \exp(\mathrm{i}p_z z/\hbar)\, \mathrm{d}p_z \qquad [24]$$

Thus, the one-dimensional Fourier transform of a Compton profile with the scattering vector along a given direction β equals $B_\beta(\boldsymbol{r})$ along a ray in that same direction in $\boldsymbol{r}$-space through the origin. $B_\beta(\boldsymbol{r})$ can thus be obtained along different rays in $\boldsymbol{r}$-space from the set of Compton profiles $J_\beta(p_z)$. It is then a simple matter to find the three-dimensional function $B(\boldsymbol{r})$ via interpolation, followed by an inverse Fourier transformation of $B(\boldsymbol{r})$,

$$\rho(\boldsymbol{p}) = (2\pi\hbar)^{-3} \int B(\boldsymbol{r}) \exp(-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar)\, \mathrm{d}\boldsymbol{r} \qquad [25]$$

to find the momentum density $\rho(\boldsymbol{p})$.

A useful theorem (the "Lock-Crisp-West" (LCW) theorem) connects the momentum density $\rho(\boldsymbol{p})$ with the occupation density $n(\boldsymbol{k})$, which represents the number of filled electron states at wave vector $\boldsymbol{k}$:

$$n(\boldsymbol{k}) = \sum_{\boldsymbol{K}} \rho(\hbar\boldsymbol{k} + \hbar\boldsymbol{K}) \qquad [26]$$

where the summation is over the reciprocal lattice vectors $\boldsymbol{K}$. The occupation density $n(\boldsymbol{k})$ is a periodic function of $\boldsymbol{k}$, that is, $n(\boldsymbol{k}) = n(\boldsymbol{k} + \boldsymbol{G})$ for any reciprocal vector $\boldsymbol{G}$. The theorem states that the occupation density at wave vector $\boldsymbol{k}$ is obtained by folding the momentum density at all corresponding points $\boldsymbol{p} = \hbar(\boldsymbol{k} + \boldsymbol{K})$ in the higher Brillouin zones back into the first zone.

Comparison with Theory

The methods discussed in the preceding section have all been used extensively in analyzing Compton data. Insight into wide classes of materials has been gained via comparisons between experimental Compton spectra and corresponding first-principles electronic structure computations. In this connection, many reliable techniques based on the local (spin) density approximation of density-functional theory have been developed for treating crystalline systems during the last few decades such as the (linearized) augmented plane wave (LAPW) method, the linearized muffin-tin orbital (LMTO) method, and the Korringa–Kohn–Rostoker (KKR) method. The KKR method is based on the use of multiple scattering theory and possesses the advantage that it can be generalized straightforwardly to treat the electronic structure and momentum densities in substitutionally disordered (random solid solution) phases of alloys

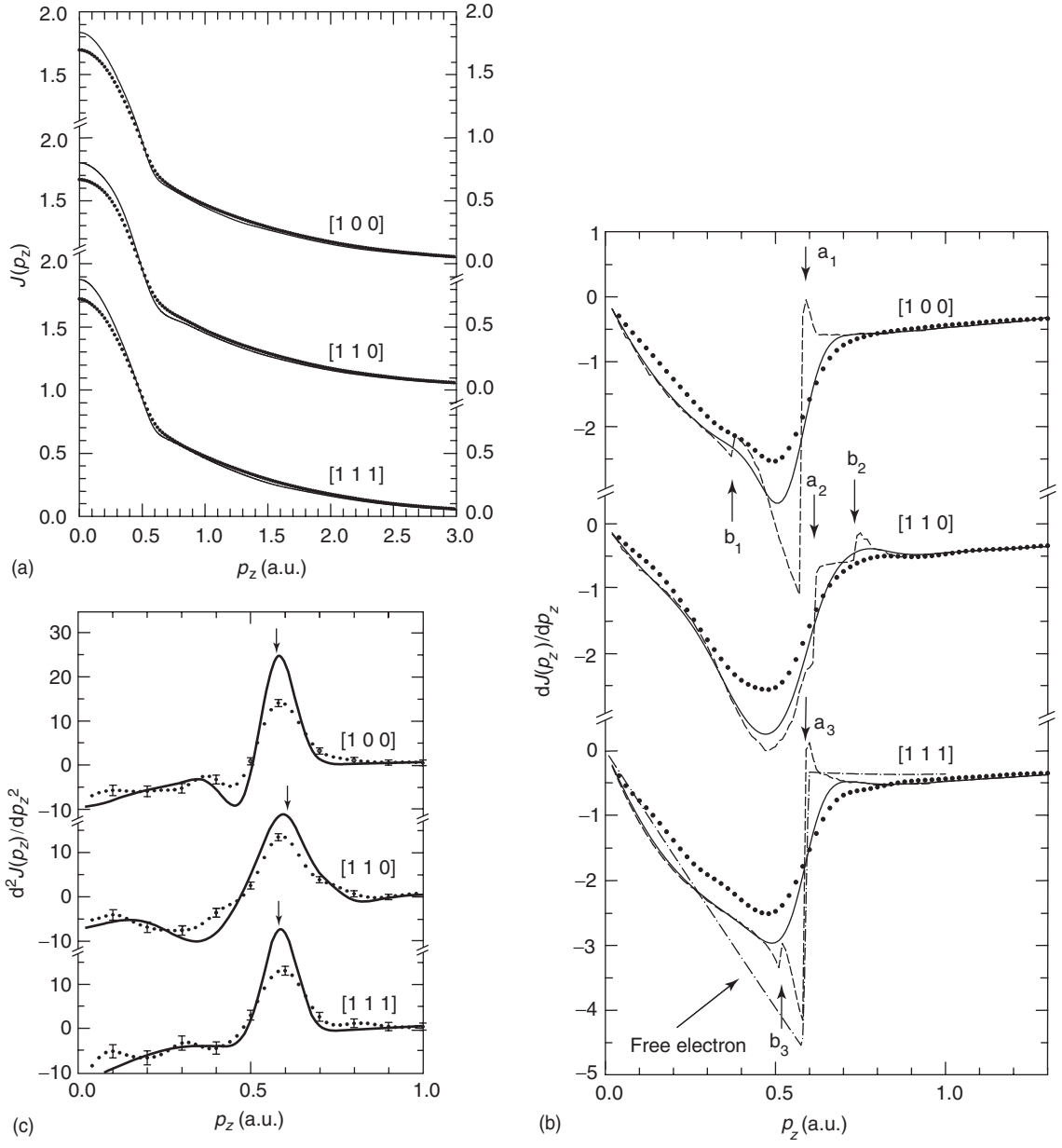


Figure 7 (a) Measured (· · · ·) and computed (——) Compton profile along the three principal symmetry directions of lithium. (b) First derivatives of the profiles. Theoretical profiles are shown with (——) and without (- - - -) resolution broadening of 0.12 a.u. The structures marked a_1–a_3 and b_1–b_3 are Fermi surface breaks. The free-electron result is given for the [1 1 1] case. (c) Second derivatives of profiles. Arrows mark the theoretical Fermi radii. (Reprinted figure with permission from Sakurai Y, Tanaka Y, Bansil A, Kaprzyk S, Stewart AT, *et al.* (1995) *Physical Review Letters* 74: 2252–2255; © American Physical Society.)

within the KKR–CPA scheme. A few illustrative examples of comparison between theory and experiment are presented now. These results will also serve to highlight some of the general features of the Compton spectra that have been discussed throughout this chapter.

Figure 7 shows the Compton profiles for lithium (Li). The measured and computed Compton profiles for the scattering vector lying along three high-symmetry directions are shown in **Figure 7a**. All profiles have been normalized so that the area under each of the profiles equals the number of electrons (three) in Li. The first derivatives of the profiles in **Figure 7b** show that unbroadened theory curves (long dashes) contain breaks at Fermi radii (a_1–a_3) and one set of associated images at $p_F + \hbar K$ (b_1–b_3). The resolution broadening of 0.12 a.u. (full-width-at-half-maximum) washes out the breaks b_1–b_3 as seen from the solid lines. The first derivative for the free-electron profile of eqn [21], shown for reference only along the [1 1 1]-direction by the dot-dashed line, makes it clear that the momentum density in even a relatively simple system such as Li is modified significantly from the free-electron behavior due to solid-state effects. In addition, note how differences in the results for the three directions, that is, the anisotropy of the profile, are more evident in **Figure 7b** compared to **Figure 7a**. The actual Fermi surface radii a_1–a_3 are more easily determined from the second derivatives of **Figure 7c** and are denoted in the theory curves by arrows. The Fermi surface of lithium is a distorted sphere, somewhat elongated along the $\langle 1\,1\,0 \rangle$ directions. The agreement between theory and experiment is good overall, although the theoretical distribution and particularly the Fermi surface breaks are broader in the experiment. This discrepancy reflects the residual electron correlation effects that are not treated accurately in the conventional band theory picture underlying the computations of **Figure 7**.

Figure 8 considers directional anisotropies with the example of an alloy. The results here are for a solid solution of $Ni_{75}Cu_{25}$. The maximum value of the anisotropy $\Delta J(p_z)$ is typically one to two orders of magnitude smaller than the Compton peak intensity. The experimental data in **Figure 8** were taken using a γ-ray source with a momentum resolution of 0.4 a.u. (FWHM), under which all Fermi surface signatures in the data are essentially washed out. Despite some differences in detail, the undulations in $\Delta J(p_z)$ are similar in theory and experiment. The presence of oscillations in $\Delta J(p_z)$ extending to rather high momenta is indicative of interactions in the solid state.

Figure 9 takes up an example where a reconstruction technique has been applied to obtain the Fermi

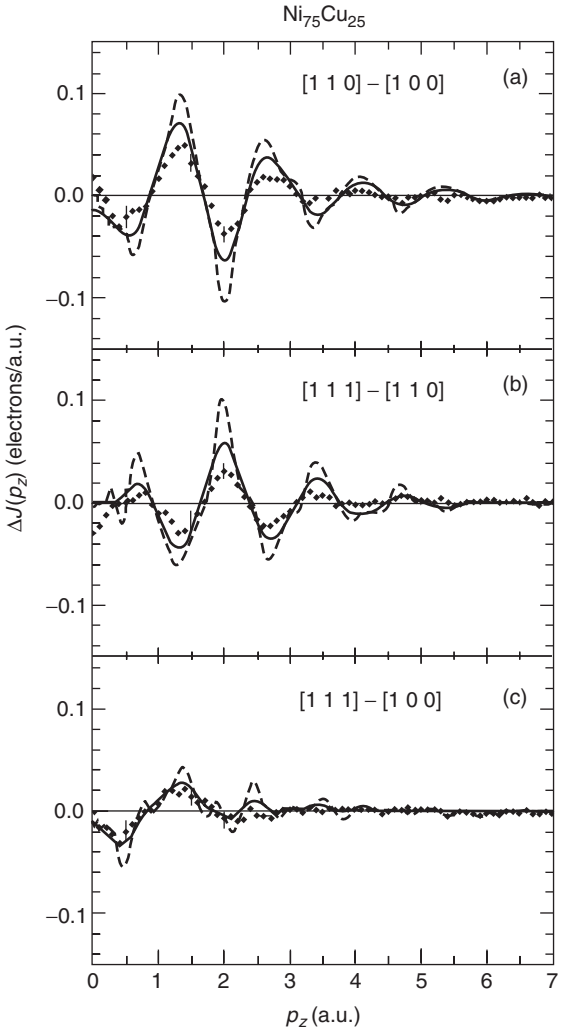


Figure 8 Directional difference profiles for the disordered alloy $Ni_{75}Cu_{25}$. The experimental (♦) results are compared with theoretical profiles with (——) and without (- - -) the experimental resolution of 0.4 a.u. (Reprinted figure with permission from Bansil A, Kaprzyk S, Andrejczuk A, Dobrzyński L, Kwiatkowska J, *et al.* (1998) *Physical Review B* 57: 314–323; © American Physical Society.)

surface of the alloy Al–3at.%Li. Here nine Compton profiles with scattering vectors in the (1 1 0) plane were measured and Fourier-transformed to obtain a projection of the momentum density $\rho(\boldsymbol{p})$ onto the (1 1 0) plane. Using the LCW folding procedure of [26] over a selected set of $\boldsymbol{K}$'s then yields the projected image of the Fermi surface shown in **Figure 9**. The presence of the Fermi surface outside the indicated geometrical boundaries of the first Brillouin zone is clearly seen. The theoretical maps in the figure have been obtained from nine corresponding computed profiles following procedures which were identical to those used to treat the experimental data.

Figure 10 presents the magnetic momentum density of the electrons with unpaired spin in iron. In this

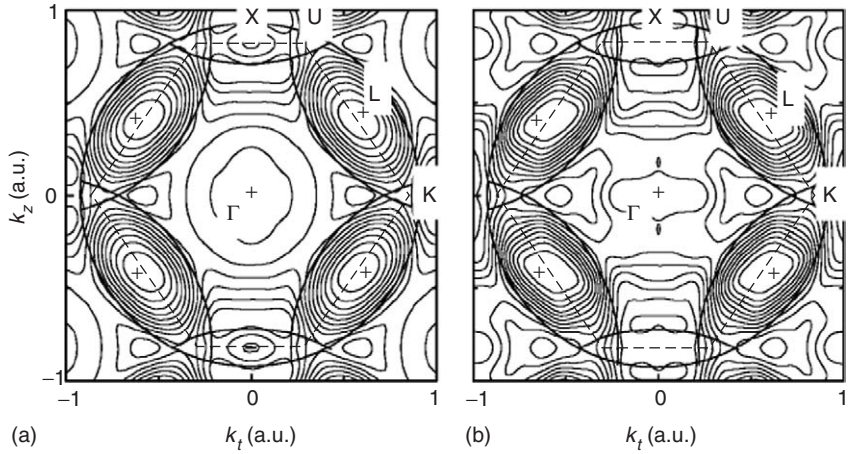


Figure 9 Contour maps of the (a) theoretical and (b) experimental (1 1 0)-plane-projected occupation number densities for Al–3at.%Li, obtained by LCW folding. Resolution broadening (0.12 a.u.) is included in the theory. The dashed lines mark the first Brillouin zone. The plus signs indicate high-density regions. (Reprinted figure with permission from Matsumoto I, Kwiatkowska J, Maniawski F, Itou M, Kawata H, *et al.* (2001) *Physical Review B* 64: 045121; © American Physical Society.)

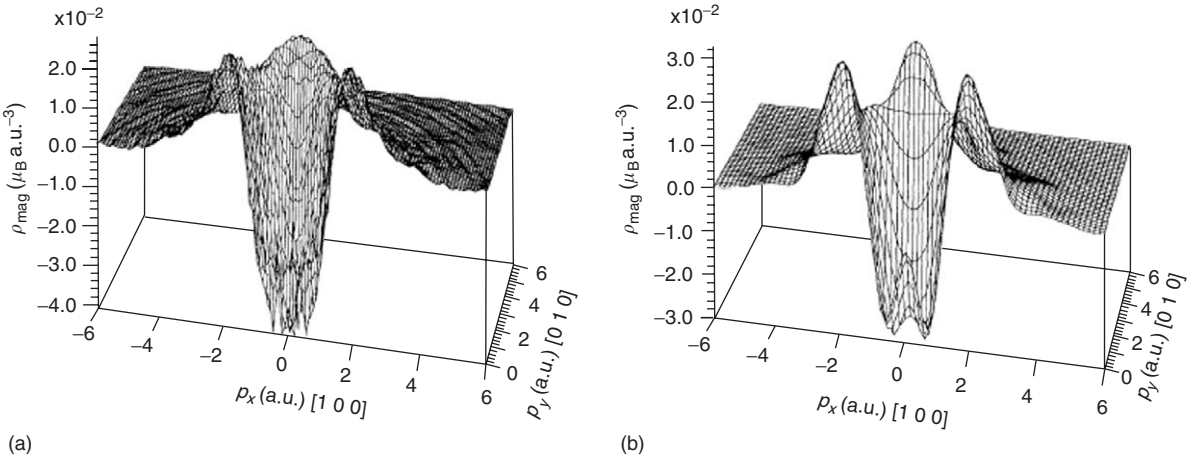


Figure 10 (a) Cross section of the (0 0 1) plane containing the point $\boldsymbol{p} = 0$ through the reconstructed experimental magnetic momentum density $\rho_{mag}(\boldsymbol{p})$ in iron. (b) Corresponding computed magnetic momentum density broadened by the resolution of 0.76 a.u. (Reprinted figure with permission from Tanaka Y, Sakai N, Kubo Y, and Kawata H (1993) *Physical Review Letters* 70: 1537–1540; © American Physical Society.)

study Compton profiles were measured for 14 orientations of an iron single crystal. By reversing the magnetization, two profiles were obtained for each of the 14 orientations; their differences yielded 14 magnetic Compton profiles. The Fourier transforms of these 14 profiles were used to map the reciprocal magnetic form factor $B_{mag}(\boldsymbol{r})$ along the lines of eqn [24]. The inverse transform of [23] was then performed to obtain the magnetic momentum density $\rho_{mag}(\boldsymbol{p})$ of **Figure 10a,** which is seen to display a large negative momentum density at low momenta. **Figure 10b** depicts the corresponding theoretical results and shows a similar general behavior of the computed momentum density.

Finally, a recent example of a complex material is given in **Figure 11**. The magnetic Compton profile from the double layer manganite $La_{1.2}Sr_{1.8}Mn_2O_7$ for the scattering vector along the [1 1 0] direction is shown. The measurements were carried out at a temperature of 5 K under an external magnetic field of 7 T along [1 1 0]. The theoretical profile computed within the standard band theory picture is seen after resolution broadening (thick line) to be in good accord with experiments. Studies such as those shown in **Figure 11** indicate the potential of magnetic Compton scattering as an emerging spectroscopic tool for investigating properties of magnetic electrons in complex materials.

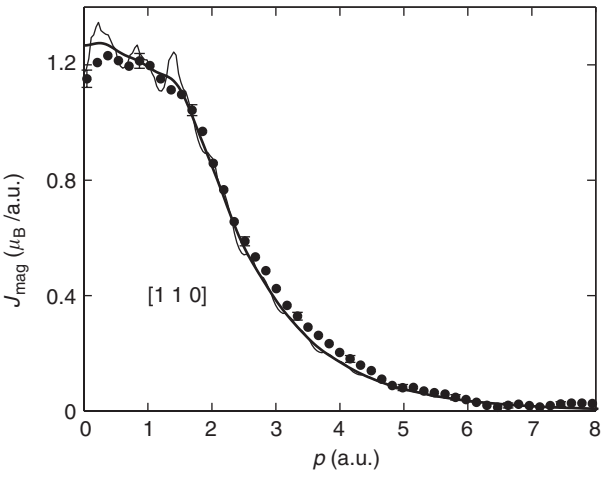


Figure 11 Experimental (•) and theoretical magnetic Compton profile along [1 1 0] in the double layer manganite $La_{1.2}Sr_{1.8}Mn_2O_7$. Measurements were carried out at 5 K under a magnetic field of 7 T. The theoretical profile is shown with (——) and without (—) resolution broadening of 0.4 a.u. All profiles are normalized such that the area gives the measured magnetic moment of 3.4 μ_B/Mn atom. (Reprinted figure with permission from Li Yinwan, Montano PA, Mitchell JF, Barbiellini B, Mijnarends PE, *et al.* (2004) *Physical Review Letters* 93: 207206; © American Physical Society.)

See also: Scattering, Elastic (General); Scattering, Inelastic: Brillouin; Scattering, Inelastic: Electron; Scattering, Inelastic: Raman; Scattering, Inelastic: X-Ray (Methods and Applications); Scattering, Rayleigh.

PACS: 78.70.Ck; 07.85.Nc; 71.20. – **b**

Further Reading

Cooper MJ, Mijnarends PE, Shiotani N, Sakai N, and Bansil A (eds.) (2004) *X-Ray Compton Scattering*. Oxford: Oxford University Press.

Lovesey SW and Collins SP (1996) *X-ray Scattering and Absorption by Magnetic Materials*. Oxford: Oxford University Press.

Williams B (ed.) (1975) *Compton Scattering*. London: McGraw-Hill.

Nomenclature

$B(\boldsymbol{r})$	reciprocal form factor (–)
c	speed of light ($m\,s^{-1}$)
e	charge of the electron (C)
E_0	energy of incident photon ($kg\,m^2\,s^{-2}$)
E_1	energy of scattered photon ($kg\,m^2\,s^{-2}$)
E_C	energy at top of Compton profile ($kg\,m^2\,s^{-2}$)
E_F	Fermi energy ($kg\,m^2\,s^{-2}$)
h	Planck's constant ($kg\,m^2\,s^{-1}$)
$\hbar$	Planck's constant/2π ($kg\,m^2\,s^{-1}$)
J	Compton profile ($kg^{-1}\,m^{-1}s$)
$\boldsymbol{k}$	wave vector (m^{-1})
$\boldsymbol{k}_0$	momentum of incident photon ($kg\,m\,s^{-1}$)
$\boldsymbol{k}_1$	momentum of scattered photon ($kg\,m\,s^{-1}$)
$\boldsymbol{K}$	reciprocal lattice vector (m^{-1})
m	electronic rest mass (kg)
n	occupation function (–)
$\boldsymbol{p}_0$	initial momentum of electron ($kg\,m\,s^{-1}$)
p_1	momentum of recoil electron ($kg\,m\,s^{-1}$)
p_x, p_y, p_z	components of electron momentum ($kg\,m\,s^{-1}$)
p_F	Fermi momentum ($kg\,m\,s^{-1}$)
P_c,P_l	Stokes parameters (–)
$\boldsymbol{q}$	scattering vector (m^{-1})
r_0	classical electron radius (m)
$t_{1/2}$	radioactive half-life (s)
Z	total number of electrons (–)
α, α_0	emission angle of recoil electron (rad)
β	sample orientation (rad)
λ	wavelength (m)
ν	frequency (s^{-1})
ρ	momentum density ($kg^{-3}\,m^{-3}\,s^3$)
σ	cross section (m^2)
ϕ	scattering angle (rad), crystal wave function ($m^{-3/2}$)
χ	momentum wave function ($kg^{-3/2}\,m^{-3/2}\,s^{3/2}$)
ψ	wave function ($m^{-3/2}$)
Ω	solid angle (sterrad)
ε_0	initial energy of electron ($kg\,m^2\,s^{-2}$)
ε_1	energy of recoil electron ($kg\,m^2\,s^{-2}$)

Scattering, Elastic (General)

R Colella, Purdue University, West Lafayette, IN, USA

Introduction

X-rays were discovered in 1895 by the German scientist W C Röntgen, who immediately realized the enormous field of applications resulting from their deep penetration in solid substances. This was dramatically demonstrated by the early photographs of human bones and small metallic objects, such as the wedding ring of Mrs. W C Röntgen. It took several years, however, before their physical nature could be really understood. Since their trajectories were not modified by a magnetic field, the only possible

alternative choice was some form of electromagnetic (e.m.) radiation. The problem was that none of the typical features of e.m. radiation – such as reflection by mirrors, refraction, polarization, and diffraction from a grating – could be reproduced. Refraction, for example, was not considered to take place because no change of angle was observed when a collimated beam of X-rays was observed to penetrate a slab of a solid substance. It was later understood that X-rays are indeed e.m. waves, of extremely short wavelength, ~ 0.1–0.2 nm. Using an oversimplified theory for the index of refraction, based on a model in which atoms are treated as harmonic oscillators, with resonant frequencies much smaller than the X-ray frequency, it is easy to show that the index of refraction is less than 1, and can be written as $n = 1 - \delta$, with $\delta \sim 10^{-5}$. Now it becomes clear why X-rays cannot be reflected by mirrors. In effect, they are reflected, provided the angle of incidence is very small, less than a critical value $\theta_c = \sqrt{2\delta} \sim 10$–$30$ arcmin. The small value of δ also explains why refraction was not observed. It was just too small to be observed with standard techniques borrowed from optics in the visible range. When the precision of the observations was improved, it could be observed and measured. If X-rays are e.m. radiation, one would expect to see polarization effects. These were indeed observed by the English scientist Ch G Barkla, who showed that X-rays scattered at 90° by a carbon block are polarized, which is what one would expect from a transverse e.m. radiation. The decisive event which proved the wave nature of X-rays unambiguously was the discovery in 1912, by the German scientist M von Laue, that X-rays are diffracted by a crystal. In this experiment, a zinc sulfide crystal was used, in which the interatomic distances are ~ 0.2–0.3 nm, supposedly of the same order of magnitude as the X-rays. It became clear that all slits normally used in optical experiments were several thousand times greater than the X-ray wavelength, and therefore no visible diffraction effects could be produced by slits. Diffraction of X-rays by crystals is a milestone in the development of condensed matter physics. In fact, von Laue's experiment not only clarified the wave nature of X-rays, but also gave the key for determining crystal structures, the internal arrangements of atoms in a periodic medium. A new chapter in the development of solid-state physics was started by the English scientist W L Bragg, who showed, in collaboration with his father W H Bragg, how the cell parameters of a crystal could be determined by comparing the intensities of the various diffracted beams in Laue's experiment. They also found, more importantly, that the atomic positions within each crystal cell could be determined. W L Bragg can certainly be considered the founder of modern crystallography.

The big question is: how can a crystal behave the same way as a ruled grating? Laue's equations are derived under the assumption that each atom acts as a source of spherical waves, which are capable of producing mutual interference. These equations show that the interference is generally negative (destructive), except for certain special directions in space. The analysis is considerably simplified if Bragg's viewpoint is adopted. Bragg's scheme is based on the notion of "atomic plane," plane surfaces populated by atoms. Each atomic plane is considered to behave like a weakly reflecting mirror. When the distance d between parallel atomic planes, the angle of incidence θ, and the wavelength λ are related so that the beams reflected by each plane are all in phase, a strong reflection is produced (**Figure 1**). Since there are many possible choices of atomic planes in a point lattice, many diffracted beams are produced in several different directions (**Figure 2**). For example, in a cubic point lattice the most populated atomic planes are defined by the faces of the cube. But one can also visualize atomic planes containing the body diagonals of the cube. Different atomic planes are identified by means of sets of three integer numbers, called "Miller indices." They are the reciprocals of the intercepts that a given atomic plane, closest to the origin, cuts through the three axes used to define the point lattice of the crystal. So, for example, the faces of the cube in **Figure 3** are the {1 0 0} planes, the planes through the body diagonals are the {1 1 0} planes, etc.

In the Laue method, the incident X-ray beam is "polychromatic." In this way, every set of atomic planes $\{h\,k\,l\}$ will utilize appropriate X-ray energy, present in the incident beam, to produce a spot on the film. So, in principle, every spot of the diffraction pattern corresponds to a given set of Miller indices. The Laue technique, the oldest experimental

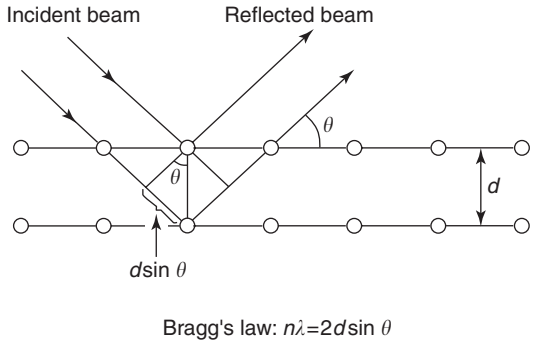


Figure 1 Reflection of waves from two adjacent parallel Bragg planes.

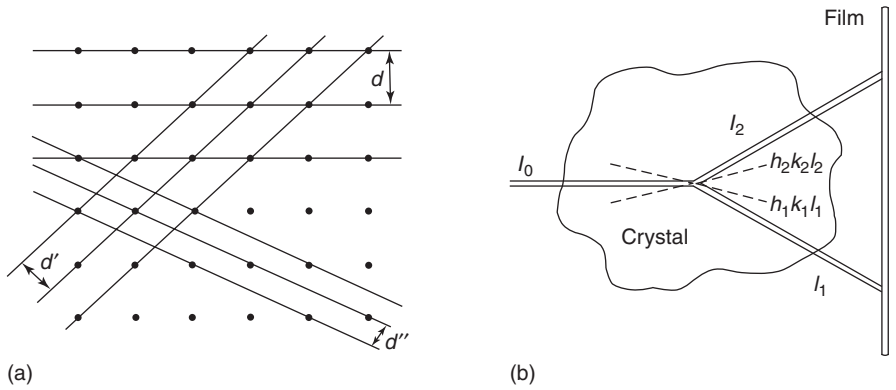


Figure 2 (a) Three sets of parallel Bragg planes with different grating spacings. (b) Principle of the Laue method. The incident beam is polychromatic. Two sets of lattice planes h_1 k_1 l_1 and h_2 k_2 l_2 reflect two monochromatic beams I_1 and I_2, with different wavelengths, according to Bragg's law.

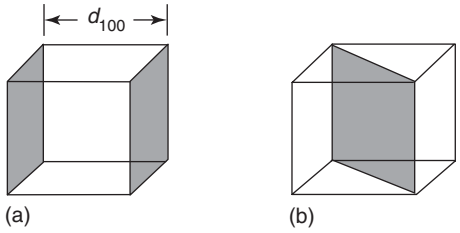


Figure 3 Lattice planes corresponding to the (a) {1 0 0} and (b) {1 1 0} Miller indices in a simple cubic lattice.

technique for X-ray diffraction, is still used for finding the orientation of single crystals. It has been resurrected as a research experimental tool for the crystal structure determination of proteins and other biological crystals. It is used at synchrotron sites, in a time-resolved mode, to study time-dependent phenomena, with resolution times $\sim 10^{-9}$ s or less.

Temperature Effects

It is already discussed in the previous section that a three-dimensional ordered array of pointlike scattering centers can produce positive interference along certain directions in space, giving rise to diffracted beams. The assumption was that these scattering centers are fixed. In reality, it is known that atoms in a solid are not stationary; they vibrate around their equilibrium positions. These thermal vibrations are the origin of thermal energy. Two identical bodies at different temperatures correspond to two identical atomic arrays with different amplitudes of thermal vibrations. If the atoms are at rest, exactly located at their lattice positions, no X-rays are scattered outside of the Bragg beams. Thermal vibrations produce attenuation of the Bragg beams, and diffuse scattering outside of the Bragg beams. Thermal vibrations are present at all temperatures, even at absolute zero. It is a phenomenon called "zero-point energy," an effect predicted by the quantum theory and amply verified by experiment. Typical frequencies of thermal vibrations are in the range of 10^{12}–10^{13} Hz. Thermal amplitudes are normally a large fraction of the interatomic distances. They can be easily $\sim 10\%$ of the distance between atoms. The time a photon takes to go through a crystal and be channeled into one of the diffracted beams is comparable with the period of thermal vibrations. So, the crystal appears to be totally disordered to an incoming photon. This fact was well recognized by the early pioneers of X-ray physics, to the point that it was believed that X-ray diffraction could not be possible as a consequence of thermal disorder. Fortunately, Laue decided to try anyway, and the experiment was a great success.

It was later understood that the process of X-ray scattering from a vibrating atom is equivalent, apart from inelastic effects, to scattering from an electronic "cloud," which represents the time average of the electronic charge. So, a thin two-dimensional sheet of atoms acquires a certain "thickness," as a result of thermal vibrations. Clearly, interfering effects over the electronic clouds decrease the scattering amplitude by a factor e^{-M} called "Debye–Waller factor." It should be emphasized that thermal disorder is not equivalent to the kind of structural randomness present in a glass or a liquid. The diffraction spots are sharp and well-defined; only they become weaker as temperature is increased. What counts is the time average of the charge density, which still preserves the same kind of long-range order present in a nonvibrating crystal.

The Debye theory gives for a monoatomic crystal

$$M = \left(\frac{\sin\theta}{\lambda}\right)^2 \frac{6h^2T}{M\kappa_B\Theta^2}\left(\varphi(x) + \frac{x}{4}\right) \qquad [1]$$

where θ is half of the scattering angle, λ the X-ray wavelength, h is the Planck constant, M the atomic mass, T the absolute temperature, κ_B the Boltzmann constant, Θ the Debye temperature, and $\varphi(x)$ is the Debye function

$$\varphi(x) = \frac{1}{x}\int_0^x \frac{y\,dy}{e^y - 1} \quad \text{with } x = \frac{\Theta}{T} \qquad [2]$$

Note that eqn [1] shows that M is never zero, not even at $T = 0$.

One important point which has been clarified only recently (say, after 1960) is that the diffraction spots of a vibrating crystal are due to purely elastic scattering events. The inelastic events produce scattering outside of the Bragg reflections. This is surprising because every scattering event involves a substantial transfer of energy from the X-ray photons to the scattering centers (the energy transfer can be positive or negative). For a large number of scattering events, the energy is transferred to the whole crystal, not just to a single atom. This important point is the physical origin of the Mössbauer effect. It turns out that the ratio of recoilless events is just given by e^{-M}, typically between 0.9 and 1.

Dynamical Diffraction

The amplitude of X-rays scattered by the unit cell in a crystal is given by

$$A(\boldsymbol{H}) = \sum_i f_i \exp(i\boldsymbol{r}_i \cdot \boldsymbol{H}) \qquad [3]$$

where f_i are the individual scattering form factors for individual atoms, $\boldsymbol{r}_i$ are the atomic sites, and $\boldsymbol{H}$ is the momentum exchanged in the scattering process. It is also called "scattering vector," and it has components h, k, l, the Miller indices of the diffracting planes. Equation [3] simply states that each atom becomes a source of radiation, with an appropriate phase factor. It is assumed in eqn [3] that each X-ray photon is scattered "only once," within each cell. This is the essence of the first Born approximation in the scattering theory. When many unit cells are considered, this assumption is not always verified. Typical crystal size dimensions are in the range of 10 μm or more. Each atom acts as an oscillator, capable of absorbing and emitting radiation. A balance between emission and absorption is established, which is consistent with the notion of multiple scattering. A condition of self-consistency is set up for all plane waves existing in the crystal. This was the doctoral thesis assigned in 1912 by A Sommerfeld to P P Ewald before the discovery of X-rays by W C Röntgen: to investigate what happens when electromagnetic radiation of short wavelength (0.1–0.5 nm) interacts with a crystal. For X-rays, the appropriate set of equations suitable for this problem are Maxwell equations; for electrons and neutrons – the Schrödinger equation. In either case, the crystal is described as a medium which is periodic in three dimensions. More specifically, the index of refraction is supposed to be periodic in three dimensions, and can therefore be Fourier analyzed. In principle, an infinite number of plane waves is considered, one for every reciprocal lattice node, and an infinite number of homogeneous linear equations is obtained for the amplitudes of the plane waves. In practice, a finite number n of plane waves is considered, where n is the number of sets of atomic planes for which Bragg's law is close to be satisfied. The number n is typically 2, which means that only one Bragg reflection is excited. It turns out that solutions can be found for two distinct values of the index of refraction (eigenvalues); in other words, the crystal becomes "birefringent" when Bragg's law is close to being satisfied for a particular set of atomic planes. An incident beam with wave vector $\boldsymbol{k}_o$ generates in the crystal two refracted beams with the same tangential components $\boldsymbol{\beta}_{o1}$, $\boldsymbol{\beta}_{o2}$, which in turn, by virtue of Bragg's law, generate two diffracted beams with wave vectors $\boldsymbol{\beta}_{H1}$, $\boldsymbol{\beta}_{H2}$, (H stands for h, k, l) . The two wave vectors $\boldsymbol{\beta}_{o1}$, $\boldsymbol{\beta}_{o2}$ only differ by a small amount along the normal to the surface. The same holds for $\boldsymbol{\beta}_{H1}$, $\boldsymbol{\beta}_{H2}$. It is expected that waves with almost identical wave vectors will generate beating effects. Such effects can only be seen in highly perfect crystals. These beating effects were predicted by M von Laue in the 1930s and named by him as "Pendellösung," but they could only be observed in the late 1950s, when the semiconductor industry was able to produce highly perfect crystals of germanium and silicon, which were needed for the production of transistors. It was also realized that the two eigenvalues for the index of refraction, corresponding to two different modes of propagation called "wave fields," acquire different absorption coefficients when Bragg's law is satisfied. One mode (can be called α) propagates through the crystal very easily, even when the average photoelectric absorption is very strong (the value in operation when Bragg's law is far from being satisfied). The α mode corresponds to a situation called 'anomalous transmission." The other mode (can be called β) corresponds to a very high absorption, higher than the average value. The physical reason for this difference in absorption coefficients is that at the Bragg condition a system of standing waves is set up in the crystal for both branches, α and β. The α branch has maxima between atomic planes; therefore, absorption is weak because there are no electrons

where the electric field has maxima. The β branch has maxima on the atomic planes; therefore, it is heavily absorbed. The standing waves maxima and minima can be shifted in a controlled fashion between the atomic planes. By looking at the X-ray fluorescence emitted by particular impurities (such as arsenic in silicon), it has been possible to confirm that the arsenic atoms lie on substitutional sites. The X-ray standing waves technique (XRSW) is also used quite often to characterize the structure of epitaxial layers deposited on a perfect crystal, such as silicon or germanium.

A perfect crystal is characterized by a high level of coherence between distant regions, which favors multiple scattering and consequently a marked reduction in scattered intensity. A strained region therefore diffracts more intensely. This is the basis of "X-ray topography," a technique in which the contrast of X-ray images is used to learn about the structural defects present in the crystal.

X-Ray Interferometry

Interference between coherent beams requires precise machining of the optical elements (lenses and mirrors) to an accuracy of the order of the wavelength of light or less. For visible light, this means an accuracy of $\sim 5000\,\text{Å}$, which is feasible. For X-rays a typical wavelength is 1 Å, which is comparable to the size of an atom. So, it would seem that X-ray interferometry is out of the question. In 1965 a report by Ulrich Bonse and Michael Hart appeared in print in which coherent interference effects were observed between spatially separated X-ray beams. Use was made of "Laue diffraction," namely, diffraction through crystal slabs, in transmission geometry, from lattice planes perpendicular to the surface. Typically, in Laue diffraction, two beams emerge from the backside of the slab, which are coherently related (**Figure 4**). By combining three silicon slabs in sequence, U Bonse and M Hart were able to observe interference fringes when a plastic wedge was inserted in one of the interfering beams. The thickness gradient was perpendicular to the plane of **Figure 5**. To guarantee a perfect alignment of the three slabs with atomic precision, they were carved out from a monolithic silicon block (**Figure 6**). A plastic wedge was introduced on one of the interfering beams in order to produce a variable phase shift, which gave rise to a number of interference fringes on the beams T and H emerging from the backside of the right slab in **Figure 5**. Since it is impossible to polish a surface to angstrom accuracy, one may wonder how X-ray interferometry can be observed in the angstrom region. The question to be asked, both for X-ray and optical interferometry, is

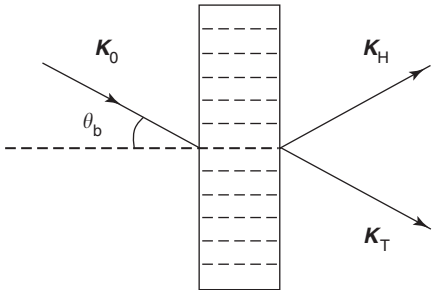


Figure 4 Geometry of Laue case diffraction. The diffracted beam propagates in a half-space which is the backside of the crystal, different from the one in which the incident beam is propagating.

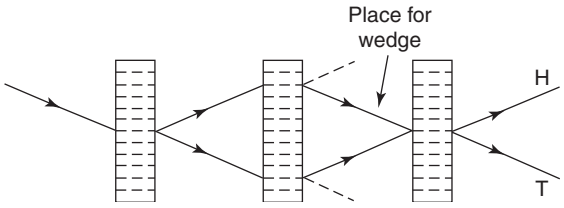


Figure 5 The three crystals are combined in one single unit. They are carved out from the same monolithic block. A plastic wedge is inserted along one of the interfering beams. The upper optical path is different from that of the lower beam. Interference fringes are expected to become visible.

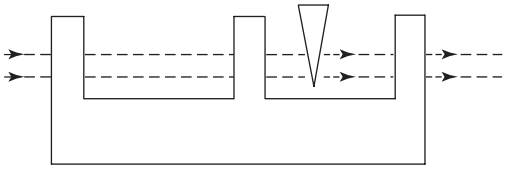


Figure 6 Side view of the Bonse–Hart interferometer. A plastic wedge is inserted to intercept one of the interfering beams. The optical path in the upper ray is smaller than that in the lower ray. Interference fringes are produced and made visible on a photographic film placed on the backside of the interferometer.

how much is the phase of a beam going to be affected by a step Δt. In the case of visible light, the dephasing effect amounts to $2\pi\ \Delta t/\lambda$, which is a big effect. In the case of X-rays, the situation is different, because X-rays are scattered by "atoms," not by the surface itself. The dephasing amounts to $2\pi\Delta t(n-1)/\lambda$, which is quite small, because $(n-1)$ is $\sim 10^{-5}$, in magnitude. So, for X-rays, relatively large values of Δt can be tolerated. To enable coherent photons to meet on the third slab on the right (**Figure 5**), the distances between the slabs should be "exactly" equal. How exactly? Within the longitudinal and transverse coherence lengths of the X-ray photons, typically in the range of 5–20 μm, which is not too hard to achieve in most cases.

A question may be asked about the requirement of monochromaticity for the incident X-ray beam. The trajectories of the X-ray photons are determined by the Bragg diffraction angles. Since the peak widths of the Bragg diffracted beams are quite narrow (seconds of arc), photons with even slightly different energies follow distinctly different paths. Since every photon can only interfere with itself, it is realized that $\Delta\lambda$ does not depend on the monochromator upstream, but only on the peak width $\Delta\theta$. So, monochromaticity is not an issue.

X-ray interferometers have been used by Deslattes and Henins in 1973 to calibrate X-ray wavelengths against optical standards. The same X-ray interferometers have been successfully used with thermal neutrons giving rise to a whole new field of "neutron interferometry." Many quantum effects have been observed for the first time by means of neutron interferometry and have been described in the book by Rauch and Werner. The list includes the following:

1. The observation of gravitationally induced quantum interference.
2. The observation of the phase shift of a neutron due to precession in a magnetic field.
3. The effect of the Earth's rotation on the quantum mechanical phase of the neutron, and many others.

Multiple Beam Diffraction

An X-ray diffraction experiment yields the magnitudes of the various Fourier components of the scattering object. The phases, however, are lost. This is a severe limitation in the ability to determine structures by means of diffraction experiments. This is the essence of the so-called "phase problem" in crystallography. One technique that provides experimental determination of phases is based on the principle of multiple Bragg scattering, a situation in which two or more Bragg reflections are excited at the same time. Suppose that a weak Bragg reflection $\boldsymbol{P}$ is excited, and that the crystal is rotated around $\boldsymbol{P}$ in such a way that the incident X-ray beam always forms the same angle with the lattice planes associated with the $\boldsymbol{P}$ reflection. A plot of the intensity of $\boldsymbol{P}$ versus the rotation angle ψ is called "azimuthal plot." When another reflection $\boldsymbol{H}$ is excited simultaneously for a particular value of ψ, a strong perturbation in the $\boldsymbol{P}$ intensity is observed, typically a sharp peak. A careful analysis of the azimuthal plot reveals that in most cases a pronounced asymmetry is present at the base of the peak. This is called "the asymmetry effect." It can be shown that the asymmetry effect is due to interference between Bragg beams, and that it carries phase information. Specifically, a three-beam experiment (which involves two Bragg reflections plus the incident beam) provides a value for the so-called "triplet invariant":

$$\delta = \varphi_{\boldsymbol{H}} + \varphi_{\boldsymbol{P}-\boldsymbol{H}} - \varphi_{\boldsymbol{P}}$$

This definition of δ makes it independent of the origin chosen for calculating structure factors. For a centrosymmetric structure, δ is restricted to two possible values, 0° or 180°. These ideas have been fruitfully applied to improve the understanding of the structure of quasicrystals. This is the name given to solid substances which are not periodic, yet they are capable of producing sharp diffraction peaks, similar to those produced by regular crystals. In structural analysis of new substances, the first question to be asked is whether the structure is centrosymmetric or not. It was generally assumed, without a strong justification, that most quasicrystals, with very few exceptions, are centrosymmetric. However, when there is no periodicity, it is difficult to understand what centrosymmetry really means. A number of three-beam diffraction experiments have been done to clarify this issue. **Figure 7** shows the azimuthal plot of Al–Pd–Mn quasicrystal, a very high-quality material, generally considered to be centrosymmetric. It is noted that the scattering vectors for quasicrystals are defined in a six-dimensional space, and therefore they have six components. It is a mathematical convenience which allows one to treat quasicrystals as periodic objects in a six-dimensional space. The asymmetry effect is clearly visible, and it is consistent

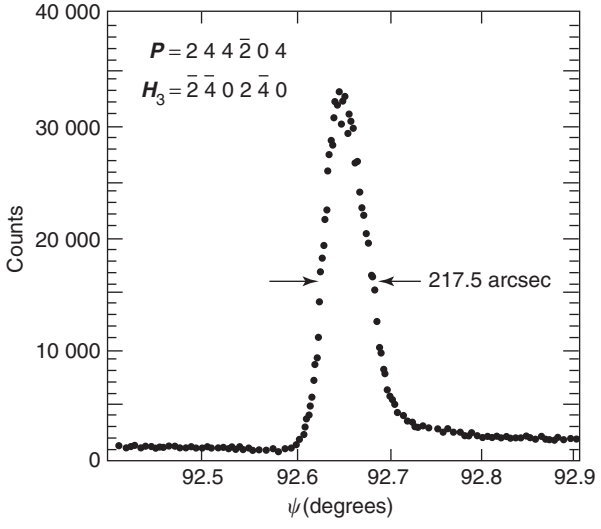


Figure 7 Azimuthal plot of the $(2\,4\,4\,\bar{2}\,0\,4)$ reflection, a three-beam case. The simultaneous reflection $\boldsymbol{H}_3$ is the $(\bar{2}\,\bar{4}\,0\,2\,\bar{4}\,0)$. The X-ray energy is 7000 eV.

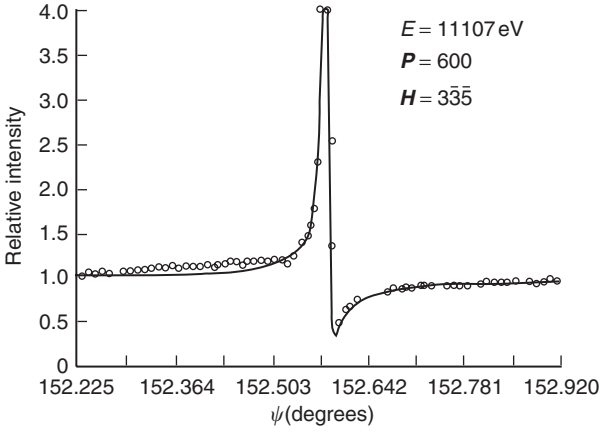


Figure 8 Azimuthal dependence of the 600 in germanium, near a peak due to three-beam diffraction (the $3\,\bar{3}\,\bar{5}$). Circles represent experimental data. Solid line is theory. The agreement is excellent.

with a triple invariant $\delta = 112.5^\circ$, a clear indication of lack of centrosymmetry. While the deviations from centrosymmetry are probably very small, it is found that multibeam diffraction tends to overemphasize even a small amount of centrosymmetry.

Another area in which three-beam diffraction is finding useful applications is that of resonant scattering. In germanium crystal, the 600 reflection is space-group forbidden. This means that the 600 cannot be turned on even for the most general charge density consistent with the symmetry properties of the Fd3*m* space group. This means that if an electron is put in a general position (x, y, z) within the unit cell, and then one electron is placed in all other 47 equivalent positions, as dictated by the symmetry properties of the space group, the 600 structure factor for these electrons is always zero. Surprisingly enough, the 600 can be turned on again if resonant scattering is used. It is found that the 600 exhibits a beautiful asymmetry effect around a three-beam situation, and that interesting phase effects can be observed, which should provide a much more detailed description of the valence charge density in Ge. **Figure 8** is an example of a three-beam experiment in Ge-600 at resonance. Again, the asymmetry effect is clearly visible, and phases can be reliably extracted from the azimuthal profiles.

See also: Crystal Structure Determination; Mössbauer Spectroscopy; X-Ray Topography.

PACS: 61.10.Nz; 61.10.Dp; 61.44.Br

Further Reading

Bonse U and Hart M (1965) An X-ray interferometer. *Applied Physics Letters* 6: 155.

Colella R (1996) X-ray and neutron interferometry. In: Authier A, Lagomarsino S, and Tanner BK (eds.) *X-Ray and Neutron Dynamical Diffraction*. New York: Plenum.

Deslattes RD and Henins A (1973) *Physical Review Letters* 31: 972.

Rauch H and Werner SA (2000) *Neutron Interferometry*. Oxford: Clarendon.

Shvyd' ko YV (2004) *X-Ray Optics. High Energy Resolutions Applications*. Berlin: Springer.

Scattering, Inelastic: Brillouin

T Blachowicz, Department of Microelectronics, Silesian University of Technology, Gliwice, Poland
M Grimsditch, Argonne National Laboratory, Argonne, IL, USA

Published by Elsevier Ltd.

Introduction

Brillouin light scattering (BLS), in its historical sense, is the inelastic scattering of light by acoustic phonons. It thus provides information on the elastic properties of the scattering medium and it has been used to study gases, liquids, crystals, polymers, glasses, semiconductors, nontransparent thin layers, and superlattices. BLS yields information similar to that obtained using ultrasonic techniques but it has certain advantages when dealing with small samples, reactive samples, or in cases where contact with a transducer poses experimental difficulties (e.g., high temperatures and pressures). The Fabry–Perot interferometer (FP) is, par excellence, the instrument of choice to achieve the 1–100 GHz resolution required in typical BLS experiments. Because of this, BLS has become synonymous for all experiments performed with an FP independent of whether the scattering is produced by acoustic phonons, by magnons, or by electronic states. Here BLS is used in this broader context.

This article provides a brief historical review and a condensed description of the origin of the effect. Also, some key experimental aspects are briefly described. The emphasis of this article is to illustrate, via specific examples, the range of problems that can

be investigated using BLS. This will include the determination of elastic constants with emphasis on its application under adverse experimental conditions (temperature, size, and pressure), as well as its application to magnetic systems and electronic levels.

Brief History of BLS

The term Brillouin light scattering (BLS) was derived from the theoretical prediction made by L Brillouin in 1914 of a possible inelastic interaction between light and sound. The description of light-scattering experiments was published in his PhD dissertation. A similar prediction was made independently in 1926 by L I Mandelstam. The first experiment was performed by E Gross in 1930. BLS spectroscopy developed rapidly in the early 1960s as lasers became available. The first light-scattering experiments in pure liquids were done by G B Benedek *et al*. in 1964. The technique remained almost an academic curiosity until the multipassed and later the tandem FP, both introduced by J Sandercock, enabled the weak Brillouin signals to be observed in the presence of the usually overwhelming intensity of the elastically scattered light. At that point, BLS became a powerful and versatile tool for material characterization. This article describes how it has made significant contributions to a wide range of problems in condensed matter physics.

Origin of BLS

Any excitation that produces a time-varying change in the polarizability of a medium (for a longitudinal sound wave this is related to the change in refractive index caused by the densification and rarefaction produced by the wave) interacts with electromagnetic radiation via the polarizability α. The polarization of the medium induced by the electromagnetic wave becomes

$$\boldsymbol{P}(\boldsymbol{r},t) = [\langle \alpha \rangle + \delta\alpha(\boldsymbol{r},t)]E_0 \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_0 t)}$$

and it can be shown that this leads to the re-radiation of electromagnetic waves with frequencies $\omega_0 \pm \omega_{\mathrm{phon}}$. The origin of the polarizability changes ($\delta\alpha$) is different for different excitations: bulk phonons modulate the polarizability via elastooptic effects (strain-induced changes in the refractive index), surface phonons produce polarizability changes as the interface moves into and out of the vacuum (known as the ripple mechanism), and magnons interact via magnetooptic contributions.

Interference between the radiation emanating from throughout the sample leads to the condition that only excitations with wave vector $\boldsymbol{q}$ that satisfy

$$\boldsymbol{q} = \boldsymbol{k}' - \boldsymbol{k}$$

where $\boldsymbol{k}$ and $\boldsymbol{k}'$ are the wave vectors of the incident and scattered radiation, respectively, will produce scattered radiation. Since $\boldsymbol{k}$ and $\boldsymbol{k}'$ are determined by the directions of the incident and scattered radiation, it is clear from the above equation that in a given experiment only excitations with a fixed wave vector will be probed. This also makes it clear that the geometry chosen for the experiment is crucial for analyzing the resulting data.

Scattering Geometries and Frequency Analysis

There are a number of scattering geometries that are often used in Brillouin experiments; these are shown schematically in **Figure 1**. Not only does each scattering geometry select different q values but it is clear that geometries (a), (b), and (c) are only suitable for transparent materials while geometries (d) and (e) can be used even if the sample is not transparent. The (e) geometry is a variant of the back-scattering geometry (d). The magnitude of the wave vector probed in each geometry, calculated from the $\boldsymbol{q} = \boldsymbol{k}' - \boldsymbol{k}$ equation, is also indicated in **Figure 1**. In these expressions, $|\boldsymbol{k}| = 2\pi/\lambda$ (with λ the wavelength of the incident radiation, and the refractive index of the medium). In opaque materials, only the component of q parallel to the surface is conserved so that the $\boldsymbol{q} = \boldsymbol{k}' - \boldsymbol{k}$ equation yields only the value of $\boldsymbol{q}$ in the surface plane $\boldsymbol{q}_s$. Typically, the choice of scattering geometry depends on the nature of the sample and the constraints of the experiment being performed; for example, cryostat, magnet, and furnace. In the laboratory, the incident radiation depicted in **Figure 1** is usually delivered by a laser. Since the frequency shifts are small, it normally requires a laser operating in a single longitudinal cavity mode. The scattered radiation is typically collected by a high-quality lens (often a camera lens). Frequency analysis of the scattered light is accomplished using an FP interferometer. In all but a few cases, experiments rely on the high contrast provided by multipassed and/or tandem instruments. These very delicate instruments are technically complex. Excellent reviews on their operation and fabrication exist in the literature (see the "Further reading" section).

Examples

To highlight BLS, a selection of examples are discussed. First, the classical application is described: the

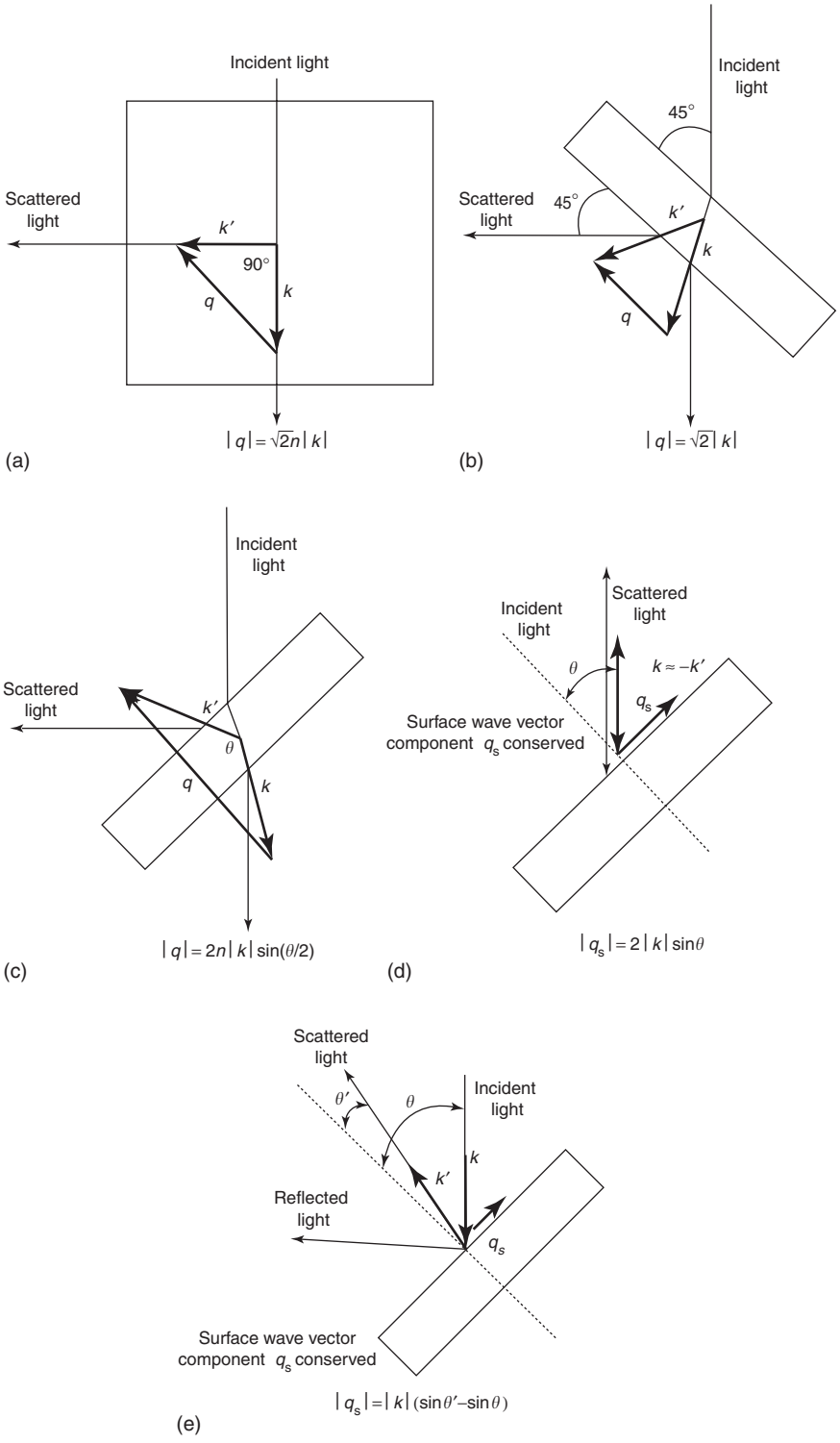


Figure 1 Scattering geometries: (a) a bulk-type analysis of excitations inside a transparent material using 90° geometry, (b) a thin transparent layer – this geometry is sensitive to excitations propagating in the plane of the sample, (c) a transparent layer with near-normal incidence – this geometry is sensitive to wave vectors perpendicular to a sample surface, (d) a back-scattering geometry in an opaque material which is sensitive to in-plane, surface wave vectors, and (e) a variant of geometry (d), where $\boldsymbol{k}'$ is not along the direction of the incident light, is also used for studying surface excitations. For geometries (a) and (c), the material refractive index n is required for data analysis.

determination of elastic constants in a bulk transparent material. In this subsection, references to examples are provided where the determination of c_{ij} has been done at high temperatures and high pressures. Then examples of BLS applied to liquids, glasses, opaque materials (surface wave BLS), magnetic materials (magnons), and electronic levels are given.

Determination of Elastic Constants of a Bulk Crystal

Figure 2 shows a Brillouin spectrum obtained from a diamond crystal in the backscattering geometry shown in **Figure 1c**. Since the diamond surface was (1 1 1) and since the phonons probed in the experiment propagate along the surface normal, the peaks observed in **Figure 2** correspond to phonons along the [1 1 1] direction. Along this direction the longitudinal sound velocity is $v_L = \sqrt{(c_{11} + 2c_{12} + 4c_{44})/3\rho}$ and the velocity of the (degenerate) transverse phonons is $v_T = \sqrt{(c_{11} - c_{12} + c_{44})/3\rho}$. The peaks in **Figure 2** have therefore been labeled as L and T, respectively.

Since diamond crystals are large enough to be cut and polished with any desired orientation, it is clear that from a combination of measurements in different scattering geometries and/or crystal orientations the whole set of elastic constants can be determined. In **Table 1**, the c_{ij} of diamond determined by both Brillouin scattering and by ultrasonic techniques are listed. The results of both techniques are in very good agreement but it is interesting to note that while the BLS measurements were performed on crystals of around 1/3 carat, the ultrasonic measurements used a 22-carat diamond.

In cases where large, transparent crystals are available, it is clear from **Table 1** that BLS has no particular advantage over ultrasonic techniques. However, when only small samples are available the BLS technique becomes the technique of choice. Single crystals of cubic boron nitride (BN), for example, cannot be grown as single crystals larger than $\approx 500\,\mu m$. In this case, BLS becomes the technique of choice and yields the c_{ij} also included in **Table 1**.

The preceding section outlines how BLS can be used to determine c_{ij}. At ambient conditions, c_{ij} can often be determined more accurately using ultrasonic techniques. In some cases, however, BLS becomes a much more versatile tool than ultrasonics. Such cases usually involve samples where transducer bonding is difficult, as was mentioned above, in small or thin samples, high-pressure and/or high-temperature experiments, or cases where chemical reactivity becomes severe. For example, the temperature dependence of the c_{11} and c_{44} elastic constants of $Ba_{1-x}La_xF_{2+x}$ over the range 300–1600 K was determined by P E Ngoepe and J D Comins who found a linear decrease that they attributed to anharmonicity. A case involving high pressures is represented by measurements of liquid and solid argon in a diamond anvil cell up to 70 GPa.

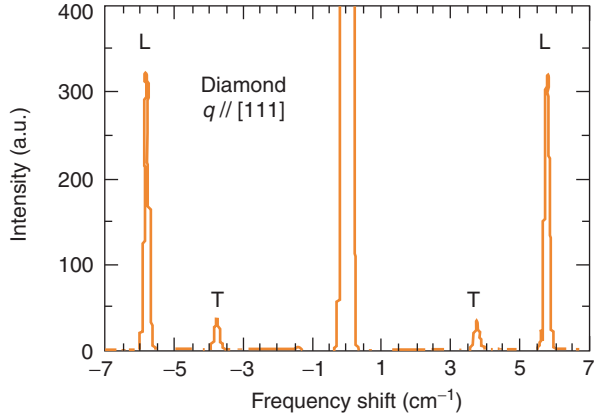


Figure 2 Brillouin spectrum from a diamond crystal in the geometry from **Figure 1c**. The longitudinal (L) and transverse (T) phonons propagate along the [1 1 1] direction.

Table 1 Elastic constants of diamond and boron nitride

Material and method	c_{11} (GPa)	c_{12} (GPa)	c_{44} (GPa)
Diamond, BLS	1076	125	577
Diamond, ultras	1079	124	578
Boron nitride, BLS	820	190	480

Measurements of Spectral Features in Liquids

BLS can also be used to determine the elastic response of liquids. In contrast to solids, however, where the unshifted peak is dominated by the effects of impurities and surface scattering, in pure liquids the central peak is due to entropy fluctuations. The intensity of the unshifted component is related to that of the shifted BLS peaks via the Landau–Placzek ratio, C_p/C_v–1, where C_p and C_v are the specific heats at constant pressure and volume, respectively. In pure liquids, therefore, BLS can be used to study the intensity and shape of the central peak. Frequency broadening of this peak can be caused by fluctuations in the orientation of anisotropically polarizable molecules or by diffusive processes governed by the equations of hydrodynamics. In liquids, it is also possible to investigate relaxation and damping effects via the phonon lifetime that translates into changes in the shape and line width of the phonon peaks. The same approach cannot usually be applied to most bulk crystals, because of a very low damping that results in line widths below the experimental resolution.

However, in temperature regions where structural phase transitions occur, damping processes influence Brillouin line widths enough so that the changes are detectable, for example, in molten and crystalline alkali halides, during order–disorder transitions in mixed crystals, or during ferroelastic phase transitions in some crystals.

Glasses

BLS has been extensively used to investigate the transition of liquid to solid in materials that do not crystallize: for example, silica, $Rb_{1-x}(NH_4)_x/H_2PO_4$ (10–300 K), $Rb_x(NH_4)_y/H_2PO_4$ (20–80 K), and $Ca(NO_3)/KNO_3$ (320–640 K). Across the transition region, the phonons in these materials become highly damped and the resulting line shapes can be interpreted in terms of the underlying mechanism of the glass transition. These types of investigations are also of practical importance, because these materials are often used in fiber optics technology and in optoelectronic devices.

Gases

BLS in gases detects local fluctuations that lead to frequency shifts in the 0.1–1 GHz range. These shifts can also be viewed as a Doppler shift from the motion of individual atoms. In this context, BLS has also been used to investigate how the phonon spectrum evolves as a gas is pressurized to become a fluid. This evolution reflects the hydrodynamic transformation as the thermal mean free path for atomic motion becomes comparable and then shorter than the wavelength at which the medium is being probed by the scattering of visible light. These experiments can also be interpreted in terms of the hydrodynamic equations of motion. Investigations of CO_2 up to 8 atm pressure, and Xe over the range 0.022–0.66 atm are two instances where this phenomenon has been investigated.

Surface Acoustic Modes in Thin Layers

Light scattering from acoustic excitations in opaque materials results from a purely surface-induced mechanism, viz. phonon-induced thermal ripples. In these cases, the scattering geometries (d) and (e) in **Figure 1** must be employed. In a semi-infinite medium, the strongest coupling is to Rayleigh waves (reminiscent of waves in the ocean) with some weaker features induced by bulk waves reflecting from the free surface. In thin films and layered structures, however, it is also possible to couple to modes that are localized within the layers or at the interfaces; these modes are generally referred to as Sezawa modes. If the medium is not completely opaque, it is possible to couple to the phonons via both the ripple and elastooptic mechanisms. In these cases, unusual interference effects are possible.

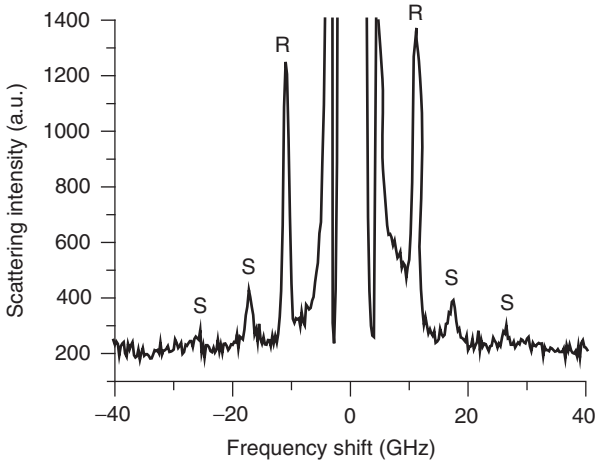


Figure 3 BLS from surface acoustic phonons in the Co/Cu superlattice. Descriptions: R – the Rayleigh surface mode, S – Sezawa or higher-order Rayleigh modes. The measurement was carried out for (*s*–*s*) polarizations of the incident and scattered light.

Figure 3 shows a BLS spectrum from a Co/Cu superlattice built from 60 bilayers of 1 nm Co/0.8 nm Cu in which the Rayleigh and Sezawa modes are observed. The experiment was carried out using a Sandercock-type tandem 3-pass interferometer. From these results, the effective elastic properties of the superlattice can be extracted.

In cases where the surface layers become much thinner than the wavelength of light, the medium again becomes equivalent to a homogeneous solid. In these cases, Brillouin scattering has been a very efficient and reliable tool for determining the effective elastic constants of the heterogeneous solid. For example, R Danner *et al.* in the Ni/V superlattice, for different layer thicknesses from the range of 1.85–63.3 nm, proved that this b.c.c./f.c.c. system reveals softening of elastic parameters. J A Bell and co-workers measured effective elastic constants in the Mo/Ta superlattice (c_{11}, c_{13}, c_{33}, c_{44}) with an elementary spatial bilayer thickness ranging from 0.7 to 20 nm.

Observation of Bulk and Surface Spin Waves in Metallic Superlattices

Brillouin scattering can also be used to investigate the magnetic properties of materials via their magnetic excitations – magnons. In this sense, the information it yields is similar to that obtained in ferromagnetic resonance (FMR) experiments. However, contrary to

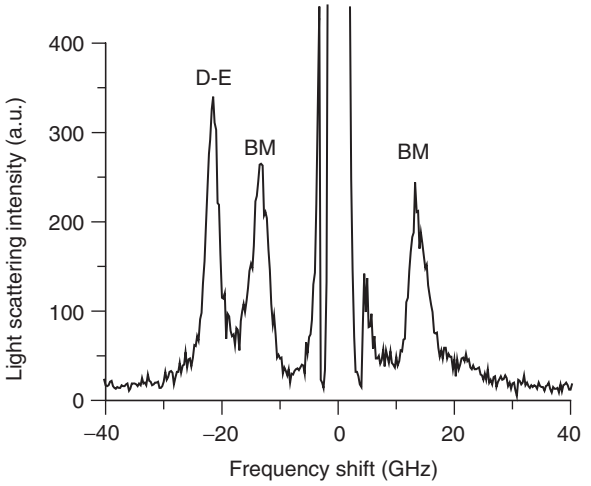


Figure 4 BLS from spin waves in a multilayered Co/Cu system. Descriptions: D-E, Damon–Eshbach surface mode, BM, (bulk magnon) collective excitations of a bulk-like nature resulting from coupling between spin waves through the nonmagnetic spacer layer. The measurement was carried out for (*s*–*p*) polarizations of the incident-scattered light.

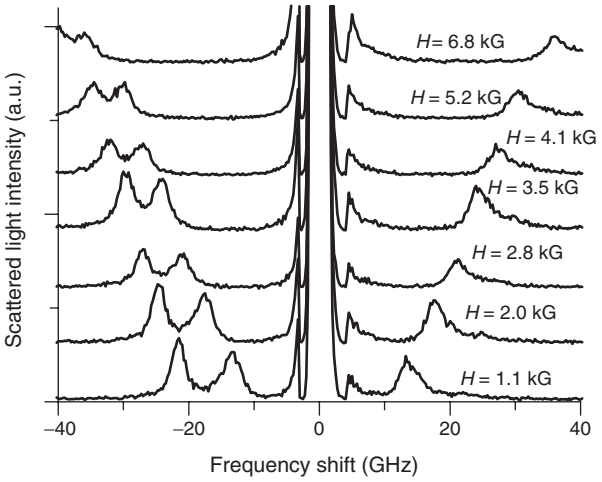


Figure 5 A set of Brillouin spectra measured for different magnetic field intensities in the Co/Cu superlattice. The peaks that appear only on the left side (Stokes signals) result from scattering from a unidirectional surface-like spin wave (D-E mode from **Figure 4**).

FMR that probes the infinite wavelength excitations, Brillouin scattering probes excitations with wavelengths comparable to the wavelength of light. As for phonons, magnons can be subdivided into surface-like and bulk-like excitations. Expressions relating the magnon frequencies to the magnetic properties of the material – magnetization, anisotropies, applied field, and gyromagnetic ratio – can be quite complex and often cannot be rendered in analytical form.

The Brillouin scattering technique has also been used to investigate the nature of magnetic excitations in nonconventional materials. In the case of the Co/Cu superlattice described in the previous paragraph, which is built from alternating magnetic and nonmagnetic materials, bulk-like collective excitations resulting from coupling between spin waves across the nonmagnetic spacer are detected (**Figure 4**). Additionally, a surface-like mode propagating on the surface of the sample was also found. The experiment was carried out with a tandem 3-pass spectrometer, with (*s*–*p*) polarization of the incident and scattered light, and a constant magnetic field lying in the sample plane. **Figure 5** provides results for different magnetic field intensities, from which the gyromagnetic ratio and saturation magnetization can be determined.

Scatterings from Electronics Levels

More recently, BLS has been used to probe the low-lying excitations of B impurity atoms in a diamond host . In this case, since the excitations are local and

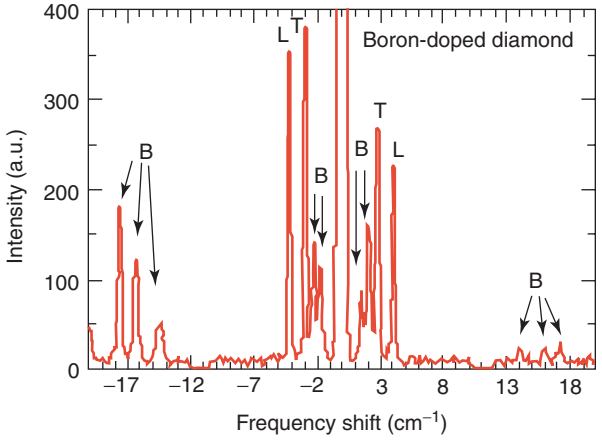


Figure 6 Brillouin scattering from B impurities in a diamond crystal as a host. Descriptions: (L) longitudinal phonons, (T) transverse phonons, and (B) electronic transitions between the hydrogen-like states surrounding the B atoms.

are not described by a wave vector, the results do not depend on the scattering geometry. **Figure 6** shows a spectrum recorded at low temperature and with an applied field of 4 T; apart from the expected longitudinal (L) and transverse (T) phonon peaks, the additional peaks labeled B can be traced to electronic transitions between the hydrogen-like states surrounding the B atoms. Under the effect of an applied field these modes split and from the measured splittings, the degeneracy and nature of the transitions is extracted. This, in turn, provides information on how B impurities in diamond lead to its semiconducting properties.

Acknowledgment

Work at ANL was supported by the US DOE, Basic Energy Sciences, Materials Sciences under contract W-31-109-ENG-38.

See also: Scattering, Inelastic: Raman; Thin Films, Mechanical Behavior of.

PACS: 43.35.G; 68.65.+g; 63.20.−e; 63.22.+m; 72.10.Di; 75.30.Ds; 78.35.+c; 81.07.−b

Further Reading

Bell JA, Bennett WR, Zanoni R, Sttegeman GI, Falco CM, *et al.* (1987) Elastic constants of Mo/Ta superlattices measured by Brillouin scattering. *Physical Review B* 35: 4127–4130.

Benedek GB, Lastovka JB, Fritsch K, and Greytak T (1964) Brillouin scattering in liquids and solids using low-power lasers. *Journal of the Optical Society of America* 54: 1284–1285.

Blachowicz T (2003) *Brillouin Spectroscopy in Crystal Lattices. Acoustic and Spin Waves*. Gliwice: Silesian University of Technology Press.

Brillouin L (1922) Diffusion de la Lumière et des Rayonnes X par un Corps Transparent Homogéne; Influence de l'Agitation Thérmique. *Annales de Physique* 17: 88.

Broz A, Harrigan M, Kasten R, and Monkiewicz A (1971) Light scattered from thermal fluctuations in gases. *The Journal of the Acoustical Society of America* 49: 950–953.

Danner R, Huebener RP, Chun CL, Grimsditch M, and Schuller IK (1986) Surface acoustic waves in Ni/V superlattices. *Physical Review B* 33: 3696–3701.

Grimsditch MH (1989) Brillouin scattering from metallic superlattices. In: Cardona M and Güntherodt G (eds.) *Light Scattering in Solids, Superlattices and Other Microstructures*, vol. 5, pp. 285–302. Berlin: Springer.

Grimsditch MH and Ramdas AK (1975) Brillouin scattering in diamond. *Physical Review B* 11: 3139–3148.

Grimsditch MH, Zouboulis ES, and Polian A (1994) Elastic constants of boron nitride. *Journal of Applied Physics* 76: 832–834.

Gross EF (1930) Change of wave-length of light due to elastic heat waves at scattering in liquids. *Nature* 126: 201.

Hyunjung K, Ramdas AK, Rodriguez S, Grimsditch MH, and Anthony TR (1999) Magnetospectroscopy of acceptors in "blue" diamonds. *Physical Review Letters* 83: 3254–3257.

Mandelstam LI (1926) K Voprosu O Rasseānii Sveta Neodnorodnoj Sredoj. *Zhurnal Russkovo Fiziko-Khimicheskovo Obshchestva* 58: 381.

Ngoepe PE and Comins JD (1986) Measurements of elastic constants in super-ionic $B_{1-x}La_xF_{2+x}$. *Journal of Physics C: Solid State Physics* 19: L267–L271.

Sandercock JR (1982) Trends in Brillouin scattering: studies of opaque materials, supported films, and central modes. In: Cardona M and Güntherodt G (eds.) *Light Scattering in Solids, Recent Results*, vol. 3, pp. 173–206. Berlin: Springer.

Shimizu H, Tashiro H, Kume T, and Sasaki S (2001) High-pressure elastic properties of solid Argon to 70 GPa. *Physical Review Letters* 86: 4568–4571.

Scattering, Inelastic: Electron

M Erbudak, Laboratorium für Festkörperphysik, Zürich, Switzerland
D D Vvedensky, Imperial College, London, UK

Introduction

Inelastic electron-scattering phenomena are used to obtain information on the electronic properties and atomic structure of materials. An interaction event between an electron and a target material is called inelastic if the electron loses some or all of its energy; otherwise, the event is called elastic. Any energy loss by the electron is equal to the energy gained by the material and is characteristic of the type of scattering event. The measurement of an inelastic event determines the energy transferred and the rate at which this transfer occurs, which provides physical information about the elementary excitations of the material.

Inelastic interactions can be grouped according to different energy-transfer regimes that form the basis of the principal experimental techniques. Historically, inelastic electron scattering was developed prior to the availability of continuous-energy photon sources from dedicated storage rings. In fact, under certain circumstances, electronic excitations are equivalent to photon excitations in that they can be described within a common conceptual and theoretical framework.

An electron is an elementary particle with a mass m and, owing to its wave nature, has a wave number $\boldsymbol{k} = \boldsymbol{p}/\hbar$, where $\boldsymbol{p} = m\boldsymbol{v}$ is the momentum, $\boldsymbol{v}$ is its velocity, and $\hbar$ is Planck's constant. The kinetic energy is expressed classically by $E = (1/2)mv^2$, where $v = |\boldsymbol{v}|$. Electrons have a negative charge e and a nonclassical angular momentum (spin). These properties play a central role in the interactions between electrons and their environment and constitute the basis of a wide range of spectroscopic techniques. Owing to their unique properties, there are both advantages and disadvantages in using electrons as a probe in analytical applications. Spin-polarized or spin-averaged beams of electrons are readily created, electrons can be accelerated to any energy, and focused on a target at will. After they interact with

the target material, electrons can be analyzed according to their intensity, momentum, energy, and spin polarization. In the final state, there are, even in an idealized single-scattering event, two electrons which arbitrarily share the initial energy. These electrons are indistinguishable, which makes the interpretation of the results especially challenging.

The Cross Section for Inelastic Scattering

Electron–atom scattering is the cradle of today's quantum mechanical understanding of many solid-state interactions. Electron–solid interactions, however, not only present a wealth of phenomena to aid the development of modern physics, but have also made spectroscopies based on electronic excitations one of the major sources of information on atomic and electronic structure and on elemental composition of materials. In order to account for all electronic interactions, a first-principles theory should be developed, though a semiclassical approach often suffices.

The strength of the electronic interaction depends on the electron energy and the material, which endows the energy dependence of a type of event with dynamical information about the interaction. If the interaction is weak, the energy $\hbar\omega$ lost by the electron and the associated momentum transfer $\boldsymbol{q}$ involve the creation of a single excitation with an energy $\hbar\omega$ and momentum $\boldsymbol{q}$. In this case, the Born approximation is valid, so the initial (i) and final (f) states of the electron are given by plane waves:

$$\psi_{\mathrm{i}} = \mathrm{e}^{\mathrm{i}\boldsymbol{k}_{\mathrm{i}}\cdot\boldsymbol{r}} \quad \text{and} \quad \psi_{\mathrm{f}} = f(\boldsymbol{k}_{\mathrm{i}})\mathrm{e}^{\mathrm{i}\boldsymbol{k}_{\mathrm{f}}\cdot\boldsymbol{r}} \qquad [1]$$

where $f(k_{\mathrm{i}})$ is the scattering amplitude. The interaction between the electron and the environment is described by a Hamiltonian containing the energy of the unperturbed sample, the energy of the electron, and the Coulomb potential $V(\boldsymbol{r})$ between the sample and the electron. During the interaction, energy and momentum are conserved: $E_{\mathrm{i}} - E_{\mathrm{f}} = \hbar\omega$ and $\boldsymbol{k}_{\mathrm{f}} - \boldsymbol{k}_{\mathrm{i}} = \boldsymbol{q}$. Using Fermi's golden rule, the scattering cross section can be expressed as

$$\mathrm{d}\omega = \frac{2\pi}{\hbar}|M^{E_{\mathrm{i}},\boldsymbol{k}_{\mathrm{i}}}_{E_{\mathrm{f}},\boldsymbol{k}_{\mathrm{f}}}|^2\delta(E_{\mathrm{i}} - E_{\mathrm{f}})\mathrm{d}^3\boldsymbol{k} \qquad [2]$$

where the matrix element $M = \langle \phi_{\mathrm{f}}\psi_{\mathrm{f}}|V(\boldsymbol{r})|\psi_{\mathrm{i}}\phi_{\mathrm{i}}\rangle$, and ϕ_{i} and ϕ_{f} are the wave functions of the target before and after the scattering.

When an electron scatters from an atom, the interaction only with the other electrons can be inelastic because no energy is transferred to the nucleus via Coulomb scattering. The most general type of energy transfer in electronic excitations produces a transition from an occupied to an unoccupied state separated by an energy equal to the energy loss of the primary electron. The differential cross section of such a transition is obtained by integrating eqn [2] over ω:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}E\,\mathrm{d}q} = 8\pi\left(\frac{e^2m^2}{\hbar^2}\right)^2\frac{1}{q^3k_{\mathrm{i}}^2}|\langle \phi_{\mathrm{f}}|\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}|\phi_{\mathrm{i}}\rangle|^2 \qquad [3]$$

where $q = |q|''$. The cross section $N(E) \propto \mathrm{d}\sigma/\mathrm{d}E$ is dominated by processes for which the momentum transfer is small. If $\boldsymbol{q}\cdot\boldsymbol{r} \ll 1$, one can write $\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \approx 1 - \mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}$, and the dipole approximation used in optics becomes applicable to inelastic electron scattering:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}E\,\mathrm{d}q} = 8\pi\left(\frac{e^2}{\hbar^2 v_i}\right)^2\frac{1}{q}|\langle \phi_{\mathrm{f}}|\boldsymbol{n}\cdot\boldsymbol{r}|\phi_{\mathrm{i}}\rangle|^2 \qquad [4]$$

where $v_{\mathrm{i}} = \hbar k_{\mathrm{i}}/m$, $\boldsymbol{q} = \boldsymbol{n}q$, and $\boldsymbol{n}$ is the unit vector along $\boldsymbol{q}$.

Interband Transitions

The validity of the dipole approximation means that the scattering cross section can be related to the dielectric properties of the target material. The dielectric response $\varepsilon(\boldsymbol{q},\omega) = \varepsilon_1(\boldsymbol{q},\omega) + \mathrm{i}\varepsilon_2(\boldsymbol{q},\omega)$ depends on the energy transfer $\hbar\omega$ and momentum transfer $\boldsymbol{q}$. Electrons interacting with the target produce a longitudinal perturbation over the long-range dipole fields. These fields attenuate the electron intensity through the absorptive part ε_2 of the dielectric constant. Since the field is screened by $1/\varepsilon$, the electron-loss function can be written as the sum of surface and bulk contributions:

$$\begin{aligned} N(E) &\propto -\operatorname{Im}\left[\frac{S}{1+\varepsilon(\omega,\boldsymbol{q})} + \frac{B}{\varepsilon(\omega,\boldsymbol{q})}\right] \\ &\equiv -\operatorname{Im}\left[\frac{1}{\tilde{\varepsilon}(\omega,\boldsymbol{q})}\right] \end{aligned} \qquad [5]$$

where S and B are weighting factors for the surface and the bulk, respectively. The cross section $N(E)$, therefore, contains fundamental information on the dielectric behavior of the bulk and the surface of the sample. A direct link between electron energy losses and the effective dielectric function $\tilde{\varepsilon}(\boldsymbol{q},\omega)$ is provided by the Kramers–Kronig relation,

$$\operatorname{Re}\left[\frac{1}{\tilde{\varepsilon}(\omega,\boldsymbol{q})}\right] = 1 - \frac{2}{\pi}\int_0^\infty \frac{\operatorname{Im}[-1/\tilde{\varepsilon}(\omega,\boldsymbol{q})]}{\omega'^2 - \omega^2}\omega'\,\mathrm{d}\omega' \qquad [6]$$

which gives

$$\tilde{\varepsilon}(\omega, \boldsymbol{q}) = \frac{\mathrm{Re}[1/\tilde{\varepsilon}(\omega, \boldsymbol{q})] - \mathrm{i}\,\mathrm{Im}[1/\tilde{\varepsilon}(\omega, \boldsymbol{q})]}{[1/\tilde{\varepsilon}(\omega, \boldsymbol{q})]^2} \qquad [7]$$

The practical advantage of energy-loss measurements over optics lies in the fact that they cover a larger energy range than experiments with photons, which facilitates a more complete evaluation of eqn [6]. The electron energy-loss function, in fact, depends on the joint density of states (DOS), defined as the convolution of the occupied initial states and the unoccupied final states that are separated by the energy loss and are coupled by the matrix element. The optical dielectric function contains the transverse properties of the solid, which may be different from the longitudinal properties accessible to electron-energy loss spectroscopy. The differences between the transverse and the longitudinal responses are negligible for a small momentum transfer $\boldsymbol{q}$ and become equal for $\boldsymbol{q} = 0$ (the random phase approximation). For higher values of $\boldsymbol{q}$, multipole transitions are well resolved in electron-induced transitions, but are strictly forbidden in optics.

Inelastic interactions are responsible for the intensity attenuation of electrons within a bulk material. The characteristic quantity for the penetration depth of the electrons is called the mean free path Λ, the value of which depends on the available excitations at a particular energy. For the range of energies $\sim$30–100 eV, Λ has values comparable to interatomic distances. Thus, the experiments contain information mainly on the surface properties (small B in eqn [5]). In optics, on the other hand, photons penetrate deep into the target material, so the information is dominated by the bulk DOS (small S in eqn [5]). Transitions that are characteristic of the surface (e.g., surface states, adsorbates) help to distinguish between bulk and surface excitations.

Plasmon Excitations

The cross section $N(E)$ in eqn [5] attains peak values for $\varepsilon_1 = 0$ or -1 and $\varepsilon_2 = 0$. This condition corresponds to charge-density fluctuations caused by the incident electrons. A longitudinal plasma wave along the crystal produces long-range Coulomb forces between positive and negative charges and excites collective oscillations. These are called plasmons in the case of a quasi-free electron gas. The plasmon energy is obtained from the Fourier modes of the electron density,

$$\rho(r) = \sum_{\boldsymbol{k}} \rho_{\boldsymbol{k}} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \qquad [8]$$

The $\rho_{\boldsymbol{k}}$ are amplitudes of harmonic density fluctuations obeying

$$\ddot{\rho}_{\boldsymbol{k}} + \omega_{\mathrm{p}}^2 \rho_{\boldsymbol{k}} = 0 \qquad [9]$$

in which ω_p is the Langmuir frequency

$$\omega_{\mathrm{p}} = \sqrt{\frac{n_e e^2}{\varepsilon_0 m}} \qquad [10]$$

n_e is the bulk charge density, ε_0 is the permittivity of free space, and the energy loss due to plasmon creation is $\hbar\omega_{\mathrm{p}}$. Plasmons are quantized excitations, so the energy loss is $n\hbar\omega_{\mathrm{p}}$, where n is an integer.

Plasmon energies can be used to extract changes in the electron density that accompany changes of the electronic properties or the volume of a solid. For example, the alloy Al–3 at.%Ag shows a transition from a high-temperature phase, where Al and Ag form a homogeneous solid solution, to a regime $\sim$400°C where Ag atoms precipitate as Ag_2Al clusters. The formation of these precipitates can be monitored by observing the plasmon energies of Al across the transition. With Ag atoms homogeneously distributed in the bulk, more volume is available to Al atoms, resulting in a lower effective electron density, thus shifting the plasmon energy to lower values. The spectra in **Figure 1** show that, at 480°C, the plasmon energy is reduced from 15.8 to 15.5 eV, signaling a 3.8% volume reduction. At temperatures below 440°C, Ag_2Al precipitates form that are largely decoupled from the aluminum matrix, thus leaving the volume and electron density unaltered from their values in pure Al. Thus, there is no change in the plasmon energy.

There are detailed investigations of the dependence of plasmon excitations on the momentum transfer. For increasing values of $\boldsymbol{q}$, the peak in the loss function moves to higher energies until the excitation decays into electron–hole pairs. In semiconductors, electron–hole pairs form a hydrogenic bound state called excitons which, for a material with an indirect bandgap, such as Si, have very long lifetimes because the recombination process requires the emission of a phonon to conserve momentum.

Surfaces have different charge distributions from those in the bulk. This moves the plasmon frequency to ω_{s}, where

$$\omega_{\mathrm{s}}^2(1 + \varepsilon) = \omega_{\mathrm{p}}^2 \qquad [11]$$

which, with $\varepsilon = 1$, yields

$$\omega_{\mathrm{s}} = \frac{1}{\sqrt{2}}\omega_{\mathrm{p}} \qquad [12]$$

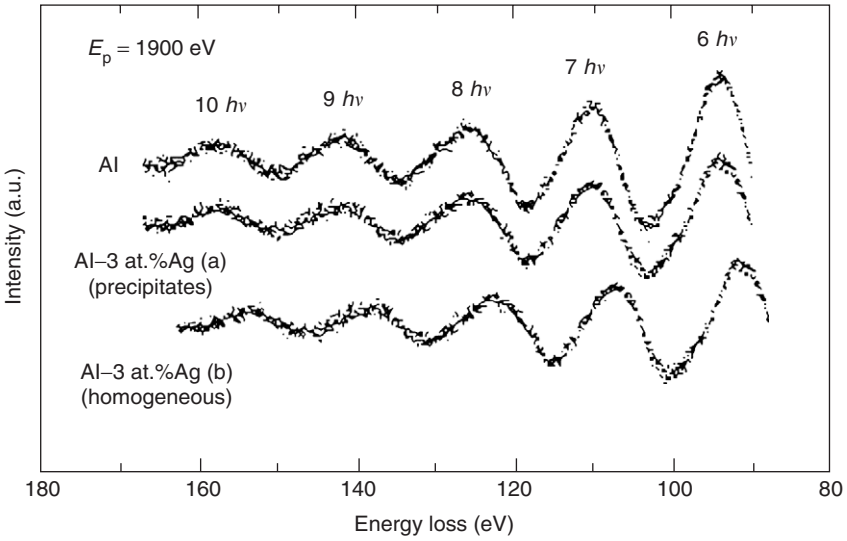


Figure 1 Energy losses in pure Al and Al–3 at.%Ag with (a) Ag_2Al precipitates, and (b) Ag atoms homogeneously distributed at 480°C. To accurately determine the plasmon energy, the measurements are extended up to the tenth plasmon loss. (Reprinted figure with permission from Wetli E, Erbudak M, and Schulthess T (1994) Segregation and dissolution of Ag-rich clusters at the (1 0 0) surface of Al–3 at.%Ag. *Physical Review B* 49: 14628–14631; © American Physical Society.)

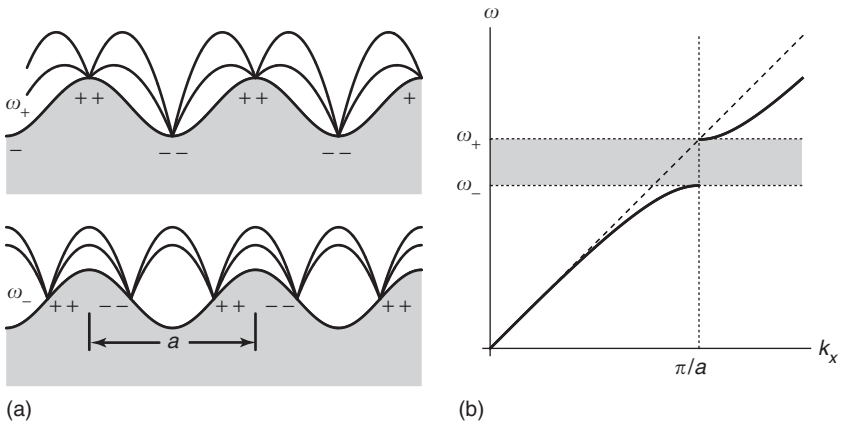


Figure 2 (a) Periodic textures on the metal surface that (b) lead to the formation of a photonic bandgap of surface plasmons. (Adapted from Barnes WL, Dereux A, and Ebbesen TW (2003) Surface plasmon subwavelength optics. *Nature* 424: 824–830.)

Surface plasmons are generated by the resonant interaction between the surface-charge oscillations and the electric field of the electrons. In electron energy-loss spectra, these plasmons reflect transitions associated with surface charges. In optics, they have potential uses in magnetooptic data storage. **Figure 2a** shows a corrugated metal surface with a period of a. If a corresponds to a surface-plasmon wavelength, there are two standing-wave solutions that have the same wavelength, but different energy, giving rise to a photonic bandgap (**Figure 2b**). At this point, surface plasmons cannot propagate, producing a significant increase in the electromagnetic field.

Vibrations at Surfaces

Electrons moving near the surface of a solid can excite localized and collective vibrations. The major experimental achievement in the measurement of these vibrations has been the development of near-monochromatic primary-electron beams and, additionally, the detection of reflected electrons with a resolution of the order of vibrational energies (a few meV). On Si(111), one of the most intensively studied surfaces, the local arrangement of atoms at and near the surface, including the stretching of Si–Si bonds, has been obtained from surface phonon

measurements and found to be consistent with the dimer-adatom-stacking-fault model.

Adsorbates are foreign atoms that are bound to sites with a particular point-group symmetry of the surface structure. If the dipole image charge of the adsorbate has a component normal to the surface, the adsorbate can be excited into vibrational states, while modes with their dipole moment parallel to the surface are screened by a factor $1/(\varepsilon+1)^2$. These selection rules are similar to those governing infrared transitions. One can determine the local symmetry by observing the number of vibrational modes. The local geometry is modeled by a rigid substrate against which the adsorbates vibrate. **Figure 3** shows some fundamental types of bonding arrangements. In (a), an adsorbate atom is placed at a top position. This site would produce a single low-energy vibrational mode. In (b), the same atom is shown in a bridge position. Energy loss would be observed at an even lower value than (a), corresponding to the normal vibrational component. If the chemical bonding were not symmetrical, then one would observe two peaks at similar energy losses. Generally, a lowering of the adsorption symmetry, such as uneven bonding, increases the number of vibrational modes. For molecular adsorbates (**Figures 3c** and **3d**), a high-energy loss would be due to the stretching mode of the atoms in the molecule, while a low-energy loss would be observed, due either to the vibration of each atom against the substrate, or the entire molecule. Hence, for molecular adsorption one always expects at least two modes. Observation of a single mode is, therefore, evidence for the dissociative adsorption of a molecule. Apart from the number of observed modes, the vibrational energy contains information about the chemical nature of the bond. An adsorption site on top of a substrate atom results in a higher force constant compared to an adsorbate at a bridge or hollow site, where the bond strength is shared by several surface atoms.

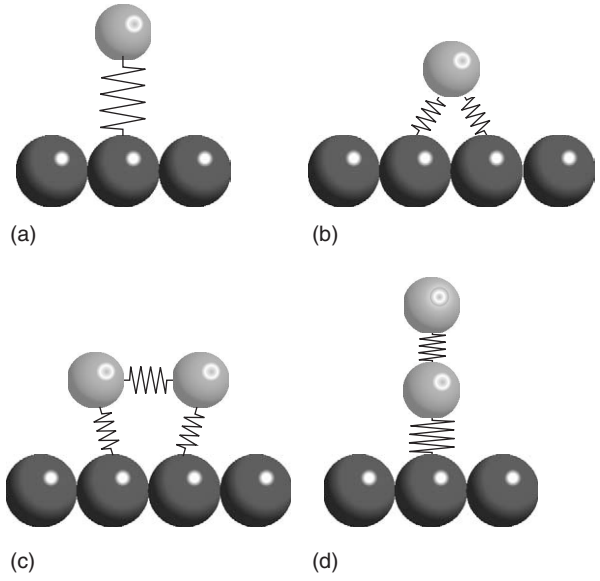


Figure 3 A schematic description of four typical adsorption states. The springs represent chemical bonds of different strengths.

Excitation of Core Electrons

A special case of interband transitions is the excitation of core electrons with a binding energy, E_B. The primary electron can transfer any amount of energy $E_L \leqslant E_i$ to the core electron. By measuring the energy loss of the primary electron, one knows the kinetic energy of the core electron after scattering, that is, how far above E_F the core electron has been excited. The excitation into states at E_F corresponds to zero kinetic energy, and hence, $E_L = E_B$. This is called the absorption edge. It is characteristic of a particular element and its chemical state, and can thereby be used as a tool for chemical analysis.

Nanotubes provide a topical illustration of such measurements. Their chemical and physical properties can be manipulated by incorporating atoms or molecules within the tubes. **Figure 4** shows a high-resolution electron micrograph of an isolated single-wall nanotube containing C_{82} fullerenes and a schematic representation of the structure. The image reveals single Gd atoms encapsulated in the fullerene cages. **Figure 5** displays a spatially resolved electron-energy-loss spectrum of C *K*-edge (1*s* initial state) and Gd *N*-edge (4*d* initial state) absorption edges. Thus, it is possible to detect and identify single Gd atoms encaged in a complex C environment. Core-electron binding energies may show a chemical shift depending on the chemical state of the element because,

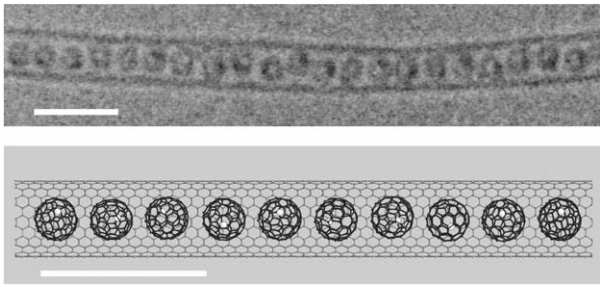


Figure 4 Electron micrograph of an isolated nanotube incorporated with a chain of Gd-metallofullerenes and a schematic representation of the structure. Horizontal bars represent a length of 3 nm. (Reprinted figure with permission from Suenaga K, Tencé M, Mory C, Colliex C, Kato H, *et al.* (2000) Element-selective single atom imaging. *Science* 290: 2280–2282; © AAAS.)

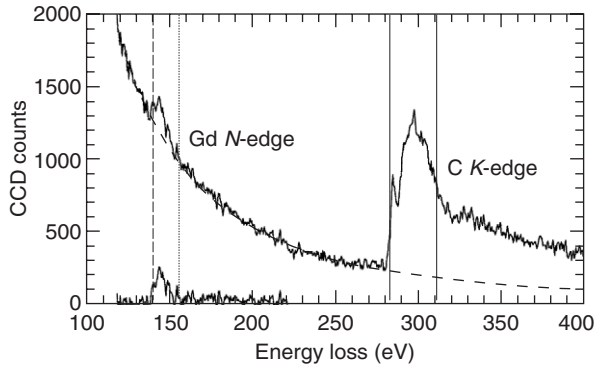


Figure 5 Electron-energy-loss spectroscopy obtained from an isolated nanotube incorporated with a chain of Gd-metallofullerenes. (Reprinted figure with permission from Suenaga K, Tencé M, Mory C, Colliex C, Koto H, *et al.* (2000) Element-selective single atom imaging. *Science* 290: 2280–2282; © AAAS.)

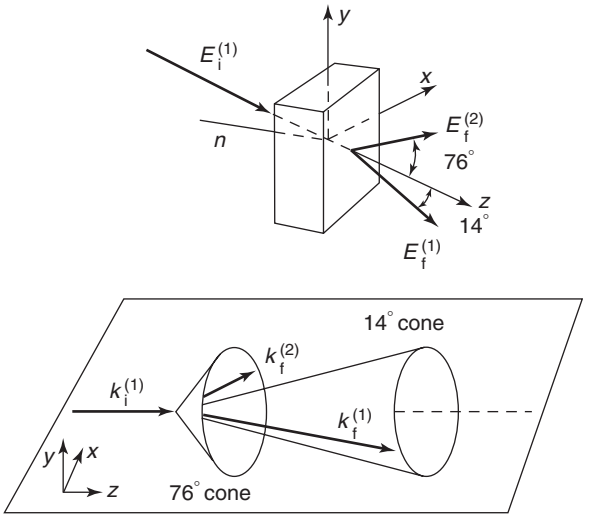


Figure 6 A schematic drawing illustrating the energy conservation in an $(e^-, 2e^-)$ experiment. (Adapted from Cai YQ, Vos M, Stoner P, Kheifets AS, MacCarthy IE, *et al.* (1995) Direct imaging of the valence electronic structure of solids by (*e*, 2*e*) spectroscopy. *Solid State Communications* 95: 25–29.)

through chemical binding, the electronic charge density is redistributed, thereby shifting the binding energy. The core-electron binding energy of Gd shows a downward shift compared to the neutral Gd metal, which suggests a charge transfer from the Gd to a C atom.

Core states have almost no overlap between neighboring atoms, leading to extremely narrow bandwidths. Therefore, the energy-loss spectrum at and above the absorption edge is indicative of the unoccupied DOS at the site of the excited atom with an orbital angular momentum determined by the symmetry of the final state. If the final states are far above E_F, the modulations in the spectra can be described within a single-scattering framework and used to extract structural information about the local environment of the excited atom. This is known as extended energy-loss fine structure (EELFS); the optical analog is extended X-ray-absorption fine structure (EXAFS). Closer to the absorption edge, the interpretation of the spectra requires a multiple-scattering treatment, as in near-edge X-ray-absorption fine-structure (NEXAFS) in optics.

e^-2e^- Scattering

A general, yet simple, case of electron–atom or electron–solid interactions is considered. The target has electrons of initial energy $E_i^{(2)}$ and momentum $\boldsymbol{k}_i^{(2)}$, while the primary-electron energy has $E_i^{(1)}$ and momentum $\boldsymbol{k}_i^{(1)}$. In the final state, there are two electrons: one with an energy $E_f^{(1)}$ and momentum $\boldsymbol{k}_f^{(1)}$, and the initially bound electron, now with energy $E_f^{(2)}$ and momentum $\boldsymbol{k}_f^{(2)}$. In the limit of large energy and momentum transfer, the $E_f^{(1)}$ and $E_f^{(2)}$ are well above the band structure regime, so the unoccupied DOS is parabolic and does not impose additional structure upon the measurement. This interaction regime is the basis of $(e^-, 2e^-)$ spectroscopy for solids. For atomic or molecular targets, $(e^-, 2e^-)$ spectroscopy provides information about the wave function density in momentum space by measuring the momentum distribution of the bound electron as a function of the momentum transfer. For crystalline targets, both the momentum density and the energy dispersion of valence electrons can be determined. Since the momentum of any ejected electron can be measured, the results are not confined to the first Brillouin zone.

In a binary collision between a high-energy electron and a band electron, the conservation of energy requires that $E_i^{(1)} = E_f^{(1)} + E_f^{(2)} + E_i^{(2)}$ and the conservation of momentum is given by $\boldsymbol{k}_i^{(1)} = \boldsymbol{k}_f^{(1)} + \boldsymbol{k}_f^{(2)} + \boldsymbol{k}_i^{(2)}$. There is, however, a continuous spectrum of single- and multiple-scattering events that leads to $E_f^{(1)} + E_f^{(2)}$. To single out the event for which both equations hold, the particles must be recorded in coincidence.

Figure 6a shows the geometry of primary electrons impinging upon a solid sample. A binary collision event is depicted in which a target electron acquires energy and is ejected from the back of the sample together with the primary electron. For simplicity, the scattering event is shown with coplanar initial and final wave vectors. In general, the 2π azimuthal distribution relative to the primary-electron incidence is measured to access the details of the initial

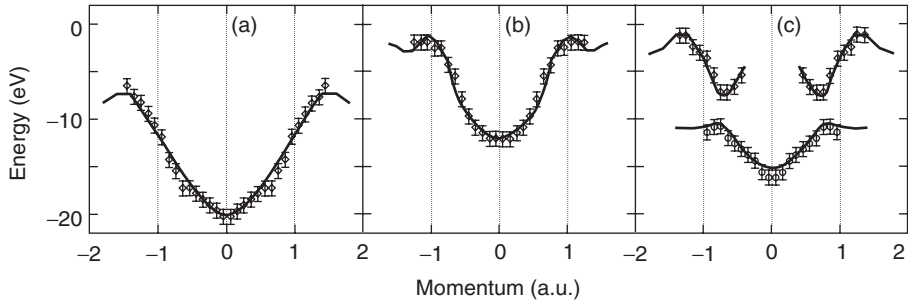


Figure 7 Measured (markers) and calculated (solid line) energy-band dispersions of (a) the amorphous C, (b) the amorphous Si, and (c) the polycrystalline SiC. (From Cai YQ, Vos M, Stoner P, Kheifets AS, MacCarthy IE, *et al.* (1995) Direct imaging of the valence electronic structure of solids by (*e*, 2*e*) spectroscopy. S*olid State Communications* 95: 25–29.)

state, as shown in **Figure 6b**. Energy conservation requires that the binary-collision event takes place as close to the surface as possible, so inelastic interactions after the indicated principal collision are not considered. Consequently, the spectroscopic results are characteristic of the near-surface region of the target material.

Figure 7 shows an energy-band dispersion in amorphous C and Si and polycrystalline SiC obtained in an $(e^-, 2e^-)$ experiment. The results compare favorably with the computational band structure spherically averaged over the cubic Brillouin zone. Thus, momentum densities determined in $(e^-, 2e^-)$ spectroscopy are associated with local order and energy-band dispersion exists in disordered and amorphous systems.

See also: Dielectric Function; Electron Gas (Theory); Electron–Phonon Interactions and the Response of Polarons; Excitons: Theory; Lattice Dynamics: Vibrational Modes; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Metals and Metallic Alloys, Optical Properties of; Optical Absorption and Reflectance; Optical Properties of Materials; Optical Sum Rules and Kramers–Kronig Relations; Plasmons; Surfaces, Optical Properties of; X-Ray Absorption Spectroscopy.

PACS: 79.20.Uv; 77.22.−d; 71.45.Gm; 73.20.Mf; 03.65.Nk

Further Reading

De Crescenzi M and Piancastelli MN (1996) *Electron Scattering and Related Spectroscopies*. Singapore: World Scientific.

Ibach H and Mills DL (1982) *Electron Energy Loss Spectroscopy and Surface Vibrations*. New York: Academic Press.

McCarthy IE and Weigold E (1991) Electron momentum spectroscopy of atoms and molecules. *Reports of Progress in Physics*. 54: 789–879.

Messiah A (1958) *Quantum Mechanics*, vols. 1 and 2. New York: Wiley.

Mott NF and Massey HSW (1965) *The Theory of Atomic Collisions*. Oxford: Clarendon.

Pines D (1965) *Elementary Excitations in Solids*. New York: Benjamin.

Platzman PM and Wolff PW (1973) *Waves and Interactions in Solid State Plasmas*. New York: Academic Press.

Prinz R and Koningsberger K (1986) *X-Ray Absorption: Principles, Applications and Techniques of EXAFS, SEXAFS and XANES*. New York: Wiley.

Raether H (1965) Solid state excitations by electrons. In: Höhler G (ed.) *Springer Tracts in Modern Physics*, vol. 38, pp. 85–157. Berlin: Springer.

Schattschneider P and Jouffrey B (1995) Plasmons and related excitations. In: Reimer L (ed.) *Energy-Filtering Transmission Electron Microscopy*, Springer Series in Optical Sciences, vol. 71, pp. 151–224. Berlin: Springer.

Schnatterly SE (1979) Inelastic electron scattering spectroscopy. In: Ehrenreich H, Seitz F, and Turnbull D (eds.) *Solid State Physics*, vol. 34. New York: Academic Press.

Vvedensky DD (1992) Theory of X-ray absorption fine structure. In: Fuggle JC and Inglesfield JE (eds.) *Unoccupied Electronic States*, Topics in Applied Physics, vol. 69. pp. 139–201. Berlin: Springer.

Wooten F (1972) *Optical Properties of Solids*. New York: Academic Press.

Scattering, Inelastic: Raman

P Lemmens, Max Planck Institute for Solid State Research, Stuttgart, Germany
K-Y Choi, Tohoku University, Sendai, Japan

Introduction

A light scattering experiment is sketched in **Figure 1a**. Monochromatic light with frequency ω_I, polarization $\boldsymbol{e}_I$, and the external electric field $\boldsymbol{E}_I(\omega, \boldsymbol{r})$ is incident on the medium and induces a polarization $\boldsymbol{P}(\omega, \boldsymbol{r})$. The medium scatters light of spectral density $\rho(\boldsymbol{q}, \omega)$ with frequency ω_S and polarization $\boldsymbol{e}_S$. This dipole-like radiation scales in intensity $I \propto \omega_S^4$. The Raman or combination scattering effect was discovered in 1928 independently by C V Raman, G S Landsberg, and L I Mandel'shtam. Raman scattering (RS) is analyzed as a function of the frequency difference or Raman shift $\omega = \omega_I - \omega_S$ with respect to the incident light and registered using a sensitive detector. Only 10^{-7}–10^{-9} of the incident photons are scattered inelastically. If the Raman shift is negligible, $\omega_S \approx \omega_I$, the process is elastic and called Rayleigh scattering, while for $\omega_S < \omega_I$ or $\omega_S > \omega_I$, it is denoted as Stokes scattering or anti-Stokes scattering, respectively. The ratio of anti-Stokes to Stokes intensities is strongly temperature dependent:

$$I_{AS}/I_{ST}(\Delta\omega, T) = \left(\frac{\omega_I + \omega}{\omega_I - \omega}\right)^4 \exp\left(-\frac{\hbar\omega}{k_B T}\right) \quad [1]$$

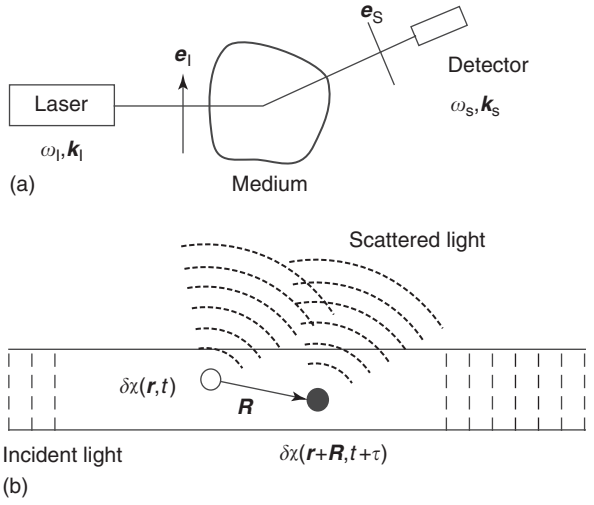


Figure 1 (a) Schematic description of light scattering geometry. (b) Interference of secondary waves from different points inside the incident beam.

The induced polarization $\boldsymbol{P}$ is related to the electronic susceptibility tensor $\chi(\omega)$, a material specific quantity, using the response relation $\boldsymbol{P}(\omega, \boldsymbol{r}) = \varepsilon_0 \chi(\omega) \boldsymbol{E}(\omega, \boldsymbol{r})$. In a homogeneous medium, χ is constant in space and time. Secondary waves from different points inside the incident beam cancel each other out according to the Huygens–Fresnel principle (see **Figure 1b**). The scattering process is therefore a result of spatial and temporal fluctuations of χ in the medium. The spectral density of light scattering is given by a correlation function of the fluctuations:

$$\rho(q, \omega) \sim \int \langle \delta\chi^*(q, t) \delta\chi(q, 0) \rangle e^{i\omega t} \, dt \quad [2]$$

with $\boldsymbol{q} = \boldsymbol{k}_I - \boldsymbol{k}_S$ and $\omega = \omega_I - \omega_S$, the change of wave vector and wave number according to their conservation in the scattering process. The fluctuation spectrum $\chi(\omega)$ in solids originates either from density fluctuations or from internal degrees of freedom, as phonons. The former leads to a weak scattering continuum that quickly decreases in intensity as a function of the Raman shift. The corresponding energy transfer from the electromagnetic field to low-energy electronic states has no well-defined eigenfrequency and this response is usually screened. The latter fluctuation contribution leads to peaks in $\rho(\omega)$ at frequencies corresponding to the energy of excitations.

Typical Raman shifts of phonons in solids range from 50 to 1000 cm^{-1}, while molecular vibrations may range up to 3500 cm^{-1}. The unit wave number (cm^{-1}) corresponds to the difference of the inverse of the corresponding wavelengths λ in centimeters and can be converted using 1 (cm^{-1}) $\approx$ 1.44 (K) $\approx$ 1/8 (meV) $\approx$ 33 (GHz) to other units used in solid-state spectroscopy.

Phonon wave vectors over a major part of the Brillouin zone are two to three orders of magnitude larger than the wave vectors of electromagnetic radiation in the visible range. Therefore, the kinematics of the scattering process restricts phonons scattering to small wave vectors close to the origin of the Brillouin zone. Acoustic modes with small wave vectors and corresponding to very small energies can only be detected in Brillouin scattering setups based on interferometers. The conservation of momentum may be lifted at surfaces, defects, in the case of multiparticle scattering, and if nonadiabatic scattering is involved. In metals, the very small penetration depth of light may lead to the observation of excitations with a broad range of momenta.

Classical Theory

RS originates from radiating dipole moments induced by an incident electric field. In a crystal, the electric field of a plane, monochromatic wave with angular frequency ω_I is given by $\boldsymbol{E}_I(\omega, \boldsymbol{r}) = \boldsymbol{E}_I \cos \omega_I t$. In general, the polarization of the crystal is a function of the instantaneous positions of the atoms in the lattice. Thus, the susceptibility can be expanded in terms of normal coordinates of the crystal:

$$\chi = \chi_0 + \sum_a \left(\frac{\partial \chi}{\partial Q_a}\right)_0 Q_a + \frac{1}{2} \sum_{a,a'} \left(\frac{\partial^2 \chi}{\partial Q_a \partial Q_{a'}}\right)_0 Q_a Q_{a'} + \cdots \quad [3]$$

The linear term in Q_a leads to first-order scattering, while the quadratic and higher terms are responsible for second- and higher–order effects. The laser power used in Raman experiments is small and the induced atomic displacements do not exceed 10^{-10} m. Therefore, higher-order derivatives that, for example, lead to hyper-RS at multiples of ω_I can be neglected. Since the lattice waves are described by $Q_a = Q_0 \cos \omega_p t$ for each normal coordinate, the induced polarization becomes

$$\boldsymbol{P} = \varepsilon_0 \left[\chi_0 + \sum_a \left(\frac{\partial \chi}{\partial Q_a}\right)_0 Q_0 \cos \omega_p t\right] \times \boldsymbol{E}_I \cos \omega_I t \quad [4]$$

where $\varepsilon_0 \chi_0 \boldsymbol{E}_I \cos \omega_I t$ corresponds to elastic scattering. The inelastic terms of the polarization can be transformed to

$$\begin{aligned} \boldsymbol{P}_i &= \boldsymbol{A} \cos \omega_p t \cos \omega_I t \\ &= \frac{\boldsymbol{A}}{2} [\cos(\omega_I + \omega_p)t + \cos(\omega_I - \omega_p)t] \\ &\text{with } \boldsymbol{A} = \varepsilon_0 \sum_a \left(\frac{\partial \chi}{\partial Q_a}\right)_0 Q_0 \boldsymbol{E}_I \end{aligned} \quad [5]$$

The frequency shift of $\omega_I + \omega_p$ and $\omega_I - \omega_p$ corresponds to anti-Stokes and Stokes scattering, respectively. The former (latter) process can be interpreted as the destruction (creation) of elementary excitations in the solid. The intensity of the inelastically scattered light is given by the amplitude factor $\boldsymbol{A}$ and especially by the derivative $\sum_a (\partial \chi / \partial Q_a)_0$. It follows that a phonon is Raman active only if the first derivative of the polarizability with respect to the phonon displacement is nonzero. This derivative depends on crystal symmetry and scattering configuration. The latter is given by the polarization of the incident and scattered light, $\boldsymbol{e}_I$ and $\boldsymbol{e}_S$, with respect to the crystal structure.

In crystals where the corresponding point group has an inversion center, a phonon mode can be either IR or Raman active. The Raman active modes are "gerade" with respect to an inversion. Furthermore, using the group theory, the number of Raman-active phonons and their selection rules (measurement geometry) can be derived. This scheme consists of determining the site symmetry of each atom in the unit cell of the investigated compound, finding the corresponding irreducible representations (phonons), and subtracting from the sum of these representations the acoustic-, silent-, and infrared-active modes. These steps and the needed tables are documented in literature. A calculation or estimation of the optical phonon frequencies can be performed using shell lattice models or a comparison with known compounds.

The modes are classified with respect to the dimensionality of their representations as 1D (A_{1g} or B_{1g}), 2D (E_{2g}), 3D (T_{1g} or F_{1g}). A and B distinguish a positive/negative value of χ with respect to rotation about the principal axis. The possible indices are g for gerade, u for ungerade with respect to the inversion center, and the numbers 2 and 3 are used to denote a symmetry with respect to other rotations or with respect to reflections, respectively. An irreducible representation of Raman-active modes may then look like the following:

$$\Gamma_{\text{Raman}} = A_{1g} + E_{1g} + 3B_{1g} \quad [6]$$

This means that, in total, five phonon modes would be expected. Experimentally, the phonon modes can be differentiated using the symmetry properties of their Raman tensor. For example, modes with A_{1g} observed with $(\boldsymbol{e}_I, \boldsymbol{e}_S)$ both parallel to one of the crystallographic directions, while, for example, B_{1g} are observed with $\boldsymbol{e}_I$ perpendicular to $\boldsymbol{e}_S$ and both polarizations turned by 45° within a crystallographic plane. These scattering geometries are abbreviated as $z(xx)\text{-}z$ and $z(x'y')\text{-}z$, with x, y, z parallel to the crystallographic axes and x' turned by 45° within the xy-plane. In this experiment the wave vectors $\boldsymbol{k}_I, \boldsymbol{k}_S$ are parallel to the z-axis. As an example, the matrix representations of a tetragonal group, D_{4h}, is used:

$$A_{1g}: \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}; \quad B_{1g}: \begin{pmatrix} c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \text{and} \quad E_g: \begin{pmatrix} 0 & 0 & d \\ 0 & 0 & e \\ d & e & 0 \end{pmatrix} \quad [7]$$

Application of Raman Scattering and Experimental Setup

Application of Raman Scattering

The most fruitful application of RS is certainly the investigation of phonons in semiconductors or insulators. The penetration depth of light in the visible range is large and the phonon scattering cross section is also normally high. **Figure 2** shows a measurement in Si with a phonon at $520\,\mathrm{cm}^{-1}$, and two-phonon scattering at $950\,\mathrm{cm}^{-1}$. Due to surface defects, a phonon density of states (DOS) is observed with a very small intensity for $\Delta\omega < 1000\ \mathrm{cm}^{-1}$.

RS may furthermore be used for the characterization of strain, interface properties, and the control of epitaxy and growth processes of thin films. Frequently, RS is applied to determine the chemical composition and doping profiles with a topographical resolution using scanning stages. This makes use of the fact that phonon modes of the respective materials sensitively depend on composition. The spacial resolution of such investigations is limited by the focus diameter, which is of the order of the light wavelength. An improvement can be gained if optical near-field techniques are used.

The strong dependence of light penetration depth on the frequency of the incident light can be used to determine depth profiles. Furthermore, the RS cross section can be enhanced using electronic resonances in the bulk (resonance scattering) or on the surface (surface-enhanced RS) of the medium. Sensitivity of single molecules can be achieved.

Experimental Setup

Setups for inelastic light scattering experiments can be optimized either for high-resolution and quasi-elastic light rejection (**Figure 3**), or for small sample and surface characterization with a microscope. They contain the following functional elements:

Monochromatic light source and beam tailoring Noble gases (helium/neon, argon, krypton) ion lasers or frequency-doubled laser diodes are used as a light source. Spurious contributions at frequency different from the desired one (plasma lines) are filtered by a prism double monochromator or a notch filter. The polarization and ellipticity of the laser beam are adjusted by transmitting through $\lambda/4$ and $\lambda/2$ wave plates. The beam is then focused on the sample. A typical light power level on the sample is 4–20 mW on a focus diameter of 50–100 μm or a smaller power, if a Raman microscope with a focus of 1–5 μm is used.

Sample and scattering configuration A backscattering configuration is frequently used as solids are usually opaque or semi-opaque. A small canting of the sample with respect to the incoming laser beam suppresses quasi-elastic light. RS allows one to change the sample environment or temperature easily using a cryostat, pressure cell, or glass fibers, as the Raman shift of the scattered light is small with respect to the used excitation frequency, $\Delta\omega/\omega_\mathrm{I} \approx 10^{-2}$. Therefore, only one set of window material (fused silica) is needed. The scattered light is collected with a lens, for example, a photographic camera lens, and filtered

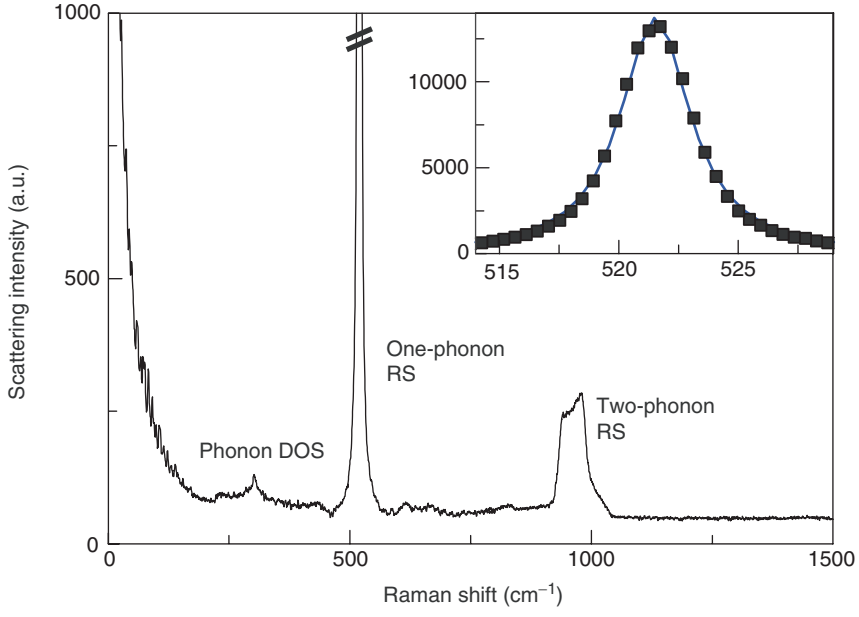


Figure 2 Raman scattering in silicon in (*xx*) polarization including one-phonon and two-phonon scattering. The inset shows a Gaussian fit (solid line) to the one-phonon scattering data (points).

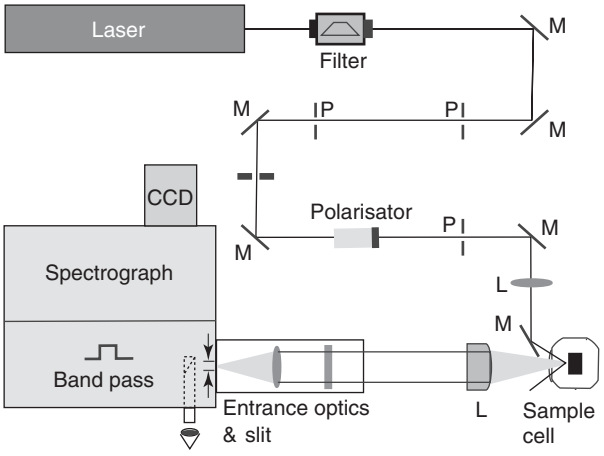


Figure 3 RS setup in quasi-backscattering geometry with a spectrometer, mirrors (M), pinholes (P), and lenses (L).

with respect to polarization. The light is then focused on the entrance slit of a monochromator. It is convenient to use an eyepiece or a camera with a view on the sample from behind the open entrance slit as an aid for the optical alignment. The definition of the focus can be improved using confocal entrance optics (spacial filtering).

Monochromator and detection Quasi-elastic light due to surface roughness or Raleigh scattering is suppressed by an interference filter or by a subtractive double monochromator used as a bandpass. The transmitted light is then dispersed on the detector with a single monochromator stage with blazed, holographic gratings used as dispersing elements. The detection of light is usually performed by a charge coupled device (CCD). This array detector allows the simultaneous detection of a broad frequency band with single-photon sensitivity. Sensors with quantum efficiencies of 80% are available. Typical accumulation times may range from 10 s to 5 min. Several of these measurements are averaged to further reduce noise and to eliminate spikes induced by cosmic muons in the detector.

Scattering Cross Section and Quantum Mechanics

The inelastic scattering process by a crystal involves incident photons of energy $\hbar\omega_I$ and momentum $\hbar\boldsymbol{k}_I$ leading to photons of energy $\hbar\omega_S$ and momentum $\hbar\boldsymbol{k}_S$. At the same time, an initial state of the crystal, i, with energy E_i transits into a final state, f, with energy E_f. The probability of scattering, P_{if}, can be formulated with the help of Fermi's golden rule of the second-order perturbation theory as

$$P_{if} = \frac{2\pi}{\hbar}\rho_{\mathrm{dos}}(\omega_S)|H(\boldsymbol{k}_I, i : \boldsymbol{k}_S, f)|^2 \times \delta(\hbar\omega_I + E_i - \hbar\omega_S - E_f) \quad [8]$$

where $\rho_{\mathrm{dos}}(\omega_S)$ is the DOS of photons, which is given by $\rho(\omega_S)\,\mathrm{d}\omega_S\,\mathrm{d}\Omega = (1/2\pi c)^3\omega_S^2\,\mathrm{d}\omega_S\,\mathrm{d}\Omega/\hbar$. Here, $\mathrm{d}\Omega$ is a solid angle. The main difficulty is to determine the Hamiltonian, $H(\boldsymbol{k}_I, i : \boldsymbol{k}_S, f)$, which describes the interaction between the electromagnetic radiation and the crystal. Taking into account that the wavelength of the incident/scattered light is much larger than the sizes of atoms, the effective Hamiltonian can be written in terms of the polarizations as

$$H(\boldsymbol{k}_I, i : \boldsymbol{k}_S, f) = \sum_{r,\alpha,\beta} E_I^{\alpha}\chi(r)^{\alpha\beta}E_S^{\beta} \quad [9]$$

Here, α and β denote Cartesian coordinates. The electric field is described in terms of the photon creation (annihilation) operator, $b^{\dagger}(k)$ $[b(k)]$:

$$\boldsymbol{E}_I = i\sqrt{2\pi\hbar\omega_I}E_I\boldsymbol{e}_I[b(\boldsymbol{k}_I) - b^{\dagger}(-\boldsymbol{k}_I)] \quad [10]$$

where $\boldsymbol{e}_I$ is the polarization vector of the photon. Using these results, one can obtain the total RS intensity

$$I_R \sim \omega_I\omega_S^3 \int \ll\chi(\boldsymbol{k}_I\boldsymbol{k}_S; t)\chi^{\dagger}(\boldsymbol{k}_I\boldsymbol{k}_S; 0)\gg \mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t \quad [11]$$

The brackets $\ll\cdots\gg$ denote a thermodynamic average, and $\omega = \omega_I - \omega_S$. Since $\omega_I \approx \omega_S$, the intensity has an approximate form of the dipole radiation as mentioned above. It is important to note that the RS involves a pair correlation functions of the fluctuating polarization. This microscopic derivation confirms the classical picture discussed in the section "Classical theory." Note that an explicit form of the susceptibility can be provided by the dipole matrix elements $\langle f|\boldsymbol{M}_\beta|i\rangle$ with the moments $\boldsymbol{M}_\alpha = e r_\alpha$ between the initial and final states with $\omega_{fi} = [E_f - E_i]/\hbar$:

$$\chi_{\alpha\beta}(\omega) = 2\sum_{i,f}\int \langle i|\boldsymbol{M}_\alpha|f\rangle\langle f|\boldsymbol{M}_\beta|i\rangle \times \left(\frac{1}{\omega_{fi} - \omega} + \frac{1}{\omega_{fi} + \omega}\right)\frac{\mathrm{d}^3k}{(2\pi)^3} \quad [12]$$

The factor 2 is due to the spin.

Phonon Scattering

The RS intensity of solids is usually dominated by phonon modes. In an adiabatic approximation, an optical phonon $v(= 1, \ldots, 3N_c)$, with N_c the number

of ions per unit cell, consists of atomic displacements with the eigenvector Q_v that can be treated as a very slow ionic displacement leading to a static potential within the unit cell $\delta U_v(r)$. This slow varying potential will (1) shift the energy of the initial state by $\delta E_{i,v} = \langle i|\delta U_v(r)|i\rangle$ and (2) change the band wave function as

$$\delta|i\rangle = \sum_{f\neq i} \frac{\langle i|\delta U_v(r)|i\rangle}{E_f - E_i}|f\rangle$$

Thus, the susceptibility tensor in [12] will be changed:

$$\delta\chi_v^{\alpha\beta} = \sum_{a=1}^{N_c} Q_{a,v} \frac{\partial\chi^{\alpha\beta}(\omega)}{\partial Q_a} \qquad [13]$$

This leads to the scattering intensity at the phonon frequency Ω_v:

$$\rho(\omega) = \frac{1}{\pi} I_R \frac{\gamma_v}{(\omega - \Omega_v)^2 + \gamma_v^2} \qquad [14]$$

where γ_v is the phonon line width and I_R is given by eqn [11]. By inserting [12] into [13], one sees that the integrated intensity of phonons is determined by two contributions. The first one is due to the dependence of the band energy on the ion configuration, $\partial\omega_{fi}/\partial Q_a$. The second one can be ascribed to the polarization of band wave functions and the change in the interionic distances $\partial\langle f|\mathbf{M}_\alpha|i\rangle/\partial Q_a$. Note here that phonon RS is not the result of a direct light–phonon interaction because of the large difference of the photon energy ($\hbar\omega_I \approx 1.5-2.5$ eV) and the phonon energy (<100 meV). Rather, one-phonon RS arises from the third-order time-dependent perturbation theory as depicted by a diagram in **Figure 4**.

The RS process includes three virtual electronic transitions, which can be understood as three steps:

1. The incident photon excites a virtual electron–hole pair state via the electron–radiation coupling.
2. The electron–phonon interaction causes a transition of the electron (Stokes process) or hole (anti-Stokes one) to a different state.
3. The transition to the electronic ground state occurs via a recombination of the electron–hole pair emitting the scattered photon.

The relevant Hamiltonian involving a phonon excitation may be written as

$$H_R = -\frac{e}{mc}\boldsymbol{p}\cdot\boldsymbol{A} + \sum_{ijk} g_{ij}(k)(b^\dagger + b)c_{ik}^\dagger c_{jk} \qquad [15]$$

where $g_{ij}(k)$ is the electron–phonon coupling between the band i and j, while $b^\dagger(b)$ and $c_{ik}^\dagger(c_{jk})$ are the phonon and electron creation (annihilation) operators, respectively. The first term in [15] is the electron–radiation interaction involving steps 1 and 3. It describes a scattering process via an intermediate state resulting in an interband transition. The second term in [15] is a contribution, which is linear in terms of the electron–phonon interaction. This term involves step 2 by coupling phonons to electrons states in two ways: (1) the phonon couples electron states with the same occupancy (both are empty or occupied) and (2) the phonon couples occupied and unoccupied states.

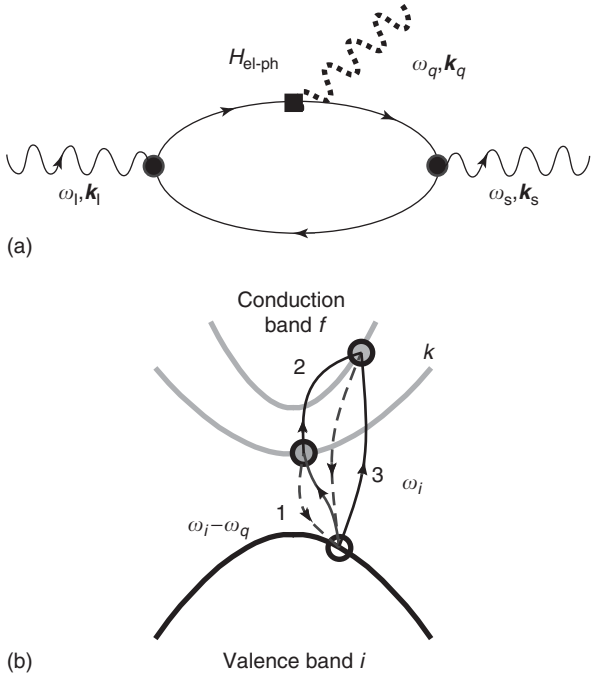


Figure 4 (a) Illustration of a first-order phonon scattering process. (b) Schematic electronic energy bands of a crystal showing the electronic transitions of the third-order light-scattering process for Stokes scattering. The pair state that is created in the first transition step transforms in the second step within the conduction band before recombining again during final transition step.

Often, higher-order phonon RS shows a substantial scattering cross section. This signal may deliver valuable information about the electronic states of semiconductors and insulators. **Figure 2** gives an example of such an observation in the semiconductor Si. The microscopic origin of the two-phonon process is considered by expanding the electron–phonon interaction up to the quadratic term:

$$\begin{aligned} H_{e-p} = &\sum_q g_1(q)(b_{-q}^\dagger + b_q)c_{k+q}^\dagger c_k \\ &+ \sum_q g_2(q)(b_{q1}^\dagger b_{q2}^\dagger + b_{q1}^\dagger b_{q2} \\ &+ b_{q1}b_{q2})c_{k+q}^\dagger c_k \end{aligned} \qquad [16]$$

Here, g_2 is the anharmonic interaction constant. The second term of the Hamiltonian [16] accounts for a simultaneous emission of two phonons due to the lattice anharmonicity. In contrast, the first term in the Hamiltonian [16] shows that two phonons are created separately in a pair of first-order electron–lattice interactions. This process contains an additional virtual electron state besides the emission of two phonons. This state can be in resonance when the energy of the incident light is suitably selected. In this case, the intensity of two-phonon scattering as a function of the incident light energy can determine interband energies.

When the two phonons are created, the energy and momentum conservation gives

$$\begin{aligned} \omega_{\sigma q} + \omega_{\sigma' q'} &= \omega_{\mathrm{I}} - \omega_{\mathrm{S}} \\ \boldsymbol{q} + \boldsymbol{q}' &= \boldsymbol{k}_{\mathrm{I}} - \boldsymbol{k}_{\mathrm{S}} \end{aligned} \qquad [17]$$

Here σ and σ' denote the branch of phonons. Thus, one can see that only the total momentum of a pair of emitted phonons $\boldsymbol{q} + \boldsymbol{q}' = 0$ should vanish. This is because phonon scattering is restricted to a regime near the Γ-point of the Brillouin zone. Then, the frequencies are the same at wave vectors $\boldsymbol{q}$ and $-\boldsymbol{q}$, and the energy conservation can be written as $\omega_{\sigma q} + \omega_{\sigma' q} = \omega_{\mathrm{I}} - \omega_{\mathrm{S}}$. To the first approximation, the second-order light-scattering spectrum of the phonons is provided by the combined DOS of pairs of phonons:

$$\rho_2(\omega) = \sum_{\sigma,\sigma'} \sum_{q} \delta(\omega - \omega_{\sigma q} - \omega_{\sigma' q}) \qquad [18]$$

In contrast to one-phonon scattering, this implies that higher-order scattering can provide indirect information about the whole Brillouin zone.

Until now, it has been assumed that electronic states of the ground and excited state do not change in the RS process by phonons. For solids and large molecules, this is largely true as **Figure 4b** depicts. For small molecules, however, the positions of vibrational functions are substantially shifted from the initial to the intermediate state. A sketch of such a situation is shown in the inset of **Figure 5**. As a result, the overlap between the vibrational wave functions of the two states is nonvanishing. This can lead to a multiphonon process near resonance via the so-called Frank–Condon mechanism. A representative example of multiphonon RS for the molecular system $(C_5H_{12}N)_2CuBr_2$ is given in **Figure 5**. Higher-order scattering appears exactly at the integer multiples of the one-phonon frequency. Furthermore, higher-phonon peaks are sharp. This behavior, arising from a local mechanism, should be contrasted with the observations in Si. For the latter, the main contribution to an n-phonon scattering comes from regions of the Brillouin zone where the phonon DOS is the largest. Moreover, on going to a higher order, the respective maxima shift to lower energy due to anharmonic phonon–phonon interactions. Note that the Frank–Condon mechanism is not restricted to small molecules. Even in solids and semiconductors, a local mechanism, such as self-trapped excitons, can lead to similar features of multiphonon scattering.

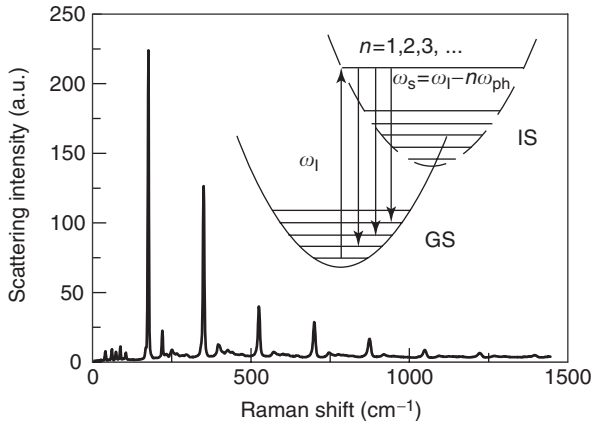


Figure 5 Multiphonon RS in the molecular system $(C_5H_{12}N)_2CuBr_4$ in (*xx*) polarization. The inset shows a configurational diagram including the electronic ground state (GS), intermediate state (IS), and phonon transitions $\omega_{\mathrm{S}} = \omega_{\mathrm{I}} - n\omega_{\mathrm{ph}}$ of different order.

Electronic Raman Scattering

Electronic systems can also contribute to the RS response. The scattering mechanism is related to fluctuations of effective charge density. If the incident energy of the photons is smaller than a bandgap, the effective charge density is expressed in terms of the curvature of the bands. This means that fluctuations of the effective mass tensor give a dominant contribution to electronic RS. Here, the effective charge density is not a conserved quantity in contrast to real charge. Thus, RS has no sum rule of spectral weights in contrast to IR absorption. In a clean metal, however, electronic RS is suppressed by screening. This is demonstrated taking the Lindhard expression in the limit of a small wave vector. For strongly correlated electron systems, the electronic Raman response contains valuable information about the lifetime or scattering rate of quasiparticles. Compounds with a metal–insulator transition show a strong spectral rearrangement of electronic RS with a characteristic frequency (the so-called isosbestic point) where RS is independent of temperature. This

behavior is common for Kondo systems, mixed-valence materials, and underdoped high-temperature superconductors. In superconductors with an anisotropic gap, this anisotropy can be determined using the polarization dependence of the electronic scattering and the respective depletion of spectral weights.

Magnetic Raman Scattering

For many materials, light scattering through magnetic excitations is relevant. Magnons induce a spatially periodic modulation of the susceptibility tensor. In this case, the modulation of $\chi^{\alpha\beta}(r)$ can be expressed in the Taylor series in terms of spin operators:

$$\chi^{\alpha\beta}(r) = \chi_0^{\alpha\beta}(r) + \sum_{\mu} K_{\mu}^{\alpha\beta}(r) S_r^{\mu} + \sum_{\mu,\nu} G_{\mu\nu}^{\alpha\beta}(r) S_r^{\mu} S_r^{\nu} + \sum_{\delta} \sum_{\mu,\nu} H_{\mu\nu}^{\alpha\beta}(r,\delta) S_r^{\mu} S_{r+\delta}^{\nu} + \cdots \quad [19]$$

Here, the coefficient tensors K, G, and H denote the strength of the coupling between light and the magnetic system. The constant term $\chi_0^{\alpha\beta}$ corresponds to an elastic scattering contribution. The linear term in S_r^{μ} can be written as a combination of the terms S_r^{+} and S_r^{-} which describe one-magnon creation and annihilation, respectively. The quadratic term in spin operators at a single ionic site r also gives rise to one-magnon scattering because this term can be rewritten in terms of a linear, transverse spin operator as $S_r^{+}S_r^{z}$, $S_r^{z}S_r^{+}$, $S_r^{-}S_r^{z}$, and $S_r^{z}S_r^{-}$. The quadratic spin term at different sites also has a contribution to the one-magnon scattering due to the terms proportional to $S_r^{+}S_{r+\delta}^{z}$, $S_r^{z}S_{r+\delta}^{+}$, $S_r^{-}S_{r+\delta}^{z}$, and $S_r^{z}S_{r+\delta}^{-}$.

This phenomenological description of one-magnon scattering can be derived from a microscopic mechanism of spin–orbit coupling. For simplicity, the ground state consisting of zero-orbital angular momentum and spin S, as displayed in **Figure 6a**, is considered. The effective magnetic field will split this state into $(2S+1)$ components. An excited orbital state with $L=1$ and the same S as the ground state is assumed. This will be split into three components corresponding to $J=S+1$, S, and $S-1$ by spin–orbit interaction $\lambda L \cdot S$, where λ is the spin–orbit coupling constant. Each excited state $|J, J^z\rangle$ is expressed as a linear combination of the unperturbed $|L^z, S^z\rangle$ states. Then, the transition from $S^z = S$ ground state to $S^z = S-1$ takes place by a successive electric dipole transition via the $L=1$ virtual intermediate state. In this way, the coupling constant K can be calculated in terms of the spin–orbit coupling λ.

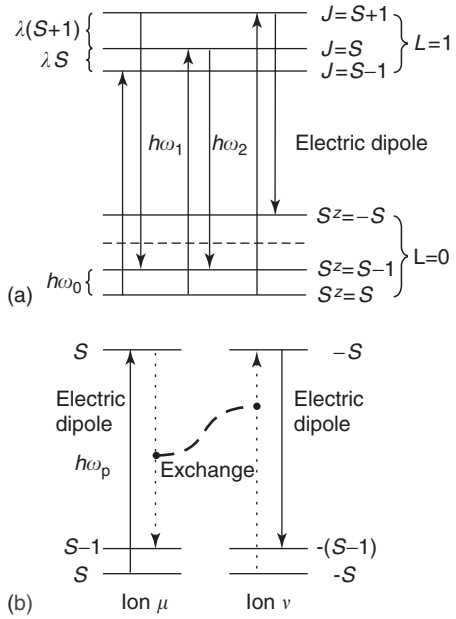


Figure 6 (a) One-magnon scattering and (b) two-magnon exchange scattering.

Two-magnon scattering arises from a fundamentally different process. Although it is a higher-order scattering process, it has generally a larger scattering cross section compared to one-magnon scattering. Two-magnon scattering is based on a double spin-flip caused by the matrix elements of the Coulomb interaction conserving the total z-component of spins. The so-called exchange scattering mechanism is illustrated in **Figure 6b**. The ground state of ion μ has an electron r_1 with $S^z = 1/2$ (denoted by $\uparrow$) in an orbital $|\mu\rangle$, while ion ν has an electron r_2 with $S^z = -1/2$ (denoted by $\downarrow$) in an orbital $|\nu\rangle$. The simultaneous change of the spin components of ions μ and ν occurs through the second order in the electric dipole interaction, and the first order in the exchange interaction between two electrons via Coulomb interaction. As a result, the effective spin Hamiltonian is given by

$$H_{2-\mathrm{mag}} \propto \sum_{ij} E_{\mathrm{S}}^{i} E_{\mathrm{I}}^{j} S_{\mu}^{-} S_{\nu}^{+} \quad [20]$$

Note that the two-magnon Hamiltonian has the same form as the Heisenberg Hamiltonian, as the same exchange mechanism involving the Coulomb interaction is relevant. Two-magnon scattering is effective for nearest-neighbor spins of an antiferromagnetic system, and even short-range correlations of exchange coupling can be detected. It allows one to determine exchange constants and magnon–magnon interactions. As an example, a two-magnon

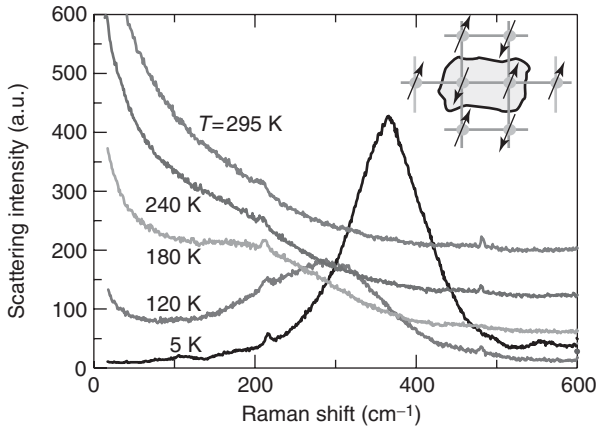


Figure 7 Two-magnon exchange scattering in ($x'y'$) polarization in $LaSrMnO_4$ with a maximum at $\Delta\omega = 360\ cm^{-1}$ as a function of temperature. The ordering temperature of the compound is $T_N = 126$ K. Curves are given an offset for clarity. The inset shows a sketch of the spins in a plane involved in an exchange scattering process.

scattering on the antiferromagnet $LaSrMnO_4$ with layers of Mn^{3+} and localized $S = 2$, as sketched in the inset of **Figure 7**, is shown. Each spin pair involved in a two-magnon process has six neighbors. A rough estimation of the energy of such a process would be $E_{2-\mathrm{mag}} = 6S \cdot J$. The broad maximum at $\Delta\omega = 360\ cm^{-1}$ shown in **Figure 7** is attributed to the two-magnon scattering. It leads to an exchange coupling constant $J = 43$ K, in good agreement with neutron scattering experiments. At elevated temperatures, the maximum is broadened but still visible for $T > T_N = 126$ K. With even higher temperatures, it is superimposed by a quasi-elastic tail of scattering that is induced by energy fluctuations.

Quasi-Elastic Scattering

In many magnetic systems, the energy of a magnetic system is not constant, but fluctuates about its mean value. This leads to quasi-elastic scattering as shown in the RS data on $LaSrMnO_4$ at $T = 295$ K. This tail of scattering around $\Delta\omega \approx 0$ on the Stokes and anti-Stokes side of the spectrum increases strongly with increasing temperature. The scattering intensity is described as a pair correlation function of the spin energy:

$$I(\omega) \propto \int_{-\infty}^{\infty} \ll E(k,t)E^*(-k,0) \gg \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}t \quad [21]$$

where $E(k,t)$ is magnetic energy density given by the Fourier transform of $E(r) = -\langle \sum_{i>j} J_{ij} S_i \cdot S_j \delta(r - r_i) \rangle$ with the position of the ith spin r_i. Under hydrodynamic conditions and for high temperatures, the scattering intensity is simplified as a Lorentzian lineshape

$$I(\omega) \propto C_\mathrm{m} T^2 \frac{D_\mathrm{T} k^2}{\omega^2 + (D_\mathrm{T} q^2)^2} \quad [22]$$

This equation implies that the integrated intensity and the half-width at half maximum of the quasi-elastic scattering is determined by the magnetic specific heat C_m and thermal diffusion constant D_T, thus enabling one to determine thermodynamic quantities of the magnetic system by a spectroscopic experiment.

See also: Group Theory in Materials Science, Applications; Interaction of Light and Matter; Scattering, Inelastic: Brillouin; Spin Density Waves and Magnons; Strongly Correlated Electron Systems.

PACS: 78.30. – j

Further Reading

Bright Wilson E, Decius JC, and Cross PC (1995) *Molecular Vibrations*. New York: McGraw-Hill.

Cardona M and Güntherodt G (eds.) (1975–2000) *Light Scattering in Solids*, vol. 1–8. Berlin, Heidelberg, New York: Springer.

Cottham MG and Lockwood DJ (1986) *Light Scattering in Magnetic Solids*. New York: Wiley-Interscience.

Fabelinskii IL (1998) Seventy years of combination (Raman) scattering. *Physics – Us-bekhi* 41(12): 1229.

Hayes W and Loudon R (1978) *Scattering of Light by Crystals*. New York: Wiley.

Kuzmany H (1988) *Solid State Spectroscopy*. Berlin and Heidelberg: Springer.

Lemmens P, Güntherodt G, and Gros C (2003) Magnetic light scattering in low-dimensional quantum spin systems. *Physics Reports* 375: 1.

Rousseau DL, Baumann RP, and Porto SPS (1981) Normal mode determination in crystals. *Journal of Raman Spectroscopy* 10: 253.

Schrader B (ed.) (1995) *Infrared and Raman Spectroscopy*. New York: VCH.

Sherman E, Ya Misochko OV, and Lemmens P (2003) What can one learn from Raman spectra of high-temperature superconductors. In: Plakida NM (ed.) *Spectroscopical Studies of High Temperature Superconductors*, ch. 2. London and New York: Taylor and Francis.

Spiro TG (ed.) (1987) *Biological Applications of Raman Spectroscopy*, vol. 2. New York: Wiley-Interscience.

Sugai S (2000) in Magneto-optics. In: Sugano S and Kojima N (eds.) *Raman Spectroscopy of Magnetic Compounds with Strong Electron-Correlation*, vol. 128, ch. 3, pp. 75–106. New York: Springer.

Turrell G and Corset J (eds.) (1996) *Raman Microscopy. Developments and Applications*. London: Academic Press.

Scattering, Inelastic: X-Ray (Methods and Applications)

M Krisch and F Sette, European Synchrotron Radiation Facility, Grenoble, France

Introduction

Inelastic X-ray scattering (IXS) experiments, including both energy and direction (momentum) analysis of the scattered photons, provide an important spectroscopic tool in the study of dynamical properties of matter. The different methods can be roughly classified according to the size of the energy (and momentum) transfer. This is schematically illustrated in **Figure 1**. IXS from collective ion excitations aims at the study of phonons and phonon-like excitations. It is complementary to inelastic neutron scattering (INS), and has proven to be particularly powerful in cases where INS techniques cannot be applied. Electron dynamics and electronic excitations are studied in various regimes. Compton scattering (which will not be further discussed in detail, see contribution by A Bansil) allows one to determine the electron momentum distribution. IXS from valence and core electrons bears a very close resemblance to electron-energy-loss spectroscopy, X-ray absorption, and photoemission techniques, and gives access to single-particle or collective electron (plasmon) excitations.

Despite the fact that most of the IXS techniques were developed several decades ago, they gained their full maturity only after the outstanding properties of storage ring-based X-ray sources, in terms of source brilliance and energy tunability, became apparent. IXS experiments are performed over a very large spectral range – from the vacuum-ultraviolet over the soft X-ray range into the hard X-ray regime. Consequently, the associated instrumentation is very diversified, and is based on diffraction gratings for the low energy range, whereas perfect Bragg crystal optics are utilized in the X-ray range.

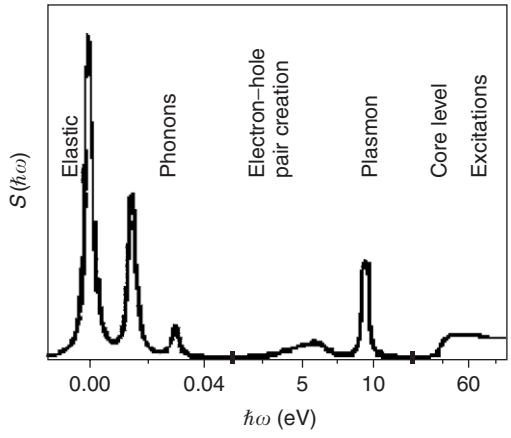


Figure 1 Schematic X-ray excitation spectrum $S(\hbar\omega)$ as a function of energy transfer.

Theoretical Aspects

The basic kinematics of the inelastic scattering process are depicted in **Figure 2**. The incident photon of energy $\hbar\omega_1$ and momentum $\hbar\boldsymbol{k}_1$ is scattered from the sample into a photon of energy $\hbar\omega_2$ and momentum $\hbar\boldsymbol{k}_2$. The process is therefore characterized by the energy transfer $E = \hbar\omega_1 - \hbar\omega_2$ and the momentum transfer $\hbar Q = \hbar\boldsymbol{k}_1 - \hbar\boldsymbol{k}_2$. Here, the polarization transfer is not considered. In cases where the energy transfer is much smaller than the incident photon energy, Q is solely controlled by the scattering angle Θ_s, and is given by the following expression:

$$Q = 2k_1 \sin(\Theta_s/2)$$

The Hamiltonian, describing the electron–photon interaction in the scattering process, is composed, in the weak relativistic limit, of four terms. Neglecting the much weaker magnetic couplings, only two terms have to be retained:

$$H_{\text{int}} = \sum_j \frac{e}{2mc^2} A_j^2 + \sum_j \frac{e}{mc} A_j p_j$$

where $r_0 = e^2/mc^2$ is the classical electron radius, $\boldsymbol{A}$ the vector potential of the electromagnetic field, and $\boldsymbol{p}$ the momentum operator of the scattering electrons. The sum extends over all the electrons in the system. Treating the scattering process in the lowest-order perturbation theory, representing the vector potential $\boldsymbol{A}$ in terms of photon creation and annihilation operators, and assuming that both the initial and the final photon states can be represented by plane waves, one obtains, for the double differential cross

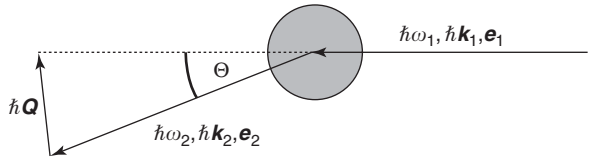


Figure 2 Kinematics of the scattering process.

section (DDCS),

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\omega_2} = r_0^2\left(\frac{\omega_2}{\omega_1}\right)\sum_F\left|\left\langle F|\sum_j \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_2-\boldsymbol{k}_1)\boldsymbol{r}_j}|I\right\rangle(\boldsymbol{e}_1\cdot\boldsymbol{e}_2)\right.$$
$$-\frac{1}{m}\sum_N\left\{\frac{\left\langle F|\boldsymbol{e}_2\cdot\sum_j \boldsymbol{p}_j\mathrm{e}^{-\mathrm{i}\boldsymbol{k}_2\boldsymbol{r}_j}|N\right\rangle\left\langle N|\boldsymbol{e}_1\cdot\sum_j \boldsymbol{p}_j\mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\boldsymbol{r}_j}|I\right\rangle}{E_N-E_I-\hbar\omega_1-0.5i\Gamma_N}\right.$$
$$\left.\left.+\frac{\left\langle F|\boldsymbol{e}_1\cdot\sum_j \boldsymbol{p}_j\mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\boldsymbol{r}_j}|N\right\rangle\left\langle N|\boldsymbol{e}_2\cdot\sum_j \boldsymbol{p}_j\mathrm{e}^{-\mathrm{i}\boldsymbol{k}_2\boldsymbol{r}_j}|I\right\rangle}{E_N-E_I+\hbar\omega_2}\right\}\right|^2$$
$$\times\,\delta(E_\mathrm{F}-E_I-\hbar\omega)$$

where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the polarization vector of the incident and the scattered photon, respectively, and Γ_N is the lifetime broadening of the intermediate state. The δ-function ensures energy conservation. The DDCS is proportional to the number of incident probe particles scattered within an energy range $\Delta E(\Delta\hbar w_2)$ and wave vector transfer variation ΔQ into a solid angle $\Delta\Omega$. If the interference terms are neglected, the cross section is composed of three contributions. The first term, derived from the A^2-term in first-order perturbation, describes nonresonant IXS. The second and the third term are derived from the $p\cdot A$-term in second-order perturbation. The second term describes an absorption process followed by an emission process, coupled via a resonant intermediate state $|N\rangle$, leading to an increase of the cross section, whenever the energy of the incident photon is close to an absorption edge of the element under study. The third term can be neglected in practical cases.

From the first term of the DDCS, which implicitly contains the correlation function of the electron density, one arrives at the correlation function of the atomic density, on the basis of the following considerations:

(1) *Validity of the adiabatic approximation*. This allows one to separate the system quantum state $|S\rangle$ into the product of an electronic part, $|S_\mathrm{e}\rangle$, which depends only parametrically on the nuclear coordinates, and a nuclear part, $|S_\mathrm{n}\rangle : |S\rangle = |S_\mathrm{e}\rangle|S_\mathrm{n}\rangle$. This approximation is particularly good for exchanged energies that are small with respect to the excitation energies of electrons in bound core states: this is indeed the case in basically any atomic species when considering phonon energies. In this approximation, one neglects the portion of the total electron density contributing to the delocalized bonding states in the valence band region.

(2) Limitation to the case for which the electronic part of the total wave function is not changed by the scattering process, and therefore the difference between the initial state $|I\rangle = |I_\mathrm{e}\rangle|I_\mathrm{n}\rangle$ and the final state $|F\rangle = |I_\mathrm{e}\rangle|F_\mathrm{n}\rangle$ is due only to excitations associated with the atomic density fluctuations.

IXS from Electronic Excitations

In the following, the various IXS processes which arise from electronic excitations (with the exception of Compton scattering) are discussed. These can be roughly divided into two main fields according to the corresponding piece of the DDCS: (1) nonresonant IXS and (2) resonant IXS. In contrast to spectroscopies utilizing electron detection, IXS experiments are bulk sensitive and, if performed in the hard X-ray regime, allow one to study systems under extreme conditions such as very high temperatures and pressures.

Nonresonant Inelastic X-Ray Scattering from Electrons

The cross section for nonresonant IXS is derived from the A^2-term of the interaction Hamiltonian, and constitutes the first term in the expression of the DDCS. The DDCS can be factorized into two parts, the first describes the coupling of the scattering probe to the system, and the second contains the properties of the system in the absence of the perturbing probe:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\omega_2} = r_0^2(\boldsymbol{e}_1\cdot\boldsymbol{e}_2)^2\left(\frac{\omega_2}{\omega_1}\right)S(\boldsymbol{Q},\omega)$$
$$= \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_\mathrm{Th}S(\boldsymbol{Q},\omega)$$

where $(\mathrm{d}\sigma/\mathrm{d}\Omega)_\mathrm{Th}$ is the Thomson cross section. The dynamical structure factor contains the information about electronic excitations via the imaginary part of the dielectric response function $1/\varepsilon(\boldsymbol{Q},\omega)$:

$$S(\boldsymbol{Q},\omega) = \left(-\frac{\hbar Q^2}{4\pi^2e^2n}\right)\mathrm{Im}\left[\frac{1}{\varepsilon(\boldsymbol{Q},\omega)}\right]$$

where n is the electron density of the system. The dielectric response function in turn is directly related to the polarization function $\chi(\boldsymbol{Q},\omega)$:

$$\frac{1}{\varepsilon(\boldsymbol{Q},\omega)} = 1+\frac{4\pi e^2}{Q^2}\chi(\boldsymbol{Q},\omega)$$

where $\chi(\boldsymbol{Q},\omega)$ connects the Fourier-transformed induced charge $n_\mathrm{ind}(\boldsymbol{Q},\omega)$ with the Fourier transform of the external potential $\Phi_\mathrm{ext}(\boldsymbol{Q},\omega)$:

$$n_\mathrm{ind}(\boldsymbol{Q},\omega) = -e\chi(\boldsymbol{Q},\omega)\Phi_\mathrm{ext}(\boldsymbol{Q},\omega)$$

The basic ingredient of the polarization function is the so-called Lindhard polarization function, which describes the polarization of a homogeneous electron system by taking into account all single-particle excitation processes (particle–hole excitations) mediated by the momentum transfer, assuming that they are allowed by Pauli's principle, but neglecting the interaction of the particle and the hole with the surrounding electron fluid as well as any interaction between the particle and the hole left behind. A first step of approximating the Lindhard polarization function is the so-called random phase approximation (RPA) where, in addition to the external potential, the potential due to the induced polarization also acts on the system.

IXS from valence electrons **Figure 3** shows three characteristic $S(\boldsymbol{Q},\omega)$ spectra along with the real and imaginary parts of the dielectric function $\varepsilon(\boldsymbol{Q},\omega)$. The evolution of the spectra can be understood by inspection of $\varepsilon(\boldsymbol{Q},\omega)$ and $\mathrm{Im}[\varepsilon(\boldsymbol{Q},\omega)^{-1}]$ for three different regimes of the momentum transfer Q. Within the RPA formalism, a sharp structure in the $S(\boldsymbol{Q},\omega)$ appears at a frequency $\omega=\omega_\mathrm{p}$, the plasma frequency of the system, when $\mathrm{Re}[\varepsilon(\boldsymbol{Q},\omega)]$ has a zero passage and simultaneously $\mathrm{Im}[\varepsilon(\boldsymbol{Q},\omega)]$ is close to zero. This behavior persists up to the so-called critical momentum transfer Q_c, for which the zero passage of $\mathrm{Re}[\varepsilon(\boldsymbol{Q},\omega)]$ coincides with the high-energy cutoff of $\mathrm{Im}[\varepsilon(\boldsymbol{Q},\omega)]$. In the above regime of $Q<Q_\mathrm{c}$, the plasmon dispersion can be determined. For $Q>Q_\mathrm{c}$, the plasmons can decay into particle–hole excitations, and a much larger spectral range yields a significant contribution to the $S(\boldsymbol{Q},\omega)$. A comparison of the experimental spectra with relevant calculations gives an insight into electron correlation and band-structure effects. Moreover, IXS allows one to determine the bandgap of semiconductors or insulators, and – in the case of an indirect gap – its relative momentum. Furthermore, IXS spectra recorded in the particle–hole excitations regime provide information about the combined electron density of states. Finally, one gains access to the off-diagonal response of the dynamical structure factor by performing the IXS experiment in the presence of a standing-wave field, generated by the coherent superposition of an incident and diffracted wave vector when the sample is positioned in Bragg reflection conditions.

IXS from core electrons (X-ray Raman scattering) Here, the final state $|F\rangle$ of the electron system is characterized by a hole in a core level and an electron in a continuum state above the Fermi level. Within the RPA, the dynamical structure factor takes the following form:

$$S(\boldsymbol{Q},\omega)=\left(\frac{\hbar Q^2}{4\pi e^2 n}\right)\sum_{k',v'}|\langle \boldsymbol{k}'v|\exp(\mathrm{i}\boldsymbol{Q}\boldsymbol{r})|c\rangle|^2 \times\delta(\hbar\omega+E_\mathrm{c}-E(\boldsymbol{k}',v))$$

where $|c\rangle$ is a core state with energy E_c and $E(\boldsymbol{k}',v)$ is the energy of a one-electron Bloch state with reduced wave vector $\boldsymbol{k}'$ and band index v. The characteristics of the electronic transition are controlled by the transition operator. XRS from core electrons of low Z materials is analogous to soft X-ray absorption spectroscopy (XAS), as long as the momentum transfer Q is small compared with the radial extent of the wave function of the core electron, involved in the inelastic scattering process. While in XAS the incident photon energy has to be tuned to the absorption edge energy under consideration, in XRS this role is taken by the energy transfer, thus leaving a certain

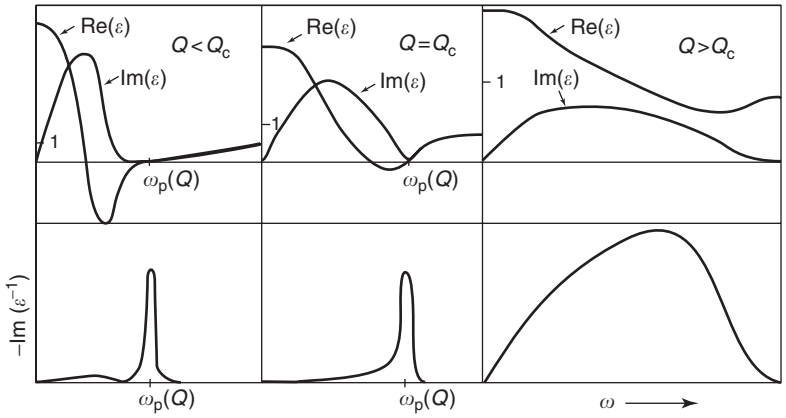


Figure 3 Upper panel: real and imaginary part of the RPA dielectric function $\varepsilon_{\mathrm{RPA}}(Q,\omega)$ for $Q<Q_\mathrm{c}$, $Q=Q_\mathrm{c}$, and $Q>Q_\mathrm{c}$. Lower panel: corresponding imaginary part of $-\varepsilon^{-1}_{\mathrm{RPA}}(Q,\omega)$. (Adapted from Schülke with permission.)

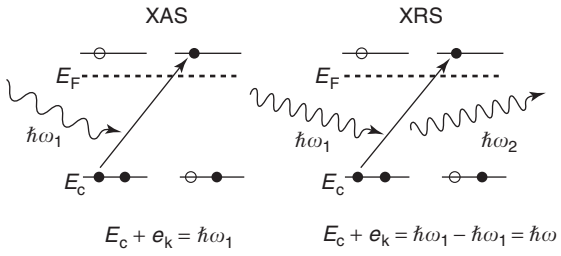


Figure 4 Schematic representation of the X-ray absorption (left) and X-ray Raman scattering (right) process. E_c denotes the binding energy of a core level, e_k the kinetic energy of the photoelectron, and E_F the Fermi energy.

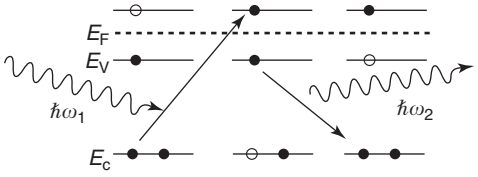


Figure 5 Schematic representation of the resonant inelastic X-ray scattering process. E_V denotes the binding energy of a (shallow) core level or the filled part of the valence band.

freedom in the choice of the incident photon energy. This is schematically illustrated in **Figure 4**. XRS therefore allows one to perform soft X-ray absorption studies in the hard X-ray regime, with the advantage of probing truly bulk properties and study systems which are not compatible with an ultrahigh vacuum environment, necessary in the soft X-ray regime. Moreover, by varying the momentum transfer $\hbar Q$, the electric dipole selection rule, defining the final state symmetry which can be reached in an absorption process, is relaxed and the full multipolar expansion of the transition operator has to be taken into account: in particular, electric monopole transitions become possible.

Resonant IXS

The cross section for resonant inelastic X-ray scattering (RIXS) is derived from the $p \cdot A$-term in the interaction Hamiltonian, and constitutes the second term of the DDCS. It is a second-order optical process in which a core electron is excited by an incident X-ray photon of energy $\hbar\omega_1$ into empty states above the Fermi level. This excited state then decays, and the deep core hole is filled by an electron from a shallower core level or the valence band under emission of a photon of energy $\hbar\omega_2$ (see **Figure 5**). In the most common case, both the excitation and deexcitation step involve electric dipolar transitions. If the intermediate core excited states are closely located in energy, interference between the absorption and emission process is possible. Dynamics in the time domain can be assessed if the intermediate state undergoes a transition (emission or absorption of a phonon) in the time interval between the excitation and the emission step. The rate for this transition can then be directly compared to the core hole decay rate, and information about the time development in the range of the core hole lifetime can be obtained. Well above the absorption threshold, the IXS process can be regarded as a decoupled two-step process, and is commonly dubbed X-ray fluorescence. In general, a simple one-electron picture is not sufficient to account for all the observed phenomena. Especially for open shell systems, such as transition metals or rare earths, a many-body description has to be invoked. Despite these complexities RIXS is a very powerful spectroscopic tool, thanks to its high degree of selectivity. It is element and angular momentum selective (via its resonant intermediate state and the transition selection rules, respectively). Site selectivity with respect to identical atomic species in different chemical environments can be obtained by choosing an excitation energy so that a core electron is promoted to an unoccupied state which has a strong localization on that particular site. Detecting the scattered X-rays with angular selectivity adds further information: in solids band-structure information can be obtained, whereas for molecules the symmetry of molecular states can be studied. Furthermore, exploiting the unique polarization properties of synchrotron-based X-ray sources, linear and circular dichroism phenomena can be studied. As in the case of nonresonant IXS, the various applications can be roughly classified according to the transferred energy.

RIXS from valence electrons These experiments aim to detect low-lying electronic excitations such as interband, *d–d*, and charge-transfer excitations. To perform these experiments resonant with a characteristic absorption edge bears several advantages: (1) the resonant enhancement of the cross section of typically two orders of magnitude allows one to study essentially all compounds of interest while nonresonant IXS is practically limited to low Z materials; (2) the presence of (an) intermediate state(s) and eventual interference effects between them make the RIXS process highly selective in terms of angular momentum selection rules and geometrical considerations when the experiment is performed on single crystals. For example, the shape of resonantly excited valence fluorescence spectra from single crystals depends strongly on the energy of the incident photon and the momentum transfer Q (both direction and

size) which is a direct consequence of momentum conservation, stating that Q must be equal to the vector difference $\boldsymbol{k}_e - \boldsymbol{k}_h$ between the Bloch vectors of the excited electron and those of the hole left behind, modulo a reciprocal lattice vector $\boldsymbol{G}$. This Bloch k-selective RIXS allows, in principle, one to obtain band-structure information although attention has to be paid to the relaxation processes in the intermediate state.

RIXS from core electrons Since the final state of the RIXS process is characterized by a hole in a shallow core level (or the valence band), spectral features narrower than the intermediate state lifetime width can be observed. The ultimate resolution is limited by the final-state lifetime width. If these final state RIXS features are associated with different excitation channels (located at different energies in the intermediate state), they will resonate at different incident photon energies. RIXS, therefore, can reveal different excitation channels that are completely obscured in the corresponding X-ray absorption spectrum due to the large core hole lifetime width. This allows, for example, one to determine the energy position and the relative strength of weak quadrupolar $(2p \rightarrow 4f)$ transitions in rare-earth compounds.

Spin-selective RIXS The shape of X-ray fluorescence spectra of elements which contain an open electronic shell (e.g., transition metals or rare earths) shows characteristic shoulders or satellites which reflect the energies of the different electronic configurations of the excited atom. Besides the Coulomb interactions between the electrons occupying different orbital configurations, there are some interactions that have a magnetic origin, in particular in cases where the excited atom has valence electrons with unpaired spin giving rise to a magnetic moment. These exchange interactions between the magnetic electrons of the valence shell and the core electrons in the orbital where the photon emission process has left a core hole lead to features in the emission spectrum of dominant spin character. The energy separation is related to the strength of these interactions, which are largest when the hole is in an orbital with the same principal quantum number n of the magnetic electrons. In a simple picture, the X-ray emission is separated into high- and low-spin states whose energy separation is proportional to the exchange integral J and $(2S+1)$, where S is the total spin moment of the magnetic shell, and whose intensity ratio is given by $S/(S+1)$. Changes in the magnetic state (induced, for example, by pressure, temperature, or doping) can therefore be witnessed by changes in the X-ray emission lineshape. By its nature, this method does not need a long-range magnetic order, and it is sensitive to the local (atomic) magnetic spin moment of a specific valence orbital. As a consequence, however, the method cannot discriminate between ferro-, antiferro-, ferri-, and paramagnetism.

Dichroism effects in RIXS Resonant IXS and fluorescence spectral shapes are sensitive to both the polarization of incident and scattered photons. In ferromagnetically aligned samples, the emission spectra therefore display dichroic effects, in close analogy to X-ray absorption spectroscopy utilizing circular or linear polarized X-rays. RIXS from valence electrons, excited with circular polarized X-rays, provides information on the occupied spin density, whereas RIXS from core electrons allows one to reveal the magnetic character of the predominant features composing the emission line. The X-ray emission linear magnetic dichroism effect represents the angular dependence of the X-ray emission cross section with respect to the magnetization axis. For a given configuration, the sizes of these dichroic effects are proportional to the magnetic moment of the involved magnetic shell.

IXS from Phonons

The study of atomic dynamics in condensed matter at momentum transfers Q and energies E, characteristic of collective motions, is traditionally the domain of neutron spectroscopy. There are, however, situations where the use of photons has important advantages over neutrons. The most significant is that there are no kinematic limitations: for IXS, any energy transfer can be reached for a given momentum transfer, whereas for INS, due to the neutron mass and typical neutron energies, the small-Q range is not accessible. This is depicted in **Figure 6** for different scattering angles Θ_s. Another advantage of IXS arises from the fact that very small beam sizes, in the order of a few tens of micrometers or less, can presently be obtained at third generation synchrotron sources (in contrast to centimeter-sized neutron beam sizes).

In its most general form, the DDCS can be written as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\omega_2\,\mathrm{d}\Omega} = r_0^2(\boldsymbol{e}_1\cdot\boldsymbol{e}_2)^2\frac{\omega_2}{\omega_1}\sum_{I_n,F_n}P_{I_n} \times\left|\left\langle F_n\left|\sum_j f_j(\boldsymbol{Q})\mathrm{e}^{\mathrm{i}\boldsymbol{Q}\boldsymbol{R}_j}\right|I_n\right\rangle\right|^2 \times\delta(\hbar\omega+\hbar\omega_2-\hbar\omega_1)$$

where $|I_n\rangle$ ($|F_n\rangle$) are the initial (final) nuclear states, P_{I_n} represents the thermal population of the initial

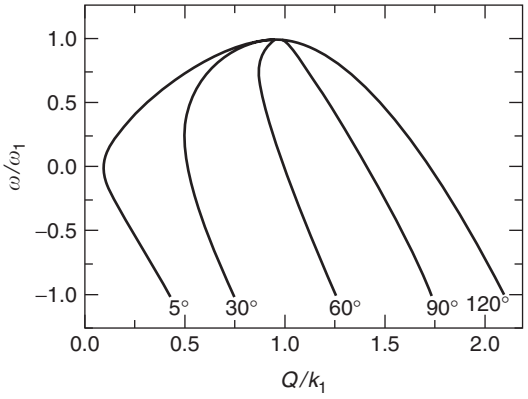


Figure 6 Accessible energy–momentum transfer regions for inelastic neutron scattering for different scattering angles Θ_s, reported normalized to the incident neutron energy $\hbar\omega_1$ and incident neutron momentum $\hbar k_1$.

states, and $f_j(\boldsymbol{Q})$ is the atomic form factor of the jth particle whose center-of-mass position is $\boldsymbol{R}_j$.

With respect to the INS cross section, the following aspects are noteworthy.

1. X-rays couple to the electrons of the system with a cross section proportional to the square of the classical electron radius, $r_0 = 2.82 \times 10^{-15}$ m, that is with a strength comparable to the neutron-nucleus scattering cross section b^2.
2. The IXS cross section is proportional to $f_j^2(\boldsymbol{Q})$. In the limit $Q \to 0$, the form factor is equal to the number of electrons in the scattering atom, Z; for increasing values of Q, the form factor decays with decay constants of the order of the inverse of the atomic wave function dimensions of the electrons in the atom.
3. For elements with $Z > 4$ and typical X-ray energies of 10–30 keV, photoelectric absorption is the main attenuation process. Far away from any absorption resonances, it is roughly proportional to Z^4, leading therefore to a reduction of the cross section for heavier elements.
4. The cross section for X-rays is highly coherent; therefore, only static or quasistatic structural and/or chemical disorder contributions have to be considered. Finally, it is noted that the instrument resolution function for IXS is independent of the energy and momentum transfer.

Phonons in Crystalline Materials

The capability of providing micrometer-sized X-ray beams opens up the possibility of studying systems available only in small quantities down to a few $10^{-6}\,\text{mm}^3$. These comprise novel or advanced materials such as superconductors for which sufficient crystalline quality and homogeneous doping can only be obtained for tiny crystals, by far too small to be studied by INS. Moreover, materials can be studied under very high pressures (>100 GPa) using diamond anvil cells thus allowing an important insight into the dynamics close to structural phase transitions and the pressure evolution of elastic properties.

IXS from phonons in single crystalline materials provides the same information as coherent INS, and all aspects of lattice dynamics in terms of symmetry considerations, phonon eigenvectors, and selection rules are identical. One exception, however, concerns the relative intensity of the phonons in a multi-species system (with j different atom species in the unit cell), which can be very different due to the different weight of the atomic form factor $f_j(\boldsymbol{Q})$ with respect to the neutron scattering length b_j.

For polycrystalline materials, only the modulus of Q is well defined, and in general the $S(\boldsymbol{Q}, \omega)$ spectrum will contain contributions from all the phonon branches. The density-of-states limit is reached, if a fine sampling of a sufficiently large Q-space is performed. On the other hand, if Q is confined within the first Brillouin zone, $Q = q$ (q is the momentum of the phonon), only phonon modes which have a longitudinal symmetry component are excited, allowing the determination of an orientation-averaged acoustic phonon dispersion.

High-Frequency Dynamics in Liquids and Glasses

The absence of kinematic limitations for IXS has allowed one to access a previously unexploited region in Q–ω space, the so-called mesoscopic region, in which the momentum transfer $\hbar Q$ approaches the inverse of the interparticle separation a. This Q-range is particularly important, because here the collective dynamics undergoes a transition from a hydrodynamic behavior to a microscopic single-particle one.

Numerous IXS studies in glasses have identified – despite qualitative differences – several common features in the high-frequency dynamics.

1. Propagating acoustic-like excitations exist up to a maximum Q-value Q_m($Q_m a \approx 1$–3), depending on the fragility of the system).
2. The excitations are well defined and display a linear dispersion $\hbar\omega(Q)$. An extrapolation toward the $Q \to 0$ limit yields the macroscopic sound velocity.
3. The width $\Gamma(Q)$ of the inelastic peaks follows a power law, $\Gamma = DQ^{\alpha}$ with $\alpha = 2$ within the current experimental accuracy.
4. The value of D does not depend significantly on temperature, indicating that this broadening in the

high-frequency region does not have a dynamic origin, but is instead due to static disorder.

For liquids, the dynamical structure factor is strongly affected, if the probed distances ($l = 2\pi/Q$) are comparable to those that characterize structural correlations among particles (over typical correlation lengths ξ), and times ($t = 1/\omega$) comparable to the lifetime τ of these correlations. In this case, one observes a positive sound dispersion $c(Q)$ from its low-frequency ($\omega = 2\pi\nu \ll 1/\tau$) adiabatic value c_0 to the high-frequency ($\omega \gg 1/\tau$) elastic value c_∞. A quantitative analysis of the $S(Q,\omega)$ can be performed within the framework of a generalized Langevin equation formalism using the memory function approach. This allows the reliable extraction of c_∞, relaxation times, and relaxation strengths.

Deep Inelastic X-Ray Scattering

Quantitative information on the atomic momentum distribution can be obtained at sufficiently large momentum transfers, where the core electrons of the system under study recoil almost freely with an effective mass equal to the atomic mass. These conditions are met, if the electron dynamics is instantaneous compared to the atomic one and if the atoms have an appropriate quantity of electronic charge on a length scale smaller than $1/Q$. Deep IXS provides a complementary technique to deep INS, and is particularly valuable for polyatomic or polyisotopic systems, for which the relative contrast between the different atomic/isotopic species in the sample is different for INS and IXS.

See also: Scattering Techniques, Compton; Scattering, Resonant.

PACS: 61.10.Dp; 61.10.Eq; 61.20.Lc; 61.43.Fs; 62.20.Dc; 62.50.+p; 63.20.−e; 71.10.−w; 78.70.−g; 78.70.Ck; 78.70.En

Further Reading

Burkel E (2000) Phonon spectroscopy by inelastic X-ray scattering. *Reports on Progress in Physics* 63: 171–232.

Gel'mukhanov F and Agren H (1999) Resonant X-ray Raman Scattering. *Physics Report* 312: 87–330.

Kotani A (2001) Resonant inelastic X-ray scattering spectra in solids. *Review of Modern Physics* 73: 203–247.

Ruocco G and Sette F (1999) The high-frequency dynamics of liquid water. *Journal of Physics: Condensed Matter* 11: R259–R293.

Ruocco G and Sette F (2001) High-frequency vibrational dynamics in glasses. *Journal of Physics: Condensed Matter* 13: 9141–9164.

Schülke W (1991) Inelastic scattering by electron excitations. In: Brown G and Moncton DE (eds.) *Handbook on Synchrotron Radiation*, vol. 3, pp. 565–637. Amsterdam: North-Holland.

Nomenclature

a	inter-particle separation (m)
$\boldsymbol{A}$	vector potential of the electromagnetic field ($\mathrm{kg\,m^2\,C^{-1}\,s^{-2}}$)
c	speed of light ($\mathrm{m\,s^{-1}}$)
c	sound velocity ($\mathrm{m\,s^{-1}}$)
e	electron charge magnitude (C)
$\boldsymbol{e}$	photon polarization vector
e_k	kinetic energy of the photoelectron (eV)
E_c	core level binding energy (eV)
E_F	Fermi energy (eV)
E_V	valence band binding energy (eV)
$f(Q)$	atomic form factor
$\boldsymbol{G}$	reciprocal lattice vector ($\mathrm{m^{-1}}$)
$\hbar$	reduced Planck's constant (J s)
J	exchange integral (eV)
k	wave vector ($\mathrm{m^{-1}}$)
m	electron mass (kg)
n	electron density ($\mathrm{m^{-3}}$)
$\boldsymbol{p}$	momentum operator of scattered electrons ($\mathrm{kg\,m\,s^{-1}}$)
Q	wave vector transfer ($\mathrm{m^{-1}}$)
$\boldsymbol{r}$	position operator of the electron (m)
r_0	classical electron radius (m)
$\boldsymbol{R}$	position operator of the ion (m)
S	total spin moment (μ_B)
$S(Q,\omega)$	dynamical structure factor ($\mathrm{eV^{-1}}$)
Z	number of electrons in an atom
Γ	lifetime broadening width (eV)
$\varepsilon(Q,\omega)$	dielectric function
θ_s	scattering angle (deg)
ξ	structural correlation length (m)
σ	cross section ($\mathrm{m^2}$)
τ	structural relaxation time (s)
χ	polarization function ($\mathrm{C^{-2}\,m^{-2}}$)
ω	angular frequency of radiation ($\mathrm{s^{-1}}$)
Ω	solid angle ($\mathrm{rad^2}$)

Scattering, Nuclear Resonant

W Sturhahn, Argonne National Laboratory, Argonne, IL, USA

Published by Elsevier Ltd.

Introduction

Nuclear resonant scattering (NRS) is a set of experimental techniques that use synchrotron radiation (SR) and well-defined resonances in atomic nuclei. Dynamics of the neutrons and protons cause a spectrum of these resonances for every nucleus except hydrogen. A nucleus can be created in an excited state by radioactive decay of a parent nucleus, or it can be brought into an excited state by collision with energetic particles (electrons, protons, neutrons, etc.) or the absorption of a photon. The latter process, a resonant excitation, becomes important if the energy of the photon is very close to the energy difference between the nuclear ground state and a nuclear excited state (see **Figure 1**.) Studies using NRS require isotopes with large nuclear resonant cross sections and transition energies that can be produced with high intensity by SR sources. NRS methods use timing techniques, and the observation of NRS from isotopes with very short lifetimes is technically difficult. Therefore, energy widths of the excited state $\sim 10^{-8}$ eV are favorable. Large values potentially result in higher counting rates in scattering experiments but also lead to shorter natural lifetimes.

Spectroscopic techniques with nuclear resonances are based on either the use of radioactive sources or parent nuclei, for example, Mössbauer spectroscopy and time-perturbed angular correlation (TPAC), or the direct excitation with photons, for example, nuclear magnetic resonance (NMR) or the SR-based NRS described here. Nuclear resonant inelastic X-ray scattering (NRIXS) for the study of lattice dynamics and synchrotron Mössbauer spectroscopy (SMS) for the study of hyperfine interactions assumed practical importance during the past decade at highly brilliant SR facilities.

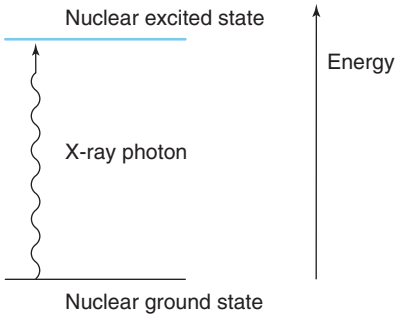


Figure 1 Resonant excitation of a nucleus. The energy width of the excited state is typically 12 orders of magnitude smaller than the energy of the photon. Nuclei show the most narrow resonances in the X-ray regime.

SMS includes scattering processes that occur without recoil, that is, without participation of lattice vibrations, and provides information about the electronic environment of the resonant nuclei analogous to conventional Mössbauer spectroscopy. The use of very collimated SR permits focusing to micrometer levels and easy manipulation of X-ray polarization. SMS has been applied very successfully to problems in high-pressure and biophysics research. In the study of magnetism, the strategic use of the isotopes ^{57}Fe (resonant) and ^{56}Fe (nonresonant) opened unique possibilities for analysis of thin films and nanostructures.

NRIXS is used for the study of lattice vibrations and provides one with a spectrum of the phonons experienced by the resonant isotope. Most importantly, one determines the phonon density of states (DOS) In contrast to other experimental techniques, NRIXS gives direct access to the partial and projected DOS of the resonant isotope only. This complete isotope selectivity is truly unique in the area of lattice dynamics research. For example, materials surrounding the sample that do not contain resonant nuclei do not produce a background, and this feature permitted high-pressure experiments in the Mbar regime that were impossible earlier. Many proteins contain only a few iron atoms at their biologically active sites; NRIXS studies using ^{57}Fe were conducted on several model compounds for such proteins, as well as on the proteins themselves. Among the thousands of atoms, NRIXS selects only the iron vibrations and thus permits detailed studies of protein dynamics related to these active sites. The same reasoning applies to studies of thin films and buried interfaces. Also, for nanoparticles and disordered alloys, the NRIXS method has been applied very successfullly. In those cases, as well as for amorphous materials, phonon dispersions are difficult or impossible to measure, but the DOS still gives valuable information.

Properties of Nuclear Resonances

X-ray scattering is generally treated solely with the electronic charge and spin. A justification is provided by classical electrodynamics in terms of the Thomson-scattering cross section, which is a factor of $(Zm/M)^2 \approx 10^{-7}$ smaller for the nucleus than for the electron shell. Here, m and M are the masses of electron

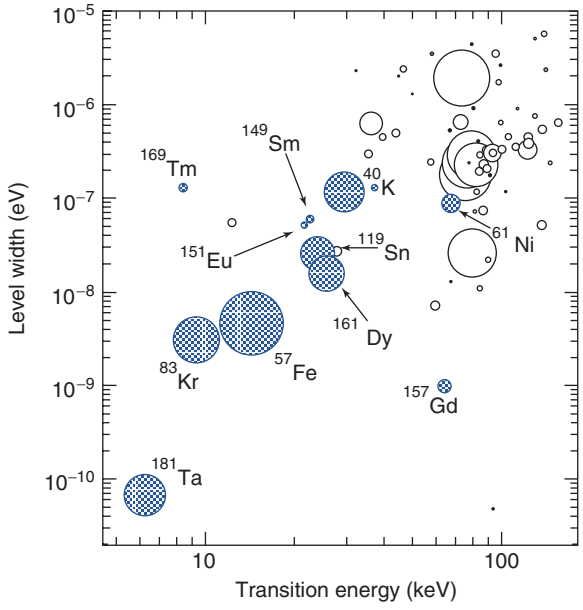


Figure 2 Level widths of nuclear resonances below 150 keV. The size of the symbols is proportional to the nuclear resonant cross section. Isotopes that have been used in NRS experiments are identified.

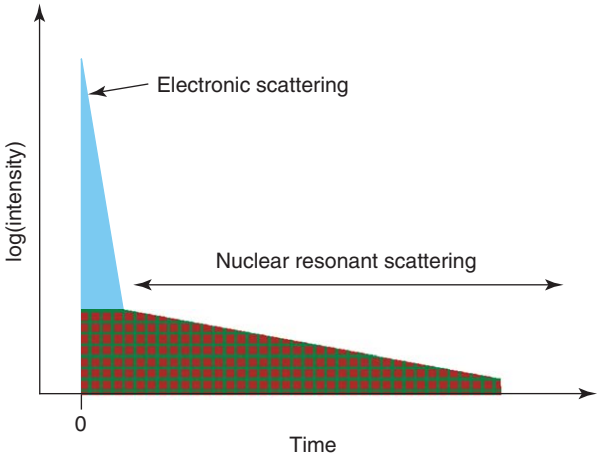


Figure 3 Scattered intensity vs. time. At zero time, an SR pulse excites a sample containing a nuclear resonant isotope. Electronic scattering is prompt whereas the response of the resonant nuclei is delayed. Time discrimination is the key to distinguishing nuclear and electronic scattering.

and nucleus respectively, and Z is the atomic number. Based on this, it seems appropriate to ignore the scattering contributions from the nucleus. However, the above argument fails if the timescale of internal nuclear dynamics matches the energy of the X-ray photon, that is, the nucleus experiences a resonant excitation. Then the nuclear resonant cross section is calculated as

$$\sigma_N = \frac{\lambda^2}{2\pi}\frac{1}{1+\alpha}\frac{2I'+1}{2I+1} \qquad [1]$$

where λ is the wavelength of the resonant X-rays, α is the internal conversion coefficient, and I,I' are the spins of nuclear, ground, and excited states. Values of α for relevant nuclear transitions range from 1 to 1000. The nuclear resonant cross section can become very large, for example, the 14.4 keV nuclear transition of ^{57}Fe gives (with $\lambda = 86$ pm $\alpha = 8.6$, $I = 1/2$, $I' = 3/2$) a value of $\sigma_N = 2.56$ Mbarn and a ration of $\sigma_N/\sigma_{pe} \approx 450$, where σ_{pe} is the photoelectric cross section. **Figure 2** displays nuclear resonant cross sections and nuclear-level widths for isotopes with resonances below 150 keV. Higher transition energies are less favorable because the intensity of SR sources decreases, and X-ray optics, as well as detectors, become less efficient. Even though the nuclear resonant cross section is very large, the extraordinarily narrow energy width of these resonances leads to a small energy-integrated cross section. For example, the best energy resolutions of X-ray optics are of the order $\delta E \approx 100\,\mu$eV, and a resonant enhancement over a neV scale would remain unresolved and unnoticed because $\sigma_N\Gamma \ll \sigma_{pe}\delta E$. Therefore, NRS experiments require special observation techniques.

A sample containing resonant nuclei responds to an SR pulse on a timescale defined by the lifetime of the nuclear state τ. The value of τ is inversely related to the nuclear-level width via $\tau\Gamma = \hbar$. Whereas values for τ are in the ns–μs range, the duration of an SR pulse is typically less than 100 ps. All electronic scattering of X-rays occurs typically on the timescale of femtoseconds, virtually instantaneous compared to nuclear resonant contributions. A "time-discrimination trick," which is illustrated in **Figure 3**, can cleanly separate nuclear resonant signals from other scattering contributions. The use of this time-discrimination trick makes NRS techniques possible.

Scattering Processes

A schematic of the excitation process of a nuclear resonance is shown in **Figure 1**. For a description of NRS, the de-excitation pathways must also be considered. The classification of elastic–inelastic and coherent–incoherent scattering processes is based on an analysis of the initial and final quantum state of the scatterer symbolized by

$$|\Phi_i\rangle|\gamma_i\rangle \rightarrow |\Phi_n\rangle \rightarrow |\Phi_f\rangle|\gamma_f\rangle \qquad [2]$$

States of the X-ray field and the scatterer are described by $|\gamma\rangle$ and $|\Phi\rangle$, respectively. The intermediate state exists without X-ray photons. An inelastic scattering process is characterized by different

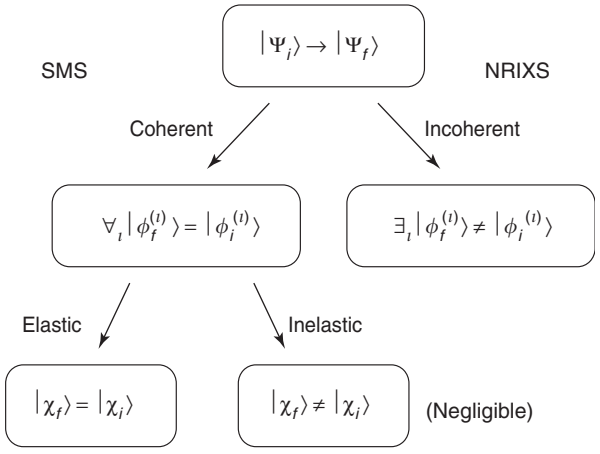


Figure 4 Classification scheme for scattering processes. Incoherent scattering implies energy transfer unless some core states are degenerate. Coherent inelastic NRS is of negligible strength.

energies of the initial and final photon states, which implies an energy transfer to the scatterer. A meaningful definition of coherent and incoherent scattering can only be given for a collection of atoms, for which

$$|\Phi_i\rangle = |\chi_i\rangle \prod_j |\phi_i^{(j)}\rangle \qquad [3]$$

Here, $|\phi^{(j)}\rangle$ symbolizes a core state, that is, the state of the nucleus and inner electron shell, of atom j. The collective state of itinerant electrons and lattice vibrations is written as $|\chi\rangle$.

A classification of scattering processes using this nomenclature is presented in **Figure 4**. Coherent processes leave all core states $|\phi^{(j)}\rangle$ unchanged. The scattering atom, in principle, cannot be identified, and the situation is analogous to a quantum mechanical multislit experiment. The change of the state of just one atom serves as an indicator, and the process will be incoherent. In particular, SMS is based on the coherent elastic process, whereas NRIXS originates from incoherent processes. Coherent inelastic scattering by resonant nuclei is of negligible strength. The almost purely incoherent character of inelastic NRS simplifies the interpretation of NRIXS data and presents a clear distinction to inelastic neutron scattering where incoherent and coherent processes often occur with similar strength.

Nuclear Resonant Inelastic X-Ray Scattering

The resonant absorption of an X-ray photon can lead to a simultaneous change of nuclear state and vibrational quantum state of the solid as illustrated in

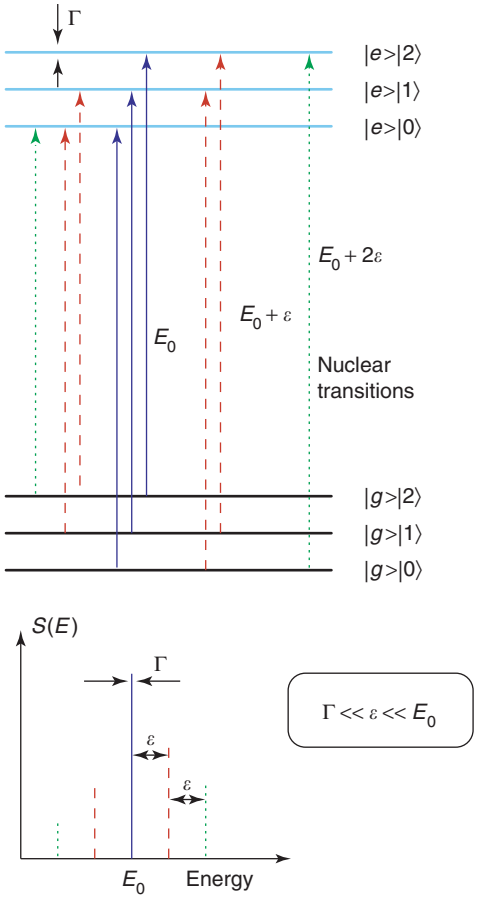


Figure 5 The influence of atomic vibrations on the nuclear level scheme. The nucleus is embedded in an Einstein solid with a specific phonon energy ε. This leads to sidebands in the excitation probability density $S(E)$ corresponding to phonon creation or annihilation.

Figure 5. Three energy scales determine the properties of this process: the transition energy of several keV; the phonon energies of the order of meV; the nuclear-level width of less than μeV. The energies are different by a many orders of magnitude, and energy eigenstates are very well described by a simple combination of nuclear states and phonon states, for example, $|g\rangle|1\rangle$ is a quantum state with the nucleus in the ground state and one phonon present. If one investigates the vibrating nucleus with monochromatic X-rays, one would observe transitions with nuclear transition energy E_0 described by $|g\rangle|n\rangle \rightarrow |e\rangle|n\rangle$ and m-phonon creation of the type $|g\rangle|n\rangle \rightarrow |e\rangle|n+m\rangle$ leading to upshifted transition energies $E_0 + m\varepsilon$, where ε is the energy of the phonon. Also, m-phonon annihilation of the type $|g\rangle|n\rangle \rightarrow |e\rangle|n-m\rangle$ with $n \geqslant m$ leads to downshifted transition energies $E_0 - m\varepsilon$. The absorption behavior is quantitatively described by the excitation probability density $S(E)$. The value of $S(E)\mathrm{d}E$ gives the probability that the

nucleus can be excited by X-rays in the energy range $[E, E+\mathrm{d}E]$. The phonon-creation/annihilation side bands of $S(E)$ are analogous to the Stokes and anti-Stokes lines observed by optical spectroscopies. The ratio $S(-E)/S(E)$ is universally given by the Boltzmann factor, that is, $S(-E) = \exp[-\beta E]S(E)$, with $\beta = 1/(k_\mathrm{B}T)$, temperature T, and Boltzmann's constant k_B. This relationship is known as "detailed balance" and is an intrinsic feature of all NRIXS spectra. The partial DOS is extracted by a mathematical inversion procedure from $S(E)$. Here, "partial" refers to the selection that has taken place by observing only vibrations of the resonant isotope.

Cross Section

Excitation probabilities can be calculated with Fermi's golden rule, which gives the transition rate Λ_{ni} for changes from the initial state $|\Phi_i\rangle = |\chi_i\rangle|\phi_i\rangle$ to an intermediate state $|\Phi_n\rangle = |\chi_n\rangle|\phi_n\rangle$ under the influence of a monochromatic X-ray field of energy $\bar{E}$, amplitude $\boldsymbol{a}$, and wave vector $\boldsymbol{k}$

$$\Lambda_{ni} = \frac{2\pi}{\hbar}\,|\langle\Phi_n|\hat{H}_\mathrm{int}|\Phi_i\rangle|^2\delta(E_{ni}-\bar{E}) \qquad [4]$$

The energy difference of the states is given by E_{ni}, and the interaction Hamiltonian can be expressed by

$$\hat{H}_\mathrm{int} = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{a}\cdot\hat{\boldsymbol{J}}(-\boldsymbol{k}) \qquad [5]$$

where $\hat{\boldsymbol{J}}$ is the spatial Fourier transform of the nuclear current operator in the center of mass frame, and the position of the center of mass is given by $\boldsymbol{r}$. One obtains for the cross section

$$\sigma(E,\boldsymbol{k}) = \frac{1}{|\boldsymbol{a}|^2}\sum_n\langle\Lambda_{ni}\rangle_i = \frac{\pi}{2}\,\sigma_\mathrm{N}\Gamma S(E,\boldsymbol{k}) \qquad [6]$$

where $E = \bar{E} - E_0$ is the energy of the incident X-rays relative to the nuclear transition energy. The angle brackets indicate an average over all initial states that may be present, for example, in a thermal distribution. The last term is defined by

$$S(E,\boldsymbol{k}) = \int\frac{\mathrm{d}t}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}Et/\hbar}\langle\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\hat{\boldsymbol{r}}(t)}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\hat{\boldsymbol{r}}(0)}\rangle \qquad [7]$$

This is the excitation probability density mentioned earlier, and here it is expressed as a self-correlation function of the position of a resonant isotope. For solid materials, $S(E,\boldsymbol{k})$ shows a sharp peak with width Γ around the nuclear transition energy E_0, that is, around $E=0$. The existence of this peak is tantamount to recoilless absorption of X-rays by the nucleus, the Mössbauer effect. In fact, $\Gamma S(0) = f$, the Lamb–Mössbauer factor or probability for recoilless absorption. The value of f varies approximately between 0.05 and 0.9 for solids at room temperature but vanishes for liquids and gases. The estimate for absorption "on resonance" is then

$$\sigma(0) \approx \frac{\pi}{2}\,\sigma_\mathrm{N} f \qquad [8]$$

With the exception of the elastic peak, $S(E)$ is expected to be a smooth function in energy, extending over an energy range that may be estimated by the Debye energy Θ. From eqn [7] one obtains $\int S(E)\mathrm{d}E = 1$, and with this normalization, one can estimate for the "off resonance" region $\Theta S(E) \approx 1-f$ providing

$$\sigma(E\neq 0) \approx \frac{\pi}{2}\,\sigma_\mathrm{N}(1-f)\frac{\Gamma}{\Theta} \qquad [9]$$

A quantitative estimate for ^{57}Fe in metallic form gives in units of the photoelectric cross section

$$\sigma(0) \approx 560\ \sigma_\mathrm{pe} \quad \sigma(E\neq 0) \approx 0.0002\ \sigma_\mathrm{pe} \qquad [10]$$

The change of cross section by six orders of magnitude when moving away from the transition energy is dramatic. The large "on resonance" value permits conventional energy-resolved Mössbauer spectroscopy. The observation of the phonon sidebands via the small "off resonance" value requires the use of time discrimination.

Observation

Imagine a resonant nucleus that was somehow excited. The nucleus will decay into the ground state either by emission of an X-ray photon or by transferring the excitation energy to the electron shell. In the latter case, an electron is expelled, and the hole is quickly filled by other electrons with the emission of fluorescence X-rays. All these decay products are emitted with some delay relative to the time of excitation of the nucleus, and the average delay time is given by the natural lifetime. **Figure 3** illustrates the time evolution of the scattered intensity of a material containing resonant nuclei after excitation with an SR pulse. If the energy of the incident X-ray is close to the nuclear transition energy, nuclei are excited, and delayed emission of X-rays can be observed. If only the delayed photons are counted while tuning the energy of the incident X-ray pulses, an event rate that is proportional to $\sigma(E)$ is expected to be measured. The normalized data then provide $S(E)$. An example of data collected from a metallic foil of ^{57}Fe is shown in **Figure 6**. The low background shows that the time-discrimination trick removes all nonnuclear scattering of the X-rays very effectively.

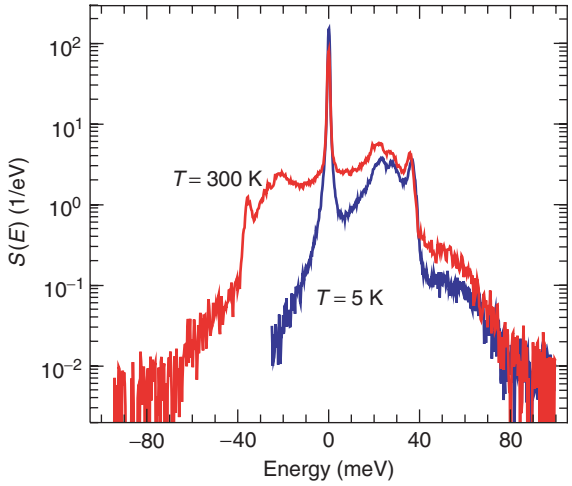


Figure 6 Phonon excitation probability density from NRIXS data of an iron foil at two temperatures. Energy zero corresponds to the nuclear transition energy of 14.4125 keV of the ^{57}Fe isotope. At low temperatures, phonon annihilation (negative energy values) is inhibited.

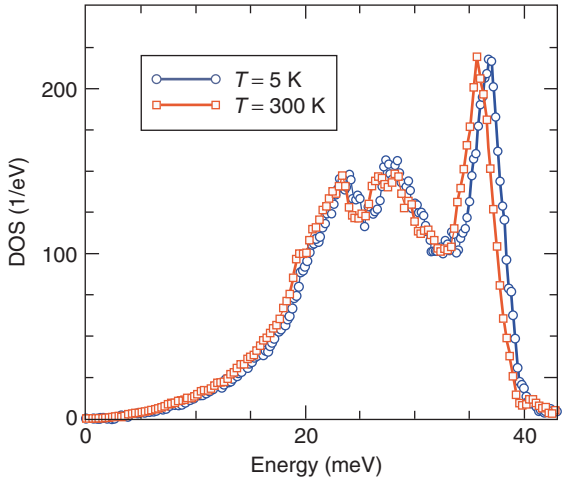


Figure 7 The DOS of iron metal (b.c.c.) at two temperatures. The curves show $3g$ calculated from the data shown in **Figure 6** using eqn [11].

Phonon Density of States

In many cases, an expansion of the interatomic potential around the equilibrium positions of the atoms in a solid is quite accurately described by a quadratic dependence on the atomic displacements. In this quasiharmonic approximation, the equation of motion for the atomic motions can be integrated. The excitation probability can then be rewritten in a physically meaningful way as

$$S(E) = f\delta(E) + \sum_{n=1} S_n(E)$$

$$S_1(E) = \frac{fE_R}{E(1-\exp(-\beta E))}\, g(|E|) \qquad [11]$$

$$S_n(E) = \frac{1}{nf} \int S_{n-1}(E')S_1(E-E')\ \mathrm{d}E'$$

where the individual terms S_n correspond to n-phonon contributions. The delta-distribution term describes elastic scattering. The probability for this process is described in terms of the Lamb–Mössbauer factor f. The probability $\mathrm{d}P_1$ to create ($E>0$) or annihilate ($E<0$) one phonon of energy E is given by the first-order term, that is, $\mathrm{d}P_1 = S_1(E)\,\mathrm{d}E$. The second-order term incorporates the participation of two phonons with energies that add to E. For a given inverse temperature β, all terms in eqn [11] are generated from the partial DOS $g(E)$. The measurement provides $S(E)$, and an inversion of eqn [11] gives $g(E)$. **Figure 7** shows the DOS of iron metal.

Synchrotron Mössbauer Spectroscopy

In solids, the probability that NRS occurs without change in the vibrational state is often appreciable and was introduced previously as the Lamb–Mössbauer factor f. In terms of the phonon excitation probability density of eqn [7], $f = \Gamma S(0)$ with corresponding transitions indicated in **Figure 5**. Now that vibrational transitions are excluded, slight energy differences of the nuclear levels become visible and measurable. The nuclear-level splitting is caused by the electronic environment and is known as hyperfine interaction. The exclusion of phonon excitations results in elastic coherent scattering. This process is described by a differential scattering cross section that is strongly peaked in certain directions and, in particular, the direction of the incident X-rays (nuclear forward scattering).

Level Splitting

Figure 8 shows transitions of a ^{57}Fe nucleus under the influence of an electric field gradient (EFG) or a magnetic field. The energy levels of the isolated nucleus are twofold and fourfold degenerate, respectively. The presence of an EFG, which is created by the electronic environment and couples to the quadrupole moment of the nucleus, partly removes the degeneracy. In the case of an axially symmetric EFG, the energy splitting for spin quantum numbers $I>1/2$ and $m = -I, \ldots, I$ is given by

$$E_m = \Delta\, \frac{3m^2 - I(I+1)}{2I(2I-1)} \qquad [12]$$

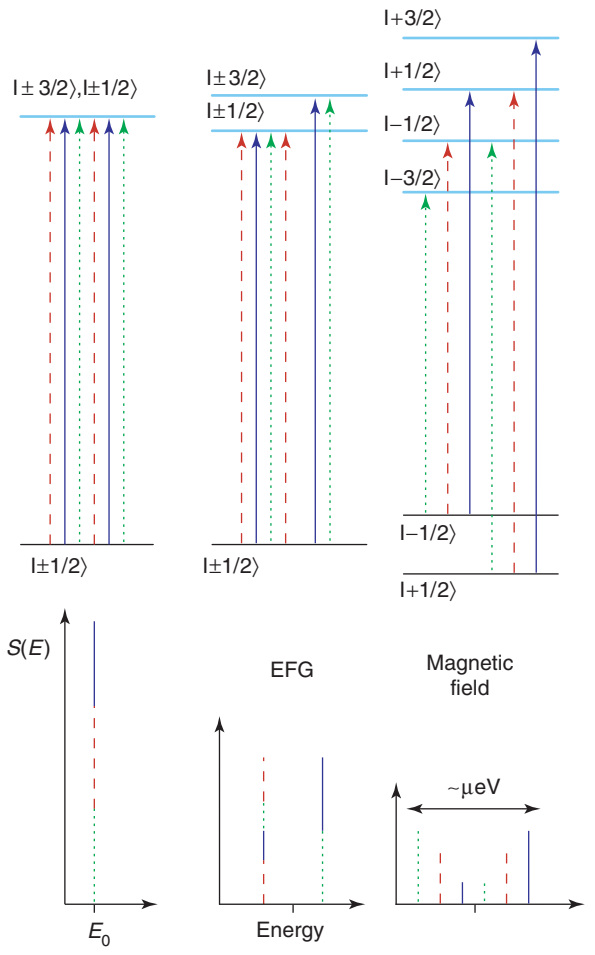


Figure 8 Nuclear level splitting of ^{57}Fe by interactions with electrons. The presence of an EFG or a magnetic field leads to several nuclear transitions separated by up to several μeV. The ground and excited state have spin quantum numbers of 1/2 and 3/2, respectively.

where Δ is called the electric quadrupole splitting. For the ^{57}Fe case, one obtains a characteristic two-line pattern with $E_{\pm 3/2} - E_{\pm 1/2} = \Delta$.

The presence of the magnetic field imprints unidirectional symmetry, which removes the degeneracy completely. The energy splittings are given by

$$E_m = -\mu B \frac{m}{I}, \quad m = -I, \ldots, I \qquad [13]$$

where μ and I are the magnetic moment, and spin of the nuclear state respectively, B is the magnetic field at the nucleus, and m is the spin-projection quantum number. The 14.4125 keV transition of ^{57}Fe has M1 multipolarity, and the individual transitions shown in **Figure** 8 follow dipole selection rules, that is, $|m_e - m_g| \leqslant 1$. This results in a characteristic six-line pattern.

In addition to the interactions discussed above, one expects a renormalization of the energy of nuclear levels due to the presence of *s*-electrons in the nuclear volume. The observable result is a slight shift of the transition energy that is proportional to the density of the *s*-electrons in the nuclear volume. This so-called isomer shift can only be measured by comparison of different materials, for example, for ^{57}Fe the isomer shift is normally given relative to ^{57}Fe nuclei in iron metal at ambient conditions.

Index of Refraction

SMS is an elastic coherent process, and all possible nuclear sublevels can be excited simultaneously by a short X-ray pulse with sufficient bandwidth. The different nuclear transitions are then analogous to a set of oscillators with slightly different energies that are excited in phase at a given instant. Shortly, after the X-ray pulse has passed, the de-excitation begins, also in phase, but soon the increasing phase difference leads to destructive interference and then again to a constructive interference. This process leads to oscillations in the emission of the scattered radiation that depend directly on the nuclear-level splitting and thus on the hyperfine interactions. In a transmission geometry, SMS is usually described by an energy-dependent contribution to the index of refraction

$$n(E) = \frac{\hbar c}{2E} \rho \sigma_N f \sum_l \frac{W_l}{z_l(E) - i} \qquad [14]$$

where ρ is the number density of resonant nuclei, σ_N is the nuclear resonant cross section, and f is the Lamb–Mössbauer factor. The sum is over all sublevels of nuclear ground and excited states. The function $z_l = 2(E_l - E)/\Gamma$ depends on the energy difference between excited and ground states E_l and the nuclear level width Γ. The weight of each resonance is given by W_l with the normalization $\sum_l W_l = 1$. The weights depend on the directions of the hyperfine fields relative to the direction and the polarization of the SR. For simplicity, their tensor character is ignored here. Then, the transmission of monochromatic X-rays of energy E through a sample of thickness D is given by

$$T(E) = T_0 \exp\left[-\eta \sum_l \frac{W_l}{z_l^2(E) + 1}\right] \qquad [15]$$

where $\eta = \rho \sigma_N f D$ is called "effective thickness" of the platelet. The factor T_0 accounts for electronic absorption. The time response to an SR pulse is expressed by

$$\frac{dI}{dt} = T_0 \tau \left| \mathscr{F}\left(\exp\left[i \frac{2\pi D}{\lambda} n(E) \right] - 1 \right) \right|^2 \qquad [16]$$

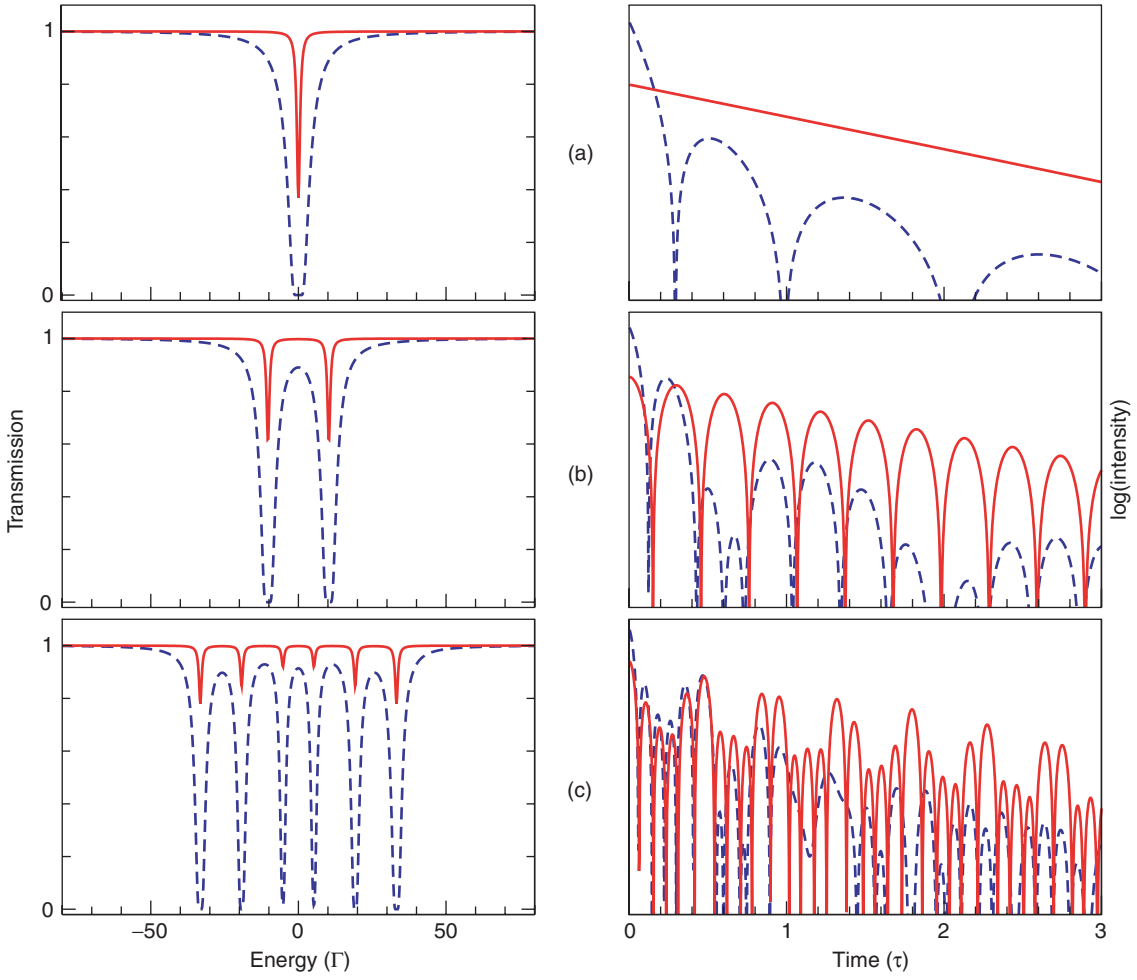


Figure 9 SMS spectra in energy and time. Calculations used eqns [15] and [16] with small ($\eta = 1$, solid lines) and large ($\eta = 50$, dashed lines) effective thicknesses. Panel (a) displays the case without hyperfine fields; panels (b) and (c) show spectra characteristic for an EFG and a magnetic field, respectively. Time spectra (right side) were normalized to identical area.

where λ is the X-ray wavelength and $\mathscr{F}$ symbolizes a Fourier transformation. The primary application of SMS is the determination of hyperfine interaction parameters. The transmission spectra are easier to interpret if the effective thickness is small. Large values of η lead to a self-absorption effect, that is, a broadening of the absorption lines in the energy spectra and a faster decay in the time spectra accompanied by aperiodic oscillations. The influence of the effective thickness is demonstrated in **Figure 9** by calculations using eqns [15] and [16].

Experimental Procedure

A schematic of the typical experimental setup that can be found at third-generation SR facilities is shown in **Figure 10**. The X-ray source consists of electron bunches that are orbiting in the storage ring and periodically pass through an undulator. The

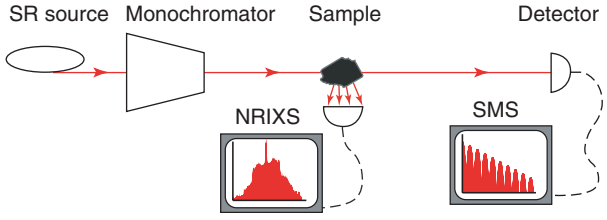


Figure 10 Experimental setup for NRIXS and SMS.

X-rays are monochromatized in two steps using a premonochromator and a high-resolution monochromator.

For NRIXS measurements, the energy bandwidth determines the resolution of the phonon spectra of the samples. The high-resolution monochromator is tuned around the nuclear transition energy, and the X-rays excite the nuclei in the sample. The re-emitted radiation is observed with an avalanche photodiode

detector that is placed as close as possible to the sample but away from any strong coherent scattering directions. The integrated delayed counting rate is recorded.

For SMS measurements, the energy bandwidth should be as small as practicably achievable with reasonable efficiency. The high-resolution monochromator is tuned to the nuclear transition energy and kept as stable as possible. X-rays that are transmitted through the sample excite the resonant nuclei coherently and are observed with an avalanche photodiode detector that is placed far enough, away from the sample to avoid contamination from incoherent scattering. The delayed events are mapped as a function of elapsed time between arrival of an SR pulse and detection of transmitted X-ray photon – this constitutes the time spectrum of the nuclei in the sample.

Acknowledgments

The author is grateful for valuable discussions with J M Jackson and T S Toellner. This work was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

See also: Scattering, Resonant; Synchrotron Radiation.

PACS: **76.80.+y; 63.20.−e; 07.85.Qe; 13.40.−f**

Further Reading

Gerdau E and de Waard H (eds.) (1999/2000) Nuclear resonant scattering of synchrotron radiation. *Hyperfine Institute*, vol. 123–125.

Sturhahn W (2004) Nuclear resonant spectroscopy. *Journal of Physics: Condensed Matter*, 16: S497–S530.

Scattering, Rayleigh

R Piazza, Politecnico di Milano, Milan, Italy
V Degiorgio, Università di Pavia, Pavia, Italy

Introduction

Rayleigh scattering (RS) is the elastic spreading of electromagnetic radiation from bound electrons, after they have been excited to virtual states far from resonances. Therefore, it differs from resonant fluorescence, molecular Raman scattering, and Brillouin scattering in condensed media, which are discussed elsewhere in this encyclopedia. Thomson scattering can be seen as a particular case of RS from free electrons, but is not further discussed in this article (neither is its relativistic Compton counterpart). Molecular RS plays a primary role in atmospheric optics accounting for the blue color of the sky and a large number of optical phenomena ranging from rainbows to fogbows, glories, halos, coronae, blue sun and moon, and the white color of clouds and snow, to cite only a few. A comprehensive discussion of atmospheric scattering has been made by M J G Minnaert in 1954 and D K Lynch and W Livingston in 1995.

Yet, probably, the most interesting applications of RS concern scattering from mesoscopic objects, either in the form of aerosols or suspensions of colloidal particles with a size ranging from a few nanometers to a few micrometers. Scattering from pollutant airborne particles has, for instance, crucial consequences on re-irradiation of solar energy from the lower atmosphere. Measurements of the scattering properties of dispersed particles yield primary information on colloid size, morphology, and interactions. Chances of obtaining photonic bandgap crystals often rely on the peculiar scattering properties of metal-coated colloids. Finally, by exploiting radiation forces associated to scattering, particles can be easily manipulated using noninvasive "optical tweezers," allowing one to design template surfaces of interest in photonics. Although, fluid systems are primarily dealt with, it is also important to point out that the attenuation due to RS sets severe limits to the useful wavelength range for light propagation in optical fibers.

This article is organized as follows. After a basic analysis of elastic molecular scattering, the key role of fluctuations in RS from condensed media is pointed out, and a fluctuation approach to shortly derive RS from simple fluids and mixtures is applied. The final and largest part is devoted to particle RS and its exploitation. Besides specific applications, however, the main message of which an attempt is made to convey is the following: "every deviation of light from rectilinear propagation, including diffraction and even geometrical refraction, can be ultimately seen as a consequence of Rayleigh scattering." As will be seen, the optical theorem, linking the refractive index to scattering properties, and the large-particle limit of Mie scattering, yielding classical diffraction and eventually geometric optics, justify this strong statement.

Molecular Rayleigh Scattering

First, recall the basic geometry of a scattering experiment. As incident field, what will always be taken is a plane wave having a wave vector $\boldsymbol{k}_\mathrm{i}$, with $|k_\mathrm{i}| = k = 2\pi/\lambda$ (where λ is the wavelength in the medium), linearly polarized along $\hat{\boldsymbol{n}}_\mathrm{i}$. Measurements consist in detecting the light scattered along the direction set by the wave vector $\boldsymbol{k}_\mathrm{s}$ that, together with $\boldsymbol{k}_\mathrm{i}$, fixes the scattering plane. Since Rayleigh scattering is elastic, one must have $|k_\mathrm{s}| = k$, and therefore the scattering vector $\boldsymbol{q} \equiv \boldsymbol{k}_\mathrm{i} - \boldsymbol{k}_\mathrm{s}$ has magnitude $q = 4\pi/\lambda \sin \vartheta/2$, where ϑ is the angle in the scattering plane between $\boldsymbol{k}_\mathrm{i}$ and $\boldsymbol{k}_\mathrm{s}$ (the azimuth angle with respect to the normal to the scattering plane will conversely be denoted by φ). In some measurements, a specific polarization component of the scattered field along $\hat{\boldsymbol{n}}_\mathrm{s}$ may be selected with a polarizer.

Probably, the simplest way to estimate elastic scattering from molecules is the elegant dimensional argument originally devised by Rayleigh in 1871 that, restated in modern terms, goes as follows: "Provided that the molecular dimensions are much smaller than λ, so that the scatterer essentially "sees" a spatially uniform, time-varying field, the scattered and incident fields will be proportional." Their dimensionless ratio $E_\mathrm{s}/E_\mathrm{i}$ may, in principle, depend on λ, the velocity of light c, the volume V of the scatterer, the dielectric constants $\varepsilon_\mathrm{P}, \varepsilon_\mathrm{M}$ of the materials constituting the particle and the surrounding medium, and the distance r of the observation point from the scattering volume. Since c is the only quantity that contains time, it should not appear in $E_\mathrm{s}/E_\mathrm{i}$. Similarly, ε_P and ε_M must appear only as a homogeneous function f of the ratio $\varepsilon_\mathrm{P}/\varepsilon_\mathrm{M}$. Since molecular polarizability α is linearly proportional to V, so must be E_s. Finally, in the radiation zone the scattered field will be a spherical wave, so that $E_\mathrm{s} \propto r^{-1}$. Therefore, the simplest dimensionless combination is $E_\mathrm{s}/E_\mathrm{i} = Vr^{-1}\lambda^{-2} f(\varepsilon_\mathrm{P}/\varepsilon_\mathrm{M})$, and for the scattered intensity one must have

$$I_\mathrm{s} = \frac{V^2}{r^2\lambda^4} f^2\left(\frac{\varepsilon_\mathrm{P}}{\varepsilon_\mathrm{M}}\right) I_0 \qquad [1]$$

where I_0 is the incident intensity, giving both a strong preferential scattering of short wavelengths (accounting for the hue of the sky and for the sun reddening at dawn and setting) and the remarkable dependence on molecular volume (yielding, for spheroidal molecules, a sixth power dependence on the radius).

For a molecular scatterer of size $d \ll \lambda$, a formal analysis can be performed by solving the Helmoltz equation for the vector potential in a far-field (radiation) zone $r \gg \lambda \gg d$, expanding the solution in multipoles, and retaining only the lowest (electric dipole) order. Within this approximation, the incident field induces a dipole $\boldsymbol{p} = \boldsymbol{\alpha}\boldsymbol{E}_\mathrm{i}$, where in general, the molecular polarizability $\boldsymbol{\alpha} = \alpha_{ij}$ is a tensor (so that $\boldsymbol{p}$ and $\boldsymbol{E}_\mathrm{i}$ are not necessarily parallel). The induced dipole irradiates a spherical wave at the same frequency and with amplitude given by

$$\boldsymbol{E}_\mathrm{s}(\boldsymbol{r}) = \frac{k^2}{4\pi\varepsilon_0} (\hat{\boldsymbol{r}} \times \boldsymbol{p} \times \hat{\boldsymbol{r}}) \frac{\exp(\mathrm{i}kr)}{r} \qquad [2]$$

If, for simplicity, one assumes that the molecule has a scalar polarizability α, so that $p \equiv |\boldsymbol{p}| = \alpha E_\mathrm{i}$, the scattered intensity (irradiance) is given by

$$I_\mathrm{s} = \frac{\varepsilon_0 c}{2} \overline{\boldsymbol{E}_\mathrm{i}^* \boldsymbol{E}_\mathrm{i}} = \frac{\pi^2\alpha^2 \sin^2\gamma}{\varepsilon_0^2\lambda^4 r^2} I_0 \qquad [3]$$

where γ is the angle between $\boldsymbol{p}$ and the direction of propagation of the scattered field. Equation [3] is consistent with the general prediction of eqn [1], since α is proportional to the scattering volume and, as later discussed in the text, depends only on the particle and solvent dielectric constants.

The simple classical analysis that was made is fully adequate for treating most of the problems involving molecular RS. For what concerns quantum effects, what is first noticed is that RS is a two-photon process requiring full quantization of the radiation field. A simple quantum analysis of RS can be made starting from the two-level Bloch equations, or from a more general multilevel perturbation theory leading to the Kramers–Heisenberg (KH) formula, which has the advantage of treating on the same footing as Rayleigh, Thomson, and inelastic Raman scattering, allowing explicit evaluation of the scattering cross sections from the matrix elements (while phenomenological coefficients are needed in the classical approach). The KH formula also yields quantum corrections to classical RS cross sections: however, far from resonances the latter are negligible, so that a fully classical approach is used in what follows.

Scattering and Fluctuations

The former single-scatterer approach to RS may however be misleading, missing a fundamental point: indeed, a uniform distribution of molecules will not scatter light at all, since the phases of the scattered fields are uniformly distributed in $[0, 2\pi]$. Therefore, the total scattered field vanishes for all except the forward direction where, adding to the incident beam, leads to a phase delay of the transmitted beam, modifying (as shall be better seen below) the refractive index of the medium. Exceptions are ordered

arrays of the scatterers, where Bragg peaks appear due to breaking of continuous translational symmetry. Scattering therefore always requires fluctuations. A very general approach to the problem can be made by treating the scattering medium as a continuum where, in order to account also for optical anisotropy effects, local dielectric (and therefore, refractive index) fluctuations are introduced by means of a tensor $\delta\boldsymbol{\varepsilon}(\boldsymbol{r},t)$, and deriving a wave-propagation equation in weakly inhomogeneous media. By assuming that the scattered field is much smaller than the incident field, and decomposing the dielectric fluctuations in spatial Fourier components as

$$\delta\boldsymbol{\varepsilon}(\boldsymbol{q},t) = \int \mathrm{d}V \delta\boldsymbol{\varepsilon}(\boldsymbol{r},t)\,\exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}) \qquad [4]$$

where r gives the position of the scattering element with respect to an arbitrary origin O in the scattering volume. The instantaneous scattered field amplitude reaching a detector placed at a distance R from O the scattering volume, after passing through a polarizer with transmitting axis set along $\hat{\boldsymbol{n}}_\mathrm{f}$, is given by

$$E_\mathrm{s}^{if}(R,t) = -\frac{k^2 E_0}{4\pi\varepsilon_0 R}\,\delta\varepsilon_{if}(q,t)\,\exp[\mathrm{i}(k_\mathrm{s}R-\omega t)] \quad [5]$$

where $\delta\varepsilon_{i,f} \equiv \hat{\boldsymbol{n}}_\mathrm{i}\cdot\delta\boldsymbol{\varepsilon}(\boldsymbol{q},\boldsymbol{t})\cdot\hat{\boldsymbol{n}}_\mathrm{f}$. The time-averaged detected intensity is

$$I_\mathrm{s}^{if}(\boldsymbol{q}) = \frac{k^4 I_0}{16\pi^2\varepsilon_0^2 R^2}\langle|\delta\varepsilon_{if}|^2\rangle \qquad [6]$$

and is therefore proportional to the mean square fluctuation of the dielectric constant with the same spatial frequency.

In order to find a connection with the molecular approach, which is particularly useful when dealing with scattering from dispersed particles, it can simply be assumed that the local dielectric constant depends explicitly only on the local molecular number density $\rho(\boldsymbol{r},t)$ (see the next section for a discussion) writing $\delta\boldsymbol{\varepsilon}(\boldsymbol{r},t) = \boldsymbol{\alpha}(t)\rho(\boldsymbol{r})$, where $\boldsymbol{\alpha}(t)$ is the molecular polarizability tensor, depending on time because of molecular rotations. Expressing the local density fluctuations as a sum of delta functions over the molecular positions $\boldsymbol{r}_i(t)$ as $\delta\rho(\boldsymbol{r},t) = \sum_i \delta(\boldsymbol{r}-\boldsymbol{r}_i) - \bar{\rho}$, where $\bar{\rho}$ is the average number density, and with $\alpha_{if}(t) \equiv \hat{\boldsymbol{n}}_f\cdot\boldsymbol{\alpha}(\boldsymbol{q},t)\cdot\hat{\boldsymbol{n}}_f$, one can immediately see that

$$\delta\varepsilon_{if}(\boldsymbol{q},t) = \alpha_{if}(t)\left\{\sum_i \exp[\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_i(t)] - N\delta(q)\right\} \qquad [7]$$

where the second term (coming from $\bar{\rho}$) is simply the $q=0$ contribution from the N molecules. Provided that molecular orientations and translations are uncorrelated (so that the amplitude and phase averages factorize), from eqn [6], one therefore obtains

$$I_\mathrm{s}^{if}(\boldsymbol{q}) = \frac{k^4 N I_0}{16\pi^2\varepsilon_0^2 R^2}\langle|\alpha_{if}|^2\rangle S(q) \qquad [8]$$

where

$$S(q) = N^{-1}\sum_{i,j}\langle\exp[\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)]\rangle \qquad [9]$$

is the static structure factor. To make things clearer, an ideal gas where particle positions are uncorrelated is considered. Then $\langle\exp[\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)]\rangle = \delta_{i,j}$, so that $S(q)=1$ and I is proportional to N. It is noticed that this apparently "trivial" scaling with N, holds true only because (Poisson) density fluctuations in an ideal gas are proportional to $\sqrt{N}$: for interacting systems, it is no longer so. It is important to recall that the typical q-range over which structural features of simple fluids develop is much larger than the range probed by light scattering, which therefore essentially measures only $S(0) = Vk_\mathrm{B}T\rho^2\chi_T$, where χ_T is the isothermal compressibility (the situation is however very different for suspensions of mesoscopic particles, where the structural spatial scales are often of the order λ).

Rayleigh Scattering from Pure Fluids and Fluid Mixtures

In liquids, fluctuations of the local dielectric constant can be expanded in terms of any two thermodynamic variables whose fluctuations are uncorrelated, for instance, temperature T and density ρ. Since experimentally, for most simple liquids, $(\partial\varepsilon/\partial T)_\rho \approx 0$, RS is in fact mostly due to density fluctuations. In order to derive RS effects in simple fluids, it is however more convenient taking pressure p and entropy S as independent variables, and using a fluctuating hydrodynamics approach. This amounts to linearizing the continuity for small fluctuations, Navier–Stokes, and heat equations and finding their normal modes. Since in the continuity equation ρ is coupled to the fluid velocity $\boldsymbol{v}$ only through $\nabla\cdot\boldsymbol{v}$, density fluctuations are uncoupled to transverse velocity modes, which, therefore, cannot be detected by light scattering. In addition to two propagating pressure modes (corresponding to absorption or excitation of sound waves, giving rise to Brillouin scattering), the solution of linearized mode equations yields a thermal diffusive mode, corresponding to the frequency-unshifted Rayleigh peak having a width $\Delta\omega = D_\mathrm{T}q^2$, where D_T is the thermal diffusivity,

and an amplitude proportional to the difference $c_p - c_v$ of the specific heats at constant pressure and volume. For many liquids, and in particular for water, $c_p - c_v$ is rather small, and therefore RS is fairly weak (at variance to a widespread belief, the bluish appearance of thick water layers on ice blocks is not due primarily to scattering, but to a much stronger absorption effect due to an overtone vibrational peak close to the red edge of the visible spectrum; D_2O, for which the absorption shifts to higher λ due to isotope effects, indeed is much more transparent).

In binary liquid mixtures new modes appear due to the presence of an additional conserved variable, namely the concentration c of one of the two species, adding a hydrodynamic mass diffusion equation. The analysis is much simpler as discussed by B J Berne and R Pecora in 1976 if pressure fluctuation, and therefore effects on Brillouin peaks, are neglected, so that the problems amount to solving by Laplace transforming the linearized form of the coupled equations for the fluctuations δc, δT

$$\begin{aligned}\frac{\partial(\delta c)}{\partial t} &= D\left[\nabla^2(\delta c) + \frac{k_T}{T}\nabla^2(\delta T)\right] \\ \frac{\partial(\delta T)}{\partial t} - \frac{1}{c_p}k_T\left(\frac{\partial\mu}{\partial c}\right)\frac{\partial(\delta c)}{\partial t} &= D_T\nabla^2(\delta T)\end{aligned} \qquad [10]$$

where D is the mass diffusion coefficient, μ is the chemical potential of the mixture, and k_T is the thermal diffusion ratio, quantifying coupling of mass and heat transport through the so-called Soret effect. Solution of eqn [10] yields two diffusive modes where thermal and concentration fluctuations are, in general, mixed: in practice, however, since $D \ll D_T$ (for most liquids $D_T \simeq 10^{-3}\ \mathrm{cm^2\,s^{-1}}$, while D varies between $10^{-5}\ \mathrm{cm^2\,s^{-1}}$ for simple mixtures and $10^{-6} - 10^{-8}\ \mathrm{cm^2\,s^{-1}}$ for macromolecular solutes) the two modes are almost decoupled. The peak due to concentration fluctuation has an extremely narrow width $\Delta\omega = Dq^2$ and an amplitude proportional to $(\partial\mu/\partial c)^{-1}$ (for dilute solutions, this means that RS is proportional to the osmotic compressibility).

Scattering from Dispersed Particles

Statement of the Problem

The important case of elastic scattering from dispersed particles is discussed here. Only particles made of an optically isotropic material, that is with a scalar polarizability are considered first. In a far field, the scattered field $E_s(r,t)$ is an amplitude-modulated spherical wave, which can be written as

$$E_s(r,t) = S(\vartheta,\varphi)\,\frac{\exp[i(kr-\omega t)]}{ikr}\,E_0 \qquad [11]$$

where a term i has been factorized out, just for convenience, and the dimensionless scattering function $S(\vartheta,\varphi)$ has been introduced, which is, in general, a complex quantity: $S \equiv F(\vartheta,\varphi)\exp[i\phi(\vartheta,\varphi)]$. The scattered intensity can, therefore, be written in terms of the incident intensity I_0 as

$$I = \frac{F(\vartheta,\varphi)^2}{k^2r^2}\,I_0 \qquad [12]$$

By integrating $r^2 I_s$ over the solid angle Ω, the scattering cross section is therefore found to be

$$\sigma_s = \frac{1}{k^2}\int \mathrm{d}\Omega F(\vartheta,\varphi)^2 \qquad [13]$$

If, besides scattering, the particle absorbs part of the radiation with an absorption cross section σ_A, the total extinction cross section will be given by $\sigma_{ext} = \sigma_s + \sigma_A$.

The previous definition can be generalized to account for scattering from optically anisotropic particles by writing

$$\begin{bmatrix} E_s^{\perp} \\ E_s^{\parallel} \end{bmatrix} = \frac{\exp[ik(r-z)]}{ikr}\begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}\begin{bmatrix} E_i^{\perp} \\ E_i^{\parallel} \end{bmatrix} \qquad [14]$$

where z is the particle coordinate along the direction of the incident beam with respect to an arbitrary origin.

Specific forms for the scattering matrix (S) are dictated by particle symmetry as discussed by H C van de Hulst in 1981: for instance, it is diagonal for spherical particles made of an optically isotropic material and symmetric for nonchiral particles.

The Optical Theorem (OT)

A simple, but extremely useful relation connects the total scattering cross section to the scattering amplitude in the forward direction (the direction of the incident beam, once again taken as z). Forward scattering is rather peculiar. At any finite angle, the total scattering from N spatially uncorrelated particles is a vector sum of random components, and therefore a Gaussian complex random variable. In full analogy with what has been seen for RS from ideal gas, the total scattered field is then proportional to $\sqrt{N}$ and the scattered intensity to N. This means that the individual particle cross sections add, $\sigma_s = \sum \sigma_s^{(i)}$. However, it is easy to show that all scattering contributions in the forward direction add in phase, so

that one must sum field amplitudes, not intensities, with $S(0) = \sum S^{(i)}(0)$. van de Hulst, in 1981, has shown that the forward scattering contribution for a thin slab of thickness δ containing ρ particles per unit volume is given by $dE_s = -[2\pi\rho\delta S_0/k^2]E_i(z,t)$, where S_0 is the value of the scattering function in the forward direction. The propagation equation for the spatial part of the total field $E = E_i + E_s$ is therefore

$$\frac{dE(z)}{dz} = \left(ik + \frac{2\pi\rho}{k^2} S_0\right) E(z) \qquad [15]$$

yielding

$$E(z) = E_0 \exp[ikz(1 + i2\pi\rho S_0/k^3)] \qquad [16]$$

Identifying eqn [16] with the general expression $E(z) = E_0 \exp(i\tilde{n}kz)$, where $\tilde{n} = n + in'$ is the complex refractive index, one has

$$\begin{aligned} n &= 1 + (2\pi\rho/k^3)\ \mathrm{Im}(S_0) \\ n' &= -(2\pi\rho/k^3)\ \mathrm{Re}(S_0) \end{aligned} \qquad [17]$$

connecting the refractive index of the medium to the presence of scattering. Provided that the particles are nonabsorbing, the transmitted intensity is given by $I(z) = I_0 \exp[-2k\ \mathrm{Im}(\tilde{n})]$, so that the extinction coefficient (power loss) per unit volume is $\gamma = 2kn'$. The extinction cross section is finally given by $\sigma_{ext} = \gamma/\rho$, one finally has

$$\sigma_{ext} = \frac{\gamma}{\rho} = -\frac{4\pi}{k^2}\ \mathrm{Re}(S_0) \qquad [18]$$

which is the usual form of the OT. Apparently, this rather curious result, summing up scattering in all directions only in terms of zero-angle scattering, has been independently "rediscovered" and extended by many authors, totally unaware of previous results, both for classical electromagnetic scattering and in quantum electrodynamics.

Scattering Regimes

The full solution of the electromagnetic scattering for particles of arbitrary size and composition is very elaborate, even for the simplest case of homogeneous spheres. It is however useful considering some specific approximation yielding simple explicit expressions for the scattering function. Both the external incident field and the internal field within the particle are required. The refractive indices of the particle and of the surrounding medium ("solvent") shall be denoted by n_P and n_s, the wavelength of the incident radiation in vacuum by λ_0 (so the wave vector is $k_{ext} = 2\pi n_s/\lambda_0$ for the external field, and $k_{int} = 2\pi n_P/\lambda_0$ within the particle), and a characteristic particle size by a.

Particles small compared to the wavelength This is what is called the Rayleigh regime in the strict sense. Both the external and the internal fields are required to be essentially uniform over the particle size. This means that the phase shifts $k_{ext}a$ and $k_{int}a$ have to be small, or $a \ll \lambda_0/2\pi n_s, a \ll \lambda_0/2\pi n_P$ (the second condition is generally more restrictive-it is to be noted in particular that it excludes strongly absorbing particles, whatever their size). The situation is then completely analogous to molecular Rayleigh scattering. If the particle is made by an optically isotropic material so that its polarizability is scalar, the amplitude of the scattered field is given by eqn [3] while, by integrating eqn [3] and dividing by the incident intensity, the scattering cross section is found as

$$\sigma_s = k^4\alpha^2/6\pi\varepsilon_0^2 \qquad [19]$$

A note of caution however concerns particle polarizability, which has to be interpreted as the excess optical polarizability with respect to an equal solvent volume. Noticing that, when the incident field is vertically polarized (perpendicular to the scattering plane), $\gamma \equiv \pi/2$, while if it is horizontally polarized $\gamma \equiv \pi/2 - \vartheta$, the full scattering matrix is found to be

$$\boldsymbol{S} = \frac{ik^3\alpha}{4\pi\varepsilon_0} \begin{bmatrix} 1 & 0 \\ 0 & \cos\vartheta \end{bmatrix} \qquad [20]$$

Specific expressions for polarizability can be found by solving the electrostatic problem for a uniform field. For instance, for dielectric spheres of radius a, made of a material with dielectric constant ε_P,

$$\alpha = 4\pi\varepsilon_0 \frac{\varepsilon_P/\varepsilon_s - 1}{\varepsilon_P/\varepsilon_s + 2} a^3 = 4\pi\varepsilon_s \frac{n_P^2 - n_s^2}{n_P^2 + 2n_s^2} a^3 \qquad [21]$$

Since $S(0) = S_{1,2} = ik^3\alpha/4\pi\varepsilon_0$, from the OT one has $\sigma_{ext} = (k/\varepsilon_0)\ \mathrm{Im}(\alpha)$, so one should have reduction of the incident power only for absorbing particles. This apparent contradiction only means that to evaluate total cross sections via the OT, one must go beyond the Rayleigh approximation, so to obtain a nonvanishing real part of $S(0)$.

The previous result can be extended to small optically anisotropic particles. For simplicity, it is assumed that the particle is made of a uniaxial birefringent material. The polarizability tensor, referred to the particle optical principal axes X, Y, Z, has only two independent components, that shall be denoted as $\alpha_{||} \equiv \alpha_X$ and $\alpha_\perp \equiv \alpha_Y = \alpha_Z$. Assuming for

convenience x as the direction of the scattered field and z as the normal to the (x, y) scattering plane, the two most common experimental configurations are the VV and VH geometries. In both of them the incident field is vertically polarized, but while in the VV scheme the vertical component I_{VV} of the scattered intensity is measured, the orthogonal ("depolarized") I_{VH} component is detected in the VH scheme. I_{VV} and I_{VH} can be found by evaluating $\hat{\boldsymbol{n}}_z \cdot \boldsymbol{\alpha} \cdot \hat{\boldsymbol{n}}_z$ and $\hat{\boldsymbol{n}}_y \cdot \boldsymbol{\alpha} \cdot \hat{\boldsymbol{n}}_z$, after the particle polarizability tensor has been re-expressed in the laboratory fixed frame, in terms of the average particle polarizability $\bar{\alpha} = (\alpha_{||} + 2\alpha_{\perp})/3$ and of the particle optical anisotropy $\beta = \alpha_{||} - \alpha_{\perp}$. Defining $\bar{n}_P^2 = \bar{\alpha}/V + 1$ and (provided that the optical anisotropy is small) $\Delta n = \beta/2\bar{n}V$, and assuming that $\bar{n}_P \approx n_s$, the final expression, after subtracting the solvent polarizability $\alpha_s = (n_s^2 - 1)V$, is

$$\begin{aligned} I_{VV} &= CV\bar{n}_P^2[(\bar{n}_P - n_s)^2 + (4/45)(\Delta n)^2] \\ I_{VH} &= CV\bar{n}_P^2(\Delta n)^2/15 \end{aligned} \qquad [22]$$

where C does not depend on particle or solvent optical properties. Notice that both I_{VH} and the second term on the right-hand side of the expression for I_{VV} do not depend on n_s, so that scattering is present even in "best index matching" conditions $n_s = \bar{n}_P$. These contributions are the direct analogs, in depolarized Rayleigh scattering, of the incoherent cross section for neutron scattering. **Figure 1** shows the experimental results for optically anisotropic particles in water.

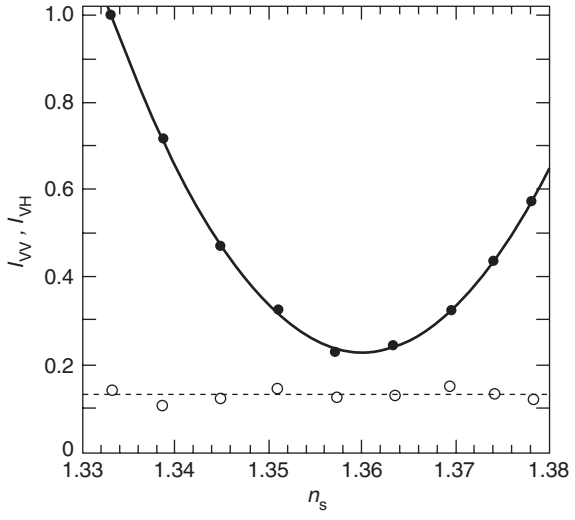


Figure 1 Experimental values for I_{VV} (•) and I_{VH} (○) (normalized to their values for water suspensions) as a function of the solvent refractive index n_s for optically anisotropic colloidal spheres made of the partially crystalline polymer perfluoro-alkylvinylether (PFA). The parabolic fit to I_{VV} yields an average particle refractive index $\bar{n}_P \simeq 1.36$ and an optical anisotropy $\Delta n \simeq 0.03$. As discussed in the text, I_{VH} does not depend on n_s.

Rayleigh–Gans–Debye scattering Most dispersed particles do not fulfill the strict requirements of the "pure" Rayleigh regime. A much more useful approximation can be obtained by letting the incident field vary, even appreciably, over the particle length scale, but requiring that "the internal field does not appreciably differ from the incident one," both in amplitude and in phase. Weak amplitude changes simply imply $|n_P - n_s| \ll 1$, while tuning of the phases requires the optical path with or without the particles to be comparable, or

$$\frac{2\pi}{\lambda}|n_P - n_s|a \ll 1 \qquad [23]$$

which, for $a > \lambda/2\pi$, is the more restrictive criterion framing the Rayleigh-Gans-Debye (RGD) regime. Notice that even particles with size $a \gg \lambda$ may satisfy eqn [23], provided that they are sufficiently "faint." Equating the internal and external fields amounts to neglecting effects on any given volume element δV of the fields scattered by the surrounding, so that the RGD regime is the equivalent of the first Born approximation in quantum scattering. Each δV can, therefore, be considered as an independent elementary Rayleigh scatterer, and the total scattered field is obtained by summing over all volume elements, taking into account the phase of each scattering contribution (note that the phase shift of a wave scattered by an element placed at $\boldsymbol{r}$ compared to the contribution from a volume element centered around an arbitrary origin is $\Delta\varphi = \boldsymbol{k}_i \cdot \boldsymbol{r} - \boldsymbol{k}_s \cdot \boldsymbol{r} = \boldsymbol{q} \cdot \boldsymbol{r}$). Summing up over all volume elements, and allowing for nonuniformity of the particle refractive index by expressing $n_P = n_P(\boldsymbol{r})$, one obtains

$$\begin{bmatrix} S_1(\vartheta) \\ S_2(\vartheta) \end{bmatrix} = \frac{\mathrm{i}n_s k^3 V}{2\pi} F(q) \begin{bmatrix} 1 \\ \cos\vartheta \end{bmatrix} \qquad [24]$$

where

$$F(q) = \frac{1}{V}\int_V \mathrm{d}^3 r[n_P(\boldsymbol{r}) - n_s]\exp(\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}) \qquad [25]$$

By introducing the "characteristic function" $\chi(\boldsymbol{r}) = 1$ if $\boldsymbol{r} \in V$ and zero otherwise, $F(q)$ is seen to be the spatial Fourier transform of the distribution of the refractive index mismatch. For the special case of a particle with uniform refractive index distribution and vertical incident polarization, the scattered

intensity is then given by

$$I(q) = I_0 \frac{V^2 k^4}{4\pi^2 R^2} (n_P - n_s)^2 P(q) \qquad [26]$$

where

$$P(V) = \frac{1}{V^2} \left| \int_V \exp(i\boldsymbol{q} \cdot \boldsymbol{r}) \right|^2 \qquad [27]$$

called the particle "form factor" is readily seen to represent an "intraparticle" structure factor. Scattering pattern in the RGD approximation can therefore be evaluated by a straight geometrical integration: for homogeneous spheres, for instance, $F(q)$ is simply proportional to the spherical Bessel function $J_{3/2}(qa)$, and an analytical expression can also be found for ellipsoids, long rods, and thin disks. RGD scattering considerably extends the class of particles for which an easy analysis is feasible. A particularly interesting situation is colloidal fractal aggregation induced by the addition of salt. Fractal clusters grow with time, soon becoming much larger than λ, but they also became more and more "tenuous." Therefore, their average refractive index becomes more and more matched with the solvent, and it can be shown that, asymptotically, RGD conditions are always fulfilled.

Mie scattering The exact solution for RS from an arbitrary sphere was obtained in 1908, with a real mathematical "tour de force," by Gustav Mie. Here only the main feature of the results are described, commenting on some general aspects. First of all, all solutions are fixed. The general solution is fixed only by two parameters, namely $m = n_P/n_s$ and the dimensionless size $x = ka = 2\pi n_s a/\lambda$, and is formally given by

$$\begin{aligned} S_1 &= -\sum_n \frac{2n+1}{n(n+1)} (a_n \pi_n + b_n \tau_n) \\ S_2 &= -\sum_n \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \pi_n) \end{aligned} \qquad [28]$$

where $\pi_n(\vartheta) = P_n^1(\cos\vartheta)/\sin\vartheta$ and $\tau_n(\vartheta) = \mathrm{d}P_n^1(\cos\vartheta)/\mathrm{d}\vartheta$ give, in terms of the associate Lagrange functions $P_n^1(\cos\vartheta)$, the angular distribution of the scattered light. The amplitude coefficients a_n and b_n are complicated functions of m and x. What is important to point out, however, is that their leading behavior is $a_n \sim x^{2n+1}$, $b_n \sim x^{2n+3}$, so that, for small particles, taking into account only the first terms is sufficient. By increasing n, the functions π_n and τ_n display in a polar diagram an increasing number of lobes, changing in direction and sign with n. For large particles, therefore, the sum of many terms tends to cancel out scattering at most angles. The only exceptions are the lobes that are present for all n around $\vartheta = 0°$ and $\vartheta = 180°$, the latter however alternating in sign. As a result, large particles tend to scatter predominantly forward, with a residual backscattering cone (so one gets dazzled by backlighted drops on a windshield). For very large particles, the forward lobe has a width $\Delta\vartheta \sim \lambda/a$, witnessing the merging of Mie scattering with the classical diffraction theory. **Figure 2** compares the full Mie and approximate RGD solutions for polystyrene latex spheres in water.

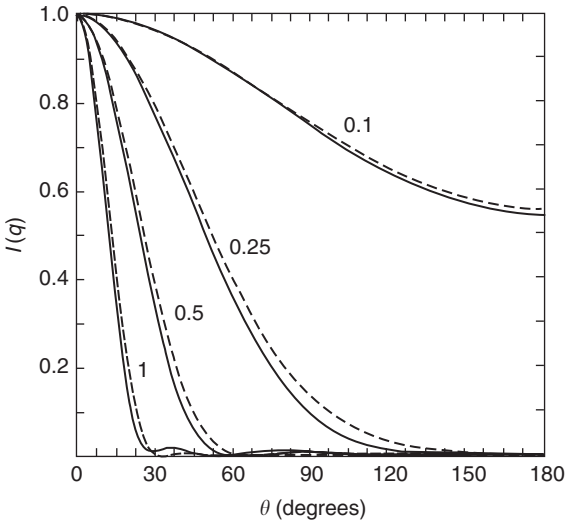


Figure 2 Normalized scattering intensity vs. scattering angle ϑ for polystyrene spheres ($n_P = 1.59$) in water for $\lambda_0 = 633$ nm. For each value of the ratio a/λ, indicated close to the curves, the full and dotted lines are, respectively, the numerical result from Mie theory and corresponding RGD analytical solution.

The coefficients a_n and b_n are finite for all real values of x. However, for strongly absorbing particles, they often present complex poles with very small values of the imaginary part. These "quasiresonances" lead to strong enhancement of specific modes, and are responsible for the beautiful colors often observed in suspensions of metallic colloids.

By exploiting the OT, the total extinction cross section can be obtained directly as a sum over the coefficients a_n, b_n:

$$\sigma_{ext} = \frac{2\pi}{k^2} \sum_n (2n+1)\, \mathrm{Re}(a_n + b_n) \qquad [29]$$

For small particles, σ_{ext} goes as λ^{-4} as in the RGD approximation, but by increasing x it tends to level off with substantial oscillations. In particular, for a specific (narrow) particles size range, it may happen that in the visible region σ_{ext} increases with λ (see **Figure 3**). This takes place for instance, when

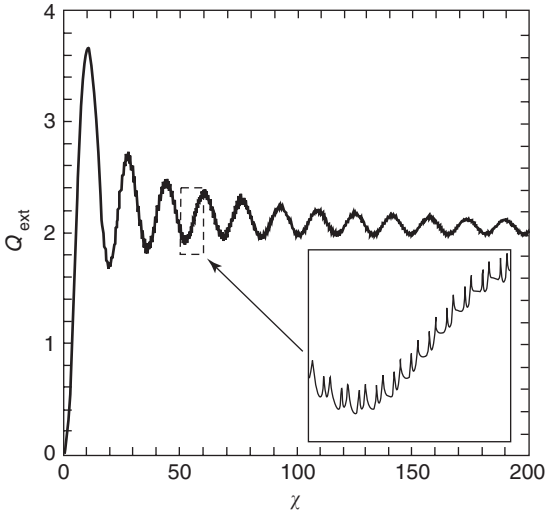


Figure 3 Extinction efficiency factor $Q_{ext} = \sigma_{ext}/\pi a^2$ for polystyrene spheres in water as a function of x (varied by fixing $\lambda =$ 560 nm and increasing a). The steep rise at small x corresponds to the RGD regime, $Q \sim x^4$, while the rich substructure at higher x, shown in the inset, derives from the large number of contributing Mie modes. "Blue moon" effects may take place in the regions where $dQ_{ext}/dx < 0$.

volcanic bursts eject huge quantities of large particles in the atmosphere, giving rise to the effect of a blue sun or moon (as the name suggests, a rare event indeed).

The strong propensity to forward scattering for large particles also allows one to explain a puzzling fact about extinction. As shown in **Figure 3**, in the limit $x \to \infty$, eqn [29] yields a value for σ_{ext} which is twice the particle geometric cross section $\sigma_g = \pi a^2$, that is, the particles subtract twice the expected power from the incident beam. This is not an artifact, but a real effect; what happens, however, is that half of the scattered power is concentrated in the forward lobe, so that it is collected by any finite-size detector as the eye. Such an intuitive geometrical optic concept as "shadow" is therefore based on subtle scattering effects.

A final consideration concerns transfer of momentum by radiation pressure, which is of primary importance both for optical levitation and for particle trapping by laser tweezers. For nonabsorbing particles, it is easy to show that the radiation pressure is given by $p_{rad} = I_0(\sigma_s/\sigma_g)(1 - \langle \cos\vartheta \rangle)$, where $\langle \cos\vartheta \rangle$ is averaged over the scattering pattern distribution. For $a \ll \lambda$, one obtains $\langle \cos\vartheta \rangle \simeq 0$, and the radiation pressure increases approximately as a^4. However, for very large particles $\sigma_s/\sigma_g \to 2$ and $\langle \cos\vartheta \to 1 \rangle$, so that radiation pressure decreases. Depending on the size and material composition of the particle, there is therefore an optimal λ for efficient momentum transfer.

The limit of geometrical optics A brilliant result by van de Hulst in 1981 yields a rigorous link between scattering and simple "ray optics": in the limit $x \gg 1$, a term of order n in the Mie expansion is mapped into a ray passing at a distance $(n+1)\lambda/2\pi$ from the particle center (the number of "distinguishable" rays must be finite: a single "ray," to be meaningful must be straight over a length $l \geqslant \lambda$, so that, because of diffraction, it must have a minimum spot area of the order of λ^2). The Mie coefficients can then be split into two equal parts leading, respectively, to diffuse the refraction–reflection effects and to the forward diffraction cone. Obviously, classic ray optics is found to be incorrect near caustics and focal point.

Dynamic Rayleigh Scattering

Rayleigh scattering from dispersed particles is not exactly elastic. Particles in suspensions undergo a Brownian motion, so that the phase (and, for optically anisotropic particles, also the amplitude see [7]) of the scattered field in eqn [2] depends on time. This slow phase modulation leads to a spectral broadening $\Delta\omega$ that is usually "extremely" small, typically $\Delta\omega/\omega \ll 10^{-9}$. Therefore, $\Delta\omega$ cannot be resolved even by sophisticated interferometric methods, since it is much smaller than the intrinsic spectral broadening on any source. A brilliant solution is looking at "intensity" fluctuations measured by a "quadratic" detector (photodiode, photomultiplier, CCD). Quadratic detection indeed introduced a "replica" of the spectrum shifted to the base band (as in homodyne radio systems) that can be easily extracted. Even more, this scheme does not require an incident light with strict temporal coherence (lasers are actually used, but only because of their brilliance). This clever idea stands at the root of "intensity correlation spectroscopy" (ICS, also called "dynamic light scattering"), allowing one to obtain the particle size from the behavior of the intensity time-correlation function $g(t) = \langle I(0)I(t) \rangle$ measured by a digital correlator. A comprehensive analysis of ICS has been made by B J Berne and R Pecora in 1976.

See also: Electrodynamics: Continuous Media; Interaction of Light and Matter; Scattering, Elastic (General); Scattering, Inelastic: Brillouin; Scattering, Inelastic: Raman; Small Particles and Clusters, Optical Properties of.

PACS: 42.25.Fx; 78.35.+c; 87.64.Cc; 94.10.Gb; 82.70.+y

Further Reading

Berne BJ and Pecora R (1976) *Dynamic Light Scattering*. New York: Wiley.

Bohren CF and Huffman DR (1983) *Absorption and Scattering of Light by Small Particles*. New York: Wiley.
Degiorgio V, Piazza R, Bellini T, and Visca M (1994) *Advances in Colloid and Interface Science* 48: 61–91.
Landau LD, Lifshitz EM, and Pitaevski LP (1990) *Statistical Physics*. New York: Pergamon Press.
Loudon R (2000) *The Quantum Theory of Light*, 3rd edn. Oxford: Oxford University Press.
Lynch DK and Livingston W (1995) *Color and Light in Nature*. Cambridge: Cambridge University Press.
Mie G (1908) *Annalen der Physik (Leipzig)* 25: 377–445.
Minnaert MJG (1954) *The Nature of Light and Colour in the Open Air*. New York: Dover.
Newton RG (1976) *American Journal of Physics* 44: 639–642.
Rayleigh L (1871) *Phil. Mag.* 41: 107–120. 274–279. (Reprinted in (1964) *Scientific Papers by Lord Rayleigh*, vol. I, pp. 1869–1881. New York: Dover.)
van de Hulst HC (1981) *Light Scattering by Small Particles*. New York: Dover.

Scattering, Resonant

C Vettier, Institut Laue Langevin, Grenoble, France

Introduction

In its simplest form, the scattering of X-rays by atoms and molecules can be approximated in terms of the classical scattering of electromagnetic radiation (X-rays) by a free electric charge which was first solved by J J Thomson. In this approximation, only electrons orbiting around the nuclei contribute to the scattering. Other charges, such as protons which make up the nuclei, are not considered here.

Electrons are generally located in atomic bound states when they form atoms. These bound states are at the origin of the resonant effects. Suppose that electrons are loosely bound to atoms and that they are distributed over a volume comparable with atomic dimensions and the wavelength of the X-ray radiation. In this case, the amplitude of scattering by an atom is given by the so-called atomic scattering factor, which contains information about the spatial density of the electronic shells around the nucleus. Atoms arranged in a crystalline array scatter coherently and the amplitudes scattered by different scattering centers can be added; the constructive interferences lead to diffraction peaks (Bragg scattering) whose positions and intensities are analyzed to reconstruct the structure of the scattering medium. If atoms are distributed at random, the scattered intensities are added; a similar situation occurs in the case of incoherent or Compton scattering where the radiation is scattered inelastically.

The following sections deal with the coherent X-ray scattering only. The general expression for elastic scattering of X-rays from a collection of atoms is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = r_\mathrm{e}^2 \left| \sum_n \mathrm{e}^{\mathrm{i}K\cdot r_n} f_n(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega) \right|^2$$

where $\boldsymbol{k}$ and $\boldsymbol{k}'$ are the incident and scattered wave vectors of the photons, $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ is the photon momentum transfer or scattering vector, and $\hbar\omega = \hbar c|\boldsymbol{k}|$ is the photon energy. $r_\mathrm{e} = e^2/mc^2 \approx 2.8 \times 10^{-15}$ m is the classical electron radius which measures the magnitude of the scattering. $f_n(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ represents the scattering amplitude of the nth atom (or atomic scattering factor) and is expressed in "electron units." A knowledge of $f_n(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ is a prerequisite to crystal structure determination when analyzing diffracted intensities. The resonant effects occur when the photon energy is varied through absorption edges, which correspond to transitions between electronic states. These transitions alter the $f_n(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ dramatically because the energy dependence of the atomic scattering factors $f_n(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ occurs in resonant denominators. The following sections describe these phenomena and their consequences for applications.

Coherent X-Ray Scattering

The complete derivation of X-ray scattering amplitudes requires the use of quantum principles. X-rays are represented by a quantized electromagnetic field. For an electron of mass m and spin $\boldsymbol{s}$ in a field of vector-potential $\boldsymbol{A}$ and scalar potential Φ, the appropriate form of the interaction Hamiltonian is

$$H = \frac{1}{2m}\left(p - \frac{e}{c}\boldsymbol{A}\right)^2 - \frac{e\hbar}{2mc}\,\boldsymbol{s}\cdot B - \frac{e\hbar}{2(mc)^2}\,\boldsymbol{s}\cdot E \times \left(p - \frac{e}{c}\boldsymbol{A}\right)$$

This Hamiltonian includes the electron kinetic energy and the interactions with the electromagnetic field, Zeeman coupling, and spin–orbit term. In the Zeeman term $B = \nabla \times \boldsymbol{A}$, while the spin–orbit term contains $E = -\nabla\Phi - (1/c)(\partial\boldsymbol{A}/\partial t)$. The vector potential $\boldsymbol{A}(\boldsymbol{r})$ is linear in the photon creation and annihilation operators. At $\boldsymbol{r}_j$, the position of the jth

electron, $\boldsymbol{A}(\boldsymbol{r})$ can be expanded in terms of the photon creation and annihilation operators, $C(\boldsymbol{k}\sigma)$ and $C^{+}(\boldsymbol{k}\sigma)$ with a quantization volume V:

$$\boldsymbol{A}(\boldsymbol{r}_j) = \sum_{q\sigma}\left(\frac{2\pi\hbar c^2}{V\omega_{\boldsymbol{k}}}\right)^{1/2}[\varepsilon_\sigma C(\boldsymbol{k}\sigma)\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} + \varepsilon_\sigma^{*}C^{+}(\boldsymbol{k}\sigma)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}]$$

where $\varepsilon_\sigma \equiv e(\boldsymbol{k}\sigma)$ denotes the two transverse unit polarization vectors (label $\sigma = 1, 2$) associated with the wave $\boldsymbol{k}$; $\hbar\omega_{\boldsymbol{k}}$ is the energy of the wave $\boldsymbol{k}$. Since a scattering event conserves the number of photons, the linear terms in $\boldsymbol{A}$ lead to scattering in the second order while only quadratic terms in $\boldsymbol{A}$ induce scattering in the first order.

The spin–orbit term in H is of the order of vc; therefore, the Hamiltonian H can be further reduced by omitting linear terms in $\boldsymbol{A}$ which contribute to the scattering in the second order.

Finally, the interaction between an electron and the electromagnetic field can be written as

$$H' = \frac{e^2}{2mc^2}\boldsymbol{A}^2 - \frac{e}{mc}p\cdot\boldsymbol{A} - \frac{e\hbar}{mc}\boldsymbol{s}\cdot\nabla\times\boldsymbol{A} - \frac{e^2\hbar}{2(mc^2)^2}\boldsymbol{s}\cdot\frac{\partial\boldsymbol{A}}{\partial t}\times\boldsymbol{A}$$

The first term, $H_1^{'}$, in the equation above corresponds to the classical Thomson scattering term. The second, $H_2^{'}$, which is linear in $\boldsymbol{A}$, is at the origin of the resonant or anomalous X-ray scattering. The last two terms, $H_3^{'}$ and $H_4^{'}$, which contain the spin of electrons lead to magnetic X-ray scattering.

Fermi's golden rule is used to calculate the scattering cross section or the scattering amplitude $f(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$. It is assumed that the total system photon + scattering medium evolves from the initial state $|i\rangle = |a\rangle|\boldsymbol{k}\sigma\rangle$ with photon $|\boldsymbol{k}\sigma\rangle$ and the medium in a state of energy E_a to a final state $|f\rangle = |b\rangle|\boldsymbol{k}'\sigma'\rangle$ with photon $|\boldsymbol{k}'\sigma'\rangle$ and the medium in a state of energy E_b. The transition probability p is given by Fermi's golden rule up to the second order:

$$p = \frac{2\pi}{\hbar}\left|\langle f|H'|i\rangle + \sum_c \frac{\langle f|H'|c\rangle\langle c|H'|i\rangle}{E_i - E_c}\right|^2 \times \delta(E_i - E_f),$$
$$E_i = E_a + \hbar\omega_{\boldsymbol{k}} \qquad E_f = E_b + \hbar\omega_{\boldsymbol{k}'}$$

As mentioned above, only the terms in H' that are quadratic in $\boldsymbol{A}$ contribute to scattering in the first-order approximation, whereas the terms linear in $\boldsymbol{A}$ contribute to the second-order perturbation. As a result, $H_1^{'}$ and $H_4^{'}$ contribute to the nonresonant scattering; in contrast, $H_2^{'}$ and $H_3^{'}$, treated in the second-order perturbation, lead to scattering terms containing energy-dependent denominators at the origin of the resonant effects. In the following, only coherent elastic scattering in which the final state $|b\rangle$ of the scattering medium is the same as the initial state $|a\rangle$ with $\hbar\omega_{\boldsymbol{k}} = \hbar\omega_{\boldsymbol{k}'} = \hbar\omega$ is considered.

The photon part of the matrix can be calculated, and by summing over all electrons at positions j, the scattering amplitude by an atom can be written as

$$f_{\sigma'\sigma}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega) = -\sum_a p_a\left[\left\langle a\left|\sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j}\right|a\right\rangle \varepsilon_{\sigma'}^{\prime *}\cdot\varepsilon_\sigma - \mathrm{i}\frac{\hbar\omega}{mc^2}\boldsymbol{B}\cdot\left\langle a\left|\sum_j \boldsymbol{s}_j\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j}\right|a\right\rangle + \frac{1}{m}\sum_c\left(\frac{\langle a|O^{+}(\boldsymbol{k}')|c\rangle\langle c|O(\boldsymbol{k})|a\rangle}{E_a - E_c + \hbar\omega - i(\Gamma_c/2)} + \frac{\langle a|O(\boldsymbol{k})|c\rangle\langle c|O^{+}(\boldsymbol{k}')|a\rangle}{E_a - E_c - \hbar\omega + i(\Gamma_c/2)}\right)\right]$$

Note that the scattering amplitude is polarization dependent. It is generally represented by 2×2 matrices with the two transverse polarization states as basis vectors. In the following, the polarization index σ is dropped for the sake of simplicity.

The sum is taken over all possible initial states $|a\rangle$ of the scattering medium with probability p_a. The $\boldsymbol{B}$ polarization factor is expressed in terms of the incident and scattered unit wave vectors and polarization vectors:

$$\boldsymbol{B} = \varepsilon'^{*}\times\varepsilon + (\hat{\boldsymbol{k}}'\times\varepsilon'^{*})(\hat{\boldsymbol{k}}'\cdot\varepsilon) - (\hat{\boldsymbol{k}}\times\varepsilon)(\hat{\boldsymbol{k}}\cdot\varepsilon'^{*}) - (\hat{\boldsymbol{k}}'\times\varepsilon'^{*})\times(\hat{\boldsymbol{k}}\times\varepsilon),$$
$$\hat{\boldsymbol{k}} = \frac{\boldsymbol{k}}{|\boldsymbol{k}|} \qquad \hat{\boldsymbol{k}}' = \frac{\boldsymbol{k}'}{|\boldsymbol{k}'|}$$

The intermediate states $|c\rangle$ correspond to an electron excited into a previously unoccupied state and leaving a hole in the core level $|a\rangle$. Such excited states have a finite lifetime $1/\Gamma_c$; Γ_c is the natural width of the intermediate state $|c\rangle$, which is determined by all radiative and nonradiative de-excitations of $|c\rangle$. The operators $O(\boldsymbol{k})$ are given by: $O(\boldsymbol{k}) = \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}\{\boldsymbol{p}_j\cdot\varepsilon - i\hbar\varepsilon\cdot(\boldsymbol{k}\times\boldsymbol{s}_j)\}$, where $\boldsymbol{p}_j$ and $\boldsymbol{s}_j$ are the momentum and the spin of the jth electron; these operators depend on the polarization geometry.

Further manipulations of $f_n^{\mathrm{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ to separate the nonresonant from the pure resonant

scattering lead to

$$f_n(\boldsymbol{k},\boldsymbol{k}',\hbar\omega) = f_n^{\text{nonres}}(\boldsymbol{K}=\boldsymbol{k}-\boldsymbol{k}') + f_n^{\text{res}}(\boldsymbol{k},\boldsymbol{k}',\hbar\omega)$$

with

$$f_n^{\text{nonres}}(\boldsymbol{K}) = -\sum_a p_a \left[\left\langle a \left| \sum_j e^{i\boldsymbol{K}\cdot\boldsymbol{r}_j} \right| a \right\rangle \varepsilon'^{*}\cdot\varepsilon - i\frac{\hbar\omega}{mc^2} \left\langle a \left| \sum_j e^{i\boldsymbol{K}\cdot\boldsymbol{r}_j} \left(\frac{i\boldsymbol{K}\times\boldsymbol{p}_j}{\hbar K^2}\cdot(\varepsilon'^{*}\times\varepsilon) + \boldsymbol{B}\cdot\boldsymbol{s}_j \right) \right| a \right\rangle \right]$$

and

$$f_n^{\text{res}}(\boldsymbol{k},\boldsymbol{k}',\hbar\omega) = -\sum_a p_a \left(-\frac{1}{m}\sum_c \frac{E_a-E_c}{\hbar\omega} \times \frac{\langle a|O^{+}(\boldsymbol{k}')|c\rangle\langle c|O(\boldsymbol{k})|a\rangle}{E_a-E_c+\hbar\omega-i(\Gamma/2)} + \frac{1}{m}\sum_c \frac{E_a-E_c}{\hbar\omega} \frac{\langle a|O(\boldsymbol{k})|c\rangle\langle c|O^{+}(\boldsymbol{k}')|a\rangle}{E_a-E_c-\hbar\omega+i(\Gamma/2)} \right)$$

Nonresonant Scattering

The first term in $f_n^{\text{nonres}}(\boldsymbol{K})$ gives the classical Thomson charge scattering from electrons. The polarization dependence of the charge part is simple, that is, $f_n^{\text{nonres charge}}(\boldsymbol{K}) = -\rho_n(\boldsymbol{K})\varepsilon'^{*}\cdot\varepsilon$, where $\rho_n(\boldsymbol{K})$ is the Fourier transform of the charge density: the atomic charge scattering is a diagonal tensor which implies that polarization of X-rays is not rotated by charge scattering. The quantity $\rho_n(\boldsymbol{K})$ can be estimated from *ab initio* calculations based on model electronic wave functions as discussed below.

The second term in $f_n^{\text{nonres}}(\boldsymbol{K})$ corresponds to magnetic X-ray scattering with two contributions, one from the electron spin and the other from the electron momentum. The matrix elements containing the electron momentum can be rewritten explicitly in terms of the Fourier transform of the atomic orbital magnetization density $\boldsymbol{L}_n(\boldsymbol{K})$. This leads to

$$f_n^{\text{nonres magnetic}}(\boldsymbol{K}) = -i\frac{\hbar\omega}{mc^2}\left(\frac{1}{2}\boldsymbol{L}_n(\boldsymbol{K})\cdot\boldsymbol{A} + \boldsymbol{S}_n(\boldsymbol{K})\cdot\boldsymbol{B}\right)$$

where $\boldsymbol{S}_n(\boldsymbol{K})$ is the Fourier transform of the spin density of the nth atom and $\boldsymbol{A}$ is another polarization vector:

$$\boldsymbol{A} = 2(1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}')(\varepsilon'^{*}\times\varepsilon) - (\hat{\boldsymbol{k}}\times\varepsilon)(\hat{\boldsymbol{k}}\cdot\varepsilon'^{*}) + (\hat{\boldsymbol{k}}'\times\varepsilon'^{*})(\hat{\boldsymbol{k}}'\cdot\varepsilon)$$

The nonresonant magnetic scattering is much weaker than the charge scattering as indicated by the prefactor $\hbar\omega/mc^2$; the ratio of the magnetic to the charge scattering is given by

$$\left|f_n^{\text{nonres magnetic}}/f_n^{\text{nonres charge}}\right|^2 \approx \left(\frac{\hbar\omega}{mc^2}\right)^2\left(\frac{\rho_n^{\text{magnetic}}}{\rho_n}\right)^2$$

where ρ_n^{magnetic} is the Fourier transform of the "magnetic" electron density. For typical scattering experiments performed with X-ray energies around $\hbar\omega \approx 5$–$10\,\text{keV}$ (compared to $mc^2 \approx 511\,\text{keV}$), this ratio is $\sim 1\times10^{-6}$. This makes the observation of nonresonant magnetic X-ray scattering difficult, even when using powerful X-ray sources such as synchrotron sources. However, it is worth noting that the charge part, the spin part, and the orbital momentum part of $f_n^{\text{nonres}}(\boldsymbol{K})$ have different polarization dependences through $\boldsymbol{A}$ and $\boldsymbol{B}$, which not only allows the discrimination of magnetic and charge scattering but permits the full separation of the spin and momentum contributions to magnetic moments. There, synchrotron sources play a key role because they produce highly polarized X-ray beams.

If resonant effects could be ignored, then the atomic scattering factor for X-rays would not vary with the photon energy and there would not be any absorption. However, this approximation cannot be justified in all cases of practical importance: even when using standard X-ray sources in laboratories, Cu K-α radiation ($\hbar\omega \approx 10\,\text{keV}$) is often used to investigate the structure and properties of materials containing other $3d$ elements of the periodic table which have absorption edges in the same energy range.

Resonant X-Ray Scattering

The denominators in $f_n^{\text{res}}(\boldsymbol{k},\boldsymbol{k}',\hbar\omega)$ which contain the photon energy $\hbar\omega$ lead to resonant effects when the photon energy is tuned near any natural absorption energy of the scattering atoms, $E_c - E_a$. A schematic single-electron energy-level diagram in **Figure 1** illustrates the resonant scattering process. The incident photon virtually excites an electron from the ground state $|a\rangle$ into an unoccupied state above the Fermi level creating a hole in the core level; the excited electron then decays by filling the core hole and emitting a photon. The same intermediate electronic states contribute to the white lines in X-ray spectroscopy.

The different atomic absorption edges are labeled according to the core level $|a\rangle$, the final state of the excited electron being ignored. In a filled atomic shell (n,l) the total momentum is zero; however, as soon as a hole is created, it acquires a momentum

$j = l \pm (1/2)$. This j value can be used as an index to label the core level $|a\rangle$. **Table 1** shows the correspondence between the accepted nomenclature and the core levels.

Further levels with $n = 4, 5, 6$ are labeled similarly with letters N, O, and P. All levels with the same n are grouped together: the core electrons in the inner shells are hydrogen-like electrons and their energies are given in a first approximation by

$$-\frac{me^4}{2\hbar^2}\frac{Z^2}{n^2} = 13.60\frac{Z^2}{n^2}$$

in eV units, and depend on n only. Relativistic spin–orbit effects induce the separation of j multiplets for a given n; for example, L_1 and L_2 have a different energy than L_3. Terms with different l values but same n and j are separated due to intra-atomic electron interactions, which modify the Coulomb potential from the nucleus. Values for the absorption edges have been tabulated for elements of the periodic table in various chemical environments and valence states. The values of the K edges vary from 16 eV for hydrogen, 7.1 keV for iron to 115 keV for uranium, while L_3 edges range from 700 eV in the $3d$ series to 17.2 keV for uranium. **Table 2** gives a list of absorption edge energies for selected chemical elements.

Typical values for the energy line width are $\sim$1–10 eV, which leads to a fast scattering process in the 10^{-16} s range but sufficiently slow so that the excited states can be considered as quasistationary.

The resonant part of the scattering amplitude has three fundamental properties:

1. First, there are additional terms in the scattering amplitude which are known as dispersion correction terms. Normally, the atomic scattering factor is written as: $f = f_0 + \Delta f' + i\Delta f''$ where f_0 is the nonresonant atomic scattering factor. $\Delta f'$ and $\Delta f''$ are the real and imaginary parts of the dispersion correction given by $f_n^{\mathrm{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$. The imaginary part of $f_n^{\mathrm{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$, $\Delta f''$ which is positive, leads to X-ray absorption by the scattering medium. Similarly, the complex refractive index for X-rays $n = 1 - \alpha - i\beta$ is related to $f = f_0 + \Delta f' + i\Delta f''$:

$$\alpha = r_e \frac{\lambda^2}{2\pi} N(f_0 + \Delta f') \quad \text{and} \quad \beta = r_e \frac{\lambda^2}{2\pi} N\Delta f''$$

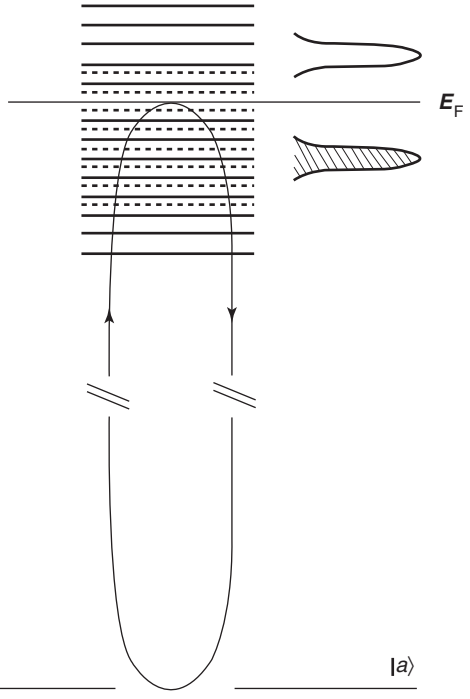


Figure 1 Schematic diagram showing the resonant process: the incident photon excites an electron in ground state $|a\rangle$ into an available state near the Fermi level $\boldsymbol{E}_F$. The intermediate state $|c\rangle$ contains a hole in $|a\rangle$ and an extra electron near $\boldsymbol{E}_F$. In an elastic scattering process, the excited electron in $|c\rangle$ decays back into $|a\rangle$, emitting a scattered photon with the same energy as the incident photon.

Table 1 Nomenclature of absorption edges

$1s_{1/2}$	$2s_{1/2}$	$2p_{1/2}$	$2p_{3/2}$	$3s_{1/2}$	$3p_{1/2}$	$3p_{3/2}$	$3d_{3/2}$	$3d_{5/2}$	…
K	L_1	L_2	L_3	M_1	M_2	M_3	M_4	M_5	…

Table 2 Selected resonances for characteristic chemical species. The transition process for dipolar ($E1$) and quadrupolar ($E2$) transitions are indicated. Resonant energies are given in keV; the corresponding X-ray wavelengths are given in Å

Chemical elements		*Resonance process*			*Resonant energy* (keV)	*Wavelength* (Å)
3*d* Transition metal	Fe	$E1$	$p \leftrightarrow d$	L_2, L_3	0.71	17.5
		$E1$, $E2$	$s \leftrightarrow p,d$	K	7.1	1.5
5*d* Transition metal	Pt	$E1$	$p \leftrightarrow d$	L_2, L_3	11.5–13	1.08–0.95
		$E1$, $E2$	$s \leftrightarrow p,d$	K	78.4	0.16
4*f* Rare earth	Gd	$E1$	$d \leftrightarrow f$	M_4, M_5	1.22–1.19	10.2–10.4
		$E1$, $E2$	$p \leftrightarrow d,f$	L_2, L_3	7.9–7.2	1.57–1.72
5*f* Actinides	U	$E1$	$d \leftrightarrow f$	M_4, M_5	3.7–3.5	3.35–3.54
		$E1$, $E2$	$p \leftrightarrow d,f$	L_2, L_3	20.9–17.2	0.59–0.72

where λ is the X-ray wavelength and N is the atomic density. Resonant effects induce large changes in the refraction of X-rays.

2. Second, resonant amplitudes depend on the polarization geometry and the local site symmetry of the "edge" atoms in crystals. The anisotropic resonant scattering amplitude must be represented by a second-rank tensor. The anisotropy of the tensorial macroscopic refraction index n gives rise to X-ray optical effects similar to those observed in the visible light range. All local symmetry-breaking effects can play a role: the local symmetry of specific sites, crystal lattice distortions, and even the onset of magnetic moments.
3. Finally, it should be noted that the resonant process is highly selective. Resonant energies characterize chemical elements in well-defined valence states (**Table 2**) and the resonances are rather narrow; therefore, only "edge" atoms in materials, that is, those atoms where edges are excited by the incident X-rays, contribute to these resonant intensities. Furthermore, as shown below, selection rules specify which electronic shells can contribute to resonances, hence the electronic shell selectivity.

Resonant Scattering Amplitudes

The magnitude of the resonant scattering is given by the matrix elements in $f_n^{\mathrm{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$. The operators $O(\boldsymbol{k})$ contain electron spin and momentum operators. However, for photon energy larger than 10–100 eV, the magnetic transitions induced by the spin operators are much weaker than the electric transitions induced by the electron momentum operators. Therefore, the spin operators can be omitted in $f_n^{\mathrm{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ and electric transitions only are involved in connecting to the intermediate states. The resonant contributions to the coherent scattering amplitude can be written in terms of various approximations following the expansion $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} \approx 1 + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j + (1/2)\left(\boldsymbol{k}\cdot\boldsymbol{r}_j\right)^2$ in the operator $O(\boldsymbol{k}) \approx \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}\,(\boldsymbol{p}_j\cdot\varepsilon)$. Dipole and quadrupole approximations are retained as the leading terms, bearing in mind that dipolar matrix elements are normally larger than the quadrupolar ones for the transitions that they allow. The importance of the transition amplitude is determined by overlap integrals of the occupied orbitals in the initial state and the new occupied orbitals in presence of the core hole in the intermediate states: these are matrix elements $\langle c|r^L|a\rangle$, where $L = 1.2$ for dipole and quadrupole transitions. Selection rules exist which restrict the possible intermediate states depending on the nature of the transition: a dipolar electric transition is associated with a change ΔL in the orbital momentum and no change in spin, whereas a quadrupolar electric transition involves $\Delta L = 2$. Returning to **Figure 1**, if a transition metal of the $3d$ series near the K edge is considered, dipolar transitions ($\Delta L = 1$) will involve electronic excited states in the $4p$ bands, whereas quadrupolar ($\Delta L = 2$) terms would involve the available $3d$ states. Similarly, resonant effects at the L_3 edge of a rare-earth element in the $4f$ series will include the $5d$ electronic bands via dipolar transitions and the more localized $4f$ states via quadrupolar transitions.

The detailed comparisons of dipolar and quadrupolar contributions to the resonance depend on the origin of the resonant process.

Sensitivity to Local Symmetry

The outer electronic shells of the intermediate states are sensitive to the local symmetry of atomic sites, and the atomic scattering amplitude depends on the relative orientation of the incident X-ray polarization with respect to local symmetry axis. Therefore, near a resonance, the atomic scattering factors contain resonant terms which reflect the local symmetry of the atomic sites; atomic scattering factors are no longer scalar and diffraction may exhibit birefringence. This anisotropy of the anomalous scattering leads to spectacular effects, such as the appearance of otherwise "forbidden" reflections as discussed below.

Although the leading resonant terms are due to multipolar electric transitions, resonant X-ray scattering is sensitive to local magnetic moments. The spin–orbit coupling, which couples the spin of the available electronic levels to the orbital momentum, makes overlap integrals depend on the net magnetic polarization of the electronic bands. As a result, resonant scattering also provides information on magnetic properties.

Dispersion Corrections

The dispersion corrections, that is, the dependence of the atomic scattering power on the photon energy, had remained a deep concern for crystallographers who had viewed these corrections as an annoyance. Indeed, dispersion effects are strongly dependent on the electronic structure of atoms and ions in solids; furthermore, close to the resonances, the local symmetry and environment can lead to new and large effects. Nevertheless, the advent of synchrotron radiation sources has opened new fields of activity in variable-wavelength methods, and crystallography experiments offer better understanding and more accurate estimates for these corrections. These developments have brought renewed interest, both from the theoretical and the experimental side, to

tackle the problem of dispersion corrections. Several attempts had been made to calculate the full atomic scattering factor using models for atomic electrons, such as the Thomas–Fermi model or self-consistent-field methods. The early methods had ignored the fact that electrons occupy definite energy levels within atoms. The presence of the absorption edges were reproduced by including damped oscillators to introduce dispersive corrections to the atomic scattering factors.

The current theoretical approaches for the calculation of corrections include classical, nonrelativistic, and relativistic theories. It should be noted that these calculations assume an isolated neutral atom with spherical symmetry. Most of these approaches ignore the changes in X-ray polarization during scattering. Refined relativistic calculations have provided a detailed understanding of the scattering process and have assessed the validity of simpler approaches. Values for $\Delta f'$ and $\Delta f''$ have been tabulated using different calculation methods and can be compared with experimental determinations.

The experimental techniques to determine the corrections rely on the measurements of macroscopic properties, such as the refractive index and the linear absorption coefficient. These methods involve dispersion corrections at zero photon momentum transfer $\boldsymbol{K} = 0$.

The imaginary part $\Delta f''$ is obtained through measurement of the linear absorption coefficient. The real part $\Delta f'$ can be obtained through the Kramers–Kronig inversion of the experimental data for $\Delta f''$ obtained over a sufficiently broad range of photon energies. More accurate methods involve the measurement of optical index X-ray interferometry, total reflection of X-rays, and refraction of X-rays through prisms.

For isolated absorption edges, $\Delta f'$ takes negative values at low energies reaching a minimum at the edge (where $\Delta f''$ reaches a maximum), and approaching small positive values at higher energies. The values for $\Delta f'$ at absorption edges can be very large: in the $3d$ series, $\Delta f'$ reaches $-9\ r_e$ at the edge (theories predict $-13.5\ r_e$) to be compared to the 29 electrons in Cu atoms.

Other techniques measure intensities of Bragg reflections from crystalline materials as a function of the photon energy or the variation of intensities within a diffraction peak due to the interferences between transmitted and reflected X-ray waves as predicted by the dynamical theory of X-ray diffraction. These measurements provide data on the dispersion corrections at the finite scattering vector $\boldsymbol{K}$. However, the $\boldsymbol{K}$ dependence of the dispersion corrections is not established.

Anisotropy of Resonant Scattering

Away from resonance, the nonresonant scattering amplitude $f_n^{\text{nonres charge}}$ does not depend on the relative orientation of the crystal axis with respect to the polarization vectors: diffracted intensities are not sensitive to rotations of the crystal about the scattering vector $\boldsymbol{K}$ (azimuthal scans or Renninger scans). This property is exploited to detect artifacts (such as multiple scattering) in collecting diffraction data. However, this property is no longer valid near absorption edges. In general, the dispersion corrections must be taken as second-rank tensors, which are compatible with the symmetry of the chemical bonding of the atomic sites in crystals. Consequently, the scattering cross section $d\sigma/d\Omega$ takes a tensorial form. In absorption and transmission experiments, this leads to birefringence and pleochroism effects. In diffraction experiments at an absorption edge, the intensity and polarization of diffracted peaks depend on the rotation of the crystal about the scattering vector $\boldsymbol{K}$ as shown in **Figure 2**.

A special consequence of the anisotropy of $f_n^{\text{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ is the violation of extinction rules due to equivalent crystal sites with different orientations of their local axis. Due to the anisotropy of $f_n^{\text{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$, edge atoms, which occupy equivalent crystal sites related by a screw-axis or glide-planes, do not have the same scattering power at resonance. As a result, the usual rules for absent reflections based on space groups symmetry are no longer

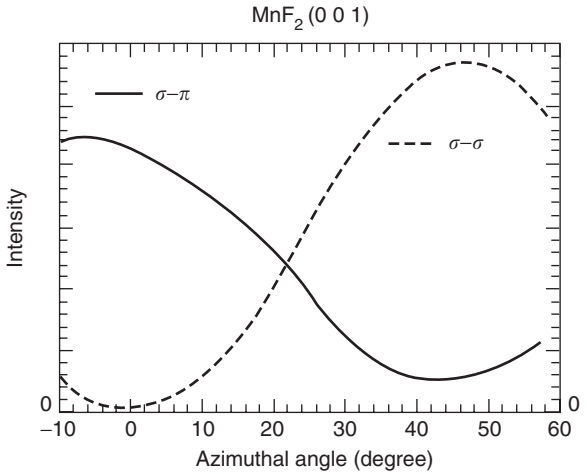


Figure 2 Rotation of the sample about the scattering vector $\boldsymbol{K}$ (azimuthal scan) at the K-edge on manganese at the forbidden reflection (0 0 1) in manganese difluoride, MnF_2. Such a reflection would be forbidden by the crystal symmetry of the MnF_2 structure if the atomic scattering factor of manganese were spherical. At resonance, atomic factors are not spherical; scattered intensities with different polarizations vary with the orientation of the crystal axis with respect to the X-ray polarization vectors (azimuthal angle).

rigorous: diffraction peaks which would be forbidden away from resonance can be observed when the photon energy is tuned near the absorption edges. The "forbidden" Bragg reflections provide information on the partial structure of the edge atoms.

Magnetic Resonant Scattering

Magnetooptics phenomena (Faraday rotation and magnetooptic Kerr effect) with visible light and infrared radiation clearly demonstrate that electromagnetic radiation interacts with magnetic moments. These phenomena involve the refraction index, which is related to the $\boldsymbol{K} = 0$ part of the scattering amplitude. However, in the X-ray range away from resonances, it can be seen that the polarization vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ which enter $f_n^{\text{nonres magnetic}}(\boldsymbol{K})$ vanish at $\boldsymbol{K} = 0$. The X-ray refraction index differs strongly from unity only near absorption edges due to the anomalous part of the scattering amplitude: this is where polarization rotation effects and magnetic scattering effects can be detected.

Resonant magnetic scattering originates from the dependence of $f_n^{\text{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ on the magnetization at site n. This happens if the overlap integrals depend on the local magnetic moment or if the energies of intermediate levels $|c\rangle$ are split by exchange interactions. The magnetic sensitivity arises from the combination of spin–orbit couplings which must be present in one of the two levels involved at the resonance and exchange interactions.

The following example illustrates the effect (**Figure 3**): if the ground state $|a\rangle$ of the atom has a single hole in the $4f$ shell ($l = 3$) and, therefore, carries a localized magnetic moment, the direct application of Hund's rule shows that the empty electronic orbital is characterized by $m_l = -2$, $m_s = -1/2$, which is also a quantum state with $j = 7/2$, $m_j = -7/2$. At the M_5 edge (initial state in the 3 $d_{5/2}$ band), dipolar transitions contribute with $\Delta L = 1$ and $\Delta M_L = -1$ only; this corresponds to a circular left-handed polarization. The other two transitions, $\Delta M_L = 1$ and $\Delta M_L = 0$, are not allowed. If all sites have their magnetic moments aligned in the same direction, left-handed circularly polarized photons with proper energy will excite this transition and will be absorbed by the assembly of magnetic atoms, whereas the other polarizations, linear and right-handed circular, would not interact. This is the origin of X-ray magnetic dichroism. Similar effects will take place in the scattering process. Estimations in this idealized case have shown that the magnetic resonance scattering amplitude could reach 100 in electron units, giving extremely large effects in X-ray absorption and diffraction.

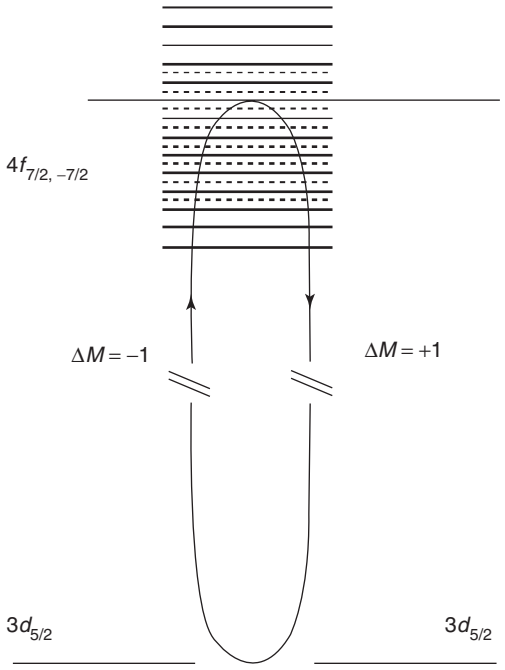


Figure 3 Schematic diagram showing magnetic sensitivity of the resonant process: only the $\Delta M = -1$ transition is allowed at the transition leading to X-ray dichroism in absorption experiments.

If the edge energy $E_c - E_a$ is split by exchange interactions, that is, depending on the spin state of the intermediate electronic level $|c\rangle$, then the transition probabilities also depend on the change in momentum ΔM_L induced by the incident radiation. As a result, X-ray dichroism and magnetic scattering will also appear.

If the momentum transfer $\boldsymbol{K}$ corresponds to the magnetic periodicity of the material, the sum over intermediate states for all atoms is nonzero, which leads to resonant magnetic Bragg peaks. The resonant scattering is superimposed on nonresonant scattering: the two scattering amplitudes interfere to produce the total cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$.

Large resonant enhancements of magnetic scattered intensities have been observed at the $L_{\text{II, III}}$ edges of rare earths (10- to 100-fold enhancements) and $M_{4,5}$ edges of actinides (several orders of magnitude enhancements) as illustrated in **Figure 4**. A fundamental difference between $f_n^{\text{nonres magnetic}}(\boldsymbol{K})$ and $f_n^{\text{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ is that $f_n^{\text{nonres magnetic}}(\boldsymbol{K})$ contains the spin and the orbital contributions to the magnetic moment, whereas $f_n^{\text{res}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ is sensitive to the orientation of the local moment only. However, there exist sum rules for $\Delta f''$ which allow one to extract values for the orbital moment and for the spin.

The largest resonant enhancements observed at the $M_{4,5}$ edges of actinides result from strong dipole transitions, which couple the d core levels directly

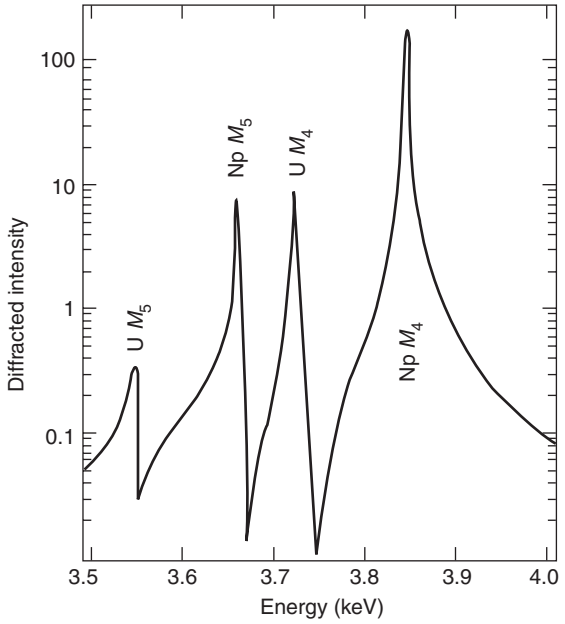


Figure 4 Photon energy dependence of magnetic diffraction intensity from a material containing two magnetic elements, uranium and neptunium. Neptunium and uranium are clearly distinguishable and well separated. Tuning the photon energy to one edge makes the diffraction sensitive to neptunium or uranium magnetic properties. The resonant intensities are enhanced by 5 orders of magnitude compared to intensities observed in the nonresonant regime far above the edges.

to the unfilled magnetic $5f$ electronic shell. The size of the magnetic resonance is a combination of the overlap integrals discussed above and the magnetic polarization of the intermediate states. The weaker enhancements observed at the $L_{2,3}$ edges of rare earths are explained by the fact that dipole transitions couple to $5d$ bands, which are more spin-polarized than the $4f$ shells, which carry the magnetization in rare earths. In such a case, quadrupolar transitions can be comparable to dipolar terms because they couple the core level ($2p$) directly to the strongly spin-polarized $4f$ states. **Table 2** lists resonant processes which are of importance for resonant magnetic scattering experiments.

The properties of resonant magnetic X-ray scattering lead to important applications. The large size of the scattering amplitude allows minute samples and even surfaces to be studied in magnetic scattering experiments. The chemical selectivity of the resonant process is illustrated in **Figure 4**: tuning the photon energy through the absorption edges of different chemical species makes it possible to distinguish the magnetic moments that reside in different elements. Similarly, the electronic sensitivity allows the distinction of the contributions from the d, f or even itinerant p electronic states to magnetization.

Corresponding effects can be observed in X-ray spectroscopy. Restricting the resonance to dipolar transitions, $f_n^{\text{res, dipole}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ can be written as

$$f_n^{\text{res, dipole}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega) = F^{(0)}\varepsilon'^{*} \cdot \varepsilon - iF^{(1)}(\varepsilon'^{*} \times \varepsilon)z_n + F^{(2)}(\varepsilon'^{*} \cdot z_n)(\varepsilon \cdot z_n)$$

where z_n is the quantization axis at the site n. The $F^{(I=0,1,2)}$ are the resonant matrix elements and contain the resonant denominators. The first term in $f_n^{\text{res, dipole}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ is independent of z_n and does not contribute to the magnetic scattering; it is the usual anomalous X-ray scattering term which contributes to white lines in the fluorescence signal. The second term in $f_n^{\text{res, dipole}}(\boldsymbol{k}, \boldsymbol{k}', \hbar\omega)$ leads to X-ray circular dichroism in X-ray spectroscopy experiments, and the third term contributes to X-ray linear dichroism.

Future Directions

Resonant scattering processes have opened new dimensions to X-ray scattering. The sensitivity to local symmetries and the selectivity properties of the resonant scattering amplitude become new tools to exploit the potential offered by the properties of synchrotron radiation: tunability of the energy and of the polarization of the X-ray beams.

See also: Scattering Techniques, Compton; Scattering, Elastic (General); Scattering, Nuclear Resonant.

PACS: 61.10.−l; 61.10.Dp; 75.25.+z; 41.50+h; 32.80.Cy; 33.55.−b

Further Reading

Blume M (1994) Magnetic effects in anomalous dispersion. In: Materlik G, Sparks CJ, and Fischer K (eds.) *Resonant Anomalous X-Ray Scattering*, pp. 495–512. Amsterdam: North-Holland.

Chaikin PM and Lubensky TC (1995) *Principles of Condensed Matter Physics*. Cambridge: Cambridge University Press.

Creagh DC and McAuley WJ (1992) X-ray dispersion corrections. In: Wilson AJC (ed.) *International Tables for Crystallography*, vol. C, pp. 206–222. Dordrecht: Kluwer.

Gibbs D, Hill JP, and Vettier C (2002) New directions in x-ray magnetic scattering. In: Mills DM (ed.) *Third-Generation Hard X-ray Radiation Sources*. New York: Wiley.

James RW (1962) *The Optical Principles of the Diffraction of X-Rays*. Woodbridge: Ox Bow Press.

Kirfel A (1994) Anisotropy of anomalous scattering in single crystals. In: Materlik G, Sparks CJ, and Fischer K (eds.) *Resonant Anomalous X-Ray Scattering*, pp. 495–512. Amsterdam: North-Holland.

Semiconductor and Metallic Clusters, Electronic Properties of

A Stella and M Patrini, Università degli Studi di Pavia, Pavia, Italy

Introduction

The clusters of atoms and molecules are somewhat intermediate in several aspects (number of constituent atoms N, size, basic properties, etc.) between simple atoms and molecules on one side and macroscopic aggregates (bulk solids) on the other side. The advent of clusters as stable or at least reproducible nanosystems composed of a number of atoms ranging roughly between 10 and 10^4 (consequently with average size up to a few nanometers, so that quite often the term nanoparticles is used to indicate them) took place gradually with the parallel development and refinements of techniques to master and study their properties.

When an increasing number of atoms are progressively aggregated to form clusters, the evolution from atomic-like to bulk-like behavior is ascertained to have happened in many different ways, depending on the kind of atoms, the growth conditions, and the environment. Useful and general schemes of classification have been proposed. Now the question may arise: are there clear-cut borderlines between single atoms or molecules and very small clusters on one side, and also between relatively bigger aggregates and bulk solids on the other side? While it is generally accepted that the main properties of more extended aggregates ($N \sim 10^3$–10^4) gradually tend to bulk values, the situation appears to be rather more complex on the former borderline. It can be safely stated that the molecules are characterized by a definite composition and also a definite structure; clusters on the contrary may contain a variable number of atoms and for a given N, may exhibit many possible structures, some of them more stable than the other ones. If N is progressively increased, it is found that more stable structures exist in correspondence to specific numbers, which are called "magic numbers."

On general grounds, it is evident that recent advances in growth techniques (lithography, film deposition, molecular beam epitaxy, etc.) have opened the way to new and rich developments of physics: this is based on the fact that size and length scales characterizing the nanosystems described here are now comparable to or smaller than typical lengths (such as, electron mean free path and exciton radius) which have been playing an essential role in the various phenomena taking place in macroscopic condensed matter.

Clusters: Size-Related Typical Quantities

The investigation of the structures in the nanometer domain presents rather challenging aspects in terms of experimental characterization and theoretical calculations for the prediction and/or interpretation of their basic properties. To summarize in one sentence, it can be stated that space confinement plays a key role in modeling and tuning the physical properties of nanocrystals. This statement is supported with a few considerations:

- limitation of the validity of effective-mass approximation below a given size, and of the *ab initio* molecular dynamics approach above a certain number of constituent atoms ($N > 10^2$);
- strong redistribution of the electronic density of states, and blue-shift of the main spectral structures when the size is smaller than the excitonic radius for semiconductors;
- discretization of the electronic density of states in metals, and modification of transport properties when the number of constituent atoms does not exceed a few tens;
- the breaking of the $k = 0$ selection rule due to the uncertainty principle $\Delta x \Delta k = 1$ (where Δx and Δk are the uncertainty on position and wave vector, respectively), and the related possibility of exploring the ω versus k dispersion curve, also away from $k = 0$;
- strong modification of the main properties when N_s/N (where N_s is the number of surface atoms, and N the overall number of atoms) becomes relevant;
- strong effects of surface scattering in the ultrafast electron dynamics, when the size is smaller than the mean free path with consequences on the thermalization times of the electron gas, and
- modification of the surface plasmon energy due to space confinement.

The appreciation of the relevant length scales (between 0.1 nm and 0.1 mm) over which the phenomena are evidenced is an important prerequisite to a more detailed understanding of the new physics. The length scales are set by different aspects:

1. there are those measured on an atomic scale, associated with the deposition of the constituent materials;

2. scales are set by the achievement of lithography and technology processes in order to stabilize and integrate nanostructures in application-oriented configurations;
3. typical lengths are set by physical principles (diameter of excitons, radius of cyclotron orbits) as cited above, and
4. other lengths are associated with scattering coherence (mean free path, surface quality, etc.), and are derived from the optical or electronic properties.

From Bulk to Nanoregime

Two completely different approaches have been adopted in the last two decades to prepare three-dimensionally confined structures or clusters. The first approach (top-down) is based on patterning and etching techniques in order to obtain space-limited systems starting from macroscopic samples. In fact, originally the fabrication methods of quantum wires and dots were based on lateral patterning of two-dimensional heterostructures, by combining fine-line lithography with wet and chemical etching. However, lithographic patterning introduces irregularities in the shape of the nanostructures, and mechanical damage cannot be avoided. The second one (bottom-up) is related to the aggregation and self-organization of single atoms and molecules to form clusters in the nanometer range via evaporation/condensation processes.

One of the most conspicuous examples of a stable cluster is certainly represented by the "buckyball" C_{60} molecule. The forces binding the clusters together are of different types, depending on the specific material; for instance, Coulomb forces are responsible for NaCl clustering, covalent forces for C_{60} and Si aggregates, van der Waals forces for rare-gas agglomeration, and metallic bonds in the case of metals. This implies, of course, that a wide variety of approaches must be adopted to grow and study them. However, instead of trying to give a detailed picture of the very ample spectrum of possible systems that can be obtained with their specific features, one can state the general principles that characterize the behavior of the clusters.

In general terms, one can summarize the evolution of the main properties with size as follows:

1. irregular or strongly oscillating size-dependence of structural, energetic, electronic, and electromagnetic observables, in connection with the existence of the magic numbers for clusters with very small N ($<10^2$);
2. if the size is increased ($N>10^2$), a smooth and continuous evolution of the properties toward bulk values is observed. The behavior of a generic property is shown in **Figure 1**.

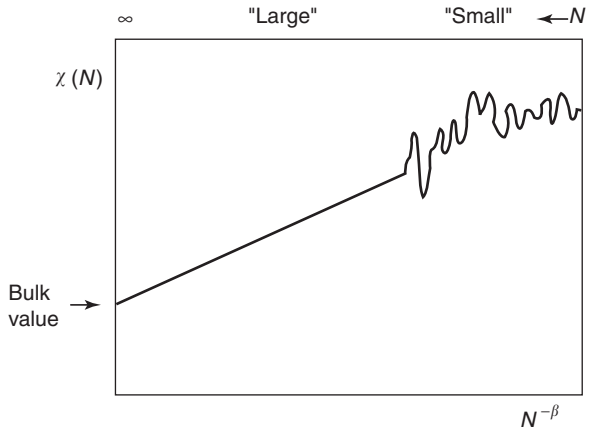


Figure 1 Schematic representation of the behavior of a general property of a cluster vs. size.

In the N region (2), a qualitative description of the size-dependence of a property can be expressed by the relationship

$$\chi(N) = \chi(\infty) + AN^{-\beta}$$

where A is a constant and β can vary between 0 and 1. $\chi(\infty)$ represents the bulk value, which is usually determined quite easily from experiments or reliable simulations. Specific examples amenable to this equation are:

1. the cohesive energy of a metal cluster in a metal-drop model,
2. the ionization potential of van der Waals clusters in the case of an impurity state in a rare-gas cluster,
3. melting temperature of a nanoparticle as compared to the bulk value, and
4. confinement of Wannier excitons where the energies of the excitonic nth levels are expressed by the above relationship.

Adequate theoretical studies of these systems require, in general, rather sophisticated and complex computational procedures and methods. In principle, at least in the simplest cases, one uses the same theoretical tools generally applied to study molecules; the quantum mechanical methods range from *ab initio* variational calculations of the entire electronic structure to more approximate methods related to the use of "effective potentials." However, recent developments are bringing important new contributions to the treatment of more complex systems. In

this respect, mention may be made of: (1) the Monte Carlo method, which allows one to explore, by statistical sampling, the accessible phase space of configurations and momenta of particles, and (2) molecular dynamics for solving, by means of numerical simulation, the simultaneous equation of motion of all the atoms of a cluster (if one has a reliable potential).

Basically, the starting approach in studying clusters and small aggregates consists in selecting a structure, and then calculating the consequent total energy. By repeating the calculations for all possible geometries, one determines the ground state, characterized by the lowest energy among all configurations. However, while these results have been quite satisfactorily obtained in molecular physics for some decades through the well-known and standard quantum mechanical method of calculating the exact wave functions of electrons and ions, nontrivial difficulties arise when the systems are given by a relatively large number of atoms; in this case, the number of local energy minima increases dramatically with size and the problem tends to be practically insoluble.

Among the several approaches adopted, it is worth stressing the methods based on a combination of density functional and molecular dynamics. The next two sections will be devoted to more specific treatments given separately to semiconductor and metallic clusters/nanosystems.

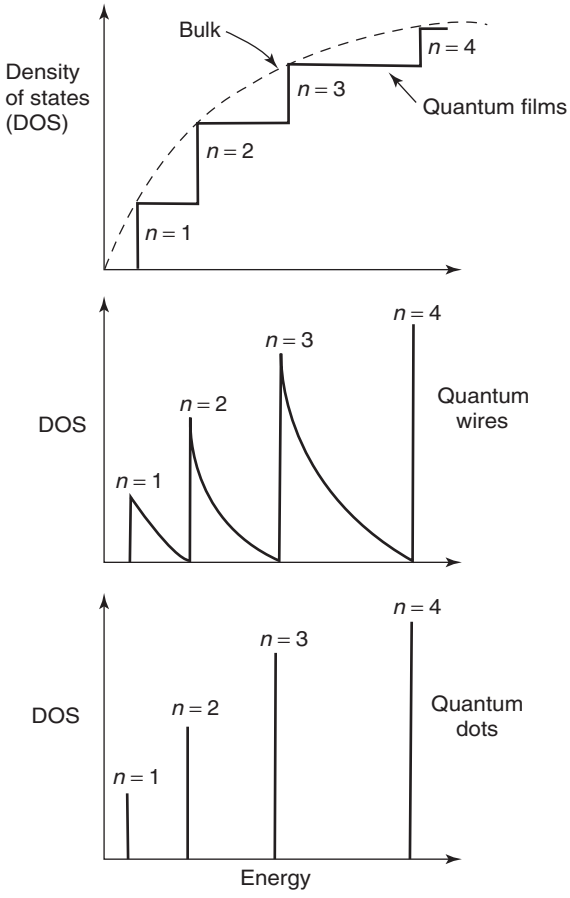


Figure 2 Effect of dimensionality on the electronic density of states: quantum films, which are confined in one dimension (as compared to bulk), quantum wires, and quantum dots (confined in two and three dimensions, respectively).

Semiconductor Clusters

Recently, a lot of emphasis has been given to semiconductor clusters or "quantum dots" (QDs, also called nanocrystals), and in particular to their quantum behavior deriving from space confinement. These nanosystems are characterized by sizeable effects on the electron density of states, with respect not only to bulk, but also to quantum wells and quantum wires.

As a matter of fact, in a more general framework three regimes of quantization can be defined depending upon whether the charge carriers are confined in one, two, or three dimensions. Confinement in one dimension creates structures that have been termed quantum wells. This is because such structures could be described by the elementary quantum mechanics of a particle in a one-dimensionally confined box. The carrier confinement in two dimensions produces quantum wires, and confinement in three dimensions produces quantum dots. The dimensionality of confinement affects many aspects of quantization; for example, **Figure 2** shows the effect on the density of states.

Remarkable developments related to applications of these type of systems have been reported in literature, and are being extensively discussed in state-of-the-art applied research.

To introduce the basic electronic properties of quantum dots (and their consequent spectroscopic behavior), it is to be assumed that one deals with a simple two-band semiconductor with parabolic, isotropic, direct bands (eventually spin degenerate). In addition, an effective mass approximation (EMA) is used: as it has been proved, this would imply that one considers dot sizes larger than a given value, which is well above 1 nm. Consider the case of spherical quantum dots with radius R. For spherical potential wells, the kinetic energy of a quantum mechanical particle takes discrete values and scales as the square of the inverse radius

$$E_{e,nl} = E_g + \frac{\hbar^2}{2m_e^*}\left[\frac{k_{nl}}{R}\right]^2$$

$$E_{h,nl} = E_g + \frac{\hbar^2}{2m_h^*}\left[\frac{k_{nl}}{R}\right]^2$$

where E_g is the fundamental energy gap, m^* is the effective mass, and the subscripts e and h refer to electrons and holes, respectively. It is usual to refer to the n, l electron or hole eigenstate as ns, np, nd, etc., where s, p, d correspond to $l = 0$, 1, 2, respectively. The lowest confined state of a single electron–hole pair has its energy increased with respect to the bulk gap by

$$\Delta E = \frac{\hbar^2}{2m_r}\left(\frac{\pi}{R}\right)^2$$

where the reduced mass m_r is given by

$$m_r = \frac{m_e^* m_h^*}{m_e^* + m_h^*}$$

To introduce scaled quantities, the exciton Rydberg energy E_R is used:

$$E_R = \frac{\hbar^2}{2m_r a_B^2}, \quad \text{with } a_B = \frac{\varepsilon' \hbar^2}{m_r e^2}$$

where a_B is the exciton Bohr radius, and ε' is the background dielectric constant. Then,

$$\Delta E = E_R\left(\frac{\pi a_B}{R}\right)^2$$

For small quantum dots, the confinement-induced energy shift is large in comparison to the exciton binding energy.

Important new aspects of quantum confinement can be evidenced by switching from the fundamental bandgap to higher-energy structures, such as E_1 and E_2. For example, a blue-shift effect is obtained in the optical absorption spectra measured on nanocrystals of germanium (see **Figure 3**), a paradigmatic case of indirect gap semiconductor, with different values of the radius R.

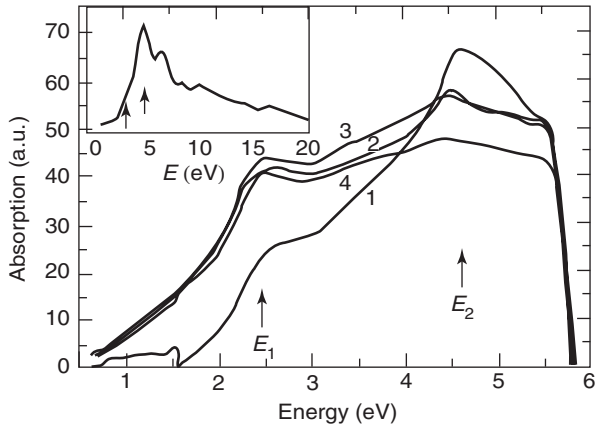


Figure 3 Optical absorption spectra of Ge nanoparticles in the E_1 and E_2 interband transition spectral region, with radii varying from 1.2 (sample 1) to 6.5 nm (sample 4). Inset: calculated $k \cdot p$ absorption spectrum for bulk Ge.

In general, E_2 transitions come from different k points in the Brillouin zone. Although a blue-shift of spectral structures is generally observed in confined systems, the quantitative explanation of the E_2 shift goes beyond the simple effective-mass models. As calculated by Wang and Zunger for Si quantum dots, the confinement shift of the spectral structures does not necessarily grow as $1/R^2$ but, in any case, it increases faster than the electron–hole Coulomb attraction, which behaves as $1/R$, so that the excitonic enhancement of absorption is expected to be less important in the quantum dots compared to the bulk.

The E_1 spectral structure presents a different behavior, characterized by the interplay of the reduction of the joint density of states (k is no more a good quantum number, the notion of interband transition between parallel bands becomes meaningless) and of the excitonic nature. The overall effect is given by a reduction of the oscillator strength accompanied by the weakening of the excitonic structure.

Recent calculations of the electronic structure of quantum dots display more and more sophisticated approaches, where the dot is viewed as a new structure rather than a perturbation on the bulk material. This is made possible computationally by a series of innovations rendering a 10^3–10^6 atom problem tractable within a pseudopotential framework. The approach starts from the bulk band structure and treats nanostructures of all dimensionalities on an equal footing; it captures the correct atomistic structures, strain and symmetry, including surface effects, and incorporates Coulomb and exchange effects without any further approximations. It can be applied to "free-standing," for example, colloidal dots as well as to embedded ("self-assembled") dots.

Since quantum dots allow one to study the electronic properties at the extreme limits of zero dimension, their behavior appears to be transferable into novel device concepts.

One of the most investigated aspects obviously concerns the optical properties of QDs that are relevant in optoelectronic devices. First, the luminescence line width from excitons in single QDs is ultranarrow up to elevated temperatures, proving that the density of electronic states is δ-function like, similar to the density of states in an atom. A potential significant point for applications is the realization of efficient QD luminescence at 1.3 μm wavelength at room temperature, and is coupled to favorable properties for QD lasers: low, temperature-insensitive threshold current, high gain, and differential gain. From the first laser presented in 1994 by Bimberg and Grundmann, which exhibited zero-dimensional

properties until 120 K, significant further progress has occurred. Laser operation at room temperature of a single sheet of dots on their ground state can be prevented by gain saturation in connection with sufficiently high optical losses – for example, in short cavities. Such an effect can be circumvented by incorporating more QDs into the cavity by vertical stacking. In vertical-cavity surface emitting lasers (VCSELs) incorporating QD stacked layers as active medium, the matching of the cavity passband with the QD optical density of states exhibits laser operation at room temperature on their ground state. When structural anisotropy of self-ordered dots is exploited, laser emission of polarized light from VCSELs can occur.

In addition, both quantum confinement and strong Coulomb interaction have strong effects on transport. The few-electron regime is accessible in vertical dot devices, which contain a dot located between source and drain contacts by means of heterostructure tunnel barriers. When a dot is coupled to a reservoir, the addition of an extra electron raises the electrochemical potential of the dot, and the one-by-one change of the number of electrons leads to conductance oscillations (Coulomb oscillations) as a function of gate voltage. The oscillation period is usually constant when a dot contains many electrons and equals e/C, where C is the total capacitance and e is the elementary electron charge. In a few-electron dot, the e–e interactions and quantum confinement effects become so strong as to modify the Coulomb oscillations remarkably. The absolute value N of the electrons can be identified by starting from the zero-current region (Coulomb-blockade region); then the current shows clear Coulomb oscillations where each period corresponds to a change of exactly one electron in the dot. When N becomes smaller than 20, an increasing irregularity is found and the oscillation period depends strongly on N.

Metal Clusters

In the case of metals also, new effects are expected due to space confinement. Already in the 1930s, Frölich had pointed out that the continuous distribution of the electronic density of states in the conduction band tends to break up into discrete states when the size of the metal becomes sufficiently small. The problem was quantitatively formulated by Kubo in the 1960s, and since then, considerable efforts have been made to get experimental evidence of the quantum effects.

In general terms, if the particles are so small that the conduction band breaks up into discrete levels separated by energies which are large compared to thermal energies, the Drude expression is no longer valid. It must be replaced by a more realistic, quantum mechanically derived dielectric function. The simplest quantum mechanical model that can be used is that of free electrons in a cubic potential well with infinite sides. The formal advantages of this model are clear: analytic expressions can be obtained at all stages in the calculation. Perhaps, a serious disadvantage is related to the high degree of symmetry and therefore of degeneracy in the electronic states, a degeneracy which probably will not be found in nature.

The one-electron wave functions of this model are well known:

$$\psi_{lmn} = \left(\frac{8}{L^3}\right)^{1/2} \sin\left(\frac{l\pi x}{L}\right) \sin\left(\frac{m\pi y}{L}\right) \sin\left(\frac{n\pi z}{L}\right)$$

where L is the length of one edge of the potential well, and l, m, n are positive integers. The corresponding energy levels are

$$E_{ijk} = E_0(l^2 + m^2 + n^2)$$

where

$$E_0 = \frac{\pi^2 h^2}{2mL^2}$$

A nontrivial aspect of the discretization of the electronic energy levels in metals is related to the gap or energy separation between two consecutive levels:

$$\Delta E = \frac{E_F}{N^{1/3}}$$

where E_F is the Fermi energy, and N the number of constituent atoms. It is evidently hard to have energy separation larger than either the thermal energies or the widths of the levels (as induced by defects, inhomogeneities, etc.), already when $N > 50$–100 atoms, that is, when the size becomes ~ 0.5–1 nm. This makes a remarkable difference with respect to semiconductors, where clearly observable quantum effects show up when the size becomes smaller than the excitonic radius a_0, that is, already ~ 5–10 nm in some cases.

The discretization of the electronic energy levels in the conduction band implies that the Drude model is no longer valid in very small nanoclusters. A clear distinction comes out with respect to large clusters or bulk metals, where plasmon-like excitations are based on almost elastic processes for the involved electrons due to field-induced coherent oscillations superimposed on their k vectors. The individual electrons can gain only small amounts of energy of the

order of $\Delta E = h\omega/N_e$, where N_e is the number of participating electrons in the cluster. This energy becomes important in small clusters ($\Delta E \sim 0.1$ eV for $N_e \sim 100$) and hot electrons are produced. The smaller the cluster, the hotter the electrons and in such a case, it is preferable to speak of collective electron excitations rather than of plasmons.

In the case of spherical clusters embedded in a dielectric matrix (ε_m), if the complex dielectric function $\varepsilon(\omega)$ of the particles is known, the average dielectric function $\langle \varepsilon \rangle$ of the composite medium describes completely the optical response of the layer. For small volume fraction of the particles, f, one can write the imaginary part of $\langle \varepsilon \rangle$ as

$$\langle \varepsilon \rangle_2 = 9f\varepsilon_m^2 \frac{\varepsilon_2(\omega)}{(\varepsilon_1(\omega) + 2\varepsilon_m)^2 + \varepsilon_2(\omega)^2}$$

which is expressed in terms of ε and ε_m.

As a consequence, a resonance can be expected at a frequency ω_s given by $\varepsilon_1(\omega_s) = -2\varepsilon_m$. This is the so-called surface plasma resonance of the effective medium embedding the small spherical particles. For small particles, scattering of electrons off the surface gives an appreciable contribution to the total reciprocal lifetime:

$$\tau^{-1} = \tau_0^{-1} + \frac{v_F}{L_{eff}}$$

where τ_0, v_F, and L_{eff} are, respectively, the electron lifetime in the bulk, the Fermi velocity, and a measure of the size of the particle. Actually, the width of the observed plasma resonances in clusters has been generally well explained by this model. The additional term on the right-hand side accounts for the behavior of the plasma resonance (at least in terms of size-related broadening) in a rather wide-size range (or number of constituent atoms $N \gg 100$). The effects due to the size variation on τ are observable also in correspondence to the size domain, where the electronic energy level distribution is quasicontinuous.

This has a remarkable effect, for example, on the ultrafast electron dynamics in metal nanoparticles. During irradiation with ultrashort laser pulses (from picoseconds to femtoseconds timescale), a change is induced in the electron energy distribution due to the formation of a low-density nonthermalized high-energetic electron population. The acquired energy is redistributed by electron–electron (e–e) and electron–phonon collisions. The electron thermalization usually occurs on a subpicosecond timescale; therefore, a hot Fermi distribution with a temperature $T_e > T_0$ (where T_0 is the initial sample temperature) is formed. The hot electron system thermalizes with the lattice through electron–phonon collisions and other interactions, mainly with the nanoparticle surface, with characteristic decay times.

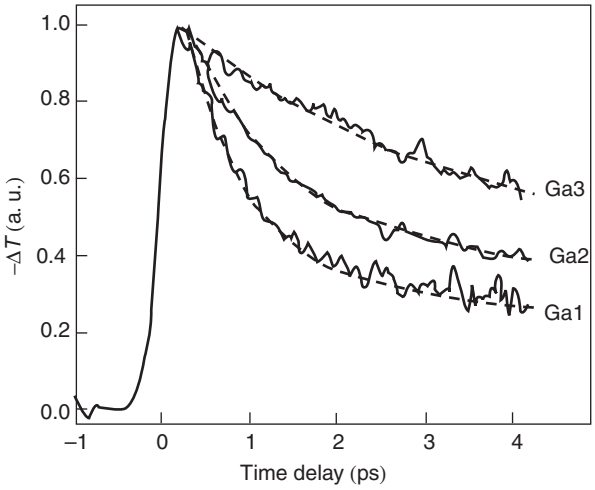


Figure 4 Time-resolved transmittivity changes ΔT in Ga nanoparticles embedded in a dielectric transparent matrix, as a function of probe time delay. Sample labels are Gan where n indicates the equivalent mass thickness of deposited Ga (in nanometers), the average radius of the nanoparticle being $R \sim 4n$.

The evident consequence is represented by a thermalization time which is size-dependent: the smaller the particle, the faster the thermalization (for example, the behavior of gallium nanoparticles, **Figure 4**). In fact, if R becomes comparable or smaller than the bulk electron mean free path, the electron oscillates inside the spherical potential well of the particle with a frequency $v = v_F/R$. In sufficiently small particles, this frequency can be larger than the Debye frequency, so that the bulk electron–phonon interaction (which is a resonant process) may become ineffective. In this case, the electrons dissipate the excess energy through generation of surface vibrational modes (currently defined "capillary waves"). In the case of gallium, the mean free path at 77 K is $l \sim 15$ nm. For smaller particles, the electron scattering with the surface takes place and the corresponding charge oscillation frequency $v = 1.82 \times 10^{14}\,s^{-1}$ turns out to be much larger than the Debye frequency $v_D = 6.6 \times 10^{12}\,s^{-1}$.

Metallic nanophase systems are also promising for applications in connection with the ease of growth combined with the wide flexibility in tuning electronic/thermodynamic properties. Applications of these systems include switching elements between different tunable states driven by an external force (optical or electrical pulses), and also the effect of phase transitions (in particular, melting) on their linear and nonlinear optical response.

See also: Metals and Metallic Alloys, Optical Properties of; Semiconductor Nanostructures.

PACS: 73.22. – f; 73.63. – b; 73.63.Bd; 73.63.Kv; 78.67. – n; 78.67.Bf; 78.67.Hc

Further Reading

Bimberg D, Grundmann M, and Ledentsov NN (1999) *Quantum Dot Heterostructures*. Chichester: Wiley.

Bonard J-M and Chatelain A (eds.) (1998) *Small Particles and Inorganic Clusters*, Proceedings of the Ninth International Symposium on Small Particles and Inorganic Clusters – ISSPIC9. Berlin: Springer.

Duncan MA (ed.) (1993–1998, 2001) *Advances in Metal and Semiconductor Clusters*, vols. 1–4. Greenwich: JAI Press; vol. 5. Amsterdam: Elsevier.

Haberland H (1994) *Clusters of Atoms and Molecules*, Springer Series in Chemical Physics 52. Berlin: Springer.

Kreibig U and Vollmer M (1995) *Optical Properties of Metal Clusters*, Springer Series in Material Science 25. Berlin: Springer.

Material Research Society (1996, 1999, 2001) *Nanophase and Nanocomposite Materials II*, *III*, and *IV*, Material Research Society Symposium Proceedings, vols. 457, 581, and 703. Warrendale, Pennsylvania: Material Research Society.

Material Research Society (2000, 2001) *Semiconductor Quantum Dots* and *Semiconductor Quantum Dots II*, Material Research Society Symposium Proceedings, vols. 571 and 642. Warrendale, Pennsylvania: Material Research Society.

Nomenclature

a_B	exciton Bohr radius (m)
C	capacitance (F)
e	electron charge (C)
E_F	Fermi energy (eV)
E_g	energy gap for semiconductors (eV)
E_R	exciton Rydberg energy (eV)
k	wave vector (m^{-1})
m^*	effective mass (kg)
R	cluster radius (m)
t	electron lifetime (s)
ν	frequency (s^{-1})
ν_D	Debye frequency (s^{-1})
ν_F	Fermi velocity (ms^{-1})
ε	dielectric function ($C^2N^{-1}m^{-2}$)
ω	angular frequency (s^{-1})

Semiconductor Compounds and Alloys, Electronic States of

J C Phillips, Rutgers University, Piscataway, NJ, USA

Like their predecessors, tetravalent elements (Si and Ge), semiconductor compounds (GaAs, InP, etc.), and alloys ((Al, Ga)As, (Ga, In) (P, As), etc.) based on tetrahedrally coordinated structures are the most studied and best understood materials in condensed matter physics. Like Si and Ge, they generally share eight valence electrons per cation/anion pair, and their formula can be written as A^NB^{8-N}. In addition to these "octet" semiconductors, there are also compounds and alloys with the formula A^NB^{10-N}, such as PbTe, that have a rock salt (octahedrally coordinated) structure, as well as many other predominantly covalent materials with more complex formulas, such as As_2Se_3, $GeSe_2$, etc. This article emphasizes the electronic states in the simplest octet case, partly because most of the applications are found there. Similar principles apply to the electronic states of all these materials, and the results are most impressive for the simplest cases. The ongoing progress of semiconductor science will probably lead to similar successes for other semiconductors as well.

Chemical Bonding in Octet Compounds

Because of their tetrahedrally coordinated structures, the electronic states of octet compounds and alloys can be described approximately as superpositions of Pauling's sp^3 hybrid atomic orbitals (**Figure 1**). In a diatomic structure there are two kinds of such hybrids, those pointing toward and away from the nearest neighbors, corresponding to bonding (valence) and antibonding (conduction) band states. When the two atoms in the unit cell are different and have different electronegativities, the covalent bonds found in elemental semiconductors (Si, Ge) become partially ionic. For many decades, the question of how to define the ionic and covalent fractions of chemical bonds was open and considered by many to be unanswerable, but this problem was solved for the octet compounds by Phillips' dielectric theory. It turns out that not only the electronic states but also almost all the physical properties of octet semiconductors are smooth functions (usually linear) of ionicity defined dielectrically, thus microscopically justifying much of Pauling's heuristic discussion of chemical bonding in molecules and solids, and explaining its many great successes.

In semiconductors, there are many energy gaps between the valence and conduction band states, the

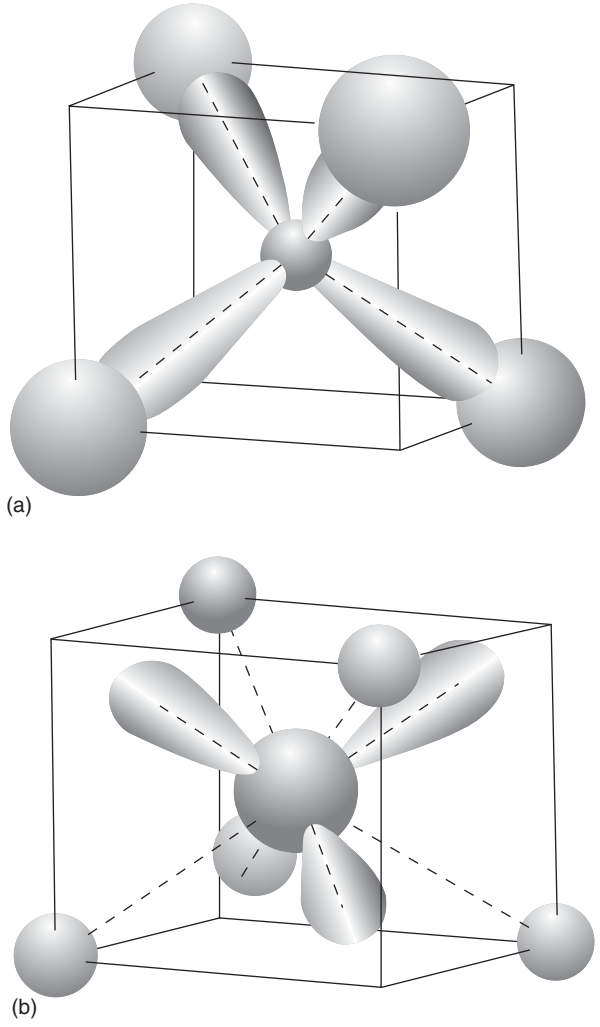


Figure 1 (a) Directed bonding and (b) antibonding orbitals in $A^N B^{8-N}$ semiconductors. The A atoms are represented by smaller spheres and are more electropositive, while the B atoms are represented by larger spheres and are more electronegative. The bonding orbitals have lower energy both because they are directed toward their nearest neighbors (covalency) and because they are centered predominantly on the more electronegative ion (ionic effect). These apparently qualitative effects are represented quantitatively (accuracy ~1%!) by the energy gaps E_h and C. (Reproduced from Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press, with permission from Elsevier.)

most important one for transport properties being the smallest gap. However, all the gaps follow chemical trends that are best described in terms of the average gap E_g defined dielectrically

$$\varepsilon(0) = 1 + A(\hbar\omega_p/E_g)^2 \qquad [1]$$

where A is a fixed number of order unity that corrects for band dispersion, and ω_p is the plasma frequency defined in terms of the valence electron density n_0 by

$$\omega_p^2 = 4\pi n_0 e^2/m \qquad [2]$$

Of course, in most cases, the average gap E_g is much larger than the smallest gap $E_{min} = E_0$. For example, in (Si, Ge) the gaps in electron volts are $E_g = (12.5, 11.5)$ and $E_{min} = (1.1, 0.7)$.

All the gaps follow chemical trends as functions of ionicity similar to those followed by the average gap, which is decomposed into its covalent or homopolar (E_h) and ionic or charge transfer (C) parts according to the Cartesian relation

$$E_g^2 = E_h^2 + C^2 \qquad [3]$$

The covalent gap E_h is a function only of the lattice constant, and is easily interpolated for compounds from the values of C, Si, Ge, and (gray) Sn. The ionic gap C is then obtained either from the known values of $\varepsilon(0)$ in the octet compounds, or from the analytic systematics of E_h and C obtained from other compounds. In horizontal series such as Ge, GaAs, and ZnSe, composed of elements from the same period, where the lattice constant changes are very small, it is found that C is proportional to $N - 4$. Thus, C behaves as a dielectrically defined electronegativity difference.

Pseudopotential Form Factors

The actual wave functions in semiconductors are not described accurately by the hybrid atomic orbitals shown in **Figure 1**, but they are described very accurately outside the atomic cores by superpositions of several hundred plane waves. The surprising aspect of these wave functions is that they can be derived from Cohen–Chelikowsky pseudopotential form factors $V_p(q)$ that depend on only a few parameters/atoms (if they are derived empirically from observed optical spectra) or that can be calculated from atomic potentials very precisely with no adjustable parameters; the values obtained empirically or from first principles are in excellent agreement. That is why one says that pseudopotentials are "transparent, reliable and transferable (TRT)." (Aside: this is a rare, virtually unprecedented, occurrence in quantum theory.) Without pseudopotentials, the knowledge and understanding of electronic states in solids would still be in a primitive, disorganized, and erratically inaccurate state.

Comparing the screened pseudopotentials for Al and Si shown in **Figure 2**, normalized to the free-electron Fermi energy E_F, it is to be noted that the form factors $V_p(q)$ of these adjacent elements in the

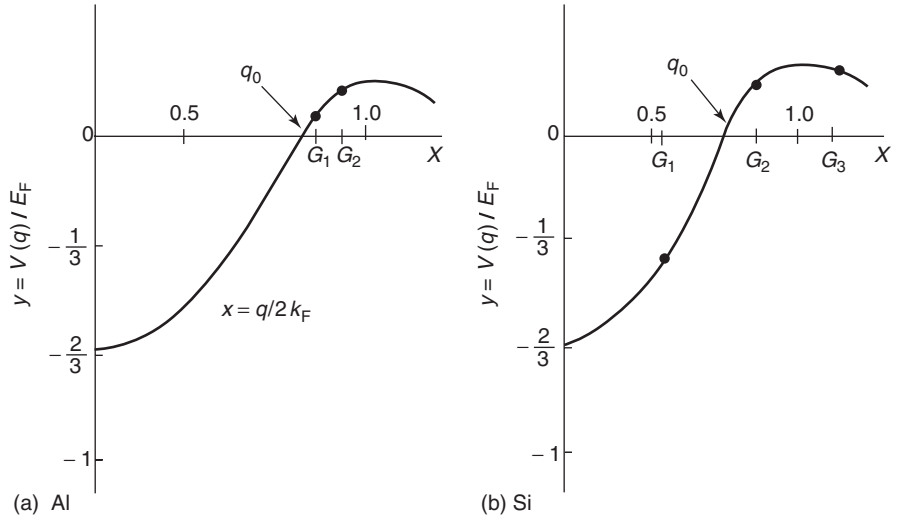


Figure 2 The electronic structures of solids arise from a complex interplay of electronic interactions at atoms and interatomic interactions between multiply scattered waves. Two such radically different materials as Al, a nearly free-electron metal, and Si, the semiconductor that is the basis of microelectronics, appear to have quite similar TRT pseudopotential form factors. The most striking differences arise because the reciprocal lattice vectors G_i are different for close-packed Al and tetrahedrally coordinated Si, leading to different values for $V(G_i)$, marked by the solid circles. However, a close examination of the form factors reveals other differences as well. Note that the number of valence electrons alone does not account for the differences in physical properties: Pb, like Si, is tetravalent, but it has the same close-packed crystal structure as Al. A close examination of the TRT pseudopotential calculations shows that the G that is largely responsible for this difference is $G_2 = (220)$, because $V(G_2)$ reverses the sign between Si and Pb. (Reproduced from Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press, with permission from Elsevier.)

periodic table as functions of $x = q/k_F$ appear at first to be quite similar, as one would expect from the fact that their atomic cores are isoelectronic. A closer inspection shows that the nodal value q_0 is smaller for Si than for Al, and the repulsive overshoot (due to exclusion of valence electrons from the atomic cores) is slightly larger. Most striking, however, are the positions of the reciprocal lattice vectors labeled G_i. In Al, with its close-packed 12-fold coordinated structure, the first two G's lie in the repulsive region beyond q_0, and covalent bonding is not possible. However, for Si q_0 lies well inside k_F and is large in magnitude and negative, favoring the formation of covalent bonds.

Thus far the atomic orbital and pseudopotential descriptions are quite similar. Now, however, a miracle happens. The covalent and ionic parts of the average energy gap are described quite accurately by

$$E_h/2 = V_s(G_2) + [V_s(G_1)]^2/\Delta T \qquad [4]$$

$$C/2 = 2V_s(G_1)V_a(G_1)/\Delta T \qquad [5]$$

where ΔT is a free-electron energy difference that depends only on the lattice constant, and $V_{s,a}(G) = V^A(G) \pm V^B(G)$. These energy gaps are central to understanding chemical trends in the electronic states of A^NB^{8-N} compound semiconductors and their alloys. They have been derived from tabulated atomic pseudopotential form factors without elaborate calculation. Their transferability has been tested elegantly for the series Mg_2X (X = Si, Ge, Sn), with a triatomic cubic fluorite structure and many new reciprocal lattice vectors, with the results shown in **Figure 3**. TRT successes of this kind are unique to pseudopotentials.

Ionicity

Given the covalent or homopolar (E_h) and ionic or charge transfer (C) energy gaps, one defines the bond ionicity as $f_i = C^2/(E_h^2 + C^2)$. Almost all bulk properties of A^NB^{8-N} compound semiconductors and their alloys are linear functions of f_i. (Strictly speaking, this statement has been tested only for 17 (!) compounds. These are: Si, Ge, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, ZnS, ZnSe, ZnTe, CdS, CdSe, CdTe, and CuCl. High-quality samples are gradually becoming available for AlN and GaN, and these should soon join the list.) The largest database against which the bond ionicity has been tested is 70 octet compounds that are either tetrahedrally or octahedrally coordinated. Success in predicting this fundamental covalent/ionic structural dichotomy had eluded all other attempts, but the parameter-free dielectric approach is completely successful, as shown in **Figure 4**.

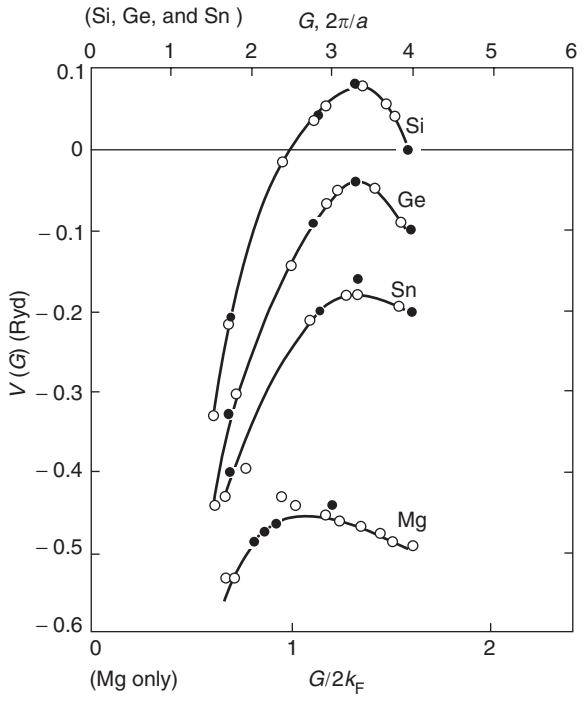


Figure 3 To test the assumption that pseudopotential form factors are transferable, self-consistent electronic structures were calculated for the X elements shown, and for the compound semiconductors Mg_2X. Values of the form factors derived from elemental band structures are represented by filled circles, while values for the compounds are marked by the empty circles. The overall consistency is spectacular, and validates the TRT description. (Reproduced with permission from Au-Yang MY and Cohen ML (1969) Electronic structure and optical properties of Mg_2Si, Mg_2Ge, and Mg_2Sn. *Physical Review* 178: 1358–1364; © American Physical Society.)

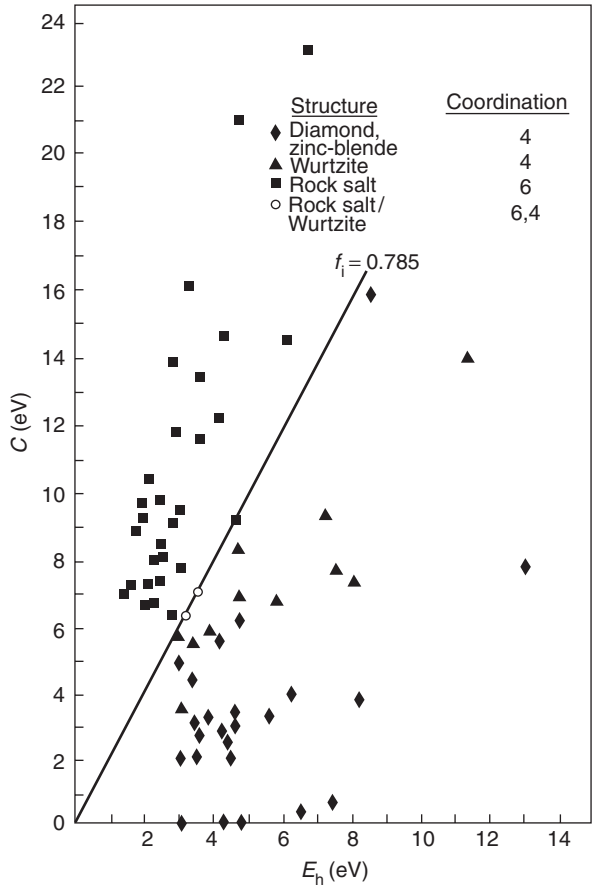


Figure 4 Values of E_h and C for ~70 octet compounds, including both tetrahedral semiconductors and octahedral insulators. The line corresponding to $f_i = C^2/(E_h^2 + C^2) = 0.785$ (5) separates the two classes exactly. (Reproduced from Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press, with permission from Elsevier.)

Charge Densities

The pseudoatom charge densities calculated using several hundred plane waves are in excellent agreement with X-ray results, after subtraction of the chemically uninteresting core background. A surprising feature of the charge densities is the appearance of a bond charge between the atoms. In elemental cases (diamond, Si, Ge, gray Sn), this charge is, of course, situated halfway between two atoms, but for compounds it moves closer to the anion. The displacement is linear in the ionicity f_i. Alternatively, one can integrate the excess charge density above the background. As shown in **Figure 5**, this excess charge extrapolates to zero as f_i approaches the critical ionicity derived in **Figure 4**.

Energy Bands of Octet Compounds

The energy bands of compound semiconductors have been calculated with great precision over a wide energy range (detailed comparisons with experiment cover an energy range of 10–20 eV) using the pseudopotential form factors illustrated in the next paragraph. The same form atomic factor occurs in many compounds, and the variations from one compound to another for the same atom are so small as to be scarcely perceptible. Two kinds of form factors have been used: semi-empirical and self-consistent first-principle, including in some cases corrections for many-electron correlation and exchange. To illustrate the differences between elemental band structures and compound band structures, one can compare (**Figure 6**) the energy bands of Ge and GaAs, which are quite similar because the main change in the crystal potential comes from the ionic potential C produced by the difference between Ga and As. The most important difference is that the smallest energy gap $E_{min} = E_0$ is indirect in the elemental semiconductors Si and Ge, that is, the valence and conduction band edges lie at different points in the crystal

momentum space k, but it is direct in most compound semiconductors, where both edges fall at $k = 0$. Almost all optical applications, including the most important one, multi-quantum well lasers, require a direct optical gap. The gap at $k = 0$ is denoted conventionally by E_0', that is associated with states near the (1 1 1) Brillouin zone face by E_1, and that is associated with states near the (2 0 0) Brillouin zone face by E_2.

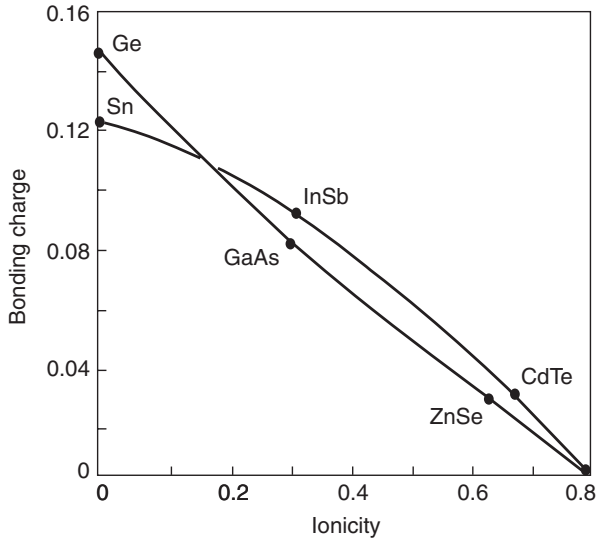


Figure 5 The bond charge as a function of dielectric ionicity for two isoelectronic series of octet semiconductors (third and fourth periods) clearly extrapolates to zero for the critical ionicity derived in **Figure 4**. (Reproduced with permission from Walter JP and Cohen ML (1971) *Physical Review* 4B: 1877; © American Physical Society.)

Elastic Constants and Phonons

The pseudopotential theory yields phonon spectra $\omega_i(q)$ for semiconductors by a convolution of the pseudopotential form factors in the conventional dynamical matrix theory, but this must be done on a case-by-case basis. More interesting are the general trends in the normalized shear modulus, shown in **Figure 7**; it is nearly linear in the ionicity f_i, except for collapsing in the Cu halides as the ionicity approaches the critical value 0.785 (**Figure 4**) for transition to the rock salt structure.

Fundamental Optical Spectra

At first sight, the optical spectra of semiconductors seem to be uninformative, because the energy bandwidth is large compared to the spacing of structures generated by crystal symmetry. However, modern TRT pseudopotential calculations yield very accurate energy bands, and oscillator strengths are routinely

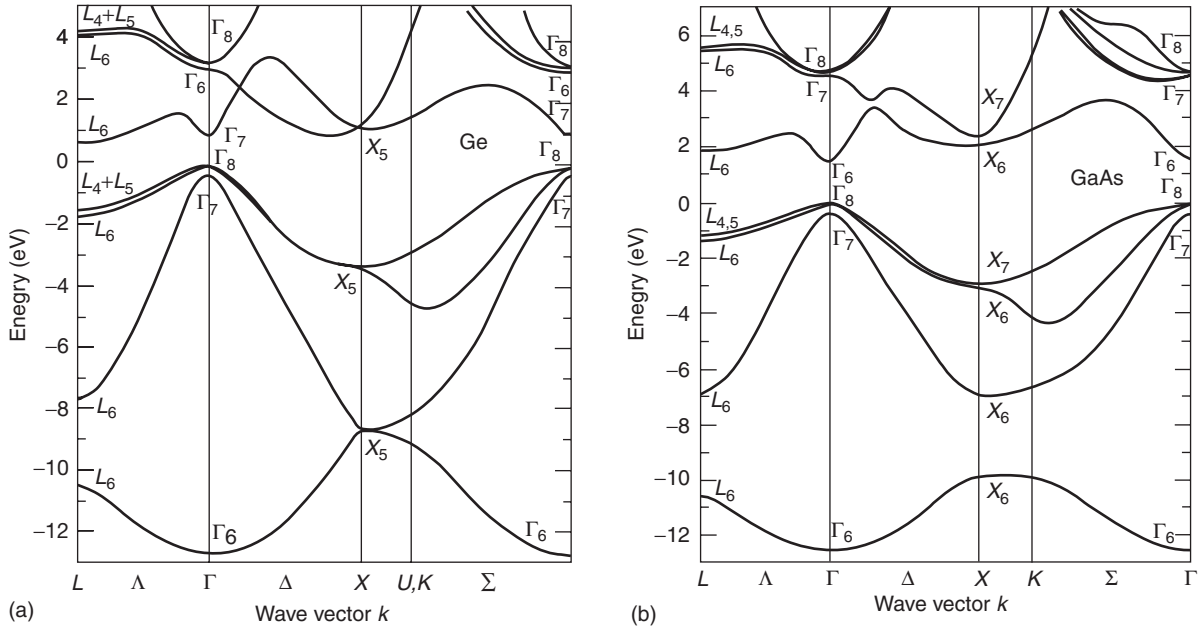


Figure 6 The energy bands of Ge and GaAs over an energy range of ~20 eV are quite similar; all the important differences in their physical properties stem from scarcely perceptible shifts in energies of their lowest conduction bands at the symmetry points Γ, *X*, and *L* in crystal momentum space *k*. These shifts, ~0.3 eV, are accurately predicted by TRT pseudopotential band structures, as these are calculated with an accuracy ~0.03 eV over the entire range of 20 eV. Accuracies of this kind are generally not achieved in quantum calculations even for polyvalent atomic spectra. (Reproduced with permission from Chelikowsky and Cohen (1973) *Physical Review Letters* 31: 1582; © American Physical Society.)

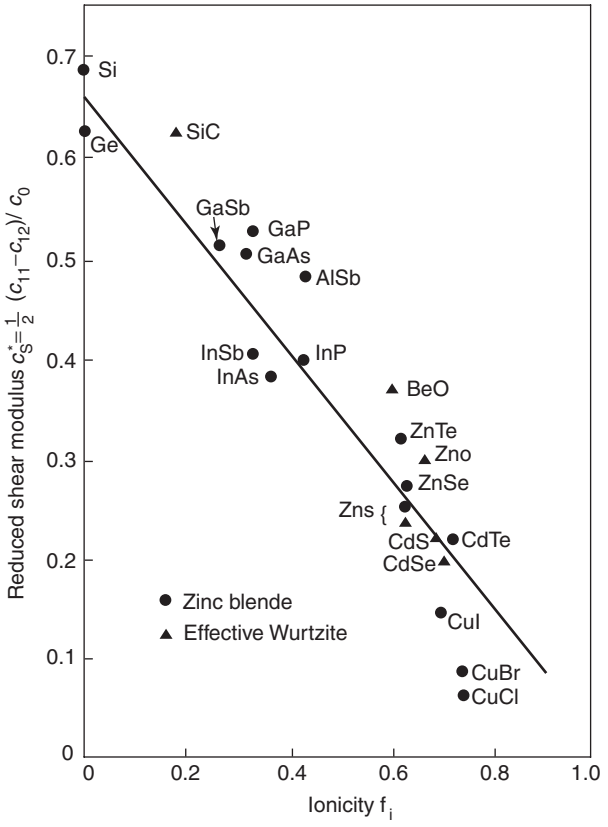


Figure 7 Increasing ionicity in octet compound semiconductors reduces the bond charge (**Figure 5**) and leads to the archetypical tetrahedral-octahedral phase transition (**Figure 4**). It also dramatically reduces the resistance of the open tetrahedral structure to shear. As shown here, this reduction appears to extrapolate to zero at $f_i = 1$, but in the Cu halides, there is a dramatic collapse as f_i approaches the critical value 0.785. This means that although there is a large volume change at the transition, so that it is technically first order, quantum mechanically it appears to be continuous, that is, second order, presumably because the bond charge (**Figure 5**) extrapolates to zero for the critical ionicity derived in **Figure 4**. (Reproduced from Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press, with permission from Elsevier.)

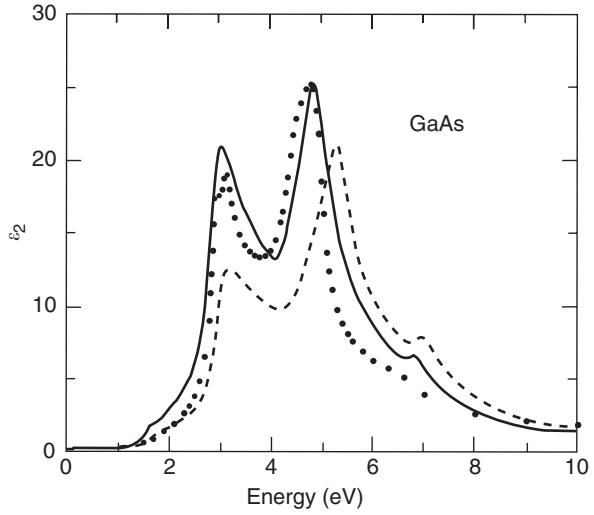


Figure 8 TRT pseudopotential calculations of the reflectance spectrum of GaAs (including oscillator strength redistribution to lower energies caused by interactions between excited electrons and holes) are in excellent agreement with experiment. (Reproduced with permission from Rohlfing M and Louie SG (1998) Electron–hole excitations in semiconductors and insulators. *Physical Review Letters* 81: 2312; © American Physical Society.)

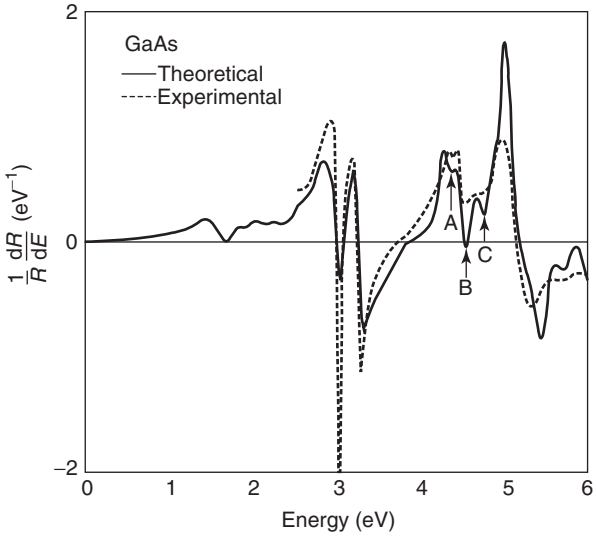


Figure 9 The derivative of the reflectance spectrum shown in **Figure 8** has also been calculated by one-electron TRT pseudopotentials; this derivative is less sensitive to many-electron interactions than the original spectrum, and the calculated results are in spectacularly good agreement with experiment. (Reproduced from Ricardo, Zucca RL, Walter JP, Shen YR, and Cohen ML (1970) *Solid State Communications* 8(8): 627–632, with permission from Elsevier.)

calculated including corrections for the Coulombic final state (excitonic) interactions. The resulting optical spectra agree so well with experiment (see **Figure 8** for GaAs) that often the discrepancies result merely from experimental problems with sample surfaces. Moreover, lock-in modulation methods uncover the fine structure in the reflectivity spectra (see **Figure 9** for GaAs) with an improvement in resolution of at least a factor of 10, as well as doubling the number of identified symmetry transitions. The static dielectric constant $\varepsilon_1(0)$ is connected to the absorptive dielectric function $\varepsilon_2(\omega)$ through the Kramers–Kronig relation as an integral over $\varepsilon_2(\omega)$. Thus if the average energy gap is $E_g^2 = E_h^2 + C^2$, then all the symmetry-derived peaks in $\varepsilon_2(\omega)$ should follow similar trends, using the same values of the dielectric electronegativity difference C. This is indeed the case; there are small downward shifts in peak energies relative to one-electron calculations, but

most of these are already incorporated into the definition of C from $\varepsilon_1(0)$, which itself includes similar corrections.

Electronic States of Defects

The most common defects in compound semiconductors are vacancies and antisite defects (i.e., A ions on B sites). The energies of formation and the electronic states associated with these defects depend on the position of the Fermi energy (the doping level) and strain. These energies can be estimated very accurately using pseudopotentials. The antisite defect energies are at least ten times smaller than the vacancy energies, and are comparable to thermal energies at growth temperatures. Thus antisite defects are the most important defects in most compound semiconductors. An example is strain at quantum well interfaces. Macroscopically, one would expect that such interfaces would be most stable when strain-free. In practice, a small amount of interfacial misfit is found to be optimal. This misfit may suppress diffusion of antisite defects; it is larger, for example, in InP-based cases (large antisite size difference) than in GaAs-based cases (small antisite size difference). Empirically, in manufacturing devices, the presence of these defects is reduced by the use of proprietary (confidential) methods. In other words, here science and technology are separated by important business considerations.

Surfaces and Interfaces

Another technologically important topic is the nature of the electronic states at surfaces and interfaces. The chemistry of these states is different in compound semiconductors than in elemental semiconductors. In the most important case – Schottky barriers between metals and compound semiconductors such as GaAs – the interface states for a midgap Fermi energy are taken mainly from the metal and the conduction band of the semiconductor.

PbS or A^NB^{10-N} Compounds

The two extra electrons outside the sp^3 octets are an extra s^2 lone pair associated with the cation: other examples of these compounds are GeTe, SnSe, SnTe, PbSe, and PbTe, as well as the elemental materials As, Sb, and Bi. The crystal structure is usually rock salt (sixfold coordinated). The weakly bound s^2 electrons produce strong absorption in the infrared region ($\sim$0.2 eV). The spin–orbit splittings of heavy elements such as Pb and Te are larger than the energy gaps, and produce unusual band structures calculated again very accurately by pseudopotential methods, in excellent

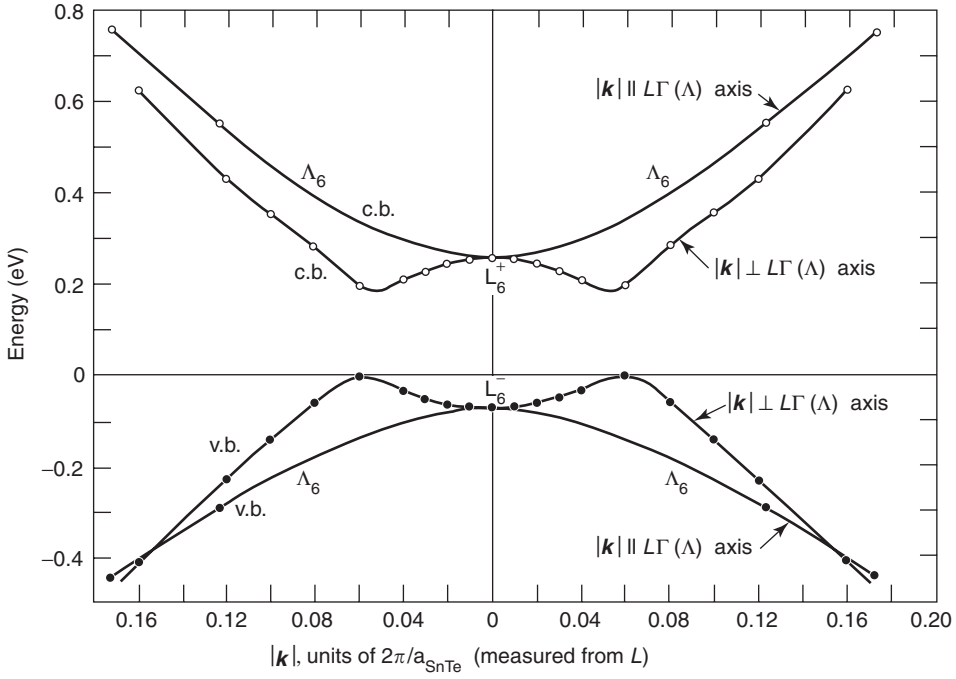


Figure 10 The structure of the energy bands of SnTe near the energy gap is complex. The complexity is caused by a combination of factors, including inversion of energy levels relative to octet compounds by the presence of lone pair electrons, and very large spin–orbit splittings. (Reproduced from Tung YW and Cohen ML (1969) *Physics Letters A* 29(5): 236–237, with permission from Elsevier.)

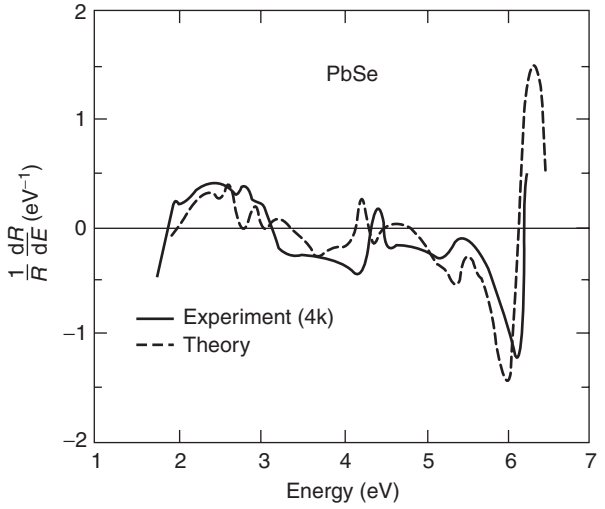


Figure 11 The modulated (derivative) reflectivity spectra of A^NB^{10-N} compounds calculated by one-electron TRT pseudopotentials are also in spectacularly good agreement with experiment. The example shown here is PbSe. (Reproduced with permission from Kohn SE, Yu PY, Petroff Y, Shen YR, Tsang Y, *et al.* (1973) *Physical Review B* 8:1477–1488; © American Physical Society.)

agreement with optical data. The band structures of these compounds can be quite complex, as illustrated for the example of SnTe in **Figure 10**. Again TRT pseudopotential derivative reflectivity spectra are in excellent agreement with experiment, as shown for the example of PbSe in **Figure 11**.

Ternary Octet Compounds

Further refinements and additional tunability (sometimes called band-structure engineering) are possible in two ways: through alloying (next paragraph) and through supercell formation (ternary compounds), the most important example being the octet chalcopyrite compounds $A^{II}B^{IV}C_2^V$ ($CuB^{III}C_2^{VI}$) which are pseudo III–V (II–VI) compounds with tetragonally distorted structures. Phase diagrams and accordingly sample preparation are complex in the chalcopyrites, in order for the cations to order on their respective sublattices in a unique polytype. Thus, the TRT pseudopotential theory has played a very important part in the development of these materials.

Electronic Structures of Alloys

The most important optical materials in practice (see following section) are semiconductor III–V alloys. The energy bands and optical transitions of these alloys are given approximately by linear interpolation between the corresponding structures of their end point pure compounds, as one would expect from a random distribution of ions on the alloying sublattices (ideal solution). (Note that ions alloyed on a sublattice are second nearest neighbors and hence interact much more weakly than atoms alloyed on the entire lattice, for instance, in (Si, Ge) alloys.) However, small curvatures are observable, and are of great practical importance for bandgap tunability; the curvatures are successfully predicted by treating the alloys as regular solutions.

Applications

As noted above, many compound semiconductors have electronic structures with valence and conduction band edges at $k = 0$ (the Brillouin zone center, often denoted conventionally by Γ), so that their lowest energy optical transitions are direct, a necessary condition for semiconductor quantum well lasers. The quantum wells themselves enhance the overlap between excited electron and hole wave functions, increasing oscillator strengths, lowering lasing thresholds, and increasing laser intensities. The transparency window of silica-based optical fibers is fixed by nature in the near infrared, and because of bandgap engineering it is possible to design semiconductor lasers that emit light at any desired wavelength in this window. The actual alloys used are quaternary (In, Ga) (P, As) alloys, and it is semiconductor multiquantum well lasers based on these alloys that have revolutionized civilization in the twenty-first century. The lifetimes of these lasers are fixed mainly by the quality of interfaces (semiconductor and semiconductor-metal). Other compound semiconductors are utilized in many other applications as far infrared detectors, and light emitting diodes are now manufactured over a wide range of colors up to the ultraviolet ((InGa)N).

See also: Elemental Semiconductors, Electronic States of; Light Emitting Diodes; Semiconductor Lasers.

PACS: 71.20.Nr, 71.15.Dx, 78.30.Fs, 73.21.Fg

Further Reading

Benedict LX, *et al.* (1999) Dielectric function of wurtzite GaN and AlN thin films. *Solid State Communication* 112: 129.

Bonapasta AA and Giannozzi P (2000) Effects of strain and local charge on the formation of deep defects in III–V ternary alloys. *Physical Review Letters* 84: 3923.

Brust D, Phillips JC, and Bassani F (1962) Critical points and ultraviolet reflectivity of semiconductors. *Physical Review Letters* 9: 94.

Cohen ML and Chelikowsky JR (1988) *Electronic Structure and Optical Properties of Semiconductors*. Heidelberg: Springer.

Kroemer H (2001) Nobel lecture: quasielectric fields and band offsets: teaching electrons new tricks. *Reviews Modern Physics* 73: 783.

Mahadevan P and Zunger A (2002) Room-temperature ferromagnetism in Mn-doped semiconducting $CdGeP_2$. *Physiological Review Letters* 88: 047205.

Maxisch T, Binggeli N, and Baldereschi A (2003) Intermetallic bonds and midgap interface states at epitaxial Al/GaAs(0 0 1) junctions. *Physical Reviews B* 67: 125315.

Nagahama S, Yanamoto T, Sano M, and Mukai T (2002) Blue-violet nitride lasers. *Physica Status Solidi A* 194: 423.

Pauling L (1960) *Nature of the Chemical Bond*. Ithaca, NY: Cornell University Press.

Wei SH, Zhang SB, and Zunger A (1999) Band structure and stability of zinc-blende-based semiconductor polytypes. *Physiological Reviews B* 59: R2478.

Semiconductor Devices

T Takahashi, University of Tokyo, Tokyo, Japan

Introduction

Semiconductor devices are key devices in modern electronics, and are roughly categorized into electronic devices and optical devices. The main applications as electronic devices are junction diodes, transistors, and memories, which are widely used in large-scale integrated circuits (LSI), as well as power devices such as thyristors and so on. As optical devices, light emitting diodes (LEDs), laser diodes (LDs), and photo detectors (PDs) are utilized in optical display and communication systems, while solar cells are important for energy conservation. In the following sections, the basic principles of these devices are described.

Junction Diodes

A junction of *p*-type and *n*-type semiconductors makes a diode structure as shown in **Figure 1**. Electrons or holes are the majority carriers in the *n*-type or *p*-type semiconductor, respectively. In the thermal equilibrium condition, the electrons in the *n*-type region near the *p–n* junction interface diffuse toward the *p*-type region because of the concentration gradient, and the holes in the *p*-type region diffuse vice versa. Then a depletion region or a space charge region, where no free carriers exist, appears. If one assumes an ideal *p–n* junction, which is very abrupt and the depletion completely occurs as shown in **Figure 2**, the built-in potential V_{bi} that appears across the junction is given by

$$V_{bi} = \tfrac{1}{2} E_m (x_p + x_p)$$

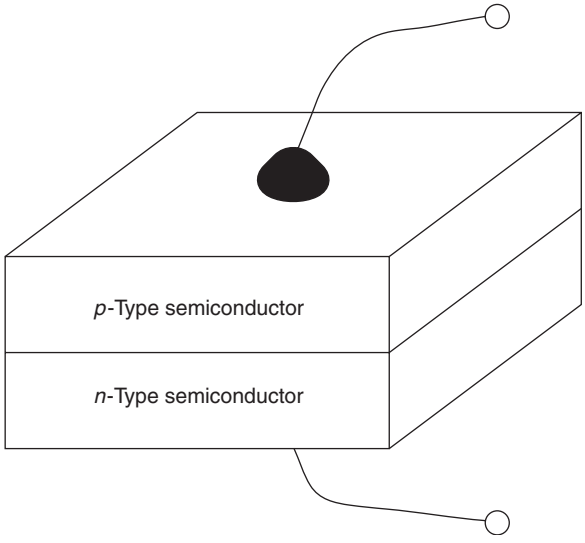


Figure 1 A schematic illustration of a *p–n* junction diode.

where

$$E_m = \frac{qN_A x_p}{\varepsilon_s} = \frac{qN_D x_n}{\varepsilon_s}$$

and the depletion region width W is given by

$$W = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{N_A + N_D}{N_A N_D}\right) V_{bi}}$$

(here, q is the unit charge and ε_s is the dielectric constant).

When a positive bias V is applied to the *p*-type region with respect to the *n*-type region, the holes in the *p*-type region easily move into the *n*-type region, and the electrons in the *n*-type region move in the opposite direction. Thus a large current flows across the junction. Such a bias condition is called a forward bias. When, on the other hand, the bias condition is inversed, which is called a reverse bias, the majority carriers are not able to move at all. Then the

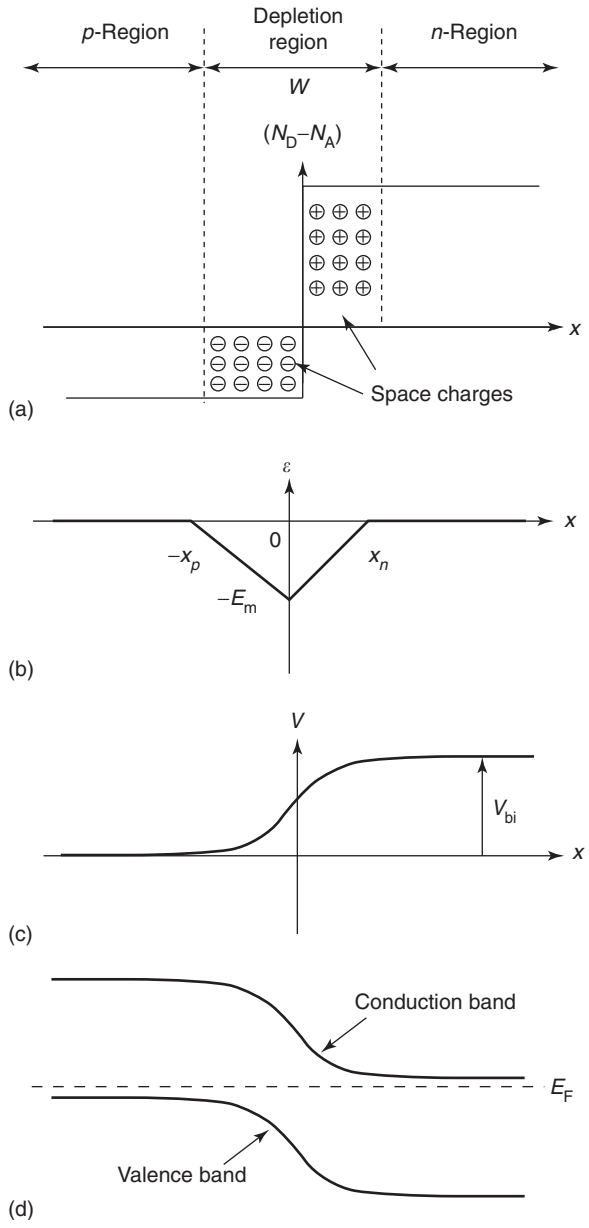


Figure 2 (a) Space charge distribution, (b) electric field distribution, (c) potential variation, and (d) energy band diagram of an ideal *p–n* junction in thermal equilibrium condition.

total current density J flowing across the junction is given by $J = J_s(e^{qV/kT} - 1)$, where J_s is the saturation current density, k is the Boltzmann constant and T is the temperature. Therefore, the diode has asymmetric current–voltage characteristics against the bias voltage V as shown in **Figure 3**, and the diode device achieves a rectification property for alternating currents. This rectification effect in the *p–n* junction is very important in most semiconductor devices because it can control the current flow by blocking undesirable current leakage. In the real *p–n* junction diode, however, a voltage breakdown occurs at a high reverse bias, which gives the reverse bias tolerance of the diode, also shown in **Figure 3**.

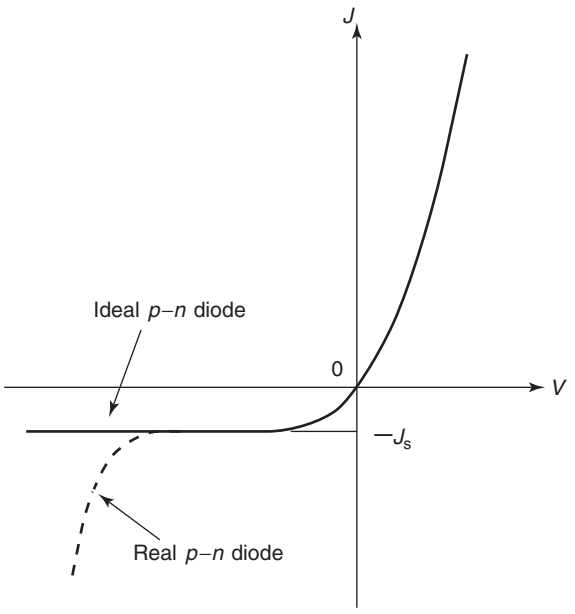


Figure 3 Current voltage characteristics in the *p–n* junction diode.

Bipolar Transistors

The bipolar transistor which was invented by a research group at Bell Laboratories in 1947 has become one of the most important semiconductor devices today. It consists of two *p–n* junctions, that is, *p–n–p* or *n–p–n* structures as shown in **Figure 4**. These three regions are called emitter, base, and collector. In the symbols of the *p–n–p* and *n–p–n* transistors, the arrow indicates the direction of current flow under normal operating conditions, where forward and reverse biases are applied to the emitter–base and collector–base junctions, respectively. The specific features in the bipolar transistor are the very thin base and the highly doped emitter. The latter brings a high injection of carriers into the base region, where the injected carriers becomes the minority, while those carriers pass through the base region into the collector with a very small recombination rate owing to the former feature. Therefore, the collector current I_C can be represented by the emitter current I_E as $I_C = \alpha I_E$, where α is slightly less than 1. In general, a forward bias V_{EB} is applied to the emitter–base junction, but the collector–base bias V_{CB} is in a reverse bias condition. Since the collector junction is highly and reversely biased, the power amplification against the input power to the emitter junction is achieved. In the following discussion, the *p–n–p* transistor is considered but an appropriate

change of polarities makes the results applicable to the *n–p–n* transistor. The basic model for the bipolar transistor is the Ebers–Moll model, in which the transistor is represented by two diodes connected face to face and two current sources. A circuit diagram in the Ebers–Moll model for the *p–n–p* transistor is illustrated in **Figure 5**. In this model, a current I_R flowing through the collector–base diode determines a current value from the current source in the emitter side, and a current I_F through the emitter–base diode determines a current from the current source in the collector side. Here, amplitudes of current gain α_R and α_F that are barely less than 1.

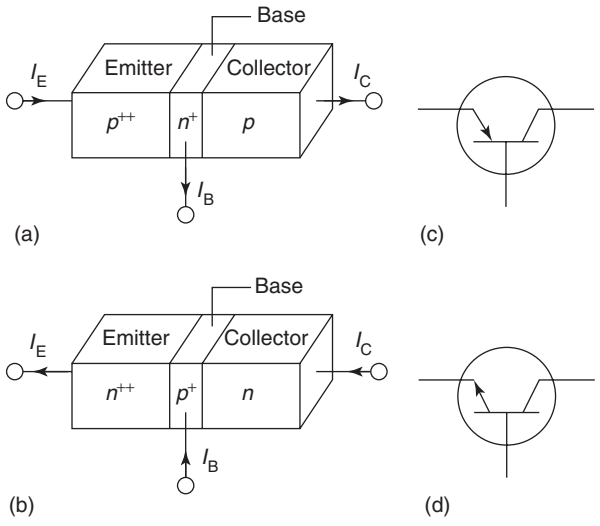


Figure 4 Schematic illustrations of (a) *p–n–p*, (b) *n–p–n* bipolar transistors and their symbols (c–d).

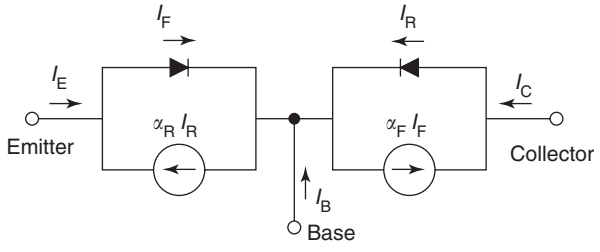


Figure 5 A circuit diagram of the Ebers–Moll model of the *p–n–p* bipolar transistor.

There are three configurations of the bipolar transistors as shown in **Figure 6**. In the common-base configuration, where the base lead is used as a common terminal for input and output, a signal is inputted to the emitter–base junction that is forwardly biased, while a load resistance is connected to a side of the reversely biased collector–base junction to extract an output signal. The common-base configuration is suited for a low input impedance circuit. According to the Ebers–Moll model, the collector current I_C is given by $-\alpha_F I_F$ because I_R is almost zero. In this configuration, therefore, only voltage amplification is obtained without current amplification. On the other hand, the common-emitter configuration, in which the emitter lead is used as a common terminal, is most widely used, and the current amplification as well as the voltage amplification is achieved because the collector current I_C is given by $I_C = \alpha_F I_B/(1-\alpha_F)$ and the collector–base junction is still reversely biased. On the contrary, the common-collector configuration gives only a current gain without the voltage amplification, which is useful as an impedance converter. The typical output characteristics for the bipolar transistor in the common-base and common-emitter configurations are illustrated in **Figure 7**.

Since a current density strongly depends on the *p–n* junction cross section, and a switching speed is determined by the base region thickness in the bipolar device, they can easily be improved by using a large junction cross section and a thin base region, respectively. Therefore, bipolar devices are also suited for high speed and/or power devices. The three *p–n* junctions in series, such as *p–n–p–n* or *n–p–n–p* structures, are used in three- or four-terminal devices, in which bistable operations with low-power dissipation is possible. These devices are called thyristors and are applicable in high-current, high-voltage, or high-power switching devices.

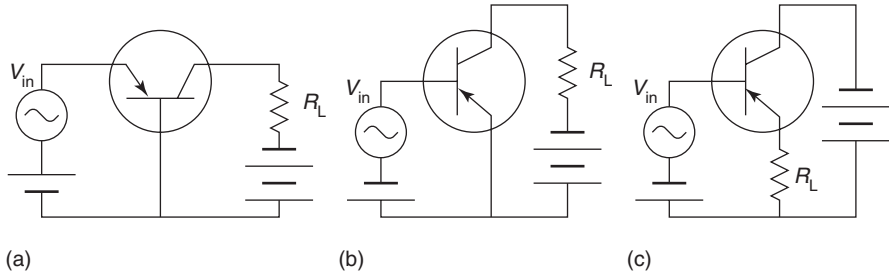


Figure 6 Configurations of the bipolar transistors; (a) common-base configuration, (b) common-emitter configuration, and (c) common-collector configuration.

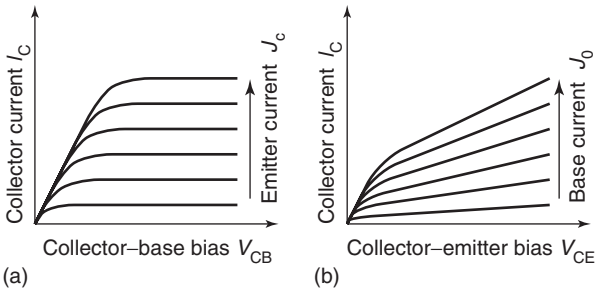


Figure 7 Typical output characteristics for the bipolar transistor in the (a) common-base, and (b) common-emitter configurations.

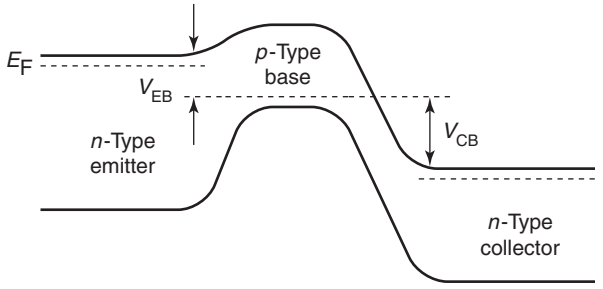


Figure 8 A schematic of the *n–p–n* heterobipolar transistor.

One may change one or all the lead materials in the bipolar devices in order to be different from others as shown in **Figure 8**; such devices are called heterobipolar devices. In heterobipolar devices, many performances are well improved. For instance, (1) high-emitter efficiency owing to the blockade of the hole injection (the minority carrier injection for the emitter) from the base into the emitter by the high barrier in the valence band, (2) low-voltage drop at the emitter–base junction, and (3) high-frequency response owing to a high-current gain and low-base resistance because heavy doping in the base region is possible without degrading emitter efficiency.

Field Effect Transistors

Similar to the bipolar transistor, a field effect transistor (FET) has three terminals, called source, gate, and drain. There are some variations of FET structures as illustrated in **Figure 9**. In the junction FET(JFET), which was the first proposed of all FET devices, the gate bias voltage varies the depletion region width at the *p–n* junction under the gate to control a channel conductance. In the metal–semiconductor FET (MESFET), a Schottky junction is formed between a gate metal and a channel semiconductor, and the depletion width is varied by the gate voltage. On the other hand, a metal-insulator-semiconductor (MIS) structure, in which an insulator

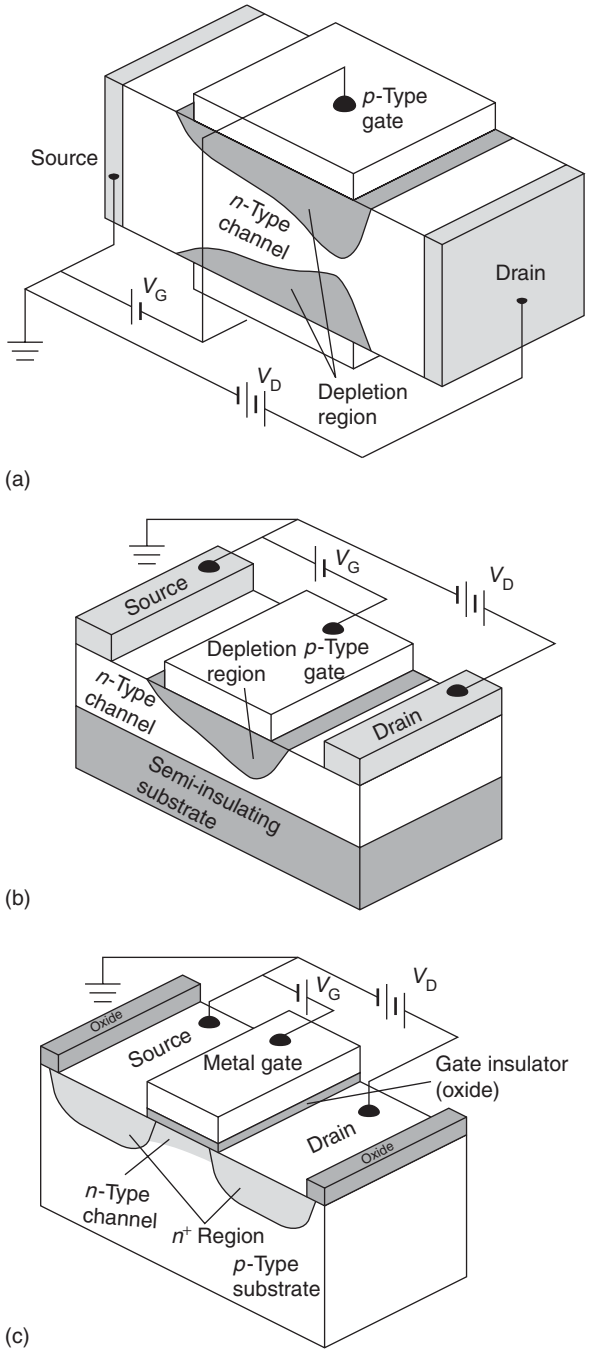


Figure 9 Some variations of the FETs; (a) a JFET, (b) a MESFET, and (c) a MISFET. In the case of using an Si dioxide film as an insulator, the FET is specially called a MOSFET.

film is inserted under the gate electrode to prevent current leakage between the gate and the channel, is suited for FET devices. Especially, silicon dioxide is widely used as an insulator in silicon FETs, which are called metal-oxide–semiconductor FET (MOSFET). In contrast to the bipolar transistor in which both positive and negative charges play important roles, only unipolar charges flow in the FET channel. Thus

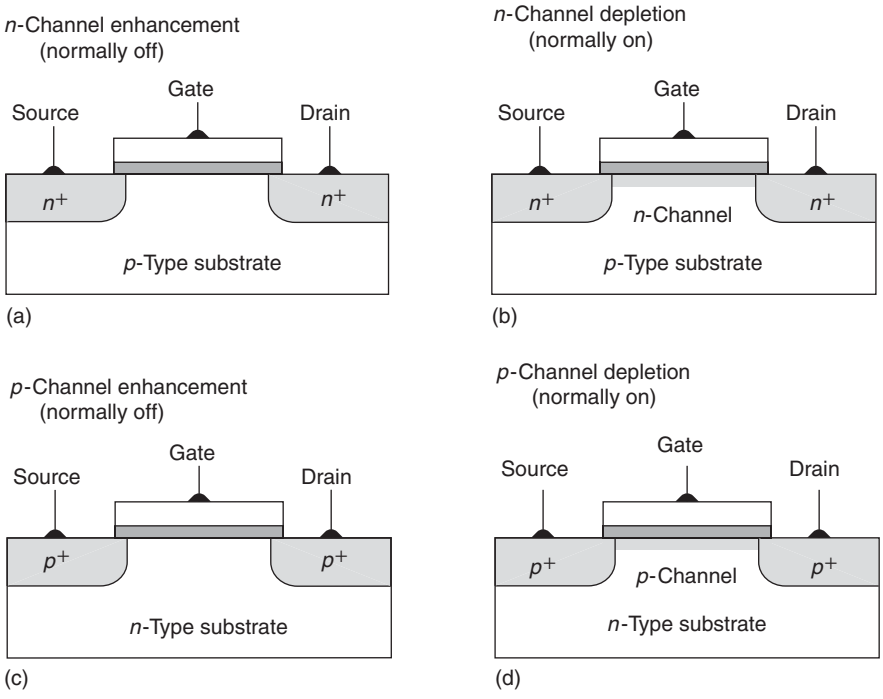


Figure 10 Four basic types of MOSFETs at zero gate voltage.

the FET is classified into either an *n*-channel or a *p*-channel FET. It is also possible to categorize MOSFET by the conductivity at zero gate bias into a normally-off type or a normally-on type as shown in **Figure 10**. If the channel conductance is very low at the zero gate bias and a positive bias to the gate is required to form a conductive *n*-channel, this type is called the normally-off (or enhancement) *n*-channel MOSFET. The MOSFET, in which the *n*-channel is sufficiently conductive even at the zero gate bias and one has to apply a negative bias to the gate to reduce the channel conductance by depleting the channel carriers, is the normally-on (or depletion) *n*-channel MOSFET. Similarly, there are the normally-off (enhancement) and normally-on (depletion) *p*-channel MOSFETs.

Now, a complementary use of the *p*-channel and *n*-channel MOSs in series, for instance, the normally-on *p*-channel and the normally-off *n*-channel MOSs, forms a complimentary MOS (CMOS) configuration as shown in **Figure 11**. In the CMOS, a low-input voltage (typically, $V_{low} = 0\,V$) keeps the *n*-channel enhancement MOS (normally-off) in the "off-state" (highly resistive) and the *p*-channel depletion MOS (normally-on) in the "on-state" (highly conductive). Therefore, an output is equal to the high source voltage. In contrast, a high input (typically, $V_{high} = 5\,V$) turns the *n*-channel MOS into an "on" state and the *p*-channel MOS into an "off" state, and therefore,

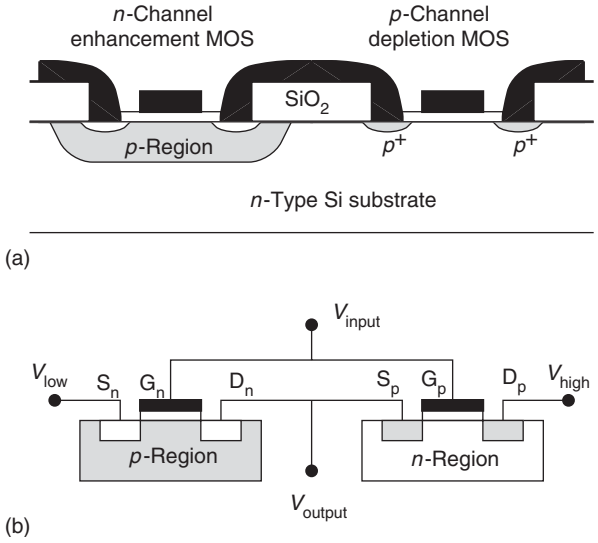


Figure 11 (a) A schematic of the CMOS structure, and (b) its equivalent electrical scheme.

the output becomes low level. Thus an inverter operation is realized by the CMOS configuration. In addition, a steady current passing through both the *n*-channel and *p*-channel MOSs can be sufficiently suppressed and the passing current only flows at the switching period, resulting in a very low power consumption in the circuits. Moreover, since the FET suits a planar configuration, CMOS transistors are mostly used in an integrated circuit (IC) or LSI.

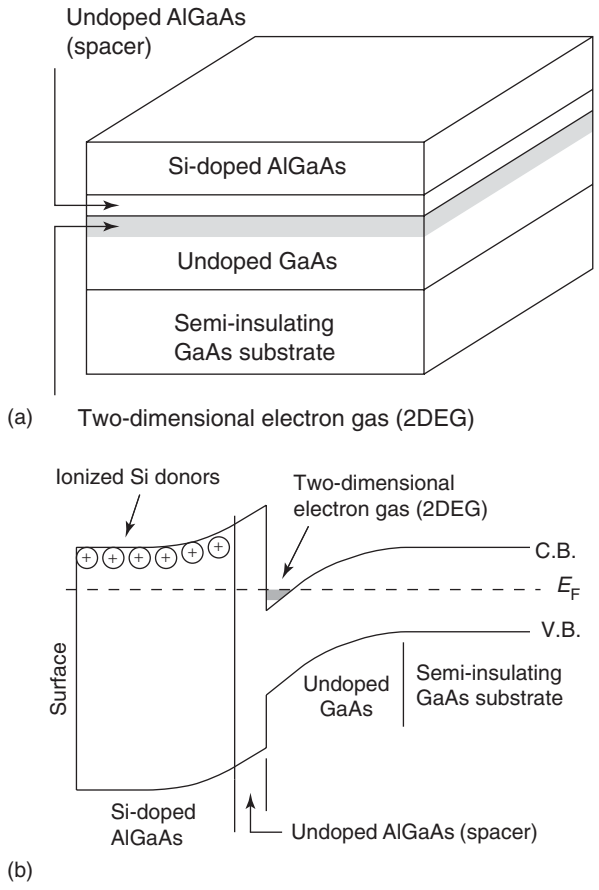


Figure 12 (a) A device structure, and (b) an energy band diagram of the HEMT realized by using the semiconductor hetero-structures.

By using compound semiconductor hetero-structures and a modulation doping technique, a high electron mobility transistor (HEMT) was realized as shown in **Figure 12**. The potential confinement of electrons at the heterointerface gives rise to a two-dimensional electron gas (2DEG). In the 2DEG system, the electron scattering probability is reduced because the states that the electrons can occupy are limited. In addition, the modulation doping also reduces the electron scattering rate by impurities because of a spatial separation between the conducting electrons and the impurities. Owing to these two effects, very high electron mobility is achieved in the HEMT device. Consequently, the HEMT device is practically used in high-speed applications.

Memories

Another key device for the LSI circuits is a memory device. A conventional memory cell needs capacitors, in which data are stored as charges, and the memory devices can be categorized into volatile and nonvolatile types. The volatile memory allows a very high access rate though it consumes a large quantity of power for keeping data, a typical device of which is a dynamic random access memory (DRAM). The DRAM, therefore, is suited for main memories in the LSI circuits. In order to integrate a lot of DRAMs in the LSIs, a deep-trench structure is used for keeping capacitance sufficiently large. In the nonvolatile one, on the other hand, the access time is relatively long but data can be kept without power consumption; typical devices are floating gate RAM (FRAM) and static RAM (SRAM), which are applied to mobile telephones, for instance.

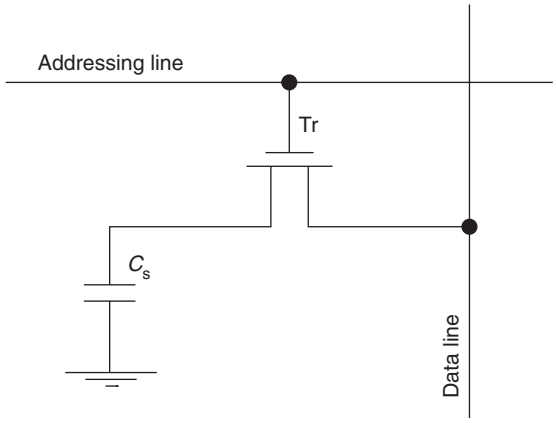


Figure 13 A schematic of a one-bit memory cell in the simplest DRAM.

In most of the cases, the memory is fabricated by the compatible process of CMOS technologies. To access the memory, charging or discharging the capacitor is done for writing or reading, respectively, through the transistor connected to an addressing line. **Figure 13** shows a one-bit memory cell in the simplest DRAM. For writing, the transistor channel is opened by an appropriate bias applied to the addressing line with a high bias applied to the data line, and subsequently the capacitor C_s is charged. For reading, on the contrary, the transistor channel is also opened and a voltage appearing in the data line is detected. If the capacitor has been charged, a certain value of voltage appears, while no voltage appears when no charge has been stored in the capacitor. These two states correspond to "1" and "0" as one bit in the binary logic.

Photodetectors and Photodiodes

Semiconductor photodetectors can detect optical signals through electronic processes. Light incident on a semiconductor excites electrons from the valence band to the conduction band, if the photon energy of the incident light exceeds a semiconductor

bandgap energy, and, then, a current or a voltage is generated in the semiconductor. Thus, the photodetector acts as a current source or a voltage source under the light illumination.

Generally, photodiodes consist of *p–n* or *p–i–n* (*i* means an intrinsic layer) structures. An electric field, existing at the depletion region around the *p–n* interface or the intrinsic region, is very important to separate photogenerated electron–hole pairs. The thin depletion region leads to a high-speed operation but low quantum efficiency (a ratio of the generated electron–hole pairs to the incident photons), while the thick one leads to the opposite. The basic operation mode of the *p–i–n* photodiode is illustrated in **Figure 14**. When the reversal bias applied to the *p–i–n* structure enhances a separation rate of the photogenerated electron–hole pairs, a high-speed operation is possible even if the depletion region is thick enough to achieve a high quantum efficiency. Especially, when a very high electric field is applied in the thick intrinsic region, secondary ionization occurs and thus the number of electrons and holes dramatically increases, resulting in quite high efficiency, which is called an avalanche photodiode (APD).

The detectable wavelength range of the photodiode is an important parameter. The long-wavelength cutoff is determined by the bandgap of the semiconductor material used in the photodiode because of the lower photon energy light, which means that a wavelength, longer than the bandgap cannot be absorbed in the semiconductor. On the other hand, the absorption coefficient for a short-wavelength light is quite high, so such a light is absorbed very near the surface. Since, the carrier recombination occurs very frequently via the surface states, most of the photocarriers recombine without either being collected in the *p–n* junction or contributing to the current. This limits the short-wavelength cutoff in photodiodes.

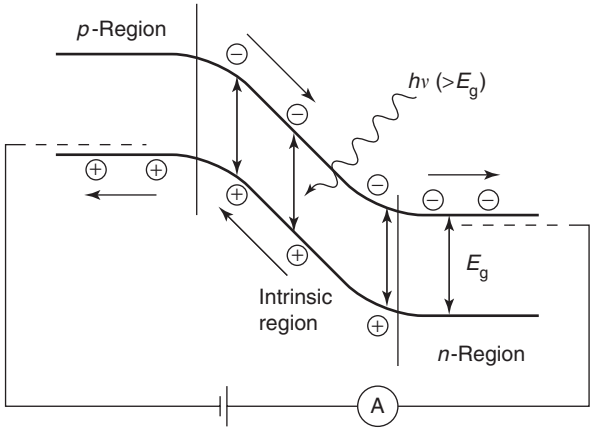


Figure 14 Basic operation mode of the *p–i–n* photodiode.

Solar Cells

A similar *p–n* diode structure in the photodiode can be used as a solar cell, which works as an electricity source. In particular, a detectable wavelength range in the solar cell is well optimized to fit the spectrum of sunlight. The photocarriers generated in the *p–n* diode are extracted as a current source, while photovoltage appears between the *p*- and *n*-type semiconductor ends due to the photocarrier separation. A current–voltage curve in the *p–n* diode changes under the light illumination as shown in **Figure 15**. The property of the solar cell is characterized by three important parameters; the short-circuit current I_{SC}, the open-circuit voltage V_{OC}, and the maximum power rectangle P_M. High I_{SC} and V_{OC}, resulting in large P_M, are desirable. Since the photoabsorption spectrum of Si material matches the incident sunlight spectrum to some extent, Si is one of the best materials for solar cells. Si solar cells have already been used as alternative energy resources in both high- and low-power applications.

Light Emitting Diodes

When the electron–hole pair disappears by recombination, it should convert its energy to other constituents because of the energy conservation law. In direct transition semiconductors such as GaAs, photons whose energy corresponds to the recombination energy are emitted, while phonons are emitted, in

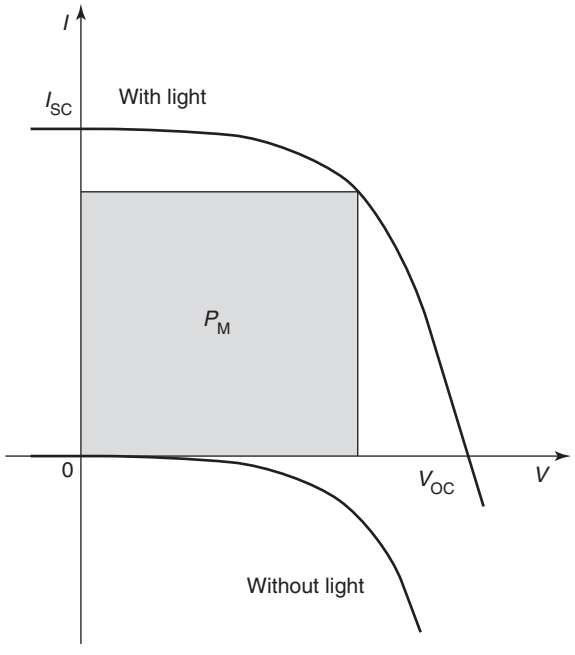


Figure 15 Typical current–voltage characteristics in the *p–n* solar cell with or without the light illumination.

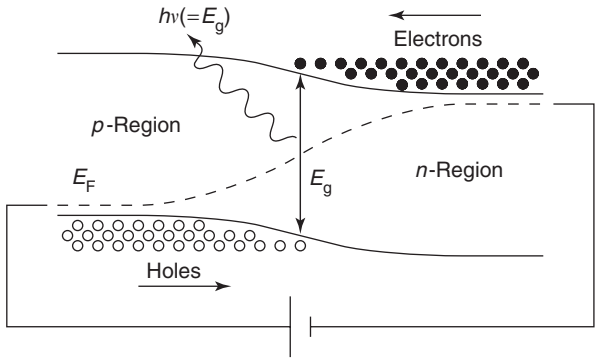


Figure 16 Electron and hole injections through *n*-type and *p*-type regions, respectively, in the LED. The emitted photon energy is almost equal to the bandgap energy.

most cases, in indirect transition semiconductors, and the semiconductor materials are heated. Especially where electrons and holes are intentionally injected from electrodes, a current–light conversion is achieved in the direct transition semiconductors. The *p–n* or *p–i–n* diode structures are suited for devices similar to photo detectors, and they are called LEDs. The electrons supplied from the *n*-type region and the holes from the *p*-type region recombine near the *p–n* interface or in the intrinsic region with light emission as illustrated in **Figure 16**. Since most electrons and holes distribute at the conduction and valence band edges, respectively, the emitted photon energy is almost equal to the bandgap energy. Therefore, the wavelength of the emitted light can be selected by an appropriate choice of the semiconductor materials. Especially by using compound semiconductors, the wavelength of the LED can be tuned in a wide range, for instance, from orange to near-infrared through red wavelengths by $A1_xGa_{1-x}As$ or $GaAs_{1-x}P_x$ materials with various mole fraction x. Very recently, green or blue LEDs were achieved by GaN-related materials. As a result, semiconductor LEDs can realize full color display devices.

Laser Diodes

In LEDs, the light is emitted through a spontaneous recombination process of the electron–hole pairs. On the other hand, the electron–hole recombination is stimulated if there exists a light whose photon energy is same as the recombination energy. This stimulated recombination process plays an important role in lasers. Since the basic structure of the semiconductor laser is almost same as that of the LED, semiconductor lasers are often called laser diodes (LDs). The difference between LDs and LEDs is that the LDs have a cavity for lasing as shown in **Figure 17**. In the cavity, the light travels back and forth and is

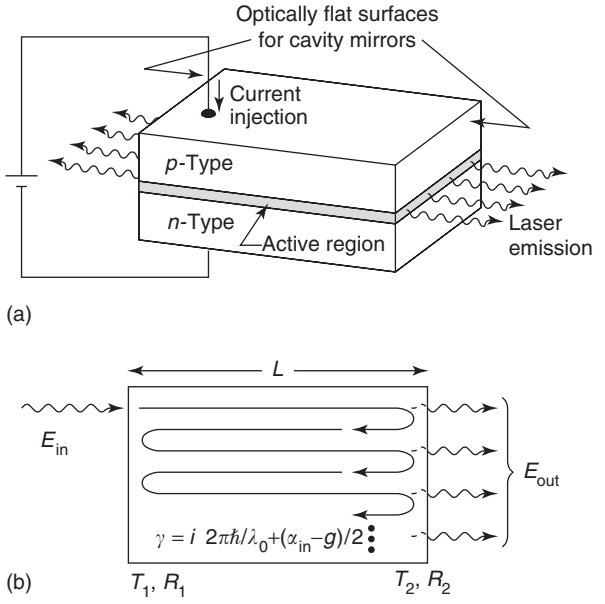


Figure 17 (a) A schematic of the LD where a laser cavity is added to the LED structure to achieve the laser oscillation, and (b) a model of light traveling in the laser cavity.

amplified by the stimulated emission. It is assumed that an incident plane wave to the cavity is $E_{in} = E_0 e^{j(\omega t - kx)}$, reflectances for input and output mirrors are R_1 and R_2, transmittances for the input and output mirrors are T_1 $(=1-R_1)$ and T_2 $(=1-R_2)$, respectively, and a parameter representing a change in both amplitude and phase for one way in the cavity is γ. Then the total output E_{out} is given by

$$\begin{aligned} E_{out} &= E_{in}T_1T_2e^{-\gamma L} + E_{in}T_1T_2R_1R_2e^{-3\gamma L} \\ &\quad + E_{in}T_1T_2R_1^2R_2^2e^{-5\gamma L} + \cdots \\ &= E_{in}T_1T_2e^{-\gamma L}(1 + R_1R_2e^{-2\gamma L} \\ &\quad + R_1^2R_2^2e^{-4\gamma L} + \cdots) \\ &= E_{in}T_1T_2e^{-\gamma L}\frac{1}{1 - R_1R_2e^{-2\gamma L}} \end{aligned}$$

If $1 - R_1R_2e^{-2\gamma L} = 0$ is satisfied, a finite output is obtained even when $E_{in} = 0$, and this is the oscillation condition for lasers. Now, γ can be rewritten as $\gamma = i2\pi\bar{n}/\lambda_0 + (\alpha_{in} - g)/2$ where $\lambda_0, \bar{n}, \alpha_{in}$, and g are the wavelength in vacuum, the refractive index, the cavity internal loss, and the optical gain in the cavity, respectively. Therefore, one can introduce the following equations from the amplitude and phase condition for lasing:

$$\begin{aligned} &R_1R_2\exp[(g - \alpha_{in})L] = 1 \\ &\quad - 2i(2\pi\bar{n}/\lambda_0)L = i2m\pi \quad (m = 1, 2, 3, \ldots) \end{aligned}$$

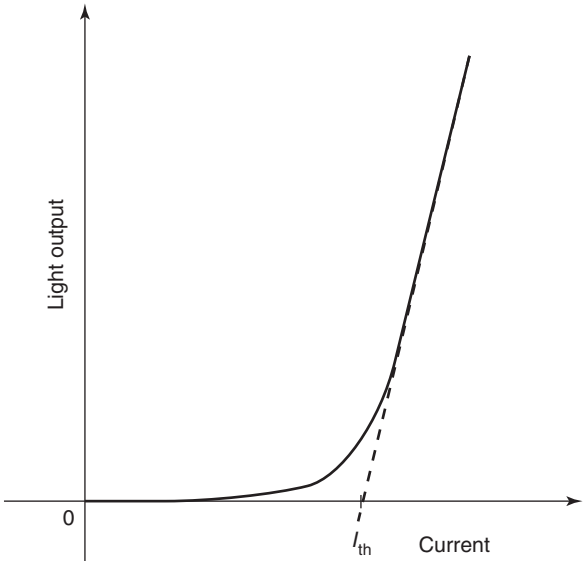


Figure 18 Typical current–light characteristics in the LD.

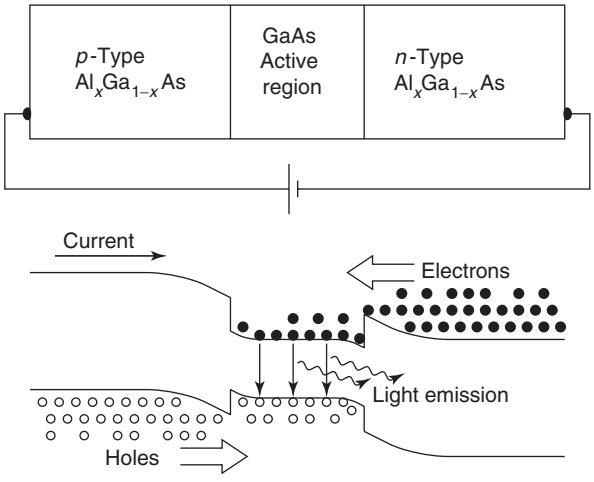


Figure 19 A double heterostructure to confine the carriers in the active region efficiently.

From the former equation, the threshold gain for lasing is given as $g = \alpha_{in} + (1/L)\ln R_1R_2$, and the longitudinal mode of lasing light wavelength is given as $\lambda = 2L/m$ from the latter one. Consequently, if the amplifying gain owing to the carrier injection from outside is high enough to exceed the cavity losses (summation of internal loss and mirror loss), a laser emission occurs, that is, some threshold current is needed to achieve the laser emission. Above the threshold, the laser output power strongly increases as an increase of the injected current, as schematically illustrated in **Figure 18**. Since, on the other hand, the lasing wavelength is determined by the cavity length, the emitted light from the LDs is very coherent, very monochromatic, and highly directional, similar to other lasers such as solid-state lasers and gas lasers. The greatest advantage in LDs is their compactness in size. In order to confine the carriers in the active region where a stimulated recombination occurs, double heterostructures are widely used as shown in **Figure 19**. The active region, in which the bandgap is tuned to be a desirable photon energy, is sandwiched by the barrier regions which have wider bandgaps than the active region and are oppositely doped to inject electrons and holes separately. Such active and barrier regions also act as core and clad regions for a waveguide structure, respectively, owing to the difference in refractive index. As a result, the traveling light wave and the injected carriers are effectively confined in the active region, which contributes to the high emission efficiency in the LDs.

When electrons are confined in very fine quantum well (QW) heterostructures, a 2DEG is formed similar to the HEMT device. As illustrated in **Figure 20**, the electron density of states in the 2DEG system looks like a staircase, so the electron distribution in the 2DEG is more concentrated near the band edge than that in the three-dimensional free electron

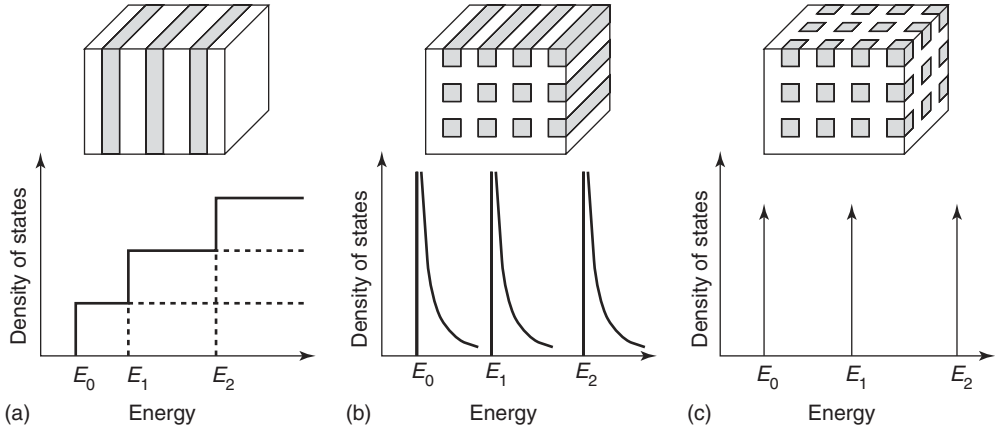


Figure 20 Schematics of the electron density of states in (a) two-dimensional, (b) one-dimensional, and (c) zero-dimensional electron gas systems in the QW, QWI, and the QD, respectively.

system, which well raises the current–light conversion efficiency in the QW. As a result, many improvements in the laser characteristics, such as low threshold current, high modulation frequency, and narrow spectrum line width, are realized. In addition, by adopting lower-dimensional electron systems in quantum wires (QWI) and quantum dots (QD), more improvements are expected owing to their sharper density of states, as illustrated in **Figure 20**.

Wide Bandgap Materials

For semiconductor devices such as Si, Ge, GaAs, InP and their compounds have widely been used since the invention of the bipolar transistor. Their bandgaps are 1–2 eV, resulting in most of the applications for optic devices being limited in the yellow, red, and infrared regions. However, according to the recent progress in crystal growth, it has been possible to fabricate the $A1_xGa_{1-x}N$ materials that have a wide bandgap, over 3 eV. By using them, green or blue light emission is realized. In addition, wide-gap materials have high voltage tolerance compared with the narrow-gap materials such that high electric fields can be applied to accelerate electrons, which brings about a possibility for very high-speed operations. Therefore, $A1_xGa_{1-x}N$ materials are very hopeful for electronic devices as well.

See also: Electrons and Holes; Elemental Semiconductors, Electronic States of; Film Growth and Epitaxy: Methods; Integrated Circuits; Light Emitting Diodes; Memory Devices, Volatile; Nanostructures, Electronic Structure of; Nanostructures, Optical Properties of; Quantum Devices of Reduced Dimensionality; Semiconductor Lasers; Semiconductors, General Properties; Semiconductors, Optical and Infrared Properties of; Transistors.

PACS: 42.55.Px; 73.40.–c; 73.61.–r; 78.40.Fy; 78.66.–w; 79.60.Bm; 85.30.–z

Further Reading

Sze SM (1981) *Physics of Semiconductor Devices*. 2nd edn. Wiley.
Yariv A (1997) *Optical Electronics*, 5th edn. Oxford: Oxford University Press.

Semiconductor Heterojunctions, Electronic Properties of

M Peressi, University of Trieste, Trieste, Italy

Introduction

Semiconductor heterojunctions are built by joining, at a planar face, two semiconductors formed by different chemical elements but usually with similar lattice structure and chemical binding, as schematically indicated in **Figure 1**. From the simple heterojunction constituted by two thick slabs of different materials, which presents only one interface, different heterostructures can be formed – from the double heterojunction showing confinement properties for the charge carriers to the repeated heterojunctions (multiple quantum wells and superlattices) and new artificial structures based on sequences of alternating different semiconductors of varying thickness (10–1000 Å).

Semiconductor heterojunctions show peculiar electronic properties giving rise to physical phenomena of a quantum mechanical nature, some of them not possible even in any "natural" bulk material. These peculiar electronic features are determined by the spatial profile of the electronic bands, which is in turn controlled both by intrinsic characteristic of constituent semiconductors, such as lattice constants, energy gaps, doping, and extrinsic features such as chemical and structural details in the interface

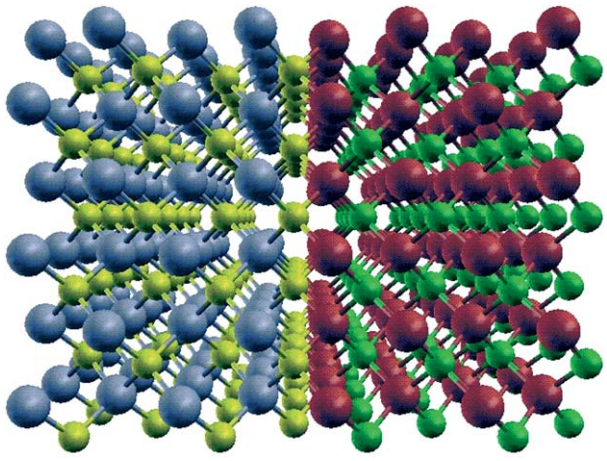


Figure 1 Stick-and-ball model of an abrupt semiconductor heterojunction between two zinc blende semiconductors with no-common ions. The interface is (0 0 1) oriented. (Adapted from Sorba L, *et al.* (1992) Structure and local dipole of Si interface layers in AlAs–GaAs heterostructures. *Physical Review B* 46: 6834.)

region, where the bands exhibit discontinuities (see **Figure 2**). The spatial profile of the electronic bands can be used to properly tailor the space-charge distribution and the behavior of the charge carriers both in the growth direction (transport properties) and in the potential wells parallel to the interface planes (confinement properties), giving the semiconductor heterojunctions a crucial role in modern electronic and optoelectronic devices.

The introduction of heterojunctions dates back to almost 50 years ago. Herbert Kroemer and Zhores I Alferov obtained the Nobel Prize in Physics in 2000 "for developing semiconductor heterostructures used in high-speed- and opto-electronics." In 1957, Herbert Kroemer published the first proposal for a heterostructure transistor. His theoretical work showed that heterostructure devices could offer superior performance compared to conventional transistors. In 1963, H Kroemer and Zhores I Alferov independently proposed ideas to build semiconductor lasers from heterostructure devices, using a double heterostructure to confine carriers. Alferov built the first semiconductor laser from gallium arsenide and aluminum arsenide in 1969.

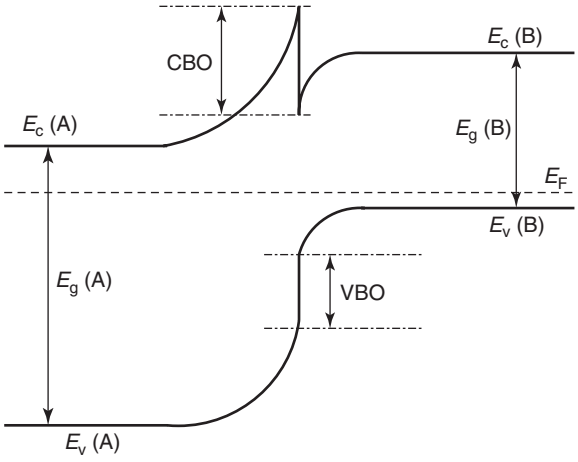


Figure 2 Schematic spatial profile of the valence and conduction bands along the growth direction for a semiconductor heterojunction, with band bending due to space-charge effects and definition of band offsets, VBO and CBO.

The traditional semiconductor heterojunctions involve elements of the central portion of the periodic table, starting from Si and GaAs. Si- and GaAs-based technology is the most mature, but many other III–V and II–VI binary compounds as well as alloys are used nowadays, the choice depending on the particular application. Among alloys, perhaps the most widely studied and used is $Al_xGa_{1-x}As$, in particular, in its interface with GaAs.

High-quality semiconductor interfaces are easily built from two semiconductors which are isovalent and lattice-matched, but more generally they can be also heterovalent (involving semiconductors of different groups) and lattice-mismatched, with slightly different lattice parameters. With the addition of nitrides of II(IV)–VI compounds and of magnetic semiconductors to the most conventional ones, the range of the lattice constants and of the energy gaps which is covered by the different constituents used to form heterostructures is very large, as indicated by **Figure 3**,

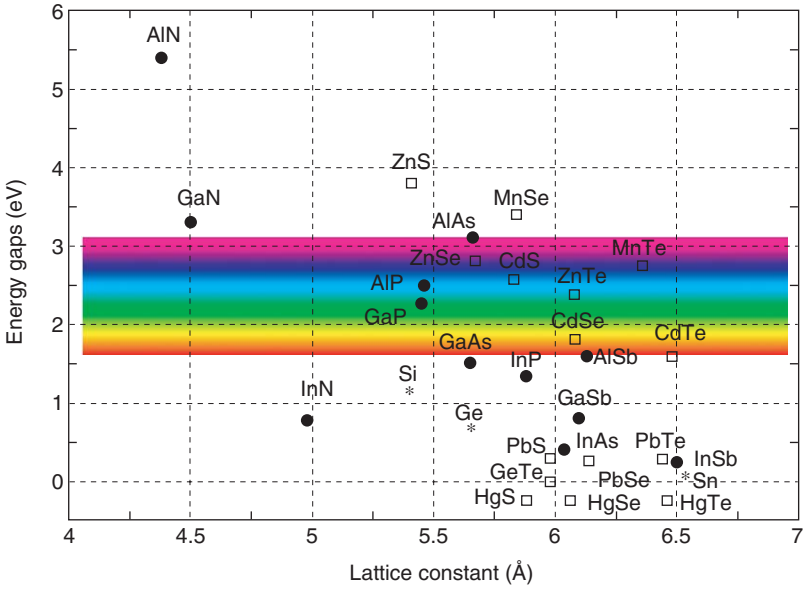


Figure 3 Energy gaps vs. lattice constants for various IV (stars), III–V (closed circles), II(IV)–VI (open squares) zinc blende semiconductors. The visible electromagnetic spectrum is also indicated.

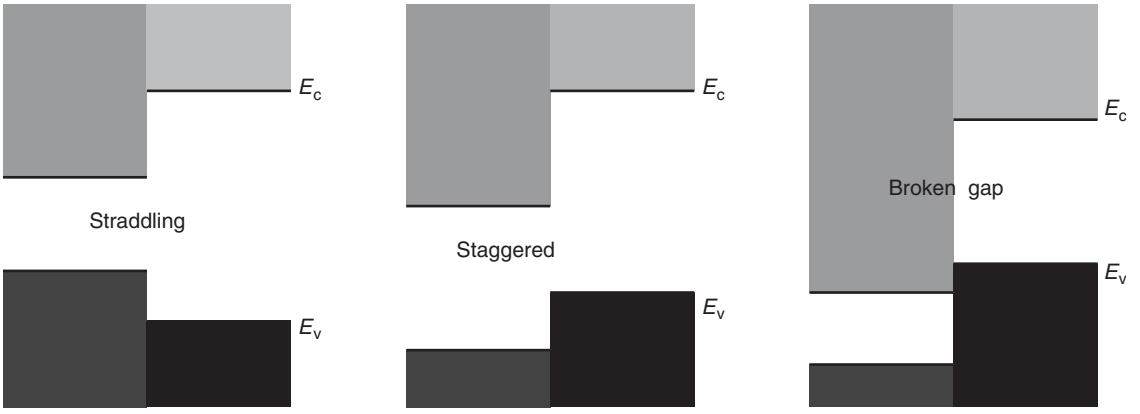


Figure 4 Schematic possible types of band alignments: (a) straddling, (b) staggered, and (c) broken gap. Flat bands are represented, as the focus is on a region which is of the order of 10 atomic units, and the band bending is negligible at this scale.

suggesting possibilities for new fundamental physics and device applications.

Thanks to the improvement of experimental techniques such as molecular beam epitaxy (MBE) and metal–organic chemical–vapor deposition (MOCVD), it is possible to grow high-quality pseudomorphic heterostructures without misfit dislocations or other defects, also with semiconductors with a rather small lattice mismatch (i.e., less than a few percent), and with different but almost commensurate structures.

A heterojunction indicated with A/B typically corresponds to a particular growth order: B is the substrate and A the epilayer. Another specific feature that is usually indicated is the crystallographic growth axis, which determines the interface plane.

Electronic Properties: Band Alignments

The spatial profile of the electronic bands is mostly determined by the band discontinuities at the interface. The bandgap difference ΔE_g, normally existing between the constituent materials, is shared among valence and conduction bands, thus giving rise to the valence and conduction band offsets (VBO and CBO):

$$\Delta E_g = \mathrm{VBO} + \mathrm{CBO} \qquad [1]$$

There are different types of band lineups which are possible, schematically indicated in **Figure 4**. The straddling or type-I lineup characterizes, for instance, the GaAs/$Ga_xAl_{1-x}As$ interface. Modern optoelectronic devices – including quantum well lasers – are based on such a lineup, typically with a double or repeated heterojunction, where the charge carriers are both confined in the slab of material with the lowest energy gap (see **Figure 5**). Other types of lineups are the staggered or type-II and the broken-gap or type-II-misaligned lineups, which could be preferred for some specific applications.

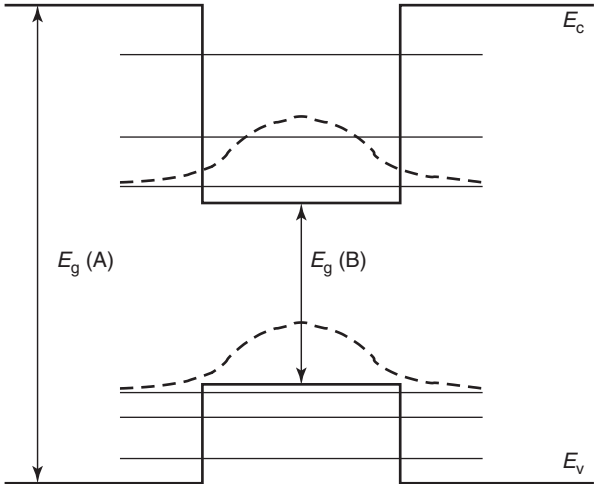


Figure 5 Confinement of electrons and holes in the slab of a material with the lower gap in a double semiconductor heterostructure with straddling-type band alignment.

Extensive theoretical and experimental work has targeted the problem of band alignments since a long time. In particular, an important and general issue widely debated is whether the band discontinuities are essentially determined by the bulk properties of the constituents, or if some interface-specific phenomena, such as crystallographic orientation and abruptness, may affect them in a significant way. A large joint experimental and theoretical/computational effort has been performed to finally clarify, quite recently, the physical mechanisms which give rise to the band alignment. Several books and review articles are available on the subject, as indicated in the "Further reading" section. Some of them are focused on the problem of band offset engineering, that is, on the possibility of controlling the interfaces thus

providing a way to manipulate the band lineups and tune the transport properties across the junctions.

Measuring Band Offsets

Several experimental techniques are available for measuring energy band discontinuities, including optical and transport measurements. Since the late 1970s, spectroscopic techniques have emerged as a fundamental tool for investigating heterojunctions on a microscopic scale. In photoemission spectroscopy, for instance, a direct measurement of the band discontinuities can be done by comparing specific core-level binding energies E_{cl} with respect to valence band top edges E_v in bulk samples with those from the heterostructure:

$$\begin{aligned}\mathrm{VBO} = &[E_{cl}(\mathrm{Ga}\ 3d) - E_v(\mathrm{GaAs})] \\ &- [E_{cl}(\mathrm{Al}\ 2p) - E_v(\mathrm{AlAs})] + \Delta E_{cl} \qquad [2]\end{aligned}$$

An example is schematically shown in **Figure 6** for the AlAs/GaAs(0 0 1) interface.

Predicting Band Offsets

Theoretical and numerical investigations based on different approaches have contributed toward predicting band offset values and identifying the basic mechanism responsible for band offsets. These investigations can be divided mainly into two classes: (1) models and approaches making simplifying and sometimes drastic approximations in describing the interface, but retaining important physical concepts, and (2) accurate computational predictions based on fully self-consistent *ab initio* approaches, which allow one to fully take into account the interface details, such as orientation, abruptness, and defects.

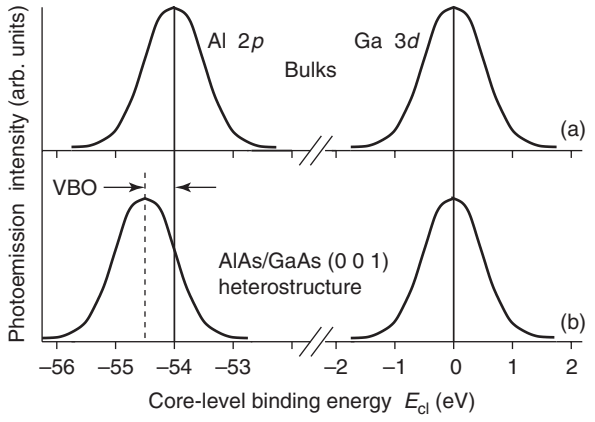


Figure 6 Example of a photoemission measurement of the band offset at the semiconductor heterojunction AlAs/GaAs(0 0 1). Core-level binding energy distribution curves for Al 2*p* and Ga 3*d* states referred to the corresponding AlAs and GaAs valence band top edges are schematically drawn. The zero of the energy scale is arbitrarily taken at the center of the Ga 3*d* emission peak. The variation of the separation ΔE_{cl} of the core level peaks from the case of the two bulks, considering the topmost valence bands of the two materials aligned (a) and the heterojunction (b) is a direct measurement of the VBO.

In general, no universal energy scale exists on which the band structures of the different semiconductors can be easily and uniquely referred, but simple and old models based on intrinsic reference levels have been widely used and have given successful predictions for a number of heterostructures: from the oldest models by Anderson, Frensley, and Kroemer, based on the difference between the electron affinities of the two semiconductors and the concept of charge neutrality levels, to Tersoff's model emphasizing the role of interface dipoles. More refined quantum mechanical approaches, based on the envelope function concept or using the tight-binding method have also been widely and successfully applied. They are reviewed in the books cited above.

The second class of theoretical/computational investigations, more recent, includes *ab initio* approaches, which are based on the solution of quantum mechanical equations for the system under consideration without any use of empirical parameters. This seemingly unaffordable task has been made feasible, thanks to the density-functional theory (DFT) proposed by Walter Kohn, who received the Nobel Prize in Chemistry in 1998 for the same, and for its approximations for practical applications, such as the local-density approximation. DFT has proven to yield reliable results, at an acceptable computational cost, on the electronic ground-state properties of complex crystalline systems, allowing a meaningful comparison with experiment or even accurate predictions of quantities not yet accessible experimentally, with an extremely useful predictive power.

In *ab initio* approaches, semiconductor heterostructures are normally modeled using periodically repeated supercells with a limited number of atoms. The relevant effects due to the presence of the interface are confined in a small region, and the bulk features of the charge density distribution and electrostatic potential are completely recovered within a few atomic units from the interface. This justifies the concept of abrupt discontinuities of the electronic bands across the interfaces which is used in the schematic diagrams of electronic band profiles.

The supercell self-consistent calculations provide the electronic charge density distribution and the corresponding electrostatic potential, as shown in **Figure 7**. Filtering out the microscopic oscillations

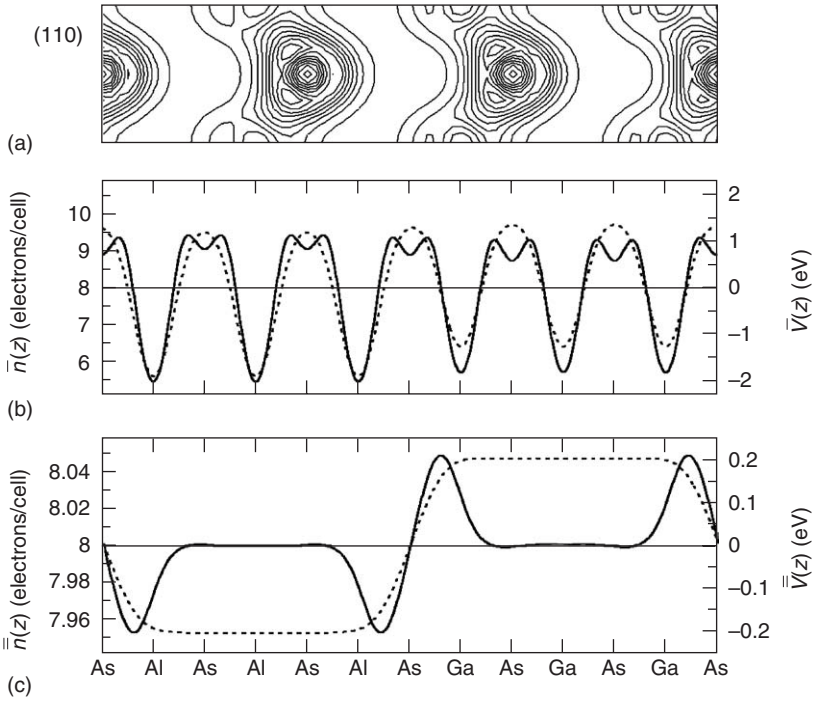


Figure 7 Contour plot of the valence electron density distribution (a) for GaAs/AlAs(0 0 1) heterojunction over a (1 1 0) plane containing the growth axis and centered on the interface anion. Profiles of the electron density and of the electrostatic potential along the growth direction averaged over planes parallel to the interface ($\bar{n}(z)$and $\bar{V}(z)$) (b), and further averaged along the growth direction with a filter equal to the periodicity of bulks ($\bar{\bar{n}}(z)$ and $\bar{\bar{V}}(z)$) (c).

with the bulk-like periodicity, one can unambiguously define an "interface dipole" and extract a potential difference across the interface, ΔV, which, in principle, depends on the interface structural and chemical details. The final band offsets VBO and CBO are determined by the difference ΔE_v and ΔE_c between the relevant band edges in the two materials, measured with respect to the average electrostatic potential in the corresponding bulk crystals:

$$\mathrm{VBO} = \Delta V + \Delta E_v \qquad [3]$$

$$\mathrm{CBO} = \Delta V + \Delta E_c \qquad [4]$$

When the two semiconductors constituting the heterostructure are rather similar (in structure and chemical composition), their differences can be quite minute with respect to the typical bulk variations, and a low-order perturbation approach is appropriate.

The state-of-the-art theoretical approaches allow one not only to compute band discontinuities accurately but also to obtain an insight into the atomic-scale mechanisms which determine the band lineups, and interpret and predict their trends. However, the predictive capability of theoretical schemes is actually limited by the difficulty in predicting the actual atomic-scale arrangement at the interfaces, and their kinetic versus thermodynamic character.

Band Offset Trends

Some peculiar trends in band offsets at semiconductor heterojunctions have been clarified, thanks to the large joint experimental and computational effort: at lattice-matched isovalent heterojunctions, the band offsets depend only on the bulk properties of the two materials, whereas at heterovalent heterojunctions they crucially depend on the interface orientation and other microscopic details. These results are briefly reviewed and justified in the following section.

Lattice-Matched Semiconductor Interfaces

Isovalent semiconductor heterojunctions The GaAs/AlAs heterojunction is the simplest prototype of lattice-matched common-anion heterojunctions. The different cations are isovalent; therefore, the induced potential drop across the interface is only due to the electronic charge. Measurements and accurate numerical investigations show that the VBO and CBO are independent of the interface orientation and abruptness. These results can also be generalized to isovalent no-common-ion heterojunctions such as InAs/GaSb and InP/$Ga_{0.47}In_{0.53}As$, where the lattice-matching conditions result from a balance between the differences of the cationic and anionic core radii, and consequently an important microscopic and configuration-dependent interfacial strain may establish.

Consequences of the bulk-like character of the VBO and CBO are the commutativity and transitivity relationships, which are valid, at least approximately, in the whole class of isovalent semiconductor heterojunctions:

$$\mathrm{VBO}(A/B) = -\mathrm{VBO}(B/A) \qquad [5]$$

$$\mathrm{VBO}(A/B) = \mathrm{VBO}(A/C) + \mathrm{VBO}(C/B) \qquad [6]$$

Heterovalent semiconductor heterojunctions For the sake of clarity, the reader is referred to the case of Ge/GaAs as the simplest prototype of heterovalent semiconductor heterojunctions, but similar considerations apply to other heterojunctions characterized by a valence mismatch between the constituent atoms, such as those among ZnSe, Ge, and GaAs.

In the (1 1 0) direction, each atomic plane is characterized by the same average ionic charge, so that an abrupt junction does not carry ionic charge contribution and the lineup is due only to the electrons. At variance, ideally abrupt interfaces along a polar direction such as (0 0 1) would be charged and hence thermodynamically unstable, as already emphasized first by Harrison in 1978. The simplest neutral and stable interfaces one can envision are terminated by one mixed plane of anions or cations, $As_{0.5}Ge_{0.5}$ or $Ga_{0.5}Ge_{0.5}$ (see **Figure 8**). These two interfaces are stoichiometrically inequivalent, and, because of ionic point charges contribution due to the different valence of the atoms involved (Ge vs. Ga, Ge vs. As), they also correspond to different band offsets. Elementary electrostatics together with a linear response theory predict that the screened ionic point-charge contribution to the offsets is equal in magnitude and opposite in sign for the two interfaces: $\Delta V_{\mathrm{ionic}} = \pm \pi e^2 / 2 a_0 \langle \varepsilon \rangle$, where a_0 is the lattice parameter involved and $\langle \varepsilon \rangle$ a proper average of the dielectric constants of the constituents. The maximum predicted possible variation with interface composition of the band offset at these interfaces is thus $\pi e^2 / a_0 \langle \varepsilon \rangle \approx 0.8$ eV for Ge/GaAs. For ZnSe/Ge, where the valence difference between the constituting elements is twice, the maximum possible variation would be $2\pi e^2 / a_0 \langle \varepsilon \rangle \approx 1.3$ eV.

In general, the atomic interdiffusion which may also occur across the interface over several atomic planes, depending on the growth conditions, reduces the point-charge contribution to the offset, and the observed variations of the offsets are smaller. However, for heterovalent heterostructures deviations from the commutativity and transitivity rule definitely beyond the experimental resolution, up to ≈ 0.5 eV, are observed, and these are a clear fingerprint of the formation of inequivalent interfaces.

Lattice-Mismatched Semiconductor Interfaces

In lattice-mismatched heterojunctions, the band offsets are affected by the stress state of the two materials, depending on the substrate. In pseudomorphically grown heterostructures, the epilayer accommodates its mismatch with the substrate with an almost homogeneous strain and a lattice constant $a_\perp$ along the growth direction, which essentially depends on its elastic properties and on the mismatch with the substrate. The strain state of the epilayer affects the band edges with a combined effect of shift and split. The former affects the averages of the band edge manifolds, calculated with respect to the reference electrostatic potential, and depends on the hydrostatic component of the strain, that is, on the relative volume change with respect to the cubic unstrained material. Splitting effects depend on the orientation and sign of the strain (tensile or compressive). The total variation of the VBO with strain also includes the variation of the electrostatic potential lineup ΔV.

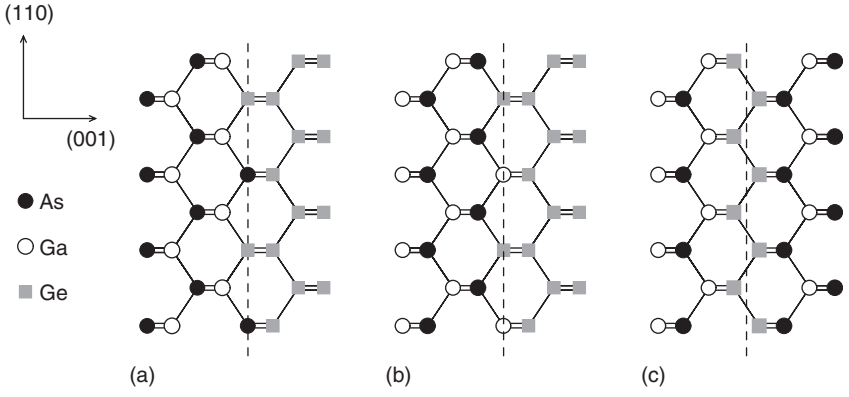


Figure 8 Atomic configurations for the two simplest neutral Ge/GaAs interfaces (a), (b), and for a Ge bilayer embedded in GaAs (c). (Adapted from Sorba L *et al.* (1992) *Physical Review B* 46: 6834.)

The simplest case of lattice-mismatched heterojunctions is Si/Ge, which is an example of isovalent homopolar interfaces. The predicted variation of the VBO at this heterojunction is ~ 0.5 eV, changing the substrate from Si to Ge. Larger effects could be, in principle, possible at heterovalent lattice-mismatched heterojunctions, such as Si/GaAs, or in presence of larger lattice mismatches. However, the difficulty of growing high-quality lattice-mismatched heterojunctions drastically limits their formation and their use in practice.

Band Offset Engineering

In bulk semiconductor technology, one of the most important and widely used features has been the possibility of intentionally varying the electronic properties by alloying and doping. For semiconductor heterojunctions, a similarly important challenge is to control and artificially modify the band offsets.

As already mentioned, the VBO and CBO at lattice-matched isovalent semiconductor heterojunctions are mostly determined by the bulk properties of the constituents, and interface details such as orientation and stoichiometry play a very minor role, whereas at heterovalent heterojunctions they are very sensitive to these extrinsic features. As a consequence, heterovalent heterojunctions seem to be ideal candidates as tunable heterojunctions, while using only isovalent materials, it seems that the only way to tune the offset is to act on their bulk properties with strain or with alloying.

Actually, the peculiarity of heterovalent interfaces leads naturally to a practical way for modifying the offset also at isovalent heterojunctions, or even for creating an offset at a homojunction. For instance, joining together the two inequivalent interfaces Ge/GaAs(0 0 1) (Ge–As and Ge–Ga terminated), one has the following sequence of atomic planes: $\cdots$As–Ga–As–$Ge_{0.5}Ga_{0.5}$–$Ge_{0.5}As_{0.5}$–Ga–As–Ga$\cdots$. Therefore, using the difference between the VBOs of these two inequivalent interfaces, a net potential drop $\Delta V = \pi e^2/a_0 \langle \varepsilon \rangle$ is predicted across the mixed bilayer. This potential drop is the same that would result from a microscopic capacitor whose plates are placed at a distance $a_0/4$, and carry a surface charge $\sigma = e/a_0^2$. The above sequence of atomic planes can also be thought of as due to the transfer of a proton per atomic pair from the As to the Ga planes.

This viewpoint is easily generalized to arbitrary concentrations of Ge in a pair of consecutive compensated GaAs planes, Ge_xGa_{1-x} and Ge_xAs_{1-x}, to ensure local charge neutrality. The maximum potential drop would be, in principle, $\Delta V_{\max} = 2\pi e^2/a_0 \langle \varepsilon \rangle$ for the case $x = 1$, that is, for an entire Ge bilayer embedded (see **Figure 8**). The above results can be further generalized to the heterojunctions, for example, doping a GaAs/AlAs interface with ultrathin layers of Si or Ge. In this case, the interlayer contribution can reduce or increase the intrinsic VBO according to the growth sequence. These predictions have been indeed experimentally confirmed for Si interlayers at GaAs/AlAs and AlAs/GaAs (0 0 1) junctions and analogous effects have also been observed in other systems.

Localized Interface States

Another important issue in the heterojunction electronic structure concerns the presence of localized interface states in heterojunctions, their origin, and their role in determining the electronic characteristics of the junctions. The existence of localized interface states is generally claimed whenever anomalous features are observed in photoemission, optical, or transport measurements. Their unambiguous identification, however, is a difficult task in experiments. Reflectance anisotropy spectroscopy has been recently applied to a few interfaces and seems to be a promising tool for more direct investigations.

The electronic states of a heterojunction can be extended everywhere, localized in one material or the other (acting as a quantum well), or localized at the interface. The true interface states are not degenerate with bulk Bloch states, decay exponentially away from the interface, and exist only in the mutual energy gaps. The comparison of the band structures of the two bulk materials with the spectrum of the heterojunction allows one to identify those states which do not belong to the bulk band structures, and are, therefore, candidates to be interface states. Localized electronic states can be successfully identified in numerical investigations also by the local density of states: far from the interface on the two sides of the junction, it yields the bulk densities of states of the two materials, whereas its deviations with respect to the constituting bulk state densities in the interface region would indicate the presence of states which are localized there.

Interface states can have different physical origins. In some cases, they are associated with defects which typically arise at junctions between two morphologically different phases (e.g., cubic and wurtzite). But localized electronic states can also occur, in principle, at semiconductor heterojunctions with constituents having the same or similar structure but different chemical properties. Simple electron-counting schemes indicate that the valence mismatch between neighboring atoms from opposite sides of the junction could result in localized interface states

corresponding to donor or acceptor bonds. Therefore, heterovalent heterojunctions, which exhibit such kind of bonds in the interface region, are natural candidates for the existence of localized interface states. Localized states have been experimentally detected at diamond/zinc blende interfaces, for example, Ge/ZnSe, at zinc blende/zinc blende heterovalent interfaces, for example GaAs/ZnSe, and even at isovalent no-common-ion interfaces, for example, BeTe/ZnSe.

Accurate calculations predict that interface electronic states critically depend not only on the type of heterojunction (i.e., heterovalent rather than isovalent), but on the particular atomic-scale morphology of the interface.

Semiconductor Heterostructures and Spintronics

The word "spintronics" denotes electronic-like heterostructures where the relevant physical quantity is the spin of the carriers and its interactions with external magnetic fields, rather than the charge of holes and electrons and the associated electronic properties.

Semiconductor heterostructures using carrier spin as a new degree of freedom offer new functionality with respect to conventional junctions. A large effort is currently devoted to integrate traditional III–V and II–VI semiconductors with ferromagnetic semiconductors or, more generally, with other ferromagnetic materials including metals and half-metals. There are, however, several advantages using all-semiconducting devices: interface properties, such as lattice matching and band offsets, are more understood and controllable, and the integration with the existing conventional semiconductor technology is easier. The diluted ferromagnetic semiconductors are semiconductor compounds where a fraction of the constituent ions is replaced by magnetic ions. This is the case, for instance, of (Ga,Mn)As, which has received particular attention: Mn is a heterovalent substitutional impurity for Ga in GaAs (valence II with respect to the valence III of Ga), and therefore, it can be used as a source of spin-polarized holes. An Mn content up to 5% is sufficient to make (Ga,Mn)As ferromagnetic.

Furthermore, there is the possibility of a good integration of half-metals with conventional semiconductors: these are materials that have only one occupied band at the Fermi level, and behave as a metal in one spin channel and as a semiconductor in the other. Current effort, both experimental and theoretical, is nowadays devoted to predict those particular heterostructures where half-metallicity is also maintained in the presence of the interface.

See also: Core Photoemission; Density-Functional Theory; Electronic Structure Calculations: Plane-Wave Methods; Epitaxy; Film Growth and Epitaxy: Methods; Nanostructures, Electronic Structure of; Pseudopotential Method; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Devices; Semiconductor Lasers; Semiconductors, General Properties; Semi-Empirical and Empirical Calculations; Surfaces and Interfaces, Electronic Structure of; Tight-Binding Method in Electronic Structure; Transmission Electron Microscopy; Tunneling Devices.

PACS: 71.15. – m; 73.20. – r; 73.20.At; 73.40.Kp; 73.40.Lq; 75.50.Pp; 78

Further Reading

Bastard G (1988) *Wave Mechanics Applied to Semiconductor Heterostructures*. Les Ulis, Cedex: Les Editions de Physique.

Brillson LJ (1992) Surfaces and interfaces: atomic-scale structure, band bending and band offsets. In: Landsberg PT (ed.) *Handbook on Semiconductors*, vol. 1, ch. 7, pp. 281–417. Amsterdam: North-Holland.

Capasso F (1987) Band-gap engineering: from physics and materials to new semiconductor devices. *Science* 235: 172–176.

Capasso F and Margaritondo G (eds.) (1987) *Heterojunction Band Discontinuities: Physics and Device Application*. Amsterdam: North-Holland.

Franciosi A and van de Walle CG (1996) Heterojunction band offset engineering. *Surface Science Reports* 25(1–4): 1–140.

Grimmeiss HG (ed.) (1996) *Heterostructures in Semiconductors*. Proceedings of the Nobel symposia in physics, Physica Scripta, vol T68, Stockholm.

Harrison WA, Kraut EA, Waldrop JR, and Grant RW (1978) Polar heterojunction interfaces. *Physical Review B* 18: 4402–4410.

http://www.nobel.se/physics/laureates/1973/esaki-lecture.html

http://www.nobel.se/physics/laureates/2000/alferov-lecture.html

http://www.nobel.se/physics/laureates/2000/kroemer-lecture.html

Margaritondo G (ed.) (1988) *Electronic Structure of Semiconductor Heterojunctions*. Dordrecht: Kluwer.

Ohno Y, Young DK, Beschoten B, Matsukura F, Ohno H, and Awschalon DD (1999) Electrical spin injection in a ferromagnetic semiconductor heterostructure. *Nature* 402(6763): 790–792.

Peressi M, Binggeli N, and Baldereschi A (1998) Band engineering at interfaces: theory and numerical experiments. *Journal of Physics D: Applied Physics* 31: 1273–1299.

Yu ET, McCaldin JO, and McGill TC (1992) Band offsets in semiconductor heterojunctions. In: Ehrenreich H and Turnbull D (eds.) *Solid State Physics*, vol. 46, pp. 1–146. Boston: Academic Press.

Vurgaftman I, Meyer JR, and Ram-Mohan LR (2001) Band parameters for III–V compound semiconductors and their alloys. *Journal of Applied Physics* 89(11): 5815–5875.

Wolf SA, Awschalom DD, Buhrman RA, Daughton JM, von Molnar S, *et al.* (2001) Spintronics: a spin-based electronics vision for the future. *Science* 294(5546): 1488–1495.

Nomenclature

a_0	lattice parameter
e	electron charge

E_c	conduction band	ΔE_g	fundamental bandgap difference
E_g	fundamental bandgap	ΔE_v	valence band difference (with respect to reference levels in the bulks)
E_v	valence band		
n	electron density	ΔV	electrostatic potential lineup
ΔE_c	conduction band difference (with respect to reference levels in the bulks)	ε	dielectric constant

Semiconductor Lasers

M-C Amann, Technische Universität Munchen, Garching, Germany

Introduction

Even though more than four decades have passed since the first demonstration of lasing in semiconductors, the pace of development of new and further improved semiconductor lasers continues to be rather high. Judging by their economic impact, semiconductor lasers can be considered today as the most important type of laser. From the beginning, the development of semiconductor lasers has been closely connected with the progress of semiconductor technology and device physics. Hence, the invention of semiconductor heterostructures as well as the straightforward implementation of quantum effects together with improved epitaxial techniques have had a decisive impact on the evolution of these lasers. Besides the physical background and the various device concepts, a proper treatment of semiconductor lasers has to particularly include a description of the relevant materials and their fabrication technology.

After a brief review of the conditions for laser action in semiconductors, this article gives an overview of the most relevant semiconductor materials and material systems. Besides the materials, it is also shown that the various choices among the various laser structures makes this laser class well suited for a vast number of applications. The stationary, dynamic, and spectral characteristics of state-of-the-art semiconductor lasers are presented in the next section and a brief discussion on single-frequency and vertical-cavity surface-emitting semiconductor lasers completes this article.

Population Inversion and Gain in Semiconductors

A fundamental concept in laser physics is to achieve population inversion between two discrete states in a nonequilibrium physical system, so that the stimulated emission can overcome absorption and losses yielding a net optical gain. In a semiconductor, however, the energy states for electrons and holes are continuously distributed within energy bands that are separated by forbidden bands as schematically shown in **Figure 1** for a direct semiconductor, where the width of the forbidden band is denoted as bandgap energy E_g. The population of these bands is governed by the Fermi distribution characterized by the Fermi level at which the population probability equals 0.5. A population inversion can be achieved by heavily pumping the semiconductor via a current injection. As shown in **Figure 1**, the nonequilibrium electron and hole distributions can be described by two equilibrium Fermi distributions with two Fermi levels E_{Fc} and E_{Fv} for the conduction and valence bands, respectively. This is possible due to the fast relaxation of the carriers within the bands,

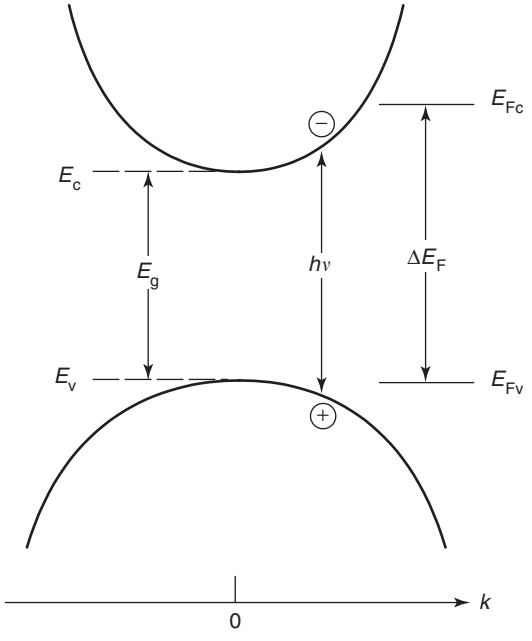


Figure 1 Schematic diagram of energy bands and quasi-Fermi levels of an inverted semiconductor illustrating the Bernard and Duraffourg condition.

hence, E_{Fc} and E_{Fv} are named quasi-Fermi levels. In 1961, Bernard and Duraffourg derived the condition for population inversion in semiconductors essentially stating that the energy difference $\Delta E_F = E_{Fc} - E_{Fv}$ needs to exceed the bandgap energy, and the photon energy $\hbar\omega$ is between ΔE_F and E_g:

$$\Delta E_F > \hbar\omega \geqslant E_g$$

The Bernard and Duraffourg condition requires that the population of at least one band is degenerate, that is, the carrier density exceeds the effective density of states N_c and N_v of the conduction and/or the valence band, respectively. This means that the semiconductors are operated in the high-injection regime, since N_c and N_v are rather large. In GaAs, for instance, at room temperature, $N_c \cong 4 \times 10^{17}$ cm^{-3} and $N_v \cong 8 \times 10^{18}$ cm^{-3}. Normally, undoped active regions are applied, so that the electron and hole densities are equal causing the conduction band to become degenerate. While, in principle, population inversion is possible in direct and indirect semiconductors, a sufficient gain to overcome internal absorption losses can hardly be achieved in indirect semiconductors.

For injection levels exceeding the threshold as defined by the Bernard and Duraffourg condition, an optical gain due to stimulated emission by the inverted band occupation occurs. The injection level at which this gain exactly compensates the material losses, such as free-carrier absorption or scattering losses, is defined as the transparency threshold. At even higher injection, the increasing gain may also compensate the mirror losses of an optical cavity enabling laser operation. It should be noted that inverted semiconductors may provide very large optical gain g up to the order of 1000 cm^{-1}, exceeding that of any other optical gain medium.

Semiconductor Heterostructures

Laser operation in homojunction semiconductors is difficult to achieve because of the high current densities required to achieve optical gain. For low-threshold operation, it is therefore necessary to reduce the volume of the active region in which the injection occurs and to exclude optical absorption in the surrounding semiconductor regions. This goal can most effectively be achieved by applying semiconductor heterostructures, in which semiconductors with different bandgap energy are combined. The most effective among them is the so-called "double heterostructure," in which the active region is completely embedded into p- and n-doped semiconductors with a larger bandgap energy as shown schematically in **Figure 2**. The double heterostructure provides two essential features relevant for laser operation:

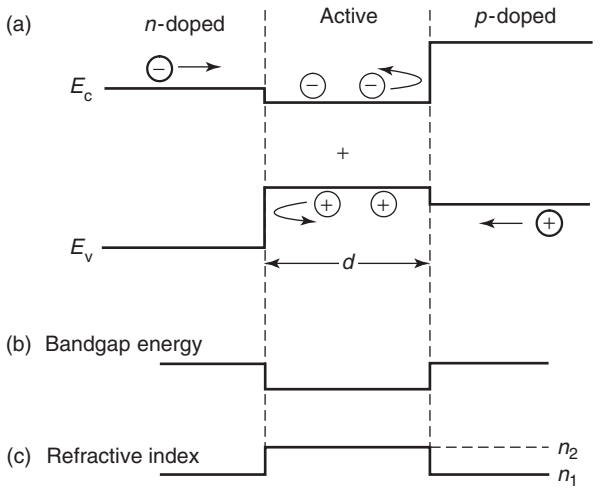


Figure 2 Schematic illustration of (a) semiconductor double heterostructure, (b) spatial profiles of bandgap energy, and (c) refractive index.

1. *Carrier confinement*. This almost completely concentrates electrons and holes within the active region even for active regions having thickness in the nm-regime.
2. *Optical confinement*. This yields optical waveguiding, establishing a slab waveguide with the active region as core because of the smaller refractive indices of the higher bandgap energy semiconductors sandwiching the active region.

Compared to the carrier confinement, the optical confinement as measured by the optical confinement factor

$$\Gamma = \frac{\text{mode power carried in active layer (region)}}{\text{total mode power}}$$

is not as complete for small dimensions, and the optimal radiation-matter interaction in the active region occurs where Γ/d is maximal. Depending on the material constants, accordingly, active region thicknesses are usually $\sim$100 nm.

Further, an improved optical gain performance shows the quantum-well (QW) structures, in which the active region thickness is reduced to below the de Broglie wavelength ($\sim$30 nm). In QWs, the quantization of the electronic states along the normal layer results in a two-dimensional (2D) electron (and hole) gas with a modified density of states that makes the inversion population possible for smaller injection levels. The density of states in the valence band can

additionally be modified by introducing elastic strain that lifts the degeneracy of light and heavy valence bands, so that only one of the two valence bands needs to be inverted yielding still smaller thresholds. Extending the quantization to the other two dimensions, one obtains quantum wire and quantum dot structures, corresponding to one-dimensional (1D) and zero-dimensional (0D) electron (hole) systems with distinct threshold reduction. In case of the quantum dot, atom-like discrete states appear to make this kind of gain medium rather similar to atomic laser media. In general, reducing the dimensionality of the carrier systems results in reduced thresholds and smaller temperature dependence of the optical gain. Because of the significant advantages achieved with double heterostructures, all modern laser diodes normally comprise a QW double heterostructure.

Semiconductor Heterostructure Material Systems

The benefits of the semiconductor heterostructure can be exploited only if the composite crystal exhibits a high crystalline quality. This generally requires that all constituents of the heterostructure have the same lattice constant a_0, while the difference in bandgap energy should be large (several hundred meV). The lattice-matching requirement strongly limits the possible semiconductor combinations because semiconductor heterostructures are almost entirely grown on binary substrates. This is because binary substrates, such as GaAs or InP, are unambigous chemical compounds in contrast to, for example, $In_xGa_{1-x}As$ and typically provide more than an order of magnitude of better thermal conductivity than ternary (or quaternary) compounds because of the nonoccurance of alloy scattering of the phonons.

Lattice-matched double heterostructures, therefore, consist of ternary and quaternary (or even quinternary) compounds grown on binary substrates. The lattice-matching condition can be found from Vegard's law which states that the lattice constant of a compound can be calculated as a linear interpolation between the constituting binaries. For a quaternary compound $A_xB_{1-x}C_yD_{1-y}$, where A and B, and C and D each share the same sublattice, the lattice constant is

$$a_0(A_xB_{1-x}C_yD_{1-y}) = xya_0(AC) + x(1-y)a_0(AD) + (1-x)ya_0(BC) + (1-x) \times (1-y)a_0(BD)$$

On the other hand, the bandgap energy of a compound shows a more complicated and nonlinear dependence on composition. As a consequence, with a proper choice of semiconductor materials, rather large bandgap energy differences can be achieved with lattice-matched compounds making numerous heterostructure combinations possible. Semiconductor lasers are nowadays made from numerous III–V and other compound semiconductors covering the entire visible as well as the infrared regime up to $\sim 100\,\mu m$ wavelength.

Stripe Laser Diodes

A schematic drawing of the basic stripe-geometry semiconductor laser structures is depicted in **Figure 3**. All of them comprise a transverse (vertical, x-axis) double heterostructure to define the vertical extension of the laser-active area with a pn-junction at or in the active layer. A laser current is fed via stripe

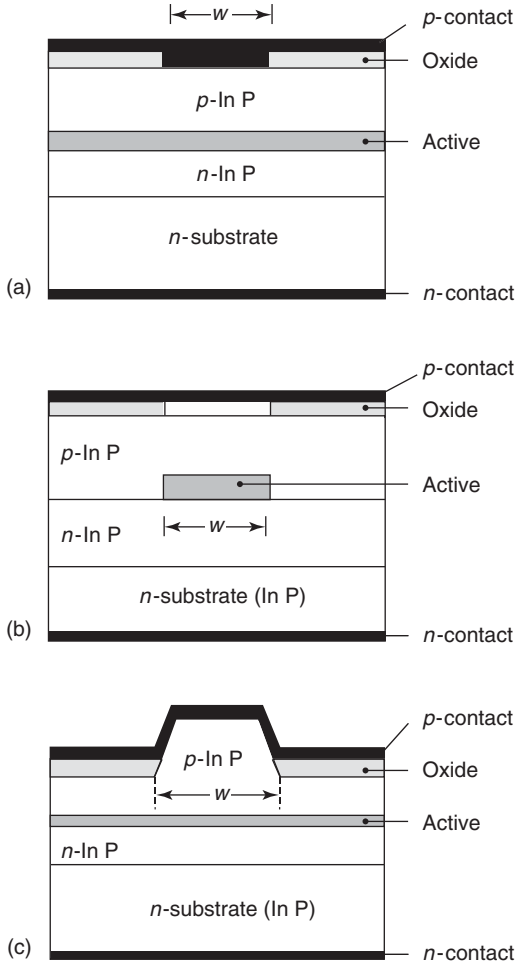


Figure 3 The three different lateral waveguiding types (a) Gain-guiding (GG), (b) index-guiding (IG), and (c) quasi-index-guiding (QIG).

contacts on the top and bottom of the laser chips. Three basic types of lasers can be distinguished: gain-guided (GG), index-guided (IG), and quasi-index-guided (QIG) laser structures, where the guiding relates to the in-plane (lateral) waveguiding along the y-axis. While the transverse waveguiding (x-axis) is accomplished by index guiding via the double heterostructure, the lateral waveguiding may either be done (1) simply by the lateral gain profile in the active layer (GG), (2) similar to the transverse direction by a lateral double heterostructure (IG), or (3) by a partial lateral replacement of part of the structure by a lower index material without attaching the active layer (QIG). While the GG lasers require the lowest technological effort and generally show an excellent long-term stability because of the nonmodified active layer, the IG lasers exhibit the smallest thresholds and optimum laser performance; however, the lateral removal of part of the active layer may worsen the reliability. Often a reasonable compromise are the QIG lasers that, in all aspects, show the behavior in between the other two structures. An optical feedback is provided by the cleaved end-facets of the laser chips, which for typical refractive indices of the semiconductors (2.5–3.5) show a reflectivity $\sim 30\%$ against air. Due to the large gain of the semiconductor laser, this comparatively small mirror reflectivity can be compensated for by laser lengths of the order of several hundred micrometers.

Waveguiding in Laser Diodes, Effective Refractive Index, and Confinement Factor

The transverse double heterostructure represents an optical slab waveguide with the schematic index profile shown in **Figure 2c**. Depending on the active layer (core) thickness d and the index profile, this slab may carry one or more guided transverse electric (TE) and transverse magnetic (TM) modes, respectively. For a sufficiently thin active layer,

$$d < \frac{\lambda}{2\sqrt{n_2^2 - n_1^2}}$$

only the fundamental TE- and TM-modes are guided, which is normally aimed for.

The optical gain which the modes acquire, that is, the mode gain g_m, is the product of the active region material gain g and the optical confinement factor Γ for the particular mode. For a given slab waveguide, the (fundamental) TE-mode exhibits a larger confinement than the (fundamental) TM-mode. Also, the facet reflectivity shows a certain polarization dependence favoring the TE-modes as well. As a consequence, semiconductor lasers normally operate in the fundamental transverse TE-mode. In a good approximation, the confinement factor for the fundamental TE-mode is

$$\Gamma \cong \left[1 + \left(\frac{d_0}{d}\right)^2\right]^{-1} \quad \text{with } d_0 = \frac{\lambda}{2\pi}\sqrt{\frac{2}{n_2^2 - n_1^2}}$$

A further mode property is the effective refractive index n_{eff}, that describes the mode propagation along the waveguide in analogy with a plane wave in a homogeneous medium, with the refractive index equal to n_{eff}. The effective refractive index, which is the ratio of the propagation constant β and the free-space wave number k_0, is between n_1 and n_2.

Considering complex refractive indices $n = n' + jn''$, where the imaginary part is related to the material gain (or loss) g as $n'' = g/2k_0$, it is obvious that the gain in the active region leading to the mode gain produces a slightly complex propagation constant and a complex effective refractive index with imaginary part equal to $n'' = \Gamma g/2k_0$.

Besides transverse profiles of the real part of n, those of the imaginary part of n may also produce optical waveguiding. This explains why the lateral gain profile constitutes a waveguide for laser radiation in the GG lasers. For common gain values, however, the GG mechanism is much smaller than IG, with the consequence that the maximum stripe width w for lateral single-mode operation is much larger in the GG lasers than in their IG counterparts.

Output Power versus Current Characteristic

A representative plot of an optical output power versus current characteristic for a semiconductor laser is shown in **Figure 4**, together with the principal current-dependent carrier density in the active region. Relevant parameters of this characteristic are threshold current I_{th}, output power P, and differential quantum efficiency η_d. The threshold current is related to the active region volume V_a and the spontaneous recombination rate R at lasing threshold carrier density N_{th}

$$I_{\text{th}} = eV_a(N_{\text{th}})$$

where e is the elementary charge. The active region gain $g(N_{\text{th}})$ at injection level N_{th} yields a mode gain that equals the total losses α_{tot} comprising internal losses (e.g., absorption and scattering) α_i and the

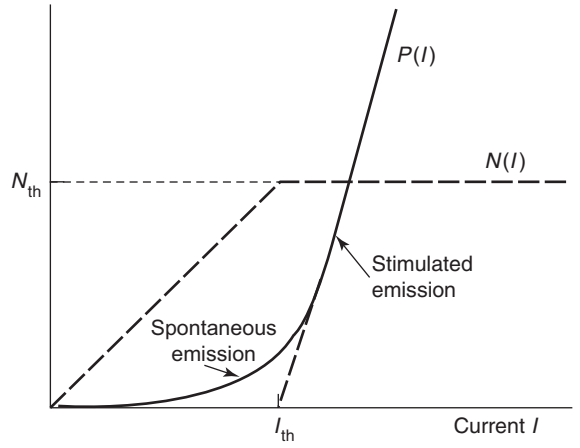


Figure 4 Schematic profile of optical output power P and carrier density in active region N vs. current in a semiconductor laser.

mirror losses α_m:

$$\Gamma g = \alpha_i + \alpha_m \text{ with the mirror losses}$$

$$\alpha_m = \frac{1}{2L} \ln \frac{1}{R_1 R_2}$$

where L is the laser length, and R_1 and R_2 are the front facet mirror reflectivities. For a symmetric laser with $R_1 = R_2 = R$, the optical output power per facet (above threshold) is

$$P(I) = \frac{h\nu}{2e} \frac{\alpha_m}{\alpha_{tot}} (I - I_{th})$$

and the differential quantum efficiency reads

$$\eta_d = \frac{dP/h\nu}{dI/e} = \frac{\ln 1/R}{2\alpha_{tot} L}$$

Emission Spectrum and Beam Properties

With state-of-the-art semiconductor technology, it is possible for most emission wavelengths to fabricate narrow-stripe laser waveguides that carry only one transverse and one lateral mode each. Even though the length of semiconductor cavities (typically 0.1–1 mm) is orders of magnitude smaller than that of most other laser types, the length-(effective)-refractive-index-product is still of the order of several thousand wavelengths. Therefore, the longitudinal mode spacing

$$\Delta\lambda = \frac{\lambda^2}{2n_g L} \quad (n_g : \text{group (effective) index})$$

is very small against the spectral width of the active region gain characteristic, which in terms of energy is of the order of the thermal energy ($k_B T$). As a consequence, the mode gains of the modes near the gain peak are only slightly different, so that the usual Fabry–Perot-type semiconductor lasers generally emit in many longitudinal modes. The purity of the spectrum can be described by the side-mode suppression ratio (SMSR) that measures the ratio of the dominant mode to the next strongest one. Given the mode gain difference Δg_m between the two strongest modes, the SMSR of a symmetric semiconductor laser ($R_1 = R_2$), under stationary conditions, is proportional to the power in the dominant mode P according to

$$\text{SMSR} = \frac{2P\Delta g}{h\nu n_{sp} v_g \alpha_{tot} \alpha_m}$$

where n_{sp} is the spontaneous emission coefficient (typically 2), that is specific for semiconductor lasers taking into account the incomplete inversion of the bands, and v_g is the group velocity. Truly single-mode lasers, that is, single-frequency lasers, show an SMSR > 30 dB, that is, three orders of magnitude, requiring a Δg of 0.01–0.1 cm^{-1}, which generally cannot be accomplished solely by the spectrally dependent active region gain. Single-frequency laser diodes, therefore, need additional wavelength selective elements such as Bragg reflectors as shown below.

An important spectral parameter, particularly for single-frequency semiconductor lasers, is the spectral line width of the mode(s). As with other lasers, the line width is obtained by the Schawlow–Townes formula; however, due to the gain-phase coupling in semiconductor lasers, the Schawlow–Townes formula must be completed yielding the so-called "Schawlow–Townes–Henry line width formula"

$$\Delta\nu = \frac{h\nu v_g^2 \alpha_{tot} \alpha_m n_{sp}}{8\pi P} (1 + \alpha_H^2)$$

where α_H is Henry's line width enhancement factor describing the gain-phase coupling accompanying carrier density changes in the active region which has the form

$$\alpha_H = -2k_0 \frac{\partial n'/\partial N}{\partial g/\partial N}$$

The physical origin for nonzero α_H is the nonsymmetric gain versus wavelength characteristic of semiconductor gain media. While the gain-wavelength characteristics of most other laser materials are symmetric, the asymmetry in semiconductors yields nonvanishing changes of the refractive index at the gain maximum, that is, at the laser wavelength,

during the occurence of gain changes, for instance, during relaxation oscillations or after each spontaneous emission event into the lasing mode(s). It should be noted that α_H is significantly smaller for QW lasers, particularly for strained QW lasers, than for bulk lasers. In case of the recently developed quantum dot lasers, α_H approaches zero because of the almost completely symmetric atom-like gain-wavelength characteristic.

The radiation beam of normal laser-diodes is transversely single-mode and nearly diffraction limited. In this case, the near-field and far-field widths depend on each other as

$$\Delta\Phi = \arctan\left(\frac{2\lambda}{\pi W}\right)$$

where $\Delta\Phi$ and W are the full width at half maximum (FWHM) of the far-field and near-field angles, respectively. Because of the narrow stripe widths of transversely single-mode semiconductor lasers (typically 0.5–3 μm at 1–2 μm emission wavelength), the diffraction-limited far-field width is of the order of several tens of degrees, in contrast to most other laser types.

Dynamic Properties and Noise

The dynamic properties of laser diodes are relevant for their use as a direct current-modulated light source in high-speed data transmission applications. Thereby, the two limiting cases of a linear response at a small-signal modulation and a nonlinear response at a large-signal (generally pulse) modulation are of great interest.

Modulating the laser around average current I and average output power P per facet, respectively, the small-signal modulation response $H(\omega)$ can be described by the transfer function of a damped second-order low-pass filter

$$H(\omega) = \frac{1}{1 - (\omega/\omega_r)^2 + j(\omega/\omega_d)}$$

where the relaxation and damping circular frequencies ω_r and ω_d are given as

$$\omega_r = \frac{1}{\tau_s}\sqrt{\frac{2a\Gamma}{\nu_g\alpha_m\alpha_{tot}h\nu V_a}P} \quad \text{and} \quad \omega_d = \frac{\tau_s}{\gamma}\omega_r^2$$

respectively. Here, the linear gain parameter $a = \partial g/\partial N$ and the photon lifetime in the cavity $\tau_s = (\nu_g\alpha_{tot})^{-1}$ appear, and γ is

$$\gamma = \frac{aS}{\alpha_{tot}} + \frac{\tau_s}{\tau_d} + \frac{n_{sp}\Gamma}{V_a S}$$

with τ_d being the differential carrier lifetime in the active region determined by the total spontaneous emission rate and nonradiative recombination processes (typically several nanoseconds). S is the photon density in the active region corresponding to output power P

$$S = \frac{2\Gamma}{\nu_g\alpha_m h\nu V_a}P$$

With an increase in the laser current or output power, respectively, the relaxation frequency as well as the damping frequency increases. The absolute value of the modulation response versus frequency f is shown in **Figure 5** for a typical 1.55 μm wavelength InGaAsP laser in comparison with an LED made from the same semiconductor material. It is usual to measure $|H(\omega)|$ in decibels, that is, $10\log|H(\omega)|$. The characteristic 3-dB bandwidth is only slightly larger than the relaxation resonance, so that the latter can be used as a reasonable measure for the modulation bandwidth.

The large-signal pulse modulation can be described by rather accurate approximations for the turn-on delay τ_{on}

$$\tau_{on} \cong \tau_d \frac{I_{on} - I_{off}}{I_{on} - I_{th}}$$

for $I_{off} < I_{th} < I_{on}$: off-state below threshold

$$\tau_{on} \cong \frac{\sqrt{2}}{\omega_r}\ln\frac{P_{on}}{P_{off}}$$

for $I_{th} < I_{off} < I_{on}$: off-state above threshold

where switching occurs from $P_{off}(I_{off})$ to $P_{on}(I_{on})$. While for lasers biased below threshold, the τ_{on} delay

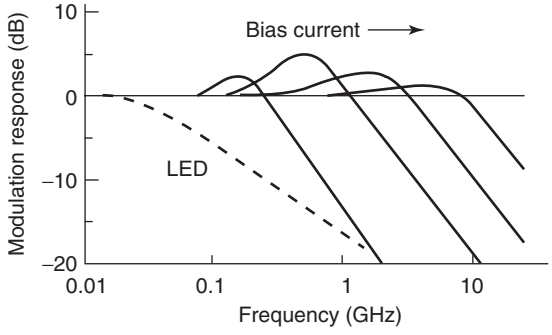


Figure 5 Absolute value $|H(f)|$ of the modulation response of a laser diode.

is of the order of the comparatively long differential carrier lifetime τ_d, τ_{on} delays ~ 0.1 ns can be obtained with an above-threshold bias.

Besides the phase noise that determines the modal line width, an intensity noise occurs in semiconductor lasers. This is usually described in terms of the relative intensity noise (RIN)

$$\mathrm{RIN(in\ dB\ Hz^{-1}}) = \frac{1}{\Delta f} \log\left(\frac{\Delta P}{P}\right)^2$$

where ΔP denotes the power fluctuations of the laser output within bandwidth Δf. The RIN shows a frequency dependence similar to the modulation response with typical low frequency values $\sim 135\,\mathrm{dB\,Hz^{-1}}$.

Single-Frequency Semiconductor Lasers

Semiconductor lasers can be made single-frequency lasers by introducing additional wavelength-selective elements into the laser cavity that select only one of the longitudinal modes. This is commonly accomplished with Bragg gratings monolithically integrated into the laser structure. As shown in **Figure 6**, the gratings may either replace the laser mirrors and are therefore put at one or both laser facets (distributed Bragg reflector (DBR) laser with a longitudinal integration of grating) or may homogeneously be collocated with the active region over the entire laser length (distributed feedback (DFB) laser with a transverse integration of grating). To avoid interference with the end facet reflection, antireflection coatings are normally applied onto those facets of DBR lasers where the gratings provide the feedback. In DFB lasers, therefore, both facets need an antireflection coating to suppress these interferences.

The Bragg gratings are made by periodic thickness variations of a grating layer sandwiched between layers of different refractive indices, so that the thickness variations result in effective refractive index variations Δn_{eff} of the same periodicity. In the DBR laser, the Bragg grating introduces a wavelength-dependent mirror loss with a minimum loss or maximum reflectivity at the (free-space) Bragg wavelength

$$\lambda_B = 2n_{eff}\Lambda$$

where Λ is the grating period, (**Figure 6b**). In the interesting regime around the Bragg wavelength, the wavelength-dependent reflectivity can be calculated by the coupled-mode theory yielding

$$R(\lambda) = \left|\frac{\kappa \sinh(sL_B)}{j\Delta\beta \sinh(sL_B) + s\cosh(sL_B)}\right|^2$$

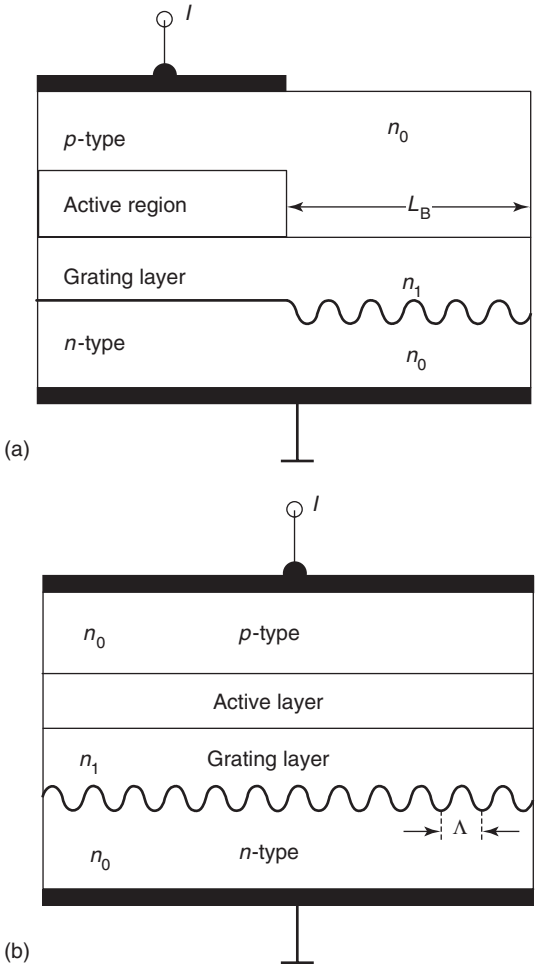


Figure 6 (a) DBR (longitudinal integration), and (b) DFB (transverse integration) single-frequency semiconductor laser structures.

where κ, the coupling coefficient of the grating describing the strength of the grating feedback has the form

$$\kappa = \frac{k_0 \Delta n_{eff}}{4}$$

and $\Delta\beta$ is the deviation of the wave vector $k_0 n_{eff}$ from the wave vector at the Bragg wavelength k_B and is thus a measure of the difference $\Delta\lambda = \lambda - \lambda_B$

$$\Delta\beta = k_0 n_{eff} - k_B \cong -\frac{2\pi n_{eff}}{\lambda_B^2}\Delta\lambda$$

L_B is the grating length, and $s = \sqrt{\kappa^2 - \Delta\beta^2}$. At the Bragg wavelength, the reflectivity is maximal and is given by

$$R(\lambda_B) = \tanh^2(\kappa L_B)$$

For most DBR semiconductor lasers, it is possible to achieve κL_B-products of the order 1 or more and thus

to realize reasonable reflectivities exceeding even the facet reflectivity of the semiconductor–air interface in the Fabry–Perot semiconductor lasers. However, the relative spectral position of the Bragg wavelength and the longitudinal mode spectrum is not controlled well in DBR lasers, particularly, if a wide operation temperature range is considered with resulting relative shifts of both against each other. The wavelength control and stability is much better in the DFB lasers because the longitudinal DFB laser modes are intrinsically related to the Bragg wavelength, and the grating and the active region are spatially not separated, so that any index changes in the active region or the grating layer may not change this relationship. As a consequence, DFB lasers have become commercial single-frequency lasers, while more sophisticated DBR devices have gained importance in the area of wavelength-tunable semiconductor lasers. Properly designed DFB lasers show excellent single-frequency performance with SMSR well exceeding 30 dB.

Vertical-Cavity Surface-Emitting Semiconductor Lasers

Vertical-cavity surface-emitting lasers (VCSELs) are relatively new semiconductor lasers in which the laser cavity is not along the wafer plane but vertical to it. Consequently, the length of the gain region along which the optical gain compensates internal and mirror losses is drastically reduced to the thickness of the active region (e.g., 50 nm) instead of its length (e.g., 500 μm). Because of the large optical gains achievable with semiconductors, it is possible to realize a sufficient gain along a gain path length of only 5–50 nm for the compensation of the losses of high-reflective mirrors. A schematic illustration of the basic VCSEL geometry is shown in **Figure 7**. Top and bottom mirrors of the VCSEL consist of several ten quarter-wavelength semiconductor layer pairs with alternating refractive index that can precisely be fabricated with molecular beam epitaxy (MBE) or metal-organic vapor-phase epitaxy (MOVPE).

The cavity length of VCSELs is of the order of the wavelength and thus orders of magnitude shorter than that of the edge-emitting lasers. Accordingly, the laser performance shows substantial differences, such as a longitudinal single-mode operation or the onset of quantization effects of the optical field.

Using thin QW active layers, the thickness d of the active region can be made much thinner than a quarter wavelength in the semiconductor, that is, $\lambda/4n$, so that by placing the active region in a maximum of the field intensity $|E|^2$ as shown in **Figure 7**, the gain is twice as large for the long active regions in edge-emitting layers, where the gain works on the

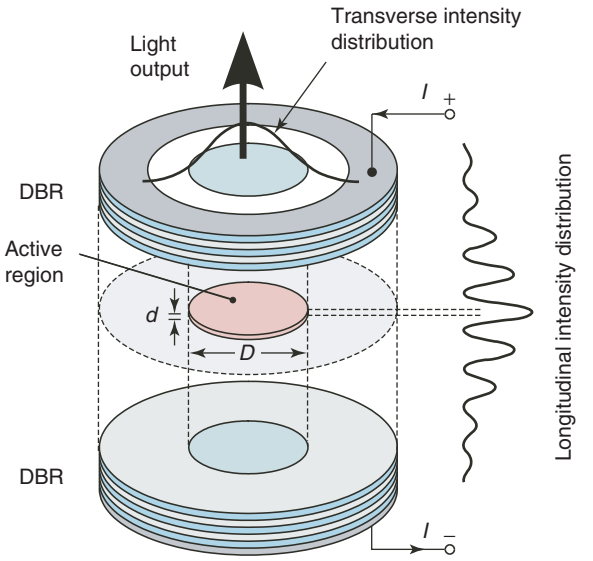


Figure 7 Schematical illustration of principal vertical-cavity surface-emitting laser (VCSEL) structure.

averaged field intensity, including nodes of $|E|^2$ as well. The threshold condition for this design, therefore, reads (note the factor of 2)

$$2dg = \alpha_i L_{\text{eff}} + \frac{1}{2} \ln \frac{1}{R_1 R_2}$$

where the effective cavity length L_{eff} is the vertical extension of the standing wave within the 1/e-decay on both ends. Due to the ultra-small d, the mirror reflectivities need to be near unity, typically $>99\%$. This can be achieved by DBRs made of alternating (and conductive) semiconductor layers of different composition yielding different refractive indices. Mathematically, the peak reflectivity of such a DBR at the Bragg wavelength is expressed as

$$R = \left[\frac{1 - (n_0/n_2)(n_2/n_1)^{2N}}{1 + (n_0/n_2)(n_2/n_1)^{2N}}\right]^2$$

where N is the number of layer pairs in the DBR, n_0 and n_2 are the refractive indices of the medium outside the VCSEL (e.g., air: $n_0 = 1$) and the inner semiconductor region respectively, and n_1 is the refractive index of the second semiconductor in the DBR. With refractive index differences up to 5–10% in usual semiconductor material systems (GaAlAs/GaAs, InGaAlAs/InP), several ten-layer pairs give reflectivities of the order 99.5–99.9% against air. With these high reflectivities, the laser threshold can be achieved with reasonable active region gains ~ 500–$1000\,\text{cm}^{-1}$. Up to now, room-temperature continuous-wave VCSELs have successfully been demonstrated for the wavelength ranges 850–1000 nm in Ga(In)AlAs/GaAs,

1.2–1.3 μm in GaInNAs/GaAs, and 1.3–2.0 μm in AlGaInAs/InP.

The lateral diameter of the active region can be kept small enough to favor single transverse-mode operations and is typically $\sim$3–10 μm. As a consequence of the small active region volume, the threshold currents of VCSELs are rather small, typically <1 mA, even though the threshold current density might be quite large to obtain the required large threshold gain. Another consequence of the small active region is that the output power of a VCSEL is accordingly small, typically of the order of a few milliwatts, which is, however, sufficient for the vast majority of applications in communications, measurements, and sensing.

See also: Quantum Cascade Lasers; Light Emitting Diodes.

PACS: 42.55.Px; 42.55.Sa; 85.35.Be; 85.60.−q

Further Reading

Casey HC and Panish MB (1978) *Heterostructure Lasers – Part A: Fundamental Principles*, 1st edn. New York: Academic Press.

Casey HC and Panish MB (1978) *Heterostructure Lasers – Part B: Materials and Operating Characteristics*, 1st edn. New York: Academic Press.

Chuang SL (1995) *Physics of Optoelectronic Devices*, 1st edn. Chichester: Wiley.

Coldren LA and Corzine SW (1995) *Diode Lasers and Photonic Integrated Circuits*, 1st edn. Chichester: Wiley.

Kapon E (1999) *Semiconductor Lasers I: Fundamentals*, 1st edn. San Diego: Academic Press.

Kapon E (1999) *Semiconductor Lasers II: Materials and Structures*, 1st edn. San Diego: Academic Press.

Li H and Iga K (2002) *Vertical-Cavity Surface-Emitting Laser Devices*, 1st edn. Berlin: Springer.

Morthier G and Vankwikelberge P (1997) *Handbook of Distributed Feedback Laser Diodes*, 1st edn. Norwood: Artech House.

Petermann K (1988) *Laser Diode Modulation and Noise*, 1st edn. Dordrecht: Kluwer Academic Publishers.

Zory PS (1993) *Quantum Well Lasers*. 1st edn. Boston: Academic Press.

Nomenclature

a_0	lattice constant (nm)
c	vacuum speed of light (cm s^{-1})
d	thickness of active region (μm)
D	diameter of circular active region in VCSEL (μm)
e	elementary charge (As)
E_c	conduction band edge energy (eV)
E_{Fc}	quasi-Fermi level in conduction band (eV)
E_{Fv}	quasi-Fermi level in valence band (eV)
E_g	bandgap energy (eV)
E_V	valence band edge energy (eV)
f	frequency (Hz)
g	optical material gain (cm^{-1})
h	Planck's constant (VAs^2)
H	modulation response
I	current (mA)
I_{th}	laser threshold current (mA)
k	wave vector (cm^{-1})
k_0	free-space wave vector $k_0 = 2\pi/\lambda$ (cm^{-1})
k_B	wave vector at Bragg wavelength (cm^{-1})
L	laser length (μm)
L_B	length of Bragg grating (μm)
$n = n' + jn''$	complex refractive index
n_{eff}	effective refractive index of a waveguide mode $n_{eff} = \beta/k_0$
n_g	group refractive index
n_{sp}	spontaneous emission coefficient (typically 1–2)
N	carrier density (cm^{-3})
N_c	effective density of states in conduction band (cm^{-3})
N_{th}	carrier density at laser threshold (cm^{-3})
N_v	effective density of states in valence band (cm^{-3})
P	optical power (mW)
R	recombination rate (cm^{-3} s^{-1}), mirror reflectivity (dimensionless)
RIN	relative intensity noise (dB Hz^{-1})
S	photon density (cm^{-3})
SMSR	side-mode suppression ratio (dB)
V_a	active region volume (cm^3)
v_g	group velocity (cm s^{-1})
w	stripe width (μm)
W	full width at half maximum (FWHM) (μm) of laser near-field
α	optical loss coefficient (cm^{-1})
α_H	Henry's line width enhancement factor
α_i	intrinsic optical loss (cm^{-1})
α_m	mirror loss (cm^{-1})
α_{tot}	total optical loss $\alpha_{tot} = \alpha_i + \alpha_m$ (cm^{-1})
β	propagation constant (cm^{-1})
Γ	confinement factor of an optical mode
ΔE_F	difference of quasi-Fermi levels, E_{Fc}–E_{Fv} (eV)
$\Delta\Phi$	FWHM of laser far-field (rad,^{o})
κ	coupling coefficient of Bragg grating (cm^{-1})
Λ	(Bragg) grating period (nm)
λ	wavelength (μm)
λ_B	Bragg wavelength (μm)
ν	frequency (Hz)
τ_d	differential carrier lifetime (s)
τ_{on}	turn-on delay time (s)
τ_s	photon lifetime (s)
ω	circular frequency (s^{-1})
ω_d	damping circular frequency (s^{-1})
ω_r	relaxation circular frequency (s^{-1})

Semiconductor Nanostructures

B Voigtländer, Forschungszentrun Jülich GmbH
Jülich, Germany

Introduction

Two conceptually different methods are used to fabricate semiconductor nanostructures. In the top–down approach, lithography is used to fabricate nanostructures. In the bottom-up approach, small nanostructures form by self-organization during epitaxial growth. The greatest advantage of the top–down approach is the large variety of different structures, which can be generated by lithographic methods. In the case of the bottom-up approach, the variety of structures is much more limited. Islands, wires, rods, and rings have been fabricated bottom-up. However, the bottom-up approach offers the unique opportunity to fabricate very small nanostructures down to the one-digit nanometer range beyond the limits of current lithography techniques. Another advantage of self-organized fabrication of semiconductor nanostructures is the inherent parallel approach, in which billions of nanostructures are generated in parallel.

The focus of this article is the self-organized growth of semiconductor structures with lateral dimensions below 100 nm. Semiconductor quantum-well structures are nanostructures in the growth directions and not lateral nanostructures, and will not be considered in this article. The actual growth techniques used to fabricate the nanostructures like molecular beam epitaxy (MBE) and chemical vapor deposition (CVD) will be considered in separate articles. The methods used to analyze the nanostructures are scanning tunneling microscopy (STM), scanning force microscopy (AFM or SFM), transmission electron microscopy (TEM), and low energy electron microscopy (LEEM). Different properties of semiconductor nanostructures, and the applications of self-organized nanostructures in devices are discussed elsewhere in this encyclopedia. As examples for the self-organized growth of nanostructures, material systems such as Si/Ge and GaAs/InAs are used predominantly.

After introducing the self-organized semiconductor nanoislands grown on planar substrates, the aligned growth on prestructured samples is discussed. Subsequently, self-organized nanowires grown on facetted surfaces and the growth of monolayer thick wires by step-flow growth are presented. Finally, the growth of nanowires by vapor liquid solid (VLS) epitaxy and hybrid systems, where nanowires aligned to lithographically defined structures are fabricated, are presented.

Physical Principles of Self-Organized Growth

Semiconductor nanostructures can be fabricated by self-organization using heteroepitaxial growth, which is the growth of a material B on a substrate of a different material A. In heteroepitaxial growth, the lattice constants of the two materials are often different. The lattice mismatch for the two most commonly used material systems Si/Ge and GaAs/InAs is 4.2% and 7%, respectively (schematically shown in **Figure 1a**). This lattice mismatch leads to a buildup of elastic stress in the initial two-dimensional (2D) growth in heteroepitaxy. In the case of Ge

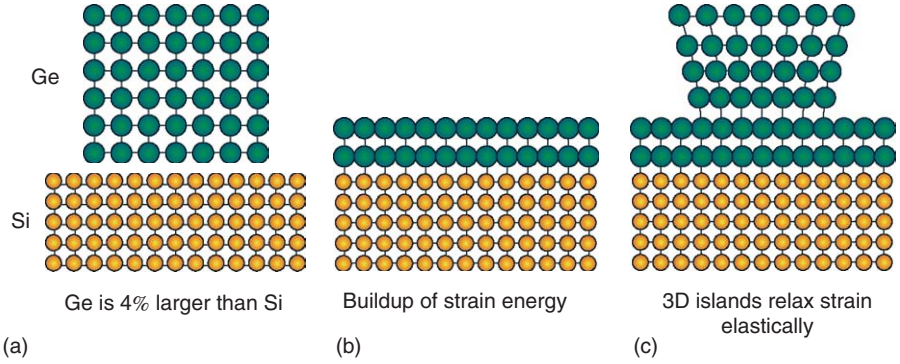


Figure 1 (a) Schematic representation of Si and Ge crystals with different lattice constants, (b) buildup of elastic strain energy during 2D growth with Ge confined to the Si lattice constant, and (c) elastic relaxation by formation of 3D islands (Stranski–Krastanov growth). In the upper part of the 3D island, the lattice constant relaxes towards the Ge bulk constant. The usual form of the 3D islands is a pyramid and not like the one shown in this schematic sketch.

heteroepitaxy on Si, the Ge is confined to the smaller lattice constant of the Si substrate, that is, the Ge is strained to the Si lattice constant. This results in a buildup of elastic stress (**Figure 1b**). One way to relax this stress is the formation of 3D Ge islands. In the 3D islands, only the bottom of the islands is confined to the substrate lattice constant. In the upper part of the 3D island, the lattice constant can relax to the Ge bulk lattice constant and reduce the stress energy this way (**Figure 1c**). This growth mode characterized by the formation of a 2D wetting layer and the subsequent growth of (partially relaxed) 3D islands is called Stranski–Krastanov growth mode.

The driving force for the formation of self-organized semiconductor nanoislands in heteroepitaxial growth is the buildup of elastic strain energy in the stressed 2D layer. As a reaction to this, a partial stress relaxation by the formation of 3D islands can lower the free energy of the system. The process of island formation close to equilibrium is a trade-off between elastic relaxation by formation of 3D islands which lowers the energy of the system and an increase of the surface area which increases the energy. Apart from the formation of 3D islands, there is another process which can partially relax the stress of a strained 2D layer: the introduction of misfit dislocations. This corresponds to the removal of one lattice plane of a compressively strained 2D layer. If a lattice plane is removed in regular distances in the 2D layer, a misfit dislocation network forms.

Depending on the growth parameters, temperature, and growth rate, the self-organized growth can be close to equilibrium or in the kinetically limited regime. Close to equilibrium, that is, at high-growth temperatures or low-deposition rates, the occurring morphology (strained layer or 3D islands or a film with dislocations) is only determined by the energies of the particular configurations. The morphology with the lowest energy will be formed. If the growth is kinetically limited, the activation barriers are important. For instance, an initially flat strained layer can transform to a morphology with 3D islands or to a film with dislocations. What actually happens depends on the kinetics of the growth process, that is, on the activation energy for formation of 3D islands compared to the activation energy for the introduction of misfit dislocations.

Semiconductor Nanoislands

Stranski–Krastanov growth occurs in InAs/GaAs growth. An example of InAs nanoislands grown on a GaAs substrate is shown in the atomic force microscopy image in **Figure 2**. The InAs islands were grown by MBE at a growth temperature of 800 K. The width

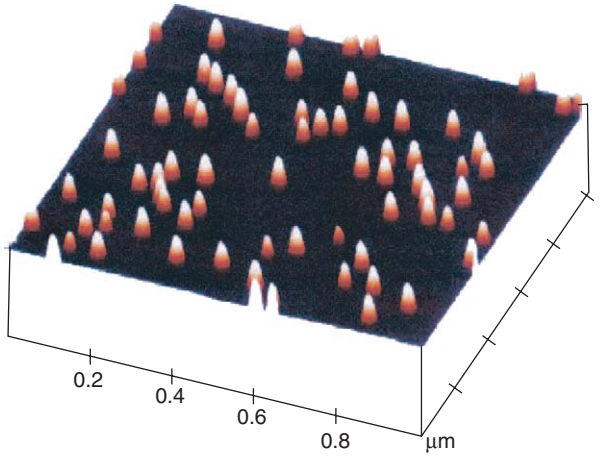


Figure 2 InAs nanoislands grown on a GaAs surface imaged by AFM. (Courtesy of A Lorke; reproduced with permission from Warburton RJ (2002) *Contemporary Physics* 44: 351; © Taylor & Francis Ltd. (http:www.tandf.co.uk).)

of the islands is 20 nm and the height 6 nm with a density of $10^{10}\,\mathrm{cm}^{-2}$. The challenges in the growth of these semiconductor islands are to grow islands of the desired size and density and with a high-size uniformity. In general, a higher growth temperature generally leads to the formation of larger islands; a higher growth rate leads to the formation of smaller islands. The size of the islands increases with the coverage. Often the density of the islands saturates at an early stage of the growth. These are general trends which may depend on the material system and the particular deposition technique. In some cases (self-limiting growth), the size of the islands saturates and the density increases with coverage. This kind of growth mode leads to a high-size uniformity of the islands. The size uniformity achieved in self-organized growth of semiconductor islands can be as small as $\sim 10\%$. For optoelectronic applications, the nanoislands have to be capped (i.e., embedded in a semiconductor to prevent oxidation of the islands under ambient conditions where the optic measurements are performed) and care has to be taken that they change their shape and composition during the capping process. The confinement of charge carriers in all three directions gives rise to atomic-like energy levels. Quantum dot lasers operating at room temperature have now been realized. The islands grown on a flat substrate are usually not ordered laterally due to the random nature of the nucleation process. In the following sections, it is shown how nucleation at specific sites can be achieved.

Lateral Positioning of Nanoislands by Growth on Templates

On homogenous substrates, the location of the nanostructured islands formed were not predictable due

to the stochastic nature of the nucleation process. It is desirable to position islands at specific (preselected) locations. This is desirable if the islands are used as functional units of nanoelectronic devices. Electric contacts can be structured easier when the location of an island is predefined. The approaches are now discussed where island nucleation occurs not randomly, but at specific locations. These predefined locations are usually specific sites on regularly prestructured substrates. Examples of such prestructured substrates are: regularly stepped substrates, facetted substrates, a regular arrangement of dislocations, or long-range surface reconstructions.

An example of ordered nucleation at a prestructured substrate is shown in **Figure 3**. Here Ge islands nucleate above dislocation lines. An SiGe film is grown on an Si(0 0 1) substrate. At the interface between the SiGe film and the substrate, dislocations form. The driving force for the formation of the dislocations is the relief of elastic strain, which arises due to the different lattice constants between the Si substrate and a Ge/Si film on this substrate. Due to the crystal structure of the substrate, the dislocations are aligned in straight lines along the two perpendicular high-symmetry directions of the substrate. During annealing, the dislocations form a relatively regular network, due to a repulsive elastic interaction between the dislocations. **Figure 3** shows growth of Ge islands on this prestructured substrate. Rows of islands grow preferentially above the dislocation lines. The nucleation of Ge islands is more favorable at the relaxed areas above the dislocations, which have a lattice constant closer to that of Ge than the strained areas of the SiGe film, where the lattice constant is confined to that of the Si substrate. Islands have a lower energy if they grow with their natural lattice constant compared to growth at a different lattice constant (strained islands). This leads to a preferred nucleation of Ge islands above the dislocations. The nucleation does not occur randomly at the surface, but simultaneously at sites which have the same structure. This can lead to a narrower size distribution than for the growth on unstructured Si(0 0 1) substrates.

In the case of the growth on a prestructured substrate, another process of self-organization occurs (before the growth of islands) during the formation of the regular arrangement of the defect structures. Here, it is often a repulsive elastic interaction, which leads to a regular distance between steps or dislocations under equilibrium conditions. Since island nucleation occurs at the defect sites, the distances are determined by the distances of the defects. Nucleation at predefined defect sites has two advantages over the random nucleation. First, the islands are located at specific sites and second, the size distribution is narrower due to simultaneous nucleation at sites with identical environment.

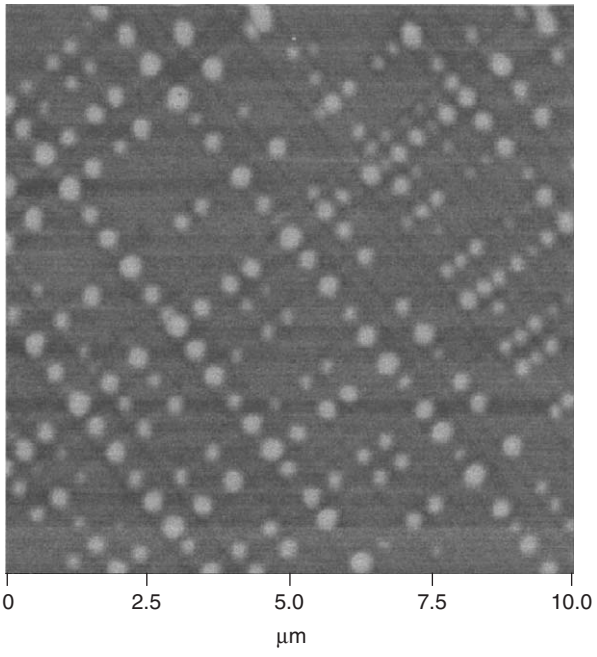


Figure 3 Ordered nucleation of Ge islands which is prestructured by an underlying network of dislocations. Image size 7 μm. (Reprinted figure with permission from Shiryaev SY, Jensen F, Lundsgaard Hansen J, Wulff Petersen J, and Nylandsted Larsen A (1997) *Physical Review Letters* 78: 503; © American Physical Society.)

Vertical Stacking of Nanoislands

A nucleation of nanoislands preferentially above existing ones can be achieved when a new layer of Ge (InAs) islands is grown on top of an Si (GaAs) spacer layer. If the spacer layer is sufficiently thin, the nucleation of islands occurs just above pre-existing islands. This vertical stacking of islands is shown in **Figure 4** and can be explained as follows: the lattice constant of the top part of the Ge islands is increased close to the bulk constant. Subsequent Si deposition leads to an increased lattice constant of the Si spacer layer just above the Ge islands. Upon the nucleation

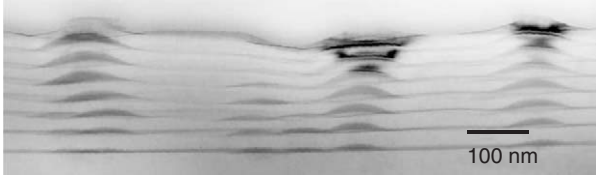


Figure 4 Cross-sectional TEM image of a stack of Ge islands in 40 nm thick Si spacer layers. The islands are aligned on top of underlying islands. (Reproduced from Vescan L (1998) *Thin Solid Films* 336: 244, with permission from Elsevier.)

of the next Ge island layer, Ge islands nucleate preferentially at locations on the Si spacer layer where the lattice constant is closest to the Ge lattice constant. This is just the case above the Ge islands. The residual strain field of the Ge islands in the lower layer seeds the nucleation in the upper layer. It was observed that the size uniformity of the 3D islands in subsequent layers is improved.

Semiconductor Nanowires

Nanowires on Facetted Surfaces

Some surfaces such as the Si(1 1 3) surface are known to facet, that is, to form alternating regions with a very high step density (step bunches) alternating with flat surface regions. Subsequent growth of Ge can result in preferential growth of nanowires in the step-bunch region. This effect is even more pronounced in Si/Ge multilayers. In **Figure 5**, a cross-sectional TEM image of a stack of five alternating depositions of Ge and Si is shown. Ge nanowires with a thickness of $\sim$4.5 nm and a width of $\sim$100 nm are formed. The formation of these nanowires is explained as a strain relaxation effect of Ge growing at the step-bunch regions where it can relax the elastic strain more effectively than on a compact 2D layer. After growth of an Si buffer layer, the next SiGe wire structure nucleates above the previous one. This is a strain-induced effect similar to the one shown for the vertical stacking of the islands.

Monolayer Thick Wires at Step Edges

Regular surface steps can be used to fabricate monolayer thick Ge wires using step-flow growth. Pre-existing step edges on the Si(1 1 1) surface are used as templates for the growth of 2D Ge wires at the step edges. When the diffusion of the deposited atoms is sufficient to reach the step edges, these deposited atoms are incorporated exclusively at the step edges and the growth proceeds by a homogenous advancement of the steps (step-flow growth mode). If small amounts of Ge are deposited, the steps advance only some nanometers, and narrow Ge wires can be grown. A key issue for the controlled fabrication of nanostructures consisting of different materials is a method of characterization, which can distinguish between the different materials on the nanoscale. In case of the important system Si/Ge, it has been difficult to differentiate between Si and Ge due to their similar electronic structure. However, if the surface is terminated with a monolayer of Bi, it is possible to distinguish between Si and Ge. **Figure 6a** shows an STM image after repeated alternating deposition of 0.15 atomic layers of Ge and Si, respectively. Due to the step-flow growth, Ge and Si wires are formed at the advancing step edge. Both elements can be easily distinguished by the apparent heights in the STM images. It turns out that the height measured by the STM is higher on areas consisting of Ge (red stripes) than on areas consisting of Si (orange stripes). The assignment of Ge and Si wires is evident from the order of the deposited materials (Ge, Si Ge, Si and Ge, respectively in this case). The initial step position is located right from the grown Ge wires. The step edge has advanced towards the left after the growth of the nanowires. The reason for the height difference in STM between Si and Ge is the different electronic structure of a Ge–Bi bond compared to an Si–Bi bond. The apparent height of Ge areas is $\sim$0.1 nm higher than the apparent height of Si wires (**Figure 6b**). The width of the Si and Ge wires is $\sim$3.5 nm as measured from the cross section (**Figure 6b**). The nanowires are 2D with a height of only one atomic layer (0.3 nm). Therefore, the cross section of a 3.3 nm wide Ge nanowire contains only 21 atoms (**Figure 6c**). The height difference arises due to an atomic layer of Bi which was deposited initially. The Bi floats always on top of the growing layer because it is less strongly bound to the substrate than Si or Ge. The Si/Ge wires are homogenous in width over larger distances and have a length of several thousand nanometers. Different widths of the wires can be easily achieved by different amounts of Ge and Si deposited.

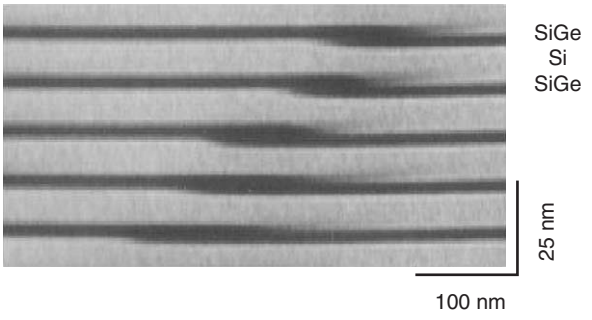


Figure 5 Cross-sectional TEM image of stacked Ge nanowires grown on a facetted Si(1 1 3) surface. (Reproduced with permission from Brunner K (2002) Si/Ge Nanostructures. *Report of Progress in Physics* 65: 27–72; © IOP Publishing.)

Free-Standing Semiconductor Nanowires Grown by Vapor Liquid Solid Growth

In the VLS method, small catalytic nanoparticles (often gold) induce the growth of a homogenous rod of a semiconductor, with the diameter of the rod determined by the nanoparticle. The gold nanoparticles can be fabricated by heating a thin evaporated gold

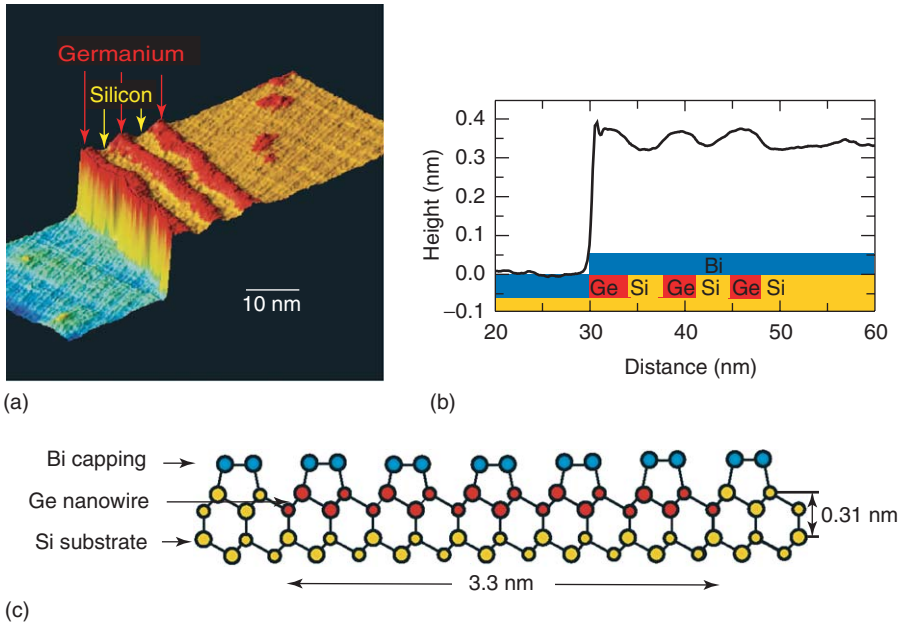


Figure 6 (a) STM image of 2D Ge/Si nanowires grown by step flow at a pre-existing step edge on a Si(1 1 1) substrate. Si wires (orange) and Ge wires (red) can be distinguished by different apparent heights. (b) The cross section shows the dimensions of the Si and Ge nanowires. The width of the wires is 3.5 nm and the height is only one atomic layer (0.3 nm). (c) Atomic structure of a 3.3 nm wide Ge wire on the Si substrate capped by Bi. The cross section of the Ge wire contains only 21 Ge atoms. (Reproduced with permission from Voigtländer B (2001) Fundamental processes in Si/Si and Ge/Si epitaxy studied by scanning tunneling microscopy during growth, *Surface Science Reports* 43: 127.)

film, which breaks up and reshapes into nanoscale droplets. Alternatively, gold aerosol nanoparticles with sizes of a few tens of nanometers are used. The gold nanoparticles form a eutectic alloy with the substrate material (for instance, GaAs). Using MBE or metallorganic vapor phase epitaxy, nanowires can be grown. These nanowires are different from the ones discussed so far, in that they are free-standing on the substrate and grow vertically. The previously considered nanowires were epitaxially grown along their length on the substrate or even embedded into the substrate during growth.

The VLS method was used for the Si/Ge and the GaAs material system. This growth method can even be used to grow semiconductor heterostructures within the nanowires with atomically sharp interfaces. **Figure 7** shows an InAs nanowire (green) with several InP barriers (red). The rapid alternation of the composition is controlled by the supply of precursor atoms to the eutectic melt supplied as molecular beams. Of particular interest is the fact, that the very small cross section allows efficient lateral relaxation of the nanowire, thereby providing freedom to combine materials with very different lattice constants to create heterostructures within the nanowire. The problem of incorporation of misfit dislocations, when a critical thickness is exceeded, does not occur due to the small lateral size of the nanowires. The

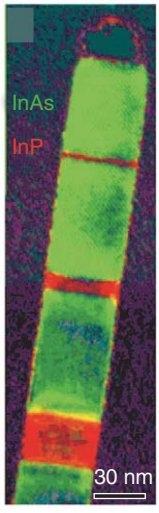


Figure 7 Composition profile of an InAs nanowire, containing several InP heterostructures, using reciprocal space analysis of lattice spacings with a TEM. InAs lattice spacings have been color-coded with green and InP spacings with red. (Reproduced with permission from Björk MT, Ohlsson BJ, Sass T, Persson AI, Thelander C, *et al.* (2002) *Nano Letters* 2: 88.)

VLS method is also used to grow semiconductor heterostructures, not along the wire, but radial heterostructures, the so-called core-shell structures.

The synthesis of Si/Ge core-shell nanowires by chemical vapor deposition is achieved by the following

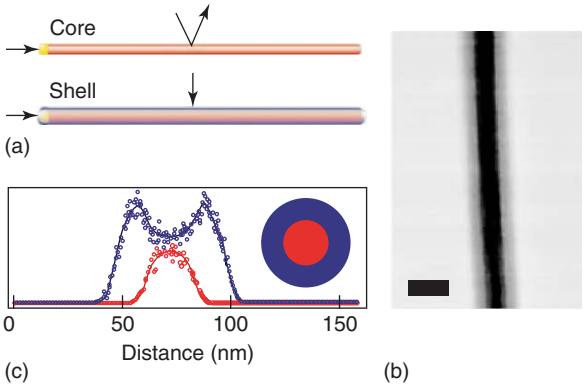


Figure 8 (a) Schematic of the core-shell growth of a radial Ge/Si structure (b) TEM image of a Ge/Si core-shell nanowire. Scale bar is 50 nm. (c) Elemental mapping along the cross section showing the Ge (red circles) and Si (blue circles) concentrations. (Reproduced with permission from Lauhon LJ, Gudiksen MS, Wang D, and Lieber CM (2002) Epitaxial core-shell and core-multishell nanowire heterostructures. *Nature* 420: 57.)

process schematically shown in **Figure 8a**. Initially a Ge nanowire is grown by vapor phase epitaxy growth from a gold nanoparticle. In this case, the substrate serves only as a support and the nanowires are not epitaxially connected to the substrate. The deposition temperature is chosen so low that no germane (GeH_4) is decomposed along the wire and radial growth is suppressed. Only axial growth occurs by the VLS mechanism at the gold nanoparticle. Subsequently, a boron-doped Si shell is grown by CVD. The addition of diborane serves to lower the decomposition temperature of silane (SiH_4) and "turns on" the radial growth. **Figure 8b** shows a TEM image of the Ge core (imaged darker) and the Si shell (lighter gray) with a diameter of $\sim$50 nm. The elemental TEM cross-sectional mapping shown in **Figure 8c** shows the radial chemical composition in a wire (core diameter 26 nm, shell thickness 15 nm). By oxidation of a shell of the radial heterostructure, it is possible to build a coaxially gated nanowire transistor.

Hybrid Systems – Combination of Lithography and Self-Organized Growth

In hybrid methods, self-organization is combined with lithographic patterning. In this approach, self-organization is used to form nanostructures on a smaller scale than the one accessible by lithography. Most importantly, the hybrid methods provide a direct contact of nanostructures formed by self-organization to lithographically patterned structures. As an example, the self-organized growth of Ge islands in oxide holes is shown in **Figures 9a–9d**. The staring

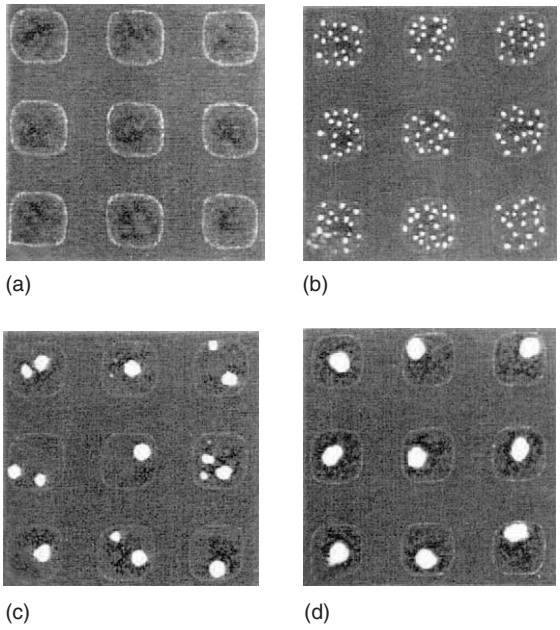


Figure 9 (a–d) Growth of Ge islands in holes on an oxidized Si substrate. (Reproduced with permission from Kim ES, Usami N, and Shiraki Y (1999) *Semiconductor Science and Technology* 14: 257; © IOP Publishing Limited.)

surface is a silicon substrate with a thin oxide layer at the surface. Electron lithography is used to remove the oxide and form holes of a diameter of 0.5 μm where the bare Si surface is exposed. Self-organized growth of Ge leads to the formation of Ge islands, which can be smaller than the size scale of the electron beam lithography. The gas phase growth of Ge is selective, that is, Ge only grows inside the holes in the oxide and not on the oxide itself. **Figure 9** shows the nucleation of Ge islands in the holes in the oxide for different growth temperatures. At lower temperatures, the island density is so large that several islands nucleate in one oxide hole. If the temperature is increased finally, only one Ge island nucleates in each oxide hole. The size of the Ge island is smaller than the lithographically defined oxide hole. This approach is called lithographic downscaling. However, the positioning of the islands is not perfect. As seen in **Figure 9d**, the position of the Ge island inside the oxide hole is not defined but is rather randomly in the center or at the corner of the oxide hole. Due to the fact that the Ge does not grow on the oxide, the edges of the oxide hole are not sinks for deposited Ge atoms. Therefore, the Ge adatom concentration is homogenous across the hole, and the nucleation of the Ge island is random within the oxide hole. If the edges of the hole were sinks for Ge atoms (for instance, if the edges of the hole consisted of Si), the adatom density would have a maximum at the center of the hole and the nucleation of the Ge islands

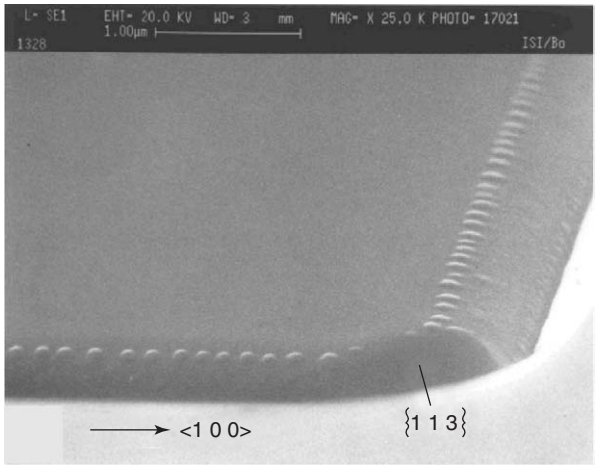


Figure 10 Regular alignment of Ge islands at the edge of a mesa which was fabricated by optical lithography. (Reproduced from Vescan L (1989) *Journal of Crystal Growth* 194: 173, with permission from Elsevier.)

would occur preferentially at the center of the oxide holes.

A simple structure formed by lithographic methods is a mesa (from the Spanish table), which is a flat terrace separated from the rest of the wafer by trenches. A silicon mesa structure is imaged by electron microscopy and shown in **Figure 10**. It has been observed that the growth of Ge on this mesa leads to preferential nucleation of Ge islands at the mesa edges. **Figure 10** shows a regular alignment of Ge islands along the mesa edges. This preferred nucleation of islands at the convex part of the surface profile is quite unexpected. Considering the surface curvature alone, the chemical potential is lowest at locations with a concave curvature. Atoms diffuse toward the lower chemical potential, that is, to the concave parts of the surface. This reasoning is true for homoepitaxy. In heteroepitaxy, another contribution is important: the strain-dependent contribution to the chemical potential. The convex regions are most favorable for strain relaxation of the strained Ge wetting layer covering the Si substrate. Therefore, the strain contribution to the chemical potential has a minimum at convex edges. In the current case, this strain-induced contribution to the chemical potential overwhelms the curvature contribution and leads to a minimum in the total chemical potential at the convex edges. Therefore, these convex edges with a minimum in the chemical potential provide a favorable nucleation site for Ge islands.

Also, nanowires can be fabricated, aligned to lithographically defined structures. V-grooves can be defined by lithography and anisotropic wet chemical etching with a low-etch speed for the crystal plane forming the side facets of the V-groove. The formation

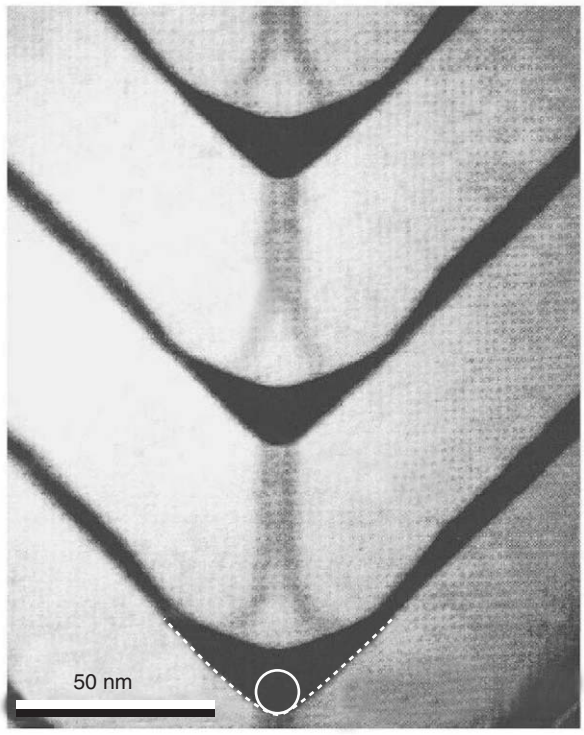


Figure 11 Cross-sectional TEM image of vertically stacked GaAs nanowires and AlGaAs barriers aligned along a lithographically defined V-groove. (Reproduced with permission from Gustafsson A, Reinhardt F, Baisiol G, and Kapon E (1995) *Applied Physics Letters* 67: 3673; © American Institute of Physics.)

of nanowires at the bottom of the V-groove is induced by the self-limiting nature of the growth front of AlGaAs on the grooved substrate. Vertically stacked arrays of GaAs/AlGaAs nanowires exhibiting 5–10% size uniformity have been demonstrated (**Figure 11**). The layer structure consists of GaAs nanowires separated by $Al_{0.42}Ga_{0.58}As$ barriers. The formation of the crescent-shaped GaAs wires is based on the formation of slow-growing side facets along the V-grooves. The migration of adatoms away from these side-walls toward the V-groove bottom produces the concave surface profiles in the corners of the V-grooves.

See also: Epitaxy; Film Growth and Epitaxy: Methods; Lithography Masks and Pattern Transfer; Luminescence; Nanostructures, Optical Properties of; Quantum Devices of Reduced Dimensionality; Scanning Probe Microscopy; Semiconductor Devices.

PACS: 68.65. – k

Further Reading

Berbezier I, Ronda A, and Portavoce A (2002) Si/Ge nanostructures: new insights into growth processes. *Journal of Physics: Condensed Matter* 14: 8283.

Bimberg D, Grundmann M, and Ledentsov NN (1989) *Quantum Dot Heterostructures*. New York: Wiley.
Gustafsson A, Reichhardt F, Baisiol G, and Kapon E (1995) Low-pressure organometallic chemical vapor deposition of quantum wires on V-grooved substrates. *Applied Physics Letters* 67: 3673.
Moriarty P (2001) Nanostructured materials. *Reports of Progress in Physics* 64: 297.
Motta N (2002) Self-assembling of Ge/Si(1 1 1) quantum dots: scanning microscopy probe studies. *Journal of Physics: Condensed Matter* 14: 8353.
Stangl J, Holy V, and Bauer G (2004) Structural properties of self-organized semiconductor nanostructures. *Reviews of Modern Physics* 76: 725.
Teichert C (2002) Self-organization of nanostructures in semiconductor heteroepitaxy. *Physics Report* 365: 335.
Yang B, Feng Liu, and Lagally MG (2004) Local strain-mediated chemical potential control of quantum dot self-organization in heteroepitaxy. *Physics Review Letters* 92: 025502.

Semiconductor Optics

G C La Rocca, Scuola Normale Superiore, Pisa, Italy

Introduction

Even though the paramount role of semiconductors, and of silicon above all, in modern technology is mainly due to their use in electronic microdevices, the optical properties of these materials are also of great interest. In particular, many applications of semiconductors are based on optoelectronic effects: the transformation of light in electric current (photovoltaic devices), and the transformation of electric current in light (semiconductor light emitting diodes and lasers). Of course, the spectroscopy of semiconductors has been, and continues to be, a subject of crucial relevance from the viewpoint of basic science as it allows a detailed study of the electronic structure of these materials. The field is still in great expansion, also due to the possibility of growing layers of different materials on top of each other by molecular beam epitaxy in order to engineer their electronic and optical properties.

The following concentrates on bulk properties rather than on heterostructures, and the focus is on the optical part of the spectrum due to interband electronic excitations (possibly assisted by phonon participation). After discussing general concepts related to the linear optical response functions, the corresponding absorption spectra are considered as they are much less sensitive than luminescence to extrinsic effects. Excitonic and polaritonic phenomena are also touched upon. Finally, the role of external perturbations on the optical properties is discussed (modulation spectroscopy).

The Dielectric Function

Semiconductors are characterized by an energy gap of the order of the electronvolt (see **Table 1**) and correspondingly by an absorption edge in the near infrared or in the visible below which they are nearly transparent and above which they are almost opaque. Actually, a direct measurement of the absorption is rather difficult as it is so strong when the frequency is higher than the threshold frequency that no radiation goes through the semiconductor, except in the case of extremely thin layers. This difficulty can be overcome when the optical quality of the surfaces makes it possible to measure by ellipsometry, both the real and the imaginary part of the dielectric function as a function of the light frequency. Alternatively, a measurement of the reflectivity over a large frequency range with sufficient accuracy and the application of the Kramers–Kronig relations allows one to determine the real and the imaginary part of the reflection coefficient. From this, the dispersion and the absorption can be obtained, establishing the bases for the experimental determination of the optical response of semiconductors and of its theoretical interpretation in terms of the electronic band structure.

It may be recalled that the normal incidence reflection coefficient, defined as the ratio between the electric field of reflected and incident wave, is related to the complex index of refraction $\tilde{n}$ by the Fresnel relation

$$r = \frac{E_r}{E_i} = \frac{\tilde{n} - 1}{\tilde{n} + 1} = |r|e^{i\theta} \qquad [1]$$

and taking the logarithm

$$\log r = \log |r| + i\theta \qquad [2]$$

Since this is an analytic function of the frequency ω in the upper complex plane $\omega + i\eta$, the following

Table 1 Energy gap values (in eV) of cubic semiconductors at room temperature

	AlSb (i)	CdTe (d)	GaAs (d)	GaP (i)	GaSb (d)
E_g	1.52	1.45	1.43	2.26	0.71
	Ge (i)	InAs (d)	InP (d)	InSb (d)	Si (i)
E_g	0.67	0.35	1.35	0.18	1.14

(d) = direct gap; (i) = indirect gap.

Kramers–Kronig relation holds true:

$$\theta(\omega) = -\frac{2\omega}{\pi}\,\mathscr{P}\int_0^\infty \frac{\log|r(\omega')|}{\omega'^2-\omega^2}\,\mathrm{d}\omega' \qquad [3]$$

and from the measured reflectivity $|r(\omega')|^2$, $\theta(\omega)$ can be obtained, and consequently the real and imaginary parts of $\tilde{n} = n + \mathrm{i}k$ can be derived from [1].

This amounts to the knowledge of dispersion and absorption since the complex dielectric function is related to $\tilde{n}$ by

$$\tilde{\varepsilon} = \varepsilon_1 + \mathrm{i}\varepsilon_2 = \tilde{n}^2 \qquad [4]$$

The dissipation is related to the imaginary parts, since the absorption coefficient is given by

$$\alpha(\omega) = \frac{2k\omega}{c} = \frac{\varepsilon_2\omega}{nc} \qquad [5]$$

while the real part $n(\omega)$ gives the dispersion.

To relate the measured optical constants to the electronic structure, the coupling dipolar interaction between the electromagnetic radiation and the electrons is to be considered:

$$H'(\boldsymbol{r}) = \mathrm{e}\boldsymbol{E}\cdot\boldsymbol{r} \qquad [6]$$

where $\boldsymbol{E} = \boldsymbol{E}_0\mathrm{e}^{-\mathrm{i}\omega t} + \boldsymbol{E}_0^*\mathrm{e}^{-\mathrm{i}\omega t}$ denotes the electric field of the radiation, and $\boldsymbol{r}$ is the position of the electron which can make a transition between different electronic states. One may recall that the transition probability per unit time between an occupied lower initial state $|i\rangle$ and an empty higher final state $|f\rangle$ is given by the Fermi golden rule

$$P^{(1)}_{i\to f} = \frac{2\pi}{\hbar}\,|\langle f|H'|i\rangle|^2\delta(E_f - E_i - \hbar\omega) \qquad [7]$$

In the case of semiconductors, one must sum over all occupied initial states (valence bands) and empty final states (conduction bands) in the unit volume to compute the electromagnetic power dissipated at a given frequency

$$W(\omega) = \frac{1}{V}\sum_{i,\,f}\hbar\omega P^{(1)}_{i\to f} \qquad [8]$$

where V is the crystal volume. The absorption coefficient is by definition

$$\alpha(\omega) = \frac{\text{dissipated power per unit volume}}{\text{incoming flux}} = \frac{W(\omega)}{(cnE^2/2\pi)} \qquad [9]$$

The imaginary part of the dielectric function, rather than the absorption coefficient, can be directly calculated as it does not depend on the real part of the index of refraction. From the above expressions, one obtains for the dissipative function

$$\varepsilon_2(\omega) = 4\pi^2\sum_{c,v}\int\frac{2\mathrm{d}\boldsymbol{k}}{(2\pi)^3}\,|\langle\psi_{ck}|e\boldsymbol{r}|\psi_{vk}\rangle|^2 \times\delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \qquad [10]$$

where the wave functions and their energies are given by a band structure calculation, the integration is over the first Brillouin zone and a factor of 2, assuming spin degeneracy has been included. Here, as the wave vector of light $2\pi/\lambda$ is much smaller than the size of the Brillouin zone π/l, being l the lattice constant, only allowed 'vertical' transitions between electronic states having the same $\boldsymbol{k}$ have been considered.

The real part $\varepsilon_1(\omega)$ is then obtained from the Kramers–Kronig dispersion relation

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi}\mathscr{P}\int_0^\infty\frac{\omega'\varepsilon_2(\omega')}{\omega'^2-\omega^2}\,\mathrm{d}\omega' \qquad [11]$$

and from the complex dielectric function, the optical dispersion and absorption properties follow. One may finally recall that since most semiconductors of interest (e.g., those in **Table 1**) are cubic, their optical response is isotropic as assumed here.

Interband Transitions

The optical absorption in crystals is mostly due to interband transitions, that is, to the promotion of electrons from occupied valence bands to empty conduction bands. Because of the energy conserving δ-function in [10], a singular behavior of the response functions is obtained in the vicinity of the critical points of the Brillouin zone (van Hove singularities), where for a given couple of conduction and valence bands

$$\boldsymbol{\nabla}_k(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k})) = 0 \qquad [12]$$

In fact, near a critical point it is often a good approximation to consider the transition dipole matrix element as a constant. Then, the integral in [10] is simply proportional to the joint density of states $J(\omega)$:

$$J(\omega) = \int\frac{2\mathrm{d}^3k}{(2\pi)^3}\,\delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) = \frac{2}{(2\pi)^3}\int_{S(\omega)}\frac{\mathrm{d}^2k}{|\nabla_k(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}))|} \qquad [13]$$

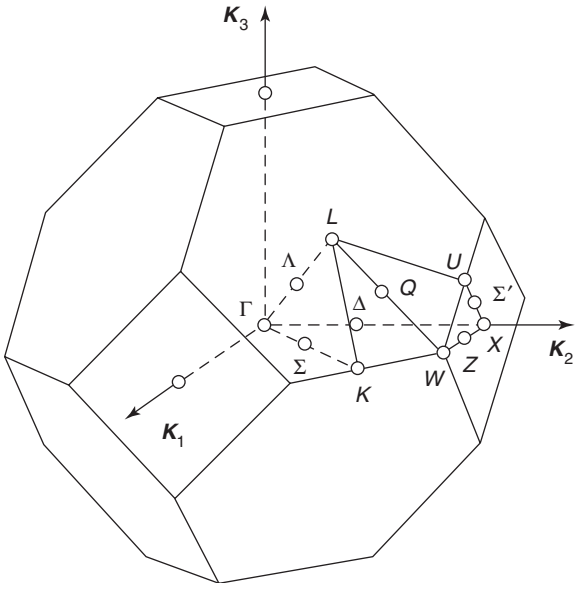


Figure 1 Brillouin zone of cubic semiconductors (f.c.c. lattice); lines and points of high symmetry are indicated.

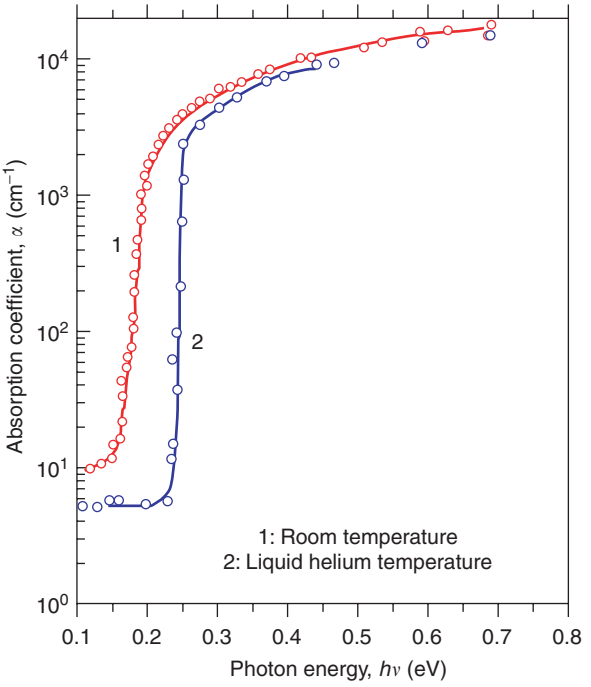


Figure 2 Absorption edge of InSb. (From the database NSM maintained by the Ioffe Institute.)

where the two-dimensional integral is over a surface $S(\omega)$ in $\boldsymbol{k}$-space for which $E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) = \hbar\omega$.

The location of critical points in the Brilluoin zone is mainly (but not completely) determined by symmetry considerations. For cubic semiconductors, the points Γ, X, L, and W in the Brilluoin zone (see **Figure 1**) are always critical points; other critical points usually occur along lines or planes of symmetry, but critical points at a generic wave vector $\boldsymbol{k}$ cannot be excluded. van Hove singularities have been found in the optical spectra of all semiconductors, and this has allowed the identification of the critical point transition peaks and the verification of the band structure.

A direct gap semiconductor is characterized by having both the conduction band minimum and the valence band maximum at the same wave vector $\boldsymbol{k}$ (typically the Γ point at $\boldsymbol{k} = 0$). When the optical transition between them is allowed, the absorption edge at a frequency corresponding to the bandgap E_g is due to the fact that the joint density of states vanishes for $\hbar\omega < E_g$ and increases sharply for $\hbar\omega \simeq E_g$. In particular, assuming a parabolic dispersion of the conduction and valence band states for $\hbar\omega \geqslant E_g$

$$J(\omega) \propto \sqrt{\hbar\omega - E_g} \qquad [14]$$

This result correctly describes the gross features of the absorption edge for allowed interband transitions in direct gap semiconductors. As an example, the absorption spectrum of InSb near its direct gap is shown in **Figure 2** exhibiting a well-pronounced absorption edge and becoming opaque at frequencies $\hbar\omega > E_g$ with an absorption coefficient $\alpha \approx 10^4$ cm^{-1}.

In **Figure 3**, the entire experimental excitation spectrum of Ge (the lowest gap of which is indirect, see below) is compared to that computed, as described above, from a pioneering band structure calculation based on empirical pseudopotentials. In this classical example, not only have the critical points been identified, but also the full band structure has been put to test obtaining a satisfactory agreement with the measured data. From the calculation of the dipole matrix elements at all points of the Brillouin zone and for all possible transitions to higher conduction bands, it can be concluded that most of the absorption is due to transitions to the lowest empty conduction bands. This fact can also be proved by verifying the f-sum rule in the optical range

$$\int_0^{\omega_M} \omega\varepsilon_2(\omega)\, d\omega \simeq \frac{2\pi^2 e^2 n_{\text{eff}}}{m} \qquad [15]$$

where in this case the frequency ω_M is just above the spectrum in the visible region, and n_{eff} is about equal to the number of valence electrons which take part in the optical transitions (8 per unit cell in elemental semiconductors such as Ge).

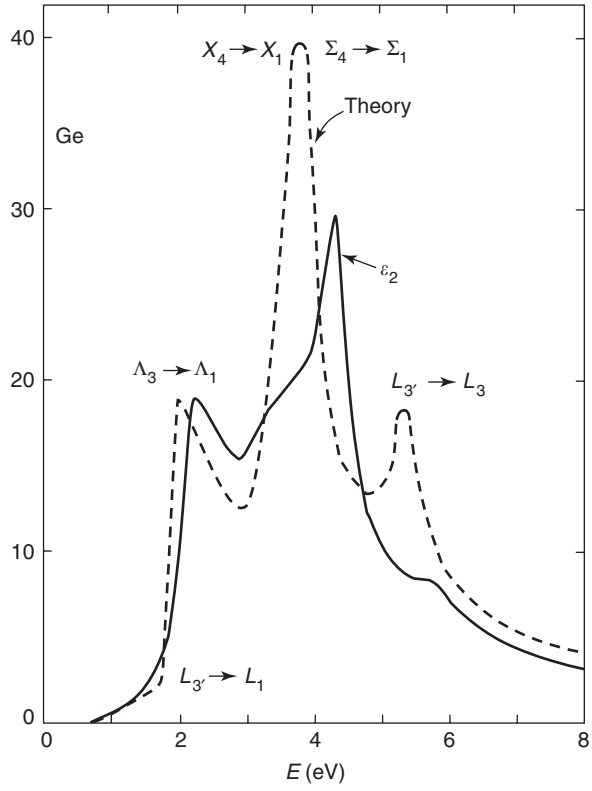


Figure 3 Imaginary part of the dielectric function of Ge: the solid line is experimental, the dashed one theoretical; optical transitions at critical points are identified. (From Brust D, Phillips JC, and Bassani F (1962) Critical points and ultraviolet reflectivity of semiconductors. *Physics Review Letters* 9: 94.)

Phonon-Assisted Transitions

Differently from the case of direct gap semiconductors such as GaAs and InSb, for Si and Ge, while the top of the valence band is still at the center of the Brillouin zone (Γ point), the minimum of the conduction band occurs away from Γ (close to the X point along the Δ line in Si, at the L point in Ge). This has been clearly observed by cyclotron resonance experiments, but it also shows in the optical properties, because the electron–phonon interaction makes an optical transition possible between different points of the Brillouin zone, the missing (pseudo) momentum being provided by the phonon emitted or adsorbed.

The probability of an indirect transition can be calculated in the second-order perturbation theory taking the perturbation interaction to be the sum of [6] and the electron–phonon interaction. Assuming constant matrix elements involving the states near the extrema which give the strongest contribution, one obtains

$$\varepsilon_2(\omega) \propto \left(n_{q_o v_o} + \frac{1}{2} \pm \frac{1}{2}\right) \int\int \frac{2\,\mathrm{d}\boldsymbol{k}_1\,\mathrm{d}\boldsymbol{k}_2}{(2\pi)^6} \times \delta(E_c(\boldsymbol{k}_1) - E_v(\boldsymbol{k}_2) \pm \hbar\omega_{v_o}(\boldsymbol{q}_o) - \hbar\omega) \quad [16]$$

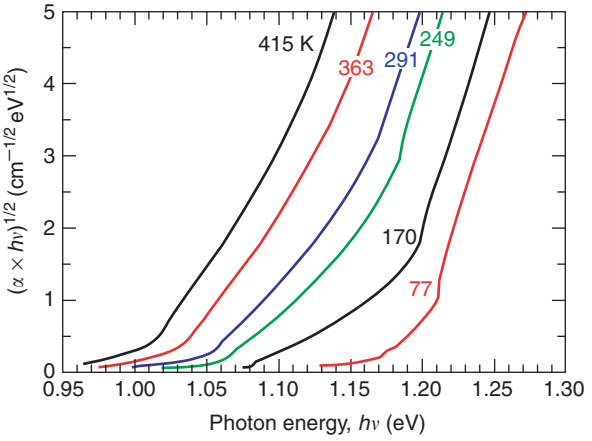


Figure 4 Indirect absorption edge of Si at various temperatures. (From the database NSM maintained by the Ioffe Institute.)

where "$\pm$" refers to the creation or destruction of a phonon of mode v_o with the wave vector $\boldsymbol{q}_o$, which connects the valence band maximum and the conduction band minimum.

The summation over the final electronic states above the edge ($\hbar\omega_o$) gives an energy dependence which is different from that of direct transitions, and precisely

$$\begin{aligned} &\varepsilon_2(\omega) = 0 \quad \text{when} \quad \omega < \omega_o \\ &\varepsilon_2(\omega) \propto (\omega - \omega_o)^2 \quad \text{when} \quad \omega > \omega_o \end{aligned} \quad [17]$$

Furthermore, different edges occur in correspondence to each phonon, and their values depend on the fact that the phonons are absorbed or emitted. At low temperature, one can only create phonons, and consequently the edge occurs at higher energy

$$\hbar\omega_o = E_g + \hbar\omega_{v_o}(\boldsymbol{q}_o) \quad [18]$$

At higher temperature, phonons can also be absorbed with a temperature-dependent probability given by the Bose–Einstein factor, and the edge is shifted to lower energy

$$\hbar\omega_o = E_g - \hbar\omega_{v_o}(\boldsymbol{q}_o) \quad [19]$$

Indirect transitions are responsible for the weak absorption tail in Si and Ge, which reflects their conduction band structures. **Figure 4** shows the absorption tail of Si; a fit to the profile of the indirect edge gives the frequency of the phonons which make the transition possible. At higher energies, when allowed direct transitions set in, the contribution of phonon-assisted indirect transitions to the absorption is negligible.

Excitons and Polaritons

The discussion above neglects the attractive interaction between the electron promoted into the conduction band and the hole left behind in the valence band. However, their motion is correlated and, in particular, they can be bound into the Wannier–Mott excitonic states with an energy below that of the single-particle energy gap. Due to the large dielectric constant ($\varepsilon_o \approx 10$) and the small reduced mass ($\mu = (m_e m_h)/(m_e + m_h) \approx 0.1\, m_o$) typical of semiconductors, the exciton radius a_o is much larger than the lattice constant and the binding energy correspondingly small. In the first approximation, an exciton turns out to be a "rescaled hydrogen atom" with binding energies depending only on the principal quantum number n:

$$E_n = E_g - \frac{R^*}{n^2},$$
$$R^* = \frac{\mu e^4}{2\varepsilon_o^2 \hbar^2} = \frac{\hbar^2}{2\mu a_o^2} = \frac{e^2}{2\varepsilon_o a_o} \qquad [20]$$

the exciton Bohr radius a_o being larger than 1 nm and the effective Rydberg R^* being typically ~ 10 meV. Thus, the exciton absorption lines are typically prominent only at low temperatures, as shown for GaAs in **Figure 5**.

The optical resonances below the bandgap corresponding to the excitonic bound states can be described by the following complex dielectric constant:

$$\varepsilon(\omega) = \varepsilon_b + \sum_n \frac{f_n}{(E_n/\hbar)^2 - \omega^2 - i\gamma_n \omega} \qquad [21]$$

where ε_b is a background dielectric constant, γ_n a phenomenological line width (e.g., due to phonon scattering), and the oscillator strength f_n is proportional to the square of the dipole matrix element between valence and conduction band extrema and inversely proportional to the exciton volume ($f_n \propto (n a_o)^{-3}$).

In correspondence to the excitonic resonances, when damping processes are small compared to the exciton–photon coupling, the electromagnetic waves propagating in the crystal hybridize with the exciton polarization giving rise to polariton states, that is, the proper modes of Maxwell's equations with the dielectric function given in [21]. Their dispersion law is obtained from

$$\omega^2 = \frac{c^2}{\varepsilon(\omega)} k^2 \qquad [22]$$

giving two branches (upper and lower polaritons) accompanied by a region of anomalous dispersion with a strongly reduced group velocity, as shown for GaAs in **Figure 6**. In polar semiconductors, having strong infrared absorption lines due to optical phonons, phonon–polariton states are similarly formed.

Modulation Spectroscopy

In an experimental spectrum, for example, of the reflectivity, the van Hove singularities such as those shown in **Figure 3** might be difficult to identify. As an example, the optical reflectivity of Si in **Figure 7** exhibits quite a smooth behavior as a function of frequency. However, the possibility of employing a phase-sensitive detection technique in conjunction with a periodic modulation of an external perturbation improves, on the one hand, the signal-to-noise

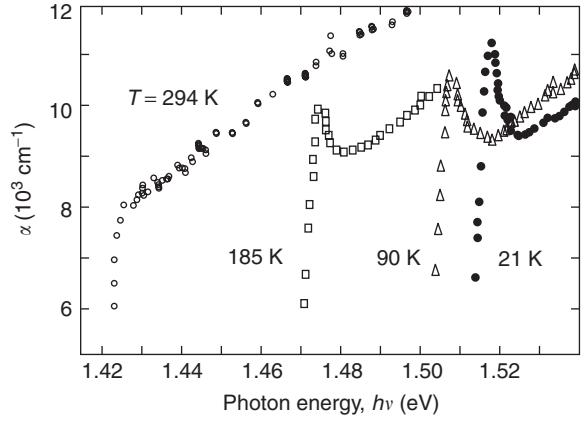


Figure 5 Excitonic absorption in GaAs, at room temperature the exciton peak is no longer observed. (From the database NSM maintained by the Ioffe Institute.)

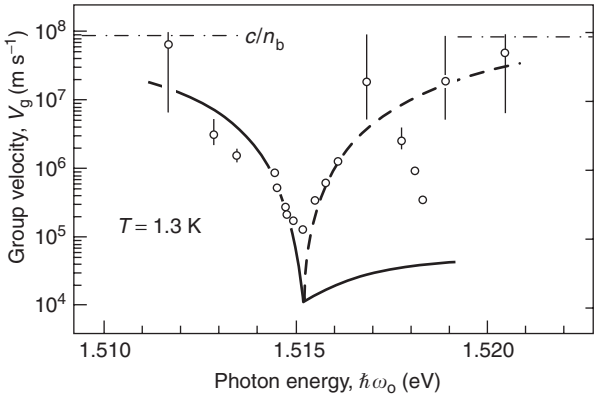


Figure 6 Exciton–polariton group velocity in GaAs, the solid and dashed lines are the theoretical values for the lower and upper polariton branch, respectively. (From Ulbrich RG and Fehrenbach GW (1979) Polariton wave packet propogation in the exciton resonance of a semiconductor. *Physics Review Letters* 43: 963.)

ratio by several orders of magnitude and, on the other hand, allows to enhance the sharp features of a spectrum with respect to a smoothly varying background. Modulation spectroscopy allows one to directly measure derivative spectra showing very clearly the stronger singularities in the first (and higher order) derivatives of, say, the reflectivity as a function of frequency.

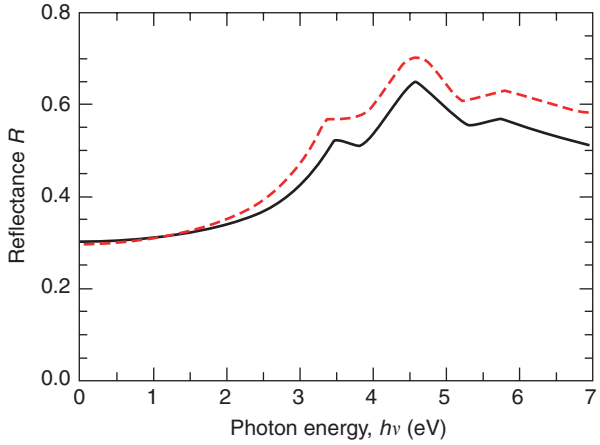


Figure 7 Reflectivity spectrum of Si: dashed line is experimental, solid line is theoretical (based on an empirical pseudopotential band structure calculation). (From the database NSM maintained by the Ioffe Institute.)

Thus, from the experimental point of view, much information on the properties of semiconductors can be obtained by measuring changes induced by externally applied perturbations. The latter can preserve the symmetry of the crystal (as for hydrostatic pressure or temperature changes) or lower it (as for uniaxial stress or electric and magnetic fields), in which case detailed information on the symmetry properties of the critical points of the electronic bands can be obtained. For instance, the reflectivity of Ge ~ 2.2 eV is dominated by interband transitions involving states along the Λ line in the Brillouin zone: uniaxial stress along the [0 0 1] direction does not split the degeneracy among the eight [1 1 1] valleys, whereas uniaxial stress along the [1 1 1] direction induces a difference between the two of them along the stress axis and the other six; in the latter case, a polarization dependence of the reflectivity is observed.

One of the most common techniques of modulation spectroscopy is electroreflectance, in which the

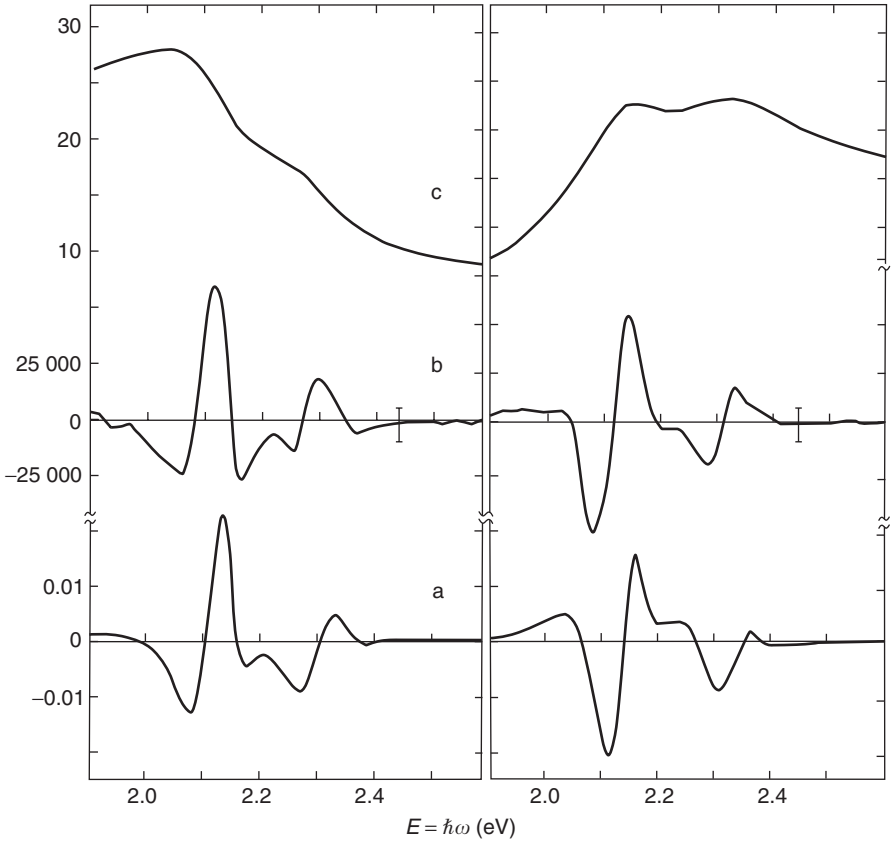


Figure 8 Real part (left panels) and imaginary part (right panels) of the dielectric function of Ge at the E_1 and $E_1 + \Delta_1$ gaps from (a) low-field electroreflectance compared to (b) the third-order derivatives as in eqn [23] calculated from (c) ellipsometric data. (From Aspnes DE (1972) Direct verification of the third-derivative nature of electroreflectance spectra. *Physics Review Letters* 28: 168.)

reflectivity is measured while applying an AC electric field parallel or perpendicular to the sample surface. The changes in optical constants measured by electroreflectance modulation spectroscopy have a very simple interpretation in the regime of low applied fields. In this case, as shown in **Figure 8**, the spectra are simply proportional to the third derivative of the respective optical constant with respect to frequency according to

$$\Delta\varepsilon(\omega) \propto \frac{e^2F^2}{\mu\hbar}\,\frac{1}{\omega^2}\,\frac{\mathrm{d}^3}{\mathrm{d}\omega^3}\,(\omega^2\varepsilon(\omega)) \qquad [23]$$

where F is the electric field and μ the reduced mass of the relevant interband transition. The electroreflectance spectra (**Figure 8a**) show much sharper features than those obtained by ellipsometry (**Figure 8c**).

See also: Elemental Semiconductors, Electronic States of; Optical Sum Rules and Kramers–Kronig Relations; Semiconductor Compounds and Alloys, Electronic States of; Semiconductor Devices; Semiconductors, Electronic Transitions in; Semiconductors, Optical and Infrared Properties of; Sum Rules and Kramers–Kronig Relations in Nonlinear Optics.

PACS: 78; 78.20.–e; 78.20.Ci; 78.40.–q; 71

Further Reading

Agranovich VM and Ginzburg VL (1984) *Crystal Optics with Spatial Dispersion and Excitons*. Berlin: Springer.
Bassani F and Parravicini P (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.
Cohen ML and Chelikowsky JR (1988) *Electronic Structure and Optical Properties of Semiconductors*. Berlin: Springer.
Klingshirn CF (1995) *Semiconductor Optics*. Berlin: Springer.
Yu PY and Cardona M (1995) *Fundamentals of Semiconductors*. Berlin: Springer.

Semiconductors, Electronic Transitions in

A Quattropani, Institute of Theoretical Physics, Lausanne, Switzerland
P Schwendimann, Defense Procurement, Bern, Switzerland

Introduction

A general introduction to semiconductors has been provided elsewhere in this encyclopedia. In order to theoretically describe electronic transitions induced in a semiconductor by a radiation field, it is first essential to model the energy eigenstates in a crystal. Therefore, this article begins with a brief review of the fundamental assumptions on which the description of the states of an ideal crystal is based. The effect of different types of impurities, such as acceptors, donors, and isoelectronic impurities, which enrich the spectrum of semiconducting crystals, is considered next. These are particularly important both from the fundamental and technological points of view. Finally, excitonic states, which appear in the crystal spectrum when electronic transitions between valence and conduction bands occur, are addressed. Once the full spectrum is understood, transitions induced by the radiation are discussed. These bear a strong analogy with atomic transitions, in particular concerning the selection rules. This article is restricted to radiation frequencies close to the fundamental gap, which, for a semiconductor, is typically of the order of a few electron volts.

Electronic Band Structures of Crystals: The Basic Approximations

A crystal consists of a very large number of atoms arranged in a regular lattice. The lattice structure of crystals implies that their properties are invariant under translations in space. In fact, the lattice of any crystal can be built from a fundamental cell, called unit cell, by spatial translations along the three fundamental axes.

In order to obtain the energy levels and the states of a solid, a number of approximations are introduced. First, the adiabatic (Born–Oppenheimer) approximation, according to which the electrons adiabatically follow the slower motion of the nuclei, allows separation of the equation of motion into two coupled problems for the nuclei and the electrons. The approximation relies on the fact that nuclei have a mass four orders of magnitude larger than the electron mass and, consequently, that the velocity of the motion of the nuclei around their equilibrium positions in the lattice is several orders of magnitude smaller than that of the electrons. The wave function of the crystal is then represented as the product of two wave functions, one for the nuclei and the other for the electrons, respectively. In the adiabatic approximation, the Schrödinger equation for electrons is solved in the potential of the nuclei, which

are frozen in their equilibrium positions. The equation for the nuclei is then averaged over the electron motion, which contributes only as a global potential term. From the viewpoint of transitions between electronic states, the equation for the electronic part of the wave function should be solved, which is a difficult many-electron problem. Therefore, a second approximation – the one-electron or Hartree–Fock approximation – is introduced. The wave function of N electrons is written as the antisymmetrized product of N one-electron wave functions, which is expressed as a determinant (the Slater determinant).

This approximation implies that the electron–electron correlations are neglected. The energies and states of electrons in the crystal are then calculated by a variational method using the Slater determinants as a basis. The corresponding Euler equations have the form of a nonlinear Schrödinger equation with two self-consistent potentials resulting from the electron–electron direct (Hartree) and exchange (Fock) interaction, respectively. The Hartree–Fock approach briefly sketched above leads to quite cumbersome calculations. Therefore, the so-called band approximation is introduced. In this case the basic equation describing the electrons in the crystal is

$$\left[\frac{p^2}{2m} + V(\boldsymbol{r})\right]\psi_n(\boldsymbol{k},\boldsymbol{r}) = E_n(\boldsymbol{k})\psi_n(\boldsymbol{k},\boldsymbol{r}) \qquad [1a]$$

where the nonlocal potential has been approximated by a local one. Here, $V(\boldsymbol{r})$ is a periodic potential, invariant under all symmetry operations of the crystal. It is the same for all states and is expressed as follows:

$$V(\boldsymbol{r}) = V_{\text{electron-lattice}} + V_{\text{Coulomb}} + V_{\text{exchange}} + V_{\text{correlation}} \qquad [1b]$$

In particular, the Coulomb and exchange components of $V(\boldsymbol{r})$ are calculated using the charge distribution of the ground state. The last term describes an approximation to the correlation energy, which is in general assumed to be proportional to a non-integral power of the charge density. In the section on excitons, the limitation of such an approximation is discussed. The solutions of eqn [1] with periodic boundary conditions are eigenfunctions of the Bloch type, that is, $\psi_n(\boldsymbol{k},\boldsymbol{r}) = e^{i\,\boldsymbol{k}\cdot\boldsymbol{r}}u_n(\boldsymbol{k},\boldsymbol{r})$, where $u_n(\boldsymbol{k},\boldsymbol{r})$ is a periodic function, with the periodicity of the lattice, and the wave vector $\boldsymbol{k}$ belongs to the first Brillouin zone. The eigenvalues $E_n(\boldsymbol{k})$ describe the energy bands of the crystal. In a semiconductor, in particular, two bands, called "valence" and "conduction" band, respectively, are well separated, in energy by the an energy gap E_g. Typically, the energy gap in a semiconductor varies between 0.5 eV and 2 eV at room temperature. A very useful model in this context is the two-band model, which neglects all bands, except the valence and conduction bands.

Impurity States

The theory sketched so far has considered the states of a perfect crystal. However, there are situations in which this picture must be refined, for example, when foreign atoms (impurities) are introduced into the crystal lattice. Two specific examples are the donor and acceptor impurities, and the isoelectronic impurities, which change the optical properties of the semiconductor without changing the carrier density.

Shallow Impurities

The effective-mass method for shallow impurities is illustrated in the case where a foreign atom (donor), which can release an electron in the conduction band of the host crystal, is introduced into the lattice. The translational periodicity of the lattice is broken and an additional potential originating from the presence of the impurity atom is introduced. In the one-electron approximation, the Schrödinger equation can be written as

$$\left[\frac{p^2}{2m} + V(\boldsymbol{r}) + U(\boldsymbol{r})\right]\Phi(\boldsymbol{r}) = E\Phi(\boldsymbol{r}) \qquad [2]$$

where $V(\boldsymbol{r})$ is, as usual, the potential of the perfect lattice and $U(\boldsymbol{r})$ is the additional potential due to the impurity. The very common case in semiconductor physics is considered where the electronic states are very weakly bound to the impurity, for example, in the case for donor impurities such as arsenic or phosphorus in germanium. The eigenvalue problem [2] is solved under the following assumptions.

1. The unperturbed conduction band $E_c(\boldsymbol{k})$ is non-degenerate, has an extremum at $\boldsymbol{k} = \boldsymbol{k}_0$, and is well separated from the other bands. The corresponding unperturbed Bloch state is denoted by $\psi_c(\boldsymbol{k},\boldsymbol{r}) = e^{i\,\boldsymbol{k}\cdot\boldsymbol{r}}u_c(\boldsymbol{k},\boldsymbol{r})$.
2. The perturbation potential $U(\boldsymbol{r})$ is weak with respect to E_g and its interband matrix element can be disregarded.
3. $U(\boldsymbol{r})$ is a slowly varying function over the crystal cell, that is, for the Fourier transform of the potential the inequality $\tilde{U}(\boldsymbol{k}) \gg \tilde{U}(\boldsymbol{k}+\boldsymbol{G})$, where $\boldsymbol{G}$ is a reciprocal lattice vector and $\boldsymbol{k}$ belongs to the first Brillouin zone, holds.

4. The periodic part of the Bloch functions is slowly varying as a function of $\boldsymbol{k}$, that is, $u_n(\boldsymbol{k},\boldsymbol{r}) \approx u_n(\boldsymbol{k}_0,\boldsymbol{r})$.

It should be noted that the usual perturbation theory cannot be used because the unperturbed eigenvalues $E_n(\boldsymbol{k})$ depend continuously on the wave vector $\boldsymbol{k}$. The solution of [2] is expanded in terms of unperturbed Bloch functions,

$$\Phi(\boldsymbol{r}) = \frac{V}{(2\pi)^3} \sum_n \int \varphi(\boldsymbol{k}) \psi_n(\boldsymbol{k},\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \qquad [3]$$

Substituting [3] into [2] and projecting the obtained equation on the unperturbed state $\psi_c(\boldsymbol{k},\boldsymbol{r})$, one obtains the following equation for the envelope function $\varphi(\boldsymbol{k})$:

$$[E_c(\boldsymbol{k}) - E]\varphi(\boldsymbol{k}) + \frac{V}{(2\pi)^3} \times \int \mathrm{d}\boldsymbol{k}' \langle \psi_c(\boldsymbol{k}) | U | \psi_c(\boldsymbol{k}') \rangle \varphi(\boldsymbol{k}') = 0 \qquad [4]$$

Substituting in the matrix element of [4], the approximation for the periodic part of the Bloch function, that is, $u_n(\boldsymbol{k},\boldsymbol{r}) \approx u_n(\boldsymbol{k}_0,\boldsymbol{r})$, as a consequence of the assumptions introduced above, one obtains $\langle \psi_c(\boldsymbol{k}) | U | \psi_c(\boldsymbol{k}') \rangle = \tilde{U}(\boldsymbol{k} - \boldsymbol{k}')$, which is the Fourier transform of the perturbing potential $U(\boldsymbol{r})$. For an isotropic and parabolic conduction band $E_c(\boldsymbol{k})$ with an effective mass m_c the Fourier transform of [4] takes the form

$$\left[-\frac{\hbar^2 k^2}{2m_c} + U(\boldsymbol{r}) \right] F(\boldsymbol{r}) = E\, F(\boldsymbol{r}) \qquad [5]$$

where $F(\boldsymbol{r})$ is the Fourier transform of the envelope function $\varphi(\boldsymbol{k})$.

For the special case of a screened Coulomb potential $U(\boldsymbol{r}) = -e^2/\varepsilon(0)r$, where $\varepsilon(0)$ is the static dielectric constant and $F(\boldsymbol{r})$ is a hydrogenic function. The binding energy in the ground state is the effective Rydberg $\mathrm{Ry}^* = \mathrm{Ry}\, m_c/m\varepsilon(0)^2$. For typical semiconductors, $\mathrm{Ry}^* \approx 10\,\mathrm{meV}$. The expression of the impurity wave function in terms of the envelope function is given by [3]. After expanding the Bloch function around the extremum $\boldsymbol{k}_0$ and performing the integration over $\boldsymbol{k}$, one obtains, in first order,

$$\Phi(\boldsymbol{r}) = F(\boldsymbol{r}) \psi_c(\boldsymbol{k}_0,\boldsymbol{r}) \qquad [6]$$

This result represents a useful introduction to the exciton problem, which is discussed in detail in the next section. In fact, the exciton, which results from the interaction of an electron in the conduction band and a hole in the valence band, is described through an extension of the approach outlined here.

Isoelectronic Impurities

Isoelectronic impurities are substitutional impurities whose valence state contains the same number of electrons as the valence state of the atom they replace. Therefore, in contrast to acceptors and donors, they do not contribute any additional charge to the host semiconductor. The isoelectronic impurities are classified by their solubility and by the presence of discrete levels. Impurities characterized by discrete levels and limited solubility are said to be of the first kind, whereas impurities that are widely soluble and continuously change the band edge are said to be of the second kind.

Examples of isoelectronic impurities of the first kind are GaP:N with N_P (nitrogen impurity in gallium phosphide, where N_P specifies that a nitrogen atom replaces a phosphorus atom), CdS:Te with Te_S, ZnTe:O with O_{Te}, and AgBr:I with I_{Br}. The last of these impurities is the basic component of a photographic film. These impurities are also called isoelectronic traps. $Ga_{1-x}In_xAs$ and $Ga_{1-x}Al_xAs$ are examples of isoelectronic impurities of the second kind. These ternary compounds are of special importance because their bandgap can have a range of values. Isoelectronic traps are used in order to enhance the optical properties of indirect-gap materials and are thus widely used in the development of LEDs and lasers.

The isoelectronic impurities are neutral impurities and the perturbing potential associated with is short range, in contrast to the long-range Coulomb potential of donors and acceptors. As a consequence, these impurities have a finite number of discrete levels. Isoelectronic impurities provide scattering and trapping centers for free carriers and excitons. Intuitively, the impurity traps a carrier (electron or hole) and the latter is then trapped by Coulomb interaction with this complex. A particularly interesting case is that of iodine, which traps a hole from an electron–hole pair, thus leaving the electron available for the creation of the latent image in photographic emulsions. The binding energy of the isoelectronic impurities is calculated in the framework of the so-called Koster–Slater formalism.

In contrast to the case of the electronic state of donors, where the perturbing potential is a slowly varying function over the crystal cell, for isoelectronic impurities the perturbing potential is short ranged. For example, in the case of AgBr, if a bromine atom is substituted by chlorine, only when the electron is very close to the impurity, it "feels" the

difference between bromine and chlorine atoms. The donor impurity model has to be modified in order to describe isoelectronic impurities. The starting point is eqn [4], as in the case of donor states. In the present case, the matrix element $\langle \psi_c(\boldsymbol{k})|U|\psi_c(\boldsymbol{k}')\rangle$ is, in first approximation, a constant independent of $\boldsymbol{k}$ and $\boldsymbol{k}'$ and is denoted by W_c; eqn [4] then reads as

$$[E_c(\boldsymbol{k}) - E]\varphi(\boldsymbol{k}) + W_c \frac{V}{N(2\pi)^3} \times \int d\boldsymbol{k}'\, \varphi(\boldsymbol{k}') = 0 \qquad [7]$$

Equation [7] for $[E_c(\boldsymbol{k}) - E] \neq 0$ is transformed into

$$1 + W_c/\varepsilon_c(E) = 0 \qquad [8a]$$

with

$$1/\varepsilon_c(E) = (2\pi)^{-3} \frac{V}{N} \int d\boldsymbol{k} \frac{1}{[E_c(\boldsymbol{k}) - E]} \qquad [8b]$$

Figure 1 plots $\varepsilon_c(E)$ for energies outside the conduction band. The horizontal lines indicate the constant W_c. The eigenvalue equation [8] has a solution only for sufficiently large $|W_c|$. In this case the energy E^* of the bound state is a solution of

$$\varepsilon_c(E) + W_c = 0 \qquad [9]$$

while the wave function of the bound state reads

$$\Phi^{(E^*)}(\boldsymbol{r}) = A \frac{V}{(2\pi)^3} \sum_n \int \frac{1}{[E_c(\boldsymbol{k}) - E^*]} \times \psi_n(\boldsymbol{k}, \boldsymbol{r})\, d\boldsymbol{k} \qquad [10]$$

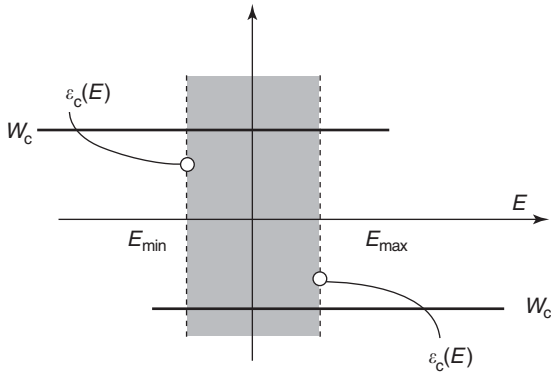


Figure 1 Graphic solution of eqn [8a] for two different values of the matrix element W_c.

For example, the above formulation yields a binding energy of iodine in AgBr to be 26 meV.

Excitons in Bulk Systems

Excitons appear when an electron excited from the valence band into the conduction band interacts with the positive excess charge (the "hole") left in the valence band after the transition. If one restricts consideration to a two-band model, the band index n indicates the valence and conduction bands, which are labeled by "v" and "c," respectively. The ground state of the crystal, in this approximation, is thus given by

$$\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \frac{1}{\sqrt{N!}} \mathrm{Det}\{\psi_v(\boldsymbol{k}_1, \boldsymbol{r}_1), \ldots, \psi_v(\boldsymbol{k}_N, \boldsymbol{r}_N)\} \qquad [11]$$

where all the electrons are assumed to occupy valence band states. It should be noted that the number of k-values in the Brillouin zone of the reciprocal space is equal to the number of electrons, N. Furthermore, the ground state of the system is always a totally symmetric state with respect to the symmetry group of the crystal lattice and, in particular, its wave vector is zero. For simplicity, the spin degree of freedom is neglected in the simplified treatment that follows.

In order to derive the excited states of the system, consider first "promoting" one electron from the valence to the conduction band. The relevant Slater determinant is

$$\Phi_{k_m, k'_m}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \frac{1}{\sqrt{N!}} \mathrm{Det}\{\psi_v(\boldsymbol{k}_1, \boldsymbol{r}_1) \cdots \psi_c(\boldsymbol{k}'_m, \boldsymbol{r}) \cdots \psi_v(\boldsymbol{k}_N, \boldsymbol{r}_N)\} \qquad [12]$$

where $\psi_c(\boldsymbol{k}'_m, \boldsymbol{r})$ is the wave function of an electron in the conduction band. The many-electron state has a total wave vector $\boldsymbol{k} = \boldsymbol{k}'_m - \boldsymbol{k}_m$, while its energy is given by $E_c(\boldsymbol{k}'_m) - E_v(\boldsymbol{k}_m)$. The state [12], however, is not a good first excited state. In fact, it has been assumed that the potential $V(\boldsymbol{r})$ appearing in [1] is constructed in a self-consistent way by minimizing the expectation value of the total Hamiltonian over the ground state [11]. When an electron is promoted from the valence band to the conduction band, the total charge density is modified with respect to the ground state and, consequently, the state [12] can be considered only as the zero-order approximation of the first excited state. The first-order state is obtained as a linear combination of Slater determinants of the

type [12] which diagonalize the total Hamiltonian. Explicitly, one can write, for an exciton trial function with zero wave vector $\boldsymbol{k}$,

$$\Phi_{\text{exc}} = \sum_{k'} \varphi(\boldsymbol{k}')\Phi_{\boldsymbol{k}',\boldsymbol{k}'} \qquad [13]$$

Normalization of [13] gives

$$\sum_{k'} |\varphi(\boldsymbol{k}')|^2 = 1 \qquad [14]$$

The coefficients $\varphi(\boldsymbol{k})$ are chosen by minimizing the expectation value of the total Hamiltonian. Using the same approximation as already used in the impurity problem, that is, $u_n(\boldsymbol{k},\boldsymbol{r}) \approx u_n(\boldsymbol{k}=0,\boldsymbol{r})$, the eigenvalue equation reads

$$[E_{\text{c}}(\boldsymbol{k}) - E_{\text{v}}(\boldsymbol{k}) - E]\varphi(\boldsymbol{k}) + \sum_{k''} \varphi(\boldsymbol{k}')V(\boldsymbol{k},\boldsymbol{k}') = 0 \qquad [15a]$$

with

$$V(\boldsymbol{k},\boldsymbol{k}') = \int \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2\, \psi_{\text{c}}^*(\boldsymbol{k},\boldsymbol{r}_1)\psi_{\text{v}}(\boldsymbol{k},\boldsymbol{r}_2) \times \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}\psi_{\text{c}}(\boldsymbol{k}',\boldsymbol{r}_1)\psi_{\text{v}}^*(\boldsymbol{k}',\boldsymbol{r}_2) \qquad [15b]$$

Here, the exchange contribution, which is small, has been neglected. Elementary solutions of [15] can only be found for two particular forms of the potential $V(\boldsymbol{k},\boldsymbol{k}') = V_1(\boldsymbol{k})V_2(\boldsymbol{k}')$ and $V(\boldsymbol{k},\boldsymbol{k}') = V(\boldsymbol{k}-\boldsymbol{k}')$. In the first case, the solution describes the localized Frenkel–Peierls excitons, while, in the second case, the solution describes delocalized Wannier excitons. Frenkel–Peierls excitons are found, for example, in solid rare gases or alkali halides. In the following, the discussion is restricted to Wannier excitons as found, for example, in copper oxide and in III–V compounds such as GaAs.

Theory of Wannier Excitons

In order to consider the case of a Wannier or weakly bound exciton in more detail, assume that its wave function extends over a region much larger than the lattice period. Equivalently, the coefficients $\varphi(\boldsymbol{k})$ in [13] can be assumed to involve only a small range of k-values around zero. This hypothesis depends, of course, on the material under consideration and may be justified only *a posteriori*. Therefore, it is convenient to use an effective-mass description of the energies and Bloch states involved in the exciton wave function, that is,

$$E_{\text{c}}(\boldsymbol{k}) = E_{\text{g}} + \frac{\hbar^2 k^2}{2m_{\text{c}}} \quad \text{and} \quad E_{\text{v}}(\boldsymbol{k}) = -\frac{\hbar^2 k^2}{2m_{\text{h}}} \qquad [16]$$

where E_{g} is the fundamental gap of the semiconductor, m_{c} and m_{h} are the effective masses associated with the conduction and valence band, respectively.

The Fourier transform of the integral eqn [15a] with the kernel $V(\boldsymbol{k},\boldsymbol{k}') = V(\boldsymbol{k}-\boldsymbol{k}') = e^2/|\boldsymbol{k}-\boldsymbol{k}'|^2$ yields a hydrogen-like Schrödinger equation for the function

$$F(\boldsymbol{r}) = \Omega^{-1/2} \sum_{k} \varphi(\boldsymbol{k})\exp[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}] \qquad [17]$$

analogous to the case of shallow impurities (see eqn [5]). In this equation, V/N is the normalization volume. In this case, the total wave vector of the exciton state is zero. Thus, under the assumptions introduced here, a simple picture of the exciton states is that of an electron which forms a hydrogenic atom with a positive charged hole in the valence band. The hole is ideally the vacancy left by the promoted electron. In reality, this derivation shows that the exciton is strictly a many-body effect, which goes even beyond the Hartree–Fock approximation. Nevertheless, the hydrogen-atom picture is very effective and is widely used. It requires, however, an important modification in order to be quantitatively predictive. In a realistic situation, the electric field between the two charges, which form the exciton, is screened by the presence of all the other electrons in the material. This screening originates from the polarization of the medium around the charges. The virtual transitions to the excited states of the system are responsible for this polarization. Only the transitions to the lower conduction band have been included in the present model. In principle, all the electron states belonging to different bands must be included to obtain the effect of screening. A quite surprising result is that this correction can be taken into account simply with the introduction of a dielectric constant $\varepsilon(0)$ screening the electron–hole Coulomb interaction term of the exciton hydrogen-like Schrödinger equation. In most of the small-gap semiconductors, the dielectric constant is rather large and takes values around 10. The two parameters that characterize a hydrogenic system that is, the effective Rydberg and the effective Bohr radius, can be written as

$$\text{Ry}^* = \mu e^4/2\varepsilon(0)^2\hbar^2, \qquad a_{\text{B}}^* = \hbar^2\varepsilon(0)/\mu e^2$$

where μ is the reduced mass of the electron–hole system. The electron–hole exchange interaction, which has been neglected so far, can be taken into account in first-order perturbation theory. This interaction is short-ranged and does not require a screening dielectric constant as in the case of the

Coulomb term. Due to the small effective masses of the conduction and valence bands in most small-gap semiconductors, and the large dielectric screening, the effective Rydberg is typically of the order of few meV only.

Optical Transitions

The interactions between radiation and electrons in the crystal may introduce transitions from the ground state to some excited electronic states. These transitions may occur to higher-energy bands for direct-gap semiconductors or the impurity or excitonic states. The threshold energy for direct interband transitions is given by the energy gap, while the transitions to impurity and exciton states are similar to atomic transitions and occur for energies smaller than the energy gap. Electronic transitions at optical frequencies are well described in the framework of the dipole approximation. The electron–field interaction has the form $e\boldsymbol{r} \cdot \boldsymbol{E}_0 \cos(\omega t)$ for a monochromatic field, where $e\boldsymbol{r}$ is the dipole operator. The contribution of the exciton or impurity states to the absorption coefficient is related to the transition probability from the ground state to the excited state induced by the radiation. The transition probability is a function of the exciting field amplitude $|\boldsymbol{E}_0|$ and is calculated with perturbation theory leading to a series expansion in powers of $|\boldsymbol{E}_0|^2$. The nth term in this expansion is denoted as n-photon process. Here only the one- and two-photon processes are considered. First consider the one-photon processes which occur under the resonance condition only. The transition probability depends on the envelope function $\varphi(\boldsymbol{k})$ and on the matrix element $\boldsymbol{M}(\boldsymbol{k}) = e\langle \psi_c(\boldsymbol{k}, \boldsymbol{r})|\boldsymbol{r}|\psi_v(\boldsymbol{k}, \boldsymbol{r})\rangle$ of the dipole operator between the valence and the conduction band wave functions taken at the same wave vector $\boldsymbol{k}$. For weakly bound impurities or Wannier excitons, the envelope function is peaked around $\boldsymbol{k} = 0$ and thus the transition probability is proportional to $|\boldsymbol{E}_0|^2|\boldsymbol{M}(\boldsymbol{k} = 0)|^2|F(\boldsymbol{r} = 0)|$, where $F(\boldsymbol{r})$ is defined in [5] and [17] for impurities and Wannier excitons, respectively. The excited states considered being hydrogen-like, the transitions are allowed to s-like states only. A typical absorption spectrum is shown in **Figure 2** for the case of GaAs.

The optical absorption and emission spectra of isoelectronic impurities show a characteristic series of peaks, which arise from the strong electron–phonon interaction. The strength of the interaction is due to the localization of the wave function of the electron in the impurity and to the strong gradient of the wave function in this region. The absorption and emission spectra have mirror-symmetric shapes and, in general, display a partial overlap. The isoelectronic trap ZnTe:O, whose absorption and emission spectra are specular, as indicated in **Figure 3**, represents an exception to this behavior.

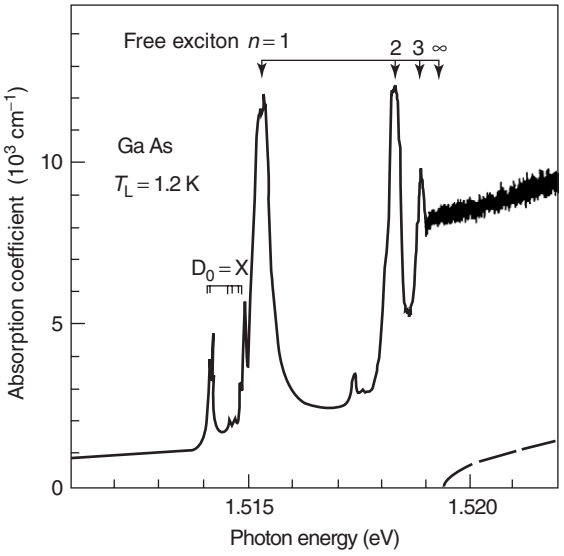


Figure 2 Absorption spectrum of GaAs at 1.2 K near the band edge. The $n = 1$, 2, 3 free-exciton peaks as well as impurity lines [D_0–X] of excitons bound to donors are shown. (Weisbuch C and Ulbrich R, unpublished, adapted from Ph.D. thesis of Weisbuch C, University of Paris 7, 1977.)

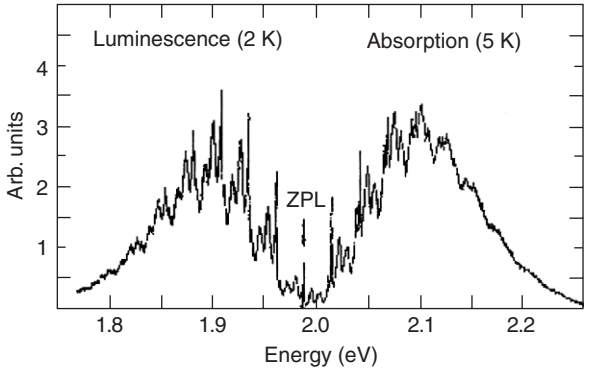


Figure 3 Absorption and luminescence spectra of CdTe:O at 2 K. ZPL indicates the zero phonon line. (Adapted from Schwendimann P, *et al.* (1992) Theoretical description of light amplification in ZnTe:O: a vibronic-laser material. *Physical Review B* 46: 7479–7485.)

When interband transitions are forbidden, that is, $\boldsymbol{M}(\boldsymbol{k} = 0) = 0$, it is necessary to take the next order in the wave vector $\boldsymbol{k}$ in the expansion of $\boldsymbol{M}(\boldsymbol{k})$ around $\boldsymbol{k} = 0$: $\boldsymbol{M}(\boldsymbol{k}) \propto \boldsymbol{k}$. In this case, the transition probability is proportional to

$$|\boldsymbol{E}_0|^2|\boldsymbol{k}|^2\left|\frac{\partial F(\boldsymbol{r})}{\partial \boldsymbol{r}}|_{r=0}\right|^2$$

and, therefore, the p-states can be reached. Indeed, the absorption spectrum of the yellow exciton in Cu_2O as presented in **Figure 4** does not contain the $n = 1$ transition line.

When one-photon transitions are forbidden, and for intense enough excitation field, two-photon transitions may occur, provided that the resonance condition $E^{\text{exciton}} = 2\hbar\omega$ is satisfied. The two-photon transition is also realized when two beams of different frequencies are used. In this case, the resonance condition $E^{\text{exciton}} = \hbar\omega_1 + \hbar\omega_2$ is fulfilled.

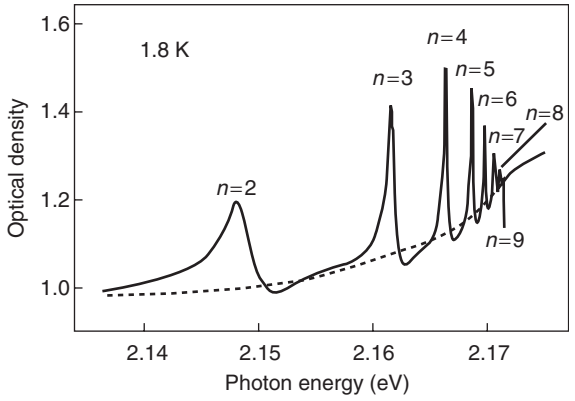


Figure 4 Absorption spectrum of the "yellow" exciton in Cu_2O at 1.8 K. Since transitions are allowed only to p-states of this exciton, the $n = 1$ state is forbidden. (Adapted form Shindo K, *et al.* (1974) Exciton-LO phonon scattering in Cu_2O. *Journal of the Physical Society of Japan* 36: 753–758.)

Cu_2O is particularly interesting because the one-photon transition can excite only the p-states, while the two-photon transition may excite s- and d-states, as indicated in **Figure 5**.

Finally, for strong photon–exciton coupling, the perturbation theory fails. The interacting photons and excitons have to be treated on the same footing, as indicated in the next section.

Polaritons

The absorption and emission spectra of semiconductor crystals have been considered in the above section. The present section outlines how the optical response of a semiconductor is modified when the propagation of the radiation field in the semiconductor is considered. Due to the presence of excitons, the propagation of radiation in a direct semiconductor shows a peculiar dispersive behavior. This is best understood through the solution of Maxwell's equations in presence of a linear polarization. The linear response of the exciton with eigenfrequencies $\omega(\boldsymbol{k})$ is given by the dielectric function $\varepsilon(\boldsymbol{k}, \omega) = \varepsilon_\infty + 4\pi f^2/(\omega^2(\boldsymbol{k}) - \omega^2)$, where ε_∞ is the background dielectric constant and f is the oscillator strength corresponding to the transition from the ground state to the 1s state. It is calculated in first-order perturbation theory corresponding to the one-photon processes discussed above. The fundamental

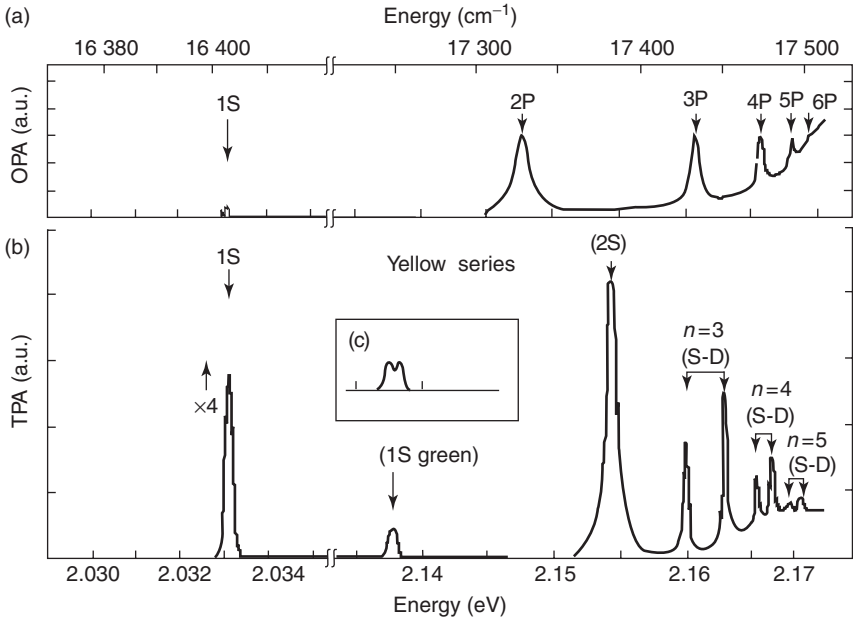


Figure 5 The one- and two-photon spectrum of Cu_2O. (a) "Yellow" exciton. One-photon data at 4.2 K. (b) "Yellow" exciton. Two-photon data at 4.5 K. (c) Splitting of the 1s green line in a magnetic field. (Adapted from Fröhlich D, *et al.* (1979) Assignment of the even-parity excitons in Cu_2O. *Physical Review Letters* 43: 1260.)

equation is then obtained via the Maxwell equations and reads $\nabla \times \nabla \times \boldsymbol{E} + \varepsilon(\boldsymbol{k}, \omega) k^2 \boldsymbol{E} = 0$. The solution of this equation for transverse modes, that is, $\nabla \cdot \boldsymbol{E} = 0$, is obtained when considering a plane-wave solution of the electric field,

$$\varepsilon(\boldsymbol{k}, \omega) = \varepsilon_\infty + \frac{\varepsilon(\boldsymbol{k}, 0) - \varepsilon_\infty}{1 - \omega^2/\omega^2(\boldsymbol{k})} = \left(\frac{c\boldsymbol{k}}{\omega}\right)^2 \quad [18]$$

For each value of k, this is a biquadratic equation whose solutions correspond to two transverse modes, called upper and lower polariton modes, propagating through the crystal. These modes, displayed in **Figure 6**, correspond to mixed exciton–photon states.

The upper-branch polariton (photon polariton) asymptotically approaches the dispersion of the free photon, whereas the lower-branch polariton (exciton polariton) similarly approaches the exciton dispersion. Polaritons are observed indirectly through nonlinear spectroscopic tools (Raman scattering, two-photon absorption). A direct observation of polariton photoluminescence is not very promising. Indeed, a polariton is a well-defined state of the crystal for any value of the wave vector, owing to its translational invariance. On the crystal surface, the translational invariance is broken and all polariton modes are mixed. Therefore, it is not possible to resolve the polariton luminescence as a function of the wave vector with sufficient accuracy.

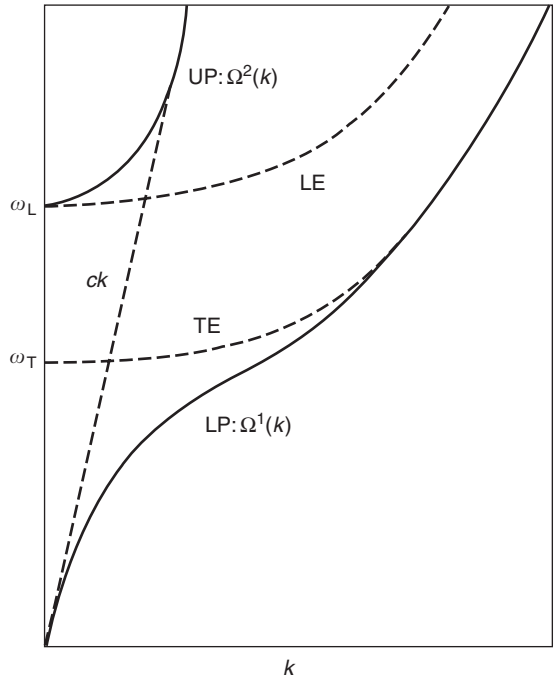


Figure 6 Schematic representation of the upper (UP) and lower (LP) polariton originated from Wannier excitons. The transverse (TE) and longitudinal (LE) excitons are also indicated as well as the radiation mode. (Adapted from Bassani F, *et al.* (1986) Microscopic quantum theory of exciton polaritons with spatial dispersion. *Il Nuovo Cimento* 7D: 700–716.)

The situation changes when lower-dimensional semiconductor structures like quantum wells, quantum wires, or quantum dots are considered. Consider the case of a quantum well. Because of the symmetry breaking in the direction of the growth of the quantum well, a single two-dimensional exciton, with a fixed in-plane momentum, acquires a radiative line width due to the coupling to the component of the three-dimensional electromagnetic field. This effect can be exploited by embedding the quantum well in a semiconductor microcavity consisting of two dielectric high-reflectivity mirrors. In this structure, only one photon mode couples to the exciton due to the large energy spacing between the cavity modes, and polaritons appear as a consequence of this interaction, which can be detected in the optical response. Microcavity polaritons form the basis of several important linear and nonlinear optical properties that have been the object of many studies in recent years.

See also: Elemental Semiconductors, Electronic States of; Excitons in Crystals; Optical Absorption and Reflectance; Polaritons; Semiconductor Compounds and Alloys, Electronic States of; Semiconductors, Impurity and Defect States in; Semiconductor Optics; Semiconductors, Optical and Infrared Properties of.

PACS: 71.22+l, 71.35.−y, 71.35Cc, 71.36.+c, 71.55.Eq, 71.55.Gs

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winstond.

Bassani F and Parravicini GP (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Bastard G (1989) *Wave Mechanics Applied to Semiconductor Heterostructures*. Paris: Les Editions de Physique.

Deveaud B, Quattropani A, and Schwendimann P (eds.) (2003) *Electron and Photon Confinement in Semiconductor Nanostructures*. Amsterdam: IOS Press.

Loudon R (1983) *The Quantum Theory of Light*. Oxford: Clarendon.

Yu PY and Cardona M (1996) *Fundamental in Semiconductors*. Berlin: Springer.

Nomenclature

a_B^*	effective Bohr radius (cm)
e	electron charge (esu $= g^{1/2} cm^{3/2} s^{-1}$)
$\boldsymbol{E}_0$	electric field (esu $= g^{1/2} cm^{-1/2} s^{-1}$)
$E_c(\boldsymbol{k})$	energy eigenvalue of the conduction band
E_g	semiconductor energy gap (eV)
$E_n(\boldsymbol{k})$	energy eigenvalue of the nth band
$E_v(\boldsymbol{k})$	energy eigenvalue of the valence band

f	oscillator strength (s^{-2})
$F(\boldsymbol{r})$	Fourier transform of the envelope function
$\boldsymbol{k}$	wave vector
m_c	conduction band effective mass (g)
m_v	valence band effective mass (g)
$\boldsymbol{p}$	moment
Ry	Rydberg (eV)
$U(\boldsymbol{r})$	impurity potential
$\tilde{U}(\boldsymbol{k})$	Fourier transform of $U(\boldsymbol{r})$
$u_c(\boldsymbol{k},\boldsymbol{r})$	periodic part of the conduction Bloch wave function
$u_n(\boldsymbol{k},\boldsymbol{r})$	periodic part of the Bloch wave function
$u_v(\boldsymbol{k},\boldsymbol{r})$	periodic part of the valence Bloch wave function
$V(\boldsymbol{r})$	periodic potential
$\varepsilon(0)$	dielectric constant
$\varepsilon(\boldsymbol{k},\omega)$	dielectric function
ε_∞	background dielectric function
$\Phi(r)$	wave function in presence of an impurity
$\varphi(\boldsymbol{k})$	envelope function
$\psi_c(\boldsymbol{k},\boldsymbol{r})$	wave function of the conduction band
$\psi_n(\boldsymbol{k},\boldsymbol{r})$	Bloch wave function of the nth band
$\psi_v(\boldsymbol{k},\boldsymbol{r})$	wave function of the valence band
μ	electron–hole reduced mass (g)

Semiconductors, General Properties

G Margaritondo, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

The main characteristics of semiconducting materials are briefly reviewed. First, semiconductors are treated in the context of the one-electron band structure defining the differences with respect to metallic conductors. The same framework is then used to present basic notions, such as the effective mass, the mobility, intrinsic and doped materials, and donors and acceptors. The fundamental mechanism of rectification by p–n junctions and a short description of the main properties of selected group IV and III–V semiconducting materials – including direct-gap and indirect-gap cases – are then presented. Finally, a short introduction of the optical properties of semiconductors and their use in light emitting devices and photon detectors is presented.

Introduction: Insulators and Semiconductors

All solids fall into broad classes as far as electrical conduction is concerned: metallic conductors, superconductors, insulators, semiconductors, etc. The insulating character can be caused by electron–electron interactions, for example, in the Mott insulators. In a one-electron theoretical framework, a solid is an insulator if the Fermi level falls within the forbidden gap that separates the valence band and the conduction band. At a temperature $T = 0\,\mathrm{K}$, no electron is found in the conduction band and the valence band is completely filled. Because of the symmetry of the valence-band states, the overall current created by the filled valence band is zero.

For $T > 0\,\mathrm{K}$, some electrons can be excited above the gap and into the conduction band leaving behind empty valence-band states. Such "free" electrons have energies close to those of empty states, so that low-energy excitations – required for electrical conduction – become possible. Furthermore, the empty valence-band states also contribute to conduction. The insulator thus becomes a conductor, although with a much lower conductivity than that of a metallic conductor.

For a finite temperature, the number of electrons thermally excited into the conduction band increases as the gap width E_g decreases. For E_g values below $\sim 2.5\,\mathrm{eV}$, the material is called a "semiconductor" rather than an insulator.

Semiconductors are the well-known protagonists of the modern microelectronics industry. This is primarily due to three factors: (1) the possibility of controlling their conduction properties by "doping" them with suitable impurities; (2) the possibility of electrically and chemically passivating their surface with ultrathin layers of insulating materials, thereby making it possible to integrate a huge number of microdevices into a single semiconductor slice; (3) other functions parallel to electrical conduction, such as light emission and absorption that can be exploited for specialized devices.

Doping is the most important of the three factors. In order to understand how it works, the conduction mechanism in all of its aspects must be analyzed first. This includes the role of both the conduction-band (free) electrons and the valence-band empty states.

Effective Mass, Electrons and Holes

Consider (**Figure 1a**) a simplified one-dimensional scheme of the band structure of a semiconductor. The electron energy E is plotted as a function of the

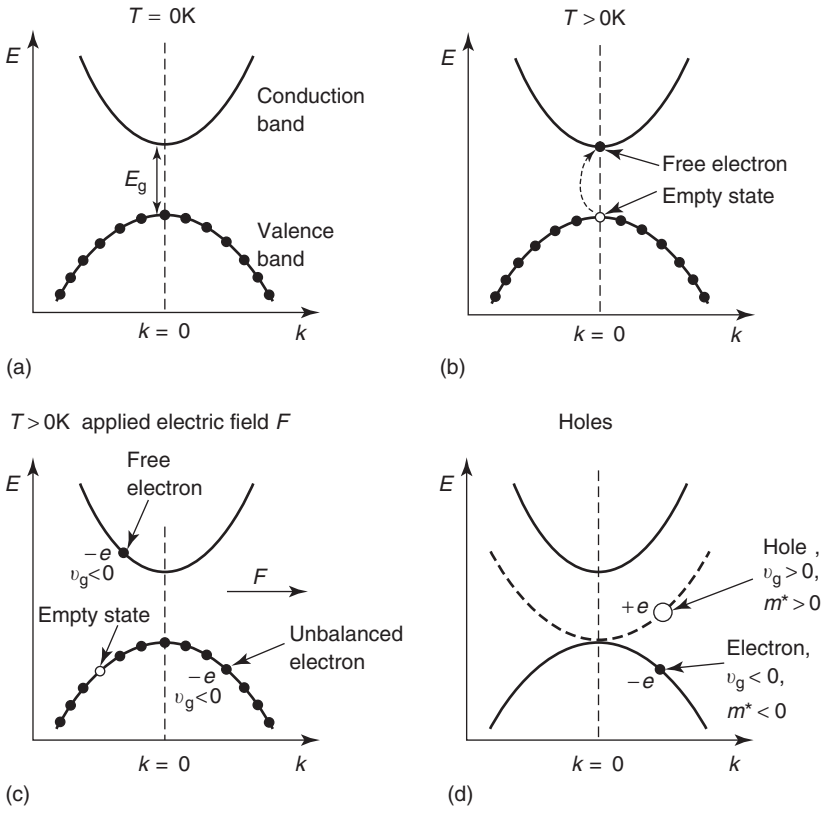


Figure 1 (a) Simplified one-dimensional scheme of the valence and conduction bands of a semiconductor: E is the electron energy and k is the wave number. At $T = 0\,\mathrm{K}$, the valence band is completely filled, the conduction band is empty, and there is no conduction. (b) Thermal excitation of an electron to the conduction band leaving an empty state in the valence band. (c) An external electric field F is applied in the positive direction of k: the conduction-band electron is subject to a negative force $-eF$ that causes a decrease in the group velocity $v_g \propto \mathrm{d}E/\mathrm{d}k$ and, therefore, a shift of the electron toward the negative side of the k-axis. The electron has negative charge and negative v_g; therefore, its contribution to the current is in the positive direction as the electric field F. In parallel, all electrons in the valence band are subject to the force $-eF$. Since their effective mass $m^* \propto 1/(\mathrm{d}^2E/\mathrm{d}k^2)$ is negative, v_g increases, and the electrons in the valence band shift in the negative k-direction, bringing with them the empty state. The "unbalanced" electron symmetric to the empty state has negative charge and negative group velocity; thus, it again produces current in the same direction as the electric field. (d) The effects of the empty valence-band state can be visualized in terms of a "hole" with positive charge $+e$, positive effective mass $-m^*$, and positive group velocity $|v_g|$: the dashed line shows its dispersion curve.

k-vector (in one dimension, the wave number k); note that $\hbar k$ is the momentum p or, more precisely, the "crystal momentum" in the Bloch theorem framework. It can be seen that at $T = 0\,\mathrm{K}$, the valence band is completely filled and the conduction band empty, corresponding to no conduction.

At $T > 0\,\mathrm{K}$, some electrons are thermally promoted into the conduction band. **Figure 1b**, for simplicity, shows only one of such electrons and the corresponding empty state in the valence band. **Figure 1c** shows the corresponding situation when an external electric field F is applied in the positive direction of k. The conduction-band electron is subject to a force $-eF$, moving to a negative value of k.

The effect of such a force can be treated based on the fundamental equation of mechanics,

$$-eF = \hbar \frac{\mathrm{d}k}{\mathrm{d}t} \qquad [1]$$

and based on the definition of group velocity,

$$v_g = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k} \qquad [2]$$

Equation [2] gives, in fact,

$$\begin{aligned}\mathrm{d}v_g/\mathrm{d}t &= (1/\hbar)(\mathrm{d}^2E/(\mathrm{d}k\ \mathrm{d}t)) \\ &= (1/\hbar)(\mathrm{d}^2E/\mathrm{d}k^2)(\mathrm{d}k/\mathrm{d}t)\end{aligned}$$

Combined with eqn [1], this becomes

$$-eF = m^*\frac{\mathrm{d}v_g}{\mathrm{d}t} \qquad [3]$$

that is, an equation of Newtonian form except that the electron mass is replaced by an "effective mass"

m^* defined as

$$m^* = \frac{\hbar^2}{(\mathrm{d}^2E/\mathrm{d}k^2)} \qquad [4]$$

In first approximation, the $E(k)$ curve near the minimum of the conduction band has a parabolic form and its second derivative is a positive constant, so that m^* is nearly constant as the real electron mass although different in value.

According to eqn [3], the effect of the force $-eF$ is a decrease of v_g that corresponds indeed to a shift of the electron toward the negative side of the k-axis, as seen in **Figure 1c**. The shift does not continue indefinitely since at a certain point, the effects of the external electric field are countered by the scattering mechanisms responsible for resistivity.

The contribution of the conduction-band electron to the overall current equals its charge $-e$ multiplied by its group velocity v_g. The electron in **Figure 1c** has negative charge and negative group velocity; therefore, its contribution to the current is in the positive direction – the same direction as the electric field F.

Consider now the empty valence-band state in **Figure 1b**. When completely filled, the valence band does not contribute to the current since the electrons moving in one direction are balanced by those moving in the symmetrically opposite direction. The $k = 0$ empty state in **Figure 1b** does not change this situation.

When an external electric field is applied (**Figure 1c**), all electrons in the valence band are subject to the force $-eF$. The effects are again described by eqn [3]. Note that near a maximum of the $E(k)$ curve, the effective mass (eqn [4]) is "negative" rather than positive. Thus, a negative force $-eF$ causes a positive derivative $\mathrm{d}v_g/\mathrm{d}t$ and an increase of the group velocity $v_g \propto \mathrm{d}E/\mathrm{d}k$. This means that all the electronic states in the valence band shift in the negative direction of the k-axis and bring the empty state with them as shown in **Figure 1c**.

The valence-band electron symmetric to the empty state is no longer balanced and produces a net current. This "unbalanced" electron has negative charge and negative group velocity; thus, it produces a positive current, again in the direction of the electric field as the conduction-band electron. The overall current is given by the sum of two contributions in the same direction, one from the conduction band and the other from empty states in the valence band.

The effects of the empty valence-band states can be treated in terms of the so-called "holes." This notion is suggested by the fact that the "unbalanced electron" in **Figure 1c** – with negative charge $-e$, negative m^*, and negative v_g – has the same effect on the current as a hypothetical particle ("hole") with positive charge $+e$, positive effective mass $-m^*$, and positive group velocity $-|v_g|$. The dispersion curve for this "hole" is shown in **Figure 1d**. The behavior of holes is easier to visualize than that of unbalanced electrons: the "hole" picture is, therefore, universally adopted when describing semiconductor properties.

In a more realistic picture, the top of the valence band and the bottom of the conduction band correspond to a maximum and a minimum of the three-dimensional surface $E(\boldsymbol{k})$, where $\boldsymbol{k}$ is the k-vector. The effective mass is not a scalar quantity but a tensor. Furthermore, several $E(\boldsymbol{k})$ surfaces can be found near the valence-band top or the conduction-band bottom, corresponding to different effective masses of the corresponding electrons and holes.

Mobility

The previous discussion leads to the following form of the total current density created in a semiconductor by an external electric field F:

$$j = -nev_e + pev_h \qquad [5]$$

where n and p are the free-electron density and the hole density; v_e is the average free-electron group velocity caused by the combined effect of the applied field F and of the scattering mechanisms responsible for the resistivity (i.e., scattering by impurities or defects and lattice vibrations); and v_h is the corresponding average group velocity for holes. Note that v_e and v_h are opposite to each other; thus, the two terms on the right-hand side of eqn [5] are in the same direction.

In first approximation, both v_e and v_h are proportional to the applied field F:

$$v_e = \mu_e F; \quad v_h = \mu_h F \qquad [6]$$

The parameters μ_e and μ_h are called "electron mobility" and "hole mobility." In a three-dimensional picture, the mobilities can have directional character.

The combination of eqns [5] and [6] gives

$$j = (-ne\mu_e + pe\mu_h)F \qquad [7]$$

The term in parentheses corresponds to the conductivity. Equation [7] implies two basic differences between a semiconductor and a simple metal. The first is the presence of two contributions to the conductivity, one from free electrons and the other from holes. The second difference is the effect of temperature. In a simple metal, the temperature primarily affects the mobility: as the temperature increases, the mobility decreases and so does the conductivity. In a

semiconductor, the temperature affects both the mobilities and the carrier densities n and p. The second effect dominates, producing an increase of the conductivity as the temperature increases.

Doping

What determines the free-electron concentration n and the hole concentration p in a semiconductor? Assume, for now, that the forbidden gap contains no impurity-related states. Each hole corresponds to an empty state in the valence-band electron that, in turn, corresponds to a free electron in the conduction band. This implies that $n = p$.

The free-electron concentration n is determined by the number of occupied states in the conduction band, determined by the Fermi–Dirac distribution and specifically by the position of the Fermi level E_F within the gap. Similarly, the position of E_F determines the hole concentration p. The condition $n = p$ requires E_F to be not too far from the middle of the gap (**Figure 2a**).

It can be shown that under rather general conditions (law of mass action), the product np is a function of the temperature, independent of the position of E_F in the gap. The conductivity of a semiconductor depends on the total concentration of available carriers, $n + p$. For a given temperature, np =constant, and the case $n = p$ corresponds to the minimum value of $n + p$.

Therefore, to increase the total carrier concentration, the condition $n = p$ must be violated. This can be done by doping the material with impurity atoms.

Consider, for example (**Figure 2b**), arsenic impurities in a silicon crystal. The chemical bonds in the pure crystal are formed by sp^3 hybridization of four $3s$ and $3p$ valence electrons for each atom. An atom forms chemical bonds with four nearest neighbors; thus, each bond involves two electrons, one for each atom. The molecular bonding states of these electrons form the valence band, whereas the antibonding states form the conduction band.

Suppose now that one silicon atom is replaced by an arsenic atom with five rather than four valence electrons (**Figure 2b**). Four of them form chemical bonds with neighboring silicon atoms and contribute to the conduction- and valence-band states. The fifth electron is not engaged in chemical bonding: its state

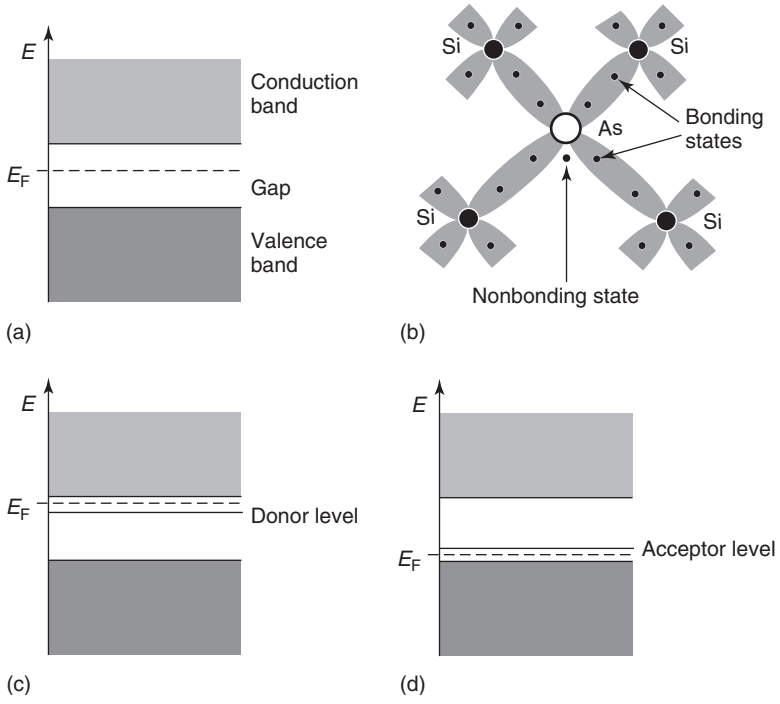


Figure 2 (a) In a semiconductor free from impurities, the condition $n = p$ implies that the Fermi level E_F is close to the middle of the gap. (b) Chemical bonds for an arsenic impurity in a silicon lattice. In pure Si, each bond engages two electrons, one per atom: all four valence electrons of a tetrahedrally coordinated Si atom are thus engaged. Arsenic has five valence electrons and the fifth one is not engaged in chemical bonds. (c) The corresponding nonbonding arsenic state has an energy level inside the gap and near the bottom of the conduction band. Electrons can be thermally promoted from this level to the conduction band: arsenic is thus a "donor" of electrons. (d) Symmetrically, a three-valent impurity, such as gallium, has an energy level near the top of the silicon valence band and can inject a large number of holes in the valence band acting as an "acceptor."

is nonbonding and its energy is intermediate between the bonding and antibonding states, that is, it falls in the forbidden gap separating the valence band from the conduction band (**Figure 2c**).

Because of the nonbonding character, this electron can be easily freed from the impurity and promoted to the conduction band. This implies that the impurity energy level is close to the bottom of the conduction band, as seen in **Figure 2c**. Thermal excitation allows the impurity to "donate" electrons to the conduction band, so that the arsenic in silicon is called a "donor" impurity.

Note that at $T = 0\,\mathrm{K}$, there is no thermal excitation at all: the donor electron states are all filled and the conduction band states are empty. This implies (**Figure 2c**) that the Fermi level is no longer close to the midgap but between the donor impurity level and the bottom of the conduction band.

Doping by donors leads to a large concentration of concentration n of impurity-donated free electrons; the $n = p$ rule is then broken and the total carrier density $n + p$ is larger than the minimum value. The material is called an "n-type" semiconductor, whereas an impurity–free sample with $n = p$ is called an "intrinsic" semiconductor.

A symmetric mechanism enables three-valent ("acceptor") impurities, such as gallium, to inject a large number of holes in the silicon valence band, producing a "p-type" sample. In this case (**Figure 2d**), the impurity has an empty nonbonding state near the top of the valence band that can capture thermally excited electrons from the valence band. This produces empty states in the valence band and, therefore, increases the hole density p. At $T = 0\,\mathrm{K}$, the Fermi level is found to be between the top of the valence band and the acceptor impurity level.

More sophisticated mechanism enables impurities with the same valence as the replaced atoms to also act as donors or acceptors. One good example of these "isoelectronic" impurities is provided by five-valent nitrogen replacing five-valent arsenic in GaAs.

The presence of impurities in a semiconductor is difficult to avoid, so that the samples naturally tend to be doped n-type or p-type. Suitable technologies must be used to control the doping level and change its character from p to n and vice versa. Particularly difficult is to obtain samples with low impurity concentrations so that $n = p$, the total carrier concentration $n + p$ is low, and the material is insulating rather than conducting. Practically speaking, the $n = p$ condition can be achieved by seeking equal concentrations of donors and acceptors rather than low donor and acceptor concentrations. A material of this kind is said to be "compensated."

Current Rectification by a *p–n* Junction

The control of transport properties by doping is a key ingredient in the manufacture of semiconductor devices. A simple example is qualitatively discussed here: the p–n junction, which is a fundamental building block of many devices. A p–n junction consists of a semiconductor sample with two different regions, one n-type and the other p-type (**Figure 3a**). In practice, such a system can be fabricated starting from a sample of one type and then changing the doping of part of it with a suitable procedure such as gas-phase exposure.

As seen in **Figure 3b**, the band diagram of the system includes an n-type region (similar to **Figure 2c**) and a p-type region (similar to **Figure 2d**), separated by an intermediate area called the "depletion" region. Elementary thermodynamics requires the Fermi level (which corresponds to the free energy function) to be the same everywhere in the sample. As a consequence, in the depletion region, E_F tends to be close to the midgap and the situation is similar to that in **Figure 2a**. This corresponds to low conductivity because of the low total carrier density: the "depletion" of carriers justifies the name of the region.

Note that free electrons trying to cross from the n-region to the p-region face an energy barrier E_B that prevents them from doing so. Similarly, holes in the p-region cannot cross into the n-region because of a symmetric barrier.

Imagine now that an external voltage bias is applied to the junction trying to create a current. A bias in the "forward" direction (**Figure 3c**), positive on the p-region and negative on the n-region, decreases the E_B barrier, making it easier for electrons to overcome it: a current is thus created. A symmetric analysis leads to the same conclusion for holes. On the contrary, a "reverse" bias (**Figure 3d**) increases the barrier and makes it even more difficult for carriers to cross the junction so that no or very little current is created.

The overall effect corresponds to the rectifying diode behavior shown in **Figure 3e**. Under rather idealized hypotheses, this behavior can be described by the equation

$$I = I_0\left[\exp\left(\frac{eV}{\eta kT}\right) - 1\right] \qquad [8]$$

where I is the current and V the applied voltage bias, k the Boltzmann constant, and η the so-called "ideality factor" ($\eta = 1$ for the most ideal case); note that for a strong reverse (negative) bias, $I \sim -I_0$, the so-called "saturation current."

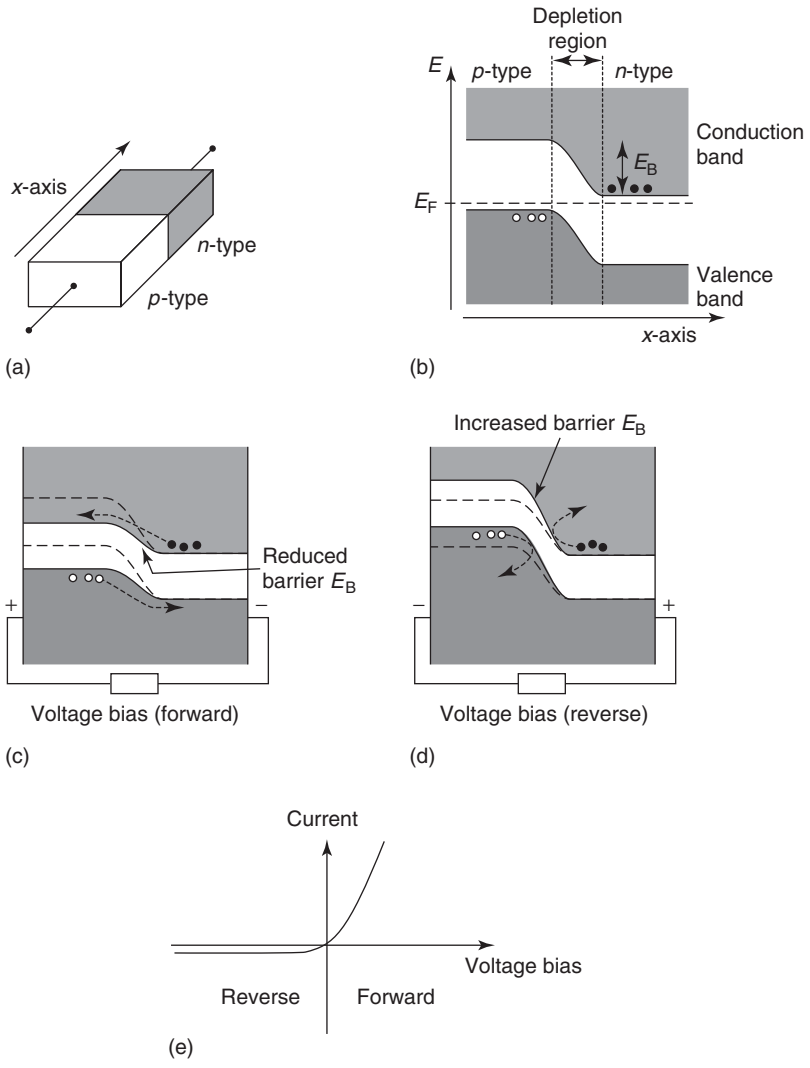


Figure 3 (a) A *p–n* junction consists of a semiconductor sample with two different regions, *n*-type and *p*-type. (b) In the corresponding band diagram, the *n*-region (similar to **Figure 2c**) and the *p*-region (similar to **Figure 2d**) are separated by a "depletion" region. There, E_F tends to be close to the midgap as in **Figure 2a**, corresponding to low conductivity. Free electrons trying to cross the depletion region are blocked by the energy barrier E_B; a symmetric conclusion is valid for holes. (c) An external voltage bias applied in the "forward" direction decreases the barrier E_B, making it easier for electrons and holes to overcome it and generating a current. (d) A "reverse" bias increases E_B and generates no current. (e) The corresponding typical current–voltage curve of a rectifying diode.

This description is, of course, oversimplified. However, it is suitable to understand the basic facts and, in particular, the crucial role of doping in the fabrication of devices. In modern microelectronics, other building-block devices are fundamentally important in addition to the *p–n* junction, in particular, the field-effect transistor.

Examples of Important Semiconductors

The most important semiconductors for industrial applications are the elemental group IV materials, such as silicon and germanium, and some III–V binary compounds, such as GaAs, GaP, InP, and GaN. Other classes of compounds include interesting semiconductors for specialized applications but none of them has a technological impact comparable to group IV or III–V materials. Furthermore, the microelectronics industry is dominated by the silicon-based technology, whose impact in market terms, is much stronger than all the other semiconductors put together. This is due to several factors – in particular passivation, crystal growth, and crystal refinement – that greatly facilitate silicon-based manufacturing and make it possible to reach advanced performance, extreme miniaturization, and massive integration.

Silicon and Germanium

Both these elemental semiconductors have the cubic structure of diamond shown at the top of **Figure 4**. Each atom is tetrahedrally bound to four nearest neighbors.

The band structures of the two materials are shown in the middle and bottom parts of **Figure 4**. Specifically, the three-dimensional dependence of the energy E on the k-vector is shown along two high-symmetry crystallographic directions of $\boldsymbol{k}$, $\langle 1\,1\,1 \rangle$ and $\langle 1\,0\,0 \rangle$.

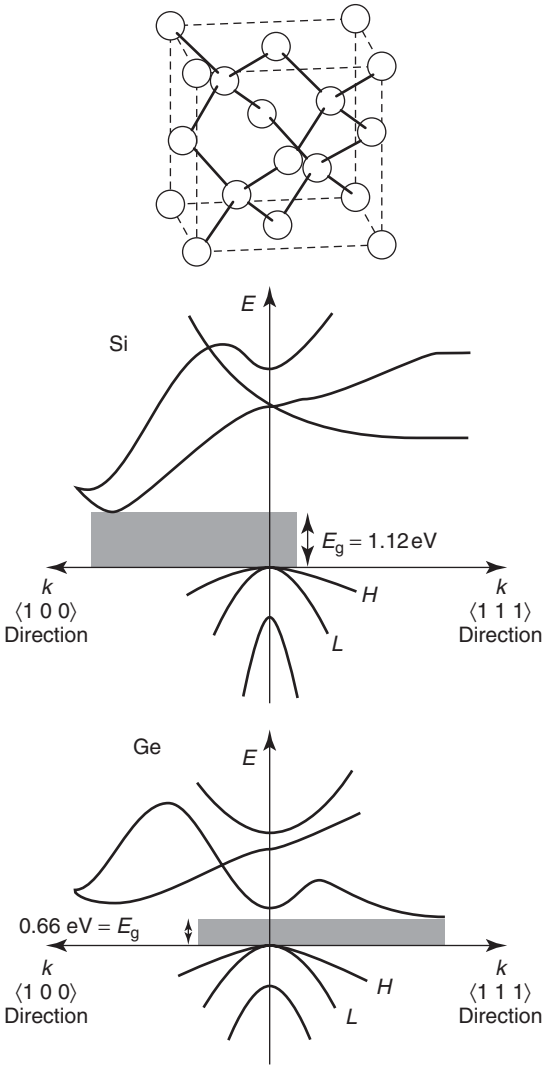


Figure 4 Top: the characteristic tetrahedrally coordinated cubic diamond crystal structure of silicon and germanium. Middle: band structure of silicon along the $\langle 1\,1\,1 \rangle$ and $\langle 1\,0\,0 \rangle$ directions of $\boldsymbol{k}$. Note that an "indirect" gap occurs when the conduction-band minimum is at a different $\boldsymbol{k}$-position than the valence-band maximum. Bottom: the band structure of germanium with its indirect gap. Also, note for both band structures the presence of a "heavy hole" band (H) and of a "light hole" band (L), with different curvatures and, therefore, different effective masses.

It is important to note that for both band structures, the bottom of the conduction band and the top of the valence band of germanium occur at two different values of the k-vector. The corresponding gaps are called "indirect," whereas, when the bottom of the conduction band and the top of the valence band occur at the same k-value, the gap is said to be "direct" as seen in the GaAs band structure in **Figure 5**.

The direct or indirect character of the gap strongly influences the physical properties of the material. For example, the excitation of electrons across the gap implies no change of the k-vector for a direct gap, whereas the contrary is true for an indirect gap. This also implies that the excitation across an indirect gap requires a change in the crystal momentum and therefore, a suitable third party, such as a lattice vibration, to guarantee the conservation of the crystal momentum.

It should also be noted that for both the Si and the Ge band structures, there are two $E(\boldsymbol{k})$ curves near the top of the valence band. Holes can belong to either curve and can have two different effective

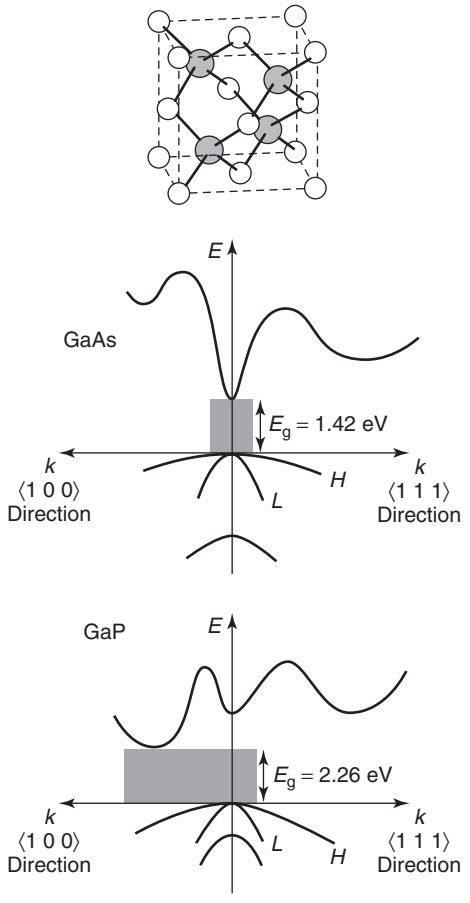


Figure 5 Top: the zinc blende crystal structure of many III–V semiconductors. Middle and bottom: the band structures of GaAs and GaP with their direct and indirect gaps, respectively.

masses. Equation [4] shows that the effective mass is proportional to the reciprocal of the curvature of $E(\boldsymbol{k})$. Therefore, a small curvature means a large effective mass and vice versa. Accordingly, the two valence-band $E(\boldsymbol{k})$ curves are called "heavy hole" band (H) and "light hole" band (L).

The following are the values of some important parameters of Si and Ge (the effective masses are expressed in terms of free electron mass m_e):

	Si	Ge
E_g (300 K)	1.12 eV	0.66 eV
Electron effective masses	0.98, 0.19	1.59, 0.08
Hole effective masses (heavy and light)	0.49, 0.16	0.33, 0.04

The two values for the electron effective masses correspond to the "longitudinal" and "transverse" effective mass. This distinction corresponds to the fact that the conduction-band minimum is not isotropic in three dimensions and that the second derivative in the denominator of eqn [4] is different in different directions. The longitudinal mass corresponds to the derivative along the crystallographic direction where the minimum occurs, whereas the transverse mass refers to directions perpendicular to it.

III–V Compounds

Materials in this class exhibit a wide spectrum of properties of interest for practical applications. **Figure 5** (top) shows the typical zinc blende crystal structure of many III–V compounds. This structure is closely related to that of diamond, with two atomic species rather than one.

The middle and bottom parts of **Figure 5** show the band structures of two particularly important III–V semiconductors, GaAs and GaP. There are some similarities with respect to the Si and Ge band structures of **Figure 4**, such as the heavy-hole band and the light-hole band. Note that the gap is direct in the case of GaAs and indirect for GaP.

There are some important parameters for GaAs, GaP, and for two other important III–V semiconductors, InP and GaN (in the zinc blende structure, GaN is also found in the wurtzite structure):

	GaAs	GaP	InP	GaN
E_g (300 K)	1.42 eV	2.26 eV	1.34	3.3 eV
Electron effective mass	0.06	0.22	0.08	0.13
Hole effective mass	0.51, 0.108	0.79, 0.14	0.6, 0.09	1.3, 0.19

Light Emission and Absorption

The presence of a forbidden gap in the band structure of a semiconductor has an immediate impact on its optical properties: without taking into account the impurities, electronic transitions cannot cause the emission or absorption of photons of energy smaller than E_g. Note, from the previous section, that the gaps of Si and Ge correspond to the energies of infrared photons, such as those of GaAs and InP. On the contrary, the GaP gap falls in the visible and the GaN gap near the border between the ultraviolet and visible regions.

Figure 6 shows the absorption coefficient of silicon plotted in the photon energy range near the gap-width value E_g. It is seen that the absorption has indeed a rapid increase at the E_g-threshold. Note that silicon is an indirect-gap material: the edge would be even sharper for a direct-gap semiconductor. The form of the absorption coefficient after the threshold can be justified in terms of the available initial and final states for the photon-excited transitions of electrons from the valence band to the conduction band.

The symmetric process of photon emission takes place if electrons are injected in the conduction band. The transition from the excited conduction-band states to unoccupied valence-band states can be visualized as the "recombination" of electrons and holes.

Real optical processes are often more complicated than suggested by the above picture. For example, this is a one-particle picture that includes no interaction between the electrons. In reality, the electrons do interact, in particular, the free electrons in the

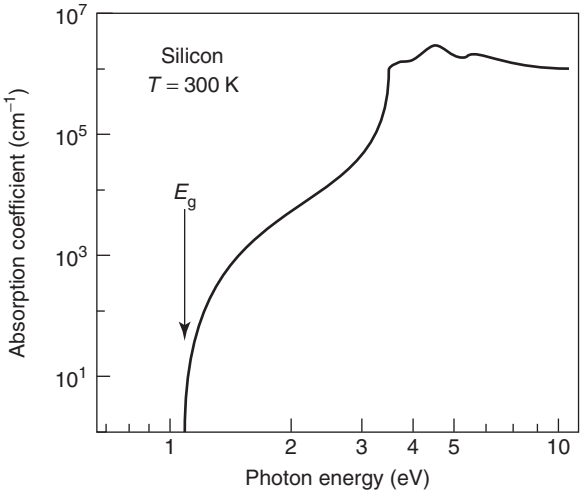


Figure 6 Absorption coefficient of silicon plotted in the photon energy range near the gap-width value E_g. Note the steep threshold related to E_g.

conduction band and the electrons in the valence band. This causes the so-called "excitonic effects."

The simplest case of the excitonic effects is the formation of "Wannier excitons," in which the interaction is visualized as the formation of bound states between free electrons in the conduction band and holes in the valence band. A Wannier exciton can thus be imagined as a hydrogen-atom system in which the proton is replaced by a much lighter hole. The Wannier electron–hole pair is sufficiently delocalized to treat the system in terms of electron and hole effective masses (the contrary is true for the opposite limit of localized "Frenkel" excitons). In the hydrogenic picture, the proton mass is replaced by the reduced-mass combination of the effective masses. Furthermore, the vacuum dielectric constant is replaced by the dielectric constant of the semiconducting material.

The excitation of an electron from the valence band can take place in two different ways: either by forming an exciton or by entirely freeing the electron into the conduction band. The latter situation corresponds to the ionization continuum of the hydrogenic exciton, whereas the former corresponds to its 1*s* ground level. Correspondingly, in the absorption spectrum, the "continuum" above the E_g edge is preceded by an exciton absorption line (or by a series of lines corresponding to different hydrogenic levels). The detection of excitonic lines requires sufficient spectral resolution and low temperature to avoid thermal broadening; otherwise, the excitonic effects are only observed as modifications of the one-electron absorption spectrum.

Figure 7a shows a typical excitonic line absorption spectrum taken on GaAs at low temperature. Note the temperature-induced shift of the E_g edge with respect to the values reported above.

The situation becomes more complicated when the gap is not direct as for GaAs, but indirect. As already mentioned, the optical transitions must be assisted by a suitable "third party" to guarantee the conservation of the (crystal) momentum. Lattice vibrations (i.e., phonons) can play such a role. **Figure 7b** shows the absorption edge of indirect-gap GaP with several features related to the intervention of specific phonons that absorb part of the photon energy and provide the necessary momentum.

Excitonic effects also affect the photon emission processes. Such processes are very important for the realization of "optoelectronic" devices, such as light-emitting diodes (LEDs) and semiconductor lasers. In most cases, in such devices the so-called "bound" excitons, those related to impurities, are the ones that are more significant.

Impurity levels in the forbidden gap allow the absorption and emission of photons of energy smaller

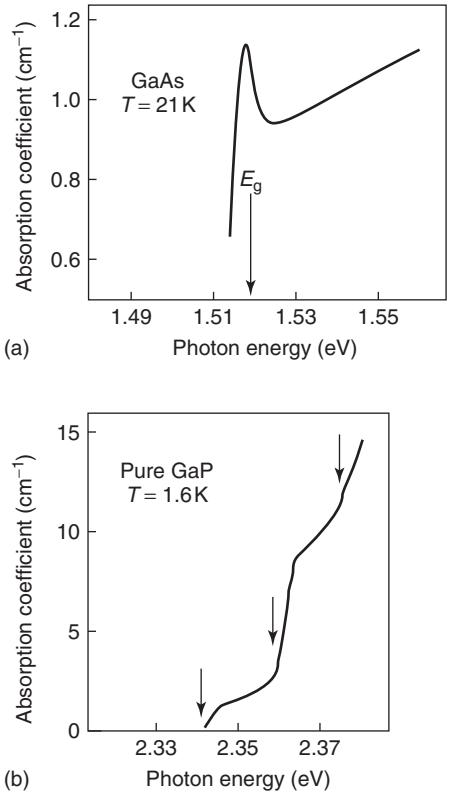


Figure 7 (a) The excitonic line in the low-temperature absorption spectrum of GaAs. (b) Absorption edge of indirect-gap GaP. The arrows mark spectral features related to specific phonons that guarantee momentum conservation.

than the gap. This has important repercussions on the corresponding optoelectronic devices. Consider, for example, the case of GaP: its room-temperature gap, 2.26 eV, would correspond to the "green" visible photons. The corresponding LEDs would produce light with excellent visibility since the peak of the human eye response occurs for such photons.

Unfortunately, the emission of photons from pure GaP is rather ineffective, since the indirect gap requires the intervention of phonons, and this leads to lower electron–hole recombination rates. Doped GaP is much more effective and is commonly used in LEDs. However, its emission no longer occurs in the green but at lower photon energies in the red, where the human eye is less efficient.

It is also possible to remove the effects of the GaP indirect gap by using an intermediate composition between GaP and direct-gap GaAs, that is, $GaAs_{1-x}P_x$. With a suitable choice of the concentration (the x-value), the material has a direct gap. However, the gap width is shifted from the "green" value of GaP toward the infrared value of GaAs and the emitted photons are once again at lower energies than the green.

From this brief discussion, it is clear that the ideal material for LEDs should have a very large gap and should be suitable for doping with different types of impurities. Such impurities would correspond to optical transitions at different photon energies smaller than the gap, ideally covering the entire visible spectrum. In this way, a single material could emit different colors and could be used in integrated devices for variable color emission.

Among the III–V semiconductors, the best candidate for this role is GaN. It is seen, indeed, that its gap falls near the visible–ultraviolet border. Doped with suitable impurities, GaN could emit blue light as well as light at lower photon energies.

The technology of GaN, unfortunately, is much less developed than that of GaAs or GaP (not to mention the extremely sophisticated technology of silicon). In recent years, however, substantial improvements have led to the fabrication of efficient GaN-based blue-emitting devices. Further improvements are foreseen so that the GaN technology is quite promising for versatile optoelectronic devices.

Figure 8a schematically shows a simple type of LED based on a forward-biased *p*–*n* junction. As noted already, the forward bias lowers the barriers for electrons and holes. Electrons from the *n*-region can cross the junction into the *p*-region, where they find a large density of holes. The electron–hole recombination causes the emission of photons.

Symmetrically, the reduced barrier for holes enables some holes to move from the *p*-region to the *n*-region. There, they find a large density of free electrons: once again, the electron–hole recombination can give rise to photon emission.

The *p*–*n* junction is also the building block of another important optoelectronic device: the light detector or photodiode. **Figure 8b** schematically shows how it operates. In this case, the *p*–*n* junction is reverse biased, so that normally very little current flows through it.

When a photon beam with photon energy larger than the gap width illuminates the junction, photon absorption can excite electrons into the conduction band of the *p*-type region. Some of these photo-injected free electrons reach the interface region where they find a suitable voltage to move into the *n*-type region. In this way, the photon beam produces a photocurrent that can reveal its arrival on the diode.

For solar cells, some photodiodes are optimized for the detection of a large part of the solar spectrum. They allow the direct conversion of solar energy into electric power, providing an interesting alternate source of energy of increasing importance.

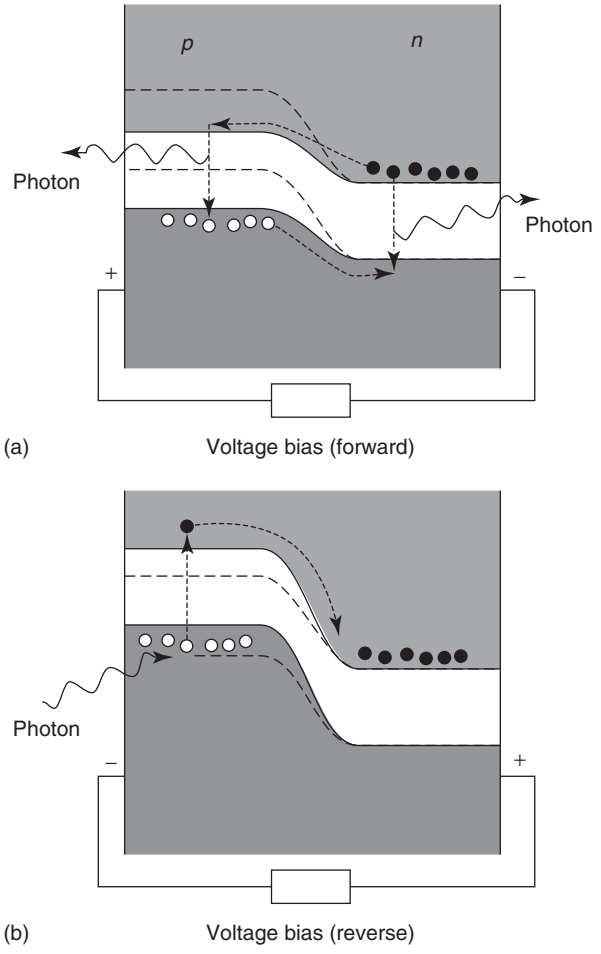


Figure 8 Simple types of semiconductor optoelectronic devices based on the *p*–*n* junction. (a) The forward bias of LED lowers the barriers for electrons and holes that can cross the junction reaching the opposite region, recombining and emitting photons. (b) Due to the reverse bias in the photodiode, very little current normally flows through the device. When photons with energy larger than the gap illuminate the junction, they can excite electrons into the conduction band of the *p*-type region. Some of these free electrons reach the interface region where they find a suitable voltage to move into the *n*-type region. The photons thus produce a current that is used to detect them.

See also: Effective Masses; Electrons and Holes; Excitons: Theory; Light Emitting Diodes; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Metals and Alloys, Impurity and Defect States in; Optical Instruments; Organic Semiconductors; Semiconductor Devices; Semiconductor Lasers; Semiconductors, Electronic Transitions in; Semiconductors, Optical and Infrared Properties of.

PACS: 61.82.Fk; 71.20.Mq; 71.20.Nr; 71.55.Cn; 72.80.Cw; 72.80.Ey; 81.05.Cy; 81.05.Ea

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Pacific Grove: Brooks/Cole.
Grosso G and Pastori Parravicini G (2000) *Solid State Physics*. Amsterdam: Academic Press.
Harrison WA (1980) *Electronic Structure and the Properties of Solids*. New York: Freeman.
Kittel C (1995) *Introduction to Solid State Physics*. Hoboken: Wiley.
Seeger K (1999) *Semiconductor Physics: An Introduction*. New York: Springer.
Sze SM (1981) *Physics of Semiconductor Devices*. Hoboken: Wiley.
Wenckebach WT (1999) *Essentials of Semiconductor Physics*. Hoboken: Wiley.
Yu PY and Cardona M (2001) *Fundamentals of Semiconductors: Physics and Materials Properties*. New York: Springer.

Semiconductors, History of

F Bassani and G C La Rocca, Scuola Normale Superiore, Pisa, Italy

Introduction

The so-called information or computer superhighway has led in modern society to a transformation comparable to the industrial revolution of the nineteenth century. This recent informatic revolution developed in the second half of the twentieth century as a consequence of the scientific knowledge of the microscopic world, conceptually based on quantum mechanics, and originated from the understanding of the electronic processes in a particular class of solids, known as semiconductors.

The existence of materials of this kind, which are poor insulators as well as poor conductors, had been noticed by Alessandro Volta (1745–1827) in the course of his studies on electricity at the end of the eighteenth century. The most common compounds of this kind were Ag_2S, PbS (galena), ZnS, CdS, ZnSe, and CdSe. Today, the most important semiconductor is silicon (Si), introduced by Antoine–Laurent Lavoisier (1743–1794) at the end of the eighteenth century as a component of sand (silex), and isolated by Joseph–Louis Gay–Lussac (1778–1850) and Jöns Jacob Berzelius (1779–1848) at the beginning of the nineteenth century.

The theoretical understanding of the physical phenomena which occur in semiconductors started in the 1930s and was developed throughout the rest of the last century, leading to new concepts of fundamental relevance, particularly in the field of many-body physics and light–matter interaction, and opening the way to new applications with the extreme miniaturization of the electronic and optoelectronic devices. At present, the semiconductor industry plays a crucial role in the economic system of every technologically advanced country, and is in constant growth because it satisfies the ever-increasing need to execute complex calculations, and to store and transmit information.

The development of the scientific and social history which led to these results is a fascinating subject and a prominent example of the strict interplay between physics and technology. An attempt has been made to briefly indicate its main points.

The Origins

The first observation of a poorly conducting behavior is due to Stephen Gray (1666–l736) who noticed how wet cork put in contact with charged glass is able to attract objects. The concept of electrical conductor (conducteur d' électricité) was subsequently introduced by Jean-Théophile Desaguiliers (1683–1744) in a communication to the Académie Royale des Belles Lettres et Arts (Bordeaux, 1742).

However, it was Volta who found, during his studies on the transport of electricity from one body to another, that a large number of substances were poor insulators and poor conductors. He introduced the notion of imperfect conductor for such materials using the term semi-insulators (semicoibenti) to define substances "which are actually permeable to the electric fluid, but oppose a strong resistance to its passage."

When intense electrical currents became available thanks to Volta's pile, Humphry Davy (1778–1829), the father of electrochemistry, could carry out experiments on the conductivity of common metals (Cu, Ag, Fe, Pt, etc.), observing that in all of them it decreases with increasing temperature. His pupil Michael Faraday (1791–1867) extended the conductivity measurements to nonmetallic materials (oxides, carbonates, sulfates) and, in particular, to HgI_2, and discovered that the increase of temperature produces in them a great increase of the conductivity, contrary to what happens in metals. The data of Faraday and later data on Ag_2S and Cu_2S by Johann

Wilhelm Hittorf (1824–1914) have been reexamined by George Busch, who found that they were in agreement with an exponential increase of conductivity with increasing temperature, as shown in **Figure 1**.

Understanding this phenomenon was the subject of a long debate. Contrary to the case of ionic crystals where the electrical conductivity was due to the flow of ions made possible by vacancies as found by Jakov I Frenkel, Carl Wagner, Walter Schottky, Walter Jost, and Robert W Pohl in the early 1920s, in the case of these materials no appreciable weight change was observed at the electrodes. It is useful to remember that the current density can be written as

$$\boldsymbol{J} = \sigma \boldsymbol{E} = ne\mu \boldsymbol{E} \qquad [1]$$

where the conductivity σ depends on the density n of the charge carriers and on their mobility μ. To understand the process of electrical conduction, it is necessary to measure with precision, besides the charge e, the density of carriers n. This became possible after 1879, thanks to the discovery by Edwin H Hall (1855–1938) that the current in the presence of a magnetic field perpendicular to the flux direction produces an electric field orthogonal to them and proportional to the current and to the magnetic field such as to compensate the deflecting Lorentz force. The proportionality constant is given by $1/nec$ and from its measurement, the sign and density of the charge carriers can be obtained, taking the elementary charge for e. This was crucial to determine that in metals and semiconductors the electrical current is carried by electrons, as shown by Carl V E Riecke (1845–1915) in 1901 at the University of Göttingen, who found that the charge-to-mass ratio obtained from the Hall effect measurements was consistent with the value e/m found by Joseph J Thompson in 1898 for electrons in vacuum. While Riecke and Paul Drude (1863–1906) formulated the model of free-electron conductivity in metals, Johann Koenigsberger (1874–1946) introduced, in 1913, the assumption that in semiconductors the conductivity is thermally activated with a material-dependent activation energy.

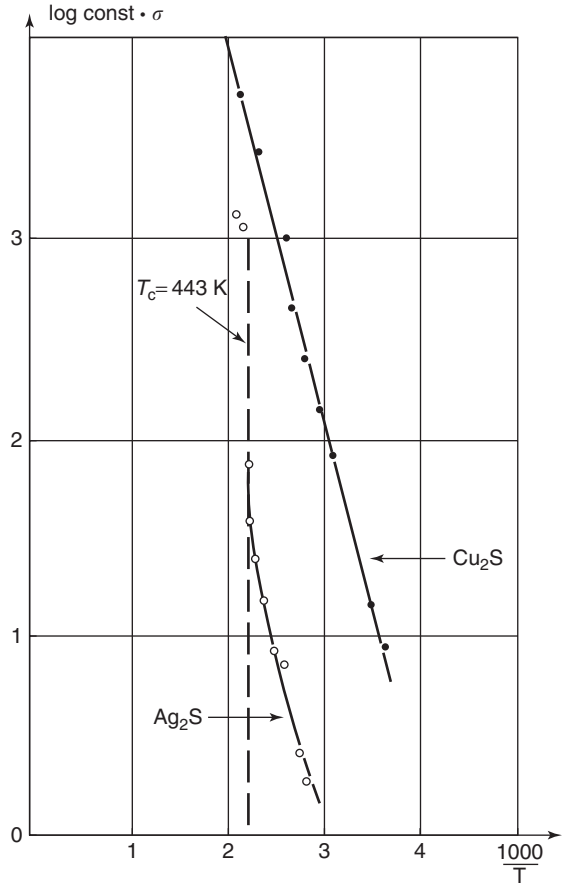


Figure 1 Conductivities of Cu_2S and Ag_2S as a function of temperature. It is to be noted that the change of slope in Ag_2S at $T_c = 443$ K is due to a phase transition.

An important and, for a long time, unexplained result was found by Karl Baedeker (1877–1914) in CuI in 1909, and later on by Orso Maria Corbino (1876–1937) in other substances: the Hall constant changed sign and showed an anomalously low concentration of carriers, yet with a large current. Now it is known that this is due to the presence of two types of carriers (negative electrons and positive holes), both of which contribute to the current, but partly compensate each other in the Hall effect.

The early studies of semiconductors were particularly challenging because the experimental results were poorly reproducible due to the uncontrolled presence of impurities. This gave the semiconductor physics a bad reputation, to the point that Wolfgang Pauli (1900–1958) called it "dirty physics." Only the possibility of producing relatively pure substances during and after the second World War allowed one to overcome such problems and start the modern physics of semiconductors.

Two important properties of semiconductors were empirically discovered in the nineteenth century. Ferdinand Braun (1850–1918), the discoverer of cathode rays and who shared the Noble prize with Marconi in 1909, observed in 1874 the rectifying behavior of galena and of other semiconducting crystals, namely the fact that in a metal–semiconductor contact, the current flows easily only in one direction and not in the opposite one. William G Adams and Richard E Day, in 1877, discovered the photovoltaic effect in Se, that is the appearance of a potential difference (open circuit) or of an electrical current (closed circuit) due to the absorption of light.

Both effects were used for important technological applications: the rectifying diodes were used as efficient detectors of radiowaves, and photovoltaic devices were used as exposimeter in the photographic process.

Detectors of Electromagnetic Waves

The electromagnetic waves, produced by Guglielmo Marconi (1874–1937) and co-workers, were initially detected by the coherer, a glass tube containing loose metal filings or mercury drops making contact with two electrodes; in the presence of the electric field of the wave, the conductivity increased and the tube acted as a detector. At the beginning of the last century, semiconductor rectifiers were used in order to obtain the electroacoustic transformation of radio waves and to allow earphones to be used. Though less sensitive than the coherer, crystal diodes soon came to be of general use as radiowave detectors. They were later substituted by vacuum tubes, which had better performances and gave more reproducible results. The vacuum tube also gave the possibility of amplifying the current, using a grid (triode). However, semiconductor diodes were never completely abandoned as they required no additional circuitry to produce the electron current.

Semiconductor diodes came to acquire utmost importance when it was necessary to detect microwaves of a few centimeters of wavelength for which vacuum tubes were not adequate. In particular, the use of silicon crystal and tungsten diodes in connection with the development of radar technology was introduced in 1939 by two young British, Denis Robinson and H W B Skinner, who also employed such diodes as nonlinear elements in a heterodyne circuit to reduce the frequency of the return signal and allow for subsequent amplification.

During the war, further progress was made in the US at the Radiation Laboratory of MIT and at the Du Pont company by producing silicon and germanium crystals of high purity doped with controlled amounts of impurities. A historical *I–V* curve of a silicon–tungsten diode is shown in **Figure 2**.

At the same time as at General Electric, at the University of Pennsylvania and at the Bell Telephone Laboratories, a research program was started on germanium, a semiconductor similar to silicon, but with a higher carrier mobility. In this case, very precise measurements of the Hall effect and the activation energy were made. The conductivity activation energy in the intrinsic regime was found to be 0.33 eV, while at lower temperatures, it was determined by the impurity content. Similar results were valid also for silicon, with an activation energy of 0.6 eV.

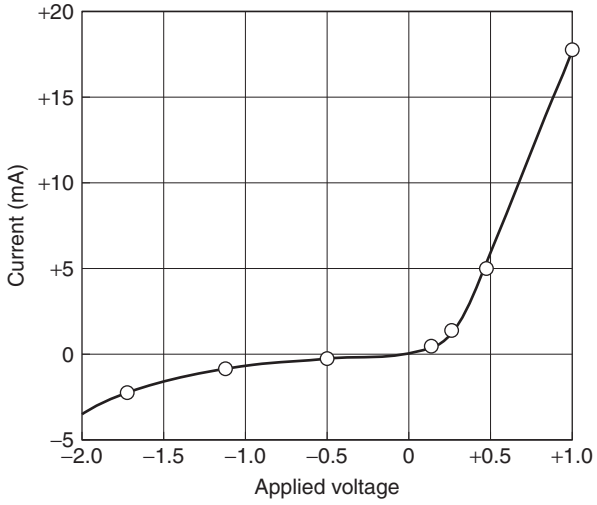


Figure 2 Historical *I–V* curve of a silicon–tungsten diode. (Adapted from Torrey HC and Whitner CA (1948) *Crystal Rectifiers*. McGraw-Hill.)

Since then, the technological interest for semiconductor crystals became predominant, and special attention was focused on crystals of group IV as well as III–V and II–VI compounds of particular simple symmetry (each atom at the center of a regular tetrahedron). Semiconductor physics after the second World War, was mainly developed by the group of Frederick Seitz at the University of Pennsylvania, and of Karl Lark-Horovitz at Purdue University, as well as at Bell Laboratories, clearly indicating the interest in the use of germanium in telecommunications.

Theory of Electronic Processes

The theoretical understanding of electronic processes in semiconductors developed gradually after the advent of quantum mechanics in the 1920s. The basic concept is that energy bands are available states for electrons in crystals. A theorem proved by Felix Bloch (1905–1983) in 1928 shows that the eigenfunctions of the Schrödinger equation in a crystal, and the corresponding available energies depend on the good quantum number $\boldsymbol{k}$ in reciprocal space. Such $\boldsymbol{k}$-vectors are uniformly distributed in the first Brillouin zone of the reciprocal space and their number equals that of the elementary cells in the crystal (half the number of atoms in Si and Ge, which have two atoms per unit cell). The energy values available to the electrons are organized in bands $E_n(\boldsymbol{k})$, each $\boldsymbol{k}$-vector corresponding to a de Broglie wavelength $\lambda = 2\pi/k$ and a momentum $\hbar\boldsymbol{k}$ of the electron. Such energy bands can be associated to the atomic states to which they would reduce if the interatomic distance were greatly increased.

The second basic concept is the exclusion principle introduced by Pauli in 1925 which states that each state can be occupied by one electron only (two electrons if spin is taken into account). Such a principle was formulated to account for the chemical properties of the atoms, but it turned out to explain the electronic properties of metals, insulators, and semiconductors, as shown conceptually by Alan H Wilson in 1931 while he was at Lipsia in the group of Werner Heisenberg. At zero temperature, in fact, electrons fill up the states of lower energy up to the Fermi level. At higher temperature, they follow the Fermi statistical distribution and some of them occupy higher electronic states. The totally filled bands cannot give any contribution to the current flow because of the exclusion principle. In metals, the Fermi level is inside a partly occupied band, and this allows the electron of that energy to transfer onto nearby empty levels with a behavior similar to that of free electrons. In insulators, totally occupied bands are separated from empty ones by an energy gap E_g, sufficiently wide to prevent the presence of electrons in higher bands at any temperature below the melting point. Semiconductors are insulating at very low temperatures, but their energy gap E_g is sufficiently narrow (typically less than 2 eV) so that at high temperatures a number of electrons per unit volume can be in the normally empty conduction band above the Fermi level. Such electrons behave as electrons in metals and are responsible for a current flow in the presence of an applied electric field.

For the purpose of illustration, **Figure 3** shows the energy bands of silicon computed with appropriate mathematical techniques in the 1950s by various authors (C Herring, F Hermann, T O Woodruff, F Bassani), and computed since then again and again with ever-increasing accuracy. One can observe that the eight valence electrons per unit cell (four for each of the two atoms) fill up all the states of the four lowest bands up to the maximum $\Gamma_{25'}$ at $\boldsymbol{k} = 0$, and the gap $E_g = 1.2$ eV separates these states from the conduction band minimum in the Δ direction at

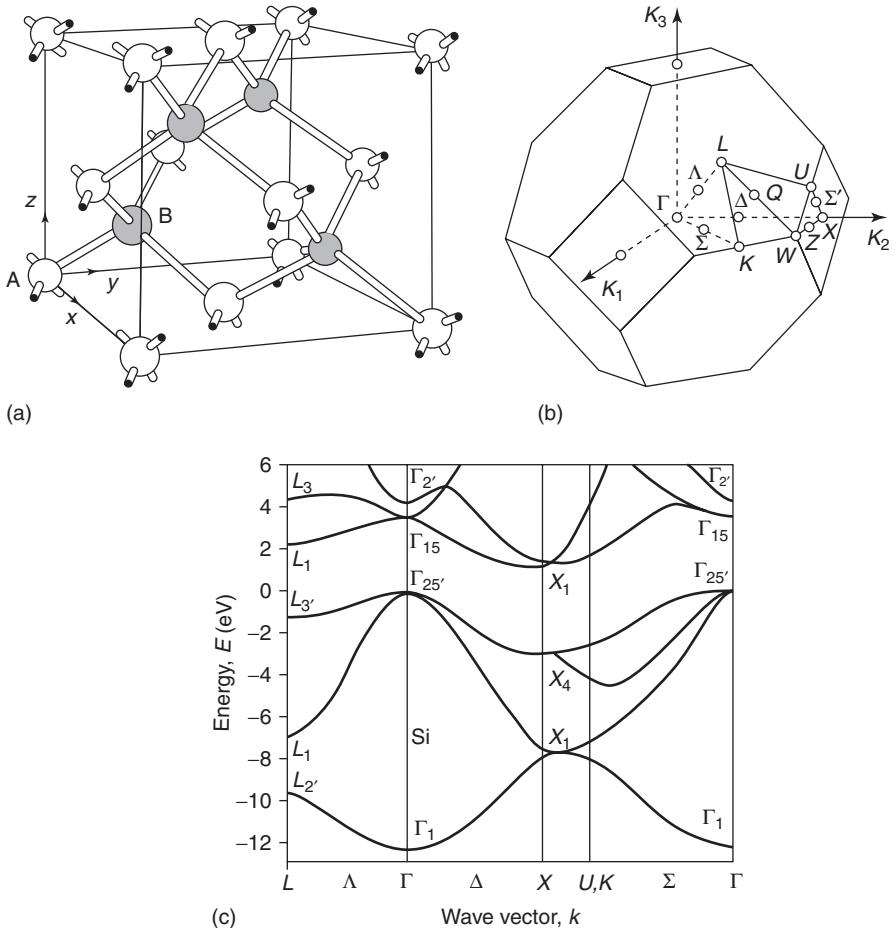


Figure 3 Structure of the zinc blende lattice: in Si both dark and light sites are occupied by Si atoms, in GaAs they are occupied by Ga and As atoms respectively; (b) Brillouin zone of the zinc blende lattice with the indication of points and lines of high symmetry; (c) Electronic band structure of silicon: the valence band maximum is the $\Gamma_{25'}$ state and the conduction band minimum is along the Δ line.

$\boldsymbol{k} = (2\pi/a)(0.8, 0, 0)$, a being the lattice constant. In this case, the gap is said to be indirect and the lifetime of electrons and holes is long because their recombination requires the participation of a phonon to preserve momentum. In other semiconductors having a direct gap, such as GaAs, optical recombination is allowed without phonon participation and has a much higher probability. Analogous results have been obtained experimentally and theoretically for all semiconductors of interest.

Another fundamental concept is that of a "hole" of electron, introduced by Heisenberg for atoms and by Rudolf Peierls for solids in 1931: an empty electronic state in a band otherwise totally occupied behaves like a positively charged particle and thus contributes to the current. The concept is analogous to that of the antiparticles of Dirac, the positrons, but in this case the effective mass m^* of electrons and holes are different as they are determined by the curvature of the bands $E_n(\boldsymbol{k})$ at the minimum of the conduction band and at the maximum of the valence band, respectively. In fact, the behavior of electrons near the band extrema is determined by the relation

$$E(\boldsymbol{k}) = E_0 \pm \frac{\hbar^2 k^2}{2m^*} \qquad [2]$$

where $+$ or $-$ refer to the minimum or maximum of the respective bands. The existence of holes explains the anomalous Hall effect mentioned above which had remained a mystery for decades.

In an intrinsic semiconductor, the concentration of electrons in the conduction band equals that of holes in the valence band, and they are given by

$$n_e = n_h = C\, e^{-E_g/(2K_B T)} \qquad [3]$$

where the prefactor depends on the effective masses and weakly on temperature. This explains the measured values of the activation energies which is half the energy gap (1.2 eV for Si, 0.66 eV for Ge).

In the case of doped semiconductors, as first suggested by Wilson, the impurities which substitute atoms of the lattice introduce additional electrons when their chemical valence exceeds that of the substituted atom, and introduce holes in the opposite case. For instance, group III impurities (B, Al) in Si and Ge introduce holes and group V impurities (P, As) electrons. Such holes or electrons are bound to the impurity atoms, but the attraction is very small because the dielectric constant screening the Coulomb interaction is large (typically >10) and the effective mass is small. Such states are then partly ionized and most of the additional electrons and holes are in the conduction or valence bands at room temperature. This explains why the amount of doping allows one to control the type of conductivity and its values. One can obtain semiconductors of n-type (with negative carriers) and of p-type (with positive carriers). The mobility μ has a much smaller influence on the conductivity with respect to the carrier density n; the latter can change by orders of magnitude and is controlled by the amount of doping (extrinsic regime) or by temperature (intrinsic regime).

This simple conceptual scheme was made available in the first half of the century, but its definite success is due to accurate research carried out in the second half of the century, particularly on the semiconductors Si and Ge. Among the main protagonists of the successful explanation of the electronic processes in semiconductors are, besides the aforesaid scientists, Walter Kohn and Charles Kittel. Kohn developed the density-functional theory to include many-body effects in the energy band calculations. Kittel developed the cyclotron resonance technique in semiconductors, which consists in producing transitions with microwaves between quantum levels induced by magnetic fields and measuring the corresponding absorption peaks. The cyclotron frequency separating the levels is eB/m^*c, which allows the experimental determination of the effective masses.

Transistor, Integrated Circuits, and Microprocessors

First of all, the band theory allowed the understanding of current rectification in the metal–semiconductor junction as due to a potential barrier at the interface produced by a migration of electrons from the side of an n-type semiconductor to the side of the metal (W Schottky, 1939). Such a potential barrier, shown in **Figure 4a** for the n-type case (an analogous situation occurring also for holes in p-type semiconductors), allows the passage through the interface of only those electrons having a higher energy. An applied external field would change the height of the potential barrier lowering it for one polarity and increasing it for the other. The number of electrons that can overcome the interface barrier varies by the factor $e^{\pm eV/K_B T}$; as a consequence, there is a great increase of the current only in the direction in which the barrier is lowered. For the same reason, as shown in **Figure 4b**, one obtains rectification at a junction between n-type and p-type semiconductors where the potential barrier is due to the fact that the Fermi level is near the conduction band or the valence band in n-type and p-type semiconductors, respectively. In 1958, Leo Esaki discovered that current can also be obtained by a tunnel effect for sufficiently high

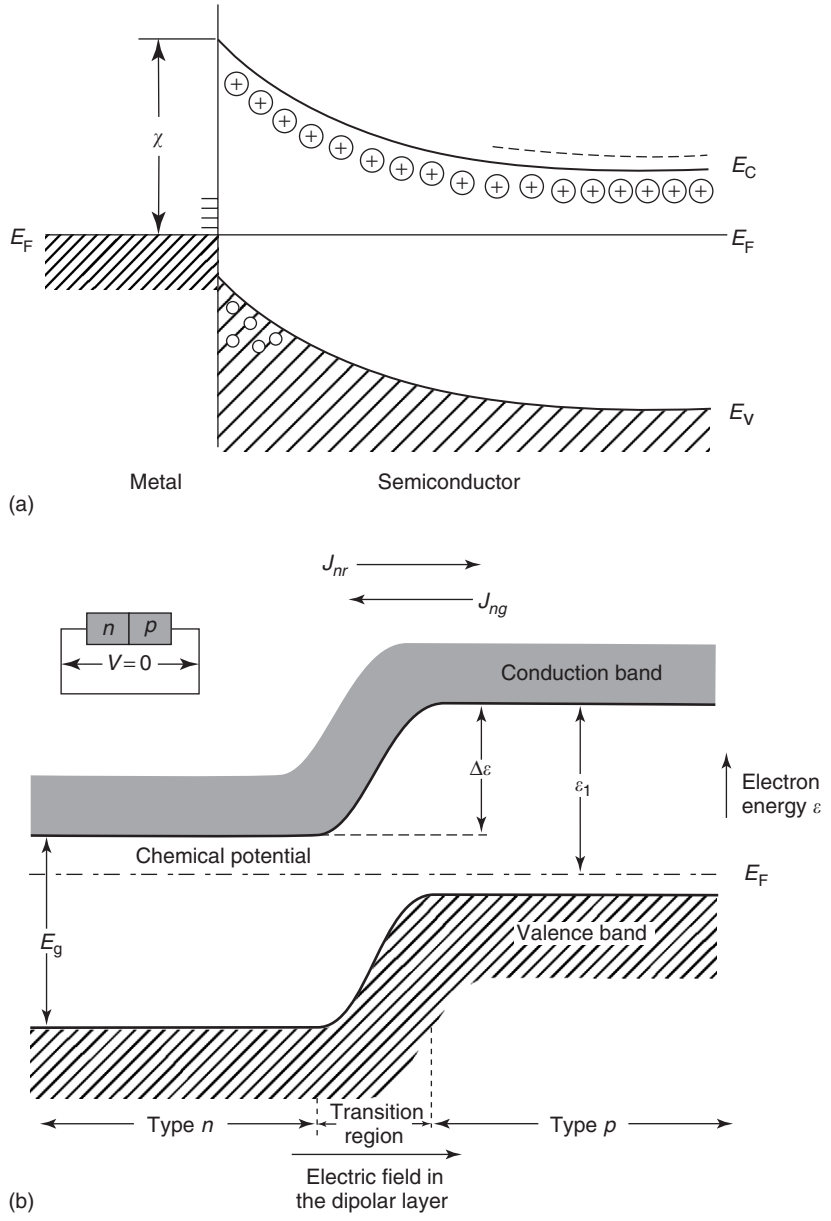


Figure 4 (a) Schematic diagrams of a metal–semiconductor Schottky barrier, and (b) schematic diagram of a *p–n* junction.

electric fields in the direction in which, normally, the diode would not conduct.

It was natural to expect that this type of knowledge could lead to a device similar to the triode, and thus allow amplification of the current. Such a device was created by J Bardeen, W H Brattain, and W Shockley in 1947 at the Bell Telephone Laboratories of Murray Hill. It consists of two junctions, one metal–semiconductor and the other semiconductor–metal, very close to each other as shown in **Figure 5a**. A small current in the base gets amplified into a large current flowing through the collector. The name "transistor" was proposed by J R Pierce of Bell Telephone Laboratories because the large swing of the collector-to-base voltage (load circuit) driven by a small variation of the base-to-emitter current implies a resistance transfer. A second type of transistor was soon invented by Shockley combining two semiconductor junctions, for example in the scheme *n–p–n* as shown in **Figure 5b**. In this case, the first *n–p* junction acts as the emitter-to-base junction and the second one as the base-to-collector junction. Again, the base current is amplified into the collector current. This bipolar junction transistor was immediately used as a substitute for amplifying vacuum tubes with hot emitting filaments in electronic circuits and computers, its main advantages being the compact dimensions and the limited power dissipation.

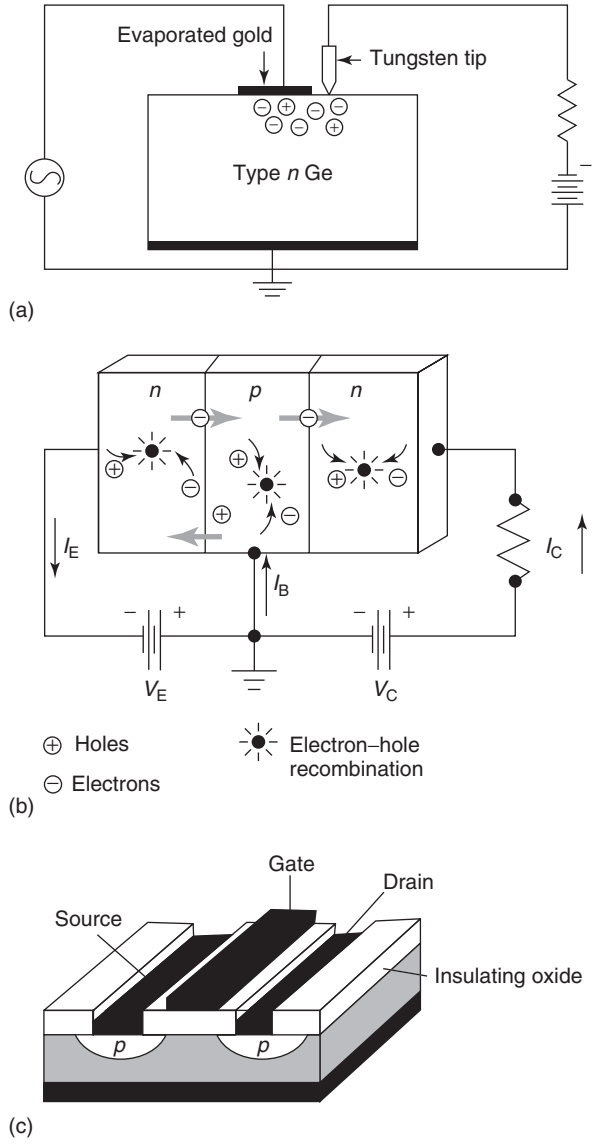


Figure 5 (a) Scheme of the point-contact transistor; (b) scheme of the *n*–*p*–*n* transistor; (c) scheme of a MOSFET.

A further development was the field-effect transistor (FET) suggested by Bardeen in 1950 and developed by G C Dacey and I M Ross in 1953. In this case, a voltage is applied to the base (also called gate) generating a field that attracts or repels charge carriers favoring or impeding the current flow from the emitter (also called source) to the collector (also called drain). A metal-oxide field-effect transistor (MOSFET) can be realized exploiting the surface oxide to isolate the contacts and building the device on a single surface of a monocrystal of Si, as shown in **Figure 5c**.

FETs, and particularly the MOSFETs, are the basic ingredients of the planar integrated circuits that have opened the way to the miniaturization of electronics. The connecting wires are substituted by the metal evaporated in the channels engraved in the oxide and capacitors, resistors, and transistors are properly inserted at the semiconductor surface.

The advantage of using a planar architecture explains the almost exclusive use of silicon as the basic material of the electronic industry as its oxide is an excellent insulator, and is not igroscopic as germanium oxide. The development and mass production of these semiconductor "chips" have led to the informatic revolution of the second half of the twentieth century. The first complete circuit system obtained by chemical processing on the surface of a silicon monocrystal eliminating the need of connecting metal wires was realized by Jack Kilby (Texas Instruments) and Robert Noyce (Fairchild Semiconductors) in 1959. In the 1970s, the first memory devices were developed by introducing on the gate of a MOSFET an isolated metallic substrate which can trap or release electrons by a tunnel effect under the application of an electric field, thus switching the device on and off. As a consequence of the above developments at Texas Instruments, at Fairchild Semiconductors and at the new Intel company, a complete microprocessor (including logical and memory devices) was realized on a single chip with a diameter of $\sim$1 inch (Federico Faggin, 1972).

Such results have been obtained through the driving force of strong industrial companies which allowed continuous technological improvements and large-scale production. The rapid growth of the field also favored the development of new companies, entirely dedicated to the production of semiconductor devices, concentrated in the region that became known as Silicon Valley in California.

These events give a vivid example of the close connection among science, technology, and industry in modern society. The development of microprocessors is still in progress, the increase of the number of transistors per unit area on a single chip following the so-called Moore's law of doubling each year, reaching more than a billion units per chip in 2002. At the same time, there has been a large increase in the cost of a single factory and a large decrease in the cost of a single chip, while the semiconductor industry production has enormously increased reaching a total worldwide market value of a thousand billion dollars per year.

Laser and Photonic Devices

Semiconductors are also very important for their optical properties which allowed the development of microlasers and other photonic devices. Optical transitions corresponding to the absorption (emission) of

light occur when an electron is excited from the valence band to the conduction band (an electron from the conduction band recombines with a hole). As mentioned before, such processes are most probable when such transitions are vertical ($\Delta \boldsymbol{k} = 0$) because the photon momentum (h/λ) is negligible with respect to that of electrons ($\hbar k \simeq \hbar/a$). As a consequence, direct gap semiconductors, such as GaAs, are the most useful ones to develop photonic devices.

The photovoltaic effect is exploited in the solar cell, which allows the transformation of photon energy into electrical power following photon absorption at the junction, as indicated in the scheme shown in **Figure 6**. One observes that the reverse process, in which a current is converted into light, takes place in a junction working as a light emitting diode (LED).

The first silicon solar cell was realized by D M Chaplin, C S Fuller, and G L Pearson (in 1954) with a conversion efficiency of the solar spectrum of 6%. Later on, the efficiency reached 24% in GaAs. The theoretical limit in bulk materials is due to the fact that photons, with an energy higher than the gap, convert only part of it into electrical power, the rest being thermally dissipated in the semiconductor. To overcome this limitation, solar cells based on graded gap materials are being developed.

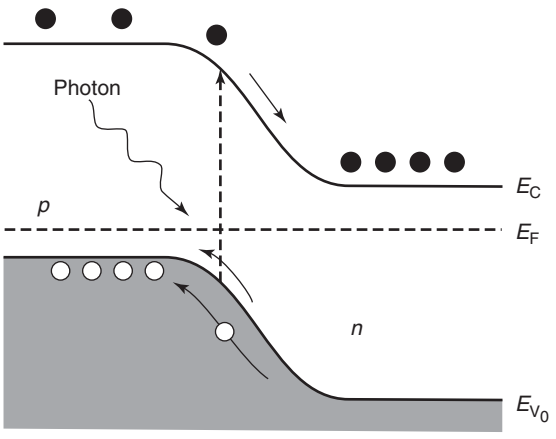


Figure 6 Scheme of a *p–n* junction solar cell.

In the 1960s, semiconductor lasers were first developed using heavily doped GaAs *p–n* junctions by R N Hall and N Holonyak at General Electric, by M I Nathan at IBM, and by T M Quist at MIT. The basic idea is to obtain population inversion in the electronic system; the applied electric field driving the current produces a spatial superposition of heavily *n*-type and *p*-type doped regions, as indicated in **Figure 7**. While electrons and holes recombine, population inversion is maintained by the continuous injection of carriers (electrical pumping). Furthermore, the surfaces of the semiconductor itself act as mirrors forming an optical cavity which allows for optical gain and laser action. An important development has been the heterostructure laser realized by Z Alferov (1969). The advantages of semiconductor lasers with respect to gas lasers are the reduced dimensions, the small power consumption, and the low cost. The frequencies of the light emitted can be chosen using semiconductors with different energy gaps, from the infrared (InSb) to the ultraviolet (GaN).

Heterostructures and Low-Dimensional Devices

In the last two decades of the last century, semiconductor physics received a new incentive from the realization of artificial crystals in which the chemical composition is controlled at the atomic level, some parts being made of one semiconductor and other parts by a different one (heterostructures). Such substances are produced by epitaxial growth from

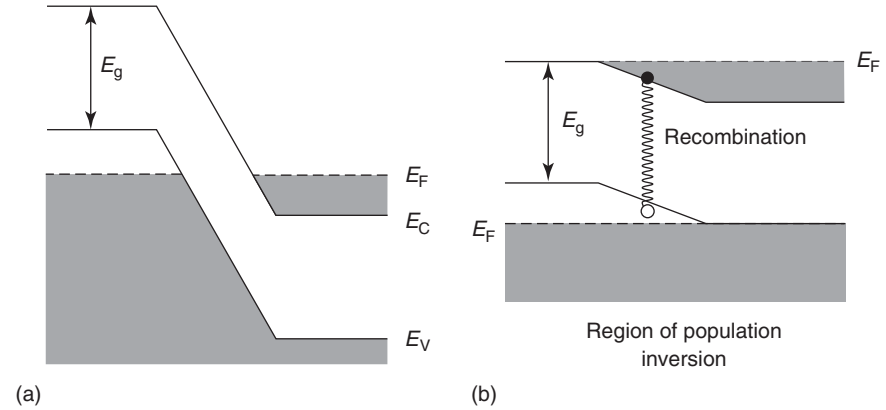


Figure 7 Scheme of a semiconductor junction laser: (a) without the applied electric field; (b) with the applied field, in working condition.

molecular beams (MBE) impinging under ultrahigh vacuum on a heated substrate, or by other growth techniques such as chemical vapor deposition (CVD). These heterostructures are not thermodynamically stable, but once produced, have a practically infinite lifetime at room temperature.

The first of such nanostructures is the superlattice, introduced by L Esaki and R Tzu in 1970, made of sequences of N atomic planes of semiconductor A with lattice constant a, and M atomic planes of semiconductor B with lattice constant b, giving rise to a large periodicity $Na + Mb$ in the growth direction. In this direction, the possible values of the wave vector $\boldsymbol{k}$ are greatly reduced and the energy bands are correspondingly narrow (minibands).

Other nanostructures can be obtained by inserting some planes of a semiconductor of a lower gap inside a semiconductor of a larger gap. Then, electrons and holes tend to be confined in the lower gap material forming two-dimensional subbands, each associated to a discrete level produced by quantum confinement (quantum well). In such a case, by controlling the thickness of the well, one can control the frequency of the optical transitions between subbands, thus obtaining lasers of desired frequency, infrared detectors, etc.

The first experimental observation of optical transitions in a quantum well is due to R Dingle (1974), and since then, there has been a continuous development in photonic nanostructures. The two-dimensional confinement of the electronic states also has a profound influence on transport properties and this has led to unexpected phenomena, such as the quantum Hall effect. At a low temperature and in a high magnetic field, the Hall resistance of a two-dimensional electron gas in the inversion layer of a silicon MOSFET or in a quantum well shows as a function of the magnetic field, well-defined plateaus, from which the universal quantum of resistance h/e^2 can be obtained (V Klitzing, 1982).

One can also realize quantum confinement in two directions producing quantum wires with one-dimensional subbands, or quantum confinement in all three directions producing quantum dots with discrete atomic-like levels (see the scheme in **Figure 8**).

A further nanostructure is the microcavity, obtained when a semiconductor of width comparable to half the wavelength of light in the material is sandwiched by totally reflecting layers (metallic mirrors or dielectric Bragg mirrors), as shown in **Figure 9**. In such a system, one obtains a concentration of light of wavelength which can be chosen in resonance, for instance, with a quantum well optical transition to which it is strongly coupled (C Weisbuch, 1992).

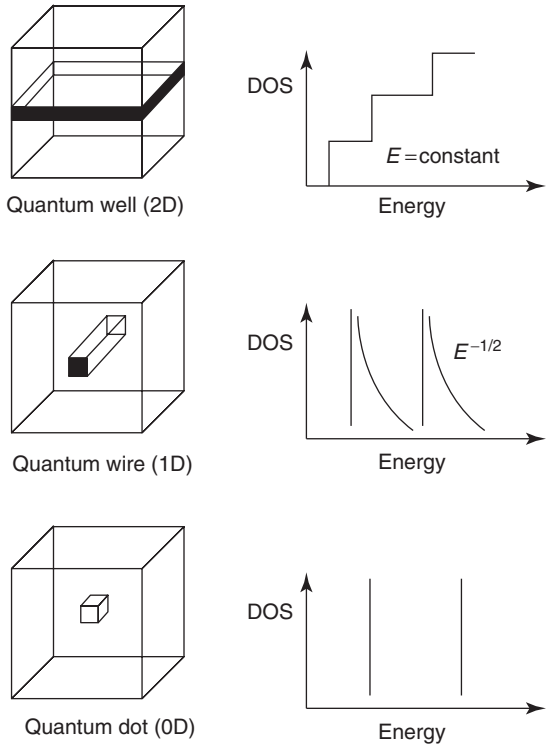


Figure 8 Nanostructures with quantum confinement in 1, 2, or 3 dimensions.

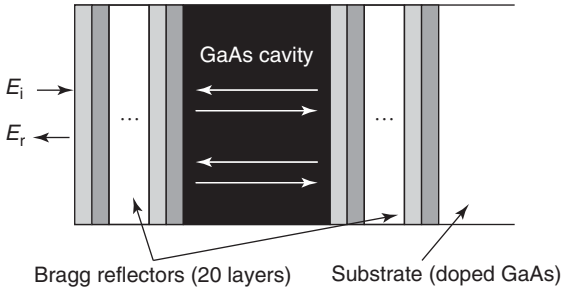


Figure 9 Semiconductor microcavity with distributed Bragg reflectors.

The use of nanostructures for the production of a number of important optical effects is now pursued. An important example is the quantum cascade laser originally proposed by R Suris and R Kazarinov in 1971 and realized for the first time by F Capasso at the Bell Telephone laboratories in 1994. In this unipolar device, quantum wells and superlattices are alternated in great numbers and displaced in energy by the applied electric field driving electrons through the structure. The complex design is such that each flowing electron makes many transitions and, in the proper regime for population inversion and laser action, emits many laser photons. Recent developments by A Tredicucci and co-workers at the Scuola

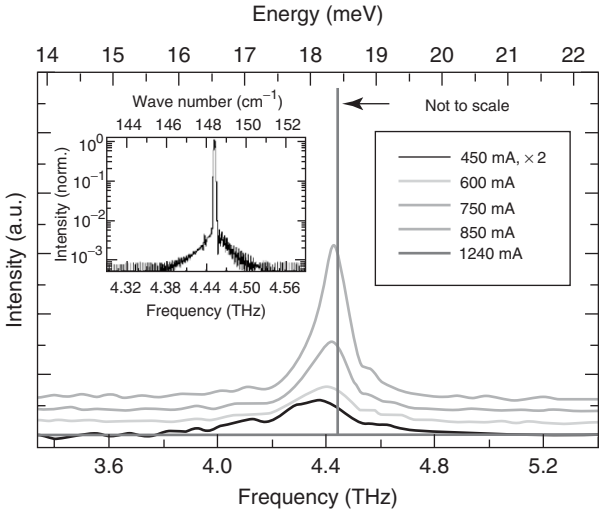


Figure 10 Quantum cascade laser emission. (From Köhler R, Tredicucci A, Beltram F, Beere HE, Linfield EH, *et al.* (2002) THz semiconductor heterostructure laser. *Nature* 417: 156.)

Normale Superiore in Pisa (2002) have led to a laser with high spectral quality in the frequency region of the terahertz (see **Figure 10**), which is of great interest for the spectroscopic detection of complex molecules and for the possibility of penetrating through most substances.

Conclusions

The history of semiconductors, from Volta's times to the present day, has witnessed many extraordinary results in fundamental and applied physics. Since the Nobel prize awarded for the discovery of the transistor, more than ten other Nobel prizes have acknowledged milestone achievements in semiconductors physics. One can anticipate that this field will continue to produce important results in the future. The theoretical limit of the miniaturization of solid-state electronic devices down to the atomic level has not yet been reached. In photonics, the new nanostructures are very promising not only for laser light production, but also for nonlinear optics and coherent control in microcavities and hybrid organic–inorganic structures. Finally, the development of computers and quantum computers based on photons as well as on electrons is a very ambitious goal that will be reached only through major developments in semiconductor physics and nanotechnology.

See also: Elemental Semiconductors, Electronic States of; Integrated Circuits; Memory Devices, Nonvolatile; Memory Devices, Volatile; Semiconductor Lasers; Semiconductor Optics; Semiconductors, General Properties; Silicon, History of; Transistors.

PACS: 01.65.+g; 72.20.−i; 71.20.−b; 78.40.Fy

Further Reading

Agrawal GP and Dutta NK (1986) *Long Wavelength Semiconductor Lasers*. New York: Van Nostrand Reinhold Co.

Bassani F and La Rocca G (1996) Semiconductors, elemental-electronic properties. In: GL Trigg (ed.) *Encyclopedia of Applied Physics*, vol. 17, pp. 403. New York: VHC Publisher American Institute of Physics.

Bassani F and Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Busch G (1993) Early history of physics and chemistry of semiconductors: from doubts to facts in a hundred years. *Condensed Matter News* 2: 15.

Kressel H and Butler JK (1977) *Semiconductor Lasers and Heterojunction LEDs*. London: Academic Press.

Runyan WR and Bean KE (1990) *Semiconductor Integrated Circuit Processing Technology*. Reading, MA: Addison-Wesley.

Seitz F and Einspruck N (1998) *Electronic Genie: The Tangled History of Silicon*. Urbana: University of Illinois Press.

Semiconductors, Impurity and Defect States in

H Ikoma, Tokyo University of Science, Tokyo, Japan

Introduction

The role of mobile charge carriers is a key factor in semiconductor device operations. These free carriers are supplied by doping semiconductors with certain kinds of impurities. The electrical properties of semiconductors can thus be controlled by adding a small amount of impurities. However, some other impurities and crystal defects often have detrimental effects on the performance of the device. These impurities and defects create discrete energy levels in the bandgap of the host semiconductor, and exert various influences (usefully or detrimentally) on the semiconductor characteristics. Therefore, a knowledge of the electronic structure of the impurities and defect states in semiconductors is very important.

Structural and Electrical Properties of Impurities and Crystal Defects

Impurity atoms normally substitute the atoms of the host semiconductor (substitutional impurity atom).

However, some of them, especially those with small atomic radii, often enter the interstitial site in the host crystal lattice (interstitial impurity atom). On the other hand, the well-known point defects – the vacancies (lattice points where atoms are lacking) and the interstitials (atoms of the host semiconductor occupying the interstitial sites) – are examples of the intrinsic crystal defects. These point defects sometimes generate a complex center such as the Frenkel defect (the pair of vacancy and interstitial atoms in the neighboring lattice sites). There are two types of impurities and defects from the electrical point of view. One of them supplies free electrons and the other provides holes in the host crystal. The former is called donor and the latter, acceptor. For example, when V-column elements such as P and As are doped in Si (IV-element), one excess valence electron is easily liberated (the impurity atom is ionized) in the Si lattice even at room temperature because of the low binding energy between the electron and the impurity atom. Also, when a III-element such as B is introduced into Si, one valence electron is deficient so that one hole is created and easily released in the Si lattice at room temperature due to their very low binding energy. These impurities, having very low binding (ionization) energies are termed shallow impurities (shallow donor and shallow acceptor). If Si is doped with VI-elements (S, Se, and so on) or II-elements (Zn, Cd, and so on), two free electrons or two holes per impurity atom, respectively, are created (double donor and double acceptor). On the other hand, some impurity elements (and crystal defects) have very high binding energies so that electrons or holes are not easily liberated in the host crystal at room temperature. These impurities with high binding energies are called deep impurities. Heavy metals such as Fe, Ni, and Cu in Si, for example, are deep impurities. In compound semiconductors, the same impurity atom behaves both as a donor and an acceptor depending on its occupied site (amphoteric impurities). The Si atom in GaAs behaves as a donor if it occupies the Ga site and as an acceptor when it substitutes the As atom. When IV-elements (Ge, C, etc.) are doped in Si, they also behave as donors or acceptors although there are cases when they are not electrically active. These impurities are called isoelectronic impurities. With doping of the impurities in semiconductors, the discrete energy levels are normally generated in the bandgap of the host semiconductor. For shallow donors and acceptors, these energy levels are located very near the bottom of the conduction band and the top of the valence band, respectively. On the other hand, the energy levels for the deep impurities occur near the middle of the bandgap.

Theory of the Electronic States of the Shallow Impurities

Shallow impurities can be theoretically treated using the effective mass approximation and the hydrogen atom model. Consider the case of the donor (the acceptor can also be considered similarly). The excess electron moves in an orbit round the impurity ion due to the Coulomb interaction between them in the host crystal, just like the hydrogen atom. The Coulomb potential is much smaller than that in the hydrogen atom due to the screening effect of the valence electrons (the change of the valence electron distributions around the impurity ion) in the crystal lattice. The screened Coulomb potential is written as $U = Zq^2/\varepsilon_s r$, where Z is the valence number of the impurity ion, q the electronic charge, ε_s the dielectric constant of the host semiconductor and, r is the distance between the electron and impurity ion. The screening effect is effectively represented by the dielectric constant ε_s. Thus, the electron is very loosely bound by the impurity ion (very low binding energy) and its wave function extends in the wide range of the crystal lattice. Hence, the electron moves feeling both the periodic potential of the host crystal and the screened Coulomb potential. The effect of the periodic potential on the motion of electrons is well described in terms of the "effective mass" m^*, that is, electrons in the periodic potential of crystal behave as if they are free electrons with their mass m^* being different from that (m_0) in a vacuum. Therefore, the electrons in the orbital round the shallow impurity ions can be considered as electrons with an effective mass m^* moving in the screened Coulomb potential U. The fundamental Schrödinger equation for these electrons is written as follows:

$$\left[\left(\frac{\hbar}{2m_0}\right)\nabla^2 + (V - U)\right]\Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r})$$

Here, $\nabla^2 = (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$ is an operator, $\Psi(\boldsymbol{r})$ the wave function of the electron, $\boldsymbol{r} = \boldsymbol{r}(x, y, z)$ the position vector of the electron, V and U the periodic potential in the perfect crystal and the screened Coulomb potential, respectively, and E is the energy of the electron. In the case of $U = 0$ (perfect crystal), the solution of the above equation is given by the Bloch wave function $\Psi n\boldsymbol{k}(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})u_{n\boldsymbol{k}}(\boldsymbol{r})$, where $\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})$ is the wave function of the free electron (plane wave) extending to the whole crystal, and $u_{n\boldsymbol{k}}(\boldsymbol{r})$ is the function having the periodicity of the crystal lattice with n and $\boldsymbol{k}$ being the quantum number specifying the energy bands and the wave number vector, respectively. For applying the effective mass approximation to the shallow

impurity problem, the wave function $\Psi(\boldsymbol{r})$ is expanded by the Bloch functions or their Fourier transforms (the Wannier functions) and substituted in the equation. Thus, the following equation is obtained:

$$\left[\frac{\hbar}{2m^*}\nabla^2 - U\right]\Phi(\boldsymbol{r}) = E\Phi(\boldsymbol{r})$$

This equation is the same as that for the hydrogen atom except for the difference in the electron mass. The wave functions of the electrons are given as the product of $\Phi(\boldsymbol{r})$ and the function which has the periodicity of the crystal lattice. The above equation is for the case of the isotropic conduction band (in which the effective mass is the same for all crystal orientations). For example, some compound semiconductors such as GaAs and InP have this type of conduction bands. On the other hand, the conduction bands are anisotropic in semiconductors such as Si, Ge, and GaP, that is, the effective mass is different in different crystal orientations. The valence bands of most semiconductors have degeneracy in which the energies of the "heavy" holes and the "light" holes coincide at the top of the band. For these cases, the above equation should be modified to take the anisotropy or degeneracy of the band structures into consideration. The details of the same are not discussed here.

The solution of the above equation is well known for the "hydrogen atom." The electronic states of the hydrogen atom are specified by the principal (Bohr) quantum number n, the angular momentum quantum number l, the magnetic quantum number m, and the spin quantum number s. The principal quantum number n is the positive integer and the angular momentum quantum number $l \leqslant n-1$ for each n value. The magnetic quantum number m is given as $m = l, l-1, l-2, 0, \ldots, -l-1, -l (l \geqslant m \geqslant -l)$ and the spin quantum number $s = \pm 1/2$. The electronic states corresponding to $l = 0$, 1, and 2 are called s, p, and d states, respectively. Then the state of $n = 1$ and $l = 0$ is called as the 1s state (the ground state: the state of the lowest energy). The higher energy states (the excited states) are thus called as 2s ($n = 2$, $l = 0$), 2p($n = 2$, $l = 1$), 3s($n = 3$, $l = 0$), 3p($n = 3$, $l = 1$), 3d($n = 3$, $l = 2$) states and so on. The energies E_n of these states obtained from the above equation are given as

$$E_n = -E_0/n^2 \quad (n = 1, 2, 3, \ldots)$$
$$E_0 = (\varepsilon_0/\varepsilon_s)^2 (m^*/m_0) E_H$$

Here E_0 is the energy of the ground (1s) state and, ε_0 and ε_s are the dielectric constants of the vacuum and semiconductor respectively. E_H is given as

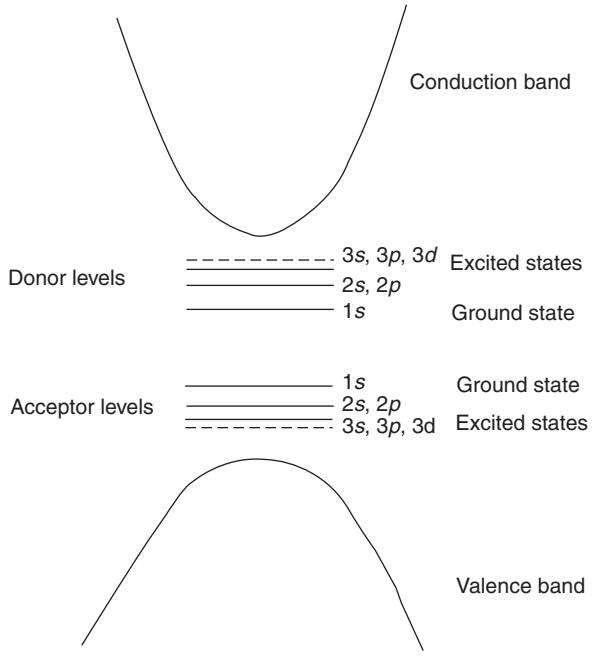


Figure 1 Shallow impurity level scheme in the wave number (k) space.

$E_H = (1/\varepsilon_0)^2 (q^4 m_0/2\hbar^2) = 13.6$ eV and is the ground-state energy of the real hydrogen atom (the Rydberg constant). From the above equations, it is obvious that the energy E_n is only dependent on the principal quantum number n (degenerate with respect to the other quantum numbers l, m, and s) in the case of the isotropic conduction band. For the donor and acceptor, respectively, the ionization energies (the binding energies) E_D and E_A from the ground state can be calculated as $E_D = E_C - E_0$ and $E_A = E_V - E_0$, where E_C and E_V are, respectively, the energies of the bottom of the conduction band and the top of the valence band. In **Figure 1**, the shallow impurity levels are schematically shown.

The wave function is $\Phi_{1s}(r)$ of the ground state is written in the following:

$$\Phi_{1s}(r) = (1/\pi)^{1/2} (1/a^*)^{3/2} \exp(-r/a^*)$$

Here, it is assumed that $Z = 1$. a^* is the Bohr radius given as $a^* = (\varepsilon_0 m^*)(\hbar^2/m_0 q^2)$ and represents the extent of the electronic wave function in the real crystal space. The wave function exhibits a sharp decrease at $r \gg a^*$. The wave functions of some excited states are as follows:

$$\Phi_{2s}(r) = (1/\pi)^{1/2} (1/2a^*)^{3/2} [2 - (r/a^*)] \exp(-r/2a^*)$$

$$\Phi_{2p}(r) = (1/\pi)^{1/2} (1/2a^*)^{3/2} (r/\sqrt{3}) \exp(-r/2a^*)$$

In these excited states, the electron wave functions show a strong decay as

$$r \gg 2a^*$$

Experimental Determination of the Shallow Energy Levels

The energy levels of shallow impurities can be determined accurately using the infrared absorption measurements at very low temperatures. It is well known that electrons bound to a hydrogen atom can be excited to higher energy levels by the absorption of light. Similarly, electrons captured by shallow impurity levels are transferred to higher energy levels by absorbing infrared light. These optical transitions are allowed only when the difference (Δl) in the angular momentum quantum numbers (l) between two related energy levels is 1 ($\Delta l = 1$) (the selection rule). Thus the electrons can be excited from the $1s$ (ground) state to the $2p$, $3p$, $4p$, … (excited) states. However, the transitions $1s \rightarrow 2s$, $1s \rightarrow 3s$, … (and so on) are not allowed. The absorption peaks of the infrared light are very sharp because these energy levels are completely discrete. By measuring the infrared absorption spectrum at very low temperature (~ 4.2 K), one can accurately determine the energies between the ground state and many excited states. The excited levels with very large principal quantum numbers n are located very close to the conduction or valence band edge so that one can determine the ground-state ($1s$) level and any excited ($2p$, $3p$, …) levels measured from the band edge (ionization energies). In **Table 1**, some examples of the measured shallow donor and acceptor levels (ionization energies) in Si are shown together with the calculated ones. The agreement between the measured and calculated values is excellent.

Table 1 Some examples of the ionization energies of the ground states of shallow donors and acceptors

Host semiconductors	*Impurities*	*Experimental value* (meV)	*Theoretical value* (meV)
Si	P (D)	45.5	44.3
	As (D)	53.7	53.1
	Sb (D)	42.7	31.7
	B (A)	45.0	31.6
GaAs	Si (D)	5.84	5.72
	Ge (D)	5.88	5.72
InP		7.14	7.14
InSb	Te (D)	0.6	0.6

D: donor; A: acceptor.

Theory of the Electronic States of the Deep Impurities (or Defects)

For the deep impurities, electrons (or holes) are very tightly bound to the impurity atom nucleus due to the highly localized core potential. The motion of the bound electrons (or holes) is strongly affected by this core potential and the wave functions of these carriers are also highly localized around the impurity atoms (or defect centers). Therefore, both the effective mass approximation and the hydrogen atom model cannot be applied to the deep level cases. The tight binding approximations, such as the linear combination of atomic orbitals (LCAO) method, are suitable to treat the deep-level problem. Since the core potentials are not small, the conventional perturbation theory is not applicable and the Green-function approach is useful to solve the fundamental Schrödinger equation. The Schrödinger equation describing the motion of the electron bound to the deep impurity is given as follows:

$$(H_0 + U)\Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r})$$

Here, H_0 is the Hamiltonian operator for the perfect crystal and U is the localized potential due to the deep impurity atom. In order to solve the above equation, the impurity potential U needs to be determined. However, this is very difficult for the following reason. In many cases of deep impurities, the impurity atoms are slightly displaced from the lattice point of the host semiconductor, and a microstrain is induced around the impurity atoms. It is assumed that if the impurity atom occupies the lattice point (as is the case for shallow impurities), a shallow energy level of E_D is generated. When this impurity atom is somewhat displaced from the lattice point (the lattice relaxation occurs), a deep energy level E_0 is assumed to be created. Also, a lattice relaxation energy (strain energy) E_L occurs. If $E_0 + E_L$ is higher than E_D, it is energetically stable to form the deep level with a small displacement of the impurity atom from the lattice point rather than the shallow levels. If the mixed semiconductor crystal AlGaAs is doped with Si, a deep-level center (called the DX center) is created while Si atoms act as shallow donors in GaAs. This DX center is considered to be as the above stated case. In order to determine the impurity potential U, the effect of the lattice relaxation should be taken into account exactly, which, however, is a very difficult problem. Many calculations have been performed by assuming various forms of the impurity potential U. The wave function $\Psi(r)$ is expanded in terms of the sp^3 atomic orbital functions (Φ_j) as $\Psi = \sum_j a_j \Phi_j$ (LCAO). Then the Schrödinger equation is

solved using the Green-function approach under the various assumptions of the impurity potentials. However, quantitative comparisons between the calculated and observed energy levels are difficult because of the difficulty in including the lattice relaxation effect in the analysis.

There are some cases of deep levels, in which the energy level merges into the conduction or the valence band of the host semiconductor. This level is called the resonant level. The existence of this level affects the band structure of the host crystal because the density of states of the impurities are superposed to those of the conduction or the valence band. With doping of the isoelectronic impurities, this resonant level is often observed. These resonant energy levels can also be calculated using the above Green-function approach.

Experimental Determination of the Deep Energy Levels

The deep level transient spectroscopy (DLTS) is the most powerful tool to determine deep energy levels. This method is based on the measurements of the transient electrical capacitances of the metal–semiconductor (Schottky) contact as the pulsed reverse bias voltage is applied. The transient Schottky capacitance $C(t)$ is written as follows.

$$C(t) = C_\infty \left[1 - \left(\frac{N}{2}\right)\exp\left(\frac{-t}{T}\right)\right]$$

Here $N = N_T/(N_D + N_T)$, N_T and N_D are the density of the deep impurity and the shallow donor, respectively. $C_\infty = \{[(\varepsilon_s\varepsilon_0/2)(V + V_D)]\ (N_D - N_T)\}^{1/2}$, ε_s and ε_0 are the dielectric constant of the semiconductor and a vacuum, respectively, and V_D is the diffusion potential of the Schottky contact. The relaxation time τ can be written as $e_n = 1/\tau$, where e_n is the electron (or hole) emission coefficient of the deep impurity and given as $e_n \propto C_n \exp[-(E_C - E_T/kT)]$ because electrons (and holes) are emitted from the deep levels by thermal activation. Here, C_n is the capture cross section of the deep impurity, E_T and E_C are the deep impurity energy level and the bottom of the conduction band, respectively. T and k are the absolute temperature and the Boltzmann constant, respectively. If the transient capacitance $C(t)$ is measured as a function of temperature T, the temperature dependence of e_n is obtained. From the plot of $\ln e_n$ versus $1/T$ (the Arrhenius plot), the deep energy level E_T can be easily determined. The capture cross section C_n and the deep impurity density N_T are at the same time determined. The famous example of the deep level is the EL2 center observed in the semi-insulating GaAs single crystal which is the cause of the high resistivity. The ionization energy of this EL2 center was measured to be $0.7 \sim 0.8$ eV (located near the middle of the bandgap of GaAs).

See also: Electrons and Holes; Insulators, Impurity and Defect States in; Metals and Alloys, Impurity and Defect States in; Semiconductors, General Properties.

PACS: 61.43.Dq; 61.72.Vv; 61.82.Fk; 71.22.+i

Further Reading

Pantelides S (1978) The electronic structure of impurity and defect states in semiconductors. *Reviews of Modern Physics* 50: 797–858.

Pantelides S (1986) *Deep Centers in Semiconductors, A State of the Art Approach*. New York: Gordon and Breach.

Yu PY and Cardona M (1996) *Fundamentals of Semiconductors*. Berlin: Springer.

Semiconductors, Optical and Infrared Properties of

L Degiorgi, ETH Zürich, Zürich, Switzerland

Introduction

An insulator is distinguished from a metal by its vanishing DC conductivity at low temperature. In contrast to what happens in metals, the electronic charge in insulators and semiconductors (and quite generally nonmetals) cannot flow freely under an applied DC field, but instead undergoes static polarization. Within classical physics, this qualitative difference is attributed to the nature of the electronic charge: either "bound" (Lorentz model for insulators) or "free" (Drude model for metals). In other words, electrons are localized in insulators and delocalized in metals. Switching to quantum physics, this clear-cut distinction is apparently lost. In most textbooks, the insulating/metallic behavior is explained by means of the band structure theory, focusing on the position of

the Fermi level of the given material: either in a bandgap (insulators), or across a band (metals). The one-electron approximation is the main paradigm of the band theory in solids. It consists of an electron assumed to be acted on by the field of the fixed atomic cores plus an average field arising from the charge distribution of all the other outer-shell electrons.

The one-electron model for solids is generally quite adequate as confirmed by several experiments on a wide variety of materials showing a good agreement with theory. This is particularly true in semiconductors and insulators when investigated by optical experiments. These experiments provide ample evidence for direct and indirect gaps, as well as for excitonic states.

This article addresses some basic features of insulators and semiconductors from the perspective of the optical and infrared properties. The article is organized as follows: starting with a broad introduction of the characteristic features of semiconductors and insulators, a few general concepts of the electronic band structure, such as the single-particle gap and the interband transitions, are addressed. The phenomenological Lorentz model is emphasized, which is a useful tool in order to describe the relevant energy scales and absorptions in semiconductors, such as the already mentioned gap and electronic transitions in general as well as the phonon modes and the impurity/defect states. In fact, optical studies can also shed light on the role of doping in semiconductors. The second part of this article then discusses, as an example, a prototype semiconducting/insulating material, focusing the attention on its experimental optical response. An outlook to some challenging issues in the present basic research of semiconductors and insulators concludes this article.

Electronic Interband Transitions: A Few Concepts

From a microscopic point of view, one can state that the excitation of one electron from the valence band (v) to the conduction band (c) (**Figure 1**) leads to an extra electron with momentum $\boldsymbol{k}_1$ in the conduction band, leaving a hole with momentum $\boldsymbol{k}_2$ in the valence band. In a crystal, momentum conservation requires that

$$\boldsymbol{k}_1 = \boldsymbol{k}_2 + \boldsymbol{K} \qquad [1]$$

where $\boldsymbol{K}$ is the reciprocal lattice vector. Neglecting Umklapp process, the so-called vertical direct transitions occur for $\boldsymbol{k}_1 = \boldsymbol{k}_2 = \boldsymbol{k}_0$ such that the excitation energy $\hbar\omega_0$ satisfies the relation:

$$\hbar\omega_0 = \varepsilon_c(\boldsymbol{k}_0) - \varepsilon_v(\boldsymbol{k}_0) \qquad [2]$$

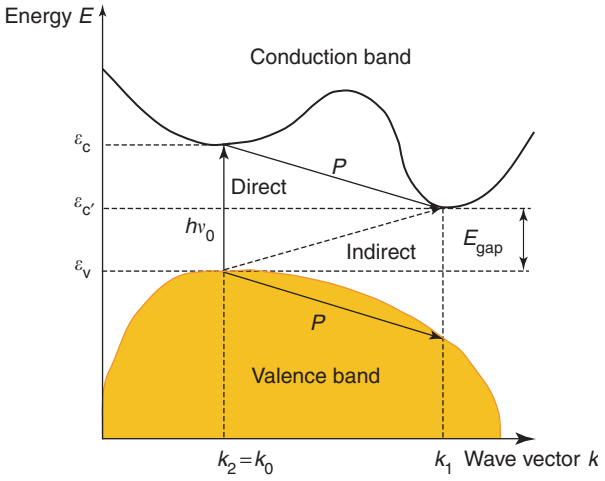


Figure 1 Transitions between occupied (valence) and unoccupied (conduction) bands in an insulating material. A direct transition at $\boldsymbol{k}_0$ and an indirect (phonon assisted) one between $\boldsymbol{k}_2$ and $\boldsymbol{k}_1$ are shown. $\boldsymbol{P}$ stays for the phonon momentum. The direct transition occurs at $\hbar\omega_0 = \varepsilon_c - \varepsilon_v$, while the indirect excitation occurs at $E_{gap} = \varepsilon'_c - \varepsilon_v$.

due to the energy conservation (**Figure 1**). The optical response is fully described by the complex dielectric function ($\tilde{\varepsilon}(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$). For direct interband transitions, it can be shown that the imaginary part ε_2 of $\tilde{\varepsilon}(\omega)$, describing the absorption spectrum, is given by

$$\varepsilon_2(\omega) = 8\left(\frac{\pi e}{m\omega}\right)^2 |p_{cv}|^2 J_{cv} \qquad [3]$$

where p_{cv} is the electric dipole transition matrix element and

$$J_{cv} = \frac{1}{8\pi^3}\int_S \frac{dS}{|\nabla_k \varepsilon_{cv}|_{\varepsilon_{cv}=h\nu}} \qquad [4]$$

is the so-called combined or joint density of states. J_{cv} accounts for both available initial and final states involved in the electronic interband transition, and can be extracted from the electronic band structure. The integral is over a surface S of $\boldsymbol{k}$ vectors for which $\varepsilon_{cv} = \varepsilon_c - \varepsilon_v = \hbar\omega$. The critical points, where

$$\nabla_k \varepsilon_{cv} = \nabla_k [\varepsilon_c(k) - \varepsilon_v(k)] = 0 \qquad [5]$$

are called van Hove singularities in the joint density of states, which usually coincide with excitations at high symmetry points in the electronic band structure. This is the case for photon energies $\hbar\omega_0$ for which the two energy bands separated by $\varepsilon_{cv} = \hbar\omega_0$ are parallel at a particular $\boldsymbol{k}$ vector. At those k-points, van Hove singularities lead to peculiar absorption peaks and band edge features in the excitation spectrum and consequently determine the prominent

structures in the imaginary part ε_2 (eqn [3]) of the dielectric function, thus of the absorption coefficient.

In a large number of semiconductors, the energy maxima in the valence and minima in the conduction band do not occur for the same momentum $\boldsymbol{k}_0$, but for different momenta $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ (**Figure 1**). Optical direct transitions between these states cannot take place due to momentum conservation; however, such transitions become possible when the excitation of phonons is involved. For instance, one scenario involves the creation of a photon for wave vector $\boldsymbol{k}_2$ and the subsequent phonon emission, which absorbs the energy and momentum, necessary to reach the conduction band at momentum $\boldsymbol{k}_1$ (**Figure 1**):

$$\boldsymbol{k}_1 = \boldsymbol{k}_2 + \boldsymbol{P} \qquad [6]$$

with $\boldsymbol{P}$ the wave vector of the phonon involved. These processes are known as indirect (phonon assisted) transitions. An opposite process can also contribute to the absorption spectrum (**Figure 1**). Such indirect transitions have several distinct characteristics when compared with direct transitions. For instance, the absorption increases with frequency as the square of the energy difference:

$$\varepsilon_2(\omega) \sim N_{\mathrm{P}}(E_{\mathrm{gap}} - \varepsilon_{\mathrm{P}} \pm \hbar\omega)^2 \qquad [7]$$

(with ε_{P} the phonon energy and N_{P} the phonon occupation number) instead of the square root dependence for direct transitions near the band edge (E_{gap} represents here the direct gap):

$$\varepsilon_2(\omega) \sim (\hbar\omega - E_{\mathrm{gap}})^{1/2} \qquad [8]$$

obviously with $\hbar\omega > E_{\mathrm{gap}} \pm \varepsilon_{\mathrm{P}}$ (eqn [7]) or $\hbar\omega > E_{\mathrm{gap}}$ (eqn [8]). Second, the indirect absorption is strongly temperature dependent, reflecting the phonon population factor (N_{P}). These differences make the distinction between direct and indirect optical transitions in analyzing measured spectra relatively straightforward.

The smallest possible energy for which an electronic transition occurs defines the insulating or semiconducting single-particle bandgap E_{gap} (**Figure 1**). Insulators can be then classified as direct or indirect gap materials depending on whether the interband transition at E_{gap} occurs at $\boldsymbol{k}_0$ or between $\boldsymbol{k}_2$ and $\boldsymbol{k}_1$ (**Figure 1**), respectively. **Table 1** gives a few examples of band semiconductors, with their direct or indirect gap values. The single-particle bandgap is therefore the central feature in insulators or semiconductors. At $T = 0$ the valence band is full, whereas the conduction band is empty (**Figure 1**), so that the Fermi level lies between these bands. A zero DC conductivity follows at $T = 0$.

Table 1 Energy gap (E_{gap}) at 300 K of selected semiconductors

Material	E_{gap} (eV)	*Gap*
Si	1.12	i
Ge	0.67	i
InSb	0.16	d
GaAs	1.4	d
GaP	2.2	d
C (diamond)	5.5	i

"i" indicates an indirect gap and "d" a direct one.

The Phenomenological Lorentz Model

The simplest and general situation that mimics the transition between different bands corresponds to the electronic direct transition between a ground state (valence band) and excited states (conduction band) of identical atoms. The excitation induced by the electromagnetic fields can be treated using the time-dependent Schrödinger equation. In a perfect crystal, the total crystal potential felt by the electrons has the periodicity of the crystal lattice and the solutions of the Schrödinger equation are the well-known Bloch functions. Thorough descriptions of this approach, finally leading to an explicit form of eqn [3], can be found in the specialized literature.

In order to highlight the relevant optical fingerprints in band semiconductors or insulators, the so-called Lorentz model is employed. While phenomenological, this model catches the main features in and allows extracting several useful parameters from the optical response of semiconductors. By assuming the response of a classical harmonic oscillator, the electrons in the Lorentz model are bound to the nucleus of the atom (like a small mass bound to a large mass by a spring). Such a classical dispersion theory leads to the following expression of the complex dielectric function:

$$\tilde{\varepsilon}(\omega) = \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega) = \varepsilon_\infty + \frac{\omega_{\mathrm{p}}^2}{(\omega_0^2 - \omega^2) - \mathrm{i}\Gamma\omega} \qquad [9]$$

where $\hbar\omega_0$ is the resonance energy of the bound electron, Γ is the damping and ω_{p}^2 is the oscillator strength. The analogy with the quantum mechanical treatment identifies $\hbar\omega_0$ with the energy difference between two electronic states involved in the transition (like the insulating gap) and ω_{p}^2 as a measure of the relative probability of the transition.

Figure 2 displays the real ε_1 (dispersive) and imaginary ε_2 (absorptive) part of $\tilde{\varepsilon}(\omega)$ for one single absorption, which exactly occurs at ω_0. The full width of the region of anomalous dispersion in ε_1 is Γ (see **Figure 2**). For $\Gamma \ll \omega_0$, there is little dispersion and in this limit the full width of the ε_2 curve at half

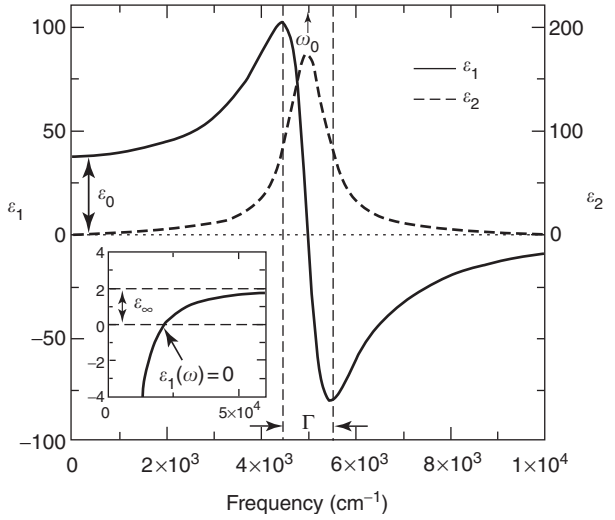


Figure 2 Real (ε_1) and imaginary (ε_2) part of the complex dielectric function calculated after the Lorentz model for one single harmonic oscillator (eqn [9]). The parameters are $\varepsilon_\infty = 2$, $\omega_0 = 5000\,\mathrm{cm}^{-1}$, $\Gamma = 1000\,\mathrm{cm}^{-1}$, and $\omega_p = 3 \times 10^4\,\mathrm{cm}^{-1}$. The static dielectric constant $\varepsilon_1(\omega \to 0) = \varepsilon_0$ (eqn [10]) is indicated. The inset is a blowup of $\varepsilon_1(\omega)$ at high frequencies, showing the zero-crossing $\varepsilon_1(\omega) = 0$ and the high-frequency limit $\varepsilon_1(\omega \to \infty) = \varepsilon_\infty$ of the real part of the dielectric function. Here and elsewhere in this article, the photon energy is expressed as frequency in a wave number (cm^{-1}) so that $1\,\mathrm{eV} = 8.065\,48 \times 10^3\,\mathrm{cm}^{-1}$.

maximum is also given by Γ. The real part ε_1 is characterized by the $\omega \to 0$ limit, the so-called static dielectric constant ε_0 (**Figure 2**):

$$\varepsilon_0 = \varepsilon_\infty + \frac{\omega_p^2}{\omega_0^2} \qquad [10]$$

and, by the $\omega \to \infty$ limit, coinciding with the optical constant ε_∞ (inset of **Figure 2**). For a material with a discrete number of Lorentz harmonic oscillators at high frequencies (representing, e.g., interband transitions), each of such oscillators will contribute to ε_0 in an additive way. Therefore, ε_∞ phenomenologically takes into account the static contribution from those high-frequency (energy) excitations.

From the well-known relations of the classical electrodynamics, one can evaluate the optical functions ($\mu = 1$):

$$\varepsilon_1 = n^2 - k^2 \qquad [11]$$

$$\varepsilon_2 = 2nk = \frac{4\pi\sigma_1}{\omega} \qquad [12]$$

where n is the refractive index, k is the absorption coefficient, and σ_1 is the real part of the optical conductivity.

From the experimental point of view, several techniques allow to reach those fundamental quantities. One way is to extract these optical functions

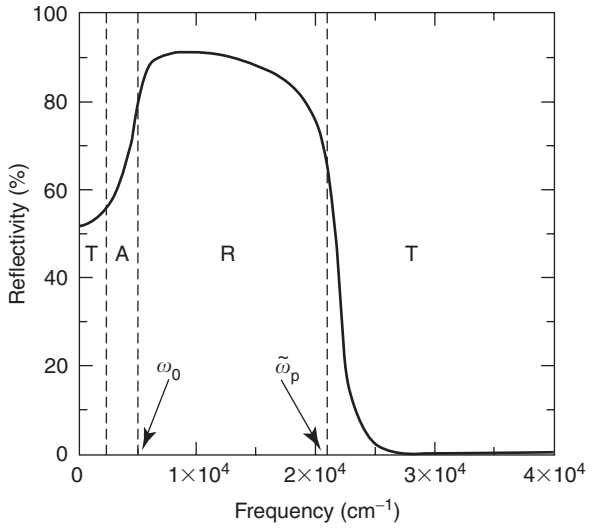


Figure 3 Optical reflectivity for an insulating material calculated with eqns [9], and [11]–[13] within the Lorentz model for one single (e.g., gap-like) excitation (same parameters as in **Figure 2**). T stands for transparent, A for absorption, and R for reflection.

from the measurement of the optical reflectivity, defined as

$$R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \qquad [13]$$

When the reflectivity is known in a broad spectral range, extending over several decades in photon energy, one can perform reliable Kramers–Kronig (KK) transformations. The KK transformations are classical dispersion relation based on the causality principle, and relate the phase Θ of the complex reflectance $\tilde{r} = \sqrt{R}e^{i\Theta}$ to the measured magnitude R. When the phase Θ is known, one can then calculate all optical functions, by solving the equation

$$\sqrt{R}e^{i\Theta} = \frac{(1-\tilde{n})}{(1+\tilde{n})} \qquad [14]$$

with respect to n and k ($\tilde{n} = n + ik$) and using eqns [11] and [12].

For the purpose of illustration, **Figure 3** displays the typical shape of $R(\omega)$ for an insulator, as calculated within the Lorentz model for one single (gap) excitation after eqns [9], and [11]–[13]. Starting from zero frequency in the region for $\omega < \omega_0$, one has $\varepsilon_2 = 0$ and $\varepsilon_1 > 1$ so that $k \sim 0$ and $\varepsilon_1 = n^2$. Thus, this is a region of high transparency (T), and no absorption takes place. By further increasing the frequency, there is, on the contrary, strong absorption (A), while for $\omega > \omega_0$ the electrons of the insulators respond as if they were free electrons (R). This is because the photon energy is

much greater than the binding energy $\hbar\omega_0$ of the electron, which is ultimately excited to the continuum. Therefore, in this latter region the $R(\omega)$ spectrum looks like the one of a metal. If $\Gamma \to 0$, it follows that $R(\omega) \to 100\%$, which is also called the "Reststrahlen" or total reflection region. At frequencies of the order of $\tilde{\omega}_p \approx \omega_p/\sqrt{\varepsilon_\infty}$, there is a sharp drop of $R(\omega)$. $\tilde{\omega}_p$ also defines the frequency where the material is again transparent (T)!

In passing, it is mentioned that the metallic behavior is obtained within the Lorentz model by setting $\omega_0 = 0$ in eqn [9]. In that case $R(\omega) \to 100\%$ (total reflection) for $\omega \to 0$ and the (sharp) drop at high frequency (i.e., $\omega \sim \tilde{\omega}_p$) of **Figure 3** would then correspond to the so-called plasma edge (see "Further reading" section for more details about the optical response of a metal).

The phenomenological presentation, adopted so far, considered one single electronic excitation coinciding with the insulating gap. In general, the absorption spectrum is further characterized by other excitations, the origin of which can be rather different, depending on the spectral range where those excitations appear. At higher energies, for instance, electronic interband transitions are possible. Below the lowest possible electronic interband transition (gap), in the so-called transparent region, one can detect various types of excitations such as the phonon modes due to the lattice vibrations or collective excitations. Excitons and impurity states can represent another important contribution to the absorption spectrum of semiconductors. Excitons are brought about by the Coulomb interaction and consist of an electron–hole pair which can be mobile. It does not contribute to the DC conductivity but can be excited by the electromagnetic field. Alternatively, Coulomb interactions between an impurity potential and electrons lead to bound states localized to the impurity potential sites. In all cases, the resulting excitations can be described simply by borrowing concepts developed for the energy levels of single atoms. These excitations appear as an absorption resonance, and similarly to the electronic interband excitations, can be described with dedicated Lorentz harmonic oscillators. Therefore, one can reconstruct the whole absorption spectrum of insulators by adding as many Lorentz harmonic oscillators as needed for each excitation.

Example of an Insulating System: EuS

Figure 4 shows the reflectivity and **Figure 5** the real part $\sigma_1(\omega)$ of the optical conductivity of EuS, which is chosen as an example for a typical insulating

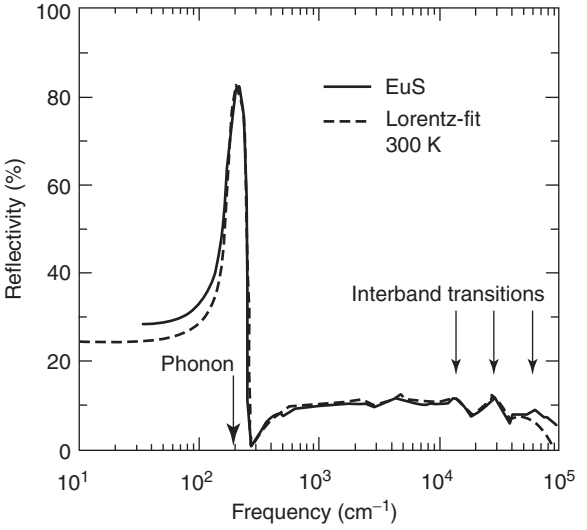


Figure 4 Optical reflectivity of EuS measured from the far infrared up to the ultraviolet (note the logarithmic energy scale). The infrared active phonon mode at 200 cm^{-1} is clearly seen. The dashed line is a Lorentz phenomenological fit (eqns [9], and [11]–[13]). The fit parameters are: $\varepsilon_\infty = 1.64$ for the optical dielectric constant, $\omega_{0ph} = 181\,\text{cm}^{-1}$, $\Gamma_{ph} = 13\,\text{cm}^{-1}$, $\omega_{p-ph} = 395\,\text{cm}^{-1}$ for the resonance frequency, damping and oscillator strength of the phonon, $\omega_{01} = 1.54 \times 10^4\,\text{cm}^{-1}$, $\Gamma_1 = 4444\,\text{cm}^{-1}$, $\omega_{p1} = 9587\,\text{cm}^{-1}$, $\omega_{02} = 3 \times 10^4\,\text{cm}^{-1}$, $\Gamma_2 = 1.2 \times 10^4\,\text{cm}^{-1}$, $\omega_{p2} = 2.24 \times 10^4\,\text{cm}^{-1}$, $\omega_{03} = 6.3 \times 10^4\,\text{cm}^{-1}$, $\Gamma_3 = 6.2 \times 10^4\,\text{cm}^{-1}$, $\omega_{p3} = 7.5 \times 10^4\,\text{cm}^{-1}$, for the resonance frequencies, dampings and oscillator strengths of the interband transitions. The frequency position for each Lorentz harmonic oscillator is indicated by arrows.

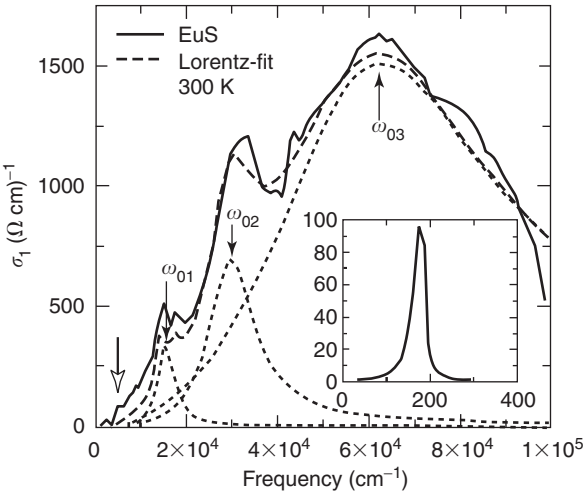


Figure 5 Real part $\sigma_1(\omega)$ of the optical conductivity of EuS obtained by KK transformation of the reflectivity $R(\omega)$ in **Figure 4**. A linear frequency scale has been used in order to emphasize the spectral range relevant for the interband transitions. The white arrow marks the onset of absorption at $\sim 5000\,\text{cm}^{-1}$. The Lorentz-fit (eqns [9] and [12]) with an appropriate number of harmonic oscillators (the single fit components are shown by dashed lines) can reproduce the interband transitions very well. The arrows identified the resonance frequencies ω_{0j} for the interband transitions within the Lorentz fit. The inset emphasizes $\sigma_1(\omega)$ in the far infrared, with the dominant phonon mode absorption. The fit parameters are the same as in **Figure 4**.

material. Eu chalcogenides have received tremendous interest from scientists all over the world. One obvious reason was the hope of using these magnetic semiconductors industrially, for example, in the form of fast, light beam addressable memory systems in computers, as magneto-optical modulators, and as magnetic-field-activated electronic switches. The hope for applications dwindled when one realized that upon doping, the magnetic-ordering temperatures could not be raised up to room temperature. However, the purely academic interest in magnetic semiconductors, especially the Eu chalcogenides, continued to be an enormous driving force, because the materials turned out to be model substances for studying magnetism. It is beyond the scope of this article to elaborate further on the interesting physics of these materials. Thus, attention is limited to the phenomenological discussion of the optical spectra, in order to illustrate the concepts developed above and also as an example for the successful application of the Lorentz model.

$R(\omega)$ (**Figure 4**) is typical for an insulator since it saturates to a constant value for $\omega \to 0$. On the one hand, one can clearly recognize the strong absorption feature at $200\,\mathrm{cm}^{-1}$ in $R(\omega)$, which is ascribed to the infrared phonon mode of the f.c.c. rocksalt structure of EuS. On the other hand, above 0.5 eV ($\sim 4000\,\mathrm{cm}^{-1}$) several absorption features can be identified. These latter absorptions can be better observed in $\sigma_1(\omega)$ (**Figure 5**). In this respect, **Figure 5** emphasizes the near infrared–ultraviolet spectral range, characterized by several peaks for the electronic interband transitions involving 4*f*, 3*p*, and 5*d* states. Of particular interest is the onset of absorption at $\sim 0.62\,\mathrm{eV}$ ($\sim 5000\,\mathrm{cm}^{-1}$) which is ascribed to the excitation between the localized $4f^7$ levels in the ^7F_J configuration and the bottom of the extended $5d(t_{2g})$ band (see white arrow in **Figure 5**).

From the phenomenological point of view, one can describe the absorptions in $\sigma_1(\omega)$ (phonon and electronic interband transitions) with *ad hoc* Lorentz harmonic oscillators. The total $\tilde{\varepsilon}(\omega)$ was obtained by adding four Lorentz harmonic oscillators (eqn [9]): one for the phonon and three for the electronic interband transitions (see parameters in the caption of **Figure 4**). **Figures 4** and **5** demonstrate the great fit quality achieved for both $R(\omega)$ and $\sigma_1(\omega)$ within the Lorentz model. The extracted parameters can then be used in order to achieve information about the intrinsic physical properties of the investigated material and can be compared with other measured quantities. For instance, the excitation energy $\hbar\omega_{0j}$ (**Figure 5**) can be compared with the allowed electronic interband transitions from the band structure calculations. The oscillator strengths ω_{pj} can be related to the transition probability and the excitation selection rules.

Insulators and Semiconductors: Topics of Interest

It should noted be that the picture considered so far for insulators applies only to a crystalline material (**Table 1**), within the independent-electron approximation. This is a very limited class of insulators indeed. In some materials, the insulating behavior is dominated by disorder (Anderson insulators), in some it is dominated by electron correlation (Mott insulators). Therefore, the band picture is grossly inadequate for a large number of insulators. To conclude this chapter, some aspects and topics are mentioned briefly, partially being still matter of debate and unsettled issues, which are of interest in semiconductors and/or insulators research. The interest from the perspective of the optical properties is particularly emphasized.

The dimensionality of the electronic structure plays an essential role in shaping the properties of solids. It is well known that metallic materials with linear-chain-like structure are characterized by a quasi-one-dimensional electronic structure and are therefore unstable. This favors a so-called nesting of the Fermi surface, which leads through a metal–insulator phase transition to a broken symmetry ground state, known as charge (**Table 2**) or spin density wave state, depending on whether the electron–phonon or the electron–electron interaction dominates. These phase transitions induce the opening of an insulating gap in the excitation spectrum, which can be detected optically.

Optical properties also reveal essential information on the nature of the phase transition and the related ground state. Indeed, the leading role played by the correlation effects in driving the phase transitions can be directly studied by analyzing the frequency dependence of the optical conductivity. For a one-dimensional band insulator, it is expected that

Table 2 Transition temparature T_c and single-particle energy gap E_{gap}, obtained from the DC resistivity, for some linear chain compunds with CDW ground state

Material	T_c(K)	E_{gap} (K)
KCP	189	1400
$K_{0.3}MoO_3$	183	920
TaS_3	215	1600
$NbSe_3$	145 and 60	700
$(TaSe_4)_2I$	263	3000

the optical conductivity will display a sharp singularity at the gap edge and will decay after the expression $\sigma_1(\omega) \sim (\omega - E_{gap})^{-1/2}$. Toward high frequencies, the decay is expected to follow the powerlaw $\sigma_1(\omega) \sim \omega^{-3}$ if the lattice is assumed to be rigid and only Umklapp scattering or of the single-period lattice potential is possible, or as $\sigma_1(\omega) \sim \omega^{-2}$ when coupling to phonons is included. A strongly frequency-dependent optical conductivity has indeed been found in several one-dimensional or nearly one-dimensional semiconductors, such as $K_2[Pt(CN)_4]Br_{0.3}3H_2O$ (better known as KCP) and $(NbSe_4)_3I$. For the latter material, a stronger dependence ($\sigma_1(\omega) \sim \omega^{-4.25}$, instead of ω^{-3}) was found, presumably due to the narrow band, with the bandwidth comparable to the optical frequencies where the measurements were made.

Electron–electron interactions may modify this picture further. Such interactions lead, in strictly one dimension, to an electron gas which is distinctively different from a Fermi liquid. In not strictly but only nearly one dimension, such interactions (Umklapp scattering process) lead to a (Mott) gap, and the features predicted by theory are expected to be observed only at high photon energies. These features include a power-law dependence of the conductivity, with the exponents on ω different from -3. Experiments on highly anisotropic materials, such as the linear chain Bechgaard salt $(TMTSF)_2PF_6$, have been interpreted as evidence for such a novel, so-called Tomonaga–Luttinger liquid.

Another topic of particular importance is the study of the optical response associated with impurity states or defects in semiconductors (as in the case of phosphorus doped silicon Si : P), mainly because of industrial purposes. Extrinsic conduction associated with such impurity states is a standard issue for solid-state physics, and transport effects which depend on the impurity concentration are also well studied. For small concentrations, these impurity states are localized to the underlying lattice, but an insulator–metal transition occurs at zero temperature as the impurity concentration increases. The optical properties are very sensitive to the defect/impurity concentration, so that one can follow very precisely concentration-induced metal–insulator transitions.

This survey is far from being exhaustive. The quoted references should allow, as "Further reading," to widen the reader's perspectives on this vast topic.

Acknowledgments

This work was financially supported by the Schweizerische Nationalfonds zur Förderung der Wissenschaftlichen Forschung. The author wishes to thank G Caimi and A Perucchi for fruitful discussions.

See also: Insulators, Optical Properties of; Nanostructures, Optical Properties of; Optical Properties of Materials; Semiconductors, Electronic Transitions in; Semiconductor Optics.

PACS: 78.20. – e; 78.40. – q; 78.30. – j

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. Philadelphia: Holt-Saunders International Editors.

Bassani GF and Pastori-Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Pergamon.

Dressel M and Grüner G (2002) *Electrodynamics of Solids*. Cambridge: Cambridge University Press.

Gruner G (1994) *Density Waves in Solids*. Addison-Wesley.

Jackson JD (1975) *Classical Electrodynamics*. New York: Wiley.

Klingshirn CF (1995) *Semiconductor Optics*. Berlin: Springer.

Kuzmany H (1998) *Solid-State Spectroscopy*. Berlin: Springer.

Mott NF (1990) *Metal–Insulator Transition*. London: Taylor and Francis.

Mott NF and Davis EH (1979) *Electronic Processes in Non-Crystalline Materials*. Oxford: Clarendon.

Pankove JI (1971) *Optical Processes in Semiconductors*. Prentice-Hall.

Seeger K (1997) *Semiconductor Physics*. Berlin: Springer.

Vescoli V, *et al.* (2000) *European Physics Journal B* 13: 503.

Vescoli V, *et al.* (2000) *Physical Review Letters* 84: 1272.

Wachter P (1979). *Handbook on the Physics and Chemistry of Rare Earth*. ch. 19, p. 507. North-Holland.

Wooten F (1972) *Optical Properties of Solids*. New York: Academic Press.

Yu PY and Cardona M (1996) *Fundamentals of Semiconductors*. Berlin: Springer.

Nomenclature

E_{gap}	energy gap (K or cm^{-1})
J_{cv}	joined density of states
$R(\omega)$	reflectivity
$\varepsilon(\omega)$	dielectric function
$\sigma_1(\omega)$	real part of optical conductivity ($(\Omega\,cm)^{-1}$)
ω	frequency in wave number ($1\,eV = 8.065\,48 \times 10^3\,cm^{-1}$)

Semi-Empirical and Empirical Calculations

M-H Whangbo, North Carolina State University, Raleigh, NC, USA

Introduction

In understanding the relationships between crystal structures and physical properties of solid-state materials, it is essential to know their electronic structures. Electronic structure calculations can be carried out either at first principles or at semi-empirical levels. What level of calculations is appropriate for a system under investigation depends on what kind of answers one seeks from calculations. For quantitative predictions, there is no substitute for first-principles electronic structure calculations. For the purpose of understanding qualitative trends, however, semi-empirical electronic structure calculations can be sufficient and expedient.

Electronic structure calculations for a many-electron system lead eventually to an effective one-electron Hamiltonian H^{eff} and its matrix representation. To construct this Hamiltonian, one needs the one-electron wave functions to be determined as eigenfunctions of H^{eff}. As a consequence, it is necessary to construct H^{eff} in terms of trial one-electron wave functions and improve H^{eff} using its eigenfunctions iteratively until a self-consistent-field (SCF) convergence is achieved. In calculating the matrix representation of H^{eff}, first-principles electronic structure calculations execute SCF adjustments of H^{eff} without simplifying approximations, while semi-empirical calculations introduce such approximations. Semi-empirical methods are divided into two groups, that is, one that requires SCF adjustments of H^{eff} and the other that does not.

The crudest, but the most useful and versatile, semi-empirical method of electronic structure calculations is the extended Hückel tight-binding (EHTB) method. This method does not take electron–electron repulsion into consideration, so that it does not require any SCF adjustment of H^{eff}. Despite this simplicity, the EHTB method has been indispensable in uncovering structure–property relationships in all kinds of materials, from molecules to solids, from organic to inorganic compounds. The EHTB method is also indispensable in understanding the trends in spin exchange interactions and magnetic structures of various magnetic solids, although these properties are a manifestation of electron correlation that is neglected in the EHTB method. This article examines the essential features of the EHTB method and discusses why this crude method can be used to describe the electronic structures of not only metallic but also magnetic insulating solids.

Extended Hückel Tight-Binding Method

For a discrete molecule with a set of valence atomic orbitals $\{\chi_1, \chi_2, \chi_3, \ldots, \chi_m\}$, the molecular orbitals ψ_i $(i = 1,2, \ldots, m)$ are expressed as linear combinations of these atomic orbitals:

$$\psi_i = \sum_{\mu=1}^{m} C_{\mu i}\chi_\mu$$

The molecular orbitals ψ_i are the eigenfunctions of the effective one-electron Hamiltonian H^{eff}:

$$H^{\text{eff}}\psi_i = \mathrm{e}_i\psi_i$$

In the EHTB method, the explicit form of H^{eff} is not specified but its matrix representation $H_{\mu\nu} = \langle \chi_\mu | H^{\text{eff}} | \chi_\nu \rangle$, in atomic orbital basis, is defined semi-empirically. The diagonal element $H_{\mu\mu}$ is approximated by the valence state ionization potential (VSIP) of the atomic orbital χ_μ, that is, $H_{\mu\mu} = -\text{VSIP}$. The off-diagonal elements $H_{\mu\nu}$ $(\mu \neq \nu)$ are approximated by the Wolfsberg–Helmholz formula, $H_{\mu\nu} = KS_{\mu\nu}(H_{\mu\mu} + H_{\nu\nu})/2$, where $S_{\mu\nu}$ is the overlap integral $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$ and $K = 1.75$. In the weighted Wolfsberg–Helmholz approximation, the coefficient K is replaced with another coefficient K', which is given by $K' = K + \Delta^2 + \Delta^4(1 - K)$, where $\Delta = (H_{\mu\mu} - H_{\nu\nu})/(H_{\mu\mu} + H_{\nu\nu})$. The weighted Wolfsberg–Helmholz approximation is used to reduce the extent of counter-intuitive orbital mixing. The energies e_i and the atomic orbital coefficients $C_{\mu i}$ of the molecular orbitals ψ_i are obtained by solving the set of simultaneous equations:

$$\sum_{\mu=1}^{m} (H_{\mu i} - \mathrm{e}_i S_{\mu i})C_{\mu i} = 0, \quad i = 1, 2, \ldots, m$$

To calculate the electronic structure of an extended solid having a set of valence atomic orbitals $\{\chi_1, \chi_2, \chi_3, \ldots, \chi_m\}$ per unit cell, it is necessary to define the corresponding set of Bloch orbitals $\{b_1(\boldsymbol{k}), b_2(\boldsymbol{k}), b_3(\boldsymbol{k}), \ldots, b_m(\boldsymbol{k})\}$ for any given wave vector point $\boldsymbol{k}$:

$$b_\mu(\boldsymbol{k}) = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}} \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{R})\chi_\mu(\boldsymbol{r} - \boldsymbol{R}), \quad \mu = 1, 2, \ldots, m$$

where N is the number of unit cells in the magnetic solid, $\boldsymbol{R}$ is the lattice vector, and $\chi_\mu(\boldsymbol{r} - \boldsymbol{R})$ is the atomic orbital χ_μ located at the lattice vector $\boldsymbol{R}$. Then,

the crystal orbitals $\psi_i(\boldsymbol{k})$ of an extended solid are written as linear combinations of the Bloch orbitals:

$$\psi_i(\boldsymbol{k}) = \sum_{\mu=1}^{m} C_{\mu i}(\boldsymbol{k}) b_\mu(\boldsymbol{k})$$

The crystal orbitals $\psi_i(\boldsymbol{k})$ are the eigenfunctions of the effective one-electron Hamiltonian H^{eff}: $H^{\text{eff}}\psi_i(\boldsymbol{k}) = \mathrm{e}_i(\boldsymbol{k})\psi_i(\boldsymbol{k})$ Once the matrix elements $H_{\mu\nu}(\boldsymbol{k}) = \langle b_\mu(\boldsymbol{k})|H^{eff}|b_\nu(\boldsymbol{k})\rangle$ and $S_{\mu\nu}(\boldsymbol{k}) = \langle b_\mu(\boldsymbol{k})|b_\nu(\boldsymbol{k})\rangle$, in Bloch orbital basis, are determined using the EHTB approximation, the energies $\mathrm{e}_i(\boldsymbol{k})$ and the Bloch orbital coefficients $C_{\mu i}(\boldsymbol{K})$ of the crystal orbitals $\psi_i(\boldsymbol{k})$ are calculated by solving the set of simultaneous equations:

$$\sum_{\mu=1}^{m}(H_{\mu i}(\boldsymbol{k}) - \mathrm{e}_i(\boldsymbol{k})S_{\mu i}(\boldsymbol{k}))C_{\mu i}(\boldsymbol{k}) = 0, \quad i = 1, 2, \ldots, m$$

To fully specify the electronic structure of a given solid, one needs to solve the above equation for a set of $\boldsymbol{k}$ points lying in the first Brillouin zone (FBZ), which gives rise to energy bands.

Atomic Orbital Parameters

In EHTB calculations, the valence atomic orbitals χ_u are approximated by Slater-type orbitals (STOs). A single-zeta (SZ) STO, χ_μ, is defined by

$$\chi_\mu \propto \boldsymbol{r}^{n-1}\exp(-\zeta \boldsymbol{r})Y(\theta, \phi)$$

where n is the principal quantum number, ζ is the exponent, and $Y(\theta, \phi)$ is the spherical harmonics. In a double-zeta (DZ) STO, a linear combination of two exponential functions is used:

$$\chi_\mu \propto \boldsymbol{r}^{n-1}[c_1\exp(-\zeta_1 \boldsymbol{r}) + c_2\exp(-\zeta_2 \boldsymbol{r})]Y(\theta, \phi)$$

Thus, the semi-empirical parameters of EHTB calculations are the ζ values for SZ STOs, the ζ_1, ζ_2, c_1, and c_2 values for DZ STOs, and the VSIP values of the valence atomic orbitals. These values can be taken from results of atomic electronic structure calculations using the Hartree–Fock method.

In general, results of EHTB calculations depend more sensitively on the exponents of STOs than on their VSIPs. For compounds consisting of transition metal elements M and main group elements L, it was customary to represent the *d*-orbitals of M by DZ STOs, the *s*/*p* orbitals of M by SZ STOs, and the *s*/*p* orbitals of L by SZ STOs (hereafter referred to as DZ/SZ calculations). However, a better agreement with experiment and with first-principles electronic structure calculations is achieved with EHTB calculations in which the *s*/*p* orbitals of main group elements L are also represented by DZ STOs (hereafter referred to as DZ/DZ calculations), regardless of whether a system under consideration is metallic or magnetic insulating. For conducting salts of organic donor molecules, which consist of only main group elements, use of DZ STOs for the main group elements also provides a better description of their electronic structures than does that of SZ STOs.

Metallic versus Magnetic Insulating States of Extended Solids

A metallic compound has one or more partially filled bands. For example, consider a one-dimensional (1D) chain with one electron and orbital per site (**Figure 1a**), which has a half-filled band. In the metallic state, the levels of the half-filled band below the Fermi level are each doubly filled while those above the Fermi level are each empty as shown in **Figure 1b**. Thus, the wave vectors of the FBZ are divided into two sets, those leading to the filled band levels and those leading to the empty levels. The boundary surface between these two sets of wave vectors are known as the Fermi surface. The transport properties of a metal are mainly governed by the

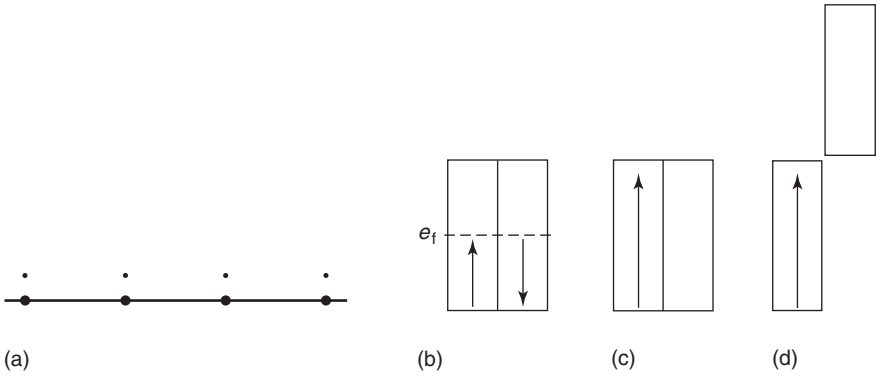


Figure 1 Metallic and magnetic insulating states of a 1D chain with one electron and one orbital per site: (a) schematic representation of a 1D chain; (b) metallic state; (c) magnetic insulating state in non-spin-polarized representation; and (d) magnetic insulating state in spin-polarized representation.

electrons around the Fermi level and hence around the Fermi surface. The charge density wave instability of a low-dimensional metal, which is analogous to the Jahn–Teller instability of a discrete molecular system, is likely to occur when its Fermi surface is nested.

A magnetic insulating state also possesses one or more partially filled bands. For example, in the magnetic insulating state of a half-filled band system, all the band levels are each singly occupied as shown in **Figure 1c**. From the viewpoint of spin-polarized representation, the latter is equivalent to filling all the up-spin band levels singly filled while leaving all the down-spin band levels unoccupied. In spin-polarized electronic band structure calculations for this formally ferromagnetic insulating state, the up-spin and down-spin bands become split in energy (**Figure 1d**). The metallic and magnetic insulating states are similar in that they possess one or more partially filled bands, but they differ in the way the band levels are occupied. Given the bandwidth of the 1D chain as W and the on-site repulsion as U, the metallic state is more stable than the magnetic insulating state if $W > U$, while the reverse is true if $W < U$.

Strengths, Weaknesses, and Utility

EHTB calculations cannot be used to predict either the optimum structure of a system or the relative energies of a system with different electron configurations. For systems of known geometry, however, EHTB calculations provide invaluable information about their electronic properties. In EHTB calculations, the total energy of a system is given by the sum of its occupied orbital energies. Thus, EHTB calculations predict that a solid with a partially filled band is always metallic, in disagreement with experiment. If EHTB calculations give a narrow partially filled band for a solid, one should expect that the solid might be a magnetic insulator rather than a metal. Similarly, for a molecule with two electrons to fill its HOMO and LUMO, EHTB calculations predict that the closed-shell singlet state is always more stable than the triplet state, in disagreement with experiment. When EHTB calculations give a small HOMO–LUMO gap for a molecule, one should expect that the ground state of the molecule might be a triplet state (see below).

Results of EHTB calculations do not depend on the number of electrons in a system, because electron–electron repulsion is not included in the EHTB method. This apparent shortcoming provides several advantages that no first-principles theory can ever match:

1. The "frontier" bands (i.e., the highest-occupied and lowest-unoccupied bands) of a complex system can be well approximated by that of its appropriate fragment. For instance, in an organic conducting salt (BEDT-TTF)$_2$X with mononegative anion X^- (BEDT-TTF–bis(ethylenedithio)tetrathiafulvalene), layers of the (BEDT-TTF)$_2^+$ cations alternate with layers of the X^- anions (**Figure 2**). The frontier bands of (BEDT-TTF)$_2$X are associated with the cation layers so that the electronic structure of a (BEDT-TTF)$_2$X salt is well approximated by that of an isolated cation layer. Similarly, in the monophosphate tungsten bronze, $(PO_2)_4(WO_3)_8$, the corner-sharing WO_6 octahedral layers are separated by PO_4 tetrahedra (**Figure 3**). The frontier bands of $(PO_2)_4(WO_3)_8$ are associated with the WO_6 octahedral layers so that the electronic structure of $(PO_2)_4(WO_3)_8$ is well approximated by that of an isolated WO_6 octahedral layer.
2. The local electronic structure of a transition metal ion M surrounded with n main group ligand atoms L in its first coordination sphere is well

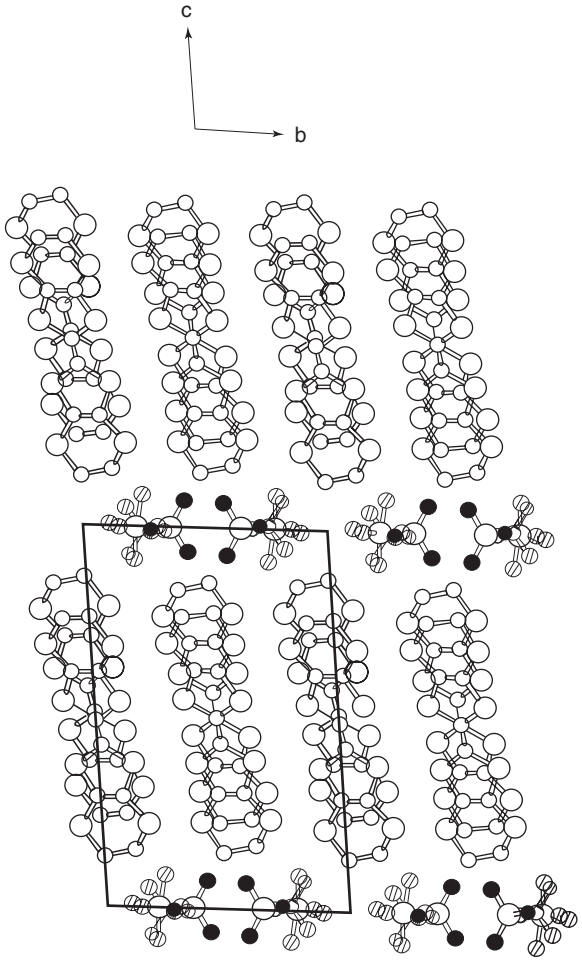


Figure 2 Projection view of the organic conducting salt, β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$, showing the alternation of the [(BEDT-TTF)$_2$]$^+$ cation and the [SF$_5$CH$_2$CF$_2$SO$_3$]$^-$ anion layers.

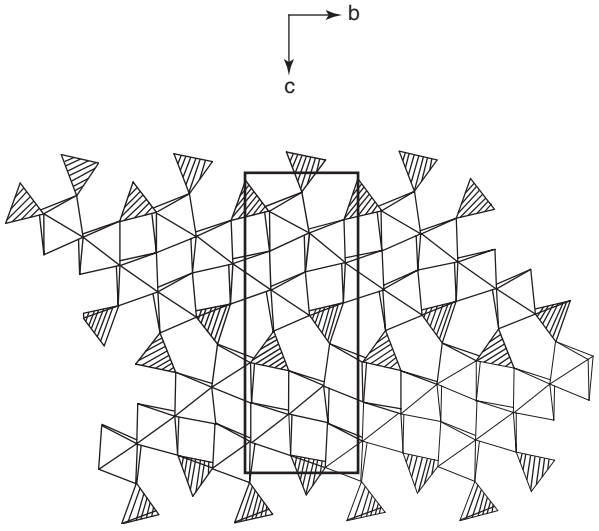


Figure 3 Projection view of the monophosphate tungsten bronze, $(PO_2)_4(WO_3)_8$, showing that the corner-sharing WO_6 octahedral layers are condensed with the PO_4 tetrahedra (shaded).

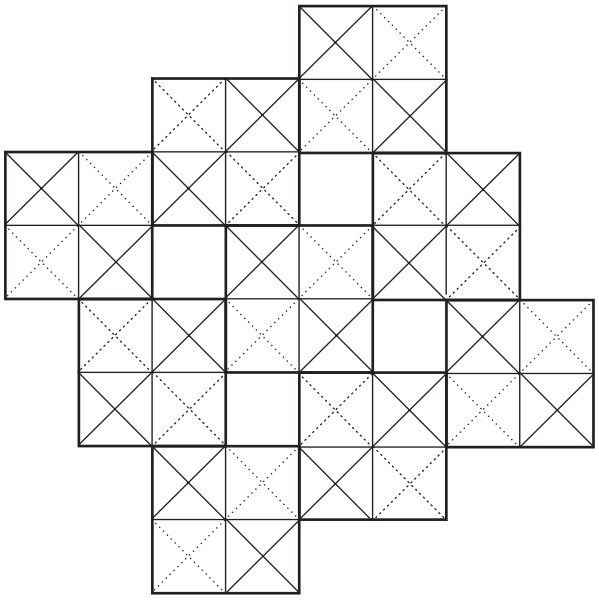


Figure 4 Condensation pattern of the V_4O_9 layers of CaV_4O_9. Each small square with solid (dotted) diagonal lines represents a VO_5 square pyramid having the apical oxygen atom above (below) the basal plane, and the large square made up of four small squares represents a plaquette.

approximated by the *d*-block electronic structure calculated for an isolated complex ML_n. For the purpose of electron counting in the *d*-block levels, each ligand is treated as a closed shell anion. For example, the magnetic oxide CaV_4O_9 is made up of edge- and corner-sharing VO_5 square pyramids (**Figure 4**), and the vanadium and oxygen atoms of CaV_4O_9 have the oxidation states $+4$ and -2, respectively. Thus, a spin monomer (i.e., a structural unit containing one spin site) is represented by the $(VO_5)^{6-}$ cluster, while a spin dimer (i.e., a structural unit containing two spin sites) is represented by the $(V_2O_8)^{8-}$ cluster for edge-sharing square pyramids and by the $(V_2O_9)^{10-}$ cluster for corner-sharing square pyramids (**Figure 5**). The highly charged anion state of these species does not present any computational problem in EHTB calculations, unlike the case of first principles and semi-empirical calculations requiring SCF adjustments of H^{eff}.

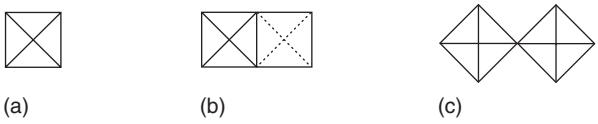


Figure 5 Spin monomer and dimer units of CaV_4O_9. (a) Spin monomer $(VO_5)^{6-}$; (b) edge-sharing spin dimer $(V_2O_8)^{8-}$; (c) corner-sharing spin dimer $(V_2O_9)^{10-}$.

3. EHTB calculations are simple and hence enable the study of molecular and extended solids that are too large to study using first-principles electronic structure calculations.

In general, there are two ways of introducing electron correlation in electronic structure calculations, namely, the configuration interaction (CI) wave function and the density-functional theory (DFT) methods. In the CI wave function method, the Hartree–Fock equations solve the exact Hamiltonian (i.e., the molecular Hamiltonian) with approximate many-body wave functions (i.e., Slater determinants). The exact solution can be achieved through systematic improvements in the form of many-body wave functions such as CI wave functions. In this approach, atomic orbitals and hence molecular orbitals play the role of basis functions used to improve the form of many-body wave functions. In the DFT method, however, approximations are introduced only in the exchange-correlation operator. The density-functional equations solve an approximate many-body Hamiltonian with exact wave functions, and DFT approaches the exact solution by improving the exchange-correlation operator. Therefore, unlike the case of the CI wave function method, orbital concepts as practiced by chemists remain valid in DFT. Results of DFT electronic structure calculations are readily interpreted by employing the concepts of orbital interactions on the basis of EHTB calculations.

Low-Dimensional Metallic Compounds

A large number of low-dimensional metallic compounds, which include transition-metal oxides, transition-metal chalcogenides, organic conducting salts,

etc., have been examined by EHTB calculations. In the following, the band dispersion relations, the density of states, and the Fermi surfaces for two examples that illustrate the utility of EHTB calculations are calculated. These calculations get the Fermi surface topology approximately correct, but underestimate the effective mass.

The cation layers of the organic conducting salt β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$ are well separated (**Figure 2**). In each cation layer, there are four BEDT-TTF molecules per unit cell. The band dispersion relations and the density of states calculated for a single cation layer using the SZ STOs for all the atoms of BEDT-TTF molecules are shown in **Figures 6a** and **6b**, respectively. For simplicity, only the four highest occupied bands, derived mainly from the HOMOs of the BEDT-TTF molecules, are shown. With the oxidation state $[(\text{BEDT-TTF})_2]^+$, there are six electrons to fill these four bands leading to partially filled bands. The Fermi surfaces associated with the partially filled bands, shown in **Figure 6c**, consist of four wavy lines. Therefore, the SZ STO calculations predict that β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$ is a 1D metal. **Figures 7a–7c** show the dispersion relations, the density of states, and the Fermi surfaces calculated for a single cation layer using the DZ STOs for the carbon and sulfur atoms of BEDT-TTF. It is noted that the bandwidth from the DZ STO calculations is about twice as wide as that from the SZ STO calculations. The Fermi surfaces of the DZ STO calculations consist of two wavy lines and a closed pocket centered at the X point. Namely, the DZ STO calculations predict that β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$ has both 1D and 2D metallic character, in agreement with the experiment.

Figure 3 shows that a unit cell of the monophosphate tungsten bronze, $(PO_2)_4(WO_3)_8$, has two corner-sharing WO_6 octahedral layers separated by PO_4 groups. **Figures 8a** and **8b** present the band dispersion relations and the density of states obtained for a single octahedral layer from the DZ/SZ calculations, respectively. For simplicity, only the t_{2g}-block bands are shown. The Fermi surfaces associated with the partially filled bands are shown in **Figure 8c**. The corresponding plots obtained for a single WO_6 octahedral layer from the DZ/DZ calculations are presented in **Figure 9**. The bandwidth from the DZ/DZ calculations is about twice as wide as that from the

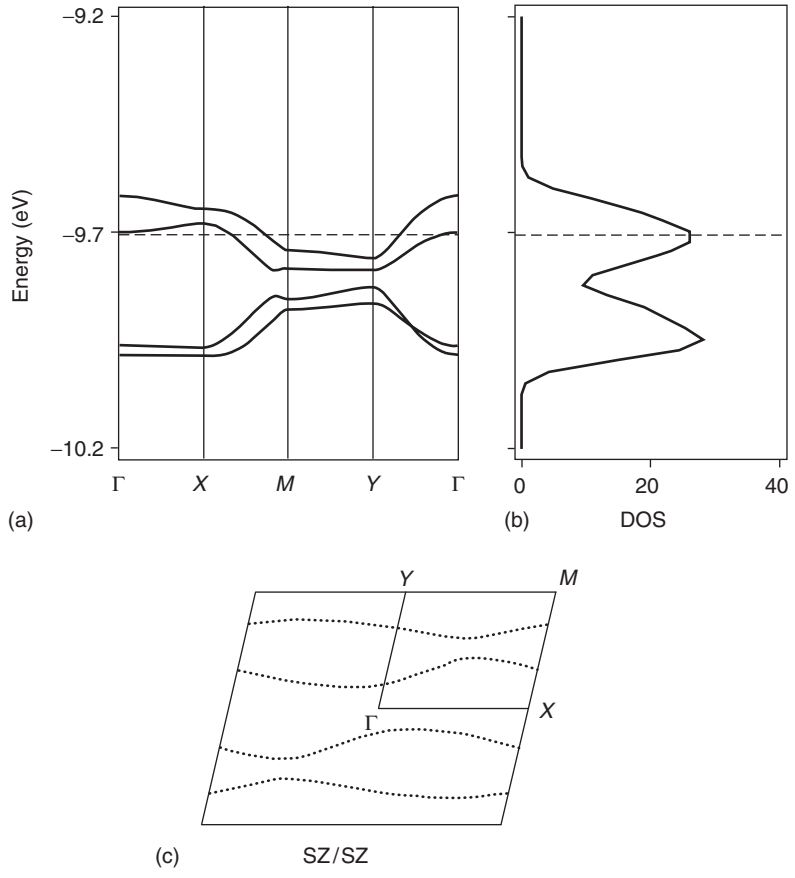


Figure 6 Electronic band structure calculated for a cation layer of β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$ using the SZ STOs. (a) Band dispersion relations; (b) density of states; (c) Fermi surface.

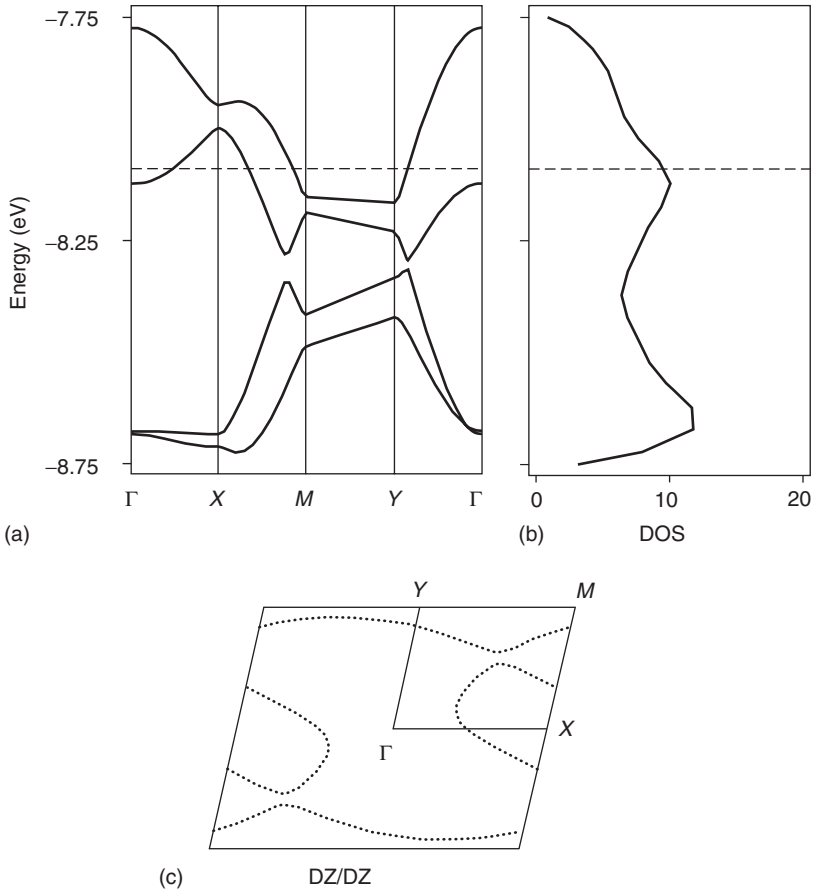


Figure 7 Electronic band structure calculated for a cation layer of β''-(BEDT-TTF)$_2$SF$_5$CH$_2$CF$_2$SO$_3$ using the DZ STOs. (a) Band dispersion relations; (b) density of states; (c) Fermi surface.

DZ/SZ calculations, but the Fermi surfaces from the two calculations are similar. The band dispersion relations, the density of states, and the Fermi surfaces obtained for the 3D lattice of $(PO_2)_4(WO_3)_8$ using the DZ/DZ calculations, presented in **Figure 10**, are very close to those obtained for a single WO_6 octahedral layer using the DZ/DZ calculations. This is so because the interaction between the adjacent WO_6 octahedral layers is weak.

Spin Dimer Analysis of Spin Exchange Interactions

Low-energy excitations of a magnetic solid with localized spins are probed by magnetic susceptibility and neutron scattering measurements, and are analyzed using a spin Hamiltonian defined in terms of a set of spin exchange parameters J. The nature and values of the spin exchange parameters deduced from this analysis depend on what spin exchange paths one includes in the spin Hamiltonian. Strongly interacting spin exchange paths of a magnetic solid are determined by the overlap between its magnetic orbitals (i.e., singly occupied molecular orbitals). Since the magnetic orbitals are anisotropic in shape, the strength of a spin exchange path is not necessarily determined by the shortness of the distance between the magnetic ions. Therefore, in interpreting results of magnetic susceptibility and inelastic neutron scattering measurements, it is crucial to employ a set of strongly interacting spin exchange paths chosen on the basis of proper electronic structure considerations. Otherwise, one may end up interpreting the magnetic properties of a system in terms of a spin lattice model quite irrelevant for the system. Spin dimer analyses based on extended Hückel tight-binding calculations provide a reliable and expedient means to estimate the relative strengths of spin exchange interactions.

In essence, electron localization is responsible for spin exchange interactions of magnetic solids. Consider that a spin dimer has two equivalent spin sites with one electron and one orbital per site (**Figure 11a**).

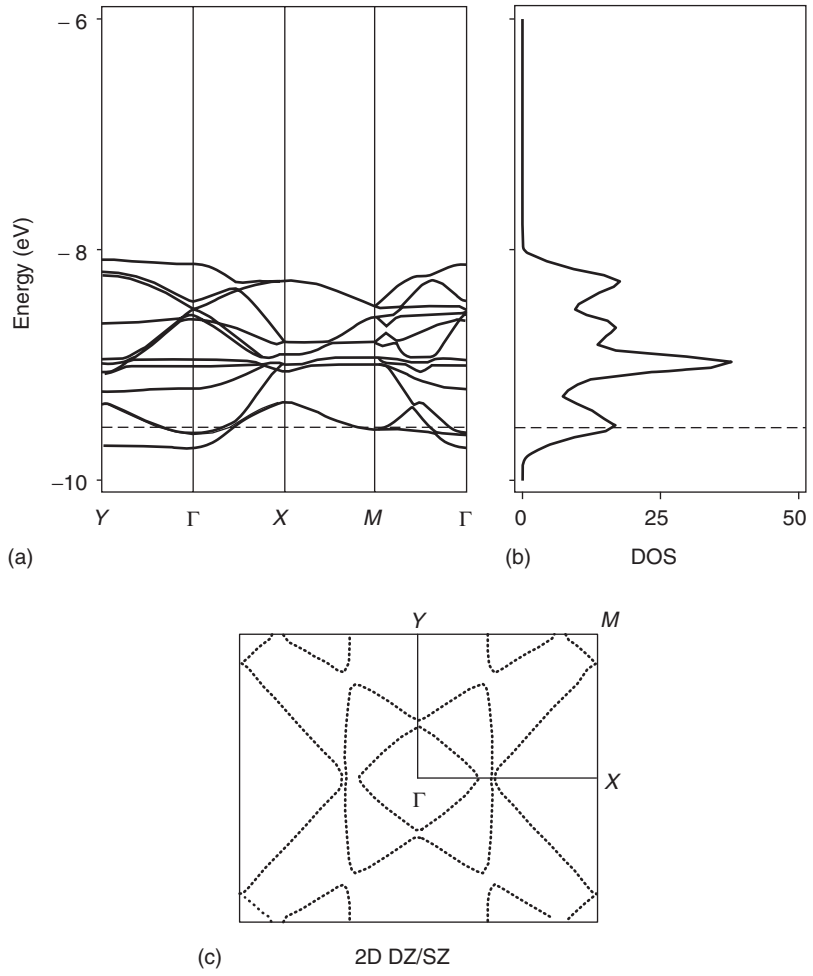


Figure 8 Electronic band structure obtained for a single WO_6 layer of $(PO_2)_4(WO_3)_8$ using the DZ/SZ-STO calculations. (a) Band dispersion relations; (b) density of states; (c) Fermi surface.

The interaction between the two magnetic orbitals χ_1 and χ_2 (representing the spin sites 1 and 2, respectively) leads to two molecular levels ψ_1 and ψ_2 with energies e_1 and e_2, respectively. Three electron configurations of interest for this spin dimer are the triplet configuration Φ_T and the singlet configurations, Φ_1 and Φ_2 (**Figures 11b–11d**). When chemical bonding interaction between the spin sites is weak, the orbital energy difference $\Delta e = e_2 - e_1$ is small, and the two electrons of the dimer are considered as localized. In this case, the two singlet configurations Φ_1 and Φ_2 become close in energy, so the CI between Φ_1 and Φ_2 becomes strong and the ground singlet state Φ_S is described by the linear combination:

$$\Phi_S = C_1\Phi_1 - C_2\Phi_2$$

where the mixing coefficients C_1 and C_2 have the same sign and are similar in magnitude. Using this CI wave function, the energy difference between the triplet and singlet electronic states, $\Delta E = E_S - E_T$, is expressed as

$$\Delta E = 2K_{12} - \frac{(\Delta e)^2}{U_{\text{eff}}}$$

where K_{12} is the exchange repulsion integral between χ_1 and χ_2, and U_{eff} is the effective on-site repulsion.

If the spin dimer is described by the Heisenberg spin Hamiltonian $\hat{H} = -J\hat{S}_1 \cdot \hat{S}_2$, it is found that the spin exchange parameter J is equal to $E_S - E_T$. Therefore,

$$J = \Delta E = 2K_{12} - \frac{(\Delta e)^2}{U_{\text{eff}}}$$

so that the spin exchange parameter J can be written as $J = J_F + J_{AF}$, where

$$J_F = 2K_{12} > 0$$

$$J_{AF} = -\frac{(\Delta e)^2}{U_{\text{eff}}} < 0$$

In general, the exchange integral K_{12} is a small positive number so that the spin exchange cannot be

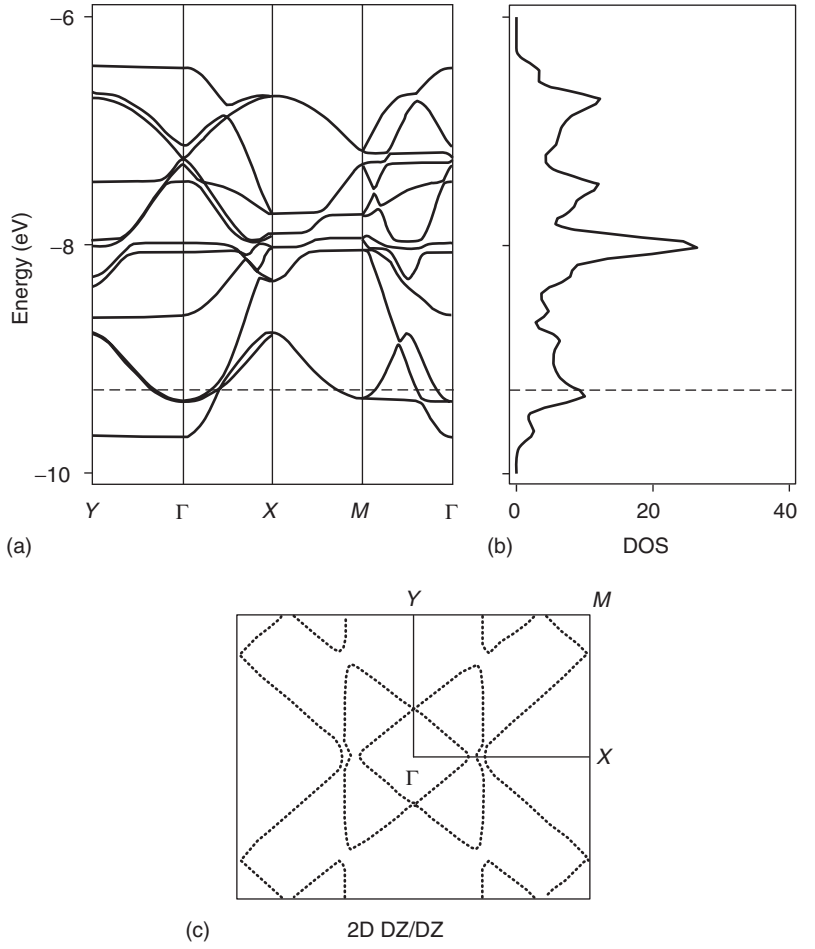


Figure 9 Electronic band structure obtained for a single WO_6 layer of $(PO_2)_4(WO_3)_8$ using the DZ/DZ-STO calculations. (a) Band dispersion relations; (b) density of states; (c) Fermi surface.

ferromagnetic (i.e., $J>0$) unless the antiferromagnetic term J_{AF} either vanishes or is very small in magnitude. For various spin exchange paths of a magnetic solid, the U_{eff} term is nearly constant so that the variation of the J_{AF} values is mainly governed by that of the $-(\Delta e)^2$ values. For an antiferromagnetic exchange interaction (i.e., $J<0$), the J_{AF} term dominates over the J_F term. Since the J_F term is a small positive number, trends in spin exchange interactions can be discussed on the basis of the trends in the corresponding $-(\Delta e)^2$ values obtained from EHTB calculations for spin dimers.

When two adjacent spin sites have M unpaired spins, the overall spin exchange parameter J is then described by

$$J = \frac{1}{M^2}\sum_{\mu=1}^{M}\sum_{\nu=1}^{M} J_{\mu\nu}$$

Therefore, the trends in spin exchange parameters can be discussed in terms of the sum of the spin orbital interaction energy squares:

$$\left\langle (\Delta e)^2 \right\rangle = \frac{1}{M^2}\sum_{\mu=1}^{M}\sum_{\nu=1}^{M} (\Delta e_{\mu\nu})^2$$

Then, the antiferromagnetic component of the overall spin exchange parameter J of the spin dimer can be written as:

$$J_{AF} = -\frac{\left\langle (\Delta e)^2 \right\rangle}{U_{eff}}$$

To illustrate the utility of the spin dimer analysis based on EHTB calculations, two examples are considered. The magnetic properties of 3D perovskites $KNiF_3$ and $KCuF_3$ as well as the layered perovskites K_2NiF_4, K_2CuF_4, and La_2CuO_4 were extensively studied both theoretically and experimentally. These compounds are made up of corner-sharing ML_6 octahedra. As summarized in **Figure 12a**, the experimental J values of these compounds are almost

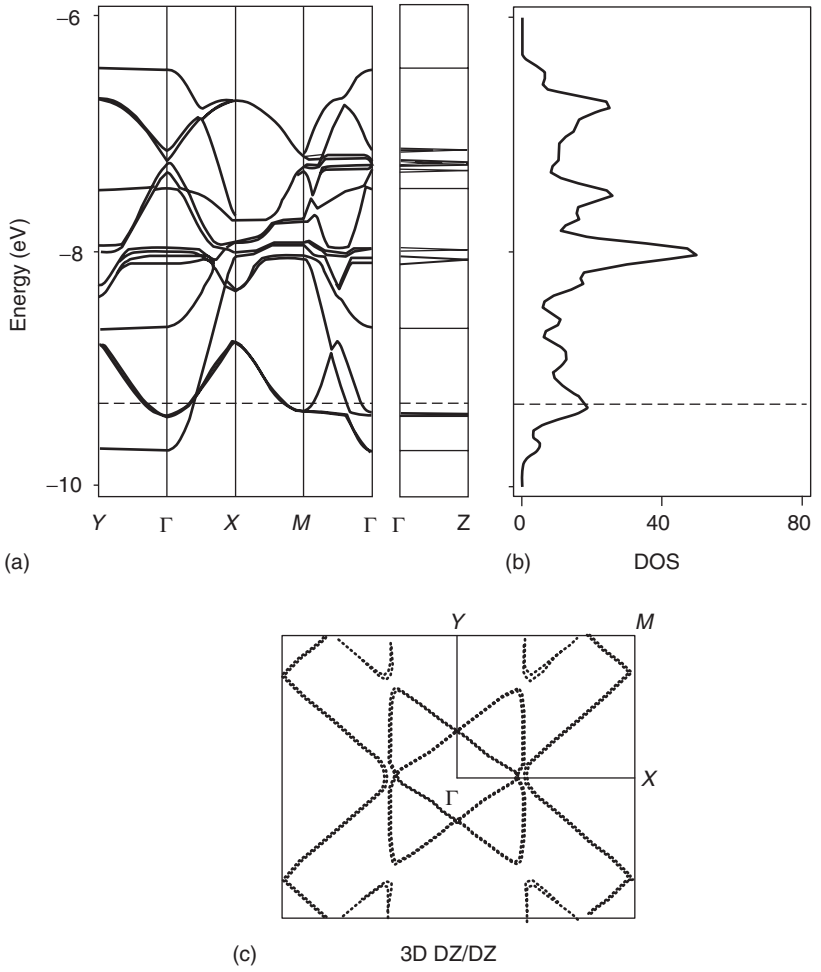


Figure 10 Electronic band structure obtained for the 3D lattice of $(PO_2)_4(WO_3)_8$ using the DZ/DZ-STO calculations. (a) Band dispersion relations; (b) density of states; (c) Fermi surface.

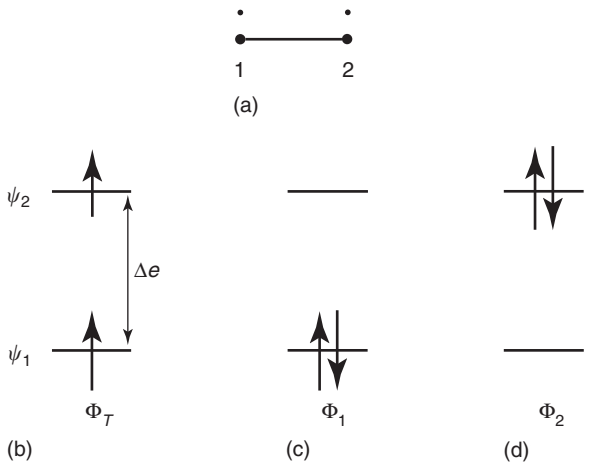


Figure 11 Three electronic configurations of a dimer with one electron and one orbital per site: (a) schematic representation of a dimer; (b) triplet state $(\psi_1)^1(\psi_2)^1$; (c) singlet configuration $(\psi_1)^2$; (d) singlet configuration $(\psi_2)^2$.

quantitatively reproduced by the J values obtained from first-principles electronic structure calculations. **Figure 12b** shows that the experimental J values and the $\langle(\Delta e)^2\rangle$ values obtained from EHTB calculations have a good linear relationship. The vanadium oxide, CaV_4O_9, consists of V_4O_9 layers that are made up of VO_5 square pyramids containing V^{4+} (d^1) ions as shown in **Figure 4**. The spin exchange parameters of CaV_4O_9 are known for four exchange paths from the neutron inelastic scattering and magnetic susceptibility studies and from first-principles electronic structure calculations (**Table 1**). From the calculated $(\Delta e)^2$ for these exchange paths, one can estimate the exchange parameters using the relationship $-(\Delta e)^2/U_{\text{eff}}$, where the U_{eff} value of 1.587 eV is chosen to reproduce the strongest J value. Clearly, the relative strengths of the four experimental J values are very well reproduced by the $-(\Delta e)^2/U_{\text{eff}}$ values (**Table 1**).

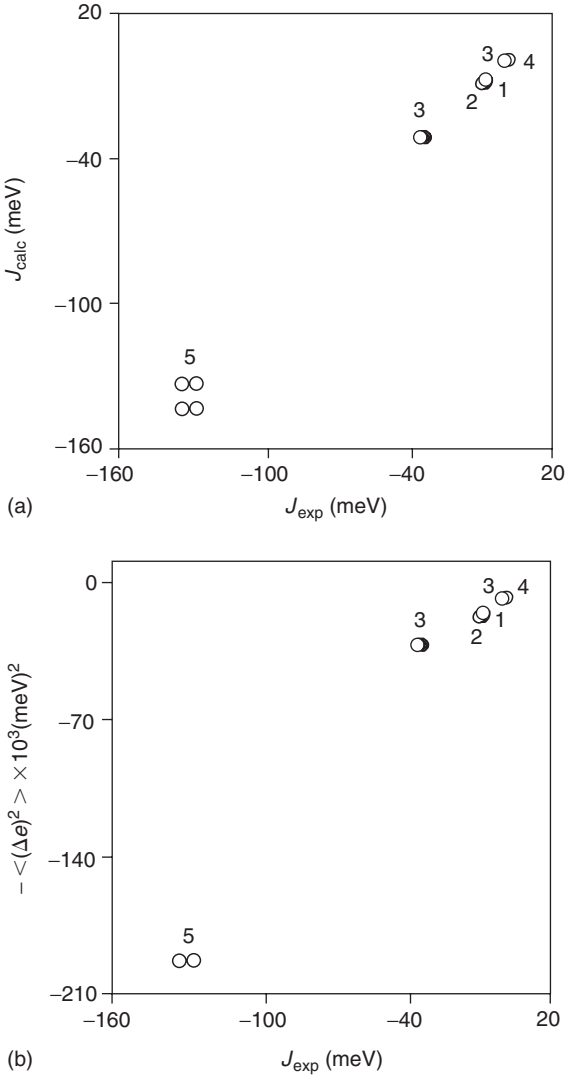


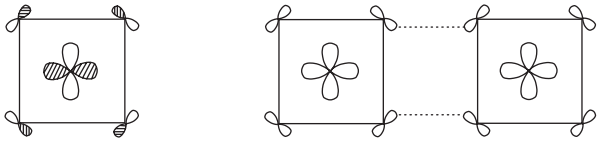
Figure 12 Comparison of the experimental spin exchange parameters J_{exp} of perovskite magnetic solids with the calculated values J_{calc} obtained from first-principles electronic structure computations and the $\langle(\Delta e)^2\rangle$ values obtained from EHTB calculations: (a) J_{exp} vs. J_{calc}; (b) J_{exp} vs. $-\langle(\Delta e)^2\rangle$ (1: $KniF_3$, 2: K_2NiF_4, 3: $KCuF_3$, 4: K_2CuF_4, and 5: La_2CuO_4).

Importance of Super-Superexchange Interactions

For extended magnetic solids in which magnetic transition metal ions M are surrounded with main elements L, the spin exchange interactions between adjacent metal ions take place via M–L–M superexchange (SE) and M–L…L–M super-superexchange (SSE) paths. On a qualitative level, the trends in the signs and magnitudes of SE interactions are discussed in terms of Goodenough–Kanamori rules. In the past, SSE interactions have been neglected because there was no simple qualitative rule by which to recognize when their strengths were significant. **Figure 13a** shows that the magnetic orbital of a V^{4+} (d^1) ion is given by the V $3d_{xy}$ orbital that makes π-antibonding interactions with the p orbitals of the four ligands L in the basal plane. As depicted in **Figure 13b**, the strength of an SSE interaction between two such magnetic orbitals increases with increasing the overlap between the p-orbital tails of the L…L contacts in the two V–L…L–V paths. This interaction becomes stronger as the L_4 ring formed from the two L…L contact units is more planar and rectangular, as the basal planes of the two spin monomers are more coplanar, and as the L…L contacts are shorter and lie within the van der Waals distance. When the latter conditions are met, an SSE interaction can be stronger than an SE interaction, as found for $(VO)_2P_2O_7$. A similar situation is also found for spin exchange interactions involving Cu^{2+} (d^9) ions. The magnetic orbital of a Cu^{2+} (d^9) ion is the Cu $3d_{x2-y2}$ orbital that makes sigma antibonding interactions with the p orbitals of the four ligands L (**Figure 14a**). As depicted in **Figure 14b**, the strength of an SSE interaction between two such magnetic orbitals increases with increasing the overlap between the two p-orbital tails of the L…L contact in one Cu–L…L–Cu path. This interaction becomes stronger as both ∠Cu–L…L bond angles become

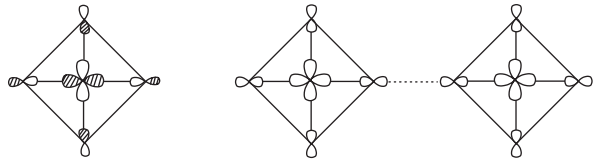
Table 1 Experimental and calculated spin exchange parameters of CaV_4O_9 (meV) for the corner- and edge-sharing exchange paths

	Intra plaquette		*Inter plaquette*	
	Edge	*Corner*	*Edge*	*Corner*
(a) Experiment				
Neutron scattering	−5.76	−1.25	−5.76	−14.73
Magnetic susceptibility	−9.3	−3.7	−9.6	−14.2
(b) First-principles electronic structure calculations				
LDA + U	−7.67	−7.84	−5.34	−12.75
LSDA	−8.9	−6.5	−1.1	−23.8
SCAD	−9.7	−3.9	−12.5	−19.3
(c) EHTB calculations				
$-(\Delta e)^2/U_{eff}$	−6.36	−4.18	−5.54	−14.73



(a) (b)

Figure 13 (a) Magnetic orbital of a VL_5 square pyramid, in which the V $3d_{xy}$ orbital makes π-antibonding with the ligand p orbitals on the basal plane; (b) arrangement of two adjacent magnetic orbitals in an SSE interaction having two V–L…L–V paths.



(a) (b)

Figure 14 (a) Magnetic orbital of a CuL_4 square plane, in which the Cu $3d_{x2-y2}$ orbital makes σ-antibonding with the ligand p orbitals on the basal plane; (b) arrangement of two adjacent magnetic orbitals in an SSE interaction having a linear Cu–L…L–Cu path.

larger and as the L…L distance becomes shorter and lies within the van der Waals distance. For instance, in the magnetic oxides, $CuWO_4$ and $CuMoO_4$-III, the strongest spin exchange interactions are the SSE interactions involving the most linear Cu–O…O–Cu paths, in which the O…O distances of these paths are in the range of 2.4 Å, which is considerably shorter than the van der Waals distance, and the two ∠Cu–O…O angles of these paths are identical and are close to 165°.

As described above, an SSE interaction can be stronger than any SE interaction. Therefore, assigning strongly interacting spin exchange paths of magnetic solids should be based on appropriate electronic structure considerations. When a magnetic solid consists of other magnetic ions with more than one unpaired spin, the relative strengths of SE and SSE interactions should be estimated with appropriate electronic structure calculations, for which EHTB calculations have been found invaluable.

Concluding Remarks

Although the semi-empirical EHTB method is based on very crude approximations, it facilitates one to extract invaluable information about metallic as well as magnetic insulating solids of known crystal structures. It is also useful in interpreting results of first-principles electronic structure calculations in terms of simple chemical concepts such as overlap, symmetry, and orbital interaction.

Acknowledgments

This work was supported by the Office of Basic Energy Sciences, Division of Materials Sciences, US Department of Energy, under Grant DE-FG02-86ER45259.

See also: Fermi Surface Measurements; Tight-Binding Method in Electronic Structure.

PACS: 71.15.−m; 71.15.Fv; 71.18.+y; 71.20.−b; 71.23.An; 71.45.Lr; 74.70.Kn; 75.00.00; 75.30.Et

Further Reading

Albright TA, Burdett JK, and Whangbo MH (1985) *Orbital Interactions in Chemistry*. New York: Wiley.

Ashcroft NW and Mermin ND (1976) *Solid State Physics*, pp. 176–190. Philadelphia: Holt, Reinhart and Winston.

Canadell E and Whangbo MH (1991) *Chemical Reviews* 91: 965–1034.

Greenblatt M (ed.) (1993) *International Journal of Modern Physics B* 7: 3937–4108.

Harrison N, Rzepniewski E, Singleton J, Gee P, Honold MM, *et al.* (1999) *Journal of Physics: Condensed Matter* 11: 7227–7242.

Ishiguro T and Yamaji K (1990) *Organic Superconductors*. New York: Springer.

Whangbo MH, Koo HJ, and Dai D (2003) Spin exchange interactions and magnetic structures of extended magnetic solids with localized spins: theoretical descriptions on formal, quantitative and qualitative levels. *Journal of Solid State Chemistry*. 176: 417–481.

Williams JM, Ferraro JR, Thorn RJ, Carlson KD, Geiser U, *et al.* (1992) *Organic Superconductors*, pp. 250–283. Englewood Cliffs: Prentice-Hall.

Nomenclature

$b_\mu(\boldsymbol{k})$	Bloch orbital at a wave vector $\boldsymbol{k}$
$C_{\mu i}$	coefficient for the atomic orbital χ_μ in molecular orbital ψ_i
$C_{\mu i}(\boldsymbol{k})$	coefficient of the $b_\mu(\boldsymbol{k})$ in crystal orbital $\psi_i(\boldsymbol{k})$
e_j	molecular orbital energy
$e_i(\boldsymbol{k})$	crystal orbital energy at a wave vector $\boldsymbol{k}$
H^{eff}	effective one-electron Hamiltonian
$H_{\mu\nu}$	interaction energy matrix element $\langle \chi_\mu \vert H^{\text{eff}} \vert \chi_\nu \rangle$
J_{AF}	antiferromagnetic component of a spin exchange parameter J
J_{F}	ferromagnetic component of a spin exchange parameter J
J, $J_{\mu\nu}$	spin exchange parameters between two magnetic orbitals
K_{12}	exchange repulsion integral
$S_{\mu\nu}$	overlap matrix element $\langle \chi_\mu \vert \chi_\nu \rangle$

U	on-site repulsion
U_{eff}	effective on-site repulsion
W	bandwidth
$Y(\theta,\phi)$	spherical harmonics
$\langle(\Delta e)^2\rangle$	sum of the spin orbital interaction energy squares
Δe, $\Delta e_{\mu\nu}$	orbital energy spilt associated with the interaction between magnetic orbitals
ζ	exponent of an exponential function of a slater type atomic orbital
χ_μ	valence atomic orbital
Φ_{i}	electron configuration
Φ_{S}	singlet state
Φ_{T}	triplet state
ψ_i	molecular orbital
$\psi_i(\boldsymbol{k})$	crystal orbital at a wave vector $\boldsymbol{k}$

Sensors

A D'Amico and C Di Natale, University of Roma "Tor Vergata," Rome, Italy

All the living beings of our planet, in different ways, display highly complex integrated biological systems (biosensors) that are used to profitably interact with the environment where they exist. The loss of efficiency of these indispensable biosystems or even their absence (lack), is reflected in a reduction of outliving capacity and, as a consequence, in a slow-down of the evolutionary processes. For this reason, these biosensors are of paramount importance. For example, in the constant effort dedicated to imitate Nature, in order to construct apparatuses of particular importance such as robots to be used for hard and precision works, man tries to construct, using available and updated technologies, sensors able to exceed the performance of the biological analogs.

In the last two decades, along with the developments in microelectronics and microsystems, a variety of sensors for the detection of physical, chemical, and biological measurands have been fabricated, and the effort is still in an evolutionary phase due to the ever-growing technological level. These are used to characterize and control the so-called generalized environment (GE) where living beings grow and develop their intrinsic knowledge. Examples of GE are: atmosphere, lands, waters, drops of biological liquids, cells, vehicles, buildings, outer space, etc.

In designing new sensors or even sensor systems, one should keep in mind the strategy adopted by Nature, especially in the case of evolved animals. Most of the sensing systems are formed by a large number of small sensors. It seems that Nature tends to reduce the dimensions of the single sensor in order to achieve, at small dimensions, one peculiar property and use a matrix of these sensors to gain another property useful at macroscopic level. As an example, consider the eye. The single-photon sensor is made small enough to determine the single-pixel resolution at a given distance; then a huge number of these high-performance sensors are used in order to have the possibility of seeing a rather large sight frame (extended field of view).

New-generation sensors are solid-state devices fabricated by microelectronic technologies able to read the GE and give responses of either the electrical or the optical type, suitable for both successive preprocessing and final actions of monitoring, storing, and control. They can measure physical parameters such as force, acceleration, strain, torque, vibration, position, velocity, pressure, temperature, ion concentration in solutions (Na^+, Ca^{++}, K^+, etc.), and a variety of volatile compounds, such as H_2, CO, CO_2, SO_2, etc. Also utilizing enzymes, antibodies, or chemically sensitive molecules designed *ad hoc*, they can improve sensitivity and resolution versus particular chemical or biological species.

Sensors being highly complex devices, their study and development can only be performed at multidisciplinary level using knowledge ranging from solid-state physics, chemistry, microelectronics, to biology, interfaces, electrochemistry, nanocorrosions, etc. This may be justified by taking into account some of the most important characteristics that must be satisfied by sensors: high sensitivity and resolution, low noise, robustness, compatibility with microelectronic technologies, and noncontaminability; finally, for all applications regarding human bodies, biocompatibility does represent the most important requirement. All these characteristics cannot, in practice, be satisfied, simultaneously, and an acceptable compromise should always be found so that sensors for specific tasks can be designed more easily.

In this article, for the sake of clarity, the following terminology is used:

1. *Physical sensors*: devices able to sense physical quantities.
2. *Chemical sensors*: devices able to sense chemical quantities.
3. *Biological sensors*: devices able to sense biological quantities.

4. *Biosensors*: devices having sensing parts made by biological materials. Here one may distinguish among:

   (a) Chemical biosensors: devices having sensing parts made by biological materials and sensitive to chemicals.
   (b) Biological biosensors: devices having sensing parts made by biological materials and sensitive to biomaterials.
   (c) Physical biosensors: devices having sensing parts made by biological materials and sensitive to physical quantities.

Numerous effects may be derived from the interactions between the chemically interactive material (CIM) and the GE. The most relevant for the CIMs are: heat generation followed by a temperature increase; changes of one of the following parameters: electronic charge, mass, conductivity, refractive index, work function, photon emission. Many transducers available today are able to utilize one or more of the above-mentioned effects.

A variety of CIMs are available for chemical and biosensing, such as: metal oxide semiconductors (SnO_x, TiO_x, Ta_2O_5, IrO_x, WO_2, etc.), metals (Pt, Pd, Ni, Ag, Cr, Sb, K, etc.), ionic conductors (ZrO_2, LaF_x, CaF_x, CeO_2, Nasicon, etc.), polymers (polypyrrole, polyphenylacetylene, cellulose, polyurethane, polycarbonate, porphyrines, phthalocyanines, polysiloxanes, etc.), enzymatic systems (glucose-oxidase, lattosio-oxidase, urease, anti-IgG, anilisteria, etc.).

It is also interesting to mention that a sensor, incorporating an appropriate transducer, can be seen as an energy-to-energy convertor, where the possible energies at the input and output, may be any of the following: electrical, magnetic, optical, chemical, mechanical, thermal, or nuclear. In this context, the following sensor conversion phenomena give us the various possibilities available for the design of new sensing devices: biochemical, bioelectric, biomagnetic, electroacoustic, elastomagnetic, elastoelectric, electrochemical, photochemical, magnetoelectric, magnetooptic, photoelastic, photoelectric, photomagnetic, thermoelectric, thermomagnetic, thermoelastic, and thermooptic.

It is also useful to mention the most important types of measurable quantities which are related to the nine application areas listed below:

1. *Acoustic*: amplitude, phase, polarization, spectra, power, and harmonic distortion.
2. *Biological*: concentration, type of molecules, functionality, and affinity.
3. *Chemical*: concentration, type of molecules, functionality, and affinity.
4. *Electric*: current, charge, potential, electric field, power, phase, polarization, resistivity, conductivity, permittivity, and permeability.
5. *Magnetic*: magnetic field, magnetic flux, permeability, and phase.
6. *Mechanical*: force, moment, torque, stress, strain, viscosity, linear position, angular position, linear velocity, angular velocity, acceleration, mass, density, flow, stiffness, and compliance.
7. *Optical*: amplitude, phase, polarization, refractive index, spectra, power, wave velocity, phase velocity, and harmonic distortion.
8. *Radiation*: frequency, wavelength, intensity, energy, power, and spectra.
9. *Thermal*: temperature, thermal conductivity, thermal resistance, specific heat, and flux.

In order to evaluate the characteristics of a sensor, suitable descriptors must be considered, e.g.:

(a) *Response curve* (*RC*). It represents the output response as a function of the measured quantity applied to its input. For instance, in the case of a chemical sensor based on change in conductivity (G) with respect to the change of the input quantity, it is recommended to use the following notations for the output response: G (conductance), G/G_0 (relative conductance), $(G - G_0)/G_0$ (relative change of the conductance).

(b) *Sensitivity* (*S*). It is defined as the derivative of the response (which is a function of the operating point) with respect to the measured quantity (M). With reference to the three above-mentioned cases, one has $S = \mathrm{d}(G)/\mathrm{d}M$, $S = \mathrm{d}(G/G_0)/\mathrm{d}M$, $S = \mathrm{d}((G - G_0)/G_0)/\mathrm{d}M$.

In the case of a linear response without offset the above three relationships of sensitivity simplify as follows: $S = G/M$, $S = (G/G_0)/M$, $S = ((G - G_0)/G_0)/M$.

In the case of a piecewise linear response, for each of the segments, the sensitivities can be further simplified as $S = \Delta G/\Delta M$, $S = \Delta(G/G_0)/\Delta M$, $S = \Delta((G - G_0)/G_0)/\Delta M$.

(c) *Resolution* (*R*). It is defined as the magnitude of the measured quantity which gives a signal-to-noise ratio equal to 1 and can be expressed in a simplified form as follows

$$R = \text{noise voltage}/S$$

which means that sensors having the same noise present a better resolution when the sensitivity is higher.

In the following, the basic principles of the most important sensors are illustrated through the use of

some figures which schematically indicate (not to scale) their shapes.

Figure 1 shows a pyroelectric sensor, which is sensitive to heat rate delivered during the sensing action. A slice of pyroelectric material, such as $LiTaO_3$ or $LiNbO_3$, is covered by two thin-film metallic contacts, and a CIM is deposited on one of them (e.g., platinum if one would like to detect ammonia). A change in dipole moment induced by external temperature changes develops either an open circuit voltage or a loaded circuit current as output signal. Temperature changes as low as 10^{-6} K are possible using this sensor.

Figure 2 schematically shows a metal oxide semiconductor (MOS) structure, which is sensitive to electronic charges, where, as an example, the gate has been covered by a thin film of palladium. In the presence of hydrogen, through a catalytic action, electrical dipoles grow at the palladium–silicon dioxide interface, which controls the source drain current of the MOS structure. In this way, the output voltage at the drain can be related to the amount of the adsorbed and absorbed hydrogen occurring during the catalytic process itself. About 10^{-9} C represents the minimum charge variation allowed at the gate to obtain the minimum readable signal.

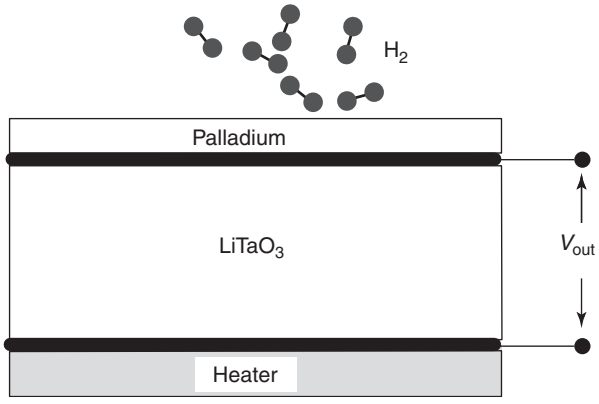


Figure 1 Schematic of a chemical sensor based on the pyroelectric effect. The interaction between airborne hydrogen and palladium results in a variation of temperature of the metal film. The underlying pyroelectric material changes its polarization state according to the temperature variation. As output, a voltage drop across the pyroelectric material is observed. To the nature of the pyroelectric effect, the output voltage is proportional to the derivative of the temperature. The heating system helps to speed up the catalithic reaction between the CIM and the volatile compound.

Figure 3 shows a differential surface acoustic wave sensor which is sensitive to mass changes occurring along the surface wave path. Usually, a CIM sensitive to one or more volatile compounds is deposited as a thin film on the surface of one of the two surface acoustic wave (SAW) devices, while the other acts as a reference. During the CIM–gas interaction, the acoustic–electric properties of the CIM change, inducing a delay in the propagating surface wave. If the SAW sensors are included in an oscillating circuit, this delay produces a change in the frequency. So, the output frequency of the sensors can be related to the concentration of the adsorbed gas. Suitable propagation modes present in the SAW structure can be used for the detection of either ions or neutral molecules even in liquids.

Figure 4 shows a bulk acoustic wave (BAW) sensor, sensitive to mass changes of the adsorbing layer, made by a quartz slice included in an oscillating circuit. The surface of this device is usually coated by a CIM. During adsorption–desorption processes, the overall quartz mass change induces a frequency variation of the oscillator which can be related to the concentration of the volatile compound. Small quantities of gas molecules (ppm), corresponding to nanograms, can be detected by this sensor. Temperature stability is the

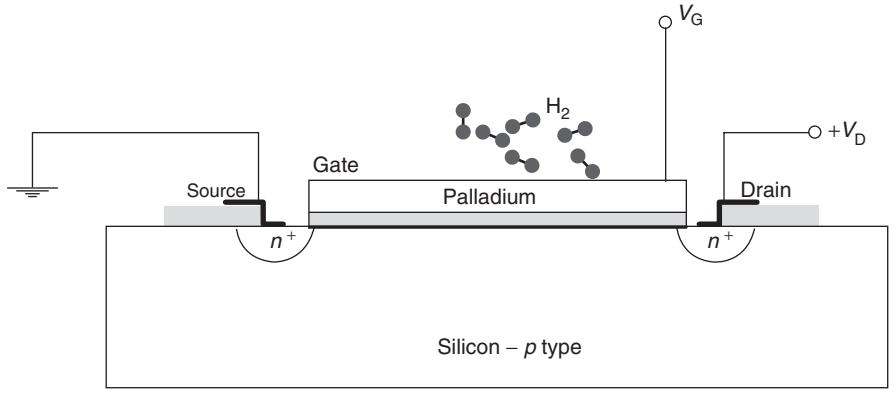


Figure 2 Conceptual scheme of a ChemFET for gas detection. In this particular example, the usual gate is replaced by a film of palladium. As a consequence of the absorption of hydrogen molecules, a number of dipoles are formed at the metal–oxide interface. This results in a detectable change of the source drain current.

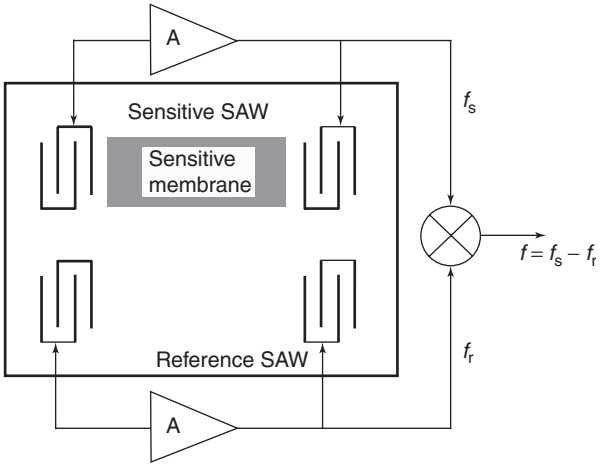


Figure 3 Schematic assembly of a SAW chemical sensor. In this configuration, two SAWs are used as delay lines in feedback configuration with two amplifiers. The propagation path of the acoustic wave in the upper SAW is modified by a chemically sensitive layer. This layer can absorb molecules from the environment and this interaction results in a modification of the acoustic wave propagation conditions. The other SAW is left uncoated and used to compensate any variation different from the chemical interaction. The output quantity is the frequency of the output signal obtained as the difference of the frequencies of the oscillators incorporating the two SAWs.

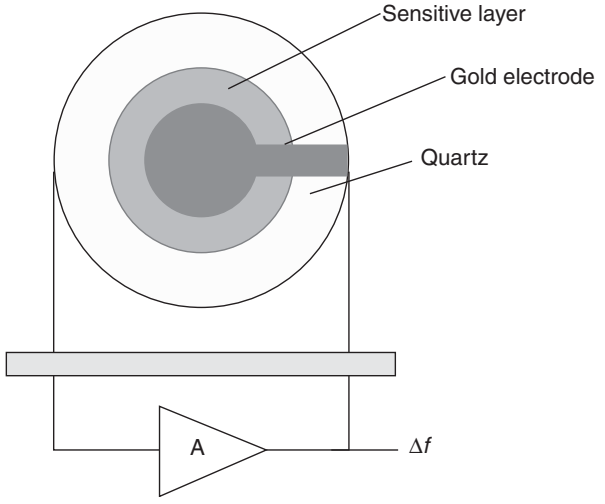


Figure 4 Scheme of a BAW sensor. The sensor is formed by a piezoelectric slab cut along some defined crystal axis. The quartz disk is electrically contacted by metallic electrodes. The mechanical oscillation frequency, sensitive to the gravitating mass, is matched by the piezoelectric effect with the electrical resonance frequency driving and electronic oscillator. The frequency of the output signal is then proportional to the gravitating mass. The chemical sensitivity is given by the sensitive layer deposited onto the quartz surface.

most important constraint to be considered and controlled during the sensing operation.

Figure 5 shows an ion-selective field effect transistor (ISFET) sensitive to the electronic charge, which

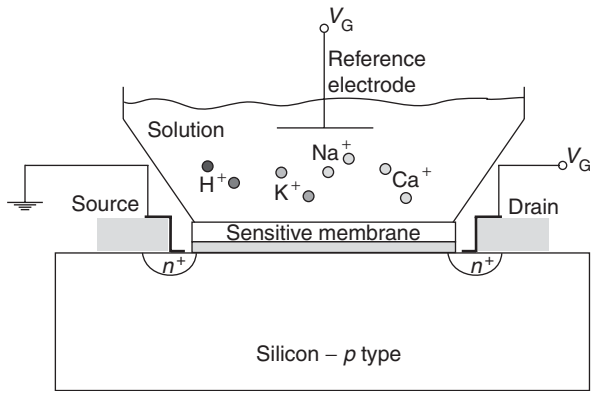


Figure 5 Schematic diagram of the ISFET sensor. In this sensor the usual gate of a MOSFET device is replaced by the sensitive membrane, the electrolytic solution, and the reference electrode. The scope of the device is, as in the case of ChemFET, to bring charged molecules at the oxide – "gate-like" interface in order to alter the current flowing from the drain to source contacts.

can detect the amount of ion concentration in a liquid. In this device, the reference gate is immersed into the liquid and the metal gate is replaced by a CIM suitable for liquids able to adsorb ions. As an example, valynomicine is a well-known CIM for the detection of potassium ions. The presence of ions inside the membrane induces a change of the source drain current which represents the useful output signal. ChemFETs (also GasFETs) belong to the same family; here the gate may be a catalytic metal able to adsorb a certain volatile compound. As an example, a palladium gate would give the possibility of detecting hydrogen or ammonia, even at room temperature. ENFETs have the gate covered by an enzyme able to adsorb biochemical species. The intrinsic operative mechanism is always the same: a charge is needed to be deposited inside the CIM or at the CIM–silicon dioxide interface in order to change the flat-band voltage of the basic device and, as a consequence, to change the output current.

Figure 6 shows a Mach–Zehnder optical interferometer used as a chemical sensor. Here, one branch of the interferometer is covered by a CIM, while the other represents the reference. During the adsorption–absorption and desorption processes, the CIM changes its refractive index and the light at the output turns out to be dependent on the changed optical path of the modified branch. In this device, changes of the refractive index of 10^{-4} can practically be detected, allowing detection capabilities, with respect to gases, of a few ppm.

Figure 7 shows the schematic diagram of a sensor based on conductivity change. These devices have the sensing part made from semiconductor oxide materials such as SnO_x. These materials, which behave

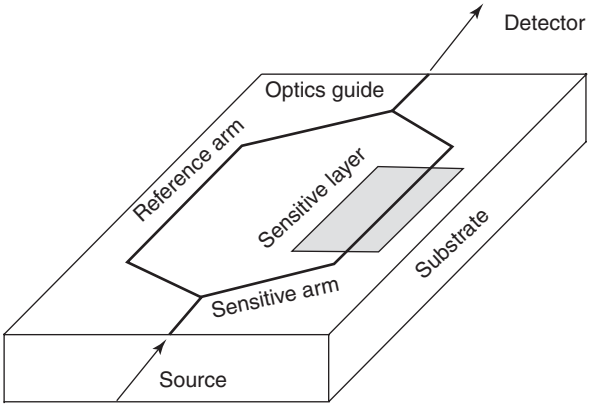


Figure 6 Schematic diagram of the Mach–Zehnder interferometer; the entering light is split into two arms and summed before the output. In order to turn the device into a chemical sensor, it is necessary to coat one of the arms with some material that after the absorption of molecules, changes its dielectric constant. It results in a modification of the optical length of the arm and, due to the interference process, in a modulation of the intensity of the output light.

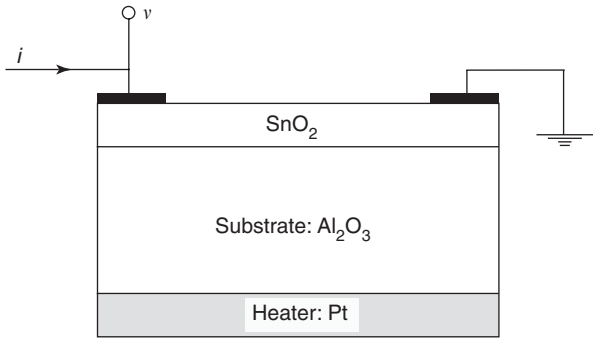


Figure 7 Schematic view of a SnO_2 based chemical sensor. The presence of the heater is necessary in order to bring and maintain the sensitive material at a temperature in the range 250–400°C. At these temperatures, the catalytic processes leading to sensitivity toward both reducing and oxidizing gases are active and the sensor can change its electric resistance.

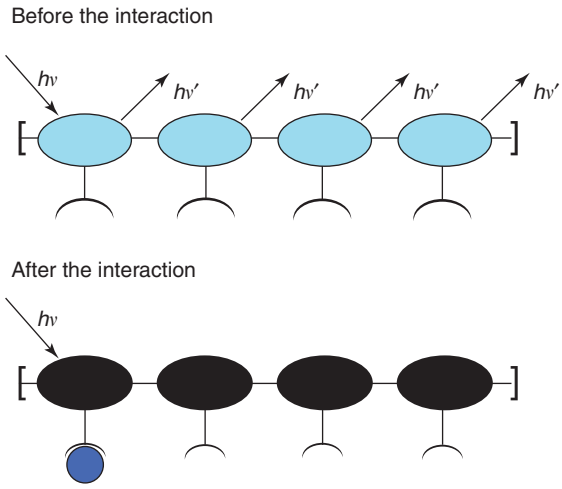


Figure 8 Effect of conjugated polymers. These polymers are formed by linking together fluorescent monomers. The effect of the polymer is that the excitation can travel along the polymer chain inducing a collective fluorescence. If each monomer is functionalized with a receptor able to bind molecules from the environment, the binding of only one molecule is enough to stop the excitation propagation along the chain, quenching the fluorescence of all the fluorescent molecules in the polymer. It provides an amplification of the binding event enhancing the total sensitivity greatly.

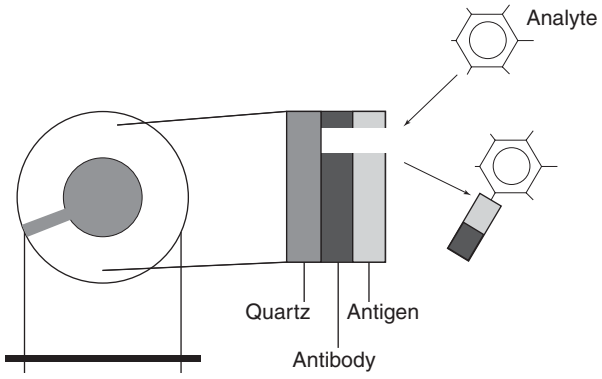


Figure 9 Biosensor based on quartz microbalance sensor. Usually, such sensors measure the mass of the absorbed analyte. In this arrangement, the action of the analyte is to remove a portion of the coating from the quartz surface. Since the removed molecules are heavier with respect to the analyte, this results in a sort of amplification of the sensor sensitivity toward the analyte. It should be remarked that successive exposure to the analyte reduces the amount of coating, and in turn the limited lifetime of the sensor.

like semiconductors, may be doped at the surface by different metals (e.g., gold, palladium, platinum, or copper) in order to enhance their sensitivity compared to volatile compounds and must operate at temperatures ranging from ~200 up to 500°C. Minimum detectable concentrations could be in the ppm range.

Figure 8 shows a chemical biosensor based on the phenomenon of luminescence quenching. A suitable biomaterial which acts as a CIM and is able to emit luminescence light in the visible range under, for instance, ultraviolet excitation, deactivates its luminescence centers by the presence of a given chemical compound. The output-quenched signal is detected by a photon-sensitive device (a solid-state photomultiplier, for instance) allowing rather high resolution values to be obtained, in the ppb range in already proven cases.

Figure 9 shows a schematic diagram of another chemical biosensor utilizing the effect of mass changes. The transducer is represented by a quartz resonator covered by an antigen–antibody structure sensitive to a given analyte, for instance, as that emitted by plastic mines. The presence of such an

analyte takes a piece of antigene out of the transducer inducing a rather strong mass variation which is detected by a frequency change. In this context, a useful strategy consisting of using two quartz resonators, one of which is considered as reference, permits one to reduce the effects of common unuseful parameters such as pressure, temperature, and humidity.

Insofar as physical sensors are concerned, a large number of them have been developed in the last few decades. They are devices able to detect physical quantities, such as photon intensity, pressure, temperature, strain, rotation angles, magnetic and electric field, and refractive index, etc.

Temperature sensors are physical sensors showing sensitivity either to temperature or to temperature changes (thermocouples and thermopiles, thermistors, diodes, and transistors), or to temperature changes only (pyroelectric devices).

Figure 10 shows a schematic of a thermopile made, for example, by two thermocouples. The thermopile output is represented by a voltage proportional to the temperature difference between the hot and cold junctions, the latter representing the reference. This type of a physical sensor is characterized by the minimum noise level, and temperature differences as low as 10^{-4}°C can be detected, even at room temperature.

Thermistors are devices where the material used (metals or semiconductors) shows an exponential resistivity change versus the temperature. Some of them display sensitivity values so high as to give the possibility of detecting even 10^{-5}°C.

Diodes and transistors have an intrinsic sensitivity to temperature; in fact, the diode currents (for the diode) and the collector currents (for the transistor) are exponential functions of the temperature. In a semilogarithmic scale, the plot of the currents versus the diode voltage (or versus the base emitter voltage) has a slope that is directly temperature-dependent. For these devices, temperature resolutions of 10^{-2}°C are possible.

Pyroelectric devices characterized by ferroelectric materials, such as $LiTaO_3$ show an output voltage proportional to the temperature rate, and may allow detection of sinusoidal temperature variation as low as 10^{-6}°C.

Figure 11 shows a schematic diagram of a pressure sensor fabricated by the micromachining technology. The basic principle relies on a membrane where four piezoresistors, used in a full-bridge configuration, are located at the position of larger strain. Once the pressure is applied, the piezoresistors are activated and a voltage can be detected at the bridge output. This voltage can be related to the pressure values.

Figure 12 shows another useful solution for a pressure-integrated sensor. Here a plate belonging to a capacitor is displaced by the applied pressure, and the detection is performed by the capacitance measurement through a suitable circuit, which is eventually fully integrated in the same substrate. Among them, infrared sensors (IR) have been the most important, especially for their strategic use, because they allow night vision with high efficiency, and also because they are useful in detecting underskin cancer. IR radiation is usually classified as near infrared

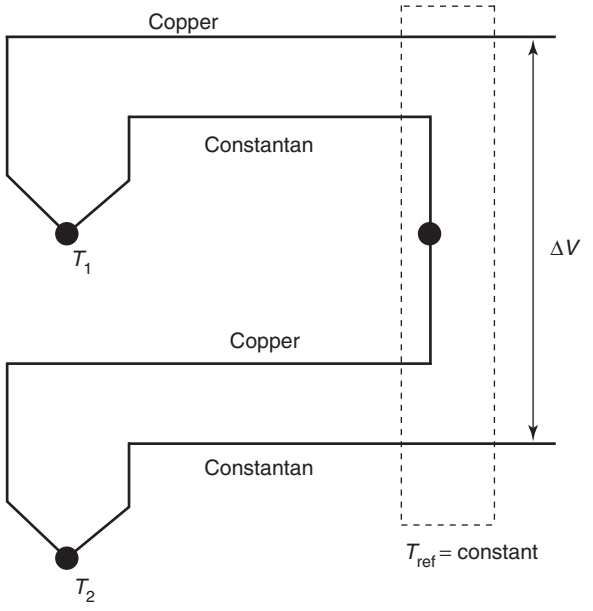


Figure 10 Two thermocouples are connected in series to show the working principle of the thermopile. The resulting voltage drop across the first and the last conductor provides a measure of the mean temperature of the hot nodes (T_1 and T_2 in figure) with a thermoelectric power, that is, the thermoelectric power of the single thermocouple multiplied by the number of couples forming the thermopile.

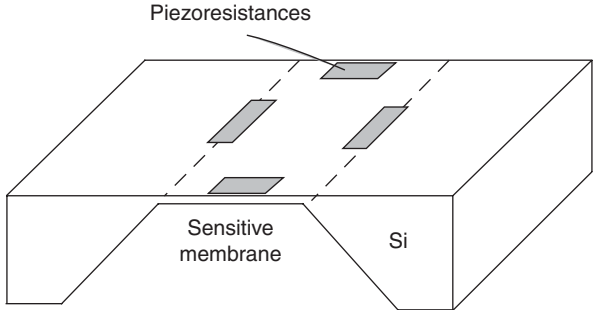


Figure 11 Arrangement of a micromachined fabrication of a pressure sensor. The sensitive membrane moves according to the pressure difference experienced by the two sides of the membrane. The four piezoresistors are assembled in order to provide negative and positive response to the same pressure signal. This arrangement takes the maximum advantage from the Wheatstone bridge configuration leading to a linear relation between output signal and pressure variation and maximum sensitivity.

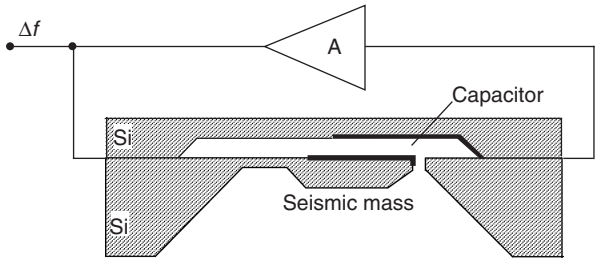


Figure 12 Scheme of an accelerometer based on the capacitive transduction of the position of a seismic mass. The value of the capacitance can then be transformed into a signal frequency by a suitable oscillator circuit.

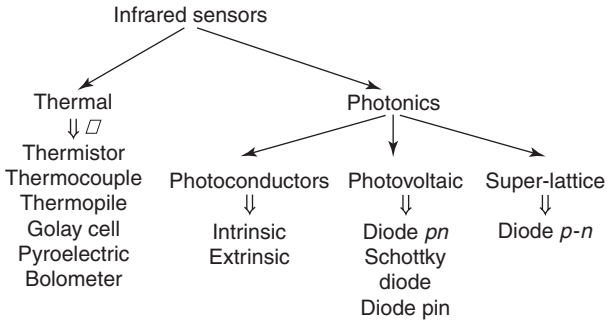


Figure 13 Classification of IR sensors according to the thermal (flat band responsivity) and photonic (narrow band responsivity) detection principle.

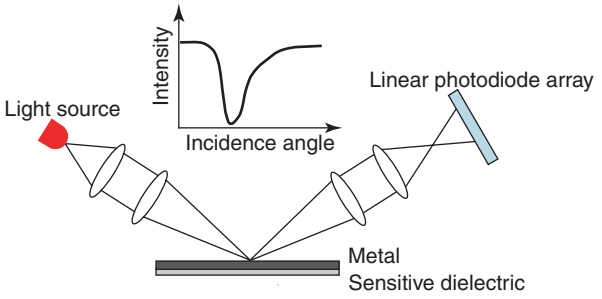


Figure 14 Schematic arrangement of a surface plasmon resonance measurement setup. The light source is condensed through a lens series in order to span a range of incidence angles. Reflected light is recollected on a linear photodiode array. In the inset, a typical response is shown.

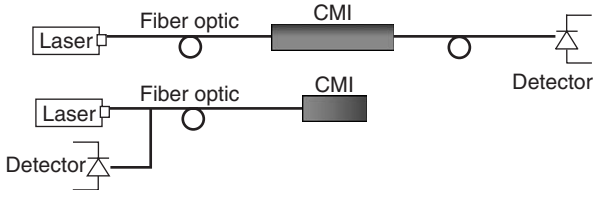


Figure 15 Two possible arrangements of fiber optic chemical sensors are shown in the figure. Both the setups are based on the interaction of chemicals with a sensitive material coating the fiber cladding. In the second case, a fiber bundle is used to collect back the light in order to keep the light source and detector in the same position.

(0.8–2 μm), medium infrared (2–14 μm), and far infrared (14–100 μm). Among these regions, the so-called IR windows are present. They represent wavelength intervals where the adsorption coefficient in air is very low. The window that ranges between 8 and 12 μm is very important, because it represents an easy propagation path of that radiation generated by a room-temperature blackbody.

Figure 13 shows some categories of the wide family of IR devices. The important difference between thermal and photonic sensors is worth mentioning. The former gives a flat response versus wavelength, while the latter does present a pass-band-like response. Photoconductor IR sensors change their conductivity in the presence of IR radiation and a current must be injected through them in order to get a voltage signal, while photovoltage IR sensors behave as IR solar cells and generate a voltage without any polarization current. Important IR photonic materials are PbS, PbTe, PbSe, CdTe, $PbSn_xTe_{(1-x)}$, and $HgCd_xTe_{(1-x)}$. Above 8 μm, many IR sensors must be cooled in order to give an acceptable signal-to-noise ratio, but some of them, such as thermopiles and pyroelectric sensors, can operate at room temperature even up to 12 μm.

Figure 14 shows a schematic diagram of a sensor able to detect the changes of refractive index of a given material. It is based on the surface plasmon resonance effect. Surface plasmons are charge waves propagating along a given metal surface. In this geometry (Kretschmann geometry), light penetrates a prism and impinges on a metal surface with a certain angle of incidence. An evanescent wave is locally generated, which propagates through the film itself. It then generates, under suitable resonance conditions related to the light frequency and incident angle, on the other side of the film, the plasmons. The presence of a liquid with a certain refractive index may be detected by analyzing the intensity change of the light out of the prism. Bioliquids have been succesfully analyzed by this method.

Optical fiber sensors also represent a powerful tool for the development of robust, reliable, low-cost, and sensitive sensors with a high degree of immunity toward electromagnetic interferences. Optical fiber sensors may be configured in order to measure physical parameters (temperature, pressure, strain, electric current, rotation, etc.) or chemical parameters (e.g., composition, concentration, etc). Different types of optical fibers can be utilized in this context: monomode, multimode, and others supporting polarized propagation light. The basic structure of an optical fiber chemical sensor is shown in **Figure 15**.

It is necessary to have a laser, an optical fiber, suitably treated, and a detector. In **Figure 15**, the central part of the fiber is covered by a CIM and, in presence of a given volatile compound (VOC) compatible with the CIM, the detector may experience a change of the light intensity. This effect can be correlated to the concentration of the VOC. In other configurations, the sensitive part is located at one of the fiber extremes, while the other extreme is used either to send a pulse of light for a certain time or to receive the response which carries information about the CIM–VOC interaction.

Artificial Olfaction Systems

One of the most ambitious applications of chemical sensors is concerned with artificial systems that are able to reproduce the functions of the natural chemical senses: olfaction and taste.

Senses necessary to life as interfaces between living beings (both animal and vegetables) and the surrounding world allow living beings to interact with the environment and adapt themselves to it. Traditionally, in humans, five senses are identified: sight, hearing, touch, taste, and olfaction. It is very interesting to note that senses can be further classified according to the nature of the things sensed. From this point of view, it is possible to distinguish the physical senses (hearing, sight, and touch) from the chemical senses (taste and olfaction). A great difference between the two groups exists, and it concerns the treatment of the perceived information. Indeed, in humans it is common to communicate hearing or visual experiences while communication of olfactive or gustative perceptions is not as common. These semantic differences have corresponding difficulties in the technological capability to reproduce the functions of natural senses. For instance, hearing and sight have largely been studied and several technological similii are currently available (such as cameras and microphones). On the other hand, only in the last few years the possibility of designing artificial systems able to mimic chemical senses began to be considered.

The first ideas about artificial olfaction came in the 1960s, but the breakthrough occurred during the 1980s when some researchers argued that the behavior of chemical sensors was rather similar to that of olfaction receptors. Then, the possibility of assembling a number of such devices allowed one to construct a sort of artificial olfaction, which was soon nicknamed "an electronic nose."

Many chemical sensors, such as those previously illustrated, share with the natural receptor an important feature: they are not selective. This means that their signal cannot be related to one specific volatile compound, but rather to the global nature of the compounds present in the air. Indeed, natural olfaction does not provide analytical information about the inhaled air, but olfaction provides signals to the brain in order to allow the individual to give a sort of qualitative description of the sniffed air. In practice, natural olfaction makes use of a large number of different receptors, where none of them is selective, but the totality of the ensemble makes the individual selective to thousands of different odors.

During the 1990s, the research and development of electronic noses grew, involving a large number of researchers. Electronic noses have been demonstrated to be able to classify and distinguish different odors (e.g., to distinguish different wines), and to identify, after proper training, unknown samples (e.g., to identify the denomination of a particular wine). A sort of standard model of an electronic nose can be found by looking at the extent of technological implementation of the original idea. This is based on an array of sensors, from a few units in the case of metal-oxide chemoresistors or quartz microbalances up to thousands of fluorescent beads micromechanically carved on a single optic fiber.

Disregarding the working principle and the technological fabrication, sensors are required to be "globally sensitive" and, at the same time, sensors with a range of selectivity properties must be produced. In this way, each sensor provides different information contributing to the global identification of samples. Another important characteristic of the electronic nose is data analysis. Multicomponent analysis techniques are utilized to infer from the sensor signals the properties of the measured sample. Among these techniques, the most utilized are those derived from chemometrics, namely, the algorithms and methods used in chemistry to analyze data of spectral

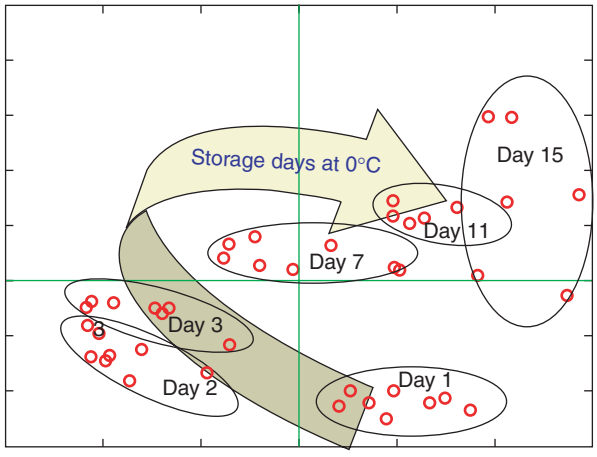


Figure 16 Result of a typical electronic-nose measurement aimed at characterizing the process of spoilage in fish.

instruments (e.g., spectrophotometry and gas chromatography). Also, neural networks have been used extensively, and these have contributed to enhancing the properties of artificial olfaction systems.

Data analysis usually results in maps where odors are plotted in order to study their similarities and differences. As an example, **Figure 16** represented below shows a typical result obtained in the analysis of food. The same principle of electronic nose has also been applied to arrays of sensors operating in liquid samples giving rise to the so-called "electronic tongue." This system has been found to be efficient in environmental (e.g., the detection of various ionic species in ground waters) and food applications (e.g., different coffee brands and spoilage processes in fruit juices).

See also: Micromechanical Devices and Systems; Semiconductor Devices.

PACS: 07.07.Df; 42.79.Pw; 87.80. – y

Further Reading

D'Amico A and Di Natale C (2001) A contribution on some basic definitions of sensor properties. *IEEE Sensors Journal* 1: 183–190.

D'Amico A, Palma A, and Verona E (1982) Surface acoustic wave hydrogen sensor. *Sensors and Actuators* 3: 31–39.

Bergveld P (1970) Development of an ion-sensitive solid-state device for neurophysiological measurements. *IEEE Transactions on Biomedical Engineering* 17: 70–71.

Frank R (1991) Pressure sensors merge micromachining and microelectronics. *Sensors and Actuators B* (28): 93–103.

Guilbault GG (1983) Determination of formaldehyde with an enzyme coated piezoelectric crystal. *Analytical Chemistry* 55: 1682–1684.

Jorgensen RC and Yee SS (1993) A fiber-optic chemical sensor based on surface plasmon resonance. *Sensors and Actuators B* (12): 213–220.

King WH (1964) Piezoelectric sorption detector. *Analytical Chemistry* 36: 1735–1739.

Kruse PW, McGlauchlin LD, and McQuistan RB (1962) *Elements of Infrared Technology.* New York: Wiley.

Liedberg B, Nylander C, and Lundstrom I (1983) *Sensors and Actuators* 4: 299–304.

Lundstrom I, Shivaram MS, and Svensson C (1975) A hydrogen–sensitive MOS field-effect transistor. *Applied Physics Letters* 26: 55.

Roylance LM and Angeil JB (1979) *IEEE Transactions on Electronic Devices* 26: 1911–1917.

Schipper EF, Bergeovet AJ, Kooyaman RP, and Greve J (1997) New detection method for atrazine pesticides with the optical waveguide Mach–Zehnder immunosensor. *Analytica Chimica Acta* 341: 171–176.

Seyama T, Kato A, Fukiishi K, and Nagatini M (1962) A new detector for gaseous components using semiconductive thin films. *Analytical Chemistry* 34: 1502–1503.

Wohltien H and Dessy R (1979) Surface acoustic probe for chemical analysis. *Analytical Chemistry* 51: 1408–1412.

Yamazoe N, Kurokawa Y, and Seyama T (1983) *Sensors and Actuators* 4: 283–289.

Zhang JS and Swager TM (2003) Fluorescent detection of chemical warfare agents: functional group specific ratiometric chemosensors. *Journal American Chemical Society* 125: 3420–3421.

Zemel JN, Keramati B, Spivak CW, and D'Amico A (1981) Non FET chemical sensors. *Sensors and Actuators* 1: 427–473.

Shubnikov–de Haas and de Haas–van Alphen Techniques

W Biberacher, Walther-Meissner-Institut, Garching, Germany

The de Haas–van Alphen (dHvA) and Shubnikov–de Haas (SdH) effects describe oscillatory components of the magnetization and conductivity, respectively, of a metal as a function of an applied magnetic field at low temperatures. The oscillations are periodic in a reciprocal magnetic field. It is the most accurate method to determine the contours of the Fermi surface of a metal or semimetal, but also gives information on other physical properties of electrons such as effective mass, scattering time, and spin-splitting factor.

Historical Introduction

The first experimental and theoretical manifestation of this effect dates back to 1930. In this year, L D Landau developed the quantum mechanical calculation of the diamagnetism of conduction electrons and mentioned the possibility of oscillatory effects there. But he believed, it was unobservable at that time. In the same year, L V Shubnikov and W J de Haas published an oscillatory behavior of the magnetoresistance in bismuth, and at the end of this year de Haas and P M van Alphen reported the observation of an oscillatory magnetization in the same material. Three years later, R Peierls developed a quantitative theory of the oscillations in Bi based on Landau's formulation of the energy levels of electrons in a magnetic field. In 1939, D Shoenberg determined the Fermi surface of Bi using the theoretical calculations of Landau. Shoenberg came to be the leading scientist in this field for several decades.

It was only around 1950 that general interest was seen in these oscillatory effects. In 1947, dHvA oscillations were detected in zinc, the first observation in a metal other than Bi. Soon after, it became clear that

this effect could be observed in nearly every polyvalent metal, if crystals of high quality were available.

In the beginning of the 1950s, three important theoretical advances were made: first, the influence of the electron spin was worked out as a damping factor of the oscillation amplitude, second, R B Dingle could show that electron scattering broadens the Landau levels. This should change the oscillations in a way similar to a rise in temperature. The most important contribution was first published by L Onsager. He showed that the dHvA frequency is connected with the geometry of the Fermi surface and gives the extremal cross section of the Fermi surface perpendicular to the magnetic field. But this breakthrough could not be exploited immediately: at that time only low-frequency oscillations of the rather complex Fermi surfaces of polyvalent metals could be experimentally observed. On the other hand, band structure calculations were not yet good enough to show all details of a Fermi surface.

Around 1950, typical laboratory magnets produced fields up to 3 T. The very few water-cooled high-power magnets available at that time suffered from their instability. It was clear that higher and/or more stable magnetic fields are needed. Shoenberg, therefore, started working with pulsed fields up to 10 T and was soon able to observe high-frequency oscillations also in polyvalent metals, and even made the first observations in monovalent metals.

On the basis of prior theoretical work, I M Lifshitz and A M Kosevich developed a complete theory of the dHvA effect in 3D metals published in 1956. Together with improved band structure calculations starting from 1960, dHvA oscillations were investigated in all metals and compared with the theoretical band structure.

Although the basic investigations of the metallic elements are more or less completed, the dHvA and SdH effects are still of high use in the investigation of new compounds. Important examples are heavy fermion systems and low-dimensional conductors such as organic metals or artificially grown 2D electron systems.

Basic Theoretical Models

Landau Quantization

The physical reason for the oscillatory effects in a magnetic field can be easily understood in the model of a free electron gas confined to a cube of length L. (The electron spin is neglected first.) The Schrödinger equation for free electrons in a magnetic field is given by

$$\frac{1}{2m_e}\left(\frac{\hbar}{i}\nabla - eA\right)^2 \Psi = E\Psi$$

where m_e is the electron mass, e is the electron charge, and $\boldsymbol{A}$ is the vector potential of the magnetic field $\boldsymbol{B}$. Landau showed in 1930, that for a magnetic field applied in the z-direction the energy is given by

$$E = \left(n + \frac{1}{2}\right)\hbar\omega_c + \frac{\hbar^2}{2m_e}k_z^2$$

with $n = 0, 1, 2, \ldots$, $\omega_c = eB/m_e$ being the cyclotron frequency and $k_z = 2\pi n_z/L$ $(n_z = 0, \pm 1, \pm 2 + \cdots)$ the wave vector in the z-direction. If one compares this to the solution of a free electron gas without a magnetic field, where $E = \hbar^2 k^2/2m_e$, it is seen that the wave vector in the x,y-plane perpendicular to the magnetic field is no longer quasicontinuous, but takes only discrete values

$$k_{x,y} = \sqrt{(n + 1/2)2m_e\omega_c/\hbar}$$

All permitted states lie on concentric tubes parallel to the magnetic field called Landau cylinders. The energy levels belonging to a certain n are highly degenerate. The degeneracy D is given by

$$D = \frac{eBL^2}{h}$$

This is the ratio between the magnetic flux BL^2 through the sample and the flux quantum h/e. The degeneracy is equal to the number of states contained in the adjacent interval of width $\hbar\omega_c$ of the electron spectrum in a zero field, that means, the adjacent states condense to the Landau tubes.

The cross-sectional areas of the Landau cylinders $S_k = \pi k_{x,y}^2$ are proportional to the applied magnetic field, that means, the Landau tube for a certain n increases as the field is increasing. In zero field, the electronic states are filled up to the Fermi level. In an increasing magnetic field, subsequent Landau levels pass through the Fermi energy and become depopulated (see **Figure 1**). This gives an oscillatory dependence to the free energy of the electronic system and to all other quantities derived from it. From the linear dependence of S_k on B, it can be easily derived that the oscillatory component must be periodic in $1/B$ with the frequency F of the oscillations given by the extremal cross section S_{extr} of the Fermi surface, as was first pointed out by Onsager:

$$F = \frac{h}{e}S_{extr}$$

Lifshitz–Kosevich (LK) Formula

In 1956, Lifshitz and Kosevich published a theory of the magnetic susceptibility of conduction electrons in a metal with an arbitrary dispersion law. The

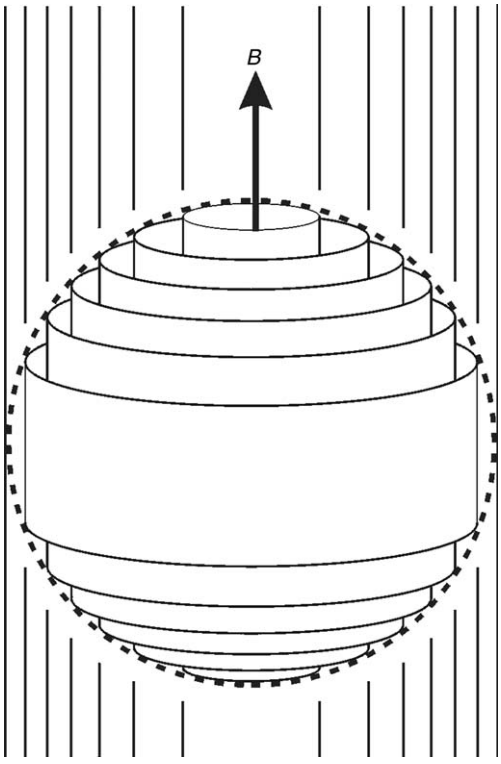


Figure 1 Schematic drawing of a spherical Fermi surface with Landau cylinders.

oscillatory part of their solution turned out to be a complete description of the experimental results of the dHvA-oscillations measured in moderate magnetic fields and is still the standard to compare experiment with theory. They calculated the thermodynamic potential Ω of a gas of quasiparticles. The magnetization is then given by the thermodynamic relation:

$$\boldsymbol{M} = -(\mathrm{grad}_{\boldsymbol{B}}\Omega)_{T,\mu}$$

The contributions to the oscillatory part come exclusively from the vicinity of the extremal orbits of the Fermi surface. For a certain extremal orbit, the oscillatory magnetization M parallel to the applied magnetic field is given by

$$M \propto \frac{F}{m}\left(\frac{B}{\partial^2 A_k/\partial k_z^2}\right)^{1/2} \sum_{p=1}^{\infty} \frac{1}{p^{3/2}} R_\mathrm{T} R_\mathrm{D} R_\mathrm{S} \times \sin\left[2\pi p\left(\frac{F}{B} - \gamma\right) \pm \frac{\pi}{4}\right]$$

F is the dHvA frequency, m is the cyclotron mass, $\partial^2 A_k/\partial k_z^2$ is the curvature of the cross-sectional area of the Fermi surface at the extremal orbit along the field direction, the factors R_T, R_D, and R_S are damping factors according to impurity scattering, temperature and spin, and γ is a value close to 0.5. The sign of the phase $\pi/4$ is dependent on the geometry: if the extremal orbit is a minimum, the sign is $+$; for a maximum, it is $-$. In normal situations, the cyclotron energy $\hbar\omega_\mathrm{c}$ is much smaller than the Fermi energy and the oscillatory magnetization is dominated by the first harmonic in the summation over p.

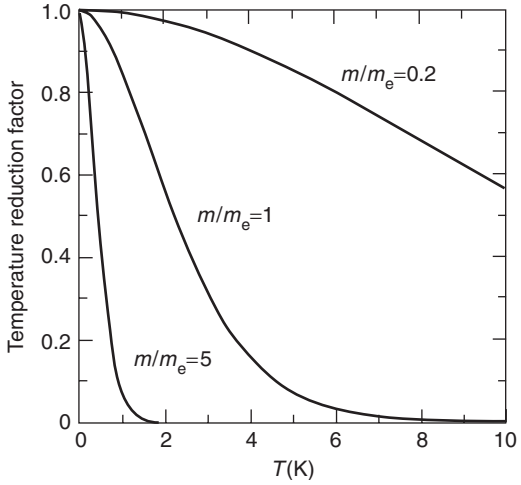


Figure 2 Temperature reduction factor of the fundamental frequency according to the Lifshitz–Kosevich theory at a magnetic field of 15 T for different values of the cyclotron mass.

In general, Fermi surfaces can be quite complicated and contain more than one extremal orbit. In that case, the total oscillatory component is a sum over all the contributions, each of which follows the LK formula, but with different parameters.

Without the reduction factors R_i the LK formula gives the idealized situation, where finite temperature, finite electron relaxation time, electron spin and other complications are neglected. At finite temperature, the occupation probability of a certain state with energy E is given by the Fermi Dirac function:

$$f(E) = \frac{1}{(1 + \exp((E - E_\mathrm{F})/kT))}$$

where k is the Boltzmann constant and E_F is the Fermi energy. This is a step function for $T = 0$ K but is smeared out by kT around E_F at finite temperature. This smearing leads to a reduction of the dHvA amplitude by the temperature reduction factor R_T:

$$R_\mathrm{T} = \frac{2\pi^2 pkmT/\hbar eB}{\sinh(2\pi^2 pmkT/\hbar eB)}$$

Typical dependencies of the reduction factor on the temperature at a magnetic field of 15 T are given in **Figure 2** for several values of the normalized cyclotron mass m/m_e.

Real crystals are always imperfect in some respect. This leads to a finite relaxation time τ of electrons, which is equivalent to a broadening of the Landau

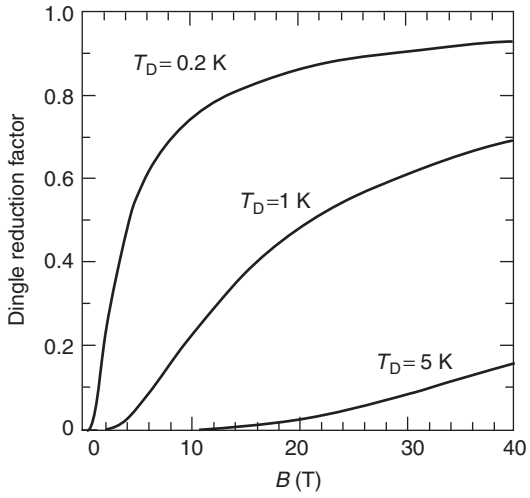


Figure 3 Dingle reduction factor of the fundamental frequency according to the Lifshitz–Kosevich theory for different values of the Dingle temperature. The cyclotron mass is set equal to the free electron mass.

levels. If this broadening is described by a Lorentzian distribution, the reduction of the oscillation amplitude is given by

$$R_D = \exp\left(\frac{-\pi p}{\omega_c \tau}\right)$$

This directly shows that the relaxation time must be long enough for the electron to complete a cyclotron orbit. The effect of impurities is similar to an increase in sample temperature. The reduction factor is generally given as

$$R_D = \exp\left(-\frac{2\pi^2 pmkT_D}{\hbar eB}\right)$$

with T_D being the so-called Dingle temperature, $T_D = \hbar/2\pi k\tau$. The influence of T_D on the oscillation amplitude is shown in **Figure 3** for free electrons ($m = m_e$). In samples with Dingle temperatures higher than a few kelvin, oscillations are hard to detect. Additional damping can be caused by sample and field inhomogeneities, mosaic spread, bending and strain on the sample, etc.

In a magnetic field, the spin degeneracy of the energy levels is lifted and each Landau level is split into two at energies $E \pm \frac{1}{2}\Delta E$ where

$$\Delta E = g\mu_B B$$

Here $\mu_B = e\hbar/2m_e$ is the Bohr magneton and g is the spin-splitting factor, which takes the value 2.0023 for free electrons. On increasing the field, the split Landau levels pass the Fermi surface at different field values. Therefore, a phase difference between these two oscillations arises and leads to a reduction of the amplitude by

$$R_S = \cos(\tfrac{1}{2}p\pi gm/m_e)$$

For free electrons the spin reduction factor is just $(-1)^p$, since the spin-splitting is equal to the Landau level spacing. In special cases, if $pgm/m_e = 1, 3, 5, \ldots, R_s$ becomes zero and the oscillation amplitude vanishes. This is called a spin-zero. The relation also shows that in the spin-zero of the fundamental frequency ($p = 1$), the second harmonic becomes maximal.

Other Oscillatory Quantities

The standard method for the determination of Fermi surfaces of metals is the dHvA effect. But as mentioned above, the dHvA oscillations were derived from oscillations of the thermodynamic potential Ω. Therefore, it is clear that all other physical properties derived from Ω show an oscillatory behavior in a high magnetic field and there are many experimental investigations of these effects. Examples are the specific heat, magnetostriction, elastic properties, and the chemical potential. Oscillations of the chemical potential play an important role in quasi-two-dimensional (Q2D) systems as discussed later. In 3D systems, they are usually weak and can be neglected in most cases. Another important quantity, which can be derived from the thermodynamic potential or calculated in a similar way, is the density of states at the Fermi energy.

There are oscillatory effects which cannot be described by thermodynamic relations alone since they involve nonequilibrium properties. The most famous example for it is the SdH effect, the oscillatory property of the magnetoresistance. As mentioned above, magneto-oscillations were first observed in the magnetoresistance of Bi in 1930. But it turned out that the SdH effect is very weak in normal metals and hard to detect there. In recent years the SdH effect has been of high importance, since it is easy to observe in Q2D electron systems such as quantum well structures or organic metals.

The theory of the SdH effect is complicated since it involves the problem of electron scattering in a magnetic field. In 1959, E N Adams and T D Holstein calculated the oscillatory conductivity for two different scattering mechanisms and found similar results in both cases. They concluded that the exact nature of scattering is not very crucial. Their result for the case of phonon scattering and high quantum number is the following:

$$\frac{\tilde{\sigma}}{\sigma} = \frac{R_T \frac{5}{2}\tilde{N}}{N_0}$$

where $\tilde{\sigma}$ is the oscillatory part of the conductivity, σ the steady conductivity, R_T the temperature reduction factor and $\tilde{N}, N_0$ are the oscillating and steady parts of the density of states at the Fermi energy, respectively. This result can be understood qualitatively by an argument of Pippard: the probability of scattering is proportional to the number of states in which the electrons can be scattered. Therefore, the conductivity is related to the oscillations of the density of states. The SdH effect is described by the same reduction factors as the dHvA effect. R_S and R_D are already involved in $\tilde{N}$. Finally, one should mention that the oscillatory part of the density of states and thus the oscillatory conductivity is proportional to the derivative of the magnetization dM/dB.

Two-Dimensional Case

The above mentioned theories for the dHvA and SdH effect were calculated for an isotropic case, that is, for a spherical Fermi surface. It turned out that they can also describe anisotropic systems with ellipsoidal or even warped cylindrical Fermi surfaces well. But in the pure 2D case, the theory is no longer applicable. From a theoretical point of view, it turned out that the 2D case is in some respect more difficult to calculate than the 3D one, and up to now, analytic solutions are available only for limiting cases.

Typical examples for Q2D electron systems are layered metallic compounds. The energy spectrum is usually given by

$$E_k = \frac{\hbar^2}{2m}(k_x^2 + k_y^2) - 2t\cos(k_z d)$$

with t being the interlayer transfer integral, k_z the wave vector perpendicular to the layers, and d the interlayer distance. This yields a cylindrical Fermi surface with a slightly oscillating cross section along the k_z direction (warping). In the case of $2t \gg \hbar\omega_c$, the system can be well described within the LK theory. In that case, there exists a minimal and a maximal cyclotron orbit and one expects two nearby frequency components in the dHvA spectrum (see **Figure 4**).

When the transfer integral becomes comparable to the Landau level separation, the LK theory is no longer appropriate and for $2t \ll \hbar\omega_c$, the system is very near to the pure 2D case. In that case, the whole Fermi surface is extremal and one expects large oscillations. It is also clear that the assumption of a field independent chemical potential is, generally speaking, no longer valid. For a constant number of electrons, one would rather expect jumps in the chemical potential periodic in $1/B$. In the extreme 2D case, the Fermi energy lies between two Landau levels. The chemical potential is fixed to the highest occupied Landau level and when this level is depopulated by increasing the field, the chemical potential will jump to the next lower Landau level. A sawtooth form of the magnetization oscillations is expected for that case and was indeed observed in artificially grown 2D electron gases. The other limiting case of a fixed chemical potential can be fulfilled by oscillations of the total number of electrons due to an infinite reservoir. This situation is probably given in some organic metals where a Q1D band is present beside the 2D one.

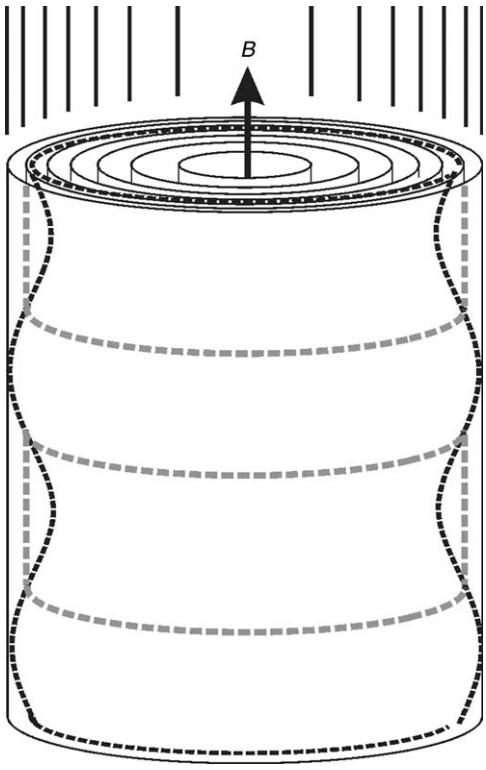


Figure 4 Slightly warped cylindrical Fermi surface (black dashed line) of a Q2D metal with Landau cylinders. The magnetic field is applied perpendicular to the conducting plane.

Analysis of Data

In a typical experimental investigation of the quantum oscillations, the first quantity to derive is the oscillation frequency. The oscillatory signal measured in a certain interval of the magnetic field and plotted against $1/B$ is transformed by a Fourier transformation into a power spectrum. This gives the frequencies of the different extremal orbits of the Fermi surface and – if present – their higher harmonics. Repeating this procedure for different field directions allows one to reproduce the complete Fermi surface. One should mention that the frequency only gives the cross section of the extremal orbit, but cannot show the deviations from a circular form.

A comparison with the LK theory gives further information. If the oscillation amplitude is measured at different temperatures, a least-square fit of the temperature reduction factor to the experimental data directly yields the effective cyclotron mass m of this extremal orbit. The magnetic field dependence of the oscillation amplitude at constant temperature is given by the product of the Dingle factor R_D and B^n, where n depends on the experimental method. Knowing the effective mass m from the temperature dependence, one can extract the Dingle temperature by a least-square fit of the field dependence.

Finally, the spin reduction factor R_S can be used to determine the spin-splitting factor g. In most cases, there is an intrinsic ambiguity in the determination of the g-factor and only a series of possible values can be derived. A special situation is given in the layered organic metals. Due to the relatively high effective mass m and the angular dependence of m, usually a whole series of angles appears, where the oscillation amplitude vanishes due to R_S. This allows one to determine the product gm/m_e without ambiguity.

Special Cases

Magnetic Interaction, Condon Domains

Normally, the oscillating part of the magnetization $\mu_0\tilde{M}$ is much smaller than the applied magnetic field B_a. Therefore, in the LK theory the inner magnetic field B_i was taken as the applied field B_a. In some cases (very pure metals with large Fermi surfaces, low temperature), this assumption is no longer valid as was first observed by Shoenberg in very pure samples of gold. To analyse his data, he started with the LK-equation for the first harmonic and replaced B by B_i. He could show that the condition

$$a = \mu_0 \left| \frac{d\tilde{M}}{dB} \right| \ll 1$$

must be fulfilled for the LK theory to be fully applicable. If the factor a is smaller than 1 but of the same order of magnitude, the observed oscillations show an increased harmonic content and, if several different orbits are involved, the occurrence of combination frequencies is expected. In these cases, one has to take into account the demagnetization factor n of the sample shape also.

If the factor a becomes larger than 1, the magnetization of the sample is no longer a single-valued quantity of the applied field. This is thermodynamically unstable and leads to jumps in the magnetization. As first proposed by Condon, under this condition a sample with a demagnetization factor $n > 0$ will split up into magnetic domains with two different values of the magnetization. The existence of these domains could be shown by NMR and muon spin rotation experiments. Up to now, clear evidences for Condon domains are reported in high-quality samples of silver, beryllium, white tin, aluminum, and lead.

Magnetic Breakdown and Quantum Interference

The Fermi surface of a metal can be completely described within the first Brillouin zone. Any constant energy surface exceeding this zone belongs to higher bands and can be folded back to the first Brillouin zone. It was therefore very exciting, when in the beginning of the 1960s, Priestly reported the observation of a dHvA orbit in Mg, which was considerably larger than the hexagonal cross section of the Brillouin zone. The explanation for this observation was given by Cohen and Falicov, who pointed out that electrons could tunnel from one band to another, if the bandgap E_g between these two bands is sufficiently small and the magnetic field strong enough. It is evident that this could happen, when the bandgap is comparable to or smaller than the Landau level separation $\hbar\omega_c$, but Blount could show that the much weaker condition

$$\hbar\omega_c \geqslant \frac{E_g^2}{E_F}$$

is valid, with E_F being the Fermi energy. Since typically $E_g/E_F \ll 1$, magnetic breakdown could happen in quite low magnetic fields.

For the theoretical description of magnetic breakdown, a simple hypothetical model of a free electron gas in a 1D crystal potential was introduced as shown in **Figure 5**. At the zone boundary, a small gap opens and the Fermi surface consists of a closed (α-orbit) and an open part. At very high magnetic fields, electrons can tunnel through the gap and a new much larger closed orbit (β-orbit) is observed. At intermediate magnetic fields, a complicated spectrum with sum and difference frequencies of α- and β-orbits will appear. The case – first thought as hypothetical – was in the meantime verified in artificial heterostructures and in organic metals such as (BEDT-TTF)$_2$Cu(NCS)$_2$ and is of high interest at present.

In systems, where magnetic breakdown is present, one often finds a striking difference in the frequency spectrum of dHvA and SdH oscillations. The SdH oscillations typically show additional frequencies not present in dHvA. These differences can often be described by quantum interference effects as introduced first by R W Stark and C B Friedberg: it is possible to have magnetic breakdown between orbits, along which electrons move in the same direction; if

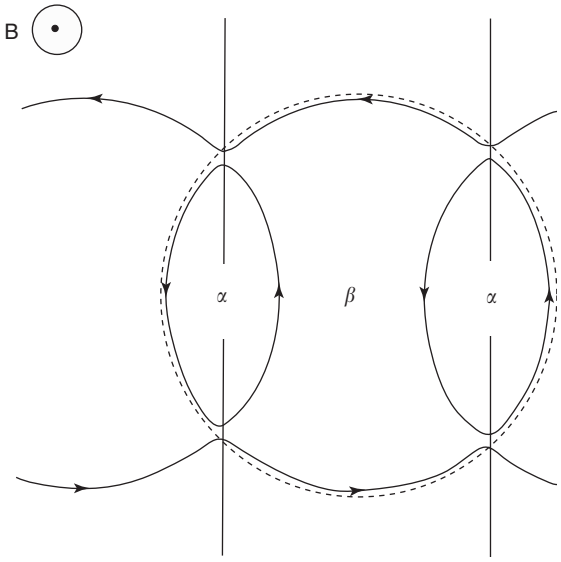


Figure 5 Model Fermi surface for magnetic breakdown: at small fields the carriers are fixed to the closed α-orbit or the open orbit above and below, at high magnetic fields the carriers can tunnel through the gap and a larger closed orbit (β-orbit, dashed line) becomes dominant.

electrons moving from point A to B can travel along two different paths connected by magnetic breakdown, interference effects yield an oscillatory contribution in transport measurements with the frequency determined by the area between the two paths. Since this represents not a closed orbit, these frequencies are not detected in thermodynamic quantities such as dHvA oscillations.

Experimental Techniques

Orders of Magnitude

Quantum oscillations are periodic in the reciprocal magnetic field. Therefore, the unit of the oscillation frequency is tesla (T). Typical values range from a few T to $\sim 50\,000$ T. Low-frequency oscillations are easier to detect and can be often observed up to temperatures of 10–20 K. The observation of high-frequency oscillations requires usually low temperatures ($T < 4$ K), and high and homogeneous magnetic fields. (The oscillation period of a 50 kT frequency amounts to 20 μT at $B = 1$ T.) The observation of oscillations in the heavy fermion compounds needs dilution fridge temperatures for detection. This can be easily verified by calculation of the temperature reduction factor R_T with high effective cyclotron masses.

Optimal relative amplitudes of the oscillatory magnetization $\mu_0 M/B$ are $\sim 10^{-5}$, which is comparable to steady susceptibilities of weakly magnetic materials. The sensitivity of magnetization experiments is often improved by measuring dM/dB (see below). The oscillations in the resistivity are experimentally observed mainly in semimetals and under magnetic breakdown conditions. In recent years, the SdH effect plays an important role in layered metals and quantum well structures. In both cases, the oscillatory amplitude can exceed the background resistance in high magnetic field considerably.

Magnetic fields are typically produced nowadays by superconducting magnets. The maximum field available in laboratories lies between 8 and 20 T. Higher steady magnetic fields can be obtained in some high-field magnet laboratories running high-power (typically, 10–25 MW)-resistive or superconducting-resistive hybrid magnets. These laboratories are usually open for users and can offer magnetic fields between 20 and 45 T. For even higher magnetic fields, some pulsed field facilities are operated. For the study of metallic samples, they can produce magnetic field pulses of duration between a few ms and 1–2 s with maximal fields between 35 and 70 T.

de Haas–van Alphen Techniques

The above mentioned magnitude of the oscillatory effect shows that the standard techniques for measuring small magnetic susceptibilities are often sufficient for the observation of the dHvA effect. But for observation in less favorable circumstances, techniques with very high sensitivity are needed. It should also be mentioned that the high degree of field homogeneity required for the observation of high-frequency oscillations makes the Faraday method less suitable. The two mostly used methods are discussed in the following.

Pickup coil techniques This technique is similar to the well-known mutual inductance method for AC-susceptibility measurements. The sample is placed in a pickup coil system consisting of two identical coils connected in series with opposite signs so that the voltage induced by a changing magnetic field is zero. If a sample is placed in one of the coils, the induced voltage will be

$$U \propto (dM/dB)(dB/dt)$$

The proportionality constant is determined by the geometry. For the arrangement of the two balanced coils, many possibilities exist. Due to limited space of high homogeneity in a magnet, the two coils are often wound one after the other on the same coil core. In the standard laboratory method, the applied magnetic field B is superimposed by an oscillatory contribution $b(t) = b_0 \cos(\omega t)$, with typical frequencies ω

lying between 20 Hz and a few kHz. The oscillation amplitude b_0 can be appropriately chosen for the system under consideration. It should be smaller or comparable to the period of the dHvA oscillation. The modulation field method allows the use of phase-sensitive detection via lock-in amplifiers, which considerably improves the signal-to-noise ratio. In addition, this allows detection at different harmonics of the modulation frequency.

The signal amplitude of the different harmonic contributions A_p of the LK theory is given by

$$U_p^{(k)} \propto k\omega J_k\left(\frac{2\pi pFb_0}{B^2}\right)A_p$$

for detection at the kth harmonic of the modulation signal, J_k is the kth order Bessel function. Some disadvantage of this method is the fact that the weight of the harmonics is changed: in case of high harmonic content, the observed waveform deviates from the real one and has to be recalculated. On the other hand, this can be used for suppressing a certain harmonic by working near the corresponding Bessel function zero, and therefore having increased sensitivity for other weak oscillations.

The changing magnetic field may be also produced by a pulsed magnetic field. In that case, the induced voltage is first amplified and then recorded by a fast data acquisition system. In both methods, a self heating of the sample due to eddy currents must be avoided.

Torque method If the Fermi surface of a given metal deviates from spherical symmetry, that is, if the characteristic frequency F of an extremal orbit is dependent on the field orientation, there exists a torque about any axis perpendicular to the magnetic field given by

$$\tau = -\frac{1}{F}\frac{\mathrm{d}F}{\mathrm{d}\theta}BVM$$

where, θ determines the angle between the magnetic field B and a characteristic direction in the plane perpendicular to the torque axis, V is the volume of the sample, and M is the parallel magnetization as given by the LK formula. The torque vanishes for some symmetry directions of the crystal, where either $\mathrm{d}F/\mathrm{d}\theta$ vanishes or two related parts of the Fermi surface cancel each other. In practice this is often not a problem, since the missing angular range is very small and can be easily interpolated. In case of large oscillations, the torque should be measured by a feedback system to avoid problems due to slight changes of the sample orientation (torque interaction). A strong advantage

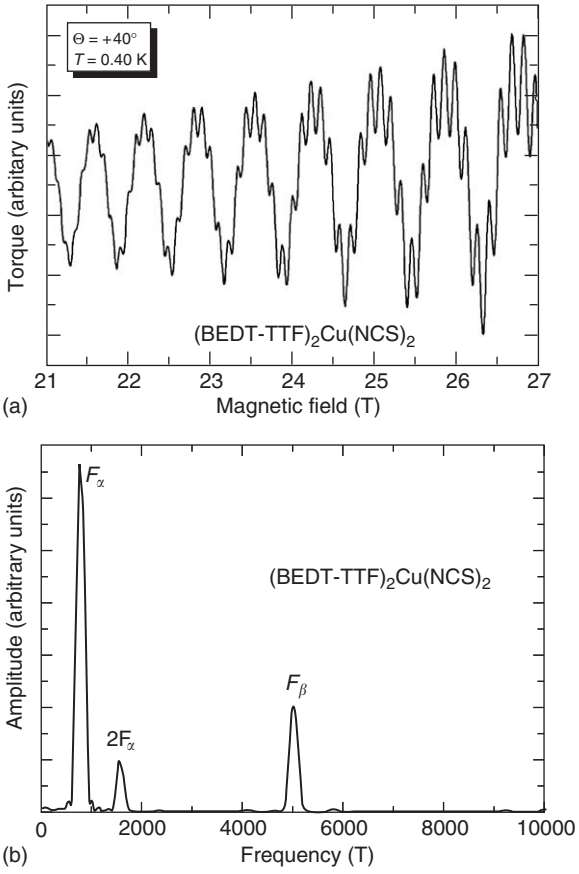


Figure 6 (a) Magnetic torque oscillations in very high magnetic fields showing magnetic breakdown. The sample is a small ($m = 80\,\mu g$) single crystal of the organic Q2D metal κ-(BEDT-TTF)$_2$Cu(NCS)$_2$. The tilt angle between the direction of magnetic field and the normal to the conducting layers is 40°. (b) Corresponding Fourier transform of the data in (a). The spectrum clearly shows the first and second harmonic of the α-orbit and the breakdown frequency F_β (see **Figure 5**).

of this method is the robust and easy set-up and the ease of calibrated measurements. Calibrated measurements are of importance for determining the curvature factor of the extremal orbit in the LK formula. A typical example of an oscillatory magnetic torque is shown in **Figure 6**.

The torque method is one of the first techniques applied by Shoenberg already in 1939, but is still in use in many cases. Whereas the first experiments used the torsion of a wire and optical detection, most of the present devices are based on capacitive detection of the torque. In a simple arrangement, the sample is placed on a small platform on a cantilever. The displacement of the platform is determined by a measurement of the capacitance between the platform and a fixed plate. Adding a calibration coil to the platform allows calibrated measurements. Nowadays, such cantilever spring torquemeters can be fabricated from a silicon wafer using lithographic

techniques. Miniature devices of this kind were even used in pulsed field experiments.

Shubnikov–de Haas Techniques

Measurements of resistance always require a four wire arrangement: a current is applied to the sample via two contacts and the voltage drop between two other contacts is measured. For absolute measurements of the resistivity, the sample should be in a well-defined geometry to assure homogeneous current flow. The application of an AC- current with typical frequencies between 10 and a few hundred Hz allows phase-sensitive detection via a lock-in amplifier. For low-temperature experiments, the current must be small enough to avoid self-heating of the sample.

As mentioned above, the SdH effect is hard to detect in pure metallic elements. At present many SdH studies are carried out in 2D systems. In artificially produced 2D layers, the geometry is easy to adjust by using suitable masks. The typical arrangement is a rectangular bar with current contacts at both short sides and typically two additional contacts on both sides for voltage measurements. This allows one to measure the resistance and Hall effect at the same time. (For the calculation of the conductivity of a pure 2D system, both quantities are needed). In bulk-layered metals the in-plane resistance is often difficult to determine, since the crystals are usually very tiny and the arrangement of well-defined contacts with a homogeneous current distribution is difficult to accomplish. Therefore, most researchers use the interplane resistance for the determination of the SdH effect in these compounds. In that case, a current and a voltage contact are placed on both the large sides of a crystal. Due to very high anisotropy in the resistance (10^3 and more) in these compounds, the interplane resistivity is measured with high accuracy.

See also: Conductivity, Electrical; Diamagnetism; Effective Masses; Electron Gas (Theory); Liquids, Theory of: Fermi Liquids; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Quantum Hall Effect.

PACS: 71.18.+y; 71.10.Ca; 75.20.−g; 72.15.Gd

Further Reading

Abrikosov AA (1988) *Fundamentals of the Theory of Metals*. Amsterdam: North-Holland.

Gordon A, Vagner ID, and Wyder P (2003) Magnetic domains in non-ferromagnetic metals: the non-linear de Haas-van Alphen effect. *Advances in Physics* 52: 385–454.

Joint R and Taillefer L (2002) The superconducting phases of UPt_3. *Reviews of Modern Physics* 74: 235–294.

Pippard AB (1989) *Magnetoresistance in Metals*. Cambridge: Cambridge University Press.

Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.

Wosnitza J (1996) *Fermi Surfaces of Low-Dimensional Organic Metals and Superconductors*. Berlin: Springer.

Silicon, History of

F Seitz, Rockefeller University, New York, NY, USA
N G Einspruch, University of Miami, Coral Gables, FL, USA

In 1781, Henry Cavendish discovered that water is not a primary chemical element, as had been believed for millennia, but a compound formed by the combination of hydrogen and oxygen. This discovery provided Antoine Laurent Lavoisier with insight concerning the true nature of the elements and stimulated him to undertake extensive experimental investigations in a search to identify as many elements as possible. His views were formalized in 1789 by the publication of a two-volume treatise (**Figure 1**) that provided the platform for the evolution of modern chemistry.

In his search for new elements, Lavoisier commonly attempted to reduce with hydrogen what were presumably compounds of the elements, a procedure that was particularly successful with the heavy metal oxides. In the course of this work, he decided that the mineral quartz was probably the oxide of an interesting element, but failed in his attempts to reduce it. A quarter of a century later, the distinguished Swedish chemist Joens Berzelius (**Figure 2**) finally achieved that goal. The newfound element ranked as an oddity. While it had some of the characteristics of a metal, such as a somewhat shiny metallic luster and the ability to conduct electricity, its conductivity was poor compared to that of silver and copper and decreased rather than increased as the temperature was lowered.

Since no immediate use was found for it, early samples of silicon were in the main assigned an honored place within exhibits of the chemical elements. However, the situation with respect to silicon changed radically during the last quarter of the nineteenth century, by which time professional chemists, well-trained

TRAITÉ
ÉLÉMENTAIRE
DE CHIMIE,
PRÉSENTÉ DANS UN ORDRE NOUVEAU
ET D'APRÈS LES DÉCOUVERTES MODERNES;
Avec Figures:
Par M. LAVOISIER, de l'Académie des Sciences, de la Société Royale de Medecine, des Sociétés d'Agriculture de Paris & d'Orléans, de la Société Royale de Londres, de l'Inſtitut de Bologne, de la Société Helvétique de Baſle, de celles de Philadelphie, Harlem, Manchester, Padoue, &c.
TOME SECOND.
A PARIS,
Chez CUCHET, Libraire, rue & hôtel Serpente.
M. DCC. LXXXIX.
Sous le Privilège de l'Académie des Sciences & de la Société Royale de Médecine.

TRAITÉ
ÉLÉMENTAIRE
DE CHIMIE,
PRÉSENTÉ DANS UN ORDRE NOUVEAU
ET D'APRÈS LES DÉCOUVERTES MODERNES;
Avec Figures:
Par M. LAVOISIER, de l'Académie des Sciences, de la Société Royale de Médecine, des Sociétés d'Agriculture de Paris & d'Orléans, de la Société Royale de Londres, de l'Inſtitut de Bologne, de la Société Helvétique de Baſle, de celles de Philadelphie, Harlem, Manchester, Padoue, &c.
TOME PREMIER.
A PARIS,
Chez CUCHET, Libraire, rue & hôtel Serpente.
M. DCC LXXXIX.
Sous le Privilège de l'Académie des Sciences & de la Société Royale de Médecine.

Figure 1 Lavoisier's two-volume treatise on chemistry that opened the doorway to modern chemistry. It was published in 1789. (Courtesy of the library of the American Philosophical Society.)

Figure 2 The Swedish chemist Joens J Berzelius, who isolated elemental silicon in 1823. (Courtesy of the Deutsches Museum, Munich.)

in scientific methods, were in charge of research and development in fields such as metallurgy, ceramics, enamels, pigments, and dyes, replacing the ancient breed of investigators who had depended with some degree of success upon tradition, hunches, and mysticism. There was, for example, much interest at this time in the study of the magnetic properties of iron and its alloys that were being used in electric generators, motors, and transformers for many power-handling applications. Systematic research disclosed the existence of a family of iron–silicon alloys whose magnetic properties were much better than those of pure iron for many purposes. Moreover, relatively modest additions of silicon to some forms of steel could confer superior properties in matters such as toughness. In any event, silicon had finally become a very useful element in technical application and was beginning to be produced commercially at a level of ~98% purity.

During this same period, two other developments that were to influence the use of silicon in fields of technology occurred. First, a young German physicist, Dr. Ferdinand Braun (**Figure 3**), decided to extend prior research carried out earlier by Michael Faraday on the electrical properties of materials that were neither good insulators nor good electrical conductors. His target group included some ionic crystal, such as salts that permit ionic conductivity, and what are now called semiconductors, which were found

Figure 3 Ferdinand K Braun who discovered rectification in semiconductors in 1874, soon after completing his doctorate. He became one of the foremost leaders in the field of radio technology during his time. He demonstrated the important role that resonant circuits can play in long-distance transmission. He shared a Nobel Prize with Marconi in 1909. (Courtesy of the Tuesday Morning Club of Tübingen, the library of the University of Tübingen.)

Figure 4 Reginald A Fessenden, the prolific Canadian inventor who invented amplitude modulation (AM) radio, introduced rectification and the heterodyne principle which made it possible to alter the frequencies of electromagnetic waves. He was also responsible for the first two-way wireless transmission across the Atlantic Ocean on a steady basis and for the invention of sonar. (Courtesy of the Fessenden collection of the Archives of the State of North Carolina in Raleigh, North Carolina.)

eventually to conduct electronically. Both types of substance are distinguished by the fact that their electrical conductivity decreases toward zero at sufficiently low temperatures. Working with specimens of semiconductors, particularly with heavy metal sulfides, Braun discovered crystal rectification in 1874, that is, asymmetrical flow of current in a crystalline specimen for a given voltage when applied in opposite directions. In the course of his investigations, he demonstrated that it was a surface effect, Ohm's law connecting the current and field strength being valid in the bulk of the specimen.

The second significant development was Heinrich Hertz's production of electromagnetic radiation in the meter and shorter range of wavelength with hand-fabricated condensers, inductance, and antennae. This generated widespread interest in the possibility of long-range transmission and reception of such waves, a goal that was led in a very dynamic way by Guglielmo Marconi.

Initially, the propagation of an electromagnetic wave was detected by means of an ingenious device known as the coherer, invented in 1890 by the French physicist Edouard Branly. The coherer contained a bundle of metallic needles that aligned parallel to one another when an electromagnetic wave passed through the needles, thereby altering the collective electrical conductivity of the bundle.

Apparently, at some time toward the end of the 1890s, the prolific Canadian inventor Reginald A Fessenden (**Figure 4**) discovered that one can hear the electromagnetic pulses with earphones of the type used in telephones, provided one rectifies the electric current produced by the signals so that the resulting

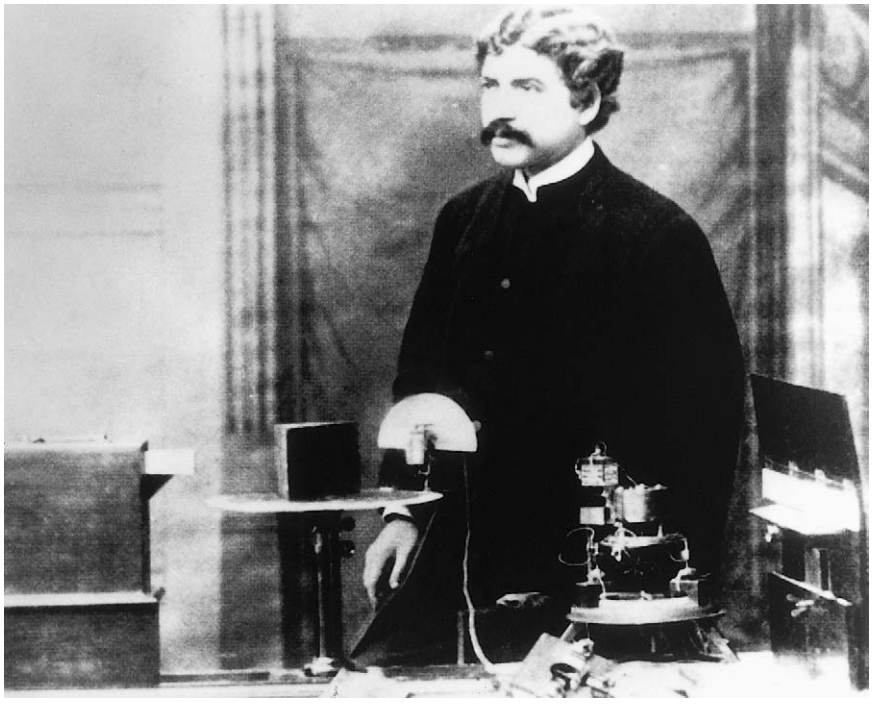


Figure 5 Jagadish C Bose who was at Cambridge University in England when Hertz published his papers on the generation of electromagnetic waves. Bose immediately began research in the centimeter and millimeter range of wavelengths. He invented many items of apparatus that were later re-invented in the 1930s, during the early days of microwave radar. He was the first person to employ a semiconductor (galena) as a rectifier. (Courtesy of the Jagadish Chandra Bose Research Institute, Calcutta, India.)

pulse contains frequencies in the audible range to which the earphones could respond. He invented a rectifier consisting of a wire dipped in an electrolytic solution. Soon after, a brilliant young Indian physicist, Sir Jagadish Chandra Bose (**Figure 5**), who was studying at Cambridge University when Hertz published his research and had done much to improve the coherer, patented the use of crystalline lead sulfide (the mineral galena), a rectifying semiconductor of the type discovered by Braun, as a replacement for the electrolytic device. This represented a very significant milestone in the development of solid-state electronics. Incidentally, Marconi was greatly assisted in his work by a very competent group of electrical engineers of the Italian Navy, who kept abreast of most technical developments in the field. He presumably used a semiconductor rectifier in his first successful transatlantic transmission in 1901.

In 1900 the American Telephone and Telegraph Company (AT&T) served only local communities because of limitations in the distance over which copper wire would transmit voice messages without amplification. The principal city was New York. The management of the company soon decided, however, that it should develop a national network in order to provide more uniform, consolidated service throughout the country. As a result, AT&T began to assemble an excellent technical staff to assist in determining the best route to follow. The range of possibilities included use of a wireless system. Dr. G W Pickard (**Figure 6**), a member of the technical staff who had become interested in the possibility of employing wireless communication, decided to see if there was a crystalline rectifier that was superior to galena. It is reported that he investigated more than a thousand compounds. In any case, carefully selected crystals of the commercially available grade of silicon turned out to have the best properties of all the specimens included in his search. The management of AT&T permitted him to set up an independent company to market the silicon rectifiers he developed.

In the meantime, many scientists began to study the properties of the electron beams (cathode rays) that could be produced in a vacuum from a negatively charged electrode. One of the investigators, John A Fleming, showed that a device containing such an electron source and a positively charged collecting electrode could serve as a rectifier. Subsequently, Dr. Lee De Forest demonstrated that one could add a third electrode to Fleming's rectifier and use it to modulate the magnitude of the collected current. With appropriate feedback, De Forest's triode could be used as the basis for creating an electronic amplifier or an oscillator. Since vacuum technology was fairly primitive at the time, the early versions of the two devices behaved in an

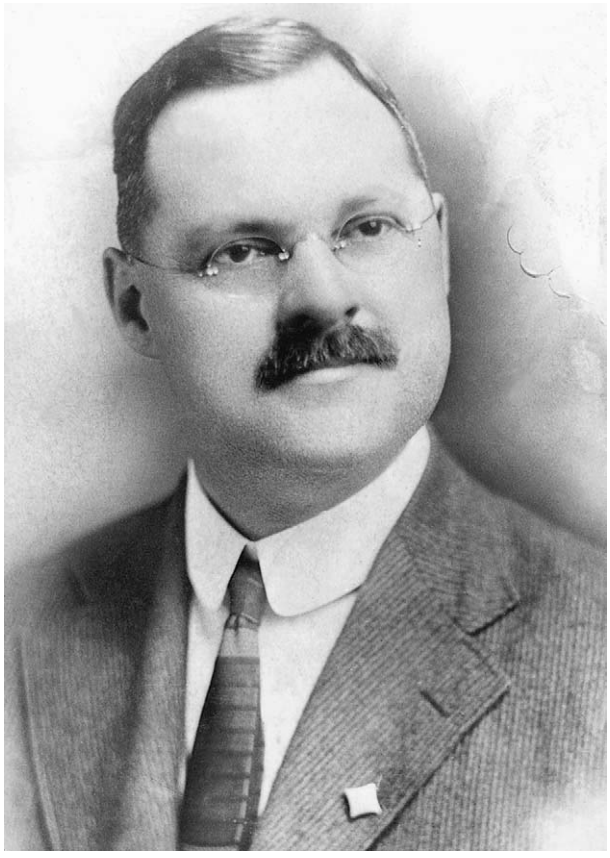


Figure 6 G W Pickard who, in 1905 while working for AT&T, made a thorough search among available semiconductors to find the best rectifier. He concluded that silicon was the preferred choice and began marketing units that contained good specimens. (Courtesy of the Center for the History of Electrical Engineering of the Institute of Electrical and Electronics Engineers, Rutgers University.)

Figure 7 Dr. Mervin J Kelly who was in charge of the production of vacuum tubes at The Western Electric Company, the manufacturing arm of AT&T, in the 1920s. (Courtesy of the Emilio Segre Visual Archives of the American Institute of Physics. Photo by Werner Wolff.)

unpredictable manner and were generally treated as interesting curiosities.

However, the technical staffs of AT&T and the General Electric Company, as well as scientists and engineers in other countries, recognized the importance of the principles involved in De Forest's invention and began to improve both vacuum tube diodes and triodes. By 1915, the telephone system established by AT&T, which incorporated greatly improved vacuum tube relay-amplifiers, was able to span the continent. By the end of World War I, essentially the entire field of electronics was in the hands of vacuum tube experts. Semiconductors were almost forgotten, although inexpensive cuprous oxide and selenium rectifiers were widely used to rectify ordinary fifty- or sixty-cycle current for special purposes, such as providing rectified current to vacuum-tube filaments. As amusement, individuals in their early teens commonly made simple, inexpensive crystal-detector radio receivers that enabled them to listen to local broadcasting stations using earphones. Galena was generally used as a rectifier in such sets. Many older boys entered the world of vacuum tube technology when they could afford to acquire such tubes.

Fortunately, semiconductors were not completely forgotten since they display two important characteristic properties: they do not need a specially designed cathode and associated energy source since some conduction electrons become free at ambient temperatures; as such, semiconductors could be fabricated into monolithic devices whose useful lifetime was not determined by burn-out or other factors that influence vacuum tubes. In continuous operation at full power, a typical vacuum tube has a lifetime of $\sim$1000 h. To overcome this major limitation, a number of individuals attempted to develop a triode analog of the De Forest tube using semiconductors prior to World War II. Some inventors succeeded in obtaining conceptual design patents for proposed devices, but none were actually "reduced to practice."

One of the individuals who took this issue very seriously in the 1920s was Dr. Mervin J Kelly (**Figure 7**), who was then in charge of the production of vacuum tubes at the Western Electric Company, the

manufacturing arm of AT&T. He was thoroughly conscious of the limitations of vacuum tubes and was captivated by the thought of replacing, or at least complementing, them by semiconductor devices. He decided that if he were ever made Head of the Bell Telephone Laboratories, he would establish a team to explore semiconductor-based technology fully. He became head of the Laboratories in the mid-1930s and, in keeping with his plans, hired Dr. William S Shockley, a doctoral graduate of the Massachusetts Institute of Technology, to join forces with Dr. Walter H Brattain, who was already at the Laboratories, in characterizing the properties of various semiconductors. While they made a preliminary study of cuprous oxide, they were diverted to other problems as World War II approached. Kelly's plan was forced to await the end of the war.

Research and development in the field of radar was a major activity among scientists and engineers during World War II. Although many significant investigations were carried out earlier, the first microsecond pulsed radar was developed in 1932 by an ingenious engineer, Dr. Hans E Hollmann, who used it to study the reflections of electromagnetic waves from the ionosphere in the upper atmosphere of the Arctic region – the layers that make transmission of radio waves in the kilohertz and lower megahertz region of the radio spectrum possible over long distances on the curved earth. He observed the reflected pulses on an oscilloscope with a synchronized time sweep and found that he could pick up features of the landscape such as mountain ranges. The German Navy used this knowledge to produce its first radar system soon thereafter. It operated with wavelengths in the vicinity of 2.4 m. Dr. Robert Page achieved the first comparable development in the US at the Naval Research Laboratory in 1939, delays in funding having retarded research that was initiated in the 1920s.

It should be added that Hollmann so distrusted Hitler and his plans for world conquest that he discontinued all research that might be employed in offensive warfare and focused on medical applications of electronics. While making the shift, however, he wrote a two-volume treatise in the field of high-frequency electronics that was by far the most advanced work of its kind when it appeared in 1936. At the end of the war, he migrated to the US at the invitation of the government and became an expert consultant to industry during the early development of transistor technology.

In the meanwhile, the British, French, and Russians were pursuing the development of radar for their own military needs. In addition to carrying on research in the meter range that was to be used in the long-range perimeter defense of their island, the British were eager to increase the resolution of radar images and decrease the size and weight of equipment for airborne and other critical uses. To achieve this purpose, they extended their research to the decimeter and centimeter range of wavelengths. Consequently, a young English engineer, Denis M Robinson (**Figure** 8), who had been employed in servicing British television transmitters, which had broadcast programs several times a week until 1939, found himself seconded to a secret laboratory in Scotland by The Telecommunications Research Establishment (TRE), the agency that was guiding radar development in Britain. He was instructed to produce a radar system that could operate at a wavelength of 10 cm.

Figure 8 Denis M Robinson who decided early in WWII, and after reading the treatise on microwave technology written by Hans Hollmann, to search for a semiconductor that could serve as a suitable heterodyne mixer in the 10 cm range of wavelengths. He joined forces with HWB Skinner who determined that commercially available silicon would serve their immediate purpose adequately. (Courtesy of the book *Five Years at the Radiation Laboratory* (Massachusetts Institute of Technology, 1946). Reprinted by IEEE in International Microwave Symposium, 1991.)

The means to generate and transmit such radiation was already available. Devising appropriate equipment to process and display the return signal, however, presented a problem. What Robinson needed most was a nonlinear device, or "mixer," that had the capability of converting the reflected return signal to a frequency that could be handled by available amplifiers and the auxiliary equipment involved in display. There was no

available vacuum tube, such as a rectifier, that could serve as a mixer in the low centimeter range. Developing one during wartime would take valuable time with an uncertain outcome. He went to a University library nearby and found Hans Hollmann's two-volume treatise mentioned above. This treatise asserted that crystal rectifiers could be useful as rectifying detectors in the centimeter range of wavelengths. Robinson remembered the use of galena from his teen years and decided to team up with Dr. H W B Skinner, previously at the University of Bristol and later also at TRE. The latter was much more familiar with the so-called solid-state devices than Robinson was. After some deliberate investigations, Skinner ascertained that selectively chosen specimens of silicon, available in its somewhat variable impure metallurgical form, would provide the answer to their immediate problem of fabricating a mixer. It is interesting to note that Dr. R S Ohl, working in the field of microwave research at the Bell Telephone Laboratories in close cooperation with Dr. G C Southworth, had recently come to the same conclusion in a search for a suitable detector, thereby reaffirming the conclusion reached by his former colleague, Dr. G W Pickard, some thirty years earlier at the start of wireless communication.

When the government of France fell before the advancing German army in May of 1940, a high-level group of French scientists and engineers had been involved in research on a very advanced form of microwave generator known as the cavity magnetron. It operated at wavelengths in the vicinity of 16 cm. At one of the last possible moments, Maurice Ponte, one of the leaders of the French group, flew their working model to England, where it added significantly to the development of the magnetron being carried on there. The French, it turned out, had not yet faced up to the problem of searching for the best heterodyne mixer in that wavelength range and marveled in their turn at the effectiveness of the silicon diodes that the British were using.

Once France had fallen and Britain was isolated from any immediate help from the continent, Winston Churchill decided that he must form the closest possible relationship with the US, which was still formally neutral but was preparing to become fully armed. As a positive step, Churchill offered to share with the US all secret technical research with which Britain was involved, including its developments in radar. President Franklin D Roosevelt readily agreed to the plan and established an Office of Scientific Research and Development within the Executive branch of the government.

In further keeping with Churchill's suggestion, Roosevelt selected two highly respected scientists to head the office, namely James B Conant and Vannevar Bush, the former from Harvard and the latter from the Massachusetts Institute of Technology (MIT). One of their first steps was to create the Radiation Laboratory, a radar research and development center, at MIT, to supplement on a large scale what the British had already been developing, with particular focus on applications in the microwave region. Dr. Lee A DuBridge, an already distinguished younger scientist-administrator, was selected to be the director; he would remain so throughout the war. The laboratory was rapidly staffed by a highly selective group of scientists and engineers, many from the top ranks of university faculties.

Most of the staff of the Radiation Laboratory were not familiar with semiconductors other than in a very peripheral way, such as the crystal radio sets of their early teen years. To them, the British-made silicon rectifier-mixers that were used for frequency conversion of the returning radar signals were both a curiosity and an abomination because of the great variability of the metallurgical-grade silicon used as starting material. Many on the staff felt that they should take the time to develop vacuum tube rectifiers that would operate in the microwave region.

Fortunately, DuBridge was familiar with research on crystalline solids and called on one of the authors of this article (FS) to see if a way to salvage the situation could be found. An excellent physical chemist, Dr. C Marcus Olson (**Figure 9**), in the Pigments Department of the DuPont Corporation near Wilmington, Delaware, found that he could produce a form of silicon that was uniformly pure to several parts in one hundred thousand by reacting purified silicon chloride with comparably pure metallic zinc. When this form of the element was properly treated with controlled additions of other elements that influenced its electrical conductivity, it could be used to produce highly satisfactory silicon diodes on a mass-production basis (**Figure 10**). Moreover, DuPont was prepared to manufacture the purer grade of silicon in practical quantities.

Once this point was reached, the Radiation Laboratory added a semiconductor section under the leadership of Dr. Henry C Torrey. It followed progress in industrial production and administered an extended research program, including research on diodes made from the semiconductor germanium, the sister-element to silicon, which found very special uses in situations in which the rectifier experienced high reverse voltages. Both semiconductors permitted conductivity by both negative and positive carriers depending upon the addition of small controlled quantities of foreign elements, known as "dopants." The positive carriers ("holes") were associated with a dearth of electrons in what normally are fully occupied shells or bands of

Figure 9 Dr. CM Olson (center), on the staff of the Pigments Department of the DuPont Company. During the early part of World War II he developed a procedure for producing elemental silicon that was sufficiently pure to permit routine mass production of silicon diodes. (Courtesy of CM Olson.)

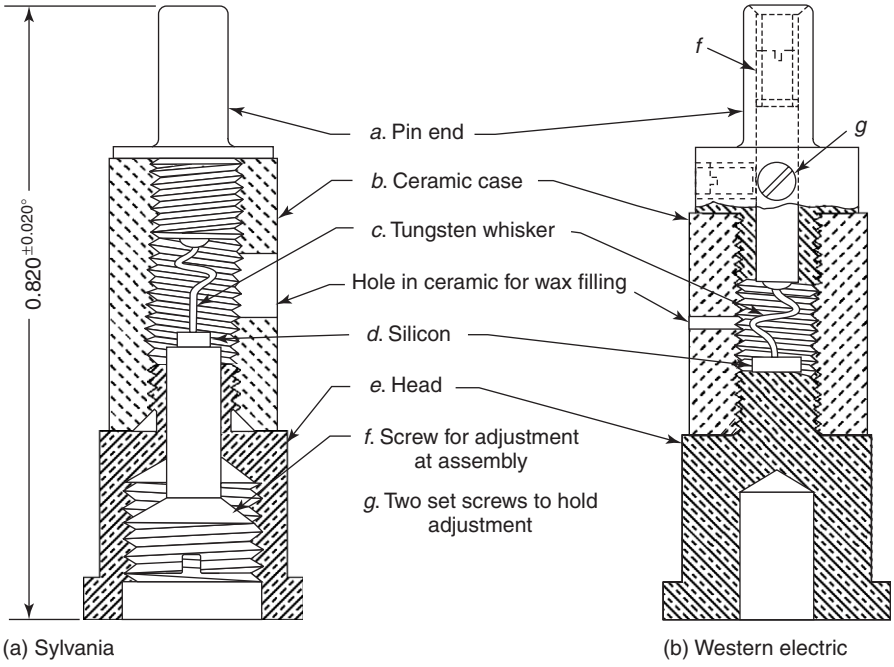


Figure 10 Two typical silicon rectifiers of the World War II period. The tip of the pointed tungsten catwhisker rests on the silicon chip. (Courtesy of archives of the Radiation Laboratory of The Massachusetts Institute of Technology. From the book *Crystal Rectifiers* by HC Torrey and CA Whitmer in the series published by the Radiation Laboratory of the Massachusetts Institute of Technology at the end of World War II under contract with the US Government.)

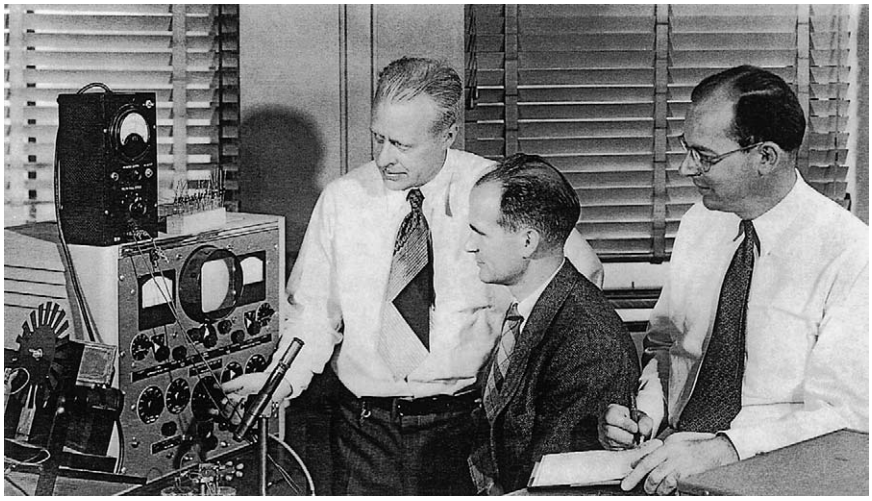


Figure 11 John Bardeen, William S Shockley, and Walter H Brattain (right to left) who developed the first transistors at the Bell Telephone Laboratories. The three shared a Nobel Prize for their discoveries. (Courtesy of Professor Lillian H Hoddeson and Lucent Technologies Bell Laboratories.)

electron states in silicon. The empty states are equivalent to bubbles in a fluid, which will move opposite to the direction of any applied forces that cause acceleration of the fluid. The two types of conductors were designated *n*-type (negative) and *p*-type (positive), respectively.

When World War II ended, Dr. Kelly returned to the secret pursuit of a semiconductor triode armed with an increased staff and the extensive new knowledge of semiconductors gained during the war. Shockley was appointed head of the group and had the good fortune to attract John Bardeen, a highly experienced engineer-physicist who possessed strong background knowledge in solid-state science, as part of the team. Brattain who knew Bardeen well joined him in a remarkably productive working partnership (**Figure 11**).

As a first step, Shockley attempted to develop what later became known as the field-effect transistor in which the density of carriers in the bulk semiconductor is altered by application of an electric field transverse to the direction of current. He failed to achieve the desired effect and turned to other matters, leaving his colleagues to carry on. Bardeen suspected that the failure was a result of a temperature-related shift in surface charge that compensated the applied field within the bulk of the semiconductor. He and Brattain demonstrated this by repeating the experiment at temperatures sufficiently low so that the surface charges would be frozen. Much later, means of stabilizing the surface charge at normal operating temperatures were found and the resulting "field effect transistor" became an important member of the transistor family.

Figure 12 Gordon K Teal who successfully pursued the development of monocrystalline silicon-based transistors, first at the Bell Telephone Laboratories and then at Texas Instruments. (Courtesy of The American Institute of Physics Emilio Segre Visual Archives.)

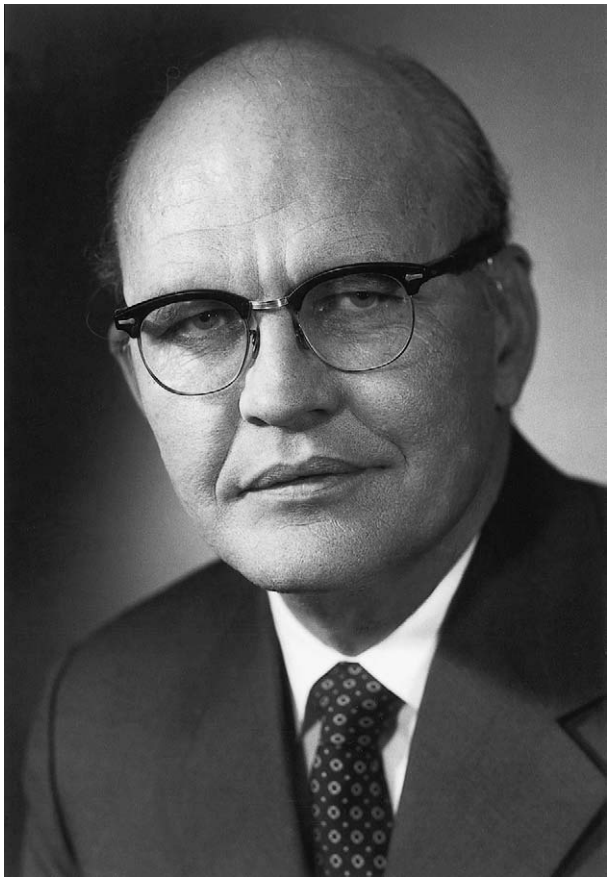


Figure 13 J S Kilby who invented the integrated circuit and received a Nobel Prize for it. (Courtesy of Kilby JS.)

While extending this research in December of 1947, Bardeen and Brattain invented, partly by chance, what became known as the "point-contact transistor." It was the first truly amplifying semiconductor triode. Shockley returned to research in this field at this point and invented what became known as the "discrete *p–n* junction transistor." The principles involved in its operation were to play a major role in the development of electronics during the next decade.

In the course of events, one of the members of the Laboratories developed a method for purifying germanium to a very high level that was particularly applicable to that element. As a consequence, Shockley mandated that all research on transistors under his direction be carried out with crystals of germanium. A courtly strong-minded Texan on the staff, Dr. Gordon K Teal (**Figure 12**), aware of the superior chemical and physical properties of silicon from wartime research, began to produce single crystals of it surreptitiously in his laboratory. Teal shifted his activity to Texas Instruments Incorporated, in Dallas, Texas, in 1953 when that company obtained a license from AT&T to produce transistors. The company gave him a warm welcome in his pursuit of the work on silicon. Methods were soon found to purify silicon to parts per billion and to improve crystal perfection, which Teal knew to be very important. The element became the mainstay of production of transistors, particularly when the beneficial characteristics of its relatively

Figure 14 A single crystal ingot of silicon having a diameter of 300 mm. Such an ingot is sliced horizontally into circular wafers whose surfaces become the site of integrated circuits. (Courtesy of H Fusstetter of Wacker and Siltronics AE.)

stable oxide were fully appreciated. Nevertheless, much research continued to be carried out with germanium. In fact, discrete transistors made of silicon or germanium were manufactured throughout much of the 1950s.

The invention of the integrated circuit by Jack Kilby (**Figure 13**) in 1958, led to previously unforeseen increases in circuit complexity, eventually at very low cost – a process requiring a good part of a decade. It was not clear at first whether such circuits, which permitted the creation and interlinking of vast arrays of transistors on a single silicon chip, would find uses other than in military applications where a premium price might be acceptable. By the mid-1960s, however, the most farsighted manufacturers began to express optimism. For example, in 1964 Patrick E Haggerty, the President of Texas Instruments, published a paper in the Proceedings of the Institute of Electrical and Electronics Engineers in which he stated that if the remaining problems of large-scale production were resolved in an economical manner, one could expect devices based on the use of integrated circuits to become "pervasive" in human affairs, creating a revolution in such matters as information processing, communications, and controls. A year later Dr. Gordon E Moore, then at Fairchild Semiconductor and later a founder of Intel, had the foresight and courage to predict that the cost of integrated circuits would soon go through a period in which the cost per transistor would be halved about every eighteen months, accelerating their use. Widespread applications of the device finally reached Moore's expansive stage by the end of the 1960s, along with the rapid growth in sophistication of the technology associated with design of operating systems and applications software, the linchpins of the computer era. The end of the evolution of this field is not yet in sight. All evidence suggests that silicon will continue to play an unending role in electronics (**Figure 14**).

See also: Semiconductors, History of.

PACS: 61.72.Tt; 01.65.+g

Further Reading

Seitz F and Einspruch NG (1998) *The Electronic Genie*. Urbana, ILL: University of Illinois Press.

Single-Molecule Methods in Biophysics

Y Ishii and T Yanagida, CREST, Suita, Osaka, Japan

Introduction

Biomolecules such as DNA and proteins assemble into molecular machines and networks in biosystems such as cells. The functions performed in molecular machines and biosystems reflect the characteristic features of the biomolecules. Recent developments in molecular cell biology have allowed the biomolecules and their roles to be identified. Advances in structural biology have allowed the basic static structures to be determined with atomic resolution. These molecules, however, act in a kinetic and dynamic fashion, when they function. To understand the underlying mechanism of the functions it is therefore essential to monitor the dynamic and kinetic behaviors of biomolecules under their normal working conditions.

In conventional ensemble measurements faint signals from individual molecules are accumulated from a large number of molecules. In the averaging process the signals are amplified and the S/N ratio increased but some of the information such as the dynamic, kinetic, and fluctuating properties of biomolecules is hidden. Recent developments in lasers, detectors, and other techniques have made it possible to detect faint signals from individual molecules. These techniques have allowed the behavior of the individual biomolecules to be detected without averaging. The single-molecule measurements have provided unique information, which was not possible to obtain in the conventional ensemble measurements. The information is important particularly in elucidating the mechanisms underlying the function of biomolecules and biosystems.

Single-Molecule Fluorescence Imaging

Single-molecule measurement techniques include the imaging and manipulation. Single molecules can be visualized by attaching a fluorescent marker. Fluorescence is the technique in which only the molecules that have been marked are visualized. The unnecessary signals such as light scattering are filtered out by using different wavelengths for the excitation and observation. Following the visualization of single

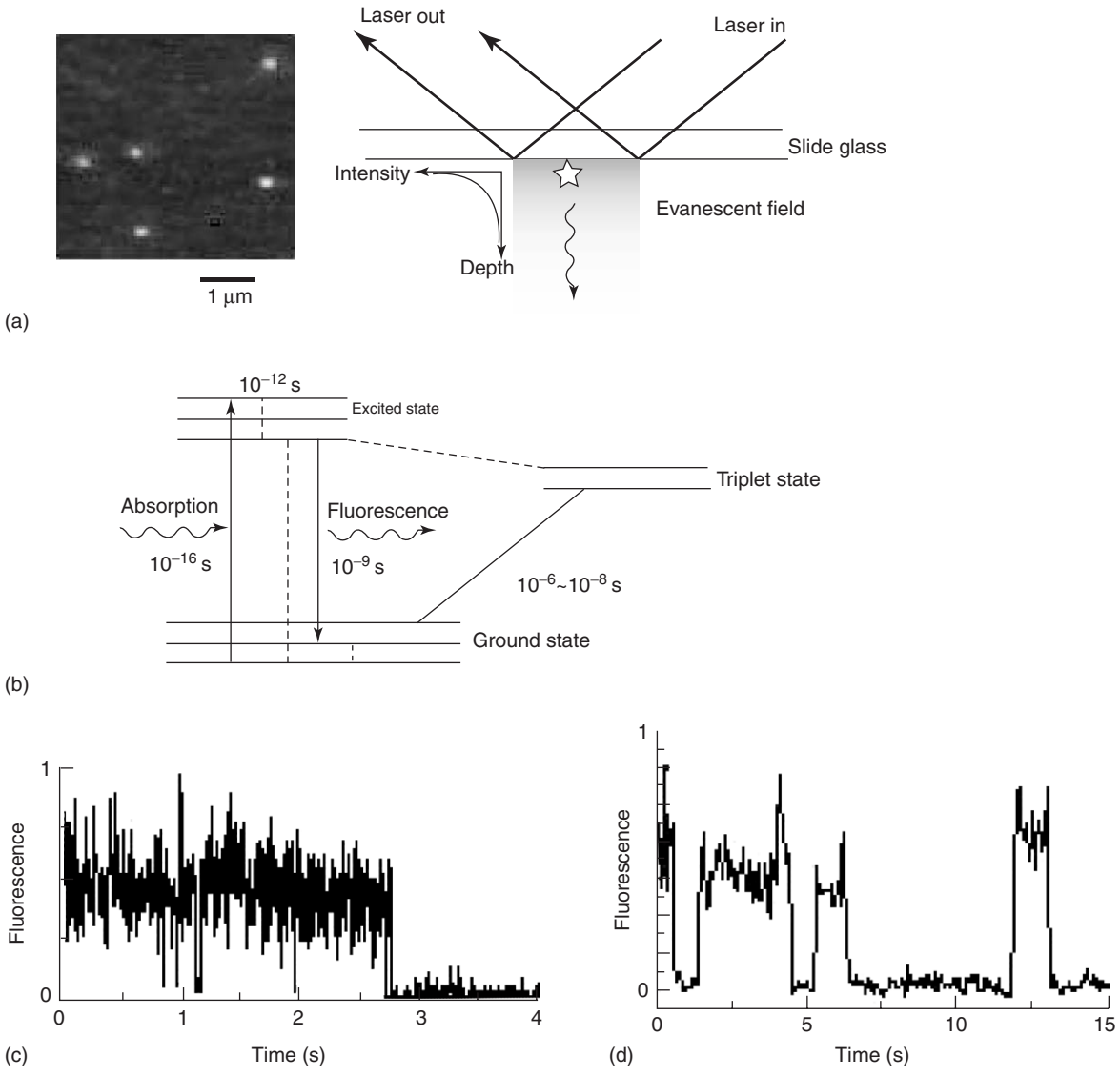


Figure 1 Imaging single fluorescent molecules. (a) Fluorescence images from single fluorophores using total internal reflection fluorescence (TIRF) microscopy: (left) single spots come from single fluorophores; (right) a schematic diagram of TIRF microscopy. (b) Mechanism of fluorescence emission. Three electronic states (ground, excited, and triplet states) contain several vibrational levels (horizontal lines). Transitions between these states (straight or dashed vertical lines) occur when fluorophores interact with light (wavy lines) or in radiationless process. The order of the time constants is indicated. (c, d) Typical time record of fluorescence from single fluorophores including photobleaching (c) and blinking (d).

fluorophores in dry air, visualizing single fluorophores attached to active biomolecules was accomplished in aqueous solution in 1995. For the measurements of active biomolecules at work, it was essential to measure biomolecules in similar condition in which they work, that is, in aqueous solution. The fluorescence in aqueous solution, is however, low mainly because of "quenching" caused by the collision of surrounding water molecules.

Since the fluorescence from single fluorophores is faint, it is essential to reduce the background noise to increase the S/N ratio. This was achieved by local illumination. Total internal reflection fluorescence (TIRF) microscopy was the most appropriate method to image single molecules (**Figure 1a**). When a laser is irradiated onto a surface between different refractive index media from the high diffractive index medium at angles greater than a critical angle, it is completely reflected. At the same time the laser generates electromagnetic fields called evanescent fields only in the vicinity of the interface to a depth of ~ 100 nm in the medium of low refractive index. For example, a laser is incident through a slideglass and the evanescent field is generated on the surface in the aqueous sample solution. The fluorophores on the glass surface can be visualized over the irradiated areas. Other scanning

probe microscopy such as confocal microscopy and near-field microscopy can also be used to visualize single molecules. TIRF microscopy is advantageous over other microscopy techniques for the systems where molecules move on the surface.

Single biomolecules are observed as fluorescent spots. Fluorescence spots are accumulations of photons emitted from single molecules. A single photon results from a single cycle of excitation and emission (**Figure 1b**). When fluorescent molecules are excited by an excitation light, the electronic state of the molecules is shifted from a ground state to an excited state using the energy of a photon of the excitation light. The molecules return to the ground state via emission of fluorescence with a lifetime in the order of nanoseconds or return via radiationless processes. The wavelength of fluorescence is longer than that of excitation light because energy is lost due to the rapid relaxation of the vibration levels within the excitation and ground states after the absorption and emission of light, respectively. The fluorescence intensity has been determined through the probabilities for adsorbing excitation light and emitting fluorescence presented by an absorption coefficient and a quantum yield, respectively. This process occurs stochastically and the emission of photons is random with time. The number of photons acquired in a measurement time or fluorescence intensity fluctuates with time. In general, the width of the fluctuation is $\sim N^{1/2}$ when the number of the photons for the signal is N. The S/N ratio increases with the number of photons.

After a fluorescent molecule emits a large number of photons, it suddenly ceases to emit photons (**Figure 1c**). This process is irreversible and is known as photobleaching. In the case of tetramethylrhodamine (TMR) or Cy3 which is used for single-molecule measurements, the photobleaching occurs after a molecule emits approximately 100 thousand to 1 million photons. The sudden drop of the fluorescence observed in the recording time is characteristic of single molecules and has been used as a test whether the fluorescence spots observed come from single molecules. The fluorescence from a single fluorophore drops in a single step and that from two molecules in two steps. Photobleaching occurs stochastically and the time for the photobleaching is distributed in an exponential manner. In agreement with this distribution, the fluorescence intensity decreases gradually with time due to the photobleaching in ensemble measurements. The photobleaching time shortens with an increase in laser power. The number of photons that single molecules emit until the photobleaching event is characteristic for each fluorescent dye. Oxygen is known to accelerate the photobleaching process so an oxygen scavenger system is useful in single-molecule measurements.

Blinking, an event in which the fluorescence is turned on and off reversibly, is more or less found in almost all fluorescent dyes (**Figure 1d**). The dark state which lasts from about a millisecond to a second is thought to be related to a triplet state. The power of the laser affects the interval time between the dark states but does not affect the duration time in the dark state. Blinking does cause problems in interpreting the data. It is difficult to discriminate blinking from the turning on–off of the fluorescence signals originating from the changes in the biomolecule systems. In general, blinking depends on the power of the laser but the fluorescence changes originated from biomolecules do not. Blinking does not depend on the solution conditions whereas the fluorescence changes originated from changes in biomolecules occur depending on the solution conditions. The blinking may or may not occur under conditions used for the ensemble measurements. In the single-molecule measurements, intense excitation light is used so that single fluorophores are excited very frequently. In the ensemble measurements, if blinking occurs, it would be hidden.

In the fluorescence images, the spots from molecules of nanometer size appear as large as the order of the wavelength (several hundred nanometers) due to the diffraction limit of the light. Molecules which exist within ~ 100 nm cannot be distinguished. For measurements, the concentration of the protein must be lower in order to distinguish individual spots. It is difficult to discriminate two separate fluorescence spots at this resolution; however, it is possible to determine the change in the position of single spots with greater resolution. Recently, the positions have been determined with an accuracy of 10 nm with the assumption of the distribution of the fluorescence around the molecules. The determination of the spatial resolution is dependent on the standard deviation of the distribution of photons.

There are only a limited number of fluorescent dyes that are suitable for the single-molecule measurements. The dyes must have a high intensity (high absorption coefficient and high quantum yield), a high stability against photobleaching, and a high stability against blinking. Many organic compounds such as TMR and Cy3 have been used to attach specifically to biomolecules. Natural fluorescent proteins have been used to fuse to proteins. Green fluorescent protein (GFP) and its mutants, which have different fluorescent colors, can also be used. Quantum dots are new materials which can prolong the photobleaching time. Further development of fluorescent dyes remains a very important research topic for single-molecule measurements.

Imaging the Association and Dissociation Events

Using single-molecule imaging it is possible to visualize the binding and dissociation of fluorescently labeled molecules to partner molecules at a fixed position. If fixed molecules are labeled with different fluorescent dyes, the position of the molecules that are fixed can be marked. At the location of the fixed molecules, fluorescent spots appear when the molecule binds and disappears when it dissociates. When the fluorescent molecules are free in solution, they undergo rapid Brownian movement. They are not observed as spots but contribute to the background noise. The concentration required to successfully image the binding of single molecules is limited below tens of nM. Thus, it is difficult to measure the binding with an affinity below $10^6\,M^{-1}\,s^{-1}$ with this assay. The duration time from the appearance of the fluorescence to its disappearance is related to the dissociation rate. The time from the disappearance of the fluorescence to the reappearance of the fluorescence signal due to the binding of the next molecule is related to the association rate. The association rate depends on the concentration of the molecules. At lower concentrations, the time between events is longer and single cycles of binding and detachment are more easily identified. The duration time varies from event to event and the decay time can be calculated from histograms. In equilibrium, the ratio of the two rates gives an association constant. The association constant can also be presented by the ratio between the sum of the duration time in the bound and unbound states.

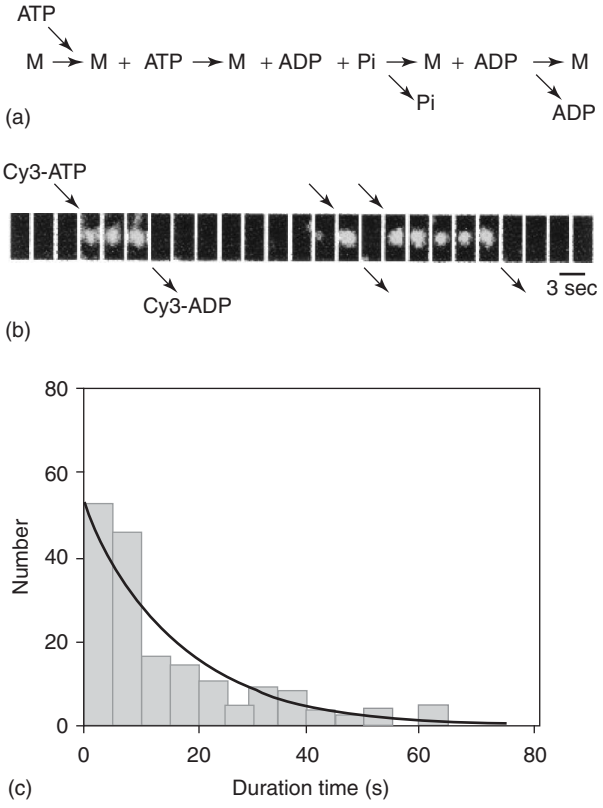


Figure 2 Imaging chemical reaction of single ATP molecules. (a) Chemical reaction of the ATP hydrolysis to ADP and inorganic phosphate (Pi) by myosin (M). (b) Time record of the ATP fluorescence image at the position of a single myosin molecule. (c) A histogram of the duration time for the ATP turnover. The data were fit to a single exponential curve with the decay time of $0.045\,s^{-1}$.

Some enzymatic reactions involve the binding and unbinding of the ligand. The ATP hydrolysis cycle includes the binding of ATP and dissociation of ADP after the hydrolysis of ATP to ADP (**Figure 2a**). Fluorescently labeled ATP (Cy3-ATP) was added to myosin molecules attached on the surface of slide-glass (**Figure 2d**). The duration time from the binding to the dissociation event could be well fitted to the single exponential decay and the average duration time gave the turnover time of ATP, which corresponds to the ATP hydrolysis rate at the limit of saturating ATP concentration (**Figure 2c**). The energy released from the ATP hydrolysis is utilized as a fuel for the function of many molecular machines including molecular motors. The visualization of this reaction is important for the studies on the function of these molecular machines.

Detecting the Dynamics of the Biomolecules

The fluorescence emission processes are sensitive to the environmental changes near the fluorophores. The techniques monitoring these changes have been developed as fluorescence spectroscopic methods. When fluorophores are attached to protein molecules, the structural changes occurring in the protein molecules are sensed by fluorescence through the environmental changes of the fluorophores. The changes in the protein structure cannot be resolved directly in fluorescence microscopy because of the diffraction limit. It is possible to monitor the changes in the structure of single molecules by combining the spectroscopic methods with single-molecule imaging. In addition to the structural changes, changes such as the interaction with other molecules, which affect the state of the fluorophores, can also be monitored.

The environmental changes alter the energy levels of the excited and ground states of the fluorophores by altering their electronic states. These changes are detected as changes in the fluorescence spectrum. For the fluorescence spectrum the number of photons emitted is measured as a function of wavelength (**Figure 3a**).

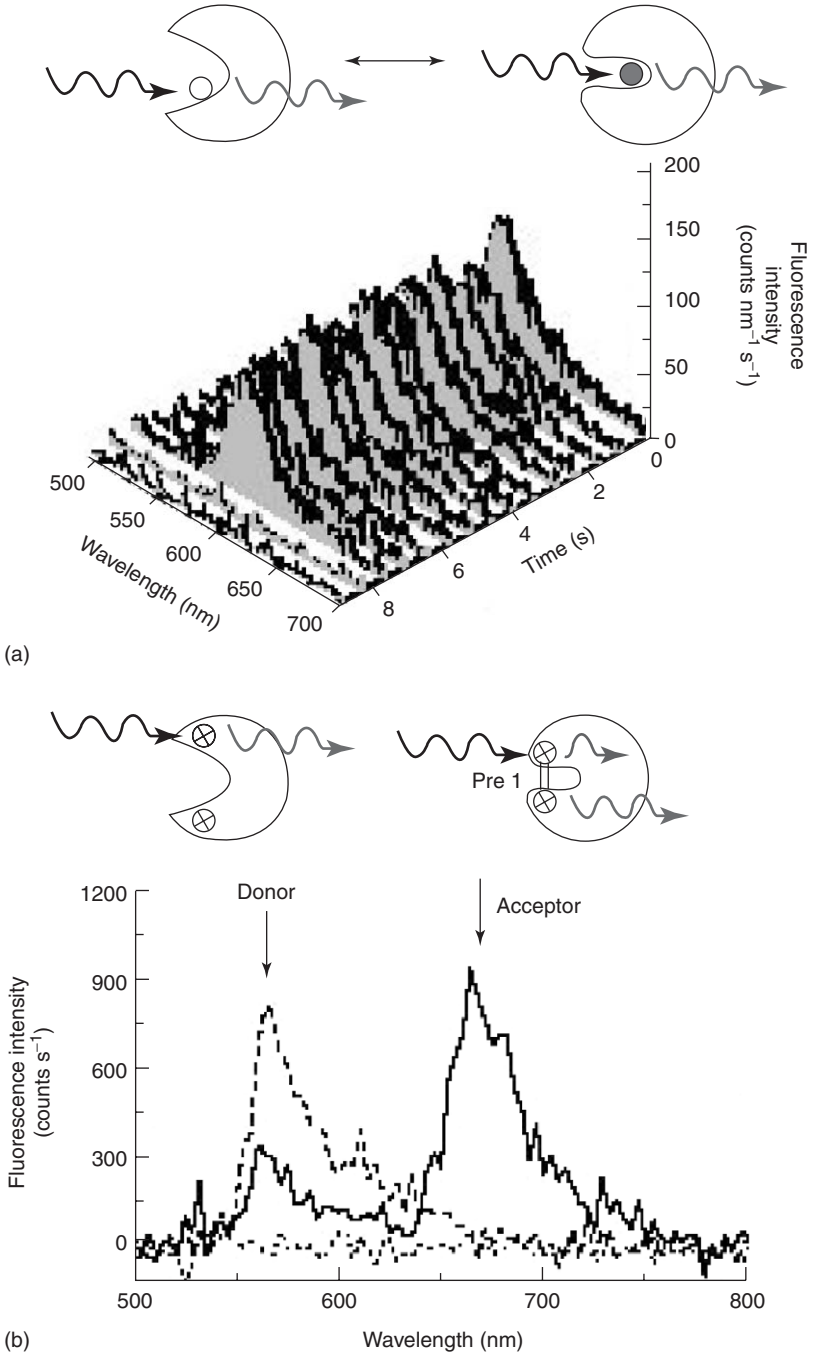


Figure 3 Spectroscopic methods to monitor the dynamic changes of single protein molecules. (a) Fluorescence spectrum method. The time record of the fluorescence spectrum from TMR attached to a single myosin molecule. (b) Fluorescence resonance energy transfer (FRET) method. The fluorescence spectra of FRET were taken from single tropomyosin molecules (straight line). The spectrum in broken line was taken after the donor was photobleached and the spectrum in dashed line was after both the donor and acceptor were photobleached.

The changes in the fluorescence spectrum are depicted by several parameters. The fluorescence intensity is one of these parameters. However, the fluorescence intensity is variant and it is difficult to compare the intensity between different measurements, because the intensity depends on the instruments, experimental conditions, and preparation of the sample. In contrast, the wavelength is an invariant number. The determination of the peak wavelength is useful when the position of the spectrum shifts. The microenvironmental changes affect the spectral changes including the hydrophobicity, the structure of water, and the

electronic potential created by the charges around the fluorophores.

Absorption and emission take place through the interaction of photons with the absorption and emission dipole moments of the fluorophores, respectively. The polarity of the fluorescence against the polarized excitation light reflects the relaxation of the dipole moment during the lifetime of the fluorescent dyes in the order of nanoseconds. When the fluorescent dyes attached to biomolecules are allowed to rotate, rotational relaxation of the dye molecules can be measured. The relaxation of the rotational Brownian motion is affected by the microenvironment and the protein structure. When the dye molecules are attached tightly to large biomolecules, the mode of the rotational motion or flexibility of the local regions of biomolecules can be measured. Fluorescent dyes attached tightly to the motor protein molecule probes can be used to detect structural changes during movement, with single-molecule measurements.

Fluorescence resonance energy transfer (FRET) is a more direct method to monitor changes in the geometrical relationship between two different fluorophores attached on the biomolecules in the nanometer order (**Figure 3b**). When the donor fluorophore is excited, the excited energy is transferred to the acceptor through the dipole–dipole interaction, resulting in the emission of fluorescence from the acceptor. The efficiency of the FRET depends on the geometry between the two dyes, but primarily on the distance between the two dyes. This method can be used to measure the structural changes in biomolecules and the changes in the interaction between two molecules. The fluorescence lifetime is also sensitive to the microenvironment of probes and the FRET efficiency.

The accumulation of a number of photons is necessary for a single data point. These data are then analyzed to obtain the fluorescent parameters. In addition to the stochastic nature of the fluorescence emission process, the process for detecting photons is stochastic. Whether the origin of the stochastic process is quantum mechanics or a statistical mechanism, the detection of a single photon does not have any physical basis. For example, in the measurements of FRET, single photons are detected as either a donor or acceptor. When a number of photons are measured, the distribution of the donor and acceptor can be obtained and the ratio of the donor and acceptor gives the FRET efficiency. The measurement time must be sufficient to acquire the appropriate number of photons for statistical analysis and at the same time short enough to monitor the dynamic changes if any exist. Changes that occur faster than the acquisition time are averaged.

For spectroscopic measurements it is possible to measure and compare more than two parameters at the same time. For example, the fluorescence spectrum includes several parameters such as the peak position, fluorescence intensity, and the half-width of the spectral band. The interpretation of the data, however, is complex. The conclusion reached will depend on the parameter being investigated. One parameter may not distinguish two states but another parameter will. The use of multiple parameters is necessary for a more solid conclusion. The measurements of the structural changes in proteins are similar. The choice of parameters and the position of the labeling is critical. There is currently no general theory that describes which parameters can describe the structural and state changes of proteins.

Studies using these technologies have shown that many proteins are in equilibrium between multiple conformations in solution; individual molecules have different conformation and the conformation of individual molecules changes spontaneously with time in the time range of 100 ms or even $\sim$1 s. Proteins are the main constituents of molecular machines and are responsible for their functions. This dynamic picture of the structure of a particular protein may play an important role in their functioning.

Detection of the Movement of Single Molecules

It is possible to trace the motion of fluorescently labeled molecules if the molecules move at an appropriate speed, that is, hundreds of $\mathrm{nm\,s^{-1}}$ to several $\mu\mathrm{m\,s^{-1}}$. Brownian movement of fluorescently labeled proteins and lipid molecules on biomembranes can be monitored for both artificial and natural cell membranes. In living cells, the movement reflects the structures of the cell membrane and their interactions with molecules of the cytoskeleton, which is important in relation to their functions.

For the advancement of studies on molecular motors, it is essential that the movement of these motors be visualized. In 1985, the sliding movement of fluorescently labeled actin filaments on myosin fixed on a slideglass was visualized for the first time using purified proteins under a microscope. Since then it has been demonstrated that single fluorescently labeled molecular motors move smoothly along protein tracks fixed on a slideglass. Kinesin and certain types of myosin can move for long distances without dissociating.

Larger and brighter probes instead of single fluorophores are advantageous for the detection of these molecules. Fluorescent beads attached to myosin and

kinesin showed that linear molecular motors move along track proteins. A fluorescently labeled actin filament attached to a rotary motor protein F1 as a marker has allowed the visualization of this rotary motion, providing the first direct evidence for the existence of rotary motors. In this motor the rotational motion is tightly coupled to the hydrolysis of ATP, and was found to be very efficient: the work done at each step was almost the same as the energy released from the hydrolysis of a single ATP molecule.

Manipulation of Molecular Motors and the Unitary Steps

The manipulation techniques have been used to carry out measurements under well-controlled conditions. Glass microneedles or beads trapped by a laser have been developed to catch single molecules and manipulate them. For example, single motor proteins attached to beads trapped by a laser can be brought into contact with single filaments of the track proteins (**Figure 4**). These systems can also be used for the mechanical measurements of molecular motors, because the displacement can be determined with high resolution and glass microneedles and laser traps have similar properties to a spring. These probes exert a spring force to pull the molecules back to the original position when they are displaced. Thus, motor molecules attached to a large probe move against the spring force until the force of molecular motors and the spring force are balanced and then the movement stops. The displacement of large probes has been measured using a divided photodiode with nanometer and milliseconds accuracy. The measurement of the displacement of the probes gives the displacement of molecular motors, if the probes attach tightly to the proteins and follow the quick movement of the molecular motors. To measure the displacement without disturbance due to Brownian motion of the probes, the stiffness of the measurement system, in general, must be high. Compliant linkages between the probes and the proteins make the system less stiff. To allow the probes to follow the rapid movement of the motors, the friction coefficient of the probes must be low.

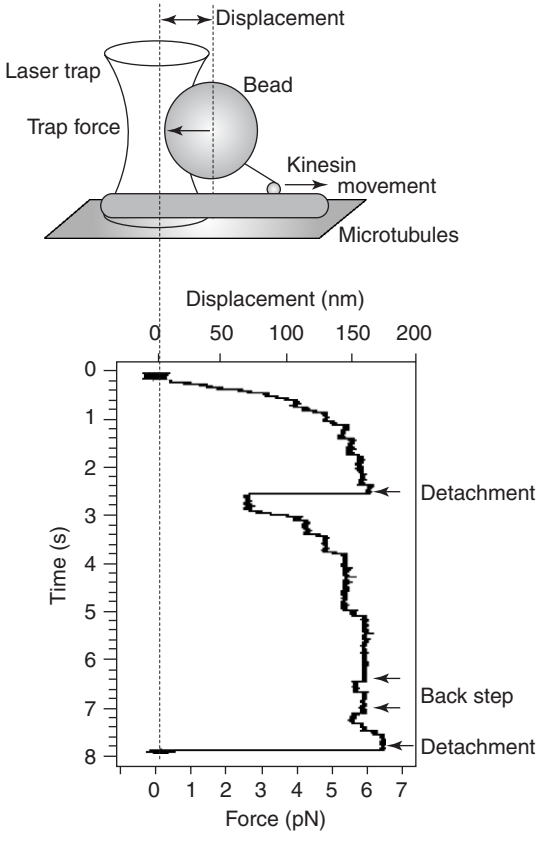


Figure 4 Measuring the stepwise movement of a single kinesin molecule: (top) a single kinesin molecule attached to a bead is trapped by a laser. The movement of single kinesin molecules along microtubules is monitored by measuring the movement of a bead; (bottom) typical record of the displacement of kinesin.

The movement of motors is fueled by the energy released from the hydrolysis of ATP. Using single-molecule measurements within an accuracy of nanometer and milliseconds, the unitary steps of single motor molecules using the energy of single ATP molecules have been identified. Kinesin and unconventional myosin V and VI move in a stepwise manner along protein tracks for long distances without dissociating. This processivity is consistent with their function of transporting cellular organelles for long distances. In the case of kinesin, the step size during hydrolysis of single ATP molecules is 8 nm, corresponding to the periodicity of the tubulin heterodimers in a microtubule (**Figure 4**). The movement occurs against an increasing load until the movement ceases at the stall force of 7 ~ 8 pN. Thus, a kinesin walks on tubulin-heterodimers using its two heads in an alternating hand-over-hand fashion.

Myosin moves along actin filaments. In the case of myosin V and VI, the processive movement has been observed with the step size of 35 nm, which corresponds to the pitch of the two-stranded helical filament of actin. Muscle myosin, in contrast to kinesin and some unconventional myosin, dissociates every ATP hydrolysis cycle. Single myosin molecules have been found to generate mean displacements of 5–25 nm at nearly zero load and mean forces of 3–5 pN at high loads (**Figure 5a**). It was difficult to determine the individual step size of muscle myosin because the starting position of the displacements cannot be determined due to the effects of Brownian movement of beads attached to actin filaments. Instead of the

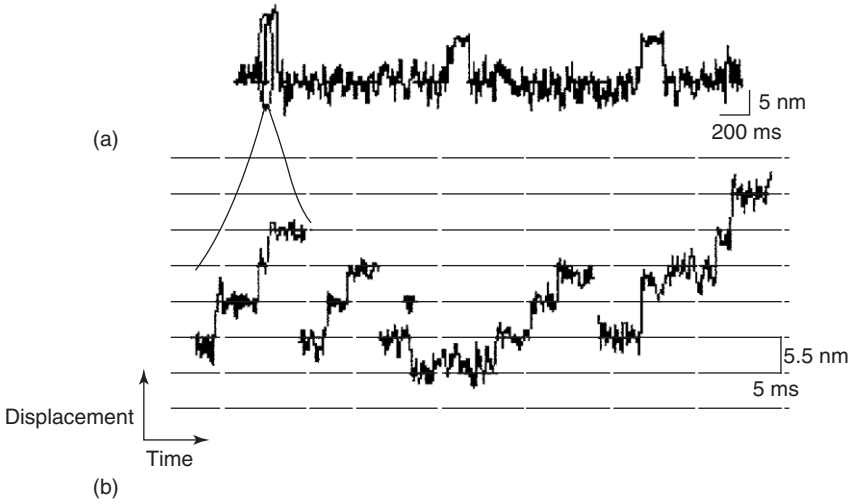


Figure 5 Stepwise movement of myosin measured using a fine scanning probe microscope. Single myosin molecules were attached to a tip of a microneedle and brought into contact with actin filaments. The trace is plotted in a different timescale. Part (b) is an expansion of (a).

manipulating actin filaments, the manipulation of single myosin molecules with a very fine scanning probe has provided a measurement system for the displacement and force generation processes with a high spatial and temporal resolution. With this measurement system, it was found that the displacements were coupled to the ATP hydrolysis and each displacement contained several distinct steps (**Figure 5b**). Each step was 5.5 nm in size, and corresponded to the distance between adjacent monomers in an actin filament. The steps took place stochastically and some of them ($<10\%$ of total steps) were in the backward direction. These results cannot be explained by the widely accepted model of structural change in which the displacement results from structural changes in the myosin coupled to the ATP hydrolysis reaction. Rather it has been suggested that a myosin head can walk along the actin monomers arranged in a filament utilizing Brownian motion. The mechanisms underlying the movement of myosin is still controversial and various measurements have been attempted from the viewpoints of the structure–function relationship.

Monitoring the Chemical Reactions

The functions in the biosystems take place in series of events. In single-molecule measurements, the series of events can be detected without interruption. In ensemble measurements, individual steps are measured separately and the entire processes reconstructed using this information. The kinetic parameters are determined by synchronizing the reactions of constituent molecules using stop flow, flash photolysis, or other transient kinetic methods. In single-molecule measurements, the reaction time varies from event to event reflecting the stochastic nature of the reactions. The rate constants, which characterize the reaction, are determined using the data from individual events.

The data obtained in the single-molecule measurements must be consistent with the data in the ensemble measurements. The data of single-molecule measurements contain more information than those of ensemble measurements. Data points from individual single molecules align with time. The time correlation data are discussed in relation to the memory effects of biomolecules. In flavoenzyme, which catalyzes the oxidation of cholesterol by oxygen, the enzymatic turnover is dependent on previous turnovers.

The series of events are described using several different parameters obtained from different measurements. It is unlikely that all the steps can be detected using single parameters obtained experimentally. In the ATP hydrolysis reaction, the fluorescence of ATP senses only the binding and dissociation events but not the hydrolysis steps or any other of the chemical steps. The mechanical measurement detects the displacement, association, and dissociation of the motor molecules but not any of the chemical steps. The chemical and mechanical parameters, which are separately measured, are however related to each other. Any change in the chemical reactions will influence the time record of the mechanical reactions. Changes in the ATP concentration change the mechanical trace. Time records of single parameters are influenced by the changes in other parameters that cannot be directly detected. Thus, the coupling between two parameters can be studied.

Direct Determination of Coupling Between Two Different Events

A more direct way of relating different events is to measure several parameters simultaneously. In single-molecule measurements, the individual measurements, in which the kinetic steps have not been averaged, can be directly compared. In the case of the molecular motors, the coupling between the ATP hydrolysis reaction as an input of the energy and mechanical work as the output has been determined (**Figure 6**). The ATP hydrolysis cycle is measured by measuring the binding and dissociation cycle of fluorescent ATP and the displacement of myosin using a laser trap. It has been suggested that the rising phase of the displacement record is not always coupled to the chemical reactions but instead is delayed. The energy released from the hydrolysis of ATP is then stored and used afterwards for the generation of the force.

With regard to the relationship between two events, the structure–function relationship is a fundamental question that needs to be addressed. In many proteins, the function can be explained by structural changes in the proteins. These changes are usually based on the atomic structures obtained from crystallographic studies. The relationship between structure and function has been studied by measuring the function using mutated proteins. Now that the structural changes of single protein molecules can be detected, the structure–function relationship will be experimentally studied in a more direct manner.

From Single Molecules to a Biosystem

Single-molecule measurements have been extended to monitoring single molecules in cells. Using TIRF microscopy it is possible to image both the basal and apical surfaces of cells. Incident laser light is totally internally reflected at the apical surface between the cytoplasm and the cell culture medium, where the diffractive indices are different. Other microscopy such as confocal microscopy has been used especially to image locations other than at the cell surfaces.

Cells perform their function in response to the external stimuli. At early processing stages, that is, from stimulation to function, the ligands bind to their receptors on the surface of cells and the signals are transmitted through a cascade of signaling proteins. In this process, many different kinds of molecules participate and chain reactions take place in a spatially organized fashion. Single-molecule measurements have been a powerful technique to explore early signal transduction systems. *Dictyostelium* amoeba exhibits chemotaxis along a gradient of adenosine 3′,5′-monophosphate (cAMP). The binding of fluorescent cAMP

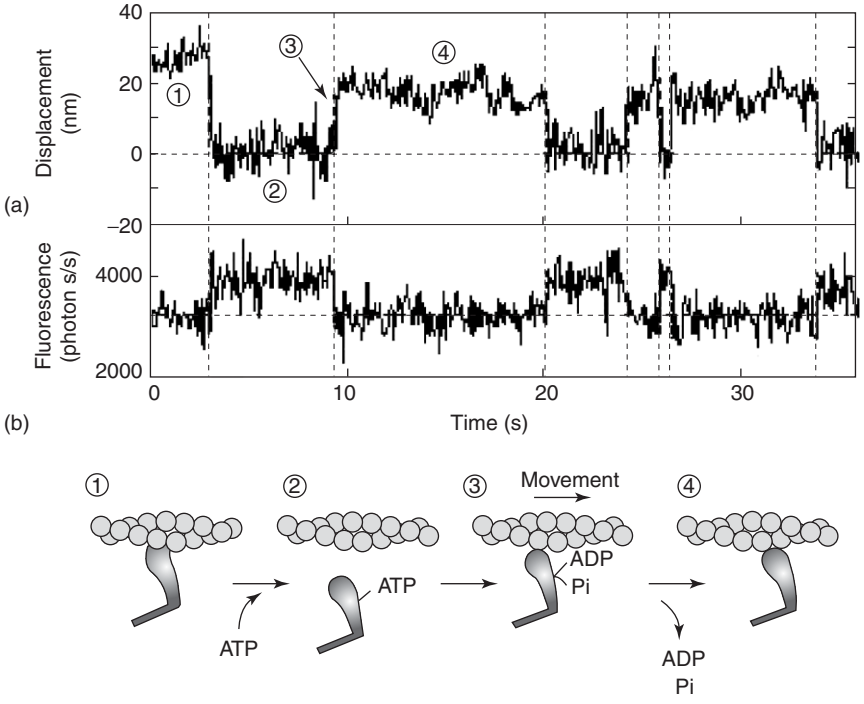


Figure 6 Direct determination of mechanochemical coupling of myosin. (a) The displacement of a single myosin molecule along an actin filament was measured using a laser trap method. (b) The fluorescence of Cy3-ATP was monitored at the same time. A schematic drawing of the interaction of myosin and actin is shown at the bottom.

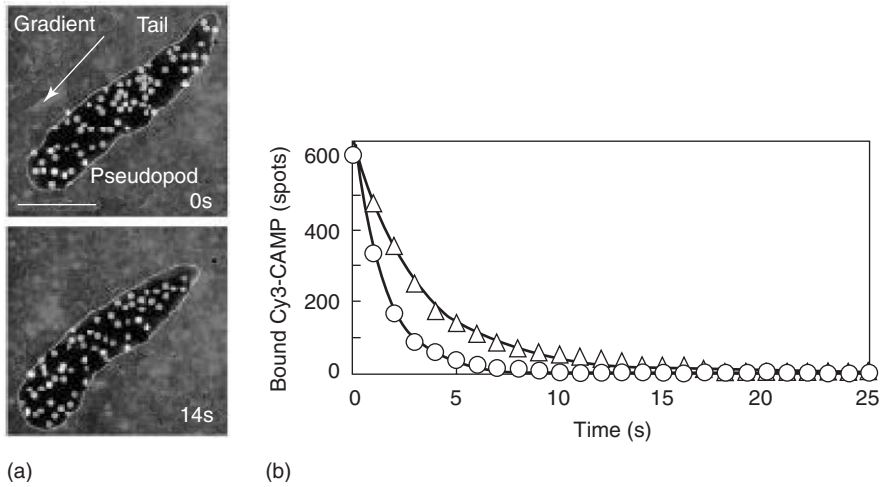


Figure 7 Single-molecule imaging of the binding of cAMP to the living cells. (a) Image of the binding to moving dictyostelium amoeba. (b) The histogram of the duration time of the cAMP binding at the pseudopod and tail halves of the moving cell.

to its receptor has been visualized at single-molecule level (**Figure 7**). It was found from observations of the dissociation rate constants that a receptor in the anterior half of the moving cells became activated when there was a difference in the cAMP concentration gradient compared to the posterior half.

In this process, in addition to the binding of the ligands, the activation, structure, and interaction of protein processes can be visualized using single-molecule imaging. Dimerization was visualized by measuring the fluorescence intensity and phosphorylation was visualized using fluorescent monoclonal antibody.

Concluding Remarks

Single-molecule measurements have allowed the measurements of the properties of biomolecules hidden in averaged measurements involving over large numbers of molecules. Dynamic, kinetic, and fluctuating properties of biomolecules, which are important in elucidating the mechanism of their function have been revealed. Using fluorescence microscopy the behavior of biomolecules at work can be measured. Both temporal and spatial information is included in the fluorescence microscopy measurements. It is possible to detect the behaviors of biomolecules in complicated systems, allowing the function of biomolecules to be imaged in living cells.

Using single-molecule measurements it is possible to monitor series of events in real time. To interpret the kinetic data, it is necessary to analyze large numbers of molecules from individual events. These complicated procedures are necessary because the energy used for the function of molecular machines is similar to the thermal energy. The energy released from the ATP hydrolysis is approximately 20 times the thermal energy. Biomolecules of nanometer size are largely disturbed by thermal energy. Thus, it is highly likely that the single-molecule measurements of biomolecules are influenced by thermal disturbances. It is an intriguing question as to how biomolecules work efficiently under situations of thermal energy disturbances. The mechanism that the biomolecules harness the thermal energy rather than being disturbed has been suggested from single-molecule measurements. These properties most likely give the systems flexibility in their function.

To measure the properties of single molecules, probes are attached. The effects of the probes on the activity can be examined by comparing the function with molecules that have no probe attached. However, it is difficult to ascertain what the effects of the probes on the properties of the biomolecules are as revealed by single-molecule measurements. *In situ*, the molecular motors move without the aid of beads or microneedles and proteins change their structure in the absence of fluorescent probes. Different probes may disturb biomolecules in different ways. Biomolecules may behave with different restrictions in different systems. Thus, measurements with different probes and in different systems may be helpful in understanding the effects of the probes.

In single-molecule measurements it is possible to monitor states that occur less frequently compared with other states. In ensemble measurements, the majority states contribute largely and the minority states are neglected. This is not to say, however, that the states that occur more frequently are most important and the states that occur less frequently are not important in the biological processes. The single-molecule measurement in cells show that the binding of a small number of ligands is sufficient to stimulate

the cells. It is possible that the minority state plays an important role even though only a small population of molecules is involved. Thus, single-molecule measurements will play an important role in the future of biosciences.

See also: Biomolecules, Scanning Probe Microscopy of; Fluorescent Biomolecules; Fluorescent Proteins.

PACS: **82.37.Np; 82.37.Rs; 82.37.Vb; 87, 87.64. – t**

Further Reading

Ishii Y and Yanagida T (2000) Single molecule detection in life science. *Single Molecule* 1: 5–16.

Marriott G and Parker I (eds.) (2003) *Biophotonics. Part B: Methods in Enzymology*, vol. 361. San Diego: Academic Press.

Schliwa M (ed.) (2003) *Molecular Motors*. Weinheim: Wiley-VCH.

Zander Ch, Enderlein J, and Keller RA (eds.) (2002) *Single Molecule Detection in Solution. Methods and Applications*. Berlin: Wiley-VCH.

Small Particles and Clusters, Optical Properties of

T Arai and K Matsuishi, University of Tsukuba, Tsukuba, Japan

Introduction

Mesoscopic particles of nanometer size are considered to be the intermediate phase which bridges between crystals and molecules and/or atoms. Appearance of various kinds of new properties – such as quantum size effects of electronic and phonon states, variations in shape and crystal structure, an increase of optical nonlinearity, a change in the selection rule of the optical transition, and an increase of surface effects – is expected in the mesoscopic particles.

This article focuses on the optical properties of small-size particles of semiconductors, ionic crystals, and metals.

Semiconductors and Ionic Crystals

Quantum Size Effects of Electronic States

The lowest excitation energy state of an electron in an ionic crystal and a semiconductor is the exciton state. The exciton can be treated as a "particle." Hence, in a first approximation the exciton will be considered as a particle in a spherical-shaped potential well. The wave function of the infinitesimally small particle with mass M in a spherical-shaped potential well with radius R is expressed by a spherical Bessel function, and the eigenenergies measured from the bottom of the potential well are given as follows:

$$\Delta E(R) = \frac{\hbar^2}{2M}\left(\frac{\phi}{R}\right)^2 \qquad [1a]$$

where $\phi/\pi = 1,\ 1.43,\ 1.83, \ldots$ give $1S$-state, $1P$-state, $1D$-state, …, respectively. Thus, the translational energies become discontinuous and increase. When the effective Bohr radius of the exciton, a_B^*, is much smaller than the radius of the particle, the exciton retains its size and can be treated as a particle. This case is called the exciton confinement or the weak confinement regime and eqn [1a] can be used with a small modification. The energy difference between the bulk exciton and the confined exciton is given by the following equation:

$$\Delta E(R) = \frac{\hbar^2}{2(m_e^* + m_h^*)} \times \frac{\phi^2}{[R - \eta(\sigma)a_B^*]^2} \qquad [1b]$$

where $\eta(\sigma)$ is a function of the ratio of the effective mass of the hole, m_h^*, to that of the electron, m_e^*. The term $\eta(\sigma)a_B^*$ expresses a dead-layer effect. That is, the exciton cannot exist within a distance less than $a_B^*/2$ from the wall of the potential well.

On the contrary, when the potential well radius R is comparable to or smaller than a_B^*, the exciton size diminishes and its internal motion is modified strongly. The exciton energy is also quantized. Therefore, the mass in eqn [1b], $m_e^* + m_h^*$, should be changed to the reduced mass of the exciton, μ, given by

$$\frac{1}{\mu} = \frac{1}{m_e^*} + \frac{1}{m_h^*}$$

This case is called the individual-confinement or the strong-confinement regime.

When the Coulomb energy in the exciton becomes equal to the energy of the relative motion of the exciton, the confinement state changes from the "exciton confinement" to the "individual confinement." The transition between the two states occurs when

$$\frac{\hbar^2\pi^2}{2\mu a_B^{*2}} \geqslant \frac{e^2}{\varepsilon R}$$

$$\frac{R}{a_B^*} = \frac{\pi^2}{2} \approx 4.9$$

This transition has been observed at $R/a_B^* \approx 4$ in a CdS microcrystal.

If a charge is placed in a small medium, a surface polarization appears, which creates a Lorentz field in the medium. Consequently, the Coulomb force and the self-energy of the exciton in the small medium vary from that of the free exciton. The Hamiltonian of an electron–hole pair which includes the effect of the Lorentz field is shown in the effective-mass approximation as follows:

$$H = -\frac{\hbar^2\nabla_e^2}{2m_e^*} - \frac{\hbar^2\nabla_h^2}{2m_h^*} - \frac{e^2}{\varepsilon_1|r_e - r_h|} + V_e(r_e) + V_h(r_h) + \frac{e^2}{2R}\sum_{n=0}^{\infty}\alpha_n\left[\left(\frac{r_e}{R}\right)^{2n} + \left(\frac{r_h}{R}\right)^{2n}\right] - \frac{e^2}{R}\sum_{n=0}^{\infty}\alpha_n\left[\left(\frac{r_e r_h}{R^2}\right)^{n} P_n(\cos\theta_{eh})\right] \quad [2]$$

where

$$\alpha_n = \frac{(n+1)(\varepsilon-1)}{\varepsilon_1(n\varepsilon+n+1)}, \quad \varepsilon = \frac{\varepsilon_1}{\varepsilon_2}$$

ε_1 and ε_2 are the dielectric constants of the small particle and the surrounding material, respectively. r_e and r_h are the coordinates of the electron and the hole. $P_n(\cos\theta_{eh})$ is the Legendre polynomial, and θ_{eh} is the angle formed by the two lines connecting the center of the particle to the electron and to the hole. The fourth and fifth terms in eqn [2] are the confinement potentials for the electron and the hole, respectively. The sixth term is the Lorentz field correction, and the seventh term is the self-energy. The calculated result of the lowest energy of the confined exciton with the wave function, which is made by the product of the 1*s*-state electron and the 1*s*-state hole wave functions, is given as follows:

$$\Delta E = \frac{\hbar^2\pi^2}{2R^2}\left[\frac{1}{m_e^*} + \frac{1}{m_h^*}\right] - \frac{1.786e^2}{\varepsilon_1 R} + \frac{e^2}{R}\sum_{n=1}^{\infty}\alpha_n\left(\frac{S}{R}\right)^{2n} \quad [3]$$

The first term represents a change in the kinetic energy, the second term the energy due to the shielding Coulomb interaction, and the third term the polarization term. The ratio of the energy shift due to the polarization term to the total energy shift is ~ 0.07 for CdS. In addition, the value of the dielectric constant decreases with a decrease in the particle size, but this effect is less than 1% in general. Equation [3] can be used in the weak-confinement regime, if the total mass of the exciton is employed instead of the reduced mass.

Figures 1 and **2** show the size dependence of excitonic transition energies for CuCl and CdSe particles.

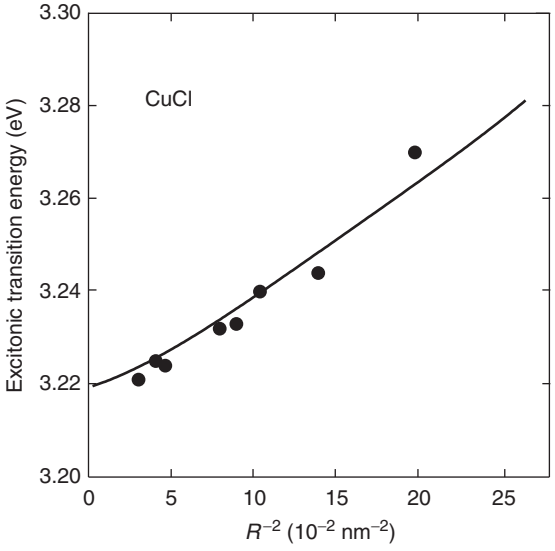


Figure 1 Size dependence of exciton energy for CuCl particles. The solid line denotes the result by the calculation assuming a dead zone with the thickness $a_B^*/2$ on the particle surface. The experimental data (solid circles) are from Itoh *et al.* (1988) *Physica Status Solidi* (b) 146:531.

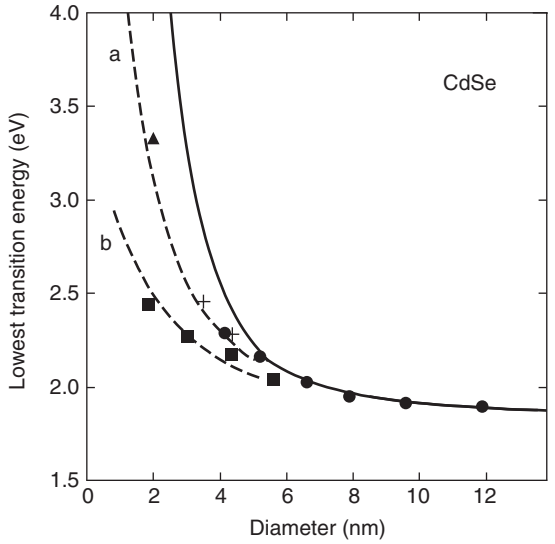


Figure 2 Size dependence of the lowest excitation energy (HOMO–LUMO transition energy) for CdSe particles. • from Arai T, *et al.* (1996) *Journal of Luminescence* 70:281; ■ from Nogami N, *et al.* (1991) *Journal of Non-Crystalline Solids* 135:182; + from Alivisatos AP, *et al.* (1988) *Journal of Chemical Physics* 89: 4001; ▲ from Chestnoy N, *et al.* (1986) *Journal of Chemical Physics* 85:2237. The solid line and the dashed lines denote the calculated results by an effective-mass approximation and by tight-binding approximations, respectively. The dashed line (a) is from Lippens PE, *et al.* (1990) *Physical Review B* 41: 6079 and (b) from Ramaniah LM, *et al.* (1993) *Physical Review B* 47: 7132.

The solid line in **Figure 1** shows the calculated result assuming a dead zone with thickness $a_B^*/2$ on the particle surface. The effective Bohr radius of the exciton in CuCl is 0.7 nm. Hence, the exciton confinement

regime is valid in the entire region of measurements. The radiation due to the dissociation of the biexciton (the exciton molecule) has been observed under intensive excitation. The binding energy of the biexciton, E_b, given by $2E_{ex} - E_{mol}$, also has size dependence, but its dependence is small compared to that of the exciton energy.

The exciton radius of CdSe is ~ 5.6 nm. The individual-confinement regime is valid in the size region, shown in **Figure 2**. The agreement between the experimental data and the effective-mass approximation calculation (solid line) is good in the region of diameters larger than ~ 6 nm. This implies that the disagreement in the small-size region is caused by the failure of the effective-mass approximation. In this region, the experimental data, however, agree with the tight-binding approximation calculations (dashed lines) (**Figure 2**).

In the above treatment, the potential well has a finite depth. Therefore, the wave function of the exciton tunnels out from the potential well. This effect becomes very important, when the small particles are arranged in a dense state.

Oscillator Strength

The coherent state of the exciton spreads over the entire region of a very small size particle; therefore, the oscillator strength increases with increasing size in the very small size region. This phenomenon is called the "giant oscillator strength" (GOS) effect. The oscillator strength of the exciton, f_{ex}, in the small particle with radius R is given by

$$f_{ex} = \frac{8}{\pi}\frac{R^3}{\nu} f_0 \propto R^3$$

where ν is the volume of the unit cell and f_0 is the oscillator strength in the bulk crystal and is given by

$$f_0 = \frac{2m^*\omega}{\hbar}|\mu_{eg}|^2 \frac{1}{\pi a_B^{*3}}$$

Here, μ_{eg} is the electric dipole moment in the transition process between e and g states.

The lifetime of an exciton is inversely proportional to the third power of the particle radius due to the GOS effect:

$$\tau = \frac{3m^*c^3}{16e^3 n\omega^2 f_0}\left(\frac{\nu}{R^3}\right) \propto \frac{1}{R^3}$$

The bandwidth of the photoluminescence spectrum, Γ_h, measured by a selective-excitation method is inversely proportional to the lifetime. Thus, Γ_h increases with an increase in the particle size. The quantum yield of photoluminescence increases with an increase in the size due to the GOS effect. For example, the quantum yield in the case of very small size for CuCl particles has been observed to be $\sim 0.1\%$. The photoluminescence spectrum shifts to the higher-energy side compared with that of the bulk. The Stokes shift has not been observed, except at very low temperatures. When the radius of the particle becomes larger than the coherent length of the exciton, the GOS effect saturates. The GOS effect is important, especially in nonlinear optical effects.

Size Effect of Phonon

The binding force between atoms is mainly a short-range force in general. Therefore, the quantum size effect on the phonon energy does not appear apparently. However, in optical transitions a relaxation of the selection rule on the wave vector (the momentum) occurs. This effect has been observed in Raman scattering and photoluminescence spectra.

The size dependence of Raman intensity $I(\omega)$ owing to optical phonon scattering is expressed essentially as follows:

$$I(\omega) \approx \int d^3q \frac{|C(0,\boldsymbol{q})|^2}{\{\omega - \omega(\boldsymbol{q})\}^2 + (\Gamma_0/2)^2}$$

Here $\omega(\boldsymbol{q})$ is the optical phonon dispersion, $C(0,\boldsymbol{q})$ the Fourier coefficient of the phonon confinement function, and Γ_0 the natural bandwidth of the Raman line. If the Gaussian function, $w(r,L) = \exp(-8\pi^2 r^2/L^2)$, is used as an attenuation function, the Fourier coefficient of the phonon confinement function is given as follows:

$$|C(0,\boldsymbol{q})|^2 = \exp\left(-\frac{|\boldsymbol{q}|^2 L^2}{16\pi^2}\right)$$

Here L is the diameter of the particle, that is $L = 2R$. The range of the contributing wave vector $\boldsymbol{q}$ to the optical excitation is inversely proportional approximately to R^2. The Raman shift decreases and the bandwidth of the spectrum becomes broader with a decrease in the size, for example, in CdS and CdSe particles. The Raman scattering owing to the spheroidal and torsional modes in a spherical particle is also expected. These modes have been observed in a very low frequency region.

Exciton–Phonon Interaction

The exciton–phonon interaction has been investigated extensively in CuCl particles. There are light- and heavy-hole bands in a CuCl crystal, and the Z_3 exciton state splits into light- and heavy-hole excitons. The coherence of both the excitons is broken by

the interband scattering between excitons through the absorption and emission of LA phonon at low temperatures, while they are broken by LO phonon scattering at high temperatures. The temperature dependence of the homogeneous bandwidth of the exciton spectrum in CuCl particles has been interpreted by the above mechanism with a consideration of the localization properties of LA and LO phonons.

Optical Nonlinearity

Under intensive laser irradiation, the imaginary part of the third-order dielectric constant, Im $\chi^{(3)}$, is proportional to the square of the oscillator strength, f, and inversely proportional to the square of the homogeneous bandwidth, Γ_h, of the absorption spectrum. The absorption intensity in the selective excitation is given by $\alpha\Gamma_h$. Here α is the absorption coefficient. $\alpha\Gamma_h$ is proportional to fN (N is the number density of particles). Therefore, the third-order electronic susceptibility, $|\chi^{(3)}|$, is given as follows:

$$|\chi^{(3)}| \propto \frac{f^2 N T_1}{\Gamma_h^2}$$

Here T_1 is the longitudinal relaxation time. In very small size particles, $|\chi^{(3)}|$ at the resonance energy for the lowest excitation state becomes larger with an increase in the size, up to a critical size. According to the theory of the GOS effect on the confined exciton, the oscillator strength of the exciton increases as R^3, up to the critical size. The oscillator strength saturates at the critical size. Then, the oscillator strength decreases with a further increase in the size. The critical size decreases with an increase in the temperature. The reasons are the following: the energy separation between the excitonic states decreases with increasing particle size. Therefore, the thermal population of the exciton in the higher excitonic states increases and the coherent length of the lowest excitonic state decreases due to an increase in the scattering between the excitonic states. Contributions from the higher excitonic states overcome the resonance increment of the oscillator strength in the lowest state. Some theories have also suggested that the enhancement of $|\chi^{(3)}|$ is larger in materials having a relatively smaller Bohr radius. The maximum values reported for $|\chi^{(3)}|$ are 2×10^{-7}, 3×10^{-8}, and 1×10^{-6} esu for CdTe-, CdS-, and CuCl-doped oxide glasses, respectively.

Taking into account the two-energy-level system, $|\chi^{(3)}|/\alpha T_1$ is proportional to f, and inversely proportional to Γ_h, as described previously. Therefore, the oscillator strength can be estimated from the measured values of $|\chi^{(3)}|/\alpha T_1$ and the homogeneous bandwidth. The estimated value of the oscillator strength is ~ 5.5 for the 4.5 nm-size CuCl particle at the wavelength where $|\chi^{(3)}|/\alpha T_1$ is maximum. Compared to the value of the oscillator strength in bulk, 5.85×10^{-3}, the enhancement factor is $\sim 10^3$.

Many other nonlinear optical phenomena, such as the superradiation, the formation of the molecular exciton (the electron–hole droplet), and the saturation of absorption (the hole burning), are enhanced strongly due to the GOS effect. For example, the hole burning in the inhomogeneous broad band has been observed easily in CuCl particles with an average size of ~ 3.6 nm at 77 K. Not only is the hole burning formed but also the absorption at the lower-energy side of the exciting energy decreases and the absorption at the higher-energy side of the exciting energy increases. These phenomena are caused by the multiselective excitations which occur due to the size distribution of particles and the production of bi- and tri-excitons.

Photoluminescence at an energy lower than that of the usual exciton luminescence has been observed in nanoparticles of some ionic crystals under irradiation of an intense laser with the resonance energy of the exciton. When this luminescence occurs, the absorption coefficient at the exciting energy is only about half of that obtained in the usual light irradiation. Therefore, this luminescence originates from the dissociation of the molecular exciton created by the two-photon absorption. The luminescence occurs in smaller excitation intensity compared to that in the bulk. It may be due to the enhancement of mutual interaction among excitons owing to the occupancy of the close position. According to the results of the research on the dynamical properties of the relaxation of a molecular exciton, the optical density, $I_S T_1$, is proportional to $1/R^3$, where I_S and T_1 are the saturation intensity of absorption and the longitudinal relaxation time of the molecular exciton, respectively.

In the highly excited Frenkel exciton system, in a small-size particle, a new elementary excitation (called an excitonic n-string) appears, and a superradiation from the n-string is expected. In the highly excited Wannier exciton system, the electron–hole droplet has been observed in Si and Ge, and the excitonic molecule has also been observed in III–V and III–VI compounds. However, the exchange interaction does not support the bound states of more than two Wannier excitons (exciton molecule). In contrast, a bound state of several Frenkel excitons (i.e., excitonic n-string) is possible in a linear chain of molecules, when the exciton accompanies the static dipole moment larger than the transition dipole moment along the chain.

The superradiation from a Wannier exciton in microcrystallites has been shown theoretically and

experimentally. New sharp superradiance with a strong peak intensity from the Frenkel exciton system is expected in intense excitation because of the high density of oscillators. According to the calculated results of the spectrum and pulse width in time, the peak intensity, I, from the n-string is given by

$$I = \tfrac{1}{2}N_s\left(\tfrac{1}{2}N_s + 1\right)\Gamma$$

where Γ is a spontaneous emission rate of a single exciton and N_s is the number of excitons in the n-string, that is, $N_s = n$. The pulse width is inversely proportional to N_s. The intensity I increases as $N_s^{1.5}$. This should be compared with the N_s^2 dependence for Diche's superradiation from atomic and/or molecular gas systems. This may be because the cooperation number is kept to the maximum for the two-level atoms, while it decreases in the emission process due to the effect of excitation transfer.

Crystals which have a zinc blende structure, such as CuCl and CuBr, have two types of excitons, Z_3 and Z_{12}, due to the splitting of the valence band. The Z_{12} exciton consists of the multicomponent exciton states due to light and heavy holes and a k-linear term in the Γ_8 valence band. The energies of both the excitons in a small particle are higher than that in the bulk crystal. In addition, the Z_{12} exciton level splits off into two levels due to a confinement effect because of the multicomponent exciton. Spectral analyses on the degenerated four-wave mixing signals have shown that the correlation curve between the luminescence signal and the delay time consists of an oscillatory structure superimposed onto the exponential decay. The oscillation period corresponds to the energy separation of the Z_{12} exciton levels in the nanoparticle. Thus, this oscillation has been interpreted as the quantum beat between the separated Z_{12} exciton levels.

Metal Particles

Absorption of Surface Plasmon

The so-called stained glass consists of glass-dispersed fine particles of metal. The optical properties of isolated fine particles in a transparent matrix were first treated by Mie. There are many different modes in the particle. Higher modes are localized on the surface, and the lowest mode, the so-called Frölich mode, polarizes uniformly in the particle. The Frölich mode is important in the small particle – it gives rise to surface plasmons which cause the colors present in a stained glass. The Frölich mode is treated in the following.

The dielectric constant of metal, ε_m, differs from that of the matrix, ε_d. Therefore, the local field in the metal particle, E_l, should differ from the field of incident radiation, E, and is given by

$$E_l = \frac{3\varepsilon_d}{\varepsilon_m + 2\varepsilon_d}E = f_l E$$

The constant f_l is termed as the factor of the local field. The averaged electric field in the material, E_{av}, and the averaged polarization of the material, P_{av}, can be defined as

$$E_{av} = (1-f)E + fE_l$$
$$P_{av} = [(1-f)(\varepsilon_d - 1)E + f(\varepsilon_m - 1)E_l]/4\pi$$

Here f is the volume fraction of the metal (i.e., the filling factor of the metal). The average dielectric constant of the material with dispersed metal particles, ε_{av}, calculated by the dipole–dipole interaction approximation is given as follows:

$$\varepsilon_{av}(\omega) = \varepsilon_d(\omega)\left\{1 + \frac{3f[\varepsilon_m(\omega) - \varepsilon_d(\omega)]}{\varepsilon_m(\omega)(1-f) + \varepsilon_d(\omega)(2+f)}\right\}$$

The absorption coefficient α is given by

$$\alpha(\omega) = \frac{\omega}{c}\frac{\mathrm{Im}\,\varepsilon_{av}(\omega)}{\mathrm{Re}\,\sqrt{\varepsilon_{av}(\omega)}}$$

The peak of absorption is given by the condition that $\mathrm{Re}\,\sqrt{\varepsilon_{av}(\omega)} = 0$. Thus, the peak of absorption is given by

$$\mathrm{Re}\,\varepsilon_m(\omega_s) = -\varepsilon_d(\omega_s)\frac{f+2}{1-f}$$

Here ω_s is the angular frequency of the surface plasmon. The above equation can be written as

$$\mathrm{Re}\,\varepsilon_m(\omega_s) = -2\varepsilon_d(\omega_s)$$

in the small-f region. The absorption spectrum in the small-f region is given as follows:

$$\alpha = \frac{18\pi f\varepsilon_d^{3/2}}{\lambda}\frac{\mathrm{Im}\,\varepsilon_m}{(\mathrm{Re}\,\varepsilon_m + 2\varepsilon_d)^2 + (\mathrm{Im}\,\varepsilon_m)^2}$$

Here λ is the wavelength of light. The results calculated by the above equation show that the absorption peak shifts to the shorter-wavelength side and the width of the absorption band becomes narrower with a decrease in the particle size. However, the experimental results do not coincide with the calculated results, because the mean free path length of the electron in a metal is generally several tens of nanometers and is limited by the particle size in small particles. If one calculates the dielectric constant including this effect, the peak value of the absorption band becomes smaller and the bandwidth of the absorption becomes wider with a decrease in the particle size.

Third-Order Electric Susceptibility

The third-order susceptibility, $\chi^{(3)}$, in the composite material which has dispersed metal nanoparticles in a transparent material is given by

$$\chi^{(3)} = f \cdot f_l^2 |f_l|^2 \chi_m^{(3)} \quad [4]$$

Here f and f_l are the volume fraction of the metal and the factor of the local field, respectively, as denoted previously. $\chi_m^{(3)}$ is the third-order electric susceptibility of the metal itself. As mentioned already, the factor of the local field increases at the surface plasma frequency. When the combination of the matrix and the particle satisfies the resonance condition, $\mathrm{Re}\,\varepsilon_m = -2\varepsilon_d$, the factor of the local field increases. As seen in eqn [4], $\chi^{(3)}$ increases as f_l^4. The susceptibility depends on the size of the metal particle. If one can choose the size that satisfies the resonance condition, $\chi^{(3)}$ increases. The enhancements of $\chi^{(3)}$ in Cu and Ag composite SiO_2 glasses are approximately 5 times and 10^4 times in the best condition, respectively.

The time dependence of a nonlinear response was measured by a pump–probe method under several $\mathrm{mJ\,cm^{-2}}$ energy density of irradiation. The bleaching of absorption at the pumping wavelength and the enhancement of absorption on either side of the pumping wavelength were observed. The shapes of the absorption bands before and after irradiating the pumping light were Lorentzian. However, the bandwidth was observed to be a few times wider on pumping. Therefore, the decay character of the absorption band gives information on the scattering mechanism of hot electrons. The decay curve is composed of fast and slow components. The decay times of these components are approximately a few picoseconds and a few hundreds of picoseconds, respectively. The fast component is due to the electron–phonon scattering and the slow component may be due to the electron–matrix scattering. The decay time of the fast component becomes shorter with a decrease in the energy density of irradiation. A similar decay character has also been observed under high-intensity radiation condition for particles of II–VI compounds such as CdS, CdSe, and CdTe.

See also: Semiconductor and Metallic Clusters, Electronic Properties of; Semiconductors, General Properties; Semiconductors, Optical and Infrared Properties of.

PACS: 36.40.Vz

Further Reading

Arai T, Mihama K, Yamamoto K, and Sugano S (eds.) (1999) *Mesoscopic Materials and Clusters – Their Physical and Chemical Properties*. Tokyo: Kodansha (Springer).

Brus EL, Efros LAl, and Itoh T (eds.) (1996) Spectroscopy of isolated and assembled semiconductor nanocrystals. *Journal of Lumiescence* 70: 1–484.

Gaponenko VS (1998) *Optical Properties of Semiconductor Nanocrystals*. Cambridge: Cambridge University Press.

Knox CR (1963) *Theory of Exciton*. London: Academic Press.

Masumoto Y and Takagahara T (eds.) (2002) *Semiconductor Quantum Dots – Physics, Spectroscopy and Applications*. Berlin: Springer.

Sugano S and Koizumi H (1998) Microcluster Physics, 2nd edn, Springer Series in Materials Science, vol. 20. Berlin: Springer.

Woggon U (1997) *Optical Properties of Semiconductor Quantum Dots*. Berlin: Springer.

Yoffe DA (1993) Low-dimensional systems: quantum size effects and electronic properties of semiconductor microcrystallites (zero-dimensional systems) and some quasi-two-dimensional systems. *Advances in Physics* 42: 173–266.

Nomenclature

a^*_B	effective Bohr radius of exciton
c	speed of light
$C(0, q)$	Fourier coefficient of phonon confinement function
E	energy, electric field
E_{av}	averaged electric field
E_b	binding energy of biexciton
E_{ex}	energy of exciton
E_l	local field of metal particle
E_{mol}	energy of exciton molecule
f	oscillator strength, volume fraction of metal (filling factor)
f_0	oscillator strength in bulk crystal
f_{ex}	oscillator strength of exciton
f_l	factor of local field
$\hbar$	$h/2\pi$ (h: Planck constant)
I	intensity (Raman, emission)
I_s	saturated intensity of absorption
m^*_e	effective mass of electron
m^*_h	effective mass of hole
M	mass
n	refractive index
N	number density of particles
N_s	number of excitons in excitonic n-string
P	polarization of material
P_{av}	averaged polarization of material
$P_n(x)$	Legendre polynomial
q	wave vector
r_e	coordinate of electron
r_h	coordinate of hole
R	radius
T_1	longitudinal relaxation time
v	volume of unit cell
α	absorption coefficient
Γ	spontaneous emission rate
Γ_h	homogeneous line width
ε	dielectric constant
ε_d	dielectric constant of matrix
ε_m	dielectric constant of metal
θ_{eh}	angle formed by the two lines connecting the center of particle to electron and hole

τ	lifetime	$\chi^{(3)}$	third-order dielectric constant (electronic susceptibility)
λ	wavelength of light	$\chi_m^{(3)}$	third-order dielectric constant of metal
μ	reduced mass of exciton	ω	angular frequency
μ_{eg}	electric dipole moment in optical transition between e and g states	ω_s	angular frequency of surface plasmon
σ	ratio of m_h^* to m_e^*		

Solid-State NMR Structural Studies of Proteins

C E Hughes and M Baldus, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

Introduction

Unlike macroscopic scattering methods such as X-ray diffraction, electron microscopy, or neutron diffraction, nuclear magnetic resonance (NMR) delivers structural information indirectly by probing the local electronic and magnetic environment of a nuclear magnetic moment (spin) or by investigating spin–spin interactions, both in the presence of a static magnetic field. The choice of NMR methods to detect such interactions can be strongly related to the mobility of the molecule of interest. In fact, molecular motion, or more generally speaking, the overall molecular correlation time τ_c, can be used to classify the type of NMR techniques for studying structure at the atomic level. Solution-state NMR studies in proteins are usually applied to systems with an overall correlation time in the nanosecond regime, while solid-state NMR, the area of research discussed in this article, traditionally relates to long or infinite correlation times (**Figure 1**).

In general, an increase in molecular correlation time may result from large molecular size, interactions with an anisotropic environment or from a decrease in temperature. In the context of this article, such phenomena may be related to large proteins, proteins interacting with membranes, or protein aggregates that are indicated in a growing number of protein folding disease states. In all cases, the molecular correlation time becomes longer than the inverse of the magnetic resonance frequency and all possible molecular orientations in the magnetic field are sampled in the NMR frequency spectrum. Unlike in solution, the spectral resolution and the overall sensitivity of solid-state NMR are furthermore influenced by the size of the nuclear spin interactions. Recent progress in NMR instrumentation, in particular the combination of ultrahigh magnetic fields (i.e., above 14 T) with rapid sample rotation (discussed in more detail below) has considerably improved the use of solid-state NMR in biophysical applications.

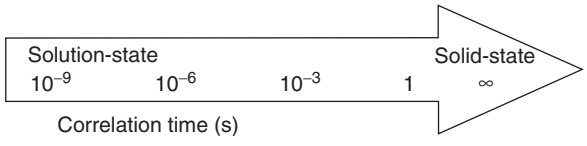


Figure 1 Typical ranges of molecular correlation times accessible to solution-state and solid-phase NMR.

So far, such conditions have not permitted a widespread application of high-resolution ^{1}H NMR spectroscopy, the most sensitive detection method in solution, but have greatly extended the possibilities for studying multiply or uniformly [^{13}C, ^{15}N]-isotope-labeled polypeptides with high sensitivity and adequate spectral resolution. In this article, an overview of the principal interactions that can be probed in NMR is provided, and a set of high-resolution solid-state NMR (HR-SSNMR) techniques that permit the detection of multiple structural parameters from a single protein sample or NMR data set, is introduced. In general, solid-state NMR spectra directly report on the conformational heterogeneity of the sample. Maximum spectral resolution is often obtained for proteins exhibiting a high degree of structural homogeneity. Hence, particular attention is placed on sample preparation and isotope labeling schemes. Finally, the possible range of applications of HR-SSNMR for the structural investigation of solid-phase (membrane) proteins is exemplified in three cases.

Interactions

The microscopic interactions that can be probed by NMR techniques are summarized in **Figure 2** and encompass the chemical shielding, the quadrupolar coupling, the dipolar coupling, and the scalar coupling. The chemical shielding describes the interaction between a nucleus and the applied magnetic field, as mediated by the environment of the nucleus. It is

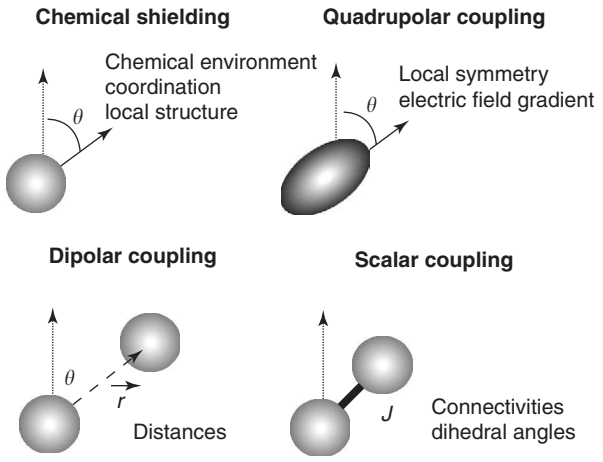


Figure 2 Summary of the four principal interactions probed by NMR. While chemical shielding and quadrupolar interactions involve only one-spin interactions, dipolar and scalar coupling refer to two-spin contacts.

observed as a shift in the resonant frequency, with isotropic and anisotropic contributions. In general, chemical shielding information can directly deliver information about the local electronic environment of a spin or the molecular coordination. In addition, for applications in polypeptides, the resulting resonance frequencies are not only diagnostic for each individual peptide residue but are also very sensitive to local backbone conformation.

The second interaction is the quadrupolar coupling, between the electric quadrupole moment of a nucleus and the electric field gradient at the nucleus. Only nuclei with a spin quantum number greater than 1/2 possess an electric quadrupole moment, making this interaction unimportant when dealing with spin-1/2 nuclei such as ^{1}H, ^{13}C, and ^{15}N. On the other hand, naturally abundant ^{14}N and deuteration with ^{2}H spins (both spin-1) may yield novel insights into the details of polypeptide structure and, in particular, dynamics. In addition to the interactions probing local chemical environments, two-spin interactions are invaluable tools for studying protein structure. For example, the dipolar coupling between pairs of nuclei, of either the same species (homonuclear) or different species (heteronuclear) directly reports on internuclear distances. This coupling is anisotropic in nature, having the same dependency upon spatial rotation as the anisotropic part of the chemical shielding. Finally, spin–spin interactions may be mediated by the through-bond coupling, also known as the scalar or *J*-coupling. This is an isotropic interaction between nuclei (homo- or heteronuclei) modulated by the bonding electrons. Both the dipolar and *J*-couplings can be exploited to transfer magnetization between nuclei, giving invaluable information about the proximity (dipolar) and bonding (*J*) of the nuclei.

In solutions, the rapid tumbling of molecules leads to an averaging out of the anisotropic interactions. In solids, these interactions remain and directly influence the detected NMR spectra. Because the dependence upon spatial rotation of both the chemical shift anisotropy (CSA) and the dipolar coupling is given by $3\cos^2\theta - 1$, where θ is the angle between the z principal axis of the interaction and the applied magnetic field, spinning a sample at an angle of 54.7° to the field axis results in a strong suppression of the anisotropic components of both interactions. The corresponding technique is known as magic-angle spinning (MAS). When MAS is combined with decoupling of the ^{1}H nuclei (whose dipolar couplings are larger and not generally removed by MAS) the line widths of many nuclei, including ^{13}C and ^{15}N, are narrowed toward values observed in solution-state NMR. By removing the dipolar and CSA interactions under ultrahigh magnetic fields (i.e., above 14 T), MAS makes the acquisition of high-resolution solid-state NMR spectra possible.

In addition, a variety of radio frequency (RF) NMR techniques have been developed to recover the structural information contained in the four interactions of **Figure 2** under MAS conditions. Some of these NMR schemes rely solely on particular settings of the MAS rate and are called rotor-driven recoupling techniques. Other techniques, known as radio-frequency-driven polarization transfer techniques, employ a series of RF pulses to overcome the effects of MAS or chemical shift perturbations and allow the detection of specific structural parameters in high resolution.

Sample Preparation

In general, for HR-SSNMR based studies on proteins, [^{13}C, ^{15}N]-isotope labeling is mandatory. Labeling patterns ranging from the incorporation of specifically [^{13}C, ^{15}N]-labeled amino acids to uniform [^{13}C, ^{15}N] substitution are often most economically generated during bacterial growth in minimal media. In addition, chemical synthesis, such as fluorenylmethoxycarbonyl (Fmoc) chemistry, or cell-free approaches can be employed to produce isotope-labeled polypeptides. In both cases, structural studies using solid-state NMR can be performed using a variety of macroscopic sample preparations. For globular proteins, NMR experiments can be conducted after lyophilization and subsequent rehydration, which often increases the structural homogeneity of the sample. For the same purpose, precipitants such as polymers (e.g., polyethylene glycol, PEG) or organic

solvents can be added that can lead to nano- or microcrystalline material. In the context of membrane proteins, structural studies have been conducted using frozen, detergent-solubilized specimens or liposomes in which the protein of interest has been reconstituted. Finally, fibrillized proteins have been extensively studied using solid-state NMR.

The highest sensitivity is usually achieved if the sample of interest is studied under MAS conditions using randomly oriented samples. In addition, macroscopically oriented samples, for example, mechanically oriented biopolymers or membrane-reconstituted peptides, aligned on thin polymer films, can be studied under fast MAS conditions. Furthermore, macroscopically oriented systems can also be investigated after orientation onto glass plates in static experiments or under MAS conditions.

Structural Analysis Using Solid-State NMR

Spectral Assignment

Of particular importance in the context of studying multiply or uniformly labeled proteins are polarization transfer methods employing scalar or dipolar two-spin interactions. Because of the isotropic character of the scalar coupling, polarization transfer mediated by through-bond interactions between two nuclei generally proceeds in an oscillatory manner, irrespective of the macroscopic orientation in the magnetic field (**Figure 3a**). In contrast, polarization transfer transmitted by through-space interactions under MAS conditions is dependent upon the orientation θ, of the dipolar vector in the static magnetic field. Integration over all possible molecular conformations can lead to a polarization transfer behavior, as depicted in **Figure 3b**, where positive and negative signal buildup corresponds to zero-quantum or double-quantum transfer, respectively.

An essential first step in NMR studies of multiply labeled proteins is the spectral assignment, based on the known amino acid sequence. Chemical shifts can be used to distinguish between general types of nuclei (e.g., ^{13}C nuclei in carbonyl (CO) and alpha (C_α) positions) but, for a full assignment, correlation experiments are necessary. These experiments generate two-dimensional (2D) spectra in which pairs of nuclei in close proximity are correlated. Two types of experiments exist, designed to establish intra- and inter-residue correlations. Different ^{13}C spin topologies of the amino acids give characteristic intra-residue correlation patterns, allowing the ^{13}C resonances to be assigned to the different residue types. Experiments

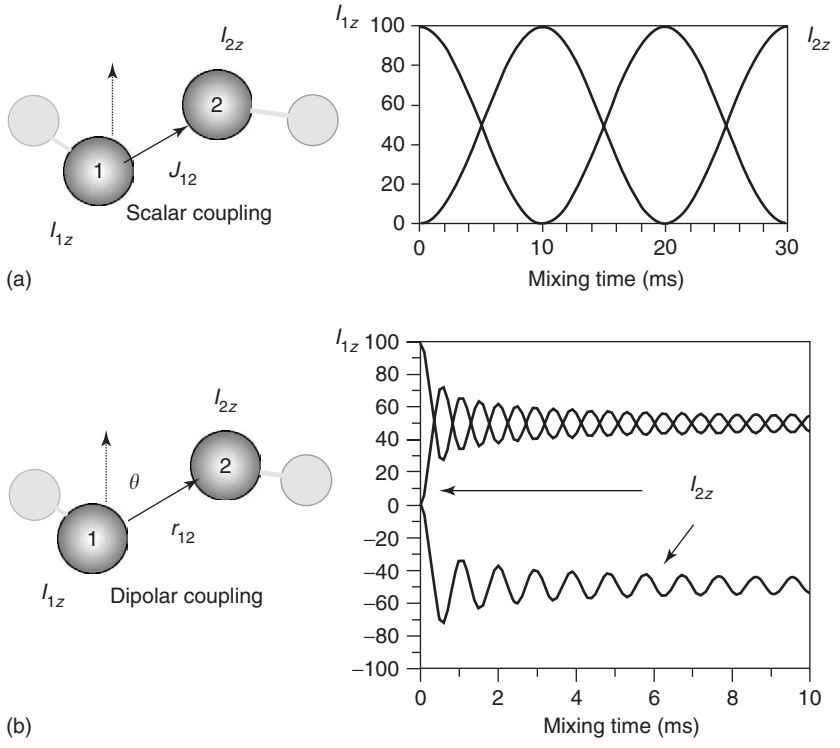


Figure 3 Analytical transfer functions for a two-spin system (denoted by the spin operators I_{1z} and I_{2z}) for the case of (a) the *J*-coupling and (b) the dipolar coupling. In both cases, the spin pair was initially prepared with magnetization only on spin 1. The transfer of polarization is subsequently monitored as a function of the mixing time (given in ms).

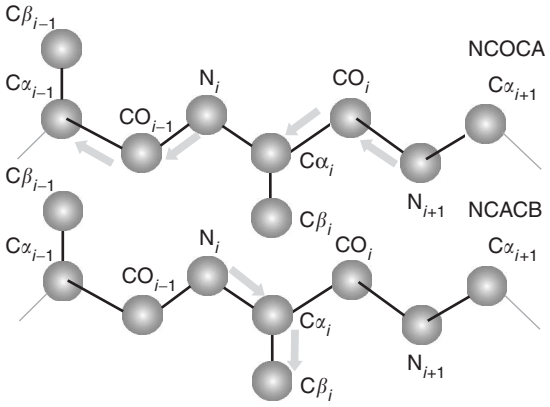


Figure 4 Graphical representation of NCOCA (upper row) and NCACB (lower row) correlation experiments that lead to spectral assignments in HR-SSNMR.

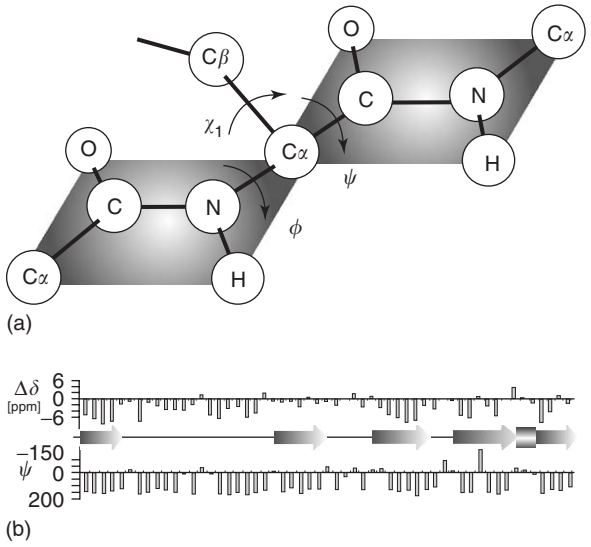


Figure 5 (a) The arrangement of atoms in the backbone of a protein. (b) The correlation between protein backbone structure, i.e., ψ, and secondary chemical shifts $\Delta\delta$ correlating C_α and C_β ^{13}C resonances obtained for the SH3 domain from α spectrin. $\Delta\delta$ is defined by the difference of observed (obs) and random coil (rc) C_α and C_β chemical shifts: $\Delta\delta = (\delta_{C\alpha}(\text{obs}) - \delta_{C\alpha}(\text{rc})) - (\delta_{C\beta}(\text{obs}) - \delta_{C\beta}(\text{rc}))$.

for acquiring these correlation spectra may use single-quantum or double-quantum signal in an indirect dimension to separate out different contributions to the simple one-dimensional spectrum.

To acquire inter-residue information, ^{13}C nuclei can be correlated with ^{15}N nuclei in the same residue and the adjacent residues. This requires selective transfer of polarization between ^{13}C and ^{15}N in order to generate the desired indirect dimension and to ensure good resolution in the direct dimension. A common method involves two experiments, carried out on uniformly ^{13}C and ^{15}N labeled samples. In the first, magnetization is transferred from the backbone ^{15}N nuclei to the adjacent carbonyl (CO) ^{13}C and then, in a subsequent step, on to the α ^{13}C nuclei. This links the ^{15}N in residue i to the ^{13}C nuclei in residue i–1 (**Figure 4**, upper row) and is, in general, referred to as an NCOCA correlation experiment. In the second experiment, magnetization is transferred from the peptide bond ^{15}N nuclei to the nearest neighbor $^{13}C_\alpha$ spin, linking the ^{15}N and ^{13}C nuclei within the same residue. A subsequent (^{13}C, ^{13}C) transfer step leads to C_β resonance of the same residue corresponding to an NCACB-type experiment (**Figure 4**, lower row). From these two experiments, all backbone ^{13}C and ^{15}N nuclei of the polypeptide can be investigated.

Secondary Structure

The secondary structure of a protein encompasses the arrangement of the backbone atoms, depicted in **Figure 5a**. As the bond lengths are invariant, the secondary structure is largely defined by the torsional angles CO($i-1$)–N(i)–C$_\alpha$(i)–CO(i) and N(i)–C$_\alpha$(i)–CO(i)–N($i+1$), known as the ϕ and ψ angles, respectively. These may be measured in a variety of ways. A strong correlation exists between the ^{13}C chemical shifts at the C_α and C_β positions and the values of ϕ and ψ. In particular, the two main backbone conformational types, the α-helix and the β-sheet, may be distinguished by comparing these ^{13}C chemical shifts to the average chemical shifts (so-called "random-coil" values) for each amino acid. An example of the strong correlation between, so-called, secondary chemical shifts $\Delta\delta$ and the torsion angle ψ is shown in **Figure 5b** for a U-[^{13}C, ^{15}N]-labeled sample of the SH3 domain from α spectrin.

Dihedral angle constraints may also be determined directly by experiments monitoring the relative orientation of the appropriate internuclear vectors. This encoding can be carried out by allowing the dipolar interactions to generate cross-peaks in a correlation spectrum characteristic of their relative orientation or by creating a multiple-quantum state that evolves under the influence of surrounding dipolar interactions. Distance measurements are also important for the determination of secondary structure. For example, the $H_\alpha(i)$–$H_N(i+1)$ distance is a sensitive measure of the ψ torsional angle and can be probed by indirect detection schemes in solid-state NMR. Specifically, ^{13}C and ^{15}N labeled samples can also be used to measure distances between residues, which are sensitive to the ϕ and/or ψ angles. In both cases, experiments are used which involve dipolar recoupling between the two nuclei of interest. The buildup of either magnetization transfer or double-quantum coherence depends upon the internuclear distance

and may be analyzed by appropriate theoretical or numerical models to give a measure of the distance. Of particular importance for structural characterization of side chains are HCCH and HNCH torsional angles, both of which can be measured using the methods outlined above.

Tertiary Structure

In order to establish the overall fold of a protein, distances must be measured between separate parts of the peptide chain. For a long time, the determination of these long-range contacts represented a major challenge for solid-state NMR due to an effect known as dipolar truncation, where the strong, short-range dipolar interactions dominate weaker, long-range interactions. As a result, relay effects over several short-range interactions may determine the spin system dynamics. These problems are most serious for uniformly [^{13}C, ^{15}N]-labeled proteins. Three approaches have been recently suggested to circumvent these difficulties. One is to dilute the isotope labeling of the protein by, for example, block labeling. A second approach involves the use of selective recoupling methods that establish polarization transfer for particular two-spin systems only. Finally, useful information may be obtained from an analysis of (^{1}H, ^{1}H) contacts, which is common practice in solution-state NMR. Because of the limited spectral resolution in solid-state ^{1}H NMR spectra, such interactions may be encoded indirectly in high-resolution ^{13}C and/or ^{15}N spectral dimensions. Correspondingly, the 2D NMR experiments have been denoted by CHHC and NHHC correlation methods, respectively.

Similar to the solution-state, NMR structure determination in the solid state subsequently involves the calculation of families of molecular conformations that are consistent with the experimentally derived distance or angle constraints. The number and precision of these parameters determine the accuracy of the resulting 3D structure.

Protein Orientations, Interactions, and Dynamics

In addition to revealing structural information about the protein in a near-biological environment, the orientation of the protein with respect to a macroscopic reference frame (for example, one defined by the membrane bilayer for membrane peptides and proteins) can be investigated. The application of such methods often involves the NMR study of macroscopically aligned polypeptides and has resulted in several high-resolution structures of membrane-embedded systems. Furthermore, interactions between different proteins, proteins–ligand interactions (e.g., involving drugs, hormones, etc.), or monomer–monomer interactions in oligomeric proteins (quaternary structure) can, in principle, be investigated by high-resolution solid-state NMR.

Applications

In the following, three examples for the range of structural problems in immobilized proteins that can be addressed by high-resolution solid-state NMR are given. In **Figure 6**, results of 2D CHHC correlation experiment conducted on a uniformly [^{13}C, ^{15}N]-labeled sample of the 2×10.4 kDa dimeric form of the regulatory protein Crh in microcrystalline form are shown. Each correlation reflects close proton–proton contacts encoded in ^{13}C resonance frequencies. Three sections are highlighted where short, inter-residue (^{1}H, ^{1}H) distances are observed. A combination of these results with NC correlation experiments discussed earlier leads to NMR assignments of the largest protein studied by HR-SSNMR thus far and provides the basis for investigating the Crh folding pathway from a monomeric, soluble protein to a domain-swapped dimer observed in the microcrystalline state. Domain-swapping is a process by which one protein molecule exchanges a domain with an identical partner, and has been discussed as a general means for creating protein complexes and as a mechanism for misfolding and aggregation.

Solid-state NMR not only can report on the complete structure of a membrane protein but can also probe dynamics in the protein of interest. Both aspects have recently been investigated in the context of a uniformly [^{13}C, ^{15}N]-labeled version of the LH2 light-harvesting complex, one of the largest membrane systems studied by solid-state NMR to date ($\sim$150 kDa). Here, the application of ultrahigh magnetic fields (up to 750 MHz) was crucial to establish high-resolution conditions in the solid state. Two-dimensional (^{15}N, ^{13}C) correlation experiments (such as shown in **Figure 7**) were conducted to assign backbone and side-chain resonances at different temperatures. At -10°C, a variety of spectral assignments within the transmembrane sections of the protein were possible. Additional correlations occur at lower temperatures consistent with the immobilization of flexible loop regions of the protein. A more detailed structural study in membrane proteins could include (^{13}C, ^{13}C) or indirect (^{1}H, ^{1}H) correlation experiments possibly in conjunction with advanced isotope-labeling approaches, to determine the complete 3D structure of a membrane protein in the solid state.

To date, no structural information of a high-affinity ligand bound to a G-protein-coupled receptor

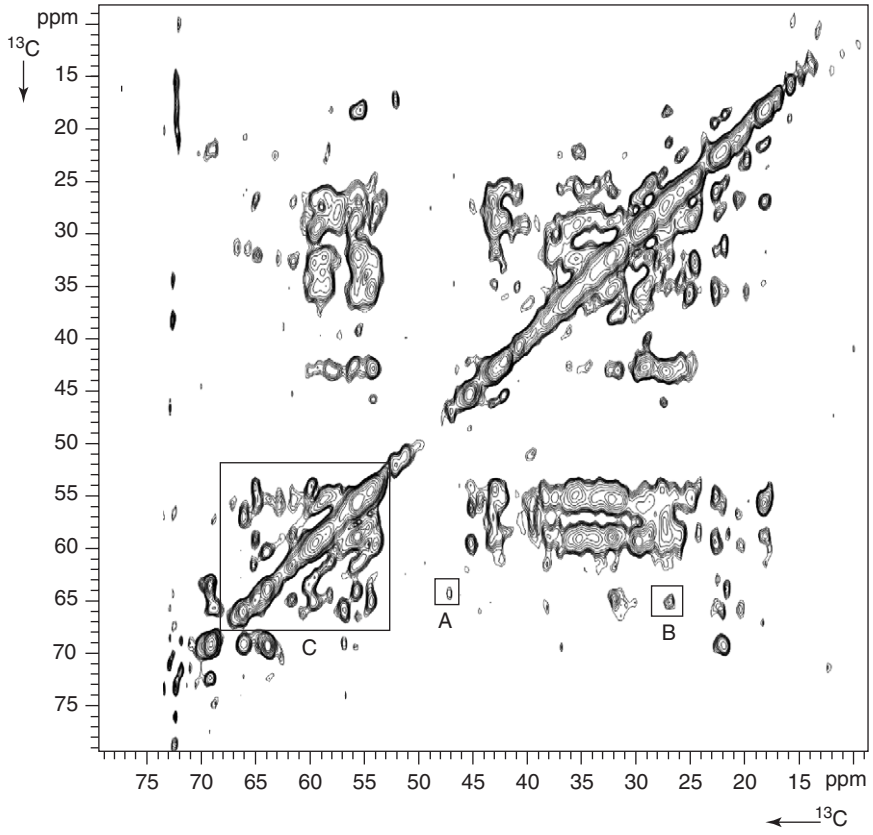


Figure 6 2D CHHC correlation experiment conducted on a uniformly [^{13}C, ^{15}N]-labeled sample of the 2 × 10.4 kDa dimeric form of the regulatory protein Crh. Each correlation reflects close proton–proton contacts encoded in ^{13}C resonance frequencies. Three sections are highlighted where short, inter-residue (^{1}H, ^{1}H) distances are observed: A – involving Gly49 and Ser46, B – Pro18 - Leu10, and C – a variety of C_α–C_α contacts.

(GPCR) is available. The recombinant expression of GPCRs in large quantities is usually difficult and must involve carefully optimized biochemical procedures. Restrictions regarding the availability of functional receptors also affect the quantities of ligand that can be studied. Moreover, the chemical environment including lipids and receptor protein can hamper the unambiguous spectral identification of a bound ligand in a solid-state NMR experiment. It has been recently shown how 2D solid-state NMR experiments can be used to detect microgram quantities of bound neurotensin, a 13-residue neuropeptide that binds in high affinity to the NTS-1 (101 kDa) receptor. For the 6-residue, biologically active, C-terminal sequence of neurotensin, a homonuclear (2Q, 1Q) correlation spectrum is sufficient to assign all C_α and C_β resonances of the uniformly [^{13}C, ^{15}N]-labeled ligand. These chemical shift assignments can be used to construct the backbone model of the ligand complexing with the receptor. Using the only high-resolution structure of a GPCR, rhodopsin, as a template, this information could be used to establish a molecular model for the binding pocket of NT(8-13) (**Figure 8**).

Conclusions and Outlook

In many research fields, NMR methods have made important contributions to the present understanding of molecular structure and function. In particular, NMR has become a standard method for the characterization of 3D structure and dynamics of proteins that undergo fast molecular reorientations in solution. Such macroscopic sample conditions are, however, difficult to establish for protein aggregates, which are associated with a growing number of protein folding diseases. Moreover, integral membrane proteins, which constitute about 30% of all proteins, are often difficult to solubilize or crystallize in functional form. In all these cases, NMR techniques particularly designed for the study of solid-phase systems ("solid-state NMR") offer unique possibilities to elucidate structural or dynamic parameters at atomic resolution.

In this article, a summary of high-resolution solid-state NMR methods that permit the detection of multiple structural parameters from a single protein sample or NMR data set is given. Recent applications in (membrane) proteins underline the growing

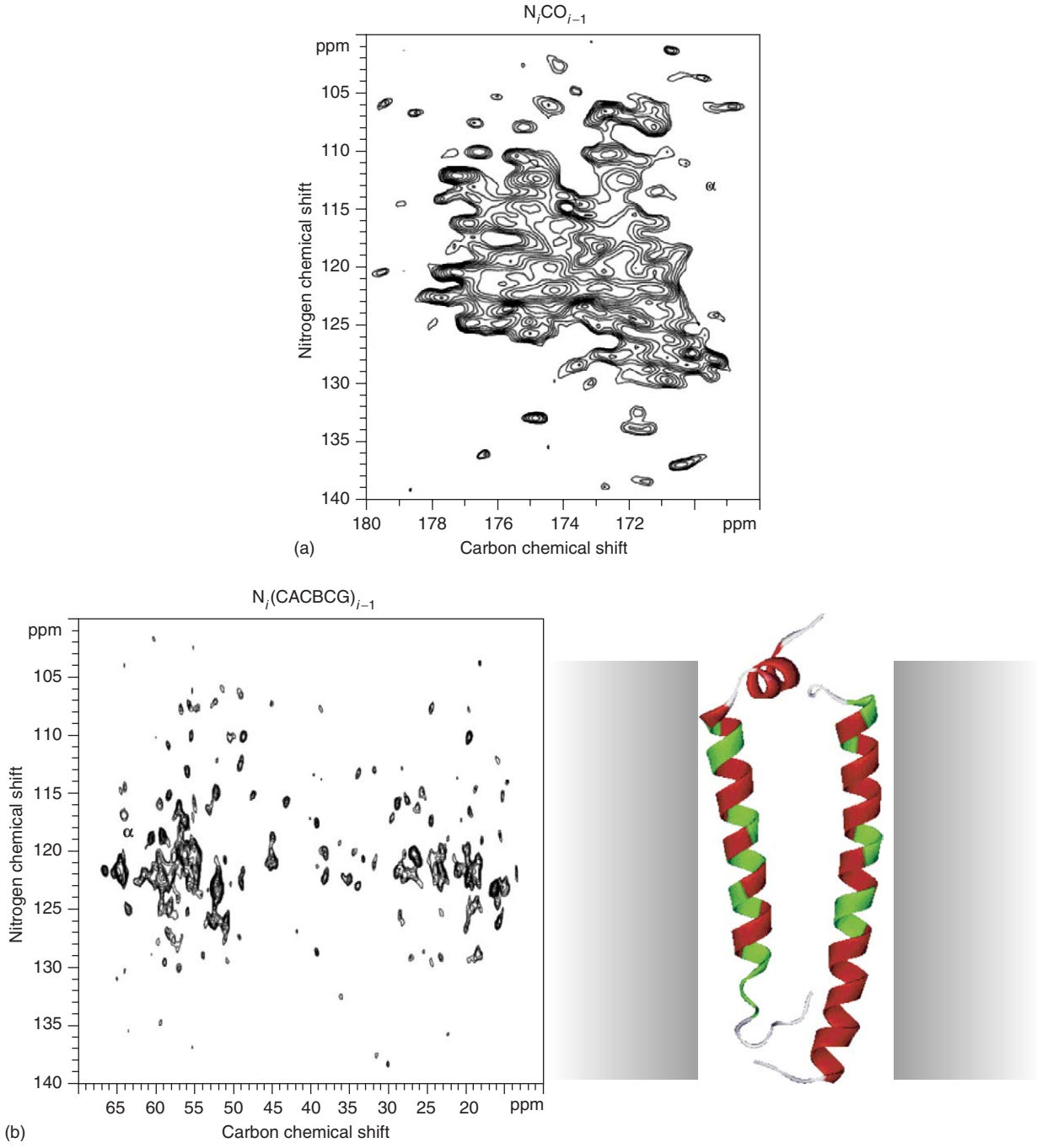


Figure 7 Combined NC-CC 2D transfer experiment on U-[^{13}C, ^{15}N]-labeled LH2 revealing a variety of N_iCO_{i-1} (a) and $N_i(CACBCG)_{i-1}$ (b) correlations. A schematic view of the backbone structure of the LH2 protomer complex embedded in a model membrane is shown. Assigned residues are indicated in green.

potential of HR-SSNMR to provide exclusive insight into the microscopic details of biological functioning. Advances regarding sample preparation (for example, including modular labeling, *in vitro* expression, and intein technology) and improvements in NMR hardware instrumentation could open up additional areas of solid-state NMR research, such as the investigation of large protein–protein complexes or the complete 3D characterization of larger membrane proteins. Solid-state NMR studies of multiply labeled biomolecules will furthermore profit from improved procedures for calculating 3D structures, in particular, in the presence of ambiguous or a limited number of structural constraints. In general, protein motion does not hinder the application of solid-state NMR methods, providing a very efficient

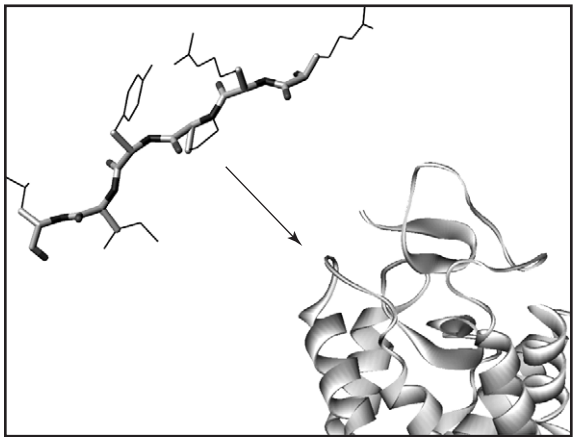


Figure 8 Schematic representation of the neurotensin NT (8-13) backbone structure, as obtained from solid-state NMR, and its interaction with a G-protein coupled receptor. For illustration, the only currently known GPCR structure (rhodopsin) is included.

means of studying protein folding, flexibility, and function under biologically relevant conditions. Complementary to solution-state techniques and crystallographic methods, solid-state NMR can give detailed insight into protein structure and function.

Acknowledgments

The authors would like to thank their collaborators and group members who contributed to the work described here. This work has been funded in part by an Alexander-von-Humboldt fellowship for C.E.H.

See also: Metalloproteins, Electron Transfer in; Metalloproteins, Structural Determination of; Solid-State NMR Structural Studies of Proteins.

PACS: 82.56.Pp; 82.56.Ub; 87.14.Ee; 87.15.Cc; 87.64.Hd

Further Reading

Baldus M (2002) Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Progress in NMR Spectroscopy* 41: 1–47.

Cross TA and Opella SJ (1994) Solid-state NMR structural studies of peptides and proteins in membranes. *Current Opinion in Structural Biology* 4: 574–581.

Ernst RR, Bodenhausen G, and Wokaun A (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford: Clarendon.

Griffin RG (1998) Dipolar recoupling in MAS spectra of biological solids. *Nature Structural Biology* 5: 508–512.

Haeberlen U (1976) High resolution NMR in solids: selective averaging. In: Waugh JS (ed.) *Advances in Magnetic Resonance*. New York: Academic Press.

Levitt MH (2002) Symmetry-based pulse sequences in magic-angle spinning solid-state NMR. In: Grant DM and Harris RK (eds.) *Encyclopedia of Nuclear Magnetic Resonance: Supplementary Volume*, pp. 165–196. Chichester: Wiley.

McDowell LM and Schaefer J (1996) High-resolution NMR of biological solids. *Current Opinion in Structural Biology* 6: 624–629.

Mehring M (1983) *Principles of High Resolution NMR in solids*, 2nd edn. Berlin: Springer.

Thompson LK (2002) Solid-state NMR studies of the structure and mechanisms of proteins. *Current Opinion in Structural Biology* 12: 661–669.

Tycko R (2000) Solid-state NMR as a probe of amyloid fibril structure. *Current Opinion in Chemical Biology* 4: 500–506.

Solidification: Models and Simulations

J A Dantzig, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Introduction

Solidification has been an important manufacturing process dating as far back as the bronze age. In more recent times, research has focused on controlling the microstructure – the internal segregation pattern which forms upon freezing – in order to control the properties of the product. This article describes how thermodynamics and heat and mass transfer interact to produce the complex patterns associated with solidification in crystalline materials. The discussion is restricted to binary alloys, which still leaves one with a rich spectrum of behavior. To begin with, the physical phenomena which occur during solidification are discussed, and then, some of the background material that supports the description is provided.

As a liquid alloy is cooled, crystallization begins at some temperature. In general, the composition of the solid is different from that of the liquid. Under the conditions normally found in conventional processes, thermodynamic equilibrium is established at the liquid–solid interface, and this provides a relationship between the interface temperature and the compositions of the phases. Microstructure formation represents the competition between heat transfer and mass transfer, mediated by thermodynamic equilibrium. The material properties which control the rates of the transport processes are the thermal diffusivities α_S and α_L, and the chemical diffusivities $\mathscr{D}_S$ and $\mathscr{D}_L$. Throughout this article, the subscripts S and L are used to indicate the solid and liquid phases,

respectively. For most materials, these quantities follow the relation that $\alpha_S \approx \alpha_L \gg D_L \gg D_S$. It is shown by scaling analysis that this hierarchy of values leads to a situation where the motion of the liquid–solid interface is controlled by heat flow, under the action of external boundary conditions, and the solute is unable to keep up with the heat transfer process. The resulting segregation pattern, called the microstructure, provides a frozen record of the results of the competition.

In the following sections, the thermodynamics of solutions is described first in order to establish the relationship between temperature and composition during solidification. Heat and mass transfer, and the unique mathematical and computational issues posed by the moving phase boundary are then taken up. Finally, some example calculations are given showing how these concepts can be used to compute microstructures.

Thermodynamics and Equilibrium Phase Diagrams

Thermodynamics tells about the equilibrium state of materials – the phases present, their compositions, and relative amounts at a given temperature – after an infinite time has passed. Although this is clearly never the case, the concepts of phase equilibria normally do apply locally, at the solidifying interface, and so an understanding of solidification begins with thermodynamics of solutions. The thermodynamic state of a mixture can be specified via the compositions of its constituents, and two of the three thermodynamic variables: temperature T, pressure p, and volume V. The third thermodynamic variable comes from an equation of state. For condensed phases, it is most convenient to choose T and p as the thermodynamic variables. It may also be noted that all but one of the compositions of the phases are independent, once the overall composition of the alloy is specified.

The free energy of the mixture can be computed from these thermodynamic quantities. Equilibrium is defined as the minimum in the free energy. Here, open systems are considered, for which the effect of pressure can be neglected, and a model which gives the free energy of the solution for any given temperature and composition is introduced. The regular solution model is a very powerful model which permits a wide range of experimentally observed behavior to be reproduced through the variation of just two parameters. Consider binary alloys, with components A and B. The composition of the alloy can be written as a mole fraction X, or as a mass fraction designated C. These measures are related through the molecular weights M_A and M_B, such that $C = M_B X/(M_A + (M_B - M_A)X)$. Solution models are more conveniently formulated in terms of X, while transport models are more compact when written in terms of C. In the regular solution model for binary alloys, the free energies of mixing G^{mix} in the solid and liquid states are given by

$$G_L^{mix}(X_L, T) = RT(X_L \ln X_L + (1 - X_L) \times \ln(1 - X_L)) + \Omega_L X_L (1 - X_L) \quad [1]$$

$$G_S^{mix}(X_S, T) = (1 - X_S)\Delta S_f^A (T - T_m^A) + X_S \Delta S_f^B (T - T_m^B) + RT\,(X_S \ln X_S + (1 - X_S)\ln(1 - X_S)) + \Omega_S X_S (1 - X_S) \quad [2]$$

where R is the ideal gas constant, ΔS_f^A and ΔS_f^B are the entropies of fusion of pure materials A and B, respectively, T_m^A and T_m^B are their melting points and Ω_S and Ω_L are the regular solution model parameters.

Three cases are considered, involving different values of Ω_S and Ω_L, and the relationship between the free energies and the equilibrium phase diagram is demonstrated. Free energy curves for the solid and liquid are shown at an interesting temperature for each case at the top of **Figure 1**. The intersection of the two free energy curves, called the T_0 point, indicates a temperature T_0 and composition X_0 which lies in a region of two-phase coexistence.

The compositions X_S and X_L of the solid and liquid phases in equilibrium are the same for any alloy in the two-phase region, including, of course, X_0. It follows from a simple mass balance that the fraction of each phase present is

$$f_S = \frac{X_L - X_0}{X_L - X_S}, \qquad f_L = \frac{X_0 - X_S}{X_L - X_S} \quad [3]$$

It follows that the total free energy G^{total} of the two-phase mixture is

$$G^{total}(x_0, T; [X_S, X_L]) = f_S G_S^{mix}(X_S, T) + f_L G_L^{mix}(X_L, T) \quad [4]$$

where the parametric dependence of G^{total} on the phase compositions X_S and X_L are explicitly indicated.

The equilibrium phase compositions are those which minimize the free energy. Thus, this may be posed as an optimization problem: find those values of X_S and X_L which minimize G^{total} in eqn [4]. The optimization problem has side constraints. For the example shown in **Figure 1a**, $0 \leqslant X_L < X_0$ and $X_0 < X_S \leqslant 1$. The equilibrium values for X_S and X_L are also the points on the two curves which share a common tangent, as indicated in the upper figure.

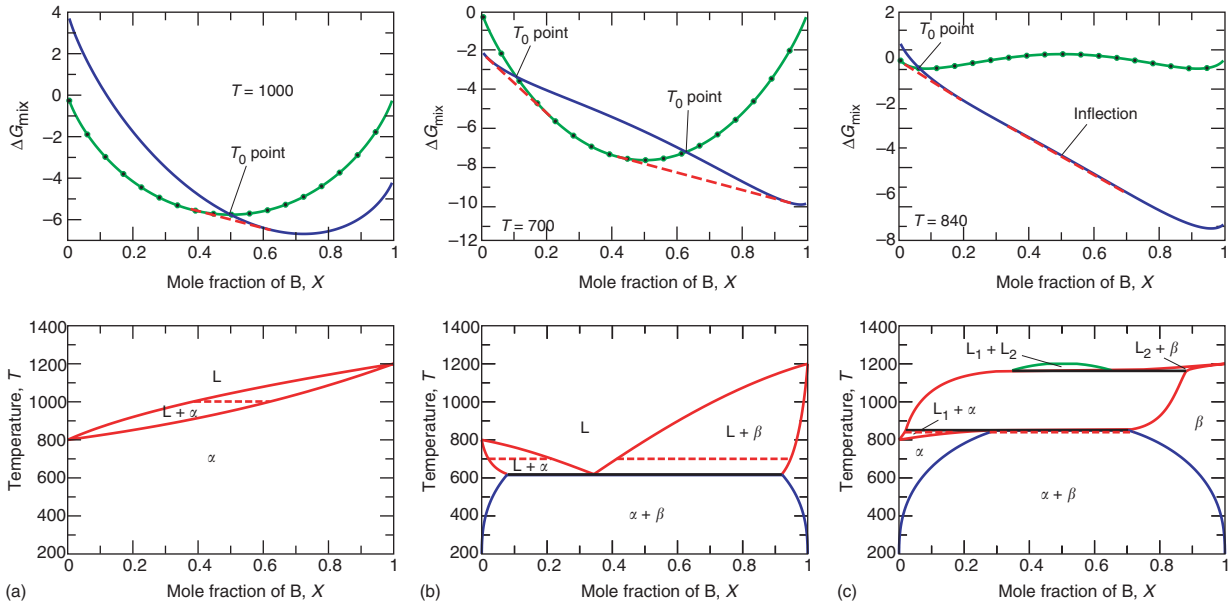


Figure 1 Correspondence between free energy and equilibrium phase diagrams for three different sets of parameters. For all three cases, $T_m^A = 800$, $T_m^B = 1200$, and $\Delta S_f^A = \Delta S_f^B = 20$. (a) An isomorphous system ($\Omega_S = \Omega_L = 0$), (b) a eutectic system ($\Omega_S = 15\,000$, $\Omega_L = -15\,000$), and (c) a system with a peritectic and a monotectic reaction ($\Omega_S = 15\,000$, $\Omega_L = 20\,000$) are shown. The dashed lines in the figures show how the common tangents in free energy correspond to two-phase equilibria. (●—●: liquid; ——: solid; and – – –: common tangent.)

The entire phase diagram can be constructed by performing a series of optimization problems over a range of temperatures, and plotting the results in the form shown in the lower part of **Figure 1**. The locus of all of the (X_L, T) pairs is called the "liquidus" curve, and the similar locus of all (X_S,T) curves is called the solidus. In the next section, these two curves are referred to as T_{liq} and T_{sol}, respectively.

The regular solution can produce a rich variety of phase diagrams, as shown in **Figure 1**. By varying only Ω_S and Ω_L, many different types of reactions can occur. For example, in **Figure 1b**, there are two points of common tangency between the liquid and solid free energy curves, and this leads to a eutectic reaction at lower temperature when the two liquid compositions coincide and there is a three-phase equilibrium. In **Figure 1c**, in addition to the liquid–solid equilibrium, a miscibility gap in the solid is seen. The solution parameters are normally found by fitting to experimental data, or from electronic structure models. While the principles are the same, the process of finding equilibrium compositions is more complicated when binary compounds A_aB_b can form, and when more than two components are used. The reader is referred to the "Further reading" section to find out more about these cases.

The thermodynamic data reveals the phases and compositions which should be in equilibrium for a given temperature and composition. However, it contains no information about either spatial distribution or the dynamics of the transformation. To address these issues, attention is drawn to heat and mass transfer.

Heat and Mass Transport

The discussion of microstructure evolution is given here in the context of the prototype experimental system called "directional solidification," illustrated in **Figure 2**. An alloy of interest is encapsulated between two microscope slides, and placed on a microscope stage with a fixed temperature gradient maintained by two temperature-controlled blocks. The alloy is frozen by translating the stage at a controlled velocity V. The experiment is normally performed in such a way that the liquid–solid interface is fixed in space, that is, the solidification velocity and the speed of traverse are the same. As suggested by the simulated microstructure shown in **Figure 2b**, a flat interface is unstable, leading to a complex pattern called the microstructure. This microstructure consists of tree-like forms called dendrites, which evolved from an initially flat interface. Primary arms are seen growing from left to right with a characteristic spacing that may be called λ_1. The goal here is to predict such structures from a knowledge of the alloy and process parameters.

To this end, the heat and mass transfer problem is taken up for the moving interface. The governing equations come from statements of the conservation

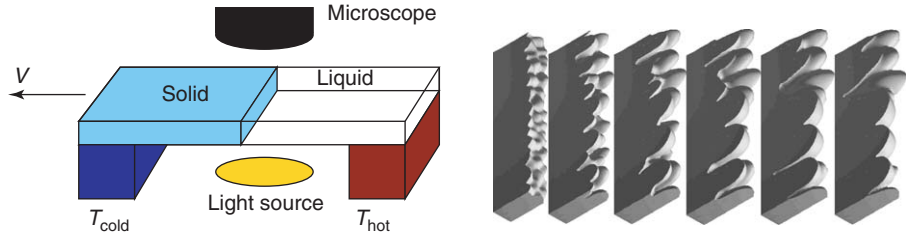


Figure 2 Schematic of (a) a directional solidification apparatus, and (b) a simulated microstructure showing the interface evolving over time. This simulation was done in three dimensions using adaptive finite elements, discussed later.

of energy and mass in the system, respectively. A linear relationship between the heat flux and temperature gradient (Fourier's law), and a similar relationship between solute flux and composition gradient (Fick's law) are assumed. It is more convenient to describe mass transport using the mass fraction C, as opposed to the mole fraction X used in the preceding section. The governing equations for heat and mass transport in the liquid and solid are

$$\rho_i c_p^i \frac{\mathrm{D}T_i}{\mathrm{D}t} = \nabla \cdot (k_i \nabla T) \qquad [5]$$

$$\rho_i \frac{\mathrm{D}C_i}{\mathrm{D}t} = \nabla \cdot (\rho_i \mathscr{D}_i \nabla C_i) \qquad [6]$$

where $i = \mathrm{S}, \mathrm{L}$ indicates the phase. The symbol $\mathrm{D}/\mathrm{D}t$ represents the material derivative $\partial/\partial t + v \cdot \nabla$ where t is time and is the mass-averaged velocity, ρ is the density, c_p is the specific heat at constant pressure, k is the thermal conductivity, C is the composition, measured in mass fraction, and $\mathscr{D}$ is the diffusion coefficient.

As the interface moves, it surrenders the latent heat of fusion L_f. The concentrations of the liquid and solid at the interface are different, given by the condition of thermodynamic equilibrium, as discussed previously. Denoting the normal velocity of the interface as V_n, energy and solute balances at the interface may be written as

$$k_\mathrm{S}\frac{\partial T_\mathrm{S}}{\partial n} - k_\mathrm{L}\frac{\partial T_\mathrm{L}}{\partial n} = \rho_\mathrm{S} L_\mathrm{f} V_\mathrm{n} \qquad [7]$$

$$\mathscr{D}_\mathrm{S}\frac{\partial C_\mathrm{S}}{\partial n} - \mathscr{D}_\mathrm{L}\frac{\partial C_\mathrm{L}}{\partial n} = (C_\mathrm{L} - C_\mathrm{S})V_\mathrm{n} \qquad [8]$$

where derivatives with respect to n represent the projection of the gradient normal to the interface. The necessary equations for solidification are completed using the assumption of thermodynamic equilibrium at the interface. This leads to a "coupling condition" between the temperature and composition, known as the Gibbs–Thomson equation:

$$T_\mathrm{S} = T_\mathrm{L} = T_\mathrm{liq}(C_\mathrm{L}) - \left(\frac{\gamma}{\Delta S_\mathrm{f}}\right)\kappa - \beta V_\mathrm{n} \qquad [9]$$

where γ is the surface tension, ΔS_f is the entropy of fusion, κ is the mean curvature, and β is a coefficient interface attachment kinetics. It is worth reiterating that $T_\mathrm{liq}(C_\mathrm{L})$ comes from thermodynamics and the equilibrium phase diagram.

In order to elucidate some important physics, consider the special case of directional solidification with a flat interface; there is no motion except for the translation of the stage at fixed velocity V in the x-direction. It is assumed that all material properties are constant, and that $\alpha_\mathrm{S} = \alpha_\mathrm{L}$. The latter assumption simplifies the presentation, and the temperature is simply written as T. The governing equations are now all one dimensional. If it is assumed that the process is steady with respect to a frame which is fixed at the interface, the governing equations reduce to

$$-V\frac{\partial T}{\partial x} = \alpha\frac{\partial^2 T}{\partial x^2}, \qquad -V\frac{\partial C_\mathrm{L}}{\partial x} = \mathscr{D}_\mathrm{L}\frac{\partial^2 C_\mathrm{L}}{\partial x^2}$$
$$-V\frac{\partial C_\mathrm{S}}{\partial x} = \mathscr{D}_\mathrm{S}\frac{\partial^2 C_\mathrm{S}}{\partial x^2} \qquad [10]$$

Even though the interface is considered to be flat (at the moment), **Figure 2b** shows clearly that there are characteristic length scales. The length is scaled on the primary dendrite spacing λ_1, defining $\xi = x/\lambda_1$. Each form in eqn [10] will then have a dimensionless ratio on the left-hand side called the Péclet number, representing the relative importance of advective to diffusive transport. The transparent organic system succinonitrile-acetone has been extensively studied experimentally, and if typical values from such experiments are taken to compute the dimensionless coefficients (see **Table 1**), one obtains

$$\frac{V\lambda_1}{\alpha} = 0.018, \qquad \frac{V\lambda_1}{D_\mathrm{L}} = 1.6$$
$$\frac{V\lambda_1}{D_\mathrm{S}} = 16\,000 \qquad [11]$$

Dropping terms that are small compared to the ones in eqn [10] leaves

$$0 \approx \frac{\partial^2 T}{\partial x^2}, \qquad -V\frac{\partial C_\mathrm{L}}{\partial x} = \frac{\partial^2 C_\mathrm{L}}{\partial x^2}, \qquad \frac{\partial C_\mathrm{S}}{\partial x} \approx 0 \qquad [12]$$

Table 1 Physical properties and experimental parameters for succinonitrile-acetone system used to evaluate dimensionless parameters

Quantity	*Value*	*Units*
α_S	1.125×10^{-7}	$m^2\,s^{-1}$
α_L	1.125×10^{-7}	$m^2\,s^{-1}$
$\mathscr{D}_L$	1.3×10^{-9}	$m^2\,s^{-1}$
$\mathscr{D}_S$	1.3×10^{-13}	$m^2\,s^{-1}$
V	3.4	$\mu m\,s^{-1}$
λ_1	610	μm

The first and third equations are easily integrated to find that the temperature profile is linear and "frozen" in space, and the composition of the solid is constant

$$T = T_{cold} + G(x - x_{cold}), \qquad C_S = \text{constant} \quad [13]$$

The composition in the liquid is obtained using eqn [8] as a boundary condition at the interface, and the far-field condition that $C_L(x \to \infty) = C_0$, the nominal composition of the alloy. As part of the solution process, one finds that $C_S = C_0$. Thermodynamic equilibrium at the interface implies that the temperature there is $T_{sol}(C_0)$, and correspondingly that $C_L(0) = C_{liq}(T_{sol})$, and so, $\Delta C_0 = C_{liq}(T_{sol}) - C_0$, and finally the solution for C_L is

$$C_L = C_0 + \Delta C_0 \exp\left(-\frac{Vx}{\mathscr{D}_L}\right) \quad [14]$$

The solution is shown schematically in **Figure 3**.

When the temperature gradient is relatively low, there is a region ahead of the interface where the local liquidus temperature is below the actual temperature, shown as a shaded region in the lower part of **Figure 3**. The existence of this region makes the interface unstable with respect to morphological disturbances. The interface will be stable when $G \leqslant m_L \partial C_L/\partial x$, where m_L is the slope of the liquidus curve on the equilibrium phase diagram at $T = T_{sol}$. This can be rewritten in terms of a dimensionless constant $\mathscr{M}$ called the morphological number:

$$\mathscr{M} = -\frac{m_L V \Delta C_0}{G \mathscr{D}_L} \leqslant 1 \quad [15]$$

To obtain information about the length scales which evolve from the unstable interface, one can perform a linear stability analysis, using the planar interface as a base state. Space does not permit reproduction of the analysis here, but there is an important piece of physics that comes out of it. When the interface is flat, the curvature term in eqn [9] plays no role. However, when the interface shape is perturbed, surface tension provides a restoring force. The strength of the surface tension is characterized

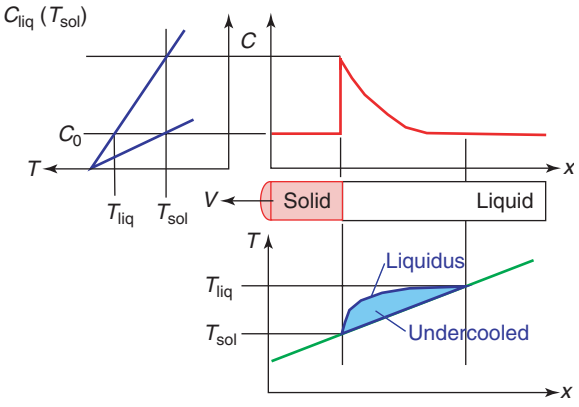


Figure 3 Pictorial view of the temperature and composition profiles during directional solidification with a flat interface. The solid and liquid compositions at the interface are given by the equilibrium phase diagram.

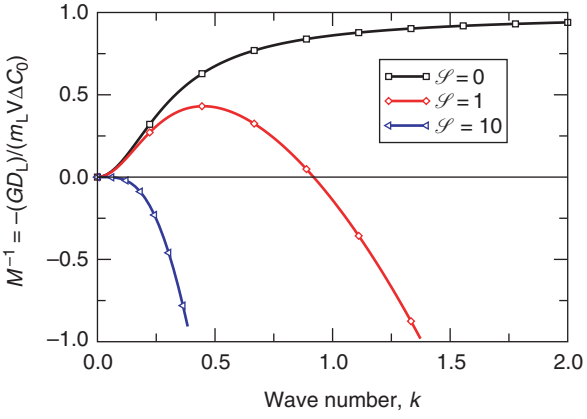


Figure 4 Neutral stability curves for plane front growth for different values of $\mathscr{S}$. Values of $1/\mathscr{M}$ above the neutral curve are stable, while values below are unstable. Negative values of $1/\mathscr{M}$ are physically unattainable.

through a second dimensionless parameter $\mathscr{S}$ given by

$$\mathscr{S} = -\frac{\gamma V}{\mathscr{D}_L m_L \Delta C_0} \quad [16]$$

Figure 4 shows the dispersion relation for various values of $\mathscr{S}$. It can be seen that for $\mathscr{S} = 1$, the presence of surface tension stabilizes small wavelength disturbances, and that only certain wave numbers are unstable. These are the ones that one expects to see in the directional solidification experiment. Note that for sufficiently high values of $\mathscr{S}$, such as $\mathscr{S} = 10$, all wavelengths are stable, a condition known as absolute stability. The linear stability analysis is useful for understanding the length scales of the initial disturbances. **Figure 2** shows, however, that a complex selection process continues after the initial instability, and this is beyond the scope of analytical treatment.

The critical question in directional solidification is to predict the microstructure which will evolve under a given set of processing conditions. The answer is understood in a semiquantitative way for a single isolated dendrite. If one imagines an isolated dendrite whose shape near the tip resembles a paraboloid of revolution, then the shape is characterized by a single parameter, the tip radius ρ_{tip}. A solution of the transport equations gives only the product $\rho_{\text{tip}}V$ as constant, leaving the exact shape undetermined. The stability analysis of the transport solution indicates that $\rho_{\text{tip}}^2 V = \sigma^\star$, a constant, and the two criteria combine to provide a unique solution or operating state. The value of the selection constant $\sigma^\star$ is determined by experiment. The theory for the isolated dendrite has been applied to directional solidification, with reasonable agreement with experiment. The dynamics of the selection process, and dependence on initial conditions must be addressed by numerical simulations, which is taken up in the next section.

Numerical Simulation of Microstructure

The numerical simulation of microstructure evolution poses some unique challenges. The governing equations of heat and mass transfer eqns [5] and [6] are to be solved, satisfying the boundary conditions given in eqns [7]–[9]. The two principal difficulties are the disparity in length and time scales in the heat and mass transfer processes, and the fact that the boundary conditions on the liquid–solid interface must be applied on a boundary whose position is unknown *a priori*.

While the length/time scale problem can be attacked by brute force, the free-boundary problem requires more care. There are two principal approaches to the problem: fixed-grid methods, which somehow spread the interface over a finite width, and recover its position from the solution; and interface-tracking methods, in which the mesh is adjusted locally to ensure that grid points coincide with the interface. In the fixed-grid methods, there is no longer a phase boundary, and the technique by which the interface is spread ensures that the boundary conditions are met. Interface-tracking methods specifically satisfy the boundary conditions. Most such methods follow a two-step procedure. First, the interface position is fixed, and the field equations are solved for temperature and composition using one of the boundary conditions, typically eqn [9]. Next, the interface is relocated using the remaining boundary conditions, typically the flux conditions. The grid is then adjusted, and the procedure begins anew. There are advantages and disadvantages to both methods, and to some extent, the choice is a matter of taste. To demonstrate this point, one example of a fixed-grid method, and another of an interface-tracking method are included.

One of the fixed-grid methods is called the phase-field method. In this approach, a continuous function ψ is introduced which has the property that $\psi = 1$ corresponds to the solid phase, $\psi = -1$ corresponds to the liquid phase, and intermediate values are associated with the interface. The phase-field function ψ satisfies an evolution equation of the form

$$\tau \frac{\partial \psi}{\partial t} = -\frac{\delta \mathscr{F}}{\delta \psi} \qquad [17]$$

where $\delta/\delta\psi$ indicates a variational derivative, and $\mathscr{F}$ is a function of the form

$$\mathscr{F} = \int_V \frac{1}{2}|w\nabla\psi|^2 + f(C, T, \psi)\, \mathrm{d}V \qquad [18]$$

where w is the interface width, and the function $f(C, T, \psi)$ is chosen such that it has a "double well" form, with minima at $\psi = \pm 1$, and contrived such that the functional dependence on C and T ensures that thermodynamic equilibrium is enforced. Flux conditions are satisfied by adding source terms proportional to $\partial\psi/\partial t$ to the field equations for T and C. There are many possible choices for f, but they should all lead to the same computational result when their parameters are chosen properly.

There are two major challenges in using the phase-field method. It is necessary to determine the relationship between the phase-field model parameters and the physical parameters, such as γ and β in eqn [9]. This usually requires an asymptotic analysis of the phase-field model, and may lead to very restrictive computational limits for some sets of parameters, most notably when the diffusion coefficients in the liquid and solid are quite different. This problem can be solved, but it always requires that computations be carried out such that the interface width w is small compared to the characteristic diffusion length, in this case $\mathscr{D}_{\text{L}}/V$. Further, the grid spacing Δ must be small compared to w in order to resolve the interface in the computations. The second difficulty comes from the need to resolve microstructural length scales as well, which implies that the computational domain characteristic length $\mathscr{L}$ must be large compared to $\mathscr{D}_{\text{L}}/V$.

Taking these two requirements together, the number of grid points in each direction in a uniform grid is $N = \mathscr{L}/\Delta = \mathscr{O}(10^3)$. This poses a severe limit on 2D computations, and makes 3D computations prohibitively expensive. One approach to deal with this restriction is to use an adaptive grid, fine near the

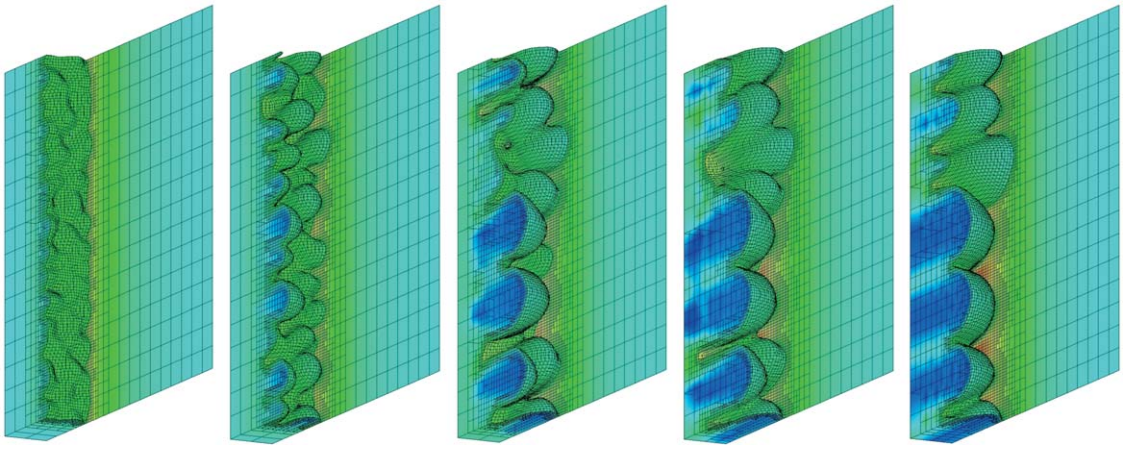


Figure 5 Time sequence of images simulating directional solidification using an adaptive finite element mesh. Color represents concentration. Note the evolution of length scales and segregation patterns during growth.

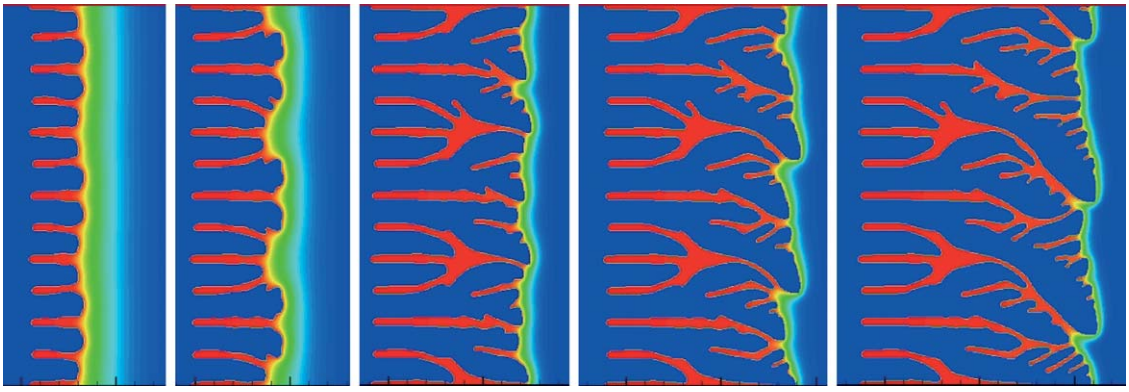


Figure 6 Sequence of images depicting the microstructure and concentration field during directional solidification, obtained using a level-set method. The crystallographic direction corresponding to the direction of fastest growth is oriented at 45° to the growth direction. Notice that the concentration field ahead of the interface is compressed as the growth velocity increases with time.

interface, and much coarser elsewhere. **Figure 5** shows an example of this technique for computing a directionally solidified microstructure. The simulation begins with a random disturbance on the interface, which evolves over time to form the pattern shown. The temperature gradient is assumed to remain constant (see eqn [13]) and the material moves through the domain at the pulling speed V. The figures have stripped most of the domain, showing the interface and concentration profiles in the solid on the front, and in the liquid on the rear of the domain. Periodic boundary conditions are applied on the narrow faces. Clearly, there are computational costs associated with the adaptive gridding procedure, but it enables these computations.

The level-set method uses front tracking to study microstructure evolution. In the level-set approach, there is an underlying fixed grid, and the interface is tracked as it moves through by defining a level-set function ϕ, contrived such that the interface corresponds to the level set $\phi = 0$, and $\phi(\boldsymbol{x})$ represents the normal distance from the interface. The field equations are first solved with the interface fixed, applying the Gibbs–Thomson condition. Determining an accurate representation for the curvature (for use in the Gibbs–Thomson equation) from the sequence of points, which define the interface, is one of the challenges of this method. In the second step, the interface is moved using a pure advection equation for ϕ:

$$\frac{\partial \phi}{\partial \mathrm{t}} + \boldsymbol{u} \cdot \nabla \phi = 0 \qquad [19]$$

The velocity $\boldsymbol{u}$ is chosen such that on the interface, $\boldsymbol{u} \cdot \boldsymbol{n} = V_{\mathrm{n}}$. Solution of eqn [19] requires great care to control oscillations and artificial diffusion due to truncation errors in the numerical schemes. Most analysts use a high-order WENO scheme to integrate eqn [19]. One of the main advantages of this method is that there is no asymptotic analysis required, and the typical condition that $\mathscr{D}_{\mathrm{S}} \ll \mathscr{D}_{\mathrm{L}}$ is readily incorporated. An example calculation of a 2D dendritic interface, obtained using level sets, is shown in **Figure 6**.

There are similar issues in the level-set method to those described for the phase-field method concerning the need for higher resolution near the interface than far away from it. However, the level-set method is less restrictive on this point, because it does not need to resolve a diffuse interface. Rather, the restriction on grid spacing comes from the need to resolve the local diffusion fields and the interface curvature. The combination of adaptive gridding with level sets for interface tracking provides a powerful combination for this type of problem.

Outlook

The modeling of microstructure development is becoming mature. The combination of more powerful computers and advanced modeling techniques has made the simulation of realistic systems accessible. However, the simulations are still limited to the size of the systems which can be modeled. Simulations in 2D are the rule, and simulations in 3D are still rather rare. Fluid flow is only occasionally included.

Extending these calculations to macroscopic-scale solidification processes requires a different approach. Consider, for example, a cast metal part which has dimensions $250 \times 25 \times 100\,\text{mm}$, and an average grain diameter of $100\,\mu\text{m}$. This object contains more than 10^8 grains. Clearly, it is neither practical nor sensible to model each one.

Instead, microstructure models, such as those presented earlier, should be used to develop scaling laws relating microstructure to process parameters, such as local thermal and flow conditions. Stochastic approaches used to predict nucleation events can be combined with the growth laws to determine the overall microstructure.

The value of such simulations is limited also by available data. Many material properties whose values are not well known appear in the simulations. In particular, the surface tension and attachment kinetics, along with their crystallographic anisotropy, are not well characterized. This topic is amenable to molecular dynamics simulations, which are now being used to address it. This is another example of the hierarchy of length scales in the problem, and how simulations at one length scale feed into those performed at larger ones.

Acknowledgments

This work has been supported by NASA under grant number NAG 8-1657, and by the National Science Foundation under grant number DMR 01-21695. The author has taken examples from work performed in his laboratory, in collaboration with Nigel Goldenfeld, Nikolas Provatas, Jun-Ho Jeong, Anthony Chang, and Badri Athreya. There are many other researchers who have made important contributions to this field, but the selection of the author's own work has been for convenience. The reader is referred to the "Further reading" section.

See also: Computer Simulation Techniques in Condensed Matter Physics; Crystal Growth, Bulk: Methods; Crystal Growth, Bulk: Theory and Models; Mass Transport; Phases and Phase Equilibrium; Phase Transformation; Thermodynamic Properties, General.

PACS: 05.70.Ln; 81.10.Aj; 81.10.Fq; 81.30.Bx; 81.30.Fb; 82.60.Lf

Further Reading

Chen S, Merriman B, Osher S, and Smereka P (1997) *Journal of Computational Physics* 135: 8.

Crank J (1984) *Free and Moving Boundary Problems*. Oxford: Oxford University Press.

Dantzig JA (2004) PDtool: Interactive calculation of binary phase diagrams. via www.http://quattro.me.uiuc.edu/~jon/PDTOOL, University of Illinois.

Dantzig JA and Tucker CL III (2001) *Modeling in Materials Processing*. New York: Cambridge University Press.

Davis SH (2001) *Theory of Solidification*. Cambridge: Cambridge University Press.

Flemings MC (1974) *Solidification Processing*. New York: McGraw-Hill.

Gaskell D (1981) *Introduction to Metallurgical Thermodynamics*. New York: Hemisphere.

Karma A and Rappel WJ (1995) Quantitative phase-field modeling of dendritic growth in two and three dimensions. *Physical Review E* 53: 3017.

Kim YT, Goldenfeld N, and Dantzig J (2000) Computation of dendritic microstructures using a level set method. *Physical Review E* 62: 2471.

Kostorz G (ed.) (2001) *Phase Transformations in Materials*. Weinheim, FRG: Wiley-VCH.

Kurz WF and Fisher DJ (1984) *Fundamentals of Solidification*. Aeder-mannsdorf: Trans-Tech.

Mullins WW and Sekerka RF (1964) Stability of a planar interface during solidification of a dilute binary alloy. *Journal of Applied Physics* 35: 444–451.

Osher S and Fedkiw R (2003) *Level Set Methods and Dynamic Implicit Surfaces*. New York: Springer.

Provatas N, Goldenfeld N, and Dantzig J (1999) Adaptive mesh refinement computation of solidification microstructures using dynamic data structures. *Journal of Computational Physics* 148: 265.

Space Groups

T Janssen, University of Nijmegen, Nijmegen, The Netherlands

Introduction

For the description or the study of properties of physical systems, symmetry is of paramount importance. The symmetry groups of crystals are the so-called space groups. Other symmetry operations, such as time reversal, are sometimes also relevant, but the most important operations are the elements of the space groups. Space group symmetry is a generalization of the property that crystals are, usually, periodic in three dimensions. This article deals with the periodicity, the other symmetry operations of crystals, the classification of the space groups, and, what is of greater relevance for the physical properties, their representations.

Lattice Periodicity

Ideally, the most common crystals consist of a periodic array of identical building blocks repeated in all three directions. There are other types of crystals also, but a discussion of these will be taken up at the end of this section. Real crystals always have shortcomings in their periodic order; they are finite, which implies that there is a limit to the periodicity. Repetition means that one may go from one block into another by a translation. Such a translation is given by a vector $\boldsymbol{n}$. All vectors of the translations form a lattice. That means there are three fundamental translations, with vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, such that each lattice translation $\boldsymbol{n}$ is a linear combination of these three with integer coefficients, and each such combination transforms a building block into another:

$$\boldsymbol{n} = n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c}, \quad \text{with integers } n_1, n_2, n_3 \qquad [1]$$

The building blocks do not have an overlap, and there are no gaps between them. Each building block consists of an arrangement of atoms or molecules. It is called a unit cell. The whole structure then remains the same if it is shifted by any of the translations of the lattice. Suppose that there are N atoms in the building block at positions $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$. Then the position of an arbitrary atom of the crystal is given by the expression

$$\boldsymbol{r}_{n,j} = \boldsymbol{r}_j + n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c} \qquad [2]$$

for specific values of j, n_1, n_2, and n_3. The vectors $\boldsymbol{r}_j$ may be chosen inside the unit cell. The vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are the basis vectors.

Because the building blocks are repeated in all directions, the three fundamental vectors are independent, which also means that every point in space can be reached from a fixed point by a linear combination, not necessarily with integer coefficients. In general, the coefficients are real numbers. In particular, the position of an arbitrary atom with respect to a chosen origin can be written as

$$\boldsymbol{r}_{n,j} = (n_1+\xi_1)\boldsymbol{a} + (n_2+\xi_2)\boldsymbol{b} + (n_3+\xi_3)\boldsymbol{c} \qquad [3]$$

Here $(n_1+\xi_1)$, etc., are real numbers. For example, for CsCl, the three basis vectors are in Cartesian coordinates $(a, 0, 0)$, $(0, a, 0)$, and $(0, 0, a)$. There is a Cs atom at $\boldsymbol{r}_1 = (0, 0, 0)$ and a Cl atom at $\boldsymbol{r}_2 = (a/2, a/2, a/2)$. With respect to the lattice basis, the Cl coordinates are $\xi_1 = \xi_2 = \xi_3 = 1/2$. The unit cell is a cube.

Distance-Preserving Transformations

Physical laws remain the same if positions are transformed into newer positions with the same distances. Therefore, such distance-preserving transformations are important for physics. Among them are translations, where all positions are shifted by the same translation vector. In addition, rotations around an arbitrary point and the full reflection of all points through an arbitrary point leave the distances invariant. Let O be an arbitrary point in space. Then the orthogonal group consists of all rotations leaving this point invariant and all products of such a rotation with the reflection through O. The group of three-dimensional orthogonal transformations is denoted by $O(3)_O$. Choosing O as origin and a basis of space consisting of three basis vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, an orthogonal transformation R corresponds to a matrix via

$$R\boldsymbol{e}_i = \sum_{j=1}^{3} R_{ji}\boldsymbol{e}_j \quad (i = 1, 2, 3) \qquad [4]$$

For a special choice of the basis, namely, with three mutually perpendicular vectors of the same length, the matrices satisfy $\sum_j R_{ij}R_{kj} = \delta_{ik}$ or $RR^{\mathrm{T}} = E$, where T indicates the transpose, and E the unit matrix. Such matrices are called orthogonal. Their determinant is either $+1$ (for rotations) or -1. If one chooses another origin, the orthogonal group around that point gives the same group of matrices, which can then be indicated by $O(3)$.

All distance-preserving transformations can be obtained as a combination of a translation and an orthogonal transformation around an origin O. The effect on a point $\boldsymbol{r}$ of an orthogonal transformation R together with a translation $\boldsymbol{t}$ is

$$\{R|\boldsymbol{t}\}\boldsymbol{r} = R\boldsymbol{r} + \boldsymbol{t} \qquad [5]$$

All these transformations again form a group, the Euclidean group $E(3)$. Its elements are called Euclidean transformations. The product of two such transformations then is the subsequent execution of the two:

$$\begin{aligned}\{R_1|\boldsymbol{t}_1\}\{R_2|\boldsymbol{t}_2\}\boldsymbol{r} &= \{R_1|\boldsymbol{t}_1\}(R_2\boldsymbol{r} + \boldsymbol{t}_2) \\ &= R_1R_2\boldsymbol{r} + R_1\boldsymbol{t}_2 + \boldsymbol{t}_1 \end{aligned} \qquad [6]$$

from which it follows that

$$\{R_1|\boldsymbol{t}_1\}\{R_2|\boldsymbol{t}_2\} = \{R_1R_2|\boldsymbol{t}_1 + R_1\boldsymbol{t}_2\} \qquad [7]$$

In particular, $\{R|\boldsymbol{t}\} = \{E|\boldsymbol{t}\}\{R|\boldsymbol{0}\}$ and $\{R|\boldsymbol{t}\}^{-1} = \{R^{-1}|-R^{-1}\boldsymbol{t}\}$.

A rotation R from $O(3)_O$ leaves the point O invariant. A translation $\boldsymbol{u}$ transfers O to $O+\boldsymbol{u}$ and this point is left invariant by $\{E|\boldsymbol{u}\}\{R|\boldsymbol{0}\}\{E|-\boldsymbol{u}\}$. This means that, the product $\{R|\boldsymbol{u}-R\boldsymbol{u}\}$ belongs to the orthogonal group $O(3)_{O+\boldsymbol{u}}$. In other words, the translation part $\boldsymbol{t}$ of a space group element $\{R|\boldsymbol{t}\}$ may be changed by an origin shift $\boldsymbol{u}$ according to

$$\boldsymbol{t} = \boldsymbol{t} + (E-R)\boldsymbol{u} \qquad [8]$$

The combination of a translation and a rotation around O is then the same as a combination of another translation and a rotation around $O+\boldsymbol{u}$ (see **Figure 1**).

There is a simple matrix formulation for the action of a Euclidean transformation on a point $\boldsymbol{r}$ with coordinates x, y, and z:

$$\{R|\boldsymbol{t}\}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} & t_1 \\ R_{21} & R_{22} & R_{23} & t_2 \\ R_{31} & R_{32} & R_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \qquad [9]$$

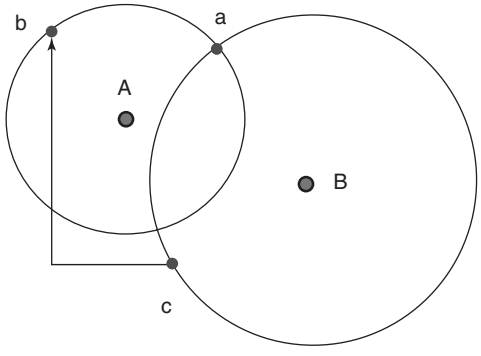


Figure 1 The 90° rotation around A brings a to b, a 90° rotation around B plus a translation brings a via c to b.

Space Groups, Plane Groups, Higher Dimensions

A collection of atoms which is invariant under the translations of a three-dimensional lattice is, in general, invariant under still more Euclidean transformations. A space group is a group of Euclidean transformations that has a translation subgroup as a lattice group, with three linearly independent basis vectors. If $\{E|\boldsymbol{a}\}$ is a translation from the translation subgroup, and $\{R|\boldsymbol{t}\}$ an arbitrary element of the space group, then the following relation holds:

$$\{R|\boldsymbol{t}\}\{E|\boldsymbol{a}\}\{R|\boldsymbol{t}\}^{-1} = \{E|R\boldsymbol{a}\} \qquad [10]$$

The first consequence is that the translation subgroup is an invariant subgroup. (A subgroup A of a group G is invariant if, for each element a of A and each g from G, the element gag^{-1} is an element of A.) Because $\{E|R\boldsymbol{a}\}$ is a translation, and therefore an element of the translation subgroup, $R\boldsymbol{a}$ belongs to the lattice. Consequently, the lattice is invariant under R, and because all R's form a subgroup of $O(3)$, this is a crystallographic point group, one of the 32 point groups. The Euclidean transformations leaving the CsCl structure invariant are the 48 elements of the cubic point group, combined with the lattice translations of the cubic lattice.

A plane group is a subgroup of the Euclidean group in two dimensions having as translation subgroup a two-dimensional lattice group. The elements R now form one of the 10 two-dimensional crystallographic point groups. The generalization to space groups in arbitrary dimensions then is straightforward. Such a group in arbitrary dimensions (including two) is also called a space group.

The Structure of a Space Group

The translation subgroup A of a space group G is invariant. (It is also called a normal subgroup.) The orthogonal transformations R appearing in the elements $\{R|\boldsymbol{t}\}$ form a three-dimensional crystallographic point group K. All the elements of K can now be numbered: $R_1 = E, R_2, \ldots, R_N$. All elements $\{R_i|\boldsymbol{t}\}$ in G form a set S_i for fixed i. One can multiply two such sets: the product S_iS_j is the set S_k if $R_iR_j = R_k$. In this

way, the sets S_i form a group, which is in fact identical to the group of all elements R_i, that is, the point group K. The group of sets S_i is called the factor group, and it is identical to (more precisely isomorphic to) the point group K. So, the space group G has an invariant subgroup A and the factor group G/A is isomorphic to the point group K.

From each set S_i, one may choose an element $\{R_i|\boldsymbol{t}_i\}$. Then, every other element from S_i may be written as the product of the representative $\{R_i|\boldsymbol{t}_i\}$ and a lattice translation $\{E|\boldsymbol{a}\}$. Of course, one could have chosen another representative element $\{R_i|\boldsymbol{t'}_i\}$. Then the two translations $\boldsymbol{t}_i$ and $\boldsymbol{t}_i'$ differ by a translation $\boldsymbol{a}$ from the lattice. The translations $\boldsymbol{t}_i$ are not necessarily lattice vectors, and they are determined by R_i only up to a lattice vector. If the representatives can be chosen such that their translation part is zero, that is, when all translations $\boldsymbol{t}_i$ are lattice vectors, the group has a simple structure: each element is a product of an element of the translation subgroup and an element of the point group. Then the space group is called symmorphic.

Consider as examples the plane groups of the two-dimensional structures for which the unit cell is given in **Figure 2**. The unit cells are rectangles, the lattices have bases vectors $(a, 0)$ and $(0, b)$. The example of **Figure 2a** has two different atoms in (0,0) and (1/2, 1/2). The point group symmetry of the lattice consists of the identity E, the two mirrors m_x and m_y, and the inversion $-E$. All four elements transform the position of an atom to a position that is related to the original position by a lattice vector. Thus, the plane group has a point group ($2mm$) with four elements. The example of **Figure 2b** has one molecule at the position $(0, 0)$. There are only two point group elements leaving the molecule invariant, and the point group is 2 with elements E and $-E$, although the lattice is the same. In example (c), there are molecules at $(0, 0)$ and $(1/2, 1/2)$. The subgroup leaving each molecule invariant has two elements: $\pm E$. The transformation m_x does not leave a molecule invariant, but the Euclidean transformation $\{m_x|(1/2,\ 1/2)\}$ does. The space group elements are the elements of the translation subgroup A and the cosets $\{-E|0\}A$, $\{m_x|(1/2,\ 1/2)\}A$, and $\{m_y|(1/2,\ 1/2)\}A$. The point group is again $2mm$.

Space Group Elements

Each element of a space group is a product of an orthogonal transformation (an element of the point group K) and a translation $\boldsymbol{t}$, the latter not always a lattice translation. The translations $\boldsymbol{t}_i$ form a vector system for the space group. The translations satisfy

$$\boldsymbol{t}_i + R_i\boldsymbol{t}_j = \boldsymbol{t}_k \quad \text{up to a lattice vector } \boldsymbol{a} \quad \text{if } R_iR_j = R_k \qquad [11]$$

The translations $\boldsymbol{t}_i$, however, depend on the choice of the origin. According to eqn [8] they change to $\boldsymbol{t}_i + (E - R_i)\boldsymbol{u}$, if the origin is shifted by a translation $\boldsymbol{u}$. This makes it necessary to review the concept of a symmorphic space group. A space group is called symmorphic if there is an origin such that all representatives $\{R_i|\boldsymbol{t}_i\}$ may be chosen with $\boldsymbol{t}_i = \boldsymbol{0}$. Then each element of G is the product of a point group element and a lattice translation. If the group is not symmorphic, then it is called nonsymmorphic.

A three-dimensional rotation has an axis and a rotation angle. If one chooses an orthonormal coordinate system and the rotation angle along the z-axis, then the rotation is given by a matrix

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

When an orthogonal transformation has a determinant equal to -1, it has the same form but with an overall minus sign. Consider the case that the orthogonal transformation is combined with a translation $\boldsymbol{t}$ with coordinates a, b, and c in the same reference system. By a shift of origin over the vector $\boldsymbol{u}$ with components x, y, and z, the translation is transformed into the translation $\boldsymbol{t} + (1 - R)\boldsymbol{u}$, with components

$$a + x(1 - \cos\phi) + y\sin\phi,$$
$$b - x\sin\phi + y(1 - \cos\phi), \quad c$$

By a proper choice of $\boldsymbol{u}$, the first two components can be eliminated, but not the third. It is the case of a screw axis, a rotation combined with a translation along the rotation axis. If the rotation is over an

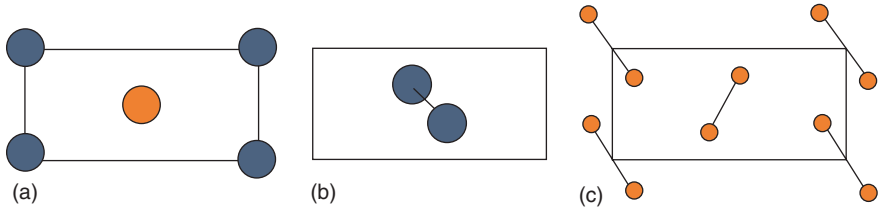


Figure 2 Three unit cells for the space groups (a) *pmm*, (b) *p2*, and (c) *pgg*.

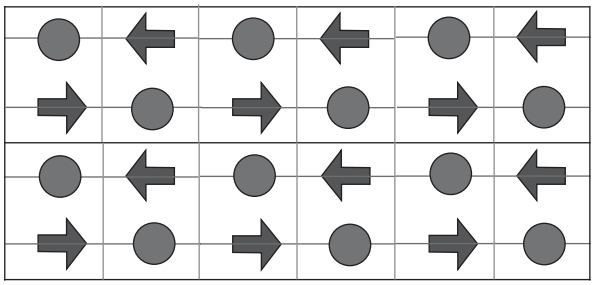


Figure 3 A two-dimensional pattern with *pgm* symmetry: rectangular lattice, horizontal mirrors, and vertical glides.

angle $\phi = 2\pi m/n$, with $n = 2,3,4$, or 6, the translation is always $1/n$ of a lattice vector. When the determinant is -1, the origin can always be shifted such that the translation vanishes, except for the case $n = 2$. Then the orthogonal transformation is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and the first two components cannot be changed. It is a mirror operation with a translation in the mirror plane, and is called a glide operation. The plane of the mirror is the glide plane (**Figure 3**).

In two dimensions, the only plane group elements with intrinsic nonzero translation components are glide operations. For a rotation with a translation, the latter can always be eliminated by a shift of the origin.

Classification

Because the orientation and the lattice constants may vary continuously, the number of space groups is infinite. However, there is good reason to identify space groups under certain conditions, for example, if they are just different orientations of the same group in space. In principle, which groups may be identified depends on the physics of the problem one wants to study by symmetry. Space groups are subgroups of the inhomogeneous affine group as well, the group of pairs of nonsingular linear transformations and translations. If a group G_1 may be transformed into group G_2 by a change of origin (translation) and a linear transformation of the lattice (homogeneous affine transformation), the two may be identified, or in other words considered as equivalent, because the choice of origin and basis does not change the physics. This is a definition inspired by physics. In mathematical language, the two groups are conjugated subgroups of the inhomogeneous affine group. This means that the groups G_1 and G_2 are considered to be equivalent if there is a nonsingular linear transformation S and a translation $\{E|\boldsymbol{t}\}$ such that

$$G_1 = \{S|\boldsymbol{t}\}G_2\{S|\boldsymbol{t}\}^{-1} \qquad [12]$$

With this relation, there are 219 equivalence classes of three-dimensional space groups. If the handedness of the basis is relevant, for instance, in the case of helical structures, there is a finer definition, which calls the groups equivalent if the conjugation is by an element $\{S|\boldsymbol{t}\}$ such that $\det(S) > 0$. Then there are 230 equivalence classes of space groups in three dimensions. With both definitions, there are 17 different plane groups in two dimensions. A theorem by Bieberbach states that conjugation in the affine group is equivalent to isomorphism.

The space groups can also be grouped into larger classes. A coarser classification uses arithmetic equivalence. Choosing an origin and a lattice basis, the point group K of a space group corresponds to a group of integer matrices $D_1(K)$. After a basis transformation, corresponding to an integer nonsingular matrix S, the same group is represented by an integer matrix group $D_2(K) = SD_1(K)S^{-1}$. Two such groups are called arithmetically equivalent. Then a space group determines an arithmetic crystal class. If one drops the condition that the conjugation matrix S has integer entries and allows real matrices, two groups conjugated by S are said to be geometrically equivalent. Arithmetic equivalence implies geometric equivalence. Space groups may be grouped into affine equivalence classes, and further into arithmetic and geometric crystal classes. In three dimensions, there are 219 affine classes, 73 arithmetic classes, and 32 geometric classes. The latter two classifications can, of course, be used for point groups as well. For each arithmetic class there is exactly one symmorphic space group. Finally, there are seven systems (triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral, hexagonal, and cubic) and six families (each three-dimensional system is a family, except the rhombohedral and hexagonal, which belong to the same family). The reader is referred to more specialized works on crystallography for a definition.

Notation

Because the full explanation of the notation and nomenclature for crystallographic groups requires much more space, only a brief discussion is presented here. There are two systems, the Hermann–Mauguin symbols, recommended by the International Union for Crystallography (IUCr), and the Schoenflies symbols. The latter are based on the symbols for the 32 point groups arranged in geometric classes. For each class there is a numbering of the corresponding space

groups. An example is C_2^2 with point group C_2 consisting of an identity and a two-fold rotation. It is the second group with this point group. The Hermann–Mauguin symbols tend to have more information. They are based on the IUCr symbols for point groups. A second ingredient are the lattices. Lattices are considered to be equivalent if the point groups that leave them invariant (which on a lattice basis may be given by groups of integer matrices) may be represented by a change of lattice basis by the same groups of matrices. These classes are the Bravais classes. In three dimensions, there are 14 Bravais classes (**Table 1**). For lattices with the same point group symmetry, there is a common conventional basis on which the point group elements are simple, but one needs additional lattice vectors. For example, there are three Bravais classes for which the symmetry is the symmetry group of the cube, with 48 elements. The lattices of one Bravais class have a basis with three mutually perpendicular basis vectors of the same lengths. For the other two, the face-centered cubic (f.c.c.) and body-centered cubic (b.c.c.), there is a sublattice of this kind, but not all lattice vectors belong to that. One needs additional vectors to obtain all the lattice translations: $(0, 1/2, 1/2)$, $(1/2, 0, 1/2)$, and $(1/2, 1/2, 0)$ for f.c.c. and $(1/2, 1/2, 1/2)$ for b.c.c (**Figure 4**). These additional vectors are given by capital letters in three dimensions, and lower case letters in two dimensions. The symbols of the symmetry groups of the three cubic lattices are $Pm\bar{3}m$, $Fm\bar{3}m$, and $Im\bar{3}m$. These are the Hermann–Mauguin symbols for the space groups of the primitive cubic, f.c.c., and b.c.c. lattices, respectively. For each of the 73 arithmetic crystal classes, and thus for all symmorphic space groups there is such a symbol, consisting of the symbol for the point group preceded by a letter indicating the additional lattice vectors (the centering).

Finally, for nonsymmorphic space groups, the nonlattice translations for the elements appearing in the symbol for the symmorphic space group are

Table 1 The 14 Bravais classes in three dimensions and their maximal symmorphic space groups

System	*Centering translations*	*Maximal symmorphic space group*	*Basis*
Triclinic	None	$P\bar{1}$	No relations
Monoclinic	None	$P2/m$	$\boldsymbol{a}.\boldsymbol{c} = \boldsymbol{b}.\boldsymbol{c} = 0$
	$(1/2,\ 0,\ 1/2)$	$B2/m$	
Orthorhombic	None	$Pmmm$	Three axes perpendicular
	$(1/2,\ 1/2,\ 1/2)$	$Immm$	
	$(0,\ 1/2,\ 1/2)$, $(1/2,\ 0,\ 1/2)$, $(1/2,\ 1/2,\ 0)$	$Fmmm$	
	$(1/2,\ 1/2,\ 0)$	$Cmmm$	
Tetragonal	None	$P4/mmm$	Three axes perpendicular
	$(1/2,\ 1/2,\ 1/2)$	$I4/mmm$	$\lvert\boldsymbol{a}\rvert = \lvert\boldsymbol{b}\rvert$
Rhombohedral	None	$R\bar{3}m$	$\lvert\boldsymbol{a}\rvert = \lvert\boldsymbol{b}\rvert = \lvert\boldsymbol{c}\rvert$ $\angle(\boldsymbol{a}, \boldsymbol{b}) = \angle(\boldsymbol{b}, \boldsymbol{c}) = \angle(\boldsymbol{c}, \boldsymbol{a})$
Hexagonal	None	$P6/mmm$	$\lvert\boldsymbol{a}\rvert = \lvert\boldsymbol{b}\rvert$, $\boldsymbol{c} \perp \boldsymbol{a}$, $\boldsymbol{c} \perp \boldsymbol{b}$ $\angle(\boldsymbol{a}, \boldsymbol{b}) = 2\pi/3$
Cubic	None	$Pm\bar{3}m$	All basis vectors mutually perpendicular and equal in length
	$(0,\ 1/2,\ 1/2)$, $(1/2,\ 0,\ 1/2)$, $(1/2,\ 1/2,\ 0)$	$Fm\bar{3}m$	
	$(1/2,\ 1/2,\ 1/2)$	$Im\bar{3}m$	

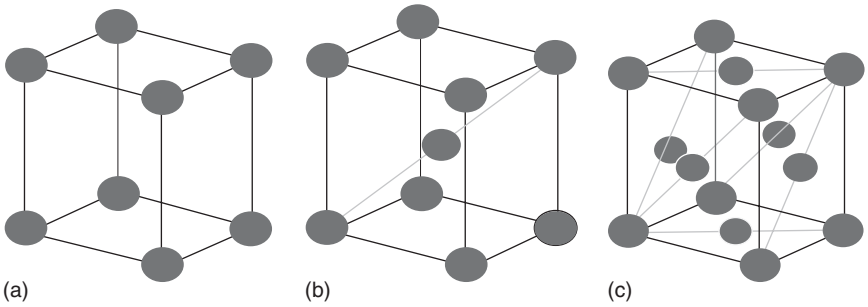


Figure 4 Unit cells for lattices from each of the three cubic Bravais classes. (a) Primitive, (b) b.c.c. with additional basis vector (1/2, 1/2, 1/2), and (c) f.c.c. with additional (0, 1/2, 1/2), (1/2, 0, 1/2), and (1/2, 1/2, 0).

indicated by either a change of the letter or by a subindex. The symbol for a symmorphic space group with point group m may be Pm. If the mirror m becomes a glide with nonlattice translation in the third direction, the symbol is Pc. The symbol $P2$ gives a symmorphic space group with point group 2, $P2_1$ indicates a nonsymmorphic group with a screw axis. All space groups, their symbols and their elements can be found in the *International Tables for Crystallography, vol. A*.

Reciprocal Lattice, Invariant Functions, Extinction Rules

A function that is invariant under translations has special properties for its Fourier transform. Suppose $\rho(\boldsymbol{r})$ is such a function (e.g., a density function of a crystal). It has the property that $\rho(\boldsymbol{r})$ and $\rho(\boldsymbol{r}+\boldsymbol{a})$ are equal for every translation $\boldsymbol{a}$ from the lattice. Writing down the Fourier decomposition yields

$$\begin{aligned}\rho(\boldsymbol{r}) &= \int \hat{\rho}(\boldsymbol{k})\, \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\mathrm{d}\boldsymbol{k} \\ \rho(\boldsymbol{r}+\boldsymbol{a}) &= \int \hat{\rho}(\boldsymbol{k})\, \exp(\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}+\boldsymbol{a}))\mathrm{d}\boldsymbol{k}\end{aligned} \quad [13]$$

from which it follows that $\boldsymbol{k}\cdot\boldsymbol{a} = 0 \pmod{2\pi}$ for each translation vector $\boldsymbol{a}$. Therefore, for each wave vector $\boldsymbol{k}$ occurring in the Fourier expansion, one has this relation. All vectors $\boldsymbol{k}$ satisfying this relation form a lattice, as is easily checked. This lattice is called the reciprocal lattice for the so-called direct lattice of the vectors $\boldsymbol{a}$.

If $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ span the direct lattice, a basis for the reciprocal lattice is givenby $\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*$ defined by

$$\begin{aligned}&\boldsymbol{a}^* = 2\pi(\boldsymbol{b}\times\boldsymbol{c})/V, \quad \boldsymbol{b}^* = 2\pi(\boldsymbol{c}\times\boldsymbol{a})/V, \\ &\boldsymbol{c}^* = 2\pi(\boldsymbol{a}\times\boldsymbol{b})/V\end{aligned} \quad [14]$$

where $V = \boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})$ is the volume of the unit cell. Then the Fourier decomposition is

$$\rho(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in\Lambda^*} \hat{\rho}(\boldsymbol{k})\, \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) \quad [15]$$

where Λ^* is the reciprocal lattice. Each vector of the reciprocal lattice can be expressed in terms of the basis of the lattice

$$\boldsymbol{k} = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^* \in \Lambda^* \quad [16]$$

The reciprocal lattice is left invariant by the point group K as well.

Apart from translations, a lattice periodic function is generally left invariant by other distance-preserving transformations as well. Suppose $g = \{R|\boldsymbol{t}\}$ is an element of the space group that leaves a function $\rho(\boldsymbol{r})$ invariant. This means that $\rho(\boldsymbol{r})$ and $\rho(g^{-1}\boldsymbol{r})$ are the same (the exponent -1 is just for convenience). Then one has the relation

$$\rho(\boldsymbol{r}) = \rho(\{R|\boldsymbol{t}\}^{-1}\boldsymbol{r}) \quad [17]$$

For the Fourier components $\hat{\rho}(\boldsymbol{k})$, this implies

$$\hat{\rho}(\boldsymbol{k}) = \hat{\rho}(R\boldsymbol{k})\, \exp(\mathrm{i}R\boldsymbol{k}\cdot\boldsymbol{t}) \quad [18]$$

This is a very interesting formula. It is known that the wave vectors occurring in the Fourier decomposition belong to the reciprocal lattice. Now, consider a reciprocal lattice vector $\boldsymbol{k}$ that is invariant under orthogonal transformation R. Then the expression becomes $\rho(\boldsymbol{k}) = \rho(\boldsymbol{k})\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t})$. This is only possible if $\rho(\boldsymbol{k}) = 0$ or if $\boldsymbol{k}\cdot\boldsymbol{t} = 0 \pmod{2\pi}$. If $\boldsymbol{t}$ is a translation from the lattice, this relation is trivially fulfilled. But if there is a $\boldsymbol{t}$ such that the second relation is not satisfied, this means that the corresponding Fourier component $\hat{\rho}(\boldsymbol{k})$ vanishes. This is an extinction rule.

An important example is the diffraction intensity. A crystal diffracts and the diffraction pattern consists of sharp Bragg peaks at the positions $\boldsymbol{k}$. The intensity of the peaks is given by the square of the absolute value of the static structure factor:

$$I(\boldsymbol{k}) = |F(\boldsymbol{k})|^2 = \left|\frac{1}{N}\sum_{j=1}^{N} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j)\right|^2 \quad [19]$$

Here j runs over the N particles in the unit cell. It is the Fourier transform of the autocorrelation function, which is invariant under space group elements. Hence, the positions of the Bragg peaks are on the reciprocal lattice, and the extinction rules apply: the intensity is zero at $\boldsymbol{k}$ if there is a space group element $\{R|\boldsymbol{t}\}$ with $R\boldsymbol{k} = \boldsymbol{k}$ and $\boldsymbol{k}\cdot\boldsymbol{t} \neq 0 \pmod{2\pi}$. Additionally, one has for arbitrary $\boldsymbol{k}$ that $I(R\boldsymbol{k}) = I(\boldsymbol{k})$. Consequently, the diffraction has the point group in its symmetry group.

Representations of Space Groups

Euclidean transformations act on positions in space. A quantum mechanical system in such a space has states on which the Euclidean transformations act as linear operators. For example, if the state is given by a wave function $\psi(\boldsymbol{r})$, the effect of a Euclidean transformation $\{R|\boldsymbol{t}\}$ gives a new wave function $\psi'(\boldsymbol{r}) = \psi(R^{-1}(\boldsymbol{r}-\boldsymbol{t}))$. This is the action of a linear operator T_g, with $g = \{R|\boldsymbol{t}\}$. Choosing a basis $\psi(\boldsymbol{r})$ in the space of states, such a linear operator corresponds to a matrix $D(g)$:

$$T_g\psi_i(\boldsymbol{r}) = \sum_{j=1}^{N} D(g)_{ji}\psi_j(\boldsymbol{r}) \quad [20]$$

(n is the dimension of the state space.) Now the matrices satisfy $D(g_1)D(g_2) = D(g_1g_2)$. This is called a (matrix) representation. If the dimension of the space is n, then the representation is said to be n-dimensional, and the matrices are $n \times n$. These representations are important for characterizing energy levels and other properties. Clearly, the space is mapped onto itself and, therefore, it is invariant under the group of transformations G. If there is no subspace (different from the origin or the whole space) that is invariant, the representation is said to be irreducible. In general, an energy level space carries an irreducible representation of the symmetry group, the dimension of the space is the level degeneracy, and there are orthogonality relations between states from the state space.

This gives a short argument why one should look at the irreducible representations of the space groups. The simplest space group is just a translation group, with three basis translations. Because the order in which translations are applied is not relevant, it is a commutative group, and according to the results of group theory the irreducible representations of commutative groups are one-dimensional, that is the matrices are just numbers. They satisfy $D(\{E|\boldsymbol{a}\})$ $D(\{E|\boldsymbol{b}\}) = D(\{E|\boldsymbol{a}+\boldsymbol{b}\})$. The solution is $D(\{E|\boldsymbol{a}\}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a})$ for some vector $\boldsymbol{k}$. Two vectors $\boldsymbol{k}$ and $\boldsymbol{k}'$ give the same representation if $\boldsymbol{k}-\boldsymbol{k}'$ belongs to the reciprocal lattice, because then $\boldsymbol{k}\cdot\boldsymbol{a} = \boldsymbol{k}'\cdot\boldsymbol{a} \pmod{2\pi}$ for all lattice vectors $\boldsymbol{a}$. Therefore, the irreducible representations are characterized by a vector from the unit cell of the reciprocal lattice. A special choice of this unit cell is the Brillouin zone. It is the unit cell of the reciprocal lattice consisting of all points in reciprocal space which are closer to the origin than to any other point of the reciprocal lattice. Of course, the unit cell of a lattice is not uniquely defined. However, the choice of the Brillouin zone is very convenient when studying electrons or elementary excitations in crystals.

A wave function $\psi(\boldsymbol{r})$ in a lattice periodic crystal which belongs to an irreducible representation of the lattice subgroup has a special form, the Bloch form, as follows from a group-theoretical argument. Under a translation $\boldsymbol{a}$, the function transforms to $\psi(\boldsymbol{r}-\boldsymbol{a})$ which should be equal to $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a})\psi(\boldsymbol{r})$ for some $\boldsymbol{k}$. Then define $U(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\psi(\boldsymbol{r})$. It follows that

$$\begin{aligned} U(\boldsymbol{r}+\boldsymbol{a}) &= \exp(\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}+\boldsymbol{a}))\psi(\boldsymbol{r}+\boldsymbol{a}) \\ &= \exp(\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}+\boldsymbol{a}))\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a})\psi(\boldsymbol{r}) \\ &= U(\boldsymbol{r}) \end{aligned}$$

which proves that $\psi(\boldsymbol{r})$ is the product of a plane wave, $\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$ and a lattice periodic function $U(\boldsymbol{r})$. This constitutes the well-known Bloch's theorem.

Consider a state space that is invariant under a space group acting on the space by operators T_g for each element g of the space group G. Suppose, furthermore, that the representation is irreducible. The translation subgroup A is a subgroup of G. So, there is a basis formed by functions ψ_i for the space consisting of the eigenvectors of the commuting operators T_a. One has

$$T_a\psi_j = \exp(\mathrm{i}\boldsymbol{k}_j\cdot\boldsymbol{a})\psi_j \qquad [21]$$

A subspace of the space is that belonging to one particular wave vector $\boldsymbol{k}$. For each state, ψ from this space it holds that $T_a\psi = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a})\psi$. Applying T_g to any state from this space, ψ transforms to $T_g\psi$. How does this transform under the translations? Applying a translation $\{E|\boldsymbol{a}\}$ to the transformed vector:

$$\begin{aligned} T_aT_g\psi &= T_g(T_{g-1}T_aT_g)\psi = T_gT_b\psi \\ &= \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{b})T_g\psi \end{aligned}$$

Here $\boldsymbol{b} = R^{-1}\boldsymbol{a}$ if $g = \{R|\boldsymbol{t}\}$. The state $T_g\psi$, therefore, acquires a phase factor $\exp(\mathrm{i}R\boldsymbol{k}\cdot\boldsymbol{a})$ and belongs to the subspace of states transforming with a wave vector $R\boldsymbol{k}$. This means that, if there are states transforming under translations with a wave vector $\boldsymbol{k}$, then there are, in the irreducible space, states transforming with the wave vector $R\boldsymbol{k}$ for every point group element. One can write the whole space as a sum of spaces each belonging to a specific $R\boldsymbol{k}$. One of these spaces is that of states transforming with $\boldsymbol{k}$.

Now define the subgroup of G consisting of all elements $\{R|\boldsymbol{t}\}$ for which $R\boldsymbol{k} = \boldsymbol{k}$ (modulo the reciprocal lattice because $\boldsymbol{k}$ and $k+\boldsymbol{K}$ give the same representation if $\boldsymbol{K}$ belongs to the reciprocal lattice). This group is called the group of $\boldsymbol{k}$: $G_{\boldsymbol{k}}$. The subspace of states transforming with $\boldsymbol{k}$ under translations is invariant under the group of $\boldsymbol{k}$. Then for an element g from the group of $\boldsymbol{k}$:

$$T_g = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t})D_R \quad (g = \{R|\boldsymbol{t}\})$$

The operators form an irreducible representation of the group of $\boldsymbol{k}$. For a symmorphic group of $\boldsymbol{k}$ the elements g can be written as a product $\{E|\boldsymbol{t}\}\{R|0\}$ of a lattice translation and a point group element. In this case, the operators D_R form an irreducible representation of the point group of the group of $\boldsymbol{k}$, denoted by $K_{\boldsymbol{k}}$. These are known, and can be labeled by a label ν. Then a basis of the space of $G_{\boldsymbol{k}}$ transforms under g as

$$T_g\psi_j = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t})D_{kj}(R)\psi_k \qquad [22]$$

The basis carries an irreducible representation ν of the point group K (simply take $\boldsymbol{t}=0$ in the last formula). Here the matrices $D(R)$ are the irreducible representation ν of the point group $K_{\boldsymbol{k}}$. For nonsymmorphic groups, the procedure is somewhat

more complicated, because not only standard representations but projective representations also occur.

Finally, a basis for the full state space can be constructed as follows. The group of $\boldsymbol{k}$ is a subgroup of the space group G. G can be decomposed according to

$$G = G_{\boldsymbol{k}} + g_2 G_{\boldsymbol{k}} + \cdots + g_s G_{\boldsymbol{k}}$$

where the space group elements g_i have homogeneous parts R_i for which $R_i\boldsymbol{k} = \boldsymbol{k}_i$. Then the basis is defined as

$$\Psi_{ij} = T_{gi}\psi_j \qquad [23]$$

The dimension of the representation is sd, where s is the number of points $\boldsymbol{k}_i$, and d the dimension of the point group representation $D(K_{\boldsymbol{k}})$. The irreducible representation carried by the state space then is characterised by the so-called "star" of $\boldsymbol{k}$ (all vectors $\boldsymbol{k}_i$), and an irreducible representation of the point group $K_{\boldsymbol{k}}$. This means that electronic states and phonons can be characterized by $\boldsymbol{k}$, v. Their transformation properties under space group transformations follows from this characterization.

Aperiodic Crystals

Apart from crystals with three-dimensional lattice periodicity, there are materials with a diffraction pattern with sharp Bragg peaks on positions

$$\boldsymbol{k} = \sum_{i=1}^{n} h_i \boldsymbol{a}_i^* \quad (\text{integer } h_i) \qquad [24]$$

When $n = 3$, the structure is periodic. If $n > 3$, the structure is aperiodic, but it is still considered as crystal, because there is long-range order. Examples are modulated phases and quasicrystals. They may be described as intersections of physical space with a higher-dimensional lattice periodic structure. The symmetry of such structures is a space group in n dimensions, and in this case the theory of space groups in arbitrary dimensions can be used.

See also: Crystal Structure; Electron–Phonon Interactions and the Response of Polarons; Group Theory; Insulators, Electronic States of; Lattice Dynamics: Vibrational Modes; Periodicity and Lattices; Point Groups; Quantum Mechanics: Foundations; Quasicrystals; Scattering, Elastic (General).

PACS: 61.50.Ah; 02.30. – a

Further Reading

Cornwell JF (1997) *Group Theory in Physics*. San Diego: Academic Press.

Hahn Th. (ed.) (1992) *Space-Group Symmetry*. In: International Tables for Crystallography. vol. A, Dordrecht: Kluwer.

Hahn T and Wondratschek H (1994) *Symmetry of Crystals: Introduction to International Tables for Crystallography. vol. A*. Sofia, Bulgaria: Heron Press.

Janssen T (1973) *Crystallographic Groups*. North-Holland: Amsterdam.

Janssen T, Janner A, Looijenga-Vos A, and de Wolff PM (1999) *Incommensurate and commensurate modulated structures*. In: Wilson AJC and Prince E, *International Tables for Crystallography, vol. C, Mathematical, physical and chemical tables*. ch. 9.8, Dordrecht: Kluwer.

Nomenclature

$\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$	lattice basis vectors
$\boldsymbol{a}^*$, $\boldsymbol{b}^*$, $\boldsymbol{c}^*$	reciprocal lattice basis vectors
A	translation subgroup of a space group
$D(R)$	matrix representation
$\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})U(\boldsymbol{r})$	Bloch form of a wave function
$E(3)$	Euclidean group
$F(\boldsymbol{k})$	structure factor
G	space group
K	point group
$\boldsymbol{n}$	lattice translation vector
$O(3)$	orthogonal group in three dimensions
$\boldsymbol{r}_j$	position of an atom in the unit cell
R	orthogonal transformation
$\{R\|\boldsymbol{t}\}$	space group element
T_g	linear operator for group element g:
$\hat{\rho}(\boldsymbol{k})$	Fourier component of $\rho(\boldsymbol{r})$
$\rho(\boldsymbol{r})$	density function
ψ_i	basis of a state space
$\psi(\boldsymbol{r})$	wave function

Specific Heat

N E Phillips and R A Fisher, University of California at Berkeley, Berkeley, CA, USA

Introduction

The specific heat of a substance is the amount of heat required to raise the temperature by one degree. When heat is introduced under certain specified conditions, it is a well-defined thermodynamic property that gives a measure of the increases in the entropy, the energy, and the enthalpy with increasing temperature. It is related to other thermodynamic properties, for example, the thermal expansion, which is a measure of the pressure dependence of the entropy. Specific-heat data make an important contribution to

the determination of the Gibbs and Helmholtz energies, the thermodynamic properties that govern the direction of spontaneous change and the equilibrium conditions for chemical reactions and phase transitions. When interpreted in terms of microscopic models and theories, they provide information on the forces and interactions at the atomic and molecular level that determine the macroscopic properties.

Thermodynamic Relations

The specific heat of a substance is the amount of heat required to increase the temperature by one degree:

$$C \equiv \frac{\delta q}{\mathrm{d}T}$$

where δq, an inexact differential, is the quantity of heat added and $\mathrm{d}T$ is the increase in temperature produced in the process. By the first law of thermodynamics, $\delta q = \mathrm{d}U - \delta w$, where U is the internal energy and δw is the work done in the process. Since U is a "thermodynamic property," a quantity that depends only on the thermodynamic state of the system, $\mathrm{d}U$ is an exact differential and depends only on the initial and final states. However, δw, and therefore both δq and C, depend on details of the way the process takes place as well as on the initial and final states. If the work is done by an external hydrostatic pressure, P, $\delta w = -P\ \mathrm{d}V$, where V is the volume. In that case,

$$C = \frac{\mathrm{d}U + P\ \mathrm{d}V}{\mathrm{d}T} = \frac{\mathrm{d}(U + PV) - V\ \mathrm{d}P}{\mathrm{d}T}$$

With the definition for the enthalpy, $H \equiv U + PV$, this leads to expressions for the constant volume and constant pressure specific heats,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V$$

and

$$C_P = \left(\frac{\partial H}{\partial T}\right)_P$$

thermodynamic properties, which can be expressed in, for example, units of $\mathrm{J\,K^{-1}\,mol^{-1}}$. If work is done by magnetic or electric forces, the relevant variables are the magnetic induction, $\boldsymbol{B}$, the magnetization, $\boldsymbol{M}$, the electric field, $\boldsymbol{E}$, and the dielectric displacement, $\boldsymbol{D}$. (In free space $\boldsymbol{B} = \mu_0 \boldsymbol{H}$, where μ_0 is the permeability of free space and $\boldsymbol{H}$ is the magnetic field.) The expressions for the work are $\delta w = \boldsymbol{H}\ \mathrm{d}\boldsymbol{M}$ and $\delta w = \boldsymbol{E}\ \mathrm{d}\boldsymbol{D}$, respectively. In other thermodynamic relations, $\boldsymbol{H}$ or $\boldsymbol{E}$ replaces the intensive variable P and $-\boldsymbol{M}$ or $-\boldsymbol{D}$ replaces the extensive variable V.

With other thermodynamic relations, based on both the first and second laws, it can be shown that

$$C_P - C_V = \frac{TV\alpha^2}{\kappa}$$

where $\alpha = V^{-1}(\partial V/\partial T)_P$ is the thermal expansion, and $\kappa = V^{-1}(\partial V/\partial P)_T$ is the isothermal compressibility. Thermodynamic stability requires that both C_V and κ be positive. Although α can be either positive or negative, it appears as the square, and $C_P - C_V$ is always positive: when the heat is absorbed at constant volume, $\delta w = 0$. In the constant-pressure process there is, generally, a change in V, $\delta w \neq 0$ and can be either positive or negative, but the increase in U ensures that the heat absorbed is always greater than in the constant-volume process. Evaluation of α and κ for an ideal gas, for which $PV = RT$, where R is the gas constant, gives $C_P - C_V = R$.

Microscopic Interpretation

Classical, High-Temperature Limit

In classical statistical mechanics, each term in the Hamiltonian for the total energy that is the square of either a momentum or a coordinate, contributes $(1/2)k_\mathrm{B}T$ to the thermal energy, where k_B is the Boltzmann constant. For a particle moving freely in space, the translational kinetic energy includes three terms in the square of a momentum, corresponding to the three dimensions in space: $U = (3/2)k_\mathrm{B}T$, and $C_V = (3/2)k_\mathrm{B}$. For Avagadro's number, N_A, of such particles, one mole of an "ideal gas," $C_V = N_A(3/2)k_\mathrm{B} = (3/2)R$. For a particle bound by harmonic forces to a lattice site in three dimensions, there are three terms in the square of a momentum in the kinetic energy and three in the square of a coordinate in the potential energy: $U = 3k_\mathrm{B}T$, and $C_V = 3k_\mathrm{B}$. At sufficiently high temperatures, quantum statistical mechanics gives the same results, the classical or high-temperature limit. At temperatures for which $k_\mathrm{B}T$ is of the order of, or smaller than, the spacing of the quantum mechanically allowed energy levels, the higher energy levels are not fully accessible, and U and C do not reach the classical limit. The energies of the allowed translational states of the particles of an ideal gas are proportional to $m^{-1/2}\,V^{2/3}$, where m is the mass of the particles and V is the volume of the container. For gases of atoms and molecules at ordinary densities, the energy levels are so closely spaced that the classical limit applies at all temperatures of interest, but for an electron gas at densities of the conduction electrons in a metal, that limit is reached only at temperatures $\sim 10^5\,\mathrm{K}$. The specific heat of a harmonic oscillator offers a typical

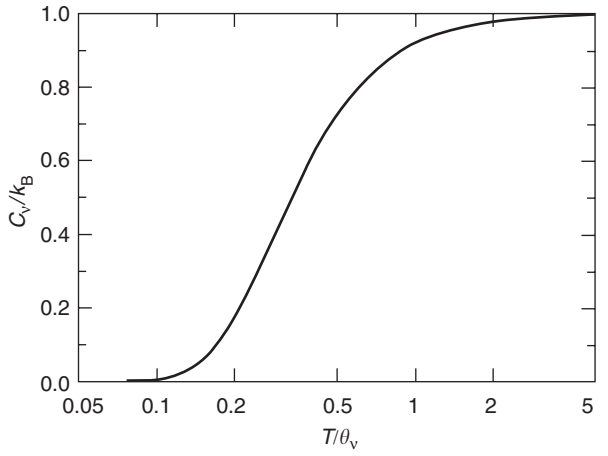


Figure 1 The contribution to the specific heat of a one-dimensional harmonic oscillator, C_v, displayed as C_v/k_B, vs. T/θ_v.

example of the effects of quantization: the energy levels are equally spaced, at intervals of $\hbar\omega \equiv k_B\theta_v$, where $\hbar$ is Planck's constant divided by 2π, T is the characteristic frequency, and θ_v is a defined characteristic temperature. As shown in **Figure 1**, the specific heat approaches k_B in the high-temperature limit and goes to zero, as $\exp(-\theta_v/T)$ at temperatures for which k_BT is less than the energy of the first excited level.

Gases

The heat capacities of gases are often treated in the ideal-gas, rigid-rotor, harmonic-oscillator approximation, in which C is the sum of contributions, C_t, C_r, and C_v, associated with translational, rotational, and vibrational motion, respectively. At sufficiently high temperatures, excited electronic states can also contribute to C_V. For a molecule with n atoms, $3n$ coordinates are required to specify the positions of the atoms: there are $3n$ "degrees of freedom." Three of these are associated with translation of the molecule as a whole; three are associated with rotation for nonlinear molecules, but only two for linear molecules (there is no rotation about the axis); the rest, $3n-6$ or $3n-5$, respectively, are associated with vibration. Each translational degree of freedom contributes one square term to the kinetic energy; each rotational degree of freedom contributes one square term, in an angular momentum, to the kinetic energy; each vibrational degree of freedom contributes one square term to the kinetic energy and one to the potential energy. In the high-temperature limit, and for one mole of gas, $C_v = 3nR$ for nonlinear molecules, but $(3n+1/2)R$ for linear molecules. The vibrational contributions decrease to zero at temperatures below the characteristic temperatures θ_v, as shown in **Figure 1**. The rotational contributions go to zero below characteristic temperatures θ_r that are defined by the allowed rotational energy levels, but there is no general expression for C_r in closed form. Some representative values of θ_v and θ_r are given in **Table 1**. **Figure 2** shows the temperature dependence of C_V for HD, an asymmetric, linear, diatomic molecule. The local maximum near 50 K is related to details of the spacing and degeneracy of the rotational levels. For symmetric molecules, certain combinations of rotational and nuclear spin states are not

Table 1 Characteristic temperatures associated with the vibrational and rotational specific heats of some diatomic molecules and one linear, symmetric, triatomic molecule, CO_2

Molecule	θ_v(K)	θ_v(K)[a]
H_2	6332	88
D_2	4487	43.8
HD	5500	66
HF	5955	30.2
HCl	4304	15.2
HBr	3812	12.2
HI	3321	9.4
N_2	3393	2.88
O_2	2274	2.08
F_2	1283	1.27
Cl_2	805	0.351
Br_2	463	0.116
CO	3122	2.78
NO	2719	2.45
CO_2	3380[b]	1.13
	1928[c]	
	960[d]	

[a] All θ_r are doubly degenerate.
[b] Antisymmetric stretch.
[c] Symmetric stretch.
[d] Perpendicular bending (doubly degenerate).

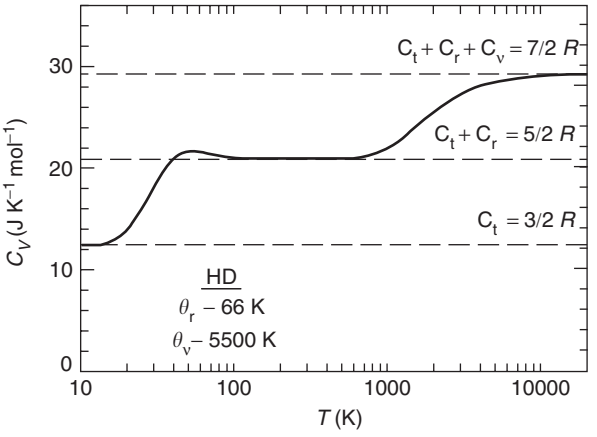


Figure 2 The specific heat at constant volume, C_V, of HD gas. C_V is the sum of translational, rotational, and vibrational contributions, C_t, C_r, and C_v, respectively. The horizontal dashed lines mark the high-temperature limits of the different contributions. (The rotational contribution was obtained from tabulations in: Mayer JE and Mayer MG (1940) *Statistical Mechanics*, p. 450. New York: Wiley.)

allowed. The prototypical examples are o-H_2, with triplet spin and even rotational states, and p-H_2, with singlet spin and odd rotational states. Depending on conditions, transitions between the two spin states can be very slow; o-H_2 and p-H_2 can exist as separate molecules with very different specific heats.

Condensed Matter

Separability of contributions The specific heat of a condensed matter phase is usually taken to be the sum of contributions arising from different modes of excitation. For a nonmagnetic metal these are the lattice, C_l, and electron, C_e, contributions that are associated with, respectively, the vibrational states of the ions that form the lattice and the energy states of the conduction electrons. The vibrational states of the lattice and the energy states of the electrons are not independent, however: The frequencies of the normal modes of the lattice are affected by the presence of the conduction electrons, and the allowed energy states of the electrons are affected by the spatially periodic potential of positive ions of the lattice and also by interaction with the quanta of vibrational energy of the lattice, the phonons. Nevertheless, to an adequate degree of approximation for most purposes, the specific heat can be taken to be the sum of separable contributions from the lattice and the conduction electrons, with the understanding that the energy levels of each system are determined in part by interaction with the other. In experimental data, the two contributions are generally separated by assigning them different, theoretically expected temperature dependences. Similar considerations apply to magnetic contributions, C_m, associated with electronic magnetic moments, and hyperfine contributions, C_h, associated with nuclear moments.

Phonon contribution For one mole of a "monatomic" solid, which has one atom per primitive cell, there are $3N_A$ vibrational degrees of freedom. In the harmonic-lattice approximation, the phonon or lattice contribution to the specific heat, C_l, is the sum of $3N_A$ harmonic-oscillator contributions, one for each of the normal vibrational modes. In the high-temperature limit, also called the Dulong–Petit limit in this context, each contributes k_B to C_l, giving $C_l = 3R$. Anharmonic effects make a contribution to C_l that becomes more important at high temperatures, and can give values of C_l that exceed the Dulong–Petit limit.

In the Einstein model, the normal modes were approximated by $3N_A$ harmonic oscillators with the same frequency, ω_E, corresponding to the characteristic, Einstein temperature, $\Theta_E \equiv \hbar\omega_E/k_B$. The Einstein expression for the specific heat is

$$\frac{C_E}{3R} = \left(\frac{\Theta_E}{T}\right)^2 \frac{\exp(\Theta_E/T)}{[\exp(\Theta_E/T) - 1]^2}$$

This model is of historical importance because it was the first to account for the decrease of C_l at low temperature, but it gives $C_l \propto \exp(-\Theta_E/T)$ in the low-temperature limit, while the experimentally observed behavior is $C_l \propto T^3$.

In the Debye model, the frequencies of the normal modes were approximated by those of the acoustic modes in an elastic continuum. A cutoff frequency, ω_D, was introduced to limit the number of modes to $3N_A$. The cutoff frequency, and the Debye temperature, $\Theta_D \equiv \hbar\omega_D/k_B$, can be calculated from the elastic constants. The Debye specific heat, given by an integral over the phonon density of states,

$$\frac{C_D}{3R} = 3\left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^4 e^x}{(e^x - 1)^2}\,dx$$

is a function of T/Θ_D. In the low-temperature limit, $C_D = (12/5)\pi^4 R(T/\Theta_D)^3$, in agreement with experiment with respect to both the T^3 temperature dependence and the relation of Θ_D to the elastic constants. **Figure 3** is a comparison of the Debye model with the lattice specific heat of Cu. The best fit to the data for $10 \leqslant T \leqslant 300\,K$ is obtained with

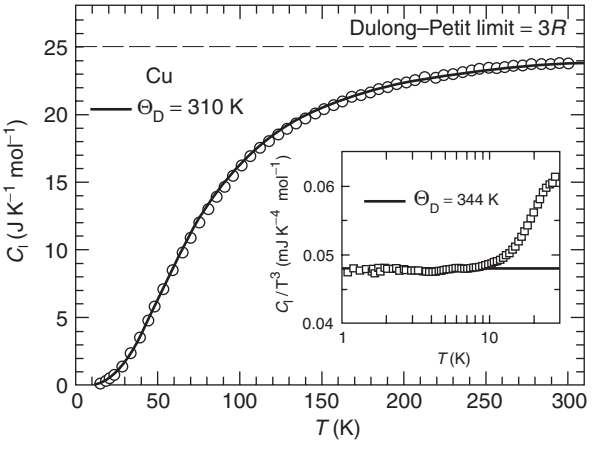


Figure 3 The lattice specific heat, C_l, of Cu: compared with the Debye model for $10 \leqslant T \leqslant 300$ K in the main panel; compared with the low-temperature, T^3 behavior predicted by the Debye model in the inset. The electron contribution to C_V has been subtracted to obtain C_l as plotted here. The data for the main panel are from: Giauque WF and Meads PF (1941) The heat capacities and entropies of aluminum and copper from 15 to 300 K. *Journal of the American Chemical Society* 63: 1897. The measurements were made at constant pressure, but corrected to C_V with the values of α and κ. (The low-temperature data in the inset are from measurements (unpublished) made in the authors' laboratory at the University of California, Berkeley, California 94720, USA.)

$\Theta_D = 310$ K, but the data in the low-temperature, T^3 region gives $\Theta_D = 344$ K, in agreement with the value calculated from the elastic constants. This comparison of data for a monatomic metal with the Debye model is fairly typical: at low temperatures, the Debye model gives the T^3 term, the contribution of the lowest-frequency acoustic phonons, correctly. Although the Debye model gives T^3 behavior for $T \leqslant \Theta_D/10$, and negative deviations from T^3 behavior at higher temperatures, experimental data typically show deviations that are positive and occur at lower temperatures. The negative deviations given by the model are a consequence of the sharp cutoff frequency; the positive deviations in a real solid are a consequence of phonon dispersion, which is neglected in the model. At higher temperatures, where the model is an approximation, experimental data cannot be fitted accurately with a single value of Θ_D. The value determined by the "best fit" in some interval of temperature is an average that reflects the structure in the phonon density of states for the real lattice at frequencies that contribute to C_l in that interval. The T^3 phonon contribution to the specific heat is also observed in liquids, as shown in an inset to **Figure 11** for ^{4}He.

Although the Debye model is a useful approximation to the lattice specific heat of many materials over a wide range of temperatures, experimental data can show substantial deviations, particularly at intermediate temperatures. For a discrete lattice with n atoms per primitive unit cell, the frequency distribution of the normal modes is that of three acoustic phonon branches and $3n - 3$ optical phonon branches. For the acoustic branches, phonon dispersion, the variation of the sound velocity with frequency, has an important effect on the phonon density of states that is not taken into account in the Debye model. The optical modes are omitted in this model. The harmonic-lattice approximation, which gives an expansion of C_l in odd powers of T that starts with the Debye T^3 term, is often used to fit experimental data at low temperatures. Sums of Einstein functions, which represent contributions from the optical modes or other peaks in the phonon density of states, and a Debye function are useful over a wide range of intermediate temperatures. **Figure 4** shows an unusual case, C_l for YbGaGe, which is well represented by a single Einstein frequency and a Debye function.

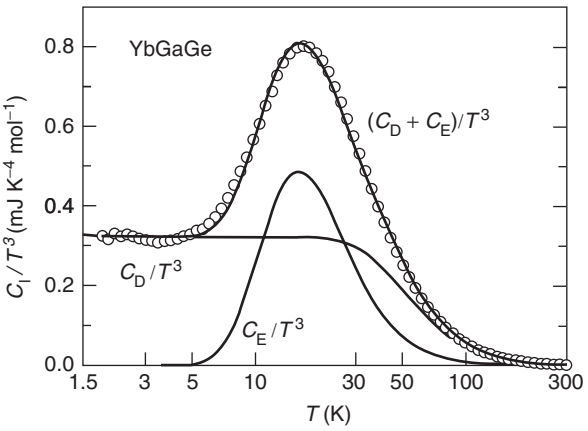


Figure 4 The lattice heat capacity of YbGaGe, as C_l/T^3 vs. log T, fitted with the sum of one Einstein function and one Debye function. The peak in C_l/T^3 centered near 20 K is an unusually prominent, but otherwise typical, signature of a "soft mode," a peak in the phonon density of states. (Specific-heat data are from measurements made at the Los Alamos National Laboratory, Los Alamos, New Mexico: Drymiotis FR, Lawes G, Migliori A, Ledbetter H, Betts JB, *et al.* (2004), unpublished. The electron contribution was subtracted from the data for this figure.)

Conduction electron contribution At ambient and low temperature, and at the densities at which they occur in metals, the conduction electrons must be treated with Fermi–Dirac statistics: there can be at most two electrons, with opposite spin, in any energy state. Their contribution to the specific heat is similar to that of the hypothetical "free electron gas," in which there are no interactions between the electrons or with the lattice. For free electrons, the allowed states have kinetic energies proportional to the sum of the squares of the wave vectors. At $T = 0$, the states are filled up to the Fermi energy,

$$E_F = \left(\frac{\hbar^2}{2m_e}\right)\left(\frac{\pi^2 z N_A}{V}\right)^{2/3}$$

where m_e is the electron mass, z is the "valence," and zN_A the number of electrons in the atomic volume, V. For values of the parameters appropriate to metals, the corresponding characteristic temperature, the Fermi temperature, $T_F \equiv E_F/k_B$, has values $\sim 10^4$–10^5 K. For $T \ll T_F$, most of the states at $E < E_F$ remain fully occupied. It is only electrons with energies within a range $k_B T$ in energy in the vicinity of E_F that can be thermally excited, and by an energy that is also of order $k_B T$. The simple argument that a fraction of the electrons, T/T_F, makes the classical contribution gives $C_e \sim RT/T_F$, which agrees both in temperature dependence and order of magnitude with experimental observations. An exact calculation gives

$$C_e = \left(\frac{\pi^2 z R}{2}\right)\left(\frac{T}{T_F}\right) = \left(\frac{\pi^2 k_B^2}{3}\right) N(E_F) T \equiv \gamma T$$

where $N(E_F)$ is the density of electron states at the Fermi level. The coefficient γ, as defined here for the free electron gas, is proportional to m_e.

The properties of the conduction electrons in a metal are modified by the interactions of the electrons with the periodic potential of the lattice, by electron–electron interactions, and by the electron–phonon interaction. However, according to Landau's Fermi-liquid theory, there is a one-to-one correspondence between excitations in the free electron gas and those in the gas with interactions; the coefficient γ is changed, with m_e replaced with an "effective mass," m^*, but the temperature dependence of C_e is unchanged. The periodic potential of the lattice modifies the relation between electron energy and wave vector. The effect on $N(E_F)$ is to replace m_e with m_{bs}, the "band-structure mass," in the expression for γ. The electron–phonon interaction, which is also responsible for the pairing of the electrons in the BCS theory of superconductivity, introduces an enhancement of γ by a factor $(1 + \lambda)$. In many cases m_{bs} can be calculated with good accuracy, and comparisons with experimental values of γ give the values of λ. For most metals, λ is in the range 0.2–1, with higher values occurring for strong-coupled superconductors such as Hg and Pb. (There is also a contribution to m^* from electron–electron interactions, but it is small, difficult to determine, and usually ignored.) The combined effect of the lattice potential and the electron–phonon interaction gives $m^* = (1 + \lambda)$ m_{bs}. Experimental values of m^*/m_e are ~ 1 for many metals, as high as 10 for some transition and lanthanide metals, and 10^2–10^3 for heavy-fermion compounds, mainly Ce and U intermetallic compounds, in which the large effective mass is associated with the Kondo effect.

Many metals undergo a transition to the superconducting state at a critical temperature, T_c. In most cases, the superconducting-state C_e is approximately exponential in temperature as predicted by the BCS theory, but T^2 and T^3 dependences are found in some "unconventional" superconductors, including heavy-fermion and cuprate superconductors. C_e for Al, a BCS superconductor with $T_c = 1.2$ K, is shown for both the superconducting and normal states in **Figure 5**. In this case the lattice contribution to the specific heat at T_c and below is relatively small, and C_e in the superconducting state is particularly well defined by the experimental data.

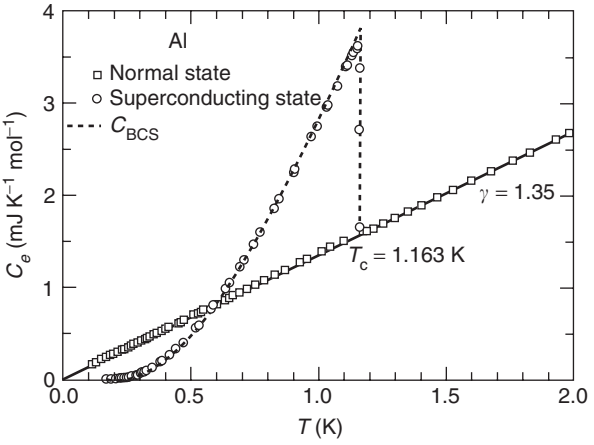


Figure 5 Conduction electron contribution to the specific heats of Al in the normal and superconducting states. The normal-state data below T_c were obtained in a magnetic field that quenched the superconductivity. The dashed curve is the theoretical specific heat of a BCS superconductor (Bardeen J, Cooper LN, and Schriffer JR (1957) Microscopic theory of superconductivity. *Physical Review* 106: 162), with numerical values from: Mühlschlegel B (1959) *Zeitschrift fur physik* 155: 313. Specific-heat data for aluminum are from: Phillips NE (1959) Heat capacity of aluminum between 0.1 and 4.0 K. *Physical Review* 114: 676.

Magnetic contribution Magnetic moments, of the order of magnitude of the Bohr magneton, μ_B, can be associated with the itinerant electrons in a metal and can appear as localized moments in both metals and insulating materials. At sufficiently high temperatures, in the paramagnetic phase, and in the absence of an applied magnetic field, they are randomly oriented and contribute $R \ln(2J + 1)$ to the entropy, S, where J is the resultant angular momentum. However, there is always some internal interaction that tends to produce an ordered structure at temperatures for which $k_B T$ is small relative to the energy of the interaction. In both metals and insulators, the possible interactions include dipole–dipole interactions and exchange interactions. In metals, the RKKY interaction that couples the localized moments via an intermediate polarization of the conduction electrons, and the Kondo effect, which leads to a compensation of the localized moments by conduction electron moments, are also important. In accordance with the third law of thermodynamics, the interactions produce a completely ordered, zero-entropy state at $T = 0$. (Possible exceptions are "frustrated" antiferromagnets, in which the exchange interactions cannot all be satisfied simultaneously.) In the temperature region in which the ordering and the associated reduction in entropy occur, there is a "specific-heat anomaly," a peak in the specific heat, which is related to the entropy by $C_P = (\partial S/\partial T)_P$. The temperature at which the ordering occurs can vary widely, reflecting differences in the energies of the interaction, for example, less than 1 mK for $Ce_2Mg_3(NO_3)_{12} \cdot 24H_2O$ and 1043 K for metallic iron.

When the ordering is driven by an external field, there are $2J + 1$ states with energies $g m_J \mu_B B$, where g is the spectroscopic splitting factor and m_J takes values from $-J$ to $+J$. The specific heat associated with

the magnetic ordering, C_m, takes the form

$$\frac{C_m}{R} = \left(\frac{g\mu_B \boldsymbol{B}}{2k_B T}\right)^2 \left\{ \operatorname{csch}^2\left(\frac{g\mu_B \boldsymbol{B}}{2k_B T}\right) - (2J+1)^2 \operatorname{csch}^2\left[(2J+1)\left(\frac{g\mu_B \boldsymbol{B}}{2k_B T}\right)\right] \right\}$$

This is an example of a Schottky anomaly, a term that is used more generally for any specific-heat anomaly associated with a finite number of energy levels. Electric fields of the crystal lattice acting on the orbital component of the angular momentum can also split the manifold of $2J+1$ states of the isolated ion to produce a Schottky anomaly, in that case called a crystal-field anomaly. For an external magnetic field and $J = 1/2$, an example of a "two-level system," the anomaly takes a particularly simple form,

$$\frac{C_m}{R} = \left(\frac{\Delta}{k_B T}\right)^2 \frac{\exp(\Delta/k_B T)}{[1+\exp(\Delta/k_B T)]^2}$$

where Δ is the energy difference between the two states, in this case $g\mu_B \boldsymbol{B}$. $CuK_2(SO_4)_2 \cdot 6H_2O$ in a field of 1.5 T provides an example of this behavior: the internal interactions, which produce an ordering near 0.05 K in the absence of an external field, can be neglected and the experimental data are in good agreement with the calculated two-level Schottky anomaly, as shown in **Figure 6**.

The internal interactions in α-$MnCl_2 \cdot 4H_2O$ produce antiferromagnetic ordering that results in a specific-heat anomaly near 1.5 K. **Figure 7** shows the zero-field anomaly, which has a shape typical of that associated with ordering by exchange interactions, and the evolution of the anomaly to a more Schottky-like form in an increasing external field. **Figure 8** shows the specific heat of $CeCu_6$, a heavy-fermion compound, in the low-temperature region where the Kondo effect produces a compensation of the Ce moments by an exchange interaction with the conduction electrons.

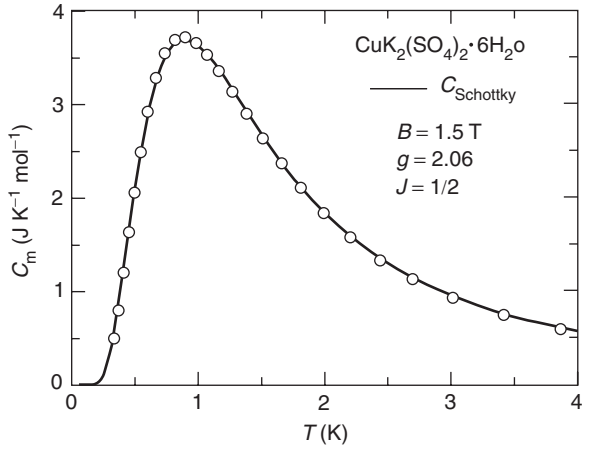


Figure 6 The magnetic specific heat, C_m, of $CuK_2(SO_4)_2 \cdot 6H_2O$ in a magnetic field of 1.5 T. At the temperatures of the measurements, the interactions between the magnetic Cu^{2+} ions have a negligible effect, and a "two-level" Schottky function for isolated ions provides an excellent fit to the data. (The data are from: Giauque WF, Brodale GE, Hornung EW, and Fisher RA (1971) Magnetothermodynamics of $CuK_2(SO_4)_2 \cdot 6H_2O$. II. Magnetic moment, heat capacity, entropy from 0.5 to 4.2 K with fields to 90 kG along the γ magnetic axis. *Journal of Chemical Physics* 54: 273.)

Nuclear moment contribution Contributions to the specific heat associated with nuclear moments, the "hyperfine" specific heat, C_h, are analogous to those associated with electronic moments. The differences arise from the small size of the nuclear magnetic moments, of the order of magnitude of the nuclear magneton, μ_N, three orders of magnitude smaller than μ_B, and the existence of nuclear electric quadrupole moments, Q, which interact with electric-field gradients in the lattice. The nuclear magnetic moments interact with external applied magnetic fields and with internal hyperfine fields produced by ordered electronic moments. They interact with each other through direct dipole–dipole interactions and, in the case of metals, through RKKY coupling. Except in the case of the internal hyperfine fields, which can be very large, it is generally only the high-temperature tail of a Schottky anomaly, associated with the interactions with magnetic fields or electric-field gradients, that is observed. For the case that the electric-field gradient and the magnetic field are in the same direction, the first two terms in the high-temperature expansion of the Schottky anomaly for C_h are D_2/T^2 and D_3/T^3, the coefficients of which have the forms

$$\frac{D_2}{R} = \left(\frac{1}{80}\right)\left[\frac{(I+1)(2I+3)}{I(2I-1)}\right]\left(\frac{e^2qQ}{k_B}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{I+1}{I}\right)\left(\frac{\mu \boldsymbol{B}}{k_B}\right)^2$$

and

$$\frac{D_3}{R} = -\left(\frac{1}{1120}\right)\left[\frac{(2I-3)(I+1)(2I+3)(2I+5)}{2I^2(2I-1)^2}\right] \times \left(\frac{e^2qQ}{k_B}\right)^3 - \left(\frac{1}{10}\right)\left[\frac{(I+1)(2I+3)}{2I^2}\right] \times \left(\frac{e^2qQ}{k_B}\right)\left(\frac{\mu \boldsymbol{B}}{k_B}\right)^2$$

where I is the nuclear spin, μ is the nuclear moment in units of μ_N, e^2q is the electric field gradient, and Q is the nuclear quadrupole moment. These terms have

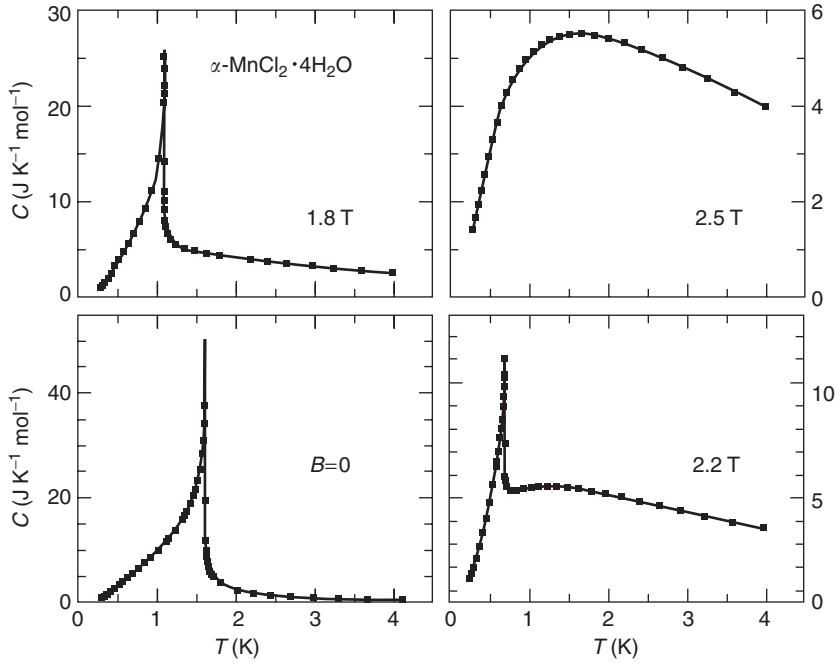


Figure 7 Specific heats for α-$MnCl_2 \cdot 4H_2O$ in various magnetic fields. For $\boldsymbol{B} = 0$, the Mn^{2+} ions order antiferromagnetically with a "lambda" anomaly at a Néel temperature of 1.6 K. As the magnetic field is increased, the antiferromagnetic ordering moves to lower temperatures and the anomaly is attenuated. For sufficiently large $\boldsymbol{B}$, the antiferromagnetic ordering is suppressed and the specific heat evolves into a more Schottky-like anomaly. (The data are from: Giauque WF, Fisher RA, Brodale GE, and Hornung EW (1970) Magnetothermodynamics of α-$MnCl_2 \cdot 4H_2O$. II. Heat capacity, entropy, magnetic moment from 0.4 to 4.2 K with fields to 90 kG along the *b* crystallographic axis. *Journal of Chemical Physics* 52: 2901.)

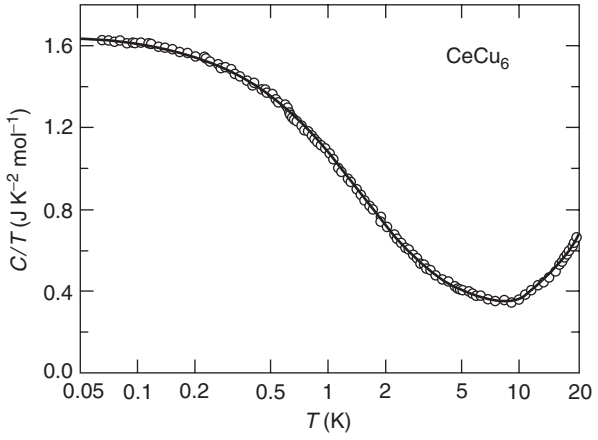


Figure 8 *C/T* vs. log *T* for the heavy-fermion compound $CeCu_6$. The high values of the heat capacity at temperatures below 5 K are a manifestation of the Kondo effect, in which the exchange interaction between the localized 4*f* moments of the Ce ions and the conduction electrons produces a compensation of the Ce moments by the conduction electrons. The associated specific heat is described by a temperature-dependent coefficient of the electron specific heat, $\gamma(T)$. At $T = 0.064$ K, $\gamma(T)$ has increased to 1.63 J K^{-2} mol^{-1}, corresponding to a conduction electron effective mass, m^*, that is three orders of magnitude greater than the free electron mass. (The data are from: Fisher RA, Lacy SE, Marcenat C, Olsen JA, Phillips NE, *et al.* (1987) Specific heat of heavy-fermion $CeCu_6$: effect of pressure and magnetic field. *Journal of Applied Physics* (*Japan*) 26: 1257.)

been identified in experimental data, usually in the vicinity of 0.1–1 K. The large internal hyperfine field, 930 T, in Ho, produces a Schottky anomaly with a maximum near 0.3 K. **Figure 9** is a comparison of the experimental data with the fitted Schottky anomaly, which includes a small quadrupole contribution. Examples of nuclear ordering at the other extreme of the temperature range in which the phenomenon is known are provided by Cu and Ag: in both cases the nuclei order antiferromagnetically, by a combination of dipole–dipole and RKKY interactions, and reductions in the nuclear entropy related to the associated specific-heat anomaly have been observed experimentally. The ordering occurs in the vicinity of 50 nK for Cu, and in the vicinity of 0.8 nK for Ag, which has a lower concentration of nuclear magnetic moments that are also smaller in magnitude.

Phase transitions Phase transitions take a variety of forms and display a corresponding variety in the associated specific-heat anomalies. Discontinuities in the entropy and the volume, the first derivatives of the Gibbs energy, *G*, are the defining characteristic of a thermodynamic first-order transition. They are related to the slope of the phase boundary in the *P*–*T* plane

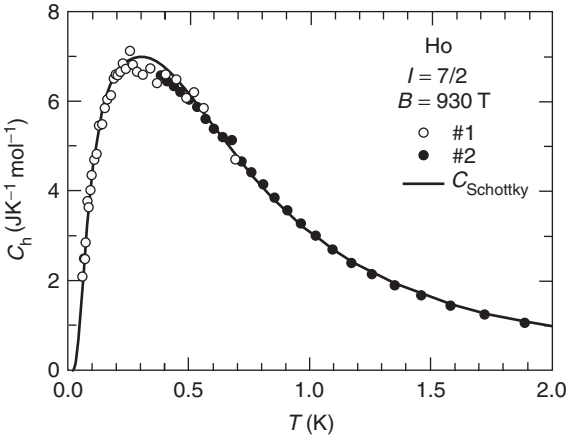


Figure 9 The hyperfine specific heat, C_h, for two samples of ferromagnetically ordered Ho metal. The splitting of the nuclear spin states arises mainly from the interaction of the nuclear magnetic moments with the effective hyperfine field created by the ordered 4*f* electron moments, but there is also a small effect of interaction of the electric quadrupole moment with an electric field gradient. The solid curve, a theoretical fit to the data, is a Schottky function calculated for the resulting, unevenly spaced levels. (The data for sample ♯1 are from: Van Kampen H, Miedema AR, and Huiskamp WJ (1964) Heat capacities of the metals terbium and holmium below 1 K. *Physica* 30: 229; for sample ♯2, from: Lounasmaa OV (1962) Specific heat of holmium metal between 0.38 and 4.2 K. *Physical Review* 128: 1136.)

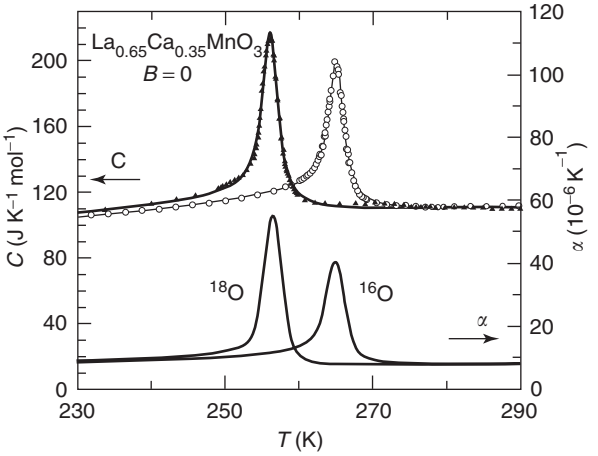


Figure 10 The specific heat, *C*, and thermal expansion, α, of $La_{0.65}Ca_{0.35}MnO_3$, a "colossal magnetoresistance" material that undergoes a first-order transition from a high-temperature, insulating, paramagnetic phase to a low-temperature, metallic, ferromagnetic phase. The transition, at the Curie temperature, shows a strong isotope effect. Data for both ^{16}O and ^{18}O samples are shown. (The data are from: Gordon JE, Marcenat C, Franck JP, Isaac I, Zhang G, *et al.* (2001) Specific heat and thermal expansion of $La_{0.65}Ca_{0.35}MnO_3$: magnetic-field dependence, isotopic effect, and evidence for a first-order phase transition. *Physical Review B* 65: 24441.)

by the Clapeyron equation, $dP/dT = \Delta S/\Delta V$. Ideally, there would be a δ function in C at the transition that would integrate to $T\Delta S$, but the transition is usually broadened by impurities, inhomogeneity, or other factors. **Figure 10** shows the specific heat at the first-order transition associated with the ferromagnetic ordering in $La_{0.65}Ca_{0.35}MnO_3$. In that case, the random mixture of La and Ca ions on their common sites accounts quantitatively for the breadth of the transition. The fitted curves correspond to broadened discontinuities in the entropy and the specific heat. The thermal-expansion data, included in the same figure, illustrate the similarity in the temperature dependence of the two properties, which is expected on the basis of simple thermodynamic models. They also give the value of ΔV, which, together with ΔS from the specific-heat data, satisfies the Clapeyron equation, confirming the first-order nature of the transition.

At an Ehrenfest second-order transition, the first derivatives of the Gibbs energy are continuous, but there are discontinuities in the second derivatives, the specific heat, the thermal expansion, and the compressibility. The Ehrenfest relations for the slope of the phase boundary are $dP/dT = \Delta\alpha/\Delta\kappa = \Delta C_P/\Delta VT\Delta\alpha$. The zero-field transition of a metal from the normal state to the superconducting state is an Ehrenfest second-order transition with a discontinuity between two finite values of the specific heat. The transition in Al, shown in **Figure 5**, is an example. For many second-order transitions, the specific heat diverges at the transition, as illustrated by the antiferromagnetic ordering in α-$MnCl_2 \cdot 4H_2O$ in **Figure 7**. They are frequently called λ transitions, as suggested by the shape of the specific-heat anomaly and the name given to the transition to the superfluid state in liquid ^{4}He at the "λ point." The λ transition in ^{4}He, for which the specific heat is shown in **Figure 11**, has been particularly intensively studied because an exceptionally high degree of sample purity and homogeneity can be attained. In the vicinity of the λ point, the specific heat varies as $\ln|T - T_\lambda|$.

Thermodynamics of Chemical Reactions

Currently, most measurements of specific heat are made for the interest in their microscopic interpretation, but historically a major motivation was to obtain data relevant to the third law of thermodynamics and to the thermodynamics of chemical reactions more generally. At constant pressure and temperature, the conditions for equilibrium between the products and reactants of a chemical reaction, and for equilibrium between two phases of the same substance, are determined by the change in Gibbs energy for the transformation. At constant volume and temperature, the change in the Helmholtz energy, *A*, takes that role. Specific-heat data contribute to the determination of

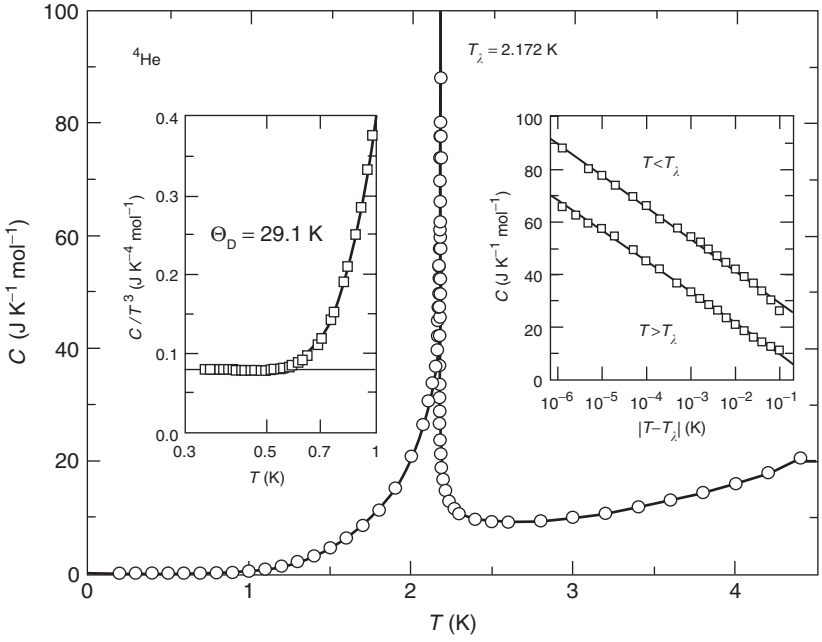


Figure 11 The specific heat of liquid ^{4}He. At T_λ, there is a transition from a low-temperature superfluid phase to a high-temperature, normal-liquid phase. The superfluid phase has properties that are analogous to those of a metal in the superconducting state. As shown in the inset on the left, $C \propto T^3$ for $T \leqslant 0.5$ K: The phonon contribution to the specific heat has the same form, and the same relation to elastic constants, as in solids. In this case, the initial deviation from the T^3 dependence includes a "roton" contribution. At T_λ, there is a logarithmetic divergence of the specific heat, as demonstrated in the inset on the right. (The data for the main panel and right inset are from: Buckingham MJ and Fairbanks WK (1961) The nature of the λ-transition in liquid helium. In: Gorter CJ (ed.) *Progress in Low Temperature Physics*, vol. 3. Amsterdam: North-Holland; for the left inset, from: Hoffer JK (1968) *PhD Thesis* (unpublished), Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA.)

the Gibbs and Helmholtz energies through the thermodynamic relations $G \equiv H - TS$, $(\partial H/\partial T)_P = C_P$, $(\partial S/\partial T)_P = C_P/T$, $A \equiv U - TS$, $(\partial U/\partial T)_V = C_V$, and $(\partial S/\partial T)_V = C_V/T$. They determine the values of the entropy, and the temperature dependences of the enthalpy, energy, and Gibbs and Helmholtz energies.

As stated in one of the many formulations of the third law of thermodynamics, for any substance in a perfect crystalline state, the entropy is zero in the limit $T \to 0$. The importance for chemical thermodynamics is that values of the entropy can be obtained from specific-heat data alone: the "third-law entropy" is obtained by extrapolating specific-heat data to 0 K, integrating C_P/T to obtain $S(T)$–S_0, and assuming, as suggested by the third law, that S_0, the entropy at the 0 K state reached by the extrapolation, is zero. Typically, the data extend to a temperature between 1 and 15 K, the extrapolation is by a semiempirical fit to the lowest-temperature data that is based on appropriate theoretical expressions, and the entropies are tabulated for 298.15 K and the standard state (the thermodynamically stable form of the substance at a pressure of one bar).

The agreement obtained in comparisons of the values of ΔS for chemical reactions calculated with and without using third-law entropies, and in comparisons of the third-law entropies with values calculated from spectroscopic data played an important role in establishing the third law. In other cases, however, similar comparisons led to the identification of substances, for example, CO, NO, and H_2O, that do not attain a "perfect crystalline state" at 0 K, and for which $S_0 \neq 0$. The existence of such materials emphasizes the importance of caution in applying the third law to the determination of entropies from specific-heat data. The third law of thermodynamics, unlike the first and second laws, cannot be expressed by a simple mathematical relation that applies rigorously in all cases. Its application requires an understanding of its nature, which can be based on a statistical mechanical expression for the entropy, the Boltzmann formula, $S = k_B \ln \Omega$ (where Ω is the number of accessible quantum states, subject to the constraint of a fixed energy), together with the principle that there is a single state of lowest energy that is reached at $T = 0$ by systems that are in thermodynamic equilibrium. The case of CO, which is well understood, is a useful illustration of the factors that determine whether or not $S_0 = 0$: In the solid, the two orientations of a molecule that differ by an interchange of the C and O atoms are two distinguishable states, but, because the molecule is

relatively symmetric with respect to both shape and electric charge distribution, the difference in energy is small. The lowest energy configuration is one in which there is a perfectly ordered arrangement of the molecules with each in a definite orientation relative to the others, but the energy is not much lower than that of the higher-entropy, random distribution of orientations that is thermodynamically stable at high temperature. Below $\sim 5\,\mathrm{K}$ the lower-energy, ordered configuration does become thermodynamically stable, but there is a potential barrier to reorientation of the molecules that exceeds the thermal energy at that temperature, and the perfectly ordered state is not attained. For kinetic reasons, there is "frozen-in disorder" at 0 K and $S_0 \neq 0$. In the completely random configuration, there are two possible states for each molecule, $\Omega = 2^{N_A}$, $S = R \ln 2 = 5.76\,\mathrm{J\,K^{-1}\,mol^{-1}}$. The experimental value is $S_0 = 4.2\,\mathrm{J\,K^{-1}\,mol^{-1}}$, indicating that the molecules are partially ordered, at least in the particular measurement that gave this result. There are other cases, for example, NO, H_2O, and the H_2O in $Na_2SO_4 \cdot 10H_2O$, in which the disorder is complete. These examples afford an understanding of the treatment of nuclear contributions to the entropy: the random distribution over the lattice sites of different isotopes of the same element persists to 0 K, the disorder is frozen-in, and there is no contribution to experimental specific-heat data. The disorder associated with nuclear spin orientation is not frozen in, but at temperatures above 1 K it is normally complete, the associated entropy is constant, and there is no contribution to the specific heat. Both of these contributions to the entropy cancel in chemical reactions, and, although they could be calculated with the Boltzmann formula, they are omitted in tabulations of the entropies of chemical substances. There are also some paramagnetic salts in which the magnet moments of the unpaired electrons, as in the case of nuclear moments, order only below 1 K. Specific-heat measurements to 1 K would suggest $S_0 \neq 0$, but measurements to sufficiently low temperature would give $S_0 = 0$.

The enthalpy, energy, and Gibbs and Helmholtz energies cannot be determined from specific-heat data alone. Tabulations of thermodynamic data give the values of these quantities for compounds in their standard states, relative to the values for the elements in their standard states, for example, as values of the "heat of formation," ΔH, for the reaction in which the compound is formed from the elements. Specific-heat data do determine the temperature dependences of these quantities, but determination of absolute values requires, in addition, the evaluation of constants of integration. For the enthalpy and energy, one constant of integration is needed, for example, one value of a heat of formation for the enthalpy; for the Gibbs and Helmholtz energies, two are needed.

See also: Bose–Einstein Condensation; Diffusionless Transformations; Incommensurate Phases; Irreversible Thermodynamics and Basic Transport Theory in Solids; Phase Transformation; Phase Transformations, Mathematical Aspects of; Phases and Phase Equilibrium; Thermodynamic Properties, General.

PACS: 51.30.+i; 63.20.−e; 65.40.Ba; 65.40.De; 65.40.Gr; 67.40.Kh; 74.25.Bt; 75.20.Hr; 75.20.−g; 75.40.Cx; 75.50.Cc; 75.50.Ee; 82.60.Fa; 75.20.Hr

Further Reading

Ashcroft NW and Mermin ND (1976) *Solid State Physics*. New York: Holt, Rinehart and Winston.

Atkins P and de Paula J (2002) *Physical Chemistry*, 7th edn. New York: W. H. Freeman.

Barron THK and White GK (1999) *Heat Capacity and Thermal Expansion at Low Temperatures*. New York: Kluwer.

Blackman M (1955) In: Flügge S (ed.) *Encyclopedia of Physics*, vol. 7, part 1, p. 325. Berlin: Springer.

Callen HB (1963) *Thermodynamics*. New York: Wiley.

Davidson N (1962) *Statistical Thermodynamics*. New York: McGraw-Hill.

Giauque WF (1969) *The Scientific Papers of William F. Giauque, Low Temperature, Chemical, and Magnetothermodynamics*, vol. 1. New York: Dover.

Giauque WF (1995) *The Scientific Papers of William F. Giauque, Low Temperature, Chemical, and Magnetothermodynamics*, vols. 2 and 3. Berkeley, California: University of California Printing Services.

Gopal ESJ (1966) *Specific Heats at Low Temperatures*. New York: Plenum.

de Jongh LJ and Miedema AR (1974) *Experiments on Simple Magnetic Model Systems*. London: Taylor and Francis.

Kittel C (1963) *Quantum Theory of Solids*. New York: Wiley.

Kittel C (1986) *Introduction to Solid State Physics*, 6th edn. New York: Wiley.

Kittel C and Kroemer H (1989) *Thermal Physics*, 2nd edn. New York: W. H. Freeman.

de Launay J (1956) The theory of specific heats and lattice vibrations. In: Seitz F and Turnbull D (eds.) *Solid State Physics*, vol. 2, p. 219. New York: Academic press.

Phillips NE (1971) Low-temperature heat capacity of metals. *CRC Critical Reviews in Solid State Sciences*, vol. 467. New York: CRC Press.

Pitzer KS (1995) Thermodynamics, 3rd edn., vol. 2. New York: McGraw-Hill.

Nomenclature

A	Helmholtz energy ($\mathrm{J\,mol^{-1}}$)
B	magnetic induction (T) (tesla $\equiv$ weber $\mathrm{m^{-2}}$)
$\boldsymbol{B}_\mathrm{h}$	hyperfine magnetic induction (T)
C	specific heat ($\mathrm{J\,K^{-1}\,mol^{-1}}$)
C_e	electronic specific heat ($\mathrm{J\,K^{-1}\,mol^{-1}}$)
C_h	hyperfine specific heat ($\mathrm{J\,K^{-1}\,mol^{-1}}$)
C_l	lattice specific heat ($\mathrm{J\,K^{-1}\,mol^{-1}}$)

C_m	magnetic specific heat ($J K^{-1} mol^{-1}$)
C_P	specific heat at constant pressure ($J K^{-1} mol^{-1}$)
C_r	rotational specific heat ($J K^{-1} mol^{-1}$)
C_t	translational specific heat ($J K^{-1} mol^{-1}$)
C_v	vibrational specific heat ($J K^{-1} mol^{-1}$)
C_{BCS}	specific heat of a BCS superconductor ($J K^{-1} mol^{-1}$)
C_D	Debye specific heat ($J K^{-1} mol^{-1}$)
C_E	Einstein specific heat ($J K^{-1} mol^{-1}$)
C_V	specific heat at constant volume ($J K^{-1} mol^{-1}$)
D	electric displacement vector ($C m^{-1}$ (coulomb m^{-1}))
e	electron charge (C (coulomb))
e^2q	electric-field gradient at the nucleus ($V m^{-2}$ (volt m^{-2}))
E	electric field ($V m^{-1}$ (volt m^{-1}))
E_F	Fermi energy ($J mol^{-1}$)
g	spectroscopic splitting factor (dimensionless)
G	Gibbs energy ($J mol^{-1}$)
$\boldsymbol{H}$	magnetic field ($A m^{-1}$ (ampere m^{-1}))
H	enthalpy ($J mol^{-1}$)
I	nuclear spin quantum number (dimensionless)
J	electronic spin quantum number (dimensionless)
k_B	Boltzmann constant ($J K^{-1}$)
m	mass (kg)
m_e	free electron mass (kg)
m_J	quantum numbers labeling sublevels of J in a magnetic field (dimensionless)
m_{bs}	electron mass from band structure calculations (kg)
m^*	effective electron mass (kg)
$\boldsymbol{M}$	magnetic moment ($J T^{-1} mol^{-1}$)
n	number of atoms (dimensionless)
N_A	Avagadro's number (mol^{-1})
$N(E_F)$	density of states at the Fermi energy (states $J^{-1} mol^{-1}$)
P	pressure ($N m^{-2}$ (newton m^{-2} = pascal))
q	heat ($J mol^{-1}$)
Q	nuclear electric quadrupole moment (m^2)
R	gas constant ($J K^{-1} mol^{-1}$)
S	entropy ($J K^{-1} mol^{-1}$)
T	temperature (K)
T_c	superconducting critical temperature (K)
T_F	Fermi temperature (K)
T_λ	lambda temperature for ^{4}He superfluid transition (K)
U, E	energy ($J mol^{-1}$)
V	volume (m^3)
w	work ($J mol^{-1}$)
z	valence (dimensionless)
α	thermal expansion (K^{-1})
Δ	energy difference between two states (J)
γ	Sommerfeld constant ($J K^{-2} mol^{-1}$)
θ_r	characteristic rotation temperature (K)
θ_v	characteristic vibration temperature (K)
Θ_D	Debye temperature (K)
Θ_E	Einstein temperature (K)
κ	isothermal compressibility ($m^{-2} kg^{-1}$)
λ	electron–phonon interaction parameter (dimensionless)
μ	nuclear magnetic moment (in units of μ_N)
μ_0	permeability of free space (henry m^{-1})
μ_B	Bohr magneton ($J T^{-1}$)
μ_N	nuclear magneton ($J T^{-1}$)
ω	frequency (s^{-1})
ω_D	Debye cutoff frequency (s^{-1})
ω_E	characteristic Einstein frequency (s^{-1})
$\hbar$	Planck's constant h divided by 2π (J s)

Spin Density Waves and Magnons

N Harrison, Los Alamos National Laboratory, Los Alamos, NM, USA

Introduction

A review of semiconductors would be incomplete without the inclusion of those classes of extremely narrow gap semiconductors and semimetals that result from microscopic interactions between itinerant (or nearly free) electrons in a correlated metal. This article is devoted to spin-density waves, which are a consequence of Coulomb interactions between electrons causing a system to become antiferromagnetic, which would otherwise be metallic. Perhaps, the most extreme example is that where the Coulomb interactions between electrons exceed the total electronic bandwidth, leading to an insulator that can be very well described by the Hubbard model. Spin-density waves are subtle by comparison, being much more dependent on details of the Fermi surface topology. It is, therefore, more convenient to consider spin-density waves as perturbations of electronic bands of quasiparticles in a metal. The gaps are

significantly smaller, enabling quasiparticles to be easily excited across the spin-density wave gap, with tiny pockets of metallic carriers surviving in many cases.

The resulting antiferromagnetism can be mapped onto the simple case of a Heisenberg antiferromagnet. The moments are nevertheless smaller with comparatively long spatial periodicities of the spin modulation projected in preferred crystallographic directions that are predetermined by the details of the band structure. The existence of thermodynamic coupling between the spin-density wave and the charge degrees of freedom of the electronic bands causes these systems to be subject to phase excitations (or phasons) in a manner similar to the charge-density wave systems. Changes in the orientation of the spins, in contrast, take place via spin-wave excitations (or magnons), which are qualitatively similar to those that occur in Heisenberg antiferromagnets.

This article begins by considering the instability at the Fermi surface responsible for the spin-density wave formation, and then continues by discussing their static and dynamic properties. These include collective excitations involving both electronic (phason) and spin (magnon) degrees of freedom. Finally, the topic of field-induced spin-density waves, in which orbital contributions to the total free energy help stabilize such ground states in the presence of a strong magnetic field, is discussed. Transitions between different field-induced spin-density wave phases involving different levels of orbital quantization give rise to a novel manifestation of the bulk quantum Hall effect.

Fermi Surface Nesting and Spin-Density Wave Formation

The term "spin-density wave" refers to a simple one-dimensional (1D) antiferromagnetic modulation of the local spin density $\boldsymbol{S} = \boldsymbol{S}_0 \cos(\boldsymbol{r} \cdot \boldsymbol{Q}_0 + \phi_0)$ condensed from electrons that may subsequently become insulating. An essential precondition for spin-density wave formation to be energetically favorable is that a large fraction of the electron and hole quasiparticles share the same group velocity $\boldsymbol{v}_F$ at the Fermi energy ε_F. The residual Coulomb potential between quasiparticles U must also be significant, but smaller than the total electronic bandwidth W. These requirements tend to be met in systems where the calculated electronic bands (1) are very narrow and (2) are highly anisotropic or have large flat sections of Fermi surface that can efficiently nest. The term "nesting" refers to the process by which the translational symmetry breaking of the crystalline lattice at a characteristic wave vector $\boldsymbol{Q}_0$ eliminates major sections of the Fermi surface upon opening a gap 2Δ. In a spin-density wave system, this gap opens up wherever spin-up and spin-down electronic bands cross. **Figure 1** shows an example of the cross section of a Fermi surface amenable to nesting. The combined process of nesting and gap formation saves energy by greatly reducing the total electronic density of states at ε_F.

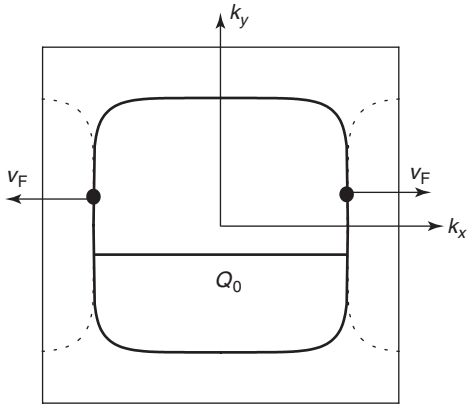


Figure 1 An example of a fictional Fermi surface susceptible to nesting. The Fermi velocity $\boldsymbol{v}_F$ is always perpendicular to the Fermi surface. Translation of the Fermi surface by the nesting vector $\boldsymbol{Q}_0$ enables sections of the Fermi surface with opposing Fermi velocities to meet each other tangentially. The nested sections become dielectric if the bands anticross upon opening of a gap. The dotted lines represent shadows of the translated Fermi surface.

Ground states of spin-density waves occur most commonly in molecular or organic metals, although classic examples occur in elemental chromium and its alloys. The antiferromagnetic ground states of heavy fermion systems, which have quasiparticle effective masses many times the free electron mass, can also sometimes be understood in terms of spin-density waves. While many of these materials have complicated electronic structures, the spin-density wave theory is most easily applied to systems with only a single electronic band,

$$\varepsilon_{k,\sigma} = \hbar^2 k_x^2/2m + \varepsilon_\perp(k_y, k_z) + \sigma_g \mu_B B \qquad [1]$$

crossing ε_F. The simplest case, by far, is that for a metal where the single electronic band has a significant dispersion only along chains parallel to $\boldsymbol{x}$, where k_x is the x-component of the momentum vector $\boldsymbol{k}$ and $\sigma = \pm 1/2$ represents spin-up and spin-down branches. A purely 1D dispersion is obtained by setting the limit $\varepsilon_\perp(k_y, k_z) \to 0$. Because spin-density wave formation only affects states that are close to ε_F, it is often convenient to consider a linear approximation of the form

$$\varepsilon_{k,\sigma} = \hbar v_F |k_x - k_F| + \sigma_g \mu_B B \qquad [2]$$

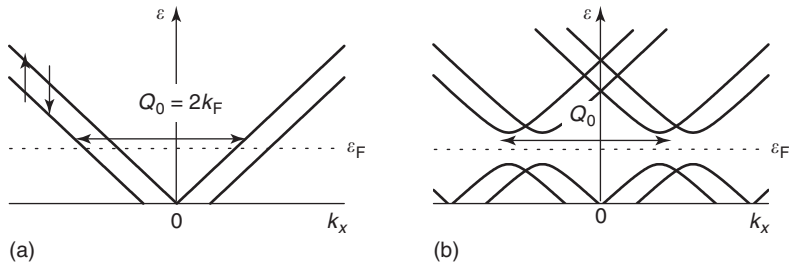


Figure 2 (a) The electronic dispersion according to eqn [2] with a finite magnetic B included to lift the spin degeneracy. The up and down arrows indicate the spin-up and spin-down branches respectively, while Q_0 denotes the nesting vector. (b) The same bands, according to eqn [3] after Fermi surface nesting and gap formation. The size of the gap at ε_F is 2Δ.

where $v_F = \hbar k_F/m$ at the 1D Fermi surface where $k = \pm k_F$. Theoretical models of spin-density waves regularly consider a dispersion of the form given by eqn [2] in conjunction with a lattice spacing along x equal to $a = 2k_F/\pi$. This situation then corresponds to a half-filled band containing one electron per unit cell. There are two reasons why theoreticians make such a choice. One is that stoichiometric metals always contain an integral number of electrons per unit cell, where 1 is then the simplest number. The other is that half-filled bands facilitate direct comparisons with the Hubbard model, which can also yield antiferromagnetic ground states. Here, eqn [2] is assumed to apply to more general band fillings. **Figure 2a** shows a sketch of this dispersion relation for which $\varepsilon_\perp(k_y, k_z) \approx 0$ and the degeneracy of the spin-up and spin-down states is lifted by a finite B in order to demonstrate the independence of $Q_0 = 2k_F$ on a magnetic field. **Figure 2b** shows this same dispersion relation after undergoing translational symmetry breaking with respect to the nesting vector and gap formation. The resulting multiple bands are obtained by solving

$$(2\varepsilon_\Delta - (\varepsilon_{k,\sigma} + \varepsilon_{nQ_0-k,-\sigma}))^2 - (\varepsilon_{k,\sigma} - \varepsilon_{nQ_0-k,-\sigma})^2 = 4\Delta^2 \quad [3]$$

for ε_Δ. The existence of a 1D Fermi surface enables perfect nesting in this case, giving rise to a fully gapped quasiparticle spectrum. Because spin-up and spin-down bands are mixed, spin is no longer a good quantum number for quasiparticle excitations across the gap.

Microscopic Theory

The opening of the gap 2Δ can be formally described in terms of the Bardeen–Cooper–Schrieffer (BCS) theory that applies to superconductivity, but with the pairing occurring at a finite momentum $Q = Q_0$ between electron and hole quasiparticles of opposite spin. The distinction between electrons and holes is purely a relative one, introduced to account for the fact that their group velocities oppose. In charge- and spin-density wave systems alike, the Fermi surface instability originates from a divergency in the Q-dependent charge and spin susceptibilities at $Q_0 = 2k_F$ which becomes logarithmic, $\chi_0(Q) \propto \ln|(Q + Q_0)/(Q - Q_0)|$, when the electronic dispersion is 1D as given by eqn [2]. Spin-density wave formation becomes increasingly favorable over the charge-density wave formation when the spin susceptibility is enhanced by a Stoner factor such that $\chi(Q) = \chi_0(Q)/(1 - U\chi_0(Q)/2\mu_0\mu_B^2)$. Because Q_0 is independent of the magnetic field, spin-density waves are particularly robust in magnetic fields, and in some cases can even be strengthened, as will be shown to be the case of the field-induced spin-density wave systems. For the dispersion relation given by eqn [2] and **Figure 2a**, the mean-field theory yields

$$\Delta_{T=0} = We^{-1/\lambda} \quad [4]$$

at zero temperature, where $\lambda = Un(\varepsilon_F)$ is the effective coupling constant, and $n(\varepsilon_F) \cong W^{-1}$ is the density of the electronic states (per unit cell) at ε_F. The effective electronic bandwidth is $W = \pi\hbar v_F/a$ for the dispersion relation given by eqn [2]. As is the case for superconductivity, the transition into a spin-density wave phase is expected to be of second order in the absence of an applied magnetic field and additional interactions, with the transition temperature T_{SDW} given by

$$2\Delta \cong 3.52k_B T_{SDW} \quad [5]$$

in the weak coupling limit. The single-particle electronic density of states becomes

$$n(\varepsilon) = W^{-1}\mathrm{Re}\left\{\frac{|\varepsilon - \varepsilon_F|}{\sqrt{(\varepsilon - \varepsilon_F)^2 - \Delta^2}}\right\} \quad [6]$$

upon opening a gap, which can be obtained by differentiating k_x in eqn [3] with respect to energy ε.

The spin-density wave condensation energy at zero temperature is given by

$$\Phi = -\tfrac{1}{2}\, n(\varepsilon_F)\Delta^2 \qquad [7]$$

which is again similar to that in superconductivity. Since the present model considers only the case of a 1D dispersion, the coherence length, $\xi_0 = \hbar v_F/\pi\Delta$, is strictly defined in one dimension.

Basic Static Properties

In high-quality samples at temperatures far below T_{SDW}, the correlation length over which the spin-density wave remains periodic greatly exceeds the coherence length. It often follows that the low-temperature, low electric field, and low-frequency conducting properties are those of a simple semiconductor or semimetal. In the case of the simple dispersion given by eqn [2], perfect nesting leads to a classic semiconductor-like behavior of the conductivity

$$\sigma_{xx} = \sigma_0 e^{-\Delta/k_B T} \qquad [8]$$

Ordinary quasiparticles must be thermally excited across the spin-density wave gap in order to contribute to the electrical conductivity, just as in a semiconductor. In metals with more complicated band structures, such as chromium, however, imperfect nesting leaves behind residual electron and hole pockets that continue to conduct in a conventional metallic fashion. In such a case, the electrical properties of the ground state are those of a semimetal, for which the conductivity contribution from quasiparticles excited across the spin-density wave gap is comparatively negligible. An important exception occurs in the case of magnetic field-induced spin-density waves, where the gap is augmented by a magnetic field as a consequence of Landau quantization.

The magnetic properties of spin-density waves are qualitatively similar to those of an antiferromagnet with a reduced moment

$$\mu = g\mu_B S = \frac{4\Delta}{U}\mu_B \qquad [9]$$

and with a long period

$$\Lambda = \frac{2\pi}{Q_0} \qquad [10]$$

that is matched to $2k_F$ rather than the crystalline lattice period a. The direction in which the spins preferentially align is determined by the effective coupling constants between them. The primary coupling constant for a simple 1D chain is given by

$$J = \frac{\hbar v_F (1-\lambda)^{1/2}}{\Lambda} \qquad [11]$$

in the weak coupling limit. If the spins order in a preferred direction orthogonal to the chains, the magnetic susceptibility $\chi_{||}$ along that preferred direction is zero at temperatures $T \ll T_{SDW}$, provided the applied magnetic field is lower than that required to flop the spin. In other directions, the spins are able to cant, giving rise to a finite spin susceptibility $\chi_\perp = Ng^2\mu_0\mu_B^2/2VJ$, where V is the volume per unit cell.

In a similar manner to charge-density waves, the commensurability of a spin-density wave is determined by the number of electrons per unit cell, $N = Q_0 a/\pi$ where $2\pi/a$ is the length of the 1D Brillouin zone prior to ordering. If N can be accurately reduced to a simple fraction of the form $N = A/B$, where A and B are integers, the superlattice period over which both the spin-density and atomic positions repeat is aB, in which case the spin-density wave is considered to be commensurate. The spin-density wave is incommensurate if N is an irrational number.

Collective Excitations: Phasons

The close correlation between the periodicity of spin-density waves and the electronic structure near the Fermi energy implies that they can undergo charge excitations like charge-density waves. These excitations then determine the electrodynamics of spin-density waves. It also serves as a useful approximation to simplify the treatment of the charge degrees of freedom by expanding the free energy in even powers of the order parameter and its spatial and temporal derivatives so as to obtain

$$\Phi = \alpha|\Delta|^2 + \beta\left(\frac{d\Delta}{dx}\right)^2 + \gamma\left(\frac{d\Delta}{dt}\right)^2 \qquad [12]$$

Compatibility with the mean-field theory is obtained by setting $\alpha = -n(\varepsilon_F)/2, \quad \beta = n(\varepsilon_F)(\hbar v_F/2\Delta)^2$, and $\gamma = n(\varepsilon_F)(\hbar/2\Delta)^2$. The order parameter itself can be expanded as

$$\Delta = (\Delta_0 + \delta)e^{i\phi} \qquad [13]$$

where δ respresents a small perturbation in its amplitude and $\phi = \boldsymbol{Q}\cdot\boldsymbol{r} = Qx$ accounts for the phase of the spin modulation described in the introduction. The dispersion relations for the excitation of the spin-density wave can then be obtained by assuming independent oscillatory variations $\delta = \delta_0 e^{i(\omega_A t - kx)}$ and $\phi = \phi_0 + \delta\phi e^{i(\omega_\phi t - kx)}$ of the amplitude and phase of the spin-density wave, respectively. Noting that the first two terms on the right-hand side of eqn [12] are potential energy terms while the third is a kinetic energy term, the application of Lagrangian mechanics yields

$$(\hbar\omega_A)^2 = 2\Delta^2 + (\hbar v_F k)^2 \qquad [14a]$$

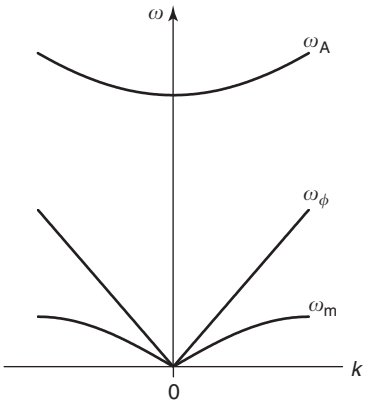


Figure 3 Approximate dispersion relations of excitations of a simple spin-density ground state at low temperatures, ignoring the effects of impurities and magnetic anisotropy. ω_A is the amplitude mode, which effectively corresponds to high-energy magnon excitations. ω_ϕ is the phason mode, corresponding to deformation and translational motion of the spin-density wave. ω_m is the magnon mode, corresponding to spin-waves involving purely rotational degrees of freedom of the spins.

and

$$\omega_\phi = v_F k \qquad [14b]$$

for the amplitude and phase modes, respectively, which are illustrated in **Figure 3**. The first mode occurs at energies comparable to the gap. Since $S \propto \Delta$, this mode is effectively a high-energy magnon mode. The second mode vanishes at $k = 0$, therefore corresponding to a collective sliding mode of the spin-density wave phase, which carries a dissipationless current analogous to the Fröhlich mode in charge-density wave systems. In practice, spin-density waves become strongly coupled to impurities and defects in the crystalline lattice, resulting in the opening of a small gap at $k = 0$. This inhibits dissipationless sliding, thereby restoring the familiar semiconducting properties of the spin-density wave state at low excitation energies.

The frequency-dependent conductivity can be estimated by considering an electric field, $E(t) = E_0 e^{i(\omega_\phi t - kx)}$ that then couples directly to the condensate charge. The equation of motion becomes

$$\frac{d^2\phi}{dt^2} + v_F^2 \frac{d^2\phi}{dx^2} = \frac{2k_F eE(\omega_\phi)}{m} \qquad [15]$$

where

$$\rho = \frac{e}{\pi}\frac{d\phi}{dx} \qquad [16]$$

is the accumulated electrical charge resulting from compression or tension of the spin-density wave and

$$j_{SDW} = \frac{e}{\pi}\frac{d\phi}{dt} \qquad [17]$$

is the electrical current resulting from its sliding motion. The solution of eqn [15] yields a frequency-dependent conductivity of the collective mode of the form

$$\sigma_{coll}(\omega_\phi) = \frac{ne^2}{m}\left(\frac{\pi}{2}\,\delta(\omega_\phi) + \frac{i}{\omega_\phi}\right) \qquad [18]$$

where n is the carrier density (becoming N/a in one dimension) and $\delta(\omega_\phi)$ is a Dirac delta singularity. Since it does not include pinning interactions with impurities, which occur in real systems, this result is somewhat academic. Such interactions have two effects: the first is to bind the spin-density wave to the lattice, causing the Dirac delta singularity to be shifted to finite frequencies, $\delta(\omega_\phi - \omega_0)$, and the second is to resist the sliding motion of the spin-density wave causing losses that broaden the collective mode. This broadening can be phenomenologically modeled by considering an effective relaxation time, τ. In spite of these complications, eqn [18] is a useful result because an integration $(2/\pi)\int \sigma_{coll}\, d\omega_\phi$ of the conductivity over the frequency yields the same total spectral weight ne^2/m as is obtained in the normal metallic state. The implication of this result is that the spin-density wave collective mode is able to account for all of the spectral weight. This contrasts with the case of a charge-density wave where coupling to the ionic lattice distortion causes the collective mode to acquire a heavy effective mass, reducing its share of the spectral weight.

A convenient means of modeling the effects of impurities on the frequency-dependent conductivity is to assume that the pinning forces dominate for small local perturbations of the spin-density wave position compared to the spin-density period $\Lambda = 2\pi/Q_0$, causing it to resonate about pinning sites at a characteristic frequency ω_0. This can be formally accomplished by inserting a pinning restoring force $\omega_0^2 x$ and a damping term $(1/\tau)d\phi/dt$ into eqn [15], and by assuming rigid motion of the spin-density wave, whereby $v_F^2(d^2\phi/dx^2) = 0$. Hence,

$$\frac{d^2\phi}{dt^2} + \frac{1}{\tau}\frac{d\phi}{dt} + \omega_0^2\phi = \frac{2k_F eE(\omega_\phi)}{m} \qquad [19]$$

which facilitates a particularly trivial solution of the form

$$\sigma_{coll}(\omega_\phi) = \frac{ne^2}{i\omega_\phi m}\left(\frac{\omega_\phi^2}{\omega_0^2 - \omega_\phi^2 - i\omega_\phi/\tau}\right) \qquad [20]$$

Owing to the simplifications made in eqn [19], eqn [20] can only be considered as an approximation. In addition to not taking into consideration deformations

of the phase, it also does not consider any details of the mechanism by which spin-density waves interact with impurities and defects. In the case of a charge-density wave, for example, the collective mode consists of a spatial modulation of the local charge, providing a direct means of coupling to charged impurities in the lattice. A spin-density wave must undergo relative phase shifts between its spin-up and spin-down components in order to realize a qualitatively similar type of coupling. In spite of these limitations, eqn [20] provides a surprisingly good account of the frequency-dependent conductivity observed in most spin-density wave systems. The conductivity at yet higher frequencies then includes an additional resonance corresponding to quasiparticle excitations across the gap 2Δ, which is typically two to three orders of magnitude larger than ω_0.

By taking appropriate limits, eqns [19] and [20] also provide one with a rough phenomenological understanding of the reaction of spin-density waves to a constant electric field. By setting $\omega_\phi = 0$ in eqn [20], for example, one can obtain zero conductivity, which is an appropriate value for small values of the electric field E_0 at $T = 0$. By seeking a time-independent solution to eqn [19], one can see that the spin-density wave will undergo a continuous sliding motion provided the electric field exceeds a threshold value $E_0 > m\omega_0^2\phi/2ek_F$, where the pinning force should eventually reverse and oscillate with x for $\phi > \pi/2$. This leads to a situation for large E_0 where the force of the electric field is balanced against the damping term, yielding a Drude-like conductivity.

Collective Excitations: Magnons

In eqn [14a], the amplitude of the spin modulation can be seen to undergo excitations, but with a rather large gap 2Δ equivalent to that for quasiparticle excitations at $T = 0$. Low-energy magnon modes are not obtained from the simple model of the free energy given by eqn [12], because it does not consider the rotational degrees of freedom of the spins that are necessary to produce spin waves. These can instead be obtained by mapping the spin degrees of freedom onto a simple 1D Heisenberg antiferromagnet Hamiltonian of the form

$$H = \left(\frac{\mu}{\mu_B}\right)^2 J \sum_p \boldsymbol{S}_{2p}\boldsymbol{S}_{2p+1} - g\mu_B \boldsymbol{B} \sum_p \boldsymbol{S}_{2p} + H_\perp \qquad [21]$$

Here, the prefactor accounts for the reduced moment, while the index p refers to a system with the periodicity Λ of the reconstructed Brillouin zone rather than the original lattice. The magnetization is expected to be isotropic for isolated chains, in which case $H_\perp = 0$, leading to a gapless spin wave behavior.

The equation of motion

$$\hbar\frac{d\boldsymbol{S}_{p+q}}{dt} = \boldsymbol{m} \times \boldsymbol{B}_{eff} \qquad [22]$$

for spin waves is obtained by equating the rate of change of the angular momentum against the torque imposed on spins rotated with respect to each other along the chain. Here, $\boldsymbol{m} = \mu g \boldsymbol{S}$ is the reduced moment expressed in vector form to account for its orientation. The total effective molecular field of an antiferromagnet,

$$\boldsymbol{B}_{eff} = (-1)^q \frac{J}{g\mu_B}(\boldsymbol{S}_{2p+q-1} + 2\boldsymbol{S}_{2p+q} + \boldsymbol{S}_{2p+q+1}) \qquad [23]$$

felt by each moment due to its two nearest neighbors alternates between odd and even q sites. The solution to eqn [22] is well known to have the form

$$\hbar\omega_m = 2JS\left|\sin\frac{\Lambda k}{2}\right| \qquad [24]$$

where, $\Lambda/2$ is the effective separation between spins along the chain. The dispersion relation given by eqn [24] is sketched in **Figure 3** along with the other modes, and can be seen to be approximately linear at low excitation energies.

In real spin-density wave systems, the crystal extends in three dimensions, leading to the possibility of additional couplings between the chains. Should the spins then have preferred axes of orientation, the antiferromagnet becomes anisotropic. The spins may no longer be free to rotate under such circumstances, causing the linear dispersion given by eqn [24] to become gapped. These effects are usually modeled by considering the anisotropy term $H_\perp$ in eqn [21] to include contributions that increase or reduce the energy D associated with the alignment of spins along certain directions. The size of the gap is typically of the order $2S\sqrt{JD}$.

Field-Induced Spin-Density Waves

Equation [3], as depicted in **Figure 2a**, yields a set of bands that change with the magnetic field, but where the gap in the density of electronic states at ε_F remains symmetric about ε_F. This occurs because the spin-density waves are the consequence of pairing between electron and hole quasiparticles of opposite spin, causing simple spin-density wave systems to be independent of a magnetic field. This situation changes, however, as soon as the nesting becomes imperfect. Imperfect nesting leaves behind small

pockets of the Fermi surface that then undergo orbital quantization in a magnetic field. Orbital quantization has a pronounced effect on the free energy of the total system at low temperatures, providing an additional term that can help make the spin-density wave ground states more stable in a magnetic field. Such effects should occur in all spin-density wave systems with imperfect nesting, including elemental chromium, but are more pronounced in systems where the transition temperature into the spin-density wave phase is low, or where the electronic bands are highly anisotropic. By far, the most striking behavior occurs in systems with an incipient spin-density wave, as has been shown to be the case in certain examples of the Bechgaard salts. "Incipient" here means that the Coulomb interactions are present to the extent that the metal is on the threshold of forming a spin-density wave phase, but where the efficiency of the nesting is insufficient to make a spin-density wave ground state stable at zero magnetic field. The system then has the potential to undergo transitions into spin-density wave phases in a magnetic field; hence, the term "field-induced spin-density waves." At zero magnetic field, the ground state can be a normal metal or a superconductor. The spin-density wave phase supersedes superconductivity in a magnetic field, because superconductivity is inhibited by magnetic fields.

One way to understand the origin of field-induced spin-density waves is through the process of one dimensionalization of the electron motion in a magnetic field. A simple model to consider is that where the k_y component of the electronic dispersion given by eqn [1] has the form

$$\varepsilon_\perp(k_y) = -2t_b \cos k_y b - 2t'_b \cos 2k_y b \qquad [25]$$

where $t_b < W/4$ and $t'_b < t_b$ are the nearest neighbor and second nearest neighbor electron transfer energies along the b-axis of a crystal. As shown in **Figure 4a**, the Fermi surface can no longer be nested by $\boldsymbol{Q}_0 = (2k_F, 0)$ or $\boldsymbol{Q}_0 = (2k_F, \pi/b)$ owing to the finite values of the transfer integrals along b. A magnetic field causes electrons on the 2D Fermi surface to follow a snake-like path similar to that in **Figure 4b**, where its real space width $4bt_b/\hbar\omega_b$ and period $\ell = 2\pi v_F/\omega_b$ decrease inversely proportionately with the frequency $\omega_b = eBv_F b/\hbar$, and hence the magnetic field B. One can make the qualitative argument that the motion becomes 1D as soon as $4t_b < \hbar\omega_b$, whereupon quasiparticles are mostly confined to single 1D chains.

From the quantum mechanical perspective, the one-dimensionalization can be understood by making a Landau–Peierls substitution, $\boldsymbol{k} \rightarrow i\nabla - e\boldsymbol{A}/\hbar$, into the

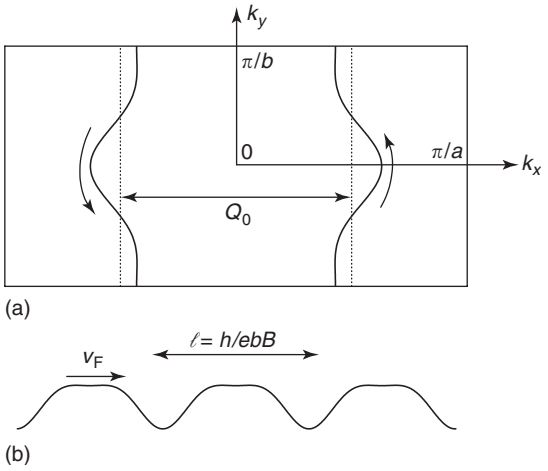


Figure 4 (a) A sketch of a Fermi surface similar to that seen in the Bechgaard salts, in which the nesting at vector $\boldsymbol{Q}_0$ is imperfect. The dotted lines show what the Fermi surface looks like when $\varepsilon_\perp(k_y) = 0$, whereupon perfect nesting is recovered. Arrows indicate the direction of motion of quasiparticles on the Fermi surface for a magnetic field pointing out of the page. (b) The trajectory of a quasiparticle in real space, obtained through the equation of motion, $\hbar\partial k/\partial t = e\boldsymbol{v}_F \times \boldsymbol{B}$, where $\boldsymbol{v}_F = \nabla_k \varepsilon$. The spatial dimensions of the real space trajectory are inversely proportional to the magnetic field strength. The product ℓb encloses a single flux quantum.

dispersion relation given by eqns [1] and [25], which yields a 1D Schrödinger equation

$$\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) - 2t_b \cos\left(k_y b - \frac{2\pi}{\ell}x\right) - 2t'_b \times \cos\left(2k_y b - \frac{4\pi}{\ell}x\right) = \varepsilon\psi(x) \qquad [26]$$

that is a function of only x where k_y then merely determines the shift in the coordinate of the center of the wave function.

The spin-density wave formation now involves the nesting of ε as determined by eqn [26]. In spite of the added complexity, the thermodynamics is shown to be surprisingly similar to that of a conventional spin-density wave ground state. Notably, eqns [4],[5], and [7] continue to be applicable but with the effective coupling constant having a modified form

$$\lambda = Un(\varepsilon_F)J_0^2\left(\frac{2t'_b}{\omega_b}\right) \qquad [27]$$

where J_0 is a zeroth-order Bessel function. The electronic density of states continues to be fully gapped, but with 2Δ also functioning as a Landau gap between Landau bands dispersed with respect to k_x, in which the orbital motion of quasiparticles is quantized. The orbital angular momentum of each quasiparticle in each Landau band is given by

$\hbar(l+1/2)$, where the quantum number has values $l=0,1,2,\ldots$, as for Landau levels in a simple 2D electron gas. The Landau bands are each separated in energy by an effective cyclotron energy $\hbar\omega_c$ that depends on the details of the nested band structure. By making an analogy with 2D electron gas systems, one can gain further insight into understanding as to why imperfectly nested spin-density waves are stabilized by a magnetic field. The free energy of a 2D electron gas experiences sharp minima as a function of B whenever ε_F is situated in a Landau gap (i.e., somewhere in the gap between two adjacent Landau levels, l and $l+1$). It is by utilizing such minima between Landau bands that field-induced spin-density waves become stable. This, therefore, exerts a constraint by which ε_F must always be situated within a Landau gap in a field-induced spin-density wave phase.

The above constraint has two important implications for field-induced spin-density wave systems. The first is that the Landau bands are always completely filled, giving rise to a type of integer quantum Hall effect where

$$\sigma_{xy} = \pm\frac{2ve^2}{h} \qquad [28]$$

per layer, v is the quantum number of the highest occupied Landau level and the sign depends on whether one or two closed pockets contain a net number of electrons or holes. The second implication is that the nesting vector is quantized, where

$$Q_0 = \left(2k_F \pm \frac{2\pi v}{\ell}, \frac{\pi}{b} + \delta k_y\right) \qquad [29]$$

in the simplest case, in order to sustain an integer filling factor v. This type of quantum Hall effect differs from that in conventional 2D electron gas systems in several ways. The first is the factor of 2 in the numerator of eqn [28], which results from the fact that up- and down-spin states are paired with no net spin quantum numbers, giving rise to completely spin-degenerate Landau bands. The second is that the net slope of the Hall resistivity $\rho_{xy} \approx 1/\sigma_{xy}$ throughout the field-induced spin-density wave phases, depicted in **Figure 5**, does not intersect $\rho_{xy} = B = 0$. Finally, the quantum Hall effect does not depend on the existence of localized states in the tails of Landau levels as it does in conventional 2D electron gas systems. The inability of ε_F to reside in a Landau band implies that a change in the value of the quantized Hall conductivity must be realized by means of a first-order phase transition between phases with different optimal values of v. An alternative possibility is that the normal metallic phase is restored between stable Hall plateaux. Orbital quantization within field-induced spin-density wave phases causes the off-diagonal components of the conductivity tensor to dominate over the diagonal ones, giving rise to a low resistivity state, $\rho_{xx} \sim \sigma_{xx}/\sigma_{xy}^2 \to 0$, as in the conventional quantum Hall effect. This contrasts with the semiconductor-like behavior seen in perfectly nested spin-density wave phases.

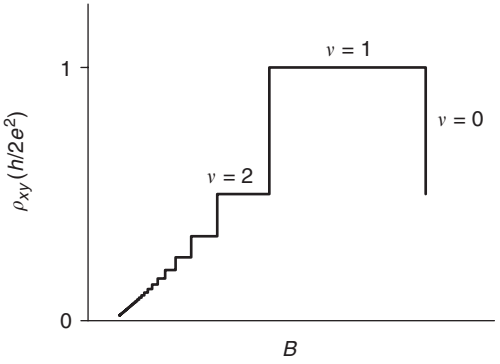


Figure 5 A plot of the Hall resistivity (per layer) of a field-induced spin-density wave system vs. magnetic field, exhibiting quantized Hall plateaux. Sharp changes are expected to occur between different states characterized by different values of v owing to the first-order nature of the transitions.

See also: Fermi Surface Measurements; Spin Density Waves and Magnons.

PACS: 75.30.Fv; 75.10.Lp; 2.15.Nj

Further Reading

Blundell S (2001) *Magnetism in Condensed Matter*. Oxford: Oxford University Press.

Chaikin PM (1996) Field induced spin density waves. *Journal Physics I France* 6: 1875–1898.

Chakraborty T and Pietiläinen P (1995) *The Quantum Hall Effects: Integral and Fractional*. Berlin: Springer.

Fawcett E (1988) Spin-density wave antiferromagnetism in chromium. *Reviews of Modern Physics* 60: 209–283.

Fawcett E, Alberts HL, Galkin VY, Noakes DR, and Yakhmi JV (1994) Spin-density wave antiferromagnetism in chromium alloys. *Reviews of Modern Physics* 66: 25–127.

Gabovich AM and Voitenko AI (2000) Superconductors with charge- and spin-density waves: theory and experiment (review). *Low Temperature Physics* 26: 305–330.

Gruner G (1994) *Density Waves in Solids, Frontiers in Physics.*, vol. 89. Reading, MA: Addison-Wesley.

Gruner G (1996) The dynamics of spin density waves. *Reviews of Modern Physics* 66: 1–24.

Kagoshima S, Nagasawa H, and Sambongi T (1988) *One-dimensional Conductors. vol. 72 of Solid State Sciences*. Berlin: Springer.

Kulikov NI and Tugushev VV (1984) Spin-density waves and itinerant antiferromagnetism in metals. *Uspehkhi Fizicheskikh Nauk* 144: 643–680.

Maki K (1986) Thermodynamics of field-induced spin-density-waves states in Bechgaard salts. *Physical Review B* 33: 4826–4829.
Roth S (1995) *One-Dimensional Metals. Physics and Materials Science*. Weinheim: VCH Verlagsges.

Nomenclature

a or b	lattice constant along $\boldsymbol{x}$ or $\boldsymbol{y}$ (m)
A or B	arbitrary integers (dimensionless)
$\boldsymbol{B}$ or B	magnetic flux density (magnetic field) (T)
$\boldsymbol{B}_{\text{eff}}$	Heisenberg molecular field (T)
D	anisotropy energy (J, eV)
e	electron charge (C)
E	electric field (Vm^{-1})
g	electronic g-factor (dimensionless)
h	Planck constant (J s)
$\hbar$	Planck constant divided by 2π(J s)
H	Heisenberg spin Hamiltonian (J, eV)
$H_{\perp}$	anisotropy Hamiltonian (J, eV)
i	square root of –1 (dimensionless)
j_{SDW}	current of collective mode (C s^{-1})
J	Heisenberg interaction (J, eV)
$\boldsymbol{k}$	electronic wave vector (m^{-1})
k_{F}	Fermi wave vector (m^{-1})
k_x, k_y, or k_z	electronic wave vector along $\boldsymbol{x}$, $\boldsymbol{y}$, or $\boldsymbol{z}$ (m^{-1})
l	Landau level index (dimensionless)
ℓ	one-dimensional Fermi surface traversal length (m)
m	quasiparticle effective mass (kg)
$\boldsymbol{m}$ or μ	moment of spin (Am^2 or units of μ_{B})
n	carrier density (3D) (m^{-3})
$n(\varepsilon)$	electronic density of states per unit cell (J^{-1}, eV^{-1})
N	number of electrons per unit cell (dimensionless)
p or q	indices of spin (dimensionless)
Q	quasiparticle excitation wave vector (m^{-1})
$\boldsymbol{Q}_0$ or Q_0	optimum nesting vector (m^{-1})
$\boldsymbol{S}$ or S	spin per unit cell (dimensionless)
t	time (s)
t_b	transfer integral along b (nearest neighbor, second nearest) (J, eV)
T	temperature (K)
U	Coulomb potential (J, eV)
$\boldsymbol{v}_{\text{F}}$ or v_{F}	quasiparticle group velocity at Fermi surface (m s^{-1})
W	electronic bandwidth (J, eV)
x	spatial dimension along $\boldsymbol{x}$ (m)
2Δ	gap energy (J, eV)
ε	electronic dispersion (J, eV)
ε_{F}	Fermi energy (J, eV)
$\varepsilon_{\perp}$	electronic dispersion orthogonal to $\boldsymbol{x}$ (J, eV)
Λ	period of spin modulation (m)
λ	effective coupling constant (dimensionless)
ρ	charge density per unit length (C m^{-1})
ν	Landau level filling factor (dimensionless)
ρ_{xy}	Hall resistivity (Ω m)
σ	electronic spin (dimensionless)
σ_{coll}	conductivity of collective mode (Ω^{-1} m^{-1})
σ_{xx}	conductivity (diagonal) ($\Omega^{-1}\text{m}^{-1}$)
σ_{xy}	Hall conductivity (Ω^{-1} m^{-1})
τ	relaxation time (transport) (s)
Φ	free energy per unit cell (J^{-1}, eV^{-1})
ϕ	phase (radians)
χ	spin susceptibility (dimensionless)
ψ	electronic wave function (dimensionless)
ω_0	pinning frequency (s^{-1})
ω_{A}	amplitudon frequency (s^{-1})
ω_b	one-dimensional Fermi surface traversal frequency (s^{-1})
ω_{m}	magnon frequency (s^{-1})
ω_{ϕ}	phason frequency (s^{-1})

Statistical Mechanics: Classical

D Cassi, Università di Parma, Parma, Italy

Introduction

Statistical mechanics was created in the second half of the nineteenth century as a branch of theoretical physics with the purpose of deriving the laws of thermodynamic systems from the equations of motion of their elementary constituents, that is, atoms and molecules.

It is commonly recognized that statistical mechanics originated principally through the efforts of Rudolf Clausius (1822–1888), who first interpreted heat as the kinetic energy of molecules, James Clerk Maxwell (1831–1879), who derived the velocity distribution of the molecules in an ideal gas and, most importantly, Ludwig Boltzmann (1844–1906), who gave a statistical meaning to entropy.

Statistical mechanics was also the field where quantum physics first revealed itself: indeed, in 1900, Max Planck proposed the quantization hypothesis and gave a statistical interpretation of the black body radiation. Only two years later, Josiah Willard Gibbs

(1839–1903) introduced the statistical-ensembles approach, which is commonly used today and allows one to extend statistical methods to a great variety of physical systems.

In the 1920s, the application of statistical methods to quantum mechanical systems gave rise to a substantially new discipline which was called quantum statistical mechanics; the theory in use up to that time, and its further developments – not involving quantum mechanics (e.g., the theory of phase transitions and critical phenomena) – were accordingly called classical statistical mechanics.

Usually, statistical mechanics deals with equilibrium phenomena and it should be called, more precisely, equilibrium statistical mechanics. Nonequilibrium phenomena are typically more difficult to study, but several important statistical treatments of them exist, starting with the famous and seminal Boltzmann equation. These topics of nonequilibrium statistical mechanics are generally discussed as specific subjects directly related to their applications (e.g., transport phenomena, relaxation phenomena). In the following sections, the basics of classical equilibrium statistical mechanics are introduced.

Macroscopic and Microscopic Variables

The physical systems studied by statistical mechanics are composed of a very large (typically 10^{23}) number (N) of particles or degrees of freedom. In classical mechanics, the (microscopic) state of a system is univocally identified by giving the generalized positions q and momenta p of all the constituents.

For the sake of clarity the following will focus on a system composed of N pointlike particles. The $6N$-tuple of all momenta and positions are denoted by (p, q) (or, when it is convenient to explicitly specify the number of particles, by $(p^{(N)}, q^{(N)})$), defining

$$\begin{aligned}(p,q) &= (p^{(N)}, q^{(N)}) \\ &= (p_{1x}, p_{1y}, p_{1z}, \ldots, p_{Nx}, p_{Ny}, p_{Nz}, \\ &\qquad q_{1x}, q_{1y}, q_{1z}, \ldots, q_{Nx}, q_{Ny}, q_{Nz})\end{aligned}$$

The microstate (p, q) is, therefore, a point in the $6N$-dimensional phase space Ξ_N, which is the set of all possible physical states of the system.

According to classical dynamics, the time evolution of a system is determined by its Hamiltonian $H(p, q)$, and its coordinates $(p(t),q(t))$ satisfy the Hamilton equations

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}, \qquad \frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}$$

On the other hand, the macroscopic state (macrostate) of a system, which is the object of thermodynamics, is described by a very small number of macroscopic variables (e.g., volume, pressure, temperature); therefore, it corresponds to many different microscopic states. The macroscopic variables appearing in thermodynamic laws can be divided into two different categories: the mechanical (or, in more general systems, the magnetic, the electric, etc.) and the statistical (or thermal) ones. The former can be calculated even for a single microstate from its coordinates (p, q), that is, they are functions, $f(p, q)$, of the points of the phase space (e.g., the internal energy). The latter cannot be determined for a single microstate, but all the microstates belonging to a given macrostate are needed in order to calculate them, that is, they are functions of subsets of the phase space (e.g., the entropy).

Time Averages

A fundamental assumption of classical statistical mechanics concerns the characteristic times of microscopic versus macroscopic dynamics. It is assumed that the microscopic dynamics is much faster than the macroscopic one, so that during a macroscopic measurement (the measurement of a macroscopic variable) the system assumes many different microstates (indeed, for continuous coordinates, infinite microstates). Therefore, the measured value of a mechanical macroscopic variable M is to be considered the time average $\bar{m}_t$ of the corresponding microscopic one m:

$$\bar{m}_t = \frac{1}{t}\int_0^t m(\tau)\,\mathrm{d}\tau \qquad [1]$$

t being the measurement duration. Considering that t is typically very large with respect to all characteristic times involved in microscopic dynamics, one takes the $t \to \infty$ limit, defining the macroscopic variable M as

$$M = \lim_{t\to\infty} \bar{m}_t \qquad [2]$$

Even with this simplifying assumption, the problem of determining M would require the solution of the equations of motion for the system, unless some further hypotheses are introduced. These are connected with the statistical equilibrium condition, that is, to the experimentally observed time independence of macroscopic variables: first, it must be supposed that the limit in [2] exists; then, M is required to be independent of the specific microstate (belonging to the given macrostate) assumed by the system at the instant $t = 0$ (formally the beginning of the measurement).

Phase-Space Averages

A crucial technical step in the development of classical statistical mechanics is the passage from time averages to phase-space averages. To introduce it, it is considered that in classical mechanics every mechanical variable m is a function, $m(p, q)$, of the phase-space coordinates p and q. Therefore, to calculate time averages one could simply consider the number of times the system has visited each point in phase space (or, more precisely, an infinitesimal volume $\mathrm{d}p\,\mathrm{d}q$ around that point); associating with every point (p, q) a weight $\rho(p, q)$ proportional to the number of visits, one can then obtain the time average of any variable m by computing its weighted average on the phase-space points. These heuristic considerations can be mathematically formalized as follows (assuming all the conditions for the interchange of the integrals and the limit hold)

$$\begin{aligned} M &= \lim_{t\to\infty} \frac{1}{t}\int_0^t m(p(\tau), q(\tau))\mathrm{d}\tau \\ &= \lim_{t\to\infty} \frac{1}{t}\int_0^t \left[\int m(p,q)\delta(p-p(\tau)) \right. \\ &\quad \left. \times\delta(q-q(\tau))\mathrm{d}p\ \mathrm{d}q\right]\mathrm{d}\tau \\ &= \int m(p,q)\left[\lim_{t\to\infty}\frac{1}{t}\int_0^t \delta(p-p(\tau))\right. \\ &\quad \left. \times\delta(q-q(\tau))\mathrm{d}\tau\right]\mathrm{d}p\ \mathrm{d}q \\ &\equiv \int m(p,q)\rho(p,q)\mathrm{d}p\ \mathrm{d}q \end{aligned} \qquad [3]$$

where

$$\rho(p,q) \equiv \lim_{t\to\infty}\frac{1}{t}\int_0^t \delta(p-p(\tau))\delta(q-q(\tau))\mathrm{d}\tau$$
$$\text{with } \mathrm{d}p \equiv \mathrm{d}p_{1x}\cdots\mathrm{d}p_{Nz} \text{ and } \mathrm{d}q \equiv \mathrm{d}q_{1x}\cdots\mathrm{d}q_{Nz} \qquad [4]$$

Notice that the introduction of $\rho(p, q)$ allows one to calculate the average values of any mechanical variable without knowing its dynamics: in other words, once $\rho(p, q)$ is known, the time can be completely neglected in all further calculations.

From the requirements introduced at the end of the last section, and from the generality of M, it follows that the function $\rho(p, q)$ only depends on, and completely determines, the macrostate it corresponds to. Therefore, it must parametrically depend only on the restricted number of macroscopic variables defining that macrostate. The same requirements would imply that, for any starting point (p, q) belonging to the macrostate, the system visits all the microstates belonging to it and that the frequency of visits is $\rho(p, q)$. This property cannot be proved for a generic dynamical system, and it has to be considered as one of the fundamental assumptions of classical statistical mechanics.

The Statistical Ensembles

The calculation of $\rho(p, q)$ from its definition would imply the complete solution of the equations of motion with any initial condition. This is in practice impossible even for simple systems, due to the large number of degrees of freedom. In fact, $\rho(p, q)$ was determined in a general form once and for all by Gibbs for the most significant cases, on the basis of physical considerations and assumptions.

Gibbs followed an alternative approach. Instead of just one system, he considered a large set of physically identical systems at once, all having the same Hamiltonian but being in different microstates at a given time t, and he assumed that the representative points (p, q) in phase space were distributed with a density $\rho(p, q)$. He called such a set a statistical ensemble and $\rho(p, q)$ its distribution. In other words, each system in the ensemble is, at time t, in a microscopic state (p, q) corresponding to the macroscopic state described by $\rho(p, q)$, with a probability given by $\rho(p, q)$ itself. Then he imposed the statistical equilibrium condition by requiring that the density of the points representing the systems in the phase space was left invariant during their time evolution. Finally, he proved that this condition is fulfilled if and only if

$$\rho(p,q) = f(H(p,q)) \qquad [5]$$

that is, if and only if $\rho(p, q)$ depends on the phase-space coordinates p and q only through the Hamiltonian of the system.

Now, considering the typical macroscopic constraints characterizing a thermodynamic system, he identified three main general kinds of ensembles and, using [5] and some specific hypotheses, he calculated their distributions. Gibbs called these ensembles microcanonical, canonical, and grand canonical, and these names have been used ever since.

The Microcanonical Ensemble

Even though Gibbs first considered the canonical ensemble and then derived the microcanonical one from it, the following formal development of statistical mechanics suggested inverting the order, to have a stronger connection with the microscopic dynamical equations.

The microcanonical ensemble describes a closed isolated system, that is, a system which cannot exchange particles and energy with the external world, confined in a fixed volume. Therefore, the macroscopic variables characterizing it are the number of particles N, the volume V, and the total energy E. Notice that all of them are mechanical variables: indeed, the microcanonical ensemble is the statistical

ensemble which is most directly connected with a classical dynamical system. The variables N, V, and E are also all extensive, that is, they are asymptotically proportional to the number of degrees of freedom (usually $3N$) for large N. Many physical properties of thermodynamic systems turn out to be simpler by taking the thermodynamic limit $N \to \infty$; in that case, it shall be required that the ratios of V and E to N (called the specific volume and the specific energy, respectively) are constant in this limit:

$$\lim_{N\to\infty} \frac{V}{N} = v, \qquad \lim_{N\to\infty} \frac{E}{N} = \varepsilon \qquad [6]$$

Due to its defining conditions, the energy of each system belonging to the microcanonical ensemble is constant during time evolution; in other words, its trajectory in phase space entirely lies on the hypersurface $H(p, q) = E$. The microcanonical distribution is obtained assuming that $\rho(p, q)$ is constant on this hypersurface and zero elsewhere:

$$\rho^{\mu C}_{N,V,E}(p,q) = \frac{\varepsilon_0 \delta(E - H(p,q))}{\Gamma(N,V,E)} \qquad [7]$$

where

$$\Gamma(N,V,E) = \varepsilon_0 \int \delta(E - H(p,q)) \mathrm{d}p \; \mathrm{d}q \qquad [8]$$

is the measure of the region of the phase space accessible to the system, that is, it measures how large the set of microstates corresponding to the given macrostate identified by N, V, and E is, and ε_0 is an arbitrary but fixed constant with the dimension of an energy, which is convenient to introduce to give Γ the dimension of a volume in phase space (without loss of generality, it can be set equal to 1).

The assumption leading to [7] means that the system, in its motion, explores "in a uniform way" the whole constant-energy hypersurface. Many attempts have been made during the last century to mathematically prove this hypothesis to be a consequence of general properties of Hamiltonian dynamical systems. However, this problem, known as the ergodic problem, has not been solved but for a very restricted class of systems (typically hard-spheres gases), and the ergodic hypothesis will be considered to be a reasonable assumption leading to experimentally verifiable results.

Knowing the microcanonical distribution $\rho^{\mu C}_{N,V,E}(p,q)$, it is possible to calculate the average of any mechanical variable according to [3]. However, there is no rule to calculate the fundamental thermal variables, such as the entropy S and the temperature T. In order to connect the microcanonical distribution with the thermodynamics, the most important postulate of all statistical mechanics has to be introduced, namely, the Boltzmann relation, which derives the entropy from Γ:

$$S(N,V,E) = k_B \ln \frac{\Gamma(N,V,E)}{\Gamma_0} \qquad [9]$$

where $k_B = 1.3807 \times 10^{-23} \mathrm{J K^{-1}}$ is the Boltzmann constant and Γ_0 is a fixed but arbitrary number with the same dimensions as Γ. It represents a unit of measure for volumes in phase space and it is necessary to make the argument of the logarithm adimensional. It is to be noted that it does not make the definition [9] ambiguous, since the entropy, like every thermodynamic potential, is always defined up to an arbitrary additive constant. As is well known, this arbitrariness can be eliminated by introducing the so-called third principle of thermodynamics, or the Nernst postulate, according to which the entropy must vanish for $T \to 0$: this requirement univocally fixes the value of Γ_0. Moreover, in order to have a coherent definition of Γ_0 for different values of N when taking the thermodynamic limit, it is customary to choose once for all the unit of measure γ for the smallest phase space, corresponding to only one particle in one dimension, and then define $\Gamma_0(N)$ for a system of N particles (in three dimensions) as

$$\Gamma_0(N) = \gamma^{3N} \qquad [10]$$

Considerations based on the semiclassical limit of quantum mechanics suggest then to choose $\gamma = h$, where h is the Planck constant, but this is outside the proper domain of classical statistical mechanics.

The Boltzmann relation cannot be mathematically proved and one has to consider it as a postulate. All that can be done is proving that, under some restrictive hypotheses, the statistical entropy defined by [9] has the same mathematical properties as the thermodynamic entropy defined by the differential relation $\mathrm{d}S = \mathrm{d}Q/T$. The fundamental of these hypotheses concerns the interactions between the particles in the system: they must decay rather fast with increasing distance. More precisely, suppose the interactions are isotropic, for the sake of simplicity, the interaction potentials $V(r)$ must decay faster than r^{-2} as $r \to \infty$. This means that equilibrium statistical mechanics cannot be applied to systems with (non-negligible) gravitational interactions, as well as to systems containing electrically charged particles and having nonzero total charge.

An important modification to [9] has to be introduced if the particles of the system are indistinguishable. Indeed, many years before the quantum concept of identical particles was introduced, Gibbs proved that Γ needs to be divided by $N!$ in order to preserve the asymptotic additivity of the entropy (his argument

is commonly know as the Gibbs paradox):

$$S(N, V, E) = k_B \ln \frac{\Gamma(N, V, E)}{\Gamma_0 N!} \quad [11]$$

In the most general case, not all the particles are indistinguishable, but one can have N_1 indistinguishable particles of one kind, N_2 of a different kind, and so on. If the system is composed of m different kinds of indistinguishable particles, with $\sum_{i=1}^{m} N_i = N$ (allowing even $N_i = 1$ for the sake of generality), the most general expression of the entropy is given by

$$S(N_1, \ldots, N_m, V, E) = k_B \ln \frac{\Gamma(N_1, \ldots, N_m, V, E)}{\Gamma_0 N_1! \cdots N_m!} \quad [12]$$

A slightly different definition of the microcanonical ensemble can be found in literature, according to which the systems belonging to the ensemble have energy between E and $E + \Delta E$ instead of exactly E. Assuming a constant distribution in the entire phase-space region satisfying that constraint, and defining Γ to be the volume of this region, one can again define S using [12]. It can be shown that in the thermodynamic limit, the two definitions of the entropy are asymptotically equivalent and the corresponding distributions give the same thermodynamic behavior.

Equation [12] is the starting point for deriving all the thermodynamics of the microcanonical ensemble. Applying basic thermodynamic relations, one obtains

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{V, N_1, \ldots, N_m} \quad [13]$$

$$\frac{P}{T} = \left.\frac{\partial S}{\partial V}\right|_{E, N_1, \ldots, N_m} \quad [14]$$

$$\frac{\mu_i}{T} = -\left.\frac{\partial S}{\partial N_i}\right|_{V, E} \quad [15]$$

giving the temperature, the pressure, and the chemical potentials, respectively.

The Canonical Ensemble

The most typical system studied in statistical mechanics is confined in a fixed volume and is closed, but it is not isolated and it can exchange energy with a thermal bath (or a reservoir) at a fixed temperature. It is described by the canonical ensemble, and the macroscopic variables characterizing it are the number of particles N, the volume V, and the temperature T.

It can be considered as a small but macroscopic subsystem of a closed isolated system, the remaining part of which acts as a reservoir. Therefore, the canonical distribution from the microcanonical one can be obtained, calculating the distribution of a closed subsystem whose volume and number of particles in the thermodynamic limit satisfy

$$\lim_{N\to\infty} \frac{V}{N} = v,$$
$$\lim_{N\to\infty} \frac{N}{N_{TOT}} = \lim_{N\to\infty} \frac{V}{V_{TOT}} = 0 \quad [16]$$

N_{TOT} and V_{TOT} being the number of particles and the volume of the whole closed isolated system. Introducing, as usual, some suitable hypotheses (e.g., the ratio between the external surface of the subsystem and its volume must vanish in the thermodynamic limit) after some calculations, one gets

$$\rho^{C}_{N,V,T}(p, q) = \frac{e^{-H(p,q)/k_B T}}{Z(N, V, T)} \quad [17]$$

where

$$Z(N, V, T) = \int e^{-H(p,q)/k_B T} dp\, dq \quad [18]$$

is called the canonical partition function.

For the canonical ensemble, Z plays a role analogous to Γ for the microcanonical one: it is a sort of normalization factor which is not directly measurable, but it is the starting point for deriving all the thermodynamic quantities. The basic relation connecting it to a thermodynamic potential is

$$F(N, V, T) = -k_B T \ln \frac{Z(N, V, T)}{Z_0} \quad [19]$$

F being the Helmholtz free energy and Z_0 an arbitrary constant with the same dimensions as Z, playing the same role of Γ_0 in [9]; indeed, F too is defined up to an additive constant like S and the arbitrariness can be eliminated in the same way by assuming the third principle of thermodynamics to hold. Moreover, the same considerations introduced for Γ_0, concerning the normalization in the thermodynamic limit and the presence of indistinguishable particles, apply. Accordingly, the general form of [19] becomes

$$F(N_1, \ldots, N_m, V, T) = -k_B T \ln \frac{Z(N_1, \ldots, N_m, V, T)}{Z_0(N) N_1! \cdots N_m!} \quad [20]$$

with

$$Z_0(N) = \gamma^{3N} \quad [21]$$

After determining F by [20], all the thermodynamics of the canonical ensemble can be easily deduced. The

main thermodynamic equations needed are

$$S = -\left.\frac{\partial F}{\partial T}\right|_{V,\, N_1,\ldots,\, N_m} \quad [22]$$

$$P = -\left.\frac{\partial F}{\partial V}\right|_{T,\, N_1,\ldots,\, N_m} \quad [23]$$

$$\mu_i = \left.\frac{\partial F}{\partial N_i}\right|_{V,\, T} \quad [24]$$

$$E = F + TS \quad [25]$$

where E denotes the internal energy and coincides with the average value of the Hamiltonian.

The Grand Canonical Ensemble

The case of a system occupying a fixed volume but exchanging particles and energy with the external world, at fixed temperature and chemical potentials, is described by the grand canonical ensemble. In the simplest situation, all the particles are of the same kind and a macrostate is as usual defined by three macroscopic variables, the volume V, the temperature T, and the chemical potential μ (representing the energy increment due to the addition of a particle). However, a microstate is no longer represented by a point in a usual phase space. Indeed, the dimension of the phase space depends on the number of particles N, and here N is variable.

To consider this problem, one has to introduce an extended space Ξ given by the union of the phase spaces Ξ_N corresponding to all possible N:

$$\Xi = \bigcup_{N=0}^{\infty} \Xi_N \quad [26]$$

A point in this space is given by the $(6N+1)$-tuple $(N, p^{(N)}, q^{(N)})$, that is, by the number of particles N and a point in the corresponding phase space Ξ_N (note that one has to include the case $N=0$, corresponding to an empty system). For any N, one has also to consider a specific Hamiltonian $H_N(p^{(N)}, q^{(N)})$ and, to have a coherent definition of volumes in phase spaces of different dimensions, the volume unit $Z_0(N) = \gamma^{3N}$ is chosen according to [21].

To obtain the grand canonical distribution, a procedure similar to the one used to derive the canonical distribution from the microcanonical one can be followed.

A small but macroscopic subsystem of a closed isolated system is considered, the remaining part of which acts as a reservoir of energy and particles. The subsystem is open (it can exchange particles and energy with the external world), but it occupies a fixed volume V in space. It is assumed that, taking the thermodynamic limit,

$$\lim_{V \to \infty} \frac{V}{V_{\mathrm{TOT}}} = 0 \quad [27]$$

where V_{TOT} is the volume of the whole closed isolated system, and that the ratio between the external surface of the subsystem and its volume also vanishes in the thermodynamic limit. After some algebra, one obtains

$$\rho^{\mathrm{GC}}_{\mu,V,T}(N, p^{(N)}, q^{(N)}) = \frac{\mathrm{e}^{(\mu N - H_N(p^{(N)}, q^{(N)}))/k_{\mathrm{B}}T}}{\gamma^{3N} N! Z^{\mathrm{GC}}(\mu, V, T)} \quad [28]$$

where

$$Z^{\mathrm{GC}}(\mu, V, T) = \sum_{N=0}^{\infty} \frac{\mathrm{e}^{\mu N/k_{\mathrm{B}}T}}{\gamma^{3N} N!} \times \int \mathrm{e}^{-(H_N(p^{(N)}, q^{(N)})/k_{\mathrm{B}}T)} \mathrm{d}p^{(N)}\, \mathrm{d}q^{(N)} \quad [29]$$

is the grand canonical partition function or grand partition function. The last expression can also be written as

$$Z^{\mathrm{GC}}(\mu, V, T) = \sum_{N=0}^{\infty} \frac{z^N Z(N, V, T)}{\gamma^{3N} N!} \quad [30]$$

in terms of the canonical partition functions $Z(N, V, T)$ and of the activity z, defined by

$$z = \mathrm{e}^{\mu/k_{\mathrm{B}}T} \quad [31]$$

In the most general case of a system composed of m different kinds of indistinguishable particles, a different chemical potential for each kind of particle has to be introduced; therefore, the macroscopic variables characterizing a macrostate are T, V, and the m chemical potentials $\mu_1, \ldots, \mu_m$.

As for the microstates, the number N_i of particles of each kind $i = 1, \ldots, m$ has to be specified, and the corresponding phase space $\Xi_{N_1,\ldots,N_m}$ has to be introduced with coordinates $(p^{(N_1,\ldots,N_m)}, q^{(N_1,\ldots,N_m)})$. Accordingly, the extended space Ξ of [26] becomes

$$\Xi = \bigcup_{N_1=0}^{\infty} \cdots \bigcup_{N_m=0}^{\infty} \Xi_{N_1,\ldots,N_m} \quad [32]$$

and a point of Ξ is specified by the coordinates $(N_1, \ldots, N_m, p^{(N_1,\ldots,N_m)}, q^{(N_1,\ldots,N_m)})$, while the Hamiltonian acting on it is $H_{N_1,\ldots,N_m}$ $(p^{(N_1,\ldots,N_m)}, q^{(N_1,\ldots,N_m)})$. Using this notation, the grand canonical

partition function is given by

$$\rho^{\mathrm{GC}}_{\mu_1,\ldots,\mu_m,V,T}(N_1,\ldots,N_m,\ p^{(N_1,\ldots,N_m)},q^{(N_1,\ldots,N_m)}) = \frac{\mathrm{e}^{\left(\sum_{i=1}^{m}\mu_i N_i - H_{N_1,\ldots,N_m}(p^{(N_1,\ldots,N_m)},q^{(N_1,\ldots,N_m)})\right)/k_{\mathrm{B}}T}}{\gamma^{3\sum_{i=1}^{m}N_i}N_1!\cdots N_m!Z^{\mathrm{GC}}(\mu_1,\ldots,\mu_m,V,T)} \quad [33]$$

with

$$Z^{\mathrm{GC}}(\mu_1,\ldots,\mu_m,V,T) = \sum_{N_1=0}^{\infty}\cdots\sum_{N_m=0}^{\infty}\prod_{i=1}^{m}\frac{z_i^{N_i}}{\gamma^{3N_i}N_i!} \times\int \mathrm{e}^{-(H_{N_1,\ldots,N_m}(p^{(N_1,\ldots,N_m)},q^{(N_1,\ldots,N_m)}))/k_{\mathrm{B}}T} \times \mathrm{d}p^{(N_1,\ldots,N_m)}\ \mathrm{d}q^{(N_1,\ldots,N_m)} \quad [34]$$

and

$$z_i = \mathrm{e}^{\mu_i/k_{\mathrm{B}}T} \quad [35]$$

The thermodynamics of the grand canonical ensemble can be derived from the basic equation

$$\Omega(\mu_1,\ldots,\mu_m,V,T) = -k_{\mathrm{B}}T\ \ln\ Z^{\mathrm{GC}}(\mu_1,\ldots,\mu_m,V,T) \quad [36]$$

where the thermodynamic potential Ω is defined by

$$\Omega = -PV \quad [37]$$

The other thermodynamic quantities are obtained from Ω applying the relations

$$S = -\left.\frac{\partial\Omega}{\partial T}\right|_{V,\mu_1,\ldots,\mu_m} \quad [38]$$

$$N_i = -\left.\frac{\partial\Omega}{\partial\mu_i}\right|_{V,T} \quad [39]$$

$$E = \Omega + TS + \sum_{i=1}^{m}\mu_i N_i \quad [40]$$

where N_i denotes the average number of particles of kind i.

The Perfect Gas

The simplest and best-known application of classical statistical mechanics concerns the so-called perfect gas, that is, a system of noninteracting particles. This is obviously an ideal case, since a complete absence of interaction would prevent a system from reaching the equilibrium, but it can be considered a good approximation for rarefied real gases, where the potential energy is negligible with respect to the kinetic energy. It also represents the natural starting point for the study of more complicated systems by perturbative techniques.

Choosing, for the sake of simplicity, a monatomic gas composed of indistinguishable pointlike particles of only one kind, the perfect gas Hamiltonian is

$$H_N(p^{(N)},q^{(N)}) = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m} \quad [41]$$

with $\boldsymbol{p}_i = (p_{ix}, p_{iy}, p_{iz})$ and m being the mass of each particle.

Microcanonical Ensemble

The case of a closed isolated system is studied first and described by the microcanonical ensemble. It is supposed that the gas is confined in a fixed space volume V, with a fixed energy E, and number of particles N.

To calculate the thermodynamic properties of such a system, first its entropy is to be determined according to [11], and to do that, the measure $\Gamma(N, V, E)$ of the region of the phase space accessible to the system is needed:

$$\Gamma(N,V,E) = \varepsilon_0\int\delta(E-H(p,q))\mathrm{d}p\ \mathrm{d}q = \varepsilon_0 V^N\int\delta\left(E-\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m}\right)\mathrm{d}\boldsymbol{p}_1\cdots\mathrm{d}\boldsymbol{p}_N \quad [42]$$

After some straightforward steps, one obtains

$$\Gamma(N,V,E) = \frac{\varepsilon_0 V^N(E)^{3N/2-1}(2m\pi)^{3N/2}}{\Gamma_{\mathrm{Eul}}(3N/2)} \quad [43]$$

and, substituting in [11] and using [10],

$$S(N,V,E) = k_{\mathrm{B}}\ \ln\frac{\varepsilon_0 V^N(E)^{3N/2-1}(2m\pi)^{3N/2}}{\Gamma_{\mathrm{Eul}}(3N/2)\gamma^{3N}N!} \quad [44]$$

which, for $N\to\infty$ and using the definitions [6], gives the asymptotic expression

$$S(N,V,E)\approx k_{\mathrm{B}}N\left(5/2+\ln\frac{(4m\varepsilon\pi/3)^{3/2}v}{\gamma^3}\right) \quad [45]$$

It is to be noted that the entropy [45] is extensive as expected, due to the Gibbs factor $N!$ introduced in [11]. Finally, applying [13]–[15], the usual thermodynamic equations for the perfect monatomic gas are

obtained:

$$\frac{1}{T} = \frac{3}{2}\frac{k_B}{\varepsilon} \quad [46]$$

$$\frac{P}{T} = \frac{k_B}{v} \quad [47]$$

$$\frac{\mu}{T} = -k_B \ln \frac{(4m\varepsilon\pi/3)^{3/2} v}{\gamma^3} \quad [48]$$

It can be noted that, due to the presence of the arbitrary normalization constant γ, the chemical potential is not univocally defined unless one introduces the third principle of thermodynamics or quantum mechanical arguments (see the comments to eqns [9] and [10]).

Canonical Ensemble

Consider the perfect monatomic gas composed of N particles in a fixed volume V, in thermal equilibrium with a reservoir at temperature T. This system is described by the canonical ensemble, and to calculate its thermodynamics the partition function Z defined in [18] is needed. Calculating Z for the Hamiltonian [41] one obtains

$$\begin{aligned} Z(N,V,T) &= \int e^{-H(p,q)/k_B T} dp\, dq \\ &= V^N \int e^{-\sum_{i=1}^{N} (p_i^2/2mk_B T)} dp_1 \cdots dp_N \\ &= V^N (2\pi m k_B T)^{3N/2} \end{aligned} \quad [49]$$

and, applying [20] for $m = 1$ and $N \to \infty$,

$$\begin{aligned} F(N,V,T) &= -k_B T \ln \frac{Z(N,V,T)}{N!\gamma^{3N}} \\ &\approx -k_B T N \ln(ev(2\pi m k_B T \gamma^{-2})^{3/2}) \end{aligned} \quad [50]$$

Again, the Helmholtz free energy is extensive due to the Gibbs factor $N!$

The main thermodynamic equations are then given by eqns [22]–[25]:

$$S = Nk_B \ln(e^{5/2} v(2\pi m k_B T \gamma^{-2})^{3/2}) \quad [51]$$

$$P = \frac{k_B T}{v} \quad [52]$$

$$\mu = -k_B T \ln(v(2\pi m k_B T \gamma^{-2})^{3/2}) \quad [53]$$

$$E = \tfrac{3}{2} N k_B T \quad [54]$$

Grand Canonical Ensemble

Finally, consider the case of a perfect monatomic gas in a fixed volume V, exchanging energy and particles with a reservoir at fixed T and μ. Using expression [30] for the grand canonical partition function, with $Z(N, V, T)$ given by [49], one obtains

$$\begin{aligned} Z^{GC}(\mu,V,T) &= \sum_{N=0}^{\infty} \frac{z^N Z(N,V,T)}{\gamma^{3N} N!} \\ &= \sum_{N=0}^{\infty} \frac{z^N V^N (2\pi m k_B T)^{3N/2}}{\gamma^{3N} N!} \\ &= e^{zV(2\pi m k_B T \gamma^{-2})^{3/2}} \end{aligned} \quad [55]$$

and, from [36] and [37],

$$\Omega = -PV = -k_B z V T^{5/2} (2\pi m k_B \gamma^{-2})^{3/2} \quad [56]$$

while the thermodynamic relations given by eqns [38]–[40] are

$$S = \left(\frac{5}{2} - \frac{\mu}{k_B T}\right) k_B z V (2\pi m k_B T \gamma^{-2})^{3/2} \quad [57]$$

$$N = zV(2\pi m k_B T \gamma^{-2})^{3/2} \quad [58]$$

$$E = \tfrac{3}{2} k_B T z V (2\pi m k_B T \gamma^{-2})^{3/2} \quad [59]$$

The thermodynamic equations derived for the three ensembles have obviously different meanings, since they refer to different physical situations. Indeed, quantities which are fixed in one of the ensembles can be fluctuating in a different ensemble (e.g., E is fixed in the microcanonical and fluctuating in the other ensembles). However, if the average values of the fluctuating quantities are considered, all these equations become equivalent from a mathematical point of view, allowing a unified thermodynamic description. From a physical point of view, this makes sense as far as the fluctuations around the average values are negligible with respect to the average values themselves or, in other words, as far as the relative fluctuations vanish.

This can be shown to happen usually in the thermodynamic limit, with the important exception of systems near critical point, where even the relative fluctuations typically diverge.

See also: Phases and Phase Equilibrium; Quantum Mechanics: Critical Phenomena; Statistical Mechanics: Quantum; Stochastic Processes in Physics and Chemistry; Specific Heat; Thermodynamic Properties, General.

PACS: 05.20. – y; 05.20.Gg

Further Reading

Gallavotti G (1999) *Statistical Mechanics*. New York: Springer.

Gibbs JW (1902) *Elementary Principles in Statistical Mechanics*. New York: Scribner's sons (free pdf version available at http://gallica.bnf.fr)

Khinchin AI (1960) *Mathematical Foundations of Statistical Mechanics*. New York: Dover.

Reichl L (1998) *A Modern Course in Statistical Physics*, 2nd edn. New York: Wiley-Interscience.
Ruelle D (1999) *Statistical Mechanics, Rigorous Results*. Singapore: World Scientific.
Shang-Keng Ma (1985) *Statistical Mechanics*. Singapore: World Scientific.
Thompson CJ (1988) *Classical Equilibrium Statistical Mechanics*. Oxford: Clarendon.
Toda M, Kubo R, and Saito N (1983) *Statistical Physics I: Equilibrium Statistical Mechanics*. Berlin: Springer.

Nomenclature

E	energy
F	Helmholtz free energy
H	Hamiltonian
k_B	Boltzmann constant
m	mass
N	number of particles
p	momenta
P	pressure
q	coordinates
S	entropy
t	time
T	absolute temperature
v	specific volume
V	volume
ε	specific energy
μ	chemical potential
Ξ	phase space

Statistical Mechanics: Quantum

R Fazio, NEST-INFM & Scuola Normale Superiore, Pisa, Italy
G Piccitto, Università di Catania, Catania, Italy

One of the fundamental results in classical statistical mechanics is the Boltzmann distribution function. Given a system formed by a large number of identical particles and in contact with a thermal reservoir, it gives the probability for the system to have a given energy E. The Boltzmann distribution can be derived by counting the number of possible microscopic configurations which lead to the same energy and assuming that they occur with the same probability. Another crucial assumption, which seems perfectly justified in the classical world, is that the particles, albeit identical, are supposed to be distinguishable. This last hypothesis does not hold in the quantum world, as one of the postulates of quantum mechanics states that identical particles should be considered as indistinguishable. The wave function of a system of identical particles should be unchanged, up to a phase factor, under a permutation of two of them. The standard argument to derive the statistics is to apply a permutation between two particles twice, that is, restoring the initial configuration. It immediately follows (the 2D case is not considered) that under a permutation, the wave function can remain unchanged (symmetric) or change sign (antisymmetric). These two cases define two different classes of particles named bosons and fermions, respectively. Bosons are particles with an integer spin, while for fermions, the spin assumes half-integer values. In both cases, the presence of a particle in a given state influences the probability of the other particles to be in that state; this is not the case in the hypothesis which leads to the Boltzmann distribution. (Indistinguishability of identical particles and then quantum mechanical effects are intimately related to the fact that the wave functions of two particles have appreciable overlap. Decreasing the temperature T or increasing the density, ρ, of an identical particle system will lead to a wave packet overlap when the thermal de Broglie wavelength $\lambda_{dB}(T) = \sqrt{2\pi\hbar^2/mk_BT}$ becomes of the order or greater than the average distance between particles, $\lambda_{dB}(T) \geqslant d \approx 1/\rho^{1/3}$.) If P_1 is the probability of adding a classical particle to a state with n particles, from the symmetry properties of the wave functions, it is easy to show that in the case of fermions or bosons, this probability is respectively suppressed by a factor $(1 - n)$, or enhanced by a factor $(1 + n)$. The probability for two bosons to be in the same quantum state is enhanced; on the contrary, fermions obey the Pauli principle which forbids two fermionic particles to have the same quantum numbers.

In order to obtain the distribution function for fermions and bosons one can proceed as follows. Consider two energy states E_1 and E_2; the principle of detailed balance states that in equilibrium the ratio of transition rates between the two states is equal to the ratio of the populations of particles in the two states or, equivalently, to the ratio of the Boltzmann factors: $\Gamma_{2\to1}/\Gamma_{1\to2} = n_1/n_2 = e^{-\beta E_1}/e^{-\beta E_2}$. The probability of having n fermions/bosons in a given state can be expressed in terms of the Boltzmann distribution if the rates are properly modified so as to account for the suppression/enhancement factor due to the occupancy of the final state. The detailed balance condition can then be rewritten as

$n_1(1 \pm n_1)/[n_2(1 \pm n_2)] = e^{-\beta(E_1 - E_2)}$ where $\beta = 1/k_B T$, T is the temperature and k_B is the Boltzmann constant. From the previous equation it follows that the distribution functions for quantum particles are

$$n_{FD}(E) = \frac{1}{e^{\beta(E-\mu)} + 1} \quad \text{for fermions} \qquad [1]$$

$$n_{BE}(E) = \frac{1}{e^{\beta(E-\mu)} - 1} \quad \text{for bosons} \qquad [2]$$

In the previous equations the chemical potential μ fixes the average total number of particles. The functions $n(E)$ are named after the scientists who formulated them, as Bose–Einstein and Fermi–Dirac distribution functions. Equations [1] and [2] are represented in **Figure 1** for different values of the temperatures. At energies E much larger than $1/\beta$, the quantum statistics approaches the Boltzmann distribution; in the limit of high temperatures, the average occupation of each energy state is much smaller than one and the quantum effects can be ignored. Indeed for $n \ll 1$, the suppression/enhancement factor related to the quantum statistics can be ignored and thus the classical distribution function is recovered. In the low temperature limit, the quantum effects become dominant. Bose–Einstein statistics lies always above the Boltzmann distribution. The Fermi–Dirac distribution function is always less than one (due to the Pauli principle), and at zero temperature, drops abruptly from one to zero at the (so-called) Fermi energy E_F (coincident in this case with the chemical potential).

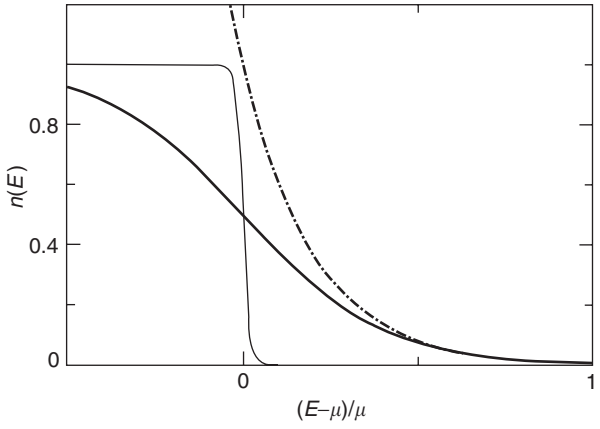


Figure 1 The Fermi–Dirac distribution function (thin solid curve) is compared to the Boltzmann distribution (dashed-dotted curve) for a temperature of $T = 0.2\,\mu/k_B$. The agreement with the classical distribution is good only for energies far above the Fermi energy. At low energies the Pauli principle prevents the double occupancy of the same state as it would follow from the classical analysis. For completeness, the Fermi–Dirac distribution is plotted also at lower temperature $T = 0.01\,\mu/k_B$; it can be noted that it approaches a step function at zero temperature.

The differences in the distribution functions will reflect in the properties of Fermi and Bose systems at low temperatures. Although it is impossible to briefly review this huge field of investigation, it is nevertheless useful to recall a few brief important examples where quantum statistics plays a dominant role. In order to keep the presentation as simple as possible, perfect gases are considered, that is, the only contribution to the energy is due to the kinetic term while interaction between particles are disregarded. The fact that a direct interaction term is not present does not mean that the particles do not feel each other. As discussed already, indistinguishability implies that the probability for a particle to occupy a given state influences the probability of the other particles being in the same state. Let fermions be considered first. At zero temperature, from the Pauli principle, it follows that all states below a certain energy are occupied while all the states above it are empty. The energy of the highest occupied level is defined as the Fermi energy. A Fermi gas is said to be degenerate when its temperature is small compared to E_F/k_B; in the opposite limit of a non-degenerate gas, the properties are those of a classical gas and therefore, they are not discussed here. For a gas of N fermions contained in a volume V, the Fermi energy is given by

$$E_F = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3} \qquad [3]$$

where $\hbar$ is the Planck constant divided by 2π, and m is the fermion mass. Another useful quantity to characterize the system is the density of states $N(\varepsilon)$, defined as the number of states which are present in an energy shell between E and $E + dE$. For a Fermi gas, counting the states in the given shell results in a density of states

$$N(E) = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2} \qquad [4]$$

The chemical potential μ coincides with the Fermi energy at zero temperatures; on increasing T, it decreases and becomes negative (a 3D system is considered) at temperatures of the order of the Fermi temperature $T_F = E_F/k_B$. Typical values of the Fermi energy for metals are such that at room temperature the ratio T/T_F is of the order of 10^{-2}. In this case, which is the most frequently encountered situation, the density of states can be assumed to be constant in an interval of the order of T centered in T_F. A good example to see deviations in the thermodynamic quantities due to quantum effects is to consider the heat capacity of a Fermi gas, such as electrons in a metal. One can recall the equipartition theorem which states that, for a monoatomic gas the heat

capacity is constant and is 3/2. That is what is expected for a Fermi gas at temperatures higher than the Fermi temperature. In the low temperature limit, the only states which contribute to the heat capacity are those in a strip of width $k_B T$ around the Fermi energy, that is, those in the energy region where the Fermi function is smeared. All the other electrons are effectively frozen due to the Pauli principle. The number of electrons thus contributing to the temperature dependence of the internal energy are of the order of NT/T_F, each of them having an energy of the order of $k_B T$. A detailed calculation leads to the result $C_e = (\pi^2/2)Nk_B T/T_F$; it goes to zero linearly with temperature. It is possible to measure the electronic contribution to the heat capacity in metals at low temperature (typically of the order of kelvin) when the lattice contribution (which depends on T^3) can be disregarded. Another interesting example of the consequence of the Pauli principle in the thermodynamic quantities is the Pauli spin susceptibility, which measures the spin response of the electron system to an applied magnetic field. Using the same line of reasoning as before, it is evident that at low-temperature spin susceptibility, $\chi = N(E_F)$. Indeed the argument goes as follows: the number of spins which can be flipped by an applied magnetic field is a fraction of T/T_F around the Fermi energy because, due to the Pauli exclusion principle, the other spins cannot be flipped since the corresponding (spin-flipped) state is occupied. The susceptibility per spin is given by the Curie law $\sim 1/T$. From here it therefore follows that the Pauli result of a spin susceptibility is independent of temperature and proportional to the density of states at the Fermi energy. In the presence of electron–electron interaction and/or external potential, instead of the particles, one can describe the properties of the Fermi liquid by means of the long-lived quasiparticle excitations close to the Fermi energy. The leading temperature dependence of the various observables is not modified (for example the heat capacity remains linear in the temperature dependence); however, the effect of the interaction enters in the renormalization of the parameters characterizing the Fermi liquid. Degenerate Fermi systems are observed in a variety of very different physical situations ranging from electrons in metals to electrons in white dwarfs. It is worth stressing at this point that the effect of interaction between electrons, or interaction of the electrons with the ions of the lattice does not always result in the renormalization of the parameters (such as effective mass).

A completely different behavior is observed in a gas of bosons at low temperatures. Here there is no Pauli blocking; on the other hand, for bosons, the probability of being in the same quantum state is enhanced as compared to the classical value. A very interesting phenomenon observed in the Bose system is that of the Bose–Einstein condensation. Below a given critical temperature, there is a macroscopic fraction of the particles which condense into the state with the lowest energy. In order to grasp the process of Bose–Einstein condensation, it is useful to understand the behavior of the chemical potential close to the transition temperature. (If the total number of bosons is not conserved, as for photons, the chemical potential is zero and the Bose–Einstein distribution coincides with the Planck distribution for the blackbody radiation.) As in the case of fermions, an ideal boson gas with a macroscopic fixed number N of particles of spin zero is considered. The density of states $N(E)$ of this system is given by eqn [4]. The factor $1/2$ is due to the fact that the density of states per spin species is being considered. In order to avoid that some occupation number could be negative, the value of chemical potential must always be less than the ground-state energy which, for simple reasons is assumed to be equal to zero, $\mu < 0$ (see eqn [2]). The chemical potential depends on the number of particles and on the temperature. At zero temperature all the N particles are in the ground state; from this fact it is easy to show that for $T \to 0$, the chemical potential tends to the ground-state energy as $\mu = -k_B T/N$. The particles are distributed among the ground state and all the excited states as

$$\begin{aligned} N &= \sum_i \frac{1}{e^{(E_{i_i}-\mu)/k_B T} - 1} = N_0(T) \\ &\quad + \int_0^\infty N(E) \frac{1}{e^{(E-\mu)/k_B T} - 1} \, dE \\ &= N_0(T) + N_{ex}(T) \end{aligned} \qquad [5]$$

The integral in eqn [5] can be calculated and the result is

$$N = N_0(T) + \frac{V}{\lambda_{dB}^3(T)} g_{3/2}(e^{\mu/k_B T}) \qquad [6]$$

where $g_\eta(z) = \sum_{t=1}^\infty z^t/t^\eta$.

At "high" temperature the value $N_{ex}(T)$ is of the same order of magnitude of N, and the population is distributed over all the states, each state being weakly occupied. Below a critical temperature $T_c \propto \rho^{2/3}$, where ρ is the particle density, the ground state becomes macroscopically occupied and the fraction of particle in the ground state is

$$\frac{N_0(T)}{N} = 1 - \left(\frac{T}{T_c}\right)^{3/2} \qquad [7]$$

The temperature T_c is the Einstein condensation temperature and satisfies the relation $N_{ex}(T_c) = N$; it

is important to note that the critical temperature is usually several orders of magnitude higher than the level spacing between the ground state and the first excited state. The occurrence of a single quantum state macroscopically occupied is the key to the Bose–Einstein condensation. The experimental observation of Bose–Einstein has been recently realized with dilute alkali gases.

Quantum statistical mechanics embraces a large class of situations where quantum effects do not necessarily stem from the quantum statistics of the constituent particles. There are a large number of situations where, though the particles are localized (therefore, the indistinguishability is not important), quantum mechanics plays a crucial role. Probably, among the most popular situations, the studies on localized spins interacting through an exchange coupling are mentioned. In this case, quantum effects emerge due to the quantum nature of the spin degree of freedom. The fact that different components of the spin do not commute reflects in a kind of quantum frustration. The impact of spin models in quantum statistical mechanics is not confined only to the theory of magnetism. More generally, pseudo-spin models should be considered as they cover a wider range of materials, for example, glasses and superconductors. They can be described in terms of fictitious spins located in a given network, in certain limits.

As it is already hinted in the previous paragraph, interaction between the particles is responsible for a huge variety of phenomena which will appear in many articles in this encyclopedia. In the brief description made of the properties of Fermi systems, it is mentioned that electron–electron interaction can be incorporated by means of renormalization of the physical parameters characterizing the fermion liquid. However, this is not always the case. Among the others, interaction between particles often leads to instabilities and consequently a phase transition to a condensed phase occurs. Examples of phase transitions are ferromagnetism, superconductivity, and charge/spin density waves, just to mention a few important cases. Phase transitions can be driven both by thermal fluctuations and by quantum fluctuations. In the former case, below a certain temperature the system orders spontaneously even in the absence of any applied field. The prototype example is the ferromagnetic transition; below the Curie temperature, the system will present a spontaneous magnetization even in the absence of any external magnetic field. A quantum phase transition instead takes place at zero temperature, and it is driven by the changing of some parameter, for example, a coupling constant or an external field, of the system. In this case the thermodynamical properties, for example, the magnetization, are deeply connected to the dynamical ones since both derive from the quantum dynamics of the many-body system. Probably, the simplest example of a quantum phase transition can be described in a model of spin interacting via an (ferromagnetic) exchange coupling (with strong anisotropy in the x-direction) J_x, and subjected to an external magnetic field B (say along the z-direction). This model Hamiltonian, known as Ising model in a transverse field, appropriate for this model is

$$H = -\frac{J_x}{2}\sum_{\langle ij\rangle}\sigma_i^x\sigma_j^x - B\sum_i \sigma_i^z \qquad [8]$$

where σ^a are the Pauli matrices ($a = x, y, z$). The interaction, in order to give an example, is chosen to be between nearest neighbors. In the limit of a large external field the spin will align along the z-direction. However, because of the quantum nature of the spin degree of freedom, the alignment along the direction of the external field prevents any kind of ordering along the x-direction due to the exchange interaction. By increasing the ratio between J_x/B, at a given critical value the spin will prefer to order along the x-direction, and the system will develop a nonzero magnetization along the x-direction. The transition to the ordered state can be induced at zero temperature solely by the variation of the parameter J_x/B and, as has already been said, is intimately related to the quantum properties of the spin. All the signals for the emergence of a phase transition are contained in the dynamical properties of the interacting spins. If the spins of the Ising model are arranged in a chain, this is one of the few exactly solvable models in quantum statistical mechanics. Here the case of a 3D lattice is considered and the quantum phase transition by means of the mean field approximation is discussed. It consists in replacing the interaction among the neighboring spins by an effective field which acts on each spin, that is, it consists in replacing $\sigma_i^x = \langle\sigma_i^x\rangle + (\sigma_i^x - \langle\sigma_i^x\rangle)$, where the angular brackets $\langle .. \rangle$ mean the quantum statistical average, disregarding the quadratic fluctuations around the average internal field. The value of $\langle\sigma^x\rangle$ is determined self-consistently by imposing that the average value of the magnetization calculated with the approximated form of the Hamiltonian, which reads $H_{\mathrm{MF}} = -(J_x z\langle\sigma^x\rangle/2)\ \sum_i\ \sigma_i^x - B\sum_i \sigma_i^z$, coincides with the mean field. In the previous equation, z is the number of nearest neighbors. This condition leads to the self-consistent equation

$$\langle\sigma^x\rangle = \langle\psi_{\mathrm{MF}}|\sigma^x|\psi_{\mathrm{MF}}\rangle \qquad [9]$$

where the $|\psi_{\mathrm{MF}}\rangle$ is the ground-state wave function of the mean field Hamiltonian (only the case of zero temperature is considered). A numerical solution of eqn [9] sketched in **Figure 2**, shows the behavior that is

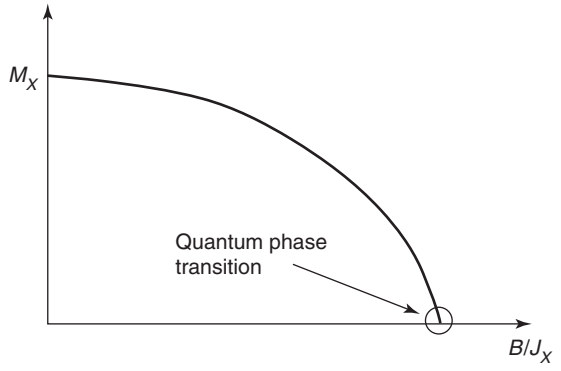


Figure 2 A sketch of the magnetization, at zero temperature, along the *x*-axis is plotted against the ratio B/J_x. The spontaneous magnetization is different from zero only at sufficiently low fields. Close to the critical point the magnetization goes as $|\lambda - \lambda_\chi|^\beta$. In the mean field analysis discussed in the text $\beta = 1/2$.

briefly outlined. As it is beautifully described by Nielsen and Chuang in 2000, thanks to universality, most of the properties of a system close to a phase transition do not depend on the microscopic details of the model. For the example discussed, while the value of the critical field is indeed dependent on the exchange coupling, the behavior of the magnetization is only related to the dimensionality of the system and to the symmetry which is spontaneously broken.

As it should emerge from these remarks, quantum mechanics plays a different role in classical and quantum phase transitions. Let superconductivity as an example be taken in order to elucidate this point. A transition to the superconducting state is induced by the condensation of a pair of electrons (the Cooper pairs). The effective electron–electron attractive interaction (which in conventional superconductors is mediated by the phonons) is responsible for the formation of the pairs. The formation of a bound state of two electrons is a Fermi surface effect, which relies on quantum mechanics. Although quantum mechanical effects are responsible for the instability to the superconducting state, the properties of the transition itself can be described by classical statistical mechanics. In a quantum phase transition, instead, the quantum fluctuations are primarily responsible for the transition itself, which is now related to a modification of the ground-state wave function.

This last paragraph represents a return to the theme in the introduction and a slight deviation from the traditional quantum statistical mechanics. The question of indistinguishability viewed from a different perspective, has stimulated a number of works in the newly born field of quantum information. The (anti-)symmetrized form of the wave function for a system composed by N (fermions) bosons implies that the many-body state cannot be written, even in the absence of any interaction, as a product of wave functions related to each single particle. The many-body wave function is said to be entangled. Entanglement is the property of quantum states to exhibit correlation which cannot be accounted for by classical physics. It arises as a consequence of the principle of superposition in the presence of a Hilbert space with a tensor product structure. Entanglement denotes the nonlocal correlations that exist, even in the absence of direct interaction, between two (spatially separated) parts of a given quantum system. Since the early days of quantum mechanics, understanding the phenomenon of entanglement has been central to the understanding of the foundations of quantum theory. Besides its fundamental importance, a great deal of interest has been brought forth by its role in quantum information. Entanglement is believed to be the main ingredient of computational speed-up in quantum information protocols. In the case being discussed here, entanglement is of statistical origin. With the development of information theory, a great deal of attention has been devoted to the characterization and quantification of entanglement contained in a given quantum state. Although the progress has been impressive, a full characterization is still far from being completed. Parallel to understanding entanglement in systems of two-level systems (qubits) used for quantum information, research has touched the question on how to characterize entanglement in "traditional" condensed matter systems with the aim of finding a bridge between information science and condensed matter. In order to understand more deeply the role of quantum mechanics in the collective properties of quantum many-body systems, probably conventional approaches in quantum statistical mechanics can be conveniently supplemented by the use of a variety of techniques developed in quantum information theory. As a final example, the attention devoted recently to the study of quantum phase transitions from a "quantum information" point of view is mentioned. Both classical and quantum phase transitions can indeed be formulated in the framework of Ginzburg–Landau approach. A crucial point that determines all the properties close to the transition is the fact that the systems behave cooperatively. Classical correlations between particles do not, however, imply that quantum correlations (entanglement) do follow the same behavior. As entanglement is one of the major resources in quantum computation, it is certainly important to be able to understand the properties and the amount of entanglement possessed by a large number of particles close to a quantum phase transition. The link between statistical mechanics and quantum information theory has been carried ahead in refining established methods

for examining many-body interacting systems. An QJ;example is the entanglement preserving Density Matrix Renormalization Group, where a deep understanding of the entanglement properties of the many-body systems has helped in reformulating and generalizing this method. A final mention is made of the attention which is being devoted to the possibility of using quantum spin networks, a traditional arena for quantum statistical mechanics, as quantum channels for transporting and/or manipulating quantum information.

See also: Bose–Einstein Condensation; Statistical Mechanics: Classical; Thermodynamic Properties, General.

PACS: 05.30. – d; 03.67

Further Reading

Abrikosov AA, Gorgov LP, and Dzyaloshinski IE (1975) *Methods of Quantum Field Theory in Statistical Physics*. New York: Dover.

Daley AJ, Kollath C, Schollwoeck U, and Vidal G condmat/0403313 at http://xxx.lanl.gov. and references therein.

Dalfovo F, Giorgini S, Pitaevskii LP, and Stringari S (1999) *Review of Modern Physics* 71: 463.

Eisberg R and Resnick R (1985) *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. New York: Wiley.

Goldenfeld N (1992) *Lectures on Phase Transitions and the Renormalization Group*. New York: Addison Wesley.

Huang K (1987) *Statistical Mechanics*. New York: Wiley.

Nielsen M and Chuang I (2000) *Quantum Computation and Quantum Communication*. Cambridge: Cambridge University Press.

Osborne TJ and Nielsen MA (2002) *Physical Review A* 66: 044301.

Osterloh A, Amico L, Falci G, and Fazio R (2002) *Nature* 416: 608.

Sachdev S (2000) *Quantum Phase Transitions*. Cambridge: Cambridge University Press.

Vedral V quant-ph/0302040 at http://xxx.lanl.gov. and references therein.

Vidal G, Latorre JI, Rico E, and Kitaev A (2003) *Physical Review Letter* 90: 227902.

Yosida K (1998) *Theory of Magnetism*. Heidelberg: Springer.

Stochastic Processes in Physics and Chemistry

C Cercignani, Politecnico di Milano, Milan, Italy

Introduction

The adjective "stochastic" goes back to the Greek word $\sigma\tau o\chi o\varsigma$, meaning "target" and is more or less equivalent to "random." Thus, a stochastic process is governed by a probabilistic rather than a deterministic law.

The typical example of a stochastic process is the Brownian motion, called after the English botanist Robert Brown, who noticed (in 1827) that pollen grains suspended in water jiggle about, following a zigzag path. Several scientists of the nineteenth century remarked that an explanation of this movement is provided by the molecular theory of matter. A suspended particle is constantly and randomly bombarded from all sides by the molecules of the liquid. If the particle is very small, the hits it takes do not balance (as is the case for a sizable object) and cause it to jump. These small random jumps are what make up the Brownian motion. The first detailed theory of this phenomenon was developed by Einstein in 1905.

Stochastic processes are involved in physics because there are many phenomena which depend on time in an extremely complicated way, far beyond any possibility of calculation, but which have some average features that obey simple laws. For example, the instantaneous value of the force exerted by the N molecules of a gas on the walls of a container varies rapidly in an unpredictable way, but when averaging over a small time interval, it becomes a simple function of the density and the average kinetic energy of a molecule. Thus, probability considerations in physics are due to the ignorance of the precise values of microscopic variables and the remarkable fact that it is still possible to detect regularities in the behavior of a few quantities that describe the macroscopic behavior in a satisfactory way.

The microscopic state (or microstate, in short) of a gas is determined by $6N$ variables, the coordinates, and the components of momentum of the N molecules. If one assigns the values of these variables at time $t = 0$, then their values at any other time instant are determined by the equations of motion, provided the forces are known as functions of position and velocity. In other words, the initial microstate x determines the microstate $y(x, t)$ at any time t.

If one considers the set of initial microstates with a probability distribution, one has a stochastic variable X, which determines the subsequent microstates, described by another stochastic variable Y; in other words, $Y = Y(X, t)$ describes a stochastic process. To find $Y(X, t)$ is as hard as finding the original motion.

From a purely Newtonian standpoint, one should not talk of a stochastic process, but rather of a stochastic model of a process. As a matter of fact, the

complex trajectory of a tossed coin, or a die, are all phenomena ruled by deterministic laws, and thus, not random at all. The stochastic description is useful due to the fact that one can obtain simplified evolution equations by averaging over the irrelevant variables. In the case of a dilute gas, this leads to the famous Boltzmann equation, which may even become exact, in a probabilistic sense, when the number of particles N tends to infinity.

Quantum mechanics brings in an additional unpredictability due to the basic theory and not the complicated nature of the system. The classical averaging over microscopic states in a region of phase space is equivalent, in quantum mechanics, to the averaging over all unit vectors in a linear subspace of the Hilbert space of the system. This is the conventional interpretation of quantum mechanics; a description of quantum processes as genuine stochastic processes has been provided by E Nelson, which is discussed at the end of this article.

Stochastic Variables

First the concept of the stochastic (or random) variable: it is a variable X which can have a value in a certain set Ω, usually called "range," "set of states," "sample space," or "phase space," with a certain probability distribution. When a particular fixed value of the same variable is considered, the small letter x is used.

The set may be discrete; the simplest example is heads or tails, where the probability distribution is 1/2 for each outcome (if one assumes an unloaded coin). More commonly, the set is continuous: for example, the kinetic energy of a particle undergoing a Brownian motion may take any real non-negative value x. The set of states may be multidimensional, in which case, X is conveniently written as a vector X. For example, X might be the velocity of a particle undergoing a Brownian motion; in this case, the set of states is a three-dimensional space and X may take any vector value x.

When the variable is continuous, the probability distribution $P(X) \geqslant 0$ is a non-negative function normalized to unity in the sense that

$$\int_{\Omega} P(x)\, \mathrm{d}x = 1 \qquad [1]$$

where the integral extends to the whole range. When not stated otherwise, the stochastic variables are assumed to be continuous. More generally, the probability can be a measure μ of the subsets of the set Ω ($\mu(\Omega) = 1$); the measure can be defined for a family F of "elementary sets" and then extended to other subsets of Ω, via the basic properties of μ (positivity, additivity for disjoint subsets).

The average or expectation value of any function $f(X)$ of the stochastic variable X is given by

$$\langle f(X) \rangle = \int_{\Omega} f(x) P(x)\, \mathrm{d}x \qquad [2]$$

In particular, $\mu_n = \langle X^n \rangle$ is called the nth moment of X, and μ_1 the average or mean of X. Also,

$$\begin{aligned} \sigma^2 &= \langle (X - \langle X \rangle)^2 \rangle = \mu_2 - \mu_1^2 \\ &= \langle X^2 \rangle - \langle X \rangle^2 \end{aligned} \qquad [3]$$

is called the variance or dispersion and is the square of the standard deviation σ.

If X is a stochastic variable having r components $X_1, X_2, ..., X_r$, its probability density $P(x) = P_r(x_1, x_2, ..., x_r)$ is also called the joint probability distribution of the r variables $X_1, X_2, ..., X_r$. If one takes a subset of $s < r$ variables $X_1, X_2, ..., X_s$, their probability distribution, regardless of the values of the remaining variables $X_{s+1}, ..., X_r$, is

$$\begin{aligned} P_s(x_1, x_2, ..., x_s) = \int_{\Omega_{r-s}} & P_r(x_1, x_2, \ldots, x_s, x_{s+1}, \ldots, x_r) \\ & \times \mathrm{d}x_{s+1}, \ldots, \mathrm{d}x_r \end{aligned} \qquad [4]$$

and is called the marginal distribution of the given subset of variables.

One can also attribute fixed values to $X_{s+1}, ..., X_r$ and consider the joint probability distribution of the s variables $X_1, X_2, ..., X_s$. This is called the conditional probability of distribution of $X_1, X_2, ..., X_s$ when $X_{s+1}, ..., X_r$ have the prescribed values $x_{s+1}, ..., x_r$. It is usually denoted by $P_{s|r-s}(x_1, x_2, ..., x_s | x_{s+1}, ..., x_r)$. This probability is related to the previous one by the Bayes rule

$$\begin{aligned} & P_{s|r-s}(x_1, x_2, ..., x_s | x_{s+1}, ..., x_r) \\ & \quad = \frac{P_r(x_1, x_2, ..., x_r)}{P_{r-s}(x_{s+1}, x_2, ..., x_r)} \end{aligned} \qquad [5]$$

Stochastic Processes

Once a stochastic variable has been defined, an infinite number of other variables can be derived from it, that is, all the quantities Y which are functions of X, $Y = f(X)$. If a family of functions indexed by a variable t is taken $Y_t = f(X, t)$, a stochastic process is obtained. Mathematically, the functions are assumed to be measurable, that is, belong to a suitable measure space.

In most applications t is the time variable, as suggested by the notation and the name process, but this is not a part of the definition. t might belong to a set T. For example in the so-called Bernoulli scheme

(or the heads and tails process), Ω is the set $[+1, -1]$, where $+1$ represents heads and -1 tails. In this case, T is the set (N) of natural numbers; and the process is the set of sequences $y = (a^1, a^2, \ldots)$ with $a^i = +1$ or -1, each with probability 1/2. If T is manifold (such as space or space–time), a stochastic process is usually called a stochastic field.

For a stochastic process, one can define joint probability densities referring to different values of t (or different instants of time), $P_n(y_1, t_1; y_2, t_2, \ldots, y_n, t_n)$. This hierarchy of functions obeys four consistency conditions:

1. $P_n \geqslant 0$,
2. P_n does not change when interchanging two pairs $(y_i, t_i$ and $y_k, t_k)$,
3. $\int P_n(y_1, t_1; \ldots, y_{n-1}, t_{n-1}, y_n, t_n)\, dy_n = P_{n-1}(y_1, t_1; \ldots, y_{n-1}, t_{n-1})$, and
4. $\int P_1(y_1, t_1)\, dy_1 = 1$.

Kolmogorov has shown that any set of functions satisfying these four conditions determines a stochastic process. One can similarly discuss the conditional probabilities.

A stochastic process is called stationary if all the joint probability densities depend on the time differences alone. A necessary, but not sufficient, condition is that $P_1(y_1)$ is independent of time.

Markov Processes

A Markov process is a stochastic process with the property that the state at a certain time t_0 determines the states for $t > t_0$ and not the states $t < t_0$. In other words,

$$P_{1|n-1}(y_n, t_n | y_1, t_1, \ldots; y_{n-1} t_{n-1}) = P_{1|1}(y_n, t_n | y_{n-1}, t_{n-1}) \quad \text{for } t_1 < t_2 \cdots < t_n \quad [6]$$

A Markov process is fully determined by the two functions $P_1(y_1, t_1)$ and $P_{1|1}(y_2, t_2 | y_1, t_1)$. Thus, for example,

$$P_3(y_1, t_1; y_2, t_2; y_3, t_3) = P_1(y_1, t_1) P_{1|1}(y_2, t_2 | y_1, t_1) \times P_{1|1}(y_3, t_3 | y_2, t_2) \quad \text{for } t_1 < t_2 < t \quad [7]$$

Integrating this identity with respect to y_2, one obtains

$$P_2(y_1, t_1; y_3, t_3) = P_1(y_1, t_1) \int_\Omega P_{1|1}(y_2, t_2 | y_1, t_1) \times P_{1|1}(y_3, t_3 | y_2, t_2)\, dy_2 \quad (t_1 < t_2 < t_3) \quad [8]$$

or

$$P_{1|1}(y_3, t_3 | y_1, t_1) = \int_\Omega P_{1|1}(y_2, t_2 | y_1, t_1) P_{1|1} \times (y_3, t_3 | y_2, t_2)\, dy_2 \quad (t_1 < t_2 < t_3) \quad [9]$$

This is called the Chapman–Kolmogorov equation. This equation and the obvious equation

$$P_1(y_2, t_2) = \int_\Omega P_{1|1}(y_2, t_2 | y_1, t_1) P_1(y_1, t_1)\, dy_1 \quad [10]$$

are both necessary and sufficient for two non-negative functions $P_1(y_1, t_1)$ and $P_{1|1}(y_2, t_2 | y_1, t_1)$ to uniquely define a Markov process.

There are two examples of the Markov process which are worth discussing in some detail. The first is the so-called Wiener process defined by

$$P_{1|1}(y_2, t_2 | y_1, t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp\left[-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}\right] \quad (t_1 < t_2) \quad [11]$$

which can be checked to satisfy the Chapman–Kolmogorov equation with $P_1(y_1, 0) = \delta(y_1)$. It describes the stochastic behavior of the position of a Brownian particle (for simplicity, the diffusion coefficient v is taken to be unity; to reinstate it, one must replace time t by vt and, as a consequence, multiply $P_{1|1}$ by v). The probability density for $t > 0$, according to eqn [10], is

$$P_1(y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{y^2}{2t}\right] \quad [12]$$

It is easy to check that

$$\lim_{\tau \to 0+} \frac{\langle [Y(t+\tau) - Y(t)]^2 \rangle}{\tau} = \lim_{\tau \to 0+} \frac{1}{\sqrt{2\pi\tau^3}} \int_{-\infty}^{\infty} y^2 \exp\left[-\frac{y^2}{2\tau}\right] dy = 1 \quad [13]$$

In the second example, $Y(t)$ takes on non-negative integer values n and $t \geqslant 0$. A Markov process is defined by

$$P_{1|1}(n_2, t_2 | n_1, t_1) = \frac{(t_2 - t_1)^{n_2 - n_1}}{(n_2 - n_1)!} e^{-(t_2 - t_1)} \quad (t_1 \leqslant t_2)$$
$$P_1(n, 0) = \delta_{n,0} \quad [14]$$

This is called a Poisson process.

When a Markov process is stationary, then $P_{1|1}$ depends on just the time difference $\tau = t_2 - t_1$ and the Chapman–Kolmogorov equation simply states that the integral operator T_τ with kernel $P_{1|1}(y_2, y_1; \tau)$ has the semigroup property $T_{\tau+\tau'} = T_\tau T_{\tau'}$.

The best known example of a stationary Markov process, after the Wiener process, is the Ornstein–Uhlenbeck process, for which

$$P_{1|1}(y_2|y_1;\tau) = \frac{1}{\sqrt{2\pi(1-e^{-2\tau})}} \exp\left[-\frac{(y_2 - y_1 e^{-\tau})^2}{2(1-e^{-2\tau})}\right]$$
$$P_1(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad [15]$$

In the case of a stationary Markov process, eqn [7] becomes

$$P_3(y_1,t_1;y_2,t_2;y_3,t_3) = P_1(y_1,t_1) \times P_{1|1}(y_2|y_1,t_2-t_1)P_{1|1}(y_3|y_2,t_3-t_2) \quad [16]$$

and, of course, it can be trivially extended from P_3 to P_n. Thus, the probability that a particle starting from $x=0$ will be found between α_1 and β_1 at time t_1, between α_2 and β_2 at time t_2, etc., and between α_n and β_n at time t_n, is given by the formula

$$\int_{\alpha_1}^{\beta_1} \dots \int_{\alpha_n}^{\beta_n} P_1(y_1,t_1)P_{1|1}(y_2|y_1;t_2-t_1)\dots P_{1|1} \times (y_n|y_{n-1};t_n-t_{n-1}) \quad [17]$$

If the Wiener process for which $P_1(y_1,0)=\delta(y_1)$ is considered, eqn [10] with $t_1=0$ becomes

$$P_1(y_2,t_2) = P_{1|1}(y_2|0;t_2) \quad [18]$$

and eqn [17] becomes

$$\int_{\alpha_1}^{\beta_1} \dots \int_{\alpha_n}^{\beta_n} P_{1|1}(y_1|0;t_1)P_{1|1}(y_2|y_1;t_2-t_1) \dots P_{1|1}(y_n|y_{n-1};t_n-t_{n-1}) \quad [19]$$

This result was first derived by Einstein and Smoluchowski for a free Brownian particle.

Time Evolution Equations for the Probability Density

A stationary Markov process can be characterized by the transition probability per unit time $W(y_3|y_1)$ of going from state y_1 to y_3. Then, the Chapman–Kolmogorov equation may be replaced by the so-called master equation

$$\frac{\partial}{\partial\tau} P_{1|1}(y_3|y_1;\tau) = \int_\Omega \{W(y_3|y_2)P_{1|1}(y_2|y_1;\tau) - W(y_2|y_3)P_{1|1}(y_3|y_1;\tau)\}dy_2 \quad [20]$$

which is sometimes written as

$$\frac{\partial}{\partial t} P(x,t) = \int_\Omega \{W(x|x')P(x',t) - W(x'|x)P(x,t)\}dx' \quad [21]$$

when it is not necessary to mention the argument y_1 of the previous equation.

If the range of Y is a discrete set of states with labels n, the master equation for the probability p_n of a state reads as follows:

$$\frac{\partial}{\partial t} p_n(t) = \sum_{n'} \{W_{nn'}p_{n'}(t) - W_{n'n}p_n(t)\} \quad [22]$$

This equation can be applied to a chemical reaction which changes the number of molecules n of a given species.

One can easily show that if the coefficients $W_{nn'}$ are non-negative, then $p_n(t)$is non-negative for $t>0$, if it is for $t=0$. In most situations of interest, there is an equilibrium solution p_n^e such that

$$\sum_{n'} \{W_{nn'}p_{n'}^e - W_{n'n}p_n^e\} = 0 \quad [23]$$

If this is the case, introduce a function $H(t)$ by

$$H(t) = \sum_n p_n^e f\left(\frac{p_n}{p_n^e}\right) = \sum_n p_n^e f(x_n) \quad [24]$$

where $x_n = p_n/p_n^e$ for convenience, and $f(x)$ is a non-negative differentiable convex function:

$$f \geqslant 0, \quad (x-y)f'(x) \geqslant f(y) - f(x) \quad (0 \leqslant x < \infty) \quad [25]$$

where the equality sign applies if and only if $y=x$. If $f(x)$ is twice differentiable, this implies (and is implied by) $f''(x) \geqslant 0$. Then

$$\frac{dH(t)}{dt} = \sum_{n,n'} W_{nn'}p_{n'}^e\{x_{n'}f'(x_n) - x_{n'}f'(x_{n'})\} \quad [26]$$

It is also remarked that for any set of numbers c_n, eqn [23] implies

$$\sum_{n'} W_{nn'}p_{n'}^e(c_n - c_{n'}) = 0 \quad [27]$$

Thus, if one chooses $c_n = f(x_n) - x_n f'(x_n)$ and adds the result to eqn [26], one obtains

$$\frac{dH(t)}{dt} = \sum_{n,n'} W_{nn'}p_{n'}^e\{(x_{n'} - x_n) + f(x_n) - f(x_{n'})\} \leqslant 0 \quad [28]$$

where [25] has been used. Thus, H decreases in time unless $x_n = 1$ or $p_n = p_n^e$. Since H cannot become negative, it must tend to zero when $t \to \infty$; if the number of states is finite, this immediately implies that p_n also tends to a limit and this limit is p_n^e. To show that this is true under more general conditions requires a more detailed argument.

Any convex function such as, $f(x) = x^{\alpha}(\alpha > 1)$ could do, but one usually chooses

$$f(x) = x \log x, \quad H = \sum_n p_n \log \frac{p_n}{p_n^e} \qquad [29]$$

for at least two reasons. One is that this choice also works for the nonlinear Boltzmann equation and the other is that this choice of H produces a function which has all the properties of entropy (except for the sign).

Similar arguments work in the continuous case as well, which is taken up again. Sometimes, it is convenient to write $W(x|x') = W(x'; r)$, where $r = x - x'$. If one assumes that $W(x; r)$ is a peaked function of r about $r = 0$ and P varies slowly with x, then a Taylor expansion to the second order in r leads to the so-called Fokker–Planck equation (also called the "Smoluchowski equation," "second Kolmogorov equation," or "generalized diffusion equation")

$$\frac{\partial}{\partial t} P(x,t) = \frac{\partial}{\partial x}[a_1(x)P] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[a_2(x)P] \qquad [30]$$

where

$$a_v(x) = \int_{\Omega} r^v W(x; r)\, \mathrm{d}r \qquad [31]$$

If $a_1 = 0$ and a_2 is independent of x, the Fokker–Planck equation becomes the standard diffusion equation describing the Brownian motion. In more than one dimension, the coefficients a_1 and a_2 become a vector and a tensor, respectively.

The Wiener Measure and the Feynman–Kaç Integral

Wiener introduced a probability measure for the trajectories of the Brownian motion. For the sample space Ω, the set of all real-valued functions $x(t)$ $(0 < t < \infty)$, normalized by the condition $x(0) = 0$, is first taken. As a family of "elementary sets," one takes sets of functions defined by conditions of the form

$$\{\alpha_1 < x(t_1) < \beta_1, \alpha_2 < x(t_2) < \beta_2, \ldots, \alpha_n < x(t_n) < \beta_n\}$$
$$0 < t_1 < t_2 \cdots < t_n \qquad [32]$$

The measure (Wiener measure) assigned to the set eqn [32] is given by the Einstein–Smoluchowski formula (eqn [19]) with $P_{1|1}(x|y;\ t_2 - t_1)$ as given by eqn [11]. It is now easily verified that this particular assignment of measures to the elementary sets satisfies the usual consistency conditions. In fact, it is seen that consistency conditions are implied by the Chapman–Kolmogorov equation [10]. Once consistency is established, one can construct (by a general theorem of Kolmogorov) a completely additive measure in the space of all real-valued functions $x(t)$ $(x(0) = 0)$. So constructed, this measure is nearly useless because many sets of direct relevance and interest are nonmeasurable. For instance, the set C of continuous functions turns out to be nonmeasurable. It has been shown by Doob that if one restricts to continuous functions and maintains the measures [19] for the sets in [32] (but now only continuous functions are allowed in these sets), then a completely additive measure μ^w with $\mu^w(C) = 1$ can be constructed. The set of differentiable functions turns out to have a vanishing measure; this means that the instantaneous velocity of the Brownian motion as described by the Wiener process is meaningless.

Having a completely additive measure in C, one can define an integral (Wiener integral) which has all the basic properties of the Lebesgue integral. For any continuous function $V(x)$ defined in $(-\infty, \infty)$ with $V(x) > 0$, Kaç considered the Wiener integral

$$I = \int_{a<x(t)<b} \exp\left[-\int_0^t V(x(\tau))\mathrm{d}\tau\right] \mathrm{d}\mu_t^{\mathrm{W}} \qquad [33]$$

which is easily shown to exist and to be equal to the ordinary integral

$$I = \int_a^b \psi(x,t)\, \mathrm{d}x \qquad [34]$$

where $\psi(x,t)$ is given by

$$\psi(x,t) = \sum_{k=0}^{\infty} (-1)^k \psi_k(x,t) \qquad [35]$$

where

$$\psi_0(x,t) = (2\pi t)^{-1/2} \exp\left(-\frac{x^2}{2t}\right)$$
$$\psi_{n+1}(x,t) = \int_0^t \int_{-\infty}^{\infty} [2\pi(t-\tau)]^{-1/2} \times \exp\left[-\frac{(x-y)^2}{2(t-\tau)}\right] V(y)\psi_n(y,\tau)\, \mathrm{d}y\, \mathrm{d}\tau \qquad [36]$$

As a consequence, $\psi(x,t)$ satisfies

$$\psi(x,t) = \psi_0(x,t) - \int_0^t \int_{-\infty}^{\infty} [2\pi(t-\tau)]^{-1/2} \times \exp\left[-\frac{(x-y)^2}{2(t-\tau)}\right] V(y)\psi(y,\tau)\, \mathrm{d}y\, \mathrm{d}\tau \qquad [37]$$

and hence, the partial differential equation

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}\frac{\partial^2 \psi}{\partial x^2} - V(x)\psi \qquad [38]$$

that is, there is a relation between the solutions of this equation and the functional integral [33].

In his research, Kaç was inspired by the famous Feynman path integral. If the connection is considered, the integral I can be seen as the limit

$$\lim_{n\to\infty}(2\pi t n^{-1})^{-n/2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left[-tn^{-1}\left\{\sum_{k=1}^{n}V(x_k)+\frac{1}{2}\sum_{k=1}^{n}\left(\frac{x_k-x_{k-1}}{tn^{-1}}\right)^2\right\}\right] \quad [39]$$

which makes it obvious that the exponent is simply a discretization of

$$-\int_0^t\left\{\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+V(x(\tau))\right\}\mathrm{d}\tau \quad [40]$$

In this form, the symbolism is physically appealing, since the expression in braces is the Hamiltonian of a particle under the action of a potential $V(x)$. Following Feynman's notation, the functional integral [33] may be written as

$$\int\exp\left[-\int_0^t\left\{\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+V(x(\tau))\right\}\mathrm{d}\tau\right]\mathrm{d}(\text{path}) \quad [41]$$

Actually Feynman, in his approach to nonrelativistic quantum mechanics, was led to considering

$$\int\exp\left[\mathrm{i}\hbar^{-1}\int_0^t\left\{\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2-V(x(\tau))\right\}\mathrm{d}\tau\right]\mathrm{d}(\text{path}) \quad [42]$$

where $\hbar = h/2\pi$ (h = Planck's constant). The functional

$$A=\int_0^t\left\{\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2-V(x(\tau))\right\}\mathrm{d}\tau \quad [43]$$

is the classical action along the path $x(t)$. Because of the factor i ($=\sqrt{-1}$) in the exponent, Feynman's theory is not easily made rigorous. One way is to replace $\hbar^{-1}$ by $\hbar^{-1}+\mathrm{i}\varepsilon$; this gives an integral which can be treated as Kaç's integral, and then one can pass to the limit when $\varepsilon\to 0+$. Of course, the partial differential equation [38] then becomes the Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2}\frac{\partial^2\psi}{\partial x^2}+V(x)\psi \quad [44]$$

When $\hbar$ tends to zero, the trajectories become very irregular and their contributions to the Feynman integral tend to cancel by interference, except for the contribution of the classical trajectory.

Stochastic Differential Equations and Stochastic Mechanics

In many cases, the coefficients a_1 and a_2 appearing in eqn [30] permit one to reconstruct the corresponding diffusion process (i.e., to find $P(y, t)$). Another way of constructing a diffusion process from these coefficients is based on the theory of the stochastic differential equations. It can be said that the stochastic process X is a solution of the stochastic differential equation

$$\mathrm{d}X=\sigma(t,X)\,\mathrm{d}W_t+b(t,X)\,\mathrm{d}t \quad [45]$$

with initial condition $X = X_0$, if

$$X=X_0+\int_0^t\sigma(t,X)\,\mathrm{d}W_t+\int_0^t b(t,X)\,\mathrm{d}t \quad [46]$$

where $b(t, x)$ is uniformly Lipschitzian in t, $W_t = W(t)$ is the Wiener process, and the first integral makes sense as a stochastic integral; in the simplest case, σ is a constant and the integral equals $\sigma(W(t) - W(0))$. The notation [45] is a purely symbolic notation in place of the rigorous one in [46] because, generally speaking, the trajectories of the processes $W(t)$ and $X(t)$ are continuous but almost nowhere differentiable, and hence, $\mathrm{d}X/\mathrm{d}t$ does not exist with probability 1. There are very general theorems which state the existence of solutions to eqn [46]. The relation between the two methods occurs through an important formula due to Itô. Frequently, the stochastic equation for a diffusion process is called the Langevin equation in physical literature.

It is clear that

$$\begin{aligned}\lim_{\tau\to 0+}\frac{\langle[X(t+\tau)-X(t)]^2\rangle}{\tau}&=[\sigma(t)]^2\\ \lim_{\tau\to 0+}\frac{\langle[W(t+\tau)-W(t)]^2\rangle}{\tau}&=[\sigma(t)]^2\end{aligned} \quad [47]$$

Thus, the coefficient $\sigma(t)$ is related to $a_2(t)$, used before, by $a_2(t) = [\sigma(t)]^2$.

Nelson has proposed a stochastic kinematics for Markov processes. He defines a forward velocity DX (the limit of the expected value of $(X(t+\tau)-X(t))/\tau$ conditioned by the past history of the process when $\tau\to 0+$) and a backward velocity D_*X (the analogous limit of the expected value of $(X(t)-X(t-\tau))/\tau$ conditioned by the future history of the process). They both coincide with $\dot{x}$ when the particular history $x(t)$ is differentiable. $v=(DX+D_*X)/2$ is called the mean velocity and $u=(Dx-D_*X)/2$ the osmotic velocity. For the acceleration, one can consider DDX, D_*D_*X, D_*DX, DD_*X, and even more choices are available for the mean acceleration a ($(DDX+D_*D_*X)/2$, $(DD_*X+D_*DX)/2$, $(DDX+D_*\,D_*X+$

$DD_*X + D_*DX)/4$). By giving examples, Nelson shows that the only choice which makes sense is $a = (DD_*X + D_*DX)/2$.

Assuming a process with forward velocity $b = b(x,t)$, backward velocity $b_* = b_*(x,t)$, and constant diffusion coefficient $\nu = \sigma^2/2$, it is easy to show that the expected value for the derivative of any function $f(X,t)$ is, by [47],

$$\langle Df \rangle = \lim_{\tau \to 0+} \left[\frac{\partial f}{\partial t} + \frac{\langle [X(t+\tau) - X(t)] \rangle}{\tau} \frac{\partial f}{\partial x} + \frac{\langle [X(t+\tau) - X(t)]^2 \rangle}{2\tau} \frac{\partial^2 f}{\partial x^2} \right] = \frac{\partial f}{\partial t} + b \frac{\partial f}{\partial x} + \nu \frac{\partial^2 f}{\partial x^2} \quad [48]$$

and, similarly

$$\langle D_* f \rangle = \frac{\partial f}{\partial t} + b_* \frac{\partial f}{\partial x} - \nu \frac{\partial^2 f}{\partial x^2} \quad [49]$$

In a similar way, one can obtain the Fokker–Planck equations satisfied by the backward and forward processes

$$\frac{\partial P}{\partial t} = -\frac{\partial (bP)}{\partial x} + \nu \frac{\partial^2 P}{\partial x^2} \quad [50]$$

$$\frac{\partial P}{\partial t} = -\frac{\partial (b_* P)}{\partial x} - \nu_* \frac{\partial^2 P}{\partial x^2} \quad [51]$$

In principle, the probability density in the second equation should be $P_* \neq P$, but Nelson says that at the basic level, there should be time symmetry and, more or less explicitly, assumes $P = P_*$, $\nu = \nu_*$. This implies

$$u = \frac{b - b_*}{2} = \nu \frac{1}{P} \frac{\partial P}{\partial x} \quad [52]$$

Since $v = (b + b_*)/2$, one also has

$$\frac{\partial P}{\partial t} + \frac{\partial (Pv)}{\partial x} = 0 \quad [53]$$

The mean acceleration a can be easily shown to be (by [48] and [49])

$$a = \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} - u \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} \quad [54]$$

If $a = F/m$, where F is the force, one obtains the dynamics of the stochastic motion, described by [53] and [54], provided u is given by eqn [52]. It is an extremely remarkable fact that, by a change of unknowns ($\psi = \sqrt{P} \exp(\mathrm{i} \int_0^x v(x')\, \mathrm{d}x'/(2\nu))$), this system becomes the Schrödinger equation (eqn [44]), provided $\nu = \hbar/2m$. In more than one dimension, this holds if and only if the force is a gradient plus (possibly) a Lorentz force.

Equations [53] and [54] (with $a = F/m$) constitute the stochastic version of quantum mechanics proposed by Nelson.

See also: Irreversible Thermodynamics and Basic Transport Theory in Solids; Statistical Mechanics: Classical; Statistical Mechanics: Quantum.

PACS: 02.50.Cw; 02.50.Ey; 02.50.Fz; 05.10.Gg

Further Reading

Kaç M (1959) *Probability and Related Topics in Physical Sciences*. London: Interscience.

van Kampen NG (1981) *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland.

Nelson E (1967) *Dynamical Theories of Brownian Motion*. Princeton: Princeton University Press.

Wax N (ed.) (1954) *Selected Papers on Noise and Stochastic Processes*. New York: Dover.

Strengthening Mechanisms in Metals *See* Mechanical Properties: Strengthening Mechanisms in Metals.

Strongly Correlated Electron Systems

H v Löhneysen, Universität Karlsruhe, Karlsruhe, Germany

Introduction

Atomic magnetic moments arise from the presence of partly filled orbitals classified by their main quantum number n and angular momentum $l \leqslant n-1$ in the atomic configuration of the elements. In addition, the electron spin also contributes to the magnetic moment. In the solid state, the electrons occupying these orbitals are subject to the electrostatic potential not only of their "own" ion core but also of the potential generated by the surrounding ions and electrons. This effect depends on the spatial extent of the

wave function of the partly filled shells. Electrons in strongly localized orbitals behave like localized moments retaining their atomic quantum numbers, with the ground-state multiplet given by Hund's rules. Hence the magnetic susceptibility χ increases with decreasing temperature T, often in a Curie–Weiss behavior $\chi \sim (T-\theta)^{-1}$, where θ is a measure of the interaction between moments. The ground-state multiplet is split by the weak crystalline electric field of the neighboring ions. On the other hand, electrons in spatially extended orbitals such as s or p electrons, in general, delocalize, resulting in a weak Pauli susceptibility independent of T. An interaction between magnetic moments leads to various types of magnetic order.

In nature, magnetic properties of solids often fall between these two extremes – localized versus delocalized moments – where the electron correlations are particularly strong and lead to a wealth of novel effects. One can order the metallic elements in a new fashion, not with increasing atomic number as in the periodic table of the elements, but with increasing spatial extent of the partly filled orbitals. The approximate sequence, $4f-5f-3d-4d-5d$, maps the increasing degree of delocalization. In intermetallic $4f$ systems, due to the strong localization of the $4f$ electrons, charges of the surrounding ligands disturb the spherically symmetric atomic potential only weakly, providing a splitting of the Hund's rule ground-state multiplet, commonly referred to as crystalline electric field (CEF) splitting. The ground state Hund's rule multiplet is determined by (1) maximizing the total spin S, (2) maximizing the total angular momentum L compatible with (1), and (3) taking account of the spin–orbit (SO) interaction with $J = |L \mp S|$ where $-(+)$ applies when the shell is less (more) than half-filled. In $3d$ systems, on the other hand, the SO coupling, depending strongly on the atomic number ($\sim Z^4$), is rather weak. Here the d-electron wave functions are rather extended, so that the ligands strongly affect the wave functions and the orbital moment ceases to be a good quantum number ("quenching" of orbital momentum). This often leads to a "spin only" magnetic moment in insulating $3d$ systems. This article focuses on intermetallic compounds and alloys where interactions between electrons lead to a wealth of new phenomena. Of particular interest are systems where the electrons in partly filled orbitals behave "between" being localized and delocalized.

Fermi Liquids

In a metal, the conduction electrons interact with each other due to their Coulomb repulsion. At the same time, there is a strong correlation by virtue of the Pauli principle, which states that two electrons (or, more generally, two indistinguishable fermions, i.e., particles with half-integer spin) cannot be in the same quantum state. The physics of a noninteracting electron gas is well known. For $kT \ll E_F$, the specific heat of a noninteracting electron system is given by $C = \gamma_0 T$ with $\gamma_0 = (\pi^2 k^2/3) N(E_F)$, where $N(E_F)$ is the density of levels at the Fermi energy E_F, and k the Boltzmann constant. For free electrons, $N(E_F) = m_0 k_F/(\hbar^2 \pi^2) = (3/2) N_A/E_F$ where m_0 is the free-electron mass, k_F the Fermi wave number, and N_A the Avogadro number. The Pauli spin susceptibility is independent of T, $\chi_0 = \mu_0 \mu_B^2 N(E_F)$ where $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{V\,s\,A^{-1}\,m^{-1}}$ and μ_B is the Bohr magneton. The contribution to the electrical resistivity due to electron–electron (e–e) scattering is given by $\rho_{e-e} = c(kT/E_F)^2$ where c is a microscopic scattering cross section. Because of the large E_F in simple metals, ρ_{e-e} is difficult to observe.

Assuming a one-to-one correspondence between the excitations of an interacting electron system, termed "quasiparticles," and of the noninteracting electron gas, Landau postulated that the low-T properties for the interacting system obey the same laws as for the Fermi gas, with a renormalized effective mass m^* (with respect to the free-electron mass m_0) and a few additional parameters taking account of the residual interactions among the quasiparticles. This special case of an interacting electron system which can be obtained from the noninteracting Fermi gas by adiabatically turning on the interaction is called a (Landau) Fermi liquid. In such a Fermi liquid (FL), the specific heat is given by $C = \gamma T = (m^*/m_0)\ \gamma_0 T$, the spin susceptibility by $\chi = \chi_0 (m^*/m_0)(1 + F_0^a)$ where F_0^a is an additional Landau parameter, and the resistivity by ρ_{e-e}, which is again proportional to $(kT/E_F)^2$ where E_F may be modified due to the interactions.

Heavy-Fermion Systems

Heavy-fermion systems (HFS) are perhaps the most remarkable manifestation of strongly correlated electron systems. These are compounds with a (often) regular sublattice of $4f$ or $5f$ atoms, notably Ce, Yb, or U. They are close to a magnetic instability giving way to long-range magnetic order arising from the interactions between the $4f$ or $5f$ magnetic moments via the conduction electrons. It is instructive to consider first the case of an isolated magnetic impurity in a nonmagnetic metallic host. The hybridization between conduction electrons and f electrons, together with a large on-site Coulomb repulsion U between f electrons, may lead to a complete screening of the magnetic moment, resulting in a singlet ground state. (Historically, this interaction was first considered

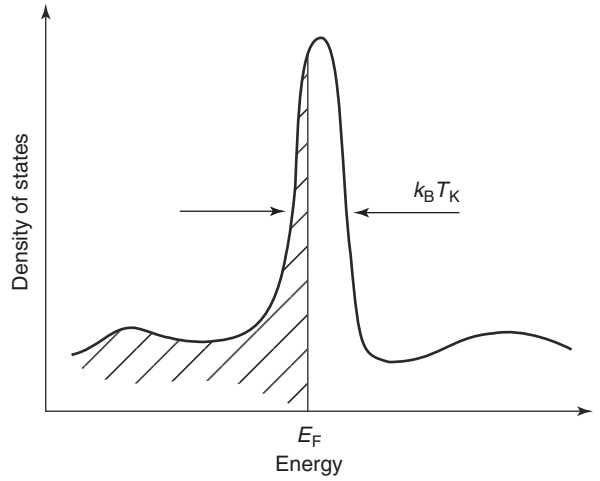


Figure 1 Local Kondo resonance at the Fermi level in the electronic density of states of a metal with a magnetic impurity. The width of the resonance is given by the Kondo temperature T_K and its height by $1/T_K$. In a simple picture, the density of states of a heavy-fermion compound may be viewed as a lattice-coherent superposition of local Kondo resonances.

between conduction electrons and localized d electrons, hence often termed as "s–d exchange" in older literature.) The interaction between conduction electrons and f electrons can be described in terms of the Kondo Hamiltonian $\mathscr{H} = -JS \cdot s$, where S is the impurity spin and s the conduction-electron spin, with an effective antiferromagnetic (AF) exchange interaction J. This screening is dynamic and can be visualized as being due to virtual transitions between the localized f level and the conduction-electron states at E_F, leading to a resonance at E_F, the so-called Kondo or Abrikosov–Suhl resonance (**Figure 1**). The "binding energy" of this singlet state is of the order kT_K. At sufficiently low T well below the characteristic Kondo temperature T_K, a local FL is formed around the magnetic impurity, exhibiting a huge specific heat $C \sim kT/T_K$. By comparison with the free-electron expression for γ_0, T_K plays the role of an effective Fermi temperature, $T_F = E_F/k$.

In a grossly simplified picture, HFS can be considered as lattice-coherent superpositions of these local FLs. Hence at low T, most HFS can be described within the framework of the FL theory with, however, a very large effective mass m^* derived from the huge linear specific heat coefficient $\gamma = C/T$ – hence the term heavy fermions – and a correspondingly large Pauli susceptibility χ, both being only weakly dependent on T. m^* exceeds m_0 by a factor of several hundreds in some cases. The contribution $\rho_{e-e} \sim (T/T_K)^2$ to the resistivity is readily observable in HFS, because the characteristic temperature scale $T_K \sim 10$ K is much smaller than T_F in typical metals.

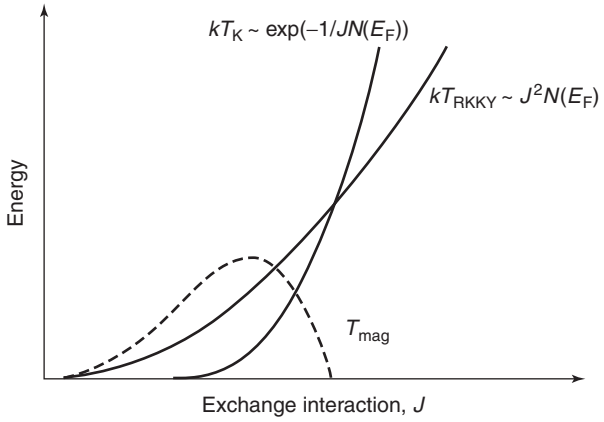


Figure 2 Competition between the on-site Kondo energy kT_K favoring singlet formation and the amplitude of the oscillating intersite RKKY interaction kT_{RKKY} as a function of conduction electron–4f electron exchange J, leading to magnetic order T_{mag} at intermediate values of J (dashed line).

The competition between the on-site Kondo interaction just discussed (quenching the 4f or 5f localized magnetic moments) and the spatially oscillating intersite Ruderman–Kittel–Kasuya–Yosida (RKKY) interaction between these moments mediated by the conduction electrons via the exchange interaction J allows for nonmagnetic or magnetically ordered ground states in HFS. This competition is, in principle, governed by a single parameter, namely the effective exchange constant J which enters the characteristic energy scales $kT_K \sim \exp(-1/N(E_F)J)$ and $kT_{RKKY} \sim J^2N(E_F)$ for Kondo and RKKY interactions, respectively. **Figure 2** illustrates the competition between T_K and T_{RKKY}, first pointed out by Doniach. The magnitude of J is usually tuned by composition or pressure. Because of the exponential dependence of T_K on J which in turn depends on the hybridization, volume changes are often the dominant effect in producing a magnetic–nonmagnetic transition. The hybridization can, of course, be strongly affected when alloying with nonisoelectronic ligands.

Magnetic Phase Transitions in Strongly Correlated Electron Systems

Magnetic phase transitions are usually observed by varying the temperature T and monitoring changes in thermodynamic quantities such as the magnetic susceptibility or the specific heat. Second-order transitions, where the first derivatives of the free energy, that is, the magnetization or entropy, vary continuously, are strongly influenced by thermal fluctuations. These fluctuations determine how at the critical temperature T_c, the system goes from a disordered state

at $T > T_c$ to a magnetically ordered state at $T < T_c$. The fluctuations diverge spatially as T_c is approached, as characterized by the correlation length $\xi \sim |T - T_c|^{-\nu}$ with a critical exponent ν, and have a typical lifetime given by $\tau \sim \xi^z$, where z is called the dynamical critical exponent. The renormalization group theory predicts that the static properties near a phase transition (PT) depend on a few parameters only, namely the spatial dimension of the system d, the symmetry of the order parameter (OP), and the range of the interaction leading to magnetic order. These parameters determine the universality class of the system. For magnetic systems, the OP is usually the spontaneous magnetization (for antiferromagnets and incommensurate magnets, the staggered magnetization), which is zero above T_c and acquires a finite value below T_c. In HFS, an AF or incommensurate order is found much more often than a ferromagnetic (FM) order.

In many cases, for example, when the magnetic moments are spatially localized, the symmetry of the OP is described completely by the number n of its components, that is, $n = 1$ (Ising model), $n = 2$ (XY model), or $n = 3$ (Heisenberg model). While the critical behavior of static properties such as the specific heat or magnetization for these examples is entirely determined by the universality classes, the dynamic behavior within a given class may differ, depending on z. Fluctuations can actually suppress a PT altogether, the more so, the easier it is to generate fluctuations, that is, the larger n is. A PT to a long-range order is only possible above a lower critical dimension d_L. $d_L = 2$ for $n = 3$, $d_L = 2$ for $n = 2$, and $d_L = 1$ for $n = 1$. (For $n = 2$ and $d = 2$, a special type of PT, the Kosterlitz–Thouless transition, exists.) On the other hand, there is an upper critical dimension d_U above which fluctuations in local-moment systems are deemed unimportant because they effectively average out, with $d_U = 4$ for a wide range of models. In general, for $d > d_U$ a mean-field theory (MFT) accurately describes the PT.

Of particular interest is the possibility that T_c may be driven to zero by some external parameter δ. This can indeed occur in HFS, where δ can be pressure or composition, thus tuning J. The picture of a "classical" phase transition (CPT) has to be modified when T_c vanishes at a critical value δ_c. For such a quantum phase transition (QPT), the relevant length scales and timescales of the critical fluctuations are coupled via the uncertainty principle. At finite temperatures, the quantum energy of fluctuations $\hbar/\tau$ becomes relevant when $\hbar/\tau > kT$, leading to the time entering as additional z dimensions. For any finite T_c, the quantum energy becomes irrelevant because $\tau \to \infty$ for $T \to T_c$ (critical slowing down). Formally, the inverse temperature corresponds to an imaginary time, hence the system acquires its full effective dimensionality $d_{eff} = d + z$ only at $T = 0$. The added time dimensions drive the system toward a mean-field behavior, where above $d_{eff} = 4$ only Gaussian fluctuations are observed and genuine critical behavior is absent. In all cases of a QPT, however, the static properties are influenced by z. This leads to a behavior even at finite T close to δ_c, which distinctly differs from that at a CPT.

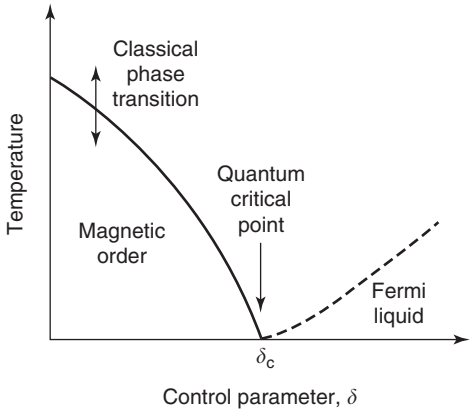


Figure 3 Schematic phase diagram of a metal with an instability towards magnetic order driven by a control parameter δ which may be pressure or composition.

A QPT scenario relevant to HFS is sketched in **Figure 3**, where an itinerant electron system is driven toward (or away from) the magnetic order by δ. While far away from δ_c, the low-T properties of the system can be described in the framework of an FL; unusual "non-Fermi liquid" (NFL) behavior is observed at finite T around δ_c. This picture assumes that the heavy quasiparticles generated by virtue of the Kondo effect retain their identity when passing through the quantum critical point (QCP), while part of the electron system goes into a spin–density wave (SDW) state. This model is in accordance with the simple Doniach picture (**Figure 2**). An alternative view, based on experimental inelastic neutron scattering (INS) data on $CeCu_{6-x}Au_x$ (see next section), suggests that the Kondo effect itself becomes inoperative at the QCP, leading to the appearance of nearly independent local moments. NFL signatures close to a QPT have been observed in weakly magnetic transition-metal compounds and in rare-earth based HFS. On the other hand, NFL behavior in metals does not necessarily arise from the proximity to a magnetic instability. Single-ion effects such as the multichannel Kondo effect or a distribution of Kondo temperatures may also lead to NFL-like properties.

When magnetic order is driven towards an instability by a control parameter, the relevant interactions also become weak or cancel each other. Hence

new low-energy scales may appear which are otherwise masked by the magnetic order. This may lead to an inherent instability of a QPT itself, in the sense that residual interactions drive the system toward a first-order transition where the coherence length and time of the critical fluctuations remain finite, or altogether new phases may appear. A prominent example of the latter is the occurrence of superconductivity (SC) near a QCP in a number of HFS (see below).

A recently discovered very interesting example of a possible novel phase is the partial order in MnSi. This compound exhibits a long-wavelength ($\lambda = 175$ Å) helical magnetic order below $T_c = 29.5$ K. Upon application of hydrostatic pressure p, T_c is reduced, with $T_c = 0$ for $p_c = 14.6$ kbar. In the vicinity of and extending considerably even above p_c, the helical order persists partially: while the λ remains practically constant, the orientation of the helix pointing toward the $\langle 1\,1\,1 \rangle$ directions at $p = 0$, is largely lost. This resembles the structure of certain liquid-crystal phases. An open question at present is how this partial order, here observed for the first time in a magnetic system, is related to the NFL resistivity $\Delta\rho \sim T^{1.5}$ found over a large T and p range.

Magnetic Instability in Heavy-Fermion Systems: $CeCu_{6-x}Au_x$, A Case Study

A large number of investigations have been carried out on HFS close to the magnetic instability. NFL behavior, that is, deviations from the FL predictions in the low-T thermodynamic and transport properties has been reported for systems too numerous to be mentioned here. The main observations are a specific heat coefficient γ exhibiting a logarithmic T dependence, saturating toward low T in a number of cases, and a T^m dependence of the electrical resistivity ρ with $m < 2$. There are only relatively few magnetization studies. The most decisive probe for magnetic fluctuations near a QPT, INS, has been employed only for a few systems so far.

Since many of the HFS are driven through the QPT by changing the composition, introducing disorder, its effect on the critical behavior is an important issue. One has to distinguish between disorder by dilution of the 4*f* or 5*f* site as, for example, in $Ce_{1-x}La_xRu_2Si_2$, or by altering the ligand configuration as, for example, in $CeCu_{6-x}Au_x$. In the first case, a "Kondo hole" introduced by dilution might lead to a substantial quasiparticle scattering and an ultimate loss of coherence. In the second case, the Ce atoms experience different local environments and this may lead to different local Kondo temperatures. In $CeCu_{6-x}Au_x$, the QPT was tuned both by composition and pressure, with no difference in the functional $C(T)$ and $\rho(T)$ dependencies. Disorder may, however, play an important role in other systems.

Table 1 Predictions of the SDW model for the temperature dependence of the linear specific heat coefficient C/T and the electrical resistivity ρ at the quantum critical point. d is the dimensionality of the system and z the dynamic critical exponent for an FM and an AF

	$d = 3$		$d = 2$	
	$z = 3$ (FM)	$z = 2$ (AF)	$z = 3$ (FM)	$z = 2$ (AF)
C/T	$\sim -\ln(T/T_0)$	$\sim \gamma_0 - \beta\sqrt{T}$	$\sim T^{-1/3}$	$\sim -\ln(T/T_0)$
ρ	$\sim T^{5/3}$	$\sim T^{3/2}$	$\sim T^{4/3}$	$\sim T$

Practically, all HFS investigated so far are close to an AF instability (as opposed to an FM instability). A number of systems, notably $Ce_{1-x}La_xRu_2Si_2$, have been compared in detail with the SDW model (see **Table 1**). A long-range AF is introduced into $CeRu_2Si_2$ by a moderate La doping because of the concomitant lattice expansion. The behavior near the QPT, occurring for $x_c \sim 0.075$, is broadly consistent with the expectations, although some differences between INS-derived and specific-heat-derived parameters remain.

In $CeCu_{6-x}Au_x$ which has been studied most extensively, the QPT can be tuned by the Au concentration x, again by virtue of a weakening of J upon expansion of the $CeCu_6$ lattice by Au doping. At the critical concentration $x_c \sim 0.1$, an NFL behavior occurs with $\gamma = a\ln(T_0/T)$ observed as temperature T changes by two orders of magnitude (see **Figure 4**), and $\rho \approx \rho_0 + AT$, between 0.02 and 0.5 K. These T dependences can be understood in terms of the SDW model when two-dimensional critical fluctuations are assumed (hence $d = 2$), which are coupled to quasiparticles with three-dimensional dynamics. Indeed, INS experiments for $x = 0.1$ reveal quasi-one-dimensional features in the dynamical structure factor $S(q; E = \text{const})$ which correspond to quasi-two-dimensional "rods" in real space. Here, the neutron energy transfer was $E = 0.15$ meV. For strictly two-dimensional fluctuations, the system would be at the upper critical dimension $d_{\text{eff}} = d + z = 4$.

The energy dependence of the critical fluctuations measured at $\mathbf{Q} = (0.8\,0\,0)$, that is, at a $\boldsymbol{q}$ point on the "rods" where $S(\boldsymbol{q}, E = \text{const})$ is maximal, has revealed an anomalous scaling with E and T of the imaginary part of the dynamical susceptibility $\chi''(\boldsymbol{q}, E, T) = S(\boldsymbol{q}, E, T) \times (1 - \exp(-E/kT))$, that is, $\chi''(E, T) = T^{-\alpha} g(E/kT)$ with $\alpha = 0.74$ (**Figure 5**). The scaling is incompatible with the SDW scenario, and the scaling function g is consistent with a simple

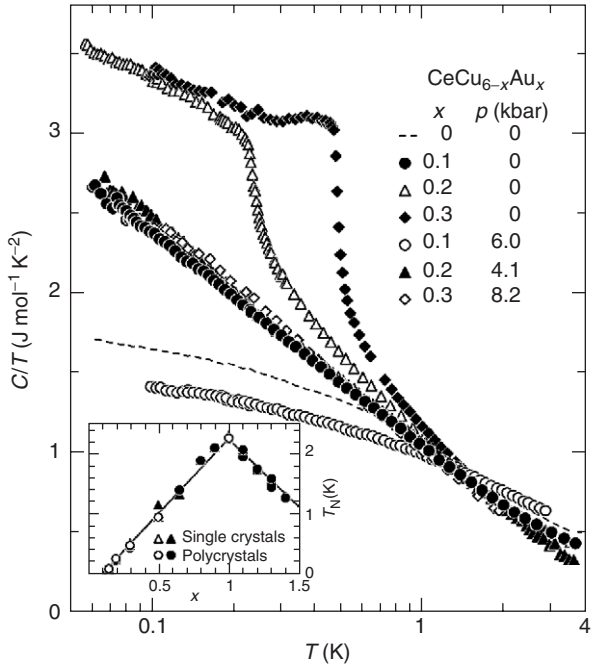


Figure 4 Specific heat C of $CeCu_{6-x}Au_x$ plotted as C/T vs. log T for different hydrostatic pressures p. Note that when $T_N = 0$, (i.e., for $x = 0.3$ at 8.2 kbar and for $x = 0.2$ at 4.1 kbar) the C/T curve falls on the same "universal curve" as for $x = 0.1$ where $T_N = 0$ at ambient pressure. Here a pressure of 6.0 kbar leads to a FL behavior $C/T \approx$ const. Dashed line indicates C/T for $CeCu_6$. Inset shows the Néel temperature T_N vs. x at $p = 0$. (Adapted from von Löhneysen H (1999) Fermi-liquid instability at magnetic-nonmagnetic phase transitions. *Journal of Magnetism and Magnetic Materials* 200: 532–551.)

form of the dynamical susceptibility

$$\chi(\boldsymbol{q}, E, T) = c(f(\boldsymbol{q}) + (-\mathrm{i}E + bT)^{\alpha})^{-1} \qquad [1]$$

where c and b are constants. The anomalous T dependence of the uniform static magnetic susceptibility can be modeled up to 7 K by $\chi = c'(\chi_0^{-1} + b''T^{\alpha})^{-1}$ with $\alpha = 0.8$. The fact that α as obtained from INS for $\boldsymbol{Q} = (0.8\,0\,0)$ is equal to α $(q = 0, E = 0)$ may imply that: (1) the nonlocal physics is contained in $f(\boldsymbol{q})$ which vanishes as $\boldsymbol{q} \to \boldsymbol{Q}$ where $\boldsymbol{Q}$ is a critical vector, that is, belongs to the quasi-one-dimensional features of $S(\boldsymbol{q}, E = \text{const})$, and (2) the anomalous scaling exponent is a local property since it appears to be independent of $\boldsymbol{q}$. This "local" behavior is, however, different from a simple local Lorentzian fluctuation spectrum that would be described by $\alpha = 1$. **Figure 5** also shows that the uniform magnetization obeys a field-temperature scaling with the same anomalous exponent α.

The QPT in $CeCu_{6-x}Au_x$ has not only been tuned by x but also by pressure. **Figure 4** shows that magnetically ordered samples with $x > x_c$ can be driven to the QPT, that is, $T_N = 0$ by appropriate hydrostatic pressure p. This proves that lattice expansion by Au is indeed the dominant effect leading to a long-range magnetic order. Near the QPT, C/T is identical for all samples indicating the equivalence of pressure and concentration tuning in this system. Likewise, the sample with $x = 0.1$ can be driven by p away from the QPT toward an FL behavior $C/T \sim$ const approximately followed by pure $CeCu_6$. The linear $T_N(p)$ and $T_N(x)$ dependencies are in agreement with the scenario of two-dimensional fluctuations. The detailed mechanism of how local moments and long-range fluctuations work together to produce the unique quantum-critical behavior of $CeCu_{6-x}Au_x$ has yet to be disentangled.

The QPT in the related system $CeCu_{6-x}Ag_x$ and also in $CeCu_{6-x}Au_x$ has been tuned by a magnetic field. The T dependencies of C/T and ρ are compatible with the SDW scenario with $d = 3$ and $z = 2$ (**Table 1**), although this scenario does not include a magnetic field.

Thorough studies of $YbRu_2Si_2$, which is close to quantum criticality at ambient pressure (AF order occurs at $T_N \sim 70$ mK), have shown not only similarities to $CeCu_{6-x}Au_x$ in that $\Delta\rho \sim T$ over an even larger T range, but also differences in the specific heat which exhibits a low-T deviation in excess of the $\gamma \sim \ln(T_0/T)$ behavior. Replacement of Si by Ge (doping of nominally 5 at. % Ge leads to $T_N = 20$ mK) amplifies these differences. The data have been suggested to support the local-criticality model introduced to account for the behavior of $CeCu_{6-x}Au_x$. Furthermore, the disparity between thermodynamic and transport data — the resistivity does not show deviations from $\Delta\rho \sim T$ — has been emphasized for $YbRh_2Si_2$. Unfortunately, neutron scattering data are not yet available for this interesting system.

Superconductivity in the Vicinity of Magnetic Order

Superconductivity (SC) was long believed to be incompatible with a long-range magnetic order. This belief was based on two facts: (1) a magnetic field drives the superconducting transition to lower temperatures, and (2) minute amounts of magnetic impurities often suppress SC. However, as **Figure 6** shows, there are a number of examples where SC exists in the proximity of a magnetic order. The common theme of all these materials is that they exhibit strong electronic correlations. These materials exhibit "unconventional" SC, in the sense that either the symmetry of the Cooper-pair wave function differs from the "conventional" s-wave pairing found for

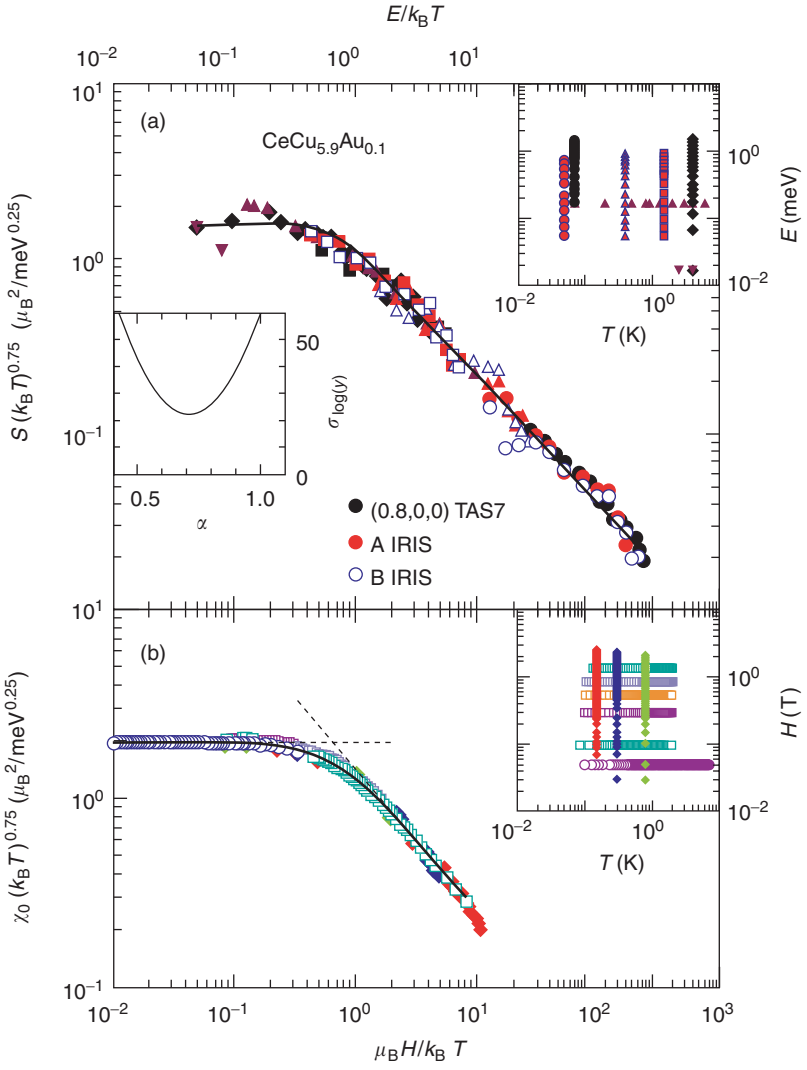


Figure 5 Scaling properties of $CeCu_{5.9}Au_{0.1}$ at quantum criticality (a) Inelastic neutron-scattering function $S(\boldsymbol{q}, E, T)$ multiplied with $T^{0.75}$ vs. E/kT. Different symbols correspond to measurements at different points in $\boldsymbol{q}$ space. The solid line presents a scaling function to eqn [1]. Upper right-hand inset shows the range in neutron energy transfer E and temperature T covered. Left inset shows the "scatter" of the scaling plot as a function of the exponent α, being minimal for $\boldsymbol{Q} = 0.72 \pm 0.05$. (b) Field-temperature scaling of the local contribution to the uniform magnetization $M(T, H)$, with $1/\chi_0 = (dM/dH)^{-1} - (4.1\mu_B^2\ \text{meV})^{-1}$. Solid line corresponds to a scaling function $f(H, T) = (1 + (g\mu_BH/kT)^2)^{-\alpha/2}$ with $\alpha = 0.75$ and $g\mu_B = 1.5\mu_B$. Inset shows range of magnetic field H and temperature T covered. (Adapted from Schröder A, Aeppli G, Coldea R, *et al.* (2000) Onset of antiferromagnetism in heavy-fermion metals. *Nature* 407: 351–355.)

elemental and many alloy superconductors, or the mechanism for an attractive interaction between electrons leading to the formation of Cooper pairs differs from the conventional electron–phonon interactions. Of course, these two aspects are related, as the $\boldsymbol{k}$ dependence of the attractive interaction determines the Cooper-pair wave function. The first "strongly correlated" superconductor, $CeCu_2Si_2$, was discovered in 1979, followed by several other systems. UPt_3 presents a particularly interesting case with several superconducting phases that differ by the spatial dependence of the superconductive order parameter. The high-temperature cuprate superconductors discovered in 1986, that is, $La_{2-x}Sr_xCuO_4$ or $YBa_2Cu_3O_{7-\delta}$, are derived from parent compounds ($x = 0$ or $\delta = 1$, respectively) that are AF insulators because of strong on-site electronic correlations. The CuO_2 planes that are the constituting feature in these materials become metallic and even superconducting upon doping with charge carriers. The interplay between SC and magnetic correlations is not completely understood at present. A particularly interesting case is $CePd_2Si_2$, which in a small p range where T_N is suppressed to zero, exhibits SC below a large T range of NFL behavior. An analogous behavior was found for $CeIn_3$ ($T_N = 10$ K at ambient pressure) and a number of

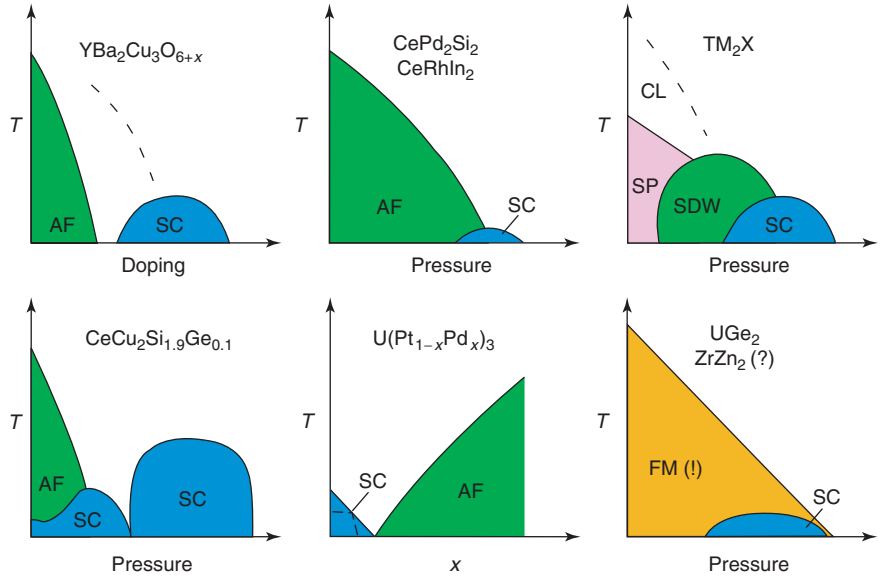


Figure 6 Qualitative picture of the proximity of SC to magnetic order in different types of strongly correlated electron systems. Upper left presents the qualitative phase diagram of cuprate high-temperature superconductors, upper right presents that of quasi-two-dimensional organic conductors. AF = antiferromagnetic order, FM = ferromagnetic order, SC = superconductivity, SP = spin-Peierls phase, SDW = spin-density wave order, CL = charge-localized phase. TM_2X is a quasi-one-dimensional organic conductor, where TM stands for TMTTF ($(CH_3)_4C_6S_4$) or TMTSF ($(CH_3)_4C_6Se_4$), and X for PF_6 or AsF_6 (so-called Bechgaard salts).

other Ce systems. This is suggestive of a spin-fluctuation-mediated SC. A very interesting new material class, based on $CeIn_3$, was recently introduced. This is $CeMIn_5$ where M = Co, Rh, Ir, that is, isoelectronic $3d$, $4d$, and $5d$ elements. While $CeRhIn_5$ is AF ($T_N = 3.8$ K), $CeIrIn_5$ ($T_c = 0.4$ K) and $CeCoIn_5$ ($T_c = 2.3$ K) are superconducting. A pressure of 16 kbar drives $CeRhIn_s$ superconducting with a maximum T_c of 2.1 K. The wealth of new phenomena in these materials is presently exploited by using pressure and composition as tuning parameters, and employing a large number of experimental techniques. However, the detailed interplay between SC and AF order in these materials is not known yet. Low-dimensional organic conductors, also shown by example in **Figure 6**, present an equally rich phase diagram of SC and magnetic order. At first sight, even more surprising, coexistence of SC with ferromagnetism, recently reported for UGe_2, URhGe and $ZrZn_2$ – theoretically anticipated more than 20 years ago – is possible if the spins of the two electrons forming a Cooper pair line up parallel, instead of antiparallel in "classical" superconductors and also in high-T_c cuprates.

Non-Fermi Liquid Behavior Due to Single-Ion Effects

Multichannel Kondo Effect

A perfect screening of the magnetic moment via the Kondo effect, leads to a local FL around a magnetic impurity in a metal as discussed above. In general, however, several conduction-electron channels N may be present and the Kondo Hamiltonian reads $\mathscr{H} = \sum_{\alpha=1}^{N} J_\alpha \boldsymbol{S}\boldsymbol{s}^\alpha$, where α is the channel index. For $N < 2S$, the screening is not complete, whereas $N = 2S$ leads to a perfect screening for $T \to 0$. In the case of degenerate channels, $J_\alpha = J$ for all α, the impurity spin will be overscreened if $N > 2S$. For the special case $S = 1/2$ and $N = 2$, an NFL behavior is predicted with $C/T \sim \chi \sim -\ln(T/T_K)$ and $\rho \sim 1 - (T/T_K)^{1/2}$ for $T \to 0$. Up to now, the experimental evidence for an NFL behavior by overscreening of the magnetic moment of an impurity has been scarce.

In a more general sense, the Kondo effect may also pertain to orbital degrees of freedom of an impurity acting as a pseudospin instead of the spin. This has been discussed for the interaction of conduction electrons with tunneling states, first in amorphous metals and more recently in metallic nanoconstrictions. In these cases, the two channels are provided by the spectator spin of the conduction electrons. For certain rare-earth and actinide metals, a quadrupolar Kondo effect (QKE) has been considered where the pseudospin of the impurity is the quadrupolar moment arising from the aspherical charge distribution of the $4f$ or $5f$ electrons. Again, the two channels are provided by the spin degeneracy of the conduction electrons. Indeed, the NFL anomalies as first reported for $U_{0.2}Y_{0.8}Pd_3$ were suggested to arise from the QKE. Here, NFL anomalies persist to rather dilute U alloys thus favoring

the single-ion QKE model. On the other hand, the NFL behavior might be due to an AF instability, supported by long-range AF order in $U_{0.4}Y_{0.6}Pd_3$. In addition, a divergence of the quadrupolar susceptibility for $T \to 0$ as predicted for the QKE scenario was not found. Hence the situation for $U_xY_{1-x}Pd_3$ is not clear at this point.

The clearest signature of a QKE is found in $U_{1-x}Th_xBe_{13}$. In $U_{1-x}Th_xPd_2Al_3$ single-ion scaling of the NFL properties compatible with the QKE is observed, while $U_{1-x}Y_xPd_2Al_3$ exhibits signatures of a magnetic instability around $x_c = 0.7$ whence the NFL properties might be due to a QPT.

Distribution of Kondo Temperatures

Features of an NFL behavior may arise from quite a different single-ion scenario, that is, a distribution of Kondo temperatures $P(T_K)$ in disordered systems. Fluctuations in J or the local $N(E_F)$ near a disorder-induced metal–insulator transition may lead to a wide distribution of $P(T_K)$. In the case of heavy-fermion solid solutions, the statistical distribution of atoms may lead to a distribution of T_K, as first put forward for the case of $UCu_{5-x}Pd_x$ ($x = 1$ and 1.5) on the basis of NMR measurements which directly showed a distribution of Knight shifts, and hence Kondo temperatures. Even for a rather narrow distribution of J and/or $N(E_F)$, a broad $P(T_K)$ with a finite value for $T_K \to 0$ may occur because of the exponential dependence of T_K on J and $N(E_F)$. A model combining the proximity to a magnetic instability and disorder, the so-called Griffiths phase model, invokes tunneling of ordered regions between different magnetic configurations.

See also: Disordered Magnetic Systems; Ferromagnetism; Kondo Effect (Theory); Liquids, Electronic Structure of; Liquids, Theory of: Fermi Liquids; Localized and Itinerant Magnetism; Quantum Mechanics: Critical Phenomena.

PACS: 71.10.Hf; 75.30.Kz; 75.30.Mb; 71.27.Ta; 74.70.Tx

Further Reading

Custers J, Gegenwart P, Wilhelm H, *et al.* (2003) The break-up of heavy electrons at a quantum critical point. *Nature* 424: 524–527.

Fulde P (1995) *Electron Correlations in Molecules and Solids*, 3rd edn. Berlin: Springer.

Grewe N and Steglich F (1991) Heavy fermions. In: Gschneidner KA Jr. and Eyring L (eds.) *Handbook on the Physics and Chemistry of Rare Earths*, vol. 14, pp. 343–474. Amsterdam: Elsevier.

Hewson AC (1993) *The Kondo Problem to Heavy Fermions*. Cambridge: Cambridge University Press.

Joynt R and Taillefer L (2002) The superconducting phases of UPt_3. *Reviews of Modern Physics* 74: 235–294.

Mathur ND, Grosche FM, Julian SR, *et al.* (1998) Magnetically mediated superconductivity in heavy-fermion compounds. *Nature* 394: 39–43.

Movshovich R, Bianchi A, Jaime M, *et al.* (2002) Unconventional superconductivity in $CeIrIn_5$ and $CeCoIn_5$. *Physica B* 312–313: 7–12.

Pfleiderer C, Reznik D, Pintschovius L, *et al.* (2004) Partial order in the non-fermi liquid phase of MnSi. *Nature* 427: 227–231.

Saxena SS, Agarwal P, Ahilan K, *et al.* (2000) Superconductivity on the border of itinerant-electron ferromagnetism in UGe_2. *Nature* 406: 587–592.

Si QM, Rabello S, Ingersent K, and Smith JL (2001) Locally critical quantum phase transition in strongly correlated metals. *Nature* 413: 804–808.

Stewart GR (2001) Non-Fermi-liquid behavior in *d*- and *f*-electron metals. *Reviews of Modern Physics* 73: 797–856.

Vojta M (2003) Quantum phase transitions. *Reports on Progress in Physics* 66: 2069–2110.

Structural Instability and Soft Modes *See* Lattice Dynamics: Structural Instability and Soft Modes.

Structure Types and Classifications

J Plévert and T J White, Nanyang Technological University, Singapore, Singapore

Introduction

The classification of structure types is an invaluable process for revealing and systematizing relationships between different compounds and for understanding basic principles governing chemical stability and physical properties. Moreover, the *ab initio* computation of atomic arrangements in solids of given stoichiometry remains a key problem in materials science, and although a general predictive method remains elusive, structural classification schemes provide guidance in constraining such calculations.

While more than 150 000 entries can be found in the two main databases of inorganic solids and alloys (ICSD and CRYSTMET, respectively), there is a striking contrast between the simplicity of some compounds such as NaCl, which is readily described by

simple hard sphere packing, and the apparently overwhelming diversity of more complex structures. It should also be added that intricate atomic patterns do not necessarily predicate involved chemistries. For example, boron offers a rich diversity of complex polymorphs.

It is well appreciated that structural diversity originates from manifold atomic interactions – ionic, covalent, van der Waals, and so on – that define bonding characteristics. Even subtle changes in atomic interactions, perhaps as a function of temperature, pressure, or chemistry, can lead to variations in the atomic arrangement that create hurdles for structural cataloging and make the development of systematic taxonomies a challenging task. Nevertheless, the goal of all classification schemes is to establish mutually exclusive categories, while at the same time, striving to be as comprehensive as possible. Practically, all sorting regimes will be complicated, and to some extent limited by intrinsic disturbances such as nonstoichiometry, solid solution series, structural or compositional modulations, polytypism, and so on.

Structure Types

Although the need for categorization was evident from the earliest days of crystallography, the "structure type" concept remains ambiguous; its meaning nuanced according to scientific discipline and the family of solids under consideration. However, the definition of structure type must be resolved as it underpins, and indeed dictates, the lexicon of any classification scheme. To address this issue the International Union of Crystallography has proposed a set of definitions to create a hierarchy-of-resemblance between structures in which:

1. isopointal structures correspond to structures possessing the same space group and Wyckoff sequence;
2. configurational isotypic structures are isopointal structures with each corresponding Wyckoff position showing a similar geometrical environment; and
3. crystal-chemically isotypic structures are configurational isotypic structures with similar physical/chemical characteristics.

It is noted that only isopointal structures are the same and allow a precise assignment of crystallographic equivalence. The next levels classify structures as similar because of the "inherent difficulty in defining *a priori* limits on the similarity," or in other words, the problem of defining mutually exclusive structure types cannot be resolved uniquely.

In the special case of strictly cubic configurational isotypic structures, the similarity ambiguity is removed as all atoms are located in fixed crystallographic positions (the coordination of corresponding atoms is identical), however there is no restriction placed on physical/chemical properties. So, for instance, NaCl (rocksalt) and PbS (galena) are strict configurational isotypes, but not crystal-chemical isotypes (**Figure 1**). Their bonding properties are different – ionic versus covalent character – leading to bandgap energies of about 0.4 eV (PbS) and 8.5 eV (NaCl) and numerous other differentiating characteristics.

Despite such obvious limitations the hierarchy-of-resemblance provides useful guidance for classification, and in-line with its principles, four main styles of representation can be selected for infinitely periodic structures (**Figure 2**). The chemical mode describes the compound as a formula, based on the stoichiometric composition only. The geometric view emphasizes the local environment of individual atoms, particularly the shape of nearest-neighbor coordination polyhedra (normally, but not exclusively cation-centered) that define the coordination number. The structural approach highlights the symmetry of the structure, including the space group, Wyckoff positions, and lattice complexes. The topological description represents structures as a net, where the atoms are vertices and bonds are the edges. These four genres are employed in various forms to systematize mineral, oxide, and alloy structural families. Although many structural classification schemes have been proposed, some are rarely used or applicable to

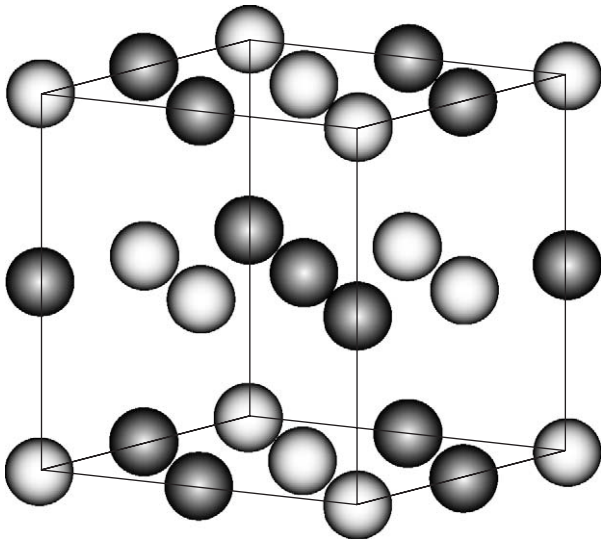


Figure 1 The atomic arrangement in the isopointal structures NaCl (rocksalt) and PbS (galena) are precisely the same although their physical and chemical characteristics are different. The crystallographic or unit cell repeat is outlined.

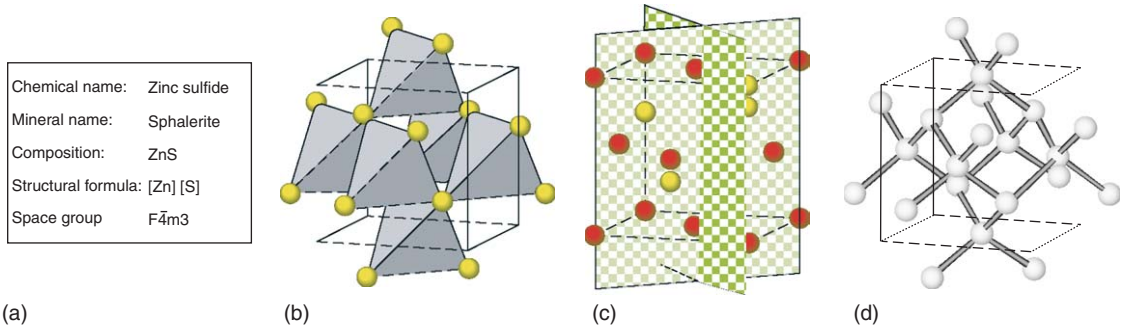


Figure 2 The cubic sphalerite form of ZnS is represented in: (a) the chemical mode where the crystallographically distinct sites are emphasized in square brackets; (b) the geometric mode where the ZnS_4 tetrahedra is shown to form a continuous motif; (c) the structural mode where set of equivalent positions occupied by the atoms Zn and S and symmetry are highlighted, in this case two orthogonal mirror planes are inserted; and (d) the topological mode where a net is constructed with Zn or S at the nodes and connections drawn between these.

a limited subset of materials. Only the better-known sorting methods are considered here.

Strukturbericht Symbols

The antecedent for grouping similar structures was developed by P P Ewald and C Hermann for arranging early crystallographic determinations collated in the journal Strukturberichte (1913–1939). This simple nomenclature was based on chemical composition (**Table 1**), with stoichiometry specified by an upper case letter: A for an element, B for binary, and C for AB_2 compounds. The letter D represents the A_mB_n family that is subdivided according to the A:B ratio such that $D0 \equiv AB_3$, $D1 \equiv AB_4$, $D5 \equiv A_2B_3$, and so on. The letters L and S represent the family of alloys and silicates respectively. A number, or sometimes a letter, follows the composition code, which reflects an entry number.

The type of crystal structure adopted by the elements depends on their position in the periodic table. Most metals are known to crystallize as a close packing of atoms because of the nondirectional character of the metallic bond. The elements in the copper group possess face-centered cubic (f.c.c.) structures, which are assigned the Strukturbericht type *A*1. The alkali metals are found in general as body-centered cubic (b.c.c.) with Strukturbericht type *A*2. Some alkaline earth metals or their polymorphic modifications, as well as zinc and cadmium, adopt the hexagonal close packing (h.c.p.), Strukturbericht symbol *A*3. The diamond type *A*4 is found in the carbon group where covalent bonding is prevalent, and graphite is assigned the Strukturbericht *A*9.

Tetragonal distortions of the f.c.c. (*A*1) may be face-centered tetragonal (f.c.t.) *A*6 or body-centered tetragonal (b.c.t.) A_a, depending on the *c*/*a* ratio of the unit-cell parameters. In this scheme, some configurational isotypes are regrouped with different Strukturbericht symbols, as in the case of $CaCl_2$ (*C*35) and marcasite FeS_2 (*C*18), the orthorhombic form of pyrite. A further complexity is that revised structures can be found with new entries as in the case of boron nitride, *B*12, and B_k.

Table 1 Strukturbericht symbols

Strukturbericht types	*Compounds*
A	Elements
B	AB
C	AB_2
D	A_mB_n
D0, D1, D5…	AB_3, AB_4, A_2B_3…
E, F, G, H, K	$A_mB_nC_o$…
L	Alloys
O	Organic solids
S	Silicates

While the Strukturbericht classification can in principle accommodate any kind of alloy (Cu_3Au, $L1_2$), mineral (olivine, $S1_2$), or molecular (NH_3, *D*1) structure, there is inherent limitations. In particular, although the designations are convenient for simple phases, they become increasingly confusing and arbitrary as structures become more complex. For this reason, the Strukturbericht designation was not maintained in editions published after 1940 (then known as Structure Reports) and to a large extent this nomenclature has fallen into disuse except for common structures.

Pearson Symbols

Fourteen different three-dimensional lattices, the Bravais lattices, are possible: *aP*, *mP*, *mS*, *oP*, *oS*, *oI*, *oF*, *tP*, *tI*, *hP*, *hR*, *cP*, *cI*, *cF*. In this nomenclature, the first lower case letter indicates the crystal family *a*, *m*, *o*, *t*, *h*, *c* (anorthic, monoclinic, orthorhombic,

Table 2 Comparison of nomenclatures for some common materials

Compound	*Strukturbericht*	*Pearson*	*Lattice complex*
NaCl (halite)	B1	cF8	$F+\frac{1}{2}\frac{1}{2}\frac{1}{2}F$
PbS (galena)	B1	cF8	$F+\frac{1}{2}\frac{1}{2}\frac{1}{2}F$
C (diamond)	A4	cF8	D
ZnS (sphalerite)	B3	cF8	$F+\frac{1}{4}\frac{1}{4}\frac{1}{4}F$
$SrTiO_3$ (perovskite)	$E2_1$	cP5	$P+\frac{1}{2}\frac{1}{2}\frac{1}{2}P+J$
$MgAl_2O_4$ (spinel)	$H1_1$	cF56	$D+\frac{1}{2}\frac{1}{2}\frac{1}{2}T+\frac{3}{4}\frac{3}{4}\frac{3}{4}F_{222}xxx$

tetragonal, hexagonal, and cubic) while the second upper case letter gives the Bravais lattice *P*, *S*, *I*, *F*, *R*. The assignments m*S* and o*S* stand for any setting of a face-centered Bravais lattice in the monoclinic and orthorhombic families respectively, independent of the unit-cell setting adopted. So, to avoid confusion, the symbol *S* designating the "side-face centered" type is used rather than *C*.

Pearson symbols unite crystal lattice symmetry with the total number of atoms in the unit-cell by combining the two-letter code of the Bravais lattice and the total number of atoms in the cell. Thus, NaCl crystallizes in a face-centered cubic unit-cell *cF* with four Na cations and four Cl anions per cell, yielding its Pearson symbol *cF*8. The Pearson classification is concerned with Bravais translations but not with symmetry operations that can be applied on the unit-cell content. For example, non-isopointal structures may possess the same Pearson symbol, as observed in the case of diamond (*A*4), NaCl (*B*1) and zinc blende ZnS (*B*3) (**Table 2**). To address such limitations, ASTM proposed a nomenclature based on the Bravais centering and the number of atoms in the unit-cell, but added a lower case letter to distinguish between different structure types: diamond – 8*aF*, NaCl – 8*bF*, ZnS – 8*cF*. A suggestion from IUPAC was to add the name of a prototype to the Pearson symbol (e.g., *cF*8, NaCl).

Pearson symbols are widely used by scientists studying intermetallic alloys and are recorded in comprehensive treatises such as "Pearson's Handbook of Crystallographic Data for Intermetallic Phases" or the "Gmelin Handbook of Inorganic and Organometallic Chemistry." Perhaps more practically, Pearson symbols are used as search codes in structure databases.

Lattice Complex Notation (Bauverband)

A lattice complex is defined as a set of all point configurations that may be generated within one type of Wyckoff sets. The same lattice complex may belong to different space groups and are designated by the symbol of the space group of the highest symmetry, followed by a Wyckoff letter. In the simplest cases, invariant lattice complexes are formed from Wyckoff sets containing fixed crystallographic positions and are represented by an upper case letter named after the Bravais lattice (*P*, *C*, *I*, *F*), or representative of the lattice of common structural types. For example, the structure of diamond corresponds to the invariant lattice complex *D* with characteristic Wyckoff position Fd-3m a, which defined the eight symmetrically equivalent points generated by the position 000 (8a) in space group Fd-3m. It follows that lattice complexes can be used for the systematic enumeration of structural elements in different crystal systems. Indeed, the combination of all 16 invariant cubic lattice complexes allows all possible invariant structures to be defined.

A nomenclature has been derived from this approach in which a structure type is represented by listing the lattice complex and the relative shifts between different lattice complexes, expressed by a vector in fractional coordinates. For example, NaCl can be represented by the combination of two lattice complexes F, shifted along the diagonal of the unit cell, as $F+\frac{1}{2}\frac{1}{2}\frac{1}{2}F$, while ZnS is described as $F+\frac{1}{4}\frac{1}{4}\frac{1}{4}F$. The lattice complex F appears in second order, as F_{222}, when the complex F can be found in a subcell with volume an eighth of the basic unit-cell.

Although precise, the description becomes cumbersome if the geometrical configuration of lattice complexes is variable in ways that lead to different coordination numbers and geometry of coordination polyhedra. In these cases, the distances between nodes of a complex depend on fractional parameters *x*,*y*,*z* and the relative values of the cell parameters a:b:c and α:β:γ for crystal systems with less than cubic symmetry. Consequently, two structures with the same combination of lattice complexes may not necessarily be configurationally isotypic. Those lattice complexes with degrees of freedom are described by adding to the invariant complex symbol the splitting number of the fixed crystallographic position and the coordinate triplet of the characteristic Wyckoff positions. For example, the univariant lattice complex P8xxx splits the invariant complex P (where x = 0) into eightfold position xxx along the diagonals of a cube.

To address the examination of more intricate structures, Bauverband terminology extends lattice

complex notation with symbolism representing the nature of coordination polyhedra around atoms, the type of linkedness (corner, edge, face sharing) and their mutual orientation by selected symmetry operations, such that configurational isotypes might be more easily recognized. In particular, different Bauverband symbols can be assigned to the same structure.

Topological Descriptions

A consideration of topology (in this instance referring to descriptions of connectedness) offers a different perspective on structure types and their classification. In this case, structures are represented by nets (infinite graphs), where each vertex corresponds to an atom and each edge corresponds to a representative bond. Individual atoms can be advantageously replaced by groups of atoms as vertices, to yield simplified descriptions. Thus, structures composed of corner-connected polyhedra are usually described as nets where the central atoms are the vertices, and the edges connect the vertices corresponding to the same coordination spheres. This approach can be extended to larger groups of atoms and is well suited for streamlining the description of complex structures as illustrated in the case of faujasite, an alumino-silicate framework $T_{192}O_{384}$ (T = Si or Al), that can be described as a diamond net of sodalite cages composed of 24 T atoms (**Figure 3**).

Topological isotypism can be recognized irrespective of differences in chemical composition or symmetry, so diamond (Struckturbericht A4, lattice complex D) is topologically isotypic to ZnS (Struckturbericht B3, lattice complex $F + \frac{1}{4}\frac{1}{4}\frac{1}{4}F$).

Often, nets are named after simple structures, such as diamond net, or the well-known Kagome net. Unnamed nets can be designated using their Schlafli symbol N^n that specifies the number n of N-gons around each vertex (e.g., diamond 6^6). However, there is no systematic nomenclature for solids, except for the family of zeolites and other microporous solids, which are identified by a three-letter code.

Prototype Assignments

Amongst mineralogists and solid-state chemists, crystal structures are often referred to by the name, or the chemical composition, of a representative solid known as the prototype or aristotype, a structural representative in the space group of highest symmetry. For example, NaCl (rocksalt, halite) is the prototype of a very common structure type, representative of more than 20 distinct chemical compositions. This approach, while embracing the concept of crystal-chemical isotypism, is broadened by excluding the requirement for similar physiochemical properties. Given this latitude, it is perhaps surprising that a small number of common prototypes of similar topology, account for several thousand common, and technologically significant structures. Although prototypes are generally simple structures of high symmetry and the more common are shown in **Table 3**, heterotypes derived from these may be chemically diverse and often have lower symmetry.

A well-known example is the perovskite-type in which the prototype is $SrTiO_3$. This cubic structure consists of corner-connected TiO_6 octahedra that create cavities containing strontium. Crucially, it is the capacity of these octahedra to twist and distort that allows perovskites to accommodate a wide range of compositions – in excess of 800 distinct

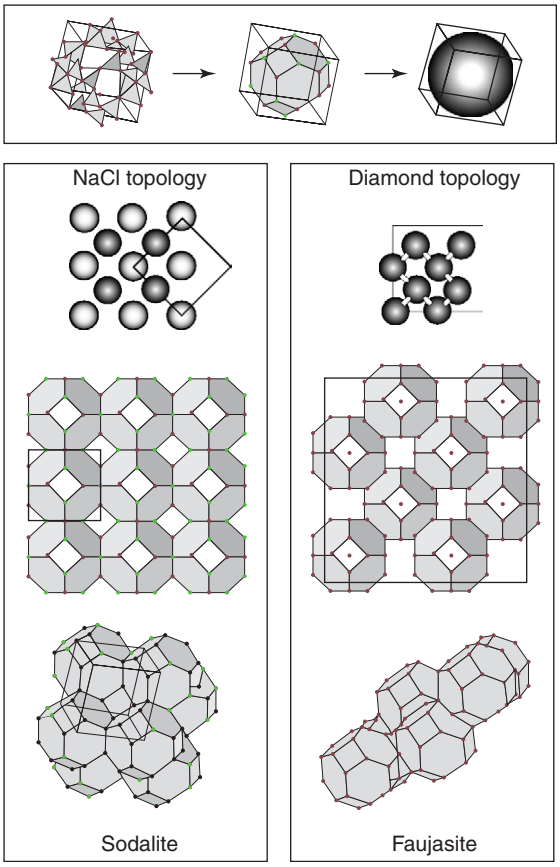


Figure 3 In complex structures, such as zeolites, relationships can be simplified by highlighting particular topologies, in which the nodes are not single atoms, as in the case of ZnS (see **Figure 2**) but groups of atoms. Uppermost is shown the progressive simplification of an atomic cluster in sodalite. Far left the AlO_4 and SiO_4 tetrahedra are highlighted in a representation that is difficult to visualize. This can be clarified by converting each tetrahedron to a single node. It is then evident that the arrangement is that of a truncated octahedron. Finally, this entire polyhedron can be regarded as a packing "sphere." In this way, it then becomes obvious that these units are packed together in NaCl topology in the case of sodalite, or in diamond topology in the zeolite faujasite.

Table 3 Frequently occurring structural prototypes

Prototype		*Approximate number of distinct species*
Name	*Compound*	
Perovskite	$SrTiO_3$	>800
Spinel	Mg_2AlO_4	>800
Apatite	$Ca_5(PO_4)_3(OH)$	>150
Olivine	Mg_2SiO_4	>400
Rocksalt	NaCl	>20

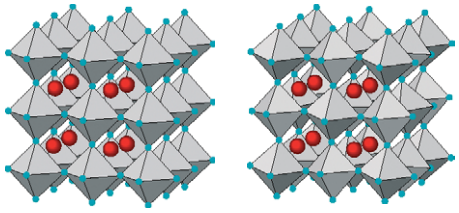


Figure 4 The *Pm-3m* perovskite prototype (or aristotype) is shown left with the TiO_6 octahedra shaded and Sr inserted large cuboctahedral cavities. $SrTiO_3$ adopts the prototype structure precisely, but many compounds belong to the perovskite family. The capacity to accommodate a wide range of chemistries arises because the octahedra are connected through their vertices in a manner that allows tilting and distortion to accommodate atoms of different sizes. The mineral perovskite $CaTiO_3$ has tilted octahedra and adopts *Pnma* symmetry.

species are known – that can display properties ranging from ion conductivity to superconductivity to semiconductor characteristics amongst others (**Figure 4**). It is interesting to note that the structure of the mineral perovskite $CaTiO_3$ is an orthorhombic derivative of the prototype.

Another useful feature of prototype classification is that an even greater variety of structures may be described as intergrowths of two prototypes. For example, the olivine-type Mg_2SiO_4 is important in geochemistry and as a refractory material. However, a different structural family – the humite family – is created by intergrowth of slabs of olivine with the prototype brucite $Mg(OH)_2$ (**Figure 5**). Thus, in addition to relating different structures, this approach may also be extended to describe, and in some cases predict, the existence of new compounds.

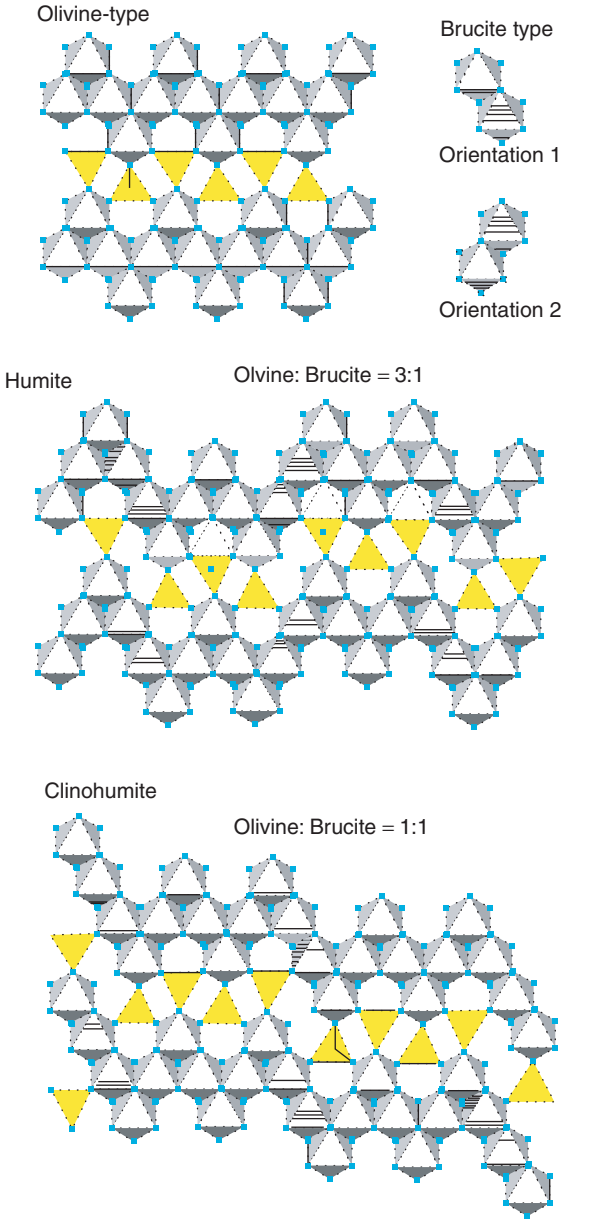


Figure 5 Structure types can be "intergrown" to build new and more complex structures. The humite structural series is derived from the olivine (Mg_2SiO_4) and brucite ($Mg(OH)_2$) prototypes through variation of the olivine:brucite ratio. In nature, four members of this structural family are known – norbergite (olivine : brucite = 1 : 1), chondrodite (2 : 1), humite (3 : 1), and clinohumite (4 : 1) – but many complex intergrowths have been observed in synthetic analogs.

Applying Structural Classification

Several different ways to describe crystallographic structures have been summarized here, but many others are known – Laves notation, Povarennykh's mineral systematization and so on. While each differs in its precision and breadth of coverage, no classification scheme is more or less correct, rather it is a question of appropriateness for a particular application, and the insights that may be drawn from their use. It is not surprising that schemes devised for comparing simple ionic solids are inadequate for complex covalent compounds, or conversely that the simplifications made for complicated systems fail to satisfy the need for exactitude in austere structures.

See also: Molecular Crystallography; Periodicity and Lattices; Point Groups; Space Groups.

PACS: 61.66.Bi; 61.66.Fn; 91.60.Ed

Further Reading

Baerlocher Ch, Meier WM, and Olson DH (eds.) (2001) *Atlas of Zeolite Structure Types*, 5th edn. (revised) Amsterdam: Elsevier.

Calvert LD (1985) *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, vols. 1–3, Metals Park, OH: American Society for Metals.

Hahn T (ed.) (2002) International tables for crystallography, Volume A. In: *Space-Group Symmetry*, 5th edn. Dordrecht: Kluwer.

http://cst-www.nrl.navy.mil/lattice/

Hyde BG and Andersson S (1989) *Inorganic Crystal Structures*. New York: Wiley.

O'Keeffe M and Hyde BG (1996) *Crystal Structures: i patterns and symmetry*. Washington, DC: Mineralogical Society of America.

Pearson WB (1967) *A Handbook of Lattice Spacings and Structures of Metals and Alloys*, vol. 2, New York: Pergamon.

Structure Reports (1940–93), Volume 8–58. Dordrecht: Kluwer.

Strukturberichte (1913–39), Volumes 1–7, Akademische Verlagsgesellschaft, Leipzig.

Wells AF (1984) *Structural Inorganic Chemistry*. Oxford: Clarendon.

Wyckoff RWG (1963–64) *Crystal Structures*, vol. 1–2. Malabar, FL: Wiley.

Sum Rules and Kramers–Kronig Relations in Nonlinear Optics

F Bassani, Scuola Normale Superiore, Pisa, Italy

General Description of Nonlinear Optical Functions

Nonlinear optics includes a number of phenomena, which are of great interest and have come to be of very common use with the availability of intense laser radiation. Among them are multiple-frequency generation, dynamic Stark effect, two-photon absorption, stimulated Raman absorption, and electromagnetically induced transparency. All these phenomena must obey the causality principle and consequently the Kramers–Kronig (KK) relations, and optical sum rules may be obtained for nonlinear optical phenomena in a way totally analogous to those used in linear optics.

One may start with the definition of the nth order polarization of a medium, expressed in terms of the external electric fields as

$$P^{(n)}(t) = \int_{-\infty}^{+\infty} \mathrm{d}t_1 \cdots \mathrm{d}t_n G^{(n)}(t_1, \ldots, t_n) \times E_1(t - t_1) \cdots E_n(t - t_n) \quad [1]$$

where the response function is symmetric under an exchange of variables, and satisfies the time causality condition

$$G^{(n)}(t_1, t_2, \ldots, t_n) = 0 \quad \text{if any } t_i < 0 \quad [2]$$

The electric fields are taken to be sinusoidal, with given frequencies, of the form

$$\boldsymbol{E}_n(\tau) = E_n(\omega_n)\exp[-\mathrm{i}\omega_n\tau] + E_n^*(\omega_n)\exp[\mathrm{i}\omega_n\tau] \quad [3]$$

where, by convention, the first term gives contributions to the susceptibility of argument ω_n and the second of argument $-\omega_n$.

Taking the Fourier transform of the polarization, the expression for the frequency component of the polarizability at any order is obtained as

$$P^{(n)}(\omega) = \chi^{(n)}(\omega_1, \ldots, \omega_n) E_1(\omega_1) \cdots E_n(\omega_n) \times \delta[\omega - (\omega_1 + \cdots + \omega_n)] \quad [4]$$

where $\chi^{(n)}(\omega_1, \ldots, \omega_n)$ is defined as the Fourier transform of $G^{(n)}(t_1, \ldots, t_n)$

$$\chi^{(n)}(\omega_1, \ldots, \omega_n) = \int_0^{\infty} \mathrm{d}t_1 \cdots \int_0^{\infty} \mathrm{d}t_n \exp[\mathrm{i}(\omega_1 t_1 + \cdots + \omega_n t_n)] \times G^{(n)}(t_1, \ldots, t_n) \quad [5]$$

and the symmetry in time implies that it is symmetric for the exchange of frequencies. The reality of $G_n^{(n)}(t_1, \ldots t_n)$, also implies

$$\chi^{(n)}(-\omega_1, \ldots, -\omega_n) = \chi^{(n)*}(\omega_1, \ldots, \omega_n) \quad [6]$$

If the fields are not sinusoidal, one can expand them in their Fourier components of type [3] and the sum over the frequencies $\omega_1, \ldots, \omega_n$ must be considered, with the appropriate field amplitudes.

One may observe that in the nonlinear case, the polarization is obtained by multiplying the nth order susceptibility by all the field amplitudes, so that its dimensionality depends on the order considered. The susceptibilities $\chi^{(n)}(\omega_1, \ldots, \omega_n)$ contain all the optical properties of the system to all orders, in the presence of a number of radiation beams.

The various terms of the susceptibilities to any order correspond to particular phenomena, and specific

expressions can be obtained in each case, but, since one is interested in the general properties, only the asymptotic behavior for $\omega \to \infty$ or $t \to 0^+$ is of interest. A simple way to obtain it in the classical model is to extend the Lorentz oscillator scheme described in the linear case by adding to the harmonic contribution a potential with all the anharmonic terms, and solving the resulting equations of motion by iteration. A more precise quantum-mechanical expression can be used, which is an extension of the expression given for the linear case

$$G^{(n)}(t_1, \ldots, t_n) = \frac{1}{n!(\mathrm{i}\hbar)^n} \sum T(1, \ldots, n) \times \{\mathrm{Tr}[P(-t_1), \ldots, [P(-t_n), \varrho_0]]P\} \quad [7]$$

where the sum is from 1 to n and T denotes the time ordering operator, and P is the dipole moment operator whose time dependence is expressed in the interaction representation in terms of the Hamiltonian H_0.

Since dissipation is proportional to the time average

$$-\left\langle \frac{\mathrm{d}P(t)}{\mathrm{d}t} E(t) \right\rangle \quad [8]$$

it vanishes when the fields and the polarization oscillate at different frequencies.

At different orders, one has different effects by considering all possible frequencies of the external electromagnetic fields. It is convenient to indicate the frequency of $P(\omega)$ and those of the external fields explicitly in the susceptibilities.

In the second order, an intense beam results in the second-harmonic generation from $\chi^{(2)}(2\omega_1; \omega_1, \omega_1)$ and optical rectification from $\chi^{(2)}(0; \omega_1, -\omega_1)$; at higher orders one has higher harmonic generation $\chi^{(n)}(n\omega_1; \omega_1, \ldots, \omega_1)$. With different beams, in the second order one has the sum and difference frequency generation $\chi^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$ and $\chi^{(2)}(\omega_1 - \omega_2, \omega_1, -\omega_2)$, and at higher order, all possible sums and differences. These do not give any absorption because the polarization frequency and the field frequency are different; the real and imaginary parts define in this case the phases of the susceptibility amplitudes.

In the third order and at higher odd orders, an important contribution to the susceptibility corresponds to the so-called pump and probe experiments, where one detects the dispersion and the absorption of a weak probe beam of frequency ω, in the presence of an intense beam of frequency ω_2. The nonlinear susceptibility $\chi^{(2n+1)}$ is proportional to the intensity of the pump beam at the nth power. The susceptibility $\chi^{(3)}(\omega; \omega, \omega_2, -\omega_2)$ and more generally $\chi^{(2n+1)}(\omega; \omega, \omega_2, -\omega_2, \ldots)$ contains a dynamical Stark effect when ω is close to a first-order resonance, a two-photon absorption when $\omega + \omega_2$ is close to a possible transition frequency, the stimulated Raman effect when $\omega - \omega_2$ is close to a transition frequency. In the pump and probe case, with a three-level system, one also obtains, under particular conditions, a coherent interference effect which produces an electromagnetically induced transparency (EIT) inside the first-order absorption line.

To consider the general properties related to pump and probe nonlinear dispersive–dissipative effects, the dielectric function which is the sum of the linear contribution and all odd nonlinear contributions with final frequency ω of the probe field may be studied

$$\varepsilon_{\mathrm{tot}}(\omega; \omega_2, E_2(\omega_2)) = \varepsilon^{\mathrm{L}}(\omega) + \varepsilon^{\mathrm{NL}}(\omega; \omega_2, E(\omega_2)) = 1 + 4\pi \sum_{n \text{ odd}} a_n \chi^{(n)} \times (\omega; \omega, \omega_2, -\omega_2, \ldots, \omega_2, -\omega_2) E^{n-1}(\omega_2) \quad [9]$$

where

$$a_n = \frac{1}{2^{n-1}} \frac{(n-1)!}{((n-1)/2)!((n-1)/2)!} \quad [10]$$

The nonlinear contribution $\varepsilon^{\mathrm{NL}}$ is given by the sum of the terms with $n > 1$, and the numerical factor is given by the invariance of the susceptibility for permutation of the frequencies. The most important nonlinear term is clearly $\chi^{(3)}(\omega; \omega, \omega_2, -\omega_2)$.

Kramers–Kronig Relations and Sum Rules

The analytic properties of the nonlinear optical susceptibilities can be examined in increasing order, and a number of sum rules can be obtained easily, when KK relations hold, by extending the procedure given for the linear case.

A first property follows immediately from causality due to the Titchmarsh theorem: the susceptibility is holomorphic in the upper half of the complex plane associated with one of the frequency variables when all the others are kept fixed. A second more general theorem due to time causality was derived by Scandolo: the nonlinear susceptibility is holomorphic in the upper half-plane of any straight line with a positive slope in the frequency space of all the ω_i. This also holds by taking the line with zero slope along a given axis, provided the other values can be kept fixed.

From the above results, it can be seen immediately that KK relations can be established in most cases

previously described, in second and higher harmonic generation along ω_1, as well as in pump and probe experiment along ω, keeping ω_2 fixed. They do not apply in optical rectification, because any straight line in the space $\omega, \omega_2 = -\omega$ cannot have positive slope; they do not apply also for $\chi^{(3)}(\omega_1; \omega_1, -\omega_1, \omega_1)$, when only one beam is present. The asymptotic behavior of the nonlinear susceptibilities is in general much faster than in the linear case, so that one may consider all integrable products of the type $\omega^k \chi^{(n)}(\omega)$, along the line of positive slope which is considered, and thus obtain a number of KK relations. From them, (using the superconvergence theorem as shown in the linear case) the asymptotic behaviors of the real and the imaginary part of the susceptibility can be obtained in terms of their moments (integrals of powers of the frequency times the susceptibility), up to a given value, above which the integral does not converge. A comparison of such asymptotic behaviors of the susceptibility with the behavior obtained from specific calculations gives a number of new sum rules for the nonlinear susceptibilities considered.

Harmonic Generation

As a first example, consider the second-harmonic generation. The asymptotic behavior of $\chi^{(2)}(2\omega, \omega, \omega)$ as $\omega \to \infty$ can be obtained easily from the anharmonic oscillator model or equivalently from the quantum theory. They give similar results, provided the appropriate third-order derivative $\partial^3 V(x)/\partial x^3$ for x-polarized radiation is substituted in the quantum theory by its expectation value on the ground-state wave function. The following second-harmonic asymptotic behavior obtains

$$\underset{\omega \to \infty}{\chi^{(2)}}(2\omega; \omega, \omega) = +\frac{1}{2!}\frac{e^3}{m^2}\varrho\left[\frac{\partial^3 V(x)}{\partial x^3}\right]_0 \times \frac{1}{4}\frac{1}{\omega^6} + o(\omega^{-6}) \qquad [11]$$

where ρ is the electron density and the potential term is given by the anharmonic coefficient in the extended Lorentz model, and by its expectation value on the ground state in the quantum-mechanical approach. In the presence of the polarized beams of the same frequency, the derivatives can refer to the polarization directions of the fields and to the polarization directions of $\boldsymbol{P}$, so that a third-order tensor component is defined. As a consequence of the asymptotic behavior [11], three independent KK relations are obtained, one for $\chi^{(2)}(2\omega; \omega, \omega)$, one for $\omega^2\chi^{(2)}$, and another one for $\omega^4\chi^{(2)}$. They can be summarized as follows in terms of their real and imaginary part χ_1 and χ_2:

$$\omega^{2\alpha}\chi_1^{(2)}(2\omega; \omega, \omega) = \frac{2}{\pi}\fint_0^\infty \frac{\omega'^{2\alpha+1}\chi_2^{(2)}(2\omega'; \omega', \omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'$$

$$\omega^{2\alpha-1}\chi_2^{(2)}(2\omega; \omega, \omega) = -\frac{2}{\pi}\fint_0^\infty \frac{\omega'^{2\alpha}\chi_1^{(2)}(2\omega'; \omega', \omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega' \qquad [12]$$

with $0 \leqslant \alpha \leqslant 2$ where the line on the integral denotes that the principal part has to be considered. By comparing the asymptotic behaviors obtained from the above KK relations, via the superconvergence theorem, which amounts to neglecting ω'^2 in the denominator of [12], with the requested one given above [11] six sum rules are obtained, which can be summarized as follows:

$$\int_0^\infty \omega^{2\alpha}\chi_1^{(2)}(2\omega; \omega, \omega)\mathrm{d}\omega = 0, \text{ for } \alpha = 0, 1, 2$$

$$\int_0^\infty \omega^{2\alpha+1}\chi_2^{(2)}(2\omega; \omega, \omega)\mathrm{d}\omega = 0, \text{ for } \alpha = 0, 1 \qquad [13]$$

$$\int_0^\infty \omega^5\chi_2^{(2)}(2\omega; \omega, \omega)\mathrm{d}\omega = -\frac{\pi}{2}\cdot\frac{e^3}{8}m^{-2}\varrho\left[\frac{\partial^3 V}{\partial x^3}\right]_0$$

The procedure given has been extended to higher harmonics and gives a large number of sum rules. All integrals of even moments of the real part of the susceptibility vanish up to the $2n$ moment, and all odd moments of the imaginary part of the susceptibility vanish, with the exception of the $2n+1$ moment which gives a result depending on the dynamics of the system.

Pump and Probe

Now, consider the pump and probe experiments, where the response to a beam of frequency ω is studied in the presence of an intense beam of frequency ω_2. The nonlinear contributions due to the probe beam alone are neglected, and the nonlinear susceptibility $\chi^{(3)}(\omega; \omega, \omega_2, -\omega_2)$, or more generally $\chi^{(2l+1)}(\omega; \omega, \omega_2, -\omega_2, \ldots, -\omega_2)$ for a very intense pump beam is considered. The asymptotic behavior can also be obtained in this case from the anharmonic oscillator model, or from the quantum-mechanical approach by using the Fourier transform [5] of the response function [7] for $t_1 \to 0^+$. One can show that the asymptotic behavior of the nonlinear susceptibility is

$$\tilde{\chi}^{\mathrm{NL}} = -\frac{C}{\omega^4} + o(\omega^{-4}) \qquad [14]$$

with

$$C = -\sum_{n\ \mathrm{odd}>1} \frac{a_n e^2}{n!(-\mathrm{i}\hbar)^{n-1} m^2} E_2^{n-1}(\omega_2) \int \mathrm{d}t_2 \cdots \int \mathrm{d}t_n \times \exp[-\mathrm{i}(\omega_2 t_2 - \omega_2 t_n)] \times \mathrm{Tr}\left\{ \frac{\partial^2 V}{\partial x^2} \sum_P T(2, \ldots, n)[P(-t_2) \cdot [\ldots P(-t_n), \varrho_0]] \right\} \quad [15]$$

and

$$a_n = \frac{1}{2^{n-1}} \frac{(n-1)!}{((n-1)/2)!((n-1)/2)!} \quad [16]$$

From this, it can be seen that two KK relations are obtained. One is the same as the linear KK relations and applies to $\chi^{\mathrm{NL}}(\omega; \omega_2,\ E_2)$. The other, typical of a nonlinear contribution for the probe–pump situation, applies to $\omega^2\ \chi^{\mathrm{NL}}(\omega; \omega_2,\ E_2)$. It is

$$\omega^2 \chi_1^{\mathrm{NL}}(\omega; \omega_2,\ E_2) = \frac{2}{\pi} \fint_0^\infty \mathrm{d}\omega' \frac{\omega'^3 \chi_2^{\mathrm{NL}}(\omega'; \omega_2, E_2)}{\omega'^2 - \omega^2}$$
$$\omega^2 \chi_2^{\mathrm{NL}}(\omega; \omega_2,\ E_2) = -\frac{2\omega}{\pi} \fint_0^\infty \mathrm{d}\omega' \frac{\omega'^2 \chi_1^{\mathrm{NL}}(\omega'; \omega_2, E_2)}{\omega'^2 - \omega^2} \quad [17]$$

These relations, appropriate to the nonlinear contribution, can be used to relate dispersive to dissipative nonlinear phenomena, as well as the original KK relation in the linear approximation.

By comparing the asymptotic expansion [14] and asymptotic expansions of the two KK relations obtained, one can derive four sum rules for the nonlinear susceptibility by following the usual procedure of imposing that their asymptotic behaviors coincide. They are the following:

$$\int_0^\infty \chi_1^{\mathrm{NL}}(\omega; \omega_2, E(\omega_2))\mathrm{d}\omega = 0$$
$$\int_0^\infty \omega \chi_2^{\mathrm{NL}}(\omega; \omega_2, E(\omega_2))\mathrm{d}\omega = 0$$
$$\int_0^\infty \omega^2 \chi_1^{\mathrm{NL}}(\omega; \omega_2, E(\omega_2))\mathrm{d}\omega = 0$$
$$\int_0^\infty \omega^3 \chi_2^{\mathrm{NL}}(\omega; \omega_2, E(\omega_2))\mathrm{d}\omega = C\frac{\pi}{2} \quad [18]$$

where C is given by [16]. The first two are analogous to the linear sum rules whose validity is thus extended to the optical response to all orders in the presence of a pump beam. The last two are typical of the nonlinear susceptibility. The last one, in particular, gives a measure of the strength of the nonlinear response through the value of the constant C.

Another simple sum rule gives immediately the nonlinear contribution of the pump field to the static susceptibility (limit $\omega \to 0$ in the KK dispersion relations)

$$\chi^{\mathrm{NL}}(0; \omega_2, E(\omega_2)) = \frac{2}{\pi} \int_0^\infty \mathrm{d}\omega' \frac{\chi_2^{\mathrm{NL}}(\omega'; \omega_2, E(\omega_2))}{\omega'} \quad [19]$$

Analogous sum rules can be obtained for all the nonlinear optical functions related to the susceptibility, as in the linear case.

It has been shown by Zimmermann that the sum rules given here are also valid when the expansion [1] in terms of the electric field does not converge, and the polarizability must be obtained in a non perturbative way, because the presence of the external field does not affect the short-time inertial character of the response function.

Examples and Applications

The Anharmonic Oscillator

One can display the relevance of the above-described sum rules by considering the simple, but physically significant anharmonic-oscillator case.

By solving the equation of motion with the probe frequency ω and the pump ω_2, the anharmonic contribution $\chi^{(3)}(\omega; \omega, \omega_2, -\omega_2)$ is obtained in the third order. The result is displayed in **Figure 1** as a function of the frequency of the probe beam, ω.

One can observe the modification of the absorption near the resonance ω_0 which is a dynamic Stark effect with a positive and a negative contribution; also a peak is present in correspondence to a two-photon absorption at $\omega_0 - \omega_2$ and another peak due to stimulated Raman absorption at $\omega = \omega_0 + \omega_2$.

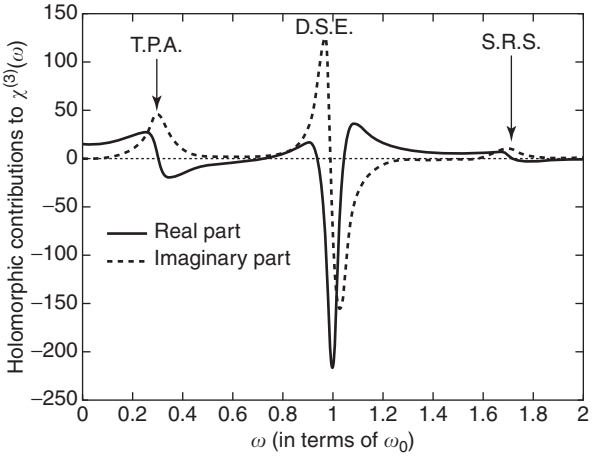


Figure 1 Nonlinear susceptibility of the anharmonic oscillator ($\omega_2 = 0.75\omega_0$.).

The sum rules [18] are all verified, but include contributions from all these effects and cannot be referred to a particular phenomenon on its own.

Analogous results are obtained from a quantum mechanical study of the oscillator, where however differences are found from the classical model in the nonlinear optical functions.

Three-Level System

The sum rules can be useful in the study of a system consisting of three levels, denoted with E_0, E_1, E_2, when the pump beam is nearly resonant with the transition between E_1 and E_2.

The two possible configurations in the three-level system are shown in **Figure 2**, and are called, respectively, the ladder and the lambda configurations. In both cases, the sum rules must be obeyed if no other levels are considered, and coherence effects must be included in studying the time dependence of the dipole moment of the system in terms of the density matrix ϱ_{ij}.

Using the Heisenberg equation for the time evolution of the density matrix and accounting for all the lifetime decay processes, one obtains in the lowest order

$$\chi^{(3)}(\omega,\omega_2,E_2) \simeq \varrho\left\{\frac{|\langle 1|ex|0\rangle|^2}{E_{10}-\hbar\omega-i\Gamma_1-\alpha_2^2/(E_{20}\pm\hbar\omega_2-\hbar\omega-i\Gamma_2)} + \frac{|\langle 1|ex|0\rangle|^2}{E_{10}+\hbar\omega+i\Gamma_1-\alpha_2^2/(E_{20}\pm\hbar\omega_2-\hbar\omega+i\Gamma_2)}\right\} \quad [20]$$

where the sign $+(-)$ refers to the ladder (lambda) configuration and $\alpha_2 = \langle 2|ex|1\rangle E_2/2$ is the Rabi frequency of the pump beam in energy units. The sum rules [18] can be immediately verified in both cases by observing that the asymptotic behavior of [20] is of order ω^{-4} as expected, and the constant C can be directly estimated.

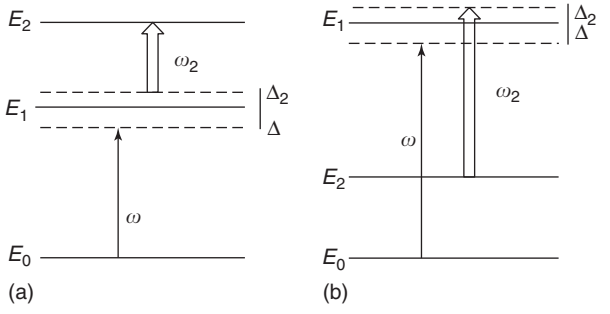


Figure 2 Three-level systems (a) in the ladder configuration or (b) lambda configuration, probe and pump beams are indicated by single and double arrows, respectively.

In the case of resonant ω_2, the typical results are given in **Figures 3** and **4** for large and small values of Γ_2, respectively. The nonlinear contributions become negative at resonance, and nearly cancel the absorption. Total transparency is obtained, when the lifetime $(\tau = (h/\Gamma))$ of level 2 is much longer than that of level 1, in particular when E_2 is a metastable state $(\Gamma_1 \gg \Gamma_2)$ and α_2 is sufficiently large. This phenomenon has been called coherent electromagnetically induced transparency (CEIT) (first observed by Alzetta *et al.* in 1976). It is of great interest because of its important applications. The CEIT effect has been so far observed in hot and cold atomic systems, but is now being studied also in solid-state materials. For

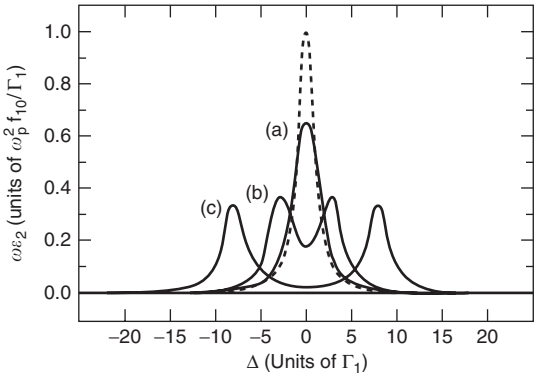


Figure 3 Probe absorption coefficient as a function of probe detuning for a resonant pump beam of increasing intensities: (a) $\alpha_2 = \Gamma_1$; (b) $\alpha_2 = 3\Gamma_1$; (c) $\alpha_2 = 8\Gamma_1$. Here, $\Gamma_2 = 2\Gamma_1$. The dotted line is the linear absorption. (From Scandolo S and Bassani F (1992) Nonlinear sum rules: the three level and the anharmonic oscillator models. *Physical Review B* 45: 5838.)

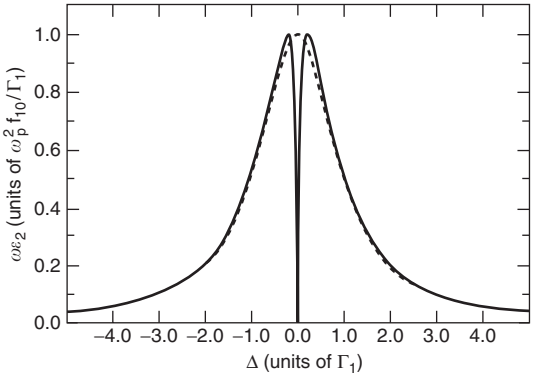


Figure 4 Probe absorption coefficient as a function of probe detuning for a resonant pump beam of moderate intensity $(\alpha_2 = 0.2\Gamma_1)$; here, $\Gamma_2 = 10^{-3}\Gamma_1$. The dotted line is the linear absorption. (From Scandolo S and Bassani F (1992) Nonlinear sum rules: the three level and the anharmonic oscillator models. *Physical Review B* 45: 5838.)

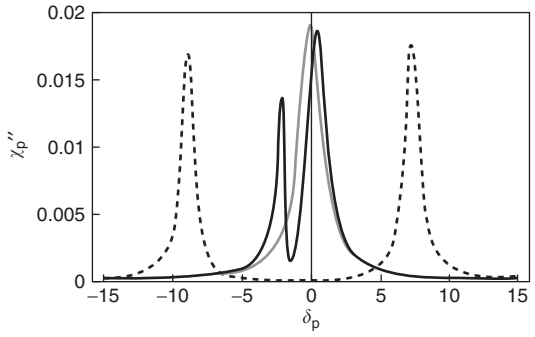


Figure 5 Imaginary part of the probe susceptibility as a function of probe detuning for slightly detuned pump beam of different intensities: $\alpha_2 = \Gamma_1$ (black solid line), $\alpha_2 = 8\Gamma_1$ (dashed line). Here, $\Gamma_2 = 0.1\Gamma_1$. The gray line is the linear absorption. (From Artoni M, Larocca GC and Bassani F (2000) Electromagnetic-induced transparency of Wannier–Mott excitons. *Europhysics Letters* (49): 445.)

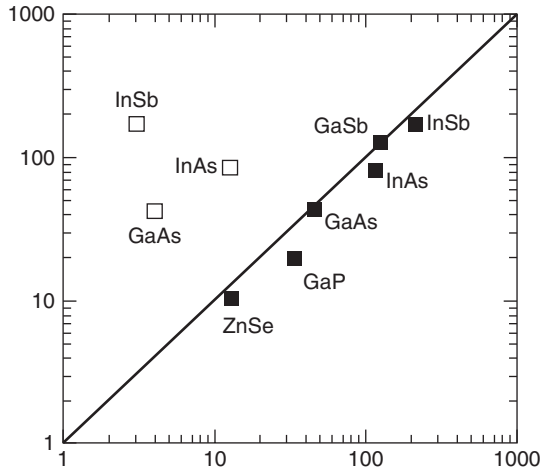


Figure 6 Second-order static susceptibility for some semiconductors. Abscissae: theoretical calculations in units 10^{-8} esu; ordinate; experimental data squares in units of 10^{-8} esu. (From Scandolo S and Bassani F (1995). Miller's rules and the static limit for second harmonic generation. *Physical Review B* 51: 6928.)

instance, **Figure 5** shows the computed absorption of a three-level system consisting of the ground state, the 1*s* metastable exciton, and 2*p* exciton in Cu_2O. These results can be seen as effects forced by the sum rules, which also have the effect of increasing the absorption away from the induced transparency in a way appropriate to verify them.

The EIT has important consequences in all phenomena related to propagation of light in dispersive media, because it produces hyperanomalous dispersion with a very large value of $dn/d\omega$. The most important effect observed is the reduction of the group velocity of light to extremely small values down to a few meters per second (slow light).

Second-Harmonic Generation

A further example of how the sum rules can be used to verify detailed theories, and to implement useful simplified models can be seen as an application to second-harmonic generation.

The quantum-mechanical calculation of $\chi^{(2)}(2\omega;\omega,\omega)$ is very complicated, because a knowledge of all excited states and a double summation involving all dipole matrix elements between them is required. A simple model is set up by considering an expansion about the main poles at ω_0 and $\omega_0/2$, ω_0 being the average frequency of the direct interband transitions in the semiconductors, which is also responsible for the static linear dielectric constant. Thus:

$$\chi^{(2)}(2\omega;\omega,\omega) = \frac{f_1}{\omega-\omega_0+i\varepsilon^+} + \frac{f_2}{(2\omega-\omega_0)+i\varepsilon^+} + \frac{f_3}{(\omega-\omega_0+i\varepsilon^+)^2} + \text{c.c}(\omega\rightarrow-\omega) \quad [21]$$

where ε^+ is a positive infinitesimal which can be set to zero. The coefficients f_i give the strength of the poles and substitute the detailed double sum of the dipole matrix elements. Imposing the sum rules, one can obtain expressions for all the coefficient f_i, in terms of the quantity C which appears in the 5th moment sum rule (see eqn [18]).

Taking the limit $\varepsilon^+ \rightarrow 0$, and expressing the first-order susceptibility in terms of the average transition frequency to the conduction band as

$$\chi = \frac{\varrho e_1^2}{m(\omega_0^2-\omega^2)} \quad [22]$$

one obtains Miller's empirical rule

$$\chi^{(2)}_{ijk}(2\omega;\omega,\omega) = \Delta_{ijk}\chi^{(1)}_{ii}(2\omega)\chi^{(1)}_{jj}(\omega)\chi^{(1)}_{kk}(\omega) \quad [23]$$

where Miller's empirical constant Δ is given by

$$\Delta_{ijk} = -\frac{1}{2e^3\varrho^2}\left\langle\frac{\partial^3 V}{\partial r_i \partial r_j \partial r_k}\right\rangle_0 \quad [24]$$

This can be easily computed for crystals from the pseudopotential theory, by means of the values of the pseudopotential parameters used in the energy band calculations. One can thus compute the static limit in a number of semiconductors and obtain values of $\chi^{(2)}(0,0)$ in good agreement with experiment, as shown in **Figure 6**.

A similar result can be obtained for third- and higher-order harmonic generation processes.

Data Analysis

The above described sum rules and KK relations can be also used to complete the experimental data available by imposing that they verify such general properties. This is particularly useful in nonlinear optics, where results on a limited frequency range are experimentally available, and only at a few fixed frequencies, both real and imaginary nonlinear susceptibilities are measured. For details on this, refer to the "Further reading" section.

Conclusions

It has been shown that the principle of time causality implies a number of KK relations between the real and the imaginary part of the nonlinear susceptibility and of all nonlinear optical constants. From them, a variety of sum rules can be obtained for the frequency moments of the real and the imaginary parts of the nonlinear susceptibilities. All frequency moments of the susceptibility vanish up to the highest integrable odd moment of the imaginary part of the susceptibility, which gives a value related to the density and to the expectation value of appropriate derivatives of the potential on the ground state.

The dispersion relations and the sum rules are shown to be useful to interpret experimental data and to verify the accuracy of the calculation models.

See also: Electromagnetically Induced Transparency; Elemental Semiconductors, Electronic States of; Interaction of Light and Matter; Nonlinear Optics; Optical Absorption and Reflectance; Optical Properties of Materials; Optical Sum Rules and Kramers–Kronig Relations; Organic Semiconductors; Semiconductor Optics.

PACS: 78; 78.20 – e; 78.40 – q; 11.55.Hx; 42.65 – k

Further Reading

Bassani F and Scandolo S (1991) Dispersion relation and sum rules in nonlinear optics. *Physical Review B* 44: 8446.

Blombergen N (1965) *Nonlinear Optics*. New York: Benjamin.

Boyd RW (1992) *Nonlinear Optics*. New York: Academic Press.

Lucarini V, Bassani F, Peiponen KE, and Saarinen JJ (2003) Dispersion theory and sum rules in linear and nonlinear optics. *La Rivista del Nuovo Cimento* 26(12): 1–120.

Lucarini V, Saarinen JJ, Peiponen KE, and Vartiainen EM (2005) *Kramers–Kronig Relations in Optical Materials Research*. Heidelberg: Springer.

Peiponen KE, Vartiainen EM, and Asakura T (1999) *Dispersion, Complex Analysis and Optical Spectroscopy*. Heidelberg: Springer.

Shen YR (1980) *The Principles of Nonlinear Optics*. New York: Wiley.

Superconductivity: Applications

D J Bishop, Bell Labs, Lucent Technologies, Murray Hill, NJ, USA

Introduction

In this section, what superconductivity is, why it is interesting from a scientific point of view, and why it has attracted the significant attention it has for applications will be discussed. This section shall lay the groundwork for the later detailed discussions of the applications of superconductivity.

The twentieth century was truly the century of the electron. At the beginning of the century, mankind had harnessed to a remarkable degree the power and benefits of machines of a great many kinds ranging from trains and cars to printing presses and steam engines. However, the electron was only beginning to be understood and used widely. However, by the end of the century we had computers, airplanes, cars, radios, TVs, EKG machines, and the Internet. None of these would have been possible without a century of extraordinary achievements in harnessing the use of the electron. However, in a century of remarkable discoveries, none is perhaps as revolutionary as the discovery of the effect called superconductivity.

One way to parametrize the nearly infinitely wide range of solid matter is by how well a given sample can carry electrical currents. At one end of the spectrum are metals such as copper, aluminum, gold, and silver that can carry electrical currents very well. These materials are conductors whose electrical resistivity is low. At the other end of the spectrum are materials that do not carry currents at all, called insulators. Examples of such insulating materials are glass, plastic, cloth, and paper. Typical electrical circuits of the kind found in homes for lighting consist of conductors – the wires (typically copper) and insulators (typically plastics) to keep the wires from making accidental connections. Despite the fact that many materials are good conductors, they all suffer from the disadvantage that the resistance is always nonzero. The resistance can be small but it is always finite for normal conductors. For example, when a house gets wired, the builder uses a large diameter wire to keep the losses small, but they are always

finite. In a house, these nonzero losses can be kept acceptably small. However, in an electric power station, these losses can be as much as 10% of the total power through the plant and represent a significant economic cost.

Before the discovery of the phenomenon of superconductivity, these losses found in normal conductors appeared to be inevitable. However, the discovery by Kamerlingh Ohnes in 1904 that below $\sim 4\,\mathrm{K}$ the resistivity of mercury drops to zero changed forever our view of electrical conduction in metals. This was one of the greatest scientific discoveries of the twentieth century, a century with many great discoveries. The discovery was the first and perhaps the finest example of condensed matter physics – the study of systems with many degrees of freedom.

The early superconductors, while scientifically very important, were impractical for use in any applied context. That is because they were what are now known as type I superconductors and these materials cannot carry any appreciable electrical currents or tolerate any sizable magnetic fields. At low temperatures type I superconductors exhibit what is known as the Meissner effect. In this remarkable effect, the magnetic field is completely expelled from the interior of the sample (not trapped, but completely expelled). However, this only works for relatively low fields (or small currents which is, in essence, the same thing) (see **Figure 1**). Typical fields that type I materials can tolerate are usually much less than 1000 G. This means type I materials are limited to use at low temperatures, low fields, and low currents which provide limited scope for applications.

The first type II superconductors were discovered by Shubnikov and co-workers in 1937. In these materials, there is also a Meissner state at low fields but there is also a new superconducting regime, a mixed state at higher fields. In this mixed state, the applied magnetic field enters the sample in the form of quantized lines of magnetic flux, magnetic vortices. This mixed state can survive to very high magnetic fields (and/or support large currents) and are the only practical superconductors. The discovery of the type II superconductors started the real search for applications for the phenomenon.

Figure 1 Magnetic levitation of a magnet about a YBCO sample.

Since the original discovery of superconductivity in mercury at 4 K, there have been literally thousands of other materials discovered with a wide range of transition temperatures (the temperature below which superconductivity occurs). However, the field of superconductivity was thrown into high gear in 1987 when a new class of materials was discovered by Bednorz and Muller, the so-called "High T_c Superconductors." These materials opened the way to the use of much less expensive cooling technologies such as liquid nitrogen ($\sim 77\,\mathrm{K}$) and closed-cycle refrigerators for cooling the materials. Much of the real work in finding applications for superconductivity was driven by the explosive growth in interest stimulated by the discovery of the high-T_c materials.

The basic phenomenology of superconductivity lends itself to application for three distinct reasons. The first is the characteristic that gives it its name, zero resistance or superconductivity. This allows conductors to be built that can carry very high current densities with, in principle, zero resistive losses and zero heating. The second phenomenon is that the materials superconduct because they are in a macroscopic quantum state that can be measured and exploited with such devices as a SQUID (described later). The third effect is that the transition into the superconducting state is what is called a phase transition or a change of state with a well-defined transition temperature such as the melting point of ice or the boiling point of water. This effect allows for applications in which the superconductor can be a very sensitive thermometer. All the applications of superconductivity discussed use one or more of these attributes to good advantage and none of these attributes can be found in conventional conductors such as copper or silver.

The final part of this introduction talks about a number of useful materials. While many thousands of superconductors have been discovered, only a few are used in practical applications. Useful materials not only need to have the characteristic of being superconductors, they also need many others as well. Additional needed characteristics include good metallic properties such as ductility, chemical stability, low thermal expansion, mechanical strength, low cost, ease of fabrication, nontoxicity, ease of joining, and others. When this long list of necessary characteristics is used to filter the many thousands of candidates, only a few are found to be practically useful. One of

the easiest and most useful materials is pure niobium. Thin films can be deposited and devices like SQUIDS and high-Q microwave cavities can be built. As a pure element, niobium is still very useful in a wide variety of applications. For wires, two alloys of niobium are widely used today. The first is NbTi and the second is Nb_3Sn. These materials are widely available commercially and still dominate the market despite 15 years of hard work on the high-T_c materials.

Among the newer, high-T_c materials, YBCO ($Y_1Ba_2Cu_3O_7$) and BSCCO ($(Bi,Pb)_2Sr_2Ca_2Cu_3O_x$ (Bi-2223)) are the materials of choice. They combine a reasonable combination of workable physical properties, extremely high critical magnetic fields (the maximum field at which they are still superconductors), and high critical currents (the maximum currents they can carry and still be superconducting). In general, the critical field is measured at very low currents and critical currents at very low applied fields and there is a tradeoff between the two quantities where finite fields lower the useful range of currents and finite currents lower the useful range of fields.

In the following sections, some of the nearly unlimited range of applications for these and other materials are discussed.

Power Generation/Transmission/Storage Applications

In this section, one of the most important application areas for superconductivity – electrical power generation and transmission – is discussed. The electrical power industry is one of the largest and most important industries in the US today. Indeed, without this industry, none of the others could exist. Currently, electrical power worth tens of billions of dollars is generated, stored, transmitted, and turned into useful work every day. The current system is the result of many years of work and is quite efficient, given the limitations of the real materials the engineers have at their disposal. The current system is roughly 90% efficient, a modern marvel. However, it is that last 10% which represents a vast economic payoff for superconducting technology.

Most power in the world is produced by large generators that are either hydroelectric (driven by the force of moving water) or steam (produced by geothermal, oil, coal, natural gas, or nuclear sources of heat). Generators consist of a rotor and a stator. Both have coils of wires. The stator is used to create a magnetic field that the rotor moves in. The rotor wires, when moved in the stator's field, have currents induced in them which is the current produced by the generator. Superconducting wires, because of their low losses in carrying electrical currents, can allow the generator to be smaller, lighter, and more electrically efficient than generators made using conventional conductors. Many such superconducting generators have been built and are nearing commercial deployment.

After generation, the electrical power must be transmitted to where it will be used. Most electrical power is sent around via the electrical power grid. This grid allows the power to be used a long way from where it is generated. In addition, the grid allows a more efficient process of supply and demand as peak demand periods move from east to west with the time zones. Also, the grid acts as a kind of buffer in case of scheduled or unscheduled outages in a particular generating station. The challenge of operating a power plant is that while the plant needs to be sized to handle peak loads and its cost is related to that peak capacity, the return to the investors is proportional to the average use of the plant. Therefore, a plant whose peak demands are much higher than the average demand is uneconomical to operate. The grid allows a much more efficient use of the country's generating capacity because power can flow to where it is needed, when it is needed, and average out the peak demands. Despite its evident value, sending electrical power long distances is neither cheap nor easy. The current way this is done is with high-voltage transmission lines that reduce the I^2R losses by increasing the voltage and reducing the currents that need to flow to send a given amount of power through the line. The current state of the art is million volt transmission lines. This is an important opportunity for superconductivity.

By building transmission lines using superconductors with very low losses instead of conventional conductors with a finite amount of loss, the efficiency of electrical transmission could be enhanced significantly. This has been an area of significant effort in the area of superconducting research for decades. Many trials have been done and with the advent of the high-T_c materials, the technology appears close to being ready for commercial deployment.

Another application for superconducting technology is in large crossbar switches used to protect and route the large currents generated and transmitted in the power grid. Superconducting switches, with their intrinsically low losses, offer a very cost-effective way to do this. In addition to using superconductivity to generate, switch, and transmit power, it can also be used for storage. The issue discussed above about the need to average out demand peaks also argues for the value of an efficient way to store large amounts of power. Power could then be generated during off-peak periods for use when the demand peaks and by doing this, the size

(and cost) of the plant could be reduced. As shown by Ohnes in his pioneering work on superconductivity, currents induced in superconducting loops can flow for very long times (almost infinitely long). Thus, large coils could store electrical power for long periods of time. Such large storage facilities are increasingly being envisioned for times in the not too distant future when the costs of fossil fuels rise as shortages appear.

Two final applications of superconductivity in the power plant in the longer term are the use of magneto-hydro-dynamics (MHD) and in fusion reactors. MHD generates power by having ions flow past low loss conductors inducing large electric fields and hence currents. Low losses associated with superconductivity make MHD technology feasible to install and use. Also, as will be discussed later, there is much basic research in building fusion reactors where atoms such as hydrogen are joined as opposed to fission where heavy atoms such as uranium or plutonium are split apart. The fusion reaction is expected to allow an unlimited amount of clean, cheap power. Such plants are quite likely to use strong superconducting magnets to contain the reactions in much the same way they are used today for the research experiments.

To sum up, there are many places in the generation, transmission, switching, and storage of electrical power where the low electrical losses allowed by using superconducting technology will drive many applications. Early field trials and tests have been completed for many of these devices and full-scale deployment is not too far off.

Medical- and Health-Related Applications

Another massive industry in the US is health and medical. Close to a trillion dollars a year is being spent in the pursuit of long and healthy lives for all the planet's inhabitants. One of the most important tools used in hospitals today for diagnostic imaging is MRI or magnetic resonance imaging. This technique uses the fact that many atoms have a magnetic spin that can be resonantly excited with RF coils in a magnetic field. By applying both a uniform field and a field gradient, density maps of the various atoms can be made in a living organism. These can be used to see with remarkable clarity bones, blood vessels, organs, and other structures within a living human body. The physics of the technique is such that the higher the field the better the signal-to-noise of the image and the better the spatial resolution. This is a perfect application for superconducting wires which enable very high field magnets to be made.

Figure 2 An MRI machine in operation. (Siemens press picture. Reproduced from Siemens AG, Munich/Berlin.)

A typical magnet may generate fields 20 000 times the earth's field in a magnet the size of a small car. The patient is usually placed on a belt that moves them into and out of the field while the image is being taken (**Figure 2**). Another place where superconductors are often used in such systems is in the RF coils used to excite the relevant magnetic atoms. The lower electrical losses of such coils give better signal-to-noise in a typical MRI system. Today MRI is being used for everything from routine hospital diagnostics to the study of how the human brain functions – not so much what you are thinking but where you are thinking.

In MRI systems superconductors are being used to generate very large magnetic fields. A completely different and new application is to use them to sense very small magnetic fields. The device is called a SQUID or superconducting quantum interference device. A SQUID uses the unique properties of the superconducting state itself – it is a macroscopic quantum state that can be "phase shifted" by the application of a small magnetic field. A typical SQUID is a ring structure with a weak link and induced currents that flow in the ring are periodic with the applied field. Such devices can sense fields in the picogauss range. The medical application for such devices is to sense the magnetic fields associated with flowing blood, magneto-cardiography, and magnetic fields associated with flowing currents in the brain, magneto-encephalography. In both cases it is possible to image defects that other techniques cannot, providing useful diagnostic information. In the brain imaging case, it is possible to find the source of localized defects that can cause tremors and, when localized, excise them noninvasively using a gamma knife. Such systems are being used in field trials today and will soon be deployed widely.

Clearly the use of superconductivity in medicine is already a large and growing industry and given the graying of the US population, shows no signs of slowing down.

Military and Homeland Security Applications

The military applications of superconductivity are as varied and important as the commercial possibilities. These application areas range from large devices such as motors and generators to microscopic ones such as would be found in a sensor network. Clearly, the military applications are well appreciated and much DOD funding has gone into superconductivity research.

For ships and submarines there are many applications for large motors and generators. Aboard ships and especially submarines, size and weight are important considerations. The same features that make superconducting motors appealing in a commercial power station make them useful on a ship or sub. The fact that superconducting wire allows for higher current densities over conventional conductors makes it possible to build smaller, lighter, and higher performance motors and generators even when the cooling plant needed for superconducting motors is taken into account.

Another area attracting attention in the military world are superconducting rail guns. A rail gun uses very high currents in an applied magnetic field to impart large velocities to projectiles, usually much greater than could be obtained with conventional explosives. In fact, rail guns are envisioned as being able to put small objects into orbit, something that cannot presently be done with a conventional, single-stage explosive shell.

The ability to use superconductivity to store large amounts of power in a compact way may be applied to military laser systems. For many applications such as missile defense, very high-power lasers are needed. Generally, these are pulsed systems which require large amounts of power for short periods of time. For the same reasons as in a commercial power plant, superconductivity may be the technology of choice for this application.

A final application area discussed here are SQUID sensor networks. Described earlier, a SQUID can be configured to allow very small magnetic fields to be sensed. Devices and networks of such devices are currently being used to sense everything from submarines in the deep ocean to perimeter defense in a tactical situation.

High current densities and sensitive magnetic field detection are but two of the characteristics that make superconductivity an appealing technology that will be applied in military and homeland security areas.

Electronic Applications for Superconductivity

The world as we know it has been changed beyond all recognition by the relentless progress in electronics. Everything from computers to air travel would be impossible without microelectronics and silicon VLSI. This progress in electronics has been well described by Moore's law, the observation that transistors get a factor-of-two smaller, cheaper, and faster every 18 months. However, it is clear that this trend cannot go on for much more than another decade or so. The race is on for what can replace conventional silicon VLSI. For high-end computing applications where the cooling will not be a problem, superconducting electronics is being actively pursued.

In conventional electronics one of the things that makes it difficult to scale transistors into the deep submicron regime is the heat they generate as they switch at high speeds. As wires, gates, sources, and drains all get smaller, the current densities they must work at get larger and this heat limits both performance and lifetime. Superconducting electronics allows one to use both switches and interconnects that dissipate much less power than their conventional counterparts. They then allow a much faster, higher-performance operation for the chip. A number of high-end superconducting processing systems have been built and made commercially available and more can be expected in the next decade or so.

Another aspect of superconductivity and computing that go naturally together is the concept of quantum computing. In quantum computing, the q-bits (the analog to the classical bit) obey the rules of quantum mechanics, not classical mechanics. For certain classes of problems such as factoring large numbers, quantum computers are expected to be much faster than conventional ones. While the theory of such computers is progressing well, the experimental work has proceeded much more slowly. One of the experimental challenges is to find systems where the quantum coherence occurs over large enough distances and for long enough times for real circuits to be built and tested. One of the best candidates for such systems are networks of Josephson Junctions built from superconducting materials. While work in this area is at an early stage, it appears likely that if a practical quantum computer is ever built, it will use superconductors for its practical implementation.

In the area of microwave electronics, superconductors have found a number of unique application areas. For microwave systems filters play a crucial role allowing the entire available electromagnetic spectrum in a given band to be used. Filters using superconducting films have properties superior to conventional filters and are beginning to be deployed in both military and commercial RF systems. Superconductors are also being used as both microwave detectors/mixers and microwave sources. As detectors and mixers, they are currently state-of-the-art in terms of sensitivity and noise and are widely used in microwave astronomy applications. As sources, they are being used as frequency standards in high-precision applications.

The economic value of high-end electronic computing is large and it is clear that as we get to the end of the Moore's law curve, superconductivity will play an ever-increasing role.

Applications of Superconductivity in Transportation

Transportation is one of the major needs any civilized society has to provide for its citizens. However, it is clear that the current arrangement of personal cars powered by fossil fuels is destined to be replaced by more efficient mass transportation. One of the leading candidates for efficient, fast, and convenient mass transportation is magnetically levitated trains or MAGLEV trains. Such trains use superconducting magnets to create large magnetic fields to levitate the train so that wheel friction is absent making the trains faster, lighter, and much more power efficient. They literally glide on a layer of magnetic field like the suspended magnet shown in **Figure 1**. Such trains are being actively developed with a number of large-scale field trials having been completed. Work in this area in Japan is particularly advanced. It is clear that in the next decade or so superconducting MAGLEV trains will enter commercial service.

Other application areas for superconductivity in transportation include large motors and generators for ships, subs, and trains and storage devices for use in clean vehicles. In particular, when a train or car stops, it turns all of the kinetic energy generated by the motor into heat. An efficient storage device such as a superconducting storage ring would allow one to convert that kinetic energy into energy usable to start the car or train moving again. Significant fuel efficiencies would result from such a system.

Superconductivity will play an important role in providing safe, inexpensive, and fast transportation in the coming decades.

Scientific Applications of Superconductivity

The business of science is a place where superconducting technology is already being widely applied. In many places, state-of-the-art measurements are being done in ways that are only possible with superconducting devices. This is then the cutting edge of applying superconducting technology.

One of the largest-scale deployments of superconducting technology is for high-energy particle accelerators. These machines apply large magnetic fields at just the right time to accelerate charged particles to nearly the speed of light. Superconductors allow for much higher fields, lower costs, smaller sizes, and better performance than anything a conventional conductor could achieve. Devices such as the TEVATRON at Fermi Lab and RICK at Brookhaven could not have been built without superconductivity.

Another key area of scientific research is in MRI imaging. For example, in functional brain imaging, the large magnetic fields that can only be attained using superconductors allow brain scientists to gain fundamental insights into how the brain works. MRI is both a key research tool as well as a key diagnostic tool used in most hospitals in the US today. NMR of complex molecules is another important scientific field which would be impossible without superconductivity.

Another area of scientific research is the use of large research magnets built using superconducting wires. Fields as high as 45 T can be obtained allowing everything from work on the quantum Hall effect to superconductivity itself to be studied.

As mentioned earlier, fusion research uses high-field superconducting magnets to generate the magnetic fields needed to contain the hot plasmas created during the fusion process. At the other end of the field spectrum, devices such as SQUIDS are being used to do precision measurements on things ranging from the magnetization of new materials to gravity wave detectors. Superconducting devices are also being used to create a new class of particle detectors that work as sensitive bolometers where the particle, when absorbed by the detector, heats up the detector and this change in temperature is sensed by an edge detector which is a superconducting film biased to be just at the transition to normal conduction; Mose resistance, therefore, is very sensitive to small changes in temperature.

The art of scientific measurement relies heavily on superconducting technology and this provides a kind of testing ground for later, large-scale commercial applications of the technology. The extensive use of such technology in science makes it clear that large-scale commercial applications are on their way.

Miscellaneous Applications of Superconductivity

The numerous applications already discussed by no means exhausts the list of active areas of superconducting development. Superconducting technology is being used to create very sensitive gravity detectors for use in prospecting for oil and gas.

Magnetic separation is another area of considerable commercial importance. These devices allow the separation of materials with differing responses to a magnetic field such as aluminum and iron alloys. In practice, a stream of particles is shot through the magnet and their trajectory is determined by their response to the applied magnetic field. The potential commercial value of such devices is very large as we increasingly focus on recycling materials as opposed to creating new ones in our manufacturing processes.

Conclusions and the Future

It should be clear that the range of potential applications for superconducting technology is truly vast. This article, being limited in scope, has only begun to scratch the surface of what is being done and what will be done in terms of applications. In almost every area of human commercial activity, superconductivity will play an important role. Power generation and transmission, transportation, electronics, military systems, scientific research, and health and medicine are some important examples but by no means exhaust the list. While superconductivity was discovered in the twentieth century, the twenty-first century will see an amazing range of applications for the technology developed and deployed. It will be an exciting time for both scientists and engineers working in the field.

See also: Integrated Circuits; Magnetic Materials and Applications; Superconductivity: BCS Theory; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductivity: Tunneling; Superconductors, High T_c; Superconductors, Metallic; Superfluidity; Tunneling Devices.

PACS: **85.25.j; 74.xxx; 74.25.Ha; 74.25.Nf**

Further Reading

Batlogg B (1991) Physical properties of high-T_c superconductors. *Physics Today* 44: 44.

Clarke J (1994) SQUIDS. *Scientific American* 271(2): 46.

Kittel C (1996) *Introduction to Solid State Physics*, 7th edn. New York: Wiley.

Lubkin GB (1996) Power applications of high-temperature superconductors. *Physics Today* 49: 48.

Ohanian H (1989) *Physics, 2E Expanded*: WW Norton.

Superconductivity: BCS Theory

K H Bennemann, Freie Universität Berlin, Berlin, Germany

Introduction

For many years superconductivity has remained one of the most interesting phenomena in physics. After its discovery by Kamerlingh–Onnes in 1911, many superconductors have been found, see **Figure 1**. The recent discovery of MgB_2 as a superconductor with a transition temperature $T_c = 39\,K$ suggests that further surprises can be expected in the future. The basic interdependence of superconductivity and lattice structure still needs to be understood. **Figure 2** shows the transition temperature of some interesting metals.

Some of the cornerstones in the history of superconductivity were the observation of (1) a vanishing electrical DC resistivity $\rho(T)$ at T_c, (2) the magnetic behavior exhibiting type-I superconductors expelling magnetic flux (Meissner–Ochsenfeld), and type-II superconductors where quantized magnetic flux can penetrate in filaments, (3) the isotope effect, $T_c \propto M^{-\alpha}$, for example, $\alpha \approx 0.5$ for Hg, suggesting that the electron–phonon coupling might be responsible for superconductivity, and (4) Josephson tunneling with an electronic current $j = j_0 + j_1 \sin(\Delta\varphi - (2e/\hbar)\, V_{21} t)$, where $\Delta\varphi$ is the phase difference and V_{21} the voltage between two superconductors 1 and 2. There are many additional interesting observations including the occurrence of a gap in the electronic density of states (DOS), gapless superconductors and the interdependence of atomic structure and T_c, seen in particular in layered structures, granular superconductors, etc.

Most important theories developed for a physical understanding of the superconductivity phenomena were: (1) the London theory in 1935 explaining the Meissner–Ochsenfeld effect by using for the superconducting current driven by the vector potential $\boldsymbol{A}$ the expression $\boldsymbol{j}_s = -(c/4\pi\lambda_L^2)\boldsymbol{A}$, with the penetration

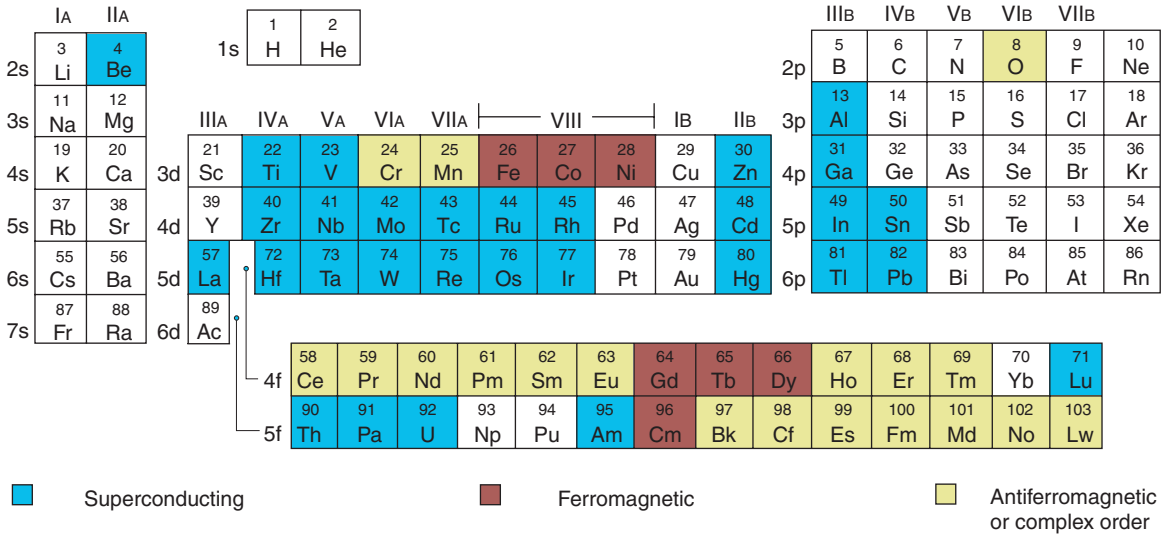


Figure 1 The occurrence of superconductivity in the periodic table is illustrated. As the history of superconductivity shows, alloys and compounds of the elements play an important role.

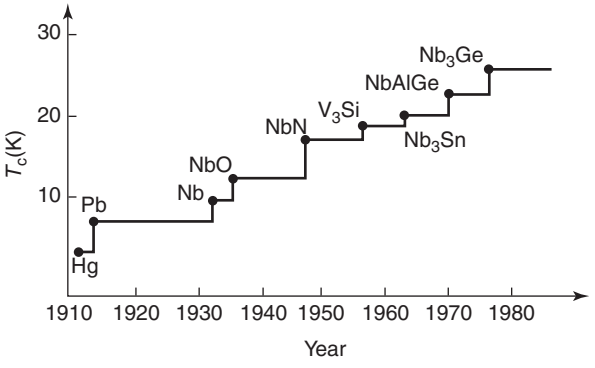


Figure 2 History of the transition temperature T_c for the first 70 years following the discovery of superconductivity in 1911. The A-15 compounds were of particular interest in the search for higher T_c superconductors.

depth $\lambda_L = (mc^2/4\pi e^2 n_s)^{1/2}$, where n_s is the superfluid density. (2) the Ginzburg–Landau (GL) theory in 1950, which extended the London theory by writing the free energy in terms of an order parameter $\psi(\boldsymbol{r}, t) = |\psi| e^{i\varphi(r)}$ as

$$F = F_0 + \int d^3r \left\{ a|\psi|^2 + \frac{1}{2} b|\psi|^4 + \frac{\hbar}{2m^*} \left| \left(\nabla - \frac{ie^*}{\hbar c} \boldsymbol{A} \right) \psi \right|^2 + \frac{h^2}{8\pi} \right\} \quad [1]$$

yielding the current via minimization with respect to ψ (or ψ^*):

$$\boldsymbol{j}(\boldsymbol{r}) = -\frac{ie^*}{2m^*} \left\{ (\psi^* \nabla \psi - \psi \nabla \psi^*) - \frac{e^{*2}}{m^* c} |\psi|^2 \boldsymbol{A}(\boldsymbol{r}) \right\} \quad [2]$$

(3) In 1956, the Bardeen–Cooper–Schriefer theory (BCS) deriving superconductivity from an electronic Hamiltonian $H = H_0 + H_{eph}$, where H_{eph} describes the electron–phonon interaction. It is to be noted that eqn [2] immediately yields the Meissner effect, by assuming that the superconducting wave function is rigid and is not affected by an electromagnetic field (thus, the first term in eqn [2] is zero), and for $\psi = |\psi| e^{i\varphi(r)}$ ($\psi\psi^* \sim n_s$ and φ are conjugate variables). The Josephson tunnel current between two superconductors 1,2 (with phase and voltage difference $\Delta\varphi$, ΔV) is given by

$$j_J = j_0 \sin\left(\Delta\varphi - \frac{e^*}{\hbar} \Delta V t \right) \quad [3]$$

Here,

$$\Delta\phi = \varphi_2 - \varphi_1 + \frac{2\pi}{\phi_0} \int_1^2 \boldsymbol{A} \cdot d\boldsymbol{l}$$

$$j_0 = \frac{e^* \hbar}{m^* d} |\psi|^2 \quad (m^* = 2m,\ e^* = 2e)$$

and d is a parameter referring to the tunnel junction, $\phi_0 = \hbar c/2e$. Assuming the single-valuedness of ψ, one obtains the quantization of flux trapped (in a filament) in a superconductor (see gauge transformation ψ):

$$\phi = \phi_0 \left(n + \tfrac{1}{2} \right), \quad n = 0, \pm 1, \ldots \quad [4]$$

These are remarkable general results which hold for different possible mechanisms causing condensation of a Fermi liquid into a superconducting state.

The BCS theory is one of the most elegant theories in physics and effectively explains superconductivity resulting from Cooper-pairs ($\boldsymbol{k}\uparrow$, $-\boldsymbol{k}\downarrow$) for electrons

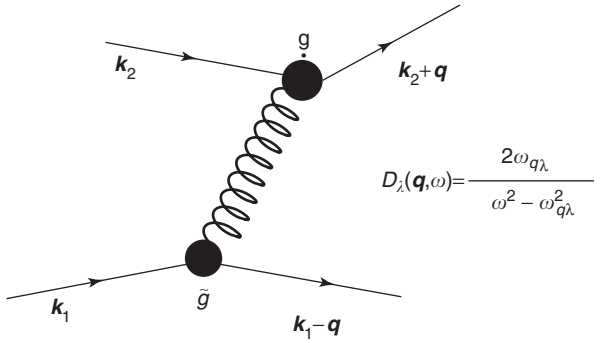


Figure 3 Illustration of the interaction between electrons with k_1 and k_2 via exchange of a phonon with polarization λ, wave vector q and energy $\omega_{q\lambda}; \omega \simeq \varepsilon_{k_1-q} - \varepsilon_{k_1}$; then an attractive effective electron–electron interaction $g^2 D < 0$ results, if $\omega < \omega_{q\lambda}$. $\tilde{g} = g\Gamma$, where Γ denotes many-body corrections to the electron–phonon coupling g ($\tilde{g} \simeq g$; see Migdal theorem).

with opposite spin and momentum k at the Fermi surface. As illustrated in **Figure 3**, the exchange of virtual phonons yields an attractive interaction and thus, a singlet Cooper pairing. This pairing causes a gap in the single-electron DOS, reflecting the binding energy of the Cooper pairs, $\Delta = \Delta(T, h, \ldots)$. The theory correctly predicts the behavior $\Delta(T = 0) \propto T_c$ and some other thermodynamical and electrodynamical quantities.

Following the BCS theory, the quantum field theory using the Green's function method, in particular by Gor'kov *et al.*, gave the theory of superconductivity a very elegant form. In general, for phase transitions (ferromagnetism, etc.), the superconducting state cannot be obtained from the normal state by using the perturbation theory. That the electron–phonon interaction described in second-quantization notation by

$$H_{\text{eph}} = \sum_{k,k',\lambda} g_{kk'\lambda} c_{k'}^{+} c_k + \cdots \qquad [5]$$

is responsible for superconductivity was suggested early by Fröhlich and others and explains, for example, the isotope effect. The electron–phonon coupling constant is given by

$$g_{kk'\lambda} = \sum_{i,\alpha} \left(\frac{\hbar N_\alpha}{2\omega_{q\lambda} M_\alpha} \right) \langle k' | \nabla_i U_{i\alpha} | k \rangle \times \mathscr{E}_{q\lambda}(\alpha) e^{iq\cdot\rho_\alpha^0} \qquad [6]$$

where $U_{i\alpha}$ is the potential of the ion i, α and $\mathscr{E}_{q\lambda}$ are the lattice polarization vectors. Since, as illustrated in **Figure 3**, emission and absorption of phonons by electrons at the Fermi surface can lead to an effective attractive interaction, it seemed possible that Cooper pairing ($k\uparrow$, $-k\downarrow$) occurs for electrons within an energy shell of $2\omega_D$, where ω_D is the Debye frequency. On general grounds, one expects, for an attractive potential with a range of the order of the interatomic distance ($a \sim p_F^{-1}$), due to the Pauli principle, singlet pairing and pairing of k and $-k$ electrons to avoid kinetic energy of the Cooper pairs and thus to obtain maximal condensation energy. To avoid the destructive repulsive electron–electron interaction, the pairing results from a retarded potential. The electron $k\uparrow$ polarizes the lattice for a time of the order of $t \sim \omega_D^{-1}$, which is slower than the electron motion, and then the other electron $-k\downarrow$, arriving at the polarized lattice within that time might feel an effective attractive electron–electron interaction due to over-screening. The resultant binding energy 2Δ causes a correlation and coherence length of the order of $\xi \sim v_F/2\Delta$. It is to be noted that Coulomb fields are screened at a distance of the order of the interatomic distance. This analysis and the study by Cooper, showing that for electrons at the Fermi surface the pair wave function $\psi = \sum_{k \sim k_F} a_k e^{ik\cdot\rho}$, $\rho = r_1 - r_2$, lowers the energy ($(H_0 + V)\psi = E\psi$ and $V < 0$), gave a more definite clue that the Cooper pairing of the electrons within an energy shell of the order of $2\hbar\omega_D$ causes superconductivity.

Cooper pairing is the essence of the electronic BCS theory, which gives the microscopic justification of the phenomenological Ginzburg–Landau theory. Its formulation was achieved with the help of quantum field theoretical methods.

BCS Theory

The BCS theory begins by using the Hamiltonian

$$H = \sum_{k,\sigma} \xi_k n_{k\sigma} + \sum_{k,k'} V_{kk'} a_{k\uparrow}^{+} a_{-k\downarrow}^{+} a_{-k'\downarrow} a_{k'\uparrow} + \cdots \qquad [7]$$

with excitation energies $\xi_k = \varepsilon_k - \mu$ for electrons near the Fermi energy (μ is the chemical potential) and the effective interaction $V_{kk'}$ between electrons with wave vector k and $-k$ and spin $\uparrow$ and $\downarrow$. Note that the remaining interactions between the electrons are assumed to be included in the energies $\xi_k . a_{k\sigma}^{+}$ and $a_{k\sigma}$ are the usual creation and annihilation operators for fermions with k and spin σ, and $n_{k\sigma}$ is the occupation operator. The BCS theory uses the wave function

$$|\psi\rangle = \prod_k (u_k + v_k a_{k\uparrow}^{+} a_{-k\downarrow}^{+}) |\psi_0\rangle, \quad |u_k|^2 + |v_k|^2 = 1 \qquad [8]$$

to calculate the free energy. The ground state is $|\psi_0\rangle$. $|v_k|^2$ gives the probability of a pair ($k\uparrow$, $-k\downarrow$). The diagonalization of H is most elegantly achieved by the

Bogoliubov and Valatin canonical transformations

$$b^+_{-k+} = u_k a_{-k\downarrow} + v_k a^+_{k\uparrow}, \quad b_{k-} = u_k a_{k\uparrow} - v_k a^+_{-k\downarrow}, \quad u_k^2 + v_k^2 = 1 \qquad [9]$$

which introduce new fermion operators $b_{k\sigma}$, $b^+_{k\sigma}$. Due to isotropy, the coefficients u_k, v_k depend only on k and can be chosen to be real and such that, for a given entropy the electronic energy E is minimal. Then one gets straightforwardly (with $V_{kk'} = -g/V$)

$$E = 2\sum_k \xi_k v_k^2 + \sum_k \xi_k (u_k^2 - v_k^2)(n_{k\uparrow} + n_{k\downarrow}) - \frac{g}{V}\left(\sum_k u_k v_k (1 - n_{k\uparrow} - n_{k\downarrow})\right)^2 \qquad [10]$$

and, from $\partial E/\partial u_k = 0$, the result

$$2\xi_k u_k v_k = \Delta_k (u_k^2 - v_k^2) \qquad [11]$$

Here, using $v_k^2 + u_k^2 = 1$ and eqn [9], the order parameter ($\Delta_k \propto \langle a_{k\uparrow} a_{k\downarrow} \rangle$)

$$\Delta_k = \frac{g}{V}\sum_k u_k v_k (1 - n_{k\uparrow} - n_{k\downarrow}) \qquad [12]$$

has been introduced. The order parameter can be expressed as

$$\Delta_k = \frac{g}{V}\sum_{k'} \frac{\Delta_{k'}}{2\mathscr{E}_{k'}} \tanh(\mathscr{E}_{k'}/2k_B T), \quad \mathscr{E}_k = \sqrt{\xi_k^2 + \Delta_k^2} \qquad [13]$$

It is of interest to note that eqn [13] has solutions not only for attractive interaction (with $g > 0$), but possibly also for repulsive ones with $g < 0$. Then the k dependence of Δ_k is important.

The new quasiparticle energies are $\mathscr{E}_k = (\mathrm{d}E/\mathrm{d}n_k)_{u_k, v_k} = \sqrt{\xi_k^2 + \Delta_k^2}$. Their distribution function is $n_k = (\mathrm{e}^{\mathscr{E}_k/kT} + 1)^{-1}$. These quasiparticles contribute to the energy the term $\sum_k \varepsilon_k n_{k\sigma}$. It is to be noted that $2\Delta_k$ is the binding energy of the correlated electrons $k\uparrow$, $-k\downarrow$. Thus, the coherence length $\xi = \hbar v_F/2\Delta$ gives the correlation distance, for most conventional superconductors, a much larger value than the interatomic distance ($\hbar/p_F$).

Note that for an isotopic gap $\Delta_k = \Delta$ one gets

$$1 = \frac{g}{V}\sum_k \frac{1}{2\sqrt{\xi_k^2 + \Delta^2}} \tanh\left(\frac{\mathscr{E}_k}{2k_B T}\right) \qquad [14]$$

and from this, for $\Delta(T)$, the expression

$$\Delta(T) = 3.2 k_B T_c \left(1 - \frac{T}{T_c}\right)^{1/2} \qquad [15]$$

The superconducting transition temperature T_c follows from $\Delta(T_c) = 0$. Thus, ($\sum_k \to \int_{-\omega_D}^{\omega_D} \mathrm{d}\varepsilon$, $\omega_D \propto M^{-1/2}$)

$$T_c \simeq 1.14\hbar\omega_D \mathrm{e}^{-1/(N(0)g)} \propto M^{-1/2}, \quad 2\Delta(T=0)/k_B T_c = 3.52 \qquad [16]$$

As shown by McMillan, Garland, Bennemann and others, eqn [16] needs to be improved to calculate realistic T_c values for simple metals, transition metals, A-15 compounds, alloys, etc., and to understand the interdependence of superconductivity and atomic structure better. Taking into account more details of the electron–phonon coupling, the phonon spectrum, the retardation of the pairing potential, and the remaining repulsive electron–electron interaction ($\propto \mu^*$, pseudo-Coulomb potential), one gets the McMillan equation:

$$T_c \simeq \frac{\omega_D}{1.45} \exp\left(-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1 + 0.62\lambda)}\right), \quad \lambda = 2\int_0^\infty \mathrm{d}\omega \frac{\alpha^2(\omega) F(\omega)}{\omega} \qquad [17]$$

Here, the effective coupling constant λ depends on the DOS, $N(0)$. First $\lambda \propto N(0)$, then λ saturates and decreases again as $N(0)$ increases such as in transition metals. Note that $\lambda = (m^*/m) - 1$. Typically $\mu^* \approx 0.1$, but μ^* may also change strongly near magnetic instabilities.

The DOS for the quasiparticles is given by

$$N_s(\varepsilon) = N_n(0) \frac{\varepsilon}{\sqrt{\varepsilon^2 - \Delta^2}} \qquad [18]$$

This follows from the one-to-one correspondence of old and new quasiparticles, $N_s(\varepsilon)\mathrm{d}\varepsilon = N_n(\xi)\mathrm{d}\xi$. For the thermodynamic potential Ω, which is the appropriate function for a system with given μ, all thermodynamic properties can be calculated using the formulas

$$\left.\frac{\partial \Omega}{\partial g}\right|_{T,V,\mu} = -V\frac{\Delta^2}{g^2}, \quad \Omega_s - \Omega_n = \int_0^\Delta \mathrm{d}\Delta\, \Delta^2 \frac{\mathrm{d}(1/g)}{\mathrm{d}\Delta} \qquad [19]$$

This gives, in particular at T_c, the specific heat (**Figure 4**)

$$c_s - c_n \simeq \frac{4 m p_F}{7\xi(3)\hbar^3} T_c, \quad c_s(T) \propto \mathrm{e}^{-\Delta_0/T} \qquad [20]$$

where $\xi(x)$ is the Riemann–Zeta function, $V = 1$. To obtain the behavior in electromagnetic fields, one uses the recipe $\boldsymbol{p} \to \boldsymbol{p} - e\boldsymbol{A}/c$, $\boldsymbol{p} \to -\mathrm{i}\hbar\nabla$, where $\boldsymbol{A}$ denotes the vector potential ($\boldsymbol{A} = \sum_q \boldsymbol{a}_q \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}$). This

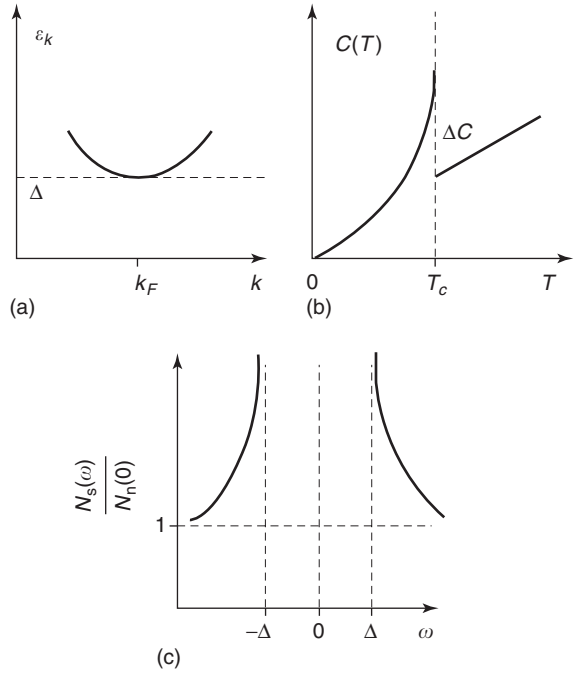


Figure 4 Illustration of (a) the gap Δ in the elementary excitation spectrum, (b) the jump $(C_S/C_n)_{T_c} = 2.43$ in the specific heat, and (c) the electronic DOS.

yields the coupling

$$H_{\text{int}} = \frac{ie\hbar}{2mc}\sum_k (\boldsymbol{v}_k \cdot \boldsymbol{A} + \boldsymbol{A} \cdot \boldsymbol{v}_k)$$
$$= -\frac{e\hbar}{mc}\sum_{k,q} \boldsymbol{k} \cdot \boldsymbol{a}_q c^+_{k+q\sigma} c_{k\sigma} \qquad [21]$$

Here, the spin is unchanged, since $\boldsymbol{A}$ couples dominantly to the orbital motion of pairing electrons. (Note that $\boldsymbol{q} \cdot \boldsymbol{a} = 0$ for tranversal fields.) Thus, one obtains, as in the GL theory for the electric current due to $\boldsymbol{A}$ ($\boldsymbol{j} = en\boldsymbol{v}$) for Cooper pairs ($m \to 2m$, $e \to 2e$):

$$\boldsymbol{j} = -\frac{ne^2}{me}\boldsymbol{A} + \boldsymbol{j}_2 \qquad [22]$$

where the paramagnetic current is $\boldsymbol{j}_2 = e\hbar/m \sum_{k,q} \boldsymbol{k} a^+_{k-q} a_k$, which would vanish if the superconducting wave function ψ is rigid, that is, not affected by $\boldsymbol{A}$.

It is to be noted that, since n_s is the density of Cooper pairs, the first term gives the famous London equation:

$$\boldsymbol{j}_L = -\frac{n_s e^2}{mc}\boldsymbol{A} = -\frac{c}{4\pi\lambda^2}\boldsymbol{A},$$
$$\lambda^2 = \frac{mc^2}{4\pi e^2 n_s} \qquad [23]$$

Hence, by determining the penetration depth λ of an external magnetic field $\boldsymbol{h}$ into a superconductor, one may calculate the superfluid density n_s, an important quantity reflecting whether all electrons at the Fermi surface pair. Evidently, if $\xi > \lambda$, eqn [22] gives the Meissner effect. Using the response theory,

$$\boldsymbol{j}_q = -\frac{c}{4\pi} K_q \boldsymbol{a}_q \qquad [24]$$

where $K_q = \lambda^{-2}(1 + \lambda^2 K'_q)$ and K'_q is calculated within the BCS theory.

The two characteristic length scales ξ and λ suggest that there exist two types of superconductors characterized by $K = \lambda/\xi$. As shown by Abrikosov, they correspond to types I ($\kappa < 1/\sqrt{2}$) and II ($\kappa > 1/\sqrt{2}$). Superconductors of type I repel an external magnetic field, which however may penetrate via an array of quantized flux filaments into type-II ones. For these, one expects vortex lines with $\Delta \simeq 0$ around them. The superconductor compromises both effects in order to profit from the Cooper-pair condensation energy and the magnetic field energy. It is to be noted that the formation of a vortex costs an energy of the order $(\xi h^2/8\pi)$ due to a variation of $\Delta(\boldsymbol{r})$ or the wave function ψ, and reduces the (diamagnetic) field energy by about $(\lambda h^2/8\pi)$. **Figure 5** characterizes the behavior of superconductors in an external magnetic field. From the general formula for the free energy,

$$F_s = F_n - \frac{h^2}{8\pi} \qquad [25]$$

assuming no field penetration into the superconductor, one gets a critical magnetic field H_{c1} below which a pure Meissner superconductor is present. Above H_{c1}, the magnetic flux ϕ (around vortices) may penetrate until $H_{c2} = \phi/\pi\xi^2$ (ϕ = flux of area $(\pi\xi^2)$). As already noted, the flux ϕ is quantized due to the macroscopic phase of the superconductor as a whole. Note that for fields h such that $H_{c1} < h < H_{c2}$ one gets a mixed state, with $H_{c2} = \sqrt{2}KH_{c1}$. Surfaces may favor Cooper-pair condensation. Then there exists another critical field h_{c3} above which superconductivity at the surface is destroyed ($H_{c3} \simeq 1.7H_{c2}$).

Obviously, strong magnetic fields may also affect Cooper pairing by destroying the singlet state. This is of relevance for the interdependence of superconductivity and magnetism in antiferromagnets or ferromagnets. As a result of internal molecular fields acting on the electron spins, one may get a spatially inhomogeneous superconducting state, $\Delta = \Delta(\boldsymbol{k}, \boldsymbol{r})$, and then again a compromise to profit from both Cooper pairing energy and magnetic energy.

The spin–orbit coupling and scattering by local spins (paramagnetic impurities) also affect the singlet pairing. Scattering by paramagnetic impurities with spins via exchange coupling (J), $H_{ex} = -\sum_i J\boldsymbol{S}_i \cdot \sigma$, will, of course, weaken the singlet Cooper pairing. The

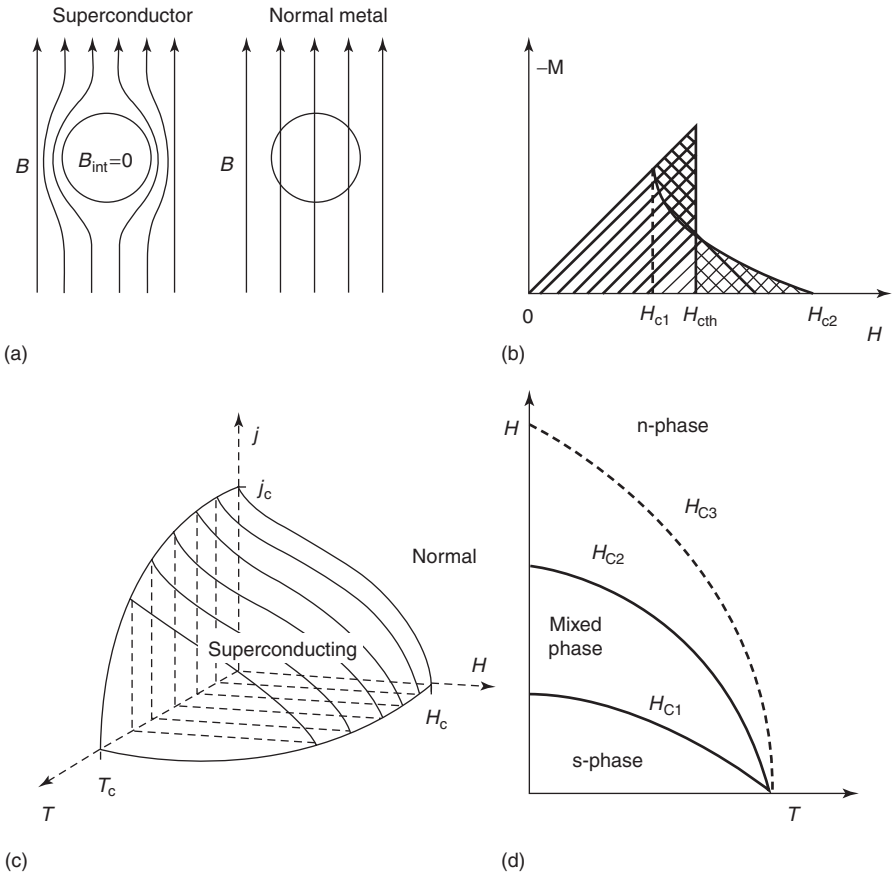


Figure 5 (a) Meissner-effect for type-I superconductors where, below the superconducting transition temperature T_c, the magnetic flux B is abruptly expelled; (b) Typical behavior of type-II superconductors, where, below B_{c2}, the flux penetrates, and perfect Meissner-effect behavior occurs below B_{c1}; (c) dependence of type-I superconductor on B and j and (d) temperature dependence of critical magnetic fields. At H_{c3} surface superconductivity disappears, $B = H + 4\pi M$.

interesting analysis by Abrikosov and Gorkov yields

$$\ln\frac{T_c}{T_{co}} = \psi\left(\frac{1}{2}\right) - \psi\left(\frac{1}{2} + \frac{\hbar}{4\pi\tau}\frac{1}{T_c}\right) \qquad [26]$$

where τ is spin scattering time and $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the digamma function. The discovery of gapless elementary excitations was unexpected. The effects on the singlet superconducting state due to spin-mixing via spin–orbit coupling and spin-flip scattering via exchange interaction are observed through the temperature dependence of the Knight shift $K(T)$, for example, and cause $K > 0$ for $T \to 0$. Clearly, the spin splitting of the Fermi surface for spin $\uparrow$ and $\downarrow$ electrons also weakens pairing of k and $-k$ electrons.

For a better understanding of the phonon and electron dynamics due to H_{eph}, one can use the Hamiltonian $H = H_0 + H_{eph} + H_{el}$, where H_0 refers to phonons and electrons in the absence of H_{eph}, and H_{el} to the electron–electron interactions. The dynamic field operators ψ, ψ^+, and φ are treated on the same level. This yields renormalizations of both phonon and electron excitations. The physical situation is illustrated in **Figure 6**. Then one obtains a frequency-dependent complex superconducting order parameter $\Delta(\omega)$ and, for the DOS (see **Figure 7**),

$$N_S(\omega) \sim \mathrm{Re}\frac{\omega}{\sqrt{\omega^2 - \Delta^2(\omega)}} \qquad [27]$$

Using matrix Green's functions, the Nambu notation $\psi_k^+ = (c_{k\uparrow}^+ c_{-k\downarrow})$, and Matsubara frequencies, one gets for $\mathscr{G} = -\langle T_\tau \psi_k(\tau)\psi_k^+(0)\rangle$, and similarly for $\mathscr{D}(q,\tau)$, the Dyson equations shown in **Figure 6** where

$$\mathscr{G} = \begin{pmatrix} G & F \\ F^+ & \tilde{G} \end{pmatrix}, \quad \mathscr{G}_0 = \begin{pmatrix} G_0 & 0 \\ 0 & \tilde{G}_0 \end{pmatrix}, \quad \mathscr{D} = \begin{pmatrix} D & 0 \\ 0 & \tilde{D} \end{pmatrix}$$

One gets (for $D \equiv (\mathrm{i}\omega_n Z)^2 - (\varepsilon_k + \xi)^2 - \phi^2$, $\omega_n = (2n+1)\pi T$)

$$G(\boldsymbol{k}, \mathrm{i}\omega_n) = \frac{(\mathrm{i}\omega_n Z + \varepsilon_k + \xi)}{D}, \qquad F = \frac{\phi(\boldsymbol{k}, \mathrm{i}\omega_n)}{D} \qquad [28]$$

for the electron (G) and Cooper pair Green's function (F), respectively. Note that $D = 0$ yields the elementary excitations and $\Delta = \phi/Z$ the order parameter

$Z \simeq 1$. After continuation to the real axis, one obtains

$$\Delta(\omega) = \frac{N(0)}{Z(\omega)} \int_{\Delta_0}^{\omega_c} d\omega' \, \mathrm{Re} \frac{\Delta(\omega')}{\sqrt{\omega'^2 - \Delta^2}} \times (K_{ph} - U_c), \quad (T \to 0) \qquad [29]$$

with $U_c = V_c/(1/N(0)V_c \ln(\omega_m/\omega_0))$ due to electron–electron interactions, ($\omega_m \sim \varepsilon_F$, ω_0 – cutoff parameter) and (from phonon spectrum $B_\alpha(q, z)$)

$$K_{ph} = \sum_\alpha \int_0^{2k_F} \frac{dg\, q}{2k_F^2} \int_0^\infty dz \, B_\alpha(q,z) |g_{q\alpha}|^2 \times \left(\frac{1}{\omega' + \omega + z - i\delta} + \frac{1}{\omega' - \omega + z - i\delta} \right) \qquad [30]$$

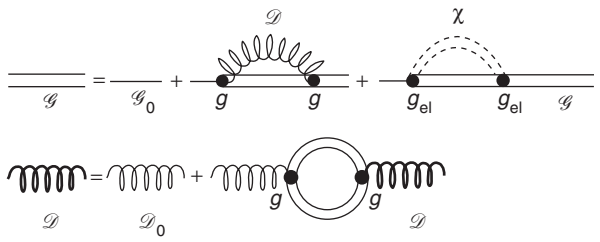


Figure 6 Illustration of the dynamics of the coupled electron and phonon systems using Feynman diagrams and thermodynamic matrix Green's functions $\mathscr{G}$, $\mathscr{D}$ (Eliashberg, McMillan type theory). The graph(= = =)refers to an additional electron–electron interaction and also to spin excitations (and impurity scattering) (g = eph., g_{el} = el−el.).

These equations are the essence of the Eliashberg-type theory (McMillan) and produce, for a simplified phonon spectrum, the results obtained by Scalapino, Schrieffer, and Wilkins (SSW) shown in **Figure 7**. This analysis is of utmost significance, since it shows definitely that for BCS-type superconductors the electron–phonon coupling causes Cooper pairing. Note that phase-coherent Cooper pairs occur at T_c. The phonon frequencies are

$$\omega_q^2 \simeq (\omega_q^0)^2 + 2\omega_q^0 \frac{|g_g|^2}{V_c(q)} \left(\frac{1}{\mathscr{E}(q,\omega)} - 1 \right) \qquad [31]$$

where $\mathscr{E}(q,\omega)$ is the dielectric function. Presumably, the Dyson equation illustrated in **Figure 6** also applies if $\mathscr{D}$ is replaced by the Green's function of spin excitations, the dynamical spin susceptibility $\chi(q, i\omega_n)$, or of other modes that possibly cause Cooper pairing. **Figure 6** describes not only singlet, but also triplet Cooper pairing such as in He^3, and also suppression of Cooper pairing due to magnetic interactions; see "Further reading" section. Superconductivity in different materials, for example, isotropic lattices, layered lattice structures, nanostructures, and amorphous metals can be described adequately by taking into account details of the coupling constant $g_{kk'\lambda}$ and its renormalization due to vertex corrections (Ward identities) (**Table 1**).

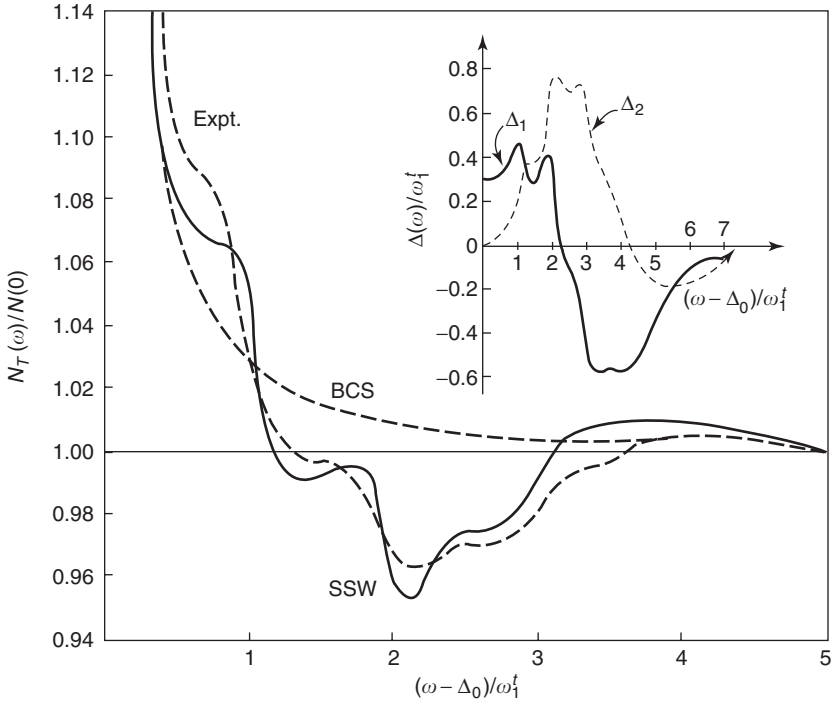


Figure 7 Tunneling DOS with structure due to electron–phonon interaction. Also $\Delta(\omega) = \Delta_1 + i\Delta_2$ is given (ω^t: phonon frequency).

Table 1 Estimates of the superconducting transition temperature T_c in disordered and amorphous metals

Material	$(T_c/T_{c0})_{expt}$	$(T_c/T_{c0})_{calc.}$
Al	~5	4.9
Pb	~0	0
Ga	~8	8
Sn	~1.3	2
In	~1.3	1.2

Courtesy of Bŭckel G *et al.*

Summary and Outlook

Both BCS theory and its extension including dynamical effects yielding the Abrikosov–Gorkov and Eliashberg theory explain superconductivity resulting from electron–phonon coupling. It still remains a challenge to calculate T_c from first principles and its dependence on atomic structure, lattice instabilities, and magnetic activity.

Changes in the atomic structure affect the phonon spectrum and the coupling constant $g_{q\lambda}$ and T_c, Δ, etc., as observed for disordered and amorphous metals. In transition metals and alloys thereof with large $N(0)$, one may attempt to relate T_c to the cohesive energy, since $g \sim \langle \nabla t_{ij} \rangle \sim t$, where t_{ij} is the hopping integral for electrons. Similarly, for transition metal alloys A_xB_{1-x}, the different bonds (AA, AB, and BB) may play a role in determining $\Delta(x)$ and $T_c(x)$. For this, the BCS theory has been extended by Bennemann and co-workers, using the coherent-potential approximation. In case of strong anisotropy of the lattice structure as in layered structures, one may expect approximately two coupling constants g_i ($g_i \sim \nabla t_i \cdot e$), different critical fields H_{c2}^{a-b}, H_{c2}^{c} and two gaps ($i \simeq \pi, \sigma$, $a-b$ in plane, $c \perp a-b$)

$$\Delta_i Z_i \simeq \pi T \sum_{l,m} \frac{\lambda_{il} - \mu_{il}^*}{\sqrt{\omega_m^2 + \Delta_l^2(m)}} \Delta_l(m) \qquad [32]$$

Figure 8, shows the results for MgB_2. Note that hybridization of π, σ electrons gives a common T_c.

The interplay of magnetism and superconductivity suggested already by the occurrence of superconductivity in the periodic table is, for several reasons, of utmost interest. This has, for example, been studied for (film) structures $N_1|N_2|N_3$. Results are given in **Figure 9**.

Organic superconductors display, in particular, interesting dimensional structural effects and those due to charge-wave and magnetic instabilities. Also, astonishingly, superconductivity may be induced by strong magnetic fields. In **Figure 10** the interesting phase diagram of κ-(BEDT-TTF)$_2$-Cu[N(CN)$_2$]Br, a typical composition of a charge-transfer salt, is

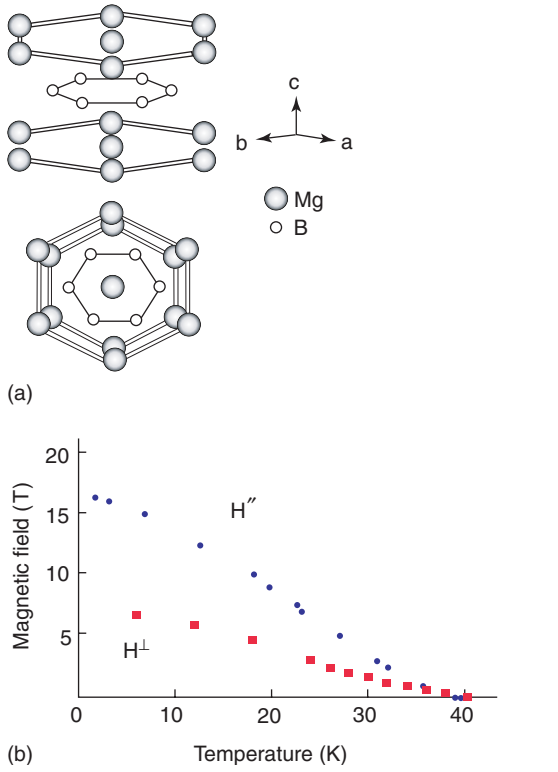


Figure 8 (a) Structure of MgB_2 (AlB_2, etc.). The boron planes seem to play an important role regarding superconductivity and Cooper pairing. The Mg–B bonds are softer than the B–B ones. (b) Critical upper magnetic fields $H''_{c2}(T)$ and $H^{\perp}_{c2}(T)$ referring to the B-planes and direction perpendicular to it, respectively. The anisotropic behavior of $H_{c2}(T)$ suggests two-gap behavior.

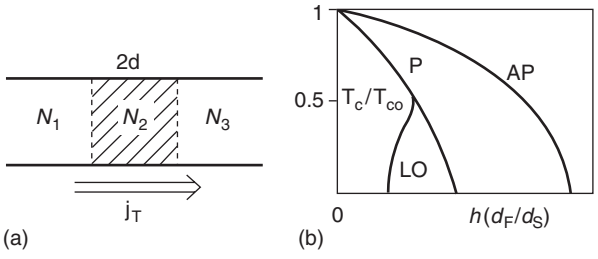


Figure 9 (a) Illustration of a tunnel junction $N_1/N_2/N_3$, where N_i may refer to singlet and triplet superconductors and to normal-state metals, including ferromagnets. The tunnel current j_T may refer to single electrons or Cooper pairs and transport of charge and spin. Depending on the thickness, $2d$, of the tunnel medium proximity effect and Andreev reflection at S/N interfaces plays a role. (b) Phase diagram of a (F_1/S/F_3) sandwich. S refers to a singlet superconductor and F_1,F_3 to ferromagnets with parallel (P) or antiparallel (AP) orientation of their magnetization. LO is the Larkin–Ovchinnikov phase, h the exchange field and T_c and T_{co}, the superconducting transition temperature in the presence and absence, respectively, of the exchange field. d_F and d_S refer to the thickness of the ferromagnetic and superconducting film, respectively.

shown. Such molecules are stacked into layers. Planes of Cu[N(CN)$_2$]Br anions separate these layers. Within these layers, the molecules form dimers and electrons or holes can then hop easily from one

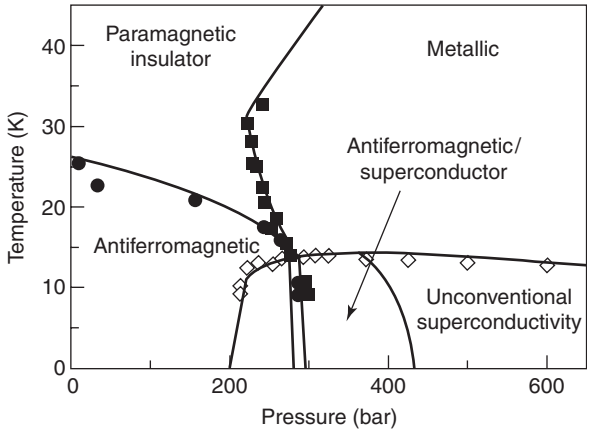


Figure 10 Phase diagram of κ-(BEDT-TTF)$_2$ Cu[N(CN)$_2$]Br, an organic superconductor, from AC-susceptibility and nuclear magnetic resonance.

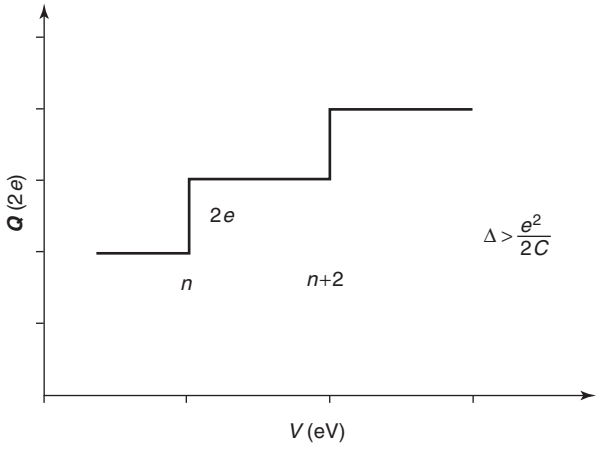


Figure 11 Charging-up behavior of a small superconductor with capacitance C and charge $Q = en$, n even, in a potential V ($\Delta > e^2/2C$). This behavior follows from $E_n = E_{n+2}$ and reflects Cooper pair formation. One also has $E_n = E_{n+1} + \Delta$, where E_{n+1} refers to a state with one unpaired electron. Then Q changes by e at $\Delta V \propto (2n+1)$.

molecule to the next one, but not perpendicular to the layers. The salts consist of conducting layers separated by a nonconducting enviroment (anions). Clearly, bond-length changes, also due to pressure, will sensitively change the electrical properties. Also, a magnetic field perpendicular to the layers (along the *c*-axis) affects the system due to the formation of Landau levels, which will pass the Fermi energy as the external magnetic field varies (cf. de Haas–van Alphen oscillations). In stronger magnetic fields quantum mechanical interband tunneling also occurs. Note that superconductivity occurs at relatively low temperatures, usually upon applying pressure. Experiments indicate unconventional Cooper pairing ($2\Delta_0 \sim 7k_B T_c$, non-*s*-symmetry of $\Delta_{\boldsymbol{k}}$). Whether phonons are involved in Cooper pairing must be clarified by further analysis.

In an external magnetic field, interesting behavior, such as Landau-level formation and spin-split bands, may result due to the layered structure. Possibly, Cooper pairs ($\boldsymbol{k}\uparrow$, $-\boldsymbol{k}+\boldsymbol{q}\downarrow$) may form. A Larkin–Ovchinnikov (LO) state with $\Delta_{\boldsymbol{k}} \sim \cos(\boldsymbol{kr})$ may occur. In general, the response to an external magnetic field is highly anisotropic. Furthermore, it is interesting that superconductivity may be induced in λ-(BETS)$_2$FeCl$_4$ by a magnetic field h, which destroys long-range magnetic order of the Fe^{3+}. Field-induced superconductivity may also occur in α-(BEDT-TTF)$_2$KHg(SCN)$_4$ by affecting the charge density-wave (CDW) with the external magnetic field h. Further experiments are needed; however, rich behavior may be expected in general, since CDW and SDW excitations, Landau levels, and 2D properties are present.

Superconductivity in small particles (radius R) due to electronic level discretisation ($\delta_2 \approx \mathscr{E}_F/n \approx (\hbar v_F/R)(k_F R)^{-2}$, $\xi \sim (R\delta/kT_c)(Rk_F)^2$), and charge fluctuations, $Q = en \leftrightarrow Q \pm e$, as a result of single electron and Cooper pair hopping, exhibit interesting behavior.

The charging of small particles is controlled by electrostatic energy. The change of the charge $Q = en \rightarrow Q \pm 1$ causes an energy change $e^2/2C$, where C is the capacitance of the small particle, leading to a Coulomb blockade (in tunneling, for example). At temperatures $T \simeq 0$, due to the electrostatic energy, it is ($\Delta Q = e\Delta n$)

$$E_n = \frac{Q^2}{2C} - \frac{Q}{C} C' V \qquad [33]$$

(C, C' are capacitances of a system including voltage contact and V is the external potential), and the blockade is periodically lifted as a function of n and V ($E_{n+1} = E_n$). In the superconducting state, one has for $T \simeq 0$ that $\Delta > e^2/2C$. The situation is illustrated in **Figure 11**. From $E_{n+1} + \Delta = E_n$ and $E_{n+2} = E_n$, n even and large sea of Cooper pairs, a period doubling results, with respect to the normal state. Note that the equation $E_{n+2} = E_n$ might not hold for smaller density of Cooper pairs. In view of the strong quantum-mechanical behavior of small particles, structures consisting of a larger number of small particles seem very interesting (mesoscopic systems).

An ensemble of grains for illustration may exhibit transitions to superconductivity at T_{c1}, where the single grains get superconducting, and at T_{c2}, where the whole ensemble becomes superconducting ($T_{c2} \leqslant T_{c1}$) globally and phase coherently; see **Figure 12**. Generally, an ensemble of grains may resemble an alloy-like array of S/S and S/N junctions (S = superconducting, N = normal state grains). The behavior of an ensemble of grains depends, of course,

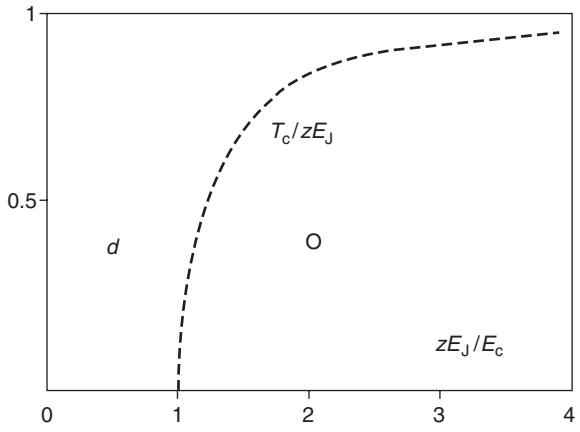


Figure 12 Transitions in an ensemble of superconducting quantum dots controlled by the Josephson-energy E_J and the Coulomb-energy $E_c = e^2/2C$. T_c refers to the global superconducting transition temperature of the ensemble. Phase O refers to a globally ordered superconducting state of the ensemble and phase *d* to superconducting grains with no global Cooper pair phase coherence, z is the coordination number.

on the coupling and electron hopping between the grains. Also the structural order of the grains plays a role. A regular lattice-like arrangement of the grains (one-dimensional or two-dimensional topology) is of particular interest. For small distances between the cluster particles, electron tunneling and inelastic Cooper pair tunneling (Josephson coupling) occur. Hence, charge fluctuations are present in such granular superconductors. The resultant interesting Josephson effect can be described by

$$H = H_T - \sum_{i,j} E_g(i,j)\cos\Phi_{ij} + \frac{2e^2}{C}\sum_{i,j}(N_i - N_j)^2 + \cdots \qquad [34]$$

The first term refers to normal electron hopping between the grains, the second term to Cooper-pairs tunneling between neighboring superconducting grains (Josephson effect), and the last term is the electrostatic energy due to different charges of neighboring grains i, j. For simplicity, the same capacitance C and E_J (Josephson energy) are used. The phase difference $\Phi_{ij} = \Phi_i - \Phi_j$ refers to the phases Φ_i of the superconducting order parameter $\Delta_i = |\Delta_i|e^{i\Phi_i}$ of grain i. It is important to note that $[\Phi, N] = i$, where N is the Cooper-pair number operator.

In conclusion, BCS superconductivity has been discussed. While the basis of the BCS theory is well developed, details of the physics may have interesting consequences that have not been understood yet. New important discoveries can be expected. The relationship of phase-coherent and, in particular, phase-incoherent Cooper-pairing with the Bose–Einstein condensation (BEC) is of fundamental interest. One expects, in particular, for more local pairing, perfect boson behavior of the two paired electrons and thus an intimate relationship between Cooper pairing and BEC. Meissner-type superconductivity and BEC should result only for phase-coherent Cooper pairing.

Acknowledgment

The author is grateful to J W Garland, W Buckel, M Avignon, L Tewordt, M Peter, R Parks, D Manske, R Schrieffer, and J Bardeen, and many more for helping him in understanding superconductivity. A helpful visit at the Aspen Institute of Physics is also gratefully acknowledged.

See also: Magnetism, History of; Quantum Mechanics: Methods; Statistical Mechanics: Classical; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductivity: Tunneling; Superconductors, High T_c.

PACS: 74.20.Fg

Further Reading

Abrikosov AA, Gor'kov LP, and Dzyaloshinski IE (1963) *Quantum Field Theory in Statistical Physics*. NJ: Englewood Cliffs.

Bennemann K and Ketterson J (2003) *The Physics of Superconductors*. Berlin: Springer.

Parks RD (1969) *Superconductivity*. New York: Marcel Dekker.

Schrieffer JR (1964) *Theory of Superconductivity*. New York: Benjamin.

Tinkham M (1975) *Introduction to Superconductivity*. McGraw-Hill.

Superconductivity: Critical Currents

A Gurevich, University of Wisconsin, Madison, WI, USA

Introduction

The superconducting state exists at temperature T below the critical temperature T_c and magnetic fields below the upper critical field $H_{c2}(T)$. A superconductor can carry any weakly dissipative current of density J smaller than the critical current density $J_c(H,T)$, in the region of the H–T–J phase diagram as shown in **Figure 1**. Both the critical current density $J_c(H,T)$ and the irreversibility field $H^*(T) < H_{c2}(T)$ at which $J_c(H)$ vanishes are essential parameters invaluable for applications of superconductors in power transmission lines, magnets for medical magnetic resonance imaging systems, etc. This article gives a background to the basic mechanisms of critical currents, which are determined not only by the physics of vortices in superconductors but also by complex multiscale defect structures of superconducting materials.

There are two types of superconductors, which exhibit very different mechanisms of current transport. In the type-I, superconductors represented mostly by pure metals, such as Al, Pb, and Sn current can only flow in a narrow surface layer, due to the magnetic flux expansion and the ideal diamagnetism of the superconducting state (Meissner effect). Once the magnetic field generated by the transport current I on the sample surface exceeds the "thermodynamic critical field" H_c, a type-I superconductor switches into a resistive state at the critical current $I_c = 2\pi R H_c$ (for a round wire of radius R). However, many superconducting alloys, intermetallic and cuprate high-T_c layered compounds are type-II superconductors, which can carry bulk nondissipative currents in a wide magnetic field range above the "lower critical field" H_{c1}, but below the "upper critical field" H_{c2}. Critical currents of type-II superconductors in the mixed state at $H_{c1} < H < H_{c2}$ are determined by the pinning of vortices by various defects of the crystalline structure. Because the upper critical fields H_{c2} of many compounds can reach very high values $\sim 20 - 100$ T, it is the high-field type-II superconductors, which are of interest in power and magnet applications. The current-carrying capability of type-II superconductors is determined by multiple mechanisms on very different length scales: from the nanometer scales set by the size of vortex cores to the micrometer and larger scales, which define the scales of current flow in polycrystals. An optimum defect structure can, in principle, be produced by appropriate treatment of the material.

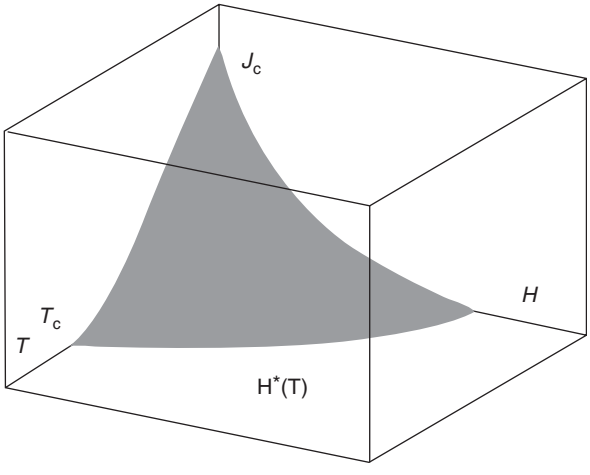


Figure 1 Typical $J_c(T,B)$ surface which defines the region $J < J_c$ on the H–T diagram where a superconductor can carry weakly dissipative currents.

Elementary Pinning Mechanisms

Under equilibrium conditions, magnetic flux penetrates into the bulk of a type-II superconductor above the lower critical field $H_{c1}(0) \sim 10-20$ mT for many high-field materials. At $H > H_{c1}$, this magnetic flux exists in the form of a hexagonal lattice of quantized vortex lines. Each vortex is a tube of radius of the London magnetic penetration depth $\lambda(T)$, in which screening currents circulate around a small nonsuperconducting core of radius $\xi(T)$, where the coherence length $\xi(T)$ at $T = 0$ quantifies the size of the Cooper pairs. Bulk vortices exist if the Ginzburg–Landau parameter $\kappa = \lambda/\xi$ exceeds $2^{-1/2}$, as characteristic of many superconducting compounds, especially such extreme type-II superconductors as Nb_3Sn or layered high-T_c cuprates with $\kappa = 20-100$. The flux produced by screening currents in a vortex equals the flux quantum $\phi_0 = 2.07 \times 10^{-15}$ Wb, so the vortex density $n = B/\phi_0$ is proportional to the magnetic induction B. Bulk superconductivity is destroyed when the normal cores overlap at the upper critical field, $H_{c2}(T) = \phi_0/2\pi\mu_0\xi^2(T)$. In isotropic superconductors such as Nb–Ti and Nb_3Sn, vortex lines are continuous, but in anisotropic layered high-T_c compounds, such as $Bi_2Sr_2Ca_2Cu_3O_x$ and $Bi_2Sr_2CaCu_2O_x$, vortex lines consist of stacks of weakly coupled "pancake" vortices whose circulating

currents are mostly confined within the superconducting CuO_2 planes.

Superconductors can carry bulk currents if there is a macroscopic vortex density gradient (or bending of vortices in thin films) defined by the Maxwell equation $\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J}$. This gradient can only be sustained by pinning the vortices (flux pinning) at microstructural defects. Flux pinning is determined by spatial perturbations of the free energy of the vortex lines due to local interactions of their normal cores and screening currents with these microstructural imperfections. In addition, the vortex structure is subjected to the Lorentz force $\boldsymbol{F}_L = \boldsymbol{J} \times \boldsymbol{B}$ of the macroscopic current density $\boldsymbol{J}$. The critical current density $J_c(T, B)$ is then defined by the balance of the pinning and Lorentz force $J_c(T, B)B = F_p(T, B)$, where F_p is the volume pinning force produced by pinning defects in the strongly interacting array of flux lines. Ideally, a type-II superconductor can carry any nondissipative current density J smaller than J_c. If J exceeds J_c, a superconductor switches into a dissipative vortex flow state driven by the Lorentz force. Flux pinning sets the upper limit to J_c, which itself cannot exceed the depairing current density J_d, the maximum current density circulating near the vortex cores:

$$J_c = \frac{\phi_0}{3\sqrt{3}\pi\mu_0\lambda^2\xi} \qquad [1]$$

Pinning defects can be classified by their spatial extent and/or by mechanisms of the elementary pinning forces f_p between vortices and defects. These defects can be either randomly distributed point pinning centers, such as nanoscale nonsuperconducting precipitates and oxygen vacancies in high-T_c superconductors, or defects extended along vortex lines (grain boundaries, dislocations or columnar radiation tracks). In turn, elementary pinning interactions can be either due to a short-range interaction of the vortex cores with point defects, or due to a more long-range magnetic interaction with extended planar defects. The core pinning results from the gain in the superconducting condensation energy as the core crosses a defect, the strongest pinning force corresponding to defect sizes $\sim\xi$ (see **Figure 2**). Magnetic pinning results from the deformation of vortex screening currents near a planar defect on the scale $\sim\lambda$, as shown in **Figure 3**. The magnetic pinning force $f(x) = \phi_0^2/4\pi\mu_0\lambda^2 x$ can be rather strong if the distance x between the vortex and the defect is smaller than the interaction length $l = \phi_0/2\pi\mu_0\lambda^2 J_b$. Here, $l < \lambda$ is the distance from the vortex core at which the circulating current density $J(x)$ equals the maximum current density J_b which can pass through the defect.

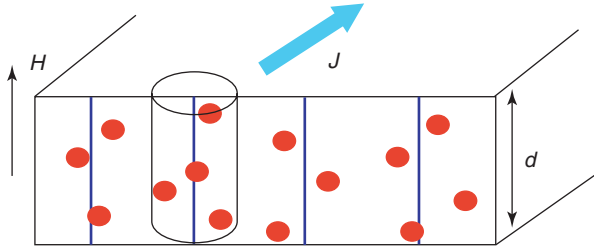


Figure 2 Pinning of vortex cores (vertical lines) by uncorrelated defects (dots) in a film of thickness d in a magnetic field H. The cylindrical region around the vortex cores depicts the region of circulating screening currents.

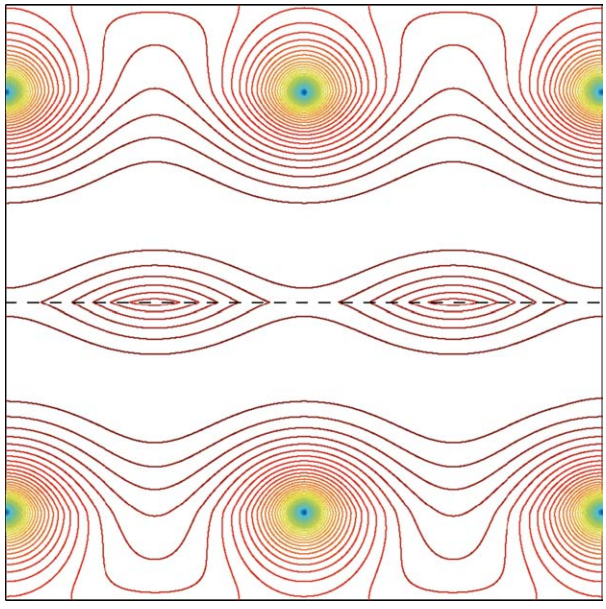


Figure 3 Deformation of vortex current streamlines near a planar defect. If the distance between normal cores of bulk vortices and the defect is smaller than the magnetic interaction length $l = \phi_0/2\pi\mu_0\lambda^2 J_b$, vortices are attracted strongly by the defect. Vortices localized on the defect have no normal cores, which turn into extended regions of length l along the defect in which the superconducting gap is not suppressed by circulating vortex currents. These elongated vortices on the defect have a highly anisotropic pinning force, which is maximum for J parallel to the defect and minimum for J perpendicular to the defect.

Identification of the most effective pinning defects in a particular superconductor is a difficult problem. The best information about the limits to flux pinning in practical superconductors is obtained from extensive studies of Nb–47%Ti, in which strong pinning by a dense ~20–25 vol.% lamellar structure of nonsuperconducting α-Ti ribbons ~1 nm (0.2ξ) thick can produce a J_c that approaches 5–10% of J_d at zero field and 4.2 K. The flux pinning mechanism is also known for Nb_3Sn, in which J_c is determined by the magnetic interaction of vortices with grain boundaries. These materials have zero-field J_c values, which can exceed 1 MA cm^{-2} at 4.2 K. Details of flux pinning

Table 1 Parameters of some type-II superconductors

Material	T_c (K)	$\lambda(0)$ (nm)	$\xi(0)$ (nm)	H_{c2} (T)	H^* (T)	J_d (4.2 K) (MA cm^{-2})	J_c (MA cm^{-2})
Nb–Ti	9	240	4	12 (4.2 K)	10.5 (4.2 K)	36	~0.4 (5 T, 4.2 K)
Nb_3Sn	18	65	3	27 (4.2 K)	24 (4.2 K)	770	~1 (0 T, 4.2 K)
$YBa_2Cu_3O_7$	92	150 (*ab* plane)	1.5 (*ab* plane)	>100 (4.2 K)	5–7(77 K)	300	~1–10(0 T, 77 K)
$Bi_2Sr_2Ca_2Cu_3O_x$	108	150 (*ab* plane)	1.5 (*ab* plane)	>100 (4.2 K)	~0.2 (77 K)	300	~1 (0 T, 77 K)

interactions in high-T_c superconductors are largely unknown, partly because ξ is so small that practically any atomic-size defects such as oxygen vacancies can pin vortices. The critical temperature of high-T_c superconductors is extremely sensitive to the carrier (hole) density, which is in turn determined by local oxygen nonstoichiometries, so that even a weak hole-depletion at crystalline defects can locally drive a high-T_c superconductor into an antiferromagnetic insulator. The proximity of the high-T_c superconducting state to the metal–insulator transition, the *d*-wave pairing symmetry, the very short in-plane coherence length $\xi(0) = 1.5-2$ nm and the much longer Debye screening length $l_D \sim \xi(0)$, all combine to produce significant suppression of superconductivity near crystal defects, such as impurities, dislocations, and grain boundaries, thus making them effective pinning centers. Characteristic parameters of some superconductors are summarized in **Table 1**.

Single-Vortex and Collective Pinning by Point Defects

Calculation of J_c requires a summation of elementary pinning forces acting on flexible, strongly interacting vortex lines. This is a very difficult theoretical problem, which has not been completely resolved so far. Here, the basic physics of the pinning force summation is discussed, by first evaluating J_c at low B for a single straight vortex interacting with randomly distributed precipitates in a film of thickness d (**Figure 2**). If d is much greater than the mean spacing l_i between the precipitates, a vortex interacts with $N \sim r_p^2 d n_p$ pinning centers, adjusting its position so that the net force vanishes. Here, r_p is the pinning interaction radius, and $n_p = l_i^{-3}$ is the density of precipitates. The transport current displaces the vortex from a local minimum of the pinning potential, so that random elementary pinning forces do not compensate each other. The balance of the total pinning force from uncorrelated defects in the volume dr_p^2 against the total Lorentz force, $f_p N^{1/2} = d\phi_0 J_c$ yields

$$J_c = \frac{1}{\phi_0}\sqrt{\frac{\gamma}{d}} \qquad [2]$$

where $\gamma = f_p^2 r_p^2 n_p$ is the pinning strength parameter. The critical current density $J_c(d)$ decreases as the film thickness increases. For $d \to \infty$, eqn [2] gives $J_c = 0$ because the net pinning force per unit length of an infinite straight vortex in a random potential averages to zero.

Calculation of bulk J_c must, therefore, take into account local bending of vortices near pinning defects, using the following approach of a collective pinning theory. Now introduce a correlation length L_c along the vortex line defined by the relation $\langle u(L_c)u(0)\rangle = r_p^2$, where $\langle u(z)u(0)\rangle$ is a correlation function of pinning-induced transverse vortex displacements. The relative displacement of the ends of a vortex segment of length L_c exceeds the pinning interaction radius r_p, so a vortex may be regarded as an array of straight segments of length L_c pinned independently. Therefore, J_c can be evaluated from eqn [2] with d replaced by L_c. In turn, L_c is evaluated from the balance of a characteristic pinning energy $\phi_0 J_c r_p$ and an elastic bending energy $\varepsilon(r_p/L_c)^2$, where $\varepsilon = (\phi_0^2/4\pi\mu_0\lambda^2)\ln\kappa$ is the vortex line tension. Hence, the condition $\phi_0 J_c r_p = \varepsilon(r_p/L_c)^2$ along with eqn [2] at $d = L_c$ yields

$$J_c = \frac{\gamma^{2/3}}{\phi_0 \varepsilon^{1/3} r_p^{1/3}}, \qquad L_c = \frac{\varepsilon^{2/3} r_p^{2/3}}{\gamma^{1/3}} \qquad [3]$$

Due to the collective interaction of a vortex with many defects, J_c increases as the line energy ε decreases, because a more flexible vortex can bend more easily to accommodate pinning centers, as depicted in **Figure 4**. Meanwhile, the correlation length L_c increases as the pinning parameter γ decreases, resulting in a "dimensional crossover" in films, from the two-dimensional pinning of rigid straight vortices at $d < L_c$ to the three-dimensional pinning of flexible vortices for $d > L_c$.

A theory of weak collective pinning in strong fields $H \gg H_{c1}$ involves two characteristic spatial scales: the correlation length along the vortex lines $L_c(T, B)$ and the correlation length $R_c(T, B)$ perpendicular to the vortex lines. These lengths define the correlation volume $V_c = R_c^2 L_c$ – a mesoscopic region in which pinned vortices still maintain the positional hexagonal short-range order characteristic of the ideal flux

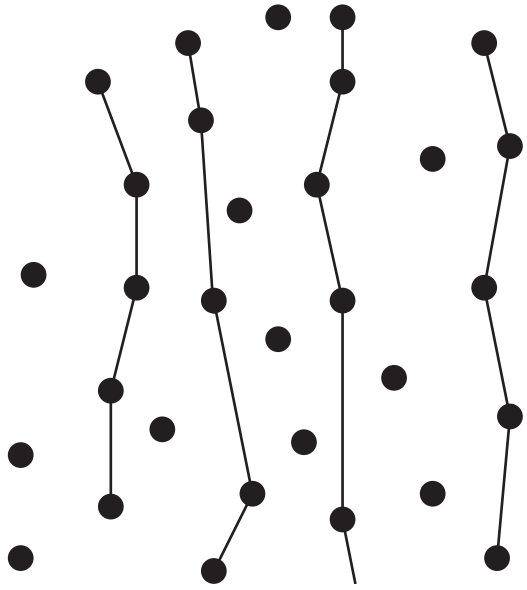


Figure 4 Vortex lines pinned by randomly distributed precipitates.

line lattice. The positional long-range order of vortices is destroyed by pinning on scales larger than the correlation lengths L_c and R_c. The critical current density is then defined by the balance of the Lorentz force applied to the correlation volume, BV_cJ_c against the pinning force from spatially uncorrelated defects, $f_pN_c^{1/2}$, where $N_c = n_pV_c$. This yields

$$J_c = \frac{f_p}{BR_c}\sqrt{\frac{n_p}{L_c}} \qquad [4]$$

where both $L_c(T,B,\gamma)$ and $R_c(T,B,\gamma)$ depend not only on T and B, but also on the pinning parameter γ. The lengths $L_c(T,B,\gamma)$ and $R_c(T,B,\gamma)$ are determined self-consistently from the conditions $\langle u(R_c,0)u(0,0)\rangle = r_p^2$ and $\langle u(0,L_c)u(0,0)\rangle = r_p^2$, where the correlation function of vortex displacements $\langle u(\boldsymbol{r})u(0)\rangle$ is calculated using a continuum elasticity theory for the vortex lattice perturbed by a weak random pinning potential. Depending on the relations between the pinning correlation lengths and the London penetration depth λ and the vortex spacing $a = (\phi_0/B)^{1/2}$, the collective pinning theory predicts a variety of behaviors of $L_c(T,B,\gamma)$ and $R_c(T,B,\gamma)$, which manifest themselves in different dependencies of J_c on T and B. For instance, a crossover from a weak collective pinning to a strong single-vortex pinning regime typically occurs at lower B for which $R_c(T,B,\gamma)$ becomes smaller than the vortex spacing. In this case, pinning-induced correlations of neighboring vortex displacements disappear and vortices become pinned independently. The collective pinning theory for layered high-T_c superconductors is more complicated because it has to take into account a highly anisotropic elastic response of the vortex lattice to pinning forces.

Pinning by Extended Defects

Pinning by extended defects, such as grain boundaries or columnar radiation defects, can be different from pinning by uncorrelated disorder, because a single extended defect parallel to B can pin a vortex segment greater than the correlation length $L_c = \xi(J_d/J_c)^{1/2}$ of the collective pinning theory. Pinning by extended defects can be rather effective and result in high J_c values. Indeed, a single vortex trapped by a cylindrical cavity of radius $\cong\xi$, has the pinning energy U_p of the order of the condensation energy of the vortex core $\phi_0^2/4\pi\mu_0\lambda^2$, yielding a very high $J_c \cong U_p/\xi\phi_0$ of the order of the depairing current density [1]. This model has been used to describe experiments on irradiation of high-T_c superconductors with 1–10 GeV heavy ions (Pb, Au, Xe, etc), which can produce 10–100 μm columnar tracks of about 5–10 nm in diameter comparable to ξ. The density of the columnar defects n_i is normally expressed in terms of the matching field $B_m = \phi_0 n_i$ which typically ranges from 0.1 to 10 T for different irradiation conditions. For $B < B_m$, each vortex is trapped by a columnar defect, resulting in enhancement of J_c and H^* in high-T_c superconducting films, as has been shown by many experimental groups.

Pinning by extended planar defects such as grain boundaries in Nb_3Sn, or by thin α-Ti ribbons in Nb–Ti can also result in high J_c values, $\sim$1–10 MA cm^{-2} at 4.2 K and $H = 0$ (see **Table 1**). Chemical nonstoichiometry, lattice disorder, strain, and charging effects associated with grain boundaries cause local suppression of superconductivity. This can result in the so-called weak link behavior of grain boundaries, which is especially pronounced in high-T_c cuprates because the critical current density through grain boundaries J_b decreases exponentially as a function of the misorientation angle θ between neighboring grains if $\theta > \theta_0 \cong 4-5°$. As a result, J_b through grain boundaries or thin α-Ti ribbons in Nb–Ti, is usually much smaller than the depairing current density [1] in the bulk. It is the suppression of J_b as compared to J_d, which causes a strong magnetic and core interaction between a vortex and a planar defect, in particular, the radical change in the structure of the vortex core as it moves closer to the defect shown in **Figure 3**. Therefore, there are two different types of vortices in polycrystals: vortices in the grains pinned by a long-range magnetic interaction with grain boundaries and elongated vortices on grain boundaries. The critical current density J_c can be evaluated by balancing the

total Lorentz force BJ_cD^3 in the grain of size $\cong D$ against the total pinning force proportional to the surface area of the grain D^2. This yields the inverse dependence of $J_c \propto D^{-1}$ on the grain size, as has been observed on Nb_3Sn.

In polycrystals, grain boundaries not only pin vortices, they can also block macroscopic current flow. Indeed, current loops can form in the grains if the critical current density J_g, due to pinning of vortices in the grains, is higher than the current density J_b, which can flow through grain boundaries. This effect of magnetic granularity can limit global critical current density J_c of superconducting polycrystals. In turn, it indicates that there is an optimum value of J_b, which maximizes J_c, because J_c vanishes in the two limiting cases of $J_b = J_d$ (no pinning, since grain boundaries do not disturb currents circulating around vortices) and $J_b = 0$ (current cannot pass through grain boundaries). Calculation of the optimum J_b, in addition to the summation of all vortex interactions, also requires the account of the grain geometry and microscopic mechanisms of current transport through the grain boundary.

Current transport through grain boundaries is very important for high-T_c polycrystals, for which the weak link behavior of high-angle grain boundaries is one of the most serious current-limiting mechanisms. This behavior is due to rapid decrease of $J_b(\theta) = J_0 \exp(-\theta/\theta_0)$ with the misorientation angle θ and the resulting increase of the vortex core size, from ξ to $l(\theta) = \xi J_d/J_b(\theta)$ for $l(\theta) < \lambda$ or to the so-called Josephson length $l(\theta) = [\xi\lambda J_d/J_b(\theta)]^{1/2}$ for $l(\theta) > \lambda$. The extended core of the grain boundary vortices leads to their weaker pinning force parallel to the boundary against the Lorentz force of the transport current flowing perpendicular to the boundary. As a result, the global J_c of polycrystals may be limited by depinning of a small fraction of vortices, which can move more easily along a network of grain boundaries. This percolative motion of grain boundary vortices is impeded by structural defects along grain boundaries (such as facets, nanoprecipitates) and by magnetic pinning shear stress as the grain boundary vortices move past more strongly pinned vortices in the grain (see **Figure 3**). The significant reduction of J_b of grain boundaries in high-T_c superconductors is due to extreme sensitivity of their critical temperature to a rather weak hole depletion, which can drive the high-T_c superconducting state into an antiferromagnetic insulator. The local hole depletion near grain boundaries can be due to local oxygen nonstoichiometry, charging and strain effects produced by the crystalline dislocations which form the grain boundary, or more macroscopic strain fields due to faceting. The values of J_b in $YBa_2Cu_3O_7$ can be improved by local overdoping with appropriate impurities (such as Ca) near grain boundaries. The current-carrying capability of $YBa_2Cu_3O_7$ polycrystals can also be improved by depositing $YBa_2Cu_3O_7$ films onto "biaxially textured" metallic substrates to eliminate the weakest high-angle grain boundaries.

Global Critical Currents of Superconductors

Multiple pinning mechanisms in superconducting materials manifest themselves in different dependencies of the critical current density on T and B. For low-T_c superconductors, the field dependence of J_c can often be described as follows:

$$J_c(h) = J_{c0}(T)h^{-p}(1-h)^q \qquad [5]$$

where $h = H/H^*(T)$ is a reduced field, $H^*(T)$ is the irreversibility field, and J_{c0}, p and q depend on a particular material. For instance, eqn [5] with $p = 0$ and $q = 1$ describes well the $J_c(B)$ for Nb–Ti alloys, while Nb_3Sn is better described by eqn [5] with $p = 1/2$ and $q = 2$, consistent with strong pinning by grain boundaries. In high-T_c superconductors, $J_c(T, H)$, at elevated temperatures $T > 20-30$ K, often exhibits an exponential field dependence, $J_c(H) = J_{c0} \exp(-H/H_0)$, where $H_0(T)$ can be much smaller than $H_{c2}(T)$, especially in layered $Bi_2Sr_2CaCu_2O_x$.

In high-T_c superconductors, the irreversibility field $H^*(T)$ can be much smaller than $H_{c2}(T)$, whereas in low-T_c superconductors, such as Nb–Ti and Nb_3Sn, or even "intermediate-T_c" compound MgB_2 with $T_c = 39$ K, the field $H_{c2}(T)$ is rather close to $H^*(T) \approx 0.8-0.9H_{c2}(T)$ as shown in **Figure 5** (see also **Table 1**). This fact has important consequences for superconductivity applications at the liquid nitrogen temperature of 77 K, because high-T_c cuprates, despite their very high H_{c2} values, can carry weakly dissipative currents only in a small part of their H–T diagram. The reason for this behavior is due to strong magnetic "flux creep," caused by thermally activated hopping of vortices between neighboring pinning positions even if the current density J is below J_c. Vortex creep velocity $\boldsymbol{v}$ is parallel to the driving Lorentz force $[\boldsymbol{J} \times \boldsymbol{B}]$, causing the electric field $\boldsymbol{E} = [\boldsymbol{v} \times \boldsymbol{B}]$ directed along J (in isotropic materials). The E–J characteristic of a superconductor for $J < J_c$ has the form

$$\boldsymbol{E} = E_c \exp\left[-\frac{U(J,T,B)}{k_BT}\right]\frac{\boldsymbol{J}}{J} \qquad [6]$$

where the Boltzmann factor $\exp(-U/k_BT)$ accounts for the fact that only a small fraction of all vortices is

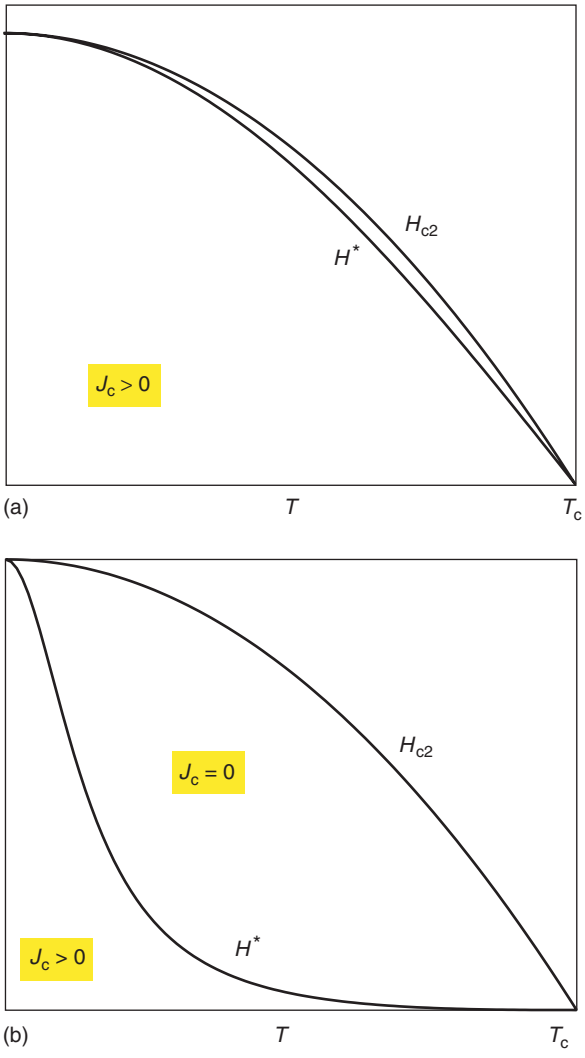


Figure 5 (a) Magnetic phase diagram of low-T_c superconductors, for which $H^*(T) \approx H_{c2}(T)$. (b) Typical magnetic phase diagram of high-T_c superconductors which cannot carry critical currents in a wide field range $H^*(T) < H < H_{c2}(T)$. Both cases (a) and (b) correspond to extreme type-II superconductors ($\kappa \gg 1$) for which the lower critical field $H_{c1} \ll H_{c2}$ is not shown.

thermally activated and thus contribute to the global flux motion. The flux creep activation energy U depends on J, which makes the E–J curve highly nonlinear at $J < J_c$. It is convenient to define U in such a way that $U(J)$ vanishes at $J = J_c$, using the following expressions for U: (1) $U = U_0(1 - J/J_c)$, (2) $U = U_0 \ln(J_c/J)$, (3) $U = U_0[(J_c/J)^{\mu} - 1]$, where $U_0(T, B)$ is a characteristic energy barrier for thermally activated hopping of a correlated group of vortices. These models give the same exponential dependence of E on J at $J \approx J_c$, but rather different behaviors of $E(J)$ at $J \ll J_c$. Typical energy barriers in low-T_c superconductors with strong pinning are $\sim$500–1000 K, whereas in high-T_c superconductors they can be considerably smaller, $U_0 \sim$ 100–400 K, because pancake vortices on neighboring *ab* planes are weakly coupled due to highly anisotropic structures of these materials.

Since current flow causes dissipation at $J < J_c$, critical currents depend on the electric field E_c at which J_c is measured. Generally, J_c is defined as a crossover current density at which a resistive transition from an exponentially small $E(J)$ to a linear flux flow portion of the E–J curve occurs (a conventional criterion of $E_c = 1\ \mu\text{V cm}^{-1}$ is often used). The relation between the critical current densities J_c and J_c' defined at different electric fields E_c and E_c' can be obtained from eqn [6] by linearizing the activation energy $U(J) \approx U_0(1 - J/J_c)$ near J_c:

$$J_c' = \left(1 + \frac{k_B T}{U_0} \ln \frac{E_c'}{E_c}\right) J_c \qquad [7]$$

For low-T_c superconductors ($k_B T_c/U_0 \sim 10^{-2}$), J_c is well defined in a practically accessible electric field window of $1\ \text{nV cm}^{-1} < E < 1\ \text{mV cm}^{-1}$, where $J_c(E_c)$ varies only by a few percent. By contrast, high-T_c superconductors with $k_B T_c/U_0 \sim 0.1$ exhibit a broader resistive transition at $J \approx J_c$ and a larger difference between J_c and J_c'. The parameter $k_B T/U_0$ also defines the rate at which magnetization currents in superconductors, $J(t) = J_c[1 - (k_B T/U_0) \ln(t/t_0)]$, decay due to thermally activated flux creep. The condition $U_0(H^*, T) \approx k_B T$ then determines the irreversibility field H^* at which a highly nonlinear $E(J)$ curve turns into a nearly linear $E(J)$ curve as shown in **Figure 6**. In high-T_c superconductors, such as $Bi_2Sr_2Ca_2Cu_3O_x$, the activation energy of weakly interacting pancake vortices on different *ab* planes is only a few times greater than $k_B T_c$. The resulting strong thermal activation of vortices yields $H^* \ll H_{c2}$, so pinning becomes ineffective in a significant part of the H–T diagram where J_c vanishes (**Figure 5b**). By contrast, in low-T_c superconductors, the reversible region $H^* < H < H_{c2}$ shown in **Figure 5a** is rather narrow as discussed above (see also **Table 1**).

In addition to multiscale pinning mechanisms, global critical current densities $J_c = I_c/A_{\text{eff}}$ of superconductors for a given critical current I_c are often limited by macroscopic scales of current flow defined by an effective current-carrying cross section A_{eff}. Here, A_{eff} is determined by current percolation through arrays of large second-phase precipitates, colonies of high-angle grain boundaries, microcracks, and other inhomogeneities much greater than the pinning correlation lengths L_c and R_c. This results in rather nonuniform current distributions in superconductors, which have been revealed by different experimental techniques, such as magneto-optical

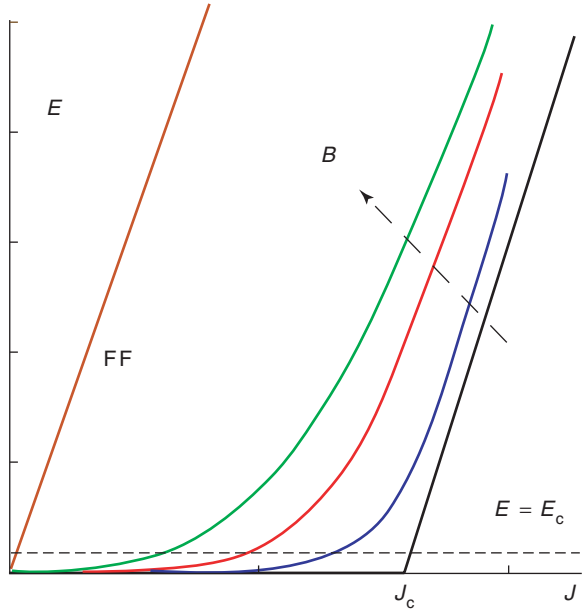


Figure 6 Broadening of the E–J characteristics as B increases. Here the green curve corresponds to B close to but below the irreversibility field B^*, while the FF line shows the Ohmic flux flow E–J characteristic for $B > B^*$. The dashed line shows the electric field criterion E_c at which J_c is defined in transport measurements.

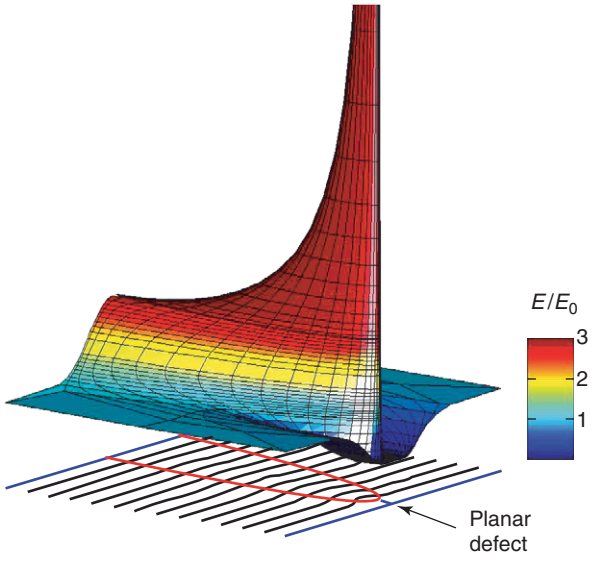


Figure 7 Spatial distribution of the electric field magnitude in a hot spot near a planar defect of length $b = 0.05w$ in a film of width w, calculated for the power-law E–J characteristic $E = E_0(J/J_0)^n$ for $n = 30$. Here, E_0 is the uniform field far away from the defect, and $J_0 = J_c(E_0)$. Below the $E(x, y)$ surface are the corresponding current streamlines in the film.

imaging or micro Hall probe arrays. Current flow in superconductors is described by a highly nonlinear Maxwell equation $\mathrm{div}\, \boldsymbol{J}(\nabla\varphi) = 0$ for the electric potential φ, where $E(J)$ is determined by eqn [6]. The nonlinearity of $E(J)$ results in extended electric field hot spots near defects on scales much greater than the defect size $\sim b$. For instance, for $E = E_c(J/J_c)^n$ characteristic of low-T_c superconductors, the electric field hot spot shown in **Figure 7** has the length $L_\perp \sim nb$ perpendicular to current flow and $L_\parallel \sim n^{1/2}b$ parallel to current flow. For $n \sim 30-50$, planar defects blocking only a few percent of the sample cross section can cause a noticeable voltage drop on the whole conductor. These effects have important consequences for current-carrying capability of superconductors.

See also: Superconductivity: Applications; Superconductivity: BCS Theory; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductors, High T_c.

PACS: 74.25 Ha; 74.25 Qt; 74.25 Sv

Further Reading

Blatter G, Feigel'man MV, Geshkenbein VB, Larkin AI, and Vinokur VM (1994) Vortices in high-temperature superconductors. *Reviews of Modern Physics* 66: 1125–1388.

Brandt EH (1995) The flux-line lattice in superconductors. *Reports of Progress in Physics* 58: 1465–1594.

Campbell AM and Evetts JE (1972) *Critical Currents in Superconductors*. London: Taylor and Frances.

Civale L (1997) Vortex pinning and creep in high-temperature superconductors with columnar defects. *Superconductor Science and Technology* 10: A11–A28.

Larbalestier D, Gurevich A, Feldmann DM, and Polyanskii A (2001) High-T_c superconducting materials for electric power applications. *Nature* 414: 368–377.

Tinkham M (1985) *Introduction to Superconductivity*. New York: McGraw-Hill.

Nomenclature

B	magnetic induction
E	electric field
E_c	electric field criterion
f_p	elementary pinning force
H	magnetic field
H^*	irreversibility field
H_{c1}	lower critical field
H_{c2}	upper critical field
J_c	critical current density
J_d	depairing current density
L_c	longitudinal pinning correlation length
n_p	density of pinning centers
r_p	pinning interaction radius
R_c	transverse pinning correlation length
T	temperature
T_c	critical temperature
$U(J)$	flux creep activation energy
$\gamma = f_p^2 n_p$	pinning parameter
λ	London penetration depth
ξ	coherence length

Superconductivity: Flux Quantization

J R Kirtley and C C Tsuei, T J Watson Research Center, Yorktown Heights, NY, USA

Introduction

According to the Bardeen–Cooper–Schrieffer (BCS) theory of superconductivity, the normal-to-superconducting phase transition temperature (T_c) of a superconductor marks the inception of a macroscopic phase-coherent condensate of Cooper pairs. An energy gap (2Δ), centered at the Fermi energy (E_F) in the density of states of the electronic excitation spectrum, opens up at T_c. This gap grows as the temperature decreases and more pairs are condensed. The current-carrying superelectrons represent a short-circuit to those normal electrons which have not yet condensed.

In conventional low-T_c superconductors, such as Al, Pb, or Sn, the size of the Cooper pairs, as measured by the zero temperature coherence length ξ_0, is $\sim$of 10^3 Å. At temperatures $T \ll T_c$, there are as many as 10^6 pairs coexisting in the so-called coherence volume, $\sim \xi_0^3$. As a consequence of the many-body interactions among these overlapping paired electrons, the Cooper pairs acquire some of the attributes of Bose particles. In particular, in the superconducting state, the Cooper pairs form a superfluid condensate in which all paired electrons occupy the lowest-energy quantum state (ground state) and long-range phase coherence is established. Thus, the whole ensemble of the Cooper pairs in the condensate can be described by a single complex wave function:

$$\Psi = |\Psi| e^{i\varphi} \qquad [1]$$

The amplitude $|\Psi|$ determines quantities like T_c and Δ, which are measures of the pairing strength. It is related to the superfluid pair density n_s by $|\Psi|^2 = n_s$. The quantity n_s can be determined by measurements of the magnetic penetration depth λ or the muon spin relaxation rate. The macroscopic phase φ contains information about static and dynamic charge transport. The well-known hallmarks of superconductivity, such as zero resistance and perfect diamagnetism, all require the establishment of long-range phase coherence among all the Cooper pairs in the condensate. Under the influence of an increasing electric current or a magnetic field, a phase gradient $\nabla\varphi$ develops and grows until the critical field (B_c) or current density (J_c) is reached. At these points the global phase coherence is lost and the system reverts to the normal state.

The stability of the superconducting state can be evaluated from the energy E_{coh} needed to convert a coherence volume of the superconductor into the normal state.

$$E_{coh} = \frac{B_c^2(0)}{8\pi}\xi_0^3 = \frac{1}{2}N(E_F)\Delta_0^2\xi_0^3 \qquad [2]$$

where $B_c^2(0)/8\pi$ is the condensation energy per unit volume, and $N(E_F)$ is the density of states of the charge carriers at the Fermi energy.

For a typical low-T_c superconductor such as Al or Sn, E_{coh} is $\sim$1 eV, which corresponds to a thermal energy of 1.16×10^4 K. On the other hand, the energy cost of removing a Cooper pair from the BCS ground state ($2\Delta_0$) for Al, for example, is only 0.36 meV, or equivalently a thermal energy of $\sim$4 K. Such a large disparity between these two energy scales emphatically underscores the collective nature of the BCS pairing state. The rigidity of the BCS macroscopic wave function can also be attested to by the extremely small effect of thermal fluctuations near T_c. Based on the Ginzburg–Landau formalism, the temperature range δT over which the statistical average of the standard deviation of the wave function is of the same size as the pair wave function itself (i.e., $\langle \Psi^2 \rangle^{\frac{1}{2}} \approx \langle |\Psi| \rangle$) can be estimated by

$$\frac{\delta T}{T_c} \propto \left(\frac{k_B T_c}{8E_{coh}}\right)^2 \text{ for } T \approx T_c \qquad [3]$$

For typical low-T_c superconductors, the values of $\delta T_c/T_c$ are in the range 10^{-8} to 10^{-10}. Indeed, in conventional superconductors, Cooper pairing and phase coherence occur simultaneously.

The anomalous superconducting and normal-state properties of high-temperature superconductors (HTS) present an entirely different scenario. The superconducting cuprates belong to a family of correlated electron systems consisting of superconducting CuO_2 layers (the *ab*-planes) Josephson coupled via the charge reservoir layers in the *c*-axis direction. These quasi-two-dimensional layered superconductors are characterized by a low charge carrier (hole or electron) density, strong anisotropy, and (at least in the underdoped regime) nanometer-scale charge inhomogeneity. In particular, the extremely short and highly anisotropic coherence lengths ($\xi_0^{ab} \approx 20$ Å, $\xi_0^c \leqslant 1$ Å), combined with low charge density ($\sim 5 \times 10^{21}/\text{cm}^{-3}$), make the coherence volume in HTS so small that only a few pairs reside there, instead of 10^6 pairs for conventional superconductors. If the BCS formalism is assumed to be valid, the condensation energy per coherence volume $(\xi_0^{ab})^2\xi_c$ in HTS (see eqn [2]) would be

at least three orders of magnitude smaller than that of the low-T_c counterparts. Furthermore, due to the low superfluid density in the underdoped HTS regime, it has been suggested that phase fluctuations of the complex order parameter (eqn [1]) play an important role in determining various properties near T_c and into the superconducting state. As a consequence, the long-range phase coherence vanishes at T_c, but pairing can persist until another temperature T^* is reached. In between T_c and T^*, there is a suppression in the density of states of the electronic spectrum which is referred to as the "pseudogap". The value of T^* is of the order of 10^3 K near the Mott metal–insulator transition in the highly underdoped region of the cuprate phase diagram. T^* decreases monotonically with increasing doping (and T_c), and is roughly equal to T_c around the optimal doping, beyond which more conventional normal and superconducting states are observed. The idea of pairing without the establishment of macroscopic phase coherence provides a natural explanation for the pseudogap and related phenomena observed in underdoped HTS. However, another school of thought, instead of emphasizing the importance of phase fluctuations, suggests that the pseudogap may stem from some nonsuperconducting mechanisms such as spin, charge, or d-density waves, which represent orders competing with the BCS pair state. Another intriguing alternative possibility is that local pairing above T_c may be derived from the nanometer charge inhomogeneity which promotes pair formation locally and dynamically.

To assess the different unconventional theoretical approaches mentioned above, it is crucial to find out to what extent standard paradigms such as the Fermi liquid formalism and BCS theory can be used for understanding the origin of high-temperature superconductivity. To achieve this goal, it is important to first probe the nature of pairing and phase coherence in HTS.

Flux quantization is one of the most striking manifestations of macroscopic phase coherence in superconductors. Flux quantization based experiments have been employed for studying fundamental aspects of the superconducting state in HTS as well as conventional low-T_c superconductors. Some examples of such studies will be presented in the following sections.

Flux Quantization

The phenomenon of flux quantization is a direct consequence of the requirement that the pair wave function Ψ for describing the superconducting state has to be single-valued at any point in a superconductor. It is a quantum mechanical effect that manifests itself on a macroscopic scale.

When a superconducting sample is placed in a magnetic field, supercurrents are generated to screen the fields from penetrating into the bulk of the sample. This is the "Meissner effect." The length over which the shielding current and magnetic fields vary is the penetration depth λ. A Abrikosov showed in 1957 that if, as in conventional superconductors (type I), the penetration depth is shorter than the coherence length ξ, then it is energetically unfavorable for the sample to break up into magnetic domains with normal regions penetrating into the bulk of the superconductor. However, for superconductors with $\xi < \lambda$ (type II) such as the high critical temperature cuprate superconductors, the opposite is true. The crossover between the two types of behavior occurs at a Ginzburg–London parameter $k = \lambda/\xi = 1/\sqrt{2}$.

Consider a hole or normal region through a type I superconductor which is cooled below T_c in the presence of a magnetic field that is less than H_c. Due to the Meissner effect, the magnetic flux Φ trapped in the hole can penetrate into the superconductor only over a distance of the London penetration depth λ at the inner surface of the hole. There is a supercurrent (with vector current density $\boldsymbol{J}_s$) induced in this layer to support the flux threading through the hole. The flow of supercurrent can be described in terms of the spatial and temporal variations of the single macroscopic pair wave function $\Psi(\boldsymbol{r}, t)$ throughout the entire superfluid condensate. This leads to the quantum mechanical expression for $\boldsymbol{J}_s$:

$$\boldsymbol{J}_s = \frac{q^*}{m^*} n_s(\boldsymbol{r}, t)(\hbar \boldsymbol{\nabla} \varphi(\boldsymbol{r}, t) - q^* \boldsymbol{A}(\boldsymbol{r}, t)) \qquad [4]$$

where $n_s(\boldsymbol{r}, t) = |\Psi(\boldsymbol{r}, t)|^2$ is the local pair density, $\varphi(\boldsymbol{r}, t)$ is the local phase, $(\boldsymbol{A}(\boldsymbol{r}, t)$ is the magnetic vector potential, and q^* and m^* are the charge and mass of the charge carriers, respectively.

Although the absolute phase φ of the wave function is not gauge-invariant, the phase change along a path in space is well-defined. The phase difference $\delta\varphi$ can be obtained by integrating $\boldsymbol{J}_s$ around a closed contour such as C or C' in **Figure 1**:

$$\begin{aligned} \delta\varphi &= \oint \delta\varphi(\boldsymbol{r}, t) \cdot \mathrm{d}\boldsymbol{l} \\ &= \frac{m^*}{\hbar n_s q^*} \oint \boldsymbol{J}_s \cdot \mathrm{d}\boldsymbol{l} + \frac{q^*}{\hbar} \int_s \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} \end{aligned} \qquad [5]$$

where $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ is the magnetic flux density, and S is the area enclosed by the contour. As a consequence of

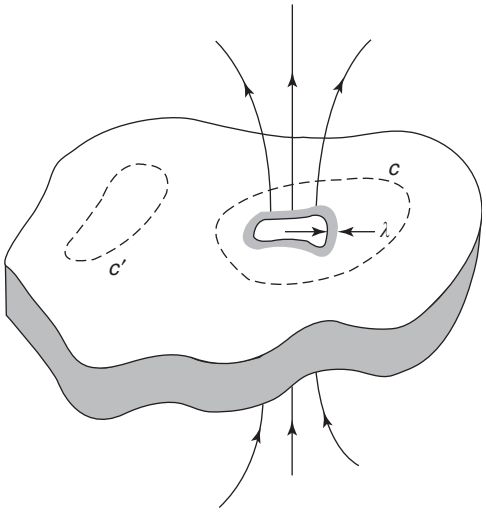


Figure 1 Schematic of magnetic flux threading through the interior of a superconductor.

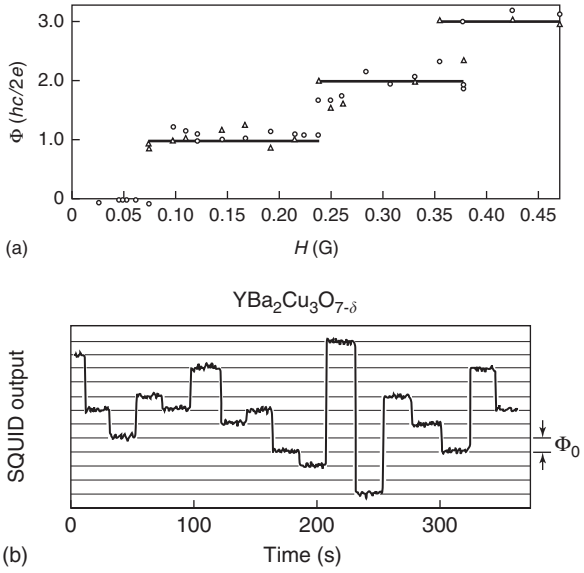


Figure 2 (a) Magnetic flux threading through a narrow superconducting tube as a function of cooling field. (Reproduced from Deaver BS Jr. and Fairbank WM (1961) *Physical Review Letters* 7: 43.) (b) Magnetic flux threading a polycrystalline $YBa_2Cu_3O_{7-\delta}$ YBCO ring as a function of time. (Reproduced with permission from Gough CE, *et al.* (1987) Quantization in a high-T_c superconductors. *Nature* 326: 855; © Nature Publishing Group.)

the single-valuedness of the pair wave function $\Psi(\boldsymbol{r}, t)$ the phase $\varphi(\boldsymbol{r}, t)$ around a closed loop can vary only by a multiple of 2π (i.e., $\delta\varphi = 2n\pi$, $n = 0, 1, 2, \ldots$). One can rewrite the above equation as follows:

$$\Phi' = \frac{4m^*}{n_s q^{*2}} \oint \boldsymbol{J}_s \cdot \mathrm{d}\boldsymbol{l} + \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = n\left(\frac{h}{q^*}\right) \qquad [6]$$

The quantity Φ' is called the fluxoid, first introduced by F London. The second term in the equation is the magnetic flux threading the hole (or normal region) and the penetration layer around it. Strictly speaking, it is the fluxoid contained in a loop so as C in **Figure 1** that is quantized in integral units of (h/q^*). However, if the dimensions of the superconductor are such that an integration contour such as C can be chosen such that $\boldsymbol{J}_s$ is always zero, then there is no difference between fluxoid and flux (i.e., $\Phi' = \Phi$). On the other hand, $\boldsymbol{J}_s$ is always zero along an integration path in a simply connected region (such as C' in **Figure 1**) and it encloses no flux. This leads to $n = 0$, a restatement of the Meissner effect.

Flux quantization was demonstrated independently by B S Deaver and W M Fairbank, and R Doll and M Näbauer in 1961. In these experiments, the flux trapped in superconducting hollow cylinders (with wall thickness much larger than the London penetration depth) was found to be quantized only in multiples of the flux quantum $\Phi_0 = h/2e = 2.07 \times 10^{-15}$ Wb $= 20.7$ G μm^2 (see **Figure 2a**). This pioneering work has provided compelling evidence for electron pairing in superconductors. Moreover, the concept of quantum phase coherence is validated eloquently through some simple classical magnetization measurements. Flux quantization in a ring made of a polycrystalline sample of the cuprate superconductor $YBa_2Cu_3O_{7-\delta}$ (YBCO) was demonstrated in 1987, shortly after the discovery of high-temperature superconductivity by Bednorz and Mueller. The magnetic flux state of the ring was monitored by using a superconducting quantum interference device (SQUID). As shown in **Figure 2b**, flux jumps, induced by a local source of electromagnetic noise, were found to occur only in integral multiples of the flux quantum Φ_0, thus providing strong evidence for Cooper pairing of charge $2e$ in HTS, as in conventional superconductors. This establishes the existence of macroscopic phase coherence in this very short coherence length ($\xi_0 \sim 1$ nm) superconductor, since the observed flux quantization is supported by a supercurrent flowing through the weak links between thousands of individual crystals around the ring. Since then, flux quantization has been observed many times in various cuprate systems.

Vortex State in Type II Superconductors

In Type II ($\kappa = \lambda/\xi > 1/\sqrt{2}$) superconductors it is energetically favorable for fields to enter in the form of flux lines when the applied field exceeds a critical value H_{c1}. Each of such flux lines, with screening currents circulating around it, is called a "vortex." One can apply the flux quantization condition previously described to show that each vortex carries exactly one flux quantum Φ_0. At the center of the

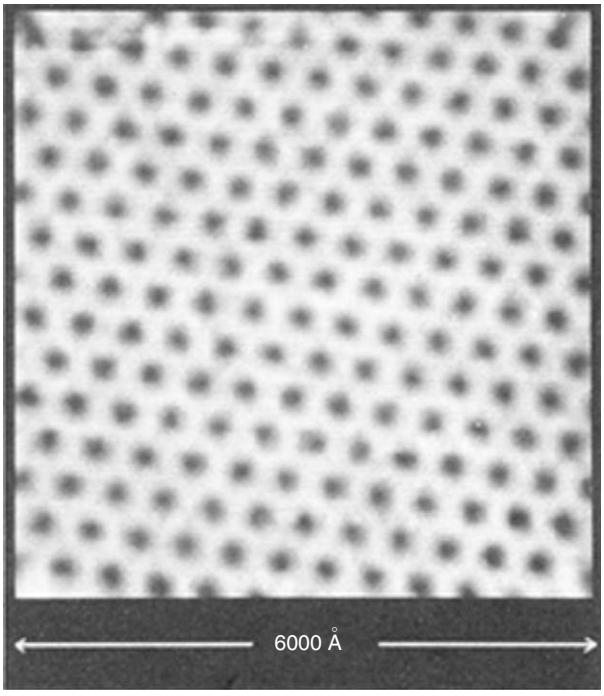


Figure 3 Scanning tunneling microscope image of a triangular lattice of vortices in $NbSe_2$. (Reproduced from Hess HF *et al.* (1989) *Physical Review Letters* 62: 214–216.)

vortex is the "vortex core," which is a region in which the superconductivity is suppressed of radius $\sim\xi$. Extending from the vortex center a distance of $\sim\lambda$ are supercurrents that circulate perpendicular to the vortex line, producing magnetic fields parallel to this line. As the applied magnetic fields increase, the vortices interact with each other more strongly. In a type II bulk superconductor, this interaction is repulsive. A Abrikosov showed in 1957 that this leads to a regular lattice of vortices. The different forms this lattice takes can vary: for conventional superconductors it is triangular. The existence of such flux line lattices have been observed and studied by a variety of techniques, including Bitter ferromagnetic powder decoration, small-angle neutron scattering (SANS), and scanning tunneling microscope (STM). As an example, an STM image of the triangular vortex lattice in $NbSe_2$ is shown in **Figure 3**.

The HTS superconductors have a large Ginzburg–London parameter $\kappa\approx 100$ and are therefore extreme type-II superconductors. The vortex state in HTS is significantly different from that of the conventional superconductors, both in their magnetic and electronic structures within the vortex core itself, and in the interaction between vortices. This leads to different vortex lattice structures and vortex dynamics. There is a vast literature on the study of the vortex state in conventional and high-temperature superconductors (see "Further reading" section). There has been a special emphasis on the study of vortex dynamics since dissipation in type II superconductors results from the motion of vortices. Recent progress in this area of research has been important for technological development, or electronic and large-scale applications of superconductivity.

Mesoscopic Effects

In type II ($\kappa=\lambda/\xi>1/\sqrt{2}$) superconductors, the magnetic fields and supercurrents of vortices spread upon reaching the surface, with a characteristic length given by the penetration depth λ. Vortices trapped in superconductors with a thickness $d\ll\lambda$ are called "Pearl vortices," after J Pearl, who first described them theoretically in 1966. These vortices have magnetic fields at short distances that diverge as r^{-1}, as opposed to $ln(r)$ for bulk vortices. They also have different vortex–vortex interactions. Vortices in highly anisotropic layered superconductors, such as the cuprates, can be described as strings of "pancake vortices," each pancake localized in a plane and interacting only weakly with its neighbors above and below in the vortex string. Vortices trapped in an interface between superconductors coupled by Josephson tunneling of Cooper pairs are called Josephson vortices. Josephson vortices can be highly anisotropic, with very strong screening and short penetration depths perpendicular to the interface, but weak screening and long penetration depths parallel to the interface.

Vortices do not normally occur in type I ($\kappa\equiv\lambda/\xi<1/\sqrt{2}$) superconductors. However, they can occur in very thin films of type I superconductors. "Giant" vortices with integer multiples of the fluxoid number can occur in superconductors with dimensions smaller as compared to the coherence length ξ. In addition, the fluxoids trapped in such samples can arrange in regular patterns to minimize their free energy. These patterns have been imaged using scanning Hall bar microscopy, and are in good agreement with predictions using the Ginzburg–Landau equations.

Spontaneous Vortex Generation

It has been proposed that phase transitions in condensed matter systems can be studied to provide clues to analogous phase transitions in the early development of the universe. One such phase transition is the normal-superconducting transition. Vortices can be generated in superconducting samples during the superconducting normal-phase transition, even in the absence of an applied magnetic field, if the cooling rate is sufficiently fast. Studies have been made of this spontaneous vortex generation in bulk

superconductors, and in rings of superconductors with and without Josephson weak links. This vortex generation is believed to result from a "freeze-out" of fluctuations in phase of the order parameter. This freeze-out can be the result of a slowing down of the communication between different sections of the superconductor or from the raising of the barrier to vortex formation as the temperature is lowered.

Half-Integer Flux Quantum Effect

The phenomena of integer flux quantization described so far are based on the requirement that the phase of a BCS pair wave function can change only by multiples of 2π when going around a closed superconducting circuit. However, a phase shift of π can be incorporated in a superconducting loop containing one or more Josephson junctions between unconventional superconductors. In a conventional (*s*-wave) superconductor, the pair wave function, which is proportional to the energy gap, always maintains the same sign, and has the full symmetry of the underlying crystal lattice. In sharp contrast, in unconventional superconductors, the wave-vector-dependent gap parameter $\Delta(\boldsymbol{k})$ is highly direction-dependent, with a symmetry lower than that of its *s*-wave counterpart. For example, it is believed that the cuprate superconductors have two-dimensional *d*-wave pair wave function which varies as $k_x^2 - k_y^2$. This wave function changes its sign across the node lines $\boldsymbol{k}_x = \pm \boldsymbol{k}_y$, where the amplitude of the pair wave function vanishes (see the four-leaf clover schematic representation in **Figure 4a**).

Consider a superconducting ring of inductance L interrupted by Josephson junctions between superconductors i and j with a critical current I_c^{ij}. The supercurrent circulating in the ring is governed by the Josephson equation for pair tunneling across the junction ij:

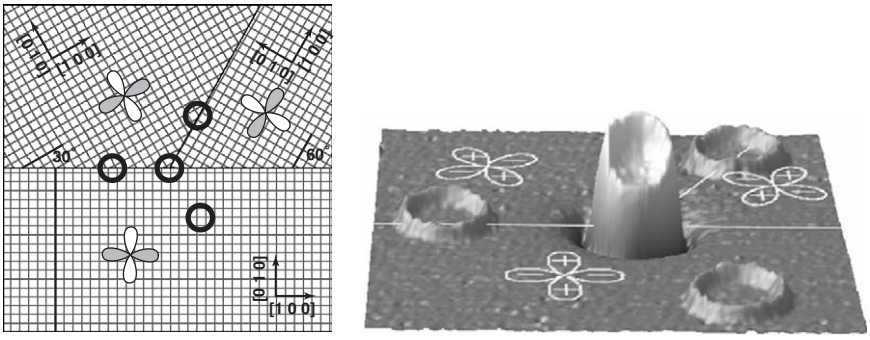
$$I_s = I_c^{ij} \sin(\delta\varphi_{ij}) \qquad [7]$$

where φ_{ij} is the phase difference between the macroscopic wave functions of the junction electrodes i and j. Since the critical current I_c^{ij} is proportional to the product of the energy gaps, or equivalently the pair wave functions, projected onto the normal to the junction interface, a negative critical current $I_c^{ij} < 0$ is therefore possible for a properly configured junction made of at least one unconventional (e.g., *d*-wave) superconductor which has nodes ($\Delta(\boldsymbol{k}) = 0$) in the gap function. A junction with $I_c < 0$ contributes a phase-shift of π to the phase-shift sum around the ring. If, for example, $I_c^{12} = -|I_c^{12}|$, then $I_s = -|I_c^{12}| \sin(\delta\varphi_{12} + \pi)$. The integer flux quantization of the ring requires that the total of the phase differences around the rings sums to an integral multiple of 2π:

$$2\pi\left(\frac{\Phi_a}{\Phi_0} + \frac{I_s L}{\Phi_0}\right) + \sum_{ij} \delta\varphi_{ij} + \alpha = n \times 2\pi \qquad [8]$$

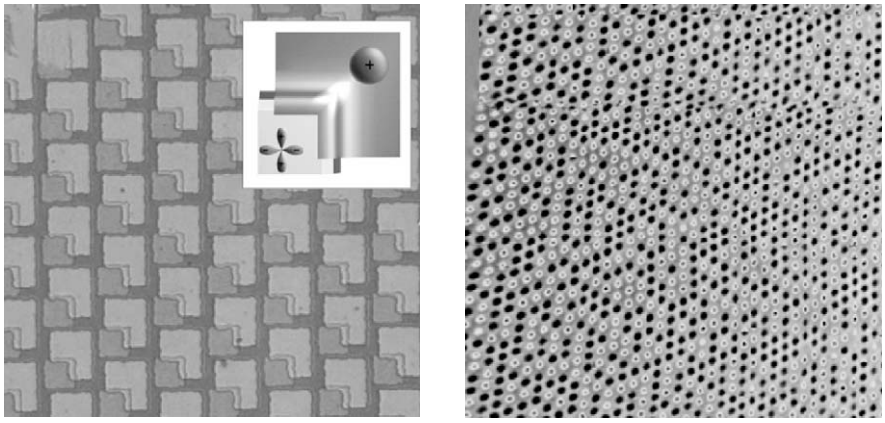
where Φ_a is the flux applied by an external field, $I_s L$ corresponds to the flux generated by the pair tunneling current circulating in the ring, and $\alpha = \pi$ for a ring with an odd number of sign changes in I_s (termed a π-ring); $\alpha = 0$ for a ring with an even number of sign changes (0-ring). The combined constraints imposed by eqns [7] and [8] lead to the conclusion that the ground state ($\Phi_a = 0$, $n = 0$) of a π-ring is characterized by a spontaneous magnetization of a half-flux quantum $\Phi = I_s L \approx (1/2)\Phi_0$, provided that $I_c^{ij} \gg \Phi_0$. A free energy consideration of the π-ring has reached the same conclusion.

The magnetic flux threading through a superconducting loop with an intrinsic π-phase shift (called a π-loop) is quantized in half-integral multiples of the



(a) (b)

Figure 4 (a) Tricrystal geometry for paring symmetry tests. (b) Scanning SQUID microscope image of spontaneously generated half-flux quantum in central ring in tricrystal experiments. (Reproduced from Tsuei CC and Kirtley JR (2000) *Review of Modern Physics* 72: 969–1016.)



(a) (b)

Figure 5 (a) Optical microscopy image of a triangular array of photolithographically patterned YBCO-Nb π-loops. (b) Scanning SQUID microscope image of such an array, cooled in zero field. Each π-loop has either a half-flux quantum with fields pointing out of the plane, or into the plane, of the array, corresponding to the time-reversed degenerate flux states described in eqn [9]. (Reproduced from Hilgenkamp H, *et al.* (2003) *Nature* 422: 50–53.)

flux quantum Φ_0:

$$\Phi = \pm\left(n + \frac{1}{2}\right)\Phi_0, \quad n = 0, 1, 2, \ldots \qquad [9]$$

where the $\pm$signs correspond to the screening supercurrents flowing respectively in the clockwise and counterclockwise directions. This half-integer flux quantum effect has been the basis for a new class of phase-sensitive tests for pairing symmetry in HTS. In particular, a series of tricrystal experiments have been used to establish the *d*-wave pairing symmetry in HTS. The basic design of the tricrystal experiments is shown in **Figure 4a**. The samples are made by depositing and patterning *c*-axis-oriented thin film rings of cuprate superconductors such as YBCO. The crystal orientations of the tricrystal substrates were chosen such that the three grain-boundary weak links in the ring encircling the tricrystal meeting point give rise to a net phase-shift of π, if the cuprate under test is a *d*-wave superconductor. **Figure 4b** shows a scanning SQUID microscope image, taken at 4.2 K and nominal zero field, of four YBCO rings on a tricrystal $SrTiO_3$ substrate with the configuration shown in **Figure 4a**. The fact that spontaneous magnetization of $\Phi_0/2$ is only observed in the 3-junction ring, but not in the 2- and 0-junction rings, represents strong evidence for *d*-wave pairing in YBCO. Additional tricrystal experiments have provided support for a *d*-wave pair state in various other HTS, including the electron-doped cuprates. Conventional techniques (e.g., angle-resolved photoemission spectroscopy, nuclear magnetic resonance, penetration depth measurement) have also produced important supporting evidence for *d*-wave pairing symmetry in HTS.

The establishment of *d*-wave pairing symmetry in HTS does not pin down a specific high-T_c mechanism. It does impose a well-defined constraint on possible models for high-temperature superconductivity. Recent research on the low-energy quasiparticle excitations near the nodes of the *d*-wave energy gap has led to the prediction and confirmation of new effects in low-temperature properties such as specific heat and thermal conductivity. These studies have produced strong evidence in support of *d*-wave pairing symmetry. Furthermore, they suggest that, in the context of the Fermi liquid formalism, BCS theory represents a valid description of the superconducting state in an unconventional superconductor.

Practical applications of *d*-wave superconductivity in areas such as π-SQUIDS and quantum computation, are still under active development. A recent work (see **Figure 5**) demonstrating fabrication of 25 000 π-loops on a single chip, and showing the presence of magnetic coupling between the individual π-loops in the array, represents an important step toward such potential applications.

See also: Superconductivity: Applications; Superconductivity: BCS Theory; Superconductivity: Critical Currents; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductivity: Tunneling; Superconductors, High T_c; Superconductors, Metallic.

PACS: 74.25.Ha; 74.25.Qt; 74.20.Rp

Further Reading

Blatter G, *et al.* (1994) *Reviews of Modern Physics* 66: 1125.
Huebener RP (1979) *Magnetic Flux Structures in Superconductors*, 2nd edn. New York: Springer.

Orland TP and Delin KA (1991) *Foundations of Applied Superconductivity*. Reading: Addison Wesley.
Timusk T and Statt B (1999) *Reports of Progress in Physics* 62: 61.
Tinkham M (1996) *Introduction to Superconductivity*, 2nd edn. New York: McGraw-Hill.
Tsuei CC and Kirtley JR (2000) Pairing symmetry in cuprate superconductors. *Reviews of Modern Physics* 72: 969–1016.

Superconductivity: General Aspects

A I Larkin, University of Minnesota, Minneapolis, MN, USA

Introduction

Superconductivity is a very exciting and unusual macroscopic quantum phenomenon. Different superconducting devices are continuing to find more and more applications around the world. There is no doubt that in the future, these applications will become even more widespread. In order to understand how these devices work, one should know the general aspects of superconductivity. A lot of detailed information on the subject can be found in the corresponding specialized articles of this encyclopedia. This article deals with only the main results of the Bardeen–Cooper–Schrieffer (BCS) theory and the Ginzburg–Landau theory as well as those aspects of superconductivity which are not mentioned in the other articles such as unconventional superconductivity, inhomogeneous states (also known as Larkin–Ovchinnikov–Fulde–Ferrell (LOFF) states), superconducting fluctuations, pinning of vortices. A special chapter is devoted to the manifestation of superconductivity in various nonmetallic systems.

Conventional Superconductors

A major success of low-temperature physics was achieved with the introduction of the notion of quasiparticles by Landau. According to the Landau Fermi-liquid theory, the properties of a many-body interacting system at low temperatures are determined by the spectrum of some low-energy, long-living excitations (called quasiparticles). Another milestone of the many-body theory is the mean field approximation (MFA). The BCS theory of superconductivity is a good example of the use of both the quasiparticle description and MFA. In 1956, L Cooper found that even a weak attraction between particles in a degenerated Fermi liquid led to the formation of bound states, now called Cooper pairs. Soon after this discovery Bardeen, Cooper, and Schrieffer proposed a microscopic theory of superconductivity as a theory describing Bose condensation of Cooper pairs. N N Bogolyubov succeeded in solving the problem of superconductivity by the method of approximate second quantization and L P Gor'kov proposed a solution of the problem in the framework of the Green functions formalism. This method permitted the use of the well known Feynman diagram technique from the quantum field theory in the theory of superconductivity. Using these methods, Gor'kov demonstrated that the phenomenological Ginzburg–Landau equations followed from the BCS theory in the limit $T \to T_c$ (i.e., in the vicinity of the transition). According to the BCS theory, the excitation spectrum $E(p)$ in a superconductor has a gap proportional to the amplitude of the Bose-condensate of Cooper pairs:

$$E(p) = \sqrt{[\varepsilon(p) - \mu]^2 + \Delta^2(p)}$$

where $\varepsilon(p)$ is the excitation spectrum in the normal state, $\Delta(p)$ is the superconducting gap, and μ is the chemical potential. In conventional superconductors, the gap has the same symmetry as the one of the lattices and in some systems does not depend on the momentum. In the latter case, both the gap and the transition temperature do not depend upon the density of nonmagnetic impurities (this statement is known as the Anderson theorem).

Unconventional Superconductors

In some systems, the effective electron–electron interaction is strongly repulsive at short distances but becomes attractive at large distances. In this case, an unconventional pairing can take place, with the gap having a symmetry group smaller than the one of the lattices: For example, in a tetragonal high-T_c superconductor, the angular dependence of the gap has the form

$$\Delta_d(p) \propto p_x^2 - p_y^2$$

In the isotropic case, the such a dependence corresponds to the orbital momentum $l = 2$ (d-wave superconductivity). In the anisotropic case, the momentum is not a good quantum number anymore.

The corresponding label just indicates the number of zeros of the function $\Delta(p)$. For example, the d-wave gap function has four zeros corresponding to the momenta $p_x = \pm p_y$. Nonmagnetic impurities destroy unconventional superconductivity if $k_B T_{c0} \cdot \tau_{tr} > \hbar$, where τ_{tr} is the transport scattering time and T_{c0} is the transition temperature in the pure system.

Inhomogeneous State

In the usual BCS state, electrons forming a Cooper pair have opposite momenta and the total momentum of a pair is zero. In ferromagnetic superconductors or layered superconductors in a strong magnetic field, the densities and Fermi-momenta of electrons with opposite spins are different. In such conditions, it is more favorable for each electron to stay close to its Fermi surface and for Cooper pairs to have nonzero total momenta, unlike in the usual BCS state. In such systems, the condensate wave function varies in space and the ground state is inhomogeneous, which may lead to the formation of a crystalline structure (LOFF state). The LOFF state is believed to exist in heavy fermion systems such as $CeCoIn_5$, $CeRu_2$, UPd_2Al_3, and in some organic compounds. The LOFF state also appears in superconductor–ferromagnet–superconductor junctions. In these systems, the wave function of a Cooper pair oscillates inside the ferromagnet and, for a certain thickness of the layer, may have opposite signs on different sides of the ferromagnetic layer. As a result, the Josephson energy of the contact changes sign and has two minima, corresponding to the phase difference $\pm\pi$. Such a device, called π-junction, has two degenerate states and can be used as a qubit for quantum computation.

Fluctuations (Paraconductivity)

Phenomena that cannot be described within the quasiparticle method or the MFA are called fluctuations. In bulk samples of traditional superconductors, the critical temperature T_c sharply divides the superconducting and normal phases. Such a behavior of the physical characteristics of superconductors is in perfect agreement with both the Ginzburg–Landau phenomenological theory and the BCS microscopic theory of superconductivity. Both of these theories can be derived in the framework of the mean field approximation. In the BCS theory, only the Cooper pairs in the Bose-condensate are considered. The fluctuation theory deals with Cooper pairs out of the condensate. Fluctuation phenomena manifest themselves much stronger in disordered low-dimensional systems. This is because the fluctuation region in disordered superconducting films, which is determined by the resistance per unit square, is typically much wider than in bulk samples. What is even more important is that fluctuation effects exist beyond the critical region and affect not only thermodynamic quantities but also the kinetic ones. The phenomenon, which is now known as paraconductivity, is the decrease of the resistance of a superconducting sample above the transition temperature (i.e., in the normal phase) due to the appearance of fluctuating Cooper pairs. The fluctuation conductance g of a superconducting film reads:

$$g - g_n = \frac{e^2}{16\hbar} \frac{T_c}{T - T_c}$$

Note that in high-temperature, organic, amorphous, and low-dimensional superconducting systems being studied presently, the fluctuation effects strongly differ from those in traditional superconductors. The transition turns out to be much more smeared out. The appearance of superconducting fluctuations above the critical temperature leads to precursor effects of the superconducting phase occurring in the normal phase, sometimes far from the transition. The conductivity, heat capacity, diamagnetic susceptibility, sound attenuation, and other properties may change considerably in the vicinity of the transition. A strong dependence of superconducting fluctuations on the temperature and magnetic field permits one to definitely separate the fluctuation effects from other contributions and to use them as a source of information about the microscopic parameters of a material. The account for fluctuations has become a necessary part in the design of superconducting devices.

Pinning of Vortices

The greatest success of the Ginzburg–Landau theory was the explanation of Shubnikov's phase by A A Abrikosov in 1957. In this phase, superconductivity and a magnetic field can peacefully coexist. Abrikosov suggested that the magnetic field can penetrate a superconductor along vortex lines, which can form a perfect lattice structure. In homogeneous superconductors, an applied current causes a drift of the vortex lattice and dissipation. When a transport current j_{tr} flows through a superconductor, the Lorentz force $F_L = (\Phi_0/c) j_{tr}$ exerted on each vortex line appears immediately ($\Phi_0 = hc/2e$ is the flux quantum). In homogeneous superconductors, the vortices start moving due to this force. The corresponding flow of the magnetic flux induces the electric field $E = (v/c) \cdot B$ and energy dissipation. In order to preserve

superconductivity, one should prevent the motion of vortices by introducing a dry friction force. This can be done by creating local inhomogeneous regions (pinning centers), which would pin the vortex lattice. Such centers may appear due to structural inhomogeneities of the initial crystalline lattice (dislocations, accumulations of impurities, etc.). In the theory of collective pinning, even small structural fluctuations can qualitatively change the properties of the vortex structure. A single weak center causes only a weak elastic deformation, but the collective effect of a large number of weak centers destroys the lattice. The long-range lattice order disappears, while the short-range order survives only at distances smaller than a correlation length L_c. Each region of size L_c finds its equilibrium in a local minimum of the random pinning potential. A force is needed in order to move the pinned regions from the energetically favorable positions. The appearance of the frictional force leads to a finite critical current and hysteresis.

Superconductivity in Nonmetallic Systems

Organic superconductors

W A Little has proposed the possibility of a nonphonon pairing mechanism in some organic superconductors. Little's work has stimulated a lot of further theoretical activity in the field of one-dimensional systems. All real organic superconductors are not truly one-dimensional but consist of long chains strongly coupled to each other, which make the system effectively two-dimensional and strongly anisotropic. This may result in an unconventional superconducting pairing and in the appearance of inhomogeneous LOFF states in such systems. High upper critical fields in some organic materials (e.g., $(TMTSF)_2ClO_4$ or $(TMTSF)_2PF_6$) support the scenario of the unconventional triplet pairing. An experimental evidence of the LOFF state has been found in the organic compounds $k-(BEDT-TTF)_2-Cu(NCS)_2$ and $\lambda-(BETS)_2FeCl_4$.

Ceramic Superconductors

Most of the high-T_c superconductors are ceramic superconductors.

Superfluidity of helium Superfluidity in He_4 is a result of Bose condensation of He_4 atoms themselves. In He_3, the atoms are spin-1/2 particles and the normal state of He_3 is thought to be a Fermi-liquid. Superfluidity in He_3 is the result of Cooper pairing in the fermion system. Cooper pairs in He_3 have a total spin $S=1$ and orbital moment $l=1$, which corresponds to the p-wave unconventional superconducting pairing.

Superfluid model of atomic nuclei Strong interactions between nucleons in heavy nuclei may yield the formation of Cooper pairs. This leads to the gap in the spectrum of excitations in the nuclei, which typically is of the order of $\Delta \sim 2\,\text{MeV}$, which is much larger than the level spacing. The appearance of the nucleon Bose-condensate leads to parity effects in atomic nuclei. The energy and mass of a nucleus containing an even number of nucleons are smaller than the ones of an odd nucleus:

$$2M_{2N+1} - M_{2N} - M_{2N+2} = 2\Delta$$

A rotating nucleus creates supercurrents, which produce an effect similar to the anomalous diamagnetism in conventional superconductors. As a result, the moment of inertia of a nucleus is smaller than the one of a solid or liquid object with the same mass and geometry.

Superconductivity in neutron stars In neutron stars (pulsars), the density of neutrons may be of the order of their density in heavy atomic nuclei. As a result, Cooper pairs are created and superconductivity can arise. Since pulsars are spinning, the neutron superfluid is threaded with a regular array of rotational vortices. The rotation frequency of the superfluid is proportional to the density of the vortices. As a pulsar's spin rate gradually decreases due to the emission of electromagnetic radiation, the vortices gradually move outwards.

Color superconductivity In the core of a neutron star, the density is so high that quark matter can appear. Strongly interacting quarks may form Cooper pairs (it is predicted that the red up-quarks are paired with the green down-ones) yielding the so-called color superconductivity. Since the quarks forming pairs have different charges and slightly different densities, the Fermi surface splits into two and an inhomogeneous crystalline state similar to the LOFF state appears in a shell, where the quark densities satisfy the appropriate conditions. The rotational vortices may be pinned by the LOFF structure leading to the pinning effects. Within this picture, the so-called glitch phenomena (sudden increases of rotation frequencies of a pulsar) are explained as the hopping of vortex bundles from one metastable minimum to another.

Superconductivity and masses of elementary particles Papers of Nambu and Jona-Lasinio and Vaks

and Larkin have introduced the mechanism of elementary particle mass generation via the dynamical symmetry breaking. This spontaneous symmetry breaking is formally very similar to the one occurring in superconductors, with the superconducting gap corresponding to the masses of the elementary particles. In superconductors and in the models considered by Nambu and Jona-Losinio and by Vaks and Larkin, the transition occurs due to the appearance of the Bose-condensate of Cooper pairs, while in the standard model the symmetry breaking is due to the Bose-condensate of the scalar Higgs bosons. Furthermore, the Meissner effect, which is characterized by a penetration length, is the origin, in the elementary particle physics language, of the masses of the gauge vector bosons. The masses correspond to the inverse of the penetration length in the conventional theory of superconductivity.

Acknowledgment

This work was supported by NSF grant DMR0120702.

See also: Superconductivity: BCS Theory; Superconductivity: Critical Currents; Superconductivity: Flux Quantization; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductors, High T_c; Superfluidity.

PACS: 74.20.−z; 74.40.+k; 74.70. Hk; 74.90.+n; 74.60.Ge

Further Reading

Abrikosov AA (1998) *Fundamentals of the Theory of Metals*. Groningen: North-Holland/Elsevier.

Barone A and Paterno GF (1982) *Physics and applications of Josephson effect*. New York: Wiley-Interscience.

Bennemann KH and Ketterson JB (eds.) (2002) *The Physics of Superconductors*. Berlin: Springer.

Blatter G, Feigel'man MV, Geshkenbein VB, Larkin AI, and Vinokur VM (1994) *Reviews of Modern Physics* 66: 1180.

Ishiguro T, Yamaji K, and Saito G (1998) *Organic Superconductors*. Heidelberg: Springer.

Langenberg DN and Larkin AI (eds.) (1986) *Nonequilibrium Superconductivity*. Amsterdam: Elsevier.

Migdal AB (1967) *Theory of Finite Fermi System and Applications to Atomic Nuclei*. New York: Interscience.

Mineev VM and Samokhin KV (1999) *Introduction to Unconventional Superconductivity*. Amsterdam: Gordon and Breach.

Tinkham M (1996) *Introduction to Superconductivity*, 2nd edn. New York: McGraw-Hill.

Nomenclature

e	electron charge
g	conductance
T_c	superconducting transition temperature
Δ	superconducting gap
μ	chemical potential

Superconductivity: Ginzburg–Landau Theory and Vortex Lattice

E H Brandt, Max-Planck-Institut für Metallforschung, Stuttgart, Germany

Introduction

After the discovery of superconductivity in 1912 by Heike Kamerlingh-Onnes in Leiden, it took almost 50 years until this fascinating phenomenon was understood microscopically, when in 1957 Bardeen, Cooper, and Schrieffer established their BCS theory. But long before this electron-pairing theory, there were powerful phenomenological theories which were able to explain most electromagnetic and thermodynamic observations on superconductors, and which are very useful today also. The London theory, conceived by Fritz and Heinz London in 1935, is particularly useful for the description of the high-T_c superconductors, which were discovered in 1987 by Bednorz and Müller in Zürich. Perhaps the most useful and elegant phenomenological theory of superconductors was established by Vitalii Ginzburg and Lew Landau in 1951. The Ginzburg–Landau (GL) theory generalizes Landau's theory of second-order phase transitions to spatially varying charged systems in a magnetic field and should thus be applicable near the superconducting transition temperature T_c, but in many cases it yields a qualitatively correct behavior also at lower temperatures. The GL theory can be derived from the microscopic BCS theory, and it reduces to the London theory in situations when the GL function (the superconducting order parameter) has nearly constant magnitude.

The GL theory has high predictive power. It predicts that superconductors can be of type-I (with positive energy of the wall between normal conducting and superconducting domains) or of type-II (with negative wall energy, pointing to an instability) and that superconductivity can be suppressed by a high magnetic field and by a high current density, namely,

by the depairing current which breaks the Cooper pairs. Its most spectacular success was the prediction of spontaneous nucleation and penetration of magnetic vortices in type-II superconductors by Alexei Abrikosov in 1957. The properties of these Abrikosov vortices (or fluxons, flux lines, carrying one quantum of magnetic flux $\Phi_0 = h/2e = 2.07 \times 10^{-15}$ T m^2) and their motion and pinning by material inhomogeneities, are important topics both in research and in the practical application of superconductors. Most applications require type-II superconductors with high critical magnetic fields and currents, but high loss-free supercurrents require that the vortices are pinned so that they cannot move under the action of this current and dissipate energy.

Ginzburg–Landau Theory

The GL theory introduces a complex, spatially varying order parameter $\psi(\boldsymbol{r})$ in addition to the magnetic field $\boldsymbol{B}(\boldsymbol{r}) = \nabla \times \boldsymbol{A}(\boldsymbol{r})$ or vector potential $\boldsymbol{A}$. The GL function $\psi(\boldsymbol{r})$ later turned out to be proportional to the BCS energy-gap function $\Delta(\boldsymbol{r})$, and its square $|\psi(\boldsymbol{r})|^2$ to the density of Cooper pairs. It defines two characteristic lengths: the superconducting coherence length ξ sets the scale over which $\psi(\boldsymbol{r})$ can vary, while λ governs the variation of the magnetic field as in the London theory. Both λ and ξ diverge at the superconducting transition temperature T_c according to $\lambda \propto \xi \propto (T_c - T)^{-1/2}$, but their ratio, the GL parameter $\kappa = \lambda/\xi$, is nearly independent of the temperature T. The GL theory reduces to the London theory (which is valid down to $T = 0$) in the limit $\xi \ll \lambda$, which means constant magnitude $|\psi(\boldsymbol{r})| =$ const (except in the vortex cores, where ψ vanishes, see below). The GL equations are obtained by minimizing a free-energy functional $F\{\psi, \boldsymbol{A}\}$ with respect to the GL function $\psi(\boldsymbol{r})$ and the vector potential $\boldsymbol{A}(\boldsymbol{r})$. With the length unit λ and magnetic field unit $\sqrt{2}B_c (B_c = \Phi_0/(\sqrt{8}\pi\lambda\xi)$ is the thermodynamic critical field), the GL functional reads

$$F\{\psi, \boldsymbol{A}\} = \frac{B_c^2}{\mu_0} \int_V \left[-|\psi|^2 + \frac{1}{2}|\psi|^4 + \left| \left(-\frac{i\nabla}{\kappa} - \boldsymbol{A} \right) \psi \right|^2 + (\nabla \times \boldsymbol{A})^2 \right] \mathrm{d}^3 r \quad [1]$$

Here the integral is over the volume V of the superconductor. The first two terms are the superconducting condensation energy, which is minimum when $|\psi|^2 = 1$. The second term is the energy cost of the spatial variation of ψ; this gauge-invariant gradient also introduces $\boldsymbol{A}$ and thus the magnetic field. The last term is the magnetic energy density $B^2/2\mu_0$.

The GL functional and the resulting GL equations may be expressed in terms of the real-valued function $|\psi|$ and the gauge-invariant supervelocity $\nabla\varphi/\kappa - \boldsymbol{A}$, where $\varphi(\boldsymbol{r})$ is the phase of $\psi = |\psi|\exp(i\varphi)$. From the variation $\delta F/\delta\psi = 0$ follows the first GL equation determining the amplitude and phase of ψ. From the variation $\delta F/\delta\boldsymbol{A} = 0$ follows the supercurrent density $\boldsymbol{J} = \mu_0^{-1}\lambda^2[(\Phi_0/2\pi)\nabla\varphi - \boldsymbol{A}]|\psi|^2$, which has to be supplemented by the Maxwell equation $\boldsymbol{J} = \mu_0^{-1}$ curl curl $\boldsymbol{A}$ to obtain an equation for $\boldsymbol{A}$ or $\boldsymbol{J}$ alone. The minimization of F yields not only the functions $\psi(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$, but also the boundary condition that the current density and supervelocity do not have a component perpendicular to the free surface of the superconductor.

In an external field H_a, one has to minimize not the free energy F but the Gibbs free energy $G = F - \bar{B}H_a$, where $\bar{B}$ is the spatial average of the magnetic induction B in the superconductor. The condition $\partial G/\partial\bar{B} = 0$ yields the equilibrium field $H_a = \partial F/\partial\bar{B}$. In sufficiently low H_a, the superconductor expels the applied magnetic field completely, that is, one has $B = 0$ inside the superconductor (Meissner state). More precisely, the parallel applied field penetrates into a thin surface layer of thickness λ. In a superconductor filling the half space $x \geq 0$, one has $B(x) = \mu_0 H_a \exp(-x/\lambda)$. This screening of H_a is caused by a surface current of density $j(x) = (H_a/\lambda) \exp(-x/\lambda)$ flowing perpendicular to H_a. Such an exponential screening, strictly spoken, occurs only for London superconductors with large GL-parameter $\kappa \gg 1$, while for smaller κ values both $|\psi|$ and B vary near the surface and have to be obtained by solving the GL equations. Moreover, for superconductors of realistic shape, such as short cylinders or rectangular plates, the penetration of a magnetic field into the surface layer has to be calculated numerically; only for spheres and long cylinders with $\kappa \gg 1$, analytical solutions were obtained by London.

Critical Fields

In type-I superconductors (defined by $\kappa < 1/\sqrt{2}$), complete screening occurs when the applied magnetic field H_a is less than the thermodynamic critical field $H_c = \mu_0^{-1}B_c$; in larger fields $H_a \geq H_c$, these superconductors are normal conducting. For type-II superconductors (defined by $\kappa \geq 1/\sqrt{2}$), full screening occurs up to the lower critical field $H_{c1} = \mu_0^{-1}B_{c1}$ where penetration of Abrikosov vortices begins. With a further increasing of H_a, more vortices penetrate in the form of a more or less regular flux-line lattice (**Figure 1**), which has an average induction $\bar{B} = n\Phi_0$, where n is the area-density of vortices and $\Phi_0 = 2.07 \times 10^{-15}$ T m^2 is the quantum of magnetic

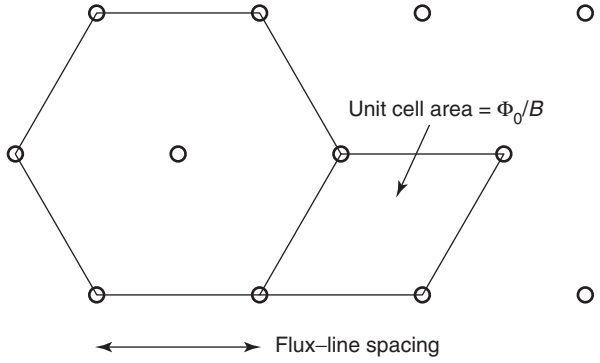


Figure 1 The triangular vortex lattice. The dots mark the positions of the vortex cores.

flux. When H_a reaches the upper critical field $H_{c2} = \mu_0^{-1} B_{c2}$, the vortex cores overlap so strongly that the superconductor turns normal and one has $\bar{B} = \mu_0 H_a \geqslant B_{c2}$. The GL theory yields

$$B_{c1} = \frac{\Phi_0(\ln\kappa + \alpha)}{4\pi\lambda^2}, \quad B_c = \frac{\Phi_0}{\sqrt{8}\pi\lambda\xi}, \quad B_{c2} = \frac{\Phi_0}{2\pi\xi^2} \quad [2]$$

with $\alpha(\kappa) \approx 0.5 + (1 + \ln 2)/(2\kappa - \sqrt{2} + 2)$. At $\kappa = 1/\sqrt{2}$, one has exactly $B_{c1} = B_c = B_{c2}$, and for type-II superconductors $B_{c1} \leqslant B_c \leqslant B_{c2}$. In both types of superconductors, B_c determines the area under the magnetization curve $\mu_0 M(H_a)$; this area equals the superconducting condensation energy B_c^2/μ_0. For ideal pin-free superconductors, the magnetization is $M = \mu_0^{-1}\bar{B} - H_a \leqslant 0$, see **Figure 2**. For type-II superconductors with an ideal planar surface, H_c is also the field up to which the so-called Bean–Livingston surface barrier may prevent a vortex penetration ("overheating"). Interestingly, as discovered theoretically by DeGennes in 1963, even above B_{c2}, a thin surface sheath of thickness ξ remains superconducting up to a third critical field, $B_{c3} = 1.695 B_{c2}$.

The above scenario applies to long superconductors in the parallel magnetic field H_a. In realistic samples like spheres, platelets, or films, demagnetization effects lead to flux penetration at lower fields. In ideal ellipsoids, the penetration fields, H_c or H_{c1}, are reduced by a factor $(1 - N)$, where N is the demagnetization factor. For spheres, one has $N = 1/3$, and for long cylinders, $N = 0$ in a parallel field and $N = 1/2$ in a perpendicular field, while for thin plates in a perpendicular field, one has $1 - N \ll 1$, and thus very small fields will force flux penetration from the edges.

In type-I superconductors with $N > 0$, demagnetization effects lead to partial penetration of flux in the form of planar domains or more complicated structures, while in type-II superconductors they just shear the magnetization curve, but the vortex lattice

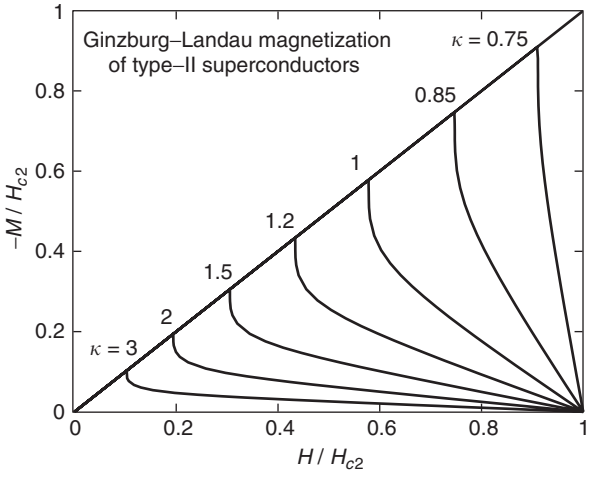


Figure 2 Magnetization curves $M(H)$ of long type-II superconductors in parallel magnetic field $H = H_a$ (demagnetization factor $N = 0$) calculated from the Ginzburg–Landau theory for various Ginzburg–Landau parameters $\kappa = 0.75 \ldots 3$. For $\kappa = 1/\sqrt{2}$, $-M$ jumps vertically from $-M = H$ to $M = 0$ at $H = H_{c1} = H_{c2}$.

remains uniform if the specimen is an ellipsoid. For other shapes such as platelets, strips, or disks with constant thickness, the demagnetization effects are more complicated and lead to a geometric barrier for flux penetration, with a penetration field which is approximately H_{c1} times the square root of the aspect ratio thickness/width. For specimens much larger than λ, this result can be derived by the continuum theory without considering individual vortex lines. But for small mesoscopic superconductors with size comparable to or smaller than λ, the full GL theory has to be solved numerically to see the detailed pattern of penetrated flux lines and the shape of the magnetization curve $M(H_a)$ that may exhibit jumps at certain values of H_a.

Vortex Lattice

In 1957 A A Abrikosov found a solution of the GL theory that exhibits a two-dimensional (2D) regular lattice of zeros in the order parameter $\psi(x,y)$ and a periodic magnetic field $B(x,y)$. This solution describes a lattice of parallel vortex lines along z. The ideal lattice is triangular, that is, each vortex has six nearest neighbors, see **Figure 1**. The vortices start penetrating at $H_a = H_{c1}$. The profiles $|\psi(r)|^2$ and $B(r)$ of one isolated vortex line are shown in **Figure 3** (r is the radial coordinate). As H_a is increased, more vortices penetrate and form a vortex lattice as shown in **Figure 4**. First, the magnetic fields of the flux lines overlap, such that the amplitude of the periodic $B(x,y)$ decreases. Then, the cores also overlap, such that the amplitude of the order parameter $|\psi|^2$ decreases until it vanishes at $H_a = H_{c2}$. For $b = \bar{B}/B_{c2} > 0.5$, good

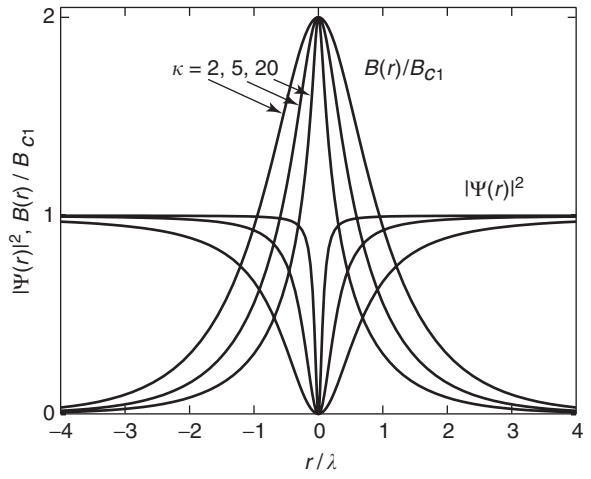


Figure 3 Magnetic field $B(r)$ and order parameter $|\psi(r)|^2$ of an isolated flux line calculated from the Ginzburg–Landau theory for Ginzburg–Landau parameters $\kappa = 2$, 5, and 20.

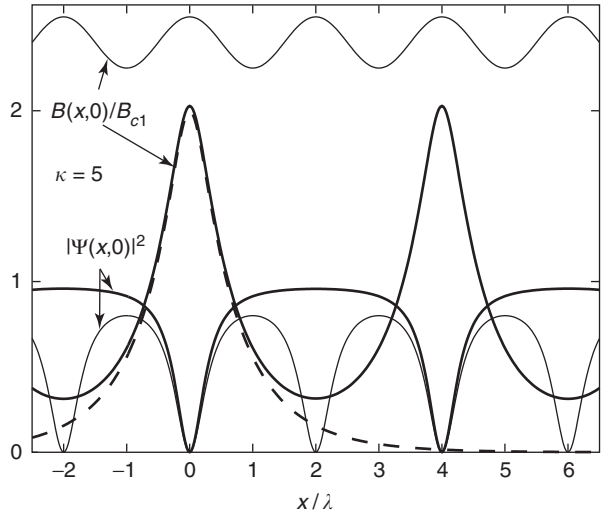


Figure 4 Two profiles of the magnetic field $B(x, y)$ and order parameter $|\Psi(x, y)|^2$ along the x-axis (a nearest-neighbor direction) for flux-line lattices with lattice spacings $a = 4\lambda$ (bold lines) and $a = 2\lambda$ (thin lines). The dashed line shows the magnetic field of an isolated flux line from **Figure 3**. Calculations from the Ginzburg–Landau theory for $\kappa = 5$.

approximations are

$$|\psi(\boldsymbol{r})|^2 = \frac{1 - \bar{B}/B_{c2}}{[1 - 1/(2\kappa^2)]\beta_A + 1} \sum_K a_K \cos \boldsymbol{K}\boldsymbol{r} \quad [3]$$

$$B(\boldsymbol{r}) = \mu_0 H_a - (\phi_0/4\pi\lambda^2)|\psi(\boldsymbol{r})|^2 \quad [4]$$

Here $\boldsymbol{r} = (x, y)$ and the sum is over all reciprocal lattice vectors $K_{mn} = (2\pi/x_1y_2)(my_2; -mx_2 + nx_1)$ of the periodic FLL with flux-line positions $R_{mn} = (mx_1 + nx_2; ny_2)$ $(m, n = \text{integer})$. The unit-cell area x_1y_2 yields the mean induction $\bar{B} = \phi_0/(x_1y_2)$. The first Brillouin zone has the area πk_{BZ}^2 with $k_{BZ}^2 = 4\pi\bar{B}/\phi_0$. For general lattice symmetry, the Fourier coefficients a_K and the Abrikosov parameter β_A are:

$$a_K = (-1)^{m+mn+n} \exp(-K_{mn}^2 x_1 y_2/8\pi) \quad [5]$$

$$\beta_A = \frac{\langle |\psi|^4 \rangle}{\langle |\psi|^2 \rangle^2} = \sum_K a_K^2 \quad [6]$$

From $\psi(\boldsymbol{r} = 0) = 0$ follows $\sum_K a_K = 0$, or with $a_{K=0} = 1$, $\sum_{K \neq 0} a_K = -1$. In particular, for the triangular vortex lattice with spacing $a = (2\phi_0/\sqrt{3}\bar{B})^{1/2}$, one has $x_1 = a$, $x_2 = a/2$, $y_2 = \sqrt{3}a/2$, $\bar{B} = 2\phi_0/(\sqrt{3}a^2)$, $a_K = (-1)^p \exp(-\pi p/\sqrt{3})$ with $p = m^2 + mn + n^2 = R_{mn}^2/R_{10}^2 = K_{mn}^2/K_{10}^2$. This yields $\beta_A = 1.15960$ (the reciprocal-lattice sum converges very rapidly) and the useful relationships $K_{10} = 2\pi/y_2$, $K_{10}^2 = 16\pi^2/3a^2 = 8\pi^2\bar{B}/\sqrt{3}\phi_0$. For the square vortex lattice, one has $x_1 = y_2 = a$, $x_2 = 0$, $\bar{B} = \Phi_0/a^2$, and $\beta_A = 1.18034$. This means the energy of the square vortex lattice is only slightly larger than that of the triangular lattice, by 2% at most. A vortex lattice with square symmetry has been observed experimentally, for example, when the underlying square symmetry of the atomic lattice couples to the vortex lattice such that the square vortex lattice has lower energy, but also other effects may stabilize the square lattice, for example, an anisotropic Fermi surface or a deviation from the GL theory.

The free energy $F(\bar{B})$ per unit volume and the negative magnetization $M = \mu_0^{-1}H_a - \bar{B} \geqslant 0$ with $H_a = \partial F/\partial \bar{B}$ are (still for $\bar{B}/B_{c2} > 0.5$ and from the GL theory):

$$F(\bar{B}) = \frac{\bar{B}^2}{2\mu_0} - \frac{(B_{c2} - \bar{B})^2/2\mu_0}{(2\kappa^2 - 1)\beta_A + 1} \quad [7]$$

$$-\mu_0 M(\bar{B}) = \frac{B_{c2} - \bar{B}}{(2\kappa^2 - 1)\beta_A + 1} = \frac{\phi_0}{4\pi\lambda^2}\langle |\psi|^2 \rangle \quad [8]$$

For the periodic field $B(\boldsymbol{r}) = \sum_K B_K \cos \boldsymbol{K}\boldsymbol{r}$, the Fourier coefficients are $B_{K \neq 0} = Ma_K$, $B_{K=0} = \bar{B}$, and $\langle |\psi|^2 \rangle = (4\pi\lambda^2/\phi_0)|M|$.

At lower induction $\bar{B}/B_{c2} < 0.5$, the vortex lattice and its magnetization curve have to be computed from the GL theory numerically, see **Figures 2–4**. However, when $\kappa \gg 1$, one may also use the London theory, which is a good approximation for small inductions $\bar{B}/B_{c2} < 0.2$.

London Theory

The London theory follows from the GL theory by putting $|\psi| = 1$, but it may also be obtained by minimizing the sum F of the energy of the magnetic

field $B(\boldsymbol{r})$ and the kinetic energy of the supercurrent density $J(\boldsymbol{r}) = \mu_0^{-1}\nabla \times B(\boldsymbol{r})$. This yields the London energy functional

$$F = \frac{1}{2\mu_0}\int_V [B^2 + \lambda^2(\nabla \times B)^2]\mathrm{d}^3 r \qquad [9]$$

Minimizing this with respect to $B = \nabla \times \boldsymbol{A}$, one obtains the homogeneous London equation $B - \lambda^2\nabla^2 B = 0$ or $J = -\mu_0^{-1}\lambda^{-2}\boldsymbol{A}$, where the Maxwell equations $\nabla \cdot B = 0$ and $\nabla \times B = \mu_0 J$ were used and the vector potential $\boldsymbol{A}$ was chosen in the "London gauge," which requires that $\nabla \times \boldsymbol{A} = 0$ and $\boldsymbol{A}$ is parallel to the surface everywhere. In the presence of vortices, one has to add singularities that describe the vortex core, which may be straight or curved.

For straight parallel vortex lines along the unit vector $\hat{\boldsymbol{z}}$, one gets the modified London equation

$$B(\boldsymbol{r}) - \lambda^2\nabla^2 B(\boldsymbol{r}) = \hat{\boldsymbol{z}}\Phi_0 \sum_v \delta_2(\boldsymbol{r} - \boldsymbol{r}_v) \qquad [10]$$

Here $\boldsymbol{r}_v = (x_v, y_v)$ are the 2D vortex positions, which now do not have to form a periodic lattice, and $\delta_2(\boldsymbol{r}) = \delta(x)\delta(y)$ is the 2D delta function. This linear equation may be solved by Fourier transform using the relations $\int \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\mathrm{d}^2 k = 4\pi^2\delta_2(\boldsymbol{r})$ and $\int \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})(k^2 + \lambda^{-2})^{-1}\mathrm{d}^2 k = 2\pi K_0(|\boldsymbol{r}|/\lambda)$. Here, $K_0(x)$ is a modified Bessel function with the limits $K_0(x) \approx \ln(1.123/x)$ for $x \ll 1$ and $K_0(x) \approx (\pi/2x)^{1/2}\exp(-x)$ for $x \gg 1$. The resulting magnetic field of any arrangement of parallel vortices is the sum of individual London vortex fields centered at the positions $\boldsymbol{r}_\mathrm{v}$,

$$B(\boldsymbol{r}) = \hat{\boldsymbol{z}}\frac{\Phi_0}{2\pi\lambda^2}\sum_v K_0\left(\frac{|\boldsymbol{r} - \boldsymbol{r}_v|}{\lambda}\right) \qquad [11]$$

The energy $F_{2\mathrm{D}}$ of this 2D arrangement of vortex lines with length L (the specimen height) is obtained by inserting eqn [10] into [9]. Integrating over the delta function, one finds that the London energy is determined by the magnetic field values at the vortex positions,

$$\begin{aligned} F_{2\mathrm{D}} &= L\frac{\Phi_0}{2\mu_0}\sum_\mu B(\boldsymbol{r}_\mu) \\ &= L\frac{\Phi_0^2}{4\pi\mu_0\lambda^2}\sum_\mu\sum_v K_0\left(\frac{|\boldsymbol{r}_\mu - \boldsymbol{r}_v|}{\lambda}\right) \end{aligned} \qquad [12]$$

This expression shows that the energy is composed of the pairwise interaction energy (terms $\mu \neq v$) and the self-energy of the vortices (terms $\mu = v$). To avoid the divergence of the self-energy, one has to cut off the logarithmic infinity of B at the vortex centers $\boldsymbol{r}_v$ by introducing a finite radius of the vortex core of order ξ, the coherence length of the GL theory. This cutoff may be achieved by replacing, in eqns [11] and [12], the distance $r_{\mu v} = |\boldsymbol{r}_\mu - \boldsymbol{r}_v|$ by $\tilde{r}_{\mu v} = (r_{\mu v}^2 + 2\xi^2)^{1/2}$, and multiplying B by a normalization factor ≈ 1 to conserve the flux Φ_0 of the vortex. This analytical expression suggested by John Clem for a single vortex and later generalized to the vortex lattice, is an excellent approximation, as was shown numerically by solving the GL equation for the periodic FLL in the entire ranges of $\bar{B}$ and κ for $0 \leqslant \bar{B} \leqslant B_{\mathrm{c}2}$ and $\kappa \geqslant 1/\sqrt{2}$.

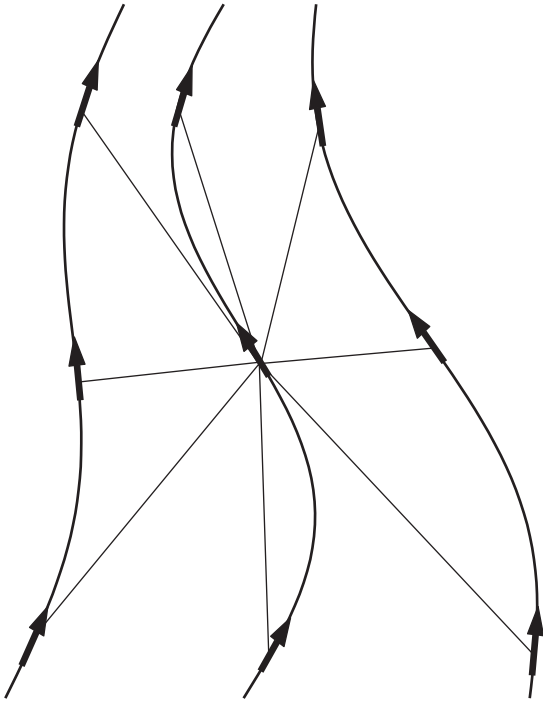
For curved vortices at arbitrary positions $\boldsymbol{r}_v(z) = [x_v(z), y_v(z), z]$, the 3D modified London equation reads

$$B(\boldsymbol{r}) - \lambda^2\nabla^2 B(\boldsymbol{r}) = \Phi_0 \sum_v \int \mathrm{d}\boldsymbol{r}_v\, \delta_3(\boldsymbol{r} - \boldsymbol{r}_v) \qquad [13]$$

Here, the integral is along the vortex lines and $\delta_3(\boldsymbol{r}) = \delta(x)\delta(y)\delta(z)$. The resulting magnetic field and energy are, with $\tilde{r}_{\mu v} = [|\boldsymbol{r}_\mu(z) - \boldsymbol{r}_v(z)|^2 + 2\xi^2]^{1/2}$,

$$B(\boldsymbol{r}) = \frac{\Phi_0}{4\pi\lambda^2}\sum_v \int \mathrm{d}\boldsymbol{r}_v \frac{\exp[-\tilde{r}_{\mu v}(\boldsymbol{r}_\mu = \boldsymbol{r})/\lambda]}{\tilde{r}_{\mu v}(\boldsymbol{r}_\mu = \boldsymbol{r})} \qquad [14]$$

$$\begin{aligned} F_{3\mathrm{D}} &= \frac{\Phi_0}{2\mu_0}\sum_\mu \int \mathrm{d}\boldsymbol{r}_\mu B(\boldsymbol{r}_\mu) \\ &= \frac{\Phi_0^2}{8\pi\mu_0\lambda^2}\sum_\mu\sum_v \int \mathrm{d}\boldsymbol{r}_\mu \int \mathrm{d}\boldsymbol{r}_v \frac{\exp(-\tilde{r}_{\mu v}/\lambda)}{\tilde{r}_{\mu v}} \end{aligned}$$



Interaction between curved flux lines

Figure 5 Pairwise interaction between all line elements (arrows) of curved flux lines within the London theory (schematic).

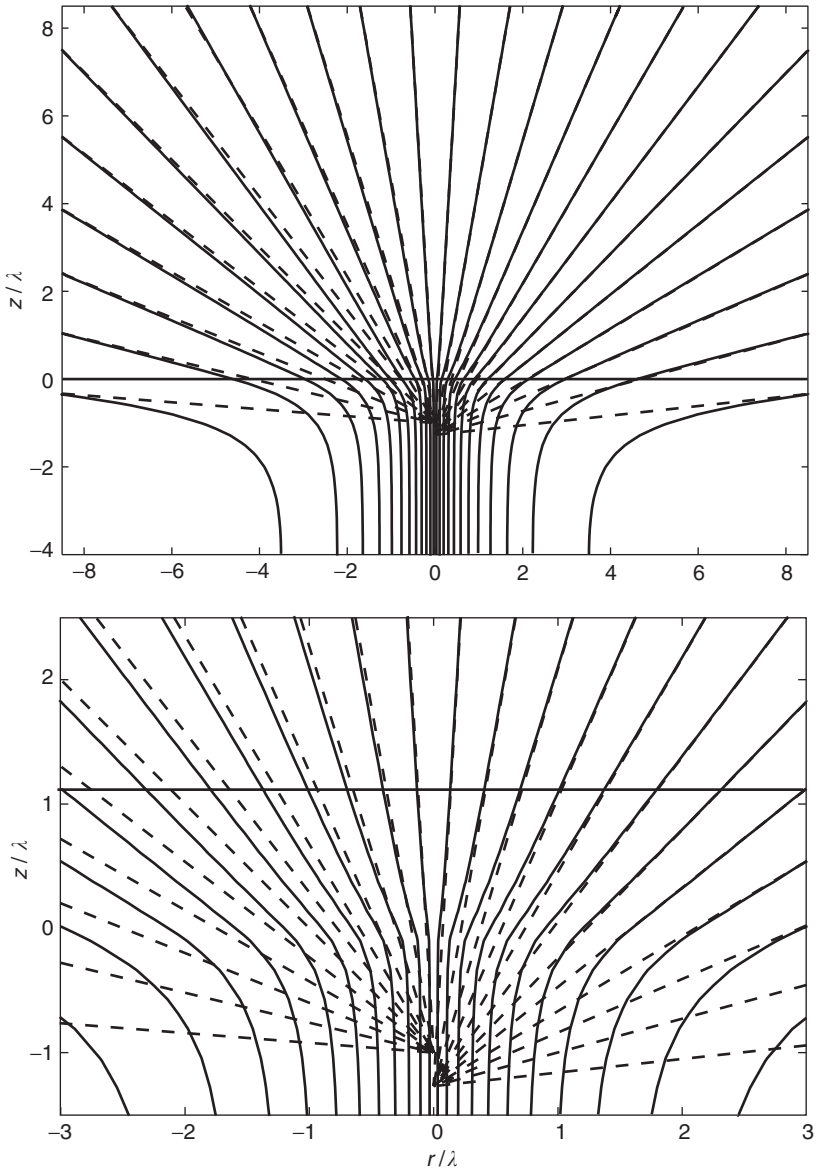


Figure 6 Magnetic field lines of a straight vortex in and near a superconducting half space. From the London theory, for $\kappa = 20$. The dashed lines give the radial field lines of a magnetic charge of size $2\Phi_0$ positioned on the vortex axis at a depth $-z_0 = \lambda$ (left half, the far field) and $-z_0 = 1.27\lambda$ (right half), which gives a better fit to the near field. The lower plot enlarges the center of the upper plot.

This means all the vortex segments interact with each other similar to magnetic dipoles or tiny current loops, but the magnetic long-range interaction $\propto 1/r$ is screened by a factor $\exp(-\tilde{r}_{\mu\nu}/\lambda)$. The 3D interaction between curved vortices is visualized in **Figure 5**.

Vortices Near Surfaces and in Films

The above London solutions apply to vortices in the bulk. Near the surface of the superconductor, these expressions have to be modified. In simple geometries, for example, for superconductors with one or two planar surfaces surrounded by vacuum, the magnetic field and energy of a given vortex arrangement is obtained by adding the field of appropriate images (in order to satisfy the boundary condition that no current crosses the surface) and a magnetic stray field which is caused by a fictitious surface layer of magnetic monopoles. This stray field makes the total magnetic field $B(\boldsymbol{r})$ continuous across the surface. **Figure 6** shows an example for this.

The magnetic field and interaction of straight vortices oriented perpendicular to a superconducting film of arbitrary thickness was calculated by Carneiro and Brandt from the London theory and by Brandt from the GL theory. In films of thickness $d \ll \lambda$ in a perpendicular magnetic field, the short vortices interact mainly via their magnetic stray field outside the

superconductor over an effective penetration depth $\Lambda = 2\lambda^2/d$. At short distances $r \ll \Lambda$, this interaction is logarithmic as in the bulk case, and at large $r \gg \Lambda$, it decreases as $\exp(-r/\Lambda)$. With decreasing thickness d, the Fourier transform of the 2D vortex interaction $V(r) = \int (\mathrm{d}^2k/4\pi^2)\tilde{V}(k)\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})$ changes from $\tilde{V}(k) = E_0\ (k^2 + \lambda^{-2})^{-1}$ for $d \gg \lambda$ to $\tilde{V}(k) = E_0(k^2 + k\Lambda^{-1})^{-1}$ for $d \ll \lambda$, where $E_0 = d\Phi_0^2/(\mu_0\lambda^2)$.

As shown by Chem, a similar (but 3D) magnetic interaction exists between the 2D "pancake vortices" in the superconducting Cu–O layers of high-T_c superconductors, $\tilde{V}(\boldsymbol{k}) = E_0 d k_3^2 k_2^{-2}(\lambda^{-2} + k_3^2)^{-1}$, where now d is the distance between the layers, $k_2^2 = k_x^2 + k_y^2, k_3^2 = k_2^2 + k_z^2$, and $\lambda = \lambda_{ab}$ is the penetration depth for the supercurrents flowing in these layers.

Elasticity of the Vortex Lattice

Small flux-line displacements caused by pinning forces or by thermal fluctuations may be calculated using the linear elasticity theory of the vortex lattice. **Figure 7** visualizes the three basic distortions of the triangular vortex lattice: shear γ, uniaxial compression ε, and tilt α, defining the three elastic moduli c_{66}, c_{11}, and c_{44} and the naive (local) elastic energy

$$F_{\mathrm{elast}} = \frac{V}{2}[c_{11}\varepsilon^2 + c_{66}\gamma^2 + c_{44}\alpha^2] \qquad [15]$$

The linear elastic energy F_{elast} of the vortex lattice is obtained by expanding its free energy F with respect to small displacements $\boldsymbol{u}_v(z) = \boldsymbol{r}_v(z) - \boldsymbol{R}_v = (u_{vx}, u_{vy})$ of the vortices from their ideal parallel lattice positions $\boldsymbol{R}_v$ and keeping only the quadratic terms. This yields

$$F_{\mathrm{elast}} = \frac{1}{2}\int_{\mathrm{BZ}} \frac{\mathrm{d}^3k}{8\pi^3} u_\alpha(\boldsymbol{k})\Phi_{\alpha\beta}(\boldsymbol{k})u_\beta^*(\boldsymbol{k}) \qquad [16]$$

where $\boldsymbol{u}(\boldsymbol{k})$ is the Fourier transform of the displacement field $\boldsymbol{u}_v(z)$, $(\alpha, \beta) = (x, y)$, and $\boldsymbol{k} = (k_x, k_y, k_z)$. The k-integral in eqn [16] is over the first Brillouin zone (BZ) of the FLL since the "elastic matrix" $\Phi_{\alpha\beta}(\boldsymbol{k})$ is periodic in the k_x, k_y plane; the finite vortex core radius restricts the k_z integration to $|k_z| \leqslant \xi^{-1}$. For an elastic medium with uniaxial symmetry, the elastic matrix reads

$$\Phi_{\alpha\beta}(\boldsymbol{k}) = (c_{11} - c_{66})k_\alpha k_\beta + \delta_{\alpha\beta}[(k_x^2 + k_y^2)c_{66} + k_z^2 c_{44}] \qquad [17]$$

Here the coefficients c_{11}, c_{66}, and c_{44} are the elastic moduli of uniaxial compression, shear, and tilt, respectively. For the vortex lattice, $\Phi_{\alpha\beta}(\boldsymbol{k})$ was calculated from the GL and London theories. The result, a sum over reciprocal lattice vectors, should coincide with expression [17] in the continuum limit, that is, for small $|\boldsymbol{k}| \ll k_{\mathrm{BZ}}$, where $k_{\mathrm{BZ}} = (4\pi\bar{B}/\Phi_0)^{1/2}$ is the radius of the circularized (actually hexagonal) Brillouin

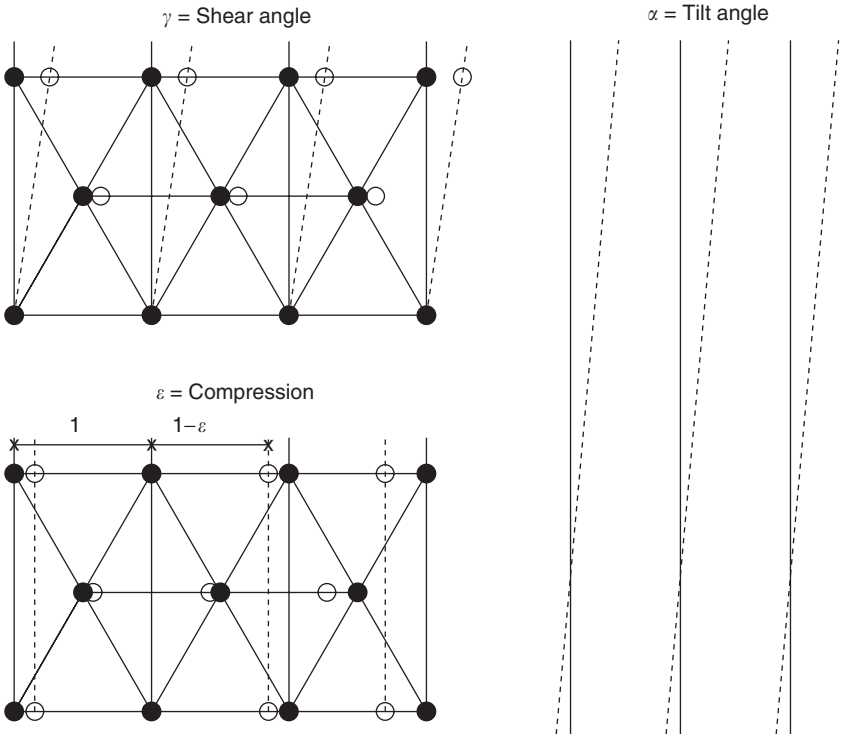


Figure 7 The three basic homogeneous elastic distortions of the triangular vortex lattice. The full dots and solid lines mark the ideal lattice, and the hollow dots and dashed lines the distorted lattice.

zone of the triangular vortex lattice with area πk_{BZ}^2. In the London limit, one finds for isotropic superconductors, the elastic moduli

$$c_{11}(k) \approx \frac{\bar{B}^2/\mu_0}{1+k^2\lambda^2}, \quad c_{66} \approx \frac{\bar{B}\Phi_0/\mu_0}{16\pi\lambda^2},$$
$$c_{44}(\boldsymbol{k}) \approx \frac{\bar{B}^2/\mu_0}{1+k^2\lambda^2} + \frac{\bar{B}\Phi_0/\mu_0}{8\pi\lambda^2} \ln \frac{\kappa^2}{1+k_z^2\lambda^2} \quad [18]$$

The GL theory yields an additional factor $(1-\bar{B}/B_{c2})^2$ in c_{66}, that is, $c_{66} \propto \bar{B}(\bar{B}-B_{c2})^2$, and replaces λ in c_{11} and c_{44} (first term) by $\lambda' = \lambda/(1-\bar{B}/B_{c2})^{1/2}$.

The $\boldsymbol{k}$ dependence (dispersion) of the compression and tilt moduli $c_{11}(k)$ and $c_{44}(\boldsymbol{k})$ means that the elasticity of the vortex lattice is nonlocal, that is, strains with short wavelengths, $2\pi/k \ll 2\pi\lambda$, have a much lower elastic energy than a homogeneous compression or tilt (corresponding to $\boldsymbol{k} \to 0$) with the same amplitude. This elastic nonlocality comes from the fact that the magnetic interaction between the flux lines typically has a range λ much longer than the flux-line spacing a_0; therefore, each flux line interacts with many other flux lines. Note that a large λ causes a small shear stiffness since $c_{66} \propto \lambda^{-2}$, and a smaller $c_{11}(k < \lambda^{-1})$ at short wavelengths, but the uniform compressibility $c_{11}(k=0)$ is independent of λ.

As a consequence of nonlocal elasticity, the vortex displacements $\boldsymbol{u}_v(z)$ caused by local pinning forces, and also the space- and time-averaged thermal fluctuations of the vortex positions, $\langle \boldsymbol{u}_v(z)^2 \rangle$, are much larger than they would be if $c_{44}(\boldsymbol{k})$ had no dispersion, that is, if it were replaced by $c_{44}(0) = \bar{B}H_a \approx \bar{B}^2/\mu_0$. The maximum vortex displacement $u(0) \propto f$ caused at $\boldsymbol{r}=0$ by a point force of density $f\delta_3(\boldsymbol{r})$, and the thermal fluctuations $\langle u^2 \rangle \propto k_B T$, are given by similar expressions,

$$\frac{2u(0)}{f} \approx \frac{\langle u^2 \rangle}{k_B T} \approx \int_{BZ} \frac{d^3k}{8\pi^3} \frac{1}{(k_x^2+k_y^2)c_{66} + k_z^2 c_{44}(\boldsymbol{k})} \approx \frac{k_{BZ}^2 \lambda}{8\pi[c_{66}c_{44}(0)]^{1/2}} \quad [19]$$

In this result, a large factor $[c_{44}(0)/c_{44}(k_{BZ})]^{1/2} \approx k_{BZ}\lambda \approx \pi\lambda/a \gg 1$ originates from the elastic nonlocality. In anisotropic superconductors with $B \| c$ (the crystalline c axis), the thermal fluctuations, eqn [19], are enhanced by an additional factor $\Gamma = \lambda_c/\lambda_{ab} \gg 1$, where λ_{ab} and λ_c are the two penetration depths of uniaxially anisotropic superconductors.

See also: Magnetism, History of; Quantum Mechanics: Methods; Statistical Mechanics: Classical; Superconductivity: BCS Theory; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductivity: Tunneling; Superconductors, High T_c.

PACS: 74.20.De; 74.25.Qt; 74.25.Op

Further Reading

Brandt EH (1986) Elastic and plastic properties of the flux-line lattice in type-II superconductors. *Physical Review B* 34: 6514–6517.

Brandt EH (1995) The flux-line lattice in superconductors. *Reports of Progress in Physics* 58: 1465–1594.

Brandt EH (2003) Properties of the ideal Ginzburg-Landau vortex lattice. *Physical Review B* 68(054506): 1–11.

Brandt EH (2005) Ginzburg-Landau vortex lattice in superconducting films of finite thickness. *Physical Review B* 71(014521): 1–21.

Brandt EH and Mikitik GP (2001) Meissner-London currents in superconductors with rectangular cross section. *Physical Review Letters* 85: 4164–4167.

Carneiro GM and Brandt EH (2000) Vortex lines in films: fields and interactions. *Physical Review B* 61: 6370–6376.

Clem JR (1991) Two-dimensional vortices in a stack of thin superconducting films: a model for high-temperature superconducting multilayers. *Physical Review B* 43: 7837–7846.

DeGennes PG (1966) *Superconductivity of Metals and Alloys*. New York: Benjamin.

Landau LD and Lifshitz EM (1960) *Theoretical Physics*, vol. 8 *Electrodynamics of Continuous Media*. New York: Pergamon.

London F (1950) *Superfluids*, Vol. 1. Chichester: Wiley.

Tinkham M (1975) *Introduction to Superconductivity*. NewYork: McGraw-Hill.

Superconductivity: Tunneling

A N Cleland, University of California, Santa Barbara, CA, USA

Introduction

The quantum mechanical tunneling of electrons and other charge carriers through an insulating or a vacuum barrier is at the heart of many electronic devices. In superconductors, the controlled tunneling of Cooper-paired electrons forms the basis for the Josephson junction and the DC and RF super-conducting quantum interference devices (RF and DC SQUIDs), as well as a number of other related devices.

The tunneling of a single particle through an energy barrier, such as an electron through an insulator or through vacuum, can be understood using a

one-dimensional quantum mechanical model. This single-particle picture can be extended to the tunneling of large numbers of electrons through a barrier separating two normal metals, a structure known as a tunnel junction, by convolving the one-particle tunneling result with the electronic distribution of states in the two metals; this can be done because the correlations between the electrons in a metal are sufficiently weak that they can be ignored. If the metals are in the superconducting state, both quasiparticles (the electron-like excitations in a superconductor) and Cooper pairs can tunnel through the barrier. The quasiparticles can be treated as independent, uncorrelated charge carriers, allowing the use of the single-particle picture. Cooper-pair tunneling involves the interference of the coherent superconducting wave functions on either side of the barrier, and, therefore, must be treated in a different fashion. This coherent process gives rise to the Josephson supercurrent. The DC Josephson effect describes the magnitude of the superconducting pair current, which is obtained for zero applied voltage and is determined by the difference in the quantum mechanical phases of the superconducting states on either side of the tunnel barrier. If a voltage V is applied between the two superconductors, this phase difference evolves in time, modulating the superconducting current; this is called the AC Josephson effect.

The article begins with the simple one-particle picture of tunneling between normal metals (NN tunneling), the discussion is then extended to include a superconductor on one side of the barrier, with a normal metal on the other (SN tunneling), and finally treats the case of superconducting metals on both sides of the barrier (SS tunneling). The article ends with a brief discussion of the effect of the electromagnetic environment on both normal and superconducting tunneling, an effect which is significant for tunnel junctions with very small geometric capacitance, where single charge effects, such as the Coulomb blockade, are relevant.

Normal Metal–Normal Metal Tunneling

Transmission through a Potential Barrier

An introductory problem in quantum mechanics is finding the transmission probability for a particle of mass m to pass through a one-dimensional potential barrier of height U and thickness d, with the barrier height greater than the particle energy E, as shown in **Figure 1a**. The particle is classically forbidden from entering the barrier region, but quantum mechanics allows a nonzero transmission probability. A solution of the one-dimensional Schrödinger equation

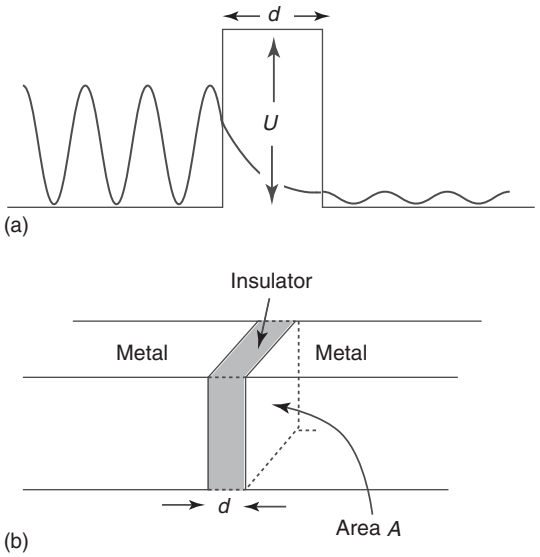


Figure 1 (a) Schematic of a one-dimensional potential barrier, with particle of energy E incident from the left and partially transmitted to the right. The barrier has height U and thickness d. (b) Schematic of metal–insulator–metal tunnel junction, of thickness d and cross-sectional area A.

yields a transmission coefficient that scales as $\exp[-2\kappa d]$, where

$$\kappa = \sqrt{\frac{2m(U-E)}{\hbar^2}} \qquad [1]$$

is the decay coefficient in the barrier region. The transmission is exponentially suppressed by increases in the barrier height and thickness. The tunneling transmission in the opposite direction, from right to left, is the same.

This simple model can be adapted to the tunneling of an electron in a metal through an insulating (or vacuum) barrier into another metal, as sketched in **Figure 1b**. The energy difference $\Delta E = U - E$ that determines the decay coefficient κ is taken to be $\Delta E \approx e(\Phi - \chi)$, the difference in the work function Φ of the metal and the insulator electron affinity χ. For a vacuum barrier, $\chi = 0$. For a characteristic energy difference $e(\Phi - \chi)$ of about 4 eV, the decay coefficient is $\kappa \approx 1 \times 10^{10}\ \mathrm{m}^{-1}$, so the transmission coefficient is reduced by a factor $1/e \approx 0.368$ for every 0.05 nm increase in the barrier thickness d. Reasonable transmission coefficients, yielding measurable rates of electron tunneling through such a tunnel barrier, require barrier thicknesses of the order $d \sim 1$ nm.

Current–Voltage Characteristic

In a real metal–insulator–metal interface, there are, of course, large numbers of electrons in each metal

electrode that can approach and interact with the tunnel barrier. The model described above is a single-particle picture, but can still be applied quite accurately to a metal as long as the Fermi exclusion principle is obeyed: there can be at most one electron in each distinct quantum mechanical state. This principle is included by allowing an electron to tunnel through the barrier only if the final state on the other side of the barrier has some probability of being empty; it is noted that, to first order, the initial and final energies of the tunneling electron must be equal.

At zero temperature, all the electron states in a metal with energy E less than the Fermi energy E_F are filled with probability unity, while all states above E_F are empty. For electrons at a nonzero temperature T, the probability that a state at energy E is occupied is given by the Fermi–Dirac distribution $f(E)$, which, for energies measured from the electrochemical potential μ_F, is given by

$$f(E) = \frac{1}{\exp(E/k_B T) + 1} \qquad [2]$$

The electrochemical potential μ_F is equal to the Fermi energy E_F at zero temperature; at a nonzero temperature, μ_F is the energy where the occupation probability is 1/2, and is very nearly equal to E_F if $k_B T \ll E_F$.

Consider applying a voltage V between the two metal electrodes, as shown in **Figure 2a**. This raises the electrochemical potential μ_L on the left electrode by an amount eV with respect to the electrochemical potential μ_R on the right, $\mu_L = \mu_R + eV$. An electron with energy E in the left electrode, measured from μ_L, will then tunnel into a state $E + eV$ in the right electrode, measured from μ_R. Using a transmission probability $|\mathscr{T}|^2$ per unit area, assumed independent of energy E, the rate at which electrons with energies in the range E to $E + dE$ will tunnel from the left to the right electrode is given by a Fermi's golden rule expression

$$\Gamma_{L\to R}(E)\,dE = \frac{2\pi A}{\hbar}|\mathscr{T}|^2 f(E)[1 - f(E + eV)]\,dE \qquad [3]$$

This includes the probability $f(E)$ of having an occupied state in the left electrode, and the probability $1 - f(E + eV)$ of a vacant state in the right electrode. There is a similar expression for electrons with energy between $E + eV$ and $E + eV + dE$ in the right electrode tunneling to the left

$$\Gamma_{R\to L}(E + eV)\,dE = \frac{2\pi A}{\hbar}|\mathscr{T}|^2 f(E + eV) \times [1 - f(E)]\,dE \qquad [4]$$

The net electrical current that flows due to these two opposing fluxes is the difference of these rates, summed over all initial energies, multiplied by the electron charge $-e$. Writing the sum over energy as an integral that includes the appropriate density of states $\mathscr{D}_{L,R}(E)$ in the left and right electrodes, the net electrical current (from right to left) is

$$\begin{aligned} I_{NN}(T) &= -e\Sigma_E(\Gamma_{R\to L}(E + eV) - \Gamma_{L\to R}(E)) \\ &= \frac{2\pi eA}{\hbar}|\mathscr{T}|^2 \int_{-\infty}^{\infty} dE\, \mathscr{D}_L(E)\mathscr{D}_R(E + eV) \\ &\quad \times \{f(E)[1 - f(E + eV)] \\ &\quad - f(E + eV)[1 - f(E)]\} \end{aligned} \qquad [5]$$

The range of integration is extended to $\pm\infty$, as the thermal occupations will limit the integrand to a range $k_B T \ll E_F$, so the increased range does not change the result. Then, the density of states is approximated on the left and right electrodes by their values at the Fermi level, $\mathscr{D}_{L,R}(E) \approx \mathscr{D}_{L,R}(0)$, and so

$$\begin{aligned} I_{NN}(T) &= \frac{2\pi eA}{\hbar}|\mathscr{T}|^2 \mathscr{D}_L(0)\mathscr{D}_R(0) \\ &\quad \times \int_{-\infty}^{\infty} dE\,(f(E) - f(E + eV)) \\ &= G_{NN} V \end{aligned} \qquad [6]$$

One thus finds a linear current–voltage relationship for a normal–insulator–normal metal tunnel junction, with a voltage- and temperature-independent electrical conductance G_{NN} given by

$$G_{NN} = \frac{2\pi e^2 A}{\hbar}|\mathscr{T}|^2 \mathscr{D}_L(0)\mathscr{D}_R(0) \qquad [7]$$

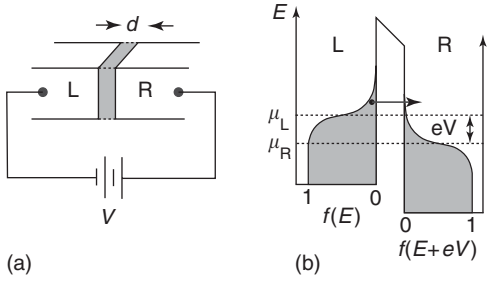


Figure 2 (a) Voltage V applied between metal electrodes on the left (L) and right (R). (b) Energy diagram for the electrons in the two electrodes with bias voltage V, showing the Fermi distributions on either electrode, with the zero of energy shifted by eV. The electrochemical potentials μ_L and μ_R are the points where the respective Fermi distributions are equal to 1/2, and are taken as the zero energy for each electrode.

Superconductor–Normal Metal Tunneling

If one of the normal metal electrodes is replaced with a superconductor, the rate of tunneling, and the form of the current–voltage characteristic, are modified

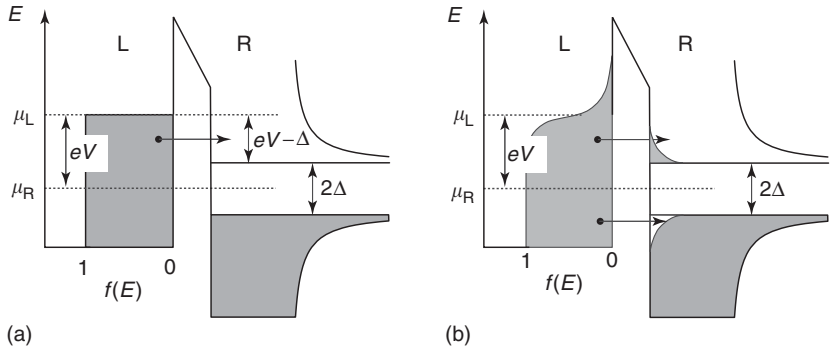


Figure 3 (a) Energy diagram at $T = 0$ for a normal metal (left) and superconductor (right) on either side of a tunnel barrier, the superconductor biased a voltage V above the normal metal. (b) Same as for (a), but at nonzero temperature, showing the distribution of occupied states in the normal metal and superconductor. Two possible tunneling processes, above and below the gap, are shown.

due to the change in the density of states in the superconducting electrode. At first, it appears that in addition, a change will occur due to the coherent quantum mechanical mixing of the electron and hole states above and below the Fermi level in the superconductor, involving the coherence factors that control this mixing. However, an electron can tunnel into either the hole-like or the electron-like excitation states, and the probabilities for these two processes add to 1. Therefore, the tunnel rate is affected only by the change in the density of states for the superconducting electrode. Choosing the right electrode as the superconductor, replace the normal metal density of states $\mathscr{D}_R(E)$ in eqn [5] with that for the superconductor,

$$\mathscr{D}_S(E) = \begin{cases} \mathscr{D}_R(0)\dfrac{|E|}{\sqrt{E^2 - \Delta^2}} & (|E| > \Delta) \\ 0 & (|E| < \Delta) \end{cases} \qquad [8]$$

where $\mathscr{D}_R(0)$ is the normal density of states at the Fermi level, and Δ is the superconducting energy gap, which depends on temperature. Energy E is measured from the electrochemical potential μ_R, which is the energy of the Cooper pairs.

At zero temperature, the quasiparticle states in the superconducting electrode are all unoccupied, as the quasiparticles are all condensed into Cooper pairs. The electron states in the left (normal metal) electrode are occupied with probability unity below the Fermi level E_F, and empty above. Applying a bias voltage V to the superconducting electrode with respect to the normal metal electrode, only electrons in the normal metal with energy E in the range $\Delta - eV < E < 0$ are able to tunnel into the superconductor (see **Figure 3a**). The current at zero temperature is given by

$$I_{SN}(0) = \frac{2\pi eA}{\hbar}|\mathscr{T}|^2 \int_{\Delta - eV}^{0} \mathscr{D}_L(E)\mathscr{D}_S(E + eV)\,dE \qquad [9]$$

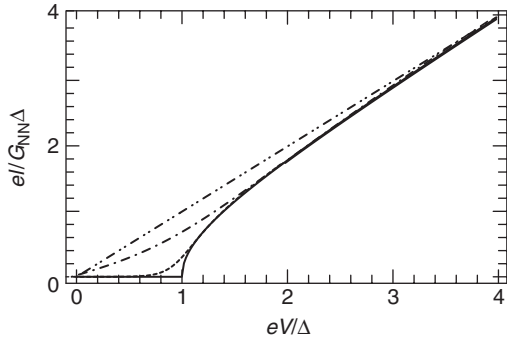


Figure 4 Current–voltage characteristic for a superconductor–normal metal tunnel junction, showing the calculated current for temperatures $k_B T = 0$, $\Delta/10$, $\Delta/2$, and for a normal–normal tunnel junction.

Inserting the density of states, eqn [8], for the superconductor, replacing the density of states for the normal metal, $\mathscr{D}_L(E)$, by the value at the Fermi level, $\mathscr{D}_L(0)$, and changing the variable of integration to $E' = E + eV$, this is

$$I_{SN}(0) = \frac{2\pi eA}{\hbar}|\mathscr{T}|^2 \mathscr{D}_L(0)\mathscr{D}_R(0) \times \int_{\Delta}^{eV} \frac{E'}{\sqrt{E'^2 - \Delta^2}}\,dE' \qquad [10]$$

$$I_{SN}(0) = \begin{cases} \dfrac{G_{NN}}{e}\sqrt{(eV)^2 - \Delta^2}, & |eV| \geq \Delta \\ 0, & |eV| < \Delta \end{cases} \qquad [11]$$

in terms of the normal–normal conductance G_{NN}, eqn [7]. The form of the zero-temperature conductance is plotted in **Figure 4**.

At nonzero temperatures, the occupation probabilities for the normal metal and superconducting quasiparticle states must be included. These are both described by the Fermi–Dirac distribution $f(E)$ given by eqn [2], measured from the local electrochemical potential; the corresponding distribution of occupied

states is shown in **Figure 3b**. The current at temperature T is the given by

$$I_{SN}(T) = \frac{G_{NN}}{e}\left(\int_{-\infty}^{-\Delta} + \int_{\Delta}^{\infty}\right)\frac{E'}{\sqrt{E'^2 - \Delta^2}} \times [f(E' - eV) - f(E')]\,dE' \qquad [12]$$

where the limits have been extended to $\pm\infty$ to include all the thermally excited states. This integral can be evaluated numerically. The current–voltage characteristics at a few temperatures are shown in **Figure 4**.

Andreev Reflection

The expressions derived for superconductor–normal metal tunneling are correct to first order in the tunneling probability $|\mathscr{T}|^2$. These predict zero current for voltages below the superconducting gap, $|V| < \Delta/e$, when at zero temperature. To second order in $|\mathscr{T}|^2$ (i.e., of order $|\mathscr{T}|^4$), however, an additional process becomes important. In this process, two electrons in the normal metal, symmetrically positioned about the electrochemical potential in the superconductor with wave vectors $\boldsymbol{k}$ and $-\boldsymbol{k}$, can tunnel simultaneously into the superconductor and form a Cooper pair. Another way to describe this process is that an electron with wave vector $\boldsymbol{k}$ tunnels into a Cooper pair, and the unfilled (hole) state in the pair tunnels simultaneously back into the normal metal, emerging with a wave vector $-\boldsymbol{k}$. The incident electron is then said to have been reflected from the superconductor as a hole, a process known as Andreev reflection. This process yields a nonzero conductance for a superconductor–normal tunnel barrier below the gap voltage.

Superconductor–Superconductor Tunneling

When both metals contacting a tunnel barrier are superconductors, there are two important effects that change the situation in comparison to the superconductor–normal metal tunnel junction: one is the appearance of the superconducting density of states on both sides of the tunnel barrier, changing the nature of the quasiparticle current–voltage characteristic. The other very significant change is the appearance of coherent Cooper-pair tunneling, present to first order in the tunneling probability $|\mathscr{T}|^2$. The theory for this process was first derived by B D Josephson, a prediction for which he was awarded the Nobel prize in 1973.

Quasiparticle Tunneling

The incoherent process of quasiparticle tunneling is first discussed, similar to that in the superconductor–normal metal tunnel junction. The rate of quasiparticle tunneling is controlled by the superconducting density of states appearing on both sides of the barrier, as well as the thermal occupation probability of these states. A sketch of the energy diagram is shown in **Figure 5**.

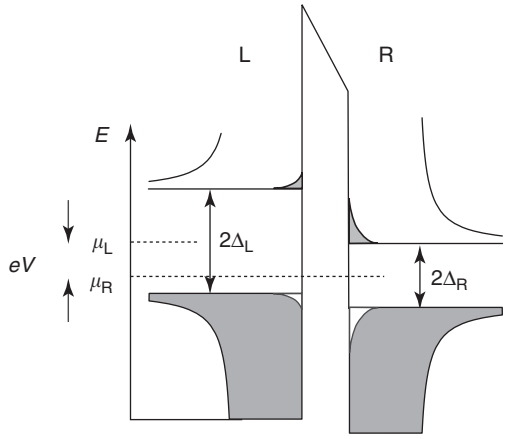


Figure 5 Quasiparticle energy diagram for two superconductors with different gaps Δ_L and Δ_R, biased with voltage V at a nonzero temperature.

The expression for the quasiparticle current $I_{SS}(T)$ at temperature T, for the right electrode biased by voltage V above the left electrode, is given by

$$I_{SS}(T) = \frac{G_{NN}}{e}\int_{-\infty}^{\infty}\frac{|E' - eV|}{\sqrt{(E' - eV)^2 - \Delta_L^2}}\frac{|E'|}{\sqrt{E'^2 - \Delta_R^2}} \times [f(E' - eV) - f(E')]\,dE' \qquad [13]$$

with energy E' measured from the electrochemical potential μ_R in the right electrode. Different superconducting gaps Δ_L and Δ_R are allowed, and it is understood that the density of states goes to zero in the gap regions on either side of the barrier, so that for energies E' in the ranges $|E'| < \Delta_R$ and $|E' - eV| < \Delta_L$, the integrand is zero. This is sketched in **Figure 5**.

At zero temperature, there is no current for voltages $|V| < (\Delta_R + \Delta_L)/e$ (for identical superconductors, this translates to $|V| < 2\Delta/e$). At the gap voltage $|V| = (\Delta_R + \Delta_L)/e$, there is a discontinuity in the current, which goes from zero to $(\pi G_{NN}/2e)(\Delta_R\Delta_L)^{1/2}$. For $T \neq 0$, an additional feature appears for voltages near the difference of the gap energies, $|V| \approx |\Delta_L - \Delta_R|/e$, due to the overlap of the discontinuity in the density of states of the larger-gap superconductor with the nonzero occupation number of the quasiparticle states in the smaller-gap superconductor. The form of the temperature-dependent current–voltage characteristic is shown schematically in **Figure 6**. For identical superconductors, with

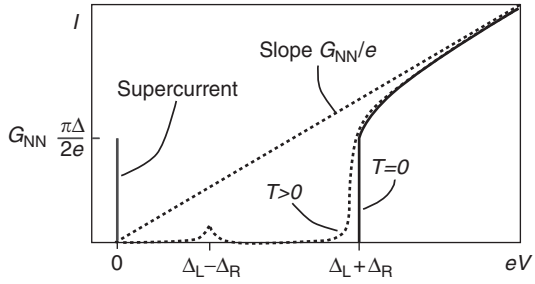


Figure 6 Sketch of a superconductor–superconductor current–voltage characteristic, showing the quasiparticle current for $T = 0$ and $T \neq 0$, for superconductors with different gaps. Also shown is the supercurrent branch at $V = 0$; the zero-temperature supercurrent amplitude I_0 and the discontinuity in the quasiparticle current at $eV = \Delta_L + \Delta_R$ are equal.

$\Delta_R = \Delta_L$, at very low temperatures $T \ll \Delta/e$, and voltages below the gap $|V| < 2\Delta/e$, an approximate analytic form for the current–voltage characteristic is given by Van Duzer and Turner

$$I_{SS}(V,T) = \frac{2G_{NN}}{e}\exp\left(-\frac{\Delta}{k_B T}\right)\sqrt{\frac{2\Delta}{eV+2\Delta}}(eV+\Delta) \times \sinh\left(\frac{eV}{2k_B T}\right)K_0\left(\frac{eV}{2k_B T}\right) \quad [14]$$

where K_0 is the zeroth-order modified Bessel function.

Josephson Effects

In addition to the incoherent tunneling of quasiparticles between two superconducting electrodes, there exist the very important and rich effects predicted by Josephson and first observed by Anderson and Rowell. A simple model for the Josephson effect is described by Feynman, and his model is reproduced here.

The superconducting state, being a macroscopic quantum condensate, can be described by a single macroscopic wave function Ψ, such that

$$\Psi(\boldsymbol{r}) = A(\boldsymbol{r})\exp[i\theta(\boldsymbol{r})] \quad [15]$$

where the real amplitude $A(\boldsymbol{r})$ is related to the Cooper-pair number density, $A(\boldsymbol{r})^2 = n(\boldsymbol{r})$, and $\theta(\boldsymbol{r})$ is the macroscopic phase of the superconducting wave function. In the simplest description, the equation of motion for the wave function of an isolated superconductor is

$$i\hbar\frac{\partial\Psi}{\partial t} = E\Psi \quad [16]$$

where E is the ground-state energy of the superconductor.

With two superconducting wave functions, Ψ_L and Ψ_R, on the left and right sides of a tunnel barrier, the tunnel barrier is assumed to give a weak coupling between the two wave functions, parametrized by the coupling constant K. The coupled equations of motion can be written as

$$i\hbar\frac{\partial\Psi_L}{\partial t} = E_L\Psi_L + K\Psi_R$$
$$i\hbar\frac{\partial\Psi_R}{\partial t} = E_R\Psi_R + K\Psi_L \quad [17]$$

The energies of the ground states E_R and E_L differ by the applied bias voltage V: $E_L - E_R = 2eV$, the factor of 2 accounting for the charge of a Cooper pair. Taking the zero of energy at the midpoint between the 2 superconductors, and separating the equations for the pair number density, $A_R = \sqrt{n_R}$ and $A_L = \sqrt{n_L}$, from that for the phases θ_R and θ_L,

$$\frac{\partial n_L}{\partial t} = +\frac{2}{\hbar}K\sqrt{n_R n_L}\sin\delta$$
$$\frac{\partial n_R}{\partial t} = -\frac{2}{\hbar}K\sqrt{n_R n_L}\sin\delta$$
$$\frac{\partial\theta_L}{\partial t} = -\frac{eV}{\hbar} - \frac{K}{\hbar}\sqrt{\frac{n_R}{n_L}}\cos\delta$$
$$\frac{\partial\theta_R}{\partial t} = +\frac{eV}{\hbar} - \frac{K}{\hbar}\sqrt{\frac{n_L}{n_R}}\cos\delta \quad [18]$$

where the phase difference $\delta = \theta_R - \theta_L$.

From these equations, it is seen that $\dot{n}_R = -\dot{n}_L$, so that the rate of increase of the pair density on the right is exactly the rate of decrease of pair density on the left; this is due to a current density $J = -2e\dot{n}_R$ flowing from right to left. In actuality, the pair density does not change. The depletion of charge in the vicinity of the barrier is replenished by a flow of current through the external circuit. To equilibrate the electrochemical potentials at zero voltage, take $n_L = n_R$.

It is thus found that at zero voltage, there is a supercurrent density $J = (2K/\hbar)\sqrt{n_R n_L}\sin\delta$. The critical current density of the barrier can be defined, $J_0 = (2K/\hbar)\sqrt{n_R n_L}$, so that $J = J_0\sin\delta$. Assuming uniform current density across the area A of the junction, the total current is $I = AJ$, and the critical current is defined as $I_0 = AJ_0$. Thus, the DC Josephson relation:

$$I = I_0\sin\delta \quad [19]$$

This shows that the current depends on the difference in the quantum mechanical phases of the superconducting states on either side of the tunnel barrier.

An equation can also be found for the evolution of the phase difference, by taking the difference of the fourth and third relations in eqn [18]. This is the AC

Josephson relation:

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = \frac{2eV}{\hbar} \quad [20]$$

(note the assumption $n_L = n_R$ has been used).

At zero voltage, the AC Josephson relation, eqn [20], shows that the phase difference δ remains fixed, and there is then a constant supercurrent $I = I_0 \sin\delta$. The phase difference can be manipulated through an external circuit, for instance, by closing a superconducting loop and applying an external magnetic field. This is the basis of the superconducting quantum interference devices, the DC and RF SQUID.

At a nonzero, fixed voltage V_{DC}, the phase evolves linearly in time

$$\delta = \delta_0 + \frac{2eV_{DC}t}{\hbar} \quad [21]$$

where δ_0 is the value at $t = 0$. Equation [19] then yields a sinusoidally oscillating current

$$I = I_0 \sin(2eV_{DC}t/\hbar + \delta_0) = I_0 \sin(\omega_J t + \delta_0) \quad [22]$$

The oscillation frequency

$$f_J = \frac{\omega_J}{2\pi} = \frac{2eV_{DC}}{h} \quad [23]$$

is called the Josephson frequency.

If one applies a voltage comprising a constant plus a sinusoidally oscillating term,

$$V = V_{DC} + v\cos\omega t \quad [24]$$

the phase evolves according to

$$\delta(t) = \delta_0 + 2eV_{DC}t/\hbar + (2ev/\hbar\omega)\sin\omega t \quad [25]$$

The Josephson supercurrent is then, for small $v \ll V_{DC}$,

$$I = I_0\left[\sin(\delta_0 + \omega_J t) + \frac{2ev}{\hbar\omega}\sin\omega t\cos(\delta_0 + \omega_J t)\right] \quad [26]$$

where $\omega_J = 2eV_{DC}/\hbar$. Averaging over times much larger than the inverse of the frequencies ω and ω_J, the first term vanishes, but the second term is nonzero for the frequencies $\omega_n = n\omega_J$, where n is an integer. This "phase-locking" of the Josephson current to an external signal gives rise to Shapiro steps, a structure in the current–voltage characteristic appearing when $V_{DC} = \hbar\omega/2ne$.

The critical current I_0 has a characteristic temperature dependence, and is related to the normal-state tunnel conductance G_{NN}, as can be worked out in the microscopic theory. The relation is

$$I_0(T) = G_{NN}\frac{\pi\Delta(T)}{2e}\tanh\left(\frac{\Delta(T)}{2k_BT}\right) \quad [27]$$

Effect of the Electromagnetic Environment

With the advent of technology capable of fabricating very small area tunnel junctions, with correspondingly small geometric capacitances, single charge effects became relevant in tunneling through both NN and SS tunnel barriers. These effects occur when the energy to charge the tunnel junction capacitance C becomes larger than the available thermal energy k_BT, creating what is called the Coulomb blockade. The simple theory described above only works well in the limit where the voltage across the tunnel barrier is fixed, that is, when the electromagnetic environment presents a very low impedance at all frequencies.

A more general treatment must be used when the voltage is not fixed, applicable to both NN and SS tunneling. For NN tunneling, the expression given by eqn [5] is modified to read

$$\begin{aligned} I_{NN}(T) = \frac{G_{NN}}{e}\int_{-\infty}^{\infty}\mathrm{d}E\int_{-\infty}^{\infty}\mathrm{d}E'\,\{f(E) \\ \times[1-f(E')]P(E+eV-E') \\ -[1-f(E)]f(E')P(E'-E-eV)\} \end{aligned} \quad [28]$$

The new term here is the appearance of the function $P(E)$, which allows the emission and absorption of energy between the tunneling electron and the electromagnetic environment during the tunneling process. This function is defined by

$$P(E) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\exp\left[J(t) + \frac{\mathrm{i}}{\hbar}Et\right] \quad [29]$$

using what is known as the equilibrium phase correlation function

$$J(t) = \langle[\phi(t) - \phi(0)]\phi(0)\rangle \quad [30]$$

In this expression, the phase $\phi(t)$ is the phase across the junction, related to the superconducting phase described above:

$$\phi(t) = \frac{e}{\hbar}\int_{-\infty}^{t}\mathrm{d}t'\,V(t') \quad [31]$$

When the electrical environment can be represented by a linear impedance $Z(\omega)$, the phase correlation function may be written as

$$\begin{aligned} J(t) = \int_0^{\infty}\frac{\mathrm{d}\omega}{\omega}\frac{\mathrm{Re}Z(\omega)}{R_Q}\left\{\coth\left(\frac{\hbar\omega}{2k_BT}\right)\right. \\ \left.\times[\cos(\omega t) - 1] - \mathrm{i}\sin\omega t\right\} \end{aligned} \quad [32]$$

where R_Q is the resistance quantum, $R_Q = h/2e^2$.

In the limit of a voltage-biased junction with a low electrical impedance, $Z(\omega) \ll R_Q$, the function $P(E)$

reduces to a delta function, $P(E) = \delta(E)$, and the results found above are obtained. More generally, however, the fluctuations presented by the real part of the environmental impedance significantly modify the current–voltage characteristic, most particularly in the limits where the tunnel barrier has a small capacitance.

The results above may be carried over to the discussion of quasiparticle tunneling in SS junctions. For the coherent Josephson processes, the effect of the electrical environment is different, but quite simple: one finds that one can get a supercurrent at nonzero voltages, and the current–voltage characteristic is given by

$$I_{SS}(V) = \frac{\pi\hbar I_0^2}{4e}[P(2eV) - P(-2eV)] \qquad [33]$$

See also: Superconductivity: Applications; Superconductivity: BCS Theory; Superconductivity: Critical Currents; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductors, High T_c; Superconductors, Metallic.

PACS: 74.20.−z; 74.45.+c; 74.50.+r; 82.25.−j; 85.25.Cp; 85.25.Dq

Further Reading

Ambegaokar V and Baratoff A (1963) Erratum. *Physical Review Letters* 11: 104.

Ambegaokar V and Baratoff A (1963) Tunneling between superconductors. *Physical Review Letters* 10: 486–489.

Anderson PW and Rowell JM (1963) Possible observation of the Josephson superconductive tunneling effect. *Physical Review Letters* 10: 230–232.

Barone A and Paterno G (1982) *Physics and Applications of the Josephson Effect*. New York: Wiley-Interscience.

Feynman RP, Leighton RB, and Sands M (1965) *Lectures on Physics*. Reading, MA: Addison-Wesley.

Grabert H and Devoret MH (eds.) (1992) *Single Charge Tunneling*. NATO ASI Series Vol. B 294. New York: Plenum Press.

Ingold G-L, Grabert H, and Eberhardt U (1994) Cooper-pair current through ultrasmall Josephson junctions. *Physical Review B* 50: 395–402.

Josephson BD (1962) Possible new effects in superconductive tunneling. *Physics Letters* 1: 251–253.

Josephson BD (1969) Weakly coupled superconductors. In: Parks RD (ed.) *Superconductivity*. New York: Dekker.

Shapiro S (1963) Josephson currents in superconducting tunneling: the effect of microwaves and other observations. *Physical Review Letters* 11: 80.

Solymar L (1972) *Superconductive Tunneling and Applications*. New York: Wiley.

Tinkham M (2004) *Introduction to Superconductivity*, 2nd edn. New York: Dover.

van Duzer T and Turner CW (1981) *Principles of Superconductive Devices and Circuits*. New York: Elsevier.

Superconductors, High T_c

M J Qin and S X Dou, University of Wollongong, Wollongong, NSW, Australia

Introduction

In 1911, Onnes discovered superconductivity in mercury with a transition temperature T_c of 4.2 K, at which the DC resistance of mercury dropped to zero. The second characteristic of a superconductor was its perfect diamagnetism, discovered by Meissner and Ochsenfeld in 1933. A theoretical understanding of superconductivity is based on the BCS microscopic theory, proposed by Bardeen, Cooper, and Schrieffer in 1957. Before 1986, the highest T_c in the now so-called low-temperature superconductors (LTSC) was found in Nb_3Ge (23 K).

In 1986, Bednorz and Muller reported superconductivity in $(LaBa)_2CuO_4$ with a T_c of 35 K, in contrast to the previous record of 23 K in LTSC. Their discovery opened up the era of high-temperature superconductivity. Shortly after that, $YBa_2Cu_3O_7$ ($T_c = 90$ K), $Bi_2Sr_2Ca_2Cu_3O_{10}$ ($T_c = 110$ K), and $Tl_2Ba_2Ca_2Cu_3O_{10}$ ($T_c = 125$ K) were discovered. Since then, high-temperature superconductivity has been the subject of intensive study. This article reviews some of the fundamental properties of high-temperature superconductor (HTSC).

Crystal Structures

Crystal structures of most HTSCs can be obtained by stacking perovskite, rock salt and/or fluorite structure along the *c*-axis. As an example, **Figure 1** shows the orthorhombic $YBa_2Cu_3O_{7-\delta}$ unit cell. Other families of HTSCs can be constructed in a similar way. The CuO chains are unique to $YBa_2Cu_3O_{7-\delta}$; however, the most important features of the crystal structures of HTSCs are the two-dimensional copper oxide CuO_2 conduction planes, which are critical for high-temperature superconductivity. Carriers are doped into the CuO_2 conduction planes by transfer of charges resulting from the introduction of foreign ions, interstitial defects, and vacancies in structural blocks between two CuO_2 planes. These blocks are called charge reservoirs.

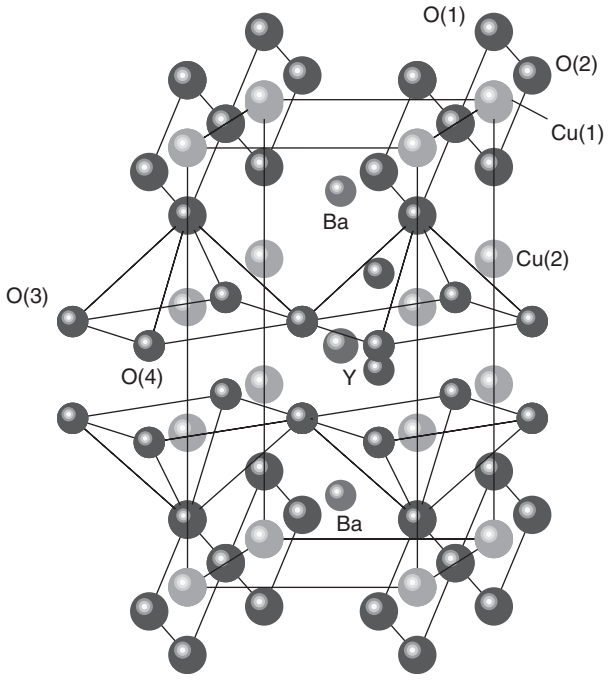


Figure 1 Crystal structure of orthorhombic $YBa_2Cu_3O_{7-\delta}$.

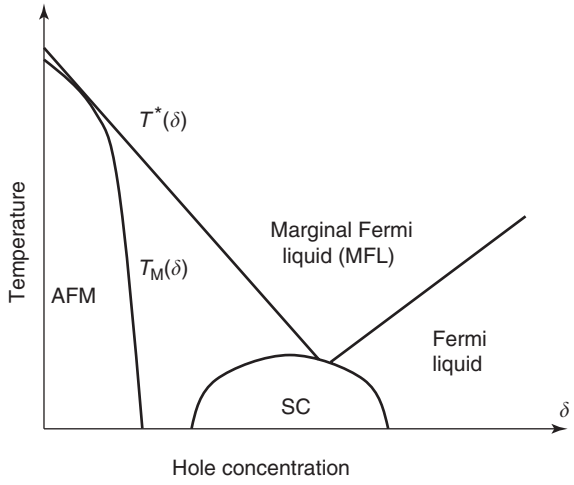


Figure 2 Generic temperature (T) vs. doping level (δ) phase diagrams of cuprates in zero magnetic field. (AFM: antiferromagnetic phase; SC: superconducting phase; T_N, and T^* are the Néel and pseudogap transition temperatures, respectively).

Phase Diagram

Depending on the temperature and the doping level, HTSCs demonstrate a wide variety of behaviors (see the phase diagram for hole-doped HTSCs in **Figure 2**). Undoped HTSCs are antiferromagnetic Mott insulators with nearest-neighbor Cu^{2+}–Cu^{2+} antiferromagnetic exchange interaction in the CuO_2 planes. The Néel temperature T_N for the antiferromagnetic-to-paramagnetic transition decreases with increasing doping level. At a certain doping level, the antiferromagnetism vanishes and one enters the pseudogap or underdoped region. The normal-state properties in this region are significantly different from those of a Fermi liquid. Some of the most interesting behaviors of HTSCs are observed in this region. Superconductivity arises upon further doping. The transition temperature T_c first increases with increasing doping level, reaching a maximum T_c at an optimal doping level, then decreases and finally vanishes with further increases in doping (overdoped regime). The non-Fermi liquid region above the superconducting region is well described by the so-called marginal Fermi liquid (MFL) hypothesis.

Superconducting Properties

The most unusual fundamental properties of HTSCs, compared to LTSCs, are high-transition temperature, short coherence length, and high anisotropy. The high anisotropy results from the layered structure shown in **Figure 1**. A phenomenological description can be achieved by means of the anisotropic Ginzburg–Landau equation.

Critical Temperature T_c

The critical temperature (i.e., transition temperature from normal state to superconducting state) is one of the most important parameters characterizing superconductors. The optimum T_c values for some of the HTSCs are listed in **Table 1**. T_c values above the liquid nitrogen temperature (77 K) open the way for many applications. However, higher operating temperatures also introduce large thermal fluctuations in HTSCs. Based on the two principal properties of superconductors (zero-resistance and perfect diamagnetism), the usual methods to obtain T_c are to measure the resistance versus temperature curve $R(T)$ or the AC susceptibility versus temperature $\chi(T)$ curve. Other methods such as measurement of DC magnetization versus temperature $M(T)$ and heat capacity versus temperature $C(T)$ are also applied to measure T_c.

Coherence Length ξ

It determines the distance over which the Cooper pairs are correlated. The short coherence lengths observed in HTSCs give rise to several unusual properties. Because the coherence length is much shorter than the penetration length, high-temperature cuprates are all type-II superconductors (see below). The short coherence length makes it very difficult to fabricate homogeneous high-temperature superconducting

Table 1 Space group, crystal structure, and optimum T_c of selected cuprate superconductors

Compound	*Space group*	*Structure*	*Optimum* T_c
La_2CuO_4	*I*4/*mmm*	214-T	40
	*P*4$_2$/*ncm*		
	Bmab	214-O	40
	Fmmm		
Nd_2CuO_4	*I*4/*mmm*	214-T	25
$(Nd,Ce,Sr)_2\ CuO_4$	*P*4/*mmm*	214-T	
$YBa_2Cu_3O_6$	*P*4/*mmm*	123-T	0
$YBa_2Cu_3O_7$	*Pmmm*	123-O	90
$YBa_2Cu_4O_8$	*Ammm*	124	80
$Y_2Ba_4Cu_7O_{15}$	*Ammm*	247	40
$(Ba,Nd)_2(Nd,Ce)_2Cu_3O_8$	*I*4/*mmm*	223	40
$Pr_2Ysr_2Cu_3O_8$	*P*4/*mmm*	2123	70
	Cmmm		
$Bi_2Sr_2CuO_6$	*Amaa*	Bi-2201	10
	*A*2/*a*		
	*C*2		
$Bi_2Sr_2CaCu_2O_8$	*Fmmm*	Bi-2212	95
	Amaa		
$Bi_2Sr_2Ca_2Cu_3O_{10}$	*I*4/*mmm*	Bi-2223	110
$Tl_2Ba_2CuO_6$	*I*4/*mmm*	Tl-2201	
	Fmmm		
$Tl_2Ba_2CaCu_2O_8$	*I*4/*mmm*	Tl-2212	115
$Tl_2Ba_2Ca_2Cu_3O_{10}$	*I*4/*mmm*	Tl-2223	125
$Tl_2Ba_2Ca_3Cu_4O_{12}$	*I*4/*mmm*	Tl-2234	
$TlBa_2CuO_5$	*P*4/*mmm*	Tl-1201	
$TlBa_2CaCu_2O_7$	*P*4/*mmm*	Tl-1212	
$TlBa_2Ca_2Cu_3O_9$	*P*4/*mmm*	Tl-1223	
$TlBa_2Ca_3Cu_4O_{11}$	*P*4/*mmm*	Tl-1234	

samples. Also associated with the short coherence length is the weak pinning of flux lines in HTSCs, compared with the pinning in LTSCs. To measure the coherence length, one usually needs to measure the upper critical fields $H_{c2}^{ab}(0)$ and $H_{c2}^{c}(0)$, then the coherence lengths can be derived according to the equations based on the anisotropic Gingzburg–Landau equation

$$\mu_0 H_{c2}^{c}(0) = \Phi_0/2\pi\xi_{ab}^2(0)$$

$$\mu_0 H_{c2}^{ab}(0) = \Phi_0/2\pi\xi_{ab}(0)\xi_c(0)$$

Here, $H_{c2}^{ab}(0)$ and $H_{c2}^{c}(0)$ indicate the upper critical fields at $T = 0\,\mathrm{K}$ for fields applied parallel to the *ab*-plane and the *c*-axis, respectively. Typical values for the coherence lengths are $\xi_{ab}(0) = 1.2$–$1.6\,\mathrm{nm}$, $\xi_c(0) = 0.15$–$0.3\,\mathrm{nm}$, for $YBa_2Cu_3O_{7-\delta}$, $\xi_{ab}(0) = 2.7$–$3.9\,\mathrm{nm}$, $\xi_c(0) = 0.045$–$1.6\,\mathrm{nm}$ for $Bi_2Sr_2CaCu_2O_{8+\delta}$, and $\xi_{ab}(0) = 2.1\,\mathrm{nm}$, $\xi_c(0) = 0.03\,\mathrm{nm}$ for $Tl_2Ba_2CaCu_2O_{8+\delta}$.

Penetration Depth λ

It determines the distance that the applied DC magnetic field can penetrate exponentially into the superconductor. There are several methods to measure the penetration depth, such as muon spin resonance (μSR), electron paramagnetic resonance (EPR), and the polarized neutron reflection techniques. The penetration depth can also be derived from measurements of susceptibility or magnetization on samples with dimensions similar to the penetration depth. Typical values for the penetration depths are $\lambda_{ab}(0) = 130$–$180\,\mathrm{nm}$, $\lambda_c(0) = 500$–$800\,\mathrm{nm}$ for $YBa_2Cu_3O_{7-\delta}$, and $\lambda_c(0) = 270$–$300\,\mathrm{nm}$, $\lambda_c(0) \geqslant 3700\,\mathrm{nm}$ for $Bi_2Sr_2Ca_2Cu_2O_{8+\delta}$. **Figure 3** shows schematically the length scales (λ and ξ) of a vortex when the applied field is parallel to the *ab*-plane.

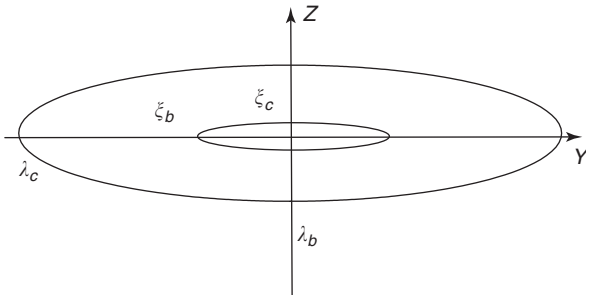


Figure 3 Schematic diagram of a single vortex when $H \| ab$, showing the anisotropic penetration depth and coherence length.

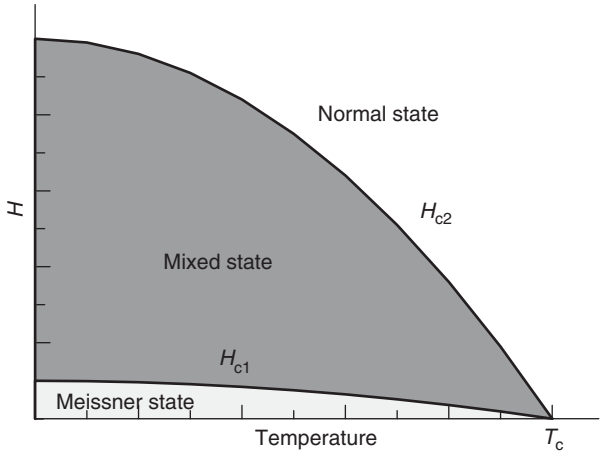


Figure 4 Phase diagram for a type-II superconductor.

Critical Fields

High-temperature cuprates are all type-II superconductors with three characteristic critical fields: the lower critical field H_{c1}, the upper critical field H_{c2}, and the thermodynamic critical field H_c (see **Figure 4**). $\mu_0 H_c^2/2$ is the free-energy difference between the superconducting and normal states. When the applied field is lower than H_{c1}, the superconductor shows perfect diamagnetism (called the Meissner state); when

the applied field is larger than H_{c1} but lower than H_{c2}, the applied field penetrates into the superconductor in the form of vortices (the so-called mixed state); when the applied field is larger than H_{c2}, the superconductor is forced into the normal state (see **Figure 4**). As the upper critical fields at low temperatures are experimentally inaccessible, one is normally interested in H_{c2} close to T_c; then using the dirty limit equation

$$H_{c2}(0) = -0.693 \times \left(\frac{dH_{c2}}{dT}\right)_{T_c} T_c$$

to obtain $H_{c2}(0)$. Typical values for $YBa_2Cu_3O_{7-\delta}$ are $\mu_0 H_{c2}^{ab}(0) = 200\,\mathrm{T}$ and $\mu_0 H_{c2}^{c}(0) = 40\,\mathrm{T}$. The lower critical field H_{c1} can be obtained from the virgin magnetization curves; typical values for $YBa_2Cu_3O_{7-\delta}$ at 4.2 K are $\mu_0 H_{c1}^{c} = 90\,\mathrm{mT}$, and $\mu_0 H_{c1}^{ab} = 20\,\mathrm{mT}$.

Pairing Symmetry

The pairing symmetry is very important for the discovery of the pairing mechanism. After the discovery of HTSCs, experiments such as those involving the observation of magnetic flux states of a polycrystalline $YBa_2Cu_3O_{7-\delta}$ ring, Andreev reflection, the Shapiro step of a Josephson junction, the high-resolution Bitter pattern technique, and Little–Parks oscillations have been performed to test whether Cooper pairing operates as in conventional superconductors. The results have confirmed electronic pairing in cuprate superconductors. In conventional superconductors, the phonon mediated electron–electron interaction gives rise to spin-singlet pairing with *s*-wave symmetry. However, the pairing symmetry in cuprate superconductors has been a controversial topic. Phase sensitive techniques (superconducting quantum interference device (SQUID) interferometry, single Josephson junction modulation, thin film magnetometry, etc.), combined with the refinement of several other symmetry-sensitive techniques (penetration depth, specific heat, thermal conductivity, angle-resolved photoemission, Raman scattering, nuclear magnetic resonance, nonlinear Meissner effect), have provided evidence in favor of *d*-wave symmetry pairing in cuprate superconductors. The energy gap, quasiparticle state, vortex state, and the surface and interfaces of *d*-wave superconductors are distinctly different from the *s*-wave case. Detailed discussions can be found in the "Further reading" section.

Energy Gap

The energy gap indicates the minimum energy required to excite a quasiparticle from a phase-coherent Cooper-pair condensate. The BCS theory predicts a universal ratio between the energy gap and the thermal energy $2\Delta(0)/k_B T_c = 3.53$, here $\Delta(0)$ is the energy gap at $T = 0\,\mathrm{K}$. For HTSCs, the energy gap can be determined from experiments such as single-electron tunneling, Andreev reflection, Raman scattering, photoemission spectra. Typical values for $YBa_2Cu_3O_{7-\delta}$ are $2\Delta_{ab}(0)/k_B T_c = 6$–$8$, $2\Delta_c(0)/k_B T_c = 3$–3.5; and for $Bi_2Sr_2CaCu_2O_{8+\delta}$ are $2\Delta_{ab}(0)/k_B T_c = 8$–$11$, $2\Delta_c(0)/k_B T_c = 5$–7. For conventional superconductors, the Cooper pairs have an isotropic *s*-wave symmetry, which means that the energy gap is 2Δ everywhere in *k*-space (see **Figure 5**). However, for a

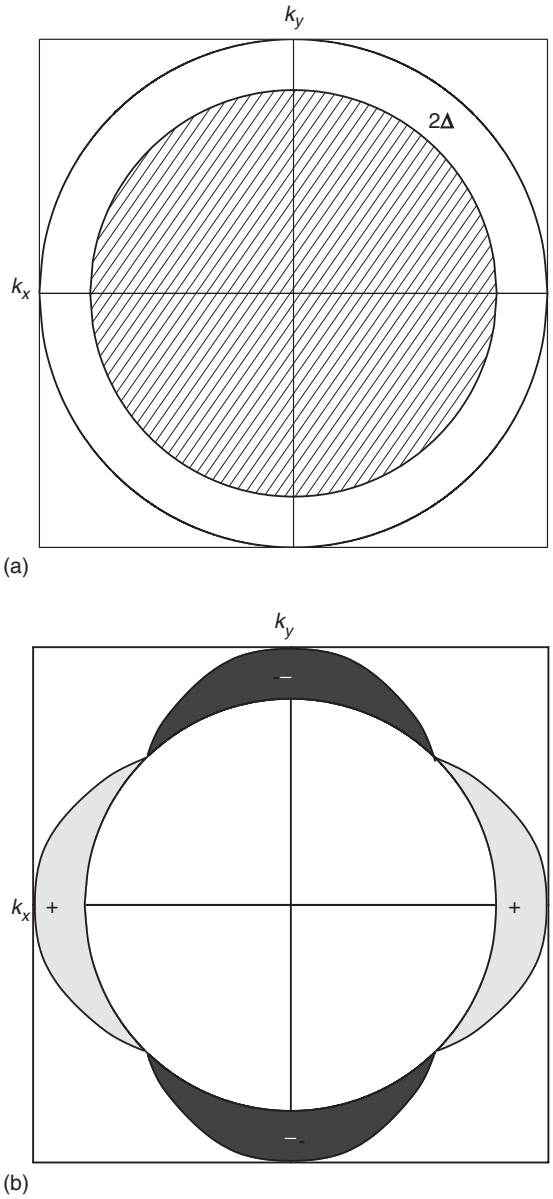


Figure 5 The energy gap for *s*-wave superconductors (upper panel) and *d*-wave superconductors (lower panel). The *s*-wave symmetry has an isotropic value of 2Δ, while for a *d*-wave symmetry the gap vanishes and changes sign at $k_x = k_y$.

d-wave symmetry superconductor, in momentum space when $k_x = \pm k_y$, the energy gap is zero, which implies that it would change sign at $k_x = \pm k_y$ (see **Figure 5**). Both properties will have important implications from both fundamental and applied points of view.

Type-II Superconductivity

HTSCs have very short coherence lengths ξ and large penetration depths, so the Ginzburg–Landau parameter $\kappa = \lambda/\xi \gg 1$, and HTSCs are typical type-II superconductors. In this section, some of the properties in the mixed state are discussed.

Vortex Structure

Due to the high anisotropy of cuprate superconductors, the vortex structures are different from those in conventional superconductors. When the applied field is along the *ab*-plane, the screening current flows both along the *ab*-plane and along the *c*-axis. There exist insulating layers alternating with conduction layers along the *c*-axis, the current is the Josephson current in the insulating layers (see **Figure 6**). When the applied field is parallel to the *c*-axis or at an angle to the *c*-axis, the vortices can be regarded as pancake vortices stacked along the *c*-axis (see **Figure 7**). Each pancake vortex is located in one of the conduction planes. The pancake vortices in the neighboring layers are coupled by means of a Josephson current, which forms a Josephson vortex. The forces between pancake vortices in the conduction plane are repulsive, while the forces between pancake vortices in the neighboring layers are attractive. Depending on the temperature and the magnetic field, the vortex system can go through a 3D to 2D transition. The magnetic properties of 2D superconductors can be described by the Lawrence–Doniach (LD) model.

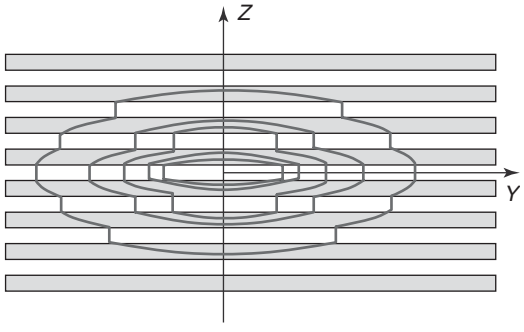


Figure 6 Schematic diagram of a single vortex when the applied magnetic field is parallel to the *a*-axis (*x*-direction). Gray areas indicate the superconducting layers.

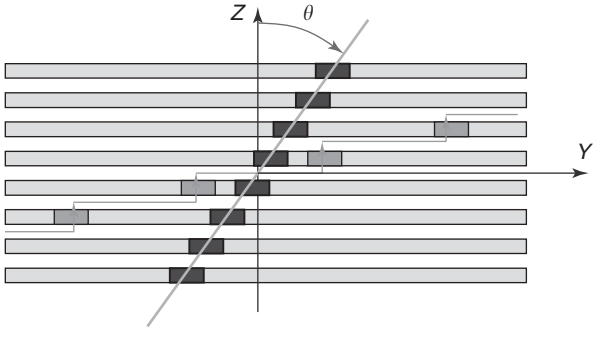


Figure 7 Schematic diagram for pancake vortices when the applied magnetic field is at an angle to the *c*-axis (*z*-direction). Gray areas indicate the superconducting layers.

Irreversibility Field

After the discovery of cuprate superconductors, it was soon reported that zero-field-cooled (ZFC) and field-cooled (FC) magnetization of La–Ba–Cu–O versus temperature at a fixed applied magnetic field can be separated into two regions. Below T_{irr}, ZFC and FC magnetization are different and irreversible, above T_{irr}, ZFC and FC magnetization are the same and reversible. T_{irr} is dependent on the magnetic field. $T_{irr}(H)$ or $H_{irr}(T)$ is called the irreversibility line. $H_{irr}(T)$ is lower than H_{c2} but larger than H_{c1} and separates the mixed state into two regions, a vortex liquid above $H_{irr}(T)$ and a vortex solid below $H_{irr}(T)$. The irreversibility line can also be obtained by hysteresis loop measurements and *I*–*V* characteristics measurements. Generally, the irreversibility line follows $1 - T_{irr}/T_{c0} = CH_{irr}^{q}$, with T_{c0} the zero-field transition temperature, and $1/2 < q < 3/4$, depending on different materials. The vortex system depends on three energy scales: the pinning potential E_{pin}, the thermal energy E_{th} and the vortex–vortex interaction energy E_{vv}. When the $E_{th} \approx E_{pin}$, the system reaches the reversible region.

Critical Currents and Flux Pinning

Due to short coherence lengths and high anisotropy in cuprate superconductors, critical currents are much more complicated than in conventional superconductors. For an anisotropic superconductor, three critical currents are defined depending on the direction of the applied magnetic field (see **Figure 8**). Critical currents result from the interaction between vortices and crystal defects (pinning effect). Twin boundaries and screw dislocations in single crystal $YBa_2Cu_3O_{7-\delta}$ and Y_2BaCuO_5 phase, as well as stacking faults and dislocations in melt-textured-growth (MTG) $YBa_2Cu_3O_{7-\delta}$, have all been reported to act as pinning centers. Heavy ion and particle irradiation have also been used to enhance the critical current density in cuprate superconductors. Due

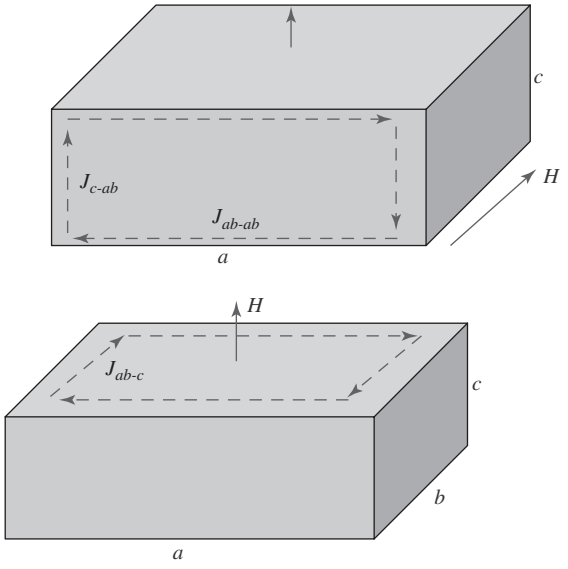


Figure 8 Schematic diagram showing the anisotropic cirtical current density in cuprate superconductors, when $H \| ab$ (upper panel) and $H \| c$ (lower panel).

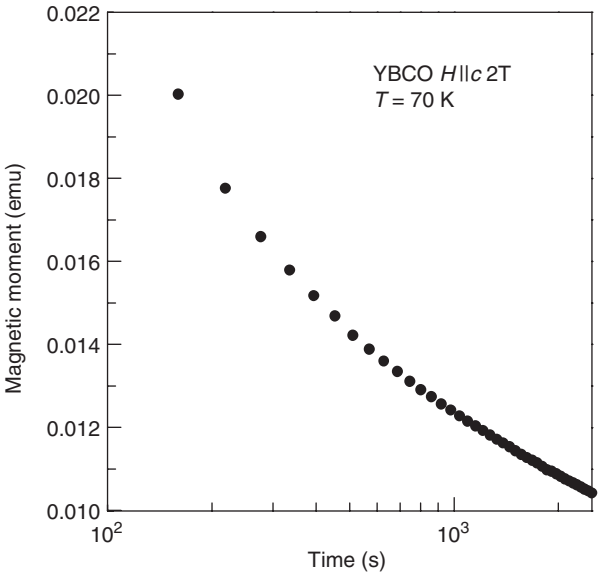


Figure 9 Magnetic relaxation of a melt-texture-growth (MTG) YBCO sample at $H \| c = 2\,\mathrm{T}$ and 70 K, showing nonlogarithmic relaxation.

to the short coherence length, point defects such as vacancies and substitution atoms may play an important role in cuprate superconductors.

Critical current can be directly obtained from transport I–V measurements. From hysteresis loop measurements, the critical state model is used to derive the current density from the measured hysteresis loop,

$$J_{ab,c} = \frac{20\Delta M}{a(1 - a/3b)} \quad (H \| c)$$

Because a larger electric field criterion is used in transport measurement, it usually provides a larger critical current density than magnetic measurements.

In polycrystalline samples, the current density across grains, called the intergrain critical current density, is very small, and shows the properties of a Josephson current. This is due to large angle grain boundaries in polycrystalline samples, and therefore these grain boundaries are called weak links. In order to have high critical current density, weak links should be avoided, by using, for example, MTG for $YBa_2Cu_3O_{7-\delta}$ bulk materials. Intragrain critical current densities up to $10^6\,\mathrm{A\,cm^{-2}}$ at 77 K and 0 T has been reported for $YBa_2Cu_3O_{7-\delta}$ thin films.

Flux Creep

In conventional superconductors, the current density was observed to decrease logarithmically with time at fixed temperature and field, an effect called magnetic relaxation. This effect was explained by vortices thermally activated out of the pinning potential (flux creep) with the hopping frequency

$$v = v_0 \exp\left[\frac{-U(j)}{k_B T}\right]$$

where v_0 is the attempt frequency and $U(j)$ the effective activation energy depending linearly on the current density,

$$U(j) = U_0\left(\frac{1 - j}{j_{c0}}\right)$$

The current density decreases logarithmically with time as

$$j(t) = j_{c0}\left[1 - \left(\frac{k_B T}{U_0}\right)\ln\left(\frac{t}{t_0}\right)\right]$$

In cuprate superconductors, strong and nonlogarithmic relaxation has been observed (see **Figure 9**). Due to high-operating temperatures and reduced pinning compared to conventional superconductors, relaxation in HTSCs is very large. Nonlogarithmic relaxation in HTSCs results from a nonlinear $U(j)$ relationship. A logarithmic $U(j)$ relationship has been proposed to account for the experimental results:

$$U(j) = U_0 \log\left(\frac{j_c}{j}\right)$$

On the other hand, according to the vortex glass theory, the $U(j)$ relationship is

$$U(j) = \left(\frac{U_0}{\mu}\right)\left[\left(\frac{j_c}{j}\right)^{\mu} - 1\right]$$

Here, the exponent μ depends on the dimensionality of the vortex system.

Magnetic relaxation measurements have been extensively performed on HTSCs to study the flux dynamics. The purpose is to determine the effective pinning potential $U(j)$ and compare it with theoretical predictions. Several methods such as the Maley method, the generalized inversion scheme, and dynamic relaxation and the AC susceptibility technique have been developed for this purpose. However, because too many parameters are used in theoretical predictions, it is impossible to unambiguously determine the effective activation energy from experiments. Most results in this area are not conclusive.

Melting of the Vortex Lattice

Thermal fluctuations in HTSCs are much more important than in conventional superconductors. According to the Lindemann melting criterion, as the displacement away from the ideal configuration by thermal fluctuation reaches 10–30% of the average flux-line separation, melting transition from vortex solid-to-vortex liquid (shear modulus equal to zero) might take place. The curve $T_m(H)$ in the H–T phase diagram at which the flux lattice melts is called the melting line. It has been suggested that in clean samples with negligible pinning the vortex solid-to-vortex liquid transition is expected to be of the first order, while static disorders drive the transition to the second order as in the vortex glass theory. Evidence for a first-order melting transition comes from a jump in the reversible magnetization, changes in latent heat, a frequency-independent peak in the AC susceptibility and a sharp, hysteretic, resistive transition. However, the effect of static disorders and their dimensionality on the order of transition is unclear as yet.

Mixed-State Phase Diagram

For a type-II superconductor, by varying the external field H, at the lower critical field H_{c1}, a phase transition from the Meissner phase to the Abrikosov phase (or mixed phase) takes place. The basic unit of the mixed state is the vortex, which contains a flux quantum given by $\Phi_0 = ch/2e = 2.07 \times 10^{-7}\,\mathrm{G\,cm^2}$, penetrating the system and forming a perfect triangular crystal (the Abrikosov lattice). At the upper critical field H_{c2}, the system is forced to undergo a second phase transition to the normal state. The simple phase diagram shown in **Figure 4** neglects disorders and thermal fluctuations, which is very serious in cuprate superconductors due to higher operating temperatures. When thermal fluctuations are taken into account, the first modification applied to the phase diagram ought to be the inclusion of the

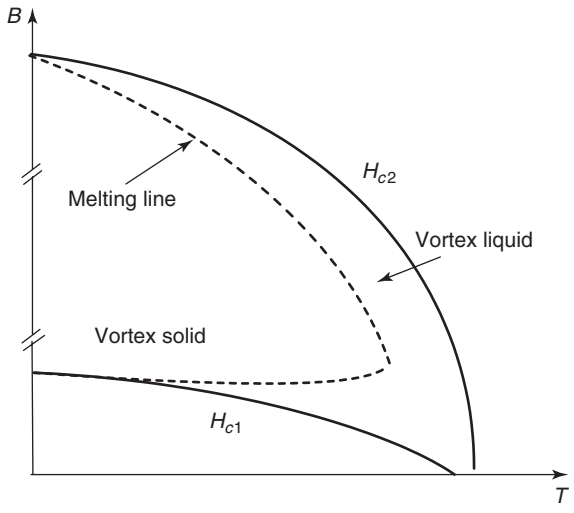


Figure 10 Schematic vortex phase diagram for high-temperature superconductors without disorders.

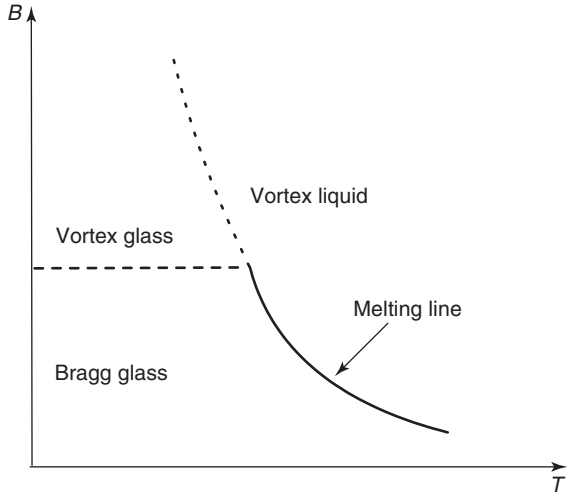


Figure 11 Schematic vortex phase diagram for high-temperature superconductors with disorders.

"water-like" melting phenomenon discussed above. In clean systems, the melting curve stretches from H_{c1} up to H_{c2} (see **Figure 10**). In case of disordered systems, the melting line terminates at much lower fields (characterized by the first-order transition). In practice, one is often far from the critical fields, leading to the unified phase diagram in **Figure 11**. Then, one has to discriminate between a sharp first-order phase transition at lower fields separating Bragg glass and vortex liquid and a more continuous second-order-like transition at higher fields separating vortex glass and vortex liquid.

Normal-State Properties

Properties in the normal state of cuprate superconductors have been found to be very useful for the

understanding of the mechanism of high-temperature superconductivity. These properties include resistivity, Hall coefficient, thermopower, magnetism, thermal conductivity, and optical properties. Although superconducting properties of cuprate superconductors are quite similar to those of conventional superconductors, the normal-state properties are unique. As an example, some general features for normal-state resistivity are listed: (1) the resistivity is anisotropic with the *c*-axis resistivity ρ_c two orders of magnitude larger than the in-plane resistivity ρ_{ab}, (2) ρ_{ab} shows metallic behavior, but ρ_c mostly shows semiconductor behavior. (3) ρ_{ab} is close to linear in T, which is the most striking normal-state property of cuprate superconductors. Theoretical explanations for these behaviors and other unique behaviors observed in optical and thermal properties remain controversial.

Mechanism of High-T_c Superconductivity

Superconductivity happens when charge carriers overcome their mutual electrostatic repulsion, bind together into Cooper pairs, and condense into a single quantum state below a certain temperature. In conventional superconductors, interactions between charge carriers and phonons (quantum vibrations of the crystal lattice) are responsible for the pairing. In HTSCs, however, the mechanism for the pairing has been controversial since the discovery of high-temperature superconductivity. The challenges are not only to find a mechanism to explain the high transition temperature in cuprate superconductors, but also to be able to account for the complex phase diagram exhibited by cuprate superconductors (see **Figure 2**). The phonon-based pairing mechanism in LTSCs is based on the so-called Landau–Fermi theory, in which the properties of single electrons are renormalized by interactions with other electrons to form quasiparticles. However, recent experimental results have provided evidence that phonons are not the principal cause of high-temperature superconductivity. There have been numerous theories proposed to understand high-temperature superconductivity. One proposal, which has attracted much attention, is based on the idea of a doped resonant valence bond (RVB) state. Other competing theories include those based on fluctuating stripes, those based on the polaron mechanism, and those based on spin fluctuations. However, a comprehensive theory, which could account for all the experimental observations, is yet to be achieved. The challenges for a successful theory are that it should predict the high transition temperature, describe the full complex phase diagram, and that it should reveal the special conditions in the cuprates that lead to this very special behavior. From this should follow some suggestions for other materials that would show similar behavior.

Material Processing

Material processing depends on the requirements of applications. Thin films, bulk materials, tapes and wires, and coated conductors have received much attention due to their applications in large- or small-scale applications.

Thin Films

$YBa_2Cu_3O_{7-x}$(YBCO) thin films are the most studied high-temperature superconducting thin films. Critical current densities up to $10^6\,A\,cm^{-2}$ have been achieved in high-quality YBCO thin films, making them very attractive for applications. There are, in general, two main methods for fabricating high-temperature superconducting thin films: *in situ* or *ex situ* film growth. *In situ* films are oxygen-treated before they are removed from the growth chamber, and therefore are superconducting after growth. *Ex situ* thin films must be annealed in an oxygen-rich environment after growth to be superconducting. Nearly every growth technique has been tried on high-temperature superconducting thin films, including evaporation, pulsed laser deposition (PLD), metal organic chemical vapor deposition (MOCVD), and on-axis and off-axis sputtering. Every technique has its merits and drawbacks, details of which can be found in the references. The search for substrate materials is an active research area. High-quality YBCO thin films have been deposited on $LaAlO_3$, $SrTiO_3$, MgO, YSZ, etc. BiSrCaCuO and TlBaCaCuO thin films have also been extensively studied. However, it is very difficult to obtain single-phase BiSrCaCuO thin films, while Tl is very toxic, creating problems in fabricating TlBaCaCuO thin films. The main application of superconducting thin films is in the area of superconducting electronics.

Bulk Materials

Many large-scale applications such as flywheel energy-storage systems need HTSCs in bulk form. Sintered YBCO polycrystalline samples used to have very low critical current densities ($100–1000\,A\,cm^{-2}$), which was attributed to weak links at grain boundaries, such as structural disorder at grain boundaries, chemical or structural variations at grain boundaries and non-superconducting materials along grain boundaries, as well as microcracking. Texture processing has been developed as a means to avoid or minimize the effects of weak links in polycrystalline YBCO. The initial success with bulks was due to the MTG process, which enhanced the critical current density two to

three orders of magnitude over that of sintered materials. Several modified MTG methods were developed later in order to further enhance the critical current density. These include the quench-and-melt-growth (QMG) process, the melt-powder-melt-growth (MPMG) process, the power-melting process (PMP), the floating-zone-melt (FZM) process, the platinum-doped melt-growth (PDMG) process, and the horizontal Bridgman method.

Tapes and Wires

The majority of the work on superconducting tapes and wires focuses on the $Bi_2Sr_2Ca_2Cu_3O_{10-x}$(BSCCO) system. The reasons are that the BSCCO system has a higher critical current at high magnetic fields, and that the BSCCO system does not suffer from the weak link problems of YBCO. The fabrication of BSCCO tapes and wires is usually carried out by means of the so-called oxide power-in-tube (OPIT) method. The powder is packed in a tube (typically Ag), then the Ag tube is drawn to a small diameter and rolled into a flat tape. The tape is then annealed, re-rolled and re-annealed again to obtain the best superconducting properties. Critical current densities up to $7000\,A\,cm^{-2}$ for a 650 m long and 19-filament-rolled tape have been achieved. Tapes and wires are widely used in applications such as motors and generators, magnet, transformers, and transmission lines.

Coated Conductors

Coated conductors are YBCO thin films deposited onto polycrystalline substrates. The films are biaxially textured, thus overcoming the phenomenon of weak links. This has been achieved by preparing a biaxially textured buffer layer, upon which the YBCO film is subsequently grown epitaxially. The two most advanced techniques are the ion beam-assisted deposition (IBAD) process, and the rolling-assisted biaxially textured substrate (RABiTS) process. In addition, inclined substrate deposition (ISD) has also emerged as a promising method for the fabrication of YBCO conductors. The coated conductors potentially have a performance superior to the first generation bismuth-based HTS compounds (BSCCO), such as high current carrying capacities, strong mechanical properties, and extraordinary magnetic field tolerances. Recently, it has been demonstrated that current densities up to $10^6\,A\,cm^{-2}$ have been achieved in conductors prepared by coating thick films of YBCO onto flexible metallic substrates. One of the biggest hurdles to widespread application of YBCO-coated conductor tape is developing a manufacturing process that will produce it in long lengths and at prices competitive to copper for applications such as motors, generators, transmission cables, and other power systems.

See also: Phases and Phase Equilibrium; Powder Processing: Models and Simulations; Structure Types and Classifications; Vapor Transport Processing: Mechanisms.

PACS: 74.20. – z; 74.25. – q; 74.70. – b; 74.72. – h

Further Reading

Blatter G, Feigel'man MV, Geshkenbein VB, Larkin AI, and Vinokur VM (1994) Vortices in high-temperature superconductors. *Reviews of Modern Physics* 66: 1125.

Campbell AM and Evetts EJ (1972) *Critical Currents in Superconductors*. London: Taylor and Francis.

Cohen LF and Jensen HJ (1997) Open questions in the magnetic behavior of high-temperature superconductors. *Reports of Progress in Physics* 60: 1581.

Gennes PGde (1966) *Superconductivity of Metals and Alloys*. New York: Benjamin.

Ginsberg DM (ed.) (1989) *Physical Properties of High Temperature Superconductors I*, *II*, *III*, *IV*. Singapore: World Scientific.

Huebener RP (1970) *Magnetic Flux Structures in Superconductors*. Berlin: Springer.

Jin SH (ed.) (1993) *Processing and Properties of High T_c Superconductors*. Singapore: World Scientific.

Senoussi S (1992) *Journal of Physics* (*III*), France vol. 2, p. 1102.

Silver T, Pan AV, Ionescu M, Qin MJ, and Dou SX (2002) Developments in high temperature superconductivity. *Annual Report Progress Chemistry* 98(sect. C): 323.

Tinkham M (1996) *Introduction to Superconductivity*, 2nd edn. New York: McGraw Hill.

Tsuei CC and Kirtley JR (2000) Pairing symmetry in cuprate superconductors. *Reviews of Modern Physics* 72: 969.

Ullmaier H (1975) *Irreversible Properties of Type II Superconductors*. Berlin: Springer.

Wordenweber R (1999) Mechanism of vortex motion in high temperature superconductors. *Reports* of *Progress in Physics* 62: 187.

Superconductors, Metallic

D J Bishop, Bell Labs, Lucent Technologies, Murray Hill, NJ, USA

Introduction and Discussion of the Scope of This Article

Superconductivity is one of the most interesting macroscopic physical manifestations of the microscopic quantum mechanical world. In this remarkable phenomenon, the resistance of a material drops to zero at a finite temperature above absolute zero. The resistance drops to not just a small value but to a number which has been measured to be zero to many significant figures. Accompanying this drop in resistance are other extraordinary phenomena such as the Meissner effect whereby all magnetic fields are expelled from the interior of the sample, flux quantization and vortex lattices, the Josephson effect, and macroscopic quantum coherence. The richness and complexity of the phenomenon of superconductivity is unparalleled in solid-state physics.

In addition to its intrinsic scientific interest, the phenomenon gives rise to a very wide range of applications ranging from generating large magnetic fields to detecting tiny ones. Logic circuits to replace Si-CMOS, superconducting power lines, large motors and generators, magnetic resonance imaging, sensitive RF detectors, and much more, are all made possible by superconductivity.

In this article, the metallic superconductors are reviewed. Although a vast and wide range of superconducting materials and systems have been discovered over the last 100 years, the metallic superconductors still reign supreme in terms of economic importance and impact. Also, these materials are the best understood. In general, much of what is really known about the theory of superconductors applies to these materials. Most of the newer, more exotic, nonmetallic systems are still largely a scientific mystery. Clearly, in terms of applications, the metallic systems NbTi and Nb_3Sn dwarf all others in terms of tons used, money created, and impact on lives made. It is these materials that have allowed many Nobel prizes to be won by generating large magnetic fields for high-energy particle accelerators and solid-state experiments probing exotic states of matter at low temperatures and high magnetic fields. Magnetic resonance imaging of the type that has saved countless lives in modern hospitals is still only possible because of the metallic superconductors. Clearly, they are the workhorse materials of the field and will remain so for many years to come. This is because they are materials that are metals with all the good things that this implies. These materials are ductile, have good mechanical and thermal properties, are relatively inexpensive, can be joined by soldering and welding, are stable and nontoxic, and can be fabricated by reasonable, low-cost industrial processes. Although a number of newer materials such as the ceramic, high-T_c superconductors hold a lot of promise, to date this potential has yet to be realized on any commercial scale and given the large number of difficult manufacturing problems that still remain with these materials, it is not clear that they will ever replace the metallic materials for widespread application.

In this article, the focus is on the metallic superconductors to the exclusion of many other types of superconducting systems. Not considered here are such exotica as organic superconductors (those containing carbon), fullerenes and carbon nanotubes, molecular salts, borocarbides, polymers, perovskites and other ceramics, ruthenates, fluoro-argentates, diamond, and the ultraconductors. Briefly considered in the section "Exotic metallic superconducting systems" are such strange metallic systems as the heavy fermions, magnetic superconductors, and MgB_2. The basic goal here is to provide a broad overview of the metallic superconductors, their properties, materials, theory, and application. Clearly, it is a vast field that has filled libraries with information. What is presented here is just the barest introduction to the subject. Some references for further study are provided in the "Further reading" section.

Properties of the Metallic Superconductors

The most fundamental property of all superconductors, the one which gives rise to their name, is the zero resistance state. **Figure 1** (the first observation of superconductivity made by Kamerlingh Onnes in 1911) shows his measurement of the resistance of the liquid metal mercury as a function of temperature. At roughly 4.2 K, the resistance drops to a value indistinguishable from zero. This zero-resistance state is the property that many of the applications seek to exploit. In systems such as superconducting power lines, electric motors and generators, and large magnets, this zero-resistance state allows large currents to flow without the normal joule heating that would occur in a conventional metal with a finite resistance.

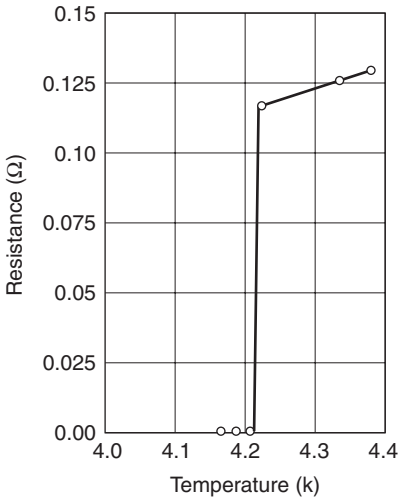


Figure 1 Data from Onnes, which is the first observation of superconductivity. It is a plot of the temperature dependence of the resistance of mercury. At 4.2 K the abrupt drop is the onset of the superconducting state.

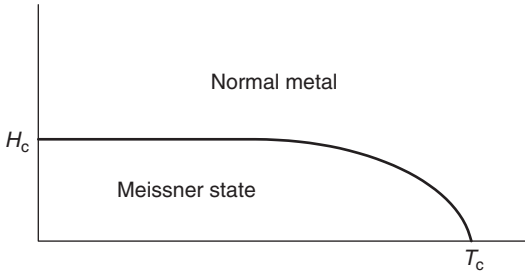


Figure 2 Phase diagram for a type I superconductor. Below T_c and H_c is the Meissner phase. For temperatures or fields above these, one regains normal metallic behavior.

All superconductors, including the metallic ones, fall into one of two broad categories. The first of these categories of superconductors are type I superconductors. Most elemental superconductors are of this type including mercury as shown in **Figure 1**. The phase diagram for these types of superconductors is given schematically in **Figure 2**. The figure shows an H–T phase diagram indicating the range of temperatures and fields for which superconductivity can occur. There are two key points on the graph, T_c and H_c. T_c is the temperature where, in zero applied magnetic field (and low measuring currents), superconductivity occurs. In general, for type I superconductors, T_c is relatively low with lead at 7.2 K being the highest known value (at ambient pressure). The other key point on the graph is H_c which is the field where, at zero (or very low) temperatures, an applied magnetic field destroys the superconductivity. For type I superconductors H_c also tends to be low, typically below 1000 G. These low critical temperatures

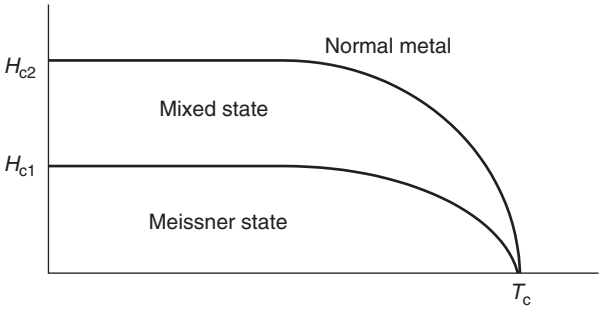


Figure 3 Phase diagram for a type II superconductor. There are two superconducting phases, the Meissner phase and the mixed state.

and low critical fields conspire to make type I superconductors not very interesting from an applications point of view.

One of the most interesting phenomena associated with type I superconductors is the Meissner effect. This effect arises because magnetic fields tend to destroy the superconducting state. This can be seen qualitatively by looking at **Figure 2** and noting that the higher the applied magnetic field, the lower the transition temperature for a type I superconductor. There are two types of superconductors, depending on how they behave under applied magnetic fields. In the first method, the system maximizes the total free energy associated with the superconducting state by totally expelling the magnetic field from the interior of the type I sample. This state of total diamagnetism was first discovered by Meissner and Ochsenfeld in 1933 and is now called the Meissner effect. The effect is large and easy to measure with relatively simple devices and is a common way to demonstrate that one has a bulk superconductor, not merely a system whose resistance has dropped to a small value which may or may not indicate superconductivity.

The other broad category of superconductors are type II superconductors. Most known superconductors, including most metallic superconductors, are of this type. Most of these superconductors are either metallic alloys or compounds. The elements vanadium, technetium, and niobium are the only known elemental type II superconducting materials (under ambient pressure). The phase diagram for type II superconductors is shown schematically in **Figure 3** and is more complicated and interesting than that found for type I superconductors. Below T_c and for fields less than a critical field called H_{c1}, there is a Meissner phase similar to that found for type I materials. However, for fields between H_{c1} and H_{c2} there is a new phase of matter called the mixed state. This arises because of the previously mentioned antipathy that the superconducting state has for applied magnetic fields. Whereas for small fields the type II material can

expel the magnetic field completely from the interior of the sample, for large fields this is no longer possible. But rather than allowing magnetic fields to enter the sample completely and uniformly as it would for a piece of copper, for example, the type II superconductor finds it energetically favorable to allow the field to penetrate in single magnetic flux lines called magnetic vortices. These lines of magnetic flux are exactly quantized to a precise value – all lines are identically the same. These lines can form a regular lattice, most typically hexagonal in symmetry although many other symmetries have been observed.

Type II superconductivity appears to be quite a robust state with observed T_c values as high as 138 K for an oxide superconductor. H_{c2}, the maximum magnetic field at which the sample is still superconducting, can also be very high with observed values in excess of a megagauss for some of the oxide superconductors. Among metallic superconducting materials such as NbTi ($T_c = 9.8$ K) and Nb_3Sn ($T_c = 18.1$ K), the critical fields are still very high allowing for a robust range of application.

Another characteristic property of superconductors that is very important is the critical current, J_c. J_c is the maximum current a sample can carry and still remain superconducting. This property is a measure of the current-carrying capacity of the superconductor. Typical values for a good system may reach one to ten million amps per square centimeter. In general, J_c tends to be not so much an intrinsic property of the material but an extrinsic one that depends on many things such as processing, thermal history, doping with impurities, and mechanical history. The larger the J_c, the more current can be carried by the system and the more useful the material is for applications such as magnets and motors. Only type II superconductors have usefully large J_cs. The physics of J_c is related to the magnetic vortex lattice mentioned above. Applied currents exert a Lorentz force on the flux lines. If they are pinned, no energy is dissipated and the superconducting state persists. However, if the lines can move, energy is dissipated as they move, heating the sample and destroying the superconducting state. The tricks used to pin the flux lines are many and complex and tend to be developed empirically. The techniques are similar to those developed to pin dislocation lines in metals to make them less ductile. For example, alloying is a common approach to both problems. Critical currents for type I superconductors tend to be very low because of the low value of H_c.

Theories of Metallic Superconductors

In general, the theoretical understanding of conventional metallic superconductors is quite good. This is to be contrasted with the rather poor state of understanding of theory for other types of superconductors such as the high-T_c cuprates. Most of the metallic superconductors (with the exotic ones discussed in the section "Exotic metallic superconducting systems" being the exceptions) are well described by the BCS theory developed by Bardeen, Cooper, and Schrieffer in 1957 and its extensions to type II materials developed by Abrikosov. All four were awarded the Nobel prize for their contributions to the understanding of these materials.

Superconductivity theory will be discussed elsewhere in this encyclopedia and so only a brief introduction is given on the most crucial concepts here. The key and basic idea behind the BCS theory is that of electron pairing. It had been understood by the London brothers in the 1930s that paired electrons could produce a gap at the Fermi surface and thereby produce superconductivity. A key intellectual challenge for the later theorists was to figure out how to get the normally repulsive electrons to pair. The BCS theory did this and showed how electrons of opposite momentum and spin could pair together giving rise to the zero-resistance state. The trick was to consider the underlying crystal lattice. In their theory, pairing of the electrons occurred because of the interaction of the electrons with the ionic crystal lattice of the material, thus creating an effective attractive interaction overcoming the normally repulsive Coulombic forces between electrons. A simple and naïve way to think about it is to imagine two bowling balls rolling around on a mattress producing an effective "attraction" between the balls. While wrong in all the details, this image is still an instructive way to understand in an intuitive way how a medium can effectively produce an attractive interaction between objects when nothing exists between them in the absence of the medium. A more correct way to think about the problem is to imagine that as the first electron moves through the lattice, its interactions with the positively charged ions produce a distortion in the crystal lattice that attracts the second electron, thereby creating an effective attractive force between the electrons. Within a few years of the publication of the BCS theory, many properties predicted and explained had been experimentally observed. Properties such as the energy gap, ultrasonic attenuation, isotope effect, the flux quantum, and Josephson tunneling (predicted by Josephson in 1962) were resounding confirmation of the theory. The isotope effect is a particularly striking example of BCS theory providing a predictive understanding. The theory predicted that changing the mass of an ion in the underlying lattice but not its electronic properties by using an isotope of the atom would change T_c in a

predictable and well-explained way. In general, these isotope effect predictions work well in the metallic superconductors, especially the simple metals, and confirm the BCS theory as the correct answer.

The work of the London brothers, Ginsburg and Landau, BCS, Abrikosov, and many others have led to the current theoretical understanding of conventional (e.g., metallic) type II superconductors. The basic idea that one needs to master to be able to understand many of the applications, which will be discussed later, is that the microscopic theory yields two important length scales in a superconductor. The first of these is called the coherence length, typically denoted by ξ. In the BCS theory, this is something like the size of the superconducting pair of electrons. Within the Ginsburg–Landau theory, it can be qualitatively understood as the spatial distance over which the superconducting order parameter can change. The second important length scale is called the penetration depth typically denoted by λ. This is defined in a number of different ways by the various theories, but in essence it is the distance over which magnetic fields can vary in a superconductor. In general, these lengths diverge (grow) with increasing temperature. For this discussion, zero-temperature values have been used. In this picture, type I and type II superconductors can be seen as the two limits of these fundamental microscopic lengths. In the limit where $\lambda > \xi/\sqrt{2}$ type II superconductivity is observed, and in the limit where $\lambda < \xi/\sqrt{2}$ type I superconductivity is observed.

An important advance was made by Josephson in 1962. He considered the situation in which one has two superconducting samples connected by a weak link through which electrons cannot normally travel. He considered the case of tunneling of the paired electrons through the junction and showed how this would produce a phase shift between the order parameters on each side of the junction. This phenomenon is the basis of an object called a superconducting quantum interference device (SQUID), the most sensitive magnetic field detector known. This effect is also the heart of the attempt to use superconducting devices to replace Si-CMOS as logic devices. The use of superconductivity in electronics generally employs some variation of the Josephson effect. This should be contrasted with large-current applications such as magnets and power lines.

This general theoretical framework outlined here very well describes the conventional metallic superconductors. In summary, the pairing mechanism is well known to be phonons (vibrations in the underlying crystal lattice), the two microscopic lengths can be related by theory to fundamental properties of the material, the various effects such as crystal structure and doping with impurities are well understood, and the microscopic understanding of the phenomena allows for a rational and predictive engineering of the materials for things like enhanced critical currents. The conventional metallic superconductors are therefore a solved theoretical problem. Understanding their properties and providing a detailed microscopic theory for their behavior has been arguably one of the greatest theoretical triumphs of physics in the twentieth century. Unfortunately, for nonmetallic superconductors such as the cuprates, this cannot also be said to be true.

Metallic Superconducting Materials

Type I Materials

Although there are over a thousand known metallic superconductors, metallic type I superconductors are only a few, several dozen under ambient pressure, with another couple of dozens that will superconduct only under high applied pressures. Essentially all the known type I superconductors are metallic elements and dilute alloys of these elements. In general, when heavily alloyed, these materials will tend to become type II superconductors. There are 27 type I superconducting elements under ambient pressure ranging from lead with a T_c of 7.2 K to rhodium with a T_c of 0.3 mK. **Table 1** lists these elements and their T_c values.

Under pressure, many more elements will superconduct. The highest known T_c for any type I material is phosphorus with a $T_c \sim 15$ K but it superconducts only under pressures of several mega-bars. While type I materials tend to be good metals, somewhat surprisingly, three of the best metals – copper, gold, and silver – do not appear to superconduct at all. This is due to the fact that their f.c.c. lattice structure is less susceptible to the kind of lattice distortion that gives rise to the BCS pairing mechanism responsible for superconductivity.

Because these type I superconductors tend to be good metals, their resistance with temperature follows the behavior shown in **Figure 1**. As the temperature is reduced, in the normal state the resistance drops due to a reduction in the thermally induced fluctuations in the ionic lattice that scatter the electrons. In the absence of superconductivity, their resistance would saturate at a nonzero value determined by the density of defects in the lattice as a result of impurities, and structural defects. However, for the materials that undergo a transition to superconductivity, this reduction of the resistance with decreasing temperature is interrupted by a sharp, precipitous drop to the zero-resistance state. Because

Table 1 List of known type I metallic superconductors (ambient pressure)

Element	*Symbol*	*Transition temperature* (K)
Aluminum	Al	1.2
Americium	Am	0.60
Beryllium	Be	0.02
Cadmium	Cd	0.52
Gallium	Ga	1.08
Hafnium	Hf	0.13
Indium	In	3.41
Iridium	Ir	0.11
Lanthanum	La	4.88
Mercury	Hg	4.15
Molybdenum	Mo	0.92
Osmium	Os	0.66
Lead	Pb	7.2
Platinum	Pt	0.002
Protactinium	Pa	1.4
Rhenium	Re	1.7
Rhodium	Rh	0.0003
Ruthenium	Ru	0.49
Tin	Sn	3.72
Tantalum	Ta	4.47
Thallium	Tl	2.38
Thorium	Th	1.38
Titanium	Ti	0.4
Tungsten	W	0.02
Uranium	U	0.20
Zinc	Zn	0.85
Zirconium	Zr	0.61

Adapted from "superconductors.org."

these materials tend to be clean metals, the superconducting transition is usually quite sharp with temperature. In fact, this very sharp, abrupt transition with temperature can be used as a very sensitive thermometer in devices such as bolometers to sense radiation. In a bolometer, the particle to be detected is absorbed in a small volume of material, the energy of the absorbed particle raises the temperature of this volume, and a small change in temperature causes a large change in resistance of the type I superconductor. Electrical circuits measure this change in resistance and "detect" the particle.

Type II Materials

The metallic type II superconductors are a much more diverse group of materials than the type I metals. Some of the simplest type II metallic materials are vanadium, technetium, and niobium. With the exception of these elements, all other metallic type II superconductors are either alloys or compounds (under ambient pressure). The earliest studied type II superconductors were lead–bismuth and lead–thallium alloys. In the early days, it was not clear that these alloys were a separate class of superconductors, the type II materials. It was thought that they were dirty type I materials that had a smeared out T_c and H_c. It was the work of Shubnikov (experimental) and Abrikosov (theory) that paved the way to the current understanding of type II materials as a separate and distinct class of materials with well-defined phase transitions into a Meissner state and a higher-field mixed state.

Table 2 List of some representative type II metallic superconductors

Superconductor	*Transition temperature* (K)
Pure elements	
Nb	9.25
Tc	7.80
V	5.40
Niobium-based compounds	
NbTi	9.8
Nb_3Sn	18.1
Nb_3Ge	23.2
NbN	16.1
Magnetic superconductors	
YPd_2B_2C	23
$ErNi_2B_2C$	10.5
$ErRh_4B_4$	10??
Heavy fermion superconductors	
UPt_3	0.48
UBe_{13}	0.87
$CeCu_2Si_2$	~0.5
URu_2Si_2	1.2
Other	
MgB_2	39

Adapted from "superconductors.org."

Table 2 contains some typical metallic type II superconductors. The number of known type II materials is too large to make it practical to compose a comprehensive list, and with new discoveries everyday such a list would be obsolete almost immediately. Before the discovery of the high-T_c materials in 1986, the record transition temperature was held by the compound Nb_3Ge with a T_c of 23 K.

The two most technologically important materials today, even years after the discovery of the high-T_c materials, are Nb_3Sn with a T_c of 18 K and NbTi with a T_c of 10 K. This is basically due to their metallic properties. While the high-T_c materials are ceramics that are quite difficult to work with, the Nb compounds are relatively simple metals. The oxide superconductors are brittle, contain many elements, and are sensitive to small changes in composition, tend to require careful oxygenization, are unstable chemically, can be chemically poisonous, and are expensive to manufacture. Since the niobium compounds are metals, they have none of the above flaws. They are strong, tough, ductile metals that contain

only two elements and are not particularly sensitive to the composition, are reasonably chemically stable, are not particularly poisonous, and are relatively cheap to manufacture. These materials can be readily engineered to have high critical currents. While their T_c and H_{c2} are not as high as the high-T_c materials, their advantages far outweigh their drawbacks for most of today's applications.

Applications of the Metallic Superconductors

Because of the many advantages of the metallic superconductors discussed above, the vast majority of the applications to date have used these materials. Wires of NbTi and Nb_3Sn are widely used in many places ranging from magnets for magnetic resonance imaging (MRI) to levitating trains. Some of these applications are discussed here.

Easily the most important application of superconductivity to date is for MRI. It is currently responsible for $\sim$75% of the worldwide market for superconductivity and this application is likely to continue to dominate the market for at least the next decade. In this technique, the spins in the object to be imaged are excited by an applied RF field in the presence of a large static uniform DC magnetic field and a magnetic field gradient. The decay of the spins in time can be sensitively detected and their presence or absence measured. Using this technique one can create a 3D map of specific elements. In hospitals, it allows images to be made with unprecedented clarity, and precision without using any harmful or invasive radiation (e.g., X-rays). By using tricks associated with the field gradients, one can image "slices" of the body to be observed. Clearly, this has become a very important technology and no modern hospital today is without an MRI machine.

Another very important application of conventional metallic superconductors is for scientific research. Research magnets of every size and shape have been built for solid-state research and numerous important, Nobel prize-winning discoveries have been made as a result. In addition to solid-state physics, conventional metallic superconductors have been used for the bending magnets in numerous particle accelerators and fusion machines around the world.

Other applications for conventional metallic superconductors include magnetic levitation (maglev) trains, superconducting power lines, motors and generators, fault limiters and transformers, superconducting computers and other circuits using Josephson junctions, SQUID low magnetic field sensors, X-ray and light detectors, energy storage systems, and systems for the magnetic separation of materials. The current market for superconductors is estimated to be $\sim$\$3 billion/year in 2003, growing to be $\sim$\$5 billion in 2010 and then to be $\sim$\$40 billion in 2020. Of this total, the current fraction of this market owned by conventional, metallic superconductors is well over 95%. Even in the year 2010, 24 years after their discovery, high-temperature superconductors will still represent less than a third of the total market. Conventional metallic superconductors will still be used in more than two-thirds of the applications and will continue to dominate the market well into the twenty-first century.

To date large-scale systems and devices such as MRI magnets and magnets for R&D have dominated the market. In the future, smaller-scale electronic applications will emerge. These include magnetic field sensors, high-speed specialized electronic circuits, magnetocardiography and magnetoencephalography. The heart of many of these devices is the Josephson junction. As for the large current applications, these new markets will also be dominated by metallic superconductors such as Niobium and NbN. It is likely to be many years before a reliable high-T_c technology evolves that will allow a large number of Josephson junctions to be used in a single device. Thus, even this new developing application for superconductivity is likely to still be dominated by the conventional metallic superconductors.

Exotic Metallic Superconducting Systems

Up until now, conventional metallic superconductors whose properties are well understood and well explained by current theories, specifically BCS, were discussed. In this section, a number of newer superconducting systems are discussed which, while certainly metallic, remain far from being understood and represent some of the interesting current frontiers of research in superconductivity. This list is by no means complete – it is an attempt to show some examples of interesting current research topics.

Heavy-Fermion Superconductivity

Heavy-fermion superconductors are metals made out of various rare earth or actinide elements. They are an extremely strange group of metals that have normal state properties that are as bizarre as their superconducting properties. First seen by Bucher in 1973, there are now roughly ten known heavy-fermion compounds. The most interesting thing about them is that in the normal state, the effective masses of the conduction electrons can be hundreds of times the

electronic mass. Thus, they have "heavy" fermions. In the superconducting state, they have been shown to be unconventional in that the pairing state is not *s*-wave, as for conventional metals. In a conventional superconductor, the superconducting gap opens up uniformly over the Fermi surface. In these materials it is known that the gap is strongly anisotropic, that is, large in some directions, small or zero in others. These materials are ones where superconductivity and magnetism, both ferromagnetic and antiferromagnetic, seem to live in an uneasy proximity. The general antipathy that the superconducting order parameter has for magnetism creates a wide range of subtle responses, and heavy-fermion superconductivity appears to be one of them. Despite more than 30 years of study, heavy-fermion superconductivity is far from being a well-understood phenomenon, either theoretically or experimentally. While few things are well understood in these systems, perhaps the best known fact is that the BCS theory does not work here as an explanation.

Magnetic Superconductors

The issues discussed in the previous section about the coexistence of magnetism and superconductivity apply here as well. There exist a variety of phases which include rare-earth elements where the materials go from normal metals to superconductors to magnets as the temperature is reduced. There are roughly half a dozen of these compounds known where, depending on the temperature and magnetic field, one of the three known ground states for metals – normal metal, superconductor, or magnet (either ferromagnet or antiferromagnet) – can be reached. Thus, the study of these materials is one of seeing the interplay of a variety of states that compete with each other giving rise to novel and unexpected phenomena. Clearly, superconductivity in these materials is unconventional in that a predictive understanding of the behavior is not there. The study of the interplay of magnetism in all its forms and superconductivity will continue to be a major theme of research in solid-state physics for decades to come.

MgB_2

Superconductivity in MgB_2 was discovered by J Akimitsu in 2001. This is an extremely interesting compound in that at first glance it would appear to be a relatively simple material. Much of the interesting physics in superconductivity arises in compounds containing rare earth, lanthanide, or actinide ions. MgB_2 contains none of those exotic ions that produce such complex magnetic behavior. Both magnesium and boron are relatively low atomic number elements that have pretty simple and well-understood properties as pure materials. However, when combined, they form a fairly high-T_c superconductor with a transition temperature of 39 K. This material is interesting for a number of reasons. It has been found to be a material with two gaps. The theory suggests that there are two separate Fermi sheets with a different gap for each of them – the first such observation in solid-state physics. Second, the superconducting mechanism appears to be BCS in origin, but with an unusual set of properties that conspire to give a very high T_c for a phonon-mediated mechanism. Finally, and perhaps most importantly, this is a relatively high-T_c material that has good metallic properties, unlike materials such as the cuprates or the fullerene superconductors that also have high T_c values but bring along a lot of baggage in terms of their mechanical and physical properties that make their application quite difficult. MgB_2 might be a material that has a high enough transition temperature to make the use of closed-cycle refrigerators feasible but also has material properties that are sufficiently forgiving to make it practical to use them for a wide range of applications. It may be the best compromise between the conventional low-T_c niobium-based materials and the high-T_c cuprates. A final and not irrelevant consideration is that both magnesium and boron are abundant and cheap and so this material should be very inexpensive to manufacture. Clearly, it will be at least a decade before all this plays out and the true impact of this material is known but it will be an interesting story to follow.

Conclusions and the Future

In this article, the metallic superconductors have been briefly reviewed. In general, these are the well-known and well-understood workhorses of the field of superconductivity. Most of the physics that is actually understood in the field of superconductivity applies to these materials. The BCS theory, the extensions to type II behavior, flux line lattices, Josephson junctions, and all the rest work very well for these materials. In terms of applications, the conventional metallic materials still own the lion's share of the market and will do so for decades to come. As an example, the fact that the cuprate superconductors are materials with anisotropic superconducting properties with short coherence lengths makes them very hard to engineer for applications – grain boundaries represent a serious impediment to the flow of critical currents. The standard materials such as NbTi and Nb_3Sn do not suffer from such problems and will always be easier to apply both for high-current applications as well as for electronic applications.

In terms of interesting new physics, the metallic superconductors remain an interesting area. New materials are being discovered at a steady rate and the field shows no signs of slowing down. Clearly, this will be an exciting and interesting field for decades to come.

See also: Superconductivity: Applications; Superconductivity: BCS Theory; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductivity: Tunneling.

PACS: 74.70.Ad; 74.70.Tx; 74.20.Fg; 74.25.-q; 74.62.Dh

Further Reading

Cardwell DA and Ginley DS (eds.) (2003) *Handbook of Superconducting Materials*. UK: IOP Publishing.
DeGennes P (1969) *Superconductivity of Metals and Alloys*. New York: Benjamin.
Evetts J (ed.) (1992) *Concise Encyclopedia of Magnetic and Superconducting Materials*. Oxford: Pergamon.
http://www.superconductors.org
Parks R (1969) *Superconductivity*. New York: Dekker.
Tampieri A and Celotti G (eds.) (2000) *Superconducting Materials*. Singapore: World Scientific.
Tilley DR and Tilley J (1990) *Superfluidity and Superconductivity*. London: IOP Publishing.
Tinkham M (1996) *Introduction to Superconductivity*. New York: McGraw-Hill.

Superfluidity

E V Thuneberg, University of Oulu, Oulu, Finland

Introduction

Fluids (gases and liquids) are distinguished from solids by the property that they can flow. In almost all cases, there is viscosity associated with the flow. Due to viscosity, the flow energy is gradually dissipated into heat. Contrary to this common situation, there is a special class of fluids, which can flow without viscosity. These are called superfluids and the phenomenon is called superfluidity. As a concrete example, consider a ring-shaped container filled with a superfluid, see **Figure 1**. Once the fluid is put into circular motion, it will continue to circulate and no energy is dissipated. The flow can continue as long as the conditions for superfluidity are satisfied.

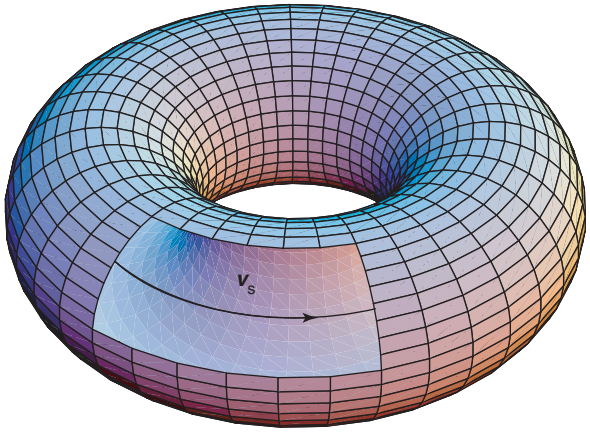


Figure 1 Once generated, the circulation of a superfluid (with velocity v_s) persists as long as the experiment can be continued.

Superfluids show many spectacular phenomena, which are discussed in sections "Hydrodynamics," "Quantization of circulation," "Rotating superfluid," and "Phase slip." Before going into these, the systems where superfluidity occurs (section "Occurrence") and the microscopic basis of superfluidity (section "Microscopic origin") are discussed. While section "Microscopic origin" gives deeper insight, it is not absolutely necessary for understanding the phenomena in the following sections.

Occurrence

Superfluidity occurs only in certain substances under special conditions. As a first case, liquid helium is discussed. Under standard pressure and temperature, helium is a gas. It liquifies at temperatures $\sim 4\,\mathrm{K}$. Cooling further down, it enters the superfluid phase at temperatures $\sim 2\,\mathrm{K}$, depending on pressure. The phase diagram of natural helium at low temperatures is shown in **Figure 2**. Natural helium consists essentially of isotope ^{4}He.

Helium has another stable isotope, ^{3}He. At temperatures below a few kelvin, its behavior is radically different from the isotope ^{4}He. It also becomes superfluid, but at temperatures that are a factor of 1000 smaller than for ^{4}He. The phase diagram of ^{3}He at low temperatures is shown in **Figure 3**. ^{3}He has three different superfluid phases, A, A_1, and B. The A_1 phase appears only in a magnetic field, and therefore is not visible in **Figure 3**.

Superfluidity is closely related to superconductivity. Superconductivity means that electric current can flow without resistance. This phenomenon appears at low temperatures in many metals like Al, Sn, and Nb. It arises from resistanceless motion of

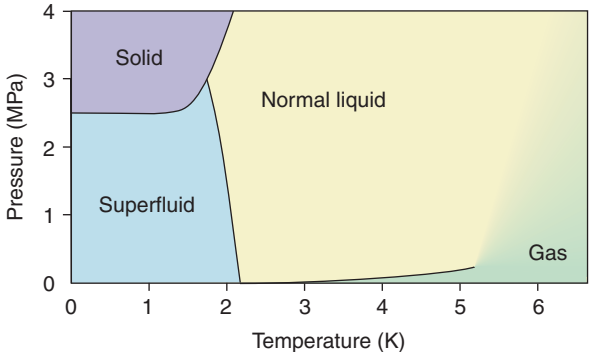


Figure 2 Phase diagram of ^{4}He at low temperatures. ^{4}He remains liquid at zero temperature if the pressure is below 2.5 MPa (~25 atm). The liquid has a phase transition to a superfluid phase, also known as He-II, at the temperature of 2.17 K (at vapor pressure).

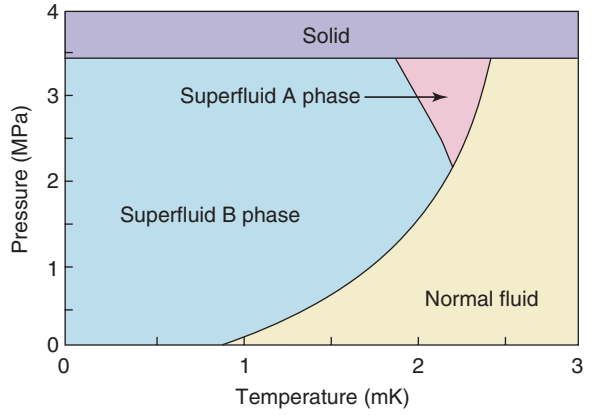


Figure 3 The phase diagram of ^{3}He at low temperatures. Note that the temperature is in units of millikelvin. Two superfluid phases of ^{3}He, A and B, are shown. (Figure based on data by DS Greywall.)

the conduction electrons in a metal. Therefore, superconductivity can be understood as superfluidity of the conduction electrons. Part of the discussion in this article applies also to superconductivity, but there are differences caused mainly by two reasons: (a) electrons have electric charge and therefore their motion is essentially coupled with magnetic fields; (b) the crystal lattice of the ions constitutes a preferred frame of reference, which does not exist for helium liquids.

With laser cooling, it is possible to cool certain atomic gases such as ^{87}Rb, ^{7}Li, ^{23}Na, and ^{1}H to very low temperatures. At low temperatures, Bose–Einstein condensation takes place in the gas. This state has many properties that are similar to superfluidity, although it is not a thermodynamically stable state, and therefore the flow cannot last for ever. Most of the discussion in this article applies also to condensed gases. An important difference is that instead of container walls for helium liquids, one has to consider the confining potential of the gas, which can be generated either magnetically (by field gradients) or optically (by laser beams).

Superfluidity is expected to occur also in astrophysical objects. The neutron liquid in a neutron star is believed to be in a superfluid state. This has been suggested as an explanation for the observed sudden changes in the rotation velocity of pulsars.

Microscopic Origin

In short, superfluidity can be explained as a quantum mechanical effect that shows up on a macroscopic scale.

Quantum mechanics is crucial in understanding the microscopic world. It explains that electrons in atoms have only discrete energies. There is no friction on the atomic scale, and the electrons can revolve around the nucleus without losing energy.

It is known that quantum mechanics rarely shows up on a macroscopic scale. Instead of quantum mechanics, macroscopic objects obey the rules of classical physics. The reason is that a macroscopic sample consists of large number of particles and, instead of individual particles, one can only observe their average behavior. Usually the particles are in different quantum states, and an average over them obeys the classical laws of physics. Examples of these laws are the Navier–Stokes equations for fluids and Ohm's law for electrical conduction.

Superfluidity is an exception to this general rule. In superfluids, a macroscopic number of particles is in the same quantum state. It follows that summing over particles does not lead to averaging, but produces a "macroscopic wave function."

Further analysis depends essentially whether the particles are bosons or fermions. ^{4}He and the gas atoms listed above are bosons, whereas ^{3}He atoms, electrons, and neutrons are fermions.

Consider a particle with mass m and momentum $\boldsymbol{p}$. Its energy is $E = \boldsymbol{p}^2/2m$. Its state is represented by the single-particle wave function

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \exp\left(\frac{\mathrm{i}}{\hbar}\, \boldsymbol{p} \cdot \boldsymbol{r}\right) \qquad [1]$$

where V is the volume of the system and $h = 2\pi\hbar$ the Planck constant. The wave function of a many-body system $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots)$ is more general and depends on the coordinates $\boldsymbol{r}_i$ of all particles.

Assume that the particles are bosons. This means that the total wave function must be symmetric when exchanging any pair of particles. In the case of two

particles, this means $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$. Further, assume that there is no interaction between the bosons. It can be shown that the occupation of the lowest-energy state becomes macroscopic, if the temperature T is less than

$$T_{\mathrm{BE}} = \frac{h^2}{2\pi m k_{\mathrm{B}}} \left(\frac{N}{2.612V} \right)^{2/3} \qquad [2]$$

where N is the number of particles and k_{B} the Boltzmann constant. This is known as Bose–Einstein condensation. The wave function [1] of the lowest-energy state ($\boldsymbol{p} = 0$) becomes macroscopic. At zero temperature, all particles are in this state.

While the ideal gas model explains Bose–Einstein condensation, it is quite insufficient in other respects. The interactions between particles are essential for the system to show superfluidity. In an interacting system, the macroscopically occupied state $\psi(\boldsymbol{r})$ need not be the lowest-energy state, and thus the macroscopic wave function can be nontrivial. In Bose gases (^{87}Rb, etc.) the interactions are weak, and a quantitative description can be achieved by the relatively simple Gross–Pitaevskii equation. In ^{4}He the interactions are much stronger, and a quantitative theory is not easily achieved.

Next, consider the fermions. Fermions have spin, which has to be described by an additional index s. Here, one needs to consider only spin-half particles, where s takes two values, $s = \pm\frac{1}{2}$. The wave function of a fermion system is $\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2, \ldots)$, and it has to be antisymmetric in the exchange of any pair of particles. For a two-particle state, this means

$$\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = -\Psi(\boldsymbol{r}_2, s_2, \boldsymbol{r}_1, s_1) \qquad [3]$$

This implies that the occupation of any single-particle state can only be zero or one. This is known as the "Pauli exclusion principle." Thus, macroscopic occupation of a single-particle state [1] is not possible.

Superfluidity in a fermion system can appear as a result of an attractive interaction between particles. Such an interaction can cause formation of pairs. Each pair has to satisfy the antisymmetry condition [3]. However, a pair is a unit that behaves like a boson. In particular, it is not excluded that several pairs are in the same pair state. Superfluidity in fermion systems can be understood as a macroscopic occupation of a single-pair state.

The spin part of the pair wave function has four different possibilities. These can be classified as a singlet state

$$\uparrow\downarrow - \downarrow\uparrow \qquad [4]$$

(which is a compact notation for $\delta_{s_1,1/2}\delta_{s_2,-1/2} - \delta_{s_1,-1/2}\delta_{s_2,1/2}$) and three triplet states, which can be chosen as

$$-\uparrow\uparrow + \downarrow\downarrow,\ i(\uparrow\uparrow + \downarrow\downarrow),\ \uparrow\downarrow + \downarrow\uparrow \qquad [5]$$

Consider the case of spin singlet [4]. The pair wave function in this case is assumed to be of the form

$$\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = \psi\left(\frac{\boldsymbol{r}_1 + \boldsymbol{r}_2}{2}\right)\chi(\boldsymbol{r}_1 - \boldsymbol{r}_2)(\uparrow\downarrow - \downarrow\uparrow) \qquad [6]$$

where the orbital wave function has been separated into a center of mass part ψ and a relative part χ. The singlet spin state [4] is antisymmetric in the exchange of the two spins. In order to satisfy pair antisymmetry [3], the corresponding orbital part $\chi(\boldsymbol{r}_1 - \boldsymbol{r}_2)$ has to be symmetric. In most superconductors, the pair wave function is of the form [6]. In majority of them (Al, Sn, Nb, …), χ is approximately independent of the direction of $\boldsymbol{r}_1 - \boldsymbol{r}_2$. This is called s-wave pairing in analogy with s, p, d, etc. atomic orbitals. In high-T_{c} superconductors, there is strong evidence of d-wave symmetry of χ.

Another alternative is that the spin state of a pair is triplet [5]. This case is realized in ^{3}He (and possibly in some superconductors). In ^{3}He, the orbital wave function is of p-type. There are three degenerate p-wave states p_x, p_y, and p_z. The pair wave function can be written as

$$\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = \sum_{j=1}^{3}\sum_{\mu=1}^{3} \psi_{\mu j}\left(\frac{\boldsymbol{r}_1 + \boldsymbol{r}_2}{2}\right) p_j(\boldsymbol{r}_1 - \boldsymbol{r}_2) i\sigma_\mu\sigma_2 \qquad [7]$$

Here $i\sigma_\mu\sigma_2$ denotes the same three spin states as in eqn [5], but expressed using Pauli spin matrices σ_i.

The macroscopic wave function of bosons is called "order parameter," since it describes ordering of the particles and it vanishes in the normal fluid phase. For fermions, the same role is played by the center of mass part of the pair function. This is the soft degree of freedom, which can change as a function of time and location, whereas the other parts in the pair wave function [6]–[7] are fixed. One can see that the order parameter in ^{4}He, and in most superconductors is a complex-valued scalar ψ, but in ^{3}He, it is a 3×3 matrix $\psi_{\mu j}$.

The quantitative theory of fermion superfluids is based on the Bardeen–Cooper–Schrieffer theory of superconductivity. Many properties can also be described by the simpler Ginzburg–Landau theory.

Hydrodynamics

Many properties of superfluids can be understood in terms of the "two-fluid model." The basic assumption

is that the liquid consists of two parts. These are called the superfluid and normal components. The current density $\boldsymbol{j}$ can be represented as a sum

$$\boldsymbol{j} = \rho_s \boldsymbol{v}_s + \rho_n \boldsymbol{v}_n \quad [8]$$

where ρ_s and $\boldsymbol{v}_s$ are the density and velocity of the superfluid component and ρ_n and $\boldsymbol{v}_n$ are the corresponding quantities for the normal part. The liquid density is the sum of the two densities, $\rho = \rho_s + \rho_n$. The superfluid component can flow without viscosity and it carries no heat or entropy. Moreover, it is curl free,

$$\nabla \times \boldsymbol{v}_s = 0 \quad [9]$$

(This is valid only in uncharged superfluids.) The normal component behaves more like a usual viscous fluid.

The two-fluid model can be justified from the microscopic theory discussed in the section "Microscopic origin." The superfluid component corresponds to particles in the macroscopic wave function, and the normal component to particles in other single-particle states. The densities of the two components depend on temperature. With increasing temperature, $\rho_s(T)$ drops continuously from $\rho_s(0) = \rho$ and vanishes at the superfluid transition temperature T_c.

The two-fluid model can explain many properties of superfluids. In particular, the existence of frictional forces depends on the type of experiment. On one hand, the flow in a ring-shaped container persists because it is only the superfluid component that flows (**Figure 1**). On the other hand, nonvanishing viscosity is measured with a rotational viscometer, where the superfluid is placed between two coaxial cylinders that rotate at different angular velocities. Here, the normal component is driven into motion and causes dissipation.

Superfluids show peculiar mixing of thermal and mechanical properties. Consider a superfluid in a channel which is heated at one end, see **Figure 4**. The superfluid component is attracted to the hot region because the chemical potential is lower there. As a consequence, a pressure difference appears. This drives the normal component in the direction of decreasing temperature and convects the heat away from the source. Assuming the geometry does not allow net mass transfer, the mass transported by the normal and superfluid components in opposite directions are equal in magnitude.

In addition to usual sound wave, superfluids have another propagating mode. This second sound is an oscillation where normal and superfluid components move in opposite directions. This leads to an oscillation of temperature whereas the density remains nearly constant. The second sound can be generated by heating the superfluid periodically, and standing waves of temperature have been demonstrated experimentally.

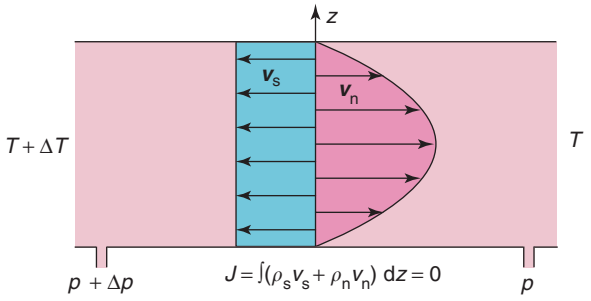


Figure 4 A difference in temperature generates flow of normal and superfluid components in opposite directions, and a pressure difference appears. The viscosity of the normal component causes $\boldsymbol{v}_n$ to vanish at walls. The superfluid velocity $\boldsymbol{v}_s(z)$ has to be constant in order to be curl free [9].

In addition to the mass current $\boldsymbol{j}$, there can be persistent spin currents. This is possible in superfluids whose order parameter is more complicated than scalar (^{3}He). The spin current is described by a tensor $j^{\text{spin}}_{\mu j}$. The index $\mu = x, y, z$ indicates the direction of the spin angular momentum that is flowing, and $j = x, y, z$ indicates the direction of the flow. Even in equilibrium, the order parameter of ^{3}He has a nontrivial spatial variation called "texture." This is associated with persistent spin currents and, in case of ^{3}He-A, also with persistent mass currents.

Quantization of Circulation

Consider a superfluid with order parameter ψ. (Assume an uncharged superfluid, ψ can be either scalar or matrix.) The superfluid velocity $\boldsymbol{v}_s$ can be expressed as a function of the order parameter as

$$\boldsymbol{v}_s = \frac{\hbar}{M} \nabla \phi \quad [10]$$

where $\phi(\boldsymbol{r})$ is the phase of the order parameter, $\psi(\boldsymbol{r}) = A e^{i\phi(r)}$, and the amplitude A is assumed to be constant. M is the boson mass, that is, the mass of a particle in a boson superfluid and the mass of a pair in a fermion superfluid. Equation [10] can be justified starting from the expression of current in quantum mechanics.

An alternative form of eqn [10] is obtained by taking the line integral along a closed path,

$$\oint \boldsymbol{v}_s \cdot d\boldsymbol{l} = N \frac{h}{M} \quad [11]$$

Here the property that ϕ is defined modulo 2π, and N is an integer, is used. Equation [11] is known as

"quantization of circulation." The curl-free condition [9] is a direct consequence of eqn [10] or [11].

Consider again a superfluid in a ring-shaped container (**Figure 1**). By applying eqn [11] to a path in the ring, one can see that, in addition to being persistent, the superfluid velocity can only have discrete values. A similar phenomenon in superconductors is flux quantization.

Rotating Superfluid

Consider a superfluid in a container that is rotated with angular velocity $\boldsymbol{\Omega}$. The normal component will follow this motion because of its viscosity. In equilibrium, it rotates uniformly with the container, $\boldsymbol{v}_n = \boldsymbol{\Omega} \times \boldsymbol{r}$. This is not possible for the superfluid component because it has to be curl free [9]. (Equation [9] should be compared to $\boldsymbol{\nabla} \times \boldsymbol{v}_n = 2\boldsymbol{\Omega}$.)

The rotating state of a superfluid is most commonly realized by "vortex lines." On a path around the vortex line, the phase ϕ changes by 2π (or an integral multiple of it). This is illustrated in **Figure 5**. Equivalently, the circulation of superfluid velocity [11] around the vortex line is h/M. Assuming cylindrical symmetry, the phase ϕ is the same as the azimuthal angle in the cylindrical coordinate system (r, ϕ, z). The velocity field can be calculated from eqn [10]:

$$\boldsymbol{v}_s = \frac{\hbar}{Mr}\,\hat{\boldsymbol{\phi}} \qquad [12]$$

where $\hat{\boldsymbol{\phi}}$ is a unit vector in the azimuthal direction.

The structure of the rotating state is determined by minimum of free energy. The rotation of the container is taken into account by minimizing $F = F_0 - \boldsymbol{L} \cdot \boldsymbol{\Omega}$, where F_0 is the free-energy functional in the stationary case and $\boldsymbol{L}$ the angular momentum. In the two-fluid model, this reduces to

$$F = \int \mathrm{d}^3 r\, \frac{1}{2}\, \rho_s (\boldsymbol{v}_s - \boldsymbol{v}_n)^2 + \text{constant} \qquad [13]$$

Thus, the optimal solution corresponds to $\boldsymbol{v}_s$ as equal as possible to $\boldsymbol{v}_n = \boldsymbol{\Omega} \times \boldsymbol{r}$, but subject to condition [10]. This is achieved by a regular array of vortex lines. The number of vortex lines n per unit area is determined by the condition that the circulations of normal and superfluid velocities are the same over an area containing many vortex lines. This yields

$$n = \frac{2M\Omega}{h} \qquad [14]$$

There are $\sim$1000 vortex lines in a circular container of radius 1 cm that is rotating 1 round per minute.

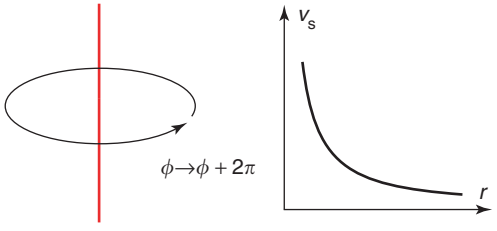


Figure 5 The vortex line and the magnitude of the velocity field [12] around it.

Vortex lines in an uncharged superfluid are analogous to flux lines, which occur in type II superconductors. Flux lines of superconductors appear in a magnetic field, which is analogous to the rotation of an uncharged superfluid.

The velocity field [12] of a vortex diverges at the vortex line. Thus, there must be a vortex core, where the two-fluid description is insufficient. A finite energy in the vortex core is achieved if the amplitude of the order parameter vanishes at the vortex line. This is the case for a scalar order parameter. For a matrix order parameter, it is not necessary that all components of the matrix vanish at the line. Such vortex lines are realized in superfluid ^{3}He-B.

The quantization of the superfluid velocity [11] is not always true for uncharged superfluids. This happens when there is an additional contribution to the superfluid velocity [10] coming from the matrix form of the order parameter. Such a case is realized in superfluid ^{3}He-A, and a careful reanalysis of the rotating state is needed. It turns out that, in addition to one-dimensional vortex lines, the vorticity may be arranged as two-dimensional vortex sheets and three-dimensional textures. All these have been confirmed experimentally. In any case, a homogeneous rotation of the superfluid is excluded.

Phase Slip

Consider superflow in a channel under thermal equilibrium ($\boldsymbol{v}_n = 0$). The maximum supercurrent is determined by a process called phase slip. Consider that a short piece of vortex line is nucleated at a surface on one side of the channel. This vortex expands, goes through the whole cross section of the channel, and finally disappears on the other side. As a result of this process, the phase difference $\Delta\phi$ between the ends of the channel has changed by 2π. Part of the superfluid kinetic energy is dissipated in the motion of the vortex. This means that the flow ceases to be dissipationless above a critical velocity for phase slips. Phase slips take first place in constrictions of the flow channel, where the superfluid velocity has its maximum value.

A special type of phase slip takes place in very short constrictions, where eqn [10] ceases to be valid. An ideally short constriction shows the Josephson

effect, where the supercurrent J_s depends on the phase difference $\Delta\phi$ as

$$J = J_c \sin(\Delta\phi) \quad [15]$$

and J_c is a constant. Moreover, the time derivative of $\Delta\phi$ is proportional to the difference of the chemical potential $\Delta\mu$ on the two sides of the constriction,

$$\frac{d\Delta\phi}{dt} = -\frac{2\Delta\mu}{\hbar} \quad [16]$$

Combining the two equations, one sees that a constant $\Delta\mu$ generates an oscillating current at the frequency $2\Delta\mu/h$.

The Josephson effect takes place in all superfluids, and has extensively been studied in superconductors. In helium superfluids, it is more difficult to fabricate constrictions that are small enough, but this has been achieved recently.

See also: Bose–Einstein Condensation; Magnetism, History of; Quantum Mechanics: Methods; Statistical Mechanics: Classical; Superconductivity: BCS Theory; Superconductivity: Flux Quantization; Superconductivity: General Aspects; Superconductivity: Ginzburg–Landau Theory and Vortex Lattice; Superconductivity: Tunneling; Superconductors, High T_c.

PACS: 67.40. – w; 67.57. – z; 74., 03.75. – b

Further Reading

Andronikashvili EL and Mamaladze YuG (1966) Quantization of macroscopic motions and hydrodynamics of rotating helium II. *Reviews of Modern Physics* 38: 567.

Davis JC and Packard RE (2002) Superfluid ^{3}He Josephson weak links. *Reviews of Modern Physics* 74: 741.

Dobbs ER (2001) *Helium Three*. Oxford: Oxford University Press.

Donnelly RJ (1991) *Quantized Vortices in Helium II*. Cambridge: Cambridge University Press.

Leggett AJ (1975) A theoretical description of the new phases of liquid ^{3}He. *Reviews of Modern Physics* 47: 331.

London F (1954) *Superfluids*, vol. II. New York: Wiley.

Lounasmaa OV and Thuneberg EV (1999) Vortices in rotating superfluid ^{3}He. *Proceedings of National Academy of Sciences USA* 96: 7760.

Tilley DR and Tilley J (1990) *Superfluidity and Superconductivity*, 3 edn. Bristol: IOP Publishing.

Vollhardt D and Wölfle P (1990) *The Superfluid Phases of Helium 3*. London: Taylor and Francis.

Wheatley JC (1975) Experimental properties of superfluid ^{3}He. *Reviews of Modern Physics* 47: 415.

Wilks J (1967) *The Properties of Liquid and Solid Helium*. Oxford: Clarendon.

Surfaces and Interfaces, Electronic Structure of

G Chiarotti and C Goletti, Università di Roma "Tor Vergata," Rome, Italy

Introduction

The presence of the surface modifies significantly the atomic and electronic structure of the first few layers of the solid. Since some of the properties of the solid are to some extent properties of its surface (e.g., the work function, the photoelectric threshold, chemisorption and physisorption, conductivity in the surface channel(s), optical reflectance, and the scattering of electrons, ions, and atoms by the solid), a detailed knowledge of the surface is important for solid-state physics and chemistry as well as for materials science and technical applications.

In this article, the discussion is limited to the so-called "clean surfaces," or surfaces that are atomically clean (i.e., with a controlled chemical composition) and crystallographically well defined. "Real surfaces" (i.e., surfaces with unspecified chemical impurities, crystallographic defects, macroscopic faults, etc.) are considered only occasionally.

Clean surfaces are obtained in ultrahigh vacuum (UHV) conditions (10^{-10} torr or better) by cleavage, ion bombardment and annealing (IBA), high-temperature annealing, molecular beam epitaxy (MBE), etc. Conversely, real surfaces are obtained and handled at ordinary pressures, mechanically ground, chemically etched, etc., and their properties often depend upon the way of preparation.

This article discusses the surface reconstruction and relaxation, electronic surface states, examples of reconstruction of relevant semiconductor surfaces and interfaces.

Surface Reconstruction and Relaxation

The structure of an "ideal surface" is that of a half crystal obtained by dividing the crystal in two halves with a lattice plane as a boundary, removing all atoms on one side and leaving the other atoms in their original positions. An ideal surface is then characterized by the Miller indices of the boundary plane.

It should be considered, however, that the atoms in the ideal surface plane experience a force different from that of their partners in the bulk. Moreover, they have a larger freedom of movement so that

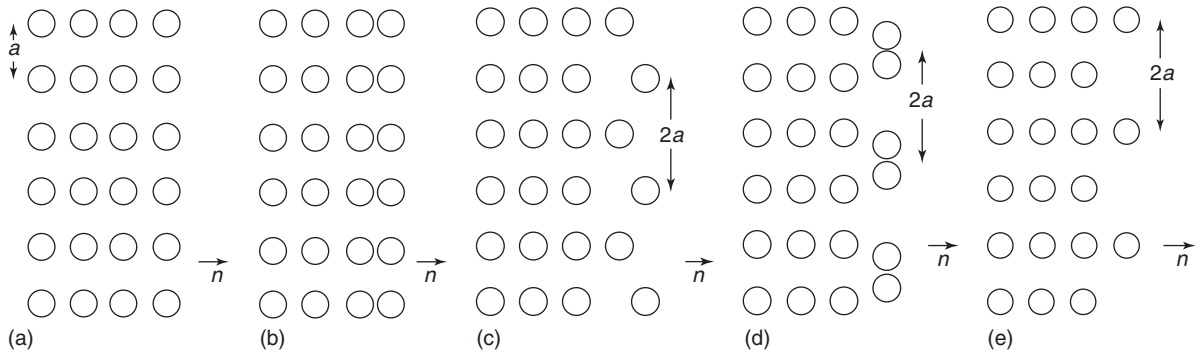


Figure 1 Schematic representation of a surface: (a) ideal surface; (b) relaxed surface; (c), (d), and (e) reconstructed surfaces (simplified examples of buckling, dimerization, and missing row models).

some rearrangement takes place due to the reduction of the free energy of the surface. Two deviations from the ideal surface are possible: relaxation and reconstruction.

Relaxation consists of a rigid displacement of all the atoms of the surface plane (in general, perpendicularly to it) that leaves unaltered the translational symmetry of the surface. On the contrary, reconstruction is a displacement of individual atoms that changes the translational symmetry in the surface plane.

The two situations are schematically shown in **Figure 1** (where it is assumed that the planes are stacked on top of each other in order to obtain the three-dimensional (3D) crystal). It is seen that in case (b), which represents relaxation, the surface unit cell is still a square of side a, while in cases (c), (d) and (e), which represent reconstruction, the surface unit cell is a rectangle of sides a and $2a$. In Wood's notations, these last reconstructions are called 2×1. It is seen that Wood notations do not univocally characterize the surface: they only give the symmetry of the reconstruction. As such, they are directly obtained from the LEED patterns.

The experimental determination of the actual coordinates of the atoms of the first few layers (that would completely characterize the reconstructed surface) is a very complex problem that may require the use of various techniques and the comparison with theoretical models. Among the most common techniques used for this purpose are: dynamical LEED, STM (scanning tunneling microscopy), ARUPS (angle-resolved ultraviolet photoemission), inverse photoemission, optical reflectance with polarized light, high-energy electron diffraction, elastic and inelastic scattering of atoms, ions, and electrons, etc.

Energies associated with reconstruction are $\sim 10^{-2}$ eVatom^{-1} for metals and $\sim 10^{-1}$ eVatom^{-1} for semiconductors. Metals are then seldom reconstructed, while semiconductors are very often so. In metals, in fact, electrons are delocalized so that the free energy does not depend much on the position of the single atoms. The opposite is true in semiconductors where directional bonding is significant.

Tables 1 and **2** list a few examples of reconstruction of metals and semiconductors. It is seen that, for a given surface, various reconstructions are possible, depending upon the method by which the surface is obtained (shown in the last column). In a few instances, reversible transformations between two phases of reconstruction have been observed as a function of temperature. On the other hand, impurities and defects may stabilize some phases.

The tables show clearly that reconstruction is a problem of extreme variety and complexity, as well as of the greatest interest. Up to now, a theory that explains the occurrence of the various types of reconstruction is not available from first principles.

Guidelines and heuristic models are discussed in the section "Examples of reconstruction of relevant semiconductor surfaces."

Electronic Surface States

The structural variations associated with reconstruction have important consequences on the electronic properties of the surface. Regardless of this, however, the reduction of the translational invariance caused by the surface is enough to introduce additional levels in the electronic structure of the solid. This was shown, for a Kronig–Penney potential, by I Tamm in 1932 and extended to a more realistic case by W Shockley in 1939.

A 1D potential with a surface at $z = 0$ is shown schematically in **Figure 2a**. Inside the solid ($z < 0$) the wave function is given, according to the Bloch theorem, by

$$\Psi(z,t) = u_k(z)\exp[\mathrm{i}(kz - \omega t)] \qquad [1]$$

Table 1 Surface structure of (some) metals

Metal	*Face*	*Reconstruction*	*Model*	*Preparation and remarks*
Ag	(1 0 0)	1×1	Relaxation	Epitaxial film on MgO. 1×1 stable for $T < 773$ K
	(1 1 0)	1×1	Relaxation (7–10%)	Roughening transition at $T \cong 700$ K
	(1 1 1)	1×1	Relaxation	$30 < T < 140$ K
Au	(1 0 0)	5×20	One layer with hexagonal symmetry	IBA + ann. at $T > 373$ K. Equilibrium configuration
		1×5	Slightly distorted hexagonal structure with buckling	Epitaxial film on MgO
	(1 1 0)	1×2	Missing row	$T < 670$ K
	(1 1 1)	23×1	Three domains rotated by 120°; Compression along [1 1 0]	22×1, 21×1 Structures also observed
Cu	(1 0 0)	1×1	Relaxation -1%	$300 < T < 1300$ K
	(1 1 0)	1×1	Relaxation $\cong 8\%$	$300 < T < 1000$ K
	(1 1 1)	1×1	Relaxation $\cong 2\%$	
Pt	(1 0 0)	1×1	Unrelaxed	Metastable, $T < 400$ K
	(1 1 0)	1×2	Missing row	
	(1 1 1)	1×1	Relaxation 2.3%	
W	(1 0 0)	$c(2 \times 2)$	In-plane zig-zag displacements	$T < 200$ K
	(1 1 0)	1×1	Unrelaxed	

Table 2 Surface structure of (some) semiconductors

Semiconductor	*Face*	*Reconstruction*	*Model*	*Preparation and remarks*
Diamond	(1 0 0)	2×1	Dimers	Polishing + ann. $T > 1300$ K
	(1 1 0)	1×1		Polishing + ann.
	(1 1 1)	2×1	π-Bonded chains	Cleavage; Polishing + ann. $T > 1200$ K
Si	(1 0 0)	2×1	Buckled dimers	IBA/MBE
		$p(2 \times 2)$	Ordered arrangement of buckled dimers	MBE; IBA. Present locally depending on the temperature
		$c(4 \times 2)$	Ordered arrangement of buckled dimers	MBE; IBA. Present locally depending on the temperature
	(1 1 0)	16×2	Dimers + adatoms (?)	Ann. $T > 1300$ K 32×2 structures also observed
		4×5	Uncertain	IBA + ann. Impurity stabilized (Ni ?)
		5×1	Uncertain	IBA + ann. Impurity stabilized (Ni ?)
	(1 1 1)	2×1	π-Bonded chains	Cleavage. Transforms into 7×7 at $T > 550$ K
		7×7	DAS (dimers–adatoms–stacking faults)	MBE; IBA + ann.; cleavage + ann. $T > 550$ K
		5×5	DAS	MBE. Present locally near steps
Ge	(1 0 0)	2×1	Dimers	IBA; MBE
		$p(2 \times 2)$	Ordered structure of dimers	MBE. Present locally
		$c(4 \times 2)$	Ordered structure of dimers	MBE. Present locally
	(1 1 1)	2×1	π-Bonded chains	Cleavage. Transforms into $c(2 \times 8)$ at $T > 400$ K
		$c(2 \times 8)$	Adatoms	MBE; cleavage + ann. at $T > 400$ K
GaAs	(1 0 0)Ga	$(4 \times 2), c(8 \times 2)$	Complex structure with dimers in the second layer	MBE; IBA
	(1 0 0)As	2×4	Dimers	MBE
	(1 1 0)	1×1	Rotation–relaxation	Cleavage; IBA
	(1 1 1)Ga	2×2	Ga vacancies	MBE
	(1 1 1)As	2×2	As trimers	MBE
InSb	(1 1 0)	1×1	Rotation–relaxation	Cleavage
	(1 1 1)In	2×2		IBA; MBE
	(1 1 1)Sb	3×3		IBA; MBE
ZnO	(0 0 0 1)Zn	1×1		Cleavage; IBA. Faceting present
	(0 0 0 1)O	1×1		Cleavage. Faceting present

ann. = annealing; IBA = ion bombardment + annealing; MBE = molecular beam epitaxy.

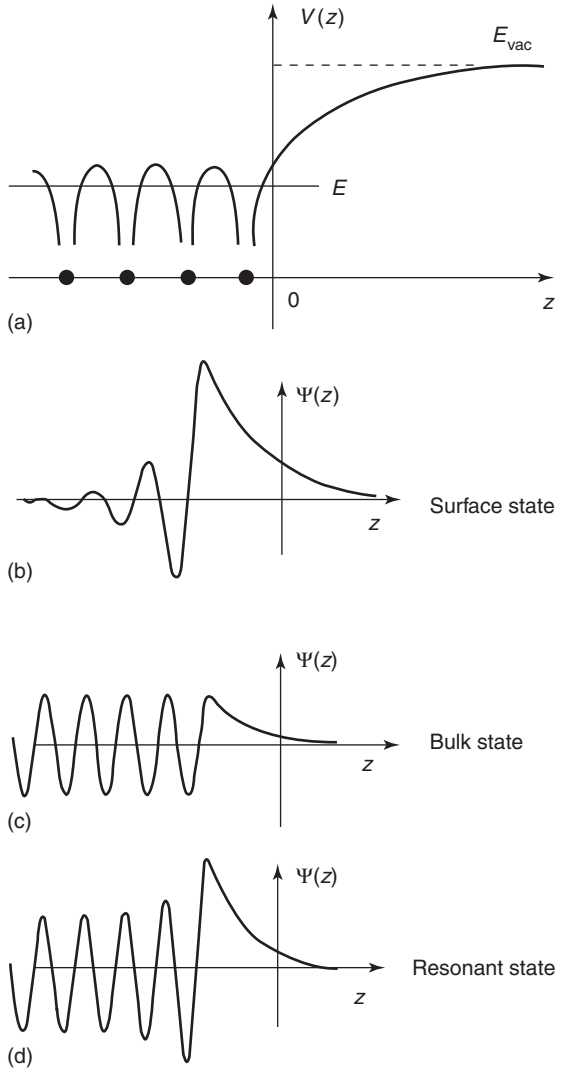


Figure 2 Schematic representation of a surface in a 1D model: (a) potential energy of an electron; (b) wave function for a surface state with energy in the gap; (c) bulk electron scattered at the surface; and (d) surface resonant state.

where $u_k(z)$ is a function with the same periodicity of the lattice, $k = 2\pi/\lambda$ is the wave number, and $E = \hbar\omega$ is the energy of the electron.

For an unbound lattice $(-\infty < z < \infty)$, k must be a real number so that eqn [1] represents a wave propagating along z. Imaginary values of k are discarded because the wave function cannot be normalized. Since the energy is a function of k, real (imaginary) values of k correspond to allowed (forbidden) bands.

In the presence of a surface, however, imaginary values of k cannot be discarded since now, the wave function can be normalized and states localized at the surface, with energies in the forbidden gaps, become possible. If it is assumed that $k = i\mu$ (with μ real), the spatial part of the wave function can be written (for $z < 0$) as

$$\psi_\mu(z) = u_\mu(z)e^{-\mu z} \qquad [2]$$

When $\mu < 0$, the wave function can be matched, at the surface, to that appropriate for $z > 0$ (a decaying exponential if the potential is given by a step function). This is shown in **Figure 2b** that represents schematically a state localized at the surface with energy in the forbidden gap. Such a state is called a "surface state." The matching, however, can also be done when k is real, giving rise to the state shown in **Figure 2c** that represents a bulk electron scattered at the surface.

In three dimensions, the translational symmetry is preserved on the surface plane so that the surface state is now represented by a Bloch wave propagating along the surface with wave vector $\boldsymbol{k}_{||}$, damped on both sides of the surface as in the 1D example. In this case, however, the energy of the surface state does not necessarily fall into the forbidden gaps because of the energy (kinetic and potential) associated to the motion along the surface.

When the degeneracy of the surface state with the bulk bands is not limited to energy but extends to $\boldsymbol{k}$-values, the state is called a "resonant state." A 1D example is shown schematically in **Figure 2d**. In a naïve picture, such a resonant state corresponds to a bulk electron that is scattered at the surface, though dwelling there for some time.

In order to present the correlation between surface and bulk states, it is customary to project the bulk bands on the surface Brillouin zone and plot the surface bands there. This is done for un-reconstructed Si(1 1 1) in **Figure 3**. Surface states are represented by dotted lines, surface resonances by full lines, while projected bulk bands are shown as hatched regions.

When the surface is reconstructed, the calculation of the bands is more difficult and can be done only on the basis of a model.

A common procedure is to pile up a number of crystallographic planes where atoms are displaced from their ideal positions according to a given model of reconstruction. The slab should be large enough (10–20 layers) so that in the interior, the structure of the bulk is resumed. Each slab exposes two surfaces. The slabs are then arranged periodically with spacing (between each two) large enough to avoid a sizable perturbation. This fictitious periodic structure allows the use of the theorems of the 3D band theory.

The results are usually presented in terms of the so-called "local density of states." As it is known, the "total density of states" $\rho^{\text{tot}}(E)$ gives the number of

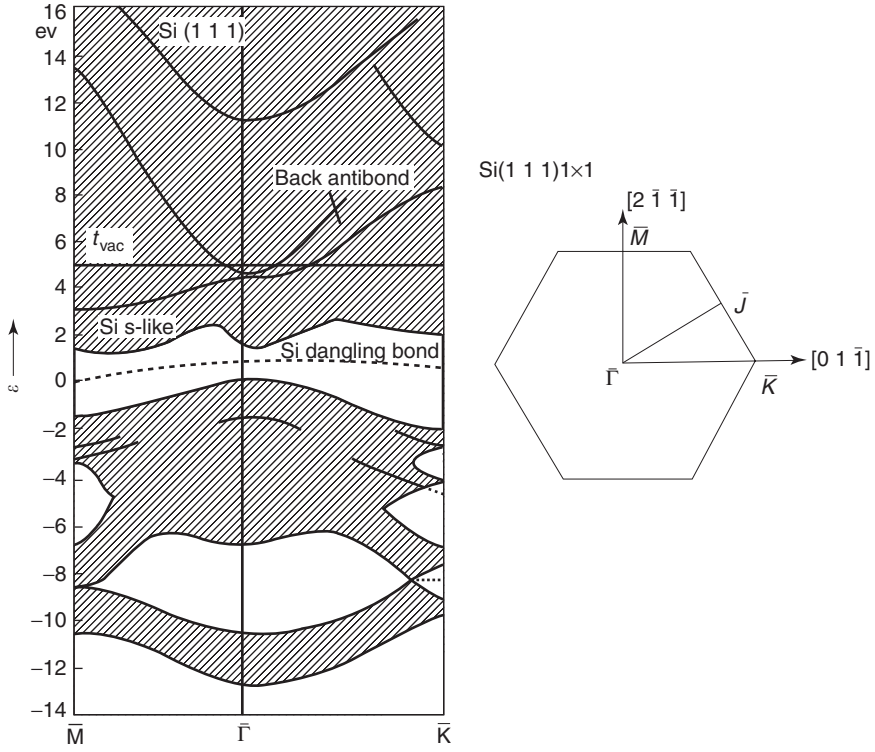


Figure 3 Bulk bands projected into the surface Brillouin zone for the ideal (un-reconstructed) surface of Si(1 1 1). Bulk bands are hatched; surface states are shown with dotted lines; surface resonances with full lines. The surface Brillouin zone is shown on the right. (After Schlüter M and Cohen ML (1978) *Physical Reviews B* 17: 716.)

states present in a unit energy range around E:

$$\rho^{\text{tot}}(E) = \sum_i \delta(E - E_i) \qquad [3]$$

where δ is the Dirac function and the sum runs over the energy states E_i.

In turn, the local density of states is defined as:

$$\rho^{\text{local}}(x, y, z; E) = \sum_i |\Psi_i(x, y, z)|^2 \delta(E - E_i) \qquad [4]$$

that is, each state is weighed with the probability of finding an electron at x, y, z.

Figure 4 shows the local density of states for the ideal surface Si(1 1 1)1 × 1. The figure clearly shows the presence of unsaturated bonds (dangling bonds, DB) on atoms of the first layer.

An important development for calculating the structural and electronic properties of a surface came from the application of the so-called Car–Parrinello method that consists of simulating the dynamical evolution of a reduced set of atoms ($\sim$100), starting from an initial configuration that can as well be the ideal surface. For each configuration, the electronic energy is calculated and,

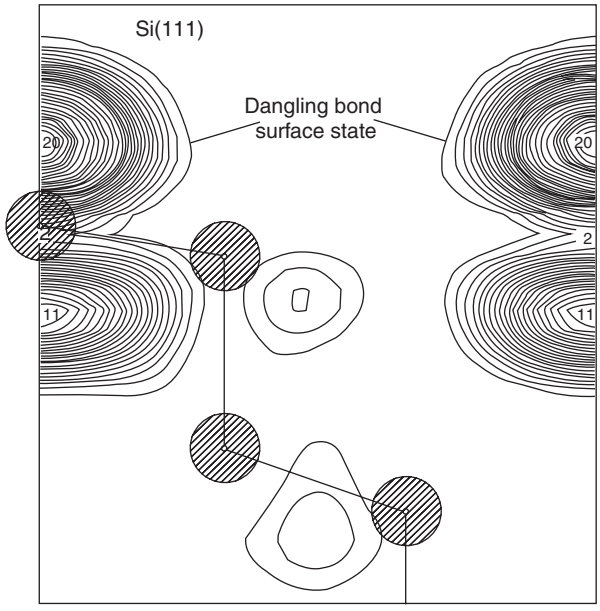


Figure 4 Charge density contours in the (1 1 0) plane (perpendicular to the (1 1 1) surface) for the un-reconstructed (ideal) surface of Si(1 1 1). The dangling bonds protruding into vacuum are clearly visible on the atoms of the surface layer. (After Schlüter M, Chelikowsky JR, Louie GS, and Cohen ML (1975) *Physical Review B* 12: 4200.)

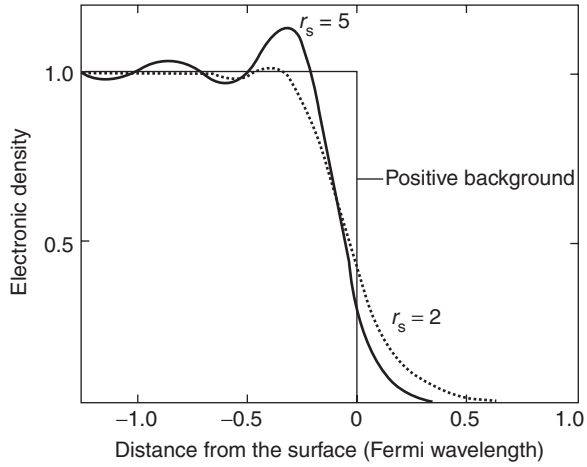


Figure 5 Surface charge density for the jellium model of a metal, plotted for two values of the electron densities. The radius r_s (in atomic units) is that of the sphere that contains one electron. $r_s = 2$, 5 correspond approximately to the cases of Al and Cs, respectively. (After Lang ND and Khon W (1970) *Physical Review B* 1: 4555.)

through the Hellman–Feynman theorem of quantum mechanics, the force on the single atoms is determined and so is the dynamical evolution of the system. Eventually, a stable configuration is reached that corresponds to the reconstructed surface. This "*ab initio*" calculation requires, however, a high computational power and some "educated guesses," especially if the reconstruction is not that of thermodynamical equilibrium.

In metals, one-electron theories are generally inadequate and calculations must be done on the basis of the many-body theory. A model that has given reliable results for the work function of several metals is the so-called "jellium," which assumes that the electrons form a gas of negatively charged particles neutralized by a continuous distribution of fixed positive charge that goes abruptly to zero at the surface. **Figure 5** shows the electron density distribution near the surface for two cases representative of Al (dashed line) and of K (continuous line). Electrons protrude from the surface giving rise to an electric dipole directed inwards, which reduces the work function. Characteristic oscillations of the electron density (Friedel oscillations) should be noticed.

A further peculiarity of metals is that the major contribution to the surface barrier comes from the image force of electrostatics. The potential energy is then, at a large distance from the surface, Coulomb-like. If the electron energy corresponds to a gap, the electron cannot penetrate the solid and remains confined in a sort of a well that extends from $z = 0$ to the Coulomb barrier. In such a well, quantized levels called "image states" exist. Each state is characterized by an integer quantum number, $n \geqslant 1$. In this scheme, the Shockley (or Tamm) states are those with $n = 0$.

The states discussed so far are called "intrinsic surface states" and can be observed only on clean surfaces. On real surfaces, electronic states associated with mesoscopic defects such as steps, kinks, islands, terraces, and pits, or with surface impurities are also present. Such states are called "extrinsic surface states" and may modify the properties of technical surfaces.

In semiconductors, the presence of surface states (both intrinsic and extrinsic), especially when localized in the gap, modifies many of the properties of the sample significantly. The accumulation of charge into surface states creates a macroscopic potential that alters the position of the Fermi level at the surface and introduces a "band bending" that extends for a length of the order of the Debye length defined as

$$L_D = \sqrt{\frac{\varepsilon_b \varepsilon_0 kT}{e^2(n_b + p_b)}} \qquad [5]$$

where $\varepsilon_b\varepsilon_0$ is the dielectric constant and n_b, p_b the densities of negative and positive carriers in the bulk. L_D spans from a fraction of an Å at metal densities to a fraction of a mm in large gap semiconductors.

Such a band bending bears a great importance on many functions of the metal oxide semiconductor (MOS), Schottky barrier, and heterostructure devices.

Examples of Reconstruction of Relevant Semiconductor Surfaces

A few examples of reconstruction of relevant surfaces of semiconductors and some of their associated electronic properties are given below. An outline of the basic experimental evidence is also presented for each surface.

Covalent or partially ionic semiconductors such as C (diamond), Si, Ge, GaAs, and InSb, crystallize in the diamond, zinc blende (or wurtzite) structures in which each atom is tetrahedrally coordinated with four neighbors. The chemical bond is essentially of the directional sp^3 type. The presence of the surface introduces a number of unsaturated bonds called dangling bonds (DBs) that play a great role in reconstruction. Since the energy associated with sp^3 bonds is rather high, the atoms on the surface tend to rearrange in order to reduce the number of DBs. This can be done in a number of ways: dimerization, formation of adatoms (i.e., atoms that have left their ordinary position to sit in the interstice on top of three atoms), π-bonding, etc.

It should be observed in this respect that the sp^3 bond is easily deformable, in the sense that little

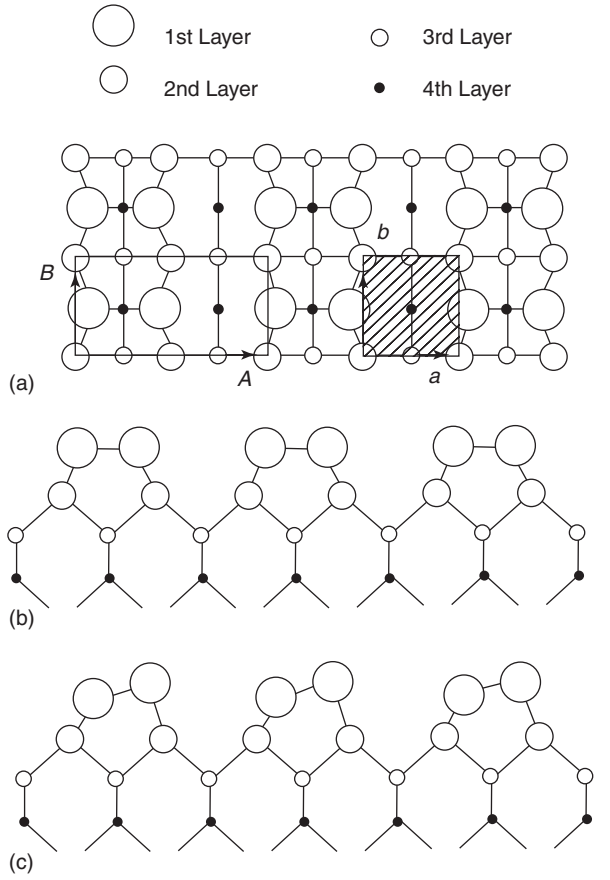


Figure 6 Stick and ball model of the surface of Si(1 0 0) reconstructed 2×1 through dimerization: (a) top view. The 1×1 unit cell (hatched, basis vectors *a*, *b*) and the 2×1 cell (basis vectors *A*, *B*) are shown; (b) side view; (c) side view with buckled dimers.

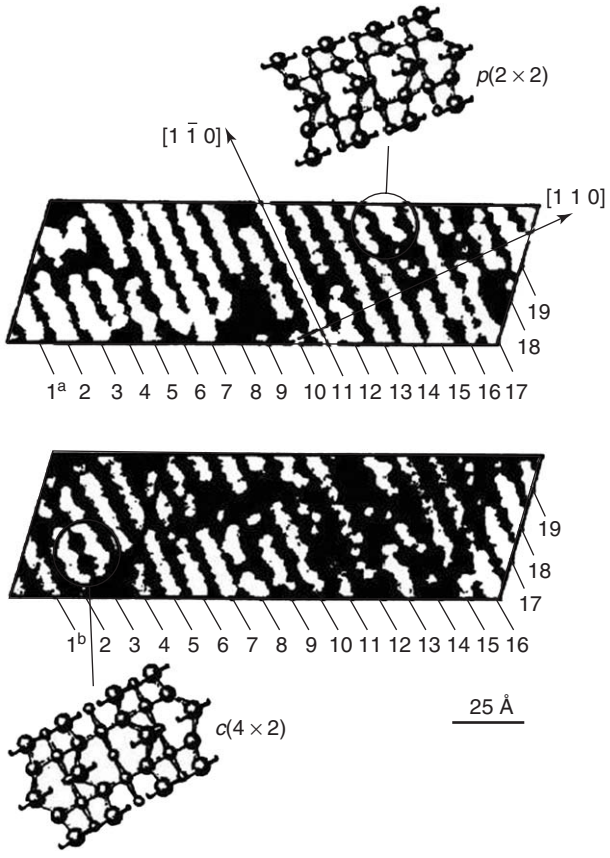


Figure 7 STM pictures of the Si(1 0 0)2×1 surface. The photographs show locally $p(2 \times 2)$ reconstruction (rows 14, 15 in the upper picture) and $c(4 \times 2)$ reconstruction (rows 2, 3 in the lower picture). (From Tromp RM, Hamers RJ, and Demuth JE (1985) *Physical Review Letters* 55: 1303.)

energy is required to change the angles among the bonds. Reconstruction takes advantage of this flexibility, allowing a number of different structures to be present even on the same surface.

Si(1 0 0)2 × 1

The ideal (1 0 0) face exposes two DBs per surface atom and is then highly unstable. Dimerization is a process that produces the 2×1 reconstruction usually observed. The 2×1 reconstructed surface is shown schematically in **Figure 6**: (a) in top view and (b), (c) in side view. The dimers are symmetric in (b) and asymmetric (buckled) in (c). A small buckling is believed to stabilize the surface.

The saturation of the DB cannot be complete so that other types of reconstruction are observed, namely $p(2 \times 2)$ and $c(4 \times 2)$ that occur on $\approx 5\%$ of the surface. Two STM pictures that show the three types of reconstruction on the same surface are presented in **Figure 7**: in the upper picture dimers are seen most distinctly in rows 7, 1, 13 while the $p(2 \times 2)$ reconstruction is seen in rows 14–15. On the other hand, the $c(4 \times 2)$ reconstruction is seen in rows 2–3 of the lower picture.

Si(1 1 1)2 × 1

The (1 1 1) face is the cleavage plane of Si and exposes one DB per surface atom. The 2×1 reconstruction is observed after cleavage at $T < 550\,\mathrm{K}$ at which temperature it transforms irreversibly into the more stable 7×7 structure. The electronic structure consists of two bands (one full and one empty) separated by a gap of $\sim 0.5\,\mathrm{eV}$, from which optical transitions are observed. The opening of the gap is probably the process that stabilizes the surface.

Several models have been proposed for this surface. The only one that satisfies the various experimental findings is that proposed in 1981 by K C Pandey, which consists of π-bonded chains along [1 1 0] directions. The model is shown in **Figure 8** in top and side views. It is seen that in order to create the chains some bonds should be broken. This, however, can be

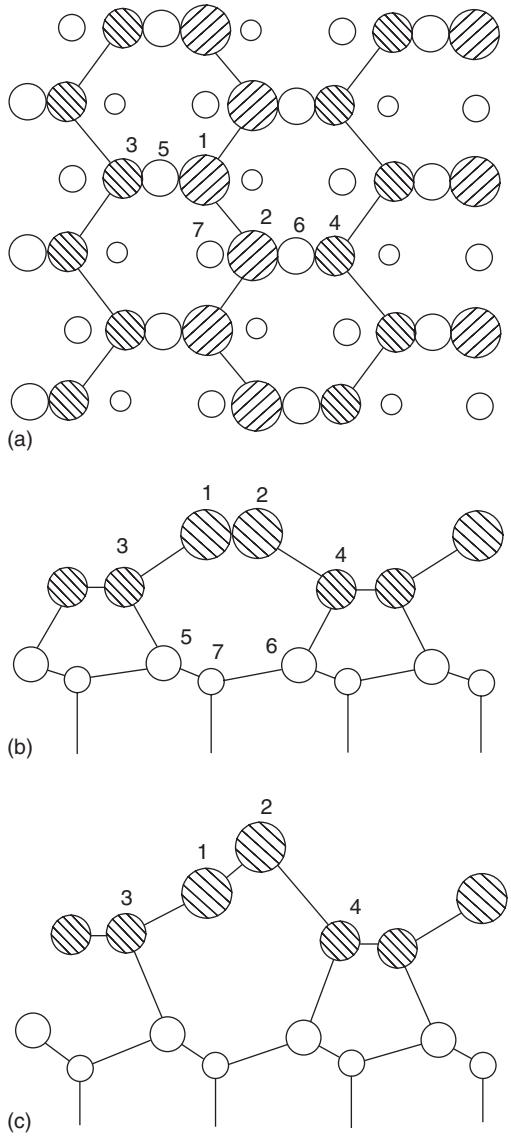


Figure 8 Stick and ball model of the Si(1 1 1)2 × 1 surface, according to the π-bonded chain reconstruction: (a) top view; (b) side view; and (c) side view for a buckled chain.

done almost adiabatically (with negligible activation energy) as can be seen from **Figure 9**, where the local density of states is plotted for various configurations starting from a slightly buckled (almost ideal) surface.

The chain model explains, among others, the following experimental results: (1) the great anisotropy of optical transitions that occur only if the electric vector lies on one of the [1 1 0] directions and (2) the strong dispersion of the occupied states observed in photoemission. The model has been subsequently verified directly by STM. A small buckling along the chain (**Figure 8c**) is necessary to explain the relatively large optical gap.

Si(1 1 1)7 × 7

As already mentioned, this is the equilibrium reconstruction of the (1 1 1) face. The elementary cell contains 49 atoms in the uppermost plane. It has then a complex structure that resisted, for at least two decades, all attempts at theoretical explanation.

By combining accurate electron diffraction at high energies (to avoid multiple scattering), STM pictures, and theoretical considerations, K Takayanagi and Coll proposed in 1985 the reconstruction model shown in **Figure 10**. The elementary cell is a rhomb with two nonequivalent halves and four vacancies at the corners. Dimers can be seen along the sides and the shorter diagonal. 12 adatoms and 6 so-called "rest atoms" (i.e., atoms that have maintained their original positions) are also visible. A stacking fault is present in the left half of the elementary cell. The model is then called DAS (dimers–adatom-stacking fault) and is universally accepted. The number of DB (which were 49 in the unreconstructed cell) has now been reduced to 19 (12 on the adatoms, 6 on the rest atoms, 1 in the vacancy). A similar 5×5 structure (with 6 adatoms, 2 rest atoms, and 4 corner vacancies) is seen in surfaces prepared by MBE and is often associated with the more common 7×7.

From the electronic point of view, the 7×7 structure has a metallic character as shown by photoemission and by the absence of optical transitions in the gap.

Interfaces

The interface is the region that connects two different physical systems, for example, a metal and a semiconductor (Schottky barrier), two different semiconductors (heterojunction) and, in an extended sense, the surface itself, which defines the boundary between the solid and the vacuum.

When the interface forms and thermal equilibrium is reached, the Fermi levels of the two sides line up. This alignment produces a charge redistribution to preserve the overall neutrality of the system, changing the electrical properties of the contact zone.

If two metals come into contact, the difference between the respective work functions results in a flow of electrons from the metal with lower work function Φ_1 into the metal with higher work function Φ_2. A contact potential difference $\Delta = (\Phi_2 - \Phi_1)/e$ develops between the metals (Volta effect), and a dipole layer appears at the interface. Due to the high charge density in metals, the thickness of the interface region is only a fraction of an angstrom.

If the junction connects a metal and a semiconductor, the Fermi level alignment in both materials results in the band scheme of **Figure 11** where the

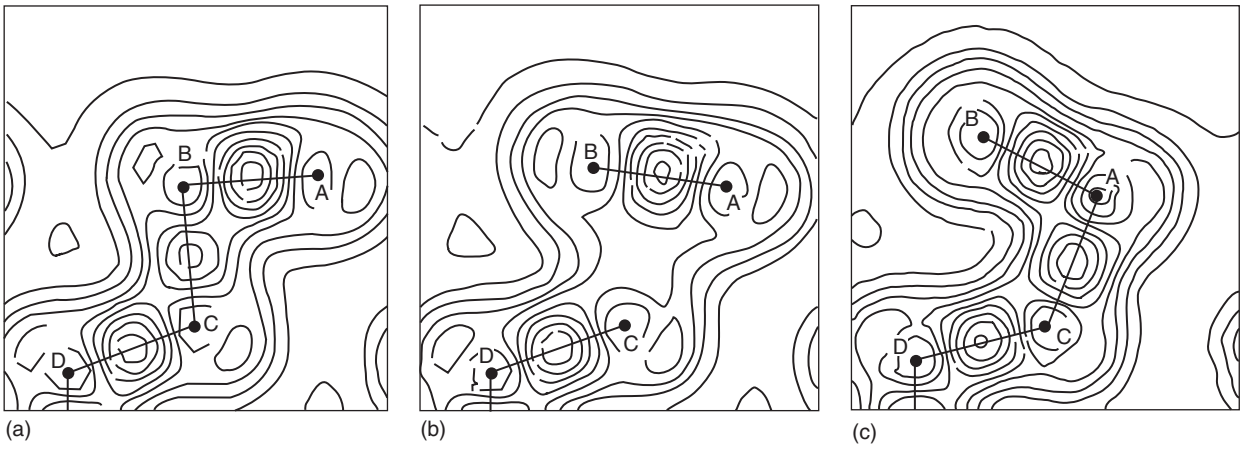


Figure 9 Theoretical contour plots of charge density in the (1 1 0) plane of Si(1 1 1)2 × 1: (a) for a slightly buckled (nearly ideal) surface; (b) for an intermediate geometry; and (c) for the chain model of **Figure 8**. (After Northrup JE and Cohen ML (1982) *Physical Review Letters* 49: 1349.)

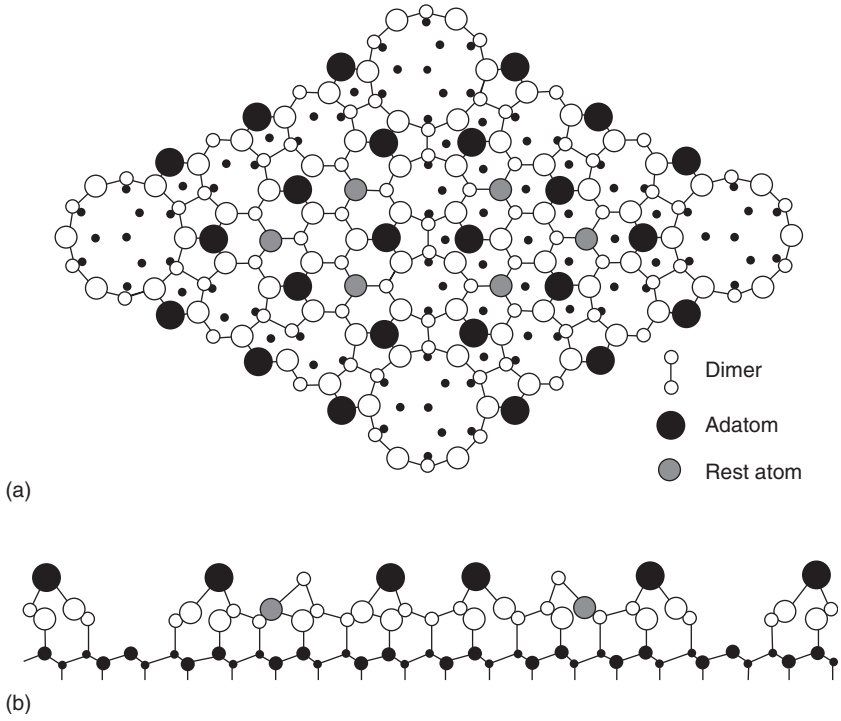


Figure 10 Structure of the Si(1 1 1)7 × 7 surface according to the DAS model: (a) top view; (b) side view. Notice the different positions of the lower neighbors of the rest atoms in the left and right half of the elementary cell, caused by the presence of the stacking fault.

band structures are shown in equilibrium: (a) before and (b) after the contact. An electron moving from the metal to the semiconductor (or vice versa) experiences a barrier (Schottky barrier, Φ_{SB}) given by

$$\Phi_{SB} = \Phi - \chi \qquad [6]$$

where χ is the electron affinity of the semiconductor. The space charge layer in the semiconductor has a much larger width than in the metal because of the difference of carrier densities and screening lengths (for the definition of Debye screening length see the section "Examples of reconstruction of relevant semiconductor surfaces", eqn [5]).

In **Figure 12**, the experimental values of the Schottky barrier for different metals deposited onto the Si(1 1 1)2 × 1 surface are shown. Though the general trend of eqn [6] is maintained, it is apparent

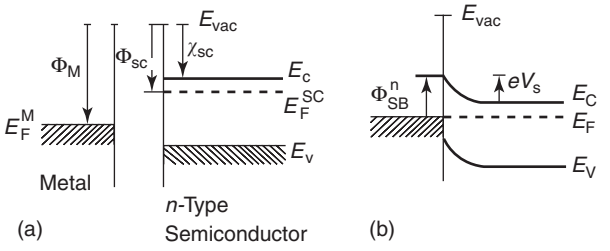


Figure 11 Band scheme of a metal–semiconductor (*n*-type) system: (a) before the contact is established; (b) in equilibrium after the contact. Φ_M (Φ_{SC}) are the work functions of the metal (semiconductor), χ_{SC} the electron affinity of the semiconductor, Φ_{SB} the Schottky barrier height, and eV_S the band bending at the surface.

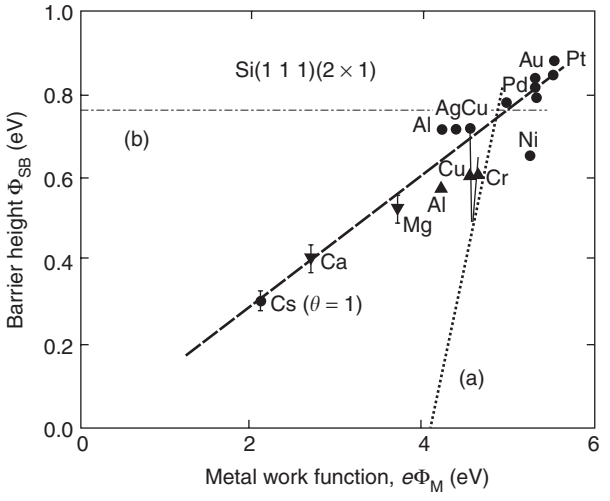


Figure 12 Experimental values of Schottky barriers for several metals on top of Si(1 1 1)2 × 1. The lines represent (a) the Schottky behavior given by eqn [6] (dotted) and (b) the case of complete pinning of Fermi level at the surface (Bardeen behavior, dash-dotted). (After W. Mönch (1986) Festkörperprobleme Vol 26.)

that the simple theory outlined above (dotted line) is inadequate to fit the experiments. The qualitative explanation is that in deriving eqn [6], the presence of electronic surface states at the interface (interface states) has been neglected. If their density is sufficiently large, they pin the Fermi level at the surface of the semiconductor, making the Schottky barrier independent of the work function of the metal (dash-dotted line in **Figure 12**). The experimental situation is intermediate between the two hypotheses (i.e., absence of interface states and complete pinning of the Fermi level).

The nature of the interface states has been the subject of extended research. An important advancement came from the introduction of the so-called metal-induced gap states (MIGS). Consider again **Figure 2c** of the section "Electronic surface states," where the scattering of a bulk electron at the surface is outlined. If the electron at the metal side of the junction has an energy corresponding to the semiconductor gap, it cannot propagate for $z > 0$. A situation similar to that of **Figure 2c** is then a reasonable approximation for the interface wave function: the electronic charge spills into the semiconductor giving rise to a density of states localized at the interface in the semiconductor side (MIGS).

On the other hand, extrinsic states due to defects, mismatching, impurities, and structural changes induced by interface chemical reactions have the same general effect as MIGS. A comprehensive theory of the electronic structure of interfaces is at present not available.

If the interface is created between two different semiconductors, it is called a heterojunction. The alignment of the Fermi levels on both sides determines the band scheme shown in **Figure 13**. The difference between the forbidden gaps of the two semiconductors is distributed between the valence-band discontinuity (ΔE_v) and the conduction-band discontinuity (ΔE_c). Moreover, band bending (eV_1 and eV_2) occurs at the two sides of the interface. Given their great influence on carrier transport, they are the important parameters determining the performance of the interface, and play a central role in the behavior of real heterojunction devices (diodes, solar cells, photodetectors, lasers, quantum devices, etc.).

The presence of electronic states at the interface modifies the distribution of charge producing the situation schematically presented in **Figure 13d**. Such states originate, in analogy to MIGS, from the tailing of the semiconductor wave function in the forbidden gaps.

In the ideal heterojunction, the two materials have the same (or slightly different) lattice parameters so that there is little strain at the interface, and the materials maintain their original properties up to the abrupt junction. If the lattice parameters are different, strain results at the junction, with the formation of dislocations, defects, and disorder. Sizable interdiffusion and chemical reactions may also occur. Such defects give rise to extrinsic electron states.

Three significant examples are given below:

1. The interface resulting after growth of GaAs (lattice parameter $a = 5.6531$ Å) onto AlAs ($a = 5.662\,2$ Å) is a representative example of a nearly ideal heterojunction between materials with a very small lattice mismatch ($\Delta a/a = 0.5\%$). The resulting interface is atomically abrupt, with a very low number of defect states.

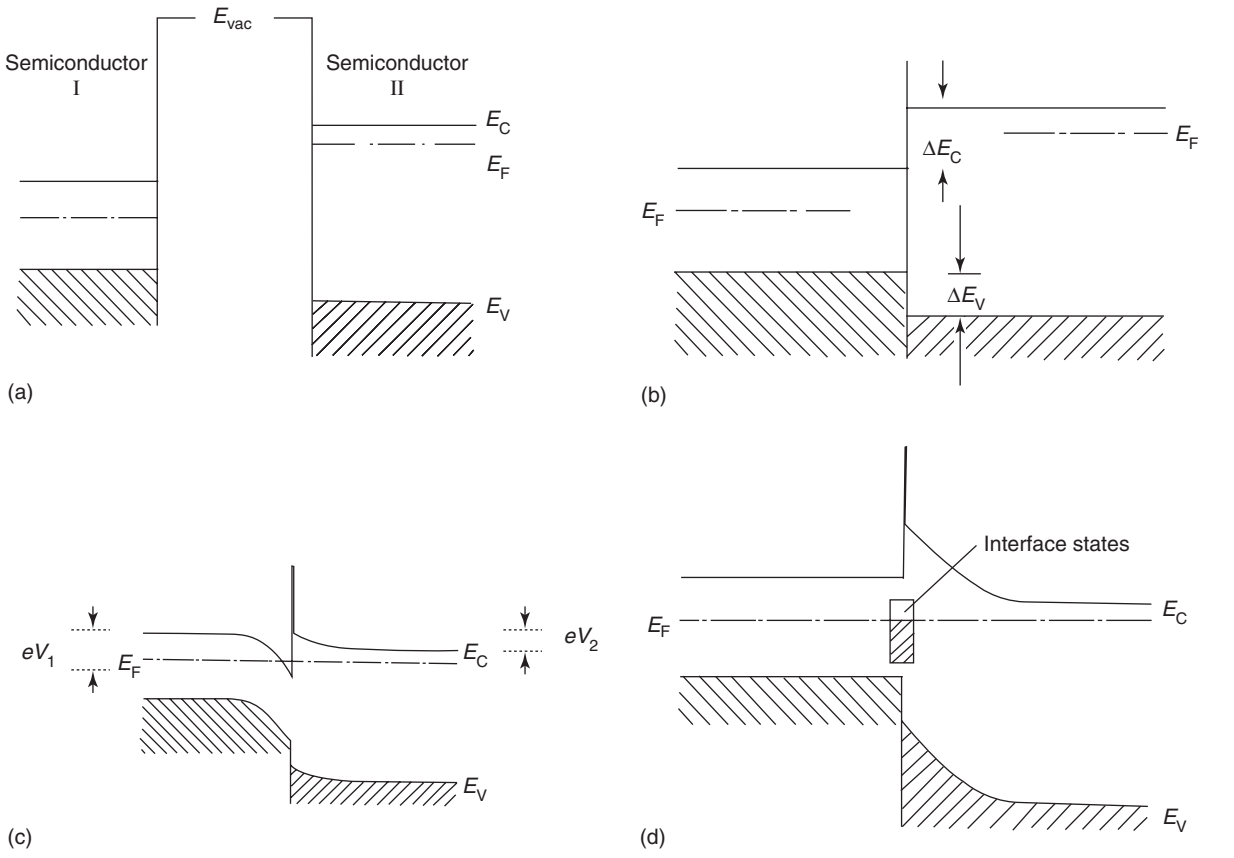


Figure 13 Band scheme of a heterojunction: (a) when the materials are not in contact; (b) after the contact is established, but before equilibrium (Fermi levels not aligned); (c) in equilibrium; (d) in equilibrium, when interface states are present.

2. The interface between Ge ($a = 5.6578$ Å) and Si ($a = 5.4307$ Å) presents a lattice mismatch of 4%, too large to be simply accommodated by a strained layer. The interface is then characterized by intermixing, with the formation of a Ge–Si alloy.
3. In the junction between Ge ($a = 5.6578$ Å) and GaAs ($a = 5.6531$ Å), the mismatch is extremely favorable (about 8×10^{-4}). Nevertheless, the chemistry of materials, that is, the high chemical affinity of Ge, Ga, and As produces a great complexity of the interface composition, with possible effects on doping. Ge acts as a donor or an acceptor in bulk GaAs, while Ga and As are, respectively, an acceptor or a donor when the Ge atom is substituted.

Heterojunctions and heterostructures (i.e., the sequence of a high number of heterojunctions, as in a superlattice) are usually grown by MBE or metal-organic chemical vapor deposition (MOCVD). These epitaxial techniques allow the production of graded junctions that are important in many semiconductor devices.

See also: Epitaxy; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Molecular Dynamics Calculations; Semiconductor Devices; Surfaces, Optical Properties of; Valence Photoemission.

PACS: 61.14.Hg; 61.16.Ch; 68.35.Bs; 73.20.−r; 73.20.At

Further Reading

Bechstedt F (2003) *Principles of Surface Physics*. Berlin: Springer.
Bechstedt F and Enderlein R (1988) *Semiconductor Surfaces and Interfaces*. Berlin: Akademic-Verlag.
Binnig G, Rohrer H, Gerber Ch, and Weikel E (1983) *Physical Review Letters* 50: 129.
Car R and Parrinello M (1985) *Physical Review Letters* 55: 2471.
Chiaradia P, Cricenti A, Selci S, and Chiarotti G (1984) *Physical Review Letters* 52: 1145.
Chiarotti G (ed.) (1993–1996) Physics of solid surfaces. *Landolt-Börnstein New Series, Group III/24*, 4 vol. Berlin: Springer.
Chiarotti G, Nannarone S, Pastore R, and Chiaradia P (1971) *Physical Review B* 4: 3398.
Echenique PM, Flores F, and Sols F (1985) *Physical Review Letters* 55: 2348.
Feenstra RM, Thompson WA, and Fein AP (1986) *Physical Review Letters* 56: 608.
Heine V (1965) *Physical Review* 138: A1689.

Himpsel FJ, Heimann P, and Eastman DE (1981) *Physical Review B* 24: 2003.
Lannoo M and Friedel P (1991) *Atomic and Electronic Structure of Surfaces*. Berlin: Springer.
Lüth H (2001) *Solid Surfaces, Interfaces and Thin Films*, 4th edn. Berlin: Springer.
Mönch W (1995) *Semiconductor Surfaces and Interfaces*, 2nd edn. Berlin: Springer.
Pandey KC (1981) *Physical Review Letters* 47: 1913.
Takayanaji K, Tanishiro, Takahashi M, and Takashi S (1985) *Journal of Vacuum Science and Technology A* 3: 1502.
Zangwill A (1988) *Physics at Surfaces*. Cambridge: Cambridge University Press.

Nomenclature

$\boldsymbol{k}$	wave vector (cm^{-1})
L_D	Debye length (cm)
ΔE_c	conduction band discontinuity (eV)
ΔE_v	valence band discontinuity (eV)
ρ_{local}	local density of states (states $atom^{-1}\,eV^{-1}$)
ρ_{tot}	total density of states (states $atom^{-1}\,eV^{-1}$)
Φ_{SB}	Schottky barrier (eV)
Φ	work function (eV)
χ	electron affinity (eV)
Ψ	wave function

Surfaces, Optical Properties of

R Del Sole and P Chiaradia, Università di Roma, Rome, Italy

Introduction

In this article, an overview of the application of linear optical spectroscopies in surface science is given. From a historical point of view, optical methods were important in the very first development of surface-state spectroscopy of clean semiconductor surfaces at the end of the 1960s, when progress in both vacuum technology and surface diagnostics made it possible to reproducibly obtain well-characterized clean surfaces. In fact, optical transitions between filled and empty surface states of intrinsic type were discovered at that time in Ge(1 1 1)2×1 and Si(1 1 1)2×1 by differential reflectance with multiple internal reflections. These results demonstrated the existence of intrinsic (electronic) surface states for the first time in a direct way rather than indirectly from electrical measurements. In the subsequent years, many other surface-sensitive spectroscopies were successfully applied to investigate surface states at the cleavage faces of semiconductors as well as at other clean surfaces.

Generally speaking, in order to study surface optical properties, one has to extract the optical signal originating from the surface from the overwhelming bulk term. For example, in the surface differential reflectivity (SDR), this is accomplished by measuring a change of reflected-light intensity induced by either contamination (usually oxidation or hydrogenation) or chemisorption of foreign atoms. Instead, in reflectance anisotropy spectroscopy (RAS), which is also named reflectance difference spectroscopy (RDS), one measures the reflectance difference for light linearly polarized along two (orthogonal) axes of symmetry at the surface of cubic semiconductors. Since, in general, the surface has a lower symmetry than the bulk, an anisotropy signal originating from the surface is often measured.

When performed with polarized light, SDR yields information on all surface optical transitions (both isotropic and anisotropic) but requires a reference surface, which is obtained by modifying the sample in an irreversible way. On the other hand, RAS does not require any reference surface but is obviously limited to detection of anisotropic transitions only.

The interpretation of the optical experiments is, in general, not straightforward. Hence, the full potentiality of surface optical spectroscopy can be exploited only by a strong interplay of experimental and theoretical work.

At the beginning of the twentieth century, the so-called three-layer model (ambient–surface layer–substrate) was introduced by Drude to describe the surface optical properties from a phenomenological point of view. However, the first steps of the modern theory of surface optical spectroscopy started in 1971, when McIntyre and Aspnes expanded the reflectivity formula for the three-layer model to first order in the surface-layer thickness, obtaining a simple result, suitable for a clear physical interpretation. Realistic and *ab initio* calculations of surface optical properties have later been carried out, even including (in recent years) the electron–hole interaction. At present, the theory is fully developed, but its application to complex surfaces is hampered by the limitations of computational capabilities.

In the following, some relevant results obtained with surface optical spectroscopies, mainly SDR, RAS, and ellipsometry, are briefly described, together with their theoretical interpretations.

Results Obtained with SDR

One of the best-known semiconductor surfaces is the cleavage face of silicon, namely Si(1 1 1)2 × 1, whose reconstruction is described by the chain model developed in 1981 by Pandey. The optical properties of this surface have been extensively investigated over the years, in particular with reflectance techniques. **Figure 1** shows the so-called absorption constant in the near infrared due to transitions between filled and empty surface states possessing a dangling-bond character, at Si(1 1 1)2 × 1. These pioneering results (1971) were obtained by applying the method of multiple internal reflections to cleaved samples in order to amplify the small surface effect. They proved the semiconducting character of the surface, in spite of the fact that the ideal (1 1 1) face should be metallic, with an energy gap of ∼0.45 eV. Similar results were obtained in 1968 for Ge(1 1 1)2 × 1.

The results of **Figure 1** were compatible with Haneman's model, in which the lack of equivalency of surface atoms was explained by a charge transfer associated with a partial dehybridization, due to their up and down displacement with respect to the surface plane. These results were indeed interpreted in the framework of that model. However, when the chain model appeared many years later, its pronounced one-dimensional character suggested a strong anisotropy of the optical properties, at variance with what was predicted by Haneman's model. Indeed, this prediction was confirmed by another experiment with linearly polarized light. **Figure 2** shows the surface reflectance peak (analogous to the absorption peak of **Figure 1**) for light polarized parallel and perpendicular to the chains of surface atoms at single domain Si(1 1 1)2 × 1. The results, recently confirmed by RAS, point out that this peak is totally anisotropic, thus providing a convincing evidence for the chain model.

In 1999, Rohlfing and Louie carried out a calculation of the optical properties of this surface, according to the method of the Bethe–Salpeter equation to account for the electron–hole interaction. The result (see **Figure 3**) has suggested that the experimental peak in **Figure 2** is due to a bound exciton, with

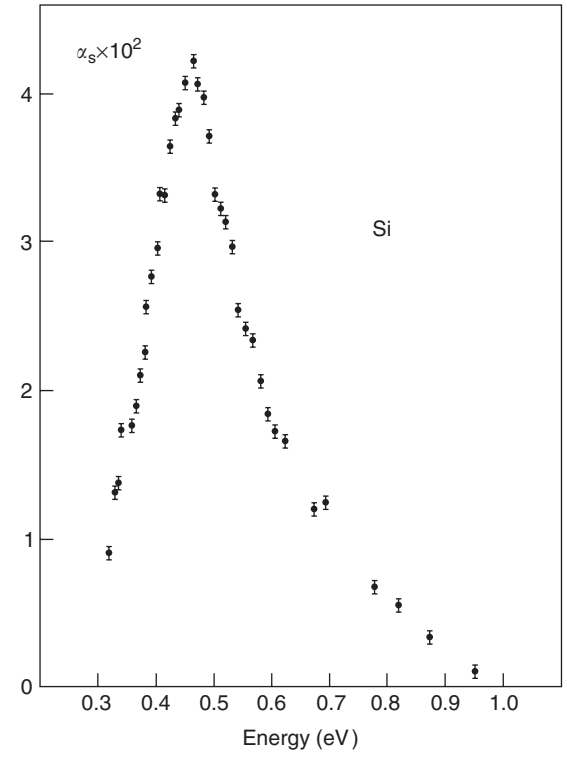


Figure 1 Absorption constant vs. photon energy for Si(1 1 1)2 × 1. (From Chiarotti G, Nannarone S, Pastore R, and Chiaradia P (1971) *Physical Review B* 4; 3398.)

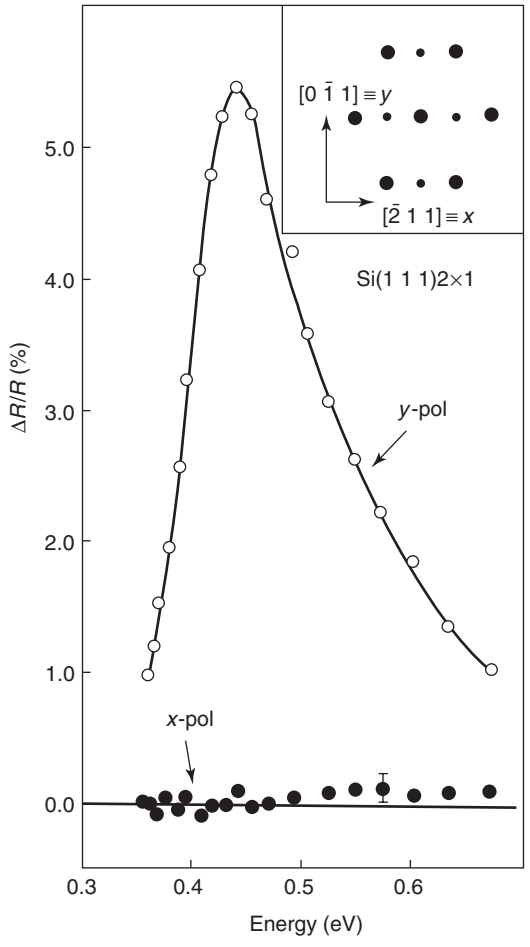


Figure 2 Differential reflectivity spectra of a single domain Si(1 1 1)2 × 1 surface, for light polarized along the $[\bar{2}\,1\,1] = x$ and $[0\,\bar{1}\,1] = y$ directions. The LEED pattern is sketched in the inset. (From Chiaradia P, Cricenti A, Selci S, and Chiarotti G (1984) *Physical Review Letters* 52: 1145.)

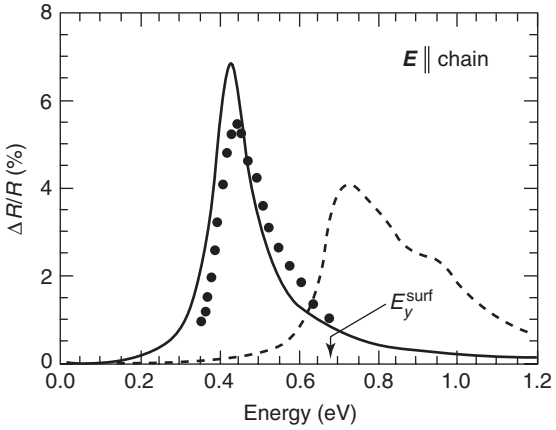


Figure 3 Surface contribution to reflectance of Si(1 1 1)2 × 1 for light polarized along the chains of surface atoms. – – – – : one-electron *GW* calculation; ——— : including excitonic and local field effects; • • • • : experiment. (Reproduced with permission from Rohling M and Louie S G (1999) *Physical Review Letters* 83: 856.)

the electron and hole in two surface-state bands with a binding energy of 0.27 eV.

Results Obtained with RAS

RAS was developed as a modulation technique distinct from SDR (in USA) by Aspnes and in Russia by Safarov, independently, in the mid-1980s.

In the beginning, Aspnes mainly studied optical anisotropies of real surfaces of elemental semiconductors, in particular, Si substrates for technological applications. These anisotropies, ∼1%, are not related to intrinsic surface states. They are, instead, associated to a lower symmetry of the subsurface region with respect to the (cubic) bulk, which perturbs the bulk-state wave functions, a situation often described as "bulk states modified by the surface." **Figure 4** shows an example of these surprising anisotropies in the case of oxidized Si(1 1 0).

It is seen in **Figure 4** that the line shape of the surface anisotropy clearly mimics the imaginary part of the bulk dielectric function of Si. While oxidized Si(1 1 0) now has become a standard for calibration of RAS apparatuses, the interpretation of spectra, such as those of **Figure 4**, is still an intriguing matter. After some controversies on the origin of the spectrum in **Figure 4**, the last calculations carried out, including the electron–hole interaction by Bechstedt and coworkers, have finally shown that it can be explained in terms of surface effects on the bulk resonant exciton underlying the E1 peak, close to 3.4 eV.

Since optical techniques can be performed in metal organic chemical vapor deposition (MOCVD) environments as well in vacuum, RA spectra during

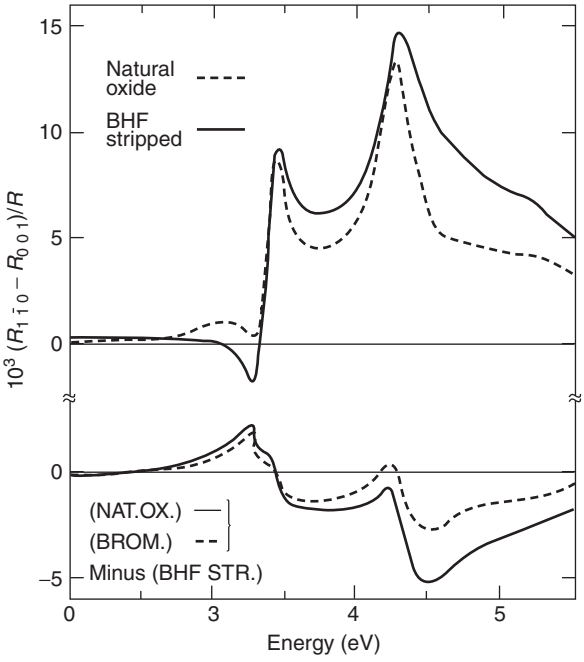


Figure 4 Top: Reflectance anisotropy spectra for a naturally oxidized Si(1 1 0) surface before and after stripping with buffered HF acid. Bottom: changes in the stripped-surface spectrum upon air oxidation and upon exposure to Br_2 in methanol. (Reproduced with permission from Aspnes DE (1985) *Journal Vacuum Science and Technology B* 3: 1498.)

growths of GaAs(0 0 1) by both molecular beam epitaxy and MOCVD have been compared. As a result, a striking correspondence has been found between RAS signals characteristic of surface reconstructions in vacuum and those obtained in an MOCVD reactor at atmospheric pressure of hydrogen in suitable conditions. **Figure 5** shows this correspondence, demonstrating that local ordering is the same in the two cases.

Calculations have been carried out for various reconstructions of GaAs(0 0 1), using both semi-empirical and *ab initio* methods, in order to understand the origin of the observed structures. A good explanation of the RAS of the As-rich reconstructions has been obtained. Most of the RAS structures arise from transitions across surface-perturbed bulk states. However, on the (2 × 4) reconstruction, a combined energy loss and RAS experimental work have recently revealed true surface-state transitions at ∼2.5 eV, well below the first bulk critical point. The optical properties of the Ga-rich (0 0 1) surface, instead, are not yet completely understood. In recent years, interesting RAS investigations have been performed on other semiconductor surfaces, both clean and covered with monolayers of alkali atoms, by the groups of Richter and Paget. Also, organic molecules deposited on different substrates have been studied

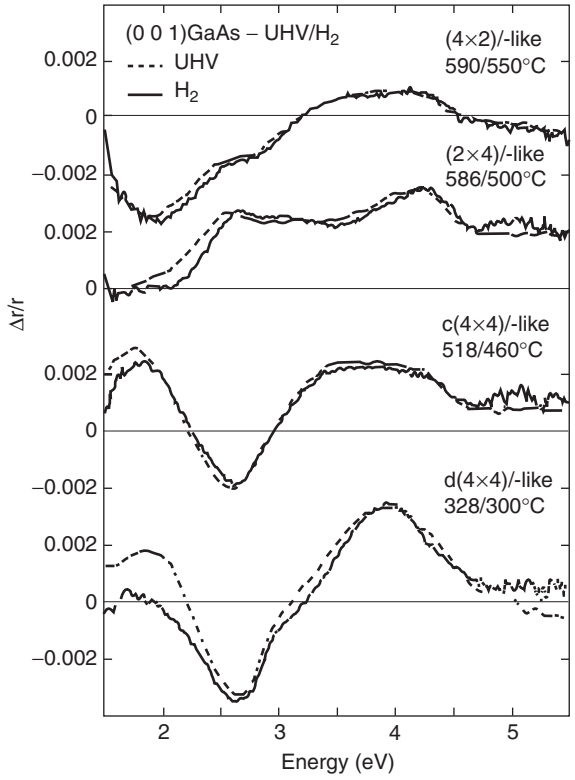


Figure 5 Reflectance anisotropy spectra of the primary reconstructions of GaAs(1 0 0) in UHV and atmospheric pressure of hydrogen, at the indicated temperatures. (Reproduced with permission from Kamiya I, Aspnes DE, Tanaka H, Florez LT, Harbison JP, and Bhat R (1992) *Physical Review Letters* 68: 627.)

by RAS, especially by Weightman and Chiaradia-Goletti. Moreover, RAS shows promise to yield valuable information on low-dimensional systems (quantum wires and dots) and nanostructures (carbon nanotubes, polymers, etc.).

Results Obtained with Ellipsometry

Ellipsometry measures changes in the polarization parameters of light upon reflection from a surface. These parameters are: the relative phase Δ and the relative amplitude ratio arctan Ψ of the two components of linearly polarized light, parallel and perpendicular to the plane of incidence,

$$r^{p}/r^{s} = \tan \Psi \exp(i\Delta) \qquad [1]$$

From the spectral dependence of Δ and Ψ, the optical constants n and k are obtained by making use of the Fresnel formulas for the reflection coefficients r^{p} and r^{s}.

When surface properties are studied by ellipsometry, a differential method has to be used, as in the case of SDR. Thus, the changes $\delta\Delta$ and $\delta\Psi$ of the

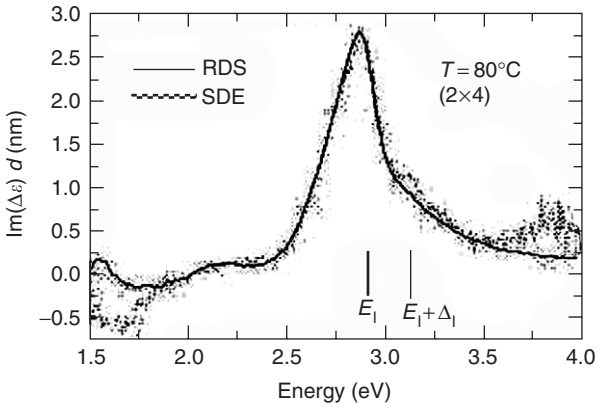


Figure 6 Dielectric anisotropy of GaAs(1 0 0) 2×4, as obtained from reflectance difference spectroscopy (RDS) and from surface differential ellipsometry (SDE). (Reproduced with permission from Wassermeier M, Behrend J, Ploog KH, Zettler JT, Stahrenberg K, and Richter W (1996) *Physical Review B* 53: 13542.)

ellipsometric parameters are recorded upon adsorption of a gas. Under simplifying hypotheses, in the framework of the usual three-layer model, the optical constants of the surface are eventually obtained.

Starting from the pioneering work of Mayer in the 1970s, ellipsometry has been applied to derive the surface optical constants of a number of clean semiconductor surfaces. Especially, various crystallographic faces of Si and Ge, as well as the cleavage faces of GaP, GaAs, and ZnO, and also a number of metallic surfaces have been investigated by this technique.

In spite of the seemingly greater complexity in the data reduction, the results obtained by ellipsometry are practically coincident with those obtained by reflection-based methods, as demonstrated, for instance, by the spectra of **Figure 6**, reporting the dielectric anisotropy obtained from both RAS and ellipsometry data for GaAs(0 0 1) 2×4.

Other Experimental Techniques

Many other optical techniques have contributed in the past to the development of surface physics. Although some of them have fallen into disuse, they deserve to be mentioned here: surface photoconductivity, surface photovoltage, photothermal displacement spectroscopy, and surface photoluminescence. A review of the main results obtained by means of these techniques, especially in the field of clean semiconductor surfaces, can be found in the literature.

Theory

Theoretical work is needed in order to fully exploit the potentiality of surface optical spectroscopy. Three

steps are involved in the theory: (1) the determination of one-electron wave functions; (2) the calculation of the dielectric susceptibility, possibly including many-body interactions; and (3) the solution of light-propagation equations. The last problem was solved in a quite simple way at the end of the 1970s. The surface contribution to reflectivity (i.e., the reflectance difference with and without the surface layer) can be calculated according to the formula:

$$\Delta R/R = 4\left(\frac{\omega}{c}\right)\cos\theta\,\mathrm{Im}\left\{\frac{\Delta\varepsilon_{yy}}{(\varepsilon_b - 1)}\right\} \qquad [2]$$

for *s*-light incident from vacuum in the *xz*-plane and polarized in the *y* direction.

Here, θ is the angle of incidence, and

$$\Delta\varepsilon_{yy} = \int_{-\infty}^{\infty} \mathrm{d}z \int_{-\infty}^{\infty} \mathrm{d}z' [\varepsilon_{yy}(z,z') - \delta(z-z')\varepsilon_0(z)] \qquad [3]$$

$\varepsilon_0(z) = \theta(-z) + \varepsilon_\mathrm{b}\theta(z)$ is the zeroth-order dielectric constant yielding Fresnel formulas, and $\varepsilon_{yy}(z,z')$ is the relevant component of the nonlocal dielectric tensor of the solid–vacuum interface.

For a microscopic evaluation of surface optical properties, one has to calculate eqn [3]. If excitonic and local-field effects are neglected, $\varepsilon_{yy}(z,z')$ can be computed from one-electron wave functions. It is usually evaluated mimicking the semi-infinite crystal by a slab of several (from 10 to 30) atomic layers. In this case, the right-hand side of eqn [3] is related to the slab dielectric function. Within the single-particle scheme, its imaginary part is simply related to the transition probability induced by the radiation between slab states:

$$\varepsilon_{yy}^{(2)}(\omega) = (8\pi^2 e^2/m^2\omega^2 A) \sum_{\boldsymbol{k}} \sum_{\mathrm{v,c}} |p_{\mathrm{v,c}}^{y}(\boldsymbol{k})|^2 \times \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \qquad [4]$$

where $p_{\mathrm{vc}}^{y}(\boldsymbol{k})$ is the matrix element of the *y*-component of the momentum operator between initial (v) and final (c) slab states at the point $\boldsymbol{k}$ in the two-dimensional Brillouin zone (BZ), and A is the sample area. The real part is computed via the Kramers–Kronig relation.

One more ingredient entering eqn [2] is the bulk dielectric function, whose imaginary part is analogous to eqn [4], where, now, eigenstates and eigenvalues of the infinite crystal, together with three-dimensional $\boldsymbol{k}$ vectors, are involved. The study of the optical response of surfaces within the single-particle picture amounts then to calculating (1) the single-particle spectrum and (2) the transition probability between occupied and unoccupied states, both for the perfect crystal and for the crystal with a surface.

If the dielectric susceptibility $\varepsilon_{yy}(z,z')$ is assumed to be local, and a three-step model is hypothesized for its z dependence, the formula of McIntyre and Aspnes mentioned in the section "Introduction" is recovered.

The one-electron states appearing in eqn [4] should be obtained, strictly speaking, within the theory of Green's functions, by solving the quasiparticle Schrödinger-like equation. In this framework, electrons move in a potential which is the sum of the ion potential, the Hartree potential due to the electronic charge density, and of a nonlocal energy-dependent self-energy, which accounts for exchange-correlation (XC) effects due to the interaction with the other electrons. The so-called *GW* approximation to the self-energy, developed by Hedin in 1965, completes the recipe for a rigorous calculation of band energies in solids. Calculations carried out starting in the 1980s have yielded band energies in agreement within 0.1 eV with experimental values for semiconductors.

GW calculations of the electronic structure of solids are computationally very demanding. Hence, as in the past, before the advent of vector computers, even now for complex systems, simpler – yet less accurate – methods are being used. Semi-empirical methods include the tight-binding and pseudopotential methods, where the matrix elements of the Hamiltonian (in the former case) or of the total potential (in the latter case) between basis functions (local orbitals in the former case, plane waves in the latter) are fitted to reproduce bandgaps at some high-symmetry points of the BZ, obtained from experiments. An *ab initio* method often used to determine the band structure of solids is the density-functional theory (DFT), where the XC self-energy is substituted by a local energy-independent XC potential, which is a functional of the electron density. Optical spectra calculated for bulk semiconductors within the one-electron approximation yield only qualitative agreement with experimental spectra, no matter which method – semi-empirical or *ab initio*, *GW*, or DFT – is used. This is because the excitonic and local-field effects are neglected. When these are considered, the discrepancies disappear, as it was shown first by Hanke and Sham in 1980 with a tight-binding calculation. *Ab initio* calculations of the optical properties of bulk solids including the electron–hole interaction and local-field effects yield quantitative agreement with experiments.

The most used methods to calculate surface optical properties are those relying on the single-particle approximation, ranging from the simplest (less accurate) ones (such as semi-empirical tight binding) to the rigorous quasiparticle *GW* method. Anyway, if the

electron–hole interaction is not included, one can expect only a qualitative description of the spectra. The inclusion of the electron–hole interaction in the calculation of excited spectra leads one to consider a two-particle problem (the Bethe–Salpeter equation) rather than a single-particle problem (the Schrödinger-like equation for quasiparticles). Hence, the computational cost increases dramatically and makes the method usable, at present, only for simple surfaces.

The semi-empirical tight-binding method, for its computational speed, has been the first method used for calculating surface optical properties in 1986, when it was successfully applied to Si(1 1 1)2 × 1 and explained the strong optical anisotropy shown in **Figure 1**. In spite of its limitations, this method has yielded, in many cases, results in qualitative agreement with experiments and has helped in clarifying the atomic and electronic structure of a number of surfaces.

Ab initio methods have been developed in the last decade within DFT local-density approximation (DFT-LDA), by expanding the wave functions in plane waves and using nonlocal norm-conserving pseudopotentials. Finally, a rigorous one-electron calculation of surface optical properties should rely on the Green's function approach, with the electron self-energy calculated according to the *GW* approximation. These approaches are much more demanding than semi-empirical tight binding from the computational point of view, since a great number of plane waves are needed to represent slab wave functions.

As an example, the case of GaAs(1 1 0) is discussed. First-principles calculations including self-energy corrections according to the *GW* approximation have been carried out for this surface and are shown in **Figure 7**. They are compared to room temperature (RT) experiments carried out by Esser and co-workers. Calculations show that the experimental peak at 2.7 eV has a substantial (∼50%) contribution arising from transitions across surface states, while all other structures are due to transitions across surface-perturbed bulk states. The reflectance anisotropy calculated in this way has the main structures in very good agreement with experimental ones as far as energy positions are concerned, while the DFT-LDA calculation, also shown in **Figure 7**, is considerably red-shifted, as a consequence of the DFT "gap problem." Differences in line shapes are still present, which are expected to be fixed by introducing excitonic and local-field effects in the calculations.

Finally, the inclusion of the electron–hole interaction can be performed by solving the Bethe–Salpeter equation at surfaces. The first calculation of this kind, carried out by Rohlfing and Louie for Si(1 1 1)2 × 1 in 1999, is shown in **Figure 3**. As it was mentioned

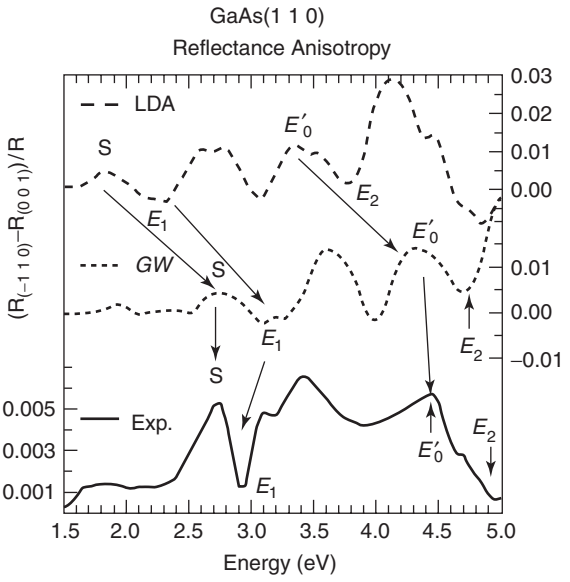


Figure 7 Reflectance anisotropy of GaAs(1 1 0). Upper curve: DFT-LDA calculation. Middle curve: *GW* calculation. Lower curve: RT experiment. (After Del Sole R, Palummo M, and Pulci O (2004) In: Cricenti A (ed.) *Epioptics*-7, p. 1. World Scientific.)

before, the SDR peak at 0.45 eV is ascribed to a surface-state bound exciton, which takes over most of the oscillator strength of the transitions across the dangling-bond (DB)-like surface states. As a consequence, these transitions become very weak and are hardly visible in the experimental spectra. Another application of this technique has been done by Bechstedt and co-workers in 2002 for Si(1 0 0):H, which explained the line shape of the RAS spectrum of **Figure 4** as due to a polarization-dependent quenching of bulk excitonic effects close to 3.4 eV.

Conclusions

Optical spectroscopy has become, in the last decades, an important tool for surface analysis, of great sensitivity to structural changes and versatility. The theoretical understanding of the results is becoming more and more satisfactory, thanks to the application of advanced many-body techniques to the simulation of surface optical properties. Nevertheless, some problems are still open: the most obvious is the extension of these techniques to complex surfaces (those with higher-order reconstructions), which is hampered by the current computational limitations. Less obvious problems are those related to the electron–phonon interaction, which affects surface optical line shapes in a drastic way, and to the complex morphological structure of surfaces, arising from steps, multiple domains, and defects, whose study is now at an initial stage.

See also: Insulators, Optical Properties of; Metals and Metallic Alloys, Optical Properties of; Small Particles and Clusters, Optical Properties of; Surfaces and Interfaces, Electronic Structure of.

PACS: 78.68.+m

Further Reading

Aspnes DE (1993) *Thin Solid Films* 233: 1–8.

Bassani F and Pastori-Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Chiaradia P (1996) In: Chiarotti G (ed.) *Physics of Solid Surfaces*, Landolt Boernstein, Group III "Condensed Matter," vol. 24, p. 29, subvolume d, (Interaction of Radiation with Surfaces and Electron Tunneling.)

Chiaradia P and Del Sole R (1999) *Surface Review and Letters* 6: 517.

Chiarotti G (1994) *Surface Science 299/300*: 541.

Halevi P (ed.) 1995 Photonic probes of surfaces. *Electromagnetic Waves: Recent Developments and Research*. Amsterdam: North-Holland/Elsevier.

Herman MA, Richter W, and Sitter H (2004) *Epitaxy*. Berlin: Springer.

Synchrotron Radiation

G Margaritondo, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Electron accelerators such as synchrotrons and storage rings are the best sources of X-rays and ultraviolet photons for research applications. The physical mechanisms for all sources in this class use relativistic effects to increase the emitted photon energies, and therefore require an electron (or positron) beam moving at relativistic speed. The specific implementation of this strategy is reviewed for different types of synchrotron or synchrotron-related sources: bending magnets, wigglers, undulators, helical wigglers, linear accelerator (LINAC) based sources, and free-electron lasers (FELs).

Introduction: X-Ray Sources

X-rays and ultraviolet photons are very extensively used in research. They are the basic ingredients of a wide variety of experimental techniques ranging from microscopy and spectroscopy to imaging and structural analysis. The reason for their usefulness is quite fundamental: they possess photon energies and wavelengths that match basic magnitudes such as the binding energy of valence or core electrons and the length of chemical bonds.

As a consequence, the development of powerful X-ray and ultraviolet sources has been for many decades a high priority for research in many different disciplines. Progress in this field, however, has been rather slow until the advent of the sources in the class broadly labeled as "synchrotrons" in the late 1960s.

A conventional, nonsynchrotron source emits X-rays or ultraviolet photons by bringing electrons to an excited state and then exploiting their radiative decay. Considering the photon energies to be emitted, the process must involve core electrons. Its effectiveness is limited: for example, pumping core electrons to an excited state requires putting a high power in the system (e.g., by bombarding it with electrons) – but the power is limited by the thermal stability of the medium.

Furthermore, the photon emission occurs over a broad angular range. This limits the collimation of the source and its usefulness in experiments that require not only a high photon flux but also a high intensity.

The best parameter to characterize the source quality is the so-called "brightness" or "brilliance," b. Roughly speaking (**Figure 1**), this parameter is proportional to the flux F and inversely proportional to the source area A and to the angular range of emission Ω:

$$b \propto \frac{F}{A\Omega} \qquad [1]$$

(more precisely, F is the number of photons per second emitted in a given bandwidth of 0.1%; A is the product of the horizontal and vertical source sizes; Ω is the product of the horizontal and vertical angular spreads of the emitted photon beam, and $b = F/(4\pi^2 A\Omega)$).

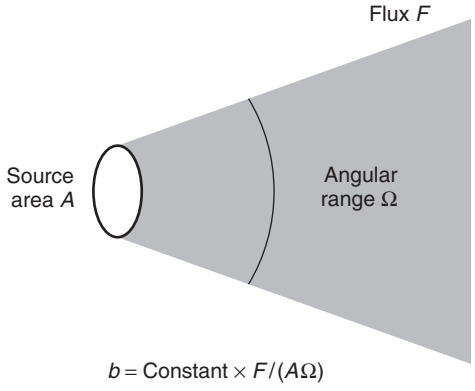


Figure 1 The brightness b is the parameter which is most often used to characterize the quality of photon sources, including synchrotron sources.

A limited flux and a large angular range result in a small brightness. Limited brightness was in fact a major problem in the experimental use of X-rays and ultraviolet photons until the advent of synchrotrons. For example, a crystallographic procedure that today is implemented with a synchrotron source in a few minutes would have required days or months – or would have been impossible.

For visible light, high brightness can be obtained from lasers. This is not easy for X-rays and ultraviolet photons. As discussed later, the effectiveness of the lasing mechanisms decreases, and the fabrication of optical cavities become problematic. Only recently have high-peak-brightness short-wavelength X-ray lasers started to appear technically feasible.

The sources in the "synchrotron" class solve the problem of increasing the flux and the brightness with an altogether different strategy with respect to classical sources. This strategy is based on a clever use of relativistic effects and therefore requires electrons or positrons moving at a speed close to the speed of light, c.

The Relativistic Basis of Synchrotron Sources

Consider (**Figure 2**) a very simple emitter of electromagnetic waves: an electron moving along a circular orbit under the action of a constant magnetic B-field (cyclotron motion). The emitted photon energy $h\nu$ is determined by the cyclotron frequency which, according to classical physics is given by $\nu = eB/(2\pi m_0)$ – where e and m_0 are the charge and the rest mass of the electron. This result derives, as it is well known, from the balance between the Lorentz force and the mass times the centripetal acceleration. The typical emitted photons have rather low energies with respect to X-rays.

Imagine, however, the same phenomenon when the speed of the circulating electron becomes close to the speed of light. In an inertial reference frame moving with the same velocity and speed (v) as the electron at a given time, the transverse Lorentz force vB of the B-field becomes a transverse electrostatic force of magnitude γvB, where the relativistic γ-factor (the energy of the moving electron divided by m_0c^2) is given by

$$\gamma = \frac{1}{\sqrt{1 - (v^2/c^2)}} \qquad [2]$$

As a consequence, relativistic dynamics predicts a cyclotron frequency $\gamma eB/2\pi m_0$, which is γ times larger than the classical result. The emitted photon energy is also multiplied by this factor in the reference frame of the electron, shifting toward the ultraviolet and X-ray spectral domains.

This shift toward shorter wavelengths and larger photon energies is further enhanced if one observes the emission in the reference frame of the laboratory. Imagine, for example, that the emitted photon energy is measured in the laboratory frame along the direction of the electron velocity. Since the source (the electron) is in motion, the photon energy is affected by the Doppler effect, which for electromagnetic waves is also a relativistic effect.

In the "forward" direction of the electron (source) velocity, the Doppler effect multiplies the photon energy by a factor $[(1 + v/c)/(1 - v/c)]^{1/2} = \gamma\ (1 + v/c)$. In the extreme relativistic case $v^2/c^2 \approx 1$, this factor is close to 2γ. Overall, therefore, the classical emitted photon energy for the cyclotron motion is relativistically increased by a factor $\approx 2\gamma^2$.

The mechanism just described is the one active in the photon emission by "bending magnets" (**Figure 3**). These are the dipole magnets that keep the electrons circulating in a closed orbit within an electron accelerator such as a storage ring. The deflection of the

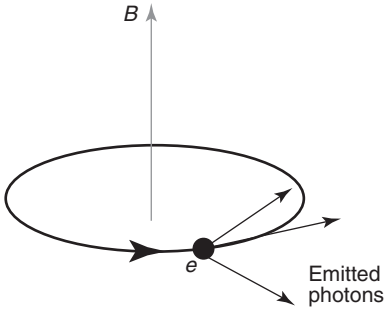


Figure 2 The emission of photons by an electron during its cyclotron motion is a good basis to analyze the mechanism of synchrotron sources. The analysis, however, must be performed for a motion at relativistic velocity, $v \approx c$.

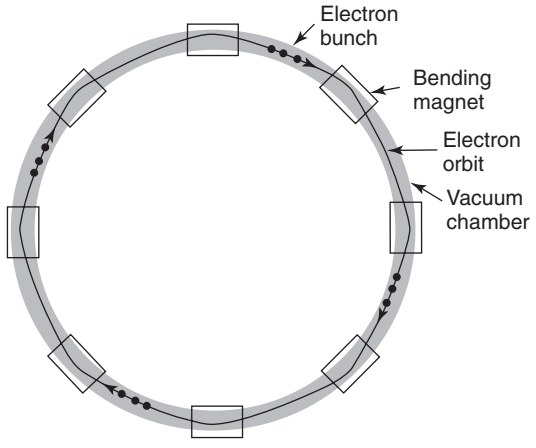


Figure 3 (Dipole) bending magnets are used to keep electrons (grouped in bunches) circulating along closed orbits in the vacuum chamber of a storage ring (top view).

electron trajectory causes a (centripetal) acceleration and therefore the emission of electromagnetic waves. The mechanism is roughly equivalent to that of a portion of a cyclotron orbit, and the characteristics of the emission are also similar – although a rigorous description requires a sophisticated electrodynamics treatment.

In the first approximation, therefore, the center of the spectrum emitted by a bending magnet is close to $2\gamma^2$ multiplied by the classical photon energy given by the cyclotron-frequency, $h\nu = heB/2\pi m_0$. For a typical B-field of 1 T, the classical photon energy would be $\sim 4 \times 10^{-23}$ J or ~ 0.2 meV, well below the ultraviolet and X-ray range. If one takes a typical electron energy of 1 GeV in a storage ring, the γ-value is $\sim 2 \times 10^3$, the factor $2\gamma^2$ becomes $\sim 4 \times 10^6$ and the relativistic emission becomes $\sim 2 \times 10^2$ eV, which falls in the X-ray range.

As discussed later, all types of synchrotron sources exploit this combination of relativistic effects to shift their emission into the ultraviolet and X-ray ranges. In addition to the spectral shift, relativity has another beneficial effect: a shrinkage of the angular range of emission and a corresponding (eqn [1]) increase in the brightness.

This angular-range shrinkage can be easily understood. As shown in **Figure 4a**, a photon emitted in the vertical plane, when observed from the reference frame of the emitting electron, can travel in a direction quite far from that of the electron velocity ("forward" direction). When seen from the laboratory frame, the photon direction changes becoming very close to the same "forward" direction.

The reason is that the relativistic Lorentz transformations include a γ-factor for the "forward" direction coordinate but not for the transverse direction coordinate. When calculating the angle between the photon direction and the "forward" direction as the ratio of the two velocity components, the γ-factor appears in the denominator but not in the numerator. This confines the emission within an angular range $\sim 1/\gamma$ from the "forward" direction. Because of the large value of γ, this relativistic effect strongly collimates the emission enhancing the brightness.

Note, however, that for bending magnets the angular collimation is really only present in the vertical direction. In the horizontal direction (**Figure 4b**), the angular range depends on the portion of the electron orbit "seen" by the detector.

Finally, it should be noted that the brightness (eqn [1]) is also enhanced because relativity boosts the flux. Electrodynamics shows that the power emitted by an accelerated charge is proportional to the square of the acceleration. In a relativistic cyclotron motion, the acceleration is roughly proportional to the speed of light multiplied by the cyclotron frequency, which in turn – as seen – is proportional to γB. Thus, the flux is proportional to $\gamma^2 B^2$, and the large γ-value enables it to reach very high levels.

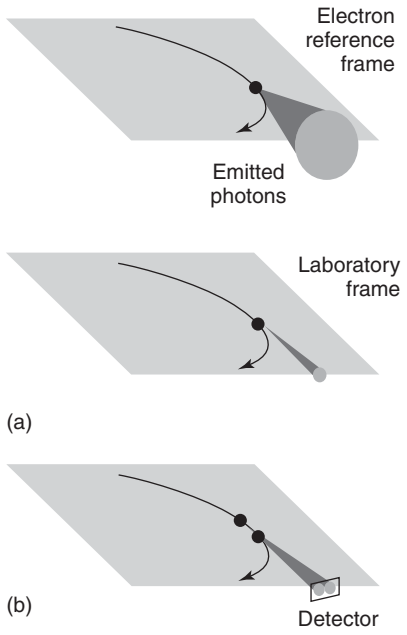


Figure 4 (a) The emission of photons by an electron, whose trajectory is deviated by a bending magnet, occurs over a wide angular range in the electron reference frame. Relativity drastically reduces this angular range when the emission is observed in the laboratory frame. (b) In practice, however, the angular collimation is only present for the vertical direction. In the horizontal direction, electrons at different points of the orbits emit in different directions and the horizontal angular range is set by the portion of the orbit "seen" by the detector.

Main Properties of Synchrotron Sources

Based on the above discussion, one can summarize the main properties of bending magnets and other synchrotron sources in the following way:

- *Spectral domain*: centered in the ultraviolet and X-ray region because of the relativistic factor $\sim 2\gamma^2$ in the emitted photon energies.
- *Bandwidth*: a bending magnet emits over a wide photon energy bandwidth. The reason can be understood by considering an emitting electron as it passes through the bending magnet. Due to the angular collimation, the electron can be imagined as a torchlight with a very narrow beam. A detector (**Figure 5a**) is illuminated by this "torchlight" for a very short time. The corresponding bandwidths of frequencies and photon energies (according to the Fourier theorem and to the uncertainty principle) are inversely proportional to the illumination time and therefore quite large.

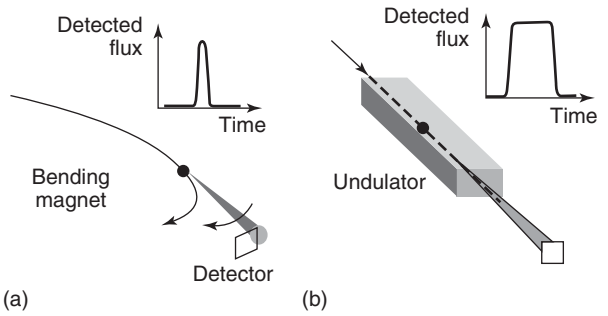


Figure 5 (a) An electron going through a bending magnet emits photons that illuminate the detector only during a short time pulse. This corresponds to a broad photon energy bandwidth. (b) An electron passing through an undulator, instead, illuminates the detector continuously during a longer pulse, producing a narrow bandwidth. The undulator bandwidth is further decreased by other factors.

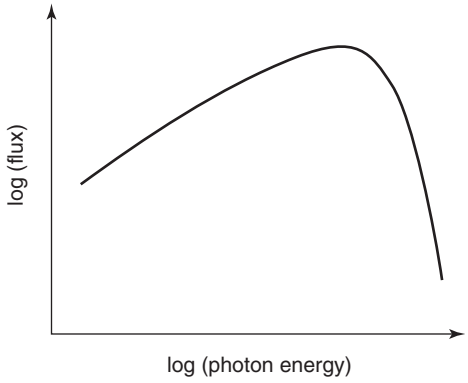


Figure 6 In first approximation, the spectrum emitted by a bending magnet can be imagined as a broad peak. The conventional way to show the spectrum, however, is a log–log plot that transforms a peak-like line shape into the familiar spectrum of bending-magnet synchrotron sources.

A detailed calculation shows that for a bending magnet the photon energy bandwidth is close to the relativistic "cyclotron" photon energy $\approx 2\gamma^2 heB/2\pi m_0$ – so that the relative bandwidth $\Delta h\nu/(h\nu) \approx 1$. In other types of synchrotron sources such as the undulators (see **Figure 5b** and the discussion later), the bandwidth is narrow rather than broad as in bending magnets. Overall, the emitted spectrum of a bending magnet can be roughly imagined as a broad peak with center close to $\sim 2\gamma^2 heB/2\pi m_0$ and $\Delta h\nu/(h\nu) \approx 1$. As shown in **Figure 6**, a log–log plot of this peak gives the familiar synchrotron spectrum.

- *Collimation*: the emission is limited to an angular range $\approx 1/\gamma$.
- *Flux*: high and proportional to $\gamma^2 B^2$.
- *Brightness*: very high because of the high flux and high collimation. In addition, the brightness is also enhanced by the small source area A (eqn [1]). The effective A is determined by the transverse section of the electron beam circulating in the accelerator. In modern storage rings as well as in other accelerators, the trajectories of the circulating electrons are controlled with excellent accuracy. This strongly limits A boosting the brightness.
- *Coherence*: the lateral coherence of a source of electromagnetic radiation increases as the product of the source size times the angular range decreases, until it reaches the "diffraction limit" (for which the product roughly equals the wavelength). Thus, the same geometric factors that influence the brightness (eqn [1]) also determine the lateral coherence. For a bending magnet, the high angular collimation in the vertical direction produces high lateral coherence in the same direction. The longitudinal coherence is determined by the relative bandwidth $\Delta h\nu/(h\nu)$ and, for a bending magnet, it is quite limited. However, the longitudinal coherence can be increased by reducing the bandwidth with a monochromator.

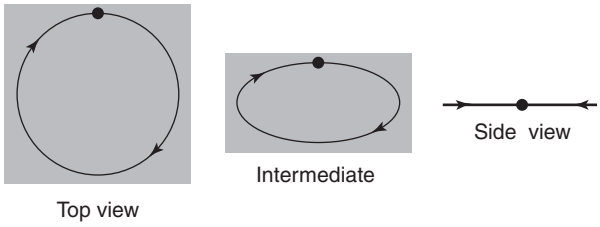


Figure 7 The polarization of the bending magnet emission can be understood by imagining to observe from different points of view an electron moving in the storage ring. It is quite clear, for example, why in the ring plane the emission is linearly polarized.

- *Polarization*: for a bending magnet, the limited angular range confines the emission to directions close to the (horizontal) plane of the storage ring. From their point of view, the emitting electrons (**Figure 7**, right) appear to move along a horizontal line, and their emission is linearly polarized with the photon electric field in the horizontal plane. Slightly out of the plane (**Figure 7**, middle), the observed electron motion is elliptical and the emitted light possesses a circularly polarized component. As discussed later, other synchrotron sources are also linearly or circularly polarized.
- *Time structure*: the electrons circulate around a storage ring grouped in bunches. Every time a bunch passes through a source point such as a bending magnet, a pulse of photons is emitted. Thus, the time structure of a synchrotron source consists of short pulses separated by longer "dark" intervals. The parameters of the time structure depend on those of the storage ring, and the pulse duration can reach in some cases the picosecond range.

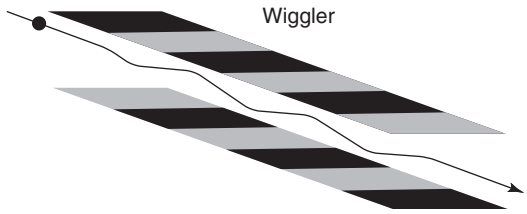


Figure 8 A wiggler is a periodic series of magnets inserted along a straight section of the storage ring. The magnets cause lateral undulations or "wiggles" of the electrons with respect to their otherwise straight trajectory. The resulting photon emission is similar to that of a series of bending magnets.

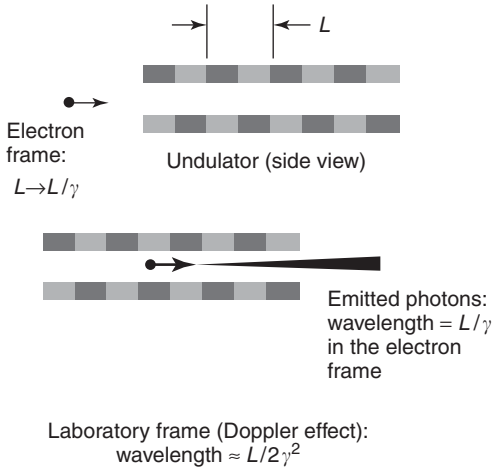


Figure 9 Schematic explanation of how the two relativistic effects (Lorentz contraction of the undulator period and Doppler shift) determine the "central" emitted wavelength and photon energy of an undulator.

Different Types of Synchrotron Sources

The previous discussion primarily dealt with bending magnet sources. Other types of synchrotron sources exist and are briefly reviewed here.

Wigglers

A wiggler (**Figure 8**) is a periodic sequence of dipole magnets inserted in an otherwise straight portion of a storage ring. It produces a series of transverse "wiggles" in the electron trajectory, each roughly equivalent to the electron trajectory in a bending magnet. The emission is thus approximately equivalent to that of a series of bending magnets.

The positive aspects of a wiggler are mainly two. First, the flux and the brightness are multiplied with respect to a single bending magnet by the number of wiggles. Second, the B-field strength can be selected to tune the emission spectrum to a desired value (e.g., to very high photon energies in the hard-X-ray range). This is not possible for bending magnets since their field strength is dictated by the task of keeping the electrons in a closed trajectory around the ring.

Undulators

An undulator is a device similar to a wiggler except that the B-field strength is more limited and the "wiggles" along the electron trajectory are less pronounced. As a consequence, the emission properties are quite different from those of wigglers or bending magnets.

Roughly speaking, the "wiggles" in this case produce angular deviations smaller than $1/\gamma$ with respect to the axis of the undulator. Since the angular spread of the emission is $\sim 1/\gamma$, the emission corresponding to one "wiggle" illuminates the subsequent wiggles, producing coherent interference. As a result, the emitted peak intensity increases with the square of the number of wiggles rather than being proportional to it.

The photon energy spectrum of an undulator is also quite different from that of a wiggler or a bending magnet. First, the bandwidth is much narrower. The reason (**Figure 5b**) is that for weak wiggles the detector is illuminated during the entire passage of the electron-torchlight through the undulator. The resulting longer pulse corresponds to a narrower photon energy bandwidth. This effect is further enhanced by other factors. In essence, therefore, an undulator concentrates the emitted flux of a wiggler into a narrow bandwidth, achieving very high brightness.

This narrow bandwidth is centered at a photon energy which is determined by the undulator period rather than by other parameters (see **Figure 9**). The physical explanation is again rather simple: in the reference frame of an electron moving at relativistic speed through the undulator, the periodic transverse B-field is observed as the combination of a periodic transverse B-field plus a transverse periodic electric field perpendicular to the B-field, both moving at a speed $\sim c$. In other words, it is observed as an electromagnetic wave.

Interacting with this wave, the electron emits waves with the same wavelength. In its reference frame, this is the undulator period L shortened by γ because of the Lorentz contraction. When observed in the laboratory frame, the wavelength is further decreased by a factor $\sim 2\gamma$ by the Doppler effect, becoming $\approx L/(2\gamma^2)$. The corresponding photon energy is $\approx 2\gamma^2 hc/L$.

In addition to this photon energy, the undulator also emits higher harmonics corresponding to its integer multiples. As the B-field increases, the sequence of higher harmonic merges into one broad spectrum corresponding to the wiggler emission.

The fundamental-harmonic photon energy $\approx 2\gamma^2 hc/L$ of an undulator is shifted when the emission is detected along a direction other than the undulator

axis, or when the B-field changes. The reason is similar in both cases: at an angle θ from the "forward" direction, the Doppler factor is no longer $\approx 2\gamma$ and becomes a function of θ. Similarly, as the B-field increases the transverse wiggles decrease the electron speed along the undulator axis and effectively modify the γ-factor in the Doppler shift.

Taking both effects into account, the first-harmonic photon energy of the undulator changes from $2\gamma^2 hc/L$ to

$$h\nu = \frac{2\gamma^2 hc}{L}\left(1 + \frac{K^2}{2} + \gamma^2\theta^2\right) \qquad [3]$$

where the K-factor depends on B: $K \propto BL$. Equation [3] has a very important impact on the operation of undulator sources: by changing the B-field (e.g., the gap between permanent magnets), one can tune the emitted photon energy to a desired value.

Contrary to bending magnets, undulators achieve high angular collimation not only on the vertical direction but also horizontally. In fact, during the entire emission process, the electron stays in a straight trajectory except for small transverse wiggles, and this limits the horizontal angular divergence. This also implies that the lateral coherence is high both vertically and horizontally.

Finally, the undulator emission is linearly polarized with the electric vector in the horizontal plane. The reason is that, when observed from the point of view of the detector, an emitting electron looks like an oscillating charge along a horizontal line.

Helical Wigglers

For many applications of synchrotron sources, it is necessary to have circular polarization. It is seen that circular polarization does exist in the out-of-plane emission of bending magnets. However, the emitted intensity strongly decreases as the direction diverges from the horizontal plane. Therefore, bending magnets can only provide circularly polarized photon beams of moderate intensity.

Dedicated sources for circular polarization are based on special magnet arrays that force the electrons to move along a periodic helical path rather than simply undulating in the transverse direction. Such devices can operate either in the wiggler mode ("helical wigglers") or as undulators.

Linear Accelerator (LINAC) Sources

Almost all synchrotron sources are based at present on electron storage rings. The reason is that electrons in such a ring can circulate after injection for a very long period of time. Photons are thus produced by the same electron beam for hours or days, and this drastically decreases the operation cost.

On the other hand, the emission of photons interferes with the operation of the storage ring limiting the capability of optimizing the electron beam geometry for photon emission. This problem is less severe for LINACs where the electron beam passes only once through the accelerator. However, the electron beam is also used only once to produce photons and this increases the operation cost (notably, the electric power) to prohibitive levels.

A new strategy is emerging in which the electron beam after passing through the accelerator is slowed down giving back most of its energy to the system. With this "energy recovery" scheme the cost of using a LINAC could be reduced to an acceptable level.

Free-Electron Lasers

Standard synchrotron sources are not based on a laser mechanism: they do possess laser-like properties such as high collimation, but they are not lasers. There exists, however, a special class of synchrotron-related sources that operate as lasers: the free-electron lasers (FELs).

A laser is based on two ingredients: optical amplification and the optical cavity. Optical amplification can be achieved in a variety of solid, gas, and liquid media. In the case of FELs, the optically amplifying medium is a relativistic electron beam during its interaction with a wiggler. Under adequate geometric conditions, the electron beam can produce sufficient optical amplification for lasing action.

This approach is implemented with different types of electron accelerators and is widely used to produce high-quality photon beams in the infrared range. However, its implementation becomes more difficult at shorter wavelengths in the ultraviolet and X-ray ranges. In fact, the effectiveness of the FEL optical amplification mechanism decreases with the wavelength. Furthermore, one cannot fabricate an optical cavity for short wavelengths since normal-incidence reflection is extremely limited. Therefore, the optical amplification must be sufficient to produce laser action without the cavity.

The most effective way to overcome these difficulties is the SASE (self-amplified spontaneous emission) mechanism, in which extremely good geometric conditions are achieved for the photon beam in a superconducting LINAC. The corresponding devices are called SASE X-FELs. The SASE mechanism has already been used to produce laser beams with increasing photon energies.

One should note that X-FELs cannot replace synchrotron sources based on storage rings. The lasing

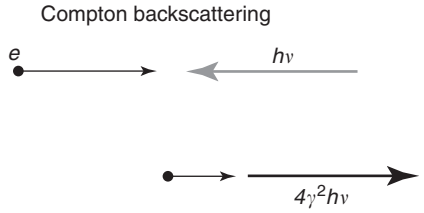


Figure 10 Compton backscattering: a laser beam of photons is sent head-on against a bunch of relativistic electrons. The laser beam is backscattered and its photon energy changes from the visible or infrared domain to X-rays.

mechanism, in fact, produces very high peak brightness over ultrashort pulses (that could reach 100 fs or less). These pulses are excellent for applications that require a high peak brightness but not useful when longer pulses or a quasicontinuum emission is required.

Inverse Compton Scattering Sources

Synchrotron sources have been often proposed for practical applications beyond research, for example in medical radiology. The major obstacle is that standard synchrotron sources have a large size and a high cost. Therefore, they cannot be simply added to the instrumentation of another facility such as a hospital.

The efforts to develop compact storage rings were not too successful, since a small storage ring of sufficiently high energy becomes vulnerable to instabilities and is generally difficult to operate. A promising alternate strategy is based on the inverse Compton scattering mechanism (**Figure 10**).

Imagine a relativistic electron beam sent head-on against a low-energy photon beam. The photons are Compton-backscattered shortening their wavelength. This phenomenon can be easily understood in terms of two Doppler shifts: first, the incoming photon beam is Doppler-shifted in the electron reference frame. Then, the electrons become the sources of the backscattered photons which appear again Doppler-shifted in the laboratory frame. This means that the overall shift factor is $\approx(2\gamma)\ (2\gamma)=4\gamma^2$: an incoming photon energy $h\nu$ in the infrared range can thus produce an X-ray beam with photon energy $4\gamma^2 h\nu$. Depending on the incoming photon energy $h\nu$, the production of X-rays can thus be achieved even with a compact accelerator with moderate γ-factor.

Compton backscattering sources will not be able to reach high brightness and coherence levels. However, they could be suitable for practical applications such as synchrotron radiology, and their size and cost could fit the use in a hospital.

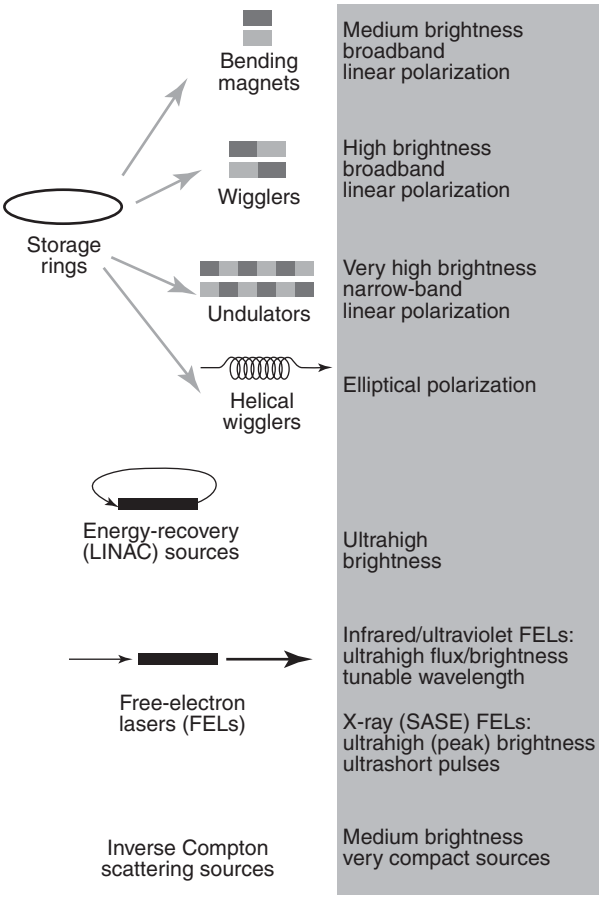


Figure 11 Schematic summary of what can be obtained from different types of synchrotron-related sources.

The Overall Picture

The different types of synchrotron or synchrotron-related sources, either already implemented or under development, provide extreme flexibility in matching the needs of a wide variety of applications. **Figure 11** presents a summary overview of the characteristics of the emission of the different machines.

Practical Use of Synchrotron Sources

Synchrotron sources ceased long ago to be exotic instruments, only used by a tiny minority of scientists. They have become instead standard instruments used by a large part of the research community in diverse disciplines. Dozens of facilities are in operation in many different countries, most of them at no cost for qualified users performing nonproprietary research. Additional programs, for example, those supported by the European Commission, provide support to users for travel and lodging expenses.

Each synchrotron facility serves its users with a variety of instruments. Roughly speaking, the core of the facility is an accelerator such as a storage ring

with its injection and control systems. The emitted photons are collected by beamlines that improve the photon beam characteristics (e.g., by selecting a narrow band of photon energies with suitable monochromators) and convey them to experimental chambers. In the accelerator, the electrons must circulate under ultrahigh vacuum. Many of the beamlines are also under vacuum since no window can be used to separate them from the accelerator without decreasing or completely eliminating the photon beam intensity.

The typical operation cycle of a synchrotron facility begins with the injection of electrons (or positrons) in the accelerator. If this is a storage ring, the electron beam keeps circulating and emitting photons for hours or days and the energy lost through photon emission is restored by radiofrequency cavities. Natural losses steadily reduce the circulating electron beam until a new injection and the beginning of a new operation cycle. In a very few cases such as the Swiss light source (SLS), the injection is performed continuously so that the electron beam can circulate indefinitely and experiments have no time limitations.

See also: Core Photoemission; Crystal Structure Determination; Molecular Crystallography; Multicircle Diffractometry Methods; Photoelectron Spectromicroscopy; Relativity; Scattering, Inelastic: X-Ray (Methods and Applications); Valence Photoemission; X-Ray Absorption Spectroscopy; X-Ray Standing Wave Techniques; X-Ray Topography.

PACS: 07.85. – m; 07.85.Qe; 41.60.Cr; 41.60. – m; 41.85.Lc

Further Reading

Carroll FE (2002) Tunable monochromatic X-rays: a new paradigm in medicine. *American Journal of Roentgenology* 179: 583–590.

D'Amico KL, Terminello LJ, and Shuh DK (eds.) (1996) *Synchrotron Radiation Techniques in Industrial, Chemical, and Materials Science*. New York: Plenum.

Duke PJ (2000) *Synchrotron Radiation: Production and Properties*. New York: Oxford University Press.

Helliwell JR (1997) *Time-Resolved Diffraction*. New York: Oxford University Press.

Lowesey SW (1996) *X-ray Scattering and Absorption by Magnetic Materials*. New York: Oxford University Press.

Margaritondo G (1988) *Introduction to Synchrotron Radiation*. New York: Oxford University Press.

Margaritondo G (2003) *Elements of Synchrotron Radiation for Chemistry, Biology and Medical Research*. New York: Oxford University Press.

Mills DM (ed.) (2002) *Third-Generation Hard* X-Ray *Synchrotron Radiation Sources: Source Properties, Optics, and Experimental Techniques*. Hoboken: Wiley.

Saisho H and Gohshi Y (1996) *Applications of Synchrotron Radiation to Materials Analysis*. Amsterdam: Elsevier.

Sham TK (2002) *Chemical Applications of Synchrotron Radiation (Advanced Series in Physical Chemistry, V 12A–12B)*. Singapore: World Scientific.

Wiedemann H (2002) *Synchrotron Radiation*. Heidelberg: Springer.

T

Tensile Properties *See* Mechanical Properties: Tensile Properties.

Thermal Conductivity at Low Temperatures

T Skośkiewicz, Polish Academy of Sciences, Warsaw, Poland

Heat transport in solids is characterized by thermal conductivity coefficient κ defined by the formula

$$q = -\kappa \nabla T \qquad [1]$$

Heat flow per unit cross-sectional area is proportional to the gradient of temperature. The "minus" sign in eqn [1] indicates that heat flows in the direction opposite to the temperature gradient. There are two main carriers involved in heat transport in solids: thermal vibrations (called phonons) and electrons. The carriers might be scattered by phonons and lattice imperfections. It is assumed that phonons as well as electrons, upon being scattered, diffuse through the material. The simple formula for thermal conductivity coefficient comes from the kinetic theory of gases:

$$\kappa = \tfrac{1}{3}(c\, l\, v) \qquad [2]$$

where v is velocity, l the mean free path of heat carriers, and c the specific heat per unit volume. In case of phonons, v is sound velocity. The typical value of sound velocity in solids is $3\text{–}5 \times 10^3\,\mathrm{m\,s^{-1}}$. Phonons are involved in heat transportation in all materials. Electrons are particularly important in metals. Thermal excitations of electrons in metals are only possible for electrons occupying energy levels close to the Fermi level. Only unoccupied electronic states available in the energy range of the order of thermal excitations participate in the heat transport process. The typical value of Fermi velocity calculated for metals is of the order $10^6\,\mathrm{m\,s^{-1}}$, which is much larger than the velocity for phonons. This is why metals are among the best solid-state heat conductors. At low temperatures, velocity of phonons and Fermi velocity of electrons do not depend on temperature. The heat capacity of electrons might be calculated by employing the free-electron theory; c is proportional to temperature. The heat capacity of phonons, according to the Debye model, is proportional to T^3. The temperature dependence of mean free path for heat carriers depends on the scattering processes. The following scattering processes will be considered; phonon–phonon, phonon–crystal defect, electron–phonon, and electron–impurity. A good assumption is that thermal resistances are additive. Scattering of heat carriers on phonons is temperature dependent because the number of phonons depends on the temperature, leading to a temperature-dependent mean free path. Scattering on defects or impurities results in a mean free path independent of temperature. The thermal conductivity of some specific types of materials is considered in the following sections.

Insulating Crystals

Thermal conductivity in insulating crystals is dominated by phonons. There are two very distinct temperature ranges. At intermediate temperatures, about one-tenth of the Debye temperature, the specific heat of phonons is temperature independent and the mean free path for phonon scattering decreases as $\sim 1/T$ with temperature. According to formula [2], one can expect similar temperature dependence of the thermal conductivity coefficient in this temperature range. At low temperatures, $T \ll \Theta$, the number of phonons is low and phonon–phonon scattering is strongly reduced. Scattering on crystal defects and grain boundaries becomes dominant. The mean free path of phonons for these scattering processes does not depend on temperature. The specific heat in this temperature range is proportional to T^3; thus the thermal conductivity coefficient is

$$K_{\mathrm{ph}} \propto T^3 \quad \text{for } T \ll \Theta \qquad [3]$$

As a result, at intermediate temperatures, a broad maximum of thermal conductivity coefficient is observed. The position of the maximum and the

thermal conductivity coefficient values at low temperatures strongly depend on the purity of the crystal and the number of defects. Examples of thermal conductivity coefficient versus temperature are given in **Figure 1a**. For some insulating crystals, such as diamond and sapphire, with very low concentration of defects and impurities, the thermal conductivity coefficient at the maximum might be larger than that for highly conductive metals such as copper and silver.

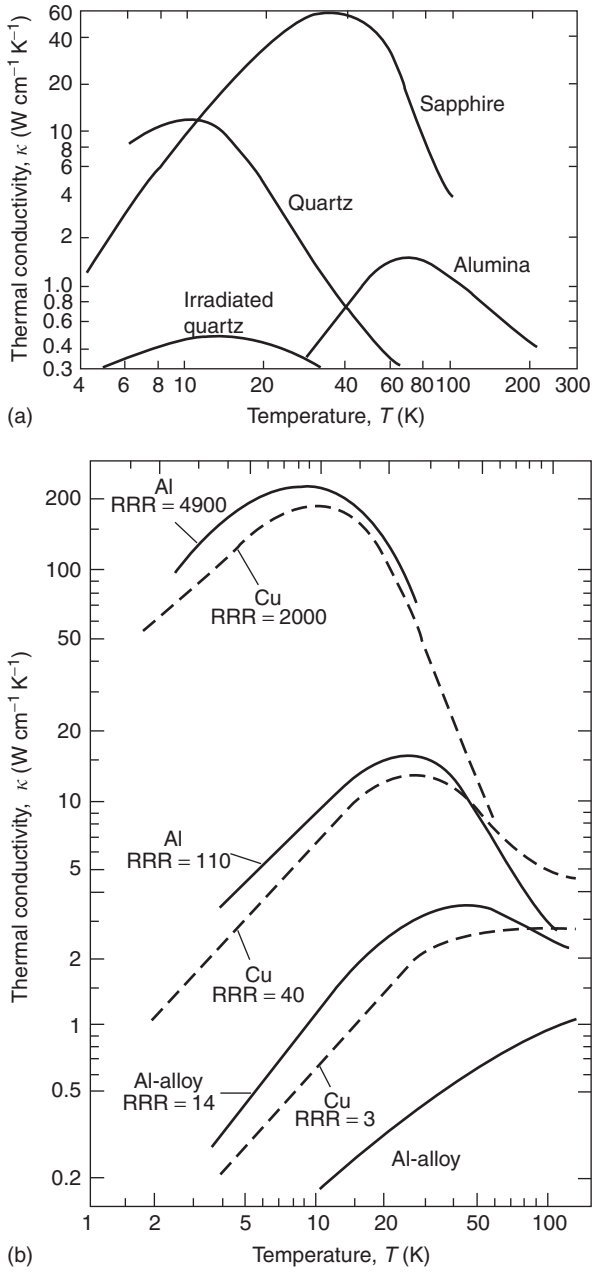


Figure 1 Temperature dependence of the thermal conductivities of (a) some dielectric solids and (b) Al and Cu of varying purity (expressed as their residual resistivity ratio RRR. (Pobell F (1992) *Matter and Methods at Low Temperatures*. Berlin: Springer.)

Metals

Heat transport in metals is mainly due to electrons. The thermal conductivity coefficient is given by formula [2], where c is the electron specific heat proportional to temperature, v the Fermi velocity (independent of temperature), and l the electron mean free path. At temperatures less than 10 K, electrons are mainly scattered at impurity atoms and crystal defects. It makes the electron mean free path to be independent of temperature. Thus the thermal conductivity coefficient K_{el} is proportional to T. At higher temperatures, scattering at phonons has to be taken into account. The number of phonons increases with temperature, which decreases the mean free path of electrons and the thermal conductivity of metals. Here again a broad maximum of thermal conductivity coefficient is observed, but at temperatures lower than that for phonons in dielectric crystals (see **Figure 1b**). In heavily disordered alloys, the electronic thermal conductivity might be reduced so strongly that it becomes comparable with lattice conductivity due to phonons.

Wiedemann–Franz Law

In metals at low temperatures, electrons carry both heat and electric current. Electrical conductivity is independent of temperature, while specific heat of electrons and thermal conductivity coefficient are proportional to temperature. The ratio of thermal conductivity coefficient and electrical conductivity is proportional to temperature

$$\kappa/\sigma = L_0 T \qquad [4]$$

where $L_0 = (\pi k_B/e)^2/3 = 2.45\,\mathrm{W\,\Omega\,K^{-2}}$ is a universal constant called the Lorenz number. The Wiedemann–Franz law [4] holds for low temperatures but is also valid for high temperatures $T > \Theta_D$, where electron transport is limited by large-angle electron–phonon scattering. It is sometimes used for estimation of thermal conductivity coefficient from electrical resistivity measurements.

Noncrystalline Solids

Thermal conductivity of noncrystalline solids is a very special case of heat transport via phonons. An almost universal T^2 temperature dependence of thermal conductivity coefficient is observed for nonmetal glassy solids. Examples are given in **Figure 2**. The

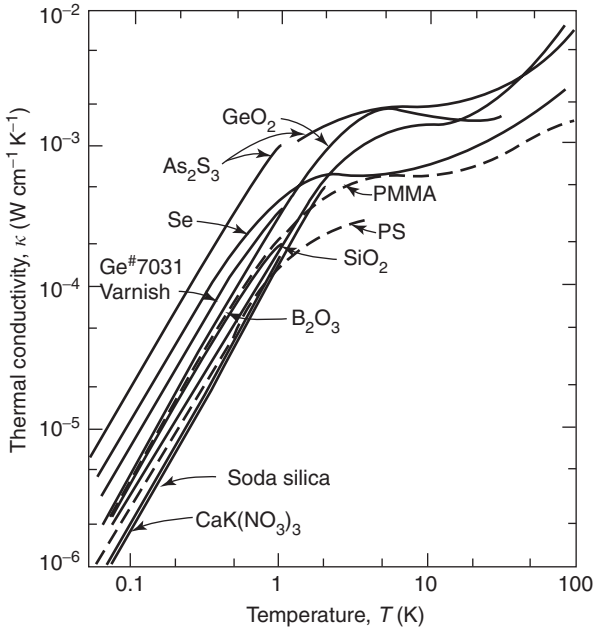


Figure 2 Thermal conductivities of various noncrystalline solids. (Pobell F (1992) *Matter and Methods at Low Temperatures*. Berlin: Springer.)

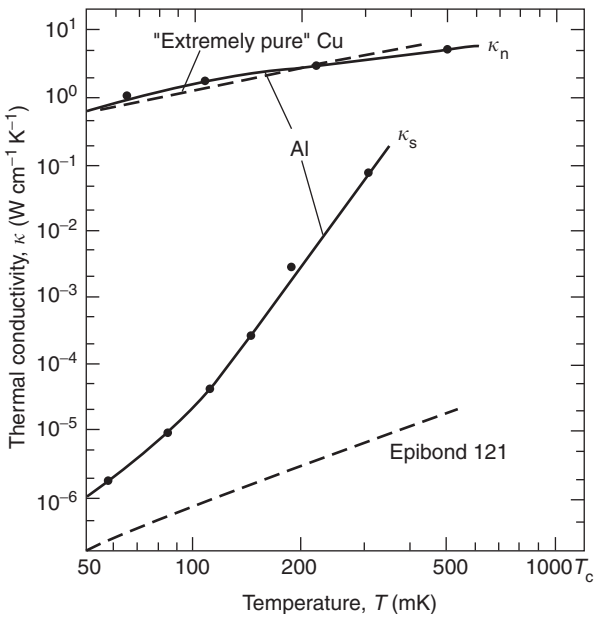


Figure 3 Thermal conductivity of Al in the normal conducting state (compared to Cu) and in the superconducting state (compared to dielectric Epibond 121) at $T > 50$ mK. (Pobell F (1992) *Matter and Methods at Low Temperatures*. Berlin: Springer.)

specific heat of amorphous solids has been observed to vary linearly with temperature below 1 K (e.g., at 25 mK, it is 10^3 times larger than Debye specific heat of crystalline solids). P W Anderson and W A Phillips independently proposed a mechanism for thermal excitations leading to an unusual temperature dependence of specific heat of glasses. They assumed that, in disordered glassy phase, many atoms have more than one position available. Almost equivalent positions are separated by an energy barrier. At low temperatures, there is a finite probability for an atom to tunnel from one position to another. The tunneling gives large contribution to the specific heat, which is proportional to temperature. The mean free path of heat carriers (phonons) is limited by such two energy level scattering centers. Universal T^2 temperature dependence and similar absolute values of thermal conductivity coefficient for various noncrystalline dielectrics show that disorder is the limiting factor for phonon heat transport. Electron heat transport in glassy metals might be negligible at low temperatures.

Superconductors

In the superconducting state of a metal, some of the electrons occupying states in the vicinity of the Fermi level form Cooper pairs. All Cooper pairs being bosons, they condense to the ground state, separated from the excited states of normal electrons by an energy gap ΔE. The pairs cannot be scattered unless the excitation is large enough to break a pair into single electrons, and this is why Cooper pairs cannot carry heat. Only the remaining unpaired electrons in the vicinity of the Fermi level might be carriers in heat transport. The number of these electrons decreases rapidly with temperature and is proportional to $\exp(-\Delta E/\ k_B T)$. The thermal conductivity of a superconductor is a product of normal-state electron thermal conductivity and the number of unpaired electrons:

$$\kappa = T \exp(-\Delta E/k_B T) \qquad [5]$$

At temperatures much lower than the critical temperature, the electronic thermal conductivity of aluminum, one of the best thermal conductors at ambient temperature, is as low as the thermal conductivity of insulators. The thermal conductivity of nonsuperconducting and superconducting aluminum in the temperature range below 1 K is given in **Figure 3**. The thermal conductivity coefficient of copper and insulator Epibond 121 epoxy resin frequently used in low-temperature experiments is also given for comparison. At 50 mK, the thermal conductivity coefficient of superconducting aluminum is less than 10^{-5} of the coefficient for nonsuperconducting aluminum. This property of superconductors opens interesting possibilities in low-temperature experiments. A superconductor might be easily switched into normal state by applying a large magnetic field. A piece of aluminum wire might serve as a thermal switch. At zero magnetic field, the switch is in a superconducting

state and behaves like a thermal insulator, while in presence of the magnetic field its thermal conductivity increases by more than 10^5 times, becoming comparable with the conductivity of the best solid-state thermal conductors.

Nanostructures

Rapid progress in the synthesis and processing of materials in the nanometer length scale has created a demand for better understanding of thermal transport in nanostructures and nanostructured materials. Heat transport in nanoscale systems is still not well understood. Unusual values of thermal conductivity coefficient were experimentally observed. Extremely high thermal conductivity coefficient $6600\,\mathrm{W\,m^{-1}\,K^{-1}}$ was predicted for an isolated carbon nanotube. The room-temperature thermal conductivity coefficient as high as $3000\,\mathrm{W\,m^{-1}\,K^{-1}}$ was measured for a multiwall carbon nanotube with a diameter of 14 nm. Some recent attempts to measure the thermal conductivity of mats of carbon nanotubes gave disappointing results. Due to weak coupling between the carbon nanotubes, the measured thermal conductivity of the mats was found to be much less than expected. On the other hand, extremely low thermal conductivity was reported for mesoscopic GaAs beams supporting a nanostructure bolometric device. The thermal conductivity coefficient for beams of dimensions $0.25 \times 0.15 \times 6\,\mu\mathrm{m}$ (see **Figure 4**) was found to be 10^4 times less than for bulk GaAs.

Superfluid ^{4}He

Among the best heat conductors is superfluid ^{4}He at temperatures above 1 K. Its thermal conductivity might exceed the thermal conductivity of the best metal heat conductors by 10^3 times. The ability of superfluid helium to transport enormous amount of heat might be well explained in terms of a simple two-fluid model. According to the model, superfluid helium consists of two inseparable fluids – normal fluid and superfluid. The equilibrium concentrations of fluids depend on temperature. At lower temperatures, the concentration of the superfluid is larger. A superfluid is a zero-entropy fluid and does not carry

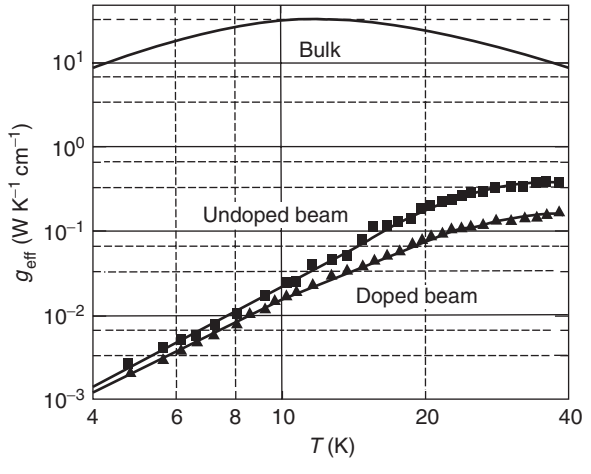


Figure 4 Comparison between the thermal conductivity of doped and undoped beams and that of bulk GaAs. (Fon W, *et al.* (2002) Phonon scattering mechanisms in suspended nanostructures from 2 to 40 K. *Physical Review B* 66: 45302–45305.)

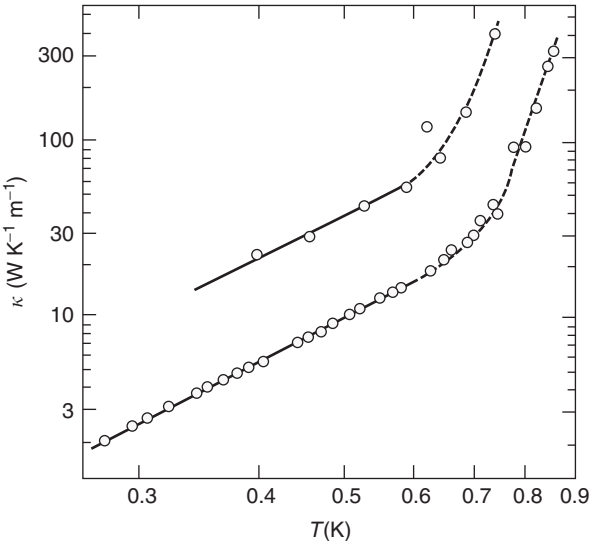


Figure 5 Thermal conductivity of liquid ^{4}He in two tubes, 0.80 mm (upper curve) and 0.29 mm (lower curve) in diameter. (Lounasmaa OY (1974) *Experimental Principles and Methods Below 1 K.* London: Academic Press.)

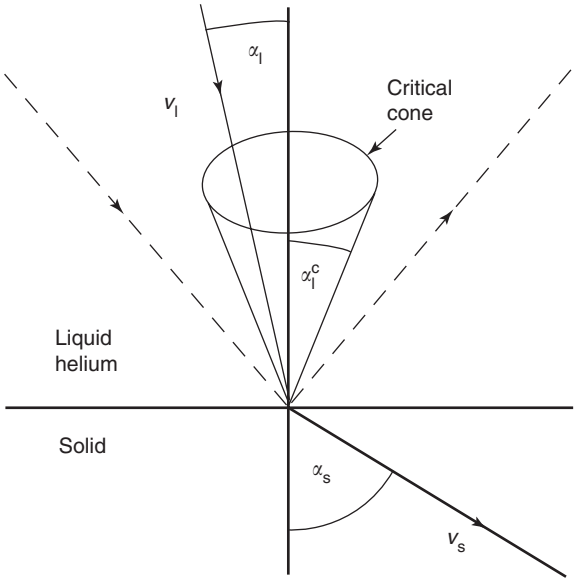


Figure 6 Phonons at the interface between liquid helium and solid body. In real systems the half angle of the critical cone α_l^c is less than 3°.

heat. If heat is supplied at some spot into superfluid helium, it must be transported away if the temperature is to remain constant. The most efficient way of doing this is converting the superfluid component into a normal one. Excess of normal fluid must flow away and be replaced by superfluid. The flow of normal component is a viscous one. Its viscosity increases and thermal conductivity of superfluid helium decreases when temperature is lowered. Experimental data on thermal conductivity of ^{4}He are given in **Figure 5**. Slopes of the curves presented change at 0.6 K, pointing that below this temperature, the process of internal convection described above becomes negligible for heat transport. Longitudinal phonons carry heat only at temperatures below 0.6 K. The specific heat of phonons is proportional to T^3. Since the mean free path of phonons and the sound velocity are independent of temperature, the thermal conductivity of superfluid helium below 0.6 K is proportional to T^3.

Kapitza Resistance

Piotr Kapitza discovered a very interesting phenomenon in 1941. He observed a thermal boundary resistance (the "Kapitza resistance") between liquid helium and a number of solids. The phenomenon, very specific for heat transport at very low temperatures, was explained in terms of acoustic mismatch theory by Kalatnikov. He considered phonon traveling in liquid helium with velocity v_l and striking the liquid–solid

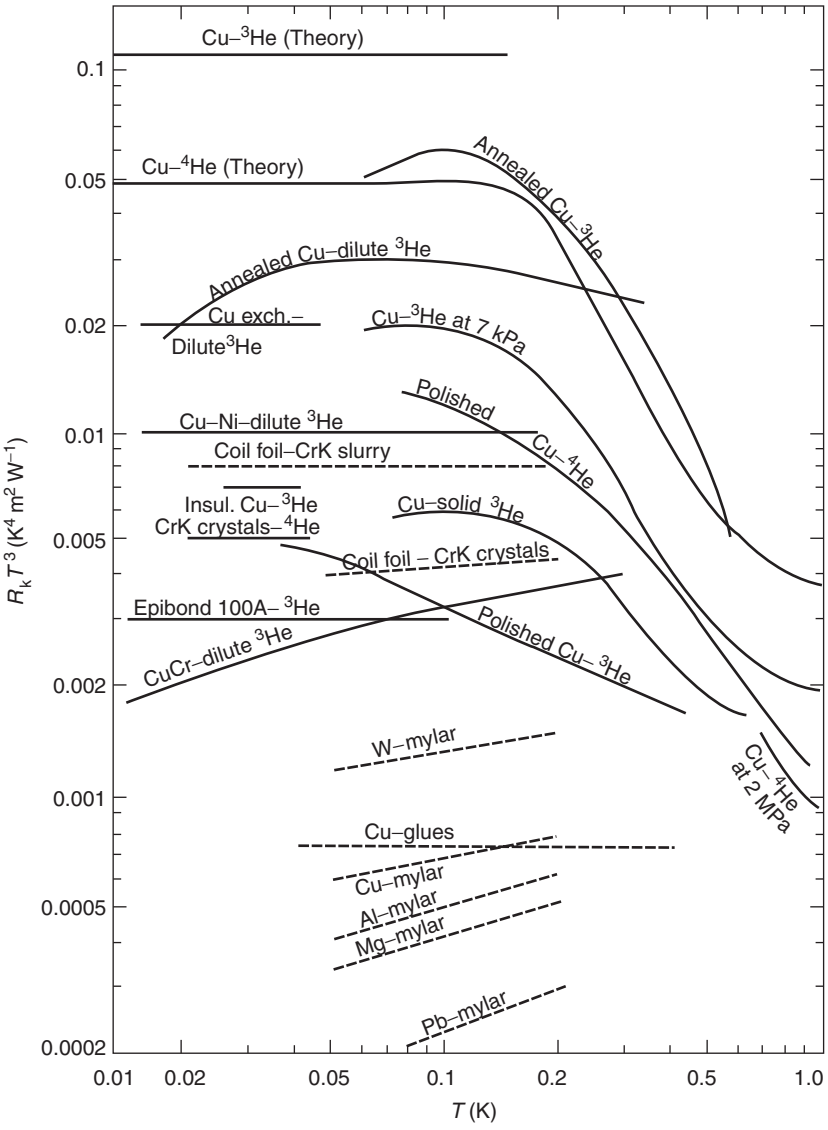


Figure 7 The Kapitza boundary thermal resistivities between liquid ^{3}He or ^{4}He and various solids. Some solid–solid boundary resistivities (dashed lines) have been included. (Lounasmaa OY (1974) *Experimental Principles and Methods Below 1 K*. London: Academic Press.)

boundary at an angle α_l (see **Figure 6**). Snell's law describes refraction of an acoustic wave at the boundary:

$$\sin \alpha_l / \sin \alpha_s = v_l / v_s \quad [6]$$

where the index s refers to the solid. The critical angle given by

$$\alpha_l^c = \arcsin v_l / v_s \quad [7]$$

defines critical cone for incident phonons entering into solid phase. Phonons striking the boundary at angles outside the critical cone are reflected back into the liquid. Below 1 K, the sound velocity in helium is $238\,\mathrm{m\,s^{-1}}$, while in solids it is typically 5×10^3 $\mathrm{m\,s^{-1}}$. It gives a critical angle less than 3° and critical cone solid angle of $2\pi \times 10^{-3}$. It means that the fraction of phonons which may enter the solid is 10^{-3} of phonons hitting the boundary. At thermal equilibrium, the same amount of energy must be transmitted from both sides of the boundary. It was calculated that Kapitza thermal boundary resistivity (resistance per unit of area) is proportional to T^3.

In **Figure 7**, the Kapitza resistivities are given for various solid–liquid and solid–solid boundaries. The Kapitza resistance might create serious problems in heat exchange between helium and solids as well as between different solids at temperatures below 50 mK.

See also: Low-Temperature Properties of Materials.

PACS: 44.10.+i; 44.20.+b; 66.60.+a; 66.70.+f

Further Reading

Cahill DG, *et al.* (2003) Nanoscale thermal transport. *Journal of Applied Physics* 93: 793–818.

Hunklinger S (1978) Anomalous properties of glasses at low temperatures. *Journal of Physique* 39: 1444–1449.

Nomenclature

c	specific heat per unit volume
e	elementary charge
k_B	Boltzmann constant
l	mean free path
L_0	Lorenz constant
q	heat
T	temperature
v	velocity
ΔE	energy gap
α	angle
σ	electrical conductivity
Θ	Debye temperature
κ	thermal conductivity
∇	gradient

Thermodynamic Properties, General

J M Honig, Purdue University, West Lafayette, IN, USA

Introduction

A general overview of thermodynamic principles is provided herein, with emphasis on topics that are useful in dealing with materials properties. Readers are referred to the "Further reading" section for an in-depth coverage of the subject matter presented here.

The Zeroth Law, Temperature, and Equations of State

The venerable field of thermodynamics rests on four basic laws that are discussed in turn.

The zeroth law asserts that "two bodies in equilibrium with a third are in equilibrium with each other." This seemingly obvious statement of transitive properties has important ramifications: consider systems 1 and 2 surrounded by walls that effectively isolate them from the rest of the universe, but allow them to be independently subject to mechanical deformations via pressures P_1 and P_2 (forces per unit area) that deform their volumes V_1 and V_2. While the systems are separated, these four quantities may be varied independently. However, when joined together and allowed to equilibrate, it is an experience of mankind that only three of the four variables may be independently altered. This is expressed mathematically by the relation $\beta_3\,(P_1, V_1, P_2, V_2) = 0$. Repeat the process for systems 1 and 3, leading to the expression $\beta_2\,(P_1, V_1, P_3, V_3) = 0$. Then by the zeroth law, $\beta_1\,(P_2, V_2, P_3, V_3) = 0$ on combining systems 2 and 3. As long as mutual equilibrium prevails after each combination, system 3 remains unaltered in its union with either system 1 or 2; one can then solve for P_3 in the functions β_2 and β_1 to express

$P_3 = \Phi_1\ (P_2, V_2, V_3) = \Phi_2\ (P_1, V_1, V_3)$, from which the following difference function is constructed: $\Phi_1\ (P_2, V_2, V_3) - \Phi_2\ (P_1, V_1, V_3) \equiv \lambda\ (P_1, V_1, P_2, V_2, V_3) = 0$. Unfortunately, a glaring inconsistency now emerges: the functional dependence of λ on V_3 is absent from the function $\beta_3 = 0$; also, it makes no sense to refer to λ in system 3 when combining systems 1 and 2. To resolve this impasse, let V_3 occur in Φ_1 and Φ_2 so as to cancel out when their difference is constructed. This may be achieved in most general terms by writing $\Phi_1 = f_2\ (P_2, V_2)\ h\ (V_3) + q\ (V_3)$ and $\Phi_2 = f_1\ (P_1, V_1)\ h\ (V_3) + q\ (V_3)$, where h and q are arbitrary functions of V_3. Substitution of the last two equations in the function λ then leads to the relation

$$f_1\ (P_1, V_1) = f_2\ (P_2, V_2) \qquad [1a]$$

Similarly, consistent with the zeroth law,

$$f_1\ (P_1, V_1) = f_3\ (P_3, V_3) \qquad [1b]$$

Reference is now made only to the variables appropriate to each separate system. Equations [1a] and [1b] thus characterize the equilibration process, and they permit one to consider system 1 as a reference standard that determines whether systems 2 and 3 are in equilibrium, according as system 1 is or is not changed when it comes successively in contact with systems 2 and 3. Clearly, the functional interrelation specified by $f_1\ (P_1, V_1)$ is of great significance; therefore, a short-hand notation symbol τ_1, termed the "empirical temperature function" is provided for this function. The relationship $\tau_i = f_i\ (P_i, V_i)$ is known as an "equation of state" for system i. One can thus characterize the empirical temperature of a given system through measurements of its pressure and/or volume. While it has been worked here with mechanical variables, other experimental measurements could have been used in a similar fashion to construct empirical temperatures. Many different methods for monitoring temperature have been employed; these are described in review articles on the subject, such as listed in the "Further reading" section. Only the use of a gas thermometer is briefly discussed here. Other commonly used techniques include measurements of electrical resistivity, thermocouple voltages, paramagnetic susceptibility, pyrometers, and thermal expansion; a description of these and other techniques would be beyond the confines of this article.

The labeling of τ_i as "temperature" obviously introduces a link between physical properties and human sensory perceptions of "hotness levels." The ordering of different hotness levels requires a quantification scheme via an empirical temperature scale; one may employ the equation of state of any suitable material as one indicator of hotness; an enormous multitude of other indicators has been applied for the same purpose (see the "Further reading" section). For present purposes, Boyle's law $f(P, V) = PV = \text{constant} \equiv \tau$ is adopted, which applies to an ideal gas, such as He at low pressure and at a constant temperature well above its critical point. The PV product can then be used as a measure of hotness. A useful quantification scheme is the so-called "Celsius" scale that assigns the values $\tau = 0°\text{C}$ (the original intent, but nowadays assigned a value 0.01°C) and $\tau = 100°\text{C}$ to water that is equilibrated, respectively, with ice and with steam at 1 bar. Let V, V_0, and V_{100} be the volume of He gas at a fixed, low pressure and at temperatures τ, 0°C and 100°C respectively; then τ is specified by

$$\tau = 100(V - V_0)/(V_{100} - V_0) = 100V/(V_{100} - V_0) - 100V_0/(V_{100} - V_0) \equiv T + T_0 \qquad [2]$$

A straight line plot of τ versus V based on the above two assigned points extrapolates to an intercept $T_0 \equiv -100V_0/(V_{100} - V_0)$ which has been found to have the value $-273.15°\text{C}$, where the volume V of the ideal gas would vanish, if the gas could be maintained at lowest temperatures. (Actually, the third law, discussed below, shows that a perfect gas cannot exist at very low temperatures. This, however, does not prevent introducing the ideal gas concept as an excellent approximation to properties of a real gas, and then extrapolating its characteristics to the ultralow temperature range, showing what its properties would be if the gas could be maintained in that range.) The extrapolation to $V = 0$ suggests that there exists a natural lower limit to the hotness levels, and that an absolute temperature scale may be set up by adding 237.15 to the Celsius scale. One thereby establishes the "thermodynamic temperature scale" $T\ (\text{K}) = \tau(°\text{C}) + 273.15$.

Force, Heat, Work, and the First Law of Thermodynamics

In preparation for the next step, let x_i represent a "deformation or thermodynamic coordinate" (such as volume, pressure, temperature, mole number, and the like) that characterizes the thermodynamic state of a given system, and let f_i represent any causative agent by which this particular state may be altered; this is taken to be the "thermodynamic force." Then, the element of work of type i performed "on a system" by its surroundings is specified by $f_i\,\mathrm{d}x_i$; if more than one force brings about a change in x_i, the various contributions must be summed. f_i must

remain sufficiently small so as to avoid turbulent or dissipative effects, otherwise, a whole host of extra-thermodynamic variables would be required to characterize the process. One thus deals with reversible processes that involve only infinitesimal displacements of a system from equilibrium. Also, the force itself may depend on the displacement and may or may not be collinear with the deformation coordinate. These situations are addressed by expressing the work performed on the system when changing its state from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ as $W_i = \int_{x_1}^{x_2} \boldsymbol{f}_i(\boldsymbol{x}_i) \cdot \mathrm{d}\boldsymbol{x}_i$, where the force is taken to be a vector and the dot vector product has been introduced. In general, such a line integral depends on the chosen path that takes the system from state $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$, whence this is also the case for the differential quantity $đW = \boldsymbol{f}_i(\boldsymbol{x}_i) \cdot \mathrm{d}\boldsymbol{x}_i$. To emphasize this point, the special symbol $đ$ has been used. It is obviously awkward to have to deal with path-dependent entities; hence, one searches for functions that depend only on the initial and final states of a given process; such entities are known as "functions of state." In the present example, a path-independent work performance may be generated if the force is representable as the gradient of a potential, so that $\boldsymbol{f} = -\nabla\varphi$, whence

$$-W_i = -\int_{x_1}^{x_2} \boldsymbol{f}_i(\boldsymbol{x}_i) \cdot \mathrm{d}\boldsymbol{x}_i = \int_{x_1}^{x_2} \nabla\phi \cdot \mathrm{d}\boldsymbol{x}_i = \phi(x_2) - \phi(x_1)$$

is indeed seen to depend only on the initial and final values of ϕ. The presence of the minus sign indicates that in going from a higher potential $\phi(\boldsymbol{x}_2)$ to a lower potential $\phi(\boldsymbol{x}_1)$, the system does work on the surroundings, the negative of work done by the surroundings on the system.

Mathematical formulations of commonly encountered types of work are briefly listed, but the list is not exhaustive: (1) mechanical pressure – volume work: $đW = -P\,\mathrm{d}V$; (2) enlargement of surface area A: $đW = -\sigma\,\mathrm{d}A$, where σ is the surface tension; (3) gravitational work involved in lifting mass $M\,\mathrm{d}n$ (M is the gram-molecular mass and n is the mole number) through a height z: $đW = gzM\,\mathrm{d}n$, where g is the gravitational constant; (4) electrostatic work (in Gaussian units): $đW = \int \mathrm{d}^3\boldsymbol{r}\boldsymbol{E}\cdot\mathrm{d}\boldsymbol{D}/4\pi = -\int_V \mathrm{d}^3\boldsymbol{r}\boldsymbol{P}\cdot\mathrm{d}\boldsymbol{E}_0$, where $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{P}$, $\boldsymbol{E}_0$ are, respectively, the electrostatic field in the presence of the sample, the electric displacement vector, the polarization vector, and the preexisting electrostatic field prior to sample insertion. The first integral must be taken over all space; the second extends over the volume of the sample, since $\boldsymbol{P}$ vanishes outside the confines of the specimen; (5) magnetic work (in Gaussian units): $đW = \int \mathrm{d}^3\boldsymbol{r}\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{B}/4\pi = \int_V \mathrm{d}^3\boldsymbol{r}\boldsymbol{H}_0\cdot\mathrm{d}\boldsymbol{M}$. Here, $\boldsymbol{H}$, $\boldsymbol{B}$, $\boldsymbol{H}_0$, and $\boldsymbol{M}$ are, respectively, the applied magnetic field in the presence of the sample, the magnetic induction, the applied magnetic field prior to insertion of the sample, and the magnetic moment vector. The above shows that work performance can always be determined from first principles; alternatively, it may be measured experimentally. This is of importance in what follows.

"The first law of thermodynamics" states in part that when work $W|_{\mathrm{a}}$ is performed adiabatically (that is, on a system impervious to changes in temperature of the surroundings) thereby changing its state from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$, then this change is independent of the type of work performed and independent of the path taken to effect this change. In these circumstances, one deals with a function of state, termed energy, here designated as U, that characterizes the $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ transformation via $\Delta U \equiv U_2 - U_1 = W|_{\mathrm{a}}$. If now, the same process is carried out under nonadiabatic conditions, the $\Delta U = W$ equivalence no longer holds, thus seemingly impairing the usefulness of the energy concept. However, in the nonadiabatic process, additional changes take place that need to be considered.

The first law of thermodynamics asserts that there exists a function of state U, termed the internal energy, such that, for any process whatsoever that takes a system from an initial state $\boldsymbol{x}_{\mathrm{n}}$ to a final state $\boldsymbol{x}_{\mathrm{f}}$, the difference $\Delta U \equiv U_{\mathrm{f}} - U_{\mathrm{n}}$ depends solely on the thermodynamic coordinates $\boldsymbol{x}_{\mathrm{n}}$ and $\boldsymbol{x}_{\mathrm{f}}$ and is independent of the path connecting these states. ΔU can be determined experimentally, or theoretically from $W|_{\mathrm{a}}$. In nonadiabatic changes between the same states, ΔU remains exactly the same, but now the work performance W is different: $\Delta U - W \neq 0$. It is then expedient to introduce a deficit function Q, termed heat transfer, that reinstates the equality in the form: $Q - (\Delta U - W) = 0$. One thus arrives at the following formulation for the first law of thermodynamics:

$$\Delta U = Q + W \quad \text{or} \quad \mathrm{d}U = đQ + đW \qquad [3]$$

where the second formulation applies to an infinitesimal stage of the process; d and $đ$ are used when one deals with infinitesimal processes that are not or are path dependent. Thus, Q and W are two different manifestations of energy in transit across the boundary separating the system from its surroundings. $Q > 0$ (< 0) increases (decreases) the energy of the system via an inflow (outflow) of heat; similarly, for the work transfer $W > 0$ (< 0). While Q, W, $đQ$, $đW$ generally are path dependent, their sum in eqn [3] is not. Thus, given a set of n thermodynamic variables $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$, the function of state U

has the differential form of an analytic function:

$$dU = \sum_{j=1}^{n} (\partial U/\partial x_j)\, dx_j \qquad [4]$$

Clearly, if processes occur totally within the system, there is no transfer of heat or work across the boundaries; then $\Delta U = 0$, which is an expression of the law of conservation of energy in an isolated system. This also disposes off the possibility of having a perpetual motion taking place.

Heat Transfer, Entropy, and the Second Law of Thermodynamics

The next law has been phrased in several different ways: the most common versions involve the statements enunciated by Kelvin and by Planck; these are derived below from more general principles. They are generally presented in elementary textbook discussions that build up to the second law. There is nothing wrong logically with this procedure. However, intellectually it is rather unsatisfactory to base a fundamental thermodynamic principle on statements concerning the operation of a heat engine.

A more general formulation was provided by Carathéodory: in every neighborhood of any state of an adiabatically isolated system, there exist other states that cannot be accessed from it by any process. This statement really provides the proper theoretical background to the second law, but considerable mathematical sophistication is needed to show how this statement leads to the concept of entropy (for a full discussion of this matter see the "Further reading" section). Here, one can simplify matters by noting that just as one connected an adiabatic performance of work to the establishment of a function of state, a transfer of heat can also be used to the same end. Effectively, the first part of the second law of thermodynamics is equivalent to the statement that a reversible infinitesimal transfer of heat to a system $đQ|_r$, though path dependent, is related by a scaling (or integrating) factor λ to a new differential function of state, the empirical entropy s, via $đs = đQ|_r/\lambda$.

The determination of λ is accomplished through a set of tight arguments by considering the process of combining two systems A and B into a compound unit C. Let the heats transferred to the surroundings in an infinitesimal process be designated by $đQ_a$ and $đQ_b$ for the component parts, and by $đQ_c$ for the compound system, $đQ_a + đQ_b = đQ_c$. Then, according to the second law,

$$ds_c = (\lambda_a/\lambda_c)ds_a + (\lambda_b/\lambda_c)ds_b \qquad [5]$$

Let the deformation coordinates of A and B be given by $x_1, x_2, \ldots, x_{n-1}$, t and $y_1, y_2, \ldots, y_{m-1}$, t respectively, where t is the common empirical temperature of the compound system. Mathematically, one can then solve for x_{n-1} in terms of $x_1, x_2, \ldots, x_{n-2}$, t, s_a and for y_{m-1} in terms of $y_1, y_2, \ldots, y_{m-2}$, t, s_b. It is now asserted that: (1) s_c must be independent of $x_1, x_2, \ldots, x_{n-2}$; $y_1, y_2, \ldots, y_{m-2}$; and t. If this were not so, eqn [5] would necessarily contain terms involving the differential quantities $dx_1, dx_2, \ldots, dy_{m-2}$, dt (see eqn [4]). (2) The ratios λ_a/λ_c and λ_b/λ_c cannot depend on these variables; otherwise, statement (1) would be contradicted. (3) λ_c cannot depend on $y_1, y_2, \ldots, y_{m-2}$; for, if it did, then λ_a in eqn [5] would have to depend on these same coordinates in such a manner as to cancel out from the ratio λ_a/λ_c, for consistency with statement (2) and (1). However, λ_a cannot possibly depend on the y-coordinates appropriate to system B. By the same reasoning, λ_c must be independent of $x_1, x_2, \ldots, x_{n-2}$. (4) Furthermore, λ_a and λ_b cannot depend on $x_1, x_2, \ldots, x_{n-2}$ and $y_1, y_2, \ldots, y_{m-2}$, respectively. For, according to (3) these variables are missing from λ_c, whence the ratios λ_a/λ_c and λ_b/λ_c must also be free of these variables, to remain consistent with (1), (2), and (3). However, this argument does not apply to t, which is common to A, B, and C. (5) As a consequence of (1)–(4), the functions λ_a, λ_b, and λ_c can involve at the most, the variables (s_a, t), (s_b, t), and (s_a, s_b, t), respectively. (6) The functional dependences mentioned in (5) must have the form $\lambda_a = \varphi_a(s_a)T(t)$, $\lambda_b = \varphi_b(s_b)T(t)$, and $\lambda_c = \varphi_c(s_a, s_b)T(t)$, in which $T(t)$ is a common though arbitrary function of the empirical temperature t; this function cancels from the ratios λ_a/λ_c and λ_b/λ_c in eqn [5], in consonance with requirement (2). In summary, it is found that

$$\begin{aligned} đQ_a &= T(t)\varphi_a(s_a)ds_a \equiv T(t)dS_a \\ đQ_b &= T(t)\varphi_b(s_b)ds_b \equiv T(t)dS_b \\ đQ_c &= T(t)\varphi_c(s_a, s_b)ds_c \equiv T(t)dS_c \end{aligned} \qquad [6]$$

where S in $dS_a \equiv \varphi_a(s_a)ds_a$, etc., is termed the metrical entropy. As described earlier, one can select any arbitrary function for T; T is thus taken to be identical with the thermodynamic temperature scale of eqn [2]. Further, equilibrium conditions have been assumed throughout, so that all processes must be executed reversibly. Thus, one is led to the result

$$đQ|_r = T\, dS \qquad [7]$$

which relates the state function S to the reversible transfer of heat across the boundaries of the system.

Now, the corollary of the second law is introduced as follows: For any process in a system in (adiabatic),

isolation, the entropy of the final state is never less than that of the intial state. Thus, $\mathrm{d}S|_{\mathrm{a}} \geqslant 0$; the equality holds at equilibrium, and the inequality applies to processes taking place totally within an isolated system; thus, the entropy increases in such processes and reaches a maximum when equilibrium is attained.

Cyclic Processes: The Clausius Inequality

The relation between entropy and heat transfer is clarified by considering a process in which a system is taken reversibly from an initial state n to a final state f, involving a heat exchange $đQ|_{\mathrm{r}}^{\mathrm{n-f}}$. This is to be compared to the heat exchange $đQ|_{\mathrm{i}}^{\mathrm{n-f}}$ accompanying an irreversible process connecting the same states. From the first law, it is found that $đQ|_{\mathrm{i}}^{\mathrm{n-f}} - đQ|_{\mathrm{r}}^{\mathrm{n-f}} = -đW|_{\mathrm{i}}^{\mathrm{n-f}} + đW|_{\mathrm{r}}^{\mathrm{n-f}}; \mathrm{d}U$, being a function of state, has been cancelled out. The above difference cannot vanish, since the two paths must necessarily differ. Thus, the algebraic sum must be either positive or negative. (1) Suppose $đQ|_{\mathrm{i}}^{\mathrm{n-f}} - đQ|_{\mathrm{f}}^{\mathrm{n\text{-}f}} \equiv đQ|_{\mathrm{i}}^{\mathrm{n-f}} + đQ|_{\mathrm{r}}^{\mathrm{f-n}} < 0$, so that $-đW|_{\mathrm{i}}^{\mathrm{n-f}} + đW|_{\mathrm{r}}^{\mathrm{n-f}} = -đW|_{\mathrm{i}}^{\mathrm{n-f}} - đW|_{\mathrm{r}}^{\mathrm{f-n}} < 0$; that is, in going around one cycle (n–f–n), work has been performed on the system and an equivalent amount of heat has been transferred by the system to the surroundings. No other changes have taken place. (2) On the other hand, if the algebraic sum $đQ|_{\mathrm{i}}^{\mathrm{n-f}} + đQ|_{\mathrm{r}}^{\mathrm{f-n}}$ is assumed to be positive, a similar reasoning shows that during the cycle, heat has been introduced into the system, and an equivalent amount of work is obtained. To choose between the alternatives, consider the n–f process for which $đQ|_{\mathrm{i}}^{\mathrm{n-f}} \equiv đQ|_{\mathrm{i}}$, $đQ|_{\mathrm{r}}^{\mathrm{n-f}} = T\,\mathrm{d}S$. Then, under alternative (1) $T\mathrm{d}S > đQ|_{\mathrm{i}}$ in all circumstances, including the case where the n–f process is performed adiabatically, in which case $T\,\mathrm{d}S|_{\mathrm{a}} > 0$. No contradictions have been encountered. By contrast, in alternative (2) $T\,\mathrm{d}S|_{\mathrm{a}} < 0$, a conclusion that is unacceptable. One obtains the important result: $đQ|_{\mathrm{i}} < T\,\mathrm{d}S = đQ_{\mathrm{r}}$. Thus, in a reversible n–f process, the entropy change is matched by a heat transfer $đQ|_{\mathrm{r}}/T$. If the same change is achieved irreversibly, the entropy change is the same but the heat transferred is smaller, the difference between $đQ|_{\mathrm{r}}$ and $đQ|_{\mathrm{i}}$ being made up by effects beyond experimental control. Also, for any process $đW|_{\mathrm{i}} < đW|_{\mathrm{r}}$; the maximum amount of work attainable involves a reversible process. Lastly, for finite changes, $S_{\mathrm{f}} - S_{\mathrm{n}} \geqslant \int_{\mathrm{n}}^{\mathrm{f}} đQ/T$; then, for a cyclic process where $S_{\mathrm{f}} = S_{\mathrm{n}}$, $\oint \mathrm{d}S = \oint đQ/T \leqslant 0$. This latter result, which is consistent with $đQ|_{\mathrm{i}} < đQ|_{\mathrm{r}}$, is known as the Clausius inequality. The above brings up the problem of defining the temperature during irreversible processes; this is generally handled in two ways. One is to assume that the system is anchored to a huge reservoir whose temperature remains essentially fixed, while assuming that the heat transfer is sufficiently rapid to have that temperature also prevail in the system. The second is to note that in irreversible processes, for which departures from equilibrium are "small" (the linear regime of irreversibility), the temperature remains locally defined but varies with position. For more drastic departures from equilibrium, the temperature concept becomes more problematic.

Efficiencies of Cyclic Engines: Kelvin and Planck Statements

One can now deal with the efficiency of any engine that operates in cycles while attached to a hot and a cold reservoir maintained at temperatures T_{h} and T_{c}, respectively. All changes in the universe then occur solely in the surroundings. Let the heat transferred per cycle at the cold and hot junctions be $Q_{\mathrm{c}} > 0$ and $Q_{\mathrm{h}} > 0$, respectively; for the engine to operate, one requires $Q_{\mathrm{h}} - Q_{\mathrm{c}} > 0$, so that less heat is rejected at the cold end than was transferred into the engine at the hot end. Correspondingly, $W = -(Q_{\mathrm{h}} - Q_{\mathrm{c}}) < 0$ represents the amount of work transferable to the surroundings. The efficiency η is defined by the ratio of work output/heat input. The Clausius inequality as applied to the present process reduces to the sum $\sum_{\mathrm{i}} Q_{\mathrm{i}}/T_{\mathrm{i}} = Q_{\mathrm{h}}/T_{\mathrm{h}} - Q_{\mathrm{c}}/T_{\mathrm{c}} \leqslant 0$; this may be converted to the inequality

$$-Q_{\mathrm{c}}/Q_{\mathrm{h}} \leqslant -T_{\mathrm{c}}/T_{\mathrm{h}}$$

Then,

$$\begin{aligned} \eta &= -W/Q_{\mathrm{h}} = 1 - Q_{\mathrm{c}}/Q_{\mathrm{h}} \leqslant 1 - T_{\mathrm{c}}/T_{\mathrm{h}} \\ &= (T_{\mathrm{h}} - T_{\mathrm{c}})/T_{\mathrm{h}} \equiv \eta_{\mathrm{Carnot}} \end{aligned} \quad [8]$$

The following are to be noted: (1) The generality of the approach; no mention is made of what types of cyclic engines are inserted between the two reservoirs. (2) The efficiency relates to the isothermal transfer of heat across the hot and cold junctions. (3) The efficiency for irreversible processes is less than that for reversible ones. (4) The efficiency is never unity since one cannot operate at the unattainable limits (see below) of $T = 0$ or $T \to \infty$. Thus, the efficiency never exceeds the Carnot efficiency η_{Carnot}. (5) η depends solely on the temperature of the "boiler" and "condenser" and, the bigger the temperature difference between them, the larger it is.

Statement (4) leads to Kelvin's statement of the second law: it is impossible to devise an engine operating in cycles that produces no effect other than to extract heat from a reservoir and to perform an equal amount of work. Also, one may infer the Clausius

statement of the second law: it is impossible to devise a machine operating in cycles that transfers heat from a colder to a hotter body without producing other changes in the universe. If this were false, one could transfer heat without loss from the cold to the hot reservoir, so that $-Q_h/T_h + Q_c/T_c \leqslant 0$ with $Q_c = Q_h$, which would lead to the contradiction that $T_h/T_c \leqslant 1$.

Deficit Functions and Their Use in Setting up Functions of State

The following discussion is much simplified by converting the Clausius inequality to an equality through the use of a deficit function $đ\varepsilon > 0$, such that $đQ|_r - đ\varepsilon = đQ|_i$; similarly, $đW|_r + đ\varepsilon = đW|_i$. In a closed universe consisting of system (energy U) and surroundings, (energy U_0) the first law leads to the expression $dU + dU_0 = 0 = đQ|_i + đW|_i + dU_0$, so that

$$T\,dS - đ\varepsilon + đW|_i - dU = T\,dS + đW|_r - dU = 0 \qquad [9]$$

One imposes small deviations from the equilibrium, under a variety of constraints, and subsequently allows the system to relax to its original state. Such small deviations are symbolized by δ. Several cases are considered here:

(1) The system is totally isolated; then $\delta Q = \delta W = \delta U = 0$, and eqn [9] reduces to $T\,\delta S = \delta\varepsilon \equiv T\,\delta\theta \geqslant 0$. Here, $\delta\theta = \delta\varepsilon/T$ has been introduced as the entropy associated with the execution of irreversible processes that are not subject to external control. It is seen immediately that in an irreversible process taking place in an isolated system, the entropy can only increase, a point made earlier. When processes cease and equilibrium prevails, the entropy has achieved a maximum value.

(2) If one allows isoenergetic processes to occur, where $\delta U = 0$ but $\delta Q = \delta W \neq 0$, then eqn [9] reads as $T\,\delta S = \delta\varepsilon + \delta Q|_i = \delta Q|_r$, which merely reinforces the fact that changes in entropy are tracked by reversible heat transfers. Also, if irreversible processes occur, the heat transfer is smaller; the entropy generated by processes outside experimental control is transferred as energy $\delta\varepsilon$.

(3) In isothermal processes for which $\delta T = 0$, eqn [9] specializes to the form

$$\delta(U - TS) \equiv \delta A = \delta W|_i - \delta\varepsilon = \delta W|_r \qquad [10]$$

where a new function of state, the Helmholtz free energy $A \equiv U - TS$ has been generated. This is tracked either by the performance of reversible work or by work over which the experimenter has control plus a component not under his control. If no work is performed, $\delta A = -\delta\varepsilon \leqslant 0$: any spontaneous process under the specified conditions reduces the Helmholtz free energy. When the processes have run their course and equilibrium prevails, A is at a minimum.

(4) One next considers the isothermal–isobaric conditions for which $\delta T = \delta P = 0$. It is now expedient to write $\delta\varepsilon = \delta\varepsilon_m + \delta\varepsilon_t$, where the two terms refer to mechanical (P–V) and nonmechanical types of work; it may be shown that $\delta\varepsilon_m \geqslant 0$ and $\delta\varepsilon_t \geqslant 0$ separately. Equation [9] now becomes

$$\delta(U - TS + PV) \equiv \delta G = \delta W_t|_i - \delta\varepsilon_t = \delta W_t|_r \qquad [11]$$

wherein one introduces a new function of state, the Gibbs free energy $G \equiv U - TS + PV$. This quantity is tracked by reversible work other than mechanical ones; in irreversible processes, the work is augmented by energy transfers not subject to control. In the absence of work, $\delta G = -\delta\varepsilon_t \leqslant 0$, showing that under the stated conditions, the Gibbs free energy diminishes in spontaneous processes until it achieves a minimum at equilibrium.

(5) Isentropic phenomena: here, S is fixed, so that from eqn [9]

$$\delta U = -T\,\delta\theta + \delta W|_i = \delta W|_r \qquad [12]$$

which merely repeats the first law, as specialized to adiabatic conditions. In the absence of work, $\delta U = -T\delta\theta \leqslant 0$, so that under the stated conditions, the energy is reduced and tends to a minimum at equilibrium.

(6) For isentropic–isobaric conditions (P and S fixed), eqn [9] specializes to

$$\delta(U + PV) \equiv \delta H = \delta W_t|_i - \delta\varepsilon = \delta W_t|_r \qquad [13]$$

where yet another function of state, the enthalpy $H \equiv U + PV$, tracked by nonmechanical work, generally augmented by irreversibility effects, has been constructed. When no work is performed under the stated conditions, $\delta H = -T\,\delta\theta \leqslant 0$, so that once more H diminishes in irreversible phenomena and achieves a minimum at equilibrium.

Among the most frequently used functions of state is the chemical potential, derived from eqn [11] as $\mu = (\partial G/\partial n_i)_{T,P,n_{j\neq i}}$, where n_i represents the mole number of species i in the system. The utility of this function arises in part because it involves T and P as independent variables; these are precisely the quantities that are readily subject to experimental control. The second useful feature hinges on the fact that at equilibrium, μ is a minimum. Different phases are associated with different chemical potentials that

vary differently with T and P. The phase among competing phases that prevails under a given set of conditions is the one with the lowest μ. As conditions change, the relative values of μ for the different possible phases are also altered, and a transition from one phase to another takes place whenever the chemical potential of a particular phase falls below that of a previously stable phase. Two phases coexist under conditions where their chemical potentials are the same. This very general subject area has vast ramifications that cannot be explored here.

Standard Formulation for Differentials of State Functions: Maxwell Relations

The above functions of state are summarized in **Table 1**. They lead to the various standard differential forms displayed in part A; note that all work other than mechanical has been excluded. If other types of work are to be included, one must adjoin to these equations, the appropriate work differentials (refer to those listed above).

The process of generating the differentials dH, dA, and dG from dU is called a Legendre transformation. These differential functions of state are very useful in determining the thermodynamic properties of systems. Taking first derivatives leads to the results shown in part B; by taking cross derivatives in either order, one obtains the so-called Maxwell relations shown in part C. Once the equation of state $P = P(T, V)$ or $V = V(T, P)$ is known, one can integrate [I.10] and [I.12] to find S as a function of T and V or T and P. Alternatively, one may employ eqn [I.7] if A or G are known, for example, from statistical mechanics.

Other deductions are found by substituting $S = S(T, V)$ or $S = S(T, P)$ into $U(S,V)$ or $H = H(S, P)$ to obtain $U = U(S(V, T), V) = U(T, V)$, and $H = H(S\ (P, T), P) = H(T, P)$, so that

$$\begin{aligned} \mathrm{d}U &= T(\partial S/\partial T)_V \mathrm{d}T + \left[T(\partial S/\partial V)_T - P\right]\mathrm{d}V \\ &= (\partial U/\partial T)_V \mathrm{d}T + (\partial U/\partial V)_T \mathrm{d}V \end{aligned} \quad [14a]$$

$$\begin{aligned} \mathrm{d}H &= T(\partial S/\partial T)_P \mathrm{d}T + \left[T(\partial S/\partial P)_T + V\right]\mathrm{d}P \\ &= (\partial H/\partial T)_P \mathrm{d}T + (\partial H/\partial P)_T \mathrm{d}P \end{aligned} \quad [14b]$$

Comparing the coefficients leads to the results

$$\begin{aligned} (\partial U/\partial T)_V &= T(\partial S/\partial T)_V \equiv C_V \\ (\partial H/\partial T)_P &= T(\partial S/\partial T)_P \equiv C_P \end{aligned} \quad [15a]$$

$$\begin{aligned} (\partial U/\partial V)_T &= T(\partial S/\partial V)_T - P \\ &= T(\partial P/\partial T)_V - P(V, T) \end{aligned} \quad [15b]$$

$$\begin{aligned} (\partial H/\partial P)_T &= T(\partial S/\partial P)_T + V \\ &= -T(\partial V/\partial T)_P + V(P, T) \end{aligned} \quad [15c]$$

The appropriate Maxwell relations have been introduced on the right-hand side of eqns [15b] and [15c]. The symbols C_P and C_V represent the heat capacities at constant pressure and volume; these are readily measurable as a function of temperature. This permits the determination of U and H as a function of T, with either V or P as a parameter. Equations [15b] and [15c] are known as caloric equations of state; insertion of the equations of state $P(T,V)$ or $V(T,P)$ into these relations, followed by integration, enables one to determine $U = U(T, V)$ and $H = H(T, P)$. Also, from [15a], after measurements of C_V or C_P as a function of T, followed by integration, one can determine $S = S(T, V)$ or $S = S(T, P)$. In a series of lengthy but elementary steps listed in all textbooks, one finds that $C_P - C_V = \alpha^2 VT/\beta$, with $\alpha \equiv V^{-1}(\partial V/\partial T)_P$ representing the isobaric expansion coefficient, and $\beta \equiv -V^{-1}(\partial V/\partial P)_T$, the

Table 1 Standard formulations for thermodynamic functions of state and resulting relations

A. Legendre transformations

From the definitions for $H = U + PV$, $A = U - TS$, $G = U - TS + PV$, and the fundamental relation for dU, one obtains

$\mathrm{d}U = T\ \mathrm{d}S - P\ \mathrm{d}V \qquad \mathrm{d}A = -S\ \mathrm{d}T - P\ \mathrm{d}V$	[I.1, I.3]
$\mathrm{d}H = T\ \mathrm{d}S + V\ \mathrm{d}P \qquad \mathrm{d}G = -S\ \mathrm{d}T + V\ \mathrm{d}P$	[I.2, I.4]

If work other than mechanical is involved, those terms must also be included.

B. First derivatives

$T = (\partial U/\partial S)_V = (\partial H/\partial S)_P \qquad V = (\partial H/\partial P)_S = (\partial G/\partial P)_T$	[I.5, I.6]
$S = -(\partial A/\partial T)_V = -(\partial G/\partial T)_P \qquad P = -(\partial U/\partial V)_S = -(\partial A/\partial V)_T$	[I.7, I.8]

C. Cross-derivatives, Maxwell relations

$(\partial T/\partial V)_S = -(\partial P/\partial S)_V \qquad (\partial S/\partial P)_T = -(\partial V/\partial T)_P$	[I.9, I.10]
$(\partial T/\partial P)_S = (\partial V/\partial S)_P \qquad (\partial S/\partial V)_T = (\partial P/\partial T)_V$	[I.11, I.12]

Equation [I.7] is particularly useful in determining the entropy when the dependence of A or G on T is known from other sources. Equations [I.10] and [I.12] permit substitution for entropy derivatives in terms of information derived from the equations of state.

isothermal compressibility of the material. This difference relation is of importance because it is easier to measure C_P experimentally, whereas the quantity useful in theoretical analysis is C_V. Lastly, insertion of eqn [I.7] into $A = U - TS$ leads to $A = U + T(\partial A/\partial T)_V$; similarly, $G = H + T(\partial G/\partial T)_P$. These are known as "Gibbs–Helmholtz equations," which may be converted into the partial differential forms involving $(\partial A/\partial T)_V = A/T - U/T$ and $(\partial G/\partial T)_P = G/T - H/T$, whose solution, after insertion for U or H yields A or G.

The Third Law of Thermodynamics: Unattainability of Zero Temperature

The "third law of thermodynamics" deals with events as $T \to 0$, where $đQ/T$ might diverge. One first demonstrates that zero temperature is unattainable. Consider the general case where z represents a deformation coordinate and Z is the conjugate variable, such that the first law assumes the form $\mathrm{d}S = T^{-1}[\mathrm{d}U + Z\,\mathrm{d}z] = T^{-1}\{(\partial U/\partial T)\,\mathrm{d}T + [Z + (\partial U/\partial z)]\,\mathrm{d}z\}$. Below, one needs the result of the cross differentiation of this expression with respect to T and z, namely $Z + (\partial U/\partial z) = T(\partial Z/\partial T)$. Special interest now attaches to the case $\mathrm{d}S = 0$, since only under adiabatic conditions, one can hope to reach the lowest possible temperature. Imposing this requirement leads to

$$\mathrm{d}T = \{[Z + (\partial U/\partial z)]/(\partial U/\partial T)\}\,\mathrm{d}z \qquad [16]$$

This relation shows that any change in deformation coordinates under isentropic (adiabatic) conditions necessarily changes the temperature of the system. Insertion of the cross differentiation result leads to the requirement

$$\begin{aligned}\mathrm{d}T &= -T[(\partial Z/\partial T)_z/(\partial U/\partial T)_z]\,\mathrm{d}z \\ &= -[T(\partial Z/\partial T)_z/C_z]\,\mathrm{d}z \\ &= -(T/C_z)(\partial S/\partial z)_T\,\mathrm{d}z, \end{aligned} \qquad [17]$$

where eqn [I.12] with the correspondences P, Z, and V; z was introduced on the right-hand side. Whether $T = 0$ can be achieved or not, thus, hinges on the properties of $(\partial Z/\partial T)_z$ and of (T/C_z) at exceedingly low temperatures. In the experience of mankind, both quantities approach zero as $T \to 0$, but $(\partial Z/\partial T)_z$ does so more rapidly than does C_z/T. Thus, the limit $T = 0$ is unattainable. Moreover, $(\partial S/\partial z)_T \to 0$ in this limit, which is an important statement that becomes part of the third law.

Based on these findings, one can now state the third law of thermodynamics which asserts that as $T \to 0$, the entropy of any system tends toward a least value when the system is in its lowest energy state and, as the thermodynamic coordinates are altered, approaches this value with zero slope. The wording implies that the system must actually have reached its equilibrium state, hence, its lowest energy, on cooling to the lowest possible temperature. Contrary to statements in some textbooks, it is not claimed that $S \to 0$ as $T \to 0$. Liquid mercury under ambient pressure, disordered solid solutions or amalgams, random elemental isotopic abundances, are common counter examples. However, if in any given process these types of ground states are left undisturbed, it is effectively permissible to set $S = 0$ at $T = 0$, much as the gravitational energy of a quiescent system at sea level is assigned an energy zero.

Thermodynamics of Anisotropic Media

Lastly, the deformation of anisotropic solids is considered. Let $\boldsymbol{s}(\boldsymbol{r})$ and $\boldsymbol{s}(\boldsymbol{r} + \mathrm{d}\boldsymbol{r})$ be the displacement vectors that move element A, originally at $\boldsymbol{r}$, to $\boldsymbol{r} + \boldsymbol{s}(\boldsymbol{r})$, and neighboring element B, originally at $\boldsymbol{r} + \mathrm{d}\boldsymbol{r}$ to new locations. For small displacements, $\boldsymbol{s}(\boldsymbol{r} + \mathrm{d}\boldsymbol{r}) \approx \boldsymbol{s}(\boldsymbol{r}) + \mathrm{d}\boldsymbol{r} \cdot \boldsymbol{\nabla}\boldsymbol{s}(\boldsymbol{r})$. Here, $\boldsymbol{\nabla}\boldsymbol{s}(\boldsymbol{r})$ is a matrix (a tensor) with columns $\partial s_x(\boldsymbol{r})/\partial x$, $\partial s_x(\boldsymbol{r})/\partial y$, and $\partial s_x(\boldsymbol{r})/\partial z$, and similarly with $s_y(\boldsymbol{r})$ and $s_z(\boldsymbol{r})$. Now, any matrix with entries M_{ij} may be rewritten as $M_{ij} = (1/2)(M_{ij} + M_{ji}) + (1/2)(M_{ij} - M_{ji})$, involving a symmetric ($\boldsymbol{e}_\mathrm{s}$) and an antisymmetric component, $\boldsymbol{e}_\mathrm{a}$. The matrix for $\boldsymbol{\nabla}\boldsymbol{s}(\boldsymbol{r}) = \boldsymbol{e}_\mathrm{s} + \boldsymbol{e}_\mathrm{a}$ is shown for $\boldsymbol{e}_\mathrm{s}$ in **Table 2**, with entries $e_{ij} = (1/2)[\partial \boldsymbol{s}_i/\partial \boldsymbol{r}_j + \partial \boldsymbol{s}_j/\partial \boldsymbol{r}_i]$. For the $\boldsymbol{e}_\mathrm{a}$ entries, the "+" sign is replaced by "−," so that the diagonal elements of $\boldsymbol{e}_\mathrm{a}$ vanish.

Consider now a dilatation of a solid in which element A at $\boldsymbol{r}$ is displaced by $\boldsymbol{s}(\boldsymbol{r})$ to A', and a neighboring element B at $\boldsymbol{r} + \mathrm{d}\boldsymbol{r}$ is displaced to B' by $\boldsymbol{s}(\boldsymbol{r} + \mathrm{d}\boldsymbol{r})$. The A'–B' separation distance is then $\mathrm{d}\boldsymbol{r}' = \mathrm{d}\boldsymbol{r} + [\boldsymbol{s}(\boldsymbol{r} + \mathrm{d}\boldsymbol{r}) - \boldsymbol{s}(\boldsymbol{r})] \approx \mathrm{d}\boldsymbol{r} + \mathrm{d}\boldsymbol{r} \cdot \boldsymbol{\nabla}\boldsymbol{s}(\boldsymbol{r}) = \mathrm{d}\boldsymbol{r} + \mathrm{d}\boldsymbol{r} \cdot (\boldsymbol{e}_\mathrm{s} + \boldsymbol{e}_\mathrm{a})$. Note that $\boldsymbol{e}_\mathrm{a}$ represents a pure rotation; on rotating the coordinate system by $-\boldsymbol{e}_\mathrm{a}$, one obtains $\mathrm{d}\boldsymbol{r}' = \mathrm{d}\boldsymbol{r} + \mathrm{d}\boldsymbol{r} \cdot \boldsymbol{e}_\mathrm{s}$ in the rotated system. Here, $\mathrm{d}\boldsymbol{r}$ is the row vector $[\mathrm{d}x, \mathrm{d}y, \mathrm{d}z]$ and the $\boldsymbol{e}_\mathrm{s}$ matrix is shown in **Table 2**.

A linear dilatation along x is then specified by $\mathrm{d}x' = (1 + e_{xx})\mathrm{d}x + e_{yx}\mathrm{d}y + e_{zx}\mathrm{d}z$, with similar expressions for $\mathrm{d}y'$ and $\mathrm{d}z'$. Thus, the diagonal element e_{xx} of **Table 2** indicates the fractional lengthening of the solid along the x-axis, and similarly, for y and z. Angular dilatations, on the other hand, involve a displacement $\mathrm{d}l$ whereby $\hat{\boldsymbol{i}}\,\mathrm{d}l$ originally perpendicular to $\hat{\boldsymbol{j}}\,\mathrm{d}l$ ($\hat{\boldsymbol{i}}$ and $\hat{\boldsymbol{j}}$ are orthogonal unit vectors along x and y) now change into $A_x \equiv \hat{\boldsymbol{i}}\,\mathrm{d}l + \hat{\boldsymbol{i}}\,\mathrm{d}l \cdot \boldsymbol{e}_\mathrm{s} = \hat{\boldsymbol{i}}(1 + e_{xx})\,\mathrm{d}l + \hat{\boldsymbol{j}}e_{xy}\,\mathrm{d}l + \hat{\boldsymbol{k}}e_{xz}\,\mathrm{d}l$, and $A_y \equiv \hat{\boldsymbol{j}}\,\mathrm{d}l + \hat{\boldsymbol{j}}\,\mathrm{d}l \cdot \boldsymbol{e}_\mathrm{s} = \hat{\boldsymbol{i}}e_{yx}\,\mathrm{d}l + \hat{\boldsymbol{j}}(1 + e_{yy})\,\mathrm{d}l + \hat{\boldsymbol{k}}e_{yz}\,\mathrm{d}l$. The angle between

Table 2 Matrix entries for deformation of anisotropic materials

A. Entries to the symmetric strain tensor

$$\boldsymbol{e} = \begin{vmatrix} e_{xx}\, e_{xy}\, e_{xz} \\ e_{yx}\, e_{yy}\, e_{yz} \\ e_{zx}\, e_{zy}\, e_{zz} \end{vmatrix} \qquad \begin{matrix} e_{xx} = (\partial s_x/\partial x) & e_{xy} = (1/2)[(\partial s_x/\partial y) + (\partial s_y/\partial x)] = e_{yx} \\ e_{yy} = (\partial s_y/\partial y) & e_{xz} = (1/2)[(\partial s_x/\partial z) + (\partial s_z/\partial x)] = e_{zx} \\ e_{zz} = (\partial s_z/\partial z) & e_{yz} = (1/2)[(\partial s_z/\partial y) + (\partial s_y/\partial z)] = e_{zy} \end{matrix}$$

B. Conventional notation for strain components

$e_1 \equiv e_{xx}$, $e_2 \equiv e_{yy}$, $e_3 \equiv e_{zz}$
$e_4 \equiv 2e_{xy} = 2e_{yx}$, $e_5 \equiv 2e_{xz} = 2e_{zx}$, $e_6 \equiv 2e_{yz} = 2e_{zy}$

C. Conventional notation for stress components

$\sigma_1 \equiv \sigma_{xx}$ pressure along x $\quad \sigma_4 \equiv \sigma_{xy} = \sigma_{yx}$ pressure $\perp x$ along y or vice versa
$\sigma_2 \equiv \sigma_{yy}$ pressure along y $\quad \sigma_5 \equiv \sigma_{xz} = \sigma_{zx}$ pressure $\perp x$ along z or vice versa
$\sigma_3 \equiv \sigma_{zz}$ pressure along z $\quad \sigma_6 \equiv \sigma_{yz} = \sigma_{zy}$ pressure $\perp y$ along z or vice versa

A_x and A_y in the x–y-plane is specified by

$$\cos\,\phi_{xy} = \frac{(\hat{\boldsymbol{i}}\,\mathrm{d}l + \hat{\boldsymbol{i}}\,\mathrm{d}l \cdot \boldsymbol{e}_s) \cdot (\hat{\boldsymbol{j}}\,\mathrm{d}l + \hat{\boldsymbol{j}}\,\mathrm{d}l \cdot \boldsymbol{e}_s)}{|\hat{\boldsymbol{i}}\,\mathrm{d}l + \hat{\boldsymbol{i}}\,\mathrm{d}l \cdot \boldsymbol{e}_s| \times |\hat{\boldsymbol{j}}\,\mathrm{d}l + \hat{\boldsymbol{j}}\,\mathrm{d}l \cdot \boldsymbol{e}_s|} \qquad [18]$$

When the appropriate entries for $\boldsymbol{e}_s$ are introduced and an expansion is carried out to first order in e_{ij}, one obtains (after a large number of elementary steps) $\cos\,\phi_{xy} \approx 2e_{xy}$. Now, introduce the deviation from the right angle by $\Omega_{xy} = \pi/2 - \phi_{xy}$, so that $\cos\,\phi_{xy} \approx \Omega_{xy} \approx 2e_{xy}$. Similarly, $\Omega_{xz} \approx 2e_{xz}$, $\Omega_{yz} \approx 2e_{yz}$.

It is conventional to introduce a more compact notation for the strain components, as shown in part B of **Table 2**; also, one sets up six components of a stress tensor $\boldsymbol{\sigma}$ that induces the strain. Here, $\sigma_i = [\partial(U/V)/\partial e_i]_{S,ej\neq i}$; these are listed in part C of **Table 2**. The element of work is then given by $V\sum_i \sigma_i\,\mathrm{d}e_i$.

With these definitions, the above concepts are embedded in thermodynamics by writing the differential of the energy (of the strained relative to the unstrained body) as

$$\mathrm{d}U = T\,\mathrm{d}S + V\sum_{i=1}^{6} \sigma_i\,\mathrm{d}e_i \qquad [19]$$

One should note the positive sign ahead of the second term: work is done by the system as the strain is diminished. Next, the enthalpy, by $H = U - \sum_{i=1}^{6} \sigma_i (Ve_i)$ is introduced. The enthalpy differential is then given by

$$\begin{aligned} \mathrm{d}H &= \mathrm{d}U - V\sum_{i=1}^{6} e_i\,\mathrm{d}\sigma_i - V\sum_{i=1}^{6} \sigma_i\,\mathrm{d}e_i \\ &= T\,\mathrm{d}S - V\sum_{i=1}^{6} e_i\,\mathrm{d}\sigma_i \end{aligned} \qquad [20]$$

From [19] and [20], Maxwell relations may be constructed in the usual manner; these are listed in **Table 3**: integration then specifies the entropy of the solid in its dependence on σ_i or on e_i under different conditions (for details, see the "Further reading" section).

Table 3 Maxwell relations for anisotropic solids; constant V

$$\frac{1}{V}\left(\frac{\partial S}{\partial e_i}\right)_T = -\left(\frac{\partial \sigma_i}{\partial T}\right)_{e_i} \qquad \frac{1}{V}\left(\frac{\partial S}{\partial \sigma_i}\right)_T = \left(\frac{\partial e_i}{\partial T}\right)_{\sigma_i} \qquad \text{[III.1, III.2]}$$

$$\frac{1}{V}\left(\frac{\partial S}{\partial \sigma_i}\right)_{e_i} = \left(\frac{\partial e_i}{\partial T}\right)_S \qquad \frac{1}{V}\left(\frac{\partial S}{\partial e_i}\right)_{\sigma_i} = -\left(\frac{\partial \sigma_i}{\partial T}\right)_S \qquad \text{[III.3, III.4]}$$

Moreover, from the definition of the stress tensor, it follows that

$$(\partial\sigma_i/\partial e_j)_{S,e_{j\neq i}} \equiv c_{ij}|_S = (\partial\sigma_j/\partial e_i)_{S,e_{j\neq i}} \equiv c_{ji}|_S \qquad [21a]$$

$$(\partial\sigma_i/\partial e_j)_{T,e_{j\neq i}} \equiv c_{ij}|_T = (\partial\sigma_j/\partial e_i)_{T,e_{j\neq i}} \equiv c_{ji}|_T \qquad [21b]$$

which are known as isentropic and isothermal elastic stiffness coefficients. The inverse $(\partial\sigma_i/\partial e_j)_{S,\sigma_{j\neq i}}$ or $(\partial\sigma_i/\partial e_j)_{T,\sigma_{j\neq i}}$ represent isentropic and isothermal compliance coefficients. There also exist six thermal strain coefficients $\alpha_i \equiv (\partial e_i/\partial T)_{\sigma j}$ and six thermal stress coefficients $\delta_i \equiv (\partial\sigma_i/\partial T)_{ej}$ that deal with the temperature variations of the respective coefficients.

This completes the brief survey of commonly used coefficients that characterize the deformation of anisotropic elastic solids.

Application of Thermodynamic Concepts

The basic concepts developed above find many applications in condensed matter physics. Several important uses are briefly listed. Prominent among these is the determination of heat capacities of materials as a function of temperature, from which the energy, enthalpy, and entropy may be found by integration of eqn [15a].

An entire subdiscipline deals with phase equilibria, phase diagrams, phase transitions, and related aspects.

All these topics are ultimately based on the properties of the chemical potential, briefly alluded to in an earlier section of the present article. Phase diagrams obtained from empirical measurements are of cardinal importance in the entire manufacturing process and industrial output of chemicals in developed countries.

Another subdiscipline deals with the properties of matter, energy, and entropy flow, as determined by irreversible thermodynamics.

Thermal properties of materials represent another area of great contemporary interest. Here, the properties of anisotropic media discussed above, represent the proper theoretical background for a rationalization of experimental results.

Results based on the third law of thermodynamics provide the basic theoretical framework for interpreting a large number of studies dealing with properties of materials at ultralow temperatures. The operation of cyclic engines, described briefly above, is another area where thermodynamics plays an important role. Lastly, by considering the various different types of work alluded to in the introduction to the first law, one can extend the above thermodynamic principles to new areas, such as electromagnetic properties, surface adsorption phenomena, curved surfaces, radiation, processes in gravitational fields, and the like. Their detailed discussion must be left to specialized monographs. They form the foundation for explaining many physical characteristics of condensed matter, such as magnetic and electrical properties, the interaction of condensed phases with electromagnetic radiation, and interfacial phenomena.

See also: Diffusionless Transformations; Incommensurate Phases; Irreversible Thermodynamics and Basic Transport Theory in Solids; Mechanical Properties: Elastic Behavior; Phase Transformation; Phase Transformations, Mathematical Aspects of; Specific Heat; Thermodynamic Properties, General.

PACS: 05.70. – a; 05.70.Ce; 65.40.Ba; 65.40.Gr; 82.60. – s

Further Reading

Buchdahl HA (1966) *The Concepts of Classical Thermodynamics*. Cambridge: Cambridge University Press.

Callen HB (1985) *Thermodynamics and an Introduction to Thermostatistics*. New York: Wiley.

Guggenheim EA (1977) *Thermodynamics, an Advanced Treatise for Chemists and Physicists*. Amsterdam: North-Holland.

Honig JM (1999) *Thermodynamics*, 2nd edn. San Diego, CA: Academic Press.

Kondepudi D and Prigogine I (1998) *Modern Thermodynamics; from Heat Engines to Dissipative Structures*. Chichester, UK: Wiley.

Penner SS (1968) *Thermodynamics for Scientists and Engineers*. Reading, MA: Addison-Wesley.

Pippard AB (1965) *Elements of Classical Thermodynamics*. London: Cambridge University Press.

Redlich O (1976) *Thermodynamics: Fundamentals, Applications*. Amsterdam: Elsevier.

Reiss H (1965) *Methods of Thermodynamics*. New York: Blaisdell.

Rock PA (1983) *Chemical Thermodynamics*. Mill Valley, CA: University Science Books.

Temperature, its Measurement and Control in Science and Industry (1941–1992), Multivolume, Multiauthor Series. New York: American Institute of Physics.

Thermoelectric and Energy Conversion Devices

J W Sharp, Marlow Industries, Inc., Dallas, TX, USA

Introduction

There are a number of ways of converting thermal to electrical energy and vice versa, and more than one of these have sometimes been referred to as "thermoelectric." Here, thermoelectric refers only to direct conversion between thermal and electrical energy by the Peltier and Seebeck effects. This article discusses the basic engineering and materials science aspects of thermoelectric devices, without proceeding to system issues such as heat exchange and feedback loops for precise temperature control. The discussion is limited to cooling and electrical power generation, setting aside aspects particular to thermoelectric sensors.

This article is composed of two main sections. The first section covers the fundamentals, analysis, and construction of thermoelectric devices. The second section focuses on thermoelectric materials, including those that are established, some that are being developed, and nanostructured formats that are the subject of research.

Thermoelectric Devices

Fundamentals

The fundamentals of a thermoelectric device are displayed in **Figure 1**. Two blocks of semiconductor, one *p*-type (*P*) and the other *n*-type (*N*), are joined by a

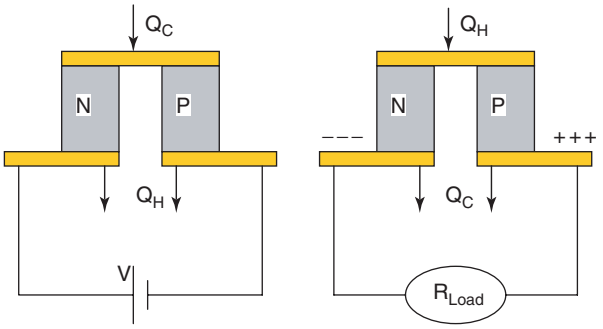


Figure 1 Schematic illustrations of thermoelectric couples in use for cooling (left) and power generation (right). In each case, Q_H and Q_C differ by the work done, either by the voltage source or on the load. For cooling, the heat sink is at the hotter temperature (T_H), while the sink is at the lower temperature (T_C) for power generation.

metal interconnect at one end and connected in series to a voltage source or load at the other end. The two blocks share a common temperature at each end, and thus are thermally in parallel across a temperature difference $\Delta T = T_H - T_C$.

Consider first the cooling mode. When a current (I) is supplied, the metal/semiconductor junctions at one end of the couple are reverse-biased and the junctions at the other end are forward-biased. These are the cooling and heating junctions respectively. At the cooling junctions, thermal energy is absorbed from the lattice to account for the mismatch in energy content of the electric current in the metal compared to the semiconductors. If the two semiconductors were joined directly, heat would be carried away from the reverse-biased junction by the creation and separation of electron–hole pairs. At the heating junctions, the energy mismatch is in the opposite sense and heat is transferred from the charge carriers to the lattice.

In the power generation mode, the applied temperature gradient is opposed by a concentration gradient of the mobile carriers. This is the Seebeck effect, and the Seebeck voltage places the hot junctions in reverse bias and the cold junctions in forward bias. By convention, P and N materials have positive and negative coefficients respectively, and the Seebeck coefficient of the couple is $S = S_P - S_N$.

The thermoelectric effect is a reversible thermodynamic process. In the absence of irreversible processes, which are ignored for the moment, the efficiency of thermoelectric energy conversion should be the Carnot limit. In a cooling device, the Peltier cooling power is ST_CI. If no ΔT is allowed to develop, then this heat is transported without energy cost and the coefficient of performance is infinite. With a temperature gradient, the voltage on the couple is $S\Delta T$, and the ratio of heat pumped to electrical power is $T_C/\Delta T$, the Carnot limit for a refrigerator. In a generation device, the thermal voltage is $S\Delta T$. If a current is flowing, the heat absorbed by the electric current at the heated junctions is ST_HI, and the efficiency is $\Delta T/T_H$, again the Carnot limit. In reality, thermoelectric device efficiencies are far from ideal because both electrical resistance and thermal conduction are significant.

Analysis

Generation Referring again to **Figure 1**, the series electrical and parallel thermal arrangement leads to $R = R_N + R_P$, and $K = K_N + K_P$ for the couple. Relative to the ideal situation, the work done on the load is reduced by the internal resistance of the couple:

$$W = IV = I(S\Delta T - IR) \qquad [1]$$

The heat absorbed to perform this work is increased by thermal conduction, but decreased by resistive heating

$$Q_H = SIT_H + K\Delta T - I^2R/2 \qquad [2]$$

The efficiency is $\eta = W/Q_H$. From eqn [1], the maximum power output occurs for $I = S\Delta T/2R$, obtained when $R_{Load} = R$, and has the value

$$W_{max} = KZ(\Delta T)^2/4 \qquad [3]$$

where $Z = S^2/RK$ is the thermoelectric figure of merit of the couple.

In general, Z for a couple is an extrinsic quantity because the geometry factors in R and K do not cancel. In the case of the P and N blocks having the same length-to-area ratio, Z is an intrinsic quantity that depends only on the Seebeck coefficient, electrical resistivity (ρ), and thermal conductivity (κ) of the P and N materials. Relative to one another, the length-to-area ratios of the P and N blocks can be optimized to give the best Z for the couple:

$$Z_{opt} = S^2/[\sqrt{\kappa_P\rho_P} + \sqrt{\kappa_N\rho_N}]^2 \qquad [4]$$

The maximum efficiency of a thermoelectric generator occurs when $R_{Load} = \gamma R$ and is

$$\eta_{max} = (\gamma - 1)\Delta T/[(\gamma + 1)T_H - \Delta T]$$
$$\gamma = \sqrt{1 + Z(T_C + T_H)/2} \qquad [5]$$

For both power generation and cooling, Z governs the maximum efficiency that is possible for a given pair of materials. This is not too surprising since Z concisely captures the requirements to maximize the

thermoelectric effect and minimize the irreversible processes involved.

ZT varies with temperature. For power generation, ΔT can be so large that the couples gain performance by the use of two or more materials in the N leg, the P leg, or both. In these segmented thermocouples, the segment materials are best chosen not just for their individual ZT values, but also for their compatibility with regard to electrical and thermal fluxes.

Commercially available generation materials have $ZT \sim 1$. For a generation device with a "cold" side near ambient and $\Delta T \sim 400^{\circ}\mathrm{C}$, the optimum efficiency is in the vicinity of 10%. The maximum power output per couple is an extrinsic quantity (eqn [3]), a design parameter that can vary greatly. For common direct generation applications (as opposed to waste heat recovery), $W_{\max}$ is ~ 1 watt per couple.

Cooling In a real thermoelectric device, Peltier cooling is offset in part by resistive heating and thermal conduction. The net rate of cooling is

$$Q_{\mathrm{C}} = ST_{\mathrm{C}}I - I^2R/2 - K\Delta T \qquad [6]$$

The work required to perform this cooling is

$$W = IV = I(IR + S\Delta T) \qquad [7]$$

and the coefficient of performance is

$$\phi = Q_{\mathrm{C}}/W \qquad [8]$$

Both Q_{C} and ϕ can be maximized by varying the current, as shown in **Figure 2**. These special current levels are

$$I_{\mathrm{Q}} = ST_{\mathrm{C}}/R \qquad [9]$$

$$I_{\phi} = S\Delta T/(\gamma - 1)R \qquad [10]$$

Use of eqn [9] in eqn [6] leads to

$$Q_{\max} = K(ZT_{\mathrm{C}}^2/2 - \Delta T) \qquad [11]$$

Similar to power generation, the maximum heat pumping capacity of a couple is an extrinsic quantity that increases with the area-to-length ratio of the thermoelectric blocks.

When $\Delta T = Z(T_{\mathrm{C}})^2/2$, no heat can be pumped. This is the largest ΔT that can be achieved with a single stage thermoelectric refrigerator. For the best available materials operating under the best conditions, the maximum ΔT is $\sim 75^{\circ}\mathrm{C}$ for heat sinking at 300 K.

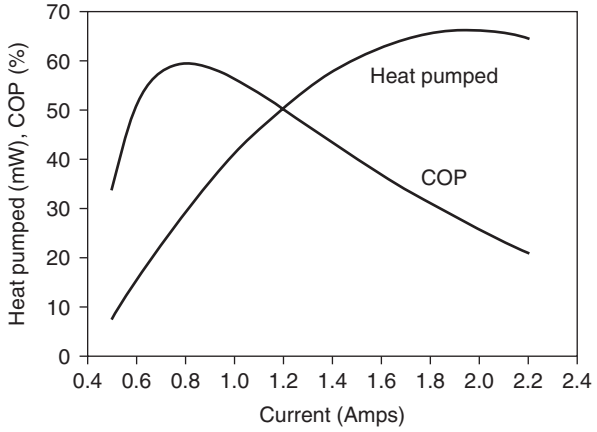


Figure 2 Dependence of heat pumped and coefficient of performance (COP) on supplied current for a typical thermoelectric cooling couple. The couple has Seebeck coefficient, electrical resistance, and thermal conductance of $420\,\mu\mathrm{V\,K^{-1}}$, $56\,\mathrm{m\Omega}$, and $10^{-3}\,\mathrm{W\,K^{-1}}$ respectively. The length-to-area ratio for each leg (P and N) is $28\,\mathrm{cm^{-1}}$. The hot and cold temperatures were taken as 300 K and 260 K.

For values of ΔT below the maximum, some heat can be pumped and the maximum coefficient of performance is given by eqns [8] and [10]:

$$\phi_{\max} = (\gamma T_{\mathrm{C}} - T_{\mathrm{H}})/(\gamma + 1)\Delta T \qquad [12]$$

The curves in **Figure 2** are plots of the current dependence of Q_{C} and ϕ with $T_{\mathrm{C}} = 260\,\mathrm{K}$, $T_{\mathrm{H}} = 300\,\mathrm{K}$, $\gamma = 1.37$, and $K = 10^{-3}\,\mathrm{W\,K^{-1}}$. This value of K is typical for a couple in a small cooling device. If such a device is operated near peak ϕ, then about 25 couples are needed to pump one watt.

Multiple stages In cooling applications, it is common to stack multiple thermoelectric devices to achieve greater temperature differences than possible with a single stage. Devices with six stages are able to reach temperatures as low as 170 K from 300 K, although with a quite low coefficient of performance.

With reference to **Figure 3**, the efficiency of a two-stage device is

$$\Phi = \frac{Q_{\mathrm{C}}}{W_1 + W_2} = \frac{Q_{\mathrm{C}}}{\dfrac{Q_{\mathrm{C}}}{\phi_1} + \dfrac{Q_{\mathrm{C}} + Q_{\mathrm{C}}/\phi_1}{\phi_2}} = \frac{1}{\dfrac{1}{\phi_1} + \dfrac{1}{\phi_2} + \dfrac{1}{\phi_1\phi_2}} \qquad [13]$$

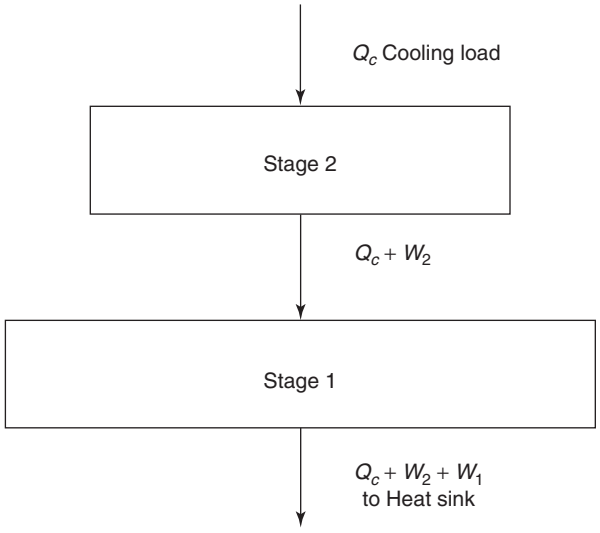


Figure 3 Schematic illustration of a two-stage thermoelectric cooling device showing how the total amount of heat increases from the cooling load to the heat sink. W_1 and W_2 are the electrical power consumed by the first and second stages respectively.

The generalization for n stages, each with the same coefficient of performance, is

$$\Phi = \frac{1}{(1+1/\phi)^n} - 1$$

Construction

Coolers **Figure 4** is a side view of a traditional thermoelectric cooler, which consists of numerous couples connected in series. Soldering the thermoelectric blocks to metal interconnects, usually copper pads that are directly bonded to ceramic plates, makes the connections. The thermoelectric materials are discussed in the next section. For now, some insight into the benefits and deficiencies of the traditional cooler design is provided.

The ceramic plates provide mechanical structure, low thermal resistance, and a flat surface for interfacing with heat transfer plates. The ceramic plate that is heat sunk also accommodates the electrical leads. The most common ceramics are Al_2O_3, AlN, and BeO. Al_2O_3 is the least expensive substrate choice, and has adequate thermal conductivity for many applications. BeO has the highest thermal conductivity of this group at $\sim 220\,W\,m^{-1}\,K^{-1}$, but is less favored due to cost and health concerns. With a thermal conductivity about half that of BeO, AlN is the most prevalent ceramic for applications with a high heat flux. AlN also has a low coefficient of thermal expansion, which is advantageous because contraction and expansion of the ceramic plates during thermal cycling eventually causes cracking within the thermoelectric material or the solder joint.

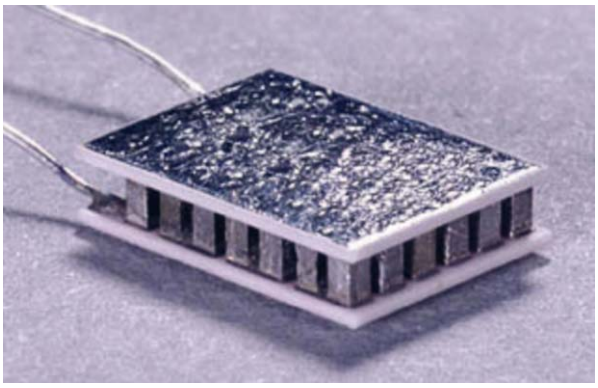


Figure 4 Side view of a commercial thermoelectric cooling module. The module dimensions are $\sim 8\,mm \times 6\,mm \times 2\,mm$ and its components and assembly are discussed in the text. Photo courtesy of Marlow Industries, Inc.

A variety of solders are used within the thermoelectric cooling industry. As a rule, these solders are the same ones used in other electronic applications and are selected for similar reasons, such as ease of manufacturing, strength, and low temperature ductility. One exception is a group of solders that are more than 90% bismuth. These solders have very good thermal cycling fatigue and can be applied directly to thermoelectric materials with acceptable results. To use other solders, it is generally necessary to protect the thermoelectric materials with a metal layer, usually nickel, to prevent interactions.

There are three main technical deficiencies of cooler design as described here. The first, susceptibility to thermally induced stress, has been mentioned already. The other two are thermal resistance of the ceramic plates, and electrical resistance of the interconnects and interfaces. Both of these problems become serious for high watt-density coolers with short elements, and reach their extreme in modules based on thin film materials.

Generators **Figure 1** emphasizes the reversibility of thermoelectric energy conversion. In principle, the same device can be used for both cooling and power generation, and in both cases only the temperatures and the material quality determine the maximum efficiency. In reality though, a cooling device is suitable for generation only if the temperature difference is relatively small and the heat source is at a modest temperature.

The construction of a device for generation under less restrictive conditions is illustrated in **Figure 5**. In this case, the thermoelectric blocks are embedded in insulation that is a component of the finished device. With regard to providing mechanical structure, this

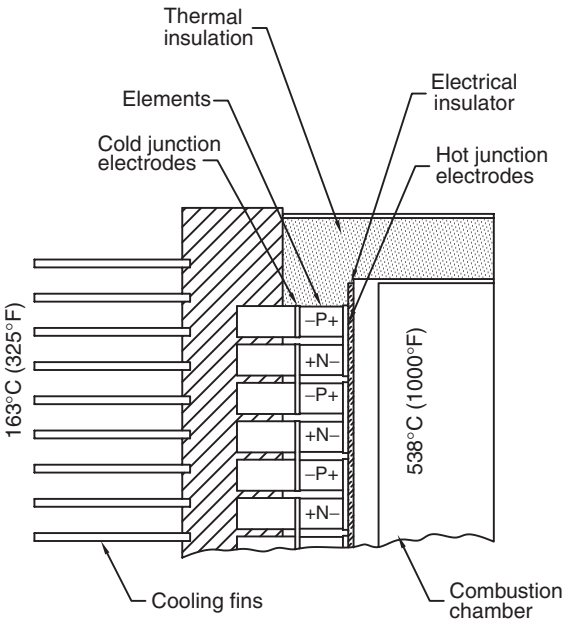


Figure 5 Illustration of the implementation of a thermoelectric generator. Each component of the assembly plays a role in thermal performance, mechanical stability, or chemical isolation. Image courtesy of Global Thermoelectric.

insulation takes the place of the ceramic plates used in cooling devices. Rigidly fixed substrates usually are not part of a generator design because thermal expansion would create destructive levels of stress within the module. A second major difference is in the method of joining the metal interconnects to the thermoelectric materials. This cannot be done with a solder. Two methods are hot pressing and thermal spraying. A third difference is that thermal interface resistances between the device and heat exchange plates are reduced with clamps or springs, rather than solder or rigid epoxies. Finally, thermoelectric generators are likely to be sealed to prevent oxidation of the thermoelectric materials.

Sound construction principles allow thermoelectric generators to survive high temperatures, and large temperature differences between the hot and cold surfaces for very long times. The attractiveness of 20 years of maintenance-free operation has led to their use in applications where maintenance is impossible (interplanetary exploration) or expensive (remote terrestrial locations), even when other technologies offer greater efficiencies.

Thermoelectric Materials

Fundamentals

It has been discussed elsewhere in this encyclopedia why good thermoelectric materials are heavily doped semiconductors with a Seebeck coefficient in the range 200–300 $\mu V K^{-1}$. Empirically, the electrical resistivity tends to lie in the range 1–3 $m\Omega cm^{-1}$ at the optimized carrier concentration. Most semiconductors do not obtain such favorable values of S and ρ simultaneously, and this is part of what makes good thermoelectric materials unique and uncommon.

Even more so, it is the lattice thermal conductivity that sets thermoelectric materials apart. The thermal conductivity values of diamond, fused silica, and common plastics are 500, 1.5, and 0.2 $W m^{-1} K^{-1}$ respectively. Good thermoelectric materials are crystals, and yet their lattice thermal conductivity is $\sim 1 W m^{-1} K^{-1}$, less than that of silica. (The charge carriers typically contribute another 0.5 $W m^{-1} K^{-1}$ to the thermal conductivity at low to moderate temperatures.) The search for new semiconductors with extreme electrical and thermal properties is a fascinating challenge that has attracted the attention of materials scientists, particularly since useful materials first were found in the 1940s and 1950s.

Almost without exception, materials with good Z values are mixtures of two or more compounds that share a crystal structure. The reason for this is the reduction of the lattice thermal conductivity by mass and elastic fluctuations. These mixtures are commonly referred to as "alloys."

Established Materials

Bi_2Te_3 and PbTe alloys dominate commercial cooling and power generation respectively. SiGe alloys and TAGS have been used in some higher temperature-generation applications. BiSb is an alloy with superior low-temperature properties, but it has found no commercial use.

Bi_2Te_3 alloys For operation between 200 and 500 K, the best P and N materials are found in mixtures of Bi_2Te_3, Sb_2Te_3, and Bi_2Se_3. These temperatures cover both cooling and generation from low-grade heat. The P material is approximately 75% Sb_2Te_3 and 25% Bi_2Te_3. A few percent of Se substitution for Te may be found to improve the properties, depending on the method of fabrication. The N material is approximately 90% Bi_2Te_3 and 10% Bi_2Se_3 for room temperature and above. Below room temperature, a formula of 90% Bi_2Te_3, 5% Bi_2Se_3, and 5% Sb_2Te_3 is preferred by some. The P material is autodoped by a slight substitution of Sb for Te, while the N material must be doped with excess Te or a halide. The melting point for all of these alloys is in the vicinity of 600°C.

The P and N materials share a crystal structure that is based on covalently bonded sheets of five atomic layers. The bonding of the atoms in the central Te/Se

layer and the adjacent Bi/Sb layers is octahedral, while the atoms of the outer Te/Se layers are bonded predominantly on one side to three Bi/Sb atoms. van der Waals bonds between outer Te/Se layers connect adjacent sheets more weakly. Not surprisingly, the layered structure results in anisotropic electrical, thermal, and mechanical properties. Electrical and thermal conductivity are higher within the plane of the covalent sheets, while the Seebeck coefficient is nearly isotropic. For *P* material, the figure of merit is essentially isotropic. For *N* material, though, the electrical anisotropy is much stronger and the figure of merit is significantly less for transport perpendicular to the covalent sheets. Single crystals or large grains of these materials exhibit mechanical cleavage that can be a hindrance to device fabrication.

Directional solidification, hot pressing, and plastic deformation can make good *P* material. Historically, directional solidification from a melt has been the most favored approach. Recently, though, researchers and production facilities have found that hot pressing and plastic deformation produce equally good, if not better, *P* material. Good *N* material can be made by directional solidification or plastic deformation. Hot pressing is not satisfactory because it does not produce adequate texture. Bi_2Te_3 and its alloys are brittle at room temperature, and so plastic deformation is performed at temperatures in excess of 400°C.

PbTe alloys PbTe and 75% PbTe + 25% SnTe are good *P* and *N* materials respectively for operation from 500 to 800 K. Direct generation of electricity from combustion heat is the most prevalent application in this temperature range. These materials have melting points ~1200 K, and are stable up to almost 800 K when used in sealed modules. The main path of degradation is sublimation of Te.

PbTe and SnTe have the cubic rock-salt crystal structure. Since the structure and properties are isotropic, hot pressing is a suitable manufacturing method. By stoichiometric variations, PbTe can be made either *P* (excess Te) or *N* (excess Pb). In neither case, though, is the doping level adequate. Higher carrier concentrations can be reached by substitution of Na or K for Pb or a halide for Te.

TAGS TAGS refers to alloys of $AgSbTe_2$ and GeTe, and is *p*-type only. (The name comes from the first letters of the chemical symbols of the constituents.) TAGS has the highest *Z* of *P* materials in the range 500–800 K and has been a mainstay of the radioisotope thermal generators (RTG) used in several space missions. $AgSbTe_2$ and GeTe do not have the same crystal structure, and this presents a problem in the utilization of TAGS. The structure of TAGS can be either rhombohedral (like the low-temperature form of GeTe) or cubic (like $AgSbTe_2$) depending on composition and temperature. Thermal cycles through the transformation temperature can lead to mechanical failure. Another problem is Ag mobility, which becomes worse at higher temperatures.

BiSb BiSb is an *N* material for low temperatures. The crystal structure is hexagonal, and the properties are anisotropic. High-quality crystals are superior to Bi_2Te_3 alloys at temperatures less than 200 K, but these crystals have cleavage planes perpendicular to the direction of current flow. BiSb made by powder metallurgy overcomes this mechanical problem, but loses *Z*. Consequently, BiSb is the best *N* material only at temperatures lower than can be reached by multistage cooling from room temperature.

Materials in Development

In this category, materials that are known to exhibit relatively good thermoelectric performance with the potential for further improvement in *ZT* or other relevant properties are placed. There are too many such materials to attempt even a cursory survey. Any recent volume of the annual International Conference on Thermoelectrics Proceedings is a good source for further reading on new thermoelectric materials. Here, two of the prominent families of new materials are highlighted.

Skutterudites Semiconducting skutterudites have the formula AB_3, where A = Co, Ir, Rh and B = P, As, Sb. The skutterudite materials of most interest for thermoelectric energy conversion are based on $CoSb_3$. Both *p*-type and *n*-type $CoSb_3$ have large values of the power factor. The lattice thermal conductivity is too high, but it can be greatly reduced by heavy doping or alloying. Another way to reduce the thermal conductivity is by introduction of rare-earth or lanthanide elements into voids in the crystal structure, and concurrent substitution of Fe or Ni for some of the Co. The La-Fe-Co-Sb and Ba-Ni-Co-Sb systems have yielded *p*-type and *n*-type formulas respectively with $ZT > 1$ at 600°C. These materials are being developed as components of segmented thermocouples for the next generation of RTGs for space missions.

Cobalt oxides In 1997, $NaCo_2O_4$ was found to be a good *p*-type thermoelectric material at high temperatures, $ZT \sim 1$ at 900–1000 K. This was a surprising discovery because oxides had not been considered good candidates. Cobalt oxides are now a major focus of thermoelectric materials development, particularly in Japan, and these explorations have uncovered a large variety of materials

based on CoO_2 layers that are charge-balanced by anions that reside between the layers. Some of these rival $NaCo_2O_4$ with regard to *ZT*. The cobalt oxides appear to be unusual materials in which a good Seebeck coefficient at nearly metallic hole concentrations might be associated with a mixed valence state of Co.

Nanoscale Materials

Present thermoelectric cooling and generation products are based on homogeneous bulk materials. The quest for higher *ZT*, though, may lead the industry toward inhomogeneous, nanostructured materials.

In this active area of research, the general rationale is to create materials with enough control of structure that S, ρ, and κ can be separately tuned. This is a difficult problem for a variety of reasons. First, the phonon mean free path is already quite low compared to the charge carrier mean free path in most good thermoelectric materials. In these cases, grain size reduction alone is unlikely to improve *ZT*. Second, very small dimensions, $\leqslant 50$ nm, are required to improve the ratio S^2/ρ by modification of the electron and hole state densities. Third, thermoelectric materials are compound semiconductors, and usually the stoichiometry and cleanliness must be controlled precisely to avoid excessive doping and carrier scattering. Also, these materials do not possess great ductility, and unsupported small-scale samples are easily damaged. Adding to these synthesis challenges, measurements, especially evaluation of thermal conductivity, are difficult because of the forms of the nanostructured materials. Nevertheless, the theoretical encouragement for pursuit of nanoscale thermoelectric materials is strong enough that a number of experimental approaches are being tested.

Superlattices Three superlattice systems serve to demonstrate the promise and peril of nanoscale thermoelectric materials. These are $PbTe/Pb_{1-x}Eu_xTe$, Bi_2Te_3/Sb_2Te_3, and $PbTe/PbTe_{1-x}Se_x$.

$PbTe/Pb_{1-x}Eu_xTe$ is a quantum-well system. The layers alloyed with Eu serve as energy barriers to the carriers in the PbTe layers. The study of this system showed, in agreement with theory, that the power factor for transport parallel to the layers could be improved in quantum wells that are sufficiently thin. Thermal conductivity was not measured, but under the safe assumption that it does not increase relative to bulk material, the *ZT* of the PbTe wells is apparently enhanced. The *ZT* of the overall superlattice is not good because the barrier layers are electrically passive but thermally active, and these layers cannot be made arbitrarily thin without losing their effectiveness as barriers.

A quite different approach is exemplified by work on Bi_2Te_3/Sb_2Te_3 superlattices. Here, there appear to be no energy barriers for *p*-type transport perpendicular to the layers. In fact, reported S^2/ρ values are larger than those of the bulk *p*-type alloy in the corresponding direction. Further, measurements by a thin-film method indicate that the cross-plane lattice thermal conductivity is significantly decreased, approaching the theoretical lower limit for a repeat thickness of ~ 4 nm. With improvements in both electrical and thermal properties, *ZT* values exceeding 2.0 are reported. Similar structures of *n*-type Bi_2Te_3/Bi_2Se_3 are not as exceptional, but estimates of the couple *ZT* exceed the bulk benchmark. To translate a high *ZT* in such structures into improved thermoelectric performance, a device design must overcome three hurdles associated with the thinness of the superlattices: contact resistance, interconnect resistance, and thermal resistance of the substrate(s).

In the case of $PbTe/PbSe_{1-x}Te_x$ ($x \approx 0.02$) superlattices, a convincing demonstration of advanced *ZT* has been made. A couple composed of a freestanding, 100-μm thick superlattice and a gold ribbon achieved a ΔT of 44°C. The inferred *ZT* value of the superlattice is 1.3–1.6. In these structures, the $PbSe_{1-x}Te_x$ layers are not continuous; they are spontaneously formed islands in a matrix of PbTe. S, ρ, and ΔT measurements indicate that both electrical and thermal properties are improved compared to bulk PbTe–PbSe alloys. The mechanisms responsible for these improvements are not known. In this case, the method of fabrication, molecular beam epitaxy, might not be viable for commercial thermoelectric applications. The laboratory results, though, have spawned interest in synthesizing similar nanostructures by scalable means.

Nanowires Under certain conditions, anodization of pure aluminum results in a layer of alumina with a dense array of cylindrical pores with their axes perpendicular to the surface and virtually parallel to one another. With pore diameters and spacings as small as 20 nm possible, anodized alumina is intriguing as a template for fabrication of nanowires of many different materials, including thermoelectric ones.

For BiSb, quantum confinement effects are expected for dimensions less than 100 nm. By filling alumina nanopores with Bi or BiSb, it is possible in principle, to test the electrical properties well into the quantum regime. Absolute measurements of the electrical resistivity are difficult, but the normalized temperature dependence can be measured more readily. Such measurements show a crossover from semimetallic to semiconducting behavior as a function of decreasing wire diameter, in support of calculations. This

transformation is advantageous because both electrons and holes contribute to conduction in bulk BiSb, which decreases the Seebeck coefficient. If the band overlap is eliminated, it might be possible to reduce the influence of the holes. Also, the lattice thermal conductivity will be reduced in the nanowire format. Whether these two benefits can overcome the expected decrease of electrical conductivity is not known.

One of the intriguing aspects of thermoelectric nanowire research is that wires with lengths greater than 50 μm have been grown by electrochemistry. If the length can be doubled, then the wires begin to reach the "bulk" regime, and it is conceivable to make couples and devices by established methods. Also, the method appears to be applicable to other materials, including Bi_2Te_3.

Inhomogeneous bulk materials The ideal thermoelectric material might be described as "kilograms of nanostructure." There is a rationale to the nanostructure approach – tailoring the material for separate control of electrical and thermal properties – that is difficult to refute. At the same time, cooling and power generation have extrinsic, peripheral, and economic aspects that require more of the material than a high *ZT*.

Nanoscale inclusions in a thermoelectric matrix, polymer/conductor composites, segmented nanowires, and thermionic/tunneling interfaces are some possible manifestations of the nanostructured bulk concept. As various factors increase motivation for alternative refrigeration and improved energy efficiency, perhaps a discovery in one of these archetypical systems will point the way to new levels of thermoelectric performance.

Acknowledgments

The author thanks Marlow Industries, Inc. for support, J Bolen for technical discussions, and S Sharp for technical assistance.

See also: Thermoelectric Effect.

PACS: 85.80.Fi; 84.60.Rb; 84.60.−b; 84.60.Bk; 81.05.Hd; 81.07.−b; 81.07.St; 81.07.Vb

Further Reading

Chen G, Dresselhaus MS, Dresselhaus G, Fleurial J-P, and Caillat T (2003) Recent developments in thermoelectric materials. *International Materials Reviews* 48(1): 45–66.

Heikes RR and Ure RW (1961) *Thermoelectricity: Science and Engineering*. New York: Interscience.

Ioffe AF (1957) *Semiconductor Thermoelements and Thermoelectric Cooling*. London: Infosearch.

Nolas GS, Sharp J, and Goldsmid HJ (2001) *Thermoelectrics: Basic Principles and New Materials Developments*. Berlin: Springer.

Nolas GS, Yang J, Hogan TP, and Johnson DC (eds.) (2003) *Thermoelectric Materials 2003–Research and Applications, MRS Proceedings* vol. 793. Pittsburgh: Materials Research Society. (This is the most recent of several MRS Proceedings on the subject.)

Rowe DM (ed.) *CRC Handbook of Thermoelectrics*. Boca Raton, Florida: CRC Press.

Wood C (1988) Materials for thermoelectric energy conversion. *Reports of Progress in Physics* 51: 459–539.

Nomenclature

I	electric current (A)
κ	thermal conductivity ($W\,m^{-1}K^{-1}$)
K	thermal conductance ($W\,K^{-1}$)
ρ	electrical resistivity (mΩ-cm)
Q	rate of heat flow (W)
R	electrical resistance (mΩ)
S	Seebeck coefficient ($\mu V\,K^{-1}$)
T	temperature (K,°C)
ΔT	temperature difference (°C)
V	voltage (V)
W	power supplied or produced (W)
Z	figure of merit (1/K)
ZT	dimensionless figure of merit
η	efficiency
ϕ,Φ	coefficient of performance

Thermoelectric Effect

G D Mahan, The Pennsylvania State University, University Park, PA, USA

Introduction

The early researchers in material science discovered a variety of phenomena that collectively were called thermoelectric effects. Now it is realized that the different properties are due to a common origin. There are only two basic thermoelectric phenomena: a volume effect described by a Seebeck coefficient (S), and an interface effect described by a boundary Seebeck coefficient S_B. The boundary Seebeck is due to electrons tunneling through barriers, but has not been well characterized.

The Seebeck coefficient is an important thermoelectric property. It has the units of volts per degree kelvin, and is usually expressed in $\mu V K^{-1}$. Early workers used other symbols, including Q and α. The symbol S is to honor Thomas Seebeck, who discovered the thermoelectric phenomena in 1823. S has a simple experimental definition. Take a bar or wire of length L, and put a slightly different temperature on each end: say $T_0 - \Delta T/2$ on the "cold" end, and $T_0 + \Delta T/2$ on the "hot" end. The temperature difference between the ends is ΔT. If the ends of the bar are electrically insulated, such that no current flows, it is found that a voltage difference ΔV is measured between the ends of the bar. The Seebeck coefficient is the ratio of these two experimental numbers

$$S(T_0) = \frac{\Delta V}{\Delta T}$$

All materials have a nonzero value for S. In many cases it is small: small means less than $1.0\,\mu V K^{-1}$. Generally, good conductors such as metals have a small value for Seebeck, while poor conductors such as insulators have a large value. The Seebeck coefficient depends upon temperature, and vanishes at the absolute zero $T = 0\,K$. The Seebeck is the entropy of the charge carriers divided by the electrical charge.

Two other historic thermoelectric coefficients are related to the Seebeck coefficient. The Peltier effect is the heating or cooling of a junction, between two different metals A and B, as a current flows through an interface between them

$$\frac{dQ}{dt} = J[\Pi_A - \Pi_B]$$

where J is the current density (amperes per area). Now it is known that $\Pi_j = TS_j$ and the Peltier effect is a manifestation of the volume effect. The Thomson effect, due to Lord Kelvin, is the bulk heating of a thermoelectric due to the temperature variations in the Seebeck

$$\frac{dQ}{dt} = \rho J^2 - \tau J \nabla T, \quad \tau = T\frac{dS}{dT}$$

where ρ is the electrical resistivity. Many solids have $S \sim T$, so the Thomson coefficient τ is as large as the Seebeck coefficient.

There are two different reasons to measure the Seebeck coefficient. The first is science. A measurement of $S(T)$ provides important information about the properties of a material. They are discussed later, but one example is the sign of the Seebeck which is determined by the sign of the charge carrier: $S < 0$ for electrons, and $S > 0$ for holes in a semiconductor.

The second reason to measure the Seebeck is to identify materials that could be used in thermoelectric refrigerators or energy conversion devices. There it is shown that the efficiency of such devices is determined by a parameter denoted as Z, which is called "the figure of merit"

$$Z = \frac{\sigma S^2}{K}$$

where σ is the electrical conductivity and K is the thermal conductivity. Each of these important transport coefficients is discussed in separate chapters in this section of the encyclopedia. The Seebeck coefficient plays a major role in the figure of merit, since it enters as the square. The efficiency of thermoelectric devices is an increasing function of Z. The material with the largest Z has the largest device efficiency. There is a continual worldwide quest to find new materials with larger values of Z. Since $Z(T)$ depends upon temperature, this search involves many different materials. Materials with the largest values of Z are narrow gap semiconductors. They typically have values of the Seebeck in the range of $S \sim 250$–$300\,\mu V K^{-1}$. Materials for thermoelectric devices should have large Seebeck coefficients, large values of electrical conductivity, and small values of thermal conductivity.

Transport Theory

This section explains the microscopic origins of the Seebeck coefficient. The most important inputs are the electronic energy bands of the crystal. Two transport coefficients σ and S can be calculated from first principles if one knows enough about the electronic energy bands. The electrical conductivity σ is needed for the calculation of the Seebeck coefficient.

Each energy band $E_j(\boldsymbol{k})$ has a band label j and a wave vector $\boldsymbol{k} = (k_x, k_y, k_z)$. If the energy bands are well separated in energy, each band j contributes to the two transport coefficients σ_j and S_j.

Let $f_j(\boldsymbol{k})$ denote the number of electrons or holes of wave vector $\boldsymbol{k}$ in a band j at temperature T. This distribution consists of an equilibrium term $f_j^{(0)}$ plus a nonequilibrium term δf_j. The equilibrium term is given by the Fermi–Dirac function

$$f_j^{(0)} = \frac{N_j}{e^{\beta[E_j(\boldsymbol{k})-\mu]} + 1}$$

$$E_j(\boldsymbol{k}) = E_{j0} + \varepsilon_j(\boldsymbol{k}), \quad \varepsilon_j(\boldsymbol{k}) = \sum_{\alpha=1}^{3} \frac{\hbar^2 k_\alpha^2}{2m_{j\alpha}}$$

where $\beta = 1/k_B T$, and k_B is Boltzmann's constant.

At low and moderate temperatures, the electrons are in thermal equilibrium near the states of lowest energy: conduction bands in n-type material, and valence bands in p-type material. E_{j0} is the energy of

the band minimum. The kinetic energy in the band is $\varepsilon_j(\boldsymbol{k})$, which has been defined for a band with different effective masses in different directions. The symbol μ denotes the chemical potential.

The numerator contains the degeneracy N_j of the band. $N_j = 2$ for a single band with only spin degeneracy. Many semiconductors have several equivalent conduction or valence bands. For example, silicon has six equivalent conduction bands, and $N_j = 12$ including spin degeneracy.

The equilibrium distribution $f_j^{(0)}$ does not cause currents. Currents are caused by the nonequilibrium term δf_j which is generated by a temperature difference, or an applied electric field. For the present discussion, it is assumed that currents are caused by an electric field $\mathscr{E}_m$ in the direction r_m. The total distribution function is

$$f_j = f_j^{(0)} + \delta f_{jm}\mathscr{E}_m$$

The particle density n_j, and two transport coefficients, are defined as

$$n_j = \int \frac{\mathrm{d}^3k}{(2\pi)^3} f_j^{(0)}(\boldsymbol{k})$$

$$\sigma_j = q_j \int \frac{\mathrm{d}^3k}{(2\pi)^3} v_{jm}\delta f_{jm}(\boldsymbol{k}, T)$$

$$T\sigma_j S_j = \int \frac{\mathrm{d}^3k}{(2\pi)^3} v_{jm}\delta f_{jm}(\boldsymbol{k}, T)[E_j(\boldsymbol{k}) - \mu]$$

where the charge is $q_j = \pm q$, and the electron has $-q$. The velocity is $v_{jm}(\boldsymbol{k}) = \nabla_k E_j(\boldsymbol{k})$.

The above formulas have the interesting feature that the Seebeck coefficient is the ratio of two integrals. The product of $\sigma_i S_i$ is one integral, and one gets S_i on dividing it by σ_i which is another integral. The two integrals differ mainly in the factor of $[E_j(\boldsymbol{k}) - \mu]$. The electrical conductivity is a measure of how much current is being carried by the electrons, while the Seebeck coefficient is a measure of how much energy is being carried by the electrons. The energy factor of $[E_j(\boldsymbol{k}) - \mu]$ is the particle energy minus the chemical potential. Rather than energy, the term "heat" is used. Electrons with $[E_j(\boldsymbol{k}) - \mu] > 0$ are called "hot," while those with $[E_j(\boldsymbol{k}) - \mu] < 0$ are called "cold." The Seebeck is nonzero when there are different densities of hot and cold electrons participating in the transport.

A formula for $\delta f_{jm}(\boldsymbol{k}, T)$ is derived by solving the Boltzmann transport equation. The simple case is when the conducting particles, electrons, or holes, have a relaxation time $\tau_j(\boldsymbol{k})$ that describes the momentum transfer to impurities and phonons

$$\delta f_{jm}(\boldsymbol{k}) = q_j v_{jm}\tau_j(\boldsymbol{k})\left(-\frac{\mathrm{d}f_j^{(0)}(\boldsymbol{k})}{\mathrm{d}E_j(\boldsymbol{k})}\right)$$

From the previous definitions of σ_j and S_j, this choice of distribution function gives simple expressions for these transport coefficients:

$$\sigma_j = q_j^2 \int \frac{\mathrm{d}^3k}{(2\pi)^3} v_{jm}^2\tau_j(\boldsymbol{k})\left(-\frac{\mathrm{d}f_j^{(0)}(\boldsymbol{k})}{\mathrm{d}E_j(\boldsymbol{k})}\right)$$

$$T\sigma_j S_j = q_j \int \frac{\mathrm{d}^3k}{(2\pi)^3} v_{jm}^2\tau_j(\boldsymbol{k})[E_j(\boldsymbol{k}) - \mu] \times \left(-\frac{\mathrm{d}f_j^{(0)}(\boldsymbol{k})}{\mathrm{d}E_j(\boldsymbol{k})}\right)$$

Both the electrical conductivity (σ_{kl}) and the Seebeck coefficient (S_{kl}) are tensor quantities. The above formulas give the *mm* components of these tensors. Other components (*mn*) have the velocity factors $v_{jm}v_{jn}$ in the integrand, instead of v_{jm}^2.

The above formulas are easily evaluated with a knowledge of the energy band dispersion $E_j(\boldsymbol{k})$. The only factor that does not depend upon $E_j(\boldsymbol{k})$, in a simple way, is the quasiparticle lifetime $\tau_j(\boldsymbol{k})$. For the electrical conductivity, a detailed calculation of the lifetime is required in order to obtain an accurate expression. Sometimes this work is also required for the Seebeck coefficient. However, the Seebeck has the advantage that it is the ratio of two integrals. If the lifetime is a constant, independent of $\boldsymbol{k}$, then it can be removed from the two integrals. It cancels from the ratio. If this approximation is valid, the Seebeck is given by

$$S(T) = \frac{k_\mathrm{B}}{q_j}\frac{L_1}{L_0}$$

$$L_j = \int \frac{\mathrm{d}^3k}{(2\pi)^3} v_{jm}^2\left(\frac{E_j(\boldsymbol{k}) - \mu}{k_\mathrm{B}T}\right)^j\left(-\frac{\mathrm{d}f_j^{(0)}(\boldsymbol{k})}{\mathrm{d}E_j(\boldsymbol{k})}\right)$$

The size of the Seebeck coefficient is determined by the ratio of the two fundamental constants $k_\mathrm{B}/|q| = 86.1734\,\mu\mathrm{V\,K}^{-1}$. The ratio L_1/L_0 is dimensionless.

An energy band calculation gives both energies $E_j(\boldsymbol{k})$ and velocities $v_{jk}(\boldsymbol{k})$. The integrals for L_j can be calculated accurately as a function of temperature. The easiest way is to first calculate the function

$D_{j,m}(E)$ defined as

$$D_{j,m}(E) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} v_{jm}^2 \delta[E_j(\boldsymbol{k}) - E]$$

$$L_{j,m} = \int \mathrm{d}E D_{j,m}(E) \left(\frac{E-\mu}{k_\mathrm{B}T}\right)^j \left(-\frac{\mathrm{d}f_j^{(0)}(E)}{\mathrm{d}E}\right)$$

Evaluating the above expression is now a routine process for computer programs that compute the energy bands of solids. It is not an exact expression for the Seebeck, because of the assumption that the lifetime cancels out from the ratio of the two integrals. However, it gives expressions that are accurate to about 10% for many solids, over a wide range of temperature. Note that $D_{j,m}(E)$ is independent of temperature, and only needs to be computed once.

The above expression will be evaluated for several different types of energy band structures: metals, insulators, and semiconductors.

Metals

Metals have a high density of electrons in their conduction band. The chemical potential μ is typically several electron volts, or more, above the bottom of the conduction band. The function $D(E)$ (its subscripts are dropped) is a smooth function of energy in the neighborhood of μ.

It is expanded in a power series around this point

$$D(E) = D(\mu) + (E-\mu)\left(\frac{\mathrm{d}D}{\mathrm{d}E}\right)_\mu + O(E-\mu)^2$$

$$L_0 = D(\mu)\left[1 + O\left(\frac{k_\mathrm{B}T}{\mu}\right)^2\right]$$

$$L_1 = \frac{\pi^2}{3} k_\mathrm{B}T \left(\frac{\mathrm{d}D(\mu)}{\mathrm{d}\mu}\right)$$

In evaluating these integrals, note that the function

$$\left(-\frac{\mathrm{d}f_j^{(0)}(E)}{\mathrm{d}E}\right) = \frac{\beta}{(\mathrm{e}^x + 1)(\mathrm{e}^{-x} + 1)},$$

$$x \equiv \frac{E-\mu}{k_\mathrm{B}T}$$

is perfectly symmetric in $E-\mu$. Let the new variable of integration be x, and the above integrals become

$$L_j = \int \mathrm{d}x D(\mu + x k_\mathrm{B}T) \frac{x^j}{2[1+\cosh(x)]}$$

The integral vanishes if there are odd powers of x in the integrand. If $j=0$, then the leading term in L_0 must be proportional to $D(\mu)$. If $j=1$, then expand $D(E)$ and take the first derivative to get even powers of x.

The Seebeck coefficient is the ratio of the above two integrals

$$S = \frac{\pi^2}{3} \frac{k_\mathrm{B}^2 T}{e} \frac{\mathrm{d}}{\mathrm{d}\mu} \ln[D(\mu)]$$

This famous formula is called the "Mott relation." It states that the leading term in $S(T)$ is linear in temperature, and the coefficient is determined by the derivative of $D(\mu)$. The formula is usually accurate, since most metals have a Seebeck coefficient that is linear in temperature, particularly at room temperature and higher. If $D(\mu) = D_0 \mu^r$ then

$$\frac{\mathrm{d}}{\mathrm{d}\mu} \ln[D(\mu)] = \frac{r}{\mu}, \quad S \sim r \frac{\pi^2}{3} \frac{k_\mathrm{B}}{q} \frac{k_\mathrm{B}T}{\mu}$$

The last ratio of $k_\mathrm{B}T/\mu \ll 1$ in metals, which makes the Seebeck have a small value.

Figure 1 shows the Seebeck coefficient, in units of $\mu\mathrm{V\,K^{-1}}$, for copper, silver, and gold. For copper, $S(T)$ is linear over a large range of temperature. For silver and gold, there is some deviation from linearity at high temperature. At room temperature, the values are smaller than $2\,\mu\mathrm{V\,K^{-1}}$. It is to be noted that the values are positive. In these metals, the current and heat are carried by electrons, so one might think the Seebeck would be given by the charge on the electron, which is negative. It is positive in these metals since the Fermi surface has regions of negative curvature. No data is shown below 100 K. Each curve has a peak at low temperatures due to phonon drag.

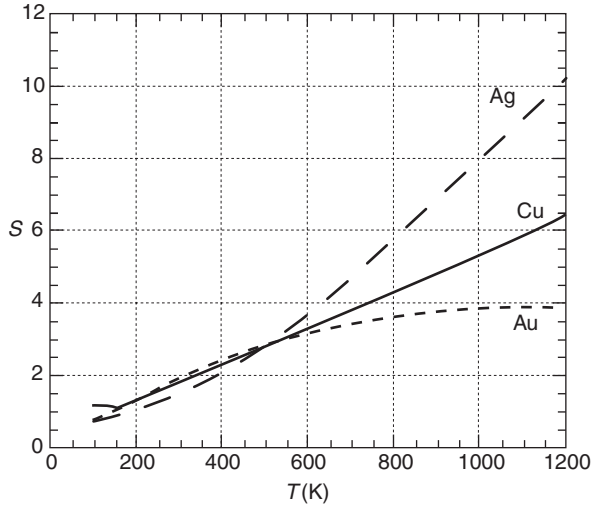


Figure 1 Seebeck coefficient of copper, silver, and gold. Units are microvolts per degree kelvin.

Insulators

Insulators tend to have large values of the Seebeck coefficient, and very small values of the electrical conductivity. There is an energy gap E_g between the occupied energy bands, called valence bands, and the unoccupied bands that are called conduction bands. The probability of exciting an electron from an occupied valence band, to an empty conduction band, is proportional to $\sim\exp(-\beta E_g)$, $\beta = 1/k_B T$. Typically $\beta E_g \gg 1$ and the density of carriers in the conduction band is small. This situation leads to a large Seebeck coefficient.

The chemical potential is located in the energy gap $E_{v0} < \mu < E_{c0}$, where E_{c0} and E_{v0} are the edges of the conduction and valence bands. For the conduction bands, since $\mu < E_{c0}$, the occupation number can be accurately approximated by the Maxwell–Boltzmann limit

$$f_c^{(0)}(\boldsymbol{k}) = N_c \exp\{\beta[\mu - E_c(\boldsymbol{k})]\}$$
$$\left(-\frac{\mathrm{d}f_c^{(0)}(E)}{\mathrm{d}E}\right) = \beta N_c \exp\{\beta[\mu - E_c(\boldsymbol{k})]\}$$

This approximation enables the integrals to be evaluated for the Seebeck coefficient: for example, for the conduction band

$$n_e = N_c e^{\beta(\mu - E_{c0})} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \, \mathrm{e}^{-\beta\varepsilon_c(\boldsymbol{k})}$$
$$= \frac{N_c}{\lambda_c^3} \, \mathrm{e}^{\beta[\mu - E_{c0}]}$$
$$\lambda_c^3 = \left[\frac{2\pi\hbar^2}{k_B T}\right]^{3/2} \frac{1}{\sqrt{m_x m_y m_z}}$$
$$L_{0c} = N_c e^{\beta[\mu - E_{c0}]} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \, v_{cl}^2 \mathrm{e}^{-\beta\varepsilon_c(\boldsymbol{k})} = \frac{k_B T}{m_l} \, n_e$$
$$L_{1c} = \left[\beta(E_{c0} - \mu) + \tfrac{5}{2}\right] L_{0c}$$

Identical expressions can be found for the valence band.

Several interesting results are apparent from these expressions:

- The ratio L_1/L_0 has a simple form

$$\frac{L_{1c}}{L_{0c}} = \frac{E_{c0} - \mu + (5/2)k_B T}{k_B T}$$

  Since $k_B T$ is a small energy, the right-hand side of this expression can be a large number such as 10 or 100. In this case, the Seebeck coefficient becomes very large.

- From the expression for the electron density, one can derive

$$\beta[\mu - E_{c0}] = -\ln(n_e) + f(T),$$
$$f(T) = \ln(N_c/\lambda_c^3)$$

  The Seebeck coefficient $S \sim (k_B/q) L_1/L_0 \sim -(k_B/e)\ln(n_e)$ is proportional to the logarithm of the density of electrons. This behavior is found experimentally.

- Similar behavior is found in *p*-type material, with the subscript c replaced by v for the valence band.

The thermal excitations of electrons, across the bandgap of the insulator, makes electron carriers in the conduction band, and hole carriers in the valence bands. Holes are the conducting states of the valence band: they have a positive mass and a positive charge. The Seebeck has to be calculated using formulas that include two conducting channels: electrons and holes

$$\sigma_T = \sigma_c + \sigma_v$$
$$S_T = \frac{\sigma_c S_c + \sigma_v S_v}{\sigma_c + \sigma_v}$$

Using the above formulas for σ_j and S_j gives the effective Seebeck coefficient

$$S_T = \frac{k_B}{q^2}\left[\frac{q_e L_{1c} + q_h L_{1v}}{L_{0c} + L_{0v}}\right]$$
$$= -\frac{k_B}{|q|}\left[\frac{L_{1c} - L_{1v}}{L_{0c} + L_{0v}}\right]$$

where $q_e = -q = -q_h$. It is useful to define the average value of L_{0j} and the difference

$$\bar{L}_0 = \tfrac{1}{2}(L_{0c} + L_{0v}), \qquad \delta L_0 = L_{c0} - L_{v0}$$
$$L_{c0} = \bar{L}_0 + \tfrac{1}{2}\,\delta L_0, \qquad L_{v0} = \bar{L}_0 - \tfrac{1}{2}\,\delta L_0$$

The effective Seebeck is now

$$S_T = -\frac{k_B}{2|q|}\left[\frac{E_g}{k_B T} + \frac{\delta L_0}{2\bar{L}_0}\frac{E_{c0} + E_{v0} - 2\mu}{k_B T}\right]$$

The first term on the right contains the energy gap $E_g = E_{c0} - E_{v0}$ divided by $k_B T$. This term is very large for most insulators. It is called the "bipolar term." It is responsible for the large value of the Seebeck coefficient in insulators.

The second term is much smaller. Usually $\delta L_0 < 2\bar{L}_0$. Divide the formula for the electron density, by the similar expression for the hole density,

and then take the logarithm

$$n_e = \frac{N_c}{\lambda_c^3} e^{\beta[\mu - E_{c0}]}, \quad n_h = \frac{N_v}{\lambda_v^3} e^{-\beta[\mu - E_{v0}]}$$

$$\frac{2\mu - E_{c0} - E_{v0}}{k_B T} = \ln\left[\frac{n_e N_v \lambda_c^3}{n_h N_c \lambda_v^3}\right]$$

Since $\beta[2\mu - E_{c0} - E_{v0}]$ is the logarithm of something, it never gets too large. It vanishes if μ is exactly in the middle of the bandgap.

The bipolar term has the following interpretation. A thermal fluctuation at one end of the material provides enough energy to excite an electron across the energy gap, and make an electron–hole pair. The pair diffuses to the other end of the material, where they recombine. The process carries the amount of heat E_g from one end of the material to the other.

The Seebeck coefficient of an insulator is a relatively large number. It is usually expressed in millivolts per degree, rather than microvolts per degree. If a material with a large Seebeck coefficient is required, there is a salt shaker on your kitchen table.

Semiconductors

Semiconductors are defined here as crystals with a small energy gap between the lowest conduction band and the highest occupied valence band. Actual semiconductors can have energy gaps from zero up to several electronvolts. The best thermoelectric materials have gaps less than 0.30 eV. Intrinsic semiconductors are those with negligible concentration of defects and impurities. They act as insulators, and typically have a high value for the Seebeck coefficient. **Figure 2** shows $S(T)$ for high-purity p-type germanium as measured by T H Geballe and G W Hull. The scale is milli-volts per degree, which is a thousand times larger than the values in **Figure 1**. The peak at very low temperature is due to phonon-drag. They do not measure below 20 K.

Thermoelectric devices must be good conductors. High values of σ are achieved in semiconductors by adding impurites that act as donors of electrons, or acceptors that make holes. Typically, the concentration of such defects is $n \sim 10^{19}\,\mathrm{cm}^{-3}$. At this concentration, at room temperature, the chemical potential $\mu(T)$ is rather close in energy to the band edge. In n-type material, it is close to the conduction band edge, and varies with temperature.

In the case of metals and insulators discussed above, accurate analytical approximations are derived in two limits. (1) For metals, $\mu \gg k_B T$, and $D(E)$ could be accurately expanded in a power series around $E \approx \mu$. (2) Insulators are in the Maxwell–Boltzmann limit, where the chemical potential is well below the conduction band edge. Neither of these two limits applies to the present problem. In this case, numerical results are obtained on the computer.

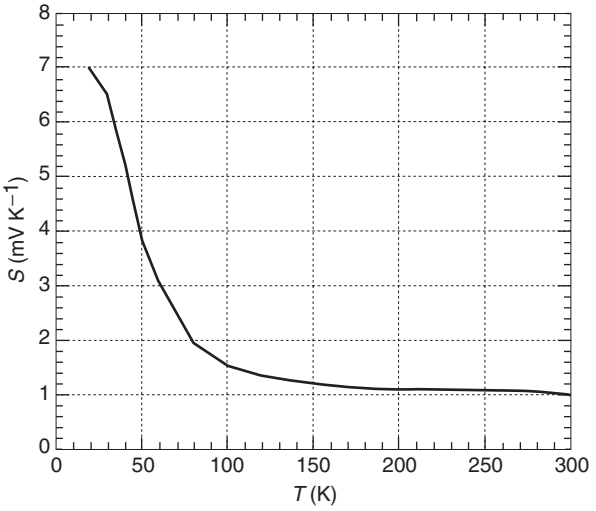


Figure 2 Seebeck coefficient for high-purity p-type germanium as measured by Geballe and Hull. Note the values are in millivolts per degree.

Consider the case in n-type semiconductors that the chemical potential is slightly higher in energy than the conduction band edge E_{c0}. Near the band edge, the kinetic energy is approximately quadratic. Choose $v_{jm} = v_{cz}$

$$E_c(\boldsymbol{k}) = E_{c0} + \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z}\right)$$

$$D_z(E) = \int \frac{d^3k}{(2\pi)^3} v_{cz}^2 \delta[E_c(\boldsymbol{k}) - E]$$

$$= D_0(E - E_{c0})^{3/2}, \quad D_0 = \frac{\sqrt{2m_x m_y m_z}}{3m_z \pi^2 \hbar^3}$$

The density of states for parabolic bands goes as $\sqrt{E - E_{c0}}$. $D(E)$ has a power of 3/2 due to the extra factor of v_{cz}^2 in the integrand. This form for $D(E)$ can be used to calculate the two integrals L_j that are needed for the Seebeck coefficient:

$$x = \frac{E - \mu}{k_B T}, \quad x_0 = \frac{\mu - E_{c0}}{k_B T}$$

$$L_j = \frac{D_0}{2}(k_B T)^{3/2} \int_{-x_0}^{\infty} dx \, \frac{x^j (x + x_0)^{3/2}}{1 + \cosh(x)}$$

Figure 3 shows a graph of L_1/L_0 as a function of x_0 determined by evaluating the integrals on the computer. At $x_0 = 0$, it has the value of 2.833. The

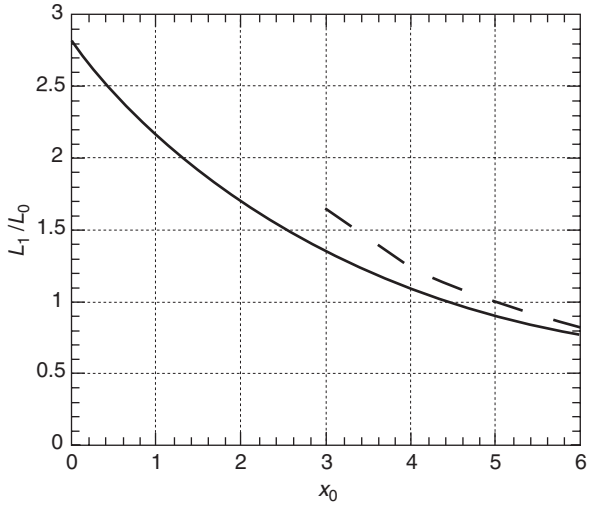


Figure 3 Solid line is the ratio L_1/L_0 when the chemical potential μ is near a band edge E_{c0}. The horizontal axis is $x_0 = (\mu - E_{c0})/k_B T$. Dashed line is the asymptotic expansion at large x_0.

Seebeck is obtained by the ratio of two fundamental constants $k_B/q = 86.2\,\mu V\,K^{-1}$. This gives $S = 244\,\mu V\,K^{-1}$. This high value is typical for the best thermoelectric materials. Their high value is achieved by having the chemical potential near the band edge at a temperature useful for device applications.

At large values of the dimensionless parameter x_0, the above integrals can be evaluated analytically by approximating $(x + x_0)^{3/2} \approx x_0^{3/2} + (3/2)x\sqrt{x_0}$ which gives the ratio

$$\lim_{x_0 \gg 1} \frac{L_1}{L_0} = \frac{\pi^2}{2x_0}[1 + O(1/x_0)]$$

$$\lim_{x_0 \gg 1} S = \frac{\pi^2}{2} \frac{k_B^2 T}{q(\mu - E_{c0})}$$

In this case, the formula derived for metals is found, with $r = 3/2$. This asymptotic expansion is graphed in **Figure 3** as the dashed line.

The largest Seebeck coefficient for a semiconductor is found when it is an insulator. It is very pure, and has a low concentration of defects. Then, the chemical potential is in the gap between the conduction and valence bands. In this case, it has a very small value for the electrical conductivity, and the pure material is not useful for devices. Impurities are intentionally added to the semiconductor to donate electrons to the conduction band, or holes to the valence band. The electrical conductivity increases, and the Seebeck decreases. The best value of the "Power factor" σS^2 is when the chemical potential is near in energy to the minimum of the energy band. The power factor is the numerator in the figure of merit Z mentioned in the "Introduction."

Phonon Drag

If a material is very pure, and free from defects, it is called "intrinsic." The thermopower of an intrinsic material always shows phonon drag at low temperatures. This phenomena happens in all dimensions. A complete theory was derived by Bailyn. It is complicated and will not be reproduced here. Instead, a physical description of the origins of this important contribution is given.

When solving the Boltzmann equation for the lifetime of an electron, the important contributions are the scattering by: defects, phonons, and boundaries. When solving the Boltzmann equation for the lifetime of the phonons, the important contributions to the scattering are from: defects, phonons, electrons, and boundaries. In an intrinsic material, scattering by defects is omitted. In large systems, edge effects and scattering by boundaries are ignored.

Phonon scattering by phonons is caused by anharmonic effects. One phonon divides into two, or two combine into one. The rate of this process increases with temperature, since the number of phonons in the system increases. At high temperature, where anharmonic effects are important, the energy dissipation path of a material in an electric field is:

- The electrons are accelerated by the electric field.
- As they accelerate, they emit phonons, thereby transferring the energy gain from the electric field into phonon energy.
- The phonons dissipate their energy through anharmonic interactions, so the energy from the electric field becomes part of the thermal bath.

At low temperatures, the anharmonic phenomena is ineffective. There are relatively few other phonons thermally excited, and three-phonon processes become rare. Then the energy dissipation path is:

- The electrons are accelerated by the electric field.
- As they accelerate, they emit phonons, thereby transferring the energy gain from the electric field into phonon energy.
- The phonons can only dissipate their energy by giving it back to the electrons.

There is a bottleneck in the dissipation of energy. The electrons and phonons form a closed system that exchange energy back and forth. As the electrons drift along in the electric field, the phonons drift along with them. The amount of energy carried by the combined system is quite large. The Seebeck is a measure of the energy content of the current, and attains high values.

Phonon drag does not affect the electrical resistivity. σ is only affected by the rate that the electrons and phonons scatter from each other. That rate is not affected by phonon drag and there is no peak in the conductivity or resistivity. The thermopower is the only measurement that shows the peak from phonon drag.

The theory of phonon drag includes electrons and phonons. The Boltzmann equations for electrons and phonons become coupled. The coupled equations are usually solved by a variational technique.

In a crystal with defects or impurities, both electrons and phonons can dissipate momentum by scattering from defects at all temperatures. The phonon drag peak is absent in impure materials.

See also: Elemental Semiconductors, Electronic States of; Semiconductor Compounds and Alloys, Electronic States of; Thermoelectric and Energy Conversion Devices.

PACS: 72.15.Jf; 72.20.Pa

Further Reading

Bailyn M (1958) Transport in metals: effect of the nonequilibrium phonons. *Physical Review* 112: 1587–1598.

Blatt FJ, Schroeder PA, Foiles CL, and Greig D (1976) *Thermoelectric Power of Metals*. New York: Plenum.

Goldsmid HJ (1986) *Electronic Refrigeration*. London: Pion Ltd.

Mahan GD (1998) Good thermoelectrics. In: Ehrenreich H and Spaepen F (ed.) *Solid State Physics*, vol. 51, ch. 2. New York: Academic Press.

Nolas GS, Sharp J, and Goldsmid HJ (2001) *Thermoelectrics*. New York: Springer.

Rowe DM and Bhandari CM (1983) *Modern Thermoelectrics*. Reston, VA: Reston Pub.

Rowe DM (ed.) (1995) *CRC Handbook of Thermoelectrics*. Boca Raton: CRC Press.

Tritt TM (2001) Recent trends in thermoelectric materials research I, II, III. In: Willardson RK and Weber ER (eds.) *Semiconductors and Semimetals*, vol. 69–71. New York: Academic Press.

Nomenclature

$D_j(E)$	transport distribution function (same as L_j)
E_j, ε_j	particle energies (J)
f_j, δf_j	distribution functions
$\hbar$	Planck's constant (J s)
J	current density ($\mathrm{A\,m^{-2}}$)
k_B	Boltzmann's constant ($\mathrm{J\,K^{-1}}$)
K	thermal conductivity ($\mathrm{W\,m^{-1}\,K^{-1}}$)
L_1, L_0, L_j	integrals defining thermopower
N_j, N_c, N_v	band degeneracy
q, q_j	charge of particle (C)
Q	heat (J)
S	Seebeck coefficient ($\mathrm{V\,K^{-1}}$)
T	temperature (K)
v_{jm}	velocity of particle ($\mathrm{m\,s^{-1}}$)
V	voltage (V)
Z	figure of merit ($\mathrm{K^{-1}}$)
β	$1/k_BT$ ($\mathrm{J^{-1}}$)
μ	chemical potential (J)
Π	Peltier coefficient (V)
ρ	resistivity (Ω m)
σ	electrical conductivity ($\mathrm{S\,m^{-1}}$)
τ	Thomson coefficient ($\mathrm{V\,K^{-1}}$)
τ_j	particle lifetime (s)

Thin Films, Mechanical Behavior of

S P Baker, Cornell University, Ithaca, NY, USA

Introduction

Materials in the form of thin films (layers having thicknesses of a few micrometers or less) are the basic building blocks of microfabricated or nanofabricated devices such as integrated microelectronic circuits. They are used as electrical conductors and insulators, magnetic layers, optical reflectors and absorbers, chemical passivations and catalysts, and as structural elements. Although often intended for nonmechanical functions, the mechanical properties of thin films are very important. Thin films are typically found attached to much more massive substrates, and the constraint of the substrate can lead to very high stresses in the film. These stresses may, in turn, lead to failure by a variety of mechanisms, including excessive elastic or plastic deformation, void formation, fracture, or delamination. Understanding the stress states in, and the mechanical response of, thin films is thus a prerequisite to optimizing reliability in microfabricated and nanofabricated devices.

Achieving this goal is complicated by the fact that the mechanical behavior of thin films often does not follow the scaling laws developed for bulk materials. Thus, the mechanical properties of thin films must be measured directly. In this article, the sources of stress in thin films, common mechanical property measurements, and some factors that determine non-bulk-like mechanical response in thin films are briefly reviewed.

Thin Film Stresses

While thin films may experience stresses due to externally applied loads, the largest source of stress in a film typically arises from an interaction between the film and substrate. The origins of these stresses can be understood by the simple thought experiment shown in **Figure 1**. In **Figure 1a**, a stress-free film is shown attached to a substrate. Because the film is stress free, one can imagine that it is removed from its substrate without changing dimensions as shown in **Figure 1b**. One can then imagine that some process occurs that changes the in-plane dimensions of the film relative to the substrate. For example, suppose the film shrinks relative to the substrate as shown in **Figure 1c**. Since there is no interaction between the film and the substrate, each attains its equilibrium dimension and remains stress free. The film can then be returned to the substrate by imposing a biaxial stress on the detached film in order to elastically deform it, until it again has the same dimension as the substrate, as shown in **Figure 1d**, and reattaching the film to the substrate, as shown in **Figure 1e**. To complete the thought experiment, the externally applied forces are removed. The final configuration is shown in **Figure 1f**. The film continues to support biaxial stresses, and the substrate distorts elastically due to the forces transmitted across the film/substrate interface.

Any process that changes the equilibrium dimensions of the film and the substrate, relative to each other after the film has been deposited on the substrate, will lead to substrate interaction stresses. These dimensional changes can arise from four different sources: (1) Thermal strains arise due to differential thermal expansion. If the elastic strain in the film is zero at some temperature, T_0, then the elastic strain required in the film to accommodate the thermal strains at any other temperature, T, is

$$\varepsilon_{\text{th}} = \int_{T_0}^{T} (\alpha_{\text{s}}(T) - \alpha_{\text{f}}(T))\, \mathrm{d}T \qquad [1]$$

where $\alpha_{\text{f}}(T)$ and $\alpha_{\text{s}}(T)$ are the (temperature dependent) thermal expansion coefficients of the film and substrate, respectively. (2) Structure evolution strains arise from structural changes that change the density of the film. Grain growth, densification, ion implantation, chemical reactions, and phase changes are common sources. The accommodation strain in such cases is given by

$$\varepsilon_{\text{evol}} = \frac{-e_{\text{T}}}{3} \qquad [2]$$

where e_{T} is the dilatational strain associated with the structure evolution. (3) Epitaxial strains arise when the film is deposited in such a way that there is some crystallographic registry between the film and substrate. If the pertinent in-plane lattice parameters, a_{s} and a_{f}, of the film and substrate, respectively, are not the same, then the elastic strain required in the film is

$$\varepsilon_{\text{epi}} = \frac{a_{\text{s}} - a_{\text{f}}}{a_{\text{f}}} \qquad [3]$$

(4) Surface (interface) stresses arise from the change in bonding at free surfaces and interfaces compared

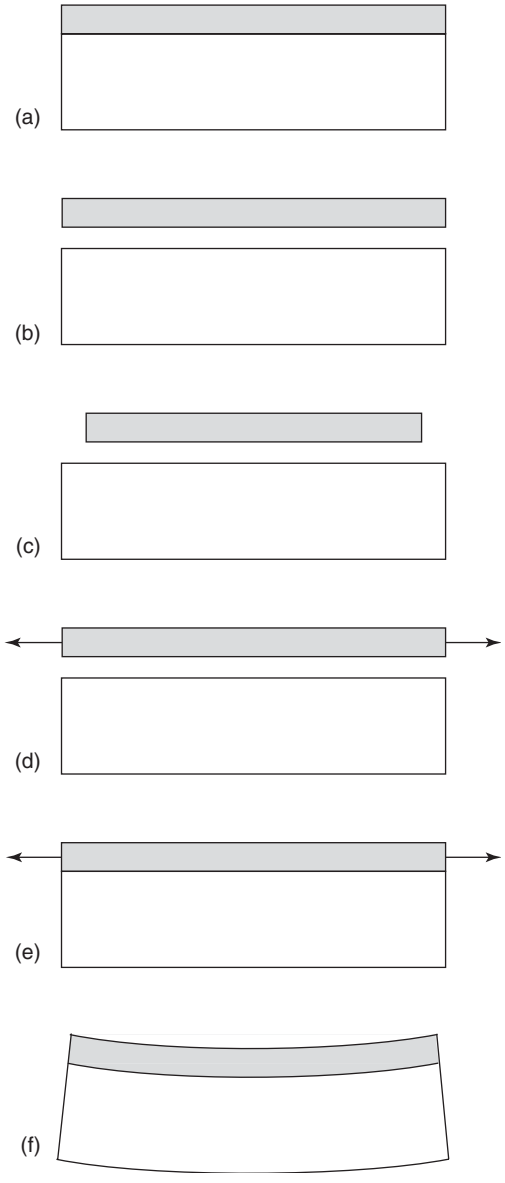


Figure 1 Thought experiment showing how an interaction between a film and substrate leads to high stresses in the film (description in text). Any process that changes the equilibrium in-plane dimension of the film relative to the substrate after the film is attached to the substrate leads to high stresses in the film.

to that in the bulk. While the forces arising from a single interface are small, surface stresses can be significant in multilayer films where the interface density is high. For a multilayer consisting of alternating layers of materials A and B having thicknesses t_A and t_B, respectively, the total stress exerted on the substrate is

$$\sigma_{int} = \frac{2(f_{A/B} + f_{B/A})}{t_A + t_B} \quad [4]$$

where $f_{A/B}$ is the interface force per unit width of A at an A/B interface, and $f_{B/A}$ is the interface force per unit width of B at a B/A interface.

Substrate interaction stresses, particularly those arising from structure evolution strains, are often referred to as "intrinsic," but this terminology should be avoided since the stresses are not intrinsic to either the film or the substrate but rather always arise from an interaction between the two. Because the structure in films can be inhomogeneous, the stresses in films can be inhomogeneous as well.

Measurements of Thin Film Deformation

Measuring mechanical properties in thin films is challenging because of the difficulty in manipulating such a thin object. This problem has been overcome by testing films while they are still attached to their substrates or by using microfabrication or nanofabrication methods to prepare samples without the need to handle the film.

The most common tests are nanoindentation and substrate curvature methods, in which the film is tested while still supported by the substrate. Commercial equipment is available for both of these tests. In nanoindentation, the load and displacement are recorded as a hard tip of a known shape is pressed into the sample surface and removed. By the use of an appropriate analysis model, a measure of elastic stiffness, hardness, and time-dependent behavior can be obtained. Nanoindentation tests are simple to perform, and quantitative information can be obtained from very small (tens of nanometers in diameter), well-located (within some tens of nanometers) volumes of materials. However, the primary disadvantage of nanoindentation lies in distinguishing film properties from substrate influences, which arise as indentations become large and surface influences (mainly due to roughness, surface layers, or particles), which become large as indentations become small.

Substrate curvature tests make use of the curvature shown in **Figure 1f**. If the film is thin and/or compliant relative to the substrate, then the stress in the film is related to the curvature induced via

$$\sigma = Y_s \frac{t_s^2}{6t_f}\left(\frac{1}{R} - \frac{1}{R_0}\right) \quad [5]$$

where R_0 and R are the radii of curvature of the substrate before and after the film is attached, respectively, t_f and t_s are the thicknesses of the film and substrate, respectively, and $Y_s = E_s/(1 - \nu_s^2)$ is the biaxial modulus of the substrate, assumed to be isotropic with Young's modulus, E_s, and Poisson's ratio, ν_s. Equation [5] is known as the Stoney equation and applies when the product of ratios $t_s Y_s / t_f Y_f$ is large, typically greater than ~ 100. R_0 and R are typically measured either by scanning a laser across the sample and observing the position of the reflected beam, or using optical interferometry. The substrate curvature method is very accurate and repeatable, as well as relatively simple and flexible. It is often used *in situ* to measure the stresses that arise in films during deposition or thermal cycling. The main disadvantage is that it is not possible to control the strain without changing the conditions that give rise to that strain (e.g., change in temperature).

Results from substrate curvature measurements conducted *in situ* during thermal cycling of 1 μm thick Cu films are shown in **Figure 2**, and provide an introduction to several unique features of thin film deformation behavior. Such tests are often used because they mimic the thermal cycles that films undergo during manufacture and use. Three sets of data are shown. The solid line indicates data obtained from a computer simulation based on steady-state deformation mechanisms for bulk Cu, including the microstructural length scales of the film. The open and filled circles are data from substrate curvature measurements. The Cu films were prepared identically except that a thin passivation layer was deposited on one, while the other was left with the Cu-free surface exposed. The shapes of the experimental stress–temperature hystereses in **Figure 2** show that the stresses can be much higher in thin films than in comparable bulk materials, particularly at high temperatures.

In recent years, a number of tests have been developed to test free-standing films or parts of films. The two most common examples are bulge tests and microtensile tests. In bulge tests, microfabrication methods are used to remove a section of the substrate from behind a film, and pressure is applied to cause the film to bulge out. With an appropriate analysis model, pressure–deflection data can be converted to stress–strain data. For microtensile tests, free-standing tensile test samples are prepared by first removing the adjacent film material and then

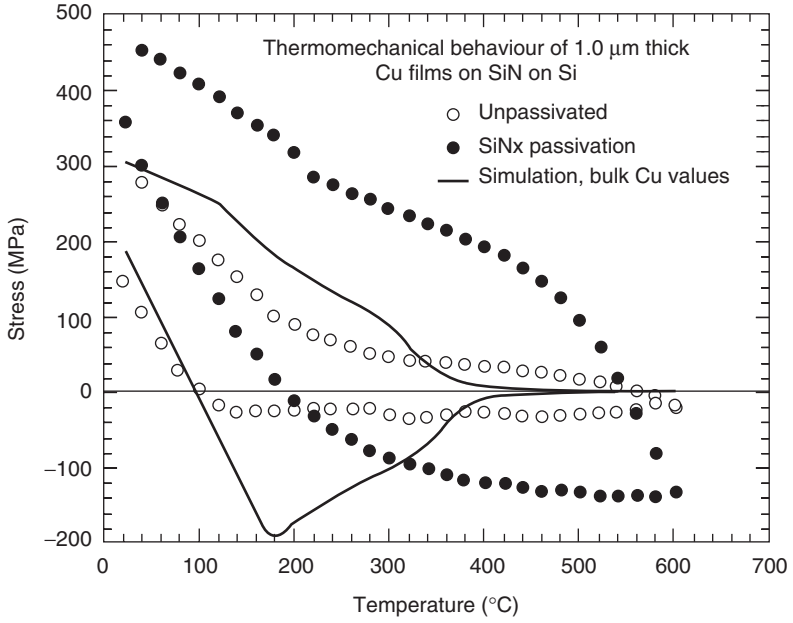


Figure 2 Stresses induced by thermal cycling in Cu films on Si substrates. The films were prepared identically except that a thin SiNx passivation layer was deposited on one. The solid line is a simulation showing the behavior expected based on scaling laws derived for bulk Cu.

removing the substrate from behind the sample, leaving a substrate frame to which the sample remains attached at both ends. Loads are applied to these substrate sections to distort the film. Microfabricated test structures are of great interest for the study of free-standing thin film structures.

Plastic Behavior

As is the case in bulk materials, the plastic behavior of thin films depends on the microstructure. Unlike bulk materials, however, the plastic behavior of thin films is often found to depend on the sample size. In addition, the plastic behavior of thin films is extremely sensitive to conditions at thin film interfaces.

The microstructures of crystalline thin films may be categorized into three groups: Single-crystal films arise primarily when films are deposited epitaxially. Nonepitaxial films are often nanocrystalline (grain sizes $\sim 10\,\text{nm}$) in their as-deposited state. The mechanical behavior of such films is similar to that of "bulk" nanocrystalline materials and is not considered here. If a fine-grained film is annealed, the mean final grain size is often found to be of the same order as the film thickness. Such films typically have "columnar" grain structure (all grain boundaries perpendicular to the plane of the film), in which certain orientations are strongly preferred. For example, in f.c.c. metal films, grains with (1 1 1) and (1 0 0) planes parallel to the plane of the film are most common as they minimize interface and strain energy, respectively.

Dislocation Mediated Plasticity

The high strength of crystalline thin films at low temperatures (**Figure 2**) is due to constraints on dislocation motion. An epitaxial thin film on a substrate provides a simple example of how the constraint of the substrate affects the mechanical behavior of thin films. A film with a capping layer is illustrated in **Figure 3**. A threading dislocation (a dislocation that runs from the top to the bottom of the film) is shown at "**a**" of **Figure 3** in an unstressed film. When a stress is applied, the threading dislocation can relax the applied strain by moving on its glide plane. However, because the film is encapsulated between the substrate and cap layer, the threading dislocation can only move ahead by leaving misfit dislocations behind at the interfaces (indicated by "**b**" in **Figure 3**). The threading dislocation moves forward only when the elastic strain energy relieved by its motion is sufficient to provide the energy required to form the misfit dislocation. Thus, there is a threshold stress (critical stress) for dislocation motion. Since the total strain energy in the film depends linearly on the film thickness and the energy per unit length of the misfit dislocation is approximately constant, the critical stress varies with reciprocal film thickness. A widely used derivation gives

$$\sigma_c = \frac{\sin\phi}{\sin\phi\cos\lambda}\frac{b}{2\pi(1-\nu)t_f} \times \left[\frac{\mu_f\mu_s}{(\mu_f+\mu_s)}\ln\left(\frac{\beta_s t}{b}\right) + \frac{\mu_f\mu_p}{(\mu_f+\mu_p)}\ln\left(\frac{\beta_p t_p}{b}\right)\right] \quad [6]$$

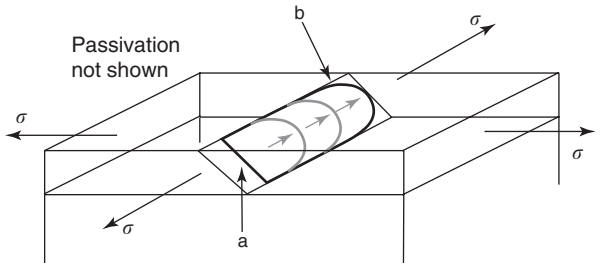


Figure 3 A dislocation gliding on its slip plane in the channel created by a thin film between a substrate and a passivation layer. (a) A threading dislocation extends from the top of the film to the bottom at zero stress. (b) In order to move through the film, the dislocation must leave misfit dislocations at the film/substrate and film/passivation interfaces.

where b is the magnitude of the Burgers vector, ϕ and λ are the angles between the film normal and the glide plane normal, respectively, μ the shear modulus, and β a constant. The subscripts f, s, and p refer to the substrate, film, and passivation layers, respectively.

The concept of critical strain is particularly important in microelectronic devices containing epitaxial semiconductor layers, as the misfit dislocations that form when the critical strain is exceeded can spoil the electronic function of the device. The reciprocal dependence of critical stress on the film thickness predicted by eqn [6] has been observed in epitaxial semiconductor films. However, higher values have also been reported. This may be due to experimental difficulties associated with observing the formation of the first misfit dislocations, or to a lack of appropriate dislocation sources. From a reliability perspective, eqn [6] provides an accurate measure of the minimum thickness at which misfit dislocations may appear in an epitaxial film.

Interestingly, an inverse dependence of strength on thickness is also seen in polycrystalline metal films, although at stress levels much higher than those predicted by eqn [6]. One can obtain such data, for example, by recording the stress at room temperature at the end of thermal cycles as in **Figure 2** for films of various thicknesses. Neither the fact that the yield stress increases as t_f decreases, nor that eqn [6] underpredicts the result, is surprising. As strain increases, a threading dislocation moving through the film must interact with misfit dislocations in its path as well as other threading dislocations and is constrained from long-range motion by grain boundaries, just as in bulk materials. That these events occur in the narrow channel of the film leads to very high stresses indeed. However, while a number of models have been presented, the precise origin of the t_f^{-1} dependence is not yet clear.

The arguments above are based on the assumption that dislocations are preserved at the interfaces. If dislocations are not preserved, then the behavior will be very different. The limiting case is that of a free-standing film, which has no critical strain threshold for dislocation motion and no misfit dislocations to block the threading dislocation motion. A film with one free surface (e.g., **Figure 2**) represents an intermediate case. However, a free surface is not needed to eliminate misfit dislocations. If the resistance to interface sliding is sufficiently low, then the cores of misfit dislocations can spread into the interface. The consequence is that the mechanical behavior of certain film systems (e.g., copper on nitride or oxide layers, as is commonly used in microelectronics) is very sensitive to any variation in interface chemistry that affects interface strength. Stress levels in fully encapsulated films can be varied over a wide range in this manner.

Another example of interface control of properties is multilayer films, in which manipulation of interface strength and density can be used to create exceptionally strong films. For example, by deposition of alternating thin layers with well-bonded, dislocation blocking interfaces, films with strengths approaching the theoretical strength can be generated.

Diffusion Mediated Plasticity

At high temperatures, the high strength of thin films is due to constraints on diffusion, and the difference between thin film behavior and that expected on the basis of bulk scaling laws is even more dramatic. For example, at the low strain rates imposed for the data in **Figure 2**, the simulation predicts that the stress would be relaxed to near zero at temperatures above 400°C, while the real films support significant stresses at these temperatures.

This discrepancy can be understood as follows: **Figure 4** shows a schematic cross section of a film on the substrate with a capping layer. A stress applied in the plane of the film can be relaxed if the material diffuses between the grain boundaries and the interfaces as shown in **Figure 4a**. The simulation in **Figure 2** was obtained by assuming that the interface diffusivity is the same as the grain boundary diffusivity. In order to achieve the stresses seen in the real capped film at high temperatures, interface diffusion must be turned completely off in the simulation and relaxation allowed to occur by other mechanisms, such as thermally activated dislocation glide. While high stresses due to limited interface diffusivity might be expected in tightly bound interfaces such as Al/SiO_2, similar high stresses are often seen in films with interfaces known to have poor adhesion and high diffusivity such as Cu/SiN_x. The reasons for this are presently unknown.

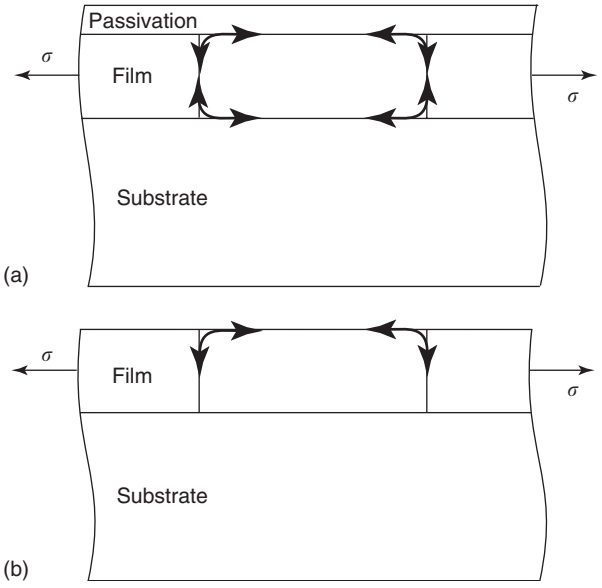


Figure 4 (a) Stresses can be relaxed in a passivated film only if diffusion is allowed along the interfaces. (b) In an unpassivated film, stresses can be relaxed by diffusion between the grain boundaries and the free surface.

An uncapped film provides an interesting case, as shown in **Figure 4b**. Here diffusion is allowed between the grain boundaries and the free surface, but not between the grain boundaries and the interface with the substrate. Tensile stresses, for example, are relaxed by diffusion of material from the free surface into the grain boundaries. Each grain boundary is filled with a wedge of material and the stresses are relaxed near the film surface, but not near the interface. This provides an inhomogeneous stress state with high shear stresses on the interface.

Elastic and Anelastic Behavior

Elastic strain is deformation that is fully recovered upon removal of the applied load. This definition encompasses both elastic deformation that arises from bond stretching and twisting, and anelastic deformation that arises from atomic reconfigurations (e.g., defect motions). Elastic deformation occurs at short timescales (transmitted at phonon velocities) while anelastic deformation is time dependent on a much longer scale.

Elastic deformation depends on the type and density of interatomic bonds in the material. One does not expect radically new types of interatomic bonds in a film just because it is thin. Indeed, the changes that can be expected in elastic constants should be of the order of $N_{\text{defect}}/N_{\text{bulk}}$, the ratio of atoms in nonbulk environments to the number in bulk environments. For example, consider a film comprised of cubic grains with side d having boundaries of width δ. The ratio of grain boundary to interior volume is $3\delta/d$. If the grain interiors have modulus E_{grain} and the boundaries have a modulus of zero (an extreme example!), then the isostrain composite modulus of the film is

$$E = (1 - 3\delta/d)E_{\text{grain}} \qquad [7]$$

For $\delta = 0.2$ nm, one would need $d = 6$ nm$(= t_f)$ to achieve a modulus reduction of only 10%. Given this expectation, a great deal of excitement has accompanied reports of much larger deviations. However, it now appears that all such reports can be attributed to experimental errors.

Anelastic behavior, on the other hand, can be much more prominent in thin films than in bulk. An example can be seen in **Figure 2**. Upon initial heating, the experimental data initially follow the predicted thermoelastic slope. At some point, the data deviate from this slope, indicating the onset of inelastic deformation. Careful inspection reveals that the inelastic strain is compressive, while the applied stress is still tensile. This "negative yielding" is similar to the well-known Bauschinger effect in bulk metals, but is much larger. As is the case with plastic deformation by both dislocation and diffusion mediated plasticity, anelastic recovery is also very sensitive to interface conditions. For example, quite spectacular anelastic recoveries of up to several tenths of a percent strain can be achieved by adjusting interface adhesion.

Two mechanisms, dislocation motion and boundary sliding, that are often used to explain anelastic behavior in bulk metals, probably also account for anelastic behavior in films. For example, consider a film that has been plastically deformed so that loops having the form shown in **Figure 3** have been generated at stress levels above σ_{ch} (eqn [6]). If the applied stress is then reduced below σ_{ch}, it becomes energetically favorable for the dislocation loops to run out of the film, increasing the strain energy (and stress) in the film, but reducing the total energy by reducing misfit dislocation line length. The quasi-2D structure of films can promote grain boundary sliding by allowing grain rotations in the plane of the film. In addition, shear stresses along the film/substrate interface generated by anisotropy or diffusional mechanisms can be relaxed by sliding along the interface, which can also lead to anelastic behavior.

The highly textured nature of films also leads to non-bulk-like scaling behavior. For example, a film composed of a cubic material has lower stiffness, but higher surface energy when (1 0 0) planes are parallel to the plane of the film compared to the case where (1 1 1) planes are parallel to the film plane. Thus, a

transition from (1 1 1) to (1 0 0) orientation (stiff to compliant) with film thickness has been predicted and experimentally verified. For Cu, Ag, and Au, for example, the difference is more than a factor 2. For a given elastic strain, the stresses are then much higher in the (1 1 1) orientations, leading to very inhomogeneous stress states and high shear stresses on the interfaces near (1 1 1)/(1 0 0) grain boundaries.

Fracture and Delamination

Brittle films with sufficiently weak interfaces may fail by delamination from the substrate (**Figure 5a**) or by fracture (**Figure 5b**). The driving force, G, for delamination can be found by comparing the elastic strain energy per unit area of the film far ahead of the delamination crack with that which is far behind the crack. For a straight crack front,

$$G = \frac{1 - \nu_f^2}{2E_f}\sigma^2 t_f \qquad [8]$$

where σ is the stress in the film before delamination.

Several tests have been developed for quantitative assessment of the interfacial fracture toughness, Γ, or the value of G at delamination. If the stress in the film is well known, the thickness at which it delaminates can be used to determine Γ from eqn [8]. In many cases, a highly stressed "superlayer" is added on top of the target film of interest to accelerate the process.

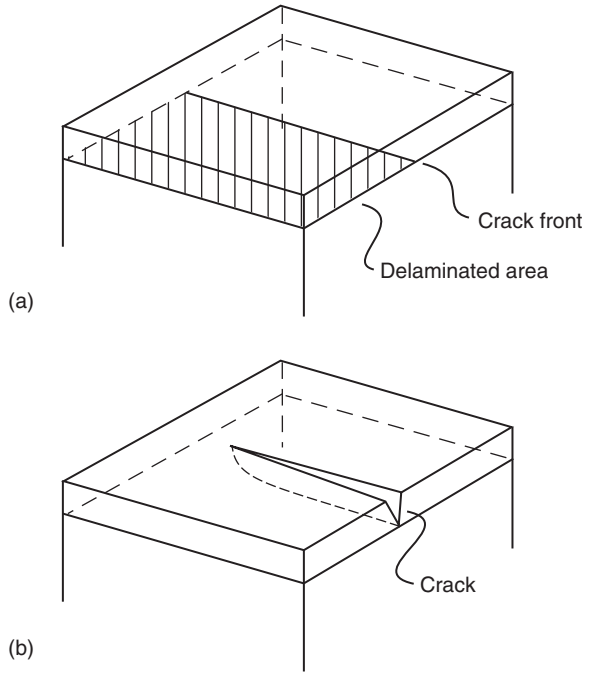


Figure 5 (a) Delamination of, and (b) through thickness (channeling) cracks in a thin film.

While the solution is more complicated, it is straightforward to find Γ for the target interface if the stresses in the layers and their thicknesses are known. In addition, several tests of film adhesion have been developed, which do not require a detailed knowledge of film stress states. In these tests, the film is bonded between two massive substrates and external loads are applied to the substrates to drive a delamination crack. These tests include the four-point bend test, the double-cantilever beam test, and several others.

Because thin films are normally under stress, subcritical delamination is an important failure mode and topic of study. Subcritical delamination occurs when chemical processes assist crack propagation so that films delaminate under conditions where G is less than the value required for rapid or "critical" crack propagation in the absence of chemical assistance. Under constant stress, subcritical delamination may occur very slowly over very long durations and is a particularly insidious failure mode. Recent work suggests that subcritical delamination of thin films is similar to subcritical cracking in glass with a region where crack velocity v depends exponentially on G when the chemical reaction rate is controlling, and a region at higher G where v is independent of G when diffusion of the chemical species to the crack tip is controlling.

Cracks that extend through the film thickness (**Figure 5b**), also known as channeling cracks, are similar to channeling dislocations (**Figure 3**) in that there is a critical stress needed for channeling crack propagation, established by the requirement that the strain energy released by the crack propagation must be sufficient to produce the new surface area and other energy costs associated with the crack. Channeling cracks, such as channeling dislocations, can also form arrays to relax strains and form configurations that depend on both interactions and sources. The mechanics of film cracking are well developed.

Conclusions

Films on substrates normally support very high stresses due to the constraint of the substrate, but do not exhibit the mechanical behavior of chemically equivalent bulk materials, and thus must be separately studied. While a number of dislocation- and diffusion-based mechanistic models for film deformation have been proposed, a complete model capable of predicting the plastic and anelastic behavior of a film based on its microstructure remains elusive. A significant factor in this is that the microstructural configurations that are assumed in such models have not been sufficiently well verified experimentally. In general, thin films follow the rule that "smaller is stronger." This is an area of ongoing development,

and new experimental methods and models are expected to extend a detailed knowledge to smaller length scales and both shorter and longer timescales.

See also: Mechanical Properties: Elastic Behavior; Mechanical Properties: Plastic Behavior; Mechanical Properties: Tensile Properties.

PACS: 62.25.+g; 68.55.−a; 68.55.Jk; 68.60.−p; 68.60.Bs; 68.65.−k

Further Reading

Arzt E (1998) Size effects in materials due to microstructural and dimensional constraints: a comparative review. *Acta Materialia* 46: 5611–5626.

Baker SP (2001) Plastic deformation and strength of materials in small dimensions. *Materials Science and Engineering A* A319–321: 16–23.

Freund LB and Suresh S (2003) *Thin Film Materials: Stress, Defect Formation and Surface Evolution*. Cambridge: Cambridge University Press.

Nix WD (1989) Mechanical properties of thin films. *Metallurgical Transactions* 20A: 2217–2245.

Ohring M (1992) *The Materials Science of Thin Films*. San Diego, CA: Academic Press.

Suo Z (2003) Reliability of interconnect structures. In: Gerberich WW and Yang W (eds.) *Interfacial and Nanoscale Fracture*. Amsterdam: Elsevier.

Vinci RP and Baker SP (eds.) (2002) *Mechanical Properties in Small Dimensions*, MRS Bulletin, vol. 27, no. 1. Warrendale, PA: Materials Research Society.

Vinci RP and Vlassak JJ (1996) Mechanical behavior of thin films. *Annual Review of Materials Science* 2: 431–462.

Tight-Binding Method in Electronic Structure

D A Papaconstantopoulos and M J Mehl, Naval Research Laboratory, Washington, DC, USA

Introduction

In the linear combination of atomic orbitals (LCAO) method, the one-electron wave function is expressed as a linear combination of Bloch sums. When such an expansion is made to the wave function in the Schrödinger equation, one obtains a set of simultaneous linear equations that has a nonzero solution if the determinant of the coefficients vanishes, that is,

$$\tilde{H} - \varepsilon\tilde{S} = 0 \qquad [1]$$

The matrix elements in this equation have the form:

$$H_{nm} = \sum_{R_j} \exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_i)] \times \int \phi_n^*(\boldsymbol{r} - \boldsymbol{R}_i)\tilde{H}\phi_m(\boldsymbol{r} - \boldsymbol{R}_j)\,\mathrm{d}^3r \qquad [2]$$

and

$$S_{nm} = \sum_{R_j} \exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_i)] \times \int \phi_n^*(\boldsymbol{r} - \boldsymbol{R}_i)\phi_m(\boldsymbol{r} - \boldsymbol{R}_j)\,\mathrm{d}^3r \qquad [3]$$

where $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ denote the positions of atoms located on orbitals ϕ_n and ϕ_m, respectively. The integrals in the above equations can, in principle, be calculated directly. However, the most common practice has been to follow Slater and Koster (SK), who suggested replacing these integrals by adjustable parameters that could be estimated from experiment, but which are generally determined by fitting to more elaborate electronic structure calculations. The size of the matrices $\tilde{H}$ and $\tilde{S}$ is determined by the number of atoms in the unit cell and the number of atomic orbitals on each site. So for face-centered cubic (f.c.c.), body-centered cubic (b.c.c.), and simple cubic (s.c.) lattices with one atom per unit cell, $\tilde{H}$ and $\tilde{S}$ are 9×9 matrices representing one *s*-function, three *p*-functions, and five *d*-functions. The *f*-states have been omitted in most works, although there have been papers that provide extensions of the SK scheme that include *f*-orbitals. For diatomic lattices, such as hexagonal close-packed (h.c.p.) and diamond and in binary compounds the *spd* model will result in an 18×18 matrix. Often in semiconductors such as Si and Ge, the *d*-states are not included and hence the size of the Hamiltonian is 8×8. The full SK Hamiltonian for a monatomic material could require the determination of 81 parameters between each pair of atoms; however, SK showed that these parameters are related by symmetry operations that considerably reduce their actual number. The resulting independent matrix elements may be found in the tables given by SK. The reader is cautioned that some typographical errors appear in these tables which have been corrected by other authors (see "Further reading" section). Since the above summations are over the interatomic distances R_j, the number of parameters increases for calculations that include more than first nearest neighbors. In practice, no more than the third nearest neighbors are included. The integrals in [2] are three-centered since they are the product of an

atomic wave function $\phi^*(\boldsymbol{r}-\boldsymbol{R}_i)$ centered on the atom at position $\boldsymbol{R}_i$, an atomic wave function $\phi(\boldsymbol{r}-\boldsymbol{R}_j)$ located on the atom at position $\boldsymbol{R}_j$, and a potential function inside the Hamiltonian H, centered on a third atom. These three-center integrals can be reduced to two-center integrals by approximating the potential energy as a sum of spherical potentials located on the same two atoms where the atomic orbitals are located. For the relationships between three- and two-center integrals, see the "Further reading" section. Although working in three-center formalism has the advantage of fitting band structures more accurately, the two-center approximation has the distinct advantage that its parameters are transferable from one structure to another. This property of the two-center scheme makes it suitable to introduce bond-length dependence in the parameters which leads one to implementing a total energy capability in the SK method.

Since the 1960s, there have been numerous applications of tight-binding (TB) theories, in particular, the SK approach, to calculate electronic energy bands, densities of states, and, more recently, total energies. There are two main uses of the TB theory. One is to determine a small number of parameters with the aim of providing a physical insight into the electronic structure without the necessity of going beyond a qualitative description. A classic example of this approach is the universal TB parameters of Harrison. Harrison's TB theory has been of great educational value for a whole generation of scientists. Very recently, extensions to this theory have been provided that greatly improve its accuracy. The other approach is to place emphasis on accurately reproducing the results of first-principles calculations by applying a least-squares procedure using a large number of parameters. The authors of this article have generally followed the second approach. A handbook by Papaconstantopoulos (see "Further reading") has provided TB parameters for most elements in the periodic table, except for those with *f*-states. The handbook provides TB parameters in both two-center and three-center bases as well as in orthogonal and nonorthogonal representations. However, these parameters are fitted to the ground state of each element and have very limited transferability to other volumes and structures. In addition, these TB Hamiltonians are not designed to compute total energies. The NRL-TB scheme on the other hand, which is discussed in the next section, is a two-center nonorthogonal TB method that uses environment-dependent parameters that capture the volume dependence of both the energy bands and the total energy.

To complete this introduction, the attention of the reader is drawn to a simple level of the TB theory, known as the second moment approximation (SMA) which originates from the Friedel model of rectangular density of states for the *d*-bands. The SMA method does not account for the band structure but, at least for the f.c.c. metals, reproduces well the total energy which has led to its wide use in molecular dynamics simulations. This method contains four adjustable parameters fitted to reproduce either experimental quantities or the total energies from first-principles calculations.

NRL Tight-Binding Method

In the (Naval Research Laboratory) NRL-TB scheme, the on-site terms h_{il} are written as a polynomial:

$$h_{il} = a_l + b_l\rho_i^{2/3} + c_l\rho_i^{4/3} + d_i\rho_i^2 \qquad [4]$$

where i labels the atom and l the angular momentum of s, p, and d characters in the present form of the method.

The quantity ρ_i that appears in [4] is an embedded-atom-like "density" per atom given by the expression

$$\rho_i = \sum \exp(-\lambda^2 R_{ij}) F(R_{ij}) \qquad [5]$$

where R_{ij} denotes the position of neighboring atoms from a central atom i, and $F(R_{ij})$ is the cutoff function. The quantities a_l, b_l, c_l, d_l, and λ are 13 parameters to be determined by a least-squares fit to the first-principles results. It is possible, within this scheme, for example, to split the *d*-onsite parameters to t_{2g} and e_g, or going even further, to have different parameters for the three *p*-orbitals and the five *d*-orbitals. It is also to be noted that this scheme has the flexibility of adding more terms in [4] to describe the density contributions from atom A to atom B (and vice versa) in the case of a binary material.

It was shown by SK that the two-center (*spd*) hopping integrals can be constructed from ten independent parameters $H_{ll'm}$, where

$$(ll'm) = ss\sigma, sp\sigma, pp\sigma, pp\pi, sd\sigma, pd\sigma, pd\pi, dd\sigma, dd\pi, \text{ and } dd\delta$$

One notes that in the case of a binary material AB, there will be a set of four additional parameters:

$$(ll'm) = ps\sigma, ds\sigma, dp\sigma, \text{ and } dp\pi$$

In the NRL-TB, each of these parameters is expressed in the form of a second-order polynomial-times an exponential function:

$$H_{ll'm}(r) = (e_{ll'm} + f_{ll'm}r + g_{ll'm}r^2) \times \exp(-t_{ll'm}^2 r) F(r) \qquad [6]$$

where r is the distance between atoms, $F(r)$ the cutoff function as in [5], and $e_{ll'm}$, $f_{ll'm}$, $g_{ll'm}$, and $t_{ll'm}$ are an additional set of 40 parameters (or 56 for the, A–B interactions) to be determined by fitting to the first-principles results. Since one usually fits to a non-orthogonal Hamiltonian, an additional set of 40 parameters is used (56 for A–B) for the overlap matrix following [6]. In cases where the validity of the parameters for interatomic distances much smaller than those in the original databases needs to be extended, one finds that the following form of overlap parameters is more successful:

$$S_{ll'm}(r) = (\delta_{ll'} + p_{ll'm}r + q_{ll'm}r^2 + r_{ll'm}r^3) \times \exp(-s^2_{ll'm}r)F(r) \quad [7]$$

where $p_{ll'm}$, $q_{ll'm}$, $r_{ll'm}$, and $s_{ll'm}$ are the corresponding parameters of the overlap matrix and $\delta_{ll'}$ is the Kronecker delta.

In most of the TB approaches, as well as in all the so-called "glue" potential atomistic methods, one writes the total energy as a sum of a band energy term (sum of eigenvalues) and a repulsive potential $G[n(r)]$ that can be viewed as replacing all the charge density-dependent terms appearing in the total energy expression of the density-functional theory. The NRL-TB method has the unique feature that eliminates G by the following ansatz: a quantity V_0 is defined as

$$V_0 = G[n(r)]/N_e \quad [8]$$

where N_e is the number of valence electrons and $n(r)$ is the charge density. All the first-principles eigenvalues $\varepsilon_i(\boldsymbol{k})$ are then shifted by the constant V_0, and the shifted eigenvalue is defined as

$$\varepsilon'_i(\boldsymbol{k}) = \varepsilon_i(\boldsymbol{k}) + V_0 \quad [9]$$

The result of this manipulation is that the first-principles total energy E is given by the expression

$$E = \sum \varepsilon'_i(\boldsymbol{k}) \quad [10]$$

where the sum is over all occupied bands and all k-points in the Brillouin zone. It is noted that the constant V_0 is different for each volume and structure of the first-principles database. The reader should recognize that each band structure has been shifted by a constant, retaining the exact shape of the first-principles bands. It should also be stressed that all this is done to the first-principles database before one proceeds with the fit that will generate the TB Hamiltonian. The next step is to use a least-squares procedure to fit this database with the shifted eigenvalues ε'_i to the TB Hamiltonian.

A typical database in the NRL-TB method contains, for example in a transition metal, five volumes each for the f.c.c. and b.c.c. structures and often the s.c. lattice to achieve better transferability to other periodic structures or defect structures.

The TB method can also be extended to cover magnetism. In this case, a set of spin-polarized parameters are constructed by fitting to spin-polarized linearized augmented plane wave (LAPW) calculations. Common hopping [6] and overlap [7] parameters are used for the majority and minority spin cases, but separate onsite [4] parameters for each channel are developed. Each channel is then diagonalized separately. A set of paramagnetic TB parameters is constructed by averaging the majority and minority spin onsite parameters. Then, the pressure-induced ferromagnetic-to-paramagnetic transitions are studied by noting whether the spin-polarized or paramagnetic parameters produce lower energies. Currently, well-tested parameters for Fe, Co, and Ni are available.

Technical Procedure

The first-principles calculations were done either with the muffin-tin potential augmented plane wave (APW) method or with the full-potential LAPW method. The small differences between these two approaches, especially for closed packed structures, are within the error one makes after fitting to the TB Hamiltonian.

For the f.c.c., b.c.c., and s.c. structures uniform k-point meshes that include the origin and contain 89, 55, and 35 k-points in the irreducible part of the Brillouin zone, respectively, are used. In most cases, the Hedin–Lundqvist parametrization of the local density approximation (LDA) to the density-functional theory (DFT) is used. In some cases, such as for spin-polarized iron, the generalized gradient approximation (GGA) has been used. Spin-polarized calculations were performed for the appropriate metals in the $3d$ transition series. For a typical transition metal, there are $\sim$4000 eigenvalues and energies in the database. An IMSL package is used, based on a finite-difference Levenberg–Marquardt algorithm, to adjust the 93 parameters involved to reproduce this database by means of a nonlinear least-squares fit, with the total energies typically weighted $\sim$200 times larger than the eigenvalues in a single band. Starting parameters are selected guided by those found in the previous work and it is ensured that the symmetry of the eigenstates is taken into account. This is an important issue because a block-diagonalization of the Hamiltonian avoids the possibility of incorrectly aligning the bands and

preserves the angular momentum character of the states. For details of this symmetrization procedure, refer to the "Further reading" section.

Another issue is the number of bands per k-point used in the fit. Both occupied and empty states are fitted. For example, in the transition metals, the lowest six bands for all k-points are fitted but the bands 7–9 are also included for the four high-symmetry points. With this choice, one obtains reliable values for the SK parameters that correspond to the p-orbitals. The fitting RMS errors are in the range 1–5 mRy for six bands and $\sim$0.5 mRy for the total energies. For the semiconductors, eight bands, divided into four valence and four conduction bands in an s–p basis are fitted. Here, the RMS error is $\sim$5 mRy for the valence bands but much higher for the conduction bands. There is a very significant improvement in the conduction bands upon inclusion of the d-orbitals. However, the s–p basis gives a very good fit to the total energy with an RMS error not exceeding 0.5 mRy.

Ground-State Behavior and Phase Stability

The NRL-TB Hamiltonians have been tested to give the correct ground state for all materials for which TB parameters have been generated. The results are summarized in **Tables 1** and **2**. **Table 1** shows the equilibrium lattice constants and bulk moduli for all of the materials studied, comparing them to first-principles LDA calculations, and to experiment. The equilibrium lattice constant is typically within 1–2% of the LDA value and correspondingly very close to the measured values. In this table, the magnetic elements are spin-polarized in a manner consistent with the observed magnetic structure. Thus, one can find that the ground state of chromium is an anti-ferromagnetic CsCl phase, which is a good approximation to the experimentally observed spin density wave.

Similarly, the equilibrium bulk moduli are in good agreement with the first-principles results, within 10% for most elements. With the exception of Ti, the h.c.p. phases were not fit to first-principles results, but instead predicted as an output of the TB scheme. As a result, the equations of state predicted by the TB model are not as accurate as for the cubic lattices. The largest error is for zirconium, where a is 7% smaller than experiment and c is 8% larger. Yttrium and hafnium also show large discrepancies between the TB model and experiment. The lattice constants for the other elements are within 2% of experiment, consistent with the errors one would find in first-principles calculations. The discrepancies from experiment found in Zr, Y, and Hf can be removed by including h.c.p. and, if necessary, the simple cubic lattice.

For Ti, the GGA calculations were fitted with the h.c.p. and sc lattices. In this more elaborate fit, it was possible to improve on the lattice constants and differentiate between the h.c.p. and omega phases. This TB method succeeds in predicting the correct ground state for all materials presented here. **Figure 1** shows the energy volume curves for a representative sample of materials that crystallize in different structures. The results for all of the elements are summarized in **Table 2**, which shows the equilibrium energy of each of these phases expressed as the difference in energy between that phase and the equilibrium energy of the experimental ground state. The energy of each crystal structure is computed by doing a conjugate-gradient minimization of the total energy with respect to all the parameters in the lattice, internal (atomic positions) as well as external (a, c, etc.). The TB method correctly predicts the ground-state structure for all metals including the ferromagnetic metals iron, cobalt, and nickel. The results are also consistent with the accepted crystal structure for the semiconductors Si and Ge, and with the more complex structures of Mn,Ga, and In.

The fact that this method finds the correct ground states of the h.c.p. metals and the other complex structures that were not included in the fit of the TB parameters should be emphasized.

The robustness of the TB parameters has also been established by studying crystal structures of lower symmetry such as the volume-conserving tetragonal strain, the so-called Bain path. The Bain path of Nb at the experimentally observed equilibrium volume of the b.c.c. phase is shown in **Figure 2**. The energy is properly a minimum for the b.c.c. lattice, where $c/a = 1$, and attains a local maximum at the f.c.c. structure ($c/a = \sqrt{2}$). This shows that the f.c.c. elastic constant associated with tetragonal shear, $C_{11} - C_{12}$, is negative; hence, the f.c.c. structure is unstable.

Elastic Constants

Elastic constants measure the proportionality between strain and stress in a crystal, provided that the strain is not so large as to violate Hook's law. Computationally, the elastic constant is determined by applying a strain to a crystal, measuring the energy versus strain, and determining the elastic constant from the curvature of this function at zero strain. A given strain is associated with a certain linear combination of elastic constants. For cubic systems,

Table 1 Equilibrium lattice constants and bulk moduli for the experimentally observed ground-state structures of the elements, comparing the results of the TB parametrization, first-principles LDA results, where available, and experiment

Element	*Structure*	*a* (Å)			*c* (Å)			B_0 (GPa)		
		TB	*LDA*	*Expt.*	*TB*	*LDA*	*Expt.*	*TB*	*LDA*	*Expt.*
C	dia	3.52	3.53	3.57	*a*	*a*	*a*	480	468	443
Mg	h.c.p.	3.22	3.16	3.21	5.26	5.02	5.21	34	39	35
Al	f.c.c.	4.00	3.99	4.05	*a*	*a*	*a*	80	70	72
Si	dia	5.43	5.40	5.43	*a*	*a*	*a*	108	96	99
Ca	f.c.c.	5.35	5.28	5.58	*a*	*a*	*a*	16	19	15
Sc	h.c.p.	3.26	3.21	3.31	4.91	5.01	5.27	63	65	44
Ti	h.c.p.	2.94	2.94	2.95	4.56	4.65	4.68	112	112	107
V	b.c.c.	2.94	2.93	3.03	*a*	*a*	*a*	211	196	162
Cr	CsCl	2.79		2.88	*a*	*a*	*a*	305		190
Mn	αMn	8.64		8.91	*a*	*a*	*a*	248		60
Fe	b.c.c.	2.84	2.83	2.87	*a*	*a*	*a*	180	197	168
Co	h.c.p.	2.54		2.51	4.01		4.07	237		191
Ni	f.c.c.	3.43	3.52	3.52	*a*	*a*	*a*	264	200	186
Cu	f.c.c.	3.52	3.52	3.61	*a*	*a*	*a*	178	188	137
Ga	αGa	4.63	4.38	4.51	4.52	4.35	4.52	651	669	613
		7.63	7.39	7.64						
Ge	dia	5.60	5.61	5.66	*a*	*a*	*a*	67	78	77
Sr	f.c.c.	5.73	5.74	6.08	*a*	*a*	*a*	15	16	12
Y	h.c.p.	3.59	3.52	3.65	5.35	5.61	5.73	46	48	37
Zr	h.c.p.	2.99	3.17	3.23	5.57	5.14	5.15	108	119	83
Nb	b.c.c.	3.25	3.25	3.30	*a*	*a*	*a*	185	193	170
Mo	b.c.c.	3.12	3.12	3.15	*a*	*a*	*a*	283	291	272
Tc	h.c.p.	2.72		2.74	4.34		4.40	304		297
Ru	h.c.p.	2.68		2.71	4.26		4.28	360		321
Rh	f.c.c.	3.77	3.76	3.80	*a*	*a*	*a*	306	309	270
Pd	f.c.c.	3.85	3.85	3.89	*a*	*a*	*a*	212	220	181
Ag	f.c.c.	4.01	4.01	4.09	*a*	*a*	*a*	142	142	101
In	b.c.t.	4.29		4.60	5.10		4.95	52		41
Sn	dia	6.48	6.47	6.49	*a*	*a*	*a*	45	45	53
Ba	b.c.c.	4.82	4.80	5.02	*a*	*a*	*a*	10	11	10
Hf	h.c.p.	3.07	3.18	3.19	5.08	5.15	5.05	111	110	109
Ta	b.c.c.	3.30	3.24	3.30	*a*	*a*	*a*	185	224	200
W	b.c.c.	3.14	3.14	3.16	*a*	*a*	*a*	319	333	323
Re	h.c.p.	2.78	2.77	2.76	4.39	4.50	4.46	371	388	372
Os	h.c.p.	2.75	2.73	2.74	4.31	4.39	4.32	441	440	418
Ir	f.c.c.	3.86	3.82	3.84	*a*	*a*	*a*	389	401	355
Pt	f.c.c.	3.90	3.90	3.92	*a*	*a*	*a*	318	305	278
Au	f.c.c.	4.06	4.06	4.08	*a*	*a*	*a*	196	191	173
Pb	f.c.c.	4.88	4.88	4.95	*a*	*a*	*a*	50	54	45
Po	sc	3.26	3.34	3.35	*a*	*a*	*a*	59	44	26

The "CsCl" structure given for Cr is described in the text. Note that gallium has an orthorhombic structure, so values for lattice constants *a* and *b* are both given in the *a* columns.

including diamond, there are three independent elastic constants, C_{11}, C_{12}, and C_{44}. One linear combination of the elastic constants is obtained from the bulk modulus

$$B(V) = VE''(V) = 1/3(C_{11} + 2C_{12}) \qquad [11]$$

where V is the unit cell volume. A second combination is most easily obtained by straining the crystal in the (1 0 0) direction while simultaneously compressing it in the (0 1 0) direction to conserve the volume, with lengths in the (0 0 1) direction remaining fixed. If x is the fractional change in the (1 0 0) direction, then,

$$C_{11} - C_{12} = E''(0)/V \qquad [12]$$

where the function, $E(x)$ measures the energy as a function of strain. Finally, one finds C_{44} by straining the crystal in the (1 1 0) direction and fixing the volume by compressing in the $(1\ \bar{1}\ 0)$ direction. In this case,

$$C_{44} = 2E''(0)/V \qquad [13]$$

Table 2 TB energies discussed in the text. The energy of the experimental ground-state structure is arbitrarily set to zero. All energies are calculated at the equilibrium volume found by the TB fit and are expressed in mRy. Below the common name of each phase is its *Strukturbericht* designation

Struct. *Struk.*	*f.c.c.* *A*1	*b.c.c.* *A*2	*h.c.p.* *A*3	*dia* *A*4	β*Sn* *A*5	*b.c.t.* *A*6	*gra.* *A*9	α*Mn* *A*12	*sc* A_h
C	306.6	302.3	307.4	1.1	1.1	120.2	0.0	41.7	206.0
Mg	0.8	3.6	0.0	74.5	19.9	2.7	68.0	5.0	28.3
Al	0.0	8.1	2.0	61.0	22.5	5.2	42.2	4.2	29.4
Si	36.4	34.8	36.6	0.0	26.5	34.8	58.6	32.2	20.4
Ca	0.0	2.2	0.8	144.1	28.2	2.3	85.9	6.0	39.9
Sc	4.7	8.6	0.0	101.1	13.5	7.8	49.8	10.3	35.5
Ti	5.1	7.5	0.0	171.6	26.9	7.0	73.8	14.1	57.9
V	19.7	0.0	20.9	180.3	48.2	13.5	112.3	11.8	76.7
Cr	28.7	0.0	30.6	228.2	78.6	26.7	134.4	20.4	119.3
Mn	7.4	14.3	2.8	171.3	60.4	7.4	80.9	0.0	90.0
Fe	9.8	0.0	11.8	76.7	34.0	8.4	61.4	8.0	47.6
Co	2.7	13.1	0.0	92.4	28.0	11.1	48.4	3.4	57.8
Ni	0.0	8.1	2.3	89.7	33.7	5.4	53.9	3.8	55.7
Cu	0.0	2.6	0.4	81.3	27.2	2.6	59.0	6.1	39.3
Ge	17.3	20.7	19.6	0.0	15.5	19.4	48.1	19.8	12.4
Sr	0.0	16.4	15.3	87.2	30.6	14.9	59.1	24.4	36.2
Y	2.3	8.8	0.0	97.9	16.7	8.0	46.9	8.6	34.3
Nb	29.6	0.0	28.8	187.7	42.8	14.0	111.6	15.4	74.1
Mo	29.7	0.0	30.6	147.0	48.4	24.9	90.8	17.2	68.8
Tc	6.1	23.4	0.0	72.9	43.4	21.8	53.3	0.2	56.8
Ru	7.5	50.9	0.0	134.5	80.8	42.0	124.8	16.4	106.2
Rh	0.0	31.6	5.0	159.7	68.0	17.2	120.5	16.2	96.2
Pd	0.0	10.2	2.6	148.8	60.9	8.6	109.1	9.3	85.0
Ag	0.0	2.8	0.6	66.1	18.2	2.5	50.2	6.0	24.6
Ba	0.8	0.0	0.0	12.0	14.1	0.8	52.4	2.8	23.9
Hf	0.9	7.1	0.0	380.3	73.8	6.9	203.9	22.9	115.9
Ta	24.6	0.0	25.4	188.7	39.8	11.8	107.4	9.0	64.6
W	35.8	0.0	38.1	159.3	76.5	31.7	124.8	18.2	115.0
Re	9.1	28.1	0.0	38.8	35.1	26.2	35.1	0.2	55.3
Os	7.9	66.0	0.0	11.8	44.4	47.5	65.5	17.8	59.0
Ir	0.0	49.9	8.3	83.7	57.7	20.2	96.6	23.9	83.4
Pt	0.0	10.0	4.8	169.0	54.8	5.0	131.9	14.5	79.1
Au	0.0	1.0	0.7	70.6	15.4	0.5	55.4	8.1	20.4
Pb	0.0	4.6	2.2	14.6	13.3	2.5	24.2	1.2	19.1
Po	21.3	14.4	5.7	23.9	7.4	22.0	14.5	1.3	0.0

Note that for diamond the calculation of C_{44} requires the minimization of the total energy with respect to an internal parameter at each strain.

The calculation of elastic constants assesses the ability of the method to determine properties not included in the fitting database. The elastic constants of the cubic materials calculated by the NRL-TB approach are compared to the experimental values in **Table 3**. These calculations were performed at the experimental volume and therefore they are not consistent with the equilibrium value of B found in **Table 1**. The deviation of the calculated C_{11} and C_{12} from the measured values is within 10–15% which is very close to the error one finds when comparing with a direct evaluation from first-principles calculations.

An inspection of **Table 3** reveals somewhat larger errors for C_{44}. A solution to this problem is to fit C_{44} either to LAPW calculations or to experiment. It should be noted here that calculations of C_{ij} in the alkaline-earth metals is problematic because the lattice is extremely soft. One can also note that the method correctly reproduces the sign of the elastic constant difference $C_{12} - C_{44}$, even in the metals rhodium and iridium, and the semiconductors Si and Ge where it is negative. It is worth mentioning that this negative sign cannot be obtained from the standard embedded-atom method.

The resulting elastic constants for the h.c.p. lattice are C_{11}, C_{12}, C_{13}, C_{33}, and C_{44}. The elastic constants found by the NRL-TB method show larger relative deviations from experiment than found in the cubic crystals. These were found from TB parameters fitted only to cubic lattices. It is shown, in the case of Ti, that if one extends the fitting database to include the h.c.p. lattice as well, there is a dramatic improvement in the comparison of TB and measured elastic constants.

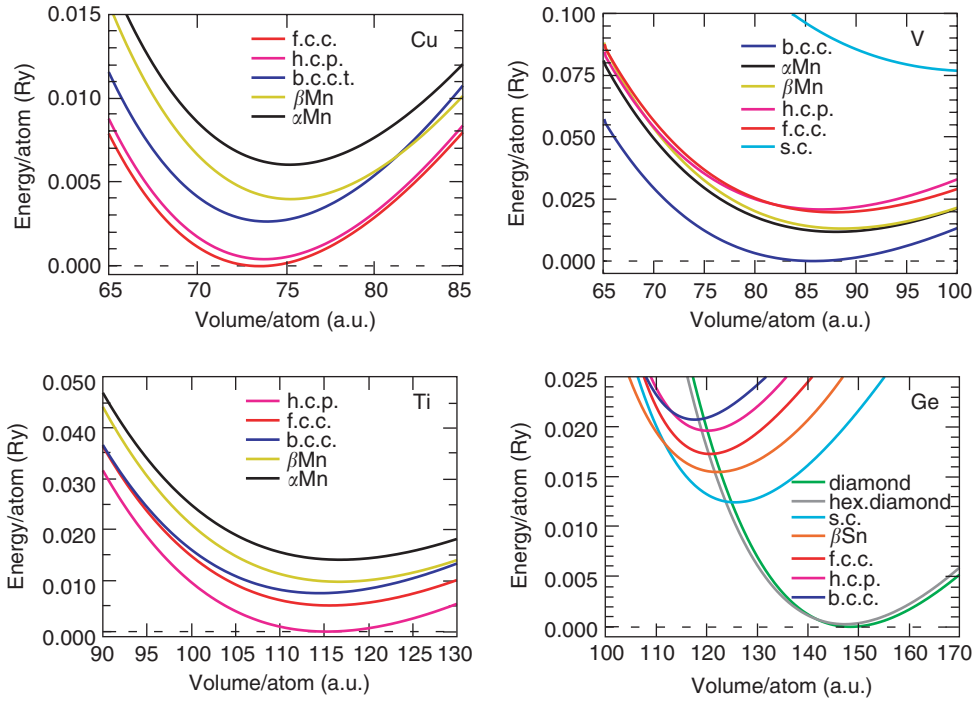


Figure 1 Energy vs. volume curves for various phases of (clockwise from the top left) copper, vanadium, germanium, and titanium, using the NRL-TB parameters described in the text. Note that the method correctly predicts the equilibrium structure.

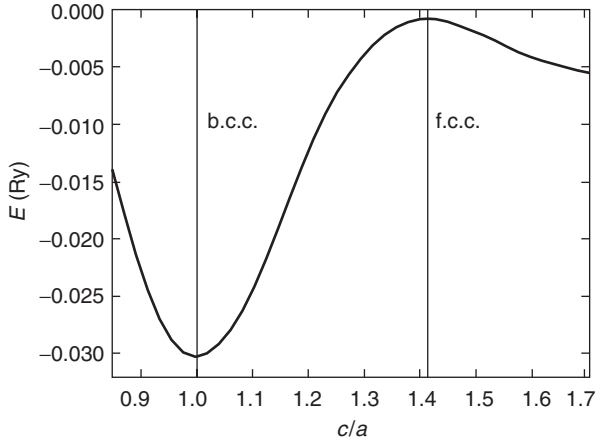


Figure 2 TB calculation of the energy of molybdenum, at the experimental equilibrium volume, under a tetragonal strain, as a function of c/a. The vertical lines denote the positions of the b.c.c. and f.c.c. lattices.

Vacancies

Vacancy formation energies have been calculated by a supercell method. One atom in the supercell is removed and neighboring atoms are allowed to relax around this vacancy while preserving the symmetry of the lattice. The great advantage of the NRL-TB method over first-principles approaches is that one

Table 3 Elastic constants for cubic elements (in GPa)

Element	*Structure*	*TB*			*Exp.*		
		C_{11}	C_{12}	C_{44}	C_{11}	C_{12}	C_{44}
C	dia	1036	48	601	1076	125	576
Al	f.c.c.	125	58	26	103	53	28
Si	dia	179	73	95	166	64	80
Ca	f.c.c.	15	10	14	16	12	8
V	b.c.c.	224	106	92	228	119	43
Fe	b.c.c.	223	95	78	237	141	69
Cu	f.c.c.	139	99	59	156	106	75
Ge	dia	133	20	107	131	49	68
Sr	f.c.c.	8	3	−3	15	6	10
Nb	b.c.c.	277	139	37	246	139	29
Mo	b.c.c.	453	147	120	450	173	125
Rh	f.c.c.	491	171	260	433	185	206
Pd	f.c.c.	233	163	63	227	176	72
Ag	f.c.c.	133	86	42	124	93	46
Sn	dia	68	30	38	67	36	30
Ta	b.c.c.	275	140	78	261	157	82
W	b.c.c.	529	170	198	523	203	160
Ir	f.c.c.	694	260	348	590	249	262
Pt	f.c.c.	380	257	71	347	251	76
Au	f.c.c.	195	174	40	189	159	42
Pb	f.c.c.	53	35	19	47	39	14
Po	sc	128	13	6			

All elements for which a set of TB parameters have been constructed and for which one has a cubic ground state (except manganese) are presented here. A comparison is made between the results of the TB parametrization and experiment. Calculations were performed at the experimental volume.

can do the calculation in a very large supercell, in a computationally efficient manner, including relaxation. It is found that a supercell containing 128 atoms is sufficient to eliminate the vacancy–vacancy interaction. The vacancy formation energy is given by

$$E_{\text{vac}}(N) = E(N-1,1) - (N-1)E(N,0)/N \quad [14]$$

where $E(M, Q)$ is the total energy of a supercell containing $M + Q$ sites, where M are occupied by atoms and Q are vacant, whence $E(N, 0)$ is N times the energy of one atom in its bulk crystal structure. The experimental lattice constant is used to set the volume of the system, since under experimental conditions the lattice constant of a metal containing isolated vacancies will be the lattice constant of the bulk metal.

Calculated values of the vacancy formation energy of the metals are shown in **Table 4**. (Note that vacancy formation energies for Si and Ge are shown in **Table 6**.) Relaxation around the vacancy was included via a conjugate gradient procedure but it introduced, not unexpectedly, a small effect. Where available, it is compared to both first-principles calculations and experiment. The results for niobium, silver, tantalum, and iridium are in excellent agreement with experiment. The vacancy formation energies for rhodium, tungsten, and platinum do not compare well with experiment. Aluminum is an intermediate case. It is possible that the measured values are not reliable. Otherwise, there may be a need to check these results against first-principles data and perhaps fit to DFT calculations.

Table 4 TB vacancy formation energies compared to first-principles calculations and experiment

Element	*TB*		*LDA*	*Exp.*
	Fixed	*Relaxed*		
Al	0.49	0.40	0.56, 0.84	0.66
Cu	1.29	1.18	1.41, 1.29	1.28 – 1.42
Nb	2.84	2.82		2.62 ± 0.03
Mo	2.63	2.46		3.0 – 3.6
Rh	3.39	3.35	2.26	1.71
Pd	2.46	2.45	1.57	1.85 ± 0.25
Ag	1.31	1.24	1.20, 1.06	1.11 – 1.31
Ta	3.17	2.95		2.9 ± 0.4
W	6.86	6.43		4.6 ± 0.8
Ir	2.19	2.17		1.97
Pt	2.79	2.79		1.35 ± 0.09
Au	1.24	1.12		0.89 ± 0.04
Pb	0.76	0.64		0.54

Energies were computed using a 128-atom supercell. Calculations with the atoms at the primitive lattice sites in the crystal (fixed) and allowing relaxation around the vacancy (relaxed) are shown. First-principles calculations of the vacancy formation energy are given in the column labeled "LDA." The experimental column shows a range of energies if several experiments have been tabulated. Otherwise, the estimated error in the experiment is given.

TB results are from Mehl MJ and Papaconstantopoulos DA (1996) *Physical Review B* 54: 4519. Experimental data are from Schaefer H-E (1987) *Physica Status Solidi A* 102: 47.

Surfaces

Surface energies are also calculated by a supercell technique. A slab of metal is formed by cleaving the crystal along the desired plane, creating two identical free surfaces. The distance between the two surfaces is increased, creating a set of slabs that repeat periodically in the direction perpendicular to the surfaces. The slabs are separated by a large region of vacuum so that the electrons on one slab cannot hop to a neighboring slab. In addition, the slabs must be thick enough, so that the atoms at the center of the slab have the electronic properties of atoms in the bulk material and the two surfaces on the same slab cannot interact with each other. It is found that these criteria are met if slabs containing 25 atomic layers and 7–13 layers between slabs depending on the direction of the surface are used. Care should be taken with respect to the k-point mesh convergence. Depending on the surface and underlying bulk structure, this requires $\sim$200 k-points in the two-dimensional Brillouin zone. The surface energy, expressed as the energy required to create a unit area of new surface, is then given by the formula

$$E_{\text{surf}} = (E_{\text{slab}} - NE_{\text{bulk}})/(2A) \quad [15]$$

where A is the area occupied by one unit cell on the surface of the slab, E_{slab} is the total energy of the slab, N is the number of atoms in the unit cell, and E_{bulk} is the energy of one atom in the bulk at the lattice constant of the atoms in the interior of the slab. In the first surface calculations, the bulk equilibrium lattice parameters with no relaxation or reconstruction at the surface were used. These calculations showed reasonable agreement with experiment indicating that relaxation does not seriously affect the surface energy. However, recently calculations with all atoms completely relaxed were performed, including charge self-consistency. The new results for Nb and several f.c.c. metals show significant differences in the interlayer separations and reconstruction.

The results for the f.c.c. metals are particularly gratifying. For the f.c.c. metals, it is found that $E(1\,1\,1) < E(1\,0\,0) < E(1\,1\,0)$, which means that close-packed surfaces are the most stable for the f.c.c. metals. For the b.c.c. metal surface energies, the

opposite inequality is satisfied, $E(1\,1\,0) < E(1\,0\,0) < E(1\,1\,1)$.

Stacking Faults

Stacking faults are introduced in a crystal by cutting a perfect crystal block along a plane and shifting the upper part with respect to the lower part by a vector $\boldsymbol{f}$, defining the generalized stacking fault energy surface. Local energy minima on this surface are called stable intrinsic stacking faults. In f.c.c. systems and for the (1 1 1) slip plane, the stable stacking fault corresponds to a slip of $a/\sqrt{6}$ in the $\langle 1\,2\,1 \rangle$ direction. The $\langle 1\,2\,1 \rangle$ slip is modeled on a (1 1 1) slip plane by a supercell consisting of nine close-packed (1 1 1) planes of atoms. The primitive vectors of the supercell are

$$\begin{aligned} \boldsymbol{a}_1 &= \frac{1}{2}\,a\hat{y} + \frac{1}{2}\,a\hat{z} \\ \boldsymbol{a}_2 &= \frac{1}{2}\,a\hat{x} + \frac{1}{2}\,a\hat{z} \\ \boldsymbol{a}_3 &= \left(3 + \frac{q}{6}\right)a\hat{x} + \left(3 + \frac{q}{6}\right)a\hat{y} - \left(3 - \frac{q}{3}\right)a\hat{z} \end{aligned} \qquad [16]$$

where q is the stacking fault displacement of the atoms in the boundary layer along the vector f in the $\langle 1\,1\,2 \rangle$ direction. At $q = 0$, the atoms are in f.c.c. positions and at $q = 1$, at the h.c.p. positions. Since the calculations are computationally intensive because of the nine planes of atoms and the 4730 k-points used in the irreducible part of the Brillouin zone, the NRL-TB method proves to be very efficient. The results of the stacking fault energy as a function of the displacement q are shown in **Figure 3**. The results correctly predict the qualitative effect that the energies for Ir are much larger than that for the noble metals and they also are in agreement with full-potential linearized muffin-tin orbital (FLMTO) results for Al, Ag and Ir. Finally, the value at the maximum energy in **Figure 3** has been used to extract the unstable stacking fault energy γ_{us}, which is used together with the surface energy γ_{surf} to establish the Rice criterion for ductility,

$$D = 0.3\gamma_{us}/\gamma_{surf} \qquad [17]$$

One finds, in agreement with experiment, that the noble metals have the largest values of D.

Phonons

Phonon frequencies at high-symmetry points in the Brillouin zone computed with the NRL-TB model using the frozen phonon approximation are compared with experimentally measured values in **Table 5**. In this table, results from metals and semiconductors have been included. The agreement with experiment is similar to that with elastic constants. For most modes the TB results are within 15–20% of experiment, which is also comparable to phonon frequencies obtained directly from first-principles calculations. For

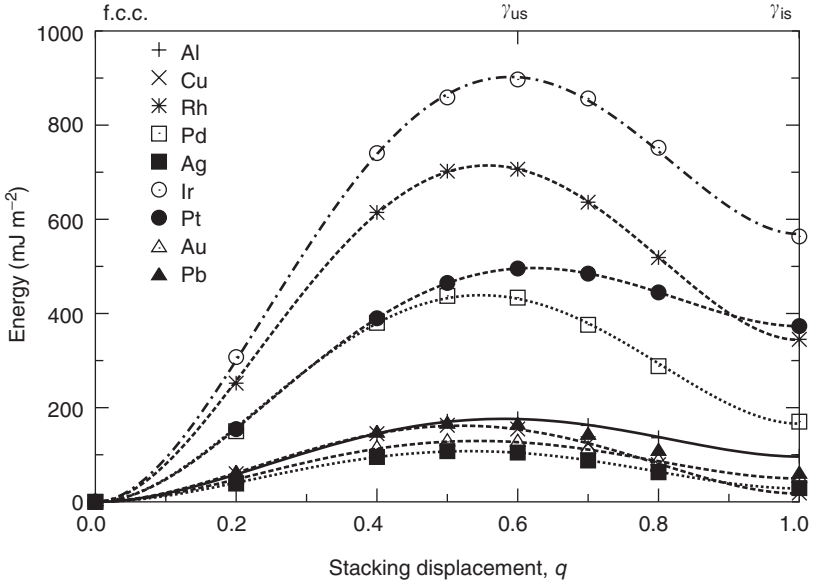


Figure 3 Stacking fault energy as a function of the parameter q in [16] for the elemental f.c.c. metals, as determined by the NRL-TB method. The labels γ_{us} and γ_{is} indicate the positions of the unstable and intrinsic stacking faults, respectively, while the f.c.c. label on the left vertical axis indicates the position of the ground-state bulk structure. (Mehl MJ, Papaconstantopoulos DA, Kioussis N, and Herbranson M (2000) *Physical Review B* 61: 4894.)

Table 5 Phonon frequencies (in THz) at high symmetry points for various cubic elements

	Sym.	*TB*	*Exp.*
C	Γ'_{25}	39.3	39.9
(dia)	X_1	34.2	35.5
	X_3	29.8	32.1
	X_4	24.7	24.2
	L_1	39.3	36.6
	L'_2	36.5	31.0
	L_3^+	17.6	16.9
	L_3^-	35.3	36.2
Mg	Γ_3^+	12.1	7.3
(h.c.p.)	Γ_5^+	4.3	3.7
	A_1	7.1	2.9
	A_3	3.9	5.2
Si	Γ'_{25}	15.9	15.5
(dia)	X_1	12.1	12.3
	X_3	15.2	13.9
	X_4	4.8	4.5
	L_3^+	3.8	3.4
	L_3^-	16.0	14.7
Ti	Γ_3^+	5.7	5.5
(h.c.p.)	Γ_5^+	4.2	4.1
	A_1	5.3	5.7
	A_3	2.9	3.0
Cu	X_3	7.2	7.3
(f.c.c.)	X_5	4.4	5.1
	L_2	7.9	7.3
	L_3	3.5	3.4
Ge	Γ'_{25}	9.7	9.1
(dia)	X_1	6.4	7.2
	X_3	8.9	8.3
	X_4	3.2	2.4
	L_1	9.3	8.7
	L'_2	5.2	6.7
	L_3^+	2.5	1.9
	L_3^-	7.3	7.3
Nb	H	6.3	6.5
(b.c.c.)	P	5.7	5.0
	N_2	4.9	5.1
	N_3	5.9	5.7
	N_4	4.4	3.9
Mo	H	4.1	5.5
(b.c.c.)			
Au	X_3	5.3	4.6
(f.c.c.)	X_5	2.9	2.7
	L_2	5.5	4.7
	L_3	1.9	2.0
Pb	X_3	1.9	1.8
(f.c.c.)	X_5	1.0	0.9
	L_2	2.2	2.2
	L_3	1.0	0.9

some materials, larger errors as for the L_1, and W_2 modes in silicon, and the P and N modes in niobium have been found. This issue has been addressed in a recent paper on Nb, where by augmenting the first-principles database by including the energies at displacement of 0.005 lattice coordinates for the high symmetry points P and N, the overall fit of the

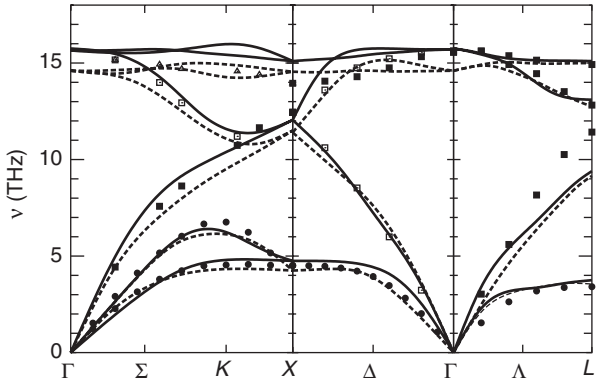


Figure 4 Phonon dispersion curves for diamond-structured Si computed from the velocity–velocity correlation function found during a molecular-dynamics simulation at 300 K (solid line) and 1500 K (dashed line). Dots are taken from experimental data. (Bernstein N, Mehl MJ, Papaconstantopoulos DA, *et al.* (2002) *Physical Review B* 62: 4477.)

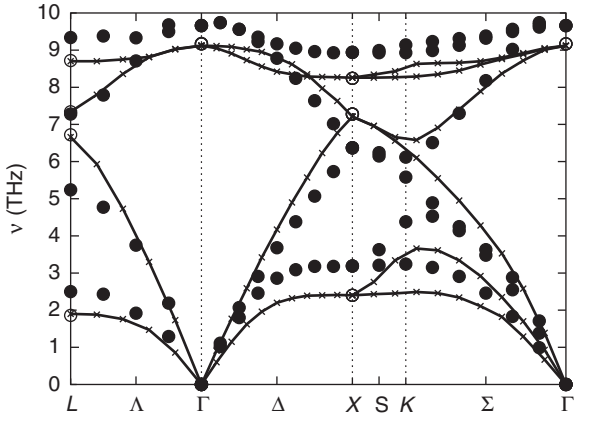


Figure 5 Phonon frequencies along high symmetry lines in Ge using the sp^3 basis TB parameters (from Bernstein N, Mehl MJ, and Papaconstantopoulos DA (2002) *Physical Review B* 66: 075212 (solid symbols) compared to DFT/LDA results (from Giannozzi P, de Gironcoli S, Pavone P, and Baroni S (1991) *Physical Review B* 43: 7231) (open symbols) and experiment at 80 K (from Nelin G and Nilsson G (1972) *Physical Review B* 5:3151).

phonon spectrum has been drastically improved. **Figures 4** and **5** show the phonon dispersion curves of Si and Ge that indicate a good agreement with the measured spectra.

Point Defects

Point defects in the semiconductors Si and Ge have been calculated. Point defects are a major source of disorder in semiconductors, since they are thermodynamically favored to occur at finite concentrations and temperatures. Because they are far more mobile than perfectly bonded atoms, they dominate

Table 6 Formation energies E_f (ideal) and relaxation energies ΔE_f (relaxed), in eV, for point defects computed using the TB model and comparison to DFT/LDA results

	TB		*LDA*		*TB*		*LDA*
	E_f^0	E_f	E_f^0	E_f	E_f^0	E_f	E_f
V	Si 4.2	3.2	3.3–4.3	2.9–3.7	Ge 4.6	3.6	1.9
I_t	4.8	4.5	3.7–4.8	3.6–4.6	3.1	2.5	3.2
I_h	5.6	5.1	4.3–5.0	3.7–3.9	5.6	4.4	2.9
⟨1 1 0⟩	3.7		3.3		4.3	3.3	2.3

Since the structure of the ideal split interstitial is not uniquely defined, the energy listed under ideal is actually the relaxed formation energy.

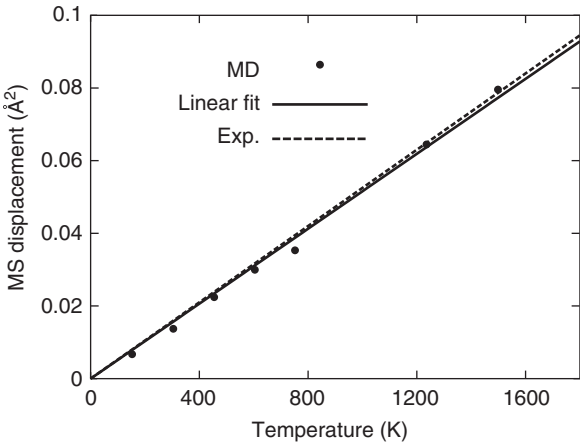


Figure 6 Mean-squared displacement of atoms in diamond-structured silicon as a function of temperature, as determined by TBMD. The dots represent the TBMD data, and the lines are linear (solid) and higher-order representations of the experimental data. (From Batterman BW and Chipman DR (1962) *Physical Review* 127: 690.)

diffusion. The formation energy of such defects strongly influences their concentrations and is, therefore, an important material property.

Table 6 lists the formation energies of three interstitial configurations: the tetrahedral, hexagonal, and 110 split, for Si and Ge. These defect energies were computed using a cubic supercell at the equilibrium lattice constant with 216 atoms, sampling the Brillouin zone at the Γ point. The relaxed configurations were obtained by a conjugate-gradient algorithm using as a criterion the force on each atom to be less than $3\,\mathrm{meVA}^{-1}$. The formation and relaxation energies of all three interstitial configurations are within 10% of the range of LDA calculations for Si but not as good for Ge. On the other hand, the relaxed geometries show more serious discrepancies from *ab initio* results.

Finite Temperature Properties from Molecular Dynamics

The TB method described above can be used with the NRL-TB molecular dynamics (TBMD) package developed by Kirchhoff *et al.* to determine thermodynamic properties. For example, the TBMD package can be used to describe computed mean square displacements, and pair correlation functions as a function of temperature, as well as thermal expansion, in silicon, germanium, and gold. It is also possible to extract phonon-dispersion curves from the velocity autocorrelation function determined in an MD simulation.

As an example of the power of the method, one can describe the calculation of the mean-squared displacement of diamond structure silicon as a function of temperature. The calculation used a 512-atom unit cell, evolving at constant energy with fixed initial kinetic energy for 2000 time steps of length 2.0 fs. As seen in **Figure 6**, results are compared with the experimentally determined values, computed from the experimental measurements of the temperature dependence of broadening of the X-ray diffraction peak. The agreement is excellent.

Very recently, de Silva and co-workers used the NRL-TB parameters for Au to present a very interesting simulation of the formation, evolution, and breaking of Au nanowires.

Multicomponent Systems

The above work was confined to single-component systems. The method is extendable to systems with more than one atomic species by making two modifications. First, one adds additional terms to the onsite energies [4] which couple the density from atoms of type *i* to atoms of type *j*. Second, one notes that the SK parameters $ps\sigma$, $ds\sigma$, $dp\sigma$, and $dp\pi$, which in the single-atom case are equivalent (within a sign) to $sp\sigma$, $sd\sigma$, $pd\sigma$, and $pd\pi$, respectively, are now independent parameters. These complications raise the total number of parameters from 93 in the single-atom case, to 330 for two-atom type systems.

This method has been applied to several compounds, including Nb–C, Pd–H, Fe–Al, Pt–O, and Cu–Au. For Cu–Au, the energy of a variety of structures was computed, and it was found that the expected ordered phases (the $L1_2$, Cu_3Au, and $L1_0$ CuAu) were stable, while other structures were not. However, the advantage of the TB model is that one can compare to structures which are not found experimentally, allowing one to explore configurations that might occur locally in an alloy but which do not form long-range ordered states. To check the correctness of the Cu–Au parameters in this case, it is compared to the cluster-expansion and first-principles

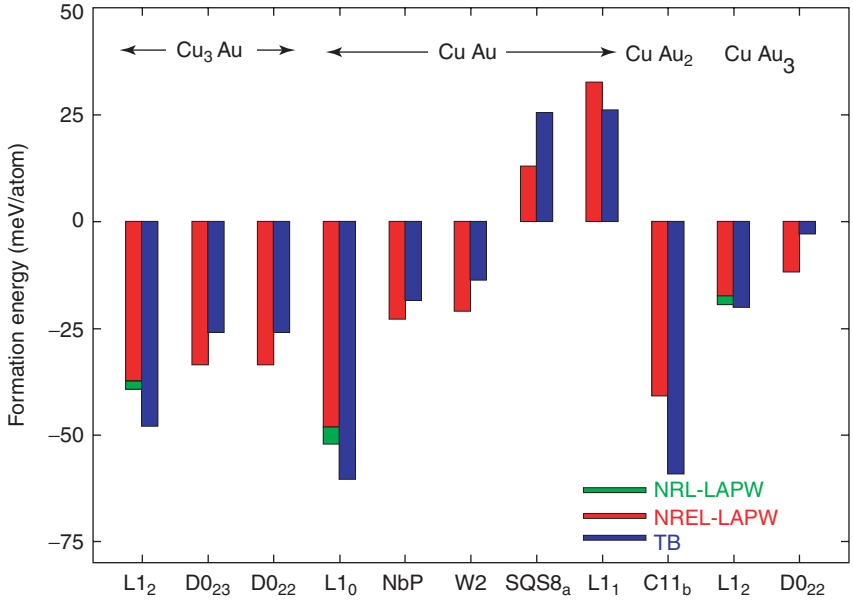


Figure 7 Formation energy of several ordered phases in the Cu_xAu_{1-x} system, calculated using the TB parameters (blue bars), and compared to first-principles calculations performed by Ozoliņš *et al.* (red bars). (The structure notation is from Ozoliņš V, Wolverton C, and Zunger A (1998) *Physical Review B* 57: 6427.) On this scale, the cluster-expansion energies are indistinguishable from the corresponding LAPW results. For comparison, the first-principles LAPW results (green bars), which were used in the Cu–Au TB fitting process are also plotted.

calculations of Ozoliņš *et al*. **Figure 7** shows formation energies computed for several ordered phases in the Cu_xAu_{1-x} system, and compares them to the energies obtained by Ozoliņš *et al*. It is found that the agreement is excellent.

Acknowledgments

The authors thank the collaborators in this work. This work was supported by the US Office of Naval Research. The development of the tight-binding codes was supported in part by the US Department of Defense Common HPC Software Support Initiative (CHSSI).

See also: Density-Functional Theory; Elemental Semiconductors, Electronic States of; Insulators, Electronic States of; Intermetallic Compounds, Electronic States of; Metals and Alloys, Electronic States of (Including Fermi Surface Calculations); Recursive Methods for Electronic States.

PACS: 71.15.Nc; 71.20. – b; 61.43.Bn

Further Reading

Bacalis NC, Papaconstantopoulos DA, Mehl MJ, and Lach-hab M (2001) *Physica B: Condensed Matter* 296: 125.

Bassani F and Pastori Parravicini G (1975) *Electronic States and Optical Transitions in Solids*. Oxford: Pergamon.

Bernstein N, Mehl MJ, Papaconstantopoulos DA, Papanicolaou NI, Bazant MZ, *et al.* (2002) *Physical Review B* 62: 4477. Erratum: *Physical Review B* 65: 249002(E).

Cleri F and Rosato V (1993) *Physical Review B* 48: 22.

Cohen RE, Mehl MJ, and Papaconstantopoulos DA (1994) *Physical Review B* 50: 14694.

Gross A, Scheffler M, Mehl MJ, and Papaconstantopoulos DA (1999) *Physical Review Letters* 82: 1209.

Groß A, Eichler A, Hafner J, Mehl MJ, and Papaconstantopoulos DA (2003) *Surface Science* 539: L542.

Haftel MI, Bernstein N, Mehl MJ, and Papaconstantopoulos DA (2004) *Physical Review B* 70: 125419.

Harrison WA (1989) *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*. New York: Dover.

Jancu JM, Scholz R, Beltram F, and Bassani F (1998) Empirical *spds** tight-binding calculation for cubic semiconductors: general method and material parameters. *Physical Review B* 57: 6493.

Kirchhoff F, Mehl MJ, Papanicolaou NI, Papaconstantopoulos DA, and Khan FS (2001) *Physical Review B* 63: 195101.

Lekka CE, Bernstein N, Mehl MJ, and Papaconstantopoulos DA (2003) *Applied Surface Science* 219: 158.

Lekka CE, Mehl MJ, Bernstein N, and Papaconstantopoulos DA (2003) *Physical Review B* 68: 035422.

McMahan AK (1998) *Physical Review B* 58: 4293.

Mehl MJ, Klein BM, and Papaconstantopoulos DA (1994) First-principals calculation of elastic properties. In: Westbrook JH and Fleischer RL (eds.) *Intermetallic compounds-principles and practice*, vol. 1, ch. 9, pp. 195–210. London: Wiley.

Mehl MJ and Papaconstantopoulos DA (2002) *Europhysics Letters* 60: 248.

Mehl MJ, Papaconstantopoulos DA, Kioussis N, and Herbranson M (2000) *Physical Review B* 61: 4894.

Papaconstantopoulos DA (1986) *Handbook of the Band Structure of Elemental Solids*. New York: Plenum.

Papaconstantopoulos DA and Mehl MJ (2003) *Journal of Physics Condensed Matter* 15: R413.

Shi L and Papaconstantopoulos DA (2004) Modifications and extensions to Harrison's tight-binding theory. *Physical Review B* 70: 205101.

Slater JC and Koster GF (1954) *Physical Review* 94: 1498.

Yang SH, Mehl MJ, and Papaconstantopoulos DA (1998) *Physical Review B* 57: R2013.

Time-Resolved Optical Spectroscopies

B Deveaud, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Introduction

Spectroscopy basically consists in retrieving information on any system through the use of optical information as a function of wavelength. The examples are countless, and the field is of major importance. For example, the structure of the hydrogen atom has been decoded through the sequence of absorption lines known as the Balmer series. Optical spectroscopy of solids is known to have provided invaluable information over the years, however, when it comes to dynamical properties, one faces right away the challenge that the typical timescales in solids can be incredibly short: 100 fs is a time that cannot be considered as a limit but rather as typical of a wide range of processes. Therefore, the development of ultrafast spectroscopic technique.

Ultrafast optical spectroscopy is particularly useful when we try to obtain information on very short timescales, where other techniques cannot reach the necessary time resolution. The very high frequency of light, 5×10^{14} Hz in the visible range, indeed allows nowadays to reach time limits of the order of a few cycles of the light field only. It is even anticipated that the reduction of the laser wavelength toward X-rays, will allow breaking the barrier of the femtosecond, and therefore entering the attosecond regime. We will not describe here the different techniques which allow to produce ultrashort light pulses, the reader only needs to know that convenient means exist that can produce pulses as short as 5 fs. Once we have such pulses at our disposal, spectroscopic techniques will allow us to probe the properties of materials with a time resolution basically given by the pulse length. Basically the spectroscopy will try to probe all properties of interest in the material, trying for example to get information on the dynamics of the distribution function of charge carriers, or of any quasiparticle in the system.

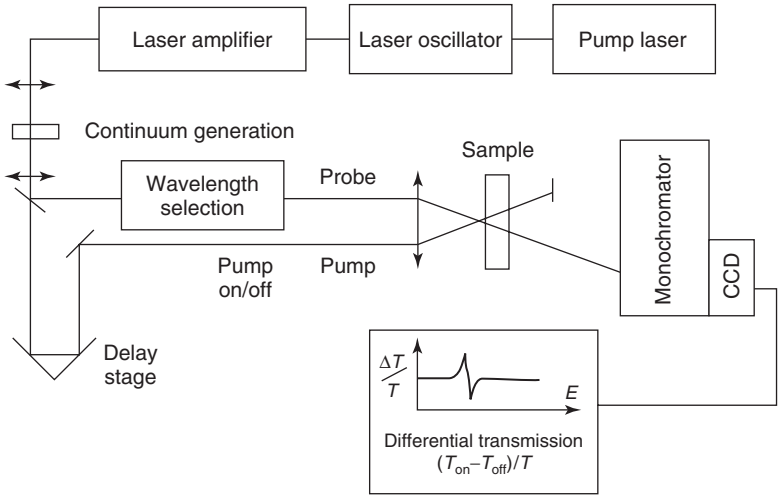


Figure 1 Schematic arrangement of a P&P experiment. A titanium sapphire laser oscillator usually creates short pulses. These pulses might be used directly to perform the experiment, leading to a measurement bandwidth limited by the Fourier transform-limited width of the pulses. It is often useful to create a larger bandwidth by focusing an intense laser pulse onto a dielectric medium. This however necessitates the use of a laser amplifier to obtain large enough pulse energy. The pump pulse is properly filtered to provide the required excitation energy and focused on the sample. The broad and less intense probe pulse is delayed by a delay stage and focused at the center of the pump spot. The changes in transmission are then recorded on the probe beam, and are usually very small so that differential techniques have to be used. The plotted quantity is usually the difference probe signal with and without the pump pulse.

Two techniques are classified along two main categories, the linear techniques, which rely on the probing of the linear response of the system, and the nonlinear techniques, which specifically aim at the probing of the nonlinear response of a given system. Most examples are taken from studies in the field of semiconductors, although the techniques may clearly be applied to all optically active materials, which basically cover more or less any kind of material system.

Linear Techniques

Absorption and Reflectivity

The first technique to be used in the field of ultrafast spectroscopy has been the pump and probe technique (P&P). This technique is indeed applicable as soon as you have a short light pulse. The idea behind the words being that you send a short light pulse on a sample, thereby modifying its properties. The changes in the properties of the system are probed in the time domain by a second pulse (the probe), time delayed with respect to the first one simply by a change in the path of this second beam. A schematic arrangement of a typical P&P setup is displayed in **Figure 1**.

Two basic arrangements may be employed; in the first one, the absorption changes are probed as a function of time after the arrival of the first pump pulse. In the second arrangement, the changes in reflectivity are measured. It is easy to show that the two types of changes are very directly linked to one another, however, the interpretation of the absorption changes being, in general, easier. The absorption P&P technique will then be preferred when usable, that is to say when the absorption of the sample is not too large. This has been applied very effectively to the understanding of semiconductor nanostructures.

An important aspect of such techniques is that, even if the pump pulse has to be as broad spectrally as imposed by the Fourier transformation of the time variation of the pulse (a 100 fs Gaussian pulse has to be 800 GHz broad), it is usually very important that absorption can be probed over a much larger spectral range, and techniques to obtain very broad pulses (so-called "white light continuum") have been developed over the years. An example of P&P absorption in a quantum well, showing the appearance of a spectral hole is shown in **Figure 2.** The spectral hole appears because of a reduction of the absorption at the energy of the exciting pulse, a large fraction of the available final states being occupied. The absorption is reduced, at the energy of the exciting pulse, over typical times corresponding to the relaxation of

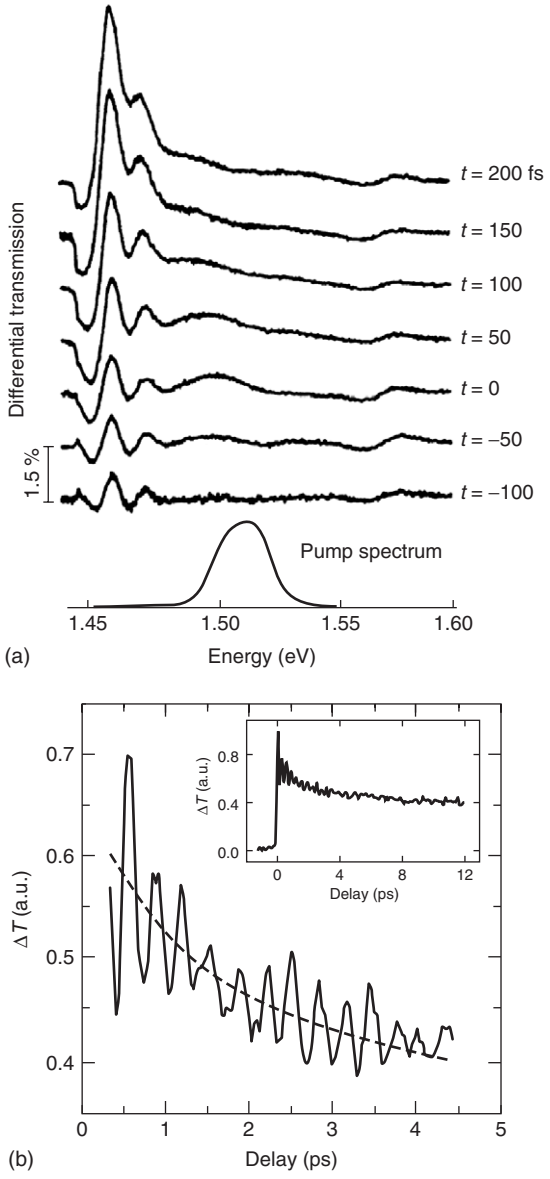


Figure 2 (a) P&P absorption spectrum of a GaAs multiquantum well system for a negligible delay between pump and probe. A clear spectral hole is observed at the energy position of the pump, which disappears with a typical time corresponding to the electron dephasing (less than 100 fs in the present case). The strong oscillations around 1.45 eV correspond to the effects of Coulomb correlations at the exciton energy. (Reproduced with permission from Knox W, Hirlimann C, Miller DAB, Shah J, Chemla DS *et al.* (1986) Femtosecond excitation of nonthermal carrier populations in GaAs quantum wells. *Physical Review Letters* 56: 1191–1193.) (b) P&P absorption signal of F centers in KBr, showing directly the oscillations of the electron wave packet after a short pulse excitation. The F center, sometimes called color center because it gives rise to the color of the crystal, is linked with an anion vacancy occupied by a single electron. The period of the oscillations is obtained from the Fourier power spectrum of the signal and corresponds to 0.33 ps. The inset shows the P&P signal over a larger time range. (Reproduced with permission from Nisoli M, DeSilvestri S, Svelto O, Scholz R, Fanciulli R, *et al.* (1996) Single-electron subpicosecond coherent dynamics in KBr F centers. *Physical Review Letters* 77: 3463–3466.)

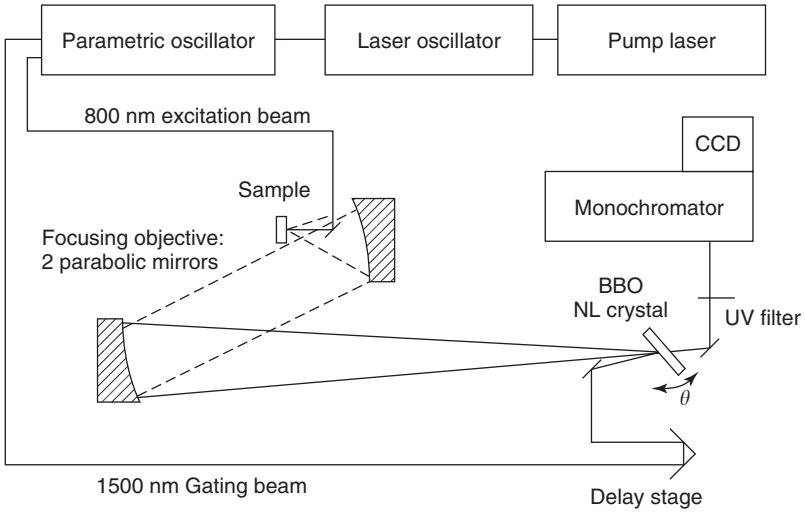


Figure 3 Schematic arrangement of an upconversion luminescence experiment. Short pulses are usually created by a titanium sapphire laser oscillator. These pulses might be used directly to perform the experiment in the case of nonresonant experiment. In the case of resonant excitation of a system, the upconverted signal can only be discriminated when the gating laser pulse is at a different energy. This can be obtained by using a parametric oscillator. The pulse from the oscillator is used to excite the sample. The luminescence is collected by reflective optics to avoid dispersion in the objective (dispersion would be a major problem in luminescence due to the very large bandwidth of the signals to be studied) and focused onto a nonlinear crystal tuned at the proper angle for phase matching with the gating pulse. This gating pulse, coming out from the parametric oscillator, is properly delayed and also focused on the nonlinear crystal. Proper phase matching allows the generation of photon addition: one photon coming from the luminescence signal, the other one from the gating beam. Thus the process is only efficient during the presence of the gating pulse, leading to a limit time resolution simply corresponding to the width of the gating pulse. After the nonlinear crystal, proper filtering and standard detection techniques may be used to only record the upconverted photons that are produced in a different wavelength range to the other signals.

the photoexcited carriers (here, less than 100 fs). Further, the changes in the electron and hole distributions can be monitored very directly, with one main difficulty due to the combined presence of electrons and holes excited in the system.

Luminescence

Another way to probe the dynamics of charge carrier changes, especially in semiconductors, is to probe the properties of the light emitted by the sample. Once again the emitted spectrum will be an image of the changes in the distribution functions of electrons and holes. Luminescence techniques have advantages and drawbacks compared to P&P. The main difficulty being that the photon detector must possess the appropriate time resolution (in P&P, the pulse itself detects the changes, and therefore has the appropriate duration). On the other hand, luminescence is a background free technique, which allows probing very weak changes usually not accessible to P&P where the signal sits on a preexisting background.

A typical femtosecond luminescence setup is sketched in **Figure 3**. The reader will be able to recognize the delay stage already present in the P&P technique. The main difference is that the pump here does not excite the sample but is used to open an optical gate pulse; this pulse is now called the gate pulse (the gate sits between two crossed polarizers). Two main kinds of optical gates are used: the first one historically was a Kerr gate, where the polarization of the signal is changed during the presence of

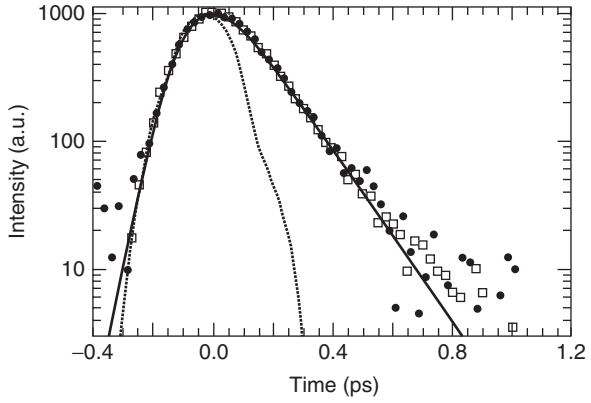


Figure 4 Upconverted luminescence signal form a diacetylene polymer chain. The chain is excited nonresonantly at 400 nm. The signal decay is very fast because of nonradiative transfer of the excitons into a triplet state (within less than 100 fs in the present case). The dotted line corresponds to the upconversion of the laser pulse, giving the time resolution of the experiment (here 70 fs) as well as a precise determination of the zero time delay. (Reproduced with permission from Haacke S, Berrehar J, Lapersonne-Meyer C, and Schott M (1999) Dynamics of singlet excitons in 1D conjugated polydiacetylene chains: a femtosecond fluorescence study. *Chemical Physics Letters* 308(5–6): 363–368.)

the gate, allowing opening this gate for a very short time only. This has the advantage of being a broadband technique (broadband meaning that the device is basically insensitive to the wavelength of the luminescence photons), but has the drawback of needing a very large pulse power for the polarization switching of the Kerr cell. The most efficient technique nowadays is the use of a nonlinear crystal, which is tuned so as to produce sum frequency photons, adding one photon from the signal and one from the gating pulse. The sum-frequency signal is only generated during the presence of the gating

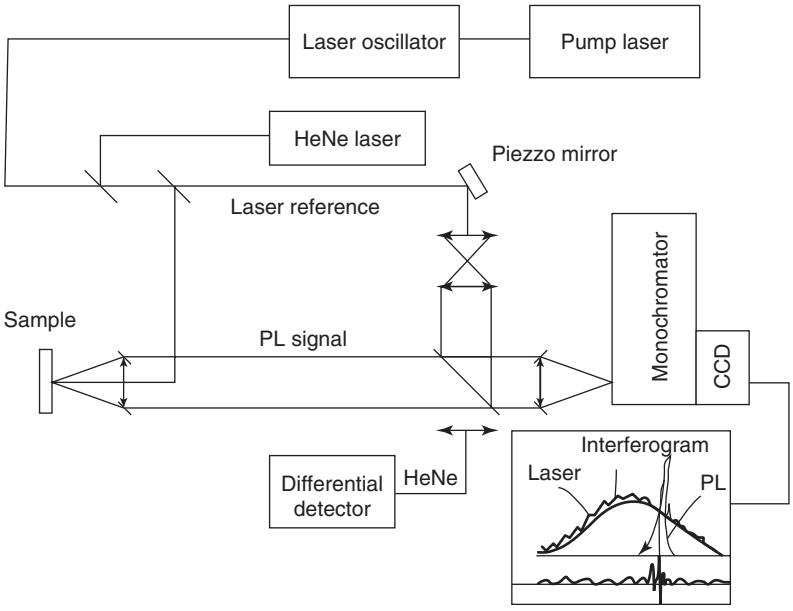


Figure 5 Schematic arrangement of a spectral interferometry experiment. Short pulses are usually created by a titanium sapphire laser oscillator. The pulses are then split into two parts within a Mach Zehnder interferometer: one part is used to excite the luminescence of the sample and the second part is used as a reference pulse. The luminescence and the reference beam are recombined at the other end of the interferometer and focused onto the entrance slit of a monochromator. One therefore observes, in the case the luminescence signal keeps some coherence with the exciting pulse, interference fringes between the luminescence signal and the reference pulse. Of course the total signal is a sum of the interference term and the two separate signals that have to be subtracted to keep only the interesting term. Being fully coherent, the interference spectrum is simply the Fourier transform of the time evolution of the coherent luminescence signal. Therefore, a reverse Fourier transform allows recovering the time behavior from the spectral behavior. Such a setup necessitates absolute stability within better than $\lambda/10$, which cannot be obtained easily with a cryostat in one of the arms of the interferometer. The whole interferometer is therefore actively stabilized through the use of a control beam coming from a HeNe laser driving a piezo mirror.

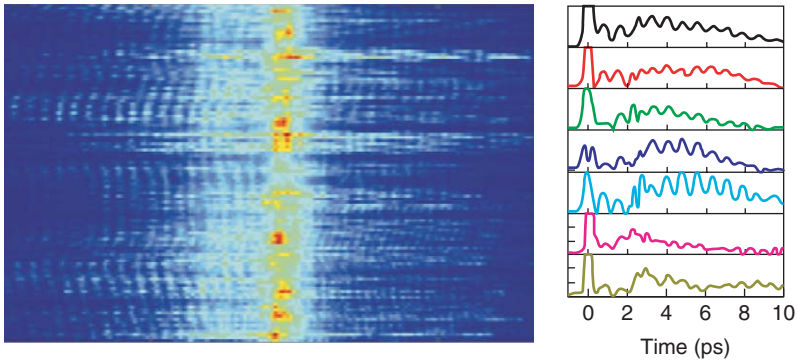


Figure 6 Spectral interferogram obtained for excitons in a GaAs quantum well. The horizontal axis corresponds to the wavelength, and the vertical axis to the angle of emission. Indeed as the signal is coherent and linked with disorder of the interfaces of the quantum well, it does give rise to speckles. The time behavior, and therefore the interferogram are different for each speckle as evidenced through the results of Fourier transformation. (Reproduced with permission from Hayes G, Deveaud B, Savona V, and Haacke S (2000) Speckle-averaged resonant Rayleigh scattering from quantum-well excitons. *Physical Review B* 62(11): 6952–6955.)

pulse, and is spectrally well separated from the luminescence photons. Once again, the time resolution of the experiment is given by the length of the gate. One difficulty, in femtosecond luminescence by upconversion, is that the wavelength of the signal and that of the pump may be very different. It is therefore very important to avoid any dispersion on the path of the signal. Reflective optics will be preferred to the simpler refractive systems, which may easily distort the timing of the signal by as much as a picosecond.

A typical luminescence signal is shown in **Figure 4** showing that, indeed, the time resolution of the experiment can be limited by the pulsewidth down to the shortest times. In this case, the short luminescence signal is just a measure of the very short lifetime of the excited state of the polymer.

Spectral Interferometry

Spectral interferometry is a heterodyne detection technique. It allows obtaining spectral and temporal information on the coherent dynamics of an optically excited system. The schematics of the experiment are displayed in **Figure 5** and shows that the properties of the signal are obtained by correlating these with the exciting beam in a Mach–Zehnder-type interferometer. The two overlapping beams are spectrally filtered in a monochromator, which Fourier transforms the time correlation into spectral oscillations (see **Figure 6**). The time behavior of the coherent part of the signal can be recovered by taking the Fourier transform of the spectral interference figure. This technique is unique in the sense that it probes the coherence of the system through a fully linear technique. It has however a major drawback which is that the interferometer has to be stabilized within much better than the wavelength of light, and that this is a large size interferometer usually incorporating a cryostat. The stabilization scheme is therefore much more difficult than for usual interferometers.

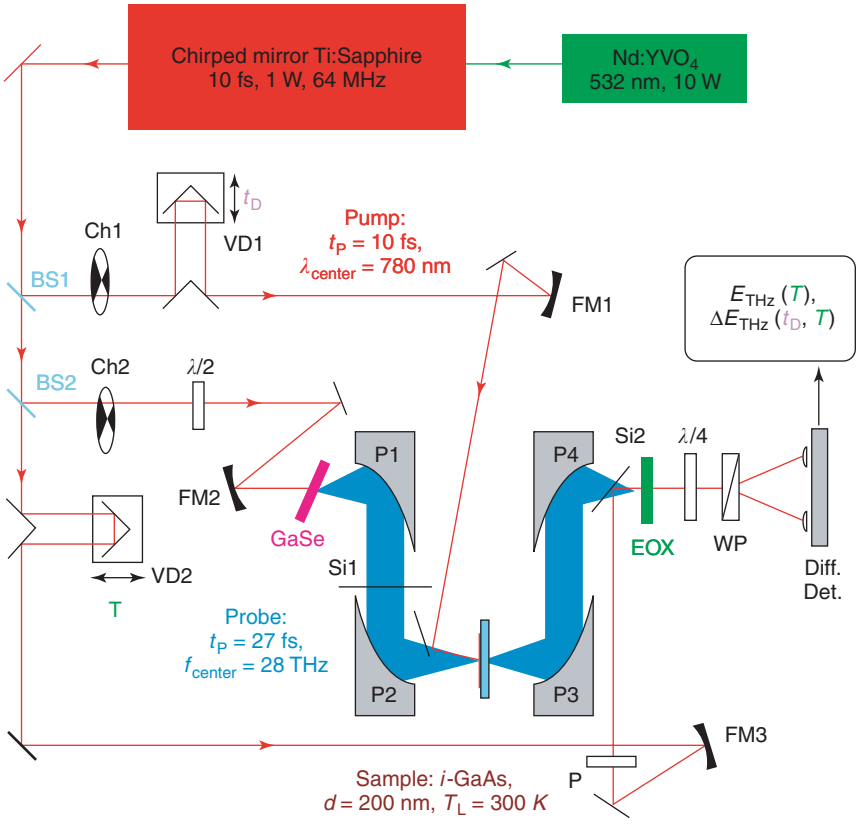


Figure 7 Schematic arrangement of a teraHertz P&P experiment. In the optimal setup, 10 fs pulses are generated by an optimized oscillator. Focusing of such short pulses onto a $GaSe_2$ crystal allows producing a single-cycle teraHertz pulse covering a large frequency range, basically extending from 1 up to 40 THz. The principle of the experiment is similar to standard P&P, which means that a pump pulse excites the sample, which is then probed after some delay by the teraHertz pulse. The detection of the changes occurring in the teraHertz pulse is performed via the use of a gating pulse onto a photoconductive switch. This has the major advantage of allowing recording the changes in the electric field of the teraHertz pulse, rather than its intensity. As a result, one may access both the changes in the real and imaginary parts of the susceptibility of the sample after the pump has hit the sample. (Courtesy A Leitenstorfer, with the permission of the author.)

TeraHertz Probing

One emerging field in ultrafast spectroscopy is very similar to P&P schemes, but utilizes as a probe teraHertz pulses. Such techniques require still rather complex setups, but are very promising in terms of the amount of information one can get. One possible design of the experiment can be found in **Figure 7**, and one can readily appreciate this complexity. This is basically a P&P system, the probe being generated by proper focusing of a laser beam onto a GaSe crystal. The transmitted teraHertz probe needs however another gate pulse in order to time resolve the variations of the probe pulse. Basically, on an experimental side, this adds up the complexity of the usual P&P techniques and of upconversion luminescence. As a result, it is possible to obtain both the real and the imaginary part of the dielectric constant as a function of time. Such a very interesting result is depicted on **Figure 8**, which evidences the build-up of screening in an optically excited semiconductor.

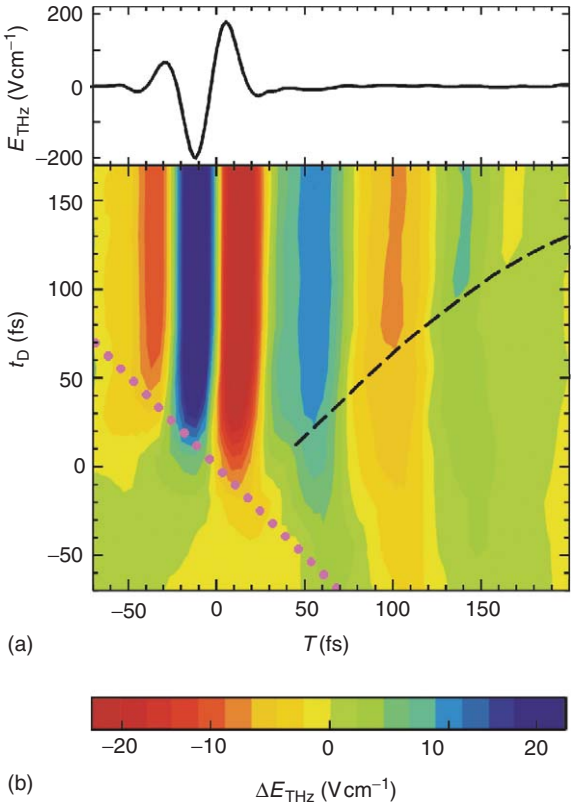


Figure 8 TeraHertz pulse changes after traveling through a sample of GaAs excited with a strong interband pulse creating a high-density plasma. Evidence for the build-up of manybody interactions is obtained from such an experiment. This is a two-time chart, T axis displays the changes in the electric field in the teraHertz pulse as it passes, the t_D axis corresponding to the delay between the arrival of the interband pulse and the teraHertz pulse on the sample. The color code corresponds to the electric field in the teraHertz pulse. (Reproduced with permission from Huber R, Tauser F, Brodschelm A, Bichler M, Abstreiter G *et al.* (2001) How many-particle interactions develop after ultrafast excitation of an electron-hole plasma. *Nature* 414(6861): 286–289.)

Nonlinear Techniques

Although P&P techniques rely on the nonlinearities of the system to obtain some signal, they are not primarily interested in getting at a measure of these nonlinearities. Nonlinear properties of optically active materials can also be used to probe the dynamics of the system, as well as to test their own dynamics. The most usually employed technique to this aim is four-wave mixing which allows a wide range of experimental configurations, each probing specific aspects of the dynamics of the system, all being labeled under the generic term "four-wave mixing" because the description of the nonlinearities require at least four photon fields. In principle, four-wave mixing relies on the creation of a polarization grating in the material with the first two pulses, and then on the diffraction of a third beam on this grating, giving rise to the fourth, wave mixing signal, highly sensitive to the nonlinearities of the system.

Two-Beam Self Diffraction

In the two-beam self diffraction configuration, only two incident beams arrive on the sample (see **Figure 9**). The first beam, with a direction corresponding to the wave vector $\boldsymbol{k}_1$, creates a polarization in the sample. The second beam, arriving after some delay τ with a direction corresponding to the wave vector $\boldsymbol{k}_2$, first generates a grating by interaction with the polarization left from the first pulse, and simultaneously self diffracts on this grating it has just created. This technique is a very simple way to get information about the coherence dynamics of the system.

The nonlinear signal, detected in the $2\boldsymbol{k}_2-\boldsymbol{k}_1$ direction does not need to be temporally resolved. In simple cases, it may simply be detected by a slow detector, leading to the signal behavior depicted in the inset of **Figure 9**. The decay of the nonlinear signal corresponds directly to the dephasing of the coherence of the polarization created by the first pulse, as the grating will not be generated if the initial polarization has been dephased. More complex systems may lead to more complex behaviors such as for example a signal at negative times, that is, when the probe arrives before the pump has even hit the sample. The study of such unexpected behavior has brought major improvements in our understanding of the nonlinearities in semiconductors.

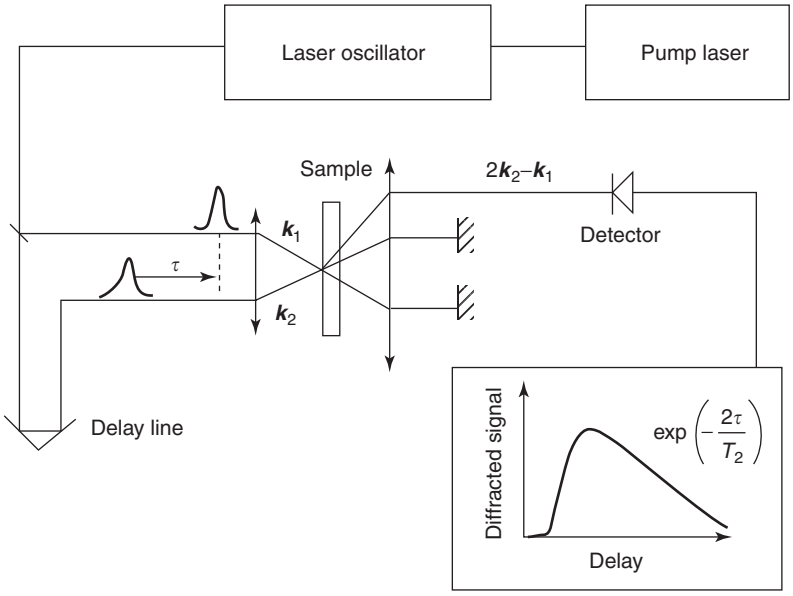


Figure 9 Schematic arrangement of a two-pulse four-wave mixing experiment (FWM). Short pulses are usually created by a titanium sapphire laser oscillator. These pulses may be used directly to perform the experiment, leading to a measurement bandwidth limited by the Fourier transform limited width of the pulses. Often, a better spectral resolution is needed, which can be obtained by pulse-shaping technique (basically filtering of the pulse) at the expense of the time resolution, as imposed by the Fourier transform of the spectrum. The first pulse of wave vector $\boldsymbol{k}_1$ is focused on the sample. The second pulse of wave vector $\boldsymbol{k}_2$ is delayed by a delay stage and focused together with the first pulse. The changes in the diffracted signal intensity, in the direction $2\boldsymbol{k}_2-\boldsymbol{k}_1$, are recorded as a function of the delay between the two pulses. The signal usually shows a clear asymmetry to long delays, the decay being exponential with a decay constant corresponding to half the dephasing time of the polarization excited in the sample.

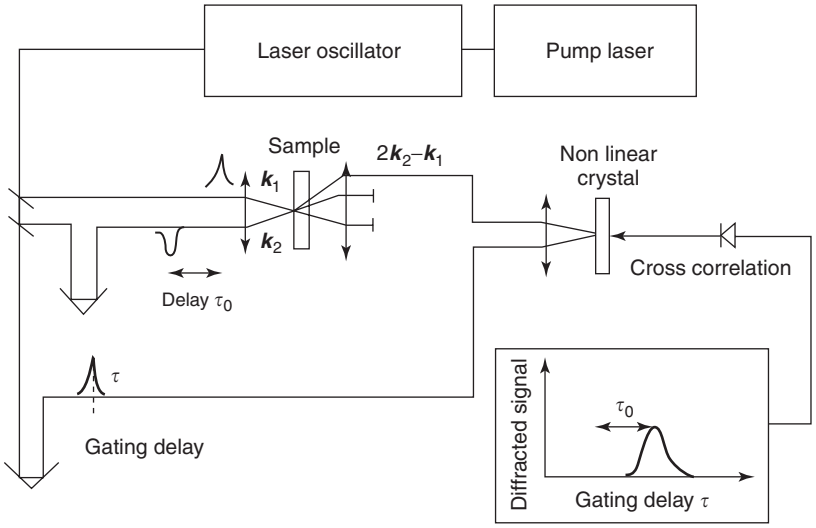


Figure 10 In the case of a resonance that is inhomogeneously broadened, the diffracted signal appears in the form of a photon echo. The diffracted pulse is emitted by the sample after the second pulse, with a delay corresponding to the delay between the two interfering pulses. Once again the intensity decays with the dephasing of the polarization, however, the decay constant is now a factor 1/4 of the dephasing time. The discrimination between homogeneous and inhomogeneous dephasing may be obtained from the time resolution of the diffracted signal, which corresponds to the addition of a gate similar to that used in upconversion luminescence.

Photon Echo

If the system would be homogeneously broadened, the integrated intensity of the four-wave-mixing signal would directly give some information on the coherence time of the polarization in the system. If, on the contrary, the system is inhomogeneously broadened, the nonlinear signal is delayed after the arrival of the second pulse. Basically, this diffracted

signal appears at a delay equal to the time separation between the first two pulses. This is a direct signature of the inhomogeneous nature of the system under study, which cannot be observed directly with the previous setup. Such an echo signal can easily be recorded by time resolving the diffracted signal. In such a case, one gets both the information about the broadening of the system, and also on the dephasing because the signal intensity decays with a characteristic time given by the coherence properties of the material. This technique gives of course more information on the system but necessitates an optical gate similar to the one described in the paragraph devoted to ultrafast luminescence (see also **Figure 10**).

Three-Beam Four-Wave Mixing

In the three-beam configuration, the two beams that create the grating and the probe beam all have different directions $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, and $\boldsymbol{k}_3$, which will generate a phase conjugate signal in the direction $\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3$. This configuration is very rich as it allows playing with any combination of delays between the different pulses. For example, if the first two pulses arrive at the same time, and the third one is delayed the situation is very similar to the self-diffraction technique, with a clear asymmetry in the process that allows distinguishing in principle homogeneous dephasing from inhomogeneous dephasing. The other configurations are very useful as, depending on the arrival sequence of the pulses on the sample, dephasing events of the ground state or of the excited state of the system can be probed selectively.

Measurement of the Pulse

Most of the techniques that have been described up to now implicitly assume that the ultrashort pulses are either perfectly characterized, or at least well behaved. One of the advantages of the Ti:Sapphire laser is that it is indeed reasonable to assume in most cases that the pulse is indeed close to a transform-limited pulse (specially in the case of a pulse duration of the order of 100 fs, in the near infrared region). When one wants to use more exotic pulses, exotic in the sense that they might be in a different wavelength window, or that they have a shorter duration, things become more complex. In particular, it becomes very difficult to keep a constant phase of the different Fourier components of the pulse over the entire wavelength range. A 10 fs pulse extends basically from 700 up to 900 nm. It becomes then necessary to characterize the pulse, the difficulty being that simple characterization schemes would need to use a pulse much shorter than the pulse to be characterized, which obviously cannot be the case anymore.

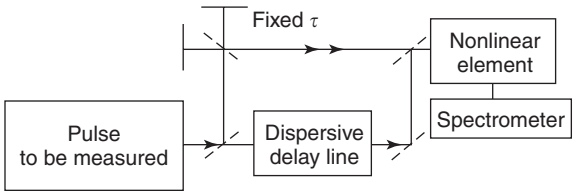


Figure 11 Sketch of the pulse measurement techniques called SPIDER or FROG. The pulse to be measured is split into two parts in a Mach Zehnder interferometer. One of the split pulses is delayed with respect to the other, and they are recombined at the surface of a nonlinear crystal. A monochromator and CCD allow recording the spectral interferogram, which allows retrieving the full information of the pulse via proper computer adjustment of the observed interferogram.

New techniques have been devised which have been given strange acronyms such as FROG (for Frequency Resolved Optical Gating) or SPIDER for (Spectral-Phase Interferometry for Direct Electric-field Reconstruction). They are somewhat based on the same principle, which is sketched in **Figure 11**, which consist in using the pulse itself as a reference for the measurement. This kind of technique is well known in the case of longer pulses, where the pulse is easier transform limited, and where the length of the pulse is the most interesting information one is trying to retrieve. Then the pulse is split into two parts, traveling different length, and recombined in front of a nonlinear crystal. The doubling efficiency is measured as a function of delay, giving rise to a symmetrical figure, the autocorrelation trace, directly related to the pulse length for the case of symmetrical pulse. In the case where the pulse would not be symmetrical, the measurement of the intensity of the frequency-doubled light is not sufficient. Basically what one needs to perform is a spectral interferometric measurement, which gives much more information than only the intensity. However, even if the spectral information contains much more than the integrated intensity, this information is not enough to retrieve directly the whole electric field, because it is obtained through autocorrelation of the pulse. The full information can however be retrieved through proper adjustment of the observed spectra, which is nowadays done through self-adjusting computing techniques.

See also: Electrons and Holes; Excitons: Theory; Harmonic Generation Frequency Conversion; Interaction of Light and Matter; Luminescence; Nonlinear Optics; Organic Semiconductors; Semiconductors, General Properties.

PACS: 78; 78.35.+c; 78.40.−q; 78.30.−j; 78.47.+p; 78.55.−m; 78.67.−n

Further Reading

Bigot JY and Deveaud-Plédran B (2001) Laser sources and spectroscopies femtoseconds: current tendencies. *Reports of the Academy of Science*. CRAS series IV, Volume 2, No. 10, pp. 1514. Paris: Elsevier.

Mukamel S (1999) Principles of nonlinear optical spectroscopy. *Oxford Series on Optical and Imaging Sciences*, 6. Oxford: Oxford University Press. ISBN 0.19.513291.2.

Shah J (1996) Ultrafast spectroscopy of semiconductors and semiconductor nanostructures. *Springer Series in Solid State Sciences, No. 115*. Berlin: Springer.

Transistors

Y Hirayama, NTT Basic Research Laboratories, Kanagawa, Japan

Introduction

A transistor is a three-terminal device, where the voltage or current applied to one terminal controls the current flow between the other two terminals. Transistors are designed to have current or voltage gain, which means a drivability of multistep transistor circuits. Submicron lithography and self-alignment have been used to improve the high-frequency operation of transistors. Transistors are key devices in cutting-edge semiconductor integrated circuits.

There are mainly two types of transistors, bipolar and field effect transistors. Transistors have been made on many systems, mainly semiconductor systems, including Si, SiGe, and III–V semiconductors. Since this encyclopedia concerns condensed matter physics, the basic principles of transistor operation and the fundamental performance of transistors, with emphasis on III–V semiconductor transistors, are discussed. Although the basic principles can be applied to many different systems, detailed discussions of transistor structures, transistor materials, processing, and circuit designs are omitted due to space constraints. (See "Further reading" section for details.)

Transistors, that is, three-terminal control devices, are important not only for electrical applications but also for physics studies. Several transistor devices often used in the study of transport physics at low temperatures are also discussed in the second half of this article.

Bipolar and Hetero-Bipolar Transistors

The principal operation of the (hetero) bipolar transistor is schematically shown in **Figure 1**. Although a bipolar transistor with *n*-type emitter, *p*-type base, and *n*-type collector (*n–p–n* transistor) is assumed in this figure, similar parameters are applicable to a

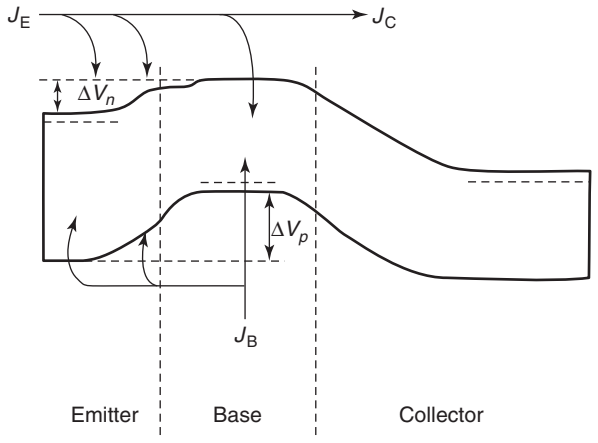


Figure 1 Schematic band diagram of a HBT with graded emitter-base junction. The schematic current flow, which determines the transistor operation, is also shown.

p–n–p transistor. In the bipolar transistor, the most important factor determining transistor operation is current gain, that is, the ratio between J_C and J_B, where J_C and J_B are corrector current and base current, respectively. When this ratio becomes larger than unity, the transistor has a gain. The current injected from the emitter passes through the base and reaches the collector. When the diffusion constant of electrons in the *p*-type base is small or the base layer is thick, the probability of electron transmission to the collector decreases. Some amount of the base current also diffuses to the emitter. Calculating all these factors results in the current gain h_{fe} determined by the emitter injection efficiency:

$$h_{fe} = \frac{n_E}{p_B}\frac{v_e}{v_h}\frac{N_{CB}}{N_{CE}}\frac{N_{VB}}{N_{VE}}\exp\left(\frac{\Delta E_g}{KT}\right) \qquad [1]$$

where n_E and p_B are, respectively, the emitter and base doping concentration, v_e (v_h) is electron (hole) effective velocity toward the collector (toward the emitter) at the corrector-side (emitter-side) base edge, and, N_{CB} and N_{CE} (N_{VB} and N_{VE}) are densities of states in the conduction (valence) band for base and emitter, respectively. $\Delta E_g = E_{ge} - E_{gb}$ is the bandgap difference between emitter and base semiconductors.

For a thick emitter, v_h is determined by $v_h = D_h/L_E$, where D_h is the hole diffusivity in the emitter region and $L_E = \sqrt{D_h \tau_E}$ is the diffusion length. τ_E is the recombination lifetime in the emitter region. In the case of the thin base with $W_B \ll L_B$, v_e is represented as

$$v_e = D_n/(L_B \sinh(W_B/L_B)) = D_n/W_B$$

where D_n, L_B, and W_B are the electron diffusivity in the base region, electron diffusion length in the base, and the base thickness, respectively. In this situation,

$$h_{fe} = \frac{n_E}{p_B}\frac{L_E}{W_B}\frac{D_n}{D_h}\frac{N_{CB}}{N_{CE}}\frac{N_{VB}}{N_{VE}}\exp\left(\frac{\Delta E_g}{kT}\right) \qquad [2]$$

In the case of a homojunction bipolar transistor, such as an Si bipolar transistor, $\Delta E_g = 0$, and h_{fe} is mainly determined by $n_E/p_B/W_B$. In order to maintain a suitable high value of h_{fe}, $n_E/p_B \sim 100$–1000 is typically chosen. The hole density of the base, p_B, usually becomes 10^{17}–$10^{18}\,cm^{-3}$ for the emitter electron density, n_E, of 10^{20}–$10^{21}\,cm^{-3}$. The current gain also increases with decreasing base layer thickness. However, the base resistance increases with decreasing p_B and W_B, so that there is a limitation to how far these parameters can be decreased in actual transistors. Si bipolar transistors are widely used in high-speed semiconductor circuits, owing to the many improvements in techniques for forming the emitter and base junctions, in isolation methods, and in the degree of self-alignment between the different device regions. Silicon is an indirect bandgap semiconductor and has a long recombination lifetime, so that the situation $W_B \ll L_B$ and a low-base resistance can be achieved even for a small p_B. However, direct bandgap semiconductors, such as GaAs, have a short recombination time, and the high n_E/p_B ratio necessary for the high-current gain is not possible. Therefore, it is difficult to build high-performance homojunction bipolar transistors from III–V semiconductors.

When heterojunctions are used, $\exp(\Delta E_g/kT)$ makes the most important contribution to h_{fe}. Improved transistor performance is expected by combining silicon-based transistors and Si/SiGe heterojunction systems. Furthermore, the large bandgap difference in III–V semiconductors enables one to use direct bandgap semiconductors in heterobipolar transistors (HBTs). This is because the term $\exp(\Delta E_g/kT)$ sufficiently compensates for the small n_E/p_B, and a thin base layer with high-carrier density becomes available. HBTs using many kinds of III–V semiconductor materials have been demonstrated.

At an abrupt heterojunction between different semiconductors, there is a discontinuity in the energy of the conduction-band minima and valence-band maxima. An energy barrier in the conduction band at the emitter–base junction tends to retard the flow of electrons from the emitter to the base. To avoid the barrier in the conduction band, a graded-alloy composition at the emitter–base junction has been used as shown in **Figure 1**. However, it is noteworthy that the barrier at the emitter–base junction is advantageous for high-speed operation because of the higher-energy electron injection from the emitter to the base. Therefore, an optimal design of the graded junction is necessary in actual HBTs.

For HBTs, many kinds of heterostructures have been studied. **Figure 2** shows a comparison of the breakdown electric field of HBTs composed of Si/SiGe, InGaP/GaAs, InP/InGaAs, and GaN/InGaN material systems. The breakdown electric field of nitride-based HBTs is much higher than those reported for HBTs composed of other material systems. Therefore, GaN-based HBTs have recently received much attention as candidates for high-power and high-speed devices in the future. Here, the typical structures and characteristics of HBTs are discussed using GaN-based HBTs as examples. **Figure 3** illustrates the schematic structure of the GaN/InGaN HBTs. In this peculiar device, the graded region is formed between the base and collector to reduce electron scattering at the collector barrier, because the same bandgap material (GaN) is used for both the emitter and corrector. A regrowth InGaN base contact is adopted to decrease the series resistance of the base. Actually, how to make a good contact to a thin base layer is a very important issue for HBTs with a thin base layer, and efforts to resolve this issue are underway for many heterostructure HBTs. The common emitter current–voltage characteristics of GaN/InGaN HBT with a $50\,\mu m \times 30\,\mu m$ emitter area are shown in **Figure 4**. The maximum collector gain exceeds 2000

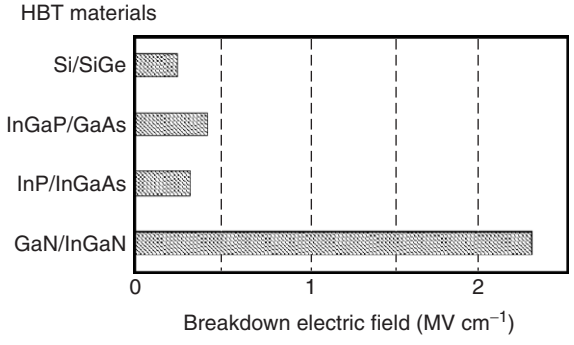


Figure 2 Comparison of breakdown electric field of HBTs composed of various material systems.

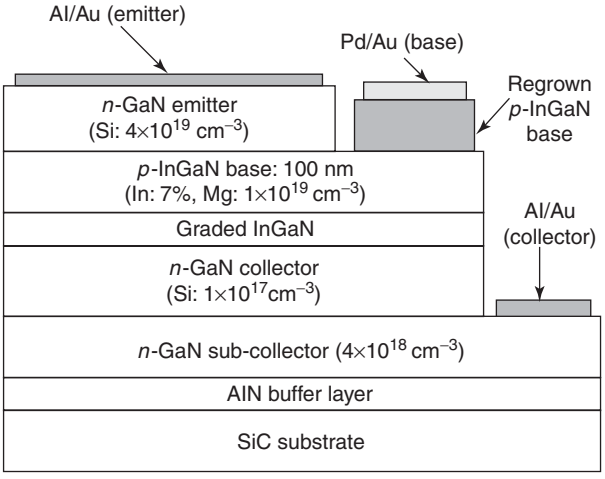


Figure 3 GaN/InGaN HBT structure with a regrown *p*-InGaN extrinsic base layer.

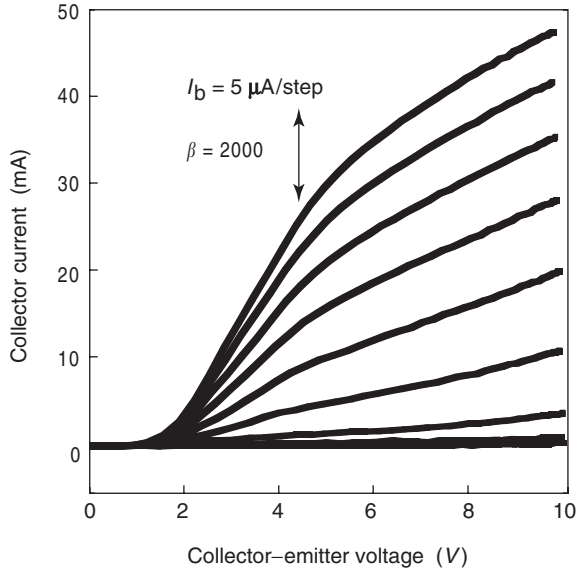


Figure 4 Common-emitter current–voltage characteristics of a GaN/InGaN HBT at room temperature. Base current is 5 μA/step.

in this device. Although this current gain is still less than the maximum gain of 5000 for Si/SiGe systems and 50 000 for InP-based systems, the current gain for GaN-based systems has increased almost two orders of magnitude from the year 2000.

Field Effect Transistors: MISFET, MESFET, and HEMT

Bipolar transistors use both majority and minority carriers in semiconductors. There is, however, another type of transistor, where only the majority carriers are used and their density is controlled by an electric field effect. This type of transistor is called a field effect transistor (FET) and is widely used in many semiconductor systems. Several types of FETs are schematically shown in **Figure 5**. Metal–insulator–controlled semiconductor FETs (MISFETs) are shown in **Figure 5a**. When a positive bias is applied to the gate, conductive electrons start accumulating at the semiconductor–insulator interface. Due to the confinement by the electric field, the accumulated electrons have a two-dimensional (2D) nature. The threshold voltage, V_{th}, where electrons start accumulating, is determined by the semiconductor properties and interface characteristics. The density of accumulated charge, Q, is proportional to the gate bias and is represented as

$$Q = ne = C_g(V_g - V_{th}) \quad [3]$$

where n, e, and C_g are the area density of electrons, electron charge, and capacitance of the gate insulator, respectively. When the bias is applied between the source and drain (V_d), the drain current (I_d) is determined by the carrier density (n) and the mobility of the carrier (μ). In the case of small V_d, where carrier density is constant over the whole device,

$$I_d = \frac{Z}{L}\mu ne V_d \quad [4]$$

where Z and L are the channel width and length of the MISFET. In the case of the MISFET, ne is expressed by eqn [3], then

$$I_d = \frac{Z}{L}\mu C_g(V_g - V_{th})V_d \quad [5]$$

With increasing V_d, the carrier density near the drain edge decreases and finally becomes zero (pinch-off) at around $V_d = V_g - V_{th}$. Further increasing V_d results in an almost constant drain current independent of V_d, because a balance is achieved between the pinch-off of the carrier and the high electric field concentrating in the pinch-off region. Therefore, I_dV_d characteristics of MISFETs, schematically shown in **Figure 6**, are separated into linear and saturated regions. The drain current in the saturated region, I_{ds}, is approximately expressed as

$$I_{ds} = \alpha\frac{Z}{L}\mu C_g(V_g - V_{th})^2 \quad [6]$$

where α is a constant close to unity.

In the case of Si, silicon dioxide (SiO_2) is used as the gate insulator. Therefore, the FET is normally called metal-oxide–semiconductor FETs (MOSFETs). It is also possible to accumulate conductive holes by applying negative bias to the gate. Combining *n*-type

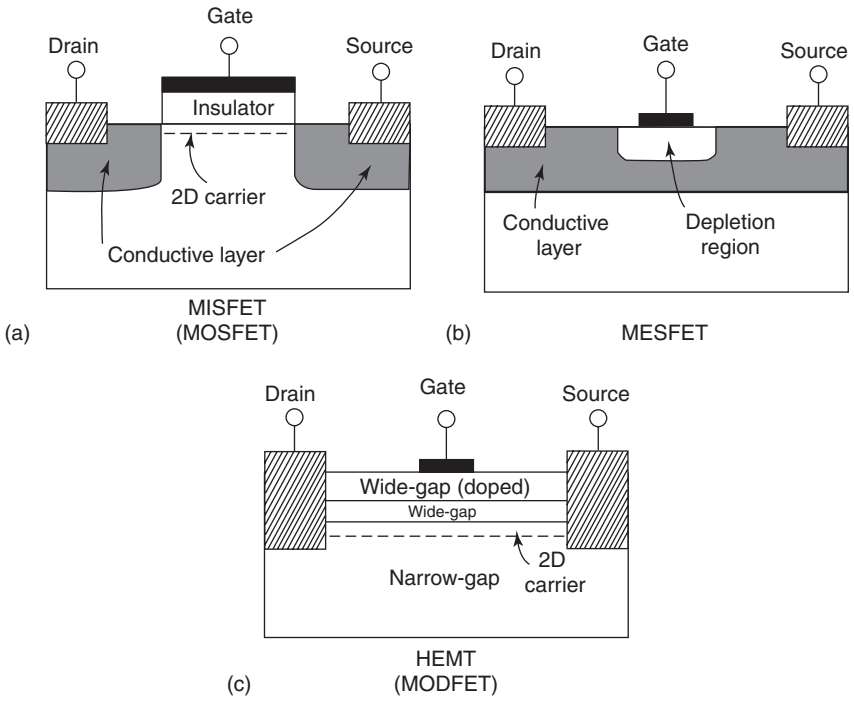


Figure 5 Schematic diagrams of field effect transistors (FETs). (a) MISFET (MOSFET), (b) MESFET, and (c) HEMT (MODFET).

(electron transport) MOSFETs and *p*-type (hole transport) MOSFETs results in complementary MOSFET devices (CMOS devices). The Si MOSFET and CMOS devices are the most important devices in the very large scale integrated circuits used in microprocessors and semiconductor memories.

In MISFET devices, some other combinations of insulators and semiconductor materials are also possible. A good example is the InP MISFET, where electrons can accumulate at the interface between the InP and insulator, such as Al_2O_3. However, the conductive carriers cannot be accumulated at the interface when the density of states of the interface defects is very large, as is the case at insulator–GaAs interface.

To avoid a direct and uncontrollable leakage current between the drain and source, *n*-type (*p*-type) MISFETs are fabricated on *p*-type (*n*-type) or semi-insulating materials. For the purpose of distinguishing these two types of MISFETs, the former is called an inversion-type MISFET and the latter an accumulation-type MISFET. For both types, the gate-induced carriers need to contact the source- and drain-conductive regions (doped regions) so that an overlap between source- (drain-) conductive regions and gate is essential for the FET operation. The self-alignment process allows one to make such an overlap without increasing the capacitance between the gate and source (drain).

Another type of FET is the metal–semiconductor FET (MESFET). **Figure 5b** is a schematic diagram of a MESFET. In this device, conductive carriers are supplied from the 3D doping layer, and the number of conductive carriers is controlled by a depletion region spreading from the Schottky metal gate. The

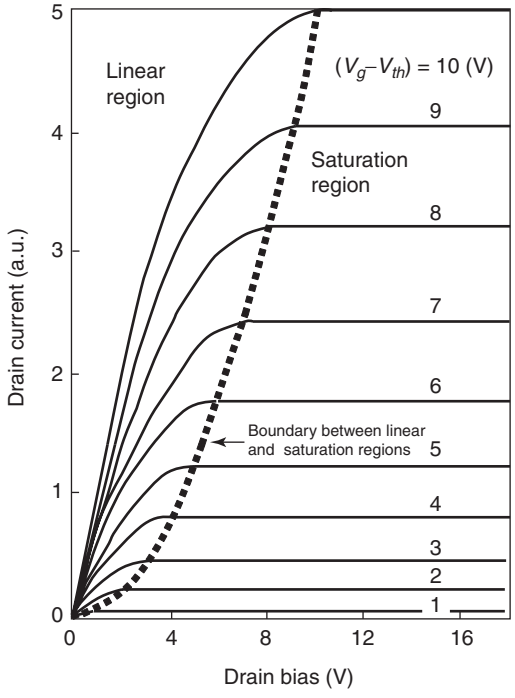


Figure 6 Schematic drain current-drain bias characteristics of a MISFET (MOSFET). Actual value of the drain current depends on device parameters and carrier mobility.

conductive layer thickness is determined by subtracting the depletion region thickness (W_d) from the total thickness of the doping layer. For the *n*-type conducting channel,

$$W_d = \sqrt{\frac{2\varepsilon(V_{bi} - V_g)}{eN_d}} \qquad [7]$$

where V_{bi}, ε, and, N_d are the built-in potential of the Schottky gate, dielectric constant of semiconductor, and donor density of the doping layer, respectively. The area density of the conductive carrier can be easily estimated from the conductive layer thickness, and the drain current is determined by eqn [4] in small V_d regions. Forming the conductive layer involves several techniques, such as ion-implantation and epitaxial growth.

Finally, there is the high-electron-mobility transistor (HEMT) which is shown in **Figure 5c**. The HEMT utilizes conductive carriers accumulated at the hetero-interface between wide- and narrow-gap semiconductors. A triangular-shaped confinement potential at the interface defines the conductive carrier in a thin 2D channel. Modulation doping, that is, doping in the wide-gap semiconductor and conductive carriers at the interface, enables one to separate conductive carriers from dopant impurities, which results in a strong suppression of impurity scattering. Therefore, HEMTs are also called modulation-doped FETs (MODFETs). The density of the 2D electron system is controlled by the gate bias and the I_dV_d characteristics are roughly represented by eqns [3]–[6]. The most widely studied HEMT devices are based on the AlGaAs/GaAs system; however, many different material combinations, such as InAlAs/InGaAs and AlGaN/GaN systems, have been developed recently.

High-frequency FETs are key devices in recent optical communications technology, where they enable the practical use of 40 Gbit s^{-1}. **Figure 7** summarizes current gain cut-off frequencies (f_T's) for various kinds of FETs. Among them, InP-based HEMTs, where a conductive two-dimensional electron gas (2DEG) channel is formed in the InAlAs/InGaAs heterostructures lattice-matched with InP substrate, have achieved record f_T's for a wide range of gate lengths (L_g's) by many research groups. The effective saturation velocity of InAlAs/InGaAs HEMTs is 2.7×10^7 cm s^{-1}, $\sim$1.3 times higher than that of GaAs-based HEMTs. Moreover, higher electron mobility and sheet-carrier density of the 2DEG at the InAlAs/InGaAs hetero-interface reduce channel and parasitic charging times, and make possible the effective use of higher saturation velocity. The current record f_T for a discrete device is beyond 400 GHz.

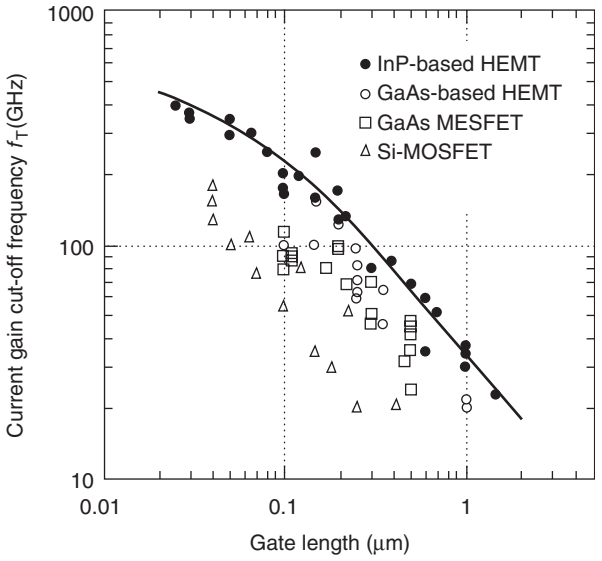


Figure 7 Current gain cut-off frequencies for various kinds of FETs.

Integrated circuits (ICs) based on InP have been developed with the aim of achieving a 100 Gbit s^{-1} class communications technology. The schematic structure of such an IC is shown in **Figure 8**. A 0.1 μm-gate InAlAs/InGaAs HEMT is integrated with InAlAs Schottky diodes, metal resistors, and metal-insulator-metal (MIM) capacitors in this example. The mushroom-shaped gate improves transistor performance.

Modulation-doped HEMT structures are important not only as cutting-edge room-temperature devices but also as key devices for studies of low-temperature physics. The 2D electron system at the AlGaAs/GaAs interface has the highest mobility at low temperatures, and has been very important in quantum Hall effect research. Mobility up to 10^7 cm^2 V^{-1} s^{-1} has been obtained in several laboratories in the USA, Europe, and Japan. The mobility is still limited by background impurities; therefore, it can be improved by increasing electron density. Recently, Bell Laboratories used a sophisticated MBE technique and two-side modulation-doping, and achieved mobility of around 3×10^7 cm^2 V^{-1} s^{-1}.

Backgated Structures

The backgated structure, where carrier density is modulated from the back side by a voltage applied to the backgate, is also attractive. Although the gate capacitance limits high-frequency applications, the density control from the back side enables one to use the free surface for many kinds of structures. When

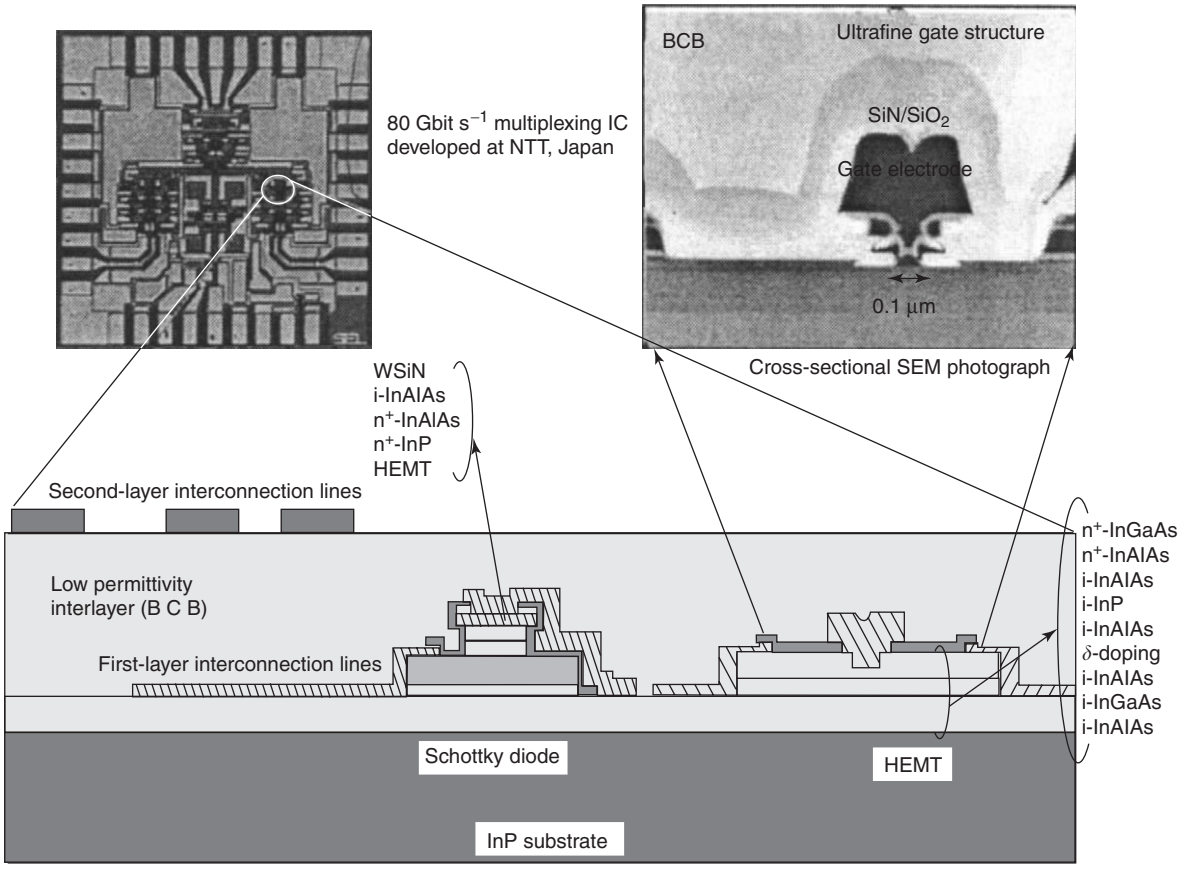


Figure 8 Schematic diagram of an InP-based HEMT IC.

fine structures are formed on the surface side, the structure behaves as a density-tunable mesoscopic system.

An example of a backgated heterostructure is shown in **Figure 9**. In this example, the 2D electron system accumulates at the hetero-interface between AlGaAs and GaAs. Conductive electrons are supplied from the ohmic contacts as in Si MOSFETs. The threshold voltage of the 2D electron accumulation is determined by the balance between surface charge and backgate bias. In case of the GaAs, the surface charge is frozen at low temperatures and the surface Fermi level is not pinned at the midgap. These features help in accumulating 2D electron systems with a small threshold voltage at low temperatures. Once the 2D electron system is accumulated at the hetero-interface, its density can be precisely controlled by the backgate bias, that is, an electric field applied between the 2D electrons and the backgate.

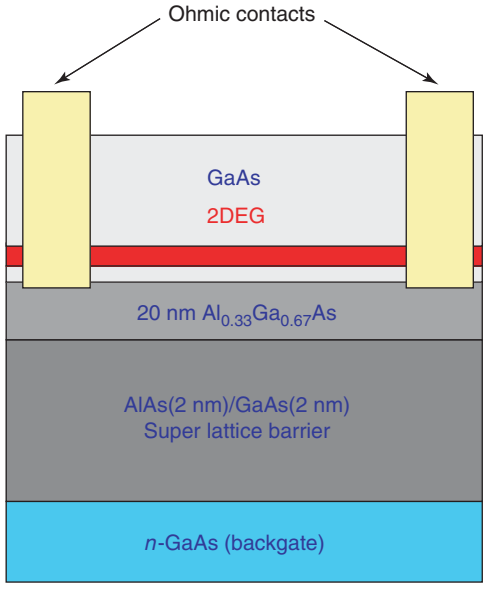


Figure 9 Example of a backgated heterostructure.

In-Plane-Gate Transistor

Transistors, that is, three-terminal devices, have been applied to confine electron systems to low dimensions, such as one and zero dimensions. When a pair of split Schottky gates is formed on a semiconductor heterostructure, such as AlGaAs/GaAs, the 2D electron system under the gates is depleted by

a negative gate bias applied to the Schottky gates (see **Figure 10a**). The small distance between the two gates results in a narrow electron channel, which shows clear 1D features at low temperatures.

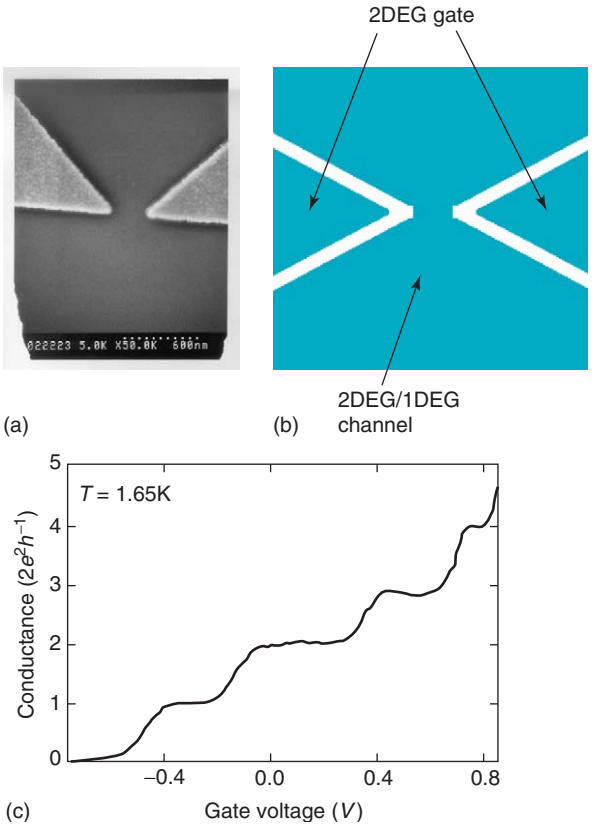


Figure 10 Quantum point contacts (QPCs) fabricated by the FET technique. (a) Split Shcottky gate QPC, (b) in-plane-gate QPC, and (c) typical characteristics of an in-plane-gate QPC fabricated on InSb quantum-well structure.

Quantized conductance characteristics have been observed for such narrow and ballistic 1D channels, which are called quantum point contacts.

A unique transistor structure, called the in-plane-gate transistor, is formed by separating a 2D electron system into three parts using insulated lines as shown in **Figure 10b**. These insulating lines are made by trench etching or ion implantation. When a negative bias is applied to two side gates against the center region, the depletion region spreads from the insulating lines and a 1D channel is formed in the center channel. As shown in **Figure 10c**, quantized conductance characteristics have been reported, reflecting the successful formation of such a channel. Although the in-plane-gate FET is limited to a 1D system from the operation principle, three-terminal structures are made without Schottky gates. Therefore, the in-plane-gate FET is beneficial for many narrow-gap semiconductors, where the formation of a good Schottky gate is difficult.

Single–Electron Transistor

In FETs, the number of carriers in the channel is determined by the gate bias through capacitance coupling. By combining FETs and nanotechnologies, it becomes possible to realize the structure schematically shown in **Figure 11a**. There is a small channel, like a quantum dot, which is separated from the source and drain by two barriers. The barrier height is set high enough to protect normal current flow, but low enough to permit tunneling coupling between the quantum dot and source (drain). The charge in the quantum dot increases linearly with the gate bias as represented by $Q = C_g V_g + Q_0$, where Q is the

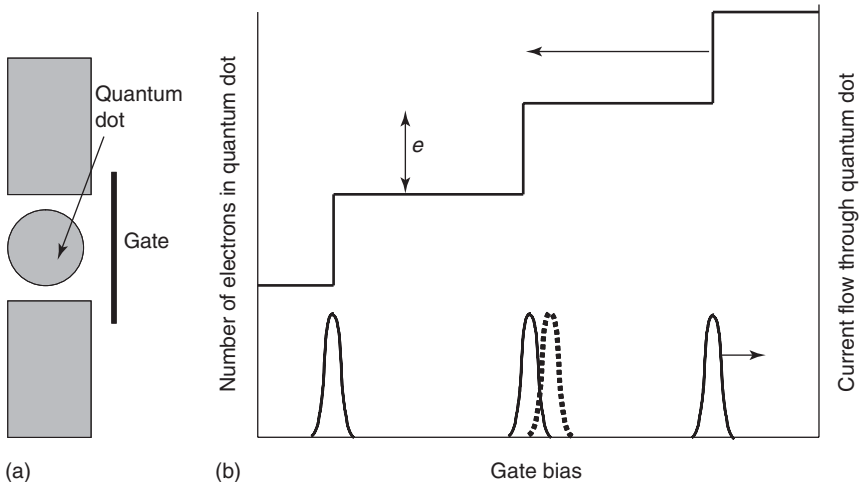


Figure 11 (a) Schematic diagram and (b) schematic operation principle of an SET. The position of the conductive peak is extremely sensitive to the charge balance in the circumstances schematically shown by the dotted lines.

charge in the quantum dot, C_g gate capacitance, V_g gate bias, and Q_0 the background charge determined by the balance of the quantum dot in the system. When electrons accumulate in the quantum dot, Q should be an integer multiple of e, the charge of a single electron. This discrete nature of the charge is not important for a large system, but becomes essential for a small system such as the quantum dot, where $\Delta V_g = e/C_g$ becomes larger than the thermal broadening, kT/e. This situation is fully satisfied for the small quantum-dot system at low temperatures, and such three-terminal structures are called single-electron transistors (SETs).

A schematic diagram of an SET operation is shown in **Figure 11**. According to the increase in the gate bias, the charge in the quantum dot increases along a step function Ne (N: integer). A current flow, that is, electron flow, should be accompanied by a change of the number of electrons in the quantum dot. Therefore, the current flow is limited only in the condition where N can change with $\Delta N = 1$. On the other hand, the current flow is blocked when N is constant (Coulomb blockade).

Some SET characteristics are shown in **Figure 12**. The quantum dot has a circular disk shape in this example. The interval of current peaks determined by $\Delta V_g = e/C_g$ decreases with increasing bias. This is because an increase in the disk size with gate bias results in increased gate capacitance. In addition to the classical charging effect, quantum effects, such as the formation of zero-dimensional energy states, become important in small quantum dots. The zero-dimensional energy levels are represented by the diagram in **Figure 12b** for the parabolic confinement, which is a good approximation of the disk-shaped quantum dot. Here, $\hbar\omega_0$ represents the zero-dimensional energy-level separation arising from the parabolic confinement. The zero-dimensional energy levels degenerate at zero magnetic field, where $\omega_c = 0$. Each energy level can include two electrons with different spins. Therefore, ΔV_g increases for $N = 2, 6, 12, \ldots$ where the zero-dimensional energy separation should be compensated by the gate bias in addition to e/C_g. Such periodic features have actually been observed as shown in **Figure 12a** for a circular quantum dot. This periodic feature comes from the same principle as the periodic table of real atoms, so that these SETs are sometimes called semiconductor artificial atoms.

At low temperatures, the conductive peak becomes sharp as experimentally shown in **Figure 12a**. Furthermore, the peak position shifts when the charge situation around the quantum dot is slightly modified as schematically shown in **Figure 11b**. Therefore, the conductance change in an SET operates as a very sensitive electrometer. Recently, the single electron motion in semiconductor nanostructures was detected by using the SET-based highly sensitive electrometer.

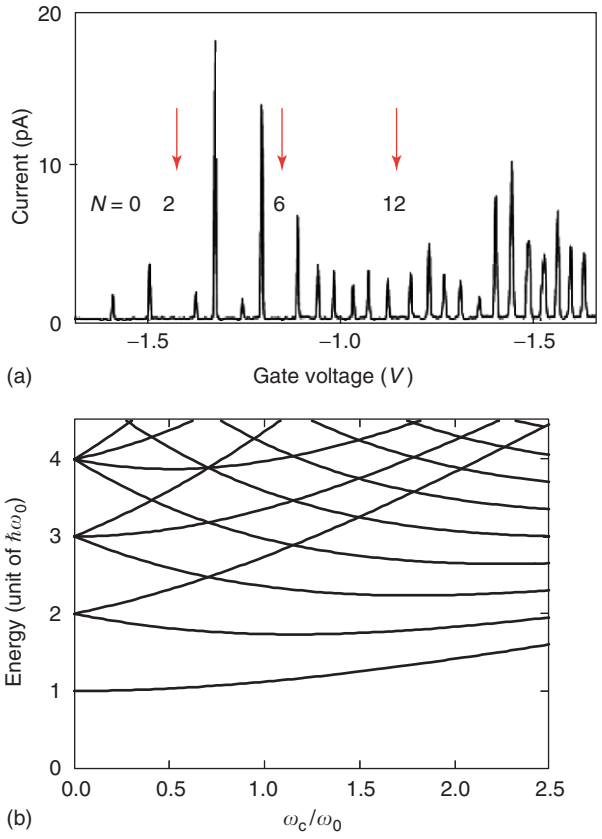


Figure 12 (a) Current vs. gate voltage characteristics experimentally observed for the disk-shaped SET. The current peak interval becomes wider for N (the number of electrons in the quantum dot) of $2, 6, 12, \ldots$, as shown by arrows. Such characteristics are explained by the zero-dimensional energy levels formed in the circular quantum dot and shown in (b). This energy diagram is calculated for a parabolic confinement potential, $V = m\omega_0^2(x^2 + y^2)/2$. $\hbar\omega_c$ is the cyclotron energy separation in the perpendicular magnetic field.

See also: Integrated Circuits; Light Emitting Diodes.

PACS: 85.30. – z; 73.40. – c; 72.20. – i; 85.35. – p; 73.23. – b; 85.40. – e; 68.65. – k

Further Reading

Adir Bar-Lev (1984) *Semiconductors and Electronic Devices*, 2nd edn. New Jersey: Prentice-Hall.

Beenakker CWJ and van Houten H (1991) Quantum transport in semiconductor nanostructures. In: Ehrenreich H and Turnbull D (eds.) *Solid State Physics*, vol. 44, p. 1. New York: Academic Press.

Enoki T, Sano E, and Ishibashi T (2001) Prospects of inp-based ic technologies for 100-Gbit/s-class lightwave communications systems. *International Journal of High Speed Electronics and Systems* 11: 137.

Fujisawa T, Austing DG, Tokura Y, Hirayama Y, and Tarucha S (2003) Electrical pulse measurement, inelastic relaxation, and non-equilibrium transport in a quantum dot. *Journal of Physics: Condensed Matter* 15: R1395 (topic review).

Kiehl RA and Sollner TCLG (1994) *High Speed Heterostructure Devices, Semiconductors and Semimetals*., vol. 41. New York: Academic Press.

Kouwenhoven LP, Markus CM, McEuen PL, Tarucha S, Westervelt RM, *et al*. (1997) Electron transport in quantum dots. In: Sohn LL, Kouwenhoven LP, and Schoen G (eds.) *Mesoscopic Electron Transport*, NATO ASI Series. Dordrecht: Kluwer Academic.

Kroemer H (1982) Heterojunction bipolar transistors and integrated circuits. *Proceedings of the IEEE*, 13.

Makimoto T, Kumakura K, and Kobayashi N (2002) *International Workshop on Nitride Semiconductors (IWN2002)*, p. 379, Aachen.

Mishra UK, Parikh P, and Wu Yi-Feng (2002) AlGaN/GaN HEMTs – an overview of device operation and applications. *Proceedings of the IEEE* 90: 1022.

Reed M (ed.) (1992) *Nanostructured Systems, Semiconductors and Semimetals*.. New York: Academic Press.

Roblin P and Rohdin H (2002) *High-Speed Heterostructure Devices: From Device Concepts to Circuit Modeling*. Cambridge: Cambridge University Press.

Sze SM (ed.) (1990) *High-Speed Semiconductor Devices*. New York: Wiley.

Sze SM (2002) *Semiconductor Devices: Physics and Technology*, 2nd edn. New York: Wiley.

Tarucha S, Austing DG, Sasaki S, Kouwenhoven LP, Reimann S, Koskinen M, and Manninen M (2001) Electronic states in circular and ellipsoidally deformed quantum dots. In: Yao T (ed.) *Physics and Applications of Semiconductor Quantum Structurech*. *9*, pp. 194–216. Bristol: IOP Publishing.

Taur Y and Ning TH (1998) *Fundamentals of Modern VLSI Devices*. Cambridge: Cambridge University Press.

Uyemura JP (2001) *Introduction to VLSI, Circuits and Systems*. New York: Wiley.

Transition-Metal Compounds, Electronic and Magnetic Properties of

J B Goodenough, University of Texas at Austin, Austin, TX, USA

Introduction

The electronic and magnetic properties of transition-metal compounds are imparted by the *d* and/or 4*f* electrons of the transition-metal compounds. (This article does not cover the actinides.) The outer *s* and *p* electrons of these compounds, which are primarily responsible for the chemical bonding, are split by an energy gap E_g between filled anionic bonding bands and empty cationic antibonding bands. The *s* and *p* states contribute to the electronic and magnetic properties of the transition-metal compounds through covalent mixing with the *d* and/or 4*f* orbitals.

The 4*f* electrons always form localized $4f^n$ configurations as on a free atom unless a $4f^n/4f^{n+1}$ redox couple is overlapped by the Fermi energy (the electrochemical potential) ε_F of a partially filled broadband of itinerant-electron states, in which case hybridization of the 4*f* and itinerant electron states creates massive itinerant electrons called heavy fermions. The *d* electrons, on the other hand, may be localized, itinerant, or vibronic depending on the relative strengths of the intra-atomic and the inter-atomic interactions and how the competition between these two is modulated by disturbances of the periodic potential of a crystalline array of the like atoms or by elastic and thermal energies.

The intra-atomic energies are those of the free transition-metal atom as modified by covalent mixing with the ligand orbitals. The Hamiltonian for the electrons of a free atom has three components of descending magnitude:

$$H = H_o + U + \lambda \boldsymbol{L} \cdot \boldsymbol{S} \qquad [1]$$

where H_o corresponds to the energy of a single electron moving in the field of a nucleus of effective charge $+(Z - \sigma_l)e$. The atomic number Z is reduced by a screening σ_l of the other atomic electrons; σ_l depends on the orbital angular momentum quantum number l of the electron being screened from the nuclear charge $+Ze$. The energy U is the electrostatic energy required to add an electron to a configuration of *d* or 4*f* orbitals and is a measure of the separation of successive ionization energies; $U = J_c - J_{ex}$ has two components, a Coulombic repulsion J_c and a quantum-mechanical correction J_{ex} that keeps electrons of parallel spin in different orbitals by the Pauli exclusion principle. The last term in eqn [1] gives rise to the multiplet splitting of a $^{2S+1}L_J$ term, where Russell–Saunders coupling applies to the d^n and $4f^n$ configurations of total angular momentum quantum number $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$; $\boldsymbol{L}$ and $\boldsymbol{S}$ are the total orbital and spin angular-momentum quantum

numbers. The atomic magnetic moment is

$$\boldsymbol{\mu}_{\mathrm{A}} = -g_{\mathrm{L}}\boldsymbol{J}\mu_{\mathrm{B}} \qquad [2]$$

where g_{L} is the Landé spectroscopic splitting factor and μ_{B} is the Bohr magneton.

In a solid, the "interatomic interactions" depend on the expectation value for an electron to transfer to a neighboring atom without a change of energy as a result of the perturbation H' of the free-atom potentials by their proximity to one another:

$$b_{ij} \equiv (\psi_i, H'\psi_j) \approx \varepsilon_b(\psi_i, \psi_j) \qquad [3]$$

where ε_b is a one-electron energy and (ψ_i, ψ_j) is the overlap integral of the initial and final atomic orbitals occupied by the electron. The subscripts refer to atoms at $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ in the crystal. Electron transfers between states of the same energy are treated in the first-order theory ($\Delta E \sim b$) and between states separated by an energy Δ or U in the second-order theory ($\Delta E \sim b^2/\Delta$ or b^2/U). The tight-binding bandwidth for electrons transferring between like atoms of a periodic array of crystallographically equivalent sites with $U = 0$ is given by

$$W_b \approx 2zb \qquad [4]$$

where z is the number of like nearest neighbors and b is the b_{ij} for these nearest neighbors. The crossover from localized-electron behavior (intra-atomic interactions dominant) to itinerant-electron behavior (interatomic interactions dominant) occurs where

$$W_b \approx U \qquad [5]$$

This crossover is known as the "Mott–Hubbard transition."

The localized electrons occupy an n-electron configuration d^n or f^n located at a single transition-metal atom and having an energy corresponding to that of a redox couple d^{n-1}/d^n or f^{n-1}/f^n. Successive redox couples are separated by an energy U. The itinerant electrons, on the other hand, belong equally to all the like atoms of a periodic array on crystallographically equivalent sites; they are described by a band of single-electron energies of width W_b that is occupied by electrons in accordance with the Pauli exclusion principle and the Fermi–Dirac distribution function. In this limit, a $U < W_b$ is assumed to be screened by the other electrons so as to make $U = 0$. For narrower bandwidths W_b, this "electronic Fermi-gas model" must be modified by the introduction of a finite $U < W_b$. This problem can be treated by mass-renormalization techniques that transform the electrons into a "quasiparticle Fermi liquid."

Impurities, lattice defects, and substitutional dopants all disturb the periodic potential of an array of transition-metal atoms in a solid. These disturbances narrow the effective width of W_b by introducing localized states (Anderson localized states) at the top and bottom of a band of itinerant-electron states above and below, respectively, a "mobility edge." If the Fermi energy ε_{F} lies in the energy range of localized states, the conductivity is described as "variable-range hopping" between these localized states.

The Mott–Hubbard transition is not smooth as originally assumed; it is a first-order transition that may give an abrupt volume change, but generally it introduces the coexistence of two different equilibrium bond lengths at crossover. In the latter case, the resulting bond-length fluctuations at high temperatures order at lower temperatures so as to give long-range-ordered atomic clusters, a charge-density wave (CDW), or a charge-ordered (CO) state in order to minimize the elastic energy. The elastic energies associated with local site deformations are also minimized by cooperative ordering. Examples of local site distortions are displacements of a cation from the center of symmetry of its interstice, cation clustering, valence disproportionation, and an orbital ordering that either lowers the local site symmetry or increases its deformation from cubic symmetry.

The electronic and magnetic properties of the d electrons of transition-metal compounds are strongly influenced not only by whether they occupy states with $W_b < U$, $W_b \approx U$, or $W_b > U$ but also by order–disorder transitions. Order–disorder transitions are driven by the heat $T\Delta S$ associated with disorder. Heat also drives disordering of long-range magnetic order, which removes magnetostrictive and/or exchange-strictive deformations of the crystal symmetry; and by increasing the crystal volume, it may induce changes from antiferromagnetic to ferromagnetic coupling, that is, "exchange inversion," below a long-range magnetic-ordering temperature.

In mixed-valent compounds, the transport properties also depend critically on whether the time $\tau_{\mathrm{h}} \approx \hbar/W_b$ for an electron to hop from one atom to a like nearest neighbor is $\tau_{\mathrm{h}} > \omega_{\mathrm{R}}^{-1}$, $\tau_{\mathrm{h}} \approx \omega_{\mathrm{R}}^{-1}$, or $\tau_{\mathrm{h}} < \omega_{\mathrm{R}}^{-1}$, where ω_{R}^{-1} is the period of an optical-mode vibration that would trap the charge carrier at a single site as a "small dielectric polaron."

In the following sections, the metal–ligand interactions are first introduced to construct ligand-field orbitals having the same symmetries as the atomic d orbitals. The interactions between the ligand-field orbitals of neighboring metal atoms are then discussed for $W_b < U$, $W_b \approx U$, and $W_b > U$. Next, the significance of the radial extension of the ligand-field orbitals and the location of the Fermi energy relative

to the edges of the broad filled and empty bands is presented in the context of illustrative examples of the varied phenomena that are encountered in these materials.

Metal–Ligand Interactions

Redox Couples f^n or d^n

The description of the outer electrons of a primarily ionic transition-metal compound begins with a point-charge model; the electrostatic Madelung energy E_M of that model generally lowers the most energetic ligand (anion) p orbitals an energy ΔE_p below the lowest unoccupied redox energy of the transition-metal atom. A $\Delta E_p \leqslant 0$ may pin the antibonding redox couple at the top of the bonding ligand-p band; in this case, the redox couple can commonly be treated by the molecular-orbital or band theory. However, back electron transfer from the anion across an energy $\Delta E_p > 0$ to an empty d or f orbital on the cation is treated in the second-order perturbation theory to give a stabilization energy $\Delta E \sim (b^{ca})^2/\Delta E_p$ of occupied bonding states that tends to compensate for the loss in electrostatic Madelung energy E_M of the point-charge model; the empty antibonding d or f orbitals of the d^n or f^n configuration are correspondingly destabilized by an energy ΔE. As noted in eqn [3], the expectation value for a back electron transfer depends on the cation–anion overlap integral in b^{ca} and, therefore, on the relative symmetries of the interacting orbitals. Consequently, symmetry arguments are used to develop a parametrized second-order description of the "ligand-field orbitals"; these orbitals take account of the interactions between the d or f electrons and the anions. Next, the metal–metal, metal–anion–metal, and/or metal–anion–anion–metal interactions are introduced; the ligand-field orbitals either remain localized or become itinerant depending on the relative strengths of these inter-atomic interactions and the intra-atomic energy U_{eff} for the atomic ligand-field orbitals.

Rare-Earth $4f^n$ Configurations

The rare-earth $4f$ electrons are tightly bound to their atomic nucleus and are largely buried within a $5s^2 5p^6$ core-electron cloud that prevents strong overlap of the $4f$ orbitals with the orbitals of the neighboring ligands. Consequently, the $4f^n$ configurations remain localized with a $U_{eff} \gg W$, and any splitting of the individual f-orbital energies by interaction with the ligands is smaller than the intra-atomic multiplet splittings by $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$. Therefore, the localized $4f^n$ configurations in a compound retain their atomic magnetic moment $\boldsymbol{\mu}_A = g_L \boldsymbol{J} \mu_B$ of eqn [2]. Nevertheless, the interactions with the ligands contribute, along with the "magnetic dipole–dipole $4f^n$ interactions," to the weak "spin–spin interatomic exchange" interactions between $4f^n$ configurations that are responsible for long-range magnetic order at low temperatures and also to the "magnetocrystalline anisotropy" that introduces a preferred orientation of $\boldsymbol{\mu}_A$ relative to the crystallographic axes and an associated "magnetostriction" below the magnetic-ordering temperature. Long-range ordering of the spins $\boldsymbol{S}$ induces long-range ordering of the orbital angular momenta $\boldsymbol{L}$ by the spin–orbit coupling $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$, and interactions with the ligands order the orbitals responsible for $\boldsymbol{L}$, and therefore $\boldsymbol{J}$, relative to the crystallographic axes and also induce the anion displacements responsible for the magnetostriction. Magnetostriction reflects the orientation of the atomic moments; "exchange striction" reflects changes in the binding energy due to magnetic order and, therefore, reflects the symmetry of the long-range magnetic order.

The d-Block d^n Configurations

The five d orbitals of a free atom are degenerate, but with more than one electron or hole in a d^n manifold, the spin degeneracy is removed by the ferromagnetic direct-exchange interactions between electron spins in orthogonal atomic orbitals. The Pauli exclusion principle keeps electrons of the same spin in different d orbitals, which reduces the energy $U = J_c - J_{ex}$ by the intra-atomic exchange energy J_{ex}. These spin–spin exchange interactions produce the "Hund intra-atomic exchange field H_{ex}" responsible for the Hund highest multiplicity rule for the free atom. The splitting between states of different spin for a given orbital will be designated Δ_{ex}.

The atomic orbitals f_m with azimuthal orbital angular momentum operator L_z, where $L_z f_m = -i\hbar \, \partial f_m/\partial\phi = \pm m\hbar f_m$ with $m = -l, \ldots, +l$, have the angular dependencies for $l = 2$:

$$f_0 \sim (\cos^2\theta - 1) \sim [(z^2 - x^L) + (z^2 - y^2)]/r^2$$

$$f_{\pm 1} \sim \sin^2\theta \, \exp(\pm i\phi) \sim [yz \pm izx]/r^2 \qquad [6]$$

$$f_{\pm 2} \sim \sin^2\theta \, \exp(\pm i2\phi) \sim [(x^2 - y^2) \pm i2xy]/r^2$$

In an isolated octahedral site, the xy and $(yz \pm izx)$ orbitals only overlap the neighboring O–2p_π orbitals while the $[(z^2 - x^2) + (z^2 - y^2)]/r^2$ and $(x^2 - y^2)/r^2$ orbitals only overlap the O–2s and O–2p orbitals. These two groupings of orbitals are distinguished by symmetry as a twofold-degenerate set of e orbitals and a threefold-degenerate set of t orbitals (**Figure 1**). The electron-transfer (resonance energy integrals

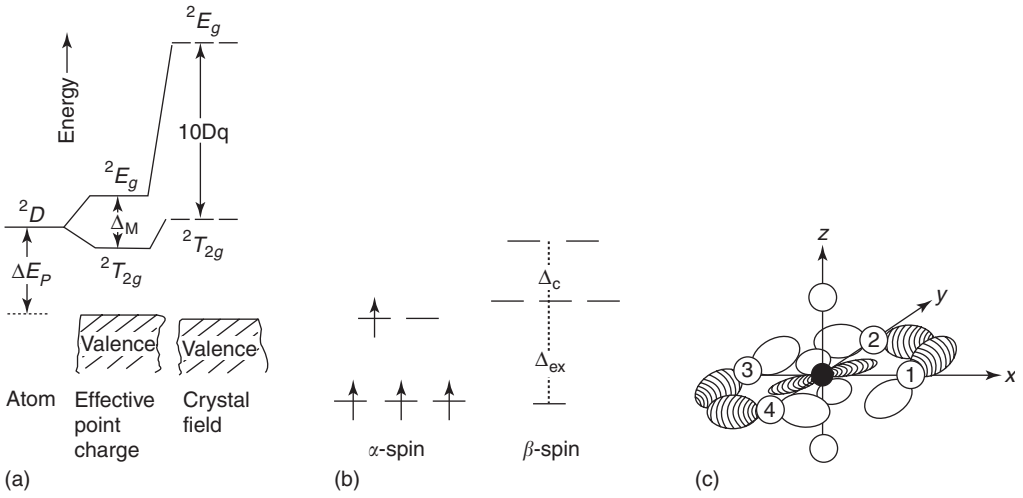


Figure 1 Schematic crystal-field splitting of the *d*-state manifold of a transition-metal atom M in an octahedral site: (a) the one-electron d^1 configuration ($\Delta_{ex} = 0$); (b) the high-spin d^4 configuration, and (c) $\phi_{xy} = p_y^{(1)} + p_x^{(2)} - p_y^{(3)} - p_x^{(4)}$ that hybridizes with a d_{xy} orbital.

$b^{ca} \equiv (f_m, H'\phi_o) \approx \varepsilon_{mo}(f_m, \phi_o)$ describing the expectation of a virtual charge transfer to an empty cation-*d* orbital from a same-symmetry sum of near-neighbor oxygen orbitals ϕ_o contain both an overlap integral (f_m, ϕ_o) and a one-electron energy ε_{mo} that are larger for σ-bonding than for π-bonding, that is, $b_\sigma^{ca} > b_\pi^{ca}$. Therefore, the antibonding *e* states of a σ-bond are raised higher by the energy $\Delta E \equiv (b^{ca})^2/\Delta E_p$ than are the antibonding *t* states of a π-bond. The resulting cubic-field splitting is

$$\Delta_c = \Delta E_\sigma - \Delta E_\pi = \Delta_M + (\lambda_\sigma^2 - \lambda_\pi^2)\Delta E_p + \lambda_s^2 \Delta E_s \quad [7]$$

where Δ_M is a purely electrostatic energy that is small and of uncertain sign due to the penetration of the ligand-p^6 electron cloud by the cation-*d* wave function. The dimensionless covalent-mixing parameters are seen to be

$$\begin{aligned} &\lambda_\sigma = b_\sigma^{ca}/\Delta E_p, \quad \lambda_\pi = b_\pi^{ca}/\Delta E_p \\ &\lambda_s = b_s^{ca}/\Delta E_s \end{aligned} \quad [8]$$

in which ΔE_p and ΔE_s are the point-charge energy gaps between the degenerate atomic *d* orbitals and the ligand *p* and *s* orbitals, respectively. The second-order ligand-field wave functions are

$$\begin{aligned} \psi_t &= N_\pi(f_t - \lambda_\pi \phi_\pi) \\ \psi_\sigma &= N_\sigma(f_e - \lambda_\sigma \phi_\sigma - \lambda_s \phi_s) \end{aligned} \quad [9]$$

where N_π and N_σ are normalization constants that constrain a one-electron state to contain a single electron.

Covalent mixing extends the ligand-field wave functions out over the anions, thereby reducing the electron–electron Coulomb energy *U* from its atomic value. For an octahedral-site cation, the ligand-field *U* is larger for electrons in the π-bonding *t*-manifold than for those in the σ-bonding *e*-manifold, that is, $U_t > U_e$.

As an aside, attention is drawn to the earlier crystal-field model in which the covalent component to the orbitals is neglected and the cubic-field splitting is simply parametrized as $\Delta_c = 10\,Dq$. In this model also, the symmetry arguments distinguish the *e* and *t* manifolds in a cubic field or other splittings in lower symmetry, but the energy of the atomic *d*-state manifold is conserved and the crystal-field wave functions are simply the atomic wave functions f_m. Although this parametrized model allows interpretation of spectroscopies that measure the energy splittings, it fails completely to account for a lowered ligand-field *U* and for the interatomic interactions between like atoms that are responsible for long-range magnetic order where a $W_b < U_{eff}$ is found or for metallic conductivity where a $W_b > U_{eff}$ occurs.

If a *d*-block transition-metal ion occupies a tetrahedral site, similar considerations apply, but in this case the *t* orbitals have a larger orbital overlap and are raised higher in energy than the *e* orbitals by the cubic-field splitting Δ_c. Early crystal-field calculations give the relative values of Δ_c for the two cases as

$$\Delta_c(\text{tetrahedral}) \approx (4/9)\Delta_c(\text{octahedral}) \quad [10]$$

The *d*-electron Hamiltonian that includes the interactions with the ligand *p* and *s* orbitals is

$$H = H_o + J_c + (\Delta_c - J_{ex}) + (\lambda \boldsymbol{L} \cdot \boldsymbol{S} + \Delta_{nc}) \quad [11]$$

where Δ_{nc} is a noncubic component of the ligand-field splitting arising from a distortion of the site symmetry. Since the octahedral-site Δ_c is the same order of magnitude as the intra-atomic exchange energy Δ_{ex}, the $d^4 - d^7$ octahedral-site configurations may be either high-spin t^3e^1, t^3e^2, t^4d^2, t^5e^2 where $\Delta_{ex} > \Delta_c$ or low-spin t^4e^0, t^5e^0, t^6e^0, t^6e^1 where $\Delta_c > \Delta_{ex}$. Moreover, lowering of the site symmetry to tetragonal or orthorhombic may stabilize an intermediate-spin d^5 or d^6 configuration t^4e^1 or t^5e^1, respectively. The larger the covalent mixing parameters $\lambda_\sigma > \lambda_\pi$, the smaller is Δ_{ex} and the larger is Δ_c. Moreover, the effective energy U_{eff} required to add an electron to a d^n manifold to make it d^{n+1} must take into account Δ_c as well as Δ_{ex}. For an octahedral-site cation, the U_{eff} for $n = 3$ and $n = 8$ is increased by Δ_c; for $n = 5$ it is increased by Δ_{ex} if $\Delta_c < \Delta_{ex}$, and for $n = 6$ by Δ_c if $\Delta_c > \Delta_{ex}$.

By splitting the $m = \pm 2$ energies of eqn [6] into real $(x^2 - y^2)$ and xy components, the operator $L_z = -i\hbar\partial/\partial\phi$ makes the azimuthal component of the orbital angular momentum imaginary, which means that the two e orbitals have $m = 0$ and the t manifold has orbitals with $m = 0, \pm 1$. It follows that configurations with zero, three, or six t electrons have their orbital angular momentum completely quenched (i.e., $L = 0$) to first order, and a spin-only atomic moment $\boldsymbol{\mu}_A = gS\mu_B$ with $g \approx 2$ is a fair approximation.

The octahedral-site splitting may leave an orbital degeneracy, as occurs with high-spin Mn(III):t^3e^1 or Cu(II):t^6e^3, for example, or with high-spin Fe(II):t^4e^2 or Co(II):t^5e^2. In this case, the cubic site symmetry is unstable relative to a site distortion that removes the orbital degeneracy. This instability is known as a "Jahn–Teller distortion." In a crystal, a long-range-cooperative site distortion, that is, orbital ordering, minimizes the elastic energy that resists the distortion and lowers the space-group symmetry of the crystal. Cooperative Jahn–Teller distortions may transform isotropic 3D interatomic spin–spin interactions into a ferromagnetic interaction in one crystallographic direction and an antiferromagnetic interaction in another as is illustrated by $LaMnO_3$ (*vide supra*). Removal of a t^n-orbital degeneracy ($n = 1, 2, 4$, or 5) may either quench the orbital angular momentum or enhance it so as to gain a further splitting by the spin–orbit coupling $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$, where σ-bonding electrons determine the magnetic order, a long-range collinear-spin configuration below a magnetic-ordering temperature introduces long-range order of the orbital angular momenta through $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$ even where the concentration of orbitally degenerate cations is not large. Therefore, below a long-range magnetic-ordering temperature, distortions that enhance $\boldsymbol{L}$ and introduce a large magnetostrictive component scale with the concentration of cations with the t-orbital degeneracy. On the other hand, spin–orbit coupling prevents a cooperative ordering of the orbital momenta in the paramagnetic state of disordered spins; so cooperative distortions setting in above a magnetic-ordering temperature quench the orbital angular momenta. The two types of distortion can be distinguished as they have opposite sign. This situation is beautifully illustrated in **Figure 2** by the spinel system $Ni_{1-x}Fe_x[Ni_xCr_{2-x}]O_4$ in which the tetrahedral-site Ni(II):e^4t^4 are the Jahn–Teller ions.

The $A[B_2]X_4$ spinel structure consists of a close-packed-cubic array of anions X; the electrostatic Madelung energy is minimized by an ordering of the B cations on half of the octahedral sites and the A cations on tetrahedral sites of the interstitial space of the $[B_2]X_4$ array. The $Ni_{1-x}Fe_x[Ni_xCr_{2-x}]O_4$ spinel also illustrates how the crystal-field splittings influence the site-preference energies of transition-metal cations. Sites that leave the higher-energy d orbitals empty are stabilized relative to those that leave them occupied. The Cr(III) ion has the configuration t^3e^0 in an octahedral site and e^2t^1 in a tetrahedral site, which gives it a strong octahedral-site preference. The high-spin Fe(III) ion with configurations t^3e^2 or e^2t^3 has little site preference. The Ni(II) ion, on the

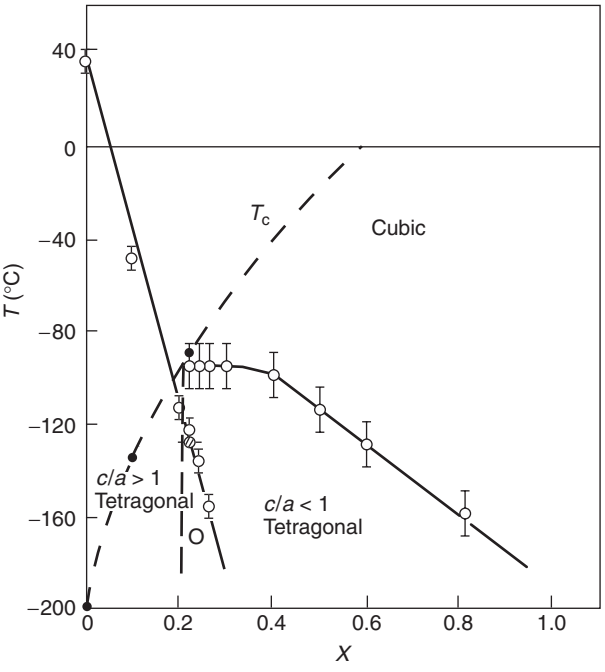


Figure 2 Phase diagram of the spinel system $Ni_{1-x}Fe_x[Ni_xCr_{2-x}]O_4$. Solid lines are tetragonal-cubic transition temperatures T_t. The O-orthorhombic phase may be a coexistence of tetragonal $c/a > 1$ and $c/a < 1$ phases. (After Arnott RJ, Wold A, and Rogers DB (1964) *Journal of the Physics and Chemistry of Solids* 25: 161.)

other hand, has an octahedral-site configuration t^6e^2 as against e^4t^4 in a tetrahedral site, which gives it a definite octahedral-site preference; but it is forced into the tetrahedral sites of the spinel by the Cr(III) ions, which have a stronger octahedral-site preference.

Ionic size also determines site preference. Smaller cations prefer tetrahedral sites, the *d*-block transition-metal cations of intermediate size, with the exception of Mn(II) and Cu(I), prefer octahedral sites, and the larger rare-earth ions prefer a greater oxygen coordination. Where a smaller cation is forced by electrostatic Coulomb energies into an octahedral site, it may be unstable relative to a cooperative displacement from the center of symmetry of its interstice. Such a displacement introduces a local electric-dipole moment, and a cooperative parallel displacement gives a "ferroelectric polarization" analogous to a "ferromagnetic magnetization." Therefore, such a displacement is referred to as a ferroelectric-type displacement even though the cooperativity may involve antiparallel displacements to give an antiferroelectric analogous to an antiferromagnet in which neighboring spins are aligned antiparallel. Lighter transition-metal cations of a given *d*-block can be stabilized in octahedral sites in their highest valence states, that is, with a d^0 configuration, but as the cation size decreases with increasing atomic number and formal valence state, the site preference changes from octahedral to tetrahedral. It is the ions with sizes at crossover to a tetrahedral-site preference that undergo ferroelectric displacements. In the 3*d* row, the T(IV) has an octahedral-site preference, but if the structure forces the Ti–O bond length to be greater than its equilibrium value, as occurs in the perovskite $BaTiO_3$, the Ti(IV) ions may undergo a cooperative ferroelectric displacement. In $BaTiO_3$, there are three successive cooperative displacements; on lowering the temperature, the Ti(IV) ions are displaced first toward one near-neighbor O^{2-} ion, then toward two, and finally toward three. The smaller V(V) cation, on the other hand, has a definite tetrahedral site preference, but when forced into an octahedral site as in V_2O_5, it forms a short double bond V═O indicating molecular-orbital formation with both p_σ and the two p_π orbitals. The V(IV) may also form a $(V{=}O)^{2+}$ vanadyl ion with an occupied d_{xy} orbital in the plane perpendicular to the d_{yz} and d_{zx} orbitals involved in the π molecular orbitals. In the 4*d* block, it is the Nb(V) ion that corresponds to the Ti(IV) and the Mo(VI) that corresponds to the V(V). In the 5*d* series, the W(VI) corresponds to the Ti(IV), and WO_3 is an antiferroelectric without a stretching of the W–O bond length by a counter cation as in $BaTiO_3$.

Finally, it is noteworthy that where a transition-metal cation is deprived of its stable oxygen coordination, as occurs in the oxygen-deficient perovskite $BaZr_{1-x}In_xO_{3-0.5x}$, water may enter the oxygen vacancies at lower temperatures ($T \leqslant 400^\circ C$) to introduce OH^- anions and protonic conduction; the Zr(IV) ion is stable in eightfold oxygen coordination and definitely prefers a sixfold over the fivefold oxygen coordination just as the large Ba^{2+} ion prefers a 12-fold over the 11-fold coordination.

Metal–Metal Interactions

A measure of the strength of the interatomic cation–cation interactions is also the energy integral b_{ij} of eqn [3]. Direct cation–cation electron transfers b^{cc} between cations sharing a common site edge or face are distinguished from cation–anion–cation transfers b^{cac} that occur across an anion. The latter interactions are possible because the shared anion-*p* component (neglecting the smaller anion-*s* component for simplicity) of the ligand-field wave functions provides a significant overlap of the wave functions on neighboring cations. For a $(180^\circ - \phi)$ cation–anion–cation interaction, the electron transfer integrals b_σ^{cae} across a common anion-p_σ orbital must be distinguished from the smaller b_π^{cac} across a common anion-p_π orbital; in a 90° cation–anion–cation interaction, the electron transfer integral is $b_{\pi\sigma}^{cac}$ as it occurs across an anion-*p* orbital that σ-bonds with the *d* orbital on one cation and π-bonds with the *d* orbital on the other cation. With the ligand-field wave functions of eqn [9], these integrals become

$$b^{cc} = \varepsilon_{cc}(\psi_i, \psi_j) \qquad [12]$$

which varies sensitively with the interatomic separation *R*,

$$\begin{aligned} b_\pi^{cac} &\approx \varepsilon_\pi \lambda_\pi^2 \\ b_\sigma^{cac} &\approx \varepsilon_\sigma \lambda_\sigma^2 \cos\phi \\ b_{\pi\sigma}^{cac} &\approx \varepsilon_{\pi\sigma} \lambda_\pi \lambda_\sigma \end{aligned} \qquad [13]$$

Two general conclusions follow. First, there is a critical cation–cation separation R_c below which the tight-binding bandwidth of eqn [4] is $W_b \approx 2zb^{cc} > U_{eff}$ and suppresses the localized-electron behavior to introduce the *d* electrons into either itinerant-electron states or cation-cluster molecular orbitals. Second, it is possible to have itinerant and localized electrons coexisting on the same atom. For example, $(180^\circ - \phi)$ cation–anion–cation interactions dominate in an AMO_3 perovskite, which has a simple cubic array of corner-sharing $MO_{6/2}$ octahedra with

a larger A-site cation in the body center. A $b_\sigma > b_\pi$ and a $U_e < U_t$ allow a $W_\sigma > U_e$ or a mixed-valent $\tau_h < \omega_R^{-1}$ of e electrons in the presence of $W_\pi < U_t$ as is illustrated by metallic and magnetic $SrFeO_3$ and the metallic ferromagnet $La_{0.7S}Sr_{0.3}MnO_3$.

The criterion for itinerant versus localized behavior in a mixed-valent compound is somewhat different from that in a single-valent compound. In a mixed-valent compound, electron transfer between atoms of a redox couple does not require overcoming an energy gap U_{eff}. However, the time $\tau_h \approx \hbar/W$ for an electron to transfer from one atom to another must be compared to the period $\omega_R^{-1} \approx 10^{-12}$ s of an optical-mode local vibration that traps a charge carrier. If the ligands have time to relax to an equilibrium position that distinguishes the smaller size of a cation of higher charge, then the stronger covalent bonding at the smaller site raises the energy of the empty antibonding orbitals above those of the occupied orbitals by ε_p as occurs with a redox couple in solution. The atomic relaxations thus self-trap a charge carrier, localizing it to a single site by introducing a small energy barrier between occupied and empty states. The self-trapped carriers are called "dielectric small polarons"; they move diffusively with a small activated mobility

$$\mu = (eD_o/kT)\exp(-\Delta G_m/hT) \qquad [14]$$

in which $\Delta G_m = \Delta H_m - T\Delta S_m$ is the free energy required to equalize the energies of the orbitals on neighboring sites. This mobility contrasts with that for an itinerant electron:

$$\mu = e\tau_s/m^* \qquad [15]$$

where τ_s is the mean free time between scatterings of the electronic charge carrier from aperiodicities in its periodic potential and m^* is the effective carrier mass. The crossover from polaronic to itinerant electronic behavior occurs where $\tau_h \approx \omega_R^{-1}$ or the polaron bandwidth is $W \approx \hbar\omega_R$. The polaronic bandwidth is reduced from the tight-binding bandwidth W_b of eqn [5] as

$$W = W_b \exp(-\lambda\varepsilon_p/\hbar\omega_R) \qquad [16]$$

where $\lambda \sim \varepsilon_p/W_b$ is a measure of the strength of the electron–lattice coupling. The transition from polaronic to itinerant electronic behavior in a mixed-valent compound would generally occur where U_{eff} remains a little larger than W_b, but Coulomb interactions between the charge carriers may slow τ_h and induce charge ordering at higher charge concentrations. Moreover, local site distortions at Jahn–Teller ions increase ε_p, which stabilizes polaronic behavior to larger values of W_b. Moreover, the virial theorem dictates a first-order transition from localized to itinerant electronic behavior, and the conditions under which the two types of crossover are found appear to be similar.

A first-order transition at $W_b \approx U_{eff}$ leads to bond-length instabilities; the two coexisting equilibrium bond lengths order at lower temperatures in a variety of ways to give such unusual phenomena as high-temperature superconductivity in copper oxides or a colossal magnetoresistance in manganese and cobalt oxides with a perovskite structure. At higher temperatures, bad-metal behavior with an unusual temperature dependence of the resistivity is found. As the temperature is increased, the bandwidth narrows and a smooth transition from itinerant to vibronic to polaronic behavior may be encountered.

The interatomic interactions of interest in the localized-electron regime $W_b < U_{eff}$ are the spin–spin interactions. The origin of the spin–spin interactions is in the quantum-mechanical exchange term J_{ex} entering U, so the spin–spin interactions are referred to as "exchange interactions."

The most general form of the bilinear interatomic spin–spin coupling between localized spins at $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ is

$$V_{ij} = S_i \cdot \boldsymbol{K}_S \cdot S_j + S_i \cdot \boldsymbol{K}_A \cdot S_j \qquad [17]$$

where $\boldsymbol{K}_S$ and $\boldsymbol{K}_A$ are symmetric and antisymmetric tensors. The first term reduces to an isotropic and an anisotropic component, $-2J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j - 2\boldsymbol{S}_i \cdot \Gamma \cdot \boldsymbol{S}_j$, where $\Gamma \sim (\Delta g/g)^2$ represents a symmetric tensor and Δg is the deviation of the spectroscopic splitting factor from $g = 2$. The antisymmetric term can be reduced to $-2\boldsymbol{D}_{ij} \cdot \boldsymbol{S}_i \times \boldsymbol{S}_j$, where the Dzialoshinskii vector $\boldsymbol{D}_{ij} \sim (\Delta g/g)$ has its axis determined by the crystal symmetry. A quantitative description of the magnetization $M(T)$ of a ferromagnetically aligned subarray may require inclusion of a biquadratic term $-2\gamma_{ij}(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2$ where the magnitude of a cation–cation interaction varies sensitively with the cation separation R. In addition, the coexistence of localized and itinerant electrons (or molecular-orbital electrons) adds an additional term H_{ex}^D, so the full phenomenological expression for the spin–spin interactions becomes

$$H_{ex} = H_{ex}^D - \sum_{ij}{}' \{J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j + \gamma_{ij}(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2 + \boldsymbol{D}_{ij} \cdot \boldsymbol{S}_i \times \boldsymbol{S}_j + \boldsymbol{S}_i \cdot \Gamma \cdot \boldsymbol{S}_j\} \qquad [18]$$

Measurement of the magnetic susceptibility $\chi_m \equiv M(T)/H_a$, where H_a is an applied magnetic

field, reveals whether localized spins are present in a compound and the character of the dominant spin–spin interactions, which may be ferromagnetic ($J_{ij}>0$) or antiferromagnetic ($J_{ij}<0$). In the absence of localized spins, a compound may be diamagnetic ($\chi_m<0$) or paramagnetic ($\chi_m>0$), but χ_m is normally small and temperature independent in this case. However, superconductors exhibit a transition below their critical temperature from a $\chi_m>0$ in the normal-metal state to a $\chi_m<0$ in the superconductive state. Localized spins, on the other hand, give a temperature-dependent paramagnetic susceptibility $\chi_m(T)>0$ that reflects not only the magnitudes of the atomic magnetic moments, but also the character of the spin–spin interactions. Moreover, the spin–spin interactions stabilize long-range magnetic order below a ferromagnetic Curie temperature T_c or an antiferromagnetic Néel temperature T_N. Several types of magnetic order are found.

"Ferromagnetism" refers to a parallel alignment of spins that gives a saturation magnetization $M_s(0)=\sum_i N_i\mu_{Ai}$ at $T=0$ K, where N_i is the number per unit volume of a transition-metal cation with moment μ_{Ai} and the sum is over all the different kinds of transition-metal cations that are present. The zero-field saturation magnetization $M_s(T)$ decreases with increasing temperature T, falling sharply to zero at $T=T_c$. In this case, the paramagnetic susceptibility obeys a Curie–Weiss law

$$\chi_m = C/(T-\theta) \qquad [19]$$

where the Curie constant for a single spin species is

$$C = Ng^2J(J+1)\mu_B^2/3k \qquad [20]$$

and the Weiss constant $\theta = CW$ reflects the strength of the ferromagnetic interactions through the W of the internal Weiss molecular field WM that aligns the spins. Short-range order above T_c makes the $\chi_m^{-1}(T)$ plot bend from linear as χ_m^{-1} approaches zero, which makes $T_c \leqslant \theta$.

"Antiferromagnetism" refers to a magnetic order that leaves the zero-field saturation magnetization $M(T)=0$ at all temperatures below a Néel temperature T_N. Néel antiferromagnetism consists of two like ferromagnetic subarrays aligned antiparallel to one another to give $M_s(T)=M_1(T)-M_2(T)=0$. In this case, a $W<0$ makes the Weiss constant of the Curie–Weiss law $\theta<0$, and χ_m^{-1} is finite at $T_N>0$. The antisymmetric term $\boldsymbol{D}_{ij}\cdot\boldsymbol{S}_i\times\boldsymbol{S}_j$ of eqn [18] may cant the spins of an antiferromagnet to give a weak ferromagnetism. That this weak ferromagnetic component is not due to an impurity is beautifully illustrated by α-Fe_2O_3, which has the crystallographic

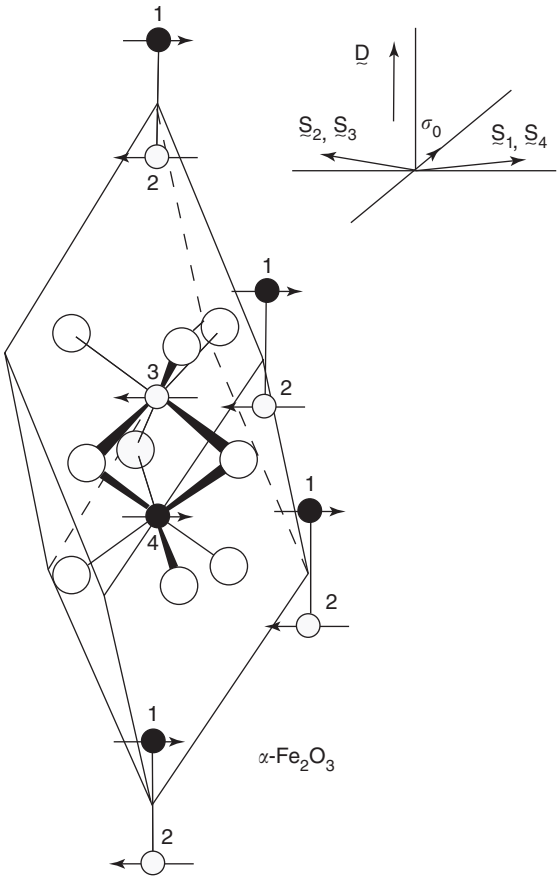


Figure 3 Antiferromagnetic spin configurations found in α-Fe_2O_3 in the temperature interval $250\,K<T<953\,K$ and the spin canting responsible for a weak ferromagnetic component σ_0. (After Shull CG, Strauser WA, and Wollan EO (1951) *Physical Review* 83: 333.)

and magnetic structure shown in **Figure 3** at room temperature. The weak ferromagnetic component has been used in magnetic recording. In this structure, the Dzialoshinskii vector $\boldsymbol{D}_{ij}$ is oriented parallel to the c-axis, so spin canting occurs if the spins lie in the basal plane. Dipole–dipole interactions orient the spins in the basal plane of α-Fe_2O_3 in the temperature interval $250\ K<T<T_N=953$ K; but below the Morin temperature $T_M=250\,K$, the spin orientation switches to the c-axis and the weak ferromagnetic component disappears abruptly.

"Ferrimagnetism" refers to an antiferromagnetic coupling of two ferromagnetic subarrays with magnetizations $M_1(T)$ and $M_2(T)$ where $M_s(T)=M_1(T)-M_2(T)\neq 0$ at all temperatures. Normally, $M_s(0)\neq 0$ in a ferrimagnet, but $Fe_2(SO_4)_3$ is an example where the Fe(III) atoms of the two subarrays are in sites that are crystallographically inequivalent, so $M_s(T)\neq 0$ although $M_s(0)=0$. The inverse paramagnetic susceptibility of a ferrimagnet is hyperbolic, approaching a Curie–Weiss law with $\theta<0$ at

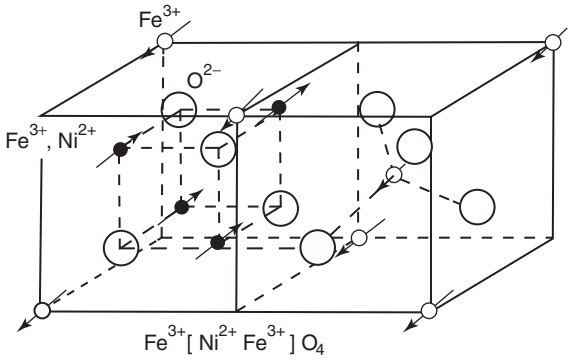


Figure 4 Ferrimagnetic order of $Fe[NiFe]O_4$. (After Hastings JM and Corliss LM (1953) *Reviews of Modern Physics* 25: 114.)

high temperatures, but falling to zero at a Curie temperature $T_c > 0$. The most celebrated ferrimagnets are the ferrospinels, which have an antiferromagnetic coupling of tetrahedral-site and octahedral-site cations. **Figure 4** illustrates the magnetic order in $Fe[NiFe]O_3$, where the octahedral-site cations are bracketed in the formula. In this case, the Fe(III) moments cancel one another and the net magnetization per formula unit is due to the nickel with $\mu_{Ni} \approx 2\mu_B$. Interestingly, substitution of Ni(II) by a nonmagnetic Zn(II) ion increases the magnetization because the Zn atoms prefer tetrahedral sites: $Zn_xFe_{1-x}[Ni_{1-x}Fe_{1+x}]O_4$ has a magnetization per formula unit for small x of $\mu_{mol} = (2 + 8x)\mu_B$. A rhombohedral component of the ligand field at an octahedral site of the spinel structure splits the t manifold into a nondegenerate a_1 orbital ($m = 0$) of lower energy and a twofold degenerate e_π-orbital manifold with $m = \pm 1$. Consequently, substitution of an Fe(II) for Ni(II) on the octahedral sites introduces mobile spins associated with the mixed iron valence, but it has little effect on the magnetocrystalline anisotropy. On the other hand, substitution of Co(II) for Ni(II) leaves the spinel an insulator but introduces a large magnetocrystalline anisotropy that increases linearly with the cobalt concentration as a result of splitting of the e_π-orbital degeneracy by $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$.

"Spin-density wave (SDW)" refers to a noncollinear spin configuration that has a propagation vector $\boldsymbol{q}$ with a wavelength that may be either commensurate or incommensurate with a lattice period. Parallel alignment of near-neighbor spins in a ferromagnet or within a subarray of a Néel antiferromagnet or ferrimagnet may oppose weaker antiferromagnetic interactions between next-near-neighbor spins of the array. In this case, the crystal compensates either by changing its structure so as to weaken the antiferromagnetic interactions or by forcing the spins to become canted or, more commonly, to form an SDW.

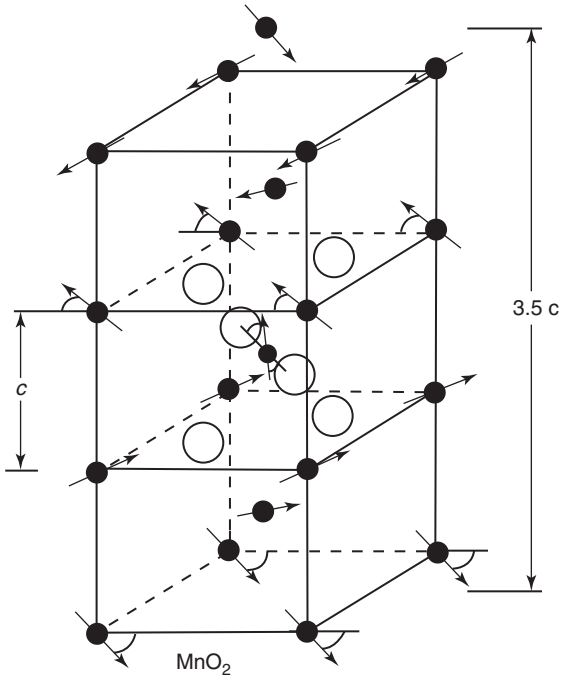


Figure 5 Spiral-spin configuration propagating parallel to the c-axis of MnO_2 (oxygen coordination shown for only one cation; cf. **Figure 9** for rutile structure). Turn angle between spins in successive (0 0 1) planes is $2\pi/7$. (After Erickson RA (1953) *Physical Review* 90: 779; Yoshimori A (1959) *Journal of the Physical Society of Japan* 14: 807.)

An SDW has an $M_s(T) = 0$ below T_N and is, therefore, antiferromagnetic. **Figure 5** illustrates a simple spiral-spin (helical) configuration found in the rutile structure of MnO_2. Canting of the octahedral-site spins is found in the tetragonal spinel $Cu[Cr_2]O_4$. **Figure 6** illustrates the spiraling of this canted-octahedral-site configuration observed in $Co[Cr_2]O_4$.

"Metamagnetism" refers to the transformation of antiferromagnetic order to ferromagnetic order in a moderate applied magnetic field. For example, an SDW having ferromagnetic nearest-neighbor interactions is transformed at 50 K in MnP to a ferromagnet on which a weak SDW component is superimposed.

An analogous, but not strictly metamagnetic, situation is found in some double perovskites $La_2M_aM_bO_6$ where ordering of the M_a and M_b cations introduces ferromagnetic M_a–O–M_b interactions within slabs, but leaves antiferromagnetic M_a–O–M_a and M_b–O–M_b interactions across antiphase boundaries. The ferromagnetic slabs are oriented antiparallel to one another in zero magnetic field, but the large magnetization of a slab allows a modest applied magnetic field to create a parallel orientation of the slabs except for spiral-spin interfaces centered at the antiphase boundaries.

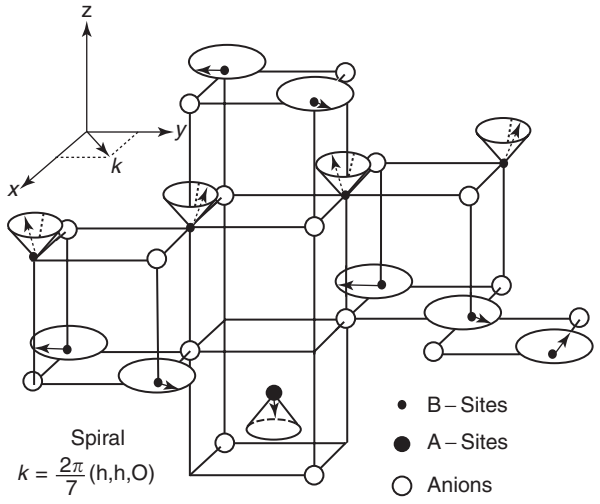


Figure 6 Complex spiral-spin configuration found in the spinel Co[Cr_2]O_4. Spiral propagates along the [1 1 0] direction. (After Dwight K and Manyuk N (1969) *Journal of Applied Physics* 40: 1156.)

"Frustrated systems" are those where strong antiferromagnetic interactions are unable to align all near-neighbor spins antiparallel to one another. This situation occurs, for example, in a close-packed plane or among the octahedral sites of a spinel. Frustration lowers the temperatures below which long-range magnetic order is found and leads to compromise spin configurations. In a close-packed plane with antiferromagnetic interactions between nearest neighbors, for example, the spins order at 120° rather than 180° with respect to one another.

"Spin glasses" are compounds containing ferromagnetic clusters in an antiferromagnetic matrix. Since the clusters are randomly distributed, the coupling between the clusters is frustrated, and the ferromagnetic regions assume random orientations below a spin-glass freezing temperature T_g. A "cluster glass" consists of antiferromagnetic clusters in a ferromagnetic matrix.

Several processes contribute to the microscopic origins of the spin–spin interactions.

"Direct exchange" between spins in orthogonal orbitals is ferromagnetic; it is a "potential exchange" as it does not involve electron transfer. The intraatomic exchange splitting Δ_{ex} responsible for Hund's highest multiplicity rule for the free atom is a direct exchange.

Interactions between spins on two atoms with overlapping orbitals is a "kinetic exchange." Two features are controlling, conservation of the spin angular momentum in the transfer and an angular dependence for electron spins having only the two directions α and β:

$$\alpha = \alpha' \cos(\theta/2) + \beta \sin(\theta/2) \qquad [21]$$

where θ is the angle between localized spins on the two neighboring atoms. Therefore, the spin-independent electron-transfer integral b_{ij} of eqn [3] is replaced by a spin-dependent integral

$$t_{ij}^{\uparrow\uparrow} = b_{ij}\cos(\theta/2) \quad \text{or} \quad t_{ij}^{\uparrow\downarrow} = b_{ij}\sin(\theta/2) \qquad [22]$$

"Superexchange" refers to interactions that require overcoming an energy barrier ΔE; in this case, the electron transfer is virtual and is treated in a second- or higher-order perturbation theory. Two cases must be distinguished:

1. If the overlapping orbitals are each half-filled, electron transfer is constrained by the Pauli exclusion principle to be antiferromagnetic and

$$\Delta E_{ex}^{S} \approx -|t_{ij}^{\uparrow\downarrow}|^2/U_{eff} = \text{const} + J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j \qquad [23]$$

   where $\Delta E = U_{eff}$ corresponds to the interaction in a single-valent compound; the $J_{ij} \sim (2b_{ij}^2)/4S^2U_{eff}$ reduces to $2b^2/U_{eff}$ for nearest neighbors with $S = 1/2$.
2. If the overlapping orbitals are "half-filled and empty" or "filled and half-filled," ΔE_{ex} favors a ferromagnetic charge transfer and

$$\Delta E_{ex}^{S} = \text{const} - J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j \qquad [24]$$

   with $J_{ij} \approx (2b_{ij}\Delta_{ex}/4S^2U_{eff}^2)$ follows from a third-order perturbation theory.

With half-filled t^3 configurations, the $(180° - \phi)$ t^3e^0–O–t^3e^0 interactions of the $RCrO_3$ (R=rare-earth) and $Sr_{1-x}Ca_xMnO_3$ perovskites are all antiferromagnetic as are the t^3e^2–O–t^3e^2 interactions in the $RFeO_3$ perovskites. An antisymmetric $\boldsymbol{D}_{ij} \cdot \boldsymbol{S}_i \times \boldsymbol{S}_j$ exchange term cants the antiparallel spins by a few degrees to give a weak ferromagnetic moment. On the other hand, the double perovskite La_2NiMnO_6 with ordered Ni(II) and Mn(IV) exhibits ferromagnetic Ni(II): t^6e^2–O–Mn(IV):t^3e^0 superexchange interactions. In the perovskite $LaMnO_3$, the cooperative Mn(III)-site distortions of **Figure 7** remove the *e*-orbital degeneracy of the t^3e^1 configuration and introduce $(180° - \phi)$ Mn(III)–O–Mn(III) superexchange interactions that are ferromagnetic in (0 0 1) planes ($\lambda_\sigma > \lambda_\pi$) and antiferromagnetic between basal planes. Here also, a spin canting gives a weak ferromagnetic component. However, dilution of the Mn(III) ions in $LaMn_{1-x}Ga_xO_3$ introduces regions of orbital disorder at lower temperatures that are ferromagnetic in three dimensions because locally cooperative distortive fluctuations favor electron transfer when a half-filled and an empty *e*-orbital overlap. This 3D

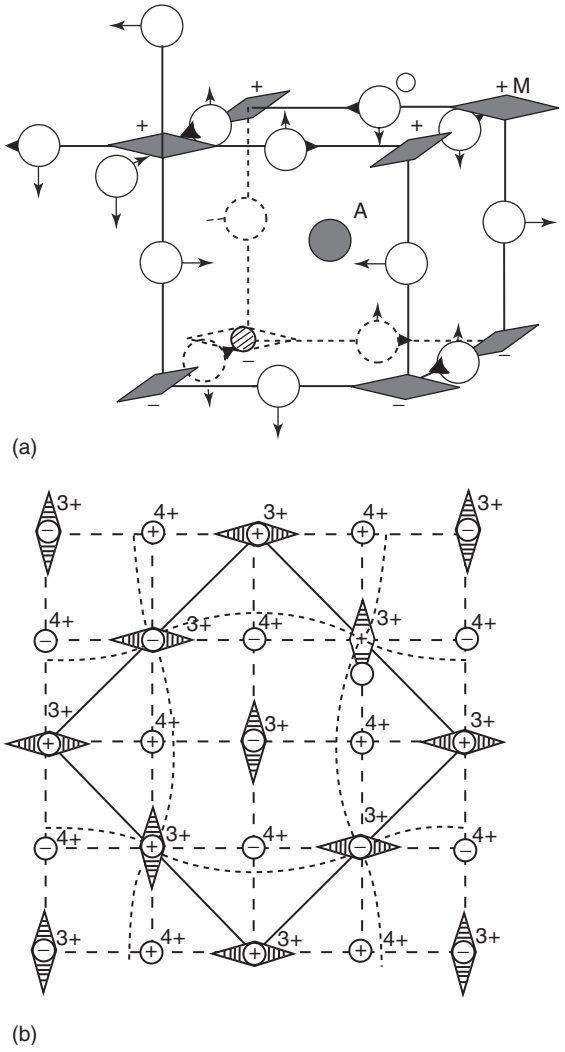


Figure 7 (a) Ordering of occupied *e* orbitals into (0 0 1) planes of $LaMnO_3$ and associated oxygen displacements. This cooperative orbital ordering is superimposed on a cooperative rotation of the $MnO_{6/2}$ octahedra (shown by arrows). Cooperative rotations bend the $(180^\circ - \phi)$ B–O–B bond angle in ABO_3 perovskites when the geometric tolerance factor $t \equiv (\text{A–O})/\sqrt{2}\,(\text{B–O}) < 1$; (A–O) and (B–O) are the equilibrium (A–O) and (B–O) bond lengths. (b) Charge, orbital, and spin [up (+), down (−)] ordering in $La_{0.5}Ca_{0.5}MnO_3$.

ferromagnetic interaction represents a "vibronic superexchange." Where the ferromagnetic Curie temperature T_c of the orbitally disordered regions is higher than the Néel temperature T_N of long-range antiferromagnetic order in the orbitally ordered matrix, an applied magnetic field stabilizes the orbitally disordered phase relative to the orbitally ordered phase to convert a spin glass into a ferromagnet.

"Semicovalent exchange" involves a two-electron transfer from a single anion-*p* orbital, one to each of the two cations on opposite sides in a $(180^\circ\text{–}\phi)$ cation–anion–cation interaction. Since the two transferred electrons have opposite spins, the rules for ferromagnetic versus antiferromagnetic coupling are the same as those for superexchange. A $U_{\text{eff}} \approx 0$ for the itinerant O–2*p* electrons gives semicovalent exchange the form $-J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j$ with $J_{ij} \approx (2b_{ij}^2/8S^2\Delta)$, where $\Delta = \Delta E_p$ is the "charge-transfer energy gap." To illustrate the significance of the semicovalent-exchange component to the total exchange, consider again the $(180^\circ - \phi)\ t^3e^0 - \text{O} - t^3e^0$ interactions in the $RCrO_3$ and $Ca_{1-x}Sr_xMnO_3$ perovskites

$$\Delta E_{\text{ex}}^S \approx -\sum_{ij}{}' (2b_{ij}^2/4S^2)[1/U_{\text{eff}}) + (1/2\Delta)] \qquad [25]$$

Whereas the superexchange interactions treat only the π-bonding t^3–O–t^3 interactions, semicovalent exchange also includes the σ-bonding interactions if a $(\Delta_{ex}/\Delta)^2$ multiplier is added. A $\Delta \leqslant U_{\text{eff}}$ in $RCrO_3$ and $Ca_{1-x}Sr_xMnO_3$ makes the semicovalent-exchange component dominant to give a total exchange-energy parameter $J_{ex} \sim \langle \cos^2\phi > \rangle$ (cf. eqn [13]). A $T_N \sim \langle \cos^2\phi \rangle$ in an RMO_3 perovskite family signals localized-electron antiferromagnetism.

"Double exchange" occurs in mixed-valent transition-metal compounds in which a mobile spin has a $\tau_h < \omega_R^{-1}$ in the presence of localized spins. The mobile spin couples to the localized spin by the Hund intra-atomic exchange field to give a ferromagnetic interaction with a spin-dependent $t_{ij}^{\uparrow,\uparrow} = b_{ij}\cos(\theta/2)$ in the tight-binding formula of eqn [4], so

$$H_{\text{ex}}^D \approx -Nc(1-c)zb_{ij}\cos(\theta/2) \qquad [26]$$

where Nc is the concentration of mobile charge carriers and $(1-c)z$ is the fraction of near-neighbor sites that are empty. Polaronic charge carriers do not introduce a double-exchange coupling as their τ_h is long compared to the time for a spin to relax to a disordered orientation. Equation [24] is the de Gennes formulation that he applied to the metallic ferromagnetic compositions of the system $La_{1-x}Sr_xMnO_3$; it assumes the coexistence of conductive $(\tau_h < \omega_R^{-1})$ σ-bond electrons in the presence of localized spins $S = 3/2$ from the π-bond t^3 manifold. The original Zener formulation for double exchange in this system constrained the $\tau_h < \omega_R^{-1}$ to a Mn^{3+}–O–Mn^{4+} two-manganese polaron that he assumed to move diffusively, but with no motional enthalpy ($\Delta H_m = 0$). In fact, the electrons of *e*-orbital parentage undergo a crossover with increasing x from polaronic to itinerant electronic behavior in this system. **Figure 8** shows a small compositional and temperature range where the Zener model is applicable; it also reveals a discontinuous change in T_c at the transition from Zener to de Gennes double-exchange as predicted from the

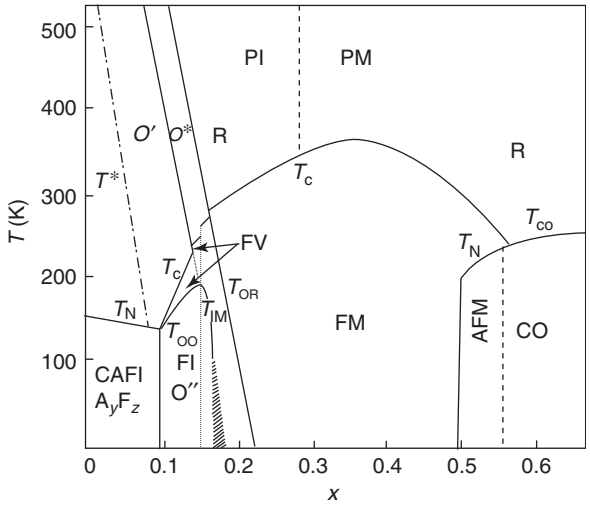


Figure 8 Phase diagram of the system $La_{1-x}Sr_xMnO_3$. CAFI=canted-spin type-A antiferromagnetic insulator [ferromagnetic (0 0 1) planes coupled antiparallel with orbital ordering of **Figure 7a**]. FI, FV, and FM=ferromagnetic insulator, vibronic conductor (Zener polarons), and de Gennes metal, respectively. AFM=type-A antiferromagnetic metal. CO=charge-ordered phase. O′ and O″ refer to different orbitally ordered orthorhombic phases, O* to an orbitally disordered orthorhombic phase, and R to a rhombohedral phase. PI and PM are, respectively, paramagnetic-insulator and paramagnetic-metal phases. (After Goodenough JB and Zhou J-S (2001) *Structure & Bonding* 98: 87.)

virial theorem for a transition from polaronic to itinerant-electron conduction. An orbital ordering that lifts the degeneracy of the σ^* band of *e*-orbital parentage stabilizes the AFM phase of **Figure 8**; metallic, ferromagnetic (0 0 1) planes are coupled antiparallel to one another along the *c*-axis by t^3–O–t^3 superexchange. However, charge and orbital ordering in $La_{0.5}$ $Ca_{0.5}$ MnO_3 traps the mobile charge carriers to give zigzag ferromagnetic chains coupled antiparallel to one another in the (0 0 1) planes, see **Figure 7b**. The Mn^{3+} ions are confined to stripes of zigzag chains, as shown in the CO phase in **Figure 8**.

The spinel Fe_3O_4 has mixed Fe(II) and Fe(III) valence states on the octahedral sites, and an antiferromagnetic Fe–O–Fe superexchange interaction couples antiparallel to the tetrahedral-site and octahedral-site spins to give a ferrimagnetic moment per Fe_3O_4 molecule of $\sim 4\,\mu_B$; the spins of the two Fe(III) ions cancel one another, and the ferromagnetic alignment of the octahedral-site spins reduces the spin-disorder scattering for motion of the minority-spin electrons in the presence of the localized majority-spin electrons to give a $\tau_h \leqslant \omega_R^{-1}$ between room temperature and a CDW Verwey transition temperature $T_V \approx 120\,K$. The mobile electrons give a ferromagnetic double-exchange interaction between the octahedral-site cations below room temperature. Fe_3O_4 represents another situation where itinerant and localized spins can coexist on the same atom. However, at higher temperatures spin-disorder scattering converts a $\tau_h \leqslant \omega_R^{-1}$ to a $\tau_h > \omega_R^{-1}$, and the mobile electrons become polaronic.

"Indirect exchange," also known as Yaffet–Kittel–Kasuya–Yosida (YKKY) exchange, is a coupling of localized spins by a partially filled band of itinerant electrons; it is an extension of the de Gennes double exchange to a broad itinerant-electron band. Intra-atomic exchange stabilizes an itinerant-electron spin density parallel to the localized spin in the vicinity of the transition-metal atom, but this itinerant-electron spin density decreases at larger distances, reversing sign in a series of oscillations. If a neighboring localized spin is close enough to be overlapped by the parallel spin density, then the localized spins are coupled ferromagnetically; but if the antiparallel spin density overlaps the neighboring spins, they couple antiferromagnetically. The extension of the parallel spin-density domain increases as the concentration of itinerant electrons decreases and their effective mass increases.

Additional Illustrative Examples

Spinels

The dominant interaction in the ferrospinel Fe[NiFe]O_4 is the antiferromagnetic 135° t^3–O–e^2 interaction between half-filled σ-bonding orbitals on the tetrahedral-site Fe(III):e^2t^3 and the octahedral-site Ni(II):t^6e^2 and Fe(III):t^3e^2 ions.

In an A[M_2]O_4 spinel in which the tetrahedral-site A cation has no spin, the only spin–spin interactions are those between octahedral-site transition-metal atoms M. The octahedral-site cations in the spinel structure of **Figure 4** share only octahedral-site edges; they interact with one another by M–M and 90° M–X–M interactions. The octahedral-site Cr(III):t^3e^0 cations in Zn[Cr_2]O_4, Zn[Cr_2]S_4, and Cd[Cr_2]S_4 all have ferromagnetic 90° Cr–X–Cr interactions by direct exchange and t^3–$p_{\pi\sigma}$–e^0 superexchange whereas the t^3–t^3 interactions are antiferromagnetic by superexchange. The antiferromagnetic t^3–t^3 interactions decrease exponentially with increasing Cr–Cr separation, that is, with increasing lattice parameter on going from Zn[Cr_2]O_4 to Zn[Cr_2]S_4 to Cd[Cr_2]S_4. The 90° Cr–X–Cr interactions, on the other hand, increase significantly on going from an oxide to a more covalent sulfide. Experimentally, Cd[Cr_2]S_4 is a ferromagnetic insulator, the Cr–Cr interactions in Zn[Cr_2]S_4 are just strong enough to convert the ferromagnetic coupling into an SDW, and in Zn[Cr_2]O_4 stronger antiferromagnetic Cr–Cr interactions are frustrated to give a low T_N.

As pointed out above, Fe_3O_4 is a good electronic conductor because of mixed Fe(III) and Fe(II) valence states on the octahedral sites. On the other hand, Mn_3O_4 and Co_3O_4 are insulators because the Mn(II) and Co(II) ions occupy tetrahedral sites having a different potential energy for the M(II)/M(III) couple than that at the octahedral sites. With Mn(III):t^3e^1 on octahedral sites, Mn_3O_4 undergoes a cooperative Jahn–Teller orbital ordering that distorts the spinel from cubic to tetragonal symmetry; the room temperature axial ratio is c/a=1.16. Fe_3O_4 is a Néel ferrimagnet with ferromagnetic double-exchange coupling among the octahedral sites whereas antiferromagnetic t^3–t^3 interactions between octahedral-site Mn(III) lead to a complex magnetic order. Co_3O_4 is a Néel antiferromagnet because the octahedral-site Co(III):t^6e^0 ions are in their low-spin state with $S = 0$ and the tetrahedral sites consist of two interpenetrating face-centered-cubic arrays as in the diamond structure with antiferromagnetic t^3–O–t^3 superexchange interactions between the Co(II):e^4t^3 ions.

Rutiles

The MX_2 compounds with X = F or O crystallize in the rutile structure of **Figure 9**. This structure has a body-centered-tetragonal array of octahedral-site cations that share edges along the c-axis; these c-axis chains share octahedral-site corners. The anions form sp^2 σ-bonds with three cation neighbors and π-bond with two quasidegenerate t orbitals, e_π; the other t orbital, $d_{||}$, is directed along the c-axis where it overlaps with the corresponding $d_{||}$ orbitals of two next-near-neighbor cations. The MO_2 oxides illustrate several features discussed above.

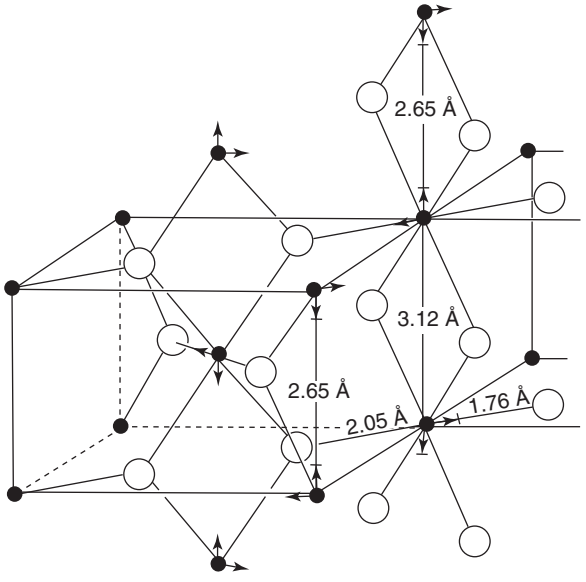


Figure 9 The tetragonal rutile structure and the cation displacements (arrows) found in VO_2 at $T < 67°C$.

The bottom of the empty 3d bands of TiO_2 lies $\sim$3 eV above the top of the filled O–2p bands, which makes TiO_2 an insulator and a crystal transparent. However, the elastic energy that would resist displacement of a Ti(IV) from the center of its interstice is not large, which is why $BaTiO_3$ and $PbTiO_3$ are ferroelectrics. In TiO_2, there is no spontaneous displacement, but displacements in an electric field give TiO_2 a large dielectric constant, which makes TiO_2 a white pigment widely used in paint and the finishing of paper. The itinerant-electron 3d bands can be doped n-type, in which case n-TiO_2 can be used for photoelectrolysis. Removal of oxygen is accommodated by an ordering of the oxygen vacancies into "shear planes" that separate slabs, n rutile cells wide, in a Ti_nO_{2n-1} family of shear structures. Ti(IV) octahedra share common faces across a shear plane, and electrostatic forces displace the cations toward opposite site faces; itinerant electrons reside on the Ti cations within the rutile slabs. The formation of shear planes allows cooperative displacements of the Ti(IV) cations, which reduces the cost in elastic energy. The shear-plane phenomena are associated with cations that also exhibit ferroelectric displacements in other contexts. At low temperatures, the itinerant electrons within the 3d bands of a slab condense into $d_{||} - d_{||}$ homopolar bonds to form pairs of Ti(III) displaced toward one another. In Ti_4O_7, these two-electron bonds are mobile and disordered in a small temperature interval above the long-range ordering temperature $T_t \approx 160\,K$. The mobile electron pairs represent bipolarons.

Above 67°C, the single 3d electron of a V(IV) cation occupies overlapping itinerant-electron bands of $d_{||}$ and e_π parentage; the metallic conductivity is nearly isotropic. The $d_{||}$ band is formed by strong V–V interactions along the c-axis: the π^* bands by strong e_π–O:2p_π–e_π interactions with nearest and next-nearest neighbors. Below 67°C, these itinerant electrons condense into $d_{||}$–$d_{||}$ homopolar bonds and the smaller V(IV) cations are also displaced toward a corner-sharing oxygen by a rocking of the pairs to form molecular clusters consisting of two vanadyl $(V{=}O)^{2+}$ cations bonded by a homopolar V–V bond. The oxygen atoms that bridge a V–V bond each belong to another such cluster. The resulting monoclinic-tetragonal first-order transition at T_t = 67°C produces an insulator–metal transition across which the formation of the V=O vanadyls lifts the π^* bands above the Fermi energy, leaving the $d_{||}$ band half-filled, and V–V pairing introduces a gap in the narrow $d_{||}$ band that localizes the electrons in homopolar bonds. This localization of electrons to a

molecular cluster is to be distinguished from formation of a 1D CDW as envisioned by a "Peierls distortion" in which the electrons of the pairs remain itinerant.

In CrO_2, the Cr–Cr separation along a *c*-axis chain is too large ($R > R_c$) for the $d_{||}$ electrons to form a band of itinerant-electron states; they form localized $d_{||}^1$ states separated from a $d_{||}^2$ configuration by a finite energy gap ($U - W_{||}$). The remaining single e_π electron occupies a π^* band of itinerant-electron states. This band is one-quarter filled, and extrapolation of the superexchange rules to band magnetism would predict ferromagnetic correlations among the e_π electrons. Alternatively, the de Gennes model of double-exchange with strong intra-atomic exchange between the localized $d_{||}$ and itinerant e_π electrons also predicts ferromagnetism. Both considerations apply, and CrO_2 is a ferromagnetic metal. Below the Curie temperature T_c, the *c*-axis expands as a result of "exchange striction"; the half-filled $d_{||}^1$ orbitals would be stabilized by antiferromagnetic superexchange, but their parallel alignment removes the *c*-axis bonding by the $d_{||}$ electrons because the Pauli exclusion principle prevents $d_{||}$-electron transfer if the spins are parallel.

In MnO_2, both the $d_{||}$ and the e_π electrons are localized and a $d_{||}^1 e_\pi^2$ manifold means that both sets of orbitals are half-filled, which leads to antiferromagnetic interactions between nearest and next-nearest neighbors by superexchange. Since both nearest and next-nearest neighbors share a common O–2p_π orbital, the antiferromagnetic interactions along the *c*-axis are stronger than those between nearest neighbors, and the competition leads to the antiferromagnetic spiral-spin configuration propagating along the *c*-axis that is illustrated in **Figure 5**.

The 4*d* electrons have a greater radial extension than the 3*d* electrons, so they have larger overlap integrals with neighboring atoms than do the 3*d* orbitals. Consequently, NbO_2 and MoO_2 do not have localized 4*d* electrons; they both form *c*-axis M–M pairs, but without the rocking of the pairs that occurs in VO_2. The homopolar-bond states of NbO_2 lie below the bottom of the π^* band of e_π-orbital parentage, and NbO_2 is a semiconductor below $T_t \approx 800^\circ C$. With one additional 4*d* electron, MoO_2 has one electron per formula unit in the π^* bands and is a metal. Low-spin Ru(IV) and Rh(IV) are also accessible in octahedral sites of an oxide, but they contain too many 4*d* electrons to allow stabilization of homopolar bonds in M–M pairs. With two and one hole per formula unit, respectively, in the $d_{||}$ and π^* bands, these oxides are metallic and retain the rutile structure to lowest temperatures.

Energy Bands and Redox Couples

Electron Energies

In an ionic model of a compound, which is illustrated in **Figure 10** for MnO, the energy E_I required to remove the last electron from the cation and place it on an O^- ion at infinite separation is more than compensated by the electrostatic Madelung energy $E_M > E_I$ that is gained by assembling the point charges in a crystalline array. Conservation of energy raises the cation energy levels and lowers the anion energies. The crystalline electric field at the octahedral-site Mn^{2+} of MnO raises the Mn:3d^5 level into the energy gap $E_g \approx 6$ eV between the filled O^{2-}:2p^6 and empty Mn^{2+}:4s^0 bands. Covalent back transfer of electrons from the O^{2-} to the Mn^{2+} ions reduces E_M, but it introduces a compensating quantum-mechanical repulsion between the bonding O-2*p* and antibonding Mn-4*s* orbitals to retain a large E_g. The covalent mixing and the like-atom interactions broaden the bonding and antibonding *s* and *p* states into bands of one-electron itinerant-electron states, and it is convenient to label the upper bonding bands as O^{2-}:2p^6 bands and the lower antibonding bands as Mn:4s^0.

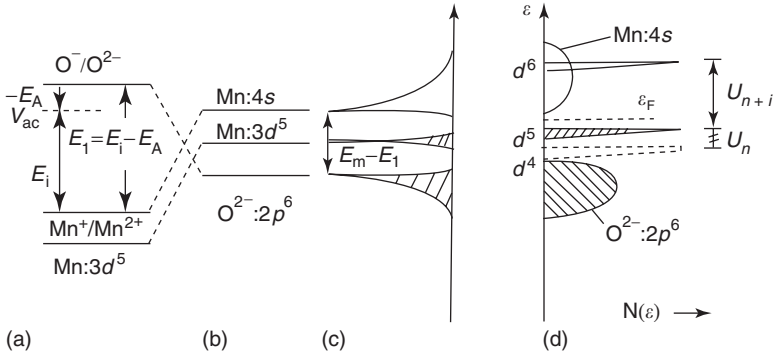


Figure 10 Schematic construction of electronic energies from the point-charge model, (a) and (b), to energy bands for Mn:4s^0 and O^{2-}:2p^6 in (c) and (d).

Where a d^n or $4f^n$ manifold falls in the energy gap E_g, as occurs in MnO, further oxidation of the cation is possible. Similarly, where an empty d^{n+1} manifold lies in the gap, reduction of the cation is possible if this manifold is not too close in energy to the bottom of the broad *s* and *p* antibonding bands. In the case of octahedral-site manganese, three valence states are stable: Mn(II), Mn(III), and Mn(IV) corresponding to Mn(III)/Mn(II) and Mn(IV)/Mn(III) redox couples lying in the gap. The splitting $U_{\text{eff}} - U_\sigma$ between the Mn(II):t^3e^2 and Mn(III):t^3e^1 energies is small; in fact, W_σ approaches U_σ in the perovskites $LaMnO_3$ and $SrMnO_3$ containing Mn(III) and Mn(IV), respectively, and a $\tau_h < \omega_R^{-1}$ for the σ^* electrons makes $La_{0.7}Sr_{0.3}MnO_3$ a ferromagnetic metal.

Since $E_M - E_I$ decreases on going from the halides (group VII) to the oxides and chalcogenides (group VI) to the pnictides (group V) and down any column of the periodic table to heavier anions, the energy gap E_g between bonding and antibonding *s* and *p* bands decreases accordingly. Consequently, the higher cation valence states accessible in an oxide may not be available in a sulfide or selenide. For example, the Mn^{2+}:d^5 manifold lies below the S^{2-}:$3p^6$ energy in the point-charge model, so the Mn(IV) state is not accessible in the chalcogenides. MnS_2 is $Mn^{2+}(S_2)^{2-}$, for example, in contrast to $Mn^{4+}(O^{2-})_2$ in MnO_2. However, where a d^n manifold lies close to the top of the bonding anion-*p* bands, covalent mixing raises antibonding states of *d*-orbital symmetry to the top of the broad bonding bands; the redox couple becomes pinned at the top of the bonding bands and changes from primarily *d* to primarily anion-*p* character as the d^n manifold falls below the anion-p^6 energy in the point-charge model. Where the anion component becomes dominant, the electronic states of the redox couple are itinerant; and where both the hole concentration and the anion character in these itinerant antibonding states are large enough for a cooperative condensation of the holes into purely anion antibonding states of anion–anion pairs, diatom anions are formed.

Rare-Earth Compounds

An energy $U > E_g \gg W$ separates the $4f^n$ and $4f^{n+1}$ localized-electron manifolds of the rare-earth ions. Therefore, they have a single R^{3+} valence state unless a $4f^n$ configuration falls within E_g, as occurs for the Ce^{3+}:$4f^1$ and Pr^{3+}:$4f^2$ or the Eu^{2+}:$4f^7$ in an oxide. However, the R-5*d* orbitals are itinerant, and covalent bonding pushes the antibonding 6*s* and 6*p* states above the bottom of the 5*d* band. Normally the 5*d* band is empty, but it can accept electrons. Doping EuO with Gd substitution for Eu, for example, creates shallow one-electron donor states below the bottom of the 5*d* conduction band just as substitution of P for Si creates *n*-type silicon. In EuO, the Eu:$4f^7$ level lies $\sim$1.1 eV below the bottom of the 5*d* band; the $4f^7$–O–$5d^0$ superexchange interactions are ferromagnetic and stronger than the antiferromagnetic $4f^7$–O–$4f^7$ interaction because of the large U_{eff} between $4f^7$ and $4f^8$ as well as the smaller overlap integrals. In the $Eu_{1-x}Gd_xO$ system, the YKKY indirect exchange by itinerant 5*d* electrons initially increases the ferromagnetic Curie temperature T_c with *x*; but as *x* increases, the extension of the ferromagnetic 5*d* volume about a rare-earth atom decreases to below a nearest-neighbor separation; the YKKY interaction becomes antiferromagnetic, and T_c decreases with a further increase of *x*. $Eu_{1-\delta}O$ contains polaronic $\tau_h > \omega_R^{-1}$ holes in the Eu(III)/Eu(II) redox couple. On the other hand, annealing in a Eu atmosphere creates $EuO_{1-\delta}$, and the oxygen vacancies trap two electrons. On cooling $EuO_{1-\delta}$ below the Curie temperature T_c, the 5*d* states with spin parallel to the $4f^7$ spin $S = 7/2$ are stabilized relative to the 5*d* states of antiparallel spin by direct intra-atomic exchange. The spin degeneracy of the two-electron trap state is also removed, the energy of the antiparallel trapped electron is raised above the bottom of the parallel-spin 5*d* conduction band. As a result, $EuO_{1-\delta}$ transforms from a semiconductor to a metal on cooling through T_c as is illustrated in **Figure 11**. Just above T_c, the carriers thermally excited into the 5*d* band stabilize regions of short-range ferromagnetic order by double-exchange electron transfer; ferromagnetic alignment reduces the spin-disorder scattering of the mobile electrons, and their stabilization by the intra-atomic exchange interaction confines the mobile electrons to the region of short-range ferromagnetic order. A mobile electron confined to clusters of short-range ferromagnetic order is called "magnetic polarons." In an external magnetic field, the size of the ferromagnetic clusters grows; at a percolation threshold, the compound undergoes a drop in electrical resistance. This magnetoresistance is large in EuO.

The valence state of a rare-earth ion is defined by its occupied $4f^n$ configuration. $Eu_{1-\delta}O$ is in a mixed Eu^{3+}/Eu^{2+} valence state since the holes occupy the Eu^{3+}/Eu^{2+} $4f^6/4f^7$ couple. GdS, on the other hand, has the Gd configuration $4f^75d^1$, and the itinerant 5*d* electrons that make GdS a metal are valence electrons. Therefore, the Gd of GdS is considered to be in the trivalent state corresponding to a localized $4f^7$ configuration. SmS has its $4f^6$ level just below the bottom of the 5*d* band; pressure increases the width of the 5*d* band, and above 6.5 kbar the 5*d* band overlaps the $4f^7$ level. A first-order contraction of the

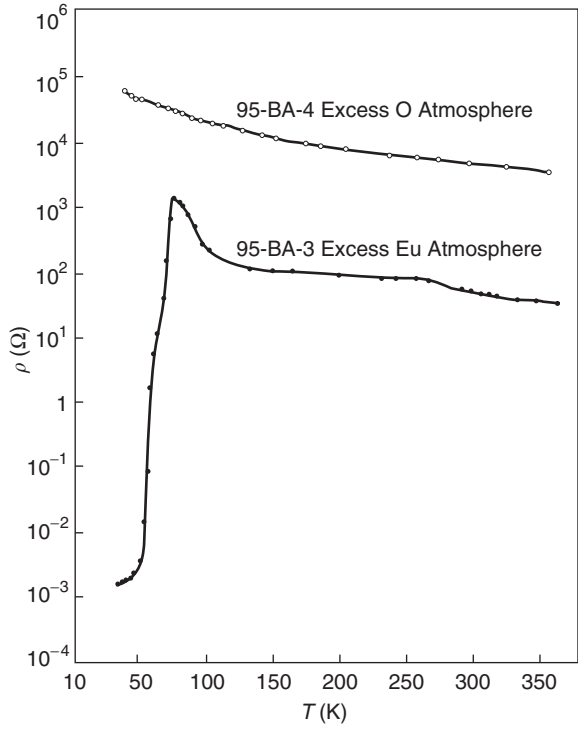


Figure 11 Resistivity vs. temperature for EuO annealed in a europium-rich atmosphere ($EuO_{1-\delta}$) and an oxygen-rich atmosphere ($Eu_{1-\delta}O$). (After Oliver MR, Dimmock JO, McWhorter AL, and Reed TB (1972) *Physical Review B* 5: 1078.)

rocksalt structure at crossover raises the $4f^6$ level into the $5d$ band to where 0.5 electrons per Sm are donated to the $5d$ band. Hybridization of $5d$ and $4f$ states creates "heavy fermions" at the Fermi energy and the compound is said to be in an intermediate-valence state.

Transition-Metal *d*-Block Compounds

The smaller U_{eff} of the d-block transition-metal compounds makes localized $4d$ and $5d$ spins relatively rare; localized spins associated with $3d$ electrons are common, but they may also be itinerant as discussed above for oxides with a rutile structure. The existence of four valence states for vanadium in V_2O_5, VO_2, V_2O_3, and VO shows that the $U_{\text{eff}} = U_\pi$ separating octahedral-site t^1, t^2, and t^3 configurations is in the neighborhood of 2 eV, which makes $W \approx U_{\text{eff}}$ in these compounds. Both VO_2 and V_2O_3 exhibit insulator–metal transitions due to cation clustering, V_2O_5 contains short V═O bonds due to displacement of the V(V) from the center of symmetry of its interstice, and VO is a semimetal exhibiting V–V bond-length fluctuations. On the other hand, the Ti–Ti interactions in metallic TiO make $W > U_\pi$; it becomes a conventional superconductor below $T_{cs} = 1$ K.

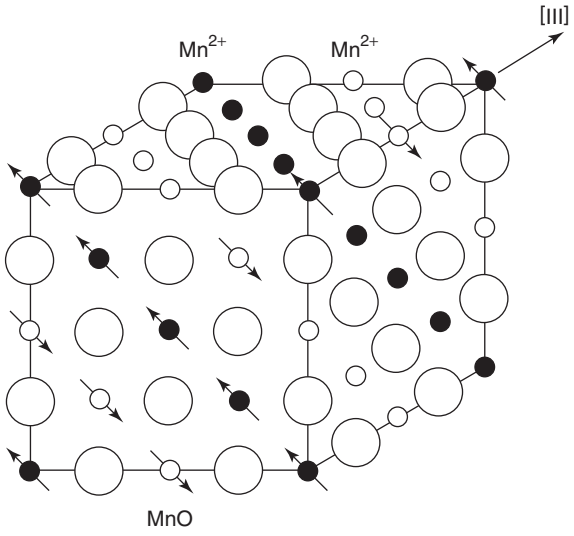


Figure 12 Spin configurations of antiferromagnetic MnO. (After Shull CG, Strauser WA, and Wollan EO (1951) *Physical Review* 83: 333.)

For a given valence state, the effective nuclear charge seen by the d electrons increases on going to heavier atoms of a given d-block and the redox energy is correspondingly stabilized. However, in an octahedral site, the splitting U_{eff} between a Cr(III):t^3e^0 and a Cr(II):t^3e^1 configuration is increased by the cubic-field splitting Δ_c, which is why CrO has not been prepared; the Cr(III)/Cr(II) level lies close to the bottom of the $4s$ band. However, CrS has been prepared; it has hexagonal rather than cubic close-packing of the anions (NiAs rather than NaCl structure) with a cooperative Jahn–Teller orbital ordering that distorts the structure from hexagonal symmetry. The d^5 configuration of MnO is more stable than the d^6 configuration of $Fe_{1-\delta}O$ because a $U_{\text{eff}} = U_\pi + \Delta_{ex}$ separates the Fe(III):t^3e^2 and Fe(II):t^4e^2 octahedral-site states. Since the majority-spin electrons occupy different d orbitals, they are more tightly bound to the nucleus than the minority-spin electron, which is why the minority-spin electron of Fe_3O_4 can have a $\tau_h \approx \omega_R^{-1}$.

The antiferromagnetic order in MnO is illustrated in **Figure 12**. The dominant interactions are 180° Mn(II)–O–Mn(II) antiferromagnetic superexchange interactions between half-filled t^3 and e^2 configurations; these interactions order parallel the spins of a (1 1 1) plane, whereas the t^3–t^3 superexchange interactions across shared octahedral-site edges within these planes are antiferromagnetic. Parallel spin alignment makes these t^3–t^3 interactions nonbonding, whereas those between neighboring, antiparallel (1 1 1) planes are bonding. Consequently, MnO distorts from cubic to rhombohedral ($\alpha > 60°$)

symmetry on cooling through T_N as a result of exchange striction. Similar antiferromagnetic 180° e^2–O–e^2 superexchange interactions are dominant in $Fe_{1-\delta}O$ and CoO, but these compounds exhibit a distortion below T_N that is due to a large magnetostriction; the orbital degeneracies of the Fe(II):t^4e^2 and Co(II):t^5e^2 configurations are removed by an enhancement of the orbital angular momentum once long-range collinear ordering of the spins allows ordering of the orbital angular momentum through the spin–orbit coupling $\lambda \boldsymbol{L} \cdot \boldsymbol{S}$. $Fe_{1-\delta}O$ becomes rhombohedral ($\alpha < 60°$) below T_N with spins oriented parallel to the [1 1 1] axis in opposition to the dipole–dipole forces; CoO becomes tetragonal ($c/a < 1$) and the dipole–dipole interactions force the spins a little away from the tetragonal c-axis.

Pinning of a redox couple at the top of an anion-p band replaces the gap U_{eff} with a charge-transfer gap $\Delta = \Delta E_p$. The Ni(II):t^6e^2 configuration of NiO is pinned at the top of the O–2p bands, and a $\Delta < U_{eff} = U_\sigma + \Delta_{ex}$ makes semicovalent exchange the dominant antiferromagnetic Ni(II):e^2–O–Ni(II):e^2 interaction responsible for its high Néel temperature. In NiS, the Ni(I):t^6e^3 configuration is at the top of the S–3p bands. At temperatures $T > 350$°C, NiS has the NiAs structure, but it transforms into the peculiar millerite structure at lower temperatures. It is possible to retain the NiAs structure at low temperatures by quenching from above 650°C. This metastable phase undergoes a first-order insulator–metal transition at a Néel temperature $T_N = 264$ K. On lowering the temperature through T_N, there is an abrupt increase in volume with only a small distortion of the structure as the width of the σ^* band of e-orbital parentage changes from $W_\sigma \approx U_{eff} = U_\sigma + \Delta_{ex}$ to $W_\sigma < U_{eff}$; the Mott–Hubbard transition is not smooth.

The lattice instabilities occurring where $W \approx U_{eff}$ are more commonly expressed by either stabilization of a CDW or, in a mixed-valent compound, by a dynamic spinodal phase separation. For example, the low-spin Ni(III):t^6e^1 configuration in the $RNiO_3$ family of **Figure 13** is pinned at the top of the O–2p bands; this single-valent family undergoes a transition from an antiferromagnetic insulator with $W < U_\sigma$ to metallic $LaNiO_3$ with $W > U_\sigma$ as the $(180° - \phi)$ Ni–O–Ni bond angle increases with increasing size of the rare-earth R^{3+} ion. The samples in the shaded area exhibit bond-length fluctuations and stabilization of a CDW below a $T_t = T_N$ that are characteristic of the lattice instabilities encountered at crossover. Below the insulator–metal transition T_{MI}, a distortion to monoclinic $P2_1/n$ symmetry reflects bond-length fluctuations that become static for R = HO,…, Lu, Y with long-range ordering of Ni (III) ions distinguished by alternating covalent and ionic Ni–O bond lengths. Moreover, applying

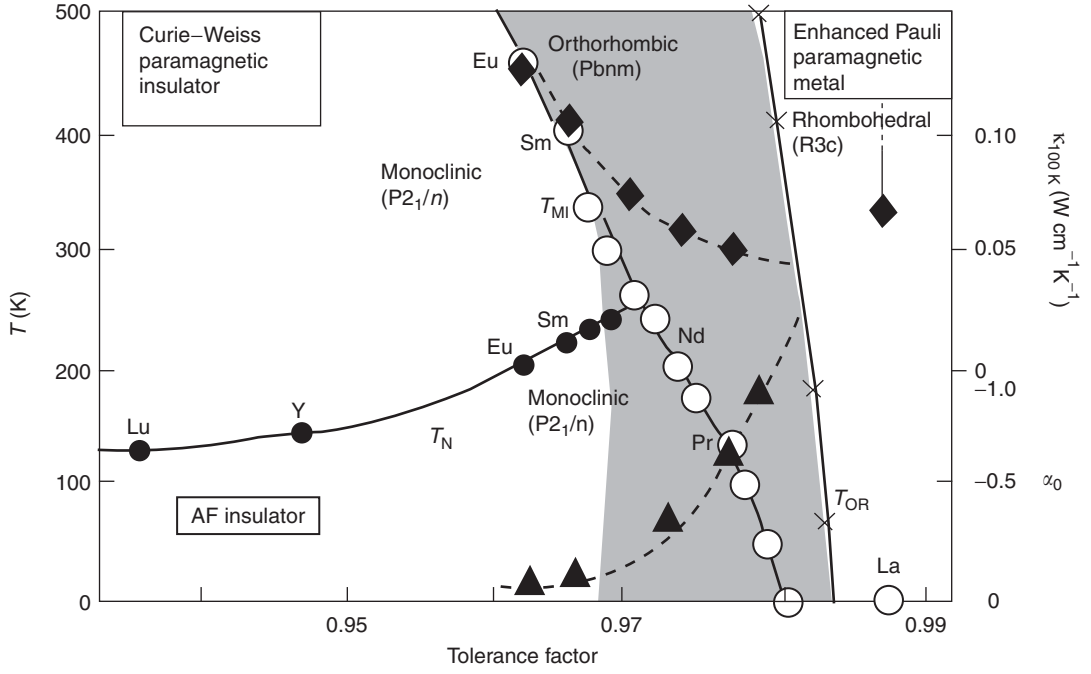


Figure 13 Schematic phase diagram of the $RNiO_3$ (R=rare-earth) family. The triangle symbol is for the $^{18}O/^{16}O$ isotope effect, the diamond is for the lattice contribution to the thermal conductivity at 100 K; both of these signal bond-length fluctuations in the shaded region. In the nonshaded region, a $T_N \sim \langle \cos^2\phi \rangle$ indicates localized-electron antiferromagnetism; $LaNiO_3$ is metallic. (Reproduced with permission from Zhou J-S, Goodenough JB, and Dabrowski B (2003) *Physical Review B* 67: 020404.)

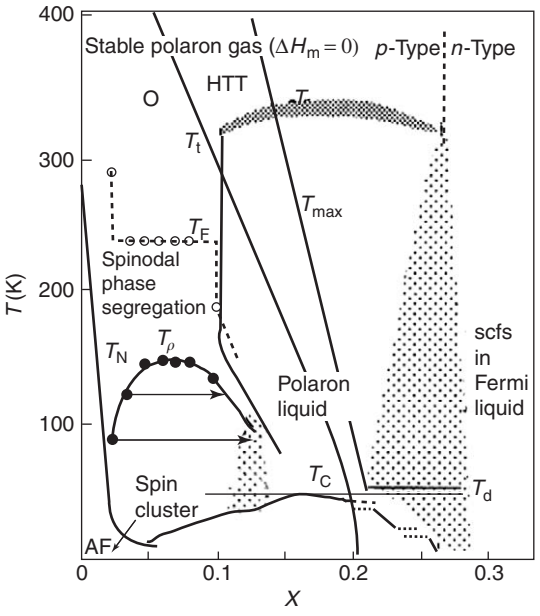


Figure 14 Phase diagram for the superconductive system $La_{2-x}Sr_xCuO_4$. (After Goodenough JB and Zhou J-S (2001) *Structure & Bonding* 98: 95.)

pressure to $PrNiO_3$ lowers $T_{MI} = T_N$ until it terminates at a first-order transition to a quantum-critical-point phase.

Pinning of a redox couple at the top of the O–2*p* bands may also be revealed by oxidation of the transition-metal array. For example, the $x = 0$ composition of the layered $Li_{1-x}CoO_2$ cathode material contains low-spin Co(III):t^6e^0 and is a charge-transfer-gap semiconductor. However, the t^5/t^4 Co(IV)/Co(III) redox couple is pinned at the top of the O–2*p* bands, and removal of Li from between adjacent CoO_2 layers of edge-shared octahedra creates a two-phase region $0.1 < x < 0.4$ in which a semiconductive phase and a metallic phase of smaller Co–Co separation coexist as a result of a $\tau_h \approx \omega_R^{-1}$. At $x > 0.6$, peroxide formation at the surface leads to the loss of oxygen as gaseous O_2.

The high-temperature copper-oxide superconductors provide an example where strong electron–lattice interactions in a mixed-valent system at the $\tau_h \approx \omega_R^{-1}$ crossover leads to a remarkable physical phenomenon. In these layered oxides, the $x^2 - y^2$ orbital of the Cu(III)/Cu(II) couple is pinned at the top of the O–2*p* bands of a CuO_2 sheet. The $(CuO_2)^{2-}$ planes of the single-valent parent compounds are antiferromagnetic insulators, but removal of electrons from these sheets induces a transition from the antiferromagnetic phase to a metallic phase that is not superconductive. Superconductivity appears in an intermediate $\tau_h \approx \omega_R^{-1}$ crossover phase where there is a strong electron coupling to bond-length fluctuations that order into a traveling CDW (or stripes) phase at low temperatures as is illustrated in **Figure 14** for the $La_{2-x}Sr_xCuO_4$ system.

See also: Crystal Field Splitting; Ferromagnetism; Jahn–Teller Effect; Magnetic Materials and Applications.

PACS: 72.80.Ga; 71.20.Be; 71.70.Ch

Further Reading

Bertaut EF (1966) Spin configuration of ionic structures. In: Rado GT and Suhl H (eds.) *Magnetism: A Treatise on Modern Theory and Materials,* vol. III, ch. 4, pp. 149–209. New York: Academic Press.

Gleitzer C and Goodenough JB (1985) Mixed-valence iron oxides. In: Clarke MJ, *et al* (eds.) *Structure & Bonding*, vol. 98, ch. 1, pp. 1–76. Berlin: Springer.

Goodenough JB (1963) *Magnetism and the Chemical Bond*. New York: Wiley and Interscience.

Goodenough JB (1966) Magnetism and crystal structure in nonmetals. In: Rado GT and Suhl H (eds.) *Magnetism: A Treatise on Modern Theory and Materials,* vol. III, ch. 1, pp. 1–62. New York: Academic Press.

Goodenough JB (1972) Theory of antiferromagnetism and ferrimagnetism. In: Hench LL and Dove DB (eds) *Physics of Electronic Ceramics*, ch. 24, pp. 777–831. New York: Dekker.

Goodenough JB (1974) Some comparisons of fluorides, oxides, and sulfides containing divalent transition metal atoms. In: Rao CNR (ed.) *Soild State Chemistry*, ch. 4, pp. 215–364. New York: Dekker.

Goodenough JB (ed.) (2001) *Structure and Bonding*, vol. 98. Berlin: Springer.

Hench LL and Dove DB (eds.) (1972) *Physics of Electronic Ceramics*. New York: Dekker.

Jona F and Shirane G (1962) *Ferroelectric Crystals*. Oxford: Pergamon.

Kanamori J (1966) Anisotropy and magnetostriction of ferromagnetic and antiferromagnetic materials. In: Rado GT and Suhl H (eds.) *Magnetism: A Treatise on Modern Theory and Materials*, vol. I, ch. 4, pp. 127–199. New York: Academic Press.

Lax B and Button KJ (1962) *Microwave Ferrites and Ferrimagnets*. New York: McGraw-Hill.

Morrish AH (1965) *The Physical Properties of Magnetism*. New York: Wiley.

Mott NF (1974) *Metal-Insulator Transitions*. London: Taylor & Francis.

Rado GT and Suhl H (eds.) (1966) *Magnetism: A Treatise on Modern Theory and Materials*, vols. I–IV. New York : Academic Press.

Rao CNR (ed.) (1974) *Solid State Chemistry*. New York: Dekker.

Rao CNR and Rao KJ (1978) *Phase Transitions in Solids*. New York: McGraw-Hill.

Reiss H (ed.) (1972) *Progress in Solid State Chemistry*, vol. 5. Oxford: Pergamon.

Smith J and Wijn HPJ (1959) *Ferrites*. New York: Wiley.

Transmission Electron Microscopy

S J Pennycook, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Published by Elsevier Ltd.

Introduction

The distinguishing feature of transmission electron microscopy (TEM) is its ability to form images of atomic arrangements at localized regions within materials. It provides a view of the microstructure, that is, the variations in structure from one region to another, and the interfaces between them. TEM plays a critical role whenever macroscopic properties are controlled or influenced by defects or interfaces, for example, in the development of advanced structural materials with their complex microstructure of second phases or electronic materials which rely on the exquisite control of interfaces and multilayers. X-ray or neutron diffraction provide quite complementary information. These techniques can determine the average structure of complex materials very precisely, but not the structure of a local region or individual nanostructure. As condensed matter physics moves toward the study of ever more complex materials, and at the same time interest in nanoscale physics and devices is increasing, TEM or its scanning counterpart STEM, is finding a rapidly increasing role in basic condensed matter physics research.

The unique role of the TEM arises because electrons are charged particles, and therefore, unlike X-rays or neutrons, are able to be accelerated and precisely focused by electromagnetic fields. The scattered beams can be collected by a lens, and refocused to form a true real space image in the manner of an optical microscope, where each point in the image corresponds to a specific point in the object. Electrons also interact much more strongly with matter and electron diffraction can be performed on materials of nanometer dimensions. For accelerating voltages of 100–1000 kV, the electron wavelength ranges from 0.004 to 0.001 nm, orders of magnitude lower than atomic spacings in materials. It would therefore appear relatively trivial to form atomic resolution images of materials. However, it is only very recently that atomic resolution imaging could be said to be at all trivial. The major limitation was realized early in the history of the microscope to be the high inherent aberrations of a round magnetic lens. Spherical aberration is the dominant aberration, and leads to a ray deviation of $C_s\alpha^3$, where α is the semiangle of the objective lens. With electron lenses, C_s is of the same order as the focal length. This means that only very small apertures could be used, and since microscope resolution is given by $0.61\lambda/\sin\alpha$, where λ is wavelength, the best resolution would be limited by diffraction to $\sim 50\lambda$.

It took ~ 50 years from the development of the first transmission electron microscopes in the 1930s to achieve sufficient electrical and mechanical stability to allow imaging of crystal lattices at 0.2–0.3 nm resolution in the 1980s. Over the next 20 years, resolution improved incrementally to 0.1–0.2 nm. Now electron microscopy is in the midst of a revolution. Thanks to the development of solid-state devices, specifically the computer and charge-coupled-device (CCD) detectors, a series of nonround magnetic lenses can be used to correct the aberrations of the (round) objective lens. The gain in resolution over the last few years is comparable to that seen in the last few decades, an extraordinary advance that has pushed TEM into the sub-Ångstrom regime for the first time in history.

Aberration-correction brings more than just resolution; better resolution brings increased sensitivity, and recent results have demonstrated the imaging of single atoms within materials and on their surfaces, together with their spectroscopic identification by electron energy loss spectroscopy (EELS). Furthermore, entirely new modes of microscopy now appear feasible. Aberration-correction is allowing the objective aperture to be opened up, and, just as in an optical microscope, the depth of focus reduces. We are seeing the beginnings of three-dimensional (3D) TEM.

Electron Optics of the Microscope: STEM Versus TEM, Reciprocity

Figure 1a shows a schematic ray diagram for conventional TEM. The specimen is illuminated through a condenser aperture and transmitted electrons are gathered by the objective lens and recombined into an image. Historically, the size of the condenser aperture has always been chosen to be as small as practically possible to provide a close approximation to parallel illumination. This is required for diffraction contrast imaging, when only one diffracted beam is passed through the objective aperture. With parallel illumination crystal defects such as dislocations, precipitates, and interfaces are observable, and many details of their atomic structure can be obtained by detailed analysis of image contrast (e.g., lattice shifts across an interface, the

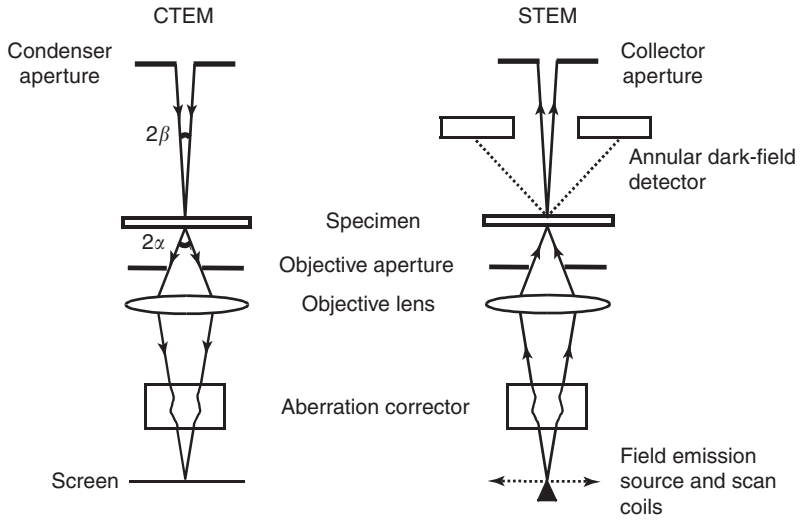


Figure 1 Ray diagrams for (a) the TEM and (b) the STEM, showing the reciprocal nature of the optical pathways. The TEM image is obtained in parallel, the STEM image pixel by pixel by scanning the probe. The STEM also provides simultaneous annular dark-field (ADF) imaging. Actual microscopes have several additional lenses and the beam limiting aperture positions may differ.

intrinsic or extrinsic nature of stacking faults, coherency strains of precipitates, Burgers' vectors of dislocations). Nearly parallel illumination provides a long coherence length in the specimen, approximately λ/β, where β is the illumination semiangle. For small β many lattice spacings are illuminated in phase, an example of coherent imaging. If the objective aperture α is opened up to include several diffracted beams, the conventional TEM lattice-imaging mode is obtained, often known as high-resolution electron microscopy (HREM), another example of coherent imaging.

In scanning transmission electron microscopy (STEM) the electron beam is focused to a small probe and scanned across the specimen. **Figure 1b** shows the ray diagram; detectors are used to pick up a signal, which is used to form an image pixel by pixel as the probe is scanned. The STEM allows complete flexibility in choice of detector, and multiple detectors are normally used giving complementary information. An annular dark field (ADF) detector collects electrons scattered out of the beam and gives an image of mass thickness, that is, a signal that increases with increasing scattering cross section and specimen thickness; a hole appears dark. A bright-field signal can also be detected using an on-axis detector (holes appear bright). If the bright-field detector subtends a small angle β at the specimen, then it is clear from comparing the two ray diagrams in **Figure 1** that the only difference in the optics of the two forms of microscope is the direction of the electron beam. The major source of image contrast is elastic scattering, for which only the scattering angle is important not the direction of propagation (an example of time-reversal symmetry). This illustrates the principle of reciprocity, where interchange of source and detector leads to the same image, and explains why images formed in bright-field STEM may show identical contrast behavior to those in TEM. If the bright-field collector aperture is small, these images are coherent images exactly equivalent to bright-field TEM images.

Reciprocity applies even if aberration correctors are added to both columns. In TEM the corrector is placed after the specimen to correct the objective lens aberrations, while in STEM it is placed before the specimen, the purpose being again to correct the aberrations of the objective lens (now used as a probe-forming lens). **Figure 1** makes clear that for the same set of aberration parameters and equivalent angles before and after the specimen, the same images are to be expected, and this is seen experimentally. However, reciprocity does not mean the images will have the same signal-to-noise ratio. Before aberration-correction, the angle β necessary to ensure good coherent imaging conditions was so small that the STEM image used only a small fraction of the electrons passing through the specimen. This resulted in noisy images and unnecessary specimen irradiation. With aberration-correction, not only can the objective aperture α be made much larger, but β can also be increased and STEM bright-field imaging becomes a useful practical possibility.

The primary motivation for the STEM has been its ability to detect signals other than the bright-field image, in particular, the ADF image. Based on the concept of a coherence width in the specimen, if the angle β is sufficiently large, the coherence width in

the specimen is below the typical atomic spacings in materials. Although the third dimension along the beam direction must also be considered, the ADF image does represent a good approximation to incoherent imaging. An incoherent image gives a simple relationship between the object and image as in a camera, and is easier to interpret than a coherent image. The very different characteristics of coherent and incoherent imaging will be discussed in the next section.

Finally, one of the major uses of STEM has been for microanalysis, which uses one or more additional signals generated as the beam passes through the specimen. Perhaps the most important of these is inelastic scattering, which allows elemental identification from inner-shell excitations. The fine structure at the absorption edges carries information about the electronic structure of the excited atom. For these purposes, EELS is comparable to X-ray absorption spectroscopy, but is available with the spatial resolution of the probe. Spectroscopy of individual atomic columns in a crystal has been demonstrated, including, recently, the spectroscopic identification of a single impurity atom within one atomic column. In the so-called spectrum-imaging mode, one complete spectrum is recorded for each pixel in the image. Similar information can be obtained on a TEM using an imaging filter, which produces 2D images corresponding to particular energy losses. A set of such images at different energy losses is somewhat equivalent to a spectrum image on the STEM. However, reciprocity does not apply to inelastic scattering and there are important differences between TEM and STEM: in the TEM the inelastically scattered electrons must be focused by the image-forming lens which leads to problems with chromatic aberration, whereas in the STEM there are no lenses after the specimen. For this reason atomic column resolution has yet to be achieved in the TEM.

Energy dispersive X-ray spectroscopy, known as EDX or EDS, collects X-rays emitted as the atom relaxes into its ground state following inner-shell excitation, and the X-ray energy is again characteristic of the element. The peak resolution is much lower than for EELS, but the peaks lie on a lower background. The major problem is the low efficiency of collection. X-rays are emitted over 4π steradians and only a small fraction can be collected by the detector, whereas the corresponding core loss electrons are forward peaked and a high fraction can be collected for EELS. This is probably the reason that atomic-resolution EDX analysis has yet to be demonstrated. Other possible imaging signals include the secondary electrons for imaging surfaces and visible photons (cathodoluminescence) for mapping optical properties.

Finally, microdiffraction, or more recently nanodiffraction patterns can be recorded with an imaging detector in STEM. This is very useful for identifying minority phases within bulk materials or new phases not possible to synthesize in large volumes, and also for quantitative measurements of charge modulation. For long wavelength modulations, electron diffraction is much more sensitive than X-ray diffraction because of the cancellation of the scattering due to the electrons and nucleus; electron scattering sees the imbalance of charge, whereas X-ray scattering sees the total electron density and is less sensitive to small changes.

For much of the history of electron microscopy it has not been possible to have high-resolution STEM and TEM imaging in a single machine. In recent years however, great advances have been made in instrument design, and it is now possible to purchase microscopes that will give good performance in any desired mode of operation, often simply at the push of a button. It is possible today to purchase an instrument with two aberration-correctors, one before the specimen to correct the probe and one after the specimen to provide parallel-detection phase contrast imaging.

Principles of Image Formation

Coherent Imaging

With parallel illumination, a crystal aligned near a zone axis will generate a set of diffracted beams leaving the sample at specific angles, as depicted in **Figure 2a**. The objective lens brings these beams to a focus in its back focal plane to form the diffraction pattern. In diffraction contrast imaging, one diffracted beam is passed through a small objective aperture and the resulting image represents a map showing the distribution of diffracted intensity from the specimen. The crystal lattice is unresolved, but the image is useful for studying defects and interfaces. To resolve the crystal lattice itself the objective aperture is opened up to include more than one diffracted beam. Now each point in the image receives contributions from multiple diffracted beams that interfere to generate image contrast, which depends on their relative phases. The principles of coherent phase contrast lattice imaging are best illustrated for the idealized case of a thin sample where multiple scattering can be ignored. The incident electron wave function takes the form of a plane wave ψ_0 which interacts with the electrostatic potential of the specimen $\phi(r)$ (ignoring magnetic fields for now). Here, lowercase letters refer to 3D variables while upper case variables refer to the two transverse dimensions

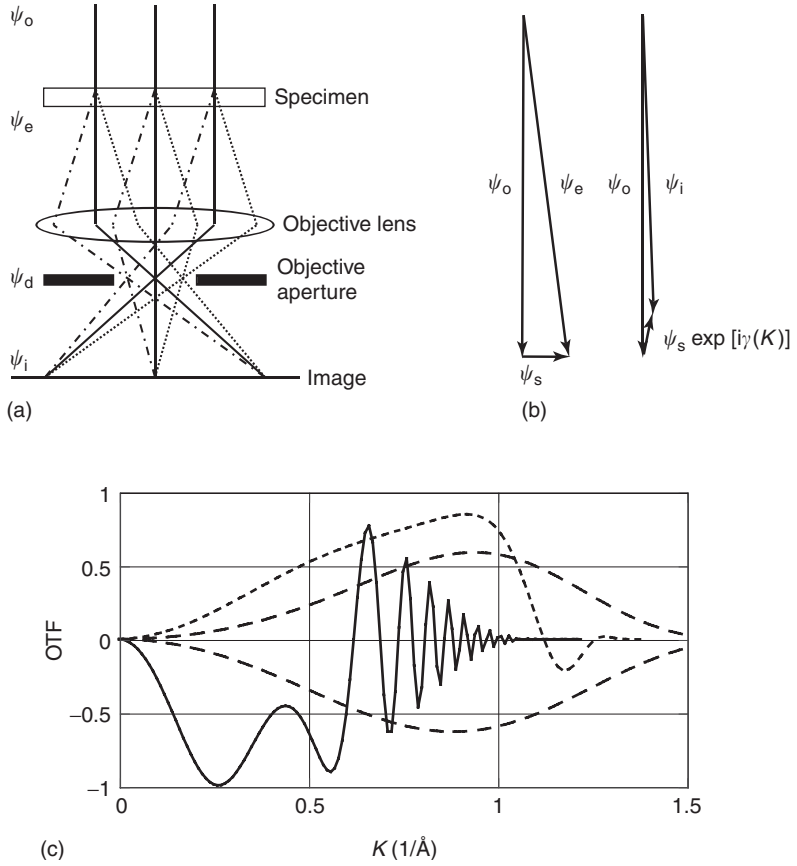


Figure 2 (a) Ray diagram for coherent phase contrast imaging of a crystal. Diffracted beams emerge from the specimen at specific angles and are brought to a focus in the back focal plane of the objective lens. They are recombined in the image plane. Imaging is therefore two consecutive Fourier transforms. (b) Vector diagram for a weak phase object showing the incident beam, a small scattered beam with $\pi/2$ phase change and the resultant transmitted beam. Lens aberrations are used to rotate the phase of the scattered beam an additional $\pi/2$ to create amplitude contrast from the phase changes. (c) Contrast transfer functions for an uncorrected 300 kV microscope, solid line ($C_s = 0.6$ mm, $\Delta f = -45$ nm, energy spread 2 nm typical for a Schottky field emission source), a corrected 300 kV microscope, dotted line ($C_s = -37\,\mu$m. $C_5 = 100$ mm, $\Delta f = 5$ nm, energy spread 1 nm typical for a cold field emission source), and a 200 kV microscope with full correction of all 5th-order aberrations, dashed line ($C_s = 23\,\mu$m. $C_5 = -2$ mm, $C_7 = 100$ mm, energy spread 1 nm). The upper trace is for $\Delta f = 1$ nm giving white atom contrast in a thin specimen, the lower one is for $\Delta f = -2$ nm giving black atoms.

normal to the direction of beam propagation, z; for example the position vector $\boldsymbol{r} = (R, z)$ and the electron wave vector $\boldsymbol{k} = (K, k)$. Assuming an incident beam of unit amplitude, a thin specimen acts as a weak phase grating, refracting the incident beam but not changing its amplitude: the exit face wave function is given by

$$\psi_e(\boldsymbol{R}) = \exp\{-i\sigma\phi(\boldsymbol{R})\} \approx 1 - i\sigma\phi(\boldsymbol{R})$$

where $\sigma = 2\pi me\lambda/h^2$ is the interaction constant and $\phi(\boldsymbol{R}) = \int \phi(r)\,dz$ is the projected potential. Phase changes in the exit wave therefore map the projected potential of the specimen. These phase changes can be represented vectorially by generation of a scattered wave $\psi_s(\boldsymbol{R}) = -i\sigma\phi(\boldsymbol{R})$ which is oriented at $\pi/2$ to the incident wave, as shown in **Figure 2b**. Entering the objective lens, the diffracted amplitude ψ_d is just the Fourier transform of the exit face wave function

$$\psi_d(\boldsymbol{K}) = \delta(0) - i\sigma\phi(\boldsymbol{K})$$

where the delta function at $\boldsymbol{K} = 0$ is just the unscattered beam. For a crystal aligned to a zone axis the diffraction pattern consists of a set of sharp diffracted beams at specific directions ($\boldsymbol{K}$ vectors). With a pure phase object, none of the individual diffracted beam intensities shows an image. To form an image of ψ_s, the diffracted beams must be passed through the objective lens and interfered with the unscattered beam. As shown in **Figure 2b**, the phase of the scattered beams should ideally be rotated an additional $\pi/2$ to convert the phase changes to perfect amplitude contrast. In this ideal case the image amplitude would become $\psi_i(\boldsymbol{R}) = 1 \pm \sigma\phi(\boldsymbol{R})$ and the intensity $I(\boldsymbol{R}) = |\psi_i(\boldsymbol{R})|^2 = 1 \pm 2\sigma\phi(\boldsymbol{R})$, depending on the

sense of rotation of the scattered wave with respect to the unscattered wave. In practice, since there are no phase plates for electrons, the only available source of additional phase changes are lens aberrations and defocus. The phase change introduced by the objective lens is referred to as the transfer function phase factor $\exp[i\gamma(\boldsymbol{K})]$ where

$$\gamma(K) = \pi\left(\Delta f \lambda K^2 + \tfrac{1}{2} C_3 \lambda^3 K^4 + \tfrac{1}{3} C_5 \lambda^5 K^6\right)$$

if only the rotationally symmetric aberrations are included so that $K = |\boldsymbol{K}|$. Here Δf is defocus, C_3 is the third-order spherical aberration coefficient, usually referred to simply as C_s before aberration-correction became viable, with typical values of the order of 1 mm. Now that C_s can be corrected, the fifth-order spherical aberration coefficient C_5 takes over as the limiting aberration. The diffraction amplitude transmitted by the lens becomes

$$\psi_d(\boldsymbol{K}) = \delta(0) - i\sigma\phi(\boldsymbol{K}) \exp[i\gamma(\boldsymbol{K})]$$

and for small phase changes the intensity becomes

$$I(\boldsymbol{R}) = 1 - 2\sigma\phi(\boldsymbol{R}) \otimes f(\sin \gamma(\boldsymbol{K}))$$

The phase changes $\phi(\boldsymbol{K})$ contribute to the image with a contrast given by $\sin \gamma(\boldsymbol{K})$, which is therefore referred to as the phase contrast transfer function, and the image is given by a convolution of the projected potential with the Fourier transform of $\sin \gamma(\boldsymbol{K})$.

The different aberrations depend on spatial frequencies to different powers and can therefore only be balanced over a finite range. **Figure 2c** shows examples for some uncorrected and aberration-corrected microscopes. Note that phase contrast imaging provides no contrast at zero spatial frequency because there is negligible additional phase change from lens aberrations. The uncorrected microscope shows rapid contrast oscillations at high spatial frequencies where the $C_s K^4$ term increases very rapidly; such oscillations are largely avoided in the corrected microscopes. The substantial drop in contrast at high spatial frequencies is due primarily to the finite energy spread of the beam. In this regard, the cold field emission gun is preferable to the Schottky thermal field emission gun. One characteristic of phase contrast is immediately apparent. The defocus phase factor is used to balance the dominant term in spherical aberration, but changes sign either side of Gaussian focus. In many cases, it is possible to reverse the image contrast on changing the focus (see **Figure 2c**) and atoms may look black or white in a phase contrast image. This can also happen if the specimen becomes thicker and additional relative phase changes occur between the diffracted beams due to dynamical diffraction (multiple scattering). The exit face wave function no longer relates simply to projected potential and becomes increasingly nonlocal. The general expression for the image intensity is

$$I(\boldsymbol{R}) = |\psi_e(\boldsymbol{R}) \otimes P(\boldsymbol{R})|^2$$

the square of a convolution of the exit face wave function with the point response function of the objective lens. The point response function is just the Fourier transform of the transfer function phase factor,

$$P(\boldsymbol{R}) = \int A(\boldsymbol{K}) e^{2\pi i \boldsymbol{K}\cdot\boldsymbol{R}} e^{i\gamma(K)}\, d\boldsymbol{K}$$

where $A(\boldsymbol{K}) = 1$ over the range of the objective aperture. Because the image is given by a square of the convolution, sum and difference spatial frequencies can appear, and for these reasons image simulations are normally required to relate a phase contrast image to an object.

Incoherent Imaging

Incoherent imaging provides a simple relationship between the object and image

$$I(\boldsymbol{R}) = O(\boldsymbol{R}) \otimes |P(\boldsymbol{R})|^2$$

where the object $O(\boldsymbol{R})$ is blurred by the resolution function of the imaging device. Strictly, this applies only when each point on the object emits independently of its neighbors so that there are no permanent phase relationships between them and no persistent interference can occur. The coherence length of the Sun on Earth is ~ 0.02 mm, so that observation on a larger length scale is effectively incoherent. This is the kind of imaging familiar from the camera, for example, where changing focus merely blurs the image but does not invert the contrast. It also applies to optical microscopy if a large aperture condenser lens is used to give effectively incoherent illumination, as first analyzed by Lord Rayleigh in 1895.

In the STEM, $|P(\boldsymbol{R})|^2$ is the intensity profile of the probe (see **Figure 1b**). Not every image takes the form of a simple convolution, as seen already for the case of the small axial detector. The requirement for an incoherent image is that all of the scattering (or other generated signal) be collected by the detector. In the case of inelastic scattering or X-ray emission, it is relatively simple to collect all of the signal, or at least a representative fraction of it. In the case of

elastically scattered electrons it is not so obvious, because collecting all of the transmitted electrons will give no image contrast; the specimen does not absorb any electrons, it only scatters them, and partial detection of the scattering is therefore necessary to generate an image.

The reason that collecting all of the scattering results in an incoherent image is that the detected signal is then insensitive to any interference that may be occurring amongst different scattered beams. Interference rearranges intensity but the total scattering is constant, by conservation of energy. The annular detector provides a close approximation to incoherent imaging because it can be arranged to average over a large number of diffracted beams. The inner detector angle should also be large to avoid being dominated by a few low-order beams around the inner hole. In practice, the image still shows strong incoherent characteristics even if the inner angle is just larger than the objective aperture angle, which is a mode that gives highly efficient dark-field imaging of single atoms and was the original motivation for the introduction of the detector by Crewe and co-workers.

High-angle scattering comes from the sharpest part of the atomic potential, the nucleus, and its intensity is proportional to the square of atomic number (Z). An annular detector with a large hole therefore gives an image showing strong Z-contrast, also referred to as a high-angle annular dark-field (HAADF) image. With a sufficiently small probe it is possible to achieve an atomic-resolution Z-contrast image in which atomic columns are seen bright (consistent with it being a dark-field image) and their relative contrast is approximately given by their projected mean square atomic number. The atomic-resolution contrast again comes from the interference of diffracted beams, except that no phase plates are necessary to create the intensity. The total intensity on the detector rises and falls as the probe scans the lattice because of interference between the diffracted beams reaching the detector, as shown in **Figure 3a**. Each beam takes the form of a disk the size of the objective aperture (the incident beam in STEM). Interference occurs wherever the disks overlap and there are two or more different pathways that can interfere. The conditions for optimum defocus are that the entire area of overlap shows the strongest possible interference as the probe scans. This is quite different from the conditions for optimum phase contrast with a (small) bright-field detector, and for the example depicted, the bright-field detector covers no overlaps and will show no image.

Smaller spacings will give diffraction disks further apart with smaller regions of overlap and lower

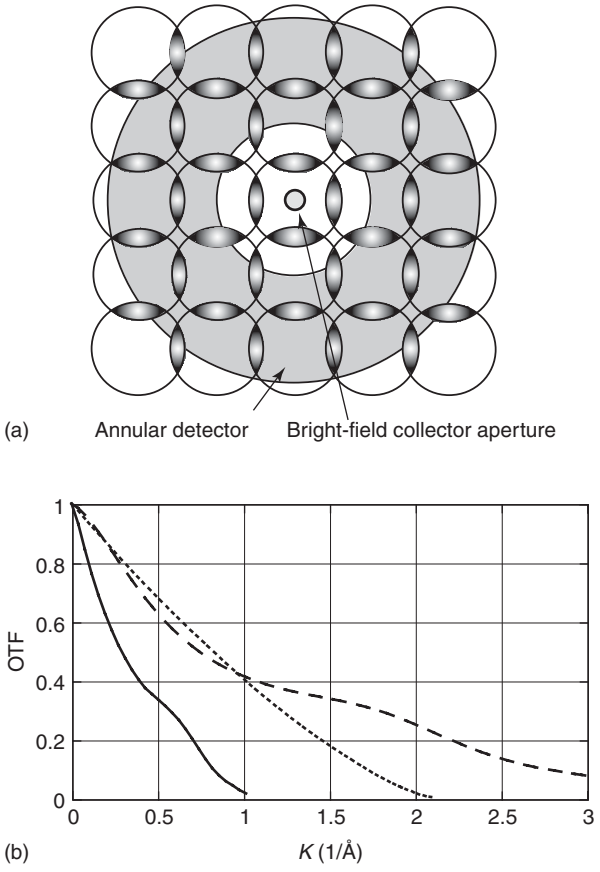


Figure 3 (a) Schematic showing overlapping diffraction disks on the STEM annular detector. The regions of overlap show interference as the probe scans leading to an atomic resolution image. Note, for the spacings depicted, the bright-field detector covers no overlaps and shows no image. (b) Incoherent object transfer functions for an uncorrected 300 kV STEM: solid line (conditions correspond to those in **Figure 2** except $\Delta f = -34$ nm), a corrected 300 kV STEM; dotted line (as **Figure 2** except $\Delta f = 2$ nm), a 200 kV STEM with correction up to 5th order; dashed line (as **Figure 2** except $\Delta f = 0$ nm).

maximum contrast. The incoherent transfer function is therefore a slowly decreasing function of spatial frequency, and is obtained from the Fourier transform of the image intensity,

$$I(\boldsymbol{K}) = O(\boldsymbol{K}) \cdot |P(\boldsymbol{K})|^2$$

The transfer function is just the Fourier transform of the probe intensity, and is referred to as a modulation transfer function, $|P(\boldsymbol{K})|^2$, or object transfer function. Generally, this remains a positive function even as defocus is varied so inverted contrast is not observed. **Figure 3b** shows incoherent transfer functions corresponding to the coherent contrast transfer functions of **Figure 2c**.

Figure 4 compares aberration-corrected coherent and incoherent images of $SrTiO_3$ in the $\langle 1\,1\,0 \rangle$

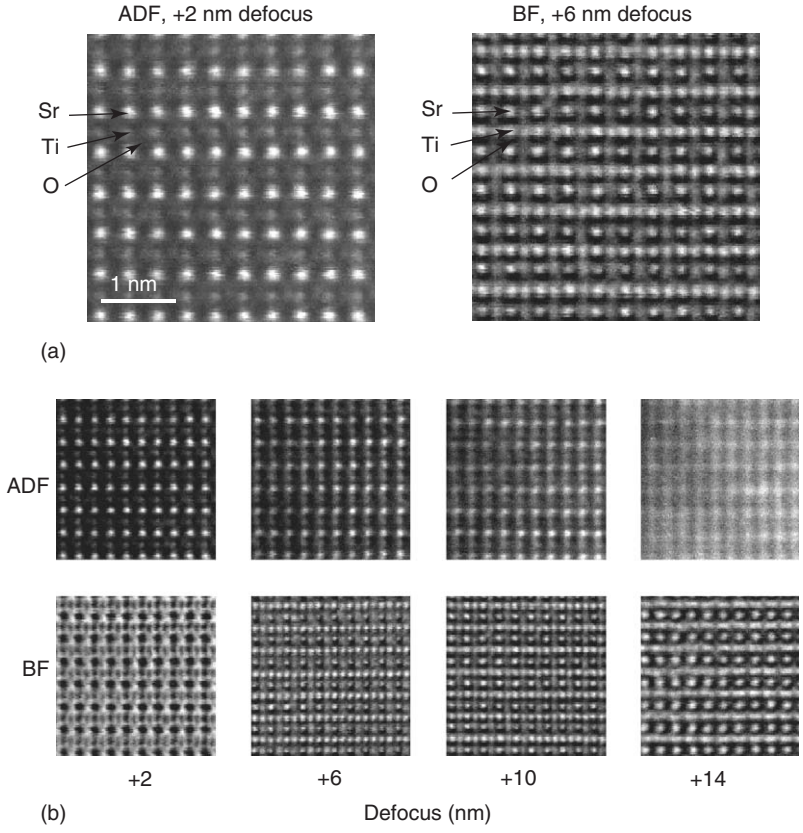


Figure 4 Comparison of incoherent and coherent imaging of $SrTiO_3$ in the $\langle 1\,1\,0 \rangle$ projection in a 300 kV aberration-corrected STEM (a VG Microscope's HB603U with Nion aberration corrector, conditions as in **Figures 2** and **3**). A defocus series was taken using ADF and coherent BF detectors simultaneously. (a) The ADF image shows good *Z*-contrast, with O barely visible. The phase contrast BF image shows the O columns with high contrast; however, the bright features between the Sr columns are double periodicity artifacts. (b) A through focal series with 4 nm defocus steps from the position of optimum ADF defocus showing how the incoherent ADF image contrast slowly blurs while the phase contrast image shows many different forms of contrast. Images are raw data and show some instabilities. (Courtesy of M F Chisholm, A R Lupini, and A Borisevich.)

projection, obtained simultaneously in an aberration-corrected STEM. There is one optimum focus for the *Z*-contrast image, where the Sr column is brightest and the TiO column less bright. A small peak is present at the oxygen column position but is difficult to detect above the noise because of its low *Z*. In the phase contrast image, a good image of the oxygen columns is obtained at a defocus of $+6$ nm, but other images are seen at other defocus values.

The TEM as a Nano-Laboratory

The TEM is a very versatile instrument and is more like a nano-laboratory than an optical instrument, with a large array of imaging and spectroscopic methods available to probe materials at the atomic scale. Imaging modes are available for magnetic fields (Lorentz imaging, Foucault imaging, holography), and movies have been taken of flux line motion in superconductors. Many applications desire measurements under conditions other than the high vacuum, room temperature environment that is standard for TEM; phase transformations can be observed with a heating holder; studies of catalysis benefit from imaging in an active gas environment; many condensed matter physics investigations require low temperatures; nanowires can be observed while being stretched with simultaneous measurement of conductance. *In situ* microscopy also benefits from aberration-correction because it is possible to maintain resolution with a larger gap between the objective lens pole pieces.

Future Directions

The achievement of aberration-correction in TEM is likely to be seen as one of the most significant events in the history of the microscope. Aberration-correction substantially improves vision at the atomic scale, through better resolution and better contrast, for the first time giving useful sensitivity to single atoms within materials and on their surfaces.

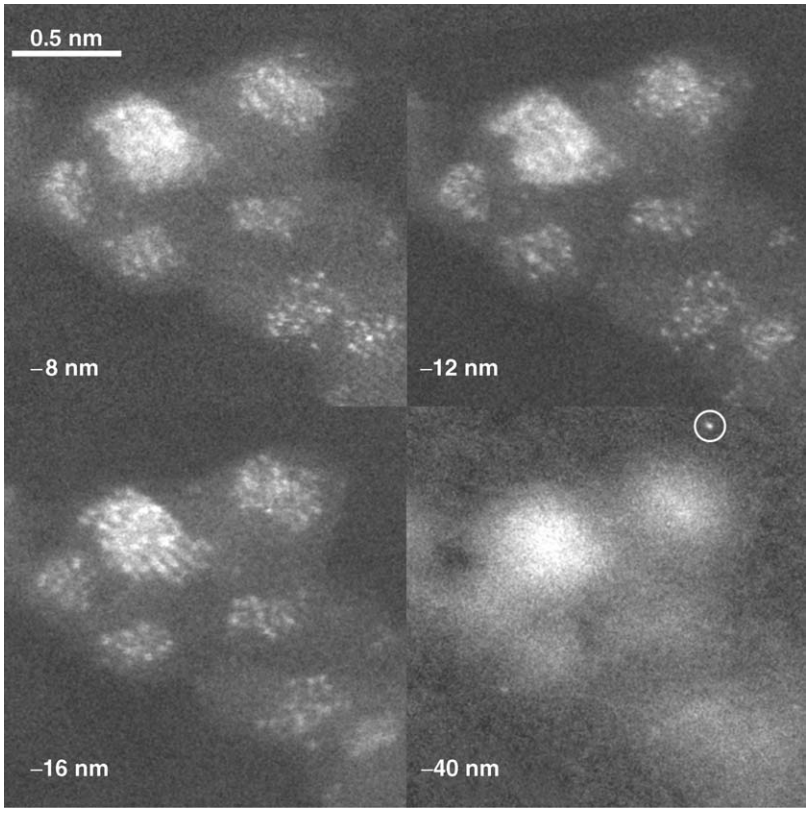


Figure 5 Four frames from a through-depth series of ADF images of a Pt_2Ru_4 catalyst supported on γ-alumina taken using a 300 kV VG Microscope's HB603U with Nion aberration corrector. The alumina is three-dimensional with thin, raft-like Pt–Ru clusters on its surface. Pt atoms are brighter spots and Ru atoms are less bright. Different clusters are resolved at different depths, until, at -40 nm defocus, the carbon support film is reached and a single Pt atom comes into focus (circled). (Image courtesy A Borisevich and A R Lupini.)

Equal improvements apply to spectroscopy. These developments promise a new level of insight into the physics of interfaces, defects, and nanostructures, issues relevant to the majority of materials and devices of a technological world.

A TEM image presents a 2D projection of a 3D specimen. Stereo techniques have been used to obtain information on the third dimension, and recently, full 3D tomography has been demonstrated at nanometer resolution from a series of images taken over a large range of sample tilts. Aberration-correction allows larger objective aperture angles to be utilized, and sub-nanometer depth resolution with sub-Ångstrom lateral resolution is likely to be achieved within the next few years. This allows the possibility of 3D STEM through depth slicing. Just as in optical confocal microscopy, a series of images taken at different focus settings can be reconstructed into a 3D object. **Figure 5** shows four sections of a 3D data set showing catalyst atoms at different depths on a high surface area alumina support.

See also: Confocal Optical Microscopy; Low-Energy Electron Microscopy; Optical Microscopy; Photoelectron Spectromicroscopy; Scanning Near-Field Optical Microscopy; Scanning Probe Microscopy; Scattering, Elastic (General).

PACS: 68.37.Lp; 61.14.−x; 61.46.+w; 61.72.−y

Further Reading

Egerton RF (1996) *Electron Energy-Loss Spectroscopy in the Electron Microscope*, 2nd edn. New York: Plenum.

Kirkland EJ (1998) *Advanced Computing in Electron Microscopy*. New York: Plenum.

Peng L-M, Dudarev SL, and Whelan MJ (2004) *High-Energy Electron Diffraction and Microscopy*. Oxford: Oxford University Press.

Reimer L (2005) *Transmission Electron Microscopy*. Berlin: Springer.

Rickerby DG, Valdré U, and Valdré G (1999) *Z-Contrast Scanning Transmission Electron Microscopy*. Dordrecht: Kluwer.

Spence JCH (2003) *High Resolution Electron Microscopy*. Oxford: Oxford.

Wang ZL (1995) *Elastic and Inelastic Scattering in Electron Diffraction and Imaging*. New York: Plenum.

Williams DB and Carter CB (1996) *Transmission Electron Microscopy*. New York: Plenum.

Zhang X-F and Zhang Z (2001) *Progress in Transmission Electron Microscopy*. Berlin: Springer.

Transport in Two-Dimensional Semiconductors

P T Coleridge, Institute for Microstructural Sciences, Ottawa, ON, Canada

Introduction

Although many characteristics of two-dimensional behavior can be seen in samples where the thickness is less than the mean free path of the carriers, true two-dimensional transport requires a system where the energy associated with motion in one dimension, normally taken as the vertical direction, is quantized. The importance of the semiconductor industry and Moore's law (the observation that the number of transistors in an integrated circuit grows exponentially with time) has meant that much of the experimental information on two-dimensional systems has been obtained using semiconductors. The lateral and vertical dimensions of semiconductor devices have both been systematically shrinking since the 1970s and vertical dimensions are now typically less than 30 nm. This is only a fraction of the de Broglie wavelength, and the electrons are therefore confined by potentials that quantize the motion in the vertical direction. For device modeling, involving room temperature operation, it is often adequate to treat the carriers classically, but at low temperatures they behave as a quantized two-dimensional layer. These systems exhibit a variety of new and interesting phenomena which are studied not only for their fundamental physical significance but also because they involve mechanisms and processes that might be exploited to overcome the fundamental limits that Moore's law is now facing.

One characteristic of two dimensions is that screening of the Coulomb interaction is significantly weaker than that in three dimensions. Surface gates can, therefore, be used very effectively to manipulate electron densities in an underlying two-dimensional semiconductor layer. This property is exploited to make devices of further reduced dimensionality, with lateral dimensions significantly less than 100 nm, and which exhibit one-dimensional and zero-dimensional behavior.

Because the semiconductor material can be grown with very high purity, and because dopants used to introduce carriers can be physically separated from the carriers, the mean free path of the carriers at low temperatures is increased by several orders of magnitude over typical room temperature values. The application of magnetic fields introduces two new length scales, the cyclotron radius and the magnetic length. In these high-purity systems, fields of only a few hundred gauss are sufficient to reach the high-field limit, and the transport properties are then almost totally dominated by magnetic field effects.

Growth of Two-Dimensional Layers

To achieve the strong confinement needed to form a two-dimensional layer requires an artificial growth technique. Molecular beam epitaxy (MBE) is commonly used: a heated, single-crystal substrate is exposed to a flux of elemental atoms in an ultrahigh vacuum (UHV) environment. Under suitable conditions, growth occurs layer by layer at a rate of typically one monolayer per second. The sources are held in heated Knudsen cells (K-cells) with shutters in front of them that can interrupt the flux and can be used to form interfaces with atomic resolution. Because of the UHV environment, there are few impurities and extremely clean layers can be grown.

An example is illustrated in **Figure 1a**: a narrow quantum well of GaAs (typically 10–20 nm) is surrounded by layers of $Ga_{1-x}Al_xAs$ with $x \sim 30\%$. The quantum well is "modulation" doped, that is, the Si donor atoms are separated from the well by undoped spacer layers, so the electrons transfer to the well leaving the remote donors ionized. Motion in the vertical or growth direction (chosen as the z-axis) is quantized with an energy E_0 but two degrees of freedom remain for motion in the x- and y-directions and

$$E(k) = E_0 + \frac{\hbar^2}{2m^*}(k_x^2 + k_y^2)$$

The potential distribution shown is a combination of conduction band offsets produced by changing the alloy concentration and electrostatic potentials produced by the ionized donor atoms and the free electrons. The most commonly used two-dimensional electron gas structures, rather than using a square quantum well, have a single GaAs/(Ga, Al) As interface with doping in the (Ga, Al)As and with the electrons then confined to a triangular shaped well. Many computer programs are available for calculating such band diagrams in terms of self-consistent solutions of the Schrödinger and Poisson equations (see "Further reading" section).

In silicon the metal-oxide-semiconductor-field effect transistor (MOSFET) structure (see **Figure 1b**) has been extensively developed for integrated circuits

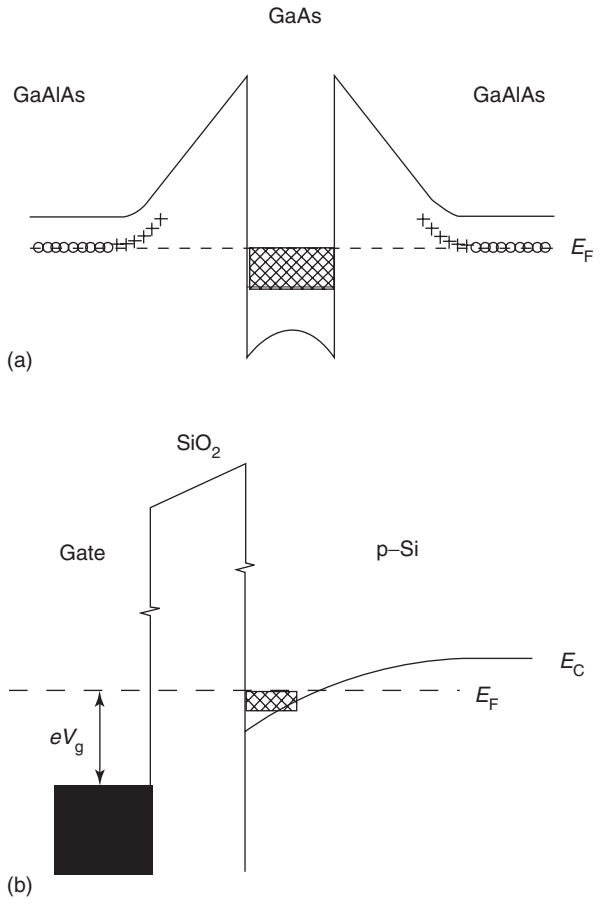


Figure 1 Schematic band diagrams: (a) modulation-doped GaAs quantum well showing ionized (+) and neutral (o) donors with electrons transferred to the quantum well; (b) Si-MOSFET showing an approximately triangular potential well and two-dimensional electron layer formed by applying a voltage V_g to the surface gate.

and also gives rise to a two-dimensional layer. In this case, a high-purity silicon oxide layer is grown on a clean surface of single-crystal silicon and then covered with a metallic gate. Applying a potential between the gate and the bulk silicon (into which ohmic contacts have been implanted) draws in electrons (or holes) and confines them in a triangular well. In this case no doping is needed, the main scattering mechanisms are unwanted defects in the oxide layer and surface roughness.

Basic Transport Properties and Characterization

Two-dimensional semiconductor samples are usually characterized in terms of a density and mobility. This implicitly assumes a simple Drude model, with only one set of carriers. For a magnetic field B, applied in the z-direction normal to the two-dimensional layer, the components of the conductivity tensor describing the response to an in-plane electric field (E_x, E_y) are then

$$\sigma_{xx} = \sigma_{yy} = \frac{\sigma_0}{1 + \mu^2 B^2}$$
$$\sigma_{xy} = -\sigma_{yx} = -\frac{\mu B \sigma_0}{1 + \mu^2 B^2} \qquad [1]$$

where the zero-field conductivity $\sigma_0 = n_s e \mu$ with n_s the sheet carrier density and μ the mobility. The mobility is related to a scattering time τ by $\mu = e\tau/m^*$, where m^* is the effective mass. The cyclotron frequency ω_c, describing motion in a magnetic field, is given by eB/m^* so $\mu B = \omega_c \tau$.

Inverting the conductivity tensor gives the resistivities that are usually measured experimentally:

$$\rho_{xx} = \rho_0 = \frac{1}{\sigma_0}, \qquad \rho_{xy} = \frac{B}{n_s e} \qquad [2]$$

The slope of the Hall resistance ρ_{xy} gives the density which can then be used with the longitudinal resistivity ρ_{xx} to determine a mobility. According to this model, ρ_{xx} is independent of the magnetic field but a significant magnetoresistance (and nonlinear increase of ρ_{xy}) is sometimes observed associated with the existence of more than one set of carriers with different mobilities. A multiband extension of the Drude model can be used to analyze such data, with separate densities and mobilities for each set of carriers, but care should be taken in interpreting these results as the model may not be valid when there is significant scattering between the several sets of carriers.

Measurement Geometries

Examples of some of the geometries used experimentally to determine resistivities are illustrated in **Figure 2**. For a Hall bar (**Figure 2a**), the resistivities are related to the measured resistances R_{ij} by $R_{xx} = (L/W)\rho_{xx}$ and $R_{xy} = \rho_{xy}$, where W and L are, respectively, the width of the Hall bar and the spacing between potential contacts. Values of ρ_{xx} are usually expressed in ohms/square to emphasize that a geometrical factor is needed. For ρ_{xy}, the Hall voltage and the current both scale with the width, so it is a particular characteristic of two dimensions that the Hall resistance should be independent of the geometry. In practice, spurious contributions from ρ_{xx} often appear, for example, if the potential probes are misaligned, but these can be eliminated by averaging results obtained with both directions of the magnetic field.

The Van der Pauw geometry (**Figure 2b**) is especially convenient because, provided the contacts are

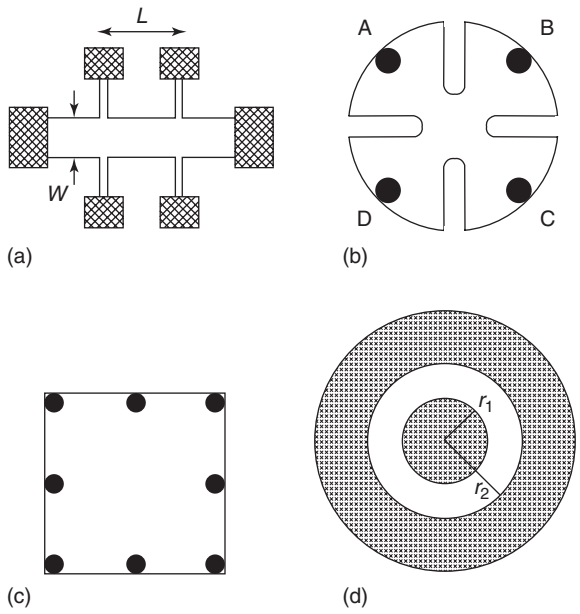


Figure 2 Resistance measuring geometries with metallic contacts shown: (a) Hall bar with source and drain and two pairs of voltage contacts; (b) Van der Pauw geometry designed to minimize errors associated with contact placement; (c) Van der Pauw geometry which can also be used as a pseudo-Hall bar; and (d) Corbino geometry used to measure conductivity rather than resistivity.

small and at the edges of the sample, ρ_{xx} can be obtained without needing to know the spacing between contacts. In this case

$$\rho_{xx} = \frac{\pi}{\ln 2} \frac{R_{\mathrm{AB,CD}} + R_{\mathrm{BC,DA}}}{2} f$$

where $R_{\mathrm{AB,CD}}$ is the resistance measured with the current through contacts A and B and voltage at contacts C and D. The factor f, which depends on the ratio $R_{\mathrm{AB,CD}}/R_{\mathrm{BC,DA}}$, is 1.0 when the ratio is 1, 0.96 with a ratio of 2, and 0.82 for a ratio of 5. The Van der Pauw method can also be used with multiple contacts on the edge of a square sample (**Figure 2c**). Comparing values obtained with different sets of contacts then provides a convenient check on the reproducibility of the results and this configuration can also be used as a pseudo-Hall bar.

In the Corbino geometry (**Figure 2d**), the circular symmetry means the two-terminal resistance R_{12} depends not on the resistivity but rather on the conductivity according to

$$R_{12} = \frac{1}{2\pi\sigma_{xx}} \ln \frac{r_2}{r_1}$$

where r_1 and r_2 are, respectively, the inner and outer radii between the two contacts. Although the Hall resistance cannot be obtained directly in this measurement, the mobility can be determined from the quadratic increase of R_{12} with the magnetic field (cf. eqn [1]) and the density then deduced from the zero-field conductivity σ_0.

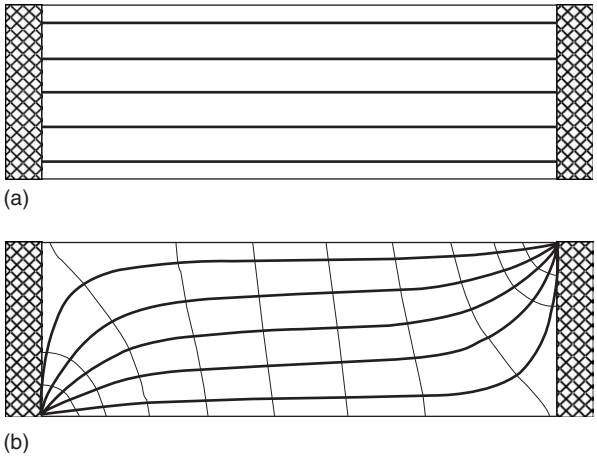


Figure 3 Current and electric field lines in a Hall bar: (a) in zero magnetic field when ***E*** and ***J*** both run parallel to the edges and (b) in a magnetic field with $\rho_{xy}/\rho_{xx} = 10$. The electric field lines meet the contacts normal to the contact edges, but because the current lines (bold) must be approximately normal to the electric field lines, they enter and exit at the corners of the contacts.

At low temperatures, the mobility of many two-dimensional systems can be very large – for example, in modulation-doped GaAs/GaAlAs heterojunctions, values of over $100\,\mathrm{m}^2(\mathrm{Vs})^{-1}$ are common. Quite modest magnetic fields are then sufficient to reach the high-field regime $\mu B \gg 1$ where the transport is dominated by the Hall resistance. In zero magnetic field, the current and electric field lines are parallel (as is illustrated in **Figure 3a**) with the electric field normal to the edges of the metallic contact pads. Inside the semiconductor, in the high-field regime, $\rho_{xy} \gg \rho_{xx}$ and the current and field lines become approximately orthogonal. At the contacts, the electric field must remain normal to the metallic edge, so the current must flow parallel to the edge of the contact and can only enter and exit at the corners of the Hall bar (see **Figure 3b**). Reversing the magnetic field moves these critical points to opposite corners and means the high current density now flows in different regions. This can be important if the semiconductor is inhomogeneous.

This behavior emphasizes, when investigating samples where $\rho_{xy} \gg \rho_{xx}$, the need to ensure that potential probes are well separated from the current contacts if reliable values of ρ_{xx} are to be obtained. It also means that the two-terminal resistance of a Hall bar is essentially just $\rho_{xy} + (L/W)\rho_{xx}$ which is dominated, at high fields, by the Hall resistance.

In the high-field regime, transport becomes very sensitive to inhomogeneities in the semiconductor. Density variations affect mainly the Hall voltage and cause currents to tend to follow the contours of constant density. If very different voltages are measured at contacts that are nominally equivalent, this usually indicates inhomogeneities. In devices where large-density gradients are intentionally introduced, for example with gates, this type of behavior dominates. Transport properties are then frequently analyzed using the Landauer–Buttiker formalism, in terms of one-dimensional current paths running along regions of high-density gradients. Under quantum Hall conditions, when Hall resistances are quantized, the transmission coefficients associated with these current paths are also quantized.

Quantum Mechanical Treatment of Magnetic Field

Semiclassically, cyclotron motion in a magnetic field occurs with an angular frequency $\omega_c = eB/m^*$ (independent of energy) and a cyclotron radius $R_{cycl} = m^* v/eB$, where v is the velocity. Quantum mechanically, magnetic fields are introduced by replacing $\hbar \boldsymbol{k}$ in the Schrödinger equation with $\hbar \boldsymbol{k} - e\mathbf{A}$ where $\boldsymbol{A}$ is the vector potential. The continuous energy distribution in zero field is replaced by a ladder of discrete Landau levels with energies $(n_L + (1/2))\hbar\omega_c$ where the Landau level index $n_L \geqslant 0$. The wave function depends on the choice of gauge. Although theoretically the Landau gauge is often used, the closest connection with the semiclassical result is made by choosing the symmetric gauge with $\boldsymbol{A} = 0$ at the center of the orbit. In this case, the wave function (without normalization) is given by

$$\psi_{n_L}(r, \phi) = \exp(\mathrm{i} n_L \phi) r^{n_L} \exp(-r^2/4l_m^2)$$

where ϕ is the azimuthal angle and the characteristic length $l_m = (\hbar/eB)^{1/2}$ is called the magnetic length. This can be considered a "race-track" orbit with a spread of l_m and a mean radius $(2n_L)^{1/2} l_m$ that is slightly less than the semiclassical radius of $(2n_L + 1)^{1/2} l_m$. This difference is small but has important consequences.

In general, the orbits need not be circular but rather follow a Fermi contour determined by the band structure of the material. This is particularly the case for holes where multiple bands may exist within a small energy range. In this case (as is discussed in more detail in the article on the de Haas–van Alphen effect in metals), the energy is still quantized into Landau levels but the cyclotron radius and effective mass of the orbit become orbital averages.

Each Landau level is highly degenerate, corresponding to orbits with different centers. The degeneracy per unit area, eB/h, corresponds to one orbit center in an area πl_m^2. In addition to this basic degeneracy, there is also a twofold spin degeneracy and (sometimes, e.g., in Si-MOSFETs) a valley degeneracy. The total degeneracy is usually expressed in terms of a filling factor ν $(= n_s h/eB)$, which relates the number of filled Landau levels in a magnetic field B (including any degeneracies) to the sheet carrier density n_s.

As illustrated in **Figure 4a**, the Landau level degeneracy means the energy of electrons in a filled Landau level is the same as that of the equivalent electrons in zero field. For electrons in a partially filled level, however, there is an excess energy (see **Figure 4b**) which can be identified, thermodynamically, with the magnetization associated with the orbital moment of the cyclotron orbits. Although the

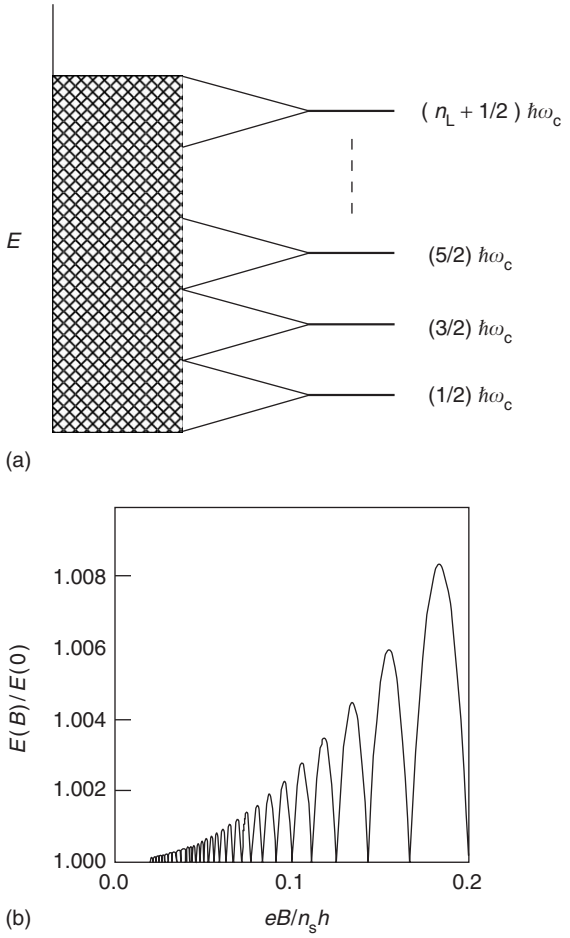


Figure 4 (a) Collapse of uniform, zero-field, energy distribution onto full Landau levels that have the same average energy. (b) Extra energy associated with partially filled Landau levels that gives rise to the Landau diamagnetism.

oscillating part of this magnetization may be washed out by temperature or disorder, the quadratic background term (ΔE_L) persists. The corresponding magnetization, $(-\partial \Delta E_L/\partial B)$, is known as the Landau diamagnetism.

The effect of uniform magnetization M distributed over an area A can be described by an equivalent current $I_s = M/A$ circulating round the edge of the sample. Any gradient of the chemical potential across the sample ($\delta\mu$), such as might be induced by the electric field in a Hall measurement, will induce a corresponding gradient in the magnetization in the bulk of the sample. This implies there is a difference, $\delta I_s = (1/A)(\partial M/\partial \mu)\delta\mu$, between the effective circulating currents on the two sides of the sample. With contacts (see **Figure 5**), $\delta\mu$ can be detected and identified as a potential difference, and δI_s can be accommodated as an extra source/drain current. The ratio $e\delta I_s/\delta\mu$ appears as a Hall conductance which, because of the special geometry associated with two dimensions, is also a Hall conductivity. Using the Maxwell relation $(\partial \boldsymbol{M}/\partial\mu)_{\mathbf{B}} = (\partial N/\partial \boldsymbol{B})_{\mu}$ (where $N = n_s A$), this can be written as $e(\partial n_s/\partial \boldsymbol{B})_{\mu}$, so the total Hall conductivity, including the usual Drude contribution term described by eqn [1], becomes

$$\sigma_{xy} = -\sigma_{yx} = -\omega_c \tau \sigma_{xx} + e(\partial n_s/\partial \boldsymbol{B})_{\mu} \qquad [3]$$

The extra "edge state" or "Streda" contribution is essential for a detailed understanding of the Hall effect in two-dimensional systems. In some instances, notably in the integer quantum Hall regime, σ_{xx} is actually zero in the bulk and current is carried only at the edges. More generally, "edge" currents and "bulk" currents both exist.

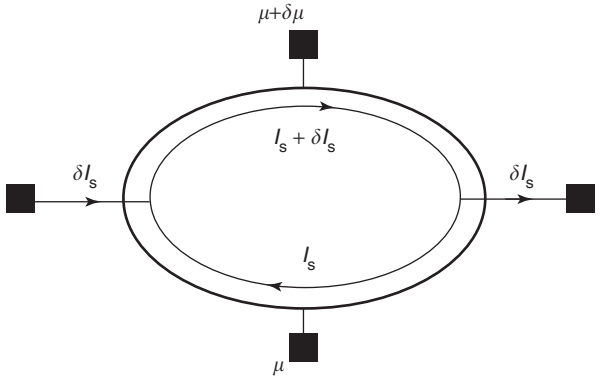


Figure 5 Schematic illustration of the origin of the "edge state" Hall conductance that results from gradients in the magnetization induced by gradients in the chemical potential. (Adapted from MacDonald AH (1994) Les-Houches lecture proceedings, with permission.)

Disorder Broadening of the Landau Levels

Although the density of states illustrated in **Figure 4a** shows infinitely sharp Landau levels, they are, in practice, broadened by disorder. At low fields and with dirty samples (see **Figure 6a**), this disorder broadening is larger than the spacing between the centers of the Landau levels, and the Landau level structure is only a minor perturbation on the constant, background, density of states. In clean systems, at higher fields, the Landau level spacing eventually becomes larger than the broadening and the levels then separate (**Figure 6b**). As noted above, each Landau level is twofold spin degenerate, and at high enough fields the Zeeman splitting associated with the spins can become large enough such that each Landau level splits into two spin-resolved levels (**Figure 6c**).

With Lorentzian broadening, the density of states associated with the ladder of Landau levels is given by the Fourier expansion

$$g(E) = g_0\left[1 + 2\sum_{s=1}^{\infty} \exp\left(\frac{-\pi s}{\omega_c \tau_q}\right) \times \cos\left(\frac{2\pi s E}{\hbar\omega_c} - s\pi\right)\right] \qquad [4]$$

where the Landau level width (Γ_L) has been expressed in terms of a single particle or quantum

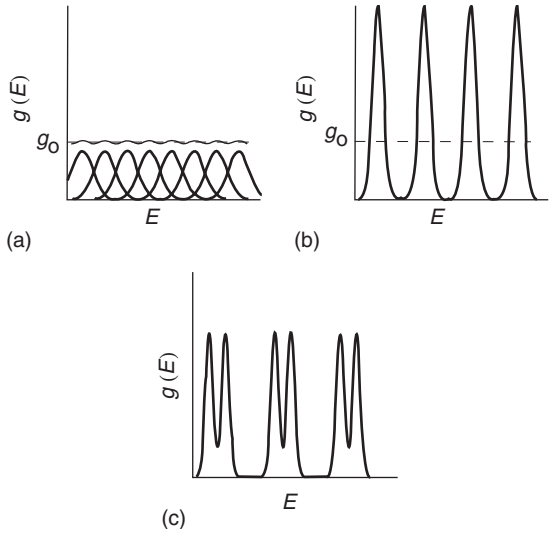


Figure 6 Landau level density of states. (a) At low fields, when the Landau level spacing is sufficiently small compared with the broadening produced by disorder that the levels overlap and the density of states deviates from the zero field value g_0 by only a small amount. (b) At higher fields, the spacing becomes large enough that the levels separate. (c) In high enough fields, the spin splitting becomes sufficiently large that each doubly degenerate Landau level splits into two spin-resolved levels.

lifetime $\tau_q (= \hbar/2\Gamma_L)$. For the situation shown in **Figure 6a**, only the fundamental term in the expansion ($s = 1$) is important and the higher harmonics are strongly damped.

The lifetime τ_q, determined by the impurity scattering, is sometimes identified (incorrectly) with the scattering time obtained from the mobility in eqn [1]. The mobility is related to a transport lifetime given by

$$\frac{1}{\tau_{tr}} = n_{imp}\frac{m^*}{\pi\hbar}\int_0^\pi (1-\cos\theta)|V(q)|^2\,d\theta \qquad [5]$$

where $q = 2k_F \sin(\theta/2)$ is the momentum change associated with scattering through an angle θ, $V(q)$ is the Fourier transform of the scattering potential, and n_{imp} is the density of impurities. The weighting factor $(1-\cos\theta)$ reflects the fact that scattering events with a large momentum change contribute more to the resistivity than small angle events.

Provided the final density of states in the scattering process remains the same, that is, for strongly overlapping Landau levels, the broadening is given by

$$\frac{1}{\tau_q} = n_{imp}\frac{m^*}{\pi\hbar}\int_0^\pi |V(q)|^2\,d\theta \qquad [6]$$

without the weighting factor. In remotely doped two-dimensional systems, where the ionized donors are displaced from the two-dimensional layer, $|V(q)|^2$ is strongly peaked near $q = 0$ and $\tau_q \ll \tau_{tr}$. In this case, the scattering time obtained from the mobility should not be used to determine the Landau level broadening.

At higher fields, when the Landau levels start to separate, the final density of states in the scattering process becomes strongly energy dependent. The scattering rates then become very dependent on where the Fermi level lies within the Landau level and the broadening has to be determined self-consistently. The Landau level shape becomes closer to Gaussian, rather than Lorentzian, with a width (Γ_G) obtained using the self-consistent Born approximation (SCBA) as

$$\Gamma_G = \left[\frac{\hbar^2\omega_c}{2\pi\tau_q}\right]^{1/2} \qquad [7]$$

τ_q is the quantity defined above in eqn [6]. The fact that Γ_G increases as $B^{1/2}$ can be understood physically in terms of a scattering rate ($\sim\Gamma_G$) proportional to the final density of states in the Landau level which is itself proportional to B/Γ_G. For Gaussian broadening, the damping of the fundamental term in the Fourier expansion analogous to eqn [4] is given by $\exp(-2\pi^2\Gamma_G^2/\hbar^2\omega_c^2)$. Using eqn [7], this is exactly the same (i.e., $\exp(-\pi/\omega_c\tau_q)$) as for Lorentzian broadening, so there is a smooth crossover from the low-field behavior, with overlapping Lorentzians, into this medium-field regime of nonoverlapping Gaussians.

In even higher fields, when the cyclotron radius becomes smaller than the characteristic correlation length of the scattering potential, it is no longer valid to use eqns [6] and [7] and the broadening is expected to become independent of B. For GaAs-based heterojunctions, with a donor spacer layer of typically 40 nm, this occurs for fields above ~ 2 T.

Transport Coefficients

The development of an oscillating Landau level structure in the density of states modifies the conductivity (and resistivity) coefficients. At low fields, this can be accounted for by replacing the density in eqns [1] and [3] by an effective density $n_{eff} = n_s g(E_F)/g_0$ and the transport time τ by $\tau_0 g_0/g(E_F)$ (where τ_0 is the zero-field value of τ_{tr}). Inverting the conductivity tensor then gives

$$\rho_{xx} = \rho_0\left(1 + 2\frac{\Delta g}{g_0}\right) \qquad [8]$$

with $\Delta g = g(E_F) - g_0$ and where a term of order $\hbar\omega_c/E_F$ has been neglected.

For ρ_{xy}, the situation is a little more complicated. In eqn [3] for σ_{xy}, oscillations appear that are both in phase with Δg and in quadrature. For the in-phase oscillations, the contributions from each of the two terms are large but of opposite sign and they substantially cancel leaving a residual component a factor of order $(\omega_c\tau_0)^{-2}$ smaller. Whether this term or the quadrature term dominates depends on the relative values of τ_q and τ_0. For $\tau_q \approx \tau_0$, the two are comparable; when $\tau_q \ll \tau_0$, the quadrature term always dominates and maxima in ρ_{xx} correspond to ρ_{xy} crossing the linear background term ($B/n_s e$).

When the Landau levels become well separated, it is still valid to take $n_{eff} \sim g(E_F)$, provided the Fermi level lies inside a Landau level. For $\omega_c\tau \gg 1$, the expression for σ_{xx} can be rewritten as

$$\sigma_{xx} = e^2 D^* g(E_F)$$

where $D^* = R_{cycl}^2/2\tau$ is a diffusion coefficient. This can be interpreted as a diffusion process with conduction taking place by electrons hopping between orbits with centers spaced, on an average, by the cyclotron radius. Taking τ as $\tau_0 g_0/g(E_F)$, one recovers eqn [8] above. More generally, τ acquires a

complicated dependence on the relative magnitudes of R_{cycl}, l_m, and the correlation length that characterizes the scattering potential.

When the Fermi level lies in the gap between Landau levels, $g(E_F) = 0$ and $\sigma_{xx} \approx 0$. The conductivity is then determined by other processes such as activation to the adjacent levels. The limit of $\sigma_{xx} \to 0$ is the integer quantum Hall regime where σ_{xy} (and ρ_{xy}) are quantized and where ρ_{xx} also tends to zero.

Typical Experimental Results

Figure 7 shows an example of the Shubnikov–de Haas and Hall effects measured in a (Ga,Al)As/GaAs heterojunction. At low fields, spin is not resolved and the minima in ρ_{xx} appear at even filling factors given by $\nu = n_s h/eB$. The density determined from the period of the oscillations agrees well with that derived from the low-field Hall coefficient. This is to be expected because the (linear) Hall resistance extrapolates accurately through the middle of the plateaus that appear at higher fields. If $\delta\rho_{xx}$ is the deviation of ρ_{xx} from the mean value, and $\delta\rho_{xy}$ the deviation of ρ_{xy} from the linear term $(B/n_s e)$, then it can be seen that $\delta\rho_{xx}$ and $\delta\rho_{xy}$ oscillate in quadrature. As noted above, this is expected for a modulation-doped structure where $\tau_{tr} \gg \tau_q$.

Above ~ 2 T, ρ_{xx} approaches zero at the minima, corresponding to well-separated Landau levels, and above ~ 4 T spin starts to be resolved and odd filling factors become visible in both ρ_{xx} and ρ_{xy}. This is initiated when the Zeeman splitting of the Landau levels becomes comparable with the broadening. Once spin splitting appears, it is strongly enhanced by the exchange interaction that results from the spin polarization associated with unequal numbers of filled spin-up and spin-down Landau levels. This induces a "boot-strapping" effect and causes the spin splitting to rapidly become fully resolved.

At low fields, the amplitude of the Shubnikov–de Haas oscillations is damped exponentially, shown in **Figure 8** in a Dingle plot. The damping can be attributed to both disorder broadening of the Landau levels and to the thermal spread of the Fermi function. For each harmonic s in the expansion given in eqn [4], there is a corresponding thermal damping term $sX_T/\sinh(sX_T)$, where $X_T = 2\pi^2 k_B T/\hbar\omega_c$. This saturates to 1 at low temperatures but at higher temperatures can be used experimentally to obtain effective masses from the temperature dependence of the oscillation amplitude. Often, only the fundamental term ($s = 1$) is relevant and in the present instance, where this is so, the data can be corrected very easily for thermal damping using the well-known value of $0.067m_e$ for the effective mass of GaAs. The corrected amplitudes are then given by (cf. eqns [4] and [8])

$$\Delta\rho_{xx} = 4\rho_0 \exp(-\pi/\mu_q B) \qquad [9]$$

Here a quantum mobility ($\mu_q = e\tau_q/m^*$) has been introduced in analogy with the standard transport mobility. The good agreement between theory and experiment, in particular the fact that the intercept of $4\rho_0$ is given correctly, confirms that in this low-field regime the theory outlined above is valid and, in particular, that σ_{xx} is proportional to the square (not the first power) of the density of states.

The slope of the Dingle plot gives a quantum mobility $\mu_q = 1.58\ m^2\ (Vs)^{-1}$, 10 times smaller than the transport mobility of $15.3\,m^2\,(Vs)^{-1}$. This confirms

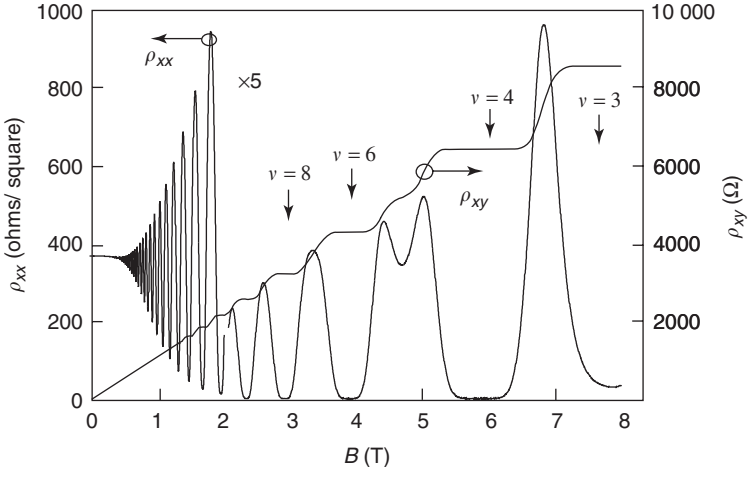


Figure 7 Shubnikov–de Haas oscillations and Hall effect in a modulation-doped GaAs/$Ga_{0.7}Al_{0.3}As$ heterojunction measured at 1.18 K. Electron density was $5.6 \times 10^{15}\,m^{-2}$ and mobility $15.3\,m^2\,(Vs)^{-1}$.

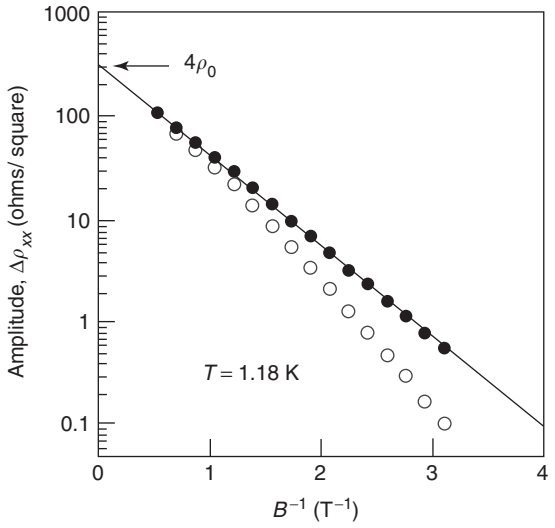


Figure 8 Dingle plot for low-field resistivity data shown in **Figure 7**. Open symbols represent amplitudes, as measured. Solid symbols have been corrected for thermal damping as described in the text.

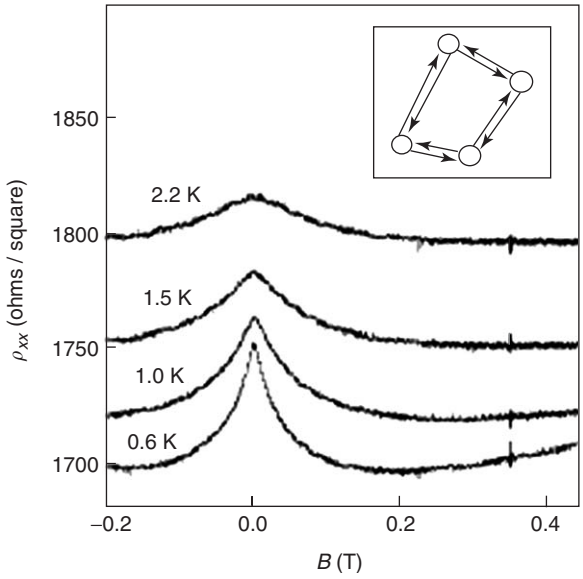


Figure 9 Weak localization in a *p*-SiGe sample with a density of $5.7 \times 10^{15}\,\mathrm{m}^{-2}$ for temperatures between 0.6 and 2.2 K. The narrowing of the peak as the temperature is lowered results from increased phase breaking times. The inset shows schematically the scattering paths, in opposite directions, that interfere coherently to increase the resistance at $B = 0$.

directly in this sample (where the distance by which the dopants are set back from the interface is 6 nm) that small angle scattering dominates. In higher-mobility GaAs/(Ga, Al)As heterojunctions where larger spacer layers are used, the ratio $\mu_{\mathrm{tr}}/\mu_{\mathrm{q}}$ is even larger, typically 20 or more.

In **Figure 7**, in the region above 2 T where the minima in ρ_{xx} approach zero, the Hall resistance develops flat plateaus at values of $\rho_{xy} = h/ve^2$. This is the integer Hall effect which appears when the Fermi level is pinned in localized states between the Landau levels. In cleaner samples when the Landau levels initially separate, there are few localized states between them and the Fermi energy can jump between Landau levels without the appearances of well-defined plateaus. In such samples, other structures also become apparent at high fields and low temperatures. This is the fractional quantum Hall effect which appears when exchange and Coulomb interaction energies, within the Landau level, become large. New gaps appear in the density of states, most strongly at filling factors 1/3 and 2/3 which give rise to new zeroes in ρ_{xx} and associated plateaus in ρ_{xy}. Both the integer and fractional quantum Hall effects are discussed in more detail elsewhere in this encyclopedia.

Weak Localization and Interaction Effects

At very low magnetic fields, additional magnetoresistance features are often seen. The example in **Figure 9**, for a two-dimensional hole gas in an (SiGe) quantum well, shows a peak around $B = 0$. This is associated with weak localization where electrons scattered around closed paths in opposite directions interfere coherently (see inset). At $B = 0$, this produces an increased back-scattering and increased resistance. With small applied magnetic fields, the flux through the closed path induces an extra Aharonov–Bohm phase, of opposite sign for the two directions, which suppresses this scattering and reduces the resistance. A value for the phase breaking time can be obtained from the width of the magnetoresistance peak. This is typically one or two orders of magnitude longer than the transport lifetime and increases with decreasing temperature as T^{-p} where p is of order 1. The weak localization correction to the conductivity

$$\Delta\sigma_{xx}^{\mathrm{WL}} = \frac{pe^2}{\pi h} \ln T$$

is of order e^2/h. Although quite obvious in the example shown, it can be difficult to detect in lower-resistivity (higher-mobility) samples where it is of the same order of magnitude but is only a very small correction to a Drude conductivity, many orders of magnitude larger.

In addition to the weak localization term, there is also a Coulomb interaction correction, of a similar order of magnitude, and also with a $\ln(T)$ dependence. This can be distinguished from the weak localization by the much smaller sensitivity to the

magnetic field. In a famous paper published by Abrahams, Anderson, Lee, and Ramakrishnan in 1981, it was shown that in two dimensions and in the absence of interactions, all systems are insulating at $T = 0$, that is, ultimately, at low T, weak localization will dominate, however large the high temperature conductivity is. More recently, it has become evident that when interactions are strong, particularly the F_0^σ Landau parameter that renormalizes the spin susceptibility, the interaction correction is of the opposite sign to the weak localization term and may become sufficiently large to overcome it. This results in metallic behavior as $T \rightarrow 0$ (i.e., a reduction in resistance) and there is the possibility of a crossover from insulating to metallic temperature dependence as a function of density. Details of this phenomenon are not yet fully worked out and it remains an active area of research.

See also: Quantum Hall Effect; Semiconductor Nanostructures; Shubnikov–de Haas and de Haas–van Alphen Techniques.

PACS: 73.20. – r; 73.40. – c; 73.50.Jt

Further Reading

Ando T, Fowler AB, and Stern F (1982) Electronic properties of two-dimensional systems. *Review of Modern Physics* 54: 437.

Davies JH (1998) *The Physics of Low-Dimensional Semiconductors: An Introduction*. Cambridge: Cambridge University Press.

Harris JJ, Pals JA, and Woltjer R (1989) Electronic transport in low-dimensional structures. *Reports of Progress in Physics* 52: 1217.

Kelly MJ (1995) *Low-Dimensional Semiconductore: Materials, Physics, Technology, Devices*. Oxford: Oxford University Press.

Kravchenko SV and Sarachik MP (2004) Metal-insulator transition in two-dimensional electron systems. *Reports of Progress in Physics* 67: 1.

Look DC (1989) *Electrical Characterisation of GaAs Materials and Devices*. Chicester: Wiley.

Morkoç H, Unlu H, and Ji G (1991) *Principles and Technology of MODFETs*. Chicester: Wiley.

Shoenberg D (1984) *Magnetic Oscillations in Metals*. Cambridge: Cambridge University Press.

Stillman GE, Bose SS, Kim MH, Lee B, and Low TS (1994) Characterisation and Properties of Semiconductors. In: Mahajan S (ed.), *Handbook of Semiconductors* (completely revised edition, 3). Amsterdam: North-Holland.

Nomenclature

B	magnetic flux density (T)
E	electric field (Vm^{-1})
E	single particle energy (J, eV)
E_F	Fermi energy (J, eV)
$g(E)$	energy density of states (Jm^{-2})
J	current density (Am^{-1})
l_m	magnetic length (m)
m^*	effective mass (kg)
m_e	free electron mass (kg)
n_L	Landau level index
n_s	sheet carrier density (m^{-2})
R_{cycl}	cyclotron radius (m)
R_{xx}, R_{xy}	resistance (Ω)
Γ_G	Landau level broadening (Gaussian) (J, eV)
Γ_L	Landau level broadening (Lorentzian) (J, eV)
μ	electrochemical potential (V)
μ	mobility (m^2Vs^{-1})
μ_q	quantum mobility (m^2Vs^{-1})
ν	filling factor
ρ_{xx}, ρ_{xy}	resistivity (ohms/square, Ω)
σ_0	zero-field conductivity (square/ohm)
σ_{xx}, σ_{xy}	conductivity (Ω^{-1})
τ_q	quantum lifetime (s)
τ_{tr}	transport lifetime (s)
ω_c	cyclotron frequency (s^{-1})

Treated Surfaces, Optical Properties of

H Ueba, Toyama University, Toyama, Japan

Introduction

> God made solids, but surfaces were the work of the Devil
> W Pauli

The ultimate goal of surface science is to unravel the atomic structure, electronic and vibrational properties of solid surfaces and adsorbates. All of them are combined into our understanding of the microscopic mechanism responsible for a wide variety of physical properties and chemical reactions taking place at surfaces. Various techniques have been employed to investigate clean and adsorbate covered surfaces using appropriate probes such as photons, electrons, ions, thermal atoms, and neutrons. The photonic probe has several advantages over other massive particles. The spectral resolutions are incomparably better than those attained by other spectroscopies. The temporal resolutions are now down to femtoseconds

that allow real-time monitoring of electronic and vibrational dynamics as well as surface chemical reactions. By using nonlinear optical processes, it is possible to eliminate the response from the bulk, and the information on the surface is selectively obtained. Thus, optical probe yields a huge amount of indispensable knowledge in surface science.

When atoms or molecules are adsorbed on metal surfaces, their electronic and vibrational properties are modified depending on the bonding sites, geometries, and the change of charge distribution and occupancy in the molecular orbitals by the chemical bond to the substrate.

In this atricle, the surface study with photonic probes is presented with a focus on nonlinear and time-resolved spectroscopies for the electronic and vibrational properties of metal surfaces and adsorbates.

Photoelectron Spectroscopy

Photoelectron spectroscopy (PES; X-ray photoemission (XPS) from deep core levels and ultraviolet photoemission (UPS) from valence state) is one of the most versatile tools to study the electronic properties of atoms, molecules and solids. As illustrated in **Figure 1a**, in photoemission, photons of well-defined energy are absorbed by the process of electronic excitation. If the photon energy is high enough to ionize the sample, the kinetic energy distribution (photoemission spectrum) of the emitted electrons provides the electronic properties of the initial electronic states in a simplified single-particle picture. Angle-resolved photoemission provides information of the dispersion curve of delocalized electrons in solids, or localized electrons in atomic and molecular orbitals in their initial states.

Core level PES is a powerful technique to study surface atomic structure and chemical reactivity of surfaces. For example, surface atoms of a clean Si(1 0 0) are reconstructed to form a buckled asymmetric dimer structure (see **Figure 2** for the top and side view of the $c(4 \times 2)$ reconstructed Si(1 0 0) surface at low temperature). With the help of the electron diffraction pattern and the atomic-resolved image using a scanning tunneling microscope, high-resolution Si $2p$ photoelectron spectroscopy, which employs a bright photon source of synchrotron radiation, enables the deconvolution of the Si $2p$ spectrum to identify the upper and lower Si atoms in the topmost dimer row as well as Si atoms in the second layer and in the bulk. By observing the change of each component upon adsorption of molecules on the surface, one is also able to study the chemical reactivity at the topmost surface layer at the atomic level.

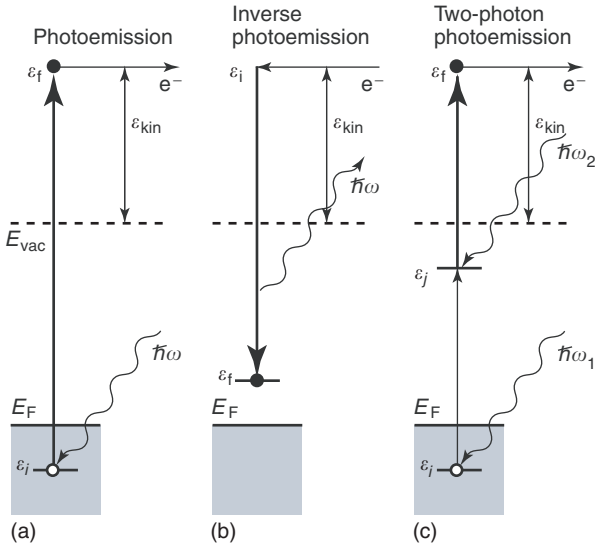


Figure 1 Schematic energy diagram for (a) photoemission, (b) inverse photoemission, and (c) two-photon photoemission. The kinetic energy ε_{kin} of photoelectron and incident electron is measured relative to the vacuum level E_{vac}.

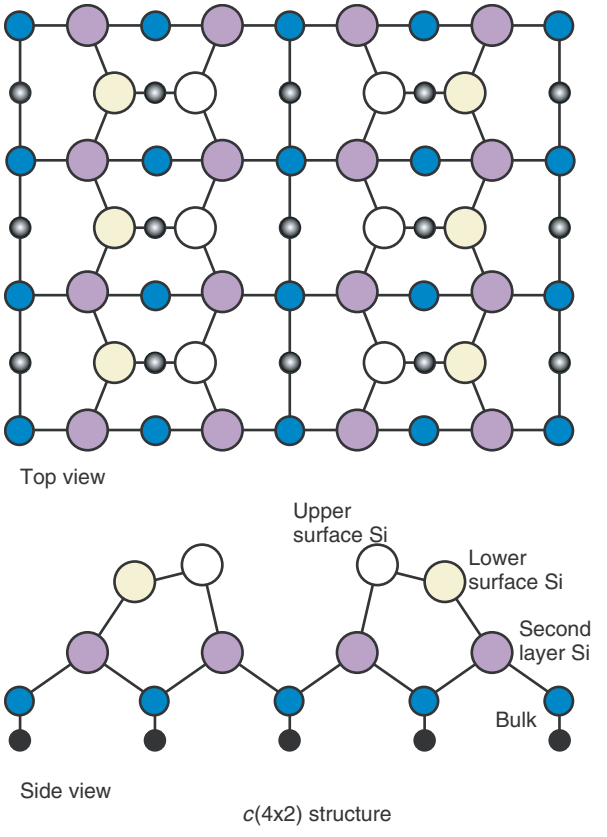


Figure 2 Top and side view of a $c(4 \times 2)$-buckled asymmetric dimer row reconstruction of a Si(1 0 0) surface at low temperature.

Photoemission is also very useful for the study of adsorbates on solid surfaces. A comparison of the photoelectron spectra before (gas-phase) and after adsorption enables one to study the change of electronic properties, in particular, the change of chemical environments and the nature of the atomic or molecular orbitals responsible for the bonding as well as the geometric arrangement at the surface.

The energy range accessible to photoemission is limited to the initial state below the Fermi level (E_F) and to final states above the vacuum level (E_{vac}). The unoccupied states between E_F and E_{vac} can be investigated by means of inverse photoemission (IPS) (**Figure 1b**). This is a time-reversal process of photoemission, which observes the radiative transition of the incident electron into the unoccupied state above E_F. For a fixed energy of the incoming electron, the energy spectrum of the emitted photons directly reflects the electronic properties of the initially unoccupied states.

X-ray absorption spectroscopy (XAS) using bright synchrotron radiation is a powerful tool to study local electronic and atomic structure of solids and adsorbates. The information on the local electronic properties can be extracted from the X-ray absorption spectrum in the vicinity of an absorption edge of an atom, whereas detailed structural properties can be determined from the extended X-ray absorption fine structure (EXAFS), having a series of oscillations due to an interference between the outgoing and backscattered photoelectron wave on the high photon energy side of the absorption edge. The kinetic energy of the photoelectron increases and the de Brogli wavelength becomes shorter with an increase in incident X-ray energy. At a certain wavelength, the interference with the scattered wave from the surrounding atoms produces a constructive or destructive effect on photoionization. This results in the oscillatory behavior in the X-ray absorption coefficient, and analyzing these oscillations close to the absorption energy of a particular atom determines the atomic number, distance, and coordination number of the atoms. In extending EXAFS to surface analysis, the weak signal and surface sensitivity is ensured by detecting the Auger electron emission (escape depth of $\sim 20\,\text{Å}$) of an atom as a function of photon energy.

XAS, in which a core-electron at ε_c is excited to an unoccupied state ε_a above E_F, is also a powerful method for studying the electronic properties of adsorbates. In particular, a sudden creation of the core-hole pulls down the energy level of the outer adsorbate orbitals by the intra-adsorbate Coulomb interaction U so that XAS measures the excitation energy,

$$E_{XAS} = \varepsilon_a - U_{ac} - \varepsilon_c = I - A - U_{ac} \qquad [1]$$

where I and A are the ionization potential and the electron affinity level of the adsorbate, respectively. In the X-ray absorption process, the charge state of the adsorbate remains neutral in the absence of the metal–adsorbate mixing so that no final relaxation shift due to the image screening contributes to E_{XAS}. However, the XAS spectrum crucially depends on the energy position of the unoccupied state relative to E_F in the presence of the core-hole. When it ends up above E_F, the excited state has a finite lifetime due to the decay into the unoccupied part of the metal and the XAS spectrum exhibits a Lorentzian shape accompanied with the so-called "Fermi edge singularity" (IR divergence) at E_F, corresponding to the onset of absorption. In the opposite case where it is pulled down below E_F, an absorption edge appears due to simultaneous excitation of a low-energy electron–hole pair (EHP) in the metal.

These optical processes have different final states for adsorbates, that is, a neutral for optical excitation in XAS, a positive ion for PES, and a negative ion for IPS. An appropriate point of view is to consider the electron states of an N electron system and the $N \pm 1$ electron system left behind after an electron is removed or added by photoemission and inverse photoemission, respectively. For a bulk solid with very large N, the final states do not differ from the initial state. However this is not the case for adsorbates. In PES, for example, a hole created in the core level or narrow valence states by the photoemission process forms a local positive charge and attracts screening electrons. In the case of XPS, one measures

$$E_{XPS} = -\varepsilon_c - \upsilon_c = I - \upsilon_c \qquad [2]$$

where υ_c, the core-level, shifts upward by the hole imaged potential. This is an extra-adsorbate screening in which the compensating charge builds up at the surface. It is also noted that there are other types of intra-adsorbate relaxation followed by an interplay of different dynamic screening processes: (1) charge transfer from the substrate into the initially unoccupied state if it shifts down below E_F by the core-hole potential, (2) extra-adsorbate screening of the core-hole by the surface response of the substrate. These dynamic core-hole screening processes play an essential role in the interpretation of the XPS spectra of adsorbates on metal surfaces. A similar argument also holds for IPS. The final states probed by IPS differs from that involved in the final states of the XAS or PES processes, unless the addition of an extra

electron does not perturb the ground states of the adsorbates. For a negatively charged adsorbate in the final state of and IPS, the induced positive charge at the metal surface binds the electron more strongly, and IPS measures,

$$E_{\mathrm{IPS}} = -\varepsilon_{\mathrm{a}} + \nu_{\mathrm{a}} = A + \nu_{\mathrm{a}} \qquad [3]$$

that is, the downward relaxation shift from ε_{a}.

In addition to XPS, the UPS study of the valence orbitals of adsorbed molecules provides rich information on the nature of bonding to the substrates. For adsorption of carbon monoxide (CO) – the most extensively studied molecule – the UPS spectrum on a Si surface exhibits a similar order of the valence orbitals (5σ, 1π, and 4σ) as in the gas phase. This indicates that CO retains much of its valence structure in the gas phase upon adsorption on a Si surface. On metal surfaces, on the other hand, the 5σ orbital belonging to the carbon atom shows a substantial downward shift relative to the 1π and 4σ orbitals when compared to their ionization energies in the gas phase. This evidences that the highest occupied 5σ (HOMO) orbital is responsible for the chemical bond formation to the substrate as schematically illustrated in **Figure 3**. Upon CO chemisorption, charge transfer (back donation) occurs from the substrate into the lowest unoccupied molecular orbital (LUMO) $2\pi^*$. Since $2\pi^*$ which can be observed by IPS orbital has an antibonding character with respect to the C–O bond, the partial occupancy weakens the C–O bond strength, thereby causing the elongation of the bond distance and the decrease in the frequency of the C–O stretch mode. It is important to note that the electronic properties of the $2\pi^*$ states, such as the position above E_{F} and broadening due to hybridization with the substrate electrons, play important roles in the de-excitation processes of a core-hole created in XAS and XPS from the C and O 1*s* level.

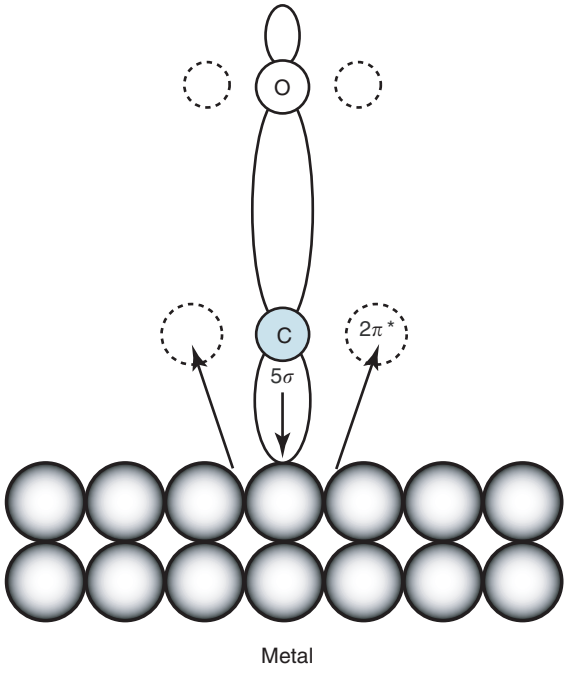


Figure 3 CO chemisorption on a metal surface.

Two-Photon Photoemission

Photoemission occurs not only by absorbing a single photon exceeding the work function ϕ but also by successive absorption of two photons of energy $\hbar\omega_1$ and $\hbar\omega_2$ under the conditions

$$\hbar\omega_1 < \phi \quad \text{and} \quad \hbar\omega_1 + \hbar\omega_2 > \phi \qquad [4]$$

Hereafter all the energies are referred to E_{F}, unless otherwise stated. Two-photon photoemission (2PPE) illustrated in **Figure 1c** is a variant of PES, which allows investigation of unoccupied electronic states located between E_{F} and the vacuum level (E_{vac}) of a metal or semiconductor. In the 2PPE process, the first pump photon $\hbar\omega_1$ excites an electron from an occupied state below E_{F} to an intermediate state between E_{F} and E_{vac}. Upon absorption of a second probe photon $\hbar\omega_2$ during the survival time of the electron in a transiently populated intermediate state (including virtual transition), electrons are excited above E_{vac} forming final photoelectron states, whose kinetic energy distribution curve gives a 2PPE spectrum. 2PPE spectroscopy thus combines both the advantages of traditional one-photon photoemission and IPS spectroscopies, and provides a unique opportunity which allows a simultaneous observation of occupied and unoccupied excited states in metals and semiconductors.

The spectral feature of the 2PPE depends not only on the electronic properties of the initial and intermediate states but also on the excitation photon energy. It has been well established that a peak in the 2PPE spectrum can be easily assigned to an initial or intermediate state, depending on the peak shift with $\hbar\omega_1$ and $\hbar\omega_2$. When a coherent two-photon excitation from an occupied state ε_i below E_{F} brings an electron above E_{vac}, the photoelectron kinetic energy E_{kin} measured from E_{F} shifts as

$$E_{\mathrm{kin}} = \hbar\omega_1 + \hbar\omega_2 + \varepsilon_i \qquad [5]$$

When it is due to two independent processes, a transient population in an intermediate state ε_j by $\hbar\omega_1$ and a subsequent excitation above E_{vac} by $\hbar\omega_2$, it

varies with $\hbar\omega_2$, that is,

$$E_{\text{kin}} = \hbar\omega_2 + \varepsilon_j \qquad [6]$$

Figure 4a schematically illustrates the possible excitation pathways leading to final photoelectron states. This energy diagram is composed of a bulk continuum and a localized surface state below E_F of a metal substrate, and an unoccupied state between E_F and E_{vac}. This diagram can be viewed as the simplest version of a clean metal surface such as Cu(1 1 1) and Ag(1 1 1) having the occupied surface state and unoccupied image potential state. **Figure 4b** depicts E_{kin} as a function of $\hbar\omega = \hbar\omega_1 = \hbar\omega_2$. The incoherent step-by-step single photon process (a) gives a slope = l, while the coherent two-photon process (b) gives a slope = 2. Their extrapolations to $\hbar\omega \to 0$ determine ε_i and ε_j. At a resonant excitation ($\hbar\omega = \varepsilon_j - \varepsilon_i$), two peaks arising from the initially occupied surface state and unoccupied intermediate states are amalgamated to a single resonance narrowing peak.

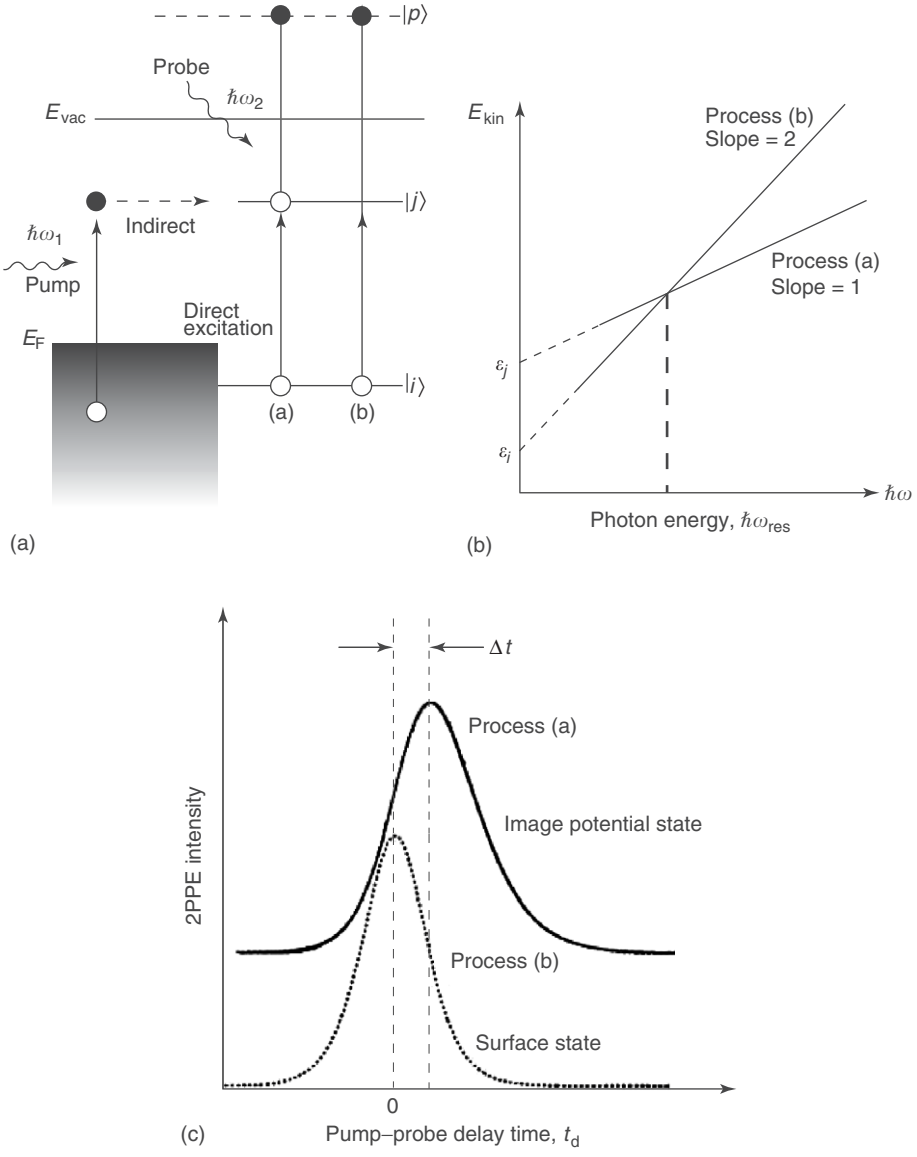


Figure 4 (a) Elementary excitation processes in two-photon photoemission. When a surface state and a continuum of the substrate are available for initial states $|i\rangle$, there are two paths leading to a temporal population in the intermediate state $|j\rangle$. One is direct excitation leading to step-by-step one-photon transition, and another is an indirect process via scattering of the photoexcited electron from a continuum. If there is no surface state like Cu(1 0 0), only the latter nonresonant process is possible. There are two paths leading to a final photoelectron state $|p\rangle$ via (a) step-by-step one-photon process, and (b) direct two-photon ionization process from $|i\rangle$. (b) $\hbar\omega$ dependence of the photoelectron kinetic energy, and $\hbar\omega_{res} = \varepsilon_j - \varepsilon_i$ is a resonant excitation. (c) 2PPE intensity (cross-correlation trace) for the occupied surface state and for the unoccupied image potential state with a finite lifetime.

Image Potential States at Metal Surfaces

One of the most successful applications of energy-, and time-resolved 2PPE is an observation of image potential states at metal surfaces and a direct determination of the lifetimes. When an electron is placed at a distance z from a metal surface, it induces an image charge as a consequence of a many-body screening by the conduction band electrons in a metal. The electron in front of a metal surface forms the Rydberg-like bound states by a potential consisting of the attractive image potential $V(z) = -(e^2/4\pi\varepsilon_0)(1/4z)$ on the vacuum side, and a repulsive crystal barrier at the image plane on the metal side. In a hydrogenic model with a finite repulsive surface barrier, the image potential states have the energies

$$E_n = -\frac{0.85}{(n+a)^2}\ (\text{eV}) \qquad [7]$$

where n is the principal quantum number and a is the so-called quantum defect due to the finite probability of the electron penetrating into the bulk.

The wave functions of these states localized at the surface decay exponentially into the bulk. Compared to the electronic states in the bulk, image potential states have relatively long lifetimes. **Figure 5a** depicts the band structure including an occupied surface

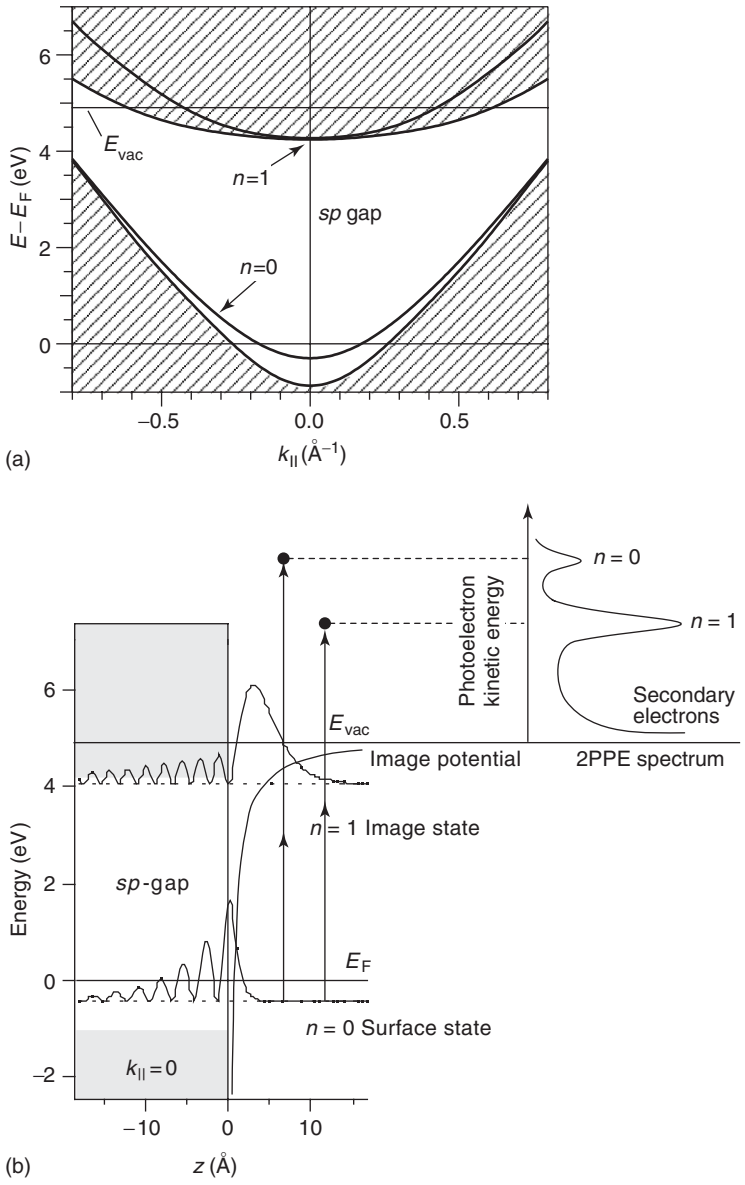


Figure 5 (a) Band structure of Cu(1 1 1) along $k_{||}$. (Modified figure with permission from Smith NV (1985) Phase analysis of image states and surface states associated with nearly-free-electron band gaps. *Physical Review B* 32: 3549–3555; © American Physical Society.) (b) Image potential state and its wave function at $k_{||} = 0$. A typical 2PPE spectrum is also shown. (Courtesy of M Wolf, Free University Berlin.)

state ($n = 0$) and an unoccupied image potential state ($n = 1$) of a Cu(1 1 1) along $k_{||}$. **Figure 5b** illustrates their wave functions at $k_{||} = 0$, whose extent of the penetrations into the bulk depend on their position relative to the projected *sp*-bandgap in the bulk band structure. A typical 2PPE spectrum including photoexcited secondary electron emission is also shown in **Figure 5b**.

Pump–Probe 2PPE Spectroscopy

A great capability of 2PPE lies in the ultrafast time-resolved pump–probe nonlinear spectroscopy to study the carrier dynamics at surfaces using laser pulses with the duration of several tens of femtoseconds. It is a natural extension of the energy-resolved 2PPE by introducing a time delay between the pump and probe photons, which allows the direct measurement of the lifetime of the intermediate state such as the image potential states on metal surfaces. The key quantity to govern the physics underlying TR-2PPE is the relative timescale among energy (population) relaxation time T_1, dephasing time T_2, and pulse duration t_p. It is well known in the context of quantum optics and ultrashort laser pulse phenomena that in the case $t_p \gg T_2$, there is no coherent superposition of the polarization and electromagnetic field oscillations. The memory of the medium is only through the change of the population. This is a case where the rate equation (no coherence effect) approach is appropriate for describing the temporal response of the excited population. In this treatment, the temporal change of the material system occurring within T_2 is averaged so that only the dynamical response beyond a timescale of T_2 is taken into consideration.

On the other hand, in a case $t_p < T_2$, namely when the exciting pulse duration becomes comparable with, or even shorter than the phase relaxation time of the excited medium, and in the presence of a pump and probe overlap during a dynamical change of photo induced polarization, a model including coherent interaction between the photon field and the transition dipole moment is required to describe the temporal response of the system. Central issues involved in the analysis of the transient 2PPE signal in order to deduce reliable T_1 of the image potential states are: (1) how to determine the lifetime T_1 of the image potential state using pulses $t_p \gg T_1$, (2) how to accurately determine the pump–probe delay time $t_d = 0$, (3) how the total dephasing time T_2 (or pure dephasing time T_2^*) plays a role, (4) how the pulse duration affects the transient signal. Some of the answers to these questions have been presented by Hertel *et al*. in 1996 in their study of the electron dynamics at a Cu(111) surface. A precise measurement of t_d can be made utilizing the non-resonant 2PPE process from the surface state as a reference. The peak position of the cross-correlation for the surface state determines the true time zero of the pump pulse with respect to the probe pulse. Due to the finite T_1 of the image potential state, on the other hand, the maximum of the cross-correlation trace appears at the positive t_d after the pump pulse has passed its maximum, as schematically illustrated in **Figure 4c**. The solution of the optical Bloch equations for the density matrix elements characterized by T_1 and T_2 allows the determination of T_1 (order of several to tens of femtoseconds), which is much shorter than t_p.

Vibrational Properties of Adsorbates

Among the various techniques to probe solid surfaces, surface vibrational spectroscopy enables one to study the chemical nature of surfaces and adsorbates. In particular, when combined with other techniques of surface analysis such as electron diffraction, infrared reflection–absorption spectroscopy (IRAS) has established its firm position to determine bonding sites and adsorbate geometry through the selection rule of a dipole transition. An incident IR beam at the angle of θ to the surface normal is decomposed into *s*- and *p*-polarized components (see **Figure 6**). Since the phase change of the *s*-component upon reflection at the surface is always near π for any θ, the total field becomes zero at the surface. On the other hand, the normal component of the *p*-polarized light increases with an increase in θ. The phase change remains small until at large θ, and rapidly drops toward $\pi/2$ at the grazing incidence. The normal component of the total *p*-polarized field reaches its maximum just before falling to zero at grazing incidence. These differences between the *s*- and *p*-polarized field gives the selection rule for the vibrational excitation of adsorbates, that is, the vibrational modes with a dipole moment normal to the surface are excited in IRAS.

Figure 7 shows several fundamental types of adsorption and their corresponding vibrational spectra

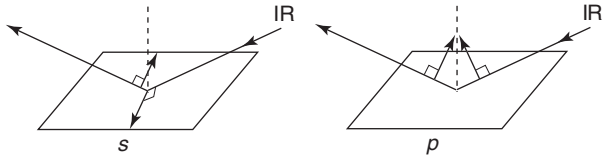


Figure 6 Electric field vector of the *s*- and *p*-polarized IR incidence and reflection.

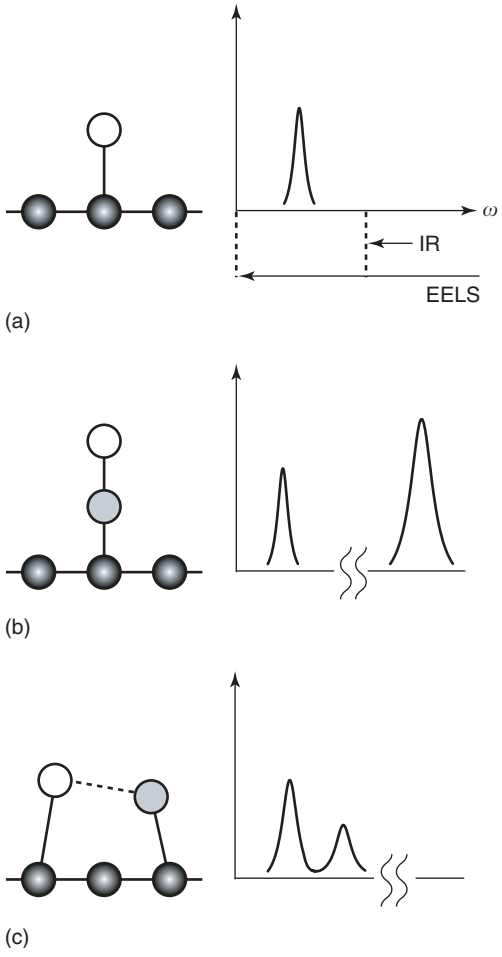


Figure 7 Schematic representation of some fundamental adsorption types of an atom and a diatomic molecule with their corresponding vibrational spectra.

observed by IRAS and energy loss spectroscopy (EELS). For simplicity, the mass of the adsorbate is assumed to be much lighter than that of the substrate atom with which the adsorbate forms the chemical bond. **Figure 7a** illustrates an atom adsorbed at on-top site of a substrate. In this simplest case, one observes a low-energy (in far infrared region) mode corresponding to the stretching vibration between substrate atom and adsorbed atom. For a doubly bonded atom in a symmetric bridge position one LEES peak at a lower energy compared to an atom adsorbed at on-top site is also observed. The single peak splits to double peaks with the different intensity for an adsorption site of lower symmetry.

Figure 7b illustrates a case of a diatomic molecule at on-top site and the molecular axis is perpendicular to the substrate (see also **Figure 3**). The low-frequency mode detected by EELS corresponds to the stretch vibration between a bonded atom of an adsorbed molecule and a substrate atom or derives from the vibration of the whole molecule against the substrate. The high-frequency mode observed by EELS as well as by IRAS is the intramolecular stretch vibration normal to the surface. For a diatomic molecule adsorbed at symmetric bridge site, these peaks shift to the lower energies. In the case of a CO molecule on metal surfaces, a larger back donation of electrons from the substrate into the $2\pi^*$ orbital of a bridge or high coordination sites such as three and four hollow sites on f.c.c.(111) and f.c.c.(100) surfaces, respectively, causes a substantial red-shift of the C–O stretch frequency. This helps a determination of the adsorption site of a CO molecule.

A fundamental issue of a diatomic molecule adsorption includes whether the molecule dissociates to two constituent atoms , each of which bonds to the substrate as a separate entity. The absence of the intramolecular stretch mode evidences a dissociation of a molecule upon adsorption. In this case, two loss peaks will be observed in the EELS spectrum, as being indicative of the vibrations between isolated two constituent atoms and the substrate, as depicted in **Figure 7c**.

Among myriad of experiments for carbon monoxide (CO) on metal surfaces, **Figure 8** shows the IRAS spectrum of CO/Pt(1 1 1) at the coverage $\theta = 0.5$ and temperature of 95 K. As illustrated in the figure, CO molecules occupy both the on-top and bridge sites and form a $c(4 \times 2)$ ordered structure at this coverage, and the corresponding peaks are observed at the red-shifted position compared to the gas-phase CO at $2130\,\mathrm{cm}^{-1}$. The Pt–C stretch mode for the on-top site was also seen in this spectrum.

EELS is also one of the most prevalent techniques in surface vibrational spectroscopy and to study electronic excitation as well as elementary excitations, such as surface phonons or surface plasmons. Vibrational EELS can be considered as the electron-analog of Raman scattering spectroscopy. A monochromatic electron beam with a primary energy of a few electron volts is incident onto adsorbate-covered surfaces, and loss features of the scattered electron are observed at the energy corresponding to the vibrational excitations. The selection rules that determine whether a vibrational band can be observed depend upon the nature of the substrate and also the experimental geometry, that is, the angles of the incident and (analyzed) scattered beams with respect to the surface. For metal substrates and a specular geometry, scattering is predominantly by a long-range dipole scattering. In this case, only those vibrational modes giving rise to a dipole change normal to the surface can be observed. In an off-specular geometry, electrons lose energy to the surface species by a short-range impact scattering. In this case, the

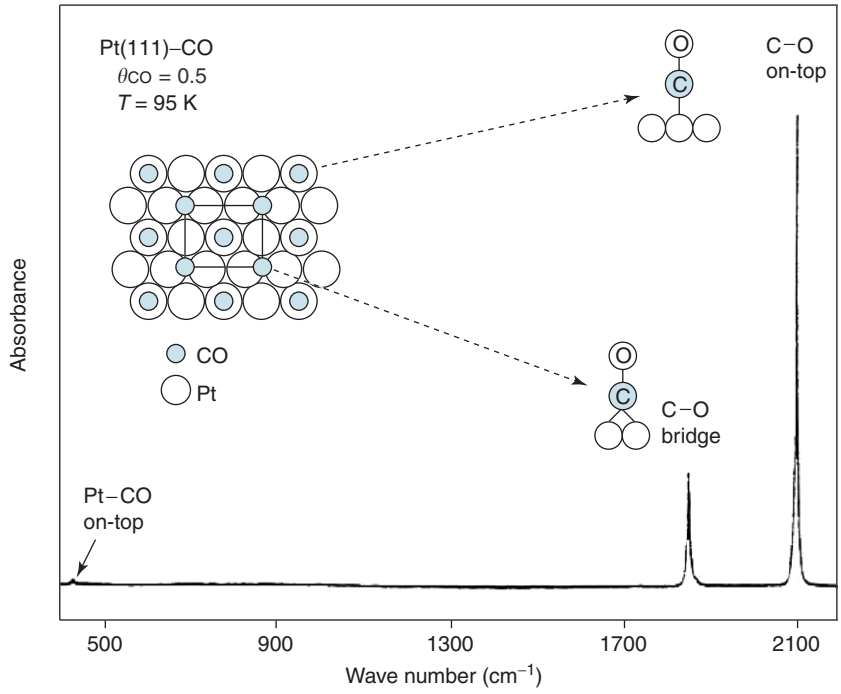


Figure 8 IR absorption spectrum of the $c(4 \times 2)$-ordered structure of CO on Pt(1 1 1) at coverage $\theta = 0.5$ and $T = 95$ K. (Reproduced from Schweizer E, Persson BNJ, Tushaus M, Hoge D, and Bradshaw AM (1989) The potential energy surface, vibrational phase relaxation and the order–disorder transition in the adsorption system Pt(1 1 1)–CO. *Surface Science* 213: 49–89, with permission from Elsevier.)

loss features are relatively weak but all vibrations may be observed. Angular distribution of the inelastically scattered electrons provides rich information of the structure and geometry of adsorbates. Thus, EELS has several advantages (high sensitivity, variable selection rules, and spectral acquisition to below $400\,\mathrm{cm}^{-1}$) and disadvantages (ultrahigh vacuum environment and poor spectral resolution) over IRAS. As a good example, **Figure 9** shows the EELS spectrum of the ordered $c(4 \times 2)$ CO adlayer at the overage of $\theta = 0.5$ on Pt(111) at 92 K (same system as **Figure 8**). It clearly shows the Pt–C stretch modes associated with the on-top and bridge-bonded CO molecule at 470 and $380\,\mathrm{cm}^{-1}$, respectively. In EELS, the Pt–C and C–O stretch modes show comparable intensity, because of the fact that the intensity of the vibrational loss in the dipole scattering mechanism is proportional to the inverse of the loss energy.

Vibrational Relaxation

The IRAS spectrum is represented by

$$I(\hbar\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\hbar\omega} \frac{\langle Q(t)Q(0) \rangle}{\langle Q^2 \rangle} \qquad [8]$$

where $Q(t)$ is the Heisenberg representation of the normal mode coordinate Q of the vibrational degree of freedom, and the brackets indicate the ensemble average of the system. In the absence of the inhomogeneous broadening associated with the structural disorder of an admolecular layer, one may assume a damped oscillator with a frequency of $\hbar\Omega$,and write

$$\begin{aligned} Q(t) &= Q_0 \exp[\mathrm{i}\hbar\Omega t - \Gamma|t|] \\ Q_0 &= 1/(2m^*\hbar\Omega)^{1/2} \end{aligned} \qquad [9]$$

which gives the Lorentzian lineshape of IRAS,

$$I(\hbar\omega) = \frac{1}{\pi} \frac{\Gamma}{(\hbar\omega - \hbar\Omega)^2 + \Gamma^2} \qquad [10]$$

where the half-width-at-half-maximum is expressed as a sum of T_1 (energy or population relaxation time caused by the energy transfer to a heat bath of the system) and T_2^* (pure dephasing time associated with a random modulation of the vibrational frequency due to anharmonic coupling to low-frequency modes);

$$\Gamma = \frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_2^*} \qquad [11]$$

The possible candidate of T_1 for adsorbates on metals is an excitation of EHPs as well as surface and bulk phonons in the substrate. The dephasing time T_2

is defined with respect to the decay of the amplitude of the vibration, and T_1, with respect to the decay of the population. This accounts for the factor 2 in the first term. The possible mechanism responsible for T_1-relaxation crucially depends on the vibrational frequency. For adsorbates whose vibrational frequency is higher than the bulk and/or surface phonons in the substrate, the energy relaxation requires multiple phonon emissions in order to satisfy the energy conservation. For a typical example of a stretching vibration of hydrogen on an Si surface, T_1 (of an order on nanosecond) due to multiple phonon emission increases with temperature. In contrast, on metal surfaces damping through excitation of an EHP in the substrate shortens T_1 down to a few picoseconds. The relative contributions of T_1, T_2^*, and inhomogeneous broadening to the line width can be estimated from the coverage and/or temperature dependence of the IRAS spectrum. For example, T_2^* becomes infinite at very low temperatures where there are no thermal fluctuations to produce pure dephasing. The effect of pure dephasing manifests itself not only in the change of the line width, but also the frequency shift with temperature.

The Lorentzian lineshape of eqn [10] is modified to an asymmetric one with a tail at the lower or higher energy side, as a result of vibrational energy damping due to EHP excitation in metals, and/or by the dipole–dipole coupling between neighboring molecules. In the case of EHP damping, it is represented by the Fano-type lineshape, characteristic to any optical transition of the discrete–continuum system. It is also remarked that in the presence of the structural disorder in the incomplete monolayer of adsorbates, the Gaussian rather than Lorentzian lineshape fits better to the IRAS spectrum.

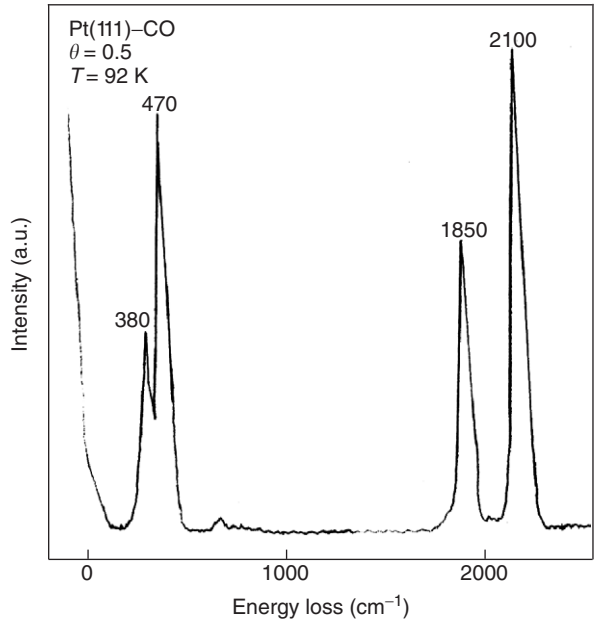


Figure 9 Electron energy loss spectrum of the ordered $c(4 \times 2)$ CO adlayer at the coverage of $\theta = 0.5$ on Pt(1 1 1) at 92 K. (Reproduced from Steiniger H, Lehward S, and Ibach H (1982) On the adsorption of CO on Pt(1 1 1). *Surface Science* 123: 264–282, with permission from Elsevier.)

Nonlinear Surface Vibrational Spectroscopy – Sum-Frequency Generation

Optical second harmonic generation (SHG) and sum-frequency generation (SFG) are the second-order nonlinear process, which forbidden in a medium with inversion symmetry. They arise from a polarization that depends quadratically on the incoming electric fields. **Figure 10** schematically illustrates the SFG process as well as a simple energy diagram for a vibrational excitation. Here a matching of the wave vector components along the surfaces $k_{||,\mathrm{sum}} = k_{||,\mathrm{ir}} + k_{||,\mathrm{vis}}$ determines the direction of the sum-frequency output. The vibrational SFG process is a multi-photon mixing process in which an IR wave of $\hbar\omega_{\mathrm{ir}}$ mixes with a visible wave of $\hbar\omega_{\mathrm{vis}}$ to yield an output at $\hbar\omega_{\mathrm{sum}}$. By scanning the IR beam, the SFG signal exhibits a similar vibrational absorption spectrum observed by IRAS. The SFG surface vibrational

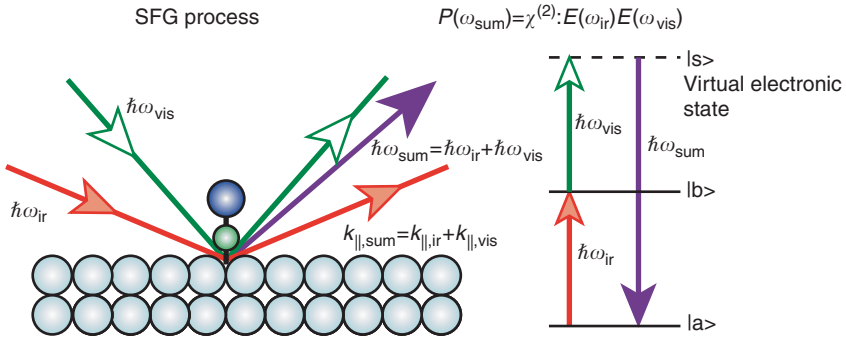


Figure 10 Elementary process of IR–visible vibrational SFG.

spectroscopy allows *in situ* studies of adsorbates with excellent spatial, spectral, and temporal resolution. It can be applied to any surfaces and interfaces accessible by light. Within the dipole approximation, the polarizability of a third-rank tensor is given by

$$P(\hbar\omega_{\text{sum}} = \hbar\omega_1 + \hbar\omega_2) = \chi^{(2)} : E(\hbar\omega_1) : E(\hbar\omega_2) \qquad [12]$$

where $\omega_1 = \omega_2$ for SHG, while $\omega_1 = \omega_{\text{ir}}$ and $\omega_2 = \omega_{\text{vis}}$ for SFG. If the medium has inversion symmetry, the fields $E(\hbar\omega)$ and $-E(\hbar\omega)$ induce $P(\hbar\omega_{\text{sum}})$ and $-P(\hbar\omega_{\text{sum}})$, respectively. This implies $\chi^{(2)} = 0$ in the case of a centrosymmetric medium. In other words, $\chi^{(2)}$ has a finite value at the surface where the inversion symmetry is broken. This is the basis for the surface sensitivity and specificity of SHG and SFG.

The second-order SFG nonlinear susceptibility $\chi^{(2)}$ at $\hbar\omega_{\text{sum}}$ is given by

$$[\chi]^{(2)}_{ijk} = \frac{\mu_k^{ba}}{\hbar\omega_{\text{ir}} - \hbar\Omega + \text{i}(1/T_2)} \times M_{ij}(\hbar\omega_{\text{sum}})(n_a - n_b) \qquad [13]$$

where $\mu_k^{ba} = \langle b|\mu_k|a\rangle$ is the vibrational transition dipole along the k-axis, and n_a and n_b are the populations in the ground and excited vibrational levels, respectively,

$$M_{ij}(\hbar\omega_{\text{sum}}) = \sum_s \left[\frac{\langle b|\mu_j|s\rangle\langle s|\mu_i|a\rangle}{\hbar\omega_{\text{sum}} + \hbar\omega_{as} + \text{i}/T_2^{sa}} - \frac{\langle b|\mu_i|s\rangle\langle s|\mu_j|a\rangle}{\hbar\omega_{\text{sum}} + \hbar\omega_{sb} + \text{i}/T_2^{sb}} \right] \qquad [14]$$

is proportional to the Raman scattering susceptibility of the vibrational transition, where the sum is over the virtual states $|s\rangle$ and $\hbar\omega_{\text{sum}}$ is assumed not to be resonant with any electronic transition. One may immediately notice that the vibrational mode must be both IR and Raman active. The matrix elements of $\chi^{(2)}$ are understood as a sequence of optical transitions starting from the vibrational ground state $|a\rangle$ to an excited state $|b\rangle$, induced by the IR photon with $\hbar\omega_{\text{ir}}$, followed by a second transition to intermediate virtual states $|s\rangle$ by the visible photon with $\hbar\omega_{\text{vis}}$. A photon with sum frequency $\hbar\omega_{\text{sum}}$ is then emitted upon de-excitation returning to the initial state. When the so-called nonresonant contributions $\chi^{(2)}_{\text{NR}}$ due to the off-resonant adsorbate or substrate response is included, the total SFG susceptibility is expressed as

$$\chi^{(2)}_{\text{tot}} = \chi^{(2)}_{\text{RES}} + \chi^{(2)}_{\text{NR}}, \quad \chi^{(2)}_{\text{NR}} = Ae^{\text{i}\phi} \qquad [15]$$

where A is the amplitude of the nonresonant susceptibility and ϕ its phase relative to the vibrational resonance. The vibrational SFG spectrum given by $|\chi^{(2)}_{\text{tot}}|^2$ is affected by $\chi^{(2)}_{\text{NR}}$. If it is negligible or the observed SFG spectrum is deconvoluted to eliminate the contribution of $\chi^{(2)}_{\text{NR}}$, the spectrum is essentially similar to the IR absorption spectrum. **Figure 11** compares the IR absorption spectrum (a) and SFG spectrum (b) of the ordered $c(2 \times 2)$ structure at the coverage of 0.5 of CO on Cu(1 0 0) at 100 K and 120 K, respectively. The IR lineshape is well reproduced by a Lorentzian (dashed line) with a tail at higher frequency due to a dipole–dipole interaction between the neighboring CO molecules. The absence

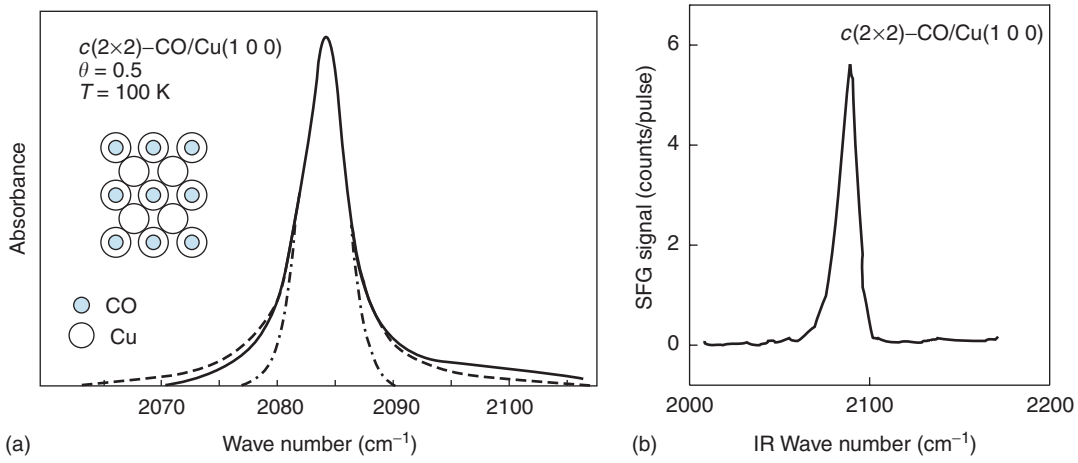


Figure 11 (a) IR absorption spectrum of the C–O stretch mode (dashed and dot-dashed curves are Lorentzian and Gaussian fit, respectively) and SFG spectrum of an ordered $c(2 \times 2)$ adlayer of CO at coverage $\theta = 0.5$ on Cu(1 0 0). (Reprinted figure with permission from Ryberg R (1985) Vibrational line shape of chemisorbed CO. *Physical Review B* 32: 2671–2673; © American Physical Society.) (b) SFG spectrum of the C–O stretch mode of $c(2 \times 2)$ CO/Cu (100) at 120 K. (Reproduced with permission from Morin M, Levinos NJ, and Harris AL (1992) Vibrational energy transfer of CO/Cu (100): nonadiabatic vibration/electron coupling. *The Journal of Chemical Physics* 96(5): 3950–3956; © American Institute of Physics.)

of the spectral change with temperature of 20–30 K reveals that the line width ($4.6\,cm^{-1}$) is predominantly determined by the vibrational energy relaxation time of $T_1 = 1.2$ ps due to EHP excitations in the substrate. On the other hand, the SFG spectrum observed using an IR pulse of 4 ps duration showed broader width than the IR absorption spectrum due to poor spectral resolution. For ω_{ir} and ω_{vis} or ω_{sum} near vibrational and electronic excitations, double (IR and visible)-resonant SFG occurs when the corresponding vibrational and electronic transitions are coupled. This reminds one of the surface enhanced Raman scattering of adsorbed molecules on metal surfaces, and enables both the vibrational and electronic properties of adsorbed molecules to be explored as a novel two-dimensional surface spectroscopy.

Pump–Probe SFG – Direct Determination of Dephasing Time

One of the important aspects of vibrational SFG spectroscopy lies in the fact that SFG response is proportional to the population difference $n_a - n_b$ between the ground and excited states. In addition to the SFG signal in the frequency domain, SFG measurement in a time-domain can be performed by introducing a time delay t_d between input pulses. The vibrational relaxation time of adsorbates on surfaces can be directly measured using different pulse sequences, as schematically shown in **Figure 12**. In the case of (a), an intense IR pump pulse (IR1) tuned to the vibrational resonance at a surface, saturates the fundamental $v = 0 \rightarrow 1$ transition. After a certain time delay t_d, a pair (IR2 and visible) of weak probe pulses generates a sum-frequency polarization proportional to the population difference $n_a(t_d) - n_b(t_d)$. The output at sum frequency decays as a function of t_d with a time constant determined by the population relaxation time T_1 of the excited state. This type of pump–probe SFG experiment was applied to a $c(2 \times 2)$ CO on Cu(1 0 0), and $T_1 = 2.0$ ps was obtained. This accounts for 60% of the line width of the IRAS spectrum shown in **Figure 11**.

A different type of transient SFG shown in (b) measures a dephasing time T_2, when t_d is set between IR1 and visible pulse without probe IR2. With this technique, the IR1 resonantly excites the vibrational polarization in a coherent manner, and this coherence decays due to dephasing as $\exp(-t/T_2)$. When the visible pulse, delayed by t_d, is less than or comparable to T_2 is incident on the surface, it probes the remaining coherence. From the decay of the SFG intensity as a function of t_d, T_2 can be deduced. This coherent decay is the so-called free induction decay

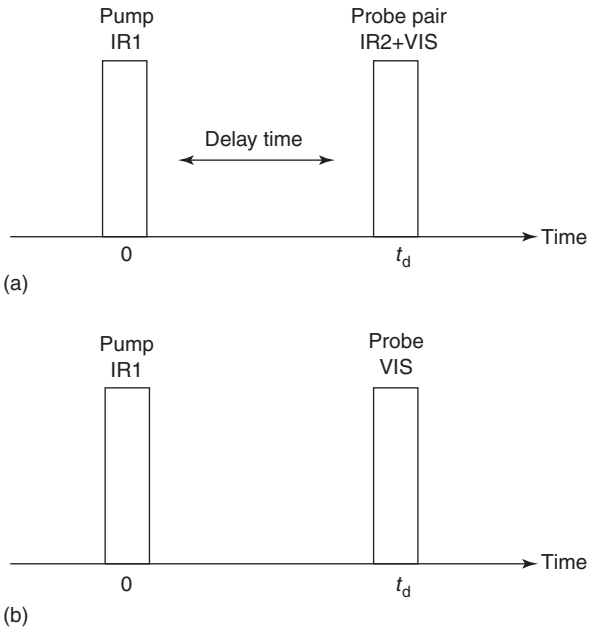


Figure 12 Pulse sequences of time-resolved SFG. (a) an intense IR pump pulse (IR1) tuned to the vibrational resonance at a surface, saturates the fundamental $v = 0 \rightarrow 1$ transition. After a certain time delay t_d, a pair (IR2 and visible) of weak probe pulses arrives at the surface. The output at sum frequency decays as a function of t_d with the population relaxation time T_1 of the excited state. (b) IR1 resonantly excites the vibrational polarization in a coherent manner, and this coherence decays due to dephasing as $\exp(-t/T_2)$. When the visible pulse, delayed by t_d, is less than or comparable to T_2 and is incident on the surface, it probes the remaining coherence. From the decay of the SFG signal, T_2 can be deduced. This coherent decay is the so-called FID.

(FID). In the absence of the coherent optical mixing, the transient SFG intensity $S(t_d)$ evaluated by the time integration of the polarization $|P(t,t_d)|^2$ decays with a time constant $-2/T_2$ as a function of t_d,

$$S(t_d) \propto e^{-2t_d/T_2} \qquad [16]$$

The coherent optical mixing effect occurs in the overlap between $E_{ir}(t)$ and $E_{vis}(t - t_d)$, and causes a shift of the maximum of the SFG intensity from $t_d = 0$. For negligible overlapping between the infrared and visible pulse, T_2 estimated from a decay slope of $S(t_d)$ does not depend on the pulse shape as far as their durations are sufficiently shorter than T_2. It is also remarked that the time resolution of SFG–FID is determined by the rising edge at the negative t_d. **Figure 13** depicts a characteristic behavior of $S(t_d)$, which explains the experimental result for the C–O stretch mode of CO/Ru (0 0 1). In this experiment, the ultrashort pump pulse (800 nm, 110 fs) was used for upconversion of the transient IR-polarization. $T_2 =$ 2 ps was deduced from the decay slopes of $S(t_d)$.

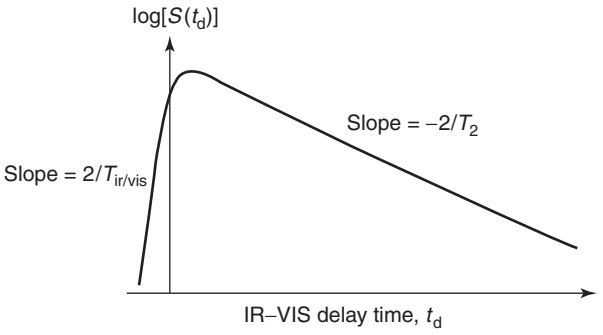


Figure 13 Overall feature of the logarithmic plot of the transient SFG intensity $S(t_d)$ as a function of a time delay between IR pump and visible probe pulse. The rising edge at a negative t_d depends on the pulse shape and determines the time resolution, while the slope at positive t_d gives $-2/T_2$ as far as $T_2 \gg T_{IR/VIS}$.

The same SFG–FID experiment at 350 K gave $T_2 = 1.2$ ps. Because of the temperature independent $T_1(\simeq 2\,\text{ps})$ due to EHP excitations in the substrate, the temperature dependence of T_2 originates from that of T_2^* due to anharmonic coupling with the low-frequency frustrated translational mode.

It should be emphasized that T_2 is reliably determined from a decay slope of the transient SFG intensity observed using sufficiently shorter IR and visible pulse than T_2. The line width of the time-resolved SFG spectra, however, depends on the delay time and is influenced by the pulse duration. When the time-resolved spectrum is observed with the pulse of duration comparable to or much shorter than T_2, no reliable T_2 determination is possible from the line width measurement. A similar phenomenon is also found in energy- and time-resolved 2PPE spectroscopy.

Perspectives

The study of electronic and vibrational dynamics at surfaces is a basis to gain deep insight into a wide variety of electronic and vibrational mediated surface reactions. In particular, nonlinear and time-resolved optical spectroscopies using ultrafast laser pulses continue to provide indispensable knowledge to ward complete understanding of microscopic characteristics of electronic and vibrational dynamics at surfaces. The recent advances in laser technology with pulse duration below several tens of femtoseconds are expected to open a novel way of quantum control of surface reactions.

See also: Core Photoemission; Interaction of Light and Matter; Lattice Dynamics: Vibrational Modes; Nonlinear Optics; Optical Absorption and Reflectance; Optical Properties of Materials; Optical Properties of Surface Layers Enhanced Raman Scattering; Photoelectron Spectromicroscopy; Polarizabilities; Surfaces and Interfaces, Electronic Structure of; Surfaces, Optical Properties of; Time-Resolved Optical Spectroscopies.

PACS: 68.47.De; 68.43.Pq; 73.20.−r; 73.20.At; 78.47.+p; 79.60.−i

Further Reading

Chabal YJ (1988) Surface infrared spectroscopy. *Surface Science Reports* 8: 211–357.

Dai H-L and Ho W (eds.) (1995) *Laser Spectroscopy and Photo-Chemistry on Metal Surfaces*. Singapore: World Scientific.

Fano U (1961) Effects of configuration interaction on intensities and phase shifts. *Physical Review* 124: 1866–1878.

Fauster T and Steinmann W (1995) Two-photon photoemission spectroscopy of image states. In: Halevi P (ed.) *Photonic Probes of Surfaces*, ch. 8, 350–411. Amsterdam: North-Holland.

Feuerbacher B, Fitton B, and Willis RF (eds.) (1977) *Photoemission from Surfaces*. New York: Wiley.

Haight R (1995) Electron dynamics at surfaces. *Surface Science Reports* 21: 275–325.

Hertel T, Knoesel E, Wolf M, and Ertl G (1996) Ultrafast electron dynamics at Cu(1 1 1): response of an electron gas to optical excitation. *Physical Review Letters* 76: 535–538.

Hess C, Wolf M, Roke S, and Bonn M (2002) Femtosecond time-resolved vibrational SFG spectroscopy of CO/Ru(0 0 1). *Surface Science* 502–503: 304–312.

Ibach H and Mills DL (1982) *Electron Energy Loss Spectroscopy and Surface Vibrations*. New York: Academic Press.

Morin M, Levinos NJ, and Harris AL (1992) Vibrational energy transfer of CO/Cu(1 0 0): Nonadiabatic vibration/electron coupling. *Journal of Chemical Physics* 96: 3950–3956.

Petek H and Ogawa S (1998) Femtosecond time-resolved two-photon photoemission studies of electron dynamics in metals. *Progress in Surface Science* 56: 239–310.

Rehr JJ and Albers RC (2000) Theoretical approaches to X-ray absorption fine structure. *Review of Modern Physics* 72: 621–654.

Shen YR (1989) Surface properties probed by second-harmonic and sum-frequency generation. *Nature* 337: 519–525.

Ueba H (1995) Optical spectroscopies of adsorbates on metal surfaces-from statics to dynamics. In: Halevi P (ed.) *Photonic Probes of Surfaces*, ch. 6, 350–411. Amsterdam: North-Holland.

Ueba H (1997) Vibrational relaxation and pump-probe spectroscopies of adsorbates on solid surfaces. *Progress in Surface Science* 55: 115–179.

Yoshinobu J (2004) Physical properties and chemical reactivity of the buckled dimers on Si(100). *Progress in Surface Science* 77: 37–70.

Tunneling Devices

D Dragoman, University of Bucharest, Bucharest, Romania
M Dragoman, National Institute for Research and Development in Microtechnology (IMT), Bucharest, Romania

Introduction

Tunneling is the phenomenon of propagation at constant energy of quantum wave functions across classically forbidden regions. In a classically forbidden region, the energy of the quantum particle is less than the potential energy so that the quantum wave function cannot penetrate the forbidden region unless its dimension is smaller than the decay length of the quantum wave function.

Tunneling plays a crucial role in almost any modern electronic and optoelectronic device. The development of tunneling devices has been intensified in the last decades and has paralleled the technological achievements that made possible the fabrication of ballistic structures. Constant-energy propagation, as that implied in the tunneling process, implies that (1) thermal excitation is negligible and (2) inelastic scattering processes are absent. Even if the first condition can be achieved at low temperatures, ballistic transport is indispensable in ensuring the second requirement.

A rough classification of tunneling devices distinguishes between single-barrier tunneling devices, double-barrier or resonant tunneling devices, superlattices, and single-electron transistors. This classification is mainly followed in the presentation of the unique features of tunneling devices and the focus is on electron tunneling.

Single-Barrier Tunneling Devices

Electron transport takes place along the direction of an applied electric field. The direction of electron propagation is denoted as x and one may consider that it is possible to separate the electron motion in the transverse (along y and z) and longitudinal (along x) directions, such that the x-dependent part of the envelope electron wave function $\Psi(x,y,z) = \psi(x)\varphi(y,z)$ satisfies in the ballistic collisionless transport regime the time-dependent Schrödinger equation

$$\left(-\frac{\hbar^2}{2}\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{m(x)}\frac{\mathrm{d}}{\mathrm{d}x} + V(x)\right)\psi(x) = E\psi(x)$$

Here m is the effective electron mass in the conduction band and V is the potential energy. Conduction electron tunneling across a single barrier is easily modeled considering the x-dependent potential energy as being predominantly due to conduction band discontinuities, which occur suddenly at the interface between two different materials and are constant in each material layer. (In general, space charge effects at the interface smoothen the discontinuity; they should be accounted for by a self-consistent solution of both Poisson and Schrödinger equations.) The tunneling structure is then composed from three regions in which both effective masses m_i and potentials V_i ($i = 1,2,3$) vary, as shown in **Figure 1**, such that V_1, $V_3 < E$ and $V_2 > E$. Under these conditions, in each layer the electron wave function can be expressed as a superposition of forward- and backward-propagating waves in the x-direction with wave number $k_i = \hbar^{-1}\sqrt{2m_i(E - V_i)}$:

$$\psi_i(x) = A_i \exp(\mathrm{i}k_i x) + B_i \exp(-\mathrm{i}k_i x)$$

The wave number $k_2 = -\mathrm{i}\gamma_2$ is purely imaginary in the barrier region for electron propagation, and takes real values in the other regions in which electron propagation is allowed. The value of the coefficients A_i, B_i is obtained by imposing the continuity condition for the wave function and $(\mathrm{d}\psi/\mathrm{d}x)/m$. If the barrier is situated between $x = 0$ and $x = L$, these conditions enable one to relate A_1, B_1 to $A_3 \exp(\mathrm{i}k_3 L)$, $B_3 \exp(-\mathrm{i}k_3 L)$ through a transfer matrix with elements M_{ij}, $i,j = 1,2$ such that

$$\begin{aligned} M_{11} &= (1/2)[(1 + v_3/v_1)\cosh(\gamma_2 L) \\ &\quad - \mathrm{i}(v_3/v_2 - v_2/v_1)\sinh(\gamma_2 L)] \\ M_{21} &= (1/2)[(1 - v_3/v_1)\cosh(\gamma_2 L) \\ &\quad - \mathrm{i}(v_3/v_2 + v_2/v_1)\sinh(\gamma_2 L)] \\ M_{22} &= M_{11}^*, \quad M_{12} = M_{21}^* \end{aligned}$$

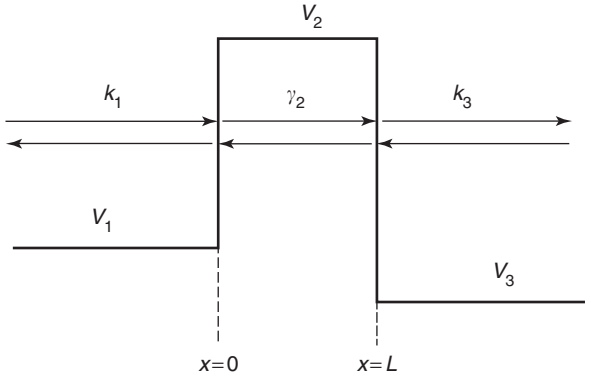


Figure 1 Geometry of a single-barrier tunneling phenomenon.

with $v_1 = k_1/m_1$, $v_2 = \gamma_2/m_2$, $v_3 = k_3/m_3$. The transmission probability of the barrier, defined as $T = v_3|A_3|^2/(v_1|A_1|^2) = v_3/(v_1|M_{11}|^2)$, is

$$T = \frac{4v_1v_3}{(v_1+v_3)^2 + [(v_1^2+v_2^2)(v_2^2+v_3^2)/v_2^2]\sinh^2(\gamma_2 L)}$$

From this expression, it follows that for $\gamma_2 L \gg 1$, the transmission probability depends exponentially on the barrier width L, $T \propto \exp(-2\gamma_2 L)$, or at low temperatures where all electrons that participate in transport have energies close to the Fermi energy E_F, $T \propto \exp\left[-2L\sqrt{2m_2(V_2 - E_F)}/\hbar\right]$. The same exponential factor appears in the expression of the current density through the barrier, which is proportional to the electron transmission probability at low temperatures. Other dependencies of the current density on the height and width of the potential barrier are found for finite temperatures or for different barrier shapes (a triangular barrier is a better description of electron tunneling at high applied voltages).

The current across a barrier connected to two electron reservoirs (contacts) at quasi-Fermi energies E_{F1} and E_{F2} is given by

$$I = \frac{2e}{h}\int T(E)[f_1(E) - f_2(E)]\mathrm{d}E$$

where $f_i(E) = 1/[1 + \exp[(E - E_{Fi})/k_B T]$ is the Fermi–Dirac distribution function at the contacts, which can be approximated at low temperatures with $f_i(E) \propto \vartheta(E_{Fi} - E)$, with ϑ being the step function. $T(E)$ in the expression above should include, if necessary, the contribution of all input and output modes of a ballistic conductor. This expression of the current is valid irrespective of the details of the region characterized by the transmission probability T, that is, it is valid for both single-barrier and multibarrier tunneling.

The electron transmission through a single-barrier model is relevant to vacuum microelectronics as well as for the scanning tunneling microscopy technique. The recently developed vacuum microelectronics studies micrometer-scale devices that operate with ballistic electrons in vacuum. These devices have the potential to operate at higher frequencies, higher power, and in a wider temperature range, even under irradiation conditions, than devices based on semiconductor technology since in vacuum, electrons travel faster and with less energy dissipation. Prompted by these advantages, vacuum microelectronic devices include flat-panel field emission displays, in which electrons tunnel from a solid surface into vacuum, electron sources for microscopes, miniaturized microwave power amplifiers, and X-ray generators. Scanning tunneling microscopy, on the other hand, is a method of mapping solid surfaces with atomic resolution. In this case, electrons tunnel from the last atom of the tip apex of the probe to single atoms of the surface under study. Scanning tunneling microscopy is also a spectroscopy method for single molecules on surfaces and a powerful technique of manipulating the surface structure.

Resonant Tunneling Devices

Resonant tunneling occurs in structures consisting of two thin barriers for electron propagation separated by a quantum well region with width L in which the electron is free to propagate with a wave vector k. The geometry of the structure is represented in **Figure 2**. The matrix formulation briefly discussed in the previous section can be employed in this case also if the potential energy has a step-like variation. The total transmission matrix M with elements M_{ij}, $i, j = 1, 2$, is obtained by multiplying the transmission matrices of the left barrier, M_l, of free propagation across the quantum well, and of the right barrier M_r. The transmission probability T of the whole structure is expressed, as above, in terms of M_{11}, that takes the expression

$$M_{11} = M_{l,11}M_{r,11}\exp(-\mathrm{i}kL) + M_{l,12}M_{r,21}\exp(\mathrm{i}kL)$$

since the free propagation matrix is diagonal with elements $\exp(-\mathrm{i}kL)$ and $\exp(\mathrm{i}kL)$. It can be shown that

$$T = \frac{T_l T_r}{(1 - \sqrt{R_l R_r})^2 + 4\sqrt{R_l R_r}\cos^2\theta}$$

with

$$\theta = kL + (\arg M_{l,12} + \arg M_{r,21} - \arg M_{l,11} - \arg M_{r,11})/2$$

where T_l, T_r are the transmission probabilities through the left and right barrier individually, and R_l, R_r the corresponding reflection probabilities. Even if T_l and T_r are small so that $1 - \sqrt{R_l R_r} \cong (T_l + T_r)/2$, the

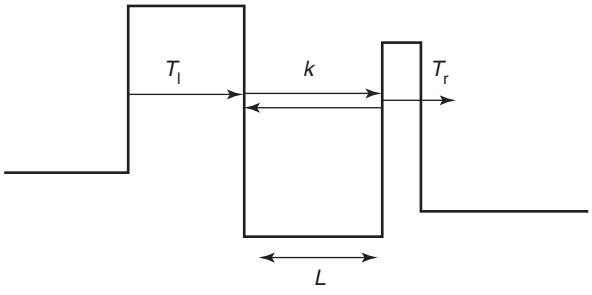


Figure 2 Geometry of the double-barrier tunneling phenomenon.

total transmission probability can still achieve large values at resonance, that is, if $\theta = (2n+1)\pi/2$ with n integer, for which $T_{res} = 4T_l T_r/(T_l + T_r)^2$. The resonant value of the transmission probability is unity if $T_l = T_r$, and can be approximated with $4T_{min}/T_{max}$ when T_l and T_r are very dissimilar, with $T_{min} = \min(T_l, T_r)$ and $T_{max} = \max(T_l, T_r)$. At resonance, that is, when the energy of the incident electrons equals one of the resonance energy values, the electrons are not only transmitted with high probability across the structure, but they are also transmitted in a shorter time than in off-resonance conditions. For this reason, resonant tunneling devices are also ultrafast devices.

Close to resonance, the transmission probability can be approximated with

$$T(E) = \frac{\Gamma_l \Gamma_r}{(\Gamma_l + \Gamma_r)^2/4 + (E - E_{res})^2}$$

where E_{res} is the resonance energy for which $\theta(E) = (2n+1)\pi/2$, and $\Gamma_l = (dE/d\theta)T_l/2$, $\Gamma_r = (dE/d\theta)T_r/2$ (divided by $\hbar$) represent the rates at which an electron placed in the well of width L leaks out through the left and right barrier, respectively. Near resonance, the transmission probability is very sensitive to the electron energy or to the position of the resonant energy level. The ease with which this last parameter can be controlled by applying a voltage across the double-barrier structure led to the development of a host of devices based on resonant tunneling.

Coherent versus Incoherent Resonant Tunneling

In the calculation of the transmission probability of the resonant tunneling structure, the coherent transport has been assumed, that is the electron wave function is assumed to be transmitted from left to right in a single quantum mechanical process. This assumption is justified if the time spent by an electron in the resonant state (the eigenstate lifetime) is much smaller than the scattering time. If this is not the case, sequential tunneling occurs, in which the electron first tunnels into the well, loses memory of its phase, and then tunnels out from the well. Despite the differences between coherent and sequential resonant tunneling, the expression of the net current flow near resonance is the same for the two cases.

Resonant Tunneling Diodes

One of the first applications of resonant tunneling structures was the resonant tunneling diode (RTD),

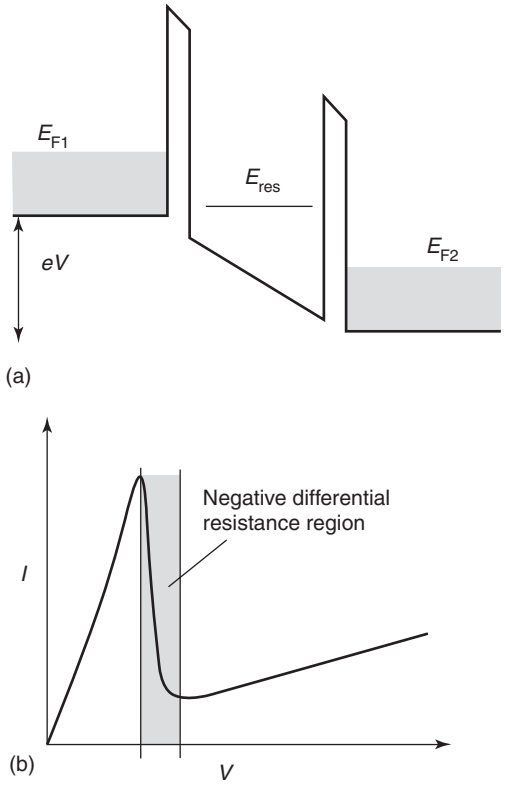


Figure 3 (a) Schematic representation of a resonant tunneling diode, and (b) its I–V characteristic.

which is a biased double-barrier structure that has only one resonance energy level on which electrons from the left contact (emitter) can tunnel (see **Figure 3a**). The main characteristic of the RTD is the occurrence of a negative differential resistance region, as shown in **Figure 3b**. Although the electrons in the emitter region have energies between the bottom of the conduction band and the Fermi energy E_{F1}, the resonance level in the well acts as a filter that allows further propagation of only those electrons with energy E_{res}; when the applied bias is sufficiently high for the resonance energy level to fall below the conduction band edge in the left emitter region, no electrons can tunnel into the right contact (collector) with a Fermi energy E_{F2}, and the current is expected to drop. Therefore, the I–V characteristic exhibits a negative differential resistance region. It is important to note that the RTD can work at room temperature, unlike most mesoscopic devices that function well only at low temperatures.

Optoelectronic Devices Based on Resonant Tunneling

Optoelectronic devices based on resonant tunneling incorporate mesoscopic structures such as quantum

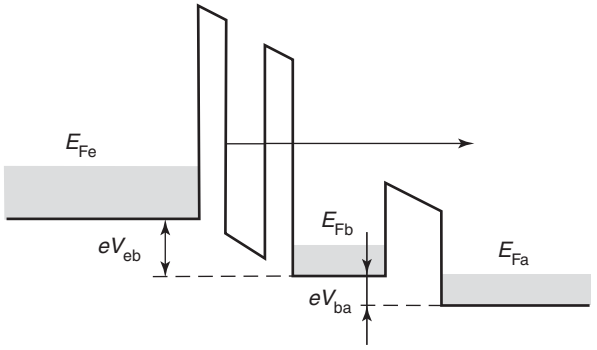


Figure 4 Schematic representation of the resonant tunneling injection hot electron laser.

wells, quantum wires, or quantum dots, which are used to generate, modulate, or detect electromagnetic radiation. In these devices, resonant tunneling is usually employed to enhance the high-speed response since the transfer rate of electrons through resonant tunneling structures is much faster than the drift and diffusion processes. Moreover, since resonant tunneling acts as an energy filter for the incident electrons, it can be used to remove the carrier speed limitations due to carrier thermalization.

In lasers, for example, based on the enhancement of the stimulated emission of radiation in an active mesoscopic medium placed inside a resonant cavity for the electromagnetic radiation, the modulation bandwidth can be improved by the injection of carriers in the active region through resonant tunneling. This tunneling injection laser is a hybrid design between quantum well tunneling lasers and active media quantum well lasers.

In more sophisticated tunneling injection lasers, as in that represented in **Figure 4**, it is possible to control the energy E of (hot) electrons in the active region through two voltages V_{eb} and V_{ba} applied, respectively, between the emitter and base regions, and between the base and active regions of the structure. The base and active regions are separated by a barrier that acts as a hot electron launcher. The Fermi energy levels in the emitter, base, and active regions are denoted in **Figure 4** by E_{Fe}, E_{Fb}, and E_{Fa}, respectively. In this device, called resonant tunneling injection hot electron laser, ultrafast gain switching is caused by the change in the injection energy rather than by carrier injection rate variations.

High modulation speeds are achieved by employing electron tunneling across the spacer region that separates the electron reservoir from the modulating structure, for example, in the device known as barrier-reservoir and quantum well electron transfer modulator (**Figure 5**). Here, however, the position of the resonant energy level E_{res} in the quantum well

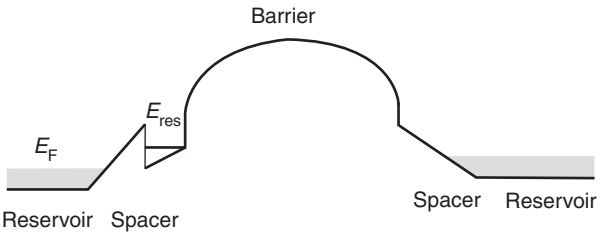


Figure 5 Schematic conduction band diagram of the barrier-reservoir and quantum well electron transfer modulator.

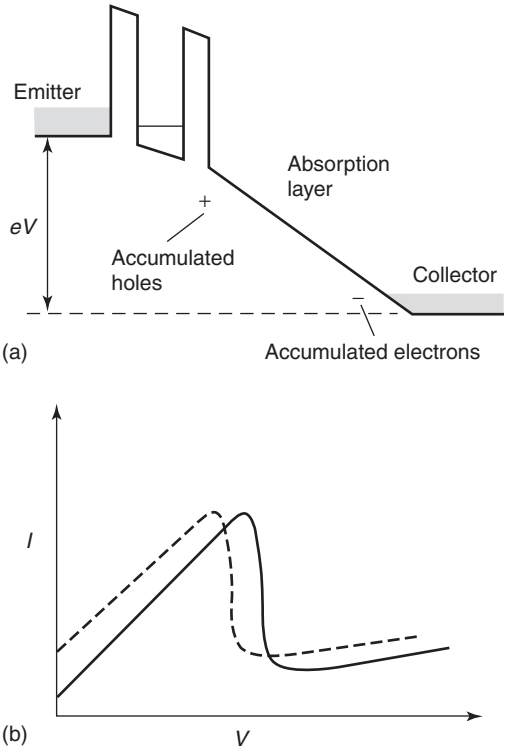


Figure 6 (a) Schematic conduction band diagram of the resonant tunneling photodetector, and (b) the I–V characteristic in the absence (—) and presence (- - - -) of light.

relative to the Fermi level E_F of the left electron reservoir is also essential. When an applied voltage drops the resonant energy level from above E_F (when it is empty) to below E_F (when it is filled), the absorption coefficient and refractive index for an incident electromagnetic radiation spectrum change, proportional to the applied voltage.

Resonant tunneling photodetectors have also been developed. They consist (see **Figure 6a**) of a resonant tunneling structure followed by a thick layer that absorbs the incident electromagnetic radiation. The photogenerated electrons and holes are separated spatially by the external voltage applied on the structure, which drops almost entirely on the absorption layer, and accumulates at the collector and the resonant barriers of the RTD, respectively.

Consequently, the applied electric field is partially screened and the *I–V* characteristic of the resonant tunneling structure, which exhibits a negative differential resistance region, is shifted to lower voltages in the presence of light absorption, as shown in **Figure 6b**. The responsivity of the photodetector is extremely high if biased near the peak voltage.

Resonant Tunneling Devices Based on Coupled Quantum Wells

A quantum resonant structure consists of a quantum well surrounded by thin barrier layers. What happens when another quantum well and another thin barrier is added to the structure? Depending on the height and width of the common barrier, the electrons confined in one quantum well can interact or not with the electrons confined in the other. The first case corresponds to coupled quantum wells, in which the electron wave function is extended throughout the entire structure, whereas the second case corresponds to a succession of noninteracting quantum wells, the electrons localized in one well being able to propagate across the structure only through sequential tunneling between adjacent wells.

In two identical coupled quantum wells, the energy degeneracy of the resonant levels in individual quantum wells is removed, the electron wave function splitting into a symmetric and an antisymmetric part, Ψ_{sym} and $\Psi_{antisym}$, respectively, as depicted in **Figure 7**. Nonidentical coupled wells, with different energy levels in the individual wells, can be brought to resonance by applying a bias that aligns two energy levels in the quantum wells. In this case, the electron wave function, which is initially confined in one well or another becomes delocalized and splits into a symmetric and an antisymmetric part. Coupled quantum wells have hosts of applications in electronics, optoelectronics, and even quantum computing.

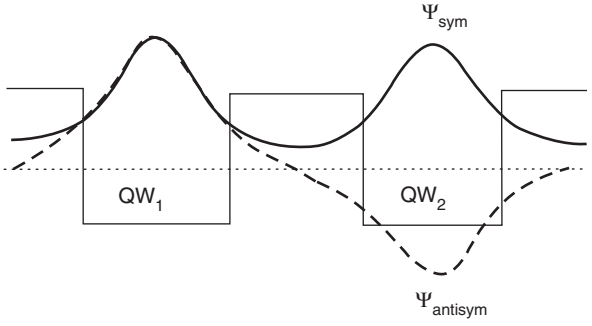


Figure 7 Splitting of the electron wave function into a symmetric and an antisymmetric part, when two identical quantum wells, QW_1 and QW_2, are brought in close proximity.

Quantum Well Tunneling Lasers

The resonant tunneling phenomenon is essential in lasers based on inter-sub-band transitions, where the radiation is caused by transitions between two resonant levels, 1 and 2, with populations N_i, $i = 1,2$, in an active quantum well in either the conduction or the valence band. In this case (see **Figure 8**), resonant tunneling assures the population inversion between the resonant levels E_1 and E_2 needed to enhance the stimulated emission of radiation. The condition $N_2 > N_1$ can only be achieved if the conduction electrons on the lower energy level can leak out from the quantum well through a resonant channel at the same energy level as E_1. Population inversion is thus created by resonant tunneling into an adjacent, coupled quantum well, which has a discrete energy level aligned to E_1. Such quantum tunneling lasers involve only the resonant transport of electrons if both active and coupled quantum wells are in the conduction band, but can also involve transport of both electrons and holes when the coupled quantum wells are situated in different bands. Often, electron injection into the active quantum well is also performed through resonant tunneling for improved efficiency and speed. Cascades of quantum well tunneling lasers can be used to enhance the power of the emitted radiation. Quantum well tunneling lasers are the future of low-energy high-wavelength radiation sources.

Modulators Based on Coupled Quantum Wells

The delocalization of electron wave function in nonidentical coupled quantum wells brought at resonance can be used to modulate an incident electromagnetic radiation. The modulation principle is based on the change of the binding energy of the excitons that can form in the system, and consequently, of the absorption coefficient and refractive index. In this type of modulator, presented in **Figure 9**, the lowest energy electron and hole states are confined in the first and second well of the conduction and valence bands,

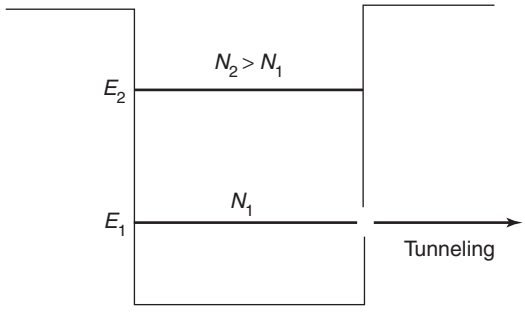


Figure 8 Population inversion mechanism in quantum well tunneling lasers.

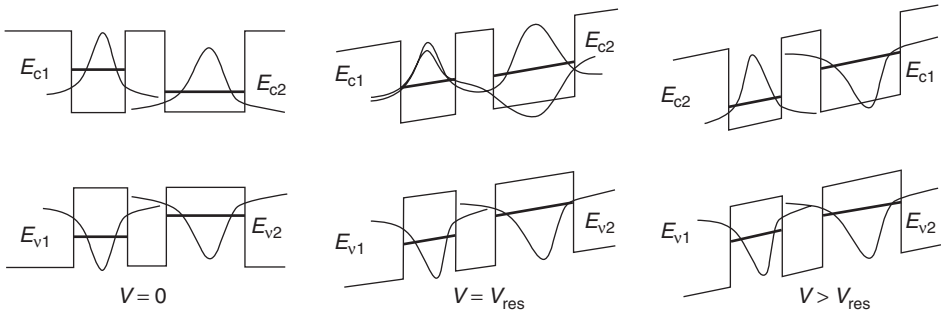


Figure 9 Dependence of the electron and hole wave functions of the applied bias in a modulator based on coupled quantum wells.

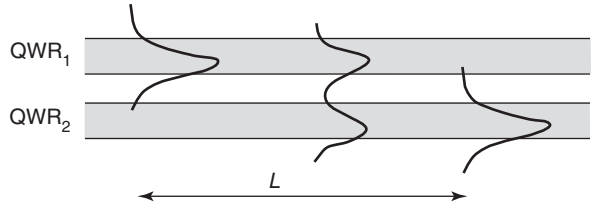


Figure 10 Evolution, in a structure consisting of coupled quantum wires, of an electron wave function initially localized in one wire.

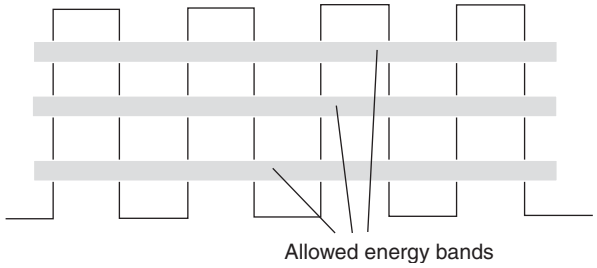


Figure 11 Conduction electron allowed and forbidden energy bands in the periodic potential of a superlattice.

respectively, when no potential is applied on the structure. In this case, excitons (bound electron–hole states) form between the electron and hole states confined in the same well, that is, between E_{c1} and E_{v1}, and E_{c2} and E_{v2}, respectively. At resonance, when the electron states are brought at the same level by an applied bias $V = V_{res}$, the electron wave function becomes delocalized, a further increase in bias inducing an interchange in the confined electron wave functions compared to the no-bias case. Assuming that the hole states in the quantum wells do not couple, the final exciton states and their respective binding energies are different from the no-bias case, since they form between the same hole levels but interchanged electron levels, that is, they form between E_{c2} and E_{v1}, and E_{c1} and E_{v2}, respectively.

Coupled Quantum Wires for Quantum Computing Applications

When two identical quantum wires, QWR_1 and QWR_2, are coupled, an electron wave function localized initially in one wire becomes confined in the other after a certain propagation length L, as shown in **Figure 10**; this situation is similar to a coupler between two parallel optical waveguides. It is then possible to identify the location of the electron wave function with one of the possible logic values 0 or 1, a superposition of the two quantum bit states occurring for propagation lengths smaller than L. Quantum operations can then be implemented by modulating the height of the common barrier by an applied bias.

Superlattices

When multiple identical quantum wells are coupled, the conduction electrons sense a periodically varying potential with a periodicity that is no longer that on the atomic scale, but on a mesoscopic scale. The periodicity of the potential energy imposes (as in bulk materials, where the periodicity is that of the lattice) the development of permitted and forbidden electron energy bands (see **Figure 11**), with positions and widths that depend on the exact form of the periodic potential and that can be controlled at will using advanced semiconductor techniques. This artificial lattice is called a superlattice. The possibility of engineering the position and width of conduction electron energy bands as well as the dispersion relation turn superlattices into extremely useful devices, with particular applications in optoelectronics, for electromagnetic fields with wavelengths comparable to the superlattice period. Then, the allowed and forbidden energy bands also develop for electromagnetic fields, similar to the allowed and forbidden energy bands of electrons in bulk materials.

Single-Electron Devices

Unlike resonant tunneling devices that are based on the discrete spectrum of resonant energy levels in a

well connected to electron reservoirs through thin barriers, single-electron devices are based on the discrete nature of electric charge that can be transferred from a conducting island (quantum dot or metallic cluster) connected through thin barriers to electron reservoirs. These devices are based on the Coulomb blockade phenomenon, which predicts the opening of a Coulomb gap in the tunneling density of states. Because of the Coulomb interaction between electrons in the island, tunneling of one electron requires an energy $e^2/2C$, where C is the capacitance between the island and the environment. This energy can be interpreted as an activation energy or, equivalently, as the opening of an energy gap of width e^2/C because, not only electrons at the Fermi level need an extra energy $e^2/2C$ to tunnel into the island but so do the holes. The contribution to this charging effect of energy quantization in ballistic islands renders the theory of single-electron tunneling more difficult.

Single-Electron Transistor

The single-electron transistor, depicted schematically in **Figure 12**, is the best-known single-electron device. It consists of a gated quantum dot or metallic cluster (generically, an island) separated by tunnel junctions from leads, and is characterized by an oscillating behavior of the conductance with density. The period of oscillations corresponds to the addition of one electron to the island. If the tunnel junctions that separate the gated island from the environment are modeled as a parallel combination of a tunneling resistance R_i and a capacitance C_i, $i = 1, 2$, the electrostatic energy of the gated island with N electrons and charge $Q = Ne$ is

$$E = -QV_gC_g/C + Q^2/2C - eV(N_1C_2 + N_2C_1)/C$$

or, to within an additive constant

$$E = (Q - Q_0)^2/2C - eV(N_1C_2 + N_2C_1)/C$$

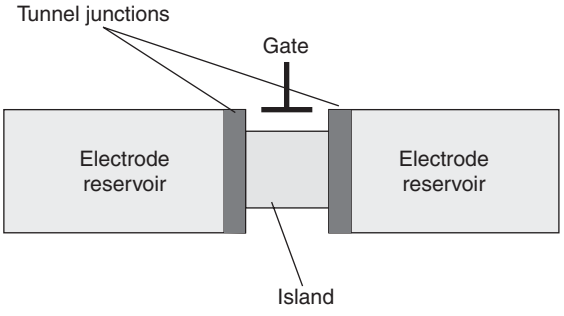


Figure 12 Schematic representation of a single-electron transistor.

where V_g is the gate voltage, V is the voltage applied between the leads, C_g is the island-gate capacitance, the total island capacitance is $C = C_1 + C_2 + C_g$, N_i is the number of electrons passed through the tunnel barrier i so that $N = N_1 - N_2$, and $Q_0 = C_gV_g$ is the external charge. Energy minimization allows only discrete values of energy for a given Q_0 due to charge quantization. At $V = 0$, when $Q_0 = Ne$ with N an integer, increasing or decreasing N by 1 requires an energy difference $e^2/2C$ and hence the opening of a Coulomb gap, whereas no energy gap in the tunneling density of states appears when $Q_0 = (N + 1/2)e$, since the states with $Q = Ne$ and $Q = (N + 1)e$ are degenerate and the charge fluctuates between them even at zero temperature. The occurrence of the energy gap at all gate voltages except those for which $Q_0 = (N + 1/2)e$ leads to the appearance of periodic sharp peaks in the conductance value at low temperatures. The spacing between the peaks is given by $\Delta V_g = e/C_g$, voltage difference that is exactly compensated for by the tunneling into or from the island of one electron.

These considerations are valid for metallic clusters isolated by thin barriers from external leads. In semiconductor quantum dots, the discretization of the energy levels inside the dots imposed by size constrictions must be accounted for in addition to charge discretization. In this case, electron tunneling implies the matching of the Fermi energy in the leads to a discrete energy level outside the Coulomb gap, both the amplitude and the position of the conductance peaks depending on the discrete levels of the dot. The gate voltage spacing ΔV_{gN} between the $(N-1)$th and the Nth peaks of the conductance increases from the e/C_g value in metallic clusters with $\Delta E_N/e : e\Delta V_{gN} = e^2/C_g + \Delta E_N$, where ΔE_N is the energy spacing of the levels of the structure; measurement of the position of conductance peaks can serve in this case as a spectrometric method for determining the discrete energy levels of the island. The amplitude of the conductance peaks depends on the discrete levels in the dot since it depends on the tunneling matrix elements, which can have different values for different energy levels in the dot.

Single-electron transistors are used as supersensitive electrometers since for applied voltages V slightly above the Coulomb blockade threshold, the current is extremely sensitive to the gate voltage, being able to measure sub-single-electron variations in the external charge. This technique of detecting subelectron charge variations has been recently proposed as the reading method for the results of quantum computations performed with solid-state Kane computers.

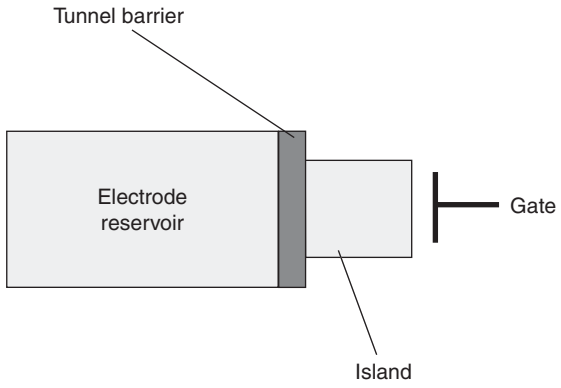


Figure 13 Schematic representation of a single-electron box.

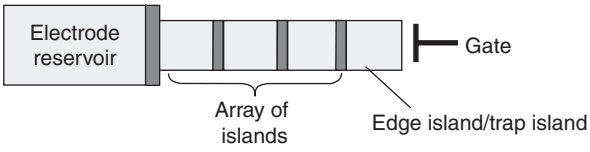


Figure 14 Schematic representation of a single-electron trap.

Other Single-Electron Devices

The simplest single-electron device is the single-electron box, which consists of a gated island separated from an electron source by a single tunnel barrier (**Figure 13**). The electrostatic energy of the box has the same expression as for the single-electron transistor but with $V = 0$, the charge of the island Q being a step-like function of Q_0, that is, of the gate voltage, the distance between neighboring steps being $\Delta Q_0 = e$. This Coulomb staircase is observable only at low temperatures; otherwise, it is smeared out by thermal fluctuations.

The single-electron trap is a generalization of the single-electron box in which the single tunnel junction is replaced by a linear array of islands separated by tunnel junctions (see **Figure 14**). Unlike single-electron boxes, where the number of electrons in the box is fixed by the gate voltage, single-electron traps may be in one of the several charged states of the edge island within certain ranges of applied gate voltages. This device is characterized by multistability, the charge state of the edge island (the trap) depending on its prehistory.

A single-electron turnstile, represented in **Figure 15**, is a linear array of islands connected through tunneling junctions to electron reservoirs, only one element of the array being gated. When no voltage is applied between the electron reservoirs, the device acts as a single-electron trap, whereas for an applied voltage an electron is always picked up from the emitter when the gate voltage V_g increases, and is delivered to the collector when V_g decreases. For a periodically varying gate voltage, an electron is transferred from emitter to collector for each period. When each island in the linear array is gated and phase-shifted RF waveforms are applied on each gate electrode, the device is called single-electron pump (see **Figure 16**). No voltage needs to be applied in this case between the electron reservoirs, the electron being transferred in a direction determined by the running wave of the electric potential.

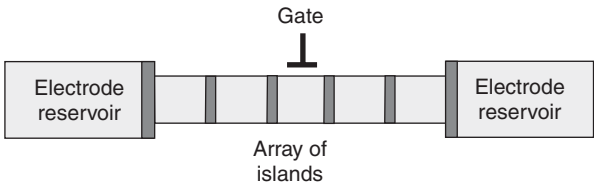


Figure 15 Schematic representation of a single-electron turnstile.

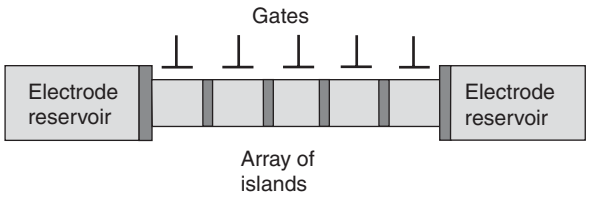


Figure 16 Schematic representation of a single-electron pump.

Single-electron tunneling devices have been proposed as standards for DC currents (especially single-electron turnstiles and pumps), temperature and resistance, for the detection of infrared radiation, and have interesting prospects for digital applications. Novel features and new applications of single-electron devices are expected to occur in systems with superconducting electrodes.

See also: Ballistic Transport; Effective Masses; Meso- and Nanostructures; Nanostructures, Electronic Structure of.

PACS: 85.30.Mn; 85.35.Be; 85.35.–z; 85.45.–w; 73.23.Hk; 73.23.–b; 7.79.Cz

Further Reading

Datta S (1997) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.

Dragoman D and Dragoman M (1999) *Advanced Optoelectronic Devices*. Berlin: Springer.

Ferry DK and Goodnick SM (1997) *Transport in Nanostructures*. Cambridge: Cambridge University Press.

Kastner MA (1992) The single-electron transistor. *Reviews of Modern Physics* 64: 849–858.
Likharev KK (1999) Single-electron devices and their application. *Proceedings IEEE* 87: 606–632.
Meyer E, Hug HJ, and Bennewitz R (2004) *Scanning Probe Microscopy*. Berlin: Springer.
Zhu W (ed.) (2001) *Vacuum Microelectronics*. New York: Wiley.

Nomenclature

C	capacitance between the island and the environment
C_g	island-gate capacitance
e	electron charge
E	energy
E_F	Fermi energy
E_{res}	resonant energy level
E_{ci}	bound state i in the conduction band
E_{vi}	bound state i in the valence band
$f(E)$	Fermi–Dirac distribution function
I	current
k	wave number
L	length
m	effective electron mass in the conduction band
M_{ij}	transfer matrix elements
N	population of an energy level
N_i	number of electrons that pass through tunnel barrier i
Q	island charge
Q_0	external charge
R	reflection probability
R_i	tunneling resistance of junction i
T	transmission probability
V	potential energy; also applied voltage
V_g	gate voltage
ΔV_{gN}	gate voltage spacing between the $(N-1)$th and the Nth peaks of the conductance
Γ	tunneling rate
ϑ	step function
$\Psi(x, y, z)$	electron wave function

V

Valence Photoemission

G Margaritondo, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

The basic features of this technique are discussed, beginning with its historical development. Angle-resolved (band mapping) and angle-integrated techniques are analyzed as well as a series of experimental approaches – photon polarization studies, experiments as a function of the photon energy, photoemission resonances and techniques affected by low signal level such as time-resolved photoemission, two-photon photoemission spectromicroscopy, and spin-polarized photoemission – requiring the use of powerful photon sources, such as undulators or ultraviolet lasers. Finally, high-resolution photoemission studies of superconductors and of other correlated systems are briefly discussed.

Historical Background: The Photoelectric Effect

The photoelectric effect, discovered by Hertz in 1887, is the basis of some of the most powerful techniques capable of analyzing the electronic structure of condensed-matter systems – the techniques collectively labeled as "photoemission spectroscopy." Even before the development of such techniques, the photoelectric effect had a fundamental (and often not correctly understood) impact on the history of modern science.

The main milestones after Hertz's work were Thompson's 1897 discovery of the electron, Einstein's 1905 hypothesis on the photon, and Millikan's 1918 reluctant experimental validation of Einstein's hypothesis. Einstein derived the photon hypothesis with a purely thermodynamic argument and used it to predict the frequency threshold in the photoelectric effect, which was not experimentally verified until the late 1910s.

The photoelectric effect provided, in this way, an experimental basis for quantum mechanics. This is undeniably a very significant contribution to the development of modern physics. The next scientific impact of the photoelectric effect – its practical use in spectroscopy – came almost 40 years later.

The reason for this long delay is the surface sensitivity of the photoelectric effect. In order to become a photoelectron, an electron in a solid must first absorb a photon to increase its energy above the photoelectric threshold. Since the threshold is of the order of several electronvolts, the photon must be in the X-ray or ultraviolet spectral range. Photons of this kind can penetrate quite deep in the solid before being absorbed. On the other hand, excited electrons can travel only over a very short distance before losing energy.

As a consequence, photoelectrons originate only from a thin slab near the surface of the sample. They carry information on the electronic structure of that slab, which typically consists of a very few atomic planes. If the surface is contaminated, the information is spurious. A suitably clean solid surface can be obtained by several different methods such as cleaving, scraping, and *in situ* growth. However, most clean surfaces do not stay clean unless under ultrahigh vacuum (pressure $< 10^{-9}$ Pa). This level of vacuum could not be routinely achieved before the 1960s, and this explains the long delay in the use of the photoelectric effect in spectroscopy.

After solving the vacuum and contamination problems, photoemission spectroscopy initiated a steady growth. This expansion was accelerated in the late 1960s by the advent of synchrotron radiation sources. With their help, photoemission became the premier tool for probing the electronic structure of solids and condensed systems in general.

The Information Content of Angle-Resolved Photoemission

The basic photoemission spectroscopy experiment adopts the following strategy (**Figure 1**): photons of a specific energy $h\nu$ are used to bombard the ultraclean surface of the solid under investigation and to extract photoelectrons. The photoelectrons are collected and their flux per unit time is analyzed as a function of the kinetic energy. This approach can be used to

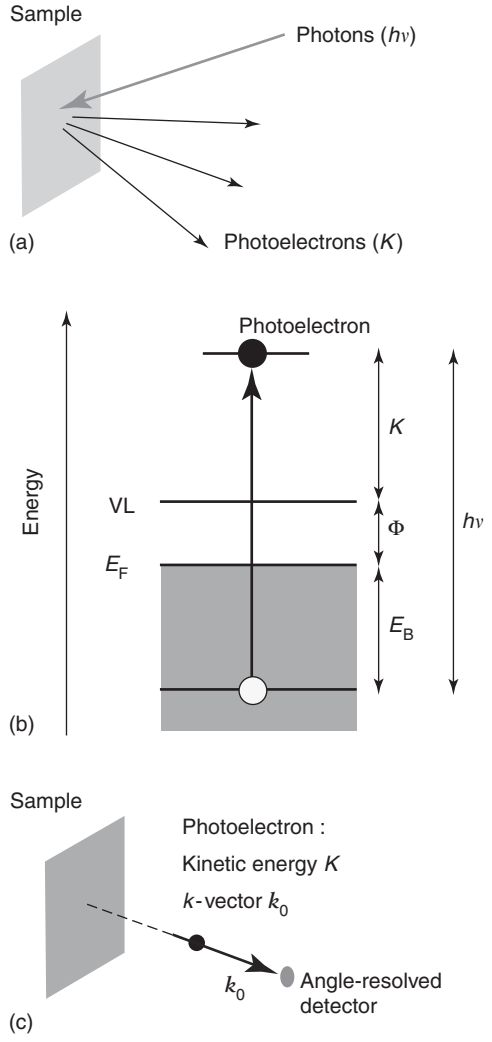


Figure 1 (a) The photoelectric effect is the foundation of photoemission spectroscopy. (b) The kinetic energy K of a photoelectron is determined by the photon energy $h\nu$, by the work function Φ, and by the initial energy of the electron in the solid or binding energy E_B; VL is the vacuum level; and E_F is the Fermi level. (c) In angle-resolved photoemission, photoelectrons are collected selecting their direction of emission. This determines not only the energy of the photoelectron but also its k-vector k_0.

study both core electrons and valence electrons: here, the analysis is limited to valence electrons. Nowadays, most photoemission experiments on valence electrons are performed by selecting the direction of emission of the collected photoelectrons. This technique is often called angle-resolved photoelectron spectroscopy (ARPES).

Figure 1a shows the essential points as far as the energy analysis is concerned in the case of a metallic solid. The zero of the photoelectron kinetic energy is called the "vacuum level" (VL). The distance between VL and the Fermi level (E_F) is the "work function," Φ. The kinetic energy K of a photoelectron is given by Einstein's equation:

$$K = h\nu - E_B - \Phi \qquad [1]$$

where E_B, the energy distance between the energy of the electron in the solid and E_F, is called the "binding energy." Equation [1] shows that by measuring K and knowing Φ and $h\nu$, the initial energy of the electron in the solid E_B can be derived.

Although this is a very valuable piece of information, the state of the electron in the solid is not completely defined by its energy. In the standard theoretical framework of solid-state physics, the state $|\boldsymbol{k}\rangle$ for a crystalline solid is defined by the crystal momentum vector $\boldsymbol{k}$, according to the Bloch theorem. The real objective of angle-resolved photoemission is thus to obtain information on $\boldsymbol{k}$.

Angle-resolved photoemission (**Figure 1b**) measures the k-vector, $\boldsymbol{k}_0$, of the free photoelectron outside the solid. In fact, $|\boldsymbol{k}_0| = (2m_0K)^{1/2}/h$, and the direction of $\boldsymbol{k}_0$ is equivalent to that of the free photoelectron. The problem, therefore, is to derive $\boldsymbol{k}$ from $\boldsymbol{k}_0$.

This is not a trivial task since many factors can influence the relation between the two vectors. In particular, the passage from the solid to vacuum changes the vector component perpendicular to the surface. Several clever schemes were developed to solve this problem.

The situation is much simpler when the system under investigation has a two-dimensional (or one-dimensional) character, for example, it is a low-dimensional crystal or a crystal surface. In that case, only the parallel component of $\boldsymbol{k}$ matters, which is in most cases equal to the measured parallel component of $\boldsymbol{k}_0$.

Angle-resolved photoemission experiments can thus experimentally probe the $E_B(\boldsymbol{k})$ function in different directions. The $E_B(\boldsymbol{k})$ function is equivalent to the so-called "band structure" of crystalline solids; therefore, the technique is known as "band mapping."

Figure 2 shows an example of band mapping, including the raw photoemission data and the corresponding plot of the $E_B(\boldsymbol{k})$ curve. It should be noted that the raw photoemission results are often plotted, not as the collected photoemission intensity versus the energy E_B, but in the form of momentum distribution curves (MDCs), that is, plots of the photoemission intensity for constant E_B as a function of the direction of the photoelectron (which coincides with the direction of $\boldsymbol{k}$). In many cases, in fact, extracting $E_B(\boldsymbol{k})$ curves is easier when the data are plotted as MDCs.

In some cases, photoelectrons are detected without selecting a specific direction and mixing together a

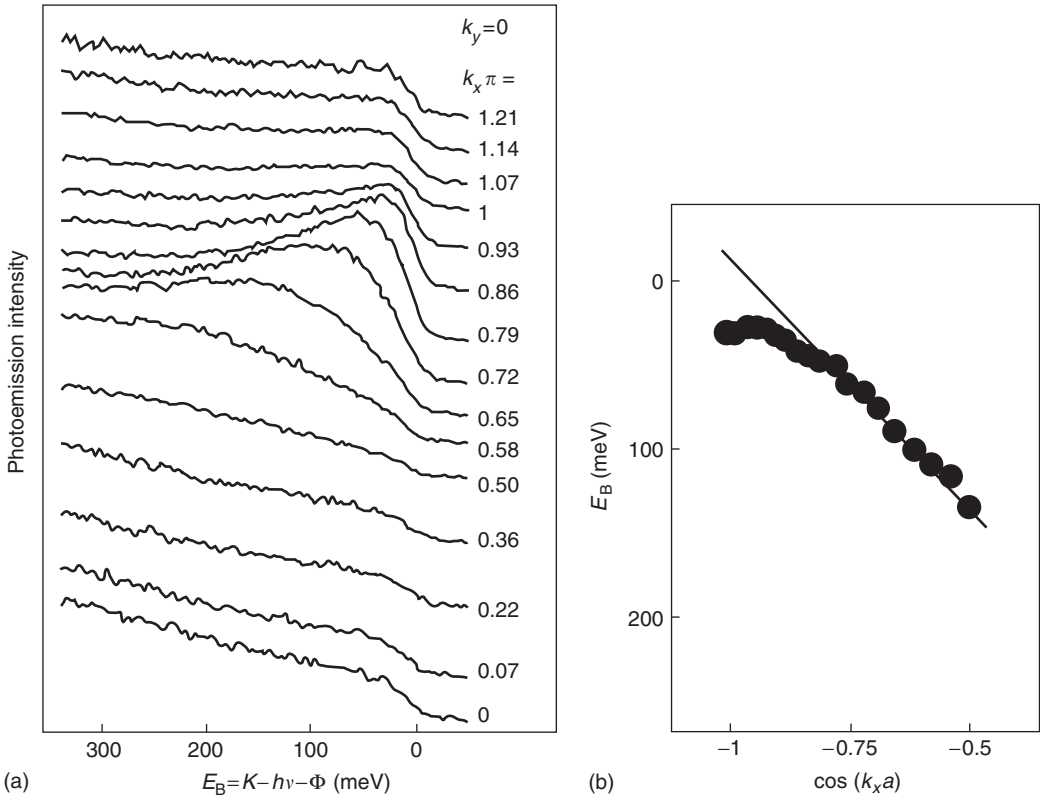


Figure 2 An example of the band-mapping technique (unpublished data from Abrecht M, Ariosa D, Cloëtta D, Margaritondo G, and Pavuna D). (a) Raw angle-resolved photoelectron energy distribution curves for an optimally doped ($x = 0.15$) strained crystalline film of LSCO ($La_{2-x}Sr_xCuO_{4-\delta}$); the different curves correspond to different *k*-vector directions along the Γ–X line. Note the dispersing feature in the low-E_B region. (b) Plot of the position of the dispersing feature, illustrating part of the "band structure" of the system.

broad range of directions. This is required when the signal level is too low and is done using angle-averaging electron energy analyzers. Furthermore, angular scrambling automatically occurs for polycrystalline samples and, to some extent, for sample surfaces cleaned by scraping.

In all these cases, plots of the photoelectron distribution in energy (energy distribution curves (EDCs)) tend to reproduce the electron "density of states." The information content is, of course, reduced with respect to that of angle-resolved photoemission.

Techniques Based on Synchrotron Sources

Several different parameters in addition to the photoelectron energy and direction can be controlled during a valence-electron photoemission experiment, simplifying the extraction of information from the data. These parameters include the photon energy, polarization and direction of incidence, as well as the photoelectron spin polarization. In most cases, the practical control of such parameters requires using a synchrotron photon source.

The superior flux and brightness of a synchrotron source enhances the signal level, facilitating all types of photoemission experiments. This is particularly important in techniques with chronically low signal levels such as spin-polarized photoemission, as discussed later.

Furthermore, synchrotron sources provide polarized photons and photons with tunable photon energy in the X-ray and ultraviolet range required for photoemission. The first feature is directly exploited to probe the symmetry of the valence electronic states.

The theoretical background of this approach is quite simple. The emission of photoelectrons requires an excitation process triggered by the photon absorption. For a given initial electronic state, the excitation may or may not be allowed depending on the symmetry of the initial and final electron state as well as on the polarization (and on the propagation direction) of the photon. By carefully analyzing the

photon polarization dependence of photoemission data, the symmetry of the initial state of the electron in the solid can thus be identified.

The final state of the photoemission process is the state of the photoelectron in vacuum. The photoemission data analysis, however, can be simplified by conceptually decomposing the photoemission process into different steps, including the optical excitation of the initial-state electron to a high-energy state in the solid. This makes the analysis of polarization effects equivalent to that of (ultraviolet or X-ray) optical absorption spectra. **Figure 3** shows an early example of this simple and powerful way to exploit polarization effects.

Interesting photoemission experiments are also performed using the circular polarization of specialized synchrotron sources, such as the "helical wigglers." These experiments provide very valuable information, notably about magnetic systems.

With the same type of analysis, one can understand the dependence of the photoemission process on the photon energy. The emission of photoelectrons depends on the optical excitation probability and, therefore, on the corresponding matrix element. The matrix element depends both on the initial state and on the final state of the excitation. When studying a given initial state with energy E_B, different photon energies correspond to different final states and, therefore, to different matrix element values. These "matrix element effects" can significantly alter the photoemission spectra, producing in some cases spurious and misleading results.

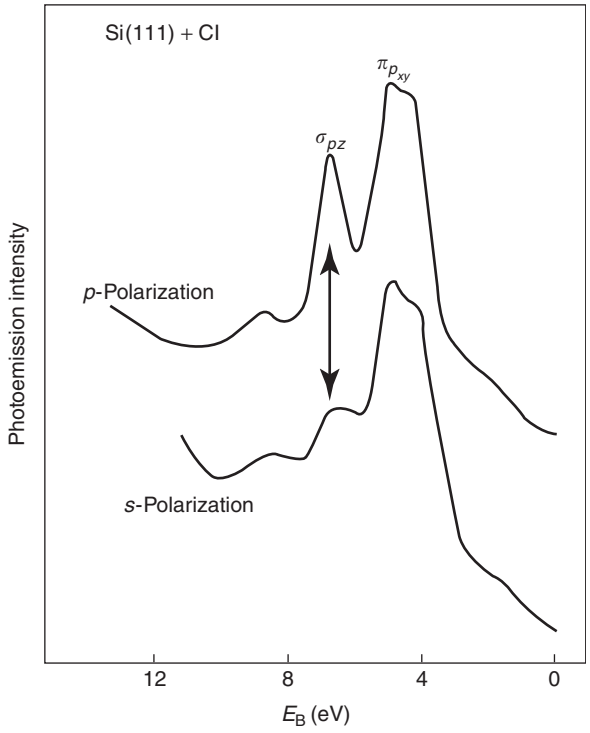


Figure 3 Photon polarization effects in the (angle-integrated) photoemission spectra of a cleaved Si(1 1 1) surface covered by chemisorbed chlorine. Note the strong decrease of the feature at $E_B = 6$–7 eV as the photon polarization changes from *p* to *s*. By analyzing the data in terms of the optical excitation induced by the photon absorption, the symmetry of the initial electronic state corresponding to this feature was identified as σ_{pz}. (Source: Schluter M, Rowe JE, Margaritondo G, Ho KM, and Cohen ML (1976) *Physics Review Letters* 37: 1632–1635.)

As a general rule, at high final-state energies, the final states are free-electron-like and the matrix element effects become negligible. In general, however, valence-electron photoemission studies are conducted with rather low photon energies (15–30 eV) and are not automatically immune from matrix element effects. The use of low $h\nu$'s is, in fact, preferable since the transition probability tends to increase for valence electrons, leading to higher signal levels. Furthermore, for a constant resolving power $h\nu/\Delta h\nu$, the absolute photon energy resolution $\Delta h\nu$ becomes better for smaller $h\nu$'s.

The best way to identify matrix element effects and to avoid mistakes in the data interpretation is to take photoemission spectra at several different photon energies and to compare the results. This procedure requires, of course, a photon source with tunable photon energy, that is, a synchrotron source. Among the suitable types of synchrotron sources, bending magnets emit a broadband of $h\nu$'s from which the desired value can be filtered with a monochromator. An undulator source automatically provides a narrow band of photon energies centered at a tunable value; however, further monochromatization is almost always required for photoemission experiments.

The possibility of controlling the photon energy offered by synchrotron sources opens up many experimental opportunities besides the identification of matrix element effects. Consider, for example, the surface sensitivity of photoemission experiments. It is seen that photoelectrons are only emitted from a thin slab near the sample surface. The thickness of this slab, called "escape depth," changes with the photoelectron kinetic energy K (see **Figure 4**), which in turn depends (eqn [1]) on the photon energy.

Therefore, by changing the photon energy, the surface sensitivity, as required for each experiment, can be increased or decreased. Roughly speaking, experiments with K-values higher than 100 eV or lower than 50 eV tend to be less surface-sensitive than experiments with K values ranging between 50 and 100 eV.

The photon energy tunability is also helpful in identifying the character of each detected initial electronic state. Matrix element effects can, in fact, be

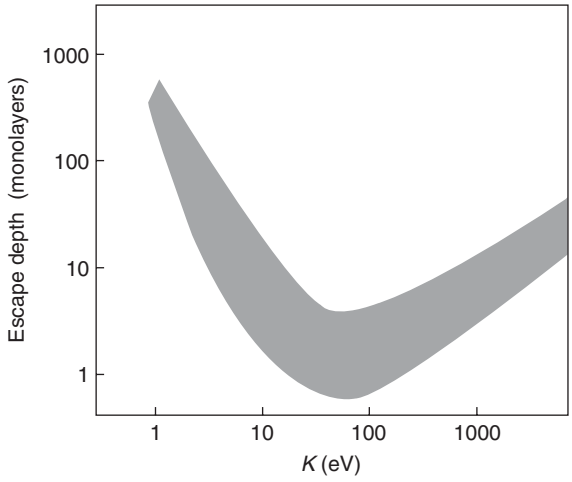


Figure 4 The escape depth of photoelectrons plotted as a function of the kinetic energy is slightly different from material to material. For all samples, however, the data fall in the shaded area and exhibit a minimum (corresponding to the maximum surface sensitivity of photoemission experiments) at $K = 50$–$100\,\text{eV}$.

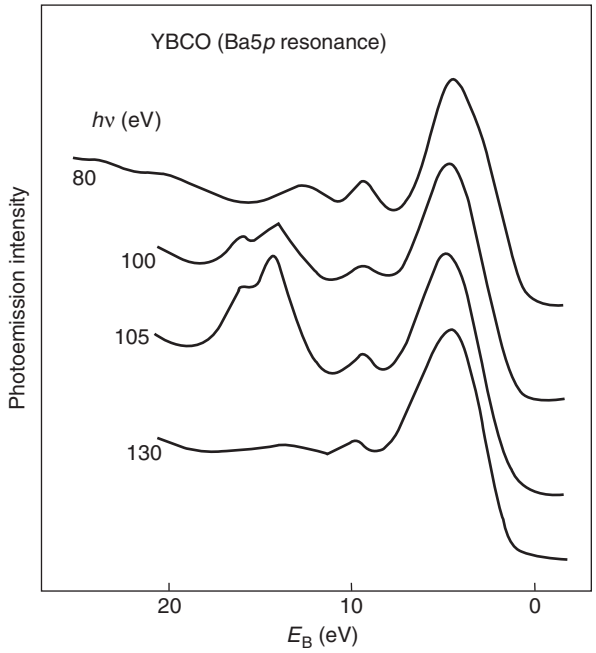


Figure 5 Photoemission spectra for the YBCO high-temperature superconductor reveal a resonating feature at $E_B = 14$–$18\,\text{eV}$ when the photon energy is near the Ba5*p* absorption edge. This photoemission resonance immediately identifies the feature as due to Ba-related states. (Source: Onellion M, Chang Y, Niles DW, Joynt R, Margaritondo G, Stoffel NG, and Tarascon JM (1987) *Physics Review* B36: 819.)

exploited for "labeling" the initial state. This is a rather difficult task except in the case of "photoemission resonances."

Several phenomena can produce very rapid and strong variations of the intensity of certain photoemission features. A resonance typically occurs for photon energies near a core-level X-ray absorption threshold of a given element and affects valence states related to the same element. Therefore, a resonating photoemission feature can be immediately attributed to the concerned element. **Figure 5** shows an example of this powerful approach.

The tunability of synchrotron sources is also exploited in specialized photoemission techniques in which the photon energy is continuously scanned rather than kept constant. One should note, in particular, the constant initial state (CIS) technique, in which both $h\nu$ and the kinetic energy K of the photoelectron are scanned while keeping their difference constant. According to eqn [1], this implies that $E_B = h\nu - K - \Phi$ stays constant; therefore, the photoemission intensity reflects the properties of the final states of the excitation more than those of the initial state of constant energy. This technique can thus be used to explore empty valence states.

Techniques Requiring High-Brightness Synchrotron Sources or Ultraviolet Lasers

The importance of the very high flux and brightness of synchrotron sources for the implementation of techniques with intrinsically low signal levels has already been mentioned. Consider, for example, the difficulties in achieving high lateral resolution in a photoemission experiment. In general, the photoelectron detector must collect the signal from a rather large sample area in order to reach a reasonable signal level. This implies that the investigated electronic structure is averaged over the same area. In the case of nonhomogeneous samples, the averaging could lead to misleading results.

High lateral resolution is often desirable, moving from "photoemission spectroscopy" to "photoemission spectromicroscopy." The corresponding decrease in signal level, however, must be avoided by using a high-brightness synchrotron source, such as an undulator.

The same requirement is present for time-resolved photoemission. In a standard photoemission experiment, the signal must be accumulated over a certain period of time to reach an acceptable signal-to-noise ratio. This is of the order of several seconds per data point and of several minutes per spectrum, making it impossible to study time-dependent phenomena faster.

The great interest of such phenomena triggered a substantial effort – based on high-brightness sources – to decrease the data-taking time. The most spectacular results are achieved with short laser pulses,

either alone or in conjunction with synchrotron sources. The required photon energies for photoemission are obtained by multiplying the photon energy of low-$h\nu$ ultrashort laser pulses with nonlinear optical devices.

Such frequency-multiplied pulses can be used to excite photoelectrons directly. Another possible approach, called "two-photon photoemission," consists of using a low-$h\nu$ laser pulse to bring valence electrons into an excited state and then to use a synchrotron pulse to extract them as photoelectrons, as shown in **Figure 6**. This approach is used to study the dynamic properties of excited states.

Techniques in this class can reach the picosecond or subpicosecond time range. Their impact is likely to be enhanced by the ongoing development of ultrabright-pulsed ultraviolet lasers based on the self-amplified spontaneous emission (SASE) mechanism.

Spin-polarized photoemission is another valence-electron technique that is chronically affected by low signal levels. Experiments discriminating the spin polarization of the photoelectrons require special detectors with typically limited effectiveness. These difficulties notwithstanding, spin-polarized photoemission has evolved to become one of the most important subfields in this domain, providing extremely valuable information, for example, on the valence states of magnetic systems.

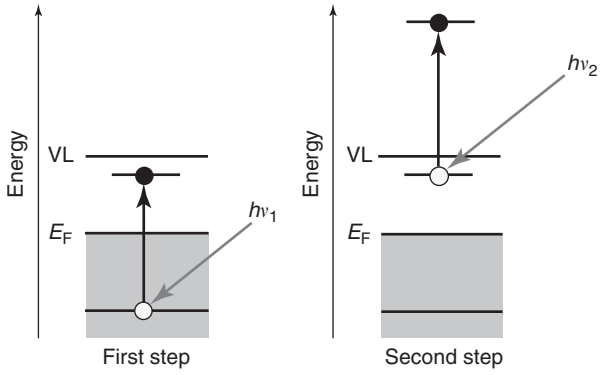


Figure 6 Schematic illustration of a two-step photoemission process involving a low-$h\nu$ laser photon ($h\nu_1$) bringing a valence electron into an excited state and a synchrotron photon ($h\nu_2$) to extract it as a photoelectron.

Valence Photoemission of Correlated Systems

The discovery of high-temperature superconductors and many subsequent experimental and theoretical

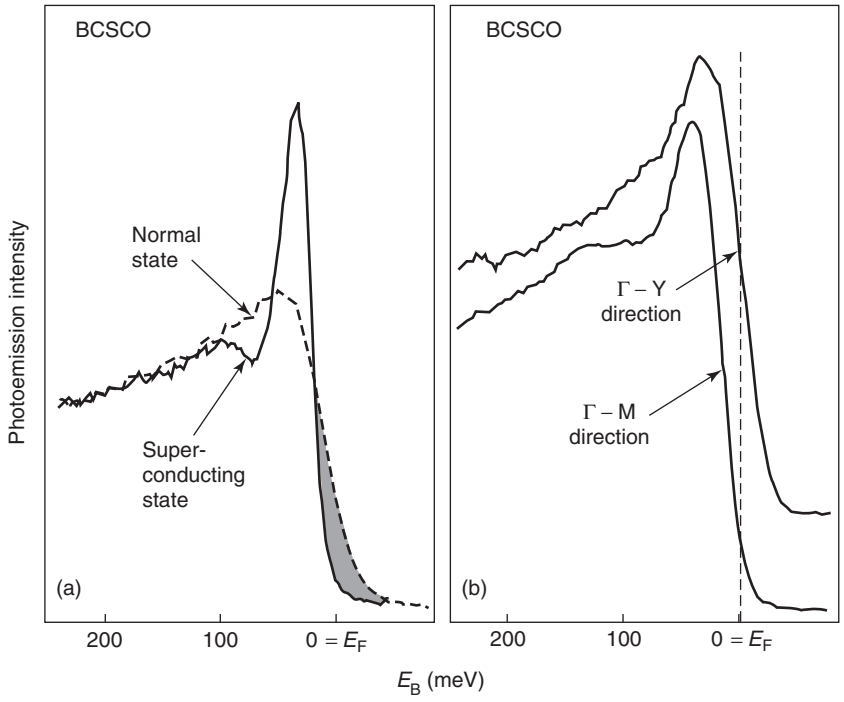


Figure 7 (a) High-resolution photoemission spectra taken on the high-temperature superconductor BCSCO at temperatures corresponding to the normal and superconducting state reveal the opening of the superconducting gap (shaded area). (Source: Hwu Y, Lozzi L, Marsi M, La Rosa S, Winokur M, Davis P, Onellion M, Berger H, Gozzo F, Lévy F, and Margaritondo G (1991) *Physics Review Letters* 67: 2573.) (b) Superconducting-state spectra taken for the same compound along different crystallographic directions reveal the anisotropy of the superconducting gap. (Source: Kelley RJ, Ma J, Quitmann C, Margaritondo G, and Onellion M (1994) *Physics Review* B50: 590.)

results increased the need of experimental probes to study electron correlation. Photoemission was able to play an important role, thanks to a substantial increase in its energy resolution levels.

Roughly speaking, phenomena related to the formation of chemical bonds affect valence electrons within 10–20 eV of the Fermi level. Except in specialized cases, their study with photoemission techniques does not require an energy resolution better than ≈ 100 meV. On the other hand, correlation phenomena occur on a smaller energy scale. A typical example is provided by the superconducting gap, which even for high-temperature superconductors does not exceed a few tens of meV. The study of such phenomena requires energy resolutions better than 50 meV, and in some cases ~ 1–2 meV.

Experimental systems capable of achieving these performances have been developed and are used to explore superconductors as well as other correlated systems. **Figure 7a** shows a typical example of these experiments: the opening of the superconducting gap during the corresponding phase transition in the high-temperature superconductor bismuth strontium copper oxide (BCSCO).

Note the high-energy edge in the normal-state spectrum, corresponding to the Fermi level. In the superconducting state, the edge recedes revealing the opening of the superconducting gap. Furthermore, the spectrum exhibits other interesting features (e.g., the strong quasiparticle peak immediately after the gap and the subsequent "dip") that have generated much debate.

Figure 7b shows an example of the valuable information that can be extracted with this approach. Superconducting-state spectra taken along different directions reveal different values of the gap, ruling out theoretical models that predict isotropy.

Correlated systems can exhibit rather "exotic" phenomena departing from the standard framework of solid-state physics. The theoretical background of almost all phenomena in solid state is provided by Landau's Fermi liquid, in which the free electrons of a Fermi gas are replaced by dressed quasiparticles with electron-like properties. The soundness of this theoretical background is evident in most photoemission results: for example, the spectrum of metallic samples (**Figure 8**) typically exhibits a Fermi-level cutoff reflecting the Fermi–Dirac distribution.

High-resolution photoemission experiments on low-dimensional systems sometime reveal phenomena departing from the Fermi-liquid framework. **Figure 8**, for example, also shows photoemission spectra of one-dimensional metallic systems that do not exhibit the Fermi-edge cutoff.

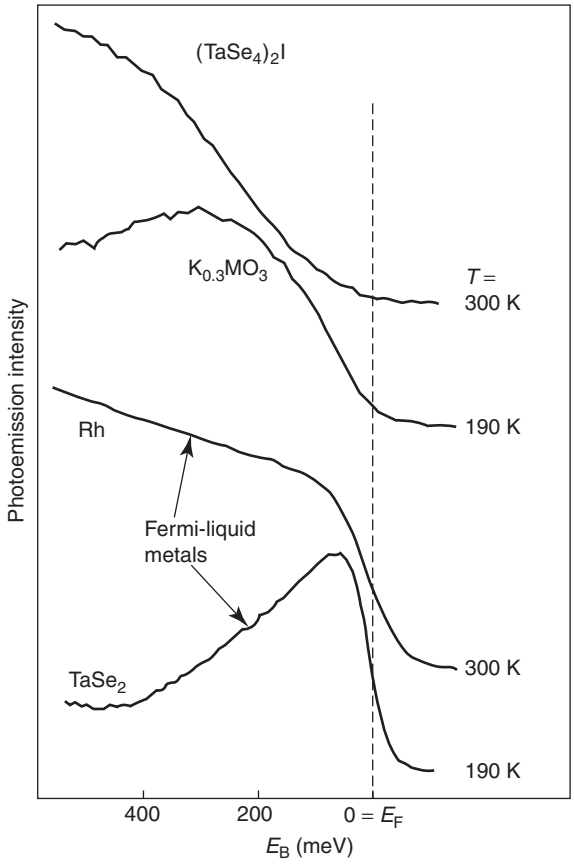


Figure 8 Bottom: photoemission spectra of three-dimensional and two-dimensional metals exhibit a clear edge at the Fermi level, as predicted for Fermi liquids. Top: one-dimensional metals often depart from this framework. (Source: Perfetti L (2002) Ph.D. thesis, EPFL: Lausanne.)

Experiments in this subdomain are trying to validate other possible theoretical frameworks to replace the Fermi liquid in highly correlated low-dimensional systems, such as the "Luttinger liquid." In extreme cases, such frameworks could lead to exotic phenomena, such as the replacement of holes as elementary excitations by spin-carrying spinons and charge-carrying holons. The validation of these theoretical hypotheses is an ambitious objective of high-resolution valence photoemission.

Correlation affects not only the initial state of the photoemission process but, in fact, the entire phenomenon also. The simplest approach to analyze photoemission spectra considers the electrons (or the Fermi-liquid quasiparticles) as a gas of independent particles. The emitted photoelectron is thus treated as an isolated entity. This is not really correct, in particular for correlated systems, and the photoemission process should be rigorously treated as the transition from a many-body correlated state to another many-body state plus a free electron. This, of course, complicates the treatment, but also enhances the

quality and quantity of information that can be extracted from the photoelectric effect.

See also: Core Photoemission; Electrons in Periodic Potentials with Magnetic Fields; Semiconductor and Metallic Clusters, Electronic Properties of; Semiconductor Heterojunctions, Electronic Properties of; Superconductors, High T_c; Synchrotron Radiation; X-Ray Absorption Spectroscopy.

PACS: 33.60.–q; 79.60.–i; 79.60.Bm; 79.60.Jv; 82.80.Pv; 87.64.Lg

Further Reading

Cardona M and Ley L (eds.) (1978–79) *Photoemission in Solids*, 2 volumes. Heidelberg: Springer.
Ghosh PK (1983) *Introduction to Photoelectron Spectroscopy*. Hoboken: Wiley.
Hufner S (2001) *Photoelectron Spectroscopy: Principles and Applications*. Heidelberg: Springer.
Hughes HP and Starnberg HI (eds.) (2000) *Electron Spectroscopies Applied to Low-Dimensional Structures*. Dordrecht: Kluwer.
Kaufmann EN (ed.) (2003) *Characterization of Materials*, 2 volumes. Hoboken: Wiley.
Lynch DW and Olson CG (1999) *Photoemission Studies of High Temperature Superconductors*. Cambridge: Cambridge University Press.
Margaritondo G (1988) *Introduction to Synchrotron Radiation*. New York: Oxford University Press.
Margaritondo G (2003) *Elements of Synchrotron Radiation for Chemistry, Biology and Medical Research*. New York: Oxford University Press.
Prince KC (2003) *Photoelectron Spectroscopy of Solid Surfaces*. Singapore: World Scientific.
Rabalais JW (1977) *Principles of Ultraviolet Photoelectron Spectroscopy*. Hoboken: Wiley.
Van Howe MA and Schattke WM (eds.) (2003) *Solid-State Photoemission and Related Methods: Principles and Practices*. Hoboken: Wiley.

van der Waals Bonding and Inert Gases

J S Rutherford, National University of Science & Technology, Bulawayo, Zimbabwe

Introduction

By van der Waals forces, one means those weak bonding interactions which occur between neutral, closed-shell atoms or molecules, and are not characterized by the contact of specific atomic groupings within these molecules, as is the case with hydrogen bonding and donor–acceptor complexation. They are called van der Waals forces because of their relation to the van der Waals equation of state for a fluid phase:

$$\left(p + \frac{a}{V^2}\right)(V - b) = nRT$$

which was the first successful attempt to model the behavior of real gases, by introducing corrections to the ideal gas law to take into account both the attractive forces between molecules (which are assumed to reduce the effective pressure by an amount a/V^2) and the finite molecular volumes, through the excluded volume b.

The van der Waals equation can be made identical for all species by choosing reduced variables based on the critical constants of the substance:

$$p_r = \frac{8T_r}{3V_r - 1} - \frac{3}{V_r^2}$$

Here $p_r = p/p_c$, where p_c is the critical pressure, and V_r and T_r are similarly defined. This result suggests that the behavior of any fluid should follow a law of corresponding states, that is, it should be identical for all species when treated in terms of the reduced variables. This law is found to be a good approximation experimentally. This observation, however, does not give specific support to the van der Waals equation, since any similar two-parameter equation would also conform to the law of corresponding states, albeit with different numerical results.

Nature of van der Waals Forces

van der Waals forces may be classified into three types: electrostatic, induction, and dispersion. Most textbooks only mention the most important interaction in each class, that is, the dipole–dipole, dipole-induced dipole, and London dispersion contributions, as these are always significant when they occur. Nevertheless, and particularly where the absence of a permanent molecular dipole arises through cancellation, by symmetry, of strong bond dipole moments, the effects of the higher electrostatic moments can make major contributions. Examples of such molecules are linear CO_2 (quadrupole) and tetrahedral CF_4 (octopole).

The energies of these main interactions at a distance r are given by the following formulas, where μ is the permanent molecular dipole moment, α_1 and α_2 are the polarizabilities of the two species a and b, I_1 and

l_2 their ionization potentials, and ε is the permittivity of vacuum:
Dipole–dipole (fixed):

$$-\frac{\mu^2}{4\pi\varepsilon_0 r^3}(2\cos\theta_a\cos\theta_b - \sin\theta_a\sin\theta_b\cos\varphi)$$

Here, θ_a and θ_b are the angles the dipoles make with the vector r, and ϕ is the angle between their projections along r.
Dipole–dipole (free) (Keesom energy):

$$-\frac{2\mu^2}{(4\pi\varepsilon_0)^2 k_B T r^6}$$

Dipole-induced dipole (Debye):

$$-\frac{\mu^2\alpha}{(4\pi\varepsilon_0)^2 r^6}$$

Dispersion (London):

$$-\frac{3\alpha_1\alpha_2 I_1 I_2}{2(I_1+I_2)(4\pi\varepsilon_0)^2 r^6}$$

Potential Energy Functions

The basic features of a bonding interatomic potential $U(r)$ are: (1) $U(r)$ is a maximum when the internuclear separation $r=0$, (2) $U(r)\to 0$ as $r\to\infty$, and (3) for some intermediate equilibrium separation r_e, U has a minimum $U(r_e)=\varepsilon$, where ε is the binding energy. For these requirements to be satisfied by a model potential function, the function must include at least two components, one a repulsive interaction which dominates when $r<r_e$, and the other an attractive interaction dominant when $r>r_e$. This in turn means that, except for the simplest model of all, the hard-sphere potential which corresponds to van der Waals original model, the potential function must consist of at least two terms with opposite signs.

The hard-sphere potential is exceptional, in that it employs a discontinuous function

$$U(r) = \infty, \qquad r<r_e$$
$$U(r) = -C_6/r^6, \quad r\geqslant r_e$$

Typically, continuous interatomic potentials have been developed from the Mie function

$$U(r) = -C_n/r^n + C_m/r^m$$

with n and m positive, and usually integers. This function has the necessary form provided $n>m$.

In the case of van der Waals forces, the attractive term may be assumed, as a first approximation, to vary as r^{-6}, in which case the resulting potential function is the Lennard–Jones (n, 6)-potential. The repulsive part of the potential is common to all types of bonding, mainly being due to the mutual repulsion of overlapping closed-shell electron clouds at shorter distances, but its effective n is not easy to estimate. In the case of the Born–Lande (i.e., simple Mie type) potential for ionic interactions, where $m=1$, n is commonly taken to be ~ 10; the results are not very sensitive to its precise value. Historically, the van der Waals interaction has commonly been treated using the Lennard–Jones (12,6)-potential

$$U(r) = -C_6/r^6 + C_{12}/r^{12}$$

alternatively written as

$$U(r) = 4\varepsilon\left\{\left(\frac{\sigma}{r}\right)^{-6} - \left(\frac{\sigma}{r}\right)^{-12}\right\}$$

where ε is the binding energy and σ the value of r (other than infinity) where $U(r)$ is zero. The popularity of the (12, 6)-potential is partly due to its mathematical simplicity, for, if one substitutes $x=(\sigma/r)^{-6}$, then

$$U(r) = 4\varepsilon(x - x^2)$$

and, for example, can readily show that $r_e = 2^{1/6}\,\sigma$.

Dipole–dipole interactions in fluid phases are readily included in the Lennard–Jones potential, since when averaged it behaves as r^{-6}; however, in solids the fixed orientations of the dipole moments require to be handled explicitly. The result is the Stockmayer potential

$$U(r,\theta_a,\theta_b,\phi) = 4\varepsilon\left\{\left(\frac{\sigma}{r}\right)^{-6} - \left(\frac{\sigma}{r}\right)^{-12}\right\} - \frac{\mu^2}{4\pi\varepsilon_0 r^3} \times (2\cos\theta_a\cos\theta_b - \sin\theta_a\sin\theta_b\cos\varphi)$$

where the angular coordinates are the same as before.

The repulsive term used in the Lennard–Jones potential is unsuitable for accurate work. First of all, because of the weak nature of the van der Waals attractive forces, the relative magnitude of the repulsive interactions at the equilibrium distance is much greater than for other types of bonding. For the (12, 6)-potential, the repulsive contribution calculates to be as large as the net binding energy, compared to only $\sim 10\%$ for the Born–Lande potential. In addition, typically 99% of an atom's electron density is contained within its van der Waals radius, and so the overlap interaction involves electronic

wave functions in their region of exponential decay toward infinite r. Hence, the repulsive contribution can be improved by replacing the r^{-n} term by an exponential, creating the (exp, 6)-potential

$$U(r) = A \exp(-Br) - C_6/r^6$$

where A and B are arbitrary fitted constants.

Potential functions such as the Lennard–Jones potential that have two adjustable parameters can be fitted to the experimental parameters r_e and ε, and are consistent with the law of corresponding states. However, if additional parameters, such as the three required by the (exp, 6)- potential are introduced, the law of corresponding states can only be satisfied if the potential functions for the various species are conformal, that is, the terms are constrained such that the potential curves have the same shape function in units of r_e and ε. Where the actual potential function can be mapped very accurately by experiment, an even more complicated theoretical form may be needed to fit it. Such is the case for the pure pairwise interactions of inert gas atoms, where Rydberg–Klein–Rees analysis may be applied to derive the form of the potential from vibrational or rotational–vibrational spectroscopic measurements on the corresponding van der Waals dimer molecules.

Such functions are useful in calculating atom–atom pair potentials for small and medium-sized molecules. However, accurate lattice energy calculations cannot ignore interactions that are not pairwise additive, although in this regard it is usually considered to be sufficient to include the Axilrod–Teller three-center term

$$\nu_{ABC} = \frac{(3 \cos \theta_a \cos \theta_b \cos \theta_c + 1)}{(r_{AB} r_{BC} r_{AC})^3}$$

where θ_a is the angle between r_{AB} and r_{AC}, etc.

van der Waals Radii

van der Waals radii can be defined in the same way as covalent, ionic, and metallic radii are for interactions. Since it is normally possible to find cases where atoms of the same element in adjacent molecules are in contact, the van der Waals radii can be defined as

$$r_A = R_{AA}/2$$

and, since the repulsive part of the interatomic potential becomes dominant at essentially the same distance for any element A regardless of the element B with which it is in contact, the van der Waals radii are effectively additive in the sense

$$R_{AB} = r_A + r_B$$

Table 1 van der Waals radii of some common species

Atom/radical	*Radius* (nm)
H	0.120
C	0.185
$-CH_3$	0.200
Si	0.210
N	0.150
P	0.190
As	0.200
Sb	0.220
O	0.140
S	0.185
Se	0.200
Te	0.220
F	0.135
Cl	0.180
Br	0.195
I	0.215

Prepared by Ladd MFC (1994) *Chemical Bonding in Solids and Fluids*. Chichester: Ellis Horwood.

Table 1 gives a number of values.

However, such values have to be treated with caution on two accounts. The first is that the minimum energy structure of a condensed phase of a complex molecular species is determined by the large number of (mainly) pairwise interactions between its constituent atoms in contact, which may require, given the relative internal rigidity of the molecules, some of these interactions not to be at their individual equilibrium distances. Because the potential well is so shallow, an individual contact can vary considerably from the predicted minimum with little energy change.

The second effect is the variation of the van der Waals radius with angle relative to an internal covalent bond. A particularly well-studied example is that of Cl–Cl interactions in crystals of both the element itself and many organochlorine compounds.

Phases and Phase Equilibria

The presence of van der Waals attractive forces for a molecular species leads to the stabilization of condensed phases at lower temperatures. However, the fact that the potential energy wells are wide and shallow, is reflected in a number of their physical properties. The low cohesive energies correspond to low melting points and boiling points, and the slow variations in pair potential with distance near the minimum result in large compressibilities and thermal expansivities.

As far as the boiling point is concerned, it is possible to make this correlation quantitative. The constants in the van der Waals equation can be related to hard sphere intermolecular potential as

$$a = 4\pi c/3d^3 \quad \text{and} \quad b = 2\pi d^3/3$$

where the attractive potential at $r > d$, the collision diameter, is given by $-C/r^6$.

Thermodynamic arguments provide an approximate value for the self-cohesive energy of a liquid at its boiling point T_b of $\sim 7RT_b$, assuming the ratio of the molar volumes V_g/V_l of the two fluid phases is $\sim 10^3$. When the work term $p\Delta V$ and the translational energy of the gas is added to this value, the enthalpy of vaporization of the liquid at the boiling point becomes roughly $9.5\,RT_b$, in line with Trouton's empirical rule. Thus, the boiling point of the liquid can be directly related to its cohesive self-energy.

Unlike the gas/liquid case, a solid and fluid in equilibrium comprise a system in which the two phases present differ in symmetry properties. The fluid is spatially homogeneous and rotationally isotropic, while the solid normally conforms, at least in ideality, to one of the space symmetry groups.

Some other properties depend on the fact that the van der Waals forces are virtually nondirectional except when fixed-orientation dipole–dipole interactions are involved. For even a small molecule, the overall energy of the system is principally determined by the net effect of a large number of these pairwise interactions at or near their minimum energy distance, and so the total volume of a particular model configuration which has been stabilized in terms of the elimination of residual forces, is a fair estimate of its relative energy. In other words, if one treats the molecules as having hard van der Waals surfaces, arrangements with similar packing fractions are likely to have similar total van der Waals energies. This tendency manifests itself in the frequent occurrence of polymorphism for molecular crystals.

An allied phenomenon is the existence of mesophases, phases intermediate in order between solids and liquids, for a number of molecular compounds. Such phases, having both enthalpy and entropy higher than the regular solid but lower than the liquid, may be stable at intermediate temperatures, and include plastic crystals and liquid crystal phases of various types.

Liquid Structure

Models of the structure of liquids are of two types. In one approach, first introduced by Bernal, a randomly packed aggregate of identical objects (ball bearings, plasticine spheres, etc.) is analyzed in terms of the distribution of nearest neighbors in number and distance. The definition of nearest neighbors here depends on the presence of a shared face of their space-filling Voronoi polyhedra. Such models typically show an average coordination number of about 13, compared to 12 for a regular close-packed array of equal spheres.

The other approach to modeling such liquids is by computer simulation. Either the Monte Carlo technique or molecular dynamics may be used. In either method, a number of molecules are given starting positions within a hypothetical volume of the bulk liquid in the form of a cube, and the density is maintained constant by imposing periodic boundary conditions. In the Monte Carlo technique, the coordinates are randomly adjusted one at a time, in such a way as to produce a series of configurations of individual energy per molecule U_N, consistent with the Boltzmann distribution $\exp(-U_N/kT)$.

In a molecular dynamics simulation, the molecules move simultaneously, each under the net force arising from the pair potentials with its neighbors, according to Newton's laws

$$m\frac{\mathrm{d}^2R}{\mathrm{d}t^2} = -\sum_{j\neq i}\frac{\partial}{\partial R_{ij}}U(R_{ij})$$

and the subsequent position and motion of the molecules is recorded over a series of very short time intervals.

Both types of simulation serve to provide estimates of thermodynamic quantities such as configurational energy and heat capacity, while the structure of any type of model may be compared to the experimental radial distribution function derived from diffraction experiments. For a monatomic species, the radial distribution function is

$$N(r) = 4\pi r^2 g(r) N/V \, \mathrm{d}r$$

where $g(r) = 1$ corresponds to a completely uniform distribution. In liquids, $n(r)$ is zero when r is less than the minimum contact distance, and fluctuates around the parabolic form corresponding to $g(r) = 1$, as shown in **Figure 1** for argon.

Structures of van der Waals Solids

Only recently there have been serious attempts to predict the crystal structures of other than the simplest molecules. The close-packed structures of the inert gases can be derived as having the maximum packing fraction for hard spheres, and other ideal

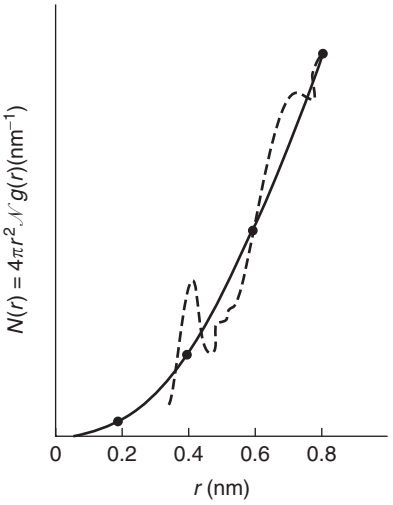


Figure 1 Radial distribution function for liquid argon. Solid line: random distribution $g(r) = 1$. Dashed line: observed distribution. (Prepared by Ladd MFC (1994) *Chemical Bonding in Solids and Fluids*. Chichester: Ellis Horwood.)

geometric forms, such as rods or disks, can be treated similarly, but for the general molecule, even assuming it has a rigid conformation and a hard van der Waals surface, it is very difficult to establish a global minimum in the lattice energy. However, the experimental frequency of occurrence of the individual space group types was interpreted by Kitaigoroskii as a preference for those symmetry operations containing a translational component (i.e., screw axes and glide planes) operations that allow the bumps of one molecular surface to be fitted into the hollows of its neighbors, and so increase the packing fraction. Arrangements of molecules in such preferred space groups are often the starting point for structure prediction using an overall minimization of the lattice energy.

Molecular Crystals of the Elements

As might be expected, the crystal structures of the solid inert gases have much in common with those of the elemental metals, despite the different cohesive forces, since both are based on the effective packing of spheres. All the inert gases have a cubic close-packed phase, while He and Ne show hexagonal close-packed structures. Helium becomes body-centered near its melting point.

Of the diatomic elements, hydrogen and nitrogen are closest to spherical, and show cubic and hexagonal close-packed phases, either in a distorted form or with freely rotating molecules. Oxygen and fluorine have similar monoclinic low- and cubic high-temperature phases, while the remaining halogens have the same layer herringbone-crystal arrangement typified by Cl_2 (s).

The other molecular elements (P_4, S_8, Se_8, etc.) have low-symmetry structures based on the packing of irregular molecules.

See also: Allotropy and Polymorphism; Crystal Structure; Liquid Crystals; Meso- and Nanostructures; Molecular Crystallography.

PACS: 61.50.Lt; 61.20.Gy; 61.20.Ja; 61.66.Bi

Further Reading

Buckingham AD, Fowler PW, and Hutson JM (1988) Theoretical studies of van der Waals molecules and intermolecular forces. *Chemical Reviews* 88: 963–988.

Chalasinki G and Gutowski M (1988) Weak interactions between small systems. Models for studying the nature of intermolecular forces and challenging problems for *ab initio* calculations. *Chemical Reviews* 88: 942–962.

Hobza P and Zahradiuk R (1988) Intermolecular interactions between medium-sized systems. Nonempirical and empirical calculations of interaction energies: successes and failures. *Chemical Reviews* 88: 871–897. Washington: American Chemical Society.

Kaplan IG (1986) *Theory of Molecular Interactions*. Amsterdam: Elsevier.

Leusen FJJ, Wilke S, Verwer P, and Engel GE (1999) Computational approaches to crystal structure and polymorph prediction. In: Howard JAK, Allen FH, and Shields GP (eds.) *Implications of Molecular and Materials Structure for New Technologies*, pp. 303–314. Dordrecht: Kluwer.

Maitland GC, Rigby M, Smith EB, and Wakeham WA (1981) *Intermolecular Forces – Their Origin and Determination*. Oxford: Clarendon.

Price SL (2000) Toward more accurate model intermolecular potentials for organic molecules. In: Lipkovitz KB and Boyd DB (eds.) *Reviews in Computational Chemistry*, vol. 14, pp. 225–289. New York: Wiley-VCH.

Rigby M, Smith EB, Wakeham WA, and Maitland GC (1986) *The Forces between Molecules*. Oxford: Oxford University Press.

Stone AJ (1996) *The Theory of Intermolecular Forces*. Oxford: Clarendon.

Verwer P and Leusen FJJ (1998) Computer simulation to predict possible crystal polymorphs. In: Lipkovitz KB and Boyd DB (eds.) *Reviews in Computational Chemistry*, vol. 12, pp. 327–365. New York: Wiley-VCH.

Wright JD (1995) *Molecular Crystals*, 2nd edn. Cambridge: Cambridge University Press.

Vapor Transport Processing: Mechanisms

D Gall, Rensselaer Polytechnic Institute, Troy, NY, USA

Introduction

The deposition of materials from the vapor phase has become a major processing technique that is essential for nearly all of today's high-technology industries. Atoms or molecules are deposited onto surfaces to form coatings or thin films ranging in thickness from one atomic layer ($\sim$0.3 nm) to hundreds of micrometers. In many applications, the coating is applied in order to alter the physical properties of the underlying bulk material, while in other cases the entire functionality of a material or device is performed by the deposited layer(s) with the bulk material simply providing mechanical support. Examples of the former include hard wear-resistant lubricious coatings on bearings and cutting tools, optical reflective and antireflective coatings on windows and lenses, thermal barrier coatings in gas-turbines, biocompatible layers for implants, decorative coatings on consumer products, barrier layers to protect against diffusion, oxidation, corrosion, and ultraviolet radiation in food packaging, automotive applications, and on various polymeric or metallic components. Examples where the primary functionality is achieved by the deposited layers themselves include integrated microelectronic devices, micro-electromechanical devices, magnetic memory disks, optical waveguide circuitry, and solid-state lasers and lighting. Such complex devices are achieved by depositing multiple layers from the vapor phase and lithographic patterning in the plane of the films, yielding a vast variety of engineered three-dimensional (3D) structures.

This article introduces the key processes that occur during thin film deposition from the vapor phase and gives an overview of the most commonly used techniques. The deposition process includes a source which provides atoms or molecules in vapor phase, the transport of the depositing species from the source to the substrate, and the actual deposition process at the substrate and growing layer.

Process Steps: Source, Transport, and Deposition

Deposition techniques are commonly referred to as physical vapor deposition (PVD) and chemical vapor deposition (CVD) based on the phase of the material at the source. In the case of PVD, solid materials are vaporized by heating them at the source using electrical heating, electron beams (e-beam evaporation), or lasers (pulsed laser deposition (PLD)). An alternative to heating the solid source material is the use of energetic ions which knock out atoms into the vapor phase by an atomic momentum transfer, which is referred to as sputtering. In CVD, the source material is a gas, an evaporated liquid, or a solid that is chemically converted to a vapor, such as the conversion from Ga (solid) to GaCl (vapor).

The second step during the deposition process is the transport of the depositing species in the vapor phase from the source to the substrate. This process is strongly influenced by the pressure at which the process is operated since the mean free path λ_{mfp} of a molecule in a gas is inversely proportional to the pressure p. At room temperature and atmospheric pressure, λ_{mfp} is $\sim$100 nm. It increases as the pressure decreases according to

$$\lambda_{mfp} = 10^{-2}/p \qquad [1]$$

Here, the units p and λ_{mfp} are Pa and m, respectively. Thus, at pressures below 10^{-2} Pa ($1\,\text{Pa} = 7.5\times 10^{-3}\,\text{Torr} = 10^{-5}\,\text{atm}$), the molecules travel more than 1 m, which typically corresponds to a direct path from chamber wall to chamber wall. The vapor transport is characterized by the Knudsen number Kn defined by

$$Kn = \lambda_{mfp}/L \qquad [2]$$

where L is the characteristic length scale, for example, the source–substrate distance or a pipe diameter through which vapor is flowing. The transport is then divided into

$$\text{molecular flow} \quad Kn > 1 \qquad [3a]$$

$$\text{intermediate flow} \quad 1 > Kn > 0.01 \qquad [3b]$$

$$\text{viscous flow} \quad Kn < 0.01 \qquad [3c]$$

Evaporation is typically done at low pressures, below $\sim 10^{-2}$ Pa, that is, $Kn > 1$. Therefore, the atoms or molecules do not scatter with other species and move on a straight path from the source to the substrate. Under such high-vacuum conditions, the geometrical arrangement of source(s) and substrate are the primary factors influencing the deposition uniformity. In contrast, CVD is mostly done at pressures above $\sim$1 Pa ($Kn < 0.01$) where the molecules experience

many collisions and the transport is described by a viscous flow with vapor-phase diffusion and more complex gas flow patterns. A third category of processes, typically those which utilize plasmas, operate in the crossover intermediate range, where molecules experience several collisions during their travel from source to substrate but do not completely thermalize. There are exceptions to this categorization, for example, gas-source molecular beam epitaxy (GS-MBE), which is a CVD technique that operates at high-vacuum ($<10^{-2}$ Pa) pressures.

The third step in the deposition process is the actual deposition, that is, the processes that occur at the substrate and the growing layer surface, including adsorption, chemical reactions, surface diffusion, and the effect of energy input in various forms. The sticking coefficient S_c is defined as the probability of an impinging atom or molecule to react with the surface and become incorporated in the growing layer. S_c is nearly unity for many PVD processes where the vapor pressure of the deposited material at the substrate temperature is small compared to the deposition flux. In contrast, S_c is often much smaller (e.g., 10^{-3}) for CVD processes since, in this case, a chemical reaction needs to occur for the deposition molecule to chemically bond with the layer. An advantage of low S_c is the ability to coat complex surface morphologies uniformly, however, at the cost of a reduced deposition rate.

The energy input onto the surface affects both the surface chemistry and the atomic arrangement which, in turn, determines the composition and microstructure. The substrate temperature T_s is a parameter that is important for all deposition techniques. In many cases, multiple thermally activated processes, including surface chemical reactions and surface diffusion, occur simultaneously. Varying the temperature typically results in large changes in layer properties, since the rate R for these processes depends exponentially on T_s:

$$R \propto \exp(-E_a/kT_s) \qquad [4]$$

Here, k is the Boltzmann constant and E_a the activation energy for a given process.

In addition to thermal activation, there are various other energy sources that affect the layer growth during deposition. Many deposition techniques use plasmas that create energetic ions that impinge onto the substrate with energies ranging from ~ 1 to ~ 1000 eV. This energy is much higher than the thermal energy ($kT = 0.025$ eV at room temperature). Other energy sources are photons during photoassisted deposition (typically 1–3 eV), and chemical energy, carried in reactive molecules (typically 1–20 eV), which are common during CVD or reactive PVD.

Evaporation

Evaporation is a vapor deposition technique where the source material is heated so that it evaporates or sublimes. This is typically done in a high vacuum with a background pressure below 10^{-2} Pa, so that the evaporated atoms do not scatter on their path from the source to the substrate. The evaporation flux can be calculated using the vapor pressure p_v, which is defined as the pressure at which either a solid or a liquid is in equilibrium with the vapor (for a given temperature and element). The Clausius–Clapeyron equation provides the following expression for the temperature dependence of p_v:

$$p_v = p_0 \exp(-\Delta H_e/kT) \qquad [5]$$

where ΔH_e and p_0 are the heat of evaporation and a prefactor, respectively. **Figure 1** shows the vapor pressures for various elements as a function of temperature, indicating the validity of eqn [5] over a large temperature range, and above and below the melting points. At equilibrium, the flux from the condensed phase (solid or liquid) to vapor and the flux from the vapor to the condensed phase are equal. Consequently, the evaporation flux Φ_e is equal to the flux that a vapor exerts on a surface at a given temperature and pressure, which is obtained from the kinetic gas theory

$$\Phi_e = \frac{N_A}{(2\pi MRT)^{1/2}} p_v \text{ molecules cm}^{-2} \text{ s}^{-1} \qquad [6]$$

Here, N_A, M, and R are Avogadro's number, the molecular weight, and the gas constant, respectively.

The heating of the source for evaporation is typically done by resistive or electron-beam (e-beam) heating. For resistive heating, the source material is placed in direct contact with a tungsten filament, in a refractory metal sheet which has various shapes and openings, or in a ceramic crucible which is heated by a tungsten wire. The disadvantage of direct heating is that heaters and support materials are required to be at high temperatures, which causes contamination and limits the evaporation of high-melting-point materials. E-beam heating eliminates these disadvantages: electrons are emitted from a filament and accelerated onto the source material by a potential of several thousand volts. This yields a high temperature at the point where the e-beam impinges while the bulk part of the source material, which is contained in a water-cooled crucible, remains cold.

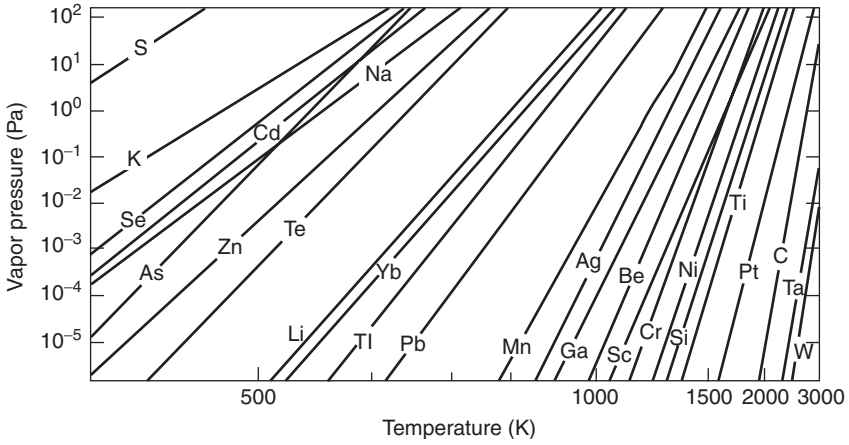


Figure 1 Vapor pressures for various materials.

A completely different approach for evaporation is employed during PLD, where high-power laser pulses are absorbed by a target material and converted into a combination of thermal, chemical, and mechanical energy. The evaporant is ejected from the target and forms a plume, which consists of energetic atoms and particles that travel through the vacuum to the substrate.

Sputter Deposition

Sputtering is the process where energetic particles impinge onto a target material and knock-out surface atoms. This process is widely used as a source for PVD. In contrast to evaporation, the source material remains cold during sputtering. Thus, sputter deposition is particularly well-suited for deposition of high-temperature materials which are hard to evaporate. In addition, sputtering typically yields a range of energetic particles that impinge onto the substrate and the growing layer and strongly influence the microstructure.

A generic schematic of a sputter deposition arrangement is shown in **Figure 2**. The target, consisting of the source material, is placed facing the substrate. A negative voltage of typically 0.2–5 kV is applied to the target (cathode). The substrate may be grounded and becomes the anode. An inert working gas, usually Ar, is introduced into the chamber and its pressure, typically between 0.1 and 10 Pa, is kept constant by controlling the inlet flow while continuously pumping. Positively charged ions are accelerated toward the target where they sputter target atoms, which then travel in the vapor phase and get deposited on the substrate. The sputter process also yields free electrons which are accelerated away from the cathode and, by collisions in the vapor phase, ionize gas atoms yielding more ions and electrons

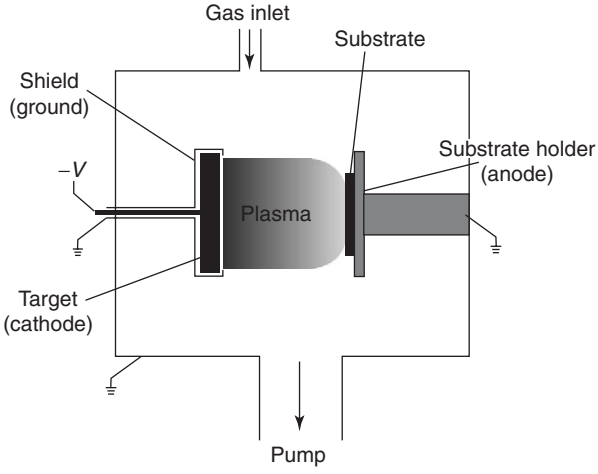


Figure 2 Generic schematic of a sputter deposition system.

which continue the sputter process. This chain reaction yields a charge neutral mixture of energetic electrons and ions between the target and the substrate, which is termed plasma or glow discharge because the excited ions often emit visible photons when relaxing to their ground state.

There is a wide range of specific arrangements and techniques to achieve the desired layer compositions and microstructures by sputtering. The following sections will introduce some common concepts and terms.

"Magnetron sputtering" refers to the use of magnetic fields to improve the ionization per electron. This technique lowers the pressure required to sustain a plasma and, in turn, yields higher deposition rates. **Figure 3** shows a schematic of a planar magnetron arrangement. Magnets are placed behind the target to form curved field lines parallel to the surface. Free electrons are accelerated by the negative target potential and, due to the Lorentz force, exhibit

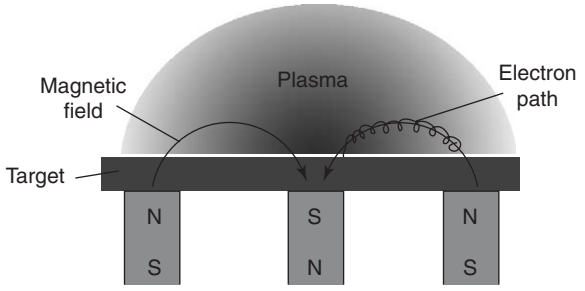


Figure 3 Planar magnetron arrangement for sputter deposition.

a spiral-motion circling along the magnetic field lines. Since the field lines bend back toward the target, the electrons are effectively trapped. They, therefore, remain much longer in the chamber than they would in the absence of magnets where they travel on straight paths from the target to the chamber wall. The confinement of electrons close to the target increases the ionization rate and, consequently, the plasma density.

"Ion-beam sputtering" utilizes a geometrically confined plasma to create energetic ions, which are directed and accelerated toward a target by grids in Kaufman-type ion guns. Such an approach allows a low pressure in the deposition chamber itself, with a mean free path of the depositing species far exceeding the target–substrate distance. This technique is less commonly used for deposition purposes but is very popular, for example, to acquire depth-profile measurement by photoelectron or Auger electron spectroscopy, to create specific 3D structures in a focused ion-beam system, or to prepare transmission electron microscopy samples by ion milling.

"Radio frequency (RF) sputtering" is primarily used to deposit insulating materials. During the above-described DC-sputtering process, ions impinge on the target and electrons are emitted, yielding an electric current which flows from the grounded substrate and chamber wall to the target. However, in order for this current to be sustained, the target material needs to be conductive. In contrast, RF electric fields cause oscillating electrons which can acquire sufficient energy for ionization, yielding a plasma that does not require electrons from the target. The electrons in such an RF-plasma have a much higher mobility than the ions, causing more electrons than ions to impinge onto the target. The insulating target, however, forces the currents of impinging electrons and ions to be equal, which yields a negative target self-bias. This (DC) self-bias causes the ions to be accelerated toward the target which results in the sputtering process.

"Compound targets" can be utilized to deposit compound coatings. The sputtering rate of the different elements in the target varies. However, this does not affect the emitted composition, since in steady state the sputtered composition corresponds to the bulk target composition. Nevertheless, unequal atomic transport efficiencies from the target to the substrate may cause the coating composition to vary from the target composition, requiring an adjustment of the target composition to achieve the desired coating.

"Co-sputtering" is the simultaneous deposition from two or more targets to create compound layers. In contrast to the use of compound targets, co-sputtering provides great flexibility in composition which is controlled by the relative power applied to each target. A disadvantage of co-sputtering is the non-normal arrangement of substrate and targets, which may yield lateral composition variations in the deposited layer.

An alternative approach for compound deposition is "reactive sputtering," which uses, in addition to the inert processing gas, a reactive gas (e.g., N_2, O_2, and CH_4), which provides elements such as C, N, and O to the layer. For example, TiN hard coatings are deposited using a metallic Ti-target and an Ar–N_2 gas mixture. The process control for reactive sputtering has its own challenges due to the surface reaction of the target, also called target poisoning. This is particularly true for the deposition of oxides where the presence of oxygen in the deposition chamber can lead to the formation of an insulating surface oxide on the target, which dramatically reduces the deposition rate and alters the layer composition.

An important aspect of sputter deposition is the presence of energetic particles that impinge on the substrate and typically cause a densification of the layer microstructure. The majority of sputtered particles have energies of a few electronvolts. However, "backscattered atoms" can have hundreds of electronvolts, which is sufficient to cause significant bulk displacements and crystalline defects. Backscattered atoms are most important during deposition of high-atomic-weight materials at low pressures. For example, in an Ar–Ta head-on atomic collision (during sputtering of a Ta-target with Ar), the momentum of the Ar is inverted while only a small amount of energy is transferred to the heavy Ta atom. If the Ar-ion loses its charge during this process (likely), it will move from the target to the substrate with an energy of several hundred electronvolts.

Bombardment of the substrate can be purposely enhanced by adding a negative "substrate bias," which accelerates ions from the plasma toward the substrate with a defined energy. In addition, the magnetic field lines during magnetron sputtering can

be designed so that some of them reach the substrate. This method is termed "unbalanced magnetron sputtering" and results in a high plasma density at the vicinity of the substrate. The ions in the high plasma density are then again accelerated toward the substrate with a bias, causing a high flux of ion irradiation with corresponding effects on the microstructure and texture of the deposited layers.

Chemical Vapor Deposition

Chemical vapor deposition (CVD) utilizes chemical reactions of volatile compounds with other gases to form nonvolatile solid deposits on a substrate. In contrast to PVD, where the vapor phase of the depositing species is obtained from a solid source, CVD employs gas-phase compounds that typically enter the deposition system through flux-controlled gas inlets. CVD is widely used in the semiconductor industry for the deposition of epitaxial (single-crystal) layers as well as various insulating and metallic thin films. Advantages of CVD include the possibility of uniformly depositing on complex-shaped surfaces, the relatively low cost for equipment and operation, and the compatibility with other processing steps. Many chemical reactions occurring during CVD can be divided into the following types:

1. *Pyrolysis*. It is the thermal decomposition of volatile compounds on hot substrates into a solid phase and gaseous by-products. An example is the decomposition of $SiH_4(g)$ into Si(s) and $H_2(g)$.
2. *Oxidation*. It is carried out using either O_2 gas or an O-containing compound, for example, during the deposition of alumina hard coatings by the reaction: $2AlCl_3(g) + 3H_2(g) + 3CO_2(g) \rightarrow Al_2O_3(s) + 3CO(g) + 6HCl(g)$.
3. *Reduction*. These reactions are typically obtained by employing H_2 gas to reduce halides and fluorides. For example, W-plugs, for contacts in integrated circuits, are deposited using $WF_6(g) + 3H_2(g) \rightarrow W(s) + 6HF(g)$.
4. *Displacement*. These reactions yield solid compounds. This category includes the very important metalorganic (MOCVD) processes such as those used for the growth of compound semiconductors. For example, GaAs is grown by the reaction: $(CH_3)_3Ga(g) + AsH_3(g) \rightarrow GaAs(s) + 3CH_4(g)$. The advantage of organic compounds as precursors is their high volatility at relatively low temperatures, which provides much better process control than techniques which require liquid or solid sources in the reactor. An example of the latter is the growth of GaAs by a reversible process that is achieved by $AsCl_3$ gas which reacts with molten Ga to form GaCl(g). This process requires the Ga source and the substrate to be kept at different temperatures within the same reactor to achieve the reaction: $As_4(g) + As_2(g) + 6GaCl(g) + 3H_2(g) \rightarrow 6GaAs(s) + 6HCl(g)$.

The thermodynamic data of the various chemical species during CVD processes are critical to assess the feasibility of new deposition schemes. In many cases, the gas and the surfaces are in quasiequilibrium; thus, the various partial pressures can be predicted from the changes in the Gibbs free energy of the relevant chemical reactions, and suitable temperatures for the substrate, chamber walls, and possible sources can be determined. Nevertheless, kinetic barriers for reactions and diffusion may limit reaction rates and lead to nonequilibrium conditions during operation since finite gas flow and deposition rates are essential for the ultimate goal of actually depositing a layer.

The understanding of gas transport during CVD processes is essential to design reactors that provide good utilization of process gases, high growth rates, and layer thickness uniformity. The latter requires an equal amount of reactant to all substrate surfaces, which is a nontrivial problem which typically requires computer modeling that includes gas transport as well as chemical reactions at the surface and in the gas phase. For most CVD systems, the operating pressures are high enough so that gas transport occurs in the viscous flow regime. The flow velocity is a critical parameter. If it is too high, it causes a turbulent flow and a corresponding nonuniform deposition. However, a low gas velocity causes a relatively thick boundary layer between the stationary substrate and the laminar stream. This reduces deposition rates since both gaseous reactants and by-products must pass through this boundary layer. Many CVD reactors exhibit zones which are kept at different temperatures and/or concentrations leading to gas transport by convection, in addition to the viscous transport and diffusion transport, which are due to pressure gradients and molecular motion, respectively.

There exists a multitude of CVD variations, as well as hybrid techniques that combine PVD and CVD processes. Some examples include low-pressure as well as atmospheric-pressure CVD, or plasma-enhanced CVD (PECVD). The latter process utilizes a glow discharge within the reactor to catalyze chemical reactions at low temperatures. Energetic collisions within the plasma cause chemically reactive ions, radicals, and molecular fragments which consecutively react at the substrate surface to result in layer growth. PECVD techniques are successfully

applied for the growth of nitrides, carbides, and amorphous group IV elements, including a-Si, and diamond-like C.

See also: Devices for Telecommunications; Diamond Anvil Cells; Film Growth and Epitaxy: Methods; Integrated Circuits; Light Emitting Diodes; Lithography Masks and Pattern Transfer; Magnetic Materials and Applications; Memory Devices, Volatile; Micromechanical Devices and Systems; Modulators, Optical; Optical Instruments; Quantum Devices of Reduced Dimensionality; Semiconductor Devices; Semiconductor Lasers; Sensors; Thermoelectric and Energy Conversion Devices; Thin Films, Mechanical Behavior of; Transistors; Vapor Transport Processing: Models and Simulations.

PACS: 81.15. – z; 81.15.Cd; 81.15 Gh

Further Reading

Galasso FS (1991) *Chemical Vapor Deposited Material*. Boca Raton, FL: CRC Press.
Glang R and Maissel LI (1970) *Handbook of Thin film Technology*. New York: McGraw-Hill.
Hoffman DM, Singh B, and Thomas JH (1997) *Handbook of Vacuum Technology*. Boston: Academic Press.
Mahan JE (2000) *Physical Vapor Deposition of Thin Films*. New York: Wiley.
Mattox DM (1998) *Handbook of Physical Vapor Deposition Processing*. Westwood, NJ: Noyes.
Ohring M (2002) *Materials Science of Thin Films*, 2nd edn. London: Academic Press.
Pierson HO (1992) *Handbook of Chemical Vapor Deposition*. Park Ridge, New Jersey: Noyes Publications.
Smith DL (1995) *Thin-Film Deposition*. Boston: McGraw-Hill.
Tsao JY (1993) *Materials Fundamentals of Molecular Beam Epitaxy*. Boston: Academic Press.
Vossen JL and Kern W (1979) *Thin Film Processes*. New York: Academic Press.
Vossen JL and Kern W (1991) *Thin Film Processes II*. New York: Academic Press.

Vapor Transport Processing: Models and Simulations

C Bernard and M Pons, LTPCM-ENSEEG, St. Martin d'Hères, France

Introduction

The aim of macroscopic modeling of reactors for growing thin films starting from a gas phase is to prepare the way for designing and optimizing them by linking the parameters of the apparatus itself to its performance (uniformity, quality, selectivity, etc. of the deposit). This involves mainly linking the properties of the film, which depend on its growth rate, composition, and microstructure, to the process parameters: activation mode, substrate temperature, pressure, composition, and flow rate of the gaseous reactant phase and geometry of the reactor. Modeling helps to understand the various phenomena that occur during thin film growth, as these phenomena cannot always be analyzed experimentally. The models are based on the fundamental laws of physics and chemistry and can be used to quantify the role of the various phenomena, visualize them, and modify them by altering operating procedures. They are a complement to experimental methods (diagnostics, preparation) and one of their objectives is to help in designing equipment.

This technique, referred to in this article as CVD (chemical vapor deposition), has been used since the early days of the microelectronics industry. Compared to other techniques (sputtering, sublimation, or evaporation) CVD is extremely flexible. It offers the possibility of checking the structure and composition of the film, obtaining good uniformity, and relatively high growth rates (a few $\mu m\ h^{-1}$). One of the important advantages of this technique is its ability to deposit "conformal" films (i.e., films of uniform thickness on irregular surfaces).

The examples of simulation presented in this article look essentially at generic cases corresponding to microelectronics applications, as this is one of the sectors in which there is the greatest demand. While the construction of new CVD reactors is still largely on the basis of trial and error, and on adjusting existing equipment to meet requirements, the present trend is to combine simulation models with an experimental knowledge in designing, optimizing, and evaluating new reactors.

After an introduction that looks at the various modeling methods, examples illustrating transfer and reactivity phenomena are discussed. The purpose of these examples is not to establish the state of the art with regard to numerical simulation, but to show various levels of complexity.

The models and simulations discussed are restricted to the thermal approach to CVD. Plasma and laser contributions are noted for information purposes only, as they would require further development for both the models and the reaction approaches.

The Various Ways of Modeling and Simulating the CVD Process

Principle

In the CVD process, the reactive molecules containing atoms of the material to be deposited are introduced in the form of gas precursors diluted in an inert carrier gas into a reaction chamber containing the substrates. The chemical reactions must be activated. Only thermal activation is discussed here.

A coating growth is the result of heterogeneous reactions. Controlling the process involves determining the species at the origin of the reactions and thus their history inside the reactor. Generally speaking, the various stages involved are as follows (**Figure 1**):

- convective and diffusive transport of the reactants from the entrance of the reactor to the reaction area,
- chemical reactions in the gas phase leading to intermediate species,
- convective and diffusive transport of the chemical species to the surface,
- adsorption of these species on the surface of the substrate,
- heterogeneous reactions catalyzed by the surface,
- desorption of the reaction products, and
- convective and diffusive transport of the products from the surface to the outlet from the reactor.

Macroscopic modeling methods draw on thermodynamic, kinetic, and transport databases and involve (1) thermodynamic equilibrium calculations, (2) kinetic calculations, and (3) heat and mass transfer calculations linked to the kinetic and/or thermodynamic calculations.

Thermodynamic Approach

The thermodynamic approach is an unavoidable stage in any method for modeling gas-phase growth processes. It is used to determine whether a material can be made from a gas phase and, based on this preliminary study, to choose the potential gas precursors and ranges of operating conditions. In most cases, thin films are prepared from the vapor phase by CVD in reactors that work at constant temperature and pressure. If it is assumed that equilibrium is reached in the reactor, it is the total Gibbs energy function of the chemical system in action that will be minimum. To determine the nature and proportions of the phases present at equilibrium, a description of the Gibbs energies of all the phases must be available. Indeed, by knowing the Gibbs energy of an element or compound, it is possible to derive all its other thermodynamic properties and in particular the most important for this type of study, namely entropy, enthalpy, and heat capacity at constant pressure (**Tables 1** and **2**).

The Gibbs energy is expressed as a function of enthalpy and entropy as follows:

$$G = H - TS$$

However, as neither can be described in an absolute manner, a reference state must be used to situate each of these state functions and describe them in the databases to be interrogated for the equilibrium calculation. The convention adopted for entropy involves taking a zero value for each phase at 0 K. For enthalpy, the convention adopted by the Scientific Group Thermodata Europe (SGTE) was chosen. The reference will be the enthalpy of the stable phase of the element in question at 298.15 K, under a pressure of 1 bar; this will be referred to as "HSER" for "enthalpy standard element reference." Hence the Gibbs energy ${}^0G_i^{\Phi}(T)$ for the pure element i, referred to as the enthalpy of its stable state Φ at 298.15 K: ${}^0H_i^{\Phi}$ (298.15 K) at temperature T will be written as

$$GHSER_i = {}^0G_i^{\Phi}(T) - {}^0H_i^{\Phi}\ (298.15\ \mathrm{K})$$

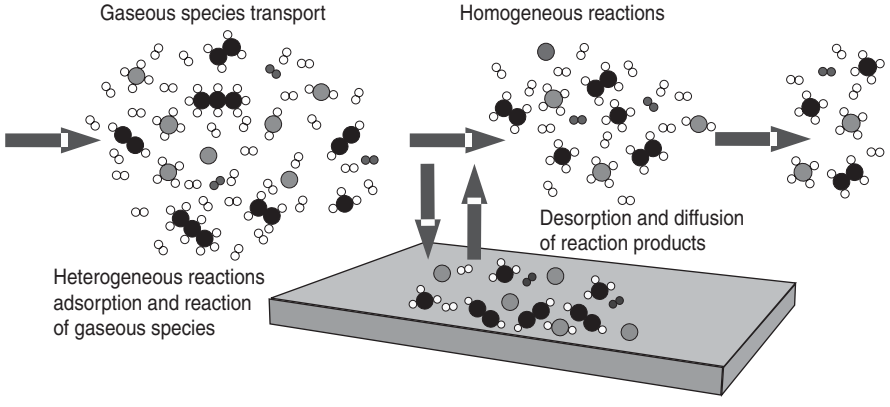


Figure 1 Schematic representation of phenomena leading to film growth.

Table 1 Thermodynamic and transport data

Transport data		*Thermodynamic data*	
λ :	Thermal conductivity ($W\,m^{-1}\,K^{-1}$)		
ρ :	Density ($kg\,m^{-3}$)		
μ :	Dynamic viscosity ($kg\,m^{-1}\,s^{-1}$)	${}^0G_i^{\alpha}(T)$:	Standard Gibbs energy of formation for species i of structure α ($J\,mol^{-1}$)
v :	Velocity ($m\,s^{-1}$)	${}^0H_i^{\alpha}$:	Standard enthalpy for species i of structure α ($J\,mol^{-1}$)
T :	Temperature (K)	c_p :	Heat capacity at constant pressure ($J\,mol^{-1}\,K^{-1}$)
P :	Pressure (Pa)	R :	Universal gas constant ($8.314\,J\,mol^{-1}\,K^{-1}$)
g :	Gravity ($g = 9.81\ m\ s^{-2}$)		
D_{ij} :	Binary diffusion coefficient ($m^{-2}\,s^{-1}$)		
D_i^T :	Multicomponent thermal diffusion coefficient ($kg\,m^{-1}\,s^{-1}$)		
ε :	Radiative emissivity		
ε/k :	Ratio of maximum energy of attraction in Lennard-Jones interaction potential and the Boltzmann constant (K)		
σ :	Collision diameter in Lennard-Jones interaction potential (A)		

Table 2 Thermodynamic data required for equilibrium modeling

Functions	*Formalism for n species*
Entropy	$S = -\left(\frac{\partial G}{\partial T}\right)_{P,n}$
Enthalpy	$H = G - T\left(\frac{\partial G}{\partial T}\right)_{P,n}$
Specific heat at constant pressure	$c_p = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_{P,n}$

Table 3 Polynomial description of the Gibbs energies of the various species involved in the CVD process

Species	*Gibbs energy referred to stable phase*
Stable structure Φ of element i	${}^0G_i^{\Phi}(T) - {}^0H_i^{\Phi}(298.15\ K) = f_i(T)$
Metastable structure φ of element i	${}^0G_i^{\varphi}(T) - {}^0H_i^{\Phi}(298.15\ K) = f_i(T)$
Pure gas i	${}^0G_i^{gas}(T) - {}^0H_i^{\Phi}(298.15\ K) = f_i(T) - RT\ln P$
Solid compound A_aB_b	${}^0G_{A_aB_b}^{\Phi}(T) - a\,{}^0H_a^{\Phi}(298.15\ K) - b\,{}^0H_b^{\Phi}(298.15\ K) = f_{A_aB_b}(T)$

The Gibbs energy function can then be stored in the databases in the form of a temperature-dependent polynomial: $f_i(T)$. The classic formula is that adopted by the SGTE:

$$GHSER_i = a + bT + cT\ln T + dT^2 + eT^3 + fT^{-1} + gT^{-7} + hT^{-9} = f_i(T)$$

The Gibbs energy function of the classic species found in CVD (elements, gases, and compounds) using the polynomial description is given in **Table 3**. In contrast, the variable-composition condensed phases, which fall into three groups (substitutional solid solutions, ordered phases, and phases with an order–disorder transformation) are described by more sophisticated models that are not discussed in this article.

Consequently, in a simple system with N constituents involving no solid solution phases, the total Gibbs energy of the system that has to be minimized can be calculated at any moment in the following way:

$$G_{tot} = \sum_{i=1}^{N} q_i(f_i(T) + d_i RT\ln P_i)$$

where q_i corresponds to the number of moles of the constituent i, and d_i has the value 1 for a gas of partial pressure P_i and 0 for a condensed phase.

This Gibbs energy, representative of all the phases present or likely to be present in the reactor, is then minimized in the space of the N constituents, taking care to maintain the mass balance of each of the elements present in the chemical system and reproducing the experimental conditions that are to be tested. This minimization provides information on the nature of the various phases present at equilibrium and their respective proportions. These calculations are all the more useful as they can be carried

out beforehand, which may give an initial idea of the feasibility of a process. Moreover, after the most promising chemical system has been adopted, these minimizations lead to the most suitable ranges of parameters.

This approach is based on restrictive assumptions. The first supposes that thermodynamic equilibrium is reached; the greater the deposition temperature and pressure or dwell time, the more this assumption is justified. The second is based on the fact that the reactor is assimilated to a closed system at fixed pressure and temperature. However, this approach means that it is not necessary to choose a set of reaction equations, in contrast to the kinetic approach. It is followed by using numerous programs associated with interactive databases.

The Kinetic Approach

Between the injection orifice and substrate in a reactor, the gas molecules interact to form new species that may react with the surface in very different ways from those of the initial molecules. In order to simulate the process, it is necessary to know the reaction paths followed by these molecules (paths consisting of elementary stages formed essentially of unimolecular or bimolecular reactions) and the associated kinetic constants. Understanding the reaction mechanisms means constructing, analyzing, and reducing the complex chemical systems involved. When the data have been obtained, for a fixed pressure and temperature, solving a set of ordinary differential equations provides a change in concentration as a function of time. It should be noted that an "infinite" interaction time gives values for the thermodynamic equilibrium representing the maximum efficiency of the process.

The main problems with this approach are validating the proposed reaction paths and the difficulty in obtaining the data. For example, for the deposition of simple elements such as silicon, several dozen simultaneous homogeneous and heterogeneous reactions have been identified and many reduced chemical models have been proposed. Setting up databases involves spectrometer measurements, quantum mechanics, and statistical thermodynamic calculations. Generally, the kinetic constants for homogeneous reactions are expressed by Arrhenius-type laws. Theoretical approaches similar to those used with homogeneous reactivity are only just emerging for heterogeneous reactions. Predicting and establishing surface reaction paths are more difficult in the heterogeneous phase than in the homogeneous phase as the surface entities are much more complex. An important step forward in the field of kinetic formalism was made with the Chemkin code from Sandia laboratories.

Heat and Mass Transfer Approach

In order to understand the dynamic nature of the process, thermodynamic and/or kinetic approaches must be linked to heat and mass transfer models. Transfer models are sets of differential equations and boundary conditions describing the flow, energy, and mass transfers and appearance/destruction of the species due to chemical reactions at a macroscopic scale. The macroscopic equations describing the various stages of the process are general and can be applied to a wide spectrum of processes. Many recent reviews have given the state of the art in this field. However, it is constantly changing in order to meet the need to understand new reactive systems. Major research is being carried out into the formulation of chemical reactivity models and transport databases. The former are independent of any particular geometry but must be completely integrated into the simulation cycle. A description of the reaction paths, in particular in plasmas, is the major factor limiting the applicability of these models. In contrast, nonreactive hydrodynamic models and purely thermal models have proved to be helpful in designing equipment.

Figure 2 gives a general view of the simulation strategies. It is clear that a single model or single modeling technique cannot provide access to the very wide field of operation of the equipment, and in particular, the plasma equipment. Generally speaking, each module provides access to a part of the physicochemical problem and the various modules must be connected with user-friendly interfaces.

The reactor models are used to calculate the velocity, temperature, and concentration fields of the neutral species associated with the experimental conditions.

Some of the aspects shown in **Figure 2** are supported by calculation codes available in the market. Calculation codes capable of simulating a large part of the transfer and reactivity models for thermally activated CVD are currently available in the market. For the present, they are an unavoidable point of departure for initiating new research in the field of CVD process engineering. Simulators can be found for plasma-assisted reactors that are adapted to experimental conditions and/or excitation modes. Linking these with simultaneous transfer models is one of the objectives of the numerous research projects in progress. While the links to be made between different codes theoretically pose no major scientific problem, the quantification of reactor

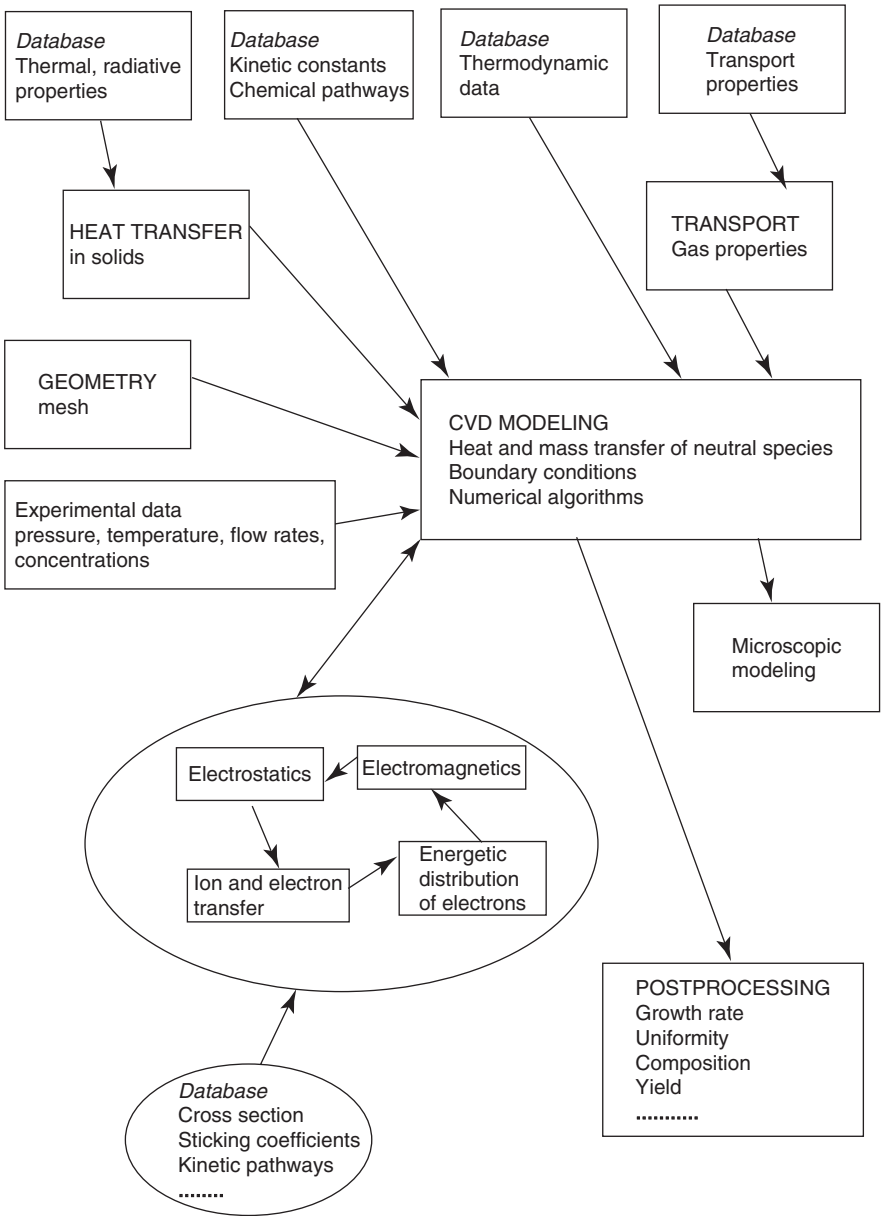


Figure 2 Schematic representation of reactor modeling strategies. The squares indicate the various models required to solve problems where only neutral species are taken into account (thermal CVD). The ellipses represent the additional modules for plasmas (plasma CVD).

performance and reactor optimization calls for close collaboration among specialists in process engineering, materials preparation, physics, and gas phase diagnostics. This is one of the reasons why few chemical systems have reached maturity. In practice, the models proposed were developed through intuitive approximations and past experience, and are based on experimental validation (*in situ* diagnostics of the gas phase and characterization of the layers obtained). However, *in situ* measurements are not usual in production reactors but can be performed in research reactors. Any quantitative or qualitative technique with enough spatial and temporal resolution can help to validate the proposed models. To make progress, it is essential to link these techniques to modeling efforts.

The core of the state of the art of CVD models includes mainly a model of transfer (or transport) phenomena, which must be supplemented by various submodels. The first is a chemical model describing the mechanisms of homogeneous and heterogeneous reactivity for a given process as a function of the preparation conditions. The second is a model based on the kinetic theory of gases for predicting the

transport properties of the gas mixture. The third is a thermal transfer model describing radiative effects coupled with the classical conductive and convective effects. Finally, the last model, which is still at the development stage, is a microscopic one linking calculated macroscopic values to film properties. All of these models require (1) thermodynamic, kinetic, thermal, and transport databases, (2) a description of the reactor geometry and operating conditions.

Heat and mass transfer and reactivity models are now well formalized. For some years now, the numerical encoding of these equations has been included in the main commercial codes that were previously dedicated to fluid mechanics calculations. Many physical and chemical models are included as standard ones. For example, the CFDACE code used for the calculations presented below includes most of the specific requirements of CVD reactor modeling. The transport properties, $\lambda, \mu, D_{ij}; D_i^T$, may be calculated directly from the kinetic theory of gases. A directly accessible thermodynamics database can be modified in order to include the missing specific gas species of the chemical system during the study. The code handles multicomponent gas mixtures, ordinary diffusion, and thermodiffusion. It is possible to describe homogeneous and heterogeneous reactions with Arrhenius-type or more complex laws, thanks to a link with the Chemkin code. Specific resolution algorithms for this type of reactivity are available. Thermal models include radiative transfer for gray bodies as well as for transparent ones. All the options are available via the user interface, which does not call for a detailed knowledge of the numerical techniques. However, when all these phenomena are added together, it is clear that there is need for caution. A knowledge of the models and their implications for the results is still essential.

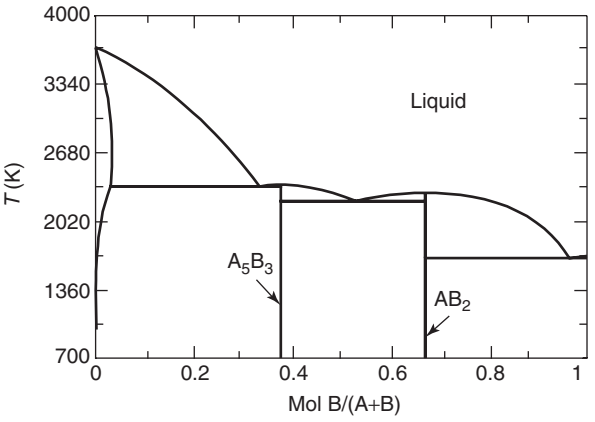


Figure 3 A–B phase diagram.

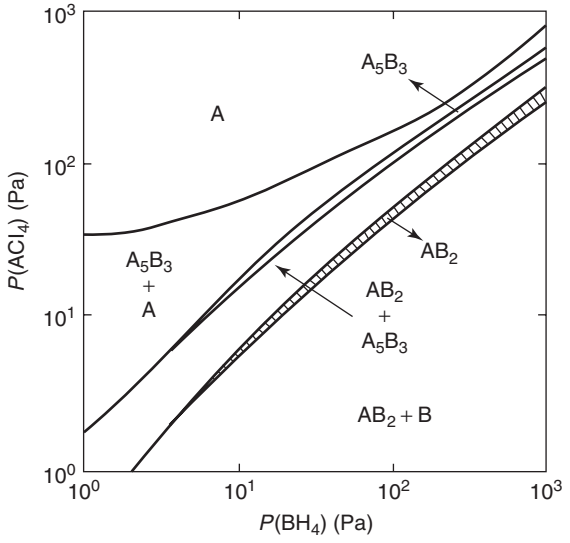


Figure 4 Calculated CVD phase diagram: $T = 1000$ K, $P_{tot} = 10^5$ Pa, $P_{Ar} = 9 \times 10^4$; initial gas mixture: $ACl_4 + BH_4 + H_2 + Ar$.

Databases and Examples

In order to be able to use the available software, the various data in the experimental temperature and composition field must be available. The data (**Table 1**) required to build a complete model are thermodynamic data (c_p, H, S), transport data $(\lambda, \mu, D_{ij}, D_i^T)$, kinetic data (k_r, k_{-r}), and radiative data (ε).

In the case of thermodynamic data, there are major databases and minimization software for calculating equilibrium in multicomponent systems. All these databases and software are interfaced with one another and can be interrogated online or installed on a PC. However, in spite of efforts by database managers, the thermodynamic values may not be completely consistent with one another, and the user will therefore always have to test their consistency at the outset in the context of his own chemical system.

As an example, the type of highly important classical information that an experimenter may hope to obtain from a simulation of his deposition experiment can be illustrated by the CVD diagram. Consider, for example, a compound AB_2 to be deposited on a substrate B using a gas mixture containing ACl_4, BH_4, H_2, and Ar (which could be the case of a deposit of silicon-rich conducting silicide on a silicon substrate). The compound AB_2 is situated in the binary diagram A–B shown in **Figure 3**, which has been optimized previously and simultaneously in the databases with the thermodynamic data of the phases concerned. Simulation of the CVD deposit at a given temperature T and fixed total pressure and neutral gas dilution leads to the diagram shown in **Figure 4**.

This indicates the nature of the phases deposited on the substrate, which is assumed to be neutral, as a function of the initial composition of the injected gas phase. A case such as this, where the domain favorable to the deposition of the phase AB_2 is fairly small, fully illustrates the interest of this type of approach.

More information on available databases and software can be obtained in particular from the (European, Canadian, and American) members of the SGTE.

With regard to transport data, most software include the calculation of these properties directly at each point of the mesh with a temperature and composition known during the calculation. The only data to be provided are the Lennard-Jones parameters $(\sigma, \varepsilon/k)$ and heat capacity c_p, which are used to calculate all the properties of the nonpolar molecules with sufficient accuracy.

The aim of the various examples presented below is to describe the transport and reactivity phenomena so as to demonstrate the potential of the simulation tools and not to propose an exhaustive overview of the problems and technological solutions. Classical geometry will be used for this purpose. It is an axisymmetrical, vertical reactor that has been widely used in research and manufacturing (**Figure 5**).

In this type of reactor, flow takes place perpendicular to the surface of the substrate. The substrate and its support may be fixed or rotating. Slow rotations ($\omega = 10$ rpm) are often used to avoid dissymmetry effects. Fast rotations ($\omega = 100-1000$ rpm) modify the transport. Reversed reactors, in which the gases enter through the bottom, are used to modify or eliminate natural convection phenomena. They are rarely used owing to the difficulties involved in installing the substrate.

The walls are usually cooled to avoid any deposition there. Lamps may heat the substrate directly in a few seconds. The various phenomena encountered when using a generic fast-heating reactor of the type shown in **Figure 6** are described.

Influence of Carrier Gas Pressure

A mixture containing silane ($100\,cm^3\,min^{-1}$ standard conditions) diluted in argon or hydrogen ($2000\,cm^3\,min^{-1}$ standard conditions) is introduced at a temperature of 300 K at the top of the reactor. The temperature of the substrate is fixed at 923 K and the walls are cooled (300 K). When the carrier gas is argon, natural convection phenomena due to gravity effects can be observed as soon as the pressure reaches 1.33×10^4 Pa (**Figure 7**). When hydrogen is used as a carrier gas, these only occur close to atmospheric pressure. This first simulation can be used to visualize and quantify the relations existing between the geometry, the nature of the gas, the temperature, and the flow facies.

The most complex thermal effects associated with the flow and reactivity phenomena may also be

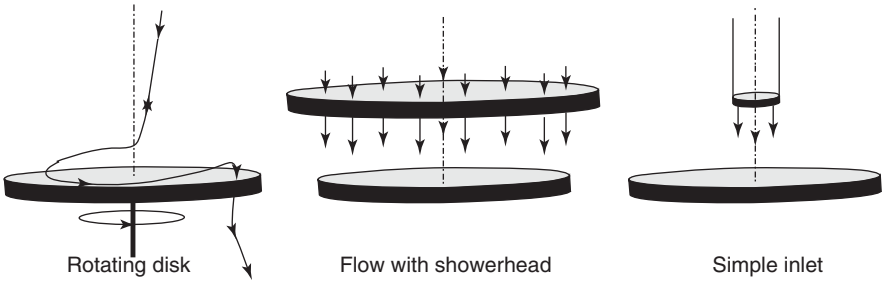


Figure 5 Schematic representation of the various technologies for axisymmetrical vertical reactors.

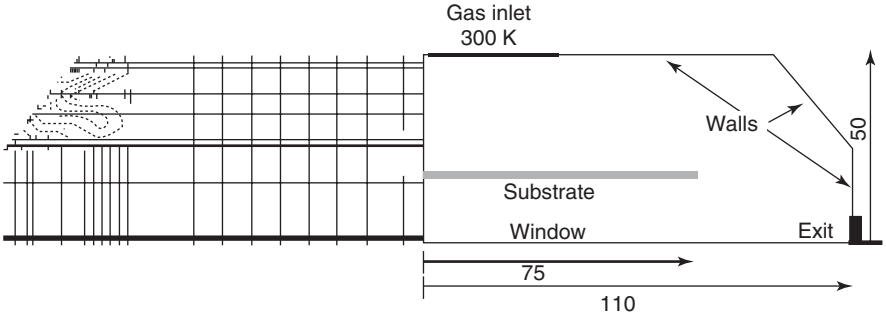


Figure 6 Schematic representation of the reactor and mesh used for the simulations (dimensions in mm).

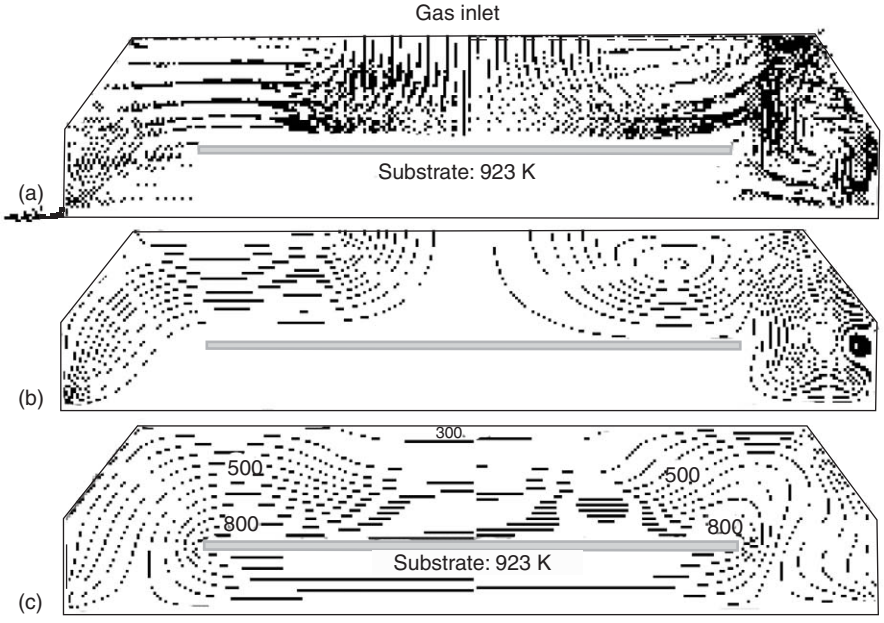


Figure 7 Influence of pressure on velocity field (a) streamlines, (b) and temperature fields, and (c) left – $P = 1333$ Pa, right – $P = 13\,333$ Pa. The initial mixture is Ar ($2000\,cm^{-3}\,min^{-1}$) and silane ($100\,cm^{-3}\,min^{-1}$).

simulated using available tools. However, it is not possible to calculate the growth rate and uniformity of the composition and thickness of the deposited film with this approach. One of the industrial challenges is to be able to check these factors. Multiconstituent mass transfers and homogeneous and heterogeneous reaction paths must therefore be added to the previous models. The deposition of silicon using silane is discussed as an example. The real chemical system has been reduced for pedagogical reasons and must not be used as such without a preliminary study.

Influence of Reactivity

The homogeneous reactions and associated kinetic constants (kmol, s, K) in the gas phase are

$$SiH_4(g) \rightarrow SiH_2(g) + H_2(g)$$
$$k_{1f} = 2.43810^{10} \exp\left(\frac{-24075}{T}\right)$$

$$SiH_2(g) + H_2(g) \rightarrow SiH_4(g)$$
$$k_{1r} = 1.80910^{5} \exp\left(\frac{3042}{T}\right)$$

$$Si_2H_6(g) \rightarrow SiH_4(g) + SiH_2(g)$$
$$k_{2f} = 1.58410^{12} \exp\left(\frac{-20063}{T}\right)$$

$$SiH_4(g) + SiH_2(g) \rightarrow Si_2H_6(g)$$
$$k_{2r} = 7.77110^{7} \exp\left(\frac{5120}{T}\right)$$

The two gas species SiH_2 and Si_2H_6 are produced by homogeneous reactions.

The heterogeneous reactions and sticking coefficients (γ) of the gas species leading to silicon deposition are

$$SiH_4(g) \rightarrow Si_s + 2H_2(g)$$
$$\gamma_1 = 1.0810^{8} T^{-0.8822} \exp\left(\frac{-19052}{T}\right)$$

$$Si_2H_6(g) \rightarrow 2Si_s + 3H_2(g)$$
$$\gamma_2 = 4.9610^{16} T^{-3.253} \exp\left(\frac{-20219}{T}\right)$$

$$SiH_2(g) \rightarrow Si_s + H_2(g) \quad \gamma_3 = 1.0$$

Figure 8 represents the temperature and mass fraction fields of silane, silylene, and disilane. Silylene is formed at the highest temperatures and as its reactive power is very great, the entire mass formed is consumed to prepare the silicon film. Disilane appears at lower temperatures and as it is less reactive, its concentration increases. The contributions of silylene and disilane are only of the order of 1% for this process. In the field of microelectronics, these variations may have a harmful effect on circuit preparation.

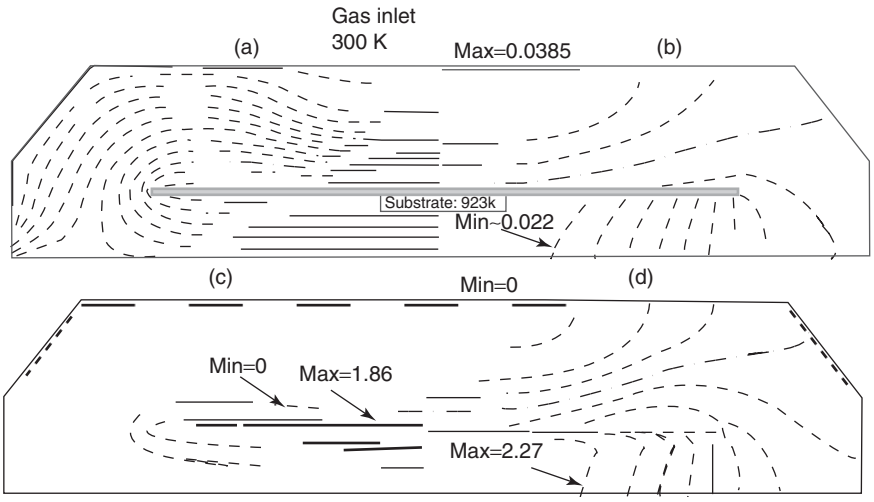


Figure 8 (a) Temperature field, (b) mass fraction field for silane, (c) silylene ($\times 10^6$), and (d) disilane ($\times 10^5$).

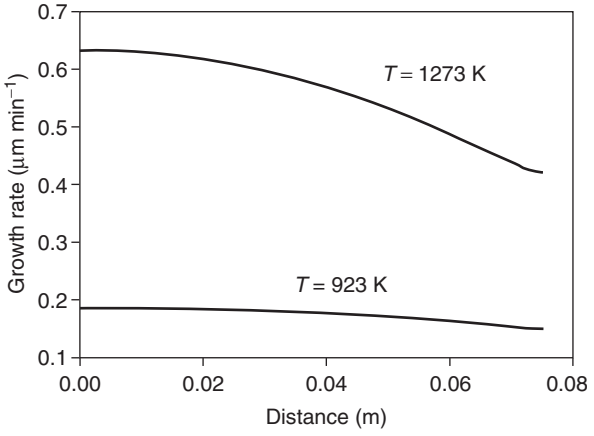


Figure 9 Growth rate as a function of substrate radius.

The process may be optimized on the basis of these results. For example, a reduction in pressure from 1333 Pa to 133 Pa leads to uniform deposits but the growth rate decreases. The same is true with a reduction in temperature (**Figure 9**).

Conclusion

The various models that can be built to simulate the phenomena occurring during film preparation by thermal CVD have been discussed. Similarly, the information required to use the models has been listed and illustrated with generic examples that reflect only a small proportion of the potential of such simulations. In practice, the phenomena are examined through a large number of interconnected submodels: thermodynamics, kinetics, heat transfers, mass transport, etc. Models of nonreactive transfers of momentum and energy are perfectly able to reveal the different modes of hydrodynamic operation of reactors. Models that take into account kinetic phenomena, even in a simplified manner, can be used for a more detailed simulation of all the mechanisms involved in the CVD process. Simulation models combining heat and mass transfers and local thermodynamic equilibrium extend thermodynamic calculations to dynamic systems. They only simulate experimental conditions for which equilibrium is reached at all points of the reactor, which is justified when the temperature is high and/or for long dwell times.

A large corpus of knowledge in terms of databases and simulations can be used regularly by specialists able to assess their scope, evaluate the contribution of the various modeling techniques, and integrate them in an overall program whereby they can be validated.

In the case of thermally activated CVD, the development of simulation tools and databases is the result of much research that has been carried out over the past 20 years. In contrast, much less attention has been paid to noncontinuous phenomena occurring at very low pressure or to modeling PECVD (plasma-enhanced chemical vapor deposition) and photo-CVD reactors. The development of these models is much more recent and simulation tools have not yet reached the same degree of sophistication.

One of the challenges for the future lies in the development of microscopic models and their links with the macroscopic approach presented here.

See also: Mass Transport; Specific Heat; Thermodynamic Properties, General; Vapor Transport Processing: Mechanisms.

PACS: 81.10.Bk; 81.15.Gh; 47.70.Fw

Further Reading

Centre de Recherches en Calcul Thermodynamique, Ecole Polytechnique de Montreal, Quebec, Canada, H3C 3A7, 1999 (http://www.crct.polymtl.ca/fact/fact.htm).

CFD-ACE is a product of CFD Research Corp., Hunstville, AL, USA, 2003 (http://www.cfdrc.com).

Chemkin-III: Reaction Design 11436 Sorrento Valley Road, San Diego, CA 92121, USA, Web: www.ReactionDesign.com, Email: chemkin@ReactionDesign.com

Clark TA (1985) *A Handbook of Computational Chemistry*. New York: Wiley.

Fluent is a product of Fluent Inc., Lebanon, NH, USA (http://www.fluent.com).

GTT, Kaiserstrasse 100 - D 52134 Herzogenrath, Germany, 1999 (http://gttserv.lth.rwth-aachen.de/).

Gurvich LV, Veyts IV, and Alcock CB (1989) *Thermodynamic Properties of Individual Substances*, 4th edn. New York: Hemisphere.

Hitchmann ML and Jensen KF (1993) *Chemical Vapor Deposition*. San Deigo: Academic Press.

Jensen KF (1989) Chemical vapor deposition, in microelectronics processing. In: Hess DW and Jensen KF (eds.) *Advances in Chemical Series*, pp. 199–263. Washington DC: American Chemical Society.

Kleijn CR (1995) Chemical vapor deposition processes. In: Meyyappan M (ed.) *Computational Modeling in Semiconductor Processing*, pp. 97–229. Boston, MA: Artech House.

Melange/Gémini is a product of LTPCM- Thermodata, BP 75 38402 Saint Martin d'Hères, France, 1999.

NIST Database Chemical Database, Version 4.0, 1992, Gaithersburg, MD, USA.

NPL, Queens Road, Teddington, Middlesex, TW11 0LW, UK, 1999 (http://www.npl.co.uk/).

Royal Institute of Technology, S - 10044, Stockholm, Sweden (http://www.mse.kth.se/tc/).

Thermodata - INPG - CNRS, BP 166, 38402 Saint Martin d'Hères, France, 1999 (http://www.cpma.u-psud.fr/therma/thermafr.html).

Vibrational Modes *See* Lattice Dynamics: Vibrational Modes.

W

Waves in Random Media

A Z Genack, Queens College of the City University of New York, Flushing, NY, USA
A A Chabanov, University of Minnesota, Minneapolis, MN, USA
P Sebbah, CNRS and Université de Nice-Sophia Antipolis, Nice, France
B A van Tiggelen, CNRS and Université Joseph Fourier, Grenoble, France

Overview

Optical scattering is ever present in the disordered world that surrounds us. It gives rise, for example, to the diffuse blue illumination of the sky and the whiteness of clouds, milk, and paint. Together with emission, absorption, and free propagation, optical scattering governs the light presented to the eye and to photographic and electrooptic imaging systems. In this article, the propagation of light and other parts of the electromagnetic spectrum within disordered media is described. Those aspects of scattering that are common to electromagnetic waves from the radio frequency to the X-rays are emphasized. Many of the illustrations are drawn from microwave measurements for which it is possible to determine both the amplitude and phase of the wave. Since the precise configuration of a random sample is by definition unknown, a statistical approach is required, as is the case in the study of electronic waves in conductors.

The observed electromagnetic field can be represented by the complex number, $A_n \exp(i\varphi_n)$, with amplitude A_n, phase φ_n, and discrete polarization index n. Though the electromagnetic field in static multiple-scattering samples varies randomly in space due to the superposition of randomly phased partial waves, the field at a given point is as temporally coherent as the source, which may be a highly coherent laser beam. The interference of scattered laser light produces a speckled intensity pattern with a correlation length of the field on the order of the wavelength, λ. The intensity and phase within a microwave speckle pattern on the output surface of a random sample is shown in **Figure 1**. Since the field at a point is the sum of a large number of statistically independent terms, associated with various scattering trajectories, it must be a Gaussian random variable, with a Gaussian distribution for the complex field. Since the intensity transmission coefficient from some incident mode a to some outgoing mode b is proportional to the square of the field, $T_{ab} = A_{ab}^2$, the probability distribution of normalized, polarized intensity may be expressed as $P(s_{ab}) = \exp(-s_{ab})$, where $s_{ab} = T_{ab}/\langle T_{ab} \rangle$ is the transmitted intensity normalized to its average value over an ensemble of random sample realizations, represented by $\langle \cdots \rangle$.

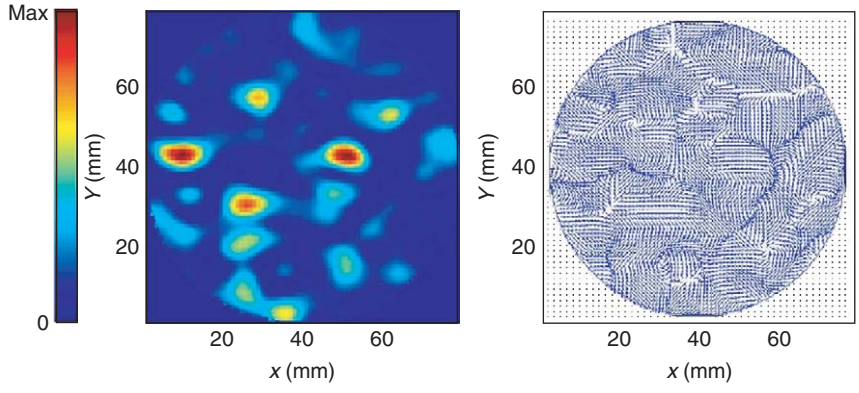


Figure 1 The near-field intensity speckle pattern of microwave radiation and the distribution of the phase at the output surface of a random sample (cylindrical waveguide filled randomly with dielectric spheres).

Early studies of multiply scattering light, a century ago, concentrated on optical flux in stellar and planetary atmospheres. These studies utilized the radiative transfer theory, which is akin to the phenomenological Boltzmann approach, used for computing atomic scattering or electron transport. Until two decades ago, optical studies dealt almost exclusively with weakly scattering natural and synthesized materials. In these media, the optical transport mean free path, ℓ, in which the direction of light is randomized, greatly exceeded the optical wavelength, λ. The evolution of photon density or intensity was shown to satisfy a diffusion equation, which gave a result for average transmission which is equivalent to Ohm's law in electron conduction. The diffusion equation also provided the photon time-of-flight distribution, which determined the sensitivity of the steady-state field to internal motion of the sample or variations of the incident frequency. Similarly, the diffusion equation determined the spatial intensity distribution on the input surface, which determined coherence of the reflected light.

The neglect of interference in calculations of average transport led to the diffusion equation, while the neglect of long-range spatial correlation by the assumption of Gaussian statistics resulted in the Rayleigh probability distribution for intensity. It has become increasingly apparent in the last two decades that the role of interference in average transport and of extended spatial correlation in intensity statistics cannot be neglected in multiple-scattering samples in which wave trajectories cross. As a result of these effects, flux does not self-average within a single sample configuration. Giant fluctuations appear in the total transmission, which is defined as the sum of transmission coefficients over the output modes for a given incident mode, $T_a = \sum_b T_{ab}$. Enhanced fluctuations of total transmission are due to the crossing of wave trajectories, which are also responsible for the phenomenon of universal conductance fluctuations. At the same time, average optical transport may be dramatically modified by the approach to photon localization, which is analogous to electron localization in many respects.

Optical and microwave measurements showed that the presence of long-range correlation of intensity within a sample led to enhanced fluctuations of total transmission over the value predicted if fluctuations in intensity were short range as are the fluctuations in the field. This represents a breakdown of Gaussian statistics of the field which leads to enhanced fluctuations in intensity. The presence of enhanced intensity fluctuations can be seen in an ensemble of quasi-1D random samples, in which the sample length is much greater than the diameter of its cross section. In such quasi-1D samples, modes are completely mixed so that the statistics of the field are the same at every point of the output surface. Though the field distribution on the output surface over a random ensemble of quasi-1D samples is no longer Gaussian, the field for the subset of samples with the same value of total transmission is a Gaussian random variable. The intensity distribution for a random ensemble with a distribution of values of total transmission can thus be regarded as a mixture of Rayleigh distributions with mixing proportions given by the distribution of transmission. As a result, non-Gaussian statistical properties of propagation can be characterized by the variance of the total transmission normalized to its ensemble average value, $\mathrm{var}(s_a) = \langle (\delta s_a)^2 \rangle$, where $\delta s_a = s_a - \langle s_a \rangle$ and $s_a = T_a / \langle T_a \rangle$. The value of $\mathrm{var}(s_a)$ is equal to the degree of correlation of transmitted intensity at remote points on the output surface or for distinct transverse transmission modes, b and b', at which the correlation in the field vanishes: $\mathrm{var}(s_a) = \langle \delta s_{ab} \delta s_{ab'} \rangle \equiv \kappa$. In systems without dissipation or gain, the degree of correlation κ is inversely proportional to the dimensionless conductance, g, which is the key parameter in electronic conductance and may be defined for classical waves in analogy with the Landauer formula for electronic conductance, $g = \sum_{ab} T_{ab} = \sum_a T_a$, $\kappa = 2/3 \langle g \rangle$.

Intensity correlation is large when the dimensionless conductance is small. The divide between extended diffusing waves and spatially localized waves is estimated to occur at $g \approx 1$ and $\kappa \approx 1$, where henceforth $\langle g \rangle$ will be denoted by g. The value of g is determined by the local strength of scattering and the dimensions of the sample. Localization occurs when wave trajectories are likely to return to a coherence volume, $V_c = (\lambda/2)^d$, in a d-dimensional sample before escaping from the sample. Wave interference then dominates transport and further enhances the return of the wave to points within the medium. Since the average number of returns of a wave to a coherence volume always exceeds unity in sufficiently large one- and two-dimensional scattering media, photon localization may be expected and indeed has been observed. In 3D samples, photon localization is more difficult to achieve. The threshold for localization is reached when $k\ell < 1$, where $k = 2\pi/\lambda$ is the wave vector. This is the same condition as the Ioffe–Regel criterion for electron localization.

Photon localization has been predicted and observed in nearly periodic 3D samples with appropriate structure such that a photonic bandgap would have been formed in the absence of disorder. Within the photonic bandgap, the wave is evanescent and propagation is forbidden in any direction. When the

sample is disordered, localized states are introduced into the bandgap, in a manner analogous to the occurrence of localized electronic states within the bandgaps of semiconductors and insulators. Key statistical properties of electromagnetic waves in samples with homogeneous disorder depend upon the closeness to the photon localization transition, which can be precisely determined using statistical measures of fluctuations and correlation of optical intensity. Observing the photon localization transition in 3D samples allows the study of the Anderson localization transition between extended and localized waves without the complication of particle interactions, which are unavoidable for electrons.

A crucial indicator of the nature of wave propagation is the variation of average transmission with a sample scale. For diffusive waves in energy-conserving systems, transmission falls inversely with the sample thickness, whereas it falls exponentially for localized waves. Contrary to the case of electronic transmission, in which the particle number is conserved, scaling of optical transmission is strongly affected by the degree of absorption or gain in a sample. When optical absorption plays a role, diffuse transmission falls exponentially as well, so that for light, scaling measurements alone are not sufficient to determine the localization threshold.

The optical speckle pattern gives a distinctive fingerprint of the coherence of the light scattered from a particular realization of a random sample (**Figure 1**). This is also revealed by an interference process called "weak localization" in which the diffuse nature of radiation is maintained, though with a suppressed diffusion coefficient. A strongly related process can be directly observed in retro-reflection: when averaged over an ensemble of configurations, a coherent backscattering peak emerges in the angular distribution of the reflected light, with a maximum intensity equal to exactly twice the background intensity. Ensemble averaging can be readily performed in colloidal samples. The angular profile of the coherent backscattered light is the Fourier transform of the intensity distribution on the input surface for point illumination, which has a spatial extent of the order of the mean free path ℓ. The diffuse random walk of photons generates a characteristic triangular coherent backscattering peak, whose width is inversely proportional to $1/k\ell$.

The study of multiple scattering in dynamic turbid samples opens up the possibility of probing the motion inside the sample. The transmitted speckle pattern is completely randomized when the change in the total path of the scattered waves is of order λ. Since the length of the wave trajectory is greatly increased within a random medium, even small rearrangements of scatterers modify the speckle pattern. Measurements of temporal correlation, called diffuse wave spectroscopy, make it possible to monitor microscopic motion on the length scale of nanometers and on the timescale of nanoseconds.

Apart from the similarities between optical and electronic transport arising from their dual wave and particle nature, there are numerous differences between light and electrons. Photons have zero mass, making their dispersion law linear, rather than parabolic. This has strong consequences for the speed of photons in resonant media. Photons also have no charge and many electronic phenomena such as the Aharonov–Bohm effect have no photonic equivalent. The absence of photon charge makes it more difficult to localize photons than electrons in random 3D samples. Nonetheless, a Hall effect for photons can be predicted and observed. Another striking difference applies to the optical scattering cross section relative to the geometric cross section, which falls rapidly with particle size for particles smaller than a wavelength. The strength of scattering is optimized at Mie resonances in larger particles. The storage of energy at wavelengths near Mie resonances slows the propagation of light. A final important difference is related to the spin. Whereas fermionic electrons are repelled by the Pauli exclusion principle, photons are bosons and can exhibit stimulated emission and lasing. The study of random lasers provides a probe of long-lived modes of a random medium in which lasing is facilitated.

Radiative Transfer and Diffusion

The phenomenological radiative transfer equation is a Boltzmann equation for the specific intensity $I_v(r,t)$ describing the flow of photons in a disordered medium at position r, at time t, and with velocity $\boldsymbol{v} = c\hat{\boldsymbol{v}}$,

$$\partial_t I_v + \boldsymbol{v}\cdot\nabla_v I_v + \left(\frac{1}{\tau_{\text{abs}}} + \frac{1}{\tau_{\text{scat}}}\right) I_v = S(\boldsymbol{r},t) + \int \mathrm{d}\hat{\boldsymbol{v}}'\, W(\boldsymbol{v},\boldsymbol{v}') I_{v'} \qquad [1]$$

This equation features all relevant processes: propagation (first two terms on the left), absorption (rate $1/\tau_{\text{abs}}$) and scattering to other directions (rate $1/\tau_{\text{scat}}$) collectively referred to as extinction, emission $S(\boldsymbol{r},t)$, and scattering from other directions (rate $W(\boldsymbol{v},\boldsymbol{v}')$ per solid angle).

Absorption and emission are incoherent processes, whereas scattering is, in principle, coherent. A classical limit of radiative transfer is to neglect scattering and to assume steady (local) thermal equilibrium between absorption and re-emission, as expressed by

Kirchhoff's theorem: $\tau_{abs} \times S = B_\nu(T)$, with $B_\nu(T)$ the Planck function for the local temperature $T(\boldsymbol{r})$. The approximate solution of the radiative transfer eqn [1] becomes,

$$I_\nu(\boldsymbol{r}) \approx B_\nu[T(\boldsymbol{r} - \boldsymbol{v} \cdot \tau_{abs})] \qquad [2]$$

showing that photons emerge from within one absorption length $v\tau_{abs}$ along the line of sight. This statement explains the limb darkening of the sun, and the solar absorption lines at the Balmer series of hydrogen. A modification of this approach also applies to photon diffusion in warm resonant gases.

The source, which may be a coherent laser beam, is often not in thermal equilibrium with absorption in the medium. It then becomes important to include scattering. Scattering integral renders the radiative transfer equation difficult to solve. A popular simplification is the diffusion approximation. The angular dependence of I_ν is expanded only up to the isotropic and dipolar terms, which describe the energy density, $E(\boldsymbol{r}, t) = c^{-1} \int \mathrm{d}\hat{\boldsymbol{v}}\, I_\nu(\boldsymbol{r}, t)$, and the energy current density, $\boldsymbol{J}(\boldsymbol{r}, t) = \int \mathrm{d}\hat{\boldsymbol{v}}\, \hat{\boldsymbol{v}} I_\nu(\boldsymbol{r}, t)$:

$$I_\nu = \frac{1}{4\pi}[Ec + 3\hat{\boldsymbol{v}} \cdot \boldsymbol{J} + \cdots] \qquad [3]$$

This leads to the following diffusion equation for the energy density,

$$\partial_t E(\boldsymbol{r}, t) - D\nabla^2 E(\boldsymbol{r}, t) + \frac{1}{\tau_{abs}} E(\boldsymbol{r}, t) = \frac{4\pi}{c} S(\boldsymbol{r}, t) \qquad [4]$$

The diffusion equation describes a random walk of photons with an average step length $\ell_{scat} = c\tau_{scat}$, and an angular redistribution described by $W(\hat{\boldsymbol{v}}.\hat{\boldsymbol{v}}')$ for each step. Scattering is now entirely described by the diffusion constant, D. It is found that $D = 1/3c^2\tau_{tr}$, with the transport mean free time,

$$\tau_{tr}^{-1} = \int \mathrm{d}\hat{\boldsymbol{v}}' W(\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{v}}')[1 - \hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{v}}'] = \tau_{scat}^{-1}[1 - \langle \cos\theta \rangle]$$

Equation [4] can only be solved if boundary conditions are specified. One usually imposes the condition that no current enters the medium. An important application is to the stationary transmission of light through a disordered, nonabsorbing slab between $z = 0$ and $z = L$ and an imbedded source layer, $S(z) = (c/4\pi)\delta(z - z_s)$. In the absence of internal reflection at the sample boundary due to a mismatch between the average refractive indices, the angular distribution of transmission is given by

$$T(\theta) = \frac{z_s + z_0}{L + 2z_0} \times \left(1 + \frac{3}{2}\cos\theta\right) \qquad [5]$$

where $z_0 = \frac{2}{3}\ell = \frac{2}{3}c\tau_{tr}$ is the extrapolation length beyond the sample interface, at which the intensity extrapolates to zero and τ_{tr} is the transport mean free time. This length increases when there is internal reflection at the sample interface due to a mismatch of effective refractive indices. The diffuse transmission decays with increasing thickness as $1/L$.

Dynamics

In weakly scattering systems, an initial ballistic burst of unscattered radiation is transmitted through the sample followed by a broad pulse with a profile described by the diffusion equation (**Figure 2**). The ballistic light, which must be filtered from the predominant multiply scattered light, is utilized in optical coherence tomography to image inhomogeneous tissues for noninvasive medical diagnostics and to monitor a wide spectrum of human pathologies based upon residual coherence in scattered light. Diffusive light rapidly reaches an asymptotic single exponential decay reflecting the sum of the leakage rate of light out of the lowest diffusion mode and absorption. In pulsed microwave measurements, in quasi-1D random dielectric media, a breakdown of the diffusion model is seen in the nonexponential decay of transmission at long time. This may be understood as the result of increasing impact of weak localization as the rate of crossing of optical trajectories within the sample increases in time. Alternatively, it may be understood as a manifestation of the increasing role of

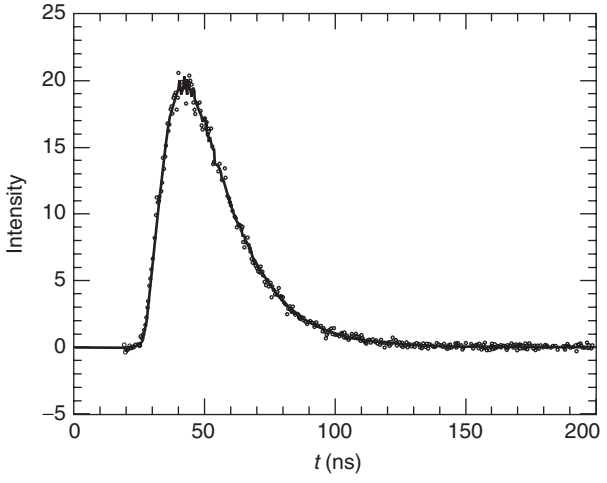


Figure 2 Time-of-flight distribution in an ensemble of quasi-1D samples of 1.3 cm diameter polystyrene spheres randomly packed in a copper tube of diameter 7.3 cm and length 100 cm with reflecting walls and open ends at a spheres volume fraction of 0.52 is obtained by averaging the temporal response to a 1 ns pulse at carrier frequency 19 GHz over 4096 sample configurations. The continuous line is the prediction of diffusion theory. (Sebbah *et al.* (2000) *Physical Review E* 62: 7348.)

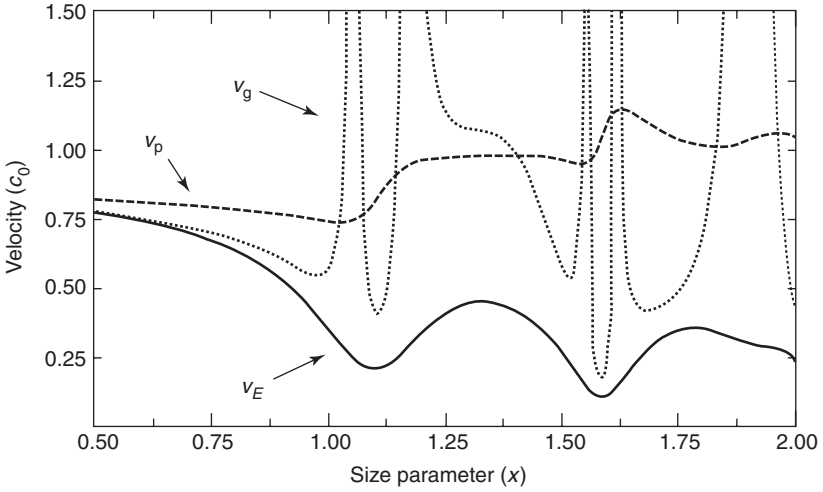


Figure 3 Energy transport velocity calculated for light propagating in a random medium filled with dielectric Mie spheres with refraction index 2.73 and volume fraction 0.36, as a function of their size a ($x = 2\pi a/\lambda$). The energy velocity v_E differs from the group velocity v_g and phase velocity v_p of the effective medium, especially near Mie resonances. (van Albada *et al.* (1991) *Physical Review Letters* 66: 3132.)

longer-lived modes at later times. Such modes tend to have greater density far from the sample surface so that the wave does not readily escape the sample.

Many wave phenomena can be explained in terms of a classical scalar wave equation for the complex wave $\psi = \mathrm{A}\exp(\mathrm{i}\varphi)$ in the presence of a source S,

$$\frac{\varepsilon(\boldsymbol{r})}{c_0^2}\partial_t^2\psi(\boldsymbol{r},t) - \nabla^2\psi(\boldsymbol{r},t) = S(\boldsymbol{r},t) \qquad [6]$$

The energy and current densities are expressed, respectively, as

$$E = \tfrac{1}{2}c_0^{-2}\varepsilon(\boldsymbol{r})|\partial_t\psi|^2 + \tfrac{1}{2}|\nabla\psi|^2$$

and

$$\boldsymbol{J} = \mathrm{Re}(\partial_t\psi\nabla\psi^*)$$

For a wave packet that is closely centered around angular frequency ω, and wave vector $\boldsymbol{k}$, $E \approx \varepsilon(\boldsymbol{r})(\omega/c_0)^2|\psi|^2$ and $\boldsymbol{J} \approx \boldsymbol{k}(\omega/c_0)|\psi|^2$, and the phase shift seems to disappear explicitly from the picture.

Though the velocity of light might be defined as a fluctuating quantity in a random medium, $c(\boldsymbol{r}) = c_0/\sqrt{\varepsilon(\boldsymbol{r})}$, where $\varepsilon(\boldsymbol{r})$ is the fluctuating local dielectric constant that gives rise to optical scattering, it is expected that, on a macroscopic level, the diffusion constant may be expressed in terms of an average transport velocity, v_E, $D = \frac{1}{3}v_E\ell$. The transport velocity may be taken as the ratio of the average current density and average energy density, calculated for a wave packet centered at frequency ω and wave number $\boldsymbol{k}$,

$$v_E = \frac{\langle J(\boldsymbol{r})\rangle}{E(\boldsymbol{r})} = \frac{kc_0^2}{\omega}\frac{\langle|\psi|^2\rangle}{\langle\varepsilon(\boldsymbol{r})|\psi|^2\rangle} \qquad [7]$$

$\boldsymbol{k}$ and ω are expected to be related by some effective medium dispersion law $\boldsymbol{k} = \omega/v_p$, featuring the phase velocity v_p. Near scattering resonances, the field inside the particles is greatly enhanced, and $v_E \ll v_p$. **Figure 3** shows a calculation of v_E for particles with an index of refraction $m = 2.73$ and volume fraction of 0.36, corresponding to TiO_2 powder. The transport velocity is seen to drop considerably near resonance.

Mesoscopic Optics

In the classical radiative transfer theory, light has been treated as photons, and not as classical electromagnetic waves. However, wave aspects of propagation are crucial if one is to understand fluctuations and correlations in multiply scattering samples in which the wave is temporally coherent. Because the regime, in which the wave is temporally coherent in electronic samples, is intermediate between the microscopic atomic scale and the macroscopic scale, it is termed mesoscopic. This term has been used to describe propagation of temporally coherent classical waves even in macroscopic samples.

Field Correlation

A full description of transport must include the interference of waves. A speckle pattern results from

the superposition of all partial waves arriving at different points of the scattered wave (**Figure 1**). The field fluctuates with the variation of some parameter p such as position, scattering angle, frequency, or time, as a result of the random phases and amplitudes of partial waves, and can be written as $\psi(p) = \psi_1(p) + \psi_2(p) + \cdots$. It is reasonable to assume that all $\psi_i(p)$ are independent, since they have visited so many different scatterers. As a result, the central limit theorem imposes complex-valued Gaussian statistics on the measured field $\psi(p)$. In addition to the Rayleigh distribution for the intensity, a second consequence of Gaussian statistics is the Siegert relation between correlations of intensity and field,

$$\begin{aligned}\langle I(p_1)I(p_2)\rangle &= \langle \psi(p_1)\psi(p_1)^*\psi(p_2)\psi^*(p_2)\rangle \\ &= \langle I(p_1)\rangle\langle I(p_2)\rangle \\ &\quad + |\langle \psi(p_1)\psi^*(p_2)\rangle|^2 \end{aligned} \quad [8]$$

The field correlation usually decays rapidly with $p_1 - p_2$. The Seigert relation is useful since intensity correlations are easier to measure than field correlations, particularly at optical frequencies.

Under a wide variety of circumstances, wave and particle aspects of propagation are Fourier transform pairs. An important example concerns the field correlation function C_E with frequency shift ν, which is the Fourier transform of the "photon time-of-flight distribution" $P(t)$ for the random ensemble,

$$P(t) = \langle |\psi(t)|^2 \rangle = \int \mathrm{d}\nu\, C_E(\nu)\exp(-\mathrm{i}2\pi\nu t) \quad [9]$$

A demonstration of this relation in microwave measurements is shown in **Figure 4**. The spectral width of the field correlation function can be identified with the width of leaky diffuse modes of the sample, and is associated via eqn [9] to the inverse of the mean photon dwell time in the sample. In a manner analogous to eqn [9], the field correlation function with a shift in wave vector, associated with an angular shift in the far field, equals the Fourier transform of the photon intensity distribution on the surface. When both the incident and scattered waves are shifted by the same angle, as when the sample is rotated, the field correlation function with angular shift equals the point spread function, which is the spatial distribution of light on the sample surface associated with excitation at a point. In this case, the field correlation in reflection gives the same angle variation found in the coherent backscattering peak. Either the intensity correlation with scattered angle or the intensity distribution with position on the input surface can be analyzed within an appropriate

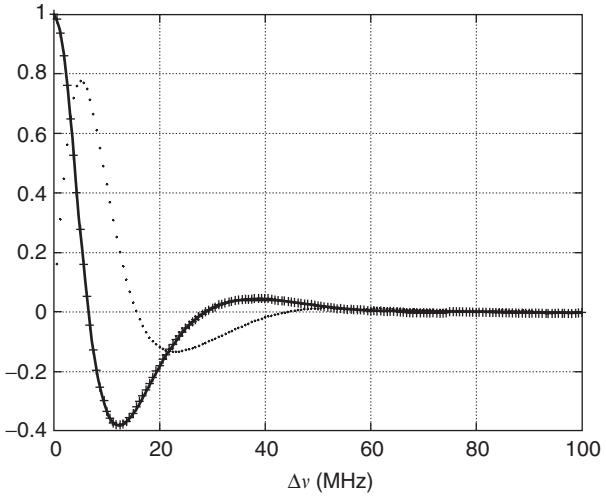


Figure 4 Real and imaginary parts of the field correlation function with frequency shift of the microwave field transmitted in an ensemble of random polystyrene samples in the range 18.5–19.5 GHz. The solid line is the real part of the Fourier transform of the time-of-flight distribution shown in **Figure 2**. (Sebbah *et al.* (2000) *Physical Review E* 62: 7348.)

transport model to give ℓ. As a third example of the relationship between photon distributions and wave correlation, the angular distribution of the scattered intensity $I(\theta)$ and the spatial correlation function of the field are Fourier conjugates. The field within a random medium becomes uncorrelated because of the random phasing of waves arriving at a point from all directions. Within the bulk of the medium, the angular distribution is isotropic, with an uncertainty of $1/\ell$ in the absolute value of the wave number. The normalized field correlation function is

$$F_E(\Delta r) = (\sin(k\Delta r)/k\Delta r)\exp(-\Delta r/2\ell) \quad [10]$$

A measurement of $F_E(\Delta r)$ is shown in **Figure 5**. The right-hand side is just the diffraction-limited Green function for the complex field. The relation between spatial correlation and the Green function is reminiscent of the fluctuation–dissipation theorem. Equation [10] has found important applications in modern seismic imaging.

A final application of Gaussian statistics is quasi-elastic light scattering in multiply scattering media. Diffuse wave spectroscopy measures the time correlation function of the intensity, $C(\tau)$, to retrieve the mean Stokes–Einstein translational diffusion constant of dynamic colloidal suspensions.

Intensity Correlation

Intensity correlations can be subject to non-Gaussian contributions that do not obey the Seigert relation

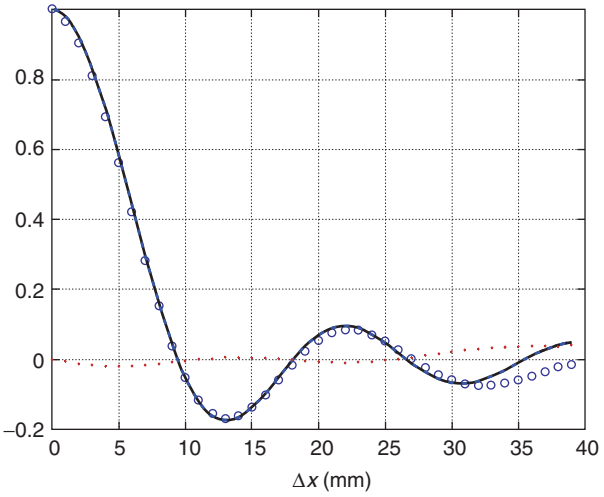


Figure 5 Real and imaginary parts of the spatial field correlation function of the microwave field transmitted in an ensemble of random polystyrene samples in the range 18.5–19.5 GHz. The solid line is the theoretical prediction. (Sebbah *et al.* (2000) *Physical Review E* 62: 7348.)

discussed earlier. These are seen to be long range. As a result, optical flux summed over different incoming or outgoing spots does not self-average. Similarly, the fluctuations of the electronic conductance are greatly enhanced over the value predicted without including correlation between values of intensity for different incident and outgoing modes. Its variance has a universal magnitude, equal to 2/15 in diffusive samples, independent of sample size and scattering strength. This is known as "universal conductance fluctuations" in electronic physics. Calculations and experiments have shown that the relative importance of non-Gaussian correlations is governed by the ensemble average of g.

The cumulant correlation function of transmitted intensity with displacement of either the source or the detector, ΔR and Δr, respectively, or rotation of the corresponding polarizations, $\Delta\Theta$ and $\Delta\theta$, can be expressed as the sum of three terms, with distinctive spatial and polarization dependencies. Let ΔQ and Δq represent the shifts of the source and the detector, respectively, with either position or polarization angle. Each term involves only the sum or the product of the square of the field correlation function, $F = |F_E|^2$ or a constant. The leading-order ("C_1") term is the product $F(\Delta Q)F(\Delta q)$ and decays rapidly with displacements of either the source or detector. The next ("C_2") term is proportional to the sum $F(\Delta Q) + F(\Delta q)$ and decays to half its value when only one of the sources and detectors is displaced. Finally, the third ("C_3") term is proportional to $F(\Delta Q)F(\Delta q) + [F(\Delta Q) + F(\Delta q)] + 1$ and correlates any state of the source to any state of the detector

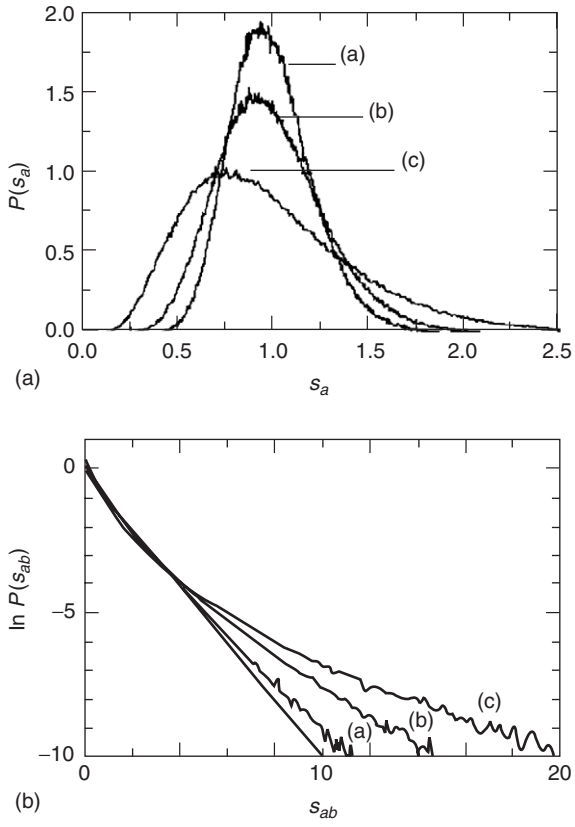


Figure 6 Distributions of total transmission (a) and intensity (b) of microwave radiation transmitted through quasi-1D samples of polystyrene spheres with different values of the variance of normalized transmission equal to the degree of nonlocal correlation: var(s_a) = 0.045 (a); 0.075 (b), and 0.30 (c). (Stoytchev and Genack (1997) *Physical Review Letters* 79: 309.)

as a result of the constant contribution. The C_2 and C_3 terms represent long-range non-Gaussian correlations. If the source is unaffected, $\Delta Q = 0$, so that $F(\Delta Q) = 1$, the cumulant intensity correlation for displacement or polarization rotation may be written as

$$C(\Delta q) = F(\Delta q) + \kappa[F(\Delta q) + 1] \qquad [11]$$

Equation [11] holds both for steady state and for pulsed excitation of a random sample. For pulsed transmission, the degree of correlation depends upon the time delay, t, and line width, σ, of the incident pulse and increases with t. For both steady-state and pulsed measurements, the probability distribution of intensity depends only upon κ. Thus, the degree of correlation determines the statistics of waves in random media (**Figure 6**).

Probability Distributions

The probability distribution of intensity or total transmission is determined exclusively by κ, since

fluctuations in average transmission occur due to extended correlation on the output surface of the sample. Since the intensity distribution is a negative exponential for a fixed value of s_a, the intensity distribution may be written as

$$P(s_{ab}) = \int_0^\infty \mathrm{d}s_a \, P(s_a) \frac{1}{\pi s_a} \exp\left(-\frac{s_{ab}}{s_a}\right)$$

This expression is equivalent to the relation between corresponding moments, $\langle s_{ab}^n \rangle = n! \langle s_a^n \rangle$.

Phase Statistics

At points in the speckle pattern at which the intensity is zero, singularities exist in the spatial distribution of the phase φ_{ab}. The phase, whose tangent equals the ratio of in- and out-of-phase components of the field for a single polarization, is undefined when the amplitude A_{ab} vanishes. It jumps by π-radians as a null in the intensity is traversed, and varies by 2π as such a phase singularity is circled (**Figure 1**). In the diffuse regime, the phase itself is uniformly distributed between $-\pi$ and π and does not provide much useful information. The time delay of a narrow-frequency-band incident pulse crossing the medium is the frequency derivative of the phase, $\tau = \varphi' \equiv \mathrm{d}\varphi/\mathrm{d}\omega$. In nonabsorbing diffusive media, the time delay fluctuates strongly with displacement and has an average value of $\langle \tau \rangle = L^2/6D$, which coincides with the time delay predicted by the diffusion theory of eqn [4]. The variance of τ diverges as a result of the large spectral derivative of the phase near phase singularities.

Both the phase delay time distribution $P(\mathrm{d}\varphi/\mathrm{d}\omega)$ and the joint distribution $P(I, \mathrm{d}\varphi/\mathrm{d}\omega)$ can be calculated for diffusive waves, assuming complex-valued Gaussian statistics. In **Figure 7**, the theoretical prediction assuming Gaussian statistics and the measurements of the joint distribution $P(I, \mathrm{d}\varphi/\mathrm{d}\omega)$ are shown. $P(I, \mathrm{d}\varphi/\mathrm{d}\omega)$ is a Gaussian function for a fixed value of I centered at $\langle \tau \rangle$, with a variance which is inversely proportional to I. For localized waves, the intensity is correlated with the delay time, since at a transmission resonance at which the intensity is high, the dwell time is long. As a result, the average delay time increases with intensity.

Coherent Backscattering

Coherent backscattering is a pertinent counter example of the statement that multiple scattering destroys all interference effects, and an illustration of how mesoscopic wave–particle duality can be

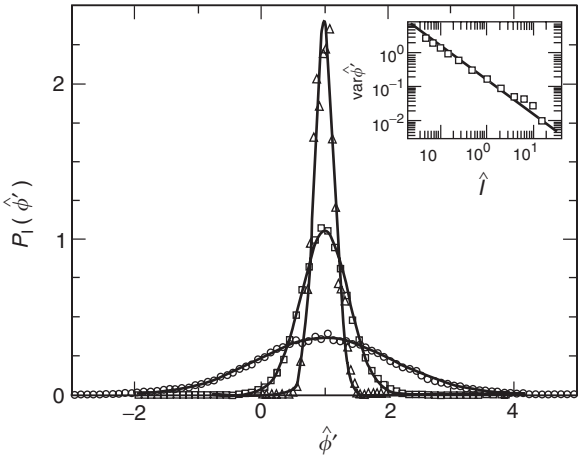


Figure 7 Probability distributions for the normalized delay time $\hat{\varphi}' \equiv \mathrm{d}\varphi/\mathrm{d}\omega / \langle \mathrm{d}\varphi/\mathrm{d}\omega \rangle$ in transmission for fixed values of the normalized intensity I. Circles: $I = 0.1$, squares: $I = 1.0$, and triangles: $I = 10$. The Gaussian curves denote the theoretical predictions. Inset: normalized variance $(\Delta\hat{\varphi}')^2$ of the delay time at fixed normalized intensity I, $(\Delta\hat{\varphi}')^2 = Q/2I$; the solid line gives the theoretical prediction with $Q = 0.31$. (Genack *et al.* (1999) *Physical Review Letters* 82: 715.)

employed. Consider a typical random walk of a wave along an arbitrary sequence of scatterers before being reflected back toward the light source. In the radiative transfer picture, all interference effects are assumed to be washed out upon averaging over all configurations of the particles, and the intensity due to waves following two paths would be given by $|\psi_1|^2 + |\psi_2|^2$. However, in the wave picture, two oppositely propagating waves will always interfere constructively at the source. The reflection of energy is then given by

$$\begin{aligned} R(\theta = \pi) &= |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2\mathrm{Re}(\psi_1 \psi_2^*) \\ &= 2(|\psi_1|^2 + |\psi_2|^2) \end{aligned} \quad [12]$$

which is twice the answer obtained from radiative transfer. Coherent backscattering is a direct consequence of the reciprocity principle. The angular line profile of coherent backscattering is the Fourier transform of the distribution of distances between the entry and exit of reflected photons, and as such a very sensitive way of probing the path length distribution of reflected photons. The algebraic long-range nature of this distribution generates a triangular cusp which is seen in **Figure 8**. It is noticed that the enhancement factor of 2 is lowered when the ℓ/λ ratio approaches unity. This has been attributed to loop-type photon paths in the random medium, which then become numerous and give a flat contribution to the background.

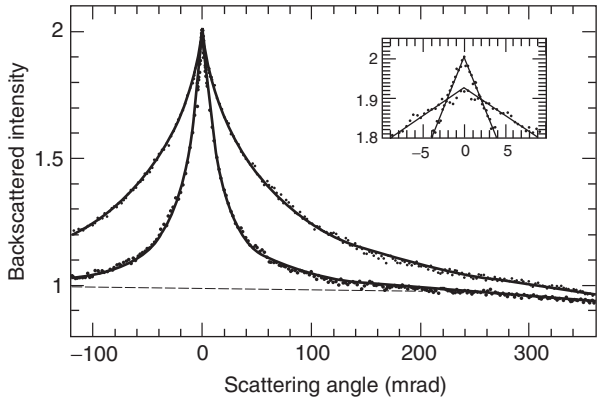


Figure 8 Coherent backscattering of light measured, with two different values for the transport mean free path ℓ. The typical angular width varies as λ/ℓ. Narrow cone: a sample of $BaSO_4$ powder with $\ell/\lambda = 4$, Broad cone: TiO_2 sample with $\ell/\lambda = 1$. The inset confirms the triangular cusp predicted by diffusion theory, and also shows that the maximum enhancement factor is lowered for the sample with small ℓ/λ. (Wiersma *et al.* (1995) *Physical Review Letters* 74: 4193.)

Photon Localization

A key indicator of the nature of propagation in a random medium is the Thouless number, δ, which is the ratio of the typical frequency width and frequency spacing between states of a random medium, $\delta = \delta\nu/\Delta\nu$. The inverse of the level spacing is the density of states of the sample. In the absence of absorption or gain, the Thouless number is equal to the dimensionless conductance, $\delta = g$, which is, in turn, essentially the inverse of κ, introduced earlier. The onset of localization occurs at the point at which $\delta\nu = \Delta\nu$, that is, $\delta = 1$. For $\delta < 1$, transport is inhibited since modes in different blocks, into which one may imagine the sample to be divided, do not overlap.

In order to localize photons, scattering can be maximized by utilizing radiation at a Mie resonance with high-index dielectric spheres. In order not to wash out the resonance, the volume fraction of scatterers should not be too high, while their density should not be too low. At the same time, localization can be facilitated by utilizing low-dimensional sample geometries. Results of microwave propagation in a waveguide containing random mixtures of alumina spheres with index of refraction $m = 3.14$ and volume fraction of 0.068 are shown in **Figure 9**. A narrow window of localization is observed just above the first Mie resonances in a range in which δ achieves its lowest value of just under unity.

The barrier to localizing radiation, even in random media in which the wave is resonant with scatterers, can be seen from a dynamical perspective. $\delta\nu$ drops sharply at resonance, as can be seen in the sharp peak in the average dwell time $\langle \tau \rangle$ (**Figure 9**). This does not result in a diminished value of the Thouless number δ, however, because it is countered by an increase in the density of states corresponding to a drop in $\Delta\nu$. In electronic systems, g can be obtained from the measurement of the conductance G, since $G = (e^2/h)g$, whereas for classical waves it can be found from the measurement of the total transmission T_a, since $g = N\langle T_a \rangle$, where N is the number of transverse modes and for a quasi-1D system given by

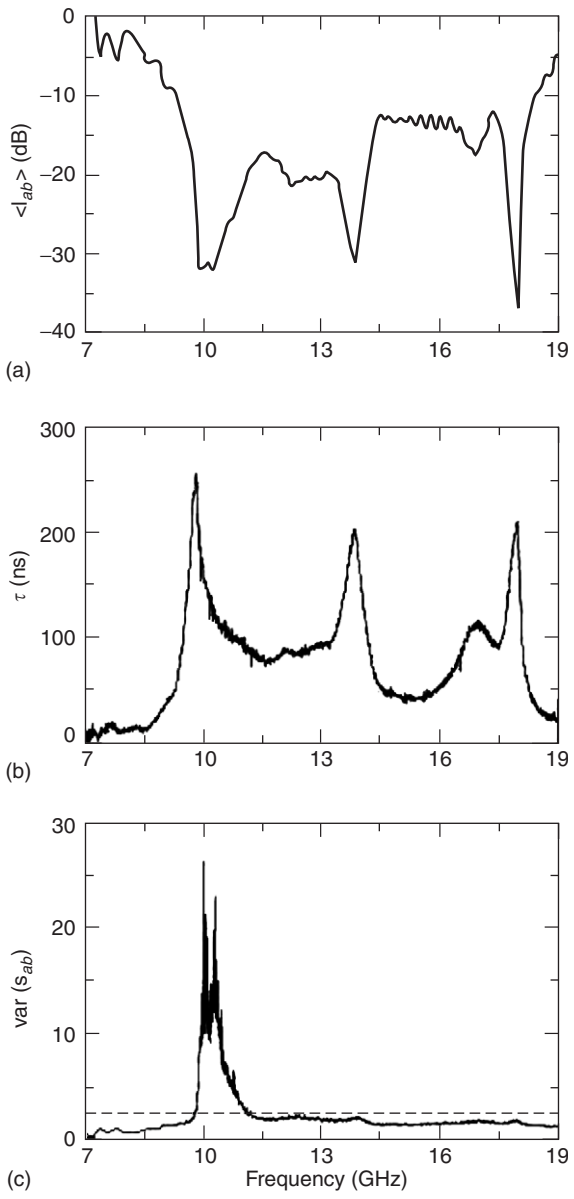


Figure 9 Frequency variation of the (a) average transmitted intensity $\langle I_{ab} \rangle$, (b) average photon transit time, $\tau = \langle s_{ab} \mathrm{d}\varphi_{ab}/\mathrm{d}\omega \rangle$, and (c) var($s_{ab}$) in an ensemble of quasi-1D samples of 0.9 cm diameter alumina spheres contained in a copper tube of diameter 7.3 cm and length 90 cm with reflecting walls and open ends at a spheres volume fraction of 0.067. The dashed line in (c) indicates the localization threshold. (Chabanov and Genack (2001) *Physical Review Letters* 87: 153901.)

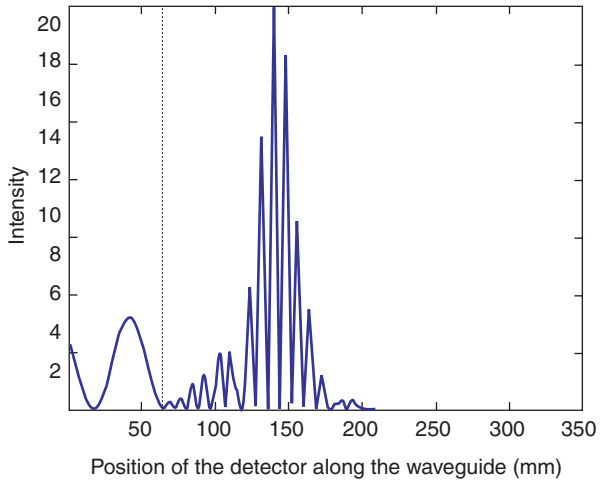


Figure 10 Intensity profile of a localized mode at 14.47 GHz inside a 1D waveguide for a particular random alternation of air and high-index dielectric slabs. The dotted line indicates the interface between the empty waveguide section (left) and the random waveguide section (right).

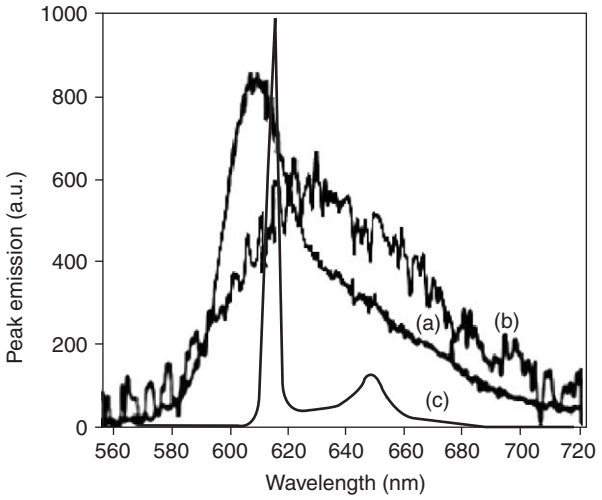


Figure 11 Emission spectra (b) and (c) of a colloidal solution of TiO_2 nanoparticles in rhodamine-640 for increasing pump power, as compared to emission spectrum *a* of pure solution. (Lawandy N M, *et al.* (1994) *Nature* 368: 436.)

$N = Ak^2/2\pi$, in terms of the cross section A of the sample and the wave number $\boldsymbol{k}$. For diffusive waves in nonabsorbing quasi-1D samples of length L, Ohm's law applies: $g = N\ell/L$.

When L exceeds the localization length, $\xi = N\ell$, g falls below the critical value of order unity. Beyond this point, the transmission falls exponentially $\langle T_a \rangle \sim \exp(-L/2\xi)$. In localized samples with $\delta < 1$, the intensity falls exponentially inside the sample between resonances with localized modes. On resonance with a localized mode, the intensity is exponentially localized around a point in the medium. An example of a single localized mode in a 1D random microwave waveguide is shown in **Figure 10**. For a wave resonant with a wave localized near the center of a sample, the wave may rise exponentially in the first half of the sample and then fall exponentially in the second half to produce a value of transmission near unity. States within a medium in which $\delta < 1$, may be coupled when two or more states overlap spatially and spectrally. In this case, the excitation in the coupled modes may extend throughout the medium and provide a path for strong transmission over a spectral range greatly exceeding the width of a single isolated localized state. Such "necklace" states may dominate transmission in many circumstances. Examples of a single localized mode and of two coupled modes in a 1D random microwave waveguide are shown in **Figure 10**.

Random Laser

Though scattering within a laser cavity generally degrades laser performance, stimulated emission can be facilitated in a multiply scattering medium with gain, since the optical dwell time within such media can be enhanced by scattering, thereby increasing the opportunity for stimulated emission. The presence of gain in random media, pumped by an external laser, favors long scattering paths within the gain medium and may facilitate laser action. In diffusive amplifying media, the line width of states of the medium greatly exceeds the typical level spacing $\Delta\nu$, so that many spectrally and spatially overlapping modes are excited simultaneously. Propagation in such systems with "nonresonant" feedback may be described in terms of photon diffusion and incoherent pumping of material energy levels. Emission spectra are narrowed and the temporal profile is greatly shortened above a threshold in pump power (**Figure 11**).

Laser emission in a series of distinct peaks has been observed in thin samples with strong scattering, in which $k\ell \sim 4$. Narrow peaks, observed in the emission spectrum of a semiconductor powder shown in **Figure 12**, are a direct manifestation of laser oscillation with resonant feedback in specific modes. In these samples, the wave is still diffusive, since $\delta > 1$, but the presence of gain allows a few longer-lived optical modes of the excited region to be selected for laser action so that distinct peaks are observed in the lasing spectrum.

The threshold for random lasing is not appreciably lowered below the one for amplified spontaneous emission in homogeneous media, since even though multiple scattering hinders the escape of emission from the medium, the excitation region is shallow

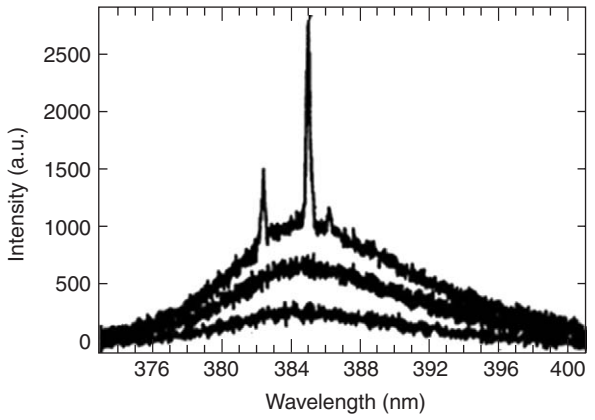


Figure 12 Spectra of emission from a 6 mm thick ZnO powder film for increasing excitation intensity (from bottom to top). (Cao H, *et al.* (1999) *Physical Review Letters* 82: 2278.)

because the flow of the incident pump light into the system is impeded. This limitation has been overcome in localized samples, such as 1D layered media, since modes localized deep within the sample can be resonantly excited by the pump laser. Emission from the excited medium will then occur in long-lived modes which overlap the excitation region. This leads to lasing at low threshold in pump power. In such systems, $\delta\nu \ll \Delta\nu$, and lasing proceeds by resonant feedback. Simulations in one and two dimensions have shown that laser modes are spatially localized and identical to the localized modes of the system without gain.

See also: Disorder and Localization Theory.

PACS: 42.25.Dd

Further Reading

Akkermans E and Montambaux G (2005) *Mesoscopic Physics of Electrons and Photons*. Cambridge: Cambridge University Press.

Chandrasekhar S (1950) *Radiative Transfer*. Oxford: Clarendon (reprinted by Dover, 1960).

Goodman JW (1985) *Statistical Optics*. New York: Wiley.

http://www.physics.utoronto.ca

http://www.tn.utwente.nl

Ishimaru A (1978) *Wave Propagation and Scattering in Random Media*. New York: Academic Press.

Lagendijk A and van Tiggelen BA (1996) Resonant multiple scattering. *Physics Report* 270: 143.

Ping Sheng (ed.) (1990) *Scattering and Localization of Classical Waves in Random Media*. Singapore: World Scientific.

POAN Research Group (ed.) (1998) New aspects of electromagnetic and acoustic wave diffusion. Heidelberg: Springer.

Sebbah P (ed.) (2001) *Waves and Imaging through Complex Media*. Dordrecht: Kluwer Academic.

van Haeringen W and Lenstra D (eds.) (1991) *Analogies in Optics and Microelectronics*. Amsterdam: North-Holland.

van de Hulst HC (1982) *Light Scattering by Small Particles*. New York: Dover.

van Rossum MCW and Nieuwenhuizen ThM (1999) Multiple scattering of classical waves: microscopy, mesocopy, and diffusion. *Reviews of Modern Physics* 71: 313.

van Tiggelen B and Skipetrov S (eds.) (2003) *Wave Scattering in Complex Media, NATO Sciences Series*. Dordrecht: Kluwer.

X

X-Ray Absorption Spectroscopy

T Jo, Hiroshima University, Higashi-Hiroshima, Japan

Introduction

X-ray absorption spectroscopy is used to measure and study the absorption intensity of incident X-rays in matter such as atoms, molecules, and condensed systems as a function of its energy or wavelength. The recent development of synchrotron radiation as a brilliant source of continuous X-ray has stimulated X-ray absorption spectroscopy. In X-ray absorption spectra, the intensity discontinuously increases at a critical incident energy and the spectra extend to higher energies. The critical incident energy is called the absorption edge or simply the edge. An absorption edge corresponds to the minimum energy needed for ionization by exciting a core electron in atoms and molecules, or that needed to excite a core electron into unoccupied states above the Fermi level in solids. In the absorption spectra, there are structures. Structures from the edge to several tens of eV are called X-ray absorption near edge structure (XANES), and those from the edge to several hundreds of eV are called extended X-ray absorption fine structure (EXAFS). The term X-ray absorption fine structure (XAFS) is also used to represent both XANES and EXAFS. In addition to the energy, X-ray is specified by the polarization, the control of which has recently become feasible in synchrotron radiation. A lot of information about electronic and atomic structures in solids is obtained from X-ray absorption spectroscopy. Here, information about electronic structures obtained from XANES is described. For information about atomic structure obtained from EXAFS, the reader is referred to the "Further reading" section. Since X-ray absorption is a quantum mechanical phenomenon, the electron–photon interaction and the derivation of the expression of the absorption spectrum with the use of Fermi's golden rule are discussed first.

Electron–Photon Interaction

For a system with an electron, the unperturbed Hamiltonian is assumed without the radiation field given by

$$\mathscr{H}_0 = \frac{p^2}{2m} + v(\boldsymbol{r})$$

where m, $\boldsymbol{r}$, and $\boldsymbol{p}$ denote the mass, the position vector, and the momentum of the electron, respectively, and $v(\boldsymbol{r})$ the potential energy acting on the electron. In the presence of the radiation field, the Hamiltonian is, with the use of the vector potential $\boldsymbol{A}(\boldsymbol{r},t)$ and the scalar potential $\phi(\boldsymbol{r},t)$, both of which are $\boldsymbol{r}$ and time (t) dependent, written as

$$\mathscr{H} = \frac{1}{2m}(\boldsymbol{p} + \mathrm{e}\boldsymbol{A}(\boldsymbol{r},t))^2 + v(\boldsymbol{r}) - e\phi(\boldsymbol{r},t)$$

with the charge of electron $-e$. Then the time-dependent perturbation due to the radiation field, that is, the electron–photon interaction $W(t)$ is, neglecting a very small $\boldsymbol{A}^2$ term, written as

$$W(t) = \frac{e}{2m}(\boldsymbol{p}\cdot\boldsymbol{A}(\boldsymbol{r},t) + \boldsymbol{A}(\boldsymbol{r},t)\cdot\boldsymbol{p}) - e\phi(\boldsymbol{r},t)$$

One adopts the Coulomb gauge $\nabla\cdot\boldsymbol{A}(\boldsymbol{r},t) = 0$ and $\phi(\boldsymbol{r},t) = 0$. Then,

$$W(t) = \frac{e}{m}(\boldsymbol{A}(\boldsymbol{r},t)\cdot\boldsymbol{p})$$

For systems with many electrons, $\boldsymbol{p}$ is replaced by the summation $\sum_i \boldsymbol{p}_i$, where $\boldsymbol{p}_i$ denotes the momentum of the ith electron. $W(t)$ represents the annihilation of a photon upon exciting an electron from the state with a lower energy to that with a higher energy (absorption) or the creation of a photon accompanied by the transition of an electron from the state with a higher energy to that with a lower energy (emission). In X-ray absorption, the state with the lower energy is a core state of an atom and that with the higher energy is an unoccupied valence state.

X-Ray Absorption Spectra

Let $|i\rangle$ be a many-body initial state of X-ray absorption, where core states and valence states up to the Fermi level are occupied by electrons. Let $|f\rangle$ be the fth many-body final state of X-ray absorption, where a core electron of an atom is transferred to an

unoccupied valence state upon absorbing a photon with energy $\hbar\omega$. Then, according to Fermi's golden rule, the transition probability per unit time from the initial state to the final state by the perturbation $W(t) = We^{-iwt}$, that is, the X-ray absorption spectrum $F(\omega)$ as a function of ω is expressed as

$$F(\omega) = \frac{2\pi}{\hbar} \sum_f |\langle f|W|i\rangle|^2 \delta(E_i + \hbar\omega - E_f)$$

where E_i and E_f denote the energy of $|i\rangle$ and $|f\rangle$, respectively. $F(\omega)$ is proportional to the ω-dependent absorption coefficient. The core hole has a lifetime due to the Auger process and decays exponentially with time. Then the δ function must be replaced by the Lorenzian $(\Gamma/\pi)/[(E_i + \omega - E_f)^2 + \Gamma^2]$ with the lifetime broadening Γ.

Spectra in a One-Electron Picture

Let the following three conditions be satisfied: (1) The matrix element $\langle f|W|i\rangle$ is dependent on ω only weakly. (2) The energy of $|i\rangle$ is expressed as the summation of occupied one-electron energies

$$E_i = \varepsilon_c + \sum_k \varepsilon_k$$

where ε_k and ε_c denote the energies of occupied conduction-band states with the wave vector k and of the core state, respectively. (3) The energy of $|f\rangle$ is also expressed as the summation of one-electron energies as

$$E_f = \varepsilon_{k'} + \sum_k \varepsilon_k$$

where $\varepsilon_{k'}$ denotes the energy of the conduction-band state with the wave vector k' unoccupied in $|i\rangle$; the one-electron states are not affected by the core hole produced in the final state of X-ray absorption. Then the absorption spectrum is expressed as

$$F(\omega) \sim \sum_{k'} \delta(\varepsilon_c + \hbar\omega - \varepsilon_{k'})$$

This means that, aside from the constant factor, $F(\omega)$ is approximately proportional to the density-of-states (DOS) of unoccupied conduction-band states, and this expression is widely used to obtain information about unoccupied states in solids.

Final-State Interaction

There are many examples where $F(\omega)$ is different from the expected DOS above the Fermi level ε_F. The attractive interaction of a core hole acting on the valence electron, which is produced in the final state, causes a reconstruction of both occupied and unoccupied valence states. This is called the final-state interaction. In the following, some examples of the same are presented.

Fermi Edge Singularity

In some metals, the core absorption spectra are known to exhibit singular behavior as a function of the incident photon energy near the absorption threshold just above ε_F. In **Figure 1**, such examples observed in alkali metals for p core absorption are shown, where the spectra exhibit a spike just above ε_F.

If the condition (3) described in the last section is not satisfied, then the photoproduced core hole acts as an attractive potential on conduction electrons. Consider a constant negative matrix element V representing the scattering of electrons, which is independent of electron wave numbers, that is, the

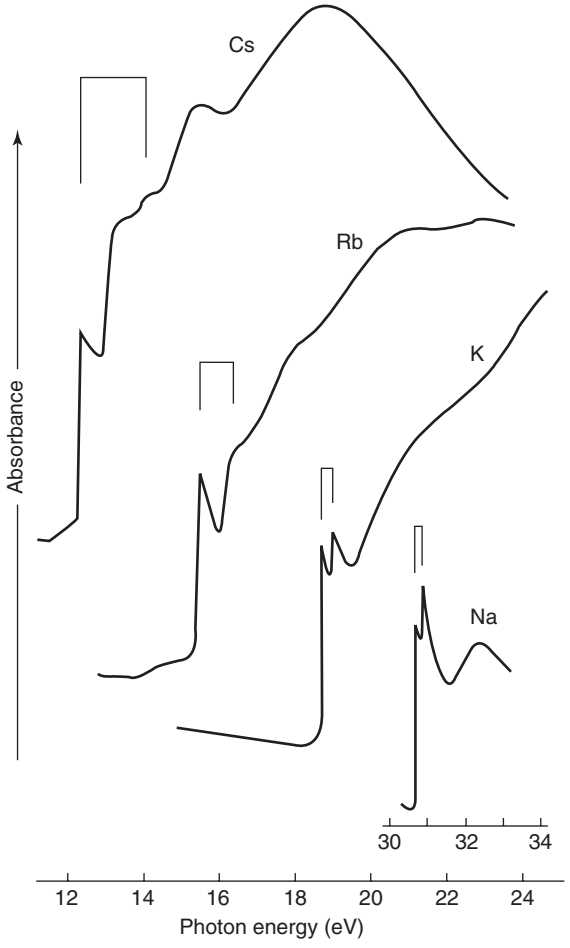


Figure 1 Optical absorption spectra of outermost *p*-shell electrons in Na, K, Rb, and Cs. Pairs of spikes linked by vertical bars correspond to the core spin–orbit splitting. (Reproduced with permission from Ishi T, *et al.* (1977) *Journal of Physical Society of Japan* 42: 876–881.)

δ function-type potential. Then $F(\omega)$ is, near ε_F, expressed as

$$F(\omega) \propto \left(\frac{1}{\hbar\omega + \varepsilon_c - \varepsilon_F}\right)^{\alpha}$$

where the index α is given by $\alpha = -2VN(\varepsilon_F)$ with the DOS at ε_F of the conduction electron $N(\varepsilon_F)$. α is a dimensionless parameter of $O(1)$. Since V is negative, α is positive and $F(\omega)$ is, therefore, enhanced for $\hbar\omega + \varepsilon_c$ near ε_F. This is called the Fermi edge singularity (or anomaly). In metals, the excitation energy starts continuously from zero due to the presence of ε_F. To the external perturbation V, the response of conduction electrons accompanied by low-energy excitations, which are peculiar to metals, causes the singularity.

Satellite Structures

The core hole produces another structure in the absorption spectrum. For example, in metallic cerium compounds, the localized $4f$-level is often just below ε_F and $|i\rangle$ is expressed as a superposition of the $4f^0$ configuration with a smaller weight and the $4f^1$ one with a larger weight due to the hybridization between the $4f$-state and the conduction band. By the $3d \rightarrow 4f$ X-ray absorption, the $4f^0$ configuration is transferred to the $4f^1$ one and the $4f^1$ one to the $4f^2$ one, both of which are accompanied by a $3d$ core hole. $V(<0)$ denotes the attractive potential of the core hole acting on the $4f$-electron and, $U(>0)$, the $4f$–$4f$ repulsion. Then the energy of $4f^1$ is $\sim V$ and that of $4f^2$ is $\sim 2V + U$ in the final state, if measured from a suitable origin of energy. The absorption spectrum is then characterized by the main peak ($4f^2$) and the satellite ($4f^1$) separated by $\sim|V + U|$. Combined with other X-ray spectroscopies, X-ray absorption gives information about U and the relative weight of $4f^0$ and $4f^1$ in $|i\rangle$. Similar satellite structures are observed in many systems, including transition-metal compounds, and they are used to extract information about electronic structures.

Multiplet Structures

In the $3p$ (or $2p$) $\rightarrow$ $3d$ X-ray absorption for insulating $3d$ transition-metal compounds, in addition to the $3d$–$3d$ multipole interaction, the strong $3d$–$3p$ (or $3d$–$2p$) multipole interactions, both of which are described by the Slater integrals, contributes to the energy splitting in the final state and the splitting is reflected in the spectrum. This is called the multiplet structure and is usually peculiar to the atomic configuration in the initial state. The multiplet structure, therefore, often plays the role of a fingerprint in identifying the configuration. Multiplet structures are also observed in the $4d$ (or $3d$) $\rightarrow$ $4f$ X-ray absorption for $4f$ rare-earth compounds.

Selection Rules for Electric Dipole Transition

Since the electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ is expressed as $\boldsymbol{E}(\boldsymbol{r}, t) = \partial\boldsymbol{A}(\boldsymbol{r}, t)/\partial t$ for $\phi = 0$, the relation $\boldsymbol{A}(\boldsymbol{r}, t) = (-\mathrm{i}/\omega)\boldsymbol{E}(\boldsymbol{r}, t)$ is obtained with noting the time-dependent factor $\mathrm{e}^{-\mathrm{i}\omega t}$. Then the time-independent part of the electron–photon interaction W is rewritten as $W = -\mathrm{i}(\mathrm{e}/m\omega)\boldsymbol{E}(\boldsymbol{r}) \cdot \boldsymbol{p}$, where $\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ represents the plane wave with the amplitude of electric field $\boldsymbol{E}$ and the wave vector $\boldsymbol{k}$. Furthermore, $\boldsymbol{p}$ in W can be replaced by $\boldsymbol{r}$ with the use of the commutation relation

$$\frac{\boldsymbol{p}}{m} = \dot{\boldsymbol{r}} = \frac{\mathrm{i}}{\hbar}[H_0, \boldsymbol{r}],$$

$$H_0|i\rangle = E_i|i\rangle, \qquad H_0|f\rangle = E_f|f\rangle$$

and the relation $\hbar\omega = E_f - E_i$ in the absorption spectrum $F(\omega)$ in terms of Fermi's golden rule. Then

$$F(\omega) = \frac{2\pi}{\hbar}\sum_f |\langle f|\boldsymbol{E}(\boldsymbol{r}) \cdot e\boldsymbol{r}|i\rangle|^2 \delta(E_i + \hbar\omega - E_f)$$

is obtained. For systems with many electrons, $e\boldsymbol{r}$ is replaced by $e\sum_i \boldsymbol{r}_i$, the electric dipole of a system. $\boldsymbol{E}(\boldsymbol{r})$ is expanded as

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}(1 + \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r} + \cdots)$$

The contributions to the spectrum of the first and second terms are called the electric dipole (E1) and the quadrupole (E2) transitions, respectively, and so on. The wave length of photons with energy from several eV to several keV is in the range of several thousand Å to several Å. It is much larger compared to the spatial extension of the atomic wave function of a core electron of a few or several tenths of Å, that is, $\boldsymbol{k} \cdot \boldsymbol{r} \ll 1$. This means that the dominant contribution to the spectrum is the E1 transition.

Let a system be in spherical symmetry, where $|i\rangle$ and $|f\rangle$ are specified by the values of total angular momenta J and J', respectively, and the values of their z components J_z and J_z', respectively. In the E1 transition, there exists an allowed relation between J and J', that is, the condition of the nonzero matrix element of the electric dipole operator. This is called the selection rule of the total angular momentum.

The inversion operator I transforms $\boldsymbol{r}$ into $-\boldsymbol{r}$. Since $I^2 = 1$, the eigenvalue of I is either $+1$ (even parity) or -1 (odd parity). In order to have a nonzero matrix element for $\langle i|\boldsymbol{r}|f\rangle$, the initial and final states must have parities different from one another. For a transition of a single electron, let l and l' be the orbital angular momentum of the initial (core) and final (valence) states, respectively. Then $l - l' = \pm 1$ and the spin states must be the same. For a system with many electrons and a spin–orbit coupling, $J - J' = 0, \pm 1$ (but $J = J' = 0$ is forbidden).

By choosing the electric polarization of an incident photon, the difference in the z component of angular momenta between the initial and final states is controlled. $\boldsymbol{E}\cdot\boldsymbol{r}$ in the above expression of $F(\omega)$ is, with their x, y, and z components, expanded as $E_x x + E_y y + E_z z$ or, with use of the spherical harmonics $Y_{lm}(\theta,\phi)$, expanded as

$$\boldsymbol{E}\cdot\boldsymbol{r} = \sum_{\mu=-1,0,1} E_\mu \sqrt{\frac{4\pi}{3}} Y_{1\mu}(\theta,\phi) r$$

where $E_\pm = (\pm E_x - \mathrm{i}E_y)/\sqrt{2}$ and $E_0 = E_z$. For the photon polarized along the z axis ($\mu = 0$), that is, the quantization axis, $J'_z - J_z = 0$. For the left (right) circularly polarized photon ($\mu = +1(-1)$), $J'_z - J_z = +1(-1)$.

Dichroism and Sum Rules

The difference in X-ray absorption between two kinds of incident polarizations is called dichroism, since the difference in absorption means the difference in the color of substances in the visible light region. Recently, due to the development of synchrotron radiation, both linearly and circularly polarized X-rays with high polarization became available and, their extensive use has led to various kinds of dichroism studies being carried out.

In the above expression of $F(\omega)$ in terms of Fermi's golden rule, $|\langle i|W|f\rangle|^2$ is rewritten as $\langle i|W^+|f\rangle\langle f|W|i\rangle$ with W^+ being the Hermite conjugate of W. If $F(\omega)$ is integrated with respect to ω in the region to cover E_f's for a given W, and $|f\rangle$'s can be regarded as a complete set, one obtains, with the closure relation $\sum_f |f\rangle\langle f| = 1$,

$$\int F(\omega)\,\mathrm{d}\omega = \frac{2\pi}{\hbar}\langle i|W^+W|i\rangle$$

that is, the integrated intensity being the expectation value of W^+W in the initial state. By employing this, some sum rules that relate the integrated intensities to physical quantities of the valence state are derived.

For ferromagnets, the difference between the right- and left-circular-polarizations with respect to the direction of magnetization is called magnetic circular dichroism. Let $I_+(I_-)$ be the integrated absorption intensity for the left (right)-circular-polarization and let I_0 be the integrated absorption intensity for $\mu = 0$ (linear polarization) from the core state with one-electron orbital angular momentum l_1 to the valence state with l_2 by the E1 transition. Then the relation

$$\frac{I_+ - I_-}{I_+ + I_0 + I_-} = \mathrm{sign}(l_1 - l_2)\frac{\langle L_z\rangle}{l_2 n_h}$$

is known as the L_z sum rule, where $\langle L_z\rangle$ and n_h denote the expectation value of the z component of the orbital angular momentum and the number of holes in the valence state, respectively. By the core spin–orbit interaction, the spectrum often splits into two regions. Let $(I_+ - I_-)_+((I_+ - I_-)_-)$ be the integrated dichroic intensity for the $j_+ = l_1 + \frac{1}{2}(j_- = l_1 - \frac{1}{2})$ region. If the separation between the two regions is sufficiently large, the relation

$$\frac{(I_+ - I_-)_+ - (l_1+1)/l_1(I_+ - I_-)_-}{I_+ + I_0 + I_-} = A(l_1,l_2)\frac{\langle S_z\rangle}{n_h} + B(l_1,l_2)\frac{\langle T_z\rangle}{n_h}$$

is satisfied and it is known as the S_z sum rule, where $\langle S_z\rangle$ and $\langle T_z\rangle$ denote the expectation values of the z components of the spin and the spin dipole defined by

$$\boldsymbol{T} = \sum_i [\boldsymbol{s}_i - 3\boldsymbol{r}_i(\boldsymbol{r}_i\cdot\boldsymbol{s}_i)/r_i^2]$$

respectively, in the valence state; $A(l_1,l_2)$ and $B(l_1,l_2)$ are the known rational functions of l_1 and l_2. A combined use of both the L_z and S_z sum rules has recently been a standard method of site-selective determination of both $\langle L_z\rangle$ and $\langle S_z\rangle$ from dichroic data. In some cases, the mixing between j_+ and j_- regions can be appreciable due to the core–valence multipole interaction, although they are apparently separated, that is, $|f\rangle$'s in each region cannot be regarded as a complete set. Then the deviation from the S_z sum rule can be appreciable. It may also be noted that there is a sum rule to relate the integrated dichroic intensity for two kinds of linear polarizations to a component of the quadrupole moment.

See also: Scattering, Inelastic: X-Ray (Methods and Applications); X-Ray Sources; X-Ray Standing Wave Techniques; X-Ray Topography.

PACS: 32.30.Rj; 61.10.Ht; 78.70.Dm

Further Reading

Carra P, Thole BT, Altarelli M, and Wang X (1993) X-ray circular dichroism and local magnetic fields. *Physical Review Letters* 70: 694–697.

Cowan RD (1981) *The Theory of Atomic Structure and Spectra*. London: University of California Press.

Gschneidner KA, Eyring L, and Hüfner S (eds.) (1987) *Handbook on Physics and Chemistry of Rare Earths*, vol. 10. Amsterdam: North-Holland.

Ishii T, Sakisaka Y, Yamaguchi S, Hanyu T, and Ishii H (1977) Threshold singularities of *p*-shell absorption in potassium, rubidium and cesium metals. *Journal of Physical Society of Japan* 42: 876–881.

Kanamori J and Kotani A (eds.) (1998) *Core Level Spectroscopy in Condensed Systems*. Berlin: Springer.

Koningsberger DC and Prins R (eds.) (1988) *X-Ray Absorption, Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES*. New York: Wiley.

Landau LD and Lifshitz EM (1977) *Quantum Mechanics*. Oxford: Pergamon.

Lovesey SW and Collins SP (1996) *X-ray Scattering and Absorption by Magnetic Materials*. Oxford: Clarendon.

Mahan GD (1990) *Many-Particle Physics*, 2nd edn. Oxford: Plenum.

Mott NF and Jones H (1958) *The Theory of Properties of Metals and Alloys*. New York: Dover.

Ohtaka K and Tanabe Y (1990) Theory of the soft X-ray edge problem in simple metals: historical survey and recent development. *Review of Modern Physics* 62: 929.

Nomenclature

$\boldsymbol{A}$	photon vector potential
e	unit of charge of an electron
$\boldsymbol{E}$	electric field
E_i, E_f	energies of the initial and final states of X-ray absorption
$F(\omega)$	X-ray absorption spectrum
$\hbar$	Planck's constant/2π
$\mathscr{H}, \mathscr{H}_0$	Hamiltonian operators with and without the radiation field
$\boldsymbol{k}$	wave vector of electric field
L_z	z component of total orbital angular momentum in units of $\hbar$
m	mass of an electron
$\boldsymbol{p}$	linear momentum of an electron
$\boldsymbol{r}$	position vector
$\boldsymbol{s}_i$	spin vector of the ith electron
S_z	z component of total spin in units of $\hbar$
$\boldsymbol{T}, T_z$	total spin dipole vector and its z component
Γ	lifetime broadening
ε_k	energy of an electron with the wave vector k
ε_F	Fermi level
ϕ	photon scalar potential
ω	photon frequency

X-Ray Sources

A Pifferi, CNR – Institute of Crystallography, Rome, Italy

X-ray is nowadays one of the most powerful tools to investigate hard and soft matter, hence the necessity to develop new sources to achieve more high-intensity beams.

Introduction

X-ray was the name given by W C Roentgen to the highly penetrating rays, which emanated when high-energy electrons struck a metal target. On 8 November 1895, he noted for the first time this phenomenon and on 28 December 1895, he astonished the scientific world with his preliminary report *Ueber eine neue Art von Strahlen* (along with experimental radiographs and by the X-ray image of his wife's hand) given to the president of the Wurzburg Physical-Medical Society. A few weeks later, C H F Muller was able to construct in his Hamburg factory the first commercial X-ray tube for one of the local hospitals. This was a milestone of what was to become a major technical industry. X-ray technology is one of the most important inventions of the nineteenth century. It plays a major role in our daily life.

Conventional Sources

X-Ray Tube

Whenever a charged particle is accelerated or decelerated, it emits energy in the form of photons. This process takes place within an X-ray tube to create X-rays. At one end of the tube is the source of electrons or the cathode, and at the other end of the tube is the metal target, or anode. To produce X-rays, electrons are accelerated from the cathode to the anode with a high voltage. When the electrons hit the anode, they are quickly stopped. This rapid change in speed causes the emission of X-ray photons. These X-rays are commonly called brehmsstrahlung or "braking radiation." Bremsstrahlung is the German word for slowing down or braking; this radiation can be considered radiation, resulting from the braking or projectile electrons by the nucleus.

An electron that completely avoids the orbital electrons on passing through an atom of the target may travel sufficiently close to the nucleus of the atom to come under its influence. Since the electron is negatively charged and the nucleus is positively charged, there is an electrostatic force of attraction between them. As the electron approaches the nucleus, it is influenced by a nuclear force much stronger than the electrostatic attraction. As it passes by the nucleus, it is slowed down and deviated in its course, leaving with reduced kinetic energy in a different direction. This loss in kinetic energy reappears as an X-ray photon.

If all the energy carried by an electron is transformed into radiation, the energy of an X-ray photon is

$$E_{\max} = h\nu_{\max} = \text{eV}$$

where $E_{\max}$ is the maximum possible energy, h the Planck's constant, $\nu_{\max}$ the photon frequency, e the charge of the electron, and V the acceleration potential. If the frequency is substituted with the wavelength, the above expression becomes

$$h\nu_{\max} = hc/\lambda_{\min} = \text{eV} \quad \text{and}$$
$$\lambda_{\min} = hc/\text{eV} = 12\,398/V$$

The energy contained in each bremsstrahlung photon emitted from an X-ray tube extends from that associated with the peak electron energy all the way down to zero. In other words, when an X-ray tube is operated at 70 kVp, bremsstrahlung photons with energies ranging from 0 to 70 keV are emitted, creating a typical continuous X-ray emission spectrum. Since total conversion of the electron energy into radiation is not a highly probable event, the radiation of the highest intensity is obtained at a longer wavelength compared to $\lambda_{\min}$ (1.5 times). **Figure 1** shows some X-ray continuous spectra as a function of the accelerating voltage.

If the bombarding electrons have sufficient energy, they can knock an electron out of an inner shell of the target metal atoms. Then electrons from higher states drop down to fill the vacancy, emitting X-ray photons with precise energies determined by the electron energy levels. These X-rays are called characteristic X-rays. In summary, characteristic X-rays are produced by transitions of orbital electrons from outer to inner shells. Since the electron binding energy for every element is different, the characteristic X-rays produced by the various elements are also different. This type of X-radiation is called characteristic radiation because it is characteristic of the target element. The effective energy characteristic X-rays increase with increasing atomic number of the target element. The characteristic lines of this type of spectrum are called K, L, and M and correspond to the transition from higher energy orbital to the K, L, and M orbital. When the two orbitals involved in the transition are adjacent then the line is called α, if they are separated by another shell the line is called β (**Figure 2**).

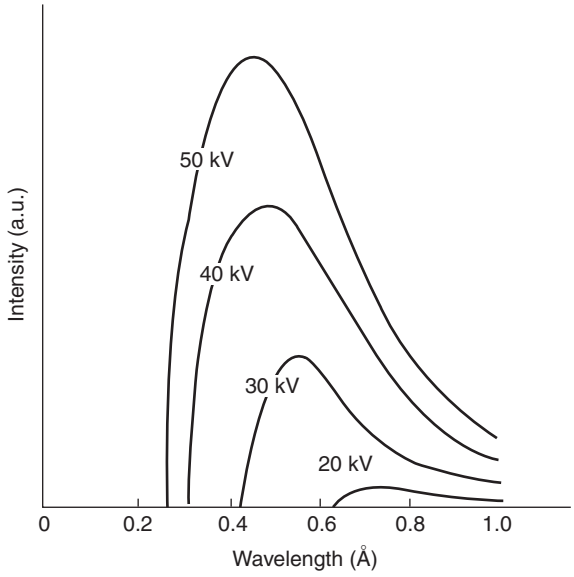


Figure 1 X-ray radiation spectra as a function of accelerating voltage.

For example, K electrons for tungsten anode have binding energies of 69.5 keV, and L electrons are bound by 12.1 keV. Therefore, the characteristic X-ray emitted has energy of

$$69.5 - 12.1 = 57.4 \text{ keV}$$

Sealed Tube and Rotating Anode Sources

The X-ray tube is a large diode valve, which is designed to produce fast moving electrons and then cause the electrons to decelerate rapidly in a vacuum environment. A conventional generator consists of a high-voltage power supply electronically controlled, applied between the anode and the cathode. In **Figure 3** the sketch of a sealed tube is showed.

It consists of a cathode with a filament that emits the electrons that are accelerated under vacuum by the high voltage. The electrons hit the anode and produce the characteristic X-ray spectrum relative to the metal of the anode. The high vacuum is necessary because the presence of gas molecules would decrease the efficiency of the tube. Normally this type of source is highly inefficient because only 0.1% of power used is transformed into X-ray, the rest is dissipated as heat. In the rotating anode the anode

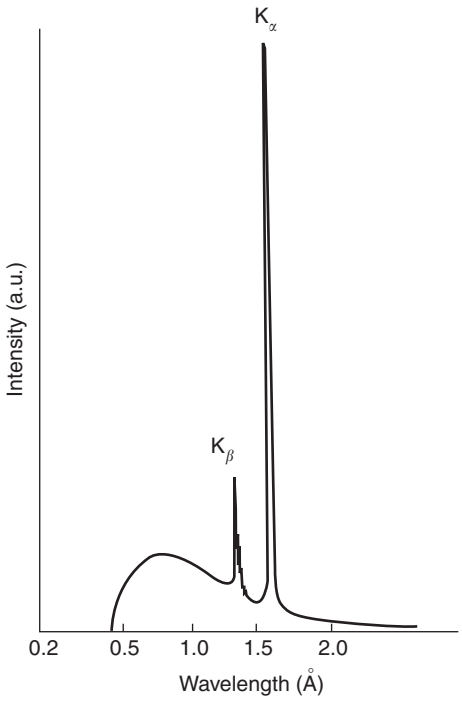


Figure 2 Characteristic lines K_α and K_β of Cu spectra superimposed on the brehmsstrahlung.

has a cylindrical shape, and to increase the efficiency of the X-ray beam the anode is continuously rotated. Thus the electron beam hits the metal of the anode in different positions permitting a better dissipation of heat. The increasing of the intensity of the beam may be quite an order of magnitude in respect to a sealed tube.

The intensity of a K line can be calculated using the equation:

$$I_K = Bi(V - V_K)^{1.5}$$

where B is a constant, i the electrical current, and V_K the excitation potential of the K series. If the accelerating potential $V = 4V_K$ is applied, the Kα line is 90 times more intense than the white radiation of the equivalent wavelength (I_W). In this condition the ratio I_K/I_W is a maximum.

Synchrotron Sources

Synchrotron radiation (SR) is emitted when charged particles moving with relativistic speeds are forced to follow curved trajectories in magnetic fields. The first visual observation of SR was in 1948 from the General Electric synchrotron in the USA, during investigations into the design and construction of accelerators suitable for the production of very high

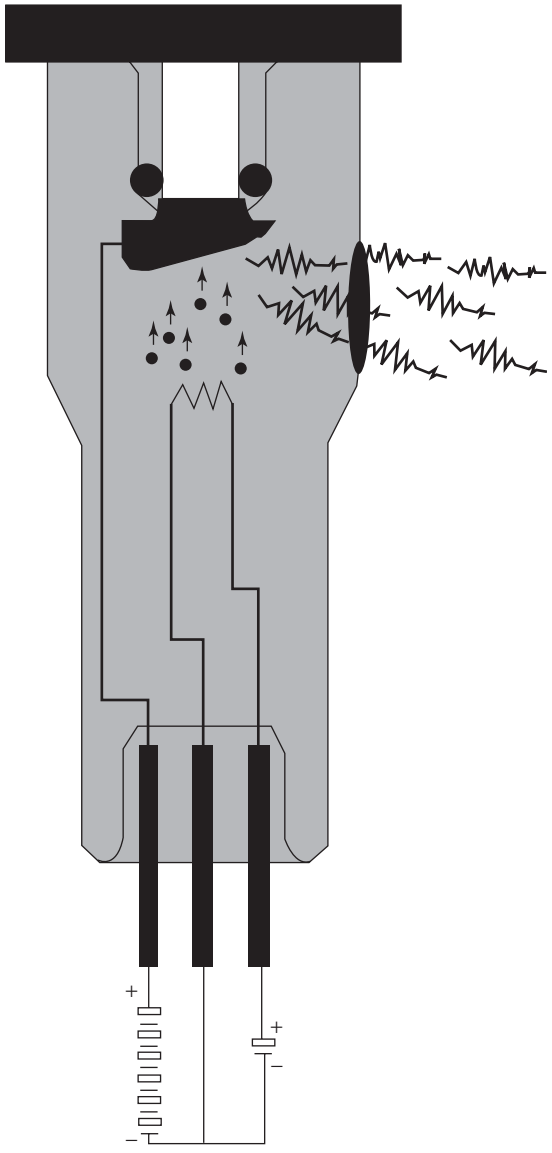


Figure 3 Sketch of a sealed X-ray tube.

energy electrons. Over the next 50 years, an explosive growth in the building of accelerators optimised for SR production has turned this interesting radiative energy into a valuable research tool.

Whenever charged particles undergo an acceleration, they emit electromagnetic radiation. An electron oscillating at radiofrequencies in an antenna emits radio waves. When electrons are subjected to an acceleration perpendicular to their velocity, for example, when they pass through a magnetic field, their direction is changed and they begin to travel in a circular path. Charged particles moving under the influence of an accelerating field emit electromagnetic radiation. The energy of this radiation is dependent on the velocity of the particle. The total instantaneous power P emitted by an electron

moving on a circular orbit is

$$P = \frac{2e^2cE^4}{3R^2(m_0c^2)^4} = \frac{2e^2c\gamma^4}{3R^2} \qquad [1]$$

where P is the energy emitted per unit time, e is the particle charge, c the speed of light, E the energy of the particle, m_0 its mass at rest, and R the bending radius of the orbit. The emitted radiation has very different characteristics in dependence of the velocity v of the particle. If $v \ll c$ the emitted pattern is identical to the well-known dipole radiation field (**Figure 4a**). When the speed increases to almost light velocity, the whole power is compressed into a narrow cone in forward direction tangentially to the orbit (**Figure 4b**). The radiation emitted is called synchrotron radiation. The quantity γ in eqn [1], which is the ratio of the total energy to the rest energy of the particle, is of considerable importance since it is related to the opening angle $\Delta\Psi$ of the cone of radiation in the perpendicular plane of the orbit by

$$\Delta\Psi \cong 1/\gamma$$

where $\Delta\Psi$ is expressed in radians.

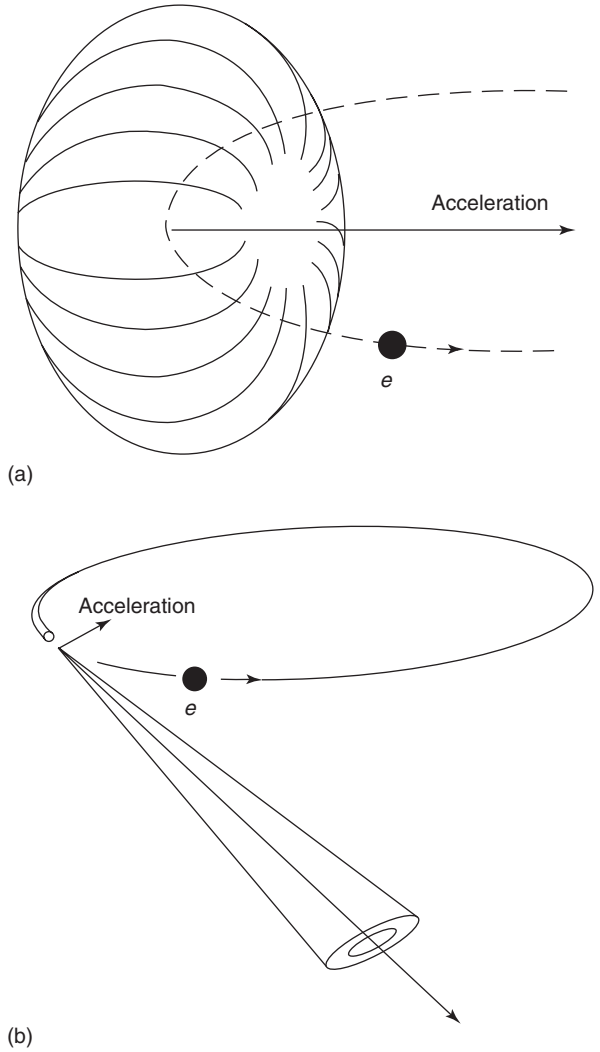


Figure 4 Emission pattern of radiation from centripetally accelerated electrons: (a) nonrelativistic case $v \ll c$, (b) relativistic case $v \cong c$.

A storage ring synchrotron consists of an arrangement of devices, where charged particles travel at essentially the speed of light in evacuated pipes under the influence of magnets, which are positioned around a closed orbit. Acceleration is achieved by the application of radio frequency electric fields at RF cavities along the circumference of the ring. The magnetic fields are increased synchronously with the acceleration in order to keep the particles on the constant radius path. Such accelerators can be used with protons or electrons, and even with heavier positive ions. The particles travel in the closed loop for periods of several hours while synchrotron radiation is emitted from all curved parts of the orbit.

In general, three kinds of magnets are used to make the necessary magnetic fields: bending magnets, wigglers, and undulators.

Bending Magnets

In bending magnets, a simple dipole structure is used to constrain the electrons in a curved path. The radiation emitted is extremely intense and extends over a broad wavelength range from the infrared through the visible and ultraviolet, and into the soft and hard X-ray regions of the electromagnetic spectrum. A typical output curve for a bending magnet source has a smooth spectral distribution with a broad maximum near the so-called critical wavelength. This critical wavelength depends on the square of the energy of the electrons and the bending radius in the dipole. In practical units of angstroms, it can be calculated as

$$\lambda c = 18.64/(B * E^2)$$

where B (the dipole field) is in tesla and E (the ring energy) is in GeV. The critical wavelength (or energy) has the property that one-half of the power is radiated above this wavelength and one-half below.

Wigglers

High-field wiggler magnets are often used as sources in order to increase the flux at shorter wavelengths. It is best to think of a wiggler as a sequence of bending magnets of alternating polarities, which gives a $2N$ enhancement in the flux, where N is the number

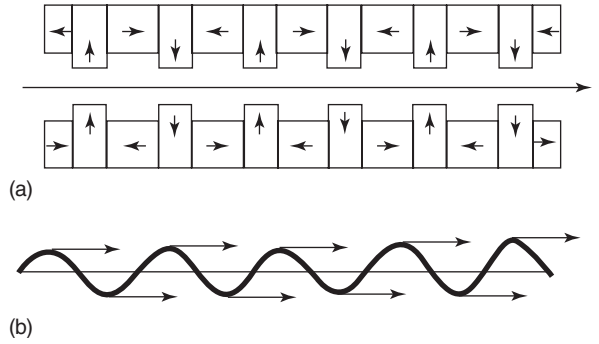


Figure 5 (a) The arrangement of magnets within a wiggler. Horizontal arrows indicate permanent magnet, vertical arrows indicate magnetic steel. (b) The trajectory of an electron beam within a wiggler. The arrows symbolize the emitted synchrotron radiation.

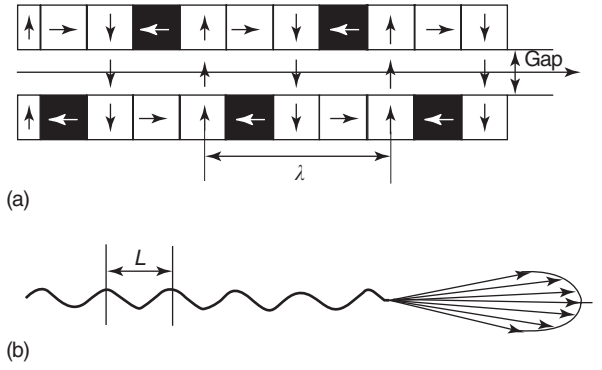


Figure 6 (a) The arrangement of magnets within an undulator. (b) The characterization of the undulator radiation. *L* is the period length.

of poles (**Figure 5**). Each magnet bends the electron beam through an angle which is large compared with mc^2/E. The properties of wiggler SR are thus very similar to that of dipole radiation with a reduction in the critical wavelength as a consequence of the higher field. For superconducting wiggler magnets a value of 6 T, as opposed to around 1.2 T for conventional dipoles, would be typical. The characteristics of the synchrotron radiation from the wiggler are the same as from the bending magnet, and so, the critical energy plays a determinant role. To reach the same spectral range as from the bending magnets, the magnetic flux density within the wigglers must be the same as within the bending magnets; if the magnetic field in the wiggler is higher than the ring bending magnet field, the wiggler spectrum extends to a higher photon energy. The photon flux as well as the central intensity of the radiation emitted by the wiggler is the same as from the bending magnet but more intensive by a factor Np, where Np is the number of poles within the wiggler.

Undulators

An undulator magnet is similar to a wiggler, in that it is also a succession of alternating magnetic poles (**Figure 6a**). However, in this case the angle of bend in each pole is of the order of mc^2/E, so that the small angular divergence due to the emission pattern of synchrotron radiation is not significantly increased. The radiation emitted at the various poles interferes coherently resulting in the emission of a pencil-shaped beam (**Figure 6b**) peaked in narrow energy bands at the harmonics of the fundamental energy. For N poles, the beam's opening angle is decreased by $N^{1/2}$ and thus the intensity per solid angle increases as N^2.

Figure 7 shows the performance in terms of brilliance of bending magnets, wigglers, and undulators for a 6 GeV energy synchrotron with a current in the ring of 100 mA.

Synchrotron radiation characteristically is highly polarized and continuous. Its intensity and frequency are directly related to the strength of the magnetic field and the energy of the charged particles affected by the field. Accordingly, the stronger the magnetic field and the higher the energy of the particles, the greater the intensity and frequency of the emitted radiation. The high intensity, broad spectral range, and other properties such as collimation, polarization, pulsed-time structure, partial coherence, make synchrotron radiation a powerful tool for basic and applied studies in biology, chemistry, medicine, and physics, as well as in applications to technology such as X-ray lithography, micromechanics, materials characterization, and analysis. Brilliance of synchrotron sources, up to $\sim 10^{10}$ larger than an X-ray tube, is mostly due to much smaller angular divergence. This is important because many experiments use small samples (e.g., protein crystals) that can use only a tiny fraction of radiation emitted by the tube. Much higher intensity makes it possible to observe very weak signals not observable with a tube source; much tighter angular collimation is especially important for small samples, for example, protein crystals, high pressure applications etc. By the use of special optical elements such as monochromators, single energies may be selected among the continuous spectrum of SR; this energy tunability is of extreme importance, especially for experiments driven near the absorption edge of specific elements.

Fourth Generation Light Sources

Future sources will be similar to existing storage ring sources. Some experiments require higher brightness,

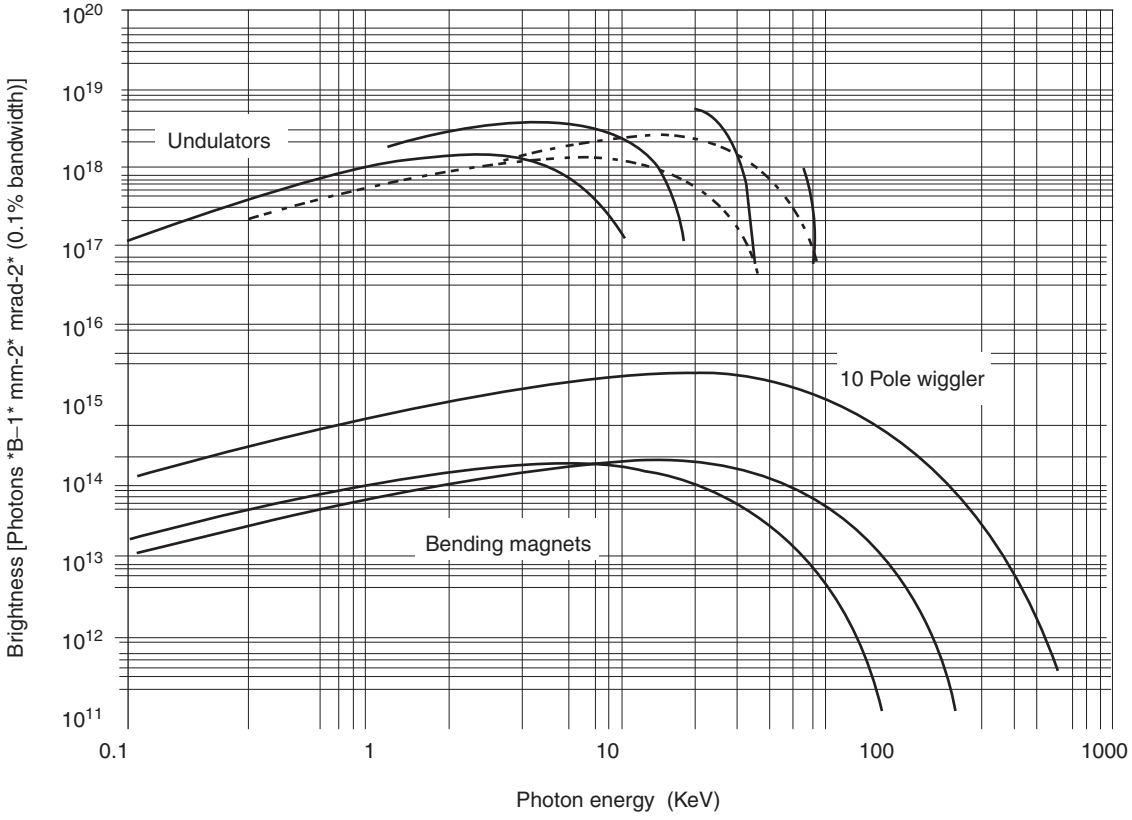


Figure 7 Plot of brilliance vs. photon energy for a storage ring of 6 GeV energy and 100 mA current.

coherence, peak power, or an extension of some other property of the radiation. Present sources are far from fundamental limits on the brightness and power of the radiation that can be produced. Concepts for future sources have been pursued at several laboratories, but the direction for the development of fourth generation light sources has been pointed out on free electron lasers (FELs). Such sources would produce radiation with brightness several orders of magnitude higher than undulators on third-generation rings. The graph in **Figure 8** shows the trend of the growth of the brilliance of X-ray sources starting from the beginning of the twentieth century.

Radiation from an FEL has much in common with radiation from a conventional optical laser, such as high power, narrow bandwidth, and diffraction limited beam propagation. One of the main differences between the two lasers is the gain medium: in a conventional laser the amplification comes from the stimulated emission of electrons bound to atoms, either in a crystal, liquid dye, or a gas, whereas the amplification medium of the FEL are "free" (unbound) electrons. The free electrons are stripped from atoms in an electron gun and are then accelerated to relativistic velocities. While the electrons are propagating in an undulator, the interaction with

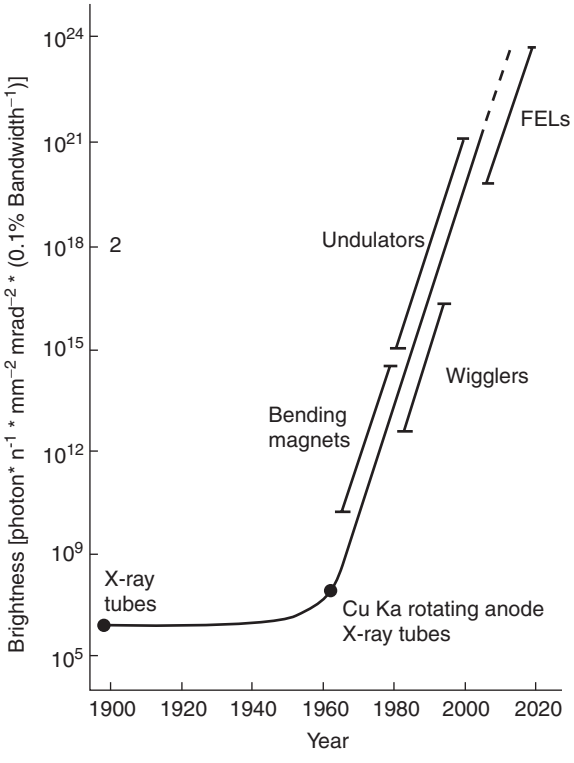


Figure 8 Trend of the brightness of X-ray sources plotted starting from the beginning of 1900.

an electromagnetic radiation field leads to an exponential growth of the radiation emitted by the electrons. This amplification of radiation is initiated by an increasingly pronounced longitudinal density modulation of the electron bunch. The initial radiation field can be an external one, for example, a seed laser, or an "internal" field, that is, the spontaneous emission of the undulator. In the latter case it is called a self amplified spontaneous emission (SASE) FEL. Oscillating through the undulator, the electron bunch then interacts with its own electromagnetic field created via spontaneous emission. Depending on the relative phase between radiation and electron oscillation, electrons experience either a deceleration or acceleration: electrons that are in phase with the electromagnetic wave are retarded, while the ones with opposite phase gain energy. Through this interaction a longitudinal fine structure, the so-called micro-bunching which amplifies the electromagnetic field is established. The longitudinal distribution of electrons in the bunch is "cut" into equidistant slices with a separation corresponding to the wavelength λ_{ph} of the emitted radiation, which causes the modulation. More and more electrons begin to radiate in phase, which results in an increasingly coherent superposition of the radiation emitted from the micro-bunched electrons. The more intense the electromagnetic field gets, the more pronounced the longitudinal density modulation of the electron bunch and vice versa. The characteristics of FEL radiation are high power, short pulse length (pico to femtoseconds), narrow bandwidth, spatial coherence, and wavelength tunability.

Laser Plasma X-Ray Sources

The concept for X-ray lasers goes back to the 1970s, when physicists realized that laser beams amplified with ions would have much higher energies than beams amplified using gases. Nuclear explosions were even envisioned as a power supply for these high-energy lasers. In X-ray lasers, a pulse of light strikes a target, stripping its atoms of electrons to form ions and pumping energy into the ions ("exciting" or "amplifying" them). As each excited ion decays from the higher energy state, it emits a photon. Many millions of these photons at the same wavelength, amplified in step, create the X-ray laser beam. Compact short-pulse high-power X-ray sources can be produced by focusing, on a solid target, ultrashort laser pulses of low or moderate energy (from several to ten or a hundred millijoules) but of high intensity ($\geqslant 10^{15}\,\mathrm{W\,cm^{-2}}$). For example, subpicosecond laser pulses can produce X-ray pulses of high power and short duration (in the picosecond range). With respect to hard and soft X-ray emission, such laboratory X-ray sources offer the following advantages over conventional X-ray sources such as synchrotrons: a greater compactness, an easier access, a lower cost, a shorter pulse length, and a smaller source size.

Conclusions

The future of the research in the field of X-ray sources will be more and more bonded to the possibility of optically manipulating the radiation, and obtaining a selective detection of the photons in order to perform even more accurate, spatial, and temporal analysis. The use of these new methodologies of detection is already paving the way to develop new X-ray sources and to study new physical phenomena which were until now, inaccessible.

See also: Scattering, Inelastic: X-Ray (Methods and Applications); X-Ray Absorption Spectroscopy; X-Ray Standing Wave Techniques; X-Ray Topography.

PACS: 29.20.Dh; 29.20.Lq; 29.27.Fh; 33.60.Fy; 41.50.+h; 41.60.Ap; 41.60.Cr; 42.55.Ve; 52.38.Ph

Further Reading

Clark GL (ed.) (1963) *The Encyclopedia of X-Rays and Gamma Rays*. New York: Reinhold Publishing Corporation.

Ewald PP (ed.) (1962) *Fifty Years of X-Ray Diffraction*. Utrecht: N.V.A. Oosthoek's Uitgeversmaatschappij.

Freund HP and Antonsen TM Jr. (1996) *Principles of Free-Electron Lasers*. London: Chapman and Hall.

Giacovazzo C (ed.) (2002) *Fundamentals of Crystallography*. New York: IUCr, Oxford University Press Inc.

Knoll GF (1989) *Radiation Detection and Measurement*. New York: Wiley.

Winick H (ed.) (1995) *Synchrotron Radiation Sources*. World Scientific.

X-Ray Standing Wave Techniques

M J Bedzyk, Northwestern University, Evanston, IL, USA

Introduction

The X-ray standing wave (XSW) technique provides an element-specific structural probe by using X-ray reflectivity to generate a "two-beam" interference field that in turn induces a spatial dependence to the X-ray spectroscopic yields from atoms within the field. The XSW technique is primarily used for resolving the atomic-scale structure at surfaces, interfaces, and thin films with applications spanning a wide scientific range including materials science, solid-state and soft-condensed-matter physics, environmental chemistry, geochemistry, biochemistry, and electrochemistry.

As a fundamental wave phenomenon, the superposition of two coherently coupled X-ray plane-waves localizes the X-ray intensity into interference fringes of an XSW field (**Figure 1**). This effect, which is produced by an X-ray reflection, makes it possible to attain a periodic structural probe with a length scale equivalent to the XSW period:

$$D = \frac{\lambda}{2 \sin \theta} = \frac{2\pi}{Q} \qquad [1]$$

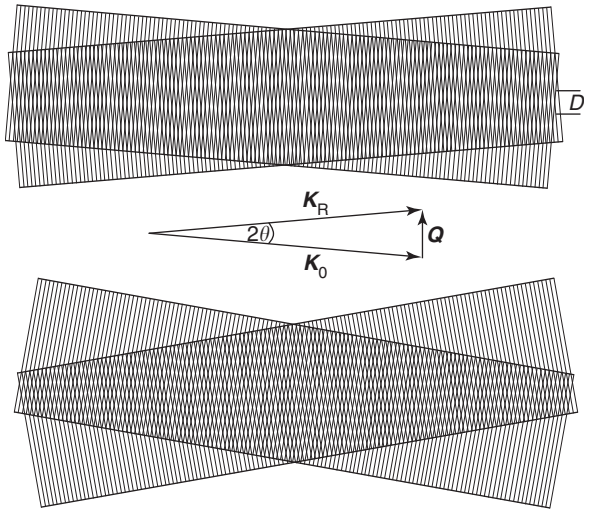


Figure 1 Top: a standing-wave field formed from the superposition of two traveling plane waves of wavelength λ and intersection angle (scattering angle) 2θ. The standing-wave period is D as defined in eqn [1]. Middle: the two traveling planes waves are represented in reciprocal space by wave vectors $\boldsymbol{K}_0$ and $\boldsymbol{K}_R$. $K_0 = K_R = 2\pi/\lambda$. The standing wave is defined by standing-wave vector $\boldsymbol{Q}$ defined in eqn [2]. Bottom: same as the top, but with a larger 2θ angle and, therefore, a smaller period D.

where λ is the X-ray wavelength, 2θ is the scattering angle between the two coherently coupled wave vectors $\boldsymbol{K}_R$ and $\boldsymbol{K}_0$, and $\boldsymbol{Q}$ is the scattering vector defined as

$$\boldsymbol{Q} = \boldsymbol{K}_R - \boldsymbol{K}_0 \qquad [2]$$

Q can also be referred to as the standing-wave vector, since it points perpendicular to the equal-intensity planes of the XSW and has a magnitude that is the reciprocal of D.

This two-beam reflection condition can be produced by: (1) strong Bragg diffraction from a single crystal, (2) strong Bragg diffraction from a periodically layered synthetic microstructure, (3) total external reflection (TER) from an X-ray mirror, or (4) weak kinematical Bragg diffraction from a single-crystal thin film. Single-crystal Bragg diffraction generates an XSW with a period equivalent to the diffraction plane spacing, typically between 0.5 and 10 Å. Cases (2) and (3) occur at small incident angles (θ) (or small Q) and therefore, according to eqn [1], will generate an XSW with a much longer period, typically 20–2000 Å. An XSW can be used as an element-specific spatial probe via the photoelectric effect, which can be observed by photoelectron, photo-ion, fluorescence, or Auger electron emission.

The XSW technique is used to investigate bulk-impurity structures in single crystals and a wide range of surface, interface, and thin film structures. These include semiconductor, metal and oxide surfaces, electrochemical interfaces, and organic membranes. In all cases, the XSW phase is directly linked to the substrate reflecting (or diffracting) lattice planes or interfaces, and can thereby be used to directly determine the absolute positions of the selected elements relative to these substrate planes. The accessibility of synchrotron radiation sources (with high brightness, energy tunability, and linear polarization) dramatically increases the sensitivity and applicability of the XSW technique.

XSWs Generated by Single-Crystal Dynamical Bragg Diffraction

An XSW generated by single-crystal Bragg diffraction can be used to determine the three-dimensional (3D) lattice location of bulk impurity atoms and surface adsorbates. Dynamical diffraction theory, which solves Maxwell's equations in a periodic dielectric with appropriate boundary conditions, is used to describe the fields inside and outside of the crystal.

Theory

Consider the two-beam Bragg diffraction condition, described in **Figure 2**, where the incident and the Bragg-diffracted X-ray plane waves are expressed as

$$\begin{aligned}\boldsymbol{E}_0(\boldsymbol{r},\ t) &= \boldsymbol{E}_0\ \exp[-\mathrm{i}(\boldsymbol{K}_0\cdot\boldsymbol{r}-\omega t)]\\ \boldsymbol{E}_H(\boldsymbol{r},\ t) &= \boldsymbol{E}_H\ \exp[-\mathrm{i}(\boldsymbol{K}_H\cdot\boldsymbol{r}-\omega t)]\end{aligned} \quad [3]$$

Here $\boldsymbol{E}_0$ and $\boldsymbol{E}_H$ are the complex E-field amplitudes associated with the incident and diffracted X-ray plane waves, $\boldsymbol{K}_0$ and $\boldsymbol{K}_H$ are the respective complex wave vectors inside the crystal, and ω is the X-ray frequency. The two wave vectors are coupled according to the Laue condition:

$$\boldsymbol{H} = \boldsymbol{K}_H - \boldsymbol{K}_0 \quad [4]$$

where $\boldsymbol{H} = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*$ is a reciprocal lattice vector. The scalar equivalent of the Laue condition reduces to Bragg's law, $\lambda = 2d_H \sin\theta_\mathrm{B}$, where $d_H = 2\pi/|\boldsymbol{H}|$ is the lattice spacing of the $H = hkl$ crystal diffraction planes, and θ_B is the geometrical Bragg angle. The interference between the incident and diffracted plane waves results in a standing-wave field. The normalized intensity of the total E-field that gives rise to the XSW field is

$$\begin{aligned}I(\theta,\ r) &= \frac{|\boldsymbol{E}_0+\boldsymbol{E}_H|^2}{|\boldsymbol{E}_0|^2}\\ &= [1+R(\theta)+2\sqrt{R(\theta)}\cos(v(\theta)-\boldsymbol{H}\cdot\boldsymbol{r})]\\ &\quad\times\begin{cases}1; & \text{above the surface}\\ \mathrm{e}^{-\mu_z(\theta)z}; & \text{at depth } z \text{ below the surface}\end{cases}\end{aligned} \quad [5]$$

where the reflectivity R is related to the E-field amplitude ratio as

$$R = \left|\frac{\boldsymbol{E}_H}{\boldsymbol{E}_0}\right|^2 \quad [6]$$

and the XSW phase, v, is identical to the relative phase between the two E-field amplitudes,

$$\frac{E_H}{E_0} = \left|\frac{\boldsymbol{E}_H}{\boldsymbol{E}_0}\right|\exp(\mathrm{i}v) \quad [7]$$

From eqns [1] and [5], one can conclude that, for Bragg diffraction, the XSW periodicity is equal to the lattice d-spacing of the hkl diffraction planes, that is, $D = d_H$.

In the following discussion, the most common case of σ-polarized symmetrical Bragg diffraction from a semi-infinite crystal with $1^\circ < \theta_\mathrm{B} < 89^\circ$ is assumed. **Figure 2** shows the case of σ-polarization with the vector directions of the two E-fields pointing perpendicular to the scattering plane defined by the two wave vectors. The incident and exit angles of the two wave vectors with respect to the surface are equivalent for a symmetric reflection.

From dynamical diffraction theory, the E-field amplitude ratio is defined as

$$\frac{E_H}{E_0} = -\sqrt{\frac{F_H}{F_{\bar{H}}}}(\eta \pm \sqrt{\eta^2-1}) \quad [8]$$

where F_H and $F_{\bar{H}}$ are the $\boldsymbol{H}$ and $-\boldsymbol{H}$ structure factors, which describe the superposition of the coherent X-ray scattering from the N atoms within the unit cell as

$$\begin{aligned}F_H &= |F_H|\exp(\mathrm{i}\phi_H)\\ &= \sum_{n=1}^{N}[f_n^0(H)+\Delta f_n'(\lambda)+\mathrm{i}\Delta f_n''(\lambda)]S_n(\boldsymbol{H})D_n(\boldsymbol{H})\end{aligned} \quad [9]$$

where $S_n(\boldsymbol{H}) = \exp(\mathrm{i}\boldsymbol{H}\cdot\boldsymbol{r}_n)$ is the geometrical phase factor for the nth atom located at $\boldsymbol{r}_n$ relative to the unit cell origin. $D_n(\boldsymbol{H}) = \exp(-M_n)$ is the Debye–Waller temperature factor for the nth atom. $\Delta f_n'$ and $\Delta f_n''$ are the real and imaginary wavelength-dependent

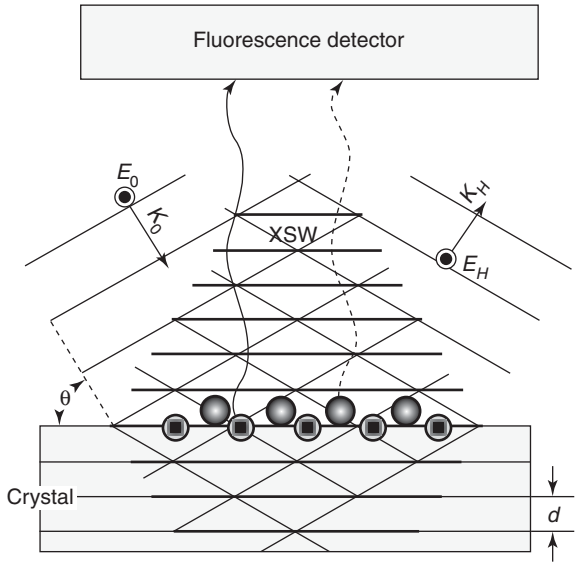


Figure 2 XSW field formed in a crystal and above its surface by the interference of incident and Bragg-diffracted X-ray plane waves. The XSW period is equal to the d-spacing d. Aligning an XSW nodal (or antinodal) plane with an atomic plane will minimize (or maximize) the characteristic fluorescence yield from that atomic plane.

anomalous dispersion corrections to the atomic form factor $f_n^0(H)$. η is the normalized angle parameter defined as

$$\eta = \frac{-\Delta\theta \sin(2\theta_B) + \Gamma F_0}{\Gamma\sqrt{F_H F_{\bar{H}}}} \qquad [10]$$

In this equation, $\Delta\theta = \theta - \theta_B$ is the relative incident angle. $\Gamma = (r_e\lambda^2)/(\pi V_c)$ is a scaling factor, where $r_e = 2.818 \times 10^{-5}$ Å is the classical electron radius and V_c is the volume of the unit cell. (To separate the real and the imaginary parts of a complex quantity A, the notation $A = A' + iA''$ is used, where A' and A'' are real quantities.) From eqns [6] to [10], it can be shown that the reflectivity approaches unity over a very small arc-second angular width w, defined as

$$w = \Delta\theta_{\eta'=-1} - \Delta\theta_{\eta'=1} = \frac{2\Gamma\sqrt{F'_H F'_{\bar{H}} + F''^2_0 - F''_H F''_{\bar{H}}}}{\sin 2\theta_B} \qquad [11]$$

This is the "Darwin width" of the reflectivity curve or "rocking curve."

Using the above dynamical diffraction theory equations (eqns [7]–[10]), one can show that the relative phase, v, of the standing-wave field decreases by π radians as the incident angle is scanned from the low-angle side to the high-angle side of the rocking curve. According to eqn [5], this causes the standing-wave antinodal planes to move by a distance of $(1/2)d_H$ in the $-\boldsymbol{H}$ direction. Also from eqn [5], if $\Delta f''_n = 0$, then $R = 1$, and the intensity at the antinode is four times the incident intensity $|\boldsymbol{E}_0|^2$, and there is zero intensity at the node. The case of $I = 4$ at the antinode assumes that the field is being examined above the surface or at a shallow depth where $\exp(-\mu_z z) \approx 1$.

The Darwin width w is dependent on both the structure factors and the wavelength of the incident X-ray beam. For a typical low-index, strong Bragg reflection from an inorganic single crystal, the angular width is within the range of $w = 5$–$100\,\mu$rad for X-rays with $\lambda = 0.5$–2 Å. **Figure 3a** shows a calculated rocking curve $R(\eta')$ and the corresponding phase $v(\eta')$ for the GaAs (1 1 1) Bragg reflection at $E_\gamma = 15$ keV. In this case, $w = 40.7\,\mu\text{rad} = 8.39$ arcsec. Semi-empirically, for the reflectivity curve, FWHM $= 1.2\,w$. Referring to the identical $R(\theta)$ curve in **Figure 3b**, notice that the center of the rocking curve is shifted slightly above the geometrical Bragg angle θ_B by $\sim 34\,\mu$rad. This shift is the result of refraction at the crystal–air interface. In general, this shift is $\Delta\theta = \Gamma F'_0/\sin 2\theta_B$. The asymmetry in the reflectivity curve, namely, the further diminishing from $R = 1$ as the angle is increased through the strong Bragg condition, is due to the movement of the XSW. On the high-angle side ($\eta' < -1$), the XSW antinodes align with the strong X-ray absorption planes in the crystal (see **Figure 4**). Therefore, absorption is higher than average on the high-angle side and weaker on the low-angle side ($\eta' > 1$) for this case.

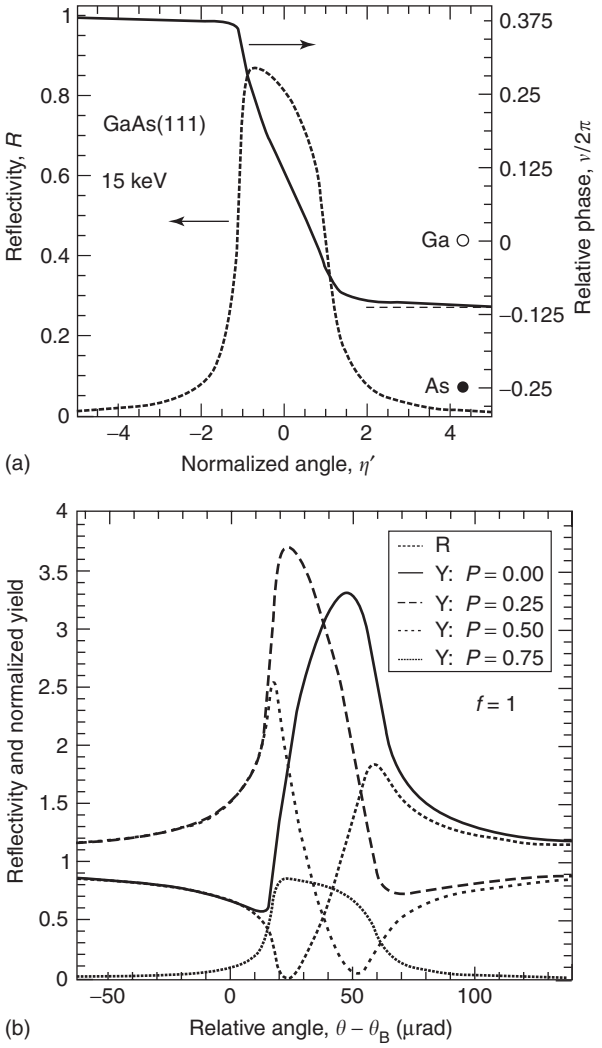


Figure 3 (a) The theoretical angle η' dependence of the reflectivity R and XSW phase $v/2\pi$ for the GaAs(1 1 1) Bragg reflection at $E\gamma = 15$ keV. (b) The corresponding theoretical angle $\Delta\theta$ dependence of the reflectivity and normalized fluorescence yields (eqn [15]) for coherent positions $P_{111} = 0.0, 0.25, 0.5$, and 0.75 with coherent fraction $f_{111} = 1$ and $Z(\theta) = 1$. Referring to **Figure 4**, this calculation has Ga at the 0, 0, 0 and As at the 1/4, 1/4, 1/4 position of the cubic zinc blende unit cell. Therefore, the $P_{111} = 0$ curve is the E-field intensity at the Ga site and the $P_{111} = 0.75$ (or $= -0.25$) curve is the E-field intensity at the As site. The dashed horizontal line in (a) represents the phase $\phi/2\pi$ of the structure factor. The respective phases of the geometrical structure factors for the Ga and As sublattices are illustrated with respect to the phase scale on the right-hand side of (a).

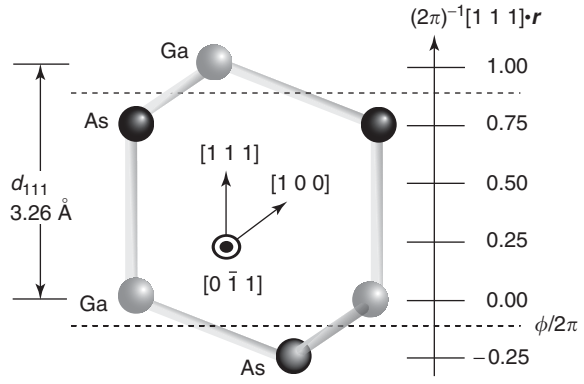


Figure 4 A [0 $\bar{1}$ 1] projection of the zinc-blende GaAs structure showing the [1 1 1] projected scale for the lattice positions in units of the d_{111} *d*-spacing. The horizontal dashed lines represent the phase of the GaAs (1 1 1) structure factor at 15 keV.

The exponential damping factor in eqn [5] accounts for attenuation effects within the crystal, in which case the effective absorption coefficient is defined as

$$\mu_z(\theta) = \frac{\mu_0}{\sin\theta_B}\left[1 + \frac{F''_{\bar{H}}}{F''_0}\left(\frac{E_H}{E_0}\right)'' + \frac{F''_{\bar{H}}}{F''_0}\left(\frac{E_H}{E_0}\right)'\right] \quad [12]$$

where $\mu_0 = (2\pi/\lambda)\Gamma F''_0$ is the linear absorption coefficient. The second and third terms in eqn [12] account for the extinction effect that strongly limits the X-ray penetration depth $1/\mu_z$ for a strong Bragg reflection. For example, the penetration depth for 15 keV X-rays at the GaAs (1 1 1) Bragg reflection ranges from 2.62 μm at off-Bragg conditions to 0.290 μm at the center ($\eta' = 0$) of the Bragg rocking curve. The general expression for this minimum penetration depth or extinction length is

$$\Lambda_{ext} = V_c\left[4d_H r_e\left(F''_0 + \sqrt{|F_H||F_{\bar{H}}|}\right)\right]^{-1} \quad [13]$$

XSW Photo-Effect Yields

The XSW field established inside the crystal and above the crystal surface induces photoelectron emission from atoms within the field. The excited ions, in turn, emit characteristic fluorescence X-rays and Auger electrons. In the dipole approximation, the photoelectric effect cross section is proportional to the *E*-field intensity at the center of the atom. (It is necessary to consider higher-order multipole terms in the photoelectric cross section under special conditions; for example, XSW induced photoemission that is not angle-integrated.) For the discussion that follows, the dipole approximation is assumed, in which case the normalized X-ray fluorescence yield is defined as

$$Y(\theta) = \int I(\theta, \boldsymbol{r})\rho(\boldsymbol{r})\,\exp[-\mu_f(\alpha)z]\mathrm{d}\boldsymbol{r} \quad [14]$$

where $\rho(\boldsymbol{r})$ is the normalized fluorescent atom distribution and $\mu_f(\alpha)$ is the effective absorption coefficient for the emitted fluorescent X-rays, which is dependent on their takeoff angle α. Upon integration, the normalized XSW yield is given as

$$Y(\theta) = [1 + R(\theta) + 2\sqrt{R(\theta)}f_H \times \cos(v(\theta) - 2\pi P_H)]Z(\theta) \quad [15]$$

where the parameters f_H and P_H are the coherent fraction and coherent position, respectively. In more general terms, f_H is the amplitude and P_H is the phase of the ***H***th-order Fourier coefficient of the normalized distribution function

$$F_H = \int_{uc} \rho(\boldsymbol{r})\exp(\mathrm{i}\boldsymbol{H}\cdot\boldsymbol{r})\mathrm{d}\boldsymbol{r} = f_H\,\exp(2\pi\mathrm{i}P_H) \quad [16]$$

$Z(\theta)$ is the effective-thickness factor, which is discussed below. $Z(\theta) = 1$ for atoms above the surface of the crystal and at a depth much less than the extinction length, $Z(\theta) \sim 1$.

Referring to **Figure 5**, in an XSW experiment, the reflectivity rocking curve $R(\theta)$ and the fluorescence yield $Y(\theta)$ are acquired simultaneously while scanning in θ through each *hkl* Bragg reflection of the sample crystal. For each *hkl* XSW data set, a rocking curve calculated according to dynamical diffraction theory and convoluted with the emission function of the upstream monochromator is fitted to the experimentally measured $R(\theta)$ to fix the absolute angle scale. This is then used for fitting the convolution of eqn [15] to the measured $Y(\theta)$. From this fit, the coherent fraction f_H and the coherent position P_H are obtained. The off-Bragg yield, which is also obtained from the fit to the measured data, is used for overall normalization of the yield, which gives the fluorescent atom concentration or surface coverage. In **Figure 3b**, calculated $Y(\theta)$ curves are shown for $Z(\theta) = 1$, $f_{111} = 1$, and $P_{111} = 0.0$, 0.25, 0.5, and 0.75. These calculated curves are for the GaAs (1 1 1) reflection at $E_\gamma = 15$ keV. The origin of the zinc blende unit cell was chosen to coincide with a Ga atom. Referring to **Figure 4**, the + (1 1 1) polarity of this noncentrosymmetric crystal structure has Ga at the top of the bilayer and As at the bottom. Therefore, the $P_{111} = 0$ curve corresponds to the expected K fluorescence yield for Ga atoms, and the $P_{111} = 0.75$ corresponds to that of As K fluorescence. The

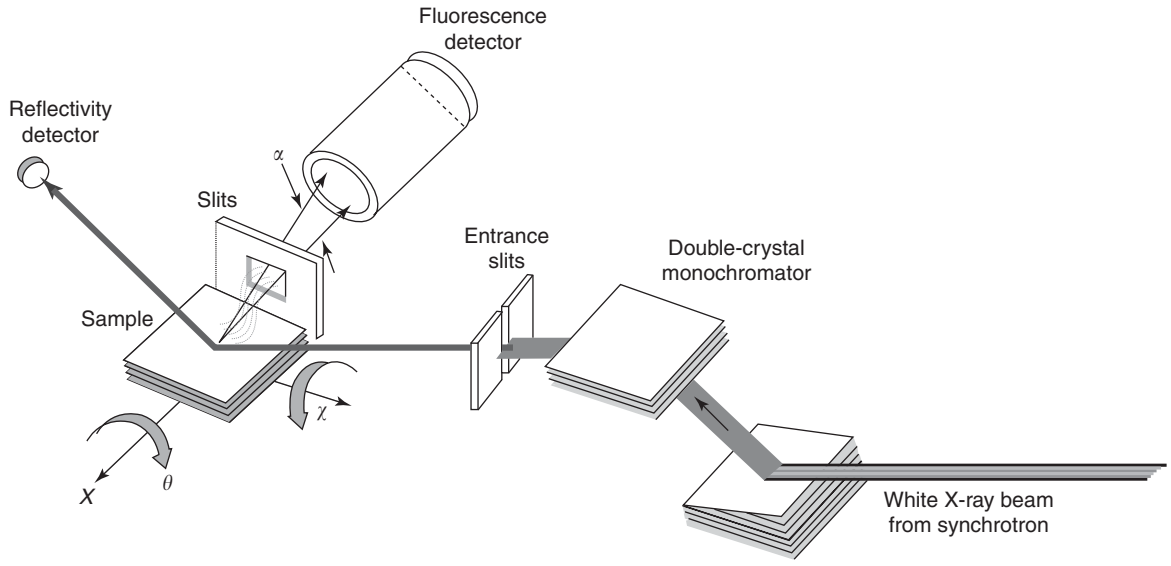


Figure 5 Depiction of a typical XSW experimental setup. The beam from a synchrotron beamline is monochromated and collimated by an asymmetrically cut, grazing incidence, Si single crystal; the second (nondispersive) symmetrically cut Si monochromator crystal is used to return the incident beam to a horizontal direction. While stepping the incident angle θ of the sample through the reflection, the reflected intensity and fluorescence spectrum are simultaneously collected by two separate detector systems.

marked change in the angular variation of $Y(\theta)$ for each increment of 0.25 in d-spacing is the basis for the spatial sensitivity of the Bragg diffraction XSW. If the GaAs(1 1 1) wafer were to be flipped over, the $(\bar{1}\,\bar{1}\,\bar{1})$ rocking curve would look essentially the same ($|F_{\bar{1}\bar{1}\bar{1}}|$ is 1.5% smaller than $|F_{111}|$ due to anomalous dispersion). However, the Ga and As XSW fluorescence yields would be essentially switched around, since the XSW antinodes pass through the Ga atomic layers when scanning in angle through the (1 1 1) reflection, and pass through the As atomic layers for the $(\bar{1}\,\bar{1}\,\bar{1})$.

Extinction Effect and Evanescent-Wave Emission

The effective-thickness factor $Z(\theta)$ in eqn [15] accounts for the θ dependence of the penetration depth of the primary X-ray field (extinction effect) in conjunction with the escape depth, Λ, of the outgoing secondary fluorescence X-rays. For atoms at the crystal surface (e.g., adsorbates) or at a depth much smaller than the extinction depth, the effective-thickness factor is constant at $Z(\theta) = 1$. For atoms evenly distributed throughout the semi-infinite crystal,

$$Z(\theta) = \frac{\mu_0(\sin\theta_B)^{-1} + \mu_f(\alpha)}{\mu_z(\theta) + \mu_f(\alpha)} \qquad [17]$$

where $\mu_z(\theta)$ is the effective absorption coefficient of the incident X-rays (eqn [12]), and $\mu_f(\alpha)$ is the effective absorption coefficient of the outgoing fluorescence X-rays from the crystal at takeoff angle α. To achieve surface sensitivity of substrate atoms, it is possible to reduce α to a value approaching the critical angle α_c of the fluorescence X-rays. Under such conditions, $\mu_f(\alpha)$ dominates over $\mu_z(\theta)$ in eqn [17] and, therefore, the effective thickness factor is constant at $Z(\theta) = 1$. The value of $\mu_f(\alpha)$ is dependent on the wavelength λ_f of the fluorescence and the index of refraction calculated at $\lambda = \lambda_f$. The index of refraction is

$$n = 1 - \delta - i\beta \qquad [18]$$

where

$$\begin{aligned} \delta &= \frac{1}{2\pi} N_e r_e \lambda^2 = \frac{1}{2}\,\Gamma F_0' \\ \beta &= \frac{1}{4\pi}\mu_0\lambda = \frac{1}{2}\,\Gamma F_0'' \end{aligned} \qquad [19]$$

and N_e is the effective electron density of the refractive medium. From the index of refraction, $\mu_f(\alpha)$ can be obtained as

$$\mu_f(\alpha) = \frac{2\sqrt{2}\pi}{\lambda_f}\left[\sqrt{(2\delta - \alpha^2)^2 + 4\beta^2} + 2\delta - \alpha^2\right]^{1/2} \qquad [20]$$

At takeoff angles much greater than the critical angle (i.e., $\alpha \gg \alpha_c = \sqrt{2\delta}$), $\mu_f = \mu_0/\sin\alpha$. Below the critical angle, refraction effects dominate and the escape depth $\Lambda = \mu_f^{-1}$ approaches $1/4(N_e r_e)^{-1/2}$.

The evanescent-wave emission effect can be used to remove the extinction effect of bulk fluorescence in XSW experiments. This is important, since for $\alpha \gg \alpha_c$, the extinction effect tends to dominate over the standing-wave effect, causing loss of phase (positional) sensitivity.

Structure Determination Using Coherent Fraction and Coherent Position

The coherent fraction f_H senses the spread in the spatial distribution of the fluorescent atoms. For simplicity, f_H is subdivided into three factors:

$$f_H = Ca_H D_H \qquad [21]$$

where C is the ordered fraction, a_H is the geometrical factor, and D_H is the Debye–Waller factor. All three factors range in value from 0 to 1.

Consider the general case of a discrete distribution having the same atom located at N different unit cell positions $\boldsymbol{r}_j = x_j\boldsymbol{a} + y_j\boldsymbol{b} + z_j\boldsymbol{c}$, plus an added random distribution of the same atoms. The ordered fraction C is the fraction of the atoms in the distribution that are coherent (or are crystallographically registered) with the substrate crystal lattice. If the occupation fractions of the atoms for the ordered positions are $c_1, c_2, \ldots, c_N$, respectively, the ordered fraction is

$$C = \sum_{j=1}^{N} c_j \qquad [22]$$

At this point, it is useful to define the normalized geometrical structure factor S_H for the ordered fluorescent-selected atoms:

$$S_H = \frac{1}{C}\sum_{j=1}^{N}[c_j \exp(\mathrm{i}\boldsymbol{H}\cdot\boldsymbol{r}_j)] \qquad [23]$$

The geometrical factor is defined as

$$a_H = |S_H| \qquad [24]$$

and the coherent position is defined as

$$P_H = \frac{\arg[S_H]}{2\pi} \qquad [25]$$

It is to be noted that the origin for the set of $\boldsymbol{r}_j$ (and, therefore, the origin of P_H) in the unit cell is the same as was arbitrarily chosen for generating the structure factor F_H used in eqns [8]–[12]. For the simple case of one atomic site, $N = 1$, $a_H = 1$, and P_H is the projected $\boldsymbol{H}\cdot\boldsymbol{r}/2\pi$ fractional d-spacing position of the atomic site. For the case of two equally occupied atom sites, the geometrical factor reduces to

$$a_H = \left|\cos\left[\frac{\boldsymbol{H}\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{2}\right]\right| \qquad [26]$$

The coherent position in this case is the averaged fractional d-spacing position of the two sites. If the two equally occupied sites have a separation of exactly one-half of a d-spacing along a particular $\boldsymbol{H}$, then $a_H = 0$ for that particular $\boldsymbol{H}$. This is analogous to a forbidden reflection in crystallography. Again, for the case of two sites separated by one-half of a d-spacing, if the site occupation fractions are 2/3 and 1/3, the geometrical fraction will be $a_H = 1/3$, and the coherent position will equal the position of the 2/3 occupied site.

In XSW analysis, the Debye–Waller factor D_H accounts for the time-averaged spatial distribution in the $\boldsymbol{H}$ direction due to thermally induced vibrations of the fluorescence-selected ordered atoms about their average lattice positions. In general, individual D_H factors can be assigned to the individual sites. In addition to including the dynamic (thermal vibration) distribution, one can also include the static (spatial disordering) distribution. Generally, D_H can be expressed in terms of the mean-square vibrational amplitude along the $\boldsymbol{H}$ direction $\langle u_H^2 \rangle$, as

$$\boldsymbol{D}_H = \exp(-M) = \exp\left(\frac{-2\pi^2\langle u_H^2\rangle}{d_H^2}\right) \qquad [27]$$

XSW measurements are not restricted to being made along the surface-normal direction (i.e., $\boldsymbol{H}$ perpendicular to the crystal surface). They can be performed with respect to any sufficiently strong Bragg reflection of a crystal. A 3D triangulation of the atom site can be obtained by combining XSW measurements by using three mutually noncollinear diffraction vectors. In triangulating an atomic site, the point symmetry of the surface can be used to reduce the number of required XSW measurements.

Direct-Space Imaging

The XSW measured coherent fraction and position are model-independent and correspond to the amplitude f_H and phase P_H of the $\boldsymbol{H}$th Fourier coefficient of the fluorescence-selected atomic distribution. Therefore, based on the Fourier inversion of eqn [16], the distribution $\rho(\boldsymbol{r})$ of each fluorescent atomic species can be synthesized directly by the Fourier summation

$$\begin{aligned}\rho(\boldsymbol{r}) &= \sum_H \rho_H = \sum_H f_H \exp[\mathrm{i}(2\pi P_H - \boldsymbol{H}\cdot\boldsymbol{r})] \\ &= 1 + 2\sum_{\substack{H\neq -H \\ H\neq 0}} f_H \cos[2\pi P_H - \boldsymbol{H}\cdot\boldsymbol{r}]\end{aligned} \qquad [28]$$

The above simplification to a summation of cosine terms makes use of $f_0 = 1$, and the symmetry relationship analogous to Friedel's law that makes $f_{\bar{H}} = f_H$ and $P_{\bar{H}} = -P_H$. Since the summation is limited to allowed bulk reflections, the image $\rho(\boldsymbol{r})$ generated

from the XSW measured f_H and P_H values, is limited to being a projection into the primitive unit cell of the bulk crystal. The quality (resolution) of the XSW image will depend on the extent in reciprocal space over which the set of f_H and P_H values were measured.

As a simple example, consider the Ge diamond-cubic crystal structure (identical to zinc blende shown in **Figure 4**, but with Ga and As replaced by Ge). In **Figure 6**, the 1D projection of the Ge Fourier components ρ_{hhh} are plotted over an interval corresponding to the d_{111} 1D unit cell and then summed according to eqn [28] to produce the 1D image. The extra oscillations are due to truncating the sum and zeroing all terms higher than $hhh = 888$.

The resolution of the image in a given [hkl] direction is equivalent to one-half of the smallest d-spacing that has been entered into the summation in that direction. This imaging method in 3D is especially useful for solving multisite surface adsorbate phases. After this typically 0.5 Å resolution model-independent analysis step, the measured f_H and P_H values can be used to refine the parameters of a structural model (suggested by the imaging method). This previously discussed model-dependent analysis method has a typical resolution of 0.03 Å.

Solution to Phase Problem

Unlike conventional diffraction methods, the XSW method does not lose phase information and can, therefore, be used directly to map the direct-space structure from the set of Fourier coefficients collected in reciprocal space. The XSW measurement does not lose phase information, because the detector of the E-field is the fluorescent atom itself, lying within the spatial region where the fields interfere coherently with each other. In contrast, in conventional diffraction measurements, the relative phase between the diffracted and incident fields is lost, because the intensities of the incident and scattered fields are detected far from this region of coherent spatial overlap.

It should be stressed that the Bragg XSW positional information acquired is in the same absolute coordinate system as used for describing the substrate unit cell. This unit cell and its origin were previously chosen when the structure factors F_H and $F_{\bar{H}}$ (as defined in eqn [9]) were calculated and used in eqns [8], [10], and [12]. The XSW phase (v) is directly linked to the phase (ϕ_H) of the structure factor. Referring to **Figures 3a** and **4**, v approaches $\phi + \pi$ at angles well below the strong Bragg condition and v approaches ϕ, far above. This corresponds precisely to the XSW antinodes shifting inward, from being halfway between the diffraction planes on the low-angle side to being coincident with the diffraction planes on the high-angle side. This XSW-antinode-asymptotic-limit definition for the hkl diffraction

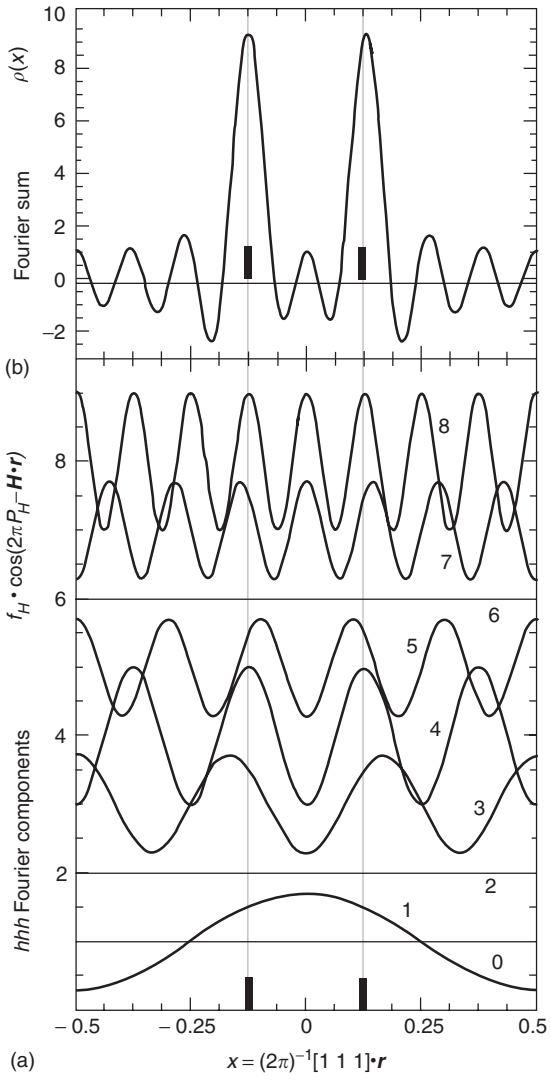


Figure 6 (a) The 1D spatial dependence of the *hhh* Fourier components (ρ_{hhh}, $h = 0-8$) for the distribution of Ge atomic centers in the diamond-cubic structure. (Each curve is given a vertical offset of $+h$.) This is identical to the GaAs zinc-blende structure shown in **Figure 4**, but with Ga and As replaced by Ge. The spatial interval corresponds to the 1D unit cell along the [1 1 1] direction that has a lattice constant corresponding to d_{111} and has two equally occupied Ge sites at $x = \pm 1/8$ (two vertical lines). The h dependence of the amplitude f_{hhh} and phase P_{hhh} for each Fourier component is calculated from eqns [23]–[26] for the case of $C = 1$ and $D_{hhh} = 1$. In this case $f_{hhh} = a_{hhh} = |cos(\pi h/4)|$. This makes $f_{000} = f_{444} = f_{888} = 1$, $f_{111} = f_{333} = f_{555} = f_{777} = 1/\sqrt{2}$, and $f_{222} = f_{666} = 0$. $P_{111} = P_{777} = P_{888} = 0$, and $P_{333} = P_{444} = P_{555} = 1/2$. (b) The calculated 1D image $\rho(x)$ of Ge atom centers as summed over the terms from $hhh = 000$ to 888 based on eqn [28]. The curves in (a) can also represent the phase and period of the XSW at the high-angle side ($\eta' < -1$) of the Bragg peak for the case of $\Delta f'' = 0$.

planes, defines the diffraction planes as coinciding with the maxima of the real part of the $\boldsymbol{H}$th Fourier component of the scattering density function, which is expressed as $\max\{\mathrm{Re}[F_H \exp(-i\boldsymbol{H}\cdot\boldsymbol{r})]\}$. Consequently, the correct values for f_0, $\Delta f'$, and $\Delta f''$ must be used to compute the structure factor F_H in order to assure an accurate determination of the absolute lattice position of an impurity or adsorbate atom. This is typically quite straightforward, unless the incident energy is near an absorption edge for one of the atoms in the bulk crystal.

"Ideal Crystals" versus "Real Crystals" in Bragg Diffraction XSW

According to eqn [11], the intrinsic Darwin width for a typical strong Bragg diffraction peak in the reflection geometry is between 5 and 100 μrad. To produce useful quantitative information from an XSW measurement, the measured reflectivity curve should reasonably match theory. Only a few exceptionally high-quality crystals qualify as ideal crystals according to this criterion, for example, Si, Ge, GaAs, and calcite. Most real single crystals contain internal imperfections that result in mosaic spreads that exceed their intrinsic Darwin widths. These mosaic spreads can reduce the reflected intensity and subsequently smear out the standing-wave effect. For these real crystals, the measured reflectivity curves cannot be fitted directly by the ideal theoretical reflectivity described by eqn [6]. To make the Bragg diffraction XSW technique adaptable to applications on real crystals, a number of theoretical and experimental methods have been developed. The most formal method is the modification of the XSW theory to include the effect of crystal imperfection on Bragg diffraction. On the other hand, one can make XSW applications under certain conditions where the stringent requirements of the conventional theory are relaxed. One such condition is when a Bragg reflection occurs near the back-reflection geometry, typically when $87^\circ < \theta_B < 90^\circ$. Under this condition, eqn [11] breaks down, and the intrinsic Darwin width is magnified to milliradians. Another condition is when the crystal is very thin; under this condition the kinematical theory can replace the dynamical theory in calculating the Bragg reflection. Both of these special cases, which are discussed briefly below, are important alternatives when the conventional XSW technique cannot be applied.

An additional practical method for applying XSW to real crystals takes advantage of the high brilliance of X-ray undulators at third-generation synchrotron sources. With such a source, sufficiently high X-ray intensities can be delivered with beams slitted down to microns in cross section, thus making it possible to illuminate one isolated "perfect" grain at the sample surface.

The Special Case of Back-Reflection XSW

As just mentioned, one means of circumventing the problem of analyzing less-than-perfect crystals is to use dynamical diffraction at Bragg angles approaching 90°. In this back-reflection geometry, the angular Darwin width is measured in milliradians instead of microradians.

For $\theta_B > 88^\circ$, the conventional dynamical diffraction theory breaks down. This is due to an approximation in the conventional theory that treats the spherical asymptotes of the dispersion surface in reciprocal space as planes. With an extended dynamical diffraction theory, the Bragg reflectivity and the standing wave E-field intensity can be properly described for $87^\circ < \theta_B < 90^\circ$. In this regime, the Bragg reflectivity has a much smaller energy width and a much broader angle width. For a given d-spacing, the largest angle width occurs at a wavelength of

$$\lambda_b = 2d_H\left(1 - \frac{\Gamma(F_0' - F_H')}{2}\right) \qquad [29]$$

In this case, $\Gamma = 4r_e d_H^2(\pi V_c)^{-1}$. For wavelengths slightly smaller, the angular width can be expressed as

$$w = \Delta\theta_{y=1} - \Delta\theta_{y=-1} \qquad [30]$$

where $\Delta\theta_{y=1}$ and $\Delta\theta_{y=-1}$ are the angular displacements from 90° of the low- and high-angle sides of the strong Bragg diffraction condition, respectively. For a symmetric reflection,

$$\Delta\theta_{y=\pm 1} = \sqrt{2(1 - \sin\theta_B) + \Gamma(F_0' \mp F_H')} \qquad [31]$$

The back-reflection XSW (BRXSW) technique has been primarily applied to metal and oxide crystals, whose rocking curves typically exhibit angular mosaicity up to ~0.1°. BRXSW is mostly used in combination with XPS (rather than XRF) for ultrahigh vacuum (UHV) surface science measurements of adsorbed molecules and for site-specific valence-band photoemission studies.

The Special Case of Thin-Film Bragg Diffraction XSW

The development of Bragg diffraction XSW from thin films is partially driven by the fact that many crystals can be grown as high-quality μm-thick films but not as large-size crystals. Similar to Bragg diffraction from a bulk crystal, Bragg diffraction from a

crystalline thin film also generates an X-ray standing-wave field. Because the thickness of the film is much less than the extinction depth, kinematical diffraction theory can be used as a good approximation to calculate the intensity of the field. The rocking curve angular width $w \sim 1/N$, where N is the atomic layer thickness of the film. The analytical procedure is fully analogous to that for bulk-reflection XSW. However, the small thickness results in very weak peak reflectivity (typically 0.01%), and consequently, very weak (typically 1%) modulation of the XSW-induced fluorescence yield from an atom within or above the film. To observe this small modulation above the statistical fluctuation, a large number of fluorescence counts must be collected at each angle step of the XSW scan.

XSWs Generated by TER

While a single-crystal XSW provides a high-resolution probe well suited for atomic-scale structural determination, this XSW period is too fine a scale to profile larger nanostructures – such as the diffuse-ion distribution at the solid–liquid interface and organic self-assembled multilayers. To extend the XSW technique to the nanoscale and beyond, one can employ a long-period XSW generated by TER from a mirror surface or generated by Bragg diffraction from a periodically layered synthetic microstructure (LSM). Because the reflection condition in these two cases occurs at much smaller angles, the XSW period, D, is much longer. For TER, in particular, D varies from 1 μm to 10 nm as θ increases through the TER condition.

TER–XSW Theory

During TER, the interference between the incident and the specular reflected X-ray plane waves produces an XSW above the mirror surface (inset to **Figure 7**). An evanescent wave is induced below the surface. The nodal (and antinodal) planes of this periodic E-field intensity pattern above the surface are parallel to the surface and have a variable period of $D = \lambda/(2\sin\theta)$, as defined in eqn [1]. The TER condition occurs between $\theta = 0$ and $\theta = \theta_c$, the critical angle. The corresponding XSW period is very long, ranging from $D = \infty$ to $D = D_c = \lambda/2\theta_c$, the critical period. From eqn [19] and $\theta_c = \sqrt{2\delta}$, the critical period can be written as

$$D_c = \frac{1}{2}\sqrt{\frac{\pi}{N_e r_e}} \qquad [32]$$

Aside from small anomalous dispersion effects, the critical period is wavelength independent. D_c is a materials property dependent on the electron density, N_e. For an Au mirror $D_c = 80$ Å, and for an Si mirror $D_c = 200$ Å. The directly related critical scattering vector is $Q_c = 2\pi/D_c$.

This TER condition is, in fact, the zeroth-order Bragg diffraction condition, or the condition when $\boldsymbol{H} = 0$ and $d_H = \infty$. Therefore, the above derivations from dynamical diffraction theory for single-crystal Bragg diffraction can be applied to the case of TER by substituting F_0 for F_H. In this case, eqn [10] becomes

$$\eta = \eta' + i\eta'' = \frac{-2\theta^2 + \Gamma F_0}{\Gamma F_0} \qquad [33]$$

with the substitutions

$$\eta' = 1 - 2x^2 \quad \text{and} \quad \eta'' = 2x^2 y \qquad [34]$$

where $x = \theta/\theta_c$ is the normalized angle parameter, and $y = \beta/\delta$ is the absorption factor. From eqn [8], the complex E-field amplitude ratio becomes

$$\begin{aligned}\frac{E_R}{E_0} &= -(\eta \pm \sqrt{\eta^2 - 1}) \\ &= \frac{x - \sqrt{x^2 - 1 - iy}}{x + \sqrt{x^2 - 1 - iy}} = \left|\frac{E_R}{E_0}\right| \exp(iv)\end{aligned} \qquad [35]$$

where E_R is the complex amplitude of the reflected plane wave. This amplitude ratio is identical to that derived from classical Fresnel theory. For the simple case of no absorption, where $\beta = 0$, the reflectivity is

$$\begin{aligned}R &= \left|\frac{E_R}{E_0}\right|^2 \\ &= \begin{cases} 1, & 0 \leqslant x \leqslant 1 \\ 8x^4 - 8x^3\sqrt{x^2 - 1} + 4x\sqrt{x^2 - 1} & \\ \quad -8x^2 + 1, & x > 1 \end{cases}\end{aligned} \qquad [36]$$

and the phase is

$$v = \begin{cases} \cos^{-1}(2x^2 - 1), & 0 \leqslant x \leqslant 1 \\ 0, & x > 1 \end{cases} \qquad [37]$$

Referring to **Figure 7a**, the phase at the mirror surface decreases from π to 0 as the mirror is tilted through the total reflection condition. Thus, at the mirror surface, where $z = 0$, the reflected plane wave is completely out of phase with the incident plane wave when $\theta = 0$. As the incident angle is increased, the phase decreases smoothly until it is completely in phase at $\theta = \theta_c$. Therefore, at $\theta = 0$, a standing-wave node is at the mirror surface, and the first antinode is at infinity. As θ increases, the first antinode moves in from infinity toward the mirror

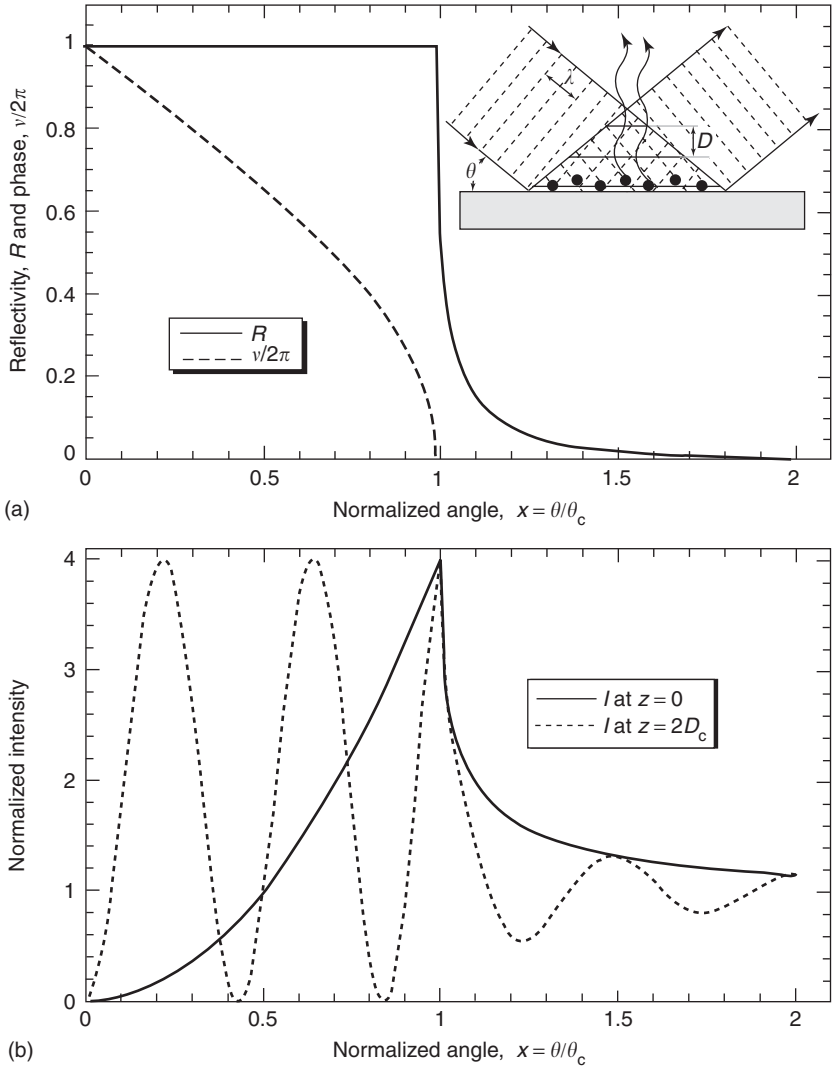


Figure 7 Under the condition of TER from an ideal mirror surface with no absorption ($\beta = 0$): (a) The angular dependence of the reflectivity R and relative phase v of the reflected X-ray plane wave (or phase of the XSW relative to the mirror surface); (b) the angular dependence of the normalized E-field intensity at the mirror surface ($z = 0$) and at a distance $z = 2D_c$ (160 Å for Au mirror, 400 Å for Si mirror). The inset depicts how the XSW field is formed above a mirror surface by the interference between an incident and a specular-reflected X-ray plane wave.

surface, until it coincides with the mirror surface upon reaching $\theta = \theta_c$. At the same time, the second-, third-, and higher-order antinodes of the standing wave also move toward the surface, as the period D decreases based on eqn [1].

The normalized E-field intensity above the mirror surface can be expressed as

$$I(\theta, z) = \frac{|\boldsymbol{E}_0 + \boldsymbol{E}_R|^2}{|\boldsymbol{E}_0|^2} = 1 + R + 2\sqrt{R}\cos(v - Qz) \qquad [38]$$

Figure 7b shows the angular dependence of the E-field intensity at $z = 0$ (the mirror surface) and at $z = 2D_c$. The fluorescence signal for an ideally narrow single atomic plane fixed at these heights would have the same angular dependence.

The normalized fluorescence yield from an arbitrary distribution of atoms $\rho(z)$ above the mirror surface can be obtained by integrating over all values of z:

$$Y(\theta) = \int_{z=0}^{z=\infty} I(\theta, z)\rho(z)\,\mathrm{d}z \qquad [39]$$

With $I(\theta, z)$ calculated from eqn [38], the atom distribution profile $\rho(z)$ can be obtained by assuming a modeled distribution and fitting it to the measured yield $Y(\theta)$. For the specific cases of δ-function atomic

distributions in a plane at $z = xD_c$, there are $x + (1/2)$ modulations between $\theta = 0$ and $\theta = \theta_c$. The extra 1/2 modulation is due to the π phase shift in v.

With TER–XSW, the Fourier transform of an atom distribution is measured over a continuous range in $Q = 2\pi/D$, with variable period D ranging from ~ 100 Å to 1 μm. Therefore, TER–XSW is ideally suited to measure surface and interface structures of length scales in the range of 50–2000 Å.

The above treatment accurately describes the X-ray E-fields at and above the mirror surface for the simple case of an interface with vacuum (or air). To apply TER–XSW as a probe for studying liquid–solid interfacial structures or organic films deposited on a solid surface, it is necessary to include reflection, refraction, and absorption effects from the layers that lie between the substrate surface and the vacuum (or air). This can be accomplished by making use of Parratt's recursion formulation to calculate the transmitted and reflected fields at any interface. These same fields can then be calculated at any point within the slab by appropriately accounting for the X-ray absorption and refraction effects on the fields as they travel from the interface to the point. The TER–XSW method is primarily used to profile nanoscale heavy-atom (or -ion) distributions in organic films and at electrified liquid/solid interfaces.

XSWs from Layered Synthetic Microstructures

For Bragg diffraction purposes, an LSM is fabricated (typically by sputter deposition) to have a depth-periodic layered structure consisting of 10–200 layer pairs of alternating high- and low-electron-density materials, such as Mo and Si. Sufficient uniformity in layer thickness is obtainable in the range between 10 and 150 Å (d-spacing of fundamental diffraction planes from 20 to 300 Å). Because of the rather low number of layer pairs that affect Bragg diffraction, these optical elements (when compared to single crystals) have a significantly wider energy band pass and angular reflection width. The required quality of an LSM is that experimental reflection curves compare well with dynamical diffraction theory, and peak reflectivities are as high as 80%. Therefore, a well-defined XSW can be generated and used to probe structures deposited on an LSM surface with a periodic scale equivalent to the rather large d-spacing. To a good approximation, the first-order Bragg diffraction planes coincide with the centers of the high-density layers of the LSM. Above the surface of the LSM, the XSW period is again defined by eqn [1]. The reflectivity can be calculated by using Parratt's recursion formulation. This same optical theory can be extended to allow the calculation of the E-field intensity at any position within any of the slabs over an extended angular range that includes TER. Then, eqn [39] is used to calculate the fluorescence yield. The LSM–XSW method is primarily used

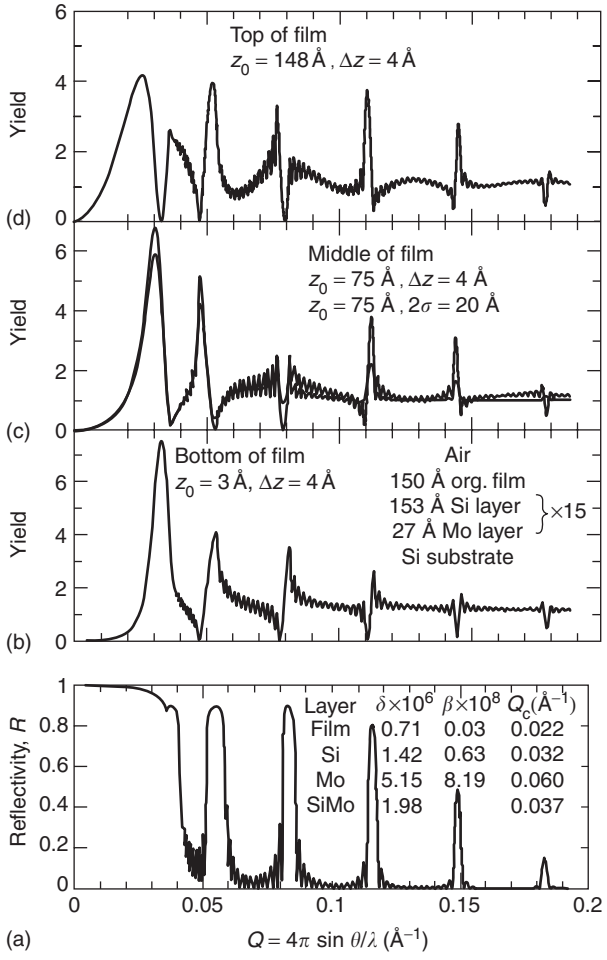


Figure 8 An XSW simulation at $E\gamma = 18.5$ keV for an Si/Mo multilayer (LSM) with $d = 180$ Å, $t_{Si} = 153$ Å, $N = 15$ layer pairs, an Si substrate and a 150 Å thick organic film on top. Each interface was given a $\sigma = 3$ Å Gaussian smear to simulate roughness. The index of refraction values for each type of layer (plus the weighted average Si/Mo LSM) are listed in (a). The angular dependence of the (a) reflectivity and (b)–(d) fluorescence yields for 4 Å thick uniform atomic distributions at the (b) bottom, (c) middle, and (d) top of the organic film are shown. Also shown in (c) is the case of a 20 Å wide Guassian distribution. This Q range includes the zeroth (TER) through 5th-order Bragg peaks. The XSW period in the air (or vacuum) is $D = 2\pi/Q$. This is also a good approximation for the XSW period inside the organic film, except when refraction effects are appreciable, i.e., $Q < 2Q_c^{Film}$. More precisely, $D_{Film} = 2\pi/(Q^2 - (Q_c^{Film})^2)^{1/2}$ for $Q > Q_c^{Film}$. The sharp drop in R at $Q = Q_c^{Si}$, corresponds to the excitation of a resonant cavity mode that has a node at the top of the organic film and a node at the bottom of the first Si layer. At this precise angle, the E-field intensity in the middle of this combined Si-film slab is much higher than $4|\boldsymbol{E}_0|^2$, as can be seen in (b).

to determine the atomic (or ionic) distributions in deposited organic films or at electrified liquid–solid interfaces. As an example, the simulated reflectivity and fluorescence yields for a $d = 180$ Å Si/Mo LSM is shown in **Figure 8**. It can be noted that the modulation pattern of the yield is very sensitive to the height and width of the fluorescent atom layer within an organic film deposited on the LSM.

See also: Core Photoemission; Crystal Structure; Crystal Structure Determination; Deep-Level Spectroscopy and Auger Spectra; Surfaces and Interfaces, Electronic Structure of; Synchrotron Radiation; Valence Photoemission; X-Ray Absorption Spectroscopy; X-Ray Sources; X-Ray Topography.

PACS: 61.10.–I; 68.49.Uv; 41.50.+h; 07.85.Qe

Further Reading

Authier A (2001) *Dynamical Theory of X-Ray Diffraction*. New York: Oxford University Press.

Batterman BW and Cole H (1964) Dynamical diffraction of X-rays by perfect crystals. *Review of Modern Physics* 36: 681–717.

Becker RS, Golovchenko JA, and Patel JR (1983) X-ray evanescent-wave absorption and emission. *Physical Review Letters* 50: 153–156.

Bedzyk MJ, Bommarito GM, and Schildkraut JS (1989) X-ray standing waves at a reflecting mirror surface. *Physical Review Letters* 62: 1376–1379.

Bedzyk MJ and Materlik G (1985) Two-beam dynamical diffraction solution of the phase problem: a determination with X-ray standing waves. *Physical Review B* 32: 6456–6463.

Born M, Wolf E, and Bhatia AB (1999) *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. New York: Cambridge University Press.

Cheng L, Fenter P, Bedzyk MJ, and Sturchio NC (2003) Fourier-expansion solution of atom distributions in a crystal using X-ray standing waves. *Physical Review Letters* 90: 255503-1-4.

Golovchenko JA, Patel JR, Kaplan DR, Cowan PL, and Bedzyk MJ (1982) Solution to the surface registration problem using X-ray standing waves. *Physical Review Letters* 49: 560–563.

Parratt LG (1954) Surface studies of solids by total reflection of X-rays. *Physical Review* 95: 359–369.

Tolan M (1999) *X-Ray Scattering From Soft-Matter Thin Films*. Berlin: Springer.

Vartanyants IA and Kovalchuk MV (2001) Theory and applications of X-ray standing waves in real crystals. *Reports of Progress in Physics* 64: 1009–1084.

Wang J, Bedzyk MJ, and Caffrey M (1992) Resonance-enhanced X-rays in thin-films – a structured probe for membranes and surface-layers. *Science* 258: 775–778.

Woodruff DP (1998) Normal incidence X-ray standing wave determination of adsorbate structures. *Progress in Surface Science* 57: 1–60.

Zegenhagen J (1993) Surface structure determination with X-ray standing waves. *Surface Science Reports* 18: 199–271.

Zegenhagen J and Kazimirov A (guest eds.) (2004) Focus on X-ray standing waves. *Synchrotron Radiation News* 17: 2–53.

Nomenclature

a_H	geometrical factor
C	ordered fraction
d	diffraction plane spacing (Å)
D	XSW period (Å)
D_c	critical period (Å)
D_H	Debye–Waller factor
$\boldsymbol{E}_H/\boldsymbol{E}_0$	E-field amplitute ratio
E_γ	X-ray photon energy (keV)
f_H	coherent fraction
$f^0(H)$	atomic form factor
F_H	normalized atomic distribution Fourier coefficient
F_H	structure factor
$\boldsymbol{H}$	reciprocal lattice vector (Å^{-1})
I	normalized E-field intensity
$\boldsymbol{K}$	wave vector (Å^{-1})
$n = 1-\delta-\mathrm{i}\beta$	index of refraction
N_e	effective electron density (Å^{-3})
P_H	coherent position
$\boldsymbol{Q}$	scattering vector (Å^{-1})
Q_c	critical scattering vector (Å^{-1})
$\boldsymbol{r}$	unit cell atomic position vector (Å)
r_e	classical electron radius (Å)
R	reflectivity
s_H	geometrical phase factor
S_H	geometrical structure factor
$\langle u_H^2 \rangle^{1/2}$	vibrational amplitude (rms) (Å)
V_c	volume of unit cell (Å^3)
w	Darwin width (radians)
Y	normalized yield
$Z(\theta)$	effective thickness factor (dimension)
α	takeoff emission angle (rad)
Γ	gamma scale factor
$\Delta f'$ and $\Delta f''$	anomalous dispersion corrections
η	normalized angle parameter
θ	incident angle (rad)
θ_c	critical angle (radians)
λ	wavelength (Å)
Λ_{ext}	extinction length (Å)
μ_0	linear absorption coefficient (Å^{-1})
μ_z	effective linear absorption coefficient (Å^{-1})
v	XSW phase
$\rho(\boldsymbol{r})$	normalized atomic distribution function
ϕ_H	phase of structure factor (radians)

X-Ray Topography

J Baruchel and J Härtwig, ESRF, Grenoble, France

Introduction

A major use of X-rays, ever since they were discovered over a century ago, has been and still is the visualization of the inside of systems, opaque for other probes. Whatever the imaging process used, it implies an inhomogeneous response of the sample to the beam. Recently, X-ray imaging techniques have dramatically progressed, in connection with the improvements in sources (synchrotron radiation), detectors, and computer image processing. Among these new features, three-dimensional imaging with spatial resolution in the micrometer range (microtomography), the sensitivity enhancement associated with phase contrast imaging, and new X-ray microbeam-based imaging approaches such as fluorescence or diffraction microtomography are discussed below.

This section is devoted to another way of obtaining X-ray images, based on Bragg diffraction: the "X-ray topographic" techniques (the "historical" name is retained: it is widely used, but misleading because it suggests that these techniques are surface techniques, whereas very often they are used to probe the bulk of single-crystal samples). They were mostly developed in the late 1950s and 1960s to visualize defects and distortions in crystals. Work performed on silicon made the routine production of large perfect crystals, as used by the microelectronics industry, possible. These techniques are nondestructive and sensitive to a large range of distortions (typically in the 10^{-3}–10^{-8} range). Understanding the main contrast mechanisms implies knowing how crystal imperfections modify the diffraction behavior of the underlying perfect crystal. The dynamical theory of diffraction describes this behavior. The specific contributions of synchrotron radiation (and, briefly, neutrons) are outlined.

Principle of the Topographic Techniques

X-ray topography is an imaging technique for single crystals based on Bragg diffraction. The local variations of the amplitude or direction of the diffracted beam provide information on crystal inhomogeneities such as defects, domains, or phases. This information is often averaged out, and lost, when using usual diffraction methods.

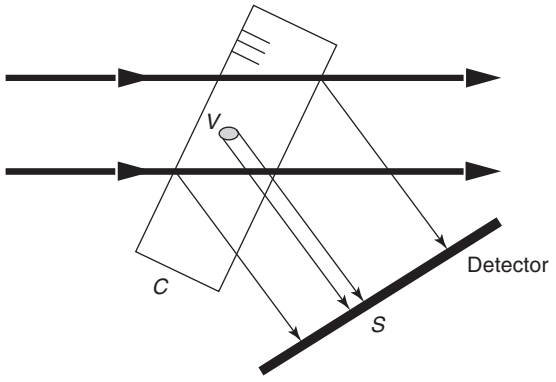


Figure 1 Principle of the topographic methods.

Figure 1 schematically represents the main idea of all the diffraction imaging methods. The single-crystal sample C is set to diffract a fraction of the incoming beam, according to Bragg's law, $2d \sin\theta_B = \lambda$. The diffracted beam is recorded on a position-sensitive detector (usually X-ray films, nuclear plates or, increasingly, CCD cameras equipped with a high-resolution scintillator). If a small volume V within C behaves differently from the matrix, intensity variations ("contrast") can be produced on an area S of the two-dimensional image produced by the diffracted beam. This contrast only occurs if the platelet-shaped crystal displays inhomogeneities.

The inhomogeneities that can be visualized using X-ray diffraction topography include defects such as dislocations (**Figure 2**), twins, domain walls, inclusions, impurity distribution, and macroscopic deformations, such as bending or acoustic waves, present within the single-crystal sample.

Figure 3 is an example of such a topographic image. It shows the image recorded using the 444 reflection from a flux-grown platelet-shaped gallium-substituted yttrium iron garnet crystal ($Y_3Ga_xFe_{5-x}O_{12}$, with $x \sim 1$, Ga–YIG) using the MoKα_1 radiation from a standard X-ray generator.

In addition to a distorted region on the upper corner, associated with the crystal fixation, various features are observable on this topograph: dislocations, growth striations and growth sectors, magnetic domains, dissolution bands, and the initial stage of a crack. These features are observed because they entail a distortion of the crystalline lattice. It is worth giving the typical value of these distortions, which include either a lattice rotation $\delta\theta$, or a lattice parameter variation $\Delta d/d$, or a combination of both. For the growth striations, which are the main observable features, this distortion is in the 10^{-4} range, whereas the variation of magnetostriction associated with the

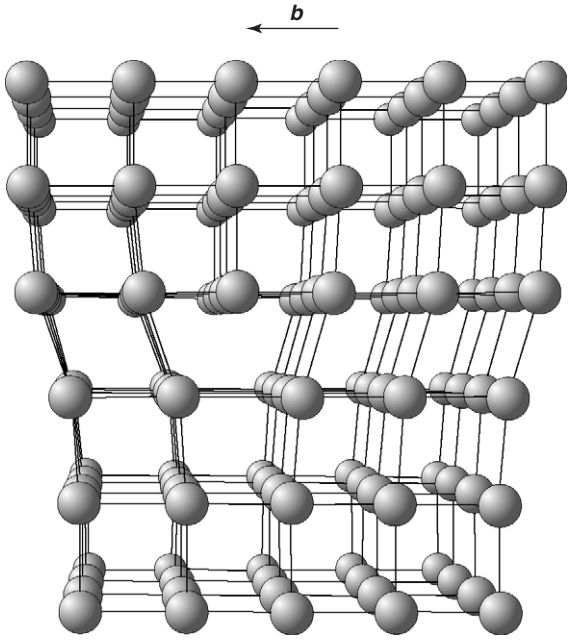


Figure 2 Schematic drawing of a dislocation, characterized by its line direction (the border of the incomplete atomic plane in this case) and the Burgers vector ***b*** (in this case – "edge" dislocation – horizontal and perpendicular to the line direction).

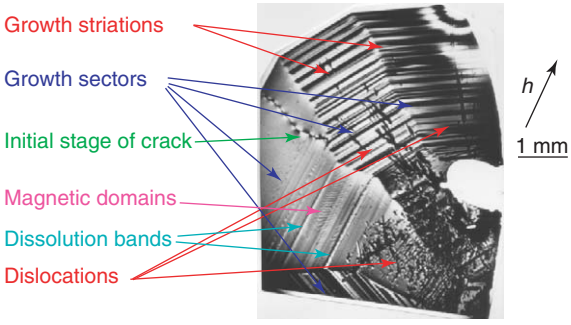


Figure 3 Topograph of a flux-grown Ga–YIG crystal.

magnetic domains, as well as the distortion between growth sectors, are in the 10^{-6} range. The distortion associated with dislocations is a function of the distance to the core; at the level of the image border, it is in the 10^{-5} range. Large contrast on the image can, therefore, correspond to rather weak distortions, showing that these techniques are very sensitive.

Some Results of Dynamical Theory

A few results of the dynamical theory, necessary to understand the contrast mechanisms, are given below. The discussion is restricted to symmetrical transmission geometry as shown in **Figure 1**.

In the case of a monochromatic plane wave beam, and a perfect crystal set in Bragg diffraction position (**Figure 1**), two "wavefields" propagate within the crystal if the direction of the incident beam propagation vector lies within an angle ω_b around the exact Bragg position (ω_b the intrinsic width of the diffraction curve, often called the Darwin width, being of the order of 10^{-5}–10^{-6} rad):

$$\omega_b = (2\lambda^2 C_p |F_b| r_0)/(\pi V_c \sin 2\theta_B) \qquad [1]$$

where C_p is the polarization factor ($C_\pi = \cos 2\theta_B$ for π polarization, i.e., for the electric field vector $\boldsymbol{E}$ in the scattering plane, and $C_\sigma = 1$ for σ polarization, i.e., for $\boldsymbol{E}$ perpendicular to this plane), r_0 is the "classical electron radius" (2.8×10^{-15} m), V_c is the unit cell volume, and F_b is the structure factor corresponding to the Bragg reflection used. A wavefield corresponds to the superposition of one diffracted and one forward-diffracted wave, coupled because the perfect crystal is set for Bragg diffraction. These waves decouple outside the crystal, producing two beams, parallel to the incident and diffracted directions, respectively (**Figure 4**).

In the case of a monochromatic spherical wave, wavefields propagate and interfere within the whole "Borrmann" fan or Borrmann triangle ABC (**Figure 4**). The characteristic coupling length between wavefields is Λ_0 (extinction length, or Pendellösung period), which corresponds, in the case considered, to the distance necessary for most of the energy participating in the diffraction process to move from the transmitted to the diffracted direction, or vice versa:

$$\Lambda_0 = (\pi V_c \cos\theta_B)/(\lambda C_p |F_b| r_0) \qquad [2]$$

Typical values of Λ_0 are in the 1–100 μm range.

If the crystal is rotated in a monochromatic beam, or if the beam is polychromatic, the diffracted power is a function of this "variable" (angle, or wavelength), leading to the integrated reflectivity ρ, defined as the area under the diffraction curve. **Figure 5** shows ρ for a nonabsorbing perfect crystal, as a function of t/Λ_0, t being the thickness of the platelet-shaped crystal, for both the transmission ("Laue") and reflection ("Bragg") cases. These curves show that

- ρ is proportional to t for $t \ll \Lambda_0$ ("kinematical" limit of the dynamical theory);
- ρ is nearly constant for $t \gg \Lambda_0$, the saturation value being proportional to $|F_b|$; and
- ρ displays an oscillatory behavior in the Laue case (the image produced by a low-absorption wedge-shaped crystal will, therefore, exhibit "equal thickness" fringes associated with the maxima and minima of the curve, and to the varying

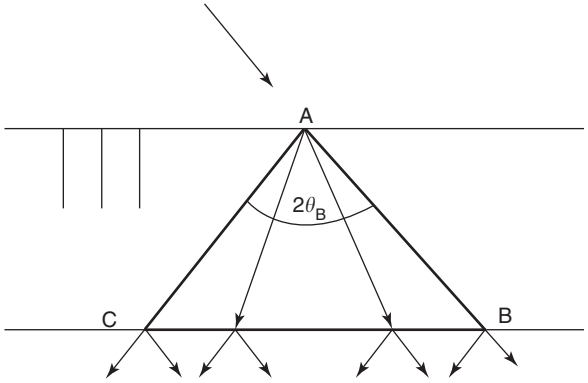


Figure 4 Propagation of wavefields within the "Borrmann" triangle ABC in the case of a perfect crystal.

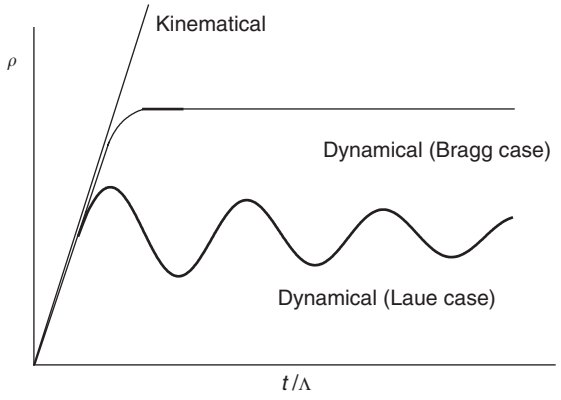


Figure 5 Reflecting power ρ of a nonabsorbing perfect crystal as a function of the reduced thickness t/Λ_0 for the transmission ("Laue") and reflection ("Bragg") cases. The kinematical approximation limit is also shown.

thickness of the crystal). Its saturation value, for $t \gg \Lambda_0$, is half of the one corresponding to the Bragg case.

Effect of Imperfections: Contrast Mechanisms

For a nonabsorbing plate-shaped crystal illuminated by a parallel and polychromatic beam (a good approximation for the simplest and presently the most usual topographic technique, white beam synchrotron radiation topography), within a very simplified approach, the direction and intensity of the locally diffracted intensity depends upon

- the structure factor F_h;
- the angle θ formed by the considered lattice planes, and the incident beam; and
- Y, a parameter that incorporates the modifications introduced by the crystal inhomogeneities on the perfect-crystal dynamical theory results.

The beams diffracted by two neighboring regions I and II of the sample produce contrast on the topograph if

1. $(F_h)_I \neq (F_h)_{II}$, where the differences can either reside in the modulus or the phase or the structure factor. This structure factor contrast is the main contrast mechanism in neutron diffraction topography of magnetic domains, the magnetic structure factor often being different for the various domains (**Figure 6**);
2. $\theta_I \neq \theta_{II}$, that is, regions I and II are misoriented – subgrains, domains, etc. – one with respect to the other: this is orientation contrast; **Figure 7** shows, for instance, white or black thick lines, corresponding to the separation, or superposition, of the images of subgrains;
3. $Y_I \neq Y_{II}$, the two regions display extinction contrast.

Imperfections modify the perfect-crystal behavior described by the dynamical theory. They can hide the interference fringes, or lead to locally enhanced or reduced diffracted intensity, depending on the absorption conditions. New images can occur through the decoupling of a wavefield and the creation of new wavefields at a defect.

In the low-absorption case, the predominant contrast mechanism is the so-called direct-image process. Assume again (**Figure 1**) that a polychromatic, extended beam impinges on a crystal such that a small crystal volume V contains a defect (inclusion, dislocation, etc.) associated with a distortion field, which decreases with growing distance from the defect core. A wavelength range $\Delta\lambda/\lambda = \omega_h/(\tan\theta_B)$ $(\sim 10^{-4})$ participates in diffraction by the "perfect" matrix crystal, whereas regions around the defect are at the Bragg position for components of the incoming beam which are outside this spectral range. The defect thus leads to additional diffracted intensity on the detector.

The regions that produce the direct image of a given defect are some distance away from the core of the defect. This distance depends not only on the nature of the defect, but also on the diffraction process itself. The lattice distortion acts on diffraction through an angle, the effective misorientation $\delta\theta$, which reflects the change in the departure from the Bragg angle that is associated with the existence of the defect:

$$\delta\theta(r) = -(\lambda/\sin 2\theta_B)\partial(\boldsymbol{h} \times \boldsymbol{u}(\boldsymbol{r}))/\partial(s_h) \qquad [3]$$

where $\boldsymbol{h}$ is the undistorted reciprocal lattice vector, $\boldsymbol{u}(\boldsymbol{r})$ is the displacement vector, and $\partial/\partial(s_h)$ is the

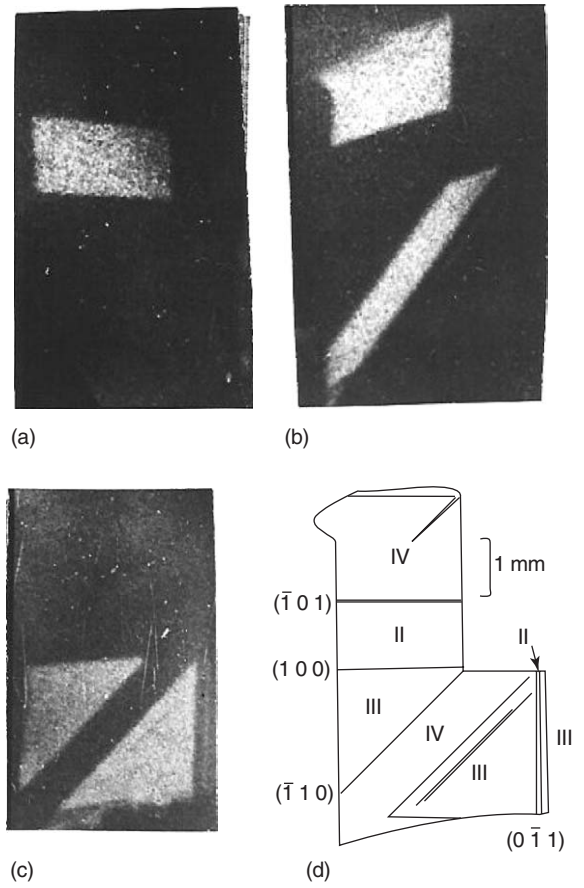


Figure 6 The observation, by neutron diffraction topography of antiferromagnetic domains in NiO is an example of "structure factor contrast." These domains are characterized by their propagation vector, which lies along one of the four equivalent 1 1 1 directions; each type of domain has a nonzero magnetic structure factor for the Bragg reflection corresponding to its propagation vector, and zero for the other three magnetic reflections: (a) 1,1, – 1, (b) – 1,1,1, (c) – 1,1, – 1, and (d) drawing of the domain configuration – type I is missing in the investigated crystal. These images were recorded – exposure times ~10 h – using a ^{157}Gd screen acting as n-β converter and X-ray film; "white" means more illuminated.

Figure 7 White beam topograph of an Fe–3%Si crystal, $E \approx 25$ keV, showing orientation contrast associated with subgrains. A fir-tree magnetic domain pattern is also visible.

differentiation along the reflected beam direction. It was shown that the width of the direct image corresponds to the contribution of regions such that $\delta\theta$ is of the order of ω_b. This can be understood when considering that the regions where the distortion is $< \omega_b$ are within the diffraction range of the perfect crystal, and that those where the distortion is $\gg \omega_b$ have a negligible contribution for the considered Bragg diffraction spot. This leads to "intrinsic" image widths, for an edge dislocation such as the one schematized in **Figure 2**, $\sim \Lambda_0 |\boldsymbol{b} \cdot \boldsymbol{h}|/4$ ranging in the 1–100 μm scale, and therefore implies a limitation on the density of defects that can be resolved on a topograph ($\sim 10^4$–$10^5\,\mathrm{cm}^{-2}$ for dislocations).

This also implies that direct images are only seen in crystals that are thick enough. Indeed, the dynamical theory predicts a width ω_b for thin ($t < \Lambda_0/\pi$) crystals inversely proportional to the thickness. If the crystal is thin, only very distorted regions with a strongly reduced volume – and consequently reduced diffracted intensity – can contribute to the direct image, which cannot therefore be observable.

The effective misorientation $\delta\theta$ allows, in addition, determining on which reflections a given defect can produce an image. Equation [3] shows that $\delta\theta = 0$ if $\boldsymbol{u}$ is perpendicular to $\boldsymbol{h}$. The defect is not visible on the corresponding topographs. A dislocation, for instance, is usually not visible when $\boldsymbol{h} \cdot \boldsymbol{b} = 0$, $\boldsymbol{b}$ being the Burgers vector. The invisibility criteria are powerful ways of characterizing the defects.

The effective misorientation can also be written as a function of the local variation in the rotation $\delta\varphi$ of the concerned lattice planes and the local relative variation in the lattice parameter, $\Delta d/d$, as

$$\delta\theta = -(\Delta d/d)\tan\theta_B \pm \delta\varphi \qquad [4]$$

When taking absorption into account, the situation can change drastically. Direct images do not occur in the high-absorption case ($\mu t > 10$), and are replaced by dips in the intensity corresponding to the disruption of a dynamical-theory-related effect, the anomalous transmission (also called the Borrmann effect). Both types of images ("additional" intensity and "dips") are simultaneously present on the topographs, as well as interference fringes ("dynamical" and "intermediary" images) for samples with intermediate absorption ($\mu t \sim 2-5$).

Diffraction Topographic Techniques

As already discussed, Bragg diffraction imaging may be used either in transmission or in reflection. In addition, the beam can be extended or restricted, and the image can be an integrated or a nonintegrated one.

Figure 1 shows an extended, divergent, and/or nonmonochromatic beam illuminating the whole sample, or a large area. This results in an image which can, to a first approximation corresponding to the predominance of the direct images, be considered as the projection of the defect distribution along the diffracted beam direction. The incident beam can also be restricted to a small width (typically 10–20 μm) in the scattering plane. In the same first approximation, this produces an image of those defects intercepted by the blade-shaped incident beam, hence the name "section" topograph.

A better definition of the extended/restricted beam distinction is based on the width of the Borrmann triangle, $w_B = 2t \sin\theta_B$ in the symmetrical Laue case, t being the crystal thickness. The extended beam, or projection topograph, situation corresponds to the case where the width w of the incoming beam in the scattering plane is such that $w \gg w_B$. The restricted beam, or section topograph, case corresponds to $w \ll w_B$.

In the section topography, the position of the direct images within the image can be used to extract the depth of the related defects in the crystal, one edge of the image corresponding to the entrance face of the sample, the other to the exit surface. In very good crystals, with small or moderate absorption, the same interference effect that leads to the oscillations shown in **Figure 5**, produces interference fringes, called Kato's fringes. These fringes are very sensitive to crystal distortion, and their modification is one of the first indications of a departure from the perfect-crystal situation (see **Figure 11**).

The use of a polychromatic beam with a sufficiently wide bandwidth often offers advantages. Several reflections are recorded simultaneously, which is very helpful when characterizing defects or phenomena, and misoriented regions within a sample diffract at once, each region finding a wavelength λ that satisfies the local Bragg condition. This "white beam" version of diffraction topography is identical to the well-known Laue technique, except that the incident beam is broad. Each Bragg spot is now a topograph. This is an integrated wave diffraction image, because it results from the superposition of contributions from a range of wavelengths. It is mainly sensitive to the variations of lattice plane orientation $\delta\varphi$, which produce a change in the diffracted beam direction, and not to $\Delta d/d$. "White beam" topography is widely used at synchrotron radiation facilities. The exposure time is often of the order of a fraction of a second, and it is possible to follow the evolution of any one of the Bragg spots in "real time," using a CCD camera equipped with a scintillator.

The incident beam on the sample can also be monochromatized. This "monochromatic wave" nevertheless displays a range of wavelengths ($\Delta\lambda/\lambda$) as well as an angular divergence ω. Let ω_c be the diffraction width of the crystal (which reduces to ω_b in the case of a "perfect" crystal). If the divergence ω (and/or $(\Delta\lambda/\lambda)\tan\theta_B$) $> \omega_c$, the image is still an integrated wave topograph.

If ω (and $(\Delta\lambda/\lambda)\tan\theta_B$) $\ll \omega_b$, the technique is called plane wave topography. It is not only suited to detect weak deformations (10^{-6}–10^{-8} range), but also to obtain images of defects such as dislocations (see **Figure 9**) with much more details than in the case of direct images. Its use allows a quantitative analysis, because the local diffracted intensity is closely related to the effective misorientation $\delta\theta$. An example of the use of "plane wave" topography is shown in **Figure 8**: the seed of the investigated quartz crystal and the various growth sectors exhibit different impurity contents, which result in small variations of the lattice parameter (in the 10^{-5}–10^{-7} range).

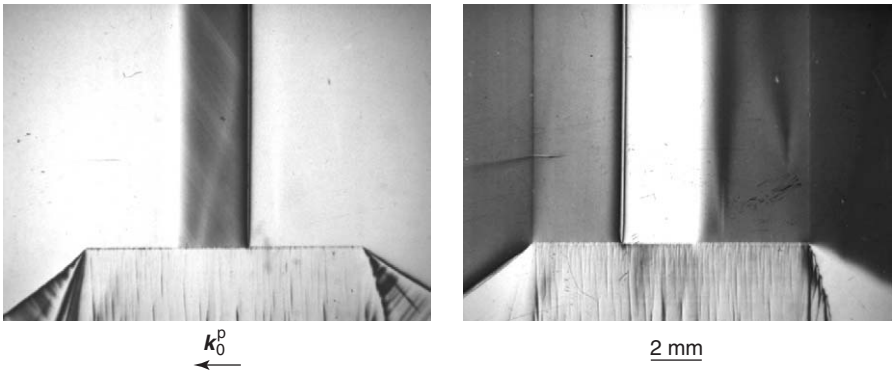


Figure 8 Bragg case plane wave topographs of a quartz plate with nonhomogeneous distributions of impurity atoms in different growth sectors and in the seed plate (rectangle in the middle part), $40\bar{4}0$ reflection, $E = 8.04778$ keV, $\boldsymbol{k}_0^p$ is the projection of the wave vector of the incident wave on the crystal surface; the topographs were recorded at ~40% of the maximum intensity on the low-angle side (left image) and ~20% of the maximum intensity on the high-angle side (right image) of the rocking curve.

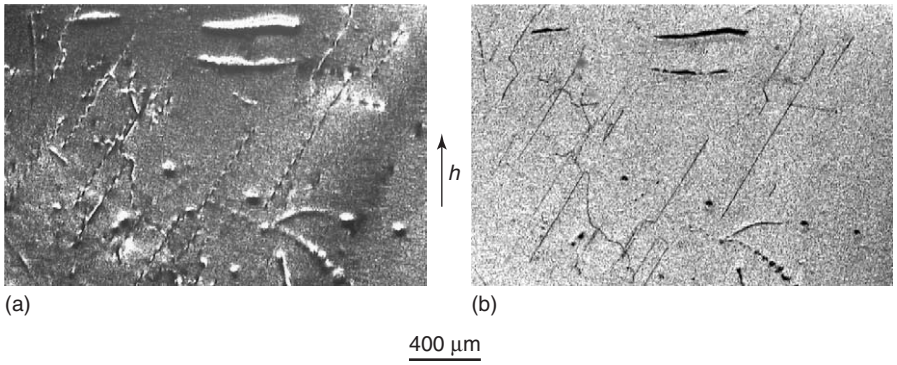


Figure 9 1 1 1 topographs of a Ge crystal recorded using the beam (35 keV) produced by an Si 1 1 1 monochromator: (a) top of the rocking curve, and (b) on the high-angle tail of the rocking curve ("weak beam" image).

If an image is recorded at an angular position very far away (typically a few ω_c) from the center of the diffraction curve, the main contribution to Bragg diffraction originates from the distorted regions, and not from the perfect-crystal matrix. This allows reducing the "intrinsic" width of the images and is known, by analogy with electron microscopy, as weak beam topography (**Figure 9**).

Synchrotron Radiation Topography: Real-Time Investigations and Coherence-Related Effects

Synchrotron radiation is now widely used for X-ray imaging. The typical exposure time to record a synchrotron radiation topograph is a fraction of a second. It is therefore possible to investigate evolving phenomena such as the movements of defects, boundaries during first-order phase transitions, or domains. Specific sample environments (cryostat, furnace, electric or magnetic field, strain, etc.) are added for each particular experiment. Image subtraction can help emphasize changes with respect to a reference state.

It is possible to investigate quicker phenomena if they are periodic, through stroboscopic imaging. Choppers located on the beam path and synchronized with periodic magnetic or electric fields were used to investigate phenomena such as the motion of magnetic domains or the evolution of conduction channels, in the 10^{-2}–10^{-3} s range. Stroboscopic experiments on bulk or surface acoustic waves, in the 10^{-8}–10^{-9} s range, were performed using the pulsed structure of the synchrotron radiation source itself (**Figure 10**).

The small source sizes and the large source-to-sample distances, which are now common features of modern synchrotron radiation facilities, lead to highly coherent X-ray beams, which can be used for diffraction topography. In this case surface inhomogeneities, or porosities, can be imaged in the diffracted beam even when they do not produce a strain field through their Fresnel diffraction image. Recording images at various sample-to-detector distances allow retrieving the phase jump involved in these images. The use of Bragg diffraction imaging with a coherent beam can even, in special cases, produce structural information at the atomic level, even though the spatial resolution of the images is on a much larger scale of micrometers. For instance, diffraction topographs of ferroelectric crystals, using a coherent beam, give access to a very elusive microstructural piece of information: how ferroelectric domains are connected across the domain wall and, in particular, how they match at the atomic level.

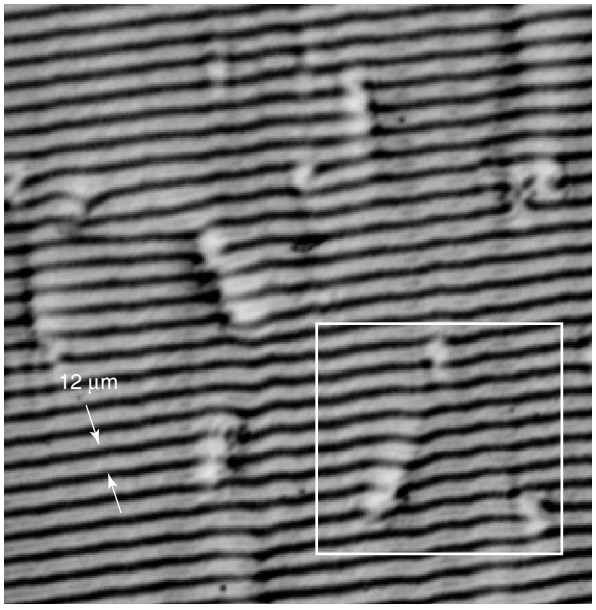


Figure 10 Stroboscopic topograph showing the propagation of surface acoustic waves – 12 μm wavelength – in an $LiNbO_3$ crystal. Wave front distortions caused by the interaction of the waves with dislocations are visible. One dislocation region is framed. (Courtesy of E Zolotoyabko.)

Simulations of X-Ray Topographs

The information provided by the topographs allows a detailed quantitative analysis. This is performed through simulations. The simulation process requires introducing a deformation field with free parameters that will be refined by comparing the simulated image to the experimental one. Depending on how rapidly the distortions (expressed by the effective misorientation $\delta\theta r$) change locally within the crystal, this may be done using diffraction theories with various complexities. In the case of **Figure 8**, the Bragg-case plane wave topographs of a quartz plate with nonhomogeneous distributions of impurity atoms, the local application of the dynamical theory for perfect crystals already provides very good results. In this case, the diffracted intensity can be approximated by the simple formula $I_h = I_M R_N(\theta_A - \delta\theta)$, where R_N is the rocking curve of the setup, I_M is a normalization factor, and θ_A is the Bragg angle in a crystal part assumed as the perfect reference one. This means that the contrast depends on the position of the "working point," which is determined, through the parameter $\delta\theta(\boldsymbol{r})$, by the local Bragg angle on the perfect reference–rocking curve. For faster varying deformations (**Figure 11**), a geometric-optical approximation may be chosen. The most general case requires a wave-optical approach, which is necessary to calculate, for instance, the contrast of a dislocation. The algorithms give very good results for the plane wave case.

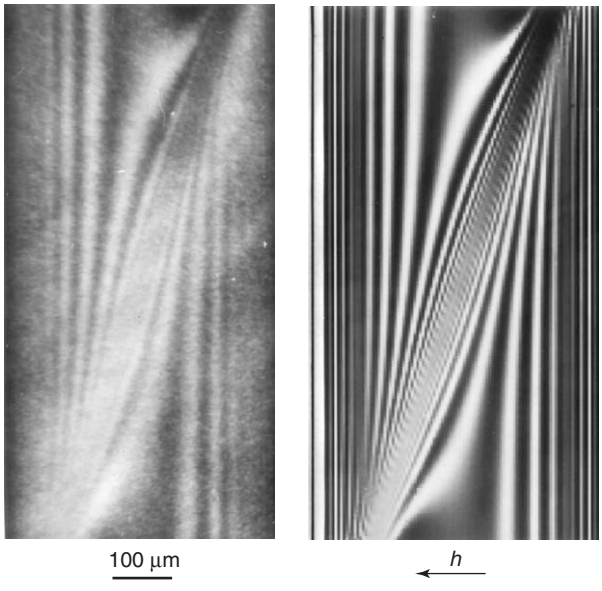


Figure 11 Measured (left) and calculated (right) 422 section topographs of a 565 μm thick silicon crystal with the edge of a 150 nm silicon oxide film, $E = 17.48$ keV.

Conclusion

X-ray diffraction topography is an invaluable tool to characterize single crystals and investigate many-crystal physics phenomena (creation of defects, domains, phase transitions, vibrations, etc.), whereas neutron topography remains a unique tool for the investigation of magnetic crystals, and more specially the antiferromagnetic ones. The scope of X-ray topography, when coupled with synchrotron radiation, extended to unexpected topics through the use of coherent beams or by *in situ* experiments, where the material, in an adequate sample environment device, is imaged while an external parameter (temperature, stress, etc.) is changed. Some of the applications, of course, require a better spatial resolution. A possibility of overcoming this limitation is to use an X-ray lens to magnify the image before the detector. Promising attempts are being made using parabolic compound refractive lenses, asymmetrically cut crystals, or bent mirrors (Kirkpatrick–Baez setup).

See also: Crystal Growth, Bulk: Theory and Models; Dislocations; Scattering, Inelastic: X-Ray (Methods and Applications); X-Ray Absorption Spectroscopy; X-Ray Sources; X-Ray Standing Wave Techniques.

PACS: 07.85; 61.10; 61.50.Cj; 61.72.y; 61.72.Dd; 61.72.Ff; 64.70.Kb; 75.60.Ch; 77.80.Dj

Further Reading

Authier A (2001) *Dynamical Theory of X-Ray Diffraction*. New York: Oxford University Press.

Authier A, Lagomarsino S, and Tanner BK (1996) *X-Ray and Neutron Dynamical Diffraction – Theory and Applications*. New York: Plenum.

Bowen DK and Tanner BK (1980) *Characterization of Crystal Growth Defects by X-Ray Methods*. New York: Plenum.

Bowen DK and Tanner BK (1998) *High Resolution X-Ray Diffractometry and Topography*. London: Taylor and Francis.

Nomenclature

$\boldsymbol{b}$	Burgers vector of a dislocation (m)
C_p	polarization factor (X-ray diffraction)
$F_{\boldsymbol{h}}$	structure factor associated with the reciprocal vector $\boldsymbol{h}$
r_0	classical electron radius (2.8×10^{-15} m)
$\boldsymbol{u}(\boldsymbol{r})$	displacement vector (m)
V_c	volume of the unit cell (m^3)
$\delta\theta(\boldsymbol{r})$	effective misorientation (rad)
θ_B	Bragg angle (rad)
λ	wavelength (m)
Λ_0	extinction distance (dynamical theory) (m)
$\omega_{\boldsymbol{h}}$	intrinsic width of the diffraction curve (dynamical theory) (rad)